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.

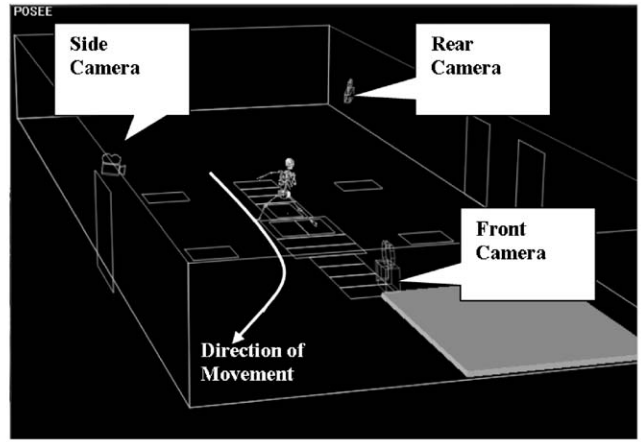

Fig. 1. Illustration of the lab setup as seen from above. The position of the three cameras is shown, as well as the direction of movement for the side step cutting maneuver.

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,

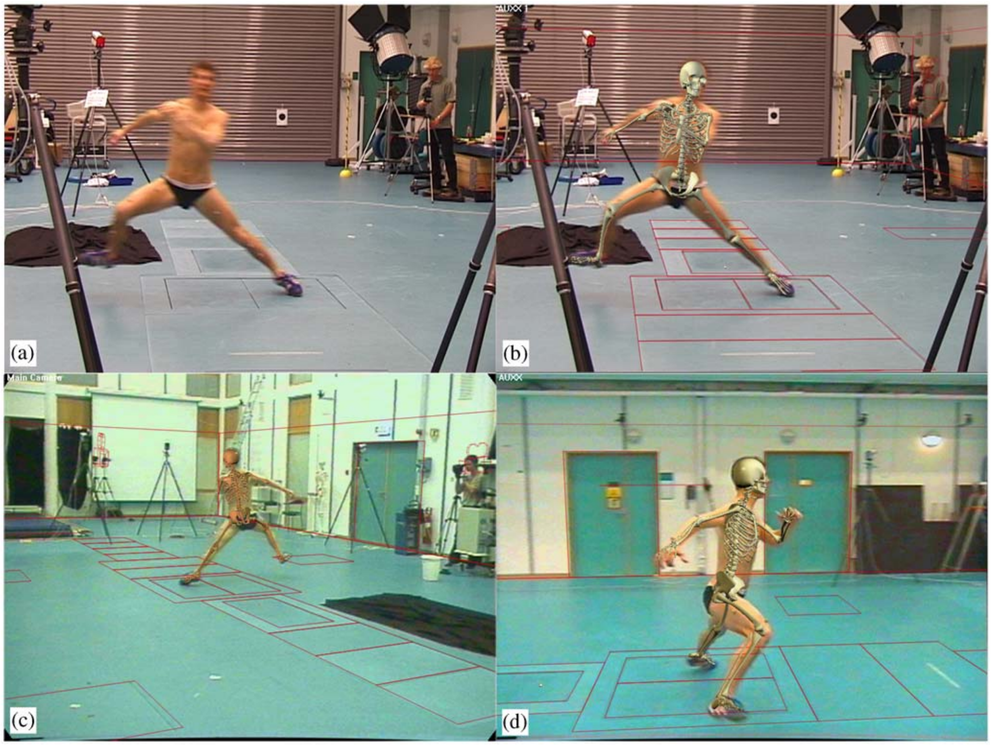

Fig. 2. Triple camera matching at toe-off. The top left panel (a) shows the original video image from the front camera. In the top right panel (b) the customized skeleton model and the laboratory model is super-imposed and matched with the same video image. In the lower left (c) and lower right (d) panels, the model is matched in the rear and side view, respectively. The Poser cameras can be seen as wire-frame models in the top right panel (one camera) and in the bottom left panel (two cameras).

Andromedia Software, Thousand Oaks, California, USA). In the cases where we used more than one camera recording, a manual synchronization was performed, by determining the time for initial contact of the left foot in each camera view.

A skeleton model within Poser (Zygote Media Group Inc., Provo, UT, USA) was used in the first stage of the matching. This model was customized to match the anthropometry of the person in the video. The model, as we used it, had 21 segments (fore foot, rear foot, lower leg, thigh, pelvis, abdomen, chest, neck, head, collar, upper arm, forearm and hand) and 57 degrees of freedom. All joints had three rotational degrees of freedom, except for the sternoclavicular, elbow and wrist joints that had two. In addition, there were three-translational degrees of freedom for the pelvis.

We then manually matched the surroundings and the skeleton model to the background video footage (Fig. 2). The frame-by-frame matching of the surroundings enables us to reconstruct camera motion for video footages shot from translating, rotating and zooming cameras. We started out with key frames, and used the cubic spline interpolation feature in Poser for the intermediate frames. The skeleton matching started with the pelvis segment. We then worked distally by matching the thighs, then the lower legs, and so on. For single camera recordings, it was also beneficial to view the skeleton from other angles (e.g. perpendicular to the original view) to more easily detect obvious errors. When the skeleton matching was completed, we replaced it with the "nude man" model in Poser. This way we could use the surface anatomy to better determine the rotation around the longitudinal axis of each segment. When it was not possible to visually detect the correct rotation about the longitudinal axis of the thigh (e.g. when the front view was not available), we assumed a zero varus/valgus knee angle to determine a unique hip rotation angle—by simply rotating the thigh until the joint center of the ankle was correctly positioned. For the tibia we simply distributed the rotation evenly between the knee and ankle joints. To further enhance the matching we generated videos with both the skeleton and the "nude male" models using different frame rates, to better sense if differences existed between the model and the video sequence(s).

