

第八讲 人脸识别交互技术

王长波



引子

- 人世间找不两张完全一样的脸！
 - 人脸是人类赖以区分不同人的基本途径
- 谁决定了你的长相？
 - 基因 + 成长环境
 - 双胞胎
 - 夫妻相
- 世间一切尽在脸上！



提 纲

- 人脸识别交互背景
- 人脸识别的基本原理
- 人脸检测典型方法
- 人脸识别典型方法

无处不在的身份验证

- 如何证明你的身份？
 - 设想你被警察拦下...
- 传统方法
 - What you have
 - 身份证，暂住证，结婚证，工作证，银行卡，钥匙...
 - What you know
 - Password, information(去清华)...
 - What you are/how you do
 - 一种新的技术手段——Biometrics



传统方法有什么问题？

- 卡、钥匙丢失
 - 青年公寓的钥匙牌...
- 密码危机
 - 密码遗忘
 - 纽约每天1000人以上忘记密码 ☹
 - 密码被猜中
 - 生日、电话号码、车号、宿舍...
 - Heavy web users have an **average of 21 passwords**; 81% of users select a **common password** and 30% **write their passwords down or store them in a file**. (2002 NTA Monitor Password Survey)
- 损失
 - 2002年，仅美国330万人次的身份盗用；670万信用卡诈骗案



一种新的技术手段

- 生物特征识别技术Biometrics
- 什么是Biometrics?
 - Bio——生物
 - Metrics——测量
- 事实含义
 - 通过人体自身的生理特征(what you are)或行为特征(how you do)进行身份验证的技术



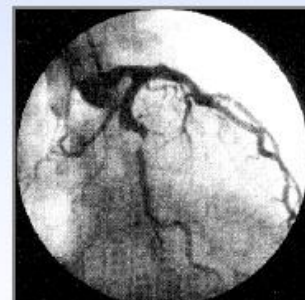
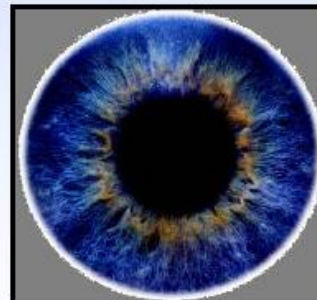
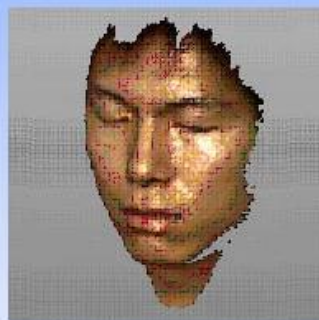
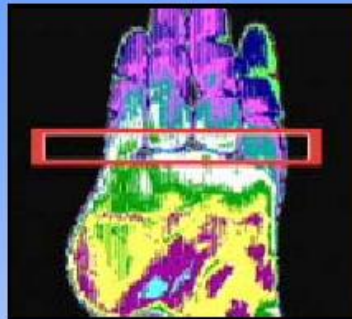
Biometrics

- 生理特征(what you are)
 - Finger, face, palm, hand, foot, iris, vein...
- 行为特征(how you do)
 - 步态, 声音, 敲击键盘, 签名...

人人拥有，人各不同！



生物特征识别技术(Biometrics)



Biometrics孰优孰劣？

- **Universality** (all users possess this biometric)
- **Uniqueness** (varies across users)
- **Permanence** (does not change over time)
- **Collectability** (can be measured quantitatively)
- **Performance** (Low error rates and processing time)
- **Acceptability** (is it acceptable to the users?)
- **Circumvention** (can it be easily spoofed?)



Biometrics比较

Biometrics	Universality	Uniqueness	Permanence	Collectability	Performance	Acceptability	Circumvention
Face	H	L	M	H	L	H	L
Fingerprint	M	H	H	M	H	M	L
Hand Geometry	M	M	M	H	M	M	M
Iris	H	H	H	M	H	L	H
Retinal Scan	H	H	M	L	H	L	H
Signature	L	L	L	H	L	H	L
Voice	M	L	L	M	L	H	L
Facial Thermogram	H	H	L	H	M	H	H

为什么要做人脸识别？

- 多学科领域的挑战性难题
 - 模式识别：最典型、最困难的模式识别问题
 - 人工智能：人类智能的基本体现
 - 计算机视觉：实现人眼的功能
 - 下一代人机交互
 - 让计算机不再“熟视无睹”
 - 让计算机具有人类的情感
- 广泛的应用前景→...
- 人脸识别相比其他生物特征识别的优势→...

应用模式	典型具体应用	特点说明	应用领域
身份识别	出入境管理	过滤敏感人物（间谍、恐怖分子等）	国家安全 公共安全
	嫌疑人照片比对	公安系统用于确定犯罪嫌疑人身份	
	敏感人物智能监控	监控敏感人物（间谍、恐怖分子等）	
	网上追逃	在PDA等移动终端上进行现场比对	
	会议代表身份识别	防止非法人员进入会场带来危险因素	
	关键场所视频监控	如银行大厅，预警可能的不安全因素	
	家政服务机器人	能够识别家庭成员的智能机器人	人机交互
	自动系统登陆	自动识别用户身份，提供个性化界面	
	智能Agent	自动识别用户身份，提供个性化界面	
	真实感虚拟游戏	提供真实感的人物面像，增加交互性	
身份验证	护照、身份证、驾照等各类证件查验	海关、港口、机要部门等查验持证人的身份是否合法	公共安全
	准考证查验	防止替考问题	教育
	机要部门物理门禁	避免钥匙和密码被窃取造成失窃	公共安全
	机要信息系统门禁	避免单纯的密码被窃取造成信息被窃	信息安全
	面像考勤系统	方便，快捷，杜绝代考勤问题	企业应用
	金融用户身份验证	避免单纯的密码被窃取造成财产损失	金融安全
	电子商务身份验证	安全可靠的身份验证手段	金融安全
	智能卡	安全可靠的授权	信息安全
	会议代表身份验证	防止非法人员进入会场带来危险因素	公共安全
	屏幕保护程序	方便快捷的允许合法用户打开屏保	人机交互

与其他生物特征识别的比较

- 生物特征识别：未来的身份验证方法！
 - 生物：指纹、虹膜、人脸、掌纹、手形、视网膜、红外温谱
 - 行为：笔迹、步态、声纹
- 人脸识别的优点
 - 可以隐蔽操作，特别适用于安全问题、罪犯监控与抓逃应用
 - 非接触式采集，没有侵犯性，容易接受
 - 方便、快捷、强大的事后追踪能力
 - 符合我们人类的识别习惯，可交互性强，无需专家评判
- 人脸识别的不足
 - 不同人脸的相似性大
 - 安全性低，识别性能受外界条件的影响非常大



技术挑战

影响人脸图像表观的因素

人与摄像设备的位置关系（距离角度等）

光照环境条件

饰物（眼镜帽子等）

摄像设备

化妆、整容

图像存储质量

精神状态

年龄变化

健康状况



意外损伤

面部毛发(头发, 胡须)



Why Face Recognition is Hard!



“The variations between the images of the same face due to illumination and viewing direction are almost always larger than image variations due to change in face identity.”

-- Moses, Adini, Ullman, ECCV '94



提 纲

- 人脸识别交互背景
- 人脸识别的基本原理
- 人脸检测典型方法
- 人脸识别典型方法

我们人类靠什么识别？

- 天赋的能力还是后天获得的？
- 脸形，面部器官结构
 - 国田由用，目甲风申
- 皮肤和肤色
 - 光滑/粗糙，黝黑/白皙
- 局部特性
 - 黑痣，刀疤，鹰勾鼻子，独眼龙
- 动态特征
 - 酒窝，皱纹

人类视觉识别系统特性简介

- 人脸识别是否是一个特定的过程？
 - 证据：“人脸识别能力缺失症（Prosopagnosia）”患者的存在，患有此症的人可以正常的识别其他物体，甚至可以正确的识别鼻子眼睛和嘴巴等面部器官，但是就是不能认出熟悉的人脸，因此有理由怀疑其人脸识别功能区遭到了破坏。
- 全局特征与局部特征孰轻孰重？
 - 全局特征主要包括人脸的肤色特征（比如白皙、黝黑）、总体轮廓（比如圆脸、鸭蛋脸、方脸、长脸等）、以及面部五官的分布特征（比如，在绘画界就有“国田由用，目甲风申”8种脸型之说），中医也将人脸按照总体结构特征划分为“金木水火土”五行（侧重人脸3D结构和肌肉凸凹情况）
 - 而局部特征则主要指面部五官的特点，比如浓眉毛、丹凤眼、鹰勾鼻、大豁嘴、八字胡须、尖下巴等，以及面部的一些奇异特征，比如黑痣、伤痕、酒窝等等
 - 二者对识别都是必要的，但全局特征一般用来进行粗略的匹配，局部特征则提供更为精细的确认。



局部特征 vs 全局特征

- 明星漫画：夸大了独特之处
 - 问题：How to find these salient features automatically?



平均脸



人脸识别的分类

- 传感器
 - 可见光——模拟眼睛
 - 红外
 - 主动
 - 被动
 - 3D人脸识别

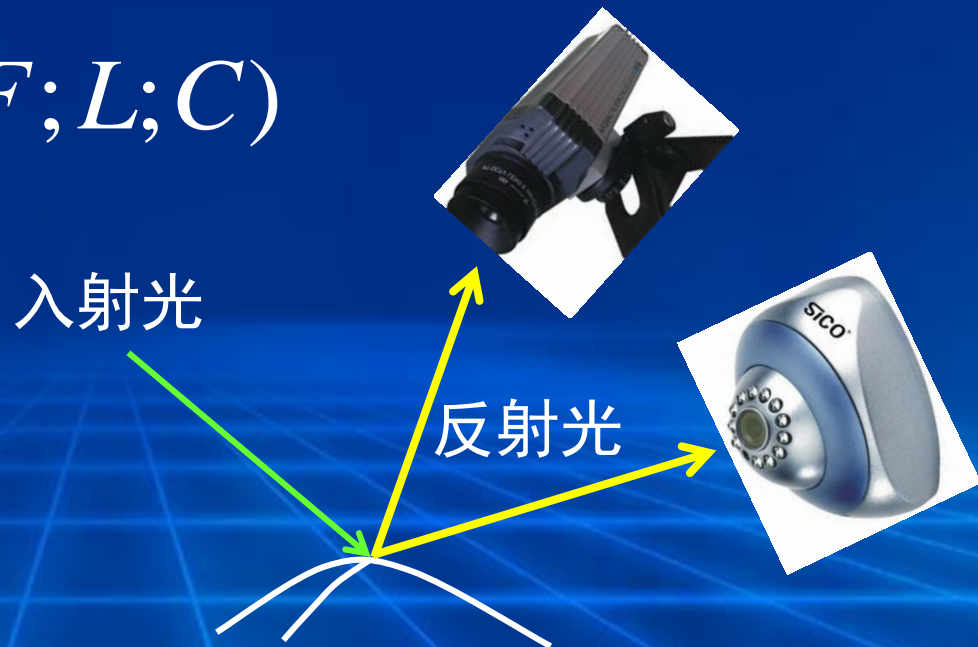


- 输入模式
 - 静态照片
 - 动态视频序列

基于可见光的人脸识别

- 基于可见光人脸图像的身份识别与验证：简称光学人脸图像
 - 是外界光源（包括太阳光、室内人造光源和其他物体表面反射而来的光线）发出的光线照射在人脸上，经人脸表面反射后传输到摄像机传感器的光线强度的度量。

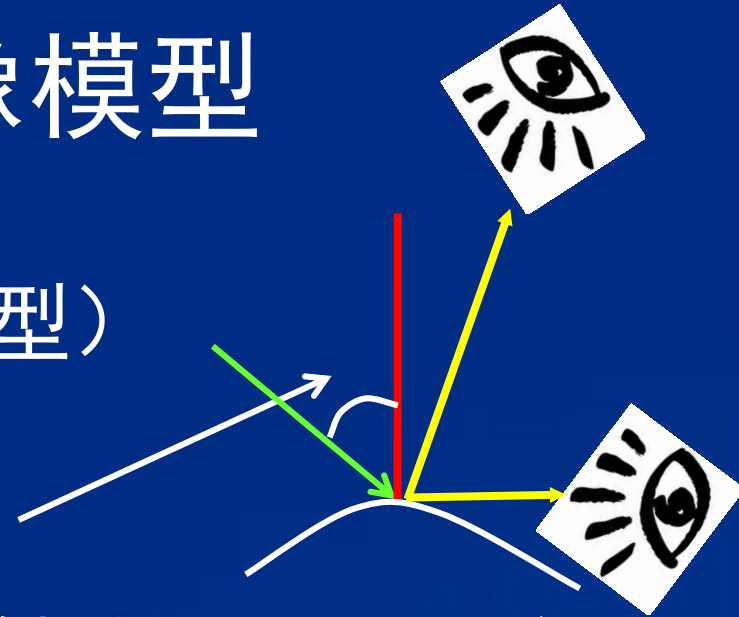
$$I = f(F; L; C)$$



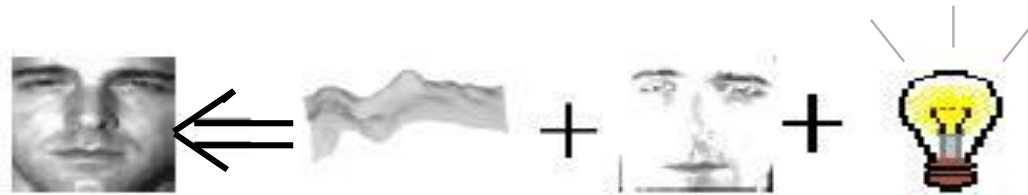
简化的人脸成像模型

- Lambert反射模型（漫反射模型）

$$I = k\rho \cos(\theta) = k\rho \vec{n} \cdot \vec{s}$$



- $\rho, \vec{n}, \vec{s}, \theta$ 分别为物体表面点漫反射系数，法向量方向，光源的方向，二者夹角； k 为入射光强度
- 与视点无关（区别镜面反射）



