Image segmentation and target tracking based on meanshift algorithm

Chen Rong¹, a

¹Department of Physics, Nanchang Normal University, Nanchang, 330032, China
aemail: xinsea@163.com

Keywords: image; segmentation; target; algorithm; tracking; meanshift

Abstract. The kernel function is proposed for the analysis of meanshift algorithm, the research of meanshift algorithm mainly analyses the two common characteristics of meanshift algorithms based on kernel function histogram and probability distribution. Image segmentation based on meanshift is proposed for the extraction of target. The meanshift tracking algorithm based on the background difference suppression adopted, thus to give greater weight to the region with significant background difference around target and make sure the goal always in the tracking box. This thesis which is based on analyzing the characteristics of moving target tracking algorithm, uses the meanshift algorithm to accomplish moving target tracking. By comparing the Bhattacharyya coefficient of similarity function between target model and measurement model, the idea of meanshift is to find the position of target through multiple iterations drift, eventually reaching an ideal result, to complete the target tracking. The algorithm of target tracking based on meanshift in continuous video sequences is extensive. Therefore, target detection and tracking in continuous video sequence has a lot of practical and applied value. Experimental results show that the algorithm can help achieve fast and effective target tracking. It could be easily realized to get accurate tracking position for low-speed targets, but not suitable for tracking fast moving target. This algorithm can solve the problem in the tracking algorithm, that is when the pedestrian moving fast, it is easy to lose tracked target.

Introduction

Computer vision is a hot research topic that has won wide attention, it has achieved rapid development in recent years. In computer vision, image processing and pattern recognition, tracking the different targets after detecting is an important content and a large field in continuous video sequences. Visual target tracking technology, the important part of computer vision, has attracted a large number of research scholars to put forward a lot of excellent tracking algorithm. Visual tracking has been extensively studied and concerned because of its importance in practical applications. The moving human tracking technology in video has become an important topic and research focus in computer vision. The detection and tracking of multiple moving targets has become the one of the hottest research topics in the field of computer vision, and its technology involve digital image processing and pattern recognition. The detection and tracking of the multiple moving targets can be summarized as follows. Firstly, using the image capture technology to extract the images from a video sequences. Secondly, using image processing techniques for each frame image denoising processing. Then, selecting a suitable method of detection to extract the target from the image. Finally, using pattern recognition technology and tracking algorithm to track the target.

The meanshift algorithm is the most commonly used in target tracking algorithm, firstly analysis the target, establishes a target model, then in the subsequent frames establish test model.

The definition of kernel function

The one-dimensional spatial kernel functions satisfy the following conditions:

\[ k_h(x) = k_h(-x) \]  \hspace{1cm} (1)

\[ k_h(x) > 0 \]  \hspace{1cm} (2)
\begin{align*}
\int k_h(x)\,dx &= 1 \\
k_h(x) &= \frac{1}{h} k\left(\frac{x}{h}\right)
\end{align*}

In the type, $h$ is the bandwidth of kernel. The following characteristics of kernel function is that: the kernel function is finite local support, symmetric, unimodal\cite{1}.

### The principle of Meanshift algorithm

The meanshift algorithm is proposed by Fukunaga and Hostetler in 1975, which is a non parametric density estimation algorithm, then the meanshift algorithm has been widely used. In modern scientific research, a large number of research institutions and researchers use kernel functions which include Uniform, Gaussian, Epanechnikov and so on. Given a sample set, using the mean shift vector adaptive search the dense regions of the sample. The mean shift vector is realized by calculating the gradient of probability density function of sample sets\cite{2}, such as in the type (5).

$$m(x) = \frac{\sum_{i=1}^{n} g\left(\frac{\|x - x_i\|^2}{h}\right)}{\sum_{i=1}^{n} g\left(\frac{\|x - x_i\|^2}{h}\right)} - x$$

In the type (5): $x$ is vector value at the center of the searching window. $x_i$ is a vector value of the first $i$ point in the searching window, $h$ is the width of the searching window.

$$g(x) = -k^\prime(x)$$

In the type (6), $k(x)$ satisfy the following formula:

\begin{align*}
&k(x) > 0 \\
k(x) &= k(-x) \\
\int k(x)\,dx &= 1
\end{align*}

The meanshift algorithm is an iterative algorithm, the specific process of meanshift algorithm is: first, calculate the average values for offset point which is to be tested, moving the point to its offset mean point, then the offset mean point is as the new test points, once again, the calculation offset mean of new point to be tested, so the iterative cycle, satisfy the corresponding conditions until the end\cite{3}. In fact, the meanshift algorithm iteratively calculate the probability density by using the gradient method along the gradient which is relatively large changes in direction. finally, the convergence to the optimum of probability density\cite{4-5}.

### Target tracking based on meanshift algorithm

The basic idea of kernel method is continuous estimation directly for similarity probability density function. Using color feature probability density describes the target motion model of the regional center based on meanshift algorithm, to measure the similarity between the maximum in the center of the target neighborhood search window, if the similarity value does not meet the condition, then it conducts original meanshift tracking in turn from other moving regions and find out the best tracking region, looking for the optimal target model, making the two moving distance is less than a threshold which is preset, the real position in order to achieve the goal. The real position of the target have been obtained\cite{6}. Thus, the meanshift algorithm can fast find back the original targets to track again once losing them during a long period tracking.
Figure 1 shows the result of image segmentation, as can be seen from the figure, removing the clouds and pollutants, the birds are extracted very clear, and easy to identify. Figure 2 shows the result of meanshift algorithm tracking, $i$ is the number of frame image. So the difference color between target and background is bigger, the extraction of target can be much better.
Fig. 2. The result of meanshift algorithm tracking

<table>
<thead>
<tr>
<th>i</th>
<th>cet</th>
<th>r1</th>
<th>r2</th>
<th>c1</th>
<th>c2</th>
<th>i</th>
<th>cet</th>
<th>r1</th>
<th>r2</th>
<th>c1</th>
<th>c2</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>213</td>
<td>371</td>
<td>202</td>
<td>224</td>
<td>366</td>
<td>3</td>
<td>214</td>
<td>370</td>
<td>203</td>
<td>225</td>
<td>365</td>
</tr>
<tr>
<td>4</td>
<td>212</td>
<td>365</td>
<td>201</td>
<td>223</td>
<td>357</td>
<td>5</td>
<td>212</td>
<td>363</td>
<td>201</td>
<td>223</td>
<td>358</td>
</tr>
<tr>
<td>6</td>
<td>212</td>
<td>362</td>
<td>201</td>
<td>223</td>
<td>357</td>
<td>7</td>
<td>213</td>
<td>357</td>
<td>202</td>
<td>224</td>
<td>352</td>
</tr>
<tr>
<td>8</td>
<td>213</td>
<td>354</td>
<td>202</td>
<td>224</td>
<td>349</td>
<td>9</td>
<td>213</td>
<td>352</td>
<td>202</td>
<td>224</td>
<td>347</td>
</tr>
<tr>
<td>10</td>
<td>213</td>
<td>347</td>
<td>202</td>
<td>224</td>
<td>346</td>
<td>11</td>
<td>213</td>
<td>346</td>
<td>202</td>
<td>224</td>
<td>341</td>
</tr>
<tr>
<td>12</td>
<td>212</td>
<td>339</td>
<td>201</td>
<td>223</td>
<td>334</td>
<td>13</td>
<td>212</td>
<td>339</td>
<td>201</td>
<td>223</td>
<td>334</td>
</tr>
<tr>
<td>14</td>
<td>211</td>
<td>332</td>
<td>200</td>
<td>222</td>
<td>327</td>
<td>15</td>
<td>212</td>
<td>332</td>
<td>201</td>
<td>223</td>
<td>327</td>
</tr>
<tr>
<td>16</td>
<td>212</td>
<td>331</td>
<td>201</td>
<td>223</td>
<td>326</td>
<td>17</td>
<td>211</td>
<td>327</td>
<td>200</td>
<td>222</td>
<td>322</td>
</tr>
<tr>
<td>18</td>
<td>209</td>
<td>325</td>
<td>198</td>
<td>220</td>
<td>320</td>
<td>19</td>
<td>209</td>
<td>324</td>
<td>198</td>
<td>220</td>
<td>319</td>
</tr>
<tr>
<td>20</td>
<td>210</td>
<td>322</td>
<td>199</td>
<td>221</td>
<td>317</td>
<td>21</td>
<td>210</td>
<td>322</td>
<td>199</td>
<td>221</td>
<td>317</td>
</tr>
<tr>
<td>22</td>
<td>210</td>
<td>320</td>
<td>199</td>
<td>221</td>
<td>315</td>
<td>23</td>
<td>210</td>
<td>319</td>
<td>199</td>
<td>221</td>
<td>314</td>
</tr>
<tr>
<td>24</td>
<td>210</td>
<td>319</td>
<td>199</td>
<td>221</td>
<td>314</td>
<td>25</td>
<td>210</td>
<td>317</td>
<td>199</td>
<td>221</td>
<td>312</td>
</tr>
<tr>
<td>26</td>
<td>210</td>
<td>317</td>
<td>199</td>
<td>221</td>
<td>312</td>
<td>27</td>
<td>210</td>
<td>317</td>
<td>199</td>
<td>221</td>
<td>312</td>
</tr>
<tr>
<td>28</td>
<td>211</td>
<td>313</td>
<td>200</td>
<td>222</td>
<td>308</td>
<td>29</td>
<td>210</td>
<td>313</td>
<td>199</td>
<td>221</td>
<td>308</td>
</tr>
<tr>
<td>30</td>
<td>210</td>
<td>313</td>
<td>199</td>
<td>221</td>
<td>308</td>
<td>31</td>
<td>211</td>
<td>313</td>
<td>199</td>
<td>220</td>
<td>308</td>
</tr>
<tr>
<td>32</td>
<td>220</td>
<td>309</td>
<td>209</td>
<td>231</td>
<td>304</td>
<td>33</td>
<td>210</td>
<td>313</td>
<td>199</td>
<td>221</td>
<td>308</td>
</tr>
<tr>
<td>34</td>
<td>211</td>
<td>313</td>
<td>200</td>
<td>222</td>
<td>308</td>
<td>35</td>
<td>210</td>
<td>313</td>
<td>199</td>
<td>221</td>
<td>308</td>
</tr>
<tr>
<td>36</td>
<td>210</td>
<td>313</td>
<td>199</td>
<td>221</td>
<td>308</td>
<td>37</td>
<td>209</td>
<td>310</td>
<td>198</td>
<td>220</td>
<td>305</td>
</tr>
<tr>
<td>38</td>
<td>209</td>
<td>309</td>
<td>198</td>
<td>220</td>
<td>304</td>
<td>39</td>
<td>209</td>
<td>307</td>
<td>198</td>
<td>220</td>
<td>302</td>
</tr>
</tbody>
</table>

Table I shows the result of target tracking, \( i \) is also the number of frame image, \( cet \) is at the center of tracking box, \( r_i \) is the maximum value of row in the tracking box, \( r_i \) is the minimum value of row in the tracking box, \( c_i \) is the minimum value of column in the tracking box, \( c_i \) is the maximum value of column in the tracking box.

**Conclusion**

Experiments show that the effect of image segmentation based on meanshift is very good. The meanshift algorithm has many advantages in target tracking:

1. Small amount of computation of meanshift algorithm, the real-time property of the algorithm is relatively high.
2. It is a non parameter estimation algorithm, it is easy to combine with other tracking algorithm.

727
(3) The meanshift algorithm have little effect on the target deformation by itself, it has little effect on background change, so meanshift algorithm have been widely applied. The results of tracking experiments show that the method computes faster in tracking.

The meanshift algorithm shortcomings are as follows:

(1) The meanshift algorithm only considered target’s color statistical information, and didn’t contain target’s space information, so when the target color closed to the background color, meanshift algorithm easily caused target’s tracking inaccurately or lost.

(2) The meanshift algorithm has problems in inaccurate tracking of fast-moving targets and easy to lose the targets enduringly tracked.

Acknowledgement

In this paper, the research was sponsored by the key discipline of Nanchang Normal University, China (Project No. NSXK20141001) and the Reform in Education of Nanchang Normal University, China (Project No. JGKT-14-18).

References


