



文章编号:1001-9081(2017)04-1189-04

DOI:10.11772/j.issn.1001-9081.2017.04.1189

## 基于图割理论的尺度自适应人脸跟踪算法

胡章芳, 秦阳鸿<sup>\*</sup>

(重庆邮电大学 光电子信息感测与传输技术重点实验室, 重庆 400065)

(\*通信作者电子邮箱 [chinaqinyh@163.com](mailto:chinaqinyh@163.com))

**摘要:**针对连续自适应的 Mean-Shift(Camshift) 算法跟踪人脸时尺度过度放缩这一问题,提出了一种基于图割的 Camshift 人脸跟踪算法。首先,在每一帧图像的 Camshift 迭代结果内建立图割区域,使用高斯肤色模型作为图割权值分割出图割区域内肤色团块;然后,计算该肤色团大小得到目标真实尺度,并比较与上一帧图像跟踪框内肤色团的尺度来判断是否需要重新跟踪目标;最后,再以该团块作为下一帧跟踪目标。实验结果表明,基于图割的 Camshift 人脸跟踪算法有效地克服了跟踪时其他肤色区域的干扰,能有效地反映人体快速运动中人脸真实尺度变化,同时防止 Camshift 算法丢失跟踪目标而陷入局部最优解,具有较好的可用性和鲁棒性。

**关键词:**图割; Camshift 算法; 目标跟踪; 真实尺度; 最大流

**中图分类号:**TP391.41    **文献标志码:**A

### Scale-adaptive face tracking algorithm based on graph cuts theory

HU Zhangfang, QIN Yanghong<sup>\*</sup>

(Key Laboratory of Optoelectronic Information Sensing and Transmission Technology,  
Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

**Abstract:** Aiming at the problem of the excessive size-changing while the tracking window is enlarged by traditional Continuously Adaptive MeanShift (Camshift) algorithm in face tracking, an adaptive window face tracking method for Camshift based on graph cuts theory was proposed. Firstly, a graph cut area was created according to the Camshift iteration result of every frame by using graph cuts theory, and the skin lump was found by using Gaussian mixture model as weights of graph cuts. As a result, the tracking window could be updated by the skin lump. Then the real size of the target was obtained by computing the size of skin lump, and whether the target needed to be re-tracked was determined by comparing the size of the skin lump in the tracking window with that in the previous frame. Finally, the skin lump in last frame was used as the tracking target of the next frame. The experimental results demonstrate that the proposed method based on graph cuts can avoid interference of other skin color targets in the background, which effectively reflects the real face size-changing of the human body in rapid movement, and prevents the Camshift algorithm from losing the tracking target and falling into the local optimal solution with good usability and robustness.

**Key words:** graph cuts; Camshift algorithm; target tracking; true measure; max-flow

### 0 引言

基于视觉的人脸识别是人机交互领域的一个重要研究课题<sup>[1]</sup>。这种交互方式自然、直观,其具有广泛的应用前景。

连续自适应的 Mean-Shift (Continuously Adaptive Mean-Shift, Camshift) 是一种常用的跟踪算法,在跟踪过程中容易出现窗口尺度过度放缩问题<sup>[2]</sup>。在 Camshift 算法中,跟踪窗口直接影响跟踪结果,如果跟踪窗口不能真实反映目标实际尺度的变化,可能会导致丢失跟踪目标。尺度控制问题决定了 Camshift 算法的鲁棒性。

国内外许多研究学者对跟踪窗口尺度控制开展了研究。由 Li 等<sup>[3]</sup>提出的 SURF (Speeded Up Robust Features) 融合进 Camshift 算法有效地跟踪复杂背景中的目标,但是在跟踪目标尺度发生剧烈变化时,容易丢失目标。还有一些基于这种思想的改进方法,如徐建军等<sup>[4]</sup>提出了融合 Adaboost (Adaptive Boostint) 算法来保证跟踪目标的真实尺度,但是该

算法要提取每一帧图像加入分类器中,导致计算量过大,跟踪效率大大降低;王鑫等<sup>[5]</sup>提出将粒子滤波融合 Camshift 算法中来计算跟踪窗口大小,该方法能较好地反映真实尺度,但是粒子数会影响跟踪效果,粒子数过低会丢失跟踪目标,而随着粒子数增加,算法运算量也会增加。

