LETTER

A Novel View of Color-Based Visual Tracker Using Principal Component Analysis

Kiyoshi NISHIYAMA[†a)] *and* Xin LU[†b)], *Members*

SUMMARY An extension of the traditional color-based visual tracker, i.e., the continuously adaptive mean shift tracker, is given for improving the convenience and generality of the color-based tracker. This is achieved by introducing a probability density function for pixels based on the hue histogram of object. As its merits, the direction and size of the tracked object are easily derived by the principle component analysis (PCA), and its extension to three-dimensional case becomes straightforward.
key words: *visual tracking, mean shift, principle component analysis, Gaussian probability density function*

1. Introduction

The continuously adaptive mean shift tracker [1] has been proposed for color-based visual tracking in a lot of application fields (e.g., user interface, surveillance, and so on). The tracker used the mean shift method [2], [3] for searching the location of the tracked region and determining its size [4]. The mean shift method calculates the centroid (center of mass) and the minimum moment of inertia of tracked region using the mechanical moments based on a mass density, and then obtains the direction and size of object by regarding the region as an ellipse with the same moment of inertia.

However, the derivation of direction and size has been considerably complicated. To solve this drawback, we introduce a novel view that the fitted ellipse is represented by a horizontal section of a Gaussian probability density function to be given for pixels. Based on the cross section, the direction and size of the object can be easily calculated by a pure statistical method — principle component analysis (PCA). Then its size can be adjusted by increasing or decreasing the probability of object, whereas the estimated size in the mechanical method tends to be smaller than the actual one. Moreover, tracking a multidimensional object by the statistical method is much easier than that by the mechanical method because the former accomplishes it by the eigenvalue decomposition. Thus replacing the mechanical approach with the statistical one improves the convenience and generality of the color-based tracker.

2. The Color-Based Tracking

Since the tracked object is defined by a region centered at (x^c, y^c) with $2l \times 2w$ pixels size at angle θ to the horizontal axis in a frame as shown in Fig. 1, the state vector for the visual tracking can be defined as $s_k = [x_k^c\ y_k^c\ l_k\ w_k\ \theta_k]^T$. The state tracker based on the hue channel of hue-saturation-value (HSV) color space is described in the following.

2.1 Tracked Region and Its Histogram in the HSV

Definitions of the tracking elements are given here. Suppose that the coordinates of a pixel is denoted by

$$\tau = (x, y),\ x = 1, \ldots, L,\ y = 1, \ldots, W \qquad (1)$$

which exists at a point written as

$$I_{HSV}(\tau) = (h(\tau), s(\tau), v(\tau)) \qquad (2)$$

in the HSV color space, where L and W are the length and width of the frame; h, s, and v are the hue, saturation, and value of a pixel at τ in the HSV color space [5]. The template of tracked object is selected manually as the following region:

$$O^{tmp} = O_0 = \{\tau_0^i,\ i = 1, \ldots, n_0\} \qquad (3)$$

in the first frame. The template consists of n_0 pixels in a rectangle with angle θ_0 to the horizontal axis, and centered at τ_0^c with $2l_0 \times 2w_0$ size. By quantizing $I_{HSV}(\tau)$ at time k as $\underline{I}_{HSV,k}(\tau) = (\underline{h}_k(\tau), \underline{s}_k(\tau), \underline{v}_k(\tau))$, the reference model of the tracked region (template) O^{tmp} in the HSV color space can be represented by its hue histogram:
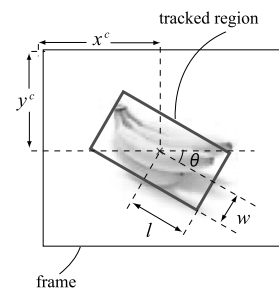
Fig. 1 Tracked region of an object.

$$H^{tmp}(\underline{h}^i) = \sum_{\tau \in O^{tmp}} 1(\underline{h}_0(\tau) - \underline{h}^i)$$

$$\underline{h}^i \in \{0, \dots, m-1\}, \quad 1(x) = \begin{cases} 1, & \text{if } x = 0 \\ 0, & \text{if } x \neq 0 \end{cases} \tag{4}$$

where m is the color level fixed to 256 (8-bit). In the subsequent frame, on updated region $O_k = \{\tau_k^i = (x_k^i, y_k^i), i = 1, \dots, n_k\}$, $k > 0$ is centered at $\tau_k^c = (x_k^c, y_k^c)$ with $2l_k \times 2w_k$ size.

2.2 Search Region and Mass Density

The search for the center τ_k^c of the current updated region O_k starts at the location of the previous updated region O_{k-1} centered at τ_{k-1}^c with $2l_{k-1} \times 2w_{k-1}$ size in the current frame. Then the search region is enlarged to $2(l_{k-1}+\delta) \times 2(w_{k-1}+\delta)$ with the fixed δ (i.e., an increasing factor for the search region) which is denoted by $\check{O}_k = \{\check{\tau}_k^j = (\check{x}_k^j, \check{y}_k^j), j = 1, \dots, \check{n}_k\}$.

For determining the current updated object region from the search region $\check{O}_k$, it is necessary to match the template O^{tmp} in the current search region. To accomplish it, we consider a weight function:

$$F_k(\check{\tau}) = H^{tmp}(\underline{h}_k(\check{\tau})), \quad \check{\tau} \in \check{O}_k \tag{5}$$

which is regarded, in a sense, as a mass density of the pixel at $\check{\tau}$.

