

文章编号:1673-5005(2007)04-0153-05

基于均值漂移算法的人脸自适应跟踪

刘宝弟, 王延江, 马 海, 王武礼

(中国石油大学 信息与控制工程学院, 山东 东营 257061)

摘要:采用自适应人脸方向模板和 YCbCr 自适应肤色模型, 提出了一种新的基于均值漂移算法的自适应人脸跟踪方法。与传统的均值漂移跟踪方法相比, 当人脸倾斜时或光线变化时, 该方法能更精确地描绘出人脸位置。实验结果表明, 在基本上不增加计算量的情况下, 该方法能对人脸的倾斜和光线的变化进行很好地自适应跟踪。

关键词:均值漂移算法; 自适应人脸倾角模板; 人脸跟踪

中图分类号:TP 391.41

文献标识码:A

Adaptive human face tracking based on mean shift algorithm

LIU Bao-di, WANG Yan-jiang, MA Hai, WANG Wu-li

(College of Information and Control Engineering in China University of Petroleum, Dongying 257061, Shandong Province, China)

Abstract: By using an adaptive facial orientation template and an adaptive YCbCr skin color model, a novel adaptive human face tracking method was proposed on the basis of mean shift algorithm. Compared with the traditional mean shift tracking method, the method can describe the face position more accurately when facial orientation changes or lighting changes. The experimental results show that the method can adapt to the facial orientation and lighting variations without increasing calculated amount.

Key words: mean shift algorithm; adaptive facial orientation template; human face tracking

均值漂移(mean shift)算法是一种有效的统计迭代算法,最早是由 K Fukunaga 等^[1]提出来的;Yizong Cheng^[2]对其进行了推广,并将其用于聚类分析;D Comaniciu 等^[3]首次将基于核的均值漂移方法应用于目标跟踪,而 Lertsudwichai 等^[4,5]将均值漂移算法用于人脸跟踪,并采用 YCbCr 颜色空间中的 C_b 和 C_r 两个分量作为颜色直方图特征,减小了光线变化对跟踪的影响。Jingping Jia 等^[6]提出了角度自适应的方法对目标的旋转运动进行跟踪,但计算量增大。笔者采用 YCbCr 颜色模型中的 C_b 、 C_r 分量对人脸肤色进行建模,提出一种新的角度自适应方法。

1 均值漂移算法

1.1 核函数的选取

对于人脸区域,由于外围的像素不可靠,很容易

受遮挡或背景的影响,离中心越远,可靠性就越差,所以选用单调递减的核函数来增加跟踪器的鲁棒性。在均值漂移中,有三类核函数(图1)经常用到。

(1) 单位均匀核函数

$$F(x) = \begin{cases} 1, & \|x\| < 1; \\ 0, & \|x\| \geq 1. \end{cases} \quad (1)$$

(2) 单位高斯核函数

$$N(x) = e^{-\|x\|^2}. \quad (2)$$

(3) Epanechnikov 核函数

$$K(x) = \begin{cases} \frac{1}{2}c_d^{-1}(d+2)(1-\|x\|^2), & \|x\|^2 \leq 1; \\ 0, & \text{其他.} \end{cases} \quad (3)$$

式中, c_d 为 d 维单位圆的体积。

本文中使用 Epanechnikov 核函数,因为其导数是均匀核函数,可以简化均值漂移向量。对核函数进

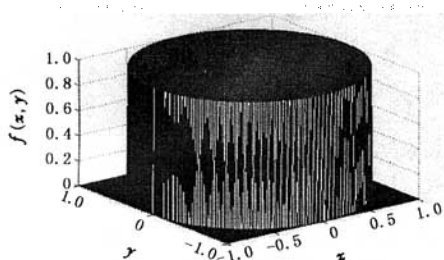
收稿日期:2006-10-10

基金项目:中国石油大学博士基金

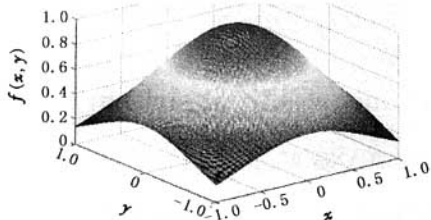
作者简介:刘宝弟(1984-),男(汉族),山东梁山人,硕士研究生,从事信号与信息处理及模式识别研究。

