

# 致 谢

随着本篇论文的完成，我在中国科学技术大学的五年学习生活也即将结束。在这里我要感谢董兰芳老师，她对我一年多来的指导，使我的专业技能和学习能力都有了很大的提高。还要感谢图形图像实验室给我提供的良好的研究环境和氛围，并感谢各位同学在我进入实验室里来对我的关心和帮助。

此外，还要感谢 9911 的多位同学为本文提供了实验素材，他们是：

朱玉华	朱丽莎	刘潇	肖玉	叶荣
许焰	章志燮	李清	范斌	张波
杜野	颜俊	陈许佳	王向晖	

感谢 523 的室友章志燮、李清、范斌陪我走过一年来的快乐时光，还要感谢朱玉华这三年多来对我的支持和帮助。最后我要感谢我的父母，感谢他们一直以来对我的关心与支持。

谢谢所有关心和帮助过我的人！

## 目 录

目 录.....	I
摘 要.....	II
Abstract.....	III
第一章 绪 论.....	1
1.1. 研究意义与背景.....	1
1.2. 国内外研究状况.....	2
1.3. 本论文的主要研究内容及组织形式.....	3
第二章 人脸特征点定位方法的概述.....	5
2.1. 基于灰度及变化信息.....	5
2.2. 基于活动轮廓线 (Active Contour Models) 模型.....	6
2.3. 基于神经网络.....	8
2.4. 基于可变形模板.....	10
2.5. 基于主动形状模型.....	10
2.6. 基于主动表观模型.....	14
第三章 对 AAM 模型的改进—基于肤色信息的 AAM 模型.....	18
3.1. 肤色模型.....	18
3.1.1. 肤色概述.....	18
3.1.2. 高斯模型.....	18
3.2. 基于肤色信息的 AAM 模型.....	20
3.2.1. 基于肤色模型的人眼定位方法.....	20
3.2.2. 肤色模型在 AAM 搜索中的应用.....	24
第四章 一个人脸肖像画生成系统.....	25
第四章 一个人脸肖像画生成系统.....	26
4.1. 系统结构.....	26
4.2. 人脸面部特征的提取.....	27
4.2.1. AAM 模型的建立.....	27
4.2.2. 根据特征点生成面部器官的肖像画效果.....	28
4.3. 头发的提取.....	28
4.4. 肖像画的合成及显示.....	30
4.5. 系统总结与分析.....	30
第五章 结论与展望.....	32
5.1. 论文的主要工作和结论.....	32
5.2. 尚需进一步研究的问题.....	32
参考文献.....	33

## 摘 要

人脸特征点定位在人脸识别、三维虚拟人脸合成、人脸表情分析与合成等方面具有非常重要的作用，是当前计算机视觉、计算机图形学、模式识别等领域的研究热点。目前，已经有很多的算法和模型提出，但是由于脸部的复杂性和人脸图像成像时环境的影响，至今仍然没有一个一般性的算法或模型能够解决这一研究难题。

本书对目前常用的人脸特征点定位算法和模型进行了总结，并在前人的基础上提出一个基于肤色模型的对 AAM 模型的改进方法。由于 AAM 模型在信息的处理上忽略了颜色信息的作用，而肤色信息在人脸检测时往往起着非常大的作用，利用肤色信息将大大减小 AAM 搜索的搜索范围，从而减小搜索时间。本文提出的方法正是基于此原理。在此方法的基础上，本文还提出并实现了一套完整的人物肖像画自动生成系统。

**关键词** 人脸特征点定位 人脸检测 AAM 模型 肤色模型 肖像画生成

## ABSTRACT

Human facial features extraction has become one of the important issues in lots of significant application, such as face recognition, virtual 3D face modeling and synthesis, facial expression analysis and synthesis, etc. It is also the focus of research in computer vision, computer graphics and pattern recognition. However, due to the structural complexity of human face and the affection of lights to the images, it still remains an unsolved problem to extract facial features exactly in all pictures and videos.

In this paper, we discuss the most commonly used models and approaches in features extraction, based on which we also bring forward a skin-color based approach to improve the Active Appearance Model in facial features extraction. Active Appearance Model works well in interpreting face images, but it still does not use all the information available, such as the skin color information, which has been widely used in face detection and preprocessing of digital images. Making full use of the chromatic information of skin will greatly reduce the window of AAM searching, as well as help to get an optional initial pose and scale of AAM model, thus reduce the time of searching. This is just the origin of our approach in this paper. Based on our novel approach to extract facial feature from color images, we bring out and implement an automated portrait generation system in this paper.

**Keywords** Facial features extraction, Face detection, Active Appearance Model, Skin-color model, Portrait generation

# 第一章 绪 论

## 1.1. 研究意义与背景

人脸作为人类最熟悉又最富多变性的一个组成部分，其复杂的结构能提供非常丰富的信息，这使得人脸特征信息在计算机中的表示形式以及这些信息的处理方式成为计算机视觉、图像处理和分析、模式识别、图像编码，计算机图形等领域的重要研究内容之一。

所谓的人脸特征定位，就是通过计算机在一幅人脸图像中自动地定位出人脸各个器官的准确位置，其中包括眼睛、鼻子、嘴巴、下巴、耳朵以及人脸外轮廓等所有需要提取特征点的位置。它为人脸图像的处理与分析提供了重要的几何信息，特征定位的准确与否直接关系到后续应用的可靠性。该方向的研究在人脸识别、基于模型的图像编码、具有真实感的三维虚拟人脸建模、人脸变形、人脸表情分析及合成、口形分析等方面有着重要的应用。其中，人脸识别是它的一个重要应用，在人脸识别算法中，一个重要的工作就是人脸图像尺寸的归一化，而这一归一化的实现都是依赖于面部特征定位的结果。另外，一些算法中需要各个器官的相对位置及形状或对面部局部区域的特征加以分析，需要在面部特征定位基础之上进行。具有真实感的三维虚拟人脸建模是它的另一个重要应用，其实质是要从人脸图像中重建出人脸的三维模型，这其中非常关键的一步就是确实各个角度中人脸器官的特征点位置，也就是要精确地提取出人脸的几何特征信息。多媒体娱乐的发展促使了人物漫画生成系统和肖像画生成系统的产生，其实质是从人脸图像是提取出人物的特征并以其它形式（比如漫画、线条画等）形象、生动甚至夸张地表现出来，而人脸特征点的准确定位正为这些表现方式提供最基础的数据。总之，人脸图像的特征定位有着广泛的作用，是许多其它的人脸图像研究工作的前提，直接关系到其它研究工作能否顺利地进行。

虽然人类可以从一幅人脸图像中很轻松地分辨出面部特征点的准确位置，但对于计算机来讲却并非一件易事。由于人脸有复杂的三维表面结构，因而对于其形成的二维图像，其变化也是非常之大的，尤其对于不同的人脸姿态、表情，不同的光照条件，得到的二维图像的差别是非常明显。因此，准确的面部特征定位是一个非常具有挑战性的课题。目前，国内外很多研究人员提出了各种方法，在

不同的领域取得了一定的成果。但是对于寻找一种能够普遍适用于各种复杂情况的准确率很高的检测算法，还有很大的探索空间。

## 1.2. 国内外研究状况

面部特征定位的研究始于基于几何特征的人脸识别方法的研究。早在上世纪六十年代末, Kanada<sup>[39]</sup>就提出了基于几何特征的人脸识别方法。其中采用的方法就是基于灰度变化信息来定位面部特征点的方法。该方法主要利用人脸不同区域灰度分布的不同特性来提取特征点的位置。该方法有与直接利用灰度的变化来分析, 因而对于光照的变化, 其鲁棒性受到限制。

Kass<sup>[4]</sup>等人在 1987 年首先提出了称为 Snake 的主动轮廓线模型 (Active Contour Model), Snake 是能量极小化的一种模型, 内力约束它的形状, 外力引导它的行为, 图像力将其拖向显著的图像特征。Snake 模型的引人之处在于它对范围广泛的一系列视觉问题给出了统一的解决方法。近年来, 它已经被越来越多的研究者成功地应用于计算机视觉的诸多领域。Won-Sook Lee<sup>[5]</sup>等 1997 年提出了一种由正面和侧面人脸图像的结构化 snake 模型来进行人脸面部特征定位。

基于可变形模板的特征提取首先是由 A. L. Yuille<sup>[6,7]</sup>等人在 1992 年提出, 用来提取眼睛和嘴巴的特征位置。针对人脸器官的复杂性, 有些区域的灰度对比度比较低, 因而无法直接检测到边缘, 即使获得了较明显的边缘, 也很难从中自动地获取精度很高的特征点位置。为此, 用弹性可变形模板进行特征提取就显示其优越性, 可变形模板用一组根据特征形状的先验知识设计的参数所定义, 这些参数是可调的, 为了定义出最符合特征形状的参数, 需要利用图像的边缘、峰值、谷值和先验知识设计合适的能量函数。参数向能量函数减小的方向做调整, 当能量函数达到最小时, 这组参数对应的模板就最符合所关心的特征形状。

Paul Debevec<sup>[8]</sup>等于 1992 年提出了利用神经网络对人脸照片中的眼睛、鼻子和嘴巴进行定位的方法。在该方法中, 对每一个器官都构造一个神经网络, 以人脸图像为网络的输入, 得到该器官在图像中的位置。

L. Wiskot<sup>[9]</sup>等人在 1997 年提出了利用 Gabor:小波变换系数进行人脸识别的弹性图匹配方法, 在该方法中利用了一组不同频率和相位下得到的 Gabor 小波变换系数作为特征, 来进行面部特征点的精细定位, 进而进行识别。V. Krüger<sup>[10]</sup>利用 Gabor 小波进行人脸图像的表达与重建。根据重建系数的相似性来进行人脸特征定位。

另一种人脸特征定位的方法是基于主动形状模型和主动表观模型的方法。主动形状模型 (ASM) 是一种基于统计模型的方法, 是由 T. F. Cootes<sup>[2, 10, 11, 12]</sup> 等人在 1995 年提出, 该方法对训练集中大量所描述的形状实例进行统计, 建立起反映目标形状变化规律的形状统计模型和反映灰度分布规律的局部灰度模型. 在搜索过程中, 首先利用训练得到的局部灰度模型进行搜索, 之后利用形状模型来对搜索到的形状进行近似表达, 同时对其合理性进行判断, 对不合理的形状进行调整以保证形状在统计意义上的合理性。通过循环迭代, 得到理想的匹配结果。在主动形状模型基础之上, T. F. Cootes<sup>[1, 13, 14]</sup> 等人在 1998 年中提出了主动表观模型 (AAM)。在此方法中, 不仅建立了反映形状变化的形状统计模型, 同时建立了反映全局纹理变化的全局纹理模型, 以充分利用全局的纹理信息。并对形状模型及纹理模型结合起来建立表观模型。得到的表观模型去除了形状和纹理之间的相关性, 可以准确地生成形状及纹理变化的目标图像。在 AAM 搜索过程中, 利用训练得到的图像灰度差值与表观参数变化的线性关系, 不断变化表观模型参数以及二维几何位置和尺度等参数, 直至得到理想的匹配效果。

ASM 和 AAM 方法可以通过对某一类特定图像进行建模, 来进行特征的提取, 具有较好的通用性与灵活性。本文所提出的人脸特征定位算法正是对 AAM 方法的一种改进。

### 1.3. 本论文的主要研究内容及组织形式

本文主要对人脸面部特征定位方法进行了研究, 并根据人脸肤色的特点提出了一种改进的 AAM 模型来进行人脸特征定位, 同时还介绍了面部特征定位的一个具体应用——人物肖像画自动生成系统的原理和实现方法。本文组织如下:

第一章是绪论, 主要阐述了人脸特征定位的研究背景、重要意义和国内外的研究动态。在第二章中介绍了几种目前比较常用的几种特征提取方法, 即基于灰度及变化信息的特征提取、基于 Snake 的特征提取、基于 ANN 的特征提取、基于可变形模板的特征提取和基于 ASM、AAM 的特征提取。

本文的第三章详细介绍了肤色模板, 并根据人脸肤色的特点提出了一种改进的 AAM 模型来进行人脸特征定位。这种改进的方法通过充分利用图像的颜色信息, 大大提高了 AAM 的搜索速率。

