Appeared in the Proceedings of the International Workshop on Image Analysis for Multimedia Interactive Services 2005

IMPERCEPTIBLE PATTERNS FOR RELIABLE ACQUISITION OF MIXED REALITY ENVIRONMENTS

Daniel Cotting, Martin Naef, Markus Gross
ETH Zurich, Switzerland
{dcotting, naef, grossm}@inf.ethz.ch

Henry Fuchs
University of North Carolina at Chapel Hill
fuchs@cs.unc.edu

ABSTRACT

Projection-based mixed and augmented reality settings often require concurrent optical camera acquisition. Unfortunately, the grabbed images frequently capture the projected imagery in addition to the desired scenery, introducing undesired interference and complicating image analysis. To efficiently improve signal-to-noise ratio, we present a method allowing the acquisition to take place under controlled illumination conditions. By exploiting the micro-mirror modulation pattern used by Digital Light Processing (DLP) projectors, a pixel-level control of light can be achieved. Since the patterns are imperceptible to the human eye and only slightly degrade the projected images, structured light techniques are introduced into human-inhabited mixed and augmented reality environments, where they often were too intrusive previously. This extended abstract gives an overview of the proposed embedding and illustrates feasibility and usefulness of the approach with representative example applications.

1. INTRODUCTION

In a variety of mixed reality systems, cameras and projectors have to run simultaneously. Unfortunately, conflicting lighting requirements have made such systems very difficult to realize: cameras need brightly lighted environments, whereas projectors need dark rooms [1, 2].

The lack of strong features in many parts of the environment has been an additional difficulty for cameras, especially for those performing 3D acquisition. Skin, clothes and furniture often appear with nearly uniform surface structure, making depth acquisition relying on stereo correspondence particularly difficult to perform. Using structured light to illuminate the scene is an option to solve this problem. However, it is highly distracting and therefore not suitable for human-populated environments.

Fortunately, if structured light is modulated sufficiently fast, the human eye is mostly unaware of it. Since micromirror based DLP data projectors modulate light in a fast way to produce a wide range of color shades, their mirror flips can be exploited to achieve precise light control in a way that is not disturbing. Lighting conditions can be specified by binary patterns, whose embedding concepts and applications will be summarized in this extended abstract.

2. RELATED WORK

Relatively little work has been done to integrate multiple cameras and projectors into a single system, allowing simultaneous display and acquisition. blue-c [1] provides an immersive environment for virtual reality and collaboration, combining projection and acquisition. The National Tele-Immersion Initiative [3] employs 3D avatars, where the geometry is acquired by trinocular stereo [4]. A combination of depth acquisition and immersive projection was proposed in the Office of the Future project [2], where a proof of concept for the embedding of invisible structured light (ISL) into DLP projections was achieved. As major restrictions, the resulting images were greyscale only, and the realization required significant modifications of the projection hardware.

As opposed to this earlier approach, our research based on the vision of ultimate light control focuses on software-based embedding [5]. We project structured light imperceptibly with conventional off-the-shelf DLP projectors, while also simultaneously displaying real-time color imagery. Our pattern embedding approach can be seen as an enabling technology for a wide range of computer graphics and vision applications. Presenting the entire field of applicable structured light algorithms is beyond the scope of this extended abstract. Several textbooks [6, 7], however, provide detailed discussions of general vision algorithms. As a proof of concept, we will present a Gray code based depth reconstruction [8] as well as a single-shot method by Vuyl-

Appeared in the Proceedings of the International Workshop on Image Analysis for Multimedia Interactive Services 2005

steke and Oosterlinck [9]. Using the novel embedding method, implementations of other algorithms are straightforward, enabling a large variety of applications in environments containing multiple projectors and cameras.

3. IMPERCEPTIBLE EMBEDDING

The underlying technology of our embedding technique is the micro-mirror based data projection technology, called Digital Light Processing (DLP). Our previously published paper [5] contains an extensive description of the DLP characteristics used for imperceptible pattern projection. In the next paragraphs we will concentrate on recapitulating the most important facts about internal projector operation and present the embedding possibilities achievable by only controlling the input data of the projector without modifying the projection hardware.