A total of seven matchings were performed for each of the two recorded motions: Three single camera matchings (rear, side and front), three double camera matchings (rear/side, rear/front and side/front) and one triple camera matching (rear/side/front). Two persons performed the matching process. One did all the single camera matchings, while the other did the double/triple camera matchings. Independently of each other, two clinicians gave their opinion on the goodness

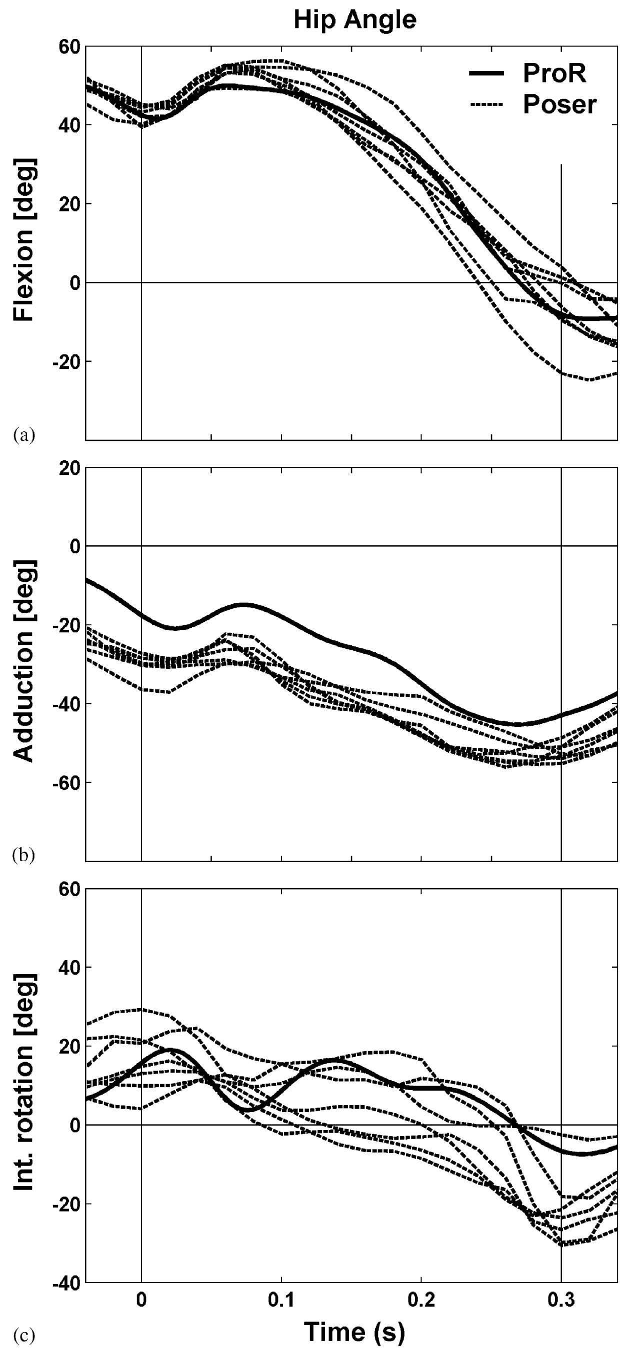

Fig. 3. Hip joint angles (°) of the support (left) leg, calculated with the Pro Reflex system (solid lines) and the model-based image-matching technique for each of the seven camera combinations (dotted lines). The vertical lines indicates initial ground contact of the support leg (time zero) and toe-off.

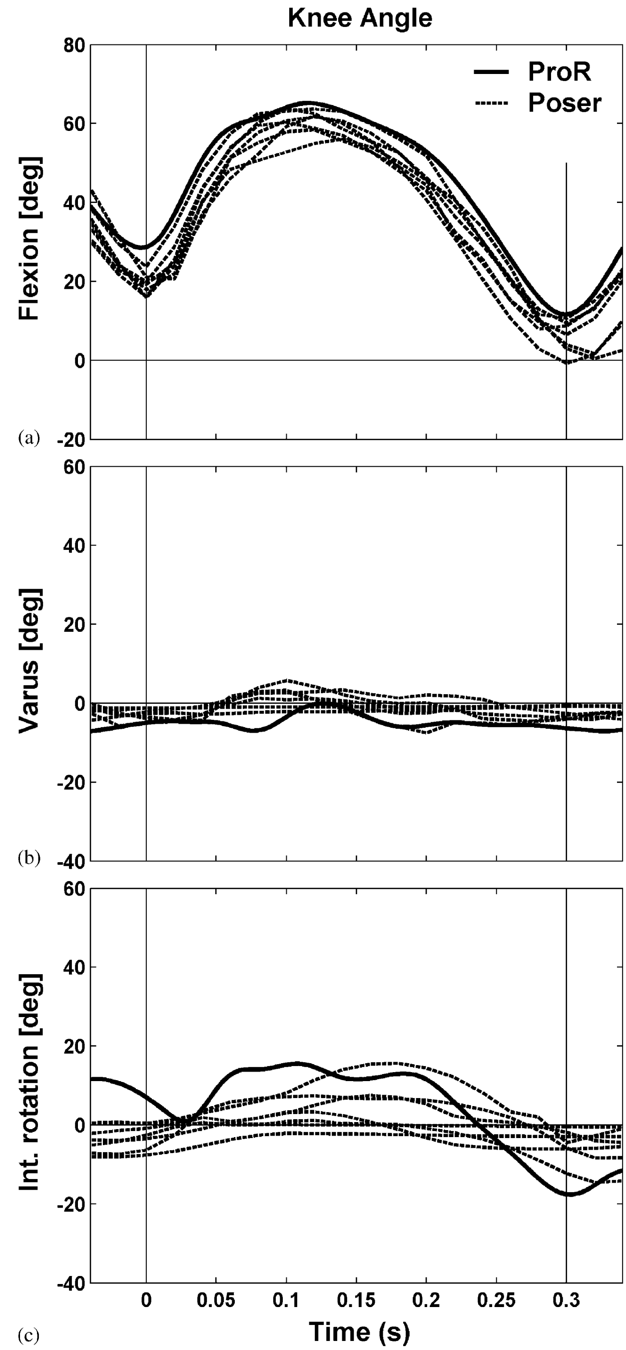

Fig. 4. Knee joint angles (°) of the support (left) leg, calculated with the Pro Reflex system (solid lines) and the model-based image-matching technique for each of the seven camera combinations (dotted lines). The vertical lines indicate initial ground contact of the support leg (time zero) and toe-off.

of the fit twice during the process, to minimize bias resulting from single-operator judgement. The matchings were then adjusted accordingly until finally an accepted motion pattern was found. When a sequence was satisfactory, the translation and joint angle time histories were read into Matlab with a customized script for further processing.

