

## 背景变化鲁棒的自适应视觉跟踪目标模型<sup>\*</sup>

王建宇<sup>1+</sup>, 陈熙霖<sup>1,2</sup>, 高文<sup>1,2</sup>, 赵德斌<sup>1,2</sup>

<sup>1</sup>(哈尔滨工业大学 计算机科学与技术学院, 黑龙江 哈尔滨 150001)

<sup>2</sup>(中国科学院 计算技术研究所, 北京 100080)

### Adaptive Appearance Model Robust to Background Variations

WANG Jian-Yu<sup>1+</sup>, CHEN Xi-Lin<sup>1,2</sup>, GAO Wen<sup>1,2</sup>, ZHAO De-Bin<sup>1,2</sup>

<sup>1</sup>(School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China)

<sup>2</sup>(Institute of Computing Technology, The Chinese Academy of Sciences, Beijing 100080, China)

+ Corresponding author: Phn: +86-10-58858300 ext 313, E-mail: jywang@jdl.ac.cn, http://www.jdl.ac.cn

Wang JY, Chen XL, Gao W, Zhao DB. Adaptive appearance model robust to background variations. *Journal of Software*, 2006,17(5):1001-1008. <http://www.jos.org.cn/1000-9825/17/1001.htm>

**Abstract:** A novel method of dynamic object modeling for visual tracking is presented. The Haar transformation is first applied on the incoming image of the video to get features, which are over-complete description of the image. Then, the Fisher criteria are employed for ranking features based on their contributions to the discrimination between the tracked objects and the background. After that, the objects are modeled by the subset of top-ranked features. During tracking, a Kalman filter is used to predict the upcoming destinations of the tracked objects and the features are re-ranked by the discrimination between the objects and predicted locations. Thereafter, objects models will be updated and only discriminative features are kept in it. This proposed strategy aims to maximally maintain the basic discrimination and reduce computational cost simultaneously. To evaluate the performance of the proposed method, several experiments have been conducted on long video sequences. The experimental results show that the proposed method can handle various uncertain factors under the real world conditions and successfully track the objects in real-time.

**Key words:** visual object tracking; online feature selection; appearance model; Haar feature; Kalman filter

**摘要:** 提出了视觉跟踪任务中目标动态建模的一种方法.该方法首先针对跟踪序列中的当前帧图像观测进行 Haar 变换,从而得到图像的过完备特征描述;然后根据 Fisher 准则,评价每个 Haar 特征对目标和当前背景的区别能力,目标模型由那些区别能力最强的 Haar 特征构成.在跟踪过程中,采用卡尔曼滤波算法预测目标下一时刻的可能位置,从而根据目标的图像观测和目标下一时刻可能的位置附近的背景图像观测,对 Haar 特征的区别能力进行动态评价.通过保留区别能力强的特征,同时淘汰区别能力弱的特征,维护目标模型的强可区分性和低计算复杂性.该方法的主要策略是,在最大程度地保持可区分性的前提下,减少计算的复杂性.实验结果表明,在存

\* Supported by the National Natural Science Foundation of China under Grant No.60332010 (国家自然科学基金); the "100 Talents Program" of the Chinese Academy of Sciences (中国科学院百人计划); the Shanghai Municipal Sciences and Technology Committee under Grant No.03DZ15013 (上海市科委项目); the ISVISION Technologies Co., Ltd. (上海银晨智能识别科技有限公司)

Received 2004-11-23; Accepted 2005-05-20

在诸多不确定性因素的真实长序列视频上,该跟踪方法能够实时地完成复杂的目标跟踪任务.

关键词: 视觉目标跟踪;在线特征选择;目标模型;Haar 特征;卡尔曼滤波

中图法分类号: TP391 文献标识码: A

自动目标跟踪是视频分析和理解的基本步骤之一,同时又是视觉监控的重要研究内容.该领域近年来引起了国内外的广泛重视,Proceedings of the IEEE 甚至在 2001 年专门出版了一个专辑<sup>[1]</sup>.对大多数跟踪问题而言,通常假定摄像机是相对背景固定的,或者可以通过全局运动进行补偿.因此,问题的关键就成为目标(前景)和背景的建模问题,所建立的模型必须保证满足以下要求:

- (1) 背景与前景的模型具有足够的可分性;
- (2) 在相邻帧之间,前景在前景模型上的测量值具有良好的连续性和稳定性;
- (3) 模型能够满足实时计算的要求,否则,运动预测模型将无法发挥其效能,导致跟踪的失败.

尽管 Mean-shift, Particle filter 和卡尔曼滤波等已成为近年来广泛使用的跟踪方法,但前景/背景的描述一直是研究的关键<sup>[2-4]</sup>.虽然 GMM 和 HMMs 像素模型、直方图、轮廓、模板、Gabor 等特征或特征的组合被大量使用,但在实际的跟踪中常常还会遇到问题,究其原因主要是模型的相对固定不能适应大幅度的光照、背景变化和物体的大范围运动等<sup>[2,5-11]</sup>.因此,一些研究者已经注意到采用自适应方法的重要性,并且在跟踪过程中加以运用<sup>[12,13]</sup>.Collins 等人通过对 RGB 色彩空间中的 3 个分量赋予不同的权重进行组合,从而可以根据背景的变化,依据目标模型可区分度,从 49 种不同的直方图进行选择所需的特征<sup>[13]</sup>.与这些工作不同,本文提出了一种新的建模策略,其基本思想是,利用 Fisher 区分性度量,在满足上述要求(1)和(2)的基础上,尽可能地选择简单的可区分特征进行前景/背景的描述,从而同时维护前景/背景的可区分性、前景描述的稳定性和计算的实时性.

## 1 方法概述

在本文中,我们提出了一种通过选择图像特征动态构建和更新目标模型,从而同时保证可区分性和低复杂性的方法.由于过完备 Haar 小波特征集在物体识别领域的卓越表现,这里,我们采用 Haar 特征作为目标的图像特征表示方法<sup>[14]</sup>.系统的框架如图 1 所示.

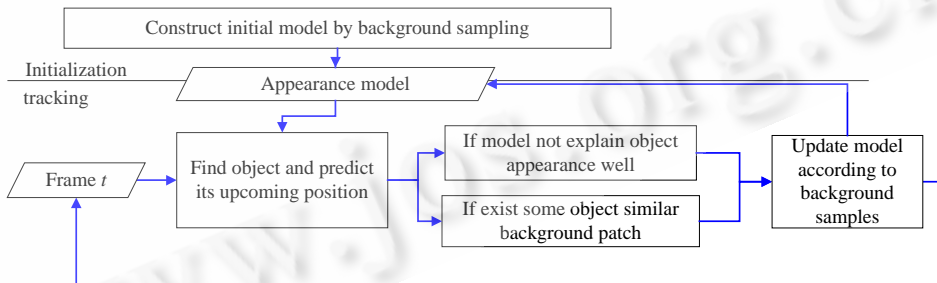


Fig.1 Flowchart of the tracking method with dynamic feature selection embedded

图 1 嵌入动态特征选择功能的跟踪方法流程图

**初始化阶段:**对背景的图像观测(以下简称背景观测)进行随机采样,将过完备 Haar 特征集在背景观测采样点上的特征值作为反例,将目标的图像观测(以下简称目标观测)上的 Haar 特征值作为正例,利用 Fisher 判别准则对每个 Haar 特征分量的分类能力进行评价,然后采用前向特征选择算法从过完备 Haar 特征集中选择当前分类能力最强的特征子集作为目标的初始模型.

**目标跟踪阶段:**对每帧新输入的图像数据,利用当前目标的位置和以往运动的轨迹作为卡尔曼滤波器的输入,从而预测目标在下一帧中可能的出现位置,并进一步利用这一位置信息在当前帧上对其周围背景进行随机采样.之后,比较当前目标模型和下一帧候选窗口之间的相似度,如果高度相似,说明在下一帧中跟踪目标可能

与背景混淆,则需要重新选择特征改变目标模型,以降低目标模型与该候选位置窗口的相似度.同时,还需要计算目标模型和当前目标观测之间的匹配度,如果匹配度过低,表明该模型对跟踪对象的描述能力不足,则需要更新目标模型,以提高目标模型和目标观测之间的匹配度,然后对下一帧图像的目标继续跟踪.

## 2 基于目标/背景差异信息的特征选择

在这一节中,我们首先描述所采用的图像特征,进而给出如何根据当前目标/背景信息选择特征来构建目标模型的方法.

### 2.1 特征集合

我们采用过完备 Haar 特征集合作为图像特征来构建目标模型.在文献[14]的工作中,基于过完备 Haar 特征集合的分类器在物体识别问题上的优异性能得到了充分的表现,而且积分图方法的引入使得 Haar 特征值的求取能够满足实时计算的要求.与跟踪研究中普遍采用的直方图、轮廓等特征相比,Haar 特征不仅具有更强的表现能力,同时还提供了多层次刻画目标的能力.

在实现中,我们采用了 3 类 Haar 特征(如图 2(a)所示).对每一类 Haar 特征有 4 个控制参数:特征在所刻画图像区域中的位置 $(x, y)$ 、特征的宽度和高度 $(w, h)$ .通过 4 个参数的不同组合,我们可以得到数以万计的特征.由于这些特征是对所刻画图像区域的过完备描述,所以,所有不同特征的总集合称为过完备 Haar 特征集合.为叙述方便,用  $F = \{f_i | i = 1, \dots, N\}$  表示特征集合,  $V(z) = \{v_i(z) | i = 1, \dots, N\}$  表示图像窗口  $z$  上求得的特征向量.

### 2.2 目标建模

对每个 Haar 特征,构造相应的弱分类器.考虑到计算复杂性的约束,我们采用如图 2(b)所示的弱分类器,

$$h_i(z) = \begin{cases} 1, & \text{if } v_i(z) \in [\bar{v}_i^{pos} - \theta_i, \bar{v}_i^{pos} + \theta_i] \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

其中:  $\bar{v}_i^{pos} = \bar{v}_i(z^{pos})$  是目标观测上 Haar 特征在特征分量  $f_i$  上的均值;  $\theta_i$  是采用学习算法确定的分类器阈值;  $v_i(z)$  是图像窗口  $z$  上 Haar 特征在特征分量  $f_i$  上的特征值.最后,根据目标模型内所有弱分类器的投票输出,系统判断该窗口为目标观测或者背景观测.

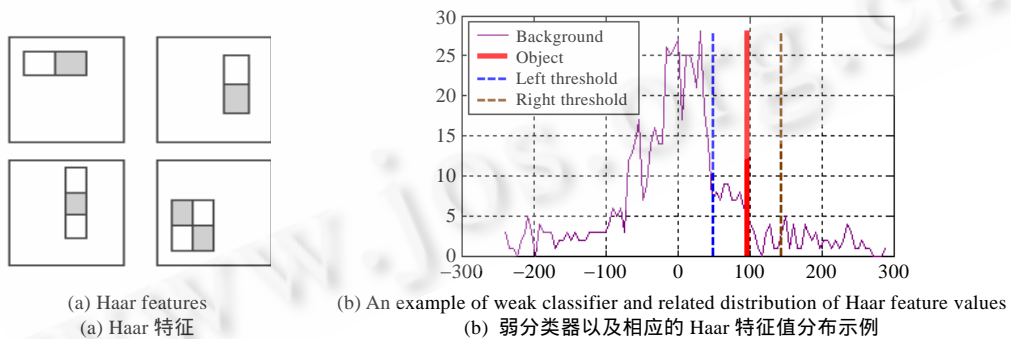


Fig.2 Haar features and corresponding weak classifier

图 2 Haar 特征和相应的弱分类器

虽然很多实验验证了 AdaBoost 是将若干弱分类器组合成强分类器的有效方法,但 AdaBoost 训练非常耗时,因而难以用于在线的特征选择任务.于是,我们提出了一种结合贪心和 Fisher 准则的快速特征选择方法.

步骤 1. 采用面积滤波和信息度量初步选择特征.对  $24 \times 24$  的基准窗口,可以利用的 Haar 特征数目有几万.如此众多的特征数量保证了 Haar 特征集的刻画能力,但同时也存在冗余性.我们首先删除那些计算面积小于一定阈值(在实验中,阈值为 16)的特征,然后对特征进行空间上等距采样.我们发现,在目标观测上求得的特征值,通常较大的特征值对应像素变化大的区域,包含了更多的纹理信息.据此,对当前余下的特征,以较大概率对特征值较大特征进行采样,进一步缩减特征集的大小.经过上述步骤,特征集中大概留有 4 000 个左右的 Haar

特征.

步骤 2. 通过对目标周围的背景进行采样,选择可区分性好的特征.在目标周围的图像上通过均匀分布做  $K$  个采样,其中  $K-L$  个图像窗口对应背景观测, $L$  个图像窗口对应目标观测(图 3(a)演示了部分采样过程).于是,对每一个特征,可以得到  $L$  个正例和  $K-L$  个反例.Fisher 线性判别准则被用来对每一个特征的分类性能进行评价<sup>[15]</sup>.

$$R_i = \frac{|\bar{v}_i^{pos} - \bar{v}_i^{neg}|^2}{S_i^{pos} + S_i^{neg}} \quad (2)$$

其中:  $\bar{v}_i^{pos}(f_i)$  和  $\bar{v}_i^{neg}$  分别是正例集合和反例集合在特征分量  $f_i$  上的均值;  $S_i^{neg} = \sum_{k=1}^{K-L} (v_i^{neg}(z_k) - \bar{v}_i^{neg})^2$  和  $S_i^{pos} = \sum_{k=1}^L (v_i^{pos}(z_k) - \bar{v}_i^{pos})^2$  分别是正例集合和反例集合的散度.

根据分类能力由强到弱对特征分量降序排列,顺序选择  $M$  个特征加入目标模型,直到满足式(3)为止.

$$\left(\sum_{m=1}^M R_m\right) > T \quad (3)$$

其中,  $T$  是保证弱分类器组合后分类能力的阈值.虽然利用现有的学习算法如 AdaBoost 等,能够自动地确定  $T$  的最优值,但对跟踪算法来说,由于  $T$  是随时间变化的,而实时计算的要求并不允许复杂的学习过程,所以我们根据实验结果经验性地将  $T$  设为 0.45,并将如何自动和实时地求取  $T$  作为未来研究的内容.通过这些弱分类器的线性组合,我们得到一个刻画目标的强分类器  $H(t,M)=\{h_m|m=1,\dots,M\}$ ,即目标模型,其中  $h_m$  是式(1)中定义的弱分类器, $t$  表明该模型从  $t$  时刻开始有效.

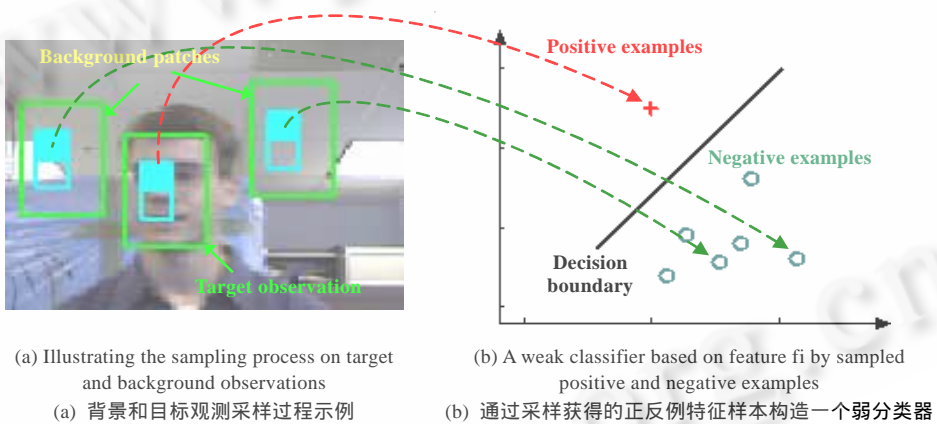


Fig.3 Demonstration of observations sampling and weak classifier training

图 3 观测采样和弱分类器训练示例

### 3 跟踪过程中目标模型的动态更新

在跟踪过程中,我们需要针对背景的变化和目标自身的变化不断调整目标模型.为了后面叙述方便,首先定义当前图像观测窗口和目标模型之间的相似性:

$$w(z) = \frac{1}{M} \sum_{m=1}^M h_m(z) \quad (4)$$

对目标模型  $H(t,M)$ ,我们可以将其分成两个子集合:  $H_1(t,M_1)$  和  $H_2(t,M_2)$ .其中:  $H_2(t,M_2)=\{h_m(z^{pos})=0|m=1,\dots,M_2\}$  包含那些在目标观测上输出为 0 的弱分类器;而  $H_1(t,M_1)=H(t,M)/H_2(t,M_2)$ .

#### 3.1 维护目标/背景差异性的模型更新

在跟踪过程中,如果在预测的下一帧候选目标位置附近的背景观测窗口和当前目标观测在当前目标模型上的相似性(利用式(4)计算)很接近,则容易造成目标歧义而使系统丢失目标.因此,我们需要更新模型,以区别未来的目标和背景观测,从而避免目标歧义现象的发生.

首先找到目标在  $t$  时刻的位置  $x(t)$ , 然后, 利用卡尔曼滤波的过程模型(process model)预测目标在  $t+1$  时刻的候选位置  $\tilde{x}(t+1)$ . 之后, 以  $\tilde{x}(t+1)$  为中心, 在其周围均匀采样背景观测中的  $K$  个窗口, 对每一个采样窗口  $z_k^{neg}$ , 利用式(4)计算其与当前目标模型的相似度  $w(z_k^{neg})$ . 进而计算

$$\Delta w_k = |w(z^{pos}) - w(z_k^{neg})| \tag{5}$$

其中,  $w(z^{pos})$  表示目标模型在目标观测上的相似度. 如果有  $\Delta w_k < T$ , 则  $z_k^{neg}$  所对应的背景观测可能造成目标混淆, 称这样的背景观测为“威胁观测”, 其中  $T$  是经验值. 如果  $T$  设得过高, 会引起模型的频繁更新, 进而导致目标模型的震荡; 如果  $T$  过低则易造成目标和背景混淆, 引起跟踪失败. 根据实验结果, 我们将  $T$  经验性地设为 0.17.

为避免当目标模型和目标观测的匹配值突然下降时, 系统因为威胁观测的存在而丢失目标, 我们收集所有“威胁观测”的特征作为反例来重新选择特征, 以改变目标模型, 使改变后的目标模型与“威胁观测”之间的相似度降低, 从而避免目标混淆现象的出现.

模型更新的步骤如下: 首先从  $H(t, M)$  中丢弃模型子集  $H_2(t, M_2)$ ; 然后, 根据每个特征的可分类性, 从小到大地从模型中移走当前分类能力最弱的子模型, 直到剩下的子模型数量降至原来总数量的 70%; 最后, 将前面收集到的“威胁观测”上得到的 Haar 特征值作为反例, 根据第 2 节描述的特征选择方法更新目标模型.

### 3.2 维护目标描述一致性的更新

在某些情况下, 仅仅更新目标模型的参数(如弱分类器的阈值)无法反映由于目标的运动而引起的目标观测的剧烈变化(比如, 目标的拓扑结构发生改变、剧烈的光照变化等等); 另一方面, 随着时间的流逝, 系统的参数会由于累计误差发生偏移, 造成目标模型与目标观测间匹配度的下降. 这时, 我们需要重新选择合适的特征对目标模型进行调整.

假设在时刻  $t$ , 如果目标模型在目标观测上的匹配度:

$$\left\{ \bar{w}(z^{pos}) = \frac{1}{L} \sum w(z^{pos}) \right\} < T \tag{6}$$

表明当前的目标模型不能合理解释目标观测的变化, 从而引起目标模型与目标观测的偏离. 其中  $T$  是经验值. 如果  $T$  设得过高, 则会引起计算量的增加, 进而导致目标模型的不稳定; 如果  $T$  过低, 则容易造成跟踪的失败. 我们根据实验结果将  $T$  经验性地设为 0.75. 于是回溯到前一帧, 从  $H(t-1, M)$  中抛弃特征子集  $H_2(t-1, M_2)$ , 并从剩余的原始 Haar 特征集中重新选择  $M_2$  个分类能力最强的特征加入目标模型中.

## 4 采用动态建模的目标跟踪

我们将前面提出的动态目标建模方法应用于目标跟踪, 这里选择卡尔曼滤波作为跟踪的一般框架, 构成一个完整的跟踪模块. 在实验中, 定义卡尔曼滤波的状态为  $x(t) = [x, y, dx, dy]^T$ , 观测为  $y(t) = [x, y]^T$ , 其中:  $x$  和  $dx$  分别是目标图像在水平方向的位置和运动速度;  $y$  和  $dy$  分别是目标图像在垂直方向的位置和运动速度. 跟踪的基本过程如下:

1) 初始化阶段: 在参考图像中定义目标观测  $z^{pos}$  以及目标的初始状态  $x(0)$ . 采用卡尔曼滤波的过程模型预测目标的可能出现区域:

$$x(t) = A \cdot x(t-1) + w(t) \tag{7}$$

其中

$$A = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \Delta t = 1, \quad p(w) \propto N(0, Q), \quad Q = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

式(7)用来预测目标的位置, 还要通过局部搜索的方法提取目标的区域, 如果预测模型的准确度比较低, 则相应扩大局部搜索的范围. 在  $\tilde{x}(1)$  所对应的背景观测周围, 采用均匀分布进行背景采样, 得到  $K$  个背景观测

$\{z_k^{neg} | k=1, \dots, K\}$ . 采用第 3 节描述的特征选择方法从完备 Haar 特征集中构建初始目标模型  $H(0, M)$ .

2) 目标观测提取和特征更新过程: 采用当前目标模型  $H(t, M)$  在卡尔曼滤波的预测区域内搜索目标的图像观测  $z^{pos}(t)$ . 根据  $z^{pos}(t)$  更新目标状态的估计  $x(t) = \tilde{x}(t) + K(t)(y(t) - H \cdot \tilde{x}(t))$ , 其中  $H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$ ,  $K(t)$  是卡尔曼增益.

3) 目标位置预测过程: 利用方程式(6)对下一帧的目标可能出现区域进行预测, 得到  $\tilde{x}(t+1)$ .

4) 根据背景进行目标模型的更新: 采用第 3.1 节所述的方法, 搜索在  $\tilde{x}(t+1)$  附近的“威胁观测”. 如果存在, 则重新选择特征更新  $H(t, M)$ , 从而保证不会出现目标歧义的现象.

5) 根据目标表面呈现的变化进行目标模型的更新: 采用第 3.2 节所述的方法, 如果目标模型和目标观测之间的相似度  $\bar{w}(z^{pos}) < T$ , 则重新选择特征更新  $H(t, M)$  保证目标模型和目标观测之间的相似度, 在新目标模型的基础上进行后续帧的目标跟踪.

## 5 实验结果

为了测试上述跟踪算法的有效性, 我们采用真实视频序列进行了测试, 选择的跟踪对象包括汽车、人体、人脸、茶杯等. 在没有进行特定参数调整的情况下, 系统都能够有效地完成长序列的目标跟踪任务.

图 4 和图 5 给出了一些在公开的 CAVIAR 数据集进行人体跟踪测试的结果<sup>[16]</sup>. 它们分别针对公共数据集中的“ThreePastShop2cor.mpg”序列(称为序列 A)和“OneStopNoEnter1cor.mpg”序列(称为序列 B).



Fig.4 Tracking results on video sequence A. Frames: 370, 469, 472, 489 (top row) and 506, 534, 653, 737 (bottom row)

图 4 视频序列 A 的跟踪结果. 帧数为 370, 469, 472, 489(上列)和 506, 534, 653, 737(下列)



Fig.5 Some results on sequence B. Frames: 239, 346, 401, 434 (top row) and 636, 650, 659, 686 (bottom row)

图 5 序列 B 的部分跟踪结果. 帧数为 239, 346, 401, 434(上行)和 636, 650, 659, 686(下行)

序列 A 共有 1 527 帧, 帧图像大小为 384×288 像素. 场景中的一个人被选作跟踪目标, 系统可以成功地跟踪

目标直到其被遮挡(参见 <http://www.jdl.ac.cn/user/jywang/index.htm> 上的演示序列).该序列的目的是测试多个目标进行相似运动时,系统对目标的稳定跟踪能力.在此序列中,有两个人与跟踪的客体沿同一个方向并肩行走,并且 3 个目标的大小和运动模式都比较相似;同时,在行进的过程中,3 个目标的交叉换位也增加了跟踪的难度.跟踪结果显示:系统通过动态调整目标模型,能够有效地区分两个与目标非常相近的“威胁观测”.从图 6 和图 7 中可以明显地看出动态特征选择的效果.在跟踪该序列的过程中,当另一个人接近目标时(图 4 中的 469 帧和 472 帧),从图 6 的相似性度量的分布上看,目标的周围出现了两个“威胁观测”,这些和目标模型相似性比较高的背景区域有可能造成目标混淆.通过引入动态建模的方法,可以计算出候选位置上的两个“威胁观测”,并通过目标模型的动态更新,降低模型在“威胁观测”上的相似性度量,并同时提高在目标观测上的相似性度量,从而可以在图 7 上正确地分出跟踪的对象.为了对比算法的性能,我们实现了文献[13]中提出的算法,并在序列 A 上采用相同的初始条件完成同样的跟踪问题.图 8 给出了文献[13]中算法的跟踪结果.从图中可以看到,采用基于目标/背景差别的直方图选择方法能够大概正确地定位目标的位置,但由于直方图本身描述能力的问题,算法无法判断目标在尺度上的变化,从而无法找到正确的图像区域更新目标模型,最终使该算法丢失目标.

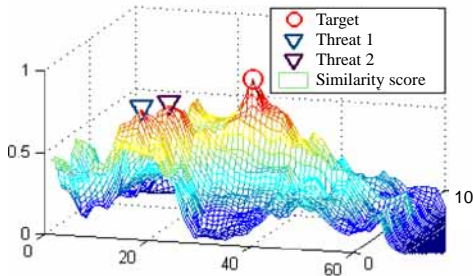


Fig.6 The distribution of similarity scores on frame 472 between the target model and image observation in sequence A

图 6 序列 A 中第 472 帧的目标模型和图像观测相似度分布

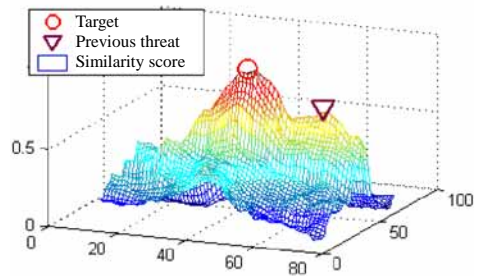


Fig.7 The distribution of similarity scores on frame 506 between the target model and image observation in sequence A

图 7 序列 A 中第 506 帧的目标模型和图像观测的相似度分布



Fig.8 Some tracking results by algorithm in Ref.[13]. Frames: 370, 469, 472, 489 (top row) and 506, 534, 653, 737 (bottom row)

图 8 采用文献[13]中算法得到的部分跟踪结果.帧数为 370,469,472,489(上行)和 506,534,653,737(下行)

序列 B 共有 724 帧,帧图像大小为 384×288 像素.由于目标行进中发生明显的旋转等姿态的改变,其图像观测的变化也很大,因此,其目的是考验长序列跟踪的能力.图 5 中的实验结果表明,系统可以成功地跟踪目标直到其步行出可见区域为止(参见 <http://www.jdl.ac.cn/user/jywang/index.htm> 上的演示序列).

在所有的实验中,帧图像都没有经过任何预处理.构成目标模型的特征大约在 120~180 个之间.实验验证:程序在没有经过特殊优化处理的情况下,能够以 10Hz 的频率完成目标跟踪任务.硬件平台为 1.5GHz Pentium 4,

内存为 256MB.

## 6 结 论

本文针对目标跟踪中的核心问题,即目标的建模与描述,提出了一种新的动态建模方法,同时维护了前景/背景的可区分性、前景描述的稳定性和计算的实时性.实验结果表明:这种方法可以有效地进行复杂背景下的长序列跟踪;同时,这种建模方法与跟踪的预测模型相对独立.尽管本文中以卡尔曼滤波作为跟踪的框架,我们在粒子滤波(particle filter)框架下的实验也得到了类似的结果.所有这些实验均表明,采用本文提出的建模方法可以全面提升跟踪算法的性能.

### References:

- [1] Regazzoni C, Ramesh V, Foresti GL. Special issue on video communications, processing, and understanding for third generation surveillance systems. Proc. of the IEEE, 2001,89(10):1355-1539
- [2] Comaniciu D, Ramesh V, Meer P. Kernel-Based object tracking. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2003, 25(5):564-577.
- [3] Gustafsson F, Gunnarsson F, Bergman N, Forssell U, Jansson J, Karlsson R, Nordlund J. Particle filters for positioning, navigation, and tracking. IEEE Trans. on Signal Processing, 2002,50(2):425-437.
- [4] Welch G, G. Bishop. An introduction to the Kalman filter. Technical Report, 95-041, University of North Carolina at Chapel Hill, 2001. [http://www.cs.unc.edu/~welch/media/pdf/kalman\\_intro.pdf](http://www.cs.unc.edu/~welch/media/pdf/kalman_intro.pdf)
- [5] Stauffer C, Grimson W. Learning patterns of activity using real-time tracking. IEEE. Trans. on Pattern Analysis and Machine Intelligence, 2000,22(8):747-757.
- [6] Stenger B, Ramesh V, Paragios N, Coetzee F, Bouhman J. Topology free hidden Markov models: Application to background modeling. In: Horaud R, ed. Proc. of the Int'l Conf. on Computer Vision. Vancouver. Vancour: IEEE Computer Society, 2001. 294-301.
- [7] Kato J, Watanabe T, Joga S, Rittscher J, Blake A. An HMM-Based segmentation method for traffic monitoring movies. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2002,24(9):1291-1296.
- [8] Rui Y, Chen Y. Better proposal distributions: Object tracking using unscented particle filter. In: Flynn P, ed. IEEE Conf. on Computer Vision and Pattern Recognition. Hawaii: IEEE Computer Society, 2001. 786-793.
- [9] Hager G, Belhumeur P. Efficient region tracking with parametric models of geometry and illumination. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1998,20(10):1025-1039.
- [10] Nguyen H, Smeulders A. Tracking aspects of the foreground against the background. In: Pajdla T, ed. European Conf. on Computer Vision, Vol.2. Prague: Springer-Verlag, 2004. 446-456.
- [11] Birchfield S. Elliptical head tracking using intensity gradients and color histograms. In: Proc. of the Computer Vision and Pattern Recognition. Santa Barbara: IEEE Computer Society, 1998. 232-237.
- [12] Stern H, Efros B. Adaptive color space switching for face tracking in multi-colored lighting environments. In: Proc. of the Int'l Conf. On Automatic Face and Gesture Recognition. Washington: IEEE Computer Society, 2002. 236-241.
- [13] Collins R, Liu Y. On-Line selection of discriminative tracking features. In: Lee SW, ed. Proc. of the IEEE Conf. on Computer Vision. Nice: IEEE Computer Society, 2003. 346-352.
- [14] Viola P, Jones M. Robust real-time face detection. Int'l Journal of Computer Vision, 2004,2(57):137-154.
- [15] Duda R, Hart P, Stork D. Pattern Classification. 2nd ed., New York: Wiley-Interscience Press, 2000.
- [16] Fisher R, et al. CAVIAR test case scenarios. 2003. <http://homepages.inf.ed.ac.uk/rbf/CAVIAR>



王建宇(1975 - ),男,辽宁营口人,博士,主要研究领域为视觉目标跟踪,目标检测,人脸识别,人脸三维合成.



高文(1956 - ),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为计算机视觉,多媒体,图像处理,多模式接口,人工智能,虚拟现实.



陈熙霖(1965 - ),男,教授,博士生导师,CCF 高级会员,主要研究领域为多模式人机接口,图像处理,模式识别,计算机视觉,数字电视技术.



赵德斌(1963 - ),男,教授,博士生导师,主要研究领域为图像/视频编解码技术及其相关应用.