

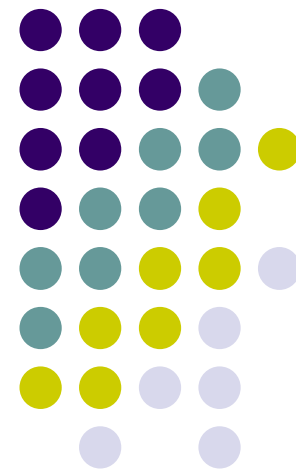


Automatic Feature Localization with Constrained Local Models

Author: David Cristinacce, Tim Cootes

LU Liang

2009.05





目录



- 作者的相关信息
- 中英文摘要
- ASM及AAM方法介绍
- 本文的主要算法：CLM
- 实验结果



作者的相关信息



- David Cristinacce
 - Research Associate in [ISBE](#) (Imaging Science and Biomedical Engineering, 图像科学与生物医学工程实验室) a department within the [University of Manchester](#).
 - Research Interests: mainly concentrated on face detection and facial feature localization
 - His PhD was supervised by Tim Cootes.
 - PhD: October 2000 to July 2004
 - MSc: Cognitive Science September 1999 to September 2000 Department of [Imaging Science and Biomedical Engineering](#), [University of Manchester](#).



作者的相关信息



- Publications

- **Facial motion analysis using clustered shortest path tree registration.** David Cristinacce, Natalie Butcher and Tim Cootes, in Workshop on Machine Learning for Vision-based Motion Analysis (MLVMA'08), part of 10th European Conference on Computer Vision ([ECCV 08](#)), Marseille, France, to appear, October, 2008.
- **Automatic Feature Localisation with Constrained Local Models.** David Cristinacce and Tim Cootes, *Journal of Pattern Recognition (PR 08)*, Volume 41, Issue 10, pages 3054-3067, October, 2008.
- **3D Facial Geometry Recovery via Group-wise Optical Flow.** Hui Fang, Nicholas Costen, David Cristinacce, John Darby, in Proceedings of 8th IEEE International Conference on Automatic Face and Gesture Recognition ([AFGR 08](#)), Amsterdam, The Netherlands, September, 2008.
- **Boosted Regression Active Shape Models.** David Cristinacce and Tim Cootes, in Proceedings of 18th British Machine Vision Conference ([BMVC 07](#)), pages 890-899, Warwick, UK, September, 2007.
- **Feature Detection and Tracking with Constrained Local Models (Best Science Paper Prize)** . David Cristinacce and Tim Cootes, in Proceedings of 17th British Machine Vision Conference ([BMVC 06](#)), pages 929-938, Edinburgh, Scotland, September, 2006.
- **Facial Feature Detection and Tracking with Automatic Template Selection** . David Cristinacce and Tim Cootes, in Proceedings of 7th IEEE International Conference on Automatic Face and Gesture Recognition, pages 429-434, Southampton, UK, April, 2006.



作者的相关信息



- Tim Cootes

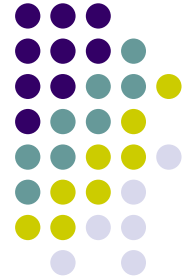
- Imaging Science and Biomedical Engineering Research Group ([ISBE](#)), University of Manchester, where he has worked since 1991. He became a Professor of Computer Vision in August 2006

- Research Interests

- Statistical models of shape and appearance (形状和表观统计模型)
- Model matching algorithms (模型匹配)
- Methods of computing correspondence across sets of shapes and images
- Applications in medical image analysis and face image interpretation(医学影像处理等)



Abstract



- We present an efficient and robust method of locating a set of feature points in an object of interest. From a training set we construct a joint model of the appearance of each feature together with their relative positions. The model is fitted to an unseen image in an iterative manner by generating templates using the joint model and the current parameter estimates, correlating the templates with the target image to generate response images and optimizing the shape parameters so as to maximize the sum of responses. The appearance model is similar to that used in the Active Appearance Models (AAM) [T.F. Cootes, G.J. Edwards, C.J. Taylor, Active appearance models, in: Proceedings of the 5th European Conference on Computer Vision 1998, vol. 2, Freiburg, Germany, 1998.].
- However in our approach the appearance model is used to generate likely feature templates, instead of trying to approximate the image pixels directly. We show that when applied to a wide range of data sets, our Constrained Local Model (CLM) algorithm is more robust and more accurate than the AAM search method, which relies on the image reconstruction error to update the model parameters. We demonstrate improved localization accuracy on photographs of human faces, magnetic resonance (MR) images of the brain and a set of dental panoramic tomograms. We also show improved tracking performance on a challenging set of in car video sequences.



Abstract



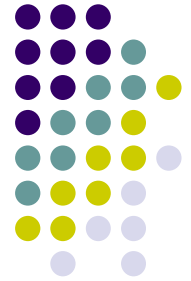
- 本文提出了一种高效鲁棒的特征点定位方法。通过训练集我们可以构建一个每个特征点**表观**（纹理模型）及其相应**位置**（形状模型）的联合模型
- 本文中采用的表观模型与Cootes提出的AAM方法[T.F. Cootes, G.J. Edwards, C.J. Taylor, Active appearance models, European Conference on Computer Vision 1998, vol. 2, 1998.]类似，不同之处在于，本文的方法中表观模型用于生成特征模板，而非直接去近似原图像中的像素
- 通过实验证明了本文的CLM方法比AAM更为鲁棒和精确，同时证明了该方法在tracking时的性能提升



Introduction to ASM and AAM



ASM: Active Shape Model (1)



- ASM模型的实现分为：
 - ASM Building 和
 - ASM Fitting 两个过程
- ASM Building —— 训练集的构建
 - 定义：训练集中第*i*个形状的第*j*个特征点的坐标记为 (x_{ij}, y_{ij})
 - 描述第*i*个形状的*n*个特征点的形状向量由特征点的坐标串接而成，记为

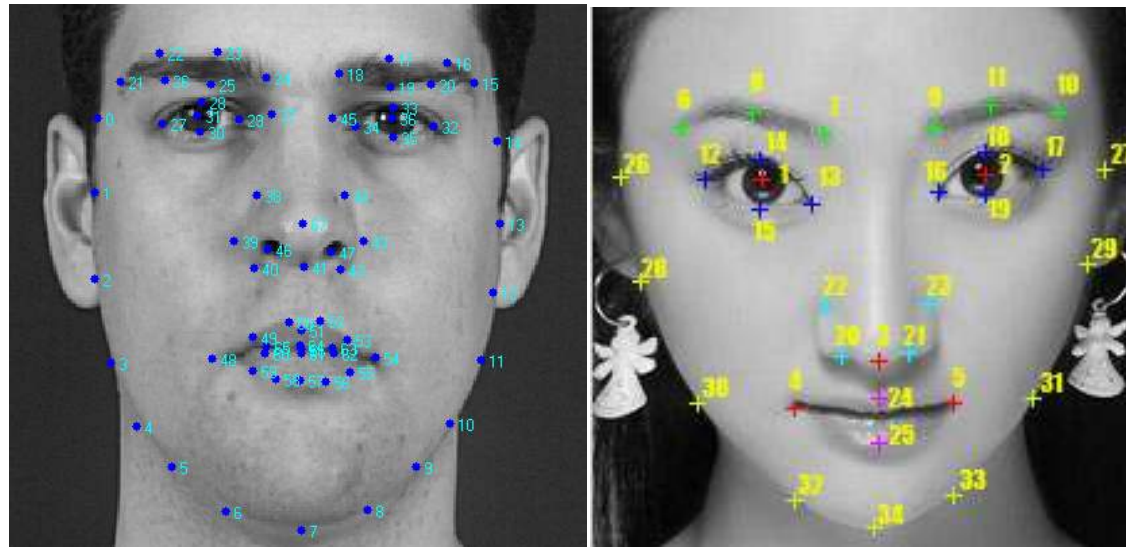
$$\mathbf{x}_i = [x_{i0}, y_{i0}, x_{i1}, y_{i1}, \dots, x_{im-1}, y_{im-1}]^T; 1 \leq i \leq N.$$



