Medical Engineering and Physics 37 (2015) 813–819

Contents lists available at ScienceDirect

Medical Engineering and Physics

journal homepage: www.elsevier.com/locate/medengphy

Technical note

Anatomical calibration for wearable motion capture systems: Video calibrated anatomical system technique

Maria Cristina Bisi*, Rita Stagni, Alessio Caroselli, Angelo Cappello

Department of Electrical, Electronic and Information Engineering – Guglielmo Marconi, University of Bologna, Viale Risorgimento 2, 40136 Bologna, Italy

ARTICLE INFO

Article history:
Received 23 January 2015
Revised 5 May 2015
Accepted 18 May 2015

Keywords:
Human movement analysis
Joint kinematics
Anatomical calibration
Wearable sensors

ABSTRACT

Inertial sensors are becoming widely used for the assessment of human movement in both clinical and research applications, thanks to their usability out of the laboratory. This work aims to propose a method for calibrating anatomical landmark position in the wearable sensor reference frame with an ease to use, portable and low cost device.

An off-the-shelf camera, a stick and a pattern, attached to the inertial sensor, compose the device. The proposed technique is referred to as video Calibrated Anatomical System Technique (vCAST). The absolute orientation of a synthetic femur was tracked both using the vCAST together with an inertial sensor and using stereo-photogrammetry as reference.

Anatomical landmark calibration showed mean absolute error of 0.6±0.5 mm: these errors are smaller than those affecting the *in-vivo* identification of anatomical landmarks. The roll, pitch and yaw anatomical frame orientations showed root mean square errors close to the accuracy limit of the wearable sensor used (1°), highlighting the reliability of the proposed technique.

In conclusion, the present paper proposes and preliminarily verifies the performance of a method (vCAST) for calibrating anatomical landmark position in the wearable sensor reference frame: the technique is low time consuming, highly portable, easy to implement and usable outside laboratory.

1. Background

In the last decades, wearable motion capture systems have become widely accepted for the assessment of human movement in both clinical application and scientific research [1–7]. Compared to traditional movement analysis systems, wearable sensors offer advantages in terms of cost, size, weight, power consumption, ease of use and, most importantly, portability. Motion capture systems based on wearable sensors frequently adopt functional methods [8–12] for the description of joint kinematics, requiring the execution of ad-hoc motor tasks to evaluate directly the mechanical rotation axes. On the other hand, recommendations for reporting kinematic data [13–15] always suggest referring to anatomical bone structure in order to allow inter-subject comparability of acquired data. Joint functional axes, estimated through a functional approach, are effective in describing subject-specific joint mechanical function, thus they can be affected by joint alterations and/or limitations in joint mobility of the specific patient and need to be referred to anatomy (e.g. the position of functional axes with respect to bone structure is indeed often used as an indicator of joint function [16]). Therefore, for the effective exploitation of wearable motion capture systems in clinical and research applications, a description of joint kinematics based on anatomical calibration techniques is needed.

Stereo-photogrammetry based motion analysis quantifies joint kinematics by reconstructing standardized bone-embedded anatomical frames [14,15,17,18] with respect to a global reference frame. The anatomical frame is usually estimated and calibrated with respect to a technical frame identified by the reflective markers positioned on the analyzed body segment: the procedure is referred to as anatomical calibration [17,19,20]. Anatomical calibration is a relatively time consuming procedure, but it is necessary for reporting kinematic data in a repeatable, standardized way, useful in research and clinical application [17].

Anatomical calibration procedures for inertial measurement units were proposed in the literature [21–23], but they showed drawbacks that limited their use to become widespread. Favre et al. [21], suggested a new technique using a high-resolution magnetic system to perform anatomical calibrations and a wearable system for the kinematics measurement. Despite the interesting potential that this mixed approach has, its usability outside laboratory is limited and expensive. Tadano et al. [22] proposed an anatomical calibration

* Corresponding author. Tel.: +39 328 4241867.
E-mail addresses: mariacristina.bisi@unibo.it (M.C. Bisi), rita.stagni@unibo.it (R. Stagni), alessio.caroselli@unibo.it (A. Caroselli), angelo.cappello@unibo.it (A. Cappello).

http://dx.doi.org/10.1016/j.medengphy.2015.05.013
1350-4533/

procedure using photos of the subject and acceleration data: inaccuracies introduced by photos and the required oversimplified assumptions resulted in a technique not suitable for an accurate joint kinematics quantification. Picerno et al. [23] suggested an anatomical calibration technique for wearable sensors using a calibration device equipped with one magnetic and inertial sensor allowing to directly measure the anatomical axes. This technique has a high portability but it becomes clumsy and expensive when extended to the whole body anatomical landmarks requiring calibration devices of different geometry, each one with a sensor mounted on.

The aim of the present work is to propose a novel anatomical calibration technique for wearable motion capture systems facilitating the achievement of a widespread use by relatively low time consumption, high portability, easy implementation, usability outside laboratory. The proposed technique exploits an off-the-shelf camera as measurement device and is referred to as video Calibrated Anatomical System Technique (vCAST) [17].

2. Methods

2.1. The vCAST device

The vCAST calibration device consists of a rigid stick with a fine tip at an end and an off-the-shelf camera at the other end. This simple stick coupled with a small planar chess pattern, applied on the wearable sensor cage, completes the device (Fig. 1a and b).