Face Image = 3D Model + Texture + Illumination

人脸图像的生成要素

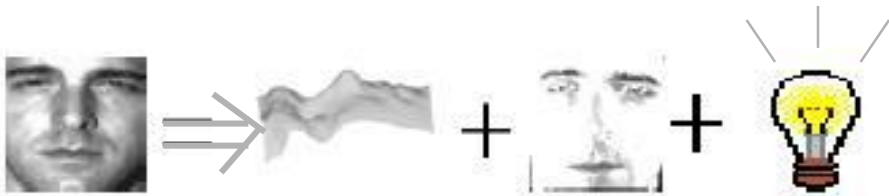
- 人脸图像实际上是三大类关键要素共同作用的结果
 - 人脸内部属性 F
 - 人脸3D形状（表面法向量方向）
 - 包括人脸表面的反射属性（包括反射系数等，通常简称为纹理）
 - 人脸表情、胡须等属性的变化；
 - 外部成像条件 L
 - 包括光源（位置和强度等）
 - 其他物体或者人体其他部件对人脸的遮挡（比如眼镜、帽子、头发）等。
 - 摄像机成像参数 C
 - 包括摄像机位置（视点）、摄像机的焦距、光圈、快门速度等内外部参数

对识别人脸无益的干扰因素！

$$I = f(F; L; C)$$

理想的识别模型

- 从人脸图像中剥离出 **遗憾 ☹ 一个病态问题!**
 - 人脸稳定不变的本质属性（3D形状与表面反射率）
 - 外界条件及其摄像参数变化导致的图像变化
- 然后，从3D形状与表面反射率属性中提取不同人脸的差异信息，馈入到后端的判别分类器中进行识别



The diagram illustrates the decomposition of a face image into three components: a 3D model, texture, and illumination. It shows a grayscale face image on the left, followed by a large right-pointing arrow. To the right of the arrow are three elements: a 3D mesh of a face, a plus sign, a grayscale texture map of a face, another plus sign, and a lit lightbulb icon. Below this visual representation is the equation: Face Image = 3D Model + Texture + Illumination.

$$\text{Face Image} = \text{3D Model} + \text{Texture} + \text{Illumination}$$

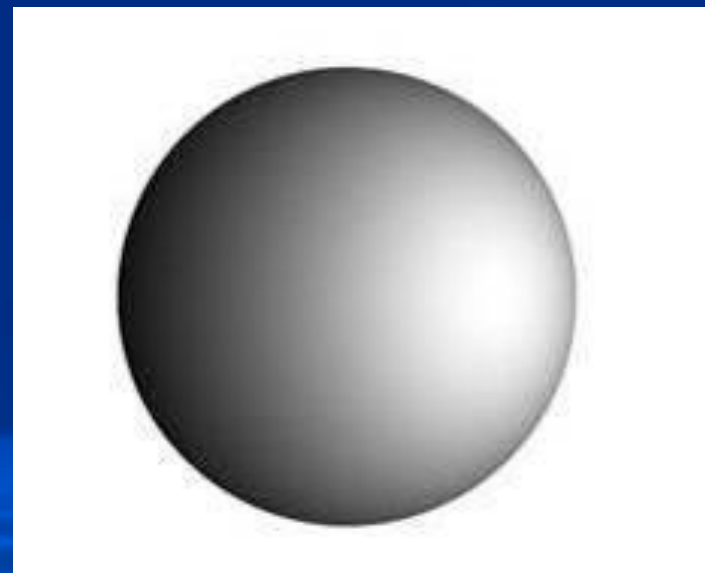
退而求其次...

- 目前的多数系统采用的人脸建模方法仍然停留在图像层面上，并没有显式地分离出3D形状和纹理
- 而是直接通过从“图像”中提取人脸表示特征并进行分类来完成识别
 - 2D 结构信息——基于几何结构的人脸特征
 - 出现了少量利用3D信息进行识别的方法：Morphable models
 - 2D 图像灰度数据统计特征——如模板匹配, Eigenface, Fisherface
 - 2D 图像变换特征——如Gabor, DCT, FFT...
 - 2D 图像的低维子空间分布
 - 线性：子空间，光照锥
 - 非线性：Kernel学习，流形等

主流：Appearance-based methods

从“图像”中识别的道理何在？

- “图像”中包含什么？
 - 亮度变化(影调, shading)、阴影等反应了3D形状
 - Albedo(ρ): 表面反射率
 - 表面反射率 \leftrightarrow 表面皮肤材质
 - 特别是五官区域

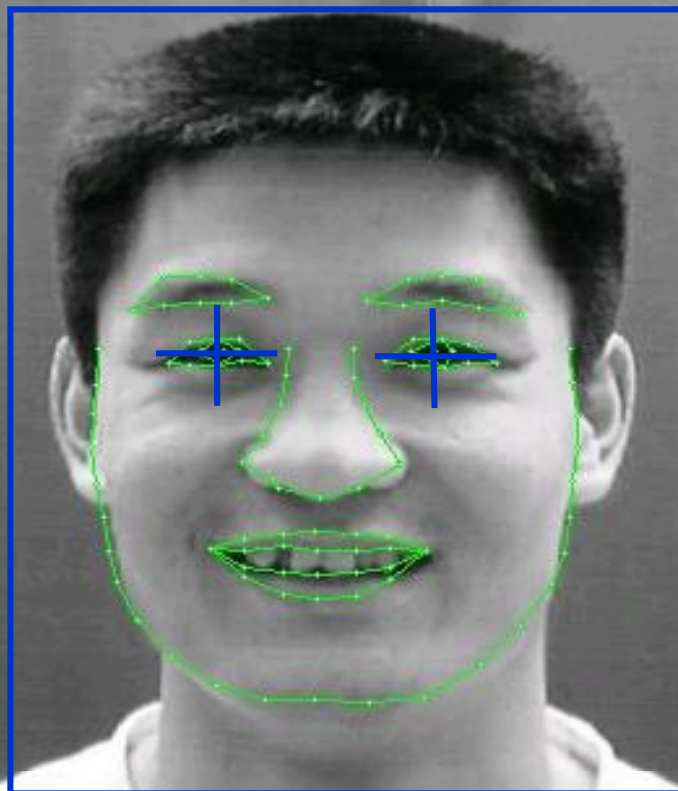


问题分解

人脸检测

面部特征定位

人脸识别/确认



$$X = (x_1, x_2, \dots, x_n)$$



国内外研究现状——研究机构

- 国外研究机构情况
 - 以美欧为主，各知名大学、研究所、企业研究院均设立了与人脸识别相关的研究组
 - 大学：CMU, MIT, UMD, USC, Michigan State University, UCLA, University of Manchester, University of Surrey...
 - 评测：FERET、MIT、PIE、Yale等
- 国内研究机构简况
 - 大学：清华大学，哈工学，中山大学，南京理工，南大，上海交大
 - 研究所：计算所，自动化所等
 - 公司：腾讯优图、商汤科技、旷视科技



人脸识别技术的研究水平

- 在比较良好的环境条件情况下，对1000人左右基本正面人脸进行识别的性能：
 - 首选识别率：99.5%以上
 - 等错误率：2%以下
- 在环境比较糟糕的情况下，对基本正面人脸进行识别的性能：
 - 首选识别率：90%以下
 - 等错误率：10%以上

人脸图像数据库

- 人脸库

- FERET人脸库，14,051张，多数人是西方人
- MIT人脸数据库，16位志愿者，2,592张
- Yale人脸数据库，15位志愿者的165张图片
- PIE人脸数据库，包含68位志愿者的41,368张
- ORL人脸数据库，剑桥大学,包含40人共400张面部图像
- BANCA人脸数据库，欧洲BANCA计划，208人
- XM2VTS人脸数据库，包含295人
- . . .

人脸识别系统

- 人脸识别的技术发展方向

- 结合三维信息：二维和三维信息融合使特征更加鲁棒
- 多特征融合：单一特征难以应对复杂光照和姿态变化
- 大规模人脸比对：面向海量数据的人脸比对与搜索
- 深度学习：在大数据条件下发挥深度神经网络的学习能力

- 国际上比较著名的人脸识别系统

- Cognitec AG——FaceVACS
- Identix (LFA) —— Facelt
- Neven Vision

提 纲

- 人脸识别交互背景
- 人脸识别的基本原理
- 人脸检测典型方法
- 人脸识别典型方法

人脸检测方法

- 基于规则/知识方法
 - 人脸模式的变化满足一定的规律，所以可以归纳描述人脸特征的规则，如灰度分布、比例关系、纹理信息等
- 基于模板的方法
 - 固定模板法，可变形模板法
- 基于不变特征的方法，如彩色信息
 - 人脸的肤色在彩色空间中的分布相对比较集中，所以可以用来检测和跟踪人脸。
- 基于外观学习的方法---**目前的主流方法**
 - 将人脸检测视为区分“非人脸样本”与“人脸样本”的模式识别问题，通过对人脸样本集和非人脸样本集的学习产生分类器



1. 基于肤色特征的检测

- 在很多人脸检测和手的跟踪应用中，人的肤色信息已被证明是一种非常有效的特征
- 有很多颜色空间可用来表征肤色，包括RGB, normalized RGB, HSV (or HSI), YIQ, YES, CIE XYZ, 和 CIE LUV 等
- 有许多建立肤色模型的方法
 - 最简单的方法是用Cr, Cb值定义肤色区域，选定阈值[Cr1,Cr2]和[Cb1,Cb2]
 - 一个像素点被归为肤色点，如果它的(Cr, Cb)值落入下面的范围： $Cr1 \leq Cr \leq Cr2$ and $Cb1 \leq Cb \leq Cb2$

多种肤色模型

- J.L.Crowley and F. Berard, “Multi-model tracking of faces for video communications”
 - 将肤色区域的RGB颜色归一化，用其中的(r,g)值的颜色直方图 $h(r,g)$ 获取肤色变量的阈值
- M. H. Yang, N. Ahuja, “Detecting human faces in color images”
 - 认为人脸肤色区域的颜色值呈高斯分布，用高斯分布的均值和方差确定肤色变量的阈值
- T. S. Jebara and A. Pentland, “Parameterized structure from motion fro 3D adaptive feedback tracking of faces”
 - 认为不同的种族和国家的人的肤色分布不同，在颜色直方图上形成多个聚类，可用高斯混合模型来表示

颜色空间

- RGB到“rg”空间

$$r = \frac{R}{R+G+B}, g = \frac{G}{R+G+B}$$

- RGB到YUV(YCrCb)空间，再转化到“FI”空间

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

$$F = \tan^{-1}(|V|/|U|)$$

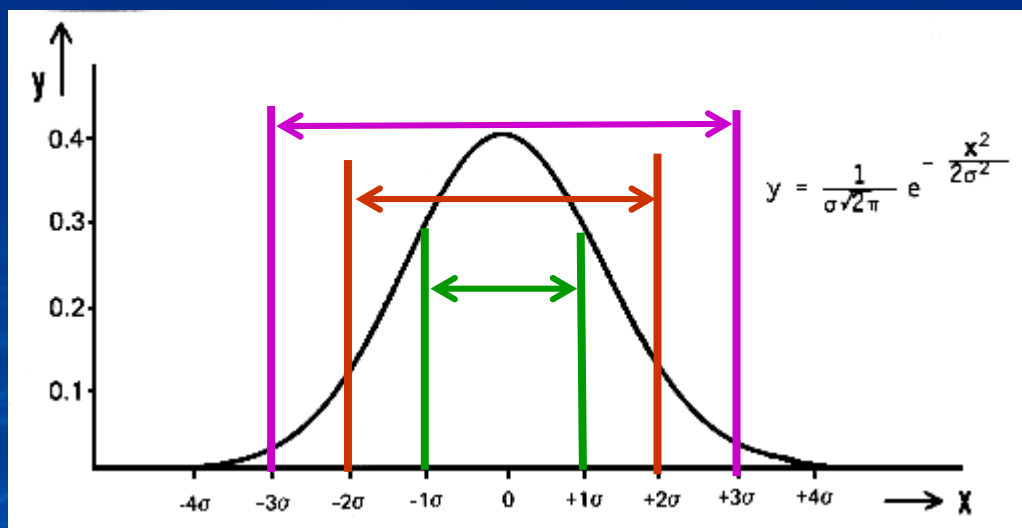
$$I = (0.596, -0.274, -0.322) \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

高斯肤色模型

- 一元正态分布肤色模型（以F颜色特征为例）

$$F \sim N(\mu, \sigma)$$

$$\mu = \frac{1}{N} \sum_{i=1}^N F_i, \quad \sigma^2 = \frac{1}{N-1} \sum_{i=1}^N (F_i - \mu)^2$$