ASM: Active Shape Model (2)



- Suitable Landmarks
 - 左图是XM2VTS定义的68个特征点的情况
 - 右图是我们定义的34个特征点的情况





ASM: Active Shape Model (3)



- Aligning the Training Set
 - 为研究训练集的图像中每个特征点的变化情况，需要将上述形状向量 \mathbf{x} 进行对齐操作
 - 对齐的优化目标：每个形状与平均形状的距离之差最小，即最小化下述等式：

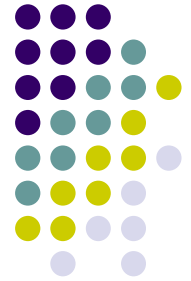
$$D = \sum |\mathbf{x}_i - \bar{\mathbf{x}}|^2$$

- 如何对齐两个形状？

给定形状 x_i 和 x_j ，寻找缩放 s ，旋转 θ 以及平移 (t_x, t_y) ，使得向量 x_j 在加权最小二乘意义下与 x_i 的距离最小
- 每个特征点的权值设定：对应大加权的特征点将更加稳定



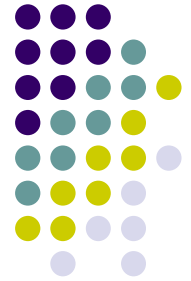
ASM: Active Shape Model (4)



- **Aligning: Iterative Approach**
 1. 将每个样本的重心平移到坐标原点
 2. 选择一个样本作为初始的平均形状估计（如第一个形状），将第一个估计的平均形状记为 \bar{x}_0
 3. 所有的形状都朝当前估计的平均形状进行仿射变换
 4. 对变换后的形状重新估计平均形状
 5. 然后将所有的形状向新的平均形状进行对齐
 6. 如果未收敛，则回到step 3，如果收敛则意味着所有的形状已经实现对齐
- 敛的标志：两次迭代之间的平均形状没有剧烈的变化，小于某一预先设定的阈值



ASM: Active Shape Model (5)



- Modeling Shape Variation

- 计算平均形状 $\bar{x} = \frac{1}{s} \sum_{i=1}^s x_i$

- 计算形状的协方差矩阵 $S = \frac{1}{s-1} \sum_{i=1}^s (x_i - \bar{x})(x_i - \bar{x})^T$

- 计算协方差矩阵S的特征向量以及特征值

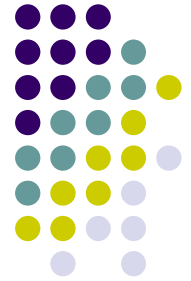
- ASM模型可以构建出训练集中出现的任何形状变化。给定一组训练集，通过PCA方法可以生成训练集中数据的形状统计模型，物体的形状可以表示成一个形状向量：

$$x \approx \bar{x} + \Phi b$$

其中， \bar{x} 是平均形状， $\Phi = (\phi_1 | \phi_2 | \phi_3 \dots | \phi_t)$ 是由训练集得到的用于描述物体形状的前t个主成分分量，即特征向量，b是权重向量



ASM: Active Shape Model (6)



- 参数向量 b 通过训练集得到，它决定了模型可以允许的形状变化
- 为了限制模型形状的变化范围，我们通常将参数向量 b 的第 i 个分量 b_i 限制在 $\pm 3\sqrt{\lambda_i}$ 之内，其中 λ_i 为对应的特征向量 ϕ_i 的特征值
- 下面一步选择模式数目，就是通过PCA方法选择前 t 个特征向量，保留形状的部分变化。其余部分可以认为是干扰噪声



ASM: Active Shape Model (7)

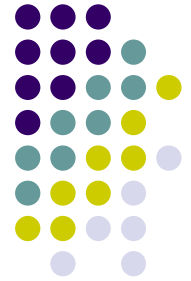


- Choice of Number of Modes (PCA)
 - 用 λ_i 表示训练数据的特征点串接向量构成的协方差矩阵的特征值
 - 每个特征向量都表示形状变化的某种模式
 - 特征值表示了数据在均值基础上的对应于相应特征向量方向上的变化
 - 训练数据的所有形变就是特征值的求和

$$V_T = \sum \lambda_i$$



ASM: Active Shape Model (8)



- Choice of Number of Modes (cont')

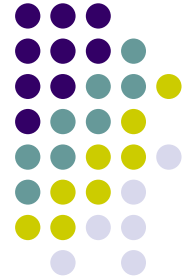
- 我们可以选择前t个最大的特征向量，使得

$$\sum_{i=1}^t \lambda_i \geq f_v V_T$$

- f_v 是0-1之间的值，定义了保留的能量的多少，如取0.98
- 这样我们便选择了部分对形状影响比较大的主成分方向，同时忽略了部分微小的变化方向（或直接认为这部分变化都是由于噪声引起的）



ASM: Active Shape Model (9)