3. Visual Tracker Using Mechanical Moments

How to calculate the location and size of tracked object using the mechanical moments in the original view is specified here. Given the centroid of the tracked object:

$$\hat{\tau}_k^c = (\hat{x}_k^c, \hat{y}_k^c), \quad \hat{x}_k^c = \frac{\mathcal{M}_k^{10}}{\mathcal{M}_k^{00}}, \quad \hat{y}_k^c = \frac{\mathcal{M}_k^{01}}{\mathcal{M}_k^{00}} \tag{6}$$

the moment of inertia of the tracked object based on the rotation axis through the centroid at the angle θ_k as shown in Fig. 2 can be written as
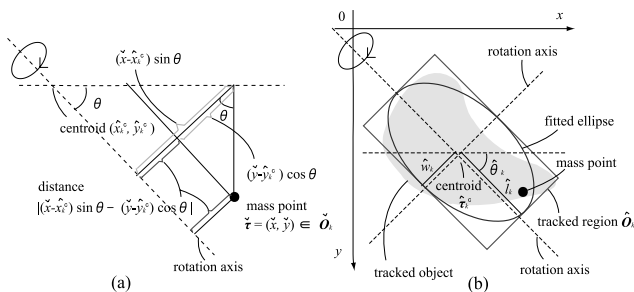
Fig. 2 Determination of the tracked region; (a) the moment of inertia of a mass point, (b) the tracked region with $2\hat{w}_k \times 2\hat{l}_k$ size based on the rotation axis through the centroid τ_k^c at the angle $\hat{\theta}_k$.

$$\mathcal{M}_k^\theta$$
$$= \sum_{\check{\tau}=(\check{x},\check{y})\in\check{O}_k} \left[(\check{x} - \hat{x}_k^c)\sin\theta_k - (\check{y} - \hat{y}_k^c)\cos\theta_k\right]^2 F_k(\check{\tau})$$
$$= \frac{1}{2}\left[\mathcal{M}_k^{02} - \frac{(\mathcal{M}_k^{01})^2}{\mathcal{M}_k^{00}} - \mathcal{M}_k^{20} + \frac{(\mathcal{M}_k^{10})^2}{\mathcal{M}_k^{00}}\right]\cos 2\theta_k$$
$$\quad - \left[\mathcal{M}_k^{11} - \frac{\mathcal{M}_k^{01}\mathcal{M}_k^{10}}{\mathcal{M}_k^{00}}\right]\sin 2\theta_k$$
$$\quad + \frac{1}{2}\left[\mathcal{M}_k^{02} - \frac{(\mathcal{M}_k^{01})^2}{\mathcal{M}_k^{00}} + \mathcal{M}_k^{20} - \frac{(\mathcal{M}_k^{10})^2}{\mathcal{M}_k^{00}}\right] \tag{7}$$

where $\mathcal{M}_k^{00}$ is the zeroth moment; $\mathcal{M}_k^{10}$ and $\mathcal{M}_k^{01}$ are the first moments; $\mathcal{M}_k^{11}$, $\mathcal{M}_k^{20}$, and $\mathcal{M}_k^{02}$ are the second moments for the search region as follows:

$$\mathcal{M}_k^{00} = \sum_{\check{\tau}\in\check{O}_k} F_k(\check{\tau}), \quad \mathcal{M}_k^{11} = \sum_{\check{\tau}=(\check{x},\check{y})\in\check{O}_k} \check{x}\check{y}F_k(\check{\tau}) \tag{8}$$

$$\mathcal{M}_k^{10} = \sum_{\check{\tau}=(\check{x},\check{y})\in\check{O}_k} \check{x}F_k(\check{\tau}), \quad \mathcal{M}_k^{20} = \sum_{\check{\tau}=(\check{x},\check{y})\in\check{O}_k} \check{x}^2 F_k(\check{\tau}) \tag{9}$$

$$\mathcal{M}_k^{01} = \sum_{\check{\tau}=(\check{x},\check{y})\in\check{O}_k} \check{y}F_k(\check{\tau}), \quad \mathcal{M}_k^{02} = \sum_{\check{\tau}=(\check{x},\check{y})\in\check{O}_k} \check{y}^2 F_k(\check{\tau}). \tag{10}$$

In order to obtain the minimum of the moment of inertia, setting the partial derivative of $\mathcal{M}_k^\theta$ with respect to θ_k to zero, we have

$$\hat{\theta}_k = \arctan\left[\frac{m_{11}}{m_{20} - m_{02}}\right]/2 \tag{11}$$

which leads to

$$\tan 2\hat{\theta}_k = \frac{m_{11}}{m_{20} - m_{02}} \tag{12}$$

where

$$m_{20} = \frac{\mathcal{M}_k^{20}}{\mathcal{M}_k^{00}} - (\hat{x}_k^c)^2, \quad m_{02} = \frac{\mathcal{M}_k^{02}}{\mathcal{M}_k^{00}} - (\hat{y}_k^c)^2$$

$$m_{11} = 2\left(\frac{\mathcal{M}_k^{11}}{\mathcal{M}_k^{00}} - \hat{x}_k^c\hat{y}_k^c\right). \tag{13}$$

