

三维人脸表情识别研究进展

北京航空航天大学 黄迪

一、引言

人脸表情是人类交流情感的主要方式，自动人脸表情识别一直以来都是模式识别和计算机视觉领域的重要研究问题。在过去的十几年中，随着便携式智能设备（例如智能手机和智能手表）的普及，人脸表情识别相关的应用层出不穷，并在人机交互、医疗健康、游戏娱乐等方面体现出巨大潜力，也因此受到了工业界的广泛关注。

对人脸表情识别的研究可以追溯到二十世纪七十年代。Ekman[1, 2]等人定义了六种普适的基本表情（即愤怒、厌恶、恐惧、高兴、悲伤与惊讶）和数十种面部动作单元，并提出将表情表示为若干面部动作单元的组合进行分析。自此之后，表情识别的研究得到大幅发展。早期的人脸表情识别方法主要基于二维图像或视频，然而受限于固有成像的特点，这些方法的准确率和鲁棒性并不理想。进入本世纪后，三维数据采集设备不断革新，三维人脸表情识别研究开始受到重视。同时，BU-3DFE、Bosphorus、BU-4DFE 等数据集的发布实质性地推动了相关研究的开展。

表情本质上是由人类面部肌肉运动引发的形状变化。三维数据能够直接记录这些形变。相比于二维数据，三维数据对于环境光照和头部姿态变化更具鲁棒性。近年来，动态三维（四维）人脸表情识别开始进入视野，序列信息的利用使三维人脸表情识别的准确率得到了进一步提升。

本文主要介绍三维人脸表情（六种基本表情）识别的核心问题及一般流程，重点阐述基于静态和动态三维数据的人脸表情识别的最新进展，并对主要方法进行比较分析和定性评价，最后指出三维人脸表情识别未来的一些可能的研究方向。

二、三维人脸表情识别背景知识

1. 处理流程

与其他三维人脸分析任务类似，传统三维人

脸表情识别方法的一般处理流程包含四个关键技术环节（如图 1）：即数据获取（Data Acquisition）、数据预处理（Data Pre-processing）、形状表示（Shape Representation）和特征分类（Feature Classification）。

数据获取是通过特定的三维采集设备记录人脸的原始形状信息。数据预处理是对原始人脸模型进行毛刺移除和孔洞填充等操作，尽可能减少噪声的影响；此外还会根据需要进行标志点定位和人脸配准等步骤。形状表示是从人脸曲面上提取几何特征，其特征的判别能力直接影响识别精度。特征分类是基于曲面表示结合特定的分类器输出表情标签。

深度学习相关理论和技术最近被成功应用于三维人脸表情识别。其流程基本沿袭了传统方法的处理框架，但将曲面表示和表情建模进行整合，以数据驱动的方式面向三维人脸表情识别任务建立更为有效的几何特征。

2. 核心问题

形状表示（几何特征提取）和表情建模三维人脸表情识别的两个核心问题。形状表示目的是获得与表情相关的几何特征。这些特征不仅要针对不同表情有高区分性，还要对身份等其他属性变化有高鲁棒性。区分性和鲁棒性相互制约，使这一问题极具挑战。此外，三维人脸数据一般保存为点云或网格形式，与二维图像和视频不同，呈现分布无规则的特点，这为几何特征提取增加了额外难度。表情建模旨在建立表情类别预测模型。这是典型的机器学习问题，一般被建模为分类任务求解。而根据数据类型不同，该问题又可以进一步细分为面向静态特征的建模与面向动态特征的建模。前者偏重分析不同表情的全局或局部形状差异，而后者则偏重分析不同表情面部曲面形变的时空特性。

