

基于 2DPCA 的红外显微图像分析

杨秀坤¹, 钟明亮^{1,2}, 景晓军², 张洁²

(1. 哈尔滨工程大学 信息与通信工程学院, 黑龙江 哈尔滨 150001; 2. 北京邮电大学 信息与通信工程学院, 北京 100876)

摘要: 为了提高光谱分解效率, 提出一种基于光谱方向的 2DPCA 的光谱分解方法, 该方法可以直接利用原始图像矩阵进行光谱分解。通过对兔子动脉和多层油漆中红外显微图像的分解实验, 验证了 2DPCA 光谱分解方法的可行性和有效性。实验表明, 当光谱数大于通道数时, 采用 2DPCA 的光谱分解效率高于 PCA 的光谱分解效率, 并且随着光谱数与通道数比值的增大, 2DPCA 的优势更明显。

关键词: 红外显微图像; PCA; 2DPCA; 光谱分解

中图分类号: O657.33; TP391.41

文献标识码: B

文章编号: 1000-436X(2012)09-0147-05

FTIR microscopic image analysis based on 2DPCA

YANG Xiu-kun¹, ZHONG Ming-liang^{1,2}, JING Xiao-jun², ZHANG Jie²

(1. College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China;

2. School of Information and Communication Engineering, Beijing University of Posts and Telecommunication, Beijing 100876, China)

Abstract: In order to improve the efficiency of spectral unmixing, a novel spectral unmixing method based on 2DPCA in the spectral direction was developed. With the proposed method, spectra can be unmixed directly using the original image matrices. Simulation experiments performed on FTIR microscopic images of a rabbit artery and a multilayer paint chip verify the feasibility and effectiveness of the proposed algorithm. Experimental results demonstrate that when the FTIR microscopic image contains more spectra than spectral channels, the proposed algorithm has higher computation efficiency than PCA algorithm for spectral unmixing. Further more, the advantage of using 2DPCA becomes more obvious with an increase in the ratio of the number of spectra to the number of spectral channels.

Key words: FTIR microscopic image; PCA; 2DPCA; spectral unmixing

1 引言

傅里叶变换红外显微成像技术将红外光谱技术和显微技术相结合, 能够在不破坏样品原始结构的前提下, 探测生物组织、高分子聚合物等样本的化学组分及其分布信息, 具有较高的空间分辨率和光谱分辨率, 是一种快速、直接、绿色的微区分析技术。因而被广泛用于生物医学、材料化学、法庭科学以及食品安全^[1-4]等领域。

由傅里叶变换红外显微成像技术获得的红外显微图像由成千上万的光谱组成, 每一个像素对应一条完整的红外光谱, 每一个波数对应一个浓度分布图。包含了海量物质的成分、浓度、分布等信息。传统的红外显微图像的分析方法主要通过选择物质的某个特征峰, 用峰高或峰面积积分成像方式来获得样品组分的分布信息^[5,6], 该方法仅使用少量的光谱数据, 无法完整有效地获取红外光谱中的大量有用信息。同时, 对于很难找到合适的特征峰的复

收稿日期: 2012-01-01; 修回日期: 2012-06-07

基金项目: 国家自然科学基金资助项目(61143008); 黑龙江省自然科学基金资助项目(ZD200915)

Foundation Items: The National Natural Science Foundation of China (61143008); The Natural Science Foundation of Heilongjiang Province (ZD200915)

杂样品或者未知组分样品，该方法则不适用。因此如何快速从海量光谱数据中得到有用信息给光谱显微图像分析技术带来了新的挑战。

近年来，许多研究一直致力于探索红外显微图像分析技术^[7~10]，其中，PCA 光谱分解方法无需先验知识，可以将数据分解为少量载荷光谱和得分图来描述原始红外显微图像，并提供可视化样品显微结构特征和化学组分信息，因而得到了广泛的应用^[11~13]。PCA 方法需要先将原始二维空间图像变成单一像素方向，才能构造协方差矩阵，增加计算复杂度，针对这一缺点，Yang 等人在 2004 年首次提出了 2DPCA 方法^[14]，该方法对 PCA 方法进行改进，可直接利用二维图像数据构造协方差矩阵。2DPCA 用于特征提取可以降低协方差矩阵的复杂度，提高运算速度，在人脸识别和图像压缩等应用中表现出了更好的性能。本文在此算法的基础上提出了一种基于光谱方向的 2DPCA 光谱分解方法，它利用红外显微图像列（或行）像素对应的二维光谱矩阵来构造广义协方差矩阵，使构造的协方差矩阵包含不同通道分布图之间的相关性，实现光谱数据的快速分解。

2 红外显微图像的分析

红外显微图像能提供被测试样品的化学组分及其空间分布信息，通常可以用数据立方体表示，其中二维是空间方向（像素位置 x, y ），一维是光谱方向（ ν ）。图像中任意一个像素的光谱浓度值由各个组分纯光谱浓度的加权和得到，每个像素的浓度值因所含成分不同而各不相同，但构成整个光谱图像组分的纯光谱都一样，因此可以通过光谱分解来获得与红外显微图像中化学组分相关的光谱及其分布图。

2.1 PCA 光谱分解

PCA 光谱分解可以去除光谱间的相关性，可以用少量的载荷光谱以及对应的得分图来描述原始红外显微图像丰富的光谱信息^[11]。获得的载荷光谱具有与纯光谱相近的特征峰位和峰形，可以看作是抽象的纯光谱；对应的得分图可以用于正确地代表与载荷光谱最相关的物质在红外显微图像中的分布趋势，可以看作抽象的浓度分布图。

PCA 的化学分析模型如下式所示：

$$X = TP^T + E \quad (1)$$

其中，矩阵 X 为红外显微图像，矩阵 T 为主分量矩阵构成的得分矩阵， P^T 为投影特征矩阵构成载荷矩

阵， E 为实验误差。载荷矩阵的行向量为载荷光谱，得分矩阵的列向量可以重新组合为得分图。

2.2 2DPCA 光谱分解

定义矩阵 $B_{m \times n \times k}$ 为红外显微图像，包含有 k 个通道，矩阵 I_j 为第 j 个通道的分布图，大小为 $m \times n$ 。Yang 等人提出的 2DPCA 利用二维空间构造图像协方差矩阵 G_i

$$G_i = \frac{1}{k} \sum_{j=1}^k (I_j - \bar{I})^T (I_j - \bar{I}) \quad (2)$$

其中， $\bar{I} = \frac{1}{k} \sum_{j=1}^k I_j$ 为分布图均值， G_i 大小为 $n \times n$ ，由 G_i 特征分解得到前 q 个最大特征值对应的特征向量构成的主分量矩阵 $U_{n \times q}$ ，将矩阵 $B_{m \times n \times k}$ 向 $U_{n \times q}$ 投影后可以得到投影特征矩阵 $V_{m \times q \times k}$ 。基于空间方向的 2DPCA，能够去除同列像素之间的相关性，在一定程度上去除噪音和冗余，但其主分量矩阵与投影特征矩阵不能直接与光谱或浓度分布图相关，因此，不适合用于红外显微图像的光谱分解。

在 2DPCA 的基础上，提出了基于光谱方向的 2DPCA 光谱分解方法，它用列（或行）像素对应的二维光谱矩阵代替只包含空间方向的图像数据来构造协方差矩阵，使构造的协方差矩阵包含不同通道间的相关性。

定义列光谱 A_i 为红外显微图像的第 i 列像素对应二维光谱数据（如图 1 所示），其大小 $m \times k$ ，用它代替分布图 I_j 构造协方差矩阵 C_i

$$C_i = \frac{1}{n} \sum_{i=1}^n (A_i - \bar{A})^T (A_i - \bar{A}) \quad (3)$$

其中， $\bar{A} = \frac{1}{n} \sum_{i=1}^n A_i$ 为列光谱的均值， C_i 大小为 $k \times k$ ，对 C_i 进行特征分解，选择前 l 个最大特征值对应的特征向量 w_1, w_2, \dots, w_l ，构成的主分量矩阵 $W_{k \times l}$ 。将矩阵 $B_{m \times n \times k}$ 向 $W_{k \times l}$ 进行投影，得到投影特征矩阵 $Z_{m \times l \times n}$ 。

因此，2DPCA 的化学分析模型为

$$B = ZW^T + E \quad (4)$$

如图 1 所示，基于光谱方向的 2DPCA 可以将红外显微图像分解为得分图和载荷图，得分图由投影特征矩阵 Z 的二维空间矩阵构成，可以用来表示物质的分布情况；载荷图由主分量矩阵 W^T 的行向量构成，它与得分图代表物质的光谱具有相近的特

征峰峰位和峰形。因此，基于光谱方向的 2DPCA 可以去掉红外显微图像不同通道间的相关性，实现光谱数据的快速分解。

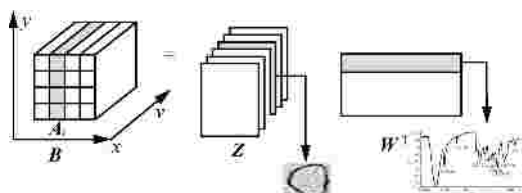


图 1 2DPCA 的化学分析模型

3 实验结果与分析

3.1 实验数据及实验平台

本文实验数据是由美国 Perkin Elmer 公司生产的傅里叶变换红外/近红外成像系统 Spotlight 400 提供的兔子动脉血管以及多层油漆片的中红外显微图像。数据处理与仿真主要基于 Matlab7.11，所有程序均在 Celeron (R) Dual-Core CPU T3 000 1.80GHz、250GB 硬盘和 1.98GB 内存的某品牌笔记本电脑上运行。

3.2 目标化合物的分析

胆固醇在动脉粥样硬化的发生与发展过程中起着重要作用。通过对动脉血管中胆固醇分布的研究，可以为医学诊断提供参考。

实验数据：兔子动脉中红外显微图像，成像面积为 $3\ 600\mu\text{m} \times 5\ 200\mu\text{m}$ ，空间分辨率 $25\mu\text{m} \times 25\mu\text{m}$ ，共有 29952 像素。光谱分辨率 4cm^{-1} ，波数范围 $4\ 000 \sim 720\text{cm}^{-1}$ ，共有 821 个通道。

实验方法：PCA 与 2DPCA 光谱分解方法。

采用 2 种光谱分解方法对兔子动脉中红外光谱进行分析。取前 8 帧得分图如图 2 和图 3 所示。可以看出图 2(f)和图 2(g)以及图 3(e)和图 3(f)均显示在动脉壁有斑块物质分布。为了确定该物质是否为胆固醇，将图 2(f)和图 2(g)以及图 3(e)和图 3(f)对应的载荷光谱与胆固醇标准谱图进行比较，结果如图 4 所示。

图 4 中的 c 为胆固醇标准谱，其中 $2\ 950 \sim 2\ 850\text{cm}^{-1}$ 区域为脂肪族甲基亚甲基的伸缩振动； $1\ 460\text{cm}^{-1}$ 处的吸收带为二甲脂 C-H 弯曲， $1\ 377\text{cm}^{-1}$ 处为 CH_3 的吸收带； $1\ 057\text{cm}^{-1}$ 为 C-O 的伸缩振动^[15]。

图 4 中的 a 和 b 为 PCA 光谱分解得到与图 2(f)和图 2(g)得分图对应的载荷光谱；图 4 中的 d 和 e 为 2DPCA 光谱分解得到与图 3(e)和图 3(f)得分图对应的载荷光谱。通过比较，可以看出载荷图 4 中的

a、b、d 和 e 具有吸收峰与胆固醇的特征峰峰位基本相同、峰形大体一致，因此可以判断分布于动脉壁上的斑块物质为胆固醇。

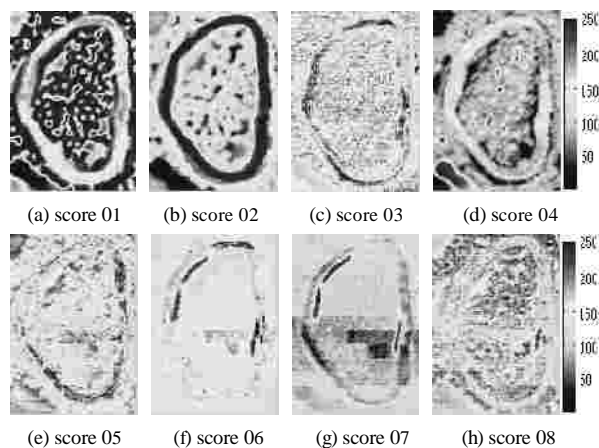


图 2 PCA 光谱分解的得分

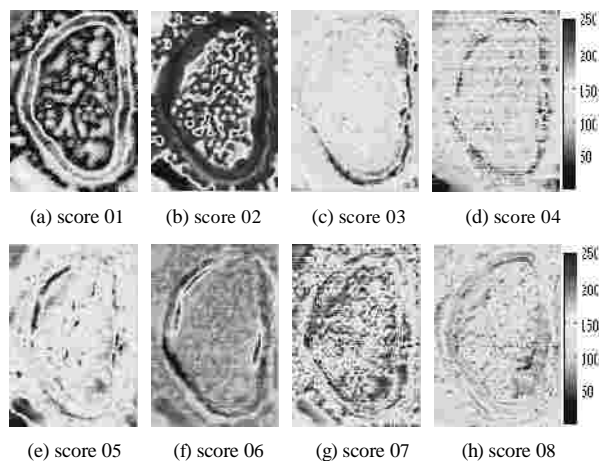


图 3 2DPCA 光谱分解的得分

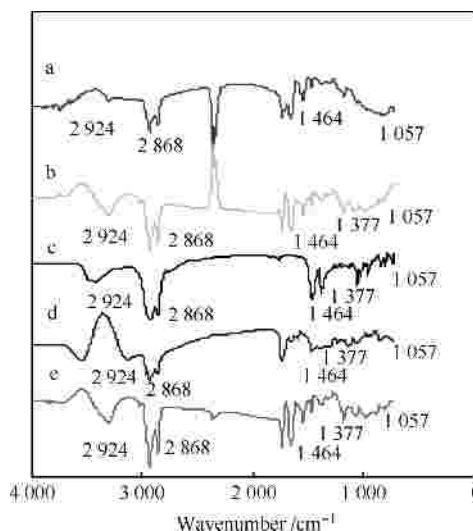


图 4 载荷光谱与胆固醇标准谱图比较

3.3 异常物质的分析

在法庭科学领域，常常需要分析少量的多层油漆样品，对于样品中异常成分的检测和分析，可以为侦查工作提供重要线索。红外显微成像技术具有较高的空间分辨率和光谱分辨率，能够提供样品的化学组分及微观结构，结合光谱分解方法可以用于异常物质的检测和分析。

实验数据：多层油漆样品中红外显微图像成像面积为 $468.7\mu\text{m}\times 250\mu\text{m}$ ，空间分辨率 $6.25\mu\text{m}\times 6.25\mu\text{m}$ ，共有 3000 个像素。光谱分辨率 8cm^{-1} ，波数范围 $4000\sim 752\text{cm}^{-1}$ ，共有 407 个通道。

实验方法：PCA 与 2DPCA 光谱分解方法。

采用 2 种光谱分解方法对多层油漆样品中红外显微图像进行分析。取前 8 帧得分图如图 5 和图 6 所示。油漆片图 5(a)与图 6(a)显示为 2 层，但在 5(f)与图 6(f)显示为 6 层，同时这 2 个得分图显示的左上角异常物质分布。通过对有异常物质分布的得分图对应载荷光谱峰位峰形的分析，可以用来确定该物质的化学组分。图 7 中的 a 和 b 为图 5(f)与图 6(f)对应的载荷光谱，图 7 中的 c 为指纹的标准谱图。

图 7 中的 c，吸收峰 3280cm^{-1} 为 N-H 的伸缩振动，吸收峰 2920cm^{-1} 为 C-H 的伸缩振动， 1655cm^{-1} 为羰基 C=O 伸缩振动峰，吸收峰 1545cm^{-1} 为 N-H 平面弯曲^[16]，这与载荷光谱图 7 中 a 和 b 的特征峰峰位基本相同、峰形大体一致，由此可以判断该异常物质为指纹。