在第四章, 利用改进的 AAM 模型, 提出了一个人物肖像画自动生成系统, 并详细介绍了该系统的结构, 以及各模块所采用的实现方法, 其中包括人脸各器官

的提取和表示、头发的提取。文中还对该系统进行了分析和讨论。

最后在第五章中对，对本文所进行的人脸面部特征定位领域所做的研究进行了总结，对研究的方法给出了相应的结论，阐述了这些方法中现存的一些问题，并指出了今后应努力的方向。



## 第二章 人脸特征点定位方法的概述

### 2.1. 基于灰度及变化信息

特征提取方法中最简单、最直观的就是通过人脸面部灰度及其变化的信息来获取特征信息。由于人脸各器官具有一些固定的灰度分布特点，比如眼睛往往是脸部灰度最深（最黑）的部分、双眼灰度分布的对称性、人的五官位置和大小存在一定的比例关系等等，所以很容易利用灰度信息来做特征的粗略定位。这类方法主要有如下三种形式：

1)、灰度的水平或垂直投影：对人脸图像来说，通常用这种方法对眼睛、嘴巴、眉毛、鼻子进行定位。因为眼睛往往是脸部区域灰度最深的位置（最黑的地方），所以可以对脸部区域同时进行灰度的水平投影和垂直投影，这两个投影曲线的交叉点就是两眼睛在脸部的位置。然后用已经知道的眼睛位置、脸部器官的先验知识和两条水平、垂直投影曲线，就可以定位出其它主要器官的位置。这种方法比较简单，对正面图像的效果比较好，但是它的结果不会很精确，而且抗干扰性比较差，稳定性不高，仅适用主要器官的粗略定位。

2)、基于边缘的特征提取：边缘提取是图像处理中一个经典的方法，它利用图像灰度变化的信息来反映图像特征。如果一个像素落在某个物体的边界上，那么它的领域将成为灰度变化带，这种变化最有用的信息是灰度的变化率和方向，它们分别用梯度的幅值和方向来表示。边缘检测方法通过检测每个像素的领域并对灰度变化率进行量化，通常也包括方向，从而提取出边缘。基于边缘的提取往往不是把提取出的边缘作为最终结果，而是在其基础上进行后处理，例如利用先验知识和规律对边缘进行连接和曲线拟合，从而形成一条为抽取物体所适用的边界。一般的拟合技术通常用通过抛物线、椭圆及高斯曲面来拟合物体。在对直线和圆进行拟合时利用 Hough 变换能达到比较好的效果，在人脸处理中也常用 Hough 变换对眼球进行检测，从而完成对眼睛的定位。常用的边缘检测算子有：梯度边缘算子、Laplace 算子、Roberts 算子、Sobel 算子、Prewitt 算子、Kirsch 算子、Cany 算子等。

3)、基于曲率的特征提取：对于物体的外形，其特征轮廓线的曲率特征比较

能够反映该物体的形状特征，所以在提取汉字特征、指纹特征和掌纹特征进行识别时，往往会提取该物体的曲率特征作为分析和识别的依据，这在汉字识别系统和指纹识别系统中已获得较大成功。在人脸图像分析时，人脸的外轮廓也能反映不同人的特征，所以偶尔也会用来做人脸识别的特征，同时因为人脸外轮廓中曲率的极值点往往也是人脸其它主要器官点的位置，所以也可以用来标定其它器官的位置。

## 2.2. 基于活动轮廓线 (Active Contour Models) 模型

Kass<sup>[4]</sup>等人在 1987 年提出了称为 Snake 的主动轮廓线模型 (Active Contour Model)。Snake 定义轮廓能量，通过极小化轮廓能量来达到轮廓曲线与图像中物体轮廓的吻合。轮廓能量的极小化是通过各种力共同作用的结果：内力约束它的形状，外力引导它的行为，图像力将其拖向显著的图像特征。Snake 模型的引人之处在于它对范围广泛的一系列视觉问题给出了统一的解决方法。

在 Snake 模型中，轮廓曲线用参数式来表示：

$$v[s] = [x(s), y(s)]^T, s \in [0,1],$$

比如椭圆曲线可以表示为：

$$v[s] = [\alpha \cos(2\pi \cdot s), \beta \sin(2\pi \cdot s)]^T, s \in [0,1]。$$

Kass 还定义了轮廓曲线的能量函数：

$$\begin{aligned} E_{snake} &= \int_0^1 E_{snake}(v(s)) ds \\ &= \int_0^1 [E_{int}(v(s)) ds + E_{image}(v(s)) ds + E_{con}(v(s)) ds] \end{aligned}$$

其中， $E_{int}(v(s)) ds$  表示轮廓曲线的内部能量， $E_{image}(v(s)) ds$  表示图像作用力产生的能量， $E_{con}(v(s)) ds$  表示外部限制用力（自定义）产生的能量。

➤ 内部能量反映的是轮廓曲线的平滑性，能量越低表示曲线越平滑：



图 1 轮廓曲线的能量

因此可以定义：

$$E_{\text{int}}(v(s))ds = \alpha(s) \left| \frac{\partial v}{\partial s} \right|^2 + \beta(s) \left| \frac{\partial^2 v}{\partial s^2} \right|^2。$$

一阶项在轮廓线存在裂口时，将有较大的值，二阶项当轮廓线曲率变化较大时有较大值，轮廓线上的任意一点的  $\alpha$  和  $\beta$  值决定了该点的延伸和弯曲程度，如果某点的  $\alpha$  为 0，则曲线在该点不连续，如果  $\beta$  为 0，则表示曲线在该点存在一个角点。

- 图像作用力能量  $E_{\text{image}}(v(s))ds$  反映的是图像中物体对轮廓曲线的吸引力，比如可以将其定义成直线、边缘和界限能量对轮廓曲线的吸引力的结果：

$$E_{\text{image}}(v(s))ds = \omega_{\text{line}} E_{\text{line}} + \omega_{\text{edge}} E_{\text{edge}} + \omega_{\text{term}} E_{\text{term}}$$

$E_{\text{line}}$  可以简单定义成  $E_{\text{line}} = f(x, y)$ ,  $f(x, y)$  表示点  $(x, y)$  像素的灰度,  $\omega_{\text{line}}$  的正负决定了轮廓曲线被吸引到深色或浅色的线;  $E_{\text{term}}$  是用高斯函数平滑过的图像中各级轮廓线的曲率;  $E_{\text{edge}}$  表示边缘对轮廓曲线的吸引:



图 2  $E_{\text{edge}}$  的物理意义

因此可以定义:

$$E_{\text{edge}} = -|\nabla f(x, y)|^2。$$

由轮廓能量的定义可以看出，轮廓能量越小，轮廓曲线与图像中的物体轮廓就越吻合。在具体实现的时候，可采用如下的离散化方法:

$$\begin{aligned} v(s) &= v(s_i), i = 1, 2, \dots, N; \\ v'(s_i) &= \bar{d} - |v_i - v_{i-1}|^2, \bar{d} \text{ 为各点间的平均距离;} \\ v''(s_i) &= |v_{i-1} - 2v_i + v_{i+1}|^2 \end{aligned}$$

Kass 等人采用各种方法来最小化轮廓的能量，其中最基本的方法是梯度下降法。首先对轮廓线进行适当的初始化，然后 snake 在图像上找到相近的能量最小值。这种方法有一个重大的缺陷，那就是它只能使用局部的信息，这样使得优化的结果非常依赖于初始化的位置，局部噪音也会对结果产生较大的影响。

为了解决 snake 存在的种种问题，研究人员提出了许多改进和延伸。在能量

函数的设计方面, Cohen<sup>[15]</sup>等引入了气球力(balloon force), Ronfard<sup>[16]</sup>等引进了基于背景和目标区域的统计模型的目标函数, Fua<sup>[17]</sup>等在 snake 模型中设置了吸引子和切线条件。在最优化过程方面, Amini<sup>[18]</sup>等使用动态编程去最小化能量函数, Gaselles<sup>[19]</sup>和 Malladi<sup>[20]</sup>分别提出用层次集(level set)方法解决 snake 的问题, Christensen<sup>[21]</sup>等将膨胀黏性液体的变形动力学中的偏微分议程应用于 snake 中, 指导模型的演化。此外, 对 snake 本身的改进方面, Menet<sup>[22]</sup>等提出了 B-snake, H. H. S. Ip 和 D. Shen<sup>[23]</sup>提出了 AI-Snake, 等等。

综合来看, Snake 模型具有一些经典方法所无法比拟的优点, 但也同时有其自身的缺点(见表 1):

优点	缺点
经适当地初始化后, 能够自主地收敛于能量极小值状态	不太稳定, 可能会收敛到局部极值点, 甚至发散
图像数据、初始估计, 目标轮廓及基于知识的约束统一于特征提取过程	对初始位置比较敏感, 需要依赖其它机制将曲线放在感兴趣的图像特征附近
适用于任意运送的物体的特征提取	

表 1 Snake 模型的优点和缺点

因此, 在人脸图像处理中 Snake 经常用于提取人脸外轮廓等形状不是很规则的特征(如图 3 所示):

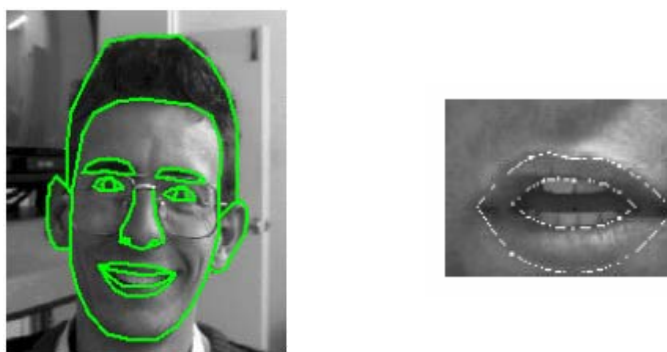


图 3 使用 Snake 提取人脸轮廓的例子

### 2.3. 基于神经网络

由于神经网络的自学习功能使得它能够较好的提取人脸的整体特征, 所以在

利用人脸的整体特征进行人脸识别和人脸检测方面，神经网络已经取得了较好的结果。在特征提取方面，Paul Debevec<sup>[24]</sup>利用神经网络对特征提取进行尝试，他用神经网络对人脸照片中的眼睛、鼻子和嘴巴进行定位，在定位时对于人脸图像中的每一个器官，都必须构造一个神经网络。对个神经网络，输入一幅人脸图像，得到该器官在人脸图像中定位出的位置。为了突出各个器官特征的差别，在训练和检测时，输入图像并不是原始图像，而是进行变换后的 log-polar 映射图。同时为加强器官间的区别和减少神经网络训练时间，则在训练时，对人脸的 log-polar 映射图进行二次抽样，抽样成  $8 \times 8 = 64$  的训练子图。

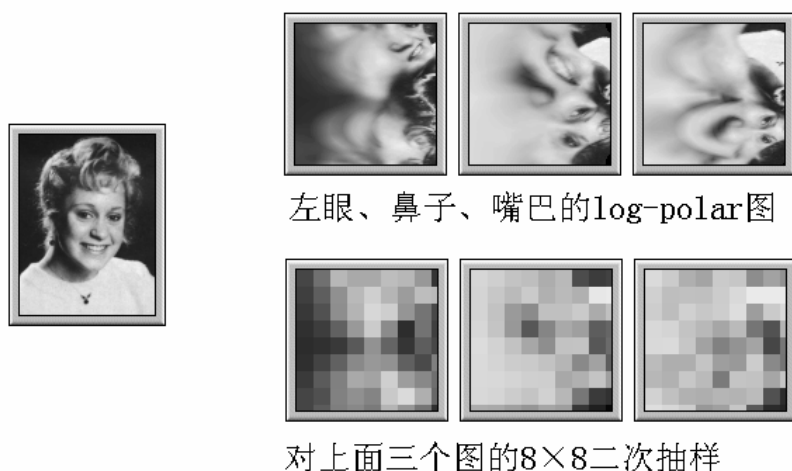


图 4 原图像、log-polar 图及二次抽样图

输入一幅人脸图像，针对不同器官的神经网络训练器，要映射成不同的训练子图。对每一个器官，输入大量的训练子图进行训练，得到稳定的神经元系数后，输入新图像就知道该器官特征在训练子图中的位置，反向映射回人脸图像，就确定了最终的人脸器官特征在图像中的位置。

