

# A Novel Inpainting-Based Layered Depth Video for 3DTV (译)

Ismaël Daribo and Hideo Saito

**摘要 (Abstract)** —基于深度图的视频 (Layered Depth Video, LDV) 被认为是具有前景的 3D 立体视频数据的表示方式, 同时支持多视点视频 (Multi-view Video, MVV) 系统的高级 3D 视频质量需求, 如 3D 电视 (Three-Dimensional Television, 3DTV)。这种表示方式包含了一个以视频加深度形式的完整或中心视点序列作为主要层 (main layer), 同时, 包括冗余的纹理和深度数据在内的额外信息用以表示边视点。因此, LDV 既是拓展, 同时也是可选择的多视点视图加深度 (Multi-view Video-plus-Depth, MVD) 的另一种表示形式, 只传输一个通道的完整的视点和相关的冗余层数据。然而, 这里存在一种问题, 冗余层数据随着主视点和边视点距离的增加而迅速增加。当边视点的部分区域在主视点中不可见时, 留下的空白点称为堵塞 (disocclusions)。这些堵塞可能会非常大, 结果大大增加了主要层需要传输的冗余层数据。在这篇文章中, 我们针对冗余层生成的问题提出了一种基于修复 (inpainting-based) 的 LDV 生成方式, 通过恢复主要层中丢失的像素来减少冗余层数据总量。在提出的方法中, 我们通过比较部分场景的前景和后景的深度信息, 实现以较低的复杂度在纹理 (texture) 和结构传播 (structure propagation) 方面的修复算法。实验结果表明了提出的方法的有效性。

**术语索引 (Index Terms)** —基于深度的图像修复 (DIBR), 修复技术 (inpainting), 深度图视频 (layered depth video), 3D 视频 (3DTV)。

## I. 介绍

3D 视频点对点通信的 3D 数据表示在近年来的研究热度大幅度增长, 尤其是 3D 电视 (3DTV) 的制造带来了被认为是下一代多媒体消费者产品。3D 电视有着很长的历史。历经多年, 学者们达成了一致的观点: 3D 广播服务只能在察觉的图像质量和观看舒适度达到传统 2D 电视 (2DTV) 的水准时, 才能被成功地进入大众的使用。对于 3D 技术的不同方面的改进增加了对多视点视频应用的兴趣, 例如 3D 电视提供了不需要额外佩戴眼镜的深度感受。立体显示技术在多个显示制造商的努力下大大发展。

虽然毫无疑问高清晰度电视 (High Definition Television, HDTV) 已经成功在电视中实现, 但它仍然缺少一个非常重要的特征: 自然深度感受的表示。目前,

3D 电视被认为是逻辑上的将 3D 感引入 HDTV 的视觉体验。多视点视频 (MVV) 系统最近引起了广泛的研究兴趣, 尤其是就视点合成技术而言。视点合成 (View synthesis) 通常分为两类: 基于几何的修复 (Geometry-Based Rendering, GBR) 和基于图像的修复 (Image-Based Rendering, IBR)。典型地, GBR 技术开发了对场景的 3D 几何纹理了解, 它需要对对象进行 3D 建模。然而, 这样的建模需要数百万的多边形, 复杂的光模型, 广泛的纹理映射和很高的运算代价。IBR 技术作为可选择的替代 GBR 技术来合成视点而受到关注。不像原始的 3D 几何技术将一系列的图像用于生成其他图像。在各种各样的 IBR 技术中, 深度层图像 (Layered Depth Image, LDI) 是其中的一个效率非常高的针对复杂 3D 场景的视点合成方法。LDI 生成首次研究是在立体场景, 之后在多视点摄像机中运用。

一个 LDI 的拓展表示法, 深度图视频 (LDV) 被提出作为 3D 视频数据的表示。LDV 被认为是通过提供一个完整或中心的视点的 3D 视频数据表示的主要层, 并且包含纹理和深度数据表示边视点的额外增强层。LDV 因而既是拓展又是多视点视频加深度 (MVD) 的一种可选表示形式: 它只传输一个通道的完整的视点 (包含相关冗余层信息), 之后, 没被传输的边视点通过视点合成而作为主视点和边视点之间的转变。主视点之后通过 IBR 转换成边视点。然而, 问题是并不是每个像素在每个视点中都是存在的, 因而造成了主视点变换时造成的空洞的出现。视点合成用于显现出部分主视点的堵塞点并使它们在边视点中可见。这个过程称为“解闭塞” (disocclusion)。

处理这些堵塞的一种方法是依靠预处理深度图视频来减少深度数据的不连续性, 而减少堵塞。然而, 这意味着引入了滤波处理而造成深度图视频的失真, 将严重降低用户对原始深度的察觉。通过更高复杂度的多维度数据表示来移除堵塞是可能的, 比如 LDV 数据表示, 允许储存额外的在主视点中堵塞的像素深度和彩色信息。这些额外的数据为修复观众视角提供了填补闭塞区的重要信息。

