

SINGLE-IMAGE 3D RECONSTRUCTION OF BALL VELOCITY AND SPIN FROM MOTION BLUR

An Experiment in Motion-from-Blur

Giacomo Boracchi, Vincenzo Caglioti and Alessandro Giusti
*Dipartimento di Elettronica e Informazione,
Politecnico di Milano, Via Ponzio, 34/5 20133 Milano, Italy*

Keywords: Motion Blur, 3D Motion Reconstruction, Single Image Analysis, Blur Analysis.

Abstract: We present an algorithm for analyzing a single calibrated image of a ball and for reconstructing its instantaneous motion (3D velocity and spin) by exploiting motion blur. We use several state-of-the-art image processing techniques for extracting information from the space-variant blurred image, then robustly integrate such information in a geometrical model of the 3D motion. We initially handle the simpler case in which the ball apparent translation is negligible w.r.t. its spin, then extend the technique to handle the most general motion. We show extensive experimental results both on synthetic and camera images. In a broader scenario, we exploit this specific problem for discussing motivations, advantages and limits of reconstructing motion from motion blur.

1 INTRODUCTION

In this paper we propose a technique for estimating the motion of a ball from a single motion blurred image. We consider the instantaneous ball motion, which can be described as the composition of 3D velocity and spin: the proposed technique estimates both these components by analyzing motion blur.

A more traditional and intuitive method consists in recovering motion by analyzing successive video frames: the expected shortcomings of such *modus operandi* in realistic operating conditions motivate our unusual approach. In fact, depending on equipment quality, lighting conditions and ball speed, a moving ball often results in a blurred image. Feature matching in successive video frames becomes very challenging because of motion blur and also because of repetitive features on the ball surface: this prevents *inter-frame* ball spin recovery. Then, it is worth considering *intra-frame* information carried by the motion blur. Our single-image approach has the further advantage of enabling the use of cheap, high-resolution consumer digital cameras, which currently provide a much higher resolution than much more expensive video cameras. High resolution images are vital for performing accurate measurements as the ball usually covers a small part of the image.

We use an alpha matting algorithm (see Section 3.1), as a preliminary step before applying a known technique for estimating the 3D position and

velocity of an uniformly colored ball from a single blurred image (Boracchi et al., 2007). This allows us to relax the uniform color assumption and consider textured balls on known background. Once the ball position and velocity are known, we analyze the blurred image of the ball textured surface: in particular, blur is characterized by smears which have varying direction and extent, resulting from the 3D motion of the ball surface. We estimate spin by analyzing such smears within small image patches, and by integrating them on a geometrical 3D model of the ball.

The blur model derived from the 3D ball motion is presented in Section 2, while in Section 3 we briefly recall the image analysis algorithms used. The proposed technique is described in Section 4. Section 5 presents experimental results and Section 6 summarizes the work and presents future research directions.

1.1 Related Works

Given a single blurred image, the most treated problem in literature is the estimation of the point spread function (PSF) that corrupted the image (Fergus et al., 2006; Levin, 2007; Jia, 2007), usually with the purpose of image restoration (*deblurring*).

Our work, on the contrary, takes advantage of motion blur for performing measurements on the imaged scene. Several other works follow a similar approach, such as (Klein and Drummond, 2005), which describes a visual gyroscope based on rotational blur

