

基于自适应活动轮廓模型的实时手势跟踪^{*}

齐苏敏^{1,2} 黄贤武¹ 孟 静¹

(苏州大学电子信息学院 苏州 215021)¹ (曲阜师范大学计算机科学学院 曲阜 273165)²

摘要 在基于视觉的手势分析与识别中,一个关键环节是手势跟踪。本文提出了基于颜色信息的自适应活动轮廓模型,并与均值漂移算法相互融合,实现图像序列的实时手势跟踪。跟踪算法分为两步进行,首先应用均值漂移算法实现手部区域的定位,然后基于自适应活动轮廓模型提取手部轮廓。在跟踪过程中,轮廓提取为下一帧的区域定位更新搜索窗口,提高了搜索效率,使目标跟踪达到实时性要求。同时,本文根据跟踪区域模板与目标模板的相似性度量 Bhattacharyya 系数给出了在跟踪目标被遮挡时的处理方法,有效地解决了这一难题。实验结果证明了在无遮挡和遮挡两种情况下算法均能实现准确、实时的手势跟踪。

关键词 手势跟踪,活动轮廓,均值漂移,遮挡

Real Time Hand Tracking Based on Adaptive Active Contour Model

QI Su-Min^{1,2} HUANG Xian-Wu¹ MENG Jing¹

(School of Electronics & Information Engineering, Soochow University, Suzhou 215021)¹

(School of Computer Science, Qufu Normal University, Qufu 273165)²

Abstract Hand tracking is an essential step for vision based gesture analysis and recognition. This paper presents an adaptive active contour model using color information, which is connected with mean shift algorithm to implement real time hand tracking in sequences. The proposed method consists of two steps: hand location using mean shift and hand extraction based on adaptive active contour model. In the process of tracking shape extraction alters the search window for hand location of next frame, which improves searching effective and makes tracking real-time. At the same time, this paper gives tracking methods in terms of the similarity measure of candidate modal and object modal i. e. Bhattacharyya coefficient while the object is occluded. Experimental results show that accurate and real-time tracking is achieved using the proposed algorithm either on the occasion of occlusion or not.

Keywords Hand tracking, Active contour, Mean shift, Occlusion

1 引言

随着计算机应用的深入,并伴随着计算机及相关学科的迅猛发展,研究符合人际交流习惯的新颖人机交互技术(HCI)变得异常活跃,并取得了一定的进展。总体来说,人机交互技术已经从以计算机为中心逐步转移到以人为中心,手势识别也逐渐成为一个研究热点^[1]。

手势识别技术按照实现途径可分为基于数据手套的手势识别和基于计算机视觉的手势识别两大类^[2]。基于视觉的手势识别是利用摄像机采集手势信息,并进行识别。该方法不需要复杂的数据手套作为输入设备,输入相对简单、方便,但识别率比较底,实时性较差,故提高识别率与实时性是其关键问题。同时在基于视觉的手势分析与识别中,手势跟踪是一个关键环节。本文研究了基于单目视觉的手势跟踪,在图像序列中实现手部区域的定位并提取手部轮廓信息。

活动轮廓模型(Snake)是1987年Kass^[3]等人提出用来处理刚性物体或非刚性物体的一种有效的分割与跟踪工具,分割与跟踪的过程是通过能量最小化的原则进行的。近年来,该模型被应用于目标边缘的检测与运动目标的跟踪,但Snake模型存在初始化轮廓问题,并对图像噪声非常敏感,不能解决快速运动的目标跟踪。本文提出一种根据目标特征自适应收缩与扩张的自适应活动轮廓模型,减小了对初始轮廓的敏感性,实现了自适应地提取手部轮廓信息。

本文提出的手势跟踪算法采用自适应的活动轮廓和均值

漂移算法。算法分为两部分,手部区域的定位与手部轮廓的提取^[4],均以肤色信息作为先验知识。首先,利用均值漂移算法实现手部区域的定位;然后,在定位处利用活动轮廓模型自适应收缩与扩张,提取手部轮廓。

