

立体元图像阵列的分布式预测编码

吕源治¹, 王世刚¹, 金福寿², 赵岩¹

(1. 吉林大学 通信工程学院, 130012 长春; 2. 吉林大学 计算机科学与技术学院, 130012 长春)

摘要: 为解决组合立体成像系统中立体元图像阵列的存储和传输问题, 采用分布式预测编码的方法对立体元图像阵列进行压缩. 利用立体元图像间的像素关系, 将立体元图像阵列转换成子图像阵列; 根据子图像阵列的规律性变化, 采用最优位移预测结合 PRISM 分布式编码结构对子图像阵列进行分级编码. 实验结果表明, 所提算法的率失真性能明显优于 JPEG 编码标准; 在立体元图像阵列分辨率一定的情况下, 组合立体成像系统的显示分辨率越高, 本方法的压缩性能越好, 对未来高分辨率显示的组合立体成像系统具有更好的适应能力.

关键词: 组合立体成像; 立体元图像阵列; 子图像阵列; 分布式编码

中图分类号: TN919.81 **文献标志码:** A **文章编号:** 0367-6234(2012)11-0087-05

Distributed prediction coding for elemental image array

LU Yuan-zhi¹, WANG Shi-gang¹, JIN Fu-shou², ZHAO Yan¹

(1. School of Communication Engineering, Jilin University, 130012 Changchun, China;
2. School of Computer Science and Technology, Jilin University, 130012 Changchun, China)

Abstract: To solve the problem of storage and transmission of elemental image array in the integral imaging, the distributed prediction coding method is used to compress the elemental image array. First, the elemental image array is transformed to sub-image array according to the pixel relationship between the elemental images, then the hierarchical coding of the sub-image array is performed via the optimal displacement prediction combined with PRISM distributed coding structure in line with the regular changes of the sub-image arrays. The experiment results indicate that the method is obviously prior to the JPEG in the rate distortion property, and when the resolution of the elemental image array is fixed, the higher the display resolution of the integral imaging, the better the compression performance of the method, which will make it have better adaptability in the high display resolution integral imaging in the future.

Key words: integral imaging; elemental image array; sub-image array; distributed coding

组合立体显示技术^[1-2]使显示器在一个广阔的视野空间内呈现出连续视差的全彩色图像, 实现了裸眼真实立体显示, 成为新一代立体显示技术的重要组成部分之一. 图 1 为组合立体成像系统示意图, 组合立体成像系统主要包括采集和显

示两部分, 在采集过程中(如图 1(a)所示), 当实际物体发出的光线通过透镜阵列而被记录在记录媒体上时, 采集设备中便获得了一个立体元图像阵列, 阵列中的每个立体元图像分别记录了被拍摄物体不同位置的图像信息. 在显示过程中(如图 1(b)所示), 将一幅立体元图像阵列放在具有相同参数的透镜阵列前面, 根据光路可逆原理, 透镜阵列将从立体元图像阵列上发出的光线汇聚成空间中真实存在的立体景象.

由于立体元图像阵列所包含的信息量远远大于普通图像, 因此, 给存储和传输带来很大困难. 目前, 已有很多针对立体元图像阵列压缩方面的

收稿日期: 2012-04-06.

基金项目: 国家自然科学基金资助项目(U0935001); 吉林省科技发展计划基础研究项目(20090506); 吉林省科技发展计划项目(20100306).

作者简介: 吕源治(1986—), 男, 博士研究生;
王世刚(1962—), 男, 教授, 博士生导师.

通信作者: 吕源治, lyz123489@163.com.

论文发表. 从是否进行预处理的角度进行分类, 主要包括: 直接对立体元图像阵列进行压缩^[3-7]以及先进行预处理, 再对生成的子图像阵列进行压缩^[8-11], 由于后者较前者能够更加充分的利用像素间的相关性, 因此, 这种预处理结构被广泛应用于各种编码方法中. 然而, 目前比较有效的压缩算法都是将经过预处理后的子图像阵列转换成视频序列, 然后, 按照运动图像的编码方法进行压缩, 这种编码方法与静止图像的压缩算法相比, 虽然提高了压缩效率, 但同时引入了运动估计和运动补偿, 当编码端的运算量或耗能受限时, 这些算法并不适用.

