

基于 MAP 准则的两步人脸图像分辨率增强算法

李 阳, 林学焱

(清华大学计算机科学与技术系普式计算教育部重点实验室, 北京 100084)

摘 要: 如何提高人脸图像分辨率是改进视觉监视系统性能需要解决的关键问题之一. 本文提出了一种基于 MAP 准则的两步人脸图像分辨率增强新算法. 第一步利用 PCA 模型和最大后验概率 (MAP) 方法计算人脸的整体图像, 第二步利用 Markov 随机场模型和 MAP 方法计算反映真实图像与整体图像差异的残差图像. 最终的结果是第一步得到的整体图像与第二步得到的残差图像之和. 利用新算法我们分别对图像分辨率增强 16 倍 (由 32 × 24 提高到 128 × 96) 和图像分辨率增强 64 倍 (由 16 × 12 提高到 128 × 96) 两种情况进行了实验, 均取得了令人满意的结果.

关键词: 人脸图像; 分辨率增强; MAP 准则

中图分类号: TP307 **文献标识码:** A **文章编号:** 0372-2112 (2004) 12A-192-04

MAP-Based Two-Step Approach to Hallucinating Faces

LI Yang, LIN Xue-yin

(Department of Computer Science and Technology, Tsinghua University Key Laboratory of Pervasive Computing, Tsinghua University, Ministry of Education, Beijing 100084, China)

Abstract: Face hallucination is to synthesize a high-resolution facial image from a low-resolution input. A novel two-step approach to hallucinating faces based on the MAP criterion is presented in this paper. First, a linear relationship between high-resolution and low-resolution facial images is established by applying PCA on both of them, and the global image, which is similar to the original high-resolution image, is reconstructed based on the MAP criterion. Second, a linear model between the residual image (the difference between the original image and the global image) and the low-resolution residual image (the difference between the low-resolution input and the manually down-sampled global image) is built, and, following a MRF prior, the optimal residual image is estimated based on the MAP criterion again. Experiments demonstrate that our approach can be applied to yield 4 - 8 fold super-resolution with high-quality hallucinated results.

Key words: face hallucination; resolution enhancement; MAP criterion

1 引言

在视觉监视系统中,受摄像机与被观测对象之间距离的限制,拍摄到的人脸图像分辨率可能较低,这给后续的人脸识别等工作带来了一定的困难.因此,如何提高人脸图像分辨率是改进视觉监视系统性能需要解决的关键问题之一.设 I_L 、 I_H 分别表示低分辨率和高分辨率的人脸图像,造成人脸图像分辨率降低的退化模型可以表示为

$$I_L = \frac{1}{s^2} \sum_{k=0}^{s-1} \sum_{l=0}^{s-1} I_H(si+k, sj+l) + n(i, j) \quad (1)$$

其中 s 是正整数,表示分辨率降低倍数, n 是随机噪声.为简单起见,采用矩阵向量形式表示,设 I_L 为 M 维列向量, I_H 为 N 维列向量 ($M = N/s^2$),则上述退化模型可简化为:

$$I_L = AI_H + n \quad (2)$$

其中 A 为 $M \times N$ 维下采样矩阵.需要解决的问题是在已知 I_L

的条件下求出最优的 I_H .

解决这一问题的常用方法是 MAP 方法,即:

$$I_H^* = \arg \max_{I_H} P(I_H | I_L) = \arg \max_{I_H} P(I_L | I_H) P(I_H) \quad (3)$$

在求解式时需要要对 I_H 建立模型,以表示其先验概率 $P(I_H)$. Schultz 和 Stevenson^[3]以及 Hardie 等人^[4]采用 Markov 随机场模型, Gunturk 等人^[6]采用 PCA 模型, Baker 和 Kanade^[2]则采用图像金字塔的方法建立了 $P(I_H)$ 的推导模型.

Freeman 和 Pasztor^[5]利用参数化 Markov 网络对高低分辨率图像建模,并提出了一种提高图像分辨率的普遍方法.这种方法适用于所有图像,但由于没用利用人脸图像的特点,因此其结果不如某些专门对人脸图像设计的算法.

Liu 等人^[1]提出了一种两步人脸图像分辨率增强算法.第一步利用 PCA 对 I_H 建模,并利用 MAP 方法计算整体图像 I_H^g ; 第二步利用非参数 Markov 网络对真实高分辨率图像 I_H 与 I_H^g

之间的残差图像 $I_H^r = I_H - I_H^g$ 建模,并利用图像块采样 (Patch-Based Sampling) 的方法计算 I_H^r . 最终结果由式 (4) 给出,这种方法取得了较好的结果. $I_H = I_H^g + I_H^r$ (4)

我们的方法是对 Liu 等人^[1]算法的改进,同样采用两步策略.在第一步中,我们分别对高低分辨率图像建立 PCA 模型,以便建立更接近实际情况的噪声模型.对第二步我们做了较大的改进,在 Liu 等人的算法中, I_H^r 某个图像块的取值受训练样本个数的限制.当训练样本数量较少或选取不合理时,这种方法无法得到合理的残差图像.我们的方法则利用了 I_H 与 $(I_L - AI_H^g)$ 存在线性关系这一特点,利用 Markov 随机场对 I_H^r 建模,并利用 MAP 方法最优计算 I_H^r . 实验结果表明,新方法的结果优于 Liu 等人的方法.

本文后面章节的内容如下:第 2 节详细说明新方法;第 3 节给出实验结果并与 Liu 等人方法的结果进行比较;最后是对我们工作的总结和对今后工作的展望.

2 方法描述

2.1 计算整体图像

主分量分析 (PCA) 是给人脸图像建立模型的有效方法,为计算整体图像,我们利用 PCA 分别对高低分辨率人脸图像建模,得

$$\begin{aligned} X_H &= B_H^T (I_H - \mu_H) & I_H &= B_H X_H + \mu_H + e_H \\ X_L &= B_L^T (I_L - \mu_L) & I_L &= B_L X_L + \mu_L + e_L \end{aligned} \quad (5)$$

其中 X_H, X_L 分别表示高低分辨率图像在 PCA 空间中的特征向量, B_H, B_L 为 PCA 空间基矩阵, μ_H, μ_L 为均值向量, e_H, e_L 为重构误差.代入式 (2),考虑到 $\mu_L = A\mu_H$,得

$$B_L X_L + e_L = AB_H X_H + Ae_H + n \quad (6)$$

根据 PCA 性质可知重构误差与 PCA 空间的基正交,即 $B_L^T e_L = 0$,式 (6) 左右同乘 B_L^T 得

$$X_L = B_L^T AB_H X_H + B_L^T (Ae_H + n) \quad (7)$$

上式表明 X_L, X_H 之间存在线性关系.对输入的低分辨率图像 I_L ,可首先由式 (5) 求出其特征向量 X_L ,再基于 MAP 准则由式 (8) 求出最优的 X_H ,最后由式 (9) 重构出整体图像 I_H^g

$$X_H^* = \arg \max_{X_H} P(X_H | X_L) = \arg \max_{X_H} P(X_L | X_H) P(X_H) \quad (8)$$

$$I_H^g = B_H X_H^* + \mu_H \quad (9)$$

为求解式 (8),需要估计先验概率 $P(X_H)$ 和条件概率 $P(X_L | X_H)$,其中 X_H 服从高斯分布

$$P(X_H) = \frac{1}{C_1} \exp \left\{ -\frac{1}{2} X_H^T \Sigma_H^{-1} X_H \right\} \quad (10)$$

这里 C_1 是常数, $\Sigma_H = \text{diag}(S_H^{2(1)}, \dots, S_H^{2(L)})$, $S_H^{2(1)}, \dots, S_H^{2(L)}$ 是计算高分辨率图像 PCA 时得到的最大的 L 个特征值.而 $P(X_L | X_H)$ 则取决于 $B_L^T (Ae_H + n)$ 的分布,令 $u = B_L^T (Ae_H + n)$,并假设其服从高斯分布 $N(\mu_u, \Sigma_u)$,参数 μ_u, Σ_u 可通过训练样本估计,则由式 (7) 得

$$P(X_L | X_H) = \frac{1}{C_2} \exp \left\{ -\frac{1}{2} (X_L - B_L^T AB_H X_H - \mu_u)^T \Sigma_u^{-1} (X_L - B_L^T AB_H X_H - \mu_u) \right\} \quad (11)$$

将式 (10) 和式 (11) 代入式 (8),经过推导可计算出

$$X_H^* = (B_H^T A^T B_L \Sigma_u^{-1} B_L^T AB_H + \Sigma_H^{-1})^{-1} B_H^T A^T B_L \Sigma_u^{-1} (X_L - \mu_u) \quad (12)$$

2.2 计算残差图像

整体图像 I_H^g 丢弃了人脸的一些细节信息,这些细节信息对于标识人的身份可能是十分重要的.为此,算法的第二步需要计算出表示整体图像与真实图像之间差异的残差图像 I_H^r .如果把 I_H^g 手工降低分辨率得到 AI_H^g ,它与 I_L 之间存在着一定的差异,这个差异反映了一定的细节特征.基于上述考虑,定义低分辨率残差图像 I_L^r

$$I_L^r = I_L - AI_H^g \quad (13)$$

将式 (2) 代入式 (13) 得

$$I_L^r = I_L - AI_H^g = AI_H + n - AI_H^g = A(I_H - I_H^g) + n = AI_H^r + n \quad (14)$$

