

虚拟现实与交互技术

## 虚实融合监控场景中的漫游路径规划方法

周 颐, 解佳琦, 吴 威, 周 忠\*

<sup>1)</sup> (北京航空航天大学虚拟现实技术与系统国家重点实验室 北京 100191)  
(zz@buaa.edu.cn)

**摘要:** 为了寻找不同视频之间的转换道路, 提升观察者的方向感, 提出一种虚实融合场景中的自动路径规划方法, 首先提出一种视点评价方法, 考虑在视频投影中的图像畸变带来的图像质量问题; 接着设计了场景的视点采样方法和道路图, 并基于视点评价结果为场景道路图的边定义权重; 在确定视频的访问顺序之后, 计算生成权重代价最小视点路径. 在 4 个视频监控系统中验证了文中提出的视点评价方法和路径生成方法的有效性, 证明其能够提升用户的虚拟观测体验.

**关键词:** 纹理畸变; 视点评估; 路径规划; 虚实融合技术  
中图法分类号: TP391.41

## Path Planning for Virtual-reality Integration Surveillance System

Zhou Yi<sup>1)</sup>, Jie Jiaqi<sup>1)</sup>, Wu Wei<sup>1)</sup>, and Zhou Zhong<sup>1)</sup>\*

<sup>1)</sup> (State key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing 100191)

**Abstract:** We seek to develop an automatic path planning method to find the way of the transition and willing to help to improve the observer's sense of direction. We firstly propose a viewpoint quality metric which evaluates the quality of viewpoints by measuring the image distortion in video projection. Then we construct a graph for the scene whose nodes denote sample viewpoints and edges denote virtual camera movements. A camera path can be generated automatically using this graph. Finally we apply our method in video surveillance and demonstrate that it improves observers' visual experience.

**Key words:** texture distortion; viewpoint evaluation; path planning; virtual-reality integration

### 1 相关工作

虚实融合技术可以将虚拟三维模型和真实的视频图像融合到统一的位置坐标系下, 在三维可视化的一体监控系统中有广泛应用<sup>[1-2]</sup>. 而增强

虚拟环境是以纹理投影为基础的虚实融合技术, 可以在三维上下文空间中以任意视角展示多路图片或者视频流<sup>[3]</sup>. 此外虚实融合技术还有将离散全景图像来拼成成条带式街景视图的谷歌街景视图<sup>[4]</sup>, 以及使用互联网照片集重建场景集合并自由浏览的照片旅游<sup>[5]</sup>等.

收稿日期: 2017-03-22; 修回日期: 2017-10-17. 基金项目: 国家“八六三”高技术研究发展计划(2015AA016403); 国家自然科学基金(61572061, 61472020). 周 颐(1988-), 男, 博士研究生, CCF 会员, 主要研究方向为三维重建、虚实融合; 解佳琦(1992-), 男, 硕士研究生, CCF 会员, 主要研究方向为路径规划、虚实融合; 吴 威(1961-), 男, 博士, 教授, 博士生导师, CCF 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 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.

目前在这些虚拟融合场景的漫游中存在诸多问题,比如用户控制虚拟视角操作复杂,需要更多人为交互;场景中存在画面质量较差的视点,用户往往会浏览至此,如何避免成为难题.以前的虚实融合系统<sup>[1,2,5]</sup>中并没有考虑在相机视点之间进行合理的转移的问题,如何设计这样的一条路径成为提升用户使用体验的关键.

面向虚实融合技术,本文分析了视频画面质量不佳的原因.提出了一种全新的视点评价方法,首次考虑在视频投影中的图像畸变带来的图像质量问题.接着设计了场景的视点采样方法和道路图,并基于视点评价结果为场景道路图的边定义权重.在确定视频的访问顺序之后,计算生成权重代价最小视点路径.

### 1.1 视点质量评价方法

视点的质量会影响观察者的视觉体验.一个好的视点往往被认为是观察者能够接受到更多的信息<sup>[6]</sup>. Adrian<sup>[7]</sup>等总结了多个常用的视点评价方法,包括视点熵<sup>[8]</sup>,网格显著度<sup>[9]</sup>,侧影轮廓稳定性<sup>[10]</sup>等.然而这些方法只考虑三维模型和构图画面等因素,不能有效地评判模型纹理具有畸变的场景视点质量.这主要是因为这些方法没有考虑纹理畸变对整个视点的质量带来的影响.因此,增强虚拟环境系统需要一种对纹理畸变评价的视点质量估计方法.

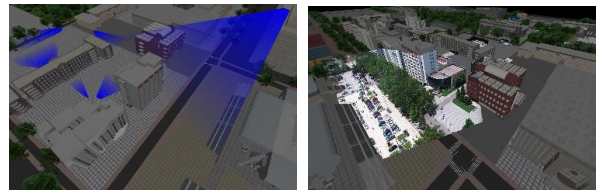
观察者在观察增强虚拟环境中的视频投影模型时,看到纹理畸变的情况是由模型深度和纹理深度不匹配引起的.因此,本文考虑从深度出发来考虑视点质量.估计单幅图像的深度是计算机视觉里一个经典任务.文献<sup>[11-13]</sup>等基于几何先验的方法在深度测试集上有着相当成功的结果.近些年来,单幅图像深度估计采用深度神经网络的结构,得到长足的进步<sup>[14-16]</sup>. Fayao<sup>[15]</sup>等将深度估计转换为一个深度连续的 CRF 模型学习的问题,不需要对场景有任何几何先验. Roy<sup>[16]</sup>等结合随机森林和深度卷积网络提出用神经随机森林来解决深度估计的问题,得到了同样优秀的结果.

