中山大学硕士学位论文

**基于组合模型学习的人脸视频**

**线描定位与传播**

**Sketch Localization and Propagation System based on Compositional Model Learning**

学位申请人： **伍少茹**

指导教师： **林倞 副教授**

专业名称： **计算机技术**

答辩委员会主席（签名）：

答辩委员会委员（签名）：

二零一三年三月八日

# 论文原创性声明

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究作出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

学位论文作者签名：

日 期：

# 学位论文使用授权声明

本人完全了解中山大学有关保留、使用学位论文的规定，即：学校有权保留学位论文并向国家主管部门或其指定机构送交论文的电子版和纸质版，有权将学位论文用于非赢利目的的少量复制并允许论文进入学校图书馆、院系资料室被查阅，有权将学位论文的内容编入有关数据库进行检索，可以采用复印、缩印或其他方法保存学位论文。

学位论文作者签名： 导师签名：

日期： 年 月 日 日期： 年 月 日

论文题目：基于组合模型学习的人脸视频线描定位与传播

专 业：计算机技术

硕 士 生：伍少茹

指导教师：林倞 副教授

# 摘 要

人脸定位一直以来都是计算机视觉、图像科学和模式识别领域的热门课题。人脸定位不仅是人脸识别和表情分析过程中的重要一环，同时也是人脸美化和肖像风格化等应用方向的基础性工作。虽然人脸定位领域的研究已有一定的历史，但是目前还没有一套完备的理论和普遍适用的算法能够处理自然的、多角度、多表情的人脸照片。因此，人脸定位算法的研究还存在广阔的研究空间。

本论文针对目前算法存在的问题进行深入研究，并把一种改进的人脸定位算法应用到视频肖像风格化系统中，使得视频剪辑中的人脸能够以线描动画的形式展示。系统采用基于层次分解模型的人脸解析算法进行人脸定位，使用自顶向下，自底向上和层间约束的三通道模型。这样的层次分解模型降低了目标照片的输入要求，使得算法能够处理带表情、多角度的自然照片，因此更适用于人脸变化较多的视频风格化系统。在线描风格处理模块中，系统根据人脸定位的结果，在数据库中按部件选取最匹配的线描模板，然后使用Warping算法把模板适应到部件的实际形状。在视频处理模块中，考虑到系统的性能，系统按照一定的时间间隔对关键帧进行人脸定位，然后把当前帧的特征点映射到下一帧来完成两个关键帧之间的帧间传播。此外，系统还对每个人脸部件的变换进行约束，定义了一系列人脸部件状态机，保证了视频传播的连贯性和流畅性。

线描动画系统能够生成一系列表情丰富的视频剪辑，使用户能够在不露面的情况下表达出自己的情感，制作出属于自己的线描艺术视频，有着广泛的应用前景。

**关键词：**人脸定位、主动形状模型、层次分解模型、模板匹配、薄板样条、线描

Title: Sketch Localization and Propagation System based on Compositional Model Learning

Major: Computer Technique

Name: Shaoru Wu

Supervisor: A/Prof. Liang Lin

# Abstract

Face alignment is one of the hotspots of computer version, image science and pattern recognition in recent years. It is not only the important part of faces recognition and expression analysis, but also the basic work of facial beautification and portrait stylization. Although face alignment has been studied for years, however, there is still no perfect algorithm for processing natural photos with multi-angle and various expression.

In this dissertation, the principle of classic face alignment is deeply analyzed and studied. To improve the aligning robustness, we adopt a novel algorithm to extract facial features and apply it in generating portrait sketches from video clips. The improved method, a hierarchical compositional model, is a three-layer searching scheme including top-down and bottom-up channels and blinding constraints。It can align faces from natural photos without a requiring initialization, so it is quite suitable for the portrait video system in which people move their heads with various facial expression. In the sketching process, the system picks the best-match template from the artistic library and warps it to the facial component. In the video propagation, faces are only aligned in the key frames within a time interval. Sketches propagate to the next frame based on the key point in the current frame. In addition, a series of state machines are defined to enforce video temporal-spatial consistency.

By using this system, users can freely generate their portrait sketch animation to express emotion behind their funny avatars. This system is quite interesting and full of wide-ranging practicality and application prospect.

**Keywords:** Face Alignment, Active Shape Model(ASM), Hierarchical Compositional Model, Template matching, Thin-plate Splines, Sketch

# 目 录

从奇数页开始，双面打印

1、目录保留2级标题

2、自动生成目录后把两个声明这两行手工删除掉，即目录从摘要开始

3、目录采用小四号宋体，不加粗

4、段落间距为1.25倍行距

摘 要 I

Abstract II

目 录 III

引 言 1

第 1 章 保持特征的网格变形方法综述 4

1.1 泊松网格编辑 4

1.2 物理似然的模拟 4

1.3 基于勾勒式界面的编辑方法 4

1.4 本章小结 4

第 2 章 网格顶点的微分坐标表示 5

2.1 网格的局部几何特征 5

2.2 顶点的离散微分坐标 5

2.3 线性约束的变形算法 6

2.4 本章小结 6

第 3 章 线性约束网格变形方法 7

3.1 保持局部特征的线性约束条件 7

3.2 仿射约束的有效性 7

3.3 局部刚性的线性约束 7

3.4 实验结果 8

3.5 本章小结 8

第 4 章 基于曲线匹配的勾勒式网格编辑 9

4.1 曲线对齐的背景 9

4.2 多分辨率曲线对齐算法 9

4.3 实验结果 9

4.4 本章小结 9

第 5 章 总结与展望 10

5.1 研究成果 10

5.2 将来的工作 10

参考文献 11

附 录 13

作 者 简 历 14

致 谢 15

（硕士的研究点、方法论创新和应用点总和不少于三个；

博士的研究点、方法论创新不少于三个，应用点不少于一个）

# 概述

在传统的互联网信息传播中，文本是人们进行信息交流中最主要的方式。但随着社会的进步和生活质量的提高，人们渴望使用更为丰富的方式进行自我表达和情感抒发，简单的文本信息并不能满足人们的通信需求。现今的网络带宽能够很好地支持语音、图像、视频等多媒体信息的传播， 互联网的主流信息正从文本向多媒体过渡。大量学者纷纷投入到图像科学和计算机视觉等领域的研究中，为学术研究注入了新的思想和新的能量，促进了该领域的蓬勃发展。其中，人脸科学是计算机视觉领域的热门话题之一。但由于人脸相对于其他自然物体拥有更为丰富和复杂的信息，不同场景下、不同表情的人脸所包含的信息也不尽相同，变化多端的人脸使得计算机难以分析和处理，因此，人脸科学的相关研究仍十分有意义。本论文从人脸定位的研究着手，力求把鲁棒的人脸定位算法应用到视频当中，并把视频进行艺术渲染，制作出具有线描风格的人脸视频动画系统。

本章首先对人脸定位的研究背景、意义和应用前景做整体概述，然后就国内外的研究现状做出分析，并介绍了本文的研究内容和主要工作，最后给出本论文的内容安排。

## **人脸定位的研究背景与意义**

* + 1. **人脸定位的概念**

人脸定位，又称人脸对齐或人脸特征点提取，是指对人脸的整体形状和局部特征进行定位，精确提取出人脸轮廓线和人脸五官(如眉毛、眼睛、鼻子和嘴巴等)的位置信息。通常用点的集合来刻画人脸对象的位置和形状信息。标定的特征点如下图 1‑1所示。



(b)PDM模型定位

(a)原始图像

图 1‑1 人脸定位示意图

人脸定位常用于人脸识别、肖像艺术化、人脸跟踪、表情分类等方向，是人脸科学中一项重要的预处理工作。例如，人脸识别的过程通常包含人脸检测、人脸定位和人脸识别三个主要环节[1]，而目前人脸检测和人脸识别模块都有独立的工作能够较好地处理自然图片，但是人脸定位在处理多角度、复杂的光照条件下的人脸，却没有一个普遍适用的算法能够输出稳定、鲁棒的结果，人脸定位这一环显得相对薄弱[2]。人脸特征点提取的准确程度直接影响后续图像处理工作的可靠性和稳定性，因此人脸定位是人脸科学应用的关键问题。

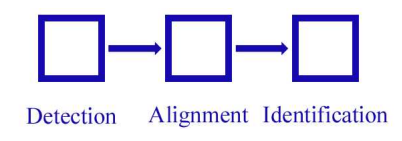


图 ‑2 人脸识别的过程

* + 1. **人脸定位的应用**

人脸定位是人脸科学及其相关领域的基础工作。目前，人脸定位已广泛应用于视觉艺术、信息安全、计算机辅助建模等多个方面，其中包括：

1. **身份认证与人脸追踪**

与指纹、视网膜、虹膜、基因、掌形等其他人体生物特征识别系统相比，人脸识别具有非侵犯性以及直接、友好和方便的特点。它是人们最容易接受的一种身份鉴别方式。由于人脸识别技术在生物特征识别技术中具有众多的优势，决定了其广泛的应用前景，如刑侦破案、证件验证、出入口控制以及视频监视等。[3]人脸识别技术是国家安全、公共安全等问题中的一项重要技术，而好的人脸定位效果能够更好地保证人脸识别结果的准确性，提高人脸追踪结果的稳定性。

1. **表情识别**

人脸表情识别是人机自然交互、计算机视觉、情感计算和图像处理等研究领域的一个热点课题，在人机交互、远程教育、安全领域、智能机器人研制、医疗领域、电脑游戏等领域有着广泛的应用[4]。美国心理学家 Ekman 和Friesen把人脸表情分为惊奇、恐惧、厌恶、愤怒、高兴、悲伤这6 种基本表情，并于1987年开发了面部动作编码系统[5]。人脸定位是表情识别过程中的重要一环，人脸特征点提取的过程相对于把丰富而复杂人脸信息进行降维处理，从人脸模型反应人脸的全局情况。

1. **肖像风格化**

在计算机数字艺术方面，人脸定位是肖像风格化的制作中必不可少的一环。由于人脸模型所在位置都是人脸中的一些边缘，包含了大量的信息，这些人脸区域与其他物体在艺术上的表现手法不同，因而做渲染处理的时候也不尽相同。而随着数码摄影设备的进步，现在几乎每人手上都有一台能拍照的相机或是一台数码相机，人们越来越渴望能够成为自己的御用绘画师，能够根据肖像照做出具有艺术风格的作品。现今已有许多电脑软件和移动终端软件能够对人像进行艺术处理，但大多数产品都基于滤波处理和边缘检测的方式，无法准确、单一地处理人脸；或是基于五官模板替换的方式，需要进行大量的交互操作，无法自动地完成渲染，手动的操作和模板数量的限制使得肖像画与照片之间的相似度下降，如Keyloft公司制作的手机软件imadeface（图 1‑2）

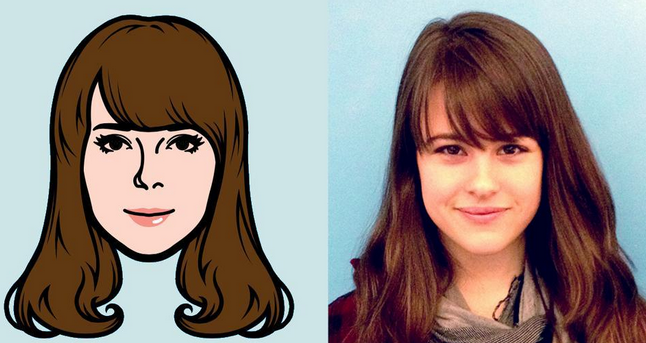


图 ‑3 imadeface软件需要用户交互地选择五官模板

1. **游戏世界的人脸建模**

游戏产业长期以来都是互联网中最为盈利的产业。为了吸引更多玩家，开发商们也越来越关注用户体验度。目前，人类面部表情捕捉已经大量应用到视频游戏产业中。比如著名的黑帮游戏《黑色洛城》（LA Noire），在其动画制作过程中采用了脸部动态捕捉技术，将演员的脸部表情和神态转换到游戏的NPC上，通过这样来完成角色建模，减少了游戏动画制作的难度。此外，微软Kinect[[1]](#footnote-1)开发的Faceshift能够捕捉面部表情并实时地映射到电玩人物上，玩家本人和其游戏中的化身拥有相似的嘴脸，使虚拟世界的人物栩栩如生，更富有生命力。