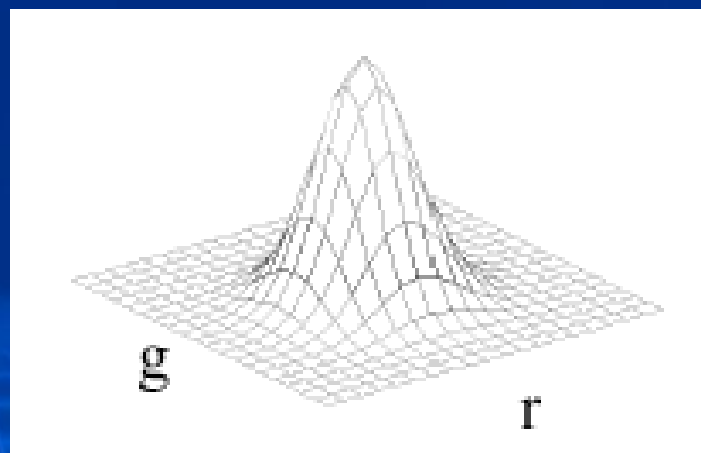
高斯肤色模型

- 二元正态分布肤色模型（以rg颜色特征为例）

$$N(\mu, \Sigma), \text{ where } \mu = (\bar{r}, \bar{g})$$

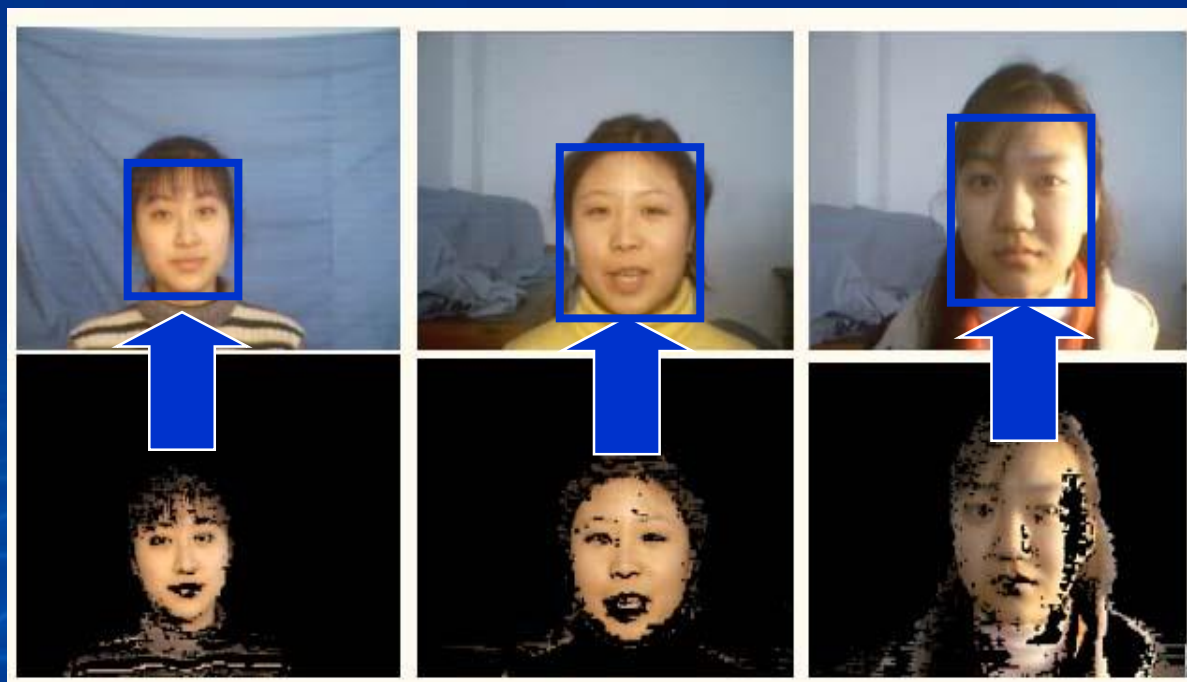
$$\bar{r} = \frac{1}{N} \sum_{i=1}^N r_i, \quad \bar{g} = \frac{1}{N} \sum_{i=1}^N g_i$$

$$\Sigma = \begin{bmatrix} \sigma_{rr} & \sigma_{rg} \\ \sigma_{gr} & \sigma_{gg} \end{bmatrix}.$$



提取肤色区域

- 对检测到的肤色区域进行分析
 - 接近椭圆形
 - 有部分非肤色区域(五官、头发)

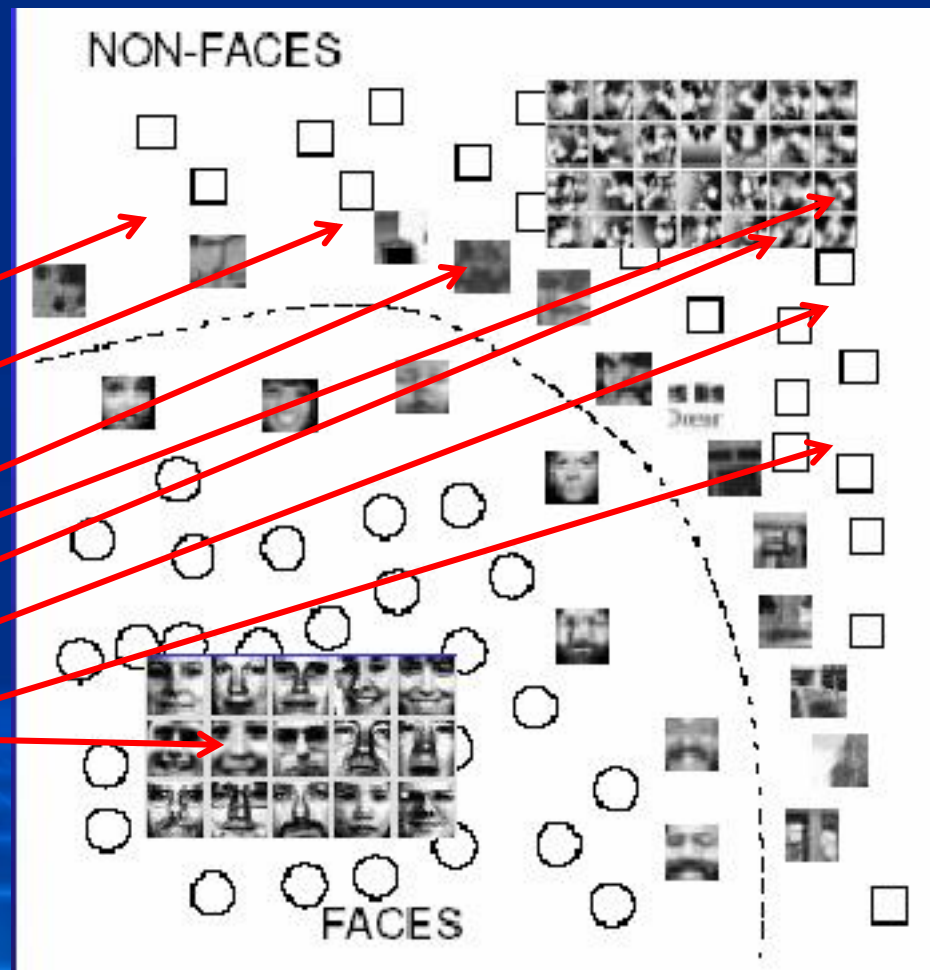
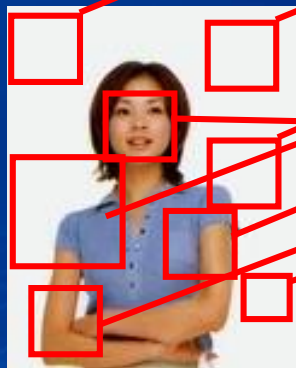


肤色模型的缺点

- 肤色模型难以适应各种环境光照变化
 - 对于背景和前景的光照变化，肤色通常不稳定
 - 单纯的肤色信息对于人脸检测通常是不充分的
- 如何适应在不同光照下的人脸跟踪
 - 如果环境光照有变化，原有的肤色模型可能不再适用，如何建立一个自适应的肤色模型需要进一步的研究

2. 基于Appearance的方法

- ANN
- SVM
- Naïve Bayes Classifier
- AdaBoost



遍历所有可能的“矩形窗口”，判断每个小窗口是否人脸？

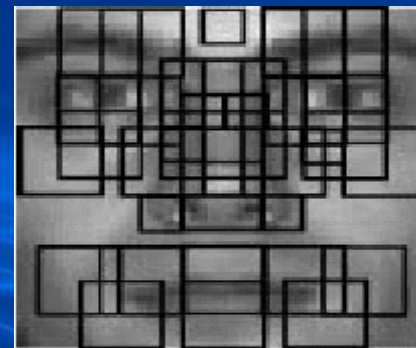
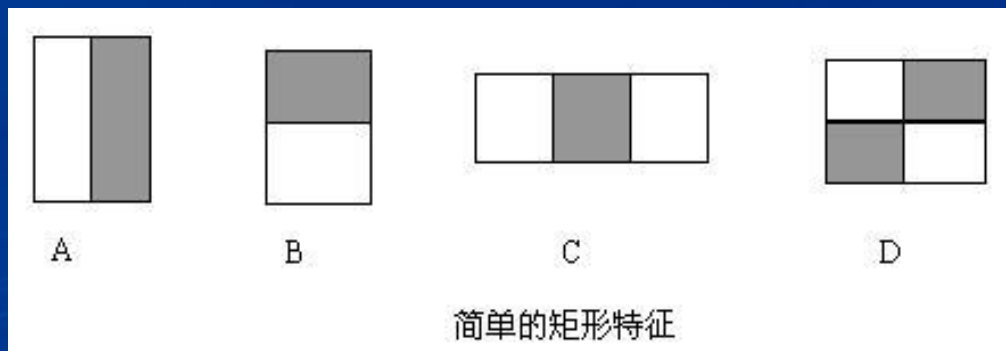


AdaBoost算法简介

- 在2001年的ICCV上，当时在Compaq的研究员Paul Viola和Michael J. Jones发表的文章介绍一个实时人脸检测系统
 - 图像大小为384x288时，其速度是平均每秒15帧
 - 第一个准实时的(准正面)人脸检测系统
 - 可以较容易的扩展到多姿态人脸检测
- 系统在技术上的三个贡献：
 - 1.用简单的Haar-like矩形特征作特征，可快速计算
 - 2.基于AdaBoost的分类器设计
 - 3.采用了Cascade(分级分类器)技术提高检测速度

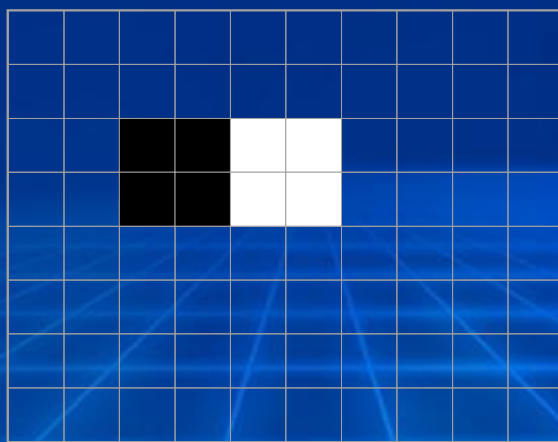
人脸的特征表示方法

- 矩形特征（Haar-like特征）
 - 矩形特征的值是所有白色矩形中点的亮度值的和减去所有灰色矩形中点的亮度值的和，所得到的差
 - 有4种类型的矩形特征



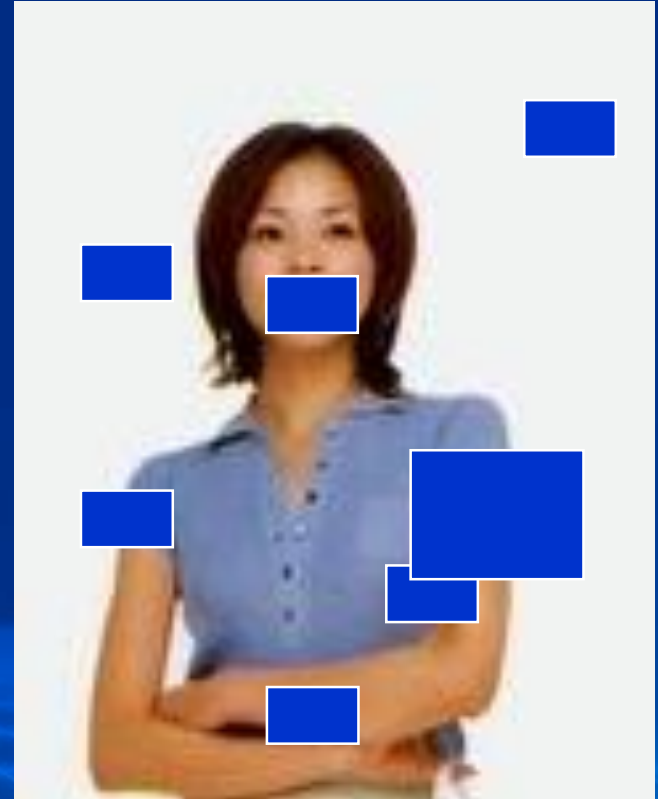
Haar-like特征的表示

- 具体特征可以用一个五元组表示
 $r(x,y,w,h,style)$
 - 比如： $r(2,2,4,2,A)$ 表示下面的特征
 - 特征值即为白色四个像素与黑色四个像素的差值



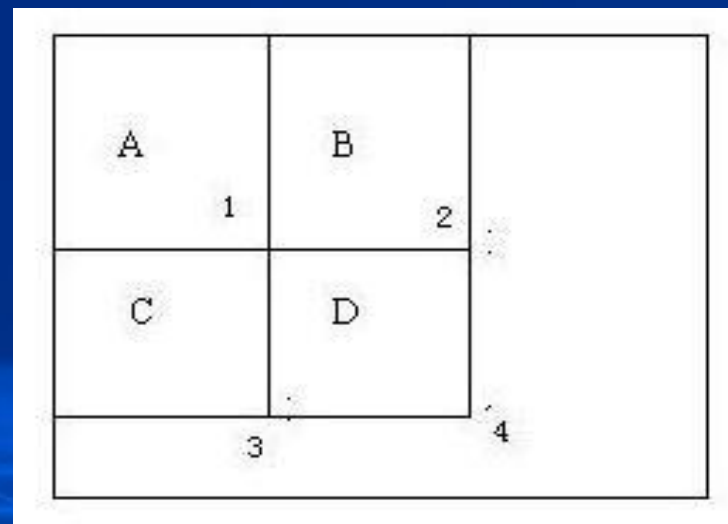
问题

- 如何快速计算任意矩形内所有像素的亮度之和？
 - 请大家设计一个算法
 - 要求计算量尽可能的小



Haar-Like特征的快速计算

- 矩形特征的计算
 - 像素点1的积分值是矩形A中所有点的亮度值的和
 - 像素点2的积分值是A+B
 - 像素点3的积分值是A+C,
 - 像素点4的积分值是A+B+C+D.
- 矩形D内像素积分值:
 - $ii(4) - [ii(2) + ii(3)] + ii(1)$