Substituting (11) into (7) provides the minimum of $\mathcal{M}_k^\theta$:

$$\mathcal{M}_k^{\hat{\theta}} = \frac{1}{2}\mathcal{M}_k^{00}(m_{20} + m_{02} - m_{22}) \tag{14}$$

which uses the following relationship:

$$\sin 2\hat{\theta}_k = \frac{m_{11}}{m_{22}}, \quad \cos 2\hat{\theta}_k = \frac{m_{20} - m_{02}}{m_{22}} \tag{15}$$

$$m_{22} = \sqrt{m_{11}^2 + (m_{20} - m_{02})^2}. \tag{16}$$

Furthermore, representing the tracked object by an ellipse whose mass density is uniform in the area of $\pi\hat{l}_k\hat{w}_k$ as

$$d_{e,k}(\tau) = \frac{\mathcal{M}_k^{00}}{\pi\hat{l}_k\hat{w}_k} = c, \quad \tau \in O_{e,k} \tag{17}$$

we can express the moment of inertia of the ellipse as

$$\mathcal{M}_{e,k}^{\hat{\theta}} = \frac{1}{4}\pi \hat{l}_k \hat{w}_k^3 c = \frac{1}{4}\hat{w}_k^2 \mathcal{M}_k^{00}. \tag{18}$$

Note that the semimajor axis $\hat{l}_k$ of the ellipse agrees with a rotation axis of the tracked object. Setting $\mathcal{M}_{e,k}^{\hat{\theta}}$ equal to $\mathcal{M}_k^{\hat{\theta}}$ (i.e., $\mathcal{M}_{e,k}^{\hat{\theta}} = \mathcal{M}_k^{\hat{\theta}}$), we obtain

$$\hat{w}_k = \sqrt{\frac{4\mathcal{M}_k^{\hat{\theta}}}{\mathcal{M}_k^{00}}} = \sqrt{2(m_{20} + m_{02} - m_{22})}. \tag{19}$$

Similarly, we can get

$$\mathcal{M}_{e,k}^{\hat{\theta}+\frac{\pi}{2}} = \frac{1}{4}\pi \hat{w}_k \hat{l}_k^3 c = \frac{1}{4}\hat{l}_k^2 \mathcal{M}_k^{00} = \mathcal{M}_k^{\hat{\theta}+\frac{\pi}{2}} \tag{20}$$

where $\mathcal{M}_{e,k}^{\hat{\theta}+\frac{\pi}{2}}$ and $\mathcal{M}_k^{\hat{\theta}+\frac{\pi}{2}}$ are the moments of inertia of the ellipse and tracked object based on the new rotation axis (i.e., the semiminor axis $\hat{w}_k$) perpendicular to the previous one. Since the following equation:

$$\mathcal{M}_k^{\hat{\theta}+\frac{\pi}{2}} = \frac{1}{2}\mathcal{M}_k^{00}(m_{20} + m_{02} + m_{22}) \tag{21}$$

is inferred from (7) recalling

$$\sin(2\hat{\theta}_k + \pi) = -\frac{m_{11}}{m_{22}}$$
$$\cos(2\hat{\theta}_k + \pi) = \frac{m_{02} - m_{20}}{m_{22}} \tag{22}$$

we obtain

$$\hat{l}_k = \sqrt{\frac{4\mathcal{M}_k^{\hat{\theta}+\frac{\pi}{2}}}{\mathcal{M}_k^{00}}} = \sqrt{2(m_{20} + m_{02} + m_{22})}. \tag{23}$$

Consequently, the tracked object is estimated by the fitted ellipse as shown in Fig. 2(b).

4. Visual Tracker Using Statistical Moments

4.1 Definition of the Pixel's Probability

To calculate the statistical moments of the object, we consider a probability measure, using $F_k(\check{\tau})$, as

$$P_k(\check{\tau}) = \frac{F_k(\check{\tau})}{\sum_{\check{\tau} \in \check{O}_k} F_k(\check{\tau})} \tag{24}$$

which denotes the likelihood that a pixel of the object appears at $\check{\tau}$ in the search region $\check{O}_k$. This probability measure is referred to as the pixel's probability. The relationship between $F_k(\check{\tau})$, $H^{tmp}(\underline{h}^i)$, and $P_k(\check{\tau})$ is illustrated in Fig. 3.

4.2 Derivation of a New Visual Tracker

Our idea is to regard the tracked object as a horizontal section of the following probability density function for continuous pixels variable:

$$p_{z_k}(z_k) = \frac{1}{(2\pi)|\Sigma_{z_k}|^{\frac{1}{2}}} e^{-\frac{1}{2}(z_k-\bar{z}_k)^T \Sigma_{z_k}^{-1}(z_k-\bar{z}_k)}. \tag{25}$$
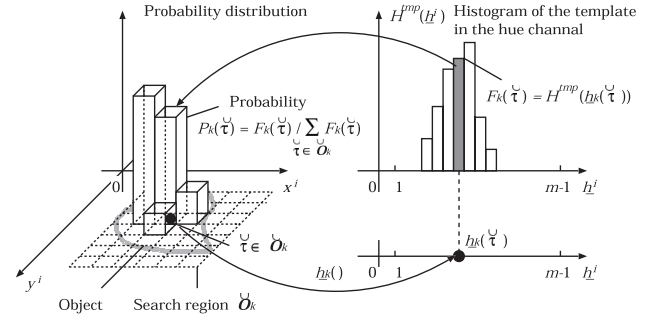
Fig. 3 Definition of the pixel's probability based on the hue histogram of object.

