

具有多峰正态分布属性的视频语义分类研究^{*}

魏 维¹ 徐 建¹ 刘凤玉¹ 许满武²

(南京理工大学计算机科学与技术系 南京 210094)¹ (南京大学计算机科学与技术系 南京 210008)²

摘 要 视频语义分类中常遇到多峰正态分布属性,如采用单峰值正态分布设计的贝叶斯分类模型会造成较大分类误差。本文采用定步长组合划分算法(FLCPA)对多峰分布属性值域按类进行划分,以留一校验法(LOOCV)估算分类错误,找出给定步长下属性的多峰分布边界点,并用监督参数估计推断出每个分段区间上的概率分布函数,从而得到整个值域上的总体分布。此外,文中给出了涉及多峰分布属性的视频语义分类器设计步骤。实验数据表明,该方法能明显降低分类错误,有效提高分类性能。

关键词 贝叶斯分类器,多峰正态分布属性,视频语义分类,留一校验,类条件概率密度函数

Study of Video Semantic-concept Classifier with Multi-normal Distribution Attribute

WEI Wei¹ XUE Jian¹ LIU Feng-Yu¹ XU Man-Wu²

(Department of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing 210094)¹

(Department of Computer Science and Technology, Nanjing University, Nanjing 210008)²

Abstract Attributes with multi-normal distribution are common in classifier design for video-semantic concept. In this case, a model assuming that the value of attributes for each class is normally distributed with some mean will lead to poor classification performance. In the paper, an approach based on fixed-length combination partition algorithm (FLCPA) is presented in the partition of attribute value-field. Leave-one-out cross-validation (LOOCV) is used to estimate classifier error. In addition, the detail of classifier design about multi-normality distribution attribute is given. The result of experiment indicate the method could reduce classifier error and improve classifier performance.

Keywords Bayesian classifier, Multi-normal distribution, Semantic-based video classifier, LOOCV, Class-conditional-density function

1 引言

用统计模式识别方法分类提取视频语义概念,能实现低层视频特征到高层语义概念的映射,在一定程度上可跨越语义鸿沟(Semantic Gap)的限制。如何用贝叶斯理论进行语义概念的分类,是目前基于语义的视频检索的重要研究内容之一[1~3]。

视频语义分类所采用的特征中,不仅大量存在单峰值的属性,而且存在一定数量多峰分布的属性。通常的贝叶斯分类方法是针对单峰的分布情况,而用于对多峰分布的情况时,效果往往欠佳。本文将多峰分布属性值域进行划分,在每个划分子区间上再按单峰正态分布情况处理,较好地解决了具有多峰正态分布属性的视频语义分类中存在的问题。

2 贝叶斯相关理论及本文解决思路

图1所示为海滨视频镜头中“边方向”特征值域上同一类语义概念出现频数分布图(总样本数为250个)[4]。纵坐标表示样本在对应值域区间出现的频数,横坐标表示特征的值域。图中统计数据表明,此数据集在特征的值域上是三峰的正态分布。如果采用常用的单峰正态分布假设来设计贝叶斯分类器,则其推断出的概率密度函数必然与图中的三峰分布存在

较大误差。

贝叶斯分类器设计过程中,通常采用监督参数估计方法由样本集推断总体概率分布。即样本所属的类别及类条件总体概率密度函数的形式为已知,而表征概率密度函数的某些参数是未知的[5]。

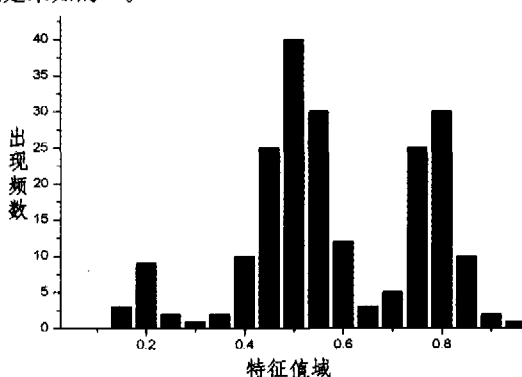


图1 典型的类属性值域多峰分布

考虑将待识样本 X 判为 c 个不同类 $\omega_1, \omega_2, \dots, \omega_c$ 中的某一类。由贝叶斯定理, X 应判为具最大后验概率的那一类。由于类条件概率分布未知,故通常假定分布为某一参数的模型,而表示分布的参数则由训练样本进行估计。采用

^{*} 本研究得到国家自然科学基金(60273035)、江苏省科技攻关项目(BE2003064)资助。魏 维 博士研究生,主要从事视频内容分析、语义视频检索研究。徐 建 博士研究生,主要研究方向为视频内容分析、语义视频检索;刘凤玉 教授,博士生导师,主要从事多媒体技术等研究;许满武 教授,博士,博士生导师,主要从事新型程序设计及多媒体方面的研究。

贝叶斯分类器时,物体是按最大后验概率进行分类的,这由一个判别函数来完成。多数情况下,该判别函数是线性的或二次的。当类服从正态分布时,要找到最优线性分类器总是不可能的。就目前所知,都是协方差矩阵相等的情况^[6]。

实际处理中,一般假定分布为(多维)正态单峰分布(当多维正态分布中均值向量和协方差矩阵已知时,由此分布得到的二次判别函数是最优的)。在此假设基础上用统计方法计算对应的方差和期望,得到条件概率密度。许多研究都集中于参数的估计方法(如文[7]中采用最大熵协方差选择方法估计贝叶斯分类器中的协方差矩阵,此方法可适用多属性,多类别的识别问题),而对不满足单峰分布假设前提条件的分类方法的探讨却不多。

对多峰分布的情况,一般采用非参数估计的方法(如Parzen窗口法, k_N 邻近法等)直接利用样本推断总体分布来设计分类器。而本文将多峰分布属性的值域按一定算法进行分划(类似思想最早出现在文[8]中,后来在文[9]用于描述子空间中子空间的划分)。在每个分划子区间上属性是单峰分布的,对应区间上可采用参数估计方法设计视频语义分类器。

3 定步长组合划分方法

对属性值域划分的本质就是选取划分的断点。如果每个子区间的划分太小,划分计算量会大大增加;如果划分的区间过大,其得到的分布必然导致语义分布的不准确。最简单的离散方法是等间隔划分法,但其效果较差,而最优断点集的求取是NP难问题,所以必须寻求适合视频语义概念分类应用的算法。本算法中将断点的选取作为定步长的组合搜索问题来处理。

3.1 留一校验

组合划分算法中用留一校验法(Leave-one-out cross-validation, LOOCV)进行分类错误估计。校验(Cross validation)是一种在概念学习算法/分类中用来估计假设错误的实验方法^[10]。将数据分为 k 个相同大小的子集,对模型一共进行 k

次训练。每次训练留出一个子集作为测试集,其余 $k-1$ 个子集作为训练集,这种估计错误的方法称作 k 重校验(K-fold cross-validation)。 k 重校验不仅适用连续错误函数值的估计,而且对非连续错误函数的估值也适合。若 k 大小不变,则为留一校验。相比较而言,留 $v(v>1)$ 校验比留一校验更精确。留一校验对贝叶斯分类的错误估计非常有效^[8]。由于留 v 校验要对测试集小于等于 v 的所有可能情况遍历,造成的计算量过大,因此本文采用留一校验估计算法中分类错误。

给定分类/学习系统A,则留一错误的定义式为^[11]:

$$R_{\text{loo}}(f_D) = \frac{1}{m} \sum_{i=1}^m l(f^i, z_i) \quad (1)$$

上式中, $l(f(x), z)$ 表示用 $f(x)$ 预测 z 的错误。 f_D 表示用训练集 $D=\{z_i, i=1, \dots, m\}$ 后,分类器进行分类的结果。留一法得到的估值是输出错误的“几乎”无偏估计。

