Advanced Materials Research Vol 186 (2011) pp 281-286 *Online: 2011-01-11*
© (2011) Trans Tech Publications, Switzerland
doi:10.4028/www.scientific.net/AMR.186.281

The Target Tracking Method Based on Camshift Algorithm Combined with SIFT

Jieyu Zhang [1, a], Haiyong Wu [2, b], Shu Chen [1, c] and Deshen Xia [3, d]

[1]Department of Basic Sciences, China Pharmaceutical University, Nanjing 211198, China

[2]Department of Physics and Electronic Engineer, Nanjing Xiaozhuang University, Nanjing 211171, China

[3] Department of Computer Science & Technology, Nanjing University of Science & Technology, Nanjing, China 210094 China

[a]zjyscy@gmail.com, [b]why@njxzc.edu.cn, [c]chenshujs@126.com, [d]deshen_x@263.net

Keywords: Camshift SIFT feature points object tracking.

Abstract. Since Camshift algorithm leads to failed tracking results when the color information of the target region is similar with the background or is not precise enough, to solve this problem a tracking method based on Camshift and SIFT was proposed in this paper. In this method, SIFT feature points, which were used to construct the color histogram and the color probability distribution, were extracted from the target region first. Then SIFT points were also extracted from the search region and these two sets of SIFT points were matched. Since the proposed method used the matched SIFT points to properly guide the location of targets, experimental results show that with the new method some more accurate and robust tracking results have been obtained.

Introduction

Object tracking, which means analyzing the image sequence obtained from sensors and computing the target location in each of the images, has received extensive attention in recent years. Many kinds of methods have been designed and studied in order to adapt different tracking environments and requirements. Nowadays, a kind of method combining regions with color models has been greatly developed, such as Mean Shift and Continuously Adaptive Mean-Shift (Camshift) [1]. Camshift algorithm, with the advantages of simple computations and good effects without any parameter, has been very widely used in object tracking [2,3,4]. However, the tracking process of Camshift will fail when the distinction of color information is not high enough. So, Guo [5] discarded interference colors before Camshift steps to improve the precise of color information. Allen [6] proposed a special method of computing weighting factor to minimize the impact of background region. Furthermore, Nouar [7] extracted facial contours with Canny operator as the initial search region, which avoided pixels in background regions participating in constructing histograms.

Since the invariant features has good quality, a new target tracking method is proposed in this paper based on combining invariant feature points with color information to get better tracking results as accurate as possible. In this new algorithm, SIFT feature points were extracted from three channels H, S, V of the target region respectively and all of those SIFT points were used to construct a color histogram in H channel. Then the color probability distribution was calculated in the next frame of the image sequence. Subsequently, SIFT points were also extracted from the H, S, V three channels of the search region. After matching the two sets of SIFT points, the image block was obtained bounded by the matched points in the search region, followed by computing the location of the next search region in the bounded block with the zero moment and the first moments. Finally, the size of the next search window was determined by the mean value of scale variable of all of the matched pairs. Since the proposed method used the matched SIFT points to properly guide the location of targets, experimental results show that some more accurate and robust tracking results have been obtained with the new method.

The Theory of Camshift

Camshift, a robust method with low-computation and insensitivity to deformation of targets, based on evaluating gradient of density function without parameters, can automatically adjust the location and the size of each search window by the trait of continuous change of probability distribution. So, this algorithm has been widely used in many tracking situations, such as human faces, hands, robot visions and so on. The detailed steps of Camshift are shown as follows:

Step1. Read the first frame and transform it from RGB color space to HSV color space.

Step2. Manually select a region of interest as the initial search window and compute the histogram $\left\{\hat{q}_u\right\}_{u=1,\ldots,n}$ in the H channel.

