



《虚空间专业概述》

|  |  |
| --- | --- |
| **专 业** | **计算机科学** |
| **姓 名** | **赵天骜** |
| **日 期** | **2022/9/8** |

**摘 要**

虚拟空间是一种对“元宇宙”更为学术性的称呼，利用科技手段进行链接与创造的，与现实世界映射与交互的虚拟世界，具备新型社会体系的数字生活空间。如今科技的发展还局限于对于图像处理和运动物体识别之中，将现实与虚拟的数据进行结合，从而指导生活和科学实践。

**目 录**

[第1章 专业研究内容介绍 - 1 -](#_Toc10873)

[1.1 Metaverse的出现及发展历史 - 1 -](#_Toc28278)

[1.2 Metaverse的技术基础 - 1 -](#_Toc15173)

[第二章 核心技术介绍 - 2 -](#_Toc12210)

[2.1 计算机视觉技术 - 2 -](#_Toc4687)

[2.1.1 图像增强 - 2 -](#_Toc31100)

[2.1.2 图像匹配 - 3 -](#_Toc27353)

[2.2 卷积神经网络 - 4 -](#_Toc26185)

[2.3 虚拟和增强现实 - 4 -](#_Toc30933)

[第三章 研究及发展现状 - 5 -](#_Toc6215)

[结 论 - 6 -](#_Toc17310)

[参考文献 - 7 -](#_Toc14280)

# 第1章 专业研究内容介绍

## Metaverse的出现及发展历史

元宇宙（ Metaverse），于1992年在一本小说《雪崩》中首次被提出。小说描绘了一个庞大的虚拟现实世界，在这里，人们用数字化身来控制，并相互竞争以提高自己的地位。时至今日，小说中描述的场景仍是一种对于未来的想象。随着计算机科学与技术的发展，虚空间，或者说元宇宙，已经成为一系列科学技术的代名词，在计算机视觉，人工智能，AI算法等一些列科学技术的帮助下，人们可以逐步将现实与虚拟数据空间进行融合，实现对现实的增强。

虚空间的发展史：

* 1983年1月1日，阿帕网和国防数据网络正式改为TCP/IP标准，互联网就此诞生。
* 1992年，Metaverse首次被提出
* 1998年，B-money出现
* 2009年，Bitcoin出现
* 2009年，区块链技术出现
* .......
* 2016年，Pokemon GO
* 2021年，微软网格
* 2021年，元平台出现（eg.FaceBook）

## Metaverse的技术基础

在技术视角下，技术意义的元宇宙包括了内容系统、区块链系统、显示系统、操作系统，最终展现为超越屏幕限制的 3D 界面，所代表的是继 PC 时代、移动时代之后的全息平台时代。支持元宇宙的技术集群包括 5 个板块：

其一，网络和算力技术：包括空间定位算法、虚拟场景拟合、实时网络传输、GPU 服务器、边缘计算，降低成本和网络拥堵；

其二，人工智能；

其三，电子游戏技术：如支持游戏的程序代码和资源（图像、声音、动画）的游戏引擎；

其四，显示技术：VR、AR、MR，特别是 XR，持续迭代升级，虚拟沉浸现实体验阶梯，不断深化的感知交互；

其五，区块链技术：通过智能合约，去中心化的清结算平台和价值传递机制，保障价值归属与流转，实现经济系统运行的稳定、高效，透明和确定性。

# 核心技术介绍

## 2.1 计算机视觉技术

计算机视觉技术（computer vision）是虚拟空间中最重要的技术之一。其核心目的与功能就是为了识别物体和人类行为。为了实现元宇宙等一系列构想，如何将现实世界虚拟化，如何将现实搬入到数据空间之中等一些列问题变得尤为重要。因此，计算机视觉技术顺理成章的成为整个科学技术集合的基石。本章主要介绍两种计算机视觉技术，图像增强和图像匹配。

### 2.1.1 图像增强

要了解图像增强技术，不得不提到图像在计算机中的存储方式。图像在计算机中实际上是一系列强度值所构成的矩阵。图像可以看成为一种映射，，其中，代表图像上每一个像素点在图像所在平面上的坐标位置，表示该点所对应的灰度值。

图像增强技术用于增强图像中的有用信息，它可以是一个失真的过程，其目的是要改善图像的视觉效果，针对给定图像的应用场合。它通过有目的地强调图像的整体或局部特性，将原来不清晰的图像变得清晰或强调某些感兴趣的特征，扩大图像中不同物体特征之间的差别，抑制不感兴趣的特征，使之改善图像质量、丰富信息量，加强图像判读和识别效果，满足某些特殊分析的需要。

图像增强可分成两大类：频率域法（频域法）和空间域法（空域法）。

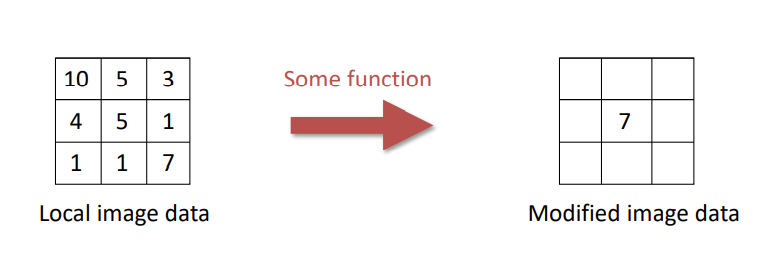
前者把图像看成一种二维信号，对其进行基于二维傅里叶变换的信号增强。采用低通滤波（即只让低频信号通过）法，可去掉图中的噪声；采用高通滤波法，则可增强边缘等高频信号，使模糊的图片变得清晰。

后者空间域法中具有代表性的算法有局部求平均值法和中值滤波（取局部邻域中的中间像素值）法等，它们可用于去除或减弱噪声。

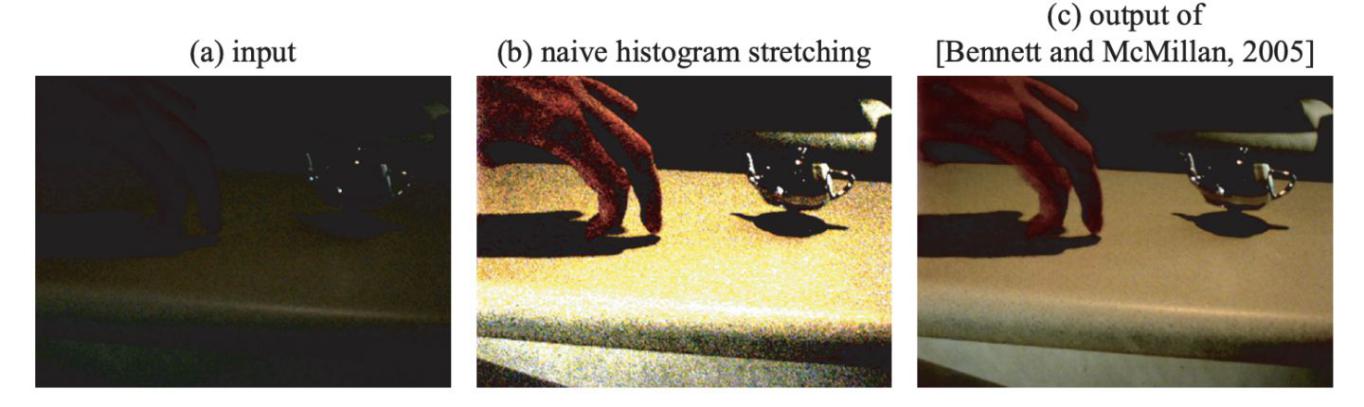
图像灰度处理：图像灰度是将彩色图像转换为灰度图像的过程。由于R、G和B三个元素在每个像素中混合了不同的颜色，并且每个元素有256 个值，因此旋转次数超过1600万次(256 x 256 x 256) 每像素。灰度图像是很特殊的彩色图像，其中每个像素大小为256 个像素，所有元素R、G和B相等。通常，许多彩色图像格式都被转换为灰度图像，以提高后续图像处理的效率。与彩色图像一样，灰度图像也可以表示图形细节特征。

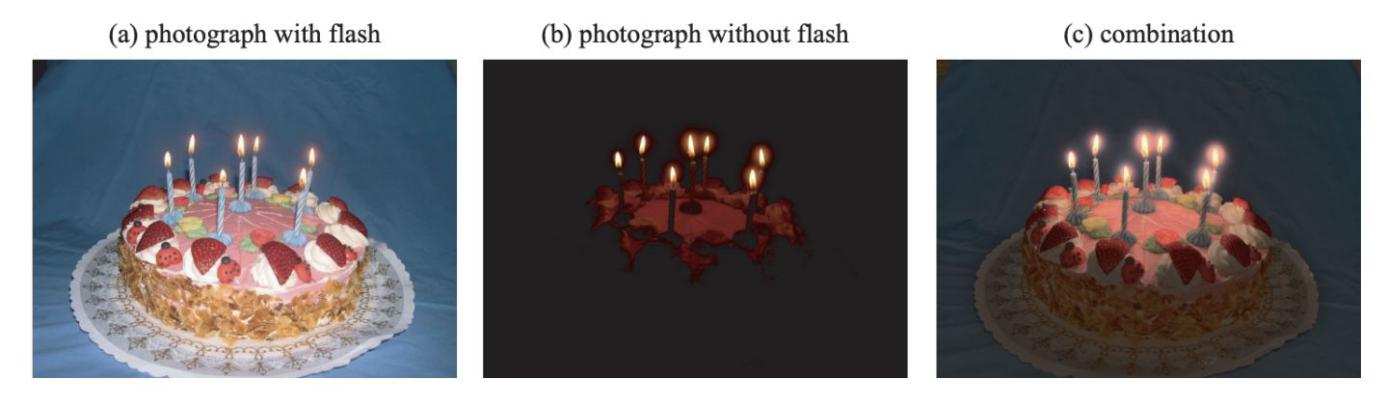
图像二值化：二值化是指通过选择合适的边界，将灵敏度锐利的灰度图像转换为只有二维灰度值的二值图像。经过这样的改变，虽然去除了背景色，但改变后的灰度图像仍然可以代表图像的一般特征。使用距离与背面的差值，使用分割阈值将不同灰度的背面和背面组合为一张图片，选择最佳阈值来区分目标和背景区域中的每个像素，然后形成二值图像，就可以得到待检测的目标。

图像滤波：图像滤波是在图像增强中常用的处理方式。图像通常有一个直观特征: 图像往往由多块组成，各块内像素相似且过渡缓慢，块与块相邻部分称作边缘。 图像滤波希望实现的是图像块内平滑降噪、去细节，并最大程度保留图像边缘。 图像滤波具备去除噪音、去除图像细节、保持图像边缘的能力，因此被广泛使用。根据每个像素的局部邻域的某些函数修改图像中的像素，如：



一些直观的例子：





### 2.1.2 图像匹配

图像匹配是指通过一定的匹配算法在两幅或多幅图像之间识别同名点，如二维图像匹配中通过比较目标区和搜索区中相同大小的窗口的相关系数，取搜索区中相关系数最大所对应的窗口中心点作为同名点。其实质是在基元相似性的条件下，运用匹配准则的最佳搜索问题。

图像匹配主要可分为以灰度为基础的匹配和以特征为基础的匹配。

灰度匹配的基本思想:以统计的观点将图像看成是二维信号，采用统计相关的方法寻找信号间的相关匹配。利用两个信号的相关函数，评价它们的相似性以确定同名点。灰度匹配通过利用某种相似性度量，如相关函数、协方差函数、差平方和、差绝对值和等测度极值，判定两幅图像中的对应关系。最经典的灰度匹配法是归一化的灰度匹配 法，其基本原理是逐像素的把一个以一定大小的实时图像窗口的灰度矩阵，与参考图像的所有可能的窗口灰度阵列，按某种相似性度量方法进行搜索比较的匹配方法，从理论上说就是采用图像相关技术。利用灰度信息匹配方法的主要缺陷是计算量太大，因为使用场合一般都有一定的速度要求，所以这些方法很少被使用。现在已经提出了一些相关的快速算法，如幅度排序相关算法，FFT相关算法和分层搜索的序列判断算法等。

特征匹配是指通过分别提取两个或多个图像的特征（点、线、面等特征），对特征进行参数描述，然后运用所描述的参数来进行匹配的一种算法。基于特征的匹配所处理的图像一般包含的特征有颜色特征、纹理特征、形状特征、空间位置特征等。特征匹配首先对图像进行预处理来提取其高层次的特征，然后建立两幅图像之间特征的匹配对应关系，通常使用的特征基元有点特征、边缘特征和区域特征。 特征匹配需要用到许多诸如矩阵的运算、梯度的求解、还有傅立叶变换和泰勒展开等数学运算。常用的特征提取与匹配方法有：统计方法、几何法、模型法、信号处理法、边界特征法、傅氏形状描述法、几何参数法、形状不变矩法等。基于图像特征的匹配方法可以克服利用图象灰度信息进行匹配的缺点，由于图像的特征点比较像素点要少很多，大大减少了匹配过程的计算量;同时，特征点的匹配度量值对位置的变化比较敏感，可以大大提高匹配的精确程度;而且，特征点的提取过程可以减少噪声的影响，对灰度变化，图像形变以及遮挡等都有较好的适应能力。所以基于图像特征的匹配在实际中的应用越来越广泛。所使用的特征基元有点特征(明显点，角点，边缘点等)，边缘线段等。

## 2.2 卷积神经网络

卷积神经网络是一类包含卷积计算且具有深度结构的前馈神经网络，是深度学习的代表算法之一 。卷积神经网络具有表征学习能力，能够按其阶层结构对输入信息进行平移不变分类，因此也被称为“平移不变人工神经网络”。对卷积神经网络的研究始于二十世纪80至90年代，时间延迟网络和LeNet-5是最早出现的卷积神经网络；在二十一世纪后，随着深度学习理论的提出和数值计算设备的改进，卷积神经网络得到了快速发展，并被应用于计算机视觉、自然语言处理等领域。卷积神经网络仿造生物的视知觉机制构建，可以进行监督学习和非监督学习，其隐含层内的卷积核参数共享和层间连接的稀疏性使得卷积神经网络能够以较小的计算量对格点化特征。

在图像处理的过程中，常常会用到卷积神经网络进行图像处理。

2.3 虚拟和增强现实

虚拟现实技术是在虚空间中的一个重要技术，是仿真技术与计算机图形学、人机接口技术、多媒体技术、传感技术、网络技术、等多种技术的集合。是一门富有挑战性的交叉技术前沿学科和研究领域。虚拟现实技术(VR)主要包括模拟环境、感知、自然技能和传感设备等方面。模拟环境是由计算机生成的、实时动态的三维立体逼真图像。感知是指理想的VR应该具有一切人所具有的感知。除计算机图形技术所生成的视觉感知外，还有听觉、触觉、力觉、运动等感知，甚至还包括嗅觉和味觉等，也称为多感知。自然技能是指人的头部转动，眼睛、手势、或其他人体行为动作，由计算机来处理与参与者的动作相适应的数据，并对用户的输入作出实时响应，并分别反馈到用户的五官。传感设备是指三维交互设备。

# **研究及发展现状**

对于虚空间的研究与发展，现如今主要集中在将虚拟数据与现实相结合，如在自动驾驶，物体识别，重识别等一系列技术之中。

在自动驾驶领域，YOLOv4算法成为一种主流：

输入:整个算法的输入。目标检测算法的输入可以是整幅图像、数据集或视频流。

Backbone:后端网络的主要功能是提取图像特征。只要图像中的浅层部分是相似的并且在公共区域,比如图像的耳朵、颜色、纹理等部分，背景网络就可以很好地利用一些内置的训练网络，比如VGG16ResNet- 50，ResNeXt-101、Darknet53或者-些轻量级骨干MobilenetV1、2、3、ShuffleNetl、 2等。通过CSDarknet53的Y0L0v4算法、触发方法和Dropblock得到的背景网络。

Neck:辅助层主要用于特征增强。之前的主干网络会将提取的一些相关浅层特征输入到Neck 层。Neck层的常见模型包括SPP、SAM、 FPN、PAN等模型。

Prediction:检测层，这部分的作用是检测提取出来的特征，输出你想要的结果。它是算法的核心部分。

但是时至今日，对于运动物体的实时监测还是存在着很大的问题，如光线，气温的影响，都会对识别提出巨大挑战，极大程度上影响识别结果。

**结 论**

虽然元宇宙一词大量出现在人们的视野中，但是其发展程度还处在一种初级阶段，所能解决的问题仍然很有限，要想真正达到小说中的情节，或者达到人们所期望的程度，还是有很大发展空间的。

**参考文献**

[1] L. Ma, F. Xu, T. Li and H. Zhang, "A Moving Object Detection Method Based on 3D Convolution Neural Network," 2020 7th International Conference on Information Science and Control Engineering (ICISCE), 2020, pp. 55-59, doi: 10.1109/ICISCE50968.2020.00022.

[2] K. R and N. S, "Pothole and Object Detection for an Autonomous Vehicle Using YOLO," 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), 2021, pp. 1585-1589, doi: 10.1109/ICICCS51141.2021.9432186.

[3] H. Wang, S. Zhang, L. Yu and C. Shi, "YOLO Target Detection Algorithm with Deformable Convolution Kernel," 2021 IEEE International Conference on Mechatronics and Automation (ICMA), 2021, pp. 768-772, doi: 10.1109/ICMA52036.2021.9512626.

[4]Paris S . A gentle introduction to bilateral filtering and its applications[C]// Acm Siggraph. ACM, 2007.