This system is aimed to allow reconstructing the stick tip position in the pattern reference frame and, in a second time, in the wearable sensor reference frame. To understand how the device works four reference frames are briefly introduced (Fig. 1b and c):

1. The image plane (I) reference frame:
   - O_I – The origin is set in the top left corner of the image.
   - u – This axis is parallel to the pixel matrix rows, pointing from left to right.
   - v – This axis is parallel to the pixel matrix columns, pointing from top to bottom.
2. The camera (C) reference frame:
   - O_C – The origin is set in the camera optical center.
   - x – This axis is parallel to the I frame u axis with opposite positive direction.
   - y – This axis is parallel to the I frame v axis with opposite positive direction.
   - z – This axis is obtained by the cross product between x and y axis versors.
3. The pattern (P) reference frame can be set arbitrarily. In this work, all the inner corners of the chess pattern lie on the XY plane and have non negative coordinates.
4. The wearable sensor (WS) reference frame axes may slightly deviate from those indicated on the sensor cage or in the data sheet.

According to the pin hole model, 3D key-points ($\mathbf{K}_i$) expressed in the pattern reference frame, for example the ith inner corner of a chessboard, generate 2D key-points ($\mathbf{k}_i$) in the image plane. The following equation is meant to highlight arguments of the function F, representing 3D points non-linear mapping into 2D points:

$$\mathbf{k}_{i,j} = F\left({}^{C}\mathbf{R}_P(j), {}^{C}\mathbf{t}_P(j), \mathbf{K}_i, f_u, f_v, o_u, o_v\right) \qquad \forall i,\ j \\ i = 1, \ldots, N_p \\ j = 1, \ldots, N_f \tag{1}$$

where

- $\mathbf{k}_{i,j}$ are 2D coordinates of the ith chessboard inner corner in the image taken at the jth instant of time.
- N_p is the number of key-points in the pattern.
- N_f is the number of sampled images.
- ${}^{C}\mathbf{R}_P(j)$ is the rotation matrix going from the pattern to the camera frame at the jth instant of time.
- ${}^{C}\mathbf{t}_P(j)$ is the position vector going from the pattern to the camera frame at the jth instant of time.
- $\mathbf{K}_i$ are the 3D coordinates of the ith chessboard inner corner in the pattern frame.
- f_u, f_v are the camera focal lengths and o_u, o_v are the principal point coordinates in the image frame.

The 3D coordinates $\mathbf{K}_i$ are set by the pattern geometry and the 2D coordinates $\mathbf{k}_{i,j}$ are obtained from the jth image using a pattern recognition algorithm [24]: the 3D camera pose with respect to the pattern reference frame can be computed if the camera parameters are known and if the pattern has at least 3 inner corners. More details of the algorithm used to estimate the camera pose are in the Appendix A. Combining the jth camera pose with the stick tip position in the camera reference frame, ${}^{C}\boldsymbol{\rho}$, allows expressing the tip position in the pattern reference frame, ${}^{P}\boldsymbol{\rho}_j$. Rigorously, only one image is needed. However, to improve robustness to light changes and since the stick tip is not moving when pointing, several images (50) are used and estimations of ${}^{P}\boldsymbol{\rho}_j$ averaged as follows:

$${}^{P}\rho = \frac{1}{50}\sum_{j=1}^{50} {}^{C}\mathbf{R}_P^T(j)\left({}^{C}\rho - {}^{C}\mathbf{t}_P(j)\right) \tag{2}$$

In addition, the ${}^{P}\boldsymbol{\rho}_j$ standard deviation along the 50 images is computed to estimate the pointing precision. All steps that lead to the tip estimation in the pattern reference frame are grouped in an algorithm called *Tip Reconstruction* (Fig. 1d.). Finally, if the mismatch between the pattern and the wearable sensor reference frame, ${}^{WS}\mathbf{R}_P$, is known, the stick tip can be expressed in the wearable sensor reference frame.

All this is summarized in two assumptions that must be valid to correctly use the vCAST device:

(a) Camera parameters and stick tip position in the camera frame must be known.
(b) Mismatch between the pattern and the wearable sensor frames must be known.

2.2. Tuning procedure

The *tuning procedure* describes the operations needed to ensure the validity of the two assumptions above: this procedure is supposed to be performed once the device is built and repeated only if the camera is detached from the stick or if a wearable sensor loses its pattern. The *tuning procedure* is divided in two parts: a) *stick and camera tuning* (ensuring the first assumption); b) *pattern tuning* (ensuring the second assumption).

(a) *Stick and camera tuning* is obtained using an algorithm, based on Cedraro et al. [25]. This iterative algorithm estimates internal camera parameters, f_u, f_v, o_u, o_v, and stick tips coordinates, ${}^{C}\boldsymbol{\rho}$, in the camera reference frame, fed with images of a 2D calibration pattern surrounded by control points (CPs). An algorithm including both these estimations, named *Camera and Stick Calibration Algorithm*, is presented in Appendix B. The calibration is performed pointing four of these control points and capturing 300 images for each point (Fig. 2a–b). The known positions of control points in the pattern reference frame, noted as ${}^{P}\bar{C}\mathbf{p}_1, {}^{P}\bar{C}\mathbf{p}_2, {}^{P}\bar{C}\mathbf{p}_3, {}^{P}\bar{C}\mathbf{p}_4$, are used to calibrate the stick tip in the camera reference frame. For each of the four groups of images 2D key-points are extracted using a pattern recognition algorithm [24]. The extracted key-points, noted as ${}^{P}\mathbf{k}_{1,i,j}, {}^{P}\mathbf{k}_{2,i,j}, {}^{P}\mathbf{k}_{3,i,j}, {}^{P}\mathbf{k}_{4,i,j}$, (where the first subscript index indicates which was the control point pointed by the stick tip), allow calibrating the internal camera parameters. Since the algorithm can get stuck in a local minimum, a spot-check is needed to verify the accuracies of the estimated parameters: