

中华博士 园地

这是本刊特为海内外正在就读和学成立业的博士、博士后青年学者们开辟的一片科普园地。深学浅著,是一门德识、慧学、素质修养的学问。你们的新知识、新调研、新观察、新目光、新展望,能够用尽可能深入浅出、通俗流畅的语言,汇报给祖国、人民、家乡父老子弟乡亲们吗?中华博士园地,乃耕耘忠孝之地、科教兴国、民族昌盛之地。要用慈母听得懂的语言,写出你们的心声!

极低码率的图象压缩编码方法

刘富强

1 引言

多媒体系统需解决的关键问题就是庞大数据量的采集、存储和传输。一般来说,文本的数据量不大,所以在传输时速率较低,存储时所占用的空间也不多;图形和声音的数据量要大一些,不过现有一般计算机系统也能处理,但是如果是图象或者是视频图象,那就不同了。如一幅中等分辨率的彩色图片,每像素 24 位,数据量为 $640 \times 480 \times 24$,能达到

7.37Mbit;若是运动的视频图象,以每秒 30 帧的速度播放,则视频信号传输速度将达到 221Mbit/s,650 兆字节的硬盘或光盘只能存储 24 秒,几乎做不了什么,更不用说观看五彩缤纷的 VCD 了。因此如果不进行数据压缩,则现有计算机系统难以处理多媒体数据特别是视频图象数据。

2 数据压缩的国际标准

图象压缩方法一般可以分成有损压缩和无损压缩两种类型。无损压缩利用数据的统计特性来进行数据压缩,典型的有霍夫曼(Huffman)编码、行程编码、算术编码等,其压缩率一般为 $2:1 \sim 5:1$;有损压缩是利用人的视觉特性使解压缩后的图象看起来与原始图象一样,不能完全恢复原始数据,主要方法有预测编码、变换编码、模型编码、分形编码、模型基编码等,压缩比随着编码方法的不同差别较大。

1969 年在美国举行了首届图象编码会议(PCS),1988 年是图象压缩编码的发展历史中极为重要的一年,集中表现在确定了 JPEG 和 H.261 两个建议的原理框架。JPEG 是由国际图象编码联合专家组于 1990 年最后确定为静止图象压缩编码标

刘富强 男,1965 年 3 月生,1987 年毕业于天津大学电子工程系,1996 年在中国矿业大学获博士学位,现在中国矿业大学从事博士后研究工作,中科院模式识别国家重点实验室客座研究人员,IEEE 会员。研究方向为多媒体技术、图象处理与模式识别、数据压缩及其应用。



准, JPEG 算法的平均压缩比为 15 : 1, 当压缩比大于 50 倍时可能出现方块效应。H. 261 覆盖了 64kbit/s ~ 1920kbit/s 的信道, 它奠定了以后 MPEG 标准的基础。

MPEG-1 是国际运动图象专家组 (MPEG) 制定的视频压缩编码标准。MPEG-1 压缩后可达到的码率约为 1.5Mbps (其中包括双通道的立体声信号), 用 MPEG-1 标准的平均压缩比为 50 : 1, 其处理能力可达 360 × 240 个像素。基于 MPEG-1 标准的 VCD 影碟机首先是在中国的万燕公司设计生产的, 而今我国成了世界上最大的 VCD 生产国。

MPEG-2 (也称 H. 262) 是 MPEG 组织制定的又一个通用压缩编码标准。MPEG-2 标准的处理能力可达广播级水平, 即 720 × 480 个像素。MPEG-2 标准兼容 MPEG-1 标准, 适应于 1.5Mbit/s ~ 80Mbit/s 编码范围。MPEG-2 标准也是高清晰度电视 (HDTV) 全数字方案所采用的数据压缩标准。目前蓄势待发的 DVD 即采用了 MPEG-2 标准。

3 极低码率的图象压缩编码方法

极低码率的压缩编码研究始于 80 年代, 当时研究的主要目的是如何利用模拟电话线路传输可视电话信号。1981 年在加拿大蒙特利尔城市召开的 PCS 会议上发表了 3 篇有关的论文。R. Wallis 等人发表了 9 600bps 的会议电视论文。1983 年 PCS 会议上,

Forchheimer 和 Fahlander 提出了一种新的压缩方案, 即模型基图象编码。在后来的 PCS 会议上 (1988 ~ 1996) 以及其它一些国际会议对极低码率的视频编码进行了深入而继续的研究工作。H. 263 是为低码率应用 (例如在公共电话网上实现可视电话或视频会议等) 而制定的视频压缩标准。为了降低码率, H. 263 在 H. 261 的压缩算法上作了一些改进。

