

MPEG-4 中基于 DT 网格编码算法的实现

韩军功, 卢朝阳, 田毅珂, 高西全

(西安电子科技大学通信工程学院 ISN 国家重点实验室, 陕西西安 710071)

摘 要: 本文按照 MPEG-4 标准对 DT 网格对象编码的要求, 设计了一套完整的基于 DT 网格的编码算法. 首先提出了网格对象编码的系统框架. 在此基础上, 着重研究了网格运动编码, 提出了一种新的三角形运动连续预测方法及相应的残差图像处理方法. 最后, 对整个数据结构进行转换, 使之完全按照标准规定的顺序进行编码. 本文算法同 H. 263 进行了全面的比较. 结果表明, DT 作为 MPEG-4 标准中的一种编码方法, 有一定的实用价值.

关键词: 网格对象; 连续预测; Delaunay 三角形; 数据结构转换

中图分类号: TP391. 4 **文献标识码:** A **文章编号:** 0372-2112 (2003) 04-0631-04

Implementation of Coding Algorithm Based on DT-mesh in MPEG-4

HAN Jun-gong, LU Zhao-yang, TIAN Yi-ke, GAO Xi-quan

(School of telecommunication engineering, ISN national key laboratory, Xidian Univ, Xi'an, Shaanxi 710071, China)

Abstract: This paper designs an integrated coding algorithm based on DT-mesh according to depiction of MPEG-4 standard. Firstly, a framework about mesh object coding is proposed, on which we study mesh motion coding, and propose a new series estimation algorithm for triangle motion and a corresponding method for residual image. Finally, data structure is transformed in order to fulfilling the coding order of MPEG-4. Compared with H. 263 standard, this algorithm has utilitarian value as a method of MPEG-4.

Key words: mesh object; series estimation; Delaunay triangulation; datastructure transformation

1 引言

MPEG-4 是 ISO/IEC 颁布的一个新的视、音频编码标准. 它是一个开放式的基于对象的编码标准, 开发者可以加入新的有效的算法模块. 综合性是其另一个特点, 除了完全支持 MPEG-x 已经提供的所有功能外, 它的基本编码结构中加入了静态纹理编码、网格对象编码等新的编码策略.

DT 网格编码是利用 Delaunay 三角形 (简称 DT) 网格优良的几何特性、连接关系的唯一性, 把图像自适应地划分成 DT 网格. 编码时充分利用三角形的顶点信息, 其视频帧间运动用三角形网格的形变来描述. 文献 [1] ~ [3] 给出了网格基运动估计和运动补偿的基本方法, 文献 [4] 给出了一种网格基编码方法. 从这些成果看, 基于 DT 网格的运动估计和运动补偿算法与传统的基于“矩形块匹配” (Block Matching Algorithm 简称 BMA) 的算法相比, 有一些明显的优越性. 本文从 MPEG-4 标准出发, 设计了完整的基于 DT 网格的编码算法, 提出了一种新的适合标准使用的运动连续预测方法. 作为参考, 本文同 H. 263 进行了全面的比较. 结果表明, 本文算法作为兼容标准的一种方法, 有一定的优越性和实用价值.

2 网格对象编码系统框架

本文的网格对象编码主要针对 MPEG-4 中的二维网格.

二维网格对象是由多个三角形拼接成的任意形状的平面图形. 一个动态的二维网格由一个二维三角形的时间序列组成, 这些网格必须有相同的拓扑结构, 但它们的顶点位置彼此不同. 图 1 是网格编码的基本框图.

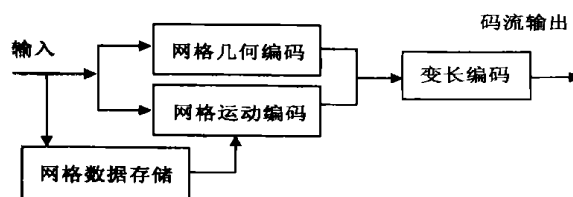


图1 网格对象编码框图

网格几何编码用于编码初始的网格对象. 在标准中分为两种情况, 一种是采用规则网格, 用五个参数来编码完整的三角形结构和节点位置; 另一种就是 Delaunay 网格, 经 DT 描述后的网格按一定顺序编码节点的位置和灰度值. 在译码端, 利用三角形的顶点信息, 使用线性插值的方法得到三角形内各点的近似灰度值. 本文第四节结合数据结构介绍网格几何编码.