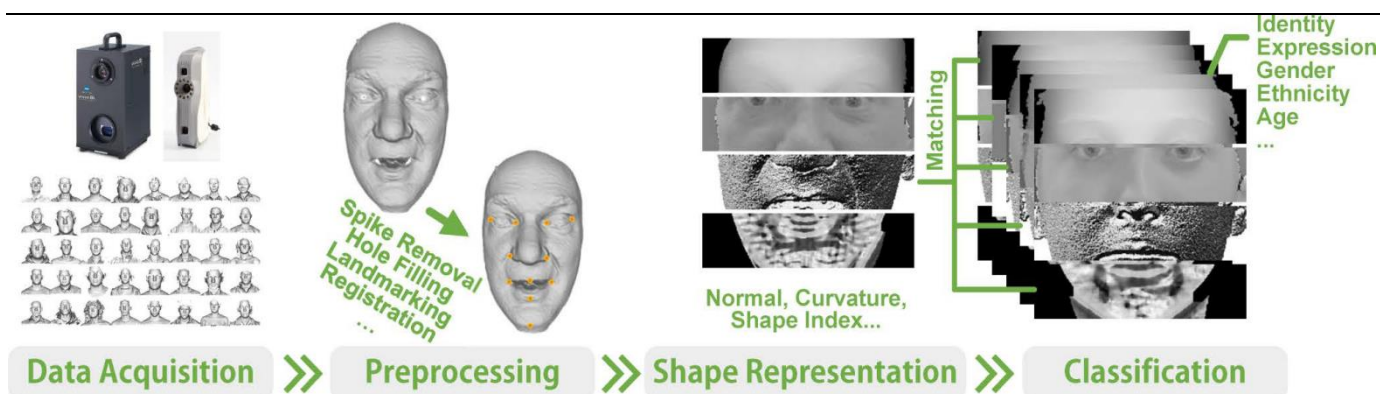


图 1 三维人脸表情识别的一般处理框架

基于深度学习的方法尝试突破传统手工设计特征的束缚，采用端到端的结构，利用随机梯度下降等优化技术，使模型同时改善特征提取和分类过程。初步的研究结果表明了这些方法的有效性，其中核心问题在于如何设计网络结构和损失函数，使模型具有更强的任务适应性和泛化性能。

三、静态三维人脸表情识别进展

在近十几年的研究中，静态三维人脸表情识别方法逐渐形成了两大类：基于模型的方法和基于特征的方法。

1. 基于模型的方法

基于模型的方法建立通用人脸形状表示模型，并拟合给定三维人脸样本将得到的系数直接作为特征或进一步构建特征用于类别表情预测。

Mpiperis 等人[3]提出了一种基于双线性模型的表情识别方法。该方法首先基于标志点将三维人脸曲面重新网格化，使所有样本拓扑相同建立点对应关系，并进一步采用主成分分析（PCA）方法将身份和表情进行分解得到双线性模型。这样每个三维人脸模型都可以计算身份系数和表情系数，其中表情系数用来进行表情识别。

Gong 等人[4]提出了一种基于共性形状表示的表情识别方法。该方法建立统一的基础面部形状部件模型，将三维人脸曲面表示成基础面部形状部件和表情形状部件，并基于原始人脸曲面和基础面部形状部件，在眼睛和嘴巴等区域提取几何特征，之后通过计算这两部分特征的差值得到表情形状部件特征，并将其用于表情分类。

Zhao 等人[5]提出了一种基于贝叶斯置信网

络的表情识别方法。该方法基于吉布斯-玻耳兹曼分布在选定的十几个人脸标志点的一系列纹理和几何特征空间建立统计学表征模型。该模型可以在拟合过程中自动定位关键点并提取相应特征进行表情识别。

Zhen 等人[6]提出了一种基于肌肉运动模型的表情识别方法。该方法结合人脸面部肌肉的组成和运动特点，采用迭代最近法相量点方法对人脸曲面进行分割。在每一个肌肉区域提取多阶几何特征，并通过遗传算法选择最优的子集作为特征表示用于表情标签预测。

基于模型的方法理论体系完备，更易于人类理解，但它们对于表情相关的形状属性（尤其是对局部细节特点）刻画能力有限，因此并未体现出明显的优势。

2. 基于特征的方法

基于特征的方法直接从人脸曲面（或相应的人脸深度图像）上提取点、线、面特征用于表情分类。

基于点特征的方法在以特定的标志点为中心的区块上计算局部形状特征，并将所有区块的特征融合形成最终的曲面表示作为表情识别的全局描述。Maalej 等人[9]在关键点周围通过同心测地圆环来构造区块，并将每个区块映射为高维流形上的样本点。他们通过计算样本点间最短的测地距离来衡量相应区块的差异，而后将人脸对应点附近的区块差异进行融合描述不同人脸形状的整体差异并用于分类。Derkach 和 Sukno[10]同样以测地同心圆环的方式提取人脸

