软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2017,28(2):292−309 [doi: 10.13328/j.cnki.jos.005047]

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

基于弱匹配概率典型相关性分析的图像自动标注*

张 博[1], 郝 杰[4], 马 刚[2,3], 史忠植[2]

[1](中国矿业大学 计算机科学与技术学院,江苏 徐州 221116)

[2](中国科学院 计算技术研究所 智能信息处理重点实验室,北京 100190)

[3](中国科学院大学,北京 100049)

[4](徐州医科大学 医学信息学院,江苏 徐州 221004)

通信作者: 郝杰, E-mail: haojie@xzmc.edu.cn

摘 要: 针对弱匹配多模态数据的相关性建模问题,提出了一种弱匹配概率典型相关性分析模型(semi-paired probabilistic CCA,简称 SemiPCCA).SemiPCCA 模型关注于各模态内部的全局结构,模型参数的估计受到了未匹配样本的影响,而未匹配样本则揭示了各模态样本空间的全局结构.在人工弱匹配多模态数据集上的实验结果表明,SemiPCCA 可以有效地解决传统 CCA(canonical correlation analysis)和 PCCA(probabilistic CCA)在匹配样本不足的情况下出现的过拟合问题,取得了较好的效果.提出了一种基于 SemiPCCA 的图像自动标注方法.该方法基于关联建模的思想,同时使用标注图像及其关键词和未标注图像学习视觉模态和文本模态之间的关联,从而能够更准确地对未知图像进行标注.

关键词: 典型相关性分析;概率典型相关性分析;弱匹配典型相关性分析;图像自动标注

中图法分类号: TP391

中文引用格式: 张博,郝杰,马刚,史忠植.基于弱匹配概率典型相关性分析的图像自动标注.软件学报,2017,28(2):292−309. http://www.jos.org.cn/1000-9825/5047.htm

英文引用格式: Zhang B, Hao J, Ma G, Shi ZZ. Automatic image annotation based on semi-paired probabilistic canonical correlation analysis. Ruan Jian Xue Bao/Journal of Software, 2017,28(2):292−309 (in Chinese). http://www.jos.org.cn/1000-9825/5047.htm

Automatic Image Annotation Based on Semi-Paired Probabilistic Canonical Correlation Analysis

ZHANG Bo[1], HAO Jie[4], MA Gang[2,3], SHI Zhong-Zhi[2]

[1](School of Computer Science and Technology, China University of Mining and Technology, Xuzhou 221116, China)

[2](Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, The Chinese Academy of Sciences, Beijing 100190, China)

[3](University of Chinese Academy of Sciences, Beijing 100049, China)

[4](School of Medicine Information, Xuzhou Medical University, Xuzhou 221004, China)

Abstract: Canonical correlation analysis (CCA) is a statistical analysis tool for analyzing the correlation between two sets of random variables. CCA requires the data be rigorously paired or one-to-one correspondence among different views due to its correlation definition. However, such requirement is usually not satisfied in real-world applications due to various reasons. Often, only a few paired and a lot of

* 基金项目: 国家重点基础研究发展计划(973)(2013CB329502); 国家自然科学基金(61035003); 国家高技术研究发展计划(863)(2012AA011003); 国家科技支撑计划(2012BA107B02); 江苏省自然科学基金(BK20160276)

Foundation item: National Program on Key Basic Research Project of China (973) (2013CB329502); National Natural Science Foundation of China (61035003); National High-Tech R&D Program of China (863) (2012AA011003); National Key Technology R&D Program of China (2012BA107B02); Natural Science Foundation of Jiangsu Province (BK20160276)

收稿时间: 2014-12-18; 修改时间: 2015-06-11, 2015-09-10; 采用时间: 2016-02-03

unpaired multi-view data are given, because unpaired multi-view data are relatively easier to be collected and pairing them is difficult, time consuming and even expensive. Such data is referred as semi-paired multi-view data. When facing semi-paired multi-view data, CCA usually performs poorly. To tackle this problem, a semi-paired variant of CCA, named SemiPCCA, is proposed based on the probabilistic model for CCA. The actual meaning of "semi-" in SemiPCCA is "semi-paired" rather than "semi-supervised" as in popular semi-supervised learning literature. The estimation of SemiPCCA model parameters is affected by the unpaired multi-view data which reveal the global structure within each modality. By using artificially generated semi-paired multi-view data sets, the experiment shows that SemiPCCA effectively overcome the over-fitting problem of traditional CCA and PCCA (probabilistic CCA) under the condition of insufficient paired multi-view data and performs better than the original CCA and PCCA. In addition, an automatic image annotation method based on the SemiPCCA is presented. Through estimating the relevance between images and words by using the labelled and unlabeled images together, this method is shown to be more accurate than previous published methods.

Key words: canonical correlation analysis; probabilistic canonical correlation analysis; semi-paired canonical correlation analysis; automatic image annotation

物联网、互联网等拥有丰富的文本、图像、视频和音频等多媒体信息资源,这些信息资源是异构的,很难直接发现它们之间的关联.目前,典型相关性分析(canonical correlation analysis,简称 CCA)作为一种分析两组随机变量之间相关性的统计分析工具,已被引入跨媒体的相关性建模中,挖掘不同模态内容特征之间潜在的统计相关性[1,2].通过特征子空间映射,将各模态的数据从原始高维特征空间映射到低维特征空间,既解决了不同类型数据间的异构性问题,消除了多模态数据间的内容鸿沟,最大程度地保持了初始的相关性不变,将不同类型的多媒体数据在特征层面上关联起来,同时也最大程度地保持初始的相关性不变.

典型相关性分析中两组相关的随机变量可以来自多种信息来源(如同一个人的声音和图像),也可以是从同一来源的信息中抽取的不同特征(如图像的颜色特征和纹理特征),但训练数据必须一对一严格匹配.很多原因造成这种严格匹配的训练数据难以获得,如:(1) 多传感器采集系统中传感器采样频率不同步或传感器故障,会造成不同通道采集来的数据不同步或丢失某一通道数据;(2) 单模态数据比较容易获得,但人工匹配却非常费时、费力.实际中,我们面对的多模态数据经常是只有少量一对一严格匹配,其余大量数据未匹配.我们称其为弱匹配多模态数据.

面向弱匹配多模态数据的典型相关性分析有两种基本方法:(1) 丢弃未匹配数据,只使用典型相关性分析处理严格匹配的多模态数据;(2) 根据特定准则,匹配多模态数据.但这两种方法都不可能获得理想的结果.

本文的主要工作包括:(1) 提出了一种全新的弱匹配概率典型相关性分析模型(semi-paired probabilistic CCA,简称 SemiPCCA).不同于以往的弱匹配典型相关性分析模型,SemiPCCA 完全基于概率典型相关性分析模型(probabilistic CCA,简称 PCCA),关注于各模态内部的全局结构,模型参数的估计受到了未匹配样本的影响,而未匹配样本则揭示了各领域样本空间的全局结构.(2) 提出了一种基于 SemiPCCA 的图像自动标注方法.该方法同时使用标注图像及其关键词和未标注图像估计隐空间的分布,学习视觉模态和文本模态之间的关联,能够较好地对未知图像进行标注.

1 相关工作

1.1 典型相关性分析

传统的特征分析方法,如 PCA(principal component analysis),ICA(independent component analysis)和 PLS(partial least squares),大多用于单模态的特征分析,实现主成分提取、去噪、维数约减和保持本征度量等目的,不能同时分析不同类型的异构特征,难以发现多种特征间的关联信息.典型相关性分析(canonical correlation analysis,CCA)是一种用来分析两组随机变量之间相关性的统计分析工具,其相关性保持特征已经在理论上得到证明,应用于经济学、气象和基因组数据分析等领域.CCA 通过统计方法找到两组异构多模态特征之间的潜在关系,从底层特征上用统一的模型将不同类型的多模态数据关联起来,同时尽可能地发现和保持数据间潜在的相关性.

维度分别为 p 和 q 的两组随机变量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$,给定均值为 0 的成对观察样本集合 $\{(\boldsymbol{x}_i,\boldsymbol{y}_i)\}_{i=1}^{n}\in R^p\times R^q$,即

$\bar{\boldsymbol{x}}=\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{x}_i=0,\bar{\boldsymbol{y}}=\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{y}_i=0$,其中,$\{\boldsymbol{x}_i\}_{i=1}^{n}$ 和 $\{\boldsymbol{y}_i\}_{i=1}^{n}$ 是多种信息来源(如爆炸的声音和图像),也可以是从同一来源的信息中抽取的不同特征(如图像的颜色特征和纹理特征).记 $\boldsymbol{X}\in R^{p\times n},\boldsymbol{Y}\in R^{q\times n}$,$n$ 表示样本数量.CCA 的目标是寻找两组投影向量 $\boldsymbol{a}_x\in R^p$ 和 $\boldsymbol{a}_y\in R^q$,使线性组合 $u=\boldsymbol{a}_x^T\boldsymbol{x}$ 和 $v=\boldsymbol{a}_y^T\boldsymbol{y}$ 之间的相关系数达到最大,即求解以下相关系数的最大值问题:

$$\rho=\max_{\boldsymbol{a}_x,\boldsymbol{a}_y}\frac{\boldsymbol{a}_x^T\boldsymbol{C}_{xy}\boldsymbol{a}_y}{\sqrt{\boldsymbol{a}_x^T\boldsymbol{C}_{xx}\boldsymbol{a}_x\boldsymbol{a}_y^T\boldsymbol{C}_{yy}\boldsymbol{a}_y}},$$

其中,$\boldsymbol{C}_{xx}=\boldsymbol{X}\boldsymbol{X}^T\in R^p\times R^p$ 和 $\boldsymbol{C}_{yy}=\boldsymbol{Y}\boldsymbol{Y}^T\in R^q\times R^q$ 表示集合内协方差矩阵(within-set covariance matrix);$\boldsymbol{C}_{xy}=\boldsymbol{X}\boldsymbol{Y}^T\in R^p\times R^q$ 表示集合间协方差矩阵(between-set covariance matrix),且 $\boldsymbol{C}_{yx}=\boldsymbol{C}_{xy}{}^T$.

常将 CCA 问题等价地描述为以下特征值问题:

$$\begin{pmatrix}&\boldsymbol{C}_{xy}\\\boldsymbol{C}_{yx}&\end{pmatrix}\begin{pmatrix}\boldsymbol{a}_x\\\boldsymbol{a}_y\end{pmatrix}=\lambda\begin{pmatrix}\boldsymbol{C}_{xx}&\\&\boldsymbol{C}_{yy}\end{pmatrix}\begin{pmatrix}\boldsymbol{a}_x\\\boldsymbol{a}_y\end{pmatrix}.$$

CCA 是一种线性数学模型,这种线性模型不足以揭示真实世界中大量存在的非线性相关现象.当用这样的线性模型来学习非线性相关现象时,将不可避免地出现欠拟合(underfitting)现象.解决这种问题目前主要有 3 种途径[3]:核方法、神经网络和局部化方法.将神经网络用于 CCA 计算是近年来非线性 CCA 的一个重要进展,通过神经网络的非线性特征揭示数据之间存在的非线性相关关系.2013 年,Andrew 等人结合深度学习,提出了 Deep Canonical Correlation Analysis 算法[4],在处理非线性相关问题时获得了优于 KCCA 的整体相关度.

CCA 对样本的类信息未予以充分利用.2008 年,孙廷凯等人引入样本的类信息,并充分考虑了同类样本之间的相关与不同类样本之间的相关关系及其对分类的影响,提出了一种新的有监督学习方法——判别型 CCA (discriminative CCA,简称 DCCA)[5],并运用核技巧,将线性的 DCCA 推广到高维特征空间,提出了核化的 DCCA (kernelized DCCA,简称 KDCCA),用来增强对线性不可分问题的分类能力.DCCA 提取的特征能够实现同类样本特征之间相关最大化,同时使得不同类样本特征之间相关最小化,这将有利于模式的分类.2011 年,Shin 等人证明了 *DCCA*(*X*,*Y*)等价于 *LDA*(*X*,*C*)+*LDA*(*Y*,*C*),并改进了 DCCA 算法,使用 *K* 近邻计算类内散布矩阵[6].类似的方法还有:2011 年 Kursun 等人提出了 WCCA(within class coupling CCA)[7];孙权森等人提出的广义典型相关分析(generalized CCA,简称 GCCA)将最小化类内散布矩阵作为目标函数之一,降低了特征的类内离散度,提高了特征表示的鉴别能力.2012 年,周旭东等人提出了增强组合特征判别性的典型相关分析(CECCA)[8].CECCA 是一种监督型降维方法,在 CCA 基础上,通过结合组合特征的判别分析,实现对组合特征相关性与判别性的联合优化,使所抽取特征更适合分类.

半监督学习是近年来机器学习领域的一个研究热点.在很多实际应用中,获取大量的无标号样本已变得非常容易,而获取有标号样本通常需要付出很大的代价.2008 年,彭岩等人在 CCA 的应用中加入了监督信息,提出一种半监督典型相关分析(semi-CCA)算法[9].该方法中利用的监督信息为样本间的成对约束信息,即已知两个样本属于同一类(称为正约束(must-link))或者不属于同一类(称为负约束(cannot-link)).在许多实际应用中,成对约束信息比类标号更容易获得,也更加实际.另外,样本之间的成对约束可以从类别标号中直接获得,反之则不可以.与 Semi-CCA 算法类似,2010 年 Hou 等人提出了 MVSSDR 算法.2010 年,Kursun 等人提出了 Semi-supervised CCA(SCCA)[10].2012 年,Chen 等人提出了统一的半匹配半监督多视图数据降维框架 S2GCA(semi-paired and semi-supervised generalized correlation analysis)[11].

1.2 概率典型相关性分析

2005 年,Bach 等人给出了 CCA 的概率解释[12],并提出了概率典型相关性分析(probabilistic CCA,简称 PCCA).PCCA 是一种线性高斯模型(linear Gaussian model),可以看作是因子分析(factor analysis,简称 FA)的一个特例,图模型如图 1 所示.

设 $\boldsymbol{X}_1=\{\boldsymbol{x}_{1n}\}_{n=1}^{N}\in R^{m_1\times N}$ 表示 m_1 维随机变量 $\boldsymbol{x}_1$ 的观察样本集合,$\boldsymbol{X}_2=\{\boldsymbol{x}_{2n}\}_{n=1}^{N}\in R^{m_2\times N}$ 表示 m_2 维随机变量 $\boldsymbol{x}_2$

的观察样本集合,N 表示样本数量,$\boldsymbol{z}$ 表示与随机变量 $\boldsymbol{x}_1,\boldsymbol{x}_2$ 相关的 d 维隐藏变量,$\boldsymbol{z}$ 的每个元素均服从独立标准正态分布.类似于因子分析,可以定义以下线性高斯模型(linear Gaussian model),即随机变量 $\boldsymbol{x}_1,\boldsymbol{x}_2$ 可以由 d 维隐藏变量 $\boldsymbol{z}$ 经过线性变换并附加一个高斯噪声生成.

$$\begin{cases}\boldsymbol{z}\sim N(0,\boldsymbol{I}_d),\ \min(m_1,m_2)\geqslant d\geqslant 1\\ \boldsymbol{x}_1=\boldsymbol{W}_1\boldsymbol{z}+\boldsymbol{\mu}_1+\boldsymbol{\varepsilon}_1,\ \boldsymbol{W}_1\in R^{m_1\times d},\ \boldsymbol{\varepsilon}_1\sim N(0,\boldsymbol{\psi}_1)\\ \boldsymbol{x}_2=\boldsymbol{W}_2\boldsymbol{z}+\boldsymbol{\mu}_2+\boldsymbol{\varepsilon}_2,\ \boldsymbol{W}_2\in R^{m_2\times d},\ \boldsymbol{\varepsilon}_2\sim N(0,\boldsymbol{\psi}_2)\end{cases} \tag{1}$$

其中,$\boldsymbol{W}_1$ 和 $\boldsymbol{W}_2$ 表示线性变换矩阵,$\boldsymbol{\varepsilon}_1$ 和 $\boldsymbol{\varepsilon}_2$ 表示高斯噪声.

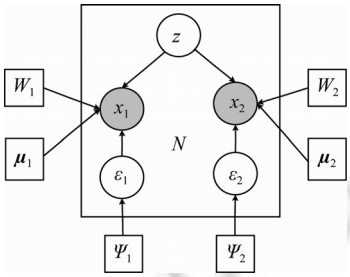

Fig.1 Graphical model for PCCA

图 1 概率典型相关性分析图模型

Bach 等人证明了存在使其似然函数最大化的参数 $\boldsymbol{W}_1,\boldsymbol{W}_2,\boldsymbol{\mu}_1,\boldsymbol{\mu}_2,\boldsymbol{\psi}_1,\boldsymbol{\psi}_2$ 解析解,即[12]

$$\hat{\boldsymbol{\mu}}_1=\tilde{\boldsymbol{\mu}}_1,\ \hat{\boldsymbol{W}}_1=\tilde{\boldsymbol{\Sigma}}_{11}\boldsymbol{U}_{1d}\boldsymbol{M}_1,\ \hat{\boldsymbol{\psi}}_1=\tilde{\boldsymbol{\Sigma}}_{11}-\hat{\boldsymbol{W}}_1\hat{\boldsymbol{W}}_1^T,\ \hat{\boldsymbol{\mu}}_2=\tilde{\boldsymbol{\mu}}_2,\ \hat{\boldsymbol{W}}_2=\tilde{\boldsymbol{\Sigma}}_{22}\boldsymbol{U}_{2d}\boldsymbol{M}_2,\ \hat{\boldsymbol{\psi}}_2=\tilde{\boldsymbol{\Sigma}}_{22}-\hat{\boldsymbol{W}}_2\hat{\boldsymbol{W}}_2^T \tag{2}$$

其中,$\tilde{\boldsymbol{\Sigma}}_{11}$,$\tilde{\boldsymbol{\Sigma}}_{22}$,$\tilde{\boldsymbol{\mu}}_1$ 和 $\tilde{\boldsymbol{\mu}}_2$ 分别表示随机变量 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$ 观察样本集合的协方差和均值,$\boldsymbol{U}_{1d}\in R^{m_1\times d},\boldsymbol{U}_{2d}\in R^{m_2\times d}$ 为观察样本集合的 d 组典型相关特征向量,$\boldsymbol{P}_d$ 为相应特征值 $\lambda_1,\lambda_2,\ldots,\lambda_d$ 组成的对角矩阵,$\boldsymbol{M}_1,\boldsymbol{M}_2$ 为任意 $d\times d$ 矩阵,且 $\boldsymbol{M}_1\boldsymbol{M}_2^T=\boldsymbol{P}_d$.$\boldsymbol{U}_{1d},\boldsymbol{U}_{2d}$ 和 $\boldsymbol{P}_d$ 对应传统 CCA 方法的结果.

降维是 CCA 的一种主要应用.PCCA 给出了随机变量 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$ 从数据空间降维到隐空间的概率解释,即后验概率 $P(\boldsymbol{z}|\boldsymbol{x}_1)\sim N\left(\boldsymbol{M}_1^T\boldsymbol{U}_{1d}^T(\boldsymbol{x}_1-\hat{\boldsymbol{\mu}}_1),\boldsymbol{I}-\boldsymbol{M}_1\boldsymbol{M}_1^T\right)$ 和 $P(\boldsymbol{z}|\boldsymbol{x}_2)\sim N\left(\boldsymbol{M}_2^T\boldsymbol{U}_{2d}^T(\boldsymbol{x}_2-\hat{\boldsymbol{\mu}}_2),\boldsymbol{I}-\boldsymbol{M}_2\boldsymbol{M}_2^T\right)$.

为了便于降维后数据的可视化,使用 $E(\boldsymbol{z}|\boldsymbol{x}_1)$,$E(\boldsymbol{z}|\boldsymbol{x}_2)$ 代替 $P(\boldsymbol{z}|\boldsymbol{x}_1)$,$P(\boldsymbol{z}|\boldsymbol{x}_2)$ 表示随机变量 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$ 从数据空间降维到隐空间后的结果,如图 2 所示.$E(\boldsymbol{z}|\boldsymbol{x}_1)$ 和 $E(\boldsymbol{z}|\boldsymbol{x}_2)$ 分别构成了样本数据空间到 PCCA 隐空间的典型投影,结果和 CCA 完全一致.

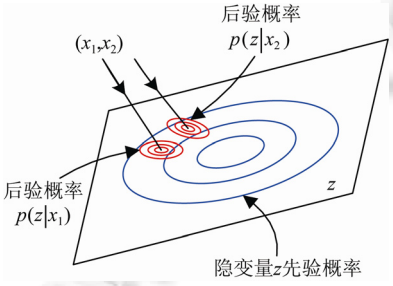

Fig.2 Projection of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ onto the mean of the posterior distribution of $\boldsymbol{z}$ in latent space of PCCA

图 2 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$ 投影到 PCCA 隐空间中 $\boldsymbol{z}$ 的后验概率期望

2006 年,Leen 等人应用非线性高斯过程给出了非线性典型相关性分析的概率模型[13].从概率密度估计的角度推导 CCA 会得到许多重要的优点,其中最为显著的优点是可以混合多个局部 PCCA 概率模型.2015 年,张博等人在 PCCA 的基础上,使用概率混合模型框架提出了混合概率典型相关性分析模型(mixture of probabilistic

CCA,简称MixPCCA)以及估计模型参数的2阶段期望最大化(expectation maximization,简称EM)算法,并给出了使用聚类融合确定局部线性模型数量的方法和MixPCCA模型应用于模式识别的理论框架[14].由于概率混合模型使用多个独立的概率分布,它可以描述一个复杂的数据分布,无论数据分布的结构如何复杂,总可以通过增加成分的方式来描述数据分布的局部特性

2007年,Klami等人提出了Bayesian CCA(BCCA)[15].同年,Wang也将变分贝叶斯方法应用于概率典型相关性分析,该方法不仅实现了模型参数的估计,同时也实现了子空间维度的自动选择[16].2010年,Viinikanoja等人实现了典型相关性分析的混合变分贝叶斯概率模型[17].然而,与因子分析和主成分分析的概率模型相比,PCCA中高斯噪声的完全方差 $\boldsymbol{\psi}_1,\boldsymbol{\psi}_2$ 使得BCCA及其扩展模型难以有效地处理高维度小样本数据,所以早期的BCCA仅仅用于10维以内的样本数据.Archambeau等人[18]以及Klami等人[19]分别于2009年和2010年通过引入额外的隐藏变量 $\boldsymbol{z}_1,\boldsymbol{z}_2$ 来解决BCCA面临的高维样本建模问题,其中,随机变量 $\boldsymbol{x}_1,\boldsymbol{x}_2$ 由隐藏变量 $\boldsymbol{z},\boldsymbol{z}_1,\boldsymbol{z}_2$ 经过线性变换并附加一个高斯噪声生成.

$$\boldsymbol{z}\sim N\left(0,\boldsymbol{I}_d\right),\ \boldsymbol{z}_1\sim N\left(0,\boldsymbol{I}_{d_1}\right),\ \boldsymbol{z}_2\sim N\left(0,\boldsymbol{I}_{d_2}\right),$$

$$\boldsymbol{x}_1=N\left(\boldsymbol{W}_1\boldsymbol{z}+\boldsymbol{V}_1\boldsymbol{z}_1,\sigma_1^2\boldsymbol{I}\right),\ \boldsymbol{W}_1\in R^{m_1\times d},\ \boldsymbol{V}_1\in R^{m_1\times d_1},$$

$$\boldsymbol{x}_2=N\left(\boldsymbol{W}_2\boldsymbol{z}+\boldsymbol{V}_2\boldsymbol{z}_2,\sigma_2^2\boldsymbol{I}\right),\ \boldsymbol{W}_2\in R^{m_2\times d},\ \boldsymbol{V}_2\in R^{m_2\times d_2}.$$

隐藏变量 $\boldsymbol{z},\boldsymbol{z}_1,\boldsymbol{z}_2$ 分别实现了随机变量之间相关的共性和随机变量自身特性的建模.但该方法在解决高维方差问题的同时也带来了新的计算问题,如,需要根据先验推理 d,d_1,d_2.2011年,Virtanen等人[20]将组稀疏(group sparsity)假设引入BCCA的ARD(automatic relevance determination)先验,只需指定 $d_c=d+d_1+d_2$ 的最大值即可实现 d,d_1,d_2 的自动选择,2015年,Virtanen等人在前期工作[20,21]的基础上,进一步将组稀疏假设引入因子分析,提出了组因子分析(group factor analysis,简称GFA).

1.3 弱匹配典型相关性分析

定义 1. 弱匹配多模态数据:设 $\boldsymbol{X}=\{\boldsymbol{x}_n\}_{n=1}^{N_1}\in R^{m_1\times N_1}$ 表示 m_1 维随机变量 $\boldsymbol{x}$ 的观察样本集合,$\boldsymbol{Y}=\{\boldsymbol{y}_n\}_{n=1}^{N_2}\in R^{m_2\times N_2}$ 表示 m_2 维随机变量 $\boldsymbol{y}$ 的观察样本集合,N_1 和 N_2 表示样本数量,其中,$\{(\boldsymbol{x}_n,\boldsymbol{y}_n)\}_{n=1}^{N_p}$ 是 N_p 对匹配样本,其余样本是否匹配未知.对于随机变量 x,$\tilde{\boldsymbol{X}}=\left[\boldsymbol{x}_1,\ldots,\boldsymbol{x}_{N_p}\right]\in R^{m_1\times N_p}$ 表示匹配样本集合.类似地可以定义 $\tilde{\boldsymbol{Y}}$.

针对弱匹配跨媒体数据问题,Blaschko等人使用流形正则化技术改进核典型相关性分析(kernel canonical correlation analysis,简称KCCA)方法,提出了SemiLRKCCA算法[22],构造了以下优化问题:

$$\max_{\boldsymbol{\alpha},\boldsymbol{\beta}}\frac{\boldsymbol{\alpha}^T\boldsymbol{K}_{X\tilde{X}}\boldsymbol{K}_{\tilde{Y}Y}\boldsymbol{\beta}}{\sqrt{\boldsymbol{\alpha}^T\left(\boldsymbol{K}_{X\tilde{X}}\boldsymbol{K}_{\tilde{X}X}+\boldsymbol{R}_X\right)\boldsymbol{\alpha}\cdot\boldsymbol{\beta}^T\left(\boldsymbol{K}_{Y\tilde{Y}}\boldsymbol{K}_{\tilde{Y}Y}+\boldsymbol{R}_Y\right)\boldsymbol{\beta}}},$$

其中,$\boldsymbol{R}_X=\varepsilon_X\boldsymbol{K}_{XX}+\frac{\gamma_X}{N_1^2}\boldsymbol{K}_{XX}\boldsymbol{L}_X\boldsymbol{K}_{XX},\boldsymbol{L}_X=\boldsymbol{D}_X-\boldsymbol{W}_X$ 为Laplacian矩阵[23],该矩阵使用集合 $\boldsymbol{X}$ 中的全部 N_1 个样本构造,$\boldsymbol{W}_{X_{ij}}$ 表示 $\boldsymbol{x}_i$ 与 $\boldsymbol{x}_j$ 之间边的权重,$\boldsymbol{D}_{X_{ii}}=\sum_{i=1}^{N_1}\boldsymbol{W}_{X_{ij}}$,核矩阵 $\boldsymbol{K}_{XX}=\phi_X\left(\boldsymbol{X}\right)^T\phi_X\left(\boldsymbol{X}\right),\boldsymbol{K}_{X\tilde{X}}=\phi_X\left(\boldsymbol{X}\right)^T\phi_X\left(\tilde{\boldsymbol{X}}\right),\boldsymbol{K}_{\tilde{X}X}=\phi_X\left(\tilde{\boldsymbol{X}}\right)^T\phi_X\left(\boldsymbol{X}\right),\boldsymbol{K}_{\tilde{X}\tilde{X}}=\phi_X\left(\tilde{\boldsymbol{X}}\right)^T\phi_X\left(\tilde{\boldsymbol{X}}\right).$

SemiLRKCCA参数过多,计算过程复杂.根据以下等式:

$$\boldsymbol{\alpha}^T\boldsymbol{K}_{X\tilde{X}}\boldsymbol{K}_{\tilde{Y}Y}\boldsymbol{\beta}=\boldsymbol{\alpha}^T\boldsymbol{X}^T\tilde{\boldsymbol{X}}\tilde{\boldsymbol{Y}}^T\boldsymbol{Y}\boldsymbol{\beta}=\boldsymbol{w}_x^T\tilde{\boldsymbol{X}}\tilde{\boldsymbol{Y}}^T\boldsymbol{w}_y,\ \boldsymbol{\alpha}^T\boldsymbol{K}_{X\tilde{X}}\boldsymbol{K}_{\tilde{X}X}\boldsymbol{\alpha}=\boldsymbol{\alpha}^T\boldsymbol{X}^T\tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}^T\boldsymbol{X}\boldsymbol{\alpha}=\boldsymbol{w}_x^T\tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}^T\boldsymbol{w}_x,$$

我们可以得到SemiLRKCCA的线性版,并重命名为SemiLRCCA.SemiLRCCA的优化问题如下:

$$\max_{\boldsymbol{w}_x,\boldsymbol{w}_y}\boldsymbol{w}_x^T\tilde{\boldsymbol{X}}\tilde{\boldsymbol{Y}}^T\boldsymbol{w}_y,$$

$$\text{s.t.}\quad\boldsymbol{w}_x^T\left(\tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}^T+\varepsilon_X\boldsymbol{I}+\frac{\gamma_X}{N_1^2}\boldsymbol{X}\boldsymbol{L}_X\boldsymbol{X}^T\right)\boldsymbol{w}_x=1,\ \boldsymbol{w}_y^T\left(\tilde{\boldsymbol{Y}}\tilde{\boldsymbol{Y}}^T+\varepsilon_Y\boldsymbol{I}+\frac{\gamma_Y}{N_2^2}\boldsymbol{Y}\boldsymbol{L}_Y\boldsymbol{Y}^T\right)\boldsymbol{w}_y=1.$$

为了解决由于一对一匹配数据过少而造成的 CCA 过拟合问题,2010 年,Kimura 等人提出了 SemiCCA 算法[24],给出了以下特征值问题:

$$\begin{pmatrix}(1-\mu)\boldsymbol{X}\boldsymbol{X}^T & \mu\tilde{\boldsymbol{X}}\tilde{\boldsymbol{Y}}^T\\ \mu\tilde{\boldsymbol{Y}}\tilde{\boldsymbol{X}}^T & (1-\mu)\boldsymbol{Y}\boldsymbol{Y}^T\end{pmatrix}\begin{pmatrix}\boldsymbol{W}_x\\ \boldsymbol{W}_y\end{pmatrix}=\lambda\begin{pmatrix}\mu\tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}^T+(1-\mu)\boldsymbol{I}_p & \\ & \mu\tilde{\boldsymbol{Y}}\tilde{\boldsymbol{Y}}^T+(1-\mu)\boldsymbol{I}_q\end{pmatrix}\begin{pmatrix}\boldsymbol{W}_x\\ \boldsymbol{W}_y\end{pmatrix},$$

其优化问题如下:

$$\max_{\boldsymbol{w}_x,\boldsymbol{w}_y} 2\mu\boldsymbol{w}_x^T\tilde{\boldsymbol{X}}\tilde{\boldsymbol{Y}}^T\boldsymbol{w}_y+(1-\mu)\left(\boldsymbol{w}_x^T\boldsymbol{X}\boldsymbol{X}^T\boldsymbol{w}_x+\boldsymbol{w}_y^T\boldsymbol{Y}\boldsymbol{Y}^T\boldsymbol{w}_y\right),$$

$$\text{s.t.}\quad \mu(\boldsymbol{w}_x^T\tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}^T\boldsymbol{w}_x+\boldsymbol{w}_y^T\tilde{\boldsymbol{Y}}\tilde{\boldsymbol{Y}}^T\boldsymbol{w}_y)+(1-\mu)\left(\boldsymbol{w}_x^T\boldsymbol{w}_x+\boldsymbol{w}_y^T\boldsymbol{w}_y\right)=1.$$

显然,SemiCCA 算法融合了 CCA 和 PCA,并通过参数 μ 调整两种方法的权重.CCA 用于匹配样本集合 $\tilde{\boldsymbol{X}}$ 与 $\tilde{\boldsymbol{Y}}$,保证了沿 $\boldsymbol{w}_x$ 和 $\boldsymbol{w}_y$ 方向投影后的匹配样本间相关性最大化,同时将 PCA 用于全部样本 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$,学习样本 $\boldsymbol{X}$ 与 $\boldsymbol{Y}$ 的全局结构信息,修正 CCA 的投影方向.

2011 年,Gu 等人针对无线传感器网定位问题中由传感器位置和信号强度构成的弱匹配跨媒体数据,提出了 PPLCA(partially paired locality correlation analysis)算法[25].PPLCA 算法分别定义随机变量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 的匹配样本与全部样本的相似性矩阵,$\boldsymbol{S}^X=\left\{\boldsymbol{S}_{ij}^X\right\}_{i,j=1}^{N_p,N_1}$ 和 $\boldsymbol{S}^Y=\left\{\boldsymbol{S}_{ij}^Y\right\}_{i,j=1}^{N_p,N_2}$. 如果 $\boldsymbol{x}_i$ 与 $\boldsymbol{x}_j$ 邻接,则 $\boldsymbol{S}_{ij}^X=\exp\dfrac{-\left\|\boldsymbol{x}_i-\boldsymbol{x}_j\right\|^2}{\sum_{i=1}^{N_p}\sum_{j=1}^{N_1}\left\|\boldsymbol{x}_i-\boldsymbol{x}_j\right\|^2/N_p(N_1-1)}$,否则 $\boldsymbol{S}_{ij}^X=0$.

PPLCA 算法使用近邻样本间的加权平均值 $\left(\sum_{j=1}^{N_1}\boldsymbol{S}_{ij}^X\boldsymbol{x}_j,\sum_{j=1}^{N_2}\boldsymbol{S}_{ij}^Y\boldsymbol{y}_j\right)$ 代替 CCA 算法中的样本均值 $(\bar{\boldsymbol{x}},\bar{\boldsymbol{y}})$ 获得以下优化问题:

$$\max_{\boldsymbol{w}_x,\boldsymbol{w}_y}\boldsymbol{w}_x^T\sum_{i=1}^{N_p}\left(\boldsymbol{x}_i-\sum_{j=1}^{N_1}\boldsymbol{S}_{ij}^X\boldsymbol{x}_j\right)\left(\boldsymbol{y}_i-\sum_{j=1}^{N_2}\boldsymbol{S}_{ij}^Y\boldsymbol{y}_j\right)^T\boldsymbol{w}_y,$$

$$\text{s.t.}\quad \boldsymbol{w}_x^T\sum_{i=1}^{N_p}\left(\boldsymbol{x}_i-\sum_{j=1}^{N_1}\boldsymbol{S}_{ij}^X\boldsymbol{x}_j\right)\left(\boldsymbol{x}_i-\sum_{j=1}^{N_1}\boldsymbol{S}_{ij}^X\boldsymbol{x}_j\right)^T\boldsymbol{w}_x=1,\ \boldsymbol{w}_y^T\sum_{i=1}^{N_p}\left(\boldsymbol{y}_i-\sum_{j=1}^{N_2}\boldsymbol{S}_{ij}^Y\boldsymbol{y}_j\right)\left(\boldsymbol{y}_i-\sum_{j=1}^{N_2}\boldsymbol{S}_{ij}^Y\boldsymbol{y}_j\right)^T\boldsymbol{w}_y=1.$$

与 PPLCA 算法使用近邻样本间相似性的思路类似,2013 年周旭东等人提出了近邻相关性分析算法(neighborhood correlation analysis,简称 NeCA)[26].

SemiCCA 算法关注于各模态内部的全局结构,而 SemiLRKCCA 算法与 PPLCA 算法均强调各模态内部的局部结构.与 SemiLRKCCA 算法相比,PPLCA 算法在目标函数和约束条件中都嵌入了样本的局部结构信息.SemiLRKCCA 算法和 SemiCCA 算法中的 Semi 不代表 Semi-supervised,而是指 Semi-paired.

2 弱匹配概率典型相关性分析模型

给定数量为 N_p 的成对观察样本集合 $\boldsymbol{X}_1^{(P)}=\left\{\left(\boldsymbol{x}_1^i\right)\right\}_{i=1}^{N_p}$ 和 $\boldsymbol{X}_2^{(P)}=\left\{\left(\boldsymbol{x}_2^i\right)\right\}_{i=1}^{N_p}$,其中,每一个样本 $\boldsymbol{x}_1^i$($\boldsymbol{x}_2^i$)代表一个 $m_1(m_2)$维向量.在成对样本数量很小的情况下,CCA 建立的相关性模型容易出现过拟合问题.下面,我们给出未匹配样本集合 $\boldsymbol{X}_1^{(U)}=\left\{\left(\boldsymbol{x}_1^j\right)\right\}_{j=N_p+1}^{N_1}$ 与/或 $\boldsymbol{X}_2^{(U)}=\left\{\left(\boldsymbol{x}_2^k\right)\right\}_{k=N_p+1}^{N_2}$,其中,$\boldsymbol{X}_1^{(U)}$ 与 $\boldsymbol{X}_2^{(U)}$ 相互独立生成.

为了解决传统 CCA 和 PCCA 模型无法直接处理未匹配样本的弊端,本文提出一种弱匹配概率典型相关性分析模型(semi-paired PCCA,简称 SemiPCCA).SemiPCCA 充分利用未匹配样本解决过拟合问题.图 3 给出了 SemiPCCA 的图模型.

$\boldsymbol{D}=\left\{\left(\boldsymbol{x}_1^i,\boldsymbol{x}_2^i\right)\right\}_{i=1}^{N_p}\cup\left\{\left(\boldsymbol{x}_1^j\right)\right\}_{j=N_p+1}^{N_1}\cup\left\{\left(\boldsymbol{x}_2^k\right)\right\}_{k=N_p+1}^{N_2}$ 表示完整的观察样本集合,包含了匹配和未匹配样本.假设样本之间相互独立,其极大似然值如下:

$$L(\theta)=\prod_{i=1}^{N_p}P\left(\boldsymbol{x}_1^i,\boldsymbol{x}_2^i;\theta\right)\prod_{j=N_p+1}^{N_1}P\left(\boldsymbol{x}_1^j;\theta\right)\prod_{k=N_p+1}^{N_2}P\left(\boldsymbol{x}_2^k;\theta\right).$$

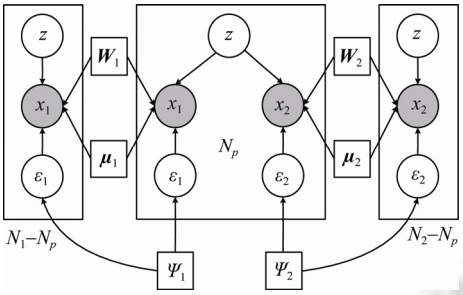

Fig.3 Graphical model for SemiPCCA

图 3 SemiPCCA 的图模型

在 SemiPCCA 模型中,对于成对样本 $\left\{\left(\boldsymbol{x}_1^i,\boldsymbol{x}_2^i\right)\right\}_{i=1}^{N_p}$, $\boldsymbol{x}_1^i$ 和 $\boldsymbol{x}_2^i$ 由相同的隐变量 $\boldsymbol{z}^i$ 生成,且 $P\left(\boldsymbol{x}_1^i,\boldsymbol{x}_2^i;\theta\right)$ 服从概率典型相关性分析(PCCA)模型,即

$$P\left(\boldsymbol{x}_1^i,\boldsymbol{x}_2^i;\theta\right)\sim N\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix},\begin{pmatrix}\boldsymbol{W}_1\boldsymbol{W}_1^T+\boldsymbol{\psi}_1 & \boldsymbol{W}_1\boldsymbol{W}_2^T\\ \boldsymbol{W}_2\boldsymbol{W}_1^T & \boldsymbol{W}_2\boldsymbol{W}_2^T+\boldsymbol{\psi}_2\end{pmatrix}\right).$$

对于未匹配样本集合 $\boldsymbol{X}_1^{(U)}=\left\{\left(\boldsymbol{x}_1^j\right)\right\}_{j=N_p+1}^{N_1}$ 和 $\boldsymbol{X}_2^{(U)}=\left\{\left(\boldsymbol{x}_2^k\right)\right\}_{k=N_p+1}^{N_2}$, $\boldsymbol{x}_1^j$ 和 $\boldsymbol{x}_2^k$ 则分别由隐变量 $\boldsymbol{z}_1^j$ 和 $\boldsymbol{z}_2^k$ 通过线性变换 $\boldsymbol{W}_1$ 和 $\boldsymbol{W}_2$ 附加高斯噪声 ε_1 和 ε_2 获得,即

$$P\left(\boldsymbol{x}_1^j;\theta\right)=\int P\left(\boldsymbol{x}_1^j|\boldsymbol{z}_1^j\right)P\left(\boldsymbol{z}_1^j\right)\mathrm{d}\boldsymbol{z}_1^j\sim N\left(\mu_1,\boldsymbol{W}_1\boldsymbol{W}_1^T+\boldsymbol{\psi}_1\right),\ P\left(\boldsymbol{x}_2^k;\theta\right)=\int P\left(\boldsymbol{x}_2^k|\boldsymbol{z}_2^k\right)P\left(\boldsymbol{z}_2^k\right)\mathrm{d}\boldsymbol{z}_2^k\sim N\left(\mu_2,\boldsymbol{W}_2\boldsymbol{W}_2^T+\boldsymbol{\psi}_2\right).$$

SemiPCCA 模型中,成对样本 $\left(\boldsymbol{x}_1^i,\boldsymbol{x}_2^i\right)$ 的投影方法类似 PCCA 模型,即

$$E\left(\boldsymbol{z}^i|\boldsymbol{x}_1^i\right)=\hat{\boldsymbol{W}}_1^T\left(\hat{\boldsymbol{W}}_1\hat{\boldsymbol{W}}_1^T+\hat{\boldsymbol{\psi}}_1\right)^{-1}\left(\boldsymbol{x}_1^i-\hat{\mu}_1\right),\ E\left(\boldsymbol{z}^i|\boldsymbol{x}_2^i\right)=\hat{\boldsymbol{W}}_2^T\left(\hat{\boldsymbol{W}}_2\hat{\boldsymbol{W}}_2^T+\hat{\boldsymbol{\psi}}_2\right)^{-1}\left(\boldsymbol{x}_2^i-\hat{\mu}_2\right).$$

$E\left(\boldsymbol{z}^i|\boldsymbol{x}_1^i\right),E\left(\boldsymbol{z}^i|\boldsymbol{x}_2^i\right)$ 构成了样本空间到 SemiPCCA 隐空间的典型投影.虽然 SemiPCCA 模型投影的结果看似与 PCCA 模型相同,但 $\hat{\boldsymbol{W}}_1$ 和 $\hat{\boldsymbol{W}}_2$ 的计算却受到了未匹配样本的影响,而未匹配样本则揭示了各领域样本空间的全局结构.同时,为使相关度最大化,不同样本空间的投影向量之间也会相互影响.

2.1 EM 算法求解 SemiPCCA

考虑到观察样本的极大似然函数 $L(\theta)$由 3 部分构成,因此 E 步骤,我们需要分别处理.

对于匹配样本集合 $\left\{\left(\boldsymbol{x}_1^i,\boldsymbol{x}_2^i\right)\right\}_{i=1}^{N_p}$ 中的第 i 对样本 $\left(\boldsymbol{x}_1^i,\boldsymbol{x}_2^i\right)$, 我们给出隐变量 $\boldsymbol{z}^i$ 的后验概率,即

$$P\left(\boldsymbol{z}^i|\boldsymbol{x}_1^i,\boldsymbol{x}_2^i;\theta\right)\sim N\left(\boldsymbol{W}^T\left(\boldsymbol{W}\boldsymbol{W}^T+\boldsymbol{\psi}\right)^{-1}\left(\begin{pmatrix}\boldsymbol{x}_1^i\\ \boldsymbol{x}_2^i\end{pmatrix}-\mu\right),\boldsymbol{I}-\boldsymbol{W}^T\left(\boldsymbol{W}\boldsymbol{W}^T+\boldsymbol{\psi}\right)^{-1}\boldsymbol{W}\right).$$

根据该后验概率 $P\left(\boldsymbol{z}^i|\boldsymbol{x}_1^i,\boldsymbol{x}_2^i;\theta\right)$, 我们计算得到 $\boldsymbol{z}^i$ 和 $\boldsymbol{z}^i\boldsymbol{z}^{iT}$ 的期望值:

$$\left\langle\boldsymbol{z}^i\right\rangle=\boldsymbol{W}^T\left(\boldsymbol{W}\boldsymbol{W}^T+\boldsymbol{\psi}\right)^{-1}\left(\begin{pmatrix}\boldsymbol{x}_1^i\\ \boldsymbol{x}_2^i\end{pmatrix}-\mu\right)\tag{3}$$

$$\left\langle\boldsymbol{z}^i\boldsymbol{z}^{iT}\right\rangle=\left\langle\boldsymbol{z}^i\right\rangle\left\langle\boldsymbol{z}^i\right\rangle^T+\boldsymbol{I}-\boldsymbol{W}^T\left(\boldsymbol{W}\boldsymbol{W}^T+\boldsymbol{\psi}\right)^{-1}W\tag{4}$$

对于未匹配样本 $\left\{\left(\boldsymbol{x}_1^j\right)\right\}_{j=N_p+1}^{N_1}$, 隐变量 $\boldsymbol{z}_1^j$ 只受 $\boldsymbol{x}_1^j$ 的影响,其后验概率的计算如下所示:

$$P\left(\boldsymbol{z}_1^j|\boldsymbol{x}_1^j;\theta\right)\sim N\left(\boldsymbol{W}_1^T\left(\boldsymbol{W}_1\boldsymbol{W}_1^T+\boldsymbol{\psi}_1\right)^{-1}\left(\boldsymbol{x}_1^j-\mu_1\right),\boldsymbol{I}-\boldsymbol{W}_1^T\left(\boldsymbol{W}_1\boldsymbol{W}_1^T+\boldsymbol{\psi}_1\right)^{-1}\boldsymbol{W}_1\right).$$

根据该后验概率 $P\left(\boldsymbol{z}_1^j|\boldsymbol{x}_1^j;\theta\right)$,我们计算得到 $\boldsymbol{z}_1^j$ 和 $\boldsymbol{z}_1^j\boldsymbol{z}_1^{jT}$ 的期望值:

$$\left\langle\boldsymbol{z}_1^j\right\rangle=\boldsymbol{W}_1^T\left(\boldsymbol{W}_1\boldsymbol{W}_1^T+\boldsymbol{\psi}_1\right)^{-1}\left(\boldsymbol{x}_1^j-\mu_1\right)\tag{5}$$

$$\left\langle\boldsymbol{z}_1^j\boldsymbol{z}_1^{jT}\right\rangle=\left\langle\boldsymbol{z}_1^j\right\rangle\left\langle\boldsymbol{z}_1^j\right\rangle^T+\boldsymbol{I}-\boldsymbol{W}_1^T\left(\boldsymbol{W}_1\boldsymbol{W}_1^T+\boldsymbol{\psi}_1\right)^{-1}\boldsymbol{W}_1\tag{6}$$

对于未匹配样本 $\left\{\left(\boldsymbol{x}_2^k\right)\right\}_{k=N_p+1}^{N_2}$,隐变量 $\boldsymbol{z}_2^k$ 只受 $\boldsymbol{x}_2^k$ 的影响,其后验概率的计算如下所示:

$$P\left(\boldsymbol{z}_2^k|\boldsymbol{x}_2^k;\theta\right)\sim N\left(\boldsymbol{W}_2^T\left(\boldsymbol{W}_2\boldsymbol{W}_2^T+\boldsymbol{\psi}_2\right)^{-1}\left(\boldsymbol{x}_2^k-\mu_2\right),\boldsymbol{I}-\boldsymbol{W}_2^T\left(\boldsymbol{W}_2\boldsymbol{W}_2^T+\boldsymbol{\psi}_2\right)^{-1}\boldsymbol{W}_2\right).$$

根据该后验概率 $P\left(\boldsymbol{z}_2^k|\boldsymbol{x}_2^k;\theta\right)$,我们计算得到 $\boldsymbol{z}_2^k$ 和 $\boldsymbol{z}_2^k\boldsymbol{z}_2^{kT}$ 的期望值:

$$\left\langle\boldsymbol{z}_2^k\right\rangle=\boldsymbol{W}_2^T\left(\boldsymbol{W}_2\boldsymbol{W}_2^T+\boldsymbol{\psi}_2\right)^{-1}\left(\boldsymbol{x}_2^k-\mu_2\right)\tag{7}$$

$$\left\langle\boldsymbol{z}_2^k\boldsymbol{z}_2^{kT}\right\rangle=\left\langle\boldsymbol{z}_2^k\right\rangle\left\langle\boldsymbol{z}_2^k\right\rangle^T+\boldsymbol{I}-\boldsymbol{W}_2^T\left(\boldsymbol{W}_2\boldsymbol{W}_2^T+\boldsymbol{\psi}_2\right)^{-1}\boldsymbol{W}_2\tag{8}$$

M 步骤,固定 E 步骤计算得到的 $P\left(\boldsymbol{z}^i|\boldsymbol{x}_1^i,\boldsymbol{x}_2^i;\theta\right),P\left(\boldsymbol{z}_1^j|\boldsymbol{x}_1^j;\theta\right)$ 和 $P\left(\boldsymbol{z}_2^k|\boldsymbol{x}_2^k;\theta\right)$,通过偏导数计算似然 $L(\theta)$最大化时参数 $\boldsymbol{W}_1,\boldsymbol{\psi}_1,\boldsymbol{W}_2,\boldsymbol{\psi}_2,\mu_1,\mu_2$ 的取值.

对于 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$ 的均值μ_1,μ_2,其取值如下:

$$\hat{\mu}_1=\tilde{\mu}_1=\frac{1}{N_1}\sum_{i=1}^{N_1}\boldsymbol{x}_1^i,\ \hat{\mu}_2=\tilde{\mu}_2=\frac{1}{N_2}\sum_{i=1}^{N_2}\boldsymbol{x}_2^i\tag{9}$$

由于 EM 算法迭代过程中,$\tilde{\mu}_1$ 和 $\tilde{\mu}_2$ 的取值不变,所以可以通过中心化样本集合 $\boldsymbol{X}_1^{(P)}\cup\boldsymbol{X}_1^{(U)},\boldsymbol{X}_2^{(P)}\cup\boldsymbol{X}_2^{(U)}$ 来避免学习过程中重复学习.为了简化描述,下文中 $\boldsymbol{x}_1^i,\boldsymbol{x}_2^i,\boldsymbol{x}_1^j$ 和 $\boldsymbol{x}_2^k$ 均表示经过中心化的向量.

对于投影向量集合 $\boldsymbol{W}_1,\boldsymbol{W}_2$,我们获得以下更新公式:

$$\hat{\boldsymbol{W}}_1=\left[\sum_{i=1}^{N_p}\boldsymbol{x}_1^i\left\langle\boldsymbol{z}^i\right\rangle^T+\sum_{j=N_p+1}^{N_1}\boldsymbol{x}_1^j\left\langle\boldsymbol{z}_1^j\right\rangle^T\right]\left[\sum_{i=1}^{N_p}\left\langle\boldsymbol{z}^i\boldsymbol{z}^{iT}\right\rangle+\sum_{j=N_p+1}^{N_1}\left\langle\boldsymbol{z}_1^j\boldsymbol{z}_1^{jT}\right\rangle\right]^{-1}\tag{10}$$

$$\hat{\boldsymbol{W}}_2=\left[\sum_{i=2}^{N_p}\boldsymbol{x}_2^i\left\langle\boldsymbol{z}^i\right\rangle^T+\sum_{k=N_p+1}^{N_2}\boldsymbol{x}_2^k\left\langle\boldsymbol{z}_2^k\right\rangle^T\right]\left[\sum_{i=1}^{N_p}\left\langle\boldsymbol{z}^i\boldsymbol{z}^{iT}\right\rangle+\sum_{k=N_p+1}^{N_2}\left\langle\boldsymbol{z}_2^k\boldsymbol{z}_2^{kT}\right\rangle\right]^{-1}\tag{11}$$

对于高斯噪声的方差 $\boldsymbol{\psi}_1,\boldsymbol{\psi}_2$,我们获得以下更新公式:

$$\hat{\boldsymbol{\psi}}_1=\frac{1}{N_1}\left\{\sum_{i=1}^{N_p}\left(\boldsymbol{x}_1^i-\hat{\boldsymbol{W}}_1\left\langle\boldsymbol{z}^i\right\rangle\right)\left(\boldsymbol{x}_1^i-\hat{\boldsymbol{W}}_1\left\langle\boldsymbol{z}^i\right\rangle\right)^T+\sum_{j=N_p+1}^{N_1}\left(\boldsymbol{x}_1^j-\hat{\boldsymbol{W}}_1\left\langle\boldsymbol{z}_1^j\right\rangle\right)\left(\boldsymbol{x}_1^j-\hat{\boldsymbol{W}}_1\left\langle\boldsymbol{z}_1^j\right\rangle\right)^T\right\}\tag{12}$$

$$\hat{\boldsymbol{\psi}}_2=\frac{1}{N_2}\left\{\sum_{i=1}^{N_p}\left(\boldsymbol{x}_2^i-\hat{\boldsymbol{W}}_2\left\langle\boldsymbol{z}^i\right\rangle\right)\left(\boldsymbol{x}_2^i-\hat{\boldsymbol{W}}_2\left\langle\boldsymbol{z}^i\right\rangle\right)^T+\sum_{k=N_p+1}^{N_2}\left(\boldsymbol{x}_2^k-\hat{\boldsymbol{W}}_2\left\langle\boldsymbol{z}_2^k\right\rangle\right)\left(\boldsymbol{x}_2^k-\hat{\boldsymbol{W}}_2\left\langle\boldsymbol{z}_2^k\right\rangle\right)^T\right\}\tag{13}$$

求解 SemiPCCA 的完整 EM 算法如下.

输入:成对样本 $\left\{\left(\boldsymbol{x}_1^i,\boldsymbol{x}_2^i\right)\right\}_{i=1}^{N_p}$,未匹配样本 $\left\{\left(\boldsymbol{x}_1^j\right)\right\}_{j=N_p+1}^{N_1}$ 和 $\left\{\left(\boldsymbol{x}_2^k\right)\right\}_{k=N_p+1}^{N_2}$,隐变量维度 d.

1: 初始化模型参数 $\theta=\{\boldsymbol{W}_1,\boldsymbol{W}_2,\boldsymbol{\psi}_1,\boldsymbol{\psi}_2\}$.

2: 使用公式(9)计算样本均值,并中心化样本集合 $\boldsymbol{X}_1^{(P)}\cup\boldsymbol{X}_1^{(U)},\boldsymbol{X}_2^{(P)}\cup\boldsymbol{X}_2^{(U)}$.

3: **repeat**

{E 步骤}

4: **for** i=1 to N_p **do**

5: 对于成对样本 $(\boldsymbol{x}_1^i,\boldsymbol{x}_2^i)$, 计算公式(3)和公式(4);

6: **end for**

7: **for** $j=N_p+1$ to N_1 **do**

8: 对于未匹配样本,计算公式(5)和公式(6);

9: **end for**

10: **for** $k=N_p+1$ to N_2 **do**

11: 对于未匹配样本 $(\boldsymbol{x}_2^k)$,计算公式(7)和公式(8);

12: **end for**

{M 步骤}

13: 使用公式(10)和公式(11)更新参数 $\boldsymbol{W}_1$ 和 $\boldsymbol{W}_2$;

14: 使用公式(12)和公式(13)更新参数 $\boldsymbol{\psi}_1$ 和 $\boldsymbol{\psi}_2$;

15: **until** 参数θ的变化小于指定阈值.

输出:模型参数θ和投影向量 $\boldsymbol{z}^i\left(i=1,\ldots,N_p\right)$.

2.2 Toy problem 实验

为了验证 SemiPCCA 模型的有效性,我们构造以下人工数据集合:样本集合 $\left\{\boldsymbol{z}^i\right\}_{i=1}^N$ 服从 $N\left(0,\boldsymbol{I}_d\right)$, 其中维度 d=2,样本数量 N=300,完整的匹配样本集合 $\left\{\left(\boldsymbol{x}_1^i,\boldsymbol{x}_2^i\right)\right\}_{i=1}^N$ 通过以下方式构造获得:

$$\boldsymbol{x}_1=\boldsymbol{T}_1\boldsymbol{z}+\varepsilon_1,\boldsymbol{T}_1\in R^{m_1\times d},\ \boldsymbol{x}_2=\boldsymbol{T}_2\boldsymbol{z}+\varepsilon_2,\boldsymbol{T}_2\in R^{m_2\times d},$$

其中,$P\left(\varepsilon_1\right)\sim N\left(0,\begin{bmatrix}0.75 & 0.5\\ 0.5 & 0.75\end{bmatrix}\right)$,$P\left(\varepsilon_2\right)\sim N\left(0,\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}\right)$,$\boldsymbol{T}_1=\begin{bmatrix}0.6 & -1/\sqrt{2}\\ 0.8 & -1/\sqrt{2}\end{bmatrix}$,$\boldsymbol{T}_2=\begin{bmatrix}0.3 & -0.7\\ 0.4 & 0.7\end{bmatrix}$, 样本维度分别设置为 m_1=2,m_2=2.

为了获得弱匹配的样本集合,我们构造一个判别函数 $f\left(\boldsymbol{x}_2\right)=\boldsymbol{a}^T\boldsymbol{x}_2-\theta$, 其中 $\boldsymbol{a}=\left(a_1,\ldots,a_{m_2}\right)^T$, θ表示判别阈值.对于样本 $\left(\boldsymbol{x}_1^i,\boldsymbol{x}_2^i\right)$, 如果其判别函数值 $f\left(\boldsymbol{x}_2^i\right)<0$, 则从 $\left\{\boldsymbol{x}_2^i\right\}_{i=1}^N$ 中移除样本.可见,θ越大,移除的样本就越多.

在比较 SemiPCCA 与传统 CCA 和 PCCA 时,我们选择了以下加权余弦距离[24]:

$$C\left(\boldsymbol{W}_x,\boldsymbol{W}_x^*,\boldsymbol{\Lambda}^*\right)=\sum_{i=1}^d\lambda_i^*\frac{\boldsymbol{w}_{x,i}^T\boldsymbol{w}_{x,i}^*}{\left\|\boldsymbol{w}_{x,i}\right\|\cdot\left\|\boldsymbol{w}_{x,i}^*\right\|},$$

其中,$\boldsymbol{W}_x^*=\left(\boldsymbol{w}_{x,1}^*,\boldsymbol{w}_{x,2}^*,\ldots,\boldsymbol{w}_{x,d}^*\right)^T$ 和 $\boldsymbol{\Lambda}^*=diag\left(\lambda_1^*,\lambda_2^*,\ldots,\lambda_d^*\right)$ 分别表示完整的匹配样本集合 $\left\{\left(\boldsymbol{x}_1^i,\boldsymbol{x}_2^i\right)\right\}_{i=1}^N$ 通过 CCA 分析后,获得的"真正"d 组典型投影向量和相关系数.使用加权余弦距离可以定量地比较投影向量偏移的程度.该加权余弦距离越大,说明相应算法求得的投影向量越接近"真正"的典型投影向量,之间的夹角越小.图 4 给出了判别阈值θ在–2~5 的取值范围内,经过 1 000 次独立实验获得的加权余弦距离平均值.图中,横坐标表示判别阈值θ,纵坐标表示加权余弦距离.实验结果表明,随着判别阈值θ的提高,匹配样本逐渐减少,CCA 和 PCCA 求得的投影向量与"真正"的典型投影向量之间的夹角在不断加大,即出现了过拟合问题.而 SemiPCCA 由于同时使用了弱匹配样本集合中的匹配样本和未匹配样本,其性能明显好于传统 CCA 和 PCCA,解决了过拟合问题,投影向量间的余弦距离相对稳定,没有随着匹配样本的减少而大幅变化.

图 5 和图 6 分别描述了当θ=–2 和θ=4 时,匹配样本(蓝色方形)、未匹配样本(红色圆形)的分布情况,以及分别由 CCA,PCCA 和 SemiPCCA 获得的 3 组典型投影向量,其中,

(1) 红色投影向量:基于完整的匹配样本集合 $\left\{\left(\boldsymbol{x}_1^i,\boldsymbol{x}_2^i\right)\right\}_{i=1}^N$,通过 CCA 或 PCCA 分析后获得的典型向量,图中以红色箭头表示.该向量是测试基准,是"真正"的典型向量.

(2) 蓝色投影向量:只考虑弱匹配样本集合中匹配样本,通过 CCA 或 PCCA 获得的典型向量,图中以蓝色箭头表示.

(3) 黑色投影向量:综合考虑了弱匹配样本集合中匹配和未匹配样本,由 SemiPCCA 获得的典型向量,图中以黑色箭头表示.

实验结果表明:

(1) 由于只使用了弱匹配样本集合中剩余的成对样本,所以蓝色投影向量严重偏离了红色箭头代表的"真正"的投影向量,即 CCA 和 PCCA 由于成对样本过少出现了过拟合问题.

(2) SemiPCCA 在参数估计的过程中,同时使用了弱匹配样本集合中的匹配样本和未匹配样本,所以相对于蓝色投影向量,其获得的黑色投影向量更加接近测试基准.

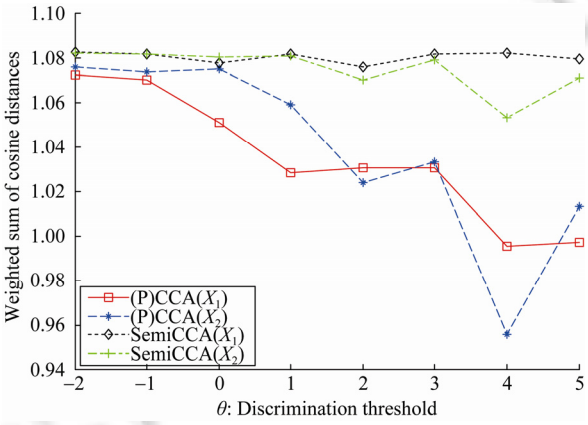

Fig.4 Weighted sum of cosine distances

图 4 加权余弦距离

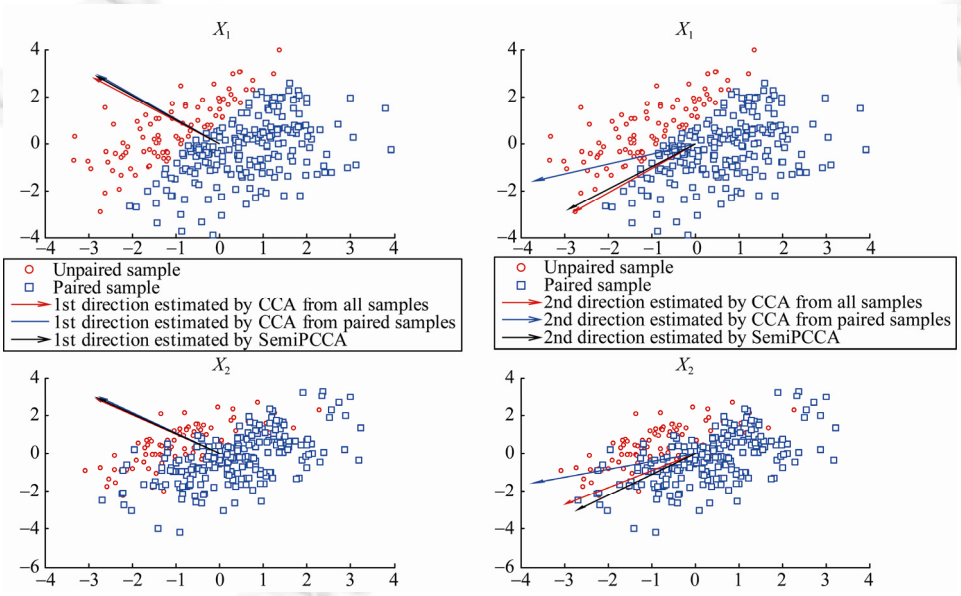

Fig.5 Distribution of canonical vectors of CCA, PCCA and SemiPCCA (a=(3,–2)^T, θ=–2)

图 5 CCA,PCCA 和 SemiPCCA 获得的典型投影向量(a=(3,–2)^T,θ=–2)

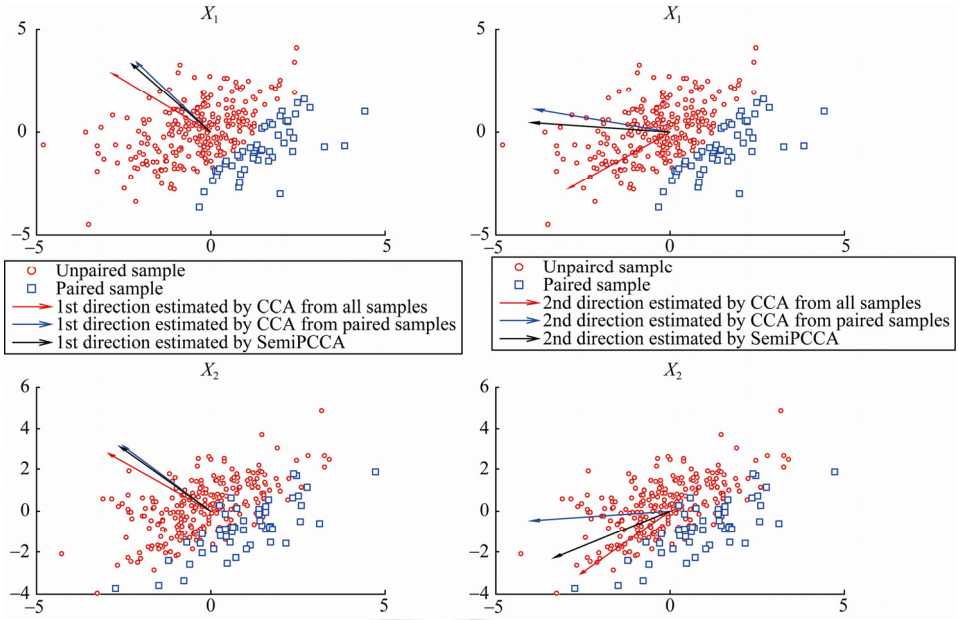

Fig.6 Distribution of canonical vectors of CCA, PCCA and SemiPCCA (a=(3,–2)^T, θ=4)

图 6 CCA,PCCA 和 SemiPCCA 获得的典型投影向量(a=(3,–2)^T,θ=4)

3 在图像语义标注领域的应用

图像检索技术包括两种主流解决方案:基于文本的图像检索和基于内容的图像检索.基于文本的图像检索利用人工对图像进行标注,并在此基础上利用传统的文本搜索引擎查询图像,这种查询方式比较直观,但是,人工标注费时、费力,使得这种检索技术不能推广到大规模的图像数据库.基于内容的图像检索采用特征提取和高维索引技术进行图像检索,它为每幅图像提取底层视觉特征,以高维形式存入数据库,通过比较这些特征的相似度来获得检索结果.这种技术在人脸识别、商标识别等某些特殊领域得到了很好的应用,但由于存在语义鸿沟,视觉特征相似的图像很可能在语义上是不相关的.为了获得语义相关的检索结果,同时避免大量的手工标注,图像自动标注成为当前关键的具有挑战性的课题[27].

图像标注方法可分为有监督的分类算法和关联建模.有监督的分类算法是一种最直接的图像标注方法.有监督的分类算法将各个语义类别(一个关键词或关键词集合)看作独立的概念,通过训练一组经过语义标注的样本图像,为每个语义类别建立各不相同的二类分类器,然后利用分类器将未标注或未归类的图像归并到某一语义类,如图 7(a)所示.最常用的有监督学习技术有贝叶斯分类器和支持向量机(support vector machine,简称 SVM).贝叶斯分类器首先选择一个图像训练集,由具有目标概念或不具有目标概念的图像组成,利用这个图像集训练一个二类贝叶斯分类器,然后将这个分类器应用到数据库中所有的图像,判断图像是否具有目标概念.Carneiro 等人对贝叶斯分类器进行改进,采用基于最小错误率的优化准则和统计分类的思想,提出一种监督多类标注算法(supervised multiclass labeling,简称 SML)[28].另一类广泛使用的分类技术是 SVM,它具有很强的理论基础,SVM 最初设计为二类分类器,在图像检索中得到了较好的应用.为了利用 SVM 学习多个语义概念,需要对每个概念单独进行训练.例如,Cusano 等人[29]将 SVM 进行推广,选择 7 类语义关键词(天空、大地、雪、建筑物等)进行实验,利用训练得到的多类 SVM 分类器对图像区域进行分类,从而产生图像的语义标注.

关联建模的方法从文本领域的研究得到启发.这类方法利用现有的标注好的图像数据集,在无监督的基础上学习图像的视觉特征和文本关键词之间的关联,然后通过统计推理的方法将这种关联应用到未标注的图像.关联建模的基本思想是引入随机变量对客观世界的隐藏状态 L 进行编码,随机变量的各个状态定义了语义关

键词和图像特征的联合分布.不同的标注方法对于隐藏状态给出了不同的定义[27].

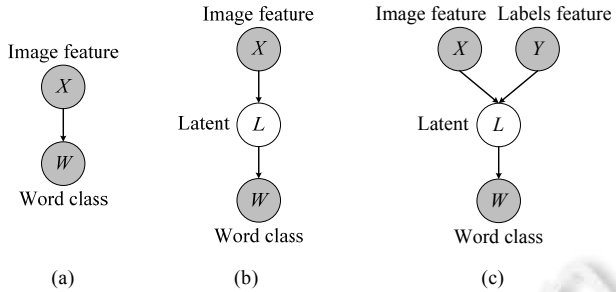

Fig.7 Approaches to the image annotation problem

图 7 图像标注方法

有些方法将图像或图像聚类与隐藏状态相联系,如机器翻译模型(translation-mode,简称 TM)[30],如图 7(b)所示.Duygulu 等人提出的机器翻译模型 TM 将图像分割为任意形状的区域,然后依据区域特征将图像区域聚类为 Blob,同时对标注关键词进行聚类,并假设图像的 Blob 与某个关键词聚类之间存在某种隐含的一一对应关系,采用 EM 算法估计图像的 Blob 和关键词的联合概率分布.借助机器翻译的概念,该模型将 Blob 和关键词看作两种对等的"语言",标注的过程可以看作是一个将 Blob 翻译为关键词的过程.类似地,还有跨媒体相关模型(cross-media relevance model,简称 CMRM)[31]、连续空间相关模型(continuous-space relevance model,简称 CRM)[32]、多贝努里相关模型(multiple bernoulli relevance model,简称 MBRM)[33].

还有方法同时使用图像和关键词估计隐藏变量的分布,实现某些模型的高层次分组(如主题)与隐藏状态相联系,如图 7(c)所示.Blei 等人使用更复杂的关联 LDA(CORR-LDA)模型为关键词和图像创建一个基于语言的关联,并在此基础上产生图像的语义标注[34].Monay 等人使用 PLSA 对跨媒体数据进行建模,并提出不对称的 PLSA 学习算法 PLSA-WORDS[35].李志欣等人在概率潜语义分析的基础上提出了融合语义主题的图像自动标注方法 PLSA-FUSION[36].李志欣等人也对传统 PLSA 模型进行改进,提出了连续 PLSA 模型处理连续量,在此基础上提出了建模连续视觉特征的图像语义标注模型 GM-PLSA[37,38].

3.1 学习与标注

类似于图 7(c)所示的关联建模方法,Harada 等人提出了基于 PCCA 的图像标注方法[39].对于已标注样本 $(\boldsymbol{x}_1,\boldsymbol{x}_2)$,隐空间中随机变量 $\boldsymbol{z}$ 的后验概率 $P(\boldsymbol{z}|\boldsymbol{x}_1,\boldsymbol{x}_2)$服从以下均值 $\hat{\boldsymbol{z}}_{12}$、方差 $\boldsymbol{\Psi}_{12}$ 的高斯分布.

$$\hat{\boldsymbol{z}}_{12}=E\left(\boldsymbol{z}|\boldsymbol{x}_1,\boldsymbol{x}_2\right)=\begin{pmatrix}\boldsymbol{M}_1\\ \boldsymbol{M}_2\end{pmatrix}^T\begin{pmatrix}\left(\boldsymbol{I}-\boldsymbol{P}_d^{\,2}\right)^{-1} & -\left(\boldsymbol{I}-\boldsymbol{P}_d^{\,2}\right)^{-1}\boldsymbol{P}_d\\ -\left(\boldsymbol{I}-\boldsymbol{P}_d^{\,2}\right)^{-1}\boldsymbol{P}_d & \left(\boldsymbol{I}-\boldsymbol{P}_d^{\,2}\right)^{-1}\end{pmatrix}\begin{pmatrix}\boldsymbol{U}_{1d}{}^T\left(\boldsymbol{x}_1-\hat{\mu}_1\right)\\ \boldsymbol{U}_{2d}{}^T\left(\boldsymbol{x}_2-\hat{\mu}_2\right)\end{pmatrix} \tag{14}$$

$$\boldsymbol{\Psi}_{12}=var\left(\boldsymbol{z}|\boldsymbol{x}_1,\boldsymbol{x}_2\right)=\boldsymbol{I}-\begin{pmatrix}\boldsymbol{M}_1\\ \boldsymbol{M}_2\end{pmatrix}^T\begin{pmatrix}\left(\boldsymbol{I}-\boldsymbol{P}_d^{\,2}\right)^{-1} & -\left(\boldsymbol{I}-\boldsymbol{P}_d^{\,2}\right)^{-1}\boldsymbol{P}_d\\ -\left(\boldsymbol{I}-\boldsymbol{P}_d^{\,2}\right)^{-1}\boldsymbol{P}_d & \left(\boldsymbol{I}-\boldsymbol{P}_d^{\,2}\right)^{-1}\end{pmatrix}\begin{pmatrix}\boldsymbol{M}_1\\ \boldsymbol{M}_2\end{pmatrix} \tag{15}$$

类似地,对于未标注的样本,隐空间中随机变量 $\boldsymbol{z}$ 在只给定样本图像特征 $\boldsymbol{x}_1$ 的情况下,其后验概率 $P(\boldsymbol{z}|\boldsymbol{x}_1)$服从以下均值 $\hat{\boldsymbol{z}}_1$、方差 $\boldsymbol{\Psi}_1$ 的高斯分布.

$$\hat{\boldsymbol{z}}_1=E\left(\boldsymbol{z}|\boldsymbol{x}_1\right)=\boldsymbol{M}_1{}^T\boldsymbol{U}_{1d}{}^T\left(\boldsymbol{x}_1-\hat{\mu}_1\right) \tag{16}$$

$$\boldsymbol{\Psi}_1=var\left(\boldsymbol{z}|\boldsymbol{x}_1\right)=\boldsymbol{I}-\boldsymbol{M}_1\boldsymbol{M}_1{}^T \tag{17}$$

根据上述结论,对于已标注图像和未标注图像,隐空间中随机变量 $\boldsymbol{z}$ 的分布情况如图 8 所示.由图 8 我们不难发现,已标注图像和未标注图像的相似性可以通过随机变量 $\boldsymbol{z}$ 的后验概率 $P(\boldsymbol{z}|\boldsymbol{x}_1,\boldsymbol{x}_2)$与 $P(\boldsymbol{z}|\boldsymbol{x}_1)$之间的 KL 距离来衡量,进而实现图像的标注.

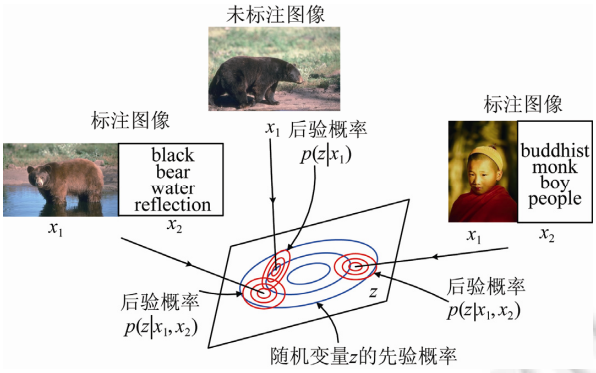

Fig.8 Posterior distribution of labelled image and unlabeled image in latent space of PCCA

图 8 在 PCCA 隐空间中标注图像和未标注图像的后验概率分布

设 $\left\{\boldsymbol{T}_i^{(P)}=\left(\boldsymbol{x}_1^i,\boldsymbol{x}_2^i\right)\right\}_{i=1}^{N_p}$ 表示已标注图像的图像特征和文本关键词集合,$\left\{\boldsymbol{Q}_j^{(U)}=\left(\boldsymbol{x}_1^j\right)\right\}_{j=N_p+1}^{N}$ 表示未标注图像的图像特征集合.Nakayam 等人[39−41]提出了一种基于 PCCA 的图像标注方法.对于给定的未标注图像 $\boldsymbol{Q}_j^{(U)}$,标注文本关键词 w 的后验概率如式(18)所示.

$$P\left(w|\boldsymbol{Q}_j^{(U)}\right)=\sum_{i=1}^{N_p}P\left(w|\boldsymbol{T}_i^{(P)}\right)P\left(\boldsymbol{T}_i^{(P)}|\boldsymbol{Q}_j^{(U)}\right) \tag{18}$$

其中,

$$P\left(\boldsymbol{T}_i^{(P)}|\boldsymbol{Q}_j^{(U)}\right)=\frac{\exp\left(-\frac{1}{8}D_{\mathrm{KL}}\left(P\left(z|\boldsymbol{T}_i^{(P)}\right),P\left(z|\boldsymbol{Q}_j^{(U)}\right)\right)\right)}{\sum_{k=1}^{N_p}\exp\left(-\frac{1}{8}D_{\mathrm{KL}}\left(P\left(z|\boldsymbol{T}_k^{(P)}\right),P\left(z|\boldsymbol{Q}_j^{(U)}\right)\right)\right)},$$

$D_{\mathrm{KL}}\left(P\left(z|\boldsymbol{T}_i^{(P)}\right),P\left(z|\boldsymbol{Q}_j^{(U)}\right)\right)$ 表示分布 $P\left(z|\boldsymbol{T}_i^{(P)}\right)$ 和 $P\left(z|\boldsymbol{Q}_j^{(U)}\right)$ 在隐空间中的 KL 距离之和,即

$$D_{\mathrm{KL}}\left(P\left(z|\boldsymbol{T}_i^{(P)}\right),P\left(z|\boldsymbol{Q}_j^{(U)}\right)\right)=\mathrm{KL}\left(P\left(z|\boldsymbol{T}_i^{(P)}\right),P\left(z|\boldsymbol{Q}_j^{(U)}\right)\right)+\mathrm{KL}\left(P\left(z|\boldsymbol{Q}_j^{(U)}\right),P\left(z|\boldsymbol{T}_i^{(P)}\right)\right).$$

根据多维高斯分布间 KL 距离的计算公式,可得:

$$\mathrm{KL}\left(P\left(z|\boldsymbol{T}_i^{(P)}\right),P\left(z|\boldsymbol{Q}_j^{(U)}\right)\right)=\frac{1}{2}\left[\mathrm{tr}\left(\boldsymbol{\psi}_1^{-1}\boldsymbol{\psi}_{12}\right)-d-\log\left(\frac{|\boldsymbol{\psi}_1|}{|\boldsymbol{\psi}_{12}|}\right)+\left(\hat{z}_q-\hat{z}_t\right)^T\boldsymbol{\psi}_1^{-1}\left(\hat{z}_q-\hat{z}_t\right)\right].$$

对于不同的样本,上式的前 3 项是常数,所以可以将以上 KL 距离简化为

$$\mathrm{KL}\left(P\left(z|\boldsymbol{T}_i^{(P)}\right),P\left(z|\boldsymbol{Q}_j^{(U)}\right)\right)=\frac{1}{2}\left(\hat{z}_q-\hat{z}_t\right)^T\boldsymbol{\psi}_1^{-1}\left(\hat{z}_q-\hat{z}_t\right) \tag{19}$$

同理,

$$\mathrm{KL}\left(P\left(z|\boldsymbol{Q}_j^{(U)}\right),P\left(z|\boldsymbol{T}_i^{(P)}\right)\right)=\frac{1}{2}\left(\hat{z}_q-\hat{z}_t\right)^T\boldsymbol{\psi}_{12}^{-1}\left(\hat{z}_q-\hat{z}_t\right) \tag{20}$$

$P\left(w|\boldsymbol{T}_i^{(P)}\right)$ 定义如下:

$$P\left(w|\boldsymbol{T}_i^{(P)}\right)=\mu\delta_{w,\boldsymbol{T}_i^{(P)}}+(1-\mu)\frac{N_w}{NW} \tag{21}$$

其中,N_w 表示标注图像集合中包含语义关键字 w 的图像数量,NW 表示语义关键字的数量.$\delta_{w,\boldsymbol{T}_i^{(P)}}=1$ 表示标注样本 $\boldsymbol{T}_i^{(P)}$ 包含语义关键字 w;否则,$\delta_{w,\boldsymbol{T}_i^{(P)}}=0$,参数 $0<\mu<1$(取μ=0.99).