- Examples of Shape Models
 - 下图是手形状变化的训练集

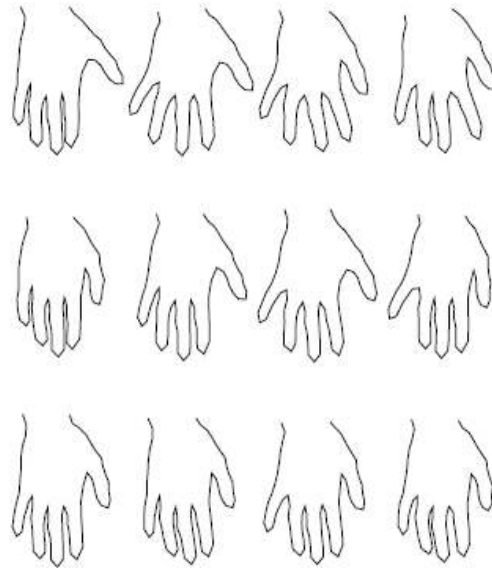


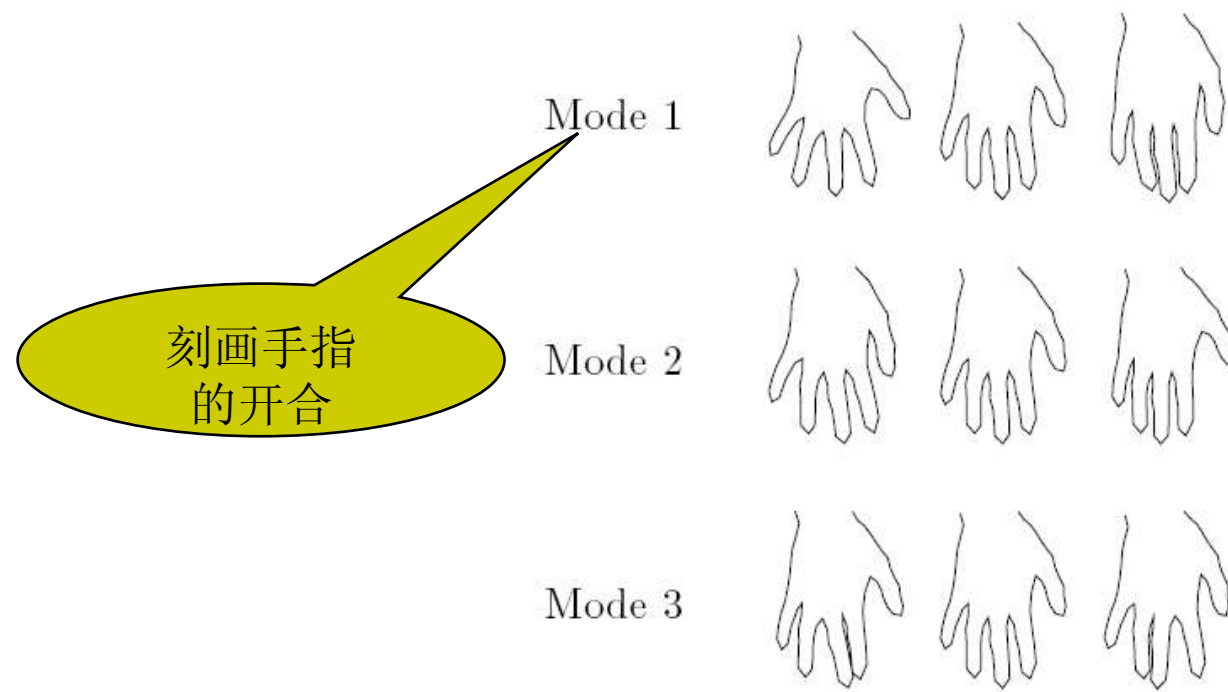
Figure 4.4: Example shapes from training set of hand outlines



ASM: Active Shape Model (10)



- 经过ASM Train之后我们得到的前三种主要变化模式如下：

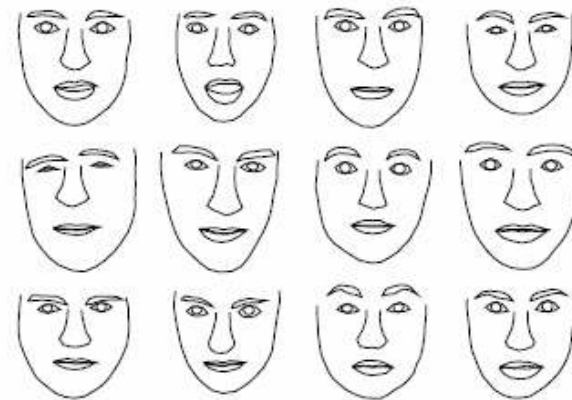
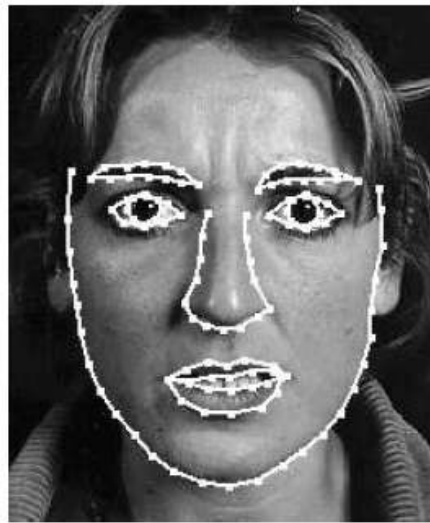




ASM: Active Shape Model (11)



- 下图是标注过的人脸数据的示例以及训练集中部分人脸图像的示例，
- 对于标注过的人脸图像训练集我们可以通过ASM训练得到主要的变化模式



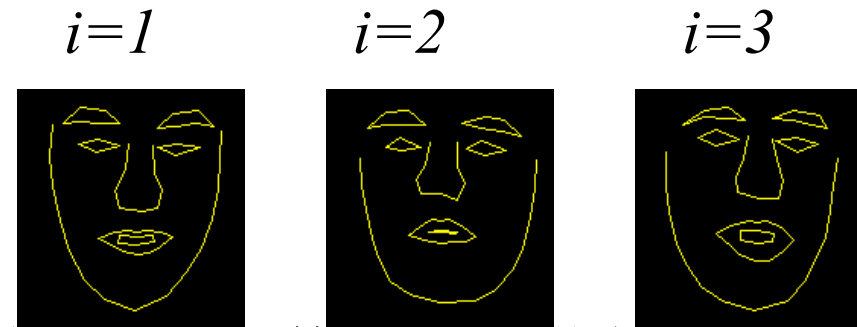
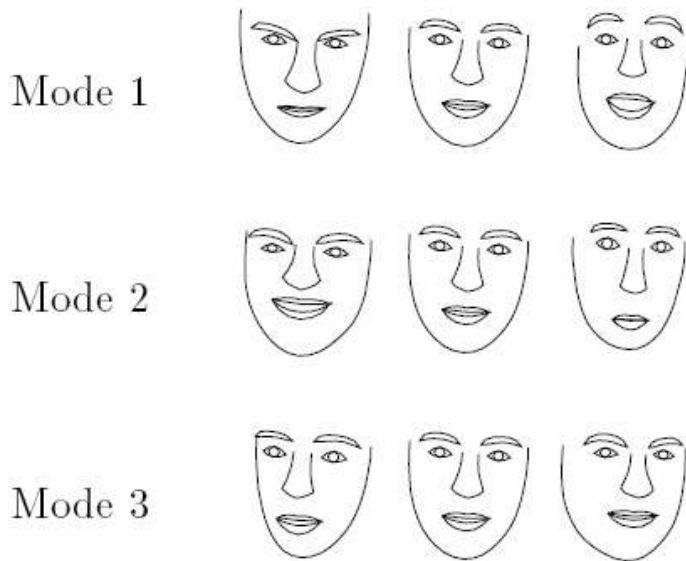
Example shapes from training set of faces



ASM: Active Shape Model (12)



- ASM训练得到的主要人脸变化模式



前3个形状模式的变化情况

$$b_s^i \in (-3\sqrt{\lambda_i}, +3\sqrt{\lambda_i})$$



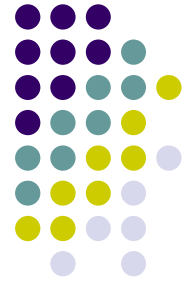
ASM: Active Shape Model (13)



- ASM Fitting----匹配过程
 - 采用迭代匹配算法，步骤如下：
 - 每个特征点局部搜索：Examine a region of the image around each point X_i to find the best nearby match for the point X'_i
 - 调整全局形状参数：Update the parameters (X_t, Y_t, s, θ, b) to best fit the new found points $X = T(X_t, Y_t, s, \theta)[\bar{x} + P_s b]$
 - Repeat until convergence.



