Chapter 16
Introduction of 3D Stereo Vision
(Theme 4)

Lin ZHANG
School of Computer Science and Technology
Tongji University

Problem definition of 3D stereo vision

- 3D stereo vision is an important branch of computer vision technology. This technology involves analyzing images or video sequences to achieve perception and understanding of the three-dimensional world
- Depending on the specific application, 3D stereo vision technology generally addresses the following three types of problems:
  - Depth Estimation: Inferring the distance from each pixel in an image to the camera, i.e., depth information, to achieve a stereoscopic perception of the scene
  - 3D Reconstruction: Recovering the 3D structure of objects from 2D images or videos, including their shape, size, and position
  - Motion Estimation: Analyzing the motion information of objects in a sequence of images, including their speed, direction, and acceleration

Problem definition of 3D stereo vision

- For a 3D stereo vision system, the typical input and output data types are described as follows:
  - Input: Usually a set of 2D images or video sequences, which may also include intrinsic and extrinsic camera parameters
  - Output: Includes 3D point clouds, 3D models, depth maps, motion trajectories, and other information used to describe the 3D structure and motion of a scene

Problem definition of 3D stereo vision

- When building a 3D stereo vision system, the following types of sensors are commonly used:
  - Cameras: Used to capture 2D images or video sequences, serving as the primary input source for 3D stereo vision technology
  - Depth sensors: Such as structured light sensors and time-of-flight sensors, used to obtain the distance information from each point in the scene to the sensor
  - Inertial Measurement Units (IMUs): Used to measure the acceleration and angular velocity of objects, assisting in motion estimation
  - LiDAR: Uses laser beams to measure the time of flight and obtain the 3D position information of objects in the scene, commonly used for 3D reconstruction in outdoor environments
  - Camera arrays: Composed of multiple cameras, used to capture multi-view image information for more accurate 3D reconstruction and stereo matching

Problem definition of 3D stereo vision

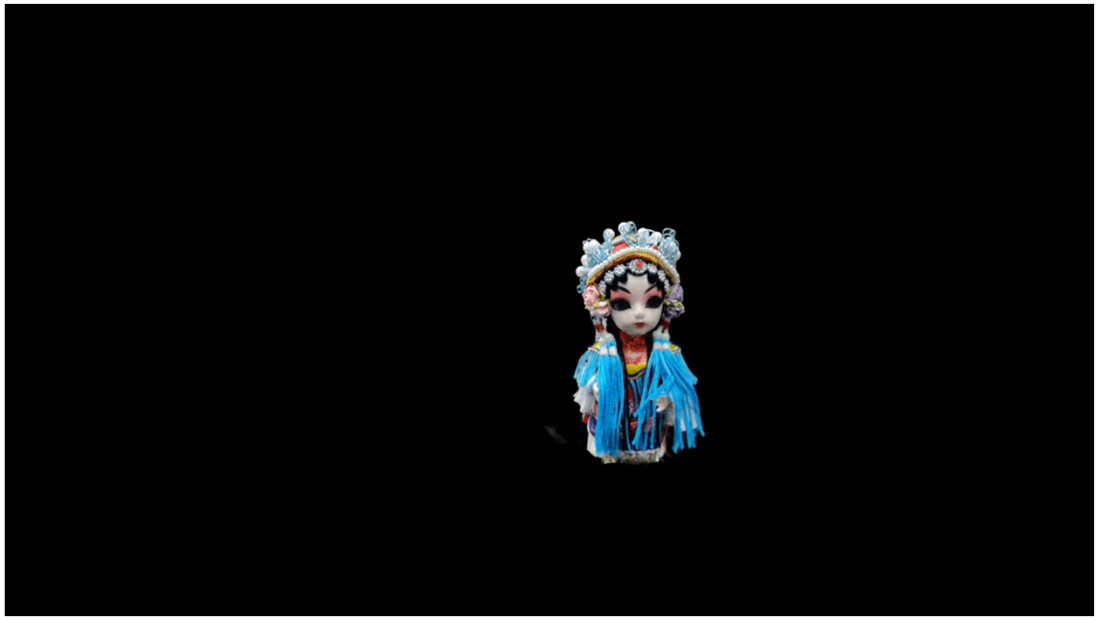
- Today, 3D stereo vision technology is widely applied in many fields, including but not limited to the following:
  - Industrial manufacturing: Through 3D reconstruction and inspection techniques, it is possible to detect the 3D shape and surface defects of products, improving product quality and production efficiency
  - Medical imaging: In medical imaging, the technology is used for the reconstruction, analysis, and diagnosis of medical images
  - Virtual reality: By tracking and reconstructing the user's viewpoint and movements, it enables free movement and interaction within the virtual environment
  - Autonomous driving: By performing 3D reconstruction and recognition of the vehicle's surroundings, it enables autonomous navigation and obstacle avoidance
  - Robotic vision: By reconstructing the environment and recognizing targets, it helps robots better understand and respond to complex environments

Contents of this theme

- ✓ Among the many 3D stereo vision techniques, binocular stereo vision is the most fundamental and important; We will detail the construction of binocular vision systems in Chapter 17
- ✓ Neural Radiance Fields (NeRF)-based 3D reconstruction technology uses neural networks to simulate radiance fields for 3D scene reconstruction and understanding. It involves two key steps: radiance field modeling and scene rendering. We will introduce NeRF-based scene rendering techniques in Chapter 18

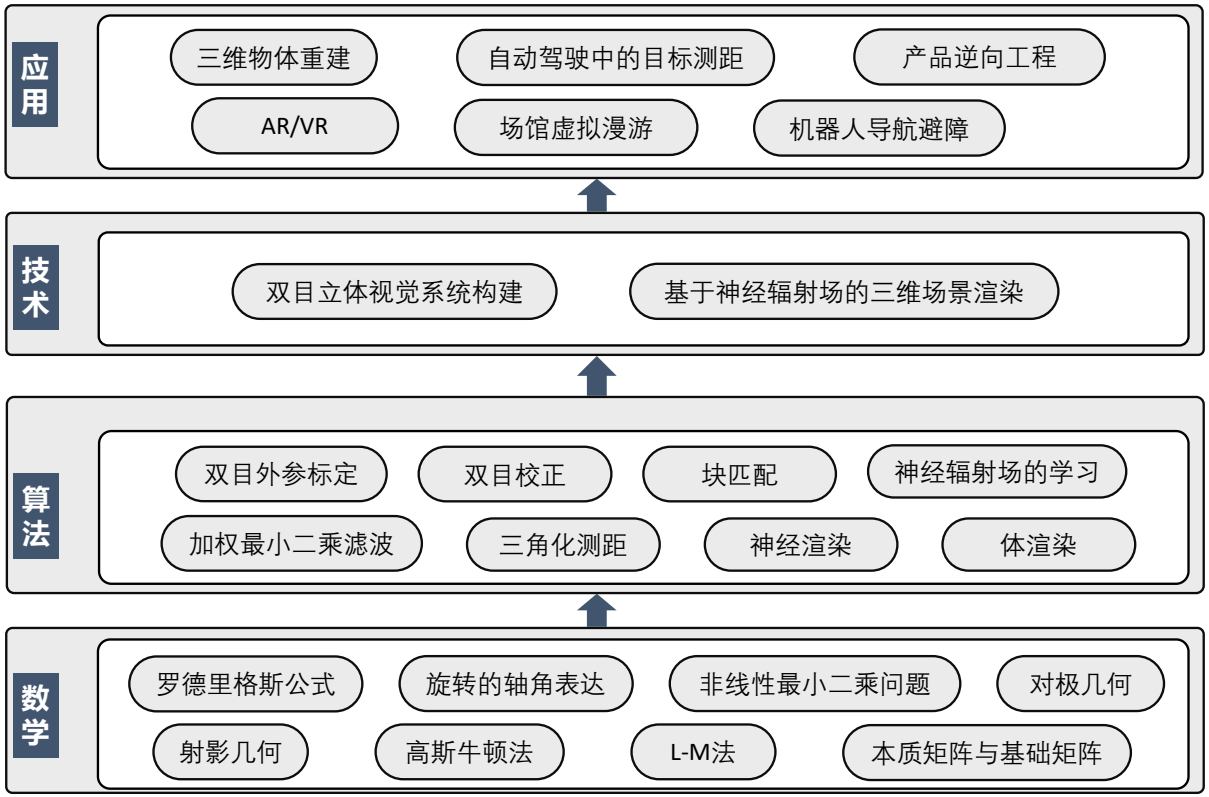
Contents of this theme

2.5D reconstruction using the binocular technique



Neural field rendering result using Instant-NGP

Contents of this theme




Q+A