**工程伦理课程论文要求**

**课程内容：**主要探讨工程伦理的基本概念、基本伦理问题，以及工程实践过程中人们将要面对的共性问题。主要包括：1.分析工程和伦理的概念，工程实践中的伦理问题，以及处理工程伦理问题的基本原则；2.分别从责任伦理与伦理责任、利益分配与公正、环境伦理与环境正义三个方面探讨所有工程实践都可能面对的一些共性问题；3.探讨工程师的职业伦理；4.分析信息工程领域面对的特殊问题，以及共性伦理问题在信息工程领域的特殊表现，分析信息工程领域的工程伦理规范。

**考核方式：**

考查。采用课程论文的形式进行考核。

提交纸质版论文**截止时间**：第16周周一16时；

**内容要求**：

一人一题。5000字左右。可结合课程内容及专业特点自由选择案例、进行分析、明确阐明个人观点（尽量选择**信息工程领域**的案例，否则酌情扣分）。

参考文献尽量在5篇以上，其中至少要有三篇近三年发表的期刊论文或电子文献。注意正文中的引文标注应从1开始顺序标注，注意相应地调整文后参考文献里的文章列表顺序。

**格式要求**（版式示例附后）：

***加试卷封面***（附后，无页眉页脚）。

***论文题目***：居中，宋体/ Times New Roman，三号加粗，单倍行距，段前段后0.5行距。

***作者信息***：居中，宋体/ Times New Roman，四号常规，单倍行距，段前段后0.5行距。

***摘要***：宋体，五号加粗；摘要和关键词左右侧各缩进2.5字符，行距20磅，无段间距。

***摘要正文***：楷体，五号常规。

**关键词**：同摘要。

***正文标题***：居左，宋体/ Times New Roman，小四号加粗，单倍行距，段前段后0.5行距。

***正文***：两端对齐，宋体/ Times New Roman，小四，行距20磅，无行间距。

***页眉***：居中，宋体，小五，单倍行距，含分隔线。

***页脚***：居中，Times New Roman，小五，单倍行距，不含分隔线。

***页面设置***：上下左右均为2cm，页眉页脚1.5cm；

*文件名****命名***：学号跟姓名-题目（同时提交Word和Pdf版本）

四 川 大 学

硕士研究生课程考试试卷

姓　　名 王祎 学　号 2020223045231

学院（所、中心） 计算机学院 专　　业 电子信息

|  |  |
| --- | --- |
| 考试课程名称 虚拟现实 | |
| 考试方式 □笔试 □口试 √撰写论文  任课教师 张严辞 | 考试成绩\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  考试时间 2020.1.5 |

四川大学研究生院制

* Techniques behind digital human “Siren”

2020223045231 王祎

**abstract：**The Siren project began as a collaboration between Epic and Tencent, with the goal of creating a proof-of-concept demonstration to show both the capabilities of UE4, and what the next generation of digital characters will look like. To handle the performance capture for the character, Epic and Tencent utilized Vicon’s Vantage optical motion capture system, Shōgun and VUE video cameras to capture precise and authentic movement and to add the character animations over the reference footage in real time. The creation of the entire siren model involves many technical content such as motion capture, facial scanning, and real-time rendering of facial information. It is the combination of these information technologies that makes the results realistic. This article will focus on the specific implementation of these technical details and mutual Combination to expand interpretation.

**keywords：**siren；motion capture；rendering

**摘 要：**Siren是Epic和腾讯共同合作的项目，目的之一是创建概念验证演示，以展示UE4的功能，另一个目的是展示下一代数字字符元素会如何发展。为了处理角色的表演，Epic和腾讯利用了Vicon的Vantage光学运动捕捉系统Shogun和VUE摄像机来捕捉精确而真实的动作，该系统在数字项目创建实时性上已经取得了极好的表现，整个siren模型的创建涉及到了动作捕获，面部扫描，面部信息实时渲染等诸多技术内容，正是这些信息技术的组合才使得其结果达到可逼真的效果，本文将针对这些技术细节的具体实现及互相组合展开解读。

**关键词：**siren；动作捕获；渲染

1. 技术框架概述

据了解《速度与激情8》在拍电影的时候演员保罗沃克已经不在人世了，为了让电影继续拍下去，电影组把这个人的形象复员了，那么这是如何实现的呢？保罗沃克非常幸运，他有一个弟弟跟他长的很像，剧组根据他弟弟的形象做了3D扫描，在原先照片的技术上做了修改，尽量看起来跟真人很像，最后才有了《速度与激情8》保罗沃克的戏。电影里面两分钟的戏最后要修的话有可能要经过几个月的时间，甚至是没有什么交互性。而现在的渲染技术siren却做得互动性，实时性，逼真性的统一，这样的技术又该如何实现？