Step3. Compute the probability distribution p_u of the current frame by the histogram of the target window $\left\{\hat{q}_u\right\}_{u=1,\ldots,n}$ obtained in the step2, shown as the Eq. 1.

$$p_u = \min\left(\frac{255\hat{q}_u}{\max(\hat{q}_u)}, 255\right)_{u=1,\ldots,n} \tag{1}$$

Step4. Iteratively compute the zero moment and the first moments, shown as the Eq. 2, using Mean Shift algorithm with the probability distribution to determine the location and the size of the next search window.

$$M_{00} = \sum_x \sum_y I(x,y), \quad M_{10} = \sum_x \sum_y xI(x,y), \quad M_{01} = \sum_x \sum_y yI(x,y). \tag{2}$$

So the center and the size of the search window can be obtained by the Eq. 3 and the Eq. 4 respectively.

$$x_c = \frac{M_{10}}{M_{00}}, \quad y_c = \frac{M_{01}}{M_{00}}. \tag{3}$$

$$s = 2\sqrt{\frac{M_{00}}{256}}. \tag{4}$$

Step5. Do not stop the iterative computation of step4 until the location of the search window converges.

In view of totally depending on the color information of targets, Camshift algorithm usually leads to failed tracking results when the color information is wrong or not accurate enough. Therefore, the new method based on combining Camshift and SIFT features is proposed in this paper.

The Tracking Method Based on Combining Camshift with SIFT

There exist two main problems in Camshift algorithm.

1) When the background contains some objects with the colors which are similar with the targets, namely, the color information is not distinctive, Camshift usually gets failed tracking results.

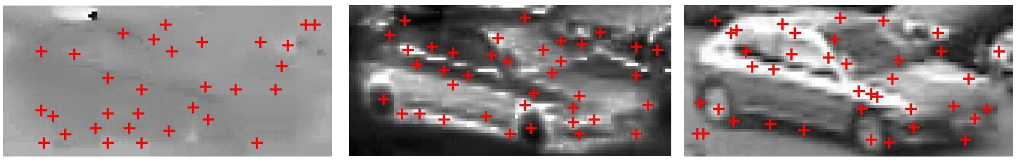
2) All of the pixels in the target window participate in construction the color histogram. However, a part of those pixels aren't located in the real targets, which may results in the inaccurate color histograms.

On the account of the problems mentioned above, the new tracking algorithm based on Camshift and SIFT features is proposed.

Computing the Color Histogram of Target Regions. Firstly, the target region selected manually is transformed from RGB color space to HSV color space. Then SIFT feature points, containing rich information of the image, are extracted from the target region in H, S, V channels respectively. Subsequently, all of the SIFT points can be used to construct the color histogram in H channel. Considering that the most of the SIFT points are located in the real targets when the texture of the background in the target region is not complicated. So the color histogram constructed by the SIFT points can more accurately reflect the color information of the real targets, which is highly advantageous to the following tracking.

Fig. 1 (a) is an original RGB image, with a white rectangle window denoting the target region and Fig. 1 (b) is the enlarged result of the target region inside Fig. 1 (a). The red crosses, in Fig. 1 (c), (d) and (e), denote SIFT points extracted from the H, S and V channels respectively. It can be seen that most of the SIFT points are located in the real target.

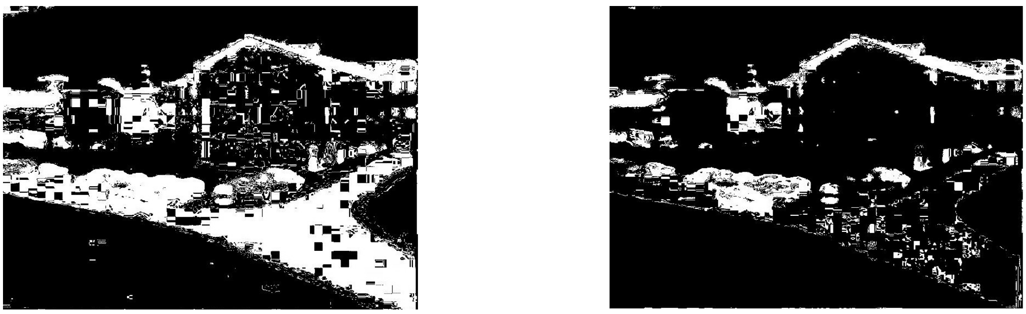
(a) An original RGB image (b) The enlarged result of the target region



(c) The SIFT points in H channel (d) The SIFT points in S channel (e) The SIFT points in V channel

Fig.1 The target region and SIFT points

Then the more accurate color histogram can be used to create the color probability distribution. In order to evaluate the superiority of the new color histogram, the new color probability distribution constructed by the new color histogram is compared with the original one in Camshift, shown in Fig. 2 (a) and (b).

