Optik 126 (2015) 2500–2507

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

A^* algorithm with dynamic weights for multiple object tracking in video sequence

Zhenghao Xi [a,*], Dongmei Xu [a], Wanqing Song [a], Yang Zheng [b]

[a] School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China
[b] School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China

ARTICLE INFO

Article history:
Received 28 April 2014
Accepted 5 June 2015

Keywords:
Multi-object tracking
A^* algorithm
Flow network model
Dynamic constraint

ABSTRACT

In the clutter environments, persistently tracking multiple objects is still a very challenging problem. In this paper, an A^* algorithm with dynamic weights is developed to solve this problem. Firstly, the problem of multiple object tracking is considered as an integer programming of flow network model. Based on this model, afterward, we relax the problem of integer programming problem to a standard linear programming and achieve the global optimal solution with the proposed algorithm, which has been proven in this work. Compared with other advanced methods employed in the modern complex environments, the proposed method is on the merit of lower calculation complexity and better tracking accuracy and robustness. Finally, the simulation result is demonstrated, which reveals that the proposed algorithm saves the time costs vastly.

Multiple object tracking is a hot issue in the field of computer vision, robust tracking of objects is important for many computer vision applications, such as human-computer interaction, video surveillance, intelligent navigation and other aspects [1,2]. Apart from the detection algorithm of high performance as an auxiliary, multi-object tracking of high quality should also track the algorithm for support, which can address certain types of complex cases, e.g., illumination, occlusion, clutter, and so on [3]. The data association (DA) method is a favorite method of multi-object tracking. The often utilized techniques include the nearest neighbor method [4], joint probability data association (JPDA) [5] and the methods based on neural networks [6] etc.

The effect of the above DA methods is closely related to the detection accuracy of the detector in the adjacent frames. These typical approaches are resilient to false positives and false negatives: if an object is not detected in a frame but is detected in previous and following frames, it is a false negative. A false positive is mistaking the tracking object 'A' as object 'B'. Although this problem can be solved using targeted design a statistical trajectory model with filtering [7,8], the estimating method exhibiting maximum posterior probability is NP-Complete.

Many recent papers proposed some approaches for this problem: Giebel et al. [9] used sampling and particle filtering to remove clutter from the same object and reduced the probability of NP-Complete. This method can obtain relatively accurate tracking trajectory but requires a sufficient sampling point. Perera et al. [10] divided a long sequence into several short ones, yielding lots of short tracking tracks, and linked them using Kalman filtering. This can avoid the NP-Complete. The accuracy of this method is inversely proportional to the length of the short one, the short track and the better tracking, but the excessive division will increase the computation time and cannot track objects for a long time. Fleuret et al. [11] processed trajectories individually over long sequences using a reasonable greedy dynamic programming (DP) to choose the order. These approaches, while effective, cannot achieve the global optimum.

Zhang's approach [12] relies on a min-cost network flow framework based optimization method to find the global optimum for multiple object tracking, but the proposed two algorithms have many defects in practice and the complexity of the algorithms is polynomial. Under this framework, Berclaz et al. [13] formulated multi-object tracking as an Integer Programming (IP) problem and reduced it to linear programming (LP). By relying on the k-shortest paths (KSP) algorithm for the optimization of the LP problem, their approach reduced the complexity to perform robust multi-object tracking in time. However, because of KSP's lack of a motion model over DP, DP's tendency to ignore fragmentary trajectories makes it more robust. Pirsiavash [14] continues the work of Zhang, his method was used to obtain the global optimal solution with the greedy algorithm for $K=1$ in $O(N)$ but only obtained the approximate solutions for $K>1$ in $O(KN)$, where K is the unknown optimal number of unique tracks.

* Corresponding author. Tel.: +86 15652932858.
E-mail address: zhenghaoxi@hotmail.com (Z. Xi).

http://dx.doi.org/10.1016/j.ijleo.2015.06.020
0030-4026/

By contrast, we effectively combine the model of Zhang and Berclaz, a more efficient A^* association algorithm with dynamic weights (A^*AADW) was developed to solve the multi-object tracking problem on this basis. The A^*AADW algorithm can directly obtain the global solution without greedy optimization, it is far better with respect to both the worst case complexity and solving time than the above state-of-the-art algorithm. The main contributions are listed as follows:

(1) A general mathematical integer programming formulation of a min-cost network flow framework is introduced for multi-object tracking, which more conveniently and naturally filters out false positives and false negatives using A^*AADW.
(2) To solve the integer programming formulation of the proposed framework and to obtain the global optimum, we propose a novel more rapid and more efficient A^*AADW algorithm, which is very robust as well.
(3) Extensive experimental validations.

The rest of this paper is organized as follows. In Section 1, we formulate an IP problem using the min-cost network flow framework and relax it to a continuous LP. Section 2 presents the proposed A^* association algorithm with dynamic weights for the relaxation of the original integer assumption. Section 3 is about the approaches of object localization and long sequence segmentation processing. Section 4 shows the experimental results and a complete evaluation metrics are also given in this section. Finally, conclusions are drawn in Section 5.

1. Network flow framework

The target motion of multi-objet tracking can be described better by the relationship of the neighborhood location between adjacent frames, which uses the DP method in a min-cost network flow framework. We define an objective function for multi-object tracking equivalent to that of [13]. The objective presence of likelihood will be estimated by the marginal posterior probability in every frame, thereby obtaining the potential object moving trajectory.

1.1. Min-cost flow model

We formulate the multi-object tracking as a whole process, in which the objective location of each time continuously and discretely changes with time. A directed 3D spatiotemporal group with random variable k is used to describe the video sequence.

$$k = (x, y, t), \quad x \in V \tag{1}$$

where k denotes any location of an object in this spatiotemporal group, V is the set of all space–time locations in a sequence, x and y are the pixel positions of the target in the transverse and longitudinal axes respectively, and t is every instant of time.

For any location k at time t, the neighborhood $N(k) \subset \{1, 2, \ldots, K\}$ denotes the locations an object can reach at time $t+1$. A single track, as an ordered set of state vectors $T = (k_1, \ldots, k_N)$ and $X = (T_1, \ldots, T_L)$ is a set of the collection of tracks. We assume that the tracking tracks independently of each other, and describe the network flow framework of multi-object tracking using the dynamic model as follows:

$$P(X) = \prod_{T \in X} P(T) \tag{2}$$

where

$$P(T) = P_{\text{source}}(k_1) \left(\prod_{n=1}^{N-1} P\left(k_{n+1} \middle| k_n\right) \right) P_{\sin k}(k_N) \tag{3}$$

$P_{\text{source}}(k_1)$ is the probability of a tracking track starting at location k_1 and $P_{\sin k}(k_N)$ is the probability of a tracking track ending at location k_N.

In the spatial coordinate set V, a binary indicator variable $\varphi_{i,k}$ for the directed flow from location i to location k, which stands for the number of objects moving from i to k. $\varphi_{i,k}$ is 1 when the space–time location i and k are included in some track, if location i at time t and k at time $t+1$, which means that an object remains at the same spatial location between times t and $t+1$. Some constraint conditions are executed for the variable $\varphi_{i,k}$.

$$\forall k, \quad \sum_{i,k \in N(i)} \varphi_{i,k} = \varphi_k = \sum_{j \in N(k)} \varphi_{k,j} \tag{4}$$

$$\forall i, k, \quad \sum_{k \in N(i)} \varphi_{i,k} \le 1 \tag{5}$$

Let a random variable M_k stands for the true presence of an object at location k in space–time. For every time t, the detector is used to check every location of the tracking zone. The marginal posterior probability of an existing object is estimated as follows

$$\rho_k = \hat{P}\left(M_k = 1 \middle| \mathbf{I}_t\right) \tag{6}$$

where $\mathbf{I}_t$ is the single image at frame t. We write $m = \{m_k\}$ for a feasible set of the existing likelihood probability distributions of objects in V by the method of Section 3.1, and M is the spatial set of M_k. The existence likelihood probability of an object in the given set of tracks X is

$$P\left(M = m \middle| X\right) = \prod_{k \in X} P\left(M_k = m_k \middle| X\right) \tag{7}$$

M_k is conditional independence in the given X, we can infer the maximum a posteriori estimate of tracks by the existing likelihood probability distributions of objects.

$$X^* = arg\max_X P(X) P\left(M = m \middle| X\right) \tag{8}$$

$$= arg\max_X \prod_{T \in X} P(T) \prod_{k \in X} P\left(M_k = m_k \middle| X\right) \tag{9}$$

$$= arg\max_X \sum_{T \in X} \log P(T) + \sum_{k \in X} \log P\left(M_k = m_k \middle| X\right) \tag{10}$$

$$= arg\max_X \sum_{T \in X} \log P(T) + \sum_{k} \left[(1 - m_k) \log P\left(M_k = 0 \middle| X\right) + m_k \log P\left(M_k = 1 \middle| X\right)\right] \tag{11}$$

$$= arg\max_X \sum_{T \in X} \log P(T) + \sum_{k} m_k \log \frac{P\left(M_k = 1 \middle| X\right)}{P\left(M_k = 0 \middle| X\right)} \tag{12}$$

$$= arg\max_X \sum_{T \in X} \log P(T) + \sum_{k} m_k \log \left(\frac{\rho_k}{1 - \rho_k}\right) \tag{13}$$

where Eq. (11) is true because m_k is 0 or 1 according to Eq. (5), for Eq. (10), and we obtain Eq. (12) by ignoring a term that does not need m_k. The cost value of a directed flow between the neighborhood locations of any adjacent frames is defined as

$$c\left(e_{k,n}\right) = -\log \left(\frac{\rho_k}{1 - \rho_k}\right) \tag{14}$$