网格运动编码用于编码当前网格相对于上一帧网格的形变信息. 编码时, 对网格节点的运动矢量按一定顺序作差, 对差值进行编码并输出码流.

3 网络运动编码

网络运动编码是利用三角形网格的运动估计和补偿技术^[1]来产生运动矢量,并对运动矢量按照一定顺序进行编码.文献[3]中给出了一种实现方法,它先进行三角形网格顶点初始运动矢量的估计,形成初始的运动矢量,然后进一步优化这个运动矢量,得到最终的运动矢量.本文在整体思路类似于文[3],但改进了文[3]中的运动估计方法,以使其兼容 MPEG 4 的要求.图 2 是本文网络运动编码的简要框图.

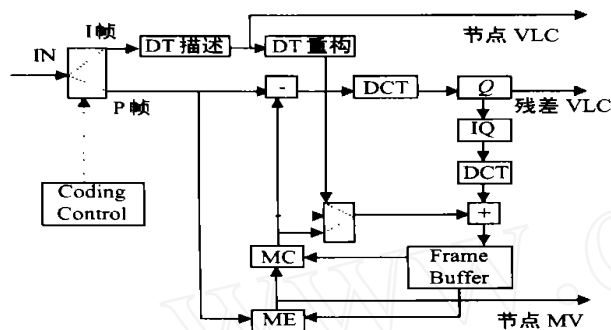


图2 网络运动编码框图

3.1 运动的连续估计方法

文献[3]中运动估计的方法是,当预测 k 帧时,首先对 $k-1$ 帧的重构图像进行 DT 描述,然后用这个 DT 结构在 k 帧中搜索最佳的匹配位置,进行形变估计.也就是说,一个参考帧仅预测一帧图像,在预测前需对参考帧进行 DT 描述,而译码端必须对每一帧图像都进行 DT 重构,故译码效率不高.并且,预测过程中的 DT 结构要经常变化,这不符合 MPEG4 标准中运动预测过程 DT 结构要保持不变的要求.本文的连续预测用一个 I 帧连续预测 10 个 P 帧,只对 I 帧进行 DT 描述,在整个预测过程中 DT 结构保持不变.在译码端只需对 I 帧的 DT 结构进行恢复即可,从而提高了译码速度.

本文运动连续预测采用的方法是:对 k 帧进行运动估计时, $k-1$ 帧为参考帧;对 $k+1$ 帧进行预测时, k 帧的恢复图像为参考帧.如图 3 所示, k 帧中的三角形 $A^1 B^1 C^1$ 是 $k-1$ 帧中的三角形 ABC 由于帧间运动产生的形变,而 $k+1$ 中的 $A^2 B^2 C^2$ 是 $A^1 B^1 C^1$ 的形变.映射变换可由六参数仿射变换进行:

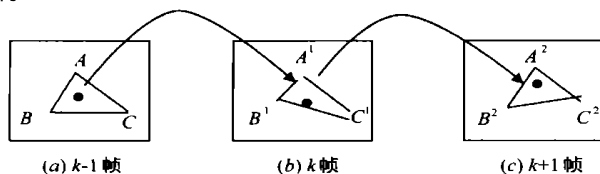


图3 连续预测示例图

$$\begin{aligned} x &= a_{11}x + a_{12}y + a_{13} \\ y &= a_{21}x + a_{22}y + a_{23} \end{aligned} \quad (1)$$

式(1)中, (x, y) 表示 ABC 内部原像素的坐标, (x, y) 表示在 $A^1 B^1 C^1$ 上的位移.把已经获得的三个顶点的 x 和 y 分别代入,解方程组即可确定 a_{11} 、 a_{12} 等六个系数.那么,

ABC 内任意一点 (x, y) 的运动矢量即可唯一地由式(1)确定.若设 (x, y) 、 (x, y) 为像素点运动前后的坐标,则 $x = x + x, y = y + y$.假设图像像素点的灰度值从一点运动到另一点保持不变,可得到预测图像的灰度值为:

$$T_k(x, y) = I_{k-1}(x, y) \quad (2)$$

若映射后的坐标值不为整数,则用双线性内插的方法从周围四个整数节点的亮度值计算新的亮度值.

