

افزایش سرعت شناسایی زمینه ویدئو با محاسبه مرحله‌ای تجزیه QR

محمود امین طوسی و سمیه پورصدیق*

دانشکده ریاضی و علوم کامپیوتر، دانشگاه حکیم سبزواری
m.amintoosi@hsu.ac.ir
pourseddigh@gmail.com

چکیده. شناسایی زمینه در ویدئو از مسائل مهم پردازش تصویر و بینایی ماشین است که تحقیقات مختلفی را به خود معطوف نموده است. در برخی از روشهای مورد استفاده در این حوزه از مقادیر قطر اصلی ماتریس R در تجزیه QR برای شناسایی فریمهای زمینه استفاده شده است. یکی از مشکلات این شیوه محاسبه مجدد تجزیه QR در بازه‌های زمانی پی‌درپی و حجم محاسباتی بالای آن است. در این مقاله با تغییر نحوه چیدمان ماتریس ورودی و محاسبه عناصری از ماتریسهای Q و R که با ورود فریم جدید ویدئو تحت تاثیر قرار می‌گیرند، حجم محاسبات کم شده و زمان پردازش از نظر تئوری به نصف کاهش پیدا می‌کند. نتایج آزمایشات انجام شده، کاهش حدود ۳۵ درصدی زمان اجرا را در عمل نشان می‌دهد.

۱. پیش‌گفتار

شناسایی نواحی متحرک در ویدئو یا پیش زمینه (Foreground) از قسمت ثابت صحنه یا زمینه (Background) یکی از مراحل اساسی در بسیاری از مسایل حوزه پردازش تصویر و بینایی ماشین می‌باشد که تحقیقات متعددی را به خود معطوف نموده است. تفاضل هر فریم ویدئو از زمینه تصویر یک راهکار معمول برای شناسایی اشیاء متحرک است. در بسیاری موارد زمینه ثابت فیلم از قبل مشخص نیست و باید تخمین زده شود. به این زمینه ثابت، تصویر زمینه یا مدل زمینه گفته می‌شود. تخمین دقیق‌تر این مدل منجر به شناسایی بهتر اشیاء خواهد شد که روشهای متعددی در این زمینه پیشنهاد شده است [۱]. یکی از روشهای مواجهه با مسئله استفاده از تجزیه QR است که در [۲] به آن پرداخته شده است. ایده اصلی در شیوه فوق‌الذکر تبدیل فریمهای ویدئو به یک ماتریس و شناسایی فریمهای زمینه از طریق میزان وابستگی خطی ستونهای ماتریس بوده است. حتی در وضعیتی که دوربین ثابت باشد و هیچ شیء متحرکی در صحنه وجود نداشته باشد، فریمهای این ویدئو کاملاً با هم برابر نبوده و مقادیر پیکسلها تا حدودی با هم متفاوت هستند. اگر هر فریم مثلاً با ابعاد ۲۴۰×۳۲۰ به صورت یک بردار ستونی با $۷۶۸۰۰ = ۳۲۰ \times ۲۴۰$ عنصر دربیاید، هر فریم را می‌توان یک بردار در فضای ۷۶۸۰۰ تصور کرد. نزدیک بودن بردارها

2010 Mathematics Subject Classification. Primary 68T45; Secondary 15A23, 68U10.

واژگان کلیدی. تجزیه QR، الگوریتم گرام اشمیت، شناسایی زمینه.

