NONRETINOTOPIC PARTICLE FILTER FOR VISUAL TRACKING

ALEXANDER A S GUNAWAN, ITO WASITO

1 Bina Nusantara University, Mathematics Department, School of Computer Science, Jakarta, Indonesia
2 Universitas Indonesia, Faculty of Computer Science, Depok, Indonesia
E-mail: 1, 2ito.wasito@cs.ui.ac.id

ABSTRACT

Visual tracking is the problem of using visual sensor measurements to determine location and path of target object. One of big challenges for visual tracking is full occlusion. When full occlusions are present, image data alone can be unreliable, and is not sufficient to detect the target object. The developed tracking algorithm is based on bootstrap particle filter and using color feature target. Furthermore the algorithm is modified using nonretinotopic concept, inspired from the way of human visual cortex handles occlusion by constructing nonretinotopic layers. We interpreted the concept by using past tracking memory about motion dynamics rather than current measurement when quality level of tracking reliability below a threshold. Using experiments, we found (i) the performance of the object tracking algorithm in handling occlusion can be improved using nonretinotopic concept, (ii) dynamic model is crucial for object tracking, especially when the target object experienced occlusion and maneuver motions, (iii) the dependency of the tracker performance on the accuracy of tracking quality threshold when facing illumination challenge. Preliminary experimental results are provided.

Keywords: Visual Tracking, Full Occlusion, Bootstrap Particle Filter, Color, Nonretinotopic, Tracking Quality Threshold

1. INTRODUCTION

In this research, we consider the problem of tracking single object in video sequences using only one camera. In particular, we focus on the cases where target object occludes by other objects, either partially or fully. Partial occlusion hides some parts of the target while complete occlusion hides the entire target for some time. Many techniques exist to handle the occlusion problem with particle filter probabilistic models, such as in [1]. But the proposed tracking method uses network of many cameras to handle this problem. The other proposed tracking method [2] adds the adaptiveness of likelihood function and invariance of color distributions to particle filtering. Note that the target object can be rigid (e.g., car) or deformable (e.g., person).

The main question of the research is how to track object in the environment are projected to photoreceptors in the retina in a similar manner as appearance models in a digital image. Nevertheless, a recent study on human vision [3] shows that the representation in higher visual areas of the visual cortex occurs in a nonretinotopic approach: visual perception seems to create dynamic layers for each moving object in the scene. This representation suggests that the appearance of the objects and their positions are marginal independent. The nonretinotopic approach describes our visual processing that always maintains the identity of observed objects across space and time [4].

The implementation of visual tracking algorithm in this research is based on Bayesian framework, mainly bootstrap particle filter. Furthermore the algorithm is modified using nonretinotopic concepts. To evaluate and compare the capabilities of two tracking algorithms: generic and retinotopic particle filter, we use various video sequences and analyze the tracking performance. The paper is organized as follows: first we discuss the Bayesian framework and its implementation based on bootstrap particle filter. For observation model in experiments, it is used color feature of target object. Then, we consider the nonretinotopic concept and
how to modify the algorithm using this concept. Finally, we discuss the experiment results dealing with occlusions in general and with motion and illumination in particular.

2. TRACKING AS BAYESIAN PROBLEM

The problem of tracking moving objects visually can be modeled as a Markov process where one wants to estimate the value of a hidden state \( x_t \) from a set of observations and a discrete time index \( t \). The main purpose of the tracker is to estimate the distribution with the state of the target \( x_t \) is the position and the velocity of the object. Based on Bayes' theorem, object tracking problems can be described as follows:

Prediction:

\[
p(x_{t,i} | z_{t-1,i}) = p(x_{t,i} | x_{t-1,i}) p(x_{t-1,i} | z_{t-1,i})
\]

Update:

\[
p(x_{t,i} | z_{t,i}) = \frac{p(z_{t,i} | x_{t,i}) p(x_{t,i} | z_{t-1,i})}{p(z_{t,i} | z_{t-1,i})}
\]