The integration $\int_D z_k p_{z_k}(z_k) dz_k$ means the probability that pixels of the object appear in the region $D \subset \mathcal{R}^2$. Here the mean and covariance of $z_k \in \mathcal{R}^2$ are given by

$$\bar{z}_k = \begin{bmatrix} \mu_{x,k} \\ \mu_{y,k} \end{bmatrix}, \quad \Sigma_{z_k} = \begin{bmatrix} \sigma_{x,k}^2 & \gamma_k \\ \gamma_k & \sigma_{y,k}^2 \end{bmatrix} \tag{26}$$

whose elements are defined with the pixel's probability as

$$\mu_{x,k} = \sum_{\check{\tau}=(\check{x},\check{y})\in\check{O}_k} \check{x} P_k(\check{\tau}), \quad \mu_{y,k} = \sum_{\check{\tau}=(\check{x},\check{y})\in\check{O}_k} \check{y} P_k(\check{\tau}) \tag{27}$$

$$\gamma_k = \sum_{\check{\tau}=(\check{x},\check{y})\in\check{O}_k} (\check{x} - \mu_{x,k})(\check{y} - \mu_{y,k}) P_k(\check{\tau}) \tag{28}$$

$$\sigma_{x,k}^2 = \sum_{\check{\tau}=(\check{x},\check{y})\in\check{O}_k} (\check{x} - \mu_{x,k})^2 P_k(\check{\tau}) \tag{29}$$

$$\sigma_{y,k}^2 = \sum_{\check{\tau}=(\check{x},\check{y})\in\check{O}_k} (\check{y} - \mu_{y,k})^2 P_k(\check{\tau}). \tag{30}$$

The first and second principal axes of the density function are obtained by solving the eigenvalue problem of the covariance Σ_{z_k}. Supposing that λ_k is the eigenvalue of Σ_{z_k}, we can reduce

$$|\Sigma_{z_k} - \lambda_k I| = \begin{vmatrix} \sigma_{x,k}^2 - \lambda_k & \gamma_k \\ \gamma_k & \sigma_{y,k}^2 - \lambda_k \end{vmatrix} = 0 \tag{31}$$

to

$$(\sigma_{x,k}^2 - \lambda_k)(\sigma_{y,k}^2 - \lambda_k) - \gamma_k^2 = 0. \tag{32}$$

From this, we obtain the eigenvalues:

$$\lambda_k = \frac{\sigma_{x,k}^2 + \sigma_{y,k}^2 \pm \xi_k}{2} \tag{33}$$

with

$$\xi_k = \sqrt{4\gamma_k^2 + (\sigma_{x,k}^2 - \sigma_{y,k}^2)^2}. \tag{34}$$

The corresponding eigenvectors, which are considered as two orthogonal principal axes and ordered based on their eigenvalues, determine a new coordinate system in the conception of the PCA. Furthermore, we have

$$\lambda_{k,max}^{-1}\|\tilde{z}_k\|^2 \le \tilde{z}_k^T \Sigma_{z_k}^{-1} \tilde{z}_k \le \lambda_{k,min}^{-1}\|\tilde{z}_k\|^2 \tag{35}$$
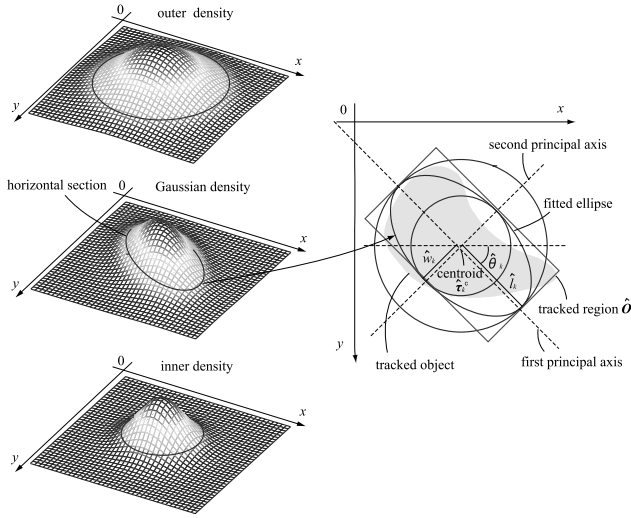
Fig. 4 Representation of the tracked object by the Gaussian density with its outer and inner bound.

according to the Rayleigh quotient with

$$\tilde{z}_k = z_k - \bar{z}_k \qquad (36)$$

where $\lambda_{k,min}$ and $\lambda_{k,max}$ are the minor and major eigenvalues of Σ_{z_k}, respectively. The relationship of (35) suggests that the Gaussian density function $p_{z_k}(z_k)$ is bounded by

$$p_{z_k,in}(z_k) \le p_{z_k}(z_k) \le p_{z_k,out}(z_k) \qquad (37)$$

where the outer and inner density functions are expressed by

$$p_{z_k,in}(z_k) = \frac{1}{(2\pi)\lambda_{k,min}} e^{-\frac{1}{2}\lambda_{k,min}^{-1}\|\tilde{z}_k\|^2} \qquad (38)$$

$$p_{z_k,out}(z_k) = \frac{1}{(2\pi)\lambda_{k,max}} e^{-\frac{1}{2}\lambda_{k,max}^{-1}\|\tilde{z}_k\|^2} \qquad (39)$$

respectively, as shown in Fig. 4.

