IEEE.org IEEE *Xplore* IEEE SA IEEE Spectrum More Sites Subscribe Subscribe Cart Create Account Perso Sign

Browse My Settings Help Institutional Sign In

Institutional Sign In

All

ADVANCED SEARCH

Conferences > Proceedings 2006 IEEE Interna...

Remote-collaboration system using mobile robot with camera and projector

Publisher: IEEE

Cite This PDF

T. Machino ; S. Iwaki ; H. Kawata ; Y. Yanagihara ; Y. Nanjo ; K.-i. Shimokura **All Authors** •••

20 Paper Citations **2** Patent Citations **324** Full Text Views

Abstract

Document Sections

I. Introduction

II. Concept for New Collaboration between Robot and Human over Network

III. Hardware Design for a Mobile-Type Scope

IV. Optomechanical Design

V. Conclusion

Authors

Figures

References

Citations

Keywords

Metrics

More Like This

Abstract: We have been studying a remote-collaboration system called SCOPE (sight collaboration by projection effect) featuring image projecting and capturing capabilities as imple... **View more**

Metadata

Abstract:

We have been studying a remote-collaboration system called SCOPE (sight collaboration by projection effect) featuring image projecting and capturing capabilities as implemented in a maintenance robot. With the help of SCOPE and a remote support person, an on-site worker can perform maintenance very efficiently. We propose a mobile SCOPE that can significantly expand the area of activity of the conventional SCOPE. To enable an on-site worker to share the field of view of a remote support person, we developed a technique for aligning the optical axes of a camera and a projector. We present experimental data that demonstrates the validity of our optomechanical design. Finally, we show that the mobile SCOPE enables workers to continuously share their field of view, no matter where the mobile SCOPE moves

Published in: Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006.

Date of Conference: 15-19 May 2006

INSPEC Accession Number: 9120248

Date Added to IEEE *Xplore*: 26 June 2006

DOI: 10.1109/ROBOT.2006.1642326

Print ISBN:0-7803-9505-0

Publisher: IEEE

Print ISSN: 1050-4729

Conference Location: Orlando, FL, USA

Contents

I. Introduction

To create a more efficient environment in which to perform remote maintenance over a network, we have been studying a new collaboration concept, called "robot-augmented communication" (RAC) [11]. In RAC, a maintenance robot and an on-site human worker, who may not be skilled in

Sign in to Optimize Reading

IEEE websites place cookies on your device to give you the best user experience. By using our websites, you agree to the placement of these cookies. To learn more, read our Privacy Policy.

Accept & Close

maintaining the target object, are assisted by a veteran operator of the object at a remote support center. RAC provides effective support for on-site workers by augmenting the communication ability of the remote support person with robotic functions that people do not have.

Authors

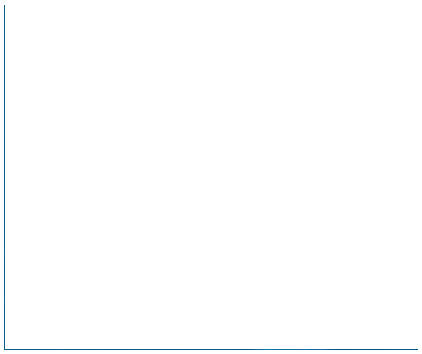
Figures

References

Citations

Keywords

Metrics



More Like This

Obstacle Avoidance System for Wheeled Mobile Robots by CMOS Image Sensors
2014 Tenth International Conference on Intelligent Information Hiding and Multimedia Signal Processing
Published: 2014

Localization and Error Correction for Mobile Robot with an Image Sensor
2006 SICE-ICASE International Joint Conference
Published: 2006

Show More

IEEE websites place cookies on your device to give you the best user experience. By using our websites, you agree to the placement of these cookies. To learn more, read our Privacy Policy.

Accept & Close

IEEE Personal Account

CHANGE USERNAME/PASSWORD

Purchase Details

PAYMENT OPTIONS

VIEW PURCHASED DOCUMENTS

Profile Information

COMMUNICATIONS PREFERENCES

PROFESSION AND EDUCATION

TECHNICAL INTERESTS

Need Help?

US & CANADA: +1 800 678 4333

WORLDWIDE: +1 732 981 0060

CONTACT & SUPPORT

Follow

About IEEE *Xplore* | Contact Us | Help | Accessibility | Terms of Use | Nondiscrimination Policy | IEEE Ethics Reporting | Sitemap | IEEE Privacy Policy

A not-for-profit organization, IEEE is the world's largest technical professional organization dedicated to advancing technology for the benefit of humanity.

© Copyright 2023 IEEE - All rights reserved.

IEEE Account

» Change Username/Password

» Update Address

Purchase Details

» Payment Options

» Order History

» View Purchased Documents

Profile Information

» Communications Preferences

» Profession and Education

» Technical Interests

Need Help?

» **US & Canada:** +1 800 678 4333

» **Worldwide:** +1 732 981 0060

» Contact & Support

About IEEE *Xplore* | Contact Us | Help | Accessibility | Terms of Use | Nondiscrimination Policy | Sitemap | Privacy & Opting Out of Cookies

A not-for-profit organization, IEEE is the world's largest technical professional organization dedicated to advancing technology for the benefit of humanity.

© Copyright 2023 IEEE - All rights reserved. Use of this web site signifies your agreement to the terms and conditions.

IEEE websites place cookies on your device to give you the best user experience. By using our websites, you agree to the placement of these cookies. To learn more, read our Privacy Policy.

Accept & Close

Amazon Ex. 1012.0147

Appendix A-6

Screen capture completed by J L Mullins from *Google Scholar* on May 9, 2023

Articles

6 results (0.02 sec)

Any time
Since 2023
Since 2022
Since 2019
Custom range...

2007 — 2009

Search

Sort by relevance
Sort by date

Create alert

Remote-collaboration system using mobile robot with camera and projector

Search within citing articles

Robots with projectors: an alternative to anthropomorphic HRI

J Park, GJ Kim - Proceedings of the 4th ACM/IEEE international ..., 2009 - dl.acm.org

Current forms of Human Robot Interaction (HRI) pursue mostly anthropomorphism and direct interaction. That is, the interaction paradigm is based on imitating how" people" interact with ...

☆ Save 99 Cite Cited by 47 Related articles All 5 versions

Framework and service allocation for network robot platform and execution of interdependent services

Y Nakamura, T Machino, M Motegi, Y Iwata... - Robotics and ..., 2008 - Elsevier

This research introduces a framework for network robot platform (NWR-PF) and a service allocation method for heterogeneous distributed robots; they cooperate to perform services ...

☆ Save 99 Cite Cited by 22 Related articles All 4 versions

Field trial of asynchronous communication using network-based interactive child watch system for the participation of parents in day-care activities

H Kawata, Y Takano, Y Iwata... - ... on Robotics and ..., 2008 - ieeexplore.ieee.org

A recent trend in society is for young parents to ask for support in performing their childcare duties. Against this background and to promote the intermixing of parents and children and ...

☆ Save 99 Cite Cited by 7 Related articles All 4 versions

Spatial memory for collaboration support system in intelligent space-User identification based on spatial memory

M Niitsuma, H Hashimoto - SICE Annual Conference 2007, 2007 - ieeexplore.ieee.org

This paper presents an implementation of the spatial memory system for supporting collaboration activity in Intelligent Space. in this paper, a design of the new spatial memory ...

☆ Save 99 Cite Cited by 2 Related articles All 2 versions

幼児施設における保護者参加型子ども見守りシステムの提案

河田博昭， 高野陽介， 岩田義行， 金丸直義... - 電子情報通信学会 ..., 2008 - search.ieice.org

近年, 核家族化, 両保護者が就業する割合の増加などの社会環境の変化に伴い, 幼児施設による育児活動へのサポートは, 保護者にとって不可欠なものとなりつつあり ...

☆ Save 99 Cite Cited by 3 Related articles All 3 versions

[CITATION] Interacting with robots “indirectly” through projected display

J Park, GJ Kim - 5th International Conference on Ubiquitous and ..., 2008

☆ Save 99 Cite Cited by 3 Related articles

Appendix A-7

Check for updates

Robots with Projectors: An Alternative to Anthropomorphic HRI

Jongkyeong Park and Gerard J. Kim
Digital Experience Laboratory
College of Information and Communication
Korea University, Seoul, Korea
82-2-3290-3196
parkjk@korea.ac.kr

ABSTRACT
Current forms of Human Robot Interaction (HRI) pursue mostly anthropomorphism and direct interaction. That is, the interaction paradigm is based on imitating how "people" interact with one another (e.g. using spoken language, gestures, facial expression, etc.). However, "direct" interaction/contact with the "robot," often causes significant inconvenience and usability problems. In this paper, we present an alternative to the anthropomorphic interface using a projected display and indirect, yet already familiar GUI based interaction. That is, the projector (on the moving robot) projects information on the nearby surface and provides a relatively large area through which indirect GUI based interaction can occur. As an instance of such an HRI paradigm, we present a moving robot kiosk that projects displays around itself and serve and interact with multiple people at once. We report our on-going development efforts and a pilot experimental study that compares it to the typical touch screen based direct HRI.

Categories and Subject Descriptors
H.5.2 [**HCI**]: User Interfaces –*Interaction Style, Input devices and strategies.*

General Terms
Design, Experimentation, Human Factors.

Keywords
Human robot interaction, Anthropomorphism, Indirect/direct interaction, Projector, Large display.

1. INTRODUCTION
Robots are finding their ways into our everyday environment [1] and as such their social aspect and interaction with humans have become a very important issue for their successful deployment and control. However, the conventional anthropomorphic and direct interaction with robots has not been so successful mainly because (1) most robots are not (or perhaps cannot be) optimized human factor-wise for the diverse users (e.g. touch screen interface, see Figure 1) and (2) recognition technologies (for language, face, etc.) for typical anthropomorphic interfaces is still limited in many ways. In fact, the problem of anthropomorphic interfaces generating false expectations of human-like responses and behaviors, and becoming a source of disappointment and confusion to users, has been well known [2].

As an alternative, "indirect" interaction mediated through a large projected display system is proposed. Figure 2 shows one possible way to employ such a framework where, in this particular case, the projector on the robot projects onto the floor around the robot and the interaction is made possible by using the camera and detection of laser pointers or software markers. We claim that such an interaction framework can alleviate much of the human factors problem and enhance the overall experience by the "enlarged" communication channel and amplified presence of the robot. The proposed framework is flexible to accommodate different types of users and tasks, one-to-many interaction, and even mixing of real and virtual objects. When combined with a hand-held interaction device, it can be extended to support a variety of interaction techniques with the consideration of robot's mobility.

In this short paper presentation, we report our progress in realizing the proposed HRI framework and a comparative experimental study to the conventional touch screen based HRI for task efficiency, usability and user satisfaction.

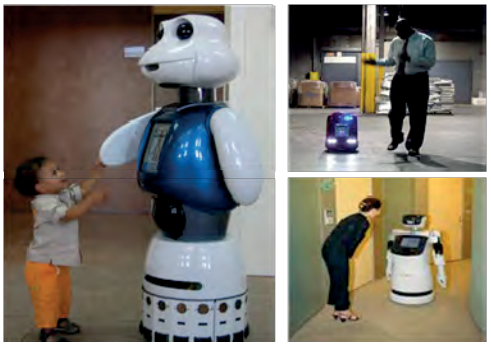


Figure 1. Interacting "directly" with robots. Robots cannot be ergonomically designed for all types of users. This makes the popular touch screen interface very difficult to use.

2. RELATED WORK
Most current HRI paradigms are anthropomorphic employing physical or graphical faces and human like behavior and recognition capabilities [3][4][5][6][7].

Copyright is held by the author/owner(s).
HRI'09, March 11–13, 2009, La Jolla, California, USA.
ACM 978-1-60558-404-1/09/03.

Amazon Ex. 1012.0151

As an alternative to conventional HRI, the idea of "projector-on-robot" was partially inspired from the movie "Star Wars" (holographic display given by R2D2) [8]. Machino et al. [9] also presented a similar idea but concentrated mostly on how to offer augmented information on the interested object in the environment. As already indicated, our framework suggests to use the projection display (and GUI) as the major mediating element between the human user and the robot. Figure 3 illustrates the idea. Note that the conventional interfaces may still be employed but perhaps in a supporting fashion due to their limited reliability. For instance, the robot can give information and guidance as it moves through the hall and interact with people mostly through the display (using hand-held interaction devices) and also by voice/sound/gestures.

3. IMPLEMENTATION & EXPERIMENT

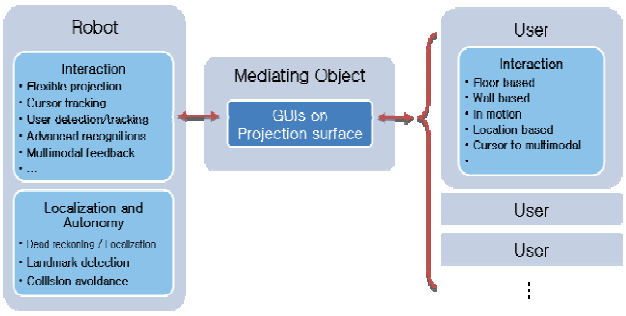


Figure 2. Major aspects of the indirect HRI framework.

Figure 2 shows several distinguishing aspects of the proposed HRI framework. On the part of the robot, aside from basic autonomy, tracking and recognition capabilities, it needs to have flexible projection capabilities (e.g. on the wall, floor, etc.), multiple user detection. In addition, a surface centered interaction framework needs to be employed for the large moving screen and to support remote interaction. Figure 3 shows our current implementation.

We conducted a pilot experimental study, comparing our HRI framework to a conventional touch screen based one. A group of subjects were tested for task completion efficiency and user experience for two different tasks, one requiring minimal amount of interaction (1~2) and another many (5~6). Contrast to our hypotheses, the proposed indirect HRI was not superior in terms of task efficiency due to the strong familiarity of the touch screen interfaces. However, it showed superiority in approachability and high satisfaction due to the large sized display. The post interviews and behavioral observation also indicated promising aspects of our framework with regards to ergonomics and ways for further improvements.

4. CONCLUSION

We have proposed a new indirect HRI framework based on the idea of mediating the interaction through projection from the robot. Our work is motivated by the lack of usability with the current mostly direct approach to interaction design for HRI. We have carried out a preliminary implementation and a comparative experiment to assess the feasibility of our approach. We are continuing to implement and complete other aspects of the proposed framework including basic robotic behaviors, middleware for multiple user and sensor/IO management, display surface selection, cursor tracking, and projection distortion correction. Based on our pilot experiment, a more formal and large scale experiment will be carried out as well.



Figure 3. The proposed HRI platform with floor projection. Bottom pictures show marker based interaction using an UMPC (left) and laser pointer based interaction (right).

5. REFERENCES

[1] Gates, B., A Robot in Every Home, Scientific American, Jan, 2007.

[2] Schneiderman B. and Plaisant, C., Designing the User Interface, pp. 484-490, Addison Wesley, 2004.

[3] Paulos, E. and Canny, J., Ubiquitous Tele-embodiment: Applications and Implications, International Journal of Human-Computer Studies, Vol. 46, No. 6, pp. 861-877, 1997.

[4] The Roboceptionist Project, http://www.andrew.cmu.edu/user/amundell/tank/project.htm

[5] The Kismet Project, http://www.ai.mit.edu/projects/humanoid-robotics-group/kismet/kismet.html

[6] Nourbakhsh, I., Course Notes for Principles of Human Robot Interaction, Unpublished Work, 2007.

[7] Jindai, M., Watanabe, T., Shibata, S., Yamamoto, T. , Development of a Handshake Robot System for Embodied Interaction with Humans, Proc. of the IEEE International Symposium on Robot and Human Interactive Communication, pp. 710-715, 2006.

[8] 20th Century Fox, Star Wars (The Movie), http://en.wikipedia.org/wiki/Star_Wars, 1977.

[9] Machino, T., Iwaki, S., Kawata, H., Yanagihara, Y., Nanjo, Y. and Shimokura, K., Remote Collaboration System Using Mobile Robot with Camera and Projector, Proc. of IEEE Intl. Conf. on Robotics and Automation, pp. 4063-4068, 2006.

Amazon Ex. 1012.0152

Appendix A-8

2008 IEEE International Conference on
Robotics and Automation
Pasadena, CA, USA, May 19-23, 2008

Field Trial of Asynchronous Communication Using Network-Based Interactive Child Watch System for the Participation of Parents in Day-Care Activities

Hiroaki Kawata*1, Yosuke Takano*2, Yoshiyuki Iwata*1,
Naoyoshi Kanamaru*1, Ken-ichiro Shimokura*3, Yoshihiro Fujita*2

*1 Network appliance and services project, NTT Cyber solutions laboratories,
*2 NEC Corporation,
*3 ATR Intelligent Robotics and Communication Laboratories

Abstract— **A recent trend in society is for young parents to ask for support in performing their childcare duties. Against this background and to promote the intermixing of parents and children and support information exchange between parents and nursery teachers, we developed the network-based interactive child-watch system called "Meru-robo Digital Log system". It is composed of the "Action Switch Platform (AcSP)" (developed by NTT Laboratories) for exchanging dialog with a "personal robot" (called "PaPeRo") (developed by NEC Corporation) by using cellular phone text messaging. A field trial showed that parents could send their intentions and receive information in a schedule-flexible manner via cellular phone text messaging. Since the robot in the nursery linked the parents to the teacher, their information exchanges did not interfere with the activities of the children. We report results of information exchanges using this system in two situations: parents are the same building (but not the same room as the children) and parents are in other locations going about their daily life. In the limited situation of children playing in a nursery, we found that parents in remote locations could get detailed information about their children's activities and expressions when desired via asynchronous communication.**

Index Terms - child-watch system, nursery facility, cellular phone, asynchronous communication

1. INTRODUCTION

More and more people are recently changing their lifestyle by having both parents working. Families now need to collaborate with the service providers taking care of their pre-school children and aged parents. To better support working parents, nursery facilities are offering extended opening hours. To encourage the elderly to lead a self-supporting life, more day service centers are offering home visit nursing. As implied, it will be more important to support collaboration between the family and the service providers.

The functionality offered by modern robots has advanced to the point that many of societies needs can be fulfilled by using these technologies. Several recent studies have targeted the effective support of the elderly and children through the use of robots. Experiments have been conducted on robots providing therapy services to elderly people and child-care in nurseries[1][2][3][4]. Several studies have been made on robotic systems and the remote control of robots for remote collaboration[5][6][7]. As mentioned above, several studies have been made on the use of robots to provide effective support for elderly people and children. However, no study has examined using robots to allow the parents to collaborate with service providers as well as their aged parents and children.

Higher levels of "safety" and "security" are being demanded, and expectations for services that allow parents to watch over children from a remote location and services that activate communication between parents and children are growing. We introduce here our concept of a robot system that establishes asynchronous communication between remote parents, teachers, and children.

We developed "Meru-robo Digital Log system", a network-based interactive child watch system for the participation of parents in day-care activities as a prototype of our concept. This system is composed of "Action Switch Platform (AcSP)" (developed by NTT Laboratory) for exchanging dialog with "PaPeRo", a personal robot developed by NEC Corporation. A key feature is that it is based on cellular-phone text messaging. The parent writes an e-mail message describing the parent's wishes and sends it to the system. PaPeRo receives the message which is passed to

978-1-4244-1647-9/08/$25.00 ©2008 IEEE.

Authorized licensed use limited to: Purdue University. Downloaded on April 25,2023 at 14:53:49 UTC from IEEE Xplore. Restrictions apply.

Amazon Ex. 1012.0154

AcSP and converted into an action command acceptable to PaPeRo. The teacher, at the appropriate time, triggers PaPeRo to process the command. PaPeRo follows the parent's instructions and collects video images of the child's response to the parent's message. The collected video data is returned to the parent.

An experiment was conducted at two child care facilities to validate our concept; 12 families participated. We report trial results and interview data collected from the families and the teachers.

We describe our concept in the second chapter. Next, we describe the"Meru-robo Digital Log system," the prototype system used to validate our concept. We report the results of the experiment in the fourth chapter; our conclusions are given in the fifth chapter.

2. Robot System with Asynchronous Communication

2.1. *Problem of nursery facilities for parents*

Given the background of working parents utilizing child-care services, parents and teachers are sharing roles for child development. Parents can watch their child activities and can listen to the voice of their child at home. This is not possible while the parents are working. If the parents have a spare moment such as a coffee break, there is no direct way of getting detailed information about their child's activities. The parents are limited to second-hand information about their children provided by their child's teacher. Since working parents are very busy, their communication time is generally short. Therefore, parents cannot learn of their children's activities in any detail. Moreover, the parents and teachers have insufficient face time to be able to share the parent's policies and requests about day-care activities, as well as learning the teacher's advice about day-care activities.

Trials have been conducted in which parents could get information about their child's activities in Japanese nursery facilities[8][9]. Past trials set a Web-camera in the play room to deliver video of the children's activities to the remote parent's PC. This arrangement is effective in providing a parent with a general sense of reassurance about the day-care activities. Such Web-cameras, however, have a fixed position (usually on the ceiling) and are wide angle. The parent may have to wait for some time to see their child actually doing anything. This can be a serious problem since working parents have so little free time. Working parents are constrained and can only peruse web-cam feeds at particular periods such as a coffee break.

Our solution is to realize a system that permits the parent to establish an immediate link with their child whenever possible and interact with the child as needed.

2.2. *Our concept*

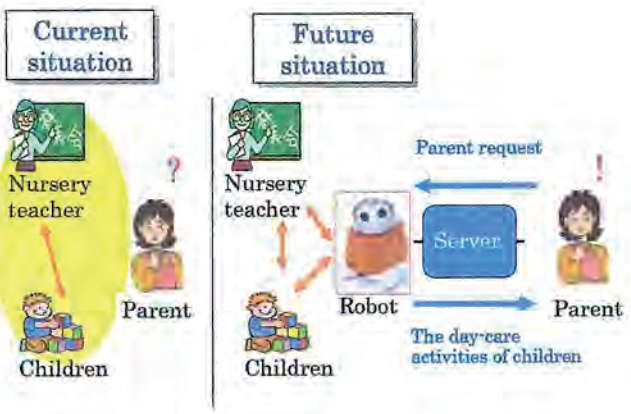


Fig. 1 Future situation in nursery facility.

Fig. 1 show the Future situation in nursery facility which we image. We think that the problems can be effectively resolved by using a robot system to link parents to the nursery teacher and child. In order to realize the proposed system, we first considered the robotic systems on remote collaboration model used to pass the information provided by a remote expert to an on-site worker and observe the results of the worker's actions.

The goal of these robotic systems are to realize an environment in which the remote expert can get feed back and so correct the worker's mistakes and the worker can get the information needed to finish the maintenance or repair actions scheduled. The remote expert and worker use video communication to share the condition of the work site in real time; a remote control robot is placed at the worker's site. The expert remotely controls the robot to assist the worker in understanding the operations needed. For example, the robot can highlight the part piece of equipment that must be tested next. This establishes intuitive communication between the remote parties.

To realize our proposed environment, a robot system is effective as the instruction medium to link remote parents with the teachers and children. But, as mentioned in the previous section, it is difficult for remote parents and nursery teacher, child in nursery facility to conduct synchronous communication. Because the parents, teacher, and child have different daily schedules.

A robot system using on asynchronous communication between parents and the teacher and child, allows us to develop new robot functionality to provide instruction channels that allow all parties to conduct comfortable communication. Our developed robot system needs to have the following capabilities.

Authorized licensed use limited to: Purdue University. Downloaded on April 25,2023 at 14:53:49 UTC from IEEE Xplore. Restrictions apply.

Amazon Ex. 1012.0155

- Make it easy for parents and teachers to asynchronously input instructions to the robot system.
- Interact with child and collect information from the child in response to the asynchronous instructions from remote parents and teachers.
- Display the information collected to remote parents and teachers via asynchronous communication.

To validate our robot system, we developed Meru-robo Digital Log system, a network-based prototype system. It is composed of the "Action Switch Platform (AcSP)" (developed by NTT Laboratories) and the personal robot "PaPeRo" (developed by NEC Corporation).

3. Our Prototype System : Meru-robo Digital Log System

3.1. Overview of Meru-robo Digital Log system

We describe here our network-based interactive child watch system for the participation of parents in the day-care activities of their children : Meru-robo Digital Log system. Figure 2 shows overview of "Meru-robo Digital Log system". In this system, the remote parent uses his/her cellular phone as the interface device, because it is used on a routine basis and can be used anywhere at anytime. The teacher and children use the personal robot "PaPeRo" as the interface device, because it was designed to be perfectly safe for children and provides the interfaces needed. In this system, the "Action Switch Platform (AcSP)" on the network links PaPeRo to the cellular phone of the parent.

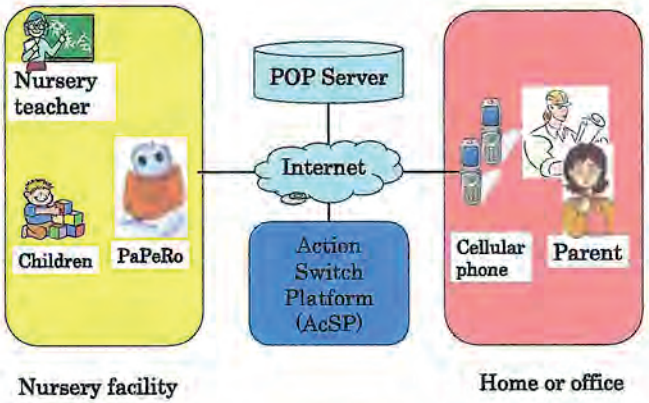


Fig. 2 Overview of "Meru-robo Digital Log system".

3.1.1Meru-robo Function

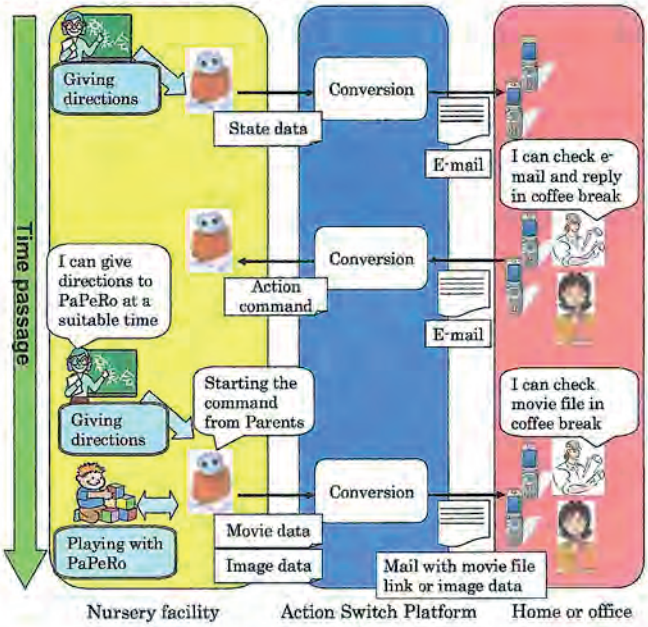


Fig.3 Overview of "Meru-robo Function".

This system has two functions but in this paper we explain only the "Meru-robo Function", an overview of which is given in Figure 3. The parent sends e-mail or reply to AcSP to access the Meru-robo Function. Using the Meru-robo function provided by this system, the parent can indicate to AcSP via cellular phone text messaging the action that their child should perform (play). The teacher gives directions to PaPeRo via touch control. When PaPeRo is triggered by the nursery teacher, PaPeRo starts the command received from AcSP and plays with the children. At the same time, PaPeRo take a movie of the child with its eye camera. The movie will include the child's facial expression, because PePeRo's eye camera is level with the child's eyes. The movie taken by PaPeRo is sent to the parent's cellular phone by Mail as a movie file link.

By using e-mail(cellular phone text messaging) as the communication tool, the parent can create and send messages when suitable for the parent and get, in return, the information desired. The parents have a direct way of getting detailed information about their child's activities. The response of the child is captured on video by PaPeRo and replayed asynchronously to the parent.

The teacher can decide what time is appropriate given the day-care schedule in the facility for starting robot activities or sending the information to remote parents.

In this way, through the robot, the parent can send a request and message to their child and the teacher, and the teacher can send details of the activities of the children to their parents at times suitable for each other.

Authorized licensed use limited to: Purdue University. Downloaded on April 25,2023 at 14:53:49 UTC from IEEE Xplore. Restrictions apply.

Amazon Ex. 1012.0156

3.2. System constitution

Meru-robo Digital Log system is composed of AcSP, PaPeRo, Cellular Phone and POP server. We describe each part as below.

3.2.1Action Switch Platform (AcSP)

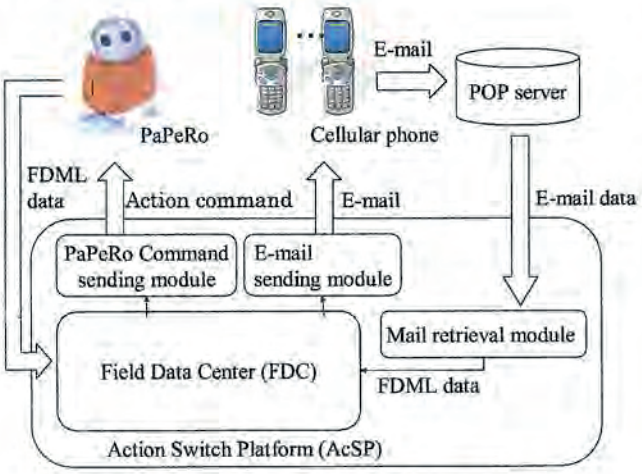


Fig.4 Outline of Action Switch Platform (AcSP).

AcSP is based on the information distribution platform "Field Data Center(FDC) " and some other modules. Figure 4 outlines AcSP. For realizing the asynchronous communication between FDC and the Personal robot, we developed a Mail retrieval module, a PaPeRo command sending module, and an E-mail sending module.

FDC was developed by other members of our laboratory[10]. The information distribution data format "Field Data markup Language(FDML) " , which is based on XML, is used to link the various devices across the network and pass information between them[11]. FDC stores and manages data in FDML format; it can check the conditions of events through the application of IF-THEN rules. FDC makes it easy to realize the functions needed for realizing link event processing with external applications. Therefore, desired events can be registered with FDC ahead of time and services and external applications can be triggered as desired by checking the conditions of events.

The Mail retrieval module converts the cellular phone text message into FDML data. It acquires e-mails from the mail server (POP server) at regular intervals. If a new e-mail is found, it extracts the body of the e-mail. This module holds pairs of keywords and FDML data. If a keyword registered in it is found in the message body, it sends the FDML data that matches the keyword to FDC.

The PaPeRo command sending module and E-mail sending module are external applications that are triggered as desired by checking the conditions of events on FDC. The PaPeRo command sending module can send an action command over HTTP protocol to PaPeRo to control its speaking function and actuators. The E-mail sending module can send one of several standard e-mails to parents.

3.2.2Personal robot "PaPeRo"

PaPeRo has several default scenarios written in a scripting language that describe basic behaviors and actions. PaPeRo's basic functions include a speaking function and an actuator control function. The software platform "robot virtual machine" in PaPeRo interprets these scenarios, and triggers the PaPeRo's functions as performance flow. This platform can switch these scenarios, received an action command over HTTP protocol, from other devices on the network. PaPeRo is able to run several scenarios in parallel such as one to interact with the child and another for receiving commands from the network. Therefore, while PaPeRo is playing with a child, PaPeRo's action and speech changed by the server. The child is unaware of any switching or activation of different scenarios.

3.2.3Processing of Meru-robo function in Meru-robo Digital Log system

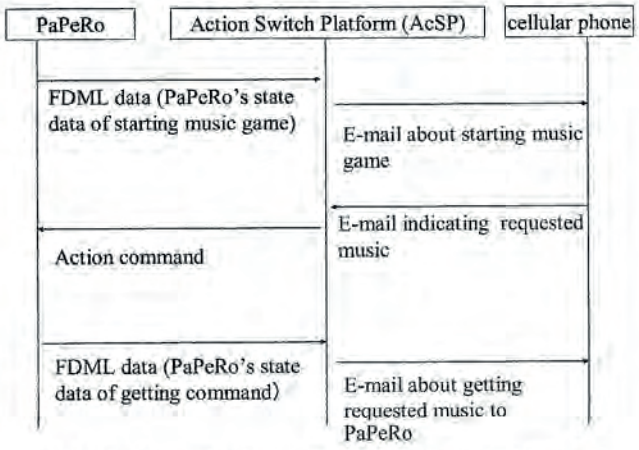


Fig.5 The example of processing flow of Meru-robo function.

We explain the key processing flow of the Meru-robo function on AcSP and PaPeRo, and the cellular phone below. Figure 5 shows an example of the processing flow of the Meru-robo function.

When the teacher triggers PaPeRo to start the music game via touch control, PaPeRo sends its FDML data to AcSP as state data corresponding to the touch command. Upon receiving the FDML data, AcSP checks the conditions of events associated with the FDML data received. If the conditions of events of "sending e-mail about starting music game to parents" are TRUE, AcSP initiates the action. After the parent's e-mail reply is received by AcSP, if the body of the e-mail contains keywords registered with AcSP, AcSP receives the FDML statements that match the keywords. Upon receiving the FDML data, AcSP checks the conditions of events associated with the FDML data received. If the

Authorized licensed use limited to: Purdue University. Downloaded on April 25,2023 at 14:53:49 UTC from IEEE Xplore. Restrictions apply.

Amazon Ex. 1012.0157

conditions of events of "an action command over HTTP protocol to PaPeRo " are TRUE, AcSP does the actions.

Through this processing flow the system establishes asynchronous communication between remote parents, teachers, and children.

4. Experiment

The experiment was designed to encourage parents to communicate asynchronously with the teachers and the children by using the Meru-robo function, and to verify whether the parents could acquire useful information.

4.1. Outline of experiment

We conducted trials in two nursery facilities, Maito-project (infant education facility,Tokyo) and Makoto-aijien (day nursery, Yamanashi Prefecture). At the Maito-project site, we examined 6 infants from 5 to 9 years old and their parents while 6 infants from 5 to 6 years old and their parents in participated in the experiment at Makoto-aijien. Trials were executed for about 2 hours on selected days at each location. Each parent was assumed to have a different situation as regards the children in the facilities. At the Makoto-aijen site, parents stood around in a room apart from their children after they received an outline of the experiment; all parents shared a cellular phone provided by us. At the Maito-project site, on the other hand, parents left the site after receiving an outline of the experiment; each were given a cellular phone (same model as above). The actual experiment used the flow in which the parent's demands were collected over a certain period and presented to the children over a 2 hour period. The period when the child care was smoothly advanced to the nursery teachers was selected. The parents answered a questionnaire after the experiment. Moreover, we collected verbal comments from the parents and the teachers.

4.2. Flows of teachers and parents during the experiment

We now explain the flow of play actions observed during the trials. Figure 6 shows the example of paying attention to the behavior of the children and parents. In this experiment, we provide four pairs of keywords and FDML data as music type requests, and five pairs of keywords and FDML data to represent the types of impression from parents.

Each PaPeRo trial began with the children receiving some instruction from the teacher. When PaPeRo was turned on, it sent an FDML statement containing its state. Figure 7 shows children with PaPeRo. AcSP then sent e-mails to the parents asking them to nominate the tune to be played. The parent replied with an e-mail containing the name of the tune. Figure 8 shows a typical e-mail from a parent and display of a movie file on the cellular phone. When the parent's e-mail was received at AcSP, AcSP sent an action command over HTTP protocol to PaPeRo.

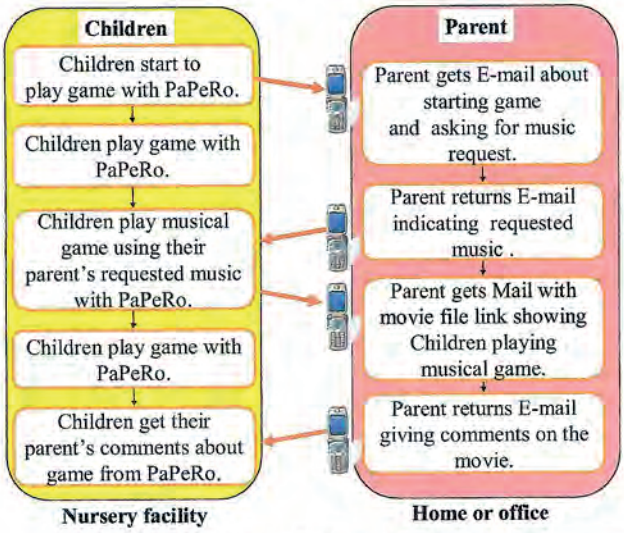


Fig.6 The example of user's flow chart.



Fig.7 Children playing with PaPeRo.

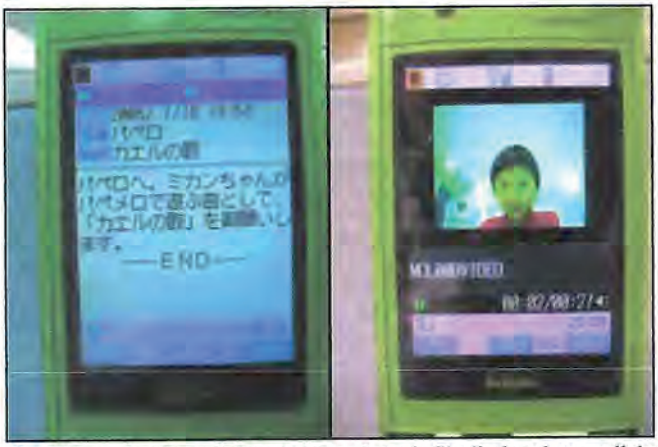


Fig. 8 Example of e-mail from parent and movie file displayed on a cellular phone. The e-mail states "To PaPeRo. my request is for the song of frogs to played in the musical game of the children with PaPeRo".

Authorized licensed use limited to: Purdue University. Downloaded on April 25,2023 at 14:53:49 UTC from IEEE Xplore. Restrictions apply.

Amazon Ex. 1012.0158

Upon receiving this instruction, PaPeRo played the tune and invited the child to sing along. While PaPeRo sang to the child, PaPeRo took a movie of the child with its eye camera. The video captured at this time was transmitted to the parents as mail with a movie file link.

In each trial, the teacher let the children play with the robot for about 20 minutes. This gave the system enough time to send the requests and collect and process the parents' requests. Afterwards, after the impression from parents is sent via mail to AcSP, the impression from parents is presented to children by PaPeRo after the teacher's approval.

4.3. Trial results

TABLE 1
QUESTIONNIRE RESULTS FROM PARENTS

Questionnaire Items	Answer	total
Could you operate the cellular phone and make e-mail with keywords ?	YES	11
	NO	1
Could you understand the appearance of your child from the movie on e-mail link ?	YES	9
	NO	3

Of the twelve parents who used the Meru-robo function, eleven were able to acquire useful information about their child. One parent was unable to reply to a teacher's e-mail asking for the parent's request. Since this parent was in a subway train, the cellular phone connection was poor which lead to a failure to receive the e-mail.