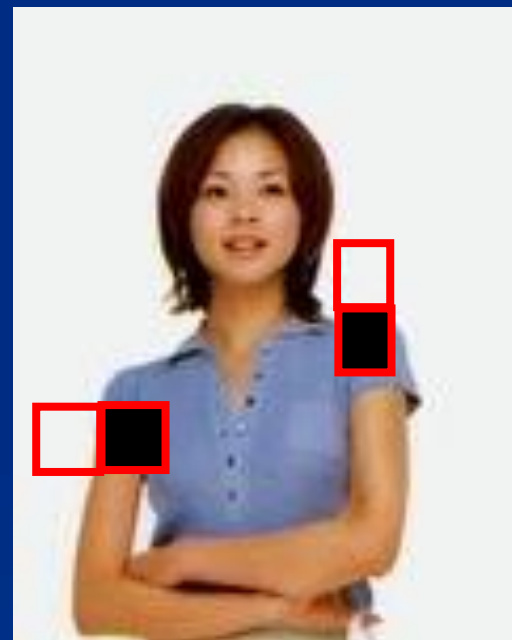




输入图像



积分图像



基于积分图像的
Haar-like特征计算

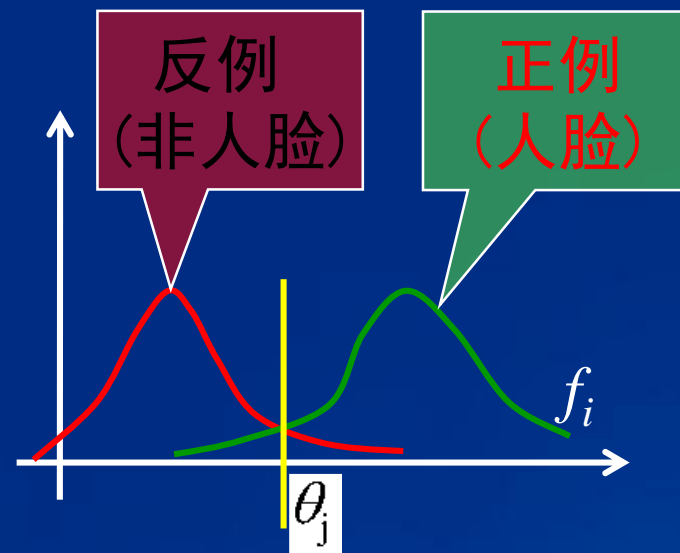
AdaBoost分类器

- AdaBoost分类器
 - Adaboost学习算法是用来提高简单分类算法的性能的
 - 通过对一些弱分类器的组合来形成一个强分类器
- 功能
 - 将分类性能不好的弱分类器提升为分类性能好的强分类器的学习算法
- 思想
 - 学习一系列分类器，在这个序列中每一个分类器对它前一个分类器导致的错误分类样例给予更大的重视

AdaBoost用于人脸模式分类

- 弱分类器

$$h_j(x) = \begin{cases} 1, & \text{if } f_j(x) < \theta_j \\ 0, & \text{otherwise} \end{cases}$$

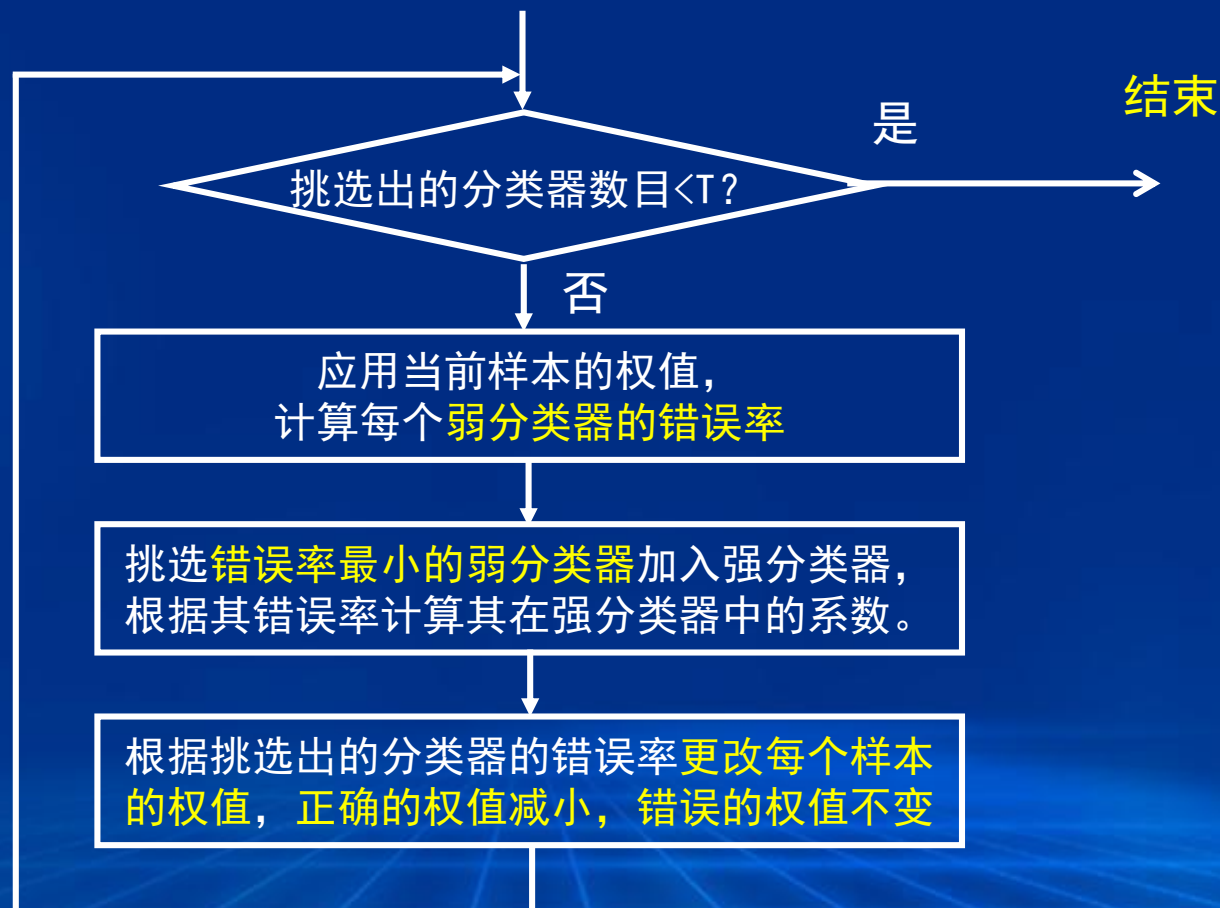


其中， h 表示弱分类器的值， θ_j 表示弱学习算法寻找出的
阈值， $f_j(x)$ 表示特征值， x 表示一个Harr-like特征。

AdaBoost用于人脸模式分类

- 输入
 - 1.训练用人脸和非脸样本
 - 2.指定要挑选出来的弱分类器的数目 T
 - 这也是程序循环的次数
 - 3.利用先验知识初始化权值向量
 - 一般可以平均设置

Adaboost学习算法流程



AdaBoost学习算法

- 学习流程

- For $t=1, \dots, T$

- 1. 归一化权重, 使得 w_t 为一个概率分布: $w_{t,i} = \frac{w_{t,i}}{\sum_{j=1}^n w_{t,j}}$
 - 2. 对每个特征 j , 训练一个弱分类器 h_j 并计算其带权重的错误率

$$\varepsilon_j = \sum_{i=1}^n w_{t,i} |h_j(x_i) - y_i|$$

- 3. 选择误差最小的弱分类器 h_t
 - 4. 更新每个样本的权重 $w_{t,i} = w_{t,i} \beta_t^{1-e_i}$, $\beta_t = \frac{\varepsilon_t}{1 - \varepsilon_t}$
其中: x_i 被正确分类, $e_i=0$, 否则 $e_i=1$

正确分类则权重下降, 否则不变;
错误率越小, Beta也越小

AdaBoost学习算法

- 最终得到的强分类器

$$H(x) = \begin{cases} 1 & \text{if } \sum_{t=1}^T \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t \\ 0 & \text{otherwise} \end{cases}$$
$$\alpha_t = \log \frac{1}{\beta_t}$$

弱分类器错误率越小， β 就越小， α 就越大

$H(x)$ 为+，则为正例（人脸）

$H(x)$ 为-，则为反例（非人脸）

检测系统的构建——遍历所有位置



检测系统的构建——遍历不同大小



如何进一步加速？

- 遍历所有可能的位置、大小，需要做多少次分类？
 - 320x240图像，20x20模板，每次走1个像素，尺度缩小因子0.9，则大约有：
 - $300 \times 220 + 278 \times 196 + 240 \times 174 + 213 \times 155 + 190 \times 137 + 169 \times 122 + 150 \times 108 + 133 \times 95 + 118 \times 83 + 103 \times 73 + 91 \times 64 + 80 \times 55 + \dots = 30\text{多万}$
- 如果再考虑3个角度的旋转，则会超过100万次分类！
- 问题
 - 如何进一步加速？

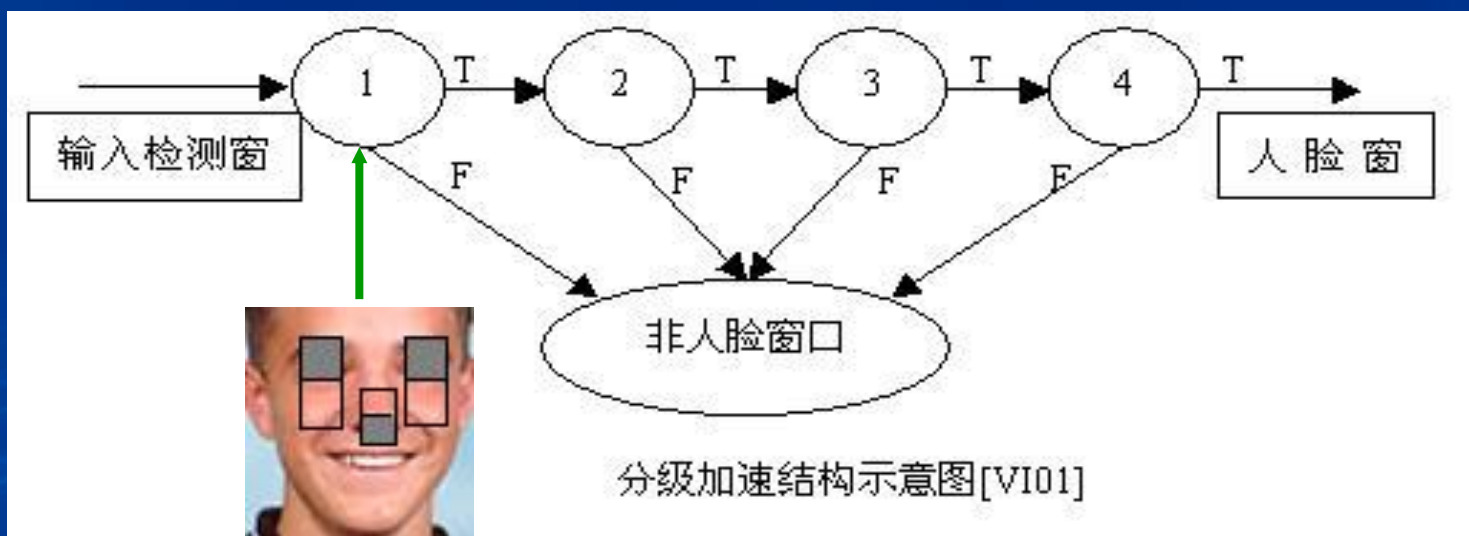
类比一下

- 分析
 - 非人脸模式里面大量与人脸极其不同！只有少量的模式很难与人脸分开！
- 类比筛沙子
 - 输入：巨石，大石头，小石头，沙砾，粗沙，细沙
 - 目标：只留下细沙（人脸）
 - 假设：大孔筛子容易做便宜，小细孔筛子难做昂贵
- 方法1
 - 只用一个筛子，网眼非常小，所有的都要过这个筛子
- 方法2
 - 用一系列筛子，网眼从大到小
 - 非常容易被截住的巨石首先被排除，然后是大石头，再是小石头，沙砾，粗沙，最后是细沙



快速人脸检测

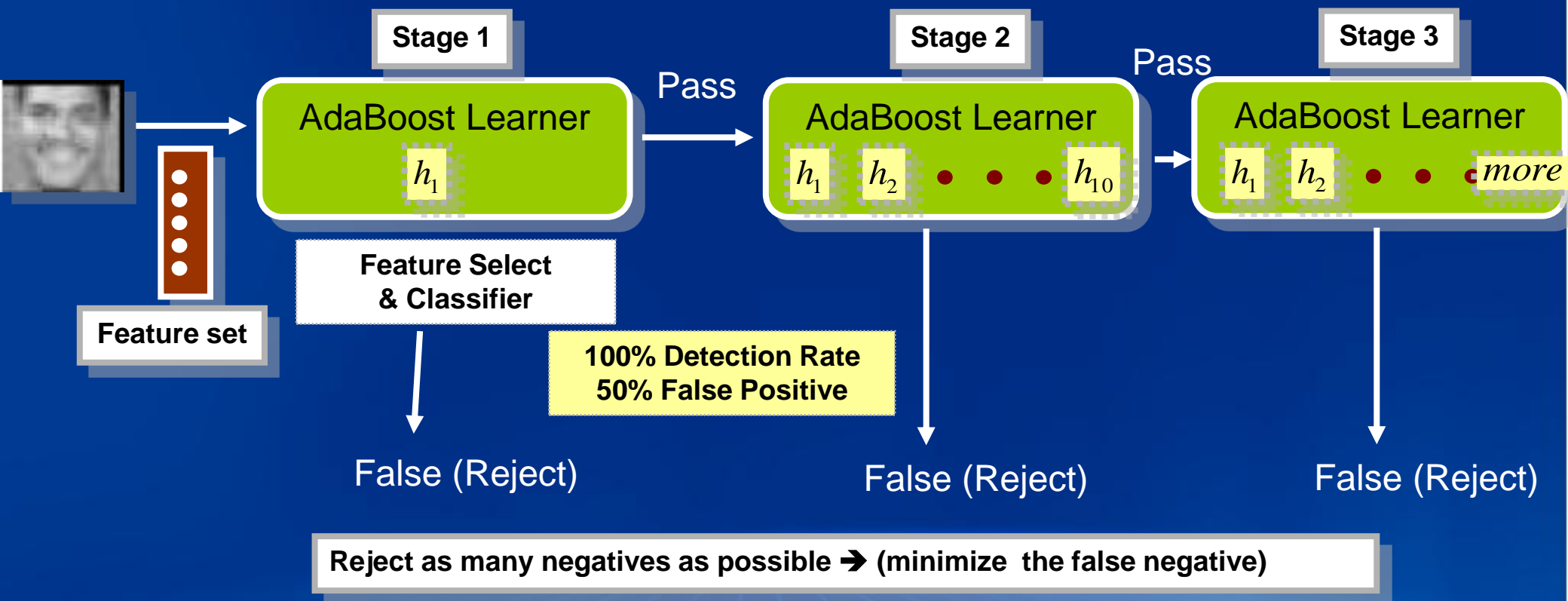
- 基于分级分类器的加速策略
 - 大量候选窗口可以利用非常少量的特征（简单快速的分类器）就可以排除是人脸的可能性！
 - 只有极少数需要大量的特征（更复杂的更慢的分类器来判别是否人脸）



分级分类器的构建

- 采用由粗到细的思想（coarse to fine）
 - 将少量区分性好的特征构成的简单分类器置于前面若干层
 - 效果：放过检测绝大多数人脸的同时，排除大量非脸
 - 后面层包含更多次重要的特征对非脸进行进一步排除
- 训练方法关键思路
 - 每层训练用非脸样本使用前面层分类器误判为人脸的那些样本

检测过程流程图

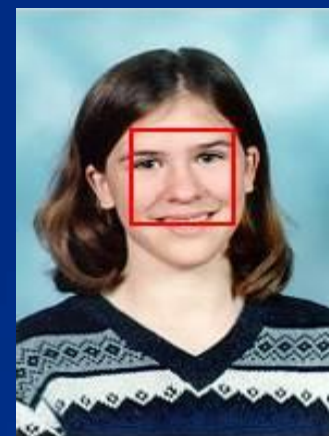


分级分类器的训练算法

1. 首先确定可以接受的每层检测器最大误检率 f , 最小正确检测率 d 然后确定最终检测的目标误检率 F_{target}
2. P 为正例样本集合 N 为反例样本集合
3. $F_0 = 1.0; D_0 = 1.0$
4. $i = 0$
- 1) While $F_i > F_{target}$
 - i) $i = i + 1$
 - ii) $n_i = 0; F_i = F_{i-1}$
 - iii) While $F_i > f \times F_{i-1}$
 - a) $n_i = n_i + 1$
 - b) 使用 P 和 N 训练一个有 n_i 个分类器的特征, 使用 AdaBoost 算法
 - c) 在测试集上测试, 得到 F_i 和 D_i
 - d) 降低这个分类器阈值, 直到这个分级分类器的检测率不小于 $d \times D_{i-1}$ (这个调整也会影响 F_i)
 - iv) 置反例集合 N 为空集
 - v) 如果 $F_i > F_{target}$ 在不包含人脸的样本集合上运行分类器, 把所有的误检为正例的样本加入到反例集合中。



一些检测结果



提 纲

- 人脸识别交互背景
- 人脸识别的基本原理
- 人脸检测典型方法
- 人脸识别典型方法

人脸识别的主要方法

- 几何特征的人脸识别方法
- 特征脸(PCA)的人脸识别方法
- 支持向量机(SVM) 人脸识别方法
- 深度神经网络的人脸识别方法

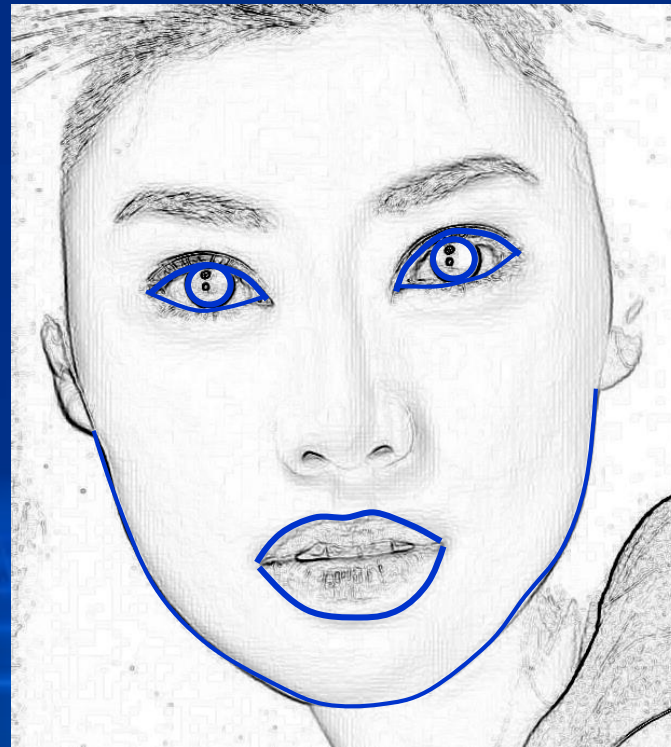
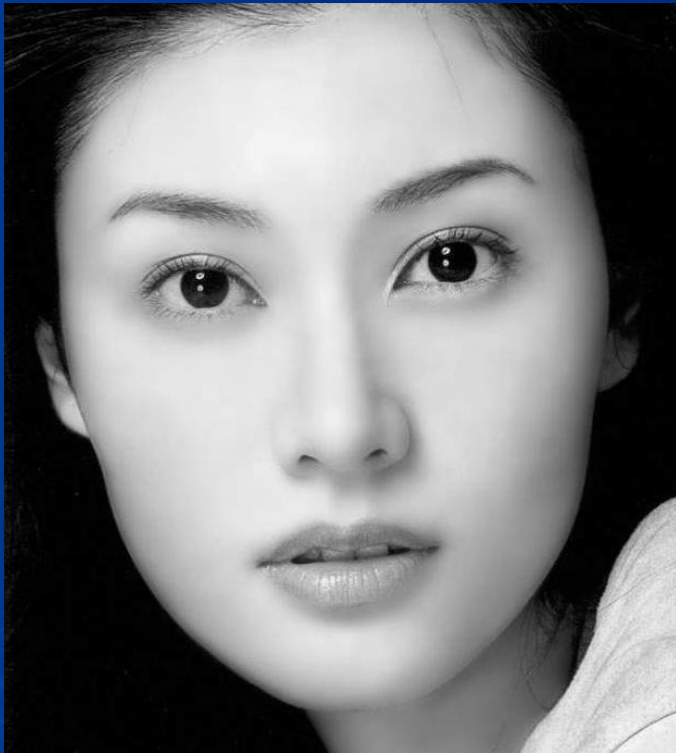


一、基于人脸特征模式匹配的识别



面部特征提取

- 可变形模板匹配 Deformable Template



用可变形模板进行面部特征定位

- 可变形模板概念要点

- 模板 T : 器官形状的参数描述形式(先验)

- 直线、圆、抛物线、四次曲线等

- 能量函数 $F = S(T_p, I)$

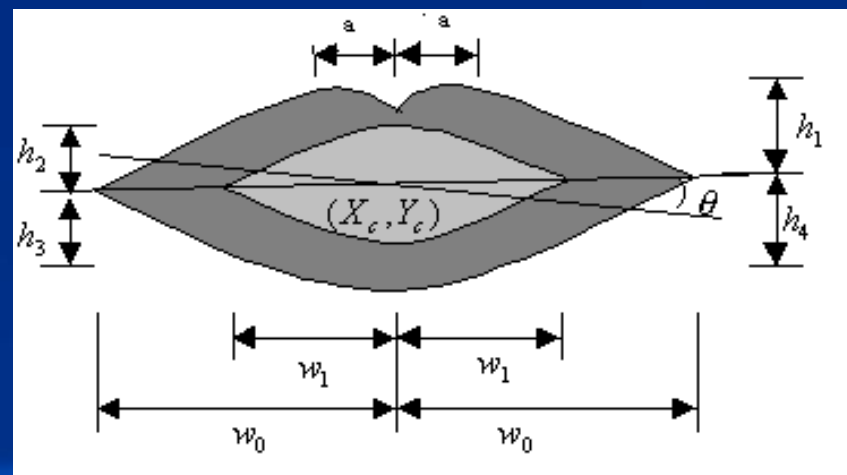
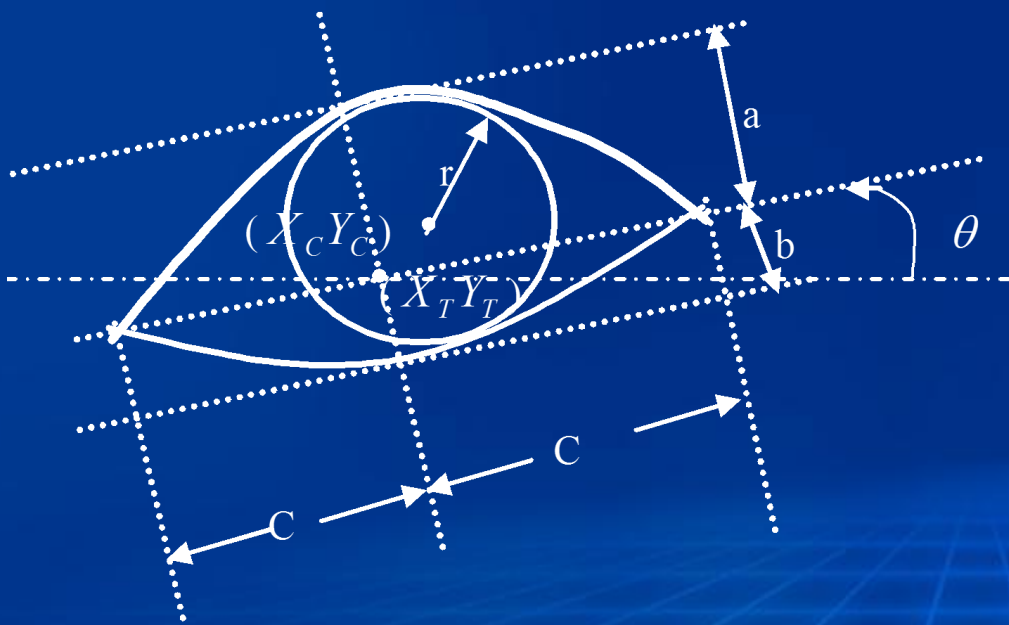
- 参数 p 定义的模板 T_p 与输入图像 I 的匹配程度

- 参数优化问题

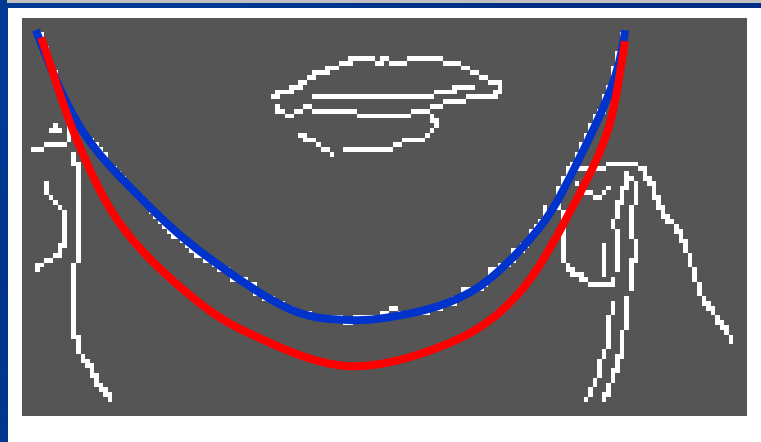
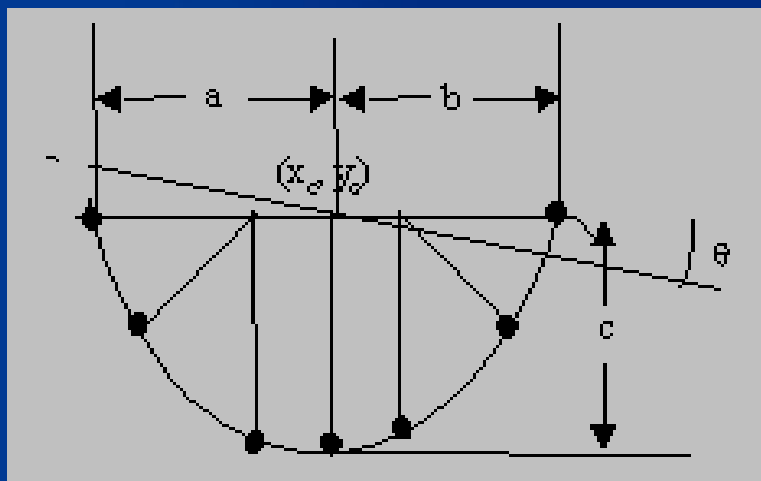
- 多参数非线性函数极小化 $p^* = \arg \min_p (F)$

