

文章编号: 1001-3806(2006)04-0429-03

采用 CASPCM 模型进行姿势鲁棒性人脸识别

赵明华, 游志胜*, 余 静, 熊运余

(四川大学 计算机学院 图形图像研究所, 成都 610064)

摘要: 针对 ASPCM 模型处理转动角度较大的人脸图像时出现的不足, 提出 CASPCM 模型。以样本与模型中心的距离为依据将训练样本分组, 为每个分组训练 ASPCM 模型; 将局部 ASPCM 模型的合成映射结果加权平均得到 CASPCM 模型的合成结果; 提出利用梯度下降法使分解映射的姿势估计逐步精确。采用精确性和概括性两个标准衡量该模型的分解性能和合成性能。实验表明, CASPCM 模型的分解性能和合成性能均优于 ASPCM 模型; 基于该模型的人脸识别系统在处理转动角度较大的人脸图像时, 识别率比 ASPCM 模型高 7%。

关键词: 图像处理; 人脸识别; 形状表示; 纹理表示; 分解合成主成分映射; 组合的分解合成主成分映射

中图分类号: TP391.41 文献标识码: A

Pose robust face recognition based on CASPCM model

ZHAO Ming-hua, YOU Zhi-sheng, YU Jing, XIONG Yun-yu

(Institute of Image & Graphic, College of Computer Science, Sichuan University, Chengdu 610064, China)

Abstract CASPCM model is proposed to make up the disadvantages of ASPCM model while dealing with faces with large angles. The training samples are grouped according to their distances to model centers and a local ASPCM model is constructed for each group. Synthesis result of CASPCM model is obtained by averaging results of the local ASPCM models with appropriate weights. Gradient descent algorithm is used to iteratively improve estimate of the head pose. Accuracy and generalization are used to gauge analysis and synthesis abilities of the model. Experiments show that the two abilities of CASPCM model are both superior to ASPCM model, recognition ratio of CASPCM model is 7% higher than ASPCM model.

Key words image processing; face recognition; shape representation; texture representation; analysis synthesis; principal component mapping (ASPCM); combined analysis synthesis; principal component mapping (CASPCM)

引 言

头部姿势、面部表情和光照等变化都会造成人脸图像变化^[1]。头部姿势由于多变且不易控制而成为影响人脸识别算法性能的重要因素。一个理想的人脸识别算法应对头部姿势变化具有较强的鲁棒性^[2]。

人脸识别算法中处理姿势变化的方法主要包括单视图法^[3]、多视图法^[4]和非最近邻域法^[5] 3种。OKADA 等提出 ASPCM (analysis synthesis principal component mapping) 模型, 并且实验证明该模型在进行头部转动不超过 20° 的人脸识别时准确度较高^[6]。作者在分析 ASPCM 模型的基础上, 提出 CASPCM (combined analysis synthesis principal component mapping) 模型, 该模型能更好地适应人脸姿势变化的非线性特征, 适用于头部转动角度较大的人脸识别。

1 ASPCM 模型

ASPCM 模型的基本思想是由 M 个训练样本 $\{(\mathbf{v}^1, \theta^1), \dots, (\mathbf{v}^m, \theta^m), \dots, (\mathbf{v}^M, \theta^M)\}$ 确定人脸图像的三维角度与面部矢量的映射关系。其中, (\mathbf{v}^m, θ^m) 表示第 m 个训练样本, \mathbf{v}^m 是面部矢量, $\theta^m = (\theta_1^m, \theta_2^m, \theta_3^m)$ 是图像的三维头部角度矢量。从三维角度到面部矢量的映射为合成映射, 反之为分解映射。

上述模型由 $\mathbf{u}_x, \mathbf{u}_j^1, \dots, \mathbf{u}_j^N, \mathbf{u}_0, Y, B^1, \dots, B^N, F, G, H^1, \dots, H^N, A, SS, TS$ 表示。其中, $\mathbf{u}_x = \frac{1}{M} \sum_{m=1}^M \mathbf{x}^m$, $\mathbf{u}_j^n = \frac{1}{M} \sum_{m=1}^M \mathbf{j}^{m,n}$, \mathbf{x}^m 是 \mathbf{v}^m 的形状表示, $\mathbf{j}^{m,n}$ 是 \mathbf{v}^m 在关键点 n 的纹理表示^[7]; $\mathbf{u}_0 = (u_{01}, u_{02}, u_{03}) = \frac{1}{M} \sum_{m=1}^M \theta^m$; Y 是形状模型, $Y = (\mathbf{y}^1, \dots, \mathbf{y}^p, \dots, \mathbf{y}^P)^T$, \mathbf{y}^p 是将 $\mathbf{x}^1, \dots, \mathbf{x}^m, \dots, \mathbf{x}^M$ 进行主成分分析得到的第 p 个形状主成分; B^n 是局部纹理模型, $B^n = (\mathbf{b}^{1,n}, \dots, \mathbf{b}^{s,n}, \dots, \mathbf{b}^{S,n})^T$, $\mathbf{b}^{s,n}$ 是将关键点 n 处的纹理表示 $\mathbf{j}^{m,1}, \dots, \mathbf{j}^{m,n}, \dots, \mathbf{j}^{m,N}$ 进行主成分分析所得到的第 s 个纹理主成分; F 是形状参

作者简介: 赵明华 (1979-), 女, 博士研究生, 主要从事数字图像处理和模式识别等研究。

* 通讯联系人。E-mail: zsyoud@mail.sc.cninfo.net

收稿日期: 2005-05-23 收到修改稿日期: 2005-07-21

数到姿势参数的转换矩阵; G 是姿势参数到形状参数的转换矩阵; H^n 是关键点 n 处纹理参数到形状参数的转换矩阵; $A: v^m \rightarrow \theta^n$ 是分解映射; $SS: \theta^n \rightarrow x^m$ 是从姿势到形状的合成映射; $TS: \theta^n \rightarrow (j^{m,1}, \dots, j^{m,N})$ 是从姿势到纹理的合成映射^[6]。

2 CASPCM 模型

ASPCM 模型用于头部转动超过 20° 的人脸识别时, 因用线性映射处理人脸姿势变化的非线性特征, 准确度较低^[6]。图 1a 中黑点表示在 N 维空间 R^N 中分布的样本, 虚线表示数据变化的非线性特征。图 1b 表示 $n=2$ 时的向量空间 R^2 , e_1 和 e_2 是两个特征向量。虽然样本可以投影到 e_1 和 e_2 张成的向量空间^[8], 但其不能精确反映样本的非线性变化特征。图 1c 表示 CASPCM 模型将数据变化的分布空间划分为局部子空间 (R_1^2, R_2^2, R_3^2), 较精确地描述了这种非线性变化。即通过组合一系列局部 ASPCM 模型以覆盖足够大的三维角度。

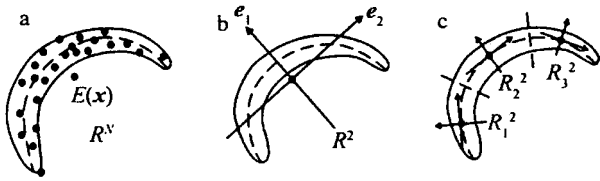


Fig 1 Non-linear sample distribution and CASPCM approach

2.1 模型 的表示

假设由一个人的 K 组样本训练出 K 个 ASPCM 模型 $\{L_1, \dots, L_{l_b}, \dots, L_K\}$ 。其中, L_k 表示第 k 个局部 ASPCM 模型。将每个局部模型的所有训练样本的平均姿势记为 $u_0^{L_k}$, 并称其为 L_k 的模型中心。CASPCM 模型的分解映射可表示为:

$$\theta = \sum_{k=1}^K w_k(\theta) \theta_k = \sum_{k=1}^K w_k(\theta) A_{L_k}(v) \quad (1)$$

式中, θ 表示由面部矢量分解得到的三维角度; $w_k(\theta)$ 表示模型 L_k 的权; θ_k 表示用模型 L_k 对样本 v 分解的结果; A 表示分解映射。同理, CASPCM 模型的合成映射可表示为:

$$\hat{v} = D \left[\sum_{k=1}^K w_k(\theta) \hat{x}_k, \sum_{k=1}^K w_k(\theta) \hat{j}_{l_b}^1, \dots, \sum_{k=1}^K w_k(\theta) \hat{j}_k^N \right] = D \left[\sum_{k=1}^K w_k(\theta) SS_{L_k}(\theta), \sum_{k=1}^K w_k(\theta) TS_{L_k}(\theta) \right] \quad (2)$$

式中, \hat{v} 表示由三维角度合成得到的面部矢量; D 表示 $(x, j^1, \dots, j^N) \rightarrow v$ 运算; \hat{x}_k 和 $\hat{j}_{l_b}^1, \dots, \hat{j}_k^N$ 分别表示由 L_k 合成的形状表示和纹理表示; SS 表示姿势到形状的合成映射; TS 表示姿势到纹理的合成映射。

构造权函数为:

$$w_k(\theta) = \tilde{n}_k(\theta - u_0^{L_k}) \prod_{k=1}^K \tilde{n}_k(\theta - u_0^{L_k}) \quad (3)$$

式中, $\tilde{n}_k(\theta) = \frac{1}{\sqrt{2\pi}\sigma_k} \exp\left[-\frac{\|\theta\|^2}{2\sigma_k^2}\right]$; σ_k 是局部模型训练样本的三维头部角度的标准方差的函数, $\sigma_k =$

$p \times \frac{1}{\sqrt{M_k - 1}} \sum_{m=1}^{M_k} \|\theta^m - u_0^{L_k}\|^2$; p 是系数; M_k 是用于训练 L_k 的样本个数。该权函数是输入图像与局部模型中心距离的函数, 且函数值随着该距离的增加而减小。考虑到局部 ASPCM 模型的性能随着输入图像与模型中心距离的增加而急剧降低^[6], 公式使用三维角度空间的标准高斯函数。

2.2 梯度下降算法

由 (3) 式知, 权函数是输入样本三维角度 θ 的函数, 故其不可用于分解映射。本文中使用梯度下降法^[9]解决此问题。用 x_i 和 θ_i 表示算法的第 i 步迭代结果。初始化头部角度和面部矢量为:

$$x_0 = u_{k_{min}}^{L_{k_{min}}}, \theta_0 = u_0^{L_{k_{min}}} \quad (4)$$

式中, k_{min} 是平均形状最接近于输入图像 x 的局部模型的序号; $u_{k_{min}}^{L_{k_{min}}}$ 和 $u_0^{L_{k_{min}}}$ 分别该局部模型的所有训练样本的平均形状和平均姿势。迭代过程如下:

$$\Delta x_i = x - x_i, \Delta \theta_i = \sum_{k=1}^K w_k(\theta_i) A_{L_k}(\Delta x_i) \quad (5)$$

$$\theta_{i+1} = \theta_i + 0.01 \times \Delta \theta_i, x_{i+1} = \sum_{k=1}^K w_k(\theta_{i+1}) SS_{L_k}(\theta_{i+1}) \quad (6)$$

迭代 (5) 式和 (6) 式, 直至均方误差 $\|\Delta x_i\|^2$ 足够小。

3 实验与分析

3.1 权函数测试

由 (3) 式知, 权的大小与 p 值有关。本文中选取特征点信息确定的人脸图像对权函数进行测试。实验结果见图 2 图 2a 和图 2b 分别表示使用 8 个形状主成分和不同 p 值确定的系统的分解和合成误差。由此可知, 当 $p=1.2$ 时, 分解误差最小; $p=0.9$ 时, 合成误差最小。本文中确定 $p=1.0$

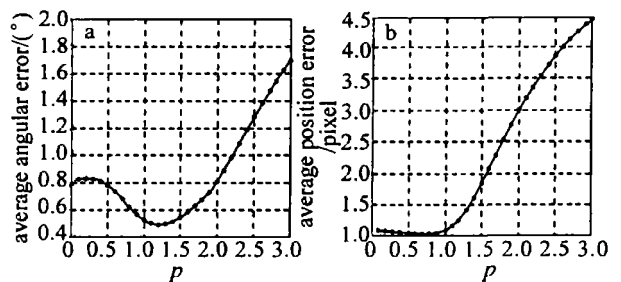


Fig 2 Comparison of weight functions determined by different p

3.2 分解合成性能测试

用 ATR 人脸库中的 20 个人脸模型测试 CASPCM 模型和 ASPCM 模型的分解、合成性能。分别以 $(0^\circ, 0^\circ, 0^\circ)$, $(\pm 30^\circ, 0^\circ, 0^\circ)$, $(0^\circ, \pm 30^\circ, 0^\circ)$, $(0^\circ, 0^\circ,$

±30°)为模型中心,为每个人产生 7 组训练样本,每组包含模型中心 15°内的 400 个训练样本。用每组训练样本构造局部 ASPCM 模型,用这些局部模型构造 CASPCM 模型;用所有训练样本构造全局 ASPCM 模型。如图 3 所示。

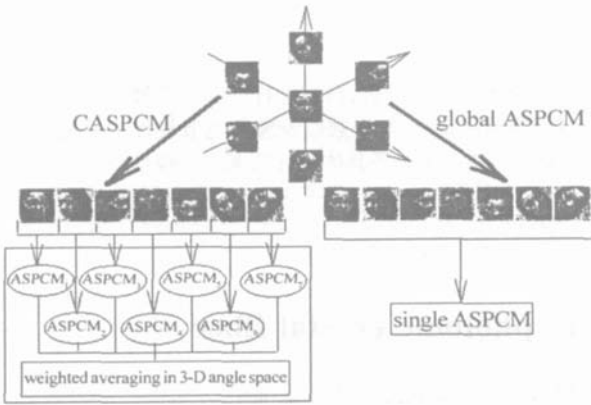


Fig 3 Sketches of CASPCM model and ASPCM model

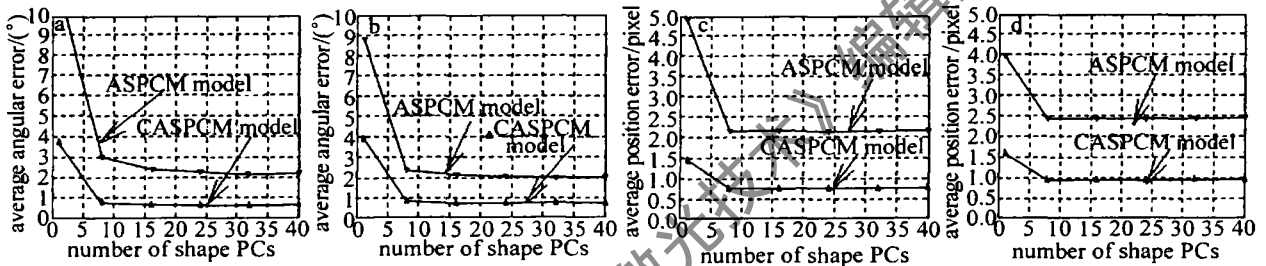


Fig 4 Comparison of the abilities of CASPCM model and ASPCM model

示合成映射时两种模型的精确性和概括性。由图可知,当用 8 个形状主成分时,分解性能和合成性能的精确性误差分别是 0.8°, 3.0°和 0.8 个像素、2.2 个像素;分解性能和合成性能的概括性误差分别是 0.9°, 2.4°和 0.9 个像素、2.4 个像素。CASPCM 模型大大提高了 ASPCM 模型的分解、合成性能。

3.3 人脸识别性能测试

每个人头部旋转角度在 ±50°范围内产生 800 个测试样本,对 3.2 节中训练的两种模型的人脸识别性能进行测试。将测试样本分成两组,第 1 组样本的头部转动角度未知,首先利用模型的分解映射从形状估计姿势参数,再利用合成映射从姿势参数估计形状参数和纹理参数;第 2 组样本的头部转动角度已知,可直接利用模型的合成映射从姿势参数估计形状参数和纹理参数。识别结果见表 1,两种情况下 CASPCM 模型的识别率均比 ASPCM 模型高 7% 以上。

Table 1 Comparison of recognition ratio of CASPCM model and ASPCM model

test set	CASPCM model/%	ASPCM model/%	improve/%
the first set	98.5	91.3	+ 7.2
the second set	99.4	92.4	+ 7.0

构造误差函数如下:

$$\epsilon_1 = \sum_{m=1}^M \left| \theta^m - \hat{\theta}^m \right| + \sum_{m=1}^M \left| v^m - \hat{v}^m \right| \quad (7)$$

$$\epsilon_2 = \sum_{i=1}^I \left| \theta^i - \hat{\theta}^i \right| + \sum_{i=1}^I \left| v^i - \hat{v}^i \right| \quad (8)$$

