Research on the Application of Machine Vision Technology in the Logistics Center Monitoring Platform System

1Shusun Zhang, 2Xing Wang and 1Kai Kang
1School of Business, Northeast Normal University, ChangChun, 130117, China
2College of Quartermaster Technology, Jilin University, ChangChun, 130061, China

Abstract: Object monitoring is one of the key technologies of logistics information platform system. The study introduced an application method of machine vision technology in the logistics center monitoring system, improved covariance matrix algorithm to monitor objects in the logistics center. Against the technical difficulties of objects detection, the covariance matrix algorithm was applied to monitor objects in the logistics center and against the shortcomings of covariance matrix algorithm in the process of monitoring objects, the study proposed a method of path prediction and template dynamic adjustment. Experiments show that the method can effectively monitor objects in the logistics center, the improved method can not only adapt quickly to pose and scale variations of objects, but also track accurately and continuously those temporarily occluded objects, has good robustness. The method provides a new solution of monitoring objects in the logistics system.

Key words: Logistics information platform, machine vision technology, tracking algorithm, logistics center

INTRODUCTION

With the development and application of information technology, the logistics information system has reached the integrated application stage of data mining, artificial intelligence, decision theory. The real-time detecting and tracking of moving objects is one of the key technologies of logistics information platform system. In the logistics system, the moving objects include freight vehicles, forklifts, pallets, moving goods on conveyor belt and workers. The application of machine vision technology to track moving objects is significant. At present, the technologies includes bar code technology, RFID, GPS/ GIS and so on (Lee et al., 2011; Zheng et al., 2010; Imran et al., 2006). Various technologies have their advantages and disadvantages for different application occasions. For example, for those freight vehicles not being recorded in the monitoring system, the above identification methods can not automatically identify them. In addition, the above methods also cannot visualize the real-time situation of moving objects and are unable to provide visual reference information for logistics decision-making. Moving objects have some features, such as a stable shape and being easy to be detected by machine vision technology but usually being in complex crowded scenes with occlusions. To overcome these shortcomings of others recognition technologies, the study introduced the machine vision into the area for tracking logistics objects.

Tracking an object in a sequence of images is currently utilized in many machine vision applications. The goal of visual tracking is to locate a region in each image that matches an appearance of a target object. The algorithm of visual object tracking can be divided broadly into two categories (Austvoll and Kwolek, 2010) feature-based and template-based. In typical scenarios, interactions between moving objects result in partial or significant occlusions, making the object tracking a highly challenging problem. Various systems and methods have been proposed to handle object tracking in complex crowded scenes with the occlusions.

The object tracking is often achieved using a single camera. However, one fundamental limitation of using one camera in the tracking of objects is dealing with object occlusions. In single-camera methods, occlusion can be identified through prediction of the object location or on a per-pixel basis. In many situations the existing algorithms do not have satisfactory tracking robustness, especially when there is a large amount of occlusion between two or more objects. Therefore, a highly efficient occlusion handling scheme through constructing a region covariance matrix is needed. We proposed the covariance matrix algorithm and its improvement method of tracking of moving objects using machine vision technology.

However, the exiting models are highly sensitive to the pose, scale and shape variations. To overcome the shortcomings of the exiting approaches, we improved the covariance matrix algorithm and applied it to detection and
tracking of moving objects. In the next section, we explain how we construct the covariance matrices, compute the distances. In section 3, we give the improved method. In section 4, we give several examples under different conditions and environment.

**COVARIANCE MATRIX ALGORITHM PRINCIPLE**

Covariance matrix method is widely researched and applied to all kinds of areas (Salmen et al., 2010; Ahmed et al., 2011). Tuzel et al. (2006) proposed covariance descriptors and applied it to object detection and texture classification problems. In this subsection, we first review the covariance matrix method.