图 ‑4 Avatar Kinect系统能够跟踪用户的头像，包括微笑、皱眉、笑容甚至是眉毛的运动

1. **计算机辅助的电影制作**

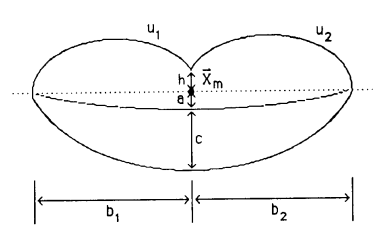
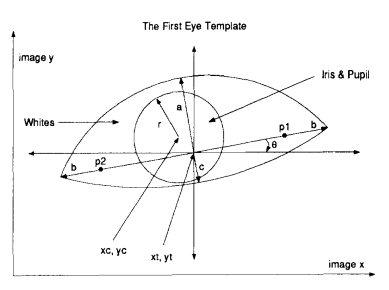
现今的电影正慢慢从2D到3D演进，一出好的电影离不开电影制作人的电脑剪辑和特效，更离不开主角丰富的表情演绎。在很多情况下，电影需要把人的表情赋予到动物或者科幻人物中，因此需要让演员预先录制一段剪辑，再通过电脑处理把人脸映射到角色上。在这过程中，准确的人脸定位效果至关重要。利用计算机辅助生成电影或动画，大大地降低了制作成本，好的剪辑技术还能够令电影生色，提供起艺术价值和观赏价值。

## **国内外研究现状**

人脸定位的研究

Kass[6]等人早在1987年提出了一种可变模型的方法来进行人脸特征点提取，这种方法称为主动轮廓模型（Active Contour Model，也称为“Snake”模型）。Snake模型的基本思想是：通过最小化能量函数来找到最优解。该能量函数是一个包含内力和外力两项的曲线函数，内力对模型的形状进行约束，促使轮廓光滑；外力根据图像的特征引导模型移动，寻找局部极值来最小化能量函数。Snake模型之所以是“主动的”，是因为它能够自主地收敛到能量极值点。但由于Snake模型不带先验信息，算法不针对某一特定的对象，在没有先验信息的指导下容易受噪声影响，陷入局部最优。

1992年，Yuilie[7]等人提出了可变模型方法来抽取脸部特征。在可变模型中，人脸模型是由一组带参数的几何模型组成，如眼睛用上下两条抛物线加上眼珠的圆形来表示，而闭合的嘴巴由四条抛物线表示(如图 1‑4)。不同的脸部特征其能量函数中项的权重都不同，定位过程通过优化能量函数来搜索最匹配的位置。但是由于自然界中人脸的五官特征并不是通过简单的几何模型就能够准确定位的，因此可变模型难以在人脸定位中展示其鲁棒性。



1. (b)

图 ‑5 可变模型中的眼睛(a)和闭合的嘴巴(b)

在Snake模型的基础上，Cootes[8]等人于1995年提出了主动形状模型(Active Shape Model, ASM)。针对Snake模型的缺点，ASM模型中加入了目标对象的先验信息，通过训练目标物体得到模型内部的约束关系，并用点分布模型(Point Distributed Model,PDM)来描述人脸特征。在搜索过程中，算法根据模型的局部纹理信息来引导模型前进，通过不断更新模型形状和多次迭代使结果收敛，搜索的大致过程如图 1‑5。ASM模型的主要思想和算法原理将在2.1节中详细介绍。

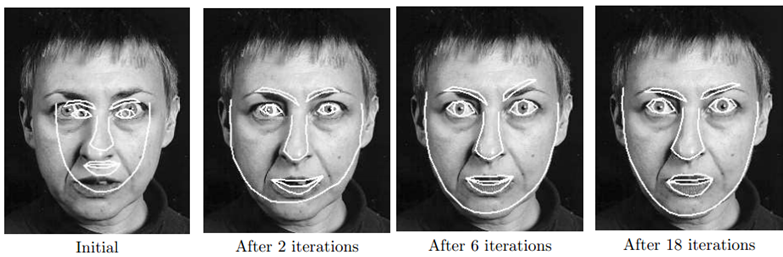


图 ‑6 ASM的搜索过程

其后于1998年，Cootes等人在ASM模型的基础上加入人脸纹理信息，提出了主动表观模型(Active Appearance Model,AAM)[9]。AAM模型包含ASM模型中的形状模型和全局的人脸纹理色彩模型，两者组合成联合纹理模型来描述人脸。类似于ASM的搜索算法，AAM的搜索过程也可看作是一个参数优化的过程，通过不断优化和更新表观参数，使得图像目标和模型当前合成的实例之间误差值最小[10]。AAM算法的搜索过程可以看作是表观模型通过扭曲变换的方式匹配目标图像的过程（图 1‑7 AAM的搜索过程）。



图 ‑7 AAM的搜索过程

ASM和AAM是人脸定位算法中的经典，在其后的十多年里，大多数人脸定位算法都是在ASM算法上改进。焦峰等人借助Gabor小波的理论来解决相应的问题，算法采用Log-Gabor小波子系数特征完全代替灰度信息对每个特征点周围纹理信息进行描述并建立纹理模型。Viola等采用AdaBoost算法构建层叠式人脸检测器，是人脸检测算法上的重大突破。随后张震球等人对这一方法加以改进，并建立了一个实时的多视角人脸检测系统。AdaBoost算法的基本思想是在给定的训练集上，反复选出关键的分类特征并扩建弱分类器集，最终强分类器由这些弱分类器线性组合构成。算法采用顺序前进法搜索策略，尽管每次迭代选择的弱分类器是局部最优，但最终构成强分类器的弱分类器及其系数并不是最优。

总体而言，目前的人脸定位算法对光照条件，人脸姿势、表情以及人脸模型初始位置有很高的要求。一般的定位算法只能在类似证件照等正脸角度、光照良好、无表情、无遮挡的照片上才有较好的效果。人脸定位仍旧面临诸多挑战性的问题，因此如何有效提高人脸定位的精确性、鲁棒性和效率仍将是相关领域内的重点研究内容。

