

## 多视角数据缺失补全\*

杨旭<sup>1,2</sup>, 朱振峰<sup>1,2</sup>, 徐美香<sup>1,2</sup>, 张幸幸<sup>1,2</sup>



<sup>1</sup>(北京交通大学 计算机科学与技术学院, 北京 100044)

<sup>2</sup>(现代信息科学与网络技术北京市重点实验室(北京交通大学), 北京 100044)

通讯作者: 朱振峰, E-mail: zhfzhu@bjtu.edu.cn

**摘要:** 随着信息技术的快速发展, 现实生活中不断涌现出大量的多视角数据, 由此应运而生的多视角学习已成为机器学习领域的研究热点. 然而, 在数据获取过程中, 由于收集的难度、高额成本或设备故障等问题, 往往导致收集到的多视角数据出现视角缺失, 这使得一些多视角学习方法无法有效进行. 为此, 提出一种基于视角相容性的多视角数据缺失补全方法. 通过监督的共享子空间学习, 获得与每类多视角数据相对应的共享子空间, 从而建立视角相容性判别模型. 与此同时, 基于共享子空间重构误差等同分布的假设, 提出了针对视角缺失的多视角数据的共享表征获取方法, 实现多视角缺失数据的预补全. 在此基础上, 进一步通过多元线性回归实现缺失视角的精确补全. 此外, 还把所提出的视角补全方法拓展到解决含有噪声的多视角数据的降噪问题. 在 UCI、COIL-20 以及人工合成数据集上的实验结果验证了所提算法的有效性.

**关键词:** 多视角学习; 缺失数据补全; 共享子空间学习; 视角相容性; 跨媒体计算

**中图法分类号:** TP391

中文引用格式: 杨旭, 朱振峰, 徐美香, 张幸幸. 多视角数据缺失补全. 软件学报, 2018, 29(4): 945-956. <http://www.jos.org.cn/1000-9825/5416.htm>

英文引用格式: Yang X, Zhu ZF, Xu MX, Zhang XX. Missing view completion for multi-view data. Ruan Jian Xue Bao/Journal of Software, 2018, 29(4): 945-956 (in Chinese). <http://www.jos.org.cn/1000-9825/5416.htm>

### Missing View Completion for Multi-View Data

YANG Xu<sup>1,2</sup>, ZHU Zhen-Feng<sup>1,2</sup>, XU Mei-Xiang<sup>1,2</sup>, ZHANG Xing-Xing<sup>1,2</sup>

<sup>1</sup>(School of Computer Science and Information Technology, Beijing Jiaotong University, Beijing 100044, China)

<sup>2</sup>(Beijing Key Laboratory of Advanced Information Science and Network Technology (Beijing Jiaotong University), Beijing 100044, China)

**Abstract:** With the rapid development of information technology, massive amounts of multi-view data are constantly emerging in people's daily life. To cope with such situation, multi-view learning has received much attention in the field of machine learning to promote the ability of data understanding. However, due to the difficulties such as high cost and equipment failure in multi-view data collection, part or all of observed values from one view can't be available, which prevents some traditional multi-view learning algorithms from working effectively as expected. This paper focuses on the missing view completion for multi-view data and proposes a view compatibility based completion method. For each class of multi-view data, a corresponding shared subspace is built by means of supervised learning. With the multiple shared subspaces, a view compatibility discrimination model is developed. Meanwhile, assuming

\* 基金项目: 国家自然科学基金(61572068, 61532005); 中央高校基本科研业务费专项基金(2015JBM039)

Foundation item: National Natural Science Foundation of China (61572068, 61532005); Fundamental Research Funds for the Central Universities (2015JBM039)

本文由“多媒体大数据处理与分析”专题特约编辑赵耀教授、李波教授、华先胜研究员、文继荣教授、蒋刚毅教授、常冬霞副教授推荐.

收稿时间: 2017-04-30; 修改时间: 2017-06-26; 采用时间: 2017-10-13; jos 在线出版时间: 2017-12-01

CNKI 网络优先出版: 2017-12-04 06:49:30, <http://kns.cnki.net/kcms/detail/11.2560.TP.20171204.0649.016.html>

that the reconstruction error of each of view of multi-view data in the shared subspace takes the independent identical distribution, an approach is put forward to seek the shared representation of multi-view data with missing view. Thus, the preliminary completion of missing view can be performed. In addition, the multiple linear regression technique is implemented to obtain a more accurate completion. Furthermore, the proposed missing view completion method is enhanced to deal with the case of the denoising of noise-polluted multi-view data. The experimental results on some datasets including UCI and Coil-20 have demonstrated the effectiveness of the proposed missing view completion method for multi-view data.

**Key words:** multi-view-learning; missing data completion; shared subspace learning; view compability; cross-media computing

随着信息技术的迅猛发展,数据的产生和收集呈现出多态性、多描述性和多源性的特点,这些数据统称为多视角数据<sup>[1]</sup>.例如:来自网页的消息,既可以由文本信息描述,也可由视频、音频、图片等信息加以描述;而在医学诊疗过程中,可以利用核磁共振(MRI)、正电子成像技术(PET)或其他生物标记(如 CSF)对阿兹海默病(Alzheimer)进行综合诊断、对数字图像利用不同技术手段提取多种特征等等.研究如何合理地利用针对同一对象广泛存在的多个视角信息,实现对由多视角所刻画的对象深度分析与理解,已成为当前机器学习领域的研究热点<sup>[1-4]</sup>.这类以面向多视角数据为研究对象的机器学习方法目前已被普遍接受为一种新的学习方式,即多视角学习(multi-view-learning,简称 MVL).

然而,在现实环境中,由于数据收集的困难性、高额成本或设备故障等,使得收集到的多视角数据中的某一视角的部分或全部属性出现缺失.多视角数据的缺失使得对由多视角所刻画的对象进行深度分析与理解变得更加困难<sup>[5]</sup>.而在一些现实问题中,则会造成对未来的天气情况无法精确预测、病人的病情无法精确诊断等.如图 1 所示,在天气预报中,往往需要对连续收集到的某一地区不同地理位置气象观测数据进行融合,以提高对该地区天气变化预测的可靠性.然而,由于受到台风或设备故障的影响,导致某一地理位置气象观测系统上的传感器在某些时刻非正常工作,从而导致如气压、温度等部分数据缺失.在现实的生物医学诊断中(如阿兹海默病),也经常面临着数据缺失问题.图 1 中,◇表示地理位置 2 在  $n+1$  时刻由于气温计等非正常工作,导致温度等特征数据缺失.图 2 为一些针对阿兹海默病的诊断方法.其中,脑脊液(CSF)会对病人产生某种不利影响或攻击性,使得病人往往拒绝该项检查,造成对病人某一项诊断数据的完全缺失.图 2 中,◇表示第  $n+1$  个病人的 CSF 诊断数据全部缺失.

