

# 一种基图像提取和内容无关图像重构方法研究

蓝章礼 申德兴 曹娟 张玉欣

重庆交通大学信息科学与工程学院 重庆 400074

(lzl17309@126.com)

**摘要** 图像作为一种典型信号,理论上可由一系列基本信号构成。为寻找一组可重构图像的基本信号,提出了基于特征的基图像提取和重构方法,使得可由任意图像集进行基图像提取并可由提取的基图像重构内容无关的任意图像。使用特征提取算法从训练集图像中分解出一系列基图像,阐述了基图像分解和提取的算法流程,通过将测试集图像投影到 $k$ 个基图像构成的空间中得到投影系数,建立由投影系数和基图像重构原图像的方法和过程。实验结果表明,通过控制基图像数量 $k$ ,图像的重构误差和质量可以达到较高要求,基图像的提取和重构的图像具备内容无关性,同时,该方法对于图像抽象特征的理解、深度神经网络应用等具有重要作用。

**关键词** 图像重构;基图像;特征提取;特征向量;内容无关

**中图法分类号** TP301

## Content-independent Method for Basis Image Extraction and Image Reconstruction

LAN Zhang-li, SHEN De-xing, CAO Juan and ZHANG Yu-xin

School of Information Science and Engineering, Chongqing Jiaotong University, Chongqing 400074, China

**Abstract** As one kind of typical signals, an image can theoretically be composed of a series of basic signals. In order to find a set of basic signals to reconstruct images, a method for obtaining basis images based on feature extraction and reconstructing images from them is proposed. It makes possible to obtain the basis images from any set of images and to reconstruct images from the obtained ones because it is content-independent. The algorithm flow of extracting a series of basis images from the training set of images by feature extraction algorithm is described. The system of reconstructing the original image from the projection coefficient and basis images by projecting the set of test images into the space formed by the  $k$  basis images is developed. The experimental results show that, by controlling the number of basis images, the error and quality of reconstructed images can achieve higher requirement, and the method for basis images extraction and image reconstruction is content-independent. At the same time, this method plays an important role in the understanding of abstract features of images and the deep neural network.

**Keywords** Image reconstruction, Basis image, Feature extraction, Feature vectors, Content-independent

## 1 引言

图像是人类从外界获取信息的最重要的途径之一<sup>[1]</sup>。面对规模日益庞大的图像数据,对图像的降维和识别变得越来越重要<sup>[2-5]</sup>。图像是一种典型的二维信号,将信号处理领域的研究方法引入图像领域是十分普遍的研究思路,类似于傅里叶级数的思想,即将一个复杂信号表示为一系列简单的正弦信号的线性组合。本文提出一种方法从任意图像集中获得基图像并重构任意图像,进而优化后续如图像识别等算法。

目前广泛使用的通过深度神经网络学习图像特征进行识别的方法缺乏较为严谨的理论支撑。以卷积神经网络为例,训练模型时需要设定大量的超参数,而对这些超参数的选择往往是基于实验结果去优化而非基于理论。当下深度神经网络面临瓶颈与人们无法理解神经网络得到的高度抽象的特征及其学习过程是分不开的。运用科学思维去分析深度神经网络

是突破的关键,物理学家 Mehta 等<sup>[6]</sup>证明特定情况下统计物理提取的相关特征和深度神经网络提取的相关特征具有一致性。本文研究图像重构方法来分析图像的基本结构,这些结构或许与神经网络习得的某些特征相似,进而对这些抽象特征做出解释。

特征提取在图像领域得到深入研究和广泛应用。常见的特征提取有5类。1)统计特征,将图像的灰度等统计数据作为特征,如 Niu 等基于灰度直方图特征识别草莓病虫害<sup>[7]</sup>。2)纹理特征,将图像中反复出现的局部模式和其排列规则作为特征,如 Lan 等提取局部纹理特征<sup>[8]</sup>来对医学图像进行识别检索。3)代数特征,将图像作为矩阵并提取其代数特征,如 Xia 等提出一种概率主成分分析算法对高光谱遥感图像进行特征提取并分类<sup>[9]</sup>。4)变换系数特征,将对图像进行傅里叶变换或小波变换等滤波变换得到的系数作为特征,如 Qayyum 等提出一种利用小波变换特征进行面部表情识别的

基金项目:重庆市教委科学技术研究项目(KJQN201800716)

This work was supported by Project of Science and Technology Research Program of Chongqing Education Commission of China (KJQN201800716).

通信作者:申德兴(493211790@qq.com)

算法<sup>[10]</sup>,Fang 等将傅里叶变换和独立成分分析结合提取木材图像特征<sup>[11]</sup>。5)机器学习特征,利用神经网络等方法自动学习图像的特征,如 Krizhevsky 等构建的 8 层卷积神经网络<sup>[12]</sup>。目前利用所取得的特征作为基图像并重构原图像的研究比较少见。

本文结合两种类型图像数据,运用特征提取中的经典算法 PCA(主成分分析)获得一组基图像并重构不同类型的原图像<sup>[13-14]</sup>。首先使用 PCA 从训练集中分解出基图像,然后选取其中  $k$  个构成一个空间,将原图像投影到该空间中得到系数,所得系数可以与基图像结合重构回原图像。为表现基图像重构原图像时具有内容无关性,将自然图像和人脸图像分别作为训练集分解出基图像进行重构,结果表明基图像具有一般性,重构的图像与训练内容无关,且能起到降维效果,有利于后续应用,如图像识别。

## 2 基于特征提取的内容无关图像重构方法

### 2.1 图像重构流程

图像重构的基本流程如图 1 所示。



图 1 图像重构流程

Fig. 1 Process of image reconstruction

先将图像数据分为训练集和测试集,比如训练集为自然图像,包含花、草、动物等,测试集为人脸图像;然后对训练集进行预处理,包括从训练集中随机截取大小为  $m \times m$  和数目  $n$  的图像块构成数据矩阵  $S$ , $S$  的每一列为一个  $m \times m$  图像块转变的  $m^2$  的向量。然后对数据矩阵  $S$  使用 PCA 算法得到基图像;将要投影到基图像构成空间上的原图像分割为  $m \times m$  的图像块并分别投影,得到系数矩阵  $W$ ;最后结合基图像和系数矩阵  $W$  重构回原图像。

随机截取的图像块一般较小,本文设为  $32 \times 32$ ,图像块数目较多,本文设为 50000;用于构成空间的基图像数目  $k$  小于图像块维数  $m^2$ , $k$  值的选择可以依需求调整。

### 2.2 算法流程

对训练集使用 PCA 得到基图像的算法如算法 1 所示,其流程如图 2 所示。

#### 算法 1

输入:数据矩阵  $S$ ,基图像的数量  $k$

输出:基图像矩阵  $B$

1. 输出预处理后得到的数据矩阵  $S$ ,需要的基图像数量为  $k$ ;
2. 计算平均基图像向量  $I_{avg}$ ;
3. 数据矩阵  $S$  中每列分别减去  $I_{avg}$ ;
4. 计算协方差矩阵

$$C = SS^T \quad (1)$$

的特征向量矩阵  $Vec$  和对应特征值组成的向量  $Val$ ,初步获得基图像;

5. 依特征值大小将  $Vec$  降序排序,图像在越大的特征值对应的特征向量方向上的投影量越多,通过排序可以方便后续对  $k$  个基图像的选择;
6. 依次遍历  $Vec$  每列,将基图像正规化;
  - 6.1. 初始化当前计算列下标  $i=1$ ,计数器  $num=0$ ;
  - 6.2. 判断是否已经遍历  $Vec$  所有列,是则转向步骤 7,否则转向步骤 6.3;
  - 6.3. 将  $Vec(i)$  正规化,即变作模长为 1 的单位向量;

- 6.4. 计数器  $num=num+1$ ;
- 6.5. 下标  $i=i+1$ ,转向 6.2;
7. 判断  $k$  是否小于  $num$ ,若是,则选取  $Vec$  前  $k$  列构成基图像矩阵  $B$ ,否则选取前  $num$  列构成基图像矩阵  $B$ ,并警告仅有  $num$  个基图像;
8. 结束。

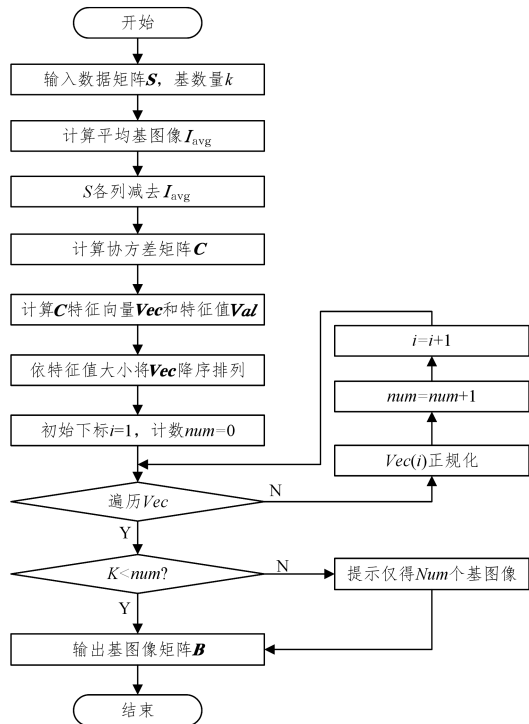


图 2 利用 PCA 得到基图像的流程

Fig. 2 Process of obtaining basis images by using PCA

(1)预处理中将每个图像块  $I_i$  转变为向量  $I_i'$  的操作如式(2)所示:

$$I_i = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mm} \end{bmatrix}_{m \times m} \xrightarrow{yields} \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \\ \vdots \\ a_{1m} \\ \vdots \\ a_{mm} \end{bmatrix}_{m^2 \times 1} = I_i' \quad (2)$$

数据矩阵  $S$  由共  $n$  个图像块向量  $I_i'$  构成,即:

$$S = [I_1' \quad I_2' \quad \cdots \quad I_n'] \quad (3)$$

(2)基图像个数  $k$  依具体需求而定,本文不做讨论。

(3)计算平均基图像向量  $I_{avg}$ ,如式(4)所示:

$$I_{avg} = \frac{1}{n} \sum_{i=1}^n I_i' \quad (4)$$

(4)数据矩阵  $S$  每列分别减去  $I_{avg}$ ,即:

$$S = [I_1' - I_{avg} \quad I_2' - I_{avg} \quad \cdots \quad I_n' - I_{avg}] \quad (5)$$

其目的在于只留下每个图像的相互间用于区分的特征,移除共有的信息。

(5)计算协方差矩阵  $C$  的意义在于,设灰度图像  $I$  是一个  $n^2$  维的向量,如一张  $32 \times 32$  的图像是一个 1024 维的向量,那么它相当于一个维度为 1024 的空间中的一个点。如果图像可以由一系列如点、直线、折线及其组合等基本结构合成,那么图像映射在这个高维空间中的点不会随机分布,进而可以用一个相对低维的子空间来描述图像。使用 PCA 即为了寻

找最能反映图像在整个高维空间内分布的向量,这些向量是与训练集图像相对应的协方差矩阵的特征向量。

(6)将  $\text{Vec}(i)$  正规化,如式(6)所示:

$$\widehat{\text{Vec}}(i) = \frac{\text{Vec}(i)}{\|\text{Vec}(i)\|} \quad (6)$$

其中,  $\text{Vec}(i)$  为特征向量矩阵  $\text{Vec}$  第  $i$  列向量,  $\|\text{Vec}(i)\|$  为  $\text{Vec}(i)$  的欧几里得范数。