2 手部区域的定位

2.1 肤色信息的获取

肤色信息可通过颜色信息^[2]获得,可通过对颜色信息进行限制,例如: $R>G>B$,但这个颜色范围还包含了很多其它颜色如红色,粉红,以及桔红色;也可从手部区域获得采样颜色信息,以此采样信息与其它颜色区域相比较,相似性较高的便认定为肤色信息。本文采用第二种方法获取肤色信息,以归一化的直方图^[5]来表示。将彩色特征空间量化为 M 级,则跟踪目标模板为:

$$q = \{q_u\}_{u=1}^M, \text{ 满足 } \sum_u q_u = 1 \quad (1)$$

其中 q_u 表示量化等级为 u 的颜色信息。实验结果证明,第二种方法获取的肤色信息更准确。

2.2 基于均值漂移算法实现手部区域的定位

均值漂移算法^[5,6]是一种有效的无参数统计迭代算法,它是使每一个点“漂移”到密度函数的局部极大值点。近年来,均值漂移算法已广泛应用于计算机视觉领域,如:跟踪、图像分割、图像平滑、聚类分析等方面。

基于均值漂移算法^[5]实现手部区域定位的步骤如下:

1) 在当前帧中初始化搜索窗口为 $W(y_0, h_0)$ (y_0 为初始

^{*}国家自然科学基金(30300088)。齐苏敏 博士生,讲师;黄贤武 博士生导师,教授;孟 静 博士生。

定位点, h_0 为窗宽), 并在提取此窗口内肤色概率分布, 称之为候选模板, 其定义为:

$$p(y_0) = \{p_u(y_0)\}_{u=1}^M, \text{ 且满足 } \sum_u p_u = 1, \quad (1)$$

$$p_u(y_0) = c_{h_0} \sum_{i=1}^{n_{h_0}} k(\| \frac{y_0 - x_i}{h_0} \|^2) \delta[b(x_i) - u] \quad (2)$$

其中 $k(\cdot)$ 是文[6]中定义的轮廓函数, $b(\cdot)$ 为映射函数, $b(x_i)$ 等于 x_i 点的量化特征值, n 为窗口 W 内的像素个数, c_{h_0} 为归一化常数。

2) 用 Bhattacharyya 系数 ρ 来度量 q 和 $p(y_0)$ 分布密度的相似性程度,

$$\rho(y_0) = \sum_{u=1}^M [q_u p_u(y_0)]^{1/2} \quad (3)$$

在 y_0 处对 $\rho(y_0)$ 进行一阶泰勒展开[5], 可得:

$$\rho(y_0) \approx c_1 + c_2 \sum_{i=1}^{n_{h_0}} \omega_i k(\| \frac{y_0 - x_i}{h_0} \|^2) \quad (4)$$

$$\omega_i = \sum_{u=1}^M [\frac{q_u}{p_u(y_0)}]^{1/2} \delta[b(x_i) - u] \quad (5)$$

其中 c_1, c_2 均为常数。根据 Cheng[6], Comaniciu[5] 等提出的均值漂移算法, 则得:

$$y_1 = \frac{\sum_{i=1}^{n_{h_0}} x_i k'(\| \frac{y_0 - x_i}{h_0} \|^2) \omega_i}{\sum_{i=1}^{n_{h_0}} k'(\| \frac{y_0 - x_i}{h_0} \|^2) \omega_i} \quad (6)$$

其中, k' 是 k 导数。

3) 判断是否满足 $\|y_1 - y_0\| > \text{thld1}$, 若是则迭代结束, 否则令 $y_0 \leftarrow y_1$, 执行 2)。其中阈值 thld1 为接近于零的较小值。

由上述算法可获得新的区域定位点 y_1 和均值漂移向量 $\overrightarrow{\text{MSV}} = y_1 - y_0$ 。

3 手部轮廓的提取

3.1 基于肤色信息的自适应轮廓模型

近年来, 活动轮廓模型(snake)被应用于目标边缘的检测与运动目标的跟踪[7]。计算机视觉跟踪就是通过序列图像对跟踪目标边界连续分割。Snake 分割过程是通过能量最小化的原则进行的, 找到能量最小的表面。其能量函数(离散状态)[8]为:

$$E_{total} = \sum_{i=1}^n (E_{int}(i) + E_{ext}(i))$$

$$= \sum_{i=1}^n [\alpha E_{con}(i) + \beta E_{cur}(i) + \gamma E_{img}(i)] \quad (7)$$