Then a horizontal section of the Gaussian density function represents the tracked object, which can be regarded as an ellipse whose semimajor axis is along the major eigenvector $v_{k,max} = [\sigma_{x,k}^2 - \sigma_{y,k}^2 + \xi_k \ 2\gamma_k]^T$ at the following angle $\hat{\theta}_k$:

$$\hat{\theta}_k = \arctan\left[\frac{2\gamma_k}{\sigma_{x,k}^2 - \sigma_{y,k}^2 + \xi_k}\right]. \qquad (40)$$

Additionally, we can obtain the length of the tracked object as

$$\hat{l}_k = \sqrt{\rho\lambda_{k,max}} = \sqrt{\rho(\sigma_{x,k}^2 + \sigma_{y,k}^2 + \xi_k)} \qquad (41)$$

where $\hat{l}_k$ is the semimajor axis of the ellipse; ρ is a scalar parameter to adjust the elliptic area of the horizontal section in the Gaussian density function. Similarly, we can obtain the width as

$$\hat{w}_k = \sqrt{\rho\lambda_{k,min}} = \sqrt{\rho(\sigma_{x,k}^2 + \sigma_{y,k}^2 - \xi_k)} \qquad (42)$$

where $\hat{w}_k$ is the semiminor axis of the ellipse. As a result,

the tracked object is obtained as an ellipse, which is centered at $\hat{\tau}_k^c = (\mu_{x,k}, \mu_{y,k})$ with $2\hat{l}_k \times 2\hat{w}_k$ size at angle $\hat{\theta}_k$ to the horizontal axis, as shown in Fig. 4. Also, the tracked region is the tangential rectangle of the ellipse, which are described as the window enclosed by red line. As the tracking process is implemented repeatedly from the first frame to the current frame, a sequence of the tracking results defined by

$$\hat{s}_i = [\hat{x}_i^c \ \hat{y}_i^c \ \hat{l}_i \ \hat{w}_i \ \hat{\theta}_i]^T, \ i = 0, \dots, k \qquad (43)$$

can be obtained, where $(\hat{x}_k^c, \hat{y}_k^c) = (\mu_{x,k}, \mu_{y,k})$.

4.3 The Relationship with the Original Tracker

The following probability of object roughly evaluates the reliability of the estimated tracked region under the Gaussian assumption.

$$
\begin{aligned}
P(\|\tilde{z}_k\| < \hat{l}_k) \\
= \int_{D_{\hat{l}_k}} p_{z_k,out}(z_k)dz_k \\
= \frac{1}{(2\pi)\lambda_{k,max}} \int_0^{2\pi} \int_0^{\hat{l}_k} e^{-\frac{1}{2}\lambda_{k,max}^{-1}r^2} r dr d\theta \\
= 1 - e^{-\frac{1}{2}\lambda_{k,max}^{-1}\hat{l}_k^2}.
\end{aligned}
\qquad (44)
$$

Recalling $\hat{l}_k = \sqrt{\rho\lambda_{k,max}}$, we have

$$P(\|\tilde{z}_k\| < \hat{l}_k) = 1 - e^{-\frac{\rho}{2}} \qquad (45)$$

which is independent of the eigenvalues. When ρ is set to 2, the proposed tracker reduces to the original one. In this time, the probability of (45) is about 63%.

5. Extension of the Tracker to a Multidimensional Case

5.1 The Mechanical Approach

We consider the tracking of a three dimensional object as shown in Fig. 5(a). In this case, we face the following two serious problems.

1. The moment along with x axis depends on the width w and thickness t of the object whose information is mixed in the resultant moment as shown in Figs. 5(b) and (c).

2. The length l of the object cannot be uniquely determined, because a rotation axis to estimate the length l is not unique as illustrated in Figs. 5(d) and (e).

5.2 The Statistical Approach

Our idea is to regard the tracked object as a section (an ellipsoid) to cross the following probability density function with a hyper plane:

$$p_{z_k}(z_k) = \frac{1}{(2\pi)|\Sigma_{z_k}|^{\frac{1}{2}}} e^{-\frac{1}{2}(z_k-\bar{z}_k)^T\Sigma_{z_k}^{-1}(z_k-\bar{z}_k)}. \qquad (46)$$
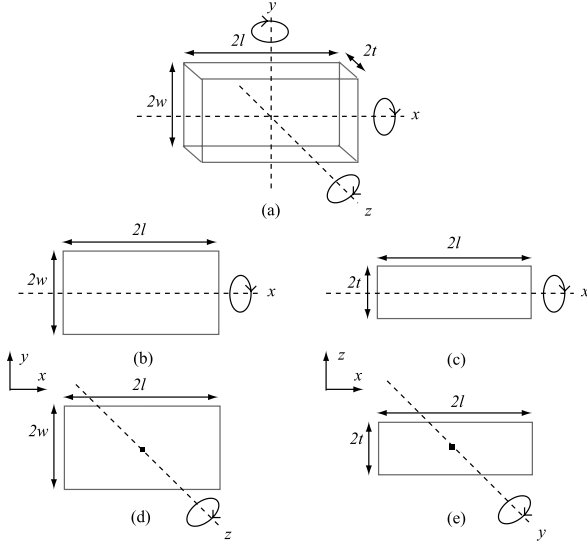
Fig. 5 Tracked object with $2w \times 2l \times 2t$ size, whose rotation axes are set to x, y, and z axes, respectively.