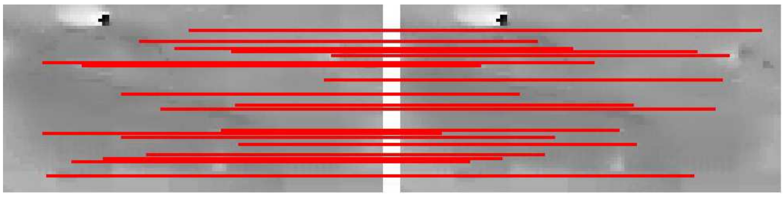


(a) The color probability distribution in Camshift (b) The color probability distribution in the new method

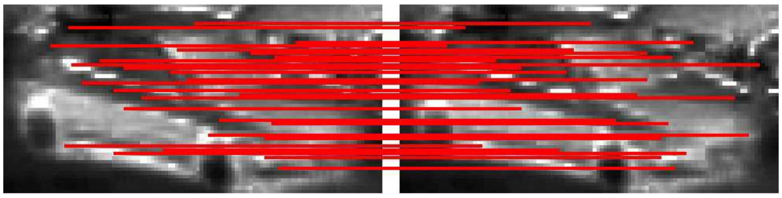
Fig.2 The comparison of two kinds of color probability distribution

The shade of the color in the probability distribution reflects the proportion of probability, namely, the higher the brightness of a certain point is, the greater the likelihood of belonging to the target region is. It can be seen that in Fig. 2 (a) many non-target blocks have high probabilities, such as all of the roads. The reason is that all of the pixels in the target region containing located not only in the real target (the blue car) but also in the background (the roads) participate in constructing the color histogram. On the contrary, in the new method, the majority of the pixels in the background which are independent of real targets are discarded. SIFT points mostly located in the real target are used to construct the new color histogram which reflect the color information of the real target more accurately. So in the Fig. 2 (b) the real target block has a higher probability and most of the non-target blocks have lower probabilities, which results in a better discrimination.

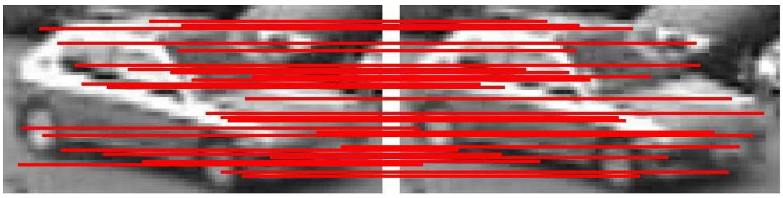
Extracting SIFT Feature Points from the Search Region. Similarly, SIFT feature points are extracted from the search region in H, S and V channels. Then the two sets of SIFT points in the target region and search region are matched by SIFT algorithm. Fig.3 (a), (b) and (c) are the matching results in three channels respectively, with red lines connecting the matching point pairs.



(a) The matching result of H channel



(b) The matching result of S channel



(c) The matching result of V channel

Fig.3 The matching results of H, S, V channels

In Camshift algorithm, all of the pixels in the last search window participate in computing the center and the size of the next search window. Those pixels, located in non-target regions with larger probability in the color probability distribution map, cause great disturbance to the calculations. But in the new tracking method, the matched SIFT points in the last search window mostly located in the real targets are used to compute the parameters of the next search window in order to reduce the error.

Determining the Parameters of the Search Window. On account of the matched SIFT points being mostly located in the real targets, the pixels in the block region bounded by the matched SIFT points in the last search window with more pixels of real targets, are used to compute the zero moment and the first moments to determine the location and the size of the next search window. This approach can not only correctly guide the location of the next search window by the matched SIFT points but also obtain the more accurate zero moment and the first moments needed by the next search window.