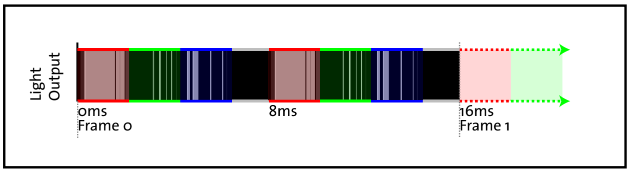

Figure 1: Mirror flip sequence for RGB value (223,47,128) during one frame. The clear segment is not used.

In DLP projectors, each displayed pixel is generated by a tiny micro-mirror, tilting towards the screen to project light and orienting towards an absorber to keep the pixel black. Gradations of intensity values are created by flipping the mirror in a fast modulation sequence, while a filter wheel rotates in the optical path to create colors. Figure 1 illustrates the flip sequence for the RGB value (223,47,128).

Figure 2 lists the mirror sequences for all 256 values of red. During the period from 0.4 ms to 0.5 ms no mirror flips occur, i.e. the mirrors are either *off* or *on* during the entire period. Thus, all 256 red values can be grouped into three classes: *off*, *on*, and *unreliable* (values that are dithered or otherwise not reliably *off* or *on*).

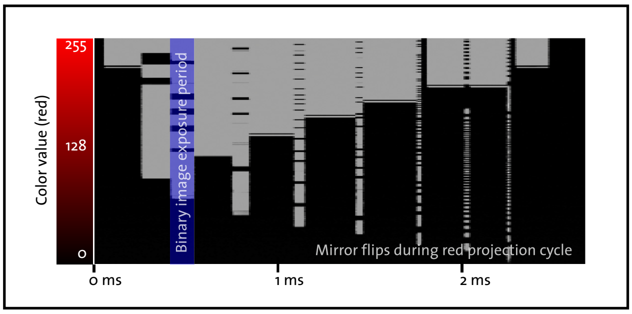

Figure 2: Measured mirror flip sequences for all 256 values of red (type of projector: projectiondesign F1). Darker gray values indicate mirrors not consistently *on* or *off*.

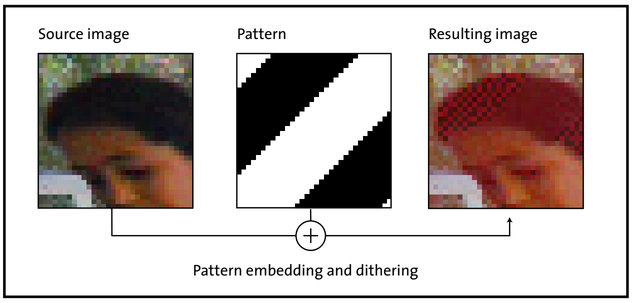

Figure 3: Embedding a binary image into a user-defined full-color image.

In case the three color channels are independent from each other for a given projector, the same thoughts apply to the green and the blue channel. Those projector models having interdependent color channels require an extended version of the classification method based on a sampling of a subspace of the entire color space.

Consider exposing a synchronized video camera during the period from 0.4 ms to 0.5 ms, called *binary image exposure period* (BIEP). Obviously, the captured image is not affected by the actual content of the projection, but only by what is displayed during the BIEP. We can now adjust each pixel in the projection image so that it comes from the *off* or *on* set, whichever is specified by the corresponding pixel in the binary control image. The color differences caused by this real-time mapping process are compensated by a slightly modified version of error-diffusion dithering, as described in many textbooks [10]. A source image and the result after embedding a specific pattern in the binary image exposure period are depicted in Figure 3. Note that other periods with similar properties exist, but in practice they are too short for reliable camera exposure.

Figure 4: Three different levels of magnification of the projection in operation. Note that the dithering is completely invisible in actual scale, and no visible hint of the embedded binary image can be found.

Since human perception is able to quickly recover color constancy, the color shift which is noticeable in Figure 3 does not appear as pronounced in actual projections (cf. Figure 4). Refer to [5] for a more thorough discussion of the various projection options and the resulting quality of perception.

Appeared in the Proceedings of the International Workshop on Image Analysis for Multimedia Interactive Services 2005

4. RESULTS

Based on the ideas presented in the previous sections, we have implemented a prototype system that concurrently displays a user application and provides acquisition using off-the-shelf components (projectiondesign F1 projectors and Point Grey Dragonfly greyscale video cameras, see Figure 5). The cameras allow external synchronization, and they provide a programmable trigger delay. Even though faster and more sensitive cameras would result in higher contrast and an increased depth of field, we consider the resulting images to be fully adequate for our purposes.

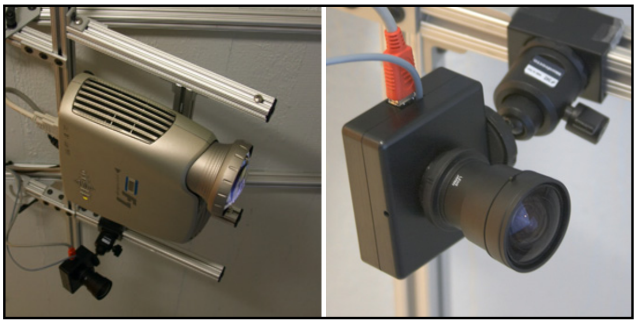

Figure 5: Hardware components: Projector and camera.

The examples presented in this section are used as illustrations of various applications that can successfully be implemented with our imperceptible embedding of patterns. We do not aim at presenting new state-of-the-art structured light algorithms, but rather limit ourselves to comprehensive and illustrative examples.

4.1. 3D acquisition

Two depth acquisition algorithms were implemented to demonstrate the suitability of our embedding for surface acquisition based on structured light: a basic Gray code surface extraction and a single-shot method based on previous work by Vuylsteke and Oosterlinck [9].

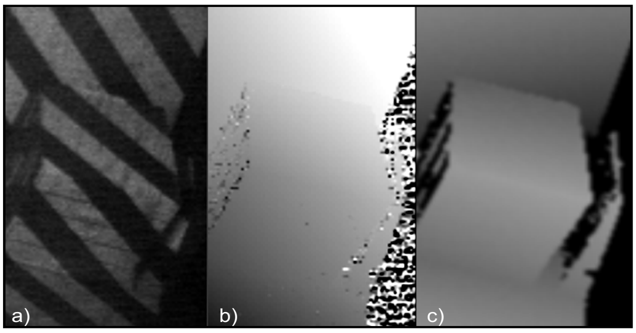

Figure 6: Depth acquisition: a) Image of a Gray code pattern. b) Extracted Gray code value. c) Depth image.

Gray code surface extraction [8] can be achieved imperceptibly by projecting the corresponding sequence patterns using our technique. Figure 6 shows a visualization of three stages in the Gray code surface extraction pipeline: a single Gray code image captured by the camera, the decoded gray code values, and the computed depths representing the 3D geometry.

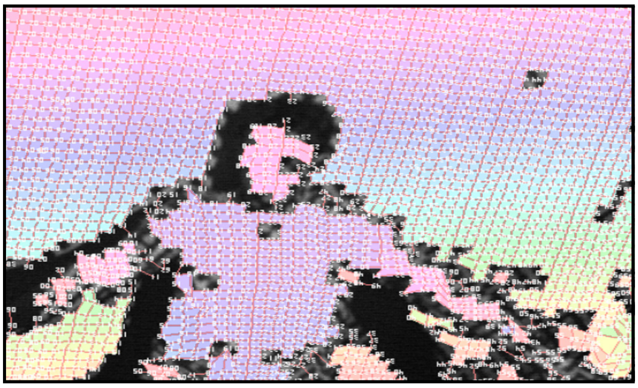

Figure 7: Color-coded depth overlaid on a capture of a moving person acquired by the chosen single-shot method.

Single-shot methods are commonly used for applications wishing to acquire dynamic, moving scenes. Figure 7 shows a color-coded output of a depth map created with an implementation of the chosen example algorithm [9].

Although we have only presented two well-known algorithms, we expect our embedding method to work with any of the contemporary state-of-the-art binary structured light methods.

4.2. Tracking

Our example tracking filter (for details see [5]) detects a user's hand and head position using a simple and efficient single-shot tracking method in real-time (at 20 fps, the

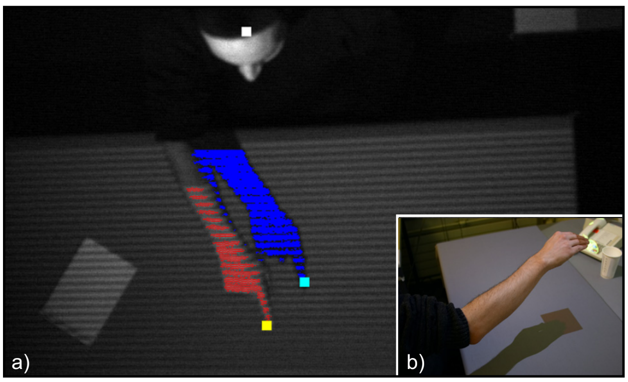

Figure 8: Tracking system at work: a) Color-coded camera image showing shadow of the hand (blue), extracted image of the hand (red), finger tips in the shadow (turquoise) and on the hand (yellow). b) Visible image of the same scene.

camera frame rate). The tracking algorithm at work is shown in Figure 8. The user's head and hand can be tracked imperceptibly, allowing encumbrance-free hand-guided interaction and encumbrance-free head-tracking. Note the lack of any discernible artifacts of the embedded binary image.

5. CONCLUSION

This extended abstract summarizes our approach for imperceptible embedding of binary patterns into conventional unmodified DLP projectors, allowing simultaneous (immersive) display and acquisition under controlled lighting conditions.

With its help, a wide range of binary structured light algorithms can be transferred to mixed and augmented reality systems, where visible pattern projections usually are too intrusive and severely limit the usability of the systems. As a proof of concept, we have presented a set of example algorithm implementations, clearly demonstrating the versatility of our imperceptible embedding and its suitability for a wide range of computer graphics and vision algorithms.

6. FUTURE WORK

Multiple projector modules are needed to achieve an increasingly immersive environment. A resulting issue is the one of light control: Each module projecting its own structured light pattern onto the scene potentially interferes with the illumination from the other projectors. For overlapping frusta, we plan to devise appropriate time-division and frequency-division multiplexing approaches. Furthermore, for the creation of large seamless displays, issues like geometry alignment, color balancing and gamut matching have to be addressed in the future, taking into account the constraints resulting from our embedding procedure. Invisible structured light enables imperceptible calibration pattern projection and point correspondence creation during system operation. Therefore, a procedure can be envisioned, which continuously refines calibration parameters in an automatic way. We are currently investigating these options and are confident that calibration will significantly increase in flexibility with our imperceptible embedding.

REFERENCES

[1] M. Gross, S. Würmlin, M. Naef, E. Lamboray, C. Spagno, A. Kunz, E. Koller-Meier, T. Svoboda, L. Van Gool, S. Lang, K. Strehlke, A. Vande Moere, and O. Staadt, "blue-c: A Spatially Immersive Display and 3D Video Portal for Telepresence," in SIGGRAPH 2003 Conference Proceedings, ACM SIGGRAPH Annual Conference Series, 2003.

[2] R. Raskar, G. Welch, M. Cutts, A. Lake, L. Stesin, and H. Fuchs, "The Office of the Future: A Unified Approach to Image-Based Modeling and Spatially Immersive Displays," Proceedings of SIGGRAPH 98, pp. 179–188, July 1998.

[3] A. Sadagic, H. Towles, J. Lanier, H. Fuchs, A. van Dam, K. Daniilidis, J. Mulligan, L. Holden, and B. Zeleznik, "National Tele-Immersion Initiative: Towards Compelling Tele-Immersive Collaborative Environments." presentation given at Medicine meets Virtual Reality 2001 Conference, Jan. 2001.

[4] J. Mulligan and K. Daniilidis, "View-Independent Scene Acquisition for Tele-Presence," in Proc. of the International Symposium on Augmented Reality, pp. 105–108, 2000.

[5] D. Cotting, M. Naef, M. Gross, and H. Fuchs, "Embedding Imperceptible Patterns into Projected Images for Simultaneous Acquisition and Display," in Proceedings of International Symposium on Mixed and Augmented Reality 2004, pp. 100–109, IEEE Computer Society Press, 2004.

[6] R. M. Haralick and L. G. Shapiro, Computer and Robot Vision (2 volumes). Reading, MA: Addison-Wesley, 1992.

[7] D. A. Forsyth and J. Ponce, Computer Vision: A Modern Approach. Prentice Hall, 2002. ISBN 0-13-085198-1.

[8] J. R. Bitner, G. Ehrlich, and E. M. Reingold, "Efficient Generation of the Binary Reflected Gray code and Its Applications," Communications of the ACM, vol. 19, no. 9, pp. 517–521, 1976.

[9] P. Vuylsteke and A. Oosterlinck, "Range Image Acquisition with a Single Binary-Encoded Light Pattern," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, pp. 148–164, Feb 1990.

[10] J. Foley, A. von Dam, S. Feiner, and J. Hughes, Computer Graphics: Principles and Practice. Reading, MA: Addison-Wesley, 1992.