ASM: Active Shape Model (14)



- 局部搜索过程：利用灰度特征，寻找Profile上的特征点在马氏距离下的最佳匹配

第*i*个特征点的Profile上的灰度采样： $s_i = [g_{i0}, g_{i1}, \dots, g_{i(n_s-1)}]^T$,

其中， n_s 为Profile的长度

Profile上灰度的梯度： $ds_i = [g_{i1} - g_{i0}, g_{i2} - g_{i1}, \dots, g_{i(n_s-1)} - g_{i(n_s-2)}]^T$

归一化的灰度梯度： $g_i = ds_i / \sum_{k=0}^{n_s-2} |ds_{ik}|$

- 最小化目标函数： $f(\mathbf{g}_s) = (\mathbf{g}_s - \bar{\mathbf{g}})^T \mathbf{S}_g^{-1} (\mathbf{g}_s - \bar{\mathbf{g}})$

便得到Profile上的特征点在马氏距离下的最佳匹配

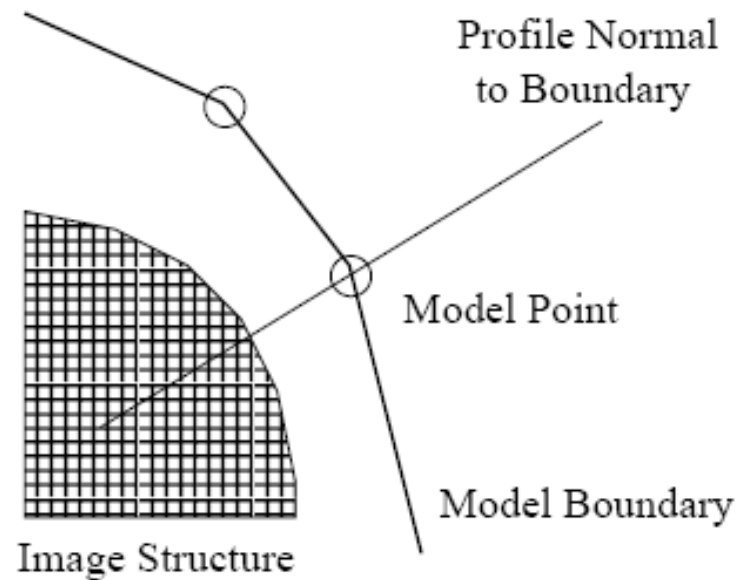
其中， S_g^{-1} 是由训练集得到的灰度协方差矩阵的逆



ASM: Active Shape Model (15)



- 局部搜索过程示意图





AAM: Active Appearance Model (1)



- Appearance = Shape + Texture
 - 每幅训练图像都被分成“a set of points”和“shape-free image patch”

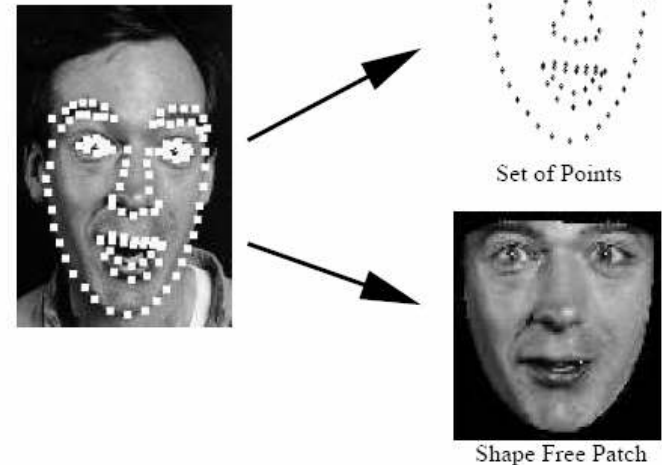
- ASM中是统计形状模型

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}_s \mathbf{b}_s$$

- AAM在统计形状模型的基础上，增加了统计灰度模型

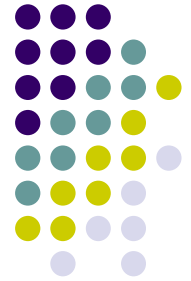
$$\mathbf{g} = \bar{\mathbf{g}} + \mathbf{P}_g \mathbf{b}_g$$

- 从而形成了复合的表观模型





AAM: Active Appearance Model (2)



- Combined Appearance Models
 - AAM将形状参数与灰度参数串接在一起，其中 W_s 是一对角阵，用于统一量纲

$$\mathbf{b} = \begin{pmatrix} W_s \mathbf{b}_s \\ \mathbf{b}_g \end{pmatrix} = \begin{pmatrix} W_s \mathbf{P}_s^T (\mathbf{x} - \bar{\mathbf{x}}) \\ \mathbf{P}_g^T (\mathbf{g} - \bar{\mathbf{g}}) \end{pmatrix}$$

- 对上述向量通过施加PCA得到 $\mathbf{b} = \mathbf{P}_c \mathbf{c}$
where \mathbf{P}_c are the eigenvectors and \mathbf{c} is a vector of appearance parameters controlling both the shape and grey-levels of the model.
- 这样AAM后续的步骤就可以按照ASM的方法进行
- 在Fitting的时候AAM也利用了纹理信息，所以AAM精度更高



AAM: Active Appearance Model (3)



- AAM示例
 - We used a training set of 400 images of faces, each labelled with 122 points around the main features
 - The model uses about 10,000 pixel values to make up the face patch.
 - 右图是训练得到形状，灰度和表观的变化

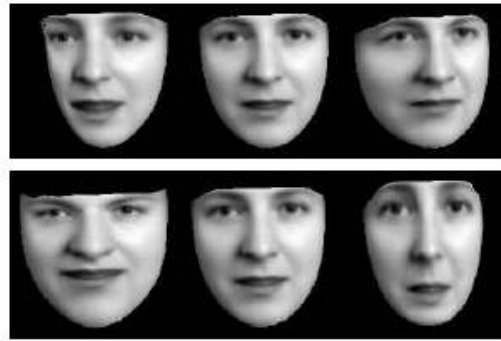


Figure 5.2: First two modes of shape variation (± 3 sd)

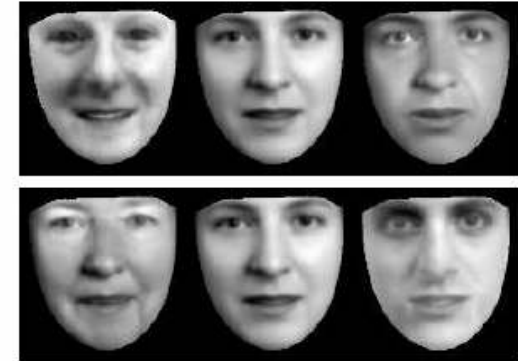


Figure 5.3: First two modes of grey-level variation (± 3 sd)



Figure 5.4: First four modes of appearance variation (± 3 sd)



AAM:

- AAM Fitting Result

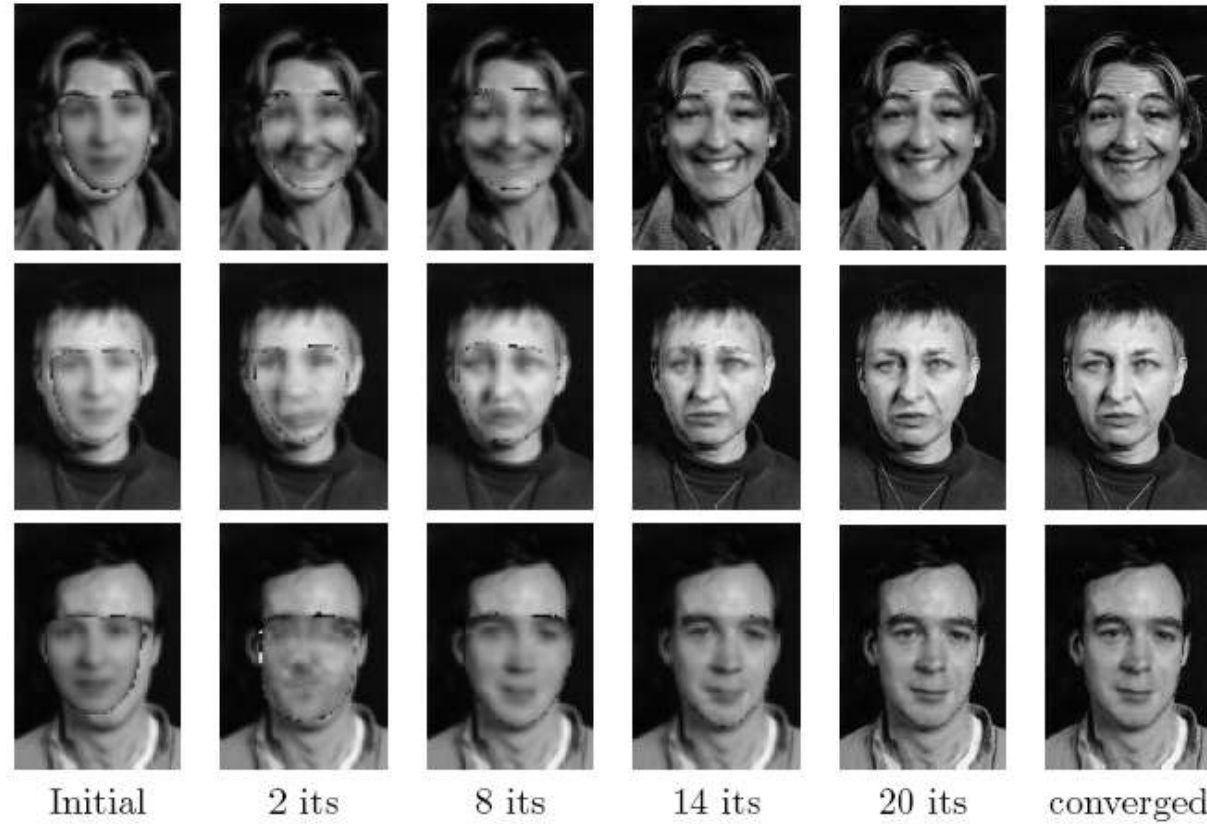


Figure 8.10: Multi-Resolution search from displaced position



Figure 8.9: Reconstruction (left) and original (right) given original landmark points



Automatic feature localization with Constrained Local Models



本文的主要算法



- CLM搜索算法
 - (i) Input an initial set of feature points.
 - (ii) Repeat:
 - (a) 使得 joint model 匹配到当前特征点集上 (see Section 3.1).
 - (b) 生成一组模板 (see Section 3.2).
 - (c) 采用本文提出的 shape constrained search 方法来预测一组新的特征点 (see Section 3.3).
 - (ii) until converged.