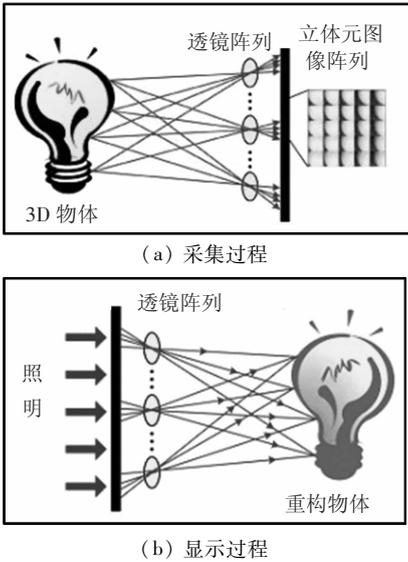


图1 组合立体成像系统

在本文的压缩算法中, 首先将立体元图像阵列预处理成多视点子图像阵列, 然后, 针对子图像阵列的特点, 采用最优位移预测结合 PRISM^[12] 分布式编码结构对子图像阵列进行分级编码. 实验结果表明, 本文所提算法优于 JPEG 编码标准, 同时, 本算法的编码复杂度远远小于运动图像编码, 因此, 特别适用于耗能受限的编码设备.

1 立体元图像阵列变换成子图像阵列

图2所示为一幅立体元图像阵列, 在变换过程中, 将每幅立体元图像同一位置的像素提取出来组成一幅子图像, 将得到的所有子图像按照组成该子图像的像素在立体元图像中的位置排列起来, 即得到子图像阵列^[13].

具体方法为: 定义一个大小为 $M \times N$ 个像素的立体元图像阵列为 $II(m, n)$. 其中, $m = 0, 1, \dots, M-1; n = 0, 1, \dots, N-1$ 为像素点在立体元图像阵列中的位置坐标. 位于第 k 行、第 l 列的一幅立体元图像定义为

$$EI_{k,l}(u, v) = II(k \times U + u, l \times V + v).$$

其中: $u = 0, 1, \dots, U-1; v = 0, 1, \dots, V-1$. 二者分别为像素点在该立体元图像中垂直和水平方向的位置坐标.

子图像阵列中, 第 u 行、第 v 列的一幅子图像定义为

$$SI_{u,v}(k, l) = II(k \times U + u, l \times V + v).$$

其中, $k = 0, 1, \dots, K-1$ 和 $l = 0, 1, \dots, L-1$ 分别为像素点在该子图像中垂直和水平方向的位置坐标. 因此, 在 $U \times V$ 个子图像中, 每幅子图像的分辨率为 $K \times L$.

图3所示为变换后的子图像阵列, 从图中可以看出相邻子图像间具有很强的相关性, 这种相关性成为压缩编码的理论依据.

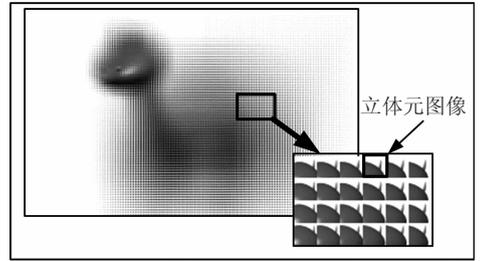


图2 立体元图像阵列

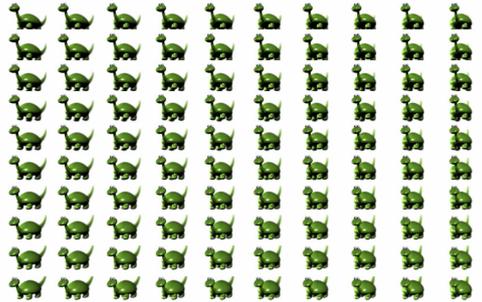


图3 子图像阵列

2 子图像阵列压缩编码

2.1 编码结构

本文提出一种分级编码结构, 如图4所示. 将子图像分为3类: K、C1和C2. 由于子图像阵列中的所有子图像间都具有很强的相关性, 因此, 本文只选择其中一幅子图像作为关键子图像, 该子图像位于子图像阵列的第一行、第一列, 称为K类子图像, K类子图像采用 H. 264 的帧内编码模式进行压缩; C1类子图像为阵列中除K类子图像以外的第一行和第一列子图像, 在编码过程中C1类子图像只使用一个参考图像进行编码; C2类子图像为阵列中除K类和C1类子图像外的其他子图像, C2类子图像在编码过程中使用两个参考图像进行编码.