基图像矩阵  $\mathbf{B}$  的每一个列向量转换为矩阵形式时外观类似于点、直线、折线等简单图像的组合,本文称这些向量为基图像。图 3 即算法所得部分基图像示意图。

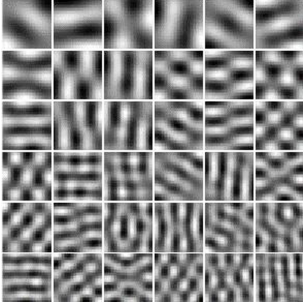


图 3 部分基图像

Fig. 3 Partial base image

### 2.3 图像的投影与重构

为了从基图像重构原图像,需要将所有基图像加权,重构图像等于所有基图像加权的线性组合。每个基图像具有相应的权重系数,这个系数表示基图像在原图像中出现的程度。如果使用从训练集中提取出的所有基图像则可以准确地重构出原图像。本文使用一部分基图像,重构图像是原图像的近似,但由于选择的是最重要的基图像,所以可以使得重构图像因仅使用部分基图像所造成的损失最小化。

将测试集的图像投影到  $\mathbf{B}$  构成的空间上时,首先需要将原图像分割为  $m \times m$  的图像块并转换为  $m^2$  的列向量,记为

$\mathbf{I}_{\text{part}}$ , 每个图像块向量  $\mathbf{I}_{\text{part}}$  减去平均基图像向量  $\mathbf{I}_{\text{avg}}$  后可以由  $\mathbf{B}$  中的基图像线性组合而成,即:

$$\mathbf{I}_{\text{part}} = \mathbf{I}_{\text{avg}} + \sum_{j=1}^k \omega_j \mathbf{B}(j) \quad (7)$$

每个基图像对应的系数  $\omega_j$  可由式(8)计算:

$$\omega_j = \mathbf{B}(j)^T (\mathbf{I}_{\text{part}} - \mathbf{I}_{\text{avg}}) \quad (8)$$

原图像块  $\mathbf{I}_{\text{part}}$  在该空间上的特征系数  $\mathbf{W}$  为:

$$\mathbf{W} = [\omega_1, \omega_2, \dots, \omega_k]^T \quad (9)$$

其中,  $k$  为基图像矩阵  $\mathbf{B}$  的列数。

由此,可通过基图像矩阵  $\mathbf{B}$  和图像的特征系数  $\mathbf{W}$  重构原图像块,再将图像块依次拼合即得到原图像,在后续的应用如图像识别、储存中可以使用维数大大降低的  $\mathbf{W}$ 。

### 3 实验结果分析

本文提出的基图像提取与内容无关图像重构方法使用 Matlab 平台实现。数据集为自然图像和谢菲尔德大学的人脸数据库。前者共有 13 张灰度图像,包含花、草、动物、山脉等;后者包含 20 位不同角度的人脸图像,共 564 张。为体现所得基图像重构为原图像时具有内容无关性,分别将自然图像和人脸图像作为训练集获得基图像并重构原图像。实验使用峰值信噪比 (PSNR) 评价重构图像的质量,其值越高越好,峰值信噪比的定义为:

$$\text{PSNR} = 10 \times \log_{10} \left( \frac{\text{MAX}_I^2}{\text{MSE}} \right) \quad (10)$$

其中,  $\text{MAX}_I$  为灰度图像的最大灰度值,即 255。MSE 为均方误差,定义为:

$$\text{MSE} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - I_r(i, j)]^2 \quad (11)$$

其中,  $I(i, j)$  为原始图像  $(i, j)$  处像素灰度值,  $I_r(i, j)$  为重构图像  $(i, j)$  处像素灰度值,  $m$  和  $n$  为图像的宽和高。

基图像数量  $k$  分别选取 10, 50, 200, 所得基图像如图 4 所示。

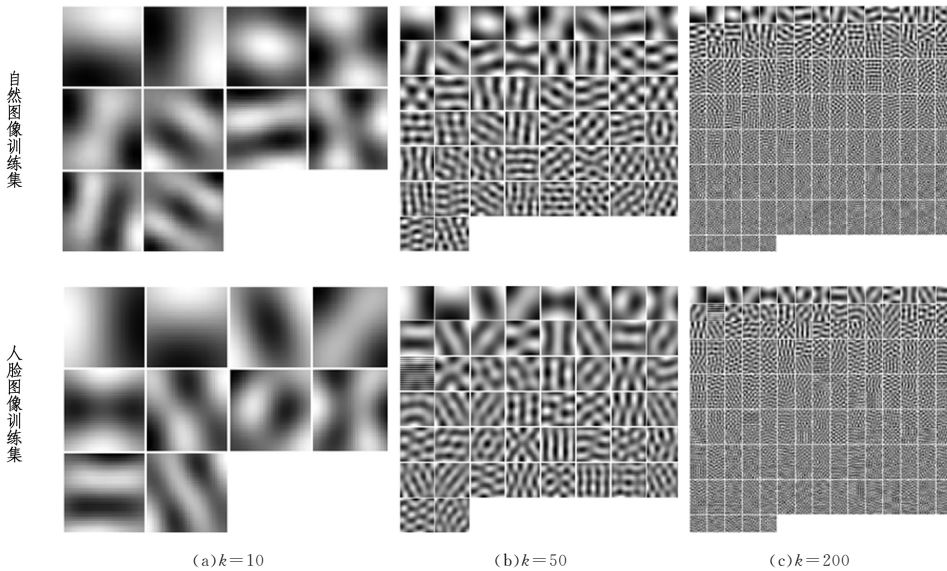


图 4 不同  $k$  值的基图像

Fig. 4 Base images with different  $k$  values

从图 4 中可以看出,基图像能量集中在点、直线、折线及其组合的简单结构上,图像内容易于理解。两种训练集的基图像直观上类似,只是在方向和细节上略有不同,基图像与

训练集中的图像内容无关。

图 5 展示了以自然图像为训练集得到的基图像对两张不同类型的图像在基图像数量  $k$  取不同值时的重构效果。随着

$k$  值增大,重构图像(图 5(b))仅具有大致轮廓,图 5(c)比较模糊,而图 5(d)与原图像基本无法区分,且自然图像和人脸图像的重构结果均表现良好。

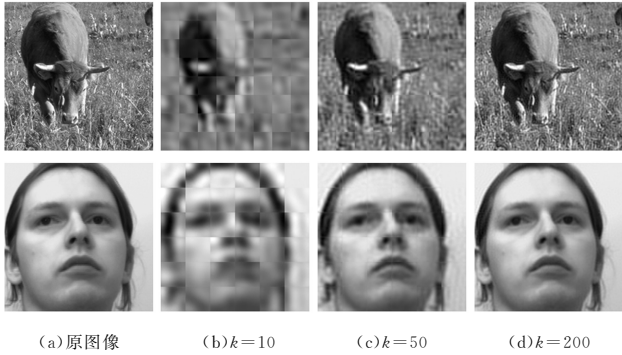


图 5 图像重构效果

Fig. 5 Effect of image reconstruction

表 1 的数据表明:不同训练集得到的基图像均可以重构不同类型的基图像,且效果相近,证明了所获得的基图像具有内容无关性,重构效果随着基图像个数的增加而变好,而与基图像由哪类图像产生无关。由于自然图像结构较人脸图像复杂,相同数量的基图像重构自然图像的效果要差于人脸图像。对于人脸图像,当  $k=200$  时重构效果已经非常接近原图像,重构质量高,同时将  $256^2$  维的原图像降维到了 12800,极大地降低了图像数据维数。

