

一种运动图像安全传输与存储平台编码算法的研究^{*}

The Research of the Encoding Arithmetic of Moving Picture's Secure Transfers and Storage Platform

胡 泉 邓 锐 江 立 肖 明

(华中师范大学计算机科学系 武汉430079)

Abstract Today, with the great progress in Wide Band, how to compress and encode moving picture, how to securely transfer it with high speed through the network, and how to store it, is still the hotspot problem in the world. This article based on the succeed in the research of a moving picture's secure transfers and storage platform, discusses a useful moving picture encoding arithmetic.

Keywords Moving picture, Secure transfers, Encoding arithmetic

1. 前言

在高速宽带网发展的今天,静态图像和运动图像的安全传输已成为现实。通过图像采集系统选择具有实用价值的运动图像,高速、安全地传输到计算机网络系统中的某个指定结点,实时或非实时地存储与显示,具有十分重要的实用价值和应用前景,其需求之大已超出人们的预料。

运动图像高速安全传输与存储平台的一个重要应用环境是社会“安保系统”。传统的安全系统是采用录像带保存技术,即将采集的运动图像保存到录像带中,再通过电视机进行播放。录像带保存影像不仅成本高、体积大,而且因换带而间断信息的捕获和存储,可能造成有用信息的丢失。另外,录像带容易变质失效,给长期保存、归档和管理带来很大困难,若改用计算机安保系统,可以将采集的运动图像通过网络传送到指定的计算机中,这样既可以无间断,也可以长期保存;既不会变质,也容易归档和管理。

但是,运动图像占用存储空间特别大,如果直接存储,一天24小时的运动图像,至少要占用几百G的磁盘空间,这样不仅成本高,而且根本不实用。因此,需要对运动图像进行压缩、编码,使其在有限的存储器中,保存更长时间的视频信息。另一方面,由于大量运动图像的网络传输造成的网络拥塞也是频繁发生,目前解决网络拥塞的有效方法是对运动图像的压缩编码。由此可见,压缩编码技术是研究运动图像高速、安全传输的核心技术之一。目前,关于压缩编码算法的研究是国内学者们的热门研究课题之一。

2. 运动图像压缩编码算法的研究

文[2]介绍了一种运动图像安全传输与存储平台的体系结构和实现算法,该平台的运动图像编码采用了运动补偿技术(Motion-Compensated),同时采用了离散余弦变换(DCT)、小波变换(Wavelet)和分形图像编码等多种压缩编码算法。

2.1 运动图像的离散余弦变换(DCT)算法

如果 $X = (x_1, x_2, \dots, x_N)^T$ 是由 N 个信号样本构成的列向量(也可以称 X 为矢量信号),

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ \vdots & \vdots & & \vdots \\ a_{N1} & a_{N2} & \dots & a_{NN} \end{bmatrix}$$

是一个 $N \times N$ 的矩阵,则

$$Y = AX \quad (1)$$

其中 A 是线性变换 X 的核矩阵,DCT 的核矩阵为

$$A = \sqrt{\frac{2}{M}} [c(k) \cos \frac{(2m+1)k\pi}{2M}]_{M \times M} \quad (2)$$

其中 k 为行, m 为列, $k, m = 0, 1, \dots, M-1$

将式2代入到式1中,得到 $M \times 1$ 矢量信号 $x = [x(0)x(1)\dots x(M-1)]^T$ 的一维离散余弦矢量信号

$$x = [x(0)x(1)\dots x(M-1)]^T,$$

即

$$X = Ax \quad (3)$$

其中

$$c(k) = \begin{cases} \frac{1}{\sqrt{2}} & k=0; \\ 1 & k=1, 2, \dots, M-1 \end{cases}$$

^{*}湖北省科技攻关项目

可以验证 A 是一个正交矩阵,但不是对称矩阵。而反变换(IDCT)矩阵根据正交性为:

$$A^{-1} = A^T = \sqrt{\frac{2}{M}} [c(k) \cos \frac{(2m+1)k\pi}{2M}]_{M \times M} \quad (4)$$