### 1.2 路径规划方法

路径规划问题在计算机图形学里被称作相机控制.大多数路径规划方法将场景建模为一个图结构,其节点表示场景中的路径驻留点,边表示可行的路径.因此,路径规划问题被转化为图搜索的问题.概率路径图方法(probabilistic roadmaps, PRM)<sup>[17]</sup>是路径搜索中的一个常用方法,其方法分为图构建步和查询步.在图构建步中,起始的空图结构被重复添加随机节点,即连接附近节点.在查询步中,搜索图生成最短路径.研究者根据选择节点方式的不同,提出了不同的路径生成方法<sup>[18-19]</sup>. Ranon<sup>[18]</sup>等使用粒子群方法选择视点,但耗时太多. Oskam<sup>[19]</sup>等使用不同半径的多重球填充场景空区域来避免漫游碰撞,进而在球的相交区域采样视点.该方法适应于封闭室内场景,不能直接适用开放的室外空间.本文针对的虚拟场景包含三维建筑模型,设计出的路径希望在漫游中避免与其碰撞,同时也能产生过渡采样视点.

## 2 自由视点下的视频画面畸变

本文首先分析视频纹理畸变的产生原因.本文针对的监控虚实融合场景由虚拟场景和实时监控视频组成.对每个视频,本文选取某一帧图像进行交互式建模<sup>[20]</sup>,生成视频模型,然后将其注册到虚拟场景上.实时的视频图像将会投影到视频模型上,即视频模型直接使用视频图像作为纹理.观测者可以在虚拟环境中观测到不同地点的三维视频.典型的虚实融合场景如图 1 所示.



a. 标识相机视域的虚拟环境 b. 视频投影融合的虚拟环境

图 1 虚实融合场景示意

受限于建模精度问题,视频模型的纹理对应的真实深度和建模深度会出现不一致的情况.当偏离真实相机视点观看视频模型时,深度不一致

反映到图像平面上就会出现视频纹理歪斜、拉伸、撕裂等畸变现象. 这一现象的产生原因示意图如图 2 所示, 而典型的歪斜和拉伸畸变示意如图 3 所示. 扭曲现象越严重, 观察者看到图像上的像素投影偏差越大.

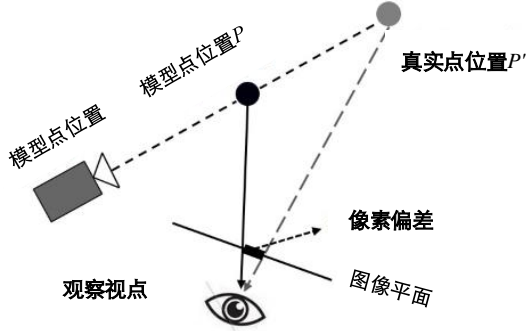


图 2 画面畸变产生原因示意



a. 真实相机视点下的画面 b. 虚拟相机视点下的画面

图 3 视频画面畸变现象

### 3 基于视频画面畸变的视点质量评价方法

基于以上的视频纹理畸变产生的原因和现象, 本文将使用累积像素投影相对误差来衡量视频纹理的畸变, 进而生成视点质量评价指标. 本文使用深度学习方法估计出的深度<sup>[15]</sup>作为参考深度. 其与视频模型的深度进行比较, 对应像素点根据深度分别投影到其图像平面上的像素点误差, 反应了扭曲现象的严重程度. 扭曲越严重, 误差越大. 本文由深度生成点云模型, 其与视频模型的对比效果如图 4 所示. 图 4a 为原始视频图片, 图 4b 为在某视点下观看视频模型的效果, 图 4c 为在相同视点下观看点云模型的效果, 图 4d, 4e, 4f 分别为图 4a, 4b, 4c 的局部放大. 从图中可以发现, 点云模型是比视频模型更精确的模型, 但是在观察点

云模型时会发生像素拉伸的现象. 在图 4d, 4e, 4f 中, 红、绿、粉、蓝四种颜色的采样点表示进行视点评价时选取的采样点, 通过计算这些采样点在屏幕上的距离差来衡量画面的畸变程度.

视点  $v$  下对单个视频模型  $V$  的画面畸变程度如式(1)计算:

$$D(v, V) = \frac{\sum_{p \in N(v, W(V))} |\mathbf{M}(v) \times R_c(p) - \mathbf{M}(v) \times R_v(p)|}{|N(v, W(V))|} \quad (1)$$

在式(1)中,  $W(V)$  为视频模型  $V$  的采样像素集合,  $N(v, W(V))$  是  $W(V)$  在视点  $v$  可见的采样点子集.  $R_c(p)$  是视频模型中像素  $p$  的空间位置,  $R_v(p)$  是点云模型中像素  $p$  的位置,  $\mathbf{M}(v)$  是视点  $v$  对应的空间到屏幕坐标的转换矩阵.  $R_c(p)$  可以直接由深度  $D_c(p)$  计算得到  $R_c(p) = I(p) + r(p) \times D_c(p)$ . 其中  $r(p)$  是由相机视点往图像点  $p$  的射线.

深度估计得到的像素深度取值范围是  $[0, 1.0]$ , 需要与视频模型中的像素深度的单位进行统一才可以进行比较. 图像中天空的深度应该是无穷远, 因此本文先去除天空以外的所有像素, 再将图像中这些像素深度的最大值和最小值分别映射到视频模型中这些像素深度的最大值和最小值, 其他数值使用线性插值的方式映射, 映射方法如式(2)所示.

$$D_c(p) = \frac{D_{v, \max} - D_{v, \min}}{D_{l, \max} - D_{l, \min}} \times (D_l(p) - D_{l, \min}) + D_{v, \min} \quad (2)$$

其中,  $D_c(p)$  为映射后的像素点  $p$  的深度,  $D_l(p)$  为使用深度估计得到的图像中像素点  $p$  的深度值,  $D_{v, \max}$  和  $D_{v, \min}$  分别为视频模型中天空以外的像素深度的最大值和最小值,  $D_{l, \max}$  和  $D_{l, \min}$  分别为图像深度中天空以外像素深度的最大值和最小值.



图 4 视频纹理畸变度量.

