

计算摄影学

课程代码： 课程名称：计算摄影学

学分：4 周学时：3

面向对象：本科生

预修课程要求：微积分或数学分析、线性代数、概率论与数理统计、数据结构基础、C语言

一、课程介绍（100-150字）

（一）中文简介：

计算摄影学是一个新兴的研究领域,旨在通过可计算的图像获取、处理和操纵技术,将软硬件有机结合起来克服传统数码相机的局限性,实现对图像能力的增强或扩展。在本课程中,我们将学习如何操纵和结合图片、视频来产生新的更好的图片、全景图、视频、3D模型、动画等等。

（二）英文简介

Computational photography is an emerging field that aims to combine software and hardware with computational image capture, processing and manipulation techniques to overcome the limitations of conventional digital cameras, so that the capabilities of digital photos can be to enhanced or extended. In this course, we will study ways of manipulating and combining photographs and videos to produce new and better pictures, panoramas, 3D models, animations, etc.

二、教学目的与基本要求

教学目的：我校目前虽然已有计算机视觉、图形学和图像处理方面的课程，但三者之间缺乏交叉和融合，而计算摄影学正好是计算机视觉、图形学和数字图像处理的深度交叉，因此开设本课程可以填补这一空白。本课程将使学生了解计算摄影的基本概念原理以及各种运用于图像和视频的计算技术，并介绍计算摄影学的最新研究方向和研究成果，可以开阔学生的研究视野，提高学生的创新能力，激发学生对科研工作的兴趣。

基本要求：掌握现代计算摄影学的最新进展和研究思路、掌握现代计算摄影学算法设计的常用技术和应用。

三、主要内容及学时分配

1. 计算摄影学概览 3 学时

主要内容：计算摄影旨在通过可计算的图像获取、处理和操纵技术,将软硬件有机结合起来克服传统数码相机的局限性,实现对图像能力的增强或扩展。本次课主要介绍传统摄像技术的局限性,计算摄影学涵盖的主题、概念和基本原理,其研究现状和发展趋势,并快速浏览一下后面的课程内容,介绍学习本课程所需要准备的东西,课后编程题和课程项目的要

求细节，考试的方式及要求等。

2. 图像的数字化、颜色空间、滤波与频域变换 3 学时

主要内容: 1) 介绍颜色的形成原理、基本属性(色调、饱和度、亮度)和颜色空间(RGB 颜色空间、CMYK 颜色空间、HSB 颜色空间、Lab 颜色空间、YUV 颜色空间、)以及不同颜色空间之间的相互转换; 2) 介绍图像的种类、图像的采样和数字化以及常见的几种图像数据格式; 3) 介绍常用的几种图像滤波(高斯滤波、均值滤波、中值滤波、双边滤波等)和对比; 4) 介绍图像的频域变换: 傅立叶变换。

学习目标: 让学生掌握颜色的形成原理、基本属性和常用的几种颜色空间和相互转换关系, 熟悉图像的种类、图像的采样和数字化以及常见的几种图像数据格式, 掌握几种常用的图像滤波方法和傅立叶变换并能编程实现。

课程作业: 编程实现几种常用的图像滤波方法和傅立叶变换。

3. 泊松图像编辑与交互式数字蒙太奇 3 学时

主要内容: 泊松图像编辑通过融合源图像梯度到目标图像的指定区域并结合目标图像的边界信息来求解泊松方程, 创造出平滑的融合图像。交互式数字蒙太奇是一个交互式、计算机辅助框架, 它将一组照片结合得到一张理想的图像。该技术整合了 Graph Cuts 和基于泊松方程的梯度域融合方法, 可以实现无缝的融合效果。本次课除了介绍这两项技术, 还将介绍常用 PDF(波方程, 热方程), 导向插值, 简单插值, 泊松方程, 稀疏线性方程组求解、Graph Cuts 数值优化方法等。

学习目标: 让学生了解图像编辑方法中常用的两种经典技术: 基于梯度域的融合和 Graph Cuts 缝隙优化方法, 以及它们的各自优缺点。了解并掌握常用 PDF(波方程, 热方程), 导向插值, 简单插值, 泊松方程, 稀疏矩阵和常用的存储格式, 稀疏线性方程组求解, Graph Cuts 数值优化方法等。

课程作业: 编程实现一个高斯赛德尔迭代法的稀疏线性方程组求解器。

4. 图像去模糊与非线性数值优化方法 3 学时

主要内容: 图像去模糊是一种消除由于场景中物体运动较快、图像的深度或者是相机的抖动问题造成图像过于模糊的方法。图像去模糊是在求解一个病态问题, 需要借助一些先验信息来求取逼近于原始图像的解。本次课将首先介绍图像去模糊的基本原理, 模糊核概念, 以及几种分类(按模糊核的性质可以划分为盲去卷积和非盲去卷积; 按模糊的区域划分可以分为全局模糊和局部模糊; 按信息源分类可以划分为单幅图像去模糊和带辅助信息的图像去模糊); 然后对一些经典和最新的图像去模糊算法进行介绍, 并详细介绍这些方法中采用的函数模型和相应的非线性数值优化方法。

学习目标: 让学生系统地了解图像去模糊的概念、基本原理、用途和分类, 了解几种经典的图像去模糊方法和它们之间的对比; 熟悉并掌握图像去模糊中常用的几种非线性数值优化方法, 比如最小二乘法、牛顿法、Levenberg-Marquardt 算法、Total Variation 方法等。

课程作业: 要求编程实现最小二乘法、牛顿法和 Levenberg-Marquardt 算法。

5. 特征匹配与光流 3 学时

主要内容: 特征匹配和光流就是对两帧图像中的对应进行检测和匹配, 根据匹配的稠密程度可以分为稀疏对应和稠密对应。本次课程会对一些较为经典和常用的特征点和描述量(例如 Harris 角点、Fast 角点、SIFT、SURF、HOG 等)进行介绍, 并介绍相对应的匹配

方法；同时会对较为经典和最新的光流估计算法进行介绍，包括 HS、LK、MDP 等。

学习目标：让学生了解特征匹配和光流的基本概念和原理，以及二者的异同之处；了解并掌握一些常用的特征点和描述量，以及相对应的匹配方法；了解并掌握最基本的光流公式，对 HS、LK、MDP 等光流算法有所了解。

6. 相机模型与运动推断结构

3 学时

主要内容：运动推断结构就是从移动摄像机拍摄的图像或视频序列中自动地恢复出场景的三维结构和摄像机的内外参数。本次课的主要内容为：1) 介绍针孔相机模型、透视相机模型、仿射变换、平面变换和 Homography；2) 介绍基本的双视图和基础矩阵求解方法；3) 介绍多视图几何原理和运动推断结构的传统框架，以及摄像机姿态五点法求解、自定义标和集束调整等内容；4) 介绍主流的几个开源或免费软件（ACTS、bundler、Visual SFM 等）并做对比。

学习目标：让学生了解并掌握常见的相机模型和平面变换，了解并掌握双视图几何和基础矩阵的求解；了解多视图几何原理以及运动推断结构的基本流程和框架；了解主流的几个开源或免费软件并做对比。

7. 同时定位与地图构建

3 学时

主要内容：同时定位与地图构建简称 SLAM (Simultaneous Localization and Mapping)，可以用于在未知环境下实时恢复摄像机运动参数和场景的三维结构。本次课的主要内容为：1) 介绍 SLAM 技术的基本原理、分类（基于滤波器的方法和基于关键帧的方法）和比较；2) 介绍开源软件 PTAM 以及主流商业软件开发包。

学习目标：让学生了解 SLAM 技术的基本原理和分类；了解基于滤波器的方法和基于关键帧的方法的各自原理、框架和对比；让学生了解目前 SLAM 方面的开源软件和主流商业软件开发包，能基于这些软件开发包开发简单的增强现实应用程序。

课程作业：要求学生使用开源或免费的 SLAM/AR SDK，写一个增强现实的 App。

8. 单视图与多视图三维重建

3 学时

主要内容：基于单视图与多视图的三维重建就是利用同一场景的一幅或多幅图像恢复出场景的三维模型。本次课的主要内容为：1) 介绍基于单视图重建的基本原理（包括灭点、灭线的概念）以及经典的求解方法；2) 介绍双视图立体匹配原理、方法分类和对比；3) 介绍多视图立体三维重建的分类和比较。

学习目标：让学生了解单视图重建的基本原理和方法，包括灭点和灭线概念；了解双视图立体匹配原理，以及常见的几种局部优化方法和全局优化方法；了解多视图立体三维重建的分类、各类方法的优缺点。

9. 图像补全

3 学时

主要内容：图像补全就是修复图像中的缺失部分，也可用于移除或替换某些选定的物体。本次课主要对一些经典和最新的图像补全算法做一下介绍，包括代表性的局部优化算法、全局优化算法和交互式的图像补全算法等。

学习目标: 让学生了解图像补全的概念、基本原理和用途, 以及常用的几种方法和各自的优缺点; 要求掌握最基本的基于样本采样的局部优化补全算法, 对全局优化方法和交互式补全做到了解即可。

10. 交互式图像分割与抠像 3 学时

主要内容: 图像分割指的是将数字图像中感兴趣的区域分割出来, 而抠像则除了要得到前景轮廓外还要进一步估计半透明通道, 以实现真实的合成。本次课主要讲述高斯混合模型、马尔科夫随机场、以及经典的 Grab Cuts、Lazy Snapping 等交互式图像分割算法, 以及贝叶斯抠像、Closed Form 抠像等算法。

学习目标: 让学生了解图像分割和抠像的概念和区别; 了解并掌握图像分割/抠像中经常用到的高斯混合模型和马尔科夫随机场; 了解经典的 Grab Cuts、Lazy Snapping、贝叶斯抠像、Closed Form 等图像分割和抠像算法。

11. 全景图拼接 3 学时

主要内容: 全景图相比于普通透视图像有更广的视角 (甚至 360 度), 可以尽可能多地表现出周围的环境。它可以由多张图像拼接而成。本次课主要讲述全景图的概念和类型 (包括柱面全景图和球面全景图), 几种有代表性的单视点全景拼接算法、自动全景图识别方法以及最新具有代表性的多视点全景图拼接算法和对比。

学习目标: 让学生了解传统全景图的概念、类型 (柱面全景图和球面全景图) 和用途, 透视图与全景图之间的坐标转换关系, 以及常见的几种全景图生成方法; 了解多视点全景图的概念, 以及几种代表性的方法和对比。

课程作业: 按照要求编程实现一个柱面全景图拼接算法, 能够将输入的图像序列自动地拼接成柱面全景图。

12. 上色与重上色 3 学时

主要内容: 上色就是给单色图片添加颜色的过程, 而重上色则是调整图片颜色。本次课的主要内容为: 1) 在了解基本颜色空间与图像种类的基础上, 介绍上色与重上色的基本概念、基本原则 (保持真实性, 自然性); 2) 介绍上色与重上色问题的难点 (光照、阴影); 3) 介绍经典的、最新的上色与重上色方法 (利用样本进行上色重上色、交互式的上色重上色、单源与多源的上色重上色)。

学习目标: 让学生掌握上色与重上色的基本原则, 熟悉不同颜色空间对上色重上色的效果, 熟悉不同上色方法的优缺点以及上色与重上色问题的难点, 掌握最经典的上色方法。

13. 纹理合成与图像缩放 3 学时

主要内容: 纹理合成就是根据样本纹理生成相似纹理的过程。经典的纹理合成算法估计纹理特征空间的概率分布, 再在需要生成新纹理的区域根据最大后验生成最佳的纹理像素。准确计算概率分布是费时的, 因此有很多经典的加速算法, 比如基于 Tree-structured Vector Quantization 的采样加速。图像缩放就是改变图像大小的过程, 为了避免缩放时发生模糊或者丢失细节, 可以采用各种图像插值来提高结果质量。另外像 Seam Carving 等算法可以在非等比缩放图像时保持图像的关键内容不发生比例失真。本次课将介绍纹理合成和图像缩放这两类问题和相关的应用, 并介绍几个经典的算法。

学习目标：让学生了解纹理具有的结构或统计特性，从而了解纹理合成的基本原理。通过学习图像缩放，学生可以了解到各种常用的图像插值算法，也将学习到非等比缩放带来的新问题。除了基本原理，学生还将学习到一些有代表性的算法，比如利用 TSVD 加速采样的纹理合成和利用动态规划求解最优 Seam 的 Seam Carving。

14. 课程讨论与项目答辩 3 学时

主要内容：每个组轮流做一下项目报告和 demo 演示。

15. 课程讨论与项目答辩 3 学时

主要内容：每个组轮流做一下项目报告和 demo 演示。

16. 课程讨论与项目答辩 3 学时

主要内容：每个组轮流做一下项目报告和 demo 演示。

四、相关教学环节安排：

1) 采用多媒体投影教学（讲授核心内容、总结、按顺序提示今后内容、答疑、公布讨论主题等）；

2) 实验室课单列，每周 2 学时；

3) 阅读与每节课内容相关的论文；

4) 布置课后思考题和课后编程作业，要求学生按时完成；

5) 编程项目展示：学生根据教学内容自选一个中等程度规模以上的 project，完成算法和程序设计(会提供程序框架和测试数据，学生只要按规定完成相应的模块即可)，最后提交完整 demo 和项目报告，要求独立或分组（不超过 3 人）合作完成。

五、教学方式：

课堂讲授为主，课程编程作业和大型实验作业为辅。

六、考试方式及要求：

本门课程的评分分为 3 个部分，每个部分分数分配如下：1) 课程作业与上课出勤：50%；2) 项目设计和演示：40%；3) 项目课堂答辩：10%。

七、参考教材及相关资料

教材：

【1】“Computer Vision: Algorithms and Applications” by Richard Szeliski. Publisher: Springer; 2011 edition (November 24, 2010)

【2】“Computational Photography: Mastering New Techniques for Lenses, Lighting, and Sensors” by Ramesh Raskar and Jack Tumblin. Publisher: A K Peters/CRC Press; 1 edition (December 26, 2014)

相关资料：

CVPR/ICCV/ECCV/Siggraph/Siggraph Asia 论文集, 2002-2014.

六、课程教学网站:

将通过校内网络提供必要的课件和文字材料链接