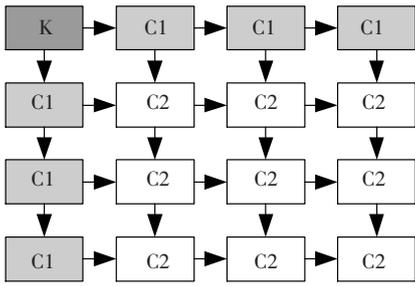


图 4 子图像阵列分级编码结构

在对子图像阵列进行编码时,本文采用从上至下、从左至右的逐行编码模式。首先,对 K 类子图像进行编码;然后,分别利用各自左侧相邻的参考图像,对第一行的 C1 类子图像进行编码,参考图像为已编码子图像在编码端的解码图像;在对第一行以外的其他子图像进行编码时,每行的第一个子图像利用其上方相邻的参考图像进行编码,该行的其他子图像为 C2 类子图像,C2 类子图像在编码时同时使用来自左侧和上方的参考图像。

2.2 最优位移预测

图 5(a)和图 5(b)中分别示出了两幅水平相邻的子图像,其中,左侧为已编码子图像在编码端的解码图像,本文将其作为右侧编码子图像的参考图像,从图 3 和图 5 中可以看出,随着子图像在子图像阵列中位置的右移,拍摄对象的成像位置逐渐左移,同时,成像角度也向左偏转。

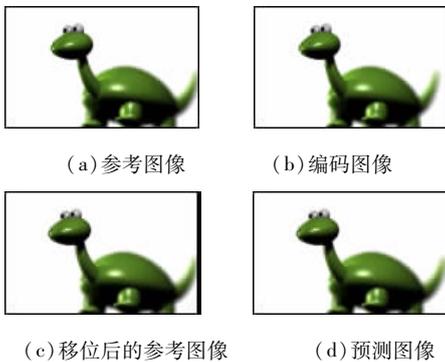


图 5 参考图像、编码图像及预测图像

针对子图像阵列的以上特点,本文提出一种简单有效的最优位移预测方法。水平相邻图像间的预测方法为:首先,将参考图像逐像素的向左平移,分别计算平移后图像与编码图像重合部分的峰值信噪比(PSNR),在完成指定范围的移动后,取 PSNR 最大时对应的移动距离作为最优位移;然后,将参考图像中的像素按照最优位移进行移动,并舍掉超出图像范围的部分,例如,由图 5(a)和图 5(b)计算得到的最优位移为 2,在进行左移操作时,舍掉图像左侧移出的 2 列像素(如图 5(c)所示)。对于移位后图像右侧出现的空隙,采用水平相邻的像素进行填充,生成的预测图像

如图 5(d)所示。比较图 5(b)和图 5(d)可以看出,经过简单的预测后,预测图像与编码图像之间达到了很高的相似度。

垂直相邻图像间进行预测时,先将两幅图像进行转置,然后,按照水平相邻图像间的预测方法进行预测。

C1 和 C2 类子图像均采用上述方法进行预测,不同之处在于,C1 类子图像只使用相邻的一幅(左侧的或上方的)参考图像进行预测;而 C2 类子图像分别使用相邻的两幅(左侧的和上方的)参考图像进行预测,比较生成的两幅预测图像与编码图像的 PSNR,取 PSNR 较大者作为该编码图像的预测图像。

2.3 K、C1 和 C2 类子图像编解码

K 类子图像采用 H. 264 的帧内编解码结构进行编码,为了保证在 C1 和 C2 类子图像的预测过程中,编码端和解码端能够使用相同的参考图像,本文所设计的编码器中都包含一个解码器。

图 6 所示为本文所提的 C1 和 C2 类子图像的编解码流程图。在编码端,首先,根据子图像在图像阵列中的位置判断子图像的类别,并为子图像选择对应的参考图像,然后,采用最优位移预测方法生成编码图像,并将计算得到的最优位移发送到解码端,接着,将编码图像与其对应的预测图像做差得到残差图像,计算残差图像中每个 16×16 块的绝对值的均值 \bar{A} ,计算方法为

$$\bar{A} = \frac{\text{sum}(\text{abs}(b))}{n_b}$$

其中, b 表示残差块中的一个子块, $\text{abs}(\cdot)$ 表示取括号内数值的绝对值, $\text{sum}(\cdot)$ 表示对括号内的所有数值求和, n_b 为 b 中数值的数目。

本文采用文献[10]中提到的分类方法:预先设定两个阈值 t_1 和 t_2 ($t_2 > t_1$),对于残差图像中某个位置的图像块,如果 $\bar{A} < t_1$,则表明编码图像中该位置的图像块与预测图像中相同位置的图像块的差异很小,编码时,对该块不进行编码操作,只发送一个 Skip 标志,解码端识别 Skip 标志后,直接将预测图像中对应位置的块作为解码图像中的块;如果 $\bar{A} > t_2$,则表明编码图像块与预测图像块的差异很大,对该图像块进行与 K 类子图像相同的帧内编码;如果 $t_1 \leq \bar{A} \leq t_2$,则对该图像块进行帧间编码,帧间编码的具体方法将在下节中进行叙述。

在解码端,首先,根据解码子图像在子图像阵列中的位置判断子图像的类别并选择对应的参考图像,然后,根据编码端发送来的最优位移生成预

测图像,最后,结合预测图像对编码码流进行解码,重建出解码后的子图像。

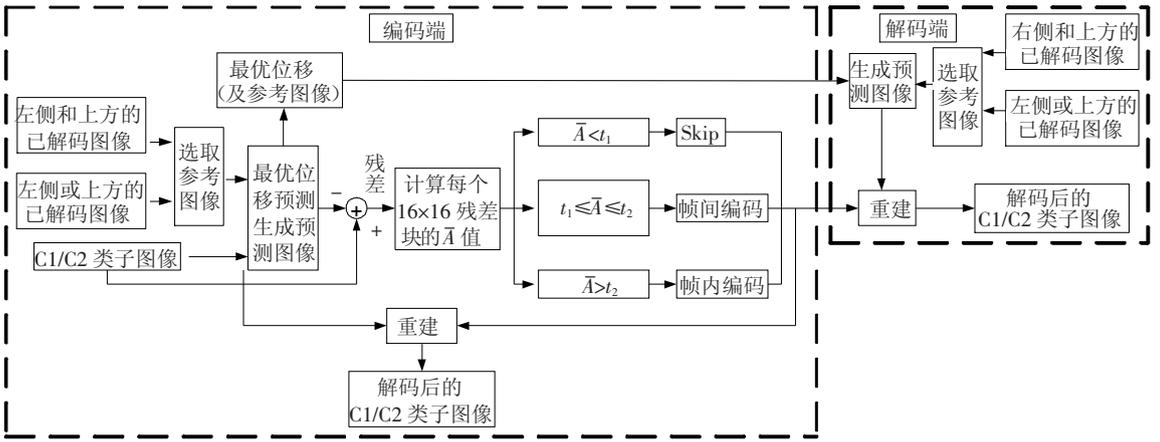


图6 C1和C2类子图像编解码流程图

2.4 帧间编码

首先,将编码图像块以及对应的预测图像块转换成4:2:0的YUV格式,这样,1个16×16的彩色图像块便转换成了4个8×8的亮度子块和2个8×8的色度子块。然后,以8×8的子块作为基本编码单元,对每个编码子块及其对应的预测子块分别进行DCT变换、量化及zigzag扫描,组成两个64×1的一维数组,将两个数组中相同位置上的系数逐对送入联合编码模块中进行编码。