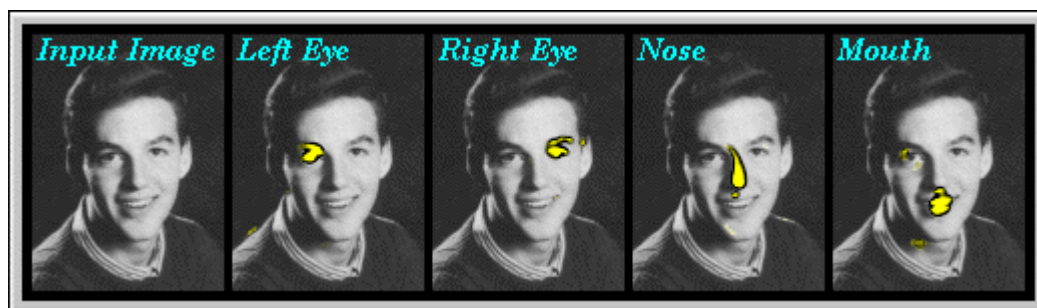


图 5 利用神经网络的脸部特征定位

分析神经网络方法的原理，可以看出其具有以下缺点：

- ✓ 自适应性不强不能适应人脸图像的旋转变化；

- ✓ 无法对人的细节特征进行提取，仅适用于主要器官位置的粗略定位。

## 2.4. 基于可变形模板

A. L. Yuille<sup>[25]</sup>等人在 1992 年提出可变形模板的方法，用来提取眼睛和嘴巴的特征位置。由于人脸器官的复杂性，加上有些地方的灰度对比度比较低甚至连边缘都检测不到，即使获得较明显的边缘，也很难从中自动地获取精度很高的特征点位置。而弹性可变形模板正好避免了这个问题。

可变形模板由一组根据特征形状的先验知识设计的可调参数所定义，为定义出符合特征形状的参数，需要利用图像的边缘、峰值、谷值和先验知识设计合适的能量函数。人脸特征定位的过程就是不断调整参数，使能量函数不断减少，当达到最小值时，这组参数对应的模板就最符合所关心的特征形状，而特征点信息可以直接从这模板中获得。

A. L. Yuille 等人首先提出的眼睛模板是由两条抛物线和一个圆构成，分别对应人眼的内眼睑和眼球，嘴巴模板由多条抛物线组成，分别对应于嘴巴的上、下嘴唇。（如图 6 所示）

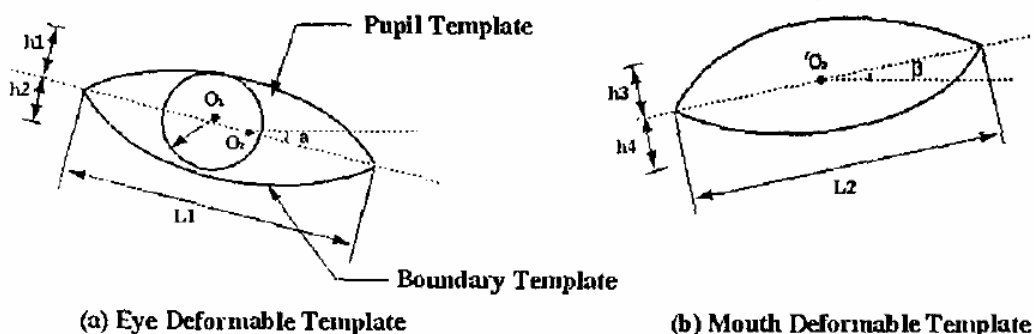


图 6 眼睛和嘴唇的可变形模板

可变形模板很重要的一个方面就是能量函数的选择，能量函数选择的好坏直接影响到可变形模板收敛的快慢和最终的匹配结果。

## 2.5. 基于主动形状模型

主动形状模型(Active Shape Models, 简称 ASM)最初是由 T. F. Cootes<sup>[2, 10, 11, 12]</sup>于 1995 年提出的，它在思想上类似于主动轮廓模型(Active Contour Models)，即定义一个能量函数，通过调整模型参数使能量函数值最小化。与 ACM 不同的是，

它是一种基于大量训练集数据的统计模型，它的动态形态能产生在训练集中出现的各种各样的物体形态，以满足在物体模型与图像匹配时的需求，同时，它又能根据训练数据对参数的调节加以限制，从而将形状的改变限制在一个合理的范围之内。由于 ASM 在特征提取方面能达到很好的效果，因而在医学图像处理和计算机礼堂的人脸图像理解方面有很大的应用和发展。

ASM 是一种基于先验知识的统计模型，其模型的建立需要一组标有特征点图像（训练集），然后对每个图像中物体的形状和局部灰度建模。特征点定位其实就是一个搜索的过程，通过不断调节形状和姿态参数从而使能量达到最小，从而使形状达到最优化。下面将就人脸特征点的定位来说明 ASM 的原理：

- ✓ ASM 中形状的表达形式：在 ASM 中，物体的形状由物体的  $n$  个二维特征点  $\{x_i, y_i\}, i \in 1, \dots, n$  来描述，并且用一个  $2n$  维的矢量来表示：

$$X = [x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_n]$$

- ✓ 训练集：训练集是由一组标有特征点的图像组成（见图 7）。ASM 是基于对训练集的统计分析的，因此训练集对 ASM 模型的好坏起着关键的作用。

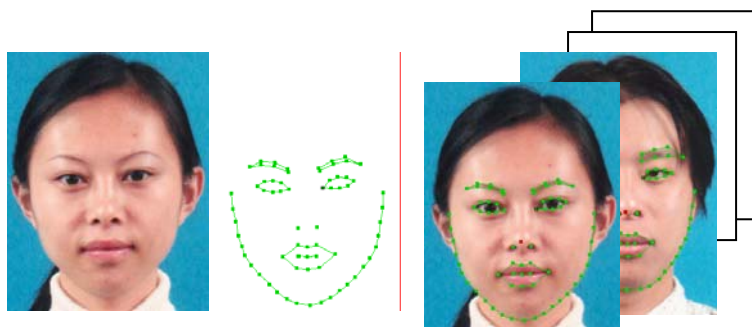


图 7 训练图像、特征点标记和训练集

- ✓ ASM 模型的建立：

ASM 模型的建立其实也就是对训练集进行统计分析的过程，由于训练集中的物体可能在大小，方向等方面存在差异，因此在进行分析之前必须先将训练集中的物体归一化（对齐），这包括缩放和旋转，然后对对齐后的形状进行 PCA 分析，即可得到 ASM 模型：

$$X = \bar{X} + Pb$$

其中， $\bar{X}$  是物体形状的平均值， $P$  是特征变换矩阵，它表示的是物体形状变化的几种主要模式， $b$  是参数向量。该式子的意义就是：将任一物体的形状看成是平均形状与各种形状变化模式合成得来的结果。通过



改变参数向量  $b$ , 就可以得到各种不同的形状。这样, 在 ASM 模型的  $\bar{X}$  和  $P$  已知的前提下, 只要知道  $b$  即可恢复出一个人脸的形状。 $P$  是通过协方差的 PCA 分析获得的, 具体步骤如下:

1)、计算训练集中形状的均值和协方差:

$$\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i$$

$$S = \frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X})(X_i - \bar{X})^T$$

$N$  是训练集中图像的个数,  $X_i$  是第  $i$  个训练图像的形状向量。

2)、求  $S$  的特征值  $(\lambda_1, \lambda_2, \dots, \lambda_m)$ , 其中  $\lambda_i \geq \lambda_{i+1}$ , 并且  $\lambda_i, i=1, 2, \dots, m$ , 取前  $t$  个特征值, 使得

$$\sum_{i=1}^t \lambda_i / \sum_{j=1}^n \lambda_j \geq 98\%$$

3)、取这  $t$  个特征值相应的特征向量 (维数是  $2n \times 1$ ), 并将其正定化成  $p_1, p_2, \dots, p_t$ , 并组成  $P$ :

$$P = (p_1, p_2, \dots, p_t)$$

因为与较大特征值  $\lambda_i$  对应的特征向量  $p_i$  反映了比较重要的形状, 所以这  $t$  个特征向量足以表示大多数形状向量  $X$ 。

- ✓ 参数向量  $b$  的分析: 由参数的概率分布模型, 可以得出参数向量第  $i$  个参数  $b_i$  统计意义上的合理变化范围是  $\pm 3\sqrt{\lambda_i}$  ( $i=1, \dots, t$ ), 这样参数向量就限制在一定变化范围内, 只有在此范围内才能用  $b$  来恢复出正常的人脸形状, 而这个变化范围是由训练集来决定的, 亦即 ASM 模型恢复出来的人脸形状是受限于训练集的。
- ✓ 局部纹理建模: 在每个特征点延外轮廓在该点的法线方向一定范围内的灰度变化做统计, 找出该特征点在此区域内的灰度分布规律, 这样, 在定位特征点时只要找到最符合该分布规律的点即是该特征点的位置。

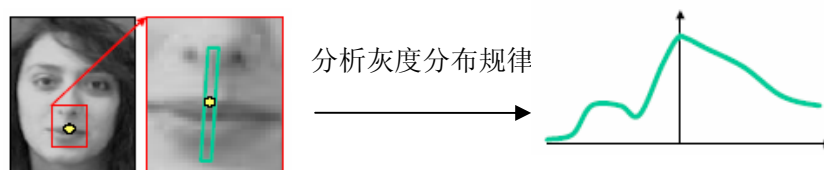


图 8 局部纹理建模过程



对每幅训练图像的特征点如上图取法线，以某特征点为中心取一定数目 ( $n_p$ ) 的像素点，其灰度值用向量表示，然后求灰度沿法线方向导数，并对其标准化处理：

$$h_{ki} = (h_{ki0}, \dots, h_{ki(n_p-1)}), \text{ 第 } k \text{ 幅图中第 } i \text{ 特征点法线方向灰度向量};$$

$$dh_{ki} = (h_{ki1} - h_{ki0}, \dots, h_{ki(n_p-1)} - h_{ki(n_p-2)}), \text{ 法线方向导数}$$

$$g_{ki} = dh_{ki} / \sum_{q=0}^{n_p-2} |h_{ki(q+1)} - h_{kiq}|, \text{ 标准化灰度导数}$$

可以认为这一标准化灰度导数向量整体上满足高斯分布，因此，对于某一候选点对应的标准化灰度导数向量  $g'_i$ ，它与平均标准化灰度导数向量的马氏距离为：

$$d = (g'_i - \bar{g}_i)^T C_i^{-1} (g'_i - \bar{g}_i)$$

$$\bar{g}_i = dh_{ki} / \sum_{q=0}^{n_p-2} |h_{ki(q+1)} - h_{kiq}|$$

$$C_i = \frac{1}{N} \sum_{k=0}^{N-1} (g_{ki} - \bar{g}_i)(g_{ki} - \bar{g}_i)^T$$

这一距离反映了该候选点是这一特征点的概率，在特征点定位过程中，可以通过寻找概率最大的点来确定该特征点的位置。

✓ ASM 目标搜索过程：

1)、目标初始定位。根据其它图像处理的方法得到目标的大致位置，并以此位置进行 ASM 模型的初始化（旋转、缩放和平移），得到初始形状：

$$X = M(s, \theta) \bar{X} + t, \quad \theta \text{ 为旋转角度, } s \text{ 为缩放系数, } t \text{ 为平移矢量}$$

2)、在各个特征点法线方向上进行搜索，即对该范围内的每一点计算标准化灰度导数向量  $g'_i$ ，计算它与  $\bar{g}_i$  的马氏距离  $d$ ，选取距离最小的候选点做最佳匹配点。对模型中每一特征点都做搜索，得到新的形状向量  $X'$ 。

