

分类号: TP311

单位代码: 10422

密 级:

学 号: 201013020



山东大学

硕士学位论文

(专业学位)

论文题目: 视频中人脸的京剧脸谱映射

**Mapping the Peking Opera Facial Makeup
onto a Human Face in Video Sequences**

作者姓名 原娜

学院名称 计算机科学与技术学院

专业学位名称 计算机技术

指导教师 屠长河 教授

合作导师

2012 年 4 月 15 日

原创性声明

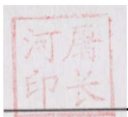
本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究所取得的成果。除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的科研成果。对本文的研究作出重要贡献的个人和集体，均已在文中以明确方式标明。本声明的法律责任由本人承担。

论文作者签名： 原娜 日期： 2012-4-20

关于学位论文使用授权的声明

本人完全了解山东大学有关保留、使用学位论文的规定，同意学校保留或向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅；本人授权山东大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或其他复制手段保存论文和汇编本学位论文。

(保密论文在解密后应遵守此规定)

论文作者签名： 原娜 导师签名：  日期： 2012-4-20



目 录

摘 要	I
ABSTRACT.....	III
第 1 章 绪论.....	1
1.1 选题背景.....	1
1.2 国内外研究综述.....	2
1.3 选题的意义.....	2
1.4 论文的主要工作.....	3
1.5 论文的组织结构.....	3
第 2 章 京剧脸谱的分析及矢量化.....	5
2.1 京剧脸谱的结构.....	5
2.2 京剧脸谱的勾画方法.....	6
2.3 京剧脸谱图像矢量化.....	7
2.4 纹样轮廓曲线细分.....	9
2.5 本章小结.....	11
第 3 章 人脸特征点的定位.....	12
3.1 特征点初始化.....	12
3.2 主动外观模型跟踪基本特征点.....	13
3.3 结合单应性跟踪特殊特征点.....	15
3.3.1 单应性矩阵	16
3.3.2 求解单应性矩阵	17
3.4 跟踪结果.....	21
3.5 本章小结.....	22
第 4 章 京剧脸谱的变形映射.....	23
4.1 自由变形技术.....	24
4.2 基于重心坐标的变形.....	26
4.2.1 均值坐标	27

4.2.2 格林坐标	29
4.3 移动最小二乘方法.....	32
4.3.1 基于点的移动最小二乘图像变形	33
4.3.2 利用控制曲线的图像变形	36
4.4 变形算法应用结果比较.....	40
4.5 本章小结.....	41
第5章 视频中京剧脸谱映射的算法实现	42
5.1 算法实现流程.....	42
5.2 算法实现结果.....	43
第6章 总结和展望.....	45
参考文献.....	46
致 谢.....	51

CONTENTS

Abstract in Chinese	I
Abstract in English	III
Chapter 1 Introduction	1
1.1 Background.....	1
1.2 Current Status	2
1.3 Significance	2
1.4 Main Work	3
1.5 Contents Organization	3
Chapter 2 Analysis and Vectorize Peking Opera facial makeup.....	5
2.1 Structure of Peking Opera facial makeup.....	5
2.2 Hand-painted Process of Peking Opera facial makeup	6
2.3 Vectorize Peking Opera facial makeup.....	7
2.4 Subdivision of Patterns' Boundary Curves.....	9
2.5 Chapter Conclusion	11
Chapter 3 Positioning Face Feature Points	12
3.1 Initialization of the Feature Points.....	12
3.2 Tracking Basic Feature Points by AAM.....	13
3.3 Tracking Special Feature Points	15
3.3.1 Homography Matrix	16
3.3.2 Calculate the Homography Matrix	17
3.4 Result of Tracking	21
3.5 Chapter Conclusion	22
Chapter 4 The Deformation.....	23
4.1 Free Form Deformation	24
4.2 Deformation based on Barycentric Coordinates.....	26
4.2.1 Mean Value Coordinates.....	27
4.2.2 Green Coordinates	29
4.3 Moving Least Square Method	32
4.3.1 MLS Deformation based on Control Points	33
4.3.2 Deformation using Control Curves	36

4.4 Compare the Different Method Results.....	40
4.5 Chapter Conclusion	41
Chapter 5 Implementation of Mapping.....	42
5.1 Process of Our Method.....	42
5.2 Result of Our Method.....	43
Chapter 6 Conclusions and Prospects.....	45
References.....	46
Acknowledgments.....	51

摘 要

京剧脸谱是一种中国戏剧中独特的人脸化妆,它源自生活有着强烈的中华民族特色,蕴含着浓厚的文化底蕴。随着演员的演出京剧脸谱一并展现在舞台上,是一门绝妙的活的艺术,散发着独特迷人魅力,其自身就有着很高的审美、文化和应用价值。随着对传统文化遗产继承、发扬和保护的重视,越来越多的人关注如何有效使用高新技术数字化脸谱这种图案艺术,研究它的建模应用。与此同时,人脸检测与跟踪除了开始的人脸识别应用外,还被广泛应用在人脸表情分析、视频监控、图片及视频检索等方面,在数字娱乐中的应用是新的应用之一,在动画及游戏设计中,丰富的人物表情更加能够表现人物场景的真实有趣性,表征人脸表情的特征点识别的研究也受到很大的关注。本文以京剧脸谱为研究对象,通过对脸谱这种传统艺术的分析,将京剧脸谱与视频中真实人脸结合,对人脸进行数字化京剧脸谱映射,其中涉及了计算机图形学技术、人脸特征点定位技术、图像变形技术等。通过研究,可以让我们更有效的认识京剧脸谱这一传统文化,并推动其在数字世界的应用发展。

目前已有的对京剧脸谱的研究是以标准人脸模型为主,没有关于对视频中真实人脸的研究。本文提出了一种使用图像变形技术映射京剧脸谱到视频序列中表情变化的人脸上的方法,是以真实人脸为对象,目标是将输入的京剧脸谱图和人脸视频转换成化妆有给定京剧脸谱的人脸视频,过程中,器官位置,如眼睛、嘴,脸谱上的和真实人脸上的相应位置密切匹配,与此同时,脸谱的艺术图案的形状能够较好的保持。主要就如何获得矢量化脸谱及其纹样的边界点、如何定位京剧脸谱和视频中每帧人脸的特征点及用何种方法进行映射等问题进行了研究。

前期分析了京剧脸谱的结构及其手工绘制的过程,对脸谱进行矢量化并用 SVG 格式文件表示,存入脸谱库。进行映射操作前首先对脸谱各纹样轮廓曲线细分获得边缘点,用这些点代表脸谱进行映射将大大减少计算时间。然后定义脸谱轮廓、眼睛、嘴边缘的点作为特征点,以及它们在视频中第一帧人脸上相对应的点,我们使用 AAM 算法跟踪人脸上的点,并通过单应性来解决额头点位置计算问题,获得以后每帧相应特征点位置。对每个视频帧图像,将脸谱图及人脸上特征点视为

变形前后的控制点，使用移动最小二乘方法映射脸谱到视频中人脸上，最终得到绘制有京剧脸谱的人脸视频输出。与其它方法相比，脸谱和真实人脸的器官边缘处匹配良好并且脸谱纹样的形状得到好的保持。

关键词：京剧脸谱；矢量化图像；特征点定位；图像变形；移动最小二乘

ABSTRACT

Peking Opera Facial Make-ups, which is originated in the life, is a special kind of makeup method with national characteristics. It is a living art appeared on the stage with the actors. It has the unique charm and itself has very high aesthetic, culture and application value. Along with the value of traditional cultural heritage protection, more and more attention is paid to the application research of this graphic art's digital modeling. At the same time, face detection and face tracking have not only confined to the application category of face recognition. They are also widely used in the picture and video retrieval, video monitoring, facial expression analysis and gender, race, age discrimination. And the use in digital entertainment is one of the new applications. Rich face expressions can express the realness and interesting of the characters and scenes in animation and game design. Feature point's identification which can show face expression also gets much attention. We take Peking Opera Facial Make-ups as research object in this paper. Through the analysis of this traditional art and the combination of the facial make-up and real face in video, we do mapping of the digital facial make-up's. It involves the computer graphics technology, face feature points positioning technology, image deformation technology and so on. Through the research, we can know this traditional culture more effectively, and promote the development of the facial make-up's application in digital world.

The existing researches of Peking Opera Facial Make-ups are based on standard face model and no study about the real face in videos. This paper we present a new method which uses image deformation technology for mapping Peking opera facial makeup onto a human face with time varying expressions in a video sequence. It takes real face as the object, and the goal is translating the input that is Peking opera facial makeup and face video to the video where makeup face has given facial makeup style. In this mapping, the organs of the facial makeup, such like eyes, mouth, are aligned with that of the human face tightly. At the same time, the shape of artistic patterns of the facial makeup is well preserved. The related programs which are discussed in this thesis include: how to get the vector facial makeup and the boundary points of its patterns, how to position the feature points of facial makeup and face in each video frame, and which kind of method we can use to do this mapping.

In the earlier stage we analyzed the structure of Peking opera facial makeup and the hand-painted process, and then used SVG format file to save them in a facial makeup library. Before mapping, we first get the boundary points by subdividing the contour curve segments of patterns. It will reduce the computation time with this points that can represent the facial makeup to do mapping. Second, we define a set of points on the boundary of organs of the facial makeup as the feature points, as well as their correspondences on the human face in the first frame of the video. We track the set of points on the human face using active appearance models (AAM) algorithm and compute the position of forehead points by using homographic matrix to get their corresponding points on each video frame. For each video frame, with the feature points as the control points, we use moving least squares method to map the facial makeup onto the human face frame by frame and get a video that the human face in the video is with the Peking opera facial makeup. Comparing with the other methods, the organs in the facial makeup and the human face are more tightly aligned and the shape of the strokes on the facial makeup is better preserved in our method.

Key words: Peking Opera Facial Make-ups; Vector Image; Feature Points Location; Image deformation; MLS

第1章 绪论

本章主要介绍视频中人脸的京剧脸谱映射的选题背景、国内外研究情况及研究的理论与实际意义，并对本文主要研究内容及组织结构进行了简单的介绍。

1.1 选题背景

京剧脸谱^[1,2,3]是在男演员脸上的彩色化妆，以人的面部为表现手段的图案艺术，是独有的、区别于其它国家任何戏剧的化妆。它可突破人物真实面目的局限，强烈鲜明的表现人物的个性特征。在长期的艺术实践中，戏曲艺术家们对生活现象的长期观察、体验并进行综合，不断分析判断剧中角色，做出评价，才逐步形成一套完整的艺术手法，造就了脸谱的演变和发展。

作为中国的国粹，京剧在二百多年的发展中，广为流传。每一种脸谱都是从人的五官部位、性格特征出发，虽然画法各有不同，但在形式、色彩和类型上都有一定的格式。有 15 个基本面部模式和超过 1000 个具体的变化，而且每一个都有其独特的个性特征。脸谱的图案纹样，有着流畅的曲线和鲜明的色彩，彰显角色个性，它用夸张、美化、变形、象征等手法来寓义褒贬，区分善恶，从而使人一目了然。京剧脸谱来源于舞台，和演员一起出现在舞台上，是一种活的艺术，有着很高的审美、文化和应用价值。尤其是近几年，我们在有些商品包装、瓷器或者穿着的衣服上都能看到风格迥异的脸谱形象。这些已经超出了原始的舞台应用范围，可见脸谱艺术在当今人们心目中所占据的地位。随着对传统文化遗产保护的重视，对脸谱这种图案艺术的数字化建模应用的研究受到了越来越多的关注。许多人出于对中国戏曲脸谱的好奇和喜爱，都在探索脸谱的奥秘。

脸谱依附于人脸，因此对脸谱应用的研究，必然离不开人脸检测^[4]、跟踪^[5]及人脸表情^[6]。近年来人脸检测和人脸跟踪已不仅局限于人脸识别的应用范畴，在图片、视频检索、视频监控、人脸表情分析、性别、种族、年龄判别等方面都有着广泛的应用，在数字娱乐^[7,8]中的应用是新的应用之一。在动画及游戏设计中，丰富的人物表情更加能够表现人物场景的真实有趣性。

1.2 国内外研究综述

2009年8月浙江大学CAD&CG实验室蔡飞龙、于金辉等通过分析手工绘制京剧脸谱的过程及京剧脸谱的结构、构成脸谱的纹样,包括眉纹、眼窝纹、嘴角纹等,使用Bezier曲线构造出脸谱的矢量化纹样库,并可根据创作需要对各个纹样按层进行组合得到不同的京剧脸谱^[9]。与手工绘制京剧脸谱相比,提出的方法可以用更灵活多变的方式,简单快捷的生成各种复杂的京剧脸谱。

为京剧脸谱赋予表情,能够使其更加生动形象并富有生命表现力,进而生成脸谱表情动画。目前已有商用的人脸表情编辑软件^[10,11,12,13],主要是基于位图变形技术来生成真实感表情。然而位图像素变形操作对于脸谱中重要面部特征且具有明显几何结构的纹样会带来不合理的形变。为了生成带有表情的脸谱需解决两个问题:怎样通过纹样变化表达不同表情,以及纹样在变化中如何保持自身结构。2010年6月,蔡飞龙、于金辉提出一个京剧脸谱表情分层驱动的矢量化建模和编辑方法^[14],在前一篇文章脸谱分析与合成的基础上对脸谱表情进行了有效的建模和编辑,他们首先根据脸部运动编码系统(FACS)标准分解脸谱表情为40个动作单元,然后合成脸谱时用户根据需要逐层选择所需要的纹样就可以得到想要的京剧脸谱图。使用自由形态变形(FFD)^[15,16]技术来作用于脸谱的表情动作单元,从而生成不同脸谱表情。它在数字博物馆、京剧教学、卡通动画、装饰设计及旅游产品开发等领域有广泛的应用前景。

同时在2010年,蔡飞龙、于金辉提出一个京剧脸谱的实时交互系统^[17],可以将脸谱映射到3D人脸模型上并进行渲染。并提出混合驱动生成京剧脸谱动画的方法^[18]。利用脸部运动编码系统,分解脸谱表情动作为基本表情、眼睛、口型、头部转动四个部分,根据四个部分各自的运动原理用不同的动画技术驱动脸部运动,这样用户通过简单调节动作参数就能生成栩栩如生的脸谱动画。

1.3 选题的意义

京剧脸谱的分析与数字化建模可以使人们深化对传统京剧艺术的认识,对于传统文化的保护和利用具有重要意义。脸谱表情的研究在京剧教学、卡通动画、视频聊天等领域有广泛的应用前景。在数字学习娱乐中,通过我们搭建的脸谱映射

平台，可以让京剧迷享受虚拟演唱体验，用户选择自己喜爱的角色脸谱，表演喜爱的曲目进行录像，最后可以欣赏到惟妙惟肖的场景和表演。这种方式可以让戏迷间交流更加真实有趣，也让年轻人或小朋友们有更直观的途径去了解这一生动的国粹。平常生活中也能够在视频聊天交流中，采用脸谱模式，使聊天更加有趣，或者从安全方面讲能够保护个人的隐私。

1.4 论文的主要工作

目前已有的对京剧脸谱的研究都是以标准人脸模型为主，没有关于到视频中真实人脸上绘制的研究。我们提出的将脸谱映射到视频中表情变化的人脸中的问题，是以真实人脸为对象。

本文以京剧脸谱图和人脸视频作为输入，研究方法来将脸谱绘制到视频序列中的人脸上。目标是当人脸运动和做各种表情时使脸谱粘附着人脸，如同人脸本身进行了脸谱的化妆。视频中，眼睛和嘴的动作需要展示出来，因此需要把普通脸谱图上眼睛和嘴处留空，并使其和视频中人脸的相应眼睛和嘴的位置处轮廓分别匹配对齐，而且，做映射的时候脸谱的纹样形状要保持。

为实现京剧脸谱的绘制目标，我们想到将脸谱图像映射到人脸上，映射需要控制点将脸谱和人脸区域对齐，而且因为眼睛、嘴的位置处要留空，其上的标记点也应该能够被检测并在视频中跟踪到，因此使用现在比较成熟的主动外观模型算法，它在视频人脸表情特征点跟踪中有着广泛的应用。对于映射，直接使用纹理映射方法必然是不可行的，了解到图像变形技术可以将一幅图像进行扭曲发生形变，在动画制作中都有很好的应用，因此我们分析了常用的变形方法，并经过试验最终使用移动最小二乘来实现这一映射。然而，对一整京剧脸谱栅格图进行变形，每个像素都需要进行计算，效率极低，因此通过对脸谱的研究，将脸谱进行矢量化，整个脸谱保存时表现为各种各样的纹样，最终纹样由其轮廓上点和填充色表示。对这些点进行计算，大大减少了计算时间。

1.5 论文的组织结构

本文以完成研究目标所需经历的三个阶段为路线，分别进行了详细的讨论分

析，其内容组织结构如下：

第一章概述了本文的研究背景、国内外研究现状、理论与实际意义及研究内容等。

第二章研究了京剧脸谱的结构以及手工绘制脸谱的过程，并对脸谱的矢量化成 SVG 格式，使用 Bezier 曲线来描述构成脸谱的纹样以及曲线细分进行了分析讨论。

第三章介绍了主动外观模型及使用它进行匹配跟踪视频中人脸特征点，并引入了单应性，通过单应矩阵来跟踪额头特征点的方法，得到较好的效果。

第四章分析了几个常用图像变形方法，比较它们的优缺点，并使用图像变形的方方法将脸谱映射到视频中跟踪到的人脸上，针对脸谱特殊性结合实验效果，找到适合的视频中人脸的京剧脸谱绘制方法。

第五章展示本文实现视频中人脸的京剧脸谱映射的算法流程，得出实验结果。

第六章对本文的研究内容进行总结与展望。

第2章 京剧脸谱的分析及矢量化

京剧脸谱化妆主要在净和丑两种角色上,用象征和夸张法来凸显人物特性,在各种戏剧化妆中拥有特殊地位。它是在京剧发展过程中,经过数代艺术家和演员的研究探索,不断归纳、演绎、创新,融合人物生理、心理、性格特征甚至社会地位、使用武器、职业等,从而形成了一套完整的化妆谱式,成就了现在我们所看到的脸谱艺术,有着变形、寓意、传神的艺术特色。

对京剧脸谱进行操作之前有一个很重要的工作就是对京剧脸谱进行分析,获得数字化脸谱模型。因此,我们首先分析京剧脸谱的结构造型、纹样的组合方式以及传统手工绘制京剧脸谱的勾画流程,对脸谱进行解析、整理和归纳,然后将脸谱图像转化成可缩放矢量图形,建立京剧脸谱数据库,用 Bezier 曲线拟合各纹样轮廓。

2.1 京剧脸谱的结构

京剧脸谱依据人物的生活对其角色进行勾绘,过程中进行了个性的夸张、放大,是对生活的一种概括取义。虽然看起来好像纷繁复杂,实际上它在类型、谱式及用色上都具有一定格式。京剧脸谱的构图多种多样,主要有整脸、三块瓦脸、十字门脸、碎脸等,并且用色也非常讲究,不同图案轮廓中绘制不同象征意义的颜色,如红色象征赤胆忠心;蓝色象征刚强勇猛;黑色象征忠耿正直;白色象征奸诈狠毒等。而且各种京剧人物都有突显自己特征性格的特定谱式及色彩,其中融入了善恶褒贬,一目了然,使观众易于辨别角色的心灵。脸谱中夸张变形的图形,用生动的曲线和特征化的色彩来展示角色,图案非常丰富,结合人脸器官从上至下大体上可以划分为额头图、眉型图、眼眶图、鼻窝图、嘴叉图、嘴下图,如图 2-1 所示。



图 2-1 （窦尔敦）京剧脸谱的基本结构

Figure 2-1 (Dou Erdun) The structure of Peking Opera facial makeup

脸谱图案有着丰富的变化、强烈的色彩，虽然有规律但也没有十分确定的论断，具有很强的寓意性。如两颊、眉毛和眼睛经常画上燕子、蝙蝠或者蝴蝶的翅膀的形状，额头纹经常用于着重强调角色特殊之处，如包拯额头纹是表示清正廉洁的白月牙，窦尔敦脸谱上画有他的兵器双钩。嘴、下巴和鼻子处也都有夸张的描绘，由此整体制造出个人特色^[1]。

2.2 京剧脸谱的勾画方法

脸谱具有很强的艺术性，勾画主要有揉、抹、勾、涂、破等几种基本技法，勾脸也具有高度的技巧性，如同绘画一般，只是画布变成了人脸，因此有了它自己的绘制方式。在人脸上手工绘制京剧脸谱的时候，首先要净脸垫底，定位和揉眼、鼻窝，然后勾白粉底，再在上面画出眉、眼窝、鼻窝、嘴角，最后在额头、脸膛等部位勾画出各种纹样图案，并在空白处填充主色。

将图 2-1 所示的《盗御马》中的窦尔敦角色脸谱作为示例，介绍京剧脸谱的勾画程序^[1]：

第一步：净脸垫底。顾名思义就是洗净脸以后用手指蘸稀淡的水白粉揉匀在整个脸上打底。

第二步：定位和揉眼、鼻窝。依然用手指蘸稀水蓝色抹出脸蛋及脑门和立柱纹处的大概轮廓，再蘸稀油墨，在留出眼窝里的白眼珠的情况下，定出眼窝及鼻窝位置。

第三步：勾白粉底。接下来程序都用笔勾。用笔蘸上较浓水白，勾画出额头、

眉窝、眉间、眼窝、鼻窝处的白纹，并把下巴勾白，把第二步中留出的眼窝内白眼珠点出。

第四步：勾黑油彩、墨线。留出白眼珠，把黑眼窝用黑油彩笔重新勾一遍，并把黑嘴窝画出来。然后用蘸墨的毛笔在眉毛中间近眼窝的地方勾双钩形纹，并勾两钩形纹在白下巴上成嘴下图。

第五步：勾红油彩。使用红油彩笔勾出位于印堂中心的胆形额头纹，留出匀称白边，并填出翻鼻孔，然后在白法令纹边缘勾两细红线，再把红嘴唇和眉子的红梢勾上。

第六步：加填水蓝膛、黄眉梢。蓝脑门和脸蛋处用比较深的水蓝色填加，把水白色眉心用黄油彩笔填上黄色。

这样才完成窦尔敦的京剧脸谱图。如同上面描述，勾画脸谱的时候由浅到深，垫底后依次绘制纹样，循序渐进的完成，最终形成一个各纹样相协调的整体。

从以上京剧脸谱的结构及其绘制过程，我们可以看出脸谱由额头、眉、眼、鼻、嘴、脸颊处的各种纹样及粉底组成，这些脸谱局部区域对应的造型各异的纹样，映射人物角色的性格特征，各纹样之间相互映衬，脸谱因纹样的适宜组合而能强烈表现人物，纹样图案也因人物而富含韵味，因此脸谱是各种颜色区域纹样的有机整体，所以我们通过分析后对纹样进行一定操作即可得到京剧脸谱的变化。对于一幅脸谱图像，直接在位图图像上对每个像素进行操作非常复杂，计算量很大，纹样区域中点的颜色一致性没有利用。而矢量图则提供了一个很好的解决方案，每个纹样区域由其轮廓边界上的控制点及其填充色即可表示，这样纹样可以真正的成为一个独立的整体。应用时只需要对其边缘点进行操作后填充颜色即可。

2.3 京剧脸谱图像矢量化

首先来了解下位图图像矢量化。一般情况下，图像以点阵的形式存储在计算机中，图像处理的时候也通常是按像素进行，然而有些情况下这种方式并不合适。按照点阵方式存储图像要对每个像素进行存储，需要很大的存储空间，对计算机要求较高，并且，对于绘图图形输出管理使用点阵图并不适宜^[19,20]，点阵图和分辨率有关，图像经过一定的放大后就会模糊失真。矢量型图像^[21]是用各种几何线、

形等来描述图像数学信息，通过基本图元记录图像。由于这些图元存储时只需要保存形状线条和填充等属性，因此矢量图具有占用存储空间小，而且放大后依然清晰即与分辨率无关的优点。位图图像矢量化就是根据具体情况将已有位图通过矢量化方法转变成矢量图像，从而解决应用中的速度及空间问题，灵活发挥矢量图像的优势。

目前存在很多的矢量图的文件格式，2000 年由 W3C 最终制定发布的 SVG (Scalable Vector Graphics, 可缩放矢量图像)^[22]是使用可扩展标记语言 XML 来描述二维矢量图的一种开放标准的图像格式。SVG 包含着矢量图形、点阵图像和文本三种类图像对象，不同图像对象可以进行变换组合，还可以修改样式。SVG 将矢量图形用文本表达，这样有较强的可读性，易于理解图像结构，而且编程简单、执行迅速，也有好的扩充性，是一种很好的图像文件格式，很多国际组织及公司都支持。

SVG 图像可以进行检索，便于修改编辑，可以查找和选择图像中的文本，这些都是位图图像无法做到的。通过观察容易看出，除了少数特殊形状和折点，京剧脸谱颜色区域的轮廓大部分都是光滑弯曲的，用贝塞尔曲线就可以进行良好的构造，因此也很适合做矢量图形。这样用少量的曲线控制点和内部填充色定义纹样，既节省脸谱的存储空间，又利于在绘制过程中降低计算量。

由于脸谱的特殊性，根据脸谱的结构和手工绘制过程，我们了解到脸谱由一块块的颜色区域组成，因此使用基于颜色的方法来进行区域划分。其过程如下图 2-2 所示：

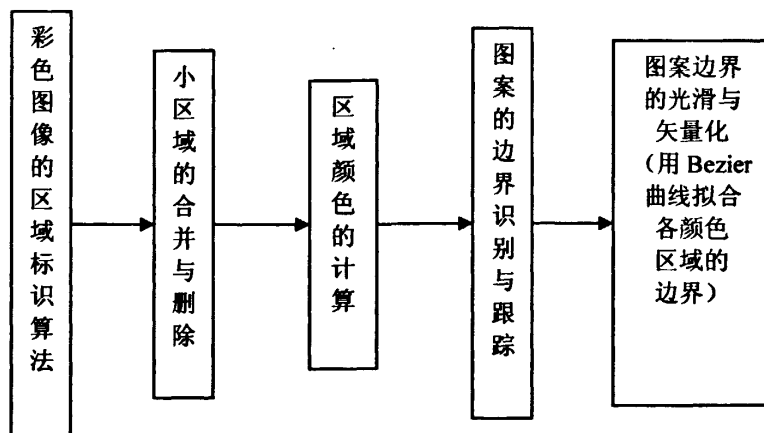


图 2-2 基于颜色进行区域划分方法进行矢量化的过程
Figure 2-2 Vectorization process based on color regional division

生成的矢量图形使用 SVG 格式存储，将每段区域的轮廓用 Bezier 曲线拟合，如图 2-3 所示的脸谱纹样，在 SVG 格式中用各项对应的 path 表示，如表 2-1 所示，此表显示了一个描述额头纹样的 SVG 文本文件，其中包含了纹样名称 id、填充色、轮廓曲线路径等属性。这样，我们整理从网络上获得的京剧脸谱位图图像，将它们进行矢量化后得到对应脸谱的 SVG 格式文件，并构建京剧脸谱库进行保存。



图 2-3 脸谱纹样（轮廓处蓝色线条表示拟合的 Bezier 曲线，端点由蓝色点表示）
Figure 2-3 Pattern of facial makeup and its boundary curve Contour line in blue show the fitting Bezier curves, and the end points are showed

表 2-1 SVG 格式的矢量图文件
Table 2-1 A vectograph file in SVG format

<pre><?xml version="1.0" encoding="UTF-8" ?> <!DOCTYPE svg PUBLIC "-//W3C//DTD SVG 1.1//EN" "http://www.w3.org/Graphics/SVG/1.1/DTD/svg11.dtd"> <svg width="574pt" height="563pt" viewBox="0 0 574 563" version="1.1" xmlns="http://www.w3.org/2000/svg"> </g> <g id="forehead"> <path id="forehead" fill="#f8e265" opacity="1.00" d=" M 279.09 35.05 L 279.96 35.00 C 281.84 67.01 285.44 98.89 289.32 130.71 C 292.75 162.25 298.48 193.48 302.09 224.99 C 304.45 243.01 306.40 261.11 309.07 279.06 L 308.18 278.83 C 300.27 252.67 291.01 226.87 280.26 201.72 C 279.52 199.87 278.29 202.17 278.07 203.05 C 272.07 217.59 266.08 232.14 261.06 247.05 C 256.82 257.13 254.45 268.43 250.03 278.19 C 259.01 197.06 273.14 116.50 279.09 35.05 Z" /> </g> </svg></pre>

2.4 纹样轮廓曲线细分

我们要按纹样映射京剧脸谱，对纹样轮廓点进行变形计算后填充颜色，以降低计算量，因此从 SVG 文件读取出的纹样轮廓曲线首先要做曲线细分操作。

计算机图形学中，贝塞尔曲线是非常重要的参数曲线，在研究中应用广泛。通常画曲线的方法是用一系列经细分后的短小线段去逼近。数值分析中的 de

Casteljau 算法是计算伯恩斯坦形式的多项式或 bezier 曲线的递归方法,它以稳定数值的方法求出 Bezier 曲线^[23,24]。此算法来源于 Paul de Casteljau 发现的性质:任意角度的曲线都能够被分割成两条角度相同的曲线。本文经由图 2-4,进行 de Casteljau 细分 Bezier 曲线算法。

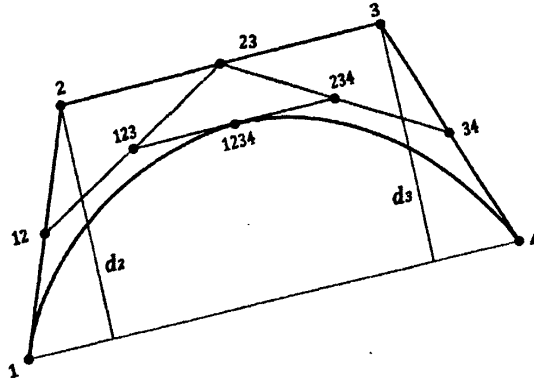


图 2-4 Bezier 曲线细分
Figure 2-4 Subdivision Bezier curves

首先有 1,2,3,4 四个控制点构成了一条贝塞尔曲线,计算出这三条线段的中点 12,23,34,并进一步得到高一阶的中点 123,234,最后一阶的中点 1234 就会落在曲线上,并由此点将曲线分成了两段贝塞尔曲线。这两段新的曲线组合起来和原来曲线完全一致,但比原来的曲线平坦许多,其中左半边曲线由 1,12,123,1234 四个点控制,右半边曲线由 1234,234,34,4 四个点控制。这样重复这个过程若干次,就可以用许多小线段来代替原贝塞尔曲线了。曲线细分的递归程序代码如表 2-2 也非常经典:

表 2-2 Bezier 曲线细分的经典递归程序代码
Table 2-2 The classic recursive code of subdivision Bezier curves

```
void subdivide_bezier (point 1, point 2, point 3, point 4)
{
    计算中点 12,23,34;
    计算高一阶中点 123,234;
    计算最后一阶中点 1234;
    If(terminate_condition)
    {
        停止细分;
        得到点 1234;
    }
    else
    {
        subdivide_bezier(1,12,123,1234);
        subdivide_bezier(1234,234,34,4);
    }
}
```

通过连接得到的点就可以绘制我们的 Bezier 曲线，现在的问题就是 `terminate_condition` 这个条件要如何设定，即何时停止细分。停止细分程度有很多，我们希望能够自适应曲率细分，曲率大的地方多细分，曲率小的地方少细分，通过少的点近似整条曲线，以控制输出点的数量引起的计算量。多次的实验之后，发现计算两段长度的和 d_2+d_3 (d_2, d_3 如图 2-4 中标示) 会更好，这种方法^[25]不需要对第一次细分进行特殊处理，并能得到一个很好的误差的估算。所以这里计算点 2、3 到线段 14 的距离 d_2, d_3 ，点 1、4 在 X、Y 轴上距离 dx, dy ，通过判断式 2-1 条件是否成立。如果满足这个条件不等式，细分停止，否则继续直到停止。

$$(d_2+d_3)*(d_2+d_3)<m_distance_tolerance*(dx*dx+dy*dy) \quad (2-1)$$

这里 `m_distance_tolerance` 是一个常量，是控制细分停止的最小距离的平方，对于一般的屏幕分辨率，可以取： $0.5*0.5=0.25$ 。细分后的纹样如下图 2-5 所示，轮廓边缘细分后得到的点蓝色点标记。



