Pitching a Baseball — Tracking High-Speed Motion with Multi-Exposure Images

Christian Theobalt Irene Albrecht Jörg Haber Marcus Magnor Hans-Peter Seidel

MPI Informatik, Saarbrücken, Germany

{theobalt,albrecht}@mpi-sb.mpg.de, haberj@acm.org, {magnor,hpseidel}@mpi-sb.mpg.de

Abstract

Athletes and coaches in most professional sports make use of high-tech equipment to analyze and, subsequently, improve the athlete's performance. High-speed video cameras are employed, for instance, to record the swing of a golf club or a tennis racket, the movement of the feet while running, and the body motion in apparatus gymnastics. High-tech and high-speed equipment, however, usually implies high-cost as well. In this paper, we present a passive optical approach to capture high-speed motion using multi-exposure images obtained with low-cost commodity still cameras and a stroboscope. The recorded motion remains completely undisturbed by the motion capture process. We apply our approach to capture the motion of hand and ball for a variety of baseball pitches and present algorithms to automatically track the position, velocity, rotation axis, and spin of the ball along its trajectory. To demonstrate the validity of our setup and algorithms, we analyze the consistency of our measurements with a physically based model that predicts the trajectory of a spinning baseball. Our approach can be applied to capture a wide variety of other high-speed objects and activities such as golfing, bowling, or tennis for visualization as well as analysis purposes.

CR Categories: I.4.1 [Image Processing and Computer Vision]: Digitization and Image Capture—Camera calibration I.4.6 [Image Processing and Computer Vision]: Segmentation—Edge and feature detection I.4.8 [Image Processing and Computer Vision]: Scene Analysis—Motion, tracking

Keywords: high-speed motion capture, multi-exposure images, pitching and flight of baseball, physically based validation

1 Introduction

Creating images of high-speed motion for analysis of the underlying action has been drawing the attention of researchers for many decades. Back in 1878, Eadweard Muybridge conducted his famous experiments to create serial images of fast motion [Muybridge 1887]. A setup of twelve cameras was used to capture different stages of a galloping horse. One of the photographs indeed showed the horse with all of its hooves off the ground, corroborating the hypothesis that had led to these experiments. In the 1930's, Harold E. Edgerton at MIT perfected the use of stroboscope photography to create multi-exposure images of high-speed motion, see for instance [Collins and Bruce 1994]. However, the acquisition process is usually constrained to actions taking place in a very limited spatial domain for which decent illumination conditions can be set up easily. In this paper, we address the problem of capturing and tracking high-speed motion sequences that cover large areas of space. To avoid the expenses of professional high-speed video cameras and high-resolution motion capture equipment, our approach is based on low-cost commodity still cameras and strobe lights.

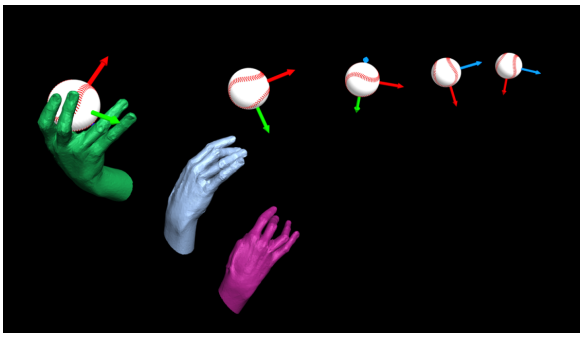

Figure 1: Visualization of ball and hand motion obtained from multi-exposure images. The hand motion during release of the baseball is captured and shown together with the resulting flight characteristics of the ball.

We demonstrate our motion capture approach for the hand motion during pitching and the flight trajectory of a baseball. Besides the popularity of baseball, there are several reasons for our choice. First of all, the underlying motion is very fast and extends over a large area of space: the speed of a pitched baseball can reach 80 mph and above, and the distance from the pitcher mound to the home base is 60.5 feet (18.44 meters). In addition, there are many different motion parameters that we would like to measure simultaneously for a variety of pitches:

- 3D positions along the trajectory of the flying ball;
- initial flight parameters of the ball: norm and direction of initial velocity, rotation axis, spin frequency;
- positions of the finger joints of the pitcher's hand before, at, and after releasing the ball.

Finally, it is possible to use a physically based model to analyze the consistency of the acquired data: if the ball's initial parameters and flight positions are reconstructed with high accuracy, they should match the results from a physically based model that predicts the flight trajectory of a spinning ball traveling through air. In summary, both the pitching and flight of a baseball turn out to be a challenging and adequate type of motion for our motion capture approach alike.

1.1 Contributions

In this paper, we introduce the following main contributions:

- an approach for capturing high-speed motion using multi-exposure images obtained with low-cost commodity still cameras and stroboscopes;

- an algorithm to automatically compute the 3D positions and the initial flight parameters of a baseball from multi-exposure images;
- a procedure to reconstruct articulated hand motion from multi-exposure images;
- validation of the approach by means of a physically based model of the flight of a baseball.

2 Background

In many sports highly competitive professional leagues have evolved where technical perfection has become essential for athletes in order to keep up with their rivals. Baseball has always been one of the most popular sports in the U.S. and many countries in Asia. Due to its variety of different elements it is technically very challenging. In particular, pitching is the single most important part of baseball. The goal for the pitcher is to throw the ball in such a way that its trajectory is as unpredictable as possible for the other team's batter. In the history of baseball a great variety of pitches has been developed. They differ in the way the pitcher's hand and fingers move during release of the ball. Different hand motions lead to different initial velocities, rotation axes, and spin frequencies of the ball, which in turn lead to different flight trajectories. The art of pitching is to be able to perform all kinds of pitches such that the ball consistently enters the strike zone near the batter in order to be valid. The pursuit of athletic perfection in baseball already starts in the minor leagues and has lead to the publication of many textbooks on specific technical aspects [Stewart 2002; House 2000].

