

国家重点基础研究发展计划（973 计划）项目
“面向公共安全的跨媒体计算理论与方法”简报
（2013 年第 1 期）

项目管理办公室编制

2013 年 3 月 25 日

- 一、973 项目“面向公共安全的跨媒体计算理论与方法”在北京举行研讨会
- 二、973 项目“面向公共安全的跨媒体计算理论与方法”在西安举行研讨会
- 三、973 项目“面向公共安全的跨媒体计算理论与方法”2012 年年度总结会议顺利召开
- 四、“跨媒体”研究进入多媒体领域顶级国际学术会议 ACM Multimedia Research Track
- 五、973 项目“面向公共安全的跨媒体计算理论与方法”2012 年年度进展整体介绍
- 六、国际相关学术热点和动态

一、973 项目“面向公共安全的跨媒体计算理论与方法” 在北京举行研讨会

2012 年 3 月 30 日，国家重点基础研究发展计划（973 计划）项目“面向公共安全的跨媒体计算理论与方法”（项目编号：2012CB316400）在北京友谊宾馆召开了学术研讨会。



973 项目北京研讨会合影（2012 年 3 月 30 日 北京）

这次会议的目的是结合当前国际学术热点，聚焦项目或课题中某个（或某些）科学性难点问题进行讨论。本次研讨会由北京交通大学赵耀教授课题组负责承担。

国家科技部基础研究管理中心李非博士、王公仆博士，项目咨询组责任专家施鹏飞教授、戴国忠教授、钱德沛教授，项目组专家高文院士、袁保宗教授、胡事民教授、庄越挺教授，项目课题负责人赵耀教授、薛建儒教授、黄庆明教授、李学龙研究员、张仲非教授及主要骨干成员等出席了本次启动会。

学术研讨会由项目首席科学家庄越挺教授主持。

国家科技部基础研究管理中心李非博士认为在 973 项目执行过程中定期召开学术交流研讨会这一形式十分重要，并借此机会对 2012 年度 973 项目申请中新现的若干特点进行说明。（1）973 项目将进一步强化组织管理，今年开始要求每个项目参与单位不超过 6 个，参与骨干成员不超过 30 人；（2）加强财务的严格管理，财务验收不通过的课题所在单位将不能继续申报，要求严格按照规划经费，人员有变动时需向科技部及时报告；（3）项目所属课题在研究过程中要

围绕整个项目的总体目标而展开，建立科学评价体系，不片面追求论文的数量。

项目首席科学家庄越挺教授对科技部领导与各位专家的光临指导表示感谢，并强调此次会议是一次头脑风暴，希望大家围绕科学问题进行深入讨论。

北京交通大学朱振峰副教授代表课题一进行了“跨媒体数据一致性表示中共享子空间技术”汇报，其内容以矩阵分解为核心展开，讨论了共享子空间建模过程中求解约束、算法收敛性以及增量在线计算等问题。

西北工大於志文教授代表课题二进行了“跨媒体属性模型感知与行为计算”汇报，其内容涉及到基于树的群体交互模式挖掘、基于跨媒体数据集成的社会关系管理以及基于手机感知数据的个体行为识别等方面。

中科院计算所刘纯熙博士代表课题三进行了“网络话题、事件检测与分析”汇报，其内容涉及到基于语义标注的跨媒体关联、融合用户意图的话题检测、基于用户信息指导的话题检测以及基于时间段的跨媒体聚类分析等方面。

浙江大学何晓飞教授代表课题四进行了“海量跨媒体数据挖掘与公共安全态势分析”汇报，其内容涉及数据降维、稀疏编码、哈希映射、基于向量场的学习方法以及多尺度结构化主题时空建模等方面。

浙江大学吴飞教授代表课题五进行了“跨媒体语义理解与检索中的结构性学习机制”汇报，其内容涉及到跨媒体数据中的结构性属性分析、基于结构性输入/输出正则化因子的学习机制、图约束机制、跨域监督式主题建模等方面。

浙江大学张仲非教授代表课题六进行了“面向公共安全的跨媒体呈现与验证和示范平台”汇报，其内容涉及关联性挖掘、规则挖掘以及示范性平台建设等三个方面。

与会人员就项目研究的执行展开了热烈讨论，项目咨询组责任专家和项目组专家也提出了诸多建议。

高文院士首先肯定了本项目的工作，认为年轻人工作扎实、印象深刻，也希望项目在进行理论攻关和深入的同时，要在“跨领域和深度搜索”这些方面多投入研究。高院士还以“李梦楠杀医生”这一事件为例，指出如果本项目能够从个体行为、社会普遍心理、社会学、精神病学、病理学等不同角度，结合来自不同领域的知识，对来自不同渠道的信息进行分析，给出这个事件的“深度分析”，这可作为本项目一个现实需求。

戴国忠教授建议项目组在埋头做自己事情的同时，也要关注正在执行中的973项目中与本项目较为相关的其他项目组，即973项目群中也要开展交流和沟

通。戴老师认为跨媒体是数据客观存在形式，本项目是对这一类型数据处理，对跨媒体数据进行处理和重新组织，而不是去产生（构建）一批跨媒体数据。在进行研究过程中，要主动与认知科学、心理学、媒体科学紧密结合起来；在重大需求应用上，公共安全并不就是舆情分析，也可以考虑环境监测、社会发展和公共卫生等应用。跨媒体呈现这个研究中，也注重研究交互，未来 25 年人机交互是主要研究方向，应把交互作为一个主要的研究内容，如美国国家科学基金会正在进行的“Technology-Mediated Social Participation”的研究。

施鹏飞教授对六个课题汇报进行了点评，希望项目组的研究成果重质量与内涵，项目本身的研究要从真正问题出发去开展研究，研究中注意数学理论的深化。

钱德沛教授建议在今后研究中将统一表达、主题检测、演化演变分析以及验证真正串联成一条线，确保每个课题不仅把自己的东西做好而且还能够给别人使用。为了更好显示项目研究的有机性，建议在项目内做一个基础平台，将各种不同的媒体结合起来，能够预测未来的数据集。

袁保宗教授认为项目在选题上很好，是时代的前沿，因此要紧密关注国际动态，如最近 IEEE Transactions on Multimedia 的专刊“Social Media as Sensors”以及美国国家标准技术研究院(NIST)正在开展的“Measurement Science for Complex Information Systems”等等。在项目开展过程中，要及时形成跨媒体(Cross-Media)的统一概念，明确与 Social Network、Social Media、Social Computing、Social Perception 以及 Network Perception 等概念的区别和联系。在公共安全这一应用对象上，可以选择我们既能做得到，又比较动人的问题。

胡事民教授认为项目研究中要突出“跨媒体”、“计算理论”和“公共安全”三个关键：在计算理论上，要关注“表示理论”、“感知理论”等问题；在跨媒体上，应当强调协同，突出“跨”字；在公共安全方面，应当体现具体应用。此外，胡事民教授建议大家分清楚每个课题的研究范围，加强“责任田”管理。

首席科学家庄越挺教授对会议进行了总结，对北京交通大学赵耀教授课题组对本次研讨会的付出表示感谢，充分肯定了本次研讨会，要求各个课题认真对待专家所提出的建议，突出和加强本课题研究。

会议讨论确定了下一次研讨会将于六月中下旬或七月初在西安召开，主要议题是各个课题围绕本课题若干重点和难点问题汇报，以及讨论项目整体技术架构和平台。

二、973 项目“面向公共安全的跨媒体计算理论与方法” 在西安举行研讨会

2012 年 7 月 9 日，国家重点基础研究发展计划（973 计划）项目“面向公共安全的跨媒体计算理论与方法”（项目编号：2012CB316400）在西安召开了项目研讨报告会。



973 项目西安研讨会合影（2012 年 7 月 9 日 西安）

项目首席科学家庄越挺教授首先介绍了这次会议的主要目的，对 973 项目过程管理中所涉及的年度总结、成果标注、中期评估、本项目关键节点等问题进行了详细说明。随后各课题组负责人以及主要成员对上阶段研究进展进行了集中汇报。

北京交通大学赵耀教授代表课题一“跨媒体统一表示和建模机制”进行了汇报。赵耀教授在重点介绍了课题一的整体研究进展以及计划后，提出了三个关键问题，即如何描述跨媒体的单一媒体、如何寻找跨媒体间的语义关联（关联检测）、如何建立跨媒体间基于关联的统一表示。针对这些关键问题，北京交通大学韦世奎副教授以及清华大学计算机系的徐昆博士分别做了报告。