图 2-5 细分后的纹样
Figure 2-5 Boundary points of a pattern after subdivision

2.5 本章小结

本章描述了实现脸谱绘制的第一阶段，即脸谱的预处理阶段：脸谱的分析及矢量化。京剧脸谱由额头、眉毛、眼窝、鼻子、嘴、脸颊等处各种各样的纹样及底色组成，把纹样单独取出，看作一个独立区域，那么每个纹样都有着自己的外边界和一个或更多的内边界。通过矢量化，每个纹样的内外边界都由许多分段平滑 Bezier 曲线段来构成，本文中我们称这些曲线为边界曲线。所有的这些曲线都以 Bezier 曲线控制点的形式存储在矢量化后的 SVG 格式文件中。在变形之前将这些曲线进行自适应细分，以少量合理的点代替整个纹样，称这些点为边界点，作为脸谱映射时进行变形的输入。

第3章 人脸特征点的定位

人脸在表达情感、交流中是最重要的角色，分析研究人脸图像是图像处理、计算机视觉等领域中活跃的分支。快速准确的从图像或视频中获得人脸特征点信息是实现好的人脸检测、跟踪、识别、表情分析等的重要前提，由于人脸是非刚体，表情的变化、光线的变化、姿态、遮挡物的干扰，都会影响对人脸特征点判断的正确性，所以人脸面部特征的定位^[26]是一个具有挑战性的课题。本章研究的内容是如何定位视频中人脸特征点，以获得用于变形的控制点，我们使用主动外观模型^[27]方法定位人脸普通特征点，由于头发等遮挡的存在，主动外观模型没有跟踪到额头处的点，对于这一问题，我们使用主动外观模型结合单应性思想方法进行处理。

3.1 特征点初始化

图像进行映射之前，首先要定义特征点以控制变形。这些点要能够表征区域形状，个数也要有一定的限制，太多的话，应用后面变形算法会比较耗时，太少，不能够反映轮廓特征，变形后的区域轮廓会出现凹凸不平滑现象。同时也需要考虑人脸器官产生表情的灵活度，现实生活中人嘴部的动作丰富，因此对应的嘴部地区特征点要分布更密集一些，眼睛次之，整个人脸的外轮廓在靠近嘴的下巴处需要较多点，其余部位可以稀疏些。

典型的 AAM 匹配到的特征点中，对眼睛、嘴轮廓及脸部外轮廓都有合理的描述，点数也比较合适，因此我们取这些位置的点，称为基本特征点。由于头发遮挡等问题的存在，AAM 没有检测额头上的特征点，而由于脸谱在整个人脸上的完整性，这些点是不容忽视的，所以我们手工标记视频中第一帧的额头点位置，称这些点为特殊特征点，所有的基本特征点和特殊特征点构成了一个完整整体，如图 3-1 (b) 所示。对于脸谱图像，则是由手工标记了对应的特征点，如图 3-1 (a) 所示。

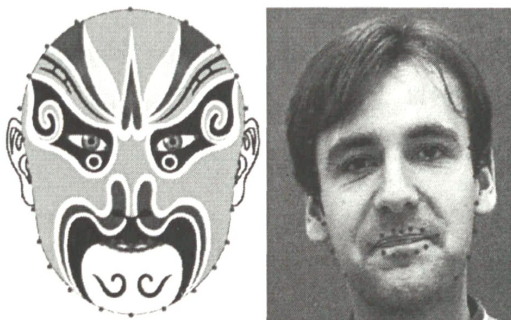


图 3-1 44 个特征点的对应位置
 (a) 脸谱上标记的特征点 (b) 视频中第一帧人脸上的特征点
 Figure 3-1 The corresponding positions of 44 feature points
 (a) Feature points in facial makeup (b) Feature points in the face of first frame

3.2 主动外观模型跟踪基本特征点

主动外观模型 (Active Appearance Models, AAMs), 是 T. F. Cootes 等^[27,28]在主动形状模型 (Active Shape Models, ASMs)^[29,30,31,32]基础上提出的另一个基于统计分析信息建立先验模型的方法, 用来对物体特征的定位。相比较于 ASM, AAM 在模型中同时包含了物体的形状 (shape) 和纹理 (texture) 信息, 对全局纹理也进行了匹配, 减少了局部收敛, 是一种建立二维模型的良好方法, 灵活并且有效, 在许多领域都有广泛的应用。人脸识别和面部特征定位中使用这一方法得到了较好的效果, 近年来, 针对不同的应用对主动外观模型方法研究者们也提出了许多的改进^[33—40]。

AAM 进行人脸特征点定位过程主要包括两部分: 外观模型的建立和模型的匹配。建立模型离不开训练集, 其中包括若干人脸图像及其对应的顶点位置文件。它定义每个人脸都包含由若干顶点通过三角化得到的基本形状, 并对人脸形状 s 用这 m 个顶点的横纵坐标描述为 $s_i = (x_{i1}, y_{i1}, \dots, x_{im}, y_{im})$, 其中 x_{ij} 、 y_{ij} 分别为训练集中第 i 幅人脸图中第 j 个特征点的坐标值。通过平移、旋转、缩放操作使训练集中人脸形状向量对齐^[42]到同一个坐标系之下, 对得到的人脸各特征点分布进行统计分析, 利用主成分分析 (PCA) 方法得到平均形状向量 S_0 和相互正交的特征形状 S_i , 这样任意形状 S 可以看成基本形状 S_0 和 n 个特征形状的线性组合, 因此对形状建模为

$$S = S_0 + \sum_{i=1}^n p_i S_i \quad (3-1)$$

基本形状 S_0 确定后, 就能确定它所包围的像素位置和数量。将所有的训练集样本人脸纹理经过分段仿射变换到 S_0 中, 使各纹理的像素数统一一致, 这样经过纹理对齐后依然使用 PCA 降维, 得到平均纹理 A_0 和相互正交的特征纹理 A_i , 同形状相似, 任意纹理 A 都可以看成由基本纹理和特征纹理的线性组合, 则相应线性纹理建模为

$$A = A_0 + \sum_{i=1}^m q_i A_i \quad (3-2)$$

将形状模型和纹理模型联合起来, 再次使用 PCA 进一步降维, 去除掉形状参数和纹理参数之间的一定相关性, 从而得到最终的组合外观模型:

$$S = S_0 + \sum_{i=1}^l c_i M S_i \quad A = A_0 + \sum_{i=1}^l c_i M A_i \quad (3-3)$$

这样, 只要操作外观模型参数 c 变化就能够同时变化形状及纹理进而控制整个人脸模型变化, 表示任意一幅人脸图像。

用 AAM 外观模型来标定任意人脸图像特征点的过程是 AAM 模型实例和特定对象的匹配过程。其思想是: 对于一幅给出的含有人脸的输入图像, 使用形状和纹理的两个基本式, 通过调整模型参数去寻找适合的组合系数, 使得合成的形状纹理图像 $A(x)$ 与已有的输入图像的人脸 $I(x)$ 区域纹理之间的差距 (如下式) 最小,

$$E = \sum_{x \in S_0} [A_0(x) + \sum_{i=1}^m q_i A_i(x) - I(W(x;p))]^2 \quad (3-4)$$

其中 x 表示 S_0 形状内的一个像素, $I(W(x;p))$ 为待匹配人脸图像的对应 S_0 形状内点集的纹理像素信息, $W(x;p)$ 是分段仿射变换。AAM 匹配迭代步骤如下:

1. 对于输入的人脸图像 I , 使用 Adaboost Haar-like Face Detector 方法找到人脸, 确定初始形状的位置。
2. 由初始位置计算该形状对应的联合模型参数 c_0 作为初始值, 根据式 3-3 计算合成外观模型的形状 S 及纹理 A ; 并对 S 内的人脸图像区域纹理采样, 对齐为 $I(W(x;p))$;
3. 计算插值图像 $E_0 = A - I(W(x;p))$, 并得误差值 $e_0 = \|E_0\|$;
4. 计算出用于预测参数 c 的调整差值 Δc ;
5. 置 $k=1.5$ 并更新外观模型参数 $c_1 = c_0 - k\Delta c$;
6. 重复步骤 2,3,4, 用新的 c_1 计算得到新的误差 e_1 ;

7. 如果 $e_1 < e_0$ ，接收 c_1 作为新的预测，当 e_1 和 e_0 的相差很小的时候，匹配结束。若非收敛，则试着用 $k=1, 0.5, 0.25, 0.125$ 重复步骤 4,5。

通过以上总结 AAM 过程为：首先为人脸的形状和纹理分别进行建模得到各自的模型，然后以一定方式结合形状和纹理模型，从而建立起可以描述整个人脸变化的统计外观模型。AAM 识别人脸图像是使用了基于优化算法进行搜索的基本思想，将合成的模型人脸去匹配给出的目标人脸，用两者的差异不断修正合成人脸，以使模型与目标人脸相适应，最终形成反映目标人脸形状和纹理的合成模型图，从而得到人脸特征点的位置。

文献^[40]对主动外观模型方法能够在视频跟踪领域中的应用进行了研究，就是以前一帧中主动外观模型的定位特征点结果做当前帧的点位置初始值，继续进行迭代匹配的过程，在视频中有较多的应用。本文使用此方法跟踪视频中人面部的基本特征点，其结果如图 3-2 所示。

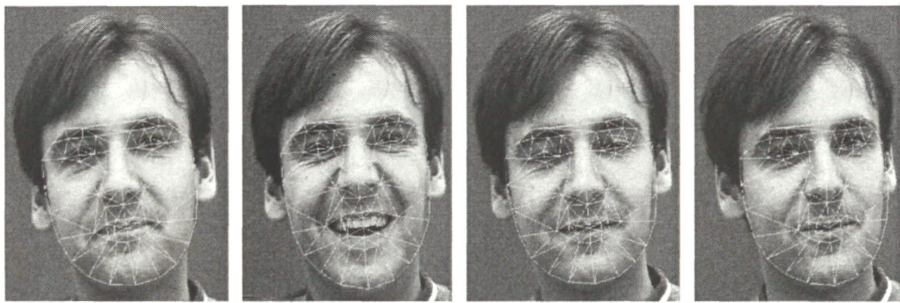


图 3-2 主动外观模型跟踪到人脸形状
Figure 3-2 The face shape tracked by AAM

3.3 结合单应性跟踪特殊特征点

因为每个人发型的不同，额头处头发遮挡情况的不同，而且视频中头发还可能被摆弄或吹动的原因，使得不同人脸图像中额头轮廓处没有一致的特征，难于建立此处外观模型，所以主动外观模型无法跟踪我们标记的额头特征点，为此我们考虑人脸整体特征及相应几何方法，研究提出在主动外观模型跟踪到的基本特征点基础上结合单应性思想计算跟踪额头处特殊特征点的方法。

单应性 (homography) 是一个从空间 P^n 到 P^n 平面间的可逆投影变换^[41,42,43]，单应性矩阵包含了摄像机内参数矩阵及旋转、平移外参数矩阵。计算机视觉领域里，经由单应性可以将同一空间平面投影形成的任两幅图像关联起来，它有着许

多的实际应用，例如图像的校正、图像的对齐以及两幅图像间的相机运动计算等。

3.3.1 单应性矩阵

根据相机模型针孔成像原理^[44,45]，三维空间中物体点通过透视变换投影到二维平面从而形成图像。用齐次坐标系表示，空间平面模板上一点 $P=[X,Y,Z]^T$ 与和它对应的二维图像上的点 $p=[u,v]^T$ 之间存在关系：

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K [R \quad T] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (3-5)$$

这里 s 是放缩比例系数， K 是相机内参数矩阵， $[R \quad T]$ 是相机外参数矩阵，其中 R 为旋转矩阵， T 表示平移距离矢量。

假设处在世界坐标系中的空间平面模板位于 $Z=0$ 的位置，旋转矩阵的 R 的列用 r_i 表示，那么有

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K [r_1 \quad r_2 \quad r_3 \quad T] \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = K [r_1 \quad r_2 \quad T] \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad (3-6)$$

将相机内参数和外参数合并一起，组合成映射矩阵，那么空间平面模板上的点和二维图像平面上的对应点之间就可以用映射矩阵 h 来表示成：

$$\overline{sm} = h \overline{M} \quad (3-7)$$

这里 $\overline{M}=[X \quad Y \quad 1]^T$ 表示空间平面模板上的点的齐次坐标， $\overline{m}$ 表示二维图像上点的奇次坐标，映射矩阵 $h=K[r_1 \quad r_2 \quad T]$ 则是一个 3×3 的矩阵，其中 r_1 、 r_2 是在世界坐标系下图像平面的两个坐标轴的方向向量。

考虑我们的视频，忽略空间中人面部器官部分较小的深度差距，认为脸部侧边轮廓点、鼻子周围点和额头的点在同一个三维平面上，这些点组成的集合设为 P ，如图 3-3 所示，其中右图表示空间中的点集 P ，左图表示映射到像平面后的点。由于人脸做表情的时候，这些位置点受到的影响较小，所以当人脸动作时，点集 P 可认为是在空间中整体做刚性变化的。

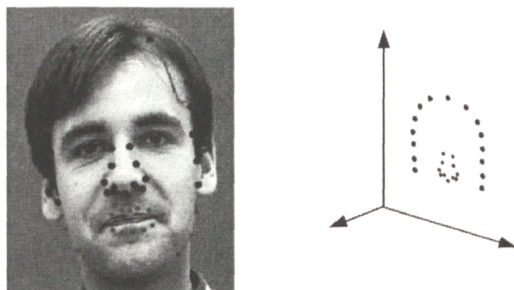


图 3-3 假设在同一三维平面中的点集（红色表示）

Figure 3-3 The point set which we suppose in the same 3D plane

这样，视频的每一帧图像中人脸相应位置点 p 就是空间中运动的三维平面的点集 P 在同一摄像机的投影。由于点集 P 是整体刚性变化，经坐标原点平移到点集中心后， P 中每个点在其三维空间平面内坐标值不变，因此对视频中图像帧，点集 p_i 的每个点 p_{ik} 对应投影关系

$$s_i \bar{p}_{ik} = h_i \bar{P}_k \quad (3-8)$$

其中， $\bar{P}_k$ 表示点集 P 中第 k 个点的齐次坐标， $\bar{p}_{ik}$ 表示第 i 帧图像中点集 p 的第 k 个点的齐次坐标。

前面描述了帧图像与空间平面点集之间的联系，借助这一空间平面，我们搭建同一视频中帧与帧之间的关联。根据式 3-8，对视频中任两帧图像 i, j ，有 $s_i \bar{p}_{ik} = h_i \bar{P}_k$ 、 $s_j \bar{p}_{jk} = h_j \bar{P}_k$ ，就相当于摄像机在空间中不同位置为同一三维平面点集拍摄成像，由此可得视频帧图像平面之间存在关系：

$$s_{ij} \bar{p}_{jk} = h_j h_i^{-1} \bar{p}_{ik} = H \bar{p}_{ik} \quad (3-9)$$

将 $h_j h_i^{-1}$ 合并在一起组成新的矩阵 H ，这里我们称 H 为帧图像平面之间单应性矩阵（homography matrix）。

帧图像平面间单应性矩阵架起了视频中帧图像之间的桥梁，这样当我们已知某一帧图像上点集 p 坐标位置的时候，就可以通过式 3-9 得到各点对应的另一帧图像坐标，因此求解此单应性矩阵 H 是非常重要的。

3.3.2 求解单应性矩阵

本文中已经由 AAM 能跟踪到人脸面部的一些基本特征点，因此上小节的点集

p 中部分点对坐标可以获得。所以我们通过获取的两前后帧 f_i 、 f_j 平面上的一些对应已知点对来解出这个单应性矩阵。因为已知点对之间必然存在误差，并不完全满足式 3-9，所以可将这个矩阵的求解化成一个非线性最小二乘问题，使用非线性优化方法中的高斯-牛顿法来解决。因此对于单应性矩阵 H ，它的极大似然估计值应满足如下最小平方和：

$$\min \sum_k \|p_{jk} - \hat{p}_{jk}\|^2 \quad (3-10)$$

这里 $\hat{p}_{jk}$ 是由 3-9 式计算得的期望图像坐标值：

$$\hat{p}_{jk} = \frac{1}{\hat{h}_3 \bar{p}_{jk}} \begin{bmatrix} \hat{h}_1 \bar{p}_{jk} \\ \hat{h}_2 \bar{p}_{jk} \end{bmatrix} \quad (3-11)$$

