Cooperatively Augmenting Smart Objects with Projector-Camera Systems

David Molyneaux and Hans-W. Gellersen
Computing Department
Lancaster University
Lancaster, England
{d.molyneaux,hwg}@comp.lancs.ac.uk

Abstract

In this paper we propose an object-centric method of augmenting mobile smart objects using projection onto the surface of the objects. We assume that the objects are both real world objects with an inherent use, and autonomous computational nodes that can communicate with the environment in real time. We investigate how an object's capability for self description and sensing can be used in cooperation with a projector-camera system's vision capability to help locate, track and display information onto object surfaces in an unconstrained environment without using markers.

1. Introduction

The interest in augmenting everyday objects with computation is growing. These smart objects retain their original appearance and uses, while embedded computing is expected to provide added value in the background [1]. There has been much work on input to smart objects, for example, on embedding sensing, location systems and methods for accumulating and distributing knowledge. However, there is little research addressing output methods to allow the user to reveal, visualise and interact with the knowledge contained within the objects.

Common methods to visualise this information in AR would be using head mounted displays or portable devices such as PDAs. However, this encumbers users and limits interaction to a single person. Instead we propose augmenting everyday smart objects using projected displays from projector-camera systems in a way that utilises the capabilities of the smart object itself.

As a smart object can contain knowledge and sensing, we use these two characteristics to help a projector camera system to locate, track and register its projection with the objects surfaces in an unconstrained environment. Our proposed system model differs from similar approaches [2][3][4][5] as the objects themselves are pro-active clients of an environment. In our model the projector-camera system is considered as infrastructure offering a location and projection service for the smart objects. The environment is able to support spontaneous interaction with objects, as no prior knowledge of the object is required to be held by the environment. Instead, we make use of the knowledge storage capability of a smart object to allow it to describe its appearance to the environment. In addition, if the object possesses sensing capability that can be correlated with a visual change in the objects appearance (such as a movement sensor) then we also can make use of this information to constrain the visual detection task and create a more robust system.

2. Problem Analysis

If we imagine a scenario similar to that presented by Strohbach et al. [6], in which a chemical handling facility stores chemical containers in a large warehouse. In this scenario the containers are augmented with computing, giving containers knowledge of their contents and safety-related handling rules (such as closest allowed proximity to other reactive chemicals). The containers have sensors for ultrasonic ranging which allow them to compute the relative distances to other containers and infer if hazard conditions exist by evaluating the embedded rules.

If steerable projectors were installed in the facility, any containers entering the warehouse could register with the system and be detected visually. As the containers move around the warehouse the cameras could track their movements and the display information or warning messages directly onto the containers when hazard situations occur, as illustrated below in Figure 2.

Figure 1 and 2: Steerable Projector-Camera System and warning message projection onto chemical containers.

To project onto the object's surfaces we first need the object location and orientation, which can be obtained using either global or local feature computer vision methods. Global methods such as template matching or colour histograms generally require training on images of the object in every pose to create a robust system, creating both a barrier to spontaneous augmentation of objects and a large database of information that must be stored in the object. In contrast, local features attempt to uniquely describe an object using just small areas around a few key points (for example - corners) extracted from the object.

Consequently, we propose a layered global and local detection method. At the bottom level where we have the least information, this approach uses global methods such as background subtraction correlated with the object's sensors just to build up a probability map of the object's location over time and disambiguate between multiple objects with similar appearances.

This is combined with a learning approach to incrementally extract more appearance knowledge, moving up the detection layers until we have a high enough detection probability to extract local features and an object model for pose computation using methods described by Rothganger et al. in [7]. This allows an object to enter the environment with any level of knowledge - from none, to a full description of local features registered to a 3D model and be successfully detected by the system. Any new appearance knowledge we extract can also be embedded back into the object for faster detection on re-entry to the environment.

Once initially located and tracked, the object can request projection onto its surfaces. However, as the appearance of the object is dynamically changed when we project onto it, we colour correct our projection for the object surface colour using the technique described by Fuji et al. in [8] and use the inverse of the correction subtracted from the camera image to obtain an image of the object without projector light for tracking.

3. Implementation

A steerable projector was constructed from a moving head display light, DLP projector and Firewire camera (see Figure 1). The system was calibrated using the method described by Ehnes et al. in [9]. Several chemical containers were augmented with Smart-Its devices [1], and an architecture developed to coordinate event messages between the physical objects and the projector-camera system. A real time GPU-based scale and rotation invariant local feature detection system was developed. For object tracking a particle-filter based location and pose engine was also developed, with support for multiple location hypotheses.

Our initial system assumes we use both planar surfaced objects and that a manufacturer has embedded an identical 3D model and local feature knowledge into each chemical container object. Using vision alone, our system cannot discriminate between the identical containers, so cannot project warning messages when hazards are detected. However, when used in combination with simple ball-switch movement sensors, the broadcast movement events can be successfully correlated with detected visual differences and objects disambiguated.

4. Conclusion

Using the self description and sensing capabilities of smart objects provides a more robust solution to object detection than pure vision based methods. However, the projection system shares many of the problems described by Bandyopadhyay et al. in [2], such as occlusion problems due to the use of a single steerable projector. In future work we intend to extend our system to support multiple projector-camera systems and non-planar object surfaces.

References

[1] L. E. Holmquist, F. Mattern, B. Schiele, P. Alahuhta, M. Beigl, H. W. Gellersen. Smart-Its Friends: A Technique for Users to Easily Establish Connections between Smart Artefacts, In Proceedings of: Ubicomp 2001, Atlanta, USA, pp. 116-122, 2001.

[2] D. Bandyopadhyay, R. Raskar, H. Fuchs. Dynamic Shader Lamps: Painting on Movable Objects, In Proceedings of: Int. Symposium on Augmented Reality, New York, 2001.

[3] J.C. Lee, P.H. Dietz, D. Maynes-Aminzade, R. Raskar, S.E. Hudson. Automatic Projector Calibration with Embedded Light Sensors, In Proceedings of: ACM Symposium on User Interface Software and Technology (UIST), pp. 123-126, October 2004.

[4] J. Summet and R. Sukthankar. Tracking Locations of Moving Hand-held Displays Using Projected Light, In Proceedings of: Pervasive 2005, Munich, Germany.

[5] R. Raskar, P. Beardsley, J. van Baar, Y. Wang, P.Dietz, J. Lee, D. Leigh, T. Willwatcher. RFIG Lamps: Interacting with a Self-Describing World via Photosensing Wireless Tags and Projectors, In: Proceedings of SIGGRAPH 2004.

[6] M. Strohbach, H.-W. Gellersen, G. Kortuem and C. Kray. Cooperative Artefacts: Assessing Real World Situations with Embedded Technology. In Proceedings of: Ubicomp 2004, Nottingham, UK

[7] F. Rothganger, S. Lazebnik, C. Schmid and Jean Ponce. Object modeling and recognition using local affine-invariant image descriptors and multi-view spatial constraints. In Proceedings of: International Journal of Computer Vision 2005.

[8] K. Fujii, M.D. Grossberg and S.K. Nayar . A Projector-Camera System with Real-Time Photometric Adaptation for Dynamic Environments. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.814-821, 2005

[9] J. Ehnes, K. Hirota, M. Hirose. Projected Augmentation - Augmented Reality using Rotatable Video Projectors. In Proceedings of: ISMAR 2004, pp. 26-35.