

基于多个颜色分布模型的 Camshift 跟踪算法

左军毅¹ 梁彦¹ 潘泉¹ 赵春晖¹ 张洪才¹

摘要 视角变化往往会引起目标外观特征的变化,传统的基于单一颜色直方图模型的 Camshift 跟踪算法往往不能适应这种变化.为此,本文从提高模型描述能力入手,提出利用目标外观的先验知识,为目标建立多个颜色模型,在此基础上设计目标函数,通过对目标函数的优化,实时地为每帧跟踪从多个模型的凸组合中选取最优模型.另外,在对 Camshift 算法深入研究的过程中,发现了概率图平均亮度和图像块颜色分布之间的一种定量关系,这种关系为进一步理解多模型算法的工作机理提供了帮助.头部跟踪的实验结果表明,与单一固定模型以及自适应单模型算法相比,多模型 Camshift 算法对目标外观的快速变化适应性很强,而且计算代价不大.

关键词 目标跟踪, 概率图, 均值漂移, 多模型

中图分类号 TP391.4

Camshift Tracker Based on Multiple Color Distribution Models

ZUO Jun-Yi¹ LIANG Yan¹ PAN Quan¹ ZHAO Chun-Hui¹ ZHANG Hong-Cai¹

Abstract Traditional Camshift tracker based on a single color histogram model is not robust to appearance changes of the target caused by changing viewpoint. To tackle the problem, a possible way is to use a model with more powerful representation ability. In this paper, we model the target with multiple color distributions according to prior knowledge of the target and then design a cost function. Through minimizing the cost function, the optimal model is selected in real time from the convex combination of model sets for tracking in the next frame. In addition, when researching Camshift tracker in detail, we find the relationship between the average intensity of probability image and the color distribution histogram of image patches, which helps to illuminate the mechanism of model selecting process. Experimental results conducted on head sequences demonstrate our tracker can deal with dramatic appearance changes of target in an elegant manner with low computational cost when compared with Camshift tracker with a single fixed model or single adaptive model.

Key words Target tracking, probability image, Camshift, multiple model

作为匹配搜索类跟踪算法的典型代表之一,均值漂移跟踪算法^[1-2]以其计算量小,对目标变形、旋转变换适应性强等特点得到普遍的重视和广泛的研究.均值漂移算法有两个基本版本: Camshift^[1]算法和标准 Meanshift^[2]算法,这两种算法都以颜色直方图为目标表示模型,通过迭代求解加权图的“质心”来跟踪目标. Camshift 在计算直方图时为每个像素赋予等权值,并用直方图反向投影来计算加权图,这种加权图又称为概率图.标准 Meanshift 在计算直方图时用满足一定条件^[3]的核函数为像素加权,通过最大化目标模型和候选目标模型之间的 Bhattacharyya 系数而得到了求取目标形心位置的迭代公式.从本质上讲,迭代的目的是获取加权图的“质心”,不过此处的加权图不同于概率图,计

算上比概率图略微复杂.

文献中介绍的均值漂移算法及其改进算法都是基于单一颜色分布模型的.由于对目标外观先验知识利用不够充分,单一颜色分布模型对目标外观的描述完备性差,当遭遇变外观的跟踪对象时效果不好.这对一种计算高效的算法来讲不能不说是一种遗憾.实际上,对变外观目标的跟踪问题长期以来都是视觉跟踪领域的一大难题.目前,解决该问题的主要途径是采用学习算法对模型进行更新^[4-9].文献 [4, 9] 用颜色直方图对目标建模,通过引入衰减因子来更新模型; Zhou^[5] 对目标建立混合高斯模型,当新的观测到来时用 EM 算法对其进行更新;文献 [6-8] 用 PCA 方法对目标建立子空间模型,再用子空间学习算法来更新模型.然而由于对模型更新过程缺少必要的监督机制,潜在的模型漂移问题使这些算法都存在丢失目标的风险.尽管人们对其进行了一些研究^[10],模型漂移问题仍没有得到彻底解决.

充分利用目标外观的先验知识来提高模型描述能力是解决变外观目标跟踪问题的一种可能途径.考虑到 Camshift 算法在目标表示上比 Meanshift 算法简单、便于进行模型的组合,本文在 Camshift 算法框架下用多个颜色分布对目标建模,在此基

收稿日期 2007-03-21 收修改稿日期 2007-07-09

Received March 21, 2007; in revised form July 09, 2007

国家自然科学基金重点项目 (60634030), 国家自然科学基金 (60372085) 资助

Supported by National Key Project of Natural Science Foundation of China (60634030) and National Natural Science Foundation of China (60372085)

1. 西北工业大学自动化学院 西安 710072

1. College of Automation, Northwestern Polytechnical University, Xi'an 710072

DOI: 10.3724/SP.J.1004.2008.00736

础上, 结合当前目标颜色分布以及目标周围背景颜色分布构建二次目标函数. 跟踪过程中, 通过最小化目标函数从模型集合中实时地选取最优模型作为当前参考模型来实现对目标的可靠跟踪. 本文对目标函数中参数取不同值的情况分别进行了研究, 得到了两种模型选择算法: 多模型切换和多模型组合算法. 最后从概率图的性质入手, 对这两种多模型算法的工作机理进行了分析.

1 概率图与 Camshift 跟踪算法

概率图是 Camshift 算法中的一个关键概念, 它是用直方图反向投影来获取的. 为了简化描述, 规定此处的直方图是经过归一化处理的, 即

$$\mathbf{q} = \{q(u)\}_{u=1,2,\dots,b} \tag{1}$$

满足

$$\sum_{u=1}^b q(u) = 1 \tag{2}$$

其中 b 为直方图的箱格数日, u 为箱格索引.

若 \mathbf{q} 是对图像块 R_q 进行统计得到的, 则

$$q(u) = \frac{1}{n} \sum_{(x,y) \in R_q} \delta[c(f_{R_q}(x,y)) - u] \tag{3}$$

式中, $f_{R_q}(x,y)$ 是图象块 R_q 的图像函数, $c(\cdot) : \mathbf{R}^3 \rightarrow \{1,2,\dots,b\}$ 是颜色空间的量化函数, $\delta(\cdot)$ 是 Kronecker 函数, n 为 R_q 中的像素个数.

定义 1. 对已知的图像块 R 和颜色分布直方图 $\mathbf{q} = \{q(u)\}_{u=1,2,\dots,b}$, 若 R 和与其同尺寸的图像块 I 的对应像素之间存在关系

$$I(x,y) = \sum_{u=1}^b q(u) \delta(c(f_R(x,y)) - u) \tag{4}$$

则称 I 为 R 对分布 \mathbf{q} 的概率图. 其中 $f_R(x,y)$ 是 R 的图象函数, 其他符号的定义与式 (3) 同. 概率图的亮度值在区间 $[0,1]$ 中 (为便于叙述, 文中的概率图直接由式 (4) 定义, 而不是常规意义上的灰度图), 且是单通道的. 有时为了显示的需要, 还需将其线性映射到 $[0,255]$ 范围内并取整. 用图 1 左侧子图中白色线框区域计算 \mathbf{q} , 并将 \mathbf{q} 向原图像反向投影计算概率图, 概率图的显示结果如图 1 右侧子图.

如果把概率图上的像素看作有质量的点, 像素的质量和其亮度成正比, 则 Camshift 算法就是通过连续迭代求解跟踪框内概率图的“质心”来跟踪目标的, 具体执行步骤见文献 [1].

Bradski^[1] 提出的完整的 Camshift 算法还通过计算概率图的二阶矩来获得目标的尺度和转角. 为简化问题, 本文将重点放在对目标外观变化的适应

性上, 在研究时仅考虑对目标形心的连续定位, 如何在多模型框架下实现尺度以及转角定位则留待以后研究.

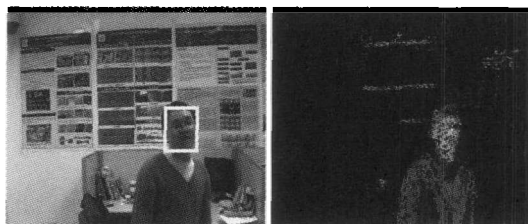


图 1 一帧图像和其对应的肤色概率图

Fig. 1 An image and its skin color probability image

2 多模型 Camshift 跟踪算法

2.1 算法方案

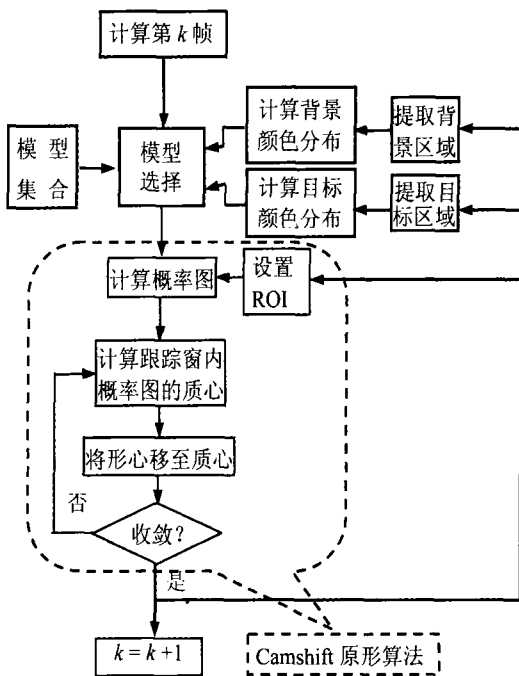


图 2 多模型 Camshift 跟踪算法流程

Fig. 2 The flow chart of multiple model Camshift

单模型 Camshift 算法跟踪变外观目标时鲁棒性不好, 其主要原因是单一颜色模型没有抓住目标外观特征的全部, 因此对外观特征的描述是不完备的. 如何提升模型描述能力以克服算法的这一缺陷呢? 一种可能的方法是在原算法中引入多模型思想, 为此本文提出图 2 所示的多模型跟踪方案. 其中模型集合中的 m 个参考模型是根据先验知识得到的 m 个颜色分布. 每帧跟踪结束时, 模型选择模块根据当前跟踪环境从 m 个模型的凸组合中选择最佳模型用于下一帧跟踪. 考虑到一个好的模型既要保证和目标模式有高的相似度, 又要保证和背景模式有较大的差异性, 因此, 这里的“当前跟踪环境”需要用两个颜色分布来表示, 即: 目标区域及目标周围背景区域的颜色分布. 图 2 中有对应的四个模块用于

计算这两个分布, 计算时先将当前目标位置信息反馈给这些模块, 获得目标区域及背景区域后再通过式 (3) 计算这两个分布. 图中虚线框内的模块组成基本 Camshift 跟踪算法, 本文对其详细算法不再赘述. 下面首先介绍建立模型集合以及选取背景区域的一些原则, 然后重点讨论模型选择算法.

2.2 模型集合的建立与背景区域的选取

1) 模型集合的建立: 首先, 模型集合是根据人们对目标外观的先验知识获取的, 这种先验知识可能来自对目标在不同视角下的外观的认识, 也可能来自目标在不同外界环境下的外观表现, 比如光照条件. 其次, 为了简化计算, 模型个数应当在保证对目标描述的完备性的前提下尽可能少. 对同一跟踪对象, 模型切换算法只能从模型集中选择其一, 而模型组合算法可以从基本模型的凸组合中选取, 模型空间比前者大, 对基本模型数量的需求也较前者少. 对头部跟踪而言, 由于头部颜色以黑色和肤色为主, 因此对于模型组合算法, 只需要两个基本模型就可实现鲁棒跟踪, 而对模型切换算法, 两个基本模型有时可能不能满足需要. 另外, 模型空间并不是越大越好, 模型空间太大会降低模型对目标和非目标的区别能力.

2) 背景区域的选取: 真正对跟踪器可能构成威胁的仅仅是目标周围的背景区域, 因此人们通常选择目标周围的一个“环”形区域来计算背景参考模型^[2,9]. 选取该区域的宽度时应综合考虑目标单帧移动距离、计算直方图所需的样本数量等因素, 若“环”形区域太宽, 则离目标较远的背景区域会淡化目标周围真正有威胁的颜色模式, 这可能会误导模型的选取; 若太窄, 则当目标运动较快时, 得到的背景颜色分布将不能反映迭代过程中目标周围真实的背景颜色模式. 在本文的实验中, 背景区域的取法与文献 [9] 相同.

2.3 模型选择

设 m 个目标模型分别为 $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_m$, 其中 $\mathbf{q}_i = [q_{i,1}, q_{i,2}, \dots, q_{i,b}]^T$, $i \in \{1, \dots, m\}$ 是关于目标的 b 维颜色分布直方图. 将这组模型用矩阵 $\mathbf{Q} \in \mathbf{R}^{b \times m}$ 来表示, 即 $\mathbf{Q} = [\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_m]$. 为了保证最优模型本身是一个分布, 假设最优目标模型来自这 m 个模型的凸组合, 即

$$\mathbf{q}_{opt} = \alpha_1 \mathbf{q}_1 + \alpha_2 \mathbf{q}_2 + \dots + \alpha_m \mathbf{q}_m = \mathbf{Q} \boldsymbol{\alpha} \tag{5}$$

其中向量 $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_m]^T$ 为最优组合系数, 满足

$$\begin{cases} \alpha_i \geq 0, & i = 1, 2, \dots, m \\ \sum_{i=1}^m \alpha_i = 1 \end{cases} \tag{6}$$

若用 $\mathbf{p}_o \in \mathbf{R}^{b \times 1}$, $\mathbf{p}_b \in \mathbf{R}^{b \times 1}$ 来表示某时刻目标区域和背景区域的颜色分布, 则 \mathbf{q}_{opt} 和 \mathbf{q}_o 以及 \mathbf{q}_b 之间欧氏距离的平方为

$$d_{ro}^2 = \|\mathbf{p}_o - \mathbf{q}_{opt}\|_2^2 \tag{7}$$

$$d_{rb}^2 = \|\mathbf{p}_b - \mathbf{q}_{opt}\|_2^2 \tag{8}$$

最优目标模型应使 d_{ro}^2 尽可能小, 同时使 d_{rb}^2 尽可能大. 这是一个典型的多目标规划问题, 用线性加权法将其转化为单目标规划问题, 即

$$d^2 = \lambda_1 d_{ro}^2 - \lambda_2 d_{rb}^2 \tag{9}$$

λ_1, λ_2 为选定的权系数, 满足

$$\lambda_1 \geq 0, \lambda_2 \geq 0, \lambda_1 + \lambda_2 = 1 \tag{10}$$

λ_1, λ_2 的大小反映了两个目标函数的重要程度. 将式 (5)、(7) 和 (8) 代入目标函数 (9), 并将目标函数看作 $\boldsymbol{\alpha}$ 的函数, 即

$$d^2(\boldsymbol{\alpha}) = (\lambda_1 - \lambda_2) \boldsymbol{\alpha}^T \mathbf{Q}^T \mathbf{Q} \boldsymbol{\alpha} - 2(\lambda_1 \mathbf{p}_o - \lambda_2 \mathbf{p}_b)^T \mathbf{Q} \boldsymbol{\alpha} + \lambda_1 \mathbf{p}_o^T \mathbf{p}_o - \lambda_2 \mathbf{p}_b^T \mathbf{p}_b \tag{11}$$

模型选择问题转化为求 $\boldsymbol{\alpha}$, 使 $d^2(\boldsymbol{\alpha})$ 最小. 略去与 $\boldsymbol{\alpha}$ 无关的项, 剩余部分用 $d_r^2(\boldsymbol{\alpha})$ 表示, 问题转化为求满足下式的 $\boldsymbol{\alpha}$

$$\boldsymbol{\alpha} = \arg \min_{\boldsymbol{\alpha}} d_r^2(\boldsymbol{\alpha}) \tag{12}$$

其中

$$d_r^2(\boldsymbol{\alpha}) = (\lambda_1 - \lambda_2) \boldsymbol{\alpha}^T \mathbf{Q}^T \mathbf{Q} \boldsymbol{\alpha} - 2(\lambda_1 \mathbf{p}_o - \lambda_2 \mathbf{p}_b)^T \mathbf{Q} \boldsymbol{\alpha} \tag{13}$$

$\boldsymbol{\alpha}$ 的约束条件由式 (6) 给出. 一般认为, 模型和背景的差异程度不会比模型和目标的相似程度更重要, 因此只考虑 $\lambda_1 \geq \lambda_2$ 的情形. 按照 λ_1 和 λ_2 的取值, 分两种情况讨论:

1) 模型切换算法

当 $\lambda_1 = \lambda_2$ 时, 关于 $\boldsymbol{\alpha}$ 的二次项为零, 令 $\mathbf{p}_{bo} = (\mathbf{p}_b - \mathbf{p}_o)$, 问题转化为求 $\boldsymbol{\alpha}$, 使

$$d_r^2(\boldsymbol{\alpha}) = \mathbf{p}_{bo}^T \mathbf{Q} \boldsymbol{\alpha} = \sum_{i=1}^m \langle \mathbf{p}_{bo}, \mathbf{q}_i \rangle \alpha_i \tag{14}$$

最小, 约束条件由式 (6) 给出, 式 (14) 中 $\langle \cdot, \cdot \rangle$ 是内积运算符. 这是一个线性规划问题, 采用单纯形法可迭代获得问题的最优解. 然而, 该问题的特殊性在于: 其可行域的全部 m 个极点可表示为 \mathbf{x}_i , $i = 1, 2, \dots, m$, 其中 $\mathbf{x}_i = [x_1, \dots, x_m]^T$ 的第 i 个分量为 1, 其余全为 0. 线性规划理论已证明: 只要线性规划问题的最优解存在, 该最优解必然在可行

域的某个极点上达到. 因此, 原问题的最优解在第 i 个极点上可以达到, 其中

$$i = \arg \min_i \langle \mathbf{p}_{bo}, \mathbf{q}_i \rangle \quad (15)$$

上式说明, 求解式 (14) 的线性规划问题可用比较内积运算代替单纯形法的迭代过程, 使运算量降低. 从极点的特殊结构还可以发现, 最优模型仅仅是在 m 个模型 $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_m$ 之间不断地切换, 即得到的是—种多模型切换 Camshift 跟踪算法. 在得到最优模型后, 跟踪过程可按文献 [1] 的步骤进行.

2) 模型组合算法

当 $\lambda_1 > \lambda_2$ 时,

$$\frac{\partial^2 d_r^2(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}^T} = 2(\lambda_1 - \lambda_2) \mathbf{Q}^T \mathbf{Q} \quad (16)$$

$\mathbf{Q}^T \mathbf{Q}$ 是 $m \times m$ 正定对称阵, $\lambda_1 - \lambda_2 > 0$, 根据凸函数的判别条件知, 式 (13) 是凸函数, 因此模型选择问题的实质就是凸二次规划问题, 约束条件仍然由式 (6) 给出. 求解这类问题已有一些标准算法可供选用, 比如内点法、外点法、可行方向法等. 求出 $\boldsymbol{\alpha}$ 后, 代入式 (5) 可计算出最优模型, 然后按文献 [1] 的方法进行跟踪.

由以上求解过程可知, 若基本模型之间存在相关性, 则对于模型切换算法而言, 只是增加了内积运算的次数, 而算法的鲁棒性还可能因此而得到一定提高. 对于模型组合算法, 相关性的出现意味着模型有冗余, 二次规划问题中的对称阵 $\mathbf{Q}^T \mathbf{Q}$ 非严格正定, 使得问题 (6)、(12) 的最优解可能不唯一, 但这并不会妨碍我们找出其中一个最优解, 所以模型有冗余对跟踪性能的影响不大. 模型冗余的缺点是使问题求解的规模增大.

2.4 对多模型 Camshift 算法的机理分析

为了进一步理解多模型 Camshift 算法的工作机理, 需先研究概率图的性质.

定理 1. 任给两个图像块 R_q 和 R_p , 假设它们各自的归—化颜色分布直方图为 $\mathbf{q} = \{q(u)\}_{u=1,2,\dots,b}$, $\mathbf{p} = \{p(u)\}_{u=1,2,\dots,b}$, 记 R_q 对于分布 \mathbf{p} 的概率图以及 R_p 对于分布 \mathbf{q} 的概率图分别为 I_{qp} 和 I_{pq} , 则 I_{qp} 和 I_{pq} 的平均亮度相等, 其值等于向量 \mathbf{q} 和 \mathbf{p} 的内积.

证明. 设 R_q 中像素的个数为 n , 图像函数为 $R_q(x, y)$, 根据式 (3)、(4), R_q 对 \mathbf{p} 的概率图的亮度均值 \bar{I}_{qp} 表示为

$$\begin{aligned} \bar{I}_{qp} &= \frac{1}{n} \sum_{(x,y) \in R_q} I_{qp}(x, y) = \\ &= \frac{1}{n} \sum_{(x,y) \in R_q} \sum_{u=1}^b p(u) \delta[c(R_q(x, y)) - u] = \end{aligned}$$

$$\begin{aligned} &= \sum_{u=1}^b p(u) \frac{1}{n} \sum_{(x,y) \in R_q} \delta[c(R_q(x, y)) - u] = \\ &= \sum_{u=1}^b p(u) q(u) = \langle \mathbf{q}, \mathbf{p} \rangle \end{aligned}$$

同理可得: $\bar{I}_{pq} = \langle \mathbf{p}, \mathbf{q} \rangle$.

由内积的交换律可知

$$\bar{I}_{pq} = \bar{I}_{qp} = \langle \mathbf{p}, \mathbf{q} \rangle \quad (17)$$

□

定理 1 是概率图的性质, 它对 Camshift 以及多模型 Camshift 算法均适用. 根据该定理, 还可以得到如下结论:

定理 2. 第 2.3 节中介绍的多模型切换算法总是从 m 个参考模型中选择使目标区域和背景区域概率图亮度均值之差最大的参考模型作为最优模型; 另外, 多模型 Camshift 算法计算出的目标形心位置等于按照各模型计算出的概率图的加权—阶矩与加权零阶矩的比值, 其中权值向量为 $\boldsymbol{\alpha}$.

证明. 对于多模型切换算法, 式 (15) 可写为

$$i = \arg \max_i (\langle \mathbf{p}_o, \mathbf{q}_i \rangle - \langle \mathbf{p}_b, \mathbf{q}_i \rangle) \quad (18)$$

可见, 定理前半部分显然成立. 特殊地, 若 R_p 中仅包含一个像素 (x, y) , 其颜色分布为 $\mathbf{p}_{(x,y)}$, 则 R_p 对 \mathbf{q} 的概率图 $I(x, y)$ 可表示为

$$I(x, y) = \langle \mathbf{p}_{(x,y)}, \mathbf{q} \rangle \quad (19)$$

于是多模型算法的目标形心横坐标可写为

$$\begin{aligned} x_c &= \frac{\sum_{(x,y) \in R_m} x \langle \mathbf{p}_{(x,y)}, \mathbf{q}_{opt} \rangle}{\sum_{(x,y) \in R_m} \langle \mathbf{p}_{(x,y)}, \mathbf{q}_{opt} \rangle} = \\ &= \frac{\sum_{(x,y) \in R_m} x \langle \mathbf{p}_{(x,y)}, \sum_{i=1}^m \alpha_i \mathbf{q}_i \rangle}{\sum_{(x,y) \in R_m} \langle \mathbf{p}_{(x,y)}, \sum_{i=1}^m \alpha_i \mathbf{q}_i \rangle} = \\ &= \frac{\sum_{i=1}^m \alpha_i \sum_{(x,y) \in R_m} x \langle \mathbf{p}_{(x,y)}, \mathbf{q}_i \rangle}{\sum_{i=1}^m \alpha_i \sum_{(x,y) \in R_m} \langle \mathbf{p}_{(x,y)}, \mathbf{q}_i \rangle} = \\ &= \frac{\sum_{i=1}^m \alpha_i m_{10}^i}{\sum_{i=1}^m \alpha_i m_{00}^i} \quad (20) \end{aligned}$$

其中, R_m 为当前跟踪框对应的图像区域. 同理, 有

$$y_c = \frac{\sum_{i=1}^m \alpha_i m_{01}^i}{\sum_{i=1}^m \alpha_i m_{00}^i} \tag{21}$$

m_{00}^i 和 m_{01}^i 、 m_{10}^i 分别是用第 i 个模型计算的目标区域概率图的零阶矩和一阶矩, 因此结论成立. \square

3 实验结果

在 P IV 1.8 GHz 256 M 内存的实验平台上编制了 C++ 代码, 针对头部视频序列执行本文提出的算法, 并与单一固定模型以及单模型自适应跟踪算法进行比较. 目标及背景的分布直方图都是在 RGB 颜色空间中构建的, 整个颜色空间被分成 $16 \times 16 \times 16$ 个量化级. 模型个数为 3, 编号为 I、II 和 III, 分别用图 3(a)、(b) 和 (c) 的三个图像块计算得到. 多模型组合算法中 $\lambda_1 = 0.7$. 由于起始阶段需要初始化三个模型, 跟踪过程是从第 56 帧开始的.

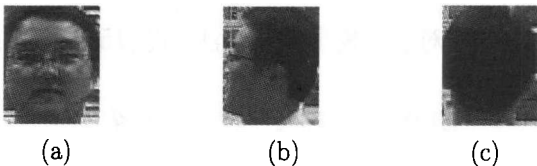


图 3 用于计算三个目标模型的三个图像块

Fig. 3 Three image pitches used to calculate three models

序列中人的运动以及摄像机的运动引起头部影像在做无规律的平移运动, 同时随着人体躯干的转动, 头部影像时而正脸, 时而侧脸, 时而背脸, 这给跟踪算法带来了很大挑战.

图 4(a) 是仅用模型 I 时的跟踪结果. 人处于正脸状态时, 算法能精确锁定目标; 第 81 帧人处于侧脸, 仅靠露出一部分脸部, 算法仍能跟上目标, 但形心定位有一定偏差; 第 119 帧人背脸时, 头部颜色模式变为以黑色为主, 算法立即失效. 即使第 172 帧人恢复正脸, 算法亦不能从错误中恢复.

图 4(b) 是自适应单模型算法的跟踪结果, 其中模型学习速率为 0.05. 和固定模型相比, 虽然自适应模型的鲁棒性有所增强, 但跟踪误差逐渐积累, 造成模型漂移, 最终在第 240 帧左右时丢失目标. 换用其他学习参数重复实验, 结果发现很难找到一个合适的参数来顺利完成跟踪 (序列长 1300 帧). 实际上, 模型的自适应性和漂移问题是一对矛盾, 很难折中, 为了适应目标外观的快速变化就必须选择大的学习速度, 而大的学习速度反过来又使模型漂移问题更加严重.

图 4(c) 是模型切换式 Camshift 算法的跟踪结果. 当人头部转动时, 算法能在 I、II 和 III 三个模

型中进行自动切换. 从第 56 帧到 119 帧, 人头部从正脸逐渐转动到背脸, 最优参考模型及时地从模型 I 经历模型 II 切换到模型 III, 很好地适应了目标外观的变化, 成功实现了对目标的持续跟踪. 该序列的模型切换规律由图 5 给出.

图 4(d) 是模型组合式 Camshift 算法的跟踪结果. 人脸转动时, 算法能从三个模型的凸组合中选取最优模型用于跟踪, 结果非常满意. 该例中, 三个模型之间是高度相关的, 但这并没有影响算法的性能. 序列前 400 帧, 跟踪器的最优模型组合系数变化情况如图 6 所示.

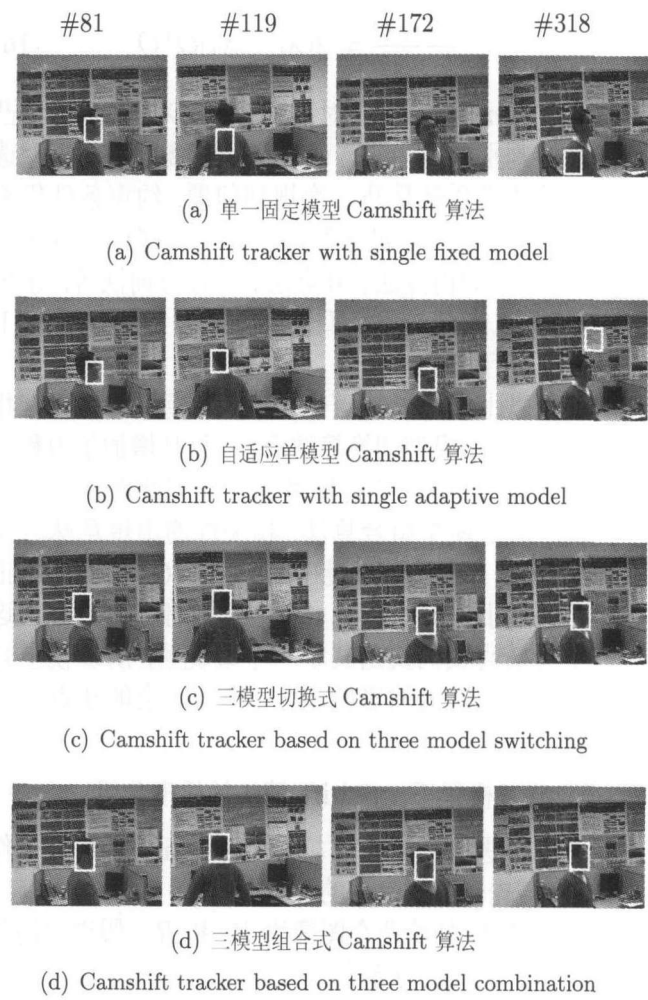


图 4 几种跟踪算法的实验结果

Fig. 4 Experimental results of several algorithms

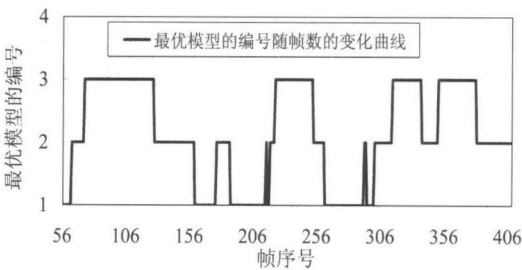


图 5 模型的最优切换规律

Fig. 5 The optimal switching rule of the model

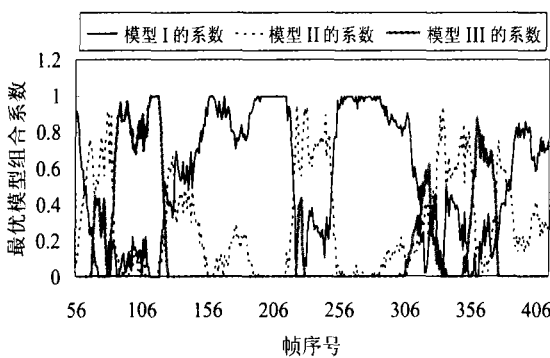


图6 模型的最优组合系数

Fig.6 The optimal combination coefficient of the model

用手工标定的方式获得目标形心位置的真值,将四种算法在垂直/水平方向的定位误差绘制成图7. 其中单一固定模型对应于图中灰色粗点线,跟踪进行到第100帧时,垂直方向定位误差已经超过30个像素(与目标沿纵轴的尺寸相当),说明此时目标已经丢失. 自适应单模型算法(灰色粗实线)在前240帧虽然和真值有相当大的误差,但还不至于失跟. 跟踪进行到250帧左右时,水平方向定位误差超过40个像素,说明目标已经丢失. 相比之下,本文的多模型切换算法(黑色细点线)和多模型组合算法(黑色细实线)误差一直比较小.

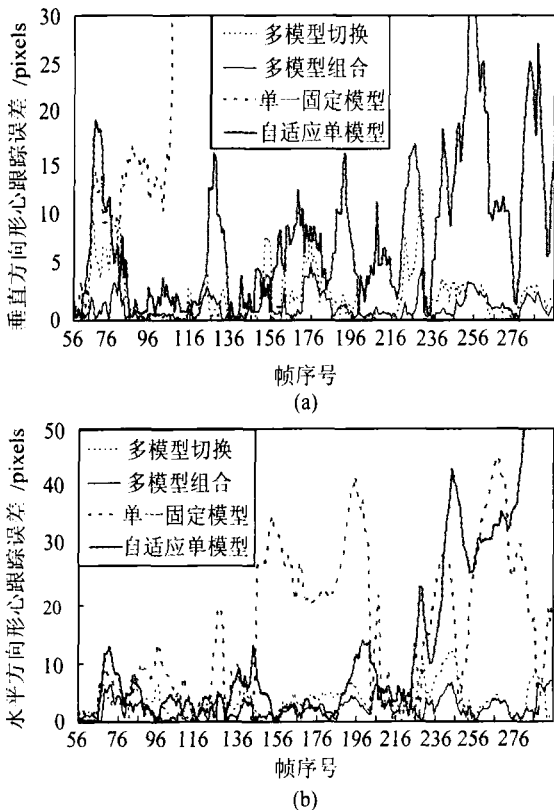


图7 四种算法的误差比较
((a) 垂直方向误差; (b) 水平方向误差)

Fig.7 Performance comparison of four algorithms
((a) Error in vertical direction; (b) Error in horizontal direction)

从图7还可以看出,虽然两种多模型算法都能实现对这个序列的跟踪,但精度不同. 我们将几种算法的性能列表1. 误差峰值是衡量跟踪性能的一个很重要的统计量,模型切换算法纵、横坐标定位的误差峰值分别为13、12,是模型组合算法(分别为6、6)的两倍. 从总体统计特性上讲,模型切换算法的纵、横坐标定位的均方误差达13.13和18.42,远大于模型组合算法的均方误差(分别为4.98、8.56). 而两种单模型算法由于目标发生丢失,因此误差峰值和均方误差都极大.

表1 不同算法的跟踪性能比较

跟踪方法	单一固定模型	自适应单模型	多模型切换	多模型组合
单帧执行时间 (ms)	6.0	6.3	6.2	18.2
X 向误差峰值 (pixels)	45	70	12	6
Y 向误差峰值 (pixels)	140	34	13	6
X 向均方误差	413	476	18.42	8.56
Y 向均方误差	4503	322	13.13	4.98

表1还列出了几种算法的单帧平均执行时间. 在多模型切换算法中,本文用比较内积运算代替了复杂的单纯形迭代求解过程,这使得模型切换的附加计算量很小,单帧执行时间和单模型算法相当. 由于求解凸二次规划问题远比比较内积运算复杂,使得模型组合算法计算开销稍大,然而,18.2毫秒/帧(相当于55帧/秒)的处理速度完全能够满足实际系统的需要. 实验中我们还充分利用了最优模型组合系数在一定程度上具有连续性的特点. 组合系数之所以表现出一定的连续性,是因为视频序列中头部的外观在邻帧之间不会有很大的变化. 因此,每次求解二次规划问题时都将前一帧最优模型系数作为初值进行迭代,这样可以减少迭代次数.

图8给出了室外场景下多目标跟踪的实例. 图像大小为640×480,模型个数仍取为3,其他参数与图4的实验相同. 跟踪起始时,我们同时选中了两个人的头部作为对象,并行执行两套多模型算法进行跟踪. 从结果来看,在目标未发生遮挡的情况下,两种多模型算法的跟踪结果都是比较稳定的. 当目标被遮挡时,算法性能会下降. 就均值漂移算法本身而言,由于它是基于局部梯度优化技术的,因此它并不擅长处理遮挡问题. 这些技术细节已不是本文讨论的重点. 另外,我们也意识到在室外场景中,背景的变化较室内要复杂得多,背景中的与目标具有相似颜色模式的区域会给跟踪器带来很大挑战,目前解决这类问题的常用办法是进行多特征融合.

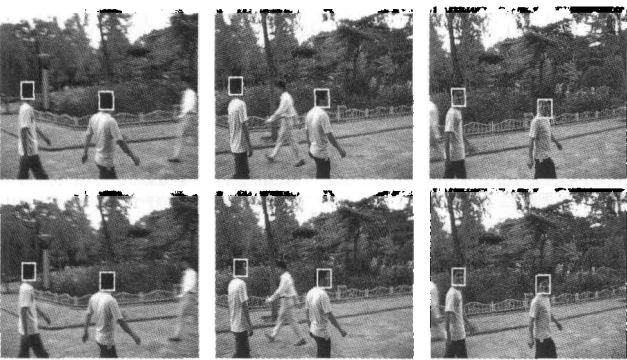


图 8 室外场景下对多个目标的跟踪结果
(第一行: 多模型切换算法; 第二行: 多模型组合算法)
Fig. 8 Multiple target tracking results in outdoor scene
(The first row: multiple model switching algorithm; The second row: multiple model grouping algorithm)

4 结论与展望

本文将多模型的思想引入 Camshift 算法中, 给出了一种多模型 Camshift 跟踪方案, 进而通过设计、优化目标函数, 成功地实现了目标模型根据当前跟踪环境的自适应切换或组合, 从而使 Camshift 算法在跟踪变外观目标时鲁棒性更好. 该方案及其相关算法简单, 机理清晰, 实验效果好. 然而, 颜色直方图本身就是一种比较粗糙的目标表示方法, 一方面, 它对目标变形, 旋转变化的鲁棒性; 另一方面, 对目标和非目标的区分能力较弱. 多模型的引入等于增大了模型空间, 这在一定程度上降低了模型对目标/非目标的识别能力. 作者认为, 在多模型框架下自适应地融合其他目标特征是提升模型识别能力的一种可行方法, 我们对此将作进一步研究.

References

1 Bradski G R. Real time face and object tracking as a component of a perceptual user interface. In: Proceedings of the 4th Workshop on Application of Computer Vision. Princeton, USA: IEEE, 1998. 214–219

2 Comaniciu D, Ramesh V, Meer P. Kernel-based object tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, **25**(5): 564–577

3 Comaniciu D, Meer P. Mean Shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, **24**(5): 603–619

4 Babu R V, Perez P, Bouthemy P. Robust tracking with motion estimation and local kernel-based color modeling. *Image and Vision Computing*, 2007, **25**(8): 1205–1216

5 Zhou S K, Chellappa R, Moghaddam B. Visual tracking and recognition using appearance-adaptive models in particle filters. *IEEE Transactions on Image Processing*, 2004, **13**(11): 1491–1506

6 Lee K C, Kriegman D. Online learning of probabilistic appearance manifolds for video-based recognition and tracking. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, USA: IEEE, 2005. 852–859

7 Nguyen H T, Ji Q, Smeulders A W M. Spatio-temporal context for robust multitarget tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007, **29**(1): 52–64

8 Ross D, Lim J, Yang M H. Adaptive probabilistic visual tracking with incremental subspace update. In: Proceedings of the 8th European Conference on Computer Vision. Prague, Czech Republic: Springer, 2004. 215–227

9 Collins R T, Liu Y X, Leordeanu M. Online selection of discriminative tracking features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, **27**(10): 1631–1643

10 Matthews L, Ishikawa T, Baker S. The template update problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, **26**(6): 810–815



左军毅 博士研究生. 主要研究方向为计算机视觉, 视频处理与图象分析. 本文通信作者. E-mail: zuojunyi@163.com
(ZUO Jun-Yi Ph.D. candidate. His research interest covers computer vision, video process, and image analysis. Corresponding author of this paper.)



梁彦 副教授, 博士. 主要研究方向为非线性滤波, 多目标跟踪, 信息融合, 图象分析. E-mail: liangyan@nwpu.edu.cn
(LIANG Yan Associate professor, Ph.D.. His research interest covers nonlinear filter, multi-target tracking, information fusion, and image analysis.)



潘泉 教授. 主要研究方向为信息融合理论及应用, 多目标跟踪, 智能信息处理, 模式识别, 图像与视频处理. E-mail: quanpan@nwpu.edu.cn
(PAN Quan Professor. His research interest covers theory and application of information fusion, multi-target tracking, intelligence information process, pattern recognition, and image and video process.)



赵春晖 讲师, 博士研究生. 主要研究方向为机器视觉, 模式识别, 图像分析. E-mail: zhaochunhui@nwpu.edu.cn
(ZHAO Chun-Hui Ph.D. candidate, lecturer. His research interest covers computer vision, pattern recognition, and image analysis.)



张洪才 教授. 主要研究方向为控制理论与应用, 多目标跟踪与识别, 信息融合, 动态系统建模. E-mail: zhc@nwpu.edu.cn
(ZHANG Hong-Cai Professor. His research interest covers control theory and application, multi-target tracking and recognition, information fusion, and dynamic system modeling.)