2.1 Recursive Bayesian Solution

A recursive solution for this estimation problem is provided by the Bayesian framework. The above recursive Bayesian solution provides the posterior conditional distribution \( p(x_t | z_{t-1,i}) \) at time \( t \), after receiving the last measurement \( z_{t-1} \). The above prediction and update stages simplify under several assumptions. First, the measurements at a given time are conditionally independent of measurements taken at other times, i.e., the measurements at time \( t \) are independent of the measurements at times \( \leq t-1 \), and that they depend only on the current states of objects via \( x_t \) and not on its entire state sequence, the measurement likelihood \( p(z_{t,i} | x_{t,i}) \) simplifies to \( p(z_{t,i} | x_t) \).

In object tracking, it is usual to make simplifying the model adequately based on real-world system. Then it is assumed that the system follow a Markov process of order one where the present state does not depend on previous states, given the last state that is \( p(x_t | x_{t-1,i}) = p(x_t | x_{t-1}) \).

The main purpose in object tracking is to estimate the distribution \( p(x_t | z_{t-1,i}) \). This knowledge can be derived from \( p(x_t | z_{t-1}) \) by integrating out objects and their state. Then the above prediction and update stages become:

Prediction:

\[
p(x_t | z_{t-1,i}) = \int p(x_t | x_{t-1}) p(x_{t-1,i} | z_{t-1,i}) \, dx_{t-1}
\]

Update:

\[
p(x_t | z_{t,i}) = \frac{p(z_{t,i} | x_t) p(x_t | z_{t-1,i})}{p(z_{t,i} | z_{t-1,i})}
\]
The BPF has the distinctive feature that the incremental weights do not depend on the past trajectory of the particles but only on the conditional likelihood of the observation \( p(z_t | x_t) \). For the BPF, sampling is very straightforward with the state transition equation used to predict new particles and is followed by the resample and move steps. The complete procedure [7] for the BPF is shown in Table 1.

### Table 1: The Bootstrap Particle Filter

<table>
<thead>
<tr>
<th>Step</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Initialize filter</strong></td>
</tr>
<tr>
<td>1. Initialize state vector particles ( x_0^{(i)} \sim q(x_0) ) ( i = 1, ..., n )</td>
</tr>
<tr>
<td>2. Initialize weights ( w_0^{(i)} = \frac{1}{n} )</td>
</tr>
<tr>
<td><strong>B. Sequential importance sampling</strong></td>
</tr>
</tbody>
</table>
| 1. Draw new particles \( v^{(i)} \sim p(v) \)  
   \( x_t^{(i)} = f(x_{t-1}^{(i)} + v^{(i)} \) \( \forall i \)  
   \( \tilde{w}_t^{(i)} = w_t^{(i)} p(z_t | x_t^{(i)}) \) |
| 2. Generate unnormalized importance weights \( w_t^{(i)} = \tilde{w}_t^{(i)} / \sum \tilde{w}_t^{(i)} \) |
| **C. Calculate \( N_{eff} \)** |
| 1/ \( \sum (\tilde{w}_t^{(i)})^2 \) |
| **D. If \( N_{eff} \ll n \)** |
| a. Resample the particles |
| b. Calculate the empirical mean \( \hat{x}_t = \frac{1}{n} \sum x_t^{(i)} \) |
| c. Regularize the resampled particles |
| d. Time step and return to B.1 |
| **else** |
| Time step and return to B.1 |
| end if |

The noise terms \( v^{(i)} \) are included in the transition equation for the state vector. Thus, there is also the unstated need for knowledge of the dynamic noise density. Thus to regularize, it can be imposed the transition equation to the resampled particles.

It needs to anticipate the degeneracy caused by migration of the spread of the discrete particle density away to specific dense distribution. A suitable measure of degeneracy for a particular application is the effective sample size \( N_{eff} \) that can be estimated from:
4. NONRETINOTOPIC PERCEPTION

The key to reverse engineering in visual perception is the understanding of visual information represented and processed in the brain. While it is clear that geometry and visual perception are closely related, very little is known about this relation beyond its early retinotopic association. In visual perception studies, retinotopic refers to the visual process where object in the environment are projected to photo-receptors in the retina in a similar manner as appearance models in a digital image. Nevertheless, a recent study on human vision [3] shows that the representation in higher visual areas of the visual cortex occurs in a nonretinotopic manner: visual perception seems to create dynamic layers for each moving object in the scene. This representation suggests that the appearance of the objects and their positions are marginal independent.

In this research, nonretinopic concept is interpreted as the filtering process use past tracking memory about motion dynamics rather than current measurement when quality level of tracking reliability below a threshold. A suitable measure of the tracking quality level for a particular application is sum of current likelihood \( Q_{lt} \) that can be estimated from:

\[
N_{eff} = \frac{1}{\sum_{i=1}^{n} \left( w_i^{t} \right)^2}
\]

4.1 Formulation of the occlusion problem

Occlusion means that there is something to see, but can't due to some event or some property of your sensor. Occlusion is a fact of visual life. Most of the objects you see are fully or at least partially obscured by other objects.

If a tracking system tracks objects (people, cars, etc) then occlusion occurs if an object is occluded by another object. For example two persons walking past each other, or a car that drives under a bridge. The problem in this case is what the tracking system does when an object disappears and reappears again [1]. In this research, it is implemented visual tracking algorithm based on bootstrap particle filter. Furthermore the algorithm is modified using nonretinotopic concepts, inspired from the way of human visual cortex handles occlusion by constructing nonretinotopic layers. Nonretinopic concept is interpreted using tracking quality level for detecting the tracking estimation of particle filter is not reliable again. Thus when the tracking quality level below a threshold, it is used past tracking memory to track the target.

5. NONRETINOTOPIC PARTICLE FILTER

In this research, we proposed an algorithm based on bootstrap particle filter using a color feature target and enhanced by nonretinotopic concept. A specific probabilistic graphical model in Figure 3 is used for defining the proposed technique.

In nonretinotopic particle filter, BPF is modified by adding the tracking estimation quality procedure after resample the particles and using past tracking memory that is the target object \( \hat{x}_t \) represented by empirical mean \( \hat{x}_t \). The modification procedure is shown in Table 2.

<table>
<thead>
<tr>
<th>D. If Neff &lt;&lt; n</th>
<th>a. Resample the particles</th>
</tr>
</thead>
<tbody>
<tr>
<td>If ( Q_{lt} &gt; \text{threshold} )</td>
<td>b. Calculate the empirical mean</td>
</tr>
<tr>
<td>\hspace{2cm}</td>
<td>c. Regularize the resampled particles</td>
</tr>
<tr>
<td>\hspace{2cm}</td>
<td>d. Calculate nonretinotopic update</td>
</tr>
<tr>
<td>\hspace{2cm}</td>
<td>e. Spread the</td>
</tr>
<tr>
<td>\hspace{2cm}</td>
<td>( \sigma_j = 3. \sigma_{r-1} )</td>
</tr>
</tbody>
</table>
particles on scene based on velocity standard deviation

end if

f. Time step and return to B.1

\[ x_{t}^{(i)} \rightarrow x_{t+1}^{(i)} \]

\[ w_{j}^{(i)} = \frac{1}{n} \]

else

Time step and return to B.1

\[ x_{t}^{(i)} \rightarrow x_{t+1}^{(i)} \]

\[ w_{j}^{(i)} = w_{j}^{(i)} \]

As a Bayesian estimator, nonretinotopic particle filter has two main steps: prediction and update. Prediction is done by propagating the particles based on the transition model. The update step is done by measuring the weight of each particles based on the observation model. The implementation of nonretinotopic particle filter can be described as follows.

1. Particle initialization.

In this research, the initialization of the particles is done by putting the particles randomly on entire scene depend on video resolution. For object target is used blue color feature in RGB format (0, 0, 255). Furthermore it is used two blue objects in experiment that is: pure blue ball and globe ball. In this initialization approach, it is used tracking quality level to determine the beginning of object tracking, when the level is more than a threshold. It can be seen on Figure 4, that pure blue ball can start to track since time step 1, indicated by yellow star.

(a) Pure blue ball  (b) Globe ball

Figure 4: Initialization Based On Random Particles

2. Prediction: Transition model.

In this research, the motion of the object is considered as the discrete time 2-dimensional (2D) motion with constant velocity assumption. The state vector at a time step t is denoted by \( X_{t} \), including horizontal position, vertical position, horizontal velocity, vertical velocity of each particle. The noise terms \( v_{t}^{(i)} \) are Gaussian noise included in the transition equation for the state vector. The noise terms provide the system with a diversity of hypotheses. Thus, each component of the particles is predicted by propagating past state vector according to the transition model. For this purpose, the equations system of transition model is described by:

\[
\begin{bmatrix}
 x_{t} \\
 y_{t} \\
 \dot{x}_{t} \\
 \dot{y}_{t}
\end{bmatrix} =
\begin{bmatrix}
 1 & 0 & 0 & 0 \\
 0 & 1 & 0 & 0 \\
 0 & 0 & 1 & 0 \\
 0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
 x_{t-1} \\
 y_{t-1} \\
 \dot{x}_{t-1} \\
 \dot{y}_{t-1}
\end{bmatrix} +
\begin{bmatrix}
 v_{t} \\
 v_{t} \\
 v_{t} \\
 v_{t}
\end{bmatrix}
\]

(11)

Furthermore, when the tracking quality level below a threshold (Qlt < threshold), it is used past tracking memory to track the target based on new dynamic transition model. The motion of the target can be modeled using a coordinated turn model [9] of the form:

\[
\begin{bmatrix}
 x_{t} \\
 y_{t} \\
 \dot{x}_{t} \\
 \dot{y}_{t}
\end{bmatrix} =
\begin{bmatrix}
 1 & \sin \omega & -\cos \omega & 0 \\
 0 & \cos \omega & \sin \omega & 0 \\
 0 & 0 & 1 & 0 \\
 0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
 x_{t-1} \\
 y_{t-1} \\
 \dot{x}_{t-1} \\
 \dot{y}_{t-1}
\end{bmatrix} +
\begin{bmatrix}
 v_{t} \\
 v_{t} \\
 v_{t} \\
 v_{t}
\end{bmatrix}
\]

(12)

where, \( \omega \in (-\pi, \pi) \) in rads/s is rate of turn which can be adjusted depend on the target motions. In the limiting case of the rate of turn parameter \( \omega \rightarrow 0 \), the coordinated turn model simplifies to equation (11).

3. Update: Observation model

The observation model is used to measure the observation likelihood of the particles. The likelihood function is based on the observation noise density and is assumed to be Gaussian. For this purpose, the likelihood function using deviation from target d is described by:

\[
p(z_{t} \mid x_{t}^{(i)}) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{d^{2}}{2\sigma^{2}}\right)
\]

(13)

In this research, the observation model is made based on color information of the target obtained by building the color histogram in the RGB color space. Moreover, it is used blue color feature in RGB format (0, 0, 255) as object target. Thus the deviation d is defined using Euclidean distance as:

\[
d = \sqrt{R^{2} + G^{2} + (B - 255)^{2}}
\]

(14)

To maintain a consistent particle, the new importance weights are set by measuring the likelihood of each particle based on the observation model. Figure 5 shows examples of object target at frame #10.
4. Resample the particles

The resample step is performed to produce new particles based on the current particles weight. During the resample step, particles with high weight possibly will be chosen several times generating identical copies, while others with relatively low weights may possibly be ignored and deleted. The resample step [7] can be done in the following steps:

a. Generate \( n \) random numbers that is uniformly distributed on \([0,1]\) that is \( u_j \sim U(0,1) \) \( j = 1 \ldots n \)

b. Calculate the cumulative probability of the normalized particles weight. The result is intervals \( w_i^{(b)} \leq x < w_i^{(a)} \) where \( i = 1 \ldots n \)

c. Find the interval where \( w_i^{(b)} \leq x < w_i^{(a)} \) and set the new particle \( x_i^{(b)} = x_i^{(a)} \)

6. EXPERIMENTAL RESULTS

In order to evaluate the proposed method, it is done the experiments using a video camera to track the objects. The experiments are implemented on Intel i3 2.53 [GHz] CPU and 2 [GB] RAM. The resolution of each frame is 552×402 [pixels] image. In each experimental result, the blue dots represent the particles, the yellow star represents object target from generic particle filter, calculated from the mean state of the particles position and the red star represents the object target from nonretinotopic particle filter.

The experiments are done using 1000 particles and in three challenges that is occlusions in general and with motion and illumination in particular. The initial position of the target object is assumed unknown. For object target is used blue color feature in RGB format (0, 0, 255). Furthermore it is used two blue objects in experiment that is: pure blue ball and globe ball. As default, tracking quality threshold is set equal to 1 except in the illumination challenge.

6.1 Object tracking with occlusion

In this experiment, it is compared generic bootstrap particle filter with nonretinotopic particle filter when the target object is fully and partially obscured by other objects. While the target object is pure blue ball, there is no different performance between two algorithms. The two algorithms can track the pure blue ball very well.

On the other hand, there is very interesting result for globe ball target. Figure 6 shows the performance of generic bootstrap particle filter in handling full occlusion (frame #210 – frame #231) using globe ball target. As shown in that figure, the algorithm fails to track the target object and traps in local area.

Figure 6: Generic Particle Filter In Handling Occlusion Challenge

Figure 8 shows the performance of nonretinotopic particle filter in handling full occlusion (frame #210 – frame #231) using globe ball target. As shown in that figure, the algorithm can track successfully the target object although the partial and full occlusion occurs.

6.2 Object tracking with different motions

In this experiment, it is observed the algorithm performance when the target object performs vertical and horizontal motions. For the horizontal motions, the tracking algorithm can track the globe ball very well. However while the target performs
vertical motions, there is interesting phenomena that is the particles sometime only follow the target object poorly. As shown in figure 10, the particles in frame #167 – frame #180 fail to follow the target object and have to spread in searching the target object.

6.3 Object tracking with different illumination

In this experiment, it is observed the algorithm performance when the object tracking is performed in different illumination environment that is outdoor, indoor and dark room. In outdoor and indoor environment, there is no out of ordinary occurrence. However, in dark room there is something remarkable that is the algorithm cannot track the globe ball target at all using default tracking quality threshold. When this threshold is decreased to 0.1, the tracking algorithm starts to track the target object in usual manner. Overall, the performance of the tracking object in dark room is decreased considerably. This illumination challenge may be treated using image processing such as Histogram equalization. As presented in Figure 11, the particles in frame #1 – frame #74 tend to spread around the target object in order to follow its motions.

7. CONCLUSIONS

This paper presented a method to track the moving object employing bootstrap particle filter based on color information. Furthermore, the algorithm is modified using nonretinotopic concepts, inspired from the way of human visual cortex handles occlusion by constructing nonretinotopic layers. There are three main conclusions: (i) the performance of the object tracking algorithm in handling occlusion can be improved using nonretinotopic concept, (ii) dynamic model is crucial for object tracking, especially when the target object experienced occlusion and maneuver motions, (iii) the dependency of the tracker performance on the accuracy of tracking quality threshold when facing illumination challenge.

The algorithm is still restricted to track single object and will extend to track multiple object and will enhance by creating the advanced dynamic model. These are remaining for our future works.

REFERENCES:


