

# 彩色立体图象编码新方案

周 军 谈 正

(西安交通大学信息工程研究所, 西安 710049)

**摘 要** 真立体显示技术是虚拟现实中的关键技术之一,而如何组织和存储其中大量的彩色立体图对,则显得尤为重要。由于立体图对之间的相关性,传统的压缩算法(JPEG、MPEG等)从速度和效率上不再适应。本文提出了一种结合 JPEG 标准并利用立体图象对相关性的压缩编码新方案,并获得了较高的压缩比及信噪比。实验结果表明了这种新方案的优越性。

**关键词** 虚拟现实,立体显示,图象编码/压缩,JPEG 压缩

## 1 概 述

随着计算机技术和人们感官需求的不断发展,彩色立体图象的显示技术已越来越被人们所需求,并已在虚拟现实技术(Virture Reality)中作为一种最基本的技术被采用。这是因为人们生活的自然环境是一个具有三维纵深的环境,人们日常所接触的、所观察的外部世界都是具有深度感的真立体世界。彩色电视技术的发展已解决了人们的远距离感知和真实感显示,但它固有的致命弱点是其所依赖的技术基础是采用了单镜头摄像后的图象,它所反映的是外部立体世界的投影,即单镜头摄取的图象丧失了深度信息。尽管人们利用想象把电视图象中的物体,通过透视、阴影、遮挡、明暗等关系理解成三维世界,但它不可能提供真正的深度信息,更不可能随人们视点的移动而改变物体的观察印象。

真立体显示技术是相对于传统平面显示技术而言,它依据视觉原理<sup>[1]</sup>,利用 2 幅视图来产生真立体效果。若我们能使人们的左、右眼能分别地观察到具有视差相关性的左右视图,那么就可以在人脑中恢复出真立体世界来。目前流行的立体电视和虚拟现实(VR)技术中的立体显示技术就是基于这一原理。为了实现这一技术,通常需有外部设备的配合,一般采用的有时间并列制和时分串行制 2 种方案。头盔

显示方案是采用左右 2 个平面显示器,左右视图并列地提供给 2 个显示器,这是当前 VR 技术中最普遍采用的手段。时分串行制方案<sup>[2]</sup>是把左右视图在单显示器上交替串行显示,并借助液晶开关眼镜,使左右眼分别观看到对应的左右视图,只要交替和同步速度足够高,其效果和并列式是完全相似的,这种方案已在立体电视技术中被采纳。本技术方案采用的是后者,虽然它对计算机系统配置要求不高,但它比起传统平面显示技术对存储、显示的要求提高了一倍,鉴于真立体显示技术的特点,传统图象压缩算法(如 JPEG, MPEG)已不再适应,有必要提出针对真立体显示的图象压缩算法。

由于左视图象和右视图象之间存在极大的相似性,若用传统的图象压缩算法分别对 2 幅图象进行压缩,就忽略了这种相似性,从而压缩数据还存在很大的冗余。为了提取并利用 2 幅图象的相关性来压缩左、右视图象,本文在时分制立体显示的基础上提出了一种基于 JPEG 标准<sup>[3,4]</sup>的块匹配压缩新方案,即把左右视图用一幅基图外加视差偏移来描述,较标准 JPEG 算法而言,该方案具有较高的压缩比及信噪比。

## 2 立体视觉原理及模型

立体视觉对人来说是极其重要的,没有深度层

次的世界将是无法想象的。当我们用眼睛观察外部世界时,外部世界在两眼视网膜上所形成的图象略微有些不同,物体在这2幅图象中的相对位置差称为视差(disparity),它和距离成反比的关系。我们的大脑能测量这种视差,经过大脑视神经的融合,从而产生深度感觉。由此可知视差对立体视觉的形成起着非常重要的作用。在左右视图象中任选一幅作为基图,而另一幅为匹配图,从而视差的测量包括下面3个步骤:(1)必须从基图中选出位于场景中一个表面上的某一特定位置(点);(2)必须从匹配图中鉴别出同一物体表面的相同位置(点),这点通常被称为“同名点”;(3)必须测量出这2个对应像点(同名点)之间的视差。实际上,对视差的测量是非常复杂的,在文献[1]中,Marr给出3个约束条件:相容性;唯一性;连续性。