其中 $\hat{h}_i$ 是指单应性矩阵 H 估计值的第 i 行向量。使用非线性优化方法对此最小二乘进行求解时首先需要一个合适的初始值以进行迭代，此初始值的获得可以使用下面方法。

对单应性矩阵进行分析，设 h_i 是单应性矩阵 H 的第 i 行向量， p_{jk} 点坐标为 $[u, v]^T$ ，

通过式 3-9 或者式 3-11 可以得到式 $h_3 \bar{p}_{jk} \begin{bmatrix} u \\ v \end{bmatrix} - \begin{bmatrix} h_1 \bar{p}_{jk} \\ h_2 \bar{p}_{jk} \end{bmatrix} = 0$ 或者 $\begin{cases} \bar{p}_{jk}^T h_1^T - u \bar{p}_{jk}^T h_3^T = 0 \\ \bar{p}_{jk}^T h_2^T - v \bar{p}_{jk}^T h_3^T = 0 \end{cases}$ ，以

矩阵相乘积形式来描述即是有下式成立：

$$\begin{bmatrix} \bar{p}_{jk}^T & 0^T & -u \bar{p}_{jk}^T \\ 0^T & \bar{p}_{jk}^T & -v \bar{p}_{jk}^T \end{bmatrix} \begin{bmatrix} h_1^T \\ h_2^T \\ h_3^T \end{bmatrix} = 0 \quad (3-12)$$

将其中的大小为 9×1 的列向量用变量 x 表示： $x = [h_1^T \ h_2^T \ h_3^T]^T$ ，则有

$$\begin{bmatrix} \bar{p}_{jk}^T & 0^T & -u \bar{p}_{jk}^T \\ 0^T & \bar{p}_{jk}^T & -v \bar{p}_{jk}^T \end{bmatrix} x = 0 \quad (3-13)$$

这是一个形似 $Lx=0$ 的矩阵方程，如果已知的对应点对有 n 个，那么就能够获得 n 个式 3-13 构成的方程组，将这些方程组合并在一起形成一个大的矩阵方程 $Lx=0$ ，其中 L 是 n 个 l 的堆积，因此是一个 $2n \times 9$ 大小的矩阵， $\|x\|=1$ 。很明显的可以看出，当有四对已知点的时候， $n=4$ ，正好可以解出这个线性齐次方程组的解 x ；少于四对点时， $n < 4$ ，无法解出；多于四对点的时候， $n > 4$ ，此时问题就转化成了

对超定方程组寻找一最优解的问题。因此，至少要有四对已知点，我们才能够解出单应矩阵，并且由于点的误差的存在，通常都用多于四对点来寻求优化的结果。

通过式 3-13 可以看出 L 矩阵组成元素的数量级差别较大，因为其中有的是图像像素位置坐标，有的项是两坐标乘积，这样矩阵中元素量级大约在 $10^4 \sim 1$ 范围，从而会导致矩阵 L 的条件数很大。而条件数能够表示矩阵运算对误差的敏感性，是稳定性分析里一个最关键因素，也就是说条件数越大则矩阵就越病态，L 中小的数据扰动就能使得结果产生很大变动，解的误差因此就可能会很大。为了解决这一问题，我们首先对矩阵 L 做一番处理改造，其中使用文献^[47]方法，该文献中讨论了对矩阵条件数如何降低问题，使用了处理图像坐标方法：坐标原点平移和归一化处理。其中坐标原点平移就是平移坐标原点使其从原来的位置到质心点位置，并换用原位置点与质心点的相对量表示坐标，这样一来坐标就形成了一个绕质心点的均衡分布，坐标值的数量级就会明显降低。对于此处平面单应性这种问题，质心就是图像中所有的坐标点求和后取得的平均值。而归一化处理分为两种：线性归一化是指处理的时候保持两坐标轴方向数据缩放的比例相同，而非线性归一化就是处理的时候两坐标轴方向数据缩放的比例不同，因此经过两种方法变换后的各点到原点的距离平均值分别是 $\sqrt{2}$ 和 1。本文中处理时采用的是非线性归一化的方法。

对于 $Lx=0$ 这样的线性齐次方程组有不同的解法^[46]，其中使用奇异值分解 (Singular Value Decomposition, SVD) 的方法求齐次解是常用方法。本文在这里单应矩阵的求解中也是首先用奇异值分解来求解，当点对数 $n>4$ 时，用此分解得到的解作为初始值。

SVD 分解是一种线性代数中重要的正交矩阵分解，经常用来解决最小二乘估计问题，在统计学及信号分析等领域都有重要应用，矩阵的奇异值分解是将一个矩阵分解成一对角矩阵与两正交矩阵： $SVD(A)=[U,S,V]$ 。对于 $m \times n (m>n)$ 的矩阵 A，设特征值为 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r \geq \lambda_{r+1} = \dots = \lambda_n = 0$ ，r 为 A 的秩，是非零对角元的个数，则有非零对角元 $\sigma_i = \sqrt{\lambda_i} \dots (i=1,2,3 \dots r)$ 为矩阵 A 的奇异值。那么存在 m 阶和 n 阶幺正矩阵 U、V，使得

$$A = U \begin{bmatrix} \sum & 0_{rx(n-r)} \\ 0_{(m-r) \times r} & 0_{(m-r) \times (n-r)} \end{bmatrix} V, \text{ 其中 } \sum = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \dots & \\ & & & \sigma_r \end{bmatrix}. \text{ 因此可以将 } A \text{ 分解为三个}$$

矩阵相乘积 $A=USV$, U 、 V 中列向量 $u_i(i=1,\dots,M)$ 、 $v_i(i=1,\dots,N)$ 分别称为矩阵 A 的左、右奇异向量, 构成一组标准正交基, S 是一非标准对角矩阵。

求单应性矩阵, 也就是求方程组 $Lx=0$ 的解 x 的过程是: 当 $n=4$ 时, L 为 8×9 矩阵, 直接对 L 进行 SVD 分解即可, 因为这种情况下只有 8 个奇异值, 但有 9 个右奇异向量, 方程组有精确的解, 此时将零也暂时认为是 L 的奇异值, 则零奇异值所对应的那个右奇异向量就是解; 当 $n>4$ 时, 如果直接对 L 做 SVD 分解, 过程中产生的 U 会是 $2n \times 2n$ 的大矩阵, 点较多时, 这一矩阵是比较庞大的, 消耗内存, 为避免大矩阵减少运算的时间, 在方程左右同乘 L^T , 转化成 9×9 的方阵 $L^T L$, 对其做奇异值分解:

$$[U, S, V] = \text{SVD}(L^T L)$$

$$S = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \sigma_9 \end{bmatrix}$$

此时 S 成为了一个标准的对角矩阵, $\sigma_i(i=1,2,\dots,9)$ 是与矩阵 L 相对应的特征值, 依然按照从大到小顺序排列, 那么与最小的特征值 σ_9 相对应的右特征向量 v_9 , 也就是矩阵 V 中的最后一列就是满足 $Lx=0$ 这一方程的初始解。最终把解出来的列向量 x 调整成 3×3 的矩阵, 其中每个元素都与矩阵中最后那个元素相除, 使该矩阵中最后一元素为 1, 即得要求的单应矩阵 H 。

将已有的图像点 (u_i, v_i) 进行的处理过程和线性方程求解过程综合, 得到求解视频中两帧图像间单应矩阵过程如下:

1. 分别求出图像所在两坐标轴上各个坐标的和, 然后求其平均值, 得到质点: $p_u = \text{mean}(u_i)$ 、 $p_v = \text{mean}(v_i)$;
2. 再求出各个点相对于质心点的差, 用相对量表示各点坐标: $\Delta u_i = u_i - p_u$ 、 $\Delta v_i = v_i - p_v$ 。由此 1、2 两步完成坐标原点的平移操作;

3. 求图像两个坐标轴上各自的归一化缩放因子:

$$s_u = 1/\text{mean}(\text{abs}(\Delta u_i))、s_v = 1/\text{mean}(\text{abs}(\Delta v_i));$$

4. 由前 3 步得到视频帧图像中坐标变换使用的关系是
$$\begin{cases} u_i' = \frac{\Delta u_i}{\text{mean}(\text{abs}(\Delta u_i))} \\ v_i' = \frac{\Delta v_i}{\text{mean}(\text{abs}(\Delta v_i))} \end{cases},$$

$$\text{表示成矩阵变换形式就是} \begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} = \begin{bmatrix} s_u & -s_u p_u & 0 \\ 0 & s_v & -s_v p_v \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix},$$

$$\text{则其反变换矩阵就是 } H_{\text{inv}} = \begin{bmatrix} 1/s_u & 0 & p_u \\ 0 & 1/s_v & p_v \\ 0 & 0 & 1 \end{bmatrix};$$

5. 用变换矩阵对已知图像点坐标进行平移与归一化处理, 然后将对应点坐标相结合构造 $Lx=0$ 方程组, 得到 L 矩阵, 并通过 SVD 分解及高斯-牛顿迭代法, 求出单应性矩阵 H_{norm} ;
6. 将这个单应性矩阵应用到图像中其它已经过坐标变换的点, 得到图像平面变换后的图像坐标, 要求出真实图像平面上点的坐标, 就需要使图像坐标反变换: $\bar{p}_i = H_{\text{inv}} \bar{p}_i'$ 。

3.4 跟踪结果

从以上对帧间单应性矩阵的分析, 我们可以使用相邻两帧 f_t 、 f_{t+1} 已知特征点求取单应性矩阵, 然后将其应用到 f_t 中额头点, 计算 f_{t+1} 中这些特殊特征点位置。然而因为存在误差累积作用, 并考虑到第一帧中我们手工标记的特征点位置的准确性, 最终本文使用第一帧作为参考帧, 因此以主动外观模型跟踪点为基础, 结合单应性思想计算跟踪特征点的过程为: 将以后每一帧由 AAM 跟踪到的脸部侧边的点和鼻子周围点与第一帧中相应的点进行计算, 得到帧间图像单应性矩阵, 然后将该矩阵作用在第一帧额头点上, 得到在当前帧的额头点位置。这样有效利用单应性思想关联视频中两帧图像, 从而克服 AAM 弱点, 产生我们所需要的特征点集, 如图 3-4 所示。

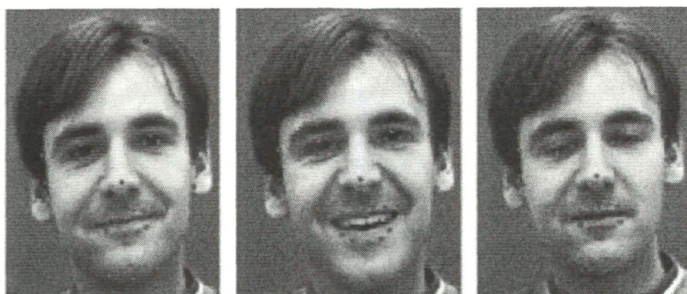


图 3-4 视频中跟踪到的人脸特征点
 (绿色标记 AAM 的跟踪结果 红色标记计算到的额头点)
 Figure 3-4 Feature points tracked in the video
 (Green marks the AAM result, Red marks the calculated forehead points)

3.5 本章小结

本章描述了实现脸谱绘制的第二阶段，视频中人脸以及器官轮廓位置的跟踪，也就是定位我们需要的人脸特征点，为后续变形提供控制点。研究了主动外观模型和单应矩阵，通过对人脸表情的分析，选择了总共 44 个点作为特征点，包括每只眼睛轮廓的 4 个点、嘴部轮廓的 12 个点以及脸部外轮廓的 24 个点（其中 9 个额头点）。在 AAM 进行跟踪的同时，提出了用单应性解决额头点位置计算的问题，得到较好的效果。但是，在推导分析使用帧间单应性矩阵过程中，我们使用了假设条件，并且 AAM 跟踪中也会出现误差，这些会影响最终结果。

第 4 章 京剧脸谱的变形映射

脸谱图如何有效的映射到人脸,是实现本文研究目标的关键问题。对于映射,首先想到图形学中纹理映射内容,它将纹理空间中的纹理像素映射到屏幕空间中像素,是真实感图像制作的一个重要部分,使用 OpenGL 就可以很容易将脸谱图纹理映射到人脸,但过程中相邻接的三角面片边缘处纹理会出现偏差走样,而且人脸外边缘及眼睛、嘴器官边缘轮廓是分段而非平滑的,使得整体脸谱纹样无法实现光滑流畅效果,明显的不真实。经过对京剧脸谱自身结构及勾画过程的研究,为了能表现较好的匹配真实感,并良好的保持其形状特征意义,我们需要将每个纹样按整体“绘制”到人脸相应位置,使它适应具体人脸的形态,将标准脸谱图看作原形状,具体人脸看作目标形状,于是就形成了一个对象形状变形问题,因此提出使用变形技术来实现京剧脸谱的映射。

空间变形技术最初是在几何造型领域作为一种造型方法提出来的,后来由于它在模拟柔性物体动画^[51]方面的良好表现,能对对象形状做扭曲变形,将其变化到画师需要的形态,逐渐地在计算机动画领域中也得到发展并广泛的应用。变形技术使用少量点控制变形整体,使用户可以用简单、灵活、直观的交互来对模型几何形状进行修改操作,便于完成造型。本章对经常使用的流行变形算法进行分析,并在研究过程中进行应用实验,比较结果,联系视频这种特殊应用对象,从轮廓的平滑、脸谱纹样的保持及计算时间方面,最终提出使用基于曲线的移动最小二乘变形实现京剧脸谱在视频中人脸的映射的方法。

空间变形^[48]技术的研究在初期的时候其控制盒为格线框(lattices),建立变形对象与相应格线框顶点的局部坐标关系,如传统的 FFD 变形技术。后来的研究中以包围盒(cage)为控制盒,包围盒在二维中就是形状与被包围在其里面的对象物体相似的多边形,就像对对象轮廓较为简单的勾勒。包围盒内的点都可以由包围盒的各顶点与相对应的特殊权重函数的乘积多项式和来表示,这里权重函数就被称为坐标(coordinates),如 Mean Value Coordinates、Harmonic Coordinates、Green Coordinates 等,包围盒的同构变化就能使其内部物体产生平滑的空间变形。基于

包围盒的空间变形技术的主要优点就是它们简单、灵活而且因为盒内对象点的重心坐标可以进行预先计算，所以比较快速。点变形算法的变形函数是完全以对离散特征点的插值为基础，这种方法提供了一个比包围盒更加直观的用户界面。

4.1 自由变形技术

自由变形 (Free Form Deformation, FFD) 是由 Sederberg 等人提出的一种被广泛使用的曲面造型技术^[15]。它的思想是先构造一个围绕物体的长方形或长方体控制格线框，为给出的区域内的每个点定义了一个局部的新位置，建立物体和控制格线框顶点之间的关系，然后移动格线框上的控制点使之变形，这样内部的物体也就会因此而变形。这里我们主要讨论二维上的变形。

首先建立一个变形域，用 $(X_{\min}, Y_{\min})$ 和 $(X_{\max}, Y_{\max})$ 来设定它的左下和右上角点，然后根据变形框位置及实际应用要求生成控制网格点，假设定义自由变形控制网格划分成 $m+1$ 个横栏和 $n+1$ 个纵栏，即 FFD 函数的度为 m, n 。初始位置时，控制网格位置 i, j 处的控制点值 P_{ij} 的真实 (x, y) 坐标为：

$$P_{ij} = (X_{\min} + \frac{i}{m}(X_{\max} - X_{\min}), Y_{\min} + \frac{j}{n}(Y_{\max} - Y_{\min})) \quad (4-1)$$

依据一个有理二变量张量积伯恩斯坦多项式来定义变形公式：

$$X(s, t) = \frac{\sum_{j=0}^n \sum_{i=0}^m w_{ij} B_i^m(s) B_j^n(t) p_{ij}}{\sum_{j=0}^n \sum_{i=0}^m w_{ij} B_i^m(s) B_j^n(t)} \quad (4-2)$$

其中， $B_i^m(s)$ 和 $B_j^n(t)$ 是贝塞尔混合函数， (s, t) 是变形框内点的局部坐标，其值在 0 和 1 闭区间内，如图 4-1 所示的那样，对于一个在变形框中的点 (x, y) 计算其 (s, t) 坐标为：

$$s = \frac{x - X_{\min}}{X_{\max} - X_{\min}}, t = \frac{y - Y_{\min}}{Y_{\max} - Y_{\min}} \quad (4-3)$$