在联合编码中,本文提出 Sign、Run、Depth、Path、Last 的编码结构。对于输入的两个不同系数,Sign 表示这两个系数是否同号;Run 表示该对系数前相同系数的数目,Depth 表示两个系数的绝对值经过二进制转换后,不同二进制位的最高位的位置,例如,设输入的两个系数分别为15和13,其中,15来自于编码块,其二进制表示为1111,13来自于预测块,其二进制表示为1101,则,Depth 的取值为2;Path 表示不同二进制位的具体取值,上例中 Path 的取值为11;Last 用来标记该对系数是否为数组中的最后一对不同系数。

在解码端,首先对预测子块进行与编码端相同的操作,得到 zigzag 扫描后的量化 DCT 系数。然后,根据接收码流,对预测子块的系数进行修正,并进行反 zigzag 扫描、反量化和反 DCT 变换得到解码后的子块。

2.5 码流结构

图7所示为每幅子图像的码流结构。首先,发送最优位移预测方法计算得到的位移,如果该子图像属于C2类子图像,则需要发送所使用的参考图像的标志,结构中为该标志分配1bit的存储空间,其中,0表示左侧图像,1表示上方图像;然后,发送帧间编码中DCT系数的量化步长;最后逐个发送16×16彩色图像块的编码码流。在发送彩色图像块的码流时,首先发送该块的编码类型,结构中为编码类型分配2bit的存储空间,其中:00表示Skip类型,01表示帧内编码,11表示帧间编码。对于帧间编码模式,每个子块的码流结构中需要发送一个Skip标志,即,如果该子块的量化DCT系数与对应的预测子块的量化DCT系数完全相同,则无需进行编码,直接跳过该块。

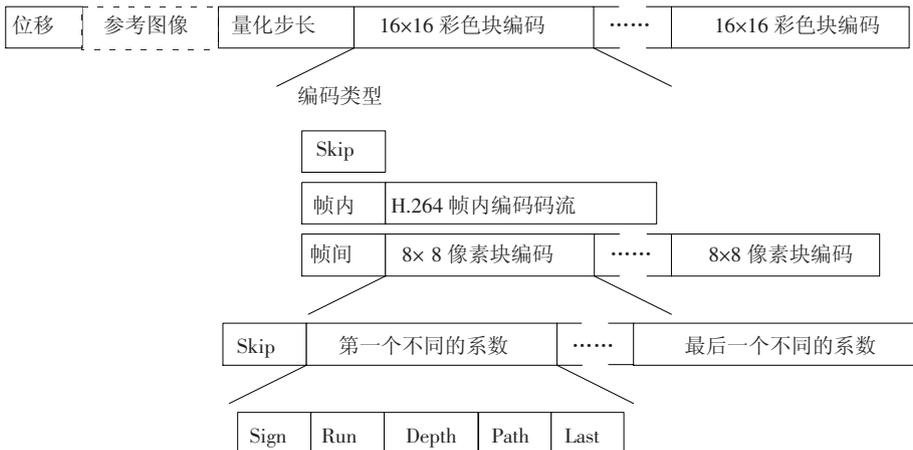


图7 子图像的码流结构

3 实验结果

实验中,分别采用本文所提算法与 JPEG 压缩标准对图 2 进行压缩,图 2 中包含 128×80 个立体元图像,每个立体元图像的分辨率为 10×10 ,两种压缩算法的率失真性能曲线如图 8 所示.从图 8 中可以看出,本文所提算法的率失真性能明显优于 JPEG 压缩标准,其原因在于:首先,本文所提算法对立体元图像阵列进行了预处理,变换后的子图像阵列中相邻像素间的相关性大大增强,在进行离散余弦变换时能够更好地将图像的能量集中在低频部分;其次,在编码过程中,本文设计了一种简单有效的最优位移预测方法,对子图像阵列进行预测编码,这种编码结构将静止图像编码转换成了一种类似于运动图像的编码方式,实现了低复杂度的立体元图像阵列预测编码.