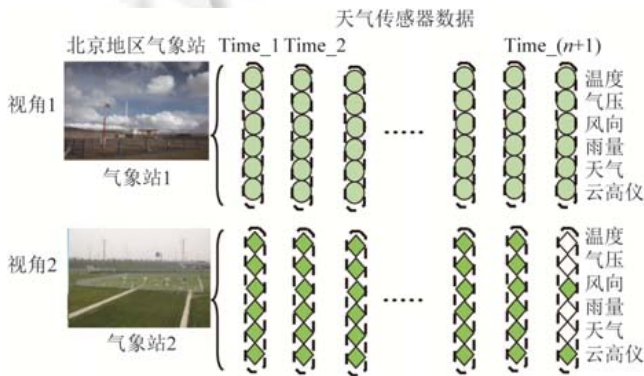


Fig.1 Illustration of the observed data from different geographic positions

图 1 不同地理位置的气象观测数据示意图

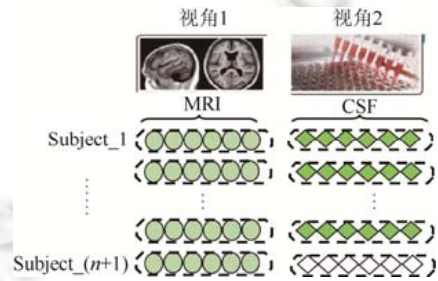


Fig.2 Illustration of different diagnoses on Alzheimer's disease patients

图 2 阿兹海默病病人的不同诊断技术图

多视角数据的缺失使得以面向多视角数据为研究对象的多视角学习方法无法有效进行,如典型相关性分析(canonical correlation analysis,简称 CCA)<sup>[6]</sup>、协同训练(co-training)<sup>[7]</sup>、多核学习(multiple kernel learning,简称 MKL)<sup>[8]</sup>等均要求来自不同视角数据的描述能够完整呈现.数据缺失使得这些多视角学习方法无法有效进行.针对多视角数据缺失问题,直观的方法是丢弃有视角数据缺失的样本,这将使得学习过程中可用于训练的样本数量减少,从而会导致泛化性能的降低.如何有效地补全缺失数据,提升对基于多视角描述所刻画的对象深度分

析与理解是目前多视角数据分析方面值得研究的课题。

近年来,补全缺失数据的方法主要针对单视角缺失数据<sup>[9,10]</sup>,如矩阵填充<sup>[11,12]</sup>,假设待填充矩阵是低秩矩阵,对于一个秩为  $r$  的矩阵,至少需要采样数大于  $2nr-r^2$  个观测值,才有可能恢复出这个矩阵.期望极大化(expectation maximization,简称 EM)算法<sup>[13]</sup>假设缺失特征与非缺失特征之间存在依赖关系,利用这种依赖关系对缺失特征进行迭代估计,进行补全,但收敛速度慢、效率低是该方法的主要问题。

相比于单视角数据缺失补全,针对多视角数据缺失补全的研究并未引起广泛的关注.对于该问题,一种方式是可直接利用现有的矩阵补全方法直接对每一视角构成的缺失数据矩阵进行补全,但是,由于此类方法并没有充分利用异构的多视角间的语义互补性,使得补全性能难以令人满意.另外一种直观的方法是可利用多视角数据视角间的近邻关系的传递性,通过使用简单的“近邻”法实现缺失数据补全,但是,由于没有充分利用多视角数据间各自所蕴含信息的互补性,同时,近邻关系传递性在异构的多视角间很难得到满足,使得补全后的数据与真实值有较大偏差.通过利用多视角间的相关性,Van de Velden 等人提出了基于广义典型相关性分析的缺失数据补全方法 (generalized canonical correlation analysis with missing values)<sup>[14]</sup>,通过迭代恢复多视角缺失数据的共享表征,但其并未实现视角缺失数据的补全.最近,Li 等人提出了基于深度学习的图像数据补全方法(deep learning based imaging data completion,简称 DLIDC)<sup>[15]</sup>.由于 DLIDC 采用卷积神经网络方法实现缺失数据的补全,因而需要大量样本数据训练网络,而对于小样本数据则易出现过拟合现象。

多视角数据的存在,体现了对同一语义对象的不同描述方式.挖掘多视角数据间的相容互补性,有利于提升对多视角涵盖的语义对象的理解与分析.针对多视角数据的缺失问题,本文提出了一种基于视角间相容性的多视角缺失数据补全方法.通过监督的共享子空间学习,获得与每类多视角数据相对应的共享子空间,从而建立视角相容性判别模型.与此同时,基于共享子空间重构误差等同分布的假设,提出了针对视角缺失的多视角数据的共享表征获取方法,实现多视角缺失数据的预补全.在此基础上,进一步通过多元线性回归实现精确补全。

本文第 1 节简要介绍所提出的基于视角相容性的多视角缺失数据补全框架.第 2 节介绍针对多视角数据的视角相容性判决模型,并在此基础上实现缺失视角的预补全.第 3 节介绍通过多元线性回归实现缺失视角的精确补全.第 4 节阐述如何把所提出的缺失视角补全方法拓展到解决含有噪声的多视角数据的降噪.第 5 节给出实验结果与分析.最后给出本文结论。

## 1 基于视角相容性的多视角缺失数据补全框架

### 1.1 符号说明

为便于阐述,首先对一些符号进行说明.令  $\{x, y | x \in R^{s \times 1}, y \in R^{r \times 1}\}$  表示来自视角  $X$  与视角  $Y$  的多视角数据(不失一般性,本文只考虑两个视角);  $X = [X_1, X_2, \dots, X_C] \in R^{s \times n}$  与  $Y = [Y_1, Y_2, \dots, Y_C] \in R^{r \times n}$  表示具有  $C$  类的多视角数据集,其中,  $X_i = [x_i^{(1)}, x_i^{(2)}, \dots, x_i^{(n_i)}] \in R^{s \times n_i}$  与  $Y_i = [y_i^{(1)}, y_i^{(2)}, \dots, y_i^{(n_i)}] \in R^{r \times n_i}$  表示第  $i$  类多视角数据,  $i=1, 2, \dots, C$ ,  $n_i$  表示第  $i$  类多视角数据的样本数目.假设  $\{x^c, y\}$  是一个视角缺失的多视角数据,令其视角  $y$  完整,  $x^c$  为对应的来自  $X$  视角的完整数据.不失一般性,对于  $x^c$  采用如下表示方式:

$$x^c = x^a + x^m \tag{1}$$

式中,  $x^a = \left[ \left( x^c [1:l] \right)^T, \overbrace{0, \dots, 0}^{s-l} \right]^T \in R^{s \times 1}$  与  $x^m = \left[ \overbrace{0, \dots, 0}^l, \left( x^c [l+1:s] \right)^T \right]^T \in R^{s \times 1}$  分别表示  $x^c$  的未缺失与缺失部分;不难看出,当  $l=0$  时,则表示视角  $x^c$  完全缺失.这里,缺失视角数据补全的目的是通过视角间的相关性获得  $x^m$  的近似  $\hat{x}^m$ 。

