

概念间关联依赖多标记视频语义概念分类方法

魏 维¹⁾ 魏 敏¹⁾ 刘凤玉²⁾

¹⁾(成都信息工程学院计算机学院, 成都 610225) ²⁾(南京理工大学计算机科学与技术学院, 南京 210094)

摘 要 一个镜头中的语义概念通常依赖于其他多个语义概念,几个同时出现的语义概念决定着另一个语义概念的出现。为此提出一种概念间关联依赖多标记视频语义概念分类方法。为得到概念间关联依赖规则,合并和修剪技术用于产生候选的项集;计算各候选项集的支持度后,得到满足最小支持度的频繁项集;经过一系列频繁项集迭代,产生具有强关联依赖关系的复合标记;在标记过程中,将具有强关联依赖关系的多个语义标记作为单标记进行标注。实验结果表明,对真实媒体数据本文方法比现有多标记分类方法更能有效进行分类。

关键词 视频概念标注 多标记分类 关联规则 语义场景分类

中图法分类号: TP391 文献标志码: A 文章编号: 1006-8961(2010)06-893-07

Inter-concepts Association and Dependency Multi-label Video Semantic Concept Classification Approach

WEI Wei¹⁾, WEI Min¹⁾, LIU Fengyu²⁾

¹⁾(College of Computer Science and Technology, Chengdu University of Information Technology, Chengdu 610225)

²⁾(College of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing 210094)

Abstract In video data, one concept in one shot are usually dependent on others concepts. Several semantic concepts appearing in one time often determine the presence of other concepts. An inter-concepts association and dependency multi-label video semantic concept classification approach is proposed in this paper. In order to generate association and dependency relation between concepts, join and prune phases are used to extract potential itemsets. After calculating the minimum support of each itemset, frequency itemsets meeting the user specified minimum support are selected. In the iteration process of generation frequency itemsets, compound labels with strong association and dependency relation of inter-concepts are obtained. Finally, compound labels are considered as a single label in the annotation step. Experiments on real-world multi-label media data show that this method beat accuracy of existing multi-label learning methods with statistically significant improvements.

Keywords video concept annotation, multi-label classification, association relation, semantic scene classification

0 引 言

随着视频内容理解技术不断深入研究,大规模语义概念的分类与标注受到越来越多的关注。多标记语义分类在现实世界中普遍存在,真实视频中

一个镜头多标记概念更常见。在多标记分类中,每一个样本都有几个标记,比如一个场景常对应于多个预先定义的视频语义概念,一个关键帧的内容同时为海滩和城市^[1]等。如何准确地将这样的场景和关键帧同时划归多个概念/主题,是视频内容理解中一个困难而重要的问题。

基金项目:成都信息工程学院引进人才启动科研项目(KYTZ200914);成都信息工程学院发展基金项目(CSRF200803);四川省教育厅青年基金项目(2006B063)

收稿日期:2008-12-23; **改回日期:**2009-02-27

第一作者简介:魏维(1976—),男,副教授。2006年于南京理工大学获博士学位。主要研究方向为视频语义分析、多媒体信息处理。

E-mail: weiwei863@hotmail.com

目前视频/媒体内容标记研究主要集中在单标记分类。在视频语义概念理解与分类领域,通常使用两种方法来识别语义概念。第 1 种是单语义概念分类与标注,即使用预先训练好的单标记分类器来对每个概念进行识别,如 Tang 在文献[2]中提出一种基于图的敏捷结构流形排序方法(SSMR)进行语义概念探测。单标记分类方法将每个语义概念看做彼此独立,忽略了视频语义概念间的依赖关联关系,因此单语义概念分类仅对一些简单独立的概念有效。第 2 种是基于多个单标记分类的融合方法。此类方法分为两步骤,第 1 步与单标记分类方法类似进行简单语义概念分类识别;第 2 步融合多个简单分类器进行高层语义的识别。这种融合在一定程度上可间接得到语义概念间的关联关系。因为两个步骤是严格的串联关系,其分类性能受第 1 步和第 2 步融合的共同影响,最终结果错误率控制较难。

