https://doi.org/10.21289/KSIC.2019.22.1.039

Multiple Human Recognition for Networked Camera based Interactive Control in IoT Space

Taeseok Jin[1*]

〈Abstract〉

We propose an active color model based method for tracking motions of multiple human using a networked multiple-camera system in IoT space as a human-robot coexistent system. An IoT space is a space where many intelligent devices, such as computers and sensors(color CCD cameras for example), are distributed. Human beings can be a part of IoT space as well. One of the main goals of IoT space is to assist humans and to do different services for them. In order to be capable of doing that, IoT space must be able to do different human related tasks. One of them is to identify and track multiple objects seamlessly. In the environment where many camera modules are distributed on network, it is important to identify object in order to track it, because different cameras may be needed as object moves throughout the space and IoT space should determine the appropriate one. This paper describes appearance based unknown object tracking with the distributed vision system in IoT space. First, we discuss how object color information is obtained and how the color appearance based model is constructed from this data. Then, we discuss the global color model based on the local color information. The process of learning within global model and the experimental results are also presented.

Keywords : Networked cameras, Object Recognition, Active model, Internet of things

1* Corresponding Author, Professor Dept. of Mechatronics
Zip Code : 47011
E-mail: jints@dongseo.ac.kr, Tel: 051-313-2001

1. Introduce

An intensive research has been going on in the field of IoT space in recent years. The IoT space has to track the objects without failure and to get the location of objects by DINDs seamlessly for these services. Seamless tracking and localization of objects must be achieved in order that the IoT space works properly[1]. In this paper, color appearance based object representation for the distributed vision system in the IoT space is described. First, distributed vision system in the IoT space will be explained. Next, color appearance based object tracking, which is currently possible in the IoT space only, will be proposed. Then, this paper will show how to obtain color information of the objects, and how to achieve the correspondence among different cameras by using the object active color models [2][3].

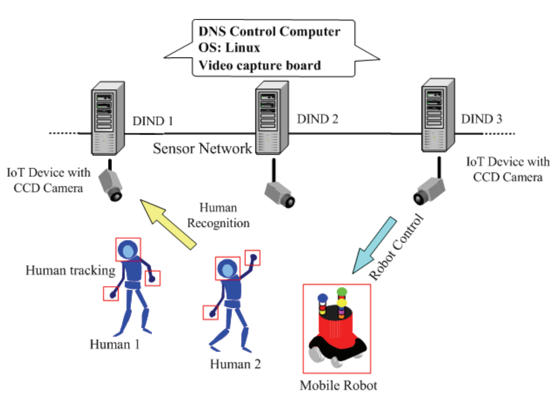


Fig. 1 Concept of IoT space.

2. Acquisition of object information

2.1 Object Finding Process

Object finding process is the process to find the new moving objects in the monitoring area of the camera module. The local color models of new objects are also acquired. Background subtraction is simple and efficient to find the new moving objects in fixed camera image. This background subtraction uses the background model updated from frame to frame adaptively. In IoT space, this background subtraction works well since the lights and the floor are configured as reducing the effects of the shadow and lighting condition. The candidate regions of moving objects are extracted after the dilation, erosion, and clustering to the binary image separated from captured image by comparison with the background image as shown in Fig.2. The small object region is removed as the noise[4][5].

The initial local color model is defined as follows. $\{x_i\}=1,...,n$ is the pixel locations in the region extracted as the object. The function b associates to the pixel at location x_i the index $b(x_i)$ of its bin in the quantized feature space. Feature space is represented by two-dimensional normalized color space, e.g. $r=R/(R+G+B)$, $g=G/(R+G+B)$. The component p_u, $u=1,...,m$ of the feature vector p in the object is then computed as

$$p_u = \frac{1}{n}\sum_{i=1}^{n}\delta[b(x_i)-u] \qquad (1)$$

where δ is the Kronecker delta function.

