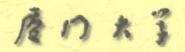
provided by Xiamen University Institutional Repository

学校编码: 10384 学号: 23020081153233 分类号\_\_\_\_密级\_\_\_\_ UDC



硕士学位论文

# 基于组件表示和协同检测的人体跟踪算法研究

Human Tracking Algorithm Based on Component Model and
Collaborative Detection

# 龚汉杰

指导教师姓名:李翠华 教授

专业名称: 计算机应用技术

论文提交日期: 2011 年 5 月

论文答辩时间: 2011 年 月

学位授予日期: 2011 年 月

2011年5月

## 厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均 在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为( )课题(组)的研究成果,获得( )课题(组)经费或实验室的资助,在( )实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名): 英汉太、 2011 年 6 月 3 日

## 厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文,并向主管部门或其指定机构送交学位论文(包括纸质版和电子版),允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索,将学位论文的标题和摘要汇编出版,采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于:

( ) 1. 经厦门大学保密委员会审查核定的保密学位论文, 于 年 月 日解密,解密后适用上述授权。

( ) 2. 不保密,适用上述授权。

(请在以上相应括号内打"√"或填上相应内容。保密学位论文应 是已经厦门大学保密委员会审定过的学位论文,未经厦门大学保密委 员会审定的学位论文均为公开学位论文。此声明栏不填写的,默认为 公开学位论文,均适用上述授权。)

声明人(签名): 龙汉杰、 2011年6月3日

## 摘要

基于视频的人的检测与跟踪是计算机视觉领域热点的研究方向之一,它在智能视频监控、军事侦察监视、交通管理、无人驾驶、视频压缩、基于内容的视频检索、人机接口、机器人控制、视频编辑等领域有着广泛的应用,并发挥着举足轻重的作用。

本文研究在背景中存在相似干扰,背景和目标的外观随着时间存在较大的变化,人体存在较大的形变等情况下的人的快速检测和跟踪:利用基于 HOG 特征和 Pictorial Structure 框架的级联分类器对视频中的人进行快速检测,并利用检测结果初始化跟踪器;通过融合人的形状约束和在线外观约束的人体模型对人体进行跟踪,并综合历史跟踪结果对目标的外观模板进行建模和更新。主要的研究工作及创新点如下:

- (1) 提出了一个基于组件表示的融合人体形状约束和外观约束的人体跟踪模型。利用 HOG 特征,基于 Pictorial Structure 框架基础上利用 SVM 离线训练出人的形状模型。在跟踪的过程中利用该形状模型约束和目标外观模板的约束来对目标的可能出现的位置进行筛选。相对传统的单靠在线维持目标的外观模板来对目标进行跟踪,本文的方法可以克服背景中跟目标具有相似外观的干扰物的干扰,确保跟踪过程中目标始终具有人的形状。另外基于组件表示的跟踪模型能够对形变较大的人体进行跟踪,对人的局部遮挡也能很好的处理。
- (2) 提出了一种综合历史跟踪结果来对目标的外观模板进行建模和更新的方法。该方法对历史跟踪结果中较好地符合人体形状约束的外观集合进行聚类,用聚类的中心来对目标的外观模板进行建模。这可以克服传统算法,单靠在线维持单个外观模板所导致模板过度更新的问题。

关键词: 人体跟踪: Pictorial Structure: 模板更新

#### **Abstract**

Video-based human detection and tracking is one of the research hotspots in the field of computer vision. It plays a very important role in many applications, such as smart surveillance, military reconnaissance and surveillance, traffic management and auto driving.

This thesis researchs how to make the machine find the human and track it automaticly in complex scene:background clutters, drastic appearance change, large deformation. We use a cascade discriminatively trained, multiscale, deformable part model to detect human in video quickly and initialize the tracker by the detection result. We propose a part based human tracking model which combine the offline learned shape constrain and appearance constrain to track the human. During tracking, we model the human appearance template by combining the history tracking results. The main researches and innovations are as follows:

- (1) We propose a part based human tracking model which combine the offline learned shape constrain and appearance constrain. Based on Pictorial Structure, we use HOG feature and Latent SVM to train the human shape model offline. During tracking, we combine the learned shape model and appearance template to get optimal configuration of human parts. Comparing to the traditional methods which only use the online maintaining appearance template, this method can overcome the distraction of background clutters and ensure that what we tracking is human body. Additionally, part based human tracking model can handle the tracking of large deformation human and partial occlusion human.
- (2) We propose a appearance themplate modeling and updating algorithm based on the history tracking results. This method collect the tracking results which is consistent with the shape constraint and then cluster the tracking results to get the cluster centers witch are used in appearance themplate modeling. This method can overcome the problem of over updating of traditional algorithm.

Key Words: human tracking, Pictorial Structure, template updating

# 目 录

摘	要		.I
Ab	str	act	II
第-	<u> </u>	章 绪论	.1
1	. 1	研究背景及意义	1
		研究现状	
		2.1 人体检测算法	
		2.2 人体跟踪算法	
1	. 3	本文的主要工作及文章结构	5
		章 人体的快速检测	
2	. 1	Pictorial Structure 模型	7
		HOG 特征	
2	. 3	人体形状模型	10
	2.	3.1 过滤器	11
	2.	3.2 可变形的组件模型	11
	2.	3.3 匹配 13	
2	. 4	人体形状模型参数的离线学习	13
2	. 5	级联的人体检测模型	14
2	. 6	级联的人体检测器阈值统计及错误率上限证明	16
2	. 7	本章实验	18
	2.	7.1 定性比较	18
			19
2	. 8	本章小结	23
第三	=1	章 形变较大的人的实时跟踪算法	24
3	. 1	基于组件表示的人体外观模型	25
		形变较大的人的跟踪模型	
3	. 3	算法加速	26
3	. 4	本章实验	29
3	. 5	本章小结	30

第四章 融合离线形状约束和在线外观约束的人体跟踪算法	31
4.1 人体跟踪模型	31
4.2 外观模板建模及更新	32
4.3 算法流程图	34
4.4 本章实验	36
4.5 本章小结	36
第五章 系统性能评测	37
5.1 本文测试视频介绍	37
5.2 评价标准	38
5.3 实验结果	38
5.4 小结	54
第六章 总结和展望	
6.1 总结	55
6.1 总结6.2 展望	56
参考文献	
研究生期间发表的论文	63
致谢	
F V V 1788	

# Contents

摘	要	·	I
Ał	str	act I	Ι
Cł	ap	ter 1 Introduction	1
	1. 1	Research Backgrounds	1
•	1. 2	Research Status	1
	1.	2.1 Human Detection Algorithm	1
	1.	2.2 Human Tracking Algorithm	4
	1. 3	Contribution and Outline	5
Cł	ap	ter2 Human Detection	7
2	2. 1	Pictorial Structure Model	7
2	2. 2	HOG Feature	8
2		Human Shape Model	
	2.	3.1 Filter1	1
		3.2 Deformable Part Model	
	2.	3.3 Match1	3
2	2. 4	Learning	3
2	2. 5	Cascade Human Detector1	4
2	2. 6	Pruning Thresholds1	6
2	2. 7	Experimental Results1	8
	2.	7.1 Qualitative Evaluation	8
	2.	7.2 Quantitative Evaluation	9
	2. 8	Summary	3
Cł	ap	ter3 Large Deformation Human Tracking Algorithm2	4
3	3. 1	Part Based Human Appearance Model 2	5
3	3. 2	Large Deformation Human Tracking Model2	5
3	3. 3	Algorithm Accelerate	6
3	3. 4	Experimental Results	9
•	3 5	Summary 3	ი

#### Contents

Chapter4 Collaborative Tracking by Detection	31	
4.1 Human Tracking Model	31	
4.2 Appearance Modeling and Updating	32	
4.3 Algorithm Flowchart	34	
4.4 Experimental Results	36	
4.5 Summary	36	
Chapter5 System Performance Evaluation		
5.1 Dataset		
5.2 Evaluation Criterion	38	
5.3 Experimental Results	38	
5.4 Summary	54	
Chapter6 Conclusions and Prospect		
6.1 Conclusions of This Thesis	55	
6.2 Suggestions in the Future	56	
References	57	
Publications		
Acknowledgement		