### 1.2 算法框架

图 3 所示为本文提出的基于视角相容性的多视角缺失数据补全框架图.其主要由两部分组成:(1) 离线的多

视角相容性判决模型学习;(2) 在线多视角缺失数据回归补全.

**离线多视角相容性判决模型学习.**针对每类多视角训练样本  $X_i$  与  $Y_i, i=1, \dots, C$ , 分别通过共享子空间学习, 获得多视角数据的共享表征  $U_i$  以及相应视角的载入矩阵(loading matrix)  $A_i$  与  $B_i$ . 在此基础上, 通过所构建的多个共享子空间, 建立基于视角相容性判决模型(详见第 2 节).

**在线多视角缺失数据补全.**针对多视角缺失数据  $\{x^a, y\}$ , 基于上述离线所建立的视角相容性判决模型, 获取其共享表征  $\hat{u}^a$  并实现多视角缺失数据的预补全. 在此基础上, 进一步通过多元线性回归实现缺失视角的精确补全(详见第 4 节).

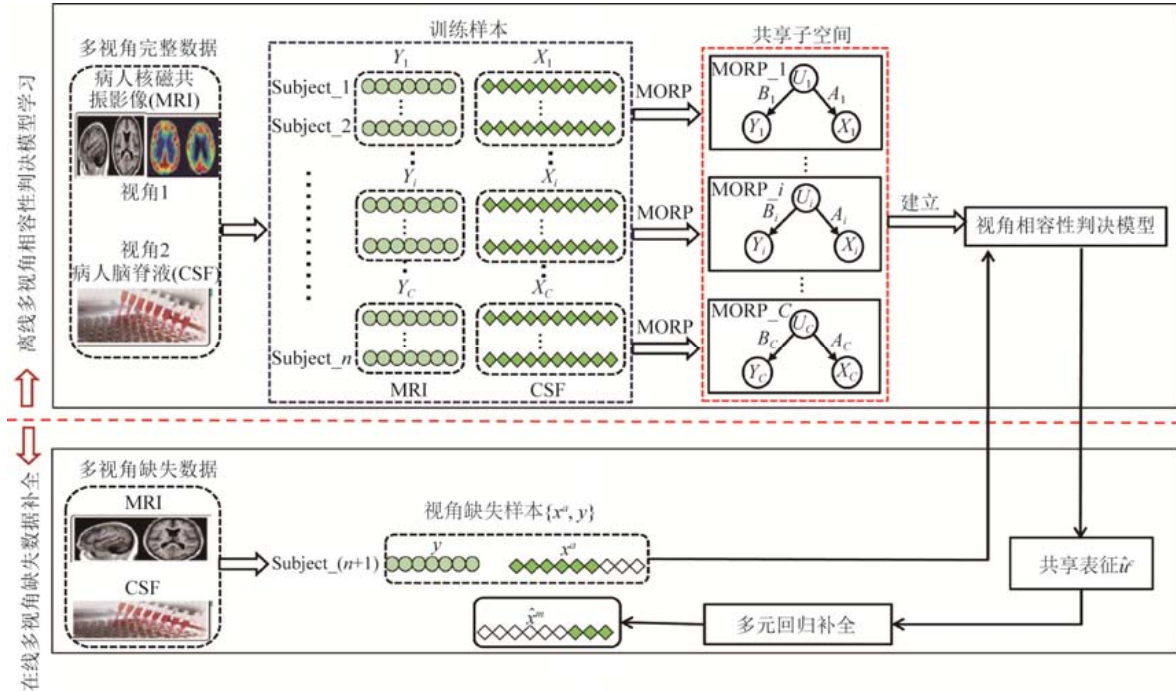


Fig.3 Multi-View missing data completion framework based on view compatibility

图 3 基于视角相容性的多视角缺失数据补全框架图

## 2 视角相容性判决模型

### 2.1 基于MORP的共享子空间学习

如上所述, 在离线多视角相容性判决模型学习中, 需利用共享子空间方法建立多视角数据视角间的关联并获得共享表征. 为此, 在本文中, 采用多输出正则投影(multi-output regularized feature projection, 简称 MORP)<sup>[16]</sup> 方法建立视角间的关联, 将多视角数据映射到共享子空间, 以此获得相应多视角数据的共享表征. MORP 的基本原理是通过因子分解将  $X$  视角数据和  $Y$  视角数据映射到共享子空间, 其优化模型为

$$\min_{A, B, U} f(A, B, U | X, Y) = \left. \begin{aligned} &(1 - \beta) \|X^T - U \cdot A\|_F^2 + \beta \|Y^T - U \cdot B\|_F^2 \\ &\text{s.t. } U^T \cdot U = I \end{aligned} \right\} \quad (2)$$

其中,  $A \in R^{p \times s}$  和  $B \in R^{p \times t}$  分别是视角  $X$  和视角  $Y$  的载入矩阵,  $U \in R^{n \times p}$  为在由载入矩阵  $\{A, B\}$  所确定的  $p$  维共享子空间中的共享表征,  $\beta (0 < \beta < 1)$  为权重系数.