对于 $k+1$ 帧中的 $A^2 B^2 C^2$,用第 k 帧恢复图像中由 $k-1$ 帧形变后产生的 $A^1 B^1 C^1$ 进行连续预测,预测方法完全一样.

本文的连续预测对节点的形变范围做了限制,限制节点的形变不能超出其所在“六边形”区域^[1].所谓“六边形”区域是指某节点和它周围节点形成的所有三角形组成的多边形区域.这样的限制是因为在图像纹理复杂区域或存在局部运动区域有许多面积较小的三角形,只有限制了节点的运动范围,才能确保在一个视频段的运动估计过程中网格的拓扑结构不变.另外,在三角形连续形变的过程中可能会出现六参数方程无法求解的情况.对式(1)中的 a_{11} 、 a_{12} 、 a_{13} 求解并写成式(3)的矩阵形式.设三个顶点的坐标为 (x_1, y_1) 、 (x_2, y_2) 、 (x_3, y_3) , x 方向上的运动矢量分别为 u_1 、 u_2 、 u_3 .

$$\begin{bmatrix} a_{11} \\ a_{12} \\ a_{13} \end{bmatrix} = \frac{1}{2K} \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \quad (3)$$

式中: $K = 1/2(a_1 b_2 - a_2 b_1)$,表示三角形的面积. $a_1 = y_2 - y_3$; $a_2 = y_3 - y_1$; $a_3 = y_2 - y_1$; $b_1 = x_3 - x_2$; $b_2 = x_1 - x_3$; $b_3 = x_2 - x_1$; $c_1 = x_2 y_3 - x_3 y_2$; $c_2 = x_1 y_3 - x_3 y_1$; $c_3 = x_2 y_1 - x_1 y_2$.

当 K 值为零时,方程就无法求解.这意味着三角形三个顶点共线,这主要是三角形的连续形变所致.本文对这种情况进行了处理,首先对这个三角形进行标记,再用三个顶点把这条直线分成两段,每一段上点的灰度预测值用它们的两个端点的灰度预测值的平均值表示.并且自动的把下一帧设为 I 帧,重新进行 DT 描述,再连续预测.本文算法同文献[3]的方法进行了对比.不难看出,本文算法有效地提高了编解码的效率.

表1 算法效率比较

Miss America 前 50 帧	平均 PSNR(dB)	平均编码时间	平均译码时间
文献[3]方法	35.628	3.34 秒/帧	0.078 秒/帧
本文方法	35.547	2.71 秒/帧	0.038 秒/帧

3.2 残差图像处理

残差图像是原始图像和预测图像的差值图像,对这部分处理的好坏严重影响图像的恢复质量.本文的方法是,按宏块统计均方误差和均值,判断宏块它们是否大于预先设的门限,如果大于,则说明该宏块内的像素预测不是很精确,需要进行残差处理.若小于预设门限,则不用处理残差数据.

把宏块残差分成四个子块并分别进行 DCT 变换和量化因子为 8 的定量化,然后检测各个子块是否为零并用 CBPY 标记.如果 CBPY 非零则编码输出,否则,跳过对该子块的处理.在解码端根据残差位置和 CBPY 恢复出残差图像.实验表

明,本文策略可以在保证图像质量的前提下提高压缩比。

4 网格几何编码及数据结构转换

本文的网格几何编码采用文献[5]中的 DT 模型基静止图像编码及其优化方法。我们改进了文[5]中 DT 描述后的数据结构,以使其符合 MPEG-4 中码流的语法规则。为了方便、高效地描述三角形的空间几何关系,文献[6]设计了一个数据表,用于存储 DT 三角形及其周边三角形的信息,结合表 2 说明其内容。