## 第一章 绪论

#### 1.1 研究背景及意义

在视频和数字图像处理和计算机视觉领域之中,人的检测和跟踪技术是一项最基本的任务,具有广泛的应用。在视频挖掘领域,以人的检测和跟踪算法为基础,可以对公共场合中,每天人的活动进行概括和整理。在人机交互(Human-Computer Interaction)领域,基于人体姿势的人机接口有助于实现智能办公和智能家居。在安全监控领域,智能监控系统使用人的检测和跟踪技术对场景中可疑的人进行自动检测、跟踪、识别和报警,从而达到对犯罪的阻止。在视频压缩和检索领域,以人的检测和跟踪技术为基础能从背景中自动分割出人,从而为人的识别和标注做准备,最终用于基于内容的视频检索和视频组织。另外一些较流行的应用还包括汽车自动辅助驾驶系统、机器人控制、视频编辑等。

正是由于人体检测和跟踪技术有如此广泛的应用范围,在计算机视觉领域已经被深入研究 30 多年。尽管这些年,人的检测和跟踪算法都有显著的进步,具有里程碑意义的比如:在人的检测方面,HOG 特征<sup>[1]</sup>被提出来描述人体的形状。在跟踪方面,粒子滤波跟踪算法框架<sup>[2]</sup>被提出来用于处理非高斯分布的背景噪声场景下的目标跟踪。但在真实的复杂场景中,现有的算法仍面临着诸多的挑战:目标外观存在较大的变化,如光照变化导致的目标外观变化剧烈,目标各个组件形变后的相互遮挡导致的组件外观变化;背景也随着时间发生比较大变化,且背景中存在很多跟目标外观相似的干扰。因此现有的人体检测和跟踪系统往往在比较多的假设(比如背景固定,目标外观固定,人体直立且形变较小)下才能取得比较好的效果,人的检测和跟踪技术仍是一个活跃的研究领域。

#### 1.2 研究现状

#### 1.2.1 人体检测算法

现有的关于人体检测的文献的贡献,可以分为以下两类: 1)使用什么样的 图像描述子或则特征集; 2)基于什么样的检测框架。我们通过这两个分类来对现 有的人体检测文献进行回顾。

#### 1.2.1.1 图像特征

我们把现有的人体检测中使用的特征集合分为 2 大类: 1) 基于点,图像块,或则目标组件的稀疏表示特征; 2) 基于图像亮度和梯度的稠密表示特征。

#### 稀疏表示

稀疏表示是基于某些相关的局部区域的描述,这些相关的局部区域可以用兴趣点检测子或则关键区域检测子得到。

#### (1) 兴趣点检测子

利用显著的局部点来做人体检测<sup>[3]</sup>是一类被实际证明有效的方法。该类方法首先用兴趣点检测子来检测局部显著点(又叫做兴趣点),然后对检测出来的局部显著点进行描述。最后从这些兴趣点的描述子中生成特征向量进而进行人体检测器的构建。该类人体检测器的检测性能,取决于所使用的兴趣点的可靠性,准确性和重现能力。经常被使用的兴趣点检测子包括: Harris<sup>[4]</sup>,Laplacian<sup>[5]</sup>,Difference of Gaussians(DoGs)<sup>[6]</sup>和尺度不变的 Harris-Laplace<sup>[7]</sup>。基于兴趣点稀疏表示的方法的一个优点在于对目标的表示比较简洁,特征向量的维度相比单纯的利用图像像素要小得多。

#### (2) 基于组件的表示

基于组件的人的表示被广泛地用在人体检测系统中<sup>[8-11]</sup>。他们使用显式的人体组件:上臂、前臂、大腿、小腿、躯干、头。每个组件用两条平行线段来表示,利用平行边缘检测器来检测组件。并且利用各个组件之间的空间关联先验关系来约束组件之间的位置。这类方法的主要缺点在于,在背景纹理比较复杂的情况下,将有很多干扰的边缘,使该类检测器容易产生误检。

#### 稠密表示

另一类方法是对整个检测窗口中的图像进行稠密(通常是挨个像素)的特征抽取。然后把该窗口里的特征值,按照一定顺序,串联成高维的特征向量用于表示目标。该类方法常采用基于图像亮度和图像梯度的稠密表示。

该类特征的典型代表是 Dalal<sup>[1]</sup>提出的 HOG 特征,首先计算整个检测窗口中的每一个像素的梯度幅度和梯度方向,然后用网格把整个检测窗口划分为一个一个正方形区域(称为 Cell),计算每个区域内梯度方向直方图。接着对每个 Cell 的梯度方向直方图进行归一化处理,最后把所有 Cell 的梯度直方图串联成一个高维特征向量用于表示人体。Dalal<sup>[1]</sup>利用 HOG 特征和 SVM 学习方法获得了

VOC2005 人体检测竞赛的第一名。该特征被后来的文献广泛引用,被认为是目前为止用于人体检测的最好特征。Felzenszwalb<sup>[12]</sup>在 Dalal<sup>[1]</sup>的基础上,采取基于组件表示的方法获得了 VOC2010 人体检测的最好结果。

#### 1.2.1.2 分类方法

人体检测问题本质上是一个把图像块分为人和非人的二分类问题。所以现有的人体检测系统都是基于分类的方法。设输入的图像特征为 $\mathbf{X}$ ,类标签为 $\mathbf{Y}$ ,则如果一个方法学习的是条件分布  $\mathbf{p}(\mathbf{y}|\mathbf{x})$ ,则该方法就是基于判别式的;如果一个方法学习的是联合分布  $\mathbf{p}(\mathbf{x},\mathbf{y})$ 则该方法就是基于产生式的。

#### 基于判别式

机器学习方法如: SVM<sup>[13, 14]</sup>, Boosting<sup>[15]</sup>是近年来两种最流行的基于判别式学习的分类方法。他们的优点在于: 1) 自动地从高维的特征向量中挑选出具有判别性的特征向量子集; 2) 他们被广大文献的实验证明分类性能优越; 3) 他们的使用相对简单。

SVM 试图寻找这样一个超平面,使得无论在输入特征的特征空间还是在输入特征被变换以后的特征空间中,目标类和非目标类之间的间隔达到最大。 Mohan<sup>[16]</sup>建立一个两级的 SVM 分类器用于人体检测:第一级利用 Haar 类特征训练人体组件分类器如:头,上臂等等;第二级基于组件分类器的输出训练得到最终的人体检测器。

AdaBoost 方法组合一组弱分类器的结果形成一个强分类器。在计算机视觉领域,它通常被用于建立一个级联的非目标否决器:在每一级挑选出最能判断输入是非目标的特征,然后依据这组特征,对输入进行级联否决。当一个输入能够通过最后一级的分类器,则判定输入为目标,否则被任何一级拒绝则输入被判断为非目标。AdaBoost 级联分类器相对 SVM 分类器的特点是:训练速度虽然慢,但是由于采用对输入的级联否决,能迅速排除非目标,所以在分类阶段相比基于SVM 的分类速度快。Viola<sup>[17]</sup>利用 AdaBoost 学习算法,训练出一个检测速度快,精度高的行人级联检测器。

#### 基于产生式

基于产生式的模型,常见的有高斯混合模型,贝叶斯产生式模型,HMM, 马尔科夫随机场等。

#### 1.2.2 人体跟踪算法

一个鲁棒的人体跟踪器需要满足以下条件: 1) 在背景固定和非固定情况下都能够自动初始化; 2) 在长时间内准确跟踪; 3)能对多个目标进行跟踪; 4)能够处理遮挡; 5)能克服背景中相似外观的物体的干扰; 6) 计算复杂度低, 能达到实时要求。但是直到目前为止还没有能够满足以上全部要求的跟踪系统。本节对现有的最好的人体跟踪算法进行回顾。

直到目前为止最成功和使用最多的人体跟踪算法<sup>[18-28]</sup>是基于隐马尔科夫模型:其中隐变量为目标的状态,用X,表示,是待估计的变量;观测量是视频中的图像序列I,。基于马尔科夫假设我们可以把联合分布分解如下:

$$P(X_{1:T}, I_{1:T}) = \prod_{t} P(X_{t} \mid X_{t-1}) P(I_{t} \mid X_{t})$$
(1.1)

其中 $X_{1:T}$ 是 $\{X_1,...,X_T\}$ 的缩写。跟踪过程就是在该概率模型上进行推理的过程。典型地,给定图像序列后和目标的初始状态后,目标状态序列的最大后验概率(MAP)可以由下面的式子得到:

$$\hat{X}_{1:T} = \underset{X_{1:T}}{\arg \max} P(X_{1:T} | I_{1:T}) = \underset{X_{1:T}}{\arg \max} P(X_{1:T}, I_{1:T})$$

$$= \underset{X_{1:T}}{\arg \max} \prod_{t} P(X_{t} | X_{t-1}) P(I_{t} | X_{t})$$
(1.2)