简单分析一下整体技术，看看为什么实时渲染出高真实度的虚拟人物是一件很困难的事情。

首当其冲的问题是渲染性能问题，离线渲染可以花费数小时渲染一帧画面，投入服务器集群的计算资源，用尽可能物理正确的光线追踪技术。但游戏中使用的实时渲染，每秒至少要渲染30帧画面，也就是说每一帧画面只有33毫秒，这比离线渲染少了五个数量级，渲染时只有一台PC的CPU和GPU资源，相比离线渲染又少了一到二个数量级，在计算资源上的差距，非常巨大。

下一个问题是渲染质量问题，毛发渲染、眼球渲染、皮肤渲染等高阶渲染技术，被大量应用在电影行业。对于动物、人物角色的衣料或者毛发，人物眼睛的表现，角色皮肤表现等等领域，传统的渲染技术还远远不能真实表现。电影行业在这些领域做了深入的开发，高质量的离线算法，虽然开销巨大，但对最终画面质量提升有着决定性的作用。实时渲染无法使用高级的渲染技术，只有高度抽象和简化过的渲染算法，这让渲染器的最终质量，比电影级的渲染低了不少。

最后一个问题是复杂的流程。为了在电影中做到如此高质量的角色，电影工业在过去20年中建立了相当复杂的制作流程。简单来说，制作一个这样的角色，离不开下面的这些技术：

(1) 高质量的扫描和建模：通过基于摄影测量的扫描技术，从多角度拍摄即将建模的物件、人体，后续通过计算机图形学，进行复杂的3D建模、Rig建立等等步骤，产生最高质量的3D模型和材质。

(2) 全身动作捕捉和还原：捕捉演员的表演动作，还原成动画数据来回放。动作捕捉技术相对比较成熟，但依然面临着捕捉精度、稳定性、设备部署维护等等巨大的挑战。这部分动画数据，后续会被用来驱动虚拟的角色，让虚拟角色能做出和演员一样的动作。

(3) 面部AI捕捉和还原：通常会和演员全身动作捕捉同期进行。演员在表演的时候，一般会佩戴特制的头盔，头盔中有一个或者数个摄像头，拍摄演员的脸部表情以及对话口型，在后期制作时再通过计算机视觉，还原成表情动画。本质上和全身动作捕捉技术有一定的相似性，但由于脸部表情对于角色形象极其重要，所以面部捕捉和还原会应用更高精度的技术来进行。

而对于siren而言，为了制作互动演示，女演员艾丽卡·李（Alexa Lee）佩戴了头戴式摄像机装备和全身运动捕捉套装。通过解决3Lateral绑定的Cubic Motion无标记跟踪，为她的脸设置了动画。使用Vicon的新Shōgun1.2软件，可以在一个屏幕上捕获她的身体和手指的运动，同时使用Vicon的新Live Link插件将数据流传输到虚幻引擎中

Siren利用了Cubic Motion的面部表情捕捉系统，3Lateral的面部操控技术，Vicon的身体动捕系统，具备逼真的面部和眼部运动，以及与参与者交互的能力，展示了面部表情捕捉系统的全部潜力，支持能够反应人类情感的人脸动画。

为了达到逼真的面部效果，siren实际脸部以60 fps实时渲染，并包含细微的面部毛发和来自UE4团队的新型高级皮肤着色器。腾讯和Epic的工程师推出了新的反向散射算法和新的双镜面波瓣。光泽镜面层表示立即在皮肤边界反射的光，第二个镜面波瓣位于主镜面波瓣下方，并被其衰减，最终通过这些技术的叠加组合达到了以假乱真的效果。

下文将就具体的技术细节展开讨论，涉及到的内容如图1-1所示：

图1-1

1. Cubic Motion 原理解读

在光学捕捉系统中，摄像机会负责“识别”演员身上的标记，然后以每秒数次的频率计算各个标记的 3D 坐标。尽管原理相对简单，但这类系统需要多台摄像机才能工作，因为每个标记在任何时间内都要至少被两台摄像机拍摄，才能计算出 3D 坐标。不过，能观察到标记的摄像机越多，计算出的位置就越精确。

摄像机看到的标记可以分为两类：主动标记和被动标记，主动标记是指能够自发光并让摄像机检测到的标记。一些主动标记能以脉冲方式闪烁，从而向摄像机发送额外的信息，例如该标记的 ID。 相反，被动标记不会自主发光，它必须借助其他途径反射光线（通常是摄像机上的环形补光灯）。被动标记不会主动发送 ID 信息，所以系统需要通过分析它相对于其他标记的运动来建立被追踪标记的数学模型，从而识别出各个标记。

主动标记的一个优点在于，它们有时可以与自然光高效地结合运用。相反，被动标记就不适用于这种情况，因为标记的亮度与环境光之间的对比度不足。可以想象，主动标记更加可靠，而且可以从更远的地方进行捕捉；然而，主动标记需要配备大量电子器件，导致它们比被动标记更为笨重，而且由于需要为标记供电，捕捉对象通常需要随身佩戴电池。由于需要大量的追踪摄像机，光学系统可能会成本高昂且不便于移动。此外，如果在现场摄影或表演转播时，观众能够看到演员，那就没必要采用动作捕捉了，可见的标记会影响到服装外观。 不过，这些系统能带来最优质的效果，并通过精确的位置数据以及受追踪的道具和摄像机，为同一场景中的多位演员提供支持。光学系统能够捕捉身体和面部的动作。面部光学系统往往需要一个头戴式摄像机（HMC）系统， 它被附加在演员戴着的定制头盔上，随后通过后续技术进行渲染，而siren的整个组成大概如下示意图2-1：

图2-1

* 1. HMC

Cubic Motion实现的重要组件之一就是头戴式摄像头系统，该系统主要负责面部内容及动作的实时捕捉，为了增加该头戴式摄像头系统的易用性和可移植性，开发者提出了以下4个特点：

一，提供三种头盔尺寸及可以选择填充物，以满足不同创作者的需求，给他们更舒适的体验。

二，HMC的设计与仅安装有一般立体摄像头的系统不同，它是通过安装最新的前置摄像头和侧面配置摄像头来完成立体动作捕获的。

三，内置照明系统采用红外灯，这种方式不仅可以保证在各种条件下也能捕获到最细微的嘴型和姿势并得到准确的恢复，还增加了舒适性，不会让演员感到厌烦和不适。

四，通过小的应用软件来管理灯光和时间码，极大的提高了易用性和准确性。

* 1. 性能管道

通过第一步捕获到了表演者的面部表情，第二步便是用性能管道解析到目标角色上。称之为性能管道的原因是该管道计算机通过一种轻巧的电池供电单元来实现无线运行，其上并装置有灵活的微型计算机系统。

每个性能管道都使用 identity system软件来进行配置和管理，该身份系统对于每个人来说都是唯一的，这样就可以保证每个演员的表演和渲染都是独一无二的，以此来传达出人物的独特性。实现方法是使用最新的机器学习算法构造每个人物身份并配备专门的演员外观和角色装备，通过系统算法将角色与角色的定制模型进行匹配和绑定。

* 1. 渲染管理器

performance manager是一个GUI应用程序，它可以让操作员方便的从远程界面配置和监视演员表演的各个方面的能力，远程监视管道HMC，包括所有摄像机，跟踪和动画源，控制演奏数据的广播，记录和导出，导入和管理Cubic Motion身份并分配给管道，可以实时读取各个部分的性能利用情况，以便于随时调整和修改。

1. 面部渲染实现

皮肤渲染技术经过数十年的发展，由最初的单张贴图+伦勃朗的渲染方式到近期的基于物理的SSSSS（屏幕空间次表面散射）。由此衍生出的皮肤渲染技术层出不穷，其中最最基础也最具代表性的是次表面散射（SSS）。在虚拟角色渲染中，皮肤的渲染尤为关键。因为皮肤是人们每天亲眼目睹的非常熟悉的东西，如果稍微渲染不好或细节处理不足，便会陷入恐怖谷（Uncanny Valley ）理论。

* 1. 皮肤构成

人类皮肤的物理构成非常复杂，其表层和内部都由非常复杂的构成物质，剖面图如下（图3-1）：

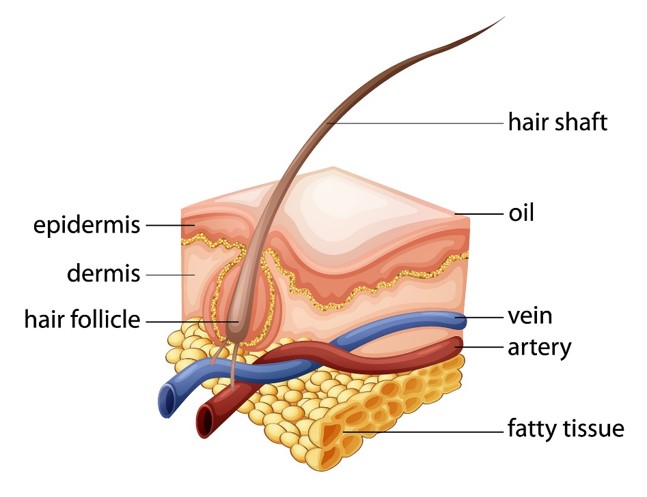


图 3-1

1. 绒毛（hair shaft）。附着于皮肤表面的细小的毛。
2. 油脂（oil）。皮肤表层有一层薄薄的油脂覆盖，是皮肤高光的主要贡献者。
3. 表皮（epidermis）。油脂层下是表皮覆盖，是造成次表面散射的物质之一。
4. 真皮（dermis）。表皮下面是真正的皮肤组织，也是造成次表面散射的物质之一。
5. 毛囊（hair follicle）。绒毛的皮下组织和根基。
6. 静脉（vein）。呈深蓝色的血管。
7. 动脉（artery）。呈暗红色的血管。
8. 脂肪组织（fatty tissue）。脂肪组织也是造成次表面散射的次要贡献物质。
9. 其它：皮肤表面的纹理、皱纹、毛孔、雀斑、痘痘、黑痣、疤痕、油脂粒等等细节。
   1. 皮肤建模

皮皮肤的渲染一直是渲染领域的难点之一：皮肤具有许多微妙的视觉特征，而观察者对皮肤的外观，特别是脸部的外观会非常敏感。皮肤的真实感渲染模型须包括皱纹，毛孔，雀斑，毛囊，疤痕等细节的模拟，而真实还原人体皮肤上的这些细节则是一个较大的挑战。

皮肤作为一种属性复杂的材质，其物理结构由多层结构组成，其表面油脂层主要贡献了皮肤光照的反射部分，而油脂层下面的表皮层和真皮层则主要贡献了的次表面散射部分。实验测试表明，光线接触到皮肤时，有大约94%被皮肤各层散射，只有大约6%被反射。虽然皮肤构成非常复杂，但图形渲染界的先贤者们利用简化的思维将皮肤建模成若干层。以下分层模型是使用广泛的模型之一：

1. 表面油脂层（Thin Oily Layer）：模拟皮肤的高光反射。
2. 表皮层（Epidermis）：模拟次表面散射的贡献层。
3. 真皮层（Dermis）：模拟次表面散射的贡献层。
   1. 渲染流程

图像渲染主要有两个方面的内容，How to Prove 和How to lmprove。How to Prove是做完一个东西之后必须要经过严格，甚至说非常变态的验证，证明做的东西是达到了照相机的水平。How to Improve，重点是在原片渲染的技术上做了哪些提升。

* + 1. 如何验证（How to Prove）

Siren技术采用了电影行业的LookDev技术，有大量的对比验证技术，在还原真实形象的过程中有很多环节，例如，相机、灯光，还有用灯光模型做验证，因为一个地方的出错会导致后面的结果越来越糟，所以要单独的验证每一个过程都是正确的。

在游戏关卡，光源的方向和朝向都跟真实的光场摆光一模一样，每一个系列从任何一个角度去比较，光照要跟真实的照片非常接近，才可以说这个东西是经过验证的。所以验证环节首先需要采集很多的照片，这是在灯光模型，可以认为是一个球体，里面会有很多盏灯，每次只亮一盏灯，其他灯都熄灭，然后依次循环下来从各个角度来拍这个人，最后要保证的结果是所有角度都是正确的。

最后是渲染的图片和灯光模型的比较，siren技术团队用了rv软件比较，通过对比的方式左右可以比较。精细拖动是更高级的方式，有一个精度拖条可以拖，一边是Rendering，最另一边对应就是渲染的照片，可以达到象素的比较，每次做完一个特写都要经过非常严格的验证，用非常严格的方式来证明做出来的东西的确是像素级。

* + 1. 如何提升（How to lmprove）

我们看一下用哪些手段对它进行提升，从最开始的“双胞胎模型”到Siren，首先很重要的技术是dual lobes，双重曝光，因为皮肤表面是很特殊的材质，它有很多层，表面是油性介质，反光非常强烈，下面还有各种油层，一层是无法表现皮肤的。观察Siren可以发现，首先是油光发亮的一层，除了油光发亮的之外，还有非常有质感的内层皮肤，两者之间有非常明显的对比。这个技术目前看来是很轻的技术，其实不然，在2007年TGDC的时候，Nvidia他们当时有一个讨论会的时候，当时的技术架构尝试用了4个lobes，于是siren模型开始也设了四层，运算指令没有比以前多一条，但因为四层的关系，调起来实在太复杂，最后siren还是放弃了就用两层，现在电影里面也是用了两层，四层的模型复杂度还是比较高的。

关于specular模型，siren一开始用的是GGX模型，效果还算不错，但是用了一段时间发现对于人物这样的材质，GGX的塑料感太强，不足以反映人类皮肤的曝光。所以借助来自电影行业的经验，切到了贝克曼模型，发现比以前的皮肤质感好很多。

脸上的皮肤，用了次表面散射模型，但是牙齿大部分游戏里面还没有用次表面散射模型，牙齿其实是透明度比皮肤还要高，用次表面散射模型来做更合理，如果牙齿和皮肤都用次表面散射模型，中间就会产生污染，因为牙齿再做次表面散射模型的时候，会把嘴唇也影响到，就会有非常明确的阴影，为了去掉阴影，牙齿原先默认点亮，一点都不像牙齿，用了次表面散射模型之后觉得这就是牙齿，为了去掉阴影，只用一个颜色。如果旁边的表现不一样，就用边缘颜色修复上去，这样相当于是把原先的替换成颜色的渐变。

眼睛，原先也是没有用次表面散射模型，也是用很近似的公式做了，并没有真正用次表面散射模型。效果跟原先的不一样，这里的profile也是用两层的。用次表面散射模型，跟牙齿存在同样的问题，它跟皮肤的Profile不一样，也会引起阴影，所以也是需要用同样的方式，对比还是挺明显的。眼睛原先看起来没有质感，之所以之前没有做，是因为原先的激光和通道不够，这次通过特殊的方式，预留出一些通道，做出来之后凹凸感比以前强了很多，体现出质感，因为眼睛是心灵的窗户，看着它眼睛的质感很好，就会觉得感情流露的很好。

脸上的皱纹，在捕捉好几个典型的哭、笑、悲哀等8种技术表情，每一个表情是有一定的区域，最后通过一定的方式形成3张贴图。材质里面会生成一个Mask，Mask有一堆的参数，做表情的时候会根据面部动产生rig的数据，控制额头的皱纹，或者是脸颊的皱纹，最后生成一个最终效果。

最后是脸上的毛发，如果要达到逼真的效果，毛发的重要性不言而喻，用到虚拟数字人上，脸上没毛平淡无聊，脸看起来光秃秃的，没有毛茸茸的质感，加了毛，跟之前的对比好了很多。关于这个毛，怎么能实时的把毛做的这么好？其实这个毛是通过非常暴力的方式做出来的，这是它的模型，麻雀虽小五脏具全，上面的毛，这个三角形算起来比脸上的三角形还要多，脸一共有4万多，这个就要超过10万多。

以上是渲染技术是根据腾讯互娱技术专家的介绍进行的总结提炼，为了保持严谨性以时间线的方式从渲染流程的角度进行了介绍。

1. 结语

本文着重介绍了siren背后的动作捕获和面部渲染技术的背后原理。对于动作捕获，近年来，光学/惯性混合系统陆续出现。这些系统保留了视觉和惯性双方各自的部分优点，同时摈除了一些缺点，比如遮蔽的问题。如果在运动期间摄像机的画面暂时屏蔽了过多的光学标记，就没有足够的数据来正确地还原演员的表演。另一方面，惯性系统就可以无视遮蔽，继续提供数据。在结合了剩余的可见标记的数据后，这些额外数据将有助于确保提供充足的追踪数据。 惯性数据也有助于减少与光学系统数据结合时的信号颤动（动作侦测时的噪点）。在每一帧寻找最接近标记的 3D 位置时，光学系统有一定程度的测量不准确性。技术标记实际上并未移动，这个“最接近”的位置也可能会在每一帧有细微变化，这是由于各种环境因素的影响，例如环境光的改变。这些细微变化会让标记计算出的位置出现颤动。在整个运动期间，颤动并不是很大的问题，它会被掩盖在更大幅度的动作下，但如果运动速度变慢，或者标记是静止的，颤动就会变得很明显。如果将惯性传感器安装到系统上，当惯性传感器的加速计侦测到标记运动缓慢或静止，光学数据就会受到额外的筛选，从而变得平滑，进而消除颤动。因此笔者希望后续开发能结合光学和惯性混合系统达到更先进的技术实现和进步。

参考文献

[1] Next-Generation-Character-Rendering (ACM Transactions on Graphics, Vol. 29(5), SIGGRAPH Asia 2010).

[2] ensen, Henrik Wann, Stephen R. Marschner, Marc Levoy, and Pat Hanrahan.2001. "A Practical Model for Subsurface Light Transport." In Proceedings of SIGGRAPH 2001.

[3] 2018TGDC解卫博：实时数字人Siren渲染技术揭秘 https://www.gameres.com/818199.html