韦世奎副教授介绍了基于词包模型及嵌入方法的媒体描述优化机制，这一机制提出了一系列优化准则来加快搜索速度和搜索精度，在视觉词典生成过程中引入一种 Partitioned K-Means Clustering 的方法，并利用 Linear Segment

Approximation 方法来产生描述子的近似嵌入码。同时,针对有监督的词典学习,课题一还提出了一个具有过完备特性的有监督 DRICA 模型,并利用学习的 DRICA 词典实现图像分类。

徐昆博士介绍了基于几何结构、相似矩阵的媒体基元的跨媒体关联分析方法。其目标在于提取和分析跨媒体三维几何单元形状和空间分布关系,提出了一种能够表达三维物体内容属性的纹理表现基元。在语音媒体分析与理解方面,提出了一种基于联合因子分析(JFA)的发音方式建模方法,构建说话人超向量及说法方式超向量,以实现对话人信息的建模。

西安交通大学刘剑毅博士代表薛建儒教授对课题二“跨媒体属性感知模型与行为计算”进行了汇报。刘剑毅博士重点介绍了基于 Web 图像挖掘的人脸属性感知方面的研究成果,该方法采用了基于多实例学习的年龄估计的方法,提高了年龄估计器的泛化能力。

其后西北工业大学王柱博士对课题二中跨媒体数据感知、关注度模型、群体行为分析(社群发现)这三方面的研究进展进行了阐述。其中特别对数据采集的方法、流程以及已经采集的数据进行了汇报。西安交通大学的杨旻博士随后汇报了针对人脸表情属性感知方面的一些进展。

中国科学院计算所黄庆明教授对课题三“跨媒体语义学习与内容理解”的研究进展进行了汇报。黄庆明教授从课题基本情况介绍、课题进展情况、下一步的规划与工作安排等几个方面对课题三的进展进行了详细介绍。内容覆盖了跨媒体语义单元学习和标注、热点话题及重大事件结构模式表示、热点话题及重大事件检测、基于群体智能的协同交互机制等四大方面。在跨媒体语义学习研究方面重点介绍了半监督多核学习、多任务多特征度量学习、海量跨媒体数据聚类学习;在跨媒体语义标注研究方面:分别介绍了层次化语义标注和字典学习、Vicept 学习、基于语义标注的行为分类识别研究;在话题检测与结构化表示方面,介绍了在线视频中的话题检测及话题摘要生成方法。

针对课题三涉及的事件和话题分析检测方面的内容,中国科学院计算所王树徽博士报告了相关进展,以及下一步的研究计划。其中下一步的重点研究内容包括面向海量数据的多学习任务共享机制,研究层次化分类和字典学习模型、跨模态相关分析模型及跨模态哈希技术,面向海量数据的聚类方法,研究数据代表性分析方法以及话题排序方法等。

浙江大学的何晓飞教授代表课题四“海量跨媒体数据挖掘与公共安全态势分

析”介绍了研究进展。何晓飞教授首先介绍了基于平行向量场的嵌入理论，借助平行向量场全局保距的性质，提出了基于平行向量场的流形排序算法，其创新点在于借助于平行向量场，并通过对向量场方向的控制，实现可以在沿测地线保序的流形排序算法。随后，何教授介绍了快速代表性采样算法，基于带限制非负矩阵分解的主题建模，社交媒体信息中的视觉和文本信息关联学习等内容。针对下一阶段目标，介绍了跨媒体数据流形对齐，海量规模跨媒体数据的快速代表性采样，多尺度社会热点事件跟踪与演化等研究思路。

首席科学家庄越挺教授对课题五“跨媒体搜索与内容整合”的相关内容进行了整体进展汇报。对课题五中涉及的跨媒体搜索、问答式检索意图理解、跨媒体排序与相关反馈与跨媒体内容整合这四方面内容，分别进行了深度剖析。并指出，要在“跨”字上下功夫，提出了在跨媒体领域应该涉及的五个方面，即：针对不同类型数据之间检索的跨媒体检索；针对不同类型数据之间索引的跨媒体索引；针对不同类型数据之间排序的跨媒体排序；针对不同类型数据之间整合的跨媒体整合；针对不同 Domain 数据之间整合的跨媒体迁移学习；并以上述五点作为课题五下一步研究的重点内容。

