

基于虚拟视点绘制失真估计的深度图帧内编码

张秋闻^{*①} 安平^{①②} 张艳^{①③} 张兆杨^{①②}

^①(上海大学通信与信息工程学院 上海 200072)

^②(新型显示技术及应用集成教育部重点实验室 上海 200072)

^③(安徽财经大学计算机科学与技术系 蚌埠 230030)

摘要: 该文针对目前压缩深度图时很少考虑直接地减少绘制失真的问题, 探索一种以绘制视点为最小化失真的方法来压缩深度图, 建立由有损深度编码引起深度变化造成的绘制视失真(不是由绘制方法引起)的估计模型。首先推导由深度编码造成每个像素深度变化在绘制上的影响, 然后计入由视频的局部特性对绘制失真的影响因素, 提出一种精确的基于区域特性的绘制失真模型。最后将视点合成失真模型引入到深度编码中率失真模型中来代替深度图像本身失真, 减小合成视失真。仿真结果表明, 所提出的方法可改进绘制虚拟视点的客观质量, 提高深度图的编码效率和减少视点绘制的冗余。可比只考虑深度图失真的基于拉格朗日优化的模式选择方法提高约 2 dB 的 PSNR 增益。

关键词: 深度编码; 虚拟视点绘制; 绘制失真估计; 率失真优化

中图分类号: TN919.8

文献标识码: A

文章编号: 1009-5896(2011)11-2541-06

DOI: 10.3724/SP.J.1146.2011.00218

Depth Map Intra Coding Based on Virtual View Rendering Distortion Estimation

Zhang Qiu-wen^① An Ping^{①②} Zhang Yan^{①③} Zhang Zhao-yang^{①②}

^①(School of Communication and Information Engineering, Shanghai University, Shanghai 200072, China)

^②(Key Laboratory of Advanced Displays and System Application, Shanghai University, Shanghai 200072, China)

^③(Department of Computer Science and Technology, Anhui University of Finance & Economics, Bengbu 230030, China)

Abstract: Since compressing depth map using existing video coding techniques yields unacceptable distortions while rendering virtual views, the depth maps need to be compressed in a way that it minimizes distortions in the rendered views. A distortion model is proposed that approximates rendering distortions caused by depth changes on depth coding. First, relationships between distortions in coded depth map and rendered view are derived. Then, a region based video characteristics distortion model is proposed for precisely estimation distortion in view synthesis. Finally, the new distortion metric is used to select encoding mode decisions for Rate Distortion (RD) optimization, while reducing rendering distortions. Simulation results illustrate that proposed techniques improve the objective quality of the rendered virtual views by up to 2 dB over the Lagrange Optimization based mode selection technique that considers the distortions only in the depth map.

Key words: Depth coding; Virtual view rendering; Rendering distortion estimation; Rate-distortion optimization

1 引言

目前, 立体/多视点视频作为一种能够提供立体感的新型视频技术, 已成为非常重要的研究方向, 3 维电视/自由视点电视(3DTV/FTV)也作为其重要应用而应运而生^[1]。与传统的平面电视相比, FTV 有很多优点, 例如: 方便用户视点交互、选择, 可

以从不同角度观看 3 维立体场景, 满足人们对场景真实和自然再现的需求。然而, 为了满足人们的这种需求, 需要多个摄像机从多个视点拍摄场景和对这些多视点视频信号进行编码传输, 在带宽受限的环境中多视点的存储或传输是不可能的。此种情况的解决方案是在自由立体视频系统的终端采用基于深度图的绘制(Depth Image Based Rendering, DIBR)技术^[2], 它是一种利用彩色图像及其对应的深度图像生成新视点图像的方法。因此, 为实现用户的任意视点选择, 除视频数据外, 还需要获取与之相应的深度数据, 而最终虚拟视点绘制质量主要依

2011-03-14收到, 2011-07-28改回

国家自然科学基金重点项目(60832003), 国家自然科学基金(60902085, 60972137), 上海市科委重点项目(10510500500)和安徽高校省级自然科学基金项目(KJ2011Z008)资助课题

*通信作者: 张秋闻 zhangqwen@yahoo.com.cn

赖于深度图的质量。这就使得高质量深度数据的压缩成为 FTV 系统中的关键技术。

深度图序列可视为一个灰度图像序列,这样就可用传统的视频编码方法直接进行深度图压缩。然而,相比于传统的视频图像,深度图呈现不同的特性。典型的深度图通常由锐边缘和低纹理的匀质区域组成,且深度图边缘与多视点视频中相关图像边缘有高相关性。还应注意的是使用深度图生成虚拟视点代替以真实视点直接地显示时,其深度图中的边缘就十分重要。针对上述深度图特点有效地压缩深度图,近年来已作了各种尝试^[3-6]。然而,这些技术中在压缩深度图时没有一种考虑直接地减少绘制失真问题。深度图并不是一种由终端用户观看的视频,而是为便于在多视点视频中绘制新视点提供帮助。以上的深度编码方法在解码后的深度图中有的会引起赝像,这些赝像将进而造成绘制视点中的失真。所以,探索一种合适的编码深度图方法压缩深度图又不会在绘制视中引入新的失真是至关重要的。

本文针对目前压缩深度图时很少考虑直接减少绘制失真的问题,提出一种不是由绘制引起而是由深度变化引起绘制失真的“失真估计模型(建立深度与视点绘制失真之间的对应关系)”,并把建立的视点绘制失真模型引入到深度编码中,提出一种可使“绘制视”为最小化失真的方法来压缩深度图,以减小视点合成失真。

2 基于深度的 FTV 系统虚拟视点绘制

基于深度图的绘制(DIBR)描述了任意3维空间点在不同成像平面上投影的对应像素的变换关系,利用深度图像,DIBR将参考图像中的像素点映射到目标图像中。

首先,参考视点上像素 (x_r, y_r) 投影到世界坐标系 (u, v, w) 中,

$$[u, v, w]^T = \mathbf{R}_r \cdot \mathbf{A}_r^{-1} \cdot [x_r, y_r, 1]^T Z_r(x_r, y_r) + \mathbf{T}_r \quad (1)$$

其中矩阵 \mathbf{A} 、 \mathbf{R} 和 \mathbf{T} 分别是摄像机的内参矩阵、旋转矩阵和平移矩阵,下标 r 表示为参考视点的摄像机参数, $Z_r(x_r, y_r)$ 是3维空间点在参考视点摄像机坐标系下的深度值。

然后,世界坐标系的点 (u, v, w) 被映射到虚拟视点 (x', y', z') 中,

$$[x', y', z']^T = \mathbf{A}_v \cdot \mathbf{R}_v^{-1} \cdot \{[u, v, w]^T - \mathbf{T}_v\} \quad (2)$$

其中下标 v 表示虚拟视点的摄像机参数,由式(1)和式(2)可推导出

$$[x', y', z']^T = \mathbf{A}_v \cdot \mathbf{R}_v^{-1} \cdot \{\mathbf{R}_r \cdot \mathbf{A}_r^{-1} \cdot [x, y, 1]^T \cdot Z_r(x_r, y_r) + \mathbf{T}_r - \mathbf{T}_v\} \quad (3)$$

这样我们可以得到对应虚拟视点的像素位置 $(x_v, y_v) = (x'/z', y'/z')$,由此,在给定参考和虚拟相机参数和深度的情况下就可以根据式(3)从一幅参考视点图像得到任意视点图像。

由于视点的变化导致场景中景物前后遮挡关系发生了变化,从而导致了在利用 DIBR 方法生成虚拟视点图像时会出现可视性的问题和空洞问题。为了减少遮挡造成的虚拟视点质量下降,MPEG 的 FTV 小组利用多幅参考图像融合的方法来提高绘制质量,即如图 1 所示的虚拟视点图像是由两个邻近视点图像分别生成的目标图像融合而成,通过融合多幅参考图像生成的目标图像能较好地填补目标图像中的较大的空洞。融合的方式如下:

$$I_V(x_V, y_V) = \begin{cases} (1-\alpha)I_L(x_L, y_L) + \alpha I_R(x_R, y_R), & \text{虚拟视点在左右相机均可见} \\ I_L(x_L, y_L), & \text{虚拟视点仅在左相机可见} \\ I_R(x_R, y_R), & \text{虚拟视点仅在右相机可见} \\ 0, & \text{其它} \end{cases} \quad (4)$$

$$\alpha = \frac{|\mathbf{T}_V - \mathbf{T}_L|}{|\mathbf{T}_V - \mathbf{T}_L| + |\mathbf{T}_V - \mathbf{T}_R|} \quad (5)$$

$I_V(x_V, y_V)$ 是虚拟视点图像坐标 (x_V, y_V) 处的像素值。 $I_L(x_L, y_L)$ 、 $I_R(x_R, y_R)$ 分别表示两个邻近相机的左右参考图像, (x_L, y_L) 和 (x_R, y_R) 分别是虚拟视点图像像素 (x_V, y_V) 在由左右参考图像分别生成的目标图像上的对应点, α 是权重因子。 \mathbf{T}_L 、 \mathbf{T}_R 和 \mathbf{T}_V 分别是左参考、右参考和虚拟视点相机外参矩阵的平移向量。

3 基于虚拟视点绘制失真估计的深度图编码

深度图并不是一种由终端用户观看的视频,而是在多视点系统绘制新视点中提供中间数据。采用

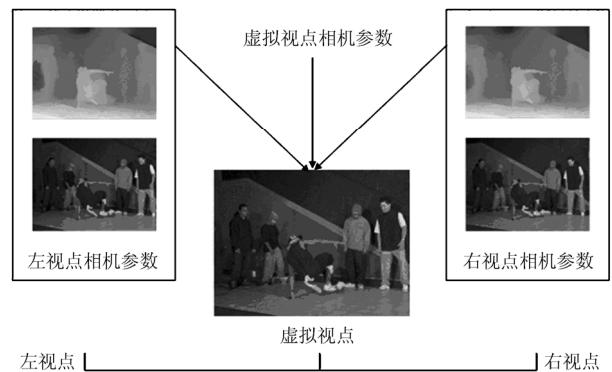


图1 基于多视点视频加深度格式的视点绘制

现有的编码器压缩深度图时会在绘制的新视点中引入失真。所以，用一种适合于绘制新视点的方法来压缩深度图是重要的。通过优化视频编码技术来编码深度图有可能提高绘制视点的质量。基于此，本文拟提出一种不是由绘制引起而是由深度变化引起绘制失真的失真估计模型。首先，推导深度变化与视点绘制几何失真之间的对应关系，然后根据视频的局部特征建立虚拟视点绘制失真模型。

3.1 深度变化引起绘制几何失真

首先，我们推导由深度编码造成每个像素深度变化在绘制上的影响，一个深度图将一幅彩色图像的每个深度信号发给绘制系统。当绘制虚拟视点时，每个像素深度值决定有多少相应的彩色像素要水平地移位，如式(3)所示。如果在深度图压缩期间某些深度像素的亮度值发生变化，就会造成相应彩色像素的移位变化，从而造成绘制视点中的彩色像素的对应是错误的。

可以将来自原始 3D 空间或相机的深度 $Z(x, y)$ 表示为真实深度值(深度图的强度) $v(x, y)$ ：

$$v(x, y) = 255 \cdot \frac{1/Z(x, y) - 1/Z_{\text{far}}}{1/Z_{\text{near}} - 1/Z_{\text{far}}} \quad (6)$$

其中 Z_{near} 和 Z_{far} 是来自摄像机场景中最近和最远截面的深度值，它分别对应深度图 v 中 255 和 0 的深度值。

因此，如果知道每个像素的深度值，我们可以在世界坐标中找到该像素点的位置，这一点可以被重新映射在另一个摄像机坐标系从而完成视点绘制。假定为水平相机阵列，在不同相机坐标系的深度值都约等于世界坐标上的深度值，即在 Z 轴不存在反转因子。这样，式(3)变为

$$\begin{aligned} [x_v, y_v, 1]^T &= \begin{bmatrix} x' & y' & z' \\ z' & z' & z' \end{bmatrix}^T \\ &= \mathbf{A}_v \cdot \mathbf{R}_v^{-1} \cdot \mathbf{R}_r \cdot \mathbf{A}_r^{-1} \cdot [x_r, y_r, 1]^T \\ &\quad + \frac{1}{Z_r(x_r, y_r)} \cdot \mathbf{A}_v \cdot \mathbf{R}_v^{-1} \cdot \{\mathbf{T}_r - \mathbf{T}_v\} \quad (7) \end{aligned}$$

为了考虑深度编码时量化带来的深度图失真，我们近似把深度编码失真表示为增加 $\Delta Z_r(x_r, y_r)$ ，这样编码后的深度图 $\hat{Z}_r(x_r, y_r)$ 可表示为

$$\hat{Z}_r(x_r, y_r) = Z_r(x_r, y_r) + \Delta Z_r(x_r, y_r) \quad (8)$$

通过 3D 变换，失真的深度图进一步造成虚拟视点合成中绘制错误。深度误差 $\Delta Z_r(x_r, y_r)$ 把参考视点中坐标点 p 的映射到虚拟视点中 p' 位置到 \hat{p}' 位置，产生了绘制几何失真 $\Delta p'$ ，即当有深度误差

$\Delta Z_r(x_r, y_r)$ 时，映射到虚拟视点 \hat{p}' 可以表示为

$$\begin{aligned} \hat{p}' &= p' + \Delta p' = [x_v + \Delta x_v, y_v + \Delta y_v, 1]^T \\ &= \left[\frac{x' + \Delta x'}{z' + \Delta z'}, \frac{y' + \Delta y'}{z' + \Delta z'}, 1 \right]^T \quad (9) \end{aligned}$$

综合式(6)-式(9)，可得出

$$\begin{aligned} [\Delta x_v, \Delta y_v, 1]^T &= \frac{\Delta v_r(x_r, y_r)}{255} \cdot \left(\frac{1}{Z_{\text{near}}} - \frac{1}{Z_{\text{far}}} \right) \\ &\quad \cdot \mathbf{A}_v \cdot \mathbf{R}_v^{-1} \cdot \{\mathbf{T}_r - \mathbf{T}_v\} \quad (10) \end{aligned}$$

这表明深度失真 Δv 和合成视点的绘制失真 $\Delta p'$ 呈现线性关系，表示如下：

$$\Delta p'(\Delta x_v, \Delta y_v) = k \cdot \Delta v_r(x_r, y_r) \quad (11)$$

其中 Δx_v 和 Δy_v 分别是绘制图像在点 (x, y) 上的 X 和 Y 方向几何误差， $\Delta v_r(x_r, y_r)$ 是参考视点深度编码后误差，比例因子 k 是由相机参数和深度范围来决定得到的。这里假设是水平平行排列摄像系统，算法生成的视差或相对位移只是水平方向的，因此合成图像在垂直方向没有失真。即 $\Delta y_v = 0$ ，式(11)可以简化为

$$\Delta p'(\Delta x_v) = k \cdot \Delta v_r(x_r, y_r) \quad (12)$$

由于比例因子 k 主要依赖于相机设定和深度范围，相同数量的深度错误会导致不同数量的绘制失真。例如：如果相机之间的距离 $|\mathbf{T}_r - \mathbf{T}_v|$ 变大，或者相机捕捉的是近距物体， $1/Z_{\text{near}}$ 变大，这样随着 k 的变大，绘制失真也变大了。这表明稠密的相机设定和远距物体能减少由深度编码失真造成绘制质量的下降。

其次，绘制几何失真大小将取决于视频的局部特征。例如，在一个视频图像的复杂纹理和对象区域，几何误差造成的失真将是巨大的，因为图像不同位置其像素值变化很大。另一方面，在简单纹理和平滑(均匀)区域，几何误差造成的失真将会很小，因为不同位置的像素值相似。因此，深度变化引起的绘制几何失真需要和视频局部的特征联合考虑。

3.2 基于视频局部特征的区域绘制失真估计

基于深度的绘制，对于深度连续的复杂纹理图像十分敏感的；而对具有弱纹理或深度不连续的图像缺乏敏感性。这表明深度图不同区域的失真对绘制质量有不同的影响，即由深度错误引起的包括几何误差在内的绘制失真可以由视频局部特征来决定。

为了精确地估计经深度编码引起的深度变化所造成的绘制失真，我们将与深度图相关联的纹理视频图像参与到失真估计中。根据纹频视频图像的局部特征把每个参考视频图像分割成若干区域，定义

每个分区的绘制失真函数。采用四叉树分解把参考视频图像分成大小可变的区域,分解后每个区域内的深度值基本相同,这样在该确定区域内所有像素的绘制几何误差几乎不变。因此,每个参考视点视频的区域失真就可使用一个绘制失真函数近似地估计。由于绘制几何误差与运动矢量存在相似性,本文使用文献[7]的频谱失真分析方法来计算每个分区 R_k 的绘制几何失真 E_{R_k} , 它可表达为

$$E_{R_k} = \|\Delta p'\|^2 \times \psi_{R_k} \quad (13)$$

这里 $\Delta p'$ 是深度编码引起的绘制位置错误,可以用节 3.1 节中的方法计算。 ψ_{R_k} 表示可参考视点视频图像每个 R_k 分区上的运动敏感度。

$$\psi_{R_k} = \frac{1}{2(2\pi)^2} \iint_{(-\pi, \pi)} S_x[\omega_1, \omega_2] (\omega_1^2 + \omega_2^2) d\omega_1 d\omega_2 \quad (14)$$

其中 $S_x(\omega_1, \omega_2)$ 为参考视点图像分区 R_k 的能量密度, $[\omega_1, \omega_2]$ 是 2 维频率向量。由于绘制位置失真和深度变化呈现线性关系,则每个分区 R_k 的绘制失真可以由式(14)计算得出。这样当我们编码深度图时,利用由式(13)表示的绘制失真模型就可使失真最小,得出的绘制质量最好。下一节将把所建立的视点绘制失真模型引入到深度编码率失真优化中,来减小视点合成失真。

4 深度图编码的率失真优化

如前所述,以“彩色+深度”格式表示的 3D 视频用于对多视视频生成虚拟视的方法,深度图在终端并不视见。而可利用它生成所需的高质量绘制视点。所以,需要一种对绘制的视点为最小化失真的方法来压缩深度图。以用此种压缩的深度图获得高质量的虚拟视点。本节将上节构建的由深度变化引起绘制失真的估计模型用在帧内编码深度图时的模式选择方法中,对深度编码的率失真进行优化。

在绘制虚拟视点过程中有两种失真:深度编码失真和绘制算法失真。由于只考虑深度编码造成的绘制失真,所以我们使用相同的绘制算法对已编码和未编码的深度图分别进行视点合成,并比较已压缩的深度图和未压缩的深度图所绘制的虚拟视质量,用每个宏块之间误差项平方和(Sum of Squares for Error, SSE)来衡量。

鉴于绘制视点质量和深度编码模式存在相关性,构建涉及此相关性的多视深度编码模式率失真模型如式(15)所示。

$$J = D_v + \lambda R_d \quad (15)$$

J 表示深度编码模式的拉格朗日代价, D_v 为绘制虚拟视质量失真, λ 和 R_d 分别是拉格朗日乘子和深度编码的比特率。在深度编码过程中,由式(15)选择

最佳模式。其中失真 D_v 是在预测的深度像素中可能的深度变化造成的绘制失真的一种度量,所以它不能单是由原始和预测深度像素之间的SSE确定。还需与第 3 节中讨论的失真估计模型与基于块的视频解码器相结合,把式(13)中的由深度变化所引起的绘制失真模型被用来计算式(15)中的 D_v 。这样当编码深度图时就可以块为基础近似得到绘制误差。

在实际的应用中,本文合成一个虚拟视点需要左右两个参考帧(视频和深度图),深度图编码的拉格朗日代价可由深度错误造成的绘制失真与上述左右摄像机的两个深度图的总比特率来计算。

$$J = \begin{cases} (1-\alpha)D_{v,L} + \alpha D_{v,R} + \lambda(R_{d,L} + R_{d,R}), & \text{虚拟视点在左右相机均可见} \\ D_{v,L} + \lambda \cdot R_{d,L}, & \text{虚拟视点仅在左相机可见} \\ D_{v,R} + \lambda \cdot R_{d,R}, & \text{虚拟视点仅在右相机可见} \\ 0, & \text{其它} \end{cases} \quad (16)$$

其中 $D_{v,L}$ 和 $D_{v,R}$ 分别表示左、右摄像机的绘制虚拟视点失真, $R_{d,L}$ 和 $R_{d,R}$ 分别表示了左、右摄像机在编码该模式下深度图所需的比特率, α 为权重因子由式(5)计算。如果假设每个视点的深度图都是独立编码的,则式(16)可以简化为

$$J = \alpha D_v + \lambda R_d \quad (17)$$

将新的绘制失真 D_v 应用到深度编码的每个宏块的帧内预测模式中,也就是说当计算拉格朗日代价 J 时,使用绘制失真估计来减少深度图编码的比特率,提高绘制质量。

5 实验结果

本文提出的方法在 H.264/AVC 参考软件 JM15.1^[8]上实现。使用 JM 帧内编码中的全部模式(Intra16×16, Intra4×4 and Transform8×8)。为了评估所提出算法的性能采用几种测试序列,即微软研究所给出的“Breakdancers”^[9]和 MPEG 提供的多视点测试序列“Newspaper”,“Lovebird1”,“Alt Moabit”^[10, 11],所有测试顺序分辨率为 1024×768。这些序列都有不同的视频局部特性和相机设置,从而可以验证所提出方法的有效性。在本文试验中,“Newspaper”,“Alt Moabit” and “Lovebird1”使用 MPEG 提供的深度估计算法 DERS5.1^[12]得到原始深度图。对于这些测试序列使用 CABAC 熵编码和以 6 个不同的 QP(22, 27, 32, 37, 42 和 47)编码深度图。将本文提出的深度帧内编码所得结果与 JM 帧内编码的块精度帧内模式选择方法进行比较。最后,采用 MPEG 视点合成参考软件 VSRS 3.5^[12]对测试序列进行虚拟视图合成。

5.1 率失真分析

图 2 分别给出了这些测试序列的率失真图，在图中，码率代表用于视点绘制的已压缩的深度图序列的总比特率(例如当绘制“Breakdancers”序列的虚拟视点 1(V1)时所用的码率是所用的视点 0(V0)和视点 2(V2)的深度图序列压缩码率之和)，视点绘制后的平均 PSNR 作为衡量两种深度编码性能的尺度。由图 2 可看出，对这 4 种测试序列的整个比特率范围内，本文提出的算法的 PSNR 明显地高于参考算法 H.264/AVC，平均高 1 dB 以上，在高比特率上增益更高，最高增益为 2 dB。还可看出，编码增益随不同序列而不同，这是因为不同序列的视频特性不同，在不同深度上对象之间的亮度变化均不同。所以，深度变化对“Newspaper”序列的绘制虚拟视点的 PSNR 上有较少的影响。进而我们可以看

到，在第 3 节中给出的失真模型对于例如“Lovebird1”序列中锐边缘深度间断处有高亮度变化的序列有更多影响。

由于优化了深度编码的性能，将其用于绘制合成视时，也就会提高绘制视点质量，这在下面的绘制视点的主观质量结果中可明显地看出。

5.2 主观结果分析

图 3 和图 4 给出分别以“Breakdancers”和“Newspaper”序列的绘制合成视主观质量结果。可以看到本文方法可明显减少模糊和重影现象，特别是在圆圈标识的区域内。其中图 3(b)好于图 3(c)，图 4(b)好于图 4(c)。因此实验结果验证了本文提出方法的有效性。

6 结论

本文提供一种基于虚拟视点绘制失真估计的深

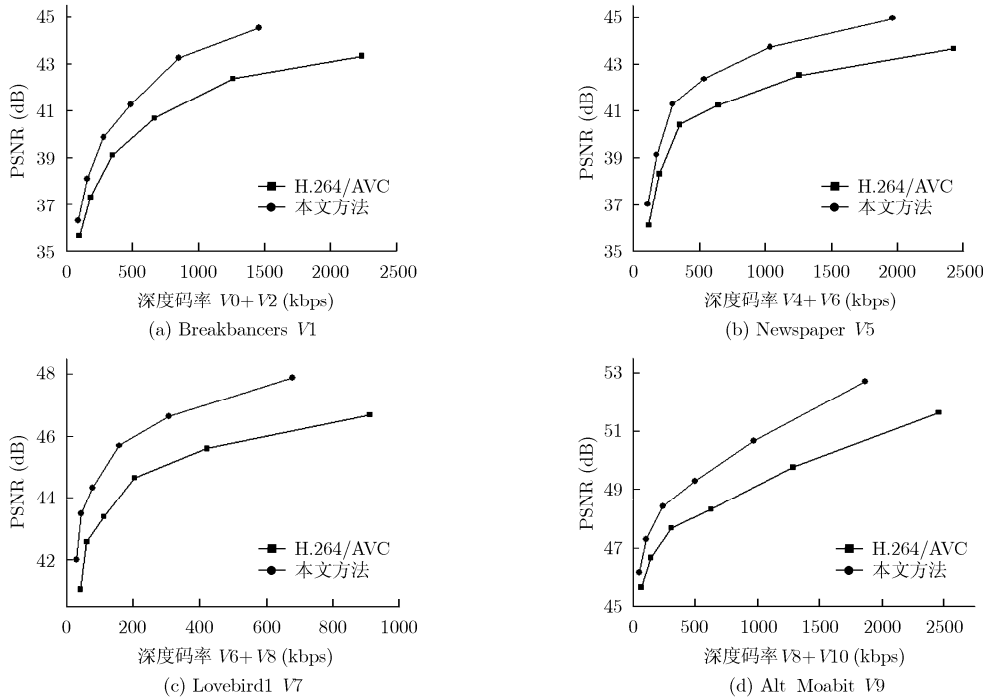


图 2 本文所提出的方法与用 H.264/AVC 的深度编码参考算法的率失真(R-D)性能比较

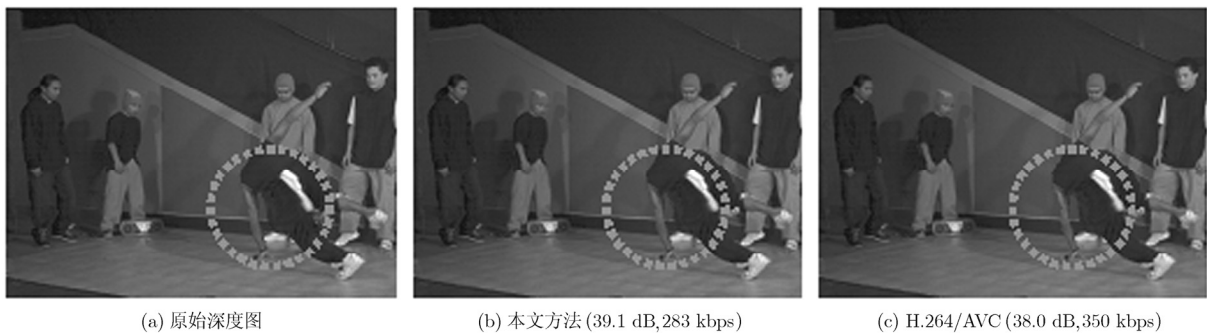


图 3 “Breakdancers”序列视点 0 和视点 2 绘制虚拟视点 1 第 16 帧的主观结果比较



图 4 “Newspaper”序列视点 4 和视点 6 绘制虚拟视点 5 第 12 帧的主观结果比较

度图编码方法, 以绘制视点质量最小化失真为目标来压缩编码深度图, 并建立深度编码引起绘制失真的失真估计模型, 最后将视点绘制失真模型引入到深度编码的率失真模型中, 以代替深度图像本身失真, 从而减小最终的视点绘制失真。实验结果表明, 本文提出的方法比由 H.264/AVC 压缩的深度图所绘制的虚拟视点最高可有 2 dB 的 PSNR 增益。就提出的方法所得到的结果而论, 我们可概括为: 比之深度图自身的最小化像素误差, 以绘制视点质量最小化失真来编码深度图是必须的。这样做, 可明显提高由压缩的深度图所绘制的虚拟视点质量。

参考文献

- [1] Müller K, Merkle P, and Wiegand T. 3-D video representation using depth maps[J]. *Proceedings of the IEEE*, 2011, 99(4): 643-656.
- [2] Fehn C. Depth-Image-Based Rendering (DIBR), compression and transmission for a new approach on 3D-TV [C]. *Proceedings in SPIE Stereoscopic Displays and Virtual Reality Systems XI*, San Jose, CA, USA, 2004: 93-104.
- [3] Merkle P, Morvan Y, Smolic A, *et al.*. The effects of multiview depth video compression on multiview rendering [J]. *Signal Processing: Image Communication*, 2009, 24(1/2): 73-88.
- [4] Maitre M and Do M N. Depth and depth-color coding using shape-adaptive wavelets[J]. *Journal of Visual Communication and Image Representation*, 2010, 21(5-6): 513-522.
- [5] Kamolrat B, Fernando W, Mrak M, *et al.*. 3D motion estimation for depth image coding in 3D video coding[J]. *IEEE Transactions on Consumer Electronics*, 2009, 55(2): 824-830.