(7)式和(8)式中, $(v^i, \theta^i) \in \{(v^1, \theta^1), \dots, (v^m, \theta^m), \dots, (v^M, \theta^M)\}$; θ^m 和 \hat{v}^m 分别表示由第 m 个面部矢量分解的三维角度和由第 m 个三维角度合成的面部矢量; ϵ_1 是将模型用于训练样本时产生的误差,表示模型的精确性; ϵ_2 是将模型用于非训练样本时产生的误差,表示模型的概括性。

取 $p=1.0$ 用平均法^[10]处理不确定关键点覆盖率为 5% 的样本,分别测试 CASPCM 模型和 ASPCM 模型的分解、合成性能,结果见图 4。图 4a 和图 4b 表示分别用 1, 8, 16, 24, 32 和 40 个形状主成分进行分解映射时两种模型的精确性和概括性;图 4c 和图 4d 分别表

4 结论

提出 CASPCM 模型并将其用于人脸识别。该模型为每个局部范围的训练样本构造 ASPCM 模型,并加权平均这些局部模型。构造权函数并确定其最优系数,使用梯度下降算法解决分解映射的姿势估计问题及选用平均法处理具有不确定关键点的人脸图像是构造该模型的关键。用精确性和概括性两个标准衡量模型的分解性能和合成性能。实验表明,在处理头部转动角度较大的人脸图像时,该模型的性能明显优于 ASPCM 模型,以该模型为基础的人脸识别系统对姿势变化具有更强的鲁棒性。

参考文献

- [1] LU Zh F, YOU Zh S, WANG Y Q. Face recognition based on PCA and ICA [J]. Laser Technology, 2004, 28(1): 78~ 81 (in Chinese).
- [2] KOUZAN I A, ZHE F, SAMMUT K. Towards invariant face recognition [J]. Information Sciences, 2000, 123(1): 75~ 101.
- [3] LEEM W, SURENDRA R. Pose-invariant face recognition using a 3-D deformable model [J]. Pattern Recognition, 2003, 36(8): 1835~ 1846.
- [4] LA I JH, YUEN P C, FENG G C. Face recognition using holistic Fourier-invariant features [J]. Pattern Recognition, 2001, 34(1): 95~ 109.
- [5] EICKELER S, MULLER S, RIGOLL G. Recognition of JPEG-compressed face images based on statistical methods [J]. Image and Vision Computing, 2001, 18(4): 279~ 287. (下转第 435 页)

的变化而发生变化,而只由相对孔径决定。色散的模拟结果如图 6 所示。

波导色散同样有一个最大值和最小值,对于一定的相对孔径 f , 空气孔间距 a 增大时,色散曲线在波长轴方向伸展,而在色散方向上收缩。同样,可以通过空气孔间距的变化来控制色散。

3 小 结

应用平面波法,结合整胞法,模拟求出了光子晶体光纤的基空间填充模有效折射率 n_{eff} , 模式参数 V , 阐明了光子晶体光纤的无尽单模性质。在此基础上利用样条插值微分,求出了群速度色散,并着重分两种情况对光子晶体光纤的群速度色散进行了详细的分析和讨论。

分析表明,在空气孔间距 a 确定的情况下,相对孔径 f 越大,有效折射率越小,而 V 参数越大。相对孔径 f 越小,有效折射率越大,而 V 参数越小。波导色散有一个最大值和最小值,相对孔径增大时,色散曲线在两个轴方向都将伸展。

在相对孔径 f 一定的情况下,空气孔间距 a 越大,有效折射率越大。 V 参数并不随空气孔间距 a 的变化而发生变化,只与相对孔径有关。随着空气孔间距的增大,色散曲线在波长轴方向伸展,而在色散轴方向上收缩。

参 考 文 献

- [1] NIGHT J C, RUSSELL P S. New way to guide light [J]. *Science* 2002, 296(5566): 276~ 277.
- [2] NIGHT J C, BRKST A, RUSSELL P J. All-silica single mode optical fiber with photonic crystal cladding [J]. *Opt Lett* 1996, 21(19): 1547~ 1549.
- [3] NIGHT J C, BROENG J. Photonic band gap guidance in optical fibers [J]. *Science* 1998, 282(5393): 1476~ 1478.
- [4] YEH P, YARIV A, MAROM E. Theory of Bragg fiber [J]. *J O S A*, 1978, 68(9): 1196~ 1201.