为了克服以上算法的不足,本文提出将图割算法融入 Camshift 中解决尺度放缩这一问题。首先利用 Camshift 跟踪目标得到跟踪目标位置,再使用跟踪框内的像素点构造图割网络,通过能量函数计算跟踪框范围内的最大流完成图割。图割算法使用高斯肤色模型作为能量函数权值,然后在跟踪窗口内分割出肤色团块,最大团块即是跟踪目标。实验结果表明,本文算法实时反映了跟踪目标的真实尺度,同时实现了实时的、快速的跟踪。

### 1 Camshift 算法

Camshift 是一种使用颜色信息的跟踪算法,通过提取

收稿日期:2016-08-31;修回日期:2016-12-23。    基金项目:重庆市教委科学技术研究项目(KJ130512)。

作者简介:胡章芳(1969—),女,重庆人,副教授,硕士,主要研究方向:通信与信息系统; 秦阳鸿(1992—),男,重庆人,硕士研究生,主要研究方向:图像处理、模式识别。



HSV 颜色空间的 H 分量分布图进行跟踪<sup>[6]</sup>。它的算法流程如下：

- 1) 初始化, 手动选择目标, 得到搜索窗 W。
- 2) 计算跟踪目标反向投影图。
- 3) 根据反向投影图和选择框进行 Meanshift 迭代, 找到跟踪目标中心。
- 4) 根据上述搜索窗内的颜色概率分布函数来重新设定搜索窗的大小。
- 5) 使用迭代出的结果作为跟踪目标在下一帧图像中继续跟踪。

## 2 基于图割的匹配算法

图割是基于图的重新组合和优化的思想<sup>[7]</sup>, 将一幅图像中的像素点映射为一个网络图, 并建立关于标号的能量函数<sup>[8]</sup>。由最大流最小割定理可知: 可以通过图像中像素所构造出来的网络最大流来得到图的最小割。首先需要构造成能量函数, 再将像素点代表的权值代入能量函数中构造图像网络才能得到最大流<sup>[9-11]</sup>。

设有向图用  $G = (V, E)$  表示, 其中  $V$  代表图中节点的集合。节点分为两类: 普通节点  $P = \{P_i | i = 1, 2, \dots, n\}$ ; 端节点  $S$  和  $T$ ,  $S$  为前景目标,  $T$  为背景。 $E = \{(u, v) | u \in P, v \in P\}$  代表图中边的集合, 每条边都设有力值  $\omega(u, v)$ 。图割就是将边的权值带入能量函数中通过计算结果使图中节点划分为两部分  $s$  和  $t$ , 而分割结果则需要满足  $s \in S, t \in T, S \cup T = E, S \cap T = \emptyset$ 。这些边组成割集  $S-T$ , 最小割就是割集  $S-T$  中边的权值之和最小的割<sup>[12]</sup>。

要分割出“目标”, 先引入能量函数:

$$E(A) = \sum_{p_i \in P} R_{p_i}(A_{p_i}) + \sum_{(p_i, p_j) \in N} \delta(A_{p_i}, A_{p_j}) \quad (1)$$

其中:  $P$  表示图像中像素点的点集;  $N$  表示  $S$  的邻域系统;  $A_{p_i}$  表示像素  $P_i$  的权值, 即若用 1 来表示属于前景目标, 那么就用 0 来表示背景;  $R_{p_i}(A_{p_i})$  和  $\delta(A_{p_i}, A_{p_j})$  分别表示将元素  $p_i$  和边  $(p_i, p_j)$  标号为目标的权值<sup>[13]</sup>。可以通过求得能量函数最小值来得到最符合真实情况的图像分割即是目标最小割<sup>[14-15]</sup>。于是构造能量函数是图割应用在目标分割上的一个关键步骤。

## 3 改进的人脸跟踪算法

在目标跟踪研究中, Camshift 算法具有很强跟踪性能, 被广泛应用于各种视频跟踪系统中。但是当跟踪目标尺度发生剧烈变化时, 跟踪窗口就可能出现过度放缩, 甚至会丢失跟踪目标。而图割理论在图像分割领域有着较强的分割性能, 可以准确分割出跟踪目标与背景。因此, 在 Camshift 跟踪出的结果中引入图割理论, 能准确分割出 Camshift 跟踪框内真实的目标, 同时计算出该目标真实大小和位置。

### 3.1 目标分割

鉴于全局图割会加大运算量, 本文采用局部图割, 使用 Camshift 迭代出的跟踪窗口  $W$  作为图割区域。这种方法不仅提高了算法效率, 同时防止了跟踪场景中肤色相近区域的干扰。

系统运行 Camshift 算法后, 会产生一个长为  $h_0$ , 宽为  $w_0$  的初始跟踪框, 其跟踪中心为  $(x_c, y_c)$ 。在跟踪中心点纵向扩展  $M$  ( $M < h_0/2$ ) 个、横向扩展  $N$  ( $N \leq w_0/2$ ) 个像素点作为图

割算法分割区域。

跟踪中心  $(x_c, y_c)$  作为源点  $S$ , Camshift 跟踪框内像素点作为有向图中普通节点, 而其他点作为背景节点, 所以图割结果就是跟踪目标和背景的两个割, 如图 1 所示。

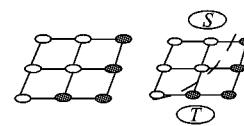


图 1 图割示意图

Fig. 1 Schematic diagram of graph-cuts

为了求取精确肤色模型来表示肤色的分布, 本文使用高斯肤色模型来作为图割算法的分割权值, 公式如下:

$$p(\mathbf{r}) = \sum_{i=1}^K \frac{W_i}{(2\pi)^{n/2} |\Sigma_i|^{1/2}} \cdot \exp\left(-\frac{1}{2} (\mathbf{r} - \boldsymbol{\mu})^T \frac{1}{\Sigma_i} (\mathbf{r} - \boldsymbol{\mu})\right) \quad (2)$$

其中:  $K$  代表高斯混合模型中高斯模型个数;  $W_i$  表示第  $i$  个高斯所占权重值, 所有  $W_i$  值之和是 1;  $\mathbf{r}$  表示高斯分布的方差;  $\boldsymbol{\mu}$  表示高斯分布的均值;  $\Sigma_i$  表示第  $i$  个高斯分布的共变异矩阵。式(2)代表参数  $(\omega, \boldsymbol{\mu}, \Sigma)$  可用  $k$  均值算法来求得, 而得到像素点的肤色判断概率值。

此时把高斯肤色赋值给从点  $S$  到每个普通节点的边  $W(p_i)$  上, 而相邻普通节点之间的权值则为  $W(p_i, p_j)$ 。因此能量函数可以构造如下:

$$E(O) = \sum_{p_i \in P} W(p_i) + \sum_{(p_i, p_j) \in N} W(p_i, p_j) \quad (3)$$

其中:  $P$  表示普通节点的点集;  $N$  表示  $S$  的邻域系统即跟踪区域;  $W(p_i)$  和  $W(p_i, p_j)$  分别表示将元素  $p_i$  和边  $(p_i, p_j)$  划分为目标的代价。计算  $W$  中的每个像素的  $W(p_i)$  和  $W(p_i, p_j)$ 。该网络中  $E(O)$  的最大流既是通割集  $S-T$  中边的权值之和最小的割, 而分割出的肤色团, 也就是跟踪目标真实位置的点集, 如图 2 所示。



图 2 肤色团分割示意图

Fig. 2 Schematic diagram of skin color segmentation

记录该帧搜索框内肤色点的数量, 并与上一帧中肤色点数量比较。当相邻两帧肤色点数差距过大, 即可判断出已丢失跟踪目标, 重新使用该帧运行 Camshift 算法搜索跟踪目标。

### 3.2 尺度提取

为了防止颈部肤色的干扰使尺度偏大, 本文选择肤色团块横坐标计算尺度。

在整个视频区域内建立坐标系, 坐标原点为视频区域中的左上角, 数字图像的列数和行数代表坐标系中 Y 轴和 X 轴。

通过像素点的四领域判断图连通性, 来判断周围像素点是否是肤色点, 并将肤色点集标记为肤色团块。而团块中肤色最多的即为最大团块。

在肤色团内找出中心点处  $(x_c, y_c)$  中  $y_c$  对应的最左侧的像素点的坐标为  $(x_1, y_c)$ , 然后取出  $y_c + i, y_c - i$  ( $i = 1, 2, \dots, h_0$ ) 在肤色团中所对应的  $2h_0$  个横坐标值, 求出  $x_1$  和这  $2h_0$  个



横坐标值的均值赋给  $x_1$ 。同时,肤色团最右边的  $x_2$  也可用此方法得出,于是得出肤色团的长度  $w = x_2 - x_1$ 。由于人脸大小的比例是不变的,本文通过计算手动标记的人脸比例大小  $h_0/w_0$ ,可以计算出肤色团的高度  $h = w(h_0/w_0)$ ,于是得到目标真实尺度。

### 3.3 本文算法基本步骤

本文提出的基于图割的人脸跟踪算法步骤如下:

- 1) 手动选取跟踪目标,得到目标的长为  $h_0$ ,宽为  $w_0$ ;
- 2) 执行 Camshift,计算目标中心  $(x_c, y_c)$ ,同时计算出当前搜索框长为  $h_1$ ,宽为  $w_1$ ;
- 3) 在搜索框区域内,由跟踪中心  $(x_c, y_c)$  上下扩展位置作为图割区域;
- 4) 比较当前帧与上一帧肤色点数量,若数量相差大于上一帧肤色点数量的  $1/4$ ,则判断丢失跟踪目标,重新执行 Camshift 搜索跟踪目标;
- 5) 当肤色点对比结果小于  $1/4$ ,则计算肤色团的长  $w =$



图3 三种算法下人脸跟踪情况

Fig. 3 Human face tracking by three algorithms

图4显示了三种算法来跟踪人脸的性能图。中心误差是跟踪结果与真实位置之间的欧氏距离,其值越小说明算法尺度控制性能越好。从图4可以看出,基于图割的 Camshift 算法在运行过程中,跟踪的结果与目标真实坐标之间的差距较小,其尺度控制的性能较为稳定;而 Camshift 算法在控制方面性能相对低,当视频中出现手的干扰时,跟踪框远远大于人脸真实尺度;而融合粒子滤波的 Camshift 算法相较本文算法在跟踪目标时的跟踪框与真实位置之间的距离也大于本文算法,因此本文算法能有效地反映人体快速运动中人脸真实尺度变化。

本文的基于图割的 Camshift 跟踪算法,由于采用了局部图割,计算量很小,跟踪目标单帧用时 0.121 s;融合粒子滤波的 Camshift 算法单帧用时相对较长,需要 0.231 s;而 Camshift 算法计算时间虽然比本文算法短,仅为 0.023 s,但是跟踪性能远不如本文算法。

$$x_2 - x_1, \text{宽 } h = w(h_0/w_0);$$

- 6) 得出当前跟踪目标真实尺度;
- 7) 继续读取下一帧,执行步骤 2) ~ 5)。

## 4 实验与分析

本实验在 Core i3-3220@3.30 GHz,内存 4.00 GB 的计算机上采用C++语言编程实现。实验采用的视频序列分辨率为  $320 \times 240$ 。实验比较 Camshift 算法、融合粒子滤波的 Camshift 算法,以及本文算法的跟踪效率和性能。

从图3所示的跟踪情况可看出不同算法的性能:图3(a)中在第46帧出现了过度放缩,算法跟踪了脸部以下的区域;在第99帧出现了手部干扰,使跟踪框放大到手部区域。图3(b)中在第112帧处发生了丢失跟踪目标。图3(c)实时地反映了目标的真实尺度,同时也避免了手出现在跟踪场景中的干扰。



度过度放缩这一问题所影响。本文提出的基于图割的 Camshift 跟踪算法根据图割出的目标来计算尺度,从而得到目标的真实尺度,并且保证了算法的运算速度。从实验结果可知:基于图割的 Camshift 跟踪算法能有效地反映跟踪目标真实的尺度变化。

#### 参考文献(References)

- [1] HU Z, WANG Y, LUO Y. An improved camshift-based particle filter algorithm for real-time hand gesture tracking [J]. Sensors & Transducers, 2014, 177(8):307–312.
- [2] 吕卓纹, 王科俊, 李宏宇, 等. 融合 Camshift 的在线 Adaboost 目标跟踪算法[J]. 中南大学学报(自然科学版), 2013, 44(7): 231–238. (LYU Z W, WANG K J, LI H Y, et al. Online AdaBoost target traking algorithm combined fused with Camshift [J]. Journal of Central South University ( Science and Technology ), 2013, 44(7):231 – 238.)
- [3] LI J, ZHANG J, ZHOU Z, et al. Object tracking using improved Camshift with SURF method[ C]// Proceedings of the 2011 IEEE International Workshop on Open-source Software for Scientific Computation. Piscataway, NJ: IEEE, 2011:136 – 141.
- [4] 徐建军, 张蓉, 毕笃彦, 等. 一种新的 AdaBoost 视频跟踪算法 [J]. 控制与决策, 2012, 27(5):681 – 685. (XU J J, ZHANG R, BI D Y, et al. An new AdaBoost video tracking algorithm[ J]. Control and Decision, 2012, 27(5):681 – 685.)
- [5] 王鑫, 唐振民. 一种改进的基于 Camshift 的粒子滤波实时目标跟踪算法[J]. 中国图象图形学报, 2010, 15(10): 1507 – 1514. (WANG X, TANG Z M. An improved camshaft based particle filter algorithm, for real-time target traking [ J ]. Journal of Image and Graphics, 2010, 15(10): 1507 – 1514.)
- [6] 厉丹, 田隽, 肖理庆, 等. 基于 Camshift 自适应多特征模板的视频目标跟踪[J]. 煤炭学报, 2013, 38(7): 1299 – 1304. (LI D, TIAN X, XIAO L Q, et al. Adaptive multi-feature template video target tracking based on Camshift algorithm[ J ]. Journal of China Coal Society, 2013, 38(7): 1299 – 1304.)
- [7] KOLMOGOROV V, BOYKOV Y, ROTHER C. Applications of parametric maxflow in computer vision[ C]// ICCV 2007: Proceedings of the 2007 IEEE 11th International Conference on Computer Vision. Piscataway, NJ: IEEE, 2007:1 – 8.
- [8] PRICE B L, MORSE B, COHEN S. Geodesic graph cut for interactive image segmentation[ C]// Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2010: 3161 – 3168.
- [9] 姜伟, 吕晓琪, 任晓颖, 等. 结合区域生长与图割算法的冠状动脉 CT 血管造影图像三维分割[J]. 计算机应用, 2015, 35(5): 1462 – 1466. (JIANG W, LYU X Q, REN X Y, et al. 3D segmentation method combining region growing and graph cut for coronary arteries computed tomography angiography image [ J ]. Journal of Computer Applications, 2015, 35(5):1462 – 1466.)
- [10] 张少娟, 邹建成. 图割综述[J]. 北方工业大学学报, 2010, 22(3): 10 – 16. (ZHANG S J, ZOU J C. Review on graph cuts [ J ]. Journal of North China University of Technology, 2010, 22(3):10 – 16. )
- [11] 张凤军, 赵岭, 安国成, 等. 一种尺度自适应的 Mean-Shift 跟踪算法[J]. 计算机研究与发展, 2014, 51(1): 215 – 224. (ZHANG F J, ZHAO L, AN G C, et al. Mean-Shift tracking algorithm with scale adaptation[ J ]. Journal of Computer Research and Development, 2014, 51(1):215 – 224.)
- [12] JUAN O, BOYKOV Y. Active graph cuts[ C]// Proceedings of the 2006 IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2006: 1023 – 1029.
- [13] STAWIASKI J, DECENCIERE E. Region merging via graph-cuts [ J ]. Image Analysis & Stereology, 2008, 27(1):39 – 45.
- [14] COOKE T. Two applications of graph-cuts to image processing [ C]// DICTA 2008: Proceedings of the 2008 Digital Image Computing: Techniques and Applications. Washington, DC: IEEE Computer Society, 2008: 498 – 504.
- [15] 段丁娜, 张欢, 邱陈辉, 等. 基于活动轮廓模型的图像分割算法综述[J]. 中国生物医学工程学报, 2015, 34(4):445 – 454. (DUAN D N, ZHANG H, QIU C H, et al. A review of active contour model based image segmentation algorithms[ J ]. Chinese Journal of Biomedical Engineering, 2015, 34(4):445 – 454.)

This work is partially supported by the Scientific and Technological Research Program of Chongqing Municipal Education Commission (KJ130512).

**HU Zhangfang**, born in 1967, M. S., associate professor. Her research interests include communication and information system.

**QIN Yanghong**, born in 1992, M. S. candidate. His research interests include image processing, pattern recognition.

**LIU Wu**, born in 1978, M. S. His research interests include electric power safety monitoring.

**ZHANG Ji**, born in 1982, M. S. His research interests include electric power video surveillance.

**ZHANG Lingchen**, born in 1989, M. S. His research interests include signal processing, pattern recognition.

**YUAN Meigui**, born in 1991, M. S. Her research interests include signal processing, pattern recognition.

**QU Lei**, born in 1980, Ph. D., professor. His research interests include signal processing, pattern recognition.

(上接第 1188 页)

- [13] SCHOLKOPFT B, MULLERT K R. Fisher discriminant analysis with kernels[ EB/OL]. [2016-03-01]. [http://courses.cs.tamu.edu/rgettler/csce666\\_f16/mika1999kernelLDA.pdf](http://courses.cs.tamu.edu/rgettler/csce666_f16/mika1999kernelLDA.pdf).

This work is partially supported by National Natural Science Foundation of China (61201396, 61301296), the National Science and Technology Support Program (2014BAH27F01), the Science and Technology Project of China State Grid Corporation (5212D01502DB).

**FANG Yiguang**, born in 1969, M. S., senior engineer. His research interests include electric power safety monitoring.