行简化,即

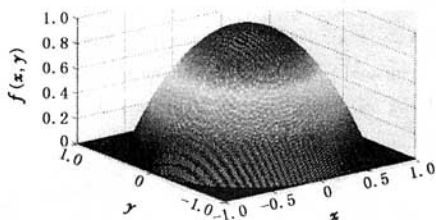
$$K(x) = \begin{cases} 1 - (\|x\|^2), & \|x\|^2 \leq 1; \\ 0, & \text{其他.} \end{cases} \quad (4)$$



(a) 单位均匀核函数



(b) 单位高斯核函数



(c) Epanechnikov 核函数

图1 常用核函数

1.2 基于核的均值漂移算法

1.2.1 人脸区域表示

为了描述人脸区域,首先要选择一个特征空间,特征空间可以选择颜色、纹理或边缘信息,本文中选用颜色概率密度函数。人脸模板可由它的概率密度函数 q 来表示,假设后续帧中人脸候选区域的位置是 y ,并用概率密度函数 $p(y)$ 来表示。同时为了减少计算量,使用了含 m 个直方图特征的直方图^[7]。

人脸模板颜色概率密度可以由下式计算:

$$q_u = C \sum_{i=1}^n K(x_i^*) \delta[b(x_i^*) - u], \quad (5)$$

其中

$$C = \frac{1}{\sum_{i=1}^n K(x_i^*)}.$$

式中, δ 为单位冲击响应;函数 $b: \mathbf{R}^2 \rightarrow \{1, \dots, m\}$ 表示位置为 x_i^* 的像素在直方图中的颜色索引。

人脸候选区域概率密度函数可以由下式计算:

$$p(y) = \{p_u(y)\}_{u=1, \dots, m} = C_h \sum_{i=1}^{n_h} K(x_i) \delta[b(x_i) - u]. \quad (6)$$

其中

$$C_h = \frac{1}{\sum_{i=1}^{n_h} K(x_i)}.$$

两者的相关程度用相似度函数 $\rho(y)$ 表示,即

$$\rho(y) = \rho[p(y), q] = \sum_{u=1}^m \sqrt{p_u(y) q_u}. \quad (7)$$

它的局部最大值表示当前帧中的人脸区域与人脸模板最相似。

1.2.2 人脸定位

为了找到人脸在当前帧中的位置,就需要找到最大的 Battacharyya 相关系数。要搜寻当前帧中人脸新的位置,首先要在上一帧中的 y_0 处开始。这样首先要计算当前帧 y_0 位置的概率密度函数 $\{p_u(y_0)\}_{u=1, \dots, m}$, 然后使用泰勒公式在 $p_u(y_0)$ 附近对式(7)展开,可得

$$\rho(y) \approx \frac{1}{2} \sum_{u=1}^m \sqrt{p_u(y_0) q_u} + \frac{1}{2} \sum_{u=1}^m p_u(y) \sqrt{\frac{q_u}{p_u(y_0)}}. \quad (8)$$

将式(6)代入式(8),可得

$$\rho(y) \approx \frac{1}{2} \sum_{u=1}^m \sqrt{p_u(y_0) q_u} + \frac{C_h}{2} \sum_{i=1}^{n_h} w_i K(x_i), \quad (9)$$

其中

$$w_i = \sum_{u=1}^m \delta[b(x_i) - u] \sqrt{\frac{q_u}{p_u(y_0)}}. \quad (10)$$

