CHAPTER 11

DATA ACQUISITION USING DUAL RGB-D CAMERAS

Fazliaty Edora Fadzli, Ajune Wanis Ismail, and Norhaida Mohd Suaib

11.1 INTRODUCTION

In the realm of computer vision and machine learning, the acquisition of high-quality data is paramount, especially when aiming for precise three-dimensional (3D) reconstruction. With the rapid advancements in technology, the integration of red, green, blue-depth (RGB-D) cameras have revolutionised the field, offering depth perception alongside traditional colour imaging. However, despite the strides made, challenges persist in capturing comprehensive data, particularly in scenarios requiring detailed spatial understanding or full-body tracking for 3D reconstruction. This chapter delves into the intricacies of data acquisition using a life-size setup of dual RGB-D cameras specifically tailored for 3D reconstruction. By employing dual cameras, researchers and practitioners can capture richer data, enabling more accurate and robust 3D models.

Moreover, utilising life-size setups facilitates the capture of human-scale interactions and environments, opening avenues for applications in various domains, including robotics, augmented reality, and healthcare. Throughout this chapter, we explore the technical considerations, methodologies, and best practises for leveraging dual RGB-D cameras in life-size setups for 3D reconstruction. From calibration techniques to synchronisation methods, each aspect is meticulously examined to ensure optimal data quality and consistency for reconstructing highly detailed and accurate 3D models. Utilise dual RGB-D cameras in life-size setups for data acquisition aimed at 3D reconstruction, empowering them to tackle complex challenges and advance the frontiers of computer vision and beyond.

11.2 UNDERSTANDING THE BASICS

The data acquisition process consists of several crucial stages, meticulously designed to ensure the capture of comprehensive and accurate data. As depicted in Figure 11.1, the initial step involves capturing and processing data inputs, including colour and depth images, from each of the RGB-D cameras deployed at the client site. This phase demands careful attention to detail, as it sets the foundation for subsequent analyses and reconstructions.

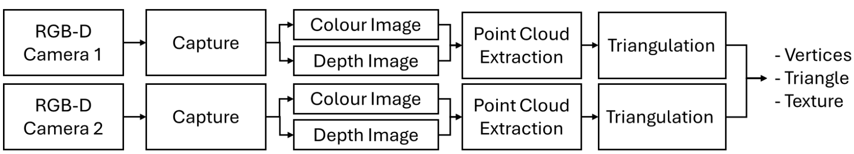

Figure 11.1 Data acquisition process

Once captured, the data streams are transmitted to the server infrastructure for further processing and analysis. This transfer is pivotal, as the raw sensory inputs are prepared for synchronisation and alignment. The integrity of the data transmission is crucial to maintain coherence across the subsequent reconstruction processes.

Upon receipt of the data at the server, synchronisation procedures are executed to minimise latency and ensure temporal coherence across the fused data streams. This synchronisation step is fundamental in achieving precise alignment between the colour and depth information from each camera, facilitating accurate 3D reconstructions.

The subsequent subsections of this chapter delve into the specifics of each stage of the data acquisition process. It details the data capture phase, exploring techniques for optimising image quality and depth accuracy. Following this are the point cloud extraction methods for transforming raw sensor data into 3D representations of the scene. Additionally, the triangulation process is described for reconstructing detailed surface geometries from the captured point clouds. These subsections serve as a comprehensive guide, offering insights into the technical intricacies and best practices essential for robust data acquisition in the context of 3D reconstruction. By understanding and mastering these processes, researchers and practitioners can effectively utilise dual RGB-D cameras in life-size setups, advancing the field of computer vision and spatial analysis.

11.2.1 Data Capture using Multiple RGB-D Cameras

This chapter employs a multi-camera setup utilising multiple RGB-D cameras to capture dynamic scenes containing moving users in real-time. The dynamic scene comprises both dynamic elements, represented by the moving users, and static elements such as tables and chairs. These static elements provide context and stability within the scene and are essential for accurate reconstruction.

The proposed fusion method aims to generate a reconstructed textured mesh of the local user by integrating data extracted from multiple RGB-D streams. These data streams are transmitted to the server and combined into a cohesive unified scene. This fusion method facilitates the extraction of additional features from multiple RGB-D cameras, enhancing the richness and detail of the reconstructed 3D scene.

Life-size capturing is possible with the Kinect V2 device situated at a height of 1.1 metres and a distance of 1.5 to 2.1 metres from the user for full-length visibility with arms extended upwards. This distance can be traversed by users between 1.75 and 2.00 metres tall. Data acquisition using depth cameras gathers the 3D data along with the texture that is used to reconstruct the user's full body.