Table 1 shows the questionnaire responses collected from the parents. From the responses, the following comments were obtained. Eleven of them could operate the cellular phone and make e-mail containing the desired keywords easily, because they often wrote e-mails of the same length. Nine of them were glad to be able to visually confirm their child's play and expression and the work performed in response to their requests by watching the movie accessed via the e-mail link. Most watched their child's appearance during spare moments such as a coffee break. Three parents indicated that they wanted to confirm not only their child's play and expression, but also their child's play with other children, an overall view. With regard to the impressions of the children, gathered from the parents, the children were very glad to receive their parents' requests and comments. In the limited situation of children playing in a nursery for this experiment, this system did not hinder the activities of the nursery teachers in terms of starting commands from parents in compliance with directions given from the teacher.

In future work, we need to improve the method of making the movie contents available to the parents and the method used by the teacher in giving directions to PaPeRo. We also need to extend the kinds of requests that the parents can issue for using this system.

5. CONCLUSON

In this paper, we proposed the concept of a robotic system that offers asynchronous communication between remote parents, nursery teachers, and the children themselves. To confirm our concept, we developed Meru-robo Digital Log system, a prototype network-based interactive child watch system that lets parents participate in day-care activities. We implemented this system in two childcare facilities and conducted a trial with the cooperation of 12 families. In the limited situation of children playing in a nursery, we found that parents in remote locations could get detailed information about their children's activities and expressions, when desired, via asynchronous communication.

ACKNOWLEDGMENT

We are grateful to two nursery facilities : Maito-project (Infant education facilities,Tokyo) and Makoto-aijien (day nursery, Yamanashi Prefecture) for cooperating with our research.

REFERENCES

[1] K. Wada, et al. "Effects of Robot -assisted Activity for Elderly People and Nurses at a Day Service Center," Proceedings of the IEEE, Volume 92, Issue 11, Nov. pp.1780 – 1788, 2004.

[2] H. ishiguro, et al. "Robovie: an interactive humanoid robot," Int. J. Industrial Robotics, vol.28, no.6, pp.498-503,2001

[3] Y. Fujita, "Personal Robot PaPeRo, "Journal of Robotics & Mechatronics, Vol.14 No.1, pp.60-63, 2002.

[4] K. dautenhahn, et al. "socially intelligent agents: creating relationships with computers and robot, " springer, 2002.

[5] S. Oyama, et al. "Development of a Mobile Robot Which Embodies a Remote Instructor, " IEEE/RSJ International Conference on Intelligent Robots and Systems, 2000, pp. 1014-1019.

[6] N.Y. Chong, et al. "Development of a Multi-telerobot System for Remote Collaboration," Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, 200, pp.1002-1007.

[7] T. Machino, et al. "Remote-Collaboration Sysytem Using Mobile Robot with Camera and Projector, " Proc. IEEE Int. Conf. on Robotics and Automation, 2006, pp. 4063- 4068.

[8] http://plusd.itmedia.co.jp/broadband/0304/08/lp16.html (in Japanese)

[9] http://www.docomo.biz/html/casestudy/kansai/024.html?_cA=site (in Japanese)

[10] J. Nakayama, et al. "Information Sharing Platform and application to remote monitoring and engineering for manufacturing industry, " JIASC2003, 2003. (in Japanese)

[11] N. Katafuchi, et al. "A Prototype System of Remote-monitoring Assistance Using Logical Sensor, " The 23th Annual Conference of the Robotics Society of Japan, 2005. (in Japanese)

Authorized licensed use limited to: Purdue University. Downloaded on April 25,2023 at 14:53:49 UTC from IEEE Xplore. Restrictions apply.

Amazon Ex. 1012.0159

Appendix A-9

PURDUE UNIVERSITY

Libraries and School of Information Studies

The following pages contain materials you requested from Interlibrary Loan via ILLiad.

As these are typically delivered without staff intervention, please report any problems with the copy to the Interlibrary Loan department as soon as possible. You may reply to the delivery email (ill@purdue.edu) or call us at 765-494-2800. In these cases, there is no need to re-request the article.

Need assistance with your research? Ask a Librarian

Copyright Statement

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted materials. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research".

Amazon Ex. 1012.0161

ExLibris RapidILL

Rapid #: -20676220

CROSS REF ID: **1708589**

LENDER: **ITD (University of Texas at Dallas) :: Eugene McDermott Library**

BORROWER: **IPL (Purdue University) :: Main Library**

TYPE:	Book Chapter
BOOK TITLE:	SICE Annual Conference 2007
USER BOOK TITLE:	SICE Annual Conference 2007
CHAPTER TITLE:	Spatial memory for collaboration support system in intelligent space -User identification based on spatial memory-
BOOK AUTHOR:	Niitsuma, Niitsuma M.
EDITION:	
VOLUME:	
PUBLISHER:	
YEAR:	2007
PAGES:	3073-3079
ISBN:	9784907764272
LCCN:	
OCLC #:	

Processed by RapidX: 4/25/2023 9:38:54 AM

This material may be protected by copyright law (Title 17 U.S. Code)

SICE Annual Conference 2007
Sept. 17-20, 2007, Kagawa University, Japan

Spatial Memory for Collaboration Support System in Intelligent Space – User Identification Based on Spatial Memory –

Mihoko Niitsuma[1,2] and Hideki Hashimoto[1]

[1]Institute of Industrial Science, the University of Tokyo, Tokyo, Japan
(Tel : +81-3-5452-6258; E-mail: niitsuma@hlab.iis.u-tokyo.ac.jp)
[2]Research Fellow of the Japan Society for the Promotion of Science

Abstract: This paper presents an implementation of the spatial memory system for supporting collaboration activity in Intelligent Space. In this paper, a design of the new spatial memory system is described based on three requirements for a collaboration system: (1) sharing information for multi-users and multi-terminals, (2) access privileges for SKTs and (3) user identification. User identification method is one of the most important functions for sharing information systems. Therefore, this paper proposes a new user identification method by using the spatial memory system. More specifically, users create their own key-Spatial-Knowledge-Tags (SKTs) as passwords to use the spatial memory system. The new user identification method is evaluated through user experiments. Experimental results show high effectiveness of the method.

Keywords: Spatial memory, Intelligent Space, human interface, user identification, collaboration support system

1. INTRODUCTION

This paper presents an implementation of the spatial memory system for supporting collaboration activity in Intelligent Space [1]. The spatial memory system enables users to store computerized information into the real world by assigning a three-dimensional position to the information. The system also allows them to retrieve the information by directly indicating the points using their own hands [2]. In the spatial memory system, what associates computerized information with a three-dimensional position is called "Spatial- Knowledge-Tag (SKT)". The user actively creates SKTs based on his/her need. As a result, embedded SKTs in a working environment correspond to activity histories of the user [3].

When using the spatial memory for collaboration, history of activities will be useful to check the progress and facilitate it. In addition, the spatial memory can deliver useful information based on the spatial feature. The spatial feature means remarkable areas which are obtained by observation of users' activities [4]. When using the spatial memory for remote collaboration, the users, including remote users, can share information based on the real world. It is important and efficient for the users to share information and the view of a working area as noted in [5].

In order to realize the services supporting collaboration activity, the spatial memory system is re-designed based on three requirements: (1) Sharing information for multi-users and multi-terminals, (2) Access privileges for SKTs and (3) User identification. User identification method is one of the most important functions for sharing information systems. Therefore, this paper proposes a new user identification method by using the spatial memory system. More specifically, users create their own key-Spatial-Knowledge-Tags (SKTs) as a password required to use the spatial memory system.

The rest of the paper is organized as follows: Section 2 introduces the concept of the spatial memory system. Section 3 and 4 describe a design and an implementation of the new spatial memory. In Section 5, we propose a new user identification method by using the spatial memory system. The new user identification method is evaluated through user experiments. The last section concludes the paper.

2. SPATIAL MEMORY

The spatial memory enables humans to store computerized information such as digital files and commands into the real world by assigning three-dimensional position as the memory address. Humans can retrieve and store such information by directly indicating the point using their own body, e.g. user's hand and user's head. That's why a point on user's body is called a human indicator. Fig. 1 shows a schematic concept of the spatial memory. The spatial memory system has the following advantages to achieve intuitively and instantaneously access and store computerized information. First, the users are able to arrange computerized information at a suitable location using their own eyes and own body action. While working in the real world, human can obtain environmental information such as arrangements of equipment including desks, file cabinets and so on. Consequently, the users can arrange information in a much easier way and memorize the whereabouts by referring to such environmental information. Also their spatial cognition capabilities such as the motion sense, are utilized for memorizing the whereabouts, and will prompt the users to recall them. Second, the users are able to store computerized information and get access to stored information without interrupting their activities to search for them. The spatial memory adopts an indication of a human body as "store" and "access" operation methods in order to achieve an intuitive and instantaneous access method that anyone can apply.

In Fig. 1, small squares represent virtual tags called Spatial-Knowledge-Tags (SKT), which associate stored

Amazon Ex. 1012.0163
Authorized licensed use limited to: Univ of Texas at Dallas. Downloaded on April 25,2023 at 15:36:41 UTC from IEEE Xplore. Restrictions apply. © 2007 SICE

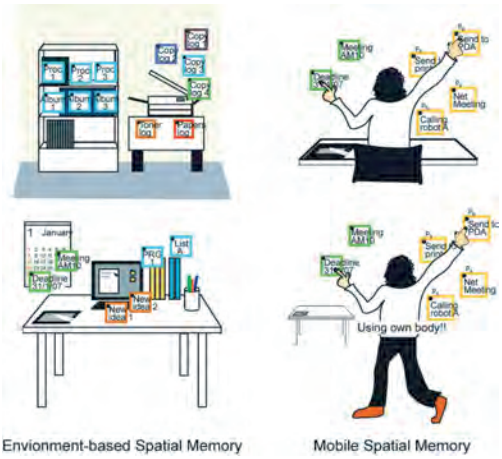


Fig. 1 Schematic concept of the spatial memory

information with a spatial location. Each SKT has three important parameters, namely stored computerized information, a three-dimensional position as a memory address, and a size of an accessible region. An accessible region is necessary for human to smoothly retrieve an SKT by using the human indicator, because he/she can not indicate the exact spatial location. The size of an accessible region should be determined according to accuracy of the human indicator [6] and the type of the human indicator. We improved the system to utilize it for a wide variety of situations.

3. DESIGN OF THE SPATIAL MEMORY SYSTEM AS COLLABORATION SUPPORT SYSTEM

In this section, a design of the spatial memory system to apply for collaboration works is described. We focus on sharing information in collaboration works. Information includes not only data and knowledge which are used for work, but also progress and status of the work. When using the spatial memory system, workers can share spatially stored information as well as the real environment as shown in Fig. 2. This kind of information sharing method for support of remote collaboration has been proposed in [5].

Considering the collaborative situations at office, laboratory, hospital and factory, information to utilize for

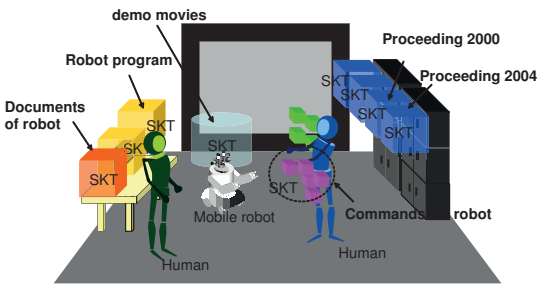


Fig. 2 Illustration of the spatial memory

works involves both of information relating to personal affairs and relating to group affairs. In order to realize sharing information with security, the spatial memory system should have three functions as follows:

1. Sharing information for multi-users and multi-terminals

The spatial memory system should be available to multi-users and multi-terminals. Multi-terminals means that users are able to retrieve spatially stored information by using multiple terminals such as personal computers and PDAs. SKTs are arranged according to setting of the environments such as the whereabouts of desks, whiteboards and meeting tables and so on. Therefore, SKTs depend on the environment, but not on terminals. In order to retrieve the information independent of the terminals, the spatial memory system provides a function to share the information.

2. Access privileges for SKTs

Ensuring security is very important issue of the spatial memory system. Access privileges for SKTs should be set so that the users can utilize the system for both of information relating to the personal affairs and relating to group affairs.

3. User identification

To set access privileges, user identification is necessary. The spatial memory has been designed to realize easily and quickly storing and retrieving method. Therefore, the system should not deteriorate the operability. In this paper, a method of user identification using the spatial memory is proposed. A set of SKTs as an authentication key is introduced. By storing and retrieving a set of SKTs, login authentication is performed. The set of SKTs is called “key-SKTs”. An arrangement of a set of key-SKTs depends on users individually. Thus, it is expected that users other than the owner hardly find his/her key-SKTs.

In next section, implementation of each function is explained.

4. IMPLEMENTATION OF THE SPATIAL MEMORY SYSTEM

The spatial memory system consists of four parts as follows: (a) Measurement unit of human indicator, (b) Spatial memory core unit, (c) Spatial memory input unit and (d) Spatial memory output unit, as shown in Fig. 3.

4.1 Measurement unit of human indicator

Measurement unit of human indicator is implemented by using an ultrasonic 3D positioning system [7] as shown in Fig. 4. The positioning system is able to measure three-dimensional positions of ultrasonic transmitters. The transmitters are activated based on time division multiple access method. Therefore, sampling frequency depends on the number of transmitters. For applying to the spatial memory, sampling frequency is important to measure motion of human indicator. From the view point of the measurement of human indicator motion, the number of transmitters is fixed to four. The position of trans-

Amazon Ex. 1012.0164
Authorized licensed use limited to: Univ of Texas at Dallas. Downloaded on April 25,2023 at 15:36:41 UTC from IEEE Xplore. Restrictions apply.

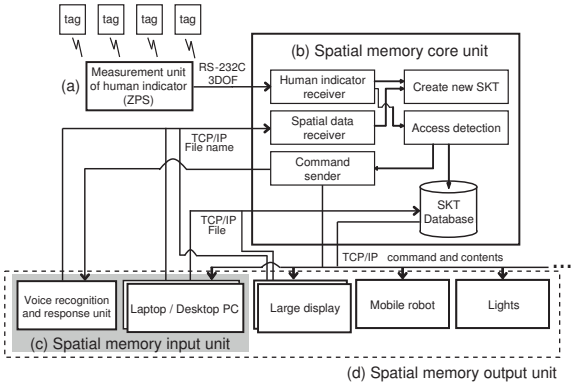


Fig. 3 System configuration

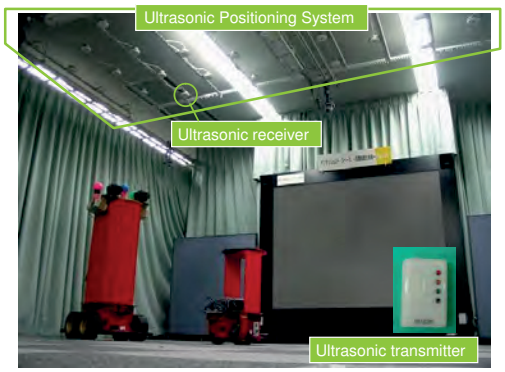


Fig. 4 Ultrasonic 3D positioning system as the measurement unit of human action

mitters is regarded as human indicator. Based on the observation of each transmitter, indication action of each user is obtained [3].

4.2 Spatial memory core unit

The spatial memory core unit implements core functions such as determination of indication action, creating SKTs, retrieving and delivering SKTs. Therefore, this unit is the only unit connected with all other units. The spatial memory core unit allows multiple spatial memory input/output units to connect with it asynchronously. When receiving information from the spatial memory input units, the core unit assigns a unique ID number to the information and creates a new SKT. SKT data structure is shown in Table. 1.

Table 1 Data structure of SKT

SKT data
ID
Spatial memory address (x, y, z)
Path of spatial memory (contents)
Size of accessible region
Number of access times
Type of contents
User name
Time stamp

When delivering an SKT to a spatial memory output unit, the spatial memory core unit chooses the nearest spatial memory output unit from the human indicator which is used to access the SKT. In addition to distance, functions of spatial memory output units should be considered, because different functions are included. For example, not only computers but also mobile robots are available as spatial memory output units. In order to discriminate functions of spatial memory output units, each output unit has a label to describe attribute identifier. The label is notified by the spatial memory output unit to the core unit. The core unit chooses a spatial memory output unit to deliver an SKT according to labels of output units, distance between a user and output units, and file extensions of information.

Fig. 5 shows principal messages between units and the sequence. The notations such as "(c1)" or "(c2)" represent the operation status and the order of processes. For example, when a spatial memory input unit sends information to the spatial memory core unit, firstly the input unit sends the message (s1) to the core unit. Then, the core unit assigns a unique ID number to the information and creates an SKT. After that, the core unit delivers the message (s2) to a spatial memory output unit. The spatial memory output unit obtains the ID number and the information. The output unit executes the information and store the information into an SKT database by assigning the ID number as the process (s3).

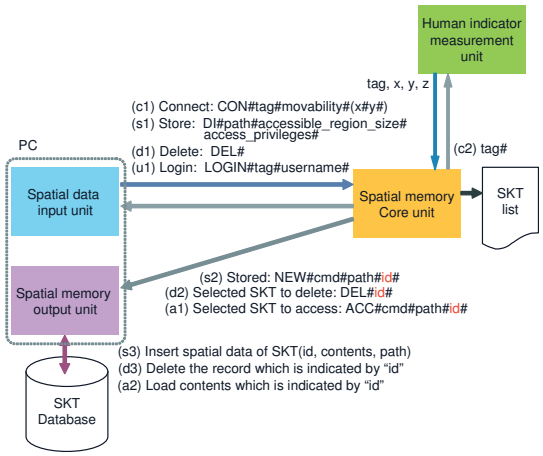


Fig. 5 Connection and messages among all units

4.3 Spatial memory input unit

Fig. 6 shows the interface of the spatial memory input unit. The interface plays a role of the front-end of the spatial memory system for users. In addition to the shown graphical interface, voice input method is also available.

When starting to use the spatial memory, user identification is required. Users choose a login mode either as anonymous or as specified user. If a user choose specified user mode, the user inputs user's name into the text field as shown in Fig. 6 (u1) area. Then, the user arranges three key-SKTs as his or her password. When the user is able to access all SKTs under specific conditions, the spa-

Amazon Ex. 1012.0165
Authorized licensed use limited to: Univ of Texas at Dallas. Downloaded on April 25,2023 at 15:36:41 UTC from IEEE Xplore. Restrictions apply.

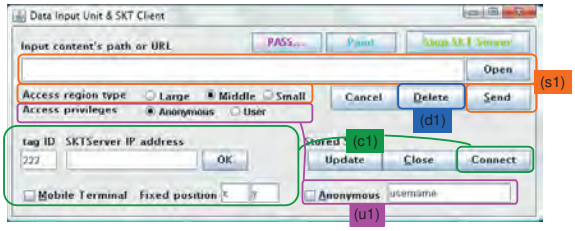


Fig. 6 Spatial data input unit –main window

tial memory system becomes available. The specific conditions and the user identification method are discussed later.

By using the interface, a user chooses or draws information which is stored as the spatial memory. After preparing the information, the user sets access privileges for the SKT. If the user sets access privileges as "anonymous" for SKTs, the SKTs can be retrieved by any users. On the other hand, if the user sets access privileges as "user" for SKTs, the SKTs can be retrieved only by the user who stored the information. Finally, the input unit sends the path of the information, accessible region size and access privileges to the core unit. Multiple input units are able to connect with the core unit asynchronously.

4.4 Spatial memory output unit

Spatial memory output unit executes the information of accessed SKTs when receiving it from the core unit. Also, the output unit operates an SKT database according to the messages from the core unit as shown in Fig. 5.

To share SKTs which are stored by different users and different input units, an architecture of sharing SKTs is needed. Namely, an SKT database which gathers the stored information (contents of SKTs) is introduced.

Contents of SKTs should be stored after assigned a unique ID number to an SKT. Therefore, a spatial memory input unit is not able to store contents into the database, because the contents have not been assigned ID.

Based on the discussion, two types of the architectures can be considered as shown in Fig. 7 (a) and (b): (a) the spatial memory core unit operates the database, (b) each spatial memory output unit operates the database.

From the view points of communications traffic and processing time, a desirable architecture is (b). If the architecture (a) is applied, input units are required to send

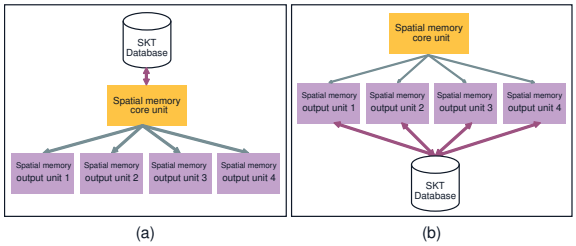


Fig. 7 Configuration of the SKT database

not only the path of contents but also data of the contents to the core unit. It is not efficient for the core unit to operate the database to deliver accessed SKTs to multiple output units. The functions of determination of human indicator and retrieving SKTs are more important for the core unit.

Therefore, if the architecture (b) is applied to share the information, communication load to retrieve information can be reduced.

5. EXPERIMENTS

5.1 User identification based on the spatial memory

To set access privileges, user identification is necessary. Requirements for a user identification method are described as follows:

1. Users are able to use the function easily.
2. The spatial memory system should prevent impersonating a user by someone else.
3. It should not make the operability of the spatial memory worse.

In order to satisfy the three requirements, this paper proposes a new user identification method using the spatial memory. The procedure of the method is described as follows: First of all, users of the spatial memory create their own "key-SKTs" as the password to use the spatial memory system. Then, when the users are able to access the all key-SKTs, the spatial memory becomes available for the users.

The spatial memory is designed to realize easily and quickly storing and retrieving information, and smoothly recalling the whereabouts of the information [2]. By using the same method of the spatial memory, the users can operate the user identification process in the same way as when using the usual spatial memory. Therefore, the method based on the spatial memory is able to satisfy the two requirements (1 and 3).

What is expected to satisfy the second requirement is the difference of actual motion among individuals. The arrangements of key-SKTs deeply depend on users, because the whereabouts of SKTs are chosen by each user based on the user's subjective visual recognition and kinesthesis. Therefore, we can assume that other people can hardly find the set of someone's key-SKTs.

Experiment to verify the assumption is performed.

5.1.1 Purpose of the experiment

The purpose of the experiment is to investigate possibility of the user identification method based on the spatial memory. Namely, we investigate if users are able to retrieve other's key-SKTs or not, if specific conditions to distinguish a target user from others can be set or not.

5.1.2 Experimental setting

The experiment is carried out by three user subjects (Subject A, Subject B and Subject C). The procedure is shown as follows:

(1) A subject creates three key-SKTs freely, while other subjects don't see the scene.

Amazon Ex. 1012.0166
Authorized licensed use limited to: Univ of Texas at Dallas. Downloaded on April 25,2023 at 15:36:41 UTC from IEEE Xplore. Restrictions apply.

(2) The subject tries to access the key-SKTs, while other subjects see the scene.
(3) The other subjects try to access the stored key-SKTs, while other two subjects see the scene.
(4) Finally, the subject who created the key-SKTs tries to access them again.

Conditions for the experiments:

Other subjects can't see the process (1), but they can see the process (2). The condition is set on the assumption that the arrangements are visually known.

The process (3) is carried out by one at a time. Therefore, another subject can see the scene twice.

The process (4) is carried out to confirm that the subject who arranged key-SKTs can retrieve them correctly.

Conditions for successful access:

When accessing an SKT, an accessible region is set for each SKT [2]. In this experiment, the size of an accessible region is defined as a sphere with radius 10 cm. When the distance between indicated position and a position of an SKT is less than 10 cm, the access to the SKT succeeds.

Time restriction to finish the processes (2) and (3) is not specified. Subjects can quit the processes when they want. And if subjects can find the all key-SKTs, each experiment will be finished.

Time restriction between accesses is also not specified. Therefore, subjects can try to access again and again.

The only condition for complete successful access is that subjects have to access all key-SKTs by following the order arranged. However, it is easy for subjects to know the order, because they can see the correct order through the process (2).

5.1.3 Experimental results

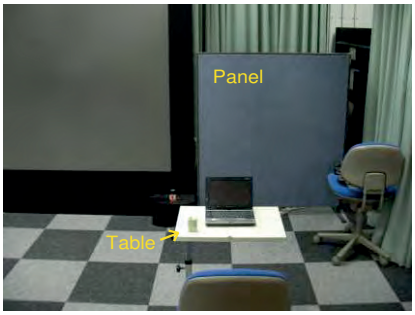
Above figure in Fig. 8 shows the experimental environment, and figures below depict the arrangements of key-SKTs which are created by each subject. Table 2 shows the order of subjects' performances. The first and last subject creates his key-SKTs, called "target subject".

Figs. 9–11 present experimental results. Filled squares (■) show successful accesses by other subjects. Plus-marks (+) and x-marks (×) represent indication actions of target subjects and other subjects respectively. Results of target subjects are shown by using their second (last) performances respectively.

As shown the results, the number of indication actions by other subjects is overwhelmingly more than the number of target subjects' actions. The case that all key-SKTs

Table 2 The order of subjects' performances

	Order of the subjects' performances
<1>	Subject A → Subject B → Subject C → Subject A
<2>	Subject B → Subject C → Subject A → Subject B
<3>	Subject C → Subject A → Subject B → Subject C



Experimental environment

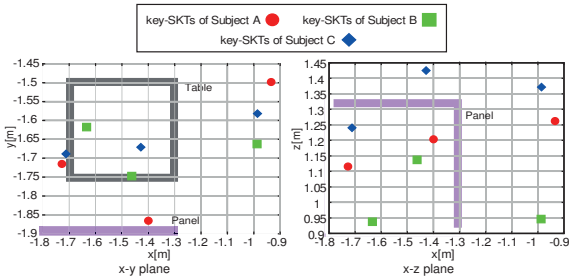


Fig. 8 Arrangements of key-SKTs

were accessed by other subjects was only the case <2> as shown in Fig. 10. Other than the case <2>, completely successful accesses were not achieved.

When we focus on time interval between accesses and total time to complete accesses, time interval took ten seconds at minimum in the case of Fig. 10 (i), and total time took twenty seconds in the same case. In other cases, for example case Fig. 9, both of Subject B and C were not able to access a key-SKT even though they tried for more than two minutes. On the other hand, all subjects were able to access their own key-SKTs in total ten seconds and with less than five indication actions.

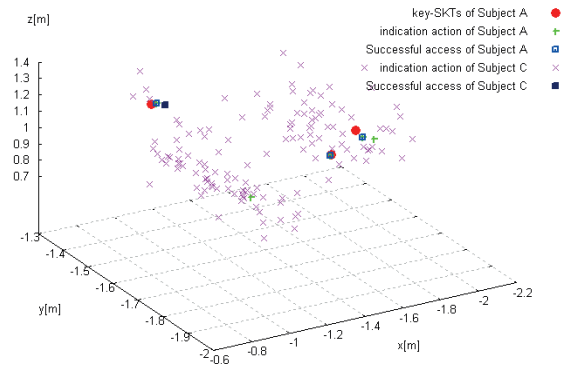
Therefore, we can specify limiting conditions for the user identification by using the parameters such as the order of access, total accessing time and the number of indication actions. Interval time can be also included, because each user will get better with increasing the number of usage of the spatial memory. Moreover, if users can indicate positions of SKTs more precisely than 10 cm, the size of an accessible region can be smaller. Table. 3 shows the distance between indicated positions and each key-SKT. All results are smaller than 10 cm. Therefore, the size of an accessible region will also be redefined according to the number of access times.

5.1.4 Discussion on the experiment

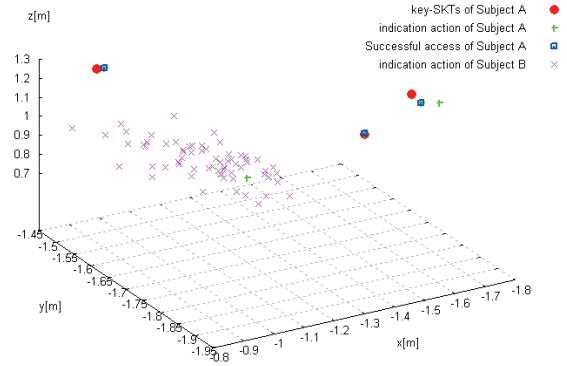
We discuss the results of the case <2> as shown in Fig. 10 which is the only case that other subjects were able to complete the accesses.

The difference between <2> and <1>, <3> is that the key-SKTs of the case <2> are arranged near objects such as a table and a laptop computer. Therefore, it would appear that other subjects could catch the positions easily by referring to the objects. That's why the key-SKTs

Amazon Ex. 1012.0167
Authorized licensed use limited to: Univ of Texas at Dallas. Downloaded on April 25,2023 at 15:36:41 UTC from IEEE Xplore. Restrictions apply.



(i) Results of Subject C → Subject A



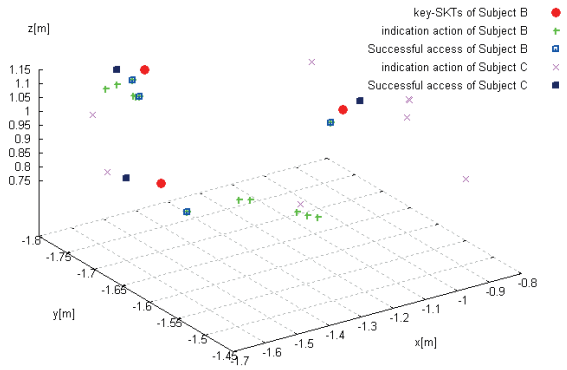
(ii) Results of Subject B → Subject A

Fig. 9 Results of the performance <1>

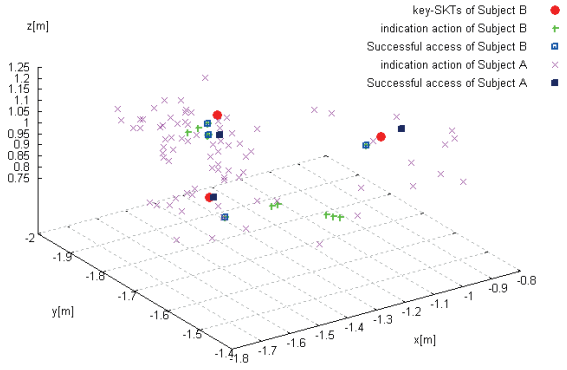
Table 3 Distance between indicated position and each key-SKT

Subject	Distance [cm]
Subject A	2.7
	3.7
	2.1
Subject B	7.7
	4.4
	7.0
Subject C	5.1
	1.9
	5.7

were completely accessed. Another possible reason is considered as subjects' body size. Table. 4 shows subjects' body height and arms length. As shown in the table, arms length of Subject B and Subject C is the same. However, Subject B was not able to access key-SKTs of Subject C even though Subject C could access key-SKTs of Subject B. Thus, we can conclude that if other users can find someone's key-SKTs depends more on the arrangement than the difference of body size.



(i) Results of Subject C → Subject B



(ii) Results of Subject A → Subject B

Fig. 10 Results of the performance <2>

Table 4 Subjects' body height and arms length

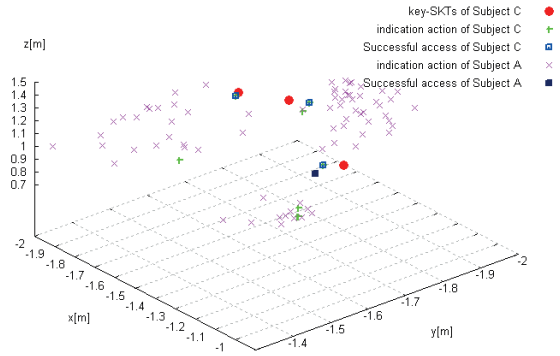
Subject	Body height [cm]	Arms length [cm]
Subject A	154.5	67
Subject B	169.9	74
Subject C	168.3	74

5.1.5 Discussion on comparison with other methods

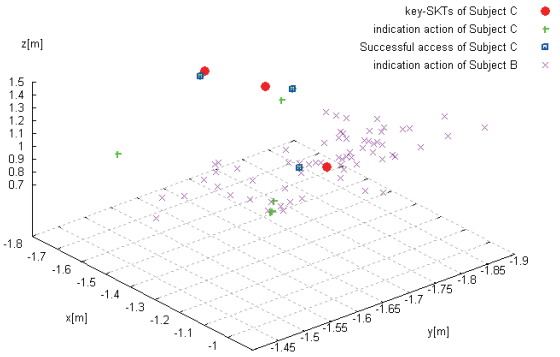
The spatial memory based user identification is compared with other methods qualitatively.

- Comparison with character password based user identification
  Character password based user identification is often used for our daily life. Users should memorize each password and corresponding target systems. Therefore, when increasing the number of character based passwords, it is difficult to distinguish corresponding target from others. If, in an effort to avoid confusion of his passwords, a user uses only one password, he will take risks that other people can impersonate him for corresponding systems when coming out it. On the other hand, key-SKTs as a password for a system are created by referring to setting of the real environment. Therefore, users can recall the set of key-SKTs according to environments.
- Comparison with biometrics based user identification

Amazon Ex. 1012.0168
Authorized licensed use limited to: Univ of Texas at Dallas. Downloaded on April 25,2023 at 15:36:41 UTC from IEEE Xplore. Restrictions apply.



(i) Results of Subject A → Subject C



(ii) Results of Subject B →Subject C

Fig. 11 Results of the performance <3>

Biometrics based user identification methods share the same problem with the case of using only one password. However, it is difficult to steal biological information. Therefore, it is suitable to apply to high confidential targets. On the other hand, key-SKT is suitable for episodic use or public spaces, because it can be easily created and changed.

- Comparison with IC tag based user identification
  IC tag based user identification is often used for our daily life. It is also convenient for storing personal information. In addition, users don't need to memorize passwords to enter target systems. However, users will have several cards when using several systems. In the case of using the same card for several systems, the risk of impersonation gets higher. And it is very easy to impersonate card holders if someone gets them. On the other hand, detection of the set of key-SKTs and complete successful access are hardly achieved by other people. In addition to the high secrecy, of course, the spatial memory system can store personal information associated with the key-SKTs.

6. CONCLUSION

This paper presented the design and the implementation of the spatial memory system for supporting collaboration activity in Intelligent Space. In this paper, the requirements for the spatial memory as a collaboration support system were described as follows: (1) Sharing information for multi-users and multi-terminals, (2) Access privileges for SKTs and (3) User identification. Then, a design of the new spatial memory system based on the requirements and the implementation were explained. This paper proposed a new user identification method by using the spatial memory system. More specifically, users create their own key-SKTs as the password to use the spatial memory system, and then access all key-SKTs. The new user identification method was evaluated through user experiments. The method may appear very simple, but the results showed high secrecy. The experimental results demonstrated that other users were not able to complete accesses to all key-SKTs as well as the owner. Based on the experimental results, we described a design of the limiting conditions for the user identification from the view points of the order of access, total accessing time and the number of indication actions. The effectiveness of the limiting conditions will be verified. For future investigation, effective and practical collaboration activity to utilize the spatial memory can be considered.

REFERENCES

[1] J.-H. Lee, H.Hashimoto, "Intelligent Space - concept and contents", Advanced Robotics, Vol.16, No.3, pp.265-280, 2002.

[2] Mihoko Niitsuma, Hiroshi Hashimoto, Hideki Hashimoto, "Spatial Memory as an Aid System for Human Activity in Intelligent Space", IEEE Trans. on Industrial Electronics, vol. 54, no. 2, pp. 1122-1131, Apr. 2007.

[3] Mihoko Niitsuma and Hideki Hashimoto, "Extraction of Space-Human Activity Association for Design of Intelligent Environment", the Proc. of 2007 IEEE Int. Conf. on Robotics and Automation, pp.1814-1819, Apr. 2007.

[4] Takeshi Sasaki, Hideki Hashimoto, "Human Observation Based Mobile Robot Navigation in Intelligent Space", Proceedings of the 2006 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, pp.1044-1049, 2006.

[5] Tamotsu Machino, Yoshito Nanjo, Yoshimasa Yanagihara, Hiroaki Kawata, Satoshi Iwaki and Ken-ichiro Shimokura, "A Remote-collaboration System using a Mobile Robot with a Camera and a Projector", Proc. 2006 IEEE Int. Conf. on Robotics and Automation, pp. 4063- 4068, 2006.

[6] M.Niitsuma, H.Hashimoto, H.Hashimoto, and A.Watanabe, "Spatial Human Interface in Working Environment –Spatial-knowledge-tags to Access the Memory of Activity–", the 30th Annual Conference of the IEEE Industrial Electronics Society (IECON 04), TC5-5, November, 2004.

[7] ZPS(Zone Positioning System)–3D, Furukawa Industrial Machinery Systems CO.,LTD, http://www.furukawakk.jp/products/ZPS_1.html

Amazon Ex. 1012.0169
Authorized licensed use limited to: Univ of Texas at Dallas. Downloaded on April 25,2023 at 15:36:41 UTC from IEEE Xplore. Restrictions apply.