3)、计算与  $X'$  最匹配的参数向量  $b$  (初始时  $b=0$ )，并且对  $b$  进行形状合理性约束，如果  $b$  不满足约束则将其变换到约束范围之内。这样就可以根据  $b$  重构出新的形状。

4)、循环执行 1) 至 3)，当两次相邻循环的形状向量的欧式距离小于一个阈值时结束，此时可以认为 ASM 模型收敛于物体的形状。

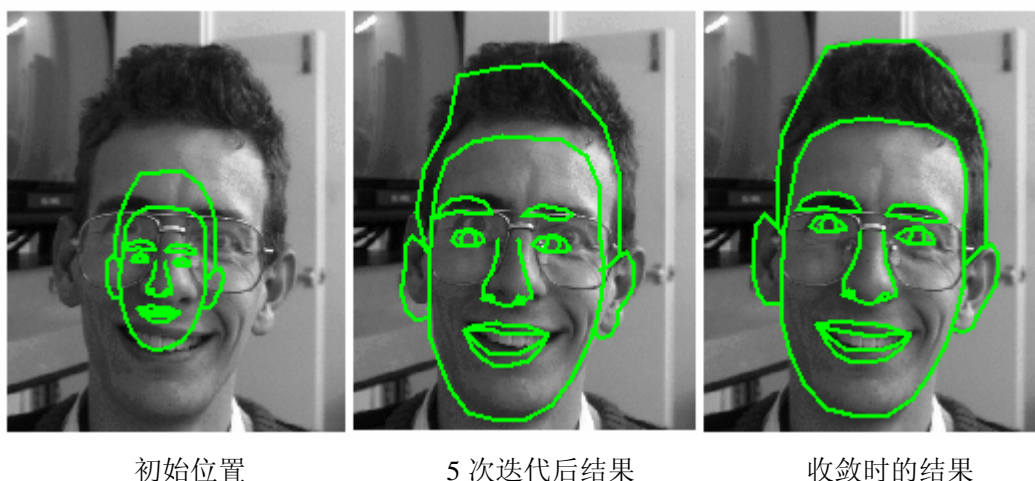


图9 ASM 搜索过程图

## 2.6. 基于主动表观模型

T. F. Cootes<sup>[1,13,14]</sup>等人在 1998 年提出了主动表观模型(Active Appearance Model, 简称 AAM)。AAM 是在 ASM 的基础上扩展得来的, 它不仅建立了反映形状变化的形状统计模型(Shape Model), 同时建立了反映全局纹理变化的全局纹理模型(Texture Model), 并将这两者结合起来建立表观模型(Appearance Model)。得到的表观模型去除了形状和纹理之间的相关性, 可以准确地生成形状及纹理变化的目标图像。在 AAM 搜索过程中, 利用训练得到的图像灰度差值与表观参数变化的线性关系, 不断变化表观模型参数以及二维几何位置和尺度等参数, 直至得到理想的匹配效果。

与 ASM 相比, 由于 ASM 只使用了物体形状轮廓的信息, 而对于灰度信息的利用并不充分, 因此鲁棒性有时不强。AAM 不仅利用了形状信息, 又充分利用了形状纹理之间的关系, 鲁棒性更强; 而且 AAM 可以对形状和纹理进行合成, 从而合成得到物体的图像, 这也是 ASM 所不具备的。从实验结果已经知道, 在人脸特征点定位方面, AAM 的效果确实比 ASM 要好。下面就以人脸特征点定位为例介绍 AAM 的原理:

### 1)、AAM 中人脸的表示形式:

在 AAM 中人脸的表示形式由两部分组成: 形状(shape)及纹理(texture), 分别以矢量  $x$  和  $g$  来表示。形状的矢量化同 ASM; 纹理的矢量化首先要将图像归一化, 亦即将图像进行扭曲, 使其特征点与平均模板中的特征点重合,

于是得到一个与形状无关的 shape-free patch，然后扫描归一化后的图像并存于矢量  $g$  中。

## 2)、Appearance Model

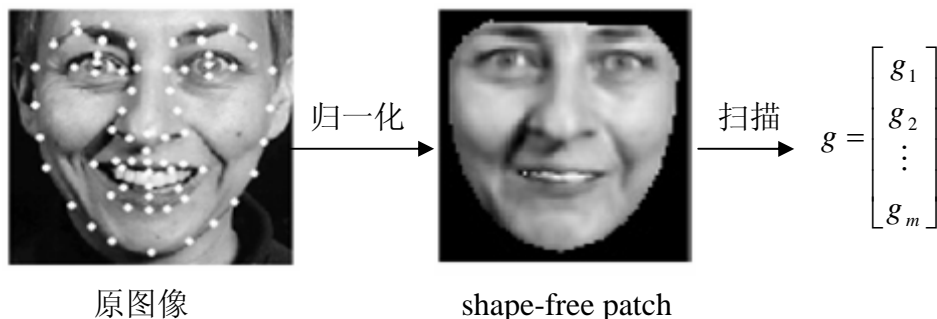


图 10 纹理矢量化过程

外观模型体现了物体的轮廓及轮廓区域内灰度的变化，它可以看成是形状模型和纹理模型的综合。在外观模型中，物体的形状（特征点的集合）和纹理（或灰度级）可以用矢量  $x$  和  $g$  表示，则外观模型可以表示如下：

$$\begin{aligned} x &= \bar{x} + Q_s c \\ g &= \bar{g} + Q_g c \end{aligned}$$

其中， $\bar{x}$ ， $\bar{g}$  分别为形状和纹理的均值， $Q_s$ ， $Q_g$  分别为描述形状和纹理的变化模式的矩阵， $\bar{x}$ ， $\bar{g}$ ， $Q_s$ ， $Q_g$  都是通过对训练集的学习得来的， $c$  为控制参数矢量，同时控制形状和纹理的变化。

这样，若已知  $c$ ，就可以通过  $g$  生成物体的纹理，然后根据  $x$  中的控制点进行扭曲，由此可合成物体的图像。因此，在已确定  $\bar{x}$ ， $\bar{g}$ ， $Q_s$ ， $Q_g$  的条件下，一个物体可以由相应的  $c$  矢量表示。调整  $c$  就可得到不同的物体。图 11 给出了人脸外观模型中对其中两个参数进行调整的例子：



图(a)和(b)中心为通过训练得到的平均模板，图(a)为对参数  $c_1$  进行调整的结果，图(b)为对参数  $c_2$  进行调整的结果。

图 11 AAM 模型参数调整效果

3)、形状模型的建立：外形模型的建立其实也就是对训练集进行统计分析的

过程，与 ASM 的建立相似，在进行分析之前必须先将训练集中的物体规一化（对齐），然后进行 PCA 分析，即可得到外形模型：

$$x = \bar{x} + P_s b_s$$

其中  $\bar{x}$  是形状向量的均值， $b_s$  是形状控制参数， $P_s$  是形状变换模式矩阵。

#### 4)、纹理模型的建立

为建立纹理的统计模型, 首先用 thin-plate spline 技术对图象进行变形, 把每一幅图象都变形到平均脸形状。将变形后的 shape-free patch 的灰度信息矢量化为  $g^m$ 。为最小化光照变化的影响, 对  $g^m$  进行标准化处理：

$$g \rightarrow (g - \mu_g \mathbf{1}) / \sigma_g, \quad \mathbf{1} \text{ 为单位向量, } \mu_g \text{ 和 } \sigma_g^2 \text{ 分别为 } g \text{ 的均值和方差}$$

标准化后,  $g^T \mathbf{1} = 0$  并且  $|g| = 1$ 。对标准化的数据进行 PCA 分析, 可以得到：

$$g = \bar{g} + P_g b_g$$

$\bar{g}$  是标准化灰度向量的均值， $b_g$  是灰度参数， $P_g$  是灰度变换模式矩阵。

#### 5)、外形模型和纹理模型的合成

现在，物体的外形和纹理可以用  $b_s$  和  $b_g$  来表示，由于外形和纹理之间存在一定的关系，可以对数据做进一步的 PCA 分析。对于每幅图象构造新的向量  $b$ ：

$$b = \begin{pmatrix} w_s b_s \\ b_g \end{pmatrix} = \begin{pmatrix} w_s p_s^T (x - \bar{x}) \\ p_g^T (g - \bar{g}) \end{pmatrix}, \quad \text{其中 } w_s \text{ 是形状参数权重系数的对角阵。}$$

对  $b$  做 PCA 分析，可以得到：

$$b = P_c c, \quad P_c = \begin{pmatrix} P_{cs} \\ P_{cg} \end{pmatrix}$$

因此，外形模型和纹理模型可以化成：

$$x = \bar{x} + P_s W_s^{-1} P_{cs} c, \quad g = \bar{g} + P_g P_{cg} c \quad \text{即：}$$

$$\begin{aligned} x &= \bar{x} + Q_s c \\ g &= \bar{g} + Q_g c \end{aligned}, \quad \text{其中 } \begin{aligned} Q_s &= P_s W_s^{-1} P_{cs} \\ Q_g &= P_g P_{cg} \end{aligned}$$

此时，AAM 模型的建立完毕。

#### 6)、AAM 搜索

✓ 目标粗略定位：使用特征脸定位方法进行目标的粗略定位，得到当前的

模型控制参数初始值  $c_0$ 。

✓ AAM 迭代过程:

令  $\delta g = g_i - g_m$  ,  $g_i$  为原图像的灰度,  $g_m$  为使用 AAM 模型合成的图像。物体的定位过程其实就是不断对控制参数  $c$  做调整 (计算  $\delta c$ ), 使  $\delta I$  最小化的过程。具体流程如下:

- a). 令  $k=1$
- b). 计算合成图像灰度与原图像的灰度之间的差异  $\delta g$  ;
- c). 根据  $\delta I$  计算  $\delta c^*$  , 得到新的控制参数:

$$c \leftarrow c - k\delta c$$

- d). 计算新的图像灰度  $g'_i$  和  $\delta g$  , 如果  $|\delta g|^2 < E_0$  ,  $E_0$  为常数, 则收敛, 此时可以由  $c$  得到图像中物体的特征点; 否则令  $k$  等于不同的值, 如 1.0、0.5、0.25 等等, 然后返回 b)。

\*假定  $\delta I$  与  $\delta c$  呈线性关系:

$$\delta c = A\delta I$$

在搜索时由  $\delta I$  来确定  $\delta c$  , 而在确定  $A$  时则反过来, 先人为地给出  $\delta c$  (人为扰动), 据此调整模型得到合成图像并计算  $\delta I$  ,  $A$  即可求出。通过制造各种扰动可以得到较准确的  $A$ 。

此外, 可以在多分辨率的情况下完成 AAM 的搜索过程, 在每一层都进行迭代直收敛, 然后将当前结果影射到下一层的模型。多分辨率的方法比单分辨率的方法效率更高并能收敛到正确结果。图 12 体现了 AAM 搜索过程中形状和纹理的变化。

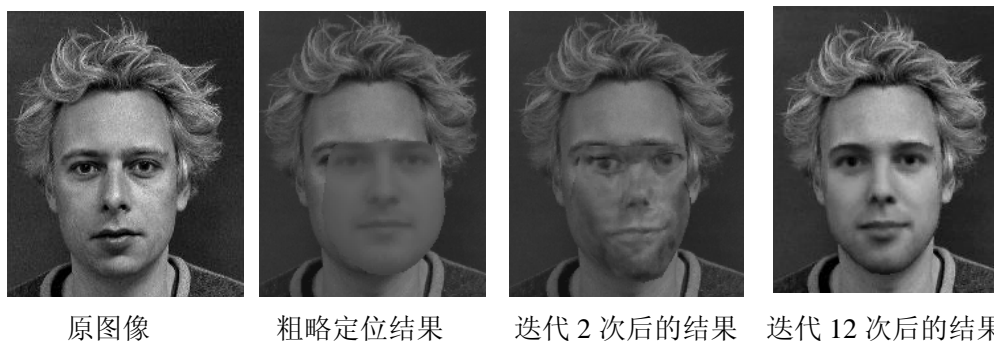


图 12 AAM 搜索过程

## 第三章 对 AAM 模型的改进—基于肤色信息的 AAM 模型

### 3.1. 肤色模型

肤色是人脸的重要信息，不依赖于面部的细节特征，对于旋转，表情等变化情况都能适用，具有相对的稳定性并且和大多数背景物体的颜色相区别。因此肤色特征在人脸检测中是最常用的一种特征，在人脸特征定位中也经常作为图像预处理的方法。肤色特征主要由肤色模型描述。

#### 3.1.1. 肤色概述

使用何种形式的肤色模型与色度空间（Chrominance Space）的选择密切相关。可以从两个方面考察某种色度空间：（1）在该色度空间中能否用给定的模型描述“肤色”区域的分布；（2）色度空间中“肤色”与“非肤色”区域的重叠有多少。人脸检测常用的色度空间主要有：RGB（红、绿、蓝三基色）、rgb（亮度归一化的三基色）[11]、SHI（饱和度、色调、亮度）、YIQ（NTSC 制的光亮度和色度模型）、YUV（PAL 制的光亮度和色度模型）、YCbCr（CCIR601 编码方式的色度模型，与 YUV 在数学上具有等价性）、CIEL\*a\*b（国际照明委员会提出的基于色度学的彩色模型）等。

常用的肤色模型有高斯模型<sup>[26]</sup>、混合高斯模型<sup>[27,28]</sup>和直方图模型<sup>[29]</sup>。除这三种肤色模型外，还有直接利用几何参数描述肤色区域分布范围的模型<sup>[30]</sup>、三维投影模型<sup>[31]</sup>、基于神经网络的肤色模型<sup>[32,33]</sup>等。此外也有同时考虑“肤色”与“非肤色”像素分布的基于贝叶斯方法的模型<sup>[34]</sup>。

#### 3.1.2. 高斯模型

本文采用的肤色模型是高斯肤色模型，并选用 YCrCb 色彩空间作为肤色统计分布的映射空间。理由是：

（1）在 YCrCb 色彩空间中，Y 代表明度的变化，Cr 和 Cb 分别代表红色和蓝色的变化，这样就把色彩中的亮度分量分离出来，分析时受到亮度影响较小。

（2）分析 Cr 和 Cb 可以将三维的 RGB 降为二维，而且 Cr 和 Cb 两维独立分

布。在这个二维空间里，肤色的分布相对集中，聚类效果较好；

(3) YCrCb 色彩格式被广泛应用于电视等领域中，采用 YCrCb 颜色空间在视频中可以避免额外色彩空间转换的计算。

令肤色像素在色度空间中表示为向量  $x = (Cr, Cb)^T$ ，这些像素在 Cr-Cb 空间中的分布相对很集中，可以用 Gauss 分布来描述：

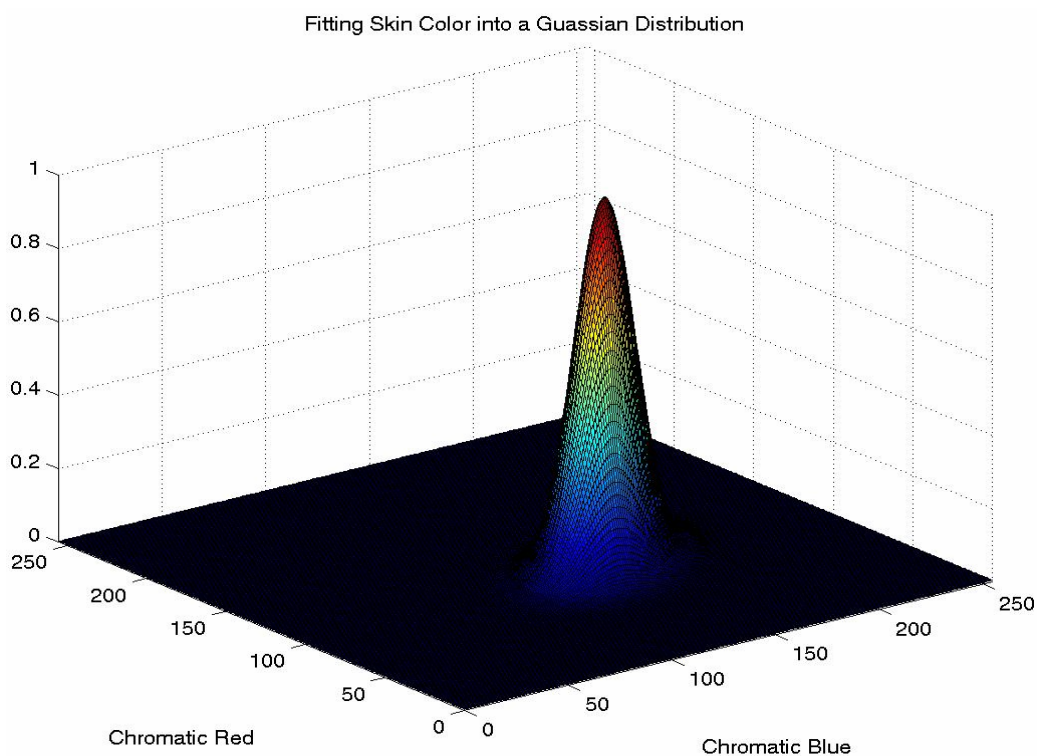


图 13 高斯分布

我们可以采用训练的方法来计算得到这个分布的中心，然后根据所考察的像素离该中心的远近来得到它和肤色的相似度，从而得到一个原图的相似度分布图，再按照一定的规则对该分布图二值化，以最终确定肤色的区域（见图 14）。计算相似度时，采用公式：

$$P(Cr, Cb) = \exp\left[-\frac{1}{2}(x - m)^T C^{-1}(x - m)\right]$$

$$x = [Cr, Cb]^T, \quad M = E(x)$$

$$C = E((x - M)(x - M)^T)$$

其中，C 和 M 分别是协方差和均值，是在训练的时候确定的。

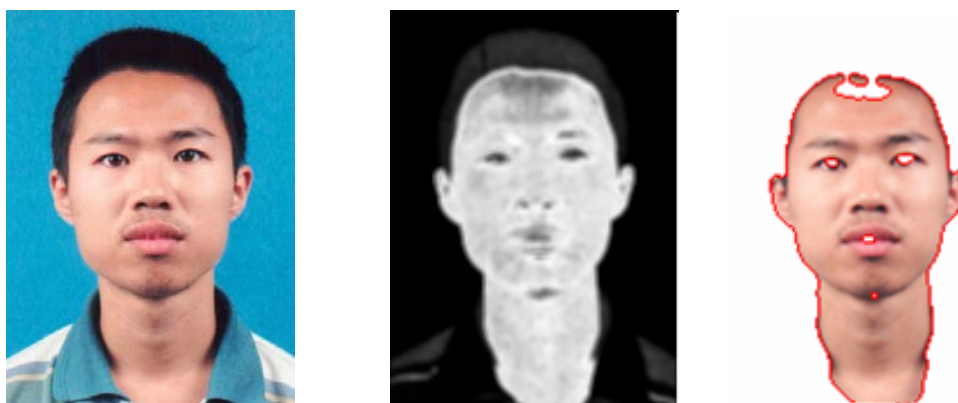


图 14 相似度图与二值化图

### 3.2. 基于肤色信息的 AAM 模型

在人脸检测方面，由于人脸肤色的 Cr, Cb 分量在二维空间上分布具有一定的规律性，所以利用肤色信息可以较容易地把脸部区域找出来。目前肤色模型在人脸检测方向的应用已经很广范。在人脸特征点定位方面，由于肤色模型只能分辨出肤色与非肤色，而不能用来检测人脸器官特征的精细位置，因此往往只用在图像的预处理阶段；至于特征点的定位则要用其它方法来解决。

在实验我们发现，由于人脸本身固有的几何特征（各器官的大小、比例和位置关系），在对人脸进行肤色分割后得到的结果往往也蕴含着这些信息。是否可以将这些信息利用起来，并以此对人脸器官做粗略地定位呢？这就是本文所提出方法的初衷。下面的内容将详细描述这一方法，并介绍其在对人脸的 AAM 搜索中的应用。

#### 3.2.1. 基于肤色模型的人眼定位方法

对人脸图像使用肤色模型计算各个像素与肤色的相似度，然后对图像进行二值化分割，可以得到一张人脸的二值图。经过分析可以看出，由于人的眼睛、眉毛、鼻孔等均属于非肤色区域，因此在二值图中这些器官所对应的区域往往会产生孔，尤其是人的眼睛区域。因此，如果在二值图中能够找在人眼所在“双孔”的位置，就可以确定人眼的位置。如图 15 所示：

由上述的方法可以确定双眼的位置，但是双眼结构的确定往往受到一些噪音“孔”的影响，比如鼻孔、眉毛所产生的孔，也就是说，除了真实的眼对外，还



存在一些可能是眼对的候选眼对，因此确定双眼结构时要将这些候选眼对排除，此时可以用眼睛的灰度分布来检测，一种比较直接的方法是模板匹配的方法。通过计算候选眼对与模板的匹配程度，从中选取最匹配的候选眼对做为最后的结果。

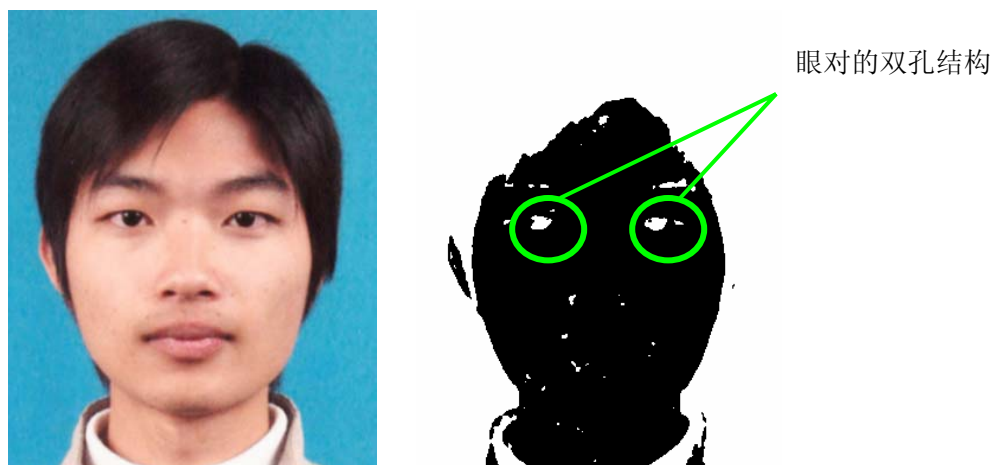


图 15 人脸二值化图中的双孔结构

下面将对上述方法的步骤做一个详细的描述：

1)、计算相似度：

使用肤色模型，计算人脸图像的肤色相似度，得到一张相似度灰度图（如图 16(b)所示），图中相似度越高像素的灰度值越大。

2)、图像分割：

用自动阈值分割算法对灰度图进行区域分割（即二值化），得到人脸二值图（如图 16（c）所示）。

3)、去噪声处理：

在完全没有先验知识的情况下，对于一些独立的肤色像素点，想将它们归为一个个人脸区域，我们很自然的会想到找连通区的方法。属于同一个人脸区域的肤色像素点应该是相互连通的；反之，一个人脸区域也应该是由一些连通的肤色像素点构成的。但由于噪声普遍存在于数字图像中；还有基于肤色模型的分割结果存在判断失误的情况，都会在连通的人脸区域中产生非肤色的噪声。所以进行人脸区域分割之前要进行去噪声或者称为滤波的处理。