2.3. *The marker-based motion analysis*

In addition to the three-ordinary video camcorders, a seven-camera infra-red, motion analysis system (ProReflex, Qualisys Inc., Gothenburg, Sweden.), and two force platforms (AMTI LG6-4-1, Watertown, MA 02472, USA) simultaneously recorded the motion at 240 and

Table 1
Root mean square and maximal difference in hip and knee joint angles (°) for flexion/extension, adduction/abduction and internal/external rotation of the support leg, between the ProReflex recordings and the estimates from our model-based matching-technique for each of the seven matchings of the plant and cut motion

		Camera views						
		Rear	Side	Front	Rear/side	Rear/front	Side/front	Rear/side/front
Hip	Flexion/extension	10.8 (18.0)	3.7 (7.9)	6.8 (13.2)	4.7 (8.4)	6.0 (10.4)	4.6 (7.8)	2.6 (6.5)
	Ab/adduction	12.1 (16.2)	13.2 (21.9)	11.3 (16.7)	13.0 (19.6)	14.0 (18.9)	13.1 (17.5)	12.6 (16.5)
	Int./ext. rotation	6.5 (16.0)	12.3 (24.4)	7.5 (13.1)	11.8 (22.5)	13.9 (25.1)	15.7 (23.6)	13.9 (23.1)
Knee	Flexion/extension	11.6 (20.9)	5.0 (11.4)	10.4 (17.0)	7.5 (12.4)	9.2 (19.2)	10.0 (16.0)	7.5 (12.8)
	Varus/valgus	3.5 (6.0)	4.4 (6.3)	3.9 (8.0)	5.3 (9.2)	3.9 (9.6)	4.7 (11.1)	4.6 (9.5)
	Int./ext. rotation	13.9 (19.5)	11.1 (16.3)	11.0 (15.9)	10.5 (16.2)	9.1 (14.7)	7.5 (13.9)	9.1 (18.8)

The maximal differences are shown in parentheses.

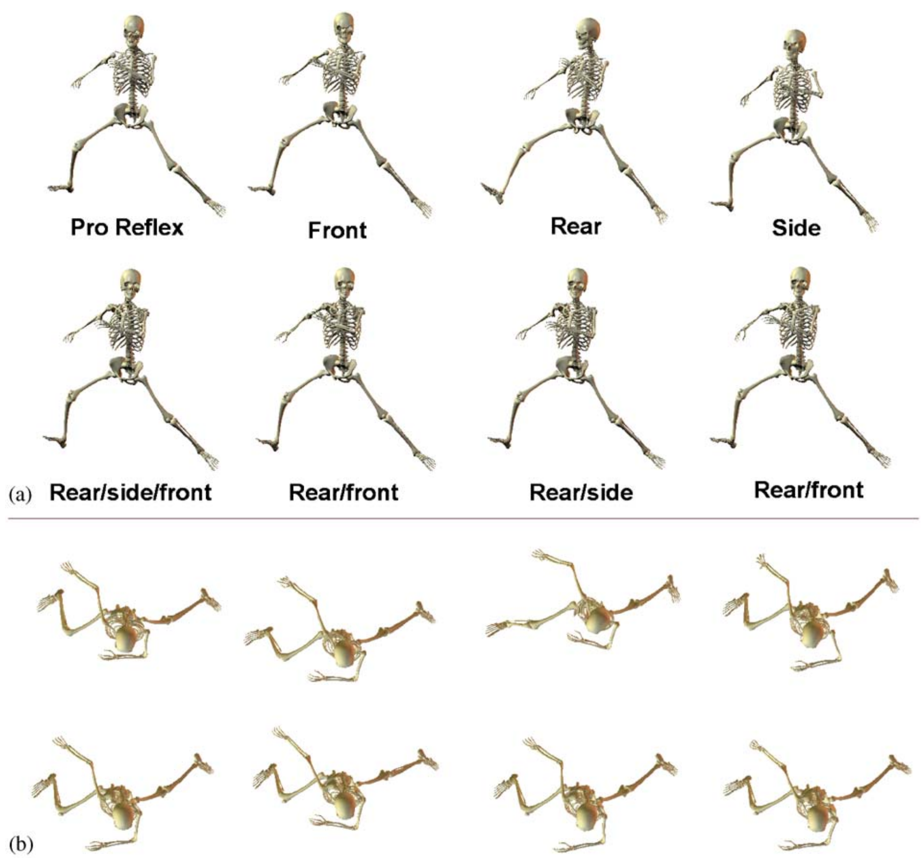

Fig. 5. Illustration of the seven different matchings and the Pro-Reflex recording at toe-off. The Pro-Reflex recorded motion was imported into Poser through a customized Matlab script. In the top panel (a) the skeletons are seen in a frontal view. In the lower panel (b), the skeletons are seen directly from above.

960 Hz, respectively. Thirty-three reflective markers (9 mm radius) were attached to the subject on the following landmarks; base of the fifth metatarsal, heel, lateral malleolus, tibial tuberosity, lateral femoral condyle, major trochanter, the anterior superior iliac spine, mid-posterior superior iliac spine, acromion, lateral epicondyle of the humerus and ulnar styloid process. In addition, we included two extra markers on each shank and three extra markers on each thigh.