Appendix A-10

Screen capture completed by J L Mullins from Western University, London, Ontario, Libraries OPAC on September 12, 2023

Western Libraries

```
leader  02521nam a2200517 a 4500
001     991587373505151
005     20210215183253.0
008     060714s2006    nju     b    101 0 eng d
020     ##$a0780395050 $q(pbk)
020     ##$a0780395069
035     ##$a(CaOLU)b41038472-01ocul_uwo
035     ##$aCaOLU41038472 $9ExL
035     ##$a(OCoLC)1069510583
040     ##$aCaOLU $beng
090     ##$aTJ210.3.I582 2006 $9local
111     2#$aIEEE International Conference on Robotics and Automation $d(2006 : $cOrlando, Fla.)
245     10$a2006 IEEE International Conference on Robotics and Automation (ICRA) : $bOrlando, Florida, 15-19 May, 2006  / $c[sponsored by the IEEE Robotics and Automation Society].
246     3#$aICRA '06
246     3#$aICRA 2006
246     3#$aIEEE ICRA-2006
246     30$aIEEE International Conference on Robotics and Automation
246     3#$aRobotics and automation, 2006
246     1#$iConference theme : $aHumanitarian robotics
260     ##$a[Pitscataway, N.J.] : $bIEEE, $cc2006.
300     ##$a10 v. (lxviii, 4275, A-62 p.) : $bill. ; $c27 cm.
500     ##$a"IEEE Catalog Number: 06CH37729"
504     ##$aIncludes bibliographical references and author index.
530     ##$aAlso available via the Internet under the title: Proceedings 2006 IEEE International Conference on Robotics and Automation.
650     #0$aRobotics $vCongresses.
650     #0$aAutomation $vCongresses.
710     2#$aIEEE Robotics and Automation Society.
730     0#$aProceedings 2006 IEEE International Conference on Robotics and Automation.
730     0#$aProceedings of the 2006 IEEE International Conference on Robotics and Automation.
856     4#$zConnect to Internet resource $xIEEE $uhttp://ieeexplore.ieee.org/servlet/opac?punumber=10932
902     ##$a191105
907     ##$a.b41038472 $br      $c-
949     ##$laoffg $cv.v.1 $i39006031697214 $g1 $0.00 $o- $t0 $y.i56798866 $s-
949     ##$laoffg $cv.v.2 $i39006031697222 $g1 $0.00 $o- $t0 $y.i56798878 $s-
949     ##$laoffg $cv.v.3 $i39006031697172 $g1 $0.00 $o- $t0 $y.i5679888x $s-
949     ##$laoffg $cv.v.4 $i39006031697180 $g1 $0.00 $o- $t0 $y.i56798891 $s-
949     ##$laoffg $cv.v.5 $i39006031697198 $g1 $0.00 $o- $t0 $y.i56798908 $s-
949     ##$laoffg $cv.v.6 $i39006031697149 $g1 $0.00 $o- $t0 $y.i5679891x $s-
949     ##$laoffg $cv.v.10 $i39006031697123 $g1 $0.00 $o- $t0 $y.i56798957 $s-
949     ##$laoffg $cv.v.7 $i39006031697156 $g1 $0.00 $o- $t0 $y.i56798921 $s-
949     ##$laoffg $cv.v.8 $i39006031697164 $g1 $0.00 $o- $t0 $y.i56798933 $s-
949     ##$laoffg $cv.v.9 $i39006031697115 $g1 $0.00 $o- $t0 $y.i56798945 $s-
976     ##$aMARS $9local
976     ##$eIEEE/EM $zA $9local
998     ##$b11 $c060714 $dm $ea $f- $g0
```

Amazon Ex. 1012.0171

Appendix B-1

RC22495 (W0205-143) May 17, 2002
Computer Science

IBM Research Report

Ubiquitous Interactive Graphics

Claudio S. Pinhanez, Frederik C. Kjeldsen, Anthony Levas,
Gopal S. Pingali, Mark E. Podlaseck, Paul B. Chou

IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598



Research Division
Almaden - Austin - Beijing - Delhi - Haifa - India - T. J. Watson - Tokyo - Zurich

LIMITED DISTRIBUTION NOTICE: This report has been submitted for publication outside of IBM and will probably be copyrighted if accepted for publication. It has been issued as a Research Report for early dissemination of its contents. In view of the transfer of copyright to the outside publisher, its distribution outside of IBM prior to publication should be limited to peer communications and specific requests. After outside publication, requests should be filled only by reprints or legally obtained copies of the article (e.g. , payment of royalties). Copies may be requested from IBM T. J. Watson Research Center , P. O. Box 218, Yorktown Heights, NY 10598 USA (email: reports@us.ibm.com). Some reports are available on the internet at http://domino.watson.ibm.com/library/CyberDig.nsf/home .

Amazon Ex. 1012.0173

Ubiquitous Interactive Graphics

Claudio Pinhanez, Rick Kjeldsen, Gopal Pingali, Anthony Levas, Mark Podlaseck, Paul Chou
IBM Research, T.J. Watson
pinhanez@us.ibm.com

Abstract

This paper presents a device that creates ubiquitous interactive graphics in a real world environment without the need for pre-wiring surfaces or requiring users to carry or wear any special devices. Called the *Everywhere Displays projector*, or *ED-projector*, this device employs *steerability* as the key principle to achieve the promise of ubiquitous computing and augmented reality. The ED-projector uses an LCD projector with a steerable mirror to project graphics onto any pre-calibrated surface, while warping the projected image to correct for oblique distortion. Three different methods to determine this warping function are given. To sense user interaction with the projected image, the ED-projector employs a pan/tilt/zoom camera and a motion-based computer vision system to track the user's hand and touch-like gestures. We propose a system software structure that ties together steering, display, and sensing. Three implemented applications demonstrate the notion of ubiquitously "painting" the real world with interactive graphics, in the context of a futuristic office, an augmented assembly task, and a ubiquitous computer game. The observation of hundreds of users performing the assembly task suggests that the traditional desktop paradigm is inadequate for ubiquitous interaction and that new conceptual widgets and interaction paradigms must be developed.

Keywords: ubiquitous computing; augmented reality; gesture recognition; interactive projections; ubiquitous games.

1 INTRODUCTION

The extraordinary success of computer graphics technology in recent years has been responsible for the creation of amazing new worlds in movies, interactive computer games, and virtual reality (VR) environments. However, this very success poses the question of when the real world, where we live and work, is going to be as rich with imagery and information as movies and games today.

This paper describes technology that can be used to ubiquitously augment, correct, and enhance reality. Although the ultimate goal is to create interactive pixels anywhere in a space as seamlessly as it happens in movies (thanks to CG special effects), this particular work focus on how to "paint" interactive graphics on the surfaces of an environment without requiring any kind of fixed display devices, special wiring of surfaces or the use of encumbering headsets.

The system described here is based on a new device, the *Everywhere Displays projector* (*ED-projector*), composed of a steerable projection system and a pan/tilt camera. A prototype of the projection system has been built using a pan/tilt mirror that directs imagery generated by a standard LCD projector. To capture user interaction with the projected graphics, the images gathered by the camera are processed using computer vision techniques. In the current prototype, it is possible to determine the user's hand position in referennce to a surface and to detect simple, touch-like gestures, effectively mapping the user's hand gestures into mouse-like events.

For instance, we have developed a computer game, called "Fro…og!", where a computer generated image of a frog initially appears "sitting" on the side of a file cabinet (see Figure 1). If a player tries to catch it, the vision system detects the player's hand approaching and then shows an image of frog jumping (see Figure 1). This is followed by moving the image of the frog to another surface, for instance the top of a table by appropriately steering the mirror. All this happens in an environment under normal lighting conditions, ultimately creating the illusion that the frog jumps around the space to avoid being caught by the player.

Our goals, however, go beyond transforming the real world into a magical entertaining place. In particular, we envision the creation of seamless interfaces to devices on surface of walls and everyday objects such as tables, chairs, couches, doors, cabinets, and carpets. Currently, only very "special" surfaces such as TVs, monitors, and projected screens support the creation of computer graphics, and usually require external devices such as mice and joysticks to make them interactive. In our paradigm, any surface is regarded as generic support for interactive graphics. By projecting imagery and detecting user interaction with a camera, it is possible to transform every surface into a virtual touch-screen.

These ideas have been explored in three application prototypes



Figure 1. Trying to catch a projected frog.

Amazon Ex. 1012.0174

described in the paper. The first in an attentive/interactive office space.The second is in the context of augmenting the environment of an assembly task by projecting interactive graphics directly onto the surfaces of the objects involved. Ultimately we aim to realize Mark Weiser's paradigm of ubiquitous computing (see [1]) without having to change the fabric of the world by embedding touch-screens on every surface. The thir application is the ubiquitous computer game "Fro…og!".

This paper starts by comparing our approach and solutions with previous work, followed by a description of the technology used in the prototypes we have built. Although initial ideas on the ED-projector have been reported in a previous publication [2], this paper presents the device comprehensively including mathematical formulation, computer vision based interaction, and three implemented applications. Specifically, this paper describes the mathematical formulation of the method for correcting oblique distortion and presents two new calibration methods for the first time. Further, in this paper we describe an integrated, working computer vision system that we implemented to track hands and detect touches, as well as a way to structure the interface to application software. While [2] ellaborates only on conceptual mock-ups, this paper describes three implemented applications using the ED-projector, followed by a discussion of how novice users — in one case, more than 600 of them — have experienced and interacted with the application prototype. We conclude by examining other applications that can take special advantage of the ED-projector and discuss the challenges these new interfaces and applications create for HCI research.

2 RELATED WORK

In his seminal paper that defined the field of ubiquitous computing, Weiser [1] envisioned computers embedded everywhere — ubiquitous and transparent. In his view, interaction with computers would be accomplished both through direct mechanisms linking sensors and actuators and by the widespread presence of cheap interactive screens (referred as "pads") embedded on the surfaces of walls, furniture, and objects. Since the publication of that paper, different devices and methods to implement the functionality envisioned for Weiser's pads have been proposed. In general, those devices and methods can be divided into four categories: head-mounted, portable, embedded, and projection-based.

A good survey of the current research on head-mounted devices and their uses in ubiquitous computing is provided by Barfield and Caudell [3], describing both wearable computers and the more traditional approach of off-body computer power. We see two main challenges in this approach First, in order to really connect to the physical world, it is necessary to precisely register the position and attitude of the head of the user with the environment in real time. An examination of the state-of-art of the research in the area shows that this is currently achievable only in very constrained environments (see, for example, Raskar et al. [4]).

The second challenge with head-mounted methods for ubiquitous displays is that they do not work well in social contexts and collaborative tasks. If multiple users are looking at the same surface (for instance, a whiteboard), it is necessary to render exactly the same graphics on each user's head-mounted display simultaneously. This creates considerable difficulties in terms of accomodating different resolutions, rendering speeds, etc. of different brands and models of head-mounted displays. In practice, it is very hard to guarantee that all participants are seeing exactly the same images.

The second method being studied to create ubiquitous displays utilizes portable devices such as laptops, PDAs, and mobile phones. Like head-mounted displays, portable devices do not work well in social and collaboration tasks (except in very unusual conditions such as in multi-player games [4]), and are a nuisance to carry and power.

The third method involves embedding screens or similar interaction devices into the objects themselves, as proposed initially by Weiser [1] and more recently by Ishii [5]. However, even if screens became extremely cheap, this approach requires a widespread change in the way everyday objects are manufactured and installed. Power and network access (or wireless bandwidth) has to be provided to even the simplest object in a home or workplace. In other words, this approach requires a drastic change of the very fabric of the real world.

Our approach follows the fourth method of creating ubiquitous displays that is based on projection devices. Interestingly, the use of projectors to change and augment reality was pioneered by media artists such as Tony Oursler [6] and Michael Naimark [7]. The use of projectors to access computers have been initially proposed by Bolt [8] and Wellner [9]. Recently, more radical applications of projectors have been proposed by Morishima et al. [10] that enhance the face of an actor, and by Raskar et al. [11] to change the color and texture of real objects such as vases and architectural models. Although our system does not currently track moving surfaces as in the former example, or project onto 3D objects as the latter, it can clearly incorporate both features in future versions.

Deviceless interaction with projections was pioneered by Krueger [12], using camera-based systems to detect the users' hand and body. The fact that most vision techniques put strict constraints on lighting conditions and on the background surfaces has restricted the use of these systems to entertainment applications such as Keays and McNeil's *metaField* [13]. To avoid lighting constraints, Omojola et al. [14], Ju [15], and the RED project at Xerox Parc [16], among others, embedded sensors on the surface where the images are projected. As in the embedded approach described before, the widespread adoption of sensor-loaded surfaces requires a dramatic change in the objects of the real world. Underkoffler and Ishii [17] and Rekimoto [18] used visual tags to detect the position and the identity of the objects on a surface and to overlay the interaction based on these attributes. To avoid the need of tagging all objects in an environment as in those two cases, a possible option is the use of object recognition techniques (such as those described by Schiele and Crowley [19]). However, those techniques currently work only in very constrained situations. We believe, though, that incorporating object recognition capabilities into our system is an important direction to follow, and we plan to do so forspecific applications.

Vision-based gesture recognition has received a great deal of attention in recent years [20], but only a little of that work has attended to the problems of gesture recognition in the context of a projected display. Because the projected image drastically changes the appearance of the hand, most approaches to hand detection are based on either background subtraction or motion (frame-frame differencing). Our experience agrees with that of von-Hardenberg and Berard [21] in that background subtraction does not give reliable results in the presence of strong projected images.

Gesture-based interaction methods tend to fall into one of two catagories: either variations of a point-and-click paradigm [21, 22] or application-specific pose or motion gestures [23]. Both these approaches have disadvantages. The point-and-click approach is often not well suited to hand pointing due to limited resolution and the lack of a natural "click". On the other hand, a proliferation of application-specific gestures will make for a complex gestural interface that a user will have to learn and execute largely without feedback, as discussed by Kjeldsen in [24].

Amazon Ex. 1012.0175

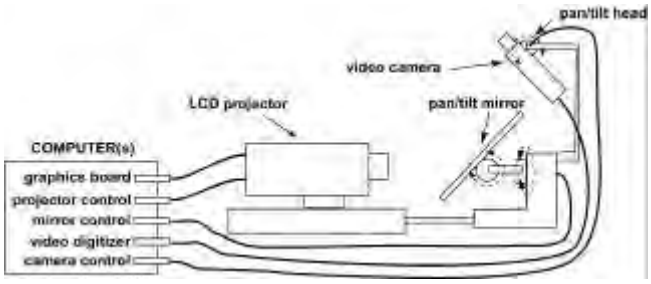


Figure 2. Diagram of the components of the ED-projector.



Figure 3. Prototype of the ED-projector: a) ED-projector without the camera mounted; b) a detail of a complete prototype, including the pan/tilt camera.

A major difference between our approach and the cited work on the use of projection-based devices to create ubiquitous interactive displays is *steerability*. By directing the projection and vision subsystems to a relevant position, it is possible to create new "interactive" surfaces in a space without any hardware installation, addition, or change. Unlike Raskar et al. [25], who use mirrors to increase the resolution of multiple-projector installations, we steer the image from a projector to realize interactive displays everywhere in the workspace. By doing so, we opportunistically use the surfaces of the objects present in the environment. We believe that steerability is a key feature that creates a real alternative to realize the promises of ubiquitous computing. However, it has constraints and technical challenges of its own, such as obstruction, oblique distortion, and surface texture interference. These issues are discussed in the next section where solutions to some of these problems are proposed.

3 THE ED-PROJECTOR

Our approach to ubiquitous interactive graphics uses a device — the *Everywhere Displays projector* (the *ED-projector*) — that can steer graphic displays onto any surface and can also sense user interaction with these steered graphical displays. The ED-projector is composed of an LCD projector, a computer-controlled pan/tilt mirror, and a pan/tilt/zoom camera. The projector is connected to the display output of a host computer that also controls the mirror and the camera (see Figure 2). The video output of the camera is sent to a digitizer board in the host computer where the imagery is processed, as described later.

Figure 3 shows a prototype of the ED-projector built with off-the-shelf components — a rotating mirror used in theatrical/discotheque lighting, a standard LCD projector, and a pan/tilt camera. The projector's light can be directed anywhere within a cone defined by approximately 60 degrees in the vertical axis and 230 degrees in the horizontal axis. The camera has a maximum vertical viewing range of 88 degrees and a horizontal range of 249 degrees, and is thus able to view the projected displays as they are steered around.

In general, a projector needs to be able to project a white pattern approximately 10 times brighter than its surroundings to create the illusion of contrast. Our current prototype employs a 3000 lumen projector, which we have found to be sufficiently bright in typical home and office lighting conditions to provide good contrast in the projected displays. Of course, the perceived brightness and constrast is heavily dependent on the color and material of the projected surface. Although our best results involve projection on white surfaces, we have also obtained good results projecting on carpets, black objects, and even translucent backlit surfaces. However, highly textured surfaces such as natural wood and surfaces with high reflectance are, in general, less suited for projected displays.

The three main issues with the ED-projector are 1) correction of the projected image to compensate for distortion due to oblique projection and surface characteristics; 2) sensing the user interactions with the projected displays; and 3) a system software structure that enables applications to easily define interactive display surfaces. These issues are discussed in detail in the rest of this section.

3.1 Correcting the Projected Image

Standard LCD projectors are designed to project light in a direction orthogonal to the projected surface. In the ED-projector, the pan/tilt mirror can deflect the light to multiple surfaces in the room, but most of the times the projected surface is oblique to the direction of projection. The result is a distorted image.

To correct the oblique projection distortion, we pre-warp the image to be projected to compensate for the distortion. It is always possible to correct the projected image, as long as the projected surface does not have a geometry that occludes the cone of projection (see [2] for details). However, the process of warping the image normally uses only a fraction of the projector's image plane, causing loss of resolution. We will discuss this issue later; first, we address the problem of determining the warping function that corrects the oblique distortion.

Although there are many different ways to determine and implement the warping function, it is important to use a method that allows fast computation. We choose to implement it using 3D computer graphics hardware to texture map the images to be displayed onto a mesh in 3D space. If the mesh is appropriately warped and/or positioned in 3D space and the virtual camera correctly positioned, the projection of the resulting rendered image is free of distortion.

The basic structure of the problem, in the case of planar surfaces, is illustrated in Figure 4. We start by defining the *application plane*, a coordinate system where images to be projected are positioned within the area of a rectangular mesh M, and the *projector image plane*, the coordinate system associated with the LCD matrix of the projector. We also define a *projected surface* to be the real world surface on to which images are projected. In addition, we have the camera image plane, the coordinate system associated with the CCD matrix of the camera.

The goal is to define a transformation H from the application plane into the projector image plane so the transformation that maps the support mesh M into the projected surface is the identity I, and, therefore, appears free of distortion from an observer orthogonal to the projected surface. Since the process of projecting the image can be modeled as a projective transformation P (as

Amazon Ex. 1012.0176

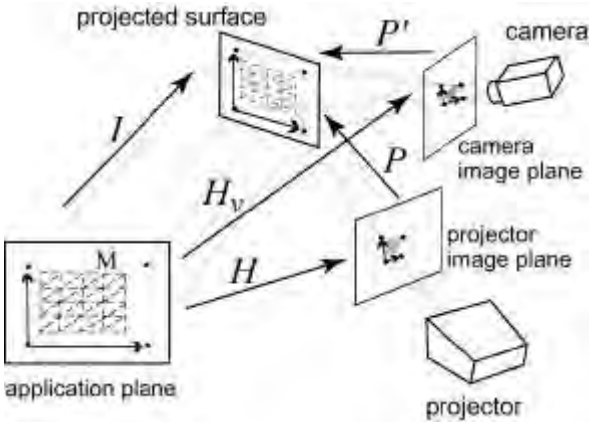


Figure 4. The transformations between the application plane, the projector's image plane, the camera's image plane and the projected surface.

observed in [26, 27]), we can represent the relationships among the planes as matrices in planar homogeneous coordinates, $I = PH$. Since H and P are projective transformations and therefore invertible, we obtain:

$$H = P^{-1}$$

Similarly if P' is the projective transformation between the camera image plane and the projected surface, and H_v is the transformation from the application plane to the camera image plane, we have $I = P'H_v$. Although there is no easy direct way to compute the projection function P, we developed three different methods to determine the actual value of H for a given surface. Of course, H and H_v are different for each planar surface in an environment.

1st method: Using an Equivalent 3D Virtual CG World

This method, first described in [2], is based on the fact that, geometrically, projection is the inverse of camera viewing. So, it is possible to model the projection by simulating the camera view by rotating, scaling, and viewing the support mesh M in 3D space.

Given the n mesh points in 2D projective space, $m_i = [m_i^x \; m_i^y \; 1]$, $1 \le i \le n$, let M be the matrix of all n points, $M = [m_1^T m_2^T \ldots m_n^T]$. This mesh can be "positioned" in the XY plane of the 3D space by multiplying it by the matrix

$$K = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad KM = \begin{bmatrix} m_1^x & m_2^x & \ldots & m_n^x \\ m_1^y & m_2^y & \ldots & m_n^y \\ 0 & 0 & \ldots & 0 \\ 1 & 1 & \ldots & 1 \end{bmatrix}$$

We can then simulate the proces of projecting the mesh as the inverse of the process of "viewing" the appropriately positioned mesh in 3D space. That is,

$$P^{-1}M = P_f R_x R_y S_x S_y T_x T_y R_z K M$$

where R_x, R_y, R_z are rotation matrixes in 3D homogeneous coordinates, S_x, S_y are scale matrixes, T_x, T_y are translation matrixes, and P_f is the 3 4 projective matrix with focus point f. Since $H = P^{-1}$, we obtain

$$H = P_f R_x R_y S_x S_y T_x T_y R_z K$$

Notice that all the matrices that compose H, with the exception of K, correspond to standard projection parameters of computer graphics boards. Therefore, to implement this method it is only necessary to determine these individual parameters, load them into the hardware, and command the graphics board to render the scene.

To determine the projection parameters for a given surface we start by texture mapping a calibration pattern to the mesh M. Starting off with default values for the projection parameters that render the mesh on the center of the projector image plane without warping (i.e., $H \approx I$), we iteratively change each of the projection parameters until the projection of the pattern on the projected surface is observed to be identical to the original mesh. We then record these parameters and reload them whenever the ED-projector is aimed at that surface. Although conceptually simple, this iterative process of determining the right projection parameters is typically very time consuming.

2nd method: Manual Determination of Corresponding Points

This and the following method are based on the fact that projective transformations can be completely determined if the corresponding coordinates of four points in each 2D projective space are known (as shown by Faugeras [28]).

In this method, four points $a_i = [a_i^x \; a_i^y \; 1]$, $i = 1,2,3,4$, are directly rendered on the projector image plane. We then iteratively move these points over the projector image space with a pointing device such as a mouse until their projections overlap the location of four known points on the projected surface, $b_i = [b_i^x \; b_i^y \; 1]$, $i = 1,2,3,4$. For greater accuracy, we can choose these four points to be the corners of the mesh M.

It is easy to see that if the projected points are aligned with the projected surface points then $PA = B$, where $A = [a_1^T \; a_2^T \; a_3^T \; a_4^T]$ and $B = [b_1^T \; b_2^T \; b_3^T \; b_4^T]$. By computing the pseudo-inverse,

$$PA = B \rightarrow PAA^T = BA^T \rightarrow P = BA^T \left(AA^T\right)^{-1}$$

Since $H = P^{-1}$, we obtain $H = AA^T (BA^T)^{-1}$. To implement this procedure, we simply construct a warped mesh M' in 3D projective space by computing $M' = KHM$, texture map the image to be projected on it, and use the graphics engine to render it in real time.

As we see, the calibration procedure of this method is quite simple and fast, requiring only a few minutes for the manual alignment of the four projected points. Although similar to the method proposed by [29], here the use of the 3D mesh M and the computer graphics hardware yields much faster rendering speeds.

3rd method: Using a Camera and a Paper Pattern

In this method, we employ the ED-projector camera so we can avoid any manual calibration procedure. In [30], Yang et al. have proposed a method to automatically determine the warping function using a projector/camera pair. Unfortunately, their approach requires precise 3D calibration of the relative attitude of the camera to the projector and, in practice, a considerable baseline distance between them. In particular, it does not work with the ED-projector that aligns the projector and the camera axis as much as possible to avoid user's obstruction of the camera's view.

Instead, we propose here a method that employs a paper pattern positioned on the surface to be calibrated so it is aligned with the desired coordinate system. No relative 3D or 2D calibration is necessary, although we assume that the camera has been positioned in such a way that it has a complete view of the projected surface. Under these conditions, we can define a projective transformation P' between the camera image and the projected surface (see Figure 4). This transformation can be determined automatically by processing an image of the pattern and extracting the camera image coordinates of the corners, $c_i = [c_i^x \; c_i^y \; 1]$, $i = 1,2,3,4$. Assuming the coordinates of the

Amazon Ex. 1012.0177

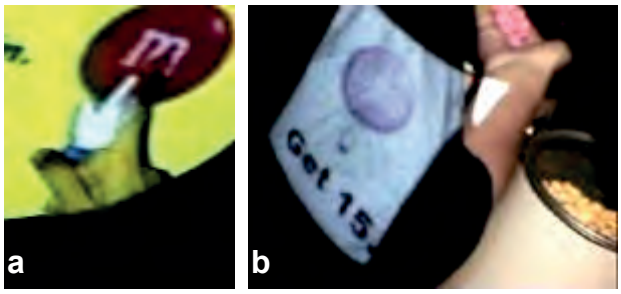


Figure 5. Difficulties faced by the vision system: a) the user's hand is partially hidden by the strong projection; b) the interactive surface is occluded by the user's shoulder as he moves to gather objects from inside a bucket.

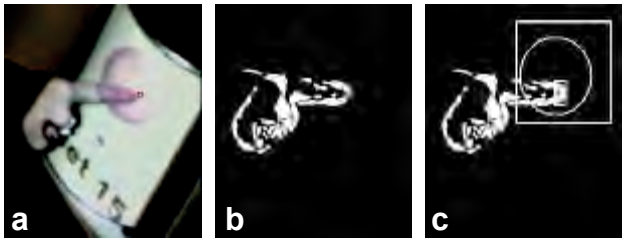


Figure 6. Detecting a touch: a) camera view of an interaction with a bucket; b) image difference data; c) overlay of search region (square), button active area (circle), and the fingertip template shown at the pointing location.

pattern corners to be identical to the mesh M in application space, $b_i = [b_i^x \; b_i^y \; 1]$, $i = 1,2,3,4$, we obtain $P'C = B$, and computing the pseudo-inverse,

$$P' = BC^T \left(CC^T\right)^{-1}$$

To complete the computation of the transformation H, it is necessary to determine the correspondence between the projector and the camera image planes. This can be done by projecting a pattern with four non-colinear points rendered on the projector image plane $d_i = [d_i^x \; d_i^y \; 1]$, $i = 1,2,3,4$, and obtaining four points on the projected surface, $e_i = [e_i^x \; e_i^y \; 1]$, $i = 1,2,3,4$. Then, by image processing the image of the pattern, it is possible to determining the coordinates of its four corners on the camera image plane, $f_i = [f_i^x \; f_i^y \; 1]$, $i = 1,2,3,4$.

Since $E = [e_1^T \; e_2^T \; e_3^T \; e_4^T]$ corresponds both to the projection of $D = [d_1^T \; d_2^T \; d_3^T \; d_4^T]$ by the projector, $PD = E$, and to the captured image of the pattern, $P'F = E$, we obtain $PD = P'F$. Computing the pseudo-inverse, we obtain $P = P'FD^T(DD^T)^{-1}$ and therefore

$$H = DD^T \left(P'FD^T\right)^{-1}$$

As in the previous method, this transformation is implemented by computing a warped mesh $M' = KHM$.

One problem with all these three distortion-correcting methods is that they project displays with resolution lower than the projector's resolution. The distortion-correction process normally fits an irregular quadrangle into the 4:3 viewing area of typical displays. A considerable amount of display area, and thus resolution, is lost in the process. We have employed standard 1024x768 XVGA projectors in our prototypes. Due to the loss of display area created by the distortion correction process, the final display resolution has approximately corresponded to VGA, i.e., 640x480 pixels.

3.2 Sensing Interaction on Any Surface

To support the ability to interact on any surface we use a pan/tilt camera feeding custom image processing and gesture recognition engines. This process has been detailed in another publication [31], but will be outlined here for completeness.

Image processing in this domain is both simplified and made more difficult by the strong light of the projected image. It is simplified in the sense that the lighting environment is very consistent because the dynamic range of the projection swamps all but major changes in ambient light. It is made more difficult because the appearance of the user's hand can be severely distorted as it passes through the projected image (see Figure 5.a).

Although parts of the hand can be hard to detect in a static image, by looking at the differences between adjacent images in the video stream, we can still detect the shape of a moving object as shown in Figure 6.a and b.The use of motion for segmentation instead of background subtraction techniques (such as [32]) also makes it easier to accommodate the variability in the projected image caused by the texture of the surface it is projected on. This motion mask forms the basis of subsequent image understanding.

To provide interaction in a flexible and consistent manner we use the notion of *interface widgets*. Each widget generates computer or application events in response to a particular user motion in a region of the image. We currently use three types of widgets: *buttons* that respond to touch-like motions; *sliders*, 1-dimensional tracking regions, and *touch pads*, which are 2-dimensional tracking regions.

These three widget types rely on tracking the user's fingertip. Convolution template matching is used to find fingertip candidates in the motion mask image described above. The candidates are evaluated using spatial clustering and domain specific heuristics, such as the distance from the user, to determine the best estimate of whether there is a fingertip in the image and where it is currently pointing. The trajectory created by this point over a sequence of video frames is examined by each widget to determine when it should trigger an event.

Touch pad widgets smooth the fingertip trajectory and map it into the touch pad coordinate system. Motion events are only generated when the fingertip is within the touch pad's borders. Sliders are simply a 1-dimensional version of touch pads.

Buttons, however, are more complex widgets. What buttons attempt to detect is when a surface in the environment is touched. Inferring the distance of the fingertip to the surface using the shadow proved to be difficult and unreliable. Fortunately, people tend to interact with real-life buttons by moving in a characteristic forward/pause/backward pattern, and this seems to carry over well to projected buttons. Our implementation of buttons takes advantage of this movementpattern.

To determine when a user has touched a button we look for a roughly linear sequence of fingertip locations that persists for 250ms or longer, with speed within a set range, that either ends or begins with a pause (i.e. the fingertip template did not find a match in the motion mask). A button generates its event when the pause in such a trajectory occurs within it (see Figure 6.c). In other words, we are looking for the user's fingertip to travel in one direction for ¼ second and stop within the button, or start moving from the button and move in a consistent direction for ¼ second. These approximations for the motion of a fingertip in the moments before and after it touches a button seem to be universal: almost every button touch either starts or ends that way. Interestingly though, many button touches only have one or the other of these motions. Of course, adjustments are needed in the case of people with tremors or other motor disabilities.

Amazon Ex. 1012.0178

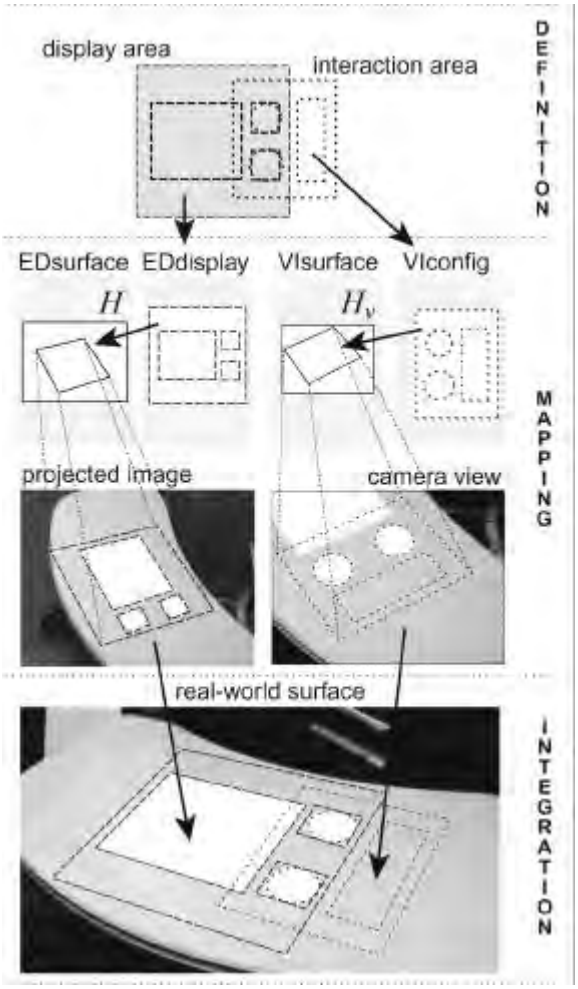


Figure 7. Process of creating an interactive surface on a table.

One of the interesting problems in this domain occurs when users occlude buttons as shown in Figure 5.b. Our experiments with users, even novices, have shown that they naturally avoid obstructing the projection if they are attending to it, but they are quite bad at staying out of the way when they are working with the physical world instead. Buttons have inherent resistance to triggering in these types of situations because any motion must pass several tests before it generates an event. It must resemble a fingertip for several frames; during that time it has to travel with a consistent direction and velocity and stop within a button. Such a sequence sometimes occur as a result of a chance combination of errors, contributing to a somewhat higher false positive rate for some buttons, particularly those which the user occludesfor a period of time rather than just passes through.

By using both the shape of the moving region as well as an analysis of its trajectory, this touch detection method gives far better accuracy than using either one alone. Because it relies directly on image data, rather than building an explicit model of the hand, it is computationally efficient enough to run at high frame rates on standard hardware.

Widgets can be combined in various ways in order to perform a complete interaction with the user. However, configurations of widgets are defined independently of the surface where they are used. That is, if an application interface using a set of buttons is moved to different surfaces, it is not necessary to define each button for each surface. Instead, the location of a widget within the camera image is determined at run time by appropriately mapping the configuration onto the camera image plane. In this case the surface refers to a set of camera parameters. These parameters, like those of the projector, have to be calibrated in advance. In particular, the calibration process has to determine the transformation H_v between the widget coordinate system and the camera image (as shown in Figure 4). Calibration of surfaces is performed using methods analogous to the 2nd and 3rd methods of projection calibration, although information about the expected location of the user, and the size of their hand, has to be obtained during an extra, special calibration step (see [31]for details).

3.3 Integrating Displaying and Sensing

The two preceding sections described the techniques and subsystems that provide the display (output) and sensing (input) functionality of the ED-projector. This section discusses the integration of these separate subsystems in a manner that affords application development and human interaction. One of our primary objectives here is to enable a specific user interaction to occur on any calibrated surface in a room without requiring the customization of this interaction for every surface. This leads to a clear separation of the abstract definition of an interaction from the actual surface upon which it occurs.

Our widgets are constructed using the standard Model View Controller (MVC) design pattern structure [33]. In our case, the projector subsystem is responsible for rendering the *view*, while the vision subsystem provides the *controller* input events. The *definition* layer of Figure 7 shows the separation of the display and interaction areas that are reponsible for the definition of the view and the interaction respectively. Display and interaction will often overlap as shown by the square overlapping buttons on the right of the *display area*. Provision is also made, however, for allowing interaction to occur in a region where nothing is displayed, as shown in the large rectangular region on the right of the *interaction area*. Notice that the display area here corresponds to the mesh M of the application plane decribed before (see also Figure 4).

Every interaction surface in a physical space is named and initially calibrated by the projection and vision subsystems yielding the transforms H and H_v respectively. H corrects the projected image while H_v provides the transformation to the camera image, for a specific surface as described in the prior sections. Figure 7 demonstrates the flow of definitions for display images and interaction data (widget data) from the *definition* layer to the *mapping* layer. During run-time, the system software automatically maps (through the warping functions H and H_v) the *definition* layer data into the projected image and the camera view associated with the current surface. The result is a projected image that generates application events when user gestures occur in the areas defined in the camera view, as shown in the *integration* layer of Figure 7.

4 PAINTING WITH INTERACTIVE GRAPHICS

To better understand how to use projected interactive graphics and how users relate to them, we have designed and implemented three application prototypes using the ED-projector. In the first application, an office space is augmented with dynamic wallpapers, visual notifications of messages, and computer access from multiple surfaces. The second prototype explores an assembly task where the instructions are projected directly onto the surfaces of the objects involved in the task. The third prototype is the "catch the frog" game described in the introduction to this paper.

Amazon Ex. 1012.0179

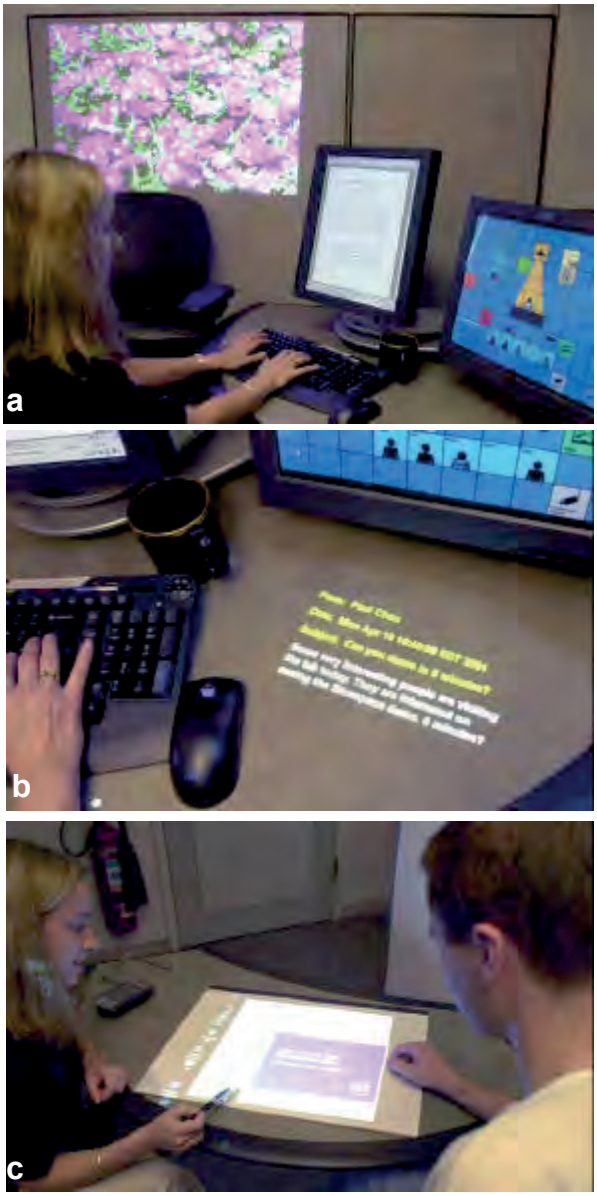


Figure 8. Uses of the ED-projector in an office application: a) to create dynamic wallpapers; b) to provide silent notification of e-mail; and c) to create a computer desktop on a working surface.

4.1 An Office Application

The first prototype application in which we have deployed the ED-projector is a futuristic personalized and context-aware workspace. The workspace has an 8' by 10' footprint and incorporates several sensors for measuring ambient lighting, temperature, humidity and noise level. Users and visitors to the workspace wear active badges that facilitate presence detection and identification. The desk chair is equipped with a pressure sensor connected to a wireless micro-controller, which detects if a person is sitting in the chair. This environment incorporates an ED-projector, strategically located to maximize coverage of the workspace. In this setting, we have used a non-interactive version of the ED-projector to facilitate greater personalization of the environment and to free its users from the confines of the desktop displays.

The presence of an ED-projector in the office tranfigures the space into an extended graphical display and enables the user to have a favorite painting on a wall, receive important notifications on surfaces that catch their attention, or use a tabletop as a display when collaborating with a colleague (see also the video figure). Figure 8.a shows an example of a dynamic wallpaper created by the ED-projector. The user can similarly have a virtual window that is connected to a live webcam looking at her favorite scenery (effectively making it a window office) or monitoring her child at home. The ED-projector can automatically switch to a neutral, less personal image when a co-worker comes to the office, based on the active badge.

Figure 8.b illustrates an urgent notification to the user appearing on a surface close to the user's current location. Here, the smart office detects that the user is seated at the desk and steers the notification right to the surface of the desk. By determining the position of the user, the system is assured that the user will see the urgent message without having to rely on the traditional sound alerts which tend to be disruptive to people working in a shared environment.

Another example of the utility of a steerable display is seen in Figure 8.c where two colleagues have created a computer desktop display on the working surface of their choice. Here, interactive capabilities, provided by a camera system, are planned to be incorporated to facilitate collaborative work.

4.2 Augmenting an Assembly Task

The ED-projector can be seen as a device for augmenting reality – a beam of light that can add information on to real world objects by highlighting them, adding text or images, flashing etc. For example, the ED-projector can show the user an object they are searching for, or indicate the procedure for an assembly task, and even where to place a part in an assembly.