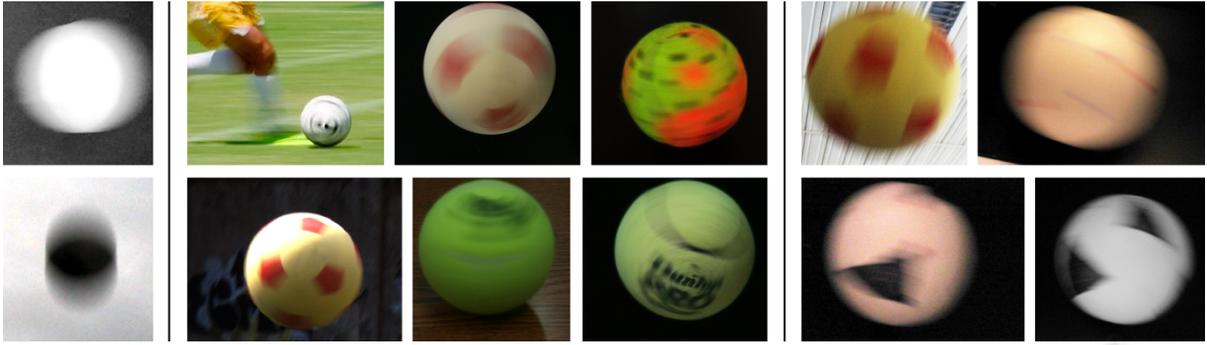


Figure 1: Some blurred ball images. Leftmost images are textureless, so their spin can not be recovered. Central images show textured balls whose spin component dominates the apparent translation. Rightmost images are the most complete case we handle, showing a significant amount of apparent translation and spin; note that the ball contours also appear blurred in this situation, whereas they are sharp in the spin-only case.

analysis, or (Levin et al., 2007), which estimates the scene depth map from an image acquired with a coded aperture camera. Also, (Rekleitis, 1996) proposes to estimate the optical flow from a single blurred image. A ball speed measurement method based on a blurred image has been proposed in (Lin and Chang, 2005). This assumes a simplified geometrical model that originates space-invariant blur and prevents the estimation of 3D motions and spin.

On the other hand, the problem of estimating the motion of a ball in the 3D space has been extensively treated in video tracking literature (Gopal Pingali and Jean, 2000; J Ren and Xu, 2004; Jonathan Rubin and Stevens, 2005). These methods assume the ball visible from multiple synchronized cameras, in order to triangulate the ball position in the corresponding frames. In (Reid and North, 1998) a method is proposed for reconstructing the ball 3D position and motion from a video sequence by analyzing its shadow. In (Kim et al., 1998; Ohno et al., 2000), a physics-based approach is adopted, to estimate the parameters of a parabolic trajectory.

Recently, some methods for estimating the 3D ball trajectory from image blur have been proposed (Caglioti and Giusti, 2006; Boracchi et al., 2007). However, these methods assume an uniformly-colored ball, and do not recover spin.

2 PROBLEM FORMULATION

Let S be a freely moving ball centered in C , whose radius R is known¹, imaged by a calibrated camera. We assume that during the exposure time T the ball

¹if the radius is not known, the whole reconstruction can be performed up to a scale factor

motion is defined by the composition of two factors:

- a linear translation with uniform velocity, \mathbf{u} . The translation distance during the exposure is therefore $T \cdot \mathbf{u}$.
- the spin around a rotation axis a passing through C , with angular speed ω . The rotation angle which occurs during the exposure is therefore $T \cdot \omega$.

From the ball localization technique (Boracchi et al., 2007) we inherit the constraint that the ball projections at the beginning and at the end of the exposure significantly overlap. Moreover, in order to recover the rotation axis and speed, we also require that spin is not too fast nor too slow w.r.t. the exposure time: $\pi/50 < \omega \cdot T < \pi/2$. In practice, these constraints allow us to use an exposure time 5 ÷ 10 times longer than the exposure time which would give a sharp image.

Our goal is to estimate the ball spin (both a and ω), velocity \mathbf{u} , and initial position by analyzing a single blurred image.

We assume that the blur on pixels depicting the ball is only due to ball motion. In practice, this can be achieved if the ball is in focus and the camera is still. Therefore the image formation model, on which our analysis is based, can be described as follows.

2.1 Blurred Image Formation

Let Z be the blurred image that depicts the moving ball and let $[0, T]$ be the exposure interval. The blurred image Z can be modeled as the integration of infinitely many (sharp) sub-images $I_t, t \in [0, T]$, each depicting the ball in a different 3D position and spin angle (see Figure 2):

$$Z(x) = \int_0^T I_t(x) dt + \eta(x), \quad x \in X. \quad (1)$$

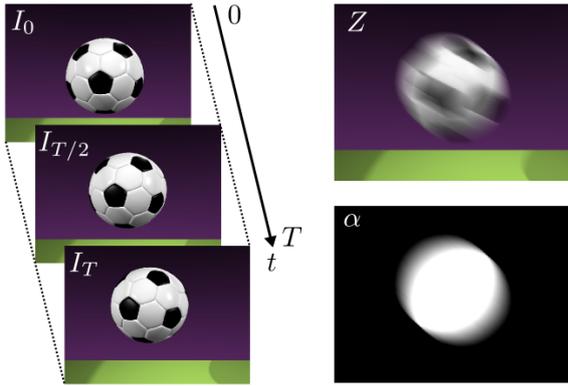


Figure 2: Blurred image formation model. The blurred image Z is obtained as the temporal integration of many still images I_t . The alpha map α of the blurred ball represents the motion of the object's contours and is used for recovering the translational motion component.

Where x represents the 2D image coordinates, $I_t(x)$ is the light intensity that reaches the pixel x at time t , and $\eta \sim N(0, \sigma^2)$ is white gaussian noise.

The ball apparent contours $\gamma_t, t \in [0, T]$ vary depending on translation only. Note that each apparent contour γ_t is an ellipse and that, in each sub-image I_t , γ_t may have a different position and also a different shape because of perspective effects (Boracchi et al., 2007). On the contrary, the spin does not affect $\gamma_t, t \in [0, T]$. In our reconstruction procedure, we will exploit the fact that the alpha map α of the blurred ball only depends on variations in γ_t , and is not affected by spin. The ball spin, combined with the translation, changes the depicted ball surface in each sub-image I_t and obviously the appearance of the ball in Z .

2.2 Blur on the Ball Surface

We treat the blur on the ball surface as locally space invariant (Bertero and Boccacci, 1998). In particular we approximate the blur in a small image region as the convolution of the sub-image I_0 with a PSF having vectorial support and constant value on it. Hence for any pixel x_i belonging to the ball image, we consider a neighborhood U_i of x_i and a PSF h_i such that

$$Z(x) = \int_X h_i(x-s)I_0(s)ds + \eta(x), \forall x \in U_i \quad (2)$$

The PSF h_i is identified by two parameters, the direction θ_i and the extent l_i .

3 IMAGE ANALYSIS

The proposed algorithm exploits two blur analysis techniques in order to separately handle the effects of

ball translation and ball spin. The ball position and velocity \mathbf{u} in 3D space are obtained by combining an alpha matting technique with the method presented in (Boracchi et al., 2007). The ball spin is computed by estimating the blur parameters within small regions on the ball image.

3.1 Alpha Matting

Alpha matting techniques have been recently applied to motion blurred images with different purposes, including PSF estimation (Jia, 2007) and blurred smear interpretation (Caglioti and Giusti, 2007). As shown in (Giusti and Caglioti,), by applying alpha matting to the motion-blurred image of an object we obtain a meaningful separation between the motion of the object's boundaries (alpha map) and the actual blurred image of the object (color map).

As we described in the previous section, in this scenario the alpha map of a blurred ball is not influenced by the spin but only by the translation: in practice, the alpha map is the image we would obtain if the background was black and the ball had a uniformly-white projection. Therefore the alpha map of the blurred ball can be used to estimate the ball position and displacement vector $T \cdot \mathbf{u}$ according to the technique presented in (Boracchi et al., 2007), even when the ball surface is textured.

On the contrary, the color map only shows the blurred ball image, as if it was captured over a black background. In the following sections, the color map will be analyzed in order to recover the ball spin.

In the general case, the matting problem is under-constrained, even if the background is known. Still, in literature many algorithms have been proposed: some of them (Smith and Blinn, 1996; Mishima, 1993) require a specific background (*blue screen matting*), whereas others, with minimal user assistance, handle unknown backgrounds (*natural image matting*) and large zones of mixed pixels ($0 < \alpha < 1$). Although none is explicitly designed for the interpretation of motion blurred images, we can get satisfactory results in our peculiar setting. In this context, however, we adopted (Giusti and Caglioti,), a fast and exact alpha matting algorithm, very suited to sport environments where its requirements on the background and foreground colors are often met.

