

基于图像重建的表情识别算法

朱明旱^{1), 2)} 罗大庸¹⁾ 王一军¹⁾

¹⁾ (中南大学信息科学与工程学院, 长沙 410083) ²⁾ (湖南文理学院电气与信息工程学院, 常德 415000)

摘要 提出了一种基于图像重建的表情识别算法。首先, 用 LE (Lipschitz embedding) 算法提取出训练集中各个对象的表情流形, 并建立流形向量与图像向量的映射关系。再通过非线性重建, 确定待测图像在流形空间中的坐标。最后, 用待测图像在各表情路径上的投影, 重建各种表情图像, 实现表情识别。该算法解决了各表情流形相互重叠的问题, 且对表情强度变化具有鲁棒性。在 Cohn-Kanade 和 CMU-AMP 人脸库上的结果实验表明, 该算法具有较好的表情识别率。

关键词 李普希茨嵌入 非线性映射 图像重建 表情识别

中图分类号: TP391.41 文献标志码: A 文章编号: 1006-8961(2010)01-0098-05

Expression Recognition Algorithm Based on Image Reconstruction

ZHU Ming-han^{1), 2)}, LUO Da-yong¹⁾, WANG Yi-jun¹⁾

¹⁾ (College of Information Science and Engineering, Central South University, Changsha 410083)

²⁾ (College of Communication and Electric Engineering, Hunan University of Arts and Science, Changde 415000)

Abstract In this paper, an expression recognition algorithm based on image reconstruction is proposed. Firstly, manifolds of different subjects were obtained singly using the Lipschitz Embedding algorithm. Mapping between manifold vectors and image vectors was established. Then, the coordinates of testing image in manifold space was confirmed by nonlinear reconstructing. Finally, expression images were reconstructed according to projecting vectors of testing image on manifold paths and expression recognition was accomplished. The algorithm solves the problem that expression manifolds overlap with one another. It is robust to the variety of expression intensity. The experiments in Cohn-Kanade and CMU-AMP face database show the algorithm has better expression recognition accuracy.

Keywords Lipschitz embedding, nonlinear mapping, image reconstruction, expression recognition

0 引言

表情是人类用来表达情绪的一种基本方式, 是非语言交流的一种有效手段, 它在沟通、社交和心理暗示上都代表着强烈的意义。Ekman 和 Friesen 提出了面部表情编码系统 (facial action coding system, FACS), 用 44 个运动单元 (action unit AU) 来描述人脸表情变化, 并定义了 6 种基本情感类别: 惊奇、

恐惧、厌恶、愤怒、高兴、悲伤^[1]。这种处理方法得到了广泛的认同, 并成为后来很多表情识别研究的基础。

针对人脸表情的识别, 研究者们已经提出了很多新的处理方法^[2-5]。最近, 部分学者又提出了一些基于流形的表情分析方法, 如: Chan 等人提出了非线性配准的李普希茨嵌入 (Lipschitz embedding LE) 表情分析法^[6-7], 先用 LE 得到每人的表情流形, 然后将所有的表情流形配准到一个标准形, 最后用 k

基金项目: 国家自然科学基金项目 (60776834); 湖南省教育厅科研项目 (08C606)

收稿日期: 2008-09-05 改回日期: 2008-11-13

第一作者简介: 朱明旱 (1974—), 男, 讲师。中南大学信息科学与工程学院模式识别与智能系统专业博士研究生。主要研究方向为模式识别与计算机视觉。Email: zhumh_123@163.com

近邻方法 (k nearest neighbor KNN) 识别表情。但是, 由于各个对象的表情流形, 在形状、伸展方向上不尽相同, 人数越多, 配准效果就越差, 识别率会降低。后来, Shan 等人提出了基于 SLPP (supervised locality preserving projections) 的表情识别方法^[8]。该方法通过加大不同表情图像间的欧氏距离, 使图像不按身份聚类, 而按表情聚类。可是保局投影不能保留全部的数据结构关系, 这使图像的聚类效果不够理想, 并且, 随着涉及对象的增加, 各种表情的流形结构间会发生越来越严重的重叠, 影响识别的准确率。因此, 该方法仅对涉及对象较少的表情识别场合, 有较好的效果。Chan-Su 等人则提出了流形学习与张量分解相结合的表情分析方法^[9], 先提取人脸图像的流形特征, 再通过张量分解进行表情识别与合成, 而张量分解的最大缺点就是运算量太大。到目前为止, 如何高效、准确地识别出表情类别, 仍然是一个有待进一步研究的课题。

本文提出了一种基于图像重建的表情识别算法, 先用 LE 算法提取训练集中每个对象的表情流形, 再通过图像重建, 确定出待测图像所在的流形空间和在此空间中的坐标向量, 最后根据重建图像与待测图像的相似度, 识别出表情。该算法为每个对象建立一个表情流形空间, 因而不存在各流形结构重叠的问题, 并且, 重建图像的表情强度会随着待测图像的表情强度而自动变化, 克服了表情强度变化对识别率的不利影响。

1 算法描述

本文的算法分 3 步实现: 1) 用 LE 算法提取出训练集中每个人的表情流形, 并建立流形空间与图像空间的映射关系; 2) 通过非线性重建 (NR), 确定待测图像在流形空间里的坐标; 3) 用流形空间中, 待测图像在各表情路径上的投影, 重建 6 种表情图像, 最终完成表情的识别。

1.1 表情流形提取

假设训练集包含 m 个对象的 6 种基本表情图像, 可用 $X(k) = \{x_1, \dots, x_n\}$ 表示第 k 个对象的所有图像向量组成的集合, $r_j (1 \leq j \leq 6)$ 表示 $X(k)$ 中的第 j 种极值表情图像向量。

对每个对象的表情图像分别用 LE 算法^[10] 处理, 共得到 m 个表情流形, 记为 $M_k (1 \leq k \leq m)$, M_k 的提取过程如下:

```

Input
     $X(k) = \{x_1, \dots, x_n\}, x_i \in \mathbf{R}^D$ 
for  $i = 1$  to  $n$ 
    for  $j = 1$  to  $6$ 
         $y_i = \|x_i - r_j\|$ 
    end
     $y_i = (y_1, \dots, y_6)^T$ 
end
Output
     $Y(k) = \{y_1, \dots, y_n\}, y_i \in \mathbf{R}^6$ 
    
```

表情流形 M_k 由 6 条从中心点向外伸展的流形路径构成, 每条路径各对应一种表情类。路径上距中心最远的点对应极值表情, 距中心越近, 表情强度越小。为了后续处理的方便, 将 M_k 进行坐标平移, 使中性表情点位于坐标原点。图 1 为 Cohn-Kanade 人脸库上^[11], 某人的表情流形 (为了可视化只取了 3 种表情, 图中“○”, “☆”, “□”各代表一种表情)。

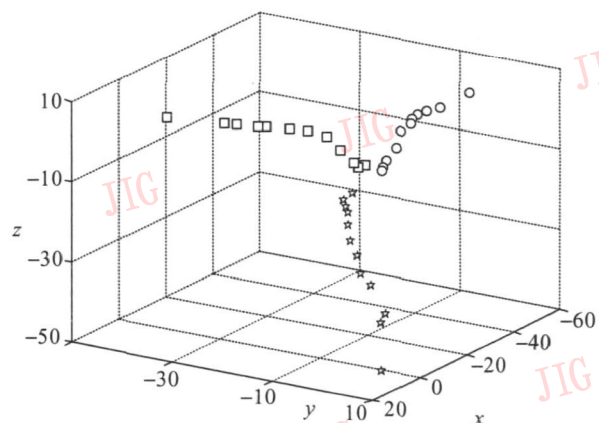


图 1 Cohn-Kanade 人脸库上某人的表情流形

Fig 1 Expression manifold of one subject from Cohn-Kanade face database

用高斯径向基函数, 建立流形向量 $\{y_1, \dots, y_n\}$ 与图像向量 $\{x_1, \dots, x_n\}$ 的非线性映射关系^[9], 核函数如式 (1) 所示:

$$\phi_k(\mathbf{y}, z_i) = \exp \left[-\frac{\|\mathbf{y} - z_i\|^2}{2\sigma_i} \right] \quad 1 \leq i \leq c \quad (1)$$

$\{z_1, \dots, z_c\}$ 是 $Y(k)$ 的 c 个聚类中心, 用 k 均值聚类

算法求出。 $\sigma_i = \frac{1}{Np} \sum_{j=1}^p (z_i - z_j)^2$, $\{z_1, \dots, z_p\}$ 为聚类中心 z_i 的 p 个近邻中心, 一般 p 取 2 定义映射

$$\phi_k(\mathbf{y}): \mathbf{R}^6 \rightarrow \mathbf{R}^D$$

$$\phi_k(\mathbf{y}) = [\phi_k(\mathbf{y}, z_1), \dots, \phi_k(\mathbf{y}, z_c)]^T \quad (2)$$

第 k 个对象的流形向量 \mathbf{y} 与它所对应的图像向

量 x 的映射关系为

$$x = \sum_{i=1}^c w_i \phi_k(y, z_i)$$

$$x = W_k \phi_k(y) \quad (3)$$

W_k 是个 $D \times c$ (D 为向量 x 的维数) 的系数矩阵, 为了求解 W_k 各元素的值, 取 c 个向量 $\{x_1, \dots, x_c\}$, 得 $D \times c$ 个线性方程

$$[x_1, \dots, x_c] = W_k [\phi_k(y_1), \dots, \phi_k(y_c)] \quad (4)$$

求解式 (4), 得 W_k 各元素的值。若矩阵 $[\phi_k(y_1), \dots, \phi_k(y_c)]$ 非奇异, 则有:

$$W_k = [x_1, \dots, x_c] [\phi_k(y_1), \dots, \phi_k(y_c)]^{-1}$$

根据所求的 W_k 和 ϕ_k , 得到由 y_i 重建图像向量 x_i 的表达式:

$$\hat{x}_i = W_k \phi_k(y_i) \quad (5)$$

1.2 待测图像嵌入

对于待测图像向量 x' , 分别用 m 个对象的参考集对它进行 LE 处理, 得到 m 个候选向量, 记为 $(y'(1), \dots, y'(m))$ 。用这 m 个候选向量, 重建出 m 个图像向量:

$$\hat{x}(k) = W_k \phi_k(y'(k)) \quad (1 \leq k \leq m) \quad (6)$$

根据自联想模型理论^[12], 定义 x' 的流形空间坐标为 $y'(k)$ 的概率:

$$P(x' | \hat{x}(k), M_k) = \exp \left\{ - \frac{\|x' - \hat{x}(k)\|^2}{t} \right\} \quad (7)$$

t 为一预先设定的常数, 可用来调节衰减速度。

最大 $P(x' | \hat{x}(k), M_k)$ 所对应的 $y'(k)$ 和 M_k , 为 x' 的投影向量 y' 和它所在的流形, 即

$$(y', M_l) = \arg \max P(x' | \hat{x}(k), M_k) \quad (8)$$

经过上述处理, 待测图像向量 x' 被嵌入到它所在流形空间, 并获得了它在此流形空间中的坐标。

1.3 表情识别

设已获得待测图像向量 x' 所在的流形 M_l 和它的流形向量 y' 。将 M_l 第 j 条路径上, 距中心最远的向量 (称为极值表情向量) 表示为 y_j^{max} 。定义流形 M_l 第 j ($1 \leq j \leq 6$) 条路径的伸展方向为

$$w_j = \frac{y_j^{max}}{\|y_j^{max}\|} \quad (9)$$

则向量 y' 向 w 投影后, 所得的向量为

$$a_j = w_j w_j^T y' \quad (10)$$

式 (10) 的推导如下, 根据 PCA (principal component analysis) 算法可知, 向量 y' 向 w_j 投影后得到的特征向量为

$$A = w_j^T (y' - \bar{y}_j)$$

这里 \bar{y}_j 为第 j 条路径上向量的平均值, 可取 $\bar{y}_j = \lambda w_j$ (λ 为一比例常数), 有

$$a_j = w_j A + \bar{y}_j$$

$$= w_j w_j^T (y' - \bar{y}_j) + \bar{y}_j$$

$$= w_j w_j^T y' - \lambda w_j (w_j^T w_j) + \lambda w_j$$

$$= w_j w_j^T y'$$

在流形 M_l 中, 用式 (10) 将流形向量 y' 分别向 6 个伸展方向投影, 得到 6 个向量 a_j ($1 \leq j \leq 6$)。图 2 展示了 y' 向 3 个伸展方向的投影 (“*” 表示 y' , “◆” 表示 a_1, a_2, a_3)。

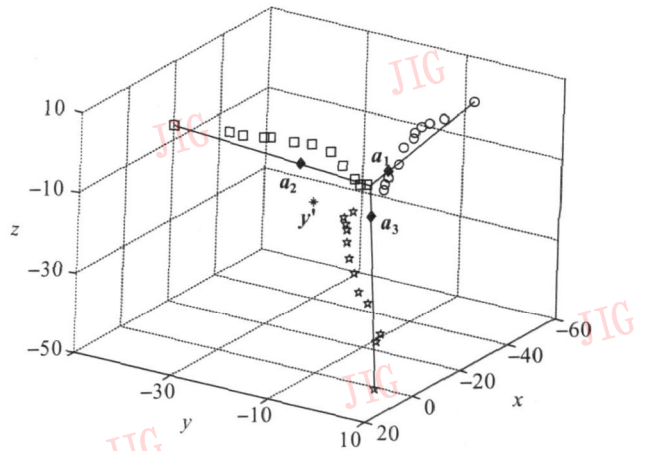


图 2 y' 向 3 条表情路径的投影

Fig 2 Projection of y' on three expression paths

与式 (6) 同理, 有重建表情图像的表达式:

$$\hat{x}_j = W_l \phi_l(a_j) \quad (11)$$

用式 (11) 重建对象 l 的 6 种表情图像 \hat{x}_j ($1 \leq j \leq 6$)。

与式 (7) 的原理相似, 定义 x' 与 \hat{x}_j 的相似度为

$$P_j' = \exp \left\{ - \frac{\|x' - \hat{x}_j\|^2}{t} \right\} \quad (12)$$

为了与概率度量一致, 将 P_j' 归一化为

$$P_j = \frac{P_j'}{\sum_{j=1}^6 P_j'} \quad (13)$$

在识别待测图像 x' 的表情类别时, P_j 就是它属于第 j 种表情的概率大小, 所以最大 P_j 所对应的 j 就是待测图像的表情类别。

在表情的识别过程中, 每个对象都有自己的流形空间, 因此各对象的流形结构不会有任何重叠。同时, 若 x' 的表情强度变大, 它的投影 a_j 距中心距离会变大。在它重建的 6 幅表情图像 \hat{x}_j ($1 \leq j \leq 6$) 中, 与 x' 同类的那幅表情图像 \hat{x}_j 的强度也会变大, 这使得该算法对表情强度的变化具有鲁棒性。

2 实验与分析

2.1 实验

实验是在 Cohn-Kanade 和 CMU-AMP^[13] 人脸库上进行的, Cohn-Kanade 人脸库包含 210 个对象的 2 000 个图像序列, 每个对象有一系列的脸部活动。从中选取了 15 个对象的 6 种表情序列 (高兴、悲伤、惊讶、生气、厌恶、恐惧), 每个表情序列选 12 幅图像, 共 1080 幅图像, 图 3 为 Cohn-Kanade 人脸库中的一些样本。CMU-AMP 人脸库由 13 人, 每人 75 幅共 975 幅 8 位灰度图像组成, 图像分辨率为 64×64 每人的 75 幅图像都包括高兴、生气、惊奇这 3 种表情, 图 4 为 CMU-AMP 人脸库中的一些样本。



图 3 Cohn-Kanade 人脸库中的图像

Fig. 3 Images from Cohn-Kanade face database



图 4 CMU-AMP 人脸库中的图像

Fig. 4 Images from CMU-AMP face database

在 Cohn-Kanade 和 CMU-AMP 人脸库上, 均采用交叉验证的方法进行实验。将 Cohn-Kanade 人脸库分为 X_1 和 X_2 两个子集, 随机抽取每人每种表情的 6 幅图像, 共 540 幅组成 X_1 , X_2 由剩下的 540 幅组成。CMU-AMP 人脸库分为 X_1 和 X_2 两个子集, 随机抽取每人每种表情的 12 幅图像, 共 468 幅组成 X_1 , X_2 由剩下的 507 幅组成。分别采用文献 [6] 的 MNA (manifold by nonlinear alignment) + KNN、本文算法步骤 1) 和 2) 的内容 + KNN (简记为 LE + NR + KNN)、本文的全部算法来识别表情。实验时, 先将 X_1 作为训练集, X_2 作为测试集, 再将 X_2 作为训练集, X_1 作为测试集, 计算出平均识别率, KNN 分类器的 $k = 5$ 各类表情的识别率如表 1 和表 2 所示。

图 5 为 Cohn-Kanade 人脸库上的两幅待测图像和它们的重建图像。图 5(a) 为待测图像 I, 图 5

(b) 为待测图像 II, 图 5(c) 为图像 I 的 6 幅重建图像, 图 5(d) 为图像 II 的 6 幅重建图像。



图 5 待测图像和它们的重建图像

Fig 5 Test images and their reconstruction images

表 1 在 Cohn-Kanade人脸库上的表情识别率

Tab 1 The recognition accuracy of expression in Cohn-Kanade face database

	%					
	高兴	悲伤	厌恶	惊讶	生气	恐惧
MNA+KNN	86.3	82.8	82.6	83.6	79.4	80.7
LE+NR+KNN	88.2	83.3	84.4	85.6	80.2	82.5
本文算法	94.3	90.4	92.2	93.5	91.4	90.1

