Jurnal Ilmiah KURSOR
Menuju Solusi Teknologi Informasi

Vol. 8, No. 2, Desember 2015

ISSN 0216 – 0544
e-ISSN 2301– 6914

PARTICLE FILTER-BASED OBJECT TRACKING USING JOINT FEATURES OF COLOR AND LOCAL BINARY PATTERN HISTOGRAM FOURIER

[a]Dewa Made Wiharta, [b]Wirawan, [c]Gamantyo Hendrantoro
[a,b,c] Department of Electrical Engineering,
Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia
[a]Department of Electrical Engineering, Universitas Udayana, Denpasar, Bali, Indonesia
Email: wiharta@unud.ac.id

Abstract

Object tracking is defined as the problem of estimating object location in image sequences. In general, the problems of object tracking in real time and complex environtment are affected by many uncertainty. In this research we use a sequensial Monte Carlo method, known as particle filter, to build an object tracking algorithm. Particle filter, due to its multiple hypotheses, is known to be a robust method in object tracking task.

The performances of particle filter is defined by how the particles distributed. The role of distribution is regulated by the system model being used. In this research, a modified system model is proposed to manage particles distribution to achieve better performance.

Object representation also plays important role in object tracking. In this research, we combine color histogram and texture from Local Binary Pattern Histogram Fourier (LBPHF) operator as feature in object tracking.

Our experiments show that the proposed system model delivers a more robust tracking task, especially for objects with sudden changes in speed and direction. The proposed joint feature is able to capture object with changing shape and has better accuracy than single feature of color or joint color texture from other LBP variants.

Keywords : Particle Filter, Object Tracking, Color Histogram, Texture, System Model

INTRODUCTION

Video tracking can be defined as the problem of estimating position of an object over time. As an estimator, the role of success of an object tracking algorithm is determined by the ability to estimate the position of target object in video frames. The estimate depends on the method being used, the object representation, and the characteristics of the object.

Many approaches have been proposed for object tracking. Babenko et.al [1] proposes to use an online Multiple Instance Learning for object tracking by choosing positive and negative examples when updating the model. Kwon and Lee [2] decompose observation model into multiple models that are produced by sparse Principal Component Analysis of a set of feature templates.

In geometric particle filter, region covariance is used to model the appearance of an object and tracking is led by particle filter [3]. Recently, [4] proposes geometric particle filtering approach to template-based visual tracking to overcome the limitation of the deterministic optimization approach.

Mean Shift algorithm as proposed in [5]–[7] locates object by the use of iterative kernel-based deterministic procedures. This methods may converge to a local maximum as it is sensitive to occlusion, background similarities and fast moving objects [3] [8]. A stochastic method, such as particle filter, can be used to mitigate those problems. Particle filter maintains multiple hypotheses in the state space and provides a robust tracking framework as they are neither limited to linear systems nor require the noise to be Gaussian [3] [8] [9].

In our previous work [10], we found that the performances of object tracking is determined by the spreading of particles, which regulated by the system model (or motion model) being used. Another thing that plays important role in object tracking is object representation. The most common feature used by many researchers to represent the object is color histogram [11]. The color histogram is produced by the use of three-dimensional color and divide the color values into bins with uniform interval.

The downside of color histogram is the loss of spatial information and poses problems from illumination and background changes, or the present of other objects with similar color. Therefore, other features are required to overcome this limitation.

Feature that gives spatial information that is not provided by color is texture. The use of simultaneous color and texture in representating object will lead to better object descriptor, and better descriptor will improve system performances.

Many researchers have already used multiple features in object tracking tasks. The work reported in [11] combines color, texture, and motion in tracking object by features matching. Similarly, [7] proposes to combine color and texture extracted by Local Binary Pattern in mean shift tracking, and [12] uses color and texture generated by DWT for tracking object using Gaussian sum particle filter.

Tekstur can be generated using Local Binary Pattern (LBP) operator [13]. LBP is known to be a simple and yet very efficient texture operator. LBP has some advantages, such as fast computation and rotation invariance properties. The extension of this operator is so called uniform rotation invariant descriptor, which reducing the number of LBP codes. Later, Ahonen et.al in [14] extends LBP operator and uses Discrete Fourier Transform to generate feature that invariant to image rotation.

In this research, we propose to use particle filter in object tracking context with modified motion model as in [15] to regulate particle distribution and use joint features of color and texture to improve the performance of particle filter. The originality of this research is the use of joint color-texture in particle filter-based object tracking, where the texture is extracted using Local Binary Pattern Histogram Fourier. We use this type of LBP since the rotation invariance is attained globally which still retain information about relative distribution of different orientations of uniform local binary.

PARTICLE FILTER

Particle filter is a sequential Monte Carlo method based on random particles representation of probability densities [16]. Particle filter is used when posterior density $p(x|z)$ and measurement density $p(z|x)$ are non-Gaussian

The posterior density $p(\mathrm{x}_k|\mathrm{z}_{1:k})$ is represented by a set of weighted sample set S $=\{(\mathrm{s}_k^{(n)},\mathrm{w}_k^{(n)}), n = 1,\ldots N\}$, where x_k denotes the state of the object at time k, $\mathrm{z}_{1:k}$ denotes all the measurements up to time k, s_k denotes sample with weight ω_k, and N denotes number of sample.

The weights are normalised such that

$$\sum_{n=1}^{N} \omega_k^{(n)} = 1 \quad {}_k \geq 0 \qquad (1)$$

The posterior density is approximated as:

$$p(\mathrm{x}_k|\mathrm{z}_{1:k}) \approx \sum_{n=1}^{N} \omega_k^{(n)}\, \delta\left(\mathrm{x}_k - \mathrm{x}_k^{(n)}\right) \qquad (2)$$

Eq (2) gives a discrete weighted approximation of the true posterior $p(\mathrm{x}_k|\mathrm{z}_{1:\mathrm{k}})$. Each sample represents one hypothetical state of the estimation, with a discrete sampling probabilityω.

COLOR HISTOGRAM

A color histogram is a representation of colors distribution of an image. This can be viewed as an approximation of an underlying continuous distribution of colors values. Color histogram is built by discretizing the colors in an image into a number of bins, and counting the number of pixels in each bin.

The color distribution $p_y = \{p_y^{(u)}\}_{u=1,\ldots m}$ at location y is calculated as [5]

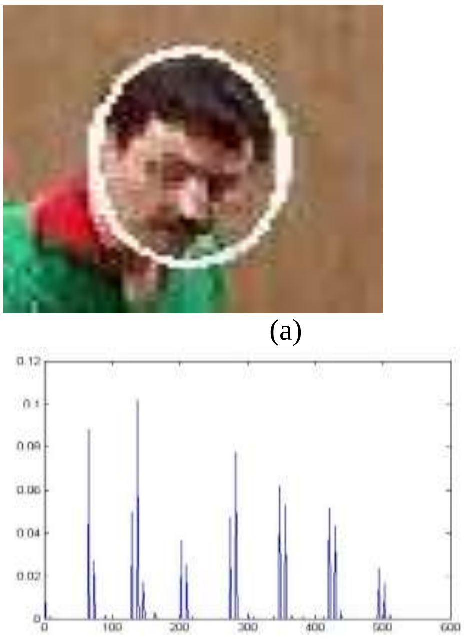
$$p_y^{(u)} = C \sum_{i=1}^{n} k\,(\|\, x_i\|^2)\delta[h(x_i - u)] \qquad (3)$$

where n is the number of pixels in the region. The result is parallel vertical bars that graphically show the frequency distribution of quantitative variables. Figure 1 displays histogram of an object in RGB color space. The histogram is constructed using 8 bin for each color component.

LOCAL BINARY PATTERN

Local Binary Pattern operator defines the description of the neighborhood of a pixel, derived from a general definition of texture in a local neighborhood.

