The Evaluation of the Gaussian Mixture Probability Hypothesis Density Approach for Multi-target Tracking

Jiandan Chen, Oyekanlu Emmanuel Adebomi, Onidare Samuel Olusayo and Wlodek Kulesza
Department of Electrical Engineering
Blekinge Institute of Technology
Karlskrona, Sweden

***Abstract*—this paper describes the performance of the Gaussian Mixture Probability Hypothesis Density (GM-PHD) filter for multiple human tracking in an intelligent vision system. Human movement trajectories were observed with a camera and tracked by the GM-PHD filter. The filter multi-target tracking ability was validated by two random motion trajectories in the paper. To evaluate the filter performance in relation to the target movement, the motion velocity and angular velocity as key evaluation factors were proposed. A circular motion model was implemented for simplified analysis of the filter tracking performance. The results indicate that the mean absolute error defined as the difference between the filter prediction and the ground truth is proportional to the motion speed and angular velocity of the target. The error is only slightly affected by the tracking targets' number.**

Keywords-Human Tracking, Probability Hypothesis Density, Performance Evaluation, Vision System

I. INTRODUCTION

The tracking of moving objects has very diverse applications in almost all endeavours of life in present-day society, including security, surveillance, robotics, aeronautics, medicine and sports. It is the fulcrum of the Intelligent Vision Agent System, IVAS [1], which is a high-performance autonomous distributed vision and information processing system. This vision sensor system consists of multiple sensors and actuators and is used for surveillance of a human activities space which includes the human being and her surrounding environment, comprising robots, household appliances, lights and so on. The object of interest in the system is the human being(s) whose state(s) at any given position in time within the activity space must be known.

The human activities space, as the target space, provides constraints to the design and planning of the active camera system. This means that the motion and possible locations of the human being need to be predicted in order to control the system and be able to track humans with the intention of keeping them in the cameras' Field of Views, FoVs.

The Bayesian's recursion provides the mathematical framework of the tracking algorithms. A typical tracking algorithm based on the Bayesian's recursion computes the posterior probability density of a process based on the prior probability density of the process and the likelihood function. In this paper, the Gaussian mixture approximation of the Probability Hypothesis Density filter, GM-PHD, is applied to track multi-target movement. The tracking information is used in an IVAS to determine the position, orientation, movement speed, focal length and baseline length of multiple cameras [1], [2].

The tracking ability is limited by the target(s)' motion speed and angular velocity. The evaluation of the filter performance for different multi-target motion parameters is useful in the design of a vision tracking system, [3].

II. RELATED WORKS

The GM-PHD filter algorithm was proposed by Vo and Ma [4]. In [5], the algorithm is used for tracking multiple targets in high clutter density. The comparison made between the algorithm and the Multiple Hypothesis Tracker (MHT) shows that the former is better in areas of high clutter density and that it initiates and eliminates targets more accurately.

Using the GM-PHD filter, Pham et al. [6] have proposed a multi-camera, multi-object tracking system that can track 3D object locations even when objects are occluded from the cameras. Also, Pham et al. [7] have demonstrated that tracking objects using multiple sensors is more efficient than using one sensor.

Furthermore, Chen and Zhu [8] have modeled the birth intensity for multiple targets tracking using the GM-PHD filter. In [9], a method of using the GM-PHD filter to track multiple objects by incorporating color representation is proposed.

Clark et al. [10] have used the GM-PHD filter in audio analysis to distinguish between measurements that are generated by actual and by false notes emanating from a piano by tracking the frequency and amplitude of the components of piano harmonics. The results obtained also show that the GM-PHD filter can be used to estimate the number of individual harmonics and maintain their identities. Clark et al [11] also provided a review of non-linear filtering techniques (Monte Carlo, second-order divided difference, Gaussian particle and unscented Kalman filter) and showed how these can be incorporated with the GM-PHD filter.

Juang et al. [12] have implemented the GM-PHD filter in the bio-medical field for tracking the movement of multiple cells and their lineages.

The work of Sidenbladh [13] is a particle PHD filter implementation for tracking multiple vehicles in terrain. The

978-1-4244-6494-4/10/$26.00 ©2010 IEEE

result of the implementation shows the robustness of the method when tracking a varying number of targets.

III. Problem Statement and Main Contributions

The IVAS must be able to predict human position to optimize the parameters of the camera. Human movement property is one of the most influential factors in tracking filter performance evaluation. When investigating how human motion speed and angular velocity affect the GM-PHD filter, tracking performance is an important issue of the filter design. The proposed model suggests the criteria for a filter's target motion speed and angular velocity in a vision system.

The main contributions of the paper can be summarized as follows:

- Two key parameters, motion speed and angular velocity, in a motion model are proposed to be used for evaluation of the GM-PHD filter tracking performance;
- A circular human motion model is introduced as a test signal for evaluation of the tracking process;
- The GM-PHD filter for multi-target tracking is modeled, implemented in Matlab and validated using two random motion trajectories;
- The filter performance evaluation using the test circular motion is verified in a simulated environment;
- The dependence of filter performance on the dynamics of target motion and the number of targets is investigated reported.

IV. Modelling

A. *Model of multi-target tracking using the GM-PHD filter*

In the framework of the multi-target tracker, the multi-target state can be described by the Random Finite Set (RFS). The multi-target state can be represented as a discrete time k set X_k defined as:

$$\mathbf{X}_k = \{\chi_{k,i} : i = 1, \cdots, M_\chi(k)\} \quad , \tag{1}$$

where $M_\chi(k)$ is the number of targets in the scene in the time k, and i is the index variable.

The multi-target measurement from the camera sensor is the set:

$$\mathbf{Z}_k = \{\zeta_{k,j} : j = 1, \cdots M_\zeta(k)\} \quad , \tag{2}$$

where $M_\zeta(k)$ is the number of observations in the time k, and j is the index variable**.**

The GM-PHD filter is based on three additional assumptions compared to the PHD filter, [4]:

(i) Each target and the sensor follow the linear Gaussian model which can be described by:

$$\begin{aligned} f_{k|k-1}(\chi|\varsigma) &= N(\chi; \mathbf{F}_{k-1}\varsigma; \mathbf{Q}_{k-1}) \quad , \\ g_k(\zeta|\chi) &= N(\zeta; \mathbf{H}_k\varphi; \mathbf{R}_k) \quad , \end{aligned} \tag{3}$$

where N is the normal or Gaussian distribution operator, $N(\cdot, m, P)$ denotes a Gaussian density with the mean m and the covariance P,
$\mathbf{F}_{k-1}$ is the state transition matrix,
$\mathbf{Q}_{k-1}$ is the process noise covariance matrix,
$\mathbf{H}_k$ is the observation matrix, and
$\mathbf{R}_k$ is the observation noise covariance matrix.

(ii) The survival and detection probabilities, p_S and p_D respectively, are state independent,
(iii) The intensity of the birth RFS is a Gaussian mixture.

Like other smoothing filters, the GM-PHD filter consists of two steps: prediction and update. The prediction equation is defined as:

$$v_{k|k-1}(\chi) = \gamma_k(\chi) + p_{S,k}\sum_{j=1}^{J_{k-1}} w_{k-1}^{(j)} N(\chi; m_{k-1}^{(j)}, P_{k|k-1}^{(j)}) \quad , \tag{4}$$

where γ is the birth intensity and w is the weight parameter.

After the object detection has been finished and Z_k is available, the state is updated according to:

$$v_k(\chi) = (1 - p_{D,k}) v_{k|k-1}(\chi) + \sum_{z \in Z_k} \sum_{j=1}^{J_{k|k-1}} w_k^{(j)}(z) N(\chi; m_{k|k}^{(j)}, P_{k|k}^{(j)}) \quad , \tag{5}$$

where v is the intensity function.

B. *Modelling of the target motion*

Assuming that a target moves in a horizontal plane, X and Y, as if on the ground, then the coordinate variable, Z, in the vertical direction, can be assumed to be a constant. The point $\mathbf{T}_{k-1}=(T_{x(k-1)}, T_{y(k-1)})$ at the time k-1 represents the target position in the plane. This point is assumed to have moved to a new position $\mathbf{T}_k$ at the time k. Fig. 1 shows a target motion illustration and can be described as:

$$\mathbf{T}_k = \mathbf{T}_{k-1} + \Delta k v_{k-1}\mathbf{a}_{k-1} \quad , \tag{6}$$

where $v_{k-1}\mathbf{a}_{k-1}$ denotes the velocity in the time k-1**.** v_{k-1} is the speed (absolute amplitude of velocity), and $\mathbf{a}_{k-1}$ is the direction vector of velocity in the horizontal plane, ($\|\mathbf{a}\| = 1$)**.** The target motion velocity direction can be presented as the vector, $\alpha_{k-1} = (\cos(\theta_{k-1}), \sin(\theta_{k-1}))$. The angle θ_{k-1} is the angle between the motion direction and the X-axis at the time k-1. The sample interval, Δk, is the observation sampling time when the target moves from position k-1 to k.

The motion angular shift, $\Delta\theta_k$ can be presented as:

$$\Delta\theta_k = \theta_k - \theta_{k-1} \quad . \tag{7}$$

The tracking performance versus the target motion speed, v, and angular velocity, ω, are studied in chapter V.

C. *Human circular motion model*

In order to easily evaluate the performance of the tracking filter, a circular motion is proposed.

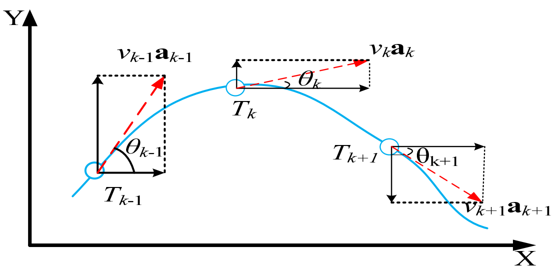

Fig. 1. Target motion illustration.

The target's initial position $\mathbf{T_0}$ in the 2D plane can be described by the function:

$$\mathbf{T_0} = \mathbf{C} + r \cdot (\cos(\alpha_0), \sin(\alpha_0)) \quad , \tag{8}$$

where $\mathbf{C}=(c_x, c_y)$ is a circle centre and r is the radius of the circle and the initial angle α_0 is an angle between the line from the circuit center to the target position and the axis X. Then, the target motion state $\mathbf{T}_k$ can be approximated by:

$$\mathbf{T}_k = \mathbf{T}_{k-1} + \omega r \Delta k \cdot (\cos(\theta_{k-1}), \sin(\theta_{k-1})) \quad , \tag{9}$$

where the angular velocity, $\omega=2\pi/(K\cdot\Delta k)$, and K is the measurement sampling rate (samples per rotation). The motion direction alteration is defined as $\Delta\theta_k = \omega\Delta k$, and the motion speed is ωr.

The observation ς_k on a camera 2D sensor plane at the time k for a target at a 3D position $\mathbf{T}_k$ can be described as:

$$\zeta_k = \mathbf{P}\mathbf{T}_k + \tau_k \quad , \tag{10}$$

where $\mathbf{P}$ is the camera projection matrix and includes the camera intrinsic and extrinsic parameters, and τ_k is the measurement noise.

The noise that is the most influenced is the spatial noise in the measurement. This noise may originate from the discrete sensor pixels, the resolution of the focal length, or camera vibration if a camera that moved dynamically was used. In the paper, a Gaussian noise distribution was used during measurements.

V. VALIDATION OF THE GM-PHD MULTI-TARGET TRACKER AND ITS PERFORMANCE ANALYSIS

The models presented in the previous chapter were implemented in Matlab and then validated by an analysis of the simulation results.

The simulated space denotes a 3D indoor environment and describes the space wherein the human body exists and activities take place; this space must be covered by the camera's Field of View. In order to simplify the analysis for the tracking process, the target is represented as a 3D point in the space. A pinhole camera is assumed in the model.

The simulation environment considers a regular room of the size 4x4x3 m. The Epipolar Geometry Toolbox, [14] was used to project the target position in the 3D space to the image plane. The camera focal length was 700 pixels, and the image plane was 1024×1024 pixels. To validate the implementation of the GM-PHD filter, and to evaluate its performance, two different human motion models are presented in the following.

A. *Model validated by random human motion trajectory*

A random walk motion was modeled and simulated. This type of human motion is used to validate the filter's multi-target tracking ability. Two targets were used in the simulation. The random trajectories can be generated by equation (6) by 50 samples. The target motion speed was generated by the Gaussian random function in Matlab, where the target movement speed, v, was in the range of 0 to 0.1 meter per sample interval. The motion direction, θ, is a Gaussian random angle in the range of 0 to 90 degrees. Two targets move within the surveillance region. The camera orientation is perpendicular to the ground plane. The same camera setting will be used in the performance analysis.

Fig. 2 indicates the tracking manner of the GM-PHD filter in the x and y motions in the sensor plane according to the time k. The two truth trajectories are presented in solid red and blue lines respectively. The filter predication is marked as the blue crosses. Observe that the filter tracking pattern, from the moment when the two targets appear until they leave the surveillance region, follows the movement pattern very well. This validates the GM-PHD filter's multi target tracking ability.

B. *Performace analysis*

The human circular motion model was applied in the performance analysis by using (7) and (8). The mean absolute error defined as the difference between the prediction and the ground truth was measured by use of the Wasserstein distance, [5]. The performance versus the target motion and the number of targets were studied.

1) *Performance versus target motion*

The target moves along a circle. The target motion speed is constant at any point on the circle for the same radius and angular velocity. In this study, three different speed levels were applied. The circular motion uses three different radiuses, r:s, 0.64m, 0.48m and 0.32m, when the measurement sampling rate, K, varies from 60 samples/rotation to 140 samples/rotation. The angular shift per sample interval is $\Delta\theta = 360°/K$. It varies in the range of 6 to 2.6 degrees per sample interval.

Fig. 3 shows that the mean absolute error varies with the sampling rate K for the different radiuses r. The blue, green and red curves correspond to the radiuses 0.64 m, 0.48 m and 0.32 m respectively. The mean absolute error is proportional to the radius and decreases exponentially when the sampling rate K increases. By the interpretation of r and K, we can also state that the mean absolute error becomes larger when the motion

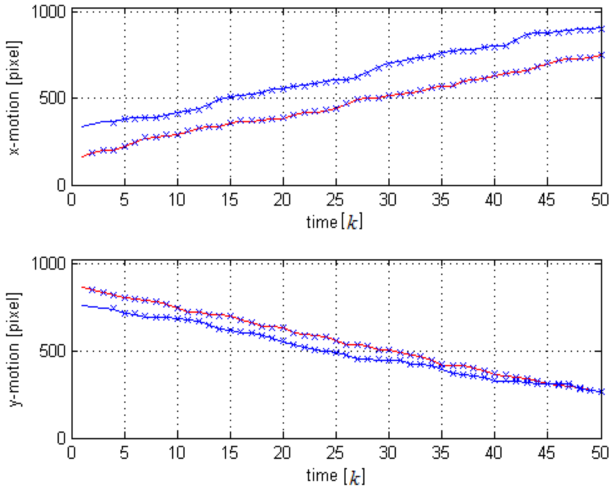

Fig. 2. The target motion path on the image plane in the x and y coordinates, the ground truth is marked as the solid line and the predictions by the GM-PHD filter are marked as crosses.

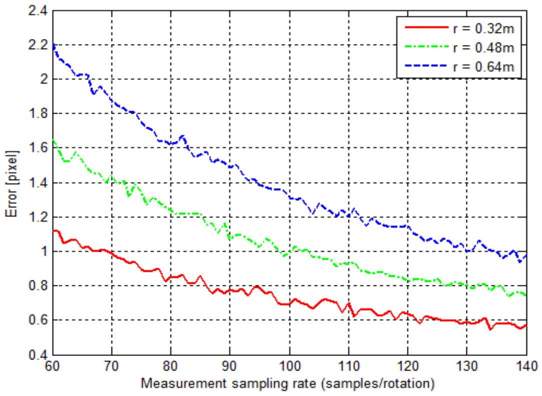

Fig. 3. Mean absolute error vs. measurement sampling rate K and circular radius r.

speed and angular velocity are higher. The error is less than 1 pixel when the sampling rate, K, is greater than 70 samples/rotation, 100 samples/rotation and 130 samples/rotation with respect to the radius 0.32 m, 0.48 m and 0.64 m respectively. The errors are in the range of the camera pixel quantization level.

2) *Performance versus number of targets*

The performance versus the number of targets was studied by a comparison between the tracking performances of the one target and of the two targets respectively for the simulated circular motion with a radius 0.48 m. Fig. 4 shows the mean absolute error as it varies with the sampling rate, K, for one-target tracking, indicated by the green curve. The error versus two-target tracking is indicated by the red curve. The mean absolute error for two-target tracking is slightly higher than that for one-target tracking.

VI. CONCLUSION

The multi-target tracking ability of the GM-PHD filter is validated by the results of two random motion trajectory simulations.

Human circular motion is modeled by two key parameters, speed and angular velocity, which are useful for the evaluation of filter tracking performance. The advantage of the circular motion model is that the parameters of the target motion and the measurement can be easily related to the radius and measurement sampling rate.

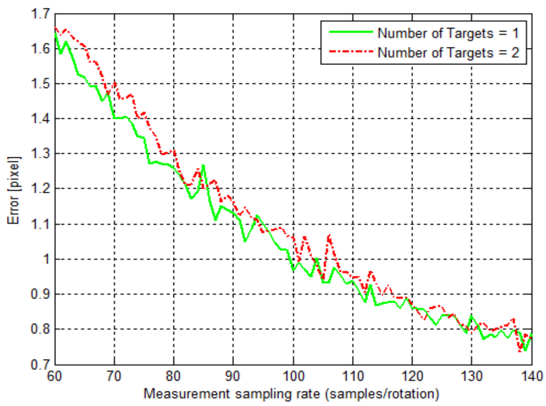

Fig. 4. Mean absolute error vs. target number.

The mean absolute error between the prediction and the ground truth in the tracking showed that the GM-PHD filter performance depends on target motion speed and angular velocity, and that the tracking error is thus proportional to motion speed and angular velocity.

The number of tracked targets only slightly affected the error level.

In future research endeavors, it would be very useful to examine the usage of multiple cameras in solving occlusion problems in motion tracking that employ the GM-PHD filter. The general relationship between the tracking error and the target motion speed and angular velocity can then be defined.

ACKNOWLEGEMENTS

The authors wish to acknowledge Dr. Anders M. Johansson for his inspiration and support, and Dr. Johan Höglund for his comments.

REFERENCES

[1]. J. Chen, S. Khatibi and W. Kulesza, "Planning of a multi stereo visual sensor system for a human activities space," in: Proc. of the 2nd Int. Conf. on Computer Vision Theory and Applications, pp. 480-485, 2007.
[2]. J. Chen, S. Khatibi and W. Kulesza, "Planning of a multi stereo visual sensor system- depth accuracy and variable baseline approach," in: Proc. of IEEE Computer Society 3DTV-Con, the True Vision Capture, Transmission and Display of 3D Video, Kos, Greece, 2007.
[3]. W. Kulesza, J. Chen, S. Khatibi, "Arrangement of a multi stereo visual sensor system for a human activities space". In: A. Bhatti (Ed.), Stereo Vision, InTech Education and Publishing, Vienna, pp. 153-172, 2008
[4]. B. Vo and W. K. Ma, "The Gaussian mixture probability hypothesis density filter," IEEE Transactions Signal Processing, vol. 54, no. 11, pp.4091-4104, 2006.
[5]. D. Clark, K. Panta and B.-N. Vo, "The GM-PHD filter multiple target tracker," Proc. Information Fusion, 2006 9th International Conference, pp. 1-8, July 2006.
[6]. N. T. Pham, W. Huang and S. H. Ong "Probability hypothesis density approach for multi-camera multi-object tracking" Lecture Notes in Computer Science, ISSN: 03029743, pp. 875-884, 2007.
[7]. N. T. Pham, W. Huang and S. H. Ong "Multiple sensor multiple object tracking with GMPHD filter" Proc. Information Fusion, 2007 10th International Conference, pp. 1-7, July, 2007.
[8]. R. Chen and M. Zhu, Birth Density Modelling in Multi-target Tracking Using the Gaussian Mixture PHD Filter, Master Thesis, Department of Electrical Engineering, Blekinge Institute of Technology, Sweden, 2008.
[9]. N.T. Pham, W. Huang and S. H. Ong "Tracking multiple objects using probability hypothesis density filter and color measurements," IEEE International Conference on Multimedia and Expo, pp. 1-4, July 2007.
[10]. D. Clark, B-T. Vo, B-N. Vo and S. Godsill, "Gaussian mixture implementations of probability hypothesis density filters for non-linear dynamical models," IET seminar on Target Tracking and Data Fusion, pp. 21-28, April, 2008.
[11]. D. Clark, A-T. Cemgil, P.Peeling and S. Godsill "Multi-object tracking of sinusoidal components in audio with the Gaussian mixture probability hypothesis density filter," IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, pp. 1-4, October 2007.
[12]. R .R. Juang, A. Levchenko and P. Burlina, "Tracking cell motion using GM-PHD," IEEE International Symposium on Biomedical Imaging, pp. 1154-1157, July 2009.
[13]. H. Sidenbladh, "Multi-Target particle filtering for the probability hypothesis density," 6th International Conference on Information Fusion. Cairns, Australia, pp. 800–806, 2003.
[14]. G. L. Mariottini, D. Prattichizzo, "The epipolar geometry toolbox: multiple view geometry and visual servoing for matlab". In: Proc. of the IEEE Int. Conf. on Robotics and Automation, 2005.