

MRMark: 面向三维视频融合的机器性能评测方法

孙奕飞^{1,2}, 蔡亮亮¹, 周颐^{1,2}, 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; 2. 北京大视界科技有限公司, 北京 100088)

摘要: 三维视频融合系统涉及视频和图形的计算, 对机器性能有较高要求, 但现有商用评测工具缺乏对应考虑, 会出现机器评分与实际运行性能不相符的情况。提出一种面向三维视频融合的机器性能评测方法, 并实现了评测工具 MRMark。该方法根据视频融合系统的特点设计了多组标准三维测试场景, 统计测试过程中的帧率、播放的视频路数等参数, 来计算 GPU、CPU 分数和耦合分数, 以及机器的整体性能分数。实验结果表明, 相比现有评测工具, 该工具的评测结果能更有效地检测机器进行三维视频融合计算的能力。

关键词: 性能评测; 评测工具; 虚拟场景; 虚实融合技术

中图分类号: TP311.56

文献标识码: A

文章编号: 1004-731X (2020) 07-1375-10

DOI: 10.16182/j.issn1004731x.joss.19-VR0442

MRMark: A Machine Performance Evaluation Method for Three-dimensional Video Fusion

Sun Yifei^{1,2}, Cai Liangliang¹, Zhou Yi^{1,2}, Zhou Zhong¹

(1. State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing 100191, China;

2. Beijing Bigview Technology Co., Ltd, Beijing 100088, China)

Abstract: The three-dimensional video fusion system needs the calculation of video and graphics, and requires high machine performance. The existing commercial evaluation tools lack the corresponding considerations, which may cause the inconsistency of a machine's performance from the score. A machine performance evaluation method for three-dimensional video fusion is proposed and an evaluation tool MRMark is implemented. Several sets of standard 3D test scenes are designed according to the characteristics of the video fusion system, the statistics on the frame rate, the number of video channels played and the other parameters during the test are collected to calculate the GPU, CPU and coupling score, as well as the overall performance score of the machine. The results show that, compared to the existing evaluation tools, the tool can more effectively detect the ability of the machine to perform the three-dimensional video fusion calculations.

Keywords: performance evaluation; evaluation tool; virtual scene; virtual-reality integration

引言

伴随着 5G 的成熟和视频图像的普及, 将虚拟

环境与真实环境进行匹配合成的虚实融合技术成为当下研究的热点, 并产生了多种新型应用, 如谷歌的实时街景地图 Instant Google Street View、微软的全息影像通话系统 Holoportation、体育赛事的虚拟现实回放系统 FreeD 等。视频融合是增强虚拟世界真实性的重要技术, 在信息世界中可以构建一个与真实环境同步演化的三维虚实融合



收稿日期: 2019-08-30 修回日期: 2019-11-05;
基金项目: 国家重点研发计划(2018YFB2100601);
作者简介: 孙奕飞(1996-), 男, 江苏泰州, 硕士生, 研究方向为虚实融合; 蔡亮亮(1996-), 男, 河南焦作, 博士生, 研究方向为图像匹配与融合。

环境, 增强虚拟世界的真实性程度, 使用户获得真实的可视感受。在监控领域, 三维视频融合系统由于能够整合所有监控相机画面, 并可以自由切换任意视角, 正得到越来越广泛的应用。在其基础上还可以整合各种视频分析, 如跨相机车辆跟踪、行人重识别等, 以及接入物联网数据来增强显示各种类型的标签。

三维视频融合系统涉及视频和图形的计算, 尤其是对包含大量视频流的场景, 计算任务繁重。为了达到良好的使用效果, 系统需要高性能的图形处理单元和逻辑运算单元的支持, 对系统运行环境做出适当的评估是必要的。目前对机器性能进行评估的工具已有不少, 如评测机器整体性能的评测工具鲁大师、PCMark, 评测游戏性能的工具 3DMark 和评测机器是否支持 VR 头显的 VRMark 等, 这些工具在对机器的整体性能或特定应用的评测是实用的, 若将它们的评测结果直接用于评估三维视频融合系统运行的机器, 会出现评测结果与实际用户体验不相符的结果。因此, 需要研究和设计专门用于评测三维视频融合系统的工具。

现有商用评测工具对三维视频融合系统缺乏良好的支持主要有以下两方面原因。第一是所评测的内容与三维视频融合系统所依赖的硬件资源不相符; 第二是运行的测试场景与三维视频融合系统所运行的场景相脱节, 无法准确地反映系统在运行过程中对机器资源的需求变化。

在具有三维场景测试的评测工具中, 它们的测试场景通常只有三维模型, 而三维视频融合系统对相机图像进行建模, 建立了视频模型, 与三维网格模型融合显示, 使得系统对各种硬件资源的占用情况都发生了变化。

