Efficient Histogram-based Occluded Object Segmentation in Tracking System

Moonyong Jin¹, Kiseon Jeong² and Dong Sun Park³
¹, ²Electronic and Information Engineering, Chonbuk National University, Jeonju-si, Jeonbuk, Korea
³IT Convergence Research Center, Jeonju-si, Jeonbuk, Korea

Abstract - Unsatisfying tracking results may be produced due to occlusion problem. Especially, system performance such as speed and accuracy in multiple-object tracking is dependent on how to deal with occlusion situation. In the paper, we proposed an efficient technique to estimate potential occlusion and to separate each occluded object based on color histogram matching between prior objects and occluded object. The results show that the proposed algorithm successfully segments objects under occlusion situation and is five times faster than template matching method.

Keywords: Multiple-object tracking, segmentation, occlusion, histogram, template matching

1 Introduction

Visual object tracking is one of the most important subjects in computer vision [1]. Many researchers have extensively studied applications dealing with human-computer interaction such as surveillance systems and augmented reality. Especially, intelligent surveillance system to prevent illegal parking and crime has attracted attention of the people. In order to analyze the behavior of people and understand the scenes, we need multiple-object tracking where targets and observations need to be correctly matched from frame to frame in a video sequence. Through object tracking, we can obtain useful information about objects.

There are many difficulties which we have to solve in the object tracking. The challenges in tracking include change of illumination, occlusion problem, scale change and object perspective. Among these difficulties, one of major challenge in multi-object tracking is occlusion. It is dependent on tracking performance. When the occlusion occurs, we can lose information about objects and fail to track. The occlusion in tracking is unavoidable challenge and we have to certainly overcome over. To deal with the occlusion problem, researchers did works and got many achievements. Qian Zhang et al. [2] use depth information, color and motion cues to segment occluded objects with multiple cameras. Bo Wu et al. [3] design the part hierarchy with silhouettes of objects and match the detection responses and object edgelets with visible silhouettes. Mohammad Azari et al. [4] present their proposed occlusion reasoning method that use template matching based on correlation coefficient method. Sohaib Khan et al. [5] presented a framework to track people during occlusion. First, they segment a person into classes of similar color using the Expectation Maximization (EM) algorithm. Then they use a maximum a posteriori probability approach to track these classes from frame to frame.

In the whole tracking system, just object tracking without occlusion is doing well. However, there are still occlusion problems. For example, template matching is time consuming because there are many computations that calculate each template and searching region to match. Besides, they need many templates to identify object since the shape of template changes when people are moving. In the case of using multiple cameras, they also are time-consuming as well as need more than 2 cameras and Graphic Processing Unit to make real time system even though is able to segment object well. For these reasons, this paper considers time performance and accuracy of object identification as well as focuses on dealing with how to segment objects under occlusion.

One of the useful features which express characteristic of object is color information. The color is very powerful information to represent objects and processing time with color is faster than method of template matching that use object identification. Hence, we present the method to match objects using color histogram as well as mean-shift clustering algorithm that helps object segmentation more robustly with color information. Also, another important feature is that we can see the whole front object under occlusion. If we know front object when objects occluded, we can easily obtain front object because color distribution of front object is not changed. We are going to use these two features to segment objects. Figure.1 is overall proposed tracking system framework.

This paper proceeds by introducing background subtraction using Gaussian mixture model in Section II. Then in Section III, the Kalman Filer to predict and update statement as well as Euclidean Distance and ratio of overlap area to track object are presented, followed by object
segmentation algorithm using color histograms of prior objects before occlusion and occlude object in Section IV. The experimental results in Section V show that the proposed algorithm is able to ensure efficient and robust object segmentation and has good performance for time and accuracy comparison to existing method. Finally, conclusion is in Section VI.

2 Background subtraction

2.1 GMM (Gaussian Mixture Model)

In the computer vision, background subtraction using GMM [6] is often used. GMM (1) is to make background model using more than two Gaussian distributions. In this method, the background model is statistically modeled on each pixel. In order to adapt to pixel value changes, we can use the GMM which update the training set by adding new samples and discarding the old ones. Also, we can estimate parameters such as weight, mean, covariance by EM algorithm [7]. In next step, the difference between the background model and the current image is evaluated.

$$p(X \mid \mu, \Sigma) = \sum_{k=1}^{K} \pi_k N(x \mid \mu_k, \Sigma_k)$$  \hspace{1cm} (1)
3 Object tracking by Kalman Filter

3.1 Kalman Filter

Kalman Filter [8] is a tool that can estimate the variables of a wide range of processes and used in the computer vision, robot, satellite and control area. Mathematically, we would say that a Kalman filter estimates the states of a linear system and minimizes error through predict and remove noise using measured prior data and new data. In other words, The Kalman filter estimates next state and is predict system based on probability.

3.1.1 State equation and measurement equation

\( x_k = Ax_{k-1} + w_{k-1} \) (2)
\( z_k = Hx_k + v_k \) (3)
\( w_{k-1} \sim N(0,Q) \) (4)
\( v_{k-1} \sim N(0,R) \) (5)

3.1.2 Prediction and Update

There are two steps in Kalman filter algorithm. First step is prediction (6), (7) and second step is update (8), (9), (10). In the prediction step, we estimate current state predict and error covariance matrix to get the a priori \( \hat{p}_x \) estimates for the next step. In the update step, we obtain Kalman gain and current state vector and the estimation error covariance matrix.