* سخنران

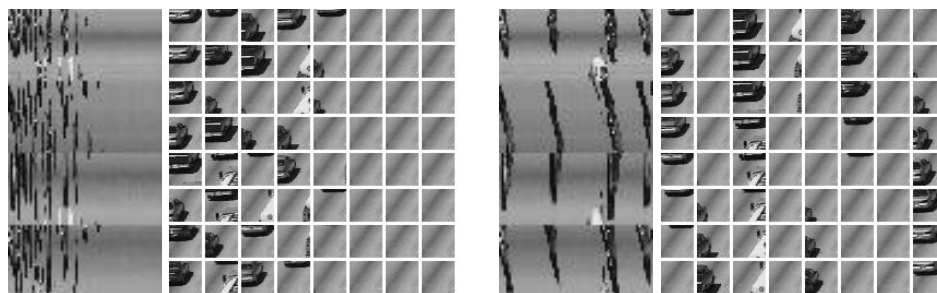


شکل ۱: یک فریم از یک ویدیوی نمونه، زمینه استخراج شده بر اساس روش مورد استفاده در مقاله و نتیجه شناسایی اشیاء متحرک مبتنی بر زمینه بدست آمده.

در فضا می‌تواند به این معنی باشد که هر دو بردار (فریم) محتوای نسبتاً یکسانی دارند. در این مقاله فرض بر این است که دوربین ثابت بوده و اشیاء متحرک در صحنه توقف طولانی ندارند؛ مانند ویدیوهای اخذ شده از دوربین‌های نظارتی اتوبانها که شکل ۱ یک نمونه را نشان می‌دهد. فرض کنیم بردارهایی که در واقع متعلق به زمینه بوده‌اند به صورت پراکنده در ماتریس A قرار گرفته‌اند. با این فرض که شیء متحرک هم‌رنگ با زمینه نیست، یکی از بردارهای متعلق به زمینه به عنوان یک پایه متعامد در ماتریس Q از تجزیه QR ماتریس A ظاهر خواهد شد و سایر ستونهای متعامد یک ماتریس Q که مرتبط با این بردارها هستند مقدار R (روی قطر اصلی) بسیار کوچکی خواهند داشت. به این ترتیب می‌توان این فریمها را شناسایی و مدل زمینه را بر اساس آنها ایجاد کرد.

در بسیاری از مواقع همانند شکل ۱ هیچ وقت کل صحنه خالی از اشیاء متحرک نیست، لذا به جای اعمال روش فوق روی کل فریم آن را روی بلاکهای کوچکتری از تصویر اعمال نموده و نتایج را در کنار هم قرار می‌دهیم. شکل ۲ یک بلاک از ۶۴ فریم متوالی ویدیوی شکل ۱ به همراه ماتریس ایجاد شده از کنار هم قرار دادن ستونی‌شده‌ی این فریمها را نشان می‌دهد. در این مثال اندازه هر بلوک ۳۰ در ۳۰ پیکسل بوده است و لذا A یک ماتریس ۹۰۰×۶۴ خواهد بود. برای نمایش بهتر فقط ۱۲۰ سطر ماتریس A (متناظر با عناصر روی قطر اصلی، قطر فرعی و سطر و ستون وسط هر بلاک) نمایش داده شده است. ترتیب جدید این بلاکها بر اساس مقادیر R تجزیه QR ماتریس مربوط به آنها در شکل ۳ نشان داده شده است. همان‌گونه که مشاهده می‌شود در ترتیب جدید بلاکهای متناظر با زمینه تصویر به انتهای لیست منتقل شده‌اند. نمایش ستونی این بلاکها در تصویر سمت چپ شکل ۳ مؤید همین مطلب است که ستونهای مرتبط با زمینه تصویر به دلیل وابستگی زیادشان به هم در انتها قرار گرفته‌اند.

عمل شناسایی فریمهای زمینه می‌تواند روی ماتریس حاصل از کل ویدیو و یا روی یک برش از ویدیو انجام شود. عموماً در کاربردهای برخط به جای اعمال روش روی کل ویدیو، یک پنجره لغزان با طول ثابت روی ویدیو در نظر گرفته می‌شود و مدل زمینه بر اساس فریمهای موجود در این پنجره برآورد می‌گردد. سپس فریمهای موجود در پنجره بروزرسانی شده و مجدداً عملیات تکرار می‌شود. به عنوان مثال اگر طول پنجره N فریم باشد، عملیات برای فریمهای ۱ تا N انجام شده (ماتریس A و مدل زمینه ساخته می‌شود) سپس با دریافت فریمهای جدید، ماتریس بر اساس فریمهای ۲ تا $N + ۱$ ساخته شده و عملیات ادامه پیدا می‌کند.