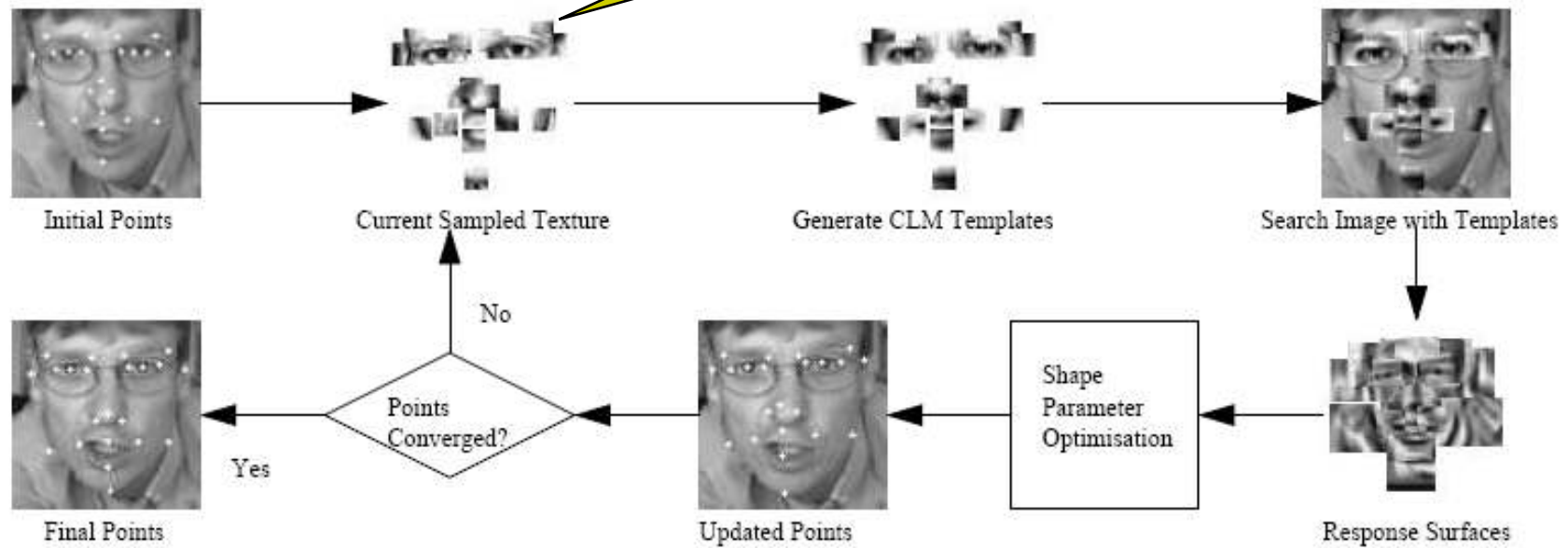


本文的主要算法



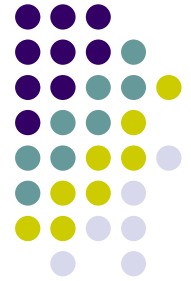
CLM搜索算法流程图

Step1: 当前的纹理采样

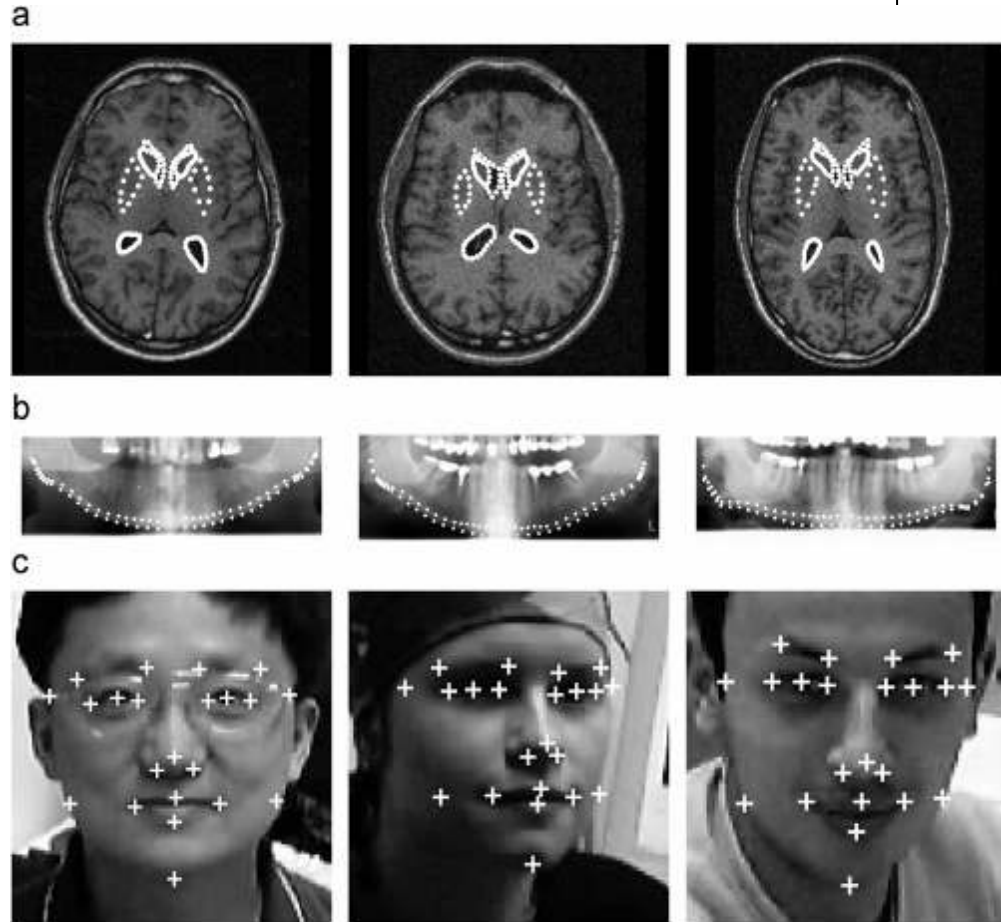




3.1. Constrained local appearance models (1)

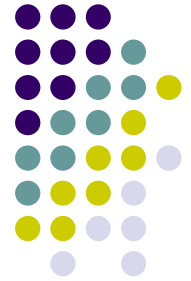


- CLM模型的构建
 - 一个 joint shape and texture model 通过手工标注的图像数据训练得到
 - 每一个 training patch 都是在特征点附近通过采样得到
 - 零均值一方差归一化处理





3.1. Constrained local appearance models (2)



- 然后将一幅训练图像的 texture patches 串接成一个向量
- 对所有向量组合而成的矩阵进行PCA处理，得到统计模型
- 经过训练得到形状和灰度的模型：

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}_s \mathbf{b}_s$$

$$\mathbf{g} = \bar{\mathbf{g}} + \mathbf{P}_g \mathbf{b}_g$$

- 然后将形状与灰度组成新的向量，并再次PCA得到 joint model θ :

$$\mathbf{b} = \mathbf{P}_c \mathbf{c}, \quad \text{其中 } \mathbf{b} = \begin{pmatrix} \mathbf{W}_s \mathbf{b}_s \\ \mathbf{b}_g \end{pmatrix}$$



3.1. Constrained local appearance models (3)



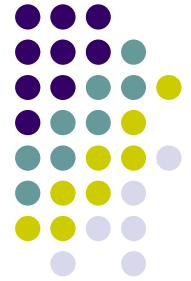
- 训练得到的模型：通过改变参数c的前两个分量得到如下模式变化



Fig. 1. PCA modes of combined shape and texture variation for CLM face model ($\pm 3\text{std}$).

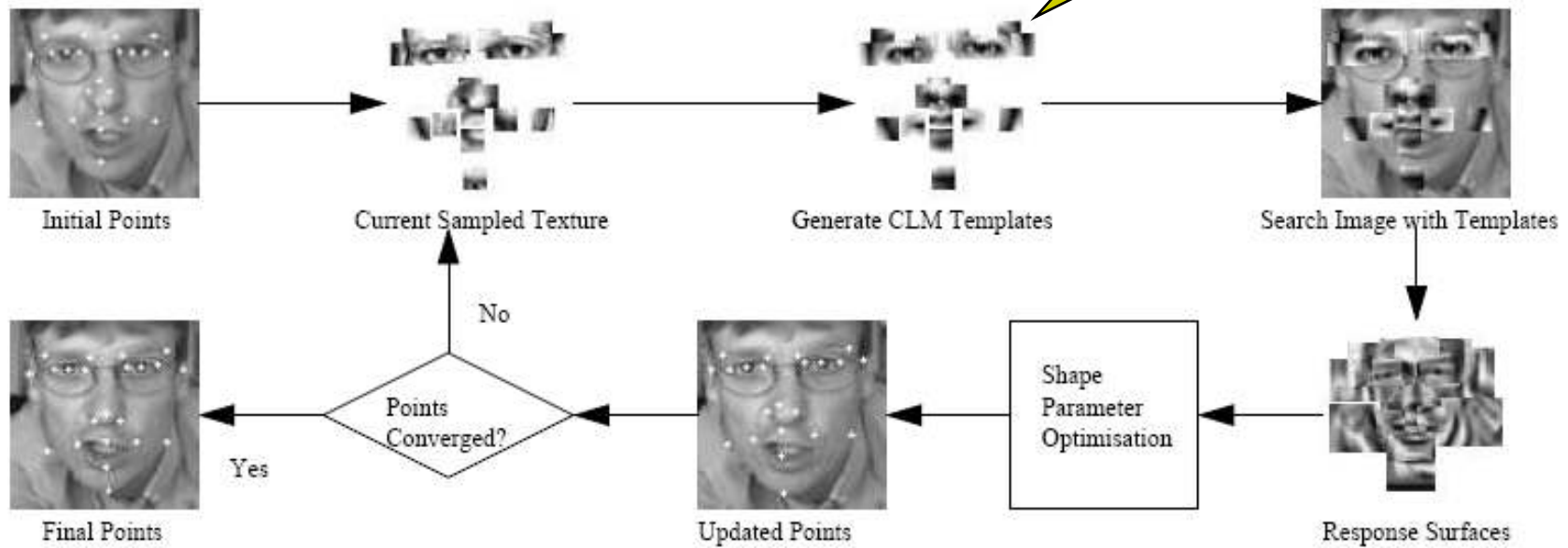


本文的主要算法



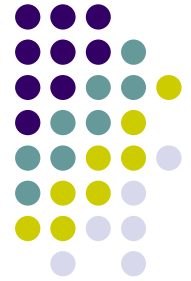
- CLM搜索算法流程图

Step 2: 生成CLM模板





3.2. Template generation (1)



- 形状模板的生成:

