

压缩域多媒体数据处理技术研究

向辉 石教英

(浙江大学 CAD & CG 国家重点实验室, 杭州 310027)

摘要 多媒体数据压缩技术是多媒体、网络通讯以及计算机等应用研究领域的一项关键技术。多媒体数据经过压缩后再进行存储或通过网络进行传输,这已经逐渐成为多媒体应用中的标准模式;同时,多媒体应用中所涉及到的各种媒体数据,如图象、视频和音频数据,往往需要进行各种灵活的操作与处理,因此,压缩域多媒体数据直接处理技术成为一项很有意义的研究领域。文中首先就压缩数据处理的基本概念进行讨论,然后集中分析了这一研究领域的重要研究成果,并指出了一些可能的研究方向。

关键词 多媒体数据压缩 压缩域数据处理 压缩比 变换域

0 引言

随着多媒体技术的飞速发展,其应用领域越来越广泛。由于多媒体数据的数据量十分巨大,对于存储器的容量以及网络通讯线路的传输带宽都提出了苛刻的要求。因此,多媒体数据压缩技术成为多媒体、网络通讯以及计算机等应用研究领域的一项关键技术。多媒体数据通过一定的压缩算法进行压缩后,以压缩格式进行存储或通过网络进行传输,这已经逐渐成为多媒体应用中的标准模式。同时,多媒体应用中所涉及到的各种媒体数据,如图象、视频和音频数据,往往需要进行各种灵活的操作与处理。例如,对于图象,我们可能希望进行诸如自由缩放、平移、旋转以及变形等几何变换或平滑滤波、边缘提取等处理;对于视频数据,可以进行视频叠加或蓝屏编辑等等操作;就音频数据而言,同样需要进行滤波、混响、音量调节等处理。另外,对于多媒体数据库中的各种数据,我们还希望进行适当的多媒体信息查询。多媒体数据的压缩格式进一步增加了多媒体数据处理过程的难度。压缩域的多媒体数据的直接处理技术正是在这种背景下出现的,其重要意义是不言而喻的。因为压缩域数据处理技术所处理的对象将是容量大大缩减过的多媒体数据,同时,也省

略了解压缩/再压缩的附加环节,可以预期,处理效率会有较大的提高。

压缩域数据处理技术近年来逐渐成为多媒体研究领域的一个热点。在1996年的图象压缩国际会议上,专门设立了一个压缩域图象视频数据处理(Compressed Domain Processing)的论文报告会^[1-3]。压缩域处理技术作为图象与视频数据处理的一个重要研究分支,正式得到了一致的认可。

文中对压缩域数据处理技术进行了综合的讨论与分析。首先,本文就压缩域数据的处理的基本概念进行讨论,然后,着重分析了这一研究领域的重要研究成果,并提出了一些很有潜力的研究方向。

1 基本概念

传统的多媒体数据压缩及处理模式如图1所示。

而在引入压缩数据操作处理之后,我们得到如图2所示模式。

由此可见,压缩域数据处理技术是直接对压缩数据之上的操作,而无须经过解压缩以及再压缩的附加环节;另外,压缩域的数据量一般远远少于原始数据的数据量,如何充分利用这些优势,实现高效的压缩域数据处理算法,正是我们研究的目标。