## **本文的研究工作**

本文将从以下几个方面进行深入研究，将改进的人脸定位算法应用到视频肖像风格化系统中，使得视频剪辑中的人脸能够以线描动画的形式展示。

* + 1. **人脸定位**

人脸定位是本文的主要研究内容。首先，本文对经典的ASM算法进行深入研究，并在此基础上实现了一种改进的基于层次分解模型的人脸解析算法[11]，该算法能够对自然的、多姿态的人脸准确定位，而且处理速度快，因此更适合应用在视频系统中。此外，系统允许用户通过微调的方式来调整人脸定位的结果，使视频在传播前有较为精确的人脸初始位置。

* + 1. **肖像线描**

系统将根据人脸定位的结果，对脸部特征点进行线描的处理，并在艺术模板库[12]中搜索最匹配的五官、头发、衣领等模板，通过形变使模板与实际人脸更为相似。系统除了能够自动生成较为相似肖像线描，还允许用户手动选择模板来定义自己的阿凡达[[2]](#footnote-2)，并以此作为动画的主角。

* + 1. **视频一致性**

线描人像生成后，系统利用薄板样条(Thin-plate Spline, TPS)变换[13]的方法实现动画的传播。为了保证视频能够流程地播放，系统设计了一系列的策略来对人脸特征的传播方式进行约束，其中包括Zijian Xu[14]提出的脸部特征状态机。状态机规定了人脸五官转换的状态，对脸部特征在视频中传播进行约束，减少了成像的抖动。此外，系统每一定的时间间隔进行人脸定位，保证了动画传播的可靠性。

## **论文结构**

以下是本文的章节安排：

1. 第一章-概述 本章介绍了人脸定位及其相关工作的研究背景和应用前景，然后就国内外的研究现状做出分析，并介绍了本文的研究内容和章节安排。
2. 第二章-相关技术的研究 本章主要对本系统所用到的相关技术进行研究。其中包括了传统的主动形状模型、改进的基于层次分解模型的人脸解析算法、sketch线描技术、薄板样条(TPS)变换、视频传播的一致性。
3. 第三章-系统的总体设计
4. 第四章-系统的详细设计
5. 第五章-实验结果和系统演示
6. 第六章-总结

# 相关技术的研究

首先介绍了传统的人脸定位算法主动形状模型，然后再详细阐述本论文的基于层次分解模型的人脸解析算法的原理，接着介绍薄板样条变换方法，最后就视频传播中约束问题的研究展开讨论。



## **人脸特征点提取**

* + 1. **主动形状模型（ASM）**

主动形状模型ASM(Active Shape Models)[8]是一个基于统计学的模型，它的形状由点的集合表示，即采用点分布模型(Point Distribution Model, PDM)[15]，以特定物体的轮廓作为训练样本学习物体变化的模式，通过模型的形变最终于目标形状相吻合[16]。ASM不仅用于人脸定位，在最初，Cootes等人利用ASM模型对电阻[8,17]、人手[8,17]、心室[8]、蠕虫[8]进行特征点提取。ASM模型具有一般性，特定物体只要经过训练，就能使用ASM进行特征抽取。

ASM进行人脸定位主要包含两个过程：模型学习和特征点定位[18]。其中，模型学习将从人脸数据集中训练出人脸形状模型和人脸纹理模型。下面先介绍ASM模型中的点分布模型。

1. **PDM模型**

早在ASM模型建立之前，在统计学领域已经有大量学者对物体形状进行统计建模[19]。PDM模型是点的集合，记录着形状轮廓特征点的坐标。一般通过手工标记特征点的方法进行PDM建模。对于同一个训练集，每个PDM模型中点的总数都是相同的，而且标记的时候顺序必须一致，这样可以保证对于同一标号的特征点都代表同一个特征部位。如下图 2‑1 中，第48个点总表示嘴唇的左嘴角。

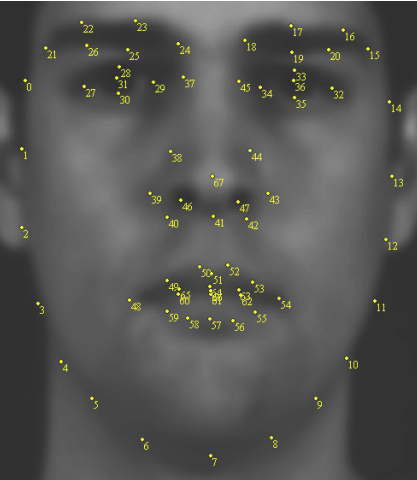


图 ‑1 PDM模型

1. **ASM模型建模**
2. **人脸对齐**

当训练集中的人脸都用PDM模型标记出来后，需要对它们进行归一化，使训练集中的样本都在同一个坐标轴上。选定一张人脸为基准（通常为样本集中的第一张），其他样本的人脸通过旋转、平移、缩放等几何变换向基准对齐（图 2‑2）。

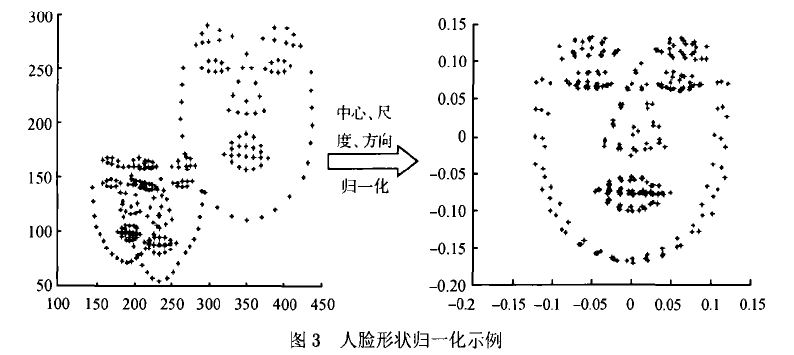
****

图 ‑2 人脸形状归一化

1. **人脸形状建模**

由于训练集的数据包含了大量人脸信息，因此需要通过主成分分析（Principal Component Analysis，PCA ）[[3]](#footnote-3)进行降维，以减少数据冗余。经过PCA降维的人脸模型*x*可以表示成：

(2‑1)

其中,是平均人脸：

(2‑2)

*P*为特征矩阵，*a*为几何变换参数，*b*为PCA形状参数，*n*是训练集样本数。公式(2‑1)表示，数据集中的任意一个人脸形状*X*都可以用平均人脸经过几何变换和PCA形状变换来表示（图 2‑3）。为了使生成的人脸不畸变，需对形状参数*b*加以约束，通常设置*bi*的范围为 ,其中为*bi*对应的特征值。



图 ‑3 黑色的平均人脸和灰色的第一主成分

1. **人脸纹理建模**

人脸纹理模型描述了脸部每个特征点附近的局部灰度信息。对于每张训练图片i上的所有特征点，收集其法线方向上左右两侧k个像素点的灰度值，得到的2k+1个点求差分，其2k个差分值组成的灰度模型即为该特征点的统计模型（图 2‑4）。

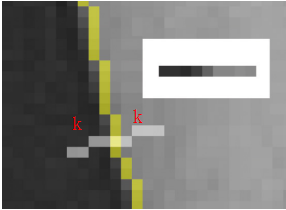


图 ‑4 纹理建模的过程

黄色线为人脸轮廓线，灰色线是特征点上的法向量，计算轮

廓法线上2k+1个点的灰度值，可以得到右上方的灰度模型。

实际上，纹理模型采用归一化梯度值，因此还要求其差分。

对所有训练图片进行标准化后得到：

(2‑3)

假设纹理模型服从高斯分布，那么人脸模型就可以用平均纹理和特征点的协方差矩阵*Sg*来表示。

(2‑4)

(2‑5)

1. **ASM的搜索策略**
2. **初始化ASM**

在使用ASM进行人脸定位之前需事先给定一个初始位置。初始位置很重要，它确定了目标人脸的大概位置。如果在搜索开始前没有给定的初始位置，或者给定的位置不正确，则算法很容易陷入局部最优，致使模型拟合到错误的位置。ASM的初始定位可以通过眼睛定位或者人脸检测得到。目前有很多鲁棒的人脸检测算法，如Rowley detector[20]，Viola Jones detector[21]，在此不一一叙述。

1. **搜索最佳位置**

在初始定位之后，算法将在模型当前所在位置附近进行搜索，寻找局部范围内最佳的位置。通常采用马氏距离 (Mahalanobis Distance) [[4]](#footnote-4)来度量采样位置与模板的匹配程度。对于目标人脸的每个轮廓点，收集沿法线方向上m个样本点的灰度特征gsj（m>k），并对其中2(m-k)+1个点计算匹配程度，定义匹配函数为：

(2‑6)

根据公式(2‑4)，公式(2‑5)和公式(2‑6)，当*f*(*gsj*)越小，匹配程度越高。因此，匹配的过程可以看作是寻找马氏距离最小的位置作为最佳位置（图 2‑5）。



图 ‑5 ASM搜索过程

(a)为灰度模型，(b)-(f)为法线上位移[-2,2]的采样点，(e)是最佳匹配位置

1. **估算最优目标和参数更新**

局部最优点引导模型前进，当使用新的位置来替代原来的位置后，需要重新计算形状参数和姿态参数。为了改变模型x使其与最优目标y最相近，我们希望通过形状变换和PCA变换改变x，因此估算最优目标的问题实际上是求,使得y与x距离的平方最小的问题。

(2‑7)

求出上式中的,，并根据新的形状参数和姿态参数计算出新的人脸模型。

至此，便完成了ASM搜索过程中的一个迭代。重复上述步骤，若两次迭代没有显著的变化，可认为算法收敛。

1. **ASM模型的缺点**

(a)在图像的匹配过程中，如果目标图像不够清晰或者模型初始位置不理想，会产生畸变的匹配结果[22]。

(b)如果待检测图像的光照条件与训练样本的光照条件相差较大，则灰度模型无法正确指导特征点搜索过程，导致算法不收敛、特征点定位不理想甚至失效[23]。

(c)在特征点搜索过程中，ASM只在法线方向进行了搜索，如果最佳位置不在法线方向上，就会导致模型不能很好地逼近目标形状[24]。灰度模型只能够提供有限的局部信息，未能有效地指导模型前进，算法需要迭代多次才能收敛。

(d)由于主动形状模型的方法是基于几何形状信息的方法，对噪声敏感[23]，因此很容易受噪声影响陷入局部最优。

* + 1. **层次分解模型**

针对上述经典ASM算法中的缺陷，Liang Lin等人提出了基于组合模型和判别搜索的人脸定位算法[25]。在该算法中，人脸模型由独立的五官、轮廓模型组成，搜索时用各自的ASM模型进行定位。在初始化时，分别对各个人脸部件进行检测，每个部件都得到一系列可能的位置，最终以分数最高的位置作为初始化位置。此外，利用Adaboost[21]对人脸的各个部件在九个方向上进行训练，得到的分类器将在搜索过程中告诉模型应该向哪个方向移动，指导模型前进。算法的大致过程如图 2‑6。

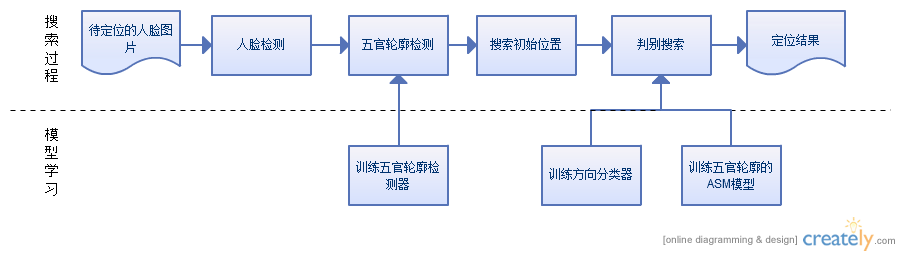


图 ‑6 组合模型定位流程

1. **模型学习**

模型学习训练出五官轮廓检测器、五官轮廓的ASM模型和方向分类器。在这个组合模型中，人脸模型被分成如图 2‑7所示11个部件。

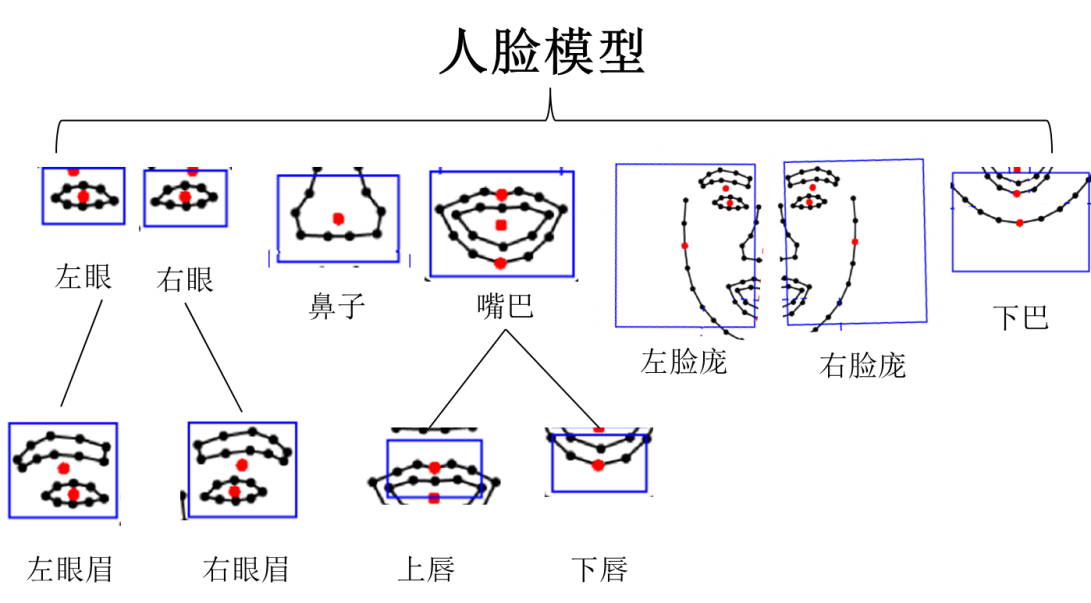


图 ‑7 人脸组合模型

五官轮廓的ASM模型训练过程就如2.1.1(2)节中所述，不同于传统的ASM模型，人脸组合模型把人脸分成各个子部件单独训练，在此不再赘述。以下介绍一下五官轮廓检测器和方向分类器。

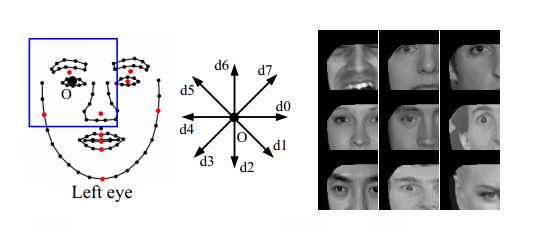
1. **五官轮廓检测器**

五官轮廓检测器通过Haar小波特征和Adaboost分类器进行训练。每个检测器都有一个置信度，定义如下：

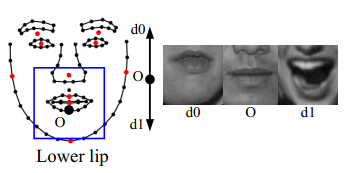
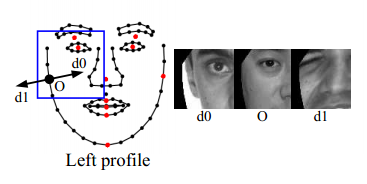
其中，r是第h层强分类器输出的得分，*Φh*是第h层中负样本与正样本的比值。

1. **方向分类器**

方向分类器的作用是在搜索过程中为模型的前进提供方向[25]，是搜索过程中的指南针。方向分类器是由基于Gabor特征的Adaboost算法训练得到。在训练阶段，眼睛、鼻子、嘴巴和下巴将分别在其9个方向上(包括不需要移动)进行训练(图 2‑8(a)); 对于左右脸庞，只需要考虑向左、向右和不需要移动这3个方向(图 2‑8(b))；上下唇和眉毛分别是嘴巴和眼睛的子部件，它们仅在其父部件定位后才进行微调，因此只需要训练向上、向下和不需要移动这3个方向(图 2‑8(c))。



（a）



1. (c)

图 ‑8 方向分类器

1. **判别搜索**

首先，对于一幅输入图片，我们使用训练阶段得到的五官轮廓检测器对各个人脸子部件进行检测，并把得分最高的位置作为各个子部件的初始位置(图 2‑9(c))。在检测时先对眼睛、鼻子、嘴巴和轮廓进行检测，当眼睛和嘴巴位置确定之后，再分别检测眉毛和上下唇。图 2‑9(a-b)中，圆点和方块最大的代表着检测得分最高。此后，输入图像将转换到平均人脸的坐标空间进行搜索(图 2‑9(d))。根据方向分类器，我们可以知道模型需要向哪个方向移动才能更好地逼近目标(图 2‑9(e))。经过数次判别搜索，模型最终收敛(图 2‑9(f))。

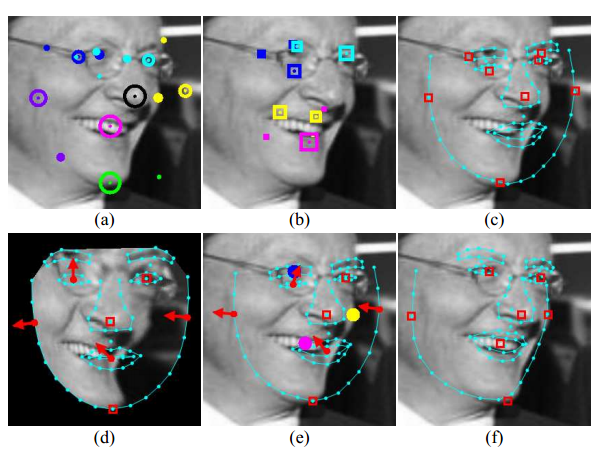
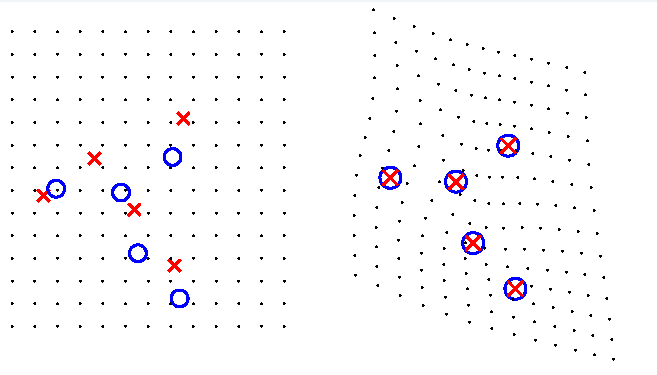