关键点特征，由于每个区块具有完全相同的网格拓扑结构，该方法可方便地采用图拉普拉斯算子计算特征基向量，并将每个区块的三维坐标在特征基向量上进行分解，得到在该基向量下的谱域特征，最终用费舍尔线性判别（LDA）和支持向量机（SVM）进行分类。

基于线特征的方法计算人脸曲面上表情相关的距离，并通过特定点间距离的变化或特定线段夹角的变化等信息判别表情。Soyel 和 Demirel[7]手动在人脸上标注了十余个关键点，并根据这些点计算眼睛开合程度、眼眉高度、嘴巴左右开合程度、嘴巴上下开合程度、唇形拉伸程度及人脸左右边缘最外侧点间跨度等六种距离特征，并将这些距离输入人工神经网络进行表情分类。Tang 等人[8]采用类似方法，但大幅增加了特征点个数（68 点），并计算了多达 24 个距离特征和 24 个角度特征用于表情预测。

基于面特征的方法通常并非只关注人脸上某些局部区块，而是将三维人脸完整划分为若干子区域，并综合考虑每个子区域的特性。Wang 等人[11]在人脸的每个点计算主曲率方向及最大最小主曲率值，进而根据这些值逐点进行曲率类型划分。他们随后将人脸根据关键点划分成七个区域，并在每个区域计算曲率方向直方图及曲率类型直方图作为形状特征。最后采用线性判别分析、二次判别分类器、朴素贝叶斯分类器及径向基核支持向量机等四种方法进行了标签预测。Sha 等人[12]利用人脸上的几何线段将曲面分成不同区域，并计算各区域的曲率信息。他们提取曲率强度较为突出的点，并将其组成一个图，通过基于图的归一化裁剪过滤器对特征点进行选择，最后采用选择的特征进行判别。Lemaire 等人[13]采用变化邻域估计平均曲率，并将其映射在二维平面生成微分平均曲率图，而后提取方向梯度直方图（HOG）特征用于表情识别。Yang 等人[14]利用散射算子在人脸深度图、一阶微分法向量图及二阶微分形状索引（Shape Index）图上计算散射特征提取更为细节的几何变化，最后基于支持向量机进行分类。

基于特征的方法处理流程直观，其精度取决于人脸曲面几何表征的鉴别性，尤其是基于深度学习的方法近期体现出了巨大潜力。例如，Li 等人[15]通过预训练卷积神经网络（VGG 模型）在[14]中的多种几何图上提取深度特征，并采用支持向量机进行分类。随后，他们设计了深度融合卷积神经网络，实现了端到端训练，进一步提升了表情识别精度。

四、 动态三维人脸表情识别进展

动态三维人脸表情识别除静态身份无关特征提取外，更强调视频序列中时序特性的建模。

在首个动态三维人脸数据集 BU-4DFE 发布时，Sun 和 Yin[17]提出了一个尝试性的动态三维人脸表情识别方法。该方法根据三维人脸曲率类别，将人脸曲面划分为六个区域，并采用线性判别分析对区域内特征进行降维。之后，该方法基于二维隐马尔科夫模型（HMM）在特征空间对时序变化进行建模，将空间区域作为第一维状态，将时序信息作为第二维状态。该方法针对每一种表情训练一个模型，并采用似然概率下的贝叶斯决策规则进行判别。

Le 等人[18]将视频各三维人脸曲面帧以鼻尖点为中心划分为一系列同心圆环，并将与参考人脸对应圆环间的倒角距离（Chamfer Distance）作为特征，经主成分分析和线性判别分析降维后训练隐马尔科夫模型建立表情识别分类器。

Sandbach 等人[19]将三维人脸帧序列在六个时空方向（ $x-y$ 、 $x-z$ 、 $y-z$ 、 $x-t$ 、 $y-t$ 及 $z-t$ ）上进行投影并计算其四叉树分解模式作为表情特征，之后采用隐马尔科夫模型进行识别。

Fang 等人[20]使用基于网格数据的方向梯度直方图（MeshHOG）方法进行曲面配准，并采用三个正交平面上的局部二值模式（LBP-TOP）特征对三维人脸序列的形状变化进行表示，最后采用支持向量机进行判别。

Hayat 等人[21]将不同位置可变长度的三维人脸序列区块映射为格拉斯曼（Grassmannian）流形上的点，并利用局部二值模式（LBP）等特征空间的谱聚类对不同表情的流形点分别计算中

心。通过匹配测试三维人脸序列的流形点与上述聚类中心得到表情标签。

Danelakis 等人[22]基于关键点构造了三种特征,包括热核特征、法向量及角度、面积和距离等几何特征,并采用动态时间对齐算法度量两个时间序列的相似性,将不同特征的加权结果生成评分,最后通过K近邻(KNN)方法进行预测。

Drira 等人[23]和 Ben Amor 等人[24]将三维人脸曲面沿以鼻尖点为中心的放射曲线重新网格化,计算帧间稠密标量场(Dense Scalar Field)作为特征,经线性判别分析降维后采用随机森林和隐马尔可夫模型识别表情。Zhen 等人[25]采用同样的特征表示和相似的分类器,提出了空间域高斯金字塔分解方法,并在时域上对形变细微的表情进行增强。此外,他们还提出了基于聚类的表情周期检测算法,增强了方法适用性。

Li 等人[26]率先提出了一种基于深度学习技术的动态三维人脸表情识别方法,采用动态几何图网络捕获三维人脸曲面序列的时序形变信息和由粗到精的二阶段划窗采样技术进行数据扩增。通过将多几何特征图的预测结果进行加权融合,提升了识别精度。

五、未来可能的研究方向

基于静态数据和动态数据的三维人脸表情识别在过去的十几年间都取得了长足的进步,但算法的准确性和实用性仍需进一步提升。

在**准确率**方面,关键挑战在于部分表情难以区分。具体原因一方面体现在一些表情(如愤怒和悲伤)驱动面部肌肉的模式有相似性,导致几何特征易混淆;另一方面体现在一些表情程度较弱,面部形变轻微,导致几何特征不明显。

在**实用性**方面,关键挑战在于应对姿态和遮挡等真实环境的变化。其中,如何从连续三维人脸视频流中检测到表情信号并正确识别成为必

须。此外,如何提升对于头部姿态变化和外部遮挡变化的鲁棒性愈发重要。

针对这些挑战,可以考虑的方案如下:

1. **通过深度学习技术提升人脸曲面(序列)的表征能力**。目前一些工作[14-16]尝试利用深度学习技术提升三维人脸表情识别的精度,并初步展示出了不错的前景。但是这些方法都是将三维人脸的深度图作为输入,本质上是复制二维数据的处理方法,因此(部分)丢掉了三维的优势。

目前,已经有一些研究者尝试提出面向三维点云和网格数据的深度学习方法[27-29],也取得了一些进展。在三维人脸表情识别问题上,Chen 等人[30]新近提出了基于快速轻量化流形卷积神经网络的表情识别方法,不但取得了非常有竞争力的结果,还体现出了对于姿态变化的良好稳定性。然而由于三维数据点分布的不规则性,对三维人脸曲面的深度表征还存在很多困难,是需要重点探索的方向。

2. **通过融合多模态数据引入辅助信息提升识别准确率**。最直接和常用的多模态融合方式就是整合二维与三维人脸数据,这也是三维人脸表情识别的一个分支。Oyedotun 等人[16]最新提出的方法针对形状信息和纹理信息,分别设计了不同的深度神经网络进行学习,之后采用全连接的方式将两种模态的深度特征进行融合,其表情识别精度明显优于只基于单一模态的方法。

另外,表情的产生常会伴随着语音、手势等信息,还会引起人体一些生理指标(如血压、心跳)的变化。一些研究表明,通过采集更多模态的数据、提取更为丰富的信息可以进一步提升表情识别的准确率。所以,如何将不同模态数据进行融合也是一个很有必要的研究方向。

(责任编辑:任传贤)

参考文献

- [1] Ekman P and Oster H. Facial Expressions of Emotion. Annual Review of Psychology, 1979(30):527-554.
- [2] Ekman P. Strong Evidence for Universals in Facial Expressions: A Reply to Russell's Mistaken Critique. Psychological Bulletin,

- 1994, 115(2):268–287.
- [3] Mpiperis I, Malassiotis S, Srinivas M G. Bilinear Models for 3-D Face and Facial Expression Recognition. *IEEE Transactions on Information Forensics and Security*, 2008, 3(3): 498-511
 - [4] Gong B, Wang Y, Liu J, et al. Automatic Facial Expression Recognition on a Single 3D Face by Exploring Shape Deformation. *ACM International Conference on Multimedia*. ACM, 2009: 569-572
 - [5] Zhao X, Huang D, Dellandréa E, et al. Automatic 3D Facial Expression Recognition based on a Bayesian Belief Net and a Statistical Facial Feature Model. *IEEE International Conference on Pattern Recognition*. IEEE, 2010: 3724-3727
 - [6] Zhen Q, Huang D, Wang Y, et al. Muscular Movement Model-based Automatic 3D/4D Facial Expression Recognition. *IEEE Transactions on Multimedia*, 2016, 18(7): 1438-1450
 - [7] Soyel H, Demirel H. Facial Expression Recognition using 3D Facial Feature Distances. *International Conference Image Analysis and Recognition*. Springer, Berlin, Heidelberg, 2007: 831-838
 - [8] Tang H, Huang T S. 3D Facial Expression Recognition based on Automatically Selected Features. *IEEE Conference on Computer Vision and Pattern Recognition Workshops*. IEEE, 2008: 1-8.[27]
 - [9] Maalej A, Amor B B, Daoudi M, et al. Shape Analysis of Local Facial Patches for 3D Facial Expression Recognition. *Pattern Recognition*, 2011, 44(8): 1581-1589
 - [10] Derkach D, Sukno F M. Local Shape Spectrum Analysis for 3D Facial Expression Recognition. *IEEE International Conference on Automatic Face and Gesture Recognition*. IEEE, 2017: 41-47
 - [11] Wang J, Yin L, Wei X, et al. 3D Facial Expression Recognition based on Primitive Surface Feature Distribution. *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2006, 2: 1399-1406
 - [12] Sha T, Song M, Bu J, et al. Feature Level Analysis for 3D Facial Expression Recognition. *Neurocomputing*, 2011, 74(12-13): 2135-2141
 - [13] Lemaire P, Ardabilian M, Chen L, et al. Fully Automatic 3D Facial Expression Recognition using Differential Mean Curvature Maps and Histograms of Oriented Gradients. *IEEE International Conference on Automatic Face and Gesture Recognition*. IEEE, 2013: 1-7
 - [14] Yang X, Huang D, Wang Y, et al. Automatic 3D Facial Expression Recognition using Geometric Scattering Representation. *IEEE International Conference on Automatic Face and Gesture Recognition*. IEEE, 2015, 1: 1-6
 - [15] Li H, Sun J, Xu Z, et al. Multimodal 2d+ 3d Facial Expression Recognition with Deep Fusion Convolutional Neural Network. *IEEE Transactions on Multimedia*, 2017, 19(12): 2816-2831
 - [16] Oyedotun O K, Demisse G G, Shabayek A E R, et al. Facial Expression Recognition via Joint Deep Learning of RGB-Depth Map Latent Representations. *IEEE International Conference on Computer Vision Workshops*. IEEE, 2017: 3161-3168
 - [17] Sun Y, Yin L. Facial Expression Recognition based on 3D Dynamic Range Model Sequences. *European Conference on Computer Vision*. Springer, Berlin, Heidelberg, 2008: 58-71
 - [18] Le V, Tang H, Huang T S. Expression Recognition from 3D Dynamic Faces using Robust Spatio-Temporal Shape Features. *IEEE International Conference on Automatic Face and Gesture Recognition*. IEEE, 2011: 414-421
 - [19] Sandbach G, Zafeiriou S, Pantic M, et al. Recognition of 3D Facial Expression Dynamics. *Image and Vision Computing*, 2012, 30(10): 762-773
 - [20] Fang T, Zhao X, Shah S K, et al. 4D Facial Expression Recognition. *IEEE International Conference on Computer Vision Workshops*. IEEE, 2011: 1594-1601
 - [21] Hayat M, Bennamoun M, El-Sallam A A. Clustering of Video-Patches on Grassmannian Manifold for Facial Expression Recognition from 3D Videos. *IEEE Workshop on Applications of Computer Vision*. IEEE, 2013: 83-88
 - [22] Danelakis A, Theoharis T, Pratikakis I, et al. An Effective Methodology for Dynamic 3D Facial Expression Retrieval. *Pattern Recognition*, 2016, 52: 174-185
 - [23] Drira H, Amor B B, Daoudi M, et al. 3D Dynamic Expression Recognition based on a Novel Deformation Vector Field and Random Forest. *IEEE International Conference on Pattern Recognition*. IEEE, 2012: 1104-1107
 - [24] Amor B B, Drira H, Berretti S, et al. 4D Facial Expression Recognition by Learning Geometric Deformations. *IEEE Transaction*