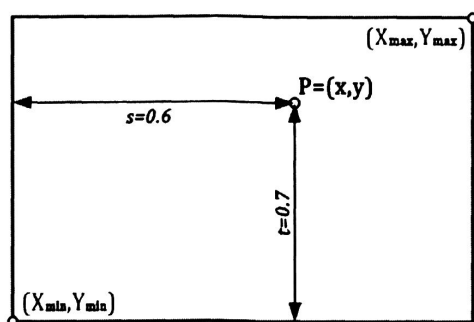


图 4-1 FFD 局部坐标
Figure 4-1 FFD local coordinates

这样自由变形的时候，首先根据式 4-3 计算出该点所对应的局部(s,t)坐标，若 s,t 的坐标数值在[0,1]区间内，通过用式 4-2 就能够得到变形后该点对应的新位置；否则，不变。FFD 变形中，一个控制点的移动就能够影响整个变形框中包含对象上的所有点，对于对象部分区域在变形框中情形，也可以通过控制变形的范围，使变形和未变形区域之间连续。

蔡飞龙的论文中使用 FFD 变形表情单元，从而使得脸谱产生各种不同的表情。如图 4-2 所示纹样的 FFD 变形。他们以 FACS 动作单元标准为参考，将脸谱做的表情看作动作单元的组合，通过纹样坐标的最大值和最小值，及纹样变形幅度设定控制点数来构造变形格线框，从而将完成表情所需要的信息分散到每个纹样的动作，例如抬右眉外侧、抬嘴角左侧等。

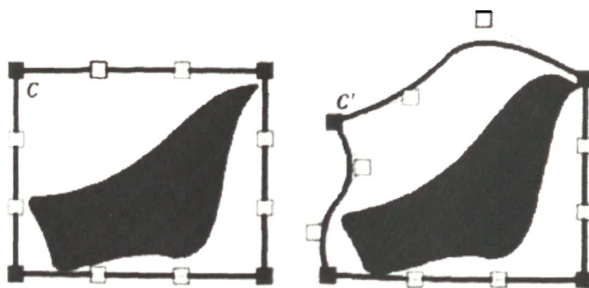


图 4-2 纹样变形前和 FFD 变形后（文献[14]）

Figure 4-2 The pattern before the deformation and after the FFD deformation[14]

虽然自由变形技术可以和任何实体一起使用，并且整体和局部都能使用，有良好的表现。但是，FFD 同时也存在一些局限性。鉴于实际模型的形状，复杂情况下控制点的放置的可控性不是很好，另外，模型表面的不连续或者洞也是比较难控制的。对复杂变形的实现非常困难，控制顶点的移动产生的变形结果难于把握，因此物体上点的准确位移不易控制，也就是对物体形状的准确控制难于实现。

直接操纵的自由变形 DFFD(Direct Manipulation of Free-Form Deformation)^[49]是对 FFD 的一种改进。思路是：对于变形对象上的一点，将其移动到要求的变形后位置，通过该点变换反求出它所在控制网格顶点的变化，然后计算变形对象上其余点位置。这是一种间接变形的方法，对于变形对象上点的精移动 DFFD 容易实现，但它也难于控制变形对象的形状。

在解决我们的问题时，原脸谱图上可以比较容易的设定包含纹样的 FFD 变形框，但是对于视频中的人脸，变形框的位置以及变形框的形变的确定都比较困难。我们只有 AAM 跟踪到的特征点信息，即使 DFFD 能通过对象上一点位置的变化，反求出变形框上控制点的变化，进而来计算对象上剩余点。然而 DFFD 使用的信息太少，它在眼睛、嘴等轮廓处的光滑配准会存在很大问题，且计算量会比较大。

4.2 基于重心坐标的变形

重心坐标的概念最早是在 1827 年由 Möbius^[50]引入，解释了计算机图形学和几何建模中的一个基本问题——多边形内部任一点能否由多边形顶点表示成一个线性组合的形式。那么，什么是重心坐标呢？给出顶点为 v_i 的闭合多边形内的一个顶点 v ，如简单的一个三角形。目标是计算一系列的权重使得能够用这个多边形的带权重的顶点集合来表示 v ，我们称这些归一化的权重为重心坐标。三角形内任一点可以由三个顶点的线性组合形式来表示，其中的三个权重系数是唯一的，就是此点在这三角形中的重心坐标。而四边形内任一点可以将进行双线性插值得到的系数的张量积看作此点在这四边形中的重心坐标。推广至平面任意多边形，问题是以何种方式定义重心坐标。在图形图像中我们应该对这种形式的坐标很熟悉，因为我们经常在应用中使用，如纹理，shading。但是重心坐标也有其他的用处，在这期间三角形的重心坐标得到延伸和发展并被引进到几何变形中并产生出一种 cage-based 的几何变形方法。对于我们的变形目标，已知前一章得到的人脸特征点，通过连接可以构成一个有两个眼睛和一个嘴巴处“洞”的复杂多边形，要变形的点位于此多边形内部，因此尝试使用基于重心坐标的变形方法。

到目前来讲，应用到变形中的主要重心坐标形式有 Floater 等提出的均值坐标 (Mean Value Coordinates) ^[51,52,53,54]，此坐标计算简单，易于变形，但在非凸形控

制盒情况下会产生负值，不满足内部局部性。对于均值坐标的这一缺点，Joshi 等在求解拉普拉斯方程基础之上提出了有着非负性的调和坐标（Harmonic Coordinates, HC）^[55,56]，调和坐标利用了非形式化的求解过程，该坐标具有非负性，但是却有很高的空间和时间复杂度，计算效率低。Lipman 等人也在此提出了正均值坐标（Positive Mean Value Coordinates, PMVC）以及格林坐标（Green Coordinates, GC）^[57,58,59]，两者也都具有非负性，其中正均值坐标使用 GPU 进行求解，虽然计算速度得到了改善，但是它不能够满足变形的光滑性，格林坐标是由格林第三公式推出，它也是全局的，不满足于内部局部性，但它是非负的保形的变换。

4.2.1 均值坐标

用数学形式描述平面多边形内任一点的重心坐标为：设平面 n 边形 P 的顶点为 $v_i, i=1, \dots, n$ 并按逆时针排列，并设此多边形内部或多边形边界上任一点为 v_0 ，如图 4-3 所示。若存在依赖于各顶点 v_i 和点 v_0 的一组非负系数 $\lambda_i \geq 0, i=1, \dots, n$ ，使得下式成立

$$\sum_{i=1}^n \lambda_i v_i = v_0 \quad (4-4)$$

$$\sum_{i=1}^n \lambda_i = 1 \quad (4-5)$$

其中式 4-4 将 v_0 表达成多边形点 v_i 的加权凸组合，那么 λ_i 即是点 v_0 的重心坐标。通常要求重心坐标 λ_i 函数是光滑的，并且在多边形边界上重心坐标要有线性性质和拉格朗日性质。

如果不要求坐标要具有的一些特定性质，问题就很容易解决了，由于 v_0 在 $v_1, \dots, v_n$ 构成的凸多边形内部，则必然在 v_0 邻域有包含 v_0 的三角形 $T=[v_{i1}, v_{i2}, v_{i3}]$ ，因此我们可以得到三角形 T 内 v_0 的重心坐标 $\lambda_{i1}, \lambda_{i2}, \lambda_{i3}$ ，然后令其它多边形顶点的系数为 0，这样 $\lambda_{i1}, \lambda_{i2}, \lambda_{i3}, 0, \dots, 0$ 就可以看做点 v_0 关于多边形的重心坐标。然而，这些坐标随机地依赖于选择的三角形，得到的重心坐标一般不连续。现在许多研究致力于构造从三角形到其他类型多边形的重心坐标的一般形式，即是寻找对 λ_i 的表示方法。首先第一个是 Wachspress^[60]，针对有限元计算，在 1975 年，他创建了一个对平面内凸多边形的求重心坐标的一般形式，对于凸多边形， W 坐标在多边形

内部是良好定义的，但在外部不是，有噪声区域，接下来 Meyer 等人^[61]使用角度的推导公式，也只是适合于凸多边形。而后，Floater 提出均值坐标（Mean Value Coordinates），被简称作 MVC，它是最经常使用的重心坐标之一。这一均值坐标满足我们需要的重心坐标性质，在凸多边形上均为正，并能推广到非凸多边形内部仍为正，可应用到星形多边形甚至扩展到三维空间。均值坐标由调和函数的均值定理来推出，应用到分段线性边界轮廓的多边形。

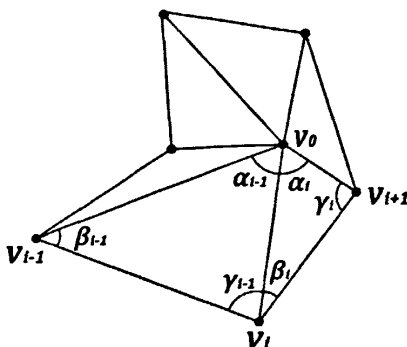


图 4-3 星形多边形

Figure 4-3 Star polygon

设 $0 < \alpha_i < \pi$ ， α_i 是在三角形 $[v_0, v_i, v_{i+1}]$ 内点 v_0 处的角，循环定义，如图 4-4 示。

权重

$$\lambda_i = \frac{w_i}{\sum_{j=1}^n w_j}$$

$$w_i = \frac{\tan(\alpha_{i-1}/2) + \tan(\alpha_i/2)}{\|v_i - v_0\|} \quad (4-6)$$

则是 v_0 关于 $v_1, \dots, v_n$ 的均值坐标具体定义形式。这个权重从均值定理对调和函数的应用中得到，因而称为均值坐标。已知圆周的均值定理：对于一个有着边界 Γ 的圆盘 $B = B(v_0, r) \subset \Omega$ ，以 v_0 为圆心 r 为半径，有 $u(v_0) = \frac{1}{2\pi r} \int_{\Gamma} u(v) ds$ ，其中 u 是平面区域 Ω 上定义的调和函数，满足拉普拉斯等式 $u_{xx} + u_{yy} = 0$ 。因此均值坐标的几何意义为：对多边形内部的点 v_0 ，有以它为中心的圆形邻域 Ω ， v_0 的坐标就可以看做是该圆形邻域里所有点的权值的平均值。通过总结，我们得到对任意平面多边形的精确的归一化均值坐标，对任意简单多边形 Ψ ，MVC 为函数 $\lambda_i: \mathbb{R}^2 \rightarrow \mathbb{R}, (i=1, \dots, n)$

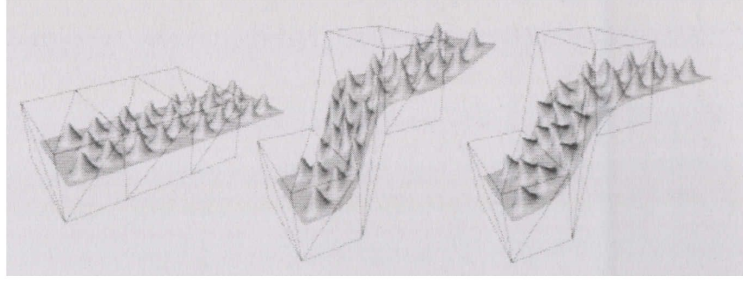
$$\lambda_i = \begin{cases} \frac{w_i(v)}{\sum_{j=1}^n w_j(v)} & \text{if}(v \notin \Psi \text{ with } w_i) \text{即} v \text{ 在多边形内且对应权重 } w_i \\ (1-\mu)\delta_{ij} + \mu\delta_{i,j+1} & \text{if}(v=(1-\mu)v_j + \mu v_{j+1} \in e_j) \text{即} v \text{ 在多边形边上} \\ \delta_{ij} & \text{if}(v=v_j) \text{即} v \text{ 与多边形顶点重合} \end{cases} \quad (4-7)$$

我们已经证明均值坐标能够自然地将三角形上的重心坐标概念推广到没有自交的任意多边形，甚至于这样一些多边形的集合。MVC 有许多重要性质并对于给出的输入多边形的顶点数据的插值特别有用。插值的行为类似于薄板样条插值，但是，它不用必须首先解决匹配问题，因此它能直接在 $O(n)$ 操作。均值坐标是全局的并且不是在任意多边形中到处都是正的，图像变形中只有当源和目标多边形都是凸多边形时，坐标变形函数才能够保证是一对一的。

4.2.2 格林坐标

之前的基于包围盒的变形方法在形状和细节的保持上没有一个好的控制，Lipman 将形状保持作为目标，从格林第三公示推导提出了格林坐标 (Green Coordinates)，引入到图像变形领域中。符合我们的研究要使变形后纹样形状得到保持的要求。

前面基于重心坐标的变形中将包围盒 P 内任意一点 η ，用顶点集 $V=\{v_i\}_{i \in I_v} \subset R^2$ 的线性和表示为： $\eta=F(\eta;P)=\sum_{i \in I_v} \phi_i(\eta)v_i$ ，求的点对应的重心坐标 $\phi_i(\cdot)$ ，这样通过对包围盒进行变形得同构多边形包围盒 P' ，就能够使对象变形： $\eta \rightarrow F(\eta;P')=\sum_{i \in I_v} \phi_i(\eta)v_i'$ 。这些被定义成坐标的变形算子是仿射不变的，所以当包围盒仿射变换时，可以通过这些算子对此变换进行重构，但作为仿射变换，其中可能出现导致无法保持形状的切变和各向异性缩放的因子。而且，从上节 MVC 变形式中可以看出，不同坐标轴在变形时都是独立的，如图 4-4 (b) 所示，这种仿射不变性使得图中细节（小的突起）在部分包围盒平移的下依然保持它原来的方向。



(a)原始变形对象 (b)MVC 变形 (c)GC 变形

图 4-4 MVC 和 GC 的变形效果示例 (文献[58])

Figure 4-4 Comparison of MVC and GC in shape keeping

要作保形变换,其雅可比矩阵应该是由旋转及各项同性的缩放构成,这就需要算子能反映坐标轴间的依赖关系,如图 4-4 (c),变形中突起细节的朝向随变形发生改变,比较真实。为了在保持包围盒算子优点的基础上,突破其局限性,这里的 Green Coordinates 在空间变形中引入适当的旋转来允许形状保持,它充分利用 cage 上的信息,不仅仅是其顶点位置,还包括其面法向,来对内部点位置进行描述。令包围盒是一个导向的简单表现(2D 时为多边形,3D 时为三角面片),用 $P=(V,T)$ 表示,其中 $V=\{v_i\}_{i \in I_v} \subset R^d$ 是顶点, $T=\{t_j\}_{j \in I_T}$ 是简单的面元素(2D 时为多边形的边,3D 时为三角面片中的三角形),并进一步定义 $n(t_j)$ 为包围盒导向简单面的朝外法向量有 $t_j(\|n(t_j)\|=1)$ 。将包围盒内部任意点 η 表示成如下顶点及法向的线性组合:

$$\eta = F(\eta; P) = \sum_{i \in I_v} \phi_i(\eta) v_i + \sum_{j \in I_T} \psi_j(\eta) n(t_j) \quad (4-8)$$

这样,由一个定义的变形后包围盒 P' 引起的变形被定义为:

$$\eta \rightarrow F(\eta; P') = \sum_{i \in I_v} \phi_i(\eta) v_i' + \sum_{j \in I_T} \psi_j(\eta) s_j n(t_j') \quad (4-9)$$

这里 v_i' 和 t_j' 分别代表 P' 的顶点和面, ϕ_i 和 ψ_j 就是格林坐标。放缩因子 $\{s_j\}_{j \in I_T}$ 对于获得像放缩不变性这样重要的性质是必要的,为了合适于形状保持变形 s_j 应该反映出包围盒变形后面 t_j 的拉伸程度,2 维中 s_j 可简单地计算 $s_j = \frac{\|t_j'\|}{\|t_j\|}$, 这里 $\|t_j\|$ 是 t_j 的长度。

由于是使用格林第三等式得到的坐标,因而称为格林坐标。格林第三公式为:令 $u(\xi) \quad \xi=(\xi_1, \dots, \xi_d)$ 是区域 $D \subset R^d$ 上一个调和函数,此区域被分段光滑边界 ∂D 来包