In recent years, the athletes' demand for tools to accurately measure and analyze their technical performance has been backed by similar interests from the media. Today, many sports enthusiasts expect concise analysis and visualization of a sports event during or after the broadcast on TV. In consequence, many researchers have approached baseball from the scientific and technological point of view. The physics of pitching and batting has been thoroughly analyzed in [Adair 2002]. Alaways examined in his PhD thesis [1998] the aerodynamics of a curve-ball. He used a system with ten high-speed video cameras operating at 240 Hz to capture the ball trajectory. Initial flight parameters of the ball were not measured but deduced from the trajectory and a physical model of ball flight. During the Summer Olympics 1996 in Atlanta, Alaways used two 120 Hz high-speed video cameras to track ball positions along the flight trajectory [Alaways et al. 2001]. The K-Zone system [Gueziec 2003; Gueziec 2002] is technically similar and designed to track the trajectory of a baseball from multiple video streams in real-time using color information and a Kalman filter. In other popular sports similar systems have been investigated. The LucentVision system [Pingali et al. 2000] enables tracking of the player positions and the ball trajectory in tennis matches from video images. The ball position is tracked using an algorithm based on ball color and frame-differencing [Pingali et al. 1998]. Rotation axis and spin are not measured. In [D'Orazio et al. 2002], a modified Circular Hough Transform is used to follow the ball in video broadcasts of a soccer game.

Due to the limited video resolution, the previously mentioned systems do not provide the same spatial accuracy as our approach, which uses high-resolution image sensors built into digital still cameras.

In this work we will demonstrate that stroboscope photography is not only an appropriate method to accurately track the ball trajectory but also to track the complex articulated motion of the human hand. Many different approaches to tracking articulated human body motion have been investigated in the past, spanning from mechanical over magnetic to optical methods that either rely on optical markers on the human body or set aside any form of intrusion into the scene [Aggarwal and Cai 1999; Gleicher et al. 2001]. Commercial optical motion capture systems typically rely on expensive high-framerate video cameras and markers on the body.

Optical approaches for tracking hand articulation usually derive the hand motion from video sequences and with the support of an explicit hand model. In [Heap and Hogg 1996], a point distribution model is used to track hand motion. Stenger *et al.* [2001] employ a kinematic model based on quadric segments and a Kalman filter to determine hand configurations from video. In [Wu et al. 2001], a 2D cardboard hand representation is used for pose computation. Other approaches that rely on an explicit hand model and image features are the Digiteyes system [Rehg and Kanade 1994] and the work in [Dorner 1993] where colored markers on the hand show the finger joint locations in the video images. A more appearance-based approach is presented in [Athitsos and Sclaroff 2003] where single hand poses are identified via comparison to a database of rendered hand models.

In contrast to video-based motion capture systems, we apply a marker-based tracking algorithm to the hand that derives hand poses from still images recorded with stroboscope photography at high accuracy.

3 Setup

We use a flexible setup to robustly acquire different types of motion data under real-world conditions. To analyze flight trajectories of different pitches we need to acquire image data that allows us to reconstruct the ball's initial flight parameters (i.e. norm and direction of its velocity, direction of its rotation axis, and spin) as well as the 3D positions of the ball along its trajectory. In addition, we want to capture the motion of the pitcher's hand and fingers before, during, and after releasing the ball.

Acquiring this type of information is very challenging since the involved speeds are considerable and the entire trajectory extends over a relatively wide area. To complicate things even further, high spatial accuracy is essential in both flight analysis and hand motion capture.

In this work we demonstrate how to acquire the necessary high-accuracy image data without resorting to expensive, specialized hardware such as high-speed video cameras. To obtain simultaneously high spatial as well as temporal resolution, we apply stroboscope photography. We capture an image of the high-speed scene in a darkened room using a regular digital still camera set to a long exposure time. The scene is illuminated with a stroboscope light that emits short light pulses at a suitable frequency. The resulting image depicts, superimposed, the dynamic scene at different, closely-spaced time instants with the temporal sampling frequency set by the frequency of the stroboscope. High spatial accuracy is easily achieved by using recent commodity digital cameras with multi-million pixel resolution.

To capture an entire baseball pitch, we set up our acquisition gear in a basement room which has a central free space area of approximately 25 m length, 4 m height, and 5 m width. This is sufficient to house the complete pitching corridor (18.44 m in length) as well as to put up the camera and lighting equipment. As imaging devices we employ consumer-market Olympus^{TM}Camedia C5050 still image cameras that provide a frame resolution of 2560x1920 pixels. This camera model features a large-aperture zoom lens that can be set to a comparatively wide angle. We use four cameras of this type in our setup. Control software was developed enabling us to control the settings of all four cameras from a single PC and to trigger all camera shutters simultaneously.

Since we intend to record a fairly wide-area scene, we need a sufficiently luminous stroboscope light source that can illuminate a large volume at high frequencies. In our setup we use two high-output strobe flashes which have an intensity of 5000 Lux each at

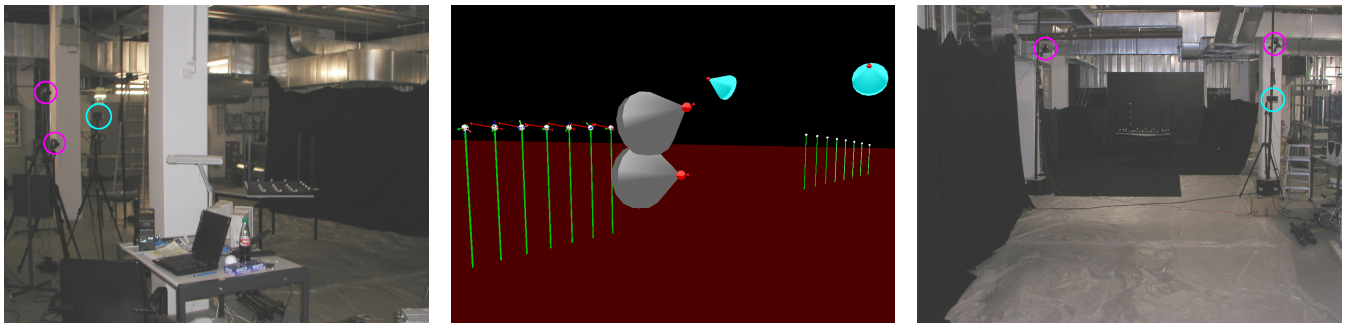

Figure 2: Ball acquisition setup. Left: a stereo camera pair (encircled in magenta) facing the black curtain on the right is capturing the ball's initial flight parameters. The ball is illuminated by a stroboscope (cyan). Middle: schematic illustration of the setup. The two grey cameras correspond to the cameras in the left image. For the initial part of the ball trajectory, the measured ball positions are shown in white, recovered rotation axes in red, and height above ground in green. The two cyan cameras observe the last third of the trajectory and correspond to the cameras in the right image. Right: a stereo camera pair (magenta) and a strobe light (cyan) facing towards the black carpet in the back are responsible for capturing the ball trajectory close to the "home base".