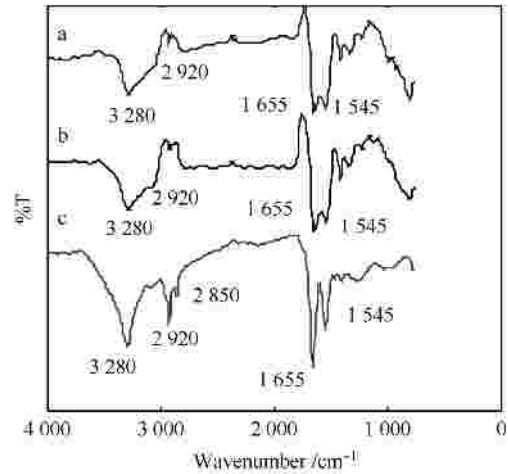


图 7 载荷光谱与指纹标准谱图比较

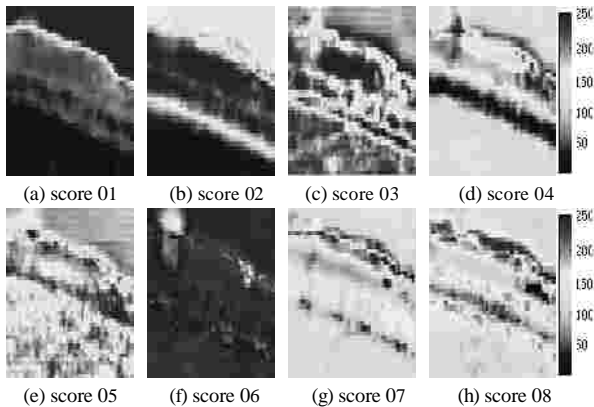


图 5 PCA 光谱分解的得分

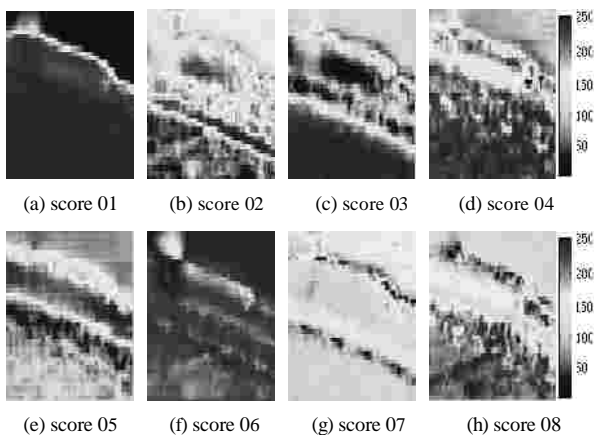


图 6 2DPCA 光谱分解的得分

3.4 算法性能比较

从以上 2 组实验可以看出 2DPCA 光谱分解与 PCA 光谱分解得到的得分图和载荷结果相似，均可以用于已知化合物分布信息的提取和异常化合物的检测。为了比较 2 组算法的性能优劣，对 2 组实验中的 2 种算法提取不同特征数的运算时间进行比较，实验结果如表 1 所示。

表 1 PCA 算法与 2DPCA 算法性能比较

运算 时间/s	实验 1		实验 2	
	PCA	2DPCA	PCA	2DPCA
1	231.130	15.016	7.156	1.907
2	232.922	15.328	7.236	2.204
3	233.594	15.734	7.282	2.328
4	234.140	15.969	7.390	2.500
5	234.938	16.106	7.469	2.609
6	236.250	16.640	7.735	2.765
7	236.750	16.890	8.094	2.938
8	237.735	17.140	8.141	3.047

由表 1 数据可以看出 2 组实验中 2DPCA 的运算时间均小于 PCA 的运算时间，其中实验 1 中 2DPCA 的运算时间远小于 PCA 的运算时间，而实验 2 中 2DPCA 的运算时间略小于 PCA 的运算时间。

这是因为 PCA 的协方差矩阵包含的是任意 2 个光谱之间的相关信息, 算法复杂度随光谱数增加而增大 2DPCA 的协方差矩阵中包含的是不同通道分布图之间的相关性, 算法的复杂度随通道数的增加而增大。所以, 当光谱数大于通道数时, 2DPCA 光谱分解的效率高于 PCA 光谱分解的效率, 并且随着光谱数与通道数比值增大, 采用 2DPCA 进行光谱分解的优势就更明显。

4 结束语

本文提出一种基于光谱方向的 2DPCA 光谱分解方法, 能够直接利用原始图像矩阵构造协方差矩阵, 将红外显微图像分解为与化学组分相关载荷光谱及得分图, 实现目标化合物的分布信息提取以及未知异常成分检测。与 PCA 光谱分解相比, 2DPCA 光谱分解更简单、直接, 对于光谱数大于通道数的红外显微图像, 采用 2DPCA 进行光谱分解的复杂度更小、运算效率更高, 并且随着光谱数与通道数比值增大, 2DPCA 的优势更明显。

参考文献:

- [1] CAINE S, HERAUD P, TOBIN M J, *et al.* The application of Fourier transform infrared micro spectroscopy for the study of diseased central nervous system tissue[J]. *NeuroImage*, 2012, 59(4):3624-3640.
- [2] HONG Z H, CONG Y H, QI Z M, *et al.* Studying deformation behavior of a single spherulite with in-situ infrared micro spectroscopic imaging[J]. *Polymer*, 2012, 53(2): 640-647.
- [3] DUBAND S, GOVIN A, DUMOLLARD J M, *et al.* Laryngeal teflonoma identified by Fourier- transform infrared micro spectroscopy after forensic autopsy: an interesting tool for foreign rial identification in forensic cases[J]. *Forensic Science International*, 2012, 214(1-3): 26-29.
- [4] ROHMAN A, SISMININDARI, ERWANTO Y, *et al.* Analysis of pork adulteration in beef meatball using Fourier transforms infrared (FTIR) spectroscopy[J]. *MeatScience*, 2011, 88(1): 91-95.
- [5] YANG T T, WENG S F, ZHENG N, *et al.* Histopathology mapping of biochemical changes in myocardial infarction by Fourier transform infrared spectral imaging[J]. *Forensic Science International*, 2011, 207(1-3): 34-39.
- [6] RAMKRISHNAN N, XIA Y, BIDTHANAPALLY A. Polarized IR microscopic imaging of articular cartilage[J]. *Phys Med Biol*, 2007, 52(15): 4601-4614.
- [7] SALZER R, SIESLER H W. *Infrared and Raman Spectroscopic Imaging*[M]. Wiley-VCH: Weinheim, Germany, 2009.
- [8] WANG D, MA Z H, PAN L G, *et al.* The application of 2nd derivative in NIR and MIR micro-imaging of two plastic materials[J]. *Advanced Materials Research*, 2011, (295-297): 2607-2614.
- [9] CARLOS C, MARETTO D A, POPPI R J, *et al.* Fourier transform infrared microspectroscopy as a bacterial source tracking tool to discriminate fecal E coli strains[J]. *Microchemical Journal*, 2011, 99(1): 15-19.
- [10] CAIRÓS C, AMIGO J M, WATT R, *et al.* Implementation of enhanced correlation maps in near infrared chemical images: application in pharmaceutical research[J]. *Talanta*, 2009, 79(3): 657-664.
- [11] KEENAN M R. Exploiting spatial-domain simplicity in spectral image analysis[J]. *Surface and Interface Analysis*, 2009, 41(2):79-87.
- [12] XUE J H, LEE C, WAKEHAM S G, *et al.* Using principal components analysis (PCA) with cluster analysis to study the organic geochemistry of sinking particles in the ocean[J]. *Organic Geochemistry*, 2011, 42(4): 356-367.
- [13] JUAN A D, MAEDER M, HANCEWICZ T, *et al.* Use of local rank-based spatial information for resolution of spectroscopic images[J]. *Journal of Chemometrics*, 2008, 22(5):291-298.
- [14] YANG J, ZHANG D, FRANGI A F, *et al.* Two-dimensional PCA: a new approach to appearance-based face representation and recognition[J]. *IEEE Trans PAMI*, 2004, 26(1): 131-137.
- [15] HOLMAN H Y, BJORNSTAD K A, MARTIN M C, *et al.* Mid-infrared reflectivity of experimental atheromas[J]. *Journal of Biomedical Optics*, 2008, 13(3):1-3.
- [16] RICCI C, PHIRIYAVITYOPAS P, CURUM N, *et al.* Chemical imaging of latent fingerprint residues[J]. *Applied Spectroscopy*, 2007, 61(5): 514-522.

作者简介:



杨秀坤(1971-),女,黑龙江鸡西人,哈尔滨工程大学教授、博士生导师,主要研究方向为图像处理、模式识别、多光谱检测和机器视觉等。

钟明亮(1986-),女,畲族,福建宁德人,哈尔滨工程大学博士生,主要研究方向为生物特征识别技术、红外显微图像处理等。

景晓军(1965-),男,北京人,北京邮电大学教授、博士生导师,主要研究方向为信息融合、模式识别和图像处理等。

张洁(1987-),女,天津人,北京邮电大学博士生,主要研究方向为模式识别、图像处理和信息融合。