



川端 勉
Tsutomu KAWABATA

研究课题

多维渐近式量子化、针对特殊概率测度的率失真理论、数据压缩算法的设计与解析

关键词

渐近式矢量量子化, 信息源编码, 格子, 微分几何学, 率失真函数, 特殊概率测度, 率失真维度, Lempel-Ziv 符号, 树源码, 上下文树加权法

所属专业	研究生院信息理工学研究科 信息·通信工学专业
研究成员	川端 勉 教授
所属学会	信息理论及其应用学会, 电子信息通信学会, IEEE, 日本应用数理学会

研究概要

信息源编码的研究

通信的目的是通过通信路径输出信息源, 并传递给收信者。作为信息源输出的案例, 有音频编码等连接时间的信号 $X(t)$ 和文本文件等离散时间的离散值数据 X_n 。

信息通信的第一阶段是采用信息源输出的二进制代码 (0 和 1) 进行编码, 其目的是压缩电子文书、图像、声音等各种数据外表的信息量。因此这个过程也被称为信息源编码或数据压缩。采用信息源编码, 通信就单纯地回归到经由通信路径的信息传输问题上。并且信息源编码的理念是包括信息传输问题的各种问题的基础。

无失真信息源编码的研究

信息源编码有两种情况, 一种是如同文本文件一样有可能进行无失真解密, 另外一种如同连续信号一样在本质上不可能进行无失真解密。因此首先要尝试考虑有可能的情况。如果 X_n 遵从概率法则, 那么就能思索作为编码输出的二进制长度的期待值。如果能将所有编码的这个期待值最小化, 可以说正是这个期待值才是信息源具有的本身的信息量吧。

那么能够从 X_n 本身估计出本质的信息量的值吗? 虽说是假设, 但如果可能的话, 采用对应算术编码法这种实数变换的技术, 就能得到估计值附带的编码方法。另一方面, X_n 为文本文件时, 可以认为本信息源这种分层式的概率法则的族是标准的。

树信息源、上下文逻辑性木加权法 (CWT)、Lempel-Ziv 编码法

该研究室针对本信息源, 就有效推定法之上下文逻辑性木加权法 (CWT) 进行相关的研究。

那么树信息源的假设中出现不合理的计算量时, 需要别的途径来达到本质信息。

他们就针对上述情况有效的 Lempel-Ziv 编码法 (辞书法) 从设计及压缩性能分析这两方面来进行研究。

以下简单介绍一下被称为增量分析法的算法设计相关的研究。这种方法是从空白辞典及用来编码的字符串出发的。方法极其简单, 从输入的字符串的开头找出辞典中没有的最短的单词, 将其剪切下来, 再将剪切的单词收入到辞典中, 这样反复操作。由于新剪切的单词是在已经收录到辞典中的单词上增加了一个字符, 所以能够进行较短的显示 (也就是数据压缩)。

他们针对这种基本方法开发出了新的技法, 能够在辞典具有的树结构中有效列举、表现新的分解单词。这种方法不仅能够以简便的方式来进行记述, 还具有良好的性质来消减原来的方法中的缺点。并且他们开发的方法超越了增量分析法, 具有广泛的应用, 目前还正在不断完善。从这点意义来看也可以说该技术是基本的技术。

该方法的提案是从他们原来进行的分析研究中产生的。

有失真信息源编码的研究

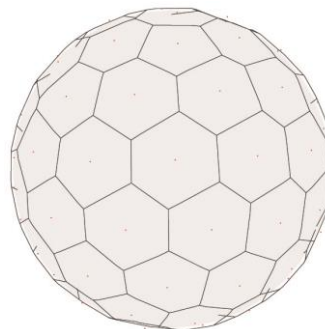
接着他们还尝试考虑允许有信息失真的情况。在信息源概率法则的假设下, 研究这种情况下编码的效果, 也就是说权衡信息压缩率与容许失真的关系是非常重要的课题。

但是这种关系很大程度上取决于信号的种类。

比如现实信号中会出现连续信号、离散信号均没有的特异信号。这种情况在理论上是空白的。

但是该研究室发现了针对包括分形信号在内的特异信号的压缩失真率的关系。

在信息理论中, 这种理论是针对连续信号、离散信号均没有的特异信号的压缩理论, 是开拓信息理论新世界的先锋。



作为二次量子化研究的一环而设计的球面量子化设备

优势

基于数理工学的独特研究

信息理论是构成信息通信理论基础的电气电子信息系统的科学, 最重视工学实现。因此致力于先进技术的研究, 基础非常重要。

该研究室研究的不仅是已经将基础体系化的信息理论, 还研究数理统计学、几何学方法等数理工学。能够将独创的算法设计及系统分析法的研究付诸实践。今后他们会不断努力, 使这些研究成果在信息理论的体系中占有一席之地。

进而他们不仅对研究生, 还对高校学生及本科一年级学生积极介绍他们的研究室。要取得很好的研究成果本身就很不容易, 在促进学生关心信息理论这种科学, 并使其产生兴趣的同时, 也提倡基础研究的重要性。

信息理论的据点: 电气通信大学

从事数字技术开发的技术人员要长时间、有意义地活跃在第一线需要有远见, 能够给予技术人员远见的基本原理就是信息理论。信息理论直接或间接地与各种科学、工学领域的基础紧密相连, 通过各种研究经验, 好不容易掌握信息理论的研究人员也很少。电气通信大学是该信息理论在日本的中心地, 在该研究室能够以本大学教授的信息理论为核心, 学到先进的信息通信技术。



在讨论会上, 不受已有概念的束缚, 自由交换各种意见

未来展望

挑战改善二次量子化方式

以下介绍再介绍一例容许失真的情况下 (矢量量子化法) 的研究。连续信号高清晰度压缩的问题是作为多次元渐近式量子化问题来进行研究的。这等同于空间的多点配置问题。

该研究室再次发掘迄今被放弃的二次量子化方式, 尝试去改善理论与方式。如果这个得以实现, 就会完全改变迄今为止使用的标准方式 (采用均匀量化法进行可变长度编码的方式) 吧。

最后还在讨论信息源编码是否有效提高通信的隐密性。可以说隐密性的提高或者信息隐藏理论是未来有希望的应用领域。