- 令 (X_i, Y_i) 表示第 i 个特征点的位置, 将其串接为一个向量:

$$\mathbf{X} = (X_1, \dots, X_n, Y_1, \dots, Y_n)^T$$

- 然后将当前形状通过仿射变换给出特征点的初始位置:

$$\mathbf{X} \approx S_t(\bar{\mathbf{x}} + \mathbf{P}_s \mathbf{b}_s)$$

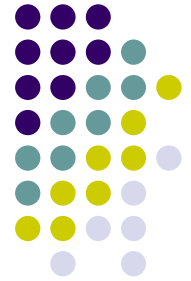
- 可以将放射变换参数与形状参数串接成一个向量, 得到如下形式:

$$\mathbf{s} = (\mathbf{t}^T | \mathbf{b}_s^T)^T$$

- 后文中可以看到, \mathbf{s} 就是优化函数的自变量



3.2. Template generation (2)

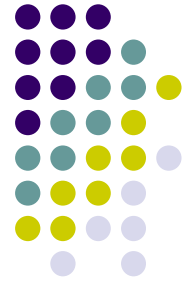


- 每个特征点的纹理模板的生成：
 - 利用图像I在点 x_i 处的texture patch的采样可以估计出joint model的最优的纹理匹配参数 b_g ，从而生成当前图像的表观模板（与Cootes的AAM中使用的方法相同）
- 将当前输入图像的形状参数与纹理参数加权之后串接起来，在进行PCA，从而可以得到当前图像模型参数c

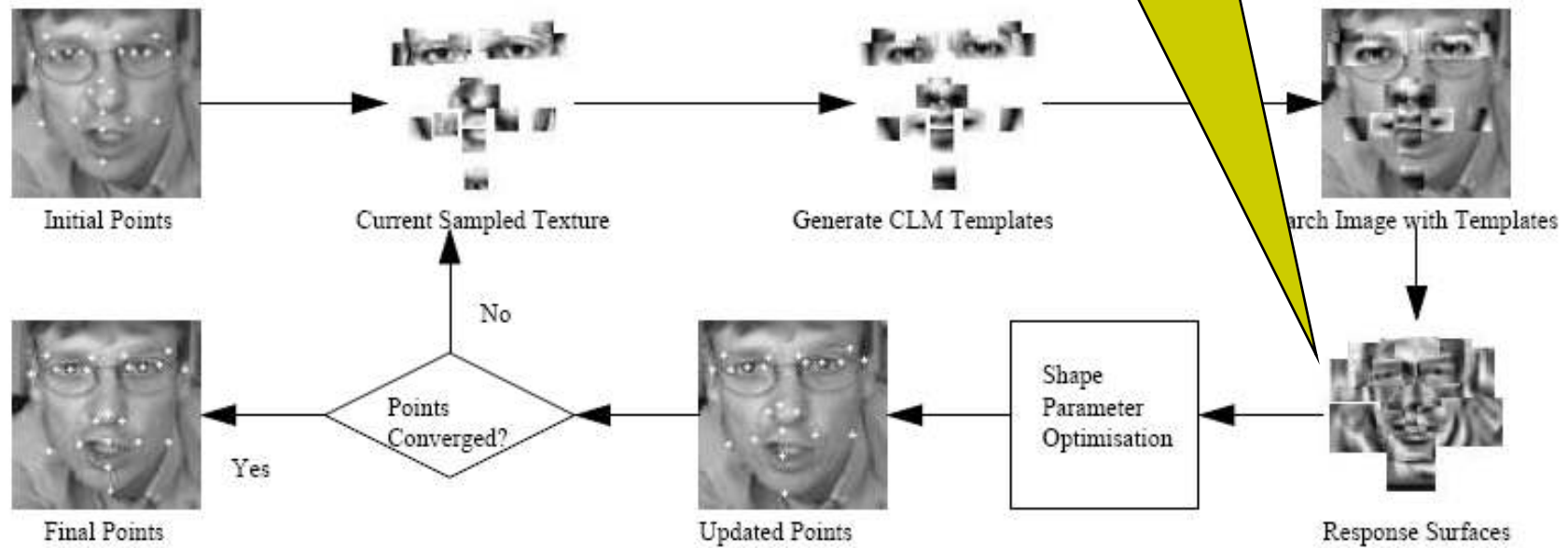
$$\mathbf{b} = \mathbf{P}_c \mathbf{c}, \text{ 其中 } \mathbf{b} = \begin{pmatrix} \mathbf{W}_s \mathbf{b}_s \\ \mathbf{b}_g \end{pmatrix}$$



本文的主要算法



- CLM搜索算法流程图





3.3. Shape constrained search update



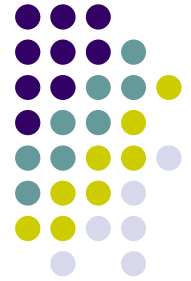
- 优化目标函数的直观解释
 - 目标函数的形式

$$f(\mathbf{s}) = \alpha \sum_{i=1}^n R_i(X_i, Y_i) - \sum_{j=1}^k \frac{b_j^2}{\lambda_j}.$$

- 解释：目标函数有两项构成，其中 α 是一个相对权重调节因子， $R_i(X_i, Y_i)$ 是第 i 个特征模板在点 (X_i, Y_i) 处的响应，可以是当前特征点的 texture patch 跟 joint model 中该点的特征模板的纹理之差，这一项刻画了模板与当前图像匹配的程度。第二项是形状参数与其对应的特征值之比的求和，可以发现该项越小则形状越接近平均形状，该项刻画了当前形状的“规则”程度。通过调节权重因子可以确定优化的过程中更注重匹配的程度还是更注重形状的“好坏”



3.4. Search algorithm



- 3.3中的公式推导主要通过Bayesian方法加Gibbs–Boltzmann分布得到的，在此不展开介绍
- 再来回顾一下刚刚介绍的整个搜索算法
- CLM搜索算法
 - (i) Input an initial set of feature points.
 - (ii) Repeat:
 - (a) 通过形状初始化，使得 joint model匹配到当前特征点集上 (see Section 3.1).
 - (b) 生成一组模板（形状模板+每个特征点的纹理模板）(see Section 3.2).
 - (c) 采用 shape constrained search 方法来预测一组新的特征点 (see Section 3.3).
 - (ii) until converged.