- [6] Oh K J, Yea S, Vetro A, *et al.*. Depth reconstruction filter and down/up sampling for depth coding in 3D video[J]. *IEEE Signal Processing Letters*, 2009, 16(9): 747-750.
- [7] Secker A and Taubman D. Highly scalable video compression with scalable motion coding[J]. *IEEE Transactions on Image Processing*, 2004, 13(8): 1029-1041.
- [8] Tourapis A M, Leontaris A, Suehring K, *et al.*. H.264/MPEG-4 AVC reference software manual. Joint Video Team (JVT) of ISO-IEC MPEG & ITU-T VCEG, JVT-AD010, Jan. 2009.
- [9] Zitnick C L, Kang S B, Uyttendaele M, *et al.*. High-quality video view interpolation using a layered representation[J]. *ACM Transactions on Graphics*, 2004, 23(3): 600-608.
- [10] Ho Y S, Lee E K, and Lee C. Multiview video test sequence and camera parameters. ISO/IEC JTC1/SC29/WG11, MPEG 2008/M15419, Archamps, France, April 2008.
- [11] Feldmann I, Müller M, Zilly F. *et al.*. HHI test material for 3D video. ISO/IEC JTC1/SC29/WG11, MPEG/M15413, Archamps, France, April 2008.
- [12] Tanimoto M, Fujii T, Suzuki, *et al.*. Reference softwares for depth estimation and view synthesis. ISO/IEC JTC1/SC29/WG11, MPEG/M15377, Archamps, France, April 2008.

张秋闻: 男, 1982年生, 博士生, 研究方向为多视点视频深度编码及应用。

安平: 女, 1968年生, 教授, 博士生导师, 研究方向为图像及视频信号处理、编码及应用。

张兆杨: 男, 1938年生, 教授, 博士生导师, 研究方向为数字视频处理及多媒体通信。