We developed a prototype assembly task to highlight these themes of augmented reality and interaction anywhere. In this prototype application, which was experienced by hundreds of novice users in a major technical exhibition, the object to be assembled is a picture made of M&M's (multi-colored sugar-coated chocolates) where each M&M is regarded as a pixel of the picture.

The theme of interactivity anywhere is demonstrated several times in this prototype application (see also the video figure). Figure 9.a shows a white fabric on a table transformed into an interactive menu for color selection. The user simply touches the color of his choice to select it. Figure 9.b shows a "clickable" button appearing on a paint bucket that contains the M&M's of the selected color. Figure 9.b and c highlight the augmented reality and user assistance aspects of the system. In Figure 9.b the system highlights the bucket that contains the M&M's of the selected color, and also provides additional information on how many M&M's to pick. In Figure 9.c, the system points the exact places on the picture board where the M&M's should be placed. Figure 9.d shows another combination of interaction and augmented reality: the user moves her hand over the surface to interactively reveal the rest of the picture being built.

More than 650 people participated in this augmented assembly over 6 days and completed 4 different M&M pictures. Considering that the vision system and the projection system were integrated for the first time only in the weeks preceding the exhibition, the combined system worked remarkably well. A sample with 130 consecutive users with 621 button touch events (touching gestures or false detections) yielded correct detection of touching gestures in 81% of the events, with 12% false negatives and 7% false positives. If the bucket events are excluded from the count, the

Amazon Ex. 1012.0180

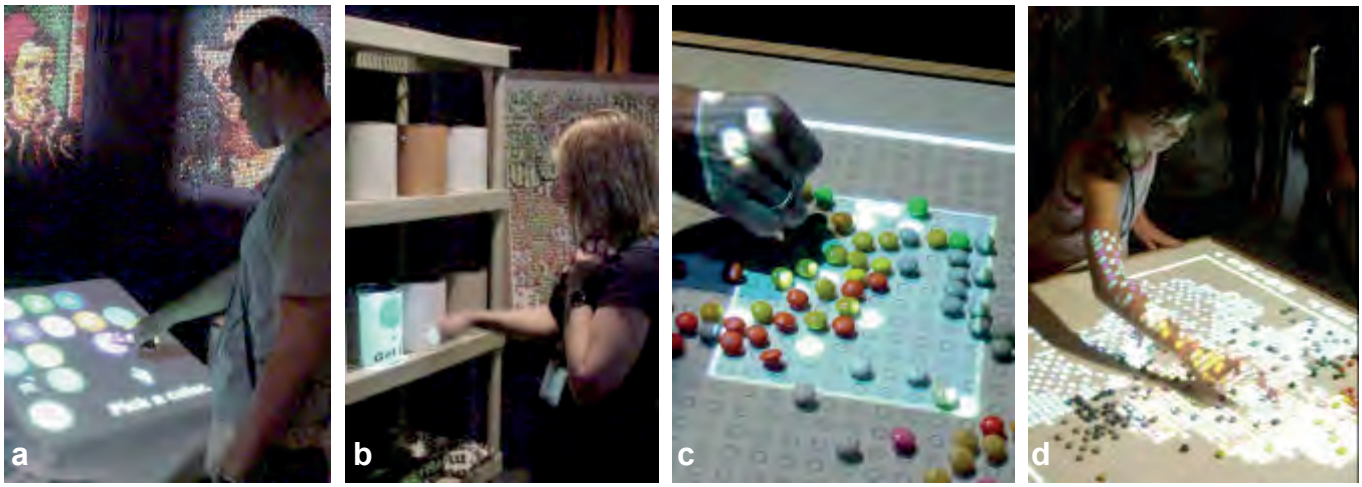


Figure 9. The M&M picture assembly task: a) choosing a color; b) "clicking" a bucket after getting the M&Ms; c) placing the M&Ms on the highlighted areas; and d) "finger painting" to reveal the complete picture.

performance exceeds 89%. A detailed account of the reasons why buckets yielded more errors is given in [31].

The experiment also revealed that users are readily able to relate to the notion of steerable interactive graphics. We also observed the value of the projection creating a shared experience between the users and the people observing the demo: in many cases, users learned the interaction process by first watching other users; and frequently, observers helped users suggesting what to do next. As mentioned before, a major advantage of projection over goggles is precisely this "social" nature of projection.

An interesting observation is that our users were less hesitant to interact with the fabric-covered interactive color selection menu (Figure 9.a) and the M&M placement board (Figure 9.c,d) than with the paint buckets (Figure 9.b). When they went to a bucket to pick up M&M's, many of them did not realize that they had to click the big "done" button on it (see also Figure 6.a). We had expected that, since this was the third occasion for touch interaction in the demo, users would quite naturally press the button on the bucket as instructed. Most of the time, this was not the case. To help the users in that situation we often gave them verbal instructions such as "Click the bucket," aiming to make clearer that there was, in fact, a button on the bucket.

4.3 "Fro…og!"

The ED-projector's unique characteristic of being able to quickly move the projected image to different surfaces is being explored in a computer game for young children. In this game, called "Fro…og!", an image of a frog is projected on a surface. Whenever the child tries to touch the frog, the camera detects her approach, and the frog "jumps" to another surface, that is, the projector moves the image to a wall or the top a chair.

Figure 1 shows two pictures of a user playing with the current prototype of "Fro…og!". The basic idea is to define surfaces in the environment where the animated frog can jump to. Whenever the user's hand approaches the frog, the next surface is randomly chosen, and an image of the frog jumping in that direction is briefly shown. After landing on the new surface, a simple animation of the frog jumping up and down is projected, so it becomes easier to find the frog in the environment.

In the next development phase of this game we plan to integrate it to a system that tracks the position of multiple children in the room, so the frog can jump to surfaces which are far from most players. Tracking is also particularly important to avoid projecting the frog directly onto the players which, in this scenario, can easily break the illusion of the game.

5 DISCUSSION

Having shown how we implemented a working prototype of an ED-projector and deployed it in three different applications, it is interesting to explore other possible applications and scenarios, and discuss how to design efficient interfaces for them.

5.1 Other Applications

We see two classes of applications of the ED-projector to create ubiquitous interactive graphics. The first involves the creation of computer interfaces anywhere, while the second deals with augmented reality in a variety of public and private spaces.

In the first class of applications, we regard the ED-projector as a generic input/output device that can replace, in many situations, current displays and interactive devices by creating computer desktop-like interactive displays on non-tethered surfaces. For example, a desktop application can be projected directly on any surface, as previously suggested by Wellner [9] and as demonstrated in subsection 4.1 . Unlike the interactive whiteboard described in [34], the projected application can be easily moved around the room, for instance, from a whiteboard on the wall to the top of a desk for more detailed reading. Similarly, the ED-projector can be used to bring information to the physical location where it is used or needed. For example, a database application managing reports can be projected on top of the file cabinet with hard copies of the reports.

In the second class of applications the ED-projector can augment reality by pointing to physical objects, showing connections among them, and projecting patterns to indicate movement or change in the real world. The prototype assembly task described in subsection 4.2 is a typical example of this class of applications. Another example applies to a library in which directions are provided to a user looking for a book by projecting arrows on the floor, walls and shelves, and finally by highlighting the desired book.

Steerable graphics produced by the ED-projector can also aid in collaborative work by providing large displays on convenient surfaces, and allowing easy reconfiguration of meeting spaces for different functions and teamwork styles, in similar ways to the office project described in subsection 4.1 .

The ED-projector can also be used to provide computer and information access in environments where traditional displays are

Amazon Ex. 1012.0181

difficult to install, secure, or maybe unsafe to operate, such as public spaces and areas subject to harsh environmental conditions. The device also permits an interactive display to be brought to the proximity of a user, eliminating the need for the user to approach a traditional, fixed display. In particular, the ED-projector can facilitate the access and use of computers by people with locomotive disabilities. For instance, it can project an interactive display on a hospital bed sheet without requiring the patient to contact any device.

We also see the ED-projector as a potential enabler of a new generation of games where interaction happens not in the virtual world but in the physical world such as the "Fro…og!" game described above. Unlike games based on phones and portable computers [4], the use of the ED-projector provides high-resolution displays where characters and fantastic objects can move from surface to surface, creating a game that surrounds the players. Using video input, various user actions can be detected depending on the game needs, for example, hand gestures, body movements, foot action, and/or facial expressions. Above all, a single ED-projector installed on the ceiling of a living room or an arcade, using just software commands, can completely reconfigure the space to create different games according to the interests, age, and motor skills of the players.

The scenarios examined so far are chiefly concerned in creating ways for humans to command computers and/or obtain information. However, we see the projecting of interactive graphics ubiquitously as a way to provide computer agents the ability to "reach and touch" the real world and the people that inhabit it. By using projected patterns, it is possible to make computers not only interact with people, but act and coordinate them. For instance, a computer agent can control people in a line by projecting red lines and green arrows around the people on the line. A steerable projector system is, thus, able to create a "body", composed of light, for an agent. From this perspective, we may be creating an interface to the real world for computers.

5.2 Human Interface Issues

As foreseen by Weiser [1], ubiquitous interactive graphics do not seem to follow the standard paradigm of interaction based on the desktop metaphor. In our M&M demonstration, we had the opportunity to observe more than 650 novice users experiencing a scenario of true ubiquitous computing, in the context of an assembly task. On one side, we observed that they quite naturally accepted the idea of augmenting reality with information directly projected on objects. Moreover, we saw they naturally change their body's position to minimize problems like the shadow of their arms and body over the board. People seemed not only to accept projected graphics naturally but also to compensate for its limitations.

However, when trying to click a button, the users almost always applied pressure onto the projected surface with their fingers. Also, whenever a button touch was not detected by the vision system, the most common reaction was to press again with more pressure and, if still not successful, using the whole hand instead of a finger. It seems that people expect buttons, even virtual buttons, to react to pressure and not to touch, and to fail if enough pressure is not applied on them.

This observation clearly illustrates the limitations of the hand-as-a-mouse paradigm used in the M&M demo. Although it is easier for users to relate the ubiquitous experience to their everyday contact with desktop interfaces, the metaphor proved to be misleading to the true nature of the interface. Unlike a mouse that has appropriate haptic feedback for clicking gestures, touching a projected button is devoid of feedback movement.

Similar observations prompt us to believe that a new paradigm for interaction with projected screens must be developed, with new kinds of widgets, tuned for the natural capabilities and limitations of projection and gesture-based interaction, just as the point and click paradigm is tuned for the characteristics of a physical mouse. Touch-screen-like interactions do translate better to the projected interface domain. Although our widgets are not based on surface contact, the widgets we have found useful so far (buttons, scroll bars and touch pads) are all based on motions people naturally make when touching.

One of the exciting aspects of this work is that it is an excellent environment to explore more adventurous gesture-based interaction styles. By giving the user feedback as to what gestural interactions the system is expecting and where in the environment they may be peformed, the cognitive load on the user can be reduced. Without such feedback a user would have to memorize what amounts to a set of complex dance steps to simply use a gesture-based interface.

This investigation of new human-computer interface paradigms is even more important in the case of multi-user interaction. Unlike desktops, wearable computers, or PDAs, projected interactive graphics naturally allow many users to simultaneously interact with an application. Although the applications we have developed so far are all single-user, it seems clear multi-user interaction will soon became an important area of research.

Another intriguing question is related to the different degrees of difficulty experienced by users when interacting with graphics projected on different supporting objects. Clicking a paint bucket seems to be harder to accept then clicking a table. Of course the design of our system and the simplicity of our experiment allow us neither to definitively state that the phenomenon does happen nor to draw conclusions about its causes. However, one hypothesis is that people seem to attribute different interactive capabilities to projected interactive graphics according to the nature of the surfaces themselves. Are the paint buckets "functionally fixed", that is, if something looks and acts like a paint bucket, is it less likely to be perceived as something else – namely, a touch-screen? We are currently preparing a study to determine the validity of this hypothesis.

Finally, our research so far has addressed only visual input and output. A promising area of research seems to be the application of the concept to steerability of other stimuli such as sound, using, for instance, the audio spotlights described by Pompei [35]. What kind of interfaces and applications can be created by a full multimodal steerable system?

6 CONCLUSION

This paper describes a new method and device for creating ubiquitously interactive graphics on non-tethered surfaces. We have shown three different methods to determine the warping function that corrects for oblique projection distortion and a computer vision system that tracks the user's hand over the projected images and detects touch-like gestures. We have also implemented three applications that realize ubiquitous computing concepts such as everywhere computer access and augmented reality. Based on the experience of running the M&M demo with hundreds of novice users, we have also detected a need to re-examine the current interaction paradigms, and possibly, investigate and propose new ones.

We see this work as an important contribution to realize the vision of ubiquitous computing, without asking people to wear headsets or wiring the surfaces of everyday objects. However, we do not advocate projection-based solutions as a panacea for ubiqutious applications: in many cases, it is more desirable and efficient to use head-mounted displays. A good analogy is to

Amazon Ex. 1012.0182

compare headsets and projectors to Walkmans and stereo sound systems. Walkmans are unsurpassable as a way to provide music when people are moving around or in public spaces such as trains and airplanes. They are very effective in creating a personal experience, albeit not able to create social interaction. Similarly, we keep loving stereo systems because they render music to a whole space, nurturing a collective experience that is attached to the space itself. We see similar advantages and limitations when comparing headsets and steerable projectors as a way to render interactive graphics onto the real world.

Acknowledgements

"M&M" is a registered trademark of Mars, Inc. "Walkman" is a registered trademark of Sony Corporation.

References

1. Weiser, M., *The Computer for the Twenty-First Century.* Scientific American, 1991. **265**(3): p. 94-100.
2. Pinhanez, C. *The Everywhere Displays Projector: A Device to Create Ubiquitous Graphical Interfaces*. in *Proc. of Ubiquitous Computing 2001 (Ubicomp'01)*. 2001. Atlanta, Georgia.
3. Barfield, W. and T. Caudell, eds. *Fundamentals of Wearable Computers and Augmented Reality.* 2001, Lawrence Erlbaum Assoc. 728.
4. Raskar, R., et al. *The Office of the Future: A Unified Approach to Image-Based Modeling and Spatially Immersive Displays*. in *Proc. of SIGGRAPH'98*. 1998. Orlando, Florida.
5. Ishii, H. and B. Ullmer. *Tangible Bits: Towards Seamless Interfaces between People, Bits, and Atoms*. in *Proc. of CHI'97*. 1997. Atlanta, Georgia.
6. Bellour, R., et al., *Tony Oursler*. 2001: Ediciones Poligrafa. 208.
7. Naimark, M., *Spatial Correspondence in Motion Picture Display.* SPIE Optics in Entertainment II, 1984. **462**: p. 78-81.
8. Bolt, R., *Put That There: Voice and Gesture at the Graphics Interface.* ACM Computer Graphics, 1980. **14**(3): p. 262-270.
9. Wellner, P., *Interacting with Paper on the DigitalDesk.* Communications of the ACM, 1993. **36**(7).
10. Morishima, S., et al. *HyperMask: Talking Head Projected Onto Real Objects*. in *Proc. of Multimedia Modeling (MMM'00)*. 2000: World Scientific.
11. Raskar, R., et al. *Shader Lamps: Animating Real Objects with Image-Based Illumination*. in *Proc. of 12th Eurographics Workshop on Rendering*. 2001. London, England.
12. Krueger, M.W., *Artificial Reality II*. 1990: Addison-Wesley.
13. Keays, B. and R. Macneil. *metaField Maze.* in *Proc. of SIGGRAPH'99*. 1999. Los Angeles, California.
14. Omojola, O., et al., *An Installation of Interactive Furniture.* IBM Systems Journal, 2000. **39**(3&4): p. 861-879.
15. Ju, W. *Origami Desk*. in *Proc. of SIGGRAPH'01 - Conference Abstracts and Applications*. 2001. Los Angeles, California.
16. Back, M., et al., *Designing Interactive Reading Experiences for a Museum Exhibition.* IEEE Computer Magazine, 2001. **34**(1): p. 1-8.
17. Underkoffler, J., B. Ullmer, and H. Ishii. *Emancipated Pixels: Real-World Graphics in the Luminous Room*. in *Proc. of SIGGRAPH'99*. 1999. Los Angeles, CA.
18. Rekimoto, J. *A Multiple Device Approach for Supporting Whiteboard-based Interactions*. in *Proc. of CHI'98*. 1998. Los Angeles, CA.
19. Schiele, B. and J.L. Crowley, *Recognition without Correspondence using Multidimensional Receptive Field Histograms.* IJCV, 2000. **36**(1): p. 31-52.
20. Wu, Y. and T. Huang, *Vision-Based Gesture Recognition: A Review.* Lecture Notes in Artificial Intelligence, 1999. **1739**.
21. Hardenberg, C.v. and F. Berard. *Bare-hand human-computer interaction*. in *Proc. of Workshop on Perceptive User Interfaces, PUI'01*. 2001. Orlando, Florida.
22. Quek, F., T. Mysliwiec, and M. Zhao. *Finger mouse: A freehand pointing interface*. in *Proc. of International Workshop on Automatic Face- and Gesture-Recognition*. 1995. Zurich, Switzerland.
23. Segen, J. *GestureVR: Vision-Based 3D Hand Interface for Spatial Interaction*. in *Proc. of ACM Multimedia Conference*. 1998. Briston, England.
24. Kjeldsen, F., *Visual Recognition of Hand Gesture as a Practical Interface Modality*. 1997, Columbia University: New York, New York.
25. Raskar, R., et al. *Multi-Projector Displays Using Camera-Based Registration*. in *Proc. of IEEE Visualization'99*. 1999. San Francisco, CA.
26. Pinhanez, C., F. Nielsen, and K. Binsted. *Projecting Computer Graphics on Moving Surfaces: A Simple Calibration and Tracking Method*. in *Proc. of SIGGRAPH'99*. 1999. Los Angeles, California.
27. Raskar, R. *Oblique Projector Rendering on Planar Surfaces for a Tracked User*. in *Proc. of SIGGRAPH'99*. 1999. Los Angeles, California.
28. Faugeras, O., *Three-Dimensional Computer Vision: A Geometric Viewpoint*. 1993, Cambridge, Massachusetts: The MIT Press.
29. Sukthankar, R., R. Stockton, and M. Mullin. *Smarter Presentations: Exploiting Homography in Camera-Projector Systems*. in *Proc. of ICCV'01*. 2001. Vancouver, Canada.
30. Yang, R. and G. Welch. *Automatic and Continuous Projector Display Surface Calibration Using Every-Day Imagery*. in *Proc. of 9th International Conf. in Central Europe in Computer Graphics, Visualization, and Computer Vision*. 2001. Plzen, Czech Republic.
31. Kjeldsen, R., et al. *Interacting with Steerable Projected Displays*. in *Submitted to Face and Gesture'02*. 2002.
32. Wren, C., et al., *Pfinder: Real-Time Tracking of the Human Body.* IEEE Trans. Pattern Analysis and Machine Intelligence, 1997. **19**(7): p. 780-785.
33. Buschmann, F., et al., *A System of Patterns - Pattern Oriented Software Architecture.* 1996, New York, New York: John Wiley & Sons.
34. Crowley, J.L., J. Coutaz, and F. Berard, *Things that See.* Communications of the ACM, 2000. **43**(3): p. 54-64.
35. Pompei, F.J. *The Use of Airborne Ultrasonics for Generating Audible Sound Beams*. in *Proc. of 105th Audio Engineering Society Convention*. 1998. San Francisco, CA.

Amazon Ex. 1012.0183

Appendix B-2

Claudio S. Pinhanez



feedback

Senior Manager, Conversational Intelligence
IBM Research - Brazil , Sao Paulo
csantosp@br.ibm.com

Professional Associations: ACM | ACM SIGCHI | Brazilian Computer Society | IBM Academy of Technology | IEI
More information: LinkedIn

2020

Creating Corpora for Seq2Seq Tone Rephrasing Using Social Media Posts
Paulo Cavalin, Marisa Vasconcelos, Marcelo Grave and Claudio Pinhanez
International Joint Conference on Neural Networks, 2020

From Disjoint Sets to Parallel Data to Train Seq2Seq Models for Sentiment Transfer
Paulo R. Cavalin, Marisa Vasconcelos, Marcelo Grave, Claudio S. Pinhanez, Victor Henrique Alves Ribeiro
Findings of the Association for Computational Linguistics: EMNLP 2020, pp. 689-698
artificial neural network, semantic similarity, machine translation, machine learning, computer science, exploit, disj
grammar, fluency, artificial intelligence

Improving Out-of-Scope Detection in Intent Classification by Using Embeddings of the Word Graph S
Paulo R. Cavalin, Victor Henrique Alves Ribeiro, Ana Paula Appel, Claudio S. Pinhanez
Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 3952-3961
Abstract graph embedding, theoretical computer science, percentage point, computer science, inverse, discrete se

HCI Research Challenges for the Next Generation of Conversational Systems
Claudio S. Pinhanez
Proceedings of the 2nd Conference on Conversational User Interfaces, 2020
psychology, human computer interaction

Amazon Ex. 1012.0185

Unveiling Practices and Challenges of Machine Teachers of Customer Service Conversational Systems
Heloisa Candello, Mairieli Santos Wessel, Claudio S. Pinhanez
Proceedings of the 2nd Conference on Conversational User Interfaces, 2020
knowledge management, customer service, computer science

Intentions, Meanings, and Whys: Designing Content for Voice-based Conversational Museum Guides
Fabrecio Barth, Heloisa Candello, Paulo R. Cavalin, Claudio S. Pinhanez
Proceedings of the 2nd Conference on Conversational User Interfaces, 2020
linguistics, computer science

Co-designing a conversational interactive exhibit for children
Heloisa Candello, Mauro Pichiliani, Claudio S. Pinhanez, Sara Vidon, Mairieli Santos Wessel
Proceedings of the 2020 ACM Interaction Design and Children Conference: Extended Abstracts, pp. 326-331
human computer interaction, computer science

2019

Can direct address affect user engagement with chatbots embodied in physical spaces
Heloisa Candello, Claudio Pinhanez, Mauro Pichiliani, Marisa Vasconcelos, Haylla Conde
Proceedings of the 1st International Conference on Conversational User Interfaces, 2019
vocative case, visitor pattern, user experience design, social relation, psychology, performing arts, observational stu
empirical research, embodied cognition

Machine Teaching by Domain Experts: Towards More Humane, Inclusive, and Intelligent Machine Lea
Claudio S. Pinhanez
arXiv preprint arXiv:1908.08931, 2019
Abstract teaching technology, open source software, mediation, intelligent machine, human computer interaction,
approach

Live It Up: Analyzing Emotions and Language Use in Tweets during the Soccer World Cup Finals
Marisa Vasconcelos, Jussara Almeida, Paulo Cavalin, Claudio Pinhanez
10th ACM Conference on Web Science , 2019

Teaching Robots to Act and Converse in Physical Spaces: Participatory Design Fictions with Museum
Heloisa Candello, Mauro Pichiliani, Mairieli Wessel, Claudio Pinhanez, Michael Muller
Proceedings of the Halfway to the Future Symposium 2019 on
visual arts, the imaginary, sociology, social dilemma, robot, participatory design, emerging technologies, dilemma, c

The Effect of Audiences on the User Experience with Conversational Interfaces in Physical Spaces
Heloisa Candello, Claudio Pinhanez, Mauro Pichiliani, Paulo Cavalin, Marisa Vasconcelos, Haylla Conde
CHI 2019

Amazon Ex. 1012.0186

2018

Recovering from Dialogue Failures Using Multiple Agents in Wealth Management Advice

Heloisa Candello, Claudio Pinhanez

Studies in Conversational UX Design, Springer, 2018

Moore, R.J., Szymanski, M.H., Arar, R., Ren, G.-J. (Eds.)

The Role of Dialogue User Data in the Information Interaction Design of Conversational Systems

Heloisa Candello, Claudio Pinhanez

DUXU'2018, Springer

Specifying and Implementing Multi-Party Conversation Rules with Finite-State-Automata

Maira Gatti de Bayser, Melina Alberio Guerra, Paulo R. Cavalin, Claudio S. Pinhanez

AAAI Workshops, *pp. 713-719*, 2018

machine learning, finite state machine, conversation, computer science, artificial intelligence

Having an Animated Coffee with a Group of Chatbots from the 19 th Century

Candello, Heloisa and Pinhanez, Claudio and Pichiliani, Mauro Carlos and Guerra, Melina Alberio and Gatti de Bayse

Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems, *pp. D206*

Abstract

Revealing Physicians Referrals from Health Insurance Claims Data

de Santana, Vagner Figueredo and Appel, Ana Paula and Moyano, Luis Gregorio and Ito, Marcia and Pinhanez, Claud

Big Data Research, Elsevier, 2018

Abstract

2017

Personality Archetypes, Pantomime, Comics: New Methods to Design Interfaces with Human Traits

Claudio S. Pinhanez

International Conference on Computer-Human Interaction Research and Applications, 1-19, 2017

Abstract service recovery, personality, human computer interaction, humanoid robot, engineering design process, comics, archetype, animation

Modeling Task Deviations as Eccentricity Distribution Peaks

de Santana, Vagner Figueredo and de Paula, Rogério Abreu and Pinhanez, Claudio Santos

Proceedings of the XVI Brazilian Symposium on Human Factors in Computing Systems, *pp. 36*, 2017

Abstract (Honorable Mention Paper)

Amazon Ex. 1012.0187

Typefaces and the Perception of Humanness in Natural Language Chatbots
Candello, Heloisa and Pinhanez, Claudio and Figueiredo, Flavio
Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, *pp. 3476--3487*
Abstract

A Hybrid Architecture for Multi-Party Conversational Systems
Gatti de Bayser, Maira and Cavalin, Paulo and Souza, Renan and Braz, Alan and Candello, Heloisa and Pinhanez, Cla
arXiv preprint arXiv:1705.01214, 2017
Abstract

Shaping the Experience of a Cognitive Investment Adviser
Candello, Heloisa and Pinhanez, Claudio and Millen, David and Andrade, Bruna Daniele
International Conference of Design, User Experience, and Usability, *pp. 594--613*, 2017

Bottester: Testing Conversational Systems with Simulated Users
Marisa Vasconcelos, Heloisa Candello, Claudio Pinhanez
IHC 2017, ACM

Evaluating the Conversation Flow and Content Quality of a Multi-bot Conversational System
Heloisa Candello, Marisa Vasconcelos, Claudio Pinhanez
HCI in Practice (IHC na Pratica) Workshop (co-located with the Brazilian Symposium about Human Factores in Comp

Computer Interfaces to Organizations: Perspectives on Borg-Human Interaction Design
Pinhanez, Claudio
arXiv preprint arXiv:1712.03012, 2017
Abstract

2016

Service Design Tools to Explore Financial Services for Poor Microbusiness Owners
Candello, Heloisa and Pinhanez, Claudio and Millen, David and Bianchi, Silvia
Service Design Geographies. Proceedings of the ServDes. 2016 Conference, *pp. 192--204*
Abstract

2015

Using Visual Analytics to Diagnose Productivity and Quality Issues on IT Service Pools
Cavalcante, Victor and Candello, Heloisa and Pinhanez, Claudio and Braz, Alan and Amorin, Franklin and Nauata, Ne
Maximizing Management Performance and Quality with Service Analytics, *pp. 207--232*, IGI Global, 2015
Abstract

Amazon Ex. 1012.0188

Understanding Fiado : Informal Credit in Brazil

Heloisa Candello, David Millen, Silvia Bianchi, Rogerio de Paula, Claudio Pinhanez

Proceedings of the 2015 Annual Symposium on Computing for Development, 147-147

Abstract northeast brazil, microfinance, marketing, financial application, database transaction, biology

A scalable architecture for real-time analysis of microblogging data

Cavalin, Paulo Rodrigo and Gatti, Ma{\'\i}ra A de C and Moraes, TGP and Oliveira, Fabio Silva and Pinhanez, Claudio S Paula, Rog{\'e}rio Abreu

IBM Journal of Research and Development 59(*2/3*), 16--1, IBM, 2015

Abstract

Collaboration and social computing in emerging financial services

Millen, David R and Pinhanez, Claudio and Kaye, Jofish and Bianchi, Silvia Cristina Sardela and Vines, John

Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing,

Abstract

2014

Borg-Human Interaction Design

Claudio S. Pinhanez

ServDes.2014 Service Future; Proceedings of the fourth Service Design and Service Innovation Conference; Lancaster April 2014 pp. 99, 100-109

Abstract service design, personality, narrative, human computer interaction, human interaction, engineering, emb methods

A CIENCIA DA INOVACAO EM SERVICOS: ESTUDO EXPLORATORIO SOBRE OS INTERESSES E PRIOR PESQUISA NO BRASIL

Paulo Cesar Calabria, Roberto Carlos Bernardes, Eduardo Raupp de Vargas, Claudio S. Pinhanez

RAI: Revista de Administracao e Inovacao 10(*4*), 110-135, 2014

Abstract

Unsupervised annotation of usability test videos

de Santana, Vagner Figueredo and Pinhanez, Claudio Santos and de Paula, Rog{'e}rio Abreu

Proceedings of the 13th Brazilian Symposium on Human Factors in Computing Systems, *pp. 207--215*, 2014

Handling Big Data on Agent-based Modeling of Online Social Networks with Mapreduce

Maira A. de C. Gatti, Marcos R. Vieira, Joao Paulo F. de Melo, Paulo Rodrigo Cavalin, Claudio Santos Pinhanez

Proceedings of the 2014 Winter Simulation Conference, *pp. 851--862*, IEEE Press

Towards Personalized Offers by Means of Life Event Detection on Social Media and Entity Matching

Paulo Rodrigo Cavalin, Maira Gatti, Claudio Santos Pinhanez

Amazon Ex. 1012.0189

HT (Doctoral Consortium / Late-breaking Results / Workshops), 2014

Abstract world wide web, software framework, social media, life events, infosphere, ibm, computer science, blosso

2013

An Awareness Raising System to Support Cross-department Coordination in Matrix Organizations

Vagner Figueredo de Santana, Rogerio Abreu de Paula, Claudio Santos Pinhanez, Silvia Cristina Sardela Bianchi, Elai Costa, Carina Missae

Proceedings of Brazilian Symposium on Collaborative Systems, *pp. 136*, 2013

Visual summarization of incidents in the context of service delivery factories

Vagner Figueredo de Santana, Rogerio Abreu de Paula, Claudio S Pinhanez, Silvia Cristina Sardela Bianchi, Elaine Na Carina Missae

Proceedings of the 12th Brazilian Symposium on Human Factors in Computing Systems, *pp. 300--303*, 2013

Data-Driven Analytical Tools for Characterization of Productivity and Service Quality Issues in IT Serv

Victor F. Cavalcante, Claudio S. Pinhanez, Rogerio A. de Paula, Carolina S. Andrade, Cleidson R. B. de Souza, Ana Pau

web science, *pp. 295-310*, 2013

Abstract service quality, service product management, service level requirement, service level objective, service g delivery framework, operations research, operations management, incident management, computer science

Reaction times for user behavior models in microblogging online social networks

Samuel Martins Barbosa Neto, Maira Athanazio De Cerqueira Gatti, Paulo Rodrigo Cavalin, Claudio Santos Pinhanez, Paula Appel

Proceedings of the 2103 workshop on Data-driven user behavioral modelling and mining from social media, *pp. 17--*

Data-driven analytical tools for characterization of productivity and service quality issues in IT service

Cavalcante, Victor F and Pinhanez, Claudio S and de Paula, Rog{'e}rio A and Andrade, Carolina S and de Souza, Cleid

Journal of Service Research, 1094670513480559, SAGE Publications, 2013

Abstract

A Simulation-based Approach to Analyze the Information Diffusion in Microblogging Online Social Ne

Maira A de C Gatti, Ana Paula Appel, Cicero Nogueira dos Santos, Claudio Santos Pinhanez, Paulo Rodrigo Cavalin, S

Winter Simulation Conference, WSC'13, 2013

accepted for publication

Large-Scale Multi-Agent-based Modeling and Simulation of Microblogging-based Online Social Netwo

Maira Gatti, Ana Paula Appel, Claudio Pinhanez, Cicero dos Santos, Daniel Gribel, Paulo Cavalin, and Samuel Barbos

The 14th International Workshop on Multi-Agent-based Simulation (MABS, AAMAS 2013)

Amazon Ex. 1012.0190

2012

Designing an enterprise social tool for cross-boundary communication, coordination, and information
de Souza, Cleidson Ronald B and Pinhanez, Claudio S and Cavalcante, Victor F and Aluani, Fernando and Daros, Vini Paula, Rog{\'e}rio A

Proceedings of the 30th ACM international conference on Design of communication, *pp. 55--64*, 2012

Abstract

Using Social Analytics for Studying Work-Networks: A Novel, Initial Approach
Rog\'erio A de Paula, Ana Paula Appel, Claudio S Pinhanez, Victor F Cavalcante, Carolina S Andrade

Collaborative Systems (SBSC), 2012 Brazilian Symposium on, *pp. 146--153*

Scheduling with Preemption for Incident Management: When Interrupting Tasks is not Such a Bad Ide
Marcos D. Assuncao, Victor F. Cavalcante, Maira A. de C. Gatti, Marco A. S. Netto, Claudio S. Pinhanez, and Cleidson

Proceedings of the Winter Simulation Conference (WSC '12), *pp. 12*, ACM, 2012

Characterizing Time-Bounded Incident Management Systems
V Cavalcante, Claudio S Pinhanez, CB de Souza, RA de Paula, Ana Paula Appel, Carolina S Andrade

SRII Global Conference (SRII), 2012 Annual, *pp. 750--759*

2011

Knowledge and information and needs of system administrators in IT service factories
De Souza, Cleidson RB and Pinhanez, Claudio S and Cavalcante, Victor

Proceedings of the 10th Brazilian Symposium on Human Factors in Computing Systems and the 5th Latin American C Interaction, *pp. 81--90*, 2011

Abstract

Information needs of system administrators in information technology service factories
C.R.B. de Souza, C.S. Pinhanez, V.F. Cavalcante

Proceedings of the 5th ACM Symposium on Computer Human Interaction for Management of Information Technology,

Human aspects of internet services: considering the needs of users and providers
C.S. Pinhanez

Journal of Internet Services and Applications 1(*3*), 155--164, Springer, 2011

2010

Designing the interaction with service systems
C. Pinhanez

Proceedings of the 28th ACM International Conference on Design of Communication, *pp. 253--253*, 2010

Amazon Ex. 1012.0191

2009

Services as Customer-Intensive Systems
Claudio Pinhanez

Design Issues 25(*2*), 3-13, 2009

Abstract type of service, service system, service provider, service product management, service design, service de process management, marketing, information system, business

A Service Science Perspective on Human-Computer Interface Issues of Online Service Applications [
C. Pinhanez

International Journal of Information Systems in the Service Sector (IJISSS) 1(*2*), 17--35, IGI Global, 2009

Humans insideâ€¦ as the key characteristic of service systems
C. Pinhanez

Proceedings of QUIS, *pp. 515--524*, 2009

Multimedia Chat for Helpdesks: A Requirements Study, a Practical SOA Architecture, and an Interface
Z.Y. Shae, T. Bergstrom, C. Pinhanez, M. Podlaseck

IEEE International Symposium on Object/Component/Service-Oriented Real-Time Distributed Computing ISORC'09, *p*

2008

Aplicando conceitos e tecnicas de servicos em IHC
Claudio Pinhanez

IHC 08 Proceedings of the VIII Brazilian Symposium on Human Factors in Computing Systems, *pp. 346-347*, 2008

Abstract philosophy, humanities

Multimedia Chat for Helpdesks: a Practical SOA Architecture
Z.Y. Shae, T. Bergstrom, C. Pinhanez, M. Podlaseck

IEEE Congress on Services. Part I, *pp. 75-77*, 2008

Service systems as customer-intensive systems and its implications for service science and engineer
C. Pinhanez

Hawaii International Conference on System Sciences, Proceedings of the 41st Annual, *pp. 117--117*, 2008

A service science perspective for interfaces of online service applications
C. Pinhanez

Proceedings of the VIII Brazilian Symposium on Human Factors in Computing Systems, *pp. 11--20*, 2008

Amazon Ex. 1012.0192

A Proposal for a Service Science Discipline Classification System
C. Pinhanez, P. Kontogiorgis
Frontiers in Service conference, 2008

2007

Ubiquitous Services
C. Pinhanez
Services Computing, 2007. SCC 2007. IEEE International Conference on, *pp. 146--153*

A Services Theory Approach to Online Service Applications
C. Pinhanez
Services Computing, 2007. SCC 2007. IEEE International Conference on, *pp. 395--402*

2006

From Users to Customers: HCI for Online Services
C. Pinhanez
Information \& Communications Technology, 2006. ICICT'06. ITI 4th International Conference on, *pp. 1--1*

2005

A study on the manipulation of 2D objects in a projector/camera-based augmented reality environme
S. Voida, M. Podlaseck, R. Kjeldsen, C. Pinhanez
Proceedings of the SIGCHI conference on Human factors in computing systems, *pp. 611--620*, 2005

To frame or not to frame: The role and design of frameless displays in ubiquitous applications
C. Pinhanez, M. Podlaseck
UbiComp 2005: Ubiquitous Computing, 903-903, Springer

Using symbiotic displays to view sensitive information in public
S. Berger, R. Kjeldsen, C. Narayanaswami, C. Pinhanez, M. Podlaseck, M. Raghunath
3rd IEEE International Conference on Pervasive Computing and Communications, PerCom, *pp. 139-148*, 2005

2004

A portable system for anywhere interactions
N. Sukaviriya, R. Kjeldsen, C. Pinhanez, L. Tang, A. Levas, G. Pingali, M. Podlaseck
CHI'04 extended abstracts on Human factors in computing systems, *pp. 789--790*, 2004

Amazon Ex. 1012.0193

Projector-camera systems for telepresence
C. Pinhanez, G. Pingali
Proceedings of the 2004 ACM SIGMM workshop on Effective telepresence, *pp. 63--66*

Dynamically reconfigurable vision-based user interfaces
R. Kjeldsen, A. Levas, C. Pinhanez
Machine Vision and Applications 16(*1*), 6--12, Springer, 2004

2003

Interval scripts: a programming paradigm for interactive environments and agents
C.S. Pinhanez, A.F. Bobick
Personal and Ubiquitous Computing 7(*1*), 1--21, Springer, 2003

An Architecture and Framework for Steerable Interface Systems
A. Levas, C. Pinhanez, G. Pingali, R. Kjeldsen, M. Podlaseck, N. Sukaviriya
UbiComp 2003: Ubiquitous Computing, *pp. 333--348*

Ubiquitous interactive displays in a retail environment
C. Pinhanez, M. Podlaseck, R. Kjeldsen, A. Levas, G. Pingali, N. Sukaviriya
Proceedings of SIGGRAPH Sketches 2(*3*), Citeseer, 2003

Applications of steerable projector-camera systems
C Pinhanez, R Kjeldsen, A Levas, G Pingali, M Podlaseck, N Sukaviriya
ICCV 2003: Proceedings of the IEEE International Workshop on Projector-Camera Systems

Augmenting a Retail Environment Using Steerable Interactive Displays
Noi Sukaviriya, Mark Podlaseck, Rick Kjeldsen, Anthony Levas, Gopal Pingali, Claudio Pinhanez
CHI '03 extended abstracts on Human factors in computing systems, *pp. 978--979*, ACM, 2003
Abstract 608035

Ubiquitous computing in the living room: Concept sketches and an implementation of a persistent use
S.S. Intille, V. Lee, C. Pinhanez
Proc. of the Video Session of Ubi-Comp3, 12--15, 2003

On interfaces projected onto real-world objects
M. Podlaseck, C. Pinhanez, N. Alvarado, M. Chan, E. Dejesus
CHI'03 extended abstracts on Human factors in computing systems, *pp. 802--803*, 2003

Creating touch-screens anywhere with interactive projected displays
C. Pinhanez, R. Kjeldsen, L. Tang, A. Levas, M. Podlaseck, N. Sukaviriya, G. Pingali
Proceedings of the eleventh ACM international conference on Multimedia, *pp. 460--461*, 2003

Amazon Ex. 1012.0194

Embedding interactions in a retail store environment: The design and lessons learned

N. Sukaviriya, M. Podlaseck, R. Kjeldsen, A. Levas, G. Pingali, C. Pinhanez

Proc. of the Ninth IFIP International Conference on Human-Computer Interaction, Zurich, Switzerland, 2003

Steerable interfaces for pervasive computing spaces

G. Pingali, C. Pinhanez, A. Levas, R. Kjeldsen, M. Podlaseck, H. Chen, N. Sukaviriya

Proceedings of the First IEEE International Conference on Pervasive Computing and Communications, PerCom, pp. 31

Fostering a symbiotic handheld environment

M. Raghunath, C. Narayanaswami, C. Pinhanez

Computer 36(*9*), 56--65, IEEE Computer Society Press, 2003

2002

HyperMask --- projecting a talking head onto a real object

T. Yotsukura, S. Morishima, F. Nielsen, K. Binsted, C. Pinhanez

The Visual Computer 18(*2*), 111-120, 2002

Abstract computer graphics, ibm, watson, computer graphics images, artificial neural network, computer vision, cc artificial intelligence

It/I: a theater play featuring an autonomous computer character

Claudio S. Pinhanez, Aaron F. Bobick

Presence: Teleoperators & Virtual Environments 11(*5*), 536-548, 2002

Abstract scripting language, virtual machine, narrative, multimedia, computer science

" It/I": A Theater Play Featuring an Autonomous Computer Character

C.S. Pinhanez, A.F. Bobick

Presence-Teleoperators and Virtual Environments 11(*5*), 536--548, Cambridge, MA: MIT Press, c1992-, 2002

HyperMask--projecting a talking head onto a real object

T. Yotsukura, S. Morishima, F. Nielsen, K. Binsted, C. Pinhanez

The Visual Computer 18(*2*), 111--120, Springer, 2002

That's entertainment! Designing streaming, multimedia web experiences

C.M. Karat, J. Karat, J. Vergo, C. Pinhanez, D. Riecken, T. Cofino

International Journal of Human-Computer Interaction 14(*3-4*), 369--384, Taylor \& Francis, 2002

Ubiquitous interactive graphics

C. Pinhanez, R. Kjeldsen, T. Levas, G. Pingali, M. Podlaseck, P. Chou

IBM Research Report RC22495 (W0205-143), 2002

Amazon Ex. 1012.0195

Creating ubiquitous interactive games using everywhere displays projectors
C. Pinhanez

Proc. of the International Workshop on Entertainment Computing (IWECâ€™02), pp. 149--156, 2002

User-Following Displays
G. Pingali, C. Pinhanez, A. Levas, F. Kjeldsen, M. Podlaseck

Proc. of the IEEE International Conference on Multimedia and Expo 2002 (ICME'02) , pp. 845-848

Interacting with Steerable Projected Displays
R. Kjeldsen, C. Pinhanez, G. Pingali, J. Hartman, A. Levas, M. Podlaseck.

Proc. of the 5th International Conference on Automatic Face and Gesture Recognition (FG'02), pp. 402--407, 2002

BlueSpace: personalizing workspace through awareness and adaptability
Jennifer Lai, Anthony Levas, Paul Chou, Claudio Pinhanez, Marisa Viveros

International Journal of Human-Computer Studies 57(5), 415--428, Elsevier, 2002

2001

"Less clicking, more watching": an option for entertainment on the web?
Claudio Pinhanez, Clare Marie Karat, John Karat, John Vergo, Renee Arora

CHI 01 Extended Abstracts on Human Factors in Computing Systems, pp. 347-348, 2001

Abstract web design, mashup, entertainment, world wide web, multimedia, narrative, computer science

Hyper Mask-Projecting Virtual Face onto a Moving Real Object
S. Morishima, T. Yotsukura, F. Nielsen, K. Binsted, C. Pinhanez

Proc. of Eurographics' 01, 2001

Can Web entertainment be passive
C. Pinhanez, C.M. Karat, J. Vergo, J. Karat, R. Arora, D. Riecken, T. Cofino

Proc, 2001

Augmenting reality with projected interactive displays
C. Pinhanez

Proceedings from Virtual and Augmented Architecture, 93--100, 2001

Less clicking, more watching: Results of the iterative design and evaluation of entertaining web exper
C.M. Karat, C. Pinhanez, J. Karat, R. Arora, J. Vergo

Proceedings of IFIP TC, pp. 455--463, 2001

Less clicking, more watching: an option for entertainment on the web?
C. Pinhanez, C.M. Karat, J. Karat, J. Vergo, R. Arora

CHI'01 extended abstracts on Human factors in computing systems, pp. 347--348, 2001

Amazon Ex. 1012.0196

Transforming Surfaces into Touch-Screens
C. Pinhanez, R. Kjeldsen, A. Levas, G. Pingali, J. Hartman, V. Kwatra, M. Podlaseck, P. Chou
submitted to CHI, Citeseer, 2001

" Less Clicking, More Watching": Results from the User-Centered Design of a Multi-Institutional Web S
J. Vergo, C.M. Karat, J. Karat, C. Pinhanez, R. Arora, T. Cofino, D. Riecken, M. Podlaseck
2001 - ERIC, ERIC

BlueSpace: Creating a personalized and context-aware workspace
Paul Chou, Marco Gruteser, Jennifer Lai, Anthony Levas, Scott McFaddin, Claudio Pinhanez, Marisa Viveros, Danny W
Relat\'orio t\'ecnico1, 10--75, 2001

Using a steerable projector and a camera to transform surfaces into interactive displays
C. Pinhanez
CHI'01 extended abstracts on Human factors in computing systems, pp. 369--370, 2001

The everywhere displays projector: A device to create ubiquitous graphical interfaces
C. Pinhanez
Ubicomp 2001: Ubiquitous Computing, pp. 315--331

2000

HyperMask: Talking Head Projected Onto Real Objects
S. Morishima, T. Yotsukura, K. Binsted, F. Nielsen, C. Pinhanez
Proc. of Multimedia Modeling (MMM'00), 2000

Perceptual user interfaces: the KidsRoom
A.F. Bobick, S.S. Intille, J.W. Davis, F. Baird, C.S. Pinhanez, L.W. Campbell, Y.A. Ivanov, A. Sch\"utte, A. Wilson
Communications of the ACM 43(*3*), 60--61, ACM, 2000

Physically interactive story environments
CS Pinhanez, JW Davis, S. Intille, MP Johnson, AD Wilson, AF Bobick, B. Blumberg
IBM Systems Journal 39(*3.4*), 438--455, IBM, 2000

1998

Scripting method based on temporal intervals for designing interactive systems
K. Mase, CS Pinhanez, AF Bobick
Trans. of IPSJ 39(*5*), 1403--1413, 1998

Amazon Ex. 1012.0197

1997

Pnf propagation and the detection of actions described by temporal intervals
C.S. Pinhanez, A.F. Bobick
Proc. of the DARPA Image Understanding Workshop, 1997

1995

Scripts in machine understanding of image sequences
C. Pinhanez, A. Bobick
AAAI Fall Symposium Series on Computational Models for Integrating Language and Vision, 1995

Year Unknown

Design Methods for Personified Interfaces
Pinhanez, Claudio S
researchgate.net, 0
Abstract

Supporting Remote and Simultaneous Participation in Meetings
C. Pinhanez
fxpal.com, 0

Share this page

Projects and Groups

- Conversational Intelligence Group in IBM Research - Brazil
- IBM Research @ CHI 2021

Technical Areas

Amazon Ex. 1012.0198

- Human Computer Interaction and Data Visualization
- Knowledge
- Mobile Computing
- Natural Language Processing PIC at IBM Research AI
- Reasoning
- Speech

Amazon Ex. 1012.0199

Appendix B-3

Claudio Pinhanez - Professional Website

Claudio Pinhanez
IBM Research - Brazil
Manager, Social Data Analytics

Ph.D. in Media Arts and Sciences
M.I.T. Media Laboratory,1999

Research interests:
Social Computing,
Service Science, HCI,
Ubiquitous Computing,
Computer Vision, AI.

E-mail:	claudio at pinhanez dot com
Twitter:	@cinhanez
Linkedin:	br.linkedin.com/in/pinhanez/
Lattes:	lattes.cnpq.br/5774595361715876



© Claudio Pinhanez, 2004

Last update: Jan 04, 2016

Appendix B-4

RC22495 (W0205-143) May 17, 2002
Computer Science

IBM Research Report

Ubiquitous Interactive Graphics

Claudio S. Pinhanez, Frederik C. Kjeldsen, Anthony Levas,
Gopal S. Pingali, Mark E. Podlaseck, Paul B. Chou

IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598



Research Division
Almaden - Austin - Beijing - Delhi - Haifa - India - T. J. Watson - Tokyo - Zurich

LIMITED DISTRIBUTION NOTICE: This report has been submitted for publication outside of IBM and will probably be copyrighted if accepted for publication. It has been issued as a Research Report for early dissemination of its contents. In view of the transfer of copyright to the outside publisher, its distribution outside of IBM prior to publication should be limited to peer communications and specific requests. After outside publication, requests should be filled only by reprints or legally obtained copies of the article (e.g. , payment of royalties). Copies may be requested from IBM T. J. Watson Research Center , P. O. Box 218, Yorktown Heights, NY 10598 USA (email: reports@us.ibm.com). Some reports are available on the internet at http://domino.watson.ibm.com/library/CyberDig.nsf/home .

Amazon Ex. 1012.0203

Ubiquitous Interactive Graphics

Claudio Pinhanez, Rick Kjeldsen, Gopal Pingali, Anthony Levas, Mark Podlaseck, Paul Chou
IBM Research, T.J. Watson
pinhanez@us.ibm.com

Abstract

This paper presents a device that creates ubiquitous interactive graphics in a real world environment without the need for pre-wiring surfaces or requiring users to carry or wear any special devices. Called the *Everywhere Displays projector*, or *ED-projector*, this device employs *steerability* as the key principle to achieve the promise of ubiquitous computing and augmented reality. The ED-projector uses an LCD projector with a steerable mirror to project graphics onto any pre-calibrated surface, while warping the projected image to correct for oblique distortion. Three different methods to determine this warping function are given. To sense user interaction with the projected image, the ED-projector employs a pan/tilt/zoom camera and a motion-based computer vision system to track the user's hand and touch-like gestures. We propose a system software structure that ties together steering, display, and sensing. Three implemented applications demonstrate the notion of ubiquitously "painting" the real world with interactive graphics, in the context of a futuristic office, an augmented assembly task, and a ubiquitous computer game. The observation of hundreds of users performing the assembly task suggests that the traditional desktop paradigm is inadequate for ubiquitous interaction and that new conceptual widgets and interaction paradigms must be developed.

Keywords: ubiquitous computing; augmented reality; gesture recognition; interactive projections; ubiquitous games.

1 INTRODUCTION

The extraordinary success of computer graphics technology in recent years has been responsible for the creation of amazing new worlds in movies, interactive computer games, and virtual reality (VR) environments. However, this very success poses the question of when the real world, where we live and work, is going to be as rich with imagery and information as movies and games today.

This paper describes technology that can be used to ubiquitously augment, correct, and enhance reality. Although the ultimate goal is to create interactive pixels anywhere in a space as seamlessly as it happens in movies (thanks to CG special effects), this particular work focus on how to "paint" interactive graphics on the surfaces of an environment without requiring any kind of fixed display devices, special wiring of surfaces or the use of encumbering headsets.

The system described here is based on a new device, the *Everywhere Displays projector* (*ED-projector*), composed of a steerable projection system and a pan/tilt camera. A prototype of the projection system has been built using a pan/tilt mirror that directs imagery generated by a standard LCD projector. To capture user interaction with the projected graphics, the images gathered by the camera are processed using computer vision techniques. In the current prototype, it is possible to determine the user's hand position in referennce to a surface and to detect simple, touch-like gestures, effectively mapping the user's hand gestures into mouse-like events.

For instance, we have developed a computer game, called "Fro…og!", where a computer generated image of a frog initially appears "sitting" on the side of a file cabinet (see Figure 1). If a player tries to catch it, the vision system detects the player's hand approaching and then shows an image of frog jumping (see Figure 1). This is followed by moving the image of the frog to another surface, for instance the top of a table by appropriately steering the mirror. All this happens in an environment under normal lighting conditions, ultimately creating the illusion that the frog jumps around the space to avoid being caught by the player.

Our goals, however, go beyond transforming the real world into a magical entertaining place. In particular, we envision the creation of seamless interfaces to devices on surface of walls and everyday objects such as tables, chairs, couches, doors, cabinets, and carpets. Currently, only very "special" surfaces such as TVs, monitors, and projected screens support the creation of computer graphics, and usually require external devices such as mice and joysticks to make them interactive. In our paradigm, any surface is regarded as generic support for interactive graphics. By projecting imagery and detecting user interaction with a camera, it is possible to transform every surface into a virtual touch-screen.

These ideas have been explored in three application prototypes



Figure 1. Trying to catch a projected frog.

Amazon Ex. 1012.0204

described in the paper. The first in an attentive/interactive office space.The second is in the context of augmenting the environment of an assembly task by projecting interactive graphics directly onto the surfaces of the objects involved. Ultimately we aim to realize Mark Weiser's paradigm of ubiquitous computing (see [1]) without having to change the fabric of the world by embedding touch-screens on every surface. The thir application is the ubiquitous computer game "Fro…og!".

This paper starts by comparing our approach and solutions with previous work, followed by a description of the technology used in the prototypes we have built. Although initial ideas on the ED-projector have been reported in a previous publication [2], this paper presents the device comprehensively including mathematical formulation, computer vision based interaction, and three implemented applications. Specifically, this paper describes the mathematical formulation of the method for correcting oblique distortion and presents two new calibration methods for the first time. Further, in this paper we describe an integrated, working computer vision system that we implemented to track hands and detect touches, as well as a way to structure the interface to application software. While [2] ellaborates only on conceptual mock-ups, this paper describes three implemented applications using the ED-projector, followed by a discussion of how novice users — in one case, more than 600 of them — have experienced and interacted with the application prototype. We conclude by examining other applications that can take special advantage of the ED-projector and discuss the challenges these new interfaces and applications create for HCI research.

2 RELATED WORK

In his seminal paper that defined the field of ubiquitous computing, Weiser [1] envisioned computers embedded everywhere — ubiquitous and transparent. In his view, interaction with computers would be accomplished both through direct mechanisms linking sensors and actuators and by the widespread presence of cheap interactive screens (referred as "pads") embedded on the surfaces of walls, furniture, and objects. Since the publication of that paper, different devices and methods to implement the functionality envisioned for Weiser's pads have been proposed. In general, those devices and methods can be divided into four categories: head-mounted, portable, embedded, and projection-based.

A good survey of the current research on head-mounted devices and their uses in ubiquitous computing is provided by Barfield and Caudell [3], describing both wearable computers and the more traditional approach of off-body computer power. We see two main challenges in this approach First, in order to really connect to the physical world, it is necessary to precisely register the position and attitude of the head of the user with the environment in real time. An examination of the state-of-art of the research in the area shows that this is currently achievable only in very constrained environments (see, for example, Raskar et al. [4]).

The second challenge with head-mounted methods for ubiquitous displays is that they do not work well in social contexts and collaborative tasks. If multiple users are looking at the same surface (for instance, a whiteboard), it is necessary to render exactly the same graphics on each user's head-mounted display simultaneously. This creates considerable difficulties in terms of accomodating different resolutions, rendering speeds, etc. of different brands and models of head-mounted displays. In practice, it is very hard to guarantee that all participants are seeing exactly the same images.

The second method being studied to create ubiquitous displays utilizes portable devices such as laptops, PDAs, and mobile phones. Like head-mounted displays, portable devices do not work well in social and collaboration tasks (except in very unusual conditions such as in multi-player games [4]), and are a nuisance to carry and power.

The third method involves embedding screens or similar interaction devices into the objects themselves, as proposed initially by Weiser [1] and more recently by Ishii [5]. However, even if screens became extremely cheap, this approach requires a widespread change in the way everyday objects are manufactured and installed. Power and network access (or wireless bandwidth) has to be provided to even the simplest object in a home or workplace. In other words, this approach requires a drastic change of the very fabric of the real world.

Our approach follows the fourth method of creating ubiquitous displays that is based on projection devices. Interestingly, the use of projectors to change and augment reality was pioneered by media artists such as Tony Oursler [6] and Michael Naimark [7]. The use of projectors to access computers have been initially proposed by Bolt [8] and Wellner [9]. Recently, more radical applications of projectors have been proposed by Morishima et al. [10] that enhance the face of an actor, and by Raskar et al. [11] to change the color and texture of real objects such as vases and architectural models. Although our system does not currently track moving surfaces as in the former example, or project onto 3D objects as the latter, it can clearly incorporate both features in future versions.

Deviceless interaction with projections was pioneered by Krueger [12], using camera-based systems to detect the users' hand and body. The fact that most vision techniques put strict constraints on lighting conditions and on the background surfaces has restricted the use of these systems to entertainment applications such as Keays and McNeil's *metaField* [13]. To avoid lighting constraints, Omojola et al. [14], Ju [15], and the RED project at Xerox Parc [16], among others, embedded sensors on the surface where the images are projected. As in the embedded approach described before, the widespread adoption of sensor-loaded surfaces requires a dramatic change in the objects of the real world. Underkoffler and Ishii [17] and Rekimoto [18] used visual tags to detect the position and the identity of the objects on a surface and to overlay the interaction based on these attributes. To avoid the need of tagging all objects in an environment as in those two cases, a possible option is the use of object recognition techniques (such as those described by Schiele and Crowley [19]). However, those techniques currently work only in very constrained situations. We believe, though, that incorporating object recognition capabilities into our system is an important direction to follow, and we plan to do so forspecific applications.

Vision-based gesture recognition has received a great deal of attention in recent years [20], but only a little of that work has attended to the problems of gesture recognition in the context of a projected display. Because the projected image drastically changes the appearance of the hand, most approaches to hand detection are based on either background subtraction or motion (frame-frame differencing). Our experience agrees with that of von-Hardenberg and Berard [21] in that background subtraction does not give reliable results in the presence of strong projected images.

Gesture-based interaction methods tend to fall into one of two catagories: either variations of a point-and-click paradigm [21, 22] or application-specific pose or motion gestures [23]. Both these approaches have disadvantages. The point-and-click approach is often not well suited to hand pointing due to limited resolution and the lack of a natural "click". On the other hand, a proliferation of application-specific gestures will make for a complex gestural interface that a user will have to learn and execute largely without feedback, as discussed by Kjeldsen in [24].

Amazon Ex. 1012.0205

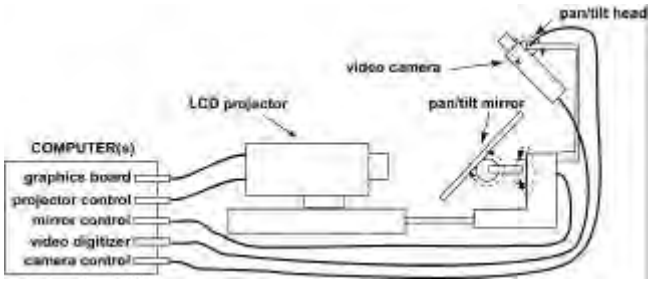


Figure 2. Diagram of the components of the ED-projector.



Figure 3. Prototype of the ED-projector: a) ED-projector without the camera mounted; b) a detail of a complete prototype, including the pan/tilt camera.

A major difference between our approach and the cited work on the use of projection-based devices to create ubiquitous interactive displays is *steerability*. By directing the projection and vision subsystems to a relevant position, it is possible to create new "interactive" surfaces in a space without any hardware installation, addition, or change. Unlike Raskar et al. [25], who use mirrors to increase the resolution of multiple-projector installations, we steer the image from a projector to realize interactive displays everywhere in the workspace. By doing so, we opportunistically use the surfaces of the objects present in the environment. We believe that steerability is a key feature that creates a real alternative to realize the promises of ubiquitous computing. However, it has constraints and technical challenges of its own, such as obstruction, oblique distortion, and surface texture interference. These issues are discussed in the next section where solutions to some of these problems are proposed.

3 THE ED-PROJECTOR

Our approach to ubiquitous interactive graphics uses a device — the *Everywhere Displays projector* (the *ED-projector*) — that can steer graphic displays onto any surface and can also sense user interaction with these steered graphical displays. The ED-projector is composed of an LCD projector, a computer-controlled pan/tilt mirror, and a pan/tilt/zoom camera. The projector is connected to the display output of a host computer that also controls the mirror and the camera (see Figure 2). The video output of the camera is sent to a digitizer board in the host computer where the imagery is processed, as described later.

Figure 3 shows a prototype of the ED-projector built with off-the-shelf components — a rotating mirror used in theatrical/discotheque lighting, a standard LCD projector, and a pan/tilt camera. The projector's light can be directed anywhere within a cone defined by approximately 60 degrees in the vertical axis and 230 degrees in the horizontal axis. The camera has a maximum vertical viewing range of 88 degrees and a horizontal range of 249 degrees, and is thus able to view the projected displays as they are steered around.

In general, a projector needs to be able to project a white pattern approximately 10 times brighter than its surroundings to create the illusion of contrast. Our current prototype employs a 3000 lumen projector, which we have found to be sufficiently bright in typical home and office lighting conditions to provide good contrast in the projected displays. Of course, the perceived brightness and constrast is heavily dependent on the color and material of the projected surface. Although our best results involve projection on white surfaces, we have also obtained good results projecting on carpets, black objects, and even translucent backlit surfaces. However, highly textured surfaces such as natural wood and surfaces with high reflectance are, in general, less suited for projected displays.

The three main issues with the ED-projector are 1) correction of the projected image to compensate for distortion due to oblique projection and surface characteristics; 2) sensing the user interactions with the projected displays; and 3) a system software structure that enables applications to easily define interactive display surfaces. These issues are discussed in detail in the rest of this section.

3.1 Correcting the Projected Image

Standard LCD projectors are designed to project light in a direction orthogonal to the projected surface. In the ED-projector, the pan/tilt mirror can deflect the light to multiple surfaces in the room, but most of the times the projected surface is oblique to the direction of projection. The result is a distorted image.

To correct the oblique projection distortion, we pre-warp the image to be projected to compensate for the distortion. It is always possible to correct the projected image, as long as the projected surface does not have a geometry that occludes the cone of projection (see [2] for details). However, the process of warping the image normally uses only a fraction of the projector's image plane, causing loss of resolution. We will discuss this issue later; first, we address the problem of determining the warping function that corrects the oblique distortion.

Although there are many different ways to determine and implement the warping function, it is important to use a method that allows fast computation. We choose to implement it using 3D computer graphics hardware to texture map the images to be displayed onto a mesh in 3D space. If the mesh is appropriately warped and/or positioned in 3D space and the virtual camera correctly positioned, the projection of the resulting rendered image is free of distortion.

The basic structure of the problem, in the case of planar surfaces, is illustrated in Figure 4. We start by defining the *application plane*, a coordinate system where images to be projected are positioned within the area of a rectangular mesh M, and the *projector image plane*, the coordinate system associated with the LCD matrix of the projector. We also define a *projected surface* to be the real world surface on to which images are projected. In addition, we have the camera image plane, the coordinate system associated with the CCD matrix of the camera.

The goal is to define a transformation H from the application plane into the projector image plane so the transformation that maps the support mesh M into the projected surface is the identity I, and, therefore, appears free of distortion from an observer orthogonal to the projected surface. Since the process of projecting the image can be modeled as a projective transformation P (as

Amazon Ex. 1012.0206

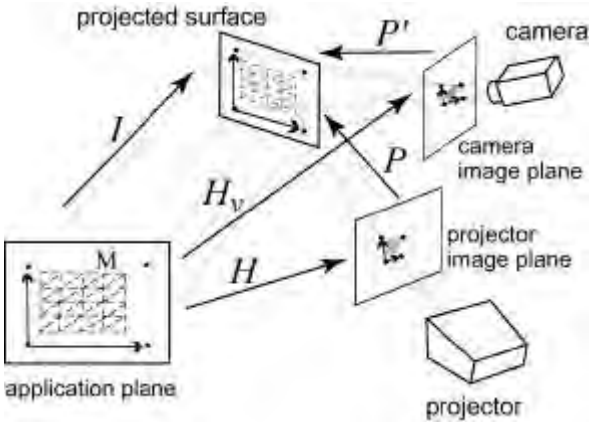


Figure 4. The transformations between the application plane, the projector's image plane, the camera's image plane and the projected surface.

observed in [26, 27]), we can represent the relationships among the planes as matrices in planar homogeneous coordinates, $I = PH$. Since H and P are projective transformations and therefore invertible, we obtain:

$$H = P^{-1}$$

Similarly if P' is the projective transformation between the camera image plane and the projected surface, and H_v is the transformation from the application plane to the camera image plane, we have $I = P'H_v$. Although there is no easy direct way to compute the projection function P, we developed three different methods to determine the actual value of H for a given surface. Of course, H and H_v are different for each planar surface in an environment.

1st method: Using an Equivalent 3D Virtual CG World

This method, first described in [2], is based on the fact that, geometrically, projection is the inverse of camera viewing. So, it is possible to model the projection by simulating the camera view by rotating, scaling, and viewing the support mesh M in 3D space.

Given the n mesh points in 2D projective space, $m_i = [m_i^x \; m_i^y \; 1]$, $1 \le i \le n$, let M be the matrix of all n points, $M = [m_1^T m_2^T \ldots m_n^T]$. This mesh can be "positioned" in the XY plane of the 3D space by multiplying it by the matrix

$$K = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad KM = \begin{bmatrix} m_1^x & m_2^x & \ldots & m_n^x \\ m_1^y & m_2^y & \ldots & m_n^y \\ 0 & 0 & \ldots & 0 \\ 1 & 1 & \ldots & 1 \end{bmatrix}$$

We can then simulate the proces of projecting the mesh as the inverse of the process of "viewing" the appropriately positioned mesh in 3D space. That is,

$$P^{-1}M = P_f R_x R_y S_x S_y T_x T_y R_z K M$$

where R_x, R_y, R_z are rotation matrixes in 3D homogeneous coordinates, S_x, S_y are scale matrixes, T_x, T_y are translation matrixes, and P_f is the 3 4 projective matrix with focus point f. Since $H = P^{-1}$, we obtain

$$H = P_f R_x R_y S_x S_y T_x T_y R_z K$$

Notice that all the matrices that compose H, with the exception of K, correspond to standard projection parameters of computer graphics boards. Therefore, to implement this method it is only necessary to determine these individual parameters, load them into the hardware, and command the graphics board to render the scene.

To determine the projection parameters for a given surface we start by texture mapping a calibration pattern to the mesh M. Starting off with default values for the projection parameters that render the mesh on the center of the projector image plane without warping (i.e., $H \approx I$), we iteratively change each of the projection parameters until the projection of the pattern on the projected surface is observed to be identical to the original mesh. We then record these parameters and reload them whenever the ED-projector is aimed at that surface. Although conceptually simple, this iterative process of determining the right projection parameters is typically very time consuming.

2nd method: Manual Determination of Corresponding Points

This and the following method are based on the fact that projective transformations can be completely determined if the corresponding coordinates of four points in each 2D projective space are known (as shown by Faugeras [28]).

In this method, four points $a_i = [a_i^x \; a_i^y \; 1]$, $i = 1,2,3,4$, are directly rendered on the projector image plane. We then iteratively move these points over the projector image space with a pointing device such as a mouse until their projections overlap the location of four known points on the projected surface, $b_i = [b_i^x \; b_i^y \; 1]$, $i = 1,2,3,4$. For greater accuracy, we can choose these four points to be the corners of the mesh M.

It is easy to see that if the projected points are aligned with the projected surface points then $PA = B$, where $A = [a_1^T \; a_2^T \; a_3^T \; a_4^T]$ and $B = [b_1^T \; b_2^T \; b_3^T \; b_4^T]$. By computing the pseudo-inverse,

$$PA = B \rightarrow PAA^T = BA^T \rightarrow P = BA^T \left(AA^T\right)^{-1}$$

Since $H = P^{-1}$, we obtain $H = AA^T (BA^T)^{-1}$. To implement this procedure, we simply construct a warped mesh M' in 3D projective space by computing $M' = KHM$, texture map the image to be projected on it, and use the graphics engine to render it in real time.

As we see, the calibration procedure of this method is quite simple and fast, requiring only a few minutes for the manual alignment of the four projected points. Although similar to the method proposed by [29], here the use of the 3D mesh M and the computer graphics hardware yields much faster rendering speeds.

3rd method: Using a Camera and a Paper Pattern

In this method, we employ the ED-projector camera so we can avoid any manual calibration procedure. In [30], Yang et al. have proposed a method to automatically determine the warping function using a projector/camera pair. Unfortunately, their approach requires precise 3D calibration of the relative attitude of the camera to the projector and, in practice, a considerable baseline distance between them. In particular, it does not work with the ED-projector that aligns the projector and the camera axis as much as possible to avoid user's obstruction of the camera's view.

Instead, we propose here a method that employs a paper pattern positioned on the surface to be calibrated so it is aligned with the desired coordinate system. No relative 3D or 2D calibration is necessary, although we assume that the camera has been positioned in such a way that it has a complete view of the projected surface. Under these conditions, we can define a projective transformation P' between the camera image and the projected surface (see Figure 4). This transformation can be determined automatically by processing an image of the pattern and extracting the camera image coordinates of the corners, $c_i = [c_i^x \; c_i^y \; 1]$, $i = 1,2,3,4$. Assuming the coordinates of the

Amazon Ex. 1012.0207

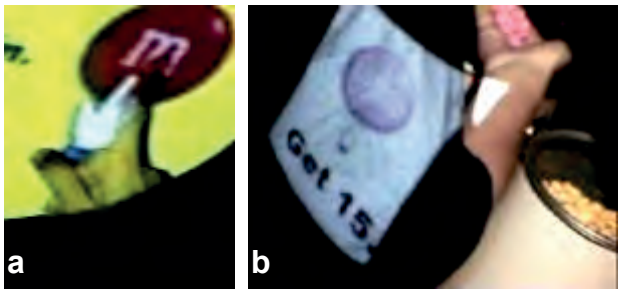


Figure 5. Difficulties faced by the vision system: a) the user's hand is partially hidden by the strong projection; b) the interactive surface is occluded by the user's shoulder as he moves to gather objects from inside a bucket.

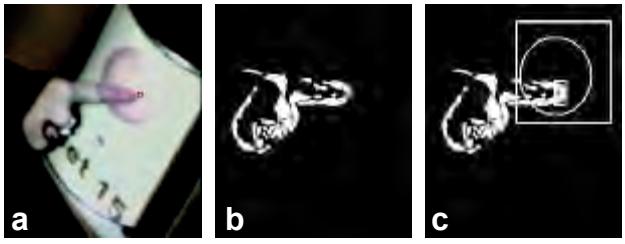


Figure 6. Detecting a touch: a) camera view of an interaction with a bucket; b) image difference data; c) overlay of search region (square), button active area (circle), and the fingertip template shown at the pointing location.

pattern corners to be identical to the mesh M in application space, $b_i = [b_i^x \; b_i^y \; 1]$, $i = 1,2,3,4$, we obtain $P'C = B$, and computing the pseudo-inverse,

$$P' = BC^T \left(CC^T\right)^{-1}$$

To complete the computation of the transformation H, it is necessary to determine the correspondence between the projector and the camera image planes. This can be done by projecting a pattern with four non-colinear points rendered on the projector image plane $d_i = [d_i^x \; d_i^y \; 1]$, $i = 1,2,3,4$, and obtaining four points on the projected surface, $e_i = [e_i^x \; e_i^y \; 1]$, $i = 1,2,3,4$. Then, by image processing the image of the pattern, it is possible to determining the coordinates of its four corners on the camera image plane, $f_i = [f_i^x \; f_i^y \; 1]$, $i = 1,2,3,4$.

Since $E = [e_1^T \; e_2^T \; e_3^T \; e_4^T]$ corresponds both to the projection of $D = [d_1^T \; d_2^T \; d_3^T \; d_4^T]$ by the projector, $PD = E$, and to the captured image of the pattern, $P'F = E$, we obtain $PD = P'F$. Computing the pseudo-inverse, we obtain $P = P'FD^T(DD^T)^{-1}$ and therefore

$$H = DD^T \left(P'FD^T\right)^{-1}$$

As in the previous method, this transformation is implemented by computing a warped mesh $M' = KHM$.

One problem with all these three distortion-correcting methods is that they project displays with resolution lower than the projector's resolution. The distortion-correction process normally fits an irregular quadrangle into the 4:3 viewing area of typical displays. A considerable amount of display area, and thus resolution, is lost in the process. We have employed standard 1024x768 XVGA projectors in our prototypes. Due to the loss of display area created by the distortion correction process, the final display resolution has approximately corresponded to VGA, i.e., 640x480 pixels.

3.2 Sensing Interaction on Any Surface

To support the ability to interact on any surface we use a pan/tilt camera feeding custom image processing and gesture recognition engines. This process has been detailed in another publication [31], but will be outlined here for completeness.

Image processing in this domain is both simplified and made more difficult by the strong light of the projected image. It is simplified in the sense that the lighting environment is very consistent because the dynamic range of the projection swamps all but major changes in ambient light. It is made more difficult because the appearance of the user's hand can be severely distorted as it passes through the projected image (see Figure 5.a).

Although parts of the hand can be hard to detect in a static image, by looking at the differences between adjacent images in the video stream, we can still detect the shape of a moving object as shown in Figure 6.a and b.The use of motion for segmentation instead of background subtraction techniques (such as [32]) also makes it easier to accommodate the variability in the projected image caused by the texture of the surface it is projected on. This motion mask forms the basis of subsequent image understanding.

To provide interaction in a flexible and consistent manner we use the notion of *interface widgets*. Each widget generates computer or application events in response to a particular user motion in a region of the image. We currently use three types of widgets: *buttons* that respond to touch-like motions; *sliders*, 1-dimensional tracking regions, and *touch pads*, which are 2-dimensional tracking regions.

These three widget types rely on tracking the user's fingertip. Convolution template matching is used to find fingertip candidates in the motion mask image described above. The candidates are evaluated using spatial clustering and domain specific heuristics, such as the distance from the user, to determine the best estimate of whether there is a fingertip in the image and where it is currently pointing. The trajectory created by this point over a sequence of video frames is examined by each widget to determine when it should trigger an event.

Touch pad widgets smooth the fingertip trajectory and map it into the touch pad coordinate system. Motion events are only generated when the fingertip is within the touch pad's borders. Sliders are simply a 1-dimensional version of touch pads.

Buttons, however, are more complex widgets. What buttons attempt to detect is when a surface in the environment is touched. Inferring the distance of the fingertip to the surface using the shadow proved to be difficult and unreliable. Fortunately, people tend to interact with real-life buttons by moving in a characteristic forward/pause/backward pattern, and this seems to carry over well to projected buttons. Our implementation of buttons takes advantage of this movementpattern.

To determine when a user has touched a button we look for a roughly linear sequence of fingertip locations that persists for 250ms or longer, with speed within a set range, that either ends or begins with a pause (i.e. the fingertip template did not find a match in the motion mask). A button generates its event when the pause in such a trajectory occurs within it (see Figure 6.c). In other words, we are looking for the user's fingertip to travel in one direction for ¼ second and stop within the button, or start moving from the button and move in a consistent direction for ¼ second. These approximations for the motion of a fingertip in the moments before and after it touches a button seem to be universal: almost every button touch either starts or ends that way. Interestingly though, many button touches only have one or the other of these motions. Of course, adjustments are needed in the case of people with tremors or other motor disabilities.

Amazon Ex. 1012.0208

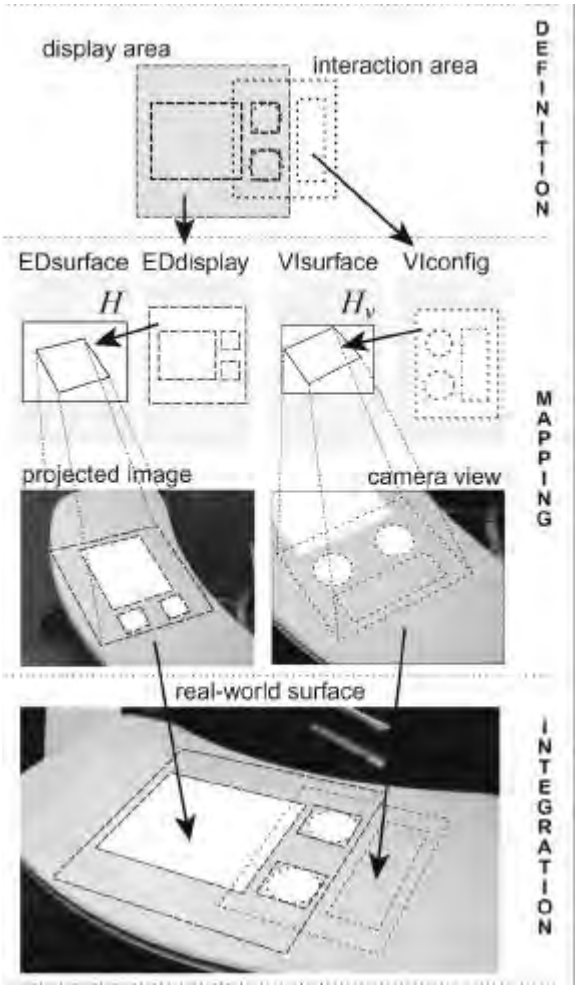


Figure 7. Process of creating an interactive surface on a table.