多标记学习是机器学习界最近的研究热点之一,源于文档分类与处理的研究^[3-4]。研究者已经探索出了两大类多标记分类学习的解决策略^[5]。一种是将多标记分类转换为传统的单标记分类问题,如 Boutella 在文献[1]中提出的“交叉训练”的方法,文献[6]中提出 SML_SVM 算法等。另一种将传统算法做一定的修改,使之能处理多标记问题,例如文献[7]把多标记学习转化为 NMF 问题,Fan 采用传统的二分方法进行多标记分类^[8]。以上两种策略常同时采用于同一种方法中,如 Zhang 等人提出的 ML- k NN 懒惰学习算法^[9]就对传统的 k NN 算法针对多标记数据进行修改,同时又使用转换策略。最近,有学者研究并提出用未标记样本来提高多标记学习性能的方法,如 Liu 等人提出的 CNMF 算法^[10],Jiang 等人提出的直推式多标记文档分类法 TML^[11]。

近年来,适合于其他领域应用的多标记分类技术不断发展,但适合视频内容多标记的分类技术研究相对较少。多标记视频标注的代表方法如 Qi 等人提出的一种能在一步中将多语义概念关系和多个语义概念同时识别的相关多标记(CML)方法^[12]。此方法中概念间语义依赖关系只局限于两个概念对之间,通过不同的语义强关联对之间的相互交叉来间接反映多个语义概念间的关联依赖关系。

一个镜头内容通常与多个语义概念相关(包括高层、中层语义概念),这些出现的视频语义概念间具有强依赖关联关系(同时出现的依赖关系称为正

关系,反之为负关系)。高识别率的单语义分类器构造十分困难,只要能找到语义概念间合理的关联特性,利用语义概念间的联系来进行多标记语义分析比单纯设计多个独立语义识别器来融合进行语义理解效果更好。

单语义概念分类与识别技术研究较多,而如何利用内部概念间关系,考虑多标记分类却关心较少。本文将设计挖掘算法获取多个语义概念间的强关联依赖关系,在分类识别过程中,将具有强关联依赖关系的多个语义概念作为“一体”进行多标记标注,然后进行转换处理。本文方法与文献[12]的共同点是都能同时用单一的步骤提取语义概念间的强关联关系和进行语义概念识别。但也有显著的异同,本文方法可直接得到多个(2个以上)语义概念间的强关联依赖关系,将具有强关联依赖关系多个语义概念作为一个分类符合标记处理,这样更符合真实的模拟视频语义概念间的依赖关联关系。

1 多标记学习

传统的机器学习主要解决单标记学习,即一个样本仅有一个标记。在许多情形中,样本所属的多个类别,即一个给定的样本点可以属于多个类别,这就是多标记学习问题,如图 1 中所示。设 $\mathcal{X} = \mathbf{R}^d$ 为输入空间,给定的训练样本集为 $\{(x_1, \mathbf{Y}_1), (x_2, \mathbf{Y}_2), \dots, (x_m, \mathbf{Y}_m)\}$,其中, $x_i \in \mathcal{X}$, x_i 对应标记为 $\mathbf{Y}_i = (y_1^{(i)}, y_2^{(i)}, \dots, y_{l_i}^{(i)})$, $y_k^{(i)} \in \mathbf{Y}$ ($k=1, 2, \dots, l_i$), l_i 为 x_i 的标记总数。 $\mathbf{Y} = \{L_1, L_2, \dots, L_{|L|}\}$ 为语义空间, $|L|$ 为语义标记的总数目。多标记学习从训练集中构造多标记分类器 $h: \mathcal{X} \rightarrow 2^{\mathbf{Y}}$ 。

