ELSEVIER

Journal of Biomechanics 38 (2005) 919–929

JOURNAL OF BIOMECHANICS

www.elsevier.com/locate/jbiomech
www.JBiomech.com

A model-based image-matching technique for three-dimensional reconstruction of human motion from uncalibrated video sequences

Tron Krosshaug*, Roald Bahr

Oslo Sports Trauma Research Center, University of Sport & Physical Education, PO Box 4014, Ullevaal Stadion, 0806, Oslo, Norway

Accepted 27 April 2004

Abstract

In many situations, e.g. sports injuries, three-dimensional kinematics cannot be obtained with traditional lab methods. However, if methods for reconstructing motion patterns from video sequences were available, our understanding of injury mechanisms could be improved. The aim of this study was to assess the accuracy of a new model-based image-matching technique for human motion reconstruction from one or more uncalibrated video sequences, using traditional motion analysis as a gold standard. A laboratory trial was conducted with one test subject performing jogging and side step cutting, while being filmed with three ordinary video cameras. This provided three single camera matchings, three double camera matchings and one triple camera matching for each of the motions. The test subject wore 33 reflective skin markers and was filmed with a seven-camera, 240 Hz motion analysis system. Root mean square (RMS) hip and knee flexion/extension angle differences were less than 12° for all the matchings. Estimates for ad-/abduction (<15°) and internal/external rotation (<16°) were less precise. RMS velocity differences up to 0.62 m/s were found for the single camera matchings, but for the triple camera matching the RMS differences were less than 0.13 m/s for each direction. In conclusion, a new model-based image-matching technique has been developed, that can be used to estimate temporal joint angle histories, velocities and accelerations from uncalibrated video recordings. The kinematic estimates, in particular for center of mass velocity and acceleration, are clearly better when two or more camera views are available. This method can potentially be used to arrive at more precise descriptions of the mechanisms of sports injuries than what has been possible without elaborate methods for three-dimensional reconstruction from uncalibrated video sequences, e.g. for knee injuries.

Keywords: Injury biomechanics; Joint motion; Image processing; Human body model

1. Introduction

Regular laboratory-based motion analysis with skin surface markers is not always feasible. An example is the motion of an actual injury situation, which for obvious ethical reasons cannot be studied in a lab experiment. Therefore, attempts to understand the mechanisms of injuries have mainly been limited to indirect methods, such as mathematical simulations (Gerritsen et al., 1996; McLean et al., 2001) or cadaver studies (Markolf et al., 1995; Bahr et al., 1997; Hame et al., 2002). In a few cases injuries have happened to take place during research experiments (Zernicke et al., 1977; Barone et al., 1999), but this is of course rare. Some research groups systematically collect video recordings of injury situations in an effort to understand the injury mechanisms, as this information is crucial to be able to prevent injuries. However, although video recordings from injury situations often exist, current methods for biomechanical analysis of these are inadequate.

Ettlinger et al. (1995), Boden et al. (2000) have published rough descriptions of the mechanisms of anterior cruciate injuries in skiing and other sports, but without providing any methodological detail. Our group has attempted a more systematic approach, by having an expert panel independently assess joint configurations at the assumed point of injury (Olsen et al., 2003). However, it is inherently difficult to interpret segment

*Corresponding author. Tel.: +47-2326-2000; fax: +47-2326-2307.
E-mail addresses: tronk@nih.no, tron.krosshang@nih.no (T. Krosshaug).

0021-9290/$ - see front matter © 2004 Published by Elsevier Ltd.
doi:10.1016/j.jbiomech.2004.04.033

attitudes and further assess joint angles in three planes simply through visual inspection. Finally, these methods cannot produce continuous estimates of joint angles and positions, which is necessary for biomechanical analyses of the injury mechanisms.

A few methods for markerless three-dimensional reconstruction from video sequences have been described in the biomechanical literature (Halvorsen, 2002; Trewartha et al., 2001). However, due to their use of edge or color tracking, these methods are only applicable under special conditions. Literature from the field of computer vision reveals that to track and reconstruct three-dimensional motion from video sequences with one or more camera views, several approaches are possible. However, these have limitations related to the movement (no camera motion, only one person in camera view at a time, subject facing camera at all times, movement parallel to camera plane, no occlusion, slow continuous movements, moving on flat ground, etc.), the appearance of the environment (constant light, static/uniform background, known camera parameters, etc.), or the appearance of the subject (known starting pose, known subject, markers, special clothing, etc.) (Gavrila, 1999; Moeslund and Granum, 2001). Also, in most cases, the ability to automate tracking and three-dimensional reconstruction is considered more important than optimal accuracy. For these reasons, none of the methods published so far seem to be suited for use in more demanding conditions with uncalibrated video sequences from one or more cameras that may simultaneously be rotating, translating and zooming. In addition, since motion patterns often are complex and the color and contrast properties of the person of interest can blend with the background, new approaches must be sought.

Thus, we wanted to develop a new model-based image-matching technique for reconstruction of human motion from uncalibrated video sequences from one or more camera views. The aim of this laboratory study was to assess the accuracy of this method using traditional motion analysis as a gold standard.

2. Materials and methods

2.1. Laboratory trials

One test subject, a 25 year old team handball player, performed trials of jogging and side step cutting maneuvers. In the side step maneuvers, the subject cut from left to right with weight bearing almost exclusively on the left leg, while the right leg just briefly contacted the ground. We recorded the trials with three ordinary cameras, two S-VHS (Blaupunkt CC695, Hildesheim, Germany) and one miniDV camera (Sony TRV900E, Tokyo, Japan) (Fig. 1). One of the S-VHS cameras filmed from the left rear, panning in the first half of the sequence (total panning angle of 5°). This camera was located 8.2 m from the test subject in mid-stance. The second S-VHS camera recorded directly from the right side during mid-stance, and was located 3.9 m from the test subject, panning during the entire motion (total panning angle of 22°). The DV camera was located in the front, 9.0 m from the test subject. The surface area of the test subject covered 2.5%, 5.1% and 6.5% of the total pixel area in the rear, side and front cameras, respectively.

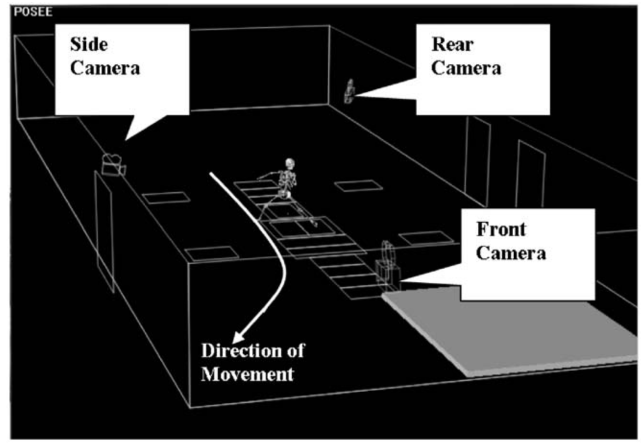

Fig. 1. Illustration of the lab setup as seen from above. The position of the three cameras is shown, as well as the direction of movement for the side step cutting maneuver.

2.2. The model-matching procedure

The model-based image-matching technique was based on the commercially available three-dimensional modeling software Poser® 4 and the Poser® Pro Pack (Curious Labs, Inc., Santa Cruz, CA, USA). This software package features several pre-built models as well as background video import, split camera view (up to four views), camera models containing six translational and rotational degrees of freedom, as well as variable focal length. The surroundings can be modeled using lines, point landmarks or more complex structures. Likewise, custom models (males/females/skeletons with different clothing and/or props) are available. We used a Pentium 4, 1800 MHz PC with three 22″ monitors for the analyses.

Before we could start the matching procedure some image processing was required. The analogue recordings were digitized and stored as uncompressed AVI sequences. We used an Adobe Photoshop (version 4.0, Adobe Systems Inc., San Jose, CA, USA) de-interlacing filter in Adobe AfterEffects (version 5.0, Adobe Systems Inc., San Jose, CA, USA) to obtain an effective frame rate of 50 Hz. Then we corrected for lens distortions by using the Andromeda LensDoc filter (version 1.1,