对单个视频模型的视点评价方法如式(3)计算:

$$s(v, V) = \left( 1 - \max \left\{ \frac{D(v, V)}{L}, 1 \right\} \right) \times P(v, V) \quad (3)$$

其中  $P(v, V)$  是视频模型  $V$  在视点  $v$  下的投影面积占屏幕的比例,  $L$  表示可接受的最大像素偏差距离, 与屏幕分辨率有关, 本文取为屏幕对角线长度的五分之一. 该式中的第一项本文记为  $D'(v, V)$ , 为对  $D(v, V)$  截断后的补值.

对于存在多个视频模型  $\{V_i\}$  的视点来说, 本文以式(4)计算:

$$s(v) = \sum_i \left( 1 - \max \left\{ \frac{D(v, V_i)}{L}, 1 \right\} \right) \times P(v, V_i) \quad (4)$$

通过这种计算方法, 将视频模型的畸变程度归一化为  $[0, 1]$  的视点质量评价. 然而对于场景中的虚拟建筑模型, 本文采用传统的视点质量评价方法, 参考视点熵<sup>[7]</sup>方法, 本文采用式(5)计算:

$$s'(v) = - \sum \frac{P(v, N_i)}{r} \lg \left( \frac{P(v, N_i)}{r} \right) \quad (5)$$

其中  $s'(v)$  表示视点评价中三维网格模型的贡献,  $P(v, N_i)$  表示在视点  $v$  下, 三维网格模型  $N_i$  的投影面积,  $r$  表示屏幕的大小. 这里的三维虚拟模型评价方法可以使用其他方法替代.

对整体视点, 本文有式(6)所示的加权评价方法:

$$S(v) = \alpha \times s(v) + (1 - \alpha) \times s'(v) \quad (6)$$

$\alpha$  为经验常量参数, 表示视频模型在整体视点评价中的重要性. 一般, 本文将其设置为 0.66.

## 4 漫游路径规划

本文将自动生成一条能够浏览所有视频模型的路径, 并希望浏览时能够在避免画面被虚拟模型完全遮挡的条件下尽可能看到更优的视频画面. 在本文的系统中, 视点  $v$  由 7 个实值参数定义, 其中 3 个参数定义虚拟相机位置, 3 个参数定义虚拟相机视点目标, 剩余的一个参数定义绕  $x$  轴旋转的翻滚角. 视场和画面长宽比都是固定的. 在 7 参数的自由空间中进行视点采样, 会有巨大的搜索空间. 如果本文采用传统 PRM 方法, 一方面路径图中将有过多的顶点, 计算代价过高, 不利于进一步的路径生成, 另一方面, 会进入过多的无意义采样视点, 比如远离视频模型的俯视视点, 实际运行中观测者不会经过该视点. 因此, 为了减少采样空间, 本文将在观测对象附近使用规则生成采样视点.

### 4.1 采样视点的选择

本文将分别对虚拟模型和视频模型周围进行视点采样. 虚拟模型包括建筑模型, 天空球和地面, 其中只需对建筑模型进行采样. 对于每个建筑模型, 本文进行多层包围球上的概率均匀的随机采样. 本文固定虚拟视点目标为朝向模型最小包围球中心. 对于球面坐标表示的视点  $(r, \theta, \phi)$ , 其中  $r$  为包围球半径,  $(\theta, \phi)$  分别为球面坐标. 本文首先对  $(\theta, \phi)$  进行均匀随机采样, 保证在单位球面积

上只有一个随机采样点. 接着对  $r$  进行均匀随机采样, 保证在单位半径距离中只有一个采样点, 其中起始距离和单位半径距离均为最小包围球半径  $r_0$ , 采样半径倍数的为  $n$ , 当  $n$  为 4 时, 采样效果如图 5 所示.

本文接着对视频模型进行视点采样. 由于视频模型与网格模型有很大的区别, 因此随机采样的两个过程也有所不同. 同样的本文使用视频拍摄点侧重的多层包围球采样, 即采样概率上偏重, 越靠近视频模型的相机视点位置, 采样视点越稠密. 这样进行采样的原因有两个: 一是在视频模型的相机视点附近, 基于视频画面畸变的视点质量较高; 二是路径规划的目标是生成连接各个视频模型的相机视点的路径, 因此在相机视点附近进行更稠密的采样有助于提高路径质量.

为了使包围球的不同区域的视点密度不同, 本文使用线性函数表示密度, 以角度  $\phi$  为例, 密度关系如式(7)所示. 本文规定在相机视点处(即  $\phi=0$  的位置)的采样密度是包围球上距离相机视点最远处(即  $\phi=\pi$  的位置)密度的  $k$  倍, 同时采样点密度在球面上进行积分的结果是 1. 这两个约束可以表达为式(8).

$$\rho(\phi) = a \times \phi + b \quad (7)$$

$$\begin{cases} \rho(0) = k \times \rho(\pi) \\ \int_0^\pi \rho(\phi) \times 2\pi \sin \phi d\phi = 1 \end{cases} \quad (8)$$

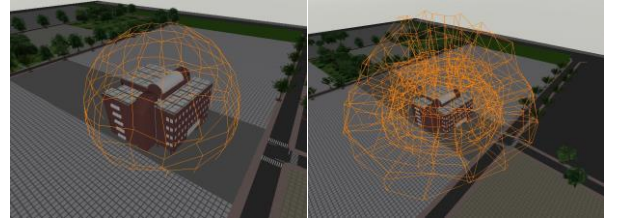
根据上式, 可以得出  $a$  和  $b$  的取值. 给定球冠上采样点占全部采样点的比例  $p$ , 本文有式(9):

$$\int_0^\phi \rho(x) \times 2\pi \sin x dx = p \quad (9)$$

通过式(9), 无法直接计算由比例  $p$  得到角度  $\phi$ . 但是可以建立查找表反向快速计算  $\phi$  的取值.

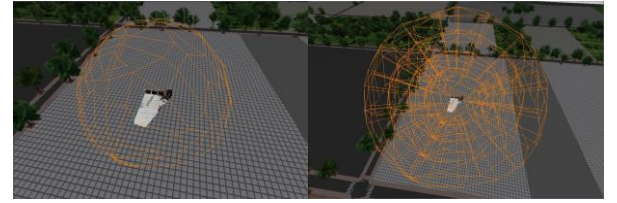
在对视频模型的  $(\theta, \phi)$  进行视频拍摄视点侧偏重采样后, 本文对  $r$  依然采用偏重采样的做法. 对于经过上一步采样得到的模型包围球上的一个视点  $(r, \theta, \phi)$ , 根据  $\phi$  的取值计算在该直线上的采样个数  $n$ , 在  $(r_0, (n+1) \times r_0)$  这个区间上随机采样  $n$  个点, 其中  $n$  的计算方法如式(10).  $\lceil \cdot \rceil$  表示向上取整.

$$n = \left\lceil \frac{1-k}{\pi} \times \phi + k \right\rceil \quad (10)$$



a. 单层包围球采样结果      b. 多层包围球采样结果

图 5 针对虚拟模型的视点采样结果示意图



a. 单层包围球采样结果      b. 多层包围球采样结果

图 6 针对视频模型的视点采样结果示意图

#### 4.2 道路图的构建

在空间中进行视点采样后, 可以构建出道路图的顶点集合. 但为了后续路径计算的过程提供一张完整的道路图, 需要构建道路图的边. 其中包括连接合理边与定义边的权重.

根据路径生成目标, 本文针对每个顶点, 判断与它邻域内的所有其他顶点是否可以构成连线, 将符合要求的连线加入道路图的边集. 判断边是否复合要求的判断标准是这条边是否与虚拟模型发生了相交, 即路径是否被模型所阻碍.

道路图中边的权重定义方式取决于系统对路径的需求. 有的路径生成系统要求路径尽可能短, 因此把移动距离作为主要权重; 有的系统需要视角旋转尽可能平稳, 因此把旋转角度作为主要权重; 有的系统需要目标物体尽量可见, 因此把目标物体的可见性作为权重的主要因素. 本文的路径结果要求用户看到的画面质量较好, 同时保证画面的平稳切换. 考虑到这些需求, 本文使用移动距离和旋转角度同时结合视点的质量评价分数来定义道路图上边的权重.

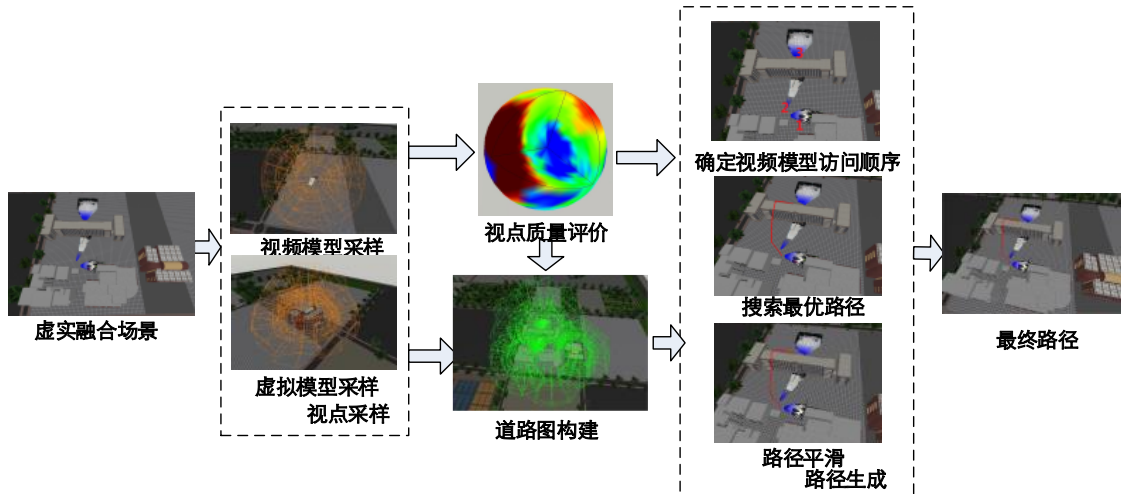


图 7 漫游路径规划方法流程图

移动距离反映了视点位置的变化,用户在漫游过程中会感受到空间位置的变化;旋转角度反映了用户画面的变化速度,过快的画面变化会使用户迷失空间感。结合视点质量,由采样视点 $(v_i, v_j)$ 构成的边对应的权重计算方式如式(11)所示。

$$f(v_i, v_j) = (w_1 \times \|v_{i, \text{pos}} - v_{j, \text{pos}}\| + w_2 \times A(v_i, v_j)) \times C(v_i, v_j) \quad (11)$$

该权重定义为距离、角度的综合评价与视点质量因素  $C(v_i, v_j)$  的乘积。 $\|v_{i, \text{pos}} - v_{j, \text{pos}}\|$  为两个视点之间的空间距离,  $A(v_i, v_j)$  为两个视点之间的旋转角度,  $w_1$  和  $w_2$  分别表示距离和角度所占的权重。 $C(v_i, v_j)$  是一个比例因子,表示视点质量对边的权重造成的影响,该因子的计算方式如式(12)所示。

$$C(v_i, v_j) = \frac{C(v_i) + C(v_j)}{2} \quad (12)$$

比例因子  $C(v_i, v_j)$  是将视点  $v_i$  和视点  $v_j$  决定的比例因子取平均值得到的。

每个视点  $v$  对应的比例因子与该视点的视点质量有关,视点质量越高,该比例因子越小,二者转化方式如式(13)所示。

$$C(v) = \exp(\beta(1 - S(v))) \quad (13)$$

### 4.3 路径生成

在道路图构建之后,场景中的所有的视频模型就构成了一张带权有向图。确定一条视频模型的观看顺序等价于旅行商问题,即在图中找到一

条连接所有顶点的路径,同时使得该路径总代价最小。该问题是个 NP 问题,本文采用近似解法解决。

本文首先人工选取一个起始点,然后根据贪心的方法选取下个距离最短且未访问的视频模型,直到所有的视频模型被选中。虽然该方法得到的不一定是最优解,但视频模型之间的访问顺序并不影响视频模型之间的转移路径的效果。根据用户的访问需求,本文会改变视频模型访问策略。

而在每两个相机之间,本文采用 A\* 方法来寻找最短路径。确保每次相机间转移生成的路径都是边权重和最小的路径。最后,本文对路径使用球面插值进行了平滑。本文路径规划方法的总体流程图如图 7 所示。

## 5 实验结果

将本文的视点质量评价与路径生成方法应用于规模不同的视频监控系统,其中包括 Campus, Railway Station, Building 以及 Traffic 等四个监控场景。所有相机均为固定的枪机监控摄像机,位姿与位置多样。其主要场景参数如表 1 所示。实验环境为 Intel(R) Xeon(R) 4 核 3.70 GHz CPU 和 12GB RAM。在视点采样阶段,  $n$  始终设置为 4,  $k$  始终设置为 5。采样相关统计结果如表 2 所示。

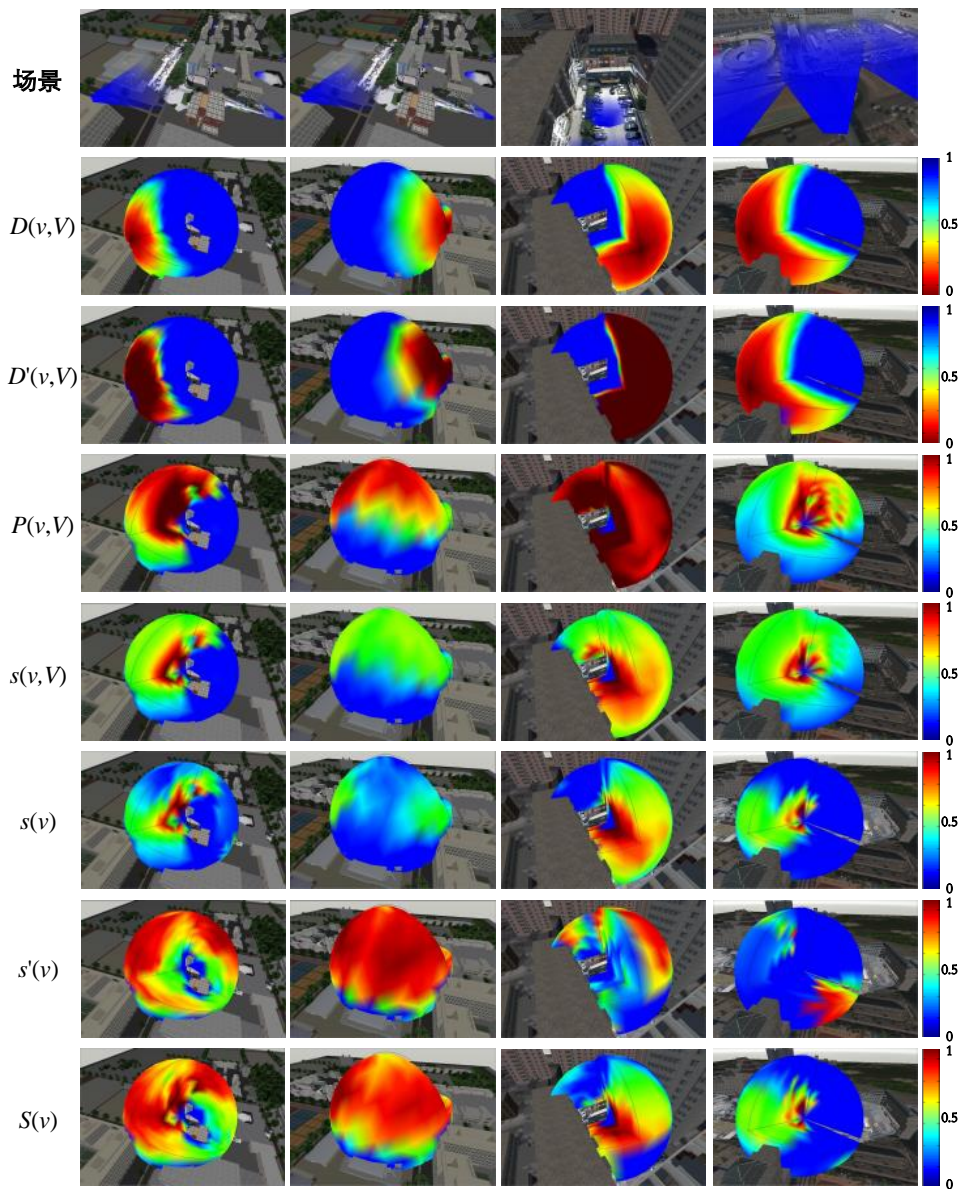


图 8 视点评价中各变量可视化结果

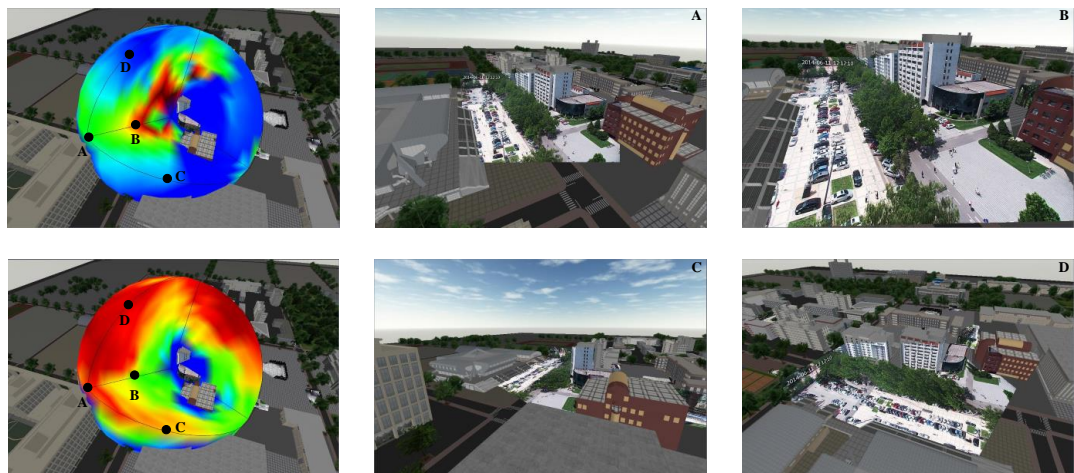


图 9 视点评价与对应视点下面图

表1 实验所用监控场景主要参数

场景名称	建筑数目	相机数目	场景大小	
			面片数	占地面积/m <sup>2</sup>
Campus	35	12	336 514	700×700
Railway Station	11	10	70 725	700×300
Building	10	6	654 166	150×150
Traffic	30	28	242 844	800×100

表2 采样结果相关参数

场景名称	有效采样点个数及比例	平均质量	平均采样时间/ms
Campus	4787(91.3%)	0.574	7.16
Railway Station	1988(93.5%)	0.594	3.15
Building	1343(95.4%)	0.603	4.45
Traffic	5466(85.4%)	0.631	1.84

### 5.1 视点质量评价效果

为了更好地可视化展示本文的结果,视点质量评价结果均以视点与视频模型中连线为半径的 $7/8$ 球表示,即限定视点目标为视频模型为中心,视点可在球内任意位置移动,如图8所示.该图中的每一行分别为视频画面的畸变程度 $D(v,V)$ ,畸变程度的截断值 $D'(v,V)$ ,投影面积比 $P(v,V)$ ,单视频模型视点评价 $s(v,V)$ ,多视频模型视点评价 $s(v)$ ,虚拟模型视点评价 $s'(v)$ 以及整体视点评价 $S(v)$ .其中第1列的图像为评价球背面视图,其余列图像为正面视图.图中的着色结果均为归一化值.

以第1列为例,从第1行可以看出,相机拍摄视点的畸变程度最小,随着偏移拍摄视点的距离和角度越远,畸变程度越高,达到一定距离和角度后,畸变程度值已与视频拍摄视点的附近视点不可相比.从第2行中可以看出截断操作对畸变程度的衡量缩减了最佳视点评价价值与其他视点评价价值的距离,体现出畸变程度随着视点移动方向的不同呈现不同的变大趋势,这与本文在实际浏览中观察到的视频画面畸变的各向异性一致,相比于 $D(v,V)$ 更加适合用于视点评价的数值计算.第3和

第4行相比,可以看出投影面积比对畸变程度的加权性,在第4行图像中,靠近球中心的A视点评价反而上升.由于其他模型的投影面积比始终很小,因此在 $s(v)$ 中其他模型对视点评价结果的加权影响不大.最后两列展示出各自评价 $s(v)$ 与 $s'(v)$ 侧重点不同,均能对整体视点的评价产生影响.

本文同时给出了球上多个视点下对应的图像,如图9所示.视频模型的视点质量评价方法 $s(v)$ 与基于视点熵评价的虚拟模型视点评价方法 $s'(v)$ 相比, $s(v)$ 在视频画面扭曲更严重的视点D上得到了更低的分数,在较少扭曲的视点B上得到了更高的分数.因此 $s(v)$ 更适合刻画视频画面中的纹理扭曲程度.这一点对于深度经常出现不匹配情况的虚实融合场景来说尤为重要.

### 5.2 路径生成效果

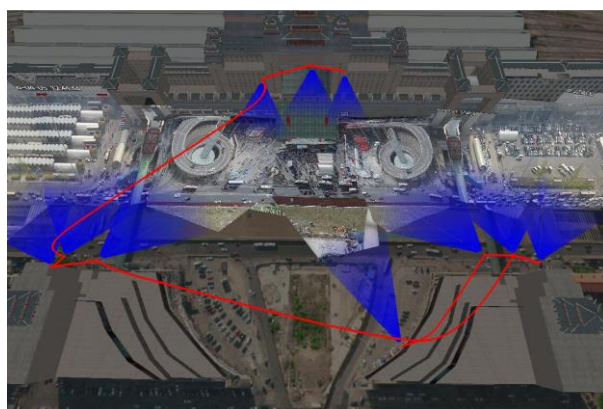
本文展示各个场景最终的漫游路径结果,如图10所示.用红色曲线表示漫游相机的行进路径<sup>②</sup>;而蓝色锥形表示相机的视场,颜色越深越靠近相机,相反越远离相机.从图中可以看出,无论是图10c的Building场景中视频围绕型分布,图10d的Traffic为长廊型分布,图10b中的Railway Station为回环型分布,还是图10a的Campus为复杂型分布,本文的方法均能产生视点转移路线,生成的路径能够遍历访问所有视频.

图11给出了漫游路径的局部图,着重展示了在两个相邻访问顺序的相机之间的转换路径.第1列第1行的图像对应于图8第1列的整体视点评价,路径穿过评价球的中部到达下一个视频,视线方向朝向右侧白色建筑的立面.第2幅图像中路径接连穿过起始和下一个视频模型的立板,然后到达下一个视频模型的视点.由于立板的建模从接地线开始折起,立起的对象只有树木,畸变不大.而地面也不存在大的畸变,因此该路线上看到的视频画面畸变均不会很大,符合本文对路径的要求.后面的七幅局部图像也同样满足要求.

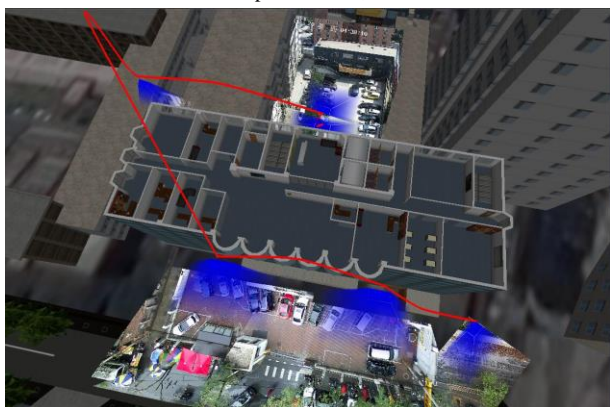




a. Campus 场景路径



b. Railway Station 场景路径



c. Building 场景路径



d. Traffic 场景路径

图 10 漫游路径整体结果

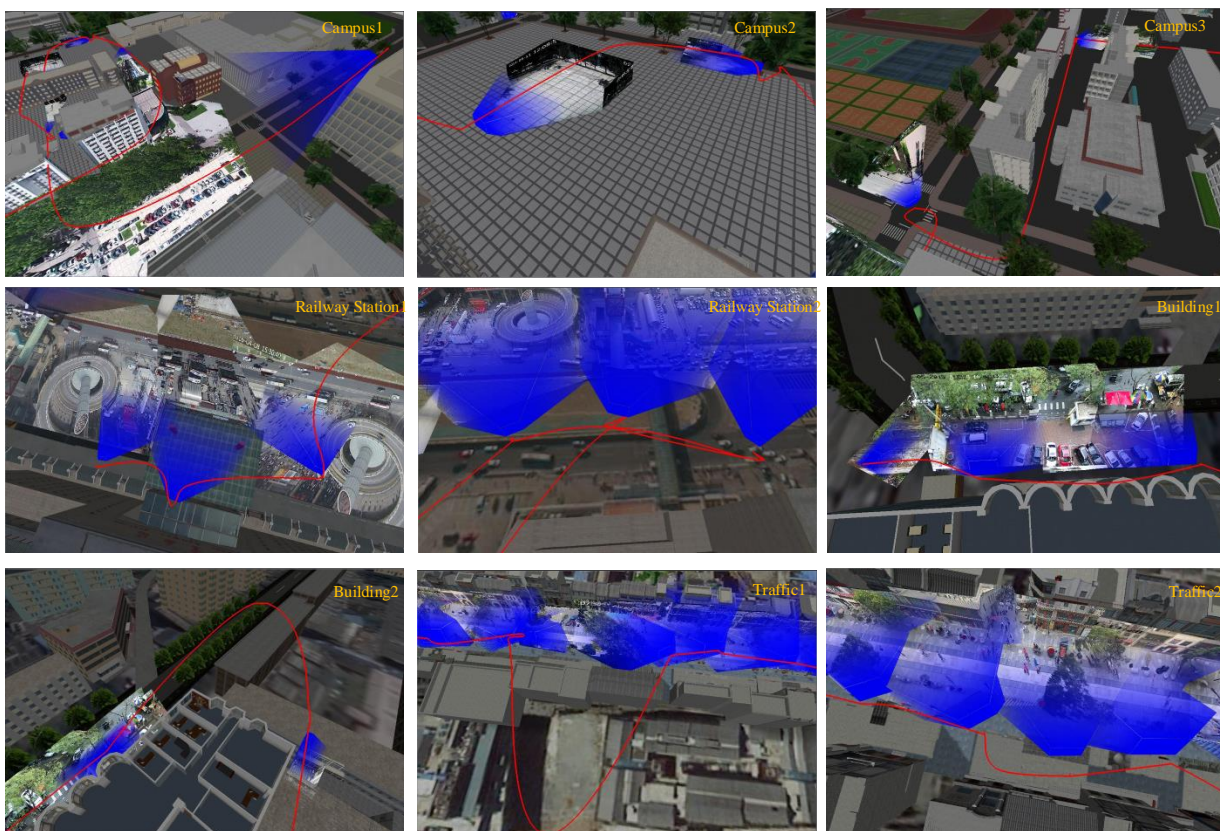


图 11 漫游路径局部结果

### 5.3 方法复杂度分析

以视点评价完成为分割线,方法前半段的复杂度是正比于采样视点个数,后半段的复杂度正比于道路图边与虚拟模型相交情况的个数.这是因为质量评价中的投影面积计算与道路图构建中的相交判断是最占用计算时间的.本文在方法实现时将视点的采样和质量评价、和道路图的构建这两步提前离线计算,而路径生成作为在线计算的步骤.确保了方法的实用性.

表 3 算法各部分耗时  $s$  及占总时间的百分比 %

场景名称	视点采样与质量评价	道路图构建	路径生成	总时间
Campus	247.6(43.3)	324.2(56.7)	0.5(0.1)	572.3
Railway Station	12.3(44.6)	14.8(53.6)	0.5(1.8)	27.6
Building	9.8(20.4)	38.0(79.2)	0.2(0.4)	48.0
Traffic	18.2(9.0)	179.7(89.3)	3.3(1.6)	201.2

### 5.4 用户研究

用户对一条漫游路径评价很大程度上受到主观性因素的影响,很难用数值指标对路径生成结果进行评价.因此,本文进行了用户研究,对用户使用该路径进行漫游的各项指标的满意程度进行了调研.

本文调研了 30 名志愿者使用该漫游路径浏览虚实融合场景的感受.这些志愿者中,有 20 名在此之前没有接触过虚实融合系统,有 5 名了解相关图形图像相关知识,但是没有使用过本虚实融合视频平台,剩余 5 名是本实验室课题组内的同学,对虚实融合视频平台较为熟悉.

每一名志愿者先通过鼠标和键盘控制视角移动,在场景中自由浏览一段时间,熟悉系统的用途与基本交互方式,接着由路径生成系统自动生成漫游路径,使用该漫游路径观察整个场景,对路径生成系统的各项指标进行满意度评价.本文共设计了 3 项评价指标:漫游过程中画面的质量、路径生成系统的操作复杂度和路径生成的交互响应及时度.用户的统计结果如表 4 所示.

表 4 用户研究的人数统计数据

评价指标	满意	一般	不满意	满意度百分比
漫游画面质量	26	2	2	86.7%
操作复杂度	29	0	1	96.7%
交互响应及时度	21	3	6	70%

由上表的统计结果可以看出,路径生成方法在各项指标上均至少有七成用户表示满意,表明路径生成方法整体来说是有效的.观察各项指标的满意度,发现本方法在漫游过程中的画面质量和简化操作复杂度方面分别获得了 86.7% 和 96.7% 的满意度,表明方法在这两个方面的效果提升较大.在交互响应及时度上,由于方法执行时间与场景大小以及模型数量有关,对于模型较多的场景所需时间较多,只获得了 70% 用户的满意,表明路径生成方法在执行效率方面仍需要进一步优化.

## 6 结 语

本文提出了一种计算自由视点下视频画面畸变程度的视点质量度量方法和基于视点质量度量的路径生成方法.基于视点质量的分布,分布为视频模型和虚拟模型设计了随机视点采样方法,接着选择邻近采样点构建连通的道路图.最后,一个在两点间累积畸变最小的路径自动的计算生成,并首尾相连组成完整路径.本文的视点评价方法,相比于其他如视点熵的视点评价方法,评价更能符合对投影纹理变化好坏的直观感受.而基于此的路径生成方法也能够实现及时交互,实用性十足.本文的实验,证明了本文的方法能够提升观测者的漫游视觉体验.

实际上,对于一个监控场景来说,漫游路径还需要根据场景的变化进行联动变化,比如监控相机发出了异常报警.此时,漫游路径需要从当前视点快速跳转到报警区域.本文的方法只需要将确定视频模型访问顺序,从最小生成树方法改成最短路径生成方法,即可满足监控场景的动态路径生成需要.

本文的方法还有值得改进的地方, 比如视点评价的参考深度来自于单幅深度估计方法, 该方法得到的深度与真实深度还有差距, 如果能够架设辅助相机, 使用双目视觉进行深度估计, 将会得到更精确的深度, 进而得到更优的视点质量评价结果. 另一方面, 本文发现在视点沿着路径的行径过程中, 空间中存在质量较优的组合视点, 如 Railway Station 场景中右方的三个相机偏后方的视点. 本文可以为邻近相机集设计并组合出特定的相机视点, 来代替拍摄视点. 这样得到的路径进行漫游时, 相机间的位置感和统揽的感觉将得到加强.

### 参考文献(References):

- [1] Sebe I O, Hu J, You S, *et al.* 3d video surveillance with augmented virtual environments[C] //Proceedings of the 1st ACM SIGMM international workshop on Video surveillance. New York: ACM Press, 2003: 107-112
- [2] 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. Los Alamitos: IEEE Computer Society Press, 2003: 61
- [3] Kim K, Oh S, Lee J, *et al.* Augmenting Aerial Earth Maps with dynamic information[C] //Proceedings of IEEE International Symposium on Mixed and Augmented Reality. Los Alamitos: IEEE Computer Society Press, 2009: 35-38
- [4] Vincent Luc. Taking online maps down to street level[J]. Computer, 2007, 40(12): 118-120
- [5] Snavely N, Seitz S M, Richard S. Photo tourism: exploring photo collections in 3D[J]. ACM transactions on graphics, 2006, 25(3): 835-846
- [6] Sbert M, Plemenos D, Feixas M, *et al.* Viewpoint quality: Measures and applications[C] //Proceedings of the 1st Eurographics Workshop on Computational Aesthetics in Graphics, Visualization and Imaging. Aire-la-Vile: The Eurographics Association Press, 2005: 185-192
- [7] Secord A, Lu J, Finkelstein A, *et al.* Perceptual models of viewpoint preference[J]. ACM Transactions on Graphics, 2011, 30(5): 109
- [8] Zquez P P, Feixas M, Sbert M, *et al.* Viewpoint selection using viewpoint entropy[C] //Proceedings of the 6th International Fall Workshop on Vision, Modeling, and Visualization. New York: ACM Press, 2001: 273-280
- [9] Lee C H, Varshney A, Jacobs D W. Mesh saliency[J]. ACM Transactions on Graphics, 2005, 24(3): 659-666
- [10] Yamauchi H, Saleem W, Yoshizawa S, *et al.* Towards Stable and Salient Multi-View Representation of 3D Shapes[C] //Proceedings of IEEE International Conference on Shape Modeling and Applications. Los Alamitos: IEEE Computer Society Press, 2006: 40
- [11] Liu B, Gould S, Koller D. Single image depth estimation from predicted semantic labels[C] //Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2010: 1253-1260
- [12] Tsai G, Xu C, Liu J, *et al.* Real-time indoor scene understanding using Bayesian filtering with motion cues[C] //Proceedings of IEEE International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2011: 121-128
- [13] Karsch K, Liu C, Kang S B. Depth Transfer: Depth extraction from video using non-parametric sampling[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 36(11): 775-788
- [14] Liu M, Salzmann M, He X. Discrete-Continuous Depth Estimation from a Single Image[C] //Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2014: 716-723
- [15] Liu F, Shen C, Lin G. Deep convolutional neural fields for depth estimation from a single image[C] //Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2015: 5162-5170
- [16] Roy A, Todorovic S. Monocular Depth Estimation Using Neural Regression Forest[C] //Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2016: 5506-5514
- [17] Kavraki L E, Švestka P, Latombe J C, *et al.* Probabilistic roadmaps for path planning in high-dimensional configuration spaces[J]. IEEE Transactions on Robotics & Automation, 1996, 12(4): 566-580
- [18] Ranon R, and Urli T. Improving the efficiency of viewpoint composition[J]. IEEE Transactions on Visualization and Computer Graphics, 2014, 20(5): 795-807
- [19] Oskam T, Sumner R W, Thurey N, *et al.* Visibility Transition Planning for Dynamic Camera Control[C] //Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation. New York: ACM Press, 2010: 55-65
- [20] Zhou Z, You J, Yang J, *et al.* Method for 3D Scene Structure Modeling and Camera Registration from Single Image: United States, 20160249041[P]. 2016-08-25

①三维路径和其朝向在二维图像中难以展示, 我们在补充视频中给出了完整路径.