This marker configuration allowed for a three-dimensional description of the lower extremities, trunk and upper arm. The head was simply assumed to be rigidly connected to the trunk. The elbow joint was modeled with one degree of freedom, and the hand was assumed to be rigidly connected with the forearm. The method of Soderkvist and Wedin (1993) was utilized to obtain the segment embedded reference frame for the thigh and shank. A static calibration recording of the

Table 2
Root mean square and maximal difference in segment orientation (°) for the pelvis, thigh and shank of the support leg, between the Pro Reflex recordings and the estimates from our model-based image matching-technique for each of the seven matchings of the plant and cut motion

		Camera views						
		Rear	Side	Front	Rear/side	Rear/front	Side/front	Rear/side/front
Pelvis	Tilt	6.0 (11.0)	9.5 (12.2)	6.4 (10.7)	10.3 (14.6)	8.8 (11.4)	9.5 (12.0)	6.8 (12.4)
	Obliquety	3.3 (6.0)	3.2 (6.3)	1.5 (3.0)	3.2 (9.7)	4.3 (8.4)	3.6 (5.8)	3.7 (8.0)
	Rotation	7.9 (13.2)	7.4 (9.8)	5.8 (9.6)	16.3 (21.2)	13.9 (20.3)	13.9 (20.3)	15.2 (21.6)
Thigh	Flexion/extension	7.4 (11.0)	7.9 (13.4)	8.7 (15.4)	8.6 (12.6)	4.0 (6.9)	8.3 (12.7)	6.9 (8.9)
	Ab-/adduction	7.0 (10.5)	7.3 (11.4)	6.4 (9.4)	7.6 (10.3)	5.4 (9.7)	5.1 (8.4)	5.6 (7.5)
	Int./ext. rotation	7.4 (15.5)	13.7 (21.3)	9.3 (15.7)	10.5 (18.8)	10.0 (19.2)	7.0 (12.9)	7.1 (14.0)
Shank	Flexion/extension	7.1 (13.5)	3.3 (7.0)	4.0 (7.3)	2.8 (5.2)	8.8 (12.9)	4.3 (8.0)	4.0 (8.4)
	Ab/adduction	2.3 (5.4)	4.8 (8.0)	2.3 (3.2)	3.3 (5.0)	2.4 (4.7)	1.0 (2.6)	1.4 (2.7)
	Int./ext. rotation	8.9 (18.2)	5.3 (10.3)	3.8 (8.8)	3.0 (6.9)	3.3 (8.3)	4.3 (7.1)	4.0 (6.8)

The maximal differences are shown in parentheses.

athlete in the anatomical position was performed to determine the anatomical axes of the three-dimensional modeled segments in the lower extremity.

To be able to make direct comparisons between the ProReflex measurements and the estimates from the Poser model, we needed to have the same definitions of the anatomical segment axes. Axis transformations were in both cases performed to make the vertical axes of the thigh and shank (**X3**) pass through the joint centers. The antero-posterior axis (**X1**) of the local axis system was defined perpendicular to the **X3** axis with no medio-lateral component. The third axis was the right-hand side cross-product of the vertical and antero-posterior axis ($\mathbf{X2} = \mathbf{X3} \times \mathbf{X1}$).

The joint centers of the ankle and knee were determined by the method of Davis et al. (1991) but the ankle joint center location was defined 1 cm distal to the lateral malleolus, as proposed by Eng and Winter (1995). The hip joint centers were estimated by the method of Bell et al. (1990). The shoulder centers were determined by the method of de Leva (1996). The elbow centers were located midway between the lateral and medial epicondyles.

To estimate the inertia parameters, we used a modified version of Yeadon's stadium solid method (Yeadon, 1990). The modifications concerned the three most distal solids defining each of the feet. These solids were modeled as for the hand, with three stadium solids lying subsequent to each other. The three stadium solids are positioned in such a way that the longitudinal axis of the solids is horizontal (from heel-to-toe), instead of vertical, for a standing person.

For both the Poser model's center of mass (COM) translation and the marker trajectories, smoothing and interpolation were performed by the generalized cross-validation package of Woltring (1986). To circumvent possible oscillation problems in interpolation, the cubic mode with a 12 Hz cut-off frequency was chosen for the marker trajectories. Velocities and accelerations were then estimated using interpolating quintic splines. The velocities and accelerations for the Poser skeleton models were calculated using the spline package in quintic mode with a 7 Hz cut-off. The joint angles presented here were calculated using the helical angular convention of Woltring (1994). To evaluate the error originating from each of the matched segments, we also calculated the segment orientation, expressed in helical angles relative to the global system. The estimated velocities from the matchings were compared with the velocities derived from the skin marker measurements. The accelerations were compared with those derived from the force plate measurements. The results are presented as the root mean square (RMS) and maximal differences.

3. Results

In both the knee and hip, the best agreement was found for the flexion angles (Figs. 3 and 4). Although hip ad-/abduction was shifted 10–15° into abduction compared with the Pro Reflex measurements (Table 1), the shapes were very similar (Fig. 3). The assessment of internal and external rotation was clearly less precise, especially for the hip (Figs. 3 and 4).