图 ‑9 判别式搜索过程

通常，该算法在2-3次迭代后可以收敛[26]。对比传统的ASM模型，带分类器的判别搜索无论是在性能还是准确性方面都比ASM要好。当模型遇到噪声时，方向分类器能够指导模型跳出局部最优，有效地避免在错误的位置上收敛。此外，传统的ASM模型对初始位置要求很高，而这个改进的算法利用了组合的人脸检测器估算人脸部件的大概位置，定位的结果不再依赖于初始化位置。由于算法使用的是组合的ASM模型，在定位的时候不受整个人脸配置的约束，即使在训练阶段没有对多角度的人脸进行学习，算法还是能够处理自然的肖像照。

## **薄板样条变形(TPS)**

薄板样条函数(Thin-plate Splines, TPS)是一种基于特征点的非刚性变形方法，可控性强，常用作图像配准。Bookstein[27]首先把薄板样条函数应用到医学图像配准。就如其名字所示，“薄板样条”代表一块金属薄片，变形时根据一组控制点对薄片进行扭曲。对于一幅输入的图片，给定一组控制点及其对应的映射集合，TPS通过最小化曲面能量从而确定整个平面的映射关系，如图 2‑6所示。



(a) (b)

图 ‑10 TPS变换

图(a)中，蓝色O是平面上的一组控制点，红色X与O中的每

个点一一对应，为使X与O重合，必须扭曲整个平面（b）。

TPS的能量函数如下：

通过最小化能量函数，得到点的映射关系如下：

(2‑8)

其中，P为源图像的空间坐标，Pi是源图像上的控制点，在薄板样条插值函数中,*r*是任意两个控制点的距离。为求薄板样条的变形系数*w1, …, wn, a1, ax, ay,*定义矩阵*W*:

(2‑9)

在矩阵*Y*中，*v*是目标图像的控制点。在矩阵*L*中，*K*是TPS的核，是一个n×n的矩阵，表示为：

*P*是一个3×n的矩阵：

把*K, P, v*代入(2‑9)可求得变形系数*W。*然后将*W*和原图像的控制点*Pi*代入(2‑8)可得到目标图像各个对应点的坐标。

## **传播约束**

# 肖像线描动画系统的总体设计

内容概括。

## **问题陈述**

## **需求分析**

* + 1. 用例析取
    2. 用例规约
    3. 补充规约

## **架构设计**

* + 1. 架构描述
    2. 架构设计
    3. 关键抽象

# 肖像线描动画系统的详细设计

内容概括

## **用例分析**

## **系统静态结构的设计**

## **系统实现**

# 系统实验与演示

## **运行环境**

## **系统的界面**

## **系统操作过程**

## **成果展示**

# 总结与展望

内容概括

## **6.1 工作总结**

## **6.2 未来的展望**

# 参考文献

1. Gary B Huang, Vidit Jain, ErikLearned-Miller, Unsupervised Joint Alignment of Complex Images, Computer Version, ICCV, 2007:1~8
2. 黄福珍，苏剑波，人脸检测，上海交通大学出版社，上海，2006,12
3. 张腾飞，闵锐，王保云，基于特征区域自动分割的人脸表情识别，计算机工 程，Vol.37,No.10，2011:146-151
4. Ekman P, Friesen W V, Facial Action Coding System(FACS): A technique for the measurement of facial movement，Consulting Psychologist Press, 1978.
5. M Kass, A Witkin, D Terzopoulos, Snakes: Active contour models, International Journal of Computer Vision, Vol.1,No.4, 1988:321~310
6. AL Yuille, PW Hallinan, DS Cohen, Feature Extraction from Faces Using Deformable Templates, International Journal of Computer Vision, Vol.8,No.2, 1992:99~111
7. TF Cootes, CJ Taylor, DH Cooper, J.Graham, Active shape models-their training and application[J]. Computer Vision and Image Understanding, Vol.61,No.1, 1995:38~59
8. TF Cootes, CJ Taylor, British Machine Vision Conference, 1992, 266~275
9. TF Cootes, GJ Edwards, CJ Taylor, Active appearance models, Pattern Analysis and Machine Intelligence, Vol.23,No.6, 2001:681~68
10. 邓梁, 基于ASM与AAM的人脸特征定位与匹配算法研究, 学士学位论文, 中南大学, 2009:1~61

[1] R. Fattal, D. Lischinski, M. Werman, Gradient domain high dynamic range compression, ACM Transactions on Graphics, vol.21,no.3, 2002: 249~256

[2] 苏步青，刘鼎元，计算几何，上海科学技术出版社，上海，1980，25~80

[3] W. M. Hsu, J. F. Hughes, H. Kaufman, Direct manipulation of free-form

deformation, In Proceedings of ACM SIGGRAPH’92, New York, ACM Press,

1992, 177~184

[4] 许栋，微分网格处理技术，博士学位论文，浙江大学，2006，1~50

注意：

(1)“参考文献”为论文中所有引文、引用观点以及对论文有重要影响和启发的文献；

(2)“参考文献”按在论文章出现的先后依次排序；

(3)“参考文献”内容一般排列在论文末尾（论文篇幅较大且引用文献较多的，可在每章末尾注出），序码与论文加注处对应；

(4)“参考文献”标注格式：

参考文献若为期刊：[序号]，著者(英文名字不要简写)，论文题目，期刊名，卷号（期号），出版年月，页码。实例见参考文献[1]

参考文献若为图书：[序号]，著者，书名，出版社，出版地，出版版次(若为二版以上)，出版年月，页码。实例见参考文献[2]

例如:

参考文献若为会议论文：[序号]，著者，论文题目，会议名称，地点，出版，卷号,年月,页码。实例见参考文献[3]

例如:

参考文献若为毕业论文：[序号]，著者，论文题目，博士（硕士）论文，校名，年份，起止页数。实例见参考文献[4]

注意：

（1）[序号]要突出.。

（2）参考文献出处最好为以下杂志期刊及级别更高的杂志期刊：

A类：中国科学、科学通报、计算机学报、软件学报、计算机研究与发展；

B类; 中大学报、小型微型计算机系统、中国图象图形学报、计算机辅助设计与图形学学报。

(3)硕士论文参考文献需要多于50篇,博士论文则需多于100篇.

注释：毕业论文中，参考所涉及到的网络文章、政府文件等非标准文献，在文中以注释形式出现。序码以圆括号放在加注处右上角，内容排在加注处所在页的页下，页下注序码每页单独排序[[5]](#footnote-5)。

**参考文献**

[1]Vidit Jain ErikLearned-Miller Gary B Huang. Unsupervised Joint Alignment of Complex Images，2007

[2]山世光, 人脸识别中若干关键问题的研究[D], 博士学位论文, 中国科学院研究生院（计算技术研究所）, 2004,

[3]黄福珍，苏剑波, 人脸检测[M], 上海交通大学出版社, 上海, 2006,

[4]张腾飞，闵锐，王保云, 基于特征区域自动分割的人脸表情识别[J], 计算机工程No.10, 2011：146-148.

[5]Paul Ekman, Wallace V. Friesen, Facial action coding system: A technique for the measurement of facial movement. Palo Alto, CA: Consulting Psychologists Press. Ellsworth, PC, & Smith, CA (1988). From appraisal to emotion: Differences among unpleasant feelings. Motivation and Emotion，1978，12：271-302

[6]Michael Kass, Andrew Witkin, Demetri Terzopoulos, Snakes: Active contour models, International Journal of Computer Vision，1988，1（4）：321-331

[7]A. L. Yuille, D. S. Cohen, P. W. Hallinan. Feature extraction from faces using deformable templates，1989

[8]Timothy F. Cootes, Christopher J. Taylor, David H. Cooperet al, Active shape models-their training and application, Computer vision and image understanding，1995，61（1）：38-59

[9]Timothy F. Cootes, Gareth J. Edwards, Christopher J. Taylor, Active appearance models, Pattern Analysis and Machine Intelligence, IEEE Transactions on，2001，23（6）：681-685

[10]邓梁, 基于ASM与AAM的人脸特征定位与匹配算法研究[D], 博士学位论文, 中南大学, 2009,

[11]Luo Ping, Wang Xiaogang, Tang Xiaoou. Hierarchical face parsing via deep learning，2012

[12]吴越, 基于人脸特征理解的肖像淡彩速写风格化系统[D], 硕士学位论文, 中山大学, 2012,

[13]Fred L. Bookstein, Principal warps: thin-plate splines and the decomposition of deformations, Pattern Analysis and Machine Intelligence, IEEE Transactions on，1989，11（6）：567-585

[14]Xu Zijian, Luo Jiebo. Accurate Dynamic Sketching of Faces from Video，2007

[15]Timothy F. Cootes, Christopher J. Taylor, David H. Cooperet al. Training models of shape from sets of examples: Citeseer，1992

[16]范玉华, 基于ASM的人脸面部关键特征点定位算法研究[D], 博士学位论文, 河南科技大学, 2008,

[17]Tim F. Cootes, Chris J. Taylor. Active shape models–smart snakes: Citeseer，1992

[18]王丽婷，丁晓青，方驰, 一种鲁棒的全自动人脸特征点定位方法[J], 自动化学报，Vol.35No.1, 2009：9-16.

[19]K. V. Mardia, I. L. Dryden. Statistical shape analysis. Chichester: John Wiley & Sons, ，1998

[20]H. A. Rowley, S. Baluja, T. Kanade, Neural network-based face detection, Pattern Analysis and Machine Intelligence, IEEE Transactions on，1998，20（1）：23-38

[21]P. Viola, M. Jones. Rapid object detection using a boosted cascade of simple features，2001

[22]张剑华，陈胜勇，刘盛等, ASM 中基于不变矩图形畸变主动检测与修正[J], 中国图象图形学报，Vol.14No.9, 2009：1886-1894.

[23]孙砚铭, 基于ASM自动人脸特征点定位及应用的研究[D], 博士学位论文, 北京交通大学, 2010,

[24]闵锋, 基于与或图表示的肖像画自动生成方法研究[D], 博士学位论文, 华中科技大学, 2009,

[25]Xu Zijian, Chen Hong, Zhu Song-Chunet al, A Hierarchical Compositional Model for Face Representation and Sketching, Pattern Analysis and Machine Intelligence, IEEE Transactions on，2008，30（6）：955-969

[26]Lin Liang, Rong Xiao, Fang Wenet al. Face Alignment Via Component-Based Discriminative Search: Springer Berlin / Heidelberg，2008

[27]Fred L. Bookstein, Principal warps: thin-plate splines and the decomposition of deformations, Pattern Analysis and Machine Intelligence, IEEE Transactions on，1989，11（6）：567-585

# 附 录

# 致 谢

由衷感谢我的导师XXX教授，本文是在他的悉心指导下完成的。

某某人

某年某月

1. Kinect是由微软开发应用于 Xbox 360 主机的周边设备，它有即时动态捕捉、影像辨识、麦克风输入、语音辨识、社群互动等功能。 [↑](#footnote-ref-1)
2. 阿凡达的英文为Avatar，源于电影《阿凡达》，在网络上有“头像、人物化身”的意思。 [↑](#footnote-ref-2)
3. 在多元统计分析中，主成分分析是一种分析、简化数据集的技术。其方法主要是通过对协方差矩阵进行特征分解，以得出数据的主成分（即特征向量）与它们的权值（即特征值）。 [↑](#footnote-ref-3)
4. 马氏距离是由印度统计学家马哈拉诺比斯(P. C. Mahalanobis)提出的，表示数据的协方差距离。它是一种有效的计算两个未知样本集的相似度的方法。 [↑](#footnote-ref-4)
5. [↑](#footnote-ref-5)