为解决以上问题, 本文结合目前前沿的三维视频融合系统的研究, 分析系统所采用的视频融合技术原理, 提取出其消耗处理器和显卡资源不同的模块来设计场景, 这样可以更准确地获知系统对各资源的需求变化, 减少硬件之间的耦合干

扰, 从而设计出合理可靠的评分方法。在此基础上, 我们实现了一个评测工具 MRMark 便于用户使用。我们将两款现有商用评测工具与本文评测工具进行了对比分析, 实验结果表明, 本文评测工具可以更准确地检测机器对三维视频融合系统的支持程度, 具有较高的实用价值。

1 相关工作

三维视频融合系统的核心技术是解决视频与虚拟环境的融合对齐, 目前国际主流的融合技术是基于视频直接投影的方法。最早由 Stephen Hsu 等^[1]提出; H.S.Sawhney 等^[2]设计的 Video Flashlights 系统中核心方法就是将视频与三维场景通过视频直接投影的方式融合; Neumann 等^[3]提出了增强虚拟环境(AVE)的概念, 并开发了系统, 该系统中提到的视频与虚拟环境相融合的方式也是视频直接投影的思想。Hu 等^[4]提出了一种基于视频图像变形的视频与虚拟场景融合的方法。DeCamp 等^[5]将鱼眼相机视频与三维场景融合, 设计了一套用于智能家居的沉浸式系统 HouseFly。Pan 等^[6]提出了一种在虚实融合环境中进行移动目标检测与跟踪的方法, 解决了视角大范围变化造成的失真问题。邓磊等^[7]提出了一种基于三维重构的交互式摄像机位姿标定方法, 减轻了传统标定方法繁杂的人工交互工作, 并且能克服主动相机或扰动带来的影响。周颐等^[8]提出了基于视频模型的虚拟现实视频融合系统, 将图片建模技术与虚实融合技术相结合。

现有商用机器评测工具主要分为两类: 第一类是对机器的整体性能进行评测, 鲁大师和 PCMark 在这方面表现优异。鲁大师是由成都奇鲁科技有限公司开发设计的对机器各硬件设备进行评测的工具, 评测内容包括处理器性能, 显卡性能, 内存性能以及磁盘性能, 并根据 4 项的综合评分来给用户参考; PCMark 是 Underwriters Laboratories 公司针对目前办公场所的各种任务而开发的测试工具, 测试内容具体分为基础任务、生产率、数字内容创作

和游戏。这些对机器综合性能评测的工具可以让用户具体了解机器各项操作的性能,但是由于缺乏针对性,而无法给出对某一具体应用的支持程度。

第二类评测工具是检测某机器是否支持某类型应用以及其支持的程度,代表如 Underwriters Laboratories 公司开发的 3DMark, VRMark 等。3DMark 主要用于检测机器的游戏性能,其测试主要包括图形测试、物理测试以及综合测试,根据测试过程中的帧率给出对机器性能的评估,游戏场景与三维视频融合系统的场景相近,都需要依赖渲染管线,但是游戏场景更多依赖显卡性能,而三维视频融合系统更加注重处理器性能,因此 3DMark 的评分无法直接迁移到三维视频融合系统

中。VRMark 则用于检测机器对 VR 头显的支持程度以及相关性能,通过客观评分和用户体验感来作为评测依据,而由于其专业性,VRMark 的评测结果对其他应用缺乏参考价值。

以上评测工具的测试项目和评测依据如表 1 所示,它们都与三维视频融合系统的特性不相符,因此使用它们的评测结果来判断某机器对系统的支持程度是不可靠的。本文的方法直接使用三维视频融合系统设计多组测试场景,考虑场景特性,结合 CPU 占用率、GPU 占用率、视频路数和帧率等参数进行评分,实验结果表明 MRMark 的测试结果比其他评测工具更加准确。

表 1 常见评测工具评测方法及评测依据
Tab. 1 Evaluation methods and evaluation basis of common evaluation tools

评测工具	测试项目	评测依据
3DMark 11	图形场景、物理场景、综合场景	平均帧率
VRMark	基准场景、进阶场景、高阶场景; 桌面模式、头显模式	用户体验感和平均帧率
鲁大师	处理器运算(整数堆排序、IDEA 加密、Huffman 压缩、傅里叶系数、神经网络、浮点运算), 显卡测试, 内存测试(内存大小、内存运行速度), 磁盘测试(随机读写、顺序读写)	处理器运算速度、显卡平均帧率、内存大小和内存运算速度、磁盘读写速度
MRMark (本文)	纯三维场景、多视频场景、综合场景	帧率、CPU 和 GPU 占用率、打开视频路数

2 基于三维视频融合的评测方法与工具

在实际三维视频融合场景中,视频模型如图 1 箭头所指区域所示,其中三维模型渲染过程与视频融合过程对处理器与显卡的依赖程度不同,我们利用三维视频融合系统的这些特性设计了 3 组对资源消耗侧重不同的标准测试场景,分析处理器和显卡的表现和它们对系统的影响程度,进而设计出科学合理的评分方法。根据这一评分方法,设计实现了相应的评测工具 MRMark。

2.1 测试场景及相关配置

建立一个区域级别的标准三维场景,如图 2(a)所示,并在其基础上在红框区域嵌入了视

频,构建了三维视频融合系统。按照测试内容对其进行不同的配置得到三组不同的场景,分别为纯三维场景、多视频场景以及综合场景。

纯三维场景中未加载视频,仅执行三维模型的加载和渲染。规划一条漫游路径,如图 2(b)所示。本场景涵盖城市场景中常见的三维模型,如建筑、树木和路灯等,均由 3ds Max 软件制作而成,总计 120K+个面片、360K+个顶点。机器在渲染模型时占用较多的显卡资源,而较少占用处理器资源,因此本场景可作为计算 GPU 分数的依据。

多视频场景中嵌入了 38 路本地相机视频,其中有 6 路分辨率为 1 308×736 的室外枪机视频,18 路分辨率为 1 536×1 536 的室内鱼眼相机视频,

14 路分辨率为 1 920×1 080 的室内外枪机视频。视点在飞到距离某相机一定角度与高度的位置时，系统对该相机视频进行解码播放。

为减少大量三维模型渲染对资源的消耗，多视频场景的漫游路径主要围绕相机视频区域，如图 2(c)所示。在漫游过程中统计随着相机视频路数的增加引起的资源消耗的变化情况，漫游结束后固定视点位置，测试机器能够同时打开的最大视频路数。本场景中系统对视频进行软解码，受处理器性能的影响更大，因此可用来计算 CPU 分数。



图 1 三维视频融合系统视频模型
Fig. 1 Video model of the 3D video fusion system

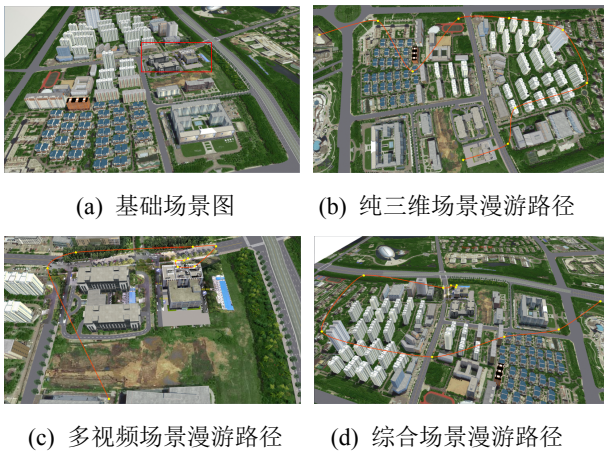


图 2 三维视频融合系统场景图及漫游路径

Fig. 2 3D video fusion system scene graphs and roaming paths

综合场景配置了 19 路本地相机视频，其中各分辨率的视频路数为多视频场景对应分辨率的视频路数的一半。我们规划的漫游路径由纯三维模型区域逐渐过渡到视频区域，如图 2(d)所示。本

场景考虑了大量三维模型和相机视频的情况，更加接近用户所需的实际场景，可以作为处理器与显卡耦合分数的依据。

为了能够充分发挥出机器的性能，我们将显卡的垂直同步功能关闭，使得场景运行时的帧率不受限制。

2.2 评分方法

三维视频融合系统依赖于多路视频投影渲染和解码，同时占用大量处理器和显卡资源，处理器和显卡紧密耦合，共同影响帧率，因此本文在公式中引入 CPU 和 GPU 占用率、打开视频路数等参数，尝试对两者解耦合，使结果更加准确。

我们采集场景运行过程中的性能数据变化，分别计算 CPU 分数、GPU 分数和耦合分数这 3 项子分数，将它们进行加权调和平均得到 MRMark 总分。

基于对该三维视频融合系统的统计分析，我们发现相比显卡，系统运行效果受处理器性能的影响更大，因此 CPU 分数的权重最高，GPU 分数和耦合分数权重相同。MRMark 总分计算方法如公式(1)所示。

$$S_{\text{MRMark}} = 1 / (0.1 / S_{\text{gpu}} + 0.8 / S_{\text{cpu}} + 0.1 / S_{\text{coupling}}) \quad (1)$$

GPU 分数：纯三维场景中，帧率主要受显卡的渲染能力影响，与渲染的面片数成负相关。为了减小处理器渲染线程处理能力对帧率的影响，将 CPU 和 GPU 占用率引入到 GPU 分数公式中是必要的。我们在每个漫游关键点处计算一次分数，最后取所有关键点处分数的平均值作为 GPU 分数，公式如下：

$$S_{\text{gpu}} = c_{\text{gpu}} \times \frac{1}{n} \sum_{i=1}^n \frac{F_i \times \sqrt{100 - C_i}}{G_i} \quad (2)$$

式中： c_{gpu} 为 30，代表缩放系数； n 为漫游关键点个数，每个场景都设置了 10 个漫游关键点； F_i 、 C_i 、 G_i 分别为视点飞到漫游关键点 i 时的帧率、CPU 占用率和 GPU 占用率。

CPU 分数：多视频场景中，系统进行三维视

频融合的过程主要包括视频解码、视频数据传输和显卡渲染, 融合效率主要取决于前两步。系统通过多线程解码视频, 多线程能力越强, 意味着可以同时打开更多的视频。解码后的视频数据通过渲染线程上传给显卡, 通过纹理投影和阴影投影的方式使视频与三维场景相融合, 达到增强虚拟环境的效果, 这一纹理上传过程占用大量时间, 因此处理器的单线程能力在很大程度上决定了系统运行的流畅程度。在该场景收集的各项数据中, 同时打开的最大视频路数是一个非常重要的指标, 刻画了处理器多线程处理能力的上限, 以之作为多线程分数; 漫游测试过程中打开的视频数量相对较少, 帧率主要受限于处理器的单线程渲染能力, 因此将该过程的平均帧率作为单线程分数。CPU 分数公式如下:

$$S_{cpu} = c_{cpu} \times \sqrt{V_{max} \times F_{cpu}} \quad (3)$$

式中: c_{cpu} 为 70, 代表缩放系数; F_{cpu} 为多视频场景漫游过程的平均帧率; V_{max} 为测试中同时打开的最大视频路数。

耦合分数: 综合场景是对处理器和显卡耦合能力的测试, 耦合能力越好, 系统运行时更稳定。平均帧率可以很好地体现出它们相互耦合的能力, 但由于对视频的处理是系统的关键技术所在, 所以帧率受处理器影响更大。耦合分数如公式(4)所示:

$$S_{coupling} = c_{coupling} \times F_{coupling} \quad (4)$$

式中: $c_{coupling}$ 为 10, 代表缩放系数; $F_{coupling}$ 为综合场景的平均帧率。

公式(2)~(4)中, 缩放系数 c_{gpu} , c_{cpu} , $c_{coupling}$ 均为经验常数, 是由不同机器测试结果归纳分析得到, 目的是使机器的 3 项子分数尽可能趋近一致。

2.3 工具设计与实现

在上述评分方法的基础上, 我们设计了评测工具 MRMark, 其执行流程如图 3 所示。

MRMark 首先会进行机器基础信息的检测,

检测用户机器是否达到运行三维视频融合系统的最低要求, 并在图 4(a)的界面中展示给用户。检测内容包括: CPU 逻辑核心数是否达到 4 核、当前使用的显卡品牌是否为 NVIDIA、显示器分辨率是否达到 800×600、是否安装图形驱动、显卡是否支持 Shader Model 4.0 或以上版本等等。若以上要求有一项不满足, 则判定该机器为未达标。

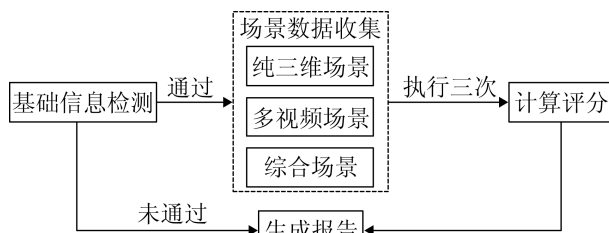


图 3 MRMark 执行流程图
Fig. 3 MRMark execution flowchart



(a) MRMark 基本信息界面



(b) MRMark 结果界面

图 4 MRMark 的界面设计
Fig. 4 Interface design of MRMark

在上述要求都满足后, 用户可以进行后续场景测试, 使用的场景即为 2.1 节介绍的 3 个场景。为减小场景运行可能不稳定而产生的误差, 我们对每个场景运行 3 次, 取 3 次得分的算数平均值作为每项子分数的最终结果。MRMark 总分和 3 项子分数将显示在图 4(b)界面上, 同时给出对用户机器的评价, 告知用户机器表现不佳时的性能瓶颈所在, 并提供相应建议。用户可以在检测结束后保存检测结果以获取更加详细的检测信息。

在实际使用过程中, 我们为 CPU 分数和 GPU 分数分别设定阈值, 若某机器至少一项没有达到阈值, 则认为该机器未达到运行三维视频融合系统的要求, 若全部达标, 则给出能在该机器上流畅运行的场景类型。

MRMark 已公开发布下载, 供业内参考。网址为 <http://prw.vr3i.com>。

3 实验结果及分析

3.1 实验环境

本文选取 6 台硬件配置不同的主流消费级计算机进行测试, 其中处理器型号包括英特尔酷睿系列和志强系列, 显卡型号包括英伟达 GTX 系列和 RTX 系列, 具体配置信息如表 2 所示。测试时关闭其他正在使用的进程, 减小对资源的消耗。本文评测工具不考虑超频的情况。为验证测试结果的准确性, 这些机器均达到了运行三维视频融合系统的要求。

3.2 不同机器测试与结果分析

实验包括两部分, 分别是不同机器测试与结果分析、不同评测工具测试结果对比。从不同场景下机器各项参数的数据对比可以分析出不同机器处理器和显卡之间的性能差异。

图 5(a)~(c)展示了纯三维场景漫游路径各关键点处的参数值。由于场景没有采用数据分页机制动态调度三维模型的加载和卸载, 也未对三维模型进行裁剪, 处理器不需要重新设置渲染状态, 且漫游

路径中需要计算的顶点数据较少, 所以处理器的 CPU 占用率较为平稳, 而显卡的 GPU 占用率和帧率会随着漫游路径中顶点和面片数的变化而产生较大波动, 因此该漫游路径的设计是合理的, 可以使处理器的影响降至最低。但我们不能因此忽略处理器的影响, 从图 5(c)中可以看出, 机器 c 的表现较为反常, 其显卡型号与机器 e 相同, 帧率与 GPU 占用率却远低于机器 e, 这是受处理器的单线程能力所限制, 处理器与显卡之间过慢的数据交互速度减少了显卡对模型的渲染次数, 因此公式(2)以 CPU 和 GPU 占用率来缩小拥有同型号显卡的机器在帧率值上的差距, 使 GPU 分数更为接近。

多视频场景在漫游测试中的参数数据如图 5(d)~(f)所示, 随着系统请求打开视频数量的增加, 处理器需要调用更多线程进行视频数据的计算, CPU 占用率的增长呈加速趋势; 同时, 由于视频帧需要作为纹理投影到三维场景中, 解码的视频帧越多, 纹理提交到显存的速度越慢, 当处理器完成纹理更新和绘制指令发送所占时间超过显卡完成栅格化等绘制操作所占时间时, 显卡陷入空闲等待状态, GPU 占用率开始下降, 所以在打开视频路数较少时帧率可以体现出处理器单线程性能强弱。

在三维视频融合场景中, 帧率反映了画面的流畅度, 不流畅的画面会给用户带来不好的视觉体验, 我们认为帧率的重要性要高于最大路数, 所以在公式(3)中单线程分数值高于多线程分数值, 对 CPU 分数的影响更大。表 3 展示了不同机器在漫游过程中的平均帧率和能够同时打开的最大视频路数, 机器 f 表现最好, 帧率最高, 且能够打开全部 38 路视频; 机器 c 最大能打开 21 路视频, 其拥有双路处理器共 32 线程, 理论上可以打开更多视频, 但其渲染线程无法处理更多的数据, 限制了多线程能力, 使得其在最大路数多于机器 a 的情况下, 帧率远低于机器 a 而 CPU 分数也低于机器 a, 因此这类服务器处理器不一定适用于三维视频融合系统。

表 2 机器配置信息
Tab. 2 Configuration information for machines

机器编号	CPU	GPU
a	Intel(R) Xeon(R) CPU E5-1620 v2 @ 3.70GHz (4 核 8 线程)	NVIDIA GeForce GTX 760
b	Intel(R) Xeon(R) CPU E5-1620 v3 @ 3.50GHz (4 核 8 线程)	NVIDIA GeForce GTX 1060 (6GB)
c	Intel(R) Xeon(R) CPU E5-2620 v4 @ 2.10GHz (8 核 16 线程) (×2)	NVIDIA GeForce GTX 1070
d	Intel(R) Core(TM) i7-6700 CPU @ 3.40GHz (6GB) (4 核 8 线程)	NVIDIA GeForce GTX 1060 (6GB)
e	Intel(R) Core(TM) i7-8700 CPU @ 3.20GHz (6 核 12 线程)	NVIDIA GeForce GTX 1070
f	Intel(R) Core(TM) i7-8700K CPU @ 3.70GHz (6 核 12 线程)	NVIDIA GeForce RTX 2070

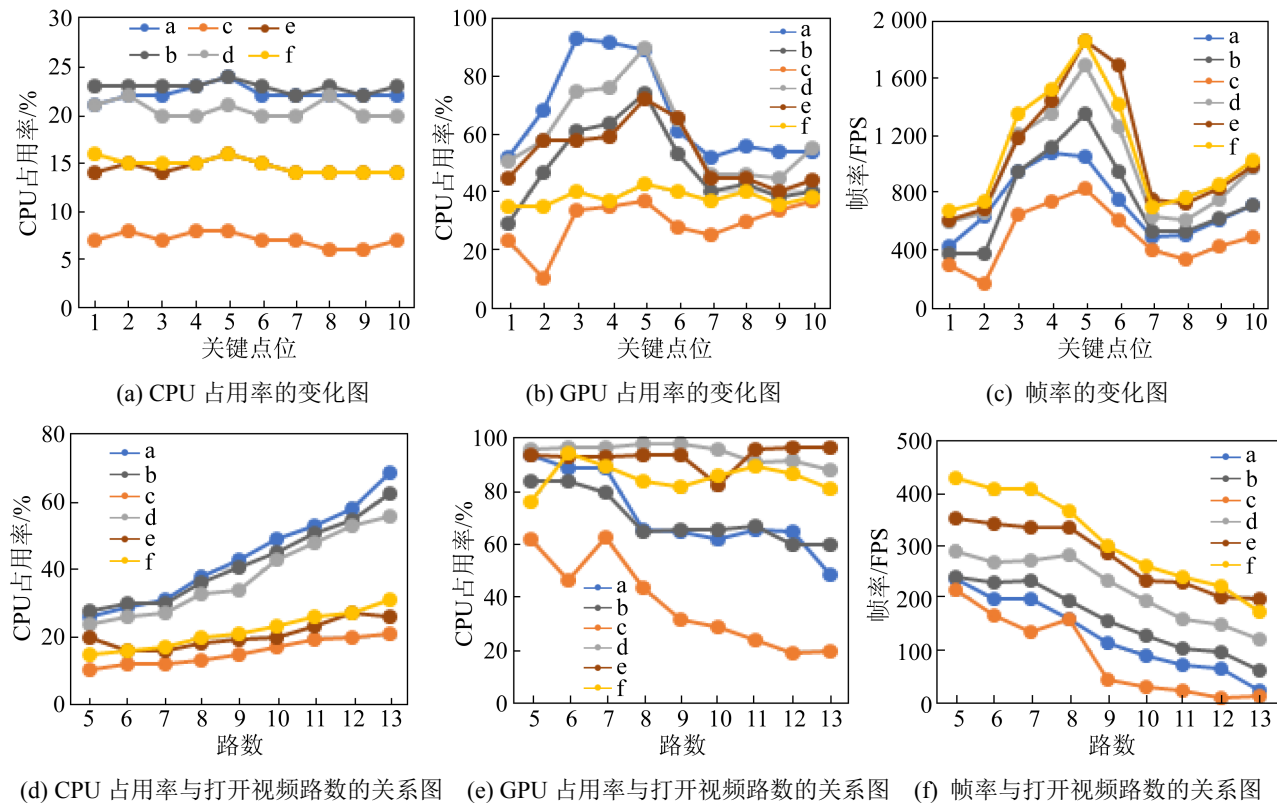


图 5 不同机器的参数对比图

Fig. 5 Comparison of parameters of different machines

表 3 不同场景的平均帧率及多视频场景最大打开视频路数

Tab. 3 Average frame rates for different scenes and maximum number of open channels for multi-video scenes

机器编号	纯三维场景 平均帧率	多视频场景 平均帧率	最大路数	综合场景 平均帧率
a	745.0	125.0	18	413.6
b	768.7	152.7	20	394.3
c	520.0	86.7	21	303.3
d	868.0	218.7	22	566.7
e	1 108.0	266.3	27	626.0
f	1 120.7	305.3	38	639.0

综合场景是纯三维场景和多视频场景的结合, 故不再列出参数对比图。不同机器在该场景中的平均帧率见表 3, 可以发现, 在处理器和显卡消耗的资源同时变化的情况下, 帧率受处理器影响更大, 其中机器 a 和 b 处理器性能相近, 而机器 a 处理器和显卡耦合性更好, 所以帧率反而略高于机器 b。

表 4 展示了每台机器的各项分数和打开最大路数。观察 CPU 分数、GPU 分数和 MRMark 分数发现, 面向三维视频融合系统, 这 6 台机器的显

卡性能排序为 $f > e > c > d > b > a$ ，处理器性能和整体性能排序均为 $f > e > d > b > a > c$ ，机器 b 与 d 、 c 与 e 的显卡型号相同，但受到处理器影响产生了一定的分数差异。由于机器运行三维视频融合系统的表现在很大程度上受到用户主观因素的影响，因此本文调研了 20 名志愿者观看这 6 台机器运行 MRMark 的感受，并让他们进行排序。根据反馈，所有志愿者的排序结果均与本文的整体性能排序结果相同，而在不考虑帧率和最大路数的情况下，18 人认为机器 e 和 f 的表现几乎分辨不出差距，都能流畅运行测试场景，2 人认为机器 f 比 e 表现略好，以上统计结果表明本文的评分方法是可靠的。

表 4 不同机器的各项分数
Tab. 4 Scores of different machines

机器编号	CPU 分数	GPU 分数	耦合分数	最大路数	MRMark 分数
a	3 289	2 819	4 137	18	3 302
b	3 777	3 787	3 943	20	3 794
c	2 986	4 947	3 033	21	3 114
d	4 818	4 156	5 657	22	4 813
e	5 936	5 474	6 260	27	5 917
f	7 473	7 730	6 390	38	7 373

3.3 不同评测工具测试结果对比

为了体现设计 MRMark 工具对三维视频融合系统进行测试的必要性和可靠性，我们将 MRMark 与市面上常用的 3DMark 11、鲁大师两款评测工具进行了对比。3DMark 11 采用了 4 个图形场景、一个物理场景和一个综合场景进行测试，

评分公式见附录，可以发现其只依赖了平均帧率这一个参数。鲁大师对处理器性能的评分是通过多种运算的时间来计算得到，对显卡性能的评分也只根据平均帧率求得。这两款工具对三维场景的考量都较为单一，无法准确反映系统的资源需求，而本文的评分方法利用帧率、视频路数、CPU 和 GPU 占用率进行综合考量，可以更准确地进行评估。

我们在相同机器上进行测试验证，测试结果如表 5 所示，其中鲁大师未测试内存分数和磁盘分数。从表中可以发现，对于显卡性能，3DMark 11 的排序为 $f > c > e > b > d > a$ ，鲁大师的排序为 $f > e > c > b > d > a$ ，这与我们的排序结果大致相同，也与显卡的理论性能相符；对于处理器性能，3DMark 11 的排序为 $f > e > d > b > c > a$ ，鲁大师的排序为 $c > f > e > d > a > b$ ，它们对低频多线程类型的处理器的评分都不符合系统运行多视频场景的帧率表现；对于机器的整体性能，3DMark 11 的排序为 $f > e > c > b > d > a$ ，更加侧重显卡的性能，鲁大师的排序为 $f > c > e > d > b > a$ ，总分只是对处理器分数与显卡分数的简单相加，没有侧重，而三维视频融合系统受处理器性能影响更大，它们的排序结果均与机器的实际表现不符。基于以上分析比较，本文的评测方法基于三维视频融合系统场景进行测试和评分，弥补了其他评测工具的不足，可以比较准确地给用户的机器做出评价，具有更好的实用性。

表 5 不同评测工具的测试结果
Tab. 5 Test results for different evaluation tools

机器编号	3DMark 11				鲁大师		
	物理分数	图形分数	结合分数	3DMark 总分	CPU 分数	GPU 分数	总分
a	8 298	3 165	3 542	3 165	70 658	92 920	163 578
b	8 674	5 397	6 831	5 733	68 485	162 967	231 452
c	8 658	7 707	7 890	7 505	145 434	205 954	351 388
d	10 287	5 318	6 648	5 707	76 959	154 637	231 596
e	12 864	7 147	9 126	7 653	123 035	223 286	346 321
f	13 682	9 444	12 124	10 248	123 960	278 940	402 900

4 结论

本文以三维城市场景为原型构建三维视频融合系统, 并设计出 3 组依赖硬件资源不同的三维场景, 依据每个场景的特点, 设计出相应的 GPU 分数、CPU 分数和耦合分数的计算方法。在此基础上, 我们实现了评测工具 MRMark。实验阶段, 我们在多台机器上进行测试比较, 分析部分机器表现较差的原因, 并将结果与其他性能评测工具对比, 结果表明本文的评测方法在评测用于运行三维视频融合系统的机器的整体性能时更加符合用户的直观感受, 也能更准确地判断出机器的性能瓶颈所在之处。本文的不足之处在于没有考虑内存、磁盘等其他硬件设备对三维视频融合系统的影响, 在未来的工作中, 我们会尝试将其他因素的影响添加到评分方法中。

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. USA: IEEE, 2000, 1: 488-495.
- [2] Sawhney H S, Arpa A, Kumar R, et al. Video flashlights: real time rendering of multiple videos for immersive model visualization[C]. Proceedings of the 13th Eurographics workshop on Rendering. Goslar, Germany: Eurographics Association, 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. USA: IEEE Computer Society, 2003: 61.
- [4] Hu J H. Integrating complementary information for photorealistic representation[D]. Los Angeles: University of Southern California, 2009.
- [5] Decamp P, Shaw G, Kubat R, et al. An immersive system for browsing and visualizing surveillance video[C]. Acm International Conference on Multimedia. USA: ACM, 2010.
- [6] Pan C, Chen Y, Wang G. Virtual-Real Fusion with Dynamic Scene from Videos[C]. 2016 International Conference on Cyberworlds (CW). USA: IEEE Computer Society, 2016.
- [7] 邓磊, 陈宝华, 赖伟良, 等. 三维监控系统中基于三

维重构的交互式标定[J]. 电子学报, 2017, 45(3): 527-533.

Deng Lei, Chen Baohua, Lai Weiliang, et al. Interactive Calibrating 3D Surveillance System based on 3D Reconstruction[J]. Acta Electronica Sinica, 2017, 45(3): 527-533.

[8] 周颐, 孟明, 吴威, 等. 基于视频模型的虚拟现实视频融合系统[J]. 系统仿真学报, 2018, 30(7): 2550-2257.

Zhou Yi, Meng Ming, Wuwei, et al. Virtual-reality Video Fusion System Based on Video Model[J]. Journal of System Simulation, 2018, 30(7): 2550-2557.

附录 A:

1. 3DMark 11 各项分数的计算公式如下。

(1) 图形分数:

$$S_{\text{graphics}} = C_{\text{graphics}} \frac{4}{\frac{1}{F_{\text{gt1}}} + \frac{1}{F_{\text{gt2}}} + \frac{1}{F_{\text{gt3}}} + \frac{1}{F_{\text{gt4}}}}$$

式中: $C_{\text{graphics}}=230$, F_{gt1} , F_{gt2} , F_{gt3} , F_{gt4} 分别为 4 个图形场景的平均帧率。

(2) 物理分数:

$$S_{\text{physics}} = C_{\text{physics}} F_{\text{physics}}$$

式中: 缩放系数 $C_{\text{physics}}=315$, F_{physics} 为物理场景的平均帧率。

(3) 综合分数:

$$S_{\text{combined}} = C_{\text{combined}} F_{\text{combined}}$$

式中: 缩放系数 $C_{\text{combined}}=215$, F_{combined} 为综合场景的平均帧率。

(4) 3DMark 总分:

$$S_{\text{3DMark}} = \frac{\frac{W_{\text{graphics}}}{S_{\text{graphics}}} + \frac{W_{\text{physics}}}{S_{\text{physics}}} + \frac{W_{\text{combined}}}{S_{\text{combined}}}}{\frac{W_{\text{graphics}}}{S_{\text{graphics}}} + \frac{W_{\text{physics}}}{S_{\text{physics}}} + \frac{W_{\text{combined}}}{S_{\text{combined}}}}$$

式中: $W_{\text{graphics}}=0.75$, $W_{\text{physics}}=0.15$, $W_{\text{combined}}=0.1$, 为各项分数在总分中所占的权重。

所有公式都只依赖了平均帧率这一个参数, 无法准确反映出三维视频融合系统对硬件资源的需求。

2. PCMark 10 各项分数的计算公式如下。

(1) 基础任务测试分数:

$$S_e = \sqrt{E_1 E_2 E_3}$$

式中: E_1 , E_2 , E_3 分别为程序启动、网页浏览及视频会议的分数。

(2) 生产力测试分数:

$$S_p = \sqrt{P_1 P_2}$$

式中: P_1 , P_2 分别为编写文档及电子表格的分数。

(3) 数字内容创作测试分数:

$$S_d = \sqrt{D_1 D_2 D_3}$$

式中: D_1 , D_2 , D_3 分别为图片编辑、视频编辑及渲染和可视化的分数。

(4) PCMark 总分公式:

$$S_{\text{PCMark}} = K \sqrt{S_e S_p S_d}$$

式中: 缩放系数 $K=0.717$ 。

3. VRMark 分数的计算公式如下:

$$S_{\text{VRMark}} = \text{average}_{\text{FPS}} \times \text{score}_{\text{Multiplier}}$$

式中: $\text{average}_{\text{FPS}}$ 为测试场景的平均帧率, 缩放系数 $\text{score}_{\text{Multiplier}}=5\ 000/109$ 。