

基于 Gabor 变换与改进 SLLE 的人脸表情识别^①

龚安, 曾雷

(中国石油大学(华东) 计算机与通信工程学院, 青岛 266580)

摘要: 本文通过 Gabor 变换进行人脸表情图像的特征提取, 并利用局部线性嵌入(LLE)系列算法进行数据降维操作. LLE 算法是一种非线性降维算法, 它可以使得降维后的数据保持原有的拓扑结构, 在人脸表情识别中有广泛的应用. 因为 LLE 算法没有考虑样本的类别信息, 因此有了监督的局部线性嵌入(SLLE)算法. 但是 SLLE 算法仅仅考虑了样本的类别信息却没有考虑到各种表情之间的关系, 因此本文提出一种改进的 SLLE 算法, 该算法认为中性表情是其他各种表情的中心. 在 JAFFE 库上进行人脸表情识别实验结果表明, 相比 LLE 算法和 SLLE 算法, 该算法获得了更好的人脸表情识别率, 是一种有效算法.

关键词: Gabor 变换; 人脸表情; 特征提取; LLE; SLLE

引用格式: 龚安, 曾雷. 基于 Gabor 变换与改进 SLLE 的人脸表情识别. 计算机系统应用, 2017, 26(9): 210-214. <http://www.c-s-a.org.cn/1003-3254/5949.html>

Facial Expression Recognition Based on Gabor Transform and Improved SLLE

GONG An, ZENG Lei

(College of Computer & Communication Engineering, China University of Petroleum (East China), Qingdao 266580, China)

Abstract: In this paper, we adopt the Gabor transform to extract features of facial expression images and use a series of locally linear embedding(LLE) algorithms to reduce the data dimension. The LLE algorithm widely used in facial expression recognition is a kind of nonlinear dimension reduction algorithm. It is able to make dimension-reduced data keep the original topology. Because the LLE algorithm does not take the category information of samples into account, the supervised locally linear embedding(SLLE) algorithm appears. But the SLLE algorithm only considers the category information of samples, and does not take the relationship among various expressions into account. Therefore, in this paper, we propose an improved SLLE algorithm, which regards the neutral expression as the center of the other expressions. The results of facial expression recognition experiments on the JAFFE database show that our algorithm obtained better facial expression recognition rate compared with the LLE algorithm and the SLLE algorithm. Our algorithm is more effective.

Key words: Gabor transform; facial expression; extract features; LLE; SLLE

人脸表情可以反映人类内心真实情感, 了解一个人内心的真实情感有利于我们更好地与这个人沟通以及为其服务. 因此人脸表情识别有很重要的研究意义, 人脸表情识别的应用领域也很广泛, 包括人机交互、机器人制造、心理咨询、智能汽车等领域. 在进行人

脸表情识别之前有两个很重要的过程: 特征提取与数据降维.

人脸表情识别中常用的特征提取方法有 ASM^[1,2]、AAM^[3-5]、LBP^[6]、弹性图匹配法、Fisher 判别、Gabor^[7,8]变换等.

^① 收稿时间: 2016-12-23; 采用时间: 2017-01-18

人脸表情识别中常用的数据降维算法分为线性降维与非线性降维。常见的线性降维算法有 PCA^[9]、LDA 等, 常见的非线性降维算法有 Isomap、LLE^[10]、LEE^[11]、MVU 等。

本文特征提取采用 Gabor 变换, 数据降维采用 LLE 系列算法。LLE 算法是 Sam T. Roweis 和 Lawrence K. Saul 在 2000 年提出的一种非线性降维方法。LLE 算法可以通过局部线性拟合来恢复全局非线性结构, 因此通过 LLE 降维后的数据会保持原有的拓扑结构。虽然 LLE 算法有很好的降维效果, 但是 LLE 算法没有考虑样本的类别信息, 因此 Dick 和 Robert 在 2003 年提出一种针对有监督的 LLE 算法, 即 SLLE^[12,13] 算法。虽然 SLLE 算法考虑到了样本的类别信息, 但是没有考虑到人脸表情识别问题上各种表情之间的关系。因此本文提出了一种改进的 SLLE 算法, 该算法在计算不同类别之间最大类间距离时, 将中性表情看成 n 维空间的中心, 其他各种表情分别在 n 个不同的维度上, 因此其他各种表情到中性表情之间的距离为它们之间的欧氏距离, 其他各种表情中的 2 种不同表情之间的距离的平方是这 2 种表情到中性距离的平方和。该算法在 JAFFE 数据库上获得了较好的实验结果, 说明该算法是一种有效的算法。

1 特征提取

Gabor 变换是 Dennis Gabor 提出的一种特殊情况下的短时傅里叶变换。它用来确定局部信号随时间变化对应的正弦频率与相位。Gabor 变换的功能是, 首先乘以一个被视为窗口的高斯函数, 然后将该结果通过傅里叶变换得到信号的时间频率分析转换。

1.1 Gabor 滤波器

在图像处理中, Gabor 滤波器是一种以 Gabor 变换为基础的用于边缘检测的线性滤波器。Gabor 滤波器中频率和方向的表现类似于人类的视觉系统, 它们被发现特别适合纹理的表示和区别。因此 Gabor 滤波器作为一种特征提取的方法, 在图像处理, 模式识别等领域得到了广泛的应用。在空间域中, 一个二维 Gabor 滤波器是一种由正弦平面波调制的高斯核函数, 具体如下:

复数形式:

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \exp\left(i\left(2\pi\frac{x'}{\lambda} + \psi\right)\right) \quad (1)$$

实部:

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi\frac{x'}{\lambda} + \psi\right) \quad (2)$$

虚部:

$$g(x, y; \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \sin\left(2\pi\frac{x'}{\lambda} + \psi\right) \quad (3)$$

其中 $x' = x\cos\theta + y\sin\theta$, $y' = -x\sin\theta + y\cos\theta$, λ 表示正弦函数的波长, θ 表示 Gabor 函数平行条纹的方向, ψ 表示相位偏移, σ 表示高斯包络的标准差, γ 表示空间的纵横比。

1.2 Gabor 特征提取

通常采用多个在不同波长和方向上的 Gabor 滤波器组成滤波器组来提取人脸图像的 Gabor 特征。本文选择 5 种波长和 8 种方向, 令 ψ 为 0, σ 为 $\sqrt{2}$, γ 为 $\sqrt{2}$ 。通常一般是选择 5 种频率尺寸与 8 种方向, 因为上面的公式(1), 公式(2), 公式(3)中用到的都是波长 λ 这个参数, 所以本文这里使用的 5 种波长, 实际上波长与频率尺度的乘积是 2π 。

5 种波长对应的 λ 为:

$$\lambda_u = fw^u \quad (u = 0, 1, 2, 3, 4) \quad (4)$$

其中 f 为 4, w 为 $\sqrt{2}$ 。

8 种方向对应的 θ 为:

$$\theta = \frac{v\pi}{8} \quad (v = 0, 1, 2, 3, 4, 5, 6, 7) \quad (5)$$

40 个滤波器的实部如图 1 所示。

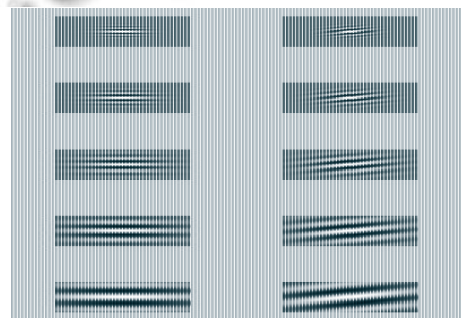


图 1 滤波器的实部

2 数据降维

Gabor 提取的特征的维度比较大, 所以需要对其进行降维操作。

2.1 LLE 算法

LLE 算法对数据进行降维的特点如下:

令高维样本集 $X=[x_1, x_2, \dots, x_N]$, X 中每一个样本为 $x_i=[x_{i1}, x_{i2}, \dots, x_{iD}]^T$. X 的低维流形表示为 $Y=[y_1, y_2, \dots, y_N]$, 其中 $y_i=[y_{i1}, y_{i2}, \dots, y_{id}]^T, d \ll D$.

LLE 算法降维流程如下所示:

步骤一. 计算每个样本点到其他样本点的距离, 求出该样本点的 k 个近邻点, 各个样本之间的距离公式如下:

$$d(x_i, x_j) = \|x_i - x_j\| \quad (6)$$

步骤二. 计算每个样本点到其近邻点的权值.

