基于视频模型的虚拟现实视频融合系统

周 颐[1)] 孟明[1)] 吴 威[1)], 周 忠[1)]

[1)] (虚拟现实技术与系统国家重点实验室, 北京航空航天大学, 北京 100191)

摘 要 虚实融合技术以减少虚拟对象和真实视频图片之间的视觉差异为目标，力图实现视觉感官上的虚实无缝结合。多视频的虚实融合技术普遍存在画面畸变与虚实对准代价大的问题。针对该问题，本文提出并设计了一种基于视频模型的虚拟现实视频融合系统。论文首先介绍了视频模型的建立原理和其对应的文件格式，阐述了系统中视频与场景融合绘制方法。在然后，本文对系统的总体结构进行了介绍。在实验阶段，我们测试了本系统的性能，并将这种虚实融合技术与国际上的虚实融合方法进行了对比分析。实验结果表明，本文方法在融合绘制效果和性能上均优于其他两类方法。

关键词 视频投影；视频模型；纹理贴图；图片建模；虚实融合技术

中图法分类号 TP391.41 DOI 号

Virtual-reality Video Fusion System based on Video Model

Zhou Yi[1)] Meng Ming[1)] Wu Wei[1)] Zhou Zhong[1)]

[1)](State key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing 100191, China)

Abstract The aim of virtual reality integration is to reduce the visual difference of virtual object and real video image. However, it still has several problems, including heavy imagery distortion when the user deviates from the source viewport, and huge modeling cost of virtual scene. For these issues, this paper proposes a novel virtual-reality video fusion system based on *Video Model*. We firstly introduce how we build video model and its file format, then present the fusion method of video and virtual scenes. On the design and implementation stage, we introduce the architecture of our system. Finally, we test our system's render performance and compare it with two state-of-the-art virtual reality integration method, which demonstrates that our fusion method is better than other two methods.

Key words video projection; video model; texture mapping; image-based modeling; virtual-reality integration

1 引言

虚拟现实(VR)作为一个新兴的科学技术领域，近年来取得了巨大的进步。由于视频图像的大量普及，作为虚拟现实技术的一个重要方向，虚实融合技术的探讨与研究更是受到了大量关注。

虚实融合技术的目标是减少虚拟物体（或者场景）和真实视频图片之间的视觉差异，通过实现视觉感官上的虚实无缝结合，为用户提供更加真实的沉浸式视觉体验。该技术实现的重点之一在于以何种方式组织虚拟物体与真实视频，为用户展现虚实一体的画面。研究者们探索了各种方式来实现虚实融合的效果，产生了多种类型的知名虚拟现实应用。例如，谷歌的实时街景地图 Google Street View；景区照片浏览系统 PhotoSynth；体育赛事 NBA 的虚拟现实回放系统 FreeD 和三维融合多视频流的监控系统 Video Flashlights[2]等。

收稿日期：年-月-日；最终修改稿收到日期：年-月-日 *投稿时不填写此项*.基金项目: 国家863计划(2015AA016403); 国家自然科学基金(61572061, 61472020, 61502020); 中国博士后科学基金(2013M540039). 周颐(1988—), 男, 博士研究生, CCF学生会员, 主要研究方向为三维重建、虚实融合; 孟明(1991-)，女，博士研究生，主要研究方向为虚实融合；吴威(1961—), 男, 博士, 教授, 博士生导师, CCF高级会员, 主要研究方向为分布式虚拟环境; 周忠(1978—), 男, 博士, 教授, 博士生导师, CCF高级会员, 论文通讯作者, 主要研究方向为网络化虚实融合、视频大数据分析、VR装置.

表1 国内外相关方法比较

对比指标/融合方法类型	基于图像变形的融合方法	基于直接视频投影的融合方法	基于图片建模的融合方法
贴图方法	纹理贴图	纹理投影	纹理投影
绘制效率	最慢（两次渲染，一次遮挡测试，绘制效率取决于基础纹理缓存-面片数目）	慢（两次渲染，一次遮挡测试，绘制效率取决于场景大小）	快（一次渲染，绘制效率高）
多视频的扩展性	支持多视频融合	支持多视频融合，但对视频最大数量有限制	支持多视频融合
主要缺点	遮挡测试与选择 buffer 内容绘制速度较慢	独占着色器，影响其他着色器的执行	需要交互式人工建模
代表性单位	南加州大学 [4]	南加州大学、美国 Sarnoff 公司等 [1][2][3]	北京航空航天大学虚拟现实国家重点实验室 [6]

将多视频流进行三维融合的虚实融合技术给用户带来更真实的体验与感知，但在该技术应用的过程中，仍存在其自身需要进一步解决的问题。问题一：虚实融合的结果容易出现视频纹理扭曲和撕裂的畸变现象，其原因来自于拍摄的图片与建模场景自身的深度不契合，若是视频投影时偏离了原拍摄视角，也会产生不尽人意的融合效果；问题二：虚实对准代价大，国际上的方法需要首先构建相当尽可能高精度的三维场景，然后才可能进行准确对齐，例如，先采用 LiDAR 和航拍等方式采集数据，然后对采集的数据进行交互式编辑等大量的后期操作。在这些问题的基础上，研究出如何以低成本的方式构建虚实对齐的三维场景模型的方法，以及研究如何进一步改善虚实融合的效果是当前这一领域需要解决的关键问题。

针对以上问题分析，本文从一个新的观点出发，将图片重建技术为基础应用到虚实融合中以期得到更好的融合效果。其中图片重建技术是利用二维图片恢复物体三维信息的数学过程和计算技术，这一技术的特点是能够满足虚实融合过程中对虚拟模型精度的要求，进而克服了视频投影本身带来的深度不匹配的问题。我们认为图片三维建模技术与虚实融合技术相结合具有广阔的应用前景，对虚实融合技术进一步的发展具有极其重要的意义。

2 国内外研究现状

目前，国际上有关三维场景虚实融合的方法已有一些相关研究，本文将虚实融合的方法分为三类，分别是基于直接视频投影的虚实融合方法、基于视频图像变形的虚实融合方法和我们提出的基于视频图片重建的虚实融合方法。

基于直接视频投影的虚实融合方法：美国 Sarnoff 公司的 Stephen Hsu 等人于 2000 年，提出以实时视频流作为纹理投影到模型上的方法[1]，使用纹理映射方法计算模型表面的颜色。美国 Sarnoff 公司的 H.S.Sawhney 等人于 2002 年提出了 Video Flashlights 系统[2]。南加州大学的 Ulrich Neumann 等人于 2003 年在 IEEE VR 上阐述了增强虚拟环境 AVE(Augmented Virtual Environment）系统的概念，同样使用了实时视频增强虚拟场景的思想[3]。

基于视频图像变形的虚实融合方法：美国南加州大学的 Hu 等人于 2009 年提出了一种基于视频图像变形(image-warping)的视频与场景融合的方法[4]。该方法首先使用航拍影像和从地面拍摄的图像对场景进行贴图，重建出具有真实纹理的三维校园场景。

该方法创建出三维场景后，每次根据视点选择一路实时视频流对场景的纹理进行更新。Hu 等人提出使用基本纹理缓存(Base Texture Buffer)来保存和更新建筑表面的纹理。在更新模型纹理时，根据视频帧变形计算出基本纹理缓存，使用该纹理缓存对相应的模型表面进行更新。对于图像会出现内容交错的问题，其使用特征匹配计算每帧视频与基础纹理缓存之间的对应关系，使用该对应关系变形后的图像再次变形得到最终使用的纹理图像。

接下来对这两类方法与我们提出的方法进行横向对比分析，结果如表 1 所示。相比而言，本文的方法只针对视频模型进行投影，拥有视频模型和视频帧的一一对应关系，已经进行了纹理指定，不

会出现融合时的纹理选择问题。我们在实验中对比了这三类虚实融合方法的绘制方法，结果证明我们的方法要优于其他两类方法。

3 视频模型原理及 IBMT 格式

3.1 建模原理

我们首先提出了视频模型的概念，即每个视频对应一个模型。融合过程中使用的模型数据不仅仅包括代表三维场景的模型，还包括通过交互式建模得到的简单视频模型两种。

多幅图像的三维场景结构建模技术，本质上是将各个图像中的二维还原与之对应的三维场景结构信息。我们在相机标定的基础上进行交互式建模操作，定义基本的操作即可通过交互式的方法输入二维图元信息，将二维图元反投影到三维空间中，形成三维体元。交互式建模需要解决的问题无非分为点从图像坐标系到相机坐标系，再到世界坐标系的问题。

首先解决如何进行将模型原点放置到场景中去的问题。我们假设，图像中的物体是垂直站立的，同时世界坐标系的Z轴是平行于相机坐标系的垂直方向。这样六自由度(6-DOF)的注册问题退化为一个 XY-平面的对齐问题。我们可以通过在建模工具的三维视图和二维视图下的分别绘制两条地面线，即可实现对齐。

给定一对 2D 线段 L^i 和 3D 线段 L^w，分别来自图像平面的图像坐标系和三维场景的世界坐标系。其中 L^w 在三维场景的世界坐标系下是位于底图之上的。线段的端点 (X_1^i, X_2^i)，(X_1^w, X_2^w) 是对应匹配的。对于 L^i，我们可以建立一个局部的世界坐标系和一个对应的 3D 线段 L^c [8]，该线段是沿着局部坐标系 X 轴的直线，其端点是 (X_1^c, X_2^c)。两种三维直线的端点一定满足变换关系：

$$X_i^c = s\mathbf{M}X_i^w, i = 1,2 \tag{1}$$

其中 s 是尺度因子，$\mathbf{M}$ 是刚性变换矩阵。我们同时计算从底图世界坐标系的 X 轴到 L^w 的旋转角度 θ。我们可以通过线段长度之比来求解 s：

$$s = \left\|L^c\right\| / \left\|L^w\right\| \tag{2}$$

而刚性变换矩阵可以表示为，

$$\mathbf{M} = \begin{bmatrix} R(L^c, \theta) & X_1^w \end{bmatrix} \tag{3}$$

其中 $R(L^c, \theta)$ 是 X 轴沿着 L^c 旋转 θ 度的旋转矩阵。因此，对于视频模型中的任意点来说，如果 s 和 $\mathbf{M}$ 已知，我们就可以将其变换到三维场景中。

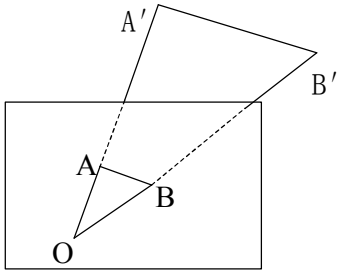


图 1 建模原理图

然后解决如何从图像二维点到相机坐标系的三维点的问题。我们通过灭点分析[8]得到平面坐标系下的单位轴向量 $\mathbf{D}_{imageplane}:\begin{bmatrix} u & v & q \end{bmatrix}$ 和其相机坐标系下对应的单位轴向量 $\mathbf{D}_{camera}:\begin{bmatrix} U & V & Q \end{bmatrix}$。我们只要知道用户绘制的点在图像的任意点B和其与参考点A的关系，即可得到B在相机坐标系下的位置。如图1所示，给定参考点A在图像平面上的坐标 $\mathbf{X}_A$ 与其在相机坐标系下的坐标 $\mathbf{X}_A^c$，则 B 在图像平面上的坐标 $\mathbf{X}_B$ 可以用 $\mathbf{X}_A$ 和 $\mathbf{D}_{imageplane}$ 表示：

$$\mathbf{X}_B = \mathbf{X}_A + \boldsymbol{\lambda}_{\mathbf{AB}}\mathbf{D}_{imageplane} = \mathbf{X}_A + \begin{bmatrix} \lambda_u & \lambda_v & \lambda_q \end{bmatrix}\begin{bmatrix} u \\ v \\ q \end{bmatrix} \tag{4}$$

其中 $\boldsymbol{\lambda}_{AB} = \begin{bmatrix} \lambda_u & \lambda_v & \lambda_q \end{bmatrix}$ 为 $\mathbf{D}_{imageplane}$ 线性表达向量 AB 的系数，即其与参考点 A 的关系。则 B 点在相机坐标系下坐标 $\mathbf{X}_B^c$ 可以用 $\mathbf{X}_A^c$ 和 $\mathbf{D}_{camera}$ 表示：

$$\mathbf{X}_B^c = \mathbf{X}_A^c + \boldsymbol{\lambda}_{\mathbf{AB}}\mathbf{D}_{camera} = \mathbf{X}_A^c + \begin{bmatrix} \lambda_u & \lambda_v & \lambda_q \end{bmatrix}\begin{bmatrix} U \\ V \\ Q \end{bmatrix} \tag{5}$$

现在图像平面上任意点B在世界坐标系下的坐标为：

$$\mathbf{X}_B^w = s^{-1}\mathbf{M}^{-1}\mathbf{X}_B^c = s^{-1}\mathbf{M}^{-1}(\mathbf{X}_A^c + \boldsymbol{\lambda}_{\mathbf{AB}}\mathbf{D}_{camera}) \tag{6}$$

3.2 IBMT文件格式

IBMT 文件格式是我们为视频模型设计的一种文件格式，以明文方式保存了单幅照片建模得到结果。其中包括建模原图片、模型文件、相机标定文件、融合显示时使用的裁剪纹理以及校正使用的纹理。因此，我们的建模方法可以应对于视频图像重叠，镜头畸变等常见现象，实现更为精细的视频融合。

4 基于视频模型的虚实融合方法

本文提出的基于视频模型的虚实融合方法，针对每个视频创建对应的视频模型[6]，然后通过纹理投影(Projective texture mapping)[5]和阴影投影的方法将视频与其模型进行融合。

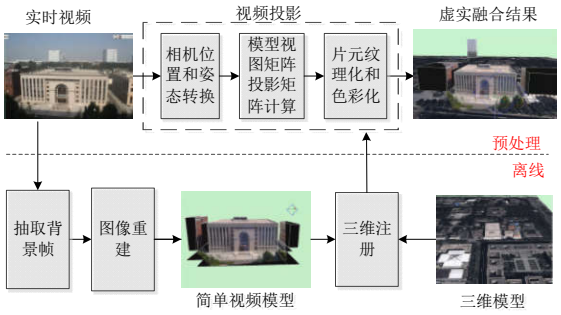


图 2 基于视频图片重建流程图

我们的方法分为两个阶段，整体流程如图 2：

(1)预处理阶段。我们在离线阶段提取视频的背景帧进行交互式建模[6]，得到简单的视频模型，然后与三维场景模型进行注册，此方法不要求与场景模型的精细程度，便于场景模型的建立与视频模型的建立分开进行。

(2)在线阶段。该阶段与直接投影的融合方法类似，但不再执行遮挡测试，其被遮挡部分在建模时被剔除，所需的深度信息在离线阶段已经计算。

然而多视频之间难免会存在重叠，我们使用透明分量混合(alpha blending)进行局部处理。如图 3 所示，视频 A 与视频 B 都包含重叠区域 C，用户可以采用手动划分的方式在重叠区域上指定分割线；将该条分割线分别投影到其中一个视频 A 的图像平面上，形成直线 L；基于直线 L 的两边扩展一定距离形成融合区，即图中由 l_1 和 l_2 以及重叠区域 C 包围形成的区域。

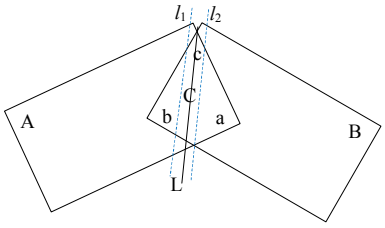


图 3 纹理混合示意图

在融合渲染时，混合颜色根据如下公式计算：

$$\mathbf{A}_{alpha}(p)\cdot\mathbf{A}_{color}(p)+\mathbf{B}_{alpha}(p)\cdot\mathbf{B}_{color}(p) \qquad (7)$$

其中 $\mathbf{A}_{color}(p)$ 与 $\mathbf{B}_{color}(p)$ 分别对应 A 图像和 B 图像的逐像素 p 的 color 值；$\mathbf{A}_{alpha}(p)$ 与 $\mathbf{B}_{alpha}(p)$ 分别为 A 图像和 B 图像的逐像素 p 的 alpha 值。当像素 p 位于融合区朝向图像 A 的区域时，alpha 值为 1，反之，为 0。当像素 p 位与融合区时其计算如下公式所示：

$$\mathbf{A}_{alpha}(p)=1-\frac{x}{d} \qquad (8)$$

图 4 纹理混合前后对比效果

其中，d 为融合区的宽度；x 为像素 p 到融合边界 l_1 的垂直距离。图像 B 的 alpha 值可同理得到。依据该方法得到的混合前后的效果对比图像，如图 4 所示。

5 系统设计与实现

我们设计研发了基于视频模型的虚拟现实视频融合系统的软硬件内，其由数据源、服务器和客户端三部分组成。数据源是该系统的输入源，主要负责视频和图片数据的采集与存储；服务器包括流媒体服务器、Web 服务器、图片服务器和权限服务器，负责控制对系统的访问并分发数据；客户端对来自流媒体服务器的视频数据进行解码、同步处理，实现视频图像与三维场景的融合绘制。其整体系统架构如图 5 所示。

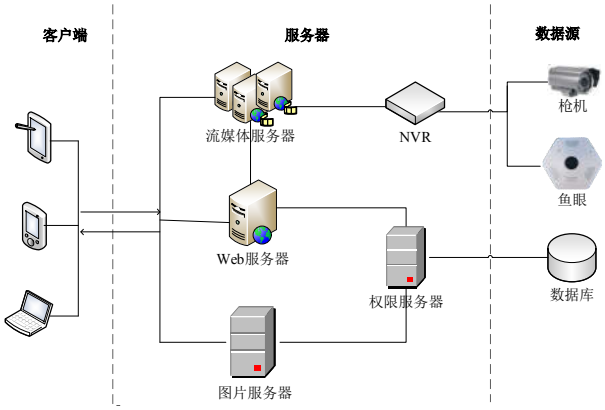


图 5 虚拟现实视频融合系统的架构

系统主要包括以下四方面的功能：(1)视频图像建模，系统提供了相机的快速标定与建模工具，支持视频图片的快速建模、场景注册、多图像的纹理融合和局部调整；(2)视频数据处理，系统设立了支持多类相机的流媒体服务器，对视频进行同步传输处理，各类视频采集设备以及离线视频被抽象成虚拟相机，动态挂载在不同的流媒体服务器之上；(3)虚实融合效果展示，系统开发了支持普通枪机以及

鱼眼相机等多种视频投影的客户端，实现了一览的虚实融合效果。

我们的虚实融合方法将视频与室内场景和室外场景进行融合，效果如图 6 所示。

(a)室外某建筑融合效果 (b)室外某广场融合效果

(c)室内某房间融合效果 (d)室内楼道B融合效果

图 6 虚实融合效果示意图

6 实验结果

6.1 实验环境

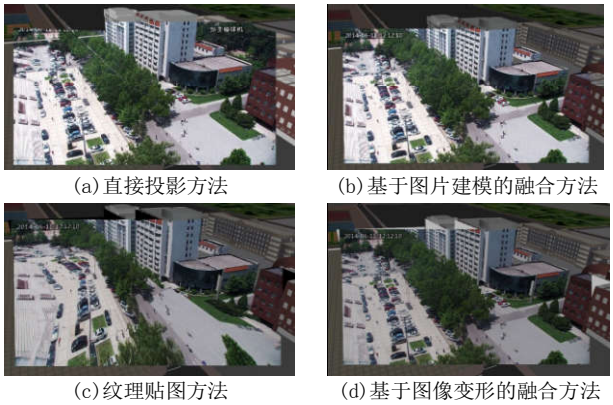
实验环境的搭建对实验的进行和分析有重要作用，它主要包括硬件环境和软件环境。其中所使用的硬件环境包括：一台服务器计算机，配置为双核 Intel(R) Core CPU 2.2GHz、2GB 内存；一台客户端计算机，配置为四核 Intel(R) Xeon(R) CPU 3.7GHz、8GB 内存，显卡型号为 NVIDIA GeForce GTX 750 Ti、2GB 显存。软件环境包括 Win7 操作系统、VS2010、C/C++和由 OSG、QT 和 FFmpeg 组成的开源库。

6.2 虚实融合效果对比

我们针对现有的直接投影方法、纹理贴图方法、基于图像变形的融合方法以及我们的基于图片建模的融合方法，进行了融合效果的对比实验，结果如图 7 所示。

图 7(a)中将图像投影至所有模型之上，随着远裁剪面的设置而在远处出现了少量图像的缺失。图 7(c)中由于采用均匀纹理采样方法的原因，生成的图像会出现扭曲。图 7(d)中纹理经过基础纹理矩阵的变换，不会出现扭曲，与图 7(b)中我们的方法效果基本一致。但是本文中的方法可以进行交互式编辑，如对右上角的不合理投影区可以进行裁剪，减少畸变。

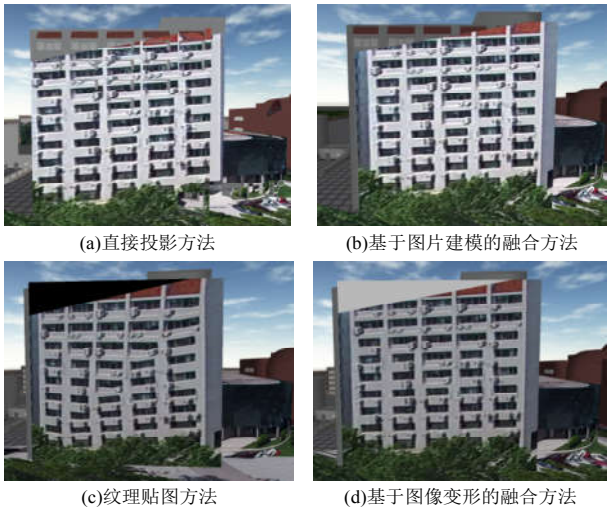
我们选择图 7 中的建筑立面进行局部放大，对比效果如图 8 所示。图 8(a)中由于图像与三维模型的不匹配，产生了撕裂现象；图 8(d)中基于图像变形的方法得到的纹理较优，与我们的方法一致。需要指出的是，我们在实现 8(d)中 Hu 等人的方法时，是将纹理贴在我们建模的视频模型之上，而不是直接贴图到建筑立面上，因此避免了可能出现的撕裂现象。由此从效果来看，本文的基于图片建模的融合方法更为合理。



(a) 直接投影方法 (b) 基于图片建模的融合方法

(c) 纹理贴图方法 (d) 基于图像变形的融合方法

图 7 各种融合方法的绘制效果对比



(a)直接投影方法 (b)基于图片建模的融合方法

(c)纹理贴图方法 (d)基于图像变形的融合方法

图 8 各种融合方法局部效果对比

6.3 虚实融合绘制性能对比

融合绘制性能对比实验包括资源占用性能和绘制帧数两大项测试。以北航场景为例进行资源占用实验，结果如表 2 所示。从表中分析得出，随着融合视频数的增加，对于 720P 和 1080P 的视频 CPU 和 GPU 的占用比增长呈加速趋势，对于常见的监控视频实时流（分辨率一般不高于 1080P），客户端可支持至少 8 路视频的融合，满足系统的基本需求。

表 2 资源占用实验结果表（合二为一）

资源类别/视频数		0	1	2	3	4	5	6	7	8
内存占用	720P	534	587	639	693	745	806	864	909	966
（MB）	1080P	538	652	790	906	1021	1140	1256	1363	1485
显存占用	720P	1018	1031	1045	1059	1073	1086	1101	1116	1130
（MB）	1080P	1075	1112	1144	1176	1208	1239	1271	1302	1333
CPU(%)	720P	25	27	29	31	32	34	36	38	40
	1080P	35	38	40	42	44	46	48	51	55
GPU(%)	720P	38	40	41	42	43	44	46	47	48
	1080P	43	44	45	46	48	49	52	54	56

同样，我们在以北航为场景的多视频监控系统上进行融合绘制性能的对比实验，分别以融合绘制帧速和各个融合算法中各子部分的时间占用进行展示。在反映融合绘制的性能时，不能仅用单一时刻的数值作为参考，所以在实验中给定多条场景浏览线路，将路径之上的动态数值作为融合绘制性能的参考进行比对实验。给定的路径示例图如图 9 所示，按照表 3 进行参数配置与设定。其中所有视频分辨率均为 1080P。

(a)配置1 (b)配置2 (3)配置3

图 9 融合绘制性能实验配置示例图

表 3 融合绘制实验配置参数表

配置序号/配置参数	视频数量	场景模型数量	路径关键点数量
1	1	7	9
2	1	25	9
3	5	7	9

图 10(a)~(c)采用曲线的方式，分别展示了直接投影方法，纹理贴图方法，基于图像变形的融合方法以及本文基于图片建模的融合方法的融合绘制帧数的对比结果。先通过在路径的起始点停留，之后再在此路径上循环的方法，得到不同融合方法在场景融合绘制时每一秒的绘制帧数。从曲线中可以看出以下三个结论：

①融合绘制帧速上，本文的方法≈纹理贴图方法>直接投影方法>>基于图像变形的方法；

②实时的性能，在视频数量增加或者场景模型复杂度增加的情况下，直接投影方法的投影目标会增加，则着色器对视频采样和深度检测的判断次数的也会增加。而基于图像变形的方法中图像变形带来的消耗也增加。存在的这些本质上难以克服的问题，使其不能达到实时的性能要求；

③融合效果，对于纹理贴图方法，虽然帧速较高，但其纹理采样结果会出现扭曲现象，融合结果并不适合于实际应用。

综上所述，本文提出的基于图片建模的融合方法最具实际应用价值。

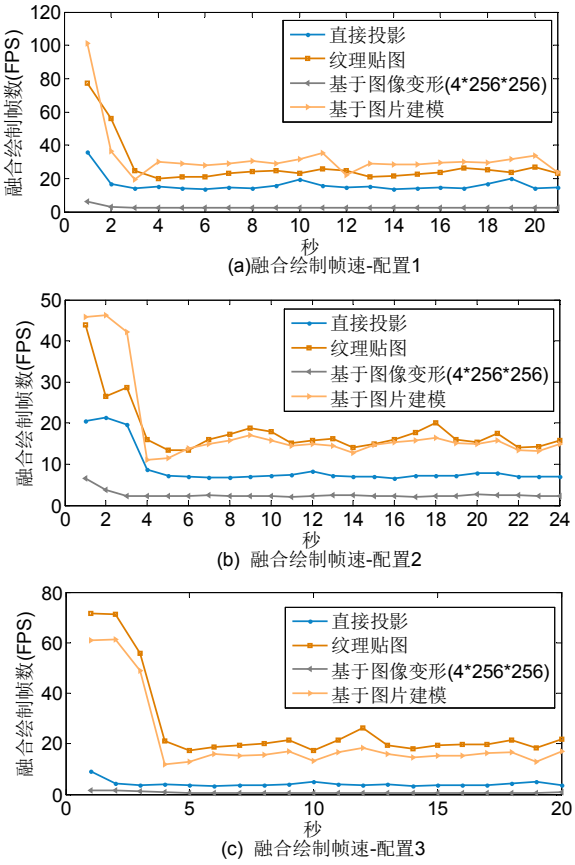


图 10 不同配置下各融合方法的绘制帧速对比

为了进一步清楚地了解各个融合方法在绘制某一帧融合图片时的时间使用情况，表 4 给出了融合绘制过程中各部分功能时间占用结果。如图 12

所示，采用柱状图对包括绘制和更新在内的时间占用结果进行可视化展示。

表 4 不同融合方法各部分功能时间(ms)占用结果表

融合方法/各个部分	渲染	裁剪	GPU	更新	事件
直接投影	15.31	6.22	15.41	0.61	0.1
纹理贴图	2.96	2.82	2.24	0.35	0.11
基于图像变形	3.02	2.63	2.22	131.89	0.11
基于图片建模	3.72	3.06	10.33	0.55	0.11

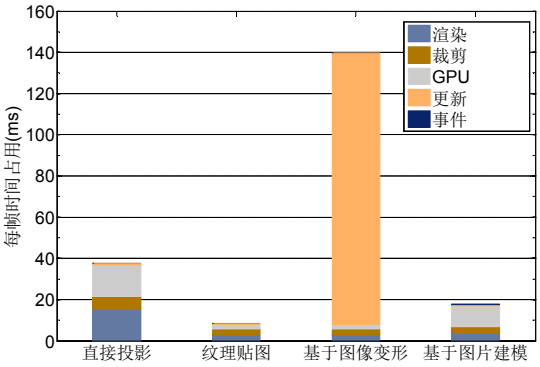


图 11 各融合方法的各部分时间占用

从图 11 中得出：(1)相对于其他两种融合方法，直接投影和本文提出的方法使用非默认渲染管线，生成纹理坐标的方式也就相对比较复杂，所以对应的 GPU 和渲染部分时间占用更多。(2)基于图像变形的方法复杂在于图像变形的处理部分，即使是低分辨率的 256*256 的基础纹理缓存也使得绘制帧数的大幅降低，所以对应的更新部分时间占用最多。(3)纹理贴图方法使用了默认渲染管线，绘制效率最高，所以该方法每个部分时间占用都相对较少，不存在瓶颈。

7 总结

本文提出并设计了一种基于视频模型的虚拟现实视频融合系统。论文首先介绍了视频模型的建立原理和其对应的文件格式，阐述了系统中视频与场景融合绘制方法。然后本文对系统的总体结构进行了介绍。在实验阶段，测试了本系统的性能，并将本文提出的虚实融合方法与国际上的虚实融合方法进行了实验对比分析，结果表明本文方法在融合绘制效果和性能上均优于其他两类方法。

参 考 文 献

[1] Hsu S, Samarasekera S, Kumar R, et al. Pose estimation, model refinement, and enhanced visualization using video[C]. IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2000, 1: 488-495.

[2] H. S. Sawhney, A. Arpa, R. Kumar, S. Samarasekera, M. Aggarwal, S. Hsu, D. Nister, and K. Hanna. Video flashlights: real time rendering of multiple videos for immersive model visualization[C]. Proceedings of the 13th Eurographics workshop on Rendering. Aire-la-Ville, Switzerland, 2002:157-168.

[3] Neumann U, You S, Hu J, et al. Augmented virtual environments(AVE): Dynamic fusion of imagery and 3d models[C]. Proceedings of IEEE Virtual Reality. IEEE Computer Society, 2003:61.

[4] Jinhui Hu. Intergrating complementary information for photorealistic representation[D]. Los Angeles: University of Southern California, 2009.

[5] Segal M, Korobkin C, Van Widenfelt R, et al. Fast shadows and lighting effects using texture mapping[C]. Conference on Computer Graphics and Interactive Techniques. 1992:249-252.

[6] Zhou Z, You J, Yang J, Zhou Y, Wu W. Method for 3D scene structure modeling and camera registration from single image, US20160249041[P], 2016.

[7] Levin D. The Approximation Power of Moving Least-Squares[M]. America: American Mathematical Society, 1998.

[8] Guillou E., Meneveaux D., Maisel E., et al. Using vanishing points for camera calibration and coarse 3D reconstruction from a single image[J]. The Visual Computer, 2000:396-410.