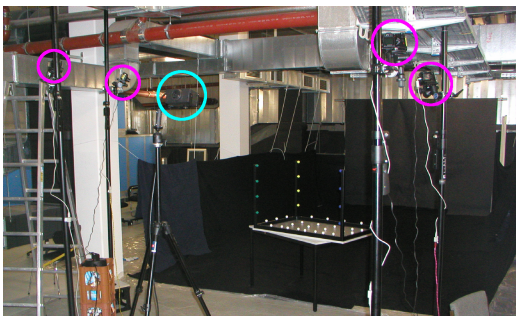

Figure 3: Two stereo camera pairs (magenta) and a strobe light (cyan) are placed in a semi-circular arrangement around the pitcher to capture the hand motion of different pitches.

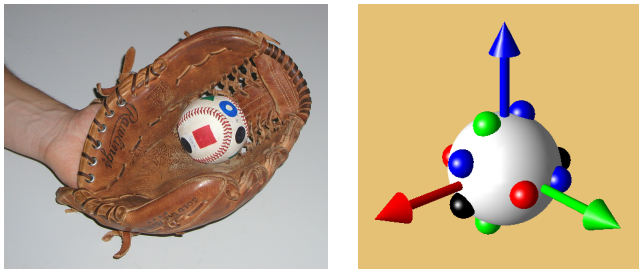

Figure 4: Left: baseball equipped with optical markers in pitcher's glove. Right: illustration of the ball coordinate system. Markers are depicted as small colored spheres on the ball.

a distance of 0.5 m from the lamp. At full intensity, the 20 μs-long flashes can be triggered up to 80 Hz which is sufficiently fast for our purposes.

During recording the floor and walls are covered with black carpet and cloth to facilitate foreground object segmentation and automatic marker tracking. Primarily, however, the dark material absorbs most light that has not hit foreground objects, preserving contrast and preventing quick saturation of the multi-exposure images. Finally, a heavy dark carpet hanging down from the ceiling at the end of the flight corridor absorbs the impact of the ball.

In our recordings, four simultaneously triggered cameras look at the scene from different positions. Two different arrangements of imaging sensors and light sources are needed to record either initial flight parameters and ball positions (see Figure 2) or the hand motion of the pitcher (see Figure 3).

To record the baseball in flight, two stereo pairs of cameras and two stroboscopes are used to capture the initial and final phase of the ball flight, respectively (Figure 2). Details about the setup for acquisition of ball motion are given in Section 4.2.

For recording the hand, the four cameras and one light source are placed in a semi-circular arrangement looking at the pitcher from behind and above, see Figure 3. Section 5.2 gives for further details about this step.

A crucial and—for a large setup like ours—challenging task is the accurate calibration of the cameras. We apply a camera model for short focal length cameras [Heikkila and Silven 1996]. Intrinsic camera parameters are estimated from images of a planar checkerboard pattern. Radial and tangential lens distortion are modeled up to second order [Jain et al. 1995] and each multi-exposure image is distortion-corrected prior to any further processing. Extrinsic camera parameters are estimated using images of our 3D calibration object, see Figure 3. Camera position and orientation are metrically calibrated.

Finally, we rely on our professional baseball pitcher who, as we have verified, performs different pitches with great faithfulness. This allows us to correlate our measurements of hand motion with the measurements of initial flight parameters and flight trajectory.

4 Tracking the Ball

4.1 Preparation of the Ball

We paint optical markers on the ball to be able to estimate its spatial orientation from multi-exposure images. Four different types of markers are used which differ in color and shape (red square, blue ring, green triangle, black circle). Over the entire surface of the ball, each marker type is used three times. Eight markers are arranged in the ball's equatorial plane, in 30°-pairs and with 60° inter-pair separation. The remaining four markers are located in a second, orthogonal plane at 30° distance from the poles. Marker types are assigned such that at least three different markers are visible from any viewpoint. In addition, the (fixed) coordinate system of the ball can be determined from the marker positions for an arbitrary viewing direction (Figure 4).

4.2 Acquisition of Ball Motion

In our experiments we focus on the fast-ball, the curve-ball, the slider, and the change-up, all of them performed as three-quarter

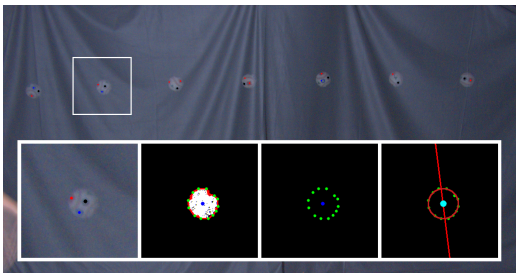

Figure 5: Multi-exposure image of the ball to recover its initial flight parameters. Automatically detected markers are shown as colored dots. Inset, left to right: magnified image region, result after background subtraction, detected ball silhouette and predicted center point, fitted circle and final center point (see Section 4.3).

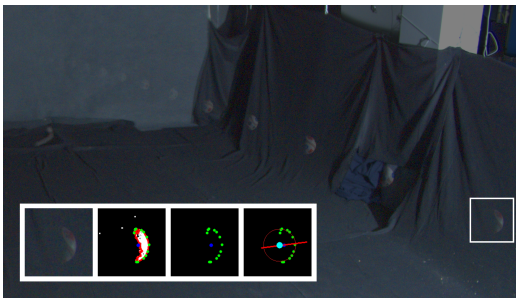

Figure 6: Multi-exposure image taken by one of the back cameras. The half-moon shape of the balls is due to the lateral position of the stroboscope illuminating the flight path. The inset shows the same processing steps as the one in Figure 5.

deliveries, i.e. with a release point above and to the right of the head. Each of these pitches was recorded multiple times. Pitches differ in the way how the hand moves during launch, giving the ball a different initial velocity, rotation axis, and spin frequency. Since these initial flight parameters completely determine the ball's trajectory, different pitches lead to different flight paths.

- The fast-ball is the fastest pitch. It has large back spin and, depending on whether it rotates over four or only two of its seams, the ball is called a 4-seamer or a 2-seamer.
- The change-up also exhibits back spin but has a lower velocity and spin frequency.
- The curve-ball is released with forward spin which makes the ball descend faster during the last phase of its flight.
- The slider is thrown with a side spin, making the ball turn to one side towards the end of the flight.