The gray value of center pixel (g_c) from the gray value of the neighborhood (g_l) with radius R is transformed into a unique $\mathrm{LBP}_{L,R}$ such as [13] [17]:



(b)

Figure 1. An object and its representation in color histogram

$$\lambda_{L,R}(x_c, y_c) = \sum_{l=0}^{L-1} r(g_l - g_c)2^L \qquad (4)$$

Eq (4) characterizes the spatial information of the local texture, where $\lambda_{L,R}$ is the LBP operator that correspond to the gray value of the center pixel (x_c,y_c) of a local neighborhood and g_l to the gray value of L equally spaced pixels on a circle with radius R. To provide better discriminator, Takala and Pietikäeinen in [13] introduce a uniformity measure U that corresponds to the number of spatial transitions in the pattern. They designated a uniform pattern with U value of at most 2 and proposed the a rotation invariant texture description:

$$\lambda_{L,R}^{\mathrm{riu2}} = \begin{cases} \sum_{l=0}^{L-1} s(g_l - g_c), & \text{if } U(LBP_{L,R}) \leq 2 \\ L + 1, & otherwise \end{cases} \qquad (5)$$

where

$$U(\lambda_{L,R}) = |r(g_{l-1} - g_c) - s(g_0 - g_c| + \sum_{l=0}^{L-1} \left| r(g_l - g_c) - r(g_{l-1} - g_{c)} \right| \qquad (6)$$

Superscript *riu2* denotes the rotation invariant uniform pattern. Another method to identify texture is by measuring it with a rotation invariant measure which contain the contrast of local image texture, defined by:

$$VAR_{L,R} = \frac{1}{L}\sum_{l=0}^{l-1}\left(g_p - \mu\right)^2 \qquad (7)$$

where $\mu = \frac{1}{L}\sum_{l=0}^{l-1} g_L$

This is the case where gray-scale invariance is not required. The resulting image, the VAR image, is invariant against shift in the gray scale.

Later, Ahonen et.al in [14] proposes a class of features that are invariant to image rotation, and computes it along the input histogram rows by using Discrete Fourier Transform (DFT).

Let $H(n,\cdot)$ be the DFT of the *n*th *row* of histogram $h_I(U_p(n,r))$, where

$$H(n,u) = \sum_{r=0}^{L-1} h_I(U_p(n,r))e^{-i2\pi ur/L} \quad (8)$$

In DFT, a cyclic shift of the input vector causes a phase shift in the DFT coefficients. If $h'\left(U_p(n,r)\right) = h(U_p(n,r-a))$, then

$$H'(n,r) = H(n,u)e^{-i2\pi ua/L} \quad (9)$$

The LBP Histogram Fourier can be constructed by [14] :

$$\mathrm{LBP}^{u2}HF(n_1,n_2,u) = H(n_1,u)\overline{H(n_2,u)} \quad (10)$$

where $\overline{H(n_2,u)}$ is complex conjugate of $H(n_2,u)$.

SYSTEM MODEL

In particle filter, the propagation of particle is performed by sample propagation through system model S_k [9][18][19]:

$$S_k = \boldsymbol{A}S_{k-1} + W_{k-1} \quad (11)$$

where A is deterministic component in form of 4 × 4 matrix and w is multivariate Gaussian random variable of size 4 × 1.

This model poses problems when dealing with object with sudden changes in speed and direction, in which no particle falls in the object region. This condition will lead to tracking failure. To mitigate this issue, we need a method to ensure most, or at least, some particles fall in the object. In this paper, we use a matrix, called propagation matrix, and the system model becomes:

$$S_k = \boldsymbol{A}S_{k-1} + \boldsymbol{B}W_{k-1} \quad (12)$$

where B is the propagation matrix with the same size as A. The value of B is set proportional to the size of the object. With this matrix, we can manage the particle distribution to be clustered in a small area or spread in a wider area.

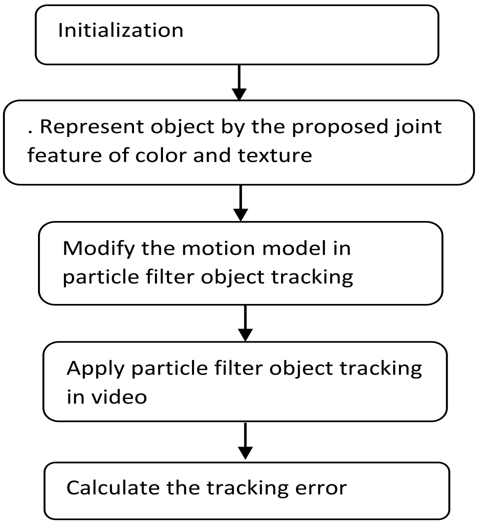

Initialization
. Represent object by the proposed joint feature of color and texture
Modify the motion model in particle filter object tracking
Apply particle filter object tracking in video
Calculate the tracking error