- 梯度下降法
 - 坐标轮换法
 - 其它优化算法

特征形状模板



可变形模板——以下巴检测为例



上：模板；下：边缘图

参数: $(x_c, y_c), a, b, c, \theta$

左侧抛物线: $L_l(x) = c \times \left(1 - \frac{x^2}{a^2}\right)$

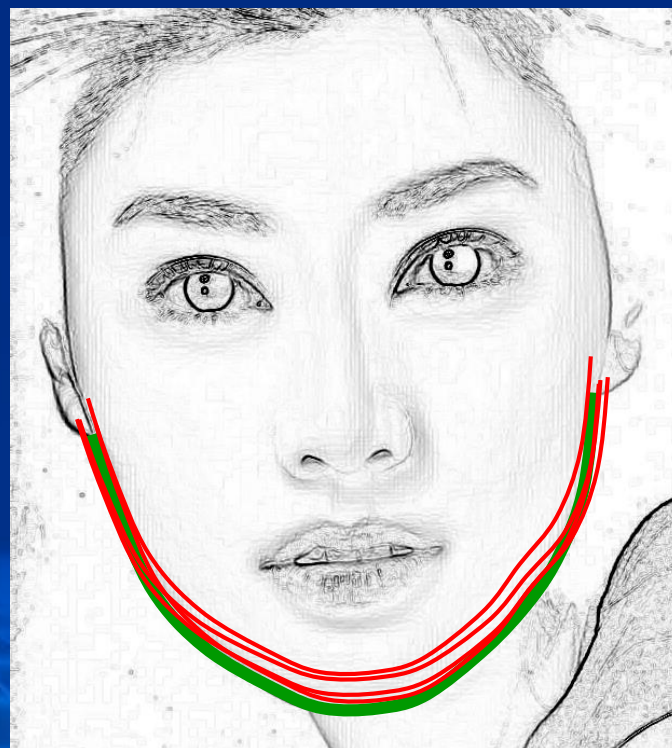
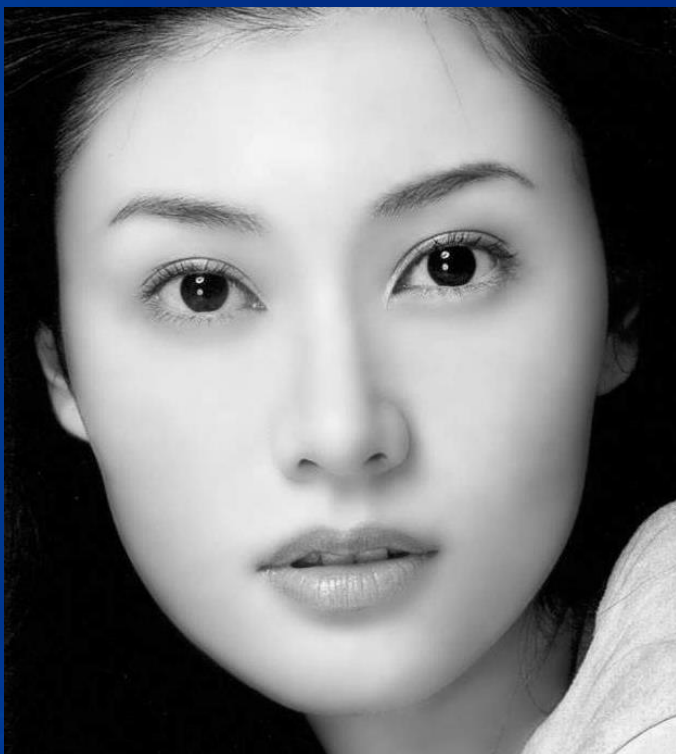
右侧抛物线: $L_r(x) = c \times \left(1 - \frac{x^2}{b^2}\right)$

能量函数: $E = w_l E_l + w_r E_r$

其中: $E_l = -\sqrt{\frac{1}{m} \sum_{(x,y) \in L_l} e^2(x,y)}$

$E_r = -\sqrt{\frac{1}{m} \sum_{(x,y) \in L_r} e^2(x,y)}$

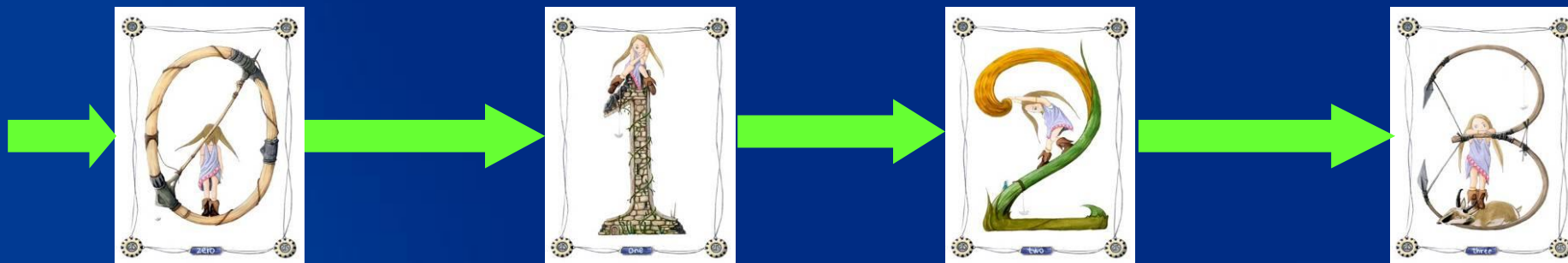
搜索流程



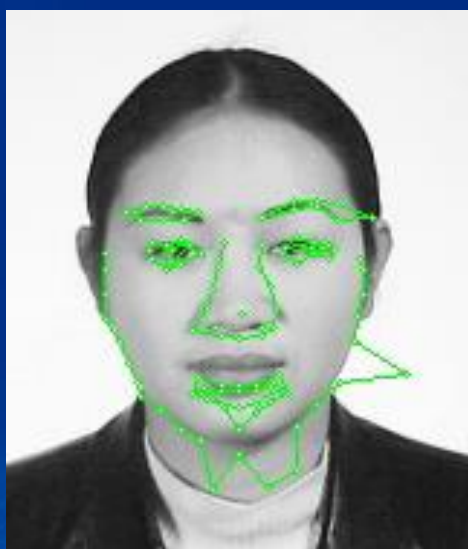


检测结果

ASM: 搜索过程图示



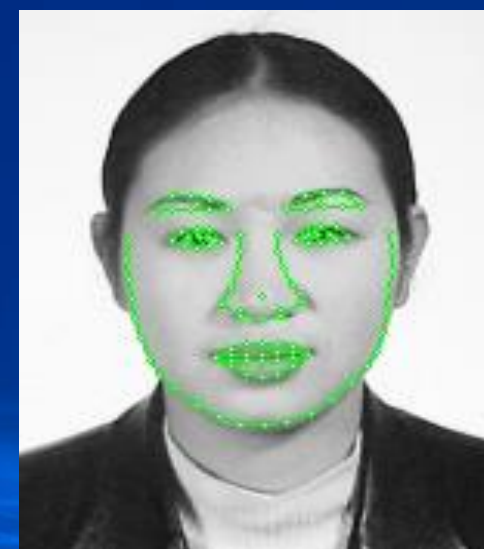
初始化(平均)



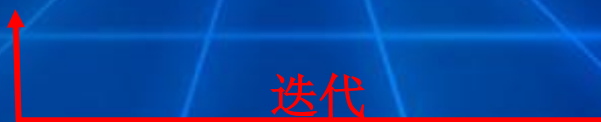
特征点局部纹理模型匹配



统计形状模型约束(PCA)



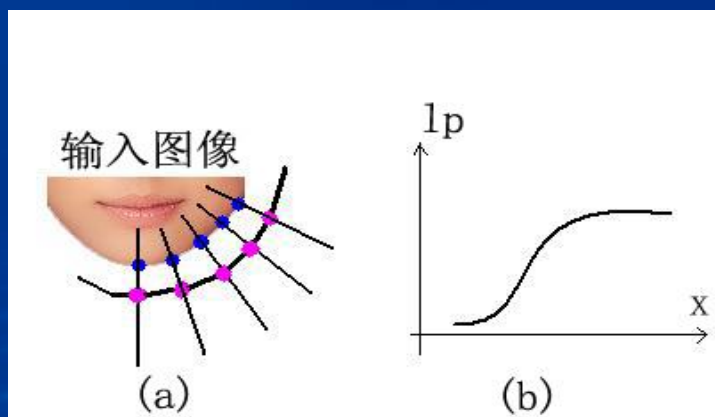
最终结果



迭代

特征点局部纹理模型匹配

- 对每个特征点，在其法线邻域内搜索，寻找局部纹理模型的最佳匹配点
 - 局部纹理模型：法线方向纹理梯度，称为Profile
 - 每个特征点的Profile模型都在训练时建立统计模型
 - 匹配方法
 - 在法线上逐点运算，选择马氏距离最小的候选点



每个特征点都是一个在法线上移动的小机器人，以发现与自己**Profile**特征最佳匹配的候选点！



基于模板匹配的识别

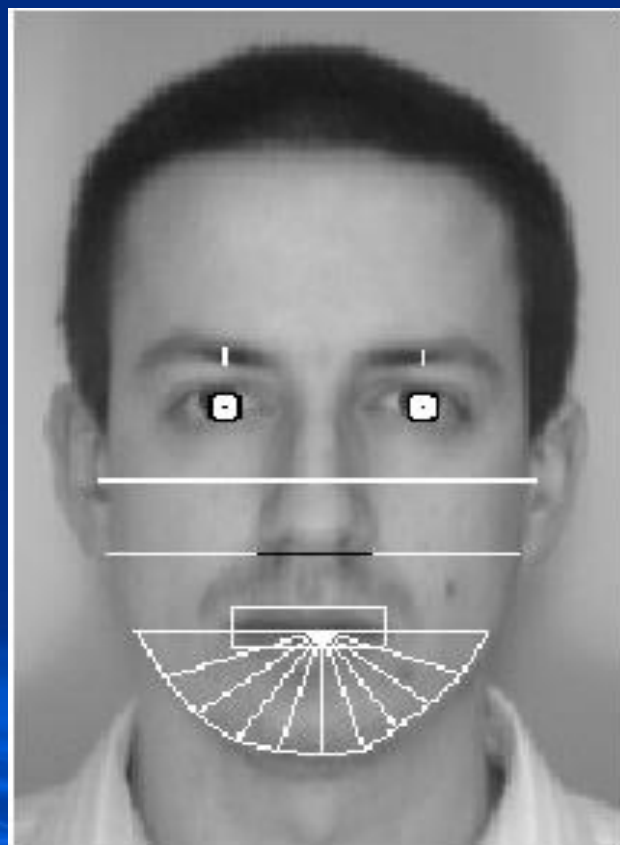
- 建模：用面部关键特征的相对位置、大小、形状、面积等参数来描述人脸
- 人脸图像 $f \rightarrow$ 特征向量 v

$$v = (x_1, x_2, \dots, x_n)$$

- 对所有已知人脸提取同样描述的几何特征

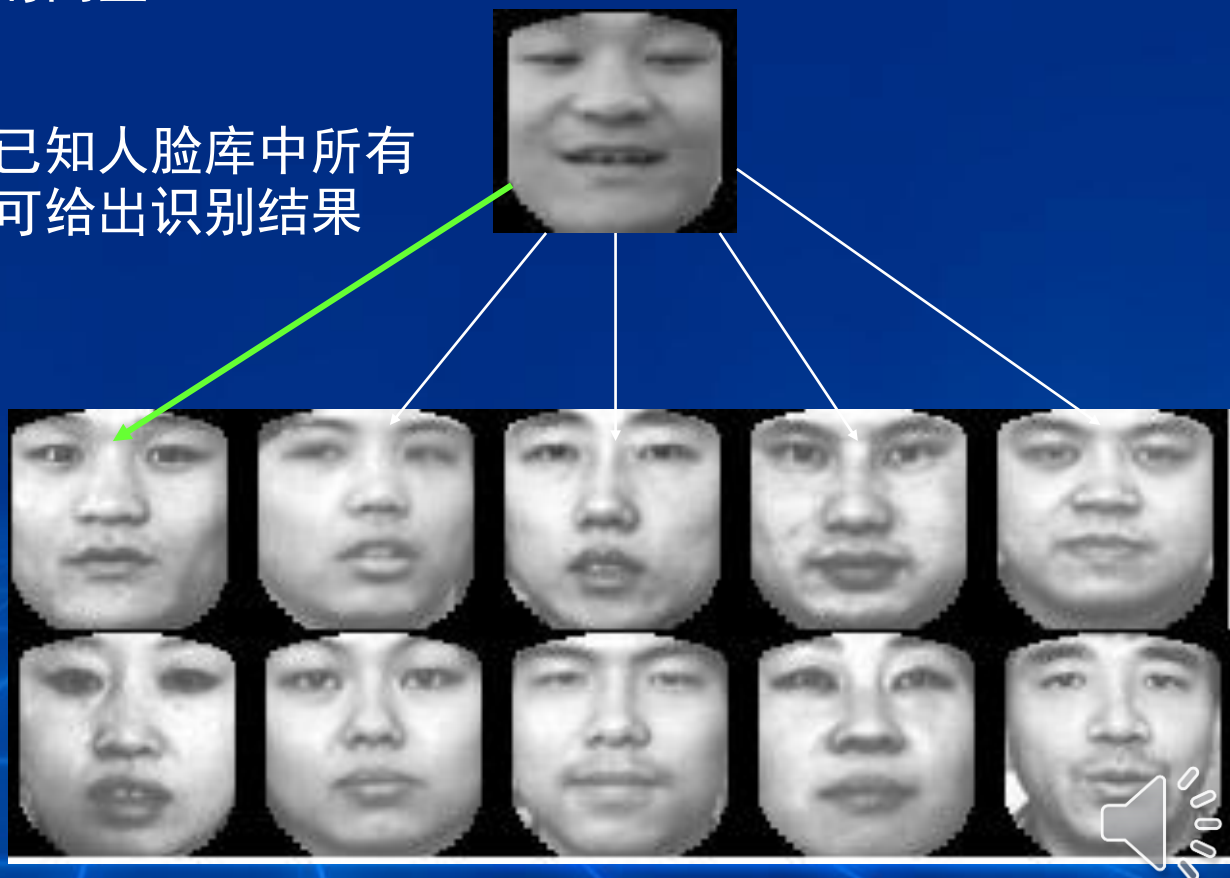
$$D = \{v_1, v_2, \dots, v_p\}$$

- 待识别的人脸 f 提取的几何特征为 v_f
- 计算 v_f 与 D 中所有 v_i 的相似度 $s(v_f, v_i)$ (比如欧式距离、 $\cosine(\cdot)$ 等)，进行排序
- 根据相似度最大的已知人脸的身份即可判断待识别人脸的身份信息