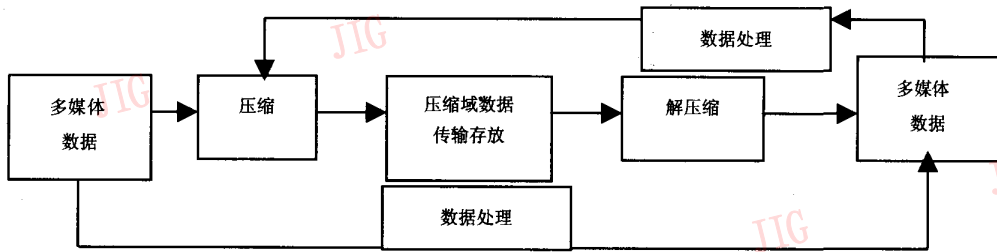


图1 传统多媒体数据处理模式

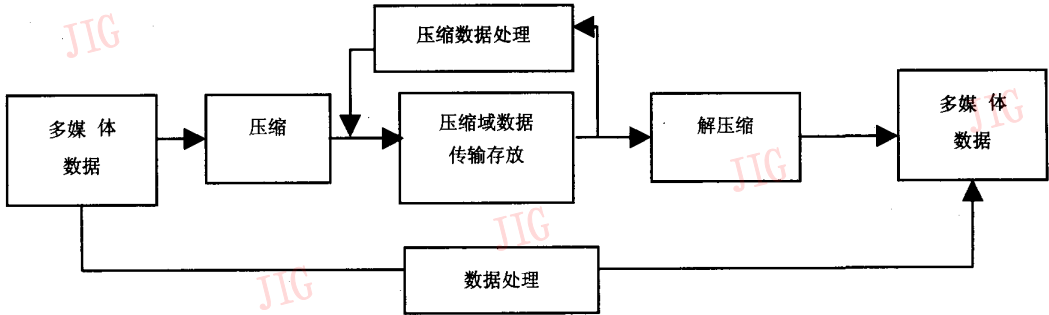


图2 压缩域多媒体数据处理模式

我们定义以下的概念来描述多媒体压缩域数据操作：

(1) 多媒体源数据 S 为从坐标集合 X 到值集合 V 的一个映射, 即 $S: X \rightarrow V$, 其中 X 为 S 的定义域, V 为 S 的值域, 即 $\text{Domain}(S) = X, \text{Range}(S) = V$ 。

(2) SC 为压缩数据, 也有 $SC: Y \rightarrow U$ 其中 Y 为 SC 的定义域, U 为 SC 的值域。

(3) 压缩变换 $T: S \rightarrow SC$ 。

(4) 压缩比 $CR = |S|/|SC|$, 其中 $|S|$ 与 $|SC|$ 为多媒体源数据集与压缩数据集的大小, 与 U 和 V 的大小也有关系。

(5) 多媒体数据操作 P 为 S 上的 n 次映射, 即 $P: S^n \rightarrow S$ 。

(6) P 所对应的压缩域对等操作 PP 定义为: $PP: SC^n \rightarrow SC$, 且满足以下条件:

$$T(P(S)) = PP(T(S))$$

(7) ϵ 邻域近似对等操作 PPA 定义为: $PPA: SC^n \rightarrow SC$, 且满足以下条件:

$$\|T(P(S)) - PPA(T(S))\| < \epsilon, \epsilon > 0$$

在上述概念的定义中, 应该说, 我们并未使用十分严格的数学定义, 而主要是为了简洁地说明问题。

根据以上的定义, 压缩域操作就是在压缩域数据集上进行源数据集操作的对等操作, 本文认为研

究的主要目标应该包括:

(1) 如何在压缩域数据集上寻找源数据集操作的对等操作;

(2) 如何降低对等操作的复杂度;

(3) 压缩域对等操作对源数据操作的加速比及其与压缩比的关系;

(4) 针对同一源数据集的不同压缩方法, 其相应的压缩域数据处理过程的加速效果是如何变化的, 也即如何选择更适合新的数据处理形式而又具有较高压缩效率的压缩方法;

(5) 对数据处理技术在源数据域与压缩数据域之间所表现出的不同性质与形式的理解;

(6) 压缩变换本身的机理。

近年来, 多媒体数据处理领域的学者围绕着以上问题, 进行了大量探讨, 取得了许多有意义的成果, 本文将作进一步讨论。

在上述问题中, 如何在压缩域数据集上寻找源数据集操作的对等操作是一个关键问题, 也是本领域研究的难点。我们知道对于某些数据操作而言, 很容易在一定的压缩域找到其对等操作, 如在源数据域的代数加减操作, 在 DCT 变换域仍然是代数加减操作; 而有些数据操作很难在一定的压缩域导出对等操作, 如图象直方图变换在 DCT 变换域就很难

处理;而对于某些数据处理技术,虽然可以严格推导出压缩域对等操作,但是其运算复杂度太高,影响其实用性。应当指出的是,这一问题的难易程度是与压缩方法和数据处理操作密切相关的,这说明,需要进一步研究新型压缩算法,并对数据处理操作作进一步的理解。

可以考虑以处理的最终效果为基础,通过多种方法在压缩域达到相同或相近的处理效果,目标是简化计算,达到提高处理速度的目的。虽然加速不是唯一目的,但是加速是压缩域操作技术具有的主要优势之一。针对这一问题,本文认为,可能的解决途径有:

- (1) 严格导出对等变换;
- (2) 在压缩域定义新的操作,达到同样效果;
- (3) 采用 ϵ 邻域近似对等操作;
- (4) 新的直接支持压缩域数据直操作的压缩算法。

2 技术分析

图象与视频数据压缩域处理技术的研究与应用比较活跃,目前的研究工作主要集中在 DCT 变换域的操作。其主要原因在于 DCT 变换是目前多媒体数据压缩标准,包括静止图象压缩标准 JPEG^[4]与视频压缩标准 MPEG^[5]等,进行压缩变换的一个主要手段。另一方面,DCT 变换域滤波技术^[6-9]早已为信号处理领域的学者所注意,对 DCT 变换的性质的研究相对比较成熟。

近年来,随着矢量量化、小波、分形等新的压缩算法的不断发展,压缩域数据操作研究的内容也大大丰富起来了。其中关于矢量量化图象的压缩域处理技术的研究工作取得了一些成果,如文献[10-13]等的工作,分别讨论了矢量量化压缩图象数据的边缘检测,纹理分析等操作,是压缩域图象数据处理领域中一个有意义的研究方向,值得我们加以注意。

文献[14]则讨论了 MPEG 音频压缩数据的直接操作,包括增益控制(Gain Control)、混响(Mixing)以及均衡化(Equalization),与传统的处理算法相比,压缩域处理算法大约快了5倍,而且得到的 MPEG 音频数据流保持了很高的质量。该算法是针对 MPEG-1层 I 和层 II 音频数据流实现的,是目前音频数据压缩域处理的一篇主要文献。

另外,一些学者在压缩算法的设计中同时也开始将压缩域操作的因素考虑在内,如下面将要介绍到的 Normal Form 算法^[15]。

下面我们将对 DCT 变换域操作的一些重要研究结果和 Normal Form 算法作进一步的分析。

2.1 DCT 变换域操作

在图象(或视频)DCT 变换域数据操作的研究中,文献[16]的影响较大。在他们的研究工作中,针对标准的 JPEG 数据流,提出并实现了诸如逐点相加、逐点相乘以及标量相加、标量相乘等的压缩域操作算法,得到了 50—100 倍的加速效果。这一工作的结果是令人鼓舞的。

不过,文献[16]所涉及到的数据操作比较简单,并且由于 DCT 变换的线性、正交的特性,这些数据操作的压缩域对等操作的推导较为直观,而且其变换形式同样也比较简单,实现也较为容易,只需要逐点的操作,充分利用了压缩域数据量大大减少的巨大优势。

CHANG 和 SHEN 则着眼于复杂的几何操作,针对图象的平移、放大与缩小、旋转等开展了一系列工作^[2,17-22]。虽然对于一些特殊的几何变换如 90° 旋转,可以方便地通过 DCT 变换矩阵的转置或符号转换来实现,但对于一般形式的几何变换,其压缩域对等操作的推导并非显而易见,而需要细致的算法设计,以达到良好的效果。CHANG 与 SHEN 分别独立地给出了自己的实现方式,二者均使用了逐行计算的方式,利用 2D 到 1D 的转换来简化计算,达到提高系统效率的目的。SHEN 将其算法称为变换域的 Scanline 算法,如对于图象任意角度的旋转,SHEN 提出将旋转过程分解成三次切变(shear),对每一次切变过程,分别在压缩域计算每根扫描线的偏移量,从而达到直接计算压缩域图象旋转的目的。他们的实验结果比传统的方法快 2 到 3 倍。

文献[3]对 2 的整数次幂的缩放操作在计算形式上进行简化,在压缩域采用了近似对等操作,在损失变换精度的前提下达到了较高的加速效果。文献[23]提出了一个不需乘法运算的近似压缩域线性操作算法,与精确运算相比,其结果图象质量几乎没有损失,而计算量大大降低。这些工作是研究近似对等操作的可喜尝试,同时也证明了本文提出的近似对等操作概念的可行性。

2.2 NORMAL FORM 算法

Normal Form 算法^[15]是考虑压缩域直接处理而设计的一个较为成功的压缩算法。传统的变换压缩方法一般通过一定的变换把图象变换到频率域,然后再进行编码处理以得到较高的压缩比,并保持较好的图象质量。如目前的 JPEG 与 MPEG 标准采用了 DCT 变换。不过,有些数据操作在 DCT 变换域的处理比较复杂,如图象的任意几何变换。Normal Form 算法试图为压缩域直接处理寻求更加自然的实现方式。由于其编码过程在几何域进行,其相应的压缩域数据操作,特别是几何变换操作的处理相当直观。这一算法的主要思路是:通过在几何域对图象进行分解,把图象表示为背景与叠加于其上的所谓规范形(Normal Form),然后通过滤波、矢量量化和信息熵编码达到最终的压缩效果。可以认为,这是一种基于模型的压缩方法。其主要特点是:

(1) Normal Form 的尺度较小,通常为几个象素点,保证了图象质量。

(2) 图象的细节使用低阶多项式来表示,计算简单。

(3) Normal Form 分为三类,每一类中用一些参数来精确定义图象结构。包括 ridge、edge 和 hill。在相应的局部坐标系中,它们分别表示为:

- ridge: $r(x, y) = a + dy^2 \quad -h < y < h$
 $-l < x < l$, 否则为零;
- edge: $e(x, y) = a + dy^3 \quad -h < y < h$
 $-l < x < l$, 否则为零;
- hill: $h(x, y) = a + d_1x^2 + d_2y^2$,
 $d_1x^2 + d_2y^2 < h^2$, 否则为零。

其中,局部坐标系中心、其对原坐标系的旋转角以及 a, d, h, l, d_1, d_2 等均为可调的参数。

我们现在来具体的看一下一些常见的图象操作是如何在 Normal Form 结构中实现的。由于编码压缩过程是在几何空间进行的,图象操作尤其是几何变换的压缩域对等操作几乎可以不加改变地直接由源数据域操作得到:

(1) 边缘检测

边缘是 Normal Form 之一,可以直接从模型数据中得到。

(2) 放大缩小

在该算法中,模型的表示是与尺度无关的。给定的具体的图象空间尺度,任意象素点的亮度值都可以直接由模型数据计算出来,因而无须进行额外的

处理。即图象缩放不占时间,同一模型可以计算任意尺度的图象。

(3) 图象旋转

图象的旋转是通过改变各类模型的角度来实现的。即图象旋转的对等操作为改变模型的参数,即局部坐标系的旋转角。

(4) 图象增强

图象的锐化或平滑是通过改变各类模型的亮度参数来实现的。

Normal Form 算法中的模型参数是需要经过进一步的压缩方法,如矢量量化,来获得高压缩比的。在实际的应用中,这一因素是必须加以考虑的。本文介绍这一算法的目的是为了介绍其思想,即强调在新一代的压缩算法的设计与实现当中,应当将压缩后的数据如何直接操作问题考虑在内。即将出台的 MPEG-4 多媒体数据压缩标准以基于内容的压缩与检索为主要特征,这为压缩域多媒体数据处理技术的研究开拓了非常良好的前景。

3 结论与展望

综上所述,压缩域数据处理技术具有以下优势:

(1) 压缩域大大减少的数据量对于总体处理效率的提高是一个有利因素;

(2) 压缩域数据,如 DCT 系数对于某些频率域的操作,适应了处理的要求;

(3) 压缩域数据处理技术省略了解压缩/再压缩的附加环节,对总体处理时间的减少是个有利因素;

(4) 压缩数据格式作为多媒体数据的主要存储和传输格式具有普遍性。

然而,压缩域数据处理技术作为新兴的技术还远未成熟,许多问题有待解决。缺乏统一的理论支持是主要问题,包括:如何有效地推导压缩域数据处理的对等变换形式,压缩比应当达到多高才能保证压缩域的处理比对原始数据的直接处理效率更高,压缩方式的分类以及数据处理过程的分类等,诸多理论问题有待解决。压缩域数据处理具有很强的实用价值。即使数据的减少被复杂的压缩域处理过程所抵消,没有达到预期的加速效果,从压缩域数据处理另一个出发点即压缩数据格式存在的普遍性考虑,只要有相近的处理速度,压缩域算法仍有其实际意义。

参 考 文 献

- 1 Boon-Lock Yeo. Visual content highlighting via automatic extraction of embeded captions on MPEG compressed video. SPIE, Digital Video Compression Algorithms and Technologies, 1996, 2668.
- 2 Shen B, Sethi I K. Scanline algorithm in compressed domain. In: Proc SPIE, Digital Video Compression Algorithms and Technologies, 1996, 2668.
- 3 Hu Q, Panchanathan S. Spatial scalability in compressed domain. In: Proc SPIE, Digital Video Compression Algorithms and Technologies, 1996, 2668.
- 4 Wallace G K. The JPEG still picture compression standard. Communications of the ACM, 1991, 34:(4).
- 5 Gall D L. MPEG: A video compression standard for multimedia applications. Communications of the ACM, 1991, 34(4).
- 6 Chen W H, Fralick S C. Image enhancement using cosine transform filtering. In: Image Sci Math Symp, Monterey, CA, Nov 1976.
- 7 Lee J B, Lee B G. Transform domain filtering based on pipelining structure. IEEE Trans on Signal Processing, 1992, 40(8):2061 ~ 2064.
- 8 Chitprasert B, Rao K R. Discrete cosine transform filtering. Signal Processing, 1990, 19:233 ~ 245.
- 9 Ngan K N, Clarke R J. Lowpass filtering in the cosine transform domain. Int Conf on Communication, Seattle, WA, June 1980, 37.7.1 ~ 37.7.5.
- 10 Cosman P C, Oehler K L, Riskin E A *et al.* Using vector quantization for image processing. In: Proc IEEE, 1993, 81(9):1326 ~ 1341.
- 11 McLean G F. Codebook edge detection. CVGIP: Graphical Models Image Process. 1993, 55(1):48 ~ 57.
- 12 McLean G F. Vector quantization for texture classification. IEEE Trans Syst Man Cybernet, 1993, 23(3).
- 13 Zhou Jie, Peng Jiaoxiong, Ding Mingyue. Improved codebook edge detection. CVGIP: Graphical Models Image Process, 1993, 57(6):48 ~ 57.
- 14 Broadhead M A, Owen C B. Direct manipulation of MPEG compressed digital audio. In: Proceedings of the ACM Multimedia '95, San Francisco, CA USA, 1995, 499 ~ 507.
- 15 Yomdin Y, Elichai Y. Normal forms representation: A technology for image compression. SPIE, Image and Video Processing, 1993, 1903.
- 16 Smith B C, Rowe L. Algorithms for manipulating compressed images. IEEE Computer Graphics and Applications, Sept 1993, 3442.
- 17 Chang S-F, Messerschmitt D G. A new approach to decoding and compositing motion-compensated DCT-based images. In: ICASSP '93, Minneapolis, MN, April 1993.
- 18 Chang S-F. Compositing and manipulation of video signals for multimedia network video services: [Dissertation]. USA: U. C. Berkeley, Aug 1993.
- 19 Chang S-F. New algorithms for processing images in the transform compressed domain. In: Proc SPIE Visual Communications and Image Processing, 1995.
- 20 Chang S-F. Compressed-domain techniques for image/video indexing and manipulation. IEEE Intern Conf on Image Processing, ICIP '95, Special Session on Digital Image/Video Libraries and Video-on-demand, Washington DC, Oct 1995.
- 21 Chang S-F, Messerschmitt D G. Manipulation and compositing of MC-DCT compressed video. IEEE JSAC Special Issue on Intelligent Signal Processing, 1995, 13(1):1 ~ 11.
- 22 Shen B, Sethi I K. Scanline algorithms in the JPEG discrete cosine transform compressed domain. Journal of Electronic Imaging, 1996, 5(2):182 ~ 190.
- 23 Merhav N. Multiplication-free approximate algorithms for compressed domain linear operations on images. Tech Reports, HP Laboratory, 1998.



向 辉 1993 年毕业于中国科技大学计算机系, 1996 年在浙江大学计算机系获硕士学位, 现在浙江大学计算机系攻读博士学位。主要研究方向为多媒体数据处理技术与信息重组理论。

石教英 毕业于前苏联列宁格勒大学。现为浙江大学计算机系教授, 博士生导师, 浙江大学 CAD&CG 国家重点实验室主任。主要研究方向为科学计算可视化、虚拟环境及多媒体计算理论。

The Study of Compressed Domain Multimedia Data Processing

Xiang Hui, Shi Jiaoying

(State Key Laboratory of CAD & CG, Zhejiang University, Hangzhou 310027)

Abstract Multimedia data compression is a key technology in the fields of Multimedia, Networking & Communication and Computer applications. Vast amounts of multimedia data like images, video or audio data make it necessary to store and transfer multimedia data in its compressed format. However, multimedia applications require technologies permitting us to process multimedia data as flexible as we can. The utilizing of compression makes the direct processing of multimedia data more difficult. So, compressed domain processing of multimedia data has become an interesting research field. In this paper, the basic concepts and theoretical topics are discussed at first and a processing model is given. Then, some important research directions are discussed and analyzed.

Keywords Multimedia data compression, Compressed domain data processing, Compression ratio, Transform domain