H-064

動画像からの樹木の葉の追跡法の基礎的検討ーカルマンフィルターとパーティクフィルターの利用法の検討

Studies of the Kalman filer and Particle filter method for tracking the moving botanical tree leaves in video sequence.

Do N. Khanh† Ohya Jun†

1. Introduction

The authors have been working on reconstructing the dynamical behavior of a moving non-rigid object such as smoke, water and botanical tree. It is very difficult to track the behavior of such objects due to their complicated structure and behavior. Among these non-rigid objects, this paper reports our recent studies of tracking botanical trees' leaves' behaviors.

Foliage can be tracked based on the well-known maximally stable extreme region (MSER) detectors as proposed in [1], but this method can only track a leafy botanical tree and branch is not tracked. When Gordon, Salmon and Smith (1993) introduced a particle filter tracking method to computer vision society, they demonstrated how to track a leaf in clutter using contour based tracking [2], but it is computationally expensive because of large dimensionality of the state space.

Our research is to develop a new algorithm which can track a non-leafy botanical tree at low computation cost. In this paper, we studies two methods based on Kalman filter and Particle filter. The Kalman filter based method uses geometric features such as bounding box, area, centroid of the object to track the dynamical behavior of a simple botanical tree moving under windy condition. The particle filter based method uses color feature and depth image data, which is acquired from a low-cost infrared camera such as Microsoft Kinect, to track the botanical tree behavior under windy condition.

The rest of this paper is organized as follows. Section 2 explains the Kalman filter based method for tracking dynamical behaviors of a botanical tree and the experiment result. Section 3 explains the Particle filter based method for tracking the dynamical behaviors of a botanical tree and the experiment result. Section 3 concludes this paper with discussion about the advantages and disadvantages of each method.

2. Kalman Filter Based Method for Tracking Dynamical Behaviors of a Botanical Tree

2.1. Overview the method

To increase the tracking accuracy and decrease the computation cost, we do not track a certain leaf by searching for the leaf all over the image. Instead, a leaf is only searched for in its region of interest (ROI). In order to do that, we use one frame delay concept and it is described as follows. An ROI, a rectangle, which is large enough to contain a leaf in the first frame and the second frame, is defined. The leaf's motion vector between the first frame and second frame is used as the motion vector of the ROI between the second and the third frame. Therefore, the motion vector of the ROI between (t-1)th frame and tth frame is defined by the motion vector of the leaf between the (t-2)th frame and (t-1)th frame, where t is the current frame.

The state vector is (x, y, x', y', h, w), where x and y are the top-left corner of the bounding box of the leaf, respectively; x' and y' are the velocity of that corner in x-axis and y-axis, respectively; h and w is the height and width of the bounding box, respectively.

The motion of a leaf is computed by process equation with process noise N(0, σ^2) and updated by measurement equation with measurement noise N(0, σ^2).

2.2 Experiment and Result

We track three different leaves of a simple botanical tree in 220 frames. Figure 1 shows that the tracker can follow the leaf during the whole sequence. Moreover, the tracker can follow the leaf even if a part of the leaf is outside of the image.

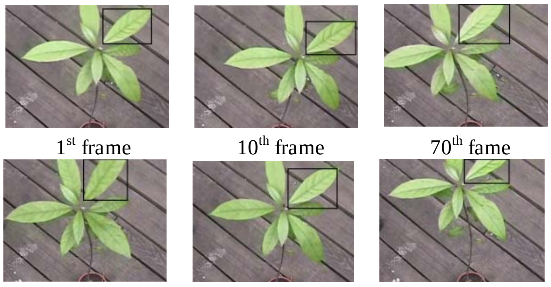

Figure 1. Leaf tracked using geometric feature. The tracker is able to follow the leaf during the whole sequence.

However, the tracker cannot correctly deal with cases in which one leaf occludes the other leaf, as shown in Fig. 2.

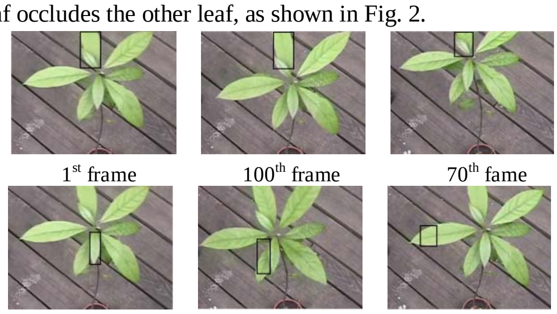

Figure 2. Leaf tracked using geometric feature. The tracker can not follow the leaf during the whole sequence.

3. Particle Filter Based Method for Tracking Dynamical Behaviors of a Botanical Tree

3.1. Overview the method

a. Initialization and Particle Definition

The object's coordinates (X, Y, and Z) as well as the height and width are obtained from analyzing the first frame. Object

† Graduate School of Global Information and Telecomunication Studies, Waseda University

thickness is given a known value. Initial color histogram of the object to be tracked is obtained by adding the histogram of a* color component and the histogram of b* color component after converting the color of the selected object in RGB color space into L*a*b color space.