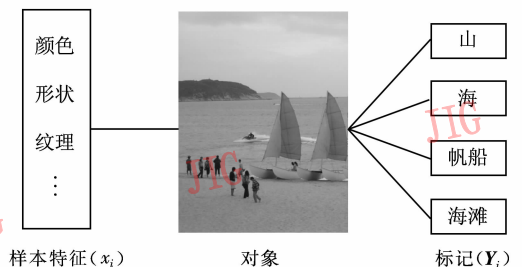


图 1 多标记学习例子

Fig. 1 Example of multi-label learning

求解多标记学习问题的策略分两类:一类是将多标记学习转化为单标记学习分类或回归问题;另一类是将传统单标记算法改造为能直接处理多标记

问题的算法。通常,将多标记学习转化为单标记学习的方法有 6 种^[5]:1) 主观或随机地选择多标记样本的某一个标记为训练标记,而丢弃该样本的其他标记,记为 TM1;2) 丢弃训练集的所有多标记样本,仅保留单标记样本,记为 TM2;3) 将具有相同标记的多标记样本组成一个新单标记类,记为 TM3,此方法的缺点是会导致分类的数目增长过多,而且对转换的新类别,样本将面临不足的境地;4) 最常用的方法是将多标记转换为一对全部的分类,即建立 $|L|$ 个不同的二元分类器 $H_{l_i}: x \rightarrow \{+1, -1\}$,一个二元分类器对应于一个标记,对第 i 个分类器,将类别(标记) i 中的点作为正样本,将不在类别 i 中的所有其他点作为负样本,记为 TM4;5) 根据样本的标记,即将样本 (x_i, y_i) 分解为 $|Y|$ 个样本 (x, l_{ab}) , $(l_{ab} \in Y_i)$,然后学习 $|Y|$ 个基于覆盖的单标记分类器,记为 TM5;6) 将每一个样本 (x_i, y_i) 分解为 $|Y| = D$ 个样本 $(x_i, l_{ab}, Y[l_{ab}])$,如果 $l_{ab} \in Y_i$,则 $Y[l_{ab}] = 1$,否则 $Y[l_{ab}] = -1$,记为 TM6。

2 强依赖关联语义关系多标记学习

在视频语义概念分析中,许多视频语义概念都同时发生在一个镜头中,充分利用语义概念间的这种依赖关联关系,有助于语义概念识别和探测。本文多标记分类方法如图 2 所示,整个多标记分类学习分为强依赖关联关系提取、单标记转换和分类学习 3 个部分。强依赖关联关系提取基于 Apriori 算法,是一个迭代过程。迭代结果为具有强关联依赖关系的复合标记。将复合标记进行标注后,再按 TM4 方法转换为单标记分类方法分类学习^[13],最后即可得到多标记分类结果。与传统学习方法相比,本文利用强关联依赖关系进行多标记识别与分类,能更有效地对镜头语义内容进行识别。

2.1 语义概念间关联关系的形式化描述

项的集合称为项集。包含 k 个项的项集称为 k -项集。项集的出现频率是包含项集的事务数,简称为项集的频率或支持计数。项集满足最小支持度 (minsup),即项集的出现频率大于或等于 minsup 与 D 中事务总数的乘积。如果项集满足最小支持度,则称它为频率项集。

设 $I = \{i_1, i_2, \dots, i_m\}$ 是由 m 个不同的项目组成的集合,给定一个事务数据集 D ,其中的每一个事务 T 是 I 中一组项目的集合,即 $T \subseteq I, T$ 有一个唯

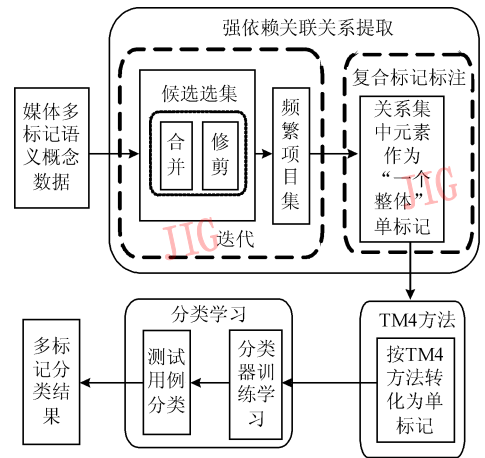


图 2 本文多标记分类方法

Fig. 2 Multi-label classification method in this paper

一的标识符 TID(即镜头编号)。若项集 $X \subseteq I$ 且 $X \subseteq T$,则 T 包含项集 X 。相联规则为 $X \rightarrow Y$,其中 $X \subseteq I, Y \subseteq I, X \cap Y = \emptyset$ 。相联规则 $X \rightarrow Y$ 成立的条件是:1) 它具有支持度 s ,即视频镜头数据库 D 中至少有 $s\%$ 的镜头包含 $X \cup Y$ 语义概念。2) 它具有依赖度置信度 c ,即在视频镜头数据库 D 中包含 x 的语义概念至少有 $c\%$ 同时也包含 Y 语义概念。支持度定义为

$$s = \frac{\text{数据项在数据库 } D \text{ 中出现的次数}}{\text{数据库 } D \text{ 的总数}} \quad (1)$$

依赖度定义为

$$c(X \rightarrow Y) = \frac{\text{数据项 } X \text{ 与 } Y \text{ 同时在数据库 } D \text{ 中出现的次数}}{\text{数据项 } X \text{ 在数据库 } D \text{ 中出现的次数}} \quad (2)$$

视频语义概念间并不是相互独立的,语义语境使许多概念存在相互关联和依赖关系。比如在同一镜头中多个关联的语义概念常同时出现,或一个语义概念出现往往伴随另一个概念出现。例如视频中“轿车”和“道路”语义概念往往同时出现在同一镜头中。当这种关联依赖关系的支持度和依赖度大于给定阈值时,称做强关联依赖关系。关联规则的挖掘问题就是在事务数据库 D 中找出具有用户给定的最小支持度和最小依赖度的关联规则。挖掘关联规则问题可以分解为“找出所有频繁项”和“由频繁项集产生强关联规则”两个子问题。镜头数据中强关联语义依赖关系的抽取过程是一个迭代的过程,每一次经历合并和修剪两个步骤^[14-15],如图 3 所示。

1) 合并 为找 L_k ,通过 L_{k-1} 与自身连接产生候选 k -项集的集合。该候选项集的集合记作 C_k 。

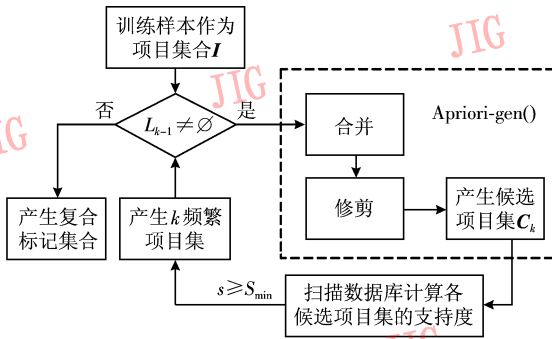


图 3 强关联依赖关系抽取原理图

Fig. 3 Principle chart of strong association and dependency relation extraction

为推导提取过程简便,将视频的语义标记用字母表示,即 A = 海景, B = 船, C = 海滩, D = 山, E = 天空。则表 1 可简化为表 2。设强依赖关联语义关系的支持度为 50%, 则最小支持数目为 2。扫描表 2, 得到候选项目集 C_1 如表 3 所示。项目集 {D} 的依赖关联支持数为 1, 小于 minsup, 所以删除此项, 得到 1-频繁项目集 L_1 如表 4 所示。

表 2 简化的视频数据集

Tab. 2 Simple video data set

	S_1	S_2	S_3	S_4
标记 (Items)	ACD	BCE	ABC	EBE

表 3 候选项目集 C_1

Tab. 3 Candidate itemset of C_1

	项目集 (Itemset)				
	{A}	{B}	{C}	{D}	{E}
支持数	2	3	3	1	3

注: {D}. 支持数 < minsup 删除

表 4 1-频繁项目集 L_1

Tab. 4 1-frequent itemsets L_1

	项目集 (Itemset)			
	{A}	{B}	{C}	{E}
支持数	2	3	3	3

注: $L_1 = \{A, B, C, E\}$

利用 $(k-1)$ -频繁项目集 (L_{k-1}) 来产生 k -候选项目集 (C_k), 其中经历合并和修剪 2 个阶段^[15]。

1) 合并 必须要有 $(k-2)$ 个项目相同, 才能将两个 $(k-1)$ -频繁项目集合并, 组合成 k -候选项目集。

2) 修剪 判断已经合并的 k -候选项目集, 其所有的子项目集集合是否都存在于 $(k-1)$ -频繁项目集中, 成立则保留此 k -候选项目集; 否则删除。

扫描数据库 D , 计算所有 k -候选项目集 (C_k) 的支持度, 若满足设定的最小支持度, 则为 k -频繁项目集 (L_k), 2-频繁项目集 L_2 的产生过程如图 4 所示。

3-频繁项目集 L_3 的产生过程如图 5 所示。对于 L_2 选取到 C_3 , 是为只要 1 项 $(k-2)$ 相同即可, 没有前后顺序差异, 所以在 C_3 合并中 ABC 是被选取的, 但是在修剪中, AB 项目没有在 L_2 中, 跟 ACE 的 AE 项目没有在 L_2 中的理由相同。

强关联依赖复合标记集合导出过程如图 6 所

2) 修剪 C_k 的成员不一定是频繁的, 但是所有的频繁 k -项集都包含在 C_k 中。扫描数据, 由 C_k 中每个候选的数目来确定 L_k 。

2.2 视频语义强依赖关联关系提取算法

给定视频数据集 $\{A_1, A_2, \dots, A_n\}$ 和关联依赖最小支持度 minsup, 找出强关联依赖符合标记集合 $S_{assoc.}$ 的算法如下:

procedure Apriori-association

$L_1 = 1$ -频繁项目集

for $(k = 2; L_{k-1} \neq \emptyset; k++)$ do begin

$C_k =$ 从 L_{k-1} 产生候选集

For each A_i do begin

$C_i = subset(C_k, A_i);$

For all candidates $c \in C_i$ do

$c.count++;$

end

$L_k = \{c \in C_k | c.count \geq minsup\}$

end

$S_{assoc} = \{s | s \in L_i, s \notin subset(c, c \in L_{i+1})\} (i = 2, \dots, k)$

end procedure

2.3 视频镜头强关联依赖关系提取过程

多标记视频镜头数据如表 1 所示。 $L_i (i = 1, \dots,$

4) 表示对应 i 镜头的多标记集合, 现以表 1 中视频数据为例来描述镜头中强关联依赖关系的提取过程。

表 1 视频数据集

Tab. 1 Video data set

镜头	标记
S_1	$L_1 = \{海景, 海滩, 山\}$
S_2	$L_2 = \{船, 海滩, 山\}$
S_3	$L_3 = \{海景, 船, 海滩, 天空\}$
S_4	$L_4 = \{船, 天空\}$

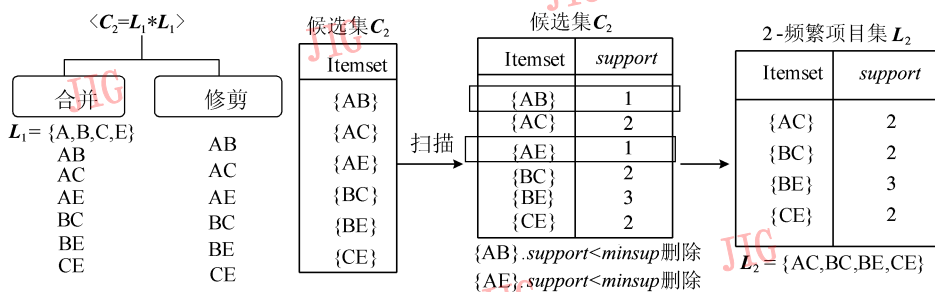


图 4 2-频繁项目集 L_2 的产生过程

Fig. 4 Generating process of 2-frequent itemsets L_2

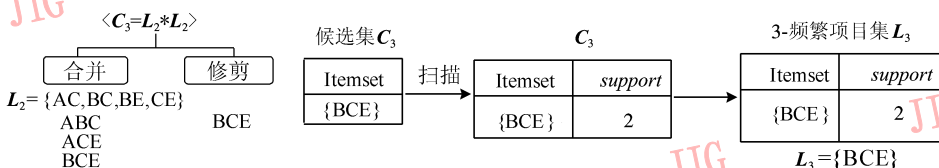


图 5 3-频繁项目集 L_3 的产生过程

Fig. 5 Generating process of 3-frequent itemsets L_3



图 6 复合标记集合 S_{assoc} 的导出过程

Fig. 6 Educing process of s compound label S_{assoc}

$$c(C \rightarrow BE) = P(BE/C) = \frac{s(BE/C)}{s(C)} = \frac{2}{3} = 67\% \quad (6)$$

$$c(CE \rightarrow B) = P(B/CE) = \frac{s(B/CE)}{s(CE)} = \frac{2}{2} = 100\% \quad (7)$$

式(6)依赖度表明了视频镜头中如果出现 3 者之一语义概念,则这个镜头中出现另外 2 个语义概念的可能性概率。式(7)则是视频镜头中如果出现任意 2 个语义概念,则这个镜头中出现另外 1 个语义概念的可能性概率。由以上 2 个式子还可以看出,式(7)的依赖度相对式(6)有提高的趋势。

3 实验与分析

视频/媒体数据包含视觉、音频、文字等模式的信息。实验采用 Scene 数据集^[1], MediaMill 数据集^[16]及 Emotions 数据集。以上 3 个数据集均来源于实际应用中的真实数据。其中, Scene 和 MediaMill 从支持向量分类数据库 LIVSVM 站点下载得到^[17], Emotions 数据集从多标记分类站点得到^[18]。3 个数据集特性如表 5 所示,数据多标记类别情况参见文献[13]。

Scene 数据集是 Boutell 等人用于语义场景多标记分类的实验数据。共有 2 407 个样本,包含了“海滩”,“日落”,“落叶”,“田野”,“城市”,“山”6 个语义类别。实验选用数据与文献[1]中实验完全相同,用 1 211 个样本数据进行训练,剩余 1 196 个样本用于测试。每个图片抽取 294 个特征。Emotions 数

示,最终得到的复合标记为 $S_{assoc} = \{海景 \wedge 海滩, 船 \wedge 海滩 \wedge 天空\}$ 。为说明强关联依赖关系集合各元素的依赖关系,按式(2)推导其相互的依赖度。AC = 海景 \wedge 海滩,其相互间的依赖度计算如下:

$$c(A \rightarrow C) = P(C/A) = \frac{s(A/C)}{s(A)} = \frac{2}{2} = 100\% \quad (3)$$

$$c(C \rightarrow A) = P(A/C) = \frac{s(C/A)}{s(C)} = \frac{2}{3} = 67\% \quad (4)$$

以上两式说明在视频镜头中如果出现包含“海景”语义概念,则这个镜头中 100% 将同时也包含“海滩”语义概念。如果出现包含“海滩”语义概念,则这个镜头中出现“海景”语义概念的可能性为 67%。由此可见,“海景”和“海滩”具有很强的相互依赖关系,该方法将“海景 \wedge 海滩”一起作为一个新的单标记来处理,目的是提高多标记分类器的性能。此外,从以上推导还可以看出:

$$c(C \rightarrow A) \neq c(A \rightarrow C) \quad (5)$$

BCE = 船 \wedge 海滩 \wedge 天空,其相互间的依赖度计算示例如下:

表 5 实验数据集特性

Tab. 5 Feature of experiment data

数据集名称	数据集所属领域	样本数目		属性数目		标记数目	标记基	标记密度
		训练样本	测试样本	数字	离散			
Emotions	多媒体	450	143	0	72	6	1.869	0.311
Scene	多媒体	1 211	1 196	0	294	6	1.074	0.179
Mediamill	多媒体	30 993	12 914	0	120	101	4.376	0.043

数据集是音频多标记分类数据,实验用 450 个训练样本,测试样本为 143 个。采样频率为 22 050 Hz,音频特征采用 MFCCs,一共有 6 个语义标记。Mediamill 数据是阿姆斯特丹大学视频语义检索原型系统所用的数据。系统用 101 个语义概念表示视频语义内容,类型包含事件(如“飞机起飞”)、对象(例如“马”)等,特征总数为 120。实验样本共有 43 907 个,实验用 30 993 个作为训练样本,12 914 作为测试样本。