Figure 2. The stages perform in the experiments

THE PROPOSED METHOD OF JOINT FEATURES

In this research, we use particle filter with resampling for object tracking in color video with combining feature of color histogram and texture from Local Binary Pattern Histogram Fourier operator. The distribution of particle is regulated by modified motion model as Eq.(12). We use 4 dimensional feature, where the first three dimensions use RGB color space with 8 × 8 × 8 bins configuration. The fourth dimension is for texture that extracted by Local Binary Pattern Histogram Fourier.

For color, the value of each pixel is assigned to the corresponding bin in the first three dimensions. For texture, first we compute the texture value by Eq.(10) and assign the result into the bin of the fourth dimension.

Similarly, the target candidate is given by:

$$p_{\mathrm{u}}(\mathrm{y}) = C_h \sum_{\mathrm{i}=1}^{n} k\left(\|y - x_i\|^2\right)\delta[h(x_i) - u] \quad (13)$$

The stages perform in the experiment is shown in block diagram (Figure 2).

Detail algorithm of object tracking in particle filter framework using our proposed method is presented below,

1. Initialization
2. Compute reference model $q = C\sum_{i=1}^{M}(k||x_i^*||^2)\delta[b(x_i^* - u)]$
3. Draw a set S_k of N unweighted particles

4. Propagate particles according to motion model
$$S_k = AS_{k-1} + Bw_{k-1}$$
5. Compute color and texture distribution of each sample $p_{s_k}^{(u)}$:
$$p_{s_k^{(n)}}^{(u)} = C_h \sum_{i=1}^{M} k(||s_k^{(n)} - x_i\ ||^2)\delta[b(x_i) - u]$$
6. Compute Bhattacharyya coefficient of each sample
$$\rho_k^{(n)} = \rho\left[p_{s_k^{(n)}}, q\right] = \sum_{u=1}^{m} \sqrt{p_{s_k^{(n)}}^{u} q^{u}}$$
7. Weight particle
$$\omega_k^{(n)} = \frac{1}{N}\left(\frac{1}{\sqrt{2\pi}\,\sigma}\right) e^{\left(\rho_k^{(n)}-1\right)/(2\sigma^2)}$$
8. Estimate object location as mean state of S_k
$E[S_k] = \sum_{n=1}^{N} \omega_k^{(n)} s_k^{(n)}$
9. Resample
10. Repeat step 4 to 9 until the last frame of video

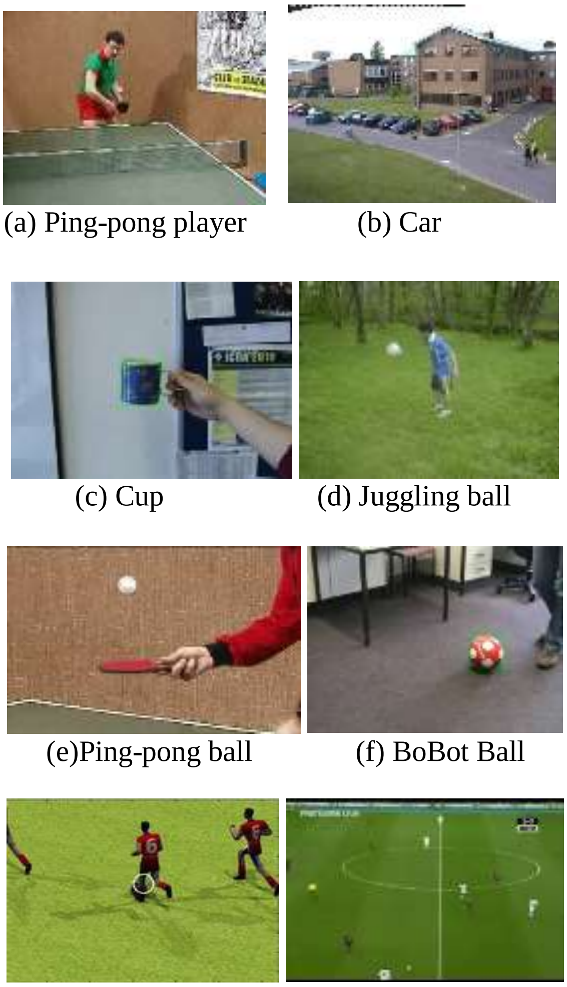
RESULT AND DISCUSSION

Test Videos

We use several videos that have been used in many literatures as test videos. The first video is a slow movement ping-pong player video used in [7] [20]. Car video from PETS database (http://ftp.pets.rdg.ac.uk/ PETS2000/) used in [17] [21] and cup video [22] are the second and third videos. There are similar objects in the car video and scaling and moving background in cup video.

For the fourth we use video of object with changing speed and direction from http://fcl.uncc.edu/nhnguye1/balltracking. html and the ping pong ball video used in various literatures [5] [20] [23] [24] as the 5th video.

Video of moving cam, rotation, and direction changes from Bonn Benchmark on Tracking (BoBot) (http://www.iai.uni-bonn.de/~kleind/tracking/) is used as the sixth test video. This video is used in [22] [25]. Tracking object with occlusion is performed by video from trictrac (http://www.multitel.be/trictrac/) [26] as the seventh video. The eighth video is of soccer player with deformable and dynamic movement. Figure 3 displays the test videos.



(a) Ping-pong player (b) Car

(c) Cup (d) Juggling ball

(e)Ping-pong ball (f) BoBot Ball

(g) Trictrac ball (h) Soccer

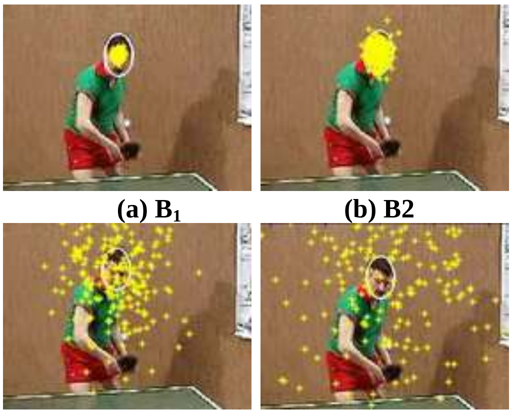
Figure 3. Test videos

Particles Distribution

The distribution of particles defines the success of tracking algorithm. Some particles should fall in the object of interest so tracking can be performed correctly. The proposed motion model acts as a regulator in spreading particles. The value of matrix B in motion model is set proportional to the object dimension.

Figure 4 displays how the motion model controls particle distributions. Small value of B causes particles to cluster in a small area and creates many effective particles, i.e. particles that fall in the object, which contribute in tracking estimate as in Figure 4 (a).

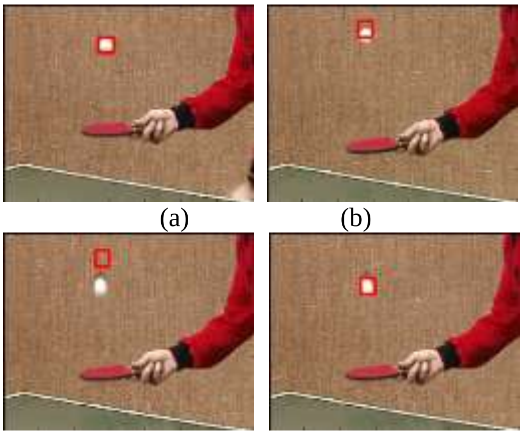
Increasing value of B helps to widen particles distribution, but over spreading will reduce the effective particles, which results in poor performances, as in Figure 4(d).



(a) B_1 **(b) B2**

(c) B_3 **(d) B_4**

Figure 4. Particle distribution in ping-pong player video



(a) (b)

(c) (d)

Figure 5. Tracking ping pong ball (a). Initial frame; (b) (c) Tracking in frame 6 and 8 with the original model (d) Frame 8 with the modified model.

In our experiments, we found that the value of B to give best results is range from ½ to 1 time object dimension. The proposed motion model is useful when dealing with object with changing direction and speed.