One of the interesting problems in this domain occurs when users occlude buttons as shown in Figure 5.b. Our experiments with users, even novices, have shown that they naturally avoid obstructing the projection if they are attending to it, but they are quite bad at staying out of the way when they are working with the physical world instead. Buttons have inherent resistance to triggering in these types of situations because any motion must pass several tests before it generates an event. It must resemble a fingertip for several frames; during that time it has to travel with a consistent direction and velocity and stop within a button. Such a sequence sometimes occur as a result of a chance combination of errors, contributing to a somewhat higher false positive rate for some buttons, particularly those which the user occludesfor a period of time rather than just passes through.

By using both the shape of the moving region as well as an analysis of its trajectory, this touch detection method gives far better accuracy than using either one alone. Because it relies directly on image data, rather than building an explicit model of the hand, it is computationally efficient enough to run at high frame rates on standard hardware.

Widgets can be combined in various ways in order to perform a complete interaction with the user. However, configurations of widgets are defined independently of the surface where they are used. That is, if an application interface using a set of buttons is moved to different surfaces, it is not necessary to define each button for each surface. Instead, the location of a widget within the camera image is determined at run time by appropriately mapping the configuration onto the camera image plane. In this case the surface refers to a set of camera parameters. These parameters, like those of the projector, have to be calibrated in advance. In particular, the calibration process has to determine the transformation H_v between the widget coordinate system and the camera image (as shown in Figure 4). Calibration of surfaces is performed using methods analogous to the 2nd and 3rd methods of projection calibration, although information about the expected location of the user, and the size of their hand, has to be obtained during an extra, special calibration step (see [31]for details).

3.3 Integrating Displaying and Sensing

The two preceding sections described the techniques and subsystems that provide the display (output) and sensing (input) functionality of the ED-projector. This section discusses the integration of these separate subsystems in a manner that affords application development and human interaction. One of our primary objectives here is to enable a specific user interaction to occur on any calibrated surface in a room without requiring the customization of this interaction for every surface. This leads to a clear separation of the abstract definition of an interaction from the actual surface upon which it occurs.

Our widgets are constructed using the standard Model View Controller (MVC) design pattern structure [33]. In our case, the projector subsystem is responsible for rendering the *view*, while the vision subsystem provides the *controller* input events. The *definition* layer of Figure 7 shows the separation of the display and interaction areas that are reponsible for the definition of the view and the interaction respectively. Display and interaction will often overlap as shown by the square overlapping buttons on the right of the *display area*. Provision is also made, however, for allowing interaction to occur in a region where nothing is displayed, as shown in the large rectangular region on the right of the *interaction area*. Notice that the display area here corresponds to the mesh M of the application plane decribed before (see also Figure 4).

Every interaction surface in a physical space is named and initially calibrated by the projection and vision subsystems yielding the transforms H and H_v respectively. H corrects the projected image while H_v provides the transformation to the camera image, for a specific surface as described in the prior sections. Figure 7 demonstrates the flow of definitions for display images and interaction data (widget data) from the *definition* layer to the *mapping* layer. During run-time, the system software automatically maps (through the warping functions H and H_v) the *definition* layer data into the projected image and the camera view associated with the current surface. The result is a projected image that generates application events when user gestures occur in the areas defined in the camera view, as shown in the *integration* layer of Figure 7.

4 PAINTING WITH INTERACTIVE GRAPHICS

To better understand how to use projected interactive graphics and how users relate to them, we have designed and implemented three application prototypes using the ED-projector. In the first application, an office space is augmented with dynamic wallpapers, visual notifications of messages, and computer access from multiple surfaces. The second prototype explores an assembly task where the instructions are projected directly onto the surfaces of the objects involved in the task. The third prototype is the "catch the frog" game described in the introduction to this paper.

Amazon Ex. 1012.0209

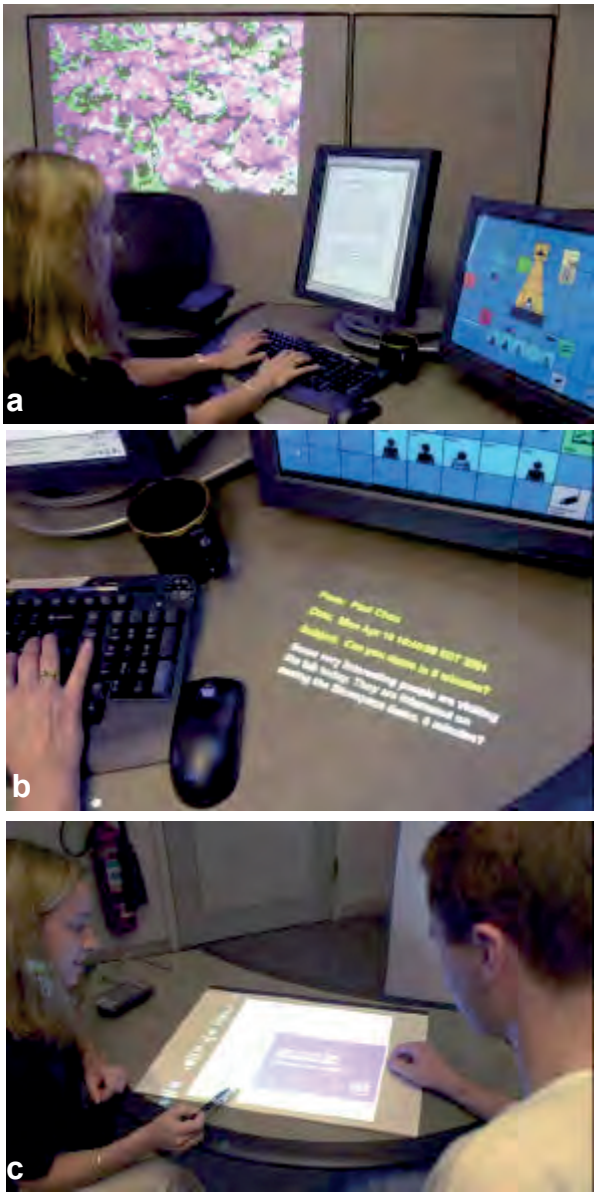


Figure 8. Uses of the ED-projector in an office application: a) to create dynamic wallpapers; b) to provide silent notification of e-mail; and c) to create a computer desktop on a working surface.

4.1 An Office Application

The first prototype application in which we have deployed the ED-projector is a futuristic personalized and context-aware workspace. The workspace has an 8' by 10' footprint and incorporates several sensors for measuring ambient lighting, temperature, humidity and noise level. Users and visitors to the workspace wear active badges that facilitate presence detection and identification. The desk chair is equipped with a pressure sensor connected to a wireless micro-controller, which detects if a person is sitting in the chair. This environment incorporates an ED-projector, strategically located to maximize coverage of the workspace. In this setting, we have used a non-interactive version of the ED-projector to facilitate greater personalization of the environment and to free its users from the confines of the desktop displays.

The presence of an ED-projector in the office tranfigures the space into an extended graphical display and enables the user to have a favorite painting on a wall, receive important notifications on surfaces that catch their attention, or use a tabletop as a display when collaborating with a colleague (see also the video figure). Figure 8.a shows an example of a dynamic wallpaper created by the ED-projector. The user can similarly have a virtual window that is connected to a live webcam looking at her favorite scenery (effectively making it a window office) or monitoring her child at home. The ED-projector can automatically switch to a neutral, less personal image when a co-worker comes to the office, based on the active badge.

Figure 8.b illustrates an urgent notification to the user appearing on a surface close to the user's current location. Here, the smart office detects that the user is seated at the desk and steers the notification right to the surface of the desk. By determining the position of the user, the system is assured that the user will see the urgent message without having to rely on the traditional sound alerts which tend to be disruptive to people working in a shared environment.

Another example of the utility of a steerable display is seen in Figure 8.c where two colleagues have created a computer desktop display on the working surface of their choice. Here, interactive capabilities, provided by a camera system, are planned to be incorporated to facilitate collaborative work.

4.2 Augmenting an Assembly Task

The ED-projector can be seen as a device for augmenting reality – a beam of light that can add information on to real world objects by highlighting them, adding text or images, flashing etc. For example, the ED-projector can show the user an object they are searching for, or indicate the procedure for an assembly task, and even where to place a part in an assembly.

We developed a prototype assembly task to highlight these themes of augmented reality and interaction anywhere. In this prototype application, which was experienced by hundreds of novice users in a major technical exhibition, the object to be assembled is a picture made of M&M's (multi-colored sugar-coated chocolates) where each M&M is regarded as a pixel of the picture.

The theme of interactivity anywhere is demonstrated several times in this prototype application (see also the video figure). Figure 9.a shows a white fabric on a table transformed into an interactive menu for color selection. The user simply touches the color of his choice to select it. Figure 9.b shows a "clickable" button appearing on a paint bucket that contains the M&M's of the selected color. Figure 9.b and c highlight the augmented reality and user assistance aspects of the system. In Figure 9.b the system highlights the bucket that contains the M&M's of the selected color, and also provides additional information on how many M&M's to pick. In Figure 9.c, the system points the exact places on the picture board where the M&M's should be placed. Figure 9.d shows another combination of interaction and augmented reality: the user moves her hand over the surface to interactively reveal the rest of the picture being built.

More than 650 people participated in this augmented assembly over 6 days and completed 4 different M&M pictures. Considering that the vision system and the projection system were integrated for the first time only in the weeks preceding the exhibition, the combined system worked remarkably well. A sample with 130 consecutive users with 621 button touch events (touching gestures or false detections) yielded correct detection of touching gestures in 81% of the events, with 12% false negatives and 7% false positives. If the bucket events are excluded from the count, the

Amazon Ex. 1012.0210

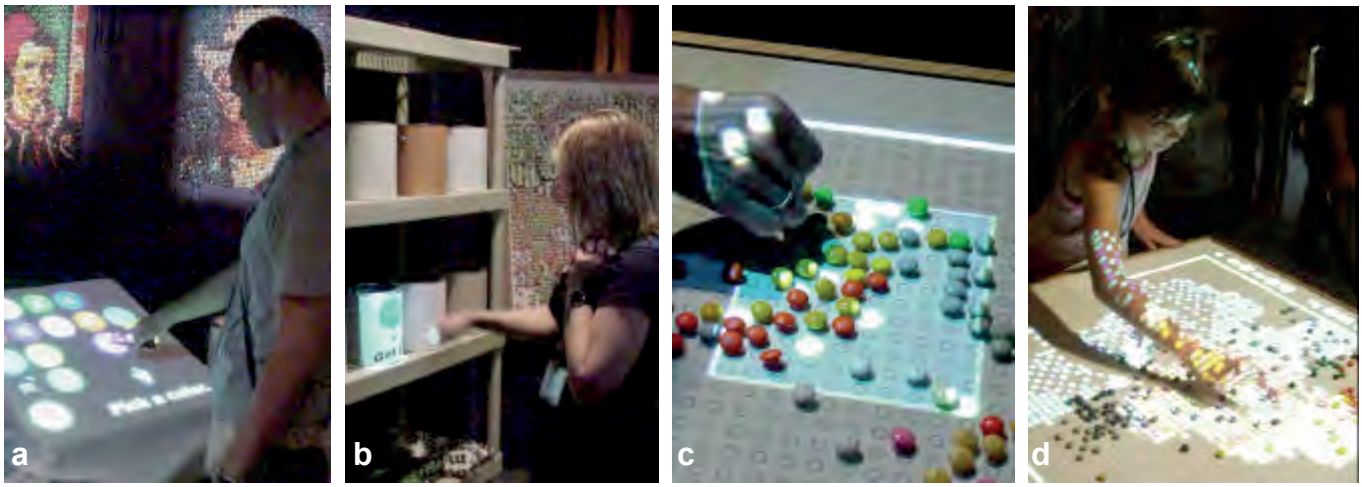


Figure 9. The M&M picture assembly task: a) choosing a color; b) "clicking" a bucket after getting the M&Ms; c) placing the M&Ms on the highlighted areas; and d) "finger painting" to reveal the complete picture.

performance exceeds 89%. A detailed account of the reasons why buckets yielded more errors is given in [31].

The experiment also revealed that users are readily able to relate to the notion of steerable interactive graphics. We also observed the value of the projection creating a shared experience between the users and the people observing the demo: in many cases, users learned the interaction process by first watching other users; and frequently, observers helped users suggesting what to do next. As mentioned before, a major advantage of projection over goggles is precisely this "social" nature of projection.

An interesting observation is that our users were less hesitant to interact with the fabric-covered interactive color selection menu (Figure 9.a) and the M&M placement board (Figure 9.c,d) than with the paint buckets (Figure 9.b). When they went to a bucket to pick up M&M's, many of them did not realize that they had to click the big "done" button on it (see also Figure 6.a). We had expected that, since this was the third occasion for touch interaction in the demo, users would quite naturally press the button on the bucket as instructed. Most of the time, this was not the case. To help the users in that situation we often gave them verbal instructions such as "Click the bucket," aiming to make clearer that there was, in fact, a button on the bucket.

4.3 "Fro…og!"

The ED-projector's unique characteristic of being able to quickly move the projected image to different surfaces is being explored in a computer game for young children. In this game, called "Fro…og!", an image of a frog is projected on a surface. Whenever the child tries to touch the frog, the camera detects her approach, and the frog "jumps" to another surface, that is, the projector moves the image to a wall or the top a chair.

Figure 1 shows two pictures of a user playing with the current prototype of "Fro…og!". The basic idea is to define surfaces in the environment where the animated frog can jump to. Whenever the user's hand approaches the frog, the next surface is randomly chosen, and an image of the frog jumping in that direction is briefly shown. After landing on the new surface, a simple animation of the frog jumping up and down is projected, so it becomes easier to find the frog in the environment.

In the next development phase of this game we plan to integrate it to a system that tracks the position of multiple children in the room, so the frog can jump to surfaces which are far from most players. Tracking is also particularly important to avoid projecting the frog directly onto the players which, in this scenario, can easily break the illusion of the game.

5 DISCUSSION

Having shown how we implemented a working prototype of an ED-projector and deployed it in three different applications, it is interesting to explore other possible applications and scenarios, and discuss how to design efficient interfaces for them.

5.1 Other Applications

We see two classes of applications of the ED-projector to create ubiquitous interactive graphics. The first involves the creation of computer interfaces anywhere, while the second deals with augmented reality in a variety of public and private spaces.

In the first class of applications, we regard the ED-projector as a generic input/output device that can replace, in many situations, current displays and interactive devices by creating computer desktop-like interactive displays on non-tethered surfaces. For example, a desktop application can be projected directly on any surface, as previously suggested by Wellner [9] and as demonstrated in subsection 4.1 . Unlike the interactive whiteboard described in [34], the projected application can be easily moved around the room, for instance, from a whiteboard on the wall to the top of a desk for more detailed reading. Similarly, the ED-projector can be used to bring information to the physical location where it is used or needed. For example, a database application managing reports can be projected on top of the file cabinet with hard copies of the reports.

In the second class of applications the ED-projector can augment reality by pointing to physical objects, showing connections among them, and projecting patterns to indicate movement or change in the real world. The prototype assembly task described in subsection 4.2 is a typical example of this class of applications. Another example applies to a library in which directions are provided to a user looking for a book by projecting arrows on the floor, walls and shelves, and finally by highlighting the desired book.

Steerable graphics produced by the ED-projector can also aid in collaborative work by providing large displays on convenient surfaces, and allowing easy reconfiguration of meeting spaces for different functions and teamwork styles, in similar ways to the office project described in subsection 4.1 .

The ED-projector can also be used to provide computer and information access in environments where traditional displays are

Amazon Ex. 1012.0211

difficult to install, secure, or maybe unsafe to operate, such as public spaces and areas subject to harsh environmental conditions. The device also permits an interactive display to be brought to the proximity of a user, eliminating the need for the user to approach a traditional, fixed display. In particular, the ED-projector can facilitate the access and use of computers by people with locomotive disabilities. For instance, it can project an interactive display on a hospital bed sheet without requiring the patient to contact any device.

We also see the ED-projector as a potential enabler of a new generation of games where interaction happens not in the virtual world but in the physical world such as the "Fro…og!" game described above. Unlike games based on phones and portable computers [4], the use of the ED-projector provides high-resolution displays where characters and fantastic objects can move from surface to surface, creating a game that surrounds the players. Using video input, various user actions can be detected depending on the game needs, for example, hand gestures, body movements, foot action, and/or facial expressions. Above all, a single ED-projector installed on the ceiling of a living room or an arcade, using just software commands, can completely reconfigure the space to create different games according to the interests, age, and motor skills of the players.

The scenarios examined so far are chiefly concerned in creating ways for humans to command computers and/or obtain information. However, we see the projecting of interactive graphics ubiquitously as a way to provide computer agents the ability to "reach and touch" the real world and the people that inhabit it. By using projected patterns, it is possible to make computers not only interact with people, but act and coordinate them. For instance, a computer agent can control people in a line by projecting red lines and green arrows around the people on the line. A steerable projector system is, thus, able to create a "body", composed of light, for an agent. From this perspective, we may be creating an interface to the real world for computers.

5.2 Human Interface Issues

As foreseen by Weiser [1], ubiquitous interactive graphics do not seem to follow the standard paradigm of interaction based on the desktop metaphor. In our M&M demonstration, we had the opportunity to observe more than 650 novice users experiencing a scenario of true ubiquitous computing, in the context of an assembly task. On one side, we observed that they quite naturally accepted the idea of augmenting reality with information directly projected on objects. Moreover, we saw they naturally change their body's position to minimize problems like the shadow of their arms and body over the board. People seemed not only to accept projected graphics naturally but also to compensate for its limitations.

However, when trying to click a button, the users almost always applied pressure onto the projected surface with their fingers. Also, whenever a button touch was not detected by the vision system, the most common reaction was to press again with more pressure and, if still not successful, using the whole hand instead of a finger. It seems that people expect buttons, even virtual buttons, to react to pressure and not to touch, and to fail if enough pressure is not applied on them.

This observation clearly illustrates the limitations of the hand-as-a-mouse paradigm used in the M&M demo. Although it is easier for users to relate the ubiquitous experience to their everyday contact with desktop interfaces, the metaphor proved to be misleading to the true nature of the interface. Unlike a mouse that has appropriate haptic feedback for clicking gestures, touching a projected button is devoid of feedback movement.

Similar observations prompt us to believe that a new paradigm for interaction with projected screens must be developed, with new kinds of widgets, tuned for the natural capabilities and limitations of projection and gesture-based interaction, just as the point and click paradigm is tuned for the characteristics of a physical mouse. Touch-screen-like interactions do translate better to the projected interface domain. Although our widgets are not based on surface contact, the widgets we have found useful so far (buttons, scroll bars and touch pads) are all based on motions people naturally make when touching.

One of the exciting aspects of this work is that it is an excellent environment to explore more adventurous gesture-based interaction styles. By giving the user feedback as to what gestural interactions the system is expecting and where in the environment they may be peformed, the cognitive load on the user can be reduced. Without such feedback a user would have to memorize what amounts to a set of complex dance steps to simply use a gesture-based interface.

This investigation of new human-computer interface paradigms is even more important in the case of multi-user interaction. Unlike desktops, wearable computers, or PDAs, projected interactive graphics naturally allow many users to simultaneously interact with an application. Although the applications we have developed so far are all single-user, it seems clear multi-user interaction will soon became an important area of research.

Another intriguing question is related to the different degrees of difficulty experienced by users when interacting with graphics projected on different supporting objects. Clicking a paint bucket seems to be harder to accept then clicking a table. Of course the design of our system and the simplicity of our experiment allow us neither to definitively state that the phenomenon does happen nor to draw conclusions about its causes. However, one hypothesis is that people seem to attribute different interactive capabilities to projected interactive graphics according to the nature of the surfaces themselves. Are the paint buckets "functionally fixed", that is, if something looks and acts like a paint bucket, is it less likely to be perceived as something else – namely, a touch-screen? We are currently preparing a study to determine the validity of this hypothesis.

Finally, our research so far has addressed only visual input and output. A promising area of research seems to be the application of the concept to steerability of other stimuli such as sound, using, for instance, the audio spotlights described by Pompei [35]. What kind of interfaces and applications can be created by a full multimodal steerable system?

6 CONCLUSION

This paper describes a new method and device for creating ubiquitously interactive graphics on non-tethered surfaces. We have shown three different methods to determine the warping function that corrects for oblique projection distortion and a computer vision system that tracks the user's hand over the projected images and detects touch-like gestures. We have also implemented three applications that realize ubiquitous computing concepts such as everywhere computer access and augmented reality. Based on the experience of running the M&M demo with hundreds of novice users, we have also detected a need to re-examine the current interaction paradigms, and possibly, investigate and propose new ones.

We see this work as an important contribution to realize the vision of ubiquitous computing, without asking people to wear headsets or wiring the surfaces of everyday objects. However, we do not advocate projection-based solutions as a panacea for ubiqutious applications: in many cases, it is more desirable and efficient to use head-mounted displays. A good analogy is to

Amazon Ex. 1012.0212

compare headsets and projectors to Walkmans and stereo sound systems. Walkmans are unsurpassable as a way to provide music when people are moving around or in public spaces such as trains and airplanes. They are very effective in creating a personal experience, albeit not able to create social interaction. Similarly, we keep loving stereo systems because they render music to a whole space, nurturing a collective experience that is attached to the space itself. We see similar advantages and limitations when comparing headsets and steerable projectors as a way to render interactive graphics onto the real world.

Acknowledgements

"M&M" is a registered trademark of Mars, Inc. "Walkman" is a registered trademark of Sony Corporation.

References

1. Weiser, M., *The Computer for the Twenty-First Century.* Scientific American, 1991. **265**(3): p. 94-100.
2. Pinhanez, C. *The Everywhere Displays Projector: A Device to Create Ubiquitous Graphical Interfaces*. in *Proc. of Ubiquitous Computing 2001 (Ubicomp'01)*. 2001. Atlanta, Georgia.
3. Barfield, W. and T. Caudell, eds. *Fundamentals of Wearable Computers and Augmented Reality.* 2001, Lawrence Erlbaum Assoc. 728.
4. Raskar, R., et al. *The Office of the Future: A Unified Approach to Image-Based Modeling and Spatially Immersive Displays*. in *Proc. of SIGGRAPH'98*. 1998. Orlando, Florida.
5. Ishii, H. and B. Ullmer. *Tangible Bits: Towards Seamless Interfaces between People, Bits, and Atoms*. in *Proc. of CHI'97*. 1997. Atlanta, Georgia.
6. Bellour, R., et al., *Tony Oursler*. 2001: Ediciones Poligrafa. 208.
7. Naimark, M., *Spatial Correspondence in Motion Picture Display.* SPIE Optics in Entertainment II, 1984. **462**: p. 78-81.
8. Bolt, R., *Put That There: Voice and Gesture at the Graphics Interface.* ACM Computer Graphics, 1980. **14**(3): p. 262-270.
9. Wellner, P., *Interacting with Paper on the DigitalDesk.* Communications of the ACM, 1993. **36**(7).
10. Morishima, S., et al. *HyperMask: Talking Head Projected Onto Real Objects*. in *Proc. of Multimedia Modeling (MMM'00)*. 2000: World Scientific.
11. Raskar, R., et al. *Shader Lamps: Animating Real Objects with Image-Based Illumination*. in *Proc. of 12th Eurographics Workshop on Rendering*. 2001. London, England.
12. Krueger, M.W., *Artificial Reality II*. 1990: Addison-Wesley.
13. Keays, B. and R. Macneil. *metaField Maze.* in *Proc. of SIGGRAPH'99*. 1999. Los Angeles, California.
14. Omojola, O., et al., *An Installation of Interactive Furniture.* IBM Systems Journal, 2000. **39**(3&4): p. 861-879.
15. Ju, W. *Origami Desk*. in *Proc. of SIGGRAPH'01 - Conference Abstracts and Applications*. 2001. Los Angeles, California.
16. Back, M., et al., *Designing Interactive Reading Experiences for a Museum Exhibition.* IEEE Computer Magazine, 2001. **34**(1): p. 1-8.
17. Underkoffler, J., B. Ullmer, and H. Ishii. *Emancipated Pixels: Real-World Graphics in the Luminous Room*. in *Proc. of SIGGRAPH'99*. 1999. Los Angeles, CA.
18. Rekimoto, J. *A Multiple Device Approach for Supporting Whiteboard-based Interactions*. in *Proc. of CHI'98*. 1998. Los Angeles, CA.
19. Schiele, B. and J.L. Crowley, *Recognition without Correspondence using Multidimensional Receptive Field Histograms.* IJCV, 2000. **36**(1): p. 31-52.
20. Wu, Y. and T. Huang, *Vision-Based Gesture Recognition: A Review.* Lecture Notes in Artificial Intelligence, 1999. **1739**.
21. Hardenberg, C.v. and F. Berard. *Bare-hand human-computer interaction*. in *Proc. of Workshop on Perceptive User Interfaces, PUI'01*. 2001. Orlando, Florida.
22. Quek, F., T. Mysliwiec, and M. Zhao. *Finger mouse: A freehand pointing interface*. in *Proc. of International Workshop on Automatic Face- and Gesture-Recognition*. 1995. Zurich, Switzerland.
23. Segen, J. *GestureVR: Vision-Based 3D Hand Interface for Spatial Interaction*. in *Proc. of ACM Multimedia Conference*. 1998. Briston, England.
24. Kjeldsen, F., *Visual Recognition of Hand Gesture as a Practical Interface Modality*. 1997, Columbia University: New York, New York.
25. Raskar, R., et al. *Multi-Projector Displays Using Camera-Based Registration*. in *Proc. of IEEE Visualization'99*. 1999. San Francisco, CA.
26. Pinhanez, C., F. Nielsen, and K. Binsted. *Projecting Computer Graphics on Moving Surfaces: A Simple Calibration and Tracking Method*. in *Proc. of SIGGRAPH'99*. 1999. Los Angeles, California.
27. Raskar, R. *Oblique Projector Rendering on Planar Surfaces for a Tracked User*. in *Proc. of SIGGRAPH'99*. 1999. Los Angeles, California.
28. Faugeras, O., *Three-Dimensional Computer Vision: A Geometric Viewpoint*. 1993, Cambridge, Massachusetts: The MIT Press.
29. Sukthankar, R., R. Stockton, and M. Mullin. *Smarter Presentations: Exploiting Homography in Camera-Projector Systems*. in *Proc. of ICCV'01*. 2001. Vancouver, Canada.
30. Yang, R. and G. Welch. *Automatic and Continuous Projector Display Surface Calibration Using Every-Day Imagery*. in *Proc. of 9th International Conf. in Central Europe in Computer Graphics, Visualization, and Computer Vision*. 2001. Plzen, Czech Republic.
31. Kjeldsen, R., et al. *Interacting with Steerable Projected Displays*. in *Submitted to Face and Gesture'02*. 2002.
32. Wren, C., et al., *Pfinder: Real-Time Tracking of the Human Body.* IEEE Trans. Pattern Analysis and Machine Intelligence, 1997. **19**(7): p. 780-785.
33. Buschmann, F., et al., *A System of Patterns - Pattern Oriented Software Architecture.* 1996, New York, New York: John Wiley & Sons.
34. Crowley, J.L., J. Coutaz, and F. Berard, *Things that See.* Communications of the ACM, 2000. **43**(3): p. 54-64.
35. Pompei, F.J. *The Use of Airborne Ultrasonics for Generating Audible Sound Beams*. in *Proc. of 105th Audio Engineering Society Convention*. 1998. San Francisco, CA.

Amazon Ex. 1012.0213

Appendix B-5

Ubiquitous interactive graphics

Search within citing articles

Projector-guided painting

M Flagg, JM Rehg - Proceedings of the 19th annual ACM symposium on …, 2006 - dl.acm.org

This paper presents a novel interactive system for guiding artists to paint using traditional media and tools. The enabling technology is a multi-projector display capable of controlling …

Save Cite Cited by 102 Related articles All 11 versions

Registration techniques for using imperfect and par tially calibrated devices in planar multi-projector displays

E Bhasker, R Juang, A Majumder - IEEE Transactions on …, 2007 - ieeexplore.ieee.org

Multi-projector displays today are automatically registered, both geometrically and photometrically, using cameras. Existing registration techniques assume pre-calibrated …

Save Cite Cited by 75 Related articles All 14 versions

A personal projected display

M Ashdown, P Robinson - Proceedings of the 12th annual ACM …, 2004 - dl.acm.org

User interfaces using windows, keyboard and mouse have been in use for over 30 years, but only offer limited facilities to the user. Conventional displays are small, at least compared …

Save Cite Cited by 50 Related articles All 15 versions

To frame or not to frame: The role and design of frameless displays in ubiquitous applications

C Pinhanez, M Podlaseck - … , UbiComp 2005, Tokyo, Japan, September 11 …, 2005 - Springer

A frameless display is a display with no perceptible boundaries; it appears to be embodied in the physical world. Frameless displays are created by projecting visual elements on a …

Save Cite Cited by 50 Related articles All 15 versions

User-centric design of a vision system for interactive applications

S Borkowski, JL Crowley, J Letessier… - … on Computer Vision …, 2006 - ieeexplore.ieee.org

Despite great promise of vision-based user interfaces, commercial employment of such systems remains marginal. Most vision-based interactive systems are one-time," proof of …

Save Cite Cited by 18 Related articles All 12 versions

The dream of Peter Pan: user interaction arts using computer vision techniques

K Kim, W Kwon, E Choi, J Park, M Jeon, H Ju… - Proceedings of the …, 2007 - dl.acm.org

In this paper, we present the user interaction arts in mixed reality environment using computer vision techniques. In most of the mixed reality environment, user has to wear …

Save Cite Cited by 4 Related articles All 3 versions

Projection-based assistance for ultrasonographic diagnosis

A Prayitno - 2004 - Citeseer

In the modern age, medical examinations tend to use visual information about the internal condition of the human body for more accurate diagnosis. This information is especially …

Save Cite Cited by 2 Related articles All 4 versions

Oil painting assistance using projected light: Bridging the gap between digital and physical art

M Flagg, JM Rehg - 2005 - smartech.gatech.edu

This paper presents a novel interactive system for guiding artists to paint using traditional media and tools. The enabling technology is a multi-projector display capable of controlling …

Save Cite Cited by 2 Related articles All 3 versions

Amazon Ex. 1012.0215

유비쿼터스 환경의 디스플레이를 위한 실시간 기하보정

경동욱, 임헌규, 정기철 - 한국정보과학회 학술발표논문집, 2005 - dbpia.co.kr

최근 프로젝터를 기반으로 사용자가 원하는 위치에 영상을 제공하기 위한 많은 연구가 진행중에 있다. 프로젝터는 투사방향이 정확하지 않으면 영상왜곡이 발생되며, 영상왜곡은 ...

☆ Save 99 Cite Cited by 2 Related articles All 2 versions

실시간 보정을 지원하는 Everywhere Display

최현철, 경동욱, 한은정, 양종열, 정기철 - 한국정보과학회 학술발표 ..., 2006 - dbpia.co.kr

최근 유비쿼터스 환경에서 프로젝터를 기반으로 사용자가 원하는 위치에 영상을 제공하는 유비쿼터스 디스플레이 연구가 진행 중이다. 프로젝터는 투사방향에 따라 영상의 왜곡이 ...

☆ Save 99 Cite Related articles All 2 versions

Amazon Ex. 1012.0216

Appendix B-6

Projector-Guided Painting

Matthew Flagg and James M. Rehg
GVU Center, College of Computing,
Georgia Institute of Technology, Atlanta, GA
{mflagg,rehg}@cc.gatech.edu

ABSTRACT

This paper presents a novel interactive system for guiding artists to paint using traditional media and tools. The enabling technology is a multi-projector display capable of controlling the appearance of an artist's canvas. This display-on-canvas guides the artist to construct the painting as a series of layers. Our process model for painting is based on classical techniques and was designed to address three main issues which are challenging to novices: (1) positioning and sizing elements on the canvas, (2) executing the brushstrokes to achieve a desired texture and (3) mixing pigments to make a target color. These challenges are addressed through a set of interaction modes. Preview and color selection modes enable the artist to focus on the current target layer by highlighting the areas of the canvas to be painted. Orientation mode displays brushstroke guidelines for the creation of desired brush texture. Color mixing mode guides the artist through the color mixing process with a user interface similar to a color wheel. These interaction modes allow a novice artist to focus on a series of manageable subtasks in executing a complex painting. Our system covers the gamut of the painting process from overall composition down to detailed brushwork. We present the results from a user study which quantify the benefit that our system can provide to a novice painter.

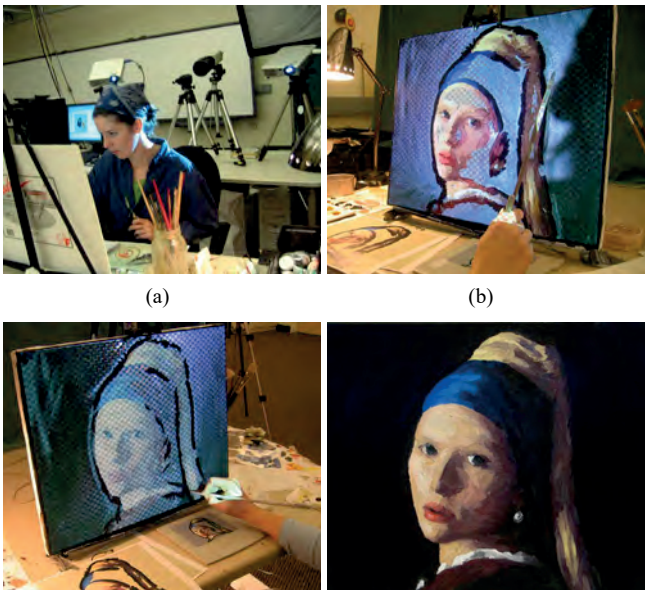
ACM Classification H5.2 [Information interfaces and presentation]: User Interfaces. - Graphical user interfaces.

General Terms Design, Human Factors, Algorithms

INTRODUCTION

In this paper, we present an interactive painting system to support novice artists in the creation of paintings using traditional media and tools. In contrast with other computer painting systems that provide tools for digital image synthesis, this work focuses on guiding users to paint according to classical techniques for creating *paintings as objects*, not just images. By employing multiple projectors to create an interactive display on the artist's canvas, our system is capable of augmenting a painting in progress with visualization features to assist with technical details such as brushstroke position, orientation and color mixing. We present a set of interaction modes designed to address common challenges for novices.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
UIST'06, October 15–18, 2006, Montreux, Switzerland..
Copyright 2006 ACM 1-59593-313-1/06/0010 ...$5.00.



(a) (b) (c) (d)

Figure 1: (a) Painting studio setup (b) Viewing a layer in *preview mode* (c) Painting the black regions of a layer in *color selection mode* (d) Final painting created using our system

Our system is part of a long history of tools for reflecting and capturing the perspective and shading of a scene as a first step in establishing the structure of a painting. Perhaps the most well-known of such tools is the *camera obscura* [22]. The modern descendant of the camera obscura is the *art projector*, a projector sold in most major art stores. The art projector is commonly used by professional painters to make accurate perspective sketches as a first step in copying originals. By replacing the art projector with video projectors and a camera, we propose using visual feedback to guide the user through the steps that follow toward creating a final desired painting.

Our system employs a multi-projector adaptive display for minimal intrusion on the traditional painter's environment in order to assist with painting technique. The display is formed on the canvas by aligning multiple overlapping front projectors. As shown in Figure 1(a), projectors are positioned on both sides of the artist to minimize the occlusion of the projected light by the artist's body. The design of our system is based on our previous work on shadow removal using redundant projection [23, 7]. Following a geometric calibration step, each pixel on the display may be illuminated by two rays of light. In addition to reducing shadows, this redundancy facilitates adaptation to the painted canvas.

Amazon Ex. 1012.0218

An issue that arises in projecting onto a painted canvas is the need to photometrically adapt the projected image to compensate for the varying reflectance properties of the paint. We use visual feedback to incrementally adjust the projected light for each canvas pixel so as to minimize the difference between the observed and desired display color. This visual feedback approach has the advantage of avoiding the need for an explicit model of paint reflectance. Our domain is a challenging one for adaptation because the canvas is continuously changing throughout the painting process and the presence of wet paint results in specular surfaces. We note that projecting onto the canvas from behind is not an option when opaque paints such as oil and acrylic are used.

This paper makes the following contributions:

- A novel interactive system for guiding artists to paint using traditional methods and tools.
- A set of interaction modes for a canvas-centric interface which supports a traditional layer-based approach to painting.
- A novel method for creating virtual paint on a previously painted canvas, which allows artists to paint each layer without being distracted by the layers underneath.

We have conducted a user study which quantifies the benefits provided by our system over a standard art projector in assisting novice painters. We show that our system provides a significantly greater improvement in the confidence of novice painters and enables them to produce higher quality artifacts as measured by a panel of experts.

Related Work

While there is a long line of work in painterly rendering for digital images [11, 15, 8, 20, 13, 12] and the simulation of traditional media and tools for art [4, 3], to our knowledge there have only been a few systems for creating art as objects using physical tools. A robotic painting system named AARON [14] has been actively developed for three decades and its paintings have been featured in several museums and art galleries. The Everywhere Display (ED) demonstration at Siggraph in 2001 [17] guided users to create "paintings" on a table out of a collection of M&Ms. In contrast, our system is designed to support a fundamentally creative process within the framework of traditional painting practice.

Other researchers have used projector-camera systems to support spatially augmented reality [18]. Shader lamps [19], for example, used multiple projectors to illuminate complex models, such as the Taj Mahal, or animate static objects, such as toy cars. As an application of real painting augmentation, [24] use projectors for restoring colors in damaged paintings. In the work of [5], a film was applied to a painting to make the surface more amenable to displaying visualization information, such as x-ray images, directly on the canvas for display in art museums. A system described in [1] used a single portable projector to create displays that spanned multiple non-coplanar surfaces. Multiple projectors were used in [6] to boost the dynamic range of projected light in a view-dependent stereoscopic display that featured photometric adaptation to the projection surface. In contrast to these applications, the goal of our system is to support a *creative process* through the construction of a dynamic display on the work of art as it is being developed. Our compensation approach is based on the method of [16], which addressed the problem of projecting onto photometrically non-uniform surfaces using feedback.

MOTIVATION AND GOALS

Existing tools that support the creation of works of art include books and manuals, demonstration videos, and art projectors. These tools have many deficiencies which limit their effectiveness. While books and manuals can demonstrate technique and illustrate the elements of painting style, it is difficult for the novice painter to bridge the gap between the printed page and the canvas. For example, painting manuals such as [2] demonstrate how to decompose a complex painting into a sequence of layers, but do not provide any direct support for the execution of each layer on the canvas. The execution of individual layers affects the overall composition in ways that are difficult for a novice painter to predict. It is also hard for novices to detect deviations between the painted layer and a reference picture.

In contrast to books and manuals, art projectors do provide direct support for the execution of the initial stages of a painting through their ability to project light directly onto the canvas. Art projectors are commonly used to trace or sketch onto the canvas the rough geometric forms that comprise a painting. Unfortunately, art projectors are unable to provide support to the artist throughout the entire painting process. They are also cumbersome to work with, as they lack the ability to calibrate and adapt automatically to the canvas and its contents.