The method for computing the location of the next search window is still following the Eq. 2 and Eq. 3. When the scale of the target changes, SIFT algorithm with scale invariance can still extract the feature points with scale parameters. In the new algorithm, the size of the next search window is determined by the mean value of scale variable of all the matched SIFT points, shown as the Eq. 5, in which n denotes the total number of the matched pairs, s_m denotes the change factor of each pair and

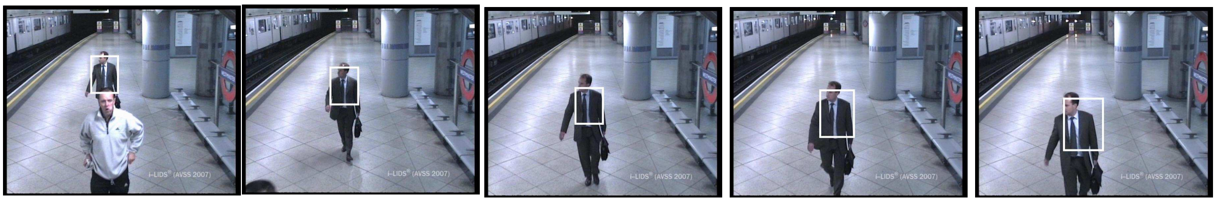
s denotes the variable of the search window respectively. So the width and the height of the next search window can be calculated by s .

$$s = \frac{1}{n}\sum_{m=1}^{n} s_m \tag{5}$$

Experimental Results and Analysis

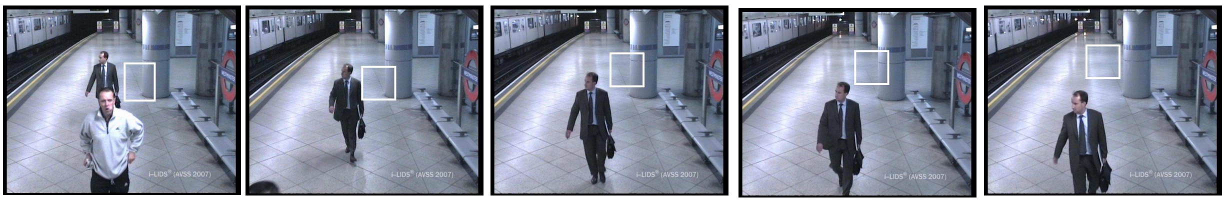
In order to evaluate the new tracking method, two image sequences are used to compare Camshift and the new method. The first image sequence contains one hundred frames about a pedestrian walking from far and near in the metro station. The second image sequence also includes one hundred frames with a man moving in a very dark room.

Fig. 4 and Fig. 5 are the tracking results of the first image sequence by the new method and Camshift, respectively. Camshift algorithm quickly lost the real target because of the disturbance of ground pixels in the target region. But the new method combining the SIFT feature points with the more accurate color information and color probability distribution obtained better tracking results.



(a) The 2nd frame (b) The 37th frame (c) The 60th frame (d) The 70th frame (e) The 85th frame

Fig.4 The tracking results of the pedestrian by the new method



(a) The 2nd frame (b) The 37th frame (c) The 60th frame (d) The 70th frame (e) The 85th frame

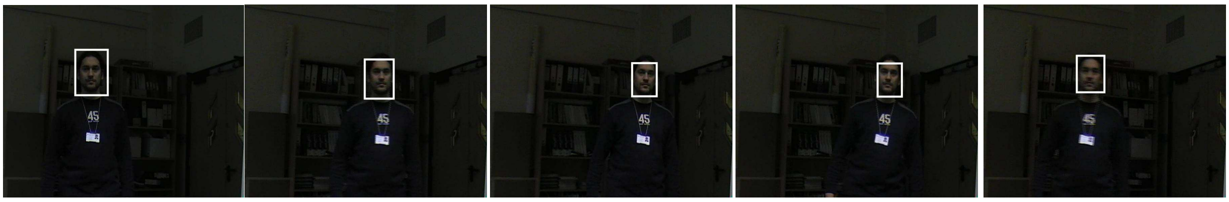
Fig.5 The tracking results of the pedestrian by Camshift

Similarly, Fig. 6 and Fig. 7 are the results of the second image sequence by the new method and Camshift respectively. The target region is the whole face moving in a very dark room. Because the color of the target region and the background is similar, the regions tracked by Camshift were smaller than the real target regions. But the tracking results of the new method obtained better effects.