**Covariance matrix:** Covariance matrix describes the object area in an image. As the object feature descriptor, covariance matrix is the most necessary features for designing the algorithm of object detecting and tracking. Denoting I is one dimensional intensity image or three dimensional color image, let F be the \(W \times H \times d\) dimensional feature image extracted from I, as in 1:

\[
F(x, y) = V(I(x, y))
\]

(1)

where, the function \(V\) may be any mapping such as color, edge magnitude, edge orientation, filter responses, etc. \(x, y\) are the coordinates of pixels. For a given rectangular window, \(R=F\), let \(\{A_k\}_{k=1,2,\ldots,n}\) be the \(d\)-dimensional feature vectors inside \(R\). We construct the feature vector \(A_k\) using two types of mappings, where they are spatial attributes and appearance attributes. These features may be associated directly to the pixel coordinates, as in 2:

\[
Z_k = [x, y, I(x, y), I_k(x, y), \ldots]
\]

(2)

Alternatively, they can be arranged in radially symmetric relationship, as in 3:

\[
Z_k^\prime = \left[\left|x, y\right|, I(x, y), I(x, y), \ldots\right]
\]

(3)

Where:

\[
\left|x, y\right| = \sqrt{x^2 + y^2}, \quad (x, y) = (x-x_0, y-y_0)
\]

They are the relative coordinates and \((x_0, y_0)\) is the coordinates of the window center.

Different associations of the spatial information to the image features enable imposing of separate blinding rules. For instance, \(Z_k\) prevails an appearance model susceptible to the object rotation with respect to window origin \((x_0, y_0)\), whereas \(Z_k^\prime\) offers invariant spatial formation of the features.

So, we represent a \(M \times N\) rectangular region \(R\) with a \(d \times d\) covariance matrix \(C_R\) of the feature point as:

\[
C_R = \frac{1}{MN} \sum_{k=1}^{MN} (f_k - \mu_A)(f_k - \mu_A)^T
\]

(4)

where, \(MN\) is the number of pixels, \(f_k\) is a property vector of the pixel \(P_{i,j}\) at location \((i, j)\), \(C_R\) is the vector of the means of the corresponding features for the points within the region \(R\).

The vector \(f_k\) can include locations of pixels and intensity values of pixels, or it can only include the red, green and blue values of pixels and their first and second derivatives. Several pixel property parameters which are defined in above Equations can be included in the \(f_k\) vector. By using combinations of these properties, different covariance descriptors for a given image region can be defined.

Covariance matrix can be invariant to rotations. Nevertheless, if information regarding the orientation of the points is embedded within the feature vector, it is possible to detect rotational discrepancies. It should be pointed that the covariance is invariant to the mean changes such as identical shifting of color values. This becomes an advantageous property when objects are tracked under varying illumination conditions.

Based on integral image representation, covariance matrix from feature can be computed. After constructing tensors of integral images corresponding to each feature dimension and multiplication of any two feature dimensions, the covariance matrix of any arbitrary rectangular region can be computed independent of the region size.

Finding the object based on the distance: The distance matrix (Forikli et al., 2006; Forstner and Moonen, 1999) uses the sum of the squared logarithms of the generalized eigenvalues to compute the dissimilarity of two covariance matrices as:

\[
L(C_O, C_T) = \sum_{i=1}^{d} \ln^2 \lambda_k(C_O, C_T)
\]

(5)

where, \(\{\lambda_k(C_O, C_T)\}\) are the generalized eigenvalues of \(C_O\) and \(C_T\), computed from \(X_k^O C_O X_k - C_T = 0\) \(k=1, \ldots, d\).

And \(X_k\) are the generalized eigenvectors. The distance measure \(L(C_O, C_T)\) satisfies the metric axioms for positive definite symmetric matrices \(C_O\) and \(C_T\). Definitions satisfy the following three properties:

- Non-negative $L(C_0, C_T) \geq 0$ and $L(C_0, C_T) = 0$ only if $C_0 = C_T$
- Symmetry $L(C_0, C_T) = L(C_T, C_0)$
- Triangle inequality

$$L(C_0, C_T) + L(C_0 + C_T) \geq L(C_T + C_1)$$

At each frame, we search the whole image to find the region which has the smallest distance from the current object model. The best matching region determines the location of the object in the current frame. According to those above properties, we can see, the more similar the features of the two objects are, the smaller the distance of the region covariance matrix is, on the contrary, the greater the distance is.

**Updating the model to adapt to variations:** For moving objects undergoing shape, size and appearance transformations over time, it is necessary to adapt to these variations. We can construct and update a temporal kernel of covariance matrices corresponding to the previously estimated object regions $R_1, ..., R_N$, we keep a set of $T$ previous covariance matrices $[C_1, ..., C_T]$ where $C_t$ denotes the current covariance matrix. From this set, we compute a sample mean covariance matrix that blends all the previous matrices. This work is computationally expensive and requires a large amount of memory to store all the previous observations. It is desirable to obtain an aggregated covariance matrix without being limited to a constant window size and keeping all the previous measurements. However, covariance matrices do not conform to Euclidean geometry. We can still compute the mean of several covariance matrices through Lie groups (Rossman, 2002; Bazzani and Castellani, 2010).

We need to focus on matrix Lie groups. The exponential map of a matrix and its inverse, log, is defined by:

$$\exp(\alpha) = \sum_{n=0}^{\infty} \frac{1}{n!} \alpha^n \log(A) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} (A - e)^n$$

We adapt the similar idea to obtain the intrinsic mean of covariance matrices. Let $C$ be a point on the Lie algebra and $\hat{C} = \exp(c)$ be its mapping to the Lie group. The distances on manifolds are defined in terms of minimum length curves between points on the manifold. The curve with the minimum length is called the geodesic and the length of the curve is the intrinsic distance. The intrinsic distance of point $C$ to the identity element of the group is given by $\|\log(C)\|$. Left multiplication by the inverse of a group element $C^{-1}$ maps the point $C$ to the identity, and tangent space at $C$ to the Lie algebra. This mapping is an isomorphism. Given $[C_t]_{t=1}^T$ as the data points on the group, taking the log of the above mapping:

$$c_t = \log(C_t^{-1}C_1)$$

The data points are mapped to the Lie algebra and $C$ to $0$. For Lie algebra, it is a vector space, we can compute a first order approximation to the intrinsic mean of the points. Starting at an initial matrix $C_t$ and iteratively computing first order approximations to the intrinsic mean, we converge to a fixed point on the group. The algorithm flow process is summarized as follows:

- Initialize $\hat{C} = C_1$
- Repeat
  - For $t = 1$ to $T$
  - Compute:

$$c_t = \log(C_t^{-1}C_1)$$

- Compute:

$$\Delta C = \exp\left(\frac{1}{T} \sum_{t=1}^{T} c_t\right)$$

- Assign $\hat{C} = \hat{C} \Delta C$
- Until $\|\log(C)\| < \varepsilon$

The error at each iterations of the algorithm can be expressed, the mapping ensures that error is minimized. At the end of the iterations, we find the intrinsic mean and use $\hat{C}$ as the current model.

In the above formulations, we considered all the previous matrices $C_1, ..., C_T$ in the set as equally influential on the result regardless of whether they are accurate or not. To prevent the model from contamination, it is possible to weight the data points proportional to its similarity to the current model. Then, the computation step on the above algorithm becomes:

$$\Delta C = \exp\left(\frac{1}{L^*} \sum_{t=1}^{T} L_t^{-1}(C_t, C^*) s_t\right)$$

where, $L$ is defined in 8:

$$L^* = \sum_{t=1}^{T} L_t^{-1}(C_t, C^*)$$

and $C^*$ is the model $\hat{C}$ computed at the previous frame.

**COVARIANCE MATRIX IMPROVEMENT**

To find the most similar region for the given object, covariance matrix algorithm requires a full scan of all
regions to compute the shortest distances between the
covariance matrices corresponding to the object window
and the candidate regions and the corresponding region
is the object we are looking for. But when pose and size of
the moving object varies greatly, or the object is
temporarily occluded, it will be misjudged, if not
improved, the found object won’t be the one that we are
looking for. Therefore, according to certain principles,
when the detected angle of the object varies, we need to
dynamically adjust the template image to accurately track
it. In addition, for illumination, machine vision sensor
noise effect and other reasons, the present covariance
matrix algorithm will wrongly identify the singular value
region as object region. To solve this question, we
introduced the path prediction algorithm to predict the
object moving path, we can solve the path drift issue and
continuously track the moving objects.

Selection and dynamic adjustment of the template image:
The template image is one of the key technologies of
stably tracking objects (Marsigilia et al., 2003). The
template image should be updated adaptively in respect
that the object observed angle, distance, machine vision
sensor noise and background will vary in the process of
object tracking.

