

ردگیری اشیاء در صحنه‌های ویدئویی شلوغ آلوده به اعوجاج کوانتیزاسیون

محمد خوانساری^۱ حمیدرضا ربیعی^۲ مجید اسدی^۱ محمد قنبری^{۱,۲,۳}
khansari@mehr.sharif.edu rabiee@sharif.edu asadi@dibe.unige.it ghan@essex.ac.uk

چکیده

این مقاله به معرفی یک روش مقاوم در برابر اعوجاج کوانتیزاسیون برای ردگیری اشیاء در صحنه‌های شلوغ می‌پردازد. الگوریتم از ضرائب تبدیل موجک بسته تلف نشده^۴ (UWPT) و تحلیل بافت برای ردگیری اشیاء دلخواه استفاده می‌کند. الگوریتم با تعیین ناحیه‌ای مستطیلی حول شیء موردنظر توسط کاربر مقداردهی اولیه می‌شود و سپس ضرائب UWPT ناحیه موردنظر برای تولید یک بردار خصیصه^۵ (FV) برای هر پیکسل درون ناحیه به کار گرفته می‌شود. جستجوی بهینه برای یافتن بهترین ناحیه معادل با استفاده از FV بردارهای خصیصه و یک پنجره جستجوی افقی صورت می‌گیرد. مقاوم بودن بردارهای خصیصه در برابر اعوجاج کوانتیزاسیون به واسطه قابی^۶ و تخمین جهت و سرعت حرکت شیء صورت می‌گیرد. مقاوم بودن بردارهای خصیصه در برابر اعوجاج کوانتیزاسیون به واسطه خاصیت نویززدایی در الگوریتم تولید خصیصه حاصل شده است.

نتایج آزمایشها نشان می‌دهد که الگوریتم پیشنهادی دارای کارایی خوبی برای ردگیری اشیاء در صحنه‌های شلوغ و دارای اعوجاج کوانتیزاسیون مثلاً بر روی پله‌ها در فرودگاه یا ایستگاه قطار با وجود جابه‌جایی، چرخش و مقیاس محدود می‌باشد.

کلمات کلیدی: ردگیری شیء، تبدیل موجک بسته تلف نشده، اعوجاج کوانتیزاسیون، صحنه‌های شلوغ، تحلیل بافت

Quantization Distortion Robust Object Tracking in Crowded Video Scenes

M. Khansari, H. R. Rabiee, M. Asadi, M. Ghanbari

Abstract

In this paper, we propose a new quantization noise robust algorithm for object tracking in the crowded video scenes. The algorithm exploits the properties of undecimated wavelet packet transform (UWPT) coefficients and texture analysis to track arbitrary objects. The algorithm is initialized by the user through specifying a region around the object of interest at the reference frame. Then, coefficients of the UWPT of the region construct a Feature Vector (FV) for every pixel in that region. Optimal search for the best match is then performed by using the generated FVs inside an adaptive search window. Adaptation of the search window is achieved by inter-frame texture analysis to find the direction and speed of the object motion. Robustness to quantization is due to denoising and edge component separation in the FV generation algorithm.

Experimental results show a good performance for object tracking in contaminated crowded scenes with quantization distortion, in particular crowds on stairs in airports or train stations in the presence of object translation, rotation, small scaling.

Keywords: Object Tracking, Crowded Scene, Undecimated Wavelet Packet Transform, Motion Direction, Texture Analysis

^۱ آزمایشگاه رسانه‌های دیجیتال، دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شریف

^۲ مرکز تحقیقات مخابرات ایران

^۳ دانشکده سیستمهای الکترونیکی، دانشگاه Essex

^۴ Undecimated Wavelet Packet Transform

^۵ Feature Vector

^۶ Interframe Texture Analysis

۱- مقدمه

یکی از مسائل عمومی در حوزه پردازش تصویر و ویدئوی کاربردی، چگونگی کاهش اثر انواع نویز است. نویز می‌تواند به شکل یک سیگنال ناخواسته در ارتباطات، اندازه‌گیری یا پردازش یک سیگنال حاوی اطلاعات ظاهر شود.

در سیستمهای ویدئویی عموماً به دو نوع نویز توجه ویژه‌ای می‌شود: نویز یا اغوج ناشی از عمل کوانتیزاسیون و نویز کانال. نویز کوانتیزاسیون معمولاً به صورت طبیعی در هنگام انجام عمل کوانتیزاسیون در کدکننده‌ها بوجود می‌آید. این نوع نویز در حالتی که نرخ ویدئو پائین باشد، خود را به خوبی به صورت بصری نشان می‌دهد. کدکننده‌های ویدئو عمدتاً از کوانتیزاسیون به عنوان شیوه‌ای برای ایجاد یک مصالحه بین نرخ فشرده‌سازی و حفظ کیفیت استفاده می‌کنند. این نوع نویز تاثیر مستقیمی بر روی کیفیت دیداری تصویر و یا ویدئوی فشرده شده دارد و در مواقعی که نرخ بیت ویدئو پائین باشد، به خوبی خود را نشان می‌دهد.

در رشته ویدئویی صحنه‌های شلوغ که در اثر عمل کوانتیزاسیون دچار این نوع نویز شده باشند، انجام عمل ردگیری اشیاء به دلیل مات شدن لبه‌ها، تغییر رنگ، حرکات غیر یکنواخت شیء و تغییر در پس‌زمینه و در نهایت وجود اشیاء متعدد مشابه (مانند سر و شانه انسان) کار بسیار پیچیده‌ای است.

لذا روشهایی که عموماً از خصوصیات مکانی^۱ تصویر مانند رنگ، بافت و لبه به تنهایی استفاده می‌کنند، معمولاً در مورد ردگیری اشیاء در صحنه‌های شلوغ با مشکلات عدیده‌ای مواجه هستند.

هیستوگرام رنگ مثالی از روشهای شناخته شده برای ردگیری اشیاء است. هیستوگرام رنگ معمولاً برای مدلسازی اشیاء برای مقابله با گرفتگی جزئی^۲ و نویز (تاحدی) و اشیاء صلب به کار می‌رود [۱، ۲]. از آنجایی که هیستوگرام فقط به توصیف توزیع کلی رنگ شیء می‌پردازد و جزئیات مکانی و چیدمان رنگها صرفنظر می‌کند، اشیایی که دارای رنگی مشابه به پس زمینه یا سایر اشیاء باشند را به سادگی گم می‌کند. علاوه بر این، این روشها نمی‌تواند به سادگی تغییر در روشنایی یا سایر اطلاعات مکانی شیء را تحمل کند. بنابراین توصیف ویژگی بر مبنای هیستوگرام رنگ برای ردگیری شیء در صحنه‌های شلوغ که در آن اشیاء مشابهی مانند سر افراد وجود دارند به احتمال زیاد جوابگو نخواهد بود.

به تازگی الگوریتمهای جابجایی میانگین^۳ که از هیستوگرام رنگ استفاده می‌کنند، تا حدی مشکلات فوق را برای ردگیری شیء کاهش داده‌اند [۱، ۳، ۴، ۵]، هرچند این شیوه‌ها نیازمند فیلتر کردن حرکت هستند که طراحی چنین فیلترهایی بسیار پیچیده است [۵].

هیستوگرام رنگ همچنین در چارچوبهای احتمالی مانند فیلترهای بیزی^۴ یا ذره‌ای^۵ به کار گرفته شده است [۸-۶، ۲]. ارزیابی

مقایسه‌ای روشهای مختلف ردگیری نشان می‌دهد که این خانواده از روشها در مقابل پدیده گرفتگی کوتاه‌مدت بهتر عمل می‌کنند [۹].

علاوه بر این، بیشتر روشهای ذکر شده معمولاً برای سناریوهای ساده‌ای آزمون شده‌اند: یک شخص از جلوی شخص دیگری عبور می‌کند [۲، ۸، ۹] یا یک بازیکن فوتبال با لباسهای رنگی شناخته شده توسط بازیکن تیم مقابل گرفته می‌شود و نه برای صحنه‌های پیچیده‌ای مانند جمعیت درون صحنه شلوغ ایستگاه مترو با اشیاء مشابه. در این صحنه‌های شلوغ، اشیاء مشابه دارای هیستوگرامهای رنگ مشابه هستند که باعث عدم دقت و اطمینان به ردگیری می‌شود. چالشهای دیگر روشهای ذکر شده مقاومت در برابر انواع نویز و پایداری خصیصه‌های انتخاب شده در حضور تبدیلات بر روی شیء است [۹].

در این مقاله به معرفی یک الگوریتم جدید مقاوم در برابر نویز کوانتیزاسیون برای ردگیری یک شیء تعریف شده توسط کاربر در صحنه‌های شلوغ با به کارگیری ویژگیهای تولید شده در حوزه UWPT می‌پردازد. الگوریتم پیشنهادی مبتنی بر بردارهای خصیصه تولید شده از طریق ضرائب تبدیل بسته موجک تلف نشده [۱۵، ۱۶] در درون یک پنجره وفقی است [۱۴-۱۲]. روش پیشنهادی مبتنی بر الگوریتم تطبیق بلوک در حوزه UWPT برای ردگیری اشیاء است که برای صحنه‌های شلوغ با حضور نویز کوانتیزاسیون تطبیق داده شده است. ویژگیهای کلیدی UWPT که آن را برای حل این مساله مناسب کرده‌اند عبارتند از افزونه بودن و در نتیجه تسهیل مقاومت در برابر نویز، جابه‌جانی پذیری^۶ و ارائه یک تقریب چگالتز نسبت به تبدیل موجک معمولی برای تبدیل موجک پیوسته [۱۰، ۱۱، ۱۷، ۱۸].

پس از این مقدمه، در ادامه الگوریتم پیشنهادی در فصل دوم تشریح شده است. بخش سوم نتایج آزمایشها و مقایسه با روشهای دیگر را نشان می‌دهد. در نهایت بخش چهارم به جمع‌بندی و ارائه کارهای آتی می‌پردازد.

۲- الگوریتم پیشنهادی

در الگوریتم پیشنهادی، ردگیری اشیاء بر مبنای ردگیری زمانی یک مستطیل محاط به شیء در قالب اصلی انجام می‌گیرد. الگوریتم از این جهت نیمه خودکار است که کاربر باید مستطیل موردنظر را که شامل شیء می‌باشد مشخص کند. دیاگرام بلوکی الگوریتم در شکل ۱ نشان داده شده است.

کاربر در ابتدا مستطیلی به شیء موردنظر مشخص محاط می‌کند، سپس بردار خصیصه برای هر پیکسل درون مستطیل با استفاده از ضرائب UWPT تولید می‌شود.

در مرحله بعد، ردگیری زمانی بردارهای خصیصه تولید شده در قالب مرجع، در قابهای بعد انجام می‌گیرد. پس از یافتن شیء در قالب جدید مکان پنجره جستجو بر مبنای تحلیل بافت بین قابی به هنگام می‌شود.

۱. تولید UWPT برای ناحیه r (توجه کنید که در صورت نیاز برای تولید UWPT، می‌تواند به اطراف ناحیه صفر اضافه شود).

۲. انتخاب پایه‌ها^۸ برای نمایش سیگنال از زیرباندهای بسط سیگنال در درخت UWPT با حرص درخت. روشهای مختلف برای حرص درخت UWPT ارائه شده است.

أ) به کاربردن الگوریتمهای مبتنی بر آنتروپی برای انتخاب پایه و حرص درخت. هدف این نوع انتخاب پایه حذف افزونگی ذاتی موجود در UWPT و فراهم کردن یک تقریب چگالتر از سیگنال اصلی است. از الگوریتمهای انتخاب پایه مبتنی بر آنتروپی معمولاً برای کاربردهای فشرده‌سازی استفاده می‌شود [۱۰، ۱۱].

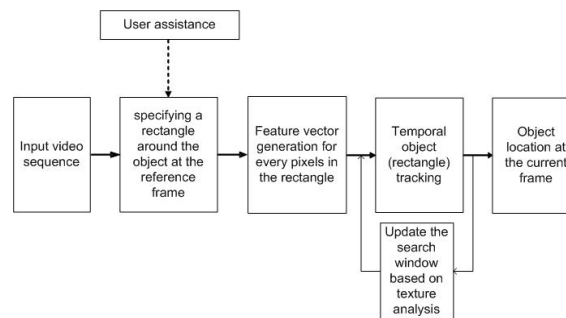
ب) انتخاب برگهای درخت بسط سیگنال برای نمایش آن. این نوع نمایش سیگنال شامل بیشترین تعداد زیرباندهاست که محاسبات سنگینی را در هنگام پردازش تحمیل می‌کند. این نمایش برای کاربرهایی که در یک ناحیه کوچک نیز به دقت بالا دارند (مانند ردگیری نقاط مرزی یک شیء) مناسب به نظر می‌رسد و برای کاربرد خاص ردگیری یک پنجره نسبتاً بزرگ توجه لازم را ندارد. برای مثال در شکل ۲،

$$x = (w_4^A, w_4^D, w_5^A, w_5^D, w_6^A, w_6^D, w_7^A, w_7^D)$$

ج) در مقوله ردگیری یک ناحیه مستطیلی، داشتن متوسط ناحیه می‌تواند یک ایده کلی برای تولید یک بردار خصیصه باشد. از آنجایی که زیرباندهای تقریب^۹، متوسطی از یک سیگنال بر مبنای تعداد سطوح یک درخت UWPT ارائه می‌دهند، حرص درخت نمایش سیگنال بر اساس این که بیشترین تعداد ضریب از زیرباند تقریب باشد، انجام می‌گیرد. این نوع از انتخاب مبنای وزن را بیشتر به ضرائب زیرباند تقریب می‌دهد که برای کاربرد مورد نظر مناسب است. برای مثال در شکل ۲، $x = (w_4^A, w_4^D)$.. برای کاربرد مورد نظر در این مقاله، این نوع از انتخاب معقول به نظر می‌رسد چراکه در بخش ردگیری زمانی مستطیل محاط بر شی در قابهای بعدی و مقایسه برای یافتن بهترین مستطیل مطابق با آن، متوسط بهترین معیار مقایسه خواهد بود.

خروجی این مرحله آرایه‌ای از شماره گره‌های درخت UWPT است که پایه‌ها را برای استفاده در قابهای بعدی مشخص می‌کند.

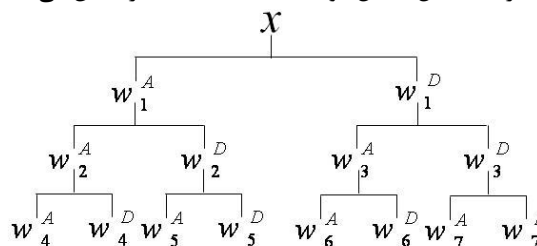
۳. بردار خصیصه برای پیکسل درون یک ناحیه مستطیلی r می‌تواند به سادگی با انتخاب ضرائب موجک در گره‌های پایه انتخاب شده در مرحله قبل تولید شود. بنابراین تعداد عناصر بردار خصیصه هر پیکسل برابر تعداد گره‌های پایه نمایش سیگنال است.



شکل ۱: دیاگرام بلوکی الگوریتم پیشنهادی برای ردگیری اشیاء در صحنه‌های دارای نویز کوانتیزاسیون

۲-۱- تولید بردار خصیصه

تبدیل موجکی که هم‌خطی است و هم خاصیت جابه‌جا ناپذیری با زمان دارد، تبدیل موجک بسته تلف نشده (UWPT) است [۱۰، ۲۰]. از دیدگاه پیاده‌سازی این تبدیل با استفاده از بانکهای فیلتر، علاوه بر فیلترکردن زیرباند پایین (همانند تبدیل موجک معمولی)، بر روی باند بالا نیز عملیات فیلتر کردن صورت می‌گیرد. به علاوه، این تبدیل به دلیل عدم وجود نمونه‌برداری به پایین^۷، یک تبدیل افزونه است. شکل ۲ یک درخت کامل تبدیل موجک بسته تلف نشده را نشان می‌دهد.



شکل ۲: درخت بسط سیگنال یک بعدی x توسط تبدیل موجک تلف-نشده

بنابراین تبدیل UWPT دارای دو خاصیت است که آن را برای تولید خصیصه‌های تغییر ناپذیر و مقاوم در حوزه پردازش تصویر و ویدئو مناسب می‌کند:

- تغییر ناپذیری، که بر مبنای آن در صورتی که یک بردار خصیصه بر مبنای ضرائب UWPT در قاب t ساخته شود، این بردار خصیصه می‌تواند در قالب $t+1$ یافت شود.
- همه زیرباندها در درخت تفکیک دارای اندازه یکسانی برابر با اندازه قاب اصلی هستند. این ویژگی باعث سادگی فرآیند استخراج خصیصه می‌شود.

همچنان که در شکل ۲ نشان داده شده است، نمایشهای متفاوت افزونه‌ای برای یک سیگنال x با ترکیبهای متفاوت زیرباندها وجود دارد و به عنوان مثال $x = (w_2^A, w_2^D, w_1^D)$ ، $x = (w_1^A, w_1^D)$ ، $x = (w_4^A, w_4^D, w_2^D, w_1^D)$ همگی نمایشهای سیگنال x هستند.

رویه تولید FV برای هر پیکسل در ناحیه r (شامل شیء) در قالب t می‌تواند در مراحل زیر خلاصه شود:

۲-۲- ردگیری زمانی

هدف از ردگیری زمانی، مکانیابی شیء مورد نظر در قابهای متوالی بر مبنای اطلاعات موجود در قالب جاری و قاب مرجع است. همچنان که در بخش قبل ذکر شد، برای هر پیکسل درون ناحیه مستطیلی حول شیء یک بردار خصیصه معادل وجود دارد. از این بردار خصیصه برای یافتن بهترین ناحیه مطابق در قالب(های) بعد استفاده می‌شود. فرآیند تطبیق ناحیه r در قاب t به ناحیه $r+1$ در قاب $t+1$ از طریق جستجوی کامل ناحیه r در پنجره جستجو در قاب $t+1$ انجام می‌شود، مکان این پنجره جستجو به صورت وقتی توسط رهیافت تحلیل بافت که در بخش ۲-۳ معرفی خواهد شد، تعیین می‌گردد [۱۵، ۱۶].

به صورت دقیقتر، هر پیکسل درون ناحیه r می‌تواند تحت تبدیلهای متفاوت و پیچیده‌ای در قابهای بعدی قرار بگیرد. در حالت کلی، یافتن مکان جدید یک پیکسل با استفاده از ویژگیهای حوزه مکانی مانند روشنایی، بافت و غیره که معمولاً متغیر و حساس هستند مشکل است. رهیافت پیشنهادی برای ردگیری ناحیه r در قاب n از بردارهای خصیصه ذکر شده برای هر پیکسل و فاصله اقلیدسی برای یافتن بهترین ناحیه مطابق با آن به شرح ذیل استفاده می‌کند:

۱. تولید بردار خصیصه برای هر پیکسل درون ناحیه مستطیلی r و پنجره جستجو توسط رویه ارائه شده در بخش ۲-۱

۲. جاروب پنجره جستجو توسط یک ناحیه مستطیلی به ابعاد ناحیه r

۳. یافتن بهترین مطابق برای r در پنجره جستجو با محاسبه حداقل جمع فواصل اقلیدسی بین بردارهای خصیصه برای هر پیکسل درون ناحیه جستجو جاری و بردار خصیصه پیکسلهای درون ناحیه r (جستجوی کامل) رویه جستجو برای یافتن بهترین ناحیه مطابق، شبیه به الگوریتم عمومی تطبیق بلوک است؛ با این تفاوت که از بردار خصیصه تولید شده برای هر پیکسل به جای روشنایی استفاده می‌شود. بنابراین، وقتی که بعضی از پیکسلها r در قاب بعدی ظاهر نشود (به دلیل گرفتگی جزئی یا تغییرات)، الگوریتم پیشنهادی می‌تواند بهترین مطابق برای r را در قاب بعدی بر مبنای رویه جستجو فوق پیدا کند.

۲-۳- مکانیزم به روزرسانی پنجره جستجو

ردگیری مکان شیء هدف نیازمند یک مکانیزم وقتی به روزرسانی پنجره جستجو به سه دلیل است:

۱. موقعیت مناسب پنجره جستجو اطمینان می‌دهد که شیء مورد نظر همواره در درون ناحیه جستجو قرار دارد.

۲. پنجره با ابعاد ثابت ولی موقعیت وقتی باعث کاهش پیچیدگی محاسباتی ناشی از وجود یک پنجره بزرگ با اندازه متغیر می‌شود [۱۳].

۳. در صورتی که شیء مورد نظر دچار گرفتگی توسط اشیاء دیگر شود، به کارگیری جهت حرکت باعث کاهش مساله گرفتگی می‌شود.

برای رسیدن به یک شیوه کارا برای به‌روزرسانی مکان پنجره جستجو روشهای متفاوتی می‌تواند به کار گرفته شود. بیشتر این شیوه‌ها از خصیصه‌های زمانی یا مکانی برای هدایت پنجره جستجو و یافتن بهترین تطبیق با حداقل محاسبات استفاده می‌کنند [۱۶]. ما دو شیوه برای به‌روزرسانی پنجره در نظر گرفتیم:

۴. به هنگام سازی مرکز پنجره جستجو در قاب‌بندی بر مبنای مکان مستطیل حول شیء در قاب جاری. در این شیوه مکان پنجره جستجو ثابت نیست و در هر قاب جدید به مرکز مستطیل مطابق در قاب قبلی تنظیم می‌شود. این شیوه ساده است ولی باعث از دست دادن شیء به دلیل انتشار خطا می‌شود [۱۴]. به علاوه در صورت وجود گرفتگی شیء در قاب جاری، در قابهای بعدی شیء به درستی قابل یافتن نخواهد بود.

۵. رهیافت دیگر تخمین جهت و سرعت حرکت شیء برای به هنگام سازی مکان پنجره جستجو است.

در این مقاله از شیوه دوم به کارگیری تکنیک تحلیل بافت بین قابی^{۱۰} استفاده شده است [۱۵، ۱۶].

برای یافتن جهت و سرعت حرکت شیء، هیستوگرام اختلاف زمانی^{۱۱} دو قاب جاری تعریف می‌شود. جهت و زبری^{۱۲} اختلاف بین دو قاب از روی هیستوگرام اختلاف زمانی استخراج است. در نهایت، جهت و سرعت حرکت با استفاده از هیستوگرام اختلاف زمانی، زبری و جهت تخمین زده می‌شود.

۲-۳-۱- هیستوگرام اختلاف زمانی

هیستوگرام اختلاف زمانی دو قاب متوالی از قدرطلق تفاضل مقادیر روشنایی پیکسلهای متناظر در دو قاب به دست می‌آید. فرض کنید $SA_t(x, y)$ پنجره جستجو در قاب t و $SA_{t+1}(x, y)$ پنجره جدید جستجو در قاب $t+1$ باشد که جابه‌جایی مرکز آن نسبت به پنجره جستجو در قاب t به اندازه $\delta = (\Delta x, \Delta y)$ باشد. لازم به توضیح است که ابعاد هر دو پنجره یکی است. قدرمطلق اختلاف زمانی ATD_δ بین دو پنجره با فرمول زیر تعریف می‌شود.

$$ATD_\delta(x, y) = \sum_{y=0}^{N_y-1} \sum_{x=0}^{N_x-1} |SA_t(x, y) - SA_{t+1}(x + \Delta x, y + \Delta y)| \quad (2.3.1.1)$$