A particle is defined as state vector (X, Y, Z, H, W, T)', where X and Y are the coordinates of the top-left corner of the bounding box in the RGB image (as well as in the depth image); Z is the distance between the camera and a point on the object closest to the camera; H and W are the height and width of the bounding box, respectively; T is the object thickness. Each particle has its weight w.

b. Motion model

First, a white Gaussian noise $N(0, \sigma^2)$ is added to Z_{t-1} to generate Z_t as indicated in Eq. (1)

$$Z_t = Z_{t-1} + N(0, \sigma^2) \tag{1}$$

Second, the depth image of the t^{th} frame is segmented to obtain all the objects' parts whose depth values are within the range between Z_t and (Z_t+T). A convenient way to describe the segmented parts is to use the binary image, because in the subsequent processes it is used to detect the part closest to the object being tracked at the previous, $(t\text{-}1)^{th}$ frame. Many objects and their parts that have similar depth range (between Z_t and (Z_t+T)) are included in the segmentation result. Therefore, blob analysis is performed to find the blob's bounding box and its centroid. After that, Euclidean distance between each blob and the tracked object's position (X_{t-1}, Y_{t-1}) in the previous frame $(t\text{-}1)^{th}$ is compared to find the blob closest to the tracked object. That blob position is (X_{blob}, Y_{blob}) and its bounding box's height and width are H_{blob} and W_{blob}, respectively. This result and Eq. (1) define the particle state x_t at frame t as indicated below.

$$x_t = (X_{blob}, Y_{blob}, Z_t, H_{blob}, W_{blob}, T)' \tag{2}$$

c. Observation model

The observation model computes the likelihood (i.e. the weight) of each particle by comparing the initial color histogram and the color histogram of cropped RGB image. The cropped RGB image is obtained from Eq. (2): its top-left point's coordinates are (X_{blob}, Y_{blob}) and its height and width are H_{blob}, W_{blob}, respectively.

The process for comparing the two color histograms is as follow. First, the image in RGB color space is converted into L*a*b color space. Second, the histogram of a* color component and the histogram of b* color component are added together to obtain one color histogram. Third, the Kullback–Leibler divergence (KLIC) [5] is used to measure the difference between the initial color histogram and the color histogram of the cropped RGB image. If the two color histograms are similar, the KLIC is small and vice-versa.

The likelihood is computed as indicated by Eq. (3).

$$p(z_t|x_t) = e^{-\lambda d} \tag{3}$$

The left side of Eq. (3) is the likelihood at the frame t, given state x_t. λ is a constant, and d is the KLIC. The likelihood value is high if two color histograms are similar and vice-verse.

The likelihood values are normalized to obtain the particle weight.

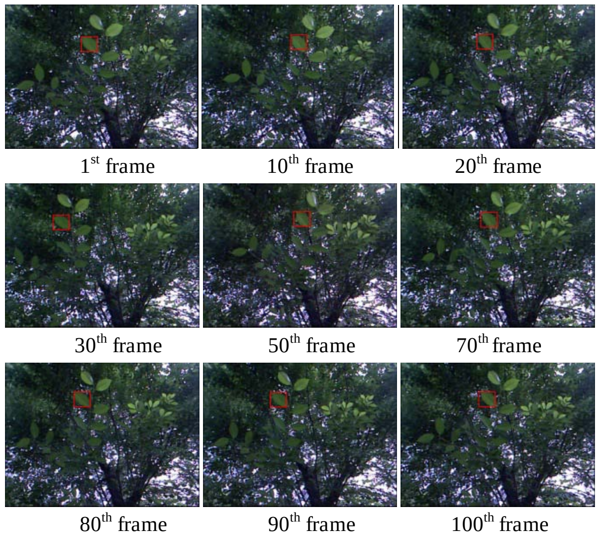

Figure 3. Tracking a leaf using proposed method. The tracker is able to follow object which has similar color to background during the whole sequence.

The resampling process is simply done as described in [2] and [4].

The mean of particles' state is the state estimation.

3.2. Experiment and Result

We use the proposed method to track a green leaf moving in a clutter of other leaves (also in green color) in 100 frames. The particle number is 100. The result is shown in Fig. 3, where the tracker can follow the object during the whole sequence.

7. Conclusion

Kalman filter based method is fast, but it could not track the leaf among a clutter of other leaves. Particle filter is more robust to this problem, but its computation cost is quite high. However, it is low cost comparing to contour based tracking particle filter as proposed in [2], because the state vector has only 6 dimensions and the number of particle filter is only 100.

Our future work is to develop a tracking method based on particle filter and utilizing depth image to track all the leaves and branches together. After that, we will reconstruct a 3D botanical tree model that behaves naturally.

References

[1] Li, C., Hall, P.M., "*Learning A Stable Structure To Describe Dynamic Texture*"; BMVC08.

[2] Gordon, N.J.; Salmond, D.J.; Smith, A.F.M.;, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation ," *Radar and Signal Processing, IEE Proceedings F* , vol.140, no.2, pp.107-113, Apr 1993.

[3] Changjiang Yang; Duraiswami, R.; Davis, L.;, "Fast multiple object tracking via a hierarchical particle filter," *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on* , vol.1, no., pp. 212- 219 Vol. 1, 17-21 Oct. 2005.

[4] Arulampalam, M.S.; Maskell, S.; Gordon, N.; Clapp, T.; , "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *Signal Processing, IEEE Transactions on* , vol.50, no.2, pp.174-188, Feb 2002.

[5] C. Bishop (2006); *Pattern Recognition and Machine Learning*; Springer; p. 55.