针对在线多视角数据  $\{x^a, y\}$ , MORP 模型并没有给出获取其共享表征的有效求解方法. 一般情况下, 需要重新求解式(2), 当  $n$  较大时, 其计算复杂度将非常高. 为此, 在我们先前的工作中提出了一种高效的在线样本扩展方

法<sup>[17]</sup>.设  $A, B$  为训练样本经过 MORP 共享子空间学习获得的相应视角载入矩阵,则对于在线的多视角数据  $\{x^c, y\}$ , 其在共享子空间的共享表征  $u^c$  可通过下式给出:

$$\min_{u^c} f(u^c | A, B, x^c, y) = (1 - \beta) \left\| (x^c)^T - u^c \cdot A \right\|_F^2 + \beta \left\| y^T - u^c \cdot B \right\|_F^2 \quad (3)$$

从而有:

$$u^c = \left[ (1 - \beta) (x^c)^T \cdot A^T + \beta y^T \cdot B^T \right] \cdot \left[ (1 - \beta) A \cdot A^T + \beta B \cdot B^T \right]^{-1} \quad (4)$$

### 2.2 视角相容性

对于多视角数据  $\{x^c, y\}$ , 由公式(4)求得的其在共享子空间的共享表征  $u^c$  实际上是对多视角间共享信息的表达. 而由共享表征  $u^c$  所反映的共享子空间的重构误差来反映信息的共享程度, 本文称其为“视角相容性”.

为对视角相容性给出定量评价, 可利用公式(3)中基于共享子空间的多视角总体重构误差函数  $f(u^c | A, B, x^c, y)$  作为相容性判据. 重构误差值越小, 表明多视角描述间的相容性越强. 与此同时, 在具有类标签信息的监督学习情况下, 来自不同类的样本在由某载入矩阵  $\{A, B\}$  所确定的共享子空间中的相容性将存在很大差异. 特别地, 对于由第  $i$  类样本  $\{X_i, Y_i\}$  所构成的共享子空间, 我们将有如下直观的观察:

$$f_i(u_i^c | A_i, B_i, x^c, y)_{\{x^c, y\} \in \{X_i, Y_i\}} < f_i(u_i^c | A_i, B_i, x^c, y)_{\{x^c, y\} \notin \{X_i, Y_i\}}$$

为此, 基于相容性判据  $f(u^c | A, B, x^c, y)$ , 可获得多视角数据  $\{x^c, y\}$  的类属判别输出:

$$z = \arg \min_{j \in \{1, \dots, C\}} f_j(u_j^c | A_j, B_j, x^c, y) \quad (5)$$

### 2.3 视角缺失的多视角相容性判据

针对多视角数据, 公式(5)给出了基于视角相容性的类判决模型. 然而, 对于具有视角缺失的多视角数据  $\{x^c, y\}$ , 公式(5)将不再适用. 引入矩阵  $A^a = \left[ A[:, 1:l], \overbrace{0, \dots, 0}^{s-l} \right] \in R^{p \times s}$  及  $A^m = \left[ \overbrace{0, \dots, 0}^l, A[:, l+1:s] \right] \in R^{p \times s}$ , 其中,  $\mathbf{0} = [0, 0, \dots, 0]^T \in R^{p \times 1}$  为全零向量, 则公式(3)中的相容性判据  $f(u^c | A, B, x^c, y)$  将变为

$$f(u^c | A, B, x^c, y) = (1 - \beta) \left\| (x^a + x^m)^T - u^c \cdot (A^a + A^m) \right\|_F^2 + \beta \left\| y^T - u^c \cdot B \right\|_F^2$$

通过简单的变换可得:

$$f(u^c | A, B, x^c, y) = (1 - \beta) \left\| (x^a)^T - u^c \cdot A^a \right\|_F^2 + (1 - \beta) \left\| (x^m)^T - u^c \cdot A^m \right\|_F^2 + \beta \left\| y^T - u^c \cdot B \right\|_F^2 \quad (6)$$

此时, 由于  $x^m$  未知, 故无法通过最小化公式(6)获得与  $\{x^c, y\}$  对应的共享表征  $u^c$ , 进而无法利用公式(6)作为多视角的相容性判据.

对于多视角数据, 假设来自每一视角局部邻域范围内的数据在共享子空间中基于共享表征的重构误差服从高斯分布, 即当用  $e_{x^c} = (x^c)^T - u^c \cdot A$  与  $e_y = y^T - u^c \cdot B$  分别表示在共享子空间中基于共享表征  $u^c$  对  $x$  与  $y$  进行重构的重构误差时, 即有  $e_{x^c} \sim G(\bar{e}_{x^c}, \sigma_{x^c}), e_y \sim G(\bar{e}_y, \sigma_y)$ , 其中,  $\bar{e}_{x^c}, \bar{e}_y$  分别表示  $X$  视角和  $Y$  视角重构误差的统计均值向量,  $\sigma_{x^c}, \sigma_y$  分别表示  $X$  视角和  $Y$  视角重构误差的统计协方差. 为此, 对于具有视角缺失的多视角数据  $\{x^a, y\}$ , 为获得视角相容性判据, 基于上述假设, 本文对公式(6)做出修正, 并有:

$$f(u^c | e_{x^a}, e_y, \bar{e}_{x^a}, \bar{e}_y) = (1 - \beta) \left\| e_{x^a} - \bar{e}_{x^a} \right\|_F^2 + \beta \left\| e_y - \bar{e}_y \right\|_F^2 \quad (7)$$

其中,

$$e_{x^a} = (x^a)^T - u^c \cdot A^a,$$

而  $\bar{e}_{x^a} = \frac{1}{k} \sum_{j=1}^k \left( \left( x_{(j)}^c [1:l]^T, \overbrace{0, \dots, 0}^{s-l} \right) - u_{(j)}^c \cdot A^a \right)_{x_{(j)}^c [1:l] \in N_k(x^c [1:l])}$  与  $\bar{e}_y = \frac{1}{k} \sum_{j=1}^k \left( \left( y_{(j)} \right)^T - u_{(j)} \cdot B \right)_{y_{(j)} \in N_k(y)}$  分别表示  $x^c [1:l]$

与  $y$  的  $k$  个近邻在共享子空间基于共享表征进行重构的平均重构误差,  $N_k(\cdot)$  表示  $k$  近邻集合.

对比公式(6)和公式(7),我们不难发现,由于存在视角缺失(即  $x^m$  未知),公式(6)中的第 2 项没有出现在公式(7)中.此外,公式(7)利用了局部邻域样本在共享子空间中重构误差具有同分布的假设,而非直接对重构误差最小化.这样有利于在视角缺失情况下,获得共享表征的最佳逼近.

进一步地,通过最小化公式(7)可得  $u^c$  的逼近:

$$\hat{u}^c = \left[ (1-\beta) \left( \left( x^a \right)^T - \bar{e}_{x^a} \right) \cdot \left( A^a \right)^T + \beta \left( y^T - \bar{e}_y \right) \cdot B^T \right] \cdot \left[ (1-\beta) A^a \cdot \left( A^a \right)^T + \beta B \cdot B^T \right]^{-1} \quad (8)$$

与此同时,基于公式(7)和公式(8),进一步可得:

$$u^c \cdot A = u^c \cdot A^a + u^c \cdot A^m \approx \hat{u}^c \cdot A^a + \bar{e}_{x^a} + \hat{u}^c \cdot A^m + \bar{e}_{x^m} = \hat{u}^c \cdot A + \bar{e}_{x^c} \quad (9)$$

式中,  $\bar{e}_{x^c} = \bar{e}_{x^a} + \bar{e}_{x^m}$ ,  $\bar{e}_{x^m} = \frac{1}{k} \sum_{j=1}^k \left( \left( \overbrace{0, \dots, 0}^l, \left( x_{(j)}^c [l+1:s]^T \right) \right) - u_{(j)}^c \cdot A^m \right)_{x_{(j)}^c [l+1:s] \in N_k(x^c [l+1:s])}$ .

### 3 多元线性回归补全

如前所述,由公式(2)确定的优化模型实际上是最小化两个视角数据在共享子空间的总体重构误差,并由  $\beta$  平衡两个视角的重构误差.不难看出,所获得的最优共享表征  $u^c$  对于单一视角的重构并不是最优的.显然,式(2)中的  $X^T$  与  $U \cdot A$  仍呈一定的线性关系,因此可进一步通过回归实现缺失视角的精确补全:

$$\min_W g(W) = \|X^T - [U \cdot A \ \mathbf{1}] \cdot W\|_F^2 + \lambda \|W\|_F^2,$$

其中,  $\mathbf{1} = [1, 1, \dots, 1]^T \in R^{n \times 1}$  为全 1 列向量,  $\lambda$  为正则参数,  $W \in R^{(s+1) \times s}$  为回归系数,则联合式(9)可进一步实现缺失数据的精确补全:

$$\left( \hat{x}^m \right)^T = \left[ \hat{u}^c \cdot A + \bar{e}_{x^c} \ \mathbf{1} \right] \cdot W^m \quad (10)$$

其中,  $W^m = \left[ \overbrace{0, \dots, 0}^l, W[:, l+1:s] \right] \in R^{(s+1) \times s}$ .

### 4 多视角噪声数据的降噪

对于多视角数据  $\{x^c, y\}$ , 在其采集过程中,除了由于视角缺失问题以外,噪声也是造成多视角数据低质的重要原因.而从某种意义上来说,噪声数据也可看作是一种视角缺失的数据(即视角真实值的缺失).为此,本文提出的视角缺失补全思想可扩展到解决含有噪声的多视角数据的降噪问题.假设多视角数据  $\{x^c, y\}$  的  $x^c$  为含有噪声的视角,为实现对  $x^c$  的去噪,首先可根据公式(4)获得多视角噪声数据  $\{x^c, y\}$  的共享表征  $u^c$ ,并可进一步直接利用回归方法获得降噪后的  $\hat{x}^c$ :

$$\left( \hat{x}^c \right)^T = \left[ u^c \cdot A \ \mathbf{1} \right] \cdot W \quad (11)$$

## 5 实验结果与分析

### 5.1 数据集说明及评价标准

#### 5.1.1 数据集

A.UCI 多特征手写体数据集<sup>[18]</sup>.

该数据集包含 0~9 共 10 个数字的多特征数据集,每类有 200 个样本,对每个二值化手写体数字图像抽取 6 个视角特征见表 1.

**Table 1** UCI hand-written dataset with multiple features

**表 1** UCI 多特征手写数据集

mfeat_fou	mfeat_fac	mfeat_kar	mfeat_pix	mfeat_zer	mfeat_mor
76 维	216 维	64 维	240 维	47 维	6 维

**B.COIL-20 数据集<sup>[19]</sup>.**

该数据集共包括 20 个如图 4 所示的物体.对于每个物体,每隔 5°进行拍照,共获得 72 张图像.在本实验中,每幅图像下采样至大小为 64×64.为构造多视角数据集,对每个物体的 72 张图像,以间隔 30°为基准,构造 36 对多视角数据集.如图 5 所示,{0°,30°}组成一组多视角数据对,以此类推,本文用 COIL-20-30°表示该数据集.此外,基于同样方法,本文还另外构造了 COIL-20-60°多视角数据集.



Fig.4 COIL-20 dataset  
图 4 COIL-20 数据集

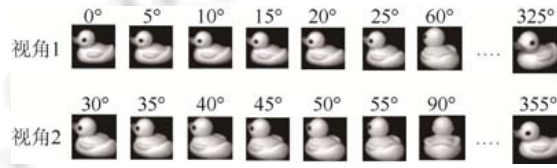


Fig.5 COIL-20-30° multiview dataset  
图 5 COIL-20-30°多视角数据集

**C.人工合成数据集.**

在上述数据集的基础上,本文还在一个人工合成的多视角数据集上进行了验证.该数据集包括 10 类,每类 200 个样本,记为  $\{X_i, Y_i\}_{i=1, \dots, 10}$ , 并有  $X_i^T = U^T \cdot A_i + \delta + \varepsilon_{i,x}, Y_i^T = U_i^T \cdot B_i + \eta + \varepsilon_{i,y}$ , 其中,  $X_i \in R^{80 \times 200}, Y_i \in R^{100 \times 200}, U_i \in R^{50 \times 200}$  为由随机函数生成的潜在共享表征矩阵.  $A_i \in R^{50 \times 80}, B_i \in R^{50 \times 100}$  为随机生成的载入矩阵,  $\delta, \eta$  为数据分布偏移常量矩阵,  $\varepsilon_{i,x} \rightarrow G(0, I_\varepsilon), \varepsilon_{i,y} \rightarrow G(0, I_\varepsilon)$  为随机扰动矩阵.重复上述过程,本文构造了 5 组人工合成的多视角数据集,以此验证本文算法的鲁棒性.

**5.1.2 评价标准**

本文采用标准化均方根误差 NRMSE<sup>[20,21]</sup>作为补全数据的评价标准:

$$NRMSE = \sqrt{\frac{\sum_{d=1}^N \frac{(x[d] - \hat{x}[d])^2}{(x^{\max}[d] - x^{\min}[d])^2}}{N}} \times 100\%,$$

其中,  $x[d]$  表示第  $d$  维变量的真实值,  $\hat{x}[d]$  表示  $x[d]$  的估计值,  $x^{\max}[d], x^{\min}[d]$  分别表示第  $d$  维变量的最大值和最小值,  $N$  为缺失属性的个数.此外,本文针对图像数据还采用了峰值信噪比(PSNR)作为图像补全、图像降噪增强的客观评价标准.

**5.2 实验结果分析**

**5.2.1 基于视角相容性的多视角数据类属判别**

为了验证基于视角相容性的多视角数据类属判别性能,本文选用 UCI 多特征手写数据集进行实验.针对该数据集,任选两个视角数据集分别作为  $X$  和  $Y$ ,这样,共有  $C_6^2 = 15$  种组合方式.对于每种组合,从每类样本中随机选择 90% 构成训练集,其余 10% 作为测试样本.

图 6 所示为组合(zer, kar)的视角相容性混淆矩阵  $C = [C_{i,j}]$ . 图中的  $C_{i,j}$  表示第  $j$  类测试样本在第  $i$  类共享子空间中的相容性平均值( $\times 10^3$ ), 即  $C_{i,j} = \sum_j f_i(u_i^c | A_i, B_i, x^c, y)_{\{x^c, y\} \in \{X_j, Y_j\}}$ , 颜色越深, 表示相容性越强, 反之则越弱. 不难看出, 基于视角相容性, 可以对多视角数据类属进行很好的判别.

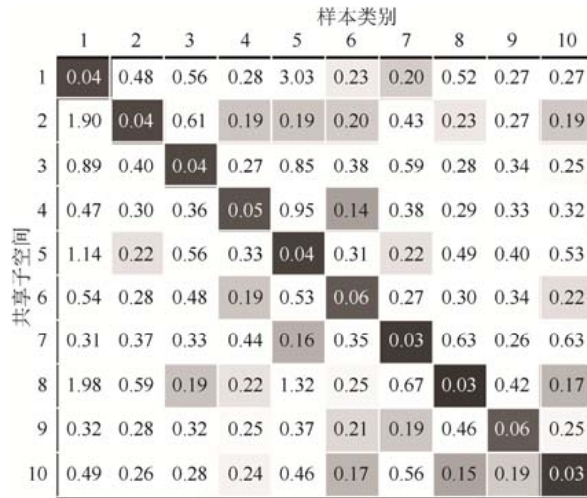


Fig.6 View compatibility based confusion matrix of group (zer, kar)

图 6 UCI 多特征手写体数据集中组合(zer, kar)的视角相容性混淆矩阵  $C = [C_{i,j}]$

在 UCI 多特征数据集上不同算法的分类性能比较见表 2. 表 2 中的  $X\_knn$ 、 $Y\_knn$ 、 $XY\_knn$  分别表示基于  $X$  视角、 $Y$  视角以及  $X$  视角和  $Y$  视角的串联利用最近邻分类器的分类性能,  $U\_knn$  为基于共享表征的最近邻分类性能, 而  $VCMC$  表示本文提出的基于视角相容性(view compatibility based multi-view classification, 简称  $VCMC$ ) 进行分类的分类性能, 每种算法的参数均由算法性能取得最优时的值来确定. 从表中可以看出, 与其他方法相比, 本文提出的  $VCMC$  方法表现出了明显的优势. 这也间接表明, 多视角数据间具有较强的共享信息, 利用视角相容性所构造的判据可以有效地提高分类性能. 同时, 本文算法也克服了最近邻分类器的度量选取问题.

Table 2 Classification comparisons of different methods on UCI dataset

表 2 不同算法在 UCI 数据集上的分类性能对比

视角 X	视角 Y	$X\_knn$ (%)	$Y\_knn$ (%)	$XY\_knn$ (%)	$U\_knn$ (%)	$VCMC$ (%)
fac	fou	96.90	79.90	<b>98.25</b>	96.40	98.00
fac	kar	96.90	96.20	97.45	96.45	<b>98.20</b>
fac	mor	96.90	66.60	<b>97.65</b>	84.75	95.70
fac	pix	96.90	97.90	97.85	94.35	<b>97.90</b>
fac	zer	96.90	79.00	97.55	97.35	<b>98.35</b>
fou	kar	79.90	96.20	96.65	92.65	<b>96.65</b>
fou	mor	79.90	66.60	81.90	82.20	<b>85.35</b>
fou	pix	79.90	97.90	<b>98.40</b>	94.35	97.90
fou	zer	79.90	79.00	83.90	81.95	<b>85.65</b>
kar	mor	96.20	66.60	96.90	91.05	<b>97.35</b>
kar	pix	96.20	97.90	97.70	95.05	<b>98.05</b>
kar	zer	96.20	79.00	96.90	96.75	<b>97.60</b>
mor	pix	66.60	97.90	<b>98.30</b>	92.65	97.40
mor	zer	66.60	79.00	82.40	73.65	<b>83.25</b>
pix	zer	97.90	79.00	97.80	97.20	<b>98.15</b>

5.2.2 多视角缺失数据补全性能分析

为了验证基于视角相容性的多视角数据缺失补全性能, 本文分别在 UCI 多特征手写数据集、COIL-20 数据集和人工合成数据集上进行了实验.

A. UCI 数据集上的视角缺失补全性能分析.

为验证补全性能, 针对该数据集的每种组合, 测试样本的缺失视角缺失部分的比例分别设为  $(s-l)/s=20\%$ ,



40%.表 3 给出了不同算法在缺失比例 $(s-l)/s=20\%$ 与 $(s-l)/s=40\%$ 下针对 UCI 数据的补全性能对比.表中的 KNN 表示利用多视角数据视角间的近邻关系的传递性,通过使用简单的“近邻”样本均值实现补全的性能;LRC 表示直接建立视角间属性回归(linear regression completion,简称 LRC)模型实现补全的性能.可以看出,与其他算法相比,本文算法具有更好的补全性能,也表明利用共享子空间建立的视角相容性判别模型可显著提高视角缺失的补全性能.