- [5] BIRK T A, ROBERT P J, RUSSELL P J. Full 2-D photonic bandgaps in silica/air structure [J]. *Electron Lett* 1995, 31(22): 1941~ 1943.
- [6] BUCZYNSKI R. Photonic crystal fibers [J]. *Acta Physica Polonica* 2004, A106(2): 141~ 167.
- [7] PAN Y Zh, ZHANG J, HU G J *et al*. Photonic crystal fiber and laser [J]. *Laser Technology*, 2004, 28(1): 48~ 51(in Chinese).
- [8] LIM H, LINDING F O, WISE F W. Femtosecond ytterbium fiber laser with photonic crystal fiber for dispersion control [J]. *Optical Express* 2002, 10(25): 1497~ 1502.
- [9] LIMPERT J, SCHREIBER T, NOLTE S. All fiber chip ed-pulse amplification system based on compression in air-guiding photonic bandgap fiber [J]. *Optical Express* 2003, 11(24): 3332~ 3337.
- [10] CHAMPERT PA, POPOV S V, TAYLOR JK. Generation of multi-watt broadband continua in holey fibers [J]. *Opt Lett* 2002, 27(2): 122~ 124.
- [11] BJARKLEV A, BROENG J, DRIDI K. Dispersion properties of photonic crystal fiber [A]. *ECOC'98 [C]*. European ECOC, 1998. 135~ 136.
- [12] NIGHT J C, ARRIAGA J, BIRK T A. Anomalous dispersion in photonic crystal fiber [J]. *IEEE Photonics Technology Letters* 2000, 12(7): 807~ 809.
- [13] MOGILEVTSYEV D, BIRK T A, RUSSELL P J. Group velocity dispersion in photonic crystal fiber [J]. *Opt Lett* 1998, 23(21): 1662~ 1664.
- [14] REN G B, LOU Sh Q, WANG Zh. Study on dispersion properties of photonic crystal fiber by equivalent index model [J]. *Acta Optica Sinica* 2004, 24(3): 319~ 323(in Chinese).
- [15] HO K M, CHAN C T, SOUKOULIS C M. Existence of a photonic gap in periodic dielectric structures [J]. *Phys Rev Lett* 1990, 65(25): 3152~ 3155.
- [16] LEUNG K M. Plane wave calculation of photonic band structures in photonic band gaps and localizations [M]. New York: Plenum Press 1993, 89~ 119.
- [17] HUANG K. Solid physics [M]. 2nd ed. Beijing: Advanced Education Press 1988, 154~ 157(in Chinese).
- [18] CHAMPENEY D C. Fourier transforms and their physical applications [M]. 10th ed. New York: Academic Press 1973, 154~ 157.
- [19] LEUNG K M, LIU Y F. Full vector wave calculation of photonic band structures in face centered cubic dielectric media [J]. *Phys Rev Lett* 1990, 65(21): 2646~ 2649.

(上接第 431 页)

- [6] OKADA K, AKAMATSU S, VON D M C. Analysis and synthesis of pose variations of human faces by a linear PCMAP model and its application for pose-invariant face recognition system [A]. *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition [C]*. USA: IEEE Computer Society 2000. 142~ 149.
- [7] TROGEN F, BULTHOFF H. Face recognition under varying poses: the role of texture and shape [J]. *Vision Research*, 1996, 36(12): 1761~ 1771.

- [8] RICHARD Q, PETER E, DAVID G. Pattern classification [M]. 2nd ed. Beijing: China Machine Press 2003, 456~ 457(in Chinese).
- [9] BIAN Z Q, ZHANG X G. Pattern recognition [M]. 2nd ed. Beijing: Tsinghua University Press 2000, 331~ 332(in Chinese).
- [10] TRESP V, HOFMANN R. Nonlinear time-series prediction with missing and noisy data [J]. *Neural Computation* 1998, 10(3): 731~ 747.