بهبود کدگذار حسگری فشرده ویدئو با استفاده از نمایش تنک

علی توانا[۱]،دانشجوی کارشناسی ارشد، عابدین واحدیان[۲]،دانشیار، احد هراتی[۳]، استادیار

۳،۲،۱- دانشکده مهندسی- دانشگاه فردوسی مشهد-مشهد-ایران

Ali.tavana68@stu.um.ac.ir[۱]، vahedain@um.ac.ir[۲]، a.harati@um.ac.ir[۳].

چکیده

بر اساس نظریه حسگری فشرده می‌توان سیگنال تنک را از روی تعداد کمی سنجش دارای ناهمدوسی به خوبی بازیابی کرد. حسگری فشرده روشی جدید برای نمونه‌برداری کمتر از نرخ نایکوئیست ارائه می‌کند که به علت همبستگی بالا بین فریم‌های ویدئو گزینه‌ی مناسبی برای فشرده‌سازی است. در این مقاله سعی شده تا روشی برای فشرده‌سازی ویدئو مبتنی بر حسگری فشرده ارائه شود. در روش پیشنهادی میزان تفاوت بین فریم‌ها و همچنین مرتب‌سازی لغت‌نامه تعیین‌کننده‌ی میزان تنکی و نرخ فشرده‌سازی در فریم‌های ویدئو می‌باشد. در این روش سعی شده به‌جای نمونه‌برداری از سیگنال اصلی, بردار ضرایب آن توسط حسگری فشرده نمونه‌برداری شود. نتایج آزمایشات نشان می‌دهد ویدئوهای بازیابی شده در این روش با نرخ فشرده‌سازی کمتر، دارای کیفیت مناسب‌تری هستند.

واژه‌های کلیدی

حسگری فشرده ویدئو, نمایش تنک, لغت‌نامه دارای افزونگی, ناهمدوسی

۱- مقدمه

در نمونه‌برداری‌های مرسوم از نرخ نایکوئیست که دو برابر بودن تعداد نمونه‌ها نسبت به فرکانس بیشینه موجود در سیگنال را ضروری می‌داند استفاده شده است. طبق نظریه حسگری فشرده اگر سیگنال را بتوان در پایه‌ای خاص به صورت تنک[۱] بیان کرد آنگاه می‌توان آن را با نرخی کمتر از نرخ نایکوئیست نمونه‌برداری کرد [۱] . نمونه‌برداری از سیگنال تنک با استفاده از ترکیب خطی تعداد کمی سنجش انجام می‌گیرد.

در حسگری فشرده ویدئو علاوه بر استفاده از لغت‌نامه‌های مرسوم می‌توان از افزونگی بین فریمی نیز برای بیان تنک بودن و ساخت لغت‌نامه استفاده کرد. در مقاله [۲] از ویژگی همبستگی زمانی فریم‌های مجاور برای به دست آوردن لغت‌نامه استفاده شده است. در حسگری فشرده برای بیان تنک از لغت‌نامه، فقط در سمت کدگشا استفاده می‌شود. در روش پیشنهادی سعی شده تا لغت‌نامه‌ای محلی به ازای هر بلوک از هر فریم ساخته شود. این لغت‌نامه به تنک شدن بلوک کمک بسیاری خواهد کرد و از آن در سمت کدگذار استفاده خواهد شد. همچنین این لغت‌نامه بدون ارسال هرگونه اطلاعات اضافی در سمت کدگشا بازسازی خواهد شد. در حسگری فشرده ویدئو معمولاً از ویژگی بلوک‌بندی برای افزایش سرعت انجام محاسبات در مسئله بهینه‌سازی استفاده می‌شود. همچنین عدم استفاده از بلوک‌بندی مسئله را نیازمند لغت‌نامه‌ای بزرگ‌تر و جامع‌تر برای نمایش تنک[۲] خواهد کرد. در روش پیشنهادی علاوه بر استفاده از بلوک‌بندی از فریم‌های کلید و غیرکلید نیز استفاده خواهد شد. فریم‌های کلیدی با استفاده از روش حسگری فشرده مرسوم و با استفاده از مرتب‌سازی لغت‌نامه کدگذاری و کدگشایی می‌شوند. فریم‌های غیرکلید نیز با توجه به تفاوتشان با فریم‌های کلید کدگذاری می‌شوند. برخلاف روش‌های حسگری فشرده مرسوم، در روش پیشنهادی از یک مسئله تنک‌سازی در سمت کدگذار استفاده و به جای نمونه‌برداری از سیگنال، از بردار ضرایب آن در لغت‌نامه, نمونه‌برداری خواهد شد. لغت‌نامه تولیدی به خاطر ترکیب لغت‌نامه انتقالی‌کسینوسی‌گسسته[۳] و بلوک‌های مجاور دارای افزونگی می‌باشد.

در ادامه این مقاله:در بخش ۲ نظریه حسگری فشرده بررسی می‌شود. در بخش ۳ به بررسی مقاله‌های انجام شده در این زمینه پرداخته خواهد شد. در بخش ۴ روش پیشنهادی مطرح می‌شود. در بخش ۵ نتایج الگوریتم پیشنهادی قرار داده شده است.و در پایان نتیجه‌گیری انجام می‌گیرد.

۲-نظریه حسگری فشرده

به علت مناسب بودن حسگری فشرده برای سخت‌افزار، اخیراً در نمونه‌برداری ویدئو نیز از آن استفاده شده است. در حسگری فشرده سعی بر آن است تا با ترکیب نمونه‌برداری و فشرده‌سازی فقط نمونه‌هایی انتخاب شوند که حاوی حداکثر اطلاعات می‌باشند; البته این شرایط زمانی درست خواهد بود که سیگنال حس شده یا از ابتدا تنک باشد یا در پایه‌ای خاص به صورت تنک قابل بیان باشد، به همین علت بیشتر روش‌های حسگری

[1] Sparse

[2] Sparse representation

[3] .Discrete cosines transform (DCT)

فشرده به دنبال استفاده از لغت‌نامه‌هایی ثابت هستند که سیگنال در آن‌ها به صورت تنک قابل بیان باشند. هرچند در اغلب روش‌های این دسته از لغت‌نامه‌های ثابت برای بیان تنک یا تنک‌سازی استفاده می‌شود؛ اما باید به این نکته هم توجه داشت که به علت همبستگی موجود بین فریم‌های ویدئو می‌توان از این ویژگی در تنک‌سازی فریم‌های غیرکلید استفاده کرد. حسگری فشرده به دو بخش کلی کدگذار و کدگشا تقسیم می‌شود.

۲-۱- حسگری فشرده درسمت کدگذار

فشرده‌سازی و نمونه‌برداری سیگنال در این قسمت انجام می‌گیرد. طبق نظریه حسگری فشرده به علت استفاده در سخت‌افزار این قسمت باید سریع و دارای حجم محاسباتی کمی باشد. روش حسگری فشرده برخلاف دیگر روش‌های نمونه‌برداری مرسوم، ازنظر ریاضی قابل اعمال بر روی بردار با ابعاد مشخص می‌باشد بنابراین برای نمونه‌برداری سیگنال گسسته‌ی x به طول n باید بتوان آن را به صورت تنک بیان کرد.

$$x = \psi\alpha \qquad (۱)$$

رابطه بالا حالتی کلی‌تر نسبت به نمایش تنک می‌باشد که در آن ψ پایه‌ای متعامد یا لغت‌نامه دارای افزونگی [۳] با ابعاد $n \times m$ و α بردار ضرایب تنکی با بعد m می‌باشد. از رایج‌ترین پایه[4]های مورد استفاده در ψ می‌توان به فوریه و موجک و DCT اشاره کرد که سیگنال را به فضایی منتقل می‌کنند که در آنجا سیگنال تنک است, پس با اینکه سیگنال اولیه تنک نبوده می‌توان آن را ابتدا با یک تبدیل تنک کرد سپس از آن در حسگری فشرده استفاده کرد. برای تنک‌سازی از همبستگی بین فریمی هم به صورت زمانی و هم به صورت مکانی می‌توان استفاده کرد.

طبق نظریه حسگری فشرده بعد از تنک‌سازی سیگنال، با استفاده از ماتریس حسگری $\emptyset$ با ابعاد $d \times n$ می‌توان سیگنال x را به صورت زیر نمونه‌برداری کرد:

$$y = \phi x \qquad (۲)$$

طبق [۴] برای تضمین یکتا بودن بردار ضرایب باید ناهمدوسی بین لغت‌نامه و ماتریس حسگری به درستی رعایت شود به همین منظور معمولاً از ماتریس حسگری تصادفی استفاده می‌شود. تصادفی بودن ماتریس حسگری به مستقل بودن این ماتریس از محیط سیگنال کمک فراوانی می‌کند. در روش پیشنهادی نیز در فریم‌های غیرکلید از ماتریس حسگری تصادفی استفاده شده است تا به علت وجود لغت‌نامه دارای افزونگی، یکتا بودن بردار ضرایب تضمین شود. در حالت کلی با جایگذاری ۱ در ۲ می‌توان رابطه زیر را نیز نتیجه گرفت:

$$y = \phi\psi\alpha \qquad (۳)$$

رابطه‌ای که بین ابعاد ماتریس و سیگنال در حسگری فشرده وجود دارد به صورت $d < n \ll m$ قابل بیان می‌باشد. اما نکته‌ای که در رابطه با ماتریس حسگری باید در نظر داشت قابلیت تطبیق آن با سخت‌افزار می‌باشد که تحقیقات گسترده‌ای درزمینهٔ طراحی و آموزش ماتریس حسگری ارائه شده است. به همین علت ماتریس‌های حسگری یا به صورت کامل تصادفی می‌باشند یا در مواردی محدود از ماتریس‌های آموزش دیده یا غیرتصادفی استفاده می‌شود[۱].

۲-۲- حسگری فشرده درسمت کدگشا

بعد از نمونه‌برداری، بردار y به سمت کدگشا ارسال می‌شود و در سمت کدگشا بردار x دوباره بازیابی می‌شود. برخلاف کدگذار که فرایند محاسباتی بسیار کمی داشت، بازیابی و کدگشایی در حسگری فشرده نیازمند حل کردن مسئله‌ی پیچیده بهینه‌سازی است. اگر سیگنال نمونه‌برداری شده در فضای لغت‌نامه به صورت تنک قابل بیان باشد, آنگاه می‌توان بردار ضرایب تنک را با حل کردن یک مسئله بهینه‌سازی به صورت زیر حل کرد:

$$\alpha^* = argmin \left|\left|\alpha\right|\right|_0 \; s.t. \; \left|\left|y - \emptyset\psi\alpha\right|\right|_2 \leq \varepsilon \qquad (۴)$$

در رابطه بالا نرم صفر به معنی تعداد غیر صفرهای بردار ضرایب می‌باشد. همچنین در رابطه بالا ε مقدار انرژی نویز تخمینی از کانال است. به علت NP-hard بودن مسئله بالا [۵] می‌توان آن را توسط روش‌های مختلف از جمله روش‌های حریصانه مثل OMP یا از روش‌های محدب آرام مثل BP که با حل کردن مسئله بهینه‌سازی بالا سعی در پیدا کردن α دارند, حل کرد. روشی که در این مقاله از آن استفاده خواهد شد BP است که نرم صفر را به صورت زیر تخمین می‌زند:

$$\alpha^* = argmin \left|\left|\alpha\right|\right|_1 \; s.t. \; \left|\left|y - \emptyset\psi\alpha\right|\right|_2 \leq \varepsilon \qquad (۵)$$

در روش‌های محدب آرام با جایگزینی $\ell 1$ به جای $\ell 0$ می‌توان رابطه کمینه‌سازی را با استفاده از بهینه‌سازی محدب حل کرد. بر اساس [۳] هزینه محاسبات در $\ell 1$ بسیار کمتر از $\ell 0$ است. بعد از مشخص شدن بردار ضرایب با رابطه زیر می‌توان سیگنال اولیه را بازیابی کرد:

$$\hat{x} = \psi\alpha^* \qquad (۶)$$

۳- مروری بر کارهای گذشته

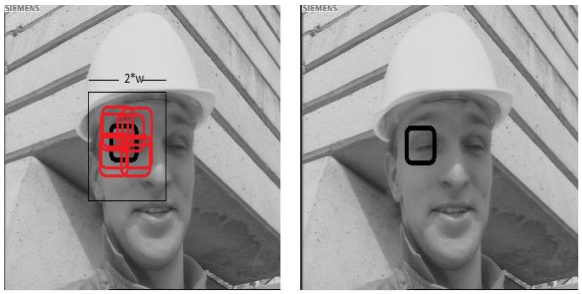
در مقاله [۶] سعی شده است لغت‌نامه‌ای کامل از ترکیب دو پایه متعامد ساخته شود. در حقیقت به ازای استفاده از یک پایه متعامد از لغت‌نامه‌ای دارای افزونگی استفاده شده است. استفاده از لغت‌نامه دارای افزونگی تنک بودن سیگنال را تضمین می‌کند. تنک بودن مهم‌ترین اصل در حسگری فشرده می‌باشد ازاین‌رو تلاش‌های بسیاری برای بیان تنک سیگنال در انواع لغت‌نامه صورت گرفته است. یکی از ویژگی‌های ویدئو افزونگی بین فریمی است که از آن می‌توان در تنک‌سازی استفاده کرد. ازاین‌رو در مقاله [۲] با در نظر گرفتن این نکته که بیشتر بار محاسباتی در بازسازی ویدئو در حسگری فشرده بر عهده‌ی کدگشا می‌باشد، از فریم‌های کلید در سمت کدگشا, برای ساخت لغت‌نامه استفاده می‌کند. این لغت‌نامه در بازنمایی تنک فریم‌های غیرکلید استفاده شده است. در این مقاله به ازای هر بلوک، از پنجره‌ای لغزان درون فریمی برای استخراج بلوک‌های مجاور در فریم‌های

[4] .basis

کلید بازسازی شده‌ی قبل استفاده می‌شود, سپس از پایه KLT برای ساخت لغت‌نامه و بازیابی بلوک جاری استفاده می‌شود. با استفاده از دانش‌های اضافی درباره تنکی سیگنال می‌توان به جای استفاده از ماتریس‌های حسگری تصادفی از بهینه‌ی آن‌ها برای بازسازی سیگنال در حسگری فشرده استفاده کرد. به همین منظور در مقاله‌ی [۷] ماتریسی حسگری بهینه شده‌ای را ارائه می‌کند که همدوسی بین این دو ماتریس را حداقل می‌سازد. .در مقاله [۸] روشی مبتنی بر حسگری فشرده ارائه شده است که ویدئو به صورت پخشی برای چند کلاینت مجزا با ظرفیت‌های کانال و ویژگی‌های مختلف ارسال می‌شود. در این مقاله بازیابی ویدئو با استفاده از الگوریتم Total variation و لغت‌نامه DCT انجام می‌شود. در این مقاله به ماتریس حسگری از دید یک پراکنده کننده اطلاعات نگاه شده است و این بحث را مطرح می‌کند که هیچ‌کدام از ابعاد ماتریس حسگری نسبت به دیگری برتری و اهمیت بیشتری ندارد. طبق نظریه حسگری فشرده هرگاه سیگنالی دارای تنکی مناسبی در لغت‌نامه خود نباشد, برای جبران عدم تنکی باید از تعداد سنجش بیشتری در نمونه‌برداری استفاده کرد,که همین امر باعث پایین آمدن نرخ فشرده‌سازی می‌شود. در [۹] سعی شده است به دنبال یک ماتریس حسگری مناسب برای لغت‌نامه باشد. و در مقاله [۱۰] سعی می‌شود با حداقل کردن فاصله‌ی بین اتم‌های[5] لغت‌نامه‌ی آموزش دیده و ماتریس حسگری, بازیابی بهتری از سیگنال ارائه دهد.