为了将式(7)最小化,由于式(9)第一项与 y 无关,就需要找到其第二项的最大值。第二项表示当前帧中含核函数的密度估计。可以通过迭代均值漂移过程找到最大值。根据

$$y_1 = \frac{\sum_{i=1}^{n_h} x_i w_i G(x_i)}{\sum_{i=1}^{n_h} w_i G(x_i)} = \frac{\sum_{i=1}^{n_h} x_i w_i}{\sum_{i=1}^{n_h} w_i} \quad (11)$$

获得新的位置 y_1 。其中 $G(x)$ 是一个新的核函数,其轮廓函数是 $g(x) (g(x) = -k'(x))$ 。

2 自适应人脸跟踪及其实现

2.1 自适应肤色模型

颜色是光谱中可视频段在人眼视网膜形成的感知现象,它是表征物体的重要特征之一。由于颜色与物体的大小、伸缩及姿态基本上无关,并且,肤色处

理速度很快,因此人的皮肤颜色可以作为实时人脸跟踪的一个重要特征。目前,在人脸检测与跟踪中,常用到的颜色空间有 RGB、HSV^[8]、YCbCr、CIE Lab 及 CIE Luv 等^[9]。本文中选取 YCbCr 颜色空间作为肤色分布统计的映射空间,该空间的优点是受亮度变化的影响较小,而且是二维独立分布,能较好地限制肤色分布区域^[10]。

2.1.1 YCbCr 颜色空间

YCrCb 颜色空间主要用于视频编码和广播中,其中 Y 表示亮度, C_r 和 C_b 是颜色差别信号,代表色度。YCrCb 颜色空间与 RGB 空间的变换关系为

RGB → YCrCb:

$$\begin{bmatrix} Y \\ C_r \\ C_b \end{bmatrix} = \begin{bmatrix} 0.257 & 0.504 & 0.098 \\ 0.439 & -0.368 & -0.071 \\ -0.148 & -0.291 & 0.439 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix},$$

YCrCb → RGB:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.164 & 1.596 & 0.000 \\ 1.164 & -0.813 & -0.391 \\ 1.164 & 0.000 & 2.018 \end{bmatrix} \begin{bmatrix} Y - 16 \\ C_r - 128 \\ C_b - 128 \end{bmatrix}.$$

将图像从 RGB 空间转换到 YCbCr 空间有以下优点:①在 YCbCr 空间中,Y 分量给出所有的亮度信息, C_b 和 C_r 分量不受亮度的影响,可有效地将亮

度分离出来;②彩色分量 Y、C_b、C_r 可以由三基色 R、G、B 经过线性变换得到,计算效率比较高,同时避免了非线性空间的奇异性。

2.1.2 YCbCr 肤色模型

研究表明^[11-12],人的肤色分布相对集中,尽管不同种族的人呈不同的肤色,但其不同主要体现在亮度上,而颜色的色度变化不是很大。在 YCbCr 颜色模型中, C_r 和 C_b 分量分别表示红色和蓝色的色度。因此,只使用 C_r 和 C_b 两个分量,就可以消除光线变化对肤色产生的影响。本文中通过阈值对肤色区域进行判决,对图像中的每个点 (x,y),用 f(x,y) 表示该点是否属于肤色像素,可得

$$f(x,y) = \begin{cases} 1, & 77 \leq C_b \leq 127, \text{ 且 } 133 < C_r \leq 173; \\ 0, & \text{其他.} \end{cases}$$