To acquire information about the flight of a baseball, two pairs of cameras are used that focus on different aspects of the ball trajectory. The front two cameras take multi-exposure pictures of the first 5 m of the baseball's trajectory right after the ball has left the pitcher's hand. The cameras are placed 3.5 m away from the flight path and are vertically aligned with a baseline of approximately 0.8 m, see Figure 2 (left). One strobe light is placed close to the cameras and illuminates the scene such that the ball silhouette appears as a circular shape in the images. In both cameras' multi-exposure images the ball is seen at several subsequent positions and orientations, flying from left to right in Figure 5. The number of visible ball positions is determined by the pulse frequency of the stroboscope. At a strobe light frequency of 80 Hz, 6–10 ball positions are captured, depending on the speed of the pitch.

The stereo camera pair in the back part of the setup records the last third of the flight trajectory close to the "home base" where the most interesting variations between different pitches occur. The cameras are placed approximately 2.8 m high and 4 m apart on either side of the flight corridor, see Figure 2 (right). A second stroboscope is located below the right camera and illuminates the ball at 50 Hz. This lower frequency is chosen to better separate the ball in the multi-exposure images. In contrast to the camera setup in the front, the illumination direction in the back setup causes partially illuminated ball silhouettes shown in Figure 6. We compensate for this before reconstructing ball positions, see Section 4.3.

During recording, the shutters of the front cameras are open for one second, while the shutters of the back cameras expose for 1.3 seconds. All cameras are triggered simultaneously. As a trade-off between image noise and brightness, we run each camera with ISO 200 sensitivity.

The 3D positions of the ball in flight are recovered via triangulation, see Figure 7, and the orientation of the ball's coordinate frame is computed. Then the ball's rotation axis and spin frequency are determined, see Figure 7 (left). At 80 Hz, a ball at a spin rate of 1600 rpm rotates by 120° between subsequent strobe flashes. Our sampling frequency is more than twice the spin frequency and therefore sufficiently high to fulfill the Nyquist criterion.

4.3 Reconstructing Flight Positions

In each multi-exposure image, the silhouettes of the ball in the foreground are separated from the background by means of a color-based background subtraction, thereby creating binary foreground masks.

In both the front and back stereo pair of images, the ball silhouettes' boundary polygons are identified via a contour finding algorithm (OpenCV [Intel 2002]). To correct small concavities at the silhouette boundaries of the balls we compute the convex set of the vertices of each boundary polygon [Slansky 1970].

First estimates of the projected ball center locations in each image are found via fitting ellipses to the silhouette boundary points. In each stereo pair of images correspondences between projected ball center locations are established via epipolar geometry [Faugeras 1993]. Approximate estimates of the 3D ball center locations are obtained by means of triangulation from corresponding ellipse centers in each stereo pair. The center estimates in the image planes are further improved by fitting implicit circle models to the silhouette boundaries by means of a Circular Hough Transform [Ballard 1981] in a local neighborhood of each ellipse center. Knowing the radii of the projected balls from the first 3D reconstruction, the Hough Transform search space reduces to two dimensions (center location). The final 3D positions are found from stereo reconstruction of the circle centers.

The whole fitting pipeline is illustrated in the insets of Figure 5 and Figure 6 for the front and back cameras, respectively. The described procedure robustly recovers 3D ball positions even if the ball silhouettes are only partially visible (Figure 6).

4.4 Reconstructing Initial Flight Parameters

After the 3D ball positions in the front and back part of a flight trajectory have been reconstructed, the initial flight parameters for that data set, i.e. velocity, rotation axis, and spin frequency, need to be determined, cf. Figures 1 (left) and 7 (left). Figure 8 gives a brief overview of the employed technique: from the reconstructed 3D marker positions, an initial guess for the flight parameters is extrapolated, which is then refined using the ball model from Figure 4 (right).

We identify the projected ball markers in the front stereo pair of images via color-based region detection and establish correspon-

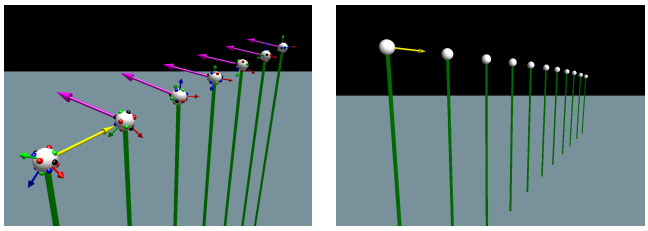

Figure 7: Reconstructed initial flight parameters (left) and flight positions (right). Distance of the balls from ground is shown in green, rotation axis in magenta, and initial velocity direction in yellow.

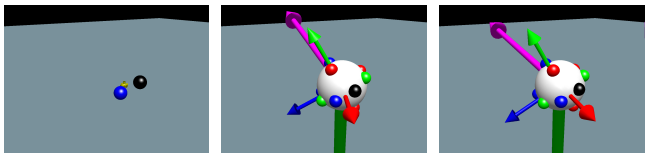

Figure 8: Stages of the fitting process. Left to right: position of markers, result from prediction, and result from final fitting.

dences across stereo images via the epipolarity constraint. The markers' 3D positions are found via triangulation and each marker is assigned to the closest ball position in 3D.

From the sequence of orientations of the ball's coordinate system immediately after release of the ball, its initial spin frequency and rotation axis are derived. In theory, it is sufficient to know the 3D positions of the ball's center and of two uniquely identified markers to determine its orientation. Unfortunately, it is impossible to decide from the color of a marker alone which one of the three instances of this marker type on the ball this is. In addition, misclassifications due to noise in the data needs to be considered. Physics tells us that the orientation of the rotation axis and the spin frequency of an ideal flying ball do not change over time (Sect. 4.5). Considering the above we determine the initial flight parameters by means of the following numerical optimization scheme.

The algorithm processes the n subsequent ball positions at the beginning of the trajectory separately and in their temporal order. The orientation of the ball at position k with respect to the world coordinate system can be represented as a rotation matrix $R(\alpha_k,\beta_k,\gamma_k)$, where $(\alpha_k,\beta_k,\gamma_k)$ are Euler angles. Our goal is to find for each subsequent pair of 3D ball positions at $k-1$ and k the rotation axis $\omega_{k-1,k}$ and rotation angle $\delta_{k-1,k}$ that correspond to the relative transformation $R_{k-1,k}$ between $R(\alpha_{k-1},\beta_{k-1},\gamma_{k-1})$ and $R(\alpha_k,\beta_k,\gamma_k)$.

At position k, the algorithm exploits temporal coherence by predicting the orientation of the ball $R(\alpha_{\text{pred}},\beta_{\text{pred}},\gamma_{\text{pred}})$ by rotating orientation $R(\alpha_{k-1},\beta_{k-1},\gamma_{k-1})$ further by $\delta_{k-2,k-1}$ around axis $\omega_{k-2,k-1}$. Starting from that parameter set $(\alpha_{\text{pred}},\beta_{\text{pred}},\gamma_{\text{pred}})$ the algorithm uses Powell's method [Press et al. 1992] to find parameters $(\alpha_k,\beta_k,\gamma_k)$ that minimize the energy function:

$$\begin{aligned} E_k(\alpha,\beta,\gamma) &= a_1E_1 + a_2E_2 \\ &= a_1\sum_{i\in M}(\Delta_m(i))^2 + a_2\sum_{j\in\{x,y,z\}}(\Delta_{ax}(j))^2\,, \quad (1) \end{aligned}$$

with a_1 and a_2 being weighting factors. M is the set of detected markers at ball position k, $\Delta_m(i)$ is the angular distance between reconstructed marker i and the closest marker of the same type in the ball model in the current orientation. $\Delta_{ax}(j)$ is the angular distance between the local coordinate axis $j\in\{x,y,z\}$ of the ball in orientation $(\alpha_k,\beta_k,\gamma_k)$ and the same axis in orientation $R(\alpha_{\text{pred}},\beta_{\text{pred}},\gamma_{\text{pred}})$.

The rotation axis $\omega_{k-1,k}$ and rotation angle $\delta_{k-1,k}$ are computed from the relative transformation $R_{k-1,k}$ between $R(\alpha_{k-1},\beta_{k-1},\gamma_{k-1})$ and $R(\alpha_k,\beta_k,\gamma_k)$ [Murray et al. 1994].

Having the sequence of rotation angles and the stroboscope frequency, the spin frequency f can be derived.

In our method we do not strictly enforce the constancy of the rotation axis and spin frequency, but instead introduce this criterion as a weighted regularization term E_2 to compensate for possible measurement errors and ball precession. For the initial rotation axis, we average the rotation axes over the sequence. The direction of the initial velocity vector coincides with the direction of the connecting line between the first two ball positions, its magnitude is computed from the strobe frequency and the Euclidean distance of the first two ball positions. For the first two ball positions the optimization is run with $a_2=0$ in Equation (1). If this initialization fails due to too few or badly located markers, a manual initialization is feasible.

In our experiments we were still able to recover valid initial flight parameters even if for some balls none or just one marker was found. We obtained almost 100 % probability of correct detection for the black markers and 90 % for the red markers. The blue and green markers were more difficult to find. In a comparative experiment it turned out that a different color scheme with more luminous marker colors significantly increases the robustness of marker detection.

4.5 Validation

For the ball flight data (3D positions and initial parameters), no ground truth information is available. To validate our acquisition setup and tracking algorithms, we show that the data obtained through our measurements and processing are consistent with the prediction of a physically based model that takes into account the dominating forces acting on a spinning ball traveling through air. In accordance to [Adair 2002] and [Alaways et al. 2001], we compute the velocity $\mathbf{v}(t)$ of a baseball with mass m using the first-order ordinary differential equation

$$m\dot{\mathbf{v}}(t) = \mathbf{F}_G + \mathbf{F}_D(\mathbf{v}(t)) + \mathbf{F}_M(\mathbf{v}(t)) \quad (2)$$

with the *gravitational force* $\mathbf{F}_G$, the *drag force* (or *air resistance*) $\mathbf{F}_D$, and the *Magnus force* $\mathbf{F}_M$ defined as:

$$\begin{aligned} \mathbf{F}_G &= m\cdot\mathbf{g}\,, \\ \mathbf{F}_D(\mathbf{v}(t)) &= -\frac{1}{2}\cdot C_D(\mathbf{v}(t))\cdot\rho\cdot A\cdot|\mathbf{v}(t)|^2\cdot\frac{\mathbf{v}(t)}{|\mathbf{v}(t)|}\,, \\ \mathbf{F}_M(\mathbf{v}(t)) &= \frac{1}{2}\cdot C_L(\mathbf{v}(t),\omega)\cdot\rho\cdot A\cdot|\mathbf{v}(t)|^2\cdot\frac{\omega\times\mathbf{v}(t)}{|\omega\times\mathbf{v}(t)|}\,, \end{aligned}$$

where $\mathbf{g}$ denotes gravity, ρ air density and A the cross-sectional area of the ball. The vector ω represents the spin axis of the ball, which is assumed to be constant during the flight of the ball[1]. To compute the *drag coefficient* $C_D(\mathbf{v}(t))$, we have fitted a polynomial curve to the data presented in [Adair 2002] and [Alaways et al. 2001]. After computing the Reynold's number $Re(\mathbf{v}(t))$ [Adair 2002] the drag coefficient is evaluated as

$$\begin{aligned} C_D(\mathbf{v}(t)) &= 2.23 - \\ &0.28342\cdot10^{-4}\cdot Re(\mathbf{v}(t)) + 0.13179\cdot10^{-9}\cdot Re(\mathbf{v}(t))^2 - \\ &0.25083\cdot10^{-15}\cdot Re(\mathbf{v}(t))^3 + 0.17083\cdot10^{-21}\cdot Re(\mathbf{v}(t))^4\,. \end{aligned}$$

[1] For a perfectly homogeneous ball, the spin axis does not change. In practice, a small precession might occur due to the inhomogeneous density of natural materials (cork, leather) used for baseballs.

pitch type	ε_{avg}	$\varepsilon_{\max}$	$\sphericalangle(\mathbf{v}_0^{\text{ref}}, \mathbf{v}_0)$	$\lvert\mathbf{v}_0^{\text{ref}}\rvert$	$\Delta(\lvert\mathbf{v}_0^{\text{ref}}\rvert, \lvert\mathbf{v}_0\rvert)$	$\sphericalangle(\omega^{\text{ref}}, \omega)$	$\lvert\omega^{\text{ref}}\rvert$	$\Delta(\lvert\omega^{\text{ref}}\rvert, \lvert\omega\rvert)$
fast ball (2 seams)	18 mm	39 mm	1.3°	63.2 mph	1.9 mph	0.4°	1596 rpm	22 rpm
fast ball (4 seams)	18 mm	41 mm	2.5°	64.2 mph	0.8 mph	0.1°	1612 rpm	17 rpm
curve ball	19 mm	39 mm	0.7°	61.9 mph	1.4 mph	0.3°	1623 rpm	7 rpm
slider	15 mm	25 mm	3.8°	65.7 mph	0.7 mph	0.4°	1491 rpm	13 rpm
change-up	13 mm	35 mm	1.4°	60.6 mph	1.1 mph	0.3°	1258 rpm	32 rpm

