

در ابتدا، روشی برای استخراج وابستگی میان فریم‌های متوالی ارائه دهید.

در این الگوریتم برای دقت بالاتر از چندین کانال از ویژگی های تصویر استفاده میکنیم مانند کانال های رنگی (11 رنگ) به منظور یافتن و تشخیص طیف رنگی جسم تعقیب شده و استفاده از hog به منظور درک بهتر ساختار و مرز بندی جسم و ایجاد پایداری در شرایط مختلف روشنایی که در فرایند تعقیب بسیار مفید خواهند بود .

برای تعلیم و آموزش فیلتر از رابطه زیر استفاده می کنیم :

$$\min_{h_k} \sum_{\mathbf{p}} |(x_k \star h_k)(\mathbf{p}) - y(\mathbf{p})|^2 + \lambda |h_k|^2$$

که در آن :

x_k : مقدار بدست آمده از هر کانال فریم اولیه (جسم تشخیص داده شده اولیه)

y_k : مقدار هدف که در اینجا از فیلتر گاوسی با قله منطبق بر مرکز جسم

h_k : فیلتر آموزش دیده

λ : پارامتری برای تنظیم میزان یادگیری فیلتر

به منظور سرعت بخشیدن به محاسبات ، محاسبات رابطه روبرو را در حوزه فرکانس انجام میدهیم یعنی :

$$\mathcal{E}(h) = |\sum_{k=1}^d (x_k \star h_k) - y|^2 + \lambda \sum_{k=1}^d |h_k|^2 \quad \text{مسئله در حوزه زمان :}$$

$$\mathcal{E}(H) = |\sum_{k=1}^d (X_k \odot H_k) - Y|^2 + \lambda \sum_{k=1}^d |H_k|^2 \quad \text{مسئله در حوزه فرکانس :}$$

ساده سازی رابطه هر فیلتر در حوزه فرکانس :

$$L = \left(\sum_k X_k H_k - Y \right)^* \left(\sum_j X_j H_j - Y \right) + \lambda \sum_k H_k^* H_k$$

مشتق گیری برای بهینه بودن پاسخ فیلتر بدست آمده برای کانال خاص $k = m$:

$$\Rightarrow \frac{\partial L}{\partial H_m^*} = X_m^* \left(\sum_j X_j H_j - Y \right) + \lambda H_m = 0$$

$$H_m = \frac{X_m^* (Y - S)}{\lambda} \Rightarrow S = \sum_j X_j H_j = \sum_j X_j \left(\frac{X_j^* (Y - S)}{\lambda} \right)$$

$$\Rightarrow S = \frac{Y - S}{\lambda} \sum_j (X_j X_j^*)$$

پس از جای گذاری و ساده سازی نهایی به رابطه نهایی زیر میرسیم :

$$H_m = \frac{X_m^* \odot Y}{\sum_j (X_j^* \odot X_j) + \lambda}$$

حال پس از بدست آمدن فیلتر ها برای یافتن نقطه مرکز جسم در فریم جدید از رابطه زیر استفاده می کنیم که در آن مختصات نقطه ماکسیمم R مرکز جسم جدید است :

$$R = \sum_k (x_k \star h_k)(x, y)$$

حاصل کانولوشن در پیکسل به مختصات (x, y)

الگوریتم بروزرسانی فیلتر

برای بروزرسانی فیلتر میتوان از رابطه زیر استفاده کرد با استفاده از پارامتر μ که مقدار 0.1 میتواند مناسب باشد برای این کار :

$$H_i = \frac{A_i}{B_i}$$

$$A_i = \mu X_i^* \odot Y + (1 - \mu)A_{i-1}$$

$$B_i = \mu \left(\sum_j (X_j^* \odot X_j) + \lambda \right) + (1 - \mu)B_{i-1}$$

بروزرسانی اندازه جعبه به صورت داینامیک :

برای این کار از روش DSST استفاده می کنیم در این روش از جسم تشخیص داده شده در فریم کنونی چند مقیاس در اندازه های مختلف مانند 1.1 و 1.2 و 0.9 و ... تولید میکنیم و از هر مقیاس مستطیلی به اندازه مشخص به مرکز جسم تشخیص داده شده جدا میکنیم و فیلتری یکسان روی همه آنها اعمال میکنیم سپس نتایج بدست آمده را با نتیجه اعمال فیلتر روی جسم در فریم قبل مقایسه میکنیم هر کدام که کمترین تفاوت را داشت نشان دهنده مقیاس جسم جدید است تنها کافی است اندازه مستطیل جدید متناسب با مقیاس بدست آمده باشد .