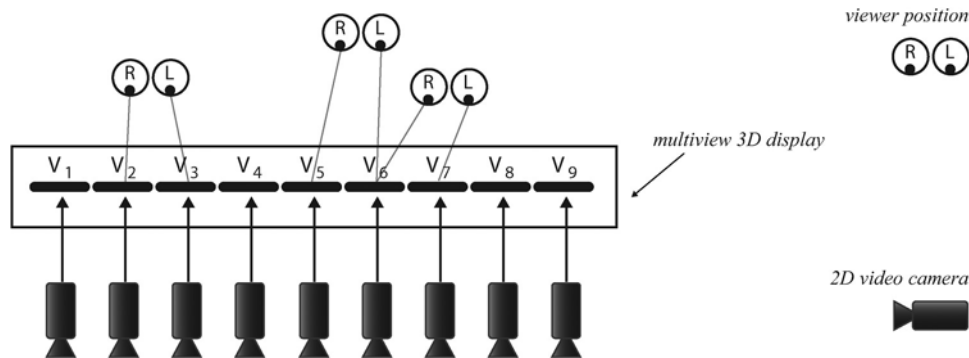


图 1. 基于 MVV 的多视点自动立体显示有效支持

Tauber 提出的一个解决方式包含了综合修复技术的 IBR 来解决大的堵塞区，主要利用了图像中损坏的空洞和视点合成中堵塞区的相似性。图像修复，同样也被认为是图像补足，通过像素附近的信息来填补大的空缺像素。图像和视频的修复有着很广的应用领域，例如移除层叠的文字和商标，重载由于刮擦和磨损造成的扫描图像的损坏，压缩图像和生成图像艺术感。艺术描述方法被广泛归类为结构性的修复或者纹理性的修复。结构性修复的重建首先假设了缺失区域的平滑度和边界情况，而纹理性修复只考虑从纹理样本或已知区域模板的可用数据。

最先由 Bertalmio 提出，结构性修复使用各向同性的扩散或者更复杂的非各向同性扩散来在同性方向扩散边界数据，另外，结构在丢失区域的平滑度也首先假设出。问理性修复使用了丢失区域的统计学或者模板特性，通常用马尔科夫随机域 (Markov random fields, MRF)。Levin 建议释放相关的图片已知部分的统计信息，然后通过 MRF 框架连接它们。

在本篇学习中，我们使用了 Criminisi 的工作成果作为起始点。它们尝试通过一个非常有深度的连接结构和纹理修复方法的优点的理论：纹理根据它各向同性方向的强度来修复。然而，这保留了纯粹的二维方法。我们在这里提出通过增加深度信息来区别一个像素属于前景还是背景的拓展性见解。清晰地表明了临近物体的空洞的轮廓，在背景邻近区域中哪一个在内容中增进了修复算法。我们同样提出以较低的额外的运算复杂度和无损优化来发挥深度信息的优势。

文章剩余部分组织如下。我们在第二部分 (Section II) 介绍了一些相关的 3D 视频格式的背景知识，并且我们在第三部分 (Section III) 描述了冗余层数据的生成。第四部分 (Section IV) 简要回顾了 Criminisi 的修复算法，第五部分 (Section V) 关注于减少需要传输的冗余层数据总量。我们最终的结论将在第六部分 (Section VII) 中阐述。

## II. 背景

为了支持 3D 电视系统的需求，许多 3D 视频数据的表示方法就复杂度，效率和效果而展开研究，根据如下几点要求：

- 可以使用尽可能多的已有的基础设施和媒介

- 对设备条件需要最小的改变
- 后向兼容性—不允许 3D 服务无法兼容现有设备
- 可以支持广泛的显示设备并允许未来的拓展性
- 高质量

立体系统是广为人知和简易的为 3D 视频数据表示的获取技术。立体视频可以通过使用左右一对视频提供一个 3D 映像，因而设计出一个立体摄像机系统，这是单通道 2D 视频无法达到。一对 2D 视频需要：一个是左眼图像，另一个是右眼图像。立体视频的标准中，MVV 可以被认为作为许多视点的立体视频数据表示的拓展。

多视点自动立体显示同时在特定视点区域处理多个视点，连续的视点犹如立体对。结果，头部运动视差视点可以被控制在有效范围内，但是，相比传统立体数据和 2D 电视，需要处理和传输的数据总量会大大增加。多视点自动显示和 MVV 的广泛应用刺激了输出视频的数量。用户因此可以选择他们的视点 (例如 超级保龄球 XXXV 运动，子弹时间效应等)。这样的高级 3D 视频应用需要一个能够修复输出视点连续性和在解码端的大量不同输出视点的 3D 视频格式。MVV 格式仍然不足以在没有额外增加输出视点和带宽的情况下支持这样的需求。



图 2. 基于视频加深度内容的自动立体显示的有效支持

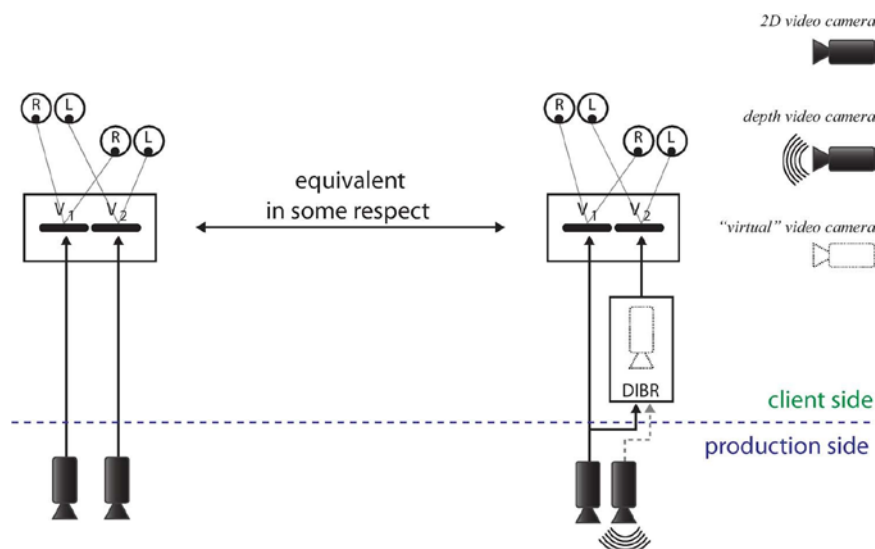


图 3. 纹理图和相应的深度图

视频加深度的数据表示形式被引入来解决这样的问题。它可以在接收端响应立体视觉，如图 2 所示，并且同时相比传统的立体视频数据表示，能大幅度减少传输带宽。通过在计算机视觉领域的初步学习，视频加深度的格式提供了一个通用的利用相关深度视频的 2D 视频（见图三）。

2D 视频提供了纹理信息，色彩浓度还有图像的结构，而深度视频标识的是每像素的 Z 轴距离，从摄像头的光学中心到视觉场景的三维点。将来，相对深度视频，2D 视频可能只表示纹理视频。

在从多个 2D 视频输入中估计深度信息已取得巨大进展。感谢近年来半导体处理的发展，因而有可能利用 TOF（光束飞行时间）摄像头来直接捕捉深度视频，也被称作深度摄像头。TOF 摄像头基于 TOF 技术，能够在真实时间里测量摄像头和场景之间的距离。这种摄像头发出的红外线光束能够被环境反射然后返回摄像头的传感器。每个像素的光束来回的时间被测量，然后用以计算场景的深度。深度视频可以被认为单色的没有纹理的视频信号。通常，深度数据会被量化为 8 比特（bits），即，最近的点会被量化为 255 而最远的点会被量化为 0。通过这种方式，深度视频被表示为一个平滑的灰度信息表示。

在客户端，通过利用深度图像修复技术（Depth Image Based Rendering, DIBR）来对第二个与次要视点相关的彩色视频进行重构。从接收端的视频加深度数据中生成立体视频是相对传统立体视频数据表示的一种拓展形式。因而，3D 的映像可以在传输后被自主调整。然后，由于视点合成引起的变化会由于修复视点的距离而快速增长，视频加深度只能支持可用的原始视点的非常有限的连续性。为了克服这个问题，MPEG 开始一个发展 3D 视频标准的活动，志在支持这些需求。这个标准给予了视频加深度的格式（MVD）格式，如图四。视频加深度可以与多视点视频（MVV）相联系，进而形成 MVD 格式，包含了多个 2D 视频，它们每一个都有各自的深度视频。

这个过程的最后一步是通过 DIBR 技术得到的数据来修复多个中间层视点。在这一点上，主视点和边视点将会被完整地处理和传输。对于除中心视点之外的，传输完整的边视点的一个可选择的方案是 LDV，可以在只考虑作为主要层的中间视点和增强层的冗余数据来减少冗余信息。这个心得表示方式会传输给目标一个高质量、高分辨率的图像，而相比那些传输多视点视频加深度（MVD），LDV 达到了更低的比特率。在下一章节，我们会讨论得到冗余层信息的重要性。