其中, m 为行, k 为列, $m, k = 0, 1, 2, \dots, M-1$

式4与式2相比较, A^{-1} 和 A^T 除了行、列序号互换外,形式上与式2完全相同。由此可以得到反变换的矢量形式

$$x = A^T X \quad (5)$$

式3的 DCT 与式5的 IDCT 也可以分别写成如下的一维求和形式

$$x(k) = \sqrt{\frac{2}{M}} c(k) \sum_{m=0}^{M-1} x(m) \cos \frac{(2m+1)k\pi}{2M}, k=0, 1, \dots, M-1 \quad (6)$$

$$x(m) = \sqrt{\frac{2}{M}} \sum_{k=0}^{M-1} c(k) X(k) \cos \frac{(2m+1)k\pi}{2M}, m=0, 1, \dots, M-1 \quad (7)$$

再将数字图像 $x(m, n)$ 看成是一个 $M \times N$ 的矩阵,借助二维 DCT,可以将图像从空间域(即 mn 平面)变换到 DCT 域(即 kl 平面)。以求和形成定义的二维 DCT 为:

$$\begin{aligned} X(k, l) &= \frac{2}{\sqrt{MN}} c(k) c(l) \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x(m, n) \cos \frac{(2m+1)k\pi}{2M} \cos \frac{(2n+1)l\pi}{2N} \\ &= \sqrt{\frac{2}{M}} c(k) \sum_{m=0}^{M-1} \left\{ \sqrt{\frac{2}{N}} c(l) \sum_{n=0}^{N-1} x(m, n) \cos \frac{(2n+1)l\pi}{2N} \right\} \cos \frac{(2m+1)k\pi}{2M} \end{aligned} \quad (8)$$

其中 $k=0, 1, \dots, M-1; l=0, 1, \dots, N-1$

式8中 DCT 的变换核可表示为:

$$g(m, n; k, l) = c(k) c(l) \cos \frac{(2m+1)k\pi}{2M} \cos \frac{(2n+1)l\pi}{2N} = u(m, k) * v(n, l) \quad (9)$$

与式8相对应的二维 DCT 标量表达式为:

$$x(m, n) = \frac{2}{\sqrt{NM}} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} c(k) c(l) x(k, l) \cos \frac{(2m+1)k\pi}{2M} \cos \frac{(2n+1)l\pi}{2N} \quad (10)$$

$m=0, 1, \dots, M-1; n=0, 1, \dots, N-1$

由此可见,二维 IDCT 的变换核也是可分离的。即二维 DCT, IDCT 都可以先逐行(或逐列),再逐行(或逐列)用一维 DCT, IDCT 直接变换。这种行列分离算法结构简单,容易实现。

2.2 运动图像小波变换算法

小波变换同时具有频率域和空间域两大能量紧缩

特性,该特性一方面使得大部分的图像能量总是集中在最低频率的子图像中,并从低频到高频呈递减分布趋势;另一方面,各子图像对应相同空间位置的像素间存在着较强的空间相关性,并且相应的系统从低频到高频呈很好的尺度级顺序递减。

对于一幅 $M \times N$ 源图作一级二维小波变换可得4幅子图像:低低频(LL)、低高频(LH)、高低频(HL)和高高频(HH),对 LL 子图像再作二级小波变换后又得4幅子图像,如此重复 K 次后可得 K 级小波变换,这些子图像可看成“金字塔”结构,并将该金字塔结构中的每个点用一个4元组定义为:

$$x(j, k, x, y)$$

其中, $j = -1, 0, 1, 2$, 用于表示频率属性或方向属性,即 LL, LH, HL, HH 等; $K = 1, 2, \dots, K_{\max}$, 用于表示分解级数; x, y 表示位置属性 Ω , 即 $(x, y) \in \Omega_k$, 而 Ω_k 随着分解级数(k)的增加逐次递减为原来的 $1/4$ 。