Fig. 5 shows a comparison of the different matchings of the side step cutting maneuver at toe-off. The configuration of the rear and side single camera matchings at this point in time, appears slightly different from the others when seen in a frontal view (Fig. 5a). The rear/side matching, on the other hand, appeared quite similar the other matchings (which all include a front camera). Although the body configurations all look reasonably similar in Fig. 5, the pelvis orientation

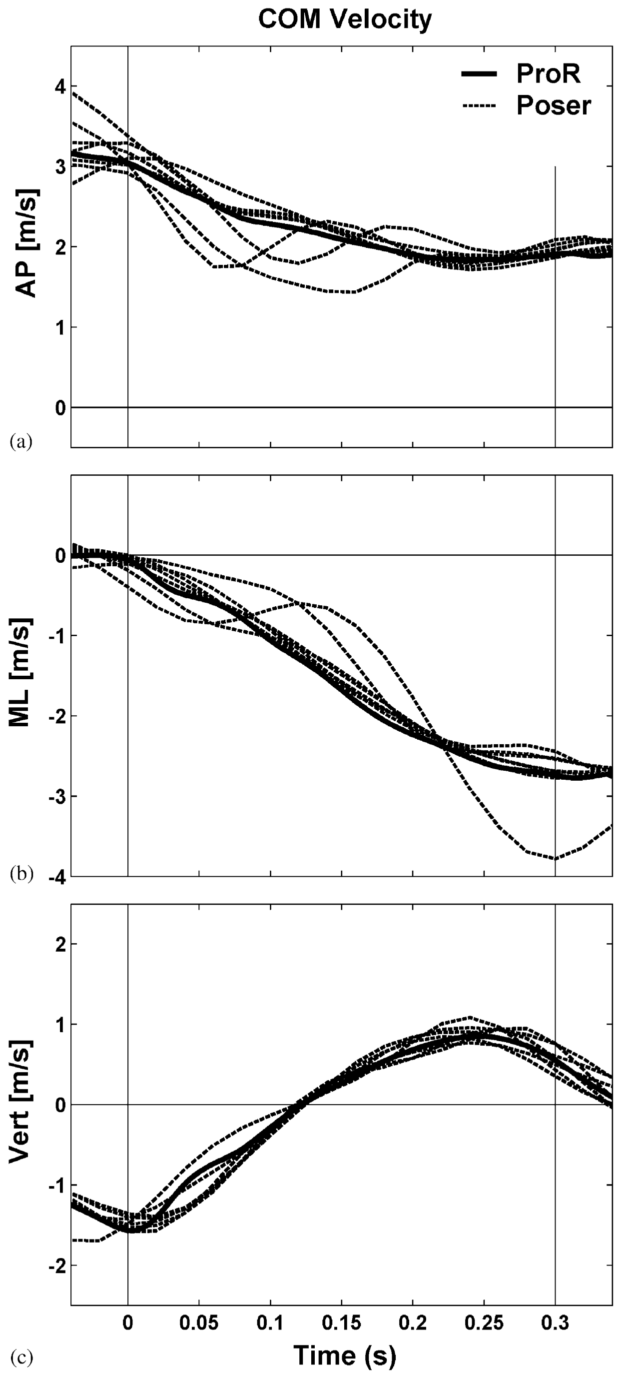

Fig. 6. COM velocity (m/s), calculated with the Pro Reflex measurements (solid lines) and the model-based image-matching technique (dotted lines) for each of the seven camera combinations. The vertical lines indicate initial ground contact of the support leg (time zero) and toe-off.

for all the matchings differed considerably from the Pro Reflex recording (also shown in Table 2). The thigh orientation differences were moderate, while the shank differences were small for all the matchings (Table 2).

Velocity estimations were generally good, especially for the vertical direction (Fig. 6). In the antero-posterior direction the errors were greater for the matchings that did not contain the side view (Table 3). Likewise, for the medio-lateral direction the largest errors were seen for the matchings that did not contain the front view.

For the accelerations, only the matchings containing at least two camera views gave reasonable estimates (Table 3). The matchings that contained perpendicular camera views (rear/side, side/front, rear/side/front) were clearly better than the matching with opposing camera views (rear/front). However, as seen in Fig. 7, the triple camera matching did not detect the initial peak force with the same sensitivity as the Pro Reflex recording.

The results were very similar for the jogging and cutting motions. For the jogging, the pelvis matching was slightly better with an RMS error for tilt ranging between 2.0° and 10.0°, for obliquety ranging between 3.1° and 6.5° and for rotation ranging between 5.4° and 12.8°. The thigh RMS errors were 2.4–9.2°, 0.8–5.0° and 3.5–9.4° for flexion, adduction and rotation, respectively. For the shank the corresponding errors were 1.5–6.3°, 3.2–4.8° and 7.8–14.1°. As for the cutting motion, the velocity errors were largest in the planes where we did not have camera information (antero-posterior direction for the front camera matching, etc.).

4. Discussion

The aim of this study was to evaluate a novel method for the reconstruction of human motion from one or more uncalibrated video sequences. We found that three-dimensional motion could be successfully reconstructed with the new method, providing kinematic estimates which until now have been unavailable using uncalibrated videos. For some research purposes, e.g. to describe knee rotations during injury situations, joint angles has until now been done by simple visual assessment (Boden et al., 2000; Olsen et al., 2003). Although we have not tested this formally, our experience during the matching procedure indicates that using a skeleton model in the assessment can substantially change our interpretation of the actual segment poses. The model-based image-matching procedure is adequate for some research purposes, e.g. to describe body velocity and knee angle patterns in injury situations, provided that at least two camera views are present, and that the video quality is good. However, joint translations and high-frequent acceleration peaks cannot be realistically estimated before video frame rate and resolution standards are significantly improved upon. We did not find any substantial differences between the quality of matchings between the complex cutting motion and the more simple jogging motion.