*** نکته قابل توجه این است که برای افزایش سرعت محاسبات و باتوجه به طولانی تر بودن فرایند کانولوشن دو بعدی نسبت به کانولوشن تک بعدی باید از ویژگی های تک بعدی جسم مانند hog استفاده کنیم که یک بردار یک بعدی به ما می دهد.

نحوه برخورد جسم با شرایطی که جسم محو می شود

برای چنین شرایطی از فیلتر کالمن استفاده میکنیم به این صورت که در فریم کنونی اطلاعات حرکتی مرکز جسم در تصویر را می یابیم و بوسیله روابط سینماتیک مرکز جسم را در فریم های بعدی پیش بینی میکنیم با این کار میتوان در صورت محو شدن جسم موقعیت تقریبی آن را حدس زد و به محض پدیدار شدن مجدد جسم باز هم آن را تعقیب کنیم.

از روابط زیر برای فیلتر کالمن استفاده میکنیم :

$$\hat{x}_k = F_k \hat{x}_{k-1}$$
$$P_k = F_k P_{k-1} F_k^T$$

که در آنها :

\hat{x}_k : ماتریس حالت در لحظه k ام

F_k : ماتریس انتقال حالت در لحظه k ام مطابق با روابط سینماتیک

P_k : ماتریس کوواریانس یا همبستگی در لحظه k ام

روابط بالا برای حالت ایده آل (بدون در نظر گرفتن عوامل خارجی) است حال اگر اقر عوامل خارجی مانند نویز سنسور ها را در نظر بگیریم ، مرکز پیش بینی شده از روابط زیر بدست می آید :

$$\widehat{\mathbf{x}}'_k = \widehat{\mathbf{x}}_k + \mathbf{K}'(\mathbf{z}_k - \mathbf{H}_k \widehat{\mathbf{x}}_k)$$
$$\mathbf{P}'_k = \mathbf{P}_k - \mathbf{K}' \mathbf{H}_k \mathbf{P}_k$$
$$\mathbf{K}' = \mathbf{P}_k \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k \mathbf{H}_k^T + \mathbf{R}_k)^{-1}$$

که در آنها :

\mathbf{K}' : ماتریس بهره کالمن

$\widehat{\mathbf{x}}'_k$: ماتریس حالت پس از بروزرسانی در لحظه k ام

H_k : ماتریس مشاهده در لحظه k ام (ماتریسی برای انتخاب داده های مورد استفاده بعد از اندازه گیری)

R_k : ماتریس کوواریانس (میزان همبستگی) مشاهدات در لحظه k ام

Z_k : ماتریس اندازه گیری های انجام شده در لحظه k ام (همان مقدار بدست آمده توسط فیلتر ها یعنی همان مقدار حقیقی)

خلاقیت ها و نکات پروژه :

- برای اعمال فیلتر روی هر فریم به جای اعمال فیلتر بر روی کل عکس ، با توجه به مرکز جسم پیش بینی شده توسط فیلتر کالمن ، فیلتر را روی مستطیلی با ابعاد 2 تا 1.5 برابر مستطیل اولیه اعمال میکنیم در این حالت هم مشکل تعقیب در حالت تغییر سرعت سریع جسم حل می شود هم محاسبات ما کاهش می یابد.
- برای قسمت محو شدن جسم ، علاوه بر پیش بینی موقعیت جسم توسط فیلتر کالمن ، در این حالت در تمام فریم ها از تشخیص دهنده مانند yolo استفاده میکنیم که به محض پدیدار شدن دوباره جسم بتوان ادامه تعقیب را انجام داد .
- برای بروزرسانی دقیق تر فیلتر در هر فریم میتوان همزمان از تشخیص دهنده *yolo* نیز استفاده کرد و در الگوریتم بروزرسانی ضربی برای خطای فیلتر روی الگوریتم بروزرسانی اعمال کرد تا بتون خطای فیلتر را نیز کاهش داد .