که در آن N_x و N_y به ترتیب عرض و ارتفاع پنجره جستجو هستند. سپس هیستوگرام مقادیر ATD_δ محاسبه می‌شود. این هیستوگرام دارای M خانه است که M تعداد سطوح خاکستری در هر قاب می‌باشد مثلاً (۲۵۶ برای یک تصویر ۸ بیتی). در نهایت، مقادیر هیستوگرام نسبت به تعداد پیکسلهای درون پنجره جستجو $(N_x * N_y)$ نرمالیزه می‌شود که در نتیجه چگالی احتمالی هر مقدار خاکستری محاسبه می‌شود $(p_\delta(i) \quad i = 0, \dots, M - 1)$.

$$TCON = \sum_{i=0}^{M-1} i^2 p_{\delta}(i) \quad (2.3.3.1)$$

که در آن M تعداد سطوح خاکستری در هر قاب مطابق با ۱، ۲، ۳ است. پارامتر $TCON$ یک معیار کمی برای زبری بافت ارائه می‌دهد و مقدار آن به مقدار تغییرات محلی در ناحیه مورد نظر بستگی دارد. وجود تغییرات محلی زیاد در یک قاب، دلیلی بر وجود فعالیت در قاب است. به این قاب در مقایسه با قابهایی با تغییرات کم، فعال^{۱۵} گفته می‌شود.

از آنجایی که قابهای فعال یک رشته ویدئو مقدار زیادی تغییرات محلی را نشان می‌دهند، کنتراست زمانی که از روی تفاضل قابها به دست می‌آید به فعالیت تصویر مرتبط است. با هدف کاهش اثر اندازه و بافت بر روی پنجره جستجو، پارامتر $TCON$ با مقدار کنتراست محلی ($LCON$) نرمالیزه می‌شود. $LCON$ که واریانس پیکسل در پنجره جستجو (SW) را تعریف می‌کند توسط فرمول زیر تعریف می‌شود:

$$LCON = \frac{1}{SW} \sum_{SW} [g(x,y) - \bar{g}]^2 \quad (2.3.3.2)$$

که $g(x,y)$ مقدار سطح خاکستری در مکان پیکسل قرار گرفته در موقعیت (x,y) است و \bar{g} متوسط مقدار سطح خاکستری در پنجره جستجو است. بر مبنای کنتراستهای محلی و زمانی، فرمول زیر تخمین خوبی از متوسط سرعت حرکت (S) در درون بلوک ارائه می‌دهد:

$$S = k \frac{TCON}{LCON} \quad (2.3.3.3)$$

که در آن K مقدار ثابت تجربی است. متوسط سرعت حرکت در معادله (۲،۳،۳،۳) تنها به اندازه شیء در حال حرکت بستگی ندارد بلکه نسبت به جهت بافت نیز تغییر ناپذیر است. مقدار S برای بخشهای ایستا در تصویر مانند پس زمینه (مستقل از بافت آن) به صفر میل می‌کند. مقدار جابه‌جایی پنجره جستجو برای قاب بعدی مطابق با فرمول زیر محاسبه می‌شود:

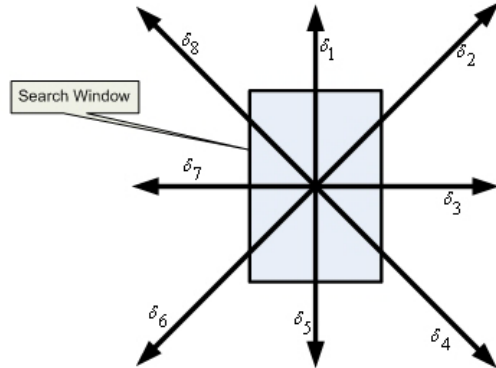
$$R_{j-1} = S_{j-1} - Disp_{j-1} \quad (2,3,3,4)$$

$$Disp_j = \lfloor S_j + R_{j-1} \rfloor$$

در بعضی از قابهای همگن است که مقدار S کمتر از یک باشد، بنابراین مقدار جابه‌جایی پنجره جستجو صفر خواهد بود. پارامتر R_{j-1} باقیمانده جابه‌جایی در قاب قبلی نشان می‌دهد. در حالتی که جابه‌جایی شیء دارای سرعت کمی باشد، پارامتر R_{j-1} کمک می‌کند که مقادیر جابه‌جایی کمتر از یک پیکسل جمع شود تا در نهایت منجر به جابه‌جایی حداقل یک پیکسلی پنجره جستجو شود.

۲-۳-۲- جهت پنجره جستجو

فرض کنیم پنجره جستجو یک بلوک مستطیلی باشد. هشت بلوک دیگر که فاصله مراکز آنها از بلوک اصلی δ_i است را مطابق با شکل ۳ در نظر می‌گیریم:



شکل ۳: بررسی فاصله‌ها در جهات مختلف برای یافتن مقدار بیشینه IDM

برای هر بلوک، p_{δ_i} (هیستوگرام اختلاف زمانی) نسبت به پنجره مرجع محاسبه می‌شود. سپس IDM_i را برای هر بلوک مطابق با فرمول (۲،۳،۲،۱) حساب می‌شود. IDM یک معیار همگن مطابق زیر است:

$$IDM = \sum_{i=0}^{M-1} \frac{p_{\delta}(i)}{i^2 + 1} \quad (2.3.2.1)$$

در یک تصویر همگن p_{δ_i} دارای تعداد کمی مقادیر بزرگ است. بنابراین IDM شامل اطلاعاتی در مورد توزیع احتمالی مقادیر غیرصفر p_{δ_i} می‌باشد و می‌تواند برای تعیین جهت بافت اصلی به کار گرفته شود. اگر یک بافت، جهت‌دار باشد در یک جهت زبرتر از بقیه باشد، درجه پخش شدن مقادیر p_{δ_i} مطابق با جهت δ_i (با فرض قرار گرفتن مقادیر آن در یک بازه مناسب) است. بنابراین جهت‌دار بودن بافت به وسیله مقایسه پخش بودن مقدار p_{δ_i} برای جهت‌های مختلف δ قابل تحلیل است.

برای استخراج جهت حرکت از جهت بافت، مقدار بیشینه IDM مطابق با فرمول زیر محاسبه شود.

$$IDM_{max} = \max \{ IDM_i \}, \quad i = 1, 2, \dots, 8 \quad (2.3.2.2)$$

مقدار بیشینه IDM_{max} نشان می‌دهد که اختلاف بین دو قاب در آن جهت از بقیه جهت‌ها همگنتر است و به طور ضمنی بیان می‌کند که بلوک متناظر در آن جهت همبستگی^{۱۳} بیشتری با بلوک اصلی دارد.

۲-۳-۳- جابه‌جایی پنجره جستجو

معیار کمی برای تعیین زبری بافت کنتراست زمانی^{۱۴} به صورت گشتاور دوم p_{δ} حول مبدا تعریف می‌شود:

۳- نتایج آزمون

در کل آزمونها فرض شده است که تغییر صحنه¹⁶ وجود ندارد؛ بدیهی است که در صورت وجود تغییر صحنه قاب مرجع و شیء هدف باید به هنگام شوند و یک تعامل جدید با کاربر الزامی است.

ارزیابیهای انجام گرفته در این مقاله بر اساس هر دو گونه معیار ذهنی¹⁷ و معیار عینی است. نتایج روش ردگیری پیشنهادی با روشهای شناخته شده ردگیری مبتنی بر هیستوگرام رنگ با دو معیار فاصله مختلف یعنی کی‌دو¹⁸ و باتاچاریا¹⁹ مقایسه شده است. بنابراین کلیه اشکال این فصل در سه بخش آمده است: (a) ردگیری مبتنی بر هیستوگرام رنگ با معیار فاصله کی‌دو (با مخفف (b, CHC)، ردگیری مبتنی بر هیستوگرام رنگ با معیار فاصله باتاچاریا (با مخفف (c, CHB)، روش پیشنهادی (با مخفف (d, WBMA)).

مجموعه کامل نتایج نیز در قالب کلیه‌های ویدئویی تهیه شده است که از طریق اینترنت²¹ قابل دسترس است. علاوه بر این معیارهای ذهنی، یک معیار عینی برای ارزیابی کیفیت ردگیری تعریف و به کار گرفته شده است. این معیار فاصله اقلیدسی بین مرکز ثقل شیء ردگیری شده و شیء واقعی در قاب جاری است. در ابتدای ردگیری و در قاب مرجع، مرکز مستطیل ردگیری حول شیء منطبق بر مرکز ثقل شیء مورد نظر است. در قابهای بعدی، فاصله بین مرکز مستطیل ردگیری شده با مرکز ثقل شیء واقعی در همان قاب که به صورت دستی تعیین و ثبت شده است، محاسبه می‌شود.

برای تولید درخت UWPT از موجکهای پایه بای‌اورتوگونال²² که برای شناسایی شیء مفید است، استفاده شده است. وجود Spike در موجک پایه بای‌اورتوگونال آن را برای کاربردهای شناسایی و ردگیری مناسب کرده است [۱۹]. در تمامی آزمایشها، از یک درخت تفکیک سه سطحی UWPT تولید شده توسط [Bior2.2] [۱۷] استفاده شده است. در N_r, N_d به ترتیب تعداد گشتاورهای صفر²³ در فیلترها سنتز و آنالیز را نشان می‌دهند [۱۷]. در پیاده‌سازی الگوریتم ردگیری مبتنی بر هیستوگرام رنگ، تعداد جعبه‌های رنگ²⁴، ۳۲ عدد فرض شده است.

برای ارزیابی الگوریتم در یک محیط واقعی، کلیه‌های ویدئویی از متروی تهران پس از اخذ مجوزهای لازم از مراجع ذیصلاح گرفته شد. این کلیه‌های ویدئویی جمعیت مردم را در بخشهای مختلف مترو تهران مانند سوار و پیاده شدن از مترو و بالا و پائین رفتن از پله برقی نشان می‌دهد. به علاوه کلیه‌های مذکور شرایط مختلف در صحنه‌های شلوغ و پرجمعیت مانند گرفتگی جزئی و کلی، سرعت کم و زیاد، گرفتگی با زمان متغیر و زوم به داخل و بیرون، تغییر شکل شیء و چرخش آن را نشان می‌دهد. در کلیه تصاویر نمونه نشان داده شده، مستطیل‌های با خطوط توپر مستطیل حول شیء و مستطیل‌های با خطوط خط‌چین نشان دهنده پنجره جستجو می‌باشد. سختی ردگیری سر افراد در صحنه‌های شلوغ به دلیل وجود سرهای مشابه در نزدیکی هم از نکات بارز در صحنه‌های انتخاب شده است.

شکل ۴ نتیجه ردگیری یک شخص را هنگام بالارفتن از پله برقی و دور شدن از دوربین را نشان می‌دهد. قاب #۹۷، قاب مرجع و آلوده به نویز است. اعوجاج کوانتیزاسیون در قاب مرجع این رشنه ویدئویی دارای PSNR=31.95 dB و نرخ ۰.۲۲ bit/pixel است برای بهتر نشان دادن مقاومت روش پیشنهادی در مقابل نویز، از شیء بزرگی با مستطیل محاط شونده ۱۶*۲۳ پیکسل و پنجره جستجوی به ابعاد ۶۹*۶۳ پیکسل استفاده شده است. پارامترهای تجربی برای یافتن جهت و سرعت و حرکت به منظور به روزسانی پنجره جستجو $d=1$ و $k=6$ هستند.

شیء هدف از پله‌ها با یک سرعت ثابت بالا می‌رود و حرکت و جابه‌جایی و با مقداری زوم به بیرون، چرخش سر و گرفتگی جزئی (برای مثال در قاب #۱۵۶) همراه است. برای نشان دادن اثر بلاکی ناشی از نویز کوانتیزاسیون و نمایش بهتری بصری نتایج، شکل‌های ردیف اول از بالا در شکل ۴-۱ از طرفین برش داده شده‌اند. همچنان که در این ریف از شکل ۴ دیده می‌شود، هر دو روش CHC و CHB هدف را در صورت وجود نویز کوانتیزاسیون در قاب گم می‌کنند و روش پیشنهادی به خوبی از عهده ردگیری شیء بر می‌آید.