در مقاله [۱۱] روشی برای فشرده‌سازی ویدئو بر اساس حسگری فشرده توزیع شده ارائه می‌شود که ماتریس تنکی بر اساس بیشینه کردن درست‌نمایی آموزش داده خواهد شد. در این روش از دو دسته فریم کلیدی و غیرکلیدی استفاده می‌شود که فریم‌های کلیدی به کمک روش مرسوم حسگری فشرده کدگذاری و کدگشایی می‌شوند اما برای کدگشایی فریم‌های غیرکلید علاوه بر آموزش لغت‌نامه از اطلاعات جانبی فریم‌های کلید بازیابی شده‌ی قبلی استفاده شده است.در [۱۲] روشی سلسله مراتبی با دسته فریم‌های متفاوت ارائه شده است که در هر دسته ماتریس‌های حسگری دارای ابعاد متفاوتی هستند. در این روش ابعاد ماتریس حسگری تصادفی با توجه به پیش‌بینی کوانتایزر و معیار بی‌نظمی از فریم‌های قبل تعیین می‌شود. سپس با استفاده از کوانتایزر توزیع داده‌های تولید شده توسط حسگری فشرده مرتب شده و کدگذاری می‌شوند و در سمت کدگشا نیز معکوس این فرآیند برای رسیدن به ویدئوی اولیه انجام می‌شود. در این الگوریتم فریم‌های غیرکلید به k گروه مختلف تقسیم می‌شوند، و در هر لایه نرخ نمونه‌برداری متفاوت است، در نتیجه کیفیت فریم‌های غیرکلید وابسته به کیفیت فریم‌های کلید و نرخ نمونه‌برداری آن دسته است. به‌طوری که هر چه فریم‌ها در لایه‌های بالاتر قرار داشته باشند دارای نرخ نمونه‌برداری بیشتری و در نتیجه کیفیت بازیابی بهتری خواهند بود. در مقاله [۱۳] هدف تنظیم نرخ نمونه‌برداری و ماتریس حسگری می‌باشد . بنابراین یکی از مهم‌ترین پارامترها در این روش، تنظیم تعداد بیت‌های مصرفی به ازای هر اندازه‌گیری برای رسیدن به بالاترین کیفیت و بهترین

[5] atom



شکل شماره ۱: روش استخراج لغت‌نامه محلی شکل سمت راست بلوک جاری در فریم غیرکلید و شکل سمت چپ دیکشنری متناظر آن در فریم کلید

نرخ بیت مصرفی است. ماتریس حسگری در اینجا با توجه به ویژگی‌های هر بلوک به صورت متفاوت تعریف می‌شود. نرخ بیت و عمق بیت نیز با توجه به الگوریتم بهینه‌سازی مطرح شده به دست خواهند آمد.

۴- روش پیشنهادی

در روش پیشنهادی ابتدا لغت‌نامه‌ای دارای افزونگی در سمت کدگذار ساخته خواهد شد و با استفاده از آن بردار ضرایب تنکی بلوک جاری تولید می‌شود. سپس بردار سنجش y گفته شده در رابطه (۳) از روی بردار ضریب تولیدی و به صورتی متفاوت از روش حسگری فشرده مرسوم فرایند فشرده‌سازی را انجام خواهد داد. در نهایت در کدگشا لغت‌نامه تولیدی در کدگذار دوباره ساخته شده و بردار ضرایب مورد نظر بازیابی می‌شود. در این روش فریم‌ها به دو دسته کلید و غیرکلید تقسیم می‌شوند که هر کدام به صورت جداگانه‌ای کدگذاری و بازیابی می‌شوند. فشرده‌سازی ویدئو در دو بخش اصلی که در کدگذار و کدگشا شرح داده خواهد شد.

