多模态张量数据挖掘算法及应用

杨琬琪[1,2] 高 阳[1,2] 周新民[3] 杨育彬[1,2] 商 琳[1,2]
(南京大学软件新技术国家重点实验室 南京 210093)[1] (南京大学江阴信息技术研究院 无锡 214433)[2]
(江苏省公安厅物证鉴定中心 南京 210024)[3]

摘 要 近年来,多模态数据挖掘技术备受关注,如何高效地挖掘大量多模态数据成为一个研究热点。其中,基于张量表示的多模态数据挖掘,即多模态张量数据挖掘,是一个重要的研究问题。综述了多模态张量数据挖掘算法进展及其在计算机视觉中的应用。首先根据算法的样本标记、任务和核心技术的不同,对这些方法进行分类,并给出了相应的介绍和分析。其次,讨论了一些多模态张量数据挖掘算法在计算机视觉问题中的典型应用。最后,就多模态张量挖掘在计算机视觉领域的研究现状与研究前景进行了简要的分析。

关键词 多模态张量,数据挖掘,张量表示,计算机视觉
中图法分类号 TP391.41 **文献标识码** A

Multi-modal Tensor Data Mining Algorithms and Applications

YANG Wan-qi[1,2] GAO Yang[1,2] ZHOU Xin-min[3] YANG Yu-bin[1,2] SHANG Lin[1,2]
(State Key Laboratory for Novel Software and Technology, Nanjing University, Nanjing 210093, China)[1]
(Jiangyin Institute of Information Technology, Nanjing University, Wuxi 214433, China)[2]
(Jiangsu Province Public Security Department, Nanjing 210024, China)[3]

Abstract Multi-modal data mining technologies have attracted many research interests in recent years. Mining large amount of multi-modal data efficiently becomes a hot spot problem. Among these multi-modal mining technologies, multi-modal data mining for tensor representation, which is also called as multi-modal tensor data mining, is one of the most significant research issues. We reviewed the state-of-the-art algorithms of the multi-modal tensor data mining and their applications in computer vision. Firstly, multi-modal tensor data mining algorithms were categorized into different classes according to the different label information, task and core technology. In addition, some analyses about these algorithms were given. Secondly, some typical multi-modal tensor mining algorithms in computer vision application were illustrated. Finally, we presented our own analyses on research status of multi-modal tensor mining algorithms, and explored some potential future issues of multi-modal tensor mining in computer vision application.

Keywords Multi-modal tensor, Data mining, Tensor representation, Computer vision

1 引言

多模态数据(视频、音频、图像等)的挖掘是机器视觉、模式识别、人工智能等领域的重要技术之一。其中,多模态数据的表示是多模态数据挖掘的核心问题。目前,多模态数据表示分为向量表示与张量表示两类。不同于成熟的基于向量表示的多模态数据挖掘方法[1],近年来基于张量表示的多模态数据挖掘方法得到了长足的发展。与向量表示相比较,张量表示的特点有:

1)张量结构直观上易于表示多模态数据,一阶表示一个模态,不同的模态维度不同,而向量表示往往忽视这一点;

2)在提取特征(尤其是图像特征)时,张量可以有效保存原始样本的结构特征;

3)在数据变换时,张量可以充分考虑不同模态之间的相关性与互补性,向量表示很难挖掘上述关系;

4)张量表示可以有效解决向量表示的疑难问题,如维度灾难、矩阵奇异等。

本文综述了当前基于张量表示的多模态数据挖掘方法。首先介绍张量代数的基础知识,其次对基于张量表示的多模态数据挖掘方法进行分类和描述,最后讨论基于张量表示的多模态数据挖掘算法在计算机视觉上的应用。

2 张量代数

张量是多向阵列,可看作向量、矩阵的高阶扩展。一个 N 阶张量 $A\in R^{I_1\times I_2\times\cdots\times I_N}$ ($I_1\times I_2\times\cdots\times I_N$ 为笛卡尔乘积),第 n 阶的维度大小是 I_n。标量 $a\in R$ 是0阶张量,矢量 $\boldsymbol{a}\in R^{I_1}$ 是1阶张量,矩阵 $A\in R^{I_1\times I_2}$ 是2阶张量。基于张量表示,可以进一步定义张量矩阵展开、模乘(Mode product)、张量积(Tensor product)和张量分解。