但是这2种算法直接用于立体图象压缩是不切合实际的,因此我们简化了视觉模型,采用了文献[2]中的视觉模型——非会聚双目成像模型(见图1),其中, $O, O'$ 分别代表左右焦点,两者间距为 $L$ ,光轴与 $z$ 轴平行,焦距为 $f$ ,左右像平面与 $xy$ 平面平行,假设三维空间任一点 $P(x, y, z)$ ,在左、右像平面分别成像为 $P_l, P_r$ , $P_r$ 是 $P_r$ 在左像平面中的相应位置点,点 $P$ 的视差 $d$ 可用 $|P_r - P_l|$ 计算,则有

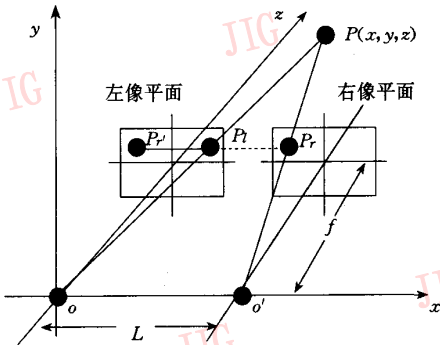


图1 非会聚双目成像模型

Fig. 1 Nonconvergent binocular imagery model

$$(x, y, z) = \frac{L}{d}(x', y', f) \quad (1)$$

其中, $(x', y')$ 是左像平面中 $P_l$ 的平面坐标,所以视差 $d$ 与 $z$ 满足:

$$d = (Lf)/z \quad (2)$$

由上式可以看到,当物体距离眼睛越远,即越大时,视差 $d$ 就越小。当物体距离眼睛太远时,视差变得非常小,就没有深度感了。

### 3 块匹配算法思想及流程图

由非会聚双目成像模型可知,我们可以将左右图象分割成 $8 \times 8$ 的块,然后扫描匹配图的各块,在基图中搜索同名块,判别准则采用相似度准则。定义两个大小为 $N \times N$ 的矩阵 $A, B$ ,其相似度 $S(A, B)$ :

$$S(A, B) = 1 - \frac{\sum_{i=1}^N \sum_{j=1}^N |A(i, j) - B(i, j)|}{\sum_{i=1}^N \sum_{j=1}^N (|A(i, j)| + |B(i, j)|)} \quad (3)$$

其中 $S \in [0, 1]$ ,当 $A$ 和 $B$ 完全相反时 $S=0$ ;当 $A$ 和 $B$ 完全相同时, $S=1$ 。根据相似度准则,我们引入相似度函数 $S(m, n)$ :

$$S(m, n) = 1 - \frac{\sum_{i=1}^N \sum_{j=1}^N |f(i, j) - f(i + m, j + n)|}{\sum_{i=1}^N \sum_{j=1}^N (|f(i, j)| + |f(i + m, j + n)|)} \quad (4)$$

其中, $f(i, j)$ 为图象信号, $0 \leq S(m, n) \leq 1$ ,对于符合图1所示关系的左右视图,其相互间的匹配关系就变得十分简易,因此匹配块的位置只存在水平视差的关系,二维相似度函数就蜕化为一维函数,即变为:

$$S(k) = 1 - \frac{\sum_{i=1}^N \sum_{j=1}^N |f(i, j) - f(i + k, j)|}{\sum_{i=1}^N \sum_{j=1}^N (|f(i, j)| + |f(i + k, j)|)} \quad (5)$$

