

基于概念对关系的视频多语义概念学习方法研究

陈丹雯¹ 刘 然³ 袁志民¹ 邓莉琼¹ 吴玲达^{1,2}

(国防科技大学信息系统工程重点实验室 长沙 410073)¹ (装备指挥技术学院 北京 100000)²

(中国电子设备系统工程公司 北京 100000)³

摘 要 多语义概念学习是视频检索的重要支持技术。针对此问题提出了基于概念对关系的视频多语义概念学习方法。首先分析大规模语义概念之间存在的概念对关系类型,并根据 TRECVID2005 标注数据以及 Columbia374 数据对其进行定量分析;然后利用概念对关系进行基于上下文关系语义概念探测器的相关概念选择,并根据探测器可靠性对相关概念赋予权重,最后根据视觉相似性和单语义概念探测器进行融合。实验证明,该方法能够取得较好的语义概念探测性能。

关键词 多语义概念学习,概念对关系,语义概念探测

中图分类号 TP391 文献标识码 A

Research on Multi-concept Learning Based on Inter-concept Relation

CHEN Dan-wen¹ LIU Ran³ YUAN Zhi-min¹ DENG Li-qiong¹ WU Ling-da^{1,2}

(Science and Technology on Information Systems Engineering Laboratory, National University of Defense Technology, Changsha 410073, China)¹

(Institute of Equipment Command & Technology, Beijing 100000, China)²

(China Electronic System Engineering Company, Beijing 100000, China)³

Abstract Multi-concept learning is a very important technology of video retrieval. This paper gave a method of multi-concept learning based on inter-concept relation. First, analysed the types of inter-concept relation in large-scale multimedia lexicon, then got quantitative analysis with TRECVID2005 annotation data and Columbia374 data. Second, selected correlative concepts to build context based conceptual fusion detectors, then got weights with reliability of detectors. At last, fused the two detectors with visually similarity. The results show that the proposed method achieves more remarkable and consistent improvement.

Keywords Multi-concept learning, Inter-concept relation, Concept detectors

1 引言

随着计算机和网络技术的快速发展,视频成为人们共享信息的重要载体,普通用户接触到越来越多的视频数据。相比于文本、语音等媒体,视频声像并茂,具有表现力强、蕴含信息量大、形象生动等优点,能够提供更多有价值的信息。但是视频媒体也具有非结构化的数据格式、巨大的数据量以及表现内容不透明等缺点,使得对视频数据的管理和分析(如视频数据的浏览、检索)面临相当大的难度。面对海量的视频数据,如何对视频进行语义标注,有效地组织和管理视频数据,以实现快速准确的存取,尽可能满足用户的查询需求,已经成为多媒体研究领域一项重要的研究课题。

以视频检索为例,发展数十年的基于内容的视频检索遇到的最大瓶颈在于底层特征(颜色、纹理等)与用户语义检索需求之间存在“语义鸿沟”(Semantic Gap)。而视频语义标注

为解决“语义鸿沟”问题提供了一种分而治之的方法。Hauptmann^[1]指出,视频语义标注可以将底层特征和用户需求之间的鸿沟切分为两个较小的鸿沟,即(1)底层特征与语义概念之间的鸿沟;(2)语义概念与用户需求之间的鸿沟。视频多语义概念学习,要解决的就是底层特征与语义概念之间的映射问题。

多语义概念学习长期以来都是多媒体内容分析和检索中的重要课题,包括两个关键的组成部分:多媒体语义概念词典和语义概念探测。本文针对语义概念探测问题进行研究,第 2 节介绍了现有的语义概念探测的相关研究工作,并指出现有方法存在的问题;第 3 节详细分析了概念间的联系,并提出了基于概念对关系的语义概念探测框架;第 4 节介绍了利用 TRECVID2005 开发数据在 LSCOM374 和 LSCOM-Lite 两个语义概念集上进行的实验,并对实验结果作出了分析;最后是总结与展望。

到稿日期:2010-05-12 返修日期:2010-08-28 本文受国家 863 高科技计划项目(2009AA01Z335),国家自然科学基金项目(60802080),国家自然科学基金(61002020)资助。

陈丹雯(1983—),女,博士生,主要研究方向为视频语义分析与检索, E-mail: danwen_chen@nudt.edu.cn; 刘 然(1982—),硕士生,主要研究方向为软件工程、图像处理;袁志民(1982—),男,博士生,主要研究方向为监控视频分析;邓莉琼(1986—),女,博士生,主要研究方向为视频关联分析;吴玲达(1962—),女,博士生导师,主要研究方向为多媒体信息系统与虚拟现实技术。

2 相关工作

要进行语义概念探测,首先必须根据人的先验知识定义出明确的语义概念集(也称为语义概念词典),并且使概念集中的每个概念都有正例标注的数据。一般而言,语义概念根据以下几个标准进行挑选:可用性、可观察性以及可行性。卡内基·梅隆大学、哥伦比亚大学、阿姆斯特丹大学以及 IBM 研究所做了很多相关工作,先后发布了多个语义概念集,并在 TRECVID2005 开发数据集上进行了标注,包括 MediaMill, LSCOM-Lite, LSCOM374。这些规范的语义概念集的发布为视频语义概念探测研究提供了公共的数据集以及良好的评测平台。本文即使用 LSCOM-Lite, LSCOM374 语义概念集展开实验。

现有的视频语义概念探测研究大多围绕以上概念集展开,归纳起来包括两类方法:单语义概念探测方法(Individual Concept Detection, ICD),以及基于上下文的多概念融合方法(Context Based Conceptual Fusion, CBCF)。ICD 即每个语义概念是在不考虑相互关系的情况下分别进行探测的。多个研究机构也先后发布了基准 ICD 以便相关领域的研究者更快地开展研究,包括 MediaMill101, Columbia374, VIREO374 以及 CU-VIREO374。本文使用的是 Columbia374 提供的 ICD。实际上,ICD 建立的模型效果并不是很好,因为现实很多事物不是独立出现的,它们都存在着相互的关联。CBCF 的基本思路就是利用概念关系、概念属性等上下文信息对独立概念探测结果进行融合,从而提升探测性能。典型的工作包括 IBM 的 TJ Waston 实验室^[2]最早在 TRECVID 2003 评测中开展的利用概念间关系提高概念探测性能的研究工作,并提出了判别模型融合方法(Discriminative Model Fusion, DMF);哥伦比亚大学的 S-F Chang 团队^[3]对 DMF 方法做了进一步的改进,提出了 CBCF 概念融合框架,并在此基础上利用主动学习(active learning)方法设计出 ACBCF 框架,以减少标注工作量并提高融合性能;微软亚洲研究院的 Hua Xiansheng 团队^[4]提出了一种关联多标注方法(Correlative Multi-Label, CML),该方法的关键在于把概念个体和它们之间的关系统一建模在一个框架中,在概念分类过程中将概念间的关联关系集成到一个概率描述框架中综合考虑,以提高概念探测的准确性。

然而,现有的许多基于 CBCF 的研究工作用到的语义概念集不超过 40 个概念,没有充分探究当概念数量增加时性能如何变化。当大规模的公共标注数据 LSCOM 可用时,我们可以进一步探究这些问题。并且现有的视频多语义概念学习研究往往利用复杂模型构建多概念间的关系,或者利用足够大的语义概念词典构建 CBCF 模型,却很少探讨如何从标注数据中获取概念间的联系进而辅助概念探测。即使有少量相关研究,也仅仅考虑了语义概念间的共现性。而实际上,概念间的关系非常复杂,详细分析概念间的联系,能够为提高多语义概念学习效率提供有益辅助。因此,本文将利用大规模语义概念集 LSCOM,研究在 CBCF 基础上,如何有效利用语义概念间联系,将其纳入语义概念探测框架,以提高语义概念探测性能。

3 基于概念对关系的多语义概念学习

3.1 语义概念对关系分析

所谓语义概念对关系,指的是两两语义概念间的联系。语义概念间的联系具有高复杂性,可以利用多元关系对其进行度量。但多元关系不仅计算量大,而且标注数据有限,可能导致过拟合问题。而二元关系最易度量,且能够较大程度上反映出变量间的关系。因此,本文着重考虑利用概念对关系来辅助构建语义概念探测模型,并利用二元关系度量概念间的联系。

大多数研究者在考虑概念间关系时仅仅利用了语义概念间的共现特性,忽略了概念间联系的其他特性。为了能够更好地衡量概念间的联系,本文定义了 3 类概念对关系特性,包括视觉相似性、共现性、互斥性,并根据 TRECVID2005 的标注数据以及 Columbia374 提供的数据对其进行了定量计算。下面对 3 类概念关系做进一步的阐述。

• 视觉相似性

视觉相似性(Visually Similarity, VS)特指基于底层特征的相似性,因为现有的视觉描述子还无法有效地和人的视觉感知相统一。可以利用多种底层特征来衡量视觉相似性,统计语义概念探测器在训练数据集上得到的预测结果,根据下式得出两个语义概念视觉上的相似度:

$$rv_{c_m c_n} = rv(c_m, c_n) + rv(c_n, c_m) \quad (1)$$

$$rv(c_m, c_n) = \sum_{i=1}^N b(y_i = c_n), m, n \in \{1, 2, \dots, M\}$$

式中, N 是整个数据集中标注为概念 c_m 的数据的总和。 $rv_{c_m c_n}$ 表示语义概念 c_m, c_n 在视觉上的相似度, $rv(c_m, c_n)$ 表示 N 个数据中被预测为概念 c_n 的数量, $b(y_i = c_n)$ 是一个二值函数;反之 $rv(c_n, c_m)$ 亦然。需要注意的是, $rv(c_m, c_n) \neq rv(c_n, c_m)$ 。若 $rv_{c_m c_n}$ 大于一定阈值,则说明两个语义概念在视觉上相似。

• 共现性

语义概念的共现性(Occurrence, OC)是现有研究中用到最多的概念间的关系。两两概念的共现性指的是两个概念经常会同时出现在同一视频段中,比如“车”、“马路”同时出现在视频里的概率比“车”、“飞机”同时出现的概念要高得多。这样的信息能够为预测语义概念是否出现提供很有价值的线索。本文利用数据集中的标注信息得到两两概念的共现度,根据下式计算得出:

$$ro_{c_m c_n} = ro(c_m, c_n) + ro(c_n, c_m) \quad (2)$$

$$ro(c_m, c_n) = \frac{\sum_{i=1}^N V_i^m V_i^n}{\sum_{i=1}^N V_i^m}, m, n \in \{1, 2, \dots, M\}$$

式中, N 是整个数据集的大小。 $ro_{c_m c_n}$ 表示语义概念 c_m, c_n 的共现度。对于任意概念 c_m , V^m 表示一个向量,向量中的每一维代表数据集中的每个数据对应概念 c_m 的标注信息。若第 i 个数据的标注中包含概念 c_m ,则 $V_i^m = 1$,反之 $V_i^m = 0$ 。若然, $ro(c_m, c_n) \neq ro(c_n, c_m)$ 。若 $ro_{c_m c_n}$ 大于一定阈值,则说明两个语义概念具有共现性。

• 互斥性

语义概念的互斥性(Mutually Exclusive, ME)和共现性

是对立的概念。两两概念间的互斥性指的是两个概念几乎不会出现在同一视频片段中。利用这种互斥性,也可以很好地辅助判断语义概念是否出现。定量衡量方面,可以根据共现度的公式,当 $ro_{c_m, n}$ 小于一定阈值时,说明两个语义概念具有互斥性。本文利用的针对 LSCOM374 语义概念集进行标注的数据 TRECVID2005 中,有大量的语义概念共现度为 0,因此直接定义共现度为 0 的语义概念具有互斥性。

3.2 基于概念对关系的语义概念探测框架

利用概念对关系构建概念探测器,是在 CBCF 框架上做进一步的改进。CBCF 框架是在不考虑概念间的共现性、互斥性等联系的情况下,直接利用语义概念集中所有概念来构建探测模型。但是,如果不考虑概念间联系,不仅会增加探测器模型的复杂度,而且使用不相关的概念也很可能会增加预测结果的不确定性。因此,谨慎选择相关语义概念来构建基于上下文关系的语义概念探测器,对提高探测器性能能够起到积极的作用。

实际上,一些概念对的相互关系非常微弱,既不是共现关系,也不存在互斥关系。也就是说,在视频片段中,一个概念出现与否似乎对另一个概念不起作用,即它们的存在几乎是相互独立的。

基于以上考虑,本文只把具有较强相关性的概念对(具有共现性、互斥性的概念对)纳入学习框架中。当然,概念对的选择可以通过专家人工定义,也可以通过数据驱动的方式。

根据上述分析,本文设计了如图 1 所示的新闻视频语义概念探测框架,包括基础语义概念探测器的构建、相关概念选择、基于上下文关系的语义概念探测器的构建以及最后的融合步骤。基础语义概念探测器的构建不是本文的重点,本文采用 Columbia374 提供的概念探测模型作为基础语义概念探测器。

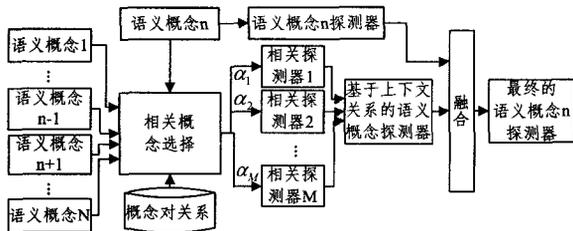


图 1 基于概念对关系的视频语义概念探测框架图

其中,相关概念选择规则通过 3.1 节中的语义概念对关系分析得来,即相关概念的选择遵循以下条件:

- 1)若概念 c 与概念 n 具有共现性,则概念 c 为概念 n 的相关概念;
- 2)若概念 c 与概念 n 具有互斥性,则概念 c 为概念 n 的相关概念。

在上述规则的基础上,本文进一步考虑了基础语义概念探测器的性能高低对构建基于上下文关系的探测器的影响。比如“人”和“沙漠”对“军事基地”的探测会有帮助,而“人”的探测准确度高,因此“人”的探测器的权重相对更高。

其中, $\{\alpha_1, \alpha_2, \dots, \alpha_M\}$ 是对应赋予各个相关概念探测器的权重, M 是相关概念的总数。 α 的值根据以下公式进行归一化处理:

$$\alpha_m = \frac{q_m}{\sum_{i=1}^M q_i}, m \in \{1, 2, \dots, M\} \quad (3)$$

式中, q_i 对应第 i 个相关单概念探测器的准确率。

由于具有共现性的概念对用到的训练数据具有高度重合性,而本文中用到的特征提取方法是针对全局考虑的,因此具有共现性的概念在很大程度上会具有视觉相似性。然而,不仅具有共现性的概念对具有视觉相似性,根据 3.1 节中对视觉相似性的分析,不具有共现性的概念对,也有很大一部分具有视觉相似性。出现这种现象的原因,主要是现有的底层特征无法很好地区分这样的概念对。因此,针对这样的概念,更依赖于基于上下文关系的语义概念探测器,而不是其自身的探测器。根据以上分析,在融合步骤中考虑上述因素,最终的语义概念探测结果根据以下公式得到:

$$P_F(y_i=1) = \lambda_i P_C(y_i=1) + (1-\lambda_i) P_I(y_i=1) \quad (4)$$

$$\lambda_i \propto \frac{1}{(VS_{c_i} - OC_{c_i})}$$

式中, $P_F(y_i=1)$ 表示对第 i 个概念 c_i 最终的预测结果, $P_C(y_i=1)$ 表示基于上下文关系的语义概念探测器的预测结果, $P_I(y_i=1)$ 表示独立语义概念探测器的预测结果。 VS_{c_i} 表示与概念 c_i 具有视觉相似性的概念的数量, OC_{c_i} 表示与概念 c_i 具有共现性的概念数量,它们的差表示与概念 c_i 发生具有视觉相似性但不具有共现性的概念的数量,也就是容易与概念 c_i 误判的概念数量。 λ_i 与其成反比,说明若易与概念 c_i 发生误判的概念数越多,则概念 c_i 的独立概念探测器的可靠性越低,因此赋予的权重也应越低,反之亦然。

4 实验结果与分析

4.1 实验数据

本文实验数据采用 TRECVID 2005 开发数据集,包含编号为 141-277 的共 137 个视频片段,共有 61901 个镜头。训练数据及测试数据的分布如表 1 所列。

表 1 实验数据分布情况

	所含视频片段	所含镜头数量
训练数据	141-220	36078
测试数据	221-277	25823

实验采用两个语义概念集进行相关概念选择,包括 LSCOM374 与 LSCOM-Lite。

LSCOM374 共包含 374 个语义概念。1.0 版本在 2006 年 3 月正式发布, LSCOM 的目标在于建立数量级为 1000 的语义概念集。1.0 版本中包含 856 个语义概念,其中 449 个给出了详细的定义。这些概念分为 6 个大类:对象、行为/事件、场景/地点、人物、图表、节目类型。而后, LSCOM 概念词典发展到包含 2600 多个概念,不过大部分概念在 TRECVID 2005 视频数据中没有正例标注或者只有很少的标注正例,其中的 374 个概念在此数据集中具有一定数量 (>10) 已标注的正例。

LSCOM-Lite 包含 39 个语义概念,它们可以看成是 LSCOM 的子集。 LSCOM-Lite 中的语义概念都包含在完整的 LSCOM 概念集中。这个概念集由 NIST 机构利用 TRECVID2005 数据集提出,并在 TRECVID 2005 评测中发布。

针对语义概念集 LSCOM-Lite 包含的 39 个语义概念的探测器性能进行了实验分析,39 个语义概念如图 2 所示。

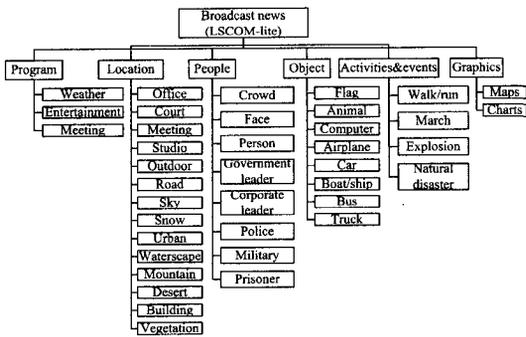


图2 LSCOM-Lite所包含的不同类别的39个语义概念

4.2 实验设置与评价标准

本文实验采用的单语义概念探测器是哥伦比亚大学发布的Columbia374探测器^[8]。对LSCOM-Lite中的39个语义概念进行分类,实际上是一个多类分类问题。实验中,多类分类器是采用多个两类SVM分类器通过一对多的方式构造的。具体来说,对每一个语义概念类别都学习来得到一个区分它和其他类别的SVM分类器。分类器模型的训练工作都在训练数据上进行,而语义概念探测则在测试数据上进行。

由于CBCF框架中采用真实标注构建分类器模型是否能够得到更好的分类效果仍存在争议,因此本文在LSCOM-Lite语义概念集上利用置信度值以及真实标注分别构建模型,并比较其预测准确率,得到的结果如图3(a)所示。

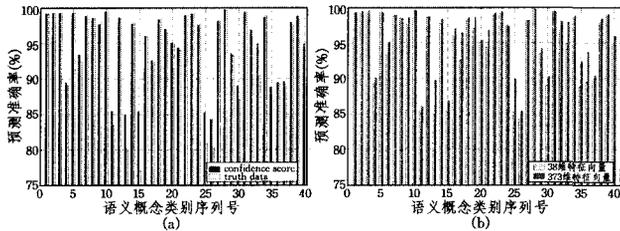


图3 标注模型与置信度模型的预测准确率比较(LSCOM374与LSCOM-Lite语义概念集上得到的基准CBCF探测器性能比较)

图3(a)中,1-39维代表的是39个语义概念分类模型分别对应的预测准确率,27个语义概念对应的置信度模型比标注模型得到的准确率更高。其中,Face高出8.4%,Entertainment高出4.7%,Outdoor高出4.4%,Studio高出4.3%。最后一维对应的是39个语义概念的平均预测准确率,置信度模型比标注模型高0.9%。因此,本文采用置信度模型进行实验。

采用置信度模型分别在LSCOM374和LSCOM-Lite两个语义概念集上,利用CBCF框架^[3]得到基准语义概念探测器。即在LSCOM374上采用373维的特征向量进行探测器模型的训练,在LSCOM-Lite上采用38维的特征向量进行探测器模型的训练。得到的基准探测器预测准确率如图3(b)所示。由图可以看出语义概念集的大小对探测器性能的影响:语义概念集更大,能够对探测器性能起到积极作用。在LSCOM374上得到的探测器有36个,多于LSCOM-Lite上得到的探测器。其中,Entertainment高出4.8%,Outdoor高出4.6%,Vegetation高出4.2%,Government-Leader高出3.8%,Urban高出3.5%,而平均预测准确率高出1.1%。这两组基准探测器作为后续实验的比较对象。

评价分类性能的方法很多,本文实验中将采用分类预测准确率即分类精度来作为评价标准:

$$\text{分类精度} = \frac{\text{分类正确的样本数目}}{\text{样本总数}}$$

分类精度是通用的衡量分类准确率的评价标准,本文考虑视频多语义概念学习,旨在解决视频检索中底层特征与语义概念之间的映射问题。

4.3 实验结果分析

本文的语义概念探测方法主要包含相关概念选择、探测器赋予权重以及单语义概念探测器与基于上下文的语义概念探测器的融合。因此,实验分为两大部分:第一部分验证利用相关概念选择来构造CBCF探测器是否能够提高探测器性能,第二部分验证综合考虑单语义概念探测器可靠性以及视觉相似性能否提高探测器性能。

4.3.1 优化相关概念选择数量

由于在相关概念选择中涉及语义概念个数的选择问题,在本实验中首先优化相关概念数量的选择。具体来说,在给定不同的相关概念个数的情况下,分别得到基于上下文关系的语义概念探测器,并利用测试数据得到相应的预测准确率。

图4给出了两个语义概念集在给定不同相关概念个数时,对探测器的分类准确率的影响情况,LSCOM-Lite用到的是(5,10,15,20,25,30,35),LSCOM用到的是(20,40,60,80,100,120,140,160,180,200)。

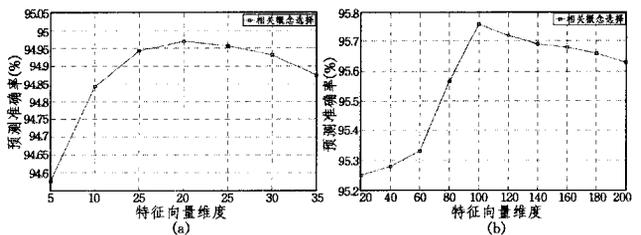


图4 不同相关概念数量时性能比较

从图4(a)可以看出,当采用LSCOM-Lite语义概念集时,相关概念个数达到20,探测器的平均分类准确率达到最大;图4(b)中,采用LSCOM374语义概念集时,相关概念个数到达100,探测器的平均分类准确率达到最大。相关概念个数过大,反而会降低分类器性能。验证了使用不相关的概念不仅会增加探测器模型的复杂度,而且会增加预测结果的不确定性,往往出现过拟合的现象。

为了更好地比较本文方法和其他方法之间的性能,在实验中实现了Xie Lexing^[5], Wei Shikui^[6]提出的方法,并与本文方法进行性能比较。文献[5]采用的是随机选择,文献[6]采用的是共现性,得到的结果如图5所示。

图5(a)所示为采用LSCOM-Lite语义概念集时的性能比较。随机选择相关概念时,预测准确率随着特征维度的增加而增加;采用共现性选择相关概念时,当相关概念个数为15,性能达到最大;随后特征维度增加,性能略有下降。而用本文方法选择相关概念时,特征维度为20,性能达到最大,并且是所有方法中性能最高的,这在一定程度上反映出本文方法能够选择出更有益于构建CBCF探测器的相关概念。图5(b)所示为采用LSCOM语义概念集时的性能比较。几类方法的情况与LSCOM-Lite大致相同,仍然是本文方法达到最高性能。并且在较大语义概念集上可以看出,本文方法能够更快地达到收敛,即本文方法在100维特征时即达到最大预测准确率,而利用共现性选择时160维特征才达到最大准确率。

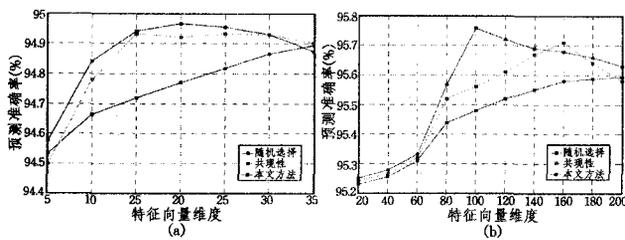


图5 与随机选择、共线性选择相关概念的性能比较

4.3.2 结合单语义概念探测器准确率以及融合后的性能比较

根据上述实验可以看出,在LSCOM-Lite概念集上,相关概念个数为20时,性能达到最佳;而在LSCOM374概念集上,相关概念个数为100时,性能达到最佳。以下实验即采用给定的特征向量维数,并以探测器detector39, detector374为基础,根据单语义概念探测器的性能优劣对相关概念根据式(3)赋予权重,重新训练,得到detector39_acu, detector374_acu;最后根据式(4)融合单语义概念探测器,得到探测器detector39_fus, detector374_fus。在39个语义概念上的预测准确率以及平均预测准确率如图6所示。

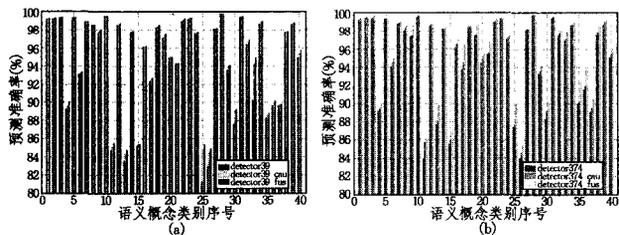


图6 综合单语义概念探测器准确率以及融合后的性能比较

如图6所示,前39维分别对应39个语义概念的预测准确率,第40维对应的是平均预测准确率。由图可以看出,平均预测准确率高低为detector39_fus > detector39_acu > detector39, detector374_fus > detector374_acu > detector374。并且39个语义概念分别的准确率高低情况也是如此,充分反映出本文提出的结合单语义概念探测器准确率以及将基于上下文关系的语义探测器和单语义概念探测器相融合对于提高语义概念探测器性能是有益的。同时可以认为,结合单语义概念探测器准确率即考虑了其可靠性,能够抑制准确率不高的单语义概念探测器给CBCF带来的负面影响;另一方面,融合单语义概念探测器,能够在一定程度上弥补某些语义概念与其他概念关系较弱,或其本身单概念探测器性能较差时导致的问题。