همچنان که در شکل ۴-۲ نشان داده شده است، در قاب #۱۰۲ هر دو روش CHC و CHB هدف را گم می‌کنند و به صورت تصادفی در قاب بعدی به آن بر می‌گردند از این قاب به بعد این دو روش مبتنی بر هیستوگرام رنگ بین اشیاء مشابه دیگر و هدف اصلی به صورت مداوم و تصادفی جابه‌جا می‌شوند. این رفتار، ناپایداری، بی‌ثباتی و عدم اطمینان خروجیهای اینگونه الگوریتمها را برای کاربرد ردگیری در صحنه‌های شلوغ نویزی نشان می‌دهد. علاوه بر این، از قاب #۱۲۸ تا پایان رشته در قاب #۱۶۰ نیز هر دو روش CHC و CHB به دلیل وجود نویز، گرفتگی جزئی و زوم به بیرون، شیء هدف را به طور کامل گم می‌کنند.

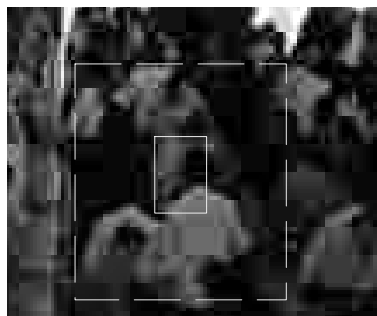
بر خلاف هر دو روش قبلی، الگوریتم پیشنهادی به خوبی از عهده ردگیری شیء حتی با حضور چرخش و گرفتگی جزئی در رشته ویدئویی مورد بررسی بر می‌آید (شکل ۴). در نتایج آزمون، الگوریتم پیشنهادی کارایی خیلی بالاتری نسبت به سایر روشهای مبتنی بر هیستوگرام رنگ نشان می‌دهد. این موضوع دلیلی بر مقاوم بودن بردارهای خصیصه پیشنهادی برای کاربردهای ردگیری اشیاء در محیطهای نویزی است. ذکر این نکته مفید است که در قابهای آخر رشته ویدئویی به دلیل بزرگ بودن پنجره جستجو و اندازه مستطیل حول شیء، گرفتگی کامل شیء و تفاوت عمده ظاهری بین شیء در قاب مرجع و قاب انتهایی، نتایج ردگیری دقیق نخواهد بود. همچنانکه در شکل ۴ نشان داده شده است، وجود نویز باعث کاهش شدید کارایی روشهای مبتنی بر هیستوگرام رنگ می‌شود، در حالی که بر روی پیشنهادی تاثیر ندارد.

(۱)

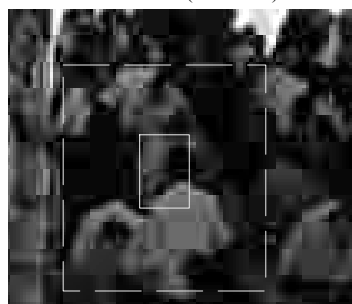


Frame #97(reference)

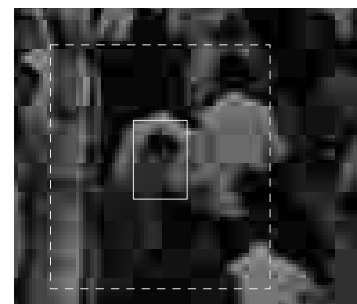
(۱)



a) Frame #108



b) Frame #108



c) Frame #108

(۱)



c) Frame #156

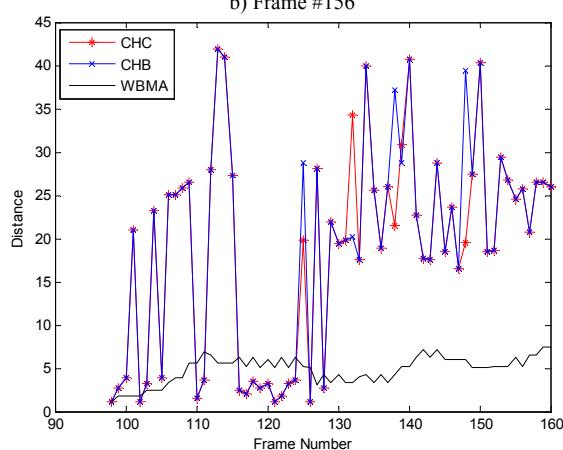


b) Frame #156



c) Frame #156

(۲)



شکل ۴:۱) ردگیری مردی که از پله‌های برقی درون جمعیت و گرفتگی جزئی و زوم به بیرون بالا می‌رود. (تصاویر دارای اعوجاج کوانتیزاسیون به مقدار PSNR=31.95dB و ۰.۲۲ CHC؛ (a) CHC، (b) CHB، (c) WBMA؛ (۲) فاصله بین مراکز مستطیل نتیجه شده از ردگیری و مرکز مستطیل مورد انتظار در هر قاب

خصیصه مقاوم برای هر پیکسل و پنجره جستجوی وقتی، روش پیشنهادی به خوبی می‌تواند تبدیلات پیچیده بر روی شیء مانند انتقال، چرخش اندک، مقیاس و گرفتگی جزئی را در تعداد معقولی قاب متوالی تحمل کند. علاوه بر این الگوریتم در مقابل نویز کوانتیزاسیون مقاوم است و امکان پیاده‌سازی بلادرنگ آن نیز وجود دارد. نتایج آزمایشها کارایی الگوریتم را در ردگیری شیء در صحنه‌های شلوغ حاوی

۴- نتیجه‌گیری و کارهای آتی

یک الگوریتم وقتی برای ردگیری شیء در صحنه‌های شلوغ با حضور نویز کوانتیزاسیون بر مبنای خصیصه‌های حوزه موجک برای پیکسل و یک شیوه سریع برای به هنگام‌سازی پنجره جستجو بر مبنای تحلیل بافت پیشنهاد شده است. بر مبنای خواص UWPT، وجود بردارهای

- [12] M. Amiri, H. R. Rabiee, F. Behazin, M. Khansari, "A new wavelet domain block matching algorithm for real-time object tracking," IEEE ICIP, September 14-17, Barcelona, Spain, pp. 961-4, 2003.
- [13] M. Khansari, H. R. Rabiee, M. Asadi, M. Ghanbari, M. Nosrati, M. Amiri, "A Semi-Automatic Video Object Extraction Algorithm Based on Joint Transform And Spatial Domain Features," CBMI, International Workshop on Content-Based Multimedia Indexing, Riga, Latvia, June 21-23, 2005.
- [14] M. Khansari, H. R. Rabiee, M. Asadi, M. Ghanbari, M. Nosrati, M. Amiri, "A Shape Tracking Algorithm Based on Generated Pixel Features by Undecimated Wavelet Packet," CSICC, 10th Annual Computer Society of Iran Computer Conference, Tehran, Iran, 2005.
- [15] M. Khansari, H. R. Rabiee, M. Asadi, P. Khadem Hamedani, M. Ghanbari, "Adaptive Search Window for Object Tracking in the Crowds using Undecimated Wavelet Packet Features," WAC, World Automation Congress, Budapest, Hungary, July 24-26, 2006.
- [16] V. E. Seferidis and M. Ghanbari, "Adaptive Motion Estimation Based on Texture Analysis," IEEE Transactions on Communications, vol. 42, no. 2-4, pp. 1277-1287, 1994.
- [17] Daubechies, Ten lectures on wavelets, CBMS, SIAM, 61, pp. 271-280, 1994.
- [18] H. Guo, "Theory and applications of the shift invariant, time-varying, and undecimated wavelet transform," Master's Thesis, Rice University, Houston, TX, 1995.
- [19] R. N. Strickland and H. I. Hahn, "Wavelet transform methods for object detection and recovery," IEEE Transaction on Image Processing, vol. 6, no. 5, pp. 724-735, May 1997.

نویز ناشی از عمل کوانتیزاسیون در مقایسه با الگوریتمهای مبتنی بر هیستوگرام رنگ نشان می‌دهد.

از مسائل باز این تحقیق، تناوب و نحوه به روزرسانی بردارهای خصیصه در هنگام ردگیری و استفاده از فیلتر کالمن برای پیشگویی ناحیه جستجو و بهبود مکانیزم به روزرسانی پنجره جستجو در مواقعی است که شیء مورد ردگیری خارج از پنجره جستجو قرار بگیرد.

سپاسگزاری

این تحقیق با حمایت مالی مرکز فناوری اطلاعات و ارتباطات پیشرفته دانشگاه صنعتی شریف (AICTC) و مرکز تحقیقات مخابرات ایران (ITRC) انجام گرفته است.

مراجع

- [1] D. Comaniciu, V. Ramesh, and P. Meer, "Kernel-Based Object Tracking," IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 5, pp. 564-577, May 2003.
- [2] P. Perez, C. Hue, J. Vermaak, and M. Gangnet, "Color-based probabilistic tracking," In Proc. European Conf. on Computer Vision, Copenhagen, Denmark, vol. 1, pp. 661-675, 2002.
- [3] D. Comaniciu and P. Meer, "Mean shift: A Robust Approach Toward Feature Space Analysis," IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 5, pp. 603-619, May 2002.
- [4] Z. Zivkovic, B. Krose, "An EM-like algorithm for color-histogrambased object tracking," Proceedings of the International Conference on Computer Vision and Pattern Recognition, 2004.
- [5] Dong Xu, Yimin Wang, and Jinwen An, "Applying a New Spatial Color Histogram in Mean-Shift Based Tracking Algorithm," Image and Vision Computing New Zealand, 2005.
- [6] A.D. Jepson, D.J. Fleet, T.F. El-Maraghi, "Robust Online Appearance Models for Visual Tracking," Proceedings of the International Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 415-422, 2001.
- [7] A. Jacquot, P. Sturm O. Ruch, "Adaptive Tracking of Non-Rigid Objects Based on Color Histograms and Automatic Parameter Selection," IEEE Workshop on Motion and Video Computing (WACV/MOTION'05), vol. 2, pp. 103-109, 2005.
- [8] B. Han, C. Yang, R. Duraiswami and L. Davis, "Bayesian Filtering and Integral Image for Visual Tracking. Invited to special session of Real-Time Object Tracking," Algorithms and Evaluation in Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), Montreux, Switzerland, 2005.
- [9] B. Deutsch, Ch. Gr̄aßl, F. Bajramovic, J. Denzler, "A Comparative Evaluation of Template and Histogram Based 2D Tracking Algorithms," Pattern Recognition, 27th DAGM Symposium, Springer, pp. 269-276, 2005.
- [10] R. R. Coifman and M. V. Wickerhauser, "Entropy-based algorithms for best basis selection," IEEE Transaction on Information Theory, Special Issue on Wavelet Transforms and Multiresolution Signal Analysis, Vol. 38, pp. 713-718, Mar. 1992.
- [11] K. Ramchandran and M. Vetterli, "Best wavelet packet bases in a rate distortion sense," IEEE Transaction on Image Processing, Vol. 2, pp. 160-175, Apr. 1993.

زیر نویس‌ها

- ¹ Spatial
² Partial Occlusion
³ Mean-Shift
⁴ Bayesian
⁵ Particle
⁶ Shift Invariant
⁷ Down sampling
⁸ Basis
⁹ Approximation
¹⁰ Inter-frame texture analysis
¹¹ Temporal Difference Histogram
¹² Coarseness
¹³ Correlation
¹⁴ Temporal Contrast
¹⁵ Active
¹⁶ Scene Cut
¹⁷ Subjective
¹⁸ Chi-Square
¹⁹ Bhattacharyya
²⁰ Wavelet Block Matching Algorithm
²¹ <http://mehr.sharif.edu/~khansari>
²² Biorthogonal
²³ Vanishing Moments
²⁴ Color Bins