شکل ۲: توالی یک بلاک تصویر در ۶۴ فریم متوالی (از چپ به راست و از بالا به پایین) و بخشی از ماتریس A حاصل از آن (اولین ستون متناظر با اولین بلاک، دومین ستون متناظر با دومین بلاک و تا آخر).

شکل ۳: بازترتیب بلاکهای شکل ۲ بر اساس مقادیر R از تجزیه QR به همراه ماتریس A که ستونهای آن بر اساس مقادیر روی قطر ماتریس R از تجزیه QR ماتریس A در شکل ۲ مرتب شده‌اند.

۲. دستاوردهای پژوهش

مشکلی که در روش فوق‌الذکر وجود دارد نیاز به محاسبه مجدد تجزیه QR در هر پنجره می‌باشد. اگر ماتریس حاصل از فریمهای یک تا N را A و ماتریس حاصل از فریمهای دو تا $N + 1$ را A_1 بنامیم، تنها اختلاف آنها ستونهای اول و آخر آنهاست. در واقع $N - 1$ ستون اول ماتریس A_1 ، همان $N - 1$ ستون آخر A هستند و آخرین ستون A_1 فریم جدیدی است که وارد شده است.

اگر بتوان از خروجی تجزیه A تجزیه A_1 را با عملیات محاسباتی کمتری بدست آورد، عمل تخمین زمینه هم با سرعت بیشتری انجام خواهد گرفت. الگوریتم‌های متعددی برای روش گرام اشمیت در منابع مختلف ذکر شده است [۳، ۴]؛ از این میان الگوریتم ارائه شده در [۳] برای منظور موردنظر ما مناسب تشخیص داده شد که در الگوریتم ۱ نشان داده شده است. برای اصلاح آن به نحوی که بتوان با داشتن تجزیه QR ماتریس A ، تجزیه ماتریس A_1 را بدست آورد، فقط کافیت مقداردهی اولیه ماتریس v را به صورت مناسب انجام داد و در اولین حلقه الگوریتم، به جای شروع از اولین ستون، از ستون i ام عملیات را ادامه دهیم. الگوریتم ۲ وضعیت جدید را نشان می‌دهد. به این ترتیب ورودی الگوریتم ۲ به جز ماتریس ورودی، ماتریسهای Q و R بدست آمده از تجزیه ماتریس فریمهای قبلی (A) و i ، اندیس آخرین فریم اضافه شده می‌باشد.

اگر فریم جدید در ستون آخر قرار گیرد و ستون اول حذف شود، مقدار i در الگوریتم ۲ همیشه باید یک باشد و لذا دو الگوریتم تفاوتی با هم نخواهند داشت، به منظور رفع این مشکل و حفظ سازگاری الگوریتم ۲ با الگوریتم ۱، ماتریس A از آخر به اول پر می‌شود، به این صورت که آخرین ستون آن، اولین فریم ویدیو و اولین ستون آن، آخرین فریم ویدیو ما در پنجره جاری خواهد بود. فریم جدید ورودی در یک صف اولین ورودی خروجی جایگزین اولین فریم قبلی (آخرین ستون ماتریس) خواهد شد.