这里定义一个误差函数, 如下:

$$\varepsilon(w) = \sum_{i=1}^n \left\| x_i - \sum_{j=1}^k w_{ij} x_{ij} \right\| \quad (7)$$

$\varepsilon(w)$ 最小时可求得权值. 其中 x_{ij} 为 x_i 的 k 个近邻点, w_{ij} 是 x_i 与 x_{ij} 之间的权值, 且要满足条件: $\sum_{i=1}^k w_{ij} = 1$, 这里求取 W 矩阵, 需要构造一个局部协方差矩阵 Q^i .

$$Q_{jm}^i = (x_i - x_{ij})^T (x_i - x_{im}) \quad (8)$$

将公式(8)与 $\sum_{i=1}^k w_{ij} = 1$ 相结合, 并采用拉格朗日乘子法, 即可求出局部最优化重建权值矩阵:

$$w_{ij} = \sum_{m=1}^k (Q_{jm}^i)^{-1} / \sum_{p=1}^k \sum_{q=1}^k (Q_{pq}^i)^{-1} \quad (9)$$

在实际运算过程中, Q^i 有可能是一个奇异矩阵, 此时必须对 Q^i 进行正则化操作, 如下所示:

$$Q^i = Q^i + rI \quad (10)$$

其中 r 是正则化参数, I 是一个 $k \times k$ 的单位矩阵.

步骤三. 将所有原始的高维样本点映射到低维空间中.

映射条件满足如下:

$$\varepsilon(Y) = \sum_{i=1}^n \left| y_i - \sum_{j=1}^k w_{ij} x_{ij} \right|^2 \quad (11)$$

其中, $\varepsilon(Y)$ 是损失函数, 当其最小时可以求得输入向量, y_i 是 x_i 的输出向量, $y_{ij}(j=1, 2, 3, \dots, k)$ 是 y_i 的 k 个近邻点, 且要满足两个条件:

$$\sum_{i=1}^n y_i = 0$$

$$\frac{1}{n} \sum_{i=1}^n y_i y_i^T = I$$

其中 I 是 $m \times m$ 的单位矩阵. 这里的 $w_{ij}(i=1, 2, 3, \dots, n;$

$j=1, 2, 3, \dots, n)$ 可以存储在 $n \times n$ 的稀疏矩阵 W 中, 当 x_j 是 x_i 的近邻点时, $W_{i,j}=w_{ij}$, 否则 $W_{i,j}=0$. 则损失函数可重写为:

$$\varepsilon(Y) = \sum_{i=1}^n \sum_{j=1}^n M_{i,j} y_i^T y_j \quad (12)$$

其中 M 是一个 $n \times n$ 的对称矩阵, 其表达式为:

$$M = (I - W)^T (I - W) \quad (13)$$

要使 $\varepsilon(Y)$ 的值最小, 则取 Y 为 M 的最小 m 个非零特征值所对应的特征向量. 在处理的过程中, 将 M 的特征值从小到大依次排列, 由于第一个特征值非常接近于零, 所以舍去第一个特征值. 一般取第 $2 \sim m+1$ 之间的特征值所对应的特征向量作为输出结果即高维样本降维后的结果.

2.2 SLLE 算法

虽然 LLE 算法在人脸表情识别中有很好的降维效果, 但是它没有考虑到样本的类别信息, 因此有了 SLLE 算法. 传统的 LLE 算法在第一步寻找样本的 k 个近邻点时采用的距离公式是公式(6). 而 SLLE 算法在处理这一步时却有所不同, 它在计算距离时考虑到了样本点的类别信息. SLLE 算法的其余步骤与 LLE 算法是一致的.

SLLE 算法在计算两个样本点之间的距离时, 采用如下公式:

$$D' = D + \alpha \max(D) \Delta \quad (14)$$

其中 D' 表示更正后的距离; D 表示通过公式(6)计算出来的两个样本点之间的距离; $\max(D)$ 是表示类与类之间的最大欧式距离; Δ 取 0 或者 1, 当两个样本点属于同一类别时, Δ 取为 0, 否则取 1; Δ 是控制样本点之间的距离参数, $\alpha \in [0, 1]$ 是一个经验参数. 当 α 取为零时, 此时的 SLLE 算法和 LLE 算法相同.

LLE 算法是将所有样本一起进行处理, 而 SLLE 算法需要先处理训练样本后处理测试样本.

对训练样本的降维过程和 LLE 算法中的降维过程类似, 不同的是第一步计算距离的公式变了, 此时计算距离的公式是公式(14), 并且在此之前要先计算出不同表情分类之间的最大距离.

对测试样本的降维过程如下:

假设训练样本为 $X_{(d \times N)}$, 训练样本降维后的输出为 $Y_{(m \times N)}$, d 为训练样本的维数, m 为训练样本的输出维数即降维后的维数, N 为训练样本的个数. 设 X' 为测

试样本的集合. 主要算法分为三步:

步骤一. 选取一个 $x_e, x_e \in X'$, 将 X_e 加入 X 矩阵中, 则 X 变为 $dx(N+1)$ 的矩阵. 在训练样本中寻找 x_e 的 k 个近邻点, 此时采用的距离公式是公式(6)而不是公式(14).

步骤二. 计算出 x_e 与其 k 个近邻点间的权值系数:

$$\varepsilon(W) = \left| x_e - \sum_{j=1}^k w_{ej} x_{ej} \right|^2 \quad (15)$$

当 $\varepsilon(Y)$ 最小且 $\sum_{j=1}^k w_{ej} = 1$ 满足条件时可以求出权值, 其中 $x_{ej}(j=1, 2, 3, \dots, k)$ 是 x_e 的 k 个邻近点 w_{ej} 是 x_e 与其邻近点 x_{ej} 之间的权值.

步骤三. 计算 x_e 降维后的输出向量 y_e .

$$y_e = \sum_{j=1}^k w_{ej} y_{ej} \quad (16)$$

其中 y_{ej} 为测试样本降维过程中 x_{ej} 的输出向量.

2.3 改进的 SLLE 算法

SLLE 算法虽然考虑到了样本的类别信息, 但是还有很多不足的地方. 比如针对人脸表情识别这个问题, SLLE 算法没有考虑到各种表情之间的关系.

本文采用 JAFFE 数据库作为实验数据, 该数据库表情有 7 种分类: 中性、生气、厌恶、恐惧、高兴、悲伤、惊讶. 在计算最大类间距离时, 中性表情可以看成是所有表情的中心, 其他 6 种表情可以看成是在 6 维空间中 6 个不同的维度上, 所有表情都是以中性表情为中心, 在 6 个不同维度上波动的.

此时对训练数据的降维过程与 SLLE 中对训练数据的降维过程中的操作类似, 不过公式(14)中的 $\max(D)$ 依然表示类与类之间的最大距离, 但是当两个类之间有一个是中性表情时, 直接计算出两个类之间的最大距离, 当两个类中都不是中性表情时, 根据两个类到中性表情的最大距离, 求出两个类之间的最大距离, 计算公式如下:

$$d = \sqrt{d_1^2 + d_2^2} \quad (17)$$

其中 d_1 表示一个表情类别到中性表情的最大距离, d_2 表示另一个表情类别到中性表情的最大距离.

对测试数据的降维过程与 SLLE 相同.

3 实验结果分析

本文对人脸表情进行识别的实验数据来自日本女

性人脸表情 JAFFE 库. JAFFE 库由 10 个人共 213 幅静态表情图像组成, 包含 7 种表情(中性、生气、厌恶、恐惧、高兴、悲伤、惊讶). 这些图像都是 256×256 的 8 位灰度图像.

采用交叉校验方式进行测试来降低测试结果的偶然性, 保证测试结果的一般性. 将所有样本随机平分为 10 份, 其中有 7 份中都是 21 张图像, 另外 3 份中都是 22 张图像. 每次依次取其中的 9 份作为训练样本而剩余的 1 份作为测试样本, 最后再将得到的 10 个识别率取平均得到整体识别率.

LLE、SLLE、改进的 SLLE 三种算法 k 取不同值时的表情识别率如图 2 所示, 纵轴表示表情的识别率, 单位是 %, 横轴表示近邻点的个数 k .

