Research on Face Feature Acquisition Based on Motion Capture

Kejian Zhu, Xiangzhen He*, Yugang Dai, Lulu Cai

Key Laboratory of China's Ethnic Languages and Information Technology of Ministry of Education, Northwest Minzu University, Lanzhou, Gansu 730000, China
hexiangzhen@163.com

Abstract. Face has rich geometrical changes because of the differences in identity and expression. This brings great inconvenience to the collection of human face features. Motion capture technology is a new data acquisition method in recent years. In the emerging fields where computer graphics, image processing, data processing and other disciplines permeate and cross each other, motion capture is a process of recording dynamic motion trajectories in real time. It tracks key points in space and measures key points in real time. The three-dimensional coordinates transform the trajectory into a mathematical representation. The Caralive and Carapost software provided by Vicon provide convenience for the collection of facial features.

1. Introduction

Human face is an important medium for human beings to communicate with the outside world and express their own emotions. Because the eyes, ears, mouth, and nose are all attached to it, the human face is the most expressive part, regardless of whether it is happy, anger, sadness or joy\(^1\). The appearance of the outside world is still a complex and changeable personality. It may also be an individual's temperament, self-cultivation, and connotation, all of which can be seen in one's face. It is because we can get a lot of useful information from the human face. For example, we can infer the basic attributes such as a person's age, race, identity, and expression. Therefore, acquiring facial features becomes crucial. It is precisely because of the large amount of information on the human face, and the faces of different people are not the same. Due to the different identities and expressions, the face has a rich geometric change, this has caused great inconvenience for us to collect facial features\(^2\).

Motion capture technology is an emerging data acquisition method in recent years. It is an emerging field where mathematics, computer graphics, image processing, data processing and other disciplines infiltrate and cross each other. It has attracted many researchers at home and abroad. In pursuit of this, a large number of relevant results have emerged constantly. The research of motion capture technology not only has certain theoretical research value, but also has a wide range of practical application value. So far, motion capture technology has been widely used in modern film and television animation, game production, medical analysis, sports science research and sports training guidance.

In the near 20 years of the development of the motion capture system, many scholars and researchers at home and abroad have carried out research work in related fields and have obtained rich theoretical and practical achievements in many aspects, including the study of motion capture methods, and the motion capture devices. Design, development and system architecture, motion capture data processing algorithms and related software system development, motion capture data management, compression, retrieval and reuse methods, motion data understanding and human motion analysis, human motion reconstruction methods, etc.
The two softwares developed by Vicon, Caralive and Carapost, are well associated with motion capture devices, facilitating the collection of facial features.

2. Motion capture system

Motion capture is a process of recording dynamic motion trajectories in real time. It measures key points in space and measures the three-dimensional coordinates of key points\[^3\]. The motion trajectory is converted into a mathematical representation. Motion capture creates a performance that extracts motion from the surface of an object. In form, it encodes moving objects in a processing and analysis method that meets human needs. Motion capture technology can solve the human body movements, analyze the facial micro-expressions, generate real three-dimensional data, and apply it to animation games to show realistic visual effects. In addition to animation, motion capture has a wide range of applications in biomedical analysis, sports performance analysis, surveillance and monitoring\[^4\].

According to different working principles, the motion capture system is divided into: machinist motion capture system, acoustic motion capture system, electromagnetic motion capture system, and optical motion capture system. This laboratory has the conditions of an optical motion capture system. Face features are captured using an optical motion capture system.

2.1. Optical Motion Capture System

The optical motion capture system is based on the principle of computer vision, and sets up multiple high-definition cameras in three-dimensional space to scan and track the feature points of the captured object from various angles to achieve motion capture. In theory, any point in the motion capture area, as long as the point can be shot by at least two cameras at the same time, we can determine the three-dimensional position of the point. When the camera configuration is high enough, the pixels are clear enough, and the capture speed is fast enough, we can quickly get the trajectory of the point\[^5\].

![Figure 2-1. Principle diagram of active optical motion capture system](image-url)
Current optical motion capture systems fall into three major categories: active motion capture systems, passive motion capture systems, and video-based markerless motion capture systems. The first two systems stick the marker point as the target sensor on the captured object. The Marker point of the active motion capture system is made of light emitting diodes\(^6\). The Marker clicks on the bones and is attached to the joints. The adjacent Marker points are connected with cables and the terminal is placed on the power supply device on the human body surface. The United States PhaseSpace system is active, it uses a bright light-emitting diode as a marker, can be used for motion capture in outdoor environments, light-emitting diodes by controlling the brightness of the pulse signal to control the brightness of the light-emitting diode time domain coding, Robust performance, increased data tracking accuracy and consistency. In the passive motion capture system, the actor wears a specially designed tights that are covered with reflective balls placed in the main joints. High resolution cameras are then strategically positioned to track these reflectors during the movement of the actors. Each camera obtains the 2D coordinates of each reflector by a segmentation step. The proprietary software is then used to analyze the data captured by all cameras to calculate the 3D coordinates of the reflector. The passive motion capture system is represented by Motion Analysis in the United States and Vicon in the United Kingdom. Its main advantage is that the technology is relatively mature, Marker points can be added or deleted as needed, the sampling rate is high, the precision is high, the motion capture is accurate, and there is a wide range of application. In particular, the latest series of Vicon products have developed advanced smart capture capabilities, which have enhanced Marker's automatic recognition capabilities and facilitate error correction. To a certain extent, real-time synchronization between live performances and animations has been achieved, and the work intensity has been greatly reduced, essentially speaking. It greatly improves the practicality of the motion capture system. The video-based unmarked point motion capture system mainly obtains spatial three-dimensional data by extracting two-dimensional coordinates of each node in the two-dimensional image and then performing three-dimensional measurement in combination with multi-camera vision. The system is based on the video image information which is complex, slow and robust. It can not be used in real-time operation and the accuracy is not high.

Although the optical motion capture system is expensive and has high requirements for the capture environment, light and noise must be eliminated. However, the sampling rate of this system is very high, and it can capture fast movements such as martial arts, acrobatics and gymnastics. The sampling rate...
usually depends on the camera used, the higher the resolution, the higher the sampling rate. A sampling rate of up to 200 frames/second is achievable. The accuracy of data acquisition is high, and a large number of marker points can be traced, and the configurations of marker points are easy to change. In addition, the optical motion capture system has a high degree of freedom. Unlike other systems, there is no cable or a limited working space, and the reflector does not have a restrictive or tedious influence on the actor. Moreover, since the reflector does not provide impedance, the number used in the capture process is virtually unlimited, which gives a very high level of theoretical detail.

2.2. Vicon Cara facial motion capture system

A complete Cara hardware system consists of Cara HeadRig, HD camera, Cara Logger, BOB, power supply, MX workstation and a range of accessories. The main components are shown in Figure 2-3.

![Figure 2-3. Main components of ViconCara facial 3D motion capture system](image)

2.2.1. Introduction to Vicon Software

Vicon Cara comes with CaraPost and CaraLive software.

Cara Live is the core software of Cara system. It is mainly used for system installation, debugging and data acquisition. It directly serves users to implement system management during the entire data collection process. Cara Live uses a new type of wireless real-time streaming media as a transmission method to continuously monitor the captured images. At the same time, a color-coded feedback system is added. Combined with the alarm system, feedback information is sent when abnormalities are captured, allowing users to take timely measures. From the standpoint of CaraLive's own advantages: users can adjust the exposure level under different lighting conditions, and can also set the exposure area according to the needs; the user can change the detector's accuracy so that the feature points can be better tracked, and can also be used for single camera views and Multi-camera view to switch freely.

The main function of CaraLive is to build and calibrate the 3D data extracted by CaraPost. It can automatically mark and trace 3D data points and export the processed data to third party software. CaraPost also offers semi-automatic and supervisory tracking to efficiently process 3D facial capture data. Compared to other data processing software, CaraPost can process data faster and more accurately.

ViconCara has several advantages compared to other motion capture systems: High-definition, four high-resolution high-speed cameras work simultaneously in the system, capturing the subtle movements of the face, custom built-in lighting, and ensuring data integrity. Unique calibration and tracking algorithms eliminate most of the interference from redundant data, and clear motion trajectories save the
resources needed for subsequent processing. The performance is good. Each component of the system is from the most basic design to the highest specification design. It captures and tracks facial movement with the best fidelity while allowing the actors to move and express freely. The Cara system design ensures reliability. Sex and actor comfort. The flexibility is strong. Cara's essence is a modular system. The hardware can be used according to the user's needs. The number of cameras can also be selected. The user can use the 3D data output from CaraPost, and can also directly extract images from the recorder and transfer them according to their purpose and output.

3. Summary
In this paper, through the understanding of the motion capture system, the optical motion capture system, the Vicon face motion capture system and the two main softwares of the icon-Caralive and Carapost, we have laid a theoretical foundation for the extraction of facial feature data and ensure the success of subsequent experiments. Go ahead and get an overview of how to extract feature data.

We should learn the Vicon Cara system, in order to build a dynamic capture experimental environment connecting hardware devices and setting camera parameters selecting actors, designing facial expressions and pasting marker points to do the preliminary work of facial capture. Then we should use CaraLive software to collect facial data and obtain four Two-dimensional motion video of road channel. Through CaraPost software, two-dimensional data points are identified and tracked to set up three-dimensional environment. The three-dimensional data points are optimized for processing so that the motion data of facial feature points are finally derived.

Acknowledgements
The research projects of the social science and humanity on Young Fund of the ministry of Education (No.17YJCZH057), This work was supported by Gansu Provincial Higher Education Research Project (No.2015B-007), Northwest MinZu University the central university of basic scientific research business funds (No.31920150228, No.31920170156) This research supported by the Fundamental Research Funds for the Central Universities (NO. Yxm2018128).

References
[1] L. Edward, S. Dakpe, P. Feissel, B. Devauchelle, F. Marin. Quantification of facial movements by motion capture[J]. Computer Methods in Biomechanics and Biomedical Engineering, 2012, 15(sup1).
[2] Haoda Huang, Jinxiang Chai, Xin Tong, Hsiang-Tao Wu. Leveraging motion capture and 3D scanning for high-fidelity facial performance acquisition[J]. ACM Transactions on Graphics (TOG), 2011, 30(4).
[3] DELPHINE SOFTWARE INTERNATIONAL. Facial motion capture imaging unit for video game generation has color video camera viewing subjects face through mirror with white light illumination[P]. FR2838904, 2003-10-24.
[4] DELPHINE SOFTWARE INTERNAT (FR). Facial motion capture imaging unit for video game generation has color video camera viewing subjects face through mirror with white light[P]. FR2838904, 2003-10-24.
[5] WATANABE Naoto, YAMAMOTO Masanobu. Body, hands and face motion capture from full high-definition television camera[J]. ITE Technical Report, 2008, 32, 58(0).
[6] Xiaoting Wang, Lu Wang, Guosheng Wu. Body and Face Animation Based on Motion Capture[J]. International Journal of Information Engineering and Electronic Business (IJIEEB), 2011, 3(2).