结束语 本文主要研究如何利用语义概念对关系来提高

语义概念探测器的性能。首先详细分析了语义概念之间各类联系,给出了定量计算方法,在此基础上设计了相关概念选择策略,综合考虑了语义概念探测器准确率以及视觉相似性,并对基于上下文关系的语义概念探测方法进行了改进,提出了基于概念对关系的语义概念探测框架。在TRECVID2005开发的数据集上进行了大量的实验验证,本文方法构建的语义概念探测模型有较好的预测准确率,与现有的方法相比,预测准确率更高。并且由于做了相关概念选择,即特征的降维,也相应地降低了探测模型的复杂度并提高了计算效率。

由于本文提出的框架基于判决式模型,因此语义概念间的层次关系无法很好地纳入此框架中,且语义概念间同时也存在大于二元关系的更高阶的多元关系。因此,在以后的研究中,如何利用语义概念的层次关系以及多元关系将是工作重点。

参考文献

- [1] Hauptmann A G. Lessons for the future from a decade of informedia video analysis research[C]// Proceedings of ACM International Conference on Image and Video Retrieval (ACM CIVR). Singapore, July 2005
- [2] Lyengar G, Nock H J, Neti C. Discriminative model fusion for semantic concept detection and annotation in video[C]// Proceeding of ACM International Conference on Multimedia (ACM MM). Berkeley, California, USA, 2003
- [3] Jiang W, Chang S-F, Loui A. Active concept-based concept fusion with partial user labels[C]// Proceedings of IEEE International Conference on Image Processing. Atlanta, GA, USA, 2006
- [4] Qi Guo-jun, Hua Xian-sheng, Rui Yong. Correlative Multi-Label Video Annotation[C]// Proceeding of ACM International Conference on Multimedia (ACM MM). Augsburg, Bavaria, Germany, 2007
- [5] Xie Le-xing, Yan Rong, Yang Jun. Multi-concept Learning with Large-Scale Multimedia Lexicons [C]// Proceedings of IEEE International Conference on Image Processing. San Diego, California, USA, 2008
- [6] Wei Shi-kui, Zhao Yao, Zhu Zhen-feng. Ontology-based Inter-concept Relation Fusion for Concept Detection[C]// Pacific-Rim Conference on Multimedia (PCM). Taiwan, 2008
- [7] 陈颀,朱福喜. 基于支持向量机的两阶段模糊聚类在视频检索中的应用[J]. 计算机科学, 2009, 36(6): 227-230
- [8] Yanagawa A, Chang S-F, Kennedy L, et al. Columbia University's Baseline detectors for 374 LSCOM Semantic Visual Concepts[R]. 222-2006-8. Columbia University ADVENT, March 2007

(上接第243页)

- [13] Xiu Ming, Shi Kai-quan, Zhang Li. P-sets and \bar{F} -data selection-discovery[J]. Quantitative Logic and Soft Computing, 2010, 2(1): 791-799
- [14] Zhang Li, Xiu Ming, Shi Kai-quan. P-sets and applications of power circle[J]. Quantitative Logic and Soft Computing, 2010, 2(1): 581-591
- [15] Lin Hong-kang, Li Yuying. P-sets and its P-separation theorems [J]. An International Journal Advances in Systems Science and Applications, 2010, 10(2): 209-215
- [16] Huang Shun-liang, Wang Wei, Geng Dian-you. P-sets and its internal P-memory characteristics [J]. An International Journal

Advances in Systems Science and Applications, 2010, 10(2): 216-222

- [17] Wang Yang, Geng Hong-qin, Shi Kai-quan. The mining of dynamic information based on P-sets and its applications[J]. An International Journal Advances in Systems Science and Applications, 2010, 10(2): 234-240
- [18] Zhang Guan-yu, Li En-zhong. Information gene and identification of its information Knock-out/Knock-in[J]. An International Journal Advances in Systems Science and Applications, 2010, 10(2): 308-315
- [19] 王亚馥,戴灼华. 遗传学[M]. 北京: 高等教育出版社, 2000: 97-110