۴-۱- روش پیشنهادی در سمت کدگذار

فرایندکدگذاری فریم‌های ویدئو به صورت بلوکی انجام می‌شود. به این صورت که تمام بلوک‌های تمام فریم‌ها به صورت غیر همپوشان کدگذاری و فشرده می‌شوند. برای کدگذاری فریم‌های کلید ابتدا لغت‌نامه‌ای مرتب، ساخته می‌شود. نحوه‌ی ساخت این لغت‌نامه به این صورت است که اتم‌های آن از پایه‌های فرکانس پایین DCT شروع و به فرکانس‌های بالا ختم می‌شود. در حقیقت این لغت‌نامه یک نوع DCT مرتب شده بر اساس فرکانس می‌باشد. این ترتیب از این نظر اهمیت دارد که سیگنال‌ها بیشترین شباهت و در نتیجه بزرگ‌ترین ضریب تولیدی در OMP را به فرکانس پایین دارند. بعد از اینکه اتم‌ها در لغت‌نامه مرتب شد، با استفاده از این لغت‌نامه بردار ضرایب تنکی بلوک جاری توسط الگوریتم OMP تولید می‌شود. فرض کنید بلوک جاری در فریم کلید x_k^i نامیده شود. انگاه بردار ضرایب α_i توسط OMP برای آن بلوک ساخته می شود. سپس حسگری فشرده بر روی آستانه‌ای مشخص از بردار ضرایب α_i تولیدی اعمال خواهد شد. چون DCT مرتب شده است بنابراین بردار ضرایب آن نیز به ترتیب از مقدار بزرگ به صفر تقریبا مرتب شده اند بنابراین حسگری فشرده فقط بر روی ۶۰ درصد اول(که شامل ضرایب بزرگ و مهم می‌باشد) اعمال خواهد شد این فرایند به گونه ایست که گویی از ضرایب کوچک نزدیک به صفر

نمونه برداری نشده است. در سمت کدگشا بعد از بازیابی بردار ضرایب مهم مابقی بردار ضرایب با صفر جایگذاری می‌شود. رابطه کدگذاری با حسگری فشرده برای فریم‌های کلید در زیر آمده است

$$y_k = \phi\hat{\alpha} \qquad (۷)$$

در این رابطه $\hat{\alpha}$ همان ۶۰ درصد اول بردار ضرایب یا مهم‌ترین ضرایب است. مزیت این کار پایین آمدن نرخ بیت مصرفی در بردار y است.

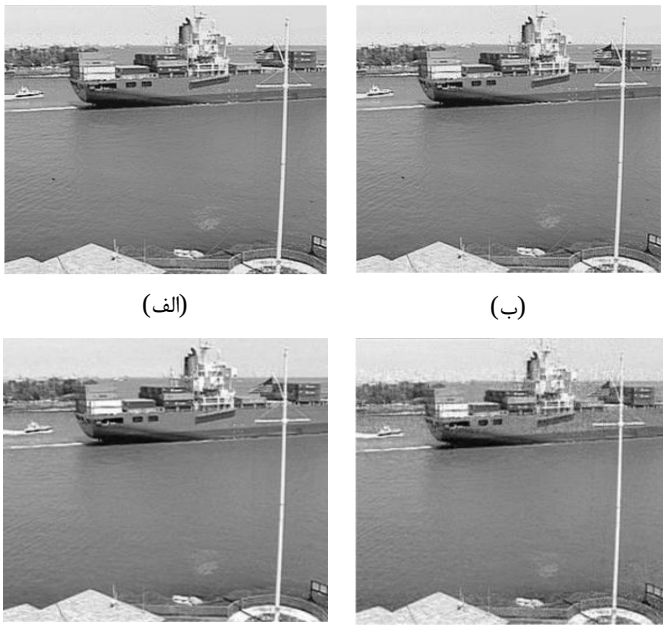
فریم‌های غیرکلیدی از یک لغت‌نامه دارای افزونگی استفاده می‌کنند. این لغت‌نامه از ترکیب DCT مرسوم و لغت‌نامه محلی ساخته شده است. لغت‌نامه محلی به صورت جداگانه به ازای هر بلوک از فریم‌های غیرکلید ساخته می‌شود. این لغت‌نامه محلی در اصل از مجموع بلوک‌های مجاور در همسایگی w پیکسلی از بلوک جاری, در فریم‌های کلید قبل و بعد ساخته شده است. نحوه‌ی ساخت این لغت‌نامه در شکل ۱ توضیح داده شده است. ما از نماد ψ برای این لغت‌نامه ترکیبی استفاده خواهیم کرد. در حقیقت مزیت این لغت‌نامه در این است که اگر بلوکی در فریم غیرکلید فقط بر اثر حرکت در بلوک‌های مجاورش (در فریم‌های کلید) تکرار شده باشد, می‌توان با یک یا چند ضریب غیر صفر به‌راحتی آن را بیان کرد که بسیار به تنکی بردار ضرایب کمک خواهد کرد در نتیجه در زمان بازیابی بلوک از کیفیت بالاتری برخوردار خواهد بود. فرایند نمونه‌برداری در سمت کدگذار به این صورت است که هر بلوک در فریم غیرکلید، توسط یک لغت‌نامه مخصوص به خود تنک خواهد شد. سپس به جای سیگنال، بردار ضرایب تنکی تولیدی نمونه‌برداری خواهد شد. رابطه تنک‌سازی به صورت زیر می‌باشد:

$$x_{nk}^{i} = \psi\alpha \qquad (۸)$$

حال می‌توان از بردار ضرایب مانند سیگنالی در حسگری فشرده استفاده کرد. پس در این روش می‌توان تنک‌سازی را به عنوان یک فرایند پیش پردازش به سیستم اعمال کرد. در این مقاله از روش OMP برای به دست آوردن بردار ضرایب در سمت کدگذار استفاده شده است. باید به این نکته توجه کرد که برخلاف تمام روش‌های مرسوم از یک الگوریتم OMP اضافی در سمت کدگذار برای تولید بردار ضرایب استفاده شده است. OMP از یک حلقه ساده در ساختار خود استفاده کرده است که این مسئله به ساده بودن کدگذار صدمه‌ای نخواهد رساند. بعد از به دست آوردن بردار ضرایب طبق نظریه حسگری فشرده می‌توان نمونه‌برداری را از بردار ضرایب انجام داد. رابطه نمونه‌برداری از بردار ضرایب در زیر آمده است:

$$y_{nk} = \phi.\alpha \qquad (۹)$$

برخلاف بردار ضرایب $\hat{\alpha}$ که در فریم کلید استفاده شد این بردار ضرایب به صورت کامل نمونه‌برداری می‌شود. فرض کنید ما همسایگی w پیکسلی را برابر دو انتخاب کرده‌ایم و از دو فریم کلید قبل و بعد برای استخراج لغت‌نامه استفاده شده است، پس در حالت کلی ۵۰ بلوک همسایه به ازای هر بلوک خواهیم داشت. حال اگر بلوک‌ها ۳۲×۳۲ باشند اتم‌های لغت‌نامه DCT نیز برابر ۱۰۲۴ خواهد بود. در نهایت لغت‌نامه نهایی که از ترکیب این دو لغت‌نامه، ساخته خواهد شد دارای ۱۰۷۴ اتم خواهد بود. درنتیجه



(الف) (ب)

(ج) (د)

شکل شماره ۲ : مقایسه روش پیشنهادی در فریم ۶ ام در ویدئوی Container. (الف) تصویر اصلی، (ب) روش پیشنهادی.(ج) تصویر به دست آمده از مقاله [۲] و (د) روش مقاله [۱۳]

بردار ضرایب α نیز ۱۰۷۴ تایی خواهد شد. طول این بردار ضرایب از سیگنال اصلی بزرگتر است اما چون α برداری تنک می‌باشد طبق تئوری حسگری فشرده در بازیابی کیفیت مناسب‌تری خواهد داشت.

۴-۲- روش پیشنهادی در سمت کدگشا

در سمت کدگشا ابتدا فریم‌های کلید بعد و قبل هر فریم جداگانه بازیابی می‌شوند. بازیابی فریم‌های کلید با استفاده از روش حسگری فشرده مرسوم انجام می‌گیرد. برای این کار ابتدا باید بردار ضرایب $\hat{\alpha}$ را بازیابی کرد بعد از بازیابی، به اندازه ۴۰ درصد باقی مانده صفر به انتهای آن اضافه کرد تا α_i به دست آید. سپس لغت‌نامه را به همان ترتیبی که در کدگذار ساخته شد، مرتب می‌شود و با استفاده از این لغت‌نامه و بردار ضرایب α_i ساخته شده، می توان سیگنال اصلی را بازیابی کرد.

در فریم‌های غیرکلید با داشتن لغت‌نامه مرسوم DCT و بلوک‌های بازیابی شده در فریم‌های کلیدی قبل و بعد، می‌توان لغت‌نامه توضیح داده شده در سمت کدگذار یا ψ را دوباره تولید کرد و با استفاده از آن و رابطه بازیابی حسگری فشرده سیگنال اصلی را بازیابی کرد. توجه شود چون فریم‌های کلید قبل و بعد پیش از این بازیابی شده‌اند بنابراین نیازی به ارسال اطلاعات اضافی از سمت کدگذار نیست.

با توجه به حسگری فشرده ارائه شده در قسمت قبل، رابطه بازیابی نیز به صورت زیر تغییر خواهد کرد. مسئله بهینه‌سازی زیر را نیز می‌توان توسط روش BPحل کرد.

$$\alpha^* = argmin\ ||\alpha||_1\ \ s.t.\ \ ||y - \emptyset\alpha||_2 \leq \varepsilon \qquad (۱۰)$$

در نهایت با داشتن بردار ضرایب به دست آمده و همچنین لغت‌نامه ساخته شده با استفاده از رابطه زیر می‌توان سیگنال اصلی را بازیابی کرد.