基于该模型的人体跟踪算法涉及到 3 个方面: 1) 推理算法是什么? 2) 动态转移模型  $P(X_t | X_{t-1})$  是什么? 3) 似然项  $P(I_t | X_t)$  怎么定义? 下面根据这 3 个方面对现有的人体跟踪算法进行回顾。

#### 1.2.2.1 推理算法

目标的状态空间(X,的可能取值空间)巨大,以至于用搜索算法(比如动态规划搜索算法)进行穷尽搜索的时间开销巨大而不能实际被使用。早期的文献使用经典的人工智能搜索算法进行近似搜索<sup>[28,29]</sup>。后面的绝大多数文献都使用卡尔曼滤波(及其变种)<sup>[18,21,24]</sup>或则粒子滤波(及其变种)<sup>[2,19,22,25-27,30-33]</sup>进行推理。推理过程为使用当前帧的状态和动态转移模型预测目标在下一帧的状态;接下来利用图像数据根据似然项来对预测的目标状态进行改善。粒子滤波跟卡尔曼滤波的差别在于:粒子滤波预测目标在下一帧出现在多个位置,而卡尔曼滤波只预测目标在下一帧出现在一个位置。粒子滤波通过对本帧得到的目标状态后验概率进行采样得到一些粒子,然后用动态转移模型对这些粒子的状态进行转移,转移之

后得到的粒子的状态就是下一帧目标可能出现的状态,多个粒子状态代表多个预测。然后利用下一帧的图像数据根据似然项对粒子的权重进行调整。

#### 1.2.2.2 动态转移模型

人体是非刚性目标,其运动包括整体的运动和各个组件(头、上臂、前臂、大腿、小腿、躯干)的运动,所以构建精确地动态转移模型是非常复杂。Sidenbladh <sup>[25,26]</sup>试图从图像序列集合中学出人的动态转移模型,但是该类方法往往只能学特定一种类型的人的运动比如:走路,跑步等。通用的参数化人体动态转移模型 <sup>[34-36]</sup>,近年来被广泛的研究,但是至今仍没有一个通用的人体动态转移模型被提出来。

#### 1.2.2.3 似然项

在跟踪过程中,当利用动态转移模型预测出下一帧目标的位置的时候,需要利用似然项 $P(I_t|X_t)$ 计算出目标出现在预测位置的概率。通常的做法是,为目标维护外观模板,根据预测位置上的图像与目标外观模板在某种度量下的距离来计算目标出现在该预测位置的概率。可以使用的目标外观信息包括目标的外轮廓形状,目标的纹理,目标的颜色等信息。通常单独使用一种信息或则混合使用多种信息来对目标外观建模。如 Gavrila [37]使用目标的外轮廓形状和目标纹理信息来对目标的外观建行建模。Ioffe [38]使用亮度和边缘信息。Mori [39,40]使用形状上下文信息。

## 1.3 本文的主要工作及文章结构

本文研究在背景中存在很多相似干扰,背景和目标的外观随着时间存在较大的变化,人体存在较大的形变情况下的人的快速检测和跟踪:利用基于 HOG 特征和 Pictorial Structure 框架的级联分类器对视频中的人进行快速检测;构建一个新的人体跟踪模型,该跟踪模型结合离线训练得到的人的形状约束和在线更新的外观模板约束,使在对目标的跟踪过程中能克服背景中与目标外观一致的干扰物的干扰,另外该跟踪模型基于组件表示,对组件的空间关系进行显示建模,能对形变较大的人体进行跟踪;在跟踪过程中收集较好满足离线人体形状约束的帧的跟踪结果,通过颜色信息聚类,最后构建基于这些类中心的外观模板。

本文共六章, 具体结构如下:

第一章: 绪论。简要地阐述了研究课题的研究背景和意义、研究现状、本文的主要工作及文章结构。

第二章:人体的快速检测。介绍了Pictorial Structure模型,HOG特征及基于这两者的人体形状模型。采用Latent SVM离线训练出人体形状模型的参数。构建级联的人体检测模型,利用正样本集统计出各级的阈值,并给出这种方法的错误率上限证明。最后利用实验定量验证了我们实现的人体检测系统具有高的检测率,低的误检率等特点。

第三章: 形变较大的人的实时跟踪算法。提出了一个基于组件表示的人体跟踪模型,该模型对各个组件之间的空间位置约束关系进行显式建模。并利用该模型对形变较大的人体进行跟踪,为了进一步提高跟踪过程的实时性,我们对算法过程进行加速。最后利用实验结果,定性地分析我们提出的人体跟踪模型对于形变较大人的跟踪的有效性。

第四章:融合离线人体形状约束和在线外观约束的人体跟踪算法。为了克服跟踪过程中背景中存在的跟目标外观相似的物体的干扰,在第三章的人体跟踪模型的基础上提出了一个融合离线人体形状约束和在线外观约束的人体跟踪模型。为了对人体进行长时间的鲁棒跟踪,提出了一个综合历史跟踪结果的外观模板构建方法,并对外观模板进行在线更新。最后通过实验定性地对本章的方法,第三章的方法进行对比。

第五章:系统性能测试。介绍了本文系统测试所用的实验视频的来源,以及每个测试视频的特点。介绍了算法性能的评价标准。最后对本文的算法与Shahed Nejhum<sup>[41]</sup>方法的性能进行定性和定量的对比。

第六章: 总结和展望。总结了本文的主要工作,并提出了下一步工作和设想。

## 第二章 人体的快速检测

基于组件表示的目标检测方法被广泛地用于处理非刚性目标的检测,它具有能处理局部遮挡和物体变形的能力。人体作为一种非刚性目标,其各个组件(头,躯干,上臂,前臂,大腿,小腿)的不同姿势组合使得人体有各种各样的形变。Pictorial Structure 模型<sup>[42]</sup>显式地对目标的各个组件的位置关系进行建模,能有效地解决形变物体的检测。Dalal<sup>[1]</sup>提出来的 HOG 特征被认为是至今为止对人体形状表述最好的特征。Felzenszwalb<sup>[12]</sup>就是在 Dalal<sup>[1]</sup>的基础上利用 Pictorial Structure 取得了目前为止人体检测的最好结果。本章对 Felzenszwalb<sup>[12]</sup>进行介绍,对其算法进行快速实现,与 Opencv2.1 的人体检测系统(对 Dalal<sup>[1]</sup>的实现)进行性能对比,试图说明采取 Pictorial Structure 模型的优点。

#### 2.1 Pictorial Structure 模型

Pictorial Structure 表示模型由 Fishchler 和 Elschlager  $^{(42)}$ 在 1973 年提出来,由 Felzenszwalb  $^{[43]}$ 在 2005 年把它引入到视觉领域,用于目标识别。他们把物体表示成一堆组件的集合,并且这些组件之间满足一定的空间位置约束关系。每个组件对应物体的一个部分,组件跟组件之间的空间位置约束关系对应物体两个相应部分的关系。如图 2.1 所示,对人来说,他的组件为:左右上臂,左右前臂,左右大腿,左右小腿,躯干,头部等 10 个组件,相连的组件围绕着连接点可以在一定角度下旋转。基于该模型可以从图片中挑出使得匹配该模型的代价函数值最小的图像块,该图像块即是由该模型刻画的目标。该模型的代价函数包括:每个组件的匹配代价和相连组件之间的变形代价。该模型的通用性很强,因为对于不同的目标,组件的外观模型可以不同,组件之间的变形模型可以不同。该模型可以表示成一个无向图 G=(V,E),其中顶点集  $V=\{v_1,...,v_n\}$  对应目标的 n 个组件,边  $(v_i,v_j)\in E$  对应组件  $v_i,v_j$  的连接关系。一个目标的实例由配置  $L=\{l_1,...,l_n\}$  刻画,其中  $l_i$  确定组件  $v_i$  在图像中的状态,包括位置,大小等。在图像中需找最优配置由最小化代价函数定义。对于一个具体的配置,它的代价取决于两个方面:1)模型中  $v_i$  的外观模板与由  $l_i$  所确定的图像块的匹配程度。2)  $l_i$  所定义的图像

Degree papers are in the "Xiamen University Electronic Theses and Dissertations Database". Full texts are available in the following ways:

- 1. If your library is a CALIS member libraries, please log on <a href="http://etd.calis.edu.cn/">http://etd.calis.edu.cn/</a> and submit requests online, or consult the interlibrary loan department in your library.
- 2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

