

doi: 10.3969/j.issn.1007-2861.2011.04.003

# 3DTV 的系统组成及其关键技术和新发展

张兆杨, 沈礼权, 张秋闻, 安平

(上海大学通信与信息工程学院, 上海 200072)

**摘要:** 综述 3DTV 系统的 3 种类型, 包括基于双目的 3DTV、基于多视的 3DTV 和由 2D 视频转换为 3D 视频的 3DTV 系统, 并在概述各自实现方法的基础上, 指出相应的优缺点. 对作为今后发展主流的多视 3DTV 系统, 分析其尚需突破的三方面关键技术, 并进一步介绍有利于解决这些关键技术的以“多视+深度”表示的 3DTV 新模式, 以及课题组对此新模式所取得的研究进展. 最后, 对 3DTV 系统的发展和应用前景进行展望.

**关键词:** 3DTV; 深度; 视频压缩

中图分类号: TN 919.81

文献标志码: A

文章编号: 1007-2861(2011)04-0353-07

## Development and Key Technology of 3DTV System

ZHANG Zhao-yang, SHEN Li-quan, ZHANG Qiu-wen, AN Ping

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

**Abstract:** This paper introduces three types of 3DTV systems including two-view 3DTV, multi-view and 2D-3D systems. Based on the description of key components, the merits and drawbacks of each system are analyzed. A new method, i. e., view-plus-depth, to represent 3D scenes, which can effectively reduce 3DTV data for transmission is described. Finally, the procedure of 3DTV development is presented and the prospect of home-oriented 3DTV application is given.

**Key words:** 3DTV; depth; video compression

目前, 二维(2D)电视由于丢失了自然场景中作为第三维的深度信息, 使人们在视觉上缺乏立体(深度)感觉, 从而与实际的自然场景不一致(失真). 人们早就盼望可恢复深度信息的三维电视(3DTV)的出现, 但由于受技术发展的限制, 直到 21 世纪, 尤其是近几年来, 相关技术才得到快速发展. 然而, 欲达到观众能真正接受的程度, 尚需解决众多的关键技术问题. 本课题组在概述 3DTV 的基本原理、系统组成和现有应用的基础上, 对正在发展中的基于“单视+

深度(single-view plus depth, SVD)”和“多视+深度(multi-view video plus depth, MVD)”新模式的 3DTV 系统研究中尚待解决的核心技术作了较深入的讨论, 并介绍了课题组已取得的部分研究成果.

### 1 不同类型 3DTV 系统的组成、关键技术及其应用

3DTV 系统由 3D 内容获取、预处理、编码、传输、解码/视合成和显示共 6 部分组成, 如图 1 所示<sup>[1-2]</sup>.

收稿日期: 2011-06-23

基金项目: 国家自然科学基金重点资助项目(60832003); 上海市科委重点资助项目(10510500500); 上海市科委世博专项基金资助项目(10DZ0580800)

通信作者: 张兆杨(1938~), 男, 教授, 博士生导师, 研究方向为图像、视频压缩编码等. E-mail: zhyzhang@shu.edu.cn

在不同的应用环境中,传输方式也不相同,在此不作讨论.图1中包括了下述3种不同类型的3DTV系统.

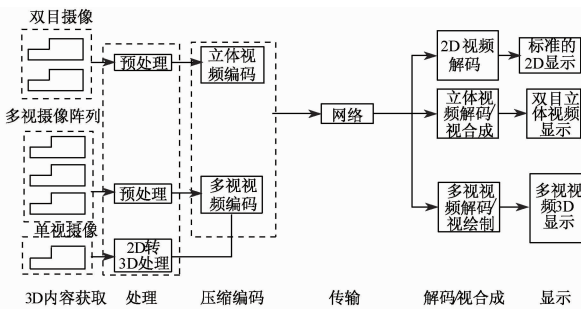


图1 3种类型3DTV系统组成

Fig. 1 Three types of 3DTV systems

### 1.1 双目(双视)3DTV系统

双目3DTV系统通常由间距为6.5 cm的2个完全相同的摄像机拍摄场景.由于2个摄像机的拍摄视角存在少许差异,因此,需对2路视频图像进行预处理校正,在几何位置、亮度以及色亮度等达到一致后,再压缩编码为视频流;然后,经传输并在解码端复原成2路视频,且合成有不同视差的“立体对”视;最后,输入立体图像显示器.为了使现有的2D电视接收机也能接收3D节目,可将解码后的1路视频送入2D显示器显示平面图像.双目3DTV系统在技术实现上相对简单,且比较成熟.目前,国外已开播的和国内不久将试播的3D频道均采用此种方式,但观看时需佩戴特殊眼镜.

### 1.2 多视3DTV系统

进入21世纪后,MPEG 3DTV工作组开始倡议发展多视3DTV技术.多视3DTV系统通常由 $N$ 个( $N$ 取8~16)相同类型的摄像机同时、同步地拍摄同一场景,所得的多个视频经几何参数校正、亮度/色度补偿等预处理后,再经多视视频编码(multi-view video coding, MVC)压缩成视频流.由于编码后重构的多个视不一定符合某个观众在显示屏前所在位置的观看要求,故解码后需通过选择一些合适的视绘制成符合需要的2个虚拟视(虚拟摄像机)后,再合成为“立体对”视.因此,这样的3DTV也被称为自由视点电视(free viewpoint television, FTV).多视3DTV系统的优点包括:在显示屏前,观众可看到立体效果的视角范围(简称为立体视角)远大于双目3DTV系统;便于发展出不用佩戴特殊眼镜而使用人眼便可直接观看的多用户瞳孔跟踪技术,使观众在屏前漫步时所看到的立体图像能够随位置的移动而

变化.如图2所示, $N$ 个观众可同时观看立体图像,并且同一观众从观众1位置移动到观众 $N$ 位置,所看到的物体的侧面是不同的,从而使观众产生看到实际景物的临场感.但是,多视3DTV系统的复杂性和实现上的困难也显著增加,主要需要突破的关键技术有以下三方面:①使 $N$ 个摄像机精确地、同时同步地变光圈、调焦、变焦和移位,这在技术和工艺实现上甚是困难,有待发展一种自动、精湛的多视深度/视差获取系统<sup>[3]</sup>;② $N$ 个视在编码后的数据量是单个视经H.264/AVC编码数据量的 $N$ 倍,这给传输和存储带来巨大的负担,而对于经MPEG-JCT的基于Hierarchical B结构的MVC压缩编码<sup>[4]</sup>,其编码效率仅比单个视的H.264/AVC提高2.5 dB的峰值信噪比(peak signal to noise ratio, PSNR)增益,而且运算量较大,即使本课题组在国家自然科学基金重点资助项目的研究中采用了自适应选择视差估计/块尺寸模式和利用宏块运动特性的快速算法<sup>[5-6]</sup>,使运算量平均降低了63%,但离实时编码仍有很大距离,尚需发展一种可进一步降低传输带宽的压缩编码方案;③多视3DTV系统由于拍摄场景的摄像机数量过多,使节目制作费用远高于二维的高清晰度电视(high definition television, HDTV).为此,需发展一种即使摄像机数量明显减少,也能保持上述多视3DTV系统优点的改进系统.

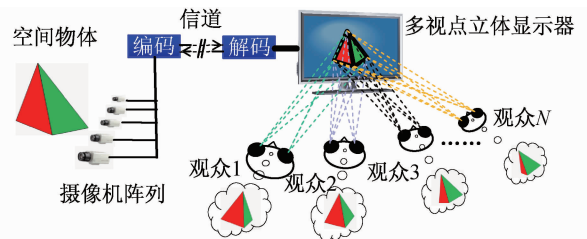


图2 多视3DTV系统立体视角宽和观看临场感的关系

Fig. 2 Relation between view angle distribution and stereoscopic feeling in multi-view 3DTV system

### 1.3 二维视频转换为三维视频的3DTV系统

如图1的左下端所示,2D视频转换为3D视频的3DTV系统是由单个高清晰度摄像机拍摄二维彩色视频后,通过2D视频转换为3D视频(简称2D转3D)的处理来得到三维视频信号(其后的流程与多视3DTV系统相同),这样大大简化了拍摄系统.虽然其立体效果比多摄像机拍摄的多视3DTV系统稍逊,但可使3D节目制作的费用显著降低.因3D节目拍摄并不像2D节目那样较单一,其立体效果随场

景的远近和类型而不同,需要交叉拍摄(对双视3DTV系统也是如此,故2010年南非世界杯现场转播时也采用专用的远近景3D拍摄机和2D转3D交叉使用).此外,现有的2D高清片源在2D转3D后可作为3D片源由DVD/机顶盒播放.

2D转3D的核心技术主要由利用边缘信息将块分组为区域、用梯度先验假设为各个区域分配深度、用双边滤波抑制块效应和基于深度图像绘制(depth image based rendering, DIBR)的多视绘制几部分组成<sup>[7]</sup>.本课题组进一步将2D转3D图像/视频转换与动画设计、3D特技效果和快速虚拟视点绘制相结合,提出了图3所示的3D视频节目制作方案.在上海市科委世博专项基金的资助下,课题组研制成4套2.616 m(103英寸)和1套2.159 m(85英寸)/2套1.168 m(46英寸)的多视3DTV系统,并已分别在中国国家馆的“低碳行动馆”和“东方足迹馆”中展出.由于实现了不戴立体眼镜直接用肉眼即可看到栩栩如生的3D动态视频,深受观众欢迎.

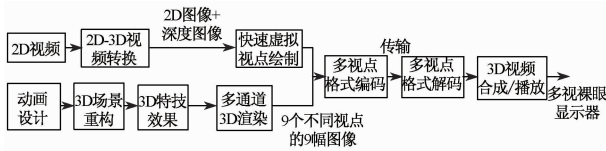


图3 上海世博会中国馆3D视频节目制作方案

Fig.3 Scheme of 3D video program generation in China Pavilion of Shanghai Expo

## 2 正在发展中的3DTV新模式

下面介绍1.2节中所提及的一种既可降低摄像机数量又能发挥多视3DTV系统优点的改进系统,并将该系统称为“多视视频+深度(MVD)”3DTV模式,其关键技术是从视频中获取深度信息<sup>[8]</sup>.目前,该模式正处在MPEG标准化过程中.除了MVD外,还有“单视+深度(SVD)”的双目3DTV模式.下面分述这2种模式,并扼要介绍本课题组在其关键技术研究上取得的部分研究成果.

### 2.1 SVD表示的3DTV系统

传统的双目3DTV系统中,2个视点的视频都需要传输,而SVD则只需传输其中1个视点的视频和以另1个视点为主的场景深度信息.深度信息是由视频图像每个像素的深度层所组成的灰度图像,称为深度图(depth map, DMP).深度值的设定与2个主要深度层限幅平面相关,将离摄像机最近的深度

层限幅平面 $Z_{near}$ (灰度层255)定义为在特定深度图中表示的最小度量深度值;将离摄像机最远的限幅平面 $Z_{far}$ (灰度层0)定义为最大度量深度值.在均匀量化下,所有 $N=2^8$ 的离散深度层均处在这2个极端值之间,由下式计算:

$$Z_v = Z_{near} \left( \frac{v}{N-1} \right) + Z_{far} \left( 1 - \frac{v}{N-1} \right), \quad (1)$$

式中, $v \in [0, 1, \dots, N-1]$ .

由于深度图在压缩编码后的比特率仅为压缩1路彩色视频比特率的15%~20%,因此,可大幅提高双目视频的编码效率.但原来的2个视在省略掉1个视后,会发生对象的某个部分原来在1个视中可见而在另1个视中不可见的“遮挡”现象,从而导致深度的不连续.而由于深度的不连续,物体边缘部分经编码后在视合成的三维映射过程中会使合成视图像形成大的空洞.为此,我们结合上海市科委重点资助项目的研究,提出如图4所示的解决方案,其中在深度估计后先对获得的深度图进行如下预处理:  
①通过边缘检测,将检测到的对象前景边缘深度向外扩展几个像素,以形成有利于图像修复<sup>[9]</sup>的带状区域,并大致保留对象边缘;  
②对扩展后的边缘进行平滑;  
③为防止垂直线条由于深度值的不同而在图像映射过程中发生扭曲,采用方向算子提取原图像的垂直边缘,进而对该方向上的深度值进行校正,以确保垂直方向上深度值的一致.

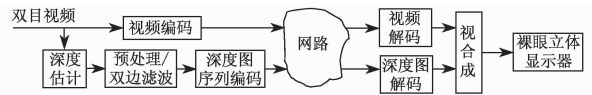


图4 基于深度的双目视频格式3DTV系统

Fig.4 Two view 3DTV system with “depth + video”

经上述预处理后,引入双边滤波器<sup>[9]</sup>消除深度图尚残存的错误.双边滤波器对深度图的像素值使用空间滤波和边界滤波的双重算子,其输出为<sup>[10]</sup>

$$B[I]_p = \frac{1}{W_p} \sum_{q \in S} f_{\sigma_s}(\|p - q\|) g_{\sigma_r}(|I_p - I_q|) I_q, \quad (2)$$

式中, $W_p$ 为归一化因子,且

$$W_p = \sum_{q \in S} f_{\sigma_s}(\|p - q\|) g_{\sigma_r}(|I_p - I_q|), \quad (3)$$

式中, $p$ 为输入图像 $I$ 上的任意像素, $S$ 为像素点 $p$ 的某个邻域, $q$ 为该邻域 $S$ 中的像素点(像素值为 $I_q$ ), $f_{\sigma_s}$ 为空间域影响函数, $I_p$ 为图像在点 $p$ 处的灰度值.由式(2)可以看出,双边滤波器的加权系数是空间邻近度因子 $\sigma_s$ 与灰度相似度因子 $\sigma_r$ 的乘积,前者随像素点与中心点之间的欧几里德距离的增大

而减小,后者随两像素灰度值之差的增大而减小.在图像变化平缓的区域,邻域内像素的灰度值相差不大,双边滤波便转化为高斯低通滤波;在图像变化剧烈的区域,滤波器利用边缘点附近那些灰度值相近像素点的平均灰度值来代替原灰度值.因此,经预处理和双边滤波后,既平滑了深度图像,又保持了物体边缘处的深度连续性,从而在如图4所示的系统中,能够在编解码后视合成的三维映射中抑制可能形成的空洞.

## 2.2 MVD 表示的 3DTV 系统

MVD 是在 2.1 节所述的 SVD 上发展而来的,即将图 4 所示 SVD 中的输入、深度估计和视频编码分别扩展为多视视频、多视深度估计和多视视频编码,如图 5 所示.

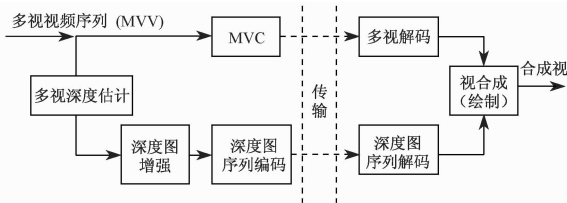


图 5 MVD 结构

Fig. 5 Structure of MVD

MVD 的最大优点是能够显著减少拍摄摄像机的数量.如图 6 所示,重建的 9 个视点 ( $S_1 \sim S_9$ ) 中,只有 3 个视点 ( $S_1, S_5$  和  $S_9$ ) 的视频内容由 3 个摄像机拍摄获得,而  $S_2 \sim S_4$  和  $S_6 \sim S_8$  是由  $S_1, S_5, S_9$  的视频和深度基于 DIBR 生成的,因此,可简化多视摄像机装置,并大幅降低编码传输的数据量.由于 MVD 极有可能成为新的 3DTV 格式的标准<sup>[11]</sup>,因此,已成为当前国际前沿的研究热点.

### 2.2.1 现有的多视深度估计方法

针对图 5 所示系统中的主要组成模块,国内外学者正在进行深入研究.对于多视深度估计模块,由于 MPEG 的深度估计软件 DERS<sup>[13]</sup>所提取出的深度图,在对象边缘和低纹理区域存在欠准确和模糊等缺陷,文献[12]在 DERS 中引入了 4 种通过辅助人工交互的视差图、边缘图、静态图和深度图对其改进,但该方法增加了运算时间.对于图 5 中的深度图增强模块,文献[14]对所提取的多视深度图序列采用后处理自适应增强方法,包括对深度边缘的中值滤波和运动的自适应处理,从而增强空间、时间和视间的深度一致性,使得在保证合成视绘制质量的同时,

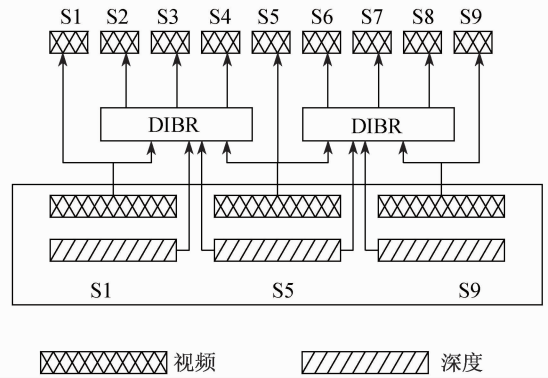


图 6 MVD 格式优于多视 3DTV 格式的原理

Fig. 6 Advantage of MVD compared with multi-view 3DTV system

时,降低了总的计算复杂度.对于深度图序列编码模块,文献[15]针对深度图编码可能会在锐边缘处造成斑点、振铃、块效应等赝像的情况,利用深度图和视图之间的几何相似性和光电度量相似性设计了三重滤波(trilateral filtering)器来取代传统的环状去块滤波器.三重滤波器由检测邻近图像位置的空间滤波、利用深度取样间相似性的范围滤波和利用深度和视频之间配对相似性的界限滤波三部分组成.在编码环路中设置三重滤波器,可有效滤除编码赝像,将此解码后的深度图用于绘制合成视时,可提高合成视的质量.

本课题组结合国家自然科学基金重点项目对图 5 所示的系统进行了综合研究.由于该 MVD 的最终目的是使绘制视合成质量尽可能逼真地重现所拍摄的 3D 场景,因此,在设计深度估计方法和深度编码时应使绘制视的失真最小.首先,通过改进多视深度估计的方法得到更准确的深度图;然后,建立视绘制中几何误差与深度图失真之间的关系,计算几何误差对每种区域所引入的失真,并在编码深度图时将此种会使合成视失真的因素计入率失真优化模型中,从而既可提高深度编码的效率,又可减少绘制误差<sup>[17]</sup>.具体过程如下所述.

### 2.2.2 改进的多视深度估计方法

对由 MPEG 给出的多视深度估计参考软件 DERS-5.1 算法所得的深度图进行增强处理,如图 7 所示,可分为以下 3 个步骤.

(1) 深度一致性检测.由从左视到右视和从右视到左视 2 个方向的交叉检测来检验已得到的像素深度估计值是否可靠.如果 2 个方向的检测结果不一致,则认为是不可靠的深度估计值.由下式计算其一致性:

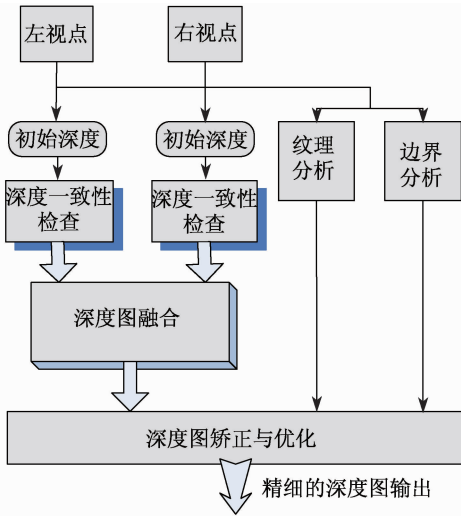


图7 深度增强处理流程图

Fig.7 Flowchart of depth enhancement

$$d(x,y) = \begin{cases} \min(d_{LR}(x_L,y), d_{RL}(x_R,y)), & |d_{LR}(x_L,y) - d_{RL}(x_R,y)| > \lambda, \\ \frac{d_{LR}(x_L,y) + d_{RL}(x_R,y)}{2}, & |d_{LR}(x_L,y) - d_{RL}(x_R,y)| \leq \lambda, \end{cases} \quad (4)$$

式中,  $x_R = x_L + d_{LR}(x_L, y)$ ,  $d_{LR}$  和  $d_{RL}$  分别表示位置  $(x, y)$  上从左到右和从右到左的视差,  $\lambda$  为预定义可变阈值. 因此, 通过左匹配误差和右匹配误差之间的差值量即可判断是否存在遮挡或伪匹配. 于是, 根据式(4), 可选择不同的阈值  $\lambda$  自适应地改变来自遮挡或伪匹配误差的检测.

(2) 多深度图融合. 在 MVD 中, 通过对多个视进行深度估计, 可将多个 (设为  $M$ ) 深度图用高斯混合模型归并为 1 个深度图, 以改进深度图的质量. 设像素块的深度为  $d_n$ ,  $C_k$  为第  $k$  个深度估计的加权参数, 则高斯混合模型可表示为

$$p(d_n) = \sum_{k=1}^M \frac{C_k}{C_0} W(d_n - d_n^k), \quad (5)$$

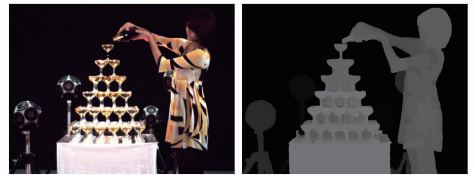
式中,

$$C_0 = \sum_{i=1}^M C_i, W(x) = 1/\sqrt{2\pi} e^{-x^2/2}. \quad (6)$$

式(5)中的  $C_k$  可通过给出的 2 个图像块相似性度量的结构相似性 (structural similarity, SSIM) 计算. 为获得目标块的最佳深度, 我们用最大似然准则得到最优深度值.

(3) 深度图优化. 由于深度图中对象边界的遮挡区等不连续区域在虚拟视绘制中会引起膺像, 因

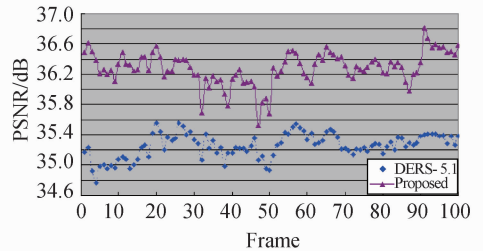
此, 需对融合后的深度图进一步优化. 由图 8 可以看出, 视频帧图像与相应的深度图在边缘之间存在几何相似性<sup>[18]</sup>, 为此, 我们利用此种相似性并结合 2.2.1 节中所述的三重滤波器对融合后的深度图进行优化 (优化过程参见文献[17]). 图 9 为采用该方法和 DERS-5.1 算法所得的深度图, 在进行视合成后结果的质量方面, 前者比后者在 PSNR 上有明显提高, 尤其对于 Book arrival 测试序列而言更加明显.



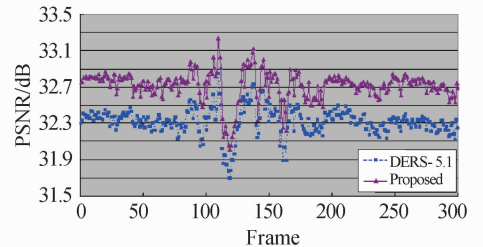
(a) 视频帧图像 (b) 相应深度图

图8 视频帧图像和相应的深度图

Fig.8 Video frame and corresponding depth map



(a) Book arrival



(b) Champagne tower

图9 虚拟视合成结果的客观质量比较

Fig.9 Objective quality comparison of synthesis views

### 2.2.3 计入视合成失真因素的深度编码率失真优化

由于图 5 所示的 MVD 最终是将经编解码的多视深度图和多视视频绘制成合成视来供显示器显示 3D 图像. 因此, 即使由深度估计所得的深度图与 3D 场景的深度十分逼近, 经压缩编码后仍会引入新的失真, 从而使绘制的合成视质量受损. 所以, 编码深度信号时不仅要考虑压缩编码效率, 而且应计入使合成视失真的因素. 下面将对该问题进行具体阐述.

(1) 深度图失真与绘制中几何误差的关系. 深

度图失真  $\Delta v$  与绘制视中的绘制位置误差  $\Delta P'$  之间有如下线性关系:

$$\Delta P'(\Delta x_v, \Delta y_v) = k \cdot \Delta v_p(x_r, y_r), \quad (7)$$

式中,  $\Delta x_v$  和  $\Delta y_v$  分别为水平误差和垂直误差,  $\Delta v_p(x_r, y_r)$  为在参考摄像机位置  $P$  上的深度图失真,  $k$  为由摄像机参数和深度范围确定的尺度因子.

(2) 利用区域视频图像估计视合成几何误差. 在基于 DIBR 的视合成绘制中, 复杂纹理图像区域、对象边界区域、低纹理区域和深度不连续区域对视合成质量的敏感性各不相同, 故在视合成失真中由深度误差造成的几何误差取决于图像的局部特性. 例如, 低纹理区域的几何误差对合成视的影响很小, 而对于对象边界区域和复杂纹理区域, 即使在几何位置上仅有少许改变, 也会对合成视有明显的影响. 因此, 我们利用与深度图相对应的参考视频图像将其划分为不同的区域  $R_k$ , 几何误差引入的每个  $R_k$  的失真  $E_{R_k}$  可表示为

$$E_{R_k} = \|\Delta P'\|^2 \times \psi_{R_k}, \quad (8)$$

式中,  $\Delta P'$  为由式(7)计算得到的因深度量化失真所造成的绘制几何位置误差,  $\psi_{R_k}$  表示在参考视频图像区域  $R_k$  中的运动敏感度因子, 其计算式参见文献[19]. 深度压缩失真即为  $R_1, R_2, \dots, R_n$  所有区域上的失真之和, 即有

$$E_v = \sum_{k=1}^n \|\Delta P'\|^2 \times \psi_{R_k}. \quad (9)$$

(3) 深度图编码的率失真优化. 绘制视的失真是由深度图编码失真和绘制过程自身的失真造成的, 通常后者造成的失真远小于前者. 因此, 假设失真是由前者造成的, 这样就可由压缩编码前的深度图与编码后的深度图分别绘制的合成视中, 取二者每个宏块之间的误差平方和 (sum of squares of error, SSE) 来度量深度图编码在虚拟视合成中所引起的失真  $D_v$ , 并将  $D_v$  计入对深度图编码的率失真优化中, 如下式所示:

$$J = D_v + \lambda \cdot R_d, \quad (10)$$

式中,  $R_d$  为编码深度图的率,  $\lambda$  为拉格朗日乘子. 由此, 以最小拉格朗日代价  $J$  选择编码模式, 即可获得计入绘制视失真的最优深度编码.  $D_v$  的计算如下:

$$D_v = \text{SSE}_v = \sum_x \sum_y (V(x_v, y_v) - V(\hat{x}_v, \hat{y}_v))^2, \quad (11)$$

式中,  $(x, y)$  为块中的像素位置,  $V(x_v, y_v)$  和  $V(\hat{x}_v, \hat{y}_v)$  分别为对深度图未进行编码和已进行编码情况下的绘制. 于是, 由式(9)和(11), 可得

$$D_v = \sum_x \sum_y (E_v(x, y))^2. \quad (12)$$

经上述的率失真优化, 所得到的深度图的压缩效率与用传统的基于 H. 264/AVC 方法来编码的深度图相比, 在所有比特率上具有平均 1 dB 的 PSNR 增益<sup>[17]</sup>. 因而, 此方法既可提高深度编码压缩效率, 又由于计入了编码对合成视的失真因素, 从而有助于提高绘制合成视的质量.

### 3 结束语

我们在介绍双目 3DTV、多视 3DTV 和 2D 视频转换 3D 视频 3 种 3DTV 系统的组成和实现原理的基础上, 重点讨论了其中作为当前发展主流的多视 3DTV 系统研究, 以及其中一些尚待突破的关键技术和有助于解决这些关键技术的“视频 + 深度”新模式, 并对新模式的核心技术进行了比较深入的研究, 取得了良好的成果. 这些成果所运用的方法中, 有的已作为国际标准化提案递交给标准化机构 ISO/IEC/JTC/MPEG.

应指出的是: 正如 2D 电视经历了黑白电视/彩色电视/数字电视/高清电视这 4 个从初级到高级的发展阶段一样, 3DTV 也将会经历双目戴立体眼镜观看/双目裸眼观看/多视裸眼观看/多视加深度裸眼观看等阶段. 目前国外已开播、国内即将试播的双目戴立体眼镜的 3DTV 尚是初级阶段, 预计发展到高级阶段还需经过约 10~20 年的历程.

### 参考文献:

- [1] 张兆杨, 安平, 张之江, 等. 二维和三维视频处理及立体显示技术[M]. 北京: 科学出版社, 2010.
- [2] SCHREER O, KAUFF P, SIKORA T. 3D video communication [M]. Chichester: John Wiley & Sons Inc, 2005:23.
- [3] OZAKTAS H M, ONURAL L. Three-dimensional television [M]. Berlin: Springer, 2008:317.
- [4] MERKLE P, SMOLIC A, MULLER K, et al. Efficient prediction structures for multiview video coding [J]. IEEE Transactions on Circuit and System for Video Technology, 2007, 17(11):1461-1473.
- [5] SHEN L Q, FENG G R, LIU Z, et al. A macroblock level adaptive search range algorithm for motion estimation in multi-view video coding [J]. J Electron Image, 2009, 18(3):033003.
- [6] SHEN L Q, LIU Z, LIU S X, et al. Selective disparity

- estimation and variable size motion estimation based on motion homogeneity for multi-view coding [J]. IEEE Transaction on Broadcasting, 2009, 55(4):761-766.
- [7] CHENG C C, LI C T, CHEN L G. A 2D-to-3D conversion system using edge information [C] // 2010 Digest of Technical Papers International Conference on Consumer Electronics. 2010:377-378.
- [8] ISO. Call for 3D test material: depth map & supplementary information [R]. ISO/IEC/JTC1/SC29/WG11, MPEG 2009/n19357, 2009.
- [9] 张倩,安平,张兆杨,等. 应用图像修复的基于深度图像绘制[J]. 光电子·激光,2009,20(10):1381-1384.
- [10] TOMASI C, MANDUCHI R. Bilateral filtering for gray and color images [C] // Proceedings of the Sixth International Conference on Computer Vision. 1998:839-846.
- [11] 张秋闻,安平,张艳,等. FTV 中面向虚拟视点合成的深度编码[J]. 应用科学学报,2011,29(3):299-307.
- [12] XIA X, LIANG J. An improved depth map estimation algorithm for view synthesis and multiview video coding [C] // Proceedings of Visual Communications and Image Processing. 2010:774429.
- [13] WILDEBOER M O, FUKUSHIMA N, YENDO T, et al. A semi-automatic multi-view depth estimation method [C] // Proceedings of Visual Communications and Image Processing. 2010:77442B.
- [14] TANIMOTO M, FUJII M T, SUZUKI T K, et al. Reference softwares for depth estimation and view synthesis [R]. ISO/IEC/JTC1/SC29/WG11, MPEG 2008/m15377, 2008.
- [15] EKMAKCIOGLU E, VELISAVLJEVIC V, WORRALL S T. Content adaptive enhancement of multi-view depth maps for free viewpoint video [J]. IEEE Journal of Selected Topics in Signal Processing, 2011, 5(2):352-361.
- [16] LIU S, LAI P, TIAN D, et al. New depth coding techniques with utilization of corresponding video [J]. IEEE Transactions on Broadcasting, 2011, 57(2):551-561.
- [17] ZHANG Q W, AN P, ZHANG Y, et al. View rendering distortion estimation for 3D video coding [R]. ISO/IEC/JTC1/SC29/WG11, MPEG 2011/m19456, 2011.
- [18] TANIMOTO M, FUJII T, FUKUSHIMA N. 1D parallel test sequences for MPEG-FTV [R]. ISO/IEC/JTC1/SC29/WG11, MPEG 2008/m15378, 2008.
- [19] SECKER A, TAUBMAN D. Highly scalable video compression with scalable motion coding [J]. IEEE Transaction on Image Processing, 2004, 13(8):1029-1041.

(编辑:赵宇)