To develop effective injury prevention measures, a more precise description of the injury mechanisms, as well as the motion preceding the injury situation

Table 3
Root mean square and maximal difference in velocity (m/s) and acceleration (m/s^2) for the center of mass (COM) between the Pro Reflex recordings (force plate recordings for acceleration) and the estimates from our model-based matching-technique for each of the seven matchings of the plant and cut motion

		Camera views						
		Rear	Side	Front	Rear/side	Rear/front	Side/front	Rear/side/front
Vel.	Antero-posterior	0.38 (0.70)	0.17 (0.39)	0.31 (0.78)	0.10 (0.18)	0.31 (0.78)	0.11 (0.28)	0.09 (0.19)
	Medio-lateral	0.34 (0.69)	0.62 (1.06)	0.15 (0.26)	0.19 (0.32)	0.12 (0.22)	0.10 (0.22)	0.11 (0.22)
	Vertical	0.10 (0.22)	0.19 (0.44)	0.14 (0.33)	0.13 (0.36)	0.15 (0.33)	0.14 (0.36)	0.13 (0.39)
Acc.	Antero-posterior	5.2 (13.7)	4.4 (12.0)	7.9 (15.7)	3.2 (7.9)	9.7 (19.3)	2.8 (7.0)	2.8 (6.4)
	Medio-lateral	6.5 (16.5)	12.7 (20.3)	3.1 (16.6)	5.8 (11.0)	2.7 (10.4)	2.6 (11.9)	3.5 (11.1)
	Vertical	5.2 (13.6)	5.9 (22.1)	4.8 (19.0)	5.9 (21.8)	5.4 (22.6)	5.4 (19.3)	4.9 (19.4)

The maximal differences are shown in parentheses.

(Ettlinger et al., 1995), is necessary. Methods have so far not been available to provide such information. Using the present method, the RMS differences in hip and knee joint angles ranged between 2.6° and 15.7° (see Table 1), and the joint angle patterns were very similar between the Poser matching and the reference method for flexion/extension, ad-/abduction, and varus/valgus motion (see Figs. 3 and 4). This indicates that the method may be adequate to describe essential kinematic characteristics associated with injuries based on video recordings, e.g. anterior cruciate ligament (ACL) injuries to the knee. The injury mechanism for non-contact ACL injuries in team handball, usually a side-step cutting maneuver (Myklebust et al., 1997, 1998; Olsen et al., 2003), appears to be consistent across nearly all of our videos (Olsen et al., 2003). Thus, it is possible that an average kinematic description based on several two- and three-camera matchings will be sufficient to produce estimates of joint forces and moments.

Nevertheless, there are some methodological limitations, which must be kept in mind. In injury research, the ultimate goal is to obtain the forces and moments causing the injury, which implies accurate estimation of accelerations. Although the error in the acceleration estimate from the triple camera matching was small during the last half of the stance (Fig. 7), we were not able to capture the high frequent dynamics of the impact, mainly because of the limited video frame rate (50 Hz). By down-sampling the original Pro Reflex recording from 240 to 50 Hz, the RMS acceleration error increased by 48%, 28% and 56% in the antero-posterior, medio-lateral and vertical directions, respectively, resulting in an acceleration estimate very similar to what was produced with the model-matching technique. The vertical impact force peak from the heel strike was, likewise, no longer detectable. Even though we have not made any attempts to calculate inverse dynamics in this study, it seems clear that this would require *a minimum* of two camera views.

As may be expected, the number of available camera views and their relationship are important. The best results were obtained with the side/front and rear/side/front cameras, in particular for velocity and acceleration (see Table 3). This is probably both due to the fact that they were nearly perpendicular to each other, and thus provided complimentary information, as well as the relatively high resolution of the test subject in the front and side views. Not surprisingly, movement perpendicular to the camera view resulted in the largest errors, e.g. antero-posterior velocity for the front camera matching (Table 3). This fact, combined with the somewhat lower resolution of the test subject in the image, made flexion/extension angle estimates for the rear camera difficult (see Table 1). However, sports events are usually taped with several cameras, and we therefore recommend that multiple, perpendicular views are used to obtain results that are more accurate.

The largest differences appeared in the pelvis matching. This is not surprising, since the shape of the pelvis makes interpretation of its attitude difficult. Even with the combined information from all three camera views we obtained up to a 22° difference in pelvis rotation, compared to the marker-based measurements (see Table 2). The matching of the shank was, on the other hand, very good. This implies that estimates for the knee joint center will probably be more accurate than for the hip joint center.

Even though we have compared this model-based image-matching method to an assumed gold standard, we are fully aware that skin marker-based optical systems may be compromised by errors—especially during high-impact motion. Skin markers can give errors of almost 15° in thigh orientation angles during slow speed running (Reinschmidt et al., 1997). From the video recordings, we observed what seemed to be quite substantial thigh marker motion relative to the underlying bone, especially at foot strike. We also measured up to 25° of varus in the right knee during the support phase (of the left leg), an unlikely result in a non-injury situation, most likely resulting from a lack of external rotation of the soft tissue of the thigh relative to the underlying bone.

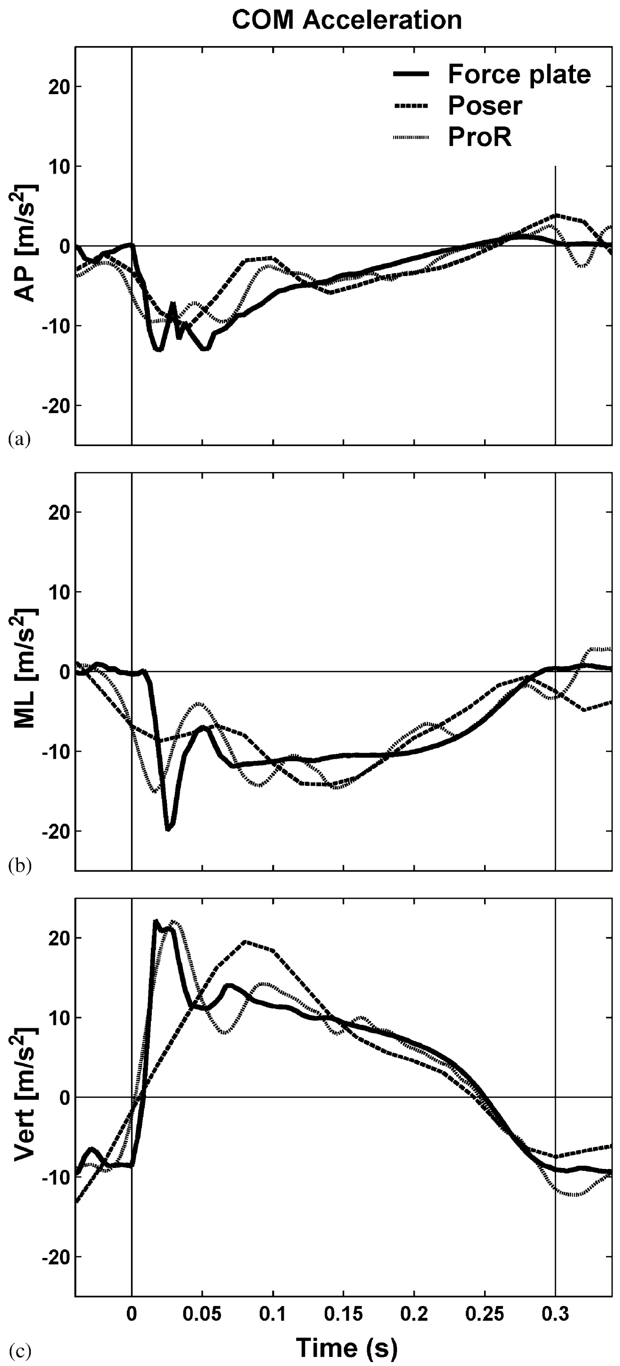

Fig. 7. COM acceleration (m/s^2), calculated from the force plate (solid lines), from the model-based image-matching technique for the triple camera matching (dotted lines), and also from with the Pro Reflex measurements (dashed lines). The vertical lines indicates initial ground contact of the support leg (time zero) and toe-off.

The ability to incorporate one, two or even more cameras for the motion reconstruction makes the proposed method versatile. In addition, the ability to reconstruct camera motion parameters (rotation, translation, and focal length) for every frame makes it possible to handle most videos from television broadcasts or amateur film. In this validation study, we had a number of background references in the surroundings suitable for modeling. This can probably be considered a best case scenario, but still representative. In many indoor sports such as basketball and European team handball, there are several lines and markings on the floor, as well as the basket, goal posts and advertising boards. In situations where there is not sufficient calibration information in the field of view to reconstruct the camera parameters, estimates of the body configuration can still be obtained by using the person as reference for the camera location(s) similar to the approach of Zheng et al. (2000). A drawback, however, is that the technique involves manual frame-by-frame matching by the operator, which means that the method is time consuming, and possibly biased by the operator's subjective judgement. Our strategy to minimize bias resulting from single-operator judgement was—as outlined in the Methods—to use a consensus procedure, whereby several experienced clinicians gave feedback on the matchings until a final motion pattern was produced that looked optimal. However, it should be noted that the inter- or intratester reproducibility of the method has not been examined.

Although the model is fitted frame-by-frame, the preceding and subsequent frames provide additional information about the segment poses at each point. Rendering videos at different frame rates gives a good indication of whether the motion is correct or not. If the motion does not appear correct, and corrections must be made at one point, the consequence is often that adjustments also must be made at different time-points in order to produce smooth, consistent motion patterns. This means that matching a video sequence of 50 frames can take more than 150 h independent of the number of camera views, even for an experienced operator. The advantage, on the other hand, is that manual assessment allows us to simultaneously utilize edge, surface, color, contrast, segment shape, texture and size as well as point landmark properties, to provide optimal motion estimates without the need to implement these complex criteria mathematically. In the computer vision literature, we also see that manual assessment of key frames is often needed to increase the accuracy of the tracking and reconstruction (Wilhelms et al., 2000; Yamamoto et al., 2000; Zheng et al., 2000; Perales and Torres, 1994; Taha et al., 1997). Furthermore, using a model is the key to reconstruct the poses of the body segments reliably. Our approach is in this way comparable to global optimization techniques as described by Lu and O'Connor (1999), where segment length constraints are utilized to make the three-dimensional reconstruction more robust. Finally, using a model is also what makes us able to obtain a unique solution of the three-dimensional reconstruction problem when only one video view is available, similar to what is seen in the method of Ambrosio et al. (2001). An anatomically correct model is therefore an important requirement for a good result.

5. Conclusion

A new model-based image-matching technique has been developed, that can produce temporal joint angle histories, velocities and accelerations from uncalibrated video recordings. The method can potentially provide knowledge about human motion in situations where traditional motion analysis is not possible, e.g. in the study of non-contact ACL injury mechanisms. The kinematic estimates, in particular, COM velocity and acceleration, are clearly better when two or more camera views are available.

Acknowledgements

The Oslo Sports Trauma Research Center has been established at the Norwegian University of Sport & Physical Education through generous grants from the Royal Norwegian Ministry of Culture, the Norwegian Olympic Committee & Confederation of Sport, Norsk Tipping AS, and Pfizer AS. We thank Lars Engebretsen, Gerald Smith, Thomas Moeslund, Svein Linge, Oddvar Hallingstad and Stein By Stroemsnes for valuable comments.

References

Ambrosio, J., Lopes, G., Costa, J., Abrantes, J., 2001. Spatial reconstruction of the human motion based on images of a single camera. Journal of Biomechanics 34, 1217–1221.

Bahr, R., Pena, F., Shine, J., Lew, W.D., Lindquist, C., Tyrdal, S., Engebretsen, L., 1997. Mechanics of the anterior drawer and talar tilt tests. A cadaveric study of lateral ligament injuries of the ankle. Acta Orthopaedica Scandinavica 68, 435–441.

Barone, M., Senner, V., Schaff, P., 1999. ACL injury mechanism in alpine skiing: analysis of an accidental ACL rupture. In: Johnson R.J. (Ed.), Skiing Trauma and Safety. 12th Edition, ASTM STP 1345, American Society for Testing and Materials, West Conshohocken, PA, pp. 63–81.