3.2 Blur Analysis

As mentioned in Section 2.2, we approximate the blur as locally shift invariant, produced by a convolution with a PSF having vector-like support. We estimate the blur direction and extent separately on N image

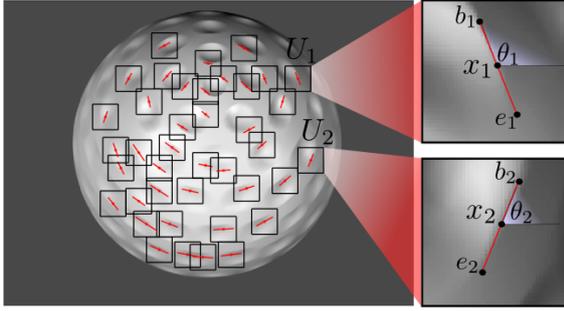


Figure 3: A synthetic image of a spinning golf ball. U_i neighborhoods and recovered blur directions and extents are shown. Each segment $b_i e_i$ represents the blur parameters θ_i , l_i within the region.

regions U_i $i = 1, \dots, N$ containing pixels which have been covered by the ball projection during the entire exposure time, i.e. $\alpha(x) = 1 \quad \forall x \in U_i, i = 1, \dots, N$.

In particular we apply the method proposed in (Yitzhaky and Kopeika, 1996) and we estimate the blur direction as the direction having minimum derivative energy. This is motivated by the fact that most of image details along the blur direction are smoothed by blur. After estimating the blur direction, the blur extent is obtained from the distance between two negative peaks in the autocorrelation of directional derivatives along the blur direction. Figure 3 shows some square regions used for blur analysis.

Other techniques for estimating the local blur directions may be used: for example, when the ball texture contains corners (like in most football balls), the method presented in (Boracchi and Caglioti, 2007) can be applied. Alternatively, blurred smears of strong features can be highlighted by applying the filtering techniques in (Caglioti and Giusti, 2007).

4 RECONSTRUCTION TECHNIQUE

For clarity purposes we illustrate the proposed technique first in the simpler case, where blur is due to ball spin only. Then, in Section 4.2.1 we cope with the most general case where the ball simultaneously translates and spins.

4.1 Null Translation

Let assume that during the exposure the ball does not translate, i.e. $\mathbf{u} = 0$, so that in the blurred image the ball apparent contour γ is an ellipse and it allows us to localize the ball in the 3D space by means of the camera calibration

parameters and knowledge of the ball radius. Points belonging to γ are easily found in the image either by ordinary background subtraction or using edge points in the alpha matte. We extract γ by fitting an ellipse to such points, enforcing the projective constraint of being the image of a sphere captured from the calibrated camera.

Then, the blur is analyzed within N regions U_i , $i = 1, \dots, N$ contained inside γ . In order to avoid uniform-color areas, we select such regions around local maxima $x_i, i = 1, \dots, N$ of the Harris corner measure (Harris and Stephens, 1988). For each of these points, a local blur direction θ_i is obtained using method presented in (Yitzhaky and Kopeika, 1996).

Such directions are now exploited in order to recover the 3D motions \mathbf{v}_i of the ball surface at points corresponding to each of the regions. Since the camera is calibrated and we know the 3D position of the sphere S , we can backproject each pixel x_i on the sphere surface. Let X_i be the intersection point, closest to the camera, between the viewing ray of x_i and sphere S : the 3D motion direction of the ball surface at X_i is described by an unit vector \mathbf{v}_i (see Figure 4 left). More precisely, let π_i be the plane tangent to S at X_i : then, \mathbf{v}_i is found as the direction of the intersection between π_i and the viewing plane of the image line passing through x_i and having direction θ_i .

As shown in Figure 4 (left), all the vectors \mathbf{v}_i $i = 1, \dots, N$ must lie on the same plane, orthogonal to the rotation axis a . Then, let $W = [\mathbf{v}_1 | \mathbf{v}_2 | \dots | \mathbf{v}_N]$, be the matrix having vectors \mathbf{v}_i as columns. The direction of a is found as the direction of the eigenvector associated to the smallest of W 's eigenvalues. This estimate is refined by iterating the procedure after removing the \mathbf{v}_i vectors that deviate too much from the plane orthogonal to a (outliers).

Note that, when the ball is not translating, the ball apparent contour γ is sharp and in this case it is easily localized by fitting an ellipse to image edge points (possibly after background subtraction) or by using a generalized Hough transform, without need of alpha matting.

Although the rotation axis can be recovered exploiting θ_i directions only, in order to estimate the angular speed we need to consider also the blur length l_i estimated within regions U_i . Each of these extents represents the length of the trajectory (assumed rectilinear) that the feature traveled in the image during the exposure. For each feature, a starting point b_i and ending point e_i are determined in the image as

$$b_i = x_i - \frac{l}{2} \cdot \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} \quad e_i = x_i + \frac{l}{2} \cdot \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} \quad (3)$$

and backprojected on the sphere surface S to points B_i and E_i , respectively. We then compute the dihe-

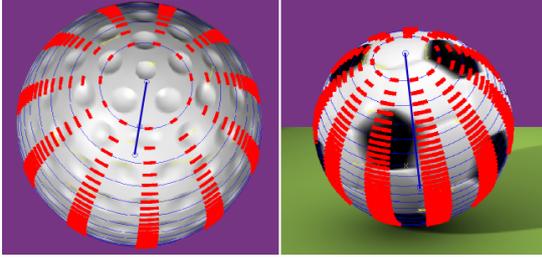


Figure 5: Reconstruction results on two synthetic images (spin only).

Table 1: Mean relative error in ω estimation, expressed as a percentage w.r.t the true value of ω . Columns where $\sigma > 0$ shows the average over ten noise realizations. Image data is in the $0 \div 255$ range.

$\omega \cdot T \setminus \sigma$	0	1	2	3
5.00	4.31	4.6222	5.0641	3.9401
6.25	2.26	2.5562	4.7898	4.3915
7.50	2.40	3.1353	2.7236	2.0544
8.75	0.75	1.5163	2.9408	5.0431
10.00	2.15	3.3975	5.3916	11.3800

5 EXPERIMENTS

We validated our technique on both synthetic and camera images.

Each synthetic image has been generated according to (1), by using the Blender 3D modeler for rendering hundreds of sharp frames each depicting the moving ball at a different t belonging to the exposure interval $[0, T]$. Each frame corresponds to a sub-image I_t , and all these sub-images are averaged together. We generated images with varying, known spin amount $\omega \cdot T$ in the $1^\circ \div 20^\circ$ range, both in the spin-only and in the spin plus translation cases. Several scenarios (some are shown in Figures 5 and 7) have been rendered with different spin axes w.r.t to the camera. Some of our test cases use a plain texture for the ball, whereas others feature a realistic ball surface with 3D details such as bumps and seams, and specular shading; this simulates difficult operating conditions, in which the ball appearance under motion is not easily defined.

Table 1 shows algorithm performance for ω estimation in an 800×600 image also accounting for noise. Figure 7 shows some of the synthetic images we used. The algorithm accuracy is reduced when the ball spin is too low as the small ball resolution does not allow reliable estimates in small regions U_i . Considered regions are disk shaped, having radius varying between 30-45 pixels according to the noise standard deviation σ , estimated using (Donoho and Johnstone,

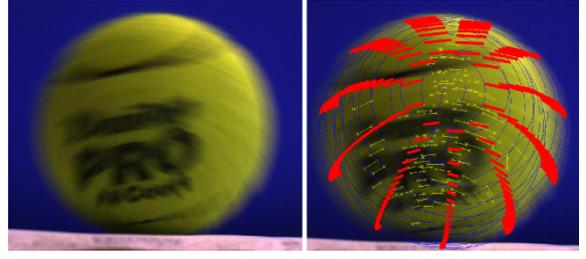


Figure 6: A real image (tennis ball) spinning and translating, and reconstructed motion (right). Note complex motion of points on the ball surface due to simultaneous spin and translation: red stripes show reconstructed motion, and correctly interpret the observed blur. Since the ball was rolling on a table (bottom of the image), features on the bottom of the ball are correctly estimated as still, and the rotation axis as coplanar with the table.