on Cybernetics, 2014, 44(12): 2443-2457

- [25] Zhen Q, Huang D, Dirra H, et al. Magnifying Subtle Facial Motions for Effective 4D Expression Recognition. IEEE Transactions on Affective Computing, 2017
 - [26] Li W, Huang D, Li H, et al. Automatic 4D Facial Expression Recognition using Dynamic Geometrical Image Network. IEEE International Conference on Automatic Face and Gesture Recognition. IEEE, 2018: 24-30
 - [27] Riegler G, Ulusoy A O, Geiger A. Octnet: Learning Deep 3D Representations at High Resolutions. IEEE Conference on Computer Vision and Pattern Recognition. 2017, 3
 - [28] Qi C R, Yi L, Su H, et al. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. Conference on Neural Information Processing Systems. 2017: 5099-5108
 - [29] Monti F, Boscaini D, Masci J, et al. Geometric Deep Learning on Graphs and Manifolds using Mixture Model CNNs. IEEE Conference on Computer Vision and Pattern Recognition. 2017, 1(2): 3
 - [30] Chen Z, Huang D, Wang Y, et al. Fast and Light Manifold CNN based 3D Facial Expression Recognition across Pose Variations. ACM International Conference on Multimedia. ACM, 2018: 229-238.
-



黄迪

博士、北京航空航天大学计算机学院院长聘副教授、博士生导师。主要研究方向涉及生物特征识别，智能图像/视频处理及情感计算等。Email: dhuang@buaa.edu.cn

面向智能无人机的计算机视觉

天津大学 朱鹏飞 胡清华 京东数科 文珑银 纽约州立大学奥尔巴尼分校 独大为

无人机云台由于搭载了不同类型的传感器，可实现对周围环境的大视角多维度感知。天津大学机器学习与数据挖掘团队依托天津市机器学习重点实验室建立了 VisDrone 无人机数据平台，覆盖了 120 米以下的无人机目标检测、目标跟踪、目标计数等任务，可见光和热红外等模式，单机和多机等配置。目前目标检测和跟踪数据采集遍布中国 14 个城市，共包含 189473 张视频帧/图像，约 250 万个标注框。图 1 为 VisDrone 检测和跟踪数据集的样例。研究团队在 ECCV2018 举办了第一届 VisDrone 无人机视觉数据竞赛，来自全球百余所大学、科研结构和企业参加了此次竞赛，目前正在 ICCV2019 举办第二届比赛。VisDrone 数据平台未来将不断拓展，服务于计算机视觉等研究领域。

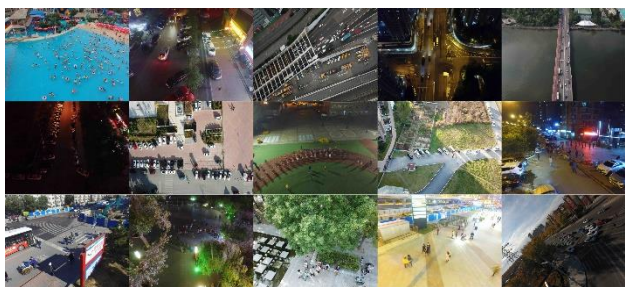


图 1 VisDrone 数据集样例



图 2 多无人机协同感知

未来的智能无人机面对复杂环境的挑战，同时环境中同时存在多源异构的多类型智能体，如何实现无人机的自主感知、学习和决策，并联合其他智能体协同学习，是值得期待的重要研究方向。图 2 展示了多无人机协同跟踪的示例。同时 5G 时代的到来使得无人机在通信和数据传输方面形成新的突破，如何开展 5G 时代无人机端的边缘计算以及云端边协同等研究也是无人机视觉研究可以关注的方向。此外无人机视觉的研究可考虑控制学、多智能体学习等结合，在交叉学科领域寻求新的突破。