Since the region extracted by background subtraction is unstable, several sets of p are required for each object in order to stabilize the initial local color model. There is also a probability that multiple objects are found simultaneously. The set of p should be clustered to some categories by the online clustering algorithm. It is decided whether obtained feature vector p belongs to any existing clusters or a new cluster is generated. The number of existing cluster is N at that time. At first, the similarity between feature vector p and each reference vector r_k of cluster is calculated to decide nearest neighbor cluster by Eq.(2). $p_{j,t}$ denotes j-th object at the current time t.

$$S(p_{j,t}, r_{k,t}) = \sum_k \min(p_{j,t},\, r_{k,t}) \qquad (2)$$

It is assumed that c represents the adequate cluster, and it is computed as

$$c = \begin{cases} \arg\max_k S(p_{j,t}, r_{r,t}) & \text{for } S(p_{j,t}, r_{k,t}) > T \\ N+1 & otherwise \end{cases} \qquad (3)$$

where, T is the threshold to evaluate the similarity between feature vectors.

The reference vector of each cluster is updated by Eq.(4). Updated vector is used as the reference vector at the next time $t+1$. α is the learning coefficient.

$$r_{k,t+1} = r_{k,t} + \alpha\, \delta_{ck}\{p_{j,t} - r_{k,t}\} \qquad (4)$$

When the vectors beyond the threshold are gathered in one cluster, object candidate which corresponds with this cluster is treated as the target object. The tracking process for each target object runs at that time as shown in Fig.2.

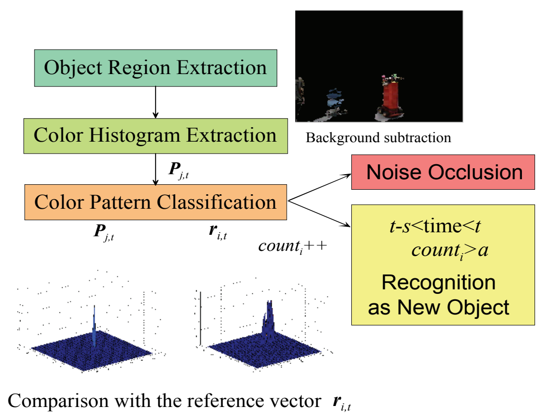


Fig. 2 Object recognition process.

2.2 Mean Shift Tracking Process

Tracking process works for tracking of object region recognized in the object finding process. Tracking process receives the local color model, initial location and size of bounding box from the object finding process.

Recently, the tracking system based on mean shift algorithm is reported that it is suitable for the color region tracking. In this system, weighted mean shift is used for multiple color region tracking. An integrated method mean shift and Kalman filter has been proposed in the previous studies. It has proven to be efficient and relatively robust to the rapid movement of the object. In addition, this method has been compensated the weakness of mean shift tracker with kalman filter. However, in case that the movement of a target is changed suddenly by collision with obstacles such a human, floor, and so on, the tracker loses the target object and there are few chances to recover [6].

By adding changes to this algorithm, we were able to deal with the above problem. Proposed new algorithm exploits color histograms which represent the target feature. The object representation based on the color histogram is relatively stable against deformation and occlusion as mentioned previously. In order to localize the target object, the Mean Shift Tracking Procedure(MSTP) with the Kalman filter application is used. In our research, we added feedback loop after the mean shift procedure as shown in Fig.3. Details of this algorithm are given by our previous paper [7].

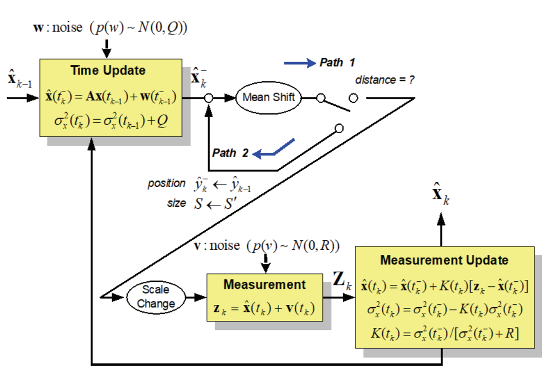


Fig. 3 Tracking process based on mean shift.

3. active model for object identification

3.1 Active Color Model

The active color model means an object model for matching of objects measured by the different camera modules as mentioned above. active model learning process runs when the occlusion among the objects doesn't happen in the tracking process. The active model is produced from local color model which has been measured since the object finding process started. This model should cancel the effects of the object posture, scaling or the direction of measurement by cameras for matching between the different camera modules. At first, it is decided whether the active model is learned, based on objects conditions such as overlapping and approaching. This condition is decided by evaluating the distance between the bounding boxes. Fig. 4 shows the details of active model learning for this process.

