Illumination Sensitive Dynamic Virtual Sets

Bert De Decker Hideaki Nii Yuki Hashimoto Dylan Moore Jay Summet Yong Zhao
Jonathan Westhues Paul Dietz John Barnwell Masahiko Inami Philippe Bekaert
Ramesh Raskar*
Mitsubishi Electric Research Laboratories (MERL)

Abstract

We describe a novel motion capture method that is sensitive to incident illumination. We will demonstrate the corresponding system at Emerging Technologies at Siggraph 2007. In this sketch, we focus on the key technical and implementation contributions involved in building a real-time special effects system where a prop is turned into a virtual sword with appropriate orientation and incident illumination.

Overview

Specifically, our system has three novel elements: (i) A photo sensing marker that can sense orientation and incident illumination (ii) Authoring of synthetic elements without explicit user interaction or color-based segmentation to identify the prop, making our system suitable for real time effects in virtual sets (i.e. automatic recovery of prop position) (iii) Correct rendering of motion and illumination blur.

We describe an economical and scalable system where the light transmitters are space-labeling beamers and beacons. Each beamer is simply an LED with a passive binary film (mask) set in front. The light intensity sequencing provides a temporal modulation, and the mask provides a spatial modulation. We use an array of such beamers, called projectors, where the binary masks of individual units are carefully chosen to exploit the epipolar geometry of the complete beamer arrangement. Each unit projects invisible (infrared) binary patterns thousands of times per second. Tags with photo sensors attached decode the transmitted space-dependent labels that enable us to compute their locations in 3D. The tags also decode the received intensities from a set of calibrated light sources, which are used to compute the orientation of the tags. Consider the light source L_i with power P_i and the vector from tag to this light source $V_i = [V_{ix}, V_{iy}, V_{iz}]$ normalized with the distance d_i between the tag and the light source. We can estimate the normal N using the intensities I_i recorded from the i-th light source by solving the following set of equations:

$$I_i = k(V_i \cdot N)\frac{P_i}{d_i^2} \quad \text{for } i = 1 \text{ to } \#\text{light sources}$$

Where k is the gain of the sensor. The location and orientation data for each unit are computed hundreds of times per second. Attaching the tags to scene points therefore yields the locations and orientations of these scene points at a very high frequency. In addition, the tags measure their incident ambient illumination. When the tagged scene points are imaged with an external camera, we can factor in real time the radiances measured at the corresponding camera pixels into the incident illuminations and the intrinsic reflectance of the corresponding scene points.

The dream of filmmakers and game developers is on-set motion capture. One of the recent examples is the motion and appearance capture from the movie Pirates of the Caribbean [2006]. Our system can support this operation with unlimited number of imperceptible, interactive and accurate tags. The top left of figure 1 shows an actor that is holding a stick with three tags attached. The left of the scene is lit with a red light source, the centre with a blue one and the right with green. In the top right of figure 1 is shown how the stick is automatically replaced with a virtual sword. For each tag our system calculates the position, orientation and incoming illumination. Using this information it is possible to draw the virtual sword with correct position, orientation and correct lighting. Notice how the color of the sword changes according to the incident illumination.

The camera captures the scene at 30 fps, while our system is able to track the marker tags at hundreds of frames per second. By rendering the synthetic elements at every frame of the motion capturing instead once for each video frame, we can obtain motion and color blur.

Applications

Fields that would immediately see the benefits of an accessible motion capture system could include independent biomedical research centers, rehabilitation clinics, and independent research scientists in many fields including physics, anthropology, and sociology. Furthermore, fields beyond human research would benefit as well: Even veterinary clinics could capitalize on the accessibility of motion tracking in examining animal gaits and behaviors for diagnosis. In this YouTube empowered world, virtual sets (such as the one we propose) at home or school may become as routine as HTML editors of yesteryear. Finally, this would be a boon to artists everywhere who have been kept out of motion capture because of price tag and technological complexity alone.

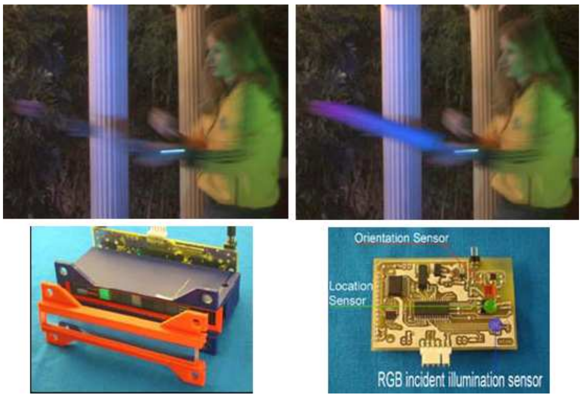

Figure 1: (Top Left) We track the tags on a stick, (Top Right) insert CG sword and change the appearance of the CG sword to match the spotlight colors. (Bottom) Our hardware is made of off-the-self components and costs 10s of US dollars.

References

ROBERTSON, B. 2006. Big moves. In *Computer Graphics World 29, 11 (Nov).*

*e-mail:raskar(at)merl.com,web:http://www.merl.com/people/raskar/