Bell, A.L., Pedersen, D.R., Brand, D., 1990. A comparison of the accuracy of several different hip center location prediction methods. Journal of Biomechanics 23, 617–621.

Boden, B.P., Dean, G.S., Feagin Jr., J.A., Garrett Jr., W.E., 2000. Mechanisms of anterior cruciate ligament injury. Orthopedics 23, 573–578.

Davis III, R.B., Ounpuu, S., Tyburski, D., Gage, J.R., 1991. A gait analysis data collection and reduction technique. Human Movement Science 10, 575–578.

de Leva, P., 1996. Joint center longitudinal positions computed from a selected subset of Chandler's data. Journal of Biomechanics 29, 1231–1233.

Eng, J.J., Winter, D.A., 1995. Kinetic analysis of the lower limbs during walking: what information can be gained from a three-dimensional model? Journal of Biomechanics 28, 753–758.

Ettlinger, C.F., Johnson, R.J., Shealy, J.E., 1995. A method to help reduce the risk of serious knee sprains incurred in alpine skiing. American Journal of Sports Medicine 23, 531–537.

Gavrila, G.M., 1999. The visual analysis of human movement: a survey. Computer vision and Understanding 73, 82–98.

Gerritsen, K.G., Nachbauer, W., van den Bogert, A.J., 1996. Computer simulation of landing movement in downhill skiing: anterior cruciate ligament injuries. Journal of Biomechanics 29, 845–854.

Halvorsen, K., 2002. Model-based Methods in Motion Capture. Ph.D. thesis, Department of Information Technology, University of Uppsala, Sweden.

Hame, S.L., Oakes, D.A., Markolf, K.L., 2002. Injury to the anterior cruciate ligament during alpine skiing: a biomechanical analysis of tibial torque and knee flexion angle. American Journal of Sports Medicine 30, 537–540.

Lu, T.W., O'Connor, J.J., 1999. Bone position estimation from skin marker co-ordinates using global optimisation with joint constraints. Journal of Biomechanics 32, 129–134.

Markolf, K.L., Burchfield, D.M., Shapiro, M.M., Shepard, M.F., Finerman, G.A., Slauterbeck, J.L., 1995. Combined knee loading states that generate high anterior cruciate ligament forces. Journal of Orthopaedic Research 13, 930–935.

McLean, S.G., Su, A., van den Bogert, A.J., 2001. Simulation of acute musculoskeletal injuries in humans: a novel approach. Proceedings of the Eight International Symposium on Computer Simulation in Biomechanics, 5–6 July, Milano, Italy, pp. 23–27.

Moeslund, T.B., Granum, E., 2001. A survey of computer vision-based human motion capture. Computer Vision and Understanding 81, 231–268.

Myklebust, G., Maehlum, S., Engebretsen, L., Strand, T., Solheim, E., 1997. Registration of cruciate ligament injuries in Norwegian top level team handball. A prospective study covering two seasons. Scandinavian Journal of Medicine & Science in Sports 7, 289–292.

Myklebust, G., Maehlum, S., Holm, I., Bahr, R., 1998. A prospective cohort study of anterior cruciate ligament injuries in elite Norwegian team handball. Scandinavian Journal of Medicine & Science in Sports 8, 149–153.

Olsen, O.E., Myklebust, G., Engebretsen, L., Bahr, R., 2003. Injury mechanisms for ACL injuries in team handball—a video analysis. American Journal of Sports Medicine, submitted for publication.

Perales, F.J., Torres, J., 1994. A system for human motion matching between synthetic and real images based on a biomechanic graphical model. In: Workshop on Motion of Non-Rigid and Articulated Objects, Austin, TX, USA, pp. 83–88.

Reinschmidt, C., van den Bogert, A.J., Nigg, B.M., Lundberg, A., Murphy, N., 1997. Effect of skin movement on the analysis of skeletal knee joint motion during running. Journal of Biomechanics 30, 729–732.

Soderkvist, I., Wedin, P.A., 1993. Determining the movements of the skeleton using well-configured markers. Journal of Biomechanics 26, 1473–1477.

Taha, Z., Brown, R., Wright, D., 1997. A neural network approach to markerless measurement of human motion. Biomedical Science and Instrumentation 33, 441–446.

Trewartha, G., Yeadon, M.R., Knight, J.P., 2001. The sensitivity of video tracking performance to the simulated environment, the human body model and the image matching algorithm. Proceedings of the Eight International Symposium on Computer Simulation in Biomechanics, 5–6 July, Milano, Italy, pp. 187–192.

Wilhelms, J., Gelder, A.V., Atkinson-Derman, L., Hong, L., 2000. Human Motion from Active Contours. Proceedings of the IEEE

Computer Society Workshop on Human Motion, December, Austin, TX, pp. 155–160.
Woltring, H.J., 1986. A Fortran package for generalized, cross-validatory spline smoothing and differentiation. Advance in Engineering Software 8, 104–113.
Woltring, H.J., 1994. 3-D attitude representation of human joints: a standardization proposal. Journal of Biomechanics 27, 1399–1414.
Yamamoto, M., Ohta, Y., Yamagiwa, T., Yagishita, K., Yamanaka, H., Ohkubo, N., 2000. Human action tracking guided by key-frames. Fourth IEEE International Conference on Automatic Face and Gesture Recognition 2000, March 26–30, Grenoble, France, pp. 354–361.
Yeadon, M.R., 1990. The simulation of aerial movement—II. A mathematical inertia model of the human body. Journal of Biomechanics 23, 67–74.
Zernicke, R.F., Garhammer, J., Jobe, F.W., 1977. Human patellar-tendon rupture. Journal of Bone and Joint Surgery American volume 59, 179–183.
Zheng, J.Y., Suezaki, S., Shiota, Y., 2000. Interactive Human Motion Acquisition from Video Sequences. IEEE Computer Society, Computer Graphics International, Silver Spring, MD, pp. 209–217.