2.2 划分算法

以 $l=(b-a)/n$ 为离散划分的基本步长,先高密度地将多峰分布的属性在值域上等间隔分为 n 段,各段的端点即为每类属性值域划分断点的候选点。选定候选断点后,划分子区间的个数可能为 $2, 3, \dots, n$ 个,总共 $n-1$ 种可能(子区间个数为1的情况即是通常的默认的单峰分布,此处不考虑)。

当划分子区间数为 $m(m=2, 3, \dots, n)$ 时,划分的断点选取方法有 $z=C_{n-1}^{m-1}$ 种($m-1$ 个断点,划分出 m 个区间)。对 z 种中的任意一断点选取方法,由留一校验法计算出贝叶斯视频语义的分类错误率 $P_{m_j}(e) (1 \leq j \leq z)$ 。取 $P_m(e) = \min_j P_{m_j}(e)$,则与 $P_m(e)$ 对应选取的断点即为子区间数为 m 时的划分断点。

按上述方法,确定属性 F_j 的 $n-1$ 种候选划分方法。在候选划分方法中对应的 $n-1$ 个错误率中选取最小值者,即 $P(e) = \min_m P_m(e)$,则 $P(e)$ 对应的划分就是处理多峰分布情况时属性 F_j 值域的最终划分。

表1 本文方法与单峰正态方法分类实验

数据集	数据特性				P(e)		
	训练样例数	测试数	类别数	n值	FLCPA 得到分段数目	单峰正态	本文方法
A	150	50	2	10	3	0.440	0.280
B	120	30	2	10	3	0.467	0.333
C	80	16	2	10	2	0.438	0.375

定步长组合划分算法(FLCPA):

1. 对属性 F_j 值域预先等间隔分为 n 段,确定步长
2. 确定划分子区间数为 m 时的划分断点:
 - A. 由留一校验法计算出 z 个贝叶斯错误率 $P_{m_j}(e)$
 - B. 取 $P(e) = \min_m P_m(e)$
 - C. $P_m(e)$ 对应的断点确定的分划则为区间数为 m 时的分划
3. 对所有划分子区间数为 $m(m=2, 3, \dots, n)$ 的情况,由2得到对应的划分
4. 由3得到的 $n-1$ 种划分方法中,最小对应的即属性 F_j 值域的最终划分

对每一类别,均用上述算法在相应属性值域上进行划分。由FLCPA算法看出,其计算量主要集中在留一校验估计上。算法中步骤1中的 n 值需预先给定,其大小直接影响到算法的计算复杂度。 n 值的选取必须适度,若其值过大,则以留一

法进行错误估算的FLCPA算法几乎不可能实现。划分算法中要采用留一校验法进行 $\sum_{i=1}^{n-1} C_{n-1}^{m-1}$ 次的贝叶斯错误率的计算。在实际的计算中还可以依据特定的视频语义概念进行适当简化,以达到降低计算复杂度的目的。比如在自然风光中,海滨视频语义分类器的EdgeHistogram特征属性峰值数一般均少于5个,因此用留一校验法进行 $\sum_{i=1}^5 C_{n-1}^{m-1}$ 次贝叶斯错误率的估算即可。相似的化简可大大减少实际应用中的计算量。

3 多峰分布的贝叶斯分类器设计

针对多峰分布属性,在第2节中的定步长组合算法得到第 $w_i (1 \leq i \leq k)$ 类相应的属性划分。 N_{ji} 表示属性 F_j 对于 w_i 类的划分子区间个数。在各个划分的分段区域中,采用监督参数估计方法推断其概率密度^[5]。具体而言,在每个子区域上假设样本服从正态分布,而正态分布的参数 μ, σ^2 是未知。

由已知类别的样本集对总体分布的 μ, σ^2 参数进行统计推断,得到本段区间上的概率密度函数。 N_μ 个分段区域上对应的概率分布用分段示性函数表示,则得到属性 F_j 上第 w_i 类的多峰概率密度函数。

分类器设计步骤:

1. 属性 F_j 第 w_i 类语义的概率密度推断:

A. 用定步长组合划分算法对属性 F_j 值域进行针对第 w_i 类视频语义概念的划分

B. 采用监督参数估计推断第 $k(1 \leq k \leq N_\mu)$ 段上的概率密度

C. 所有的 N_μ 段重复B,得到由 N_μ 个分段概率密度函数表示的多峰概率密度

2. 对所有 w_i 类($1 \leq i \leq c$),按1分别求得多峰概率密度函数

3. 对测试样例进行分类测试

4 实验及计算量分析

k 重校验不仅适用连续错误函数值的估计,而且对非连续错误函数的估值也适合。由于 $k=10$ 时,留一法能兼顾估计结果的质量和计算量^[11],因此本文实验采用 $k=10$ 的参数。将数据分为10个不相交子集,每次用9个子集进行训练,剩下的1个子集用作测试子集。最后将10次结果的均值作为最终实验结果。

表1是用本文方法与用单峰正态假设分类方法对三种典型的多峰分布视频语义概念分类的对比实验。A数据集是海滨景色的镜头,B是山峰的镜头集,C则是河流视频镜头集。此三类视频语义概念分类特征EdgeHistogram中的边方向特征均是多峰分布。分析实验结果可以看出,以存在多峰分布的属性作为分类特征时,若按通常的单峰正态分布建立贝叶斯分类模型,分类错误率 $P(e)$ 偏大(达到0.4以上)。而采用本文中的方法处理,分类错误率大大降低。

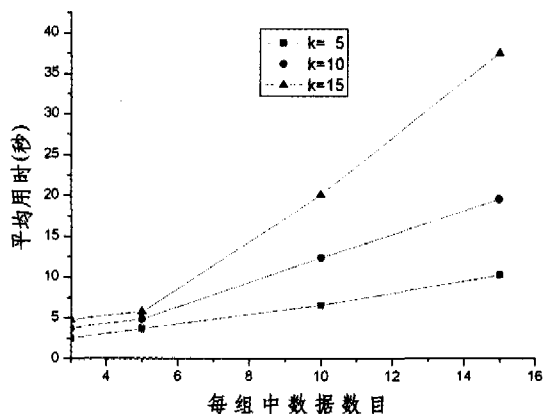


图2 FLCPA算法计算时间(效率)分析

当 n 为定值前提下,FLCPA算法的计算时间长短与数据集规模大小成正比。分组数 k 固定(即用留一校验法进行贝叶斯错误率计算的次数不变)时,每组中数据个数增加,导致每次训练中训练样本数加大,此时计算量的增加主要是训练时间变长所致;而每组中数据个数固定时,随着分组数目增加,用留一校验法进行贝叶斯错误率计算次数的增加则是计算量增大的主要原因。以上两个变量的乘积正是数据集中数据的总样本数目,这表明本算法计算效率与数据集的规模大小成反比。当数据集超过一定规模,计算时间将加长到不能

容忍的程度。具体数据集规模最大上限值的定量选取由硬件和应用中最大允许的计算时间两因素确定。

上述定性分析的直观量化实验如图2所示。此实验针对表1中A类型语义进行,其中 n 取固定值10,进行分组个数分别为 $k=5, 10, 15$ 。横坐标表示每个分组中数据数目,纵坐标是CPU平均计算用时。由图中数据分析可以得出:随着 k 值以及每组中数据数目(数据集规模)的增加,计算量增大。

图3所示是定步长组合划分算法分类错误率 $P(e)$ 与 n 的分组实验情况。此实验采用表1中数据集,数据集特性与表1中描述的完全相同。图中横坐标表示算法中属性值域的预划分数目 n ,纵坐标表示分类器的分类错误率 $P(e)$ 。由图中数据分析可以看出:随着 n 值的增加, $P(e)$ 值降低。由于定步长组合划分算法的 n 取值也将影响进行留一校验的次数,因此为确保可容忍的计算规模, n 的取值不能太大(表1实验中 $n=10$)。由于 n 的取值不能过大,因此决定边界点的步长粒度较大,这使得多峰分布特征的值域划分边界点不能精确选定。由于上述边界点选取存在一定误差,因此虽然在涉及多峰分布的分类器中采用本文方法后分类错误大幅降低,但 $P(e)$ 往往都不低于0.2。