在这里我们采用了基于像素“密度”的滤波方法来去除噪声。这是一种类似“卷积”的滤波方法。直观的讲，就是在以每一个肤色像素为

中心的  $5 \times 5$  邻域内统计肤色像素的个数, 超过半数时中心点保留为肤色, 否则认为是非肤色 (见图 16(d))。

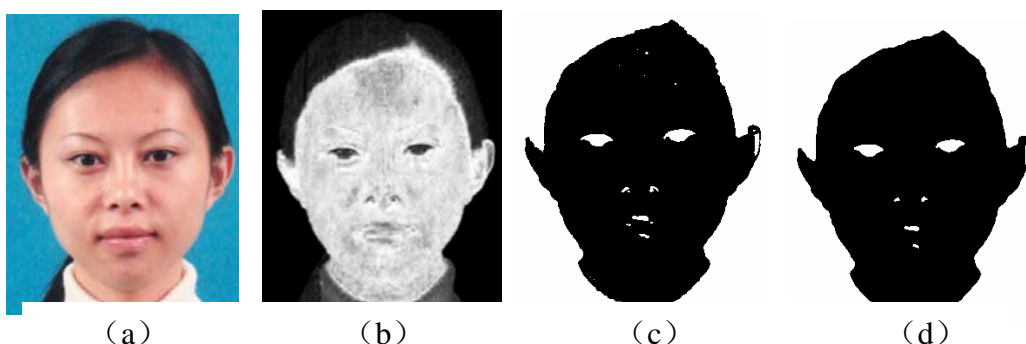


图 16 计算相似度, 二值化及去噪声处理

#### 4)、获取二值图中的所有“孔”:

要获取二值图中的孔有很多方法, 我采用的方法是: 先得到一个二值图的“掩盖图”, 该图把二值图中所有孔都填充掉; 然后求这个“掩盖图”与原二值图的异或, 结果中值为 1 的区域即是原二值图的孔。步骤如下:

(a)、用轮廓跟踪算法求得二值图的轮廓 (见图 18(b)):

按照从左到右, 从下到上的顺序搜索, 找到的第一个黑点必是最左下方的边界点, 记为 A, 并将其标记位设 1 (初始时所有点的标记位为 0)。它的右、右上、上、左上四个邻点中至少有一个是边界点, 记为 B, 并将其标记位设 1。从 B 开始找起, 按右、右上、上、左上、左、左下、右下的顺序找相邻点中的边界点 C。如果 C 标记位为 1, 则表明已经转了一圈, 找到一条轮廓; 否则从 C 点继续找, 直到找到轮廓为止。找到一条轮廓后再对图遍历, 以任一个标记位为 0 的黑点开始跟踪轮廓。如此循环直至所有黑点的标记位都为 1。

由于二值图中可能会出现如下图所示的情况, 在跟踪边界时会出现异常中止的情况, 因此必须将过细的线条去掉。可以用数学形态学的方法, 对滤去噪音后的二值图像 A 进行腐蚀运算:

$$A \leftarrow (A \ominus B_1) \ominus B_2$$

其中  $B_1$  和  $B_2$  分别为  $3 \times 1$  和  $1 \times 3$  的结构元素:

$$B_1 = [1 \ 1 \ 1]^T, \quad B_2 = [1 \ 1 \ 1]$$

通过腐蚀运算, 图像中宽度为 1 的连通线将被切断, 如图 17 所示。



图 17 细连接线所造成轮廓跟踪时的错误及经过腐蚀运算后的轮廓跟踪

(b)、对轮廓进行种子填充, 得到二值图的“掩盖图”(见图 18(c))。

(c)、对原二值图和“掩盖图”求异或, 得到二值图中的所有孔  $E$ :

$E = \{e_1, e_2, \dots, e_m\}$ ,  $e_i$  表示第  $i$  个孔,  $m$  为所有孔的个数(见图 18(d))。

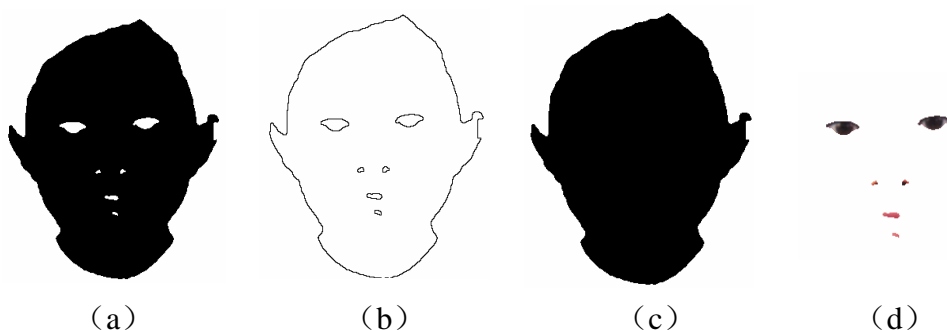
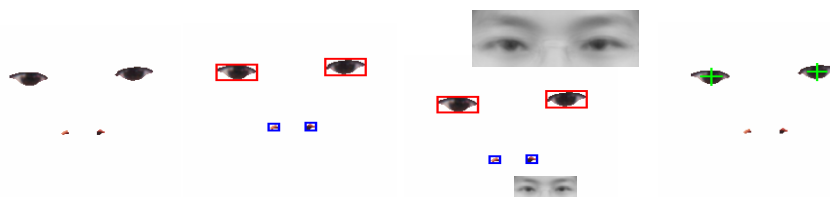


图 18 轮廓、掩盖图、二值图中的孔

5)、对  $E$  中挑选出候选眼  $E' = \{e'_1, e'_2, \dots, e'_n\}$ ,  $e'_i$  表示第  $i$  个候选眼,  $n$  为候选眼的个数。当孔满足以下条件时为候选眼:

a)、 $e_i$  所在区域像素的  $Y$  值的平均值  $\bar{Y} < Y_0$ ,  $Y_0$  为常数(可定为 80)。这是因为眼球所在区域的  $Y$  值一般在 80 以下。

b)、 $e_i$  的高度和宽度分别大于固定值  $H_0$  和  $W_0$ 。



(a) 候选眼 (b) 候选眼对 (c) 模板匹配 (d) 眼睛定位

图 19 眼睛的定位

6)、将  $E'$  中的候选眼两两配对, 若  $e_i$  和  $e_j$  满足下列条件, 则它们构成候选眼对, 由此得到候选眼对集  $M = \{m_1, m_2, \dots, m_k\}$ ,  $m_i$  表示第  $i$  个候选眼对,  $k$  为候选眼对个数:

- a)、 $e_i$  和  $e_j$  的包围盒不相交;
- b)、 $e_i$  和  $e_j$  的中心连线的斜率  $l$  满足  $|l| < l_0, l_0$  为常数;
- c)、 $e_i$  和  $e_j$  的大小之比  $k_0 \leq size(e_i)/size(e_j) \leq \frac{1}{k_0}, k_0$  为常数。

7)、对每个候选眼对  $m_i$ ，根据两个候选眼区域中心点的距离和角度，对平均眼对模板进行缩放和旋转，然后计算模板与相应区域的匹配程度  $p_i$ 。选择最大匹配程度值  $p_{\max} = \max\{p_1, \dots, p_k\}$  对应的候选眼对为结果眼对。若  $p_{\max} < p_0$  (常数, 比如 0.6), 则人眼定位失败。

### 3.2.2. 肤色模型在 AAM 搜索中的应用

AAM 模型能够有效地利用物体的形状信息和灰度变化信息来进行特征点的定位, 但是在对彩色图像进行特征点定位时无疑忽略了色彩信息的作用。通过上述的方法可以确定人眼的粗略位置和方向, 由于人脸五官位置分布的规律性, 可以由此确定人脸的位置及方向, 由此可以为 AAM 模型的搜索提供一个较好的初始位置和大小, 并且可以缩小 AAM 搜索窗口的大小, 从而减少 AAM 的搜索时间。下面将详细介绍该方法的具体步骤:

- a)、使用 3.2.1 节所描述的方法确定人脸图像中双眼的位置;
- b)、根据双眼瞳孔连线的角度和长度, 确定 AAM 模型的初始大小和方向。只要已知人脸的 AAM 模型中表示眼睛的特征点的编号, 即可对模型进行缩放和旋转, 使之达到较好的初始位置;
- c)、根据双眼与面部区域的关系, 确定 AAM 模型的搜索窗口。双眼与面部区域的关系是通过对 20 多张人脸图像的分析所得, 我们将这些图像根据眼睛位置



图 20 平均脸, 眼睛与面部区域的关系

对齐，再求平均人脸，得到平均人脸中双眼与面部区域的几何关系如图 20 所示：

d)、使用 AAM 模型进行搜索，得到人脸特征点的精确位置。

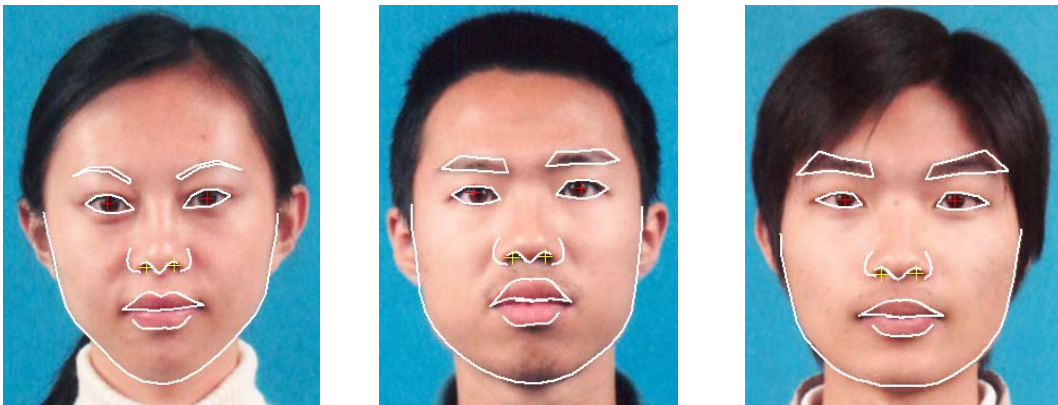


图 21 使用 AAM 搜索得到的人脸特征点

## 第四章 一个人脸肖像画生成系统

人脸特征点定位在各个领域都已经有了相当的应用，如身份认证、安全保卫自动监控系统、可与人交流的智能计算机、可视电话、电视会议等等。随着多媒体娱乐业的发展，人脸特征点定位也渐渐在多媒体智能娱乐和游戏领域有相当应用，其中比较典型的是人物肖像画的自动生成，而且主要是以漫画的形式，包括线条画和油彩漫画。

在线条画方面，MIT 的 Brennan<sup>[35]</sup>最先提出了一个交互式的夸张漫画生成系统，Murakami<sup>[36]</sup>等人提出了一种基于模板的线条画生成方法并开发了 PICASSO 系统，在该方法中，漫画的夸张程度是可调整的。Rein-Lien Hsu<sup>[37]</sup>等人提出了一种交互式的 Snake 模型来生成人脸的线条画（不带夸张效果），并把它应用到人脸匹配中。Li 等人提出了一种自动的人脸线条画生成系统，他们主要使用了对称算子、矩形滤波器和特征轮廓来检测和定位人脸的特征点。与其它方法不同，微软的陈洪等人提出了一种基于实例的线条画生成方法，他们通过对艺术家的作品的学习得到人脸照片到线条画的对应关系，当学习的作品包含夸张效果时，生成的线条画也具有夸张效果。

与线条画漫画的生成不同，Pei-Ying Chiang<sup>[38]</sup>等人提出基于 MPEG-4 的彩色夸张肖像漫画的半自动生成系统，他们的系统以已有的肖像漫画作品为原型，通过夸张化的扭曲使之与待处理的人脸相匹配，从而得到该人脸的漫画。通过使用不同风格的作品作为原型，可以得到不同风格的结果。

