

CVPR 2009 Quick Review: Misalignment in Face Recognition

讲解人: 李安南

中科院计算所JDL人脸组Seminar

2009年9月11日

讲解的文章

讲解三篇CVPR 2009 上的文章:

- ▶ Sanqiang Zhao and Yongsheng Gao. Textural Hausdorff Distance for Wider-Range Tolerance to Pose Variation and Misalignment in 2D Face Recognition. In Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1629-1634, 2009.
- ▶ Peng Li and Simon J. D. Prince. Joint and Implicit Registration for Face Recognition. In Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1510-1517, 2009.
- ▶ John Wright and Gang Hua. Implicit Elastic Matching with Random Projections for Pose-Variant Face Recognition. In Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1502-1509, 2009.

#1 第一篇文章

文章标题: Textural Hausdorff Distance for Wider-Range Tolerance to Pose Variation and Misalignment in 2D Face Recognition

作者: Sanqiang Zhao and Yongsheng Gao

文章ID: 0328

相关文献:

- ▶ S. Zhao and Y. Gao, “Significant Jet Point for Facial Image Representation and Recognition” , The IEEE International Conference on Image Processing (ICIP), pp.1664-1667, San Diego, California, USA, October 12 - 15, 2008.

#1: 第一作者

Zhao Sanqiang 赵三强

PhD Student

School of Engineering

Griffith University, Nathan Campus, Australia



学术背景:

- ▶ 2004年在北京工业大学多媒体技术与图形学实验获得硕士学位。
- ▶ 2004年至今在澳大利亚格里菲斯大学(Griffith University)攻读博士学位。

#1: 第二作者

Gao Yongsheng

Associate Professor

School of Engineering

Griffith University, Nathan Campus, Australia

<http://maxwell.me.gu.edu.au/yg/>



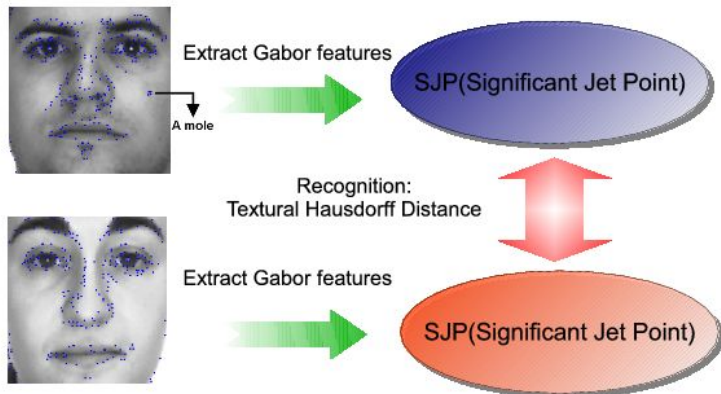
学术背景:

- ▶ 分别在1985和1988年在浙江大学获电子工程学士和硕士学位。后在新加坡南洋理工大学获得博士学位。
- ▶ 目前是澳大利亚格里菲斯大学副教授。

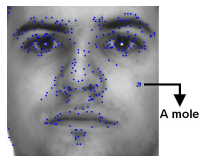
发表的文章:

- ▶ Y. Gao and M.K.H. Leung, "Face Recognition Using Line Edge Map," IEEE Trans. PAMI, 24(6): 764-779, 2002.

#1: 本文的识别流程



#1: 第一步,显著点检测



- ▶ 用参考文献[24]中的方法提取边缘。
- ▶ 然后用参考文献[25]中的方法提取显著点。
- ▶ 点的数目不固定。

- 24 R. Nevatia and K.R. Babu, "Linear Feature Extraction and Description," Computer Graphics and Image Processing, 13(3) 257-269, 1980.
- 25 M.K. Leung and Y.-H. Yang, "Dynamic Two-Strip Algorithm in Curve Fitting," Pattern Recognition, 23(1-2): 69-79, 1990.

#1: Felix Hausdorff



Felix Hausdorff (1868 - 1942) 德国数学家

#1: 豪斯道夫距离

- ▶ 对两个点集 $A = \{a_1, a_2, \dots, a_m\}$ 和 $B = \{b_1, b_2, \dots, b_n\}$ 它们之间的豪斯道夫距离(Hausdorff Distance)定义为:

$$HD(A, B) = \max(d_{HD}(A, B), d_{HD}(B, A))$$

- ▶ 其中, d_{HD} 为有向豪斯道夫距离(directed Hausdorff distance), 其定义为:

$$d_{HD}(A, B) = \max_{a \in A} \min_{b \in B} d(a, b) = \max_{a \in A} \min_{b \in B} \|a - b\|$$

- ▶ 也就是对集合A中的每个点求其到集合B中点的最短距离, 在这些最短距离中取最大的一个。

#1: 改进豪斯道夫距离

- ▶ 原始豪斯道夫距离对野点(outlier)很敏感。
- ▶ 针对这个问题，本文应用了参考文献[14]中的改进豪斯道夫距离(modified Hausdorff distance(MHD))。

$$MHD(A, B) = \max(d_MHD(A, B), d_MHD(B, A))$$

- ▶ 其中， d_MHD 为有向MHD定义为：

$$d_MHD(A, B) = \frac{1}{N_A} \sum_{a \in A} \min_{b \in B} \| a - b \|$$

其中 $N_A = m$ 是集合A中点的数目。

- ▶ 这种取平均的做法可降低对野点(outlier)的敏感性。

#1: 纹理豪斯道夫距离

- ▶ 类似改进豪斯道夫距离，对于两个有限的SJP集合

$$M = \{SJP_1^M, SJP_2^M, \dots, SJP_P^M\}$$

和 $T = \{SJP_1^T, SJP_2^T, \dots, SJP_Q^T\}$ 所谓纹理豪斯距离定义为:

$$THD(M, T) = \max(d_THD(M, T), d_THD(T, M))$$

- ▶ 其中， d_THD 为定义为:

$$d_THD(M, T) = \frac{1}{P} \sum_{1 \leq p \leq P} \min_{1 \leq q \leq Q} (d(x_p^M, x_q^T) + \lambda(1 - S(J_p^M, J_q^T)))$$

- ▶ 其中 $d(x, x') = \|x - x'\|$ 是空间位置距离，而 S 定义为:

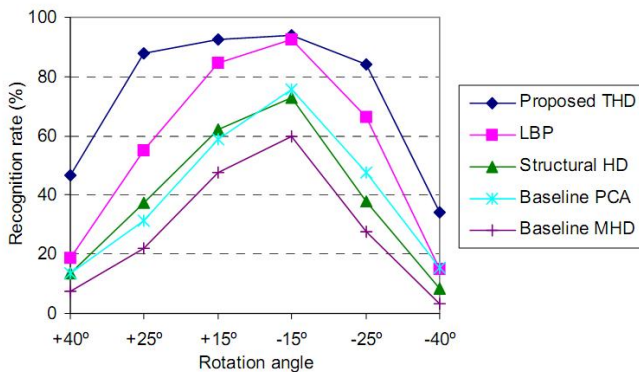
$$S(J, J') = \frac{\sum_j a_j a'_j}{\sqrt{\sum_j a_j^2 \sum_j a'_j{}^2}}$$

这里 a 是某点 Gabor Jet 的幅值。

- ▶ 就是在空间距离基础上加入了局部纹理距离。

#1: 实验结果

- ▶ 在FERET数据库的多姿态b集合验证了对姿态变化的容忍度



- ▶ 实验性能并不高，但考虑到这一方法不假设姿态角度已知这一结果还是能证明方法的有效性。

#2 第二篇文章(Oral)

文章标题: Joint and Implicit Registration for Face Recognition

作者: Peng Li and Simon. J. D. Prince

文章ID: 1029

相关文献:

- ▶ S. J. D. Prince and J. H. Elder. Probabilistic Linear Discriminant Analysis for Inferences About Identity. In ICCV, 2007.
- ▶ S. J. D. Prince, J. H. Elder, J. Warrell, and F. M. Felisberti. Tied Factor Analysis for Face Recognition across Large Pose Differences. PAMI, 30(6):970 - 984, 2008.

#2 第一作者

Li Peng

Research Fellow

Department of Computer Science

University College of London

<http://www.cs.ucl.ac.uk/staff/P.Li/>



学术背景:

- ▶ 分别在1993年和1998年在华北电力大学获得学士和硕士学位。2006年在新加坡南洋理工大学获博士学位。
- ▶ 1993-1995年保定电力研究所教师
- ▶ 1998-2002年华北电力大学讲师。
- ▶ 2007年布里斯托尔大学(University of Bristol)博士后。
- ▶ 2008年至今在伦敦大学学院任Research Fellow。

#2 第二作者

Simon J.D. Prince

Senior Lecturer

Department of Computer Science

University College of London

<http://www.cs.ucl.ac.uk/staff/s.prince/>



经历丰富的学术背景:

- ▶ S.Prince在伦敦大学学院(UCL)获**心理学**学士，1999年在剑桥大学获得**实验心理学**博士。然后在剑桥大学生理学实验室进行了两年博士后研究。
- ▶ 在2000年回到UCL攻读计算机视觉硕士。获得学位后，在新加坡国立大学电子和计算机工程系做两年半博士后研究。然后前往加拿大约克大学做博士后到2005年。
- ▶ 现在为UCL计算机系的高级讲师。是BMVC 2009的两位主席(General Chairs)之一。

#2: 第一个问题

- ▶ 否通过探索gallery图像来在probe图像上寻找特征点？—本文的答案是肯定的。
- ▶ 本文先用参考文献[16]的方法得到一个概率模型，给定一对分别在Gallery和Probe对应点上提取的特征向量 (x_g, x_p) 可以得到他们属于同一个人 \mathcal{M}_s 的概率：

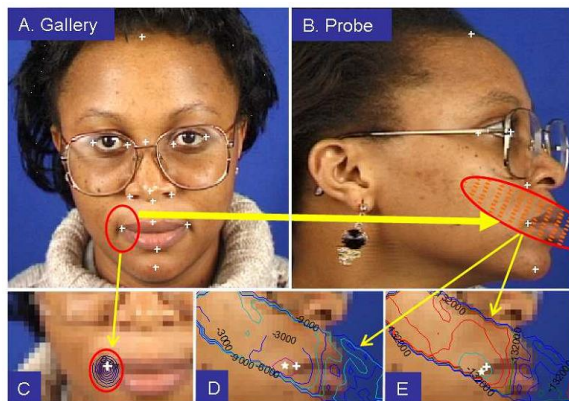
$$Pr(x_g, x_p | \mathcal{M}_s) = Pr(x' | \mathcal{M}_s)$$

- ▶ 其中 $x' = [x_g; x_p]$ 是串联起来的长向量。
- ▶ 如果假设gallery图像和probe图像属于同一个人那么所求的特征点可以通过下式求解：

$$t_p^*, t_g^* = \arg \max_{t_p, t_g} Pr(x' | \mathcal{M}_s, t_p, t_g) Pr(t_p, t_g)$$

- ▶ 如果假设不属于同一个人，则和只使用单幅图像的方法相同。
- ▶ 注意：定位的目标函数由降低通用的特征点模型误差变为了最大化属于同一人的概率。

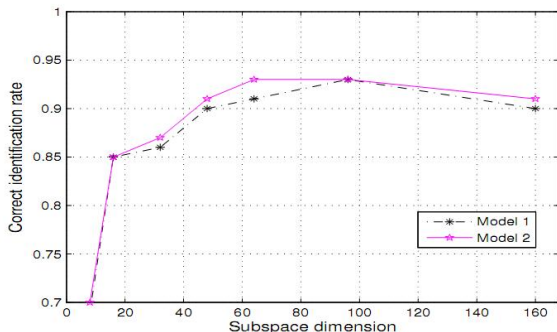
#2: 第一个问题(续)



A gallery图像 B probe图像 C gallery中嘴角点的分布 D 单独使用probe图像估计出的嘴角点分布 E 引入了gallery信息后估计的最焦点分布。五星点为估计的分布中心，十字点是真实位置。

#2: 第一个问题(续)

- ▶ 增加gallery图像的信息对在probe图像上定位特征点是有增益的。



- ▶ 上图model1使用单幅probe图像，model2中加入了gallery图像。

#2: 第二个问题

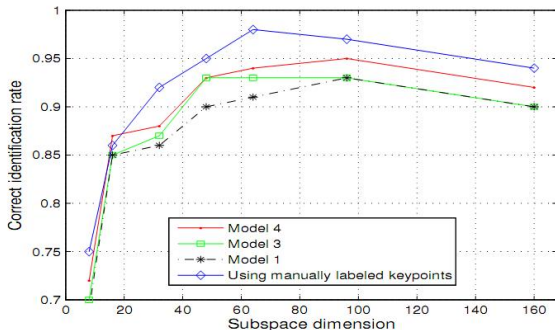
- ▶ 定位特征点对于识别而言是否是必需的？——本文的答案是否定的。
- ▶ 本文提出可以将特征点位置看做隐含的变量嵌入到识别过程中。
- ▶ 那么对于一对特征点对应的特征他们属于同一个人的概率为：

$$Pr(x_p, x_g) = \int \int Pr(x' | \mathcal{M}_s, t_p, t_g) Pr(t_p, t_g) dt_p dt_g$$

- ▶ 特征 x 和位置 t 是相关的，本文中 t 的先验分布是通过手工标定的训练集合上统计得到的。
- ▶ 本文的方法就是累加各个点上属于同一人的概率。

#2: 第二个问题(续)

- ▶ 通过下图正面人脸图像上的实验可以看出(显式的)定位特征点并不是必须的。



- ▶ 上图model1只考虑单幅图像,model3引入gallery信息显式估计特征点位置, model4没有显式估计特征点位置。
- ▶ 如果把参考文献[16]的PLDA模型替换成参考文献[17]的针对姿态的模型, 就可以变成一种不需要显式估计特征点的跨姿态人脸识别方法。

#3 第三篇文章(Oral)

文章标题: Implicit Elastic Matching with Random Projections for Pose-Variant Face Recognition

作者: John Wright and Gang Hua

文章ID: 0403

相关文献:

- ▶ Moosman, B. Triggs, and F. Jurie. Randomized clustering forests for building fast and discriminative visual vocabularies. In Proc. Neural Information Processing Systems.

◇ 这篇文章的工作是第一作者在08年暑期在Redmond的MS live Lab做实习时完成的。

#3: 第一作者

John Wright

PhD student

University of Illinois at Urbana-Champaign
will join the Visual Computing Group at MSRA

<http://perception.csl.illinois.edu/~jnwright/>



学术背景:

- ▶ 分别在2004和2007年在UIUC获计算机工程学士和电子工程硕士学位。
“将于” 2009年春在UIUC获得博士学位。
- ▶ 他的硕士和博士导师是马毅教授。

发表的文章:

- ▶ John Wright, Allen Yang, Arvind Ganesh, Shankar Sastry and Yi Ma.
"Robust Face Recognition via Sparse Representation," IEEE Transactions on Pattern Analysis and Machine intelligence, February 2009.

#3: 第二作者

Gang Hua

Researcher

Nokia Research Center, Hollywood

<http://www.ece.northwestern.edu/~ganghua/>



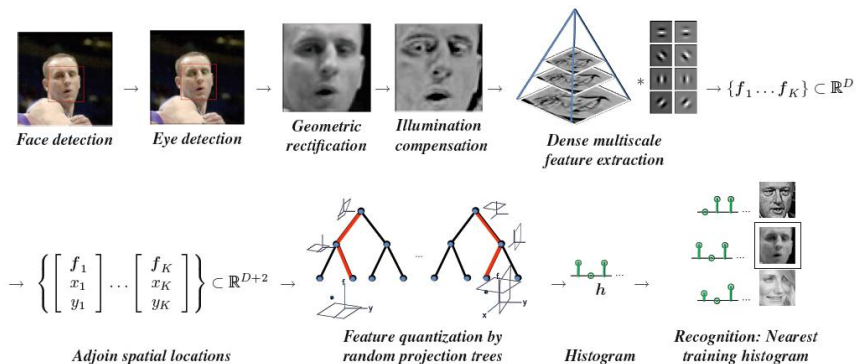
学术背景:

- ▶ 分别在1999和2002年在西安交通大学学士和硕士学位。于2006年在美国西北大学(Northwestern University)获博士学位。硕士和博士的导师分别是郑南宁院士和吴郢教授。
- ▶ 2006年到2009年4月在MS live lab。2009年4月到8月在"MS Bing Multimedia"。2009年8月至今在位于好莱坞的诺基亚研究中心，新用户界面组(New Interface Group)。

发表的文章:

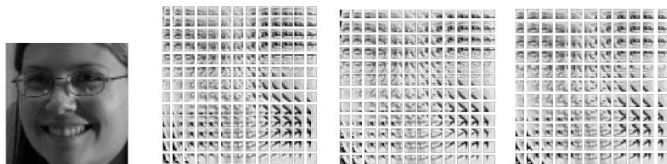
- ▶ 共发表过10篇期刊文章和28篇会议文章，其中有6篇PAMI。

#3: 本文的人脸识别流程



#3: 局部特征表示

- ▶ 首先对图像做高斯金字塔分解，稠密的采集多尺度的图像块(457个patch)。

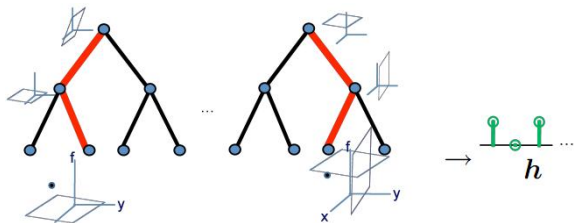


- ▶ 对每个patch用参考文献[28]的方法提取特征，得到一个400维的特征向量。
- ▶ 每个特征向量串联(x,y)坐标位置变成402维的向量。
- ▶ 这样对每幅 32×32 的人脸图像得到一个 402×457 的过完备表示。

[28] S. Winder and M. Brown. Learning local image descriptors. In Proc. CVPR, pages 1 - 8, 2007.

#3: 特征和空间位置联合量化

- ▶ 每个人脸对应的的一组特征向量通过若干决策树被量化成直方图。



- ▶ 树的每一个节点 v 表示一个随机的特征向量和一个阈值。
- ▶ 每一个输入的特征向量最终都会落到一棵树的叶子节点上。
- ▶ 每个叶子节点作为一个Bin，输入人脸就被量化成一个直方图。

#3: 构建随机投影树

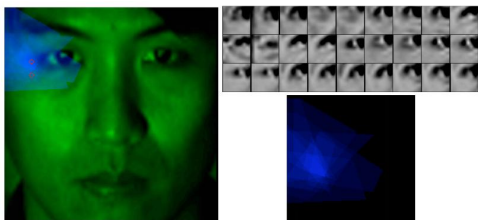
- ▶ 目标生成一个随机树构成的森林 $T_1 \dots T_k$ 。
- ▶ 每一个树独立的生成，具有确定的最大深度。
- ▶ 对随机树上的每个节点 v 生成一个随机向量 $w_v \in \mathbb{R}^{D+2}$ 和一个阈值 $\tau_i \leftarrow \text{median}\{\langle w_i, \tilde{f} \rangle | \tilde{f} \in \mathcal{X}\}$ 。
- ▶ 随机投影 w_v 从一个各项异性的高斯中采集。
 $w_v \sim \mathcal{N}(0, \text{diag}[\sigma_f^{-2} I_{D \times D}, \sigma_x^{-2} I_{2 \times 2}])$
- ▶ 其中 $\sigma_f^{-2} = \text{trace} \hat{\Sigma}(f)$, $\sigma_x^{-2} = \text{trace} \hat{\Sigma}(x, y)$ 。 $\hat{\Sigma}$ 是整个训练集合上的经验协方差、
- ▶ 选择这种分布等同对 \tilde{f} 进行重新加权，有助于平衡特征和空间位置之间的额关系。
- ▶ 通过**递归**的方式在训练集合上生成投影树。

#3: 构建随机投影树的算法流程

Algorithm: Tree induction (rptree)