$$\hat{x} = \psi\alpha^* \qquad (11)$$

۵- آزمایشات

در این بخش قصد داریم آزمایشات انجام شده بر روی روش پیشنهادی را مورد بررسی قرار دهیم. معیار کیفیت مطرح شده در این مقایسات PSNR و میزان نرخ بیت مصرفی می‌باشد .در فریم‌های کلید, اندازه بلوک برابر ۳۲×۳۲ و غیر همپوشان و همچنین از DCT مرتب شده به عنوان لغت‌نامه و از Noiselet به عنوان ماتریس حسگری استفاده شده است. در اینجا ما به صورت ثابت به فاصله هر ۱۰ فریم یک فریم کلید قرار داده‌ایم. ماتریس حسگری استفاده شده برای فریم‌های غیرکلید از نوع گوسی تصادفی انتخاب شده است.ویدئوهای مورد استفاده در مقایسات دارای نرخ ۳۰ فریم در ثانیه می‌باشند. در شکل شماره ۲ تصویر تولیدی در فریم ۶ام روش پیشنهادی با مقاله‌های [۱۳] و [۲] در ویدئوی Container با فرمت CIF مقایسه شده است. همین مقایسه با ویدئوی Foreman با فرمت CIF در شکل شماره ۳ آمده است.در نمودار شماره ۱ و ۲ به ترتیب مقایسه فشرده‌سازی ویدئوی Container و News با استفاده از روش پیشنهادی با مقاله [۲] و [۱۳] در نرخ بیت‌های مختلف آمده است. نرخ بیت‌های مختلف در حسگری فشرده با توجه به تغییر ابعاد ماتریس حسگری به دست می‌آید. ستون مشخص کننده کیفت بازیابی برحسب PSNR و سطر میزان بیت مصرفی می‌باشد. در نمودار شماره ۳ ویدئوی Foreman با فرمت CIF مورد مقایسه بین هر سه روش پیشنهادی، [۱۳] و [۲] قرار گرفته است. نتایج مورد بررسی با جدیدترین مقالات حسگری فشرده ویدئو نشان از برتری روش پیشنهادی در فشرده‌سازی و کیفیت بازیابی است. همان‌طور که از نمودارها پیداست و اینکه در عوض نمونه‌برداری سیگنال بردار ضرایب آن در لغت‌نامه نمونه‌برداری شده است، بنابراین اگر نرخ نمونه‌برداری پایین باشد به جای اینکه سیگنال تخریب شود بردار ضرایب آن تخریب شده است در نتیجه کیفیت بسیار افت خواهد کرد. اما همان‌طور که در نمودارها مشاهده می‌کنید با افزایش نرخ نمونه‌برداری چون α .نسبت به x تنک‌تر می‌باشد در نتیجه کیفیت بازیابی آن نیز بهتر خواهد بود.

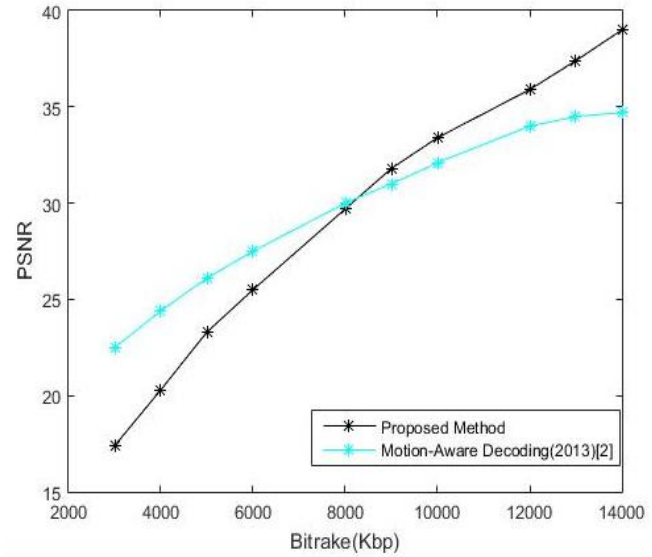
۶- نتیجه‌گیری

در این مقاله سعی شده است روشی نوین برای حسگری فشرده ویدئو ارائه شود که به جای نمونه‌برداری از اصل سیگنال از بردار ضرایب آن در فضای تنکی نمونه‌برداری کرده است. همچنین در فریم‌های کلید به جای استفاده از تمام اتم‌های دیکشنری DCT از اتم‌های مهم نمونه‌برداری شده است. در روش پیشنهادی در فریم‌های غیرکلید از دو لغت‌نامه به صورت همزمان استفاده شده است. لغت‌نامه محلی به صورت متغیر و به ازای هر بلوک جداگانه ساخته می‌شود و مزیت آن کمک فراوان به تنک‌تر کردن بردار ضرایب در حسگری فشرده می‌باشد و لغت‌نامه سراسری که متعامد و دارای قدرت بالایی در بازیابی است و در مواقعی که لغت‌نامه محلی به‌تنهایی قادر به نمایش تنک سیگنال نیست، استفاده خواهد شد.

(الف) (ب)

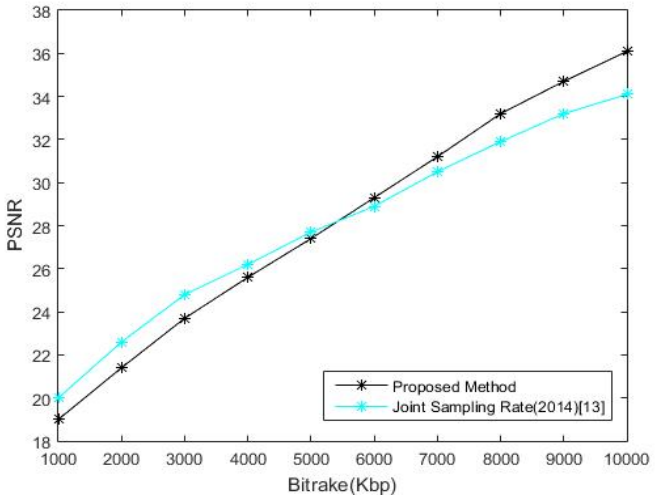
(ج) (د)