随后，浙江大学吴飞教授对课题五所在团队的具体研究进展进行了汇报。吴飞教授的汇报内容包括基于非凸正则化因子与图模型的特征选择、高维特征索引、跨媒体排序与搜索以及跨媒体内容整合。重点介绍了课题五在结构性选择中结构性 Non-Convex、基于图结构正则化约束的图像区域标注、在多任务中引入图约束的图像相关属性迁移学习、基于跨媒体索引结构化稀疏谱哈希索引、基于超图的谱哈希索引、基于样条回归哈希的图像检索以及跨域主题建模、监督式 LDA 跨域学习等方面取得的进展与成果。汤永川副教授随后也汇报了其关于跨媒体语义理解研究的相关进展。

浙江大学张仲非教授对课题六“面向公共安全的跨媒体呈现与验证和示范平台”进行了汇报，重点讲解了本课题最大的难点和挑战。并详细介绍了课题六在跨媒体数据的多视图、多标签学习；非完美标注的跨媒体数据的多视图学习；基于矩阵分解的主题回归；大规模非负密集二值矩阵分解等方面的取得的研究成果与进展情况。

课题六的浙江大学于慧敏教授、钱运涛教授、廖备水副教授也分别针对课题六所涉及到的跨媒体话题网络学习及贝叶斯推理、社会事件因果推演可控仿真、事件演变因果关系的表示与推理分别进行了相关介绍。

随后，与会人员对如课题执行中遇到的问题进行讨论，特别是针对数据集的集中采集，最终演示平台搭建的具体形式与细节进行了细致深入的探讨。

之后，与会人员就下列五个方面的问题进行了深入讨论：1）统一项目所涉及的基本概念的表述。不同课题间对同一概念有多种表述，建议减少表述方式，在项目内部尽量保持统一；2）建立共享数据。在数据采集过程中应超越中英文语言支持问题，论文实验可以偏英文数据集，示范应用可以偏中文数据集。数据共享可根据使用目的不同进行区分，针对研究实验目的，可收集国内外的公共标准数据集，方便算法对比；针对集成测试目的，可以实时抓取主流社交媒体站点，构建海量数据集进行集成测试与验证；建议建立站点实现项目内各课题间共享数据；3）项目集成展示。要求调研现有相关系统和工具，与之对比，突出本项目演示系统的创新性；4）进行国际交流与合作。建议举办相关研讨会或期刊专刊，对项目所涉及内容进行深入讨论和研究；5）加强项目管理问题。建议成立项目管理办公室，建立项目网站，撰写工作简报。

最后，首席科学家庄越挺教授对会议进行了总结，充分肯定了每个课题的表现，对每个子课题组所取得的现阶段成果与进展给予了高度的赞扬与鼓励，认为六个课题的总体状态很不错。相比而言，不同课题在研究内容理解与思路上各有深浅，研究方案的详略也不一致，要求各个课题加强从本课题角度考虑与其他课题之间的有机联系与支撑关系，并强调项目是一个整体，要合力构建项目演示平台。此外，还进一步希望各个课题组长及早布局，配备优势队伍，对指标完成预先建有规划。

会议确定了下一次讨论会的地点，并初步商量了下一次研讨会时间，要求下次会议每个课题就本课题如何突出跨媒体中的“跨”以及与整个项目之间的关系重点考虑；同时，为了扩大交流，促进青年学者成长，建议多种交流形式，加强课题内部与课题间的交流。同时，建议成立项目管理办公室，建立网站，介绍项目与各个课题，建立季度简报制度与会议纪要制度。

三、973 项目“面向公共安全的跨媒体计算理论与方法” 2012 年年度总结会议顺利召开

2012 年 11 月 24 日，国家重点基础研究发展计划（973 计划）项目“面向公共安全的跨媒体计算理论与方法”在杭州召开年度总结会。



973 项目杭州年度总结会议合影（2012 年 11 月 24 日 杭州）

项目咨询组责任专家施鹏飞教授，项目专家袁保宗教授、高文院士，项目首席庄越挺教授，项目课题负责人薛建儒教授、黄庆明教授、李学龙研究员及主要骨干成员出席了本次研讨会。

研讨会由首席科学家庄越挺教授主持。会议先由首席科学家汇报项目总体进展情况，再由各个课题分别汇报进展，会议过程中专家进行点评。

庄越挺教授先对项目总体进展进行了汇报。在汇报中，回顾了本项目研究动机与科学问题、研究方案与研究重点，提醒项目人员要围绕项目核心需求、围绕“跨媒体”开展研究工作；接着，简要介绍了各个课题的主要研究进展与整个项目研究成果，让专家和与会人员对项目的整体情况有一个概括性了解；最后，对项目后期工作进行了安排。

专家对项目总体进展情况进行了点评。施鹏飞教授充分肯定了项目的进展情况，同时也建议在论文上可以加强英文专著的撰写，在人才培养上可以争取青年 973 及重点项目的申请，在经费使用上要注重规划。袁保宗教授提到“简报”制度，加强成果的归纳与宣传；此外，也提醒大家重视财务问题，建议由专门会计进行管理。

课题一“跨媒体数据统一表示与建模机制”由北京交通大学、清华大学共同承担，由朱振峰博士汇报。工作进展主要是四个方面：（1）在跨媒体基元提取与描述上，主要开展了基于多标签联合优化的多类基元词典和基于拓扑非负矩阵分解（TNMF）基元提取这两项工作；（2）在跨媒体语义一致性描述方面，提出了基于共享子空间学习的跨媒体一致性描述方法；（3）在跨媒体关联建模方面，主要开展了跨模态的媒体数据关联、跨维度的媒体数据关联等工作；（4）在增量整合和高效索引方面，提出了基元空间分割的词典构造及映射方法。

专家对课题一的工作进行了点评。高文院士指出共享子空间方法的有效性，实际上是依赖于多态跨媒体数据之间存在关联可信这一假设。施鹏飞教授建议进一步区分“统一表示”和“一致性描述”这两个概念。而袁保宗教授则指出，需要对“统一表示”进行全面深入理解，“统一”有些地方可行，有些地方不行，在底层特征层面上的统一适用场景有限，统一在语义空间才是可行的。

课题二“跨媒体属性感知模型与行为计算”由西安交通大学与西北工业大学共同承担，由薛建儒教授汇报。工作进展主要是四个方面：（1）在数据采集及有效性辨识方面，构建了 60 万微博数据，并提出了基于颜色滤波阵列（CFA）的篡改检测算法；（2）在属性感知方面，开展了属性定义、度量与关联的工作，提出了显著语义单元提取算法以及基于 Web 图像的年龄/表情属性感知，支持 Who、Where、What 的推断；（3）在社会关系网络分析，提出基于多维度联合聚类重叠式社群发现方法；（4）在社会关注度模型，提出通过时序分析和数据挖掘等技术手段，结合社会动力学方法建立关注度模型。

专家对课题二的工作进行了点评。施鹏飞教授建议对关注度模型进行进一步深入研究，需要有一个好的定义。袁保宗教授认为课题二的“属性感知”与“行为计算”这两条线抓得很准，建议突出重点成果，从概念定义、方法阐述、创新点、论文发表情况、引用情况等几个方面逐一展开。

课题三“跨媒体语义学习与内容理解”由中科院计算所与西安交通大学共同承担，由黄庆明教授汇报。黄教授先就跨媒体语义单元学习与标注、话题及事件的结构模式表示与检测、基于群体智能的协同交互机制等三方面研究内容，介绍其存在的挑战及总体进展情况。接着，重点介绍代表性工作，主要包括：（1）基于混合范式正则化的图像语义关联描述；（2）网络视频热点事件检测。

专家对课题三的工作进行了点评。施鹏飞教授充分肯定了课题三的工作，考虑到集成问题，建议加强数据集的建设。袁保宗教授与大家交流讨论了“混合范

式”的概念。首席庄越挺教授强调大家在论文中需要突出跨媒体的概念，要跟总项目的目标保持一致。

课题四“海量跨媒体数据挖掘与公共安全态势分析”由中科院西安光机所和浙江大学共同承担，由李学龙教授汇报。工作进展主要是三方面：（1）在流形学习理论方面，提出了基于平行向量场的降维、排序、多任务学习方法；（2）在主题建模理论方面，提出了基于带限制的非负矩阵分解的图像表示理论与基于 A-最优非负投影的图像表示理论；（3）在语义关联分析方面，提出了基于超图表达的语义分析方法。

专家对课题四的工作进行了点评。高文院士提出了流形学习中为什么需要保持“光滑性”的问题，袁保宗教授提出了“Low Rank”与“稀疏”两者之间对比的问题，大家进行了深入的交流与讨论。此外，施鹏飞教授建议何晓飞等人可以将平行向量场中的理论问题进一步深化并争取更大的突破。

课题五“跨媒体搜索与内容整合”由浙江大学承担，由吴飞教授汇报。主要进展包括四个方面：（1）在高维异构特征选取方面，分别提出了基于结构稀疏正则化因子与结构性输入/输出正则化因子的高维异构特征选择算法，充分利用结构先验知识加强特征选择，此外，还提出了非凸结构组稀疏方法，保证选择结果的一致性；（2）在跨媒体度量学习与哈希索引方面，提出了结合视觉字典学习和文本主题建模的耦合子空间学习方法，基于高维稀疏系数的跨媒体哈希索引方法；（3）在跨媒体搜索与排序方面，提出了利用支持向量机实现结合结构学习和隐性模型的跨媒体检索排序方法；（4）跨域主题建模与内容合成，提出了监督式 LDA 跨域学习方法和基于 Entity-Aspect LDA 的主题合成方法。

专家对课题五的工作进行点评。高文院士指出在内容整合上可引入更多的线索，充分体现跨媒体特色。袁保宗教授对正则化因子的解析求解进行了讨论，高文院士提出非凸结构组稀疏方法（NOVA）中分段凸如何求解问题，并与大家一起进行了讨论。

课题六“面向公共安全的跨媒体呈现与验证和示范平台”由浙江大学承担，于慧敏教授汇报。工作进展主要是两方面：（1）在因果推演方面，提出了跨媒体的非完美标注方法，为大规模数据跨媒体事件因果分析问题提供有效的工具；提出了跨媒体话题网络学习与推理方法和基于因果关系的信息演化图，为主题的演变/演化与因果推断提供了理论模型和可能的解决方案；提出了基于抽象论、辩理论的事件因果关系的辩护与解释，支持因果关系的建立；（2）在跨媒体计算验证

平台方面，确定跨媒体数据源，采集构建跨媒体数据集，搭建跨媒体计算验证平台，初步支持跨媒体数据的表达分析、挖掘推理、搜索展现的一体化处理。

专家对课题六进行了点评。施鹏飞教授指出需要加强汇报内容的凝练，此外验证平台实现中需要明确应用场景。

经过各个课题的进展汇报和点评之后，各位专家对整个项目的执行情况进行了总结点评。

施鹏飞教授点评：（1）需要进一步加强各个课题之间、课题内部的合作和交流；（2）中期考核时间很紧，需要考虑如何在相关 973 项目群中突出特色，争取好的成果；（3）团队以中青年人才为主，需要重视年轻人才的培养；（4）建议要有国际化的视野，加强国际交流，提升项目国际影响力。

高文院士点评：在短期中取得不错的成果，令人鼓舞；但是为了争取中期拿到好的分数，结题取得优，还得加把劲，具体建议有二：（1）示范平台需要加强：将前五个课题成果通过平台整合进来；（2）各个课题之间的结合需要梳理清楚，搞清楚输入输出，在汇报中需要体现出来。

袁保宗教授点评：（1）中期考核看两件事，一是状态，二是前景，建议汇报中强化梯队建设内容，强化课题间交流内容；成果汇报既要通俗又要深入，即使听不懂但能看得懂；（2）整个项目前期构思很好，但是成果表现需要与构思更好的对应起来，需要下功夫加强成果提炼与组织，可以考虑对每个研究成果提炼几个 **Keyword**，对算法整理一个 **List**，重点讲清楚一两个；（3）加强跨媒体数据集的建设，数据建设越来越重要，需要加强标注，汇集算法程序库；（4）国际文章发表的不错，但是也需要注重中国科学等国内杂志的发表。

最后，庄越挺教授对会议进行了总结，认为经过各个课题的努力研究，目前取得了不少成果。经过专家的点评与指导，收获很大。中期考核前，后续工作主要包括：（1）加强课题组之间的联系，理顺输入输出，努力争取做到无缝衔接，突出跨媒体特色；（2）加强验证平台建设，中期考核前出 **DEMO**；（3）年度报告中加强成果整理与组织。

四、“跨媒体”研究进入多媒体领域顶级国际学术会议

ACM Multimedia Research Track

当前,来自微博、手机、社交网站、新闻网站以及多媒体共享网站中的文本、图像、视频和空间属性等数据迅猛增长。有别于传统的结构化和非结构化数据,这些从不同渠道获取的文本、图像和视频等不同类型媒体数据具有丰富的关联关系,对它们的分析与处理需要跨越媒体类型和跨越媒体来源,即“跨媒体(Cross-Media)”,如何突破单一类型媒体数据研究的理论和方法,实现跨媒体数据分析与理解,是处理这些具有“跨媒体”特性的数据亟待解决的巨大挑战。

近年来,跨媒体相关领域的研究已经引起了国际学术界同行的关注,多媒体领域顶级学术会议 ACM Multimedia 2013 在其 **Multimedia Analysis** 这个 Track 中列出了 **Cross-Media analysis** 的方向,这预示着“跨媒体”研究已经逐渐被学术界所认同。

为了更好在国际学术界介绍跨媒体研究,项目组在庄越挺教授和张仲非教授领导下,与包括 ACM Multimedia 2013 大会联合主席 Alejandro (Alex) Jaimes (Yahoo!, Spain) 博士等国际同行进行了交流和沟通,对“跨媒体”这一研究进行了如下形式定义,以理清“跨媒体”研究与“多媒体”研究的不同:

Nowadays, there are lots of heterogeneous and homogeneous media data from multiple sources, such as news media websites, microblog, mobile phone, social networking websites, and photo/video sharing websites. These media data are integrated together to reflect our real-world. Consequently, it is impossible to conceive our real-world without exploiting the data available on these different sources of rich multimedia content simultaneously.

Cross-media is a research area in the general multimedia field which focuses on the utilization of data with different modalities from multiple sources to discover knowledge and understand the real-world.

Specifically, we emphasize two essential elements in the study of cross-media that help differentiate cross-media from the rest of the research in the general field of multimedia.

The first is the simultaneous co-existence of data from two or more different data sources. This element indicates the concept of “cross”, e.g., cross-modality, cross-source, and cross cyberspace to reality. Cross-modality means heterogeneous

features are obtained from data in different modalities; cross-source means the data may be obtained across multiple sources (domains or collections); cross-space means the virtual world (cyberspace) and the real world (reality) complement each other.

The second is the leverage of different kinds of data across multiple sources for discovering knowledge, for example, discovering the (latent) correlation between data with different modalities across multiple sources, borrowing the strength from data with different modalities to better understand cross-media, generating a summary with data from multiple sources.

Cross-media learning is beneficial for many applications in data mining, cross-media retrieval, causal inference, and public security.

There two essential elements help promote cross-media as a new, emerging, and important research area in today's multimedia research.

五、973 项目“面向公共安全的跨媒体计算理论与方法”

2012 年年度进展整体介绍

2012 年度围绕“跨媒体数据统一表示与建模方法”、“跨媒体数据关联推理与深度挖掘”、“跨媒体数据综合搜索和内容合成”等关键科学问题，从科学问题解决、关键技术攻关、实验数据获取等几个方面着手，在跨媒体数据统一表示和建模机制、跨媒体属性感知模型与行为计算、跨媒体语义学习与内容理解、海量跨媒体数据挖掘与公共安全态势分析、跨媒体搜索与内容整合和面向公共安全的跨媒体呈现与验证和示范平台等方面进行了深入研究，取得了一批研究成果，已完成该年度计划的各项预期目标。

项目组在显著性基元特征提取、基于向量场的跨媒体表达、跨媒体关联分析、跨媒体一致性属性获取、稀疏结构特征选择、跨媒体结构语义理解、跨媒体搜索与排序、跨媒体跨域内容整合以及跨媒体因果推断等方面取得若干突破性进展。

2012 年项目启动以来，已发表学术论文 123 篇，其中在 T-PAMI、T-IP、T-KDE、T-CSVT、T-MM、Pattern Recognition、Artificial Intelligence 等国际权威期刊上发表论文 31 篇、在 ACM Multimedia、CVPR、NIPS 等计算机学会 A 类会议上发表论文 15 篇，并获得国家发明专利 23 项。

项目组赵耀教授领衔“数字媒体信息处理”教育部创新团队(2012 年度)、李学龙研究员、何晓飞教授获得 2012 年度国家自然科学基金杰出青年基金资助，於志文教授获得 2012 年度国家自然科学基金优秀青年基金资助，於志文教授获得 2012 年度中国计算机学会青年科学家奖。

杨易获得 2012 年度教育部全国百篇优秀博士论文（博士论文题目：跨媒体检索与智能处理关键技术研究，导师：潘云鹤院士、庄越挺教授）、韦世奎（博士论文题目：基于信息融合的多媒体内容搜索，导师：赵耀教授）和韩亚洪（博士论文题目：基于图模型表达和稀疏特征选择的图像语义理解，导师：庄越挺教授）均获得中国计算机学会优秀博士论文。

“大规模网络视频内容分析关键技术及应用”成果获得 2012 年度中国计算机学会科学技术奖（获奖人：黄庆明教授、蒋树强副研究员等）。

整体而言，项目已形成了结构合理、有较强攻关能力的稳定研究团队，这为本项目后续研究工作的深入奠定了坚实基础。

六、国际相关学术热点和动态

1) 群体参与的计算模式成为当前一种计算基本范式变革

美国国家科学基金会于 2012 年 3 月资助加州大学伯克利分校一个经费总额为 1000 万美元、名为“计算之探究 (Expeditions in Computing)”的项目，这一项目包括了来自伯克利计算机系和统计系等知名教授，主要负责人 (Principal Investigator) 伯克利分校计算机系 Michael J. Franklin 教授和联合负责人 (Co-Principal Investigators) Michael Jordan 教授等人指出：该项目将从算法、架构和众包 (Crowdsourcing) 三个方面的有机集成来推进大数据研究，即设计新的大规模机器学习算法、研制支持并行和集群计算能力的软硬件基础架构以支持大规模数据计算、引入群体智慧计算来弥补单纯依赖机器来提高传统计算的“人工智能”。“计算之探究”项目认为，这三个方面的有机集成将彻底革新大数据时代的计算模式，为我们处理大数据提供基础性软件平台、硬件架构和核心算法。

这一观点也得到了学术界的呼应，如 AAAI 于 2012 年举行了名为“Wisdom of the Crows”的 Workshop，涵盖了 Reward Strategies、Integration of Crowdsourcing and AI、Game Theory 等主题；美国斯坦福大学人机交互研究所 Michael Bernstein 博士提出了“Crowd-Powered Systems”的概念；美国东北大学研究人员在 IEEE computer 上面发表了“The Future of Human-in-the-Loop Cyber-Physical Systems (2013,46(1):36-45)”的论文，提出建立“Human-in-the-Loop”的计算模式。

2) 加州大学伯克利分校举办 2013 年度 Theoretical Foundations of Big Data Analysis 会议

2013 年 8 月 22 日到 12 月 20 日之间，加州大学伯克利分校将举办“Theoretical Foundations of Big Data Analysis”的会议，这一会议将分为如下四个专题进行：Succinct Data Representations and Applications、Parallel and Distributed Algorithms for Inference and Optimization、Unifying Theory and Experiment for Large-Scale Networks、Big Data and Differential Privacy。

3) 谷歌公司加强深度学习理论研究与应用

识别与分类是机器学习的终极目标之一，也是极具挑战的难点问题。纽约时报 2012 年 8 月发表了篇名为“多少台机器才能识别一只猫?(How many computers to identify a cat? 16,000)”的文章，该文章指出：谷歌公司和斯坦福大学研究人员使用了 16,000 台计算机，通过对 13,026 张猫的图像以及 23,974 不是猫的图像进行“机器学习”，成功让计算机学会了如何识别“猫”！。据报道，相关研究

人员表示：在互联网时代，只要拥有足够多的数据，总是可以通过特征学习 (Feature Learning) 的方法，学习高层语义在底层特征上表现出来的模式 (Pattern)，克服底层特征和高层语义之间的“鸿沟”。深度学习这一方向的领军人物、加拿大多伦多大学计算机科学教授 Geoffrey Hinton 于近日加盟谷歌，一方面推动谷歌“拓展现实”方面应用——谷歌眼镜 (Google Project Glass) 的进展，同时将加强谷歌在知识图谱 (Knowledge Graph) 方面的研究。