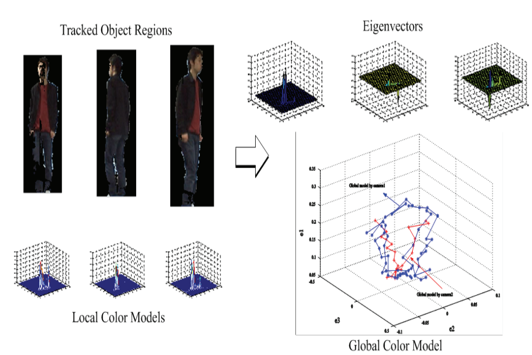


Fig. 4 Active model learning process.

active color model g_k of the object O_k at time t is acquired as follows. The covariance matrix of local color models is computed as,

$$Q = E\{(l_k - \bar{l})(l_k - \bar{l})^T\} \tag{5}$$

where l_k is the set of the local color models obtained until t and $\bar{l}$ is the mean vector for l_k. d eigenvectors $e_1, e_2, \cdots, e_d$ $(\lambda_1 > \cdots > \lambda_d > \cdots > \lambda_m)$ are determined by solving eigenvalues problem:

$$\lambda_k e_k = Q e_k \ . \tag{6}$$

The d-dimensional subspace spanned by these d eigenvectors corresponding to d large eigenvalues is called the eigenspace. By ignoring the small eigenvalues, dimension of the local color model data is reduced. The cumulative proportion of eigenvalues in Eq.(7) is evaluated in order to determine the effective dimension.

$$W_d = \frac{\sum_{i=1}^{d} \lambda_i}{\sum_{i=1}^{m} \lambda_i} > T_s \tag{7}$$

Then, one local color model is projected onto the eigenspace by

$$z_{k,t} = [e_1, \cdots, e_d]^T l_{k,t} \tag{8}$$

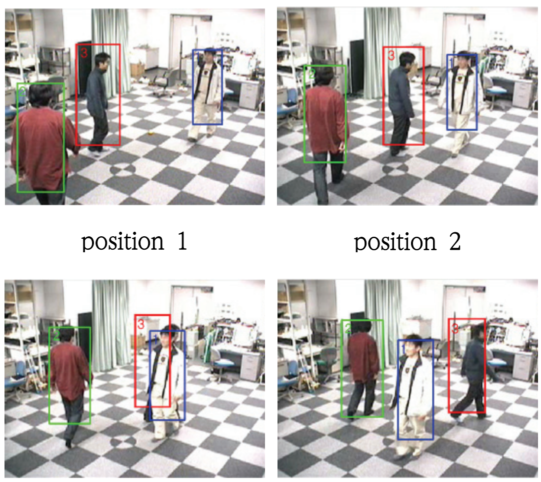
$z_{k,t}$ is a point that the local model $l_{k,t}$ at t is mapped to the eigenspace. The local models to each object can be represented as a manifold in the eigenspace. This manifold includes the local color models changing according to the posture of the object. active color model g_k is represented as this manifold.

3.2 active Model Learning Process

Figure 4 shows the relationship between the acquisition of object information, which includes the object finding process and tracking process, and the process for leaning the active color models. This active model learning process is also included in the DIND. As we have said above, local color models of each object are acquired in every frame that occlusion between objects isn't found. Because the local color models are updated depending on the appearance of the object, changed color models can be obtained in each image frame. The set of these local color models is stored into the active color model learning process. Eigenvalues and eigenvectors for the set of the local models of each object are calculated individually in this process.

4. Experiments

We implemented the proposed method to represent walking human position as shown in Fig.5, used a active color model to tracking the top positions of each person and implemented a MSTP proposed in section III. Also, tracking experiments in which our method coped with the sudden change of the object movement were performed. Figure 6 demonstrates a typical eigenvector of three walking humans in a convoy, where the active color models of position are shown.



position 1 position 2

position 3 position 4

Fig. 5 Multiple human tracking results.

Figure 7 shows the routes of three humans tracking, respectively. Objects, three humans, were tracked by the tracking process which was performed independently with the other process. In Fig. 7, the active color model obtained in the different camera is compared with the tracked humans as shown in Fig. 7. A lot of local color models are projected onto the eigenspace spanned by three eigenvectors.