\( \hat{x}_k = Ax_{k-1} + w_k \) (6)
\( \hat{p}_x = AP_{k-1}A^T + Q \) (7)
\( K_k = P_kH^T(HP_kH^T + R)^{-1} \) (8)
\( x_k = x_{k-1} + (z_k - Hx_k) \) (9)
\( P_k = (I - K_kH)P_k \) (10)

3.2 Kalman Filter in tracking system

In our tracking system, we use position \((a_k, b_k)\), velocity \((\dot{a}_k, \dot{b}_k)\) and size \((w_k, h_k)\) of object as state vector \(x_k = [a_k, b_k, \dot{a}_k, \dot{b}_k, w_k, h_k]^T\) and position and size of object as measurement vector \(y_k = [a_k, b_k, w_k, h_k]^T\). Hence, the matrix and are defined as

\[
A = \begin{bmatrix}
1 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\]

\[
H = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}
\]

3.3 Occlusion detection

Step 1. By using Kalman filter prediction of position and size of objects with Euclidean distance between prior objects and current object, we predict whether occlusion happens or not. If the predicted state estimates occluded state, we predict that there is possible occlusion in the next frame.

Step 2. If occlusion is detected in the next frame, we make histogram of the occluded object.

4 Histogram based occluded object segmentation algorithm

Histogram is a useful tool to analyze images, objects and video information. The histogram can be used to represent the color distribution of an object. In the tracking, we should choose invariant or regular pattern as feature. One of powerful feature is color to represent objects since objects have own color information. Also, processing time using color is faster than other methods such as template and stereo matching. We already extracted objects through background subtraction and can express color information as histograms that uses Hue of HSV color space to make robust illumination change. Also, another important feature is to use that we can see whole front object even though occlusion occurs. It is useful to get object segmentation under occlusion since color distribution of front object is not changed. By using that, we can automatically segment front and back object.

![Figure 5. Proposed Algorithm to segment objects](image-url)
4.1 Comparison between histograms

Through comparison between Bhattacharyya distance (11) of previous objects and occluded object, the front object is found. The distance $d_{Bhattacharyya}$ between m bins histogram $H_1$ and $H_2$ is computed as

$$d_{Bhattacharyya}(H_1, H_2) = \sqrt{1 - \frac{\sum_i \sqrt{H_1(i)H_2(i)}}{\sum_i H_1(i)\sum_i H_2(i)}}$$ (11)

If Bhattacharyya distance of matching scores is low, it indicates good matches. The range of matching score is from 0 to 1.

4.2 Back projection

Back projection is a way of recording how well the pixels fit the distribution of pixels in a histogram model. We use the back projection of the color histograms.

4.3 Mean shift filtering

Mean shift algorithm [9] is a robust method of finding local maxima of the probability density modes given by samples. Mean shift runs as follows.

**Mean Shift Procedure**

1. Choose a search window
2. Determine a center of the data
3. Center the window at the center of the data
4. Return until convergence

Mean shift filter finds the peaks of color distributions over space. In other words, this filter can find the highest density “clumps” of data in the space by scanning a window over the space. Next, all the points traversed by the windows that converge at a peak in the data become connected by that peak. Therefore, radiating out from the densest peaks forms the segmentation of the image.

4.4 Post-processing

Finally, we segment object after postprocessing such as morphology and threshold method.

5 Experimental results and analysis

We use PETS 2010 datasets [10] in our experiments and have used OpenCV [11] library. In order to make robust illumination change, we obtain color histogram using Hue after converting RGB to HSV and we use mean-shift clustering algorithm as preprocessing to effectively segment occluded objects. In this paper, the proposed algorithm has 95% average identification rate of front object and 77.80% average precision rate between pixels of original object and segmented object and is five times faster than template matching method.
comparing color histograms of objects. Secondly, through back projection technique after mean-shift clustering, we obtain the front object. In the template matching method case, it is time consuming since there are many computations that calculate each template and searching region to match when exploration window is large. Besides, they need some templates to identify object because the shape of template changes. On the other hand, the experimental results have demonstrated the efficiency and effectiveness of our method. Time performance of the proposed method is better than template method approximately 5 times as well as identification rate of front object and segmentation precision rate are 95% and 77.80% respectively.

However, the color feature we use is unreliable sometimes. This method may produce wrong correspondence when different objects have similar color. In our future work, we are going to study a new feature that represents object appearance and make more powerful object segmentation accuracy with time performance.

7 Acknowledgement

This research was financially supported by the Ministry of Education, Science Technology (MEST) and National Research Foundation of Korea (NRF) through the Human Resource Training Project for Regional Innovation and by Business for Cooperative R&D between Industry, Academy, and Research Institute funded Korea Small and Medium Business Administration in 2011 (Grants No. 00047720).

8 References


6 Conclusions

In this paper, we have proposed a method which can segment histogram-based occluded object segmentation. First, we find that object is a front object under occlusion by

**Figure 9. Experimental Results**

<table>
<thead>
<tr>
<th>Method</th>
<th>Time Average</th>
<th>Histogram Matching</th>
<th>Template Matching</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average rate</td>
<td>0.026 seconds</td>
<td>0.105 seconds</td>
<td></td>
</tr>
</tbody>
</table>