2.1.3 C_b、C_r 颜色直方图

为了实现跟踪,首先需要用颜色信息表征人脸。选用 YCbCr 颜色空间中的 C_b、C_r 两个分量的直方图。

图 2 中给出了人脸区域 C_b、C_r 直方图及采用了 Epanechnikov 核函数后的人脸区域 C_b、C_r 的概率分布情况。

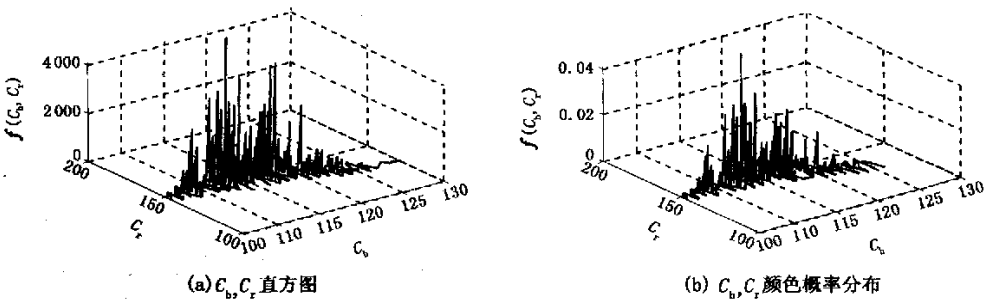


图 2 人脸区域颜色分布

2.2 自适应人脸方向模板

在传统的均值漂移算法中,窗口方向总是垂直方向的。当人脸发生倾斜时,ρ[p(y_i),q] 将会变小,导致跟踪结果变差,甚至跟踪丢失,同时会使人脸定位不准确,不能满足人脸跟踪的目的。为此,采用自适应角度的模板进行跟踪。

在模板中加入人脸倾角 θ,其计算公式为^[13]

$$\theta = \frac{\arctan \left(\frac{2 \left(\frac{M_{11}}{M_{00}} - x_c y_c \right)}{\left(\frac{M_{20}}{M_{00}} - x_c^2 \right) - \left(\frac{M_{02}}{M_{00}} - y_c^2 \right)} \right)}{2}. \tag{12}$$

其中

$$x_c = \frac{M_{10}}{M_{00}}, y_c = \frac{M_{01}}{M_{00}},$$
$$M_{00} = \sum_x \sum_y f(x,y), M_{10} = \sum_x \sum_y x f(x,y),$$
$$M_{01} = \sum_x \sum_y y f(x,y), M_{11} = \sum_x \sum_y xy f(x,y),$$
$$M_{20} = \sum_x \sum_y x^2 f(x,y), M_{02} = \sum_x \sum_y y^2 f(x,y).$$

式中,点 (x_c,y_c) 表示质心。

计算倾角 θ 时要注意,计算区域是一个正方形,其中心为人脸区域矩形的中心,边长为人脸区域矩形对角线的长。

2.3 算法步骤

① 在初始化帧中定位人脸区域,假设人脸区域中心点位置为 x_c ,根据式(12)计算出人脸倾角 θ_0 ,并根据式(5)计算出人脸模板颜色分布, $q = \{q_u\}_{u=1, \dots, m}$;

② 在当前帧中,给定人脸区域的中心位置为 y_0 ,倾角为 θ_{y_0} ,根据式(12)和式(6)计算出人脸倾角 θ_{y_0} 和人脸区域的颜色分布, $p(y_0) = \{p_u(y_0)\}_{u=1, \dots, m}$;

③ 根据式(7)计算出 q 和 $p(y_0)$ 之间的 Battachryya 系数;

④ 根据式(10)计算出权重 $\{\omega_i\}_{i=1, \dots, n_k}$;

⑤ 根据式(11)计算出人脸候选区域位置 y_1 ;

⑥ 根据式(12)和式(6)计算出 θ_{y_1} 和 $\{p_u(y_1)\}_{u=1, \dots, m}$, 并估计出 Battachryya 系数

$$\rho[p(y_1), q] = \sum_{u=1}^m \sqrt{p_u(y_1) q_u};$$

⑦ 当 $\rho[p(y_1), q] < \rho[p(y_0), q]$ 时,用 $\frac{1}{2}(y_0 + y_1)$ 来代替 y_1 ,回到步骤⑥;