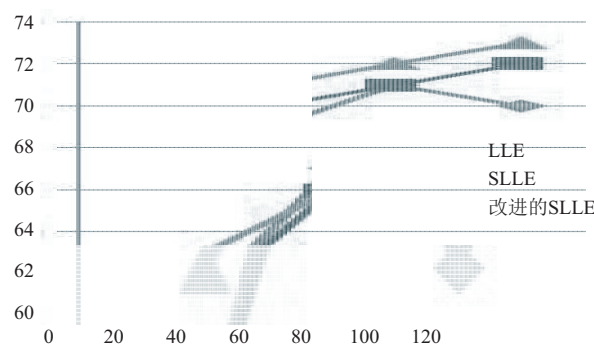


图 2 不同 k 值下面的表情识别率

基于 Gabor 特征提取的 PCA, LLE, SLLE, 改进的 SLLE 4 种降维算法的实验结果对比如表 1.

表 1 4 种降维算法的识别率

数据降维算法	识别率 (%)
PCA	56
LLE	71
SLLE	72
改进的SLLE	73

实验结果表明改进的 SLLE 算法比 PCA, LLE, SLLE 有更好的识别率, 是一个有效的算法.

4 结语

本文利用 Gabor 变换来进行人脸表情图像的特征提取, 利用 LLE 系列算法来进行数据降维. LLE 算法在人脸表情识别中有很好的降维效果, 但是它没有考虑到样本的类别信息. SLLE 算法虽然考虑到了样本的类别信息, 但是针对人脸表情识别这个问题, 没有考虑

到各种表情之间的关系. 本文针对人脸表情识别这个问题, 提出了一种改进的 SLLE 算法, 该算法假定了中性表情是所有表情的中心. 该算法在 JAFFE 库上获得了比 LLE 算法与 SLLE 算法更好的表情识别率, 它是一种有效的算法.

参考文献

- 1 Kim DJ. Facial expression recognition using ASM-based post-processing technique. *Pattern Recognition and Image Analysis*, 2016, 26(3): 576–581. [doi: [10.1134/S105466181603010X](https://doi.org/10.1134/S105466181603010X)]
- 2 Lozano-Monator E, López MT, Fernández-Caballero A, *et al.* Facial expression recognition from webcam based on active shape models and support vector machines. Pecchia L, Chen LL, Nugent C, *et al.* *Ambient Assisted Living and Daily Activities*. Cham, Switzerland: Springer, 2014, (8868): 147–154.
- 3 Wang L, Li RF, Wang K. A novel automatic facial expression recognition method based on AAM. *Journal of Computers*, 2014, 9(3): 608–617.
- 4 Edwards GJ, Cootes TF, Taylor CJ. Face recognition using active appearance models. Burkhardt H, Neumann B. *Computer Vision-ECCV'98*. Berlin, Heidelberg, Germany: Springer, 1998. 581–595.
- 5 左坤隆, 刘文耀. 基于活动外观模型的人脸表情分析与识别. *光电子·激光*, 2004, 15(7): 853–857.
- 6 Zhao XM, Zhang SQ. Facial expression recognition based on local binary patterns and kernel discriminant isomap. *Sensors*, 2011, 11(10): 9573–9588.
- 7 朱健翔, 苏光大, 李迎春. 结合 Gabor 特征与 Adaboost 的人脸表情识别. *光电子·激光*, 2006, 17(8): 993–998.
- 8 刘晓旻, 章毓晋. 基于 Gabor 直方图特征和 MVBoost 的人脸表情识别. *计算机研究与发展*, 2007, 44(7): 1089–1096.
- 9 Chakrabarti D, Dutta D. Facial expression recognition using PCA and various distance classifiers. In: Sengupta S, Das K, Khan G, eds. *Emerging Trends in Computing and Communication*. New Delhi, India: Springer, 2014. 79–85.
- 10 Roweis ST, Saul LK. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 2000, 290(5500): 2323–2326. [doi: [10.1126/science.290.5500.2323](https://doi.org/10.1126/science.290.5500.2323)]
- 11 Shao H, Chen S, Zhao JY, *et al.* Face recognition based on subset selection via metric learning on manifold. *Frontiers of Information Technology & Electronic Engineering*, 2015, 16(12): 1046–1058.
- 12 de Ridder D, Kouropyteva O, Okun O, *et al.* Supervised locally linear embedding. In: Kaynak O, Alpaydin E, Oja E, *et al.*, eds. *Artificial Neural Networks and Neural Information Processing-ICANN/ICONIP 2003*. Berlin, Heidelberg, Germany: Springer, 2003. 333–341.
- 13 应自炉, 李景文, 张有为, 等. 基于表情加权距离 SLLE 的人脸表情识别. *模式识别与人工智能*, 2010, 23(2): 278–283.