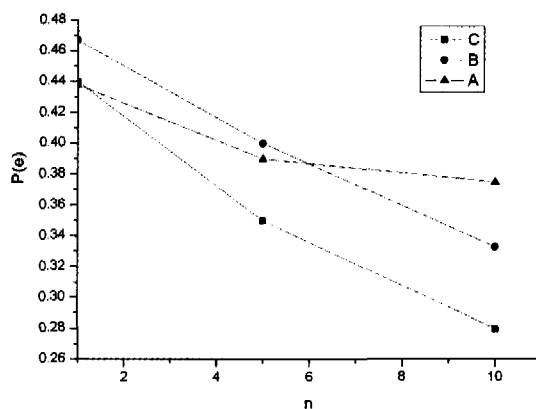


图3 FLCPA算法分类错误率 $P(e)$ 与 n 分组实验

结论与将来研究工作 视频语义分类中常遇到多峰分布属性作为分类特征的情况,本文采用定步长划分算法对多峰分布属性值域进行划分,用参数估计方法推断出与实际多峰分布吻合较好的分段概率密度函数。在涉及多峰分布属性的视频语义分类中,本方法能使分类错误率 $P(e)$ 得到降低,较好地解决了与多峰分布属性相关的语义概念分类。

由于本算法中计算量与数据集规模成正比,因此在数据集规模较大时会导致计算时间过长。如何结合本算法特点来降低LOOCR计算量,加快分类错误计算,使之适合大数据集,是将来仍需深入研究的内容。此外,在留一校验方法中错误接受率FAR(False Acceptance Rate)与错误拒绝率FRR(False Reject Rate)的分析与估计也是今后本算法进一步分析提高的基础。

参考文献

- 1 Zhou L, et al. Applying the Naive Bayes Classifier to Assist Users in Detecting Speech Recognition Errors [C]. In: Proceedings of the 38th Annual Hawaii International Conference on System Sciences, 2005. HICSS '05, Hawaii, USA, 2005. 183b~183b
- 2 Rui S, Jin W, Chua T-S. A Novel Approach to Auto Image Annotation Based on Pairwise Constrained Clustering and Semi-Native Bayesian Model [C]. In: Proceedings of the 11th International

AWGN 信道下 Turbo 码误比特率模型^{*}

冯 镔 刘文予 朱光喜 马 展

(华中科技大学 武汉光电国家实验室 武汉 430074)

摘 要 本文研究信道编解码及信道特征对误比特率性能的影响。在实验基础上,通过曲线拟合等数学工具,得到加性高斯白噪声信道条件下,信道编码采用 Turbo 码、通信系统端到端的误比特率经验模型。通过该模型,可以在发送端预测出在一定信道条件及系统参数设置下接收端的误比特率。实验表明,该模型具有极高的精度,平均误差在 3% 左右。该模型可以应用在信源信道联合编码及端到端的率失真优化中。

关键词 Turbo 码,经验模型,曲线拟合,误比特率

A BER Model for Turbo Code on AWGN Channel

FENG Bin LIU Wen-Yu ZHU Guang-Xi MA Zhan

(Wuhan Optical-Electronic National Laboratory, Huazhong Univ. of Sci. & Tech. Wuhan 430074)

Abstract Bit error rate(BER) characteristics is researched in this paper under the influence caused by the channel codes and channel statistic. Using mathematic tool such as curve fitting etc., an empirical model on BER is derived on the basis of extensive experiments. The channel model used is AWGN channel. Turbo codes are used as channel codes. Given certain channel statistic and parameter set, the BER in receiver end can be estimated in sender end using this model. Extensive experimental results demonstrate that this model bears high accuracy, whose average error is around 3%, thus can be used in joint source-channel coding and end-to-end rate-distortion optimization.