الگوریتم ۲ الگوریتم گرام اشمیت مرحله‌ای	الگوریتم ۱ الگوریتم گرام اشمیت
Input: A_1, Q, R, i Output: Q, R 1: $v \leftarrow [Q(1:m, 1:i-1) A(1:m, i:n)]$ 2: for $k=i, \dots, n$ do 3: for $j=1, \dots, k-1$ do 4: $r_{jk} \leftarrow \langle v_k, v_j \rangle$ 5: $v_k \leftarrow v_k - v_j r_{jk}$ 6: end for 7: $r_{kk} \leftarrow \ v_k\ _2$ 8: $v_k \leftarrow v_k / r_{kk}$ 9: end for	Input: $A_{m \times n}$ Output: Q, R 1: $v \leftarrow Q$ 2: for $k=1, \dots, n$ do 3: for $j=1, \dots, k-1$ do 4: $r_{jk} \leftarrow \langle v_k, v_j \rangle$ (همان q_j هست v_j) 5: $v_k \leftarrow v_k - v_j r_{jk}$ 6: end for 7: $r_{kk} \leftarrow \ v_k\ _2$ 8: $v_k \leftarrow v_k / r_{kk}$ 9: end for

در یک فرآیند دوری، هر زمان که فریم جدیدی وارد می‌شود، در سمت چپ ستون قبلی در ماتریس قرار می‌گیرد و پس از رسیدن به ابتدای ماتریس، فریم جدید در آخرین ستون قرار می‌گیرد. اگر مرتبه زمانی محاسبه هر بردار پایه متعامد ماتریس Q را $O(1)$ فرض کنیم، مرتبه زمانی الگوریتم ۱ برای محاسبه N ستون ماتریس Q ، $O(N)$ و لذا مرتبه زمانی محاسبه ماتریس Q برای تمام N فریم یک پنجره، $O(N^2)$ می‌باشد. حال اگر محاسبات مورد نیاز بر طبق الگوریتم اصلاح شده ۲ انجام شود، اگر فریم وارد شده i امین فرم ورودی باشد (البته پس از یک بار پر شدن پنجره)، مرتبه زمانی الگوریتم $O(i)$ خواهد بود. چون i در هر دور پر شدن پنجره مقادیر یک تا N را بخود می‌گیرد، پس مرتبه زمانی انجام محاسبات برای N فریم یک پنجره با الگوریتم دوم، $\sum_{i=1}^N i = N(N+1)/2$ خواهد بود که همان‌گونه که دیده می‌شود، معادل نصف زمان لازم برای انجام محاسبات با الگوریتم ۱ است.

نتایج برنامه‌های نوشته شده در MATLAB روی چند نمونه ویدیو نشان داد که در عمل به دلیل بالاسری مربوط به سایر دستورات برنامه شناسایی زمینه، روش استخراج مدل زمینه با الگوریتم دوم به صورت میانگین ۳۵ درصد نسبت به الگوریتم ۱ کاهش پیدا کرده است. خروجی شناسایی زمینه و شناسایی اشیاء متحرک برای یک فریم از یک ویدیوی نمونه در شکل ۱ آمده است.

هنوز در الگوریتم ۲ محاسبات صرفاً بر اساس فریم جدید و تجزیه ماتریس قبلی انجام نمی‌شود و به قسمتی از ماتریس ورودی هم وابسته است، به عنوان کارهای آتی می‌توان به حذف وابستگی محاسبات به ماتریس ورودی و لذا کاهش مرتبه زمانی آن از $O(i)$ به $O(1)$ اشاره نمود.

مراجع

- [1] Sobral, A., Vacavant, A.: A comprehensive review of background subtraction algorithms evaluated with synthetic and real videos Computer Vision and Image Understanding, **122**, 4 - 21, (2014).
- [2] Amintoosi, M., Farbiz, F., and Fathy, M.: A QR Decomposition based mixture model algorithm for background modeling. ICICS2007, Sixth International Conference on Information, Communication and Signal Processing, 1-5, (2007).
- [3] Watkins, D. S.: Fundamentals of Matrix Computations, (3rd ed.), (2010).
- [4] Golub, G. H.; Van Loan, C. F.: Matrix Computations, (3rd ed.), Johns Hopkins, (1996).