表 2 DT 描述中的数据结构表

三角形序号	顶点序号			相邻三角形序号		
0	1	2	5	7	5	- 1
1	0	1	7	0	9	- 1
2	2	3	6	4	7	- 1
3	3	0	4	9	4	- 1
4	3	4	6	6	2	3
5	1	5	7	8	1	0
6	4	5	6	7	4	8
7	5	2	6	2	6	0
8	5	4	7	9	5	6
9	4	0	7	1	8	3

第一列表示按生成顺序记录的三角形序号,第 2~4 列表示这个三角形中按照逆时针排列的三个顶点的编号,第 5~7 列分别表示 2~4 列顶点的对边相邻的三角形序号。(表中 - 1 表示对边在图像的边界上)。这个表便于三角形或节点的增删,很好地描述了 DT 三角形及其周边信息。但是,它的结构和顺序并不符合 MPEG-4 所规定的格式,所以必须对数据结构进行转变。

定义 1 三角形节点中具有最小 $x + y$ 值的节点 N 称为基点(若多个点 $x + y$ 值相同,则 y 值小的为基点)。

定义 2 由基点和边界上顺时针方向上下一个节点组成的边所在的三角形为初始三角形。

根据定义 1、2 可以对数据结构进行重新初始化,找出初始的三角形,并编号为 0。

定义 3 基边,当前三角形和已经标记的邻居三角形中具有最小序号的三角形连接的边。

本文按照标准的规定^[7]结合表 2 对三角形进行了重新标记。流程如下:假设当前三角形已经被标记过,记为 m 。首先检查是否有未标记的三角形和当前三角形共享它的右边,若有则标记为 $m + 1$ 。若无则跳过,不予处理;再检查是否有未标记的三角形和当前三角形共享它的左边,若有则在前一个序号的基础上加 1,若无则跳过,不用处理。在与没有标记的三角形相邻的所有已标记的三角形中找到序号最小的三角形,这个三角形作为当前三角形重复上边的过程,直到所有三角形都被标记。标记完三角形后,按序号从小到大对其中的节点进行标记。同一个三角形中,按顺时针顺序编号节点。表 3 是用上述算法进行变换后的结果,它的第 5、6 列分别记录与当前三角形右、左邻边三角形的序号。

表 3 转换后的标准的数据结构

三角形序号	顶点序号			右、左三角形序号	
0	0	1	2	9	4
1	0	3	2	1	8
2	1	2	4	6	2
3	0	5	3	- 1	5
4	3	6	2	5	6
5	2	6	4	8	7
6	1	4	7	7	- 1
7	3	5	6	0	8
8	4	6	7	0	2
9	6	5	7	- 1	7

结构变换后的数据可以按照 MPEG-4 的要求进行编码。对第一个节点,直接编码运动矢量和灰度值;其它节点,对它们相对于其上一个节点运动矢量和灰度的差值进行编码。重要的是,这种编码顺序实质上是对网格结构进行了编码,在解码端可以根据传输节点的顺序和三角形顺序很容易的恢复 DT 网格,有效地提高了解码效率。

5 实验结果及分析

为了测试算法的性能,本文将它同 H. 263 进行了全面的比较。使用的图像序列是灰度 Miss America 的前 50 帧(其它序列的实验结果基本相同)。实验环境 PIII866 + Win98 + VC6. 0。实验比较了全部由 I 帧组成(前 20 帧)和使用运动估计(前 50 帧)两种情况下在相近压缩率时二者的峰值信噪比(主要反映客观质量)和编码效率(反映算法的复杂程度和实用价值)。表 4 给出了本实验的结果。分析实验结果,得出以下结论:

(1) 本文算法在大致相同的压缩比下,平均 PSNR 一般比 H. 263 高。这反映了基于内容的运动估计和补偿方法有一定优势,它比矩形分割能更合理地跟踪物体运动。

(2) 从图 4 看,传统的 BMA 方法在压缩比较高时“方块效应”特别明显;而本文方法,在较高压缩比时恢复图像的均匀度、自然度好于前者,主观上更容易接受。因为这种基于图像内容的分割形式是根据图像的纹理自适应地划分图像,这样进行运动估计更合理一些。

表 4 实验结果

	本文基于 DT 的 MPEG-4 算法				适于低比特率传输的 H. 263 标准			
	平均 PSNR (dB)	总比特数 (bits)	压缩比	编码效率(总时间/帧数)	平均 PSNR (dB)	总比特数 (bits)	压缩比	编码效率(总时间/帧数)
全部使用 I 帧	36.91	327680	50.5	123.50s/20	35.12	327390	50.7	26.87s/20
	38.31	453280	37	128.80s/20	37.84	426760	38.8	27.03s/20
DT 的	36.12	163840	252	152.92s/50	35.64	165478	250	54.30s/50
	36.58	204800	202	156.30s/50	36.08	207887	199	55.70s/50
运动估计	36.95	245760	168	157.52s/50	36.23	250724	165	55.83s/50
	37.79	368640	112	158.78s/50	36.85	376087	110	56.01s/50
	37.94	409601	101	159.44s/50	37.06	435469	95	56.78s/50

(3) 从算法的效率上看(总时间/帧数),本文算法在不同压缩比下的效率比较接近,说明算法有比较好的稳定性。比较

而言,总体效率无法和 H. 263 相比,说明本文算法的复杂度较 H. 263 更高.但是,从解码效率看(每秒 20~35 帧),这种“非对称”的图像编解码算法完全可以胜任大部分应用场合.



图 4 对 Miss America 第 33 帧进行运动估计结果

6 结束语

本文对 MPEG4 中 DT 网格对象编码算法进行了实现,提出了一种新的适合 MPEG4 标准的运动连续预测方法.同其它标准相比,本文算法具有一定的优越性.但是,算法中用线性插值恢复的图像误差还比较大、运动估计的效率不是很高,这些都严重的影响了整个算法的质量和效率.文献[8]提出了一种网格运动估计的快速算法,改进效果比较明显,运动估计快速算法的研究是本文下一步的工作重点.相信如果在上述两个方面加以改善,这种算法完全可以满足实时编码的要求.

参考文献:

- [1] Brusewitz H. Motion Compensation with Triangles [A]. Proc. 3rd Inter. Conf. on 64 kbit Coding of Moving Video [C]. Rotterdam, Switzerland, 1990.
- [2] Nakaya Y, Harashima H. Motion Compensation based on spatial transformation [J]. IEEE Trans, 1994, CSTV-4(3): 339 - 356.
- [3] 颜尧平, 卢朝阳, 吴成柯. 基于 DT 模型的运动估计和补偿 [J]. 西安电子科技大学报, 1997, 24(3): 357 - 362.
- [4] 吴枫, 高鹏, 高文. 基于模型的编码 [J]. 计算机学报, 1999, 22(12): 1239 - 1245.
- [5] 卢朝阳. 模型优化及其在 DT 模型基图像编码中的应用 [J]. 通信学报, 1997, 18(6): 1 - 7.
- [6] 卢朝阳, 吴成柯, 陆心如. 用 DELAUNAY 三角形化实现的矩形边界表面描述算法 [J]. 计算机学报, 1992, 15(3): 161 - 170.
- [7] ISO/IEC JTC 1/SC 29/WG11 N3056, Maui, December 1999, Information technology - Generic coding of audio-visual objects - Part 2: Visual [S]. ISO/IEC 14496-2 FDAM 1, Final Draft Amendment 1.
- [8] 吴枫, 高鹏, 高文. 基于网格模型的运动估计技术 [J]. 电子学报, 2000, 28(5): 1 - 5.

作者简介:



韩军功 男, 1977 年生于陕西省西安市, 西安电子科技大学通信与信息系统博士研究生, 主要研究领域为图像处理和立体图像编码.



卢朝阳 男, 1963 年 10 月出生于陕西华县, 1990 年获西安电子科技大学通信与电子系统专业博士学位, 现为西安电子科技大学教授、博士生导师, 目前主要研究领域为图像传输与处理、低比特率图像编码、图像识别和图文信息处理等, 近年来发表学术论文 60 余篇.