The integration $\int_V z_k p_{z_k}(z_k) dz_k$ means the probability that pixels of the object appear in the region $V \subset \mathcal{R}^3$. Here the mean and covariance of $z_k \in \mathcal{R}^3$ are given by

$$\bar{z}_k = \begin{bmatrix} \mu_{x,k} \\ \mu_{y,k} \\ \mu_{z,k} \end{bmatrix}, \Sigma_{z_k} = \begin{bmatrix} \sigma_{x,k}^2 & \gamma_k^{110} & \gamma_k^{101} \\ \gamma_k^{110} & \sigma_{y,k}^2 & \gamma_k^{011} \\ \gamma_k^{101} & \gamma_k^{011} & \sigma_{z,k}^2 \end{bmatrix} \tag{47}$$

whose elements are defined with the pixel's probability as

$$\mu_{x,k} = \sum_{\check{\tau}=(\check{x},\check{y},\check{z})\in\check{O}_k} \check{x} P_k(\check{\tau}), \quad \mu_{y,k} = \sum_{\check{\tau}=(\check{x},\check{y},\check{z})\in\check{O}_k} \check{y} P_k(\check{\tau})$$

$$\mu_{z,k} = \sum_{\check{\tau}=(\check{x},\check{y},\check{z})\in\check{O}_k} \check{z} P_k(\check{\tau}) \tag{48}$$

$$\sigma_{x,k}^2 = \sum_{\check{\tau}=(\check{x},\check{y},\check{z})\in\check{O}_k} (\check{x} - \mu_{x,k})^2 P_k(\check{\tau}) \tag{49}$$

$$\sigma_{y,k}^2 = \sum_{\check{\tau}=(\check{x},\check{y},\check{z})\in\check{O}_k} (\check{y} - \mu_{y,k})^2 P_k(\check{\tau}) \tag{50}$$

$$\sigma_{z,k}^2 = \sum_{\check{\tau}=(\check{x},\check{y},\check{z})\in\check{O}_k} (\check{z} - \mu_{z,k})^2 P_k(\check{\tau}) \tag{51}$$

$$\gamma_k^{011} = \sum_{\check{\tau}=(\check{x},\check{y},\check{z})\in\check{O}_k} (\check{y} - \mu_{y,k})(\check{z} - \mu_{z,k}) P_k(\check{\tau}) \tag{52}$$

$$\gamma_k^{101} = \sum_{\check{\tau}=(\check{x},\check{y},\check{z})\in\check{O}_k} (\check{x} - \mu_{x,k})(\check{z} - \mu_{z,k}) P_k(\check{\tau}) \tag{53}$$

$$\gamma_k^{110} = \sum_{\check{\tau}=(\check{x},\check{y},\check{z})\in\check{O}_k} (\check{x} - \mu_{x,k})(\check{y} - \mu_{y,k}) P_k(\check{\tau}). \tag{54}$$

The first, second, and third principal axes of the density function are obtained by solving the eigenvalue problem of Σ_{z_k}. Supposing that λ_k is the eigenvalues of Σ_{z_k}, we have the characteristic equation:

$$|\Sigma_{z_k} - \lambda_k I| = \begin{vmatrix} \sigma_{x,k}^2 - \lambda_k & \gamma_k^{110} & \gamma_k^{101} \\ \gamma_k^{110} & \sigma_{y,k}^2 - \lambda_k & \gamma_k^{011} \\ \gamma_k^{101} & \gamma_k^{011} & \gamma_{z,k}^2 - \lambda_k \end{vmatrix}$$

$$= 0. \tag{55}$$

According to Appendix A, we can determine the three eigenvalues as

$$\lambda_k \in \{\lambda_{1,k}, \lambda_{2,k}, \lambda_{3,k}\}, \quad \lambda_{1,k} \geq \lambda_{2,k} \geq \lambda_{3,k}. \tag{56}$$

The corresponding eigenvectors $v_{1,k}$, $v_{2,k}$, and $v_{3,k}$, which are considered as three orthogonal principal axes and ordered based on their eigenvalues, determine a new coordinate system in the conception of the PCA. Here each eigenvector is defined as

$$v_{i,k} = [v_{x,i,k} \ v_{y,i,k} \ v_{z,i,k}]^T, \quad i = 1, 2, 3. \tag{57}$$

Then the tracked object can be regarded as an ellipsoid whose axes are along the eigenvectors $v_{1,k}$, $v_{2,k}$, and $v_{3,k}$, respectively. Given the projection of $v_{1,k}$ onto the plane x-y as $\tilde{v}_{1,k}$, the angle between the x axis and $\tilde{v}_{1,k}$, and the angle between $v_{1,k}$ and $\tilde{v}_{1,k}$ are given by

$$\hat{\theta}_k = \arctan \frac{v_{y,1,k}}{v_{x,1,k}}, \quad \hat{\phi}_k = \arcsin \frac{v_{z,1,k}}{\|v_{1,k}\|}. \tag{58}$$

Both of them direct $v_{1,k}$ in the three-dimensional space $\mathcal{R}^3$. Additionally, we can obtain the length, width, and thickness of the tracked object as

$$\hat{l}_k = \sqrt{\rho \lambda_{k,1}}, \quad \hat{w}_k = \sqrt{\rho \lambda_{k,2}}, \quad \hat{t}_k = \sqrt{\rho \lambda_{k,3}} \tag{59}$$

where ρ is a scalar parameter to adjust the probability $P(\|\bar{z}_k\| < \hat{l}_k)$.