**Table 3** Performance comparisons of different completion methods on UCI dataset with  $(s-l)/s=20\%$ , 40%

**表 3** 针对 UCI 数据集,不同算法在 $(s-l)/s=20\%$ ,40%下的补全性能对比(NRMSE)

视角 $x^a$	视角 $y$	$(s-l)/s=20\%$			$(s-l)/s=40\%$		
		KNN (%)	LRC (%)	本文算法(%)	KNN (%)	LRC (%)	本文算法(%)
Kar	fac	9.97	10.84	<b>8.74</b>	10.12	11.04	<b>8.92</b>
Kar	zer	11.91	10.83	<b>8.94</b>	12.06	11.06	<b>9.96</b>
Kar	fou	13.22	12.91	<b>9.56</b>	13.38	12.22	<b>11.15</b>
Kar	mor	13.96	11.99	<b>10.64</b>	14.12	12.84	<b>11.01</b>
Kar	pix	9.35	10.59	<b>6.66</b>	9.49	10.84	<b>6.83</b>
Fac	zer	12.00	8.81	<b>5.89</b>	12.05	8.85	<b>6.17</b>
Fac	fou	14.24	11.24	<b>5.29</b>	14.29	11.31	<b>5.72</b>
Fac	mor	15.13	12.21	<b>8.33</b>	15.18	12.27	<b>8.42</b>
Fac	pix	7.82	8.60	<b>3.97</b>	7.87	8.65	<b>4.17</b>
Fac	kar	8.46	7.10	<b>5.43</b>	8.50	7.17	<b>5.52</b>
Zer	fac	8.80	8.60	<b>5.18</b>	8.94	15.67	<b>6.55</b>
Zer	fou	12.43	11.54	<b>5.11</b>	12.63	11.75	<b>7.44</b>
Zer	mor	12.81	12.68	<b>9.02</b>	13.01	13.32	<b>9.81</b>
Zer	pix	8.30	11.97	<b>4.14</b>	8.41	12.23	<b>5.01</b>
Zer	kar	8.86	8.60	<b>5.18</b>	8.98	8.73	<b>6.76</b>
Fou	fac	13.56	13.55	<b>13.40</b>	13.67	19.94	<b>13.50</b>
Fou	zer	13.85	13.47	<b>12.69</b>	13.98	14.25	<b>12.87</b>
Fou	mor	14.90	14.21	<b>13.12</b>	15.02	14.31	<b>13.27</b>
Fou	pix	13.56	14.19	<b>13.43</b>	13.67	15.07	<b>13.53</b>
Fou	kar	13.82	14.44	<b>13.38</b>	13.94	14.11	<b>13.56</b>
Pix	fac	22.33	19.54	<b>18.55</b>	22.39	21.88	<b>18.68</b>
Pix	zer	29.42	24.23	<b>17.72</b>	29.47	24.35	<b>18.25</b>
Pix	fou	33.67	32.42	<b>17.20</b>	33.69	32.16	<b>20.21</b>
Pix	mor	36.06	31.20	<b>22.86</b>	36.06	31.24	<b>23.04</b>
Pix	kar	21.92	16.05	<b>15.68</b>	21.96	<b>16.09</b>	16.25

**B.COIL-20 数据集上的视角缺失补全性能分析.**

在 COIL-20-30°及 COIL-20-60°数据集上,不同算法在缺失视角缺失部分的比例分别为 $(s-l)/s=20\%$ ,60%的补全性能对比见表 4.通过本实验可进一步说明利用多视角数据间的相容互补性有助于视角缺失数据的补全.此外,为从视觉上更为直观地观察视角缺失的补全效果,图 7 给出了本文算法和 KNN 在 COIL-20-30°数据集上针对视角完全缺失( $(s-l)/s=100\%$ )情况下进行补全后的结果.图中的第 1 行为基准图像(即完全缺失的视角),第 2 行与第 3 行分别为基于本文算法与 KNN 方法进行补全后的结果.

**C.人工数据集上的视角缺失补全性能分析.**

为了验证本文算法的有效性,本文还在人工合成的多视角数据集上进行了实验.在人工合成数据集上对于 $(s-l)/s=20\%$ ,60%的情况,不同算法的补全性能对比见表 5.通过在人工合成的数据集上进行实验,充分验证了本文提出的补全算法的有效性.

**Table 4** Performance comparisons of different completion methods on COIL-20 dataset with  $(s-l)/s=20\%$ , 60%

**表 4** 针对 COIL-20 数据集,不同算法在 $(s-l)/s=20\%$ ,60%下的补全性能对比(NRMSE)

多视角数据集	$(s-l)/s$	KNN (%)	LRC (%)	本文算法(%)
COIL-20-30°	20%	8.61	44.18	<b>6.23</b>
	60%	8.64	44.23	<b>6.38</b>
COIL-20-60°	20%	7.92	56.05	<b>5.72</b>
	60%	7.93	56.22	<b>5.85</b>



Fig.7 Performance comparisons of different completion methods on COIL-20-30° dataset with  $(s-l)/s=100\%$

图7 对于 COIL-20-30°数据集,不同算法在 $(s-l)/s=100\%$ 情况下的补全性能的对比情况

**Table 5** Performance comparisons of different completion methods on synthetic dataset with  $(s-l)/s=20\%$ ,  $60\%$

表5 针对人工合成数据,不同算法在在 $(s-l)/s=20\%$ , $60\%$ 下的补全性能比较(NRMSE)

$(s-l)/s$	KNN (%)	LRC (%)	本文算法 (%)
20%	1.95	0.63	<b>0.14</b>
60%	1.97	0.64	<b>0.18</b>

从上述针对 UCI 数据集、COIL-20 数据集以及人工合成数据集视角缺失补全的实验可以看出,本文算法明显优于其他算法的补全性能.由于多视角数据间的异构性,使得多视角数据间的  $k$  近邻不具有可逆性,以及视角间不再存在简单的线性映射关系,导致基于 KNN 和 LRC 补全方法的性能较差.这也表明,本文算法引入共享子空间建立视角相容性判据是必要的.同时,也间接表明,本文算法更好地反映了多视角数据间具有较强的相容互补性.

### 5.2.3 含有噪声的多视角数据的去噪性能分析

为验证第 4 节中提出的基于视角缺失补全思想的含有噪声多视角数据的降噪性能,本实验针对 COIL-20-30°以及 COIL-20-60°两个多视角数据集,从每类样本中随机选择 90% 构成无噪声训练集用于训练由公式(3)给出的多视角相容性模型,而剩余的 10% 作为测试集.此外,对于测试集中的  $X$  视角数据(图像)添加幅度为 0.02 的高斯噪声,以形成含有噪声的多视角数据.表 6 给出了在 COIL-20-30°以及 COIL-20-60°两个多视角数据集上的降噪性能比较.从表中可以看出,本文算法具有很好的去噪能力.

为观察噪声图像经过去噪后的视觉效果,图 8 给出了不同算法在 COIL-20-30°数据集上的去噪性能对比情况.从图中可以直观地看出本文算法在图像增强方面的优势.在本实验中,中值滤波、高斯滤波均是基于邻域像素值处理的去噪方法,而本文算法通过利用视角间相容互补性更好地增强了图像质量.这也进一步表明,多视角数据间的相容互补性有助于提高图像质量.