表 1 重构图像的 PSNR 值

Table 1 PSNR values of reconstructed images

原图像类型	基图像来源	不同 $k$ 值时重构图像 PSNR 值/dB		
		$k=10$	$k=50$	$k=200$
自然图像	自然图像所得基图像	17.1776	19.5116	23.1088
	人脸图像所得基图像	17.1269	19.4955	22.7186
人脸图像	自然图像所得基图像	23.8023	31.8224	37.8885
	人脸图像所得基图像	24.1410	32.3080	39.8005

**结束语** 本文提出了一种基图像提取与内容无关图像重构方法,使构成图像的基本信号更易理解,并起到了降维作用。实验表明,通过 PCA 获得的基图像具有内容无关性,可以重构不同类型的图像。所得基图像能够体现图像构成基本要素的特征,类似于特定训练集下深度学习所得的抽象特征,同时还可以降低高维图像数据进而方便后续应用,如图像识别、储存等。本文中,对基图像数量  $k$  的选择依不同应用场景而定,在以后的研究中可以研究选择最佳  $k$  值的方法。

参考文献

[1] BAO M, HUANG C B, WANG L, et al. Visual information processing and its brain mechanism [J]. Science & Technology Review, 2017, 35(19): 15-20.

[2] WANG R, WU X J. Manifold Dimensional Reduction Algorithm Based on Tangent Space Discriminant Learning [J]. Journal of Software, 2018, 29(12): 196-208.

[3] PETSCHARNIG S, LUX M, CHATZICHRISTOFIS S. Dimensionality reduction for image features using deep learning and autoencoders[C]//Proceedings of the 15th International Workshop on Content-Based Multimedia Indexing. ACM, 2017: 23.

[4] MINGRU K, ZHENG Q, YAN S K, et al. Medical image classification algorithm based on principal component feature dimensionality reduction [J]. Future Generation Computer Systems, 2019, 98: 627-634.

[5] JIANG S Q, MIN W, WANG S H. Survey and Prospect of Intelligent Interaction-Oriented Image Recognition Techniques [J]. Journal of Computer Research and Development, 2016, 53(1): 113-122.

[6] MEHTA P, SCHWAB D J. An exact mapping between the variational renormalization group and deep learning [J]. arXiv preprint arXiv: 1410. 3831, 2014.

[7] NIU C, NIU Y G, LI H, et al. Recognition of diseases and insect pests of strawberry based on image gray histogram feature [J]. Jiangsu Agricultural Sciences, 2017, 45(4): 169-172.

[8] LAN R, ZHONG S, LIU Z, et al. A simple texture feature for retrieval of medical images [J]. Multimedia Tools and Applications, 2018, 77(9): 10853-10866.

[9] XIA J, CHANUSSOT J, DU P, et al. (Semi-) Supervised Probabilistic Principal Component Analysis for Hyperspectral Remote Sensing Image Classification [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2014, 7(6): 2224-2236.

[10] QAYYUM H, MAJID M, ANWAR S M, et al. Facial expression recognition using stationary wavelet transform features [J/OL]. Mathematical Problems in Engineering, <https://doi.org/10.1155/2017/9854050>.

[11] FANG Y M, ZHENG H P, FENG H L. Feature extraction and recognition of wood micrograph based on FFT and ICA [J]. Journal of Zhejiang A & F University, 2010, 27(6): 826-830.

[12] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [C]// International Conference on Neural Information Processing Systems, 2012.

[13] HYVÄRINEN A, HURRI J, HOYER P O. Natural Image Statistics [M]// Natural Image Statistics, 2009.

[14] GRAHAM D B, ALLINSON N M. Characterizing virtual eigen-signatures for general purpose face recognition [C]// Face Recognition: From Theory to Applications, NATO ASI Series F, Computer and Systems Sciences, 1998: 446-456.



**LAN Zhang-li**, born in 1973, Ph.D, professor, master supervisor, is a member of China Computer Federation. His main research interests include informatization and intelligentization of transportation, solar energy and image processing.



**SHEN De-xing**, born in 1994, postgraduate. His main research interest include image processing and so on.