(责任编辑：杨巨峰)



朱鹏飞

天津大学副教授，主要研究方向是无人机视觉、机器学习。

Email: zhupengfei@tju.edu.cn



胡清华

天津大学教授，主要研究方向是机器学习。

Email: huqinghua@tju.edu.cn



文珑银

京东数科美国研究院高级研究员，主要研究方向是计算机视觉和机器学习。

Email: longyin.wen@jd.com



独大为

纽约州立大学奥尔巴尼分校研究助理，主要研究方向是计算机视觉。

Email: ddu@albany.edu

多媒体检索技术中的主动学习方法

新加坡资讯通信研究院 张立宁

近年来，随着 Weibo、Facebook 等社交媒体的流行，图像、视频等异构数据的数量每天都在以惊人的速度增长。当人们面对一个超大型多媒体信息库时，单单凭借关键字很难对多媒体数据库进行有效的描述和检索。如何有效地帮助人们快速而准确地找到所需信息，成为多媒体信息领域需要解决的核心问题。多媒体信息检索可以分析图像内容，提取图像颜色、形状、纹理以及图像空间关系等信息，并对图像数据库建立索引。

目前各大主流网站有一些针对具体场景的应用，但依然存在算法处理速度慢、漏检率高、检索效果差等问题。多媒体检索技术的核心是系统既可以从图像自身提取图像特征，又可以通过与用户交互获得更多的信息，并将其用于图像间相似性的计算。系统和用户之间的关系是双向的：用户可以向系统提出查询要求，系统针对用户的查询要求返回相应的查询结果；系统可以通过用户查询结果的相关反馈来改进系统检索的性能。

我们提出了一种新颖的主动学习方法用于多媒体检索系统中选择最具信息量的样本数据由用户标记反馈，称为流形正则实验设计方法。

传统的主动学习方法，目前最大的挑战是如何有效地度量无标签样本的信息量。支撑向量机主动学习借助自身最优超平面来选择最不确定的样本由用户去标记。而要获得支撑向量机的最优超平面，系统往往需要用户大量的反馈信息，

通常用户会厌倦和疲劳，因此系统的检索准确度很难提高。

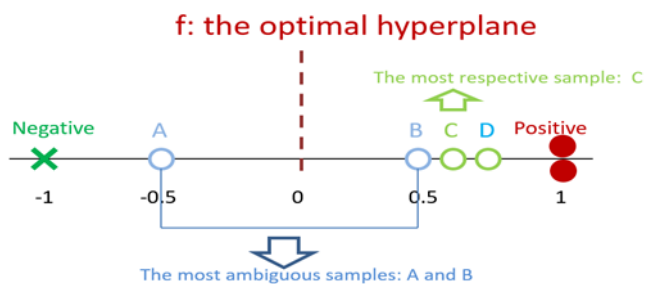
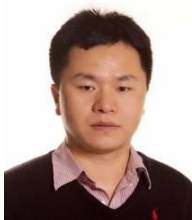


图 1 多长度哈希联合学习框架图

如图 1，新方法可以使检索系统选择多媒体数据库中最具代表性的样本集 B 和 C，而不是选择最不确定的样本集 A 或 B。我们的方法通过极小化目标函数在测试样本上的期望平均预测方差，可以使得被选择的样本具有清楚的几何空间解释，即可以选择数据库中最具代表性的样本，同时还可以全局最优地迭代选择最具信息量的样本进行给用户标记。因此可以极大地减少用户的标记工作，同时还可以保持系统更高的检索精确度。提出方法并不依赖于任何预选取的样本标签信息，可以有效地避免传统支撑向量机主动学习方法中由不充分、不准确样本标签数据引起的潜在问题。在大量真实数据库上的实验结果验证了流形正则实验设计方法在多媒体检索应用中的有效性。以上工作发表于 TIP 2017。

(责任编辑：邓成)



张立宁

新加坡资讯通信研究院研究员。主要研究方向为机器学习、计算机视觉、多媒体计算、新材料器件设计，电路设计等人工智能应用领域。

Email: zhangl1@i2r.a-star.edu.sg