شکل شماره ۳ : مقایسه روش پیشنهادی در فریم ۶ ام در ویدئوی Foreman. (الف) تصویر اصلی، (ب) روش پیشنهادی.(ج) تصویر به دست آمده از مقاله [۲] و (د) روش مقاله [۱۳]

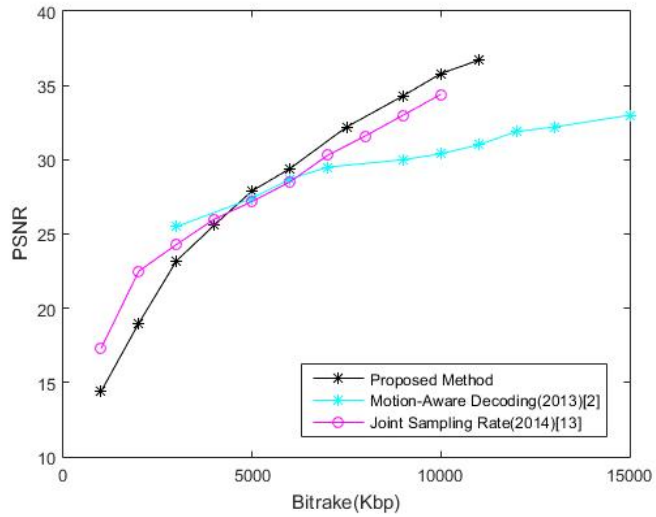


نمودار شماره ۱: مقایسه روش پیشنهادی در ویدئوی Container با مقاله [۲] با توجه میزان بیت مصرفی کیفیت بازیابی

[2] Y. Liu, M. Li, and D. A. Pados, "Motion-aware decoding of compressed-sensed video," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 23, pp. 438-444, 2013.
[3] W. Yan, Q. Wang, and Y. Shen, "Shrinkage-based alternating projection algorithm for efficient measurement matrix construction in compressive sensing," *Instrumentation and Measurement, IEEE Transactions on*, vol. 63, pp. 1073-1084, 2014.
[4] M. F. Duarte and Y. C. Eldar, "Structured compressed sensing: From theory to applications," *Signal Processing, IEEE Transactions on*, vol. 59, pp. 4053-4085, 2011.
[5] A. M. Tillmann and M. E. Pfetsch, "The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing," *Information Theory, IEEE Transactions on*, vol. 60, pp. 1248-1259, 2014.
[6] H. Rauhut, K. Schnass, and P. Vandergheynst, "Compressed sensing and redundant dictionaries," *Information Theory, IEEE Transactions on*, vol. 54, pp. 2210-2219, 2008.
[7] M. Elad, "Optimized projections for compressed sensing," *Signal Processing, IEEE Transactions on*, vol. 55, pp. 5695-5702, 2007.
[8] C. Li, H. Jiang, P. Wilford, Y. Zhang, and M. Scheutzow, "A new compressive video sensing framework for mobile broadcast," *Broadcasting, IEEE Transactions on*, vol. 59, pp. 197-205, 2013.
[9] W. Chen and M. R. Rodrigues, "Dictionary learning with optimized projection design for compressive sensing applications," *Signal Processing Letters, IEEE*, vol. 20, pp. 992-995, 2013.
[10] J. M. Duarte-Carvajalino and G. Sapiro, "Learning to sense sparse signals: Simultaneous sensing matrix and sparsifying dictionary optimization," DTIC Document2008.
[11] H. Liu, B. Song, H. Qin, and Z. Qiu, "Dictionary learning based reconstruction for distributed compressed video sensing," *Journal Of Visual Communication And Image Representation*, vol. 24, pp. 1232-1242, 2013.
[12] X. Gao, F. Jiang, S. Liu, W. Che, X. Fan, and D. Zhao, "Hierarchical frame based spatial–temporal recovery for video compressive sensing coding," *Neurocomputing*, vol. 174, pp. 404-412, 2016.
[13] H. Liu, B. Song, F. Tian, and H. Qin, "Joint sampling rate and bit-depth optimization in compressive video sampling," *Multimedia, IEEE Transactions on*, vol. 16, pp. 1549-1562, 2014.



نمودار شماره ۲: مقایسه روش پیشنهادی در ویدئوی News با مقاله [۱۳] با توجه میزان بیت مصرفی کیفیت بازیابی



نمودار شماره ۳: مقایسه روش پیشنهادی در ویدئوی Foreman با مقالات [۱۳] و [۲] با توجه میزان بیت مصرفی کیفیت بازیابی

مراجع

[1] U. Dias and M. E. Rane, "Comparative analysis of sensing matrices for compressed sensed thermal images," in *Automation, Computing, Communication, Control and Compressed Sensing (iMac4s), 2013 International Multi-Conference on*, 2013, pp. 265-270.