在改进 AAM 模型的基础上，我们设计并实现了一个人物肖像画生成系统，以下内容各节是对该系统的详细介绍。

### 4.1. 系统结构

人物肖像画生成系统的结构图如下：

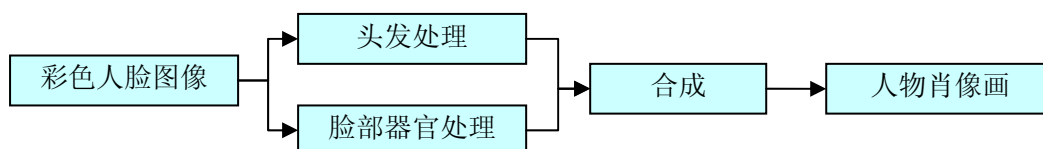


图 22 人物肖像画生成系统结构图

本系统是用 Visual C++ 6.0 开发的，系统的运行界面如下图所示：



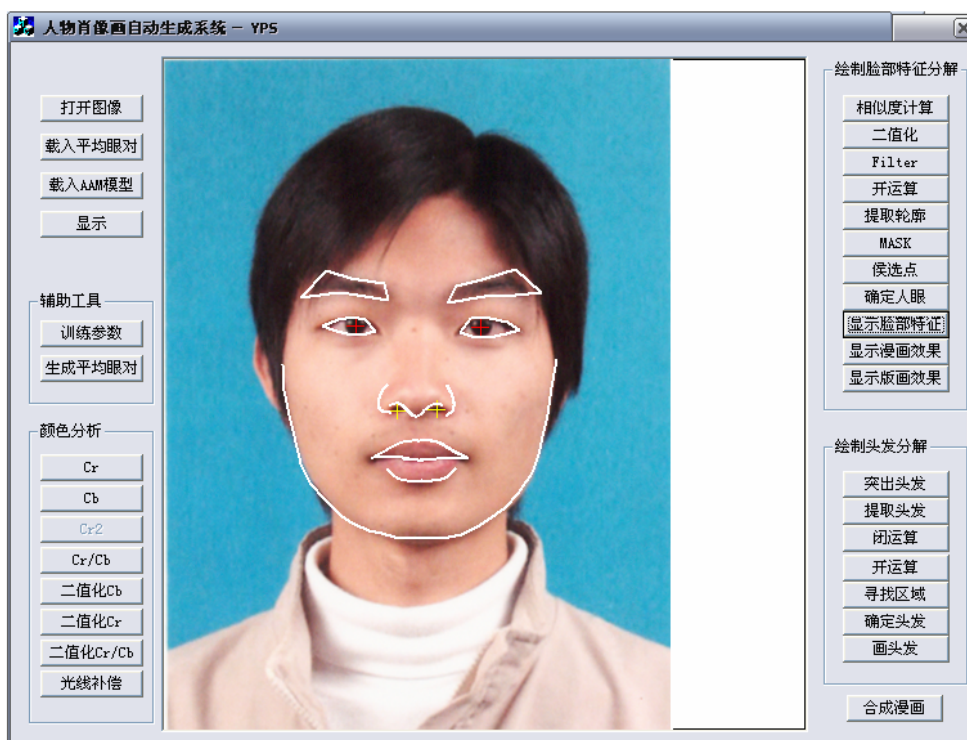


图 23 系统运行界面

## 4.2. 人脸面部特征的提取

使用第三章所介绍的方法，可以获取人脸面部各个特征点的精确定位，然而如何根据这些特征点的信息来生成肖像画？下面将对具体的细节做详细的描述：

### 4.2.1. AAM 模型的建立

我们使用 20 张人脸正面图像进行 AAM 模型的建立。在 AAM 模型中，人脸的特征点一共有 84 个（如图 24 所示）：

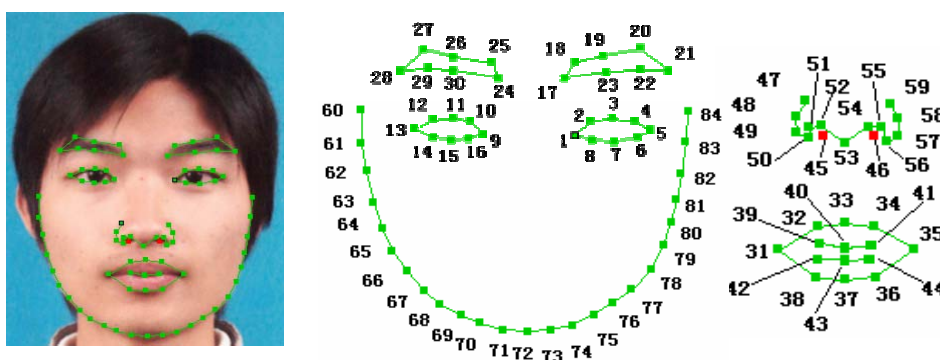


图 24 系统所使用的 AAM 模型中的特征点定义

### 4.2.2. 根据特征点生成面部器官的肖像画效果

特征点信息的使用在肖像画生成系统是非常重要的，我们能想到的最直接的方法是将特征点根据所属脸部器官的不同进行分类，然后用插值曲线把它们连接起来，这样可以得到一幅线条画（见图 25 (a)）；另一种方法则是间接地使用这些特征点的信息，因为人脸器官的灰度分布总是异于脸部的其它区域，而且有明显的边界，因此可以直接从灰度数据里获取信息，本文根据面部各器官特征点来得到各器官的包围盒，然后对该包围盒内的灰度图进行自动阈值分割，从而得到各器官的画像（见图 25 (b)）。

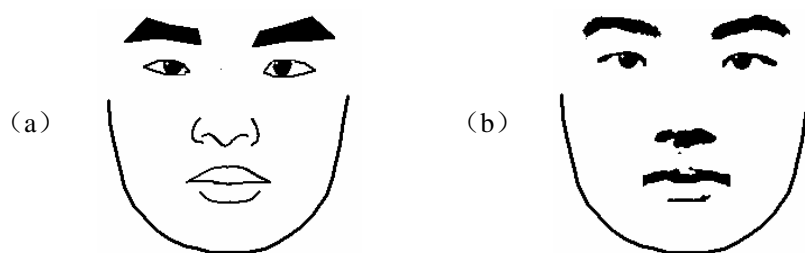


图 25 面部器官的线条画和本版画效果

### 4.3. 头发的提取

由于深色头发的颜色具有特殊的性质，即头发像素的 Y 值满足：

$$Y \leq Y_0, Y_0 \text{ 为常数, 在实验中设为 } 80 \quad (10)$$

因此，可以将不满足此性质的像素去掉，头发必然落在剩下的区域内，从中将头发所在的区域提取出来，跟踪该区域的轮廓即可得到头发的轮廓。具体的步骤如下：

(1)、计算人脸图像中各个像素的 Y 值，将不满足 (10) 条件的像素去除，得到二值化图 G。

(2)、对 G 进行数学形态学中的闭运算，将头发中的噪音（白色小孔）滤去，得到二值图 A，其中结构元素为：

$$A \leftarrow (G \oplus B) \ominus B \quad B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

(3)、可以用数学形态学的方法，对 A 进行两次腐蚀运算：



$$A \leftarrow (A \ominus B_1) \ominus B_2$$

其中  $B_1$  和  $B_2$  分别为  $3 \times 1$  和  $1 \times 3$  的结构元素:

$$B_1 = [1 \ 1 \ 1]^T, \quad B_2 = [1 \ 1 \ 1]$$

通过腐蚀运算，图像中宽度为 1 的连通线将被切断。

(4) 根据 3.1.1 中所介绍的轮廓跟踪算法跟踪图 A 中的轮廓线，每一轮廓线对应一个矩形区域  $r_i$ ，得到一个区域列表

$$R = \{r_0, r_1, \dots, r_m\}, \quad m \text{ 为轮廓线个数}$$

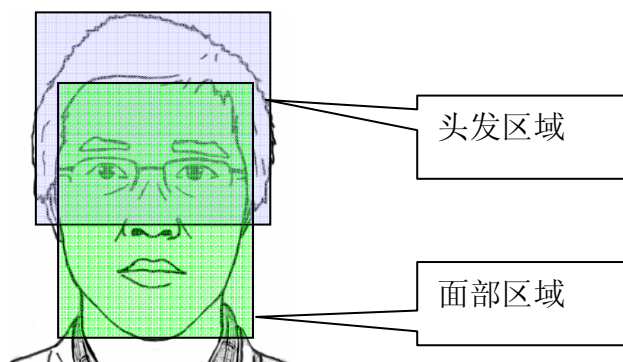


图 26 头发区域与面部区域的关系

(5) 对  $R$  中任一区域  $r_i$ ，如果  $r_i$  满足如图 26 所示的条件（即头发区域处于脸部区域上方，而且与脸部区域相交），因为满足这样条件只可能有一个，故  $r_i$  为头发所在区域。

各阶段的结果如图 27 所示：

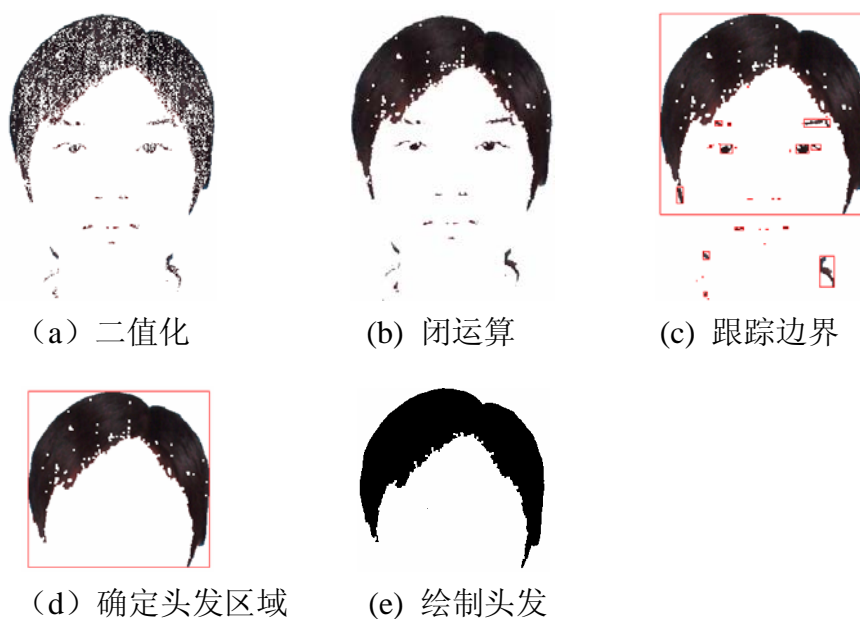


图 27 头发的提取过程

#### 4.4. 肖像画的合成及显示

将 4.2. 所得的脸部器官图与 4.3. 的结果进行组合，便可合成人脸的肖像画（如图 28 所示）。

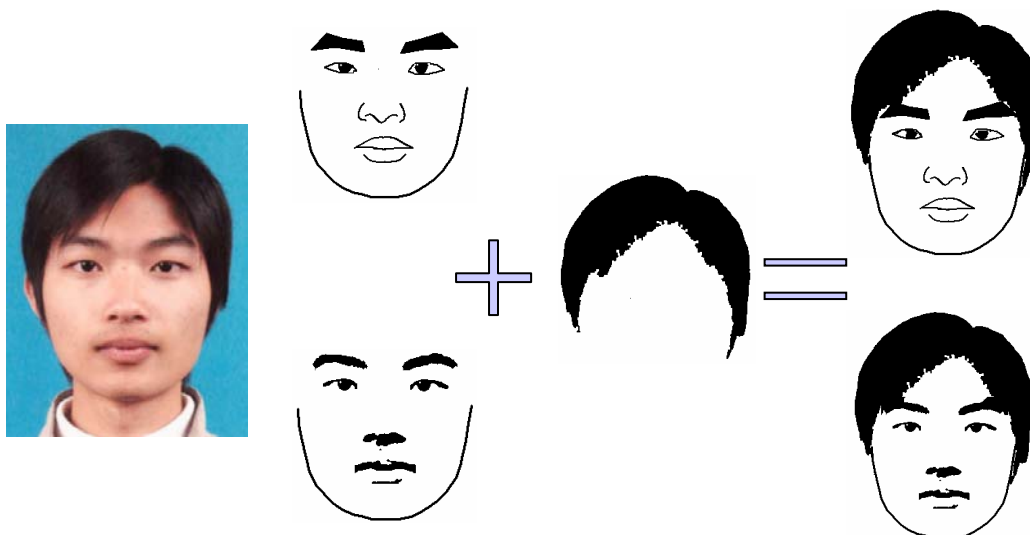


图 28 肖像画的合成

此外，我们不但可以得到人物的肖像画，还可以通过对特征点进行移动，然后使用图像变形的方对肖像画进行扭曲变形，从而可以得到不同表情的肖像。

#### 4.5. 系统总结与分析

从系统的结果来看，对于大多数人脸图像，都可以得到相应的肖像图，但是在实验当中我们也发现的系统的一些问题和待改进的地方：

(1)、系统没有对耳朵进行处理，这主要是因为人脸图像中并不总是能看到耳朵，经常会出现耳朵被头发（全部或部分）遮住的情况，这导致处理时比较难找到一个全面的解决方法；

