



CS 330 MIP – Lecture 01

多媒体信息处理介绍

Multimedia Information Processing Introduction

Jimmy Liu 刘江

2025-02-19

CS 330 Module History

The multimedia information processing course is a professional course offered by the Department of Computer Science and Engineering at the Southern University of Science and Technology of China for undergraduate students majoring in Computer Science and Engineering. Since the spring semester of 2000, It has been offered for 6 semesters.

CS 330 Module Coverage

What Will CS 330 Cover?

MM
Concepts

MM
Algorithms

MM
Technology

MM
Applications

MM
Projects

TextBook



MOOC

Med 南方科技大学

多媒体信息处理

七个里程碑

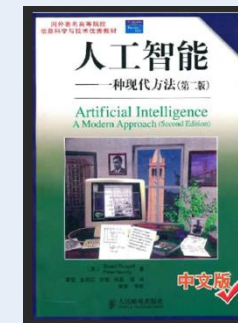
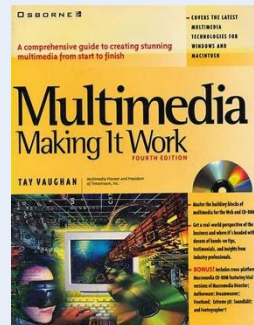
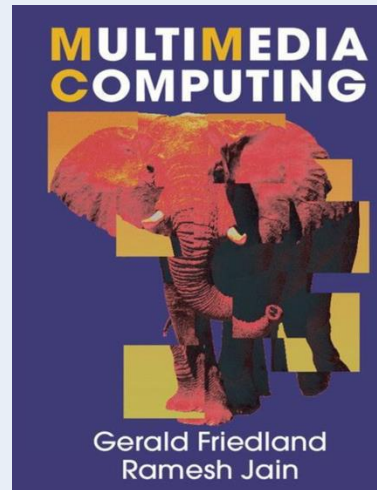
年份	概念/方法
1954年	词袋模型 (Bag-of-Words) 概念
1972年	词频-逆文档频率 (TF-IDF) 计算方法
1990年	潜在语义分析 (LSA) 算法

MOOC



长按识别二维码

Study References



2025 CS330 TAs



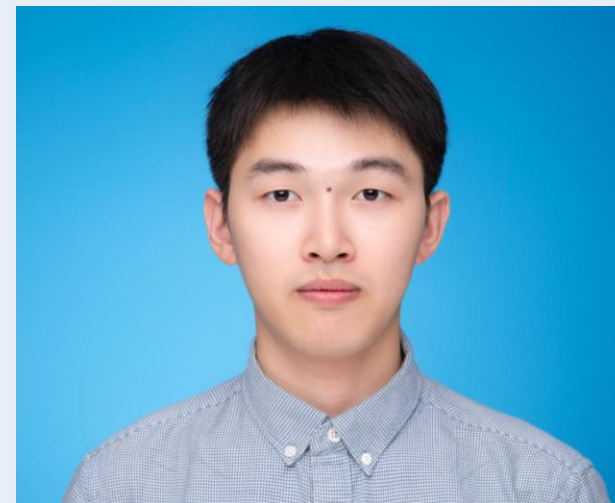
聂秋实



巫晓



魏嘉祺



曾令玺

Study Objectives and Assessment

Help students understand

the

Concepts,

Algorithms

Applications

of

multimedia technologies



 **bjectives**

1

Study Objectives and Assessment

Learn media modalities

Text

Digital audio,

Image

Animation

Graphics

Digital video



bjectives

2

Study Objectives and Assessment

Learn the
Algorithms
Math foundation
Compression
Processing algorithms
for
MM processing