Figure 5 displays the robustness of this method in object tracking. In Figure 5, object initially moves upward and suddenly moves downward in frame 5. Original model starts to fail in tracking the object from frame 6 (Figure 5(b) and 5(c)) when object changes its direction. The modified system model tracks object correctly as in Figure 5(d).

JOINT FEATURE OF COLOR HISTOGRAM AND TEXTURE

In this section, we present our experimental results of object tracking in particle filter framework with joint feature of color and texture. Textures are generated using Local Binary Pattern Histogram Fourier, and denoted as PF-LBP-HF. For comparison, we also perform tracking using another variant of LBP. We denote PF-LBP as texture generated by Eq.(5) and PF-LBP-VAR as texture generated by Eq.(7).

The number of particles is 200 and tracking performances are defined by Root Mean Square Error (RMSE).

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}\left((x_i - x_i')^2 + (y_i - y_i')^2\right)}{N}} \qquad (14)$$

Table 1 shows the results. In general, textures that generated from VAR image (PF-VAR) gives the best tracking results, in term of RMSE and computation time. PF-VAR requires less computation time since textures are computed directly from invariant rotation image, compare to PF-LBP that requires mapping beforehand. The property of invariant to shift in gray scale makes PF-LBP-VAR outperforms other textures.

PF-LBP-HF which incorporates DFT in computing texture along the input histogram row contributes to better performance for object that undergoes shape change, such as in ping pong player and car video. In those videos, the objects i.e. ping pong player head and car, changing shape from frame to frame.

The value of propagation matrix is set proportional to object dimension, and depended on the characteristics of the test video. Small value of propagation matrix delivers best results for small movement objects. Larger value of propagation matrix is required when object moves quickly and has abrupt direction changes, to ensure particles fall on the object. Empirically, we found that the value of propagation matrix to give best results is half to one time object dimension.

Combining color and texture is aimed to capture spatial information of object that is not provided by color. The experiments show that in general, texture from VAR image performs better than the one from original LBP and PF-LBP-HF contributes to better performance for object that undergoes shape change.

Table 1. RMSE and computation time of tracking

Video	PF-LBP		PF-LBP (VAR)		PF-LBP-HF	
	RMSE	Time(s)	RMSE	Time (s)	RMSE	Time (s)
Ping pong player	1.8031	11.5145	1.7466	8.257	1.5153	9.8853
Juggling Ball	4.2336	13.2610	1.9006	12.7147	2.1625	13.4110
Ping pong ball	2.5622	9.0219	2.5192	8.647	2.7127	8.1255
Tritrac Ball	6.0012	121.1016	6.0912	102.8547	6.5758	122.2327
BoBoT Ball	4.1773	109.8159	3.6207	94.2532	4.3665	108.9466
Cup	4.1779	157.0013	3.4431	99.1841	21.0241	126.8949
Car	7.1117	28.0786	15.4378	23.9263	3.2144	31.939
Soccer	2.0043	23.6151	2.1604	18.8088	1.7417	23.7578

PF-VAR requires less computation time since textures are computed directly from invariant rotation image, compare to PF-LBP that requires mapping beforehand.

The proposed PF-LBP-HF yields better performance for object that undergoes shape change, such as in ping pong player and car video. In those videos, the objects i.e. ping pong player head and car, changing shape from frame to frame**.**

REFERENCES

[1] Babenko B, Yang MH, and Belongie S, "Visual Tracking with Online Multiple Instance Learning," *IEEE Trans. Pattern Anal. Mach. Intell. PAMI*, 2011.

[2] Kwon J. and Lee K., "Visual tracking decomposition," presented at the Computer Vision and Pattern Recognition, 2010, pp. 1269–1276.

[3] Liu Y., Li G. and Shi Z, "Covariance Tracking via Geometric Particle Filtering," *EURASIP J. Adv. Signal Process.*, vol. 2010, no. 1, p. 583918, Jul. 2010.

[4] Kwon J, Lee H.S, Park F.C, and Lee KM., "A Geometric Particle Filter for Template-Based Visual Tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, Sep. 2013.

[5] Comaniciu D, Ramesh V, and Meer P, "Kernel-based object tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, pp. 564–577, 2003.

[6] Comaniciu D, Ramesh V, and Meer P, "Real-Time tracking of non-rigid objects using mean shift," presented at the IEEE CVPR, 2000, pp. 142–149.

[7] Ning J, Zhang L, Zhang D, and Wu C, "Robust Object Tracking Using Joint Color-Texture Histogram," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 23, no. 7, pp. 1245–1263, 2009.

[8] Teulière C, Marchand E, and Eck L, "A combination of particle filtering and deterministic approaches for multiple kernel tracking.," in *IEEE Int. Conf. on Robotics and Automation, ICRA'09*, 2009, pp. 3948–3954.

[9] Nummiaro K, Koller-Meier E, and Van Gool L, "A color-based particle filter," in *First International Workshop on Generative-Model-Based Vision*, 2002, vol. 2002, p. 01.

[10] Wiharta DM, Hendrantoro G, and Wirawan, "Tracking Fast Moving Object

in Particle Filter Framework," presented at the International Conference on Communications and Networking Application (ICCNA), Indonesia, 2011.

[11] Takala V and Pietikainen M, "Multi-object tracking using color, texture and motion," in *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, 2007, pp. 1–7.

[12] Brasnett P.A, Mihaylova L, Canagarajah N, and Bull D, "Particle filtering with multiple cues for object tracking in video sequences," in *Electronic Imaging 2005*, 2005, pp. 430–441.

[13] Ojala T, Pietikäeinen M, and Mäenpää T, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, 2002.

[14] Ahonen T, Matas J, He C, and Pietikäinen M, "Rotation Invariant Image Description with Local Binary Pattern Histogram Fourier Features," in *Image Analysis*, A.-B. Salberg, J. Y. Hardeberg, and R. Jenssen, Eds. Springer Berlin Heidelberg, 2009, pp. 61–70.

[15] Wiharta DM, Wirawan, and Hendrantoro G, "On the Accuracy of Particle Filter Object Tracking," *Int. J. Multimed. Ubiquitous Eng.*, vol. 10, no. 11, 2015.

[16] Arulampalam M, Maskell S, Gordon N, and Clapp T, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *Signal Process. IEEE Trans. On*, vol. 50, no. 2, pp. 174–188, 2002.

[17] Heikkila M and Pietikainen M, "A texture-based method for modeling the background and detecting moving objects," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 4, pp. 657–662, 2006.

[18] Yang C, Duraiswami R, and Davis L, "Fast multiple object tracking via a hierarchical particle filter," in *Tenth IEEE International Conference on Computer Vision.*, 2005, vol. 1, pp. 212–219.

[19] Li W, Cao J, and Wu D, "Multi-feature Fusion Tracking Based on A New Particle Filter," *J. Comput.*, vol. 7, no. 12, pp. 2939–2947, 2012.

[20] Ning J, Zhang L, Zhang D, and Wu C, "Robust mean-shift tracking with corrected background-weighted histogram," *IET Comput. Vis.*, vol. 6, no. 1, pp. 62–69, Jan. 2012.

[21] Bousetouane F, Dib L, and Snoussi H, "Improved mean shift integrating texture and color features for robust real time object tracking," *Vis. Comput.*, vol. 29, no. 3, pp. 155–170, Feb. 2012.

[22] Klein D, Schulz D, Frintrop S, and Cremers A, "Adaptive real-time video-tracking for arbitrary objects," presented at the International Conference on Intelligent Robots and Systems, 2010, pp. 772–777.

[23] Maggio E. and Cavallaro A, *Video tracking: theory and practice*. John Wiley & Sons, 2011.

[24] Naeem A, Pridmore T, and Mills S, "Managing Particle Spread via Hybrid Particle Filter/Kernel Mean Shift Tracking," presented at the British Machine Vision Conference, 2007.

[25] García GM, Klein DA, Stückler J, Frintrop S, and Cremers AB, "Adaptive Multi-cue 3D Tracking of Arbitrary Objects," in *Pattern Recognition*, 2012, pp. 357–366.

[26] Desurmont X, Hayet J-B, Machy C, Delaigle J-F, and Macq B, "On the performance evaluation of tracking systems using multiple pan-tilt-zoom cameras," in *Videometrics IX*, 2007.