Keywords Turbo codes, Empirical model, Curve fitting, BER

1 引言

自从 1993 年 C. Berrou 等人提出 Turbo 码以来^[1], Turbo 码凭借其逼近 Shannon 限的优异性能得到了学术界的广泛重视,成为信道编码领域新的研究热点。Turbo 码满足 Shannon 定理中随机编译码的条件,并利用了迭代译码算法,这使得它不仅能够抵抗恶劣的噪声环境,而且具有很强的抗干扰和衰落的能力,因此在无线通信系统中有很大的应用潜力。从第三代移动通信系统候选方案来看,普遍要求提供中速或者高速的数据业务,比较第二代移动通信中的级联 RS 加卷积码, Turbo 码性能可提高 1dB 以上。在 IMT-2000 多种方

案中,已经将 Turbo 码作为传输高速数据的信道编码标准之一^[2,3]。因此,研究 Turbo 码在无线通信中的应用是有实用价值的^[4]。

但是,由于 Turbo 码的理论尚不完善,特别是由于编码器中交织器的存在,使得对 Turbo 码的性能进行数学分析显得非常困难,国内外在这方面的研究成果也少见报道。有很多学者致力于 Turbo 码误比特率上限的研究,但是这种上限仅在高信道信噪比条件下与实际性能相接近。例如,Barbulescu 等人的研究表明,对于码率为 1/3 的 Turbo 码,信噪比门限为 2.03dB;当信道信噪比低于这个门限时,性能上限与 Turbo 码实际性能之间存在相当大的差距^[5]。而且,上限

^{*}基金项目:国家自然科学基金资助项目——无线信道下的信道广义率失真模型研究(No. 60372067)、国家 863 资助项目——数字视音频编码、传输、测试于应用示范系统(No. 2002AA119010)。冯 镔 博士研究生,主要研究方向为多媒体通信;刘文予 教授,博士生导师,主要研究方向为多媒体通信、计算机视觉;朱光喜 教授,博士生导师,主要研究方向为多媒体通信与计算机图像图形处理;马 展 硕士研究生,主要研究方向为信源信道联合编码。

Multimedia Modelling Conference, MMM 2005, Melbourne, Australia, 2005. 322~327

- 3 Wang F, et al. A Generic Framework for Semantic Sports Video Analysis Using Dynamic Bayesian Networks [C]. In: Proceedings of the 11th International Multimedia Modelling Conference, Melbourne, Australia, 2005. 115~122
- 4 ISO/IEC JTC1/SC29/WG11. Int'l Organization for Standardisation, Overview of the MPEG-7 Standard [S]. 2000
- 5 边肇祺,张学工,等. 模式识别[M]. 北京:清华大学出版社, 2000
- 6 卢力,田金文,柳健. 统计模式识别研究进展. <http://www.ci-oworld.net/cgi-new/violethtml/2/2004-04-24/20040424165850.html>, 2005
- 7 Thomaz C, D G, Q F. A New Covariance Estimate for Bayesian Classifiers in Biometric Recognition [J]. IEEE Trans. on circuits

and Systems for video technology, 2004, 14(2): 214~223

- 8 Pazzani J M. An iterative improvement approach for the discretization of numeric attributes in Bayesian classifiers [C]. In: Int'l Knowledge discovery and mining, Montreal, Canada, 1995. 228~233
- 9 Mittal A L-F C. Addressing the problems of Bayesian network classification of video using high-dimensional features [J]. IEEE Trans. on Knowledge and Data Engineering, 2004, 16(2): 230~244
- 10 What are cross-validation and bootstrapping. <http://www.faqs.org/faqs/ai-faq/neural-nets/part3/section-12.html>
- 11 Elisseeff A. Leave-one-out error and stability of learning algorithms with applications. <http://www.cs.ucl.ac.uk/staff/M.Pontil/reading/natoasi.pdf>, 2005