Development of Dual Video Acquisition and Parallel AVI Player with Camera Calibration System for Supporting Dual View Motion Detection

Amar Vijai Nasrulloh, Institut Teknologi Bandung, Indonesia and Universitas Lambung Mangkurat, Indonesia
Tati L. R. Mengko, Institut Teknologi Bandung, Indonesia
Tommy Apriantono, Institut Teknologi Bandung, Indonesia

ABSTRACT

Human motion capture consists of the recording of human body movements for immediate or delayed analysis and playback. Development of the dual video acquisition methods and continued with dual and parallel AVI player display can visually display dual view motion detection and successfully identify the depth of the motion detection. Times and frames are played together in dual AVI player display. The mouse cursor tags, using markers, every motion manually and records the two view graphics movement of the joints represented by the markers. The software module to find Ln of two camera calibration that connects a 2 dimensional image plane and 3 dimensional world coordinates has been built using Direct Linear Transformation equation with distance error \( \pm 0.44 \) pixels for video camera 1 and \( \pm 1.03 \) pixels for video camera 2.

Keywords: Camera Calibration, Direct Linear Transformation, Markers, Motion Detection, Parallel AVI Player, Video Acquisition

INTRODUCTION

Video technology has now reached a level of sophistication that allows easy digitization. Digital video can be easily edited, reproduced, incorporated into databases, and posted on intra- and Internet sites for clinical use and demonstration purposes. Numerous methods exist for the production of digital video (Mandar & Grantier, 2001)

Conventional two–dimensional video is a mature technology in both professional environments and home entertainment. A multitude of analog and digital video formats is available
today, tailored to the demands and requirements of different user groups. Efficient coding schemes have been developed for various target bit rates, ranging from less than 100 kilobit per second for video conferencing applications to several megabit per second for broadcast quality TV (Wurmlin, Lamboray, Staadt, & Gross, 2002). Multi-camera acquisition methods can successfully identify the states of the repetitive motion and capture 3D objects by using a static installation of cameras surrounding the motion volume (Matusik, Buehler, Raskar, Gortler, & Mcmillan, 2000).

Human motion capture consists of the recording of human body movements for immediate or delayed analysis and playback. The information captured can be as simple as the body position in space or as complex as the deformations of the face and muscle masses. Optical tracker systems typically use small markers attached on the body of the subject and a set of two or more cameras focused on the subject to capture its motions (Kolahi, Hoviattalab, Rezaeian, Alizadeh, Bostan, & Mokhtarzadeh, 2007).

Calibration of cameras is considered as an important issue in computer vision. Accurate calibration of cameras is especially crucial for applications that involve quantitative measurements such as dimensional measurements, depth from stereoscopy, or motion from images (Weng, Cohen, & Herniou, 1992).

Camera calibration in the context of three dimensional (3D) machine vision in the process of determining the internal camera geometric and optical characteristic and/or the 3D position and orientation of the camera frame relative to a certain world coordinate system (Tsai, 1987). The most commonly used camera calibration method is perhaps the DLT (direct linear transformation) method originally reported by Abdel-Aziz and Karara. The DLT method uses a set of control points whose object space/plane coordinates are already known. The control points are normally fixed to a rigid frame, known as the calibration frame. The problem is essentially to calculate the mapping between the 2D image space coordinates and the 3D object space coordinates (Pourcelot, Audigie, Degueurce, Geiger, & Denoi, 2000).

The Direct Linear Transformation (Abdel-Aziz & Karara, 1971) equations are:

\[
\begin{align*}
u &= \frac{L_1 x + L_2 y + L_3 z + L_4}{L_9 x + L_{10} y + L_{11} z + 1} \\
v &= \frac{L_6 x + L_7 y + L_8 z + L_9}{L_9 x + L_{10} y + L_{11} z + 1}
\end{align*}
\]

in which \((u,v)\) are the digitized co-ordinates of a point, \((X,Y,Z)\) are the object’s space co-ordinates, and \((L_1, L_2, L_3)\) the unknown DLT parameters of each camera. DLT equation could be extended for matrices of camera calibration and 3D reconstruction equation.

Camera calibration is done to find mapping variables between 2D coordinate \((u,v)\) at the image plane, which is the pixel frame, and the 3D coordinate in the real world \((X,Y,Z)\). In general, at least 6 predefined reference points are needed (Kolahi, Hoviattalab, Rezaeian, Alizadeh, Bostan, & Mokhtarzadeh, 2007).

The equation for the accuracy of camera calibration is as follows:

\[
\begin{align*}
\varepsilon_{ui} &= u_i - \frac{L_1 x_i + L_2 y_i + L_3 z_i + L_4}{L_9 x_i + L_{10} y_i + L_{11} z_i + 1} \\
\varepsilon_{vi} &= v_i - \frac{L_6 x_i + L_7 y_i + L_8 z_i + L_9}{L_9 x_i + L_{10} y_i + L_{11} z_i + 1} \\
\varepsilon_c &= \frac{1}{n} \sum_{i=1}^{n} \varepsilon_{ui}^2 + \varepsilon_{vi}^2
\end{align*}
\]

Based on the matrix in equation (1), there are 11 unknown variables, \(L_1\) to \(L_{11}\), so 11 equations are needed to solve it. Equation (3) is derived from equation (1) for camera calibration.

\[
\begin{align*}
A \cdot L &= b \\
(A^\dagger \cdot A) \cdot L &= A^\dagger \cdot b \\
\left(A^\dagger \cdot A\right)^{-1} \cdot (A^\dagger \cdot A) \cdot L &= \left(A^\dagger \cdot A\right)^{-1} \cdot A^\dagger \cdot b \\
L &= \left(A^\dagger \cdot A\right)^{-1} \cdot A^\dagger \cdot b
\end{align*}
\]
Reframing Dichotomies: Human Experiential Design of Healthcare Technologies
www.igi-global.com/chapter/reframing-dichotomies-human-experiential-design/50781?camid=4v1a