The active model is represented as the data sets of the local models compressed in this case. The cumulative proportion of eigenvalues in Eq.(8) and the threshold generally set to 0.8 or 0.9 to determine the effective dimension. This model can represent the change of the color appearance of the tracked human. Although the local color models change according to the difference of the camera, correspondence among cameras can be evaluated by the comparison of the manifold shape. The active models by camera2 and camera3 are in process of the complete active model.

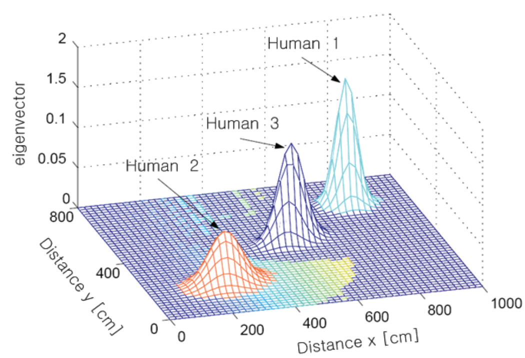


Fig. 6 Active color models of three humans.

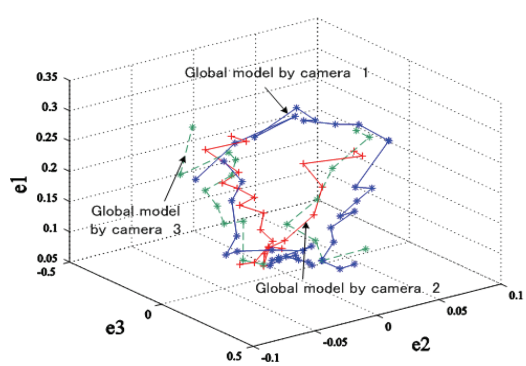


Fig. 7 Comparison of the active Models.

5. Conclusion

IoT space using network-based-visions techniques provide promising ways to human-computer interaction through understanding human and object movements from visual data. An important step in achieving this goal is the robust and accurate tracking of the moving objects such as walking human and mobile agents. However, cluttered backgrounds, unknown lighting conditions and multiple moving objects make the tracking tasks challenging. This paper mainly concentrated on active color model based on the local color information for the objects tracking by addressing these difficulties.

Acknowlegements

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education(No. 2016R1D1A1B04932124).

References

[1] Xie, B., Kumar, A., Chen, C.-M., Pandey, S., Nepal, S., Gubbi, J., Buyya, R., Marusic, S. and Palaniswami, M. "2013. Internet of Things (IoT): A vision, architectural elements, and future directions, "Future Generation Computer Systems. vol.29, no.7, pp.1645–1660, (2013).

[2] T.Matsuyama and N.Ukita, "Real-Time Multi-Target Tracking by a Cooperative Distributed Vision System," Proc. IEEE, vol.90, No.7, pp.1136-1150, (2002).

[3] H.Hashimoto, "intelligent space -How to Make Spaces Intelligent by using DIND," Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (SMC'02), pp.14-19, (2002).

[4] P. Balraj Balla, K.T. Jadhao, "IoT Based Facial Recognition Security System," 2018 International Conference on Smart City and Emerging Technology (ICSCET), pp.1-4, (2018).

[5] S. Amendola, R. Lodato, S. Manzari, C. Occhiuzzi, G. Marrocco, "RFID technology for iot-based personal healthcare in smart spaces," Internet of Things Journal, IEEE, vol.1, no.2, pp. 144–152, (2014).

[6] M.A. Hossain, A.A.N. Shirehjini, A.S. Alghamdi, A. El Saddik, "Adaptive interaction support in ambient-aware environments based on quality of context information," Multimedia Tools and Applications, vol.67, no.2, pp.409–432, (2013).

[7] B. Lavanya, G.S. Gayathri, "Exploration and Deduction of Sensor-Based Human Activity Recognition System of Smart-Phone Data," Computational Intelligence and Computing Research (ICCIC)IEEE International Conference on, pp.1-5, (2017).

(Manuscript received September 28, 2018;
revised December 31, 2018; accepted January 5, 2019)