1)张量矩阵展开

到稿日期:2011-02-23 返修日期:2011-05-19 本文受江苏省社会发展项目(BE2010638)资助。
杨琬琪(1988—),女,硕士生,主要研究方向为多模态数据挖掘与社会网络分析,E-mail:nju.yangwanqi@gmail.com。

张量矩阵展开($N>2$)是将张量中的元素重新排列构成矩阵的过程。对张量 $\boldsymbol{A}\in R^{I_1\times I_2\times\cdots\times I_N}$ 进行第 n 阶矩阵展开后,形成二阶张量展开矩阵 $\boldsymbol{A}_{(n)}\in R^{I_n\times\overline{I_n}}$,其中 $\overline{I_n}$ 是其他阶维度大小的乘积,即 $\overline{I_n}=I_1\cdots I_{n-1}I_{n+1}\cdots I_N$。以 3 阶张量 $\boldsymbol{A}\in R^{I_1\times I_2\times I_3}$ 为例,其张量矩阵展开过程如图 1 所示。

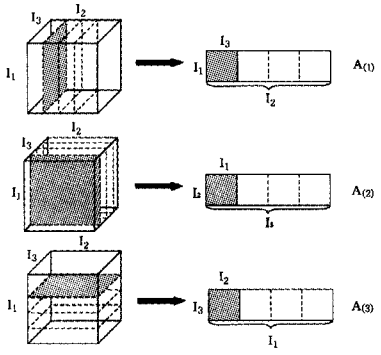

图 1 张量矩阵展开示意图

2)模乘

模乘定义张量与矩阵的乘法。张量 $\boldsymbol{A}\in R^{I_1\times I_2\times\cdots\times I_N}$ 与矩阵 $U\in R^{J_n\times I_n}$ 的 n 模乘积表示为 $\boldsymbol{A}\times_n U$,通过模乘运算将 $\boldsymbol{A}$ 的第 n 阶维度 I_n 用 U 的第 1 阶维度 J_n 代替,因而 n 模乘积 $\boldsymbol{A}\times_n U$ 是 $I_1\times I_2\times\cdots\times I_{n-1}\times J_n\times I_{n+1}\times\cdots\times I_N$ 阶张量,具体的运算定义为:

$$(\boldsymbol{A}\times_n U)_{i_1i_2\cdots i_{n-1}j_ni_{n+1}\cdots i_N}=\sum_{i_n}a_{i_1i_2\cdots i_{n-1}i_ni_{n+1}\cdots i_N}u_{j_ni_n} \tag{1}$$

另外,n 模乘积 $\boldsymbol{B}=\boldsymbol{A}\times_n U$ 可以由矩阵乘法 $\boldsymbol{B}_{(n)}=U\boldsymbol{A}_{(n)}$ 得到。

3)张量积

张量积是对多个张量的不同阶数进行笛卡尔积,形成新的高阶张量,张量积用 $\otimes$ 表示。以人脸特征张量 $\boldsymbol{A}\in R^{I_1\times I_2\times\cdots\times I_P}$ 和外在因素张量 $\boldsymbol{B}\in R^{J_1\times J_2\times\cdots\times J_Q}$ 为例,其张量积 $\boldsymbol{A}\otimes\boldsymbol{B}$ 是一个 $I_1\times I_2\times\cdots\times I_P\times J_1\times J_2\times\cdots\times J_Q$ 阶的张量,具体运算定义为:

$$(\boldsymbol{A}\otimes\boldsymbol{B})_{i_1\times i_2\times\cdots\times i_P\times j_1\times j_2\times\cdots\times j_Q}=a_{i_1\times i_2\times\cdots\times i_P}b_{j_1\times j_2\times\cdots\times j_Q} \tag{2}$$

因而高阶张量可由两个低阶张量的张量积得到。

4)张量分解

Tucker 分解是 1966 年最早提出的一种模型[4],对 3 阶张量 $\boldsymbol{A}\in R^{I\times J\times K}$ 使用 3 阶因子分析。分解公式为:

$$\boldsymbol{A}=\boldsymbol{G}\times_1U\times_2V\times_3W \tag{3}$$

式中,$U\in R^{m\times r_1}$,$V\in R^{n\times r_2}$,$W\in R^{p\times r_3}$ 是张量 $\boldsymbol{A}$ 的成分矩阵,$\boldsymbol{G}\in R^{r_1\times r_2\times r_3}$ 为核心张量。

文献[5]中 L. De Lathauwer 首次提出 HOSVD 的概念,对 Tucker 分解的成分矩阵附加正交限制,是矩阵 SVD 的多线性泛化。HOSVD 将张量 $\boldsymbol{A}\in R^{I_1\times I_2\times\cdots\times I_N}$ 分解为 $\boldsymbol{A}=\boldsymbol{T}\times_1U_1\times\cdots\times_NU_N$,简写成 $\boldsymbol{A}=\boldsymbol{T}\prod_{n=1}^{N}\times_nU_n$。其中,成分矩阵 U_n 是 $I_n\times I_n$ 阶正交阵,$\boldsymbol{T}$ 为核心张量。

3 多模态张量数据挖掘算法分类

将张量的每一阶看作是不同的模态,多模态张量数据作为样本数据,则张量分析(多维数组分析)[24]、多维子空间学习[11]等数据挖掘方法很容易推广到多模态数据挖掘任务中。我们从算法的样本性质、任务与核心技术 3 个方面,将基于张量表示的多模态数据挖掘算法按表 1 进行分类。下面针对每一类的具体算法作详细介绍。

表 1 多模态数据挖掘算法分类

样本性质	任务	核心技术	具体算法
训练样本有标记	预测分类	判别式分析	2DLDA, 2DFDA, GTDA, TFDA, MLDA, CTA, DTROD, UMLDA
		支持张量机	STM, TSTM, MSTMs, RMSTMs
		生成式分析	TMPM, BTA
训练样本无标记	聚类、降维	PCA系列	2DPCA, GPCA, MPCA, RMPCA, NMPCA, UMPCA
		ICA系列	MICA, DICA
		张量秩近似	TROD, GLAM, CSA

3.1 多模态张量的预测分类算法

3.1.1 判别式分析算法

经典的 Fisher 线性判别式算法思想是将高维数据投影到低维子空间,通过对投影后的样本最大化类间距离和最小化类内距离,使不同类别的样本最佳可分离。基于上述思想的多模态张量数据挖掘算法见表 2。

表 2 判别式分析算法

特点	相关算法
一阶线性变换	2DFDA[13]
二阶线性变换	2DLDA[10]
高阶线性变换	GTDA[14]、TFDA[15]、MLDA[16]、CTA[9]
相关性约束	UMLDA[17]
张量分解	DTROD[6]

表 2 中,2DFDA 算法与 2DLDA 算法比较简单,都是针对矩阵表示的样本。具体地说,前者对矩阵的行模式进行投影,列模式不变,属于一阶线性变换;后者对矩阵样本的行、列模式均进行投影,属于二阶线性变换。

对于高阶张量来说,算法比较复杂,需要进行高阶线性变换,即对张量样本 $\boldsymbol{X}$ 的每一阶均进行模乘,变换公式为:

$$\boldsymbol{Z}=\boldsymbol{X}\prod_{n=1}^{N}\times_nU_n^T \tag{4}$$

式中,U_n 是第 n 阶线性变换矩阵。具体的算法有 MLDA、GTDA、TFDA 和 CTA 算法。其中,前 3 个算法在判别式分析中具有代表性,MLDA 算法是 2DLDA 算法的张量扩展;GTDA 算法使用微分散度判别准则(DSDC),优点是解决了 2DLDA 和 MLDA 算法的优化过程不稳定的问题,保证结果收敛;TFDA 算法是在监督张量学习框架[15](STL)中 GTDA 算法的具体实现;不同于前 3 个重要算法,最后 1 个 CTA 算法引入类相关图来保存原有数据分布的流形结构。

表 2 中,最后两个算法 UMLDA 与 DTROD 的线性变换不同于式(4),它们都是张量与向量的模乘,将张量样本 $\boldsymbol{X}$ 通过向量 $\boldsymbol{u}_n|_{n=1}^{N}$ 变换到标量 z,变换公式为:

$$z=\boldsymbol{X}\prod_{n=1}^{N}\times_n\boldsymbol{u}_n^T \tag{5}$$

这两个算法的区别是前者对变换进行相关性约束,将样本 $\boldsymbol{X}$ 进行 K 次不相关的线性变换,分别投影到 K 个标量特征上,构成 K 维向量,如图 2 所示,因为不相关的判别式特征能够保证数据的最小冗余性和特征间的独立性;而后者从张

量分解的角度考虑，通过上述变换将样本 $\boldsymbol{X}$ 分解为多个两向量乘积的带权和。

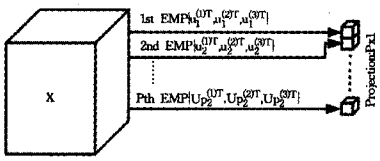

图 2　张量到向量的投影(EMP:初等多线性投影)

3.1.2　支持张量机算法

支持张量机算法[15]是指在监督张量学习的框架下训练已标记的张量样本，在样本空间寻找一个最优超平面，使分类间隔最大化，多用于两类预测。Soft-margin STM 算法优化问题为：

$$\min_{w_j,b,\xi_i} \frac{\eta}{2} \| w_j \|_F^2 + C\sum_{i=1}^{N} \xi_i$$

$$\text{s.t. } y_i(w_j^T(\boldsymbol{X}_i \overline{\times}_j w_j)+b) \geqslant 1-\xi_i, \xi_i \geqslant 0, 1 \leqslant i \leqslant N \quad (6)$$

式中，$\boldsymbol{X}_i|_{i=1}^N$ 表示张量样本，y_i 为类标记，分类超平面是 $\boldsymbol{X}_i \prod_{k=1}^{M} \times_k w_k + b = 0$。关于支持张量机的具体算法见表 3。

表 3　支持张量机算法

特点	相关算法
支持张量	STM[15,20]
半监督式	TSTM[21]
矩阵分解	MSTMs[22]
相对边界	RMSTMs[23]

表 3 中，TSTM、MSTMs 与 RMSTMs 算法均是基于 STM 算法的；具体来讲，TSTM 算法先对已标记样本训练 STM 分类超平面，再不断修改超平面两侧未标记样本的可能标记，在尽可能正确划分已标记样本的基础上，将所有样本(未标记和已标记)分类间隔最大化；MSTMs 将 STM 看作是矩阵分解的过程，对其使用乘法更新规则，收敛速度比标准 STM 算法快；RMSTMs 算法最大化分类间距时使用相对边界的限制，目的是解决 STM 算法忽视数据分布在不同投影空间变化很大的问题。

3.2　生成式分析算法

生成式分析算法一般使用概率模型描述多模态张量数据的真实分布。具体算法见表 4。

表 4　模型分布算法

特点	相关算法
MPM 的张量泛化	TMPM[15]
概率图模型	BTA[36]

表 4 中，TMPM 算法是针对样本的数据分布寻找一个超平面，最大概率地划分正、负样本，提高正确预测分类的概率；BTA 算法构建概率图模型，通过最大化预测概率密度计算投影矩阵，且自动确定投影矩阵的维度来选择合适的图模型。

3.3　多模态张量的降维算法

3.3.1　PCA 系列算法

主成分分析(PCA)算法思想是对复杂数据降维，将其投影到样本方差尽可能大的特征子空间，找出样本中最“主要”的元素和结构。PCA 系列的算法见表 5。

同 2DFDA 与 2DLDA 算法的关系类似，表 5 中 2DPCA 算法与 GPCA 算法都是针对于矩阵样本，前者只进行行方向的投影，列不变，后者进行二阶线性变换。

表 5　PCA 系列算法

特点	相关算法
一阶线性变换	2DPCA[25]
二阶线性变换	GPCA[27]
高阶线性变换	MPCA[28]、RMPCA[29]、NMPCA[30]、UMPCA[31]

表 5 中，MPCA 算法是 PCA 系列算法中处理高阶张量样本的最具代表性的算法。MPCA 算法对输入张量进行高阶变换，且在投影空间保持原有张量样本的变化性。另外，RMPCA 算法、NMPCA 算法与 UMPCA 算法是 MPCA 算法的变种。具体地说，RMPCA 算法是基于拉格朗日乘子的迭代算法，用于发现样本离异点和内部样本离群值；NMPCA 算法限制投影后的矩阵为非负值，保存原有样本的非负性；与 UMLDA 类似，UMPCA 算法提取 K 个不相关的多线性特征，将张量投影到向量，使样本方差最大化。

3.3.2　ICA 系列算法

独立主成分分析(Independent Component Analysis, ICA)算法是在输入源独立的假设下，挖掘复杂数据的独立成分。多模态张量的 ICA 系列算法见表 6。

表 6　ICA 系列算法

特点	相关算法
多线性 ICA	MICA[34]、DICA[37]

表 6 中，MICA 算法与 DICA 算法均是对高阶张量进行多线性 ICA 分析，而且后者是对前者的改进。其中，MICA 算法使用 n 阶 ICA 算法($\boldsymbol{X}=\boldsymbol{S}\prod_{n=1}^{N}\times_n C_n$)计算各模态矩阵 $C_n|_{n=1}^N$，保证它们之间的统计独立性。但在 n 阶矩阵展开时忽略了当前阶与其他阶的关系，因而影响了分类效果。为了解决这一问题，DICA 算法对图像样本 n 阶矩阵展开后，重新采样与组合，得到 n 阶方向性的图像，分类准确率有所提高。

3.3.3　张量秩近似算法

张量秩从矩阵秩推广而来，第 n 阶张量秩表示为 $R_n(\boldsymbol{X})$，是指张量第 n 阶展开矩阵的秩，即 $R_n(\boldsymbol{X})=R(\boldsymbol{X}_{(n)})$。张量 $rank-(r_1,\cdots,r_n,\cdots,r_N)$ 近似是指将一个张量近似分解为一个第 n 阶张量秩为 r_n 的张量或多个这种张量的组合($n=1,\cdots,N$)。多模态张量秩近似的算法见表 7。

表 7　张量秩近似算法

特点	相关算法
rank-one	TROD[32]
$rank-(r_1,r_2)$	GLRAM[26]
$rank-(r_1,\cdots,r_n,\cdots,r_N)$	CSA[33]

表 7 中，TROD 算法与 GLRAM 算法均是对矩阵样本 $\boldsymbol{X}$ 进行分解。然而前者分解成两个向量乘积的权重和，即 $\boldsymbol{X}=\sum\lambda_i u_i v_i^T$，向量的秩由于均为 1，因此属于 rank-one；后者将样本 $\boldsymbol{X}$ 近似为低秩矩阵 Z 的二阶线性变换，即 $\boldsymbol{X}=LZR^T$，矩阵 L 与 R 的秩由于分别为 r_1 和 r_2，因此属于 $rank-(r_1,r_2)$。从低阶到高阶演化，表 7 中的 CSA 算法是对高阶张量样本进行切片投影，$\boldsymbol{X}=\boldsymbol{Z}\prod_{n=1}^{N}\times_n U_n$，属于 $rank-(r_1,\cdots,r_n,\cdots,r_N)$，是广义上的 GLRAM。

4　多模态张量数据挖掘在计算机视觉上的应用

多模态张量数据挖掘在计算机视觉领域的应用有很多，

如图像分类[3]、图像分割[19]、3D 人脸建模[36]、目标跟踪[7]等。针对计算机视觉的应用研究者需要考虑两个方面，即数据表示形式与数据挖掘算法。第一，数据表示形式。对于图像，若只考虑特征的空间性，用矩阵表示，x 和 y 坐标分别为两个模态，可保存图像的结构特征；若考虑不同的时间、光照与视角等模态，可用张量表示图像样本，每一阶表示一个模态。对于视频数据，一般用三阶张量表示，x 坐标、y 坐标和视频帧，然而对于具体的应用会有所不同。第二，用于特征提取、预测分类等任务的数据挖掘算法。下面针对图像与视频压缩、人脸识别与步态识别、图像/视频检索 3 个应用作具体介绍。

图像/视频压缩是将数据样本压缩到低维子空间，使重构误差最小，可以通过降维、张量近似等方法来实现。对于图像矩阵，一般使用 2DPCA[25]、GLRAM[26]、2DSVD[38] 几种算法。对于视频序列，可使用张量 rank-one 近似[38]，也可使用张量 rank-R 近似[39]，即近似张量的每一阶展开矩阵的秩均为 R，这样对每一个模态都提取了 R 个成分，压缩原有样本。图像/视频压缩用于解决大量的图像/视频数据不便存储或处理的问题。

人脸识别和步态识别是目标识别中目前研究最热的两个应用。由于人们可以通过人脸与步态特征远距离地识别一个人，因此其多用于视频监控。人脸识别应用中，人脸数据可表示为 5 阶张量，分别是 x 坐标、y 坐标、光照、表情、视角和身份 5 个模态[12]。或者将人脸特征用向量表示，加入变化的外部条件，如颜色和红外线特征，共同构成多模态张量[18]。TensorFace[12]对这个多模态样本进行多线性分析。通过 HOSVD 张量分解得到模态矩阵，模态矩阵的列向量构成单模态的基向量，即特征脸，张量分解的核心张量可以很好地体现模态间的交互。步态指人走路的姿势，是具有复杂时空特征的视频序列，在人员自动识别系统中广泛应用。关于步态的识别可使用 STM[23]、GTDA[14]、MTP[2]来实现，并使用步态能量图 GEI 来衡量两个步态序列的距离。MTP 从视角、鞋型、公文包、时间等多个模态考虑，使用多线性张量投影，自动选择最优子空间维度，无参数调整。

由于用户对网络检索的高需求，图像/视频检索的应用非常广泛。视频包含图像、音频、文字，具有时序相关特性。为了能够高精度地检索，视频数据可表示为三阶张量，分别为图像、音频、文字这 3 个模态。视频语义检索应用中，文献[20，21]使用 STM、TSTM 算法进行分类预测，将降维与流形学习算法结合，寻找语义子空间，挖掘视频的内在语义信息。文献[8,20]使用 TensorShot 方法对张量数据降维，保存其流形结构。文献[8]对子空间的张量数据用信任力传播（Affinity Propagation）算法进行聚类。文献[35]提出多模态潜在语义索引（MMLSI）和多模态概率潜在语义分析（MMpLSA）方法，它融合了可视特征和文本等多模态信息，提高了检索的效率。文献[19]实现一个多媒体检索框架，即将图像特征与注释文本特征进行张量积，使之不可分割，然后遵照量子测量法计算查询与文档之间的相似度，给图像打分排序，实现检索过程。

5 讨论与展望

通过以上对多模态张量数据挖掘算法及其计算机视觉应用的分析，笔者认为，多模态张量数据挖掘未来的研究方向将体现在以下几个方面：

(1)研究对监督张量学习机（STL）算法的改进与扩展。对于 STM，考虑核化（kernelize）目标函数，将数据变换到非线性空间进行学习；可以改进交替投影优化方法[15]，解决其优化效率慢、有震荡的缺点。

(2)研究多个模态间的权重学习方法。张量表示多模态数据通常把每个模态看成是同等的一阶，这样在对以图像为主的视频（或以音频为主的语音）来说，多模态特征将会减弱其重要的描述作用，所以，可以通过添加权重来克服这一缺点。

(3)在张量子空间学习中，关于非线性子空间的张量数据挖掘算法很少[19]，缺乏系统的研究。但在实际应用中不同模态特征之间非线性关系的挖掘非常重要，因而可以着重研究张量非线性子空间的学习，如对 HOSVD 算法或 STM 算法使用核函数等。

(4)研究如何将原有的张量学习算法很好地用于多模态数据挖掘，是非常有意义的重要研究问题。这需要考虑多模态样本采集、多模态自身特性（如模态间的相关性与互补性）以及实际应用需要。

(5)目前张量表示的半监督学习算法很少[21]，可以考虑结合实际应用进行研究。

结束语 关于多模态数据挖掘的研究，相对于比较成熟的向量表示方法，张量表示多模态数据是近几年才出现的研究思路。多模态张量数据挖掘有很大的研究空间，可以扩展到很多多模态应用领域，如多目标跟踪、多行为检测、多特征生物识别、多模态人机交互、多模态 Web 网页挖掘等。

本文首先指出张量表示多模态数据的优点，介绍张量代数的基础，然后根据算法的样本特性、任务、核心技术等分类介绍了多模态张量数据挖掘算法，具体阐述了在计算机视觉下多模态张量数据挖掘的广泛应用。最后分析总结多模态张量数据挖掘算法，探讨其未来的研究方向。

参 考 文 献

[1] Wu Y, Chang E Y, Chang K, et al. ptimal Multimodal Fusion for Multimedia Data Analysis[J]. ACM Multimedia, 2004

[2] Chen C, Zhang J, Fleischer R. Multilinear Tensor-Based Nonparametric Dimension Reduction for Gait Recognition[J]. Advances In Biometrics, 2009, 5558: 1030-1039

[3] Liu Y, Liu Y, Chan K. Tensor-Based Locally Maximum Margin Classifier For Image And Video Classification[J]. Computer Vision and Image Understanding (CVIU), 2011, 115: 300-309

[4] Tucker L R. Some Mathematical Notes on Three-Mode Factor Analysis[J]. Psychometrika, 1966, 31: 279-311

[5] Lathauwer L D, Moor B D, Vandewalle J. A Multilinear Singular Value Decomposition[J]. SIAM Journal of Matrix Analysis and Applications, 2000, 21(4): 1253-1278

[6] Wang Y, Gong S. Tensor Discriminant Analysis for View-Based Object Recognition[C] // Proceedings of International Conference on Pattern Recognition, 2006, 3: 33-36

[7] Wu J, Trivedi M. An Eye Localization, Tracking and Blink Pat-

tern Recognition System: Algorithm and Evaluation[J]. ACM Trans Multimed. Comput. Commun. Appl. ,2010,6(2):1-23

[8] Liu Y,Wu F. Multi-Modality Video Shot Clustering with Tensor Representation[J]. Multim. Tools. Appl. ,2009,41(1):93-109

[9] Fu Y,Huang T. Image Classification Using Correlation Tensor Analysis[J]. IEEE Trans. Image Processing,2008,17(2):226-234

[10] Ye J, Janardan R, Li Q. Two-dimensional Linear Discriminant Analysis[J]. Advances in Neural Information Processing Systems,2005,17:1569-1576

[11] Lu H,Plataniotis K N,Venetsanopoulos A N. A Survey of Multilinear Subspace Learning for Tensor Data[J]. Pattern Recognition,2011,44(7):1540-1551

[12] Vasilescu M A O,Terzopoulos D. Multilinear Analysis of Image Ensembles[C] // Tensorfaces, Proc. Seventh European Conf. Computer Vision (ECCV 02'). 2002,2350:447-460

[13] Kong H, Teoh E, Wang J, et al. Two Dimensional Fisher Discriminant Analysis: Forget about Small Sample Size Problem[C]// Proc. IEEE Conf. Acoust. , Speech, Signal Process. Mar. 2005: 761-764

[14] Tao D,Li X,Wu X,et al. General Tensor Discriminant Analysis and Gabor Features for Gait Recognition[J]. IEEE Trans. Pattern Anal. Mach. Intell. ,2007,29:1700

[15] Tao D, Li X, Wu X, et al. Supervised Tensor Learning[C]// Proc. IEEE Conf. Data Mining (ICDM). 2005:450-457

[16] Yan S,Xu D,Yang Q,et al. Multilinear Discriminant Analysis for Face Recognition[J]. IEEE Trans. Image Processing,2007, 16:212

[17] Lu H, Plataniotis K N, Venetsanopoulos A N. Uncorrelated Multilinear Discriminant Analysis with Regularization and Aggregation for Tensor Object Recognition[J]. IEEE Trans. Neural Networks,2009,20(1):103-123

[18] Jia K, Gong S. Generalized Face Super-Resolution[J]. IEEE Transactions on Image Processing,2008,17:873-886

[19] Li Y,Du Y,Lin X. Kernel-Based Multifactor Analysis for Image Synthesis and Recognition[C]// Proc. Int'l Conf. Computer Vision (ICCV '05). 2005,1:114-119

[20] Liu Y N,Wu F,Zhuang Y T,et al. Active Post-refined Multi-Modality Video Semantic Concept Detection with Tensor Representation[C]// Proc. ACM Multimedia. 2008:91-100

[21] Wu F,Liu Y,Zhuang Y. Tensor-based Transductive Learning for Multimodality Video Semantic Concept Detection[J]. IEEE Transactions on Multimedia,2009,11(5):868-878

[22] Kotsia I,Patras I. Multiplicative Update Rules for Multilinear Support Tensor Machines[C]// 2010 20th International Conf. on Pattern Recognition (ICPR). 2010:33-36

[23] Kotsia I,Patras I. Relative Margin Support Tensor Machines for Gait and Action Recognition[C]// Proceedings of the ACM International Conference on Image and Video Retrieval (CIVR). 2010

[24] Acar E, Yener B. Unsupervised Multiway Data Analysis: A Literature Survey[J]. IEEE Trans. Knowl. Data Eng. ,2009,21:6-20

[25] Yang J,Zhang D,Frangi A,et al. Two-dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2004,26(1):131-137

[26] Ye J. Generalized low rank approximations of matrices[J]. Machine Learning,2005,61(1-3):167-191

[27] Ye J,Janardan R,Li Q. GPCA: An Efficient Dimension Reduction Scheme for Image Compression and Retrieval[C]// The Tenth ACMSIGKDD International Conference on Knowledge Discovery and Data Mining. 2004:354-363

[28] Lu H,Plataniotis K,Venetsanopoulos A N. MPCA: Multilinear Principal Component Analysis of Tensor Objects[J]. IEEE Trans. Neural Netw. ,2008,19:18

[29] Inoue K, Hara K, Urahama K. Robust Multilinear Principal Component Analysis[C]// Proceedings of IEEE Conference on Computer Vision. 2009:591-597

[30] Panagakis Y,Kotropoulos C,Arce G. Non-Negative Multilinear Principal Component Analysis of Auditory Temporal Modulations for Music Genre Classification[J]. IEEE Transactions on Audio,Speech,and Language Processing,2010,18(3):576-588

[31] Lu H, Plataniotis K N, Venetsanopoulos A N. Uncorrelated Multilinear Principal Component Analysis for Unsupervised Multilinear Subspace Learning[J]. IEEE Transactions on Neural Networks,2009,20(11):1820-1836

[32] Shashua A, Levin A. Linear Image Coding for Regression and Classification Using The Tensor-Rank Principle[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. 2001,1:42-49

[33] Xu D,Yan S,Zhang L,et al. Reconstruction and Recognition of Tensor-Based Objects with Concurrent Subspaces Analysis[J]. IEEE Transactions on Circuits and Systems for Video Technology,2008,18(1):36-47

[34] Vasilescu M A O,Terzopoulos D. Mulitlinear Independent Component Analysis[C]// Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR05'). 2005:547

[35] Chandrika P,Jawahar C V. Multi Modal Semantic Indexing for Image Retrieval[C]// Proceedings of the ACM International Conference on Image and Video Retrieval (CIVR). 2010

[36] Tao D,Song M,Li X,et al. Bayesian Tensor Approach for 3-D Face Modeling[J]. IEEE Transactions on Circuits and Systems for Video Technology,2008,18(10):1397-1410

[37] Zhang L,Gao Q,Zhang D. Directional Independent Component Analysis with Tensor Representation[C]// IEEE Conf. Computer Vision and Pattern Recognition (CVPR08'). 2008

[38] Wang H,Ahuja N. Compact Representation of Multidimensional Data Using Tensor Rank-One Decomposition[C]// Proc. 17th Int'l Conf. Pattern Recognition (ICPR '04). 2004,1:44-47

[39] Wang H,Ahuja N. A Tensor Approximation Approach to Dimensionality Reduction[J]. Int'l J. Computer Vision,2008,76 (3):217-229