

FFT 协同特征空间的图像表示方法

孙 玲¹, 陈德运¹, 李 骛¹, 付立军¹, 杨 润², 于 梁³

(1. 哈尔滨理工大学 计算机科学与技术学院, 哈尔滨市 150080;

2. 哈尔滨工业大学 计算机科学与技术学院, 哈尔滨市 150001;

3. 酒泉卫星发射中心, 敦煌 736200)

摘 要: 快速傅里叶变换(FFT)方法已经在图像识别中有广泛的应用。但是,FFT 方法面临一些挑战,比如:不同角度的遮挡、变化的光照和多变的面部表情等。将快速傅里叶变换和特征空间的图像表示方法融合起来解决上述问题。有以下阶段:①使用 FFT 从原始图像中提取频谱特征。②利用高斯核方法在特征空间中获取新的特征。新的特征和原始图像的训练样本分别使用稀疏表示来获得稀疏解。新的特征和原始图像的测试样本可以使用上述稀疏解及其训练样本来分别计算得分。随后,可以利用得分和新得分进行图像分类。这一方法在图像分类上具有稀疏性和鲁棒性,非常容易实现。实验结果表明,本文提出的方法在图像分类上具有高的准确率。

关键词: 图像识别; FFT; 基于特征空间方法

DOI: 10.15938/j.jhust.2020.06.020

中图分类号: TP317.4 文献标志码: A 文章编号: 1007-2683(2020)06-0137-05

FFT Consolidated Feature Space Methods for Image Representation

SUN Ling¹, CHEN De-yun¹, LI Ao¹, FU Li-jun¹, YANG Run², YU Liang³

(1. School of Computer Science and Technology, Harbin University of Science and Technology, Harbin 150080, China;

2. School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China;

3. Jiuquan Satellite Launch Center, Dun Huang 736200, China)

Abstract: Fast Fourier Transform (FFT) method has been widely used in image recognition. However, FFT method faces some challenges, such as occlusion of different angles, varying lighting and changeable facial expression. We combine fast Fourier transform and feature space image representation method to solve the above problems. There are the following stages: ①Use FFT to extract spectral features from the original image. ②Use Gaussian kernel method to obtain new features in the feature space. The new features and the training samples of the original image use sparse representations to obtain sparse solutions. The new features and the test samples of original image can use the aforementioned sparse solution and its training samples to calculate scores respectively. Subsequently, the score and the new score can be used for image classification. This method has sparseness and robustness in image classification, and is very easy to implement. Experimental results show that the method proposed in this paper has high accuracy in image classification.

Keywords: image recognition; FFT; feature space-based method

收稿日期: 2019-05-15

基金项目: 国家自然科学基金(61501147); 黑龙江省自然科学基金优秀青年项目(Grant YQ2019F011); 黑龙江省青年创新人才计划(Grant UNPYSCT-2018203); 黑龙江省高等学校基本科研业务专项(Grant LGYC2018JQ013)。

作者简介: 孙 玲(1983—),女,硕士研究生;

陈德运(1962—),男,教授,博士研究生导师。

通信作者: 付立军(1985—),男,博士研究生,讲师,E-mail: fulijun85@163.com。

0 引言

图像表示是一项重要的识别技术,图像表示被广泛地应用到人脸识别、掌纹识别、遥感和医学诊疗等多个领域^[1]。在过去几年中,学者们已经提出了多种图像分类方法^[2],尤其在人脸识别方面。然而,图像分类技术在现实的应用中仍然面临诸多挑战,如:不同角度的遮挡、变化的光照和多变的面部表情等^[3-4]。因此,如何更好地表示图像和提高图像分类性能成为研究的热点。

图像分类需要更多的训练样本来挖掘在不同条件下的图像特征。然而,到目前为止,所收集的图像受到采集的空间和环境的限制,导致在实际中有用的训练样本数量非常有限。尤其,在某些环境下,每类的训练样本数量只有一个,如:护照系统。以上例子说明不充足的训练样本限制了人脸识别技术的发展^[5]。

为了提高人脸识别技术,近年来一些学者提出了不同的方法来解决上述难题^[6]。合成图像是增强样本的最常用的方法。例如,Sharma等^[7]提出在变化的照明和不固定的姿势下使用单个2D图像来生成虚拟图像。该方法利用合成训练样本来对测试样本(也称为识别面)进行分类。Beymer等^[8]提出利用不同姿势的面部图像生成虚拟图像,然后利用原始图像和虚拟图像来识别身份;Tang等^[9]采用光流和表达定量图像构建虚拟图像,扩展了训练样本数量,提高了图像分类的正确率;Jung等^[10]使用噪音来获取损坏的图像;Thian等^[11]利用简单的几何变换构建虚拟图像;Thomes^[12]采用单个正面人脸来生成3D人脸模型。生成的虚拟图像用于处理一个训练样本问题^[13-14]。单一的方法在图像特征提取过程中,在不同场景下会遗漏一些重要特征,因此,利用多种方法融合来表示图像已成为近年来的研究热点。

以上的研究表明训练样本问题在图像处理上已经引起了很多的关注,许多学者提出了不同的方案来解决这个问题^[15-16]。例如,Xu等^[17]利用所获得的镜面和原始面部图像整合以识别人脸。Wang等^[18]将Gabor和二维主成分分析(2DPCA)融合用于人脸识别,其获得的精度高于单一2DPCA和(最近邻)NN以及NN和Gabor的组合。Yang等^[19]利

用完整的字典来获得稀疏系数表示原始图像。该方法利用选择最大融合规则来组合所获得的系数,通过组合系数构建新图像。它可以融合新图像和原始图像来获得分数并使用稀疏分类器来分类图像,对图像识别有很好的作用。Kong等^[20]使用生化离子交换模型来解决多焦点图像融合技术的问题。

我们注意到图像的多种表示不仅能扩充训练样本,而且提出一个强鲁棒性的图像识别方法^[21]。因此,我们提出一种FFT协同特征空间的图像表示方法。这种方法首先可以使用FFT算法从原始图像中提取频率特征。并且,我们可以将频率特征和原始图像分别分为训练数据集和测试数据集两部分。可以利用高斯核方法在特征空间中获得新的训练数据集和测试数据集,这对应于原始图像中的原始图像和频率特征。使用所有训练样本来线性表示测试样本并获得线性系数,分别获得与原始图像和频率特征相对应的分数。融合上述操作所获得的分数以获得新分数用于分类图像。这种新方法首先集成了FFT和基于特征空间的表示,这对图像分类更有效。它提取频率特征与原始图像互补,该方法是非常容易实现的。同时,具有稀疏性,这使得图像表示有更高的准确性。为了测试所提方法的性能,我们选择公共数据集Georgia Tech(GT)^[22]和(Alex Martinez与Robert Benavente)(AR)^[23]来进行实验。

1 提出的方法

1.1 获取原始图像的特征

快速傅里叶变换(FFT)是离散傅里叶变换(DFT)的快速算法^[24]。它已广泛应用于信号处理、计算机视觉和模式识别领域。本文使用FFT算法从原始图像中提取频率特征。DFT和FFT算法^[25]描述如下。我们假设有限序列的长度是 $x(c)$ 。DFT算法由等式(1)表示:

$$X(c) = \text{DFT}[x(n)] = \sum_{n=0}^{l-1} x(n) w_N^{cn} = \sum_{n=0}^{l/2-1} x(n) w_N^{cn} + \sum_{n=N/2}^{l-1} x(n) w_N^{cn}, c=0,1,\dots,l-1 \quad (1)$$

当 $n' = n - \frac{l}{2}$ 将式(1)转换成式(2):

$$X(c) = \sum_{n'=0}^{l/2-1} [x(n') + w_l^{c* l/2} x(n' + l/2)] w_l^{n'c} \quad (2)$$

当 c 为偶数时, $(-1)^c$ 为 $1, c = 2r (r = 0, 1, \dots, l/2 - 1)$

当 c 为奇数时, $(-1)^c$ 为 $-1, c = 2r + 1$ 。

因此,我们可以将(2)简化为式(3)和(4)

$$X(2r) = \left[\sum_{n'=0}^{l/2-1} x(n') + x(n' + l/2) \right] w_l^{2nr} = \sum_{n'=0}^{l/2-1} x_1(n') w_{l/2}^{nr} = \text{DFT} [x_1(n')]_{l/2} \quad (3)$$

$$X(2r+1) = \sum_{n'=0}^{l/2-1} [x(n') - x(n' + l/2)] w_l^{2nr} = \sum_{n'=0}^{l/2-1} x_2(n') w_{l/2}^{nr} = \text{DFT} [x_2(n')]_{l/2} \quad (4)$$

因为图像是二维的,所以我们可以模仿一维 FFT 来获得二维 FFT。假设 $f(x, y)$ 它代表一个原始图像,其大小是矩阵 $M \times N$ 。并用来表示要获得的频率变量 $F(u, v)$ 。

当 $0 \leq x \leq M-1, 0 \leq y \leq N-1, u = 0, 1, 2, \dots, M-1$ 和 $v = 0, 1, 2, \dots, N-1$ 时

$$F(u, v) = \text{DFT} [f(x, y)] = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(ux/M + vy/N)} \quad (5)$$

1.2 获得新的表示方法

本文利用 FFT 提取频率特征,利用核方法在核空间提取特征,并利用稀疏方法分别获得分数。最后,融合以上获得的分数并使用新分数对图像进行分类。我们可以利用 Ref^[23] 分别获得原始图像和频率特征的误差 $e_k = \|K_Y - g_k\|^2$ 。此外, $e_k = \|K_Y - g_k\|^2$ 可用来评价原始图像和频率特征图像分类的误差效果。

2 本文表示方法的优点

本文表示方法在图像分类上具有稀疏性和高准确率。在数据集 GT 上验证其性能。

图1表示当训练样本的数量 $1 \sim 100$ 时特征空间中的原始图像的值。图2表示当训练样本数量 $1 \sim 100$ 时在特征空间中获得的特征和原始图像的值。如图1和图2所示,特征空间中获得的特征和原始图像的值比特征空间中的原始图像的值更接近0。

当训练样本的数量 $1 \sim 100$ 时所提出的方法具有部分稀疏性,这使得它具有更高的图像识别精确率。

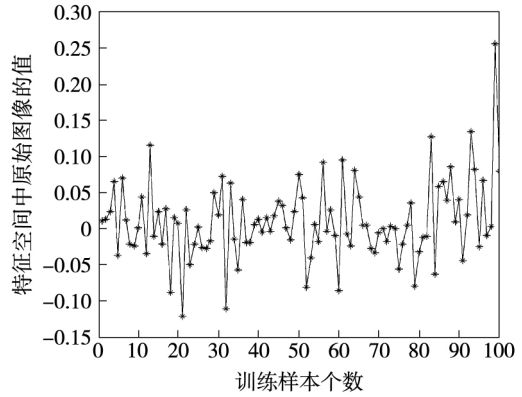


图1 当训练样本的数量 $1 \sim 100$ 时,特征空间中的原始图像的值

Fig. 1 The values of original images in the feature space when the number of training samples varies from 1 to 100

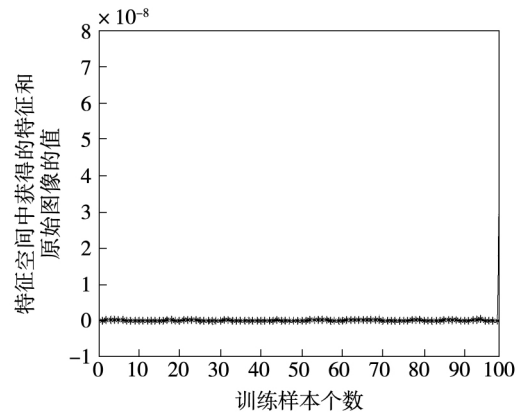


图2 当训练样本的数量 $1 \sim 100$ 时,在特征空间中获得的特征和原始图像的值

Fig. 2 The values of obtained features and original images in the feature space when the number of training samples varies from 1 to 100

3 方法性能验证

使用公开的 GT 和 AR 数据集来进行实验并验证所提出方法的性能。同时,在特征空间中使用单个获得的频率特征和单个原始图像进行对比实验。

3.1 GT 数据集

GT 人脸数据集从 50 个不同的测试者中获得,每个测试者具有 15 个样本。这些图像是佐治亚理工学院 1990 年 7 月 1 日至 1999 年 11 月 15 日之间收集的。这些图像包括不同的面部表情和不同的光

照条件。图像保存为“.jpg”格式。图3显示了来自GT面部数据集的部分面部图像。

使用每个类中不同数量的图像作训练样本,并且每个类的其他图像被用作测试样本。如表1所示,显示了图像在GT数据集中分类的错误率。在表1中,获得的特征频率和原始图像被融合并用于特征空间中的图像进行分类,从而降低错误率。

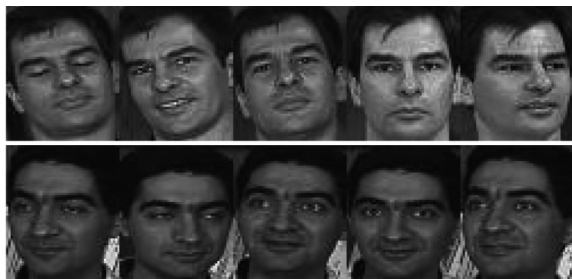


图3 来自GT面部数据集的部分面部图像

Fig. 3 Partial face images from the GT face dataset

表1 GT图像分类的错误率

Tab. 1 Errors rates of image classification on the GT %

每类训练样本的个数	6	7	8	9	10
本文表示方法	37.78	26.25	23.71	22.67	20.00
快速傅里叶变换和核表示方法	30.00	28.25	26.00	25.00	23.20
核表示方法	37.11	33.00	30.86	28.33	27.20

3.2 AR数据集

AR人脸数据集用于设计实验。AR数据集来自126个不同的测试者。这些照片由Aleix Martinez和Robert Benavente收集。这些图像包括不同的面部表情和不同的遮挡。图像保存为“.bmp”格式。图4示出了来自AR面部数据集的部分面部图像。



图4 来自AR面部数据集的部分面部图像

Fig. 4 Partial face images from the AR face dataset

使用每个类中不同数量的图像用作训练样本,并且每个类的其他图像被用作测试样本。如表2所示,显示了图像在AR人脸数据集中分类的错误率。在表2中,获得的特征频率和原始图像被融合并用于对特征空间中的图像进行分类,从而降低错误率。

表2 AR图像分类的错误率

Tab. 2 Errors rates of image classification on the AR %

每类训练样本的个数	14	15	16
提出的方法	8.13	7.20	7.00
快速傅里叶变换和核表示方法	38.13	40.53	43.17
核表示方法	37.22	39.62	41.58

4 结论

提出了一种基于快速傅里叶变换和特征空间的图像分类方法,使得特征与原始图像互补。此外,该方法简单且容易实现。本文提出的方法具有稀疏性和鲁棒性,这是提高图像分类准确率的一个重要原因。实验证明,该方法具有良好的实际应用价值。

参考文献:

- [1] XU Yong, ZHU Xingjie, LI Zhengming, et al. Using the Original and 'Symmetrical Face' Training Samples to Perform Representation Based Two-step Face Recognition [J]. Pattern Recognition, 2013, 46(4): 1151.
- [2] TIAN Chunwei, ZHANG Qi, SUN Guanglu, et al. Multiple Methods for Wechat Identification [C] // 2016 6th International Conference on Advanced Design and Manufacturing Engineering (ICADME 2016), Zhuhai, China. July 23 - 24 2016, 2016: 598.
- [3] TIAN Chunwei, ZHANG Qi, SUN Guanglu, et al. Linear Discriminant Analysis Representation and CRC Representation for Image Classification [C] // 2016 2nd IEEE International Conference on Computer and Communications, Chengdu, China. October 14 - 17, 2016: 755.
- [4] XU Yong, ZHANG Bob, ZHONG Zuofeng. Multiple Representations and Sparse Representation for Image Classification [J]. Pattern Recognition Letters, 2015, 68(S1): 9.
- [5] PISHCHULIN L, GASS T, DREUW P. et al. Image Warping for Face Recognition: From Local Optimality Towards Global Optimization [J]. Pattern Recognition, 2012, 45(9): 3131.
- [6] WEN Jie, FANG Xiaozhao, XU Yong, et al. Low-rank Representation with Adaptive Graph Regularization [J]. Neural Networks, 2018, 108: 83.

- [7] SHARMA A, Dubey P, Tripathi, et al. Pose Invariant Virtual Classifiers from Single Training Image Using Novel Hybrid-eigen-faces [J]. Neurocomputing, 2010, 73(10/12): 1868.
- [8] BEYMER D, POGGIO T. Face Recognition from One Example View [C] // IEEE International Conference on Computer Vision, Cambridge, USA. June 20 – 23, 1995: 500.
- [9] TANG Bin, LUO Siwei, HUANG Hua. High Performance Face Recognition System by Creating Virtual Sample [C] // Proceedings of 2003 International Conference on Neural Networks and Signal Processing, 2003: 972.
- [10] JUNG H, HWANG B, LEE S. Authenticating Corrupted Face Image Based on Noise Model [C] // Proceedings-Sixth IEEE International Conference on Automatic Face and Gesture Recognition, Seoul, Korea. May 17 – 19, 2004: 272.
- [11] THIAN N P H, MARCEL S, BENGIO S. Improving Face Authentication Using Virtual Samples [C] // 2003 IEEE International Conference on Accoustics, Speech, and Signal Processing Hong Kong, China, April 6 – 10, 2003: 233.
- [12] THOMAS V. Synthesis of Novel Views from a Single Face Image [J]. International Journal of Computer Vision, 1998, 28(2): 103.
- [13] PARTHA N, FEDERICO G, TOMASO P. Incorporating Prior Information in Machine Learning by Creating Virtual Examples [J]. Proceedings of the IEEE. 1998, 86(11): 2196.
- [14] MARTINEZ A M. Matching Expression Variant Faces [J]. Vision Research, 2003, 43(9): 1047.
- [15] FEI Luke, XU Yong, TANG Wenliang, et al. Double-orientation Code and Nonlinear Matching Scheme for Palmprint Recognition [J]. Pattern Recognition, 2016, 49: 89.
- [16] TIAN Chunwei, ZHANG Qi, SUN Guanglu, et al. Weighted Nearest Neighbor Algorithm and Collaborative Representation for Image Classification [J]. Journal of Computational and Theoretical Nano-science, 2016, 13(12): 9065.
- [17] XU Yong, LI Xuelong, YANG Jian, et al. Integrate the Original Face Image and its Mirror Image for Face Recognition [J]. Neuro-computing, 2014, 131: 191.
- [18] WANG Jian, CHENG Jian. Face Recognition Based on Fusion of Gabor and 2DPCA Features [C] // ISPACS 2010 – 2010 International Symposium on Intelligent Signal Processing and Communication Systems, 2010: 5704688.
- [19] YANG Bin, LI Shutao. Multifocus Image Fusion and Restoration with Sparse Representation [J]. IEEE Transactions on Instrumentation and Measurement, 2010, 59(4): 884.
- [20] KONG Weiwei, LEI Yang. Multi-focus Image Fusion Using Bio-chemical Ion Exchange Model [J]. Applied Soft Computing Journal 2017, 51: 314.
- [21] XU Yong, FEI Lunke, ZHANG David. Combining Left and Right Palmprint Images for More Accurate Personal Identification [J]. IEEE Transactions on Image Processing, 2015, 24(2): 549.
- [22] GOEL N, BEBIS G, NEFIAN A. Face recognition experiments with random projection. Proceedings of SPIE –The International Society for Optical Engineering, 2005, 5779: 426.
- [23] MARTINEZ A, BENAVENTE R. The AR face database [J]. CVC Technical Report #24, 1998: 1.
- [24] FIALKA O, CADIK M. FFT and Convolution Performance in Image Filtering on GPU [C] // Proceedings of the International Conference on Information Visualisation, London, United Kingdom. July 5 – 7, 2006: 609.
- [25] TIAN Chunwei, ZHANG Qi, ZHANG Jian, et al. 2D-PCA Representation and Sparse Representation for Image Recognition [J]. Journal of Computational and Theoretical Nanoence, 2017, 14(1): 829.

(编辑: 温泽宇)