式(7)中的 Snake 模型, 在内部能量 E_{int} 的控制下, 轮廓点由初始轮廓向中心点收缩, 在外部能量 E_{ext} 影响下吸引轮廓点停留在目标边界。但式(7)中的模型依赖于图像中细微的变化, 存在初始化轮廓问题, 并对图像噪声非常敏感, 不能解决快速运动的目标跟踪。Cohen 提出了一种可膨胀的轮廓线[9], 降低了对轮廓线初始化的敏感性。Won 等人提出了 Snake 的跳跃模型[8]来解决快速运动目标的跟踪问题。当序列图像中连续两帧图像不存在目标重叠现象或重叠不多时, Snake 的跳跃模型也可以用来跟踪物体, 这种 Snake 的跳跃模型理论是假设在图像流的处理中能够获得运动方向的基础上的, 从先前每一帧图像中获得的 Snake 的节点跳跃到目标的边界, 在确定目标分割半径的基础上, 位移到另一个区域

并且重新初始化。

基于式(1)的目标模板, 本文提出一种改进的自适应活动轮廓模型, 使轮廓线能够自适应地收缩或膨胀, 在视觉跟踪中不需在当前帧中重新初始化, 只需进行目标定位即可, 其能量函数为:

$$E_{total} = \sum_{i=1}^n [\alpha E_{con}(i) + \beta E_{cur}(i) + \gamma E_{img}(i) + \kappa E_{rad}(i)] \quad (8)$$

$E_{con}(i), E_{cur}$ 及 $E_{img}(i)$ 如文[3, 8]定义, $E_{rad}(i)$ 表示点 v_i 收缩或膨胀的能量函数, 定义如下:

$$\frac{\partial E_{rad}(i)}{\partial t} = \text{sign}(v_i) \vec{d} \quad (9)$$

$$\text{sign}(v_i) = \begin{cases} 1 & q_{b(v_i)} > \text{thld2} \\ -1 & \text{其它} \end{cases} \quad (10)$$

其中 $\vec{d} = v_i - v_c$, v_c 为轮廓区域的中心点, $b(v_i)$ 如式(2)中定义, $q_{b(v_i)} \in q$ 表示点 v_i 的颜色信息在目标模板 q 中的值, 阈值 thld2 为接近于零的较小值。在式(9)的作用下, $E_{rad}(i)$ 能量函数沿着向量 \vec{d} 的方向随着跟踪目标的颜色特征变化, 使轮廓线自适应地放缩, 降低了对初始轮廓的敏感性。式(8)定义的自适应模型较好地解决了 Snake 模型(式(7))和 Won 等人提出的 Snake 跳跃模型的不足, 保证了视觉跟踪中对目标轮廓的准确提取。

3.2 用自适应活动轮廓模型提取手部轮廓

手部轮廓提取的目的是能够找到分割手部区域与背景的闭合曲线, 应用式(8)的自适应活动轮廓可较好地完成这一过程。活动轮廓模型的实现算法有多种, 如 Kass 算法, Dynamic programming 算法, Greedy 算法等, 由于 Greedy 算法[10]计算复杂度低, 实现简洁, 因此本文采用 Greedy 算法, 具体步骤略。

假设初始化轮廓点为 m 个, 表示为 v_1, v_2, \dots, v_m , 则其轮廓区域中心点 v_c 表示为:

$$v_c.x = M_{10}/M_{00}; v_c.y = M_{01}/M_{00} \quad (11)$$

其中 M_{ab} 是轮廓线(Contour)的 $(a+b)$ 阶矩。由于自适应轮廓模型能够自适应地收缩与扩张, 初始轮廓的选择可在物体内部, 也可在物体外部, 而且距离物体可远可近, 有较强的鲁棒性。对图 1 所示的手部轮廓分别应用式(7)与式(8)进行轮廓提取, 其结果分别如图 1(黑线代表初始轮廓, 白线代表曲线演化后轮廓)所示, 图 1(a)~(c)为文献应用式(7)轮廓提取结果, 图 1(d)~(f)为应用式(8)轮廓提取结果, 参数设置为: $\alpha = \beta = \gamma = \kappa = 1$ 。