### III. 冗余层生成

我们将会用三个视频摄像头系统来描述冗余层信息的生成，也就是一个中心视点和两个边视点（左右两个视点），正如图 4 所示。

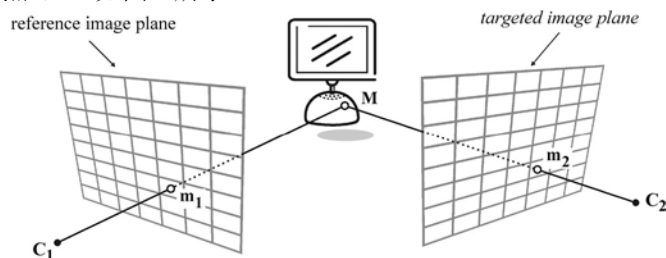


图 4. 多视点自动立体显示效率支持

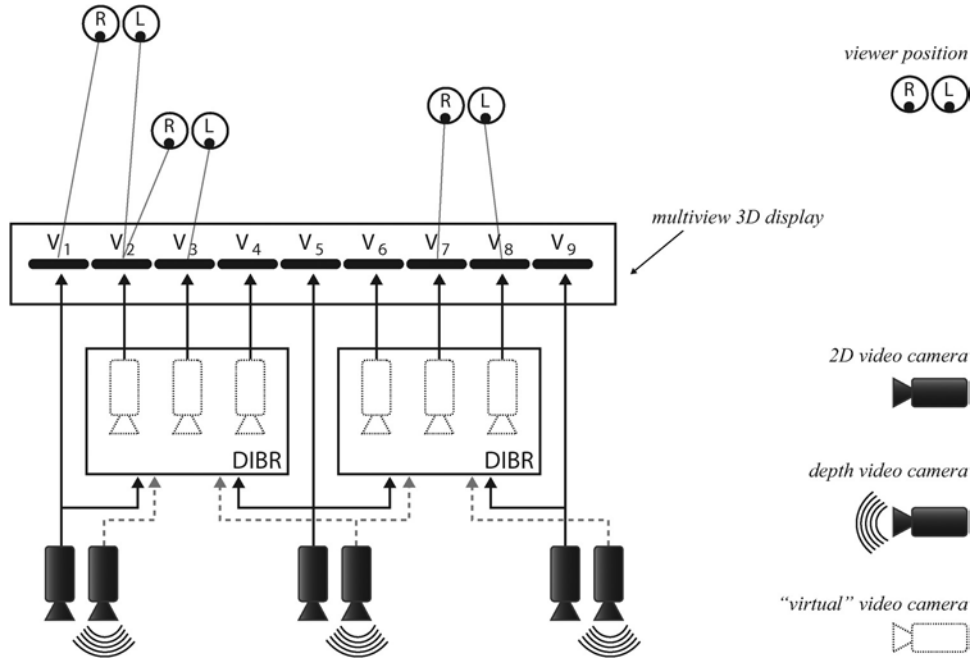


图 5. 3D 图像扭曲，齐次坐标系下 3D 点的映射

生成过程可以分为两个主要部分。首先，中心视点通过 DIBR 技术，利用给定的深度信息转换为每一个边视点。这个过程称为 3D 扭曲 (3D warping)。下一步，通过做减法，是可以知道边视点的哪些部分是存在于中心视点的。这些部分之后被标记为纹理和深度冗余层数据，而余下的部分被忽略。这个过程包括了一个将中心视点（参考图像坐标系）和对应边视点（目标图像坐标系）的映射过程。如图 5 所有，也将在下一章节中介绍。

#### A. 3D 扭曲

首先，我们将会介绍一些标记信息。参考视点图像  $I_1$  的浓度在像素坐标系  $(u_1, v_1)$  中被表示为  $I_1(u_1, v_1)$ 。通过给定的深度数据  $Z(u_1, v_1)$ ，针孔摄像机模型用语将  $I_1$  转变到第二个视角  $I_2(u_2, v_2)$ 。概念上，3D 图像扭曲过程可以被分为两个步骤：后向的把参考图像转变到 3D 世界，然后把后向转变的 3D 场景转变到目标图像坐标系中。如果我们看一下坐标系的位置  $(u_1, v_1)$ ，首先，一个对每像素的后向转变是从 2D 参考摄像头图像坐标的  $I_1$  到 3D 世界坐标系的转变。接下来，第二个转变是从 3D 世界像图像坐标系  $I_2$  在像素位置  $(u_2, v_2)$  的转变，诸如此类。为了实现这些操作，需要三个常量：  $K_1$ ,  $R_1$  和  $t_1$ ，表示的是一个  $3 \times 3$  的本征矩阵，  $3 \times 3$  的转置正交矩阵和一个  $3 \times 1$  的参考视点的转换向量。3D 世界的后向转变点被表示为非齐次的坐标 (公式 1)：

$\lambda_1$  这里是一个正值的标量。

再来看看目标摄像机的参数  $K_2$ ,  $R_2$  和  $t_2$ 。后向转变的 3D 世界的点之后通过齐次坐标被映射为目标 2D 图像坐标系 (公式 2)：

我们因而可以用参考坐标系来表示目标坐标系的方程 (公式 3)：

通常把世界坐标系系统用第一摄像机系统表示，因而  $R_3 = I_3$ ,  $t_3 = o_3$ ，可以被简化如下 (公式 4) 其中  $(u'_2, v'_2, w'_2)^T$  是 2D 图像点  $m_2$  的齐次坐标点，并且正值的标量等于 (公式 5)