⑧ 如果 $\|y_1 - y_0\| < \varepsilon$, 停止;否则 $y_0 \leftarrow y_1, \theta_{y_0} \leftarrow \theta_{y_1}$, 回到步骤②,继续后续帧的计算。

3 实验结果

采用 YCbCr 颜色空间中的 C_b 和 C_r 值作为颜色特征,并将它们量化为 16×16 个直方图特征。在 2.4 GHz PC 机及 VC 环境下,对 320×240 的视频序列跟踪速度能达到 18 帧/s。

图 3 为含自适应角度模板的跟踪,作为比较,同时使用传统的均值漂移方法和含角度自适应模板的均值漂移方法进行跟踪。从图中可以看出,传统的方法虽然能够进行跟踪,但是不能精确地刻画人脸的位置。

图 4 表示光线发生变化时的人脸跟踪结果(图 3,4 中,由左至右、由上至下分别是第 60, 80, 100, 120, 140, 160, 180, 200 帧)。

图 5 中分别给出了带角度自适应模板跟踪、垂直模板跟踪及光线变化时的 Battachryya 系数(第 128 帧时光线变暗)。

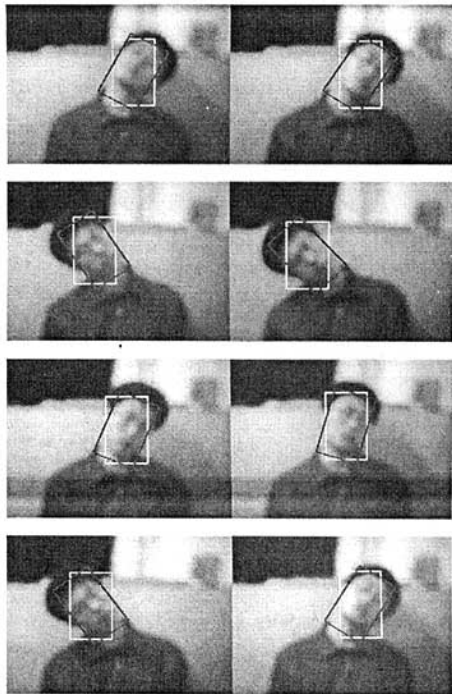


图 3 含自适应角度模板人脸序列跟踪结果

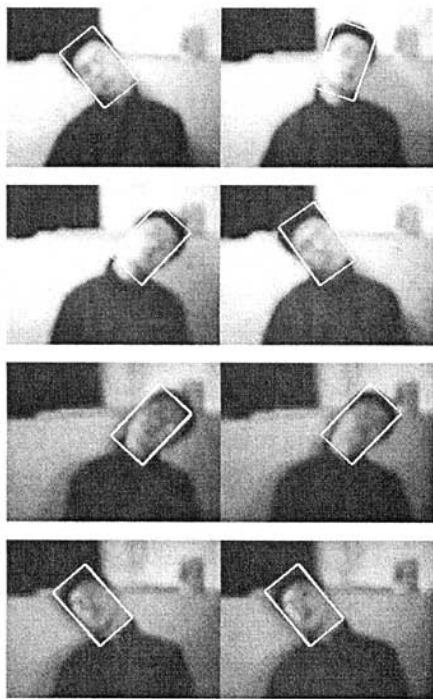


图 4 光线变化时人脸序列跟踪结果

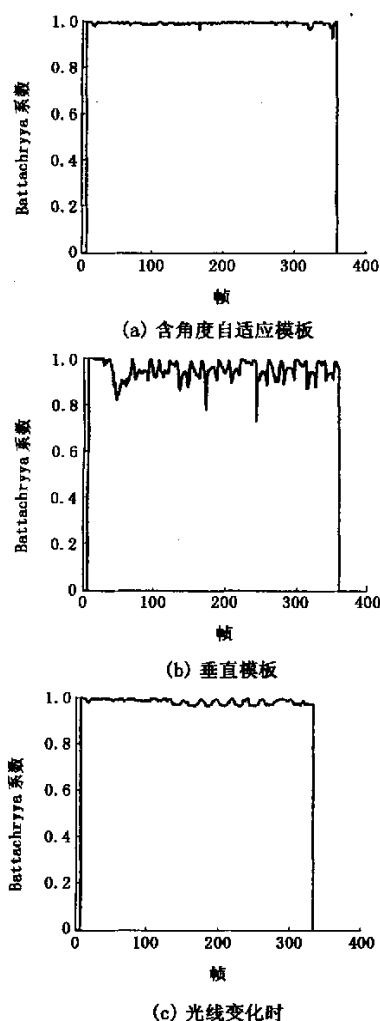


图5 跟踪中的 Battachryya 系数

4 结束语

传统的均值漂移算法采用的是垂直模板和 RGB 肤色模型,当人脸倾斜时不能够准确地表示人脸区域,而且光线变化时会导致跟踪丢失。针对上述缺点进行改进,提出了角度自适应模板和 YCbCr 自适应肤色模型。角度自适应模板能够精确地描绘出人脸的位置,而 YCbCr 肤色模型能够消除光线变化对跟踪的影响。为后续的人脸区域处理如人脸动态识别等奠定了好的基础。

参考文献:

[1] FUKUNAGA K, HOSTETLER L D. The estimation of the gradient of a density function with applications in pattern recognition[J]. IEEE Trans Information Theory, 1975, 21(1):32-40.

[2] CHENG Y. Mean shift, mode seeking, and clustering. IEEE Trans [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1995, 17(8):790-799.

[3] COMANICIU D, RAMESH V, MEER P. Kernel-based object tracking[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2003, 25(5):564-577.

[4] LERDSUDWICHAI C, ABDEL-MOTTALEB M. Algorithm for multiple faces tracking[C]//Proceedings of 2003 IEEE International Conference on Multimedia and Expo (ICME'03), Baltimore:IEEE Press, c2003:777-780.

[5] LERDSUDWICHAI C, ABDEL-MOTTALEB M, ANSARI A. Tracking multiple people with recovery from partial and total occlusion[J]. Pattern Recognition, 2005, 38: 1059-1070.

[6] JIA J, ZHANG Y, ZHAO R, et al. Tracking of objects in image sequences using multi-freeness mean shift algorithm[C]//Proceedings of the Fourth International Conference on Machine Learning and Cybernetics (ICMLC'05). Guangzhou: IEEE Press, c2005:18-21.

[7] SCOTT D W. Multivariate density estimation: theory, practice, and visualization, wiley series in probability and mathematical statistics [M]. New York, London, Sydney: John Wiley & Sons, 1992.

[8] 王延江. 基于遗传算法的人脸跟踪研究[J]. 石油大学学报:自然科学版, 2002, 26(4):92-94.
WANG Yan-jiang. Study on humankind face tracking with enetic algorithm[J]. Journal of the University of Petroleum, China (Edition of Natural Science), 2002, 26(4):92-94.

[9] 王延江. 新一代(第四代)人机交互的概念框架及其关键技术研究[D]. 北京:北方交通大学信息科学研究所, 2001.

[10] GONG Y, SAKAUCHI M. Detection of regions matching specified chromatic features[J]. Computer Vision and Image Understanding, 1995, 61(2):263-269.

[11] YANG J, WAIBEL A. Skin-color modeling and adaptation[J]. Proceedings of ACCV 98, 1998, 2:687-694.

[12] YANG J, WAIBEL A. Tracking human faces in real time [R]. CMU, Carnegie Mellon University, CA: Technical Report CMU-CS-95-210, 1995.

[13] WANG Y, YUAN B. Human face tracking using gaussian mixture models and fuzzy shape analysis [C]//7th International Conference on Signal Processing (ICSP'04), Beijing: IEEE Press, c2004:1322-1325.

(编辑 修荣荣)