- 1 输入: 增广特征向量 $\mathcal{X} = \{\tilde{f}_1 \dots \tilde{f}_m\}$, $\tilde{f}_i = (f_i, x_i, y_i) \in \mathbb{R}^{D+2}$
- 2 计算特征和坐标的方差 σ_f^2 和 σ_x^2
- 3 依照分布生成 $p \geq D + 2$ 个随机投影:
 $W \sim_{iid} \mathcal{N}(0, \text{diag}(\sigma_f^{-2} \dots \sigma_f^{-2}, \sigma_x^{-2}, \sigma_x^{-2}))$
- 4 重复k次:
- 5 采样一个投影 $i \sim \text{uni}(1 \dots p)$
- 6 计算阈值 $\tau_i \leftarrow \text{median}\{\langle w_i, \tilde{f} \rangle | f \in \mathcal{X}\}$
- 7 用这个阈值划分左右子树 $\mathcal{X}_L \leftarrow \{\tilde{f} | \langle w_i, \tilde{f} \rangle < \tau_i\}$, $\mathcal{X}_R \leftarrow \mathcal{X} \setminus \mathcal{X}_L$
- 8 计算 $r_i \leftarrow |\mathcal{X}_L| \text{diameter}^2(\mathcal{X}_L) + |\mathcal{X}_R| \text{diameter}^2(\mathcal{X}_R)$
- 9 end循环结束
- 10 选择使得 r 最小的 (w^*, τ^*)
- 11 用其作为根节点 $\text{root}(T) \leftarrow (w^*, \tau^*)$
- 12 用根节点划分左右子树: $\mathcal{X}_L \leftarrow \{\tilde{f} | \langle w_i, \tilde{f} \rangle < \tau_*\}$, $\mathcal{X}_R \leftarrow \mathcal{X} \setminus \mathcal{X}_L$
- 13 对左子树递归: $\text{leftchild}(T) \leftarrow \text{rptree}(\mathcal{X}_L)$
- 14 对右子树递归: $\text{rightchild}(T) \leftarrow \text{rptree}(\mathcal{X}_R)$
- 15 输出最终结果 T .

#3: 隐含的匹配弹性



- ▶ 左图:在PIE数据库上学出的10棵树, 对于其中的一个Bin。将左眼角点变换到图像的每一个点。
- ▶ 蓝色灰度表示和左眼角点匹配上的树的数目。
- ▶ 右上图:在最小的三棵树中量化到同一个Bin中的patch。
- ▶ 右下图:Bin的数目。
- ▶ 量化器具有"软的"空间约束, 使得落入bin中的patch都是眼角区域。
- ▶ 同时显示使用多个树构成的森林有平滑的作用。

#3: 实验结果

- 在几个比较简单的数据集上有不错的效果。

	ORL	Ext. Yale B	PIE	MultiPIE
PCA	88.1%	65.4%	62.1%	32.6%
LDA	93.9%	81.3%	89.1%	37.0%
LPP	93.7%	86.4%	89.2%	21.9%
This work	96.5%	91.4%	94.3%	67.6%

Table 3. Recognition rates across various datasets.

- 本文方法比较突出的是具有良好的推广性。跨数据库测试具有比较好的效果。

	PIE → ORL	ORL → PIE	PIE → MultiPIE
PCA	85.0%	55.7%	26.5%
LDA	58.5%	72.8%	8.5%
LPP	17.0%	69.1%	17.1%
This work	92.5%	89.7%	67.2%

Table 4. Recognition rates for transfer across datasets.

#3: 实验结果(续)

- ▶ 在LFW数据库上的实验表明本文方法具有良好的推广性。

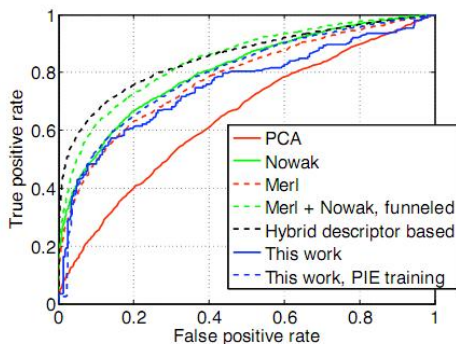


Figure 6. Receiver Operating Characteristic for Labeled Faces in the Wild. “Nowak” refers to [20]. “Hybrid descriptor based” refers to [30].

- ▶ 本文的方法可以接近目前在这个数据库上比较好的结果。尤其是在用PIE数据库做训练的情况下。这显示该方法对数据的依赖性不高。