1994).

Both in synthetic and camera images, the blur estimates show a variable percentage of outliers ($5\% \div 50\%$), which are correctly discarded in most cases. Outliers are more frequent in noisy images, with smaller spin amounts and where the ball texture shows strong, straight edges.

We found that, in general, the estimation of blur extents l_i is much more error-prone than the estimation of blur angles θ_i , without significant differences between real and synthetic images. Together with the orientation problem (see Section 4.2.1), this makes the analysis of the general motion case much more challenging than the spin-only case (where the extents are only used to estimate $\omega \cdot T$).

6 DISCUSSION, CONCLUSIONS AND ONGOING WORKS

We propose a technique for reconstructing the velocity and spin of a moving ball from a single motion-blurred image, highlighting advantages and disadvantages of the approach. Our tests show very promising results, both in synthetic and real images, especially if the ball's apparent translation during the exposure is negligible (i.e. the ball contours are sharp). This scenario, which is not unusual in practice, has several important practical advantages such as no need of alpha matting, limited reliance on the estimation of blur lengths – which we have found to be quite unreliable – and irrelevance of the orientation problem.

In a broader view, our technique solves a nontrivial motion estimation problem from the motion blur in a single image. Although unusual, this approach may result successful in situations where traditional

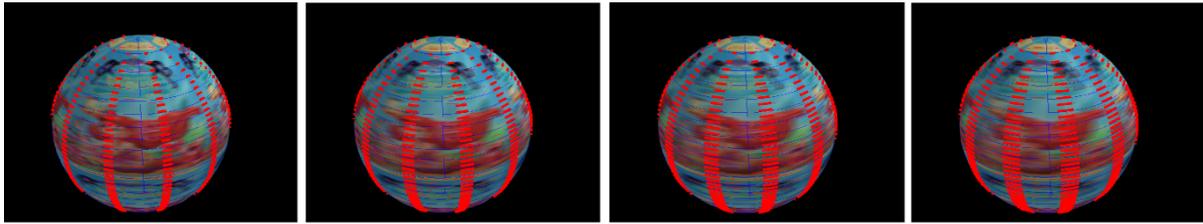


Figure 7: Synthetic image of textured ball at different ω values. From left to right $\omega \cdot T = 5^\circ; 6.875^\circ; 8.125^\circ; 10^\circ$.

