With the progress of object detection techniques, Multiple Object Tracking (MOT) has developed rapidly. Tracking by detection methods based on network flow attract much attention in the MOT area.

Let \( L = \{l_i\} \) be a set of object detections, where \( l_i \) denotes the ith detection, \( f \) is the frame index, and let \( L = \{l_i\} \) be a set of the trajectories. Tracking by detection MOT problem can be formulated as equation (1).

\[
\hat{L} = \arg\max P(L | \hat{L}) \log q(L)
\]

Fig. 1. Tracking system

1. Raw detections as inputs are preprocessed firstly. In this process, abnormal bounding boxes will be deleted, which are too large or small and appear in unreasonable positions.
2. Preprocessed detections are linked into tracks in the Tracklet Generation module.
3. We extract information and feature from the tracklets set.
4. Occlusions and missing-detections are modeled in the Virtual-Missing-detection and Occlusion Model (VMOM) which we proposed. Then the model is embedded in a network.
5. We develop an efficient Counter Embedded Iterative Shortest Paths (CEISP) algorithm which is suitable for our tracking model to get final trajectories.

**Approach**

**Tracklet Generation**

\( L = \{l_i\} \) is treated as preprocessed detections set. Tracklets are generated by direct-link method. Link probability between two detections is based on distance, size, appearance, Bhattacharyya coefficient and Gaussian function are used to get the similarity.

**State Judgment**

\( T = \{t_i\} \) is the set of tracklets generated by last step. A forward-backward searching method (2) is proposed to judge state in the step function, \( A_{i,j} \) denotes the affinity between \( i \) and \( j \), [0.1] indicates start, [1.1] terminal one, [1,1] intermediate one and [0,0] a complete one.

\[
\text{state}(t_i) = \{l_i \sum_{y=3} M_{y-2} \sum_{N} \log(l_i, \sum_{j=3} M_{y-2} \sum_{N} \log(l_i, A_{i,j} \log q(L))
\]

Experiments and Conclusion

We utilize the open Multiple Object Tracking Benchmark to evaluate our method. MOT-2015 and MOT-2016 are the datasets. Table 1 presents tracking performance. Average tracking speeds are 43.9fps and 37.4fps respectively.

Table 1. Tracking performance

<table>
<thead>
<tr>
<th>Dataset</th>
<th>MOT-2015</th>
<th>MOT-2016</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOT-2015</td>
<td>43.9</td>
<td>51.0</td>
</tr>
<tr>
<td>MOT-2016</td>
<td>37.4</td>
<td>51.0</td>
</tr>
</tbody>
</table>

Fig. 4. Samples of tracking results

Conclusion: Local occlusions can be processed in our method as shown in Fig. 4. Virtual nodes recover the occlusions accurately. They also bring some mistakes. From the tracking results as shown in Table 1, our tracker can track objects in these scenes very accurately. Tracklets with single detection makes some confusions during the occlusion processing period, which are presented by the indicators such as PP, FN and IDs. Our tracker needs further research to be improved.