MPEG-4 (H. 264) 是正在讨论的极低码率应用的视频压缩编码标准。与 JPEG、MPEG-1、MPEG-2 等其他标准所采用的基本压缩算法不同, 该标准拟采用基于模型的编码、分形编码等方法, 以获得极低码率的压缩效果。MPEG-4 比 MPEG-2 具有更广泛的应用, 它不同于过去的 H. 26X 系列标准, 它具有 3 个最新的技术特征是: (1) 基于内容的数据压缩 (2) 更高的压缩比 (3) 时空可伸缩性。

MPEG-4 可望在 1998 年形成国际标准。

现有的极低码率下图象压缩编码方法较多地考虑了模型基编码。模型基编码与以前的编码原理有所不同, 它重点分析景物中的物体结构模型, 从视觉心理上分析, 这是符合人眼特性的。以可视电话为例, 其原理如图 1 所示, 因为电视电话的景物结构性很强, 内容较为固定, 背景基本不变, 因此我们可以将运动部分即人物部分提取出来, 并建立人物眼、嘴、鼻等运动部分模型, 从编码的一端往接收端发送运动模型参数, 在接收端根据其参数重新合成图象, 压缩效率非常高。



图 1 模型基编码在可视电话中的应用

4 特定环境下极低码率的图象编码方法

图象数据压缩编码的国际标准, 虽已基本满足了低码率的图象编码, 但在一定的场合下不一定就是效率最高的。如在工业等环境电视中仍有许多问题, 主要原因是没有考虑到工业电视图象特有的性质和特殊的传输、应用环境, 故在应用于工业电视图象压缩取得的性能指标受到了很大的限制。

长期以来, 我们一直在对工业电视图象的压缩编码进行研究和分析。我们认为在工业环境中, 模型基编码可以得到更好地应用。这是基于以下原因:

(1) 工业环境中的背景一般不变, 这些数据在传输过一次并在接收端存储之后便不在传输, 从而大大减少数据传输量。

(2) 在所监视设备停止运行时, 可等同背景对待, 此时要传输的信息量很小。

(3) 工业环境中的运动部分通常有很好的规律性, 如曲线运动、直线运动、圆周运动或椭圆运动等。

这种规律性可帮助我们建立相应的模型,利用其规律性来预测下一运动点,从而减少要传输的数据。

(4)即使同样是运动部分,既有我们关心的运动部分,又有我们不关心的部分,而那些我们不关心而且又不重要的运动部分就可以不予以考虑,在传输时不参加传输。例如在煤矿一些地方的运输皮带上不允许载人,从图象处理的角度来看,如果遇到皮带载人,则尽快予以传输,并进行报警,自动或提醒值班人员采取措施;如果遇到皮带以外的物体或人员通过,不影响值班人员判断,则不予传输。

(5)从信息获取的角度来说,就工业设备的正常运转而言信息量并不是很大,而我们最关心的信息是工业设备的运转出现异常的情况,如火灾、冒烟等。因此在正常情况下,需要传输的图象数据量不是很大,只要够构成解码端模型参数的数据量就可以了。举例而言,“狗咬人”算不得什么新闻,但“人咬狗”则可能算是一条新闻了。这说明同样的数据量所携带的信息是不一样的,那么无用的数据量当然可

以被压缩掉了。

我们在接收端和发送端同时提取图象动态部分的模型,存储监视点工业设备正常运转一周期提供的所有图象,包括设备正常运转摄像头摄取的图象序列,每幅图象的顺序编号等知识的数据库。图 2 是龙东煤矿一号皮带机头一段时间内的图象序列,显示出运煤皮带的煤块被运输到储煤仓的过程。白方框的移动显示的是运动物体从图象上方移动到下面的一个循环过程,椭圆框内显示的是机械的圆周或椭圆运动。

对于正常画面,进入该系统后,可在知识库模型中搜索与该帧图象均方差最小的图象,将其编号和正常标记一同传输至接收端,接收端调用知识库中相同编号的图象显示。

当出现异常情况时,采用成熟的技术,如基于 MPEG-1、MPEG-4 或其它标准适合软硬件处理的方法进行处理。

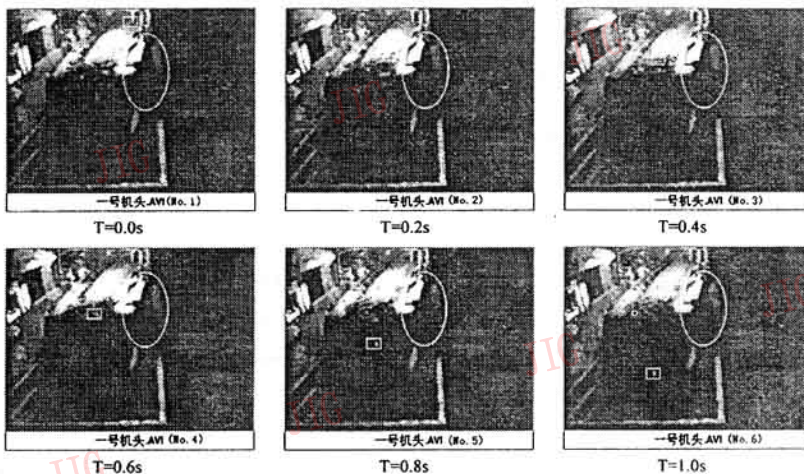


图 2 某矿一号机头 1 秒内的视频图象

以上的编码方法可以大大降低对视频服务器的要求,从而使得低成本多媒体系统成为可能。这种编码方法的数据传输率小于 1kb/s,其数据压缩比可达 10^4 至 10^5 数量级。

5 应用前景

极低码率的模型基图象编码是一种很有研究价值的实用图象压缩方法,尤其适用于工业生产中的

电视监视系统或是银行、监狱等地方的电视安全监控系统。1998 年 MPEG-4 将形成标准,符合其标准的硬件和软件将大量面市,将来我们在现有的电话线上即可以开电视会议,打可视电话,点播电视收看,进行远程图象控制等,谁能肯定说那时的极低码率编解码设备赶不上如今的 VCD 呢? 在工业领域内,图象传输和控制再也不是影响调度和安全生产的瓶颈了。

参考文献

- 1 Fabio Lavagetto, Sergio Curinga. Object-oriented scene modeling for interpersonal video communication at very low bit-rate. Signal Processing: Image Communication, 1994, 6(3).
- 2 Li H, Lundmark A, Forchheimer R. Image sequence coding at very low bitrates: a review. IEEE Trans on Image Processing, 1994, 3(5).
- 3 徐孟侠. 图象编码的进展. 通信学报, 1993. 3.
- 4 刘富强, 于洪珍. 基于模型的数据压缩编码方法在工业电视系统中的应用研究. 东南大学学报, 1997. 5B.
- 5 Liu Fuqiang, Qian Jiasheng, Yu Hongzhen et al. Research On Multimedia Optic-fiber Industrial TV Monitoring System. Hong Kong-Beijing International Computer Conference'97. 北京: 清华大学出版社, 1997, 33.
- 6 刘富强, 赵鸿萍, 于洪珍. 工业电视图象数据压缩的混合编码方法研究. 南京大学学报, 1997. 10.
- 7 胡栋. 适应于极低传输码率的物体基图象编码. 电信科学, 1995. 10.
- 8 刘富强. 煤矿多媒体综合调度指挥系统研究[博士论文]. 徐州: 中国矿业大学, 1996.

联想补天工作站图形能力令人刮目相看

联想集团推出的联想补天 2100 功能强大, 不仅是因为它采用了 Pentium II 处理器 (主频 223~333MHz), 以及 Intel 82440LX 芯片组, 能够支持最大 384MB SRAM 或 768MB EDO DIMM, 更值得一提的是, 它采用了国际知名的专业图形加速卡制造厂商艾尔莎的高档产品——具备超群的图形处理能力的 ELSA GLoria Synergy。这是联想补天 2100 的图形性能得以完美体现的关键所在。

ELSA GLoria Synergy 是一种采用 AGP 总线、3Dlabs 的 Permedia 2 处理器和 8MB SGRAM 的高性能图形加速卡。在高达 230MHz 象素时钟下, 它能使色彩分辨率高达 1920×1200; 它也可在多屏方式下运行, 并支持 16:10 宽屏显示器。ELSA GLoria Synergy 能在硬件中提供双线性纹理映射, 充分支持 OpenGL、HEIDI 和 Direct3D 等流行 3D 标准, 同时提供实用程序, 针对不同应用软件自动进行系统优化设置, 方便用户使用。

对于联想与艾尔莎的此次合作, 艾尔莎科技股

份有限公司总经理张国龙先生表示: “联想电脑是国内业界的第一大品牌。1997 年, 联想电脑在国内的市场占有率位居第一, 超过了 IBM、康柏等世界知名品牌。联想与全球著名图形加速卡领导厂商艾尔莎在产品 & 经销渠道方面的结盟, 对双方的市场营销具有实质意义。通过合作, 联想进一步坚定了拓展工作站新产品, 进军国际舞台的信心; 对艾尔莎来说, 其品质与服务又一次得到了肯定。”

联想电脑公司产品部副总经理刘旦先生表示: “德国艾尔莎公司在专业图形处理领域具有悠久的历史, 其技术力量雄厚, 产品性能优异。联想电脑公司与艾尔莎公司在电脑图形领域将进行长期的合作, 不断提高国内 NT 图形工作站的档次, 为中国用户带来超卓的性能和完善的服务。”

有理由相信, 两强的携手合作将在国内及国际市场上产生良好的反响, 并将为图形工作站市场带来一番新气象。

(文淼)