基于模板匹配的识别

- 建模
 - 2D 灰度矩阵，按行向量化为1D向量
 - 所有图像均表示为这样的向量
- 识别
 - 计算输入图像的向量与已知人脸库中所有向量的相似度，排序即可给出识别结果



二、基于特征脸的识别方法



特征脸 (Eigenface)

- PCA
 - Principal Component Analysis
- 也称为Hotteling变换或者Karhunen-Loeve变换 (KLT) .
- 寻求一个正交的坐标系统，使得不同坐标轴之间的相关性最小！




Eigenface人脸识别方法

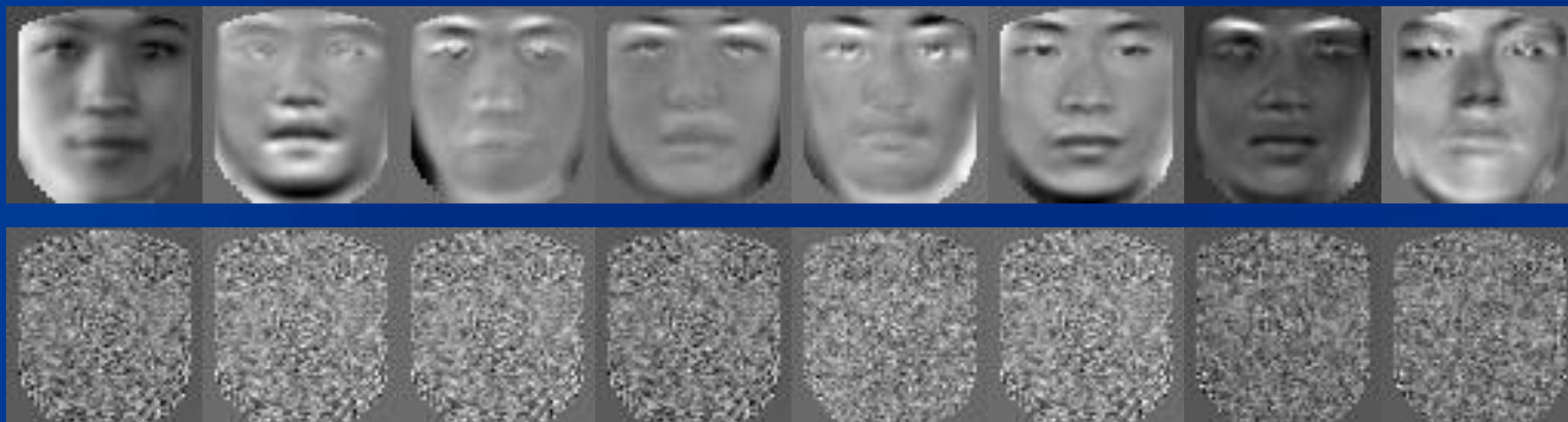
- x 为输入图像
- y 作为提取的特征
- $S(x_1, x_2) = S(y_1, y_2)$
 - 可以采用欧式距离
 - 也可以采用Cosine

$$y = W^T (x - \bar{x})$$

$$x = \bar{x} + Wy$$

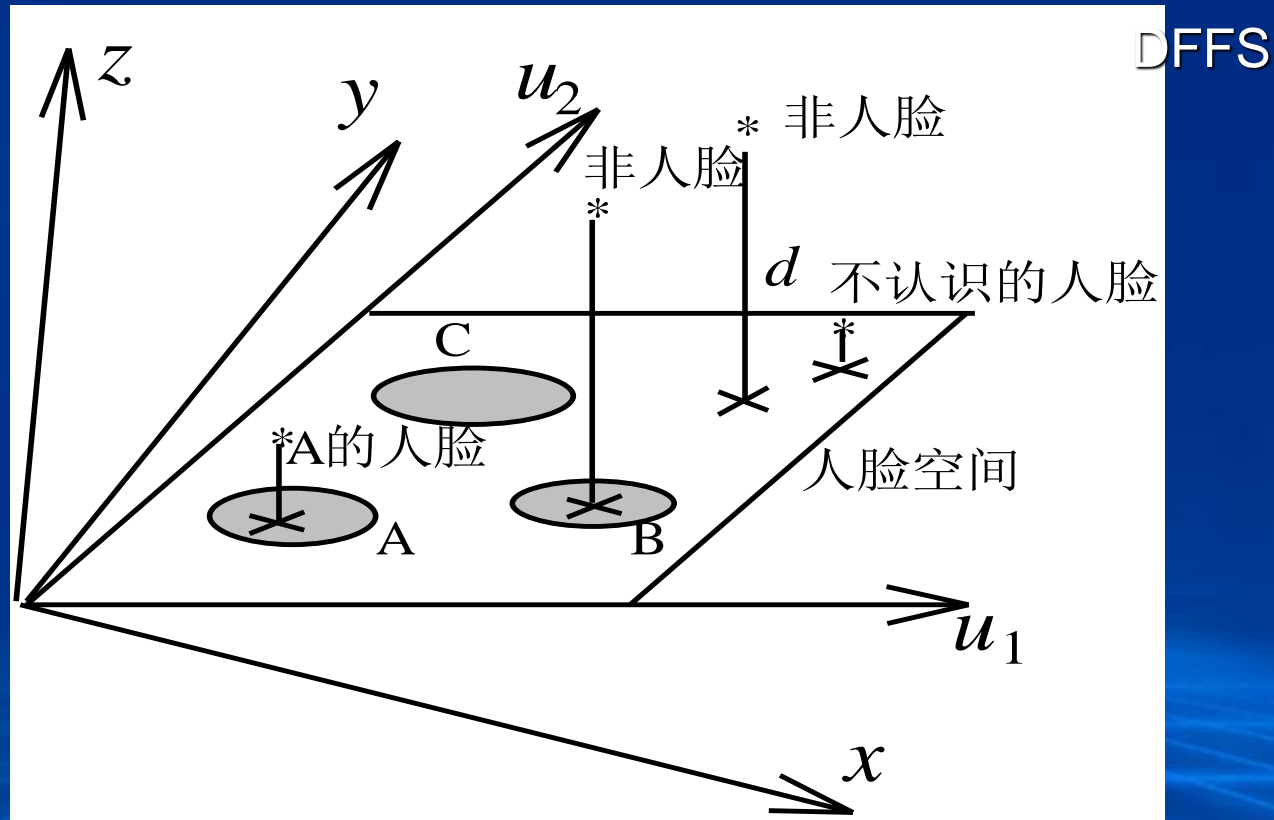

$$x \approx \bar{x} + y_1 * \bar{x} + \dots + y_m * \bar{x} + \delta(I)$$

Eigenvectors of the Covariance matrix



- 可视化的“特征脸”
 - Leading 8D: valid signal
 - Last 8D: Difference and Noise

EigenFace-based face recognition method



基本实现步骤

步骤一：获取包含 M 张人脸图像的集合 S 。



步骤二：在获取到人脸向量集合 S 后，计算得到平均图像 Ψ

步骤三：计算每张图像和平均图像的差值 Φ

步骤四：找到 M 个正交的单位向量 u_n



步骤五：识别人脸。两者通过特征脸权重来表示，寻找欧氏距离最小。



关于识别的好坏问题

- 关于特征提取与分类器
 - **理想的特征提取**过程可使得后续的分类器设计成为小菜一碟，**万能的分类器**似乎也应该使得设计者不必刻意设计特征提取过程！
 - 没有万能的分类器！
 - 甚至没有所谓最好的分类器，不同的分类器有不同的适用范围！
 - 但对具体问题，可能会有最适合的特征表示方法，而不同的特征表示也需要采用不同的分类器

识别算法没有理论上的好坏，只有识别率的差别！

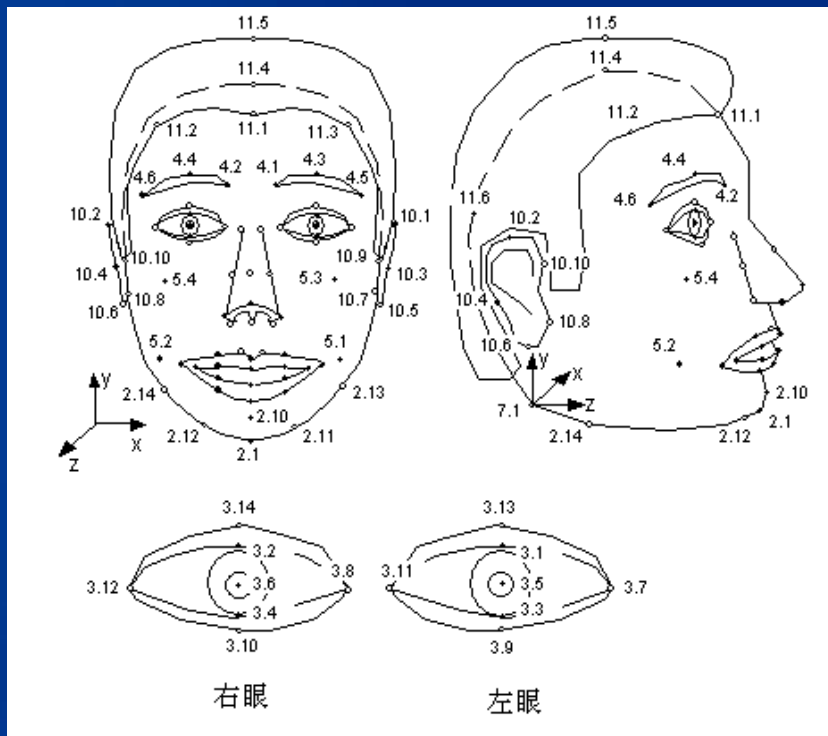


应用案例



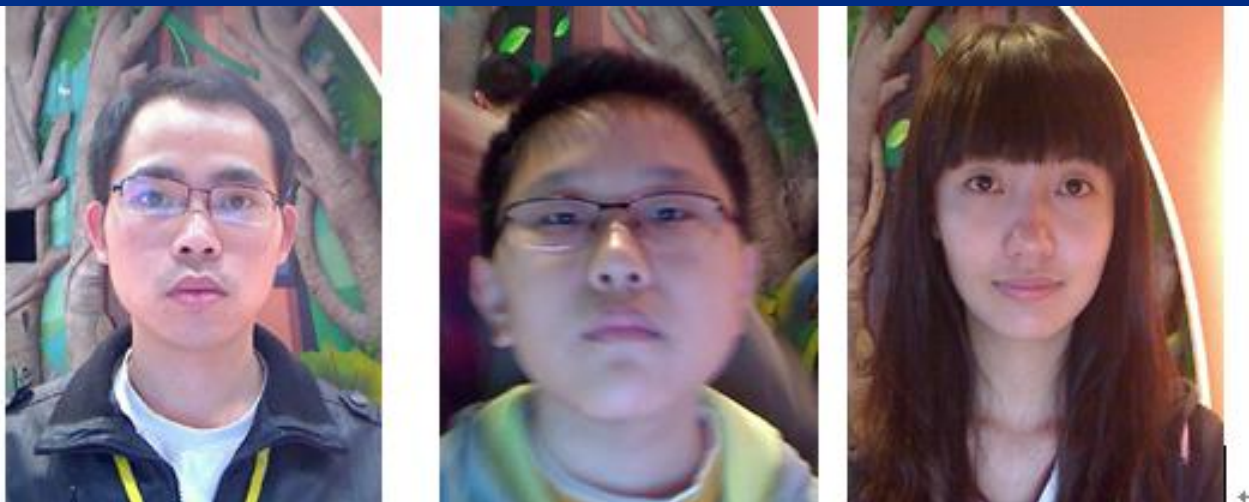
完美脸型虚拟交互展示项目（中国科技馆）

应用案例



完美脸型虚拟交互展示项目（中国科技馆）

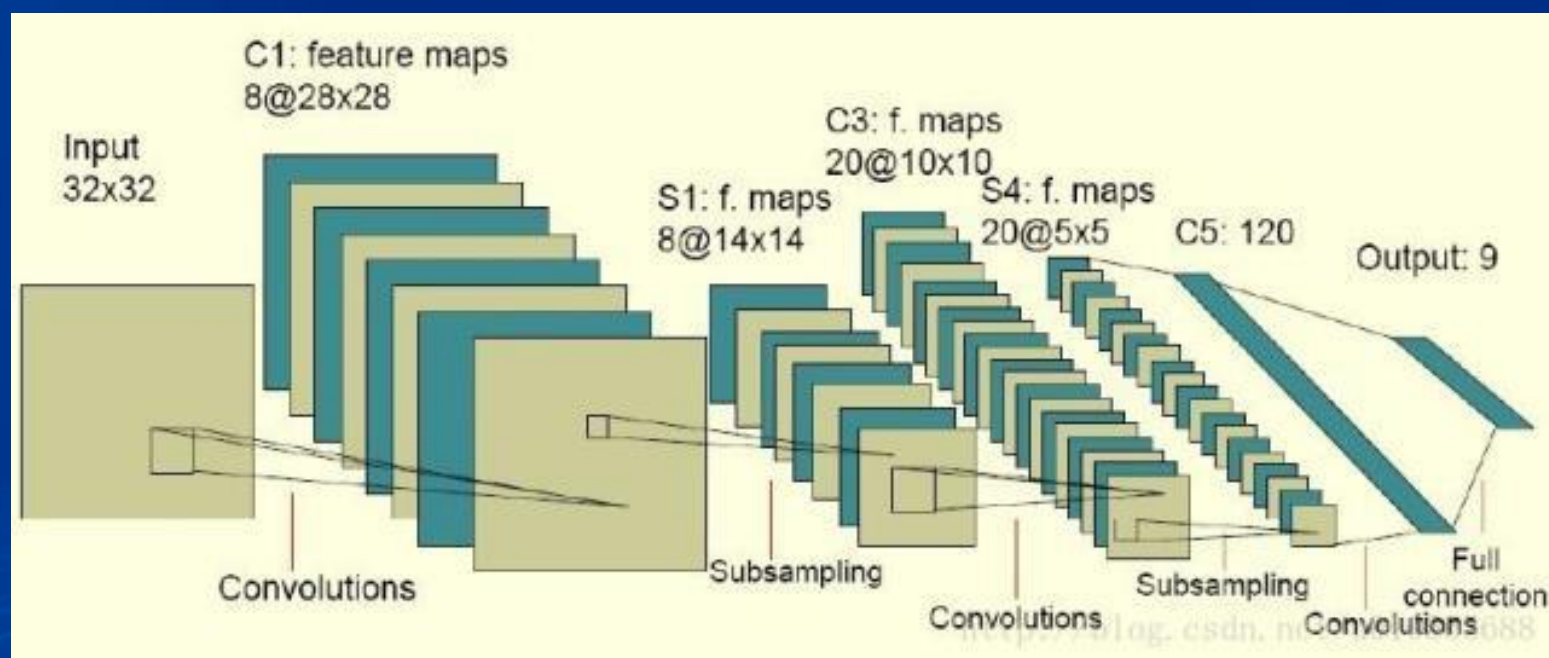
应用案例



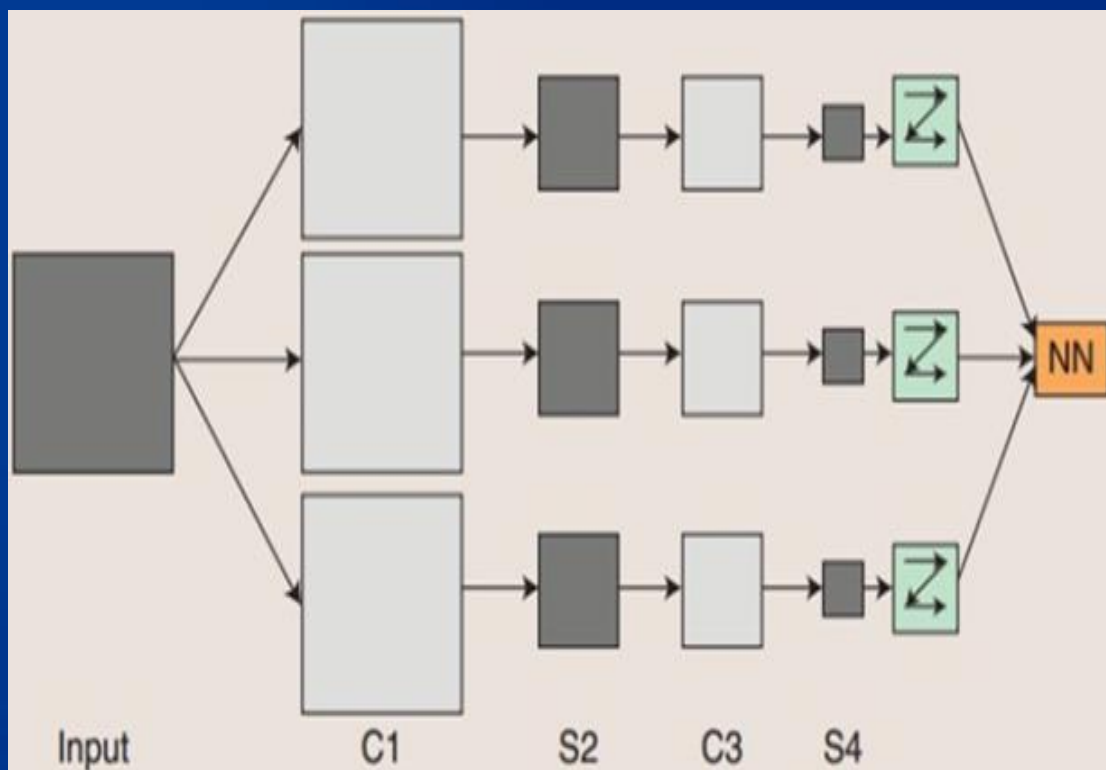
(a) 采集的原始图像 ↵



三、基于深度学习的人脸识别



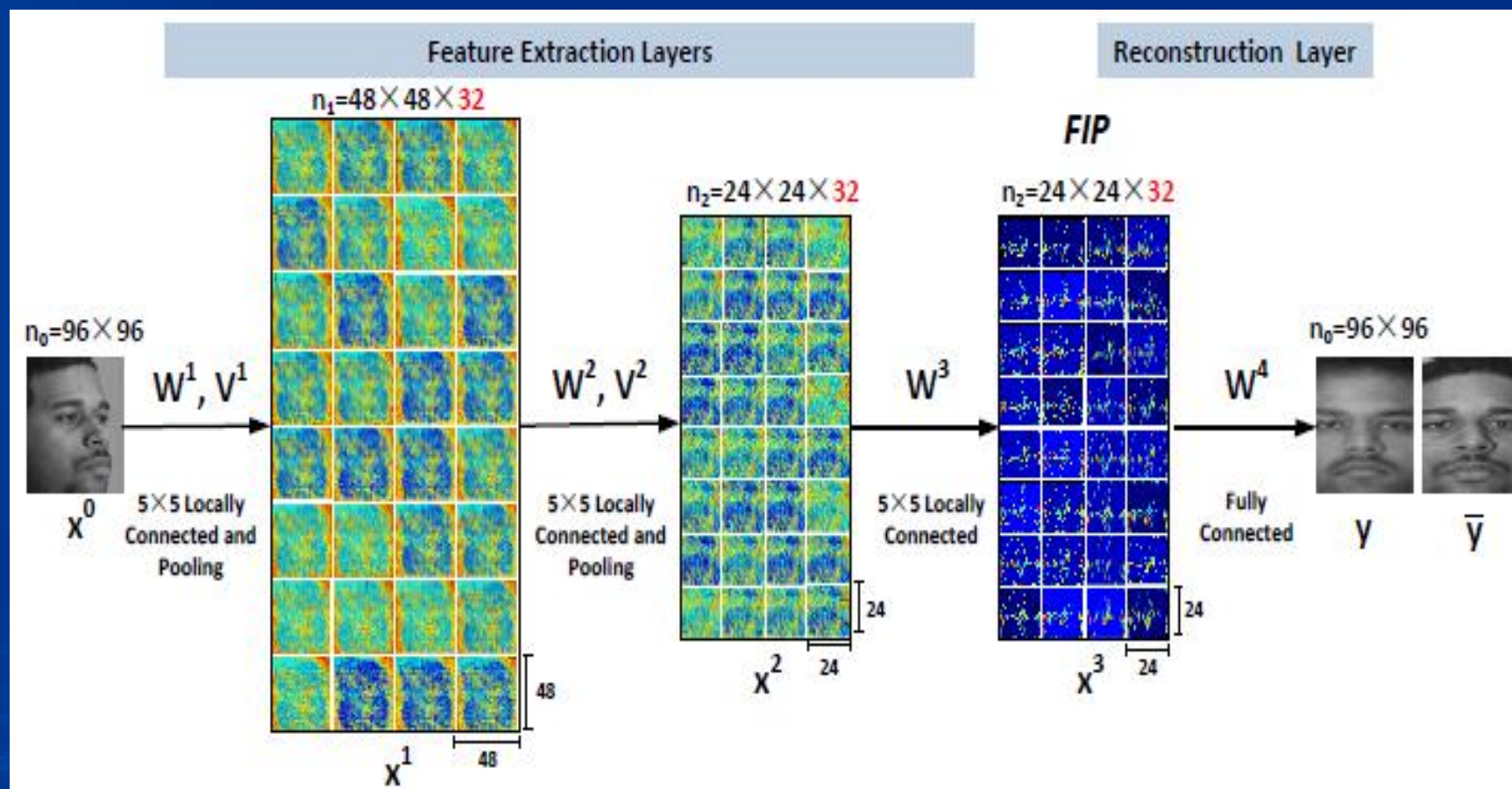
卷积波尔兹曼机 (Convolutional RBM)



- 输入图像通过与 m 个可训练的滤波器和可加偏置进行卷积，在C1层产生 m 个特征映射图。
- 然后特征映射图中每组的 n 个像素再进行求和，加权值，加偏置，通过一个Sigmoid函数得到 m 个S2层的特征映射图。
- 这些映射图再经过滤波得到C3层。这个层级结构再和S2一样产生S4。
- 最终这些像素值被光栅化，并连接成一个向量输入到传统神经网络，得到输出。



深度模型网络架构



$$x_i^1 = V^1 \sigma(W_i^1 x^0)$$

$$x^2 = V^2 \sigma(W^2 x^1)$$

$$x^3 = \sigma(W^3 x^2)$$

$$y = \sigma(W^4 x^3)$$



几种深度学习人脸识别系统

■ Face++

- 网络上搜集了5million张人脸图片用于训练深度卷积神经网络模型，在LFW数据集上可以达到99.5%的精度

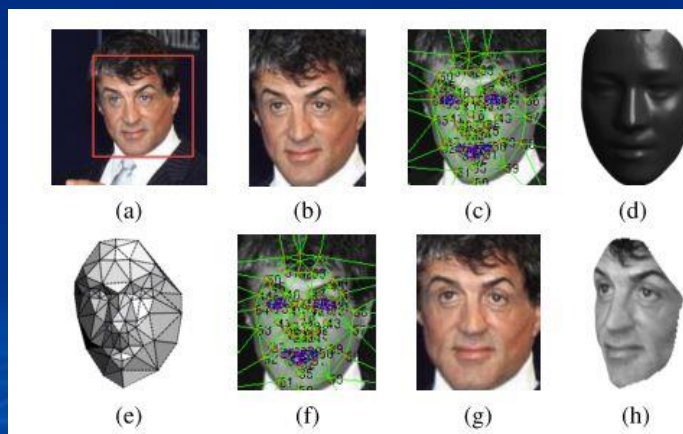


- 常用数据集：LFW数据集，共13233幅图像，其中5749个人，其中1680人有两幅及以上的图像，4069人只有一幅图像。

几种深度学习人脸识别系统

■ DeepFace (0.9735) :

- 通过额外的3d模型改进了人脸对齐的方法。1通过6个特征点检测人脸; 2剪切; 3, 建立Delaunay triangulation; 4, 参考标准3d模型; 5, 将3d模型比对到图片上; 6, 进行仿射变形; 7, 最终生成正面图像。



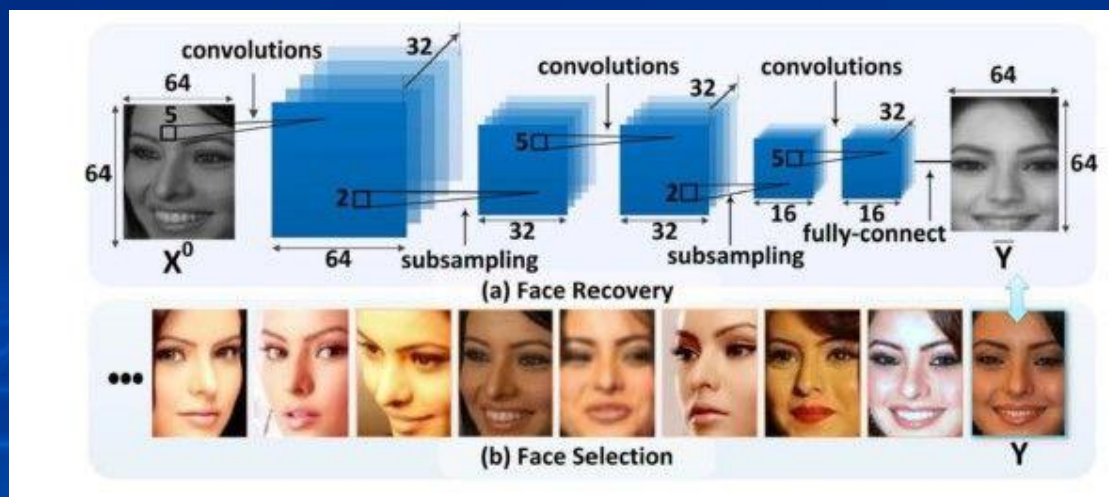
- 大数据训练的人工神经网络: 通过基于4million人脸图像 (4000个个体) 训练的一个9层的人工神经网络来进行人脸特征表达。



几种深度学习人脸识别系统

■ FR+FCN (0.9645) :

- 一是一种新的基于矩阵排序和对称性的人脸正面图像检测方法；二是训练深度神经网络来重建人脸正面标准图片。
- 通过训练深度神经网络来进行人脸重建。



● DeepID (0.9745) 、 FaceNet(0.9963)

<https://blog.csdn.net/tuuzhang/article/details/83541501>

课外实践练习 3

- 人脸检测识别交互

- 输入一个带人脸的视频，要求从视频中跟踪人脸的位置，有能力的可以识别人脸。
- 可以用外部的函数库或开源平台数据库，要求有结果的分析。

问题？