Meanwhile, the time complexity of the proposed algorithm is higher than the Camshift algorithm because of introducing SIFT. For the two image sequences above mentioned, the average tracking time of the new method and the Camshift is 0.11287 second/frame and0.04409 second/frame respectively, using Matlab7.0 on the PC with Intel Pentium4 2.93GHz and 512M memory. But it should be noted that the new algorithm combining feature points and color information can get more accurate tracking results than Camshift algorithm. So, the new algorithm proposed in the paper has practical value for the target tracking systems with low real-time demand but high processing accuracy.

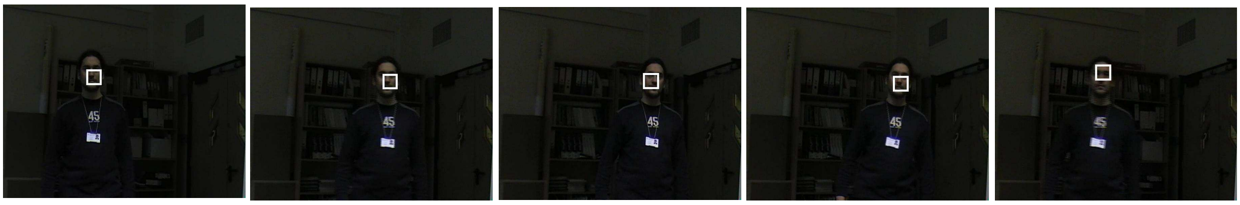
Summary

In order to cope with the problem of inaccurate color information leading failed tracking results, a new tracking method combining Camshift and SIFT feature points is proposed. In the new algorithm, SIFT feature points with good invariance are extracted and used to correctly guide the target position. This kind of invariant features makes up for the inaccurate color information. Experimental results show that the new algorithm can obtain more accurate and robust tracking results than Camshift.



(a) The 5th frame (b) The 33rd frame (c) The 49th frame (d) The 68th frame (e) The 82nd frame

Fig.6 The tracking results of the human face by the new method



(a) The 5th frame (b) The 33rd frame (c) The 49th frame (d) The 68th frame (e) The 82nd frame

Fig.7 The tracking results of the human face by Camshift

References

[1] G.R.Bradski, S.Clara. *Computer vision face tracking for use in a perceptual user interface*. Intel Technology Journal, 1998, 2 (2): 1-15.

[2] J.Biemond, P.M.B Roosmalen, R.L.Lagendijk. *Improved blotch detection by postprocessing*. IEEE International Conference on Acoustics and Signal Processing, 1999, 3101-3104.

[3] Guo Shilong, Li Wenfeng, Li Bo and Jin Jiagen. *Camshift algorithm based mobiled robots visual tracing system*. Journal of Huazhong University of Science and Technology(Nature Science Edition), 2008, 36 (Sup):156-158.

[4] Xu Kun, He Yuyao and Wang Weiya. *Object tracking algorithm with adaptive color space based on Camshift*. Journal of Computer Applications, 2009, 29(3): 757-760.

[5] Kou Chao, Bai Cong and Chen Quanlin. *An improved back-projection based Camshift algorithm for face tracking*. Computer simulation, 2009, 26(6): 228-231.

[6] J. G. Allen, R. Y. D. Xu, J. S. Jin. *Object tracking using camshift algorithm and multiple quantized feature spaces*. Proceedings of the Pan-Sydney area workshop on Visual information processing, Darlinghurst, Australia, 2004, 2-7.

[7] O. D. Nouar, G. Ali, C. Raphael. *Improved object tracking with camshift algorithm*. Proceedings of International Conference on Acoustics, Speech, and Signal Processing. Toulouse, France, 2006, 2: II – II.

New Trends and Applications of Computer-aided Material and Engineering
10.4028/www.scientific.net/AMR.186

The Target Tracking Method Based on Camshift Algorithm Combined with SIFT
10.4028/www.scientific.net/AMR.186.281

DOI References

[4] Xu Kun, He Yuyao and Wang Weiya. Object tracking algorithm with adaptive color space based on Camshift. Journal of Computer Applications, 2009, 29(3): 757-760.
doi:10.3724/SP.J.1087.2009.02411