围。当一个标量函数 u 是拉普拉斯方程的一个解时, u 被称为调和函数, 就是说 u 满足 $\Delta u = \frac{\partial^2 u}{\partial \xi_1^2} + \frac{\partial^2 u}{\partial \xi_2^2} + \cdots + \frac{\partial^2 u}{\partial \xi_d^2} = 0$ 。进而假设 R^d 中拉普拉斯方程的基础解为 $G(\cdot, \cdot)$, 也即 $\Delta_\xi G(\xi, \eta) = \delta(\xi - \eta)$, 其中 $\delta(\cdot)$ 为 Dirac delta 函数。那么, 对区域 D 内任何一个点 η , $u(\eta)$ 可以由边界值和边界法向导数表示为:

$$u(\eta) = \int_{\partial D} (u(\xi) \frac{\partial G(\xi, \eta)}{\partial n(\xi)} - G(\xi, \eta) \frac{\partial u(\xi)}{\partial n(\xi)}) d\sigma_\xi \quad (4-10)$$

这里 n 是 ∂D 的导向外法向, $d\sigma_\xi$ 是 ∂D 上的体元, 这一积分正好符合想要推导出来的顶点及法向的线性组合思想。

R^d 中拉普拉斯方程的基本解为:

$$G(\xi, \eta) = \begin{cases} \frac{1}{(2-d)\omega_d} \|\xi - \eta\|^{2-d} & d \geq 3 \\ \frac{1}{2\pi} \log \|\xi - \eta\| & d=2 \end{cases} \quad (4-11)$$

这里 ω_d 是 R^d 中一个单位球的体积。

现在令区域 D 为包围盒 P 内包围的区域, 并使 $u(\eta) = \eta$, 即坐标函数代替 4-10 中调和函数 u 。将积分写成包围盒面上积分之和, 因为每个面 t_j 的法向 $n(t_j)$ 是常量: $\partial \xi / \partial n(\xi) = \partial \xi / \partial n(t_j) = n(t_j)$, 得出:

$$\eta = \sum_{j \in I} \left(\int_{t_j} \xi \frac{\partial G}{\partial n} d\sigma_\xi - \int_{t_j} G n(t_j) d\sigma_\xi \right), \eta \in D^{\text{in}} \quad (4-12)$$

用 $N\{v_i\}$ 表示顶点 v_i 周围一环邻域的所有面集合, 并令函数 Γ_i 是定义在 $N\{v_i\}$ 上的分段线性帽函数, 即在点 v_i 时为 1, 而在一环邻域内所有其他点处为 0, 且在每个面上是线性的。那么将 ξ 写成在单独面 t_j 上的重心组合 $\xi = \sum_{k=1}^d \Gamma_k(\xi) v_k$, 这里 v_k 是面 t_j 的顶点, 从 4-12 可以得到

$$\eta = \sum_{i \in I} \phi_i(\eta) v_i + \sum_{j \in I} \psi_j(\eta) n(t_j), \eta \in D^{\text{in}} \quad (4-13)$$

其中的格林坐标函数 ϕ_i 和 ψ_j 分别是

$$\begin{aligned}\phi_i(\eta) &= \left(\int_{\xi \in N\{v_i\}} \Gamma_i(\xi) \frac{\partial G(\xi, \eta)}{\partial n(\xi)} d\sigma_\xi \right) \quad i \in I_v \\ \psi_j(\eta) &= - \left(\int_{\xi \in t_j} G(\xi, \eta) d\sigma_\xi \right) \quad j \in I_t\end{aligned}\quad (4-14)$$

变形如图 4-5 所示。比较格林坐标和均值坐标，格林坐标具有保形的性质是最重要的优点。格林坐标对包围盒内的扭曲要远小于均值坐标，格林坐标比均值坐标更能保持方形的轮廓。在一定扩展下，格林坐标可以只对局部对象包围盒做形变。但是，在保形的情况下，格林坐标变换后的对象可能会出包围盒现象，这在一些应用中是不被允许的。

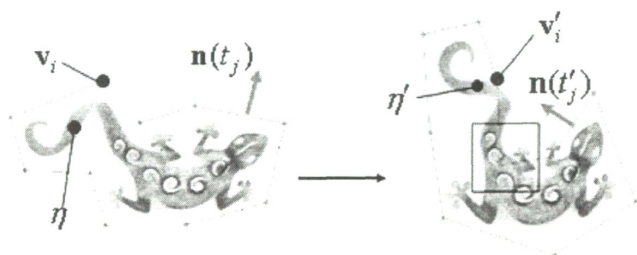


图 4-5 GC 变形效果（文献[58]）

变形对象的形状得以保持，但变形对象部分超出了控制轮廓的边界

Figure 4-5 GC deformation effect (reference [58])

The shape of deformed object is kept, but parts of the object extend out the cage

4.3 移动最小二乘方法

经过理论分析与具体实验，均值坐标和格林坐标方法都存在不足之处，对简单多边形 cage 情形变形良好，但是对于本文研究中形成的复杂多边形，存在变形后轮廓边缘不匹配、不够光滑等问题，需要适用范围更广的变形思想。而且，通过控制包围盒来实现变形的的方法，虽然计算变形速度快，但是编辑不够直观，如何构造好的包围盒也是一个研究问题，难以实现精确的控制。而点变形中特征点可以比较随机，算法的变形函数是以对离散特征点的插值为基础，对点直接控制，较为方便直观。

Schaefer 等^[62]在 SIGGRAPH 上提出使用移动最小二乘（Moving Least Squares, MLS）方法实现图像变形，通过建立基于控制点、控制直线段的图像变换映射函数，经优化求解实现图像不同的变形控制。并在 2010 年提出了 Moving Least Squares Coordinate^[63]实质上也是重心坐标家族中的一员，同样具有包围盒变形中各项性

质。它对任意的二维多边形均使用，易于计算，对开多边形也有线性精确度。

4.3.1 基于点的移动最小二乘图像变形

对图像进行变形可看作是对原图像上的每个点 v 应用变形函数 f 进行计算，将原图像变换到目标图像的过程。设原图像中各控制点组合成的集合是 p ， q 为 p 对应的目标图像各控制点组成的集合。根据 Levin 等人^[64]提出的移动最小二乘理论模型，必定存在一个变形函数 f ，对图像中任意一个像素点 v ，能够使得下面式子取得最小值：

$$\sum_i w_i |f(p_i) - q_i|^2 \quad (4-15)$$

这里， p_i 和 q_i 是对应图像控制点的坐标。 w_i 是权重函数，它的表达式是：

$$w_i = \frac{1}{|p_i - v|^{2\alpha}}, \text{ 其中 } \alpha \text{ 是调节参数，一般取 } 1。$$

此最小二乘问题中权重函数值 w_i 取决于 v 在图像中的位置。很明显，当 v 趋近于点 p_i 时， w_i 趋近于无穷，此时可定义 $f(p_i) = q_i$ ；而且，若 $p_i = q_i$ ，则定义 $f(p_i) = p_i = q_i$ 。最终，变形函数 f 在各处平滑。

通常，变形函数 $f(x)$ 可以包含两部分：线性变换矩阵 M 和平移变换矩阵 T 。

$$f(v) = vM + T \quad (4-16)$$

将上式代入到式 4-15 求其最小值，得到下面表达式：

$$T = q - p \cdot M \quad (4-17)$$

$$\text{其中, } p = \frac{\sum_i w_i p_i}{\sum_i w_i}, \quad q = \frac{\sum_i w_i q_i}{\sum_i w_i}。$$

这样，可以将 T 用式 4-17 代替，依据线性矩阵 M 重写 $f(x)$ ，得到变形函数的一般形式：

$$f(v) = (v - p)M + q \quad (4-18)$$

因此，式 4-15 的最小二乘问题就可以改写成：

$$\sum_i w_i |\bar{p}_i M - \bar{q}_i|^2 \quad (4-19)$$

其中， $\bar{p}_i = p_i - p$ ， $\bar{q}_i = q_i - q$ 。这里的矩阵 M 可以看作是一个普通的仿射矩阵，也可以对不同类型的仿射矩阵 M 进行研究，得到仿射变换变形、相似变换变形以

及刚性变换变形三种不同的变形效果。

1. 仿射变换

使用经典的标准方程解最小化等式 4-19

$$M = (\sum_i \mathbf{p}_i^T \mathbf{w}_i \mathbf{p}_i)^{-1} \sum_j \mathbf{p}_j^T \mathbf{q}_j \quad (4-20)$$

尽管这个解要求一个矩阵的逆，但是这个矩阵的大小是常量(2x2)，可以快速求逆。有了这个 M 闭合形式解，我们就可以将仿射变换的变形函数用 $f_s(v)$ 表示，形成简单表达式：

$$f_s(v) = (v - \mathbf{p}_0) (\sum_i \mathbf{p}_i^T \mathbf{w}_i \mathbf{p}_i)^{-1} \sum_j \mathbf{p}_j^T \mathbf{q}_j + \mathbf{q}_0 \quad (4-21)$$

对图像中每个点应用这个变形函数进行计算就能产生一个新的变形了的图像。

当用户通过操作点集 $\mathbf{q}$ 来进行变形时，点集 $\mathbf{p}$ 是固定不变的，因此，式 4-21 中的部分项可以先进行预计算从而减少变形时间。重写等式 4-21 为

$$f_s(v) = \sum_j A_j \mathbf{q}_j + \mathbf{q}_0 \quad (4-22)$$

其中 A_j 是一个单独的标量，有 $A_j = (v - \mathbf{p}_0) (\sum_i \mathbf{p}_i^T \mathbf{w}_i \mathbf{p}_i)^{-1} \mathbf{p}_j^T$ 。对于任一点 v ， A_j 可以别预计算产生一个简单的加权和。

仿射变换包含平移、旋转、非均匀缩放和切变，因此变形后的图像看起来不真实。

2. 相似变换

相似变换是一种只包含有平移、旋转和同比缩放的特殊仿射变换形式。变形函数要做到相似变换，需要约束矩阵 M ，使其具有性质 $M^T M = \lambda^2 I$ (λ 为标量)。如果 M 是块矩阵，形如： $M = (M_1 \ M_2)$ ，其中 $M_1 \ M_2$ 是长度为 2 的列向量，那么当 $M_1^T M_1 = M_2^T M_2 = \lambda^2$ 并且 $M_1^T M_2 = 0$ 时， M 为相似变换矩阵，此约束意味着 $M_2 = M_1^\perp$ ， $\perp$ 是一个向量操作符， $(x, y)^\perp = (-y, x)$ 。式 (4-25) 的最小化问题于是变成找到列向量 M_1 使下式最小：

$$\sum_i w_i \left\| \begin{pmatrix} \mathbf{p}_i \\ -\mathbf{p}_i^\perp \end{pmatrix} M_1 - \mathbf{q}_i \right\|^2 \quad (4-23)$$

由此得到变换矩阵 M 为：

$$M = \frac{1}{\mu_s} \sum_i w_i \begin{pmatrix} \bar{p}_i \\ \bar{p}_i^\perp \end{pmatrix} \begin{pmatrix} \bar{q}_i^\perp & \bar{q}_i^{\perp T} \end{pmatrix} \quad (4-24)$$

其中, $\mu_s = \sum_i w_i \bar{p}_i \bar{p}_i^T$ 。

当原图像控制点集 p 固定的时候, 通过操纵 q 就可以产生不同的相似变形。定义量 A_i :

$$A_i = w_i \begin{pmatrix} \bar{p}_i \\ -\bar{p}_i^\perp \end{pmatrix} \begin{pmatrix} v \cdot p_i \\ -(v \cdot p_i)^\perp \end{pmatrix}^T \quad (4-25)$$

则变形函数 f 转化成了相似变换变形函数 $f_s(v)$:

$$f_s(v) = \sum_i \bar{q}_i \left(\frac{1}{\mu_s} A_i \right) + q. \quad (4-26)$$

式中, q_i 和 A_i 只依赖于原图像的 p_i 、 v , 可以进行提前计算, 提高变形速率。

图形相似变换不改变图形中角的大小, 过程中保持形状不变, 相似 MLS 变形是一种保形变换, 在应用中有好的表现。

3. 刚性变换

刚性变换也是仿射变换一种特殊形式, 相比相似变换, 刚性变换不包含同比例缩放, 近几年来一些研究表示有些情况下, 对真实感的形状变形应尽量刚性。变形上的传统研究使用非线性约束 $M^T M = I$ 来满足刚性变换问题, Horn^[66]证明了最理想的刚性变换可在包含点 p_i 和 q_i 的协方差矩阵的特征值和特征向量中找到。我们通过接下来的定理证明了刚性变换与前一节的相似变换有关。

定理: 令 C 是使下面相似函数最小的矩阵

$$\min_{M^T M = \lambda^2 I} \sum_i w_i |\bar{p}_i M - \bar{q}_i|^2$$

如果 C 被写成形式 λR , 其中 R 是一个旋转矩阵, λ 是一个放缩量, 则旋转矩阵 R 能最小化下面刚性函数

$$\min_{M^T M = I} \sum_i w_i |\bar{p}_i M - \bar{q}_i|^2$$

任意维这一定理都是成立的, 然而在二维中易于应用。使用这一定理, 我们发现刚性变换与等式 4-30 非常相像, 除了不同的常量 μ_s 使得 $M^T M = I$, 表示成

$$\mu_r = \sqrt{(\sum_i w_i \bar{q}_i \bar{p}_i^T)^2 + (\sum_i w_i \bar{q}_i \bar{p}_i^{\perp T})^2} \quad (4-27)$$

不同于相似变换中的 $f_s(v)$ 那样，无法预计算刚性变换函数 $f_r(v)$ 中很多内容，因此要比相似变换慢。令 $\bar{f}_r(v) = \sum_i \bar{q}_i A_i$ ，这里 A_i 同式 4-31 中定义相同，向量 $\bar{f}_r(v)$ 是对向量 $v-p$ 的一个旋转放缩量。为了计算 $f_r(v)$ ，对 $\bar{f}_r(v)$ 做标准化处理，并用 $v-p$ 的长度对它进行放缩，然后依据 q 平移，得 $f_r(v)$ 的表达式为：

$$f_r(v) = |v-p| \frac{\bar{f}_r(v)}{|\bar{f}_r(v)|} + q. \quad (4-28)$$

在普通图像变形应用中刚性变换能够比相似变换得到较真实的结果，但差别没有特别明显。考虑到视频这种实际应用效率并经过试验对比，本文中采用相似 MLS 变形，保持变形过程中脸谱纹样形状不变。图 4-6 显示的是基于点的移动最小二乘相似变换的变形结果，从细节放大的小图中容易看出眼睛和嘴处脸谱与真实人脸有大致的配准，但是轮廓变形不流畅，边缘控制点处出现不光滑凹凸状。

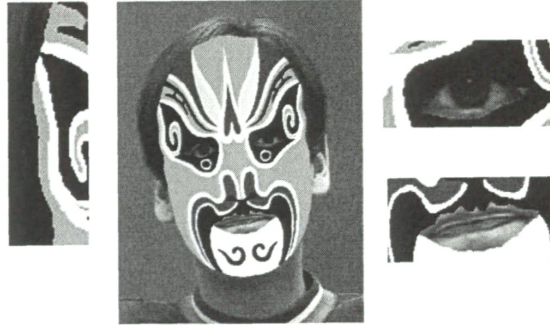


图 4-6 基于点的移动最小二乘相似变换结果
Figure 4-6 Result of MLS similarity transformation based on points

4.3.2 利用控制曲线的图像变形

在某些实际应用中，基于控制点集的变形不能实现所希望达到的变形目的，如图像中的轮廓，应用过程中轮廓变形后不光滑，如图 4-7 所示。Scheafer 提出过基于直线段的变形，通常针对用线段能够表示轮廓的情况下。考虑到大部分图像中的形状、轮廓信息，如人脸，通常由曲线可以很好拟合图像中的形状拓扑关系或轮廓信息，因此利用曲线优势控制变形操作，能够得到更加细致的变形效果。

假设 $p_i(t)$ 是原图像上的第 i 条控制曲线， $q_i(t)$ 是相应的变形后曲线， $t \in [0,1]$ 。

根据式 (4-21) 通过在每条控制曲线上进行积分可以得到二次函数。

$$\sum_i \int_0^1 w_i(t) |p_i(t)M + T - q_i(t)|^2 \quad (4-29)$$

权重 $w_i(t) = \frac{|p_i'(t)|}{|p_i(t) - v|^{2\alpha}}$, $p_i'(t)$ 是 $p_i(t)$ 的切矢量。式 (4-35) 对 T 依然是二次的,