上式表明 I_L^r 与 I_H^r 之间存在线性关系.在算法的第二步中,可首先由式 (13) 计算出 I_L^r ,然后基于 MAP 准则由式 (15) 计算出最优的 I_H^r .

$$I_H^r = \arg \max_{I_H^r} P(I_H^r | I_L^r) = \arg \max_{I_H^r} P(I_L^r | I_H^r) P(I_H^r) \quad (15)$$

这里需要估计先验概率 $P(I_H^r)$ 和条件概率 $P(I_L^r | I_H^r)$.由式 (14) 可知, $P(I_L^r | I_H^r)$ 取决于 n 的分布,设 n 服从高斯分布 $N(\mu_n, \Sigma_n)$,参数 μ_n, Σ_n 可通过训练样本估计,则由式 (14)

$$P(I_L^r | I_H^r) = \frac{1}{C_3} \exp \left\{ -\frac{1}{2} (I_L^r - AI_H^r - \mu_n)^T \Sigma_n^{-1} (I_L^r - AI_H^r - \mu_n) \right\} \quad (16)$$

其中 C_3 是常数.进一步假设 I_H^r 是一个 Markov 随机场,则 I_H^r 服从 Gibbs 分布^[4],即

$$P(I_H^r) = \frac{1}{C_4} \exp \left\{ -\frac{1}{2T} E(I_H^r) \right\} \quad (17)$$

其中 C_4 与 T 是与 I_H^r 无关的常数, $E(\cdot)$ 是能量函数.由平滑性约束,能量函数可定义为

$$E(I_H^r) = \sum_{v=0}^{N-1} (I_H^r d_v) (d_v^T I_H^r) = I_H^r \begin{pmatrix} N-1 \\ v=0 \end{pmatrix} d_v d_v^T I_H^r$$

$$\text{其中 } d_{v,i} = \begin{cases} 1, & i = v \\ -1/4, & i \text{ 是 } v \text{ 的邻域像素} \\ 0, & i \neq v \text{ 且 } i \text{ 不是 } v \text{ 的邻域像素} \end{cases} \quad (18)$$

令 $\begin{pmatrix} I \\ H \end{pmatrix}^{-1} = \sum_{v=0}^{N-1} d_v d_v^T$,将式 (18) 代入式 (17) 得

$$P(I_H^r) = \frac{1}{C_4} \exp \left\{ -\frac{1}{2T} I_H^r \begin{pmatrix} I \\ H \end{pmatrix}^{-1} I_H^r \right\} \quad (19)$$

我们在实验中取 T 为 150,将式 (19) 和式 (16) 代入式 (15),经过推导可计算出

$$I_H^r = \left[A^T \Sigma_n^{-1} A + \frac{1}{T} \begin{pmatrix} I \\ H \end{pmatrix}^{-1} \right]^{-1} A^T \Sigma_n^{-1} (I_L^r - \mu_n) \quad (20)$$

3 实验结果

我们采集了 500 个人的 1000 幅正面人脸图像,手工标定并对齐了眼睛中心、鼻尖和嘴角共五个特征点,然后把图像大小统一到 128×96 ,并利用式将所有样本手工降低分辨率得到 32×24 和 16×12 的低分辨率图像.我们将其中 980 组图像作

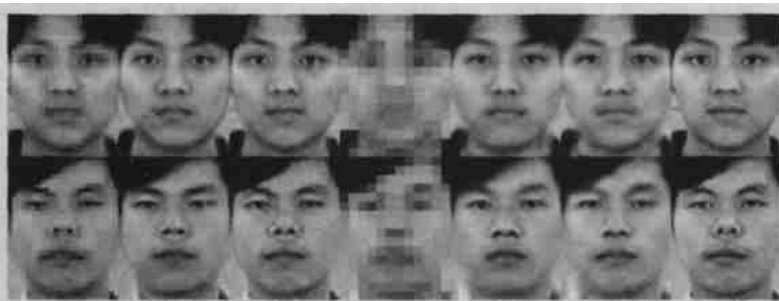


图 1 实验结果。(a)、(d)输入的低分辨率图像(32×24; 16×12); (b)、(e)整体图像(128×96); (c)、(f)最终结果(128×96); (g)真实图像(128×96)