The motivation for this work is the opportunity to leverage recent advances in projection and camera-based sensing technologies which make it possible to create a new projector-based painting aid. Our painting system has a number of novel features: it supports interaction directly on the canvas, it incorporates a model of the painting process, and it supports adaptation to an evolving painting through sensing and analysis of the painted content.

The target audience for our system is novices who would like to have an enjoyable painting experience while creating a high-quality work of art. Our system is designed to enable people who have never painted before to create a piece of artwork that they can be proud of. We accomplish this task by addressing, in the design of our system, the primary difficulties and pitfalls which novices encounter when working with oil paints. Note that we are not claiming to teach people how to paint or to provide a substitute for formal art instruction. Our system is part of a spectrum of painting aids which include traditional art projectors as well as instruction books and videos. Unlike these passive aids, however, our system is interactive and adaptive and incorporates a computational model of the painting process.

LAYER-BASED PROCESS MODEL FOR PAINTING

The versatility of oil and acrylic paint as an artistic medium has led to the development of highly refined processes and techniques for expressing a wide variety of artistic styles. These processes range from the quick freeform application of paint in impressionist styles to the painstaking 7-layer Flem-

Amazon Ex. 1012.0219

ish method, a classically realist style with drying periods of months between each layer [21]. Coupled to these wide variations in the strategy for building up a painting is an equally wide range of methods for applying paint to canvas, from precisely controlled strokes or dots to the 'drip and splash' approach that characterized Jackson Pollock's work.

The design of our system is based on the following two assumptions about the process of painting in acrylics and oils:

- We assume that paintings will be planned out in advance of execution and decomposed into a series of layers that can be applied to the canvas in linear order (back to front).
- When painting each layer, we assume that the artist will want to apply paint to the canvas in an orderly, sequential manner and will base their palette on the color scheme for the layer and the desired amount of manipulation.

Under these assumptions, a painting consists of the execution of a series of layers. Therefore, our system adapts a layer-based representation. In painting each layer we assume that the artist will mix one or more desired paint colors, select an appropriate brush or knife, and load it with paint. The artist then chooses a region of the layer corresponding to the selected color, applies the paint, and manipulates it until the desired effect has been achieved. Thus a layer is painted region-by-region until it is complete. The interaction modes described on the next page are designed to support each of these tasks. While our approach can apply equally to acrylic paints with minimal modification, for the purpose of our discussion and experiments we will confine ourselves to painting in oil.

Our model of the painting process is consistent with a number of classical and popular painting styles. Traditional layer-based oil painting methods are characterized by long drying times, which can be accelerated by thinning the paint heavily with turpentine, as in the "fat over lean" approach [10]. In these layer-based methods, wet paint is applied on top of existing layers of dry paint. An alternative is the Alla Prima method of wet-into-wet, which was popularized by Impressionist painters like Edouard Manet. In this approach paints are applied in rapid succession without intervening drying times. The goal is to achieve decisive, clean strokes which remain visible in the final painting. Although the Alla Prima method is not limited to a layer-based painting approach, it is compatible with our layer-based computational representation. Brushstrokes will often follow a progression in both size and color as early layers tend to consist of broad general strokes, followed by a progression of detail. Cézanne's advice, cited in [10], was to "start with the broom and end with the needle!". Similarly, the more opaque light paints are frequently applied after the darks. For the purpose of testing and evaluating our painting system, we have adopted the Alla Prima method, as the shorter drying times facilitate experimentation and the immediacy of the technique is appealing to beginners.

The layers which our painting model requires can come from a variety of sources. One possibility is to capture the painting process of an experienced artist. In addition, layers can be scanned in from painting manuals and registered. We can also use painterly rendering techniques to create layers automatically from photographs or 3-D models. One of the example paintings from our video submission uses the method of Hays and Essa [12] to generate layers from an input photograph, by rendering brushstrokes along contours which are generated by a multi-scale edge detector. Low frequency edges are rendered with wider brushstrokes and high frequencies are rendered with narrower ones.

CHALLENGES FACED BY NOVICE PAINTERS

Novice painters face a variety of challenges in executing a complex painting, including the difficulty of positioning and sizing the elements of the composition on the canvas, the challenge of executing the required brushstrokes, and the complexity of mixing paints to achieve the desired colors. For example, as described in [10], inexperienced painters often make the mistake of working on one small area of a painting until it is finished and then move on to the next area. This can result in a "confused and disjointed image because each area of tone and colour is separate and unrelated to its neighbors." Our adoption of a layer-based approach and the development of interaction modes which support the execution of individual layers addresses this concern.

The desired look of the painting is achieved through careful control of the brush, by regulating the pressure, movement, speed, and distance of each stroke. Impasto, for example, is a technique for producing a "fresh, vigorously textured look" [21] using thick, expressive brushstrokes. Smith [21] recognizes the difficulty of applying brushstrokes to generate the characteristic impasto texture in a uniform manner throughout a layer. This is an example of the larger class of problems novices may face at the brushstroke application level. Our system can provide guidance in the application of brushstrokes to create a desired effect.

Color mixing is one of the most difficult tasks faced by a novice painter. The proportions of paints which must be combined will vary dramatically from one shade of color to the next, and most beginning painters are unfamiliar with the subtleties of the subtractive color space that governs the paint mixing process [9]. In our observations, users can easily spend 20 minutes or longer in an attempt to reproduce a specified color for a particular painting. These long delays are a source of frustration, particularly when users must re-mix a given color after running out of it. Our system includes support for color mixing by automatically identifying the components that are required for a particular color and adaptively guiding the mixing process.

INTERACTION MODES

Our system is designed around a set of interaction modes that assist the artist in painting the ordered set of layers that comprise a finished painting. We support hands-free switching between interaction modes using a foot pedal (see the included video for example usage). Each interaction mode is realized by projecting light onto the surface of the painting itself and adapting the projection to existing paint on the canvas in creating the desired display. We describe our adaptive projected display in the following section. Our system provides five interaction modes to support the painting process:

- **Preview Mode** to display layers

Amazon Ex. 1012.0220



(a) (b) (c)

Figure 2: (a) Layer 4 out of 4 in preview mode, with highlighted region around lips. (b) Closeup of lip region. (c) Lip region following color selection. Note that all images are photographs of the actual canvas under compensation.

- **Color Selection Mode** to paint with specific colors
- **Blank Mode** to assess the state of the painting
- **Orientation Mode** to display brushstroke guidelines
- **Color Mixing Mode** to aid in creating the desired colors

Preview Mode: A visual guide for painting a given layer is provided by preview mode, which creates an overlay on the canvas that displays the current layer in progress while compensating for the underlying paint. The use of this mode is depicted in Figs. 1(b) and 2(a). In preview mode, areas of the canvas that are not to be painted in the current layer are hidden from the artist using a checkered pattern. This pattern is easily distinguished from the foreground elements of the layer. This reduces the chance that regions outside the current layer will distract the artist from a particular execution step.

Furthermore, the clear visibility of all layer regions and "invisibility" of off-layer regions provided in preview mode helps the artist devise a strategy for selecting a color palette and set of brushes that are used to paint the layer. For example, it may be the case that each layer corresponds naturally to one specific brush size. It is easy to select a brush in preview mode by physically comparing brushes to their previewed strokes on the canvas.

Color selection mode: Another challenge in executing a particular layer is the efficient use of paint. An artist with a loaded brush would like to quickly identify regions of the canvas that need that particular color. To address this need, we designed a color selection mode that highlights all strokes of a chosen color in the current layer. The artist uses a Wacom tablet with a pen attached to the opposite end of the paintbrush to select a specific color in the layer. Figs. 1(c) and 2 illustrate the use of this mode. The white cross in Fig. 2(b) indicates the selection of a particular shade of red in the layer containing the girl's lips. We employ a user-specified tolerance parameter to determine which parts of the layer match the selected color. The tolerance is adjusted interactively via the foot pedal, to expand or contract the color region of interest.

Following color selection, all other colors are hidden from view so that the artist can apply the color where it is needed. A comparison between Fig. 2(b) and 2(c) demonstrates that color selection mode effectively masks out regions in the layer which do not correspond to the selected color. For areas like these where it is difficult to distinguish between various shades of red, the color selection tool is useful for clearly marking where paint should be added.

Blank mode: The artist can systematically complete each layer through the use of the preview and color selection modes. In order to assess the progress made in painting a layer, along with the current state of the painting, the artist may switch the projectors off at any time using blank mode. The included video shows the utilization of preview, color selection, and blank mode in completing a layer.

Orientation mode: While preview and color selection modes show the artist *where* to apply paint, orientation mode shows the artist *how* to apply it. A set of guidelines overlaid on the layer indicate the orientation, width, and extent of brushstrokes that should be applied to achieve a desired brushstroke texture. For example, long, linear guidelines would indicate the need for bold, expressive strokes, while a short hatching pattern would indicate the use of a dabbing technique. Guidelines can also be used to indicate directions for blending paints, as depicted in Figures 4(b) and 4(c). Orientation guidelines must be generated along with their associated layer images. We note that one benefit of using the painterly rendering method of Hays and Essa [12] to produce layers is that it automatically generates contours that can be used as orientation guidelines.

Color mixing mode: Support for the task of combining paints to achieve a desired color is provided by color mixing mode. Fig. 3 shows the interface, which depicts a standard color wheel along with a camera view of the artist's palette. Using the Wacom tablet, the user identifies an area on the palette where paint is to be mixed. A desired color from the current layer can be specified via color selection mode. A precalibrated table maps the target color to a specific location on the color wheel, known as the target point. Labeled source points on the color wheel correspond to fully-saturated source pigments (labeled "18" and "24" in Figure 3).

The first step in mixing the target color is to identify the fully saturated colors that may be combined to make the target color. For any point on the color wheel, the system selects the two colors that lie on both the outer boundary of the color wheel and on a line connecting the center of the wheel to the

Amazon Ex. 1012.0221



Figure 3: Interface for *color mixing mode*, which employs automated capture of and localization of palette colors on a color wheel to guide the paint mixing process. The blue arrow illustrates the movement of the current mixture color, which is the selected region in the camera image, to the desired target.

target color's location on the wheel. These two fully saturated components of the target color are complementary colors. Following identification of these two ingredient colors, the user may put a sample of the color on the palette directly from the tube or by means of another mixture process. The interface may be used to help mix these source ingredients if needed. After both ingredient colors have been placed on the palette, one color is gradually mixed with the other. Following each step of paint addition, the camera image is analyzed to identify the point on the model color wheel which corresponds to the mixture in progress. Estimation of the mixture point is accomplished using a pre-calibrated set of histograms obtained for a representative set of pigments. These points chart the geometric progress of moving the color along a gradient on the color wheel. The system guides the user to slowly add one ingredient color to the other to achieve the desired target point.

CREATING A DISPLAY ON A PAINTED CANVAS

We employ a system of two projectors and one camera to create an adaptive display directly on the canvas. The projector outputs and camera input can be automatically warped and aligned with the surface of the canvas using standard calibration methods [7].

Multi-Projector Adaptive Display: By placing the two projectors at an oblique angle to the canvas and overlapping and aligning their outputs, we gain the ability to illuminate each pixel on the canvas with two distinct sources of light. This has two benefits. First, by projecting simultaneously from both projectors we gain robustness to shadows and occlusions, as described in [7]. Second, by adapting with both projectors during compensation, we can increase the dynamic range of the display on the canvas. This is particularly useful in trying to adapt dark colors to make them lighter.

Photometric Adaptation: In order to control the appearance of the canvas, our system must account for the spatially varying reflectance of a painted canvas. Any image projected onto the canvas will be modulated by the reflectance of the paint and therefore create an undesired appearance as shown in Fig. 5(d). To achieve a desired appearance, we adapt a

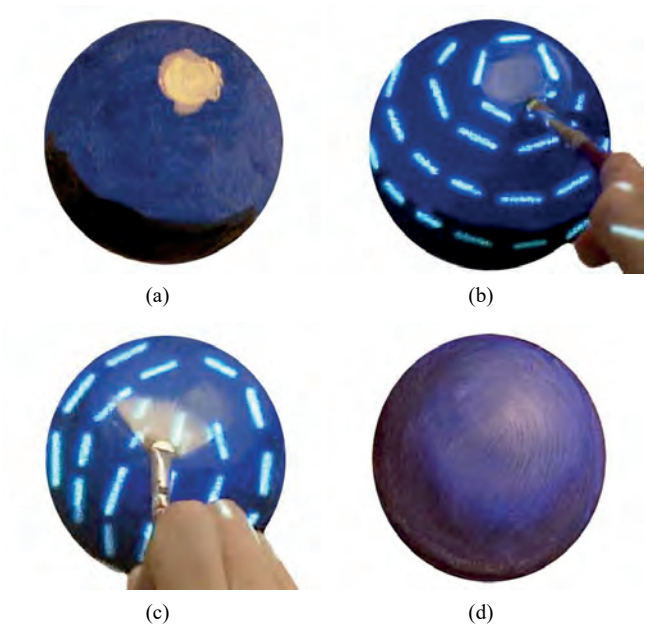


Figure 4: Painting a sphere using orientation mode. (a) Sphere with highlight and shadow; (b) Coarse blending with bristle brush in orientation mode; (c) Fine blending with fan brush in orientation mode; (d) Finished result.

feedback approach in which a captured camera image of the canvas is analyzed to create a compensation image for each projector to output.

Our compensation approach is based on the method of Nayar et. al. [16] for adapting the output of a single projector to a surface with varying reflectance. Our system employs multiple projectors to increase the dynamic range of the compensation method and supports the positioning of the projectors beside the artist instead of directly in front of the canvas. The algorithm works by projecting an initial input image and iteratively adjusting image intensities using visual feedback. Starting with all projectors projecting black, each projector iteratively reduces compensation error, the difference between desired and measured intensities for each pixel and channel independently, according to the proportional feedback law $P_{i+1} = k(I - C_i) + P_i$, where P_i is the compensation image for iteration i, I is the desired appearance image, C_i is the surface-aligned camera image for iteration i, and k is the gain. Once the error has dropped below a threshold tolerance, a completion mask is toggled for the finished pixel to avoid further adjustment of its compensation intensity. Adaptation halts after the change in error for all pixels has stabilized or a maximum number of iterations have occurred. Following adaptation, the compensation image for the current projector is fixed and adaptation begins for the next projector. Each subsequent projector reduces error for the remaining pixels that have not met the target intensities for each channel. Figure 5 illustrates the compensation process for 2 projectors positioned as in Figure 1(a). Clearly, the appearance of the canvas is closer to the desired image with two projectors than with one. These results establish the effectiveness of compensation in the case of oil paints.

Amazon Ex. 1012.0222



Figure 6: 3 paintings made using our system by 3 painters who had never painted before.

RESULTS

We present the results of a formal user study of novice painters along with informal observations on the usage of our system by several subjects.

User Study

We conducted a formal user study to quantify the benefit that our system can provide to a novice painter, relative to a standard art projector. The study had two components. In the first component, a population of twenty subjects was asked to paint copies of the model painting illustrated in Fig. 8(b).[1] Each of the subjects was given a canvas containing an initial sketch of the desired shape (as produced by a standard art projector) and printouts of the model painting. Ten of the subjects served as a control group and produced their paintings without the use of our system. The other ten subjects had full access to our system.

Prior to their painting experience, each subject filled out a questionnaire in which they were asked to rate their confidence in their ability, on a scale of 1 to 5 where 1 means "not confident at all" and 5 means "very confident", to paint a set of three paintings which were identified by an art expert to be of low, medium, and high difficulty.[2] The questionnaire also contained questions to assess whether the subject had any previous painting experience.[3] (Subjects with prior experience were allowed to complete their paintings but one was dropped from the analysis due to experience.) Each subject was then given a brief orientation session and then asked to paint a short warm-up painting which took approximately 15 minutes.[4] Each subject then completed the model painting, which took three hours on average.

Following their painting experience, the subjects filled out a second survey which again asked them to rate their ability to execute a different set of three paintings of equivalent levels of difficulty. In addition, subjects were asked to rate their level of frustration, level of boredom, likelihood to paint again, assessment of the quality of their painting, and overall satisfaction with the experience, on a scale of 1 to 5.

We conducted an analysis of the self-assessed ability ratings for each subject population before and after their painting experience. The average prior confidence rating for the painting of Medium difficulty was 2.2 for the control group and 2.3 for the group that used our system. Following their painting experience, the average rating for the Medium difficulty painting dropped to 1.6 for the control group, and rose slightly to 2.5 for the group that used our system. We ran a t-test and Mann-Whitney-U nonparametric test for the null hypothesis that there was no significant difference in the dependent variables between the two groups. The difference in assessed ratings was found to be statistically significant (at level $p < 0.05$) using both tests. We conclude that our system has a quantifiable, positive benefit over what an art projector alone can provide in improving the confidence of novice users.[5]

In the second component of the study, a group of three experienced painting professors filled out two questionnaires to evaluate each of the twenty paintings completed by the novice users (see Fig. 8). The first questionnaire asked respondents to rate 6 attributes of painting quality on a scale of 1 to 7 and the second questionnaire asked the professors to rank their top ten paintings in terms of overall quality and overall similarity to the model painting.[6] Fig. 7 reports the sum of the ranks of expert ratings for six survey attributes over the two populations of paintings, where 1 = "poor quality", 4 = "medium quality" and 7 = "high quality." The p-values for the Mann-Whitney-U test are shown for each attribute. A statistically significant difference (at $p < 0.05$) was observed for every attribute except "quality of color matching to original painting." Furthermore, a majority of the top 10 paintings ranked by each professor were painted with the use of our system. These results serve to further validate our

[1] An experienced artist created the barn image in two layers with associated orientation mode guidelines. The captured layers and guidelines were used by our system.

[2] The paintings used in our survey were as follows: Before painting, difficulty *Low* "Composition with Red, Blue, and Yellow" (Mondrian), *Medium* "The Lemon" (Manet), and *High* "Floor Scrapers" (Caillebotte). After painting, difficulty *Low* "Number 10" (Rothko), *Medium* "Dusk" (Monet), and *High* "The Music Lesson" (Vermeer).

[3] All questionnaires and raw data are available at http://www.cc.gatech.edu/cpl/projects/painting

[4] The warm-up session was conducted to help eliminate any negative effects of using basic painting tools for the first time.

[5] The average confidence ratings for the Low and High difficulty paintings did not show a statistically significant variation between the two groups. However the mean ratings for both attributes were higher for the group that used our system. Note that the model painting was designed to be of medium difficulty.

[6] In addition to ranking the paintings, the second questionnaire asked for open comments on standard attributes they have used in the classroom for comparing paintings and assessing their relative quality. These comments were found to validate our choice of survey attributes for the quality of the painting and its similarity to the model.

Amazon Ex. 1012.0223

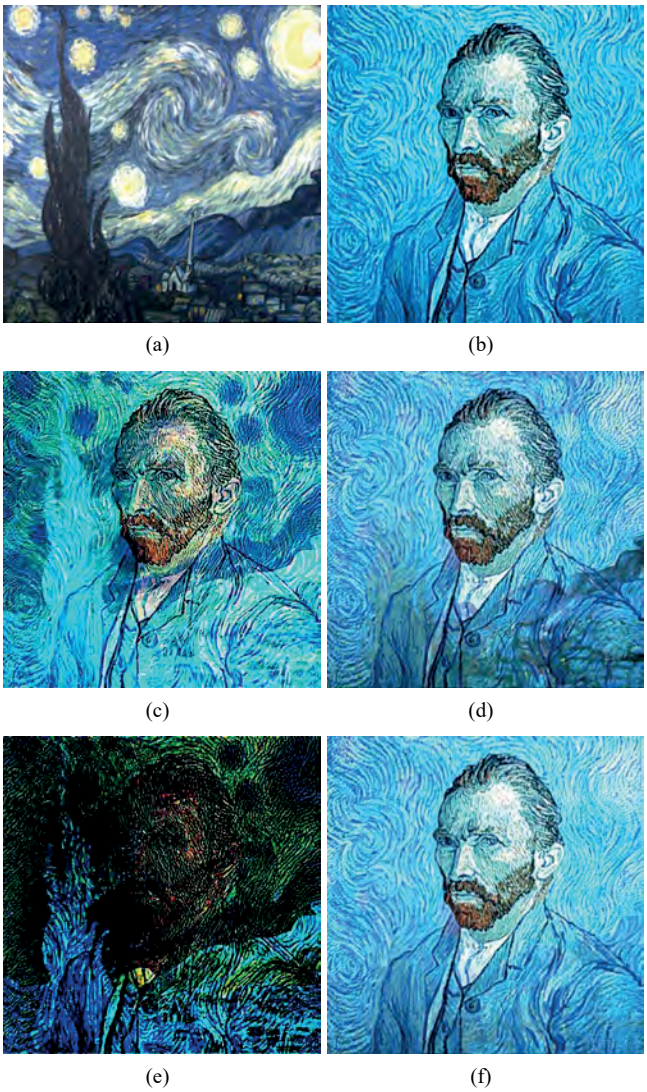


Figure 5: (a) Canvas before compensation; (b) Desired appearance; (c) Projector 1 compensation image; (d) Compensation result for projector 1 alone; (e) Projector 2 compensation image; (f) Final canvas appearance under the combination of projectors 1 and 2.

claim that our system can help novice painters to produce higher-quality artifacts than an art projector alone.

Usage Trends

Figure 6 shows three paintings created using our system by three novice artists who had never painted in oils before. Each painting took an average of 5 hours to create using our system. We observed the painting process for these artists and identified the following trends:

First, two of the subjects exhibited initial hesitation in applying the brushstrokes directly on top of the displayed virtual brushstrokes. For the two impressionist style paintings in Figure 6, for example, both artists applied the paint in very small and slow brushstrokes despite the implicit instruction to paint with quicker and more sweeping brushstrokes as visually suggested in preview, color selection and orientation modes.

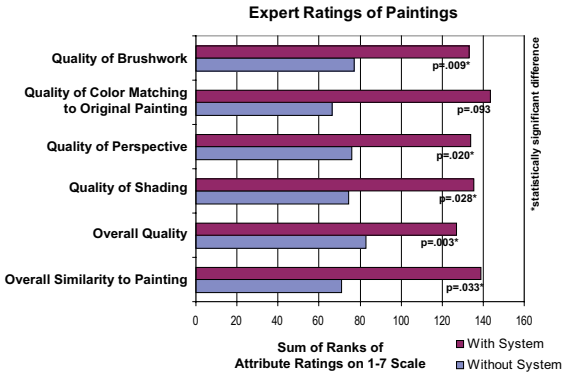


Figure 7: Aggregate quality ratings reported by three painting professors, for the two sets of paintings shown in Fig. 8.

Second, as the novices' hesitations faded and they switched between interaction modes more frequently, task execution speed increased. This indicates that they learned how to apply regions of color in a layer in a more effective manner and, as a result, became more confident in their ability to accomplish painting goals.

Third, as the paintings progressed, all three users switched to color selection mode more frequently. In one case, the artist repeatedly re-adjusted the color similarity tolerance for each color selection before applying a new color to the canvas. This enabled him to grow or shrink the layer segmentation until it appeared to conform to contiguous brushstroke areas as opposed to pieces of brushstrokes resulting from color variation across the brushstroke. This behavior suggests the need for a more sophisticated means of detecting and segmenting brushstrokes for color selection mode.

We were initially concerned about the potential negative effects of glare and shadows on the subjects in our experiment. Glare caused by projecting light onto wet oil paint could both hinder the ability to control the appearance of the canvas using photometric compensation and annoy users of our system. Shadowing on the display due to projector occlusions by the artist could prevent the effective use of interaction modes for painting guidance at the color application and manipulation level. We found that by positioning our projectors at oblique angles, we were able to reduce both of these effects in practice. None of the artists reported any discomfort caused by glare or shadows.

CONCLUSION AND FUTURE WORK

We have described a novel interactive system for guiding novice artists to paint using traditional methods and tools. We have demonstrated how a multi-projector adaptive display can support artistic painting by guiding users according to traditional painting techniques taught in artist's manuals. Furthermore, we presented a set of five interaction modes for a canvas-centric interface which supports a traditional layer-based approach to painting. We evaluated the effectiveness of our system in meeting the goals of improved quality of experience and high quality of output. The results of our user study indicate that these goals were met.

Amazon Ex. 1012.0224



(a)





(b)



(c)

Figure 8: **(a)**: *Top two rows:* Paintings made with the system. *Lower two rows:* Corresponding zoomed-in regions for the paintings in the top rows. **(b)**: Model painting all subjects in the user study were told to copy **(c)**: *Top two rows:* Paintings made without the system. *Lower two rows:* Corresponding zoomed-in regions for the paintings in the top rows. Notice the larger deviation in shading from the model for the zoomed in portion between the two groups.

Amazon Ex. 1012.0225

Our system has the flexibility to handle a wide range of painting styles and approaches. Portraiture, still life, and landscape paintings are examples of domains that are supported by our process model. However, Jackson Pollock's paintings provide an example of a less traditional means of applying paint that does not fall within the range of techniques that we currently support.

In future work, we are considering extending the painting interaction techniques presented in this paper to other areas of the fine arts. Multiple projectors could guide a sculptor in subtracting material from a block of marble, for example. Just as layers were used to build up the composition of a painting in the current system, a sculpting system could guide an artist in a hierarchical broad-to-fine chiseling approach.

ACKNOWLEDGEMENTS

We would like to thank Roxanne Flagg for providing expert painting demonstrations and making helpful comments that were instrumental to improving the painting interface. Special thanks to Lan Guo for her assistance in conducting the user study and thanks to our experts Philip Carpenter, Daryl White, and Tina Simonton. We also thank professors Jessica Hodgins, Irfan Essa and Michael Cohen for their helpful comments on the paper. This material is based on work supported in part by the National Science Foundation under NSF Grant IIS-0205507.

REFERENCES

1. M. Ashdown, M. Flagg, R. Sukthankar, and J. Rehg. A flexible projector-camera system for multi-planar displays. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, pages 165–172, 2004.
2. Barron. *All About Techniques in Oil*. Barron's Educational Series, Inc., 1997.
3. W. Baxter, V. Scheib, M.C. Lin, and D. Manocha. dab: Interactive haptic painting with 3d virtual brushes. In *Proceedings of Siggraph*, pages 461–468, New York, 2001.
4. W. Baxter, J. Wendt, and M. Lin. Impasto: A realistic, interactive model for paint. In *Proceedings of the 3rd International Symposium on Non-Photorealistc Animation and Rendering*, pages 45–56, New York, June 2004.
5. O. Bimber, F. Coriand, A. Kleppe, E. Bruns, S. Zollmann, and T. Langlotz. Superimposing pictorial artwork with projected imagery. *IEEE MultiMedia*, pages 16–26, January-March 2005.
6. O. Bimber, G. Wetzstein, A. Emmerling, and C. Nitschke. Enabling view-dependent stereoscopic projection in real environments. In *In proceedings of International Symposium on Mixed and Augmented Reality (ISMAR'05)*, pages 14–23, 2005.
7. T.-J. Cham, J.M. Rehg, R. Sukthankar, and G. Sukthankar. Shadow elimination and occluder light suppression for multi-projector displays. In *Proceedings of Computer Vision and Pattern Recognition*, pages 513–520, 2003.
8. C.J. Curtis, S.E. Anderson, J.E. Seims, K.W. Fleischer, and D.H. Salesin. Computer-generated watercolor. In *Proceedings of Siggraph*, pages 421–430, 1997.
9. B. Edwards. *Color: A Course in Mastering the Art of Mixing Colors*. Penguin Group (USA) Inc., 2004.
10. A. Gair. *The Artist's Handbook: A Step-by-Step Guide to Drawing, Watercolor and Oil Painting*. MetroBook, 1998.
11. P.E. Haeberli. Paint by numbers: Abstract image representations. In *Proceedings of Siggraph*, pages 207–214, 1990.
12. J. Hays and I. Essa. Image and video based painterly animation. In *Proceedings of the 3rd Symposium on Non-Photorealistc Animation and Rendering*, pages 113–120, New York, 2004.
13. A. Hertzmann. Painterly rendering with curved brush strokes of multiple sizes. In *Proceedings of Siggraph*, pages 453–460, 1998.
14. P. McCorduck. *Aaron's Code: Meta-Art, Artificial Intelligence and the Work of Harold Cohen*. W H Freeman and Co., 1990.
15. B.J. Meier. Painterly rendering for animation. In *Proceedings of Siggraph*, pages 477–484, 1996.
16. S. Nayar, H. Peri, M. Grossberg, and P. Belhumeur. A projection system with radiometric compensation for screen imperfections. In *First IEEE International Workshop on Projector-Camera Systems (PROCAMS-2003)*, 2003.
17. C. Pinhanez, R. Kjeldsen, T. Levas, G. Pingali, M. Podlaseck, and P. Chou. Ubiquitous Interactive Graphics. Technical Report RC22495 (W0205-143), IBM Research Report, May 2002.
18. R. Raskar, G. Welch, and H. Fuchs. Spatially augmented reality. In *First IEEE Workshop on Augmented Reality (IWAR'98)*, 1998.
19. R. Raskar, G. Welch, and K.-L. Low. Shader Lamps: Animating real objects with image-based illumination. In *Proceedings of Eurographics Workshop on Rendering*, 2001.
20. M.P. Salisbury, M.T. Wong, J.F. Hughes, and D.H. Salesin. Orientable textures for image-based pen-and-ink illustration. In *Proceedings of Siggraph*, 1997.
21. R. Smith, M. Wright, and J. Horton. *An Introduction to Art Techniques*. DK Publishing, Inc., 1995.
22. Philip Steadman. *Vermeer's Camera: Uncovering the Truth behind the Masterpieces*. Oxford University Press, 2002.
23. J. Summet, M. Flagg, T.-J. Cham, J.M. Rehg, and R. Sukthankar. Shadow elimination and blinding light suppression for interactive projected displays. *IEEE Transactions on Visualization and Computer Graphics*, Accepted for publication in 2006.
24. T. Yoshida, C. Horii, and K. Sato. A virtual color reconstruction system for real heritage with light projection. In *Proceedings of International Conference on Virtual Systems and MultiMedia*, 2003.

Amazon Ex. 1012.0226

Appendix B-7

Registration Techniques for Using Imperfect and Partially Calibrated Devices in Planar Multi-Projector Displays

Ezekiel Bhasker, Ray Juang, *Student, IEEE*, and Aditi Majumder, *Member, IEEE*

Abstract— Multi-projector displays today are automatically registered, both geometrically and photometrically, using cameras. Existing registration techniques assume pre-calibrated projectors and cameras that are devoid of imperfections such as lens distortion. In practice, however, these devices are usually imperfect and uncalibrated. Registration of each of these devices is often more challenging than the multi-projector display registration itself. To make tiled projection-based displays accessible to a layman user we should allow the use of uncalibrated inexpensive devices that are prone to imperfections. In this paper, we make two important advances in this direction. First, we present a new geometric registration technique that can achieve geometric alignment *in the presence of severe projector lens distortion* using a relatively inexpensive low-resolution camera. This is achieved via a closed-form model that relates the projectors to cameras, in planar multi-projector displays, using rational Bezier patches. This enables us to geometrically calibrate a 3000×2500 resolution planar multi-projector display made of 3×3 array of nine severely distorted projectors using a low resolution (640×480) VGA camera. Second, we present a photometric self-calibration technique for a projector-camera pair. This allows us to photometrically calibrate the same display made of nine projectors using a *photometrically uncalibrated* camera. To the best of our knowledge, this is the first work that allows geometrically imperfect projectors and photometrically uncalibrated cameras in calibrating multi-projector displays.

Index Terms—Geometric calibration, photometric calibration, tiled displays.

1 Introduction

Multi-projector displays are popular today for many applications such as visualization, entertainment, training, and simulation. Contemporary researchers envision future workspaces with *ubiquitous pixels* rendered by multi-projector displays of various scales and forms aiding users in collaboration, interface, and visualization [9, 12, 16, 17, 24, 25, 30, 33, 36].

To enable such readily accessible multi-projector displays, automated camera-based registration techniques that register the imagery coming from multiple projectors, both geometrically and photometrically, have been proposed [5, 6, 13, 19, 20, 21, 23, 26, 27, 28, 29, 34, 37, 38]. More recently, Bhasker, et al. [2], proposed a distributed network of compact projector-camera units that can self-register without any user input, relieving the user of installation and maintenance responsibilities. However, all these methods do not handle imperfections, such as lens distortions, that are common in inexpensive commodity devices. They assume pre-calibrated devices (*i.e.,* devices whose geometric and photometric parameters are known) leaving the responsibility of device calibration to the user. Device calibration is often more challenging than multi-projector display registration and demands a very educated user. This difficulty in deployment has quarantined tiled projection-based displays to national laboratories and universities despite huge advances in automated camera-based registration techniques.

The work in this paper makes initial advances by developing geometric and photometric registration techniques that allow imperfect uncalibrated devices to be used for the most common cases of planar displays. The ultimate goal would be to allow the use of projectors with *severe* geometric and photometric distortion and still achieve acceptable geometric and photometric registration. We achieve this goal using cameras that are neither geometrically nor photometrically registered. The paper by Majumder and Stevens [21] is the only work to date in this direction and presents initial methods to handle large spatial variation in intensity within and across projectors. In this paper, we make further advances to these unexplored domains. These are: (a) achieving geometric registration while using projectors that have considerable geometric imperfections (*e.g.,* lens distortions); (b) achieving photometric registration with a photometrically uncalibrated camera.

- *Ezekiel Bhasker is with University of California, Irvine, E-mail: ebhasker@ics.uci.edu.*
- *Ray Juang is with University of California, Irvine, E-mail: rjuang@ics.uci.edu.*
- *Aditi Majumder is with University of California, Irvine, E-mail: majumder@ics.uci.edu.*

Manuscript received 31 March 2007; accepted 1 August 2007; posted online 27 October 2007. Published 14 September 2007.
For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.

1.1 Using Geometrically Imperfect Projectors

We present a new geometric registration technique that can correct for severe lens distortion using a relatively low-resolution camera for planar multi-projector displays. This method relies on a new parametric geometric model relating the projector to the camera used for registering the display. The relationship is modeled in a compact and efficient manner using a rational Bezier patch.

Projection technology is advancing at a tremendous pace. Pocket projectors that can fit on one's palm are no longer a myth but a reality [1]. Industry initiatives are finding ways to build short-throw, unbreakable projectors that use LEDs as light sources so that they can be embedded in cell phones. It is evident that lens distortion will be significant in such low-cost projectors. Our method will enable the use of such low-cost projectors to build commodity tiled displays.

Further, in designing a compact multi-projector display or changing an existing setup to increase space utility, a short-throw lens is required to decrease the throw distance of the projectors. These lenses currently require several optical elements to reverse the high lens distortion created by short focal lengths. Hence, they are cost-prohibitive (around \$2000 - \$6000) and at least 5-6 times the cost of regular projectors (around \$800 - \$1500). Our work handles severe lens distortions, thus paving the way for using inexpensive lenses on projectors, just as they are used on cameras today. This allows for even more compact multi-projector display designs.

1.2 Using Photometrically Uncalibrated Cameras

Existing methods for photometric calibration of a multi-projector display often use a camera that is first calibrated using high dynamic range images [8, 26]. This involves taking images in an outdoor environment where one does not have control of motion, illumination, etc. This calibrated camera is then used to recover the photometric proper-

1077-2626/07/$25.00 © 2007 IEEE Published by the IEEE Computer Society

Amazon Ex. 1012.0228
Authorized licensed use limited to: Purdue University. Downloaded on April 25,2023 at 14:08:57 UTC from IEEE Xplore. Restrictions apply.

ties of the projectors which are then modified appropriately to achieve a photometrically seamless multi-projector display [21, 26].

In this paper, we present a photometric self-calibration technique for a projector-camera pair that allows us to register a multi-projector display photometrically using an uncalibrated camera. To the best of our knowledge, this is the first work that estimates both the intensity transfer function and the spatial variation in intensity of projectors simultaneously using a photometrically uncalibrated camera.

2 Related Work

In this section, we review existing geometric and photometric registration techniques and relate them to the techniques we present in this paper.

2.1 Geometric Registration

Geometric registration of planar multi-projector displays entails reconstructing two functions: (a) the function that relates the individual projector coordinates to the camera coordinates; and (b) the function that relates the camera coordinates to the global screen coordinates. Most geometric registration techniques devised for planar displays so far fall into the following categories.

Linear methods assume linear models for cameras and projectors. They relate the projector, camera, and the screen coordinates by linear matrices called homographies [5, 34]. However, linear models do not account for geometric non-linearities, such as lens distortion, and cannot achieve accurate alignment in their presence. To achieve acceptable alignment in such cases, projectors are manually adjusted to the 'sweet-spot' in zoom level where the non-linear distortions become negligible.

Piecewise linear methods address these geometric non-linearities by using a piecewise linear function, essentially a triangulation, to relate the projector, camera and screen coordinates to each other. Though this achieves reasonable geometric accuracy, a dense triangulation is required to sample the parameter space adequately [37]. Thus, a high-resolution (5-6 Megapixel) camera is required to register even relatively small displays made of 4-6 projectors.

Hereld, et. al. [13], use a *non-linear model* where a closed-form cubic function handles some of the geometric non-linearities. This method has two limitations. First, a simple cubic polynomial cannot capture the perspective projection between the projector and the camera adequately. As a result, this method assumes a close-to-rectangular array of projectors resulting from on-axis projection. Such an arrangement is relatively easy to achieve manually in a rear-projection system, but not in front-projection system where the projector is on the ceiling projecting on a screen placed on the ground. Second, this work is not motivated by accurate analysis of the kind of non-linearities seen in projectors and hence cannot handle higher order non-linearities.

Our method incorporates the non-linear distortions of the projector into the function that relates the projector to other entities, such as the camera and screen, using a single accurate closed-form function – the rational Bezier patch. This allows combination of distortions arising from completely different reasons (*e.g.,* lens distortion and keystoning) into a single unified representation. Further, it results in an efficient camera-based geometric registration for planar, multi-projector displays, that can handle significant lens distortion.

2.2 Photometric Registration

Photometric properties of a camera/projector involve its intensity transfer function, or ITF (more commonly called the gamma function), and the spatial intensity fall-off in its field-of-view (more commonly called the vignetting effect). *Photometric registration* of projection-based displays requires estimation of the photometric properties of the projectors using a camera. To do this, the camera is first photometrically registered. The high dynamic range imaging technique [8, 22] can be used to estimate the camera's ITF by taking images outdoors. Currently, there is no simple automated way of estimating the camera's vignetting effect. The registered camera is usually set to a narrow aperture when using it to estimate the photometric properties of the projectors in a tiled display. The techniques in the work by Raij, et al. [26], are then applied to estimate the projectors' ITF using the registered camera. Using the projectors with known ITF, the techniques described in the papers by Majumder and Stevens [20, 21] are then used to estimate the spatial intensity fall-off of the projector, and to register the multiple projectors photometrically using the estimated photometric properties. No method currently exists that can use a photometrically uncalibrated camera and recover both the ITF and the spatial intensity fall-off of the projectors simultaneously. Our method in this paper estimates all the different photometric parameters of both devices in a photometrically uncalibrated projector-camera pair. It is accomplished in a controlled indoor setting without using any other physical props. Photometric registration techniques that use these parameters [20, 21, 38] can then be utilized to achieve photometric seamlessness in the multi-projector display.

3 Using Geometrically Imperfect Projectors

In this section, we present our geometric registration technique that corrects for projector geometric imperfections, such as lens distortion. We use a camera that is corrected for lens distortion [3]. First, we develop a new model that relates a single projector to the camera in a planar multi-projector display using a rational Bezier patch. The projector lens distortion is modeled by an affine-invariant *non-rational* Bezier patch. For a projector-camera setup with a planar display, this function is followed by a homography that relates the projector to the camera and to the screen linearly. We use a perspective-invariant *rational* Bezier patch to model the combination of these non-linear and linear functions. Thus, all the different distortions in a planar projection-based display (*e.g.,* key-stoning, radial, and tangential distortions) are combined into a single closed-form function. To the best of our knowledge, this is the first work that models both linear and non-linear distortions of a projector-camera system in a unified manner.

To geometrically register multiple distorted projectors, we sparsely sample the rational Bezier function relating the projector coordinates with the camera and screen coordinates, and estimate the function parameters. These parameters are then used to warp the input image to the projector appropriately resulting in a geometrically aligned image. The sparse sampling proves adequate for acceptable geometric registration since the rational Bezier offers a piecewise curve representation (which is more suited for lens distortions that transform lines to curves), rather than a piecewise linear representation. This sparse sampling allows us to use a low-resolution camera to achieve geometric registration for even medium sized displays made of 6-16 projectors.

3.1 Model for Single Projector

We first consider the simple case of a single projector projecting onto a planar screen and being observed by a single camera corrected for lens distortion. Our goal is to design an accurate parametric function $\mathscr{F}$ that defines the relationship between the projector pixels (x,y) and the camera pixels (u,v), *i.e.,* $(u,v) = \mathscr{F}(x,y)$. Ideally, the function $\mathscr{F}$ has two components. The first is the *lens distortion function* $\mathscr{R}$ that relates the undistorted projector coordinates (x,y) to the distorted coordinates on the display screen (x_d,y_d). The second component is the well-known *homography* $\mathscr{H}$ that relates the distorted coordinates to camera coordinates (u,v).

$$(u,v) = \mathscr{H}(x_d,y_d) = \mathscr{H}(\mathscr{R}(x,y)) = \mathscr{F}(x,y) \tag{1}$$

We would like to model $\mathscr{F}$ by a single compact parametric function that accurately captures the effect of the concatenation of these two functions, $\mathscr{H}$ and $\mathscr{R}$, and is amenable to efficient data fitting computation.

Modeling $\mathscr{R}$ with Non-rational Bezier Patches: The classic lens distortion model [4, 10] consists of two independent distortions: *radial distortion* (*e.g.,* barrel or pin-cushion) and *tangential distortion*.

Radial distortion is usually modeled by

$$x_d = x + (x - x_p)(k_1 r^2 + k_2 r^4 + k_3 r^6) = x + \rho_x, \tag{2}$$

$$y_d = y + (y - y_p)(k_1 r^2 + k_2 r^4 + k_3 r^6) = y + \rho_y, \tag{3}$$

Amazon Ex. 1012 0229
Authorized licensed use limited to: Purdue University. Downloaded on April 25,2023 at 14:08:57 UTC from IEEE Xplore. Restrictions apply.

Case No.	Lens Distortion Parameters			% Error For Bezier of degree					
	Center (x_c, y_c)	Radial (k_1, k_2, k_3)	Tangential (p_1, p_2)	2	3	4	5	6	7
(a)	$(0.5, 0.5)$	$(-0.35, 0, 0)$	$(0, 0)$	5.78	ε				
(b)	$(0.5, 0.5)$	$(0, 0, 0)$	$(0.1, 0.1)$	ε					
(c)	$(0.5, 0.5)$	$(-0.35, -0.35, 0)$	$(0, 0)$	9.00	0.91	0.56	ε		
(d)	$(0.6, 0.55)$	$(-0.35, -0.13, 0)$	$(0.05, 0.05)$	1.9	0.14	0.03	ε		
(e)	$(0.6, 0.55)$	$(-0.35, -0.13, -0.016)$	$(0.05, 0.05)$	2.0	0.17	0.038	0.0012	0.00022	ε

Table 1. This table shows the accuracy of our fit for lens distortion functions using non-rational Bezier patches of different degree. Due to the large number of lower degree terms, note that in case (e) distortion of higher degrees (*e.g.,* degree 7) can be fitted to reasonable accuracy by a lower degree Bezier (*e.g.,* degree 4). ε represents the convergence threshold used which is strictly less than 10^{-11}.

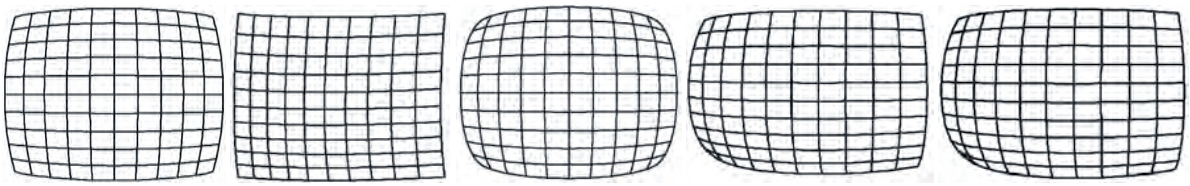


Fig. 1. Images showing the fives cases of distortion mentioned in Table 1.

where (x_d, y_d) and (x, y) are the distorted and undistorted image coordinates respectively, $r = \sqrt{(x-x_p)^2 + (y-y_p)^2}$ is the radial distance from the principal center (x_p, y_p) in the undistorted image, and k_i, $1 \leq i \leq 3$ are the radial distortion coefficients. Negative values of k_i result in barrel distortion while positive values result in pin-cushioning. The principal center is a point in the image that is unchanged by radial distortion. In general, the principal center (x_p, y_p) need not be at the center of the image but is usually close to the center.

The tangential distortion is modeled by

$$x_d = x + 2p_1xy + p_2r^2 + 2p_2x^2 = x + \tau_x, \quad (4)$$

$$y_d = y + 2p_2xy + p_1r^2 + 2p_1y^2 = y + \tau_y \quad (5)$$

where $r = \sqrt{(x-x_p)^2 + (y-y_p)^2}$ and p_1 and p_2 are the tangential distortion parameters.

Radial and tangential distortion are independent of each other and hence their effects can be combined into a comprehensive lens distortion equation, as

$$(x_d, y_d) = (x + \rho_x + \tau_x, y + \rho_y + \tau_y). \quad (6)$$

Estimating the unknown radial and tangential distortion coefficients involves solving complex non-linear optimization problems. Hence, several simplifying assumptions are usually made which are often not true in a real system. The most common simplifications are to assume that the principal point coincides with the center of the image and to ignore the higher degree terms [3, 7, 15, 32, 35, 39].

We model this lens distortion by *non-rational Bezier patches* instead. A non-rational Bezier patch of degree k and dimension d is defined by $(k+1)^d$ control points in a d-dimensional space. In our case, we consider a two-dimensional patch where $d = 2$. The Bezier patch is a tensor product patch where the control points, P_{ij}, are weighted by blending functions that are products of independent one-dimensional Bernstein polynomials. There are $k+1$ Bernstein polynomials of degree k given by

$$b_{nk}(t) = \frac{k!}{n!(k-n)!} t^n (1-t)^{k-n} \quad (7)$$

where $0 \leq n \leq k$ and $0 \leq t \leq 1$. The non-rational Bezier patch can then be defined from these polynomials as

$$\mathscr{NB}(x,y) = \sum_{i=0}^{k} \sum_{j=0}^{k} P_{ij} b_{ik}(x) b_{jk}(y) \quad (8)$$

$$= \sum_{i=0}^{k} \sum_{j=0}^{k} P_{ij} B_{ijk}(x,y), \quad (9)$$

where B_{ijk} is the tensor product of the Bernstein polynomials b_{ik} and b_{jk} and $0 \leq (x,y) \leq 1$. Hence, in the context of the single projector-camera pair for planar displays, the projector lens distortion function $\mathscr{R}(x,y)$ is represented by

$$(x_d, y_d) = \mathscr{R}(x,y) = \mathscr{NB}(x,y) \quad (10)$$

Our choice of using a non-rational Bezier patch is guided by the following rationale.

1. The non-rational Bezier patch provides a unified framework to represent a large range of lens distortions. More severe lens distortions can be included by increasing the *degree* of the Bezier patch. For example, a bicubic Bezier patch may be sufficient for standard narrow FOV lenses, but a Bezier patch of degree 4 may be required for the more severe distortions of wide FOV lenses. All of these are handled by the non-rational Bezier patch without affecting the underlying computational framework.

2. When Equation 10 is expanded and compared with Equation 6, one finds that the non-rational Bezier patch has many more higher-order cross terms, making it a more flexible model. For example, it is possible for short-throw lens projectors to have a mix of both barrel and pin-cushion distortions. Such cases are better handled by a more flexible model such as the Bezier patch.

3. The principal center creates a particular nuisance in traditional lens distortion estimation techniques, and forces most practical systems to assume it to be at the center of the image. When using a non-rational Bezier patch, the principle center is encoded within the control points themselves, and need not be estimated seperately.

4. Like lens distortion, non-rational Bezier patches are also affine-invariant and hence provide an accurate model.

Table 1 and Figure 1 present results on the accuracy of our Bezier patch model. We consider a dense sampling of the (x,y) space, distort it using various traditional lens distortion parameters, and fit a non-rational Bezier patch to the distorted points. To evaluate the accuracy of the fitted Bezier, we find the average Euclidian distance error between fitted and the distorted data as a percentage of the distance from

Amazon Ex. 1012.0230
Authorized licensed use limited to: Purdue University. Downloaded on April 25,2023 at 14:08:57 UTC from IEEE Xplore. Restrictions apply.

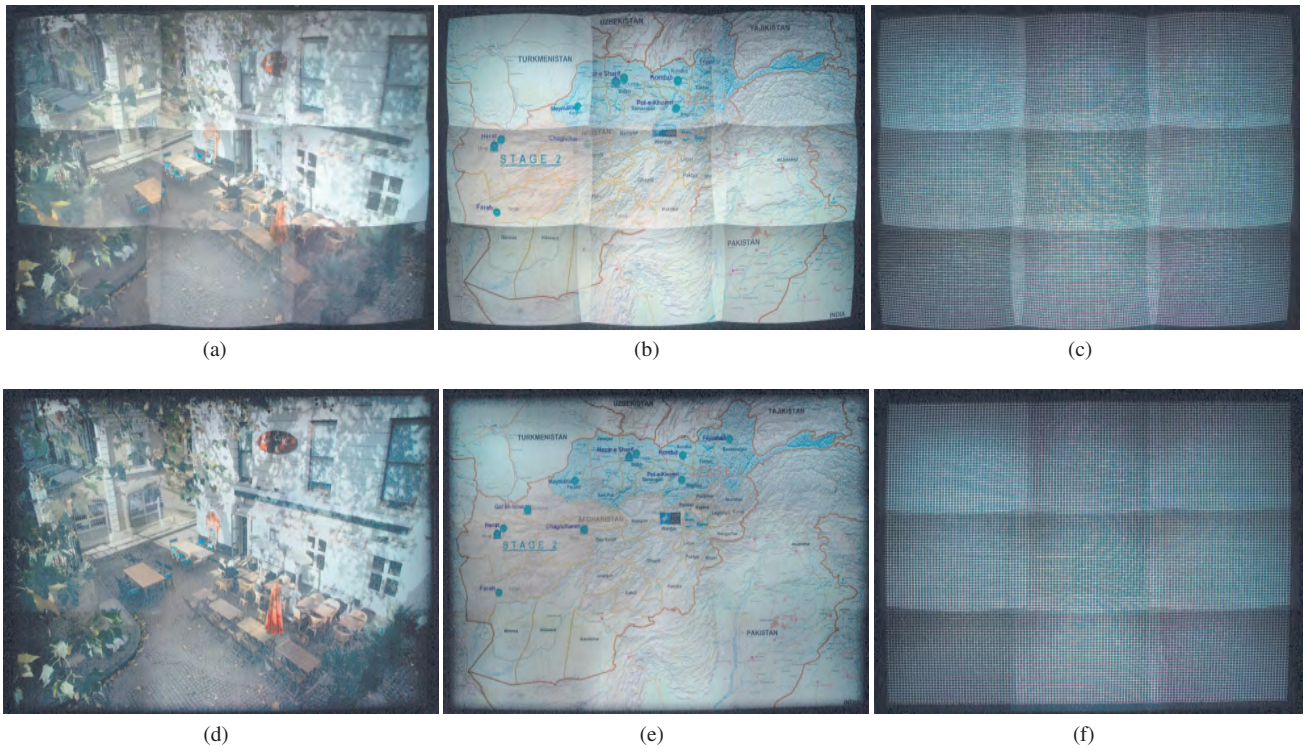


Fig. 2. Results of geometric registration shown on our display which is a 3×3 array of nine projectors. The geometric registration is achieved using bicubic rational Bezier patches with sparse correspondences (8×6 for a 1024×768 projector). (a, b, c) Before registration; (d, e, f) After registration. The grid image in the last column is made of lines that are one pixel wide. To retain clarity of the grids in the overlap region, photometric registration is not applied to the grid image in (c) when generating the registered image (f).

the principal center. Note that even when higher order distortions are present (*e.g.,* x^7), lower degree Bezier patches (*e.g.,* bicubic) can provide a good fit. This is due to the presence of a large number of lower-order cross-terms, which are absent in the traditional lens distortion model.

Modeling $\mathscr{H}\mathscr{R}$ with Rational Bezier Patches: Non-rational Bezier patches are affine-invariant but not perspective-invariant. Relating the projector to the camera not only involves a lens distortion but also a perspective projection. In the case of planar displays, this is a homography $\mathscr{H}$, a simpler 3×3 transformation induced by the planar display screen. So, to model the combined effects of $\mathscr{H}$ and $\mathscr{R}$ in Equation 1, we desire a function that retains all the properties of the Bezier and is also perspective-invariant. Retaining all the properties of Bezier patch assures retaining the accurate model of non-linearities due to distortion. Invariance under perspective projection assures the additional modeling of the perspective projection between the projector and the camera. To our advantage, a generalized form of Bezier patches called rational Bezier patches offer us a function that has all the properties of a non-rational Bezier and is also perspective-invariant. In addition to the 2D control points of Bezier patches, rational Bezier patches have a scalar weight associated with each control point, and the patch is defined as a weighted sum of the blended control points:

$$\mathscr{B}(x,y) = \frac{\sum_{i=0}^{k}\sum_{j=0}^{k} P_{ij} w_{ij} B_{ijk}(x,y)}{\sum_{i=0}^{k}\sum_{j=0}^{k} w_{ij} B_{ijk}(x,y)}. \tag{11}$$

The weights provide this function with the capability to sample the parameter space in a non-uniform (but controlled) manner. This makes it perspective-invariant. Hence, we propose the use of a rational Bezier patch to estimate the function $\mathscr{F} = \mathscr{H}\mathscr{R}$.

3.2 Geometric Registration of Multiple Projectors

Geometric registration of a multi-projector display requires estimating two functions. The first function maps the projector coordinates (x,y) to the camera coordinates (u,v). We use a rational Bezier patch $\mathscr{B}$ for this purpose. The second function maps the camera coordinates (u,v) to the normalized screen coordinates (p,q). The presence of a planar screen allows us to use a homography $\mathscr{H}_S$ for this mapping. These functions are estimated as follows.

Parameter Estimation: To estimate $\mathscr{H}_S$, we require four correspondences between the camera and screen coordinates, (u,v) and (p,q). The four correspondences result in a system of linear equations that can be solved to recover the homography matrix. To find the correspondences, we identify four points in the camera image that correspond to the four corners of the screen. This can be done automatically by attaching fiducials to the display. Since our display is nearly rectangular, we set all the projectors to display white light and use simple image processing to detect the extreme corners of the lighted regions to find the camera coordinates that correspond to the four corners of the normalized screen space.

To estimate the parameters of the rational Bezier patch, $\mathscr{B}$, we find correspondences between projector pixels (x,y) and camera pixels (u,v). Let us assume N correspondences, *i.e.,* (x_c,y_c) corresponding to (u_c,v_c), $0 \le c \le N-1$. The unknowns are the control points in the camera's space, $P_{ij} = (u_{ij}, v_{ij})$, and their scalar weights w_{ij}. From the known correspondences, Equation 11 can be written for each correspondence as follows:

$$(u_c,v_c)\sum_{i=0}^{k}\sum_{j=0}^{k} w_{ij} B_{ijk}(x_c,y_c) - \sum_{i=0}^{k}\sum_{j=0}^{k} (u_{ij},v_{ij}) w_{ij} B_{ijk}(x_c,y_c) = 0. \tag{12}$$

Assuming $U_{ij} = u_{ij} w_{ij}$ and $V_{ij} = v_{ij} w_{ij}$, Equation 12 yields two equations for each correspondence, as follows:

$$u_c \sum_{i=0}^{k}\sum_{j=0}^{k} w_{ij} B_{ijk}(x_c,y_c) - \sum_{i=0}^{k}\sum_{j=0}^{k} U_{ij} B_{ijk}(x_c,y_c) = 0. \tag{13}$$

$$v_c \sum_{i=0}^{k}\sum_{j=0}^{k} w_{ij} B_{ijk}(x_c,y_c) - \sum_{i=0}^{k}\sum_{j=0}^{k} V_{ij} B_{ijk}(x_c,y_c) = 0. \tag{14}$$

Amazon Ex. 1012.0231
Authorized licensed use limited to: Purdue University. Downloaded on April 25,2023 at 14:08:57 UTC from IEEE Xplore. Restrictions apply.

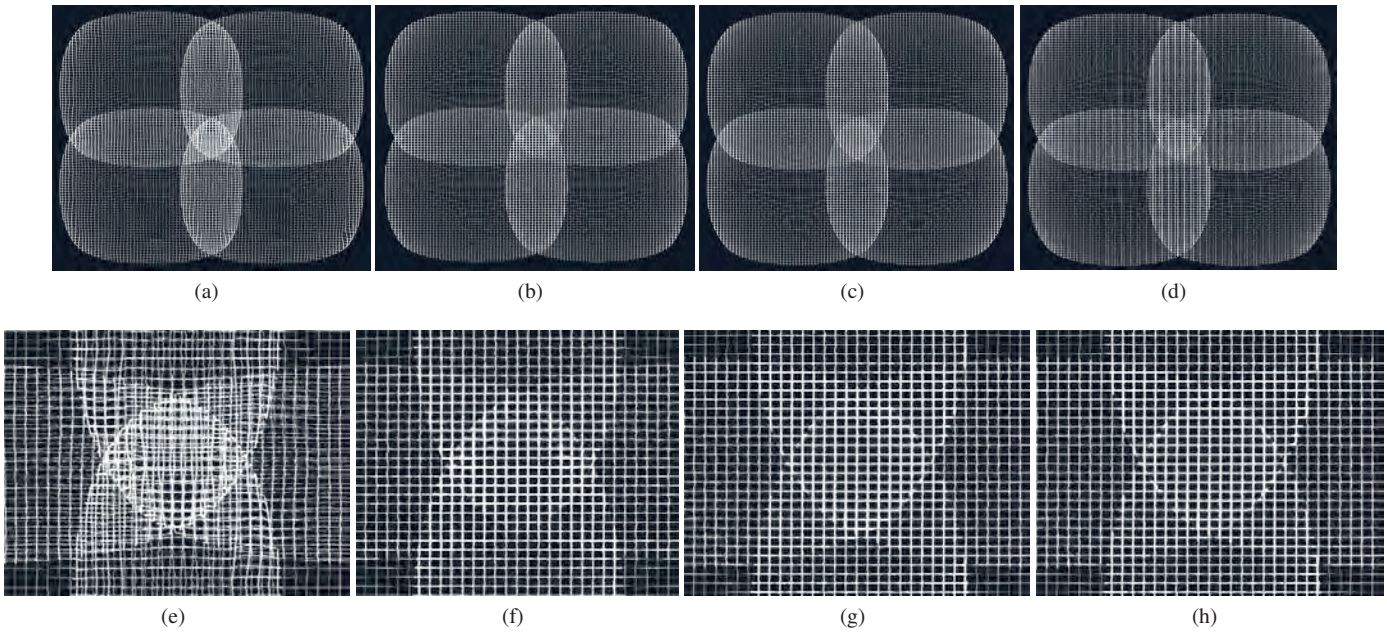


Fig. 3. Geometric registration of projectors with severe *barrel* distortion: using a linear piecewise method with sparse - 48 samples (a) and dense - 432 samples (b) correspondences; using rational *bicubic* Bezier patches with sparse - 48 samples (c) and dense - 432 samples (d) correspondences. (e,f,g,h) provides a zoomed in view of the projector overlap regions in (a,b,c,d) respectively. It is evident that piecewise linear methods with sparse correspondences lead to severe mismatches. The results using rational Bezier are similar, even when the sampling is reduced by an order of magnitude. In fact, at places, the results from using rational Bezier patches with sparse sampling is better than the results from using a dense sampling for the piecewise linear method. This is evident as the grid lines are thicker in the overlap region in (b) due to slight errors.

The unknowns for estimating a two-dimensional rational Bezier patch of degree k are $(k+1)^2$ control points and weights, which result in a total of $3(k+1)^2$ unknowns. The weights are variables that depend on the control points in a non-linear fashion. So, the above equations result in a non-linear least-squares fitting problem that can be solved efficiently using the Levenberg-Marquardt gradient descent optimization technique [31]. As an initial guess to this optimization, we fit the control polygon tightly around the set of correspondences (u_c, v_c) in the camera space and constrain the weights to be strictly greater than 0.

To find the correspondences between projector and camera coordinates, (x,y) and (u,v), we display a rectangular grid of Gaussian blobs with known coordinates on each projector. These are captured by a calibrated camera in a series of non-overlapping images. We use a 2D-stepping procedure that depends on a user to identify the top-left blob and its immediate right and bottom neighbors in camera space. With this minimal three point user input, we can design a stepping procedure that (a) estimates the rough position of the next blob in scan-line order, and (b) searches for the correct blob position using the nearest windowed center-of-mass technique [11]. If this is not possible for more extreme projector distortions, one can binary-encode the blobs and project them in a time sequential manner to recover the exact ids of the detected blobs and find the correspondences [28, 37].

Image Correction: To create a geometrically seamless and undistorted display, we warp the image for each projector as follows: For each projector coordinate (x,y), we first compute the corresponding camera coordinate (u,v) using the rational Bezier patch $\mathscr{B}$. Next, we compute the corresponding normalized screen coordinates (p,q) using the homography $\mathscr{H}_S$. The generated coordinates (p,q) are used as indices into the image being rendered on the tiled display. Bilinear interpolation is used to assign the final value of each projector pixel. Thus, the correction is achieved by the mapping

$$(p,q) = \mathscr{H}_S(u,v) = \mathscr{H}_S(\mathscr{B}(x,y)). \tag{15}$$

3.3 Implementation and Results

We have applied our method on our 8×6-foot display made of 3×3 array of nine projectors. Our projectors have relatively large throw ratios and hence do not reveal major lens distortions. Instead of choosing the cost prohibitive option of buying nine short-throw lenses, we chose to simulate the distortion digitally by distorting the input images to the projectors.

The registration is done offline and takes a couple of minutes. This generates the parameters of the rational Bezier for each projector, which are then stored to be used during image correction. We have implemented the image correction in real-time using modern GPUs through Chromium – an open-source distributed rendering engine for PC clusters [14]. A module for Chromium is written that first precomputes the coordinate-mappings of all pixel coordinates using the stored parameters of the rational Bezier. This results in a per-pixel projector to screen lookup table. A fragment program for the pixel shader quickly maps pixels from the projector coordinate space to the screen coordinate space during rendering.

Figure 2 shows the results. The advantage of using an appropriate parametric function lies in an accurate alignment even with a sparse sampling of the function. We can achieve accurate alignment of our relatively high resolution display (3000×2100) by using a low-resolution VGA (640×480) camera. Notice that even with a low-resolution camera we get sub-pixel accuracy for fine grids (one pixel wide lines) and text. Figure 3 shows that we can achieve accuracy comparable to existing methods (that use much higher resolution cameras) even when our sampling is sparser by an order of magnitude. To achieve perceptual photometric seamlessness, we used the work by Majumder, et al. [21]. To make the quality of the geometric registration visible in the grid images, we did not register the display photometrically for Figure 2(c) and Figure 3. So, these images, though geometrically aligned, still show photometric seams. The real-time rendering is demonstrated in the submitted video. However the video, being in VGA resolution, cannot capture the high resolution details.

4 Using Photometrically Uncalibrated Cameras

In this section, we present a method for photometric registration of multi-projector displays using a photometrically uncalibrated camera. Projector photometric parameters that include its ITF and spatial intensity variation, need to be estimated to instrument existing photometric registration techniques [21]. We use an uncalibrated camera (whose

Amazon Ex. 1012.0232
Authorized licensed use limited to: Purdue University. Downloaded on April 25,2023 at 14:08:57 UTC from IEEE Xplore. Restrictions apply.

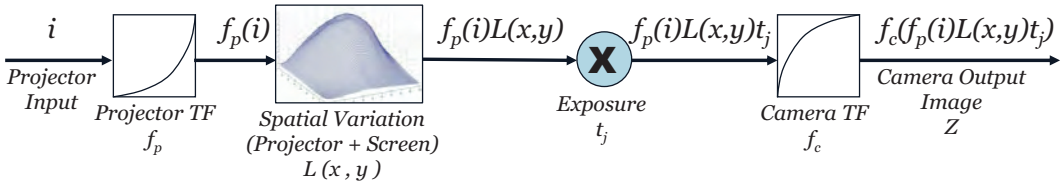


Fig. 4. The transformation process of an image as it passes through a projector-camera system.

ITF is unknown) to extract these projector photometric parameters. In the process, we also recover the camera ITF and hence calibrate the camera photometrically. To suppress the spatial intensity variation, such as vignetting effect, of the camera, like most previous work in scene reconstruction [8, 23], we operate the camera at a narrow aperture setting where it approaches the ideal pinhole model.

4.1 Model for a Single Projector

We assume a geometrically calibrated projector-camera system where a pixel (u,v) in the camera coordinate system is related to a pixel (x,y) in the projector coordinate system by a linear or non-linear warp $G(x,y) = (u,v)$. G can be determined by any geometric calibration method, including the one presented in the previous section [38].

Consider a spatially uniform grayscale input to the projector. Let the grayscale level be denoted by i. As per the model presented in the paper by Majumder and Gopi [18], the uniform image is first transformed by a spatially invariant input transfer function of the projector, f_p, to create a spatially uniform output, $f_p(i)$. Next, the projector optics introduce a spatially-dependent but input-independent intensity variation. This is further modulated by the screen reflectance/transmittance function. We denote this combined spatially dependent attenuation factor due to projector lens and screen as $L(x,y)$. Thus, the image produced after passing through the projector lens and screen is $f_p(i)L(x,y)$. The light from the screen then reaches the camera. The amount of light accepted by the camera is scaled by its exposure time t_j, where j indexes different exposure times. The different exposures are instrumented by changing the shutter speed of the camera. This produces an image $f_p(i)L(x,y)t_j$ that passes through the camera optics. Assuming negligible camera vignetting at narrow apertures, the image is transformed only by the input transfer function of the camera, f_c, to generate the grayscale value recorded by the camera Z. Thus, Z is a function of the input i, the exposure time index j, and the spatial coordinates (x,y). This is illustrated in Figure 4. The final equation is

$$Z(i,j,x,y) = f_c(f_p(i)L(x,y)t_j). \tag{16}$$

For cameras, the intensity transfer function is monotonic [8], and hence invertible. Note that the same is not true for projectors [20]. Assuming invertible f_c, the above equation becomes

$$f_c^{-1}(Z(i,j,x,y)) = f_p(i)L(x,y)t_j. \tag{17}$$

Taking the natural logarithm of both sides we get,

$$lnf_c^{-1}(Z(i,j,x,y)) = lnf_p(i) + ln(L(x,y)) + ln(t_j). \tag{18}$$

To simplify the notation, we define $h_c = lnf_c^{-1}$ and $h_p = lnf_p$, The above equation then becomes

$$h_c(Z(i,j,x,y)) = h_p(i) + ln(L(x,y)) + ln(t_j) \tag{19}$$

where i ranges over the grayscale inputs, j ranges over the exposure times, and (x,y) ranges over the spatial coordinates of the projector. In this equation, Z and t_j are known while h_p, h_c and L are unknown. Varying i and t_j results in different values of Z for each pixel (x,y). We can use this to setup a system of linear equations. We recover h_p, h_c and L that best satisfy Equation 19 in a least-squares sense. Note that recovering h_p and h_c involves solving the functions for a finite number of samples in the complete range of input values. The recovered parameters are illustrated in Figure 5.

Improving Efficiency: The system of linear equations achieved by Equation 19 is very large. Let R_i and R_Z each be the cardinality of the set of input samples for h_p and h_c respectively, P be the total number of projector pixels, and T be the number of exposures. This results in PTR_i equations and $P + R_i + R_Z$ unknown variables for the system of equations defined by Equation 19. Typically $R_i = R_Z = 256$ and $P = 1,000,000$ (assuming common XVGA projector resolution). Since there are multiple exposures for each input in the range R_i, the size of the linear system is on the order of at least a few million equations and cannot be solved efficiently.

To address this inefficiency we use a limited number of pixels in the projector space to solve for h_p and h_c. This is possible since h_p and h_c are both spatially constant. To ensure a sufficiently over-determined system, the criteria $PTR_i > P + R_i + R_z$ should be satisfied. For $R_i = R_Z = 256$ and $T = 6$, a choice of 100 for P is more than adequate. So, we first subsample L to a resolution of 10×10 pixel and use this to setup a smaller linear system of equations to solve for h_p and h_c.

With the estimated h_p and h_c, we substitute these into Equation 19 and quickly solve for $L(x,y)$ at the various projector coordinates, by rewriting Equation 19 as

$$ln(L(x,y)) = h_c(Z(i,j,x,y)) - h_p(i) - ln(t_j) \tag{20}$$

Handling Noise: Our system does exhibit considerable random noise arising not only from the devices (camera and projector) but also from the screen. In particular, we use a rear projection screen with relatively high gain which has been shown to generate considerable random noise [20]. This has an adverse effect on the signal to noise ratio, especially for low projector input i or low camera output Z. Thus, estimating h_p and h_c directly from Equation 19 and L from Equation 20 results in noisy parameter estimation. We take several measures to reduce the effect of this random noise.

To assure smooth h_p and h_c while solving the system of equations, we minimize the error function

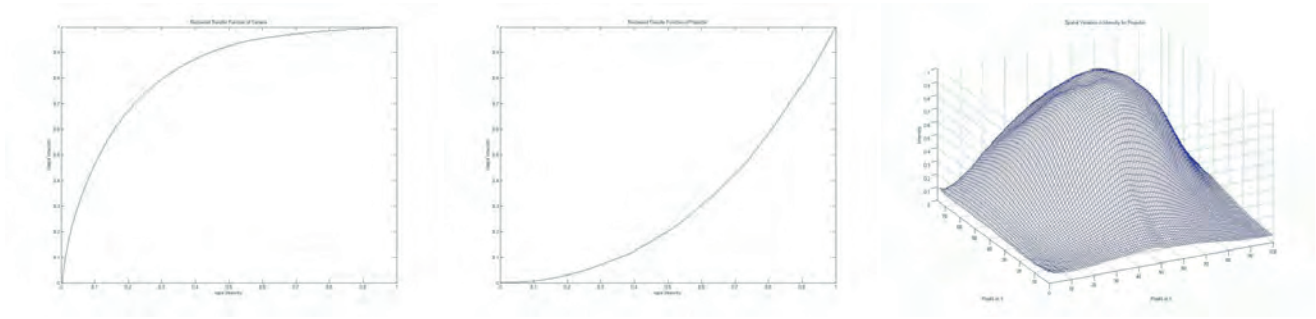
$$E = \sum_{j\in T} \sum_{(x,y)\in P} \sum_{i\in R_i} [(h_c(Z) - h_p(i) - ln(L(x,y)) - ln(t_j)]^2 \tag{21}$$

$$+\lambda \left(\sum_{Z\in R_Z} h_c''(Z)^2 + \sum_{i\in R_i} h_p''(i)^2 \right). \tag{22}$$

The first term assures that the solution arises from the set of equations given by Equation 19 in a least-squares sense. The second term is a smoothing constraint on the curvature of h_p and h_c, given by their second derivatives. In the discrete domain, we use the Laplacian operator to find the curvature of h_p and h_c, *i.e.*, $h_p''(i) = h_p(i-1) - 2h_p(i) + h_p(i+1)$. The scale factor λ weights the smoothness term relative to the data fitting term and should be chosen based on the amount of noise in Z. We choose $\lambda = 10$.

However, note that the first term in Equation 22 gives equal weights to all recorded camera values Z, and the second term gives equal weights to all i and Z. For the first term, we want to emphasize images with higher signal to noise ratio i.e. images for higher i with higher recorded values Z. For the second term, we want to emphasize smoothing of h_p and h_c in the lower ranges of i and Z respectively where low signal to noise ratio would result in a noisier estimate. To

Amazon Ex. 1012.0233
Authorized licensed use limited to: Purdue University. Downloaded on April 25,2023 at 14:08:57 UTC from IEEE Xplore. Restrictions apply.



(a) (b) (c)

Fig. 5. (a) The estimated camera input transfer function f_c. (b) The estimated projector input transfer function f_p. (c) The estimated spatial intensity variation due to projector, screen, and camera L.

achieve this, we weight the first and second term of the error function in Equation 22 as follows,

$$E = \sum_{j\in T}\sum_{(x,y)\in P}\sum_{i\in R_i}[w_c(Z)(h_c(Z) - h_p(i) - ln(L(x,y)) - ln(t_j)]^2 \quad (23)$$

$$+\lambda\left(\sum_{Z\in R_Z}[(Z_{max} - w_c(Z))h_c''(Z)]^2 + \sum_{i\in R_i}[(i_{max} - w_p(i))h_p''(i)]^2\right). \quad (24)$$

where w_p and w_c are the weighting functions corresponding to the projector input and the recorded camera values, respectively, and Z_{max} and i_{max} are the maximum camera output and projector input respectively. To give the higher intensities greater confidence we use linear weighting functions, $w_c(Z) = Z$ and $w_p(i) = i$.

Similarly, noise can have adverse effects when estimating L using Equation 20 after h_p and h_c are recovered. Usually the image captured for input i at a particular exposure does not yield unsaturated outputs at all spatial pixel locations (x,y). So, we need to use different images for reconstruction of L at different spatial locations. This will yield an L with different signal to noise ratio at different spatial regions depending on the image that was used to reconstruct L in any spatial region. Even if a rare situation occurs when one finds a single image (at a particular input and exposure) where all the pixels are unsaturated, L will still be noisy due to the inherent random noise of the projector-screen-camera system.

Averaging multiple observations is the standard way to reduce random noise significantly [11]. Using this fact, we reduce noise in L at any pixel (x,y) by solving it from multiple images and finding a weighted mean of the multiple solutions. The weighting function should emphasize higher i and Z to reduce the effect of random noise. We achieve this by a weighting function $w_L(Z,i) = w_c(Z)w_p(i)$. The solution can now be derived by modifying Equation 20 as

$$ln(L(x,y)) = \frac{\sum_{j\in T}\sum_{i\in R_i} w_L(Z,i)[h_c(Z) - h_p(i) - ln(t_j)]}{\sum_{j\in T}\sum_{i\in R_i} w_L(Z,i)}, \quad (25)$$

resulting in clean estimated parameters as illustrated in Figure 5.

4.2 Photometric Registration of Multiple Projectors

For photometric registration of a multi-projector display, we estimate the f_p and $L(x,y)$ of each projector independently using the same camera. Note that in order to recover the relative intensity levels of the projectors, we use the unnormalized $L(x,y)$ estimated independently from each projector. Using the extracted parameters from each projector, we use the photometric registration method proposed by Majumder and Stevens [21] to generate a perceptually seamless display.

4.3 Implementation and Results

We used a Kodak DCS ProSLR/n camera and projected 32 flat grayscale fields with intensity levels uniformly sampled from 0 to 255. For each intensity level, 15 exposures were taken. The camera was set to a narrow aperture, f/32. The data collection took 25 minutes. To reduce the collection time, we tried reducing the number of exposures. Empirically, the minimum number of exposures needed before quantization effects became visible was eight. The exposures used, however, need to be well distributed amongst the range of available camera exposures. The program was written in C++ and utilizes Matlab and the OpenCV library. It takes about 11-15 minutes to register each projector on a Pentium 4, 2.8 GHz PC.

The photometric registration in Figure 2 (d) and (e) are achieved by using a photometrically uncalibrated camera. Additional results with projectors that do not exhibit lens distortion are shown in Figure 6. Here a linear geometric registration method is used [5]. These results show that a photometrically uncalibrated camera can be used irrespective of the kind of geometric registration method being used. When using the photometric registration technique of Majumder and Stevens [21] we use a low photometric uniformity parameter to retain the high contrast of the display as much as possible. This results in some perceivable seams near the boundaries of the bottom left and right projectors, which are significantly dimmer than the other projectors in our display. However, these seams exist even when the photometric registration is carried out with a calibrated camera, showing that they are artifacts of the photometric registration method itself and are not due to imperfect projector parameter estimation.

5 Conclusion

This paper presents new registration techniques to handle imperfect and uncalibrated devices in planar multi-projector displays. The four relevant cases include allowing (a) photometrically imperfect projectors, (b) geometrically imperfect projectors, (c) a photometrically uncalibrated camera, and (d) a geometrically uncalibrated camera. Our work addresses (b) and (c) in particular, while (a) is addressed by existing literature [21]. Our work allows severely distorted projectors to be used which opens up the possibility of mounting inexpensive lenses on projectors, just as with cameras, leading to very short-throw projectors. Using a photometrically uncalibrated camera allows the photometric registration to be conducted within a controlled indoor environment. Further, we have proposed the first method that estimates all projector/camera photometric parameters simultaneously. However, case (d) above of achieving geometric registration in multi-projector displays using a geometrically uncalibrated camera is still a challenging problem.

Further, our methods are within the realm of centralized architecture. Current distributed architectures [2] are very amenable to easy deployment, simple maintenance and hence more affordable multi-projector displays. Our goal is to extend our developed methodologies in this paper to the distributed framework of a network of projector-camera systems [2]. This poses a significant challenge which, if resolved, can be instrumental in realizing commodity tiled projection-based displays. All of these methodologies are stepping stones towards the vision of *ubiquitous pixels* