BIOSYSTEMS ENGINEERING 151 (2016) 409–424

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/issn/15375110

Research Paper

Extrinsic calibration of a multi-Kinect camera scanning passage for measuring functional traits in dairy cows

*Jennifer Salau**, *Jan H. Haas*, *Wolfgang Junge*, *Georg Thaller*

Institute of Animal Breeding and Husbandry, Christian-Albrechts-University Kiel, Olshausenstraße 40, 24098 Kiel, Germany

ARTICLE INFO

Article history:
Received 5 August 2016
Received in revised form 30 September 2016
Accepted 12 October 2016

Keywords:
3D camera
Calibration
Image processing
Monitoring dairy cattle

Camera based systems in dairy cattle have been intensively studied over the last years. Past studies have concentrated on single camera systems with a limited range of applications. Here the development of a camera system comprising multiple 3D cameras (six Microsoft Kinect cameras) for monitoring purposes in dairy cows is presented. A recording unit was constructed, and software for recording, synchronising, sorting and segmenting images, and transforming the 3D data in a joint coordinate system was implemented. The latter is called extrinsic calibration and was dealt with in this study using a 3D calibration object. The internal estimations of acceleration data and coefficients for the parametrised plane describing the floor of the scenery obtained by the Kinect camera were tested and discussed with regard to their usage in the calibration process. The presented approach accurately determined whether a Kinect device was a side view camera or one of two types of top view cameras (accuracy 92.9%) and specified its mounting position in the system (accuracy 95.8%). Furthermore, rotations and translations between the cameras and a reference coordinate system were estimated from 3D point clouds of the calibration object. For cameras in upright position (side view cameras) the rotational accuracies were satisfactory, showing 4.4° ($\pm 2.9^\circ$) deviation from the reference. For the cameras in top view position larger errors ($25.5^\circ \pm 6.4^\circ$) were measured due to them being mounted in an inclined way. Nevertheless, a basis for registration methods was achieved.

1. Introduction

Several successful camera based studies have been published in the last years with respect to lameness detection (Pluk et al., 2012; Song et al., 2008; Viazzi et al., 2013) or body condition scoring (Azzaro et al., 2011; Bercovich et al., 2012; Halachmi, Klopcic, Polak, Roberts, & Bewley, 2013) using digital 2D or thermal cameras. To overcome problems with image segmentation (Van Hertem et al., 2013) and light, different 3D cameras like Time-of-Flight (TOF) cameras (Krukowski, 2009; Salau et al., 2014b; Weber et al., 2014), the Microsoft Kinect camera (Viazzi et al., 2014), or ASUS Xtion Pro (Kuzuhara et al., 2015) were introduced in the field. The above mentioned studies using a TOF camera recorded only animals in

* Corresponding author.
http://dx.doi.org/10.1016/j.biosystemseng.2016.10.008
1537-5110/

Abbreviation	Declaration
2D/3D	Two/three dimensional
TOF	Time-of-Flight, principle for the measurement of 3D data
RGB	Designation for the additive colour model based on red, green, and blue
LED	Light emitting diode
FOV	Field of view of a camera
LOS	Line of sight of a camera
X, Y, Z	Designation of the axes of the coordinate system in which the Kinect internally measures 3D data
A_K, B_K, C_K, D_K	Coefficients for the parametrisation of the floor plane estimated by the Kinect camera. A point $u=(u_1,u_2,u_3)\in\mathbb{R}^3$ lies on the plane $\Leftrightarrow A_K\times u_1+B_K\times u_2+C_K\times u_3+D_K=0$
SDK	Software Development Kit
$\mathbb{R}$	Real numbers
$\mathbb{R}^3$	Three dimensional real vector space
kdm	File name extension for "Kinect depth map" streams holding 3D data of a Kinect camera
acc	File name extension for text files holding data of the Kinect accelerometer and A_K, B_K, C_K, D_K
euclD	Designation for the Euclidean distance between two $\mathbb{R}^3$-vectors
angle	Designation for the angle between two $\mathbb{R}^3$-vectors given in degree
minD/maxD	Designation for the minimal/maximal difference in components between two $\mathbb{R}^3$-vectors
"S", "N", "U"	Designations of the camera types (side view, normally mounted top view, upside down top view)
"S0", "N0", "U0", "S1", "N1", "U1"	Designations of the camera positions in the system, 0 or 1 indicates the side of the system
x, y, z	Designation of the axes of the joint coordinate system defined for the six Kinect cameras in the system
PCL	Point Cloud Library, open C++ library for 2D and 3D image processing
RANSAC	Random Sample Consensus, iterative method for parameter estimation
A_R, B_R, C_R, D_R	Coefficients for the parametrisation of the floor plane approximated using RANSAC on the 3D data. A point $u=(u_1,u_2,u_3)\in\mathbb{R}^3$ lies on the plane $\Leftrightarrow A_R\times u_1+B_R\times u_2+C_R\times u_3+D_R=0$
$d_R(u)$	Minimal distance between the point $u\in\mathbb{R}^3$ and the plane parametrised by the coefficients A_R, B_R, C_R, D_R
$d_K(u)$	Minimal distance between the point $u\in\mathbb{R}^3$ and the plane parametrised by the coefficients A_K, B_K, C_K, D_K
r, l, u, d	Right, left, up-, and downwards pointing axis of the calibration object as shown in the depth map
accData2UNS	The step in the algorithm where the type of the camera is determined from the accelerometer data
ZeroThreshold	A threshold for perpendicularity used in the estimation of rotational matrices, values below ZeroThreshold were considered zero
incl	Inclination angle (degree) of the camera

standstill, because the TOF cameras were prone to motion artefacts, as the distances between camera and objects were calculated from a series of signals (Hansard, Lee, Choi, & Horaud, 2012). The Kinect camera ("PrimeSense Supplies 3-D-Sensing Technology to 'Project Natal' for Xbox 360, 2010; Kinect for Windows, 2014") calculates depth values using deformations in a predefined infrared pattern which is projected onto the scenery ("Structured Light"; Fofi, Sliwa, & Voisin, 2004). Thus, the Kinect camera is less susceptible to motion artefacts. Furthermore, the significant differences in the measurements between black and white fur (Salau et al., 2015b) observed with the above mentioned TOF camera could not be noticed when working with the Kinect camera. Another option to measure 3D data is stereo vision, which has been used in dairy science. Shelley (2016) implemented a two camera stereo vision system with the purpose of determining body condition in dairy cows.

With a setup which was designed for one specific aspect only, the above mentioned studies focused on monitoring of either lameness or body condition. As necessity for a system that provides holistic monitoring possibilities was given, the concept of a 3D cow scanner based on multiple Microsoft Kinect cameras was investigated (Salau, Haas, Junge, & Thaller, 2014a). In a passage for the cows to walk through, the fields of view of the cameras were combined to deliver 3D information on a high percentage of the surface of the animals, which enables various applications. The objective evaluation of the animal body measurements is one specific goal of the development of the cow scanner. Precise conformation recording is the basis for breeding values in dairy cattle and necessary for the selection of the animals best suited for breeding. According to the "International Committee for Animal Recording – Conformation recording dairy and beef cattle (2015)" the linear description is meant to score traits individually and measure degree rather than desirability. As the traits are lengths and angles along the cow's body, measuring them is well suited to be carried out with methods of computer vision, and the influence of the classifying person would be removed. Except from analysing singled out body parts (i.e. Zwertvaegher, Baert, Vangeyte, Genbrugge, & Weyenberg, 2011), to the authors best knowledge no approach to automate conformation recording has been made so far. Important feasibility steps have been handled and the applicability for the desired monitoring purposes has been shown reaching promising precision. In Salau, Haas, Junge, Leisen, and Thaller (2015a) claws were determined automatically with 1.2% error rate in recordings of cows passing through the system. Later the authors measured lengths and angles from 3D recordings captured with the system after manual specification of the points of interest (Thaller, Salau, Haas, & Junge, 2015). Traits computed from standing and walking animals were compared. For example teat lengths could be measured with standard errors ranging from 0.7 mm to 1.6 mm and from 1.8 mm to 3.2 mm in standstill and movement, respectively. Standard errors in the heights of ischeal tuberosities varied between 2.5 mm and