The study introduced a threshold value and useful
template image updating strategy. The mean difference
can be used as the threshold parameters of a template
image update. The best matching region for the current
given image and the template image is defined as \( f^*(x, y) \)
with \( m \times n \) size, the mean difference of the matching image
region and the template image \( g(x, y) \) uses the following
formula to compute:

\[
D(i, j) = \frac{1}{nm} \sum_{x=1}^{m} \sum_{y=1}^{n} | f^*(x+i, y+j) - g(x, y) |
\]  

(9)

where, \( i = 0, 1, ..., M - m + 1, j = 0, 1, ..., N - n + 1 \)

While \( D(i, j) \) is less than an experience threshold, the
template image mustn’t be updated, otherwise, the
template image must be updated according to the above
method.

The prediction algorithm of moving objects tracking
path: When a frame is lost and the moving objects are
occluded and it cannot be tracked, we need to predict the
object position to prevent its position from drift. In this
study, a linear logistic regression method was used to
predict the object path as a reference of the logistics
object next position.

Where \( X(t) \) is position function of the logistics
object, \( X(t_i)(i = 1, 2, ..., n) \) is the position coordinate in
i-point sequences, using:

\[
X(t_i) = a + bt_i - \left[ \begin{array}{c} q_i \\ b_i \\ \end{array} \right]
\]
as its best linear approximation value. The error between
the real position and approximation value of the logistics
object can be computed by \( \Delta q_i = X(t_i) - a - bt_i \). The sum
of squared error estimation of \( n \) points are defined as:

\[
E(\Delta q_i^2) = \sum_{i=1}^{n} [X(t_i) - a - bt_i]^2
\]

if the optimal solution of:

\[
\min \left( E(\Delta q_i^2) \right)
\]
is denoted by \( (a^*, b^*) \) then:

\[
\begin{bmatrix} a^* \\ b^* \end{bmatrix} = \arg \min \left( E(\Delta q_i^2) \right)
\]

can be obtained by the least squares algorithms as:

\[
\begin{bmatrix} a^* \\ b^* \end{bmatrix} = \left[ \frac{\sum_{i=1}^{n} X(t_i) - a \sum_{i=1}^{n} t_i}{\sum_{i=1}^{n} t_i^2} - b \right] \left( \frac{\sum_{i=1}^{n} t_i^2}{\sum_{i=1}^{n} t_i} \right) \left( \frac{n}{n} - \frac{1}{n} \sum_{i=1}^{n} t_i \right)
\]

(10)

where, \( n \) is the number of frames used to predict and:

\[
\bar{X}(t_i) = \frac{1}{n} \sum_{i=1}^{n} X(t_i), \quad \bar{t}_i = \frac{1}{n} \sum_{i=1}^{n} t_i
\]

(11)

In the process of the tracking position prediction, \( n \)
should be selected based on specific circumstances. In
this study, \( n = 3 \), when the current frame is lost, the object
is occluded or its position is obviously unreasonable, the
current frame position may be predicted based on the
previous three frames position, the position of the frame
t+1 can be predicted by the frame t-2, t-1 and the frame t.

EXPERIMENTAL AND PERFORMANCE ANALYSIS

We assessed the performance of the method based
on machine vision sensor. The study implemented the
that the improved covariance model and template update mechanism can successfully detect and adapt models to various changes, several sample tracking results are given in Fig. 1 and 2. Figure 1 are several camera sequences in a logistics center, there are two tracked objects, one of the tracked objects is very fast, in this case, the objects can still be accurately tracked.

Figure 2 are several camera sequences in a logistics center of a large household electrical appliance enterprises group in China under normal daylight condition, the detected and tracked moving objects are the forklifts. The moving objects can accurately be tracked.

CONCLUSIONS

Through the application of machine vision technology, the study successfully applied it to the field of tracking moving objects in the logistics system. The template image selection has a great influence on the effectiveness and efficiency of detecting and tracking objects, since observed angle of moving objects will vary over time, the template image update strategy can implement that the template image automatically updates and is well adapt to the tracking requirements of moving objects. By the combination of covariance matrix algorithm and tracking path prediction algorithm, the improved covariance matrix method can track moving objects continuously and accurately, experiments show that the performance of the algorithm is robust. The research results provides the technical foundation to build the logistics information platform system.

ACKNOWLEDGMENTS

The Project was supported by the Fundamental Research Funds for the Central Universities (NENU-12QN048, JLU-45006048106), we would like to thank editorial team for support.

REFERENCES