图 2 实验结果比较。(a)、(d)输入的低分辨率图像(32×24; 16×12); (b)、(e)新方法的结果(128×96); (c)、(f)Liu 等人方法的结果(128×96); (g)真实图像(128×96)

为训练样本,另外 20 组用作测试,部分实验结果如图 1 所示。为了比较新方法与 Liu 等人^[1]的方法,我们用相同的训练

和测试样本实验了 Liu 等人的方法,图 2 比较了两种方法的结果,图 3 给出了两种方法对 20 幅测试图像得到的结果与真实图像做匹配所计算出的相关系数(相关系数越大表明越接近真实图像)。通过实验比较,我们认为新方法与 Liu 等人的方法相比有两个优势,一是新方法结果更接近真实图像,二是新方法计算一幅图像所需的时间更短(在 P4 1.7GHz 的 PC 机上计算一幅图像, Liu 等人的方法大约需 30 秒,新方法只需 2~3 秒)。

新方法在我们采集的数据集上取得了令人满意的结果,为了验证其在集外图像上的分辨率增强效果,我们另外采集了 36 幅退化模型是未知的低分辨率人脸图像用于测试,训练集仍为前面提到的 980 组图像,部分实验结果如图 4 所示。从图 4 可以看出,尽管第一步重构出的整体图像与真实图像不是十分相似,但算法的第二步可以对结果起到明显的修正作用使其更加接近于真实图像。这个实验同时说明我们的方法对其它退化模型造成的分辨率图像也是适用的。对 36 个集外的人,实验获得了 36 幅整体图像和 36 幅最终结果。为了量化新方法第二步对细节信息的修正作用,我们对每幅整体图像、最终结果以及真实图像手工标定了 21 个特征点,特征点位置如图 5 所示,进而对每幅图像提取了 840 维的 Gabor 特征向量(关于 Gabor 特征向量的计算请参照^[7])。我们分别计算了整体图像以及最终结果的 Gabor 特征与真实图像的 Gabor 特征之间的相关系数,结果如图 6 所示。从图 6 可以看出,最终结果的 Gabor 特征与真实图像的 Gabor 特征更相似,这说明通过算法第二步的修正可以使最终结果

的细节信息更接近于真实图像。

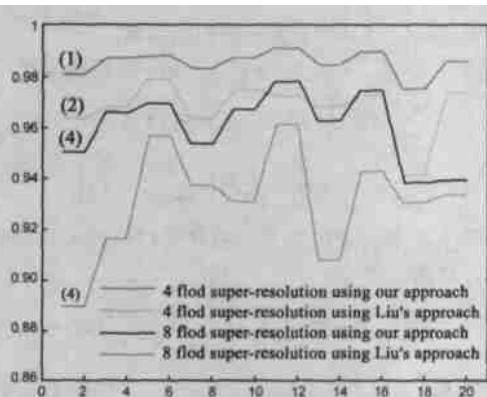


图 3 相关系数比较。横坐标为测试图像编号,纵坐标为相关系数。(1)、(2)分别是我们的方法和 Liu 等人的方法在分辨率增强 16 倍(由 32×24 增强到 128×96)时得到的结果;(3)、(4)分别是我们的方法和 Liu 等人的方法在分辨率增强 64 倍(由 16×12 增强到 128×96)时得到的结果



图 4 集外测试图像分辨率增强结果。(a)是整体图像;(b)是最终结果;(c)是真实图像

4 总结与展望

实验证明,新方法在我们采集的人脸库上已经取得了令人满意的结果,我们还需要在更大型的人脸库上进行训练和测试,以进一步验证新方法的适用性。此外,如何更加准确地得到人脸的细节信息将是我们今后工作中希望着重解决的问题。

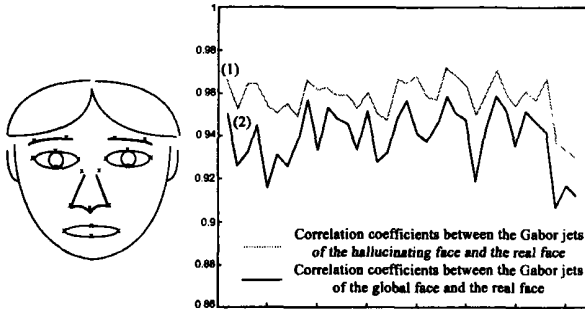


图 5 21 个人脸特征点 图 6 相关系数比较. 横坐标为测试图像编号, 纵坐标为相关系数. 曲线 (1) 描述了最终结果的 Gabor 特征与真实图像的 Gabor 特征之间的相关系数; 曲线 (2) 描述了整体图像的 Gabor 特征与真实图像的 Gabor 特征之间的相关系数

参考文献:

- [1] C Liu, H Y Shum, C-S Zhang. A two-step approach to hallucinating faces: Global parametric model and local nonparametric model [A]. In Proc of IEEE Conf on Computer Vision and Pattern Recognition [C]. Hawaii, USA, Dec 2001.
- [2] S Baker, T Kanade. Hallucinating faces [A]. In Proc of the 4th International Conf on Automatic Face and Gesture Recognition [C]. Grenoble, France, Mar. 2000.
- [3] R Schultz, R Stevenson. Extraction of high-resolution frames from video sequences [J]. IEEE Transactions on Image Processing, 1996, 5 (6): 996 - 1011.
- [4] R Hardie, K Barnard, E Armstrong. Joint MAP registration and high-resolution image estimation using a sequence of undersampled images [J]. IEEE Transactions on Image Processing, 1997, 6 (12): 1621 - 1633.
- [5] W Freeman, E Pasztor. Learning low-level vision [A]. In Proc of IEEE Conf on Computer Vision [C]. Kerkyra, Corfu, Greece, Sep 1999.
- [6] B K Gunturk, A U Batur, Y Altunbasak, M H Hayes, R M Mersereau. Eigenface-domain super-resolution for face recognition [J]. IEEE Transactions on Image Processing, 2003, 12 (5): 597 - 606.
- [7] L Wiskott, J M Fellous, N Kruger, C von der Malsburg. Face recognition by elastic bunch graph matching [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19 (7): 775 - 779.

作者简介:



李 林 2002 年获清华大学计算机科学与技术系工学学士学位, 目前于清华大学计算机科学与技术系人机交互与媒体集成研究所攻读硕士学位, 研究方向包括计算机视觉、机器学习、模式识别和图像处理. E-mail: li_yang98@mails.tsinghua.edu.cn.



林学 清华大学计算机科学与技术系教授, 博士生导师, 1962 年毕业于清华大学自动控制系, 现研究方向为计算机视觉、图像处理与模式识别, 曾负责与参与多项国家 863 高科技项目, 国家自然科学基金项目, 在国内外期刊、国际学术会议发表论文 100 多篇. Email: lxy-dcs@mail.tsinghua.edu.cn.

(上接第 206 页)

- [6] WANG Y F, Xue Y J, Zhang Y, Zhu S Y, Qian L Q. A matching model for software component classification in faceted scheme [J]. Journal of Software, 2003, 14 (3): 401 - 408.
- [7] Hai-Feng Guo, Miao Liu, Jiexiong Pi. Precise specification matching for automated component retrieval and adaptation [A]. Proceedings of the 2003 IEEE International Conference on Information Reuse and Integration (IRI-2003) [C]. Las Vegas, USA, Oct. 2003.
- [8] Jilani L L, Desharnais J, Frappier M, Mili R, Mili A. Retrieving software components that minimizing adaptation effort [A]. Automated Software Engineering, 1997 Proceedings, 12th IEEE International Conference [C]. Kluwer Academic Publisher, Dordrecht, Netherlands. 1 - 5 Nov. 1997.
- [9] Eduardo Ostertag, James Hendler, Ruben Prieto-Diaz, Christine Braun. Computing similarity in a reuse library system: An AF-based approach [J]. ACM Trans on Software Engineering and Methodology, 1992, 1 (3): 205 - 28.
- [10] M R Gardi, B Ibrahim. Automatic indexing of software artifacts [A]. Proc. 3rd Int'l Conf Software Reuse: Advances in Software Reusability [C]. IEEE Computer Society Press, Rio de Janeiro, Brazil, Nov. 1994. 24 - 32.
- [11] Dieter Merkl, A Min Tjoa, Certi Kappel. Learning the semantic similarity of reusable software components [A]. Proc. 3rd Int'l Conf. on Software Reuse: Advances in Software Reusability, IEEE Computer Society [C]. Rio de Janeiro, Brazil, 11/1994. 33 - 41.
- [12] Tomas Isakowitz, Robert J Kauffman. Supporting search for reusable software objects [J]. IEEE Trans on Software Engineering, 1996, 22 (6): 407 - 423.
- [13] Li Keqin, Guo Lifeng, Mei Hong, Yang Fuqing. An overview of JB (Jade Bird) component library system JBCL [A]. Proc TOOLS Asia '97 [C]. Beijing, China, Sep. 1997.
- [14] W B Frakes, T P Pble. Proteus: A reuse library system that supports multiple representation methods [J]. ACM SIGIR Forum, 1990. 24: 43 - 55.