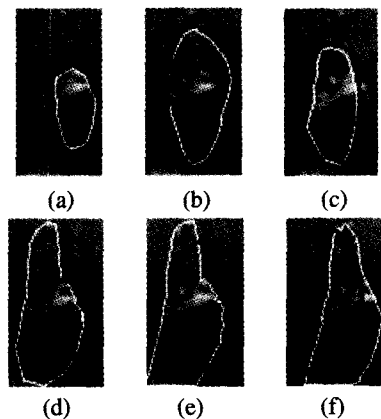


图 1 hand1 轮廓提取比较

4 实时手势跟踪算法

4.1 算法实现

融合自适应活动轮廓和均值漂移算法实现图像序列的实时手势跟踪,算法实现步骤如下:

1) 对肤色信息采样,得到式(1),并初始化轮廓为 $\Gamma_0(u_c)$,其中 u_c 为式(11)所得轮廓中心点。

2) 在跟踪首帧中,根据式(8)的自适应轮廓模型执行 Greedy 算法,得最终轮廓 $\Gamma_1(v'_c)$,其中 v'_c 为式(11)所得新的轮廓中心点;同时取得其外接窗口 $W(y_0, h_0)$,其中 y_0 为窗口中心, h_0 为窗宽。

3) 在跟踪当前帧中,执行下面两步操作:

① 由前一帧所得窗口 $W(y_0, h_0)$ 作为初始跟踪窗口,执行 2.2 节的均值漂移算法,获得新的跟踪窗口 $W(y_1, h_0)$ 。

② 平移前一帧所得轮廓 $\Gamma_1(v'_c)$ 至点 y_1 , 表示为 $\Gamma_1(y_1)$, 基于式(8)的自适应轮廓模型得新的轮廓 $\Gamma'_1(v'_c)$ 和 Γ'_1 外接窗口 $W(y_1, h_1)$ 。

4) 令 $y_0 \leftarrow y_1, h_0 \leftarrow h_1, \Gamma_1 \leftarrow \Gamma'_1$, 获取下一帧执行 3) 直到序列结束。

在上述算法中,步骤 3) 是关键,均值漂移算法 (Mean Shift 简称 MS) 与自适应活动轮廓 (Adaptive Active Contour 简称 AAC) 相互融合,均值漂移算法的准确定位为手部轮廓的提取提高了准确性,而轮廓准确提取的同时更新了跟踪窗口,使跟踪窗口的大小随跟踪目标的变形而变化,提高了均值漂移算法的搜索效率,二者相辅相成,实现序列图像中变形物体(手)的实时跟踪。它们的相互关系如图 2 所示。

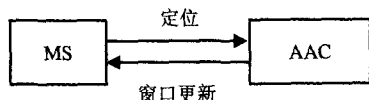


图 2 算法关系

4.2 跟踪算法对遮挡的处理

在目标跟踪中,均值漂移算法对物体的部分遮挡有较强的鲁棒性,但活动轮廓模型对遮挡非常敏感,所以当跟踪目标被其它物体遮挡时,需要进行判断与处理,以减轻或消除遮挡对跟踪效果的影响。两个分布密度的相似性程度可用这两个分布函数之间的 Bhattacharyya 系数 ρ 来度量^[5], ρ 越大,则表明相似性越大。利用这一特点对遮挡进行判断与处理,具体做法如下:

1) 对遮挡进行判断:在当前帧中,应用式(3)求得最终跟

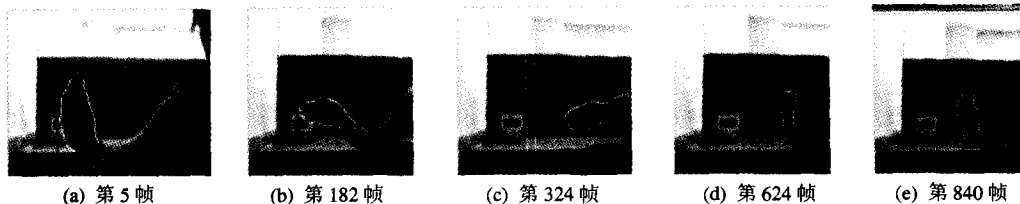


图 4 hand2 序列手势跟踪效果

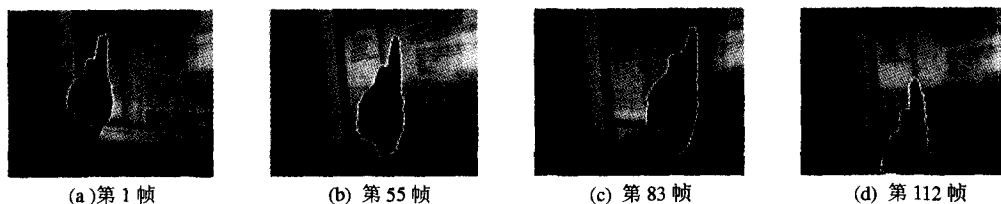


图 5 hand3 序列手势跟踪效果

踪窗口 $W(y_1, h_1)$ 内概率分布 $p(y_1)$ 与式(1)中 q 的 Bhattacharyya 系数 ρ 。设定当 $\rho < \text{thld3}$ 时,出现了轻微遮挡;当 $\rho < \text{thld4}$, 出现严重遮挡。其中阈值 thld4 约为 0.5 的实值,并且满足 $\text{thld4} < \text{thld3} < 1$ 。

2) 出现遮挡后对各种参数的重新设置:当出现轻微遮挡时,则采用均值漂移算法进行区域定位至点 y_1 ,但在当前帧中保持前一帧的提取轮廓 Γ'_1 不变,只平移到点 y_1 ,则此时 $h_1 = h_0$;当出现严重遮挡时,则会出现不准确定位,此时定位点 y_1 由跟踪目标的移动速度近似获得。

图 3 表示在一个出现遮挡的序列中, Bhattacharyya 系数的变化情况,可以看出图中出现了较大波谷,反映了 1) 中所示的两种遮挡情况。

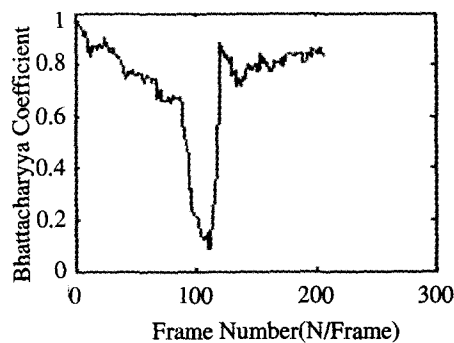


图 3 一个遮挡的序列中巴士系数的变化情况

5 实验结果与分析

实验中,在序列图像的 HSV 空间中实现算法,针对不同序列在不同的情况下进行跟踪,均取得了不错的跟踪效果。总的来说,可按无遮挡和遮挡两种情况分析。

5.1 无遮挡的图像序列的手势跟踪效果

在 hand2 序列中手的位置和形状快速变化,共有 920 帧,帧率为 25HZ。图 4 显示第 5、182、324、624、840 帧的跟踪效果,可以看出算法能够实时、准确地跟踪其轮廓,其中黑色窗口为最终跟踪窗口,白线代表手部轮廓。图 5 显示了由于显示器的快速闪动造成光线变化较大的情况下手势跟踪的效果,表明算法对光线的变化有较强的鲁棒性。本文的算法对凸性和轻微凹性的手部轮廓均有较好的效果,但对严重凹性的手部轮廓不能达到准确的跟踪,图 6 表明此种情况的跟踪效果,这需在今后的研究工作中努力克服。

cation based on maximum model distance algorithm. In: ICASSP, 2004. 125~128

- 4 Barron J L, Fleet D J, Beauchemin S S. Performance of optical flow techniques. IJCV, 1994, 12(1): 43~77
- 5 Hong Hai, Neven H, von der Malsburg C. Online Facial Express-

sion Recognition Based on Personalized Galleries. In: Proceedings of Third International Conference on Automatic Face and Gesture Recognition, April 1998

(上接 194 页)