Table 1: Comparison of our measurements with reference trajectories obtained from a physically based model (Section 4.5). For a variety of pitches, the average error ε_{avg} and the maximum error $\varepsilon_{\max}$ between the reference trajectory and our measured ball positions are given (Euclidean distance between trajectory and center of ball). The precision of our measured initial flight parameters is specified by: $\sphericalangle(\mathbf{v}_0^{\text{ref}}, \mathbf{v}_0)$ (angle between reference and measured velocity direction), $\Delta(\lvert\mathbf{v}_0^{\text{ref}}\rvert, \lvert\mathbf{v}_0\rvert)$ (difference between reference and measured initial speed), $\sphericalangle(\omega^{\text{ref}}, \omega)$ (angle between reference and measured spin axis direction), and $\Delta(\lvert\omega^{\text{ref}}\rvert, \lvert\omega\rvert)$ (difference between reference and measured spin frequency). Absolute values of reference initial speed $\lvert\mathbf{v}_0^{\text{ref}}\rvert$ and spin frequency $\lvert\omega^{\text{ref}}\rvert$ are given for the sake of completeness.

According to [Alaways et al. 2001], the *lift coefficient* C_L can be computed as $C_L(\mathbf{v}(t), \omega) = 1.5 \cdot r \cdot |\omega| / |\mathbf{v}(t)|$. For the special case of a fastball across two or four seams, better approximations of C_L can be obtained from the diagrams in [Alaways et al. 2001]. Given the initial ball position $\mathbf{p}_0 = \mathbf{p}(0)$, the initial velocity $\mathbf{v}_0 = \mathbf{v}(0)$, as well as the initial spin axis ω and frequency $f = |\omega|$, the flying ball's position $\mathbf{p}(t)$ at time t is computed via integrating $\mathbf{v}(t)$ over time. Using the Runge-Kutta-Fehlberg integration scheme `DOPRI5` from [Hairer et al. 1993] we solve ODE (2) for $\mathbf{v}(t)$.

Finally, we can compute the reference trajectory of a baseball for a given set of initial flight parameters $\mathbf{p}_0$, $\mathbf{v}_0$, and ω and compare it to our measurements. Since the trajectory computed from the ODE (2) is quite sensitive with respect to variations in the initial flight parameters, we search for an exact solution of (2) that minimizes the error both for the measured ball positions and for the measured initial flight parameters using Powell's optimization method [Press et al. 1992]. The resulting *optimal reference trajectory* is then used to compute the measurement error (Table 1).

The comparatively low average speed of the pitches is due to the high number of pitches per recording session which exceeded the usual training pensum of a baseball professional by far.

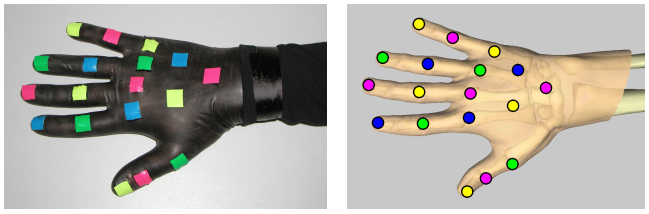

Figure 9: Left: Markers for tracking are attached to the pitcher's hand. Right: Corresponding marker positions on the personalized hand model.

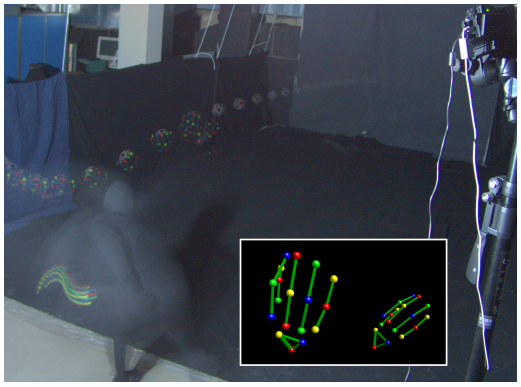

Figure 10: Multi-exposure image of one camera recording the hand motion during pitching. Inset: reconstructed hand marker positions for two hand poses.

5 Tracking the Hand

5.1 Preparation of the Pitcher's Hand

In order to determine the locations of the finger joints in the recorded images we have to mark them on the pitcher's hand. The pitcher wears a thin, transparent rubber glove onto which colored markers made of reflective tape are glued, see Figure 9 (left). The markers are placed on the joint positions, on the finger nails, and on three distinct positions on the back of the hand. Four different marker colors are distributed such that the distance between any two markers of the same color is maximized. In total 18 positions on the hand are tagged and assigned a unique position label. To facilitate identification of the markers in the multi-exposure images, the skin underneath the glove is painted with black make-up. During recordings the pitcher wears black clothes and a black face mask to prevent misclassifications of moving body parts other than the pitching hand.

5.2 Acquisition of Hand Motion

For acquisition of hand motion, all four cameras and one stroboscope are positioned in a semi-circular arrangement behind the pitcher, see Figure 3. In front of the pitcher, the walls and the floor of the flight corridor are covered with black cloth. All cameras are focused on the region where the pitcher releases the ball. The camera positions are chosen in such a way that two cameras observe the hand motion from the left and two from the right side of the pitcher's location. This way occlusions of the hand markers during the complex pitching movement are minimized and sufficiently separated exposures of the hand in the images could be obtained. The strobe light is located directly behind the pitcher such that the focus of illumination coincides with the release position of the ball. During recordings the stroboscope operates at 75 Hz, a frequency that leads to a high number of visible hand positions sufficiently separated in the images for all pitch types. For recording all four cameras are triggered synchronously with an exposure time of one second. We have recorded the same four pitches as for the trajectory measurements. Again, all pitches were performed as three-quarter deliveries.

5.3 Tracking the Hand Position

The first step in reconstructing the hand motion is to separate the marker positions from the background in each of the four multi-exposure images using background subtraction. Since all unimportant parts of the scene are colored black, the reflective markers emerge very brightly in the images, see Figure 10. The locations of the markers in each photograph can be identified with a similar technique as described in Section 4.4 for the ball markers. For each marker type a color interval is defined. Connected image regions above a minimum size whose pixels fall into one of the intervals are considered as projected marker locations. The projected centers of the markers are approximated as the centers of gravity of the marker regions. The correspondences over different camera views are established via epipolar geometry. Technically, the left and the right camera pair are treated as separate stereo pairs. In a first step the positions of visible markers are triangulated in each stereo pair separately. If a marker position is reconstructed from both stereo pairs, its position in 3D space is averaged.