Consequently, the tracked object is obtained as an ellipsoid, which is centered at $\hat{\tau}_k^c = (\mu_{x,k}, \mu_{y,k}, \mu_{z,k})$ with $2\hat{l}_k \times 2\hat{w}_k \times 2\hat{t}_k$ size. The direction of the tracked object is along the semimajor axis of the ellipsoid which is determined by $\hat{\theta}_k$ and $\hat{\phi}_k$. As the tracking process is implemented repeatedly from the first tracked frame to the current tracked frame, a sequence of the tracking results defined by

$$\hat{s}_i = [\hat{x}_i^c \ \hat{y}_i^c \ \hat{z}_i^c \ \hat{l}_i \ \hat{w}_i \ \hat{t}_i \ \hat{\theta}_i \ \hat{\phi}_i]^T, \quad i = 0, \ldots, k \tag{60}$$

can be obtained, where $(\hat{x}_k^c, \hat{y}_k^c, \hat{z}_k^c) = (\mu_{x,k}, \mu_{y,k}, \mu_{z,k})$.

For checking the accuracy of the method, suppose that a object to be tracked is a $16 \times 8 \times 4$ cuboid which is centered at $\tau^c = (0, 0, 0)$ and rotated $45°$ around z axis. Each voxel in the cuboid randomly take a value in the range of 0 to 255. From (47) to (54), we can evaluate the mean $\bar{z}$ and covariance Σ_z. Hence, using the eigenvalues and eigenvectors of Σ_z, we obtain

$$\hat{\theta} = 43.8°, \quad \hat{\phi} = 0.200°$$
$$\hat{l} = 7.79, \quad \hat{w} = 4.08, \quad \hat{t} = 2.23$$

by (58) and (59). Here ρ is set to 2.5.

6. Conclusion

We have extended the continuously adaptive mean shift tracker to a more general one, using a view that a horizontal section of Gaussian probability density function for pixels can represent the tracked object. This approach allows the direction and size of object to be easily derived by the PCA. Additionally, it is possible to propose a new color-based visual tracker for multidimensional objects, providing new applications of the visual tracker. For instance, tracking the growing of a cancer of interest from a three-dimensional medical image sequence obtained by CT or MRI could help doctors to diagnose.

References

[1] G.R. Bradsky, "Computer vision face tracking for use in a perceptual user interface," Intel Technology Journal, Quarter 2, pp.1–15, 1998.

[2] K. Fukunaga and L.D. Hostetler, "The estimation of the gradient of a density function, with applications in pattern recognition," IEEE Trans. Inf. Theory, vol.21, no.1, pp.32–40, 1975.

[3] Y. Cheng, "Mean shift, mode seeking, and clustering," IEEE Trans. Pattern Anal. Mach. Intell., vol.17, no.8, pp.790–799, 1995.

[4] B.K. P. Horn, Robot vision, The MIT Press, 1986.

[5] R. Gonzalez and R.E. Woods, Digital Image Processing, 2nd ed., Prentice Hall Press, 2002.

Appendix: The Eigenvalues of 3×3 Covariance Σ

The eigenvalues λ of 3×3 covariance Σ:

$$\Sigma = \begin{bmatrix} a & d & f \\ d & b & e \\ f & e & c \end{bmatrix} \tag{A·1}$$

can be calculates as

$$\lambda_1 = \frac{p^{1/3}}{6} - \frac{6q}{p^{1/3}} + \frac{a+b+c}{3} \tag{A·2}$$

$$\lambda_2 = -\frac{p^{1/3}}{12} + \frac{3q}{p^{1/3}} + \frac{a+b+c}{3}$$
$$+ \frac{\sqrt{3}}{2}\left(\frac{p^{1/3}}{6} + \frac{6q}{p^{1/3}}\right)i \tag{A·3}$$

$$\lambda_3 = -\frac{p^{1/3}}{12} + \frac{3q}{p^{1/3}} + \frac{a+b+c}{3}$$
$$- \frac{\sqrt{3}}{2}\left(\frac{p^{1/3}}{6} + \frac{6q}{p^{1/3}}\right)i \tag{A·4}$$

where

$$p = 216\,d\,f\,e + 48\,a\,b\,c - 72\,a\,e^2 - 72\,d^2\,c - 72\,f^2\,b$$
$$-12\,a^2\,c + 36\,a\,d^2 + 36\,a\,f^2 - 12\,a^2\,b - 12\,a\,b^2$$
$$-12\,a\,c^2 - 12\,b^2\,c - 12\,b\,c^2 + 36\,e^2\,b + 36\,e^2\,c$$

$$+36\,d^2\,b + 36\,f^2\,c + 8\,b^3 + 8\,c^3 + 8\,a^3$$
$$+12(-6\,a^2\,b^2\,f^2 - 6\,a^3\,b\,c^2 - 24\,a^2\,b^2\,d^2$$
$$-6\,a^2\,b^2\,e^2 - 24\,a\,b\,e^4 + 30\,a\,b\,d^4 - 60\,a^2\,b\,e^2\,c$$
$$+30\,a\,b^2\,e^2\,c + 6\,a\,b\,d^2\,e^2 + 114\,a\,b\,f^2\,e^2$$
$$+30\,a^2\,b\,c\,d^2 + 30\,a^2\,b\,c\,f^2 + 30\,a\,b^2\,c\,d^2$$
$$-60\,a\,b^2\,c\,f^2 + 30\,a\,b\,e^2\,c^2 + 6\,a\,b\,d^2\,f^2$$
$$-60\,a\,b\,d^2\,c^2 + 30\,a\,b\,f^2\,c^2 + 114\,a\,c\,d^2\,e^2$$
$$+6\,a\,c\,f^2\,e^2 + 6\,a\,c\,d^2\,f^2 + 6\,b\,c\,d^2\,e^2$$
$$+6\,b\,c\,f^2\,e^2 + 114\,b\,c\,d^2\,f^2 + 6\,a^4\,b\,c - 6\,a^2\,b\,c^3$$
$$-6\,a\,b^3\,c^2 + 6\,a\,b^4\,c - 6\,a\,b^2\,c^3 + 24\,a^3\,b\,e^2$$
$$-6\,a\,b^3\,e^2 + 6\,a^3\,b\,d^2 + 6\,a\,b^3\,d^2 - 6\,a^3\,b\,f^2$$
$$+24\,a\,b^3\,f^2 - 24\,a\,b\,f^4 + 6\,a\,b\,c^4 - 6\,a^2\,c^2\,e^2$$
$$-6\,a^2\,c^2\,d^2 - 24\,a^2\,c^2\,f^2 - 24\,a\,c\,e^4 - 24\,a\,c\,d^4$$
$$-6\,a^3\,c\,b^2 - 6\,a^2\,c\,b^3 + 24\,a^3\,c\,e^2 - 6\,a\,c^3\,e^2$$
$$-6\,a^3\,c\,d^2 + 24\,a\,c^3\,d^2 + 6\,a^3\,c\,f^2 + 6\,a\,c^3\,f^2$$
$$+30\,a\,c\,f^4 - 24\,b^2\,c^2\,e^2 - 6\,b^2\,c^2\,d^2 - 6\,b^2\,c^2\,f^2$$
$$+30\,b\,c\,e^4 - 24\,b\,c\,d^4 - 3\,a^4\,b^2 - 3\,a^2\,b^4$$
$$+6\,a^3\,c^3 - 3\,a^4\,c^2 - 3\,a^2\,c^4 + 6\,b^3\,c^3 - 3\,b^4\,c^2$$
$$-3\,b^2\,c^4 - 36\,e^4\,d^2 - 36\,e^4\,f^2 - 36\,e^2\,d^4$$
$$+24\,e^4\,a^2 - 3\,e^4\,b^2 - 3\,e^4\,c^2 - 36\,e^2\,f^4$$
$$-12\,e^2\,a^4 - 36\,d^4\,f^2 - 3\,d^4\,a^2 - 3\,d^4\,b^2$$
$$+24\,d^4\,c^2 - 36\,d^2\,f^4 - 12\,d^2\,c^4 + 6\,a^3\,b^3$$
$$+144\,d\,f\,e\,a\,b\,c - 36\,d\,f\,e\,a^2\,c - 36\,d\,f\,e\,a^2\,b$$
$$-36\,d\,f\,e\,a\,b^2 + 6\,b^3\,c\,e^2 + 6\,b\,c^3\,e^2 - 6\,b^3\,c\,d^2$$
$$+24\,b\,c^3\,d^2 + 24\,b^3\,c\,f^2 - 6\,b\,c^3\,f^2 - 24\,b\,c\,f^4$$
$$+252\,e^2\,d^2\,f^2 - 60\,e^2\,d^2\,a^2 - 6\,e^2\,d^2\,b^2$$
$$-60\,e^2\,d^2\,c^2 - 60\,e^2\,f^2\,a^2 - 60\,e^2\,f^2\,b^2$$
$$-6\,e^2\,f^2\,c^2 - 6\,d^2\,f^2\,a^2 - 60\,d^2\,f^2\,b^2 - 60\,d^2\,f^2\,c^2$$
$$+18\,a^2\,b^2\,c^2 - 12\,e^6 - 12\,d^6 - 216\,d\,f\,e^3\,a$$
$$-216\,d^3\,f\,e\,c - 216\,d\,f^3\,e\,b + 108\,d^3\,f\,e\,a$$
$$+108\,d\,f^3\,e\,a + 108\,d\,f\,e^3\,b + 108\,d\,f\,e^3\,c$$
$$+108\,d^3\,f\,e\,b + 108\,d\,f^3\,e\,c + 24\,d\,f\,e\,b^3$$
$$+24\,d\,f\,e\,c^3 + 24\,d\,f\,e\,a^3 - 3\,f^4\,a^2 + 24\,f^4\,b^2$$
$$-3\,f^4\,c^2 - 12\,f^2\,b^4 - 12\,f^6 - 36\,d\,f\,e\,a\,c^2$$
$$-36\,d\,f\,e\,b^2\,c - 36\,d\,f\,e\,b\,c^2)^{1/2} \tag{A·5}$$

and

$$q = \frac{a\,b + a\,c + b\,c}{9} - \frac{e^2 + d^2 + f^2}{3}$$
$$- \frac{a^2 + b^2 + c^2}{9}. \tag{A·6}$$

These equations are deduced from the software MAPLE. The eigenvectors can be calculated based on the eigenvalues by (A·2) to (A·4).