图 6 hand4 序列手势跟踪效果

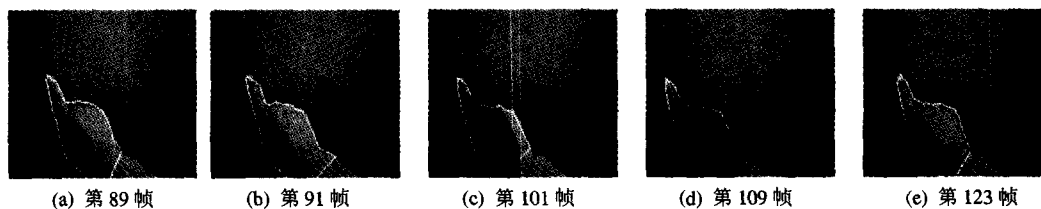


图 7 hand5 序列手势跟踪效果

5.2 出现遮挡的图像序列的手势跟踪效果

在图像序列中,当出现遮挡时,应用 4.2 节的算法进行处理,图 7 显示了这一处理效果,此时设 $thld3=0.7$, $thld4=0.5$ 。在第 89 帧中,手部区域将被其它物体遮挡,此时 Bhattacharyya 系数为 0.7045,此时不需微遮挡处理;在第 91 帧中, Bhattacharyya 系数为 0.5809,按 4.2 节第二步骤的轻微遮挡处理;在第 101 帧中, Bhattacharyya 系数为 0.2632,按 4.2 节第二步骤的严重遮挡处理;对于第 109 帧与 101 帧进行同样的处理;在第 123 帧中, Bhattacharyya 系数为 0.7896,大于 0.7,遮挡消除。实验表明该处理方法能克服运动目标的遮挡问题。

结束语 本文提出基于颜色特征的自适应活动轮廓模型,克服了以往活动轮廓模型对初始轮廓过于敏感的缺点,并与均值漂移相互融合实现手势跟踪算法,该算法首先利用均值漂移算法实现手部区域定位,然后基于自适应活动模型提取准确的手部轮廓,二者相辅相成实现手势跟踪。本文提出跟踪过程中对遮挡的处理方法,较好地解决了目标跟踪中遮挡这一难题。实验结果表明在图像序列中目标跟踪实时,轮廓跟踪准确,尤其是光线变化较大和出现遮挡时,仍能达到较好的跟踪效果。不过本文的算法对严重凹性的手部轮廓不能达到准确的跟踪效果,在今后的研究工作中应该努力克服。

参考文献

- 1 Shan C, Wei Y, Tan T, et al. Real time hand tracking by combining particle filtering and mean shift [C]. In: Proceedings Sixth IEEE International Conference on Automatic Face and Gesture

Recognition, Seoul, Korea, 2004. 669~674

- 2 Chen F S, Fu C M, Huang C L. Hand gesture recognition using a real-time tracking method and hidden Markov models [J]. Image and Vision Computing, 2003, 21(8): 745~758
- 3 Kass M, Witkin A, Terzopoulos D. Snakes: Active contour models [C]. In: Proceedings of First International Conference on Computer Vision, London, 1987. 259~269
- 4 Chang J S, Kim E Y, Jung K C, et al. Real time hand tracking based on active contour model [C]. In: Computational Science and Its Applications - ICCSA 2005, Proceedings, Part IV (Lecture Notes in Computer Science Vol. 3483), Singapore, 2005. 999~1006
- 5 Comaniciu D, Ramesh V, Meer P. Kernel-based object tracking [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(5): 564~577
- 6 Cheng Yizong. Mean shift, mode seeking, and clustering [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995, 17(8): 790~799
- 7 周志宇, 汪亚明, 黄文清. 基于动态图像序列的运动目标跟踪 [J]. 浙江工程学院学报, 2002, 19(3): 165~170
- 8 Kim W, Lee C Y, Lee J J. Tracking moving object using Snake's jump based on image flow [J]. Mechatronics, 2001, 11(2): 119~216
- 9 Cohen L D. On active contour models and balloons [J]. Computer Vision, Graphics and Image Processing (CVGIP); Image Understanding, 1991, 53(2): 211~218
- 10 Williams D J, Shah M. A fast algorithm for active contours and curvature estimation [J]. CVGIP; Image Understanding, 1992, 55(1): 14~26