当 $S(k)$ 达最大值时,则认为当前两块达到匹配,记下当前的 $k$ 值作为两块之间的视差。这样,匹配图完全可以用基图加视差偏移量来描述。

对于彩色图象,RGB三基色之间存在极大的相关性,因此常用的彩色图象表示方式有YUV方式,其中分量Y表示象点的亮度,分量UV表示象点的色度。由于人眼具有一种特性,对象点的亮度分辨力较强,而对象点色度分辨力较弱,因此将亮度分量和色度分量分别处理,便可达到更高的压缩比。在本方案中,我们可以对分量UV进行抽样(2:1或4:1),对抽样图进行块匹配,而对分量Y不抽样,直接进行块匹配。这样,不仅具有很高的压缩比,而且具有较高的信噪比。

在立体图象对匹配过程中,常会由于遮挡、视差等原因而找不到左右视图的同名点(区),为此我们

的方案是:给相似度函数取一门槛,一旦相似度函数最大值小于该门槛,则认为当前块在基图中找不到与之匹配的同名块,必须对该块单独进行编码(我们采用 DCT 编码),同时记视差偏移为 0,并参与视差偏移编码。而匹配块的视差偏移值加 1,以避免与非匹配块混淆。对于视差偏移编码,由 Marr 的第 3 个约束条件—连续性,我们对视差偏移差分值先进行游程编码再将其送入 Huffman(或算术)编码器,这样会进一步提高压缩比。算法框图如图 2 所示。

结果数据见表 1。当  $q=50, S=0.93$  时匹配图(左图)压缩数据量仅为 3 481byte,而信噪比仅下降 2dB;而用 JPEG 压缩,匹配图压缩数据量为 11 229 byte。

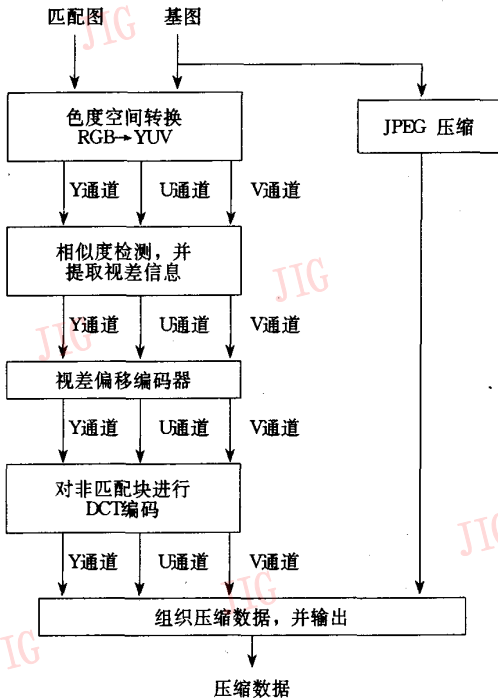


图 2 算法框图

Fig. 2 Block diagram of the method

### 4 实验结果及结论

在前面理论的基础上,我们做了一系列的实验。测试图为  $320 \times 200$  真彩色图象,存储格式为 BMP 格式,数据量为 192 054byte。本文中均以右视图为基图,左视图为匹配图。

$q$  为 JPEG 压缩的质量因子, $Cr$  为压缩比, $S$  为相似度门槛, $SNR$  为信噪比,计算公式为:

$$SNR = 10 \log_{10}(\sigma^2 / \sigma_e^2) \quad (6)$$

其中,  $\sigma^2$  和  $\sigma_e^2$  分别为信号功率和误差功率。

图 3 是利用立体摄像机拍摄的立体图象对,我们将实验结果与标准 JPEG 算法(对 2 幅图分别用标准 JPEG 算法压缩)作了比较(YUV 均未抽样),



(a) 左视图象  
(a) Left-eye image



(b) 右视图象  
(b) Right-eye image

图 3 立体摄像机拍摄的立体图象对

Fig. 3 Stereo image pair token by the stereo camera

表 1 两种算法的实验对比结果( $S=0.93$ )

Table 1 The experimental result of the two algorithms( $S=0.93$ )

	本方案算法		标准 JPEG 算法	
	$Cr$	$SNR(dB)$	$Cr$	$SNR(dB)$
$q=20$ 匹配图	35 : 1	24.41	8 : 1	26.91
基图	8 : 1	27.15	8 : 1	27.15
整体	14 : 1		8 : 1	
$q=20$ 匹配图	55 : 1	21.27	17 : 1	23.84
基图	17 : 1	24.10	17 : 1	24.10
整体	29 : 1		17 : 1	

从上述的实验示例可以证实,本文首次提出的彩色立体图象压缩编码新方案是相当成功的。

(1)在质量因子  $q$  为定值时,与采用标准 JPEG 压缩复原图象相比,在信噪比下降不大(仅下降 2~4dB)的条件下,匹配图压缩比提高了数倍,实现了高效的压缩。

(2)实践表明,立体图对的数据压缩量和匹配相似度门限  $S$  以及质量因子  $q$  的关系比较密切,相似度门限  $S$  越小,匹配图压缩比越大,但信噪比则越小;质量因子  $q$  越大,匹配图压缩比越小,而信噪比越大。

(3)采用本方案 2 幅立体图对的解压速度和用标准 JPEG 解压一幅图的速度相近,因而在解压显示成图这一环节上效率提高了近一倍,这也是本方案的优越处之一。在合适的软硬件支持下,可以保证立体图象复原显示的实时性。

(4)由于立体图对编码过程中涉及到同名块的搜索,因此比标准 JPEG 编码的耗时要大,这一点在

软件算法上还可得到改善。由于块匹配用硬件极易实现,故编码过程可采用硬件电路实现,并且完全可达到实时性。

(5)本方案思想和算法完全可推广到 MPEG 中,因而对动态立体图象序列的压缩极富有吸引力。

### 参考文献

- 1 马尔 D 著. 姚国正等译. 视觉计算理论, 北京: 科学出版社, 1988.
- 2 徐伟忠, 谈正. 三维立体显示系统的开发研究. 第 8 届全国图象图形学术会议论文集, 1996. 368~373.
- 3 ISO/IEC 10918, Digital Compression and Coding of Continuous-tone Still Images; Requirements and Guidelines, 1994.
- 4 马小虎等著. 多媒体数据压缩标准及实现, 北京: 清华大学出版社, 1996.
- 5 高文著. 多媒体数据压缩技术, 北京: 电子工业出版社, 1994.
- 6 李星, 吴元清著. C 语言在多媒体关键技术中的实现与范例. 北京: 学苑出版社, 1993.



周军, 男, 1995 年毕业于西安交通大学信息与通信工程系, 获工学学士学位。现在该校信息工程研究所攻读图象处理专业硕士学位。主要研究方向有: 图象纹理的提取与综合, 多媒体数据压缩, 立体图象对的压缩及立体图象序列的压缩等。

谈正, 男, 教授, 博士生导师, 现为西安交通大学 3DTV-VR 研究中心负责人。主要研究领域为数字图象处理与分析、三维立体显示技术、图象图形融合。近年来在国内外期刊、会议上发表 30 余篇论文, 并编译《模式识别与神经网络》一书(机械工业出版社)。

## The New Scheme for The Colour Stereo Image Compression Coding

Zhou Jun, Tan Zheng

(Information Engineering Institute, Xi'an JiaoTong University, Xi'an 710049)

**Abstract** Stereo display is one of the key technology adopted by VR (Virture Reality), but it is very important how to organize and store a large number of the colour stereo image pairs. Because of the correlation between the pair, the classic coding algorithm (JPEG, MPEG etc.) is not suitable. The paper brings forward the new scheme for the coding, which use the correlation between the stereo image pair and combines JPEG, and achieves the higher compression ratio and the higher SNR. The result shows the advantage of the new scheme.

**Keywords** Virture reality, Stereo display, Image coding/compression, JPEG