最后一步，齐次结果被转换为像素点的位置。

最后，3D 世界点的最后一个元素表明了图像中像素点的深度信息。这个数据被认为是关键的边界信息，用以恢复相应的其他图像的像素位置。

#### B. 冗余层

正如之前所提到的，将中心视点 3D 扭曲为两个边视点显示出了重叠部分 (如图 6 所示)，之后将需要作为中心视点的额外部分而被传输。这些堵塞区域主要集中在深度不连续区域的前景物体。这些边视点之后降低为冗余层，如图 7 所示，通过从一个给定的视点中减去转换的中心视点。结果将会大大减少数据速率。

在用户端，中心视点和冗余数据被解压用语重构原始的边视点 (如图 8 所示)，导致了新的视觉体验和更高的用户交互性。

### IV. Criminisi 的修复算法

Criminisi 第一次报告了基于模板的纹理合成包含了重复纹理和结构的必要处理。它们使用了 Efros 和 Leung 的采样概念，并且展现出了输出图像合成的质量被修复过程的顺序所大大影响。



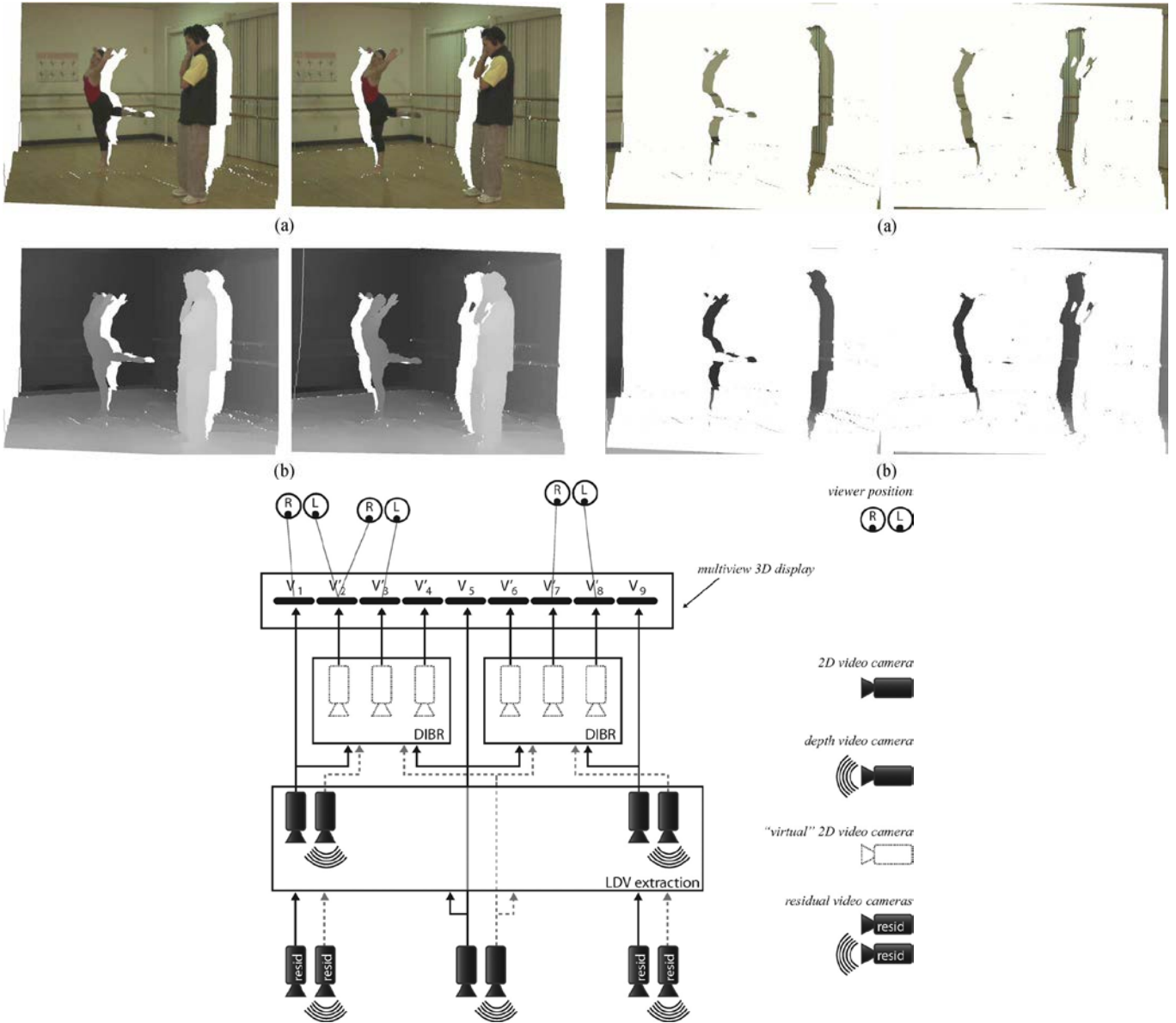


图 8. 基于 LDV 内容的多视点自动显示支持

Fig. 6. 将中心视点 3D 扭曲为两个边视点

Fig. 7. 在两个边视点冗余层数据.

通过输入图像和缺失区域，源区域 $\phi$ 被定义为 $\phi = 1 - \Omega$ （如图 9）。算法通过最优填补的方式展示了合成效果，并且优化程度依赖于边界信息 $\delta\Omega$ 。对于给定的填补包 $\psi_p$ 位于 $p \in \delta\Omega$ 的区域内（如图 9），被两个项所决定（公式 6）：

其中 $C(p)$ 表示当前补丁和数据项的可信度，从而在隔绝区给定了特别的优先度。这些项被定义为（公式 7）：

该区域为标准化的因数（ $\alpha$ 为典型的 255 阶灰度图）， $C(p)$ 表示了非丢失区域像素的百分比并被在起始为缺失区域，这对所有地方都有相同的结论。一旦优先度被计算得

出，一个基于块的匹配算法将会衍生出最好的用于填补像素的高优先权补丁，如下（公式 8）

这里的 $d$ 代表两个补丁之间的距离，使用均方差的值决定。对于已经发现的 $\Psi_q$ 的值，每个像素需要填补的关联像素将会被使用。在补丁填补之后，可信项 $C(p)$ 值将会更新如下（公式 9）：

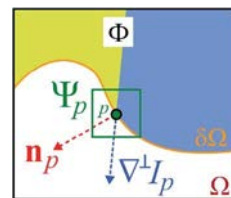


图 9. 标记图 [9].

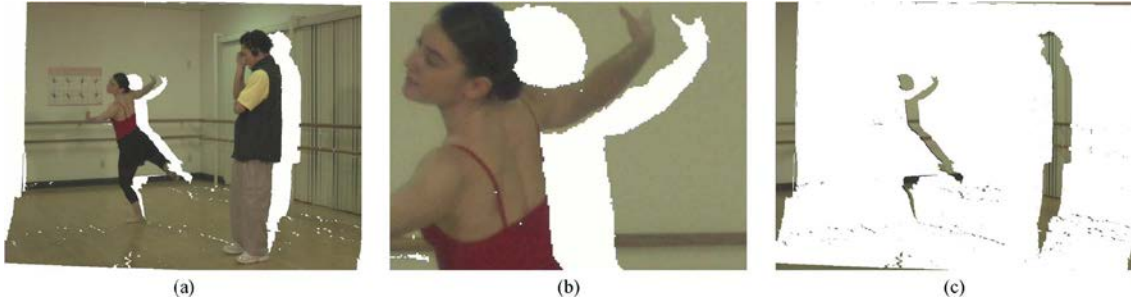


图 10. 解闭塞区用以填补空洞而产生冗余层信息



图 11. 空洞填补的例子: (a) 原始 Criminisi 算法, (b) Oh 结合 Criminisi 的算法, (c) 通过扭曲深度信息增强后的效果

(a) Criminisi 修复算法, (b) Oh 结合 Criminisi 的修复算法, (c) 提出的修复算法

### V. 基于深度的纹理和结构性蔓延

虽然现有的基于 3D 图像的修复和解闭塞技术有着概念上的相似性,但是我们仍然对解闭塞有着更为优先的了解。解闭塞的情况是部分背景区域被前景物体遮挡。使用背景填补闭塞区域因此会对前景物体更为敏感。为了实现这种填补,Cheng 发展了视点合成框架,框架使深度信息强制了纹理匹配的搜索范围并且依靠空间和深度信息来对纹理图信息进行滤波,因而增强了视点合成的质量。在类似的学习领域里,Oh 提出了从相反的一遍利用背景物体来替代前景物体。它们首先处理了闭塞区域的边界因而只保留了背景的部分像素,然后利用现有的修复技术进行修复。

基于这些已有的成果,我们提出了一个相面深度的纹理修复技术,该技术使用了 Criminisi 的算法原理:相比前景物体,给予背景像素更高的优先权。

#### A. 优先权计算 $n$

在我们的优先权计算中,目标图像坐标轴里给定了相关联的深度补丁  $Z$ 。通过增加第三个相乘项,我们提出使用之前的优先度计算权重,如公式 10。

其中  $L(p)$  指的是等级规则性项,被定义为深度补丁的相反项  $Z_p$ ,如公式 11。

其中  $Z_p$  指的是  $Z_p$  的区域(就像素数而言),而  $\bar{Z}_p$  指的是平均值。我们基于覆盖背景的补丁包以相同的深度等级,是一种自然情况下的认知,即背景像素优先前景像素。

#### B. 补丁包匹配

考虑到深度信息,我们将公式 8 更新如下:

基于块的匹配算法将会在纹理和深度两个域中进行,其中以来的参数为  $\beta$ 。该参数允许我们控制给定的深度距离最小化的重要性。通过更新距离的衡量方式,我们遍历搜索了相同深度等级的补丁

基于提出的基于深度的修复技术,我们可以通过利用中心视点中可用的纹理信息从某种程度上恢复缺失的像素。这就是减少冗余数据。最后一步包括了决定映射的中心视点的哪些部分将会被传输。我们提出了思考修复技术的人为性作为新的冗余层数据,因为相比图 12 的原始闭塞区域,闭塞区域明显大大减少。除此之外,不仅大大减少了冗余区域,我们还克服了 Muller 所提出的问题。从两个边视点中合成冗余数据到一个缓冲器中以便更深一步减少数据速率。



图 12. 冗余数据的例子

## VI. 实验结果

多视点视频加深度的序列“芭蕾”是微软公司提供用来测试所提出的方法的。校准参数也随序列一同提供。每个摄像头提供的深度视频被通过一个基于色彩的位算法来估计。由于中间摄像头和边摄像头的间距大基线近距离，3D 扭曲过程会出现更多的堵塞。

正如我们从图 10 所见，堵塞边界即属于场景的部分前景也属于部分背景。这使得传统的修复方法效率低下。我们现在将比较我们提出的方法和 Oh 的工作成果。Oh 提出了一个预滤波的步骤，针对既属于前景又属于后景的堵塞的边缘。为了提供一个公平的比较方式，我们在前后景复制粘贴后使用了 Criminisi 的修复算法。我们将这种联合称为“Oh+Criminisi”。多数相关工作将会集中在至少两个参考视点用来扭曲到虚拟视点的框架中。结果是出现更少的堵塞区，堵塞区也同时变得更小。我们的问题有一点非常值得注意，只有一个参考视点是可用的（既中心视点），，因此导致了较大的堵塞区，因而传统的修复方法显得无力。

使用区域填补的方法，原始的 Criminisi 算法以及 Criminisi 结合 Oh 提出的方法结果见图 11。比较三个方法，很显然的是我们的算法能更好地捕捉前景物体的轮廓，并且增强修复图像的视觉质量。这是利用了纹理和背景区域的拓展，而 Crimini 的算法并没有区分这两者。尽管 Oh 加上 Crimini 的方法可以从背景中依靠堵塞边界来区分前景，这个方法不适合于大的堵塞区，而且主要依赖于计算得比较好的堵塞边界。

我们可以在图 13 中看到为我们提出方法对质量的提升，并且在图 12 中，可以看见冗余数据会显著减少。冗余数据的减少是非常重要的，如图 12。并且图 13 中我们的方法更清晰可见。通过加入一个新的基于深度的项来填补优先的计算，我们没有包含任何计算上的增加，或者是预处理的需求。由于之前的很多方法都需要预处理过程或者必将导致额外的计算量。

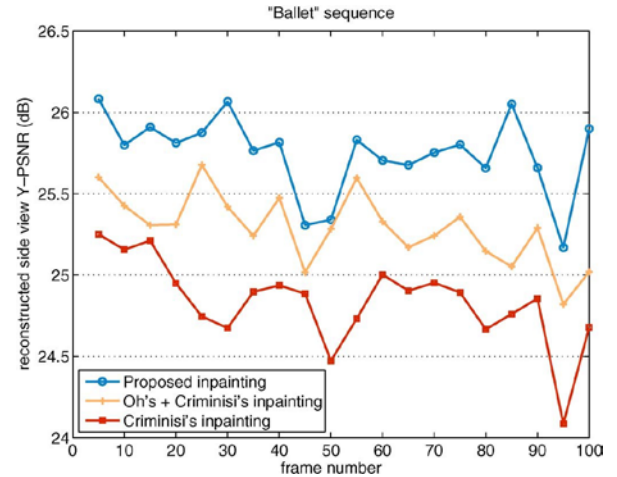


图 13. 对象的 PSNR 结果

## VII. 结论

在这篇文章里，我们通过提出一个后置的基于图像修复的堵塞的解决方法，从而得到 LDV 数据表示中的冗余层数据生成，通常这是通过沿着空洞的轮廓来对纹理和结构进行扩展。提出的方法使用了 Criminisi 的算法，并且深度信息加入了优先计算和补丁匹配。提出的算法依赖于纹理和结构的扩展，同时也考虑了部分图像前景与背景的辨识。

在不提高复杂度的情况下是可以实现视觉质量的改进的，并且没有预处理的需要。我们利用中心视点中可用的纹理来填补这些空洞区域。实验结果展现了修复图像的视觉质量尤其可以在保留前景轮廓的条件下提升。因此，除了传输完整的中心视点外，减少了冗余层数据的总量。

几个问题被留下作为更深入的研究。在未来学习中，我们希望关注对短暂一致性的研究来保证空间和短暂的稳定性，以及对抗前景深度图中的错误的鲁棒性。

## 引用

- [1] N. A. Dodgson, "Autostereoscopic 3D displays," *Comput.*, vol. 38, pp. 31–36, Aug. 2005.
- [2] Y. Zhu and T. Zhen, "3D multi-view autostereoscopic display and its key technologies," in *Asia-Pacific Conf. Inf. Process. (ACIP)*, 2009, vol. 2, pp. 31–35.
- [3] J. W. Shade, S. J. Gortler, L.-W. He, and R. Szeliski, "Layered depth images," in *Computer Graphics*, Jul. 1998, vol. 32, Annual Conference Series, pp. 231–242 [Online]. Available: <http://grail.cs.washington.edu/projects/ldi/>

- [4] X. Cheng, L. Sun, and S. Yang, "Generation of layered depth images from multi-view video," in *Proc. of the IEEE Int. Conf. Image Process. (ICIP)*, San Antonio, TX, 2007, vol. 5.
- [5] Z. Tauber, Z.-N. Li, and M. Drew, "Review and preview: Disocclusion by inpainting for image-based rendering," *IEEE Trans. Syst., Man, Cybern., Part C: Appl. Rev.*, vol. 37, no. 4, pp. 527–540, Jul. 2007.
- [6] B. Furht, *Encyclopedia of Multimedia*, 2nd ed. New York: Springer, 2008.
- [7] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in *Proc. of the annual Conference on Computer graphics and interactivetechniques(SIGGRAPH)*, New Orleans, USA, Jul. 2000, pp. 417–424.
- [8] A. Levin, A. Zomet, and Y. Weiss, "Learning how to inpaint from global image statistics," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Nice, France, Oct. 2003, vol. 1, pp. 305–312.
- [9] A. Criminisi, P. Perez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Trans. Image Process.*, vol. 13, no. 9, pp. 1200–1212, 2004.
- [10] T. Oggier, M. Lehmann, R. M. Rolf Kaufmann, M. S. , P. Metzler, G. Lang, F. Lustenberger, and N. Blanc, "An all-solid-state optical range camera for 3D real-time imaging with sub-centimeter depth resolution (swissranger)," in *Proc. SPIE Conf. Opt. Syst. Design*, 2003, vol. 5249, pp. 634–645.
- [11] L. McMillan, Jr., "An image-based approach to three-dimensional computer graphics," Ph.D. dissertation, University of North Carolina at Chapel Hill, Chapel Hill, NC, 1997.
- [12] W. Mark, "Post-rendering 3D image warping: Visibility, reconstruction, and performance for depth-image warping," Ph.D. dissertation, University of North Carolina at Chapel Hill, , NC, Apr. 1999.
- [13] M. M. Oliveira, "Relief texture mapping," Ph.D. dissertation, University of North Carolina at Chapel Hill, , NC, 2000.
- [14] A. Smolic, K. Mueller, P. Merkle, N. Atzpadin, C. Fehn, M. Mueller, O. Schreer, R. Tanger, P. Kauff, and T. Wiegand, "Multi-view video plus depth (MVD) format for advanced 3D video systems," in *Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, JVT-W100 doc.*, San Jose, CA, Apr. 2007.
- [15] A. Efros and T. Leung, "Texture synthesis by non-parametric sampling," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Kerkyra, Greece, Sep. 1999, vol. 2, pp. 1033–1038.
- [16] C.-M. Cheng, S.-J. Lin, S.-H. Lai, and J.-C. Yang, "Improved novel view synthesis from depth image with large baseline," in *Proc. Int. Conf. Pattern Recog.*, Tampa, Finland, Dec. 2008, pp. 1–4.
- [17] K.-J. Oh, S. Yea, and Y.-S. Ho, "Hole filling method using depth based in-painting for view synthesis in free viewpoint television and 3-D video," in *Proc. Picture Coding Symp. (PCS)*, Chicago, IL, USA, May 2009, pp. 1–4.
- [18] K. Muller, A. Smolic, K. Dix, P. Kauff, and T. Wiegand, "Reliabilitybased generation and view synthesis in layered depth video," in *Proc. IEEE Workshop Multimedia Signal Process. (MMSP)*, Cairns, Queensland, Australia, Oct. 2008, pp. 34–39.
- [19] "Sequence microsoft ballet and breakdancers," 2004 [Online]. Available: <http://research.microsoft.com/en-us/um/people/sbkang/3dvideodownload/>
- [20] C. L. Zitnick, S. B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski, "High-quality video view interpolation using a layered representation," in *Proc. Annu. Conf. Comput. Graph. Interactive Tech. (SIGGRAPH)*, Aug. 2004, vol. 23, pp. 600–608, no. 3.