BallCam! Dynamic View Synthesis from Spinning Cameras

Kris M. Kitani
Carnegie Mellon University
5000 Forbes Ave, Pittsburgh,
PA 15213 USA
kkitani@cs.cmu.edu

Kodai Horita
UEC Tokyo
1-5-1 Chofugaoka, Chofu,
Tokyo, JAPAN
horita@vogue.is.uec.ac.jp

Koike Hideki
UEC Tokyo
1-5-1 Chofugaoka, Chofu,
Tokyo, JAPAN
koike@vogue.is.uec.ac.jp

ABSTRACT
We are interested in generating novel video sequences from a ball's point-of-view to enhance an spectators experience in sports domains. Despite the challenge of extreme camera motion, we show that we can leverage the periodicity of spinning cameras to generate a stabilized ball point-of-view video. We present preliminary results of image stabilization and view synthesis from a single camera being hurled in the air at 600 RPM.

Author Keywords
BallCam!, View synthesis, Image stitching.

ACM Classification Keywords
H.5.1 Information Interfaces and Presentation: Artificial, augmented, and virtual realities.

General Terms
Algorithms; Design; Human Factors.

INTRODUCTION
The current spectator paradigm for ball-based sports (e.g., volleyball, basketball, football) revolves around the third-person point-of-view (POV). Typically cameras (both static and dynamic) are placed strategically around the playing field to capture player motion. In our work, we explore the possibility of augmenting the current spectator viewing paradigm with a dynamic ball's POV. What if you could experience the flight of a ball as it spins towards a receiver? What if you could watch players scramble to grab an airborne ball, from the ball's perspective?

Recent advances in portable camera technology indicates that cameras are becoming more robust to dynamic motion (e.g. sky diving, slope-style biking) and durable in extreme environments (e.g. motorcross, F1). We believe that these advances in rugged, high-quality, low-cost cameras is creating new opportunities to enhance spectator sports through the use of first-person video media.

We are currently working on a prototype ball-camera system, which we call the 'BallCam!'. It is composed of a wearable camera embedded in an American football (Figure 1). Our preliminary results show that despite the extreme camera motion that is induced by throwing the camera in the air, we are still able to generate a stable ball's POV video sequence, that mimics a bird's eye POV of the playing field. In contrast to the motionSphere [1] or the throwable panoramic ball camera [2], our system has only one camera and is able to deal with extremely fast camera motion (600 RPM).

Figure 1. Our prototype BallCam! system

GENERATING A BALL'S POV VIDEO
Due to the extreme spinning motion of the camera during flight, it is very challenging to apply standard motion stabilization algorithms directly to the image sequence. In particular, there are large displacements between frames which make it difficult to accurately align frames using feature matching or optical flow. The alignment problem is further compounded by the introduction of both motion blur and the rolling shutter effect, which makes the application of precise structure-from-motion algorithms extremely challenging. To overcome these challenges we leverage the periodic rotational motion of ball-shaped projectiles and propose an approach that utilizes the mean image intensity as a rough estimate of the ball's rotation angle. Furthermore, we show how simple image compositing techniques and motion interpolation can be used to generate visually plausible bird's eye POV videos from a ball in flight (Figure 2).

Estimating Rough Camera Rotation
The first step in our approach requires that we establish a rough virtual camera motion path from a sequence of spinning camera images. As mentioned previously, since it is difficult to apply standard motion estimation techniques, we make two domain assumptions that allows us to roughly estimate the rotation of the camera: (1) the ball-camera system rotation axis is roughly orthogonal to the flat ground plane surface normal and (2) the sky is much brighter than the ground. In an American football scenario, (1) will hold

UIST'12, October 7–10, 2012, Cambridge, Massachusetts, USA.
ACM 978-1-4503-1582-1/12/10.

Figure 2. BallCam Results: (top) our approach, (bottom) naive sub-sampling

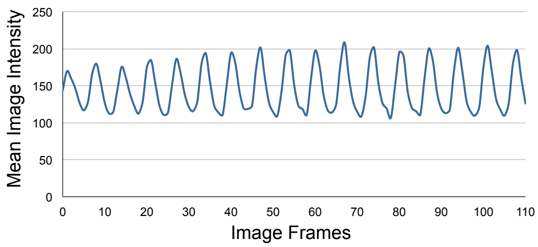

Figure 3. Plot of image intensity (vertical axis) over time (horizontal axis)

true for a ball throw with a clean spiral and (2) is true for most outdoor football fields.

In Figure 3, we show the mean intensity of each image frame over time. We observe that the mean intensity changes between bright intensities (camera is pointed to the sky) to dark intensities (camera is pointed down towards the field) with a very distinctive periodic cycle as the ball spins through the air. The image intensity feature captures only global motion and is very robust to localized motion or changes in the environment.

Virtual camera path

We generate a sequence of images from a stable virtual camera path (e.g. downward facing POV) by subsampling image frames that have the same orientation. We do this by subtracting a given intensity value s_0 from the mean intensity signal $s(t)$ and extract image frames near the zero-crossing points. This process yields a set of images that are roughly oriented in the same direction.

Motion Interpolation and Image Stitching

Since the images extracted over the virtual camera path are only roughly oriented in the same direction, we use linear motion interpolation and image stitching to smooth the video sequences. To synthesize novel views between image frames, we first compute the image transformation between images. We do this by computing the image homography using local salient feature matching combined with RANSAC-based homography estimation [3]. However, since each frame is quite distorted, we allow for significant projection error. Since we push the limits of the fundamental assumptions of homography estimation, naively blending images using these transformations causes images with significant noise (see Figure 4). To 'cover up' the artifacts introduced by the rough image alignment, we use a shortest path algorithm to find an optimal path through overlapping image regions. This approach has been used in computer graphics (e.g. texture synthesis [4], image seam carving [5]) to create seamless mosaics of images.

Figure 4. Comparing image stitching. Naive alpha blending (left), Seam carving (right)

Visualization

We show image frame results of our BallCam system in Figure 2. The motion between successive frames has been minimize. The resulting video has synthesized a plausible ball's POV perspective during it's time of flight and simulates a realistic downward looking view.

CONCLUSION

We have proposed a robust method for generating novel ball's point-of-view video sequences from a spinning camera. Our BallCam prototype has shown the potential for sensible image processing using dynamic POV video for ball-based sports domains. It is our hope that this work will act as an impetus for more dynamic viewing experiences for ball-based sports.

Acknowledgements

We thank Serge Belongie, Eric Christiansen and the UCSD Pixel Cafe for useful feedback and support.

REFERENCES

1. H. Mori, M. Inami, F. Matsuno, R. Miyauchi, H. Nii, M. Sugimoto and S. Kuwashima, MotionSPHERE. In *Proceedings of SIGGRAPH Emerging technologies*, 2005.
2. J. Pfeil, K. Hildebrand, C. Gremzow, B. Bickel and M. Alexa. Throwable Panoramic Ball Camera. In *Proceedings of SIGGRAPH Emerging technologies*, 2011.
3. R. Hartley and A. Zisserman. Multiple View Geomerty in Computer Vision. Cambridge University Press, Second Edition, 2003.
4. V. Kwatra, A. Schödl, I. Essa, G. Turk and A. Bobick. Graphcut Textures: Image and Video Synthesis Using Graph Cuts. In *Proceedings of SIGGRAPH*, 2003.
5. S. Avidan and A. Shamir. Seam Carving for Content-Aware Image Resizing. In *Proceedings of SIGGRAPH*, 2007.