Currently each of the 18 markers at each hand position is associated with the correct position label in an interactive procedure. An automatic approach that clusters 3D marker positions into separate hand clusters and assigns the marker labels in each cluster according to the colors of their neighboring markers is also feasible.

For motion reconstruction we limit ourselves to those hand positions in which the three markers on the back of the hand are visible for at least two cameras. Only then are the position and orientation of the hand root fully determined. Our setup is arranged such that this condition is fulfilled for an average of four hand positions around the release point. These hand positions are also the most interesting ones in terms of their motion characteristics since they represent that part of the motion cycle in which the hand and finger movements determine the specific rotation axes and spin frequencies of different pitch types. For some pitches it is not possible to reconstruct the position of all finger joints in each reconstructed hand position. This can happen for those pitch types where a finger is required to be ahead of the ball in the release moment such that it is occluded from all cameras.

5.4 Visualization of the Hand

For visualizing the movement of hand and fingers we use an animatable hand model. In particular, our hand model is composed of a skin mesh and the underlying bone structure, see Figure 9 (right). Animation of the hand model is controlled by joint rotation parameters specified over time. We employ a physics-based approach to compute the deformation of the skin tissue for a given configuration of the bones inside the hand. The skin mesh is identified with a mass-spring network with biphasic stiffness coefficients computed according to [Van Gelder 1998]. For the sake of brevity, we refer to the approach presented in [Albrecht et al. 2003] for a detailed description of this physically based animation technique.

To animate our hand model we have to make sure that it matches the pitcher's hand in size and proportions. To this end we apply a radial basis warping function as described in [Albrecht et al. 2003] to create a "personalized" hand model that matches the size and proportions of the pitcher's hand. The warped model is then equipped with markers at the same positions as on the glove, cf. Figure 9 (right).

Finally, the personalized hand model is animated using joint rotation parameters that have been computed automatically from the marker positions obtained from the tracking process. This conversion from marker positions to joint rotations proceeds as follows. First, we compute the position and orientation of the back of the hand by aligning the three markers on the back of the (personalized) hand model to the corresponding tracked marker positions. Next, we traverse the (anatomical) hierarchy of the hand model along each finger. For each joint, we compute the rotation angle that minimizes the distance between the position of the next marker along the hand model's hierarchy and its corresponding tracked marker. After traversing each finger up to its tip, all joint rotations are specified. We use key frame interpolation for the joint rotation parameters to compute smooth animations.

6 Results

For validation of our acquisition setup and tracking algorithms, we have performed the consistency check described in Section 4.5. As a result, we conclude that our measurements are very accurate. Average errors between the measured 3D ball position and the predicted flight trajectory are as low as 13–19 mm, which corresponds to about 18–25 % of the diameter of the baseball.

The calibration error for the camera setup was on average below one pixel in the image plane. This assures that a high-accuracy 3D reconstruction for the ball and the hand markers is feasible.

For the ball, the average distance between a measured feature in the image plane and its reprojected 3D location is below two pixels. The reprojection error for the center of the ball is about one pixel. Part of the deviation between measured and predicted ball positions might result from small inaccuracies in feature localization in the image plane.

Due to the lack of ground truth data for the hand motion we cannot assess the reconstructed hand motion data directly. The reprojection errors of the reconstructed marker positions of the hand are similarly small as those obtained for the ball measurements.

The high-quality data we acquired from different baseball pitches permit new ways of visualization that provide interesting feedback to the athlete, the coach, and the sports enthusiast. The flight of the baseball can be visualized from any camera perspective, see Figure 11. In particular, the ball's initial flight parameters and their relation to the flight trajectory can be rendered into instructive sequences as is demonstrated in the accompanying movie. Visualizing the hand motion during release of the ball in slow motion provides a new type of visual feedback for the performing pitcher. Figure 12 depicts two snapshots of such an animation.

The multi-exposure images acquired for tracking the hand motion show both the hand poses and the ball markers. We have thus reconstructed hand motion and flight parameters from the same set of stroboscope photographs. This way it is possible to visualize the influence of finger motion on the flight parameters of the ball. In Figure 1, the characteristic finger motion applied to add the necessary spin to a slider is clearly visible. In particular, the middle finger exerts high pressure on the ball to build up high spin. Due to the acceleration of the middle finger during the pitch, this finger moves further than the other fingers after release of the ball. The rotation of the ball in Figure 1 is consistent with the movement of the fingers.

7 Conclusion and Future Work

We have introduced a setup to capture high-speed, large scale motion via stroboscope photography using off-the-shelf digital still cameras. A method for automatic reconstruction of the 3D positions and initial flight parameters of a baseball from multi-exposure images is described and validated. It is shown that from the same type of images it is also possible to reconstruct complex and fast articulated hand motion.

Our system provides comprehensive and precise measurements of both pitching motion and flight parameters for a variety of baseball pitches. In combination with our visualization techniques, these measurements lead to a better understanding of the characteristics of baseball pitches and resulting flight trajectories. Thus,

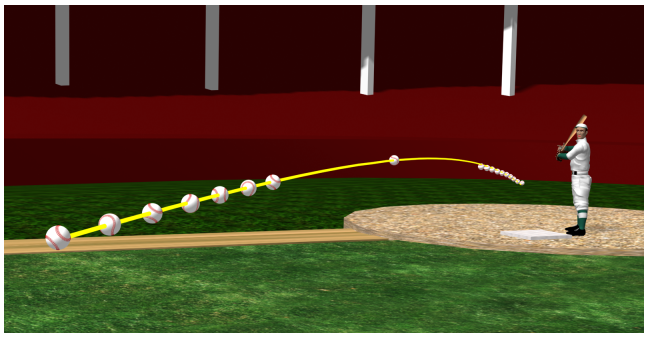

Figure 11: Visualization of a change-up trajectory in a stadium. The yellow path shows the reference trajectory obtained from the physical model. The average offset of the measured ball positions to this reference path is as low as 13 mm.

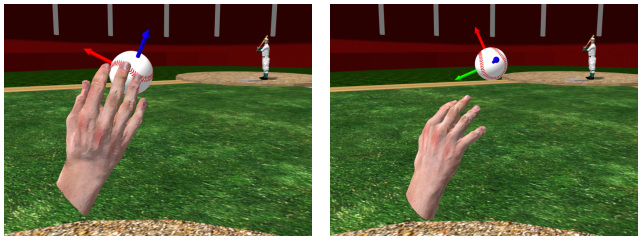

Figure 12: Visualization of hand and fingers during and after release of the ball. In this change-up pitch, the ball is spinning backwards about a rotation axis orthogonal to the flight direction. This can be seen by comparing the direction of the axes of the ball's local coordinate frame.

the system provides an instructive tool for pitchers and coaches, enabling them to improve their pitching technique through precise visual feedback. The demand of such visual feedback was confirmed during our recordings, when we experienced that the athlete's personal estimate of his performance sometimes deviates from the measured data.

We aimed at a high accuracy system that can be used to analyze high-speed motion on a limited spatial and temporal scale. We don't see our approach as a replacement for traditional motion capture techniques, but as a cost effective supplement which can be used in cases where traditional techniques fail. The baseball pitch is just an example for the application scenarios we have in mind. Other possible scenarios are tennis serves or the athlete's motion in several track and field events such as javelin or discus. Tennis players, for instance, would benefit from a precise analysis of the correlation between the movement of their racket, speed and spin of the ball, and the resulting ball trajectory during a serve.

In the future we plan to extend our framework to enable capturing human full-body motion by means of stroboscope photography.

8 Acknowledgements

The authors would like to thank their pitcher Thorsten Dehm from the *Saarlouis Hornets* for his patience while performing dozens of pitches and for his enthusiasm in our work. Furthermore, we are grateful to Marcus Weber and our housekeepers Axel Köppel and Michael Laise who helped us a lot during setting up the acquisition equipment. To Kristina Scherbaum we are indebted for lending a hand with the nastier aspects of visualization.

References

ADAIR, R. K. 2002. *The Physics of Baseball*, 3rd ed. HarperCollins, New York, NY.

AGGARWAL, J. K., AND CAI, Q. 1999. Human motion analysis: A review. *CVIU 73*, 3, 428–440.

ALAWAYS, L. W., MISH, S. P., AND HUBBARD, M. 2001. Identification of Release Conditions and Aerodynamic Forces in Pitched-Baseball Trajectories. *Journal of Applied Biomechanics 17*, 63–76.

ALAWAYS, L. W. 1998. *Aerodynamics of the Curve-Ball: An Investigation of the Effects of Angular Velocity on Baseball Trajectories*. PhD thesis, University of California, Davis, Davis, CA.

ALBRECHT, I., HABER, J., AND SEIDEL, H.-P. 2003. Construction and Animation of Anatomically Based Human Hand Models. In *Proc. Symposium on Computer Animation (SCA '03)*, 98–108.

ATHITSOS, V., AND SCLAROFF, S. 2003. Estimating 3d hand pose from a cluttered image. In *Proc. of CVPR'03*, vol. 2, 432.

BALLARD, D. 1981. Generalizing the hough transform to detect arbitrary shapes. *Pattern Recognition 13(2)*, 111–122.

COLLINS, D., AND BRUCE, R. R., Eds. 1994. *Seeing the Unseen: Dr. Harold E. Edgerton and the Wonders of Strobe Alley*. MIT Press.

D'ORAZIO, T., ANCONA, N., CICIRELLI, G., AND NITTI, M. 2002. A ball detection algorithm for real soccer image sequences. In *Proc. of ICPR'02*, vol. 1, 10210.

DORNER, B. 1993. Hand shape identification and tracking for sign language interpretation. In *IJCAI Workshop on Looking at People*.

FAUGERAS, O. 1993. *Three-dimensional computer vision : a geometric viewpoint*. MIT Press.

GLEICHER, M., FERRIER, N., GARDNER, A., SHIN, S., TOLLES, T., AND WILSON, T. 2001. Making motion capture useful. In *SIGGRAPH Course Notes*.

GUEZIEC, A. 2002. Tracking pitches for broadcast television. *IEEE Computer 35(3)*, 38–43.

GUEZIEC, A. 2003. Tracking a baseball for broadcast television. In *SIGGRAPH Course Notes*.

HAIRER, E., NØRSETT, S. P., AND WANNER, G. 1993. *Solving Ordinary Differential Equations I: Nonstiff Problems*, 2nd ed. Springer–Verlag, New York.

HEAP, T., AND HOGG, D. 1996. Towards 3D hand tracking using a deformable model. In *Proc. of 2nd Intl. Conf. on Automatic Face and Gesture Recognition*, 140.

HEIKKILA, J., AND SILVEN, O. 1996. Calibration procedure for short focal length off-the-shelf ccd cameras. In *Proc. of 13th International Conference on Pattern Recognition.*, 166–170.

HOUSE, T. 2000. *The Pitching Edge*, 2nd ed. Human Kinetics, Champaign, IL.

INTEL, 2002. Open source computer vision library. Available from `http://www.sourceforge.net/projects/opencvlibrary`.

JAIN, R., KASTURI, R., AND SCHUNCK, B. 1995. *Machine Vision*. McGraw Hill International.

MURRAY, R. M., LI, Z., AND SASTRY, S. S. 1994. *A mathematical introduction to robotic manipulation*. CRC Press.

MUYBRIDGE, E. 1887. *Animal Locomotion: An Electro-Photographic Investigation of Consecutive Phases of Animal Movements 1872–1885*. University of Pennsylvania, Philadelphia, PA.

PINGALI, G., JEAN, Y., AND CARLBOM, I. 1998. Real time tracking for enhanced tennis broadcasts. In *Proc. of CVPR*, 260–265.

PINGALI, G., YVES, J., OPALACH, A., AND CARLBOM, I. 2000. Lucentvision: Converting real world events into multimedia experiences. In *Proc. of Intl Conf. on Multimedia and Expo (ICME)*, 1433–1436.

PRESS, W. H., TEUKOLSKY, S. A., VETTERLING, W. T., AND FLANNERY, B. P. 1992. *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. Cambridge University Press, Cambridge, MA.

REHG, J. M., AND KANADE, T. 1994. Visual tracking of high DOF articulated structures: an application to human hand tracking. In *ECCV (2)*, 35–46.

SLANSKY, J. 1970. Recognition of convex blobs. *Pattern Recognition 2*, 3–10.

STENGER, B., MENDONÇA, P. R. S., AND CIPOLLA, R. 2001. Model based 3D tracking of an articulated hand. In *Proc. of CVPR*, vol. II.

STEWART, J. 2002. *The Pitching Clinic*. Burford Books, Inc., Short Hills, NJ.

VAN GELDER, A. 1998. Approximate Simulation of Elastic Membranes by Triangulated Spring Meshes. *Journal of Graphics Tools 3*, 2, 21–41.

WU, Y., LIN, J., AND HUANG, T. 2001. Capturing natural hand articulation. In *Proc. of ICCV*, 426–432.