**Table 6** PSNR (dB) based denoising performance comparisons of different completion methods on COIL-20 dataset  
表6 不同算法在 COIL-20 数据集上去噪性能对比 PSNR(dB)

多视角数据集	中值滤波	高斯滤波	本文算法
COIL-20-30°	23.92	21.46	<b>26.39</b>
COIL-20-60°	23.93	21.46	<b>26.78</b>

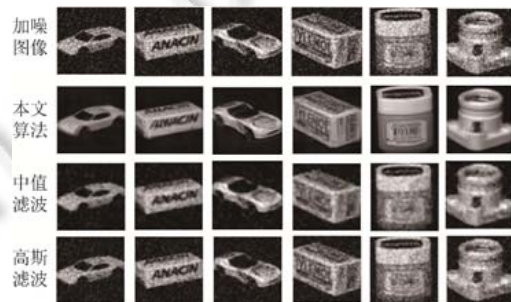


Fig.8 De-Noising performance comparisons of different completion methods on COIL-20-30° dataset

图8 不同算法在 COIL-20-30°数据集上的去噪性能对比

## 6 结束语

针对多视角数据缺失问题,本文提出一种基于视角相容性的补全方法,该方法通过共享子空间的监督学习,建立视角相容性判别模型.基于共享子空间重构误差等同分布假设,提出了针对视角缺失的多视角数据的共享表征方法,在此基础上,通过进一步的多元线性回归,实现多视角数据缺失的精确补全.

### References:

- [1] Xu C, Tao D, Xu C. A survey on multi-view learning. arXiv:1304.5634v1, 2013.
- [2] Sun S. A survey of multi-view machine learning. *Neural Computing and Applications*, 2013,23(7-8):2031–2038. [doi: 10.1007/s00521-013-1362-6]
- [3] Huang QM, Chen XL. From multimedia to cross-media. *Communications of China Computer Federation*, 2014,10(7):6–7 (in Chinese with English abstract).
- [4] Wei YC, Zhao Y, Lu CY, Wei SK, Liu LQ, Zhu ZF, Yan SC. Cross-Modal retrieval with CNN visual features: A new baseline. *IEEE Trans. on Cybernetics*, 2017,47(2):449–460. [doi: 10.1109/TCYB.2016.2519449]
- [5] Zhu X. Editorial: Special issue on mining low-quality data. *Knowledge and Information Systems*, 2007,11(2):131–136. [doi: 10.1007/s10115-006-0058-y]
- [6] Hardoon DR, Szedmak S, Shawe-Taylor J. Canonical correlation analysis: An overview with application to learning methods. *Neural Computation*, 2004,16(12):2639–2664. [doi: 10.1162/0899766042321814]
- [7] Blum A, Mitchell T. Combining labeled and unlabeled data with co-training. In: *Proc. of the 11th Annual Conf. on Computational Learning Theory*. 1998. 92–100. [doi: 10.1145/279943.279962]
- [8] Lanckriet GRG, Cristianini N, Bartlett P. Learning the kernel matrix with semidefinite programming. *Journal of Machine Learning Research*, 2004,5(1):27–72.
- [9] Little RJA, Rubin DB. *Statistical Analysis with Missing Data*. New York: John Wiley & Sons, 2014.
- [10] Ghahramani Z, Jordan MI. Learning from incomplete data. 1997. [https://www.researchgate.net/publication/2353392\\_Learning\\_From\\_Incomplete\\_Data](https://www.researchgate.net/publication/2353392_Learning_From_Incomplete_Data)
- [11] Candès EJ, Recht B. Exact matrix completion via convex optimization. *Foundations of Computational Mathematics*, 2009,9(6):717–772. [doi: 10.1007/s10208-009-9045-5]
- [12] Cai JF, Candès EJ, Shen Z. A singular value thresholding algorithm for matrix completion. *SIAM Journal on Optimization*, 2010,20(4):1956–1982. [doi: 10.1137/080738970]
- [13] Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Methodological (Series B)*, 1977,39(1):1–38.
- [14] Van de Velden M, Takane Y. Generalized canonical correlation analysis with missing values. *Computational Statistics*, 2012,27(3):551–571. [doi: 10.1007/s00180-011-0276-y]
- [15] Li RJ, Zhang WL, Suk HI, Wang L, Shen DG, Ji SW. Deep Learning Based Imaging Data Completion for Improved Brain Disease Diagnosis. Springer Int'l Publishing, 2014. 305–312. [doi: 10.1007/978-3-319-10443-0\_39]
- [16] Yu SP, Yu K, Tresp V. Multi-Output regularized feature projection. *IEEE Trans. on Knowledge and Data Engineering*, 2006,18(12):1600–1613. [doi: 10.1109/TKDE.2006.194]
- [17] Du LL, Zhu ZF, Duan HS. LSPSA: Local structure preserves shared-subspace analysis. *Computer Science*, 2014,41(10):67–71 (in Chinese with English abstract). [doi: 10.11896/j.issn.1002-137X.2014.10.015]
- [18] Xu L, Krzyzak A, Suen CY. Methods of combining multiple classifiers and their applications to handwriting recognition. *IEEE Trans. on Cybernetics*, 1992,22(3):418–435.
- [19] Nene SA, Nayar SK, Murase H. Columbia object image library (COIL-20). Technical Report, CUCS-005-96, Columbia University, 1996.
- [20] Troyanskaya O, Cantor M, Sherlock G, Brown P, Hastie T, Tibshiran R, Botstein D, Altman RB. Missing value estimation methods for DNA microarrays. *Bioinformatics*, 2001,17(6):520–525. [doi: 10.1093/bioinformatics/17.6.520]

- [21] Oba S, Sato M, Takemasa I. A Bayesian missing value estimation method for gene expression profile data. *Bioinformatics*, 2003,19(16):2088–2096. [doi: 10.1093/bioinformatics/btg287]

附中文参考文献:

- [3] 黄庆明,陈熙霖.从多媒体到跨媒体.计算机学会通讯,2014,10(7):6–7.
- [17] 杜琳琳,朱振峰,段红帅.LSPSA:基于局部结构保持的共享子空间分析.计算机学报,2014,41(10):67–71. [doi: 10.11896/j.issn.1002-137X.2014.10.015]



杨旭(1990—),男,山东青岛人,硕士,主要研究领域为机器学习.



徐美香(1988—),女,硕士,主要研究领域为机器学习.



朱振峰(1974—),男,博士,教授,博士生导师,CCF 专业会员,主要研究领域为图像、视频分析与理解,机器学习,网络信息挖掘与分析.



张幸幸(1993—),女,学士,主要研究领域为机器学习.