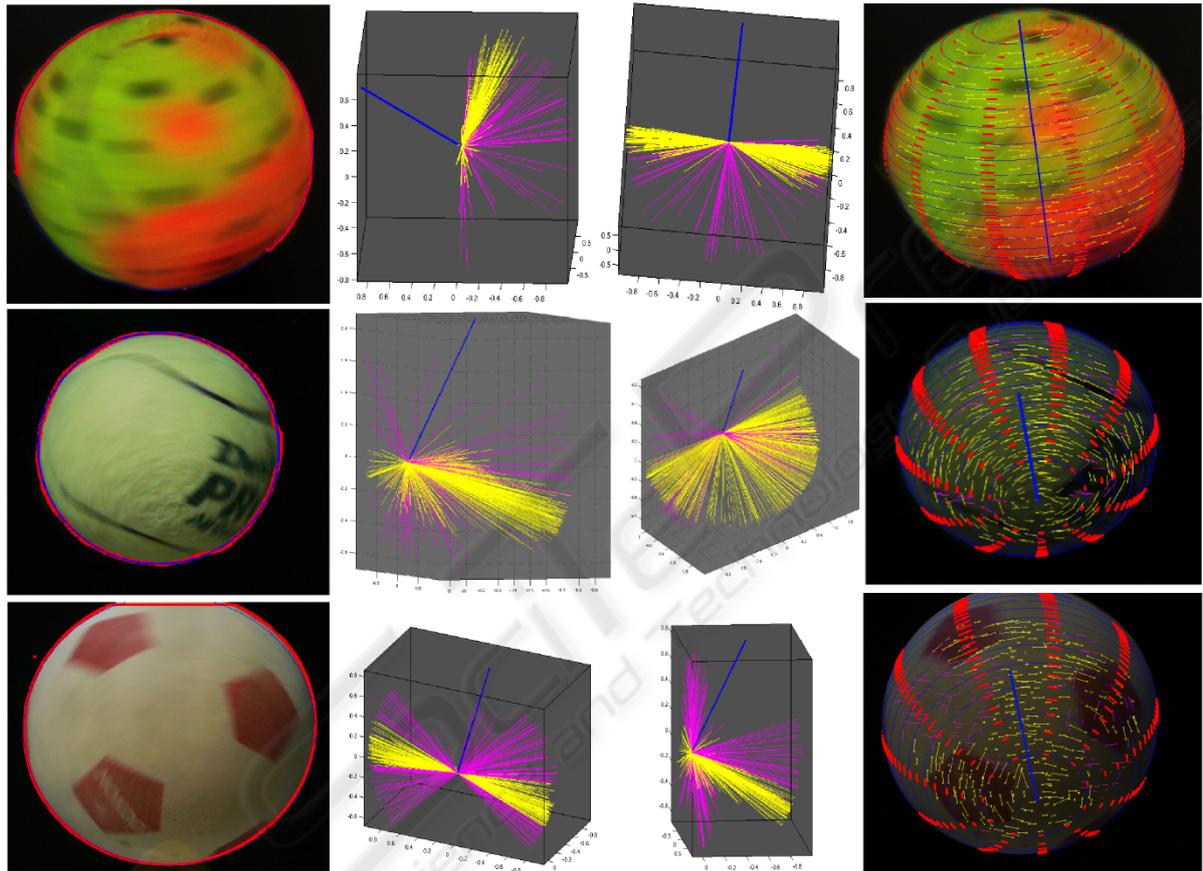


Figure 8: Real images (spin only). Central columns shows axis (blue) and v_i directions from different viewpoints: yellow ones are inliers, magenta are outliers. Corresponding blur estimates are shown in the rightmost image as yellow segments. Reconstructed spin axes and speeds correctly explain the blurred image: for example, the spin axis passes through the sharpest parts of the ball image.

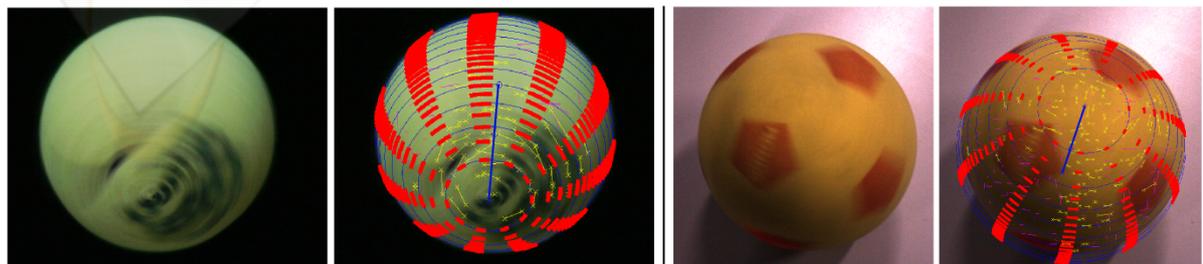


Figure 9: Other real images, and reconstructed motions (spin only). Display colors are the same as in previous figures.

video-based methods fail; target applications include training support and match analysis in sport environments.

Other than further improving our software toward a more robust implementation, we are currently investigating the practical possibility of estimating spin axis and velocity without the need of blur extents. Moreover, we are testing other techniques for blur estimation and alpha matting.

REFERENCES

- Bertero, M. and Boccacci, P. (1998). *Introduction to Inverse Problems in Imaging*. Institute of Physics Publishing.
- Boracchi, G. and Caglioti, V. (2007). Corner displacement from motion blur. In *proc. of ICIAP 2007 Conference*.
- Boracchi, G., Caglioti, V., and Giusti, A. (2007). Ball position and motion reconstruction from blur in a single perspective image. In *Inproceedings of ICIAP 2007, Modena*.
- Caglioti, V. and Giusti, A. (2006). Ball trajectory reconstruction from a single long-exposure perspective image. In *Proc. of Workshop on Computer Vision Based Analysis in Sport Environments (CVBASE)*.
- Caglioti, V. and Giusti, A. (2007). On the apparent transparency of a motion blurred object. In *Proc. of ICCV workshop on Photometric Analysis in Computer Vision (PACV) 2007*.
- Donoho, D. L. and Johnstone, I. M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455.
- Fergus, R., Singh, B., Hertzmann, A., Roweis, S. T., and Freeman, W. T. (2006). Removing camera shake from a single photograph. In *ACM SIGGRAPH 2006 Papers*.
- Giusti, A. and Caglioti, V. Isolating motion and color in a motion blurred image. In *Proc. of BMVC 2007*.
- Gopal Pingali, A. O. and Jean, Y. (2000). Ball tracking and virtual replays for innovative tennis broadcasts. In *proc. of ICPR 2000 Conference*, page 4152, Washington, DC, USA. IEEE Computer Society.
- Harris, C. and Stephens, M. (1988). A combined corner and edge detector. In *Proceedings of the 4th Alvey Vision Conference*, pages 147–151.
- J Ren, J. Orwell, G. J. and Xu, M. (2004). A general framework for 3d soccer ball estimation and tracking. In *Proc. of ICIP 2004 Conference*.
- Jia, J. (2007). Single image motion deblurring using transparency. In *Inproceedings of CVPR 2007, Minneapolis*.
- Jonathan Rubin, Burkhard C. Wuensche, L. C. and Stevens, C. (2005). Computer vision for low cost 3-d golf ball and club tracking. In *Proc. of Image and Vision Computing New Zealand*.
- Kim, T., Seo, Y., and Hong, K.-S. (1998). Physics-based 3d position analysis of a soccer ball from monocular image sequences. In *proc of ICCV 1998 Conference*, pages 721–726.
- Klein, G. and Drummond, T. (2005). A single-frame visual gyroscope. In *Proc. British Machine Vision Conference (BMVC'05)*, volume 2, pages 529–538, Oxford. BMVA.
- Levin, A. (2007). Blind motion deblurring using image statistics. In Schölkopf, B., Platt, J., and Hoffman, T., editors, *Advances in Neural Information Processing Systems 19*. MIT Press, Cambridge, MA.
- Levin, A., Fergus, R., Durand, F., and Freeman, W. T. (2007). Image and depth from a conventional camera with a coded aperture. *ACM Trans. Graph.*, 26(3):70.
- Lin, H.-Y. and Chang, C.-H. (2005). Automatic speed measurements of spherical objects using an off-the-shelf digital camera. In *Mechatronics, 2005. ICM '05. IEEE International Conference on*, pages 66–71.
- Mishima, Y. (1993). Soft edge chroma-key generation based upon hexoctahedral color space. U.S. Patent 5,355,174.
- Ohno, Y., Miura, J., and Shirai, Y. (2000). Tracking players and estimation of the 3d position of a ball in soccer games. In *ICPR*, pages 1145–1148.
- Reid, I. D. and North, A. (1998). 3d trajectories from a single viewpoint using shadows. In *proc. of BMVC 1998 Conference*.
- Rekleitis, I. M. (1996). Steerable filters and cepstral analysis for optical flow calculation from a single blurred image. In *Vision Interface*, pages 159–166, Toronto.
- Smith, A. R. and Blinn, J. F. (1996). Blue screen matting. In *SIGGRAPH '96: Proc. of the 23rd annual conference on Computer graphics and interactive techniques*, pages 259–268.
- Yitzhaky, Y. and Kopeika, N. S. (1996). Identification of blur parameters from motion-blurred images. In *Proc. SPIE Vol. 2847*, pages 270–280.