小波系数在点 $X = (j, k, x, y)$ 的值记为 $C(x)$ 或 $C = (j, k, x, y)$, 系数图像记为

$$C_{j,k} = \{C(X) | X = (j, k, x, y), (x, y) \in \Omega_k\} \quad (11)$$

金字塔的分解过程可表示为:

$$W(C_{-1,k}) = \{C_{j,k+1} | j = -1, 0, 1, 2\} \quad (12)$$

其逆变换为

$$W^{-1}(\{C_{j,k} | j = -1, 0, 1, 2\}) = C_{-1,k-1} \quad (13)$$

相对于点 $X = (j, k, x, y)$ 分解后的子图像为

$$C(X) = \begin{cases} \{(j, k-1, 2x, 2y), (j, k-1, 2x, 2y+1), \\ (j, k-1, 2x, 2y+1), (j, k-1, 2x+1, 2y)\} \\ \text{空集 } \Phi \end{cases} \quad (14)$$

若递归地定义

$$D(X) = \{O(X), D(X') | X' \in O(X)\}$$

为点 $x(j, k, x, y)$ 的家族(即 x 的子图像及其后代全体),且在给定阈值 T_b 的条件下以 x 为树根的结构若满足

$$\max(|C(x')|) < T_b$$

则称

$$Tree(x) = [x, D(x)]$$

为关于 T_b 的零树。我们利用零树的特性可以设计出多种压缩比很高的编码算法。

2.3 运动图像分形编码思想

分形是一些简单空间上点的集合,称之为“分形集”。分形集具有任意小尺度下的不变形性,即无论尺度如何变化,景物的任一小部分的形状都与较大部分的形状相似,即分形集具有显著的自相似性或自仿射性。图像压缩编码技术正是最大限度地利用了分形集的这种自相似性。利用相似性也可以设计出多种压缩

(下转第77页)

对应的矩阵列上,这 m 个关键词对应的元素的分值增加。这样这个矩阵中的每一列就反映了一个搜索引擎的综合查询能力。系统对增加新的关键词、合并重复的关键词,进行统一管理,保证各子系统统一。

(4)数据库更新 首先,要保证数据库内信息的有效性(信息源是否被更新或取消)。这只需要令 Robot 定期将数据库内的信息与信息源进行核对即可。

其次,系统在索引本地资源的同时,对符合用户信息需求的外部信息进行缓存(即存储用户点击的外部页面索引),出于对系统性能的考虑,每一个本地数据库中的信息量都不能超过一定的上限,因此本地搜索引擎只能保存一部分网络信息,那么最优的选择就是保留那些本地用户最需要的网络资源,然而很难在系统刚刚开始使用时就准确地知道本地用户最需要哪些信息。而且系统在使用过程中,由于网络资源的不断变化,用户组成的不断变化,用户兴趣的不断变化等等,用户最需要的信息也会不断变化,因此需要不断更新数据库内存储的信息,以便跟上用户需求的变化。

利用用户的反馈来跟踪用户需求的变化,并不断淘汰数据库内过时的信息,添加新的信息。具体的做法是:当一条新的信息被存入数据库时,它被赋予一定的“生命力”;当它符合某个用户的查询要求(被点击)时,便增加其生命力;生命力随着时间的流逝而减小,根据系统容量,系统保留生命力大的记录。这种机制保证了数据库只存储用户经常需要的信息。

(上接第73页)

比高的编码算法。例如:

首先将原始图像划分为值域块和定义域块两种大小的块:值域块是完全无重叠的 n^2 小块 $A_1, A_2, A_3, \dots, A_{n^2}$; 定义域块是允许有部分重叠的 m^2 大块 $B_1, B_2, B_3, \dots, B_{m^2}$ (其中 $m > n$)。

然后寻找合适的仿射变换 ω_k 和定义域块 B_k ,使得 $R_i = \omega_k(B_k)$, 其中三维仿射变换 $\omega: R^2 \rightarrow R^2$ 的简化形式为:

$$\omega \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & u \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} e \\ f \\ g \end{pmatrix} \quad (15)$$

在实际应用中,用式16的等价组合变换来代替式15才具有实现的可行性:

$$\omega_i = G_i \sigma_i \rho_i \Psi_i \quad (16)$$

其中, G_i 为灰度处理算子,包括比例因子 u 和灰度补偿因子 g ; σ_i 为1个对折和旋转变换(通常 $i=8$); Ψ_i 是 (x, y) 平面上的紧缩变换,即将大小为 m^2 的 D_i 映射成大小为 n^2 的块。这里

$$K_i = G_i \sigma_i \rho_i \Psi_i(D_i) \quad (17)$$

最后,只要找到了仿射变换 ω 和定义域块 D_i ,并

搜索引擎数据库也存在对搜索引擎的更新的问题,距今越久远的记录,对系统决策的影响应该越小。由于搜索引擎数据库内存贮的是每个搜索引擎对每个关键词的查询能力,都是0~1之间的数值,数值越大表示查询能力越强。因此通过定期让所有数值衰减,达到更新数据库的目的。

结论 采用分布式的结构设计购书系统的搜索引擎,各个节点搜索引擎之间的竞争能使它们不断完善自己,并形成各自的特色,合作又使它们可以互相补充、互相依赖。提出的数据库小型化、本地化等设计原则注重于提高系统的性能与服务质量,特别是充分利用了用户的反馈信息,可以使系统根据用户的不同需求及时调整自身的运作,以提供最符合用户需求的购书信息,从而提高了系统的服务质量。

参考文献

- 1 汪晓岩,胡庆生,李斌,庄镇泉.面向 Internet 的个性化智能搜索引擎.计算机研究与发展,1999,36(9):1039~1044
- 2 邹涛,王继成,朱华宇,金翔宇,张福炎 WWW 上的信息挖掘技术及实现.计算机研究与发展,1999,36(8):1019~1023
- 3 张卫丰,徐宝文.Web 搜索引擎框架研究.计算机研究与发展,2000,37(3):376~378
- 4 王军玲,赵沁平.一种基于类比的启发式搜索方法.计算机科学,1998,25(5):33~37
- 5 常桂然,张晓辉.Web 信息检索服务系统与搜索引擎.计算机科学,1998,25(5):86~90
- 6 文栋辉,李光亚,赵振西.支持开放性的超媒体引擎.计算机科学,1998,25(5):107~109

分别将其参数存储保存,待所有的值域子块都编码之后,就可以完成运动图像的整体分形编码。

结束语 本文在文[2]的基础上介绍了一种运动图像安全传输与存储平台的实现技术,重点研究了实现该平台的压缩编码算法。该平台采用现代高速网络技术,安全传输运动图像;采用国际上最先进的计算机存储技术,智能化地长期保存运动图像信息;采用国际上最新的软、硬件关键技术开发整个系统。其突出的优点是可靠性高、体积小、成本低,可智能化地、长期不间断地保存运动图像信息,而且不会发生影像失真或变质,所以归档和管理容易,适合在我国大范围推广使用。

参考文献

- 1 Tekalp A M. Digital Video processing. Prentice-Hall International, 1997
- 2 胡泉,邓锐,江立,肖明.一种运动图像安全传输与存储平台的研究.华中师范大学学报(自然科学版),2001,8
- 3 钟玉琢,乔秉新,祁卫.运动图像及其伴音通用编码国际标准——MPEG2.清华大学出版社,1997
- 4 钟玉琢,沈伟铨,沈洪.多媒体技术基础及应用.清华大学出版社,2000
- 5 吴乐南.数据压缩.东南大学出版社,2000
- 6 杨福生著.小波变换的工程分析与应用.科学出版社,2000