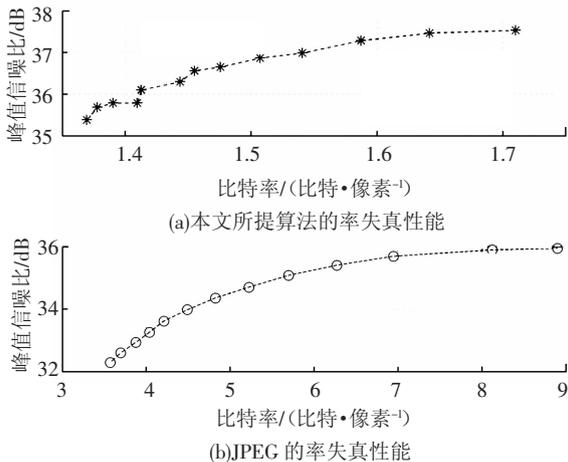


图 8 本文所提算法与 JPEG 标准的率失真性能比较

显示分辨率的不断提高是图像显示技术发展的普遍规律,组合立体成像系统的显示分辨率取决于立体元图像阵列中立体元图像的数量,图 9 所示为一幅与图 2 的拍摄对象相同的立体元图像阵列,不同的是图 9 中包含 160×100 个立体元图像,每个立体元图像的分辨率为 8×8 ,也就是说图 9 的显示分辨率为 160×100 ,而图 2 的显示分辨率为 128×80 .

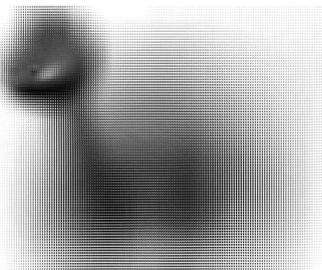


图 9 显示分辨率为 160×100 的立体元图像阵列

实验中,使用本文所提方法对图 9 中的立体元图像阵列进行编码,并与图 8 中的结果进行比

较,比较结果如图 10 所示,从图中可以看出,本文所提算法对高分辨率显示的立体元图像阵列具有更好的率失真性能,其原因在于,随着显示分辨率的提高,每个子图像中拍摄对象的像素分布更加柔和,相邻子图像间的相关性更强,从而提高了预测精度,减少了编码码流.

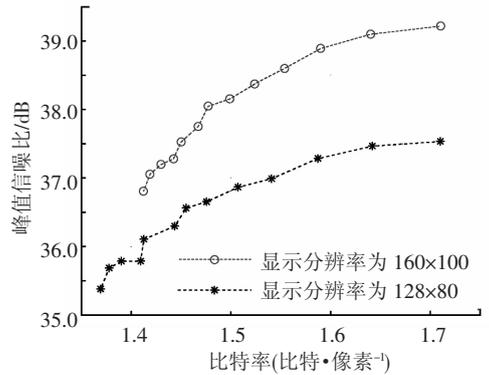


图 10 本文所提算法在不同显示分辨率下的性能比较

4 结论

1) 本文在 PRISM 分布式编码结构的基础上提出了一种适合于立体元图像的编码方法. 首先,对立体元图像阵列进行预处理,将其转换成子图像阵列;然后,根据子图像阵列中每个子图像间的高度相关性,提出一种简单有效的最优位移预测方法,实现了低复杂度的立体元图像预测编码.

2) 仿真结果表明:本文所提算法的编码性能优于 JPEG 标准,同时,随着组合立体成像系统显示分辨率的提高,本方法表现出了更好的率失真性能,对未来组合立体成像系统的发展具有很好的适应能力.

参考文献:

- [1] LVES H E. Optical properties of a lippmann lenticulated sheet[J]. Journal of the Optical Society of America, 1931, 21(3):171 - 176.
- [2] PARK J H, HONG K, LEE B. Recent progress in three-dimensional information processing based on integral imaging[J]. Applied Optics, 2009, 48(34): H77 - H94.
- [3] FORMAN M C, AGGOUN A, MCCORMICK M. A novel coding scheme for full parallax 3D - TV pictures [C]//Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing. Munich, Germany: IEEE, 1997:2945 - 2947.
- [4] AGGOUN A. A 3D DCT compression algorithm for omnidirectional integral images [C]//Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing. Toulouse, France: IEEE, 2006: II - 517 - II - 520.