所以 T 仍可由矩阵 M 表示为:

$$T = q_* - p_* M$$

只是, p_* 和 q_* 的表达式变为:

$$p_* = \frac{\sum_i \int_0^1 w_i(t) p_i(t) dt}{\sum_i \int_0^1 w_i(t) dt}, \quad q_* = \frac{\sum_i \int_0^1 w_i(t) q_i(t) dt}{\sum_i \int_0^1 w_i(t) dt}$$

因此, 依据 M 将式 (10) 重写为:

$$\sum_i \int_0^1 w_i(t) |\bar{p}_i(t)M - \bar{q}_i(t)|^2 \quad (4-30)$$

其中, $\bar{p}_i(t) = p_i(t) - p_*$, $\bar{q}_i(t) = q_i(t) - q_*$ 。

为了表示光滑的轮廓, 对于已有的轮廓上的控制点, 我们用曲线对它们进行分段拟合, 曲线需要通过控制点, 并且在控制点处二阶连续, 因此选择三次参数样条曲线来作为控制曲线。这样本文中使用的 $p_i(t)$ 和 $q_i(t)$ 为三次参数样条曲线, 设 a_{i-1} 和 a_i 是 $p_i(t)$ 的起止点坐标, a'_{i-1} 和 a'_i 是 $p_i(t)$ 起止点的一阶导数 (由自由端点条件求出)。那么 $p_i(t)$ 对应的曲线方程表达成矩阵形式是:

$$p_i(t) = T A_i^T \quad (4-31)$$

其中, $T = [t^3 \ t^2 \ t \ 1]$,

$$A_i^T = [A_{3i} \ A_{2i} \ A_{1i} \ A_{0i}]^T = \begin{bmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} a_{i-1} \\ a_i \\ a'_{i-1} \\ a'_i \end{bmatrix}$$

同理, 用 b_{i-1} 和 b_i 表示 $q_i(t)$ 的起止点坐标, b'_{i-1} 和 b'_i 表示 $q_i(t)$ 起止点处的一阶导数值, 那么 $q_i(t)$ 对应的曲线方程表达成矩阵形式是:

$$q_i(t) = T B_i^T \quad (4-32)$$

其中, $T = \begin{bmatrix} t^3 & t^2 & t & 1 \end{bmatrix}$,

$$B_i^T = \begin{bmatrix} B_{3i} & B_{2i} & B_{1i} & B_{0i} \end{bmatrix}^T = \begin{bmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} b_{i-1} \\ b_i \\ b_{i-1}' \\ b_i' \end{bmatrix}$$

通过 4-31 和 4-32 两式能够拟合出三次样条曲线, 由此便可以构造控制变形曲线集 P 和 Q。

根据得到的曲线集, 有两种方法进行接下来的变形。最简单的是将 P 和 Q 上的所有对应点作为控制点, 然后利用前一节的基于控制点集变形的方法进行图形变形, 由于控制点很多, 此方法计算量很大效率低。我们在每段曲线上设置间隔数, 控制间隔个数, 如设间隔数为 5, 则每条小曲线分为了 5 段 6 个控制点, 能够减少变形计算量。

另一种方法就是如 Scheafer 推导的基于直线段的变形方法, 使用推导的基于控制曲线段的变形方法^[68]。依然使用相似变换, 那么就有对特定值 λ , $M^T M = \lambda^2 I$ 成立, 与前面基于点的移动最小二乘相似变形类似。

从前面的过程中, 可以看出 $\bar{p}_i(t)$ 和 $\bar{q}_i(t)$ 也可以是写成曲线表达形式:

$$\begin{aligned} \bar{p}_i(t) &= (t^3 \ t^2 \ t \ 1) \cdot (\bar{A}_{3i} \ \bar{A}_{2i} \ \bar{A}_{1i} \ \bar{A}_{0i})^T \\ \bar{q}_i(t) &= (t^3 \ t^2 \ t \ 1) \cdot (\bar{B}_{3i} \ \bar{B}_{2i} \ \bar{B}_{1i} \ \bar{B}_{0i})^T. \end{aligned}$$

将它们代入式 4-23, 则形成使下式取最小值情形,

$$\sum_i \int_0^1 w_i(t) \left[\begin{bmatrix} t^3 & 0 & t^2 & 0 & t & 0 & 1 & 0 \\ 0 & t^3 & 0 & t^2 & 0 & t & 0 & 1 \end{bmatrix} \begin{bmatrix} \bar{A}_{3i} \\ -\bar{A}_{3i}^\perp \\ \bar{A}_{2i} \\ -\bar{A}_{2i}^\perp \\ \bar{A}_{1i} \\ -\bar{A}_{1i}^\perp \\ \bar{A}_{0i} \\ -\bar{A}_{0i}^\perp \end{bmatrix} M_i - \begin{bmatrix} \bar{B}_{3i}^\perp \\ \bar{B}_{2i}^\perp \\ \bar{B}_{1i}^\perp \\ \bar{B}_{0i}^\perp \end{bmatrix} \right]^2 dt \quad (4-33)$$

按照式 4-24 可以得出此处的 M 表达式:

$$M = \frac{1}{k_s} \sum_i \begin{pmatrix} \bar{A}_{3i} \\ -\bar{A}_{3i}^\perp \\ \bar{A}_{2i} \\ -\bar{A}_{2i}^\perp \\ \bar{A}_{1i} \\ -\bar{A}_{1i}^\perp \\ \bar{A}_{0i} \\ -\bar{A}_{0i}^\perp \end{pmatrix} W_i \begin{bmatrix} \bar{B}_{3i}^T & \bar{B}_{3i}^{\perp T} \\ \bar{B}_{2i}^T & \bar{B}_{2i}^{\perp T} \\ \bar{B}_{1i}^T & \bar{B}_{1i}^{\perp T} \\ \bar{B}_{0i}^T & \bar{B}_{0i}^{\perp T} \end{bmatrix} \quad (4-34)$$

这里的权重矩阵 W_i 是对称矩阵:

$$W_i = \begin{bmatrix} \delta_i^{00} & \delta_i^{01} & \delta_i^{02} & \delta_i^{03} \\ & \delta_i^{00} & \delta_i^{01} & \delta_i^{02} & \delta_i^{03} \\ & & \delta_i^{11} & \delta_i^{12} & \delta_i^{13} \\ & & & \delta_i^{11} & \delta_i^{12} & \delta_i^{13} \\ \text{对} & & & & \delta_i^{22} & \delta_i^{23} \\ & & & & & \delta_i^{22} & \delta_i^{23} \\ & & & & & & \delta_i^{33} \\ & & & & & & & \delta_i^{33} \end{bmatrix} \quad (4-35)$$

δ_i 是权重 $w_i(t)$ 和多项式的乘积, 然后进行积分计算得到的积分值:

$$\begin{aligned} \delta_i^{00} &= \int_0^1 w_i(t) t^6 dt & \delta_i^{01} &= \int_0^1 w_i(t) t^5 dt & \delta_i^{12} &= \int_0^1 w_i(t) t^3 dt \\ \delta_i^{11} &= \int_0^1 w_i(t) t^4 dt & \delta_i^{02} &= \int_0^1 w_i(t) t^4 dt & \delta_i^{13} &= \int_0^1 w_i(t) t^2 dt \\ \delta_i^{22} &= \int_0^1 w_i(t) t^2 dt & \delta_i^{03} &= \int_0^1 w_i(t) t^3 dt & \delta_i^{23} &= \int_0^1 w_i(t) t dt \\ \delta_i^{33} &= \int_0^1 w_i(t) dt \end{aligned}$$

最终推导出的基于曲线的移动最小二乘相似变换变形函数 $f_{cs}(v)$

$$f_s(v) = \left(\sum_i (\bar{B}_{3i} \ \bar{B}_{2i} \ \bar{B}_{1i} \ \bar{B}_{0i}) \left(\frac{1}{k_s} A_i \right) \right) + q_s \quad (4-36)$$

$$\text{其中 } A_i = W_i \begin{pmatrix} A_{3i} \\ -A_{3i}^\perp \\ A_{2i} \\ -A_{2i}^\perp \\ A_{1i} \\ -A_{1i}^\perp \\ A_{0i} \\ -A_{0i}^\perp \end{pmatrix} \begin{bmatrix} v - p_* \\ -(v - p_*)^\perp \end{bmatrix}^T, \quad k_s = \sum_i \int_0^1 w_i(t) p_i(t) p_i^T(t) dt.$$

这种基于曲线集合的变形函数计算涉及大量积分运算，为方便计算，对积分进行数值拟合，使用复化辛普森公式，以一定步长划分积分区间，将连续积分转化成了离散区间积分之和。

4.4 变形算法应用结果比较

我们使用了不同的变形算法对脸谱的映射进行试验，将京剧脸谱与视频中人脸上定位的 44 个点作为变形前后图像的控制点，结果如图 4-7 所示，从左到右依次是均值坐标、格林坐标、利用控制曲线的移动最小二乘结果。

进行 MVC 和 GC 变形时，我们将特征点连接，其中 24 个外轮廓特征点按逆时针顺序连接在一起形成一个多边形，包含其中的眼睛和嘴按顺时针序分别连接在一起，形成一个“带洞”的复杂多边形包围盒。从图(a)、(b)中可以看出，使用 MVC 得到的结果不能和器官有好的配准，纹样的形状也发生了一定的扭曲变化，而且脸谱的边缘不平滑。GC 虽然是一种保形变换，在这样一个非简单多边形中，它的器官位置也没有对齐，这跟它为了保持形状而从包围盒中“外扩”是有关系的。

用曲线可以更好的描述变形区域的轮廓信息，因此首先通过点拟合曲线，使用控制曲线的 MLS 图像变形算法来获得更好的效果。图 4-7(c)中结果表明，此时脸谱和人脸准确对齐，而且纹样的形状得到保持，能够实现我们的研究目标。

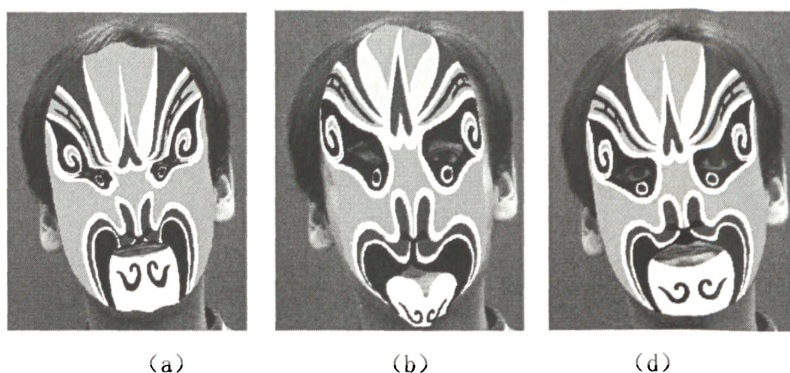


图 4-7 变形结果比较

(a) MVC 映射结果, (b) GC 映射结果, (c) 利用曲线的 MLS 映射结果

Figure 4-7 Experimental comparisons

(a) Result of using MVC, (b) Result of using GC, (c) Result of using proposed method

4.5 本章小结

本章对实现脸谱绘制的最后一个阶段：脸谱映射至真实人脸进行了描述。首先对图像变形技术进行了解并研究分析了几种比较流行的变形技术，然后将这些方法应用到我们的具体映射问题中，比较得到的实验结果。根据本文要实现的目标，人脸与脸谱的对准及形状的保持，选择使用基于曲线的移动最小二乘相似变形方法。

第 5 章 视频中京剧脸谱映射的算法实现

前面几章介绍了完成整个研究目标所经历的三个阶段，对每个问题都进行了具体分析，提出了解决方案。本章我们综合前述内容，对整体算法实现流程及实验结果进行说明展示。

5.1 算法实现流程

以京剧脸谱图和有表情变化的人脸视频作为输入，目标输出戴有给定京剧脸谱图的人脸视频，实现视频中人脸的京剧脸谱映射。如图 5-1 所示。

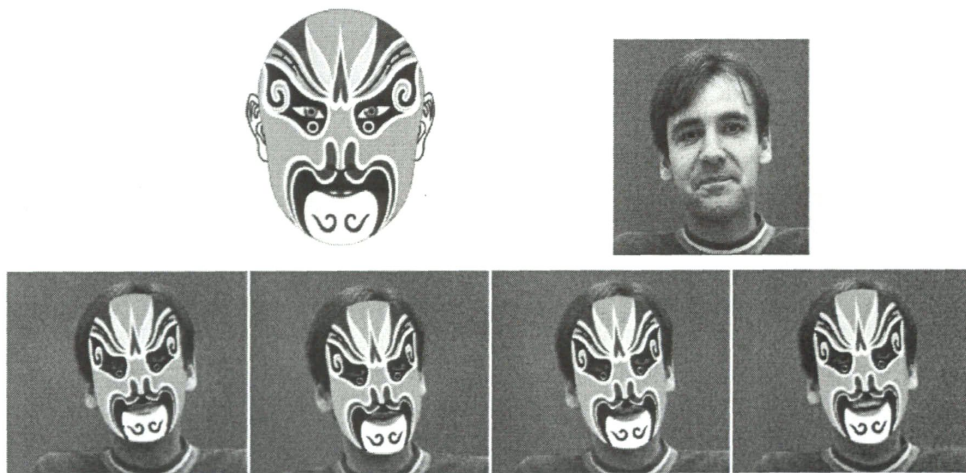


图 5-1 第一行表示输入，第二行表示输出

Figure 5-1 The first line is the input, the second line is the output

算法流程概述为：

预处理：搜索京剧脸谱图，进行矢量化并存储成 SVG 格式文件，入脸谱库。

开始

步骤 1：读取 SVG 格式京剧脸谱文件，提取每个纹样中的分段曲线进行细分，计算得到纹样边界点，用 P 表示整幅脸谱的边界点集；

步骤 2：在脸谱上定义特征点集，由 S 表示，对每个 S 中的点 p_i ；得到在人脸上的对应点 q_i ，并定义这些点的集合为 T ；

步骤 3：对视频中的每一帧

开始

- 步骤 3.1: 跟踪 T 中点 q_i 的新位置;
- 步骤 3.2: 使用移动最小二乘, 得到相似变换函数 f ;
- 步骤 3.3: 对 P 中每个点应用变换函数 f , 得到新点集 P' ;
- 步骤 3.4: 根据变换后的点集 P' , 重新绘制相应纹样的边界曲线,
然后填充相应颜色;

结束

结束

5.2 算法实现结果

对特征点的跟踪过程中, 我们通过曼彻斯特 Tim Cootes 的网页上下载的 am_tools_win_v3 工具来建立 AAM 模型, 使用 IMM 人脸数据库^[67]做训练集。所有的实验在实验室 PC 机上完成, 其配置为: Intel(R) Core 2 Duo CPU, 2.33GHz, 2G 内存。实验工具为: Microsoft Visual Studio 2008, OpenGL 动态链接库。表 5-1 显示了本文方法的时间性能, 虽然使用了纹样轮廓自适应细分点, 大大减少了参与变形计算的点个数, 并对脸谱图进行了某些变形算子的预计算, 但是仍然有较多的点参与较大量的运算, 而且由于控制曲线的使用, 拟合取点, 使控制点增多, 计算量增大。因此变形运行时间长没有实现实时。图 5-2、5-3 中给出了本文方法的更多实验结果。

表 5-1 脸谱映射的执行时间
Table 5-1 Running time of our deformation

Video Facial Makeup	Tim Coote's (s/f)	Qingxiang Wang's (s/f)	Yu Liu's (s/f)
Erdun Dou	8.55	8.55	8.51
Yan Wei	8.21	8.23	8.19
Han Wu	3.41	3.34	3.34
GoldEye Monk	8.58	8.70	8.67
Biao Liu	7.90	7.97	7.98
Yu Xiang	11.52	11.67	11.65
Chenghei Wu	13.47	12.51	12.47

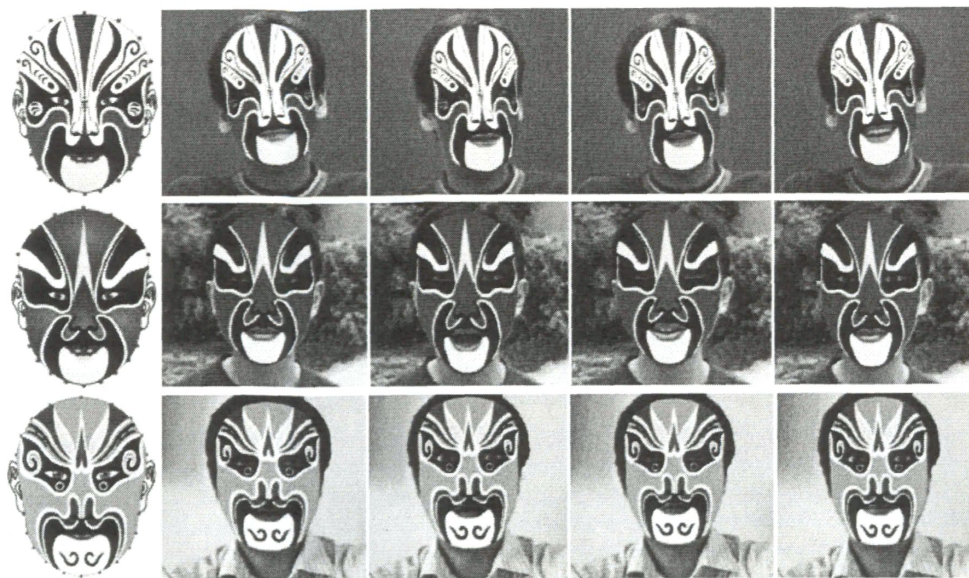


图 5-2 不同脸谱在不同人脸视频中的映射效果

Figure 5-2 Results with different facial masks on different persons



图 5-3 不同脸谱在同一人脸上的映射效果

Figure 5-3 Results with different facial masks on one person

第 6 章 总结和展望

本文主要是针对视频中人脸的京剧脸谱映射的研究,将传统的文化艺术—京剧脸谱与计算机图形图像技术相融合,结合图像矢量化、人脸识别以及图像变形等知识,提出了一中使用图像变形技术映射京剧脸谱到视频中真实人脸的方法,实现了针对不同人脸绘制相应的京剧脸谱的目标。

过程中我们首先对京剧脸谱夸张写实的艺术含义进行了解,探讨了课题研究的背景及价值意义,进而分析京剧脸谱的结构、勾画过程,将脸谱进行矢量化存储,并从 Bezier 曲线拟合的纹样获得细分后的脸谱点,此矢量化的京剧脸谱图存储量小,描述图形语义明确,同时能够减少后面操作的计算量。第二部分对我们所需要的变形控制点进行获取,将人脸上的特征点作为目标,使用 AAM 算法匹配跟踪到脸部、眼睛、嘴巴处的特征点,对于额头上的点 AAM 无法跟踪到的情况下,提出结合单应性,利用单应矩阵求解的方式来获得。第三部分对几种常用的图像变形方法进行分析,比较了它们的优缺点,根据我们此项研究的具体对象情况,应用移动最小二乘方法相似变换来做变形映射,由控制曲线较好的描述区域的轮廓,使绘制得到的脸谱光滑流畅,形状得到良好的保持,而且方法直观简单。

由于使用 AAM 得到脸部特征点,所以在特征点跟踪方面有些来自于 AAM 的限制。当面部不是正向位置的时候,AAM 跟踪到的点的结果会出现较大的误差,AAM 的误差会导致单应矩阵的计算结果与真实情况有误差,从而额头特征点的跟踪位置准确性出现偏差,得到的映射后脸谱会出现抖动情况。另外,实时性问题,也是我们未来研究的一个方向。

参考文献

- [1] 赵梦林 中国京剧脸谱 (Facial Makeup In Beijing Opera Of China) 北京: 朝华出版社, 2003
- [2] 中国戏曲脸谱网 <http://www.xqlpw.com/>.
- [3] 中国脸谱网官方网站 <http://www.lianpu.com.cn/>.
- [4] Erik Hjelmas, Boon Kee. Low Face detection: A Survey. Computer Vision and Image Understanding, Volume 83 Issue 3, September 2001:p236-274.
- [5] W. Zhao, R. Chellappa, P. J. Phillips, A. Rosenfeld. Face Recognition: A Literature Survey. ACM Computing Surveys (CSUR), Journal, volume 35, Issue 4, December 2003:p399-458.
- [6] B. Fasel, Juergen Luetin. Automatic facial expression analysis: a survey. Pattern Recognition, Volume 36 Issue 1, January 2003:p259-275.
- [7] 李思屈 数字娱乐产业 四川: 四川大学出版社, 2006.
- [8] 秦可. 数字娱乐若干关键技术研究[学位论文]. 杭州: 浙江大学, 2005.
- [9] Feilong Cai, Ren Peng, Jinhui Yu. Analysis and Synthesis of Peking Opera Facial Make-ups. Journal of Computer-aided Design & Computer Graphics, Vol 21, No 8, Aug 2009:p1092-1097.
- [10] FaceFilter Studio 2. <http://www.reallusion.com/facefilter>.
- [11] Makeup Pilot. <http://www.colorpilot.com/makeup.html>.
- [12] VirtualFasion Professional. <http://virtual-fashion.com>.
- [13] Virtual MakeOver. <Http://www.dailymakeover.com/makeover/virtualMakeover>.
- [14] Feilong Cai, Jinhui Yu. Modeling and Editing of Peking Opera Facial Makeup Expressions. Journal of Image and Graphics. June, 2010: 958-963.
- [15] Sederberg T W, Parry S R. Free-form Deformation of Solid Geometric Models. ACM SIGGRAPH Computer Graphics, 1986, 20(4): 151-160.
- [16] Sabine Coquillart, Pierre Jancene. Animated free-form deformation: an interactive animation technique. ACM SIGGRAPH Computer Graphics. July 1991:23-26.

- [17] FeiLong Cai, JinHui Yu. A Real-Time Interactive System for Facial Makeup of Peking Opera[J]. Transactions on Edutainment IV, Lecture Notes in Computer Science, 2010: 256-265.
- [18] 蔡飞龙,于金辉,李娟. 基于混合驱动的 2.5 维脸谱动画[C]. 全国第七届智能 CAD 与数字娱乐会议, 2010.
- [19] Kasturi R, Fernandez R, Amlani M L, et al. Map Data Processing in Geographic Information System. Computer, 1989, 22(12):10-21.
- [20] Kasturi R, Alemany J. Information Extractiong from Images of Paper-based Maps. IEEE Trans Software Engineering, 1988, 4(5):671-675.
- [21] 夏彦刚. 图像的矢量化及保边缘图像失量化方法[学位论文]. 城都: 四川大学, 2005.
- [22] W3C. Scalable Vector Graphics (SVG) 1.0 Specification. W3C Recommendation 04 September 2001.
- [23] POLynomials, POLar Forms, and InterPOLation. Mathematical methods in computer aided geometric design II. Academic Press Professional, Inc. September 1992.
- [24] Gerald E. Farin, Dianne Hansford. The essentials of CAGD. Natic MA:A K Peters Ltd, ISBN 1-56881-123-3.
- [25] Paul Bourke. Adaptive Subdivision of Bezier Curves. http://www.antigrain.com/research/adaptive_bezier/index.html.
- [26] 王巍. 人脸面部特征定位与人脸识别方法的研究[学位论文]. 北京: 北京工业大学, 2003.
- [27] T F Cootes, G J Edwards, C J Taylar. Active Shape Models. Proc. European Conf. Cmputer Vision, 1998:p484-498.
- [28] Timonthy F. Coontes, Gareth J. Edwards, Christopher J. Taylor. Active Appearance Models. IEEE Transactions on Pattren Analysis and Machine Intelligence. Jun 2001: 681-685.
- [29] T F Cootes, C J Taylar, D H Cooper, J Haslam. Training Models of Shape from Sets of Examples. in Proc. British Machine Vision Conference, Springer-Verlag, 1992:p9-18.
- [30] T F Cootes, C J Taylar. Active Shape Models—"smart snakes". Proceeding of the British Vision Conference, Springer-Verlag, September, 1992:p266-275.

- [31] T F Cootes, C J Taylar, A Lanitis. Multi-Resolution Search with Active Shape Models. Proc. ICPR 1994.
- [32] T F Cootes, C J Taylar, D H Cooper, J Graham. Active Shape Models—their training and application. Computer vision and image understanding,61(1). 1995:p38-59.
- [33] K N Walker, T F Cootes, and C J Taylor. Automatically building appearance models from image sequences using salient features. Image Vis. Comput., vol. 20, no. 5, pp. 435–440, 2002.
- [34] D Cristinacce, T F Cootes. A comparison of shape constrained facial feature detectors. in Proc. Int. Conf. Autom. Face Gesture Recog.,2004, pp. 375–380.
- [35] I Matthews, S Baker. Active appearance models revisited. International Journal of Computer Vision, 2004,60(2), p135–164.
- [36] J Xiao, S Baker, I Matthews, and T. Kanade. Real-time combined 2D+3D active appearance models. in Proc. IEEE Conf. Comput. Vis.Pattern Recog., 2004, vol. 2, p 535–542.
- [37] X Liu. Generic face alignment using boosted appearance model. in Proc. IEEE Conf. Comput. Vis. Pattern Recog., 2007, p 1–8
- [38] P. Viola, M. J. Jones. Robust real-time face detection. International Journal of Computer Vision, vol. 57, no. 2, pp.137–154, 2004.
- [39] R. Gross, I. Matthews, and S. Baker. Constructing and fitting active appearance models with occlusion. in Proc. IEEE Conf. Comput. Vis. Pattern Recog, Workshops, 2004, vol. 5, p. 72.
- [40] T. F. Cootes, G. V. Wheeler, K. N.Walker, and C. J. Taylor. View-based active appearance models. in Proc. IEEE Int. Conf. Autom. FaceGesture Recog.2000, pp. 227–232.
- [41] E Malis, R Cipolla. Multi-view constraints between collineations: application to self-calibration from unknown planar structures. European Conference on Computer Vision, 2000.
- [42] E Malis, R Cipolla. Self-calibration of zooming cameras observing an unknown planar structure. 15th International Conference on Pattern Recognition(ICPR'00), September 2000:p85-88.
- [43] 谭晓波, 摄像机标定及相关技术研究, 国防科技大学, 2004.

- [44] 马颂德, 张正友. 计算机视觉——计算理论与算法基础. 北京:科学出版社, 1997.
- [45] H Bakstein. A complete DLT-based camera calibration with a virtual 3D calibration object. CPRW'2000.
- [46] A Criminisi, I Reid, A Zisserman. A plane measuring device. The 8th British Machine Vision Conference(BMVC'97), September 1997.
- [47] R I Hartley. In defence of the 8-point algorithm. 5th International Conference on Computer Vision. June 1995:p1064-1070.
- [48] 鲍虎军, 金小刚, 彭群生. 计算机动画的算法基础[M]. 杭州:浙江大学出版社, 2000.310-367.
- [49] Hsu W M, Hughes J F, Kaufman H. Direct manipulation of Free-form Deformation. Computer graphics, 1992, 26(2):p177-184.
- [50] Mobius A F. Der Barycentrische Calcul. Johann Ambrosius Barth, Leipzig, 1827.
- [51] Michael S.Floater. Mean Value Coordinates. Computer Aided Geometric Design. March 2003: 19-27.
- [52] Kai Hormann, Michael S. Floater. Mean Value Coordinates for Arbitrary Planar Polygons. ACM Transactions on Graphics. 2006: 1424-1441.
- [53] FLOATER, M. S., HORMANN, K., AND K'OS, G. 2006. A general construction of barycentric coordinates over convex polygons. Adv. Comp. Math. 24, 1-4, 311-331.
- [54] FLOATER, M.S., K'OS, G., AND REIMERS, M. 2005. Mean value coordinates in 3D. Comput. Aided Geom. Des. 22, 7, 623-631.
- [55] Tao Ju, Scott Schaefer, Joe Warren. Mean Value Coordinates for Closed Triangular Meshes. ACM Transaction on Graphics. July 2005:p561-566.
- [56] Pushkar Joshi, Mark Meyer, Tony DeRose. Harmonic Coordinates for Character Articulation. ACM Transaction on Graphics. July 2007, article 71:1-9.
- [57] Yaron Lipman, David Levin. Derivation and analysis of Green Coordinates.
- [58] Yaron Lipman, David Levin, Daniel Cohen-Or. Green Coordinates. ACM Transaction on Graphics. August 2008: 78:1-10.
- [59] Ofir Weber, Hybrid Methods for Interactive Shape Manipulation. Research Thesis. Israel Institute of Technology. July 2010.

- [60] Wachspress E. A Rational Finite Element Basis. Academic Press, New York, 1975.
- [61] Meyer M, Lee H, Bar A, Desbrun M. Generalized Barycentric Coordinates for Irregular Polygons. *Journal of Graphics Tools* 7, 1, 2002:p13–22.
- [62] Scott Schaefer, Travis McPhail, Joe Warren. Image Deformation Using Moving Least Squares. In *Proceedings of ACM SIGGRAPH*, Boston, 2006:p533-540.
- [63] Josiah Manson, Scott Schaefer. Moving Least Squares Coordinates. *Eurographics Symposium on Geometry Processing*, 2010:p1517-1524.
- [64] Levin D. The Approximation Power of Moving Least Squares. *Mathematics of Computation*, 1998, 67(224):p1517-1531.
- [65] Alexa M, Cohen-Or D, Levin D. As rigid as possible shape interpolation. *ACM Press/Addison-Wesley Publishing Co.* 2000:p157-164.
- [66] Horn B. Closed_form solution of absolute orientation using quaternions. *Journal of the Optical Society of America*, April 1987:p29-42.
- [67] M. M. Nordstrøm, M. Larsen, J. Sierakowski, and M. B. Stegmann. The IMM Face Database - An Annotated Dataset of 240 Face Images. Technical reportt, Informatics and Mathematical Modelling, Technical University of Denmark, May 2004.

致 谢

首先感谢我的导师屠长河教授。屠老师在我研究生阶段的学习和研究工作都给予了我极大的指导和帮助。感谢老师为我提供了良好的学习科研环境，无私的传授自己的知识和敬仰，悉心指导我对课题的研究。屠老师专业知识渊博、学术态度严谨、平易近人、言传身教，是我将来工作、学习的榜样。正是因为屠老师严格的要求和悉心的指导，才使我的工作和学习得以顺利开展，老师为我指点迷津，提出宝贵意见，鼓励我支持我，从而才有了研究目标的逐步实现以及最终本文的顺利完成，在此我对屠老师表示由衷的感谢。

同时，感谢人机交互与虚拟现实实验中心的全体教职工，在工作和生活中他们热情的帮助和大力的支持帮助我克服困难。中心的学术讨论班以及专家教授的讲座提供给了我了解领域研究动态情况以及学术交流的机会，拓展了知识并提升了能力，使我得到许多启迪，最终顺利地完课题，祝他们工作顺利、生活美满。

同样感谢所有我们组的同学们以及帮助我的其他师兄师姐们，他们刻苦的钻研精神激励我不断前进，在生活和科研中他们也给了我许多的帮助，与他们共同学习交流的时光是美好而珍贵的，虽然短暂，却是宝贵的回忆，积极影响我以后的生活。祝他们学业有成、一切顺利。

感谢我的父母和家人，他们一如既往的关心我支持我，快乐的时候为我高兴，失意的时候努力安慰我，使我更加坚强。我衷心的感谢，我会更加努力，不辜负他们的期望，同时也祝愿他们幸福平安。

感谢山东大学计算机科学与技术学院对我的培养，以及各位老师的辛勤工作。我衷心的祝愿所有帮助过我的人健康、平安、快乐！

攻读学位期间发表的学术论文

Mapping the Peking Opera Facial Makeup onto a Human Face in Video Sequences,
Asia-Pacific Signal and Information Processing Association 2011, 西安, 2011 年 10
月 20 日, 第一作者

学位论文评阅及答辩情况表

论文评阅人	姓 名	专业技术 职 务	是否博导 (硕导)	所 在 单 位	总体评价 ※	
	匿名评阅人 1	-	-	-	B	
	匿名评阅人 2	-	-	-	B	
答辩委员会成员	姓 名	专业技术 职 务	是否博导 (硕导)	所 在 单 位		
	主席 冯志全	教授	是	济南大学		
	委员	崔立真	副教授	是	山东大学	
		张瑞华	副教授	是	山东大学	
		赵合计	副教授	是	山东大学	
答辩委员会对论文的 总体评价※		A	答辩秘书	孙建东	答辩日期	
备注		2012.5.23				

※优秀为“A”；良好为“B”；合格为“C”；不合格为“D”。