8.9M



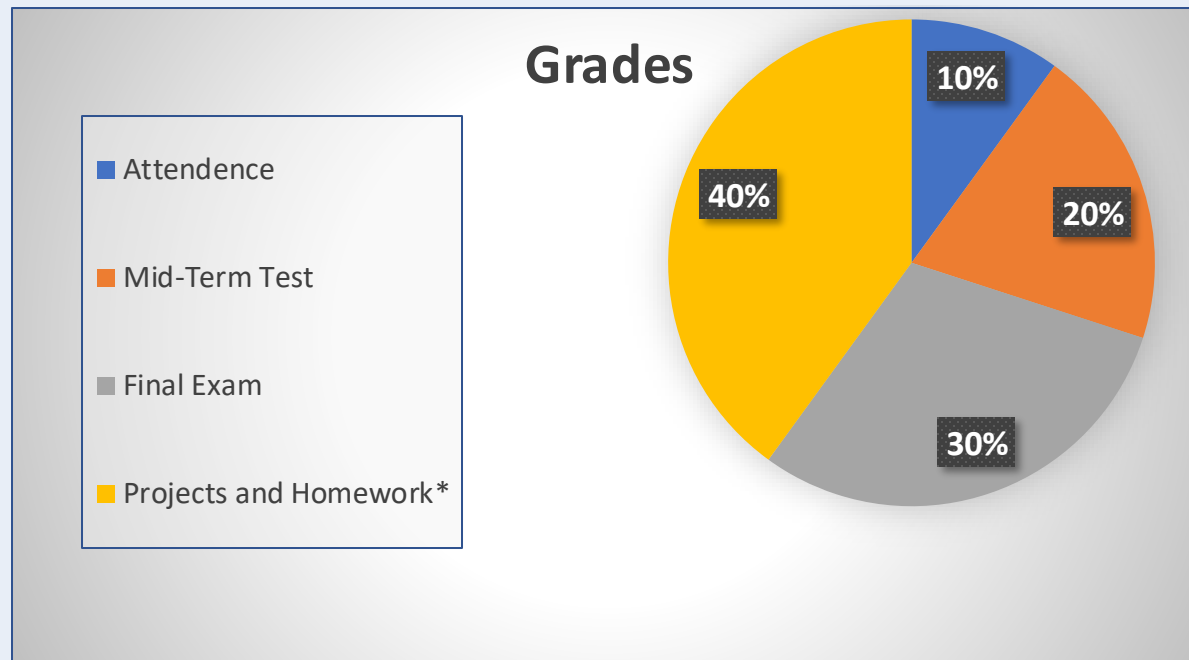
68.34K



bjectives

3

Study Objectives and Assessment



Extra Bonus up to 10% maybe given to students done excellent in the project or homework

* Late submission will not be accepted.

Assessment - Attendance and Schedule

Multimedia Information Processing

The Course is conducted at:

The First Teaching Building, Room 324
(第一教学楼324教室)

Every Wednesday at morning section 3-4
10:20am-12:10am

Experiment at:

The First Teaching Building, Room 324
(第一教学楼324教室)

Every Wednesday at morning section 5-6
14:00pm-15:50pm



Assessment - Plagiarism

1

No Plagiarism is allowed

2

If plagiarism on project is found for the first time:

- the plagiaristic part is graded as 0 and
- warning is given to the students

3

If plagiarism is found for the second time

- the course is graded as 0



Assessment - Plagiarism

4

For project report, any sentence that is copied from other paper or article should cite the original source as the reference, otherwise, the report is considered as plagiarism

5

Please read the computer science department document on the Blackboard CS 330 Course Website, we adopt this rule in our class.



Rules on Plagiarism

已附加文件:  计算机系本科生作业和学位论文抄袭认定标准及处理办法.docx (87.098 KB)

Please read the rule on plagiarism of this course

How Are Lectures Organized:

Systematic Study From Easy to Difficult Rather Than Uniform



First 3 Weeks: General Introduction

- – Build Your Interest and Inspiration



4-7th Weeks: Fundamental Concepts

- – Study The Building Blocks of MM



8th Week Mid-Term Test

- – Time to Rethink



9-16th Weeks: MM Technology and Algorithms

- – Time to Study Hard

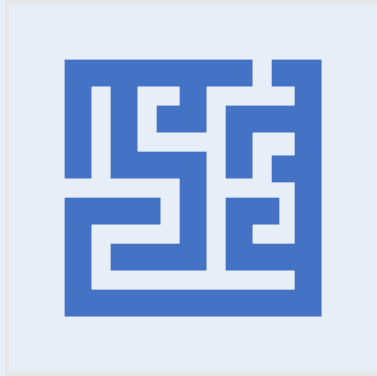
Let us form a Wechat Group

群聊：CS 330 2025 多媒体信息
处理课程群



How To Study OUR CS 330 Module

- How Much You Are Inspired to Learn by Yourself



AI+ and Multimedia System Literacy



Collaborative Study

How To Study OUR CS 330 Module?



IP+



M

How To Study OUR CS 330 Module



M+



IP

How To Study OUR CS 330 Module

-Suggested Way of Studying CS 330

Study in advance








Reading even a page in advance prevents you from cramming and rushing in learning your lessons.



**PROJECT
DRIVEN
CREATION**




Message 1: Study in Advance- CS 330 Blackboard Website and Syllabus

-  Lecture slides and additional notes
-  Course Notice
-  Assessments, coursework and exam
-  Textbooks, reading list
-  Other resources
-  ...
-  Check Before Lectures to Study in Advance

多媒体信息处理 (2025春)
CS330 Multimedia Information Processing 多媒体信息处理

CS330 Multimedia Information Processing 多媒体信息处理
CS330 Lecturer 课程教师-刘江 Jimmy Liu
CS330 教材与慕课
CS330 Reference Book 课程参考书
诚信承诺书 (Rules on Plagiarism)
课程内容 Course Contents
互动专区

CS330 Multimedia Information Processing 多媒体信息处理


Course Introduction 课程介绍

This is the course website of 2025 CS 330 "[Multimedia Information Processing 多媒体信息处理](#)".

The multimedia information processing course is a professional course offered by the Department of Computer Science and Engineering at the South University of Science and Technology of China for undergraduate students majoring in computer-related majors. Since the spring semester of 2000, six semesters have been offered.

The Course is conducted at: First Teaching Building, Room 324 (第一教学楼324教室)
Every Wednesday at morning section 3-4, 10:20am-12:10pm
experiment at: First Teaching Building, Room 324 (第一教学楼324教室)
Every Wednesday at morning section 5-6, 14:00am-15:50pm

The [course syllabus](#) will cover:

Message 1: Study in Advance-

CS 330 MOOC



长按识别二维码

Message 2: Do 2025 MIP Projects Well

AR探月 梦想启航 | 我校本科生团队与四川美院共同研发载人月球车增强现实显示系统

SUSTech-CSE 南科大计算机系

2024年06月13日 21:28 5人

中国探月工程的进展带动了大学生对月球的兴趣，南方科技大学四位本科生—高祖卿、游俊涛、施永祺、纪可鸣在《多媒体信息处理》课程上选择了探月主题的课程



Project 9: 载人月球车增强现实显示系统



项目背景：与四川美术学院合作，利用AR模拟在月球上行驶的月球车，需要显示的内容包括电池信息，行驶信息，车辆故障信息，路面障碍等，模拟现场环境和数据源。

项目目的：基于二维图像利用AI生成3D场景，并结合AR进行展示和交互。月球车的数字模型已初步完成。

项目要求：

- 利用AI实现升维，基于二维图像生成三维场景
- 采用Unity+Vuuforia平台搭建
- 可在手机、PC、HoloLens等多平台通用，并实现交互功能



图1. 月球车展示图



图2. 360全景视野



图3. 场景模拟预期

2025 CS330 项目

1. 低视力视觉增强辅助：Retissa 视网膜投影显示技术的创新应用（睿翎）
2. 低视力视觉增强辅助：HoloLens MR 头显技术的创新应用（睿翎）
3. 面向低视力患者的视功能评估技术（姚成林）
4. 基于图像的室内建图与导航（载道）
5. 盲人出行的导盲设备人机交互方式研究（新博）
6. 基于连续视频信息的盲人出行障碍物目标跟踪与检测（新博）
7. 基于MOOC的个性化教育智能体（佳璐）
8.

课题5：盲人出行的导盲设备人机交互方式研究

- 背景：盲人出行时如何从导盲设备中获取相关环境信息，并能通过语音/遥控方式完成与导盲设备的交互过程。
- 任务目标：设计一种盲人出行时导盲设备与盲人的交互方法，要求：导盲设备通过相机/雷达来获取周围环境信息，通过语音播报提示的方式来告知盲人周围环境变化；盲人可通过语音/遥控的方式来输入需求指令，完成与导盲设备的交互过程。

课题5：盲人出行的导盲设备人机交互方式研究

- 需求分析：
 1. 导盲设备自动检测障碍物并识别其类别
 2. 障碍物信息（种类、方向）转换为语音进行播报
 3. 针对不同障碍物进行避障策略提示（向左/向右/等待）
 4. 盲人通过输入设备（遥控/语音）与导盲设备进行简单交互

课题6：基于连续视频信息的盲人出行障碍物目标跟踪与检测

- 背景：导盲是一个动态过程，导盲设备与周围障碍物位置会实时产生变化。静态的图像只能捕获当前时刻的障碍物，可能存在视野盲区，而这对于盲人的出行安全是至关重要的。因此，在导盲过程，需要基于连续的视频信息对障碍物进行跟踪与检测，以实现全目标的综合检测，避免离散静态图像信息下导盲方案存在的安全隐患。
- 任务目标：基于导盲设备采集的连续视频进行障碍物检测并进行跟踪，利用帧间障碍物移动距离以及盲人自身的移动距离等轨迹预测，来综合判断障碍物在当前帧是否仍然存在以及是否被正确检出，如果未被检出，需要在当前帧进行补全。

课题6：基于连续视频信息的盲人出行障碍物目标跟踪与检测

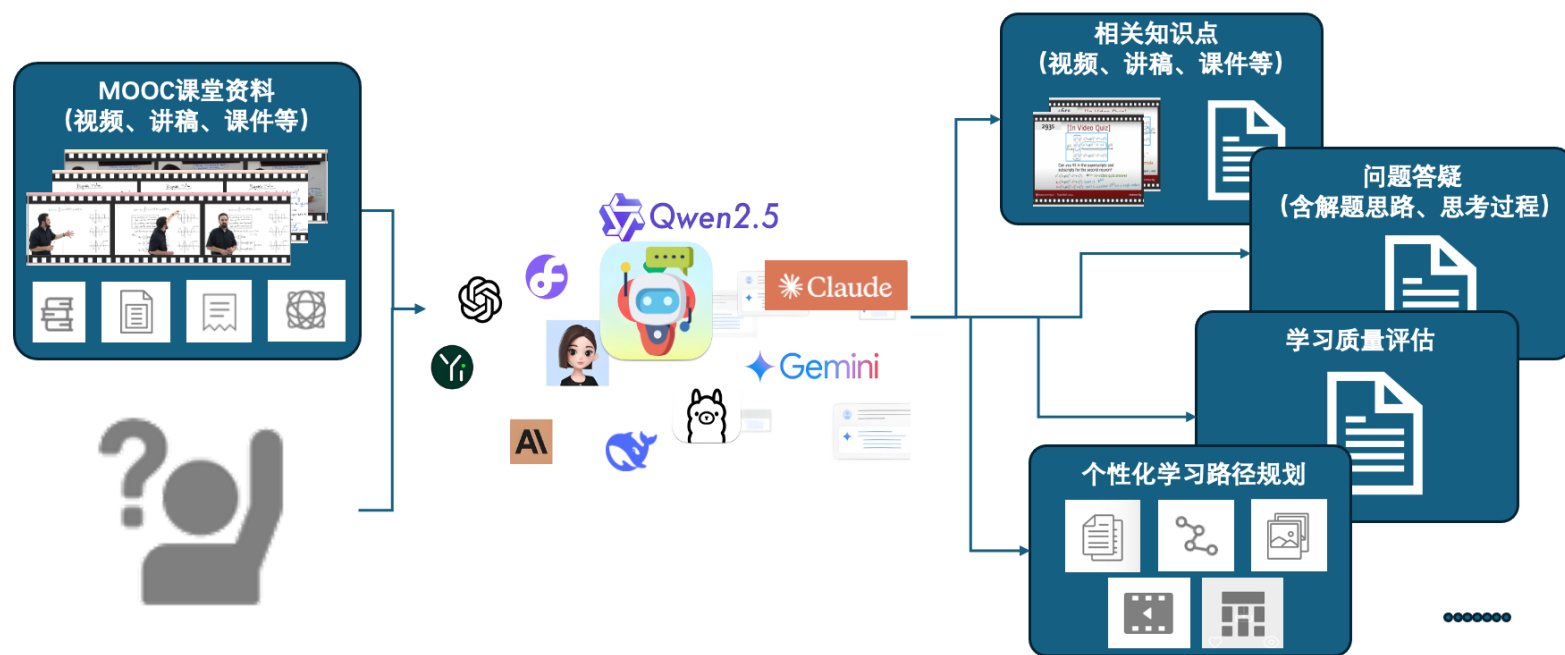
- 需求分析：
 1. 实现盲人出行障碍物的检测
 2. 对画面中出现的障碍物进行目标跟踪，获知其运动方向与轨迹，以补缺当前帧丢失的障碍物目标
 3. 利用目标跟踪结果对目标运动轨迹进行轨迹预测
 4. 根据轨迹预测结果，对可能与盲人发生碰撞的障碍物进行避障提示

课题8： 基于MOOC的个性化教育智能体

- 背景：在数字化教育迅猛发展的时代背景下，MOOC（大规模开放在线课程）以其开放性、普惠性和可扩展性，成为全球教育革新的重要载体。然而，当前在线MOOC教育存在“重资源供给、轻学习支持”的问题。比如，现有MOOC平台主要依赖于异步交互模式（如讨论帖），学生的问题不能及时得到解答。再比如，虽然课程讲稿、课件等静态文本可实现知识点溯源，但视频作为MOOC核心载体，知识点定位困难，学生需反复跳转寻找目标片段，耗时低效，学习效率低下。

课题8： 基于MOOC的个性化教育智能体

- 任务目标：开发多模态在线教育智能体，通过融合视频、课件、习题等多源数据，借助多模态大模型、知识图谱与推理增强技术，构建具备实时答疑与动态辅导能力的智能学习支持系统，全面优化MOOC学习体验，提升学习效果。



Message 3: Master the Latest MIP Technologies



Cool Multimedia Technology Areas



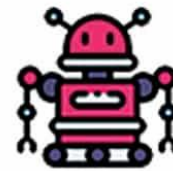
Artificial
intelligence



Augmented
reality



Blockchain



Chatbots



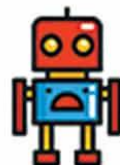
Digital
assistants



Face
recognition



Machine
learning



Robotics



Sensors/IoT



Virtual reality



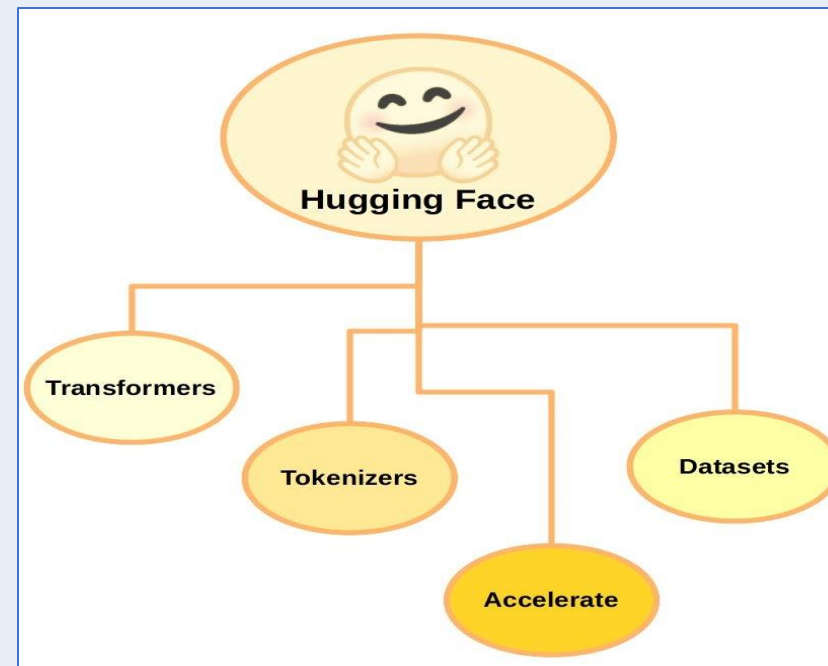
Voice
interfaces



4DX

1: Multimedia LLM AI

- Generative AI, especially large language models with multimedia capabilities, have expanded their application range from text generation to image, audio, and video generation. This technology is transforming fields such as education, entertainment, healthcare, and industry by providing more complex interaction scenarios.
- Hugging Face Transformers, which provides tools for developing and training multimodal AI models. DeepSeek, GPT-4 and potential future versions like GPT-5 (if officially released) also provide rich multimedia support

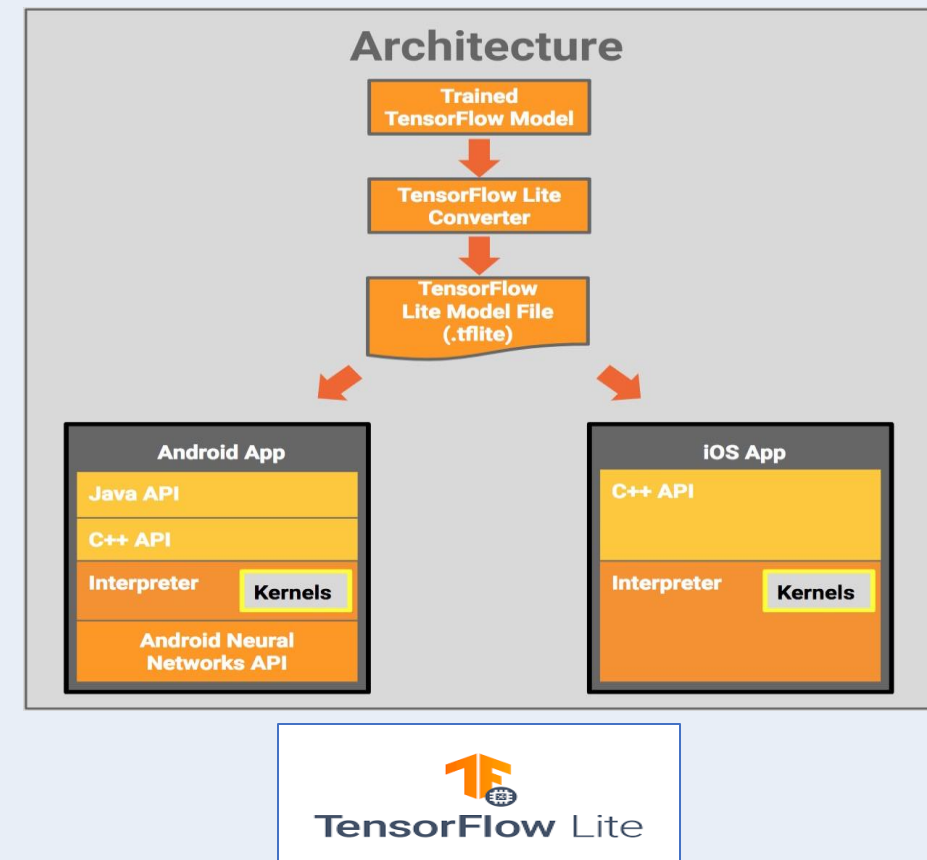


```
# Load model directly
from transformers import AutoTokenizer, AutoModelForCausalLM

tokenizer = AutoTokenizer.from_pretrained("meta-llama/Llama-3.3-70B-Instruct")
model = AutoModelForCausalLM.from_pretrained("meta-llama/Llama-3.3-70B-Instruct")
```

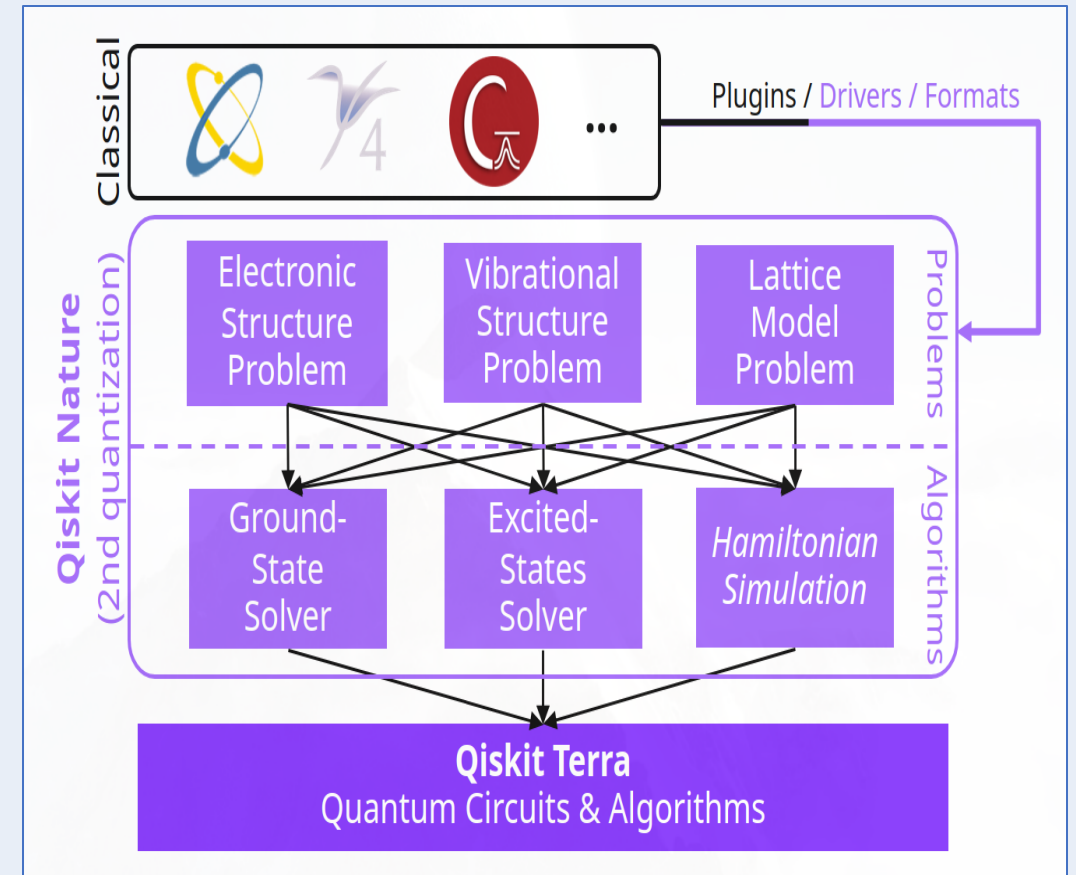
2: Multimedia Edge Intelligence (Edge AI)

- Edge AI brings the power of AI closer to the data source, enabling real-time multimedia data processing and reducing latency. This is particularly critical for IoT devices, enhancing user experience.
- TensorFlow Lite, an AI framework that supports edge devices, is a pivotal technology in this domain.



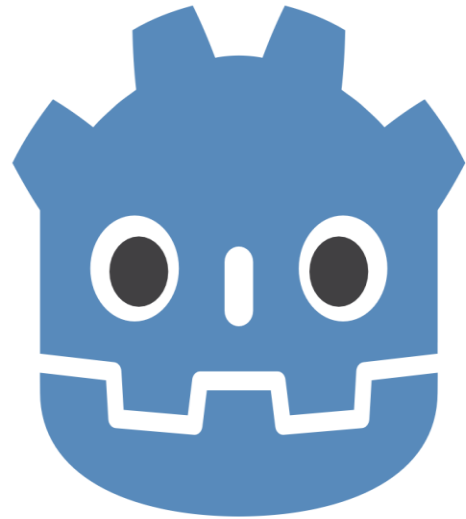
3: Multimedia Quantum Computing

- Quantum computing has achieved breakthroughs in algorithm optimization and hardware performance, demonstrating unique advantages in fields such as multimedia chemical simulation and financial risk analysis.
- Frameworks like Qiskit (provided by IBM) and Cirq (open-sourced by Google) are driving the development of quantum computing algorithms.



4: Virtual Reality (VR) and Augmented Reality (AR)

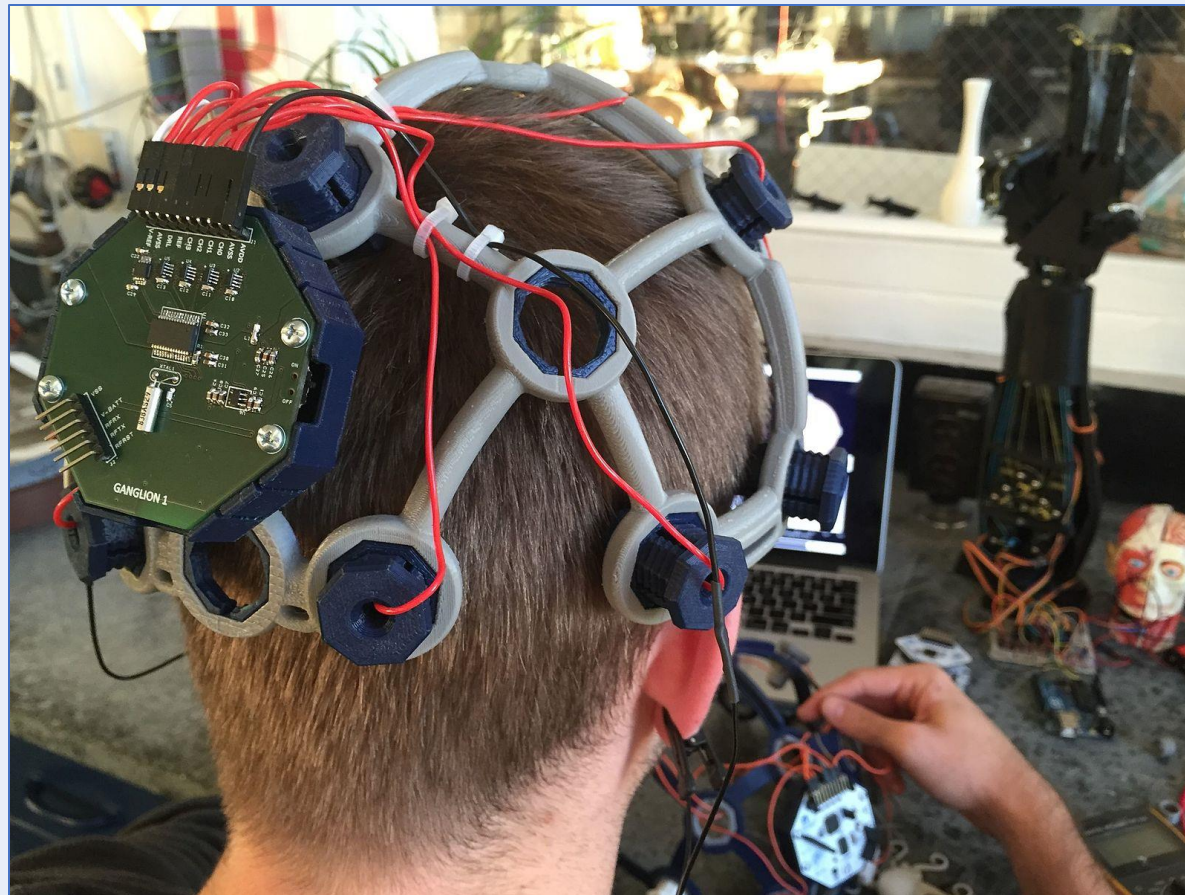
- VR/AR technologies are revolutionizing multimedia content creation by providing immersive experiences with AI-Powered Avatars. These technologies are widely used in video production, education, and entertainment.
- Tools and engines such as free, open-source 2D and 3D Godot game Engine, which supports AR/VR development, are essential for creating these immersive experiences.



GODOT
Game engine

5: Multimedia BCI

- BCI technology as a new multimedia type, has the potential to revolutionize human-computer interaction by allowing users to control devices using their brain signals.
- OpenBCI, which provides open-source hardware and software tools for multimedia BCI development, it can be used to measure and record multimodal electrical activity produced by the brain (EEG), muscles (EMG), and heart (EKG),



Homework 01 (PPT) Submitted By Sunday Night— Template Will be Provided on BB (中英文都可以)

1

Tell me about yourself, what and how do you like to study from MIP 330

2

Produce one your favorite AIGC content with prompts



CS 330 MIP – Lecture 01

多媒体信息处理介绍

Multimedia Information Processing Introduction

Jimmy Liu 刘江

2025-02-19