Experiments



- 误差度量准则：Distance error
(Distance metric)

$$m_e = \frac{1}{nS} \sum_{i=1}^{i=n} d_i$$

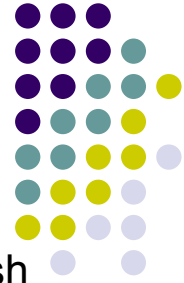
欧式距离下定位点
与标注点的误差

双眼
间距

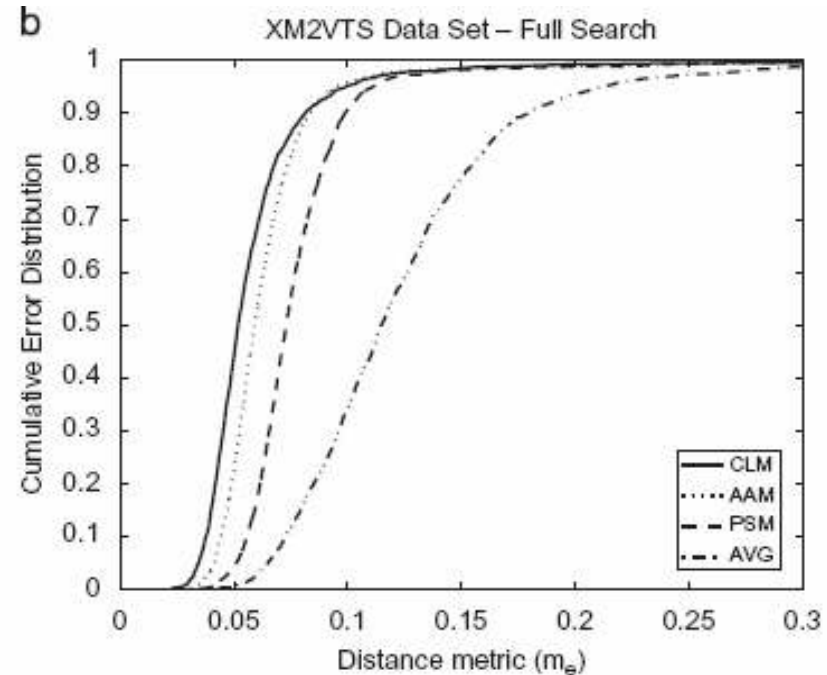
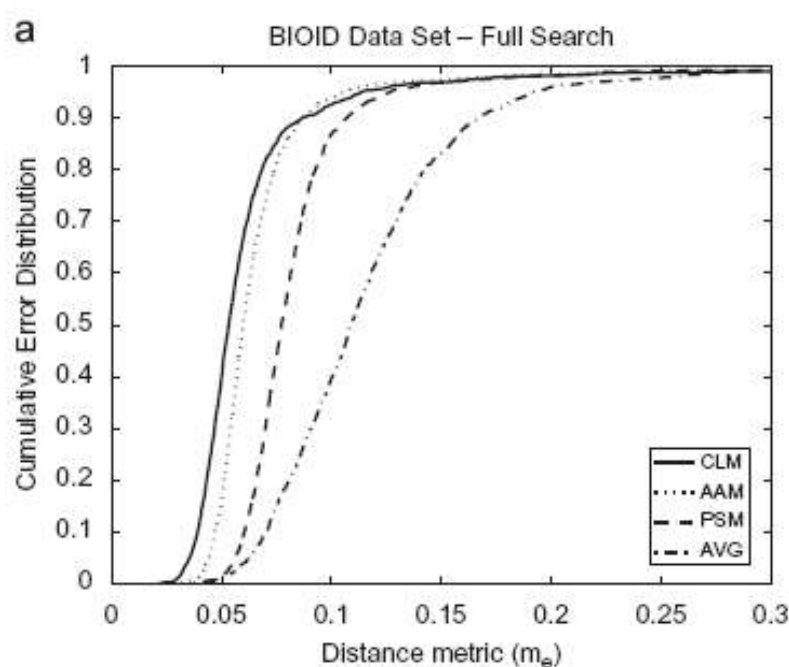
- 训练集：标注了22个点的1052幅图像，与测试集独立
- 测试集：publicly available BIOID[25] and XM2VTS[26]



Experiments



- **AVG**—Average points within the global Viola and Jones face detector[22] (dot-dash line).
- **PSM**—Pictorial Structure Matching points found within the Viola and Jones candidate face region[1] (dashed line).
- **AAM**—Active Appearance Model algorithm initialized with PSM points[3](dotted line).
- **CLM**—Constrained Local Model initialized with the PSM points (solid line).





Experiments

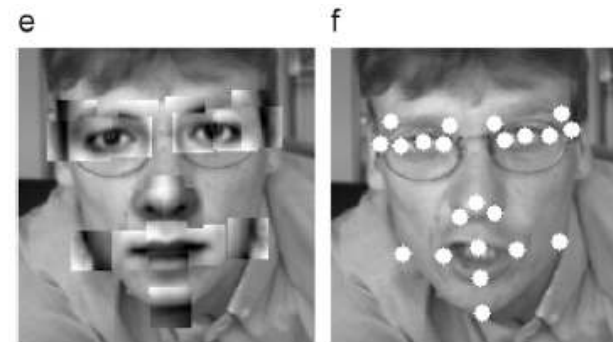
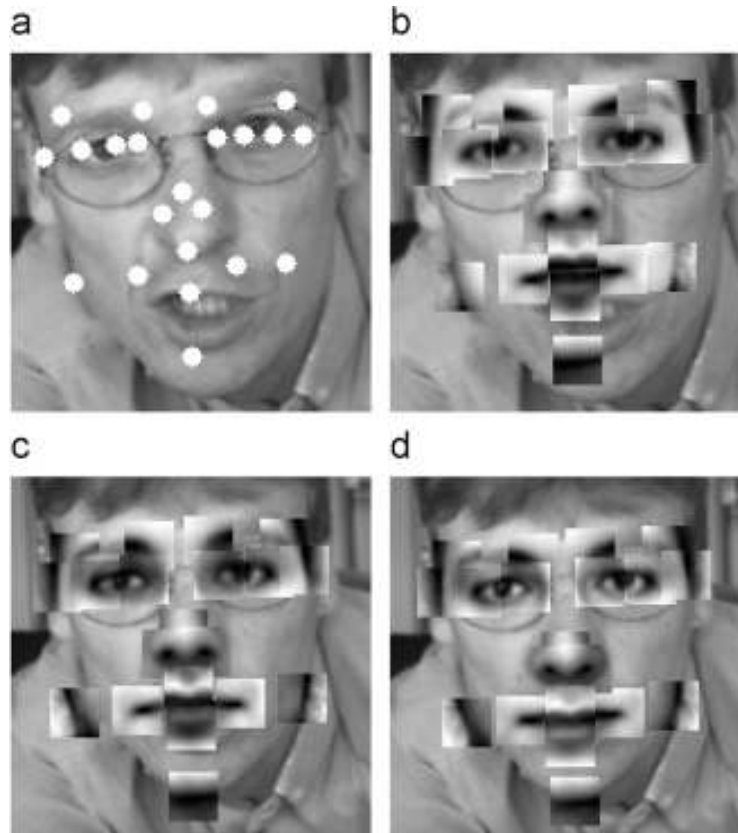


Fig. 9. Evolution of CLM templates when searching a static image. (a) Start points, (b) start templates, (c) after iteration 1, (d) after iteration 2, (e) after iteration 3 and (f) final points.

- CLM迭代的结果
 - (a) Start points
 - (b) start templates
 - (c) after iteration 1
 - (d) after iteration 2
 - (e) after iteration 3
 - (f) final points



谢谢大家!