多标记数据集特性各不相同。有的数据集中的每个样本语义标记的数都远小于 $|L|$,有的却与 $|L|$ 接近,这种特性可用“标记基”和“标记密度”来描述,标记基和标记密度详细描述参看文献[5]。

实验先采用 ML- k NN^[9] 和 SVM 按 TM4 方法进行多标记分类学习,然后再用本文方法对比实验。通过主观衡量不同支持度对应所得到的强关联依赖标记,实验得到强依赖关联语义关系支持度的较优值为 50%,实验中支持度设为 50%。先找出强关联依赖复合标记集合 S_{assoc} ,然后将集合 S_{assoc} 中的复合标记作为单标记,加上原有不属于集合 S_{assoc} 的标记后一起采用 TM4 方法转换为多标记分类学习。在 Emotions 数据集、Scene 数据集和 Mediamill 数据集上的实验结果分别如表 6、表 7 和表 8 所示。其中, Average precision 值越大表示学习性能越好。而其余 4 个指标实验值越小,则学习性能越好。表中第 1 例的 5 个指标从不同角度对学习性能进行评价,具体含义和计算方法参见文献[19]。

表 6 Emotions 数据集实验结果

Tab. 6 Experiment results of Emotions data

	直接多标记分类		本文方法	
	ML- k NN	SVM	ML- k NN	SVM
Hamming loss ↓	0.142	0.144	0.133	0.139
One-error ↓	0.213	0.224	0.204	0.217
Ranking loss ↓	0.212	0.236	0.198	0.228
Average precision ↑	0.721	0.739	0.742	0.746
Coverage ↓	6.312	6.412	6.212	6.329

表 7 Scene 数据集实验结果

Tab. 7 Experiment results of Scene data

	直接多标记分类		本文方法	
	ML- k NN	SVM	ML- k NN	SVM
Hamming loss ↓	0.124	0.139	0.124	0.135
One-error ↓	0.288	0.214	0.285	0.215
Ranking loss ↓	0.276	0.277	0.275	0.274
Average precision ↑	0.712	0.741	0.718	0.747
Coverage ↓	6.912	7.010	6.910	7.008

表 8 Mediamill 数据集实验结果

Tab. 8 Experiment results of Mediamill data

	直接多标记分类		本文方法	
	ML- k NN	SVM	ML- k NN	SVM
Hamming loss ↓	0.142	0.141	0.140	0.136
One-error ↓	0.216	0.207	0.212	0.200
Ranking loss ↓	0.191	0.189	0.187	0.182
Average precision ↑	0.780	0.788	0.791	0.795
Coverage ↓	7.681	7.679	7.678	7.675

从实验结果可以看出,本文方法的各评价标准实验结果比直接用 TM4 方法的实验结果要好。分析其中原因,主要是这两个媒体数据集中存在有较大比例相互关联依赖的语义概念标记。而本文方法将这种强多标记依赖语义在分类过程中作为一个“整体的”符合单标记处理,其实验结果也证明是合理的。综合以上实验,从 5 个指标实验数据来看,本文方法对视频多标记概念分类性能较好。

4 结 论

本文提出一种视频语义概念关联依赖关系的多标记分类学习方法。本分类方法不仅适用于视频中多标记语义概念的分类,也可用于其他具有较强关联依赖关系数据的多标记分类。实验表明本文方法对视频内容分析中涉及的多标记语义概念分类较有效。视频长时间段场景中关联关系的抽取有待今后进一步深入研究。

参考文献 (References)

- [1] Boutella M R, Luo J, Shena X, et al. Learning multi-label scene classification [J]. Pattern Recognition, 2004, 37(9): 1757-1771.
- [2] Tang J, Hua X S, Qi G J, et al. Structure-sensitive manifold ranking for video concept detection [C]//Proceedings of International Multimedia Conference Proceedings of the 15th International Conference on Multimedia. New York: ACM, 2007: 852-861.
- [3] Mccallum A K. Multi-label text classification with a mixture model trained by EM [C]//Proceedings of AAAI'99 Workshop on Text Menlo Park. California, USA: AAAI Press, 1999: 45-52.
- [4] Schapire R E, Singer Y. Boostexter: a boosting-based system for text categorization [J]. Machine Learning, 2000, 39(2/3): 135-168.
- [5] Tsoumakas G, Katakis I. Multi-label classification: an overview [J]. International Journal of Data Warehousing and Mining, 2007, 3(3): 1-13.
- [6] Chen Xiaofeng, Wang Shitong, Cao Suqun. Gene function analysis of semi-supervised multi-label learning [J]. CAAI Transactions on Intelligent Systems, 2008, 3(1): 83-90. [陈晓峰, 王士同, 曹苏群. 半监督多标记学习的基因功能分析 [J]. 智能系统学报, 2008, 3(1): 83-90.]
- [7] Liu Y, Jin R, Yang L. Semi-supervised multi-label learning by constrained non-negative matrix factorization [C]//Proceeding of the Twenty-First National Conference on Artificial Intelligence. Menlo Park, California; AAAI Press, 2006: 421-426.
- [8] Fan R E, Lin C J. A study on threshold selection for multi-label classification [EB/OL]. [2009-12-28]. <http://www.csie.ntu.edu.tw/~cjlin/papers/threshold.pdf>.
- [9] Zhang M L, Zhou Z H. MI-knn: a lazy learning approach to multi-label learning [J]. Pattern Recognition, 2007, 40(7): 2038-2048.
- [10] Liu Y, Jin R, Yang L. Semi-supervised multi-label learning by constrained non-negative matrix factorization [C]//Proceedings of Twenty-First Conference on Artificial Intelligence (AAAI'06). Menlo Park, California, USA: AAAI Press, 2006: 421-426.
- [11] Jiang Yuan, She Qiaoqiao, Li Ming, et al. A transductive multi-label text categorization approach [J]. Journal of Computer Research and Development, 2008, 45(11): 1817-1822. [姜远, 余俏俏, 黎铭, 等. 一种直推式多标记文档分类方法 [J]. 计算机研究与发展, 2008, 45(11): 1817-1822.]
- [12] Qi G J, Hua X S, Rui Y, et al. Correlative multi-label video annotation [C]//Proceedings of ACM International Conference on Multimedia (ACM MM). New York, USA: ACM, 2007: 17-26.
- [13] Wei W, Liu W Q, Li L L. Multi-label annotation study in video semantic content analysis [C]//Proceedings of 2009 International Conference on Machine Learning and Cybernetics. Baoding, China: Hebei University, 2009: 2453-2457.
- [14] Liu K H, Weng M F, Tseng C Y, et al. Association and temporal rule mining for post-filtering of semantic concept detection in video [J]. IEEE Transactions on Multimedia, 2008, 10(2): 240-251.
- [15] Wei W, Bin Y, Qing L. Study of association relation mining in multimedia data [C]//Proceedings of the 4th International Symposium on Intelligence Computation and Applications. Berlin, Germany: Springer, 2009: 180-187.
- [16] Snoek C G M, Worring M, Gemert J C v, et al. The challenge problem for automated detection of 101 semantic concepts in multimedia [C]//Proceedings of ACM Multimedia. New York: ACM, 2006: 421-430.
- [17] Chang C C, Lin C J. LIBSVM Data: Multi-label Classification [EB/OL]. [2009-12-28]. <http://www.csie.ntu.edu.tw/~7Ejlin/libsvmtools/datasets/multilabel.html>.
- [18] Trohidis K, Tsoumakas G, Kalliris G, et al. Learning from Multi-Label Data [EB/OL]. [2009-12-28]. <http://mlkd.csd.auth.gr/multilabel.html>.
- [19] Godbole S, Sarawagi S. Discriminative methods for multi-labeled classification [C]//Proceedings of Proceedings of the 8th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2004). Berlin, Germany: Springer, 2004: 22-30.