(2)、将系统的生成的肖像画与原图做比较，我们发现两者的相似性并不是很高，特别是线条画的结果。原因其实很简单，就象我们能够很容易地对一幅人脸图象手工标记各个特征点，但是要一般人用线条来描绘出人脸的特征却很难。也就是说，从特征点集及人脸灰度分布到肖像画之间其实存在着一种映射关系（如图 29 如示），一般人并不能很好地把握这些关系，而画家却能很好地把握，而且不同风格的画家所使用的映射关系不同。

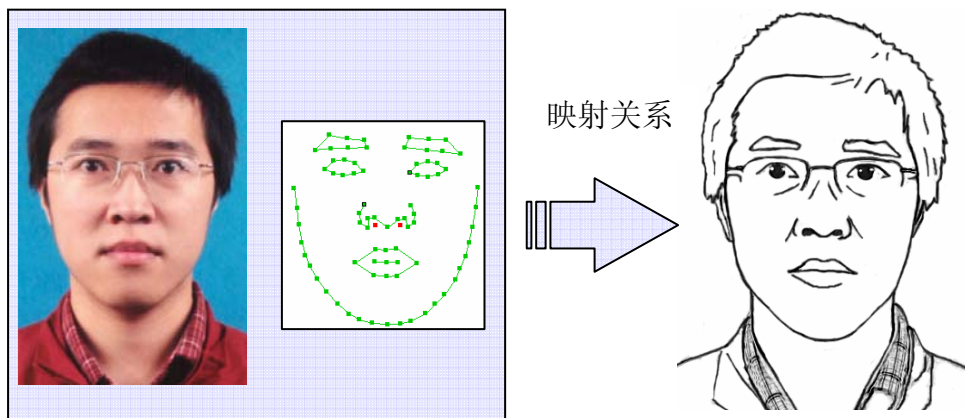


图 29 特征点集及人脸灰度到肖像画之间的映射关系

(3)、系统对头发的处理并没有考虑如下图例子，亦即头发被分割成若干区域的情况(见下图)。



图 30 头发被分割成多区域的情况

(4)、头发边界的处理太过粗糙(如图 31 所示)，常出现锯齿状边缘。针对这种情况可以使用 Snake 模型获取头发区域的外轮廓，这是以后待改进的地方之一。

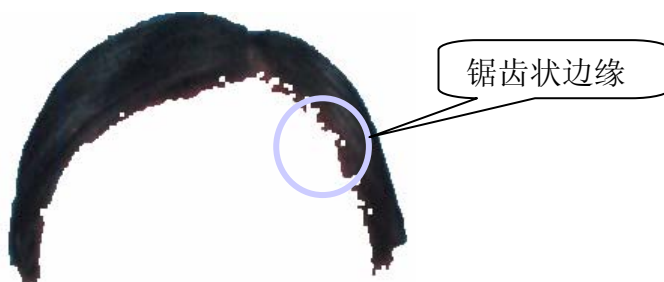


图 31 头发的锯齿状边缘

## 第五章 结论与展望

本文主要研究了人脸图像的特征点定位方法，总结了目前常用的一些特征点定位算法，提出了对 AAM 模型的一种基于肤色模型的改进方法，并将其应用在计算机肖像画生成之中。初步提出并设计实现了一整套较为完整的人物肖像自动生成系统。

### 5.1. 论文的主要工作和结论

(1)、总结了常用的特征点定位算法，包括基于灰度及变化信息的方法、基于神经网络的方法、Snake 模型、可变形模板方法、ASM 模型和 AAM 模型。

(2)、提出了对 AAM 模型的一种基于肤色模型的改进方法，将其应用到人脸特征点定位上，并用 Visual C++ 6.0 实现了该方法。

(3)、设计一个完整的人物肖像自动生成系统，并用 Visual C++ 6.0 实现该系统。并且对该系统进行了分析和总结。

### 5.2. 尚需进一步研究的问题

(1)、寻找一种能够在各种情况下都有效的获取耳朵轮廓的方法；

(2)、研究人脸图像及其特征点与画家作品之间的映射关系，从而得到与原图像比较想像的人物肖像画；

(3)、针对头发被分割成多区域的情况，研究能够识别头发整体区域的方法。并对头发的边缘进行平滑处理，避免锯齿状轮廓的出现；

(4)、本文提取的肖像画生成方法主要是针对正面图像，如何生成任意姿态和角度的人物肖像画仍有待研究；

(5)、如何绘制不同人种的人(比如, 亚洲人、非洲人和欧洲人的面部结构的几何细节是有差异的)；

(6)、如何生成夸张效果的肖像像。

## 参考文献

- [1] G.J. Edwards, C.J. Taylor, T. Cootes, Interpreting Face Images using Active Appearance Models. *Proc. of the 3rd Int. Conf. on Automatic Face and Gesture Recognition*, Nara, Japan, pp 300-305, 1998
- [2] Cootes T, Taylor C, Cooper D, Graham J. Active Shape Models – Their Training and Application. *Computer Vision and Image Understanding*, January 1995, Vol. 61, No. 1, pp. 38-59
- [3] Vladimir Vezhnevets, Vassili Sazonov, Alla Andreeva, A Survey on Pixel-Based Skin Color Detection Techniques.
- [4] M.Kass, A. Witkin, and D.Terzopoulos, Snakes: Active contour models. *IJCV*, pp.321-331, 1988
- [5] LeeW. S, Kalra P, Magnenat-Thalmann , Model Based Face Reconstruction for Animation, *Proc. Multimedia Modeling (MMM'97)*, Singapore, pp. 323-338, 1997
- [6] A.L.Yuille, Deformable templates for face detection. *J. Cogn. NeuroSci.* 3, 59-70, 1991
- [7] A.L.Yuille, P.Hallinan, D.S.Cohen, Feature extraction from faces using deformable templates. *IJCV*, 8 August 1992, pp. 99-11.
- [8] L. Wiskot, J.M. Fellous, N. Kruger, C.v.d.Malsburg, Face Recognition by Elastic Bunch Graph Matching. *IEEE Trans. On PAMI*, Vol. 19, No. 7, pp775-779, 1997
- [9] V. Kr"uger and G. Sommer. Gabor wavelet networks for object representation. Technical Report 2002, Institute of Computer Science, University of Kiel, 2000
- [10] T.F. Cootes, C.J. Taylor, D.H.Cooper, J.Haslam, Training Models of Shape from Sets of Examples. *Proc. British Machine Vision Conference*. Springer-Verlag, 1992, pp. 9-18
- [11] T.F. Cootes, C.J.taylor, Active shape models -"smart snakes". *Proceeding of the British Vision Conference*, Leeds, UK, 22-24 September 1992, Lodon,1992, Springer-verlag, p266-275.
- [12] T.F. Cootes, C.J. Taylor, A. Lanitis, Multi-Resolution with Active Shape Models *Proc. ICPR* 1994.

- [13] T.F. Cootes, G.J. Edwards, C.J. Taylor, Active Appearance Models. *Proc. European Conf Computer Vision*, vol. 2, pp. 484-498, 1998.
- [14] T.F. Cootes, C.J. Taylor. "Statistical models of appearance for computer vision". <http://www.isbe.man.ac.uk/-bin/refs.html>, 2001
- [15] L. Cohen, Note on active contour models and balloons. *CVGIP: Image Understanding* 53(2), 1991, 221-218.
- [16] R. Ronfard, Region-based strategies for active contour models. *IJCV* 13(2), October 1994, 229-51.
- [17] P. Fua, C. Brechbuhler, Imposing hard constraints on soft snakes. *Proc. Eur. Conf. On Computer Vision*, Vol. II, Cambridge, UK, 1996, pp. 495-506.
- [18] A. Ameni, T. Weymouth, R. Jain, Using dynamic programming for solving variational problems in vision. *IEEE Trans, Pattern Anal. Mach. Intell.*, 12(9), 1990, 8555-8567.
- [19] V. Casselles, F. Catte, T. Coll, F. Dibos, A geometric model for active contours in image processing. *Numberische Mathematik*, 66(1993), 1-31.
- [20] R. Malladi, J. Sethian, B. Vemuri, Shape modeling with front propagation: A level set approach. *IEEE Trans, Pattern Anal. Mach. Intell.*, 17(2), 1995, 158-175.
- [21] G.E. Christensen, R.D. Rabbitt, M.I. Miller, Deformable templates using large deformation kinematics. *IEEE Trans. Image Process*, 5(10), 1996, 1435-1447.
- [22] S. Menet, P.Sant-Marc, G. Medioni, B-snakes: Implementation and application to stereo. *DARPA Image Understanding Workshop*, 1990, pp. 720-726.
- [23] Horace H. S. Ip and D. Shen, An affine-invariant active contour model (ai-snake) for model-based segmentation. *Image and Vision Computing*, 16(2), Feb. 1998.
- [24] Paul Debevec, A Neural Network for Facial Feature Location. UC Berkeley CS283 Project Report, December 1992, <http://athens.ict.usc.edu/FaceRecognition/>
- [25] A.L. Yuille, Peter W. Hallinan, David S. Cohen, Feature extraction from faces using deformable templates. *IJCV*, 8(2): 99-111, 1992.
- [26] Menser, B., Wien, M., Segmentation and tracking of facial regions in color image sequences. *Proc. SPIE Visual Communications and Image Processing 2000*, 731-740.
- [27] Terrillon, J.-C., Shirazi, M. N., Fukamachi, H., Akamatsu S., Comparative performance of different skin chrominance models and chrominance spaces for the

automatic detection of human faces in color images. *Proc. of the International Conference on Face and Gesture Recognition*, 54–61.

[28]Yang, M., AHUJA, N. Gaussian mixture model for human skin color and its application in image and video databases. In *Proc. of the SPIE: Conf. on Storage and Retrieval for Image and Video Databases (SPIE 99)*, vol. 3656, 458–466, 1999.

[29]M.J. Jones, J.M. Rehg. Statistical color models with application to skin detection. CVPR, 1999

[30]Garcia C, Zikos G, Tziritas G, Face detection in color images using wavelet packet analysis, In: *Proc. Multimedia Computing and Systems, Centro Affari, Florence, Italy*, 1999, Vol.1: 703-708

[31]Abdel-Mottaleb M, Elgammal A. Face detection in complex environments from color images. In: *Proc. IEEE Conf. on Image Processing, Kobe, Japan*, 1999, Vol.3 (11-13): 622-626

[32]Karlekar J, Desai U B. Finding faces in color images using wavelet transform. In: *Proc. IEEE Conf. on Image Analysis and Processing, Venice, Italy*, 1999, 1085-1088

[33] Ishii H, Fukumi M, Akamatsu N. Face detection based on skin color information in visual scenes by neural networks. In *Proc. Systems, Man, and Cybernetics, Tokyo, Japan*, 1999 Vol.5: 350-355

[34]Zarit B D, Super B J, Quek F., Comparison of five color models in skin pixel classification. In: *Proc. Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time systems, Corfu, Greece*, 1999, 58-63

[35]Brennan SE., Caricature generator [MS. Thesis]. Cambridge, MA: MIT, 1982.

[36]Tominaga M, Fukuoka S, Murakami K, Koshimizu H. Facial caricaturing with motion caricaturing in PICASSO system. In: *Proceedings of the IEEE/ASME International Conference on Advanced Intelligent Mechatronics'97*. 1997,30-37.

[37]R. L. Hsu and A. K. Jain, "Generating Discriminating Cartoon Faces Using Interacting Snakes", *IEEE Transactions on PAMI*, vol. 25, pp. 1388-1398, Nov. 2003.

[38]Pei-Ying Chiang, Wen-Hung Liao, Tsai-Yen Li, AUTOMATIC CARICATURE GENERATION BY ANALYZING FACIAL FEATURES, *Proceeding of 2004 Asia Conference on Computer Vision (ACCV2004)*, Korea, 2004

[39] 边肇祺, 张学工, 模式识别, 北京:清华大学出版社, 2000