在 SemiPCCA 的基础上,我们改进了上述基于 PCCA 的图像标注方法,其建模和标注过程如下.

训练阶段,首先提取训练集中每幅图像(包括标注图像和未标注图像)的视觉特征 $\boldsymbol{x}_1$,即将一幅图像的视觉

信息表示为一个连续特征向量.然后,基于标注图像的文本关键词信息$\boldsymbol{x}_2$,拟合一个SemiPCCA模型.由于在图像标注问题中只存在未标注的图像集合$\boldsymbol{X}_1^{(U)}$,而不存在未匹配的文本标注集合$\boldsymbol{X}_2^{(U)}$,所以EM算法求解SemiPCCA模型的过程中需要使用以下更新公式估计模型参数$\boldsymbol{W}_1,\boldsymbol{W}_2,\boldsymbol{\psi}_1,\boldsymbol{\psi}_2$.

$$\hat{\boldsymbol{W}}_1=\left[\sum_{i=1}^{N_p}\boldsymbol{x}_1^i\left\langle\boldsymbol{z}^i\right\rangle^T+\sum_{j=N_p+1}^{N}\boldsymbol{x}_1^j\left\langle\boldsymbol{z}_1^j\right\rangle^T\right]\left[\sum_{i=1}^{N_p}\left\langle\boldsymbol{z}^i\boldsymbol{z}^{iT}\right\rangle+\sum_{j=N_p+1}^{N}\left\langle\boldsymbol{z}_1^j\boldsymbol{z}_1^{jT}\right\rangle\right]^{-1} \tag{22}$$

$$\hat{\boldsymbol{W}}_2=\left[\sum_{i=1}^{N_p}\boldsymbol{x}_2^i\left\langle\boldsymbol{z}^i\right\rangle^T\right]\left[\sum_{i=1}^{N_p}\left\langle\boldsymbol{z}^i\boldsymbol{z}^{iT}\right\rangle\right]^{-1} \tag{23}$$

$$\hat{\boldsymbol{\psi}}_1=\frac{1}{N_1}\left\{\sum_{i=1}^{N_p}\left(\boldsymbol{x}_1^i-\hat{\boldsymbol{W}}_1\left\langle\boldsymbol{z}^i\right\rangle\right)\left(\boldsymbol{x}_1^i-\hat{\boldsymbol{W}}_1\left\langle\boldsymbol{z}^i\right\rangle\right)^T+\sum_{j=N_p+1}^{N}\left(\boldsymbol{x}_1^j-\hat{\boldsymbol{W}}_1\left\langle\boldsymbol{z}^j\right\rangle\right)\left(\boldsymbol{x}_1^j-\hat{\boldsymbol{W}}_1\left\langle\boldsymbol{z}^j\right\rangle\right)^T\right\} \tag{24}$$

$$\hat{\boldsymbol{\psi}}_2=\frac{1}{N_p}\left\{\sum_{i=1}^{N_p}\left(\boldsymbol{x}_2^i-\hat{\boldsymbol{W}}_2\left\langle\boldsymbol{z}^i\right\rangle\right)\left(\boldsymbol{x}_2^i-\hat{\boldsymbol{W}}_2\left\langle\boldsymbol{z}^i\right\rangle\right)^T\right\} \tag{25}$$

根据得到的模型参数,可以得到隐空间中随机变量$\boldsymbol{z}$在给定已标注图像$(\boldsymbol{x}_1,\boldsymbol{x}_2)$时的后验概率$P(\boldsymbol{z}|\boldsymbol{x}_1,\boldsymbol{x}_2)$,以及随机变量$\boldsymbol{z}$在只给定未标注图像$\boldsymbol{x}_1$时的后验概率$P(\boldsymbol{z}|\boldsymbol{x}_1)$,即

$$P\left(\boldsymbol{z}|\boldsymbol{x}_1,\boldsymbol{x}_2;\theta\right)\sim\mathcal{N}\left(\boldsymbol{W}^T\left(\boldsymbol{W}\boldsymbol{W}^T+\boldsymbol{\psi}\right)^{-1}\left(\begin{pmatrix}\boldsymbol{x}_1\\\boldsymbol{x}_2\end{pmatrix}-\mu\right),\boldsymbol{I}-\boldsymbol{W}^T\left(\boldsymbol{W}\boldsymbol{W}^T+\boldsymbol{\psi}\right)^{-1}\boldsymbol{W}\right) \tag{26}$$

$$P\left(\boldsymbol{z}|\boldsymbol{x}_1;\theta\right)\sim\mathcal{N}\left(\boldsymbol{W}_1^T\left(\boldsymbol{W}_1\boldsymbol{W}_1^T+\boldsymbol{\psi}_1\right)^{-1}\left(\boldsymbol{x}_1-\boldsymbol{\mu}_1\right),\boldsymbol{I}-\boldsymbol{W}_1^T\left(\boldsymbol{W}_1\boldsymbol{W}_1^T+\boldsymbol{\psi}_1\right)^{-1}\boldsymbol{W}_1\right) \tag{27}$$

标注阶段,对于每幅未标注测试图像,提取图像视觉特征$\boldsymbol{x}_1$后,可以根据公式(27)计算其投影到隐空间后随机变量$\boldsymbol{z}$的后验概率$P(\boldsymbol{z}|\boldsymbol{x}_1)$,每个文本关键词的后验概率可以通过公式(18)计算获得.与其他典型的标注模型类似,SemiPCCA为每幅图像选取5个具有最大后验概率的关键词作为其语义标注.

3.2 实验过程和结果

3.2.1 实验数据

我们采用文献[30]使用的Corel5K数据集和文献[28]使用的Corel30K数据集进行实验.Corel5K数据集包含5 000幅图像,来自50个Corel库存图像CD,每张CD包含同样语义内容的100幅图像,每幅图像标注1个~5个关键词.Corel5k共有371个关键词,将至少标注了8幅图像的关键词选入词汇表,合计260个关键词.整个数据集分为两部分:4 500幅标注图像作为训练集,500幅图像作为测试集.Corel30K数据集与Corel5K类似,但包含31 695幅图像和5 587个关键词,将至少标注了10幅图像的关键词选入词汇表,合计950个关键词.

实验中使用Corel5K测试集的500幅图像作为测试图像,从Corel5K训练集中的分别选择1 500幅、2 250幅和4 500幅图像作为标注图像,其余图像与Corel30K数据集中的31 695幅图像作为未标注图像,参与SemiPCCA模型的学习.

3.2.2 图像特征

本实验中采用基于图像颜色的高阶局部自相关特征(color higher order local auto-correlation,简称Color-HLAC).HLAC使用模板匹配的方法快速计算二值图像相邻像素点的自相关特征,能够很好地提取图像的局部信息,描述空间上的相关关系[42].*m*-th HLAC表示*m*阶HLAC特征.随着阶数的增加,HLAC特征的表示能力增强,但同时计算量也在增加,所以通常使用一阶HLAC特征(1st HLAC)或二阶HLAC特征(2nd HLAC特征).HLAC已经被广泛地用于图像识别.Color-HLAC特征是HLAC特征在RGB图像上的扩展,分别计算RGB各层的HLAC特征,然后"串行融合".一阶Color-HLAC特征为45维.二阶Color-HLAC特征为714维.

$\boldsymbol{x}=\left\{\boldsymbol{x}_1,\boldsymbol{x}_{1/2},\boldsymbol{x}_1',\boldsymbol{x}_{1/2}'\right\}$表示图像特征,其中,$\boldsymbol{x}_1$表示原始图像的二阶Color-HLAC特征,$\boldsymbol{x}_{1/2}$表示原始图像缩小一半后的二阶Color-HLAC特征,$\boldsymbol{x}_1'$和$\boldsymbol{x}_{1/2}'$分别表示上述图像的RGB层经过式(28)中的sigmoid函数二值化后的二阶Color-HLAC特征.

$$v_{\text{new}}=\frac{255}{1+\exp\left(-k\times\left(v-f_t\right)\right)} \tag{28}$$

二阶 Color-HLAC 特征经 PCA 降维后各保留 80 维,最终得到的图像特征共 320 维.

3.2.3 图像自动标注结果

本节中,使用平均精度和平均召回率比较若干图像自动标注方法的性能,包括机器翻译模型(translation-mode,简称 TM)[30]、跨媒体相关模型 CMRM[31]、连续空间相关模型 CRM[32]、多贝努里相关模型 MBRM[33]、PLSA-WORDS[35]、GM-PLSA[37,38]、PCCA[39]和本文的方法.SemiPCCA 中,隐变量 $\boldsymbol{z}$ 的维度 d=50,平滑参数 k=0.3,阈值 f_t=80.

图像标注的性能通过比较测试集的图像自动标注结果与原始标注进行评价.类似于文献[32],本文只取前 5 个后验概率最大的关键词作为每幅图像的标注结果,并计算测试集中每个关键字的精度(也称查准率)、召回率(也称查全率)及其综合评价指标 $F1$ 值.对于一个关键词 w,精度 $P=B/A$,召回率 $R=B/C$,综合评价指标 $F1=2\times P\times R/(P+R)$,其中,$A$ 表示所有自动标注了 w 的图像个数,B 表示正确标注 w 的图像个数,即这些图像的原始标注和自动标注都包含 w,C 表示原始标注中包含 w 的图像个数.计算精度和召回率的平均值可用来评价系统的标注性能.此外,本文也考虑了召回率大于 0 的关键词个数,这个值可以代表系统能够有效学习的关键词个数.

表 1 给出了 PCCA 和 SemiPCCA 在 Corel 图像库的标注性能比较,包括性能最佳的 49 个关键词的平均召回率和平均精度,以及全部 260 个关键词的平均召回率和平均精度,训练集分别选择 Corel5k 中的 1 500 幅、2 250 幅和 4 500 幅标注图像.从表 1 中数据可以看出,随着标注图像的减少,PCCA 标注图像的性能快速降低,而 SemiPCCA 的性能却相对稳定,并持续优于 PCCA.

Table 1 Performance comparison of PCCA and SemiPCCA on Corel5k dataset

表 1 PCCA,SemiPCCA 在 Corel5k 图像库上的图像自动标注性能比较

模型		PCCA			SemiPCCA		
训练样本		1 500	2 250	4 500	1 500	2 250	4 500
召回率>0 的关键词个数		99	126	150	113	132	151
性能最佳的 49 个关键词	平均召回率 R	0.61	0.74	0.89	0.71	0.85	0.94
	平均精度 P	0.56	0.65	0.7	0.60	0.72	0.77
	$F1$ 值	0.58	0.69	0.78	0.65	0.78	0.85
全部 260 个关键词	平均召回率 R	0.16	0.24	0.30	0.20	0.27	0.32
	平均精度 P	0.13	0.18	0.22	0.15	0.20	0.24
	$F1$ 值	0.14	0.21	0.25	0.17	0.23	0.27

表 2 给出了 TM,CMRM,CRM,MBRM,PLSA-WORDS,GM-PLSA 和本文提出的 SemiPCCA 的标注性能对比.为了与过去的模型进行比较,训练集采用 Corel5k 中的 4 500 幅标注图像,其中同样报告了两种标注结果:性能最佳的 49 个关键词的平均召回率和平均精度与全部 260 个关键词的平均召回率和平均精度.从表 2 中数据可以看出,SemiPCCA 的性能大幅度优于 TM,CMRM,CRM 和 PLSA-WORDS,也稍优于 MBRM 和 GM-PLSA.

Table 2 Performance comparison of SemiPCCA and other automatic image annotation models on Corel5k dataset

表 2 SemiPCCA 与其他模型在 Corel5k 图像库上的图像自动标注性能比较

模型		TM	CMRM	CRM	MBRM	PLSA-WORDS	GM-PLSA	Semi-PCCA
召回率>0 的关键词个数		49	66	107	122	105	125	151
性能最佳的 49 个关键词	平均召回率 R	0.34	0.48	0.70	0.78	0.71	0.79	0.94
	平均精度 P	0.20	0.40	0.59	0.74	0.56	0.76	0.77
	$F1$ 值	0.25	0.44	0.64	0.76	0.63	0.77	**0.85**
全部 260 个关键词	平均召回率 R	0.04	0.09	0.19	0.25	0.20	0.25	0.32
	平均精度 P	0.06	0.10	0.24	0.24	0.14	0.26	0.24
	$F1$ 值	0.05	0.09	0.21	0.24	0.16	0.25	**0.27**

4 总 结

针对弱匹配多模态数据的相关性建模问题,本文提出了一种全新的弱匹配概率典型相关性分析模型(SemiPCCA).不同于以往的弱匹配典型相关性分析模型,SemiPCCA 完全基于概率典型相关性分析模型(PCCA),关注于各模态内部的全局结构,模型参数的估计受到了未匹配样本的影响,而未匹配样本则揭示了各模态样本空间的全局结构.在人工弱匹配多模态数据集上的实验结果表明,SemiPCCA 可以有效地解决传统 CCA 和 PCCA 在匹配样本不足的情况下出现的过拟合问题,取得了很好的效果.接着,本文提出了一种基于 SemiPCCA 的图像自动标注方法.该方法基于关联建模的思想,同时使用标注图像及其关键词和未标注图像估计隐空间的分布,学习视觉模态和文本模态之间的关联,从而能够较好地对未知图像进行标注.在 Corel 数据集上进行的实验结果表明,SemiPCCA 比几种典型的图像标注方法具有更高的标注精度和更好的检索效果.

致谢 在此,我们向对本文工作予以支持和建议的老师和同学表示感谢,并向对本文工作不足之处提出评审意见的老师表示衷心的感谢.

References:

[1] Zhang H, Wu F, Zhuang, YT, Chen JX. Cross-Media retrieval method based on content correlations. Chinese Journal of Computers, 2008,31(5):820−826 (in Chinese with English abstract).

[2] Rasiwasia N, Pereira JC, Coviello E, Doyle G, Lanckriet GRG, Levy R, Vasconcelos N. A new approach to cross-modal multimedia retrieval. In: Proc. of the 18th ACM Int'l Conf. on Multimedia (MM 2010). New York: ACM, 2010. 251−260. [doi: 10.1145/1873951.1873987]

[3] Sun TK. Research on enhanced canonical correlation analysis with applications [Ph.D. Thesis]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2006 (in Chinese with English abstract).

[4] Andrew G, Arora R, Bilmes J, Livescu K. Deep canonical correlation analysis. In: Dasgupta S, McAllester D, eds. Proc. of the 30th Int'l Conf. on Machine Learning (ICML 2013). Atlanta: IMLS, 2013. 1247−1255.

[5] Sun TK, Chen SC, Yang JY, Shi PF. A novel method of combined feature extraction for recognition. In: Proc. of the 8th IEEE Int'l Conf. on Data Mining (ICDM 2008). Los Alamitos: IEEE Press, 2008. 1043−1048. [doi: 10.1109/ICDM.2008.28]

[6] Shin YJ, Park CH. Analysis of correlation based dimension reduction methods. Int'l Journal of Applied Mathematics and Computer Science, 2011,21(3):549−558. [doi: 10.2478/v10006-011-0043-9]

[7] Kursun O, Alpaydin E, Favorov OV. Canonical correlation analysis using within-class coupling. Pattern Recognition Letters, 2011, 32(2):134−144. [doi:10.1016/j.patrec.2010.09.025]

[8] Zhou XD, Chen XH, Chen SC. Combined-Feature-Discriminability enhanced canonical correlation analysis. Pattern Recognition and Aitificial Intelligence, 2012,25(2):285−291 (in Chinese with English abstract).

[9] Peng Y, Zhang DQ. Semi-Supervised canonical correlation analysis algorithm. Ruan Jian Xue Bao/Journal of Software, 2008,19 (11):2822−2832 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/19/2822.htm [doi: 10.3724/SP.J.1001.2008.02822]

[10] Kursun O, Alpaydin E. Canonical correlation analysis for multiview semi-supervised feature extraction. In: Rutkowski L, Scherer R, Tadeusiewicz R, Zadeh L, Zurada JM, eds. Proc. of the 10th Int'l Conf. on Artificial Intelligence and Soft Computing (ICAISC 2010). Heidelberg: Springer-Verlag, 2010. 430−436. [doi: 10.1007/978-3-642-13208-7_54]

[11] Chen XH, Chen SC, Xue H, Zhou XD. A unified dimensionality reduction framework for semi-paired and semi-supervised multi-view data. Pattern Recognition, 2012,45(5):2005−2018. [doi: 10.1016/j.patcog.2011.11.008]

[12] Bach FR, Jordan MI. A probability interpretation of canonical correlation analysis. Technical Report, 688, Berkeley: Department of Statistics, University of California, Berkeley, 2005.

[13] Leen G, Fyfe C. A Gaussian process latent variable model formulation of canonical correlation analysis. In: Proc. of the 14th European Symp. on Artificial Neural Networks (ESANN 2006). 2006. 413−418.

[14] Zhang B, Hao J, Ma G, Yue JP, Zhang JH, Shi ZZ. Mixture of probabilistic canonical correlation analysis. Journal of Computer Research and Development, 2015,52(7):1463−1476 (in Chinese with English abstract).

[15] Klami A, Kaski S. Local dependent components. In: Proc. of the 24th Int'l Conf. on Machine Learning (ICML 2007). New York: ACM, 2007. 425−432. [doi: 10.1145/1273496.1273550]

[16] Wang C. Variational Bayesian approach to canonical correlation analysis. IEEE Trans. on Neural Networks, 2007,18(3):905−910. [doi: 10.1109/TNN.2007.891186]

[17] Viinikanoja J, Klami A, Kaski S. Variational Bayesian mixture of robust CCA models. In: Proc. of the 2010 European Conf. on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD 2010). Heidelberg: Springer-Verlag, 2010. 370−385. [doi: 10.1007/978-3-642-15939-8_24]

[18] Archambeau C, Bach FR. Sparse probabilistic projections. In: Advances in Neural Information Processing Systems 21 (NIPS 2008). Vancouver: MIT Press, 2009. 73−80.

[19] Klami A, Virtanen S, Kaski S. Bayesian exponential family projections for coupled data sources. In: Proc. of the 26th Conf. on Uncertainty in Artificial Intelligence (UAI 2010). Corvallis: AUAI Press, 2010. 286−293.

[20] Virtanen S, Klami A, Kaski S. Bayesian CCA via group sparsity. In: Proc. of the 28th Int'l Conf. on Machine Learning (ICML 2011). Bellevue: IMLS, 2011. 457−464.

[21] Virtanen S, Klami A, Khan S, Kaski S. Bayesian group factor analysis. In: Proc. of the 15th Int'l Conf. on Artificial Intelligence and Statistics (AISTATS 2012). La Palma: JMLR, 2012. 1269−1277.

[22] Blaschko M, Lampert C, Gretton A. Semi-Supervised Laplacian regularization of kernel canonical correlation analysis. In: Proc. of the 2008 European Conf. on Machine Learning and Knowledge Discovery in Databases (ECML PKDD 2008). Heidelberg: Springer-Verlag, 2008. 133−145. [doi: 10.1007/978-3-540-87479-9_27]

[23] Belkin M, Niyogi P, Sindhwani V. Manifold regularization: A geometric framework for learning from labeled and unlabeled examples. Journal of Machine Learning Research, 2006,7:2399−2434.

[24] Kimura A, Kameoka H, Sugiyama M, Nakano T. SemiCCA: Efficient semi-supervised learning of canonical correlations. In: Proc. of the 20th Int'l Conf. on Pattern Recognition (ICPR 2010). Los Alamitos: IEEE Press, 2010. 2933−2936. [doi: 10.1109/ICPR.2010.719]

[25] Gu JJ, Chen SC, Sun TK. Localization with incompletely paired data in complex wireless sensor network. IEEE Trans. on Wireless Communications, 2011,10(9):2841−2849. [doi: 10.1109/TWC.2011.070511.100270]

[26] Zhou XD, Chen XH, Chen SC. Neighborhood correlation analysis for semi-paired two-view data. Neural Process Letter, 2013,37(3):335−354. [doi: 10.1007/s11063-012-9251-z]

[27] Li ZX. Research on semantic image annotation and retrieval [Ph.D. Thesis]. Beijing: Graduate University, The Chinese Academy of Sciences, 2010 (in Chinese with English abstract).

[28] Carneiro G, Chan AB, Moreno PJ, Vasconcelos N. Supervised learning of semantic classes for image annotation and retrieval. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2007,29(3):394−410. [doi: 10.1109/TPAMI.2007.61]

[29] Cusano C, Ciocca G, Schettini R. Image annotation using SVM. In: Proc. of the SPIE, Vol.5304. San Jose: SPIE, 2003. 330−338. [doi:10.1117/12.526746]

[30] Duygulu P, Barnard K, de Freitas JFG, Forsyth D. Object recognition as machine translation: Learning a lexicon for a fixed image vocabulary. In: Proc. of the 7th European Conf. on Computer Vision (ECCV 2002). Heidelberg: Springer-Verlag, 2002. 97−112. [doi: 10.1007/3-540-47979-1_7]

[31] Jeon J, Lavrenko V, Manmatha R. Automatic image annotation and retrieval using cross-media relevance models. In: Proc. of the 26th Annual Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR 2003). New York: ACM, 2003. 119−126. [doi: 10.1145/860435.860459]

[32] Lavrenko V, Manmatha R, Jeon J. A model for learning the semantics of pictures. In: Advances in Neural Information Processing Systems 16 (NIPS 2003). Vancouver, Whistler: MIT Press, 2003. 553−560.

[33] Feng SL, Manmatha R, Lavrenko V. Multiple Bernoulli relevance models for image and video annotation. In: Proc. of the 2004 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR 2004). Los Alamitos: IEEE Press, 2004. 1002−1009. [doi: 10.1109/CVPR.2004.1315274]

[34] Blei DM, Jordan MI. Modeling annotated data. In: Proc. of the 26th Annual Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR 2003). New York: ACM, 2003. 127−134. [doi: 10.1145/860435.860460]

[35] Monay F, Gatica-Perez D. Modeling semantic aspects for cross-media image indexing. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2007,29(10):1802−1817. [doi: 10.1109/TPAMI.2007.1097]

[36] Li ZX, Shi ZP, Liu X, Li ZQ, Shi ZZ. Fusing semantic aspects for image annotation and retrieval. Journal of Visual Communication and Image Representation, 2010,21(8):798−805. [doi: 10.1016/j.jvcir.2010.06.004]

[37] Li ZX, Shi ZP, Liu X, Shi ZZ. Modeling continuous visual features for semantic image annotation and retrieval. Pattern Recognition Letters, 2011,32(3):516−523. [doi:10.1016/j.patrec.2010.11.015]

[38] Li ZX, Shi ZP, Liu X, Shi ZZ. Semantic image annotation by modeling continuous visual features. Journal of Computer-Aided Design & Computer Graphics, 2010,22(8):1412−1420 (in Chinese with English abstract).

[39] Harada T, Nakayama H, Kuniyoshi Y. Image annotation and retrieval based on efficient learning of contextual latent space. In: Proc. of the 2009 IEEE Int'l Conf. on Multimedia and Expo (ICME 2009). Los Alamitos: IEEE Press, 2009. 858−861. [doi: 10.1109/ICME.2009.5202630]

[40] Nakayama H, Harada T, Kuniyoshi Y. Canonical contextual distance for large-scale image annotation and retrieval. In: Proc. of the 1st ACM Workshop on Large-Scale Multimedia Retrieval and Mining (LS-MMRM 2009). New York: ACM, 2009. 3−10. [doi: 10.1145/1631058.1631062]

[41] Nakayama H, Harada T, Kuniyoshi Y. Evaluation of dimensionality reduction methods for image auto-annotation. In: Proc. of the 21st British Machine Vision Conf. (BMVC 2010). British Machine Vision Association (BMVA), 2010. 1−12. [doi: 10.5244/C.24.94]

[42] Nakayama H, Harada T, Kuniyoshi Y, Otsu N. High-Performance image annotation and retrieval for weakly labeled images using latent space learning. In: Proc. of the 9th Pacific Rim Conf. on Multimedia (PCM 2008). Heidelberg: Springer-Verlag, 2008. 601−610. [doi: 10.1007/978-3-540-89796-5_62]

附中文参考文献:

[1] 张鸿,吴飞,庄越挺,陈建勋.一种基于内容相关性的跨媒体检索方法.计算机学报,2008,31(5):820−826.

[3] 孙廷凯.增强型典型相关分析研究与应用[博士学位论文].南京:南京航空航天大学,2006.

[8] 周旭东,陈晓红,陈松灿.增强组合特征判别性的典型相关分析.模式识别与人工智能,2012,25(2):285−291.

[9] 彭岩,张道强.半监督典型相关分析算法.软件学报,2008,19(11):2822−2832. http://www.jos.org.cn/1000-9825/19/2822.htm [doi: 10.3724/SP.J.1001.2008.02822]

[14] 张博,郝杰,马刚,岳金朋,张建华,史忠植.混合概率典型相关性分析.计算机研究与发展,2015,52(7):1463−1476.

[27] 李志欣.图像语义标注和检索的研究[博士学位论文].北京:中国科学院研究生院,2010.

[38] 李志欣,施智平,刘曦,史忠植.建模连续视觉特征的图像语义标注方法.计算机辅助设计与图形学学报,2010,22(8):1412−1420.

张博(1981－),男,陕西延川人,博士,讲师,CCF 专业会员,主要研究领域为人工智能,机器学习,云计算.

马刚(1986－),男,博士生,CCF 学生会员,主要研究领域为人工智能,机器学习,神经计算,复杂网络.

郝杰(1980－),女,副教授,主要研究领域为人工智能,机器学习,医学影像处理.

史忠植(1941－),男,研究员,博士生导师,CCF 会士,主要研究领域为人工智能,机器学习,神经计算,认知科学.