表 2 在 CMU-AMP人脸库上的表情识别率

Tab 2 The recognition accuracy of expression in CMU-AMP face database

	%		
	高兴	生气	惊讶
MNA+KNN	82.5	81.4	82.3
LE+NR+KNN	85.1	83.8	81.5
本文算法	91.7	90.4	89.2

2.2 实验分析

用图 5 展示了表情识别中, 某 2 幅待测图像和它们的重建图像。不难发现, 图 5(c) 中的第 2 幅图像与待测图像 I 最相似, 图 5(d) 中的第 2 幅图像与待测图像 II 最相似。且这两幅图像的表情强度随待测图像的表情强度而变化, 表现了该算法对表情强度变化的适应性。

综合表 1 和 2 的数据, 得到 MNA+KNN 的平均表情识别率为 82.40%, LE+NR+KNN 的为 83.8%, 本文算法的为 91.47%。这表明本文的算法较好地实现了对人脸表情的识别。

3 结 论

本文提出了一种基于图像重建的表情识别算法。它的表情识别处理是在各个对象的流形空间里完成的, 因此, 没有各对象表情流形重叠的情况发生, 也无需对各对象的流形进行配准处理, 这使该算法的表情识别率, 不会因为对象的增加而受到较大的影响。由于重建图像的表情强度会随待测图像而自动变化, 使该算法的识别率对表情强度的变化有较好的鲁棒性。最终的表情识别通过比较待测图像与重建图像的相似度而实现, 与流形分析后用 KNN 算法来识别表情相比, 在一定程度上也会降低训练样本数的多少对识别率的影响。

参考文献 (References)

- [1] Ekman P, Friesen W V. Facial action coding system: A Technique for the Measurement of Facial Movement[M]. Palo Alto, CA, USA: Consulting Psychologists Press, 1978.
- [2] Tian Y, Kanade T, Cohn J. Recognizing action units for facial expression analysis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(2): 97-115.
- [3] Franco L, Treves A. A neural network facial expression recognition system using unsupervised local processing[C]//Proceedings of International Symposium on Image and Signal Processing and Analysis. Piscataway, NJ, USA, IEEE Service Center, ETATS-UNIS, 2001: 628-632.
- [4] Cohen I, Sebe N, Garg A, et al. Facial expression recognition from video sequences: Temporal and static modeling[J]. Computer Vision and Image Understanding, 2003, 91(1): 160-187.
- [5] Fasel R, Luetttin J. Automatic facial expression analysis: A survey[J]. Pattern Recognition, 2003, 36(1): 259-275.
- [6] Hu C, Chang Y, Feris R, et al. Manifold based analysis of facial expression[C]//Proceedings of Computer Vision and Pattern Recognition Workshop on Face Processing in Video. Washington, USA: IEEE Computer Society, 2004, 5: 81-87.
- [7] Chang Y, Hu C, Turk M. Probabilistic expression analysis on manifolds[C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, USA, IEEE Computer Society, 2004, 2: 520-527.
- [8] Shan C F, Gong S G, Peter W M. Appearance manifold of facial expression[C]//Proceedings of IEEE International Workshop on Human-Computer Interaction. Berlin: Springer, 2005, 3766: 221-230.
- [9] Lee Chan-su, Elgammal A. Facial expression analysis using nonlinear decomposable generative models[C]//Proceedings of International Workshop on Analysis and Modeling of Faces and Gestures. New York, NY, USA: Springer, 2005, 3723: 17-31.
- [10] Bougain J. On Lipschitz embedding of finite metric space in Hilbert space[J]. Israel Journal of Mathematics, 1985, 52(1): 46-52.
- [11] Cohn J, Kanade T. The Cohn-Kanade AU-Coded Facial Expression Database[EB/OL]. [2007-06-15] <http://www.cs.cmu.edu/~face/index2.htm/1999>
- [12] Junping Z, Stan Z L. Adaptive nonlinear auto-associative modeling through manifold learning[C]//Proceedings of Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining. Berlin: Springer, 2005, 3518: 599-604.
- [13] Advanced Multimedia Processing Lab. CMU-AMP Face Expression Database[EB/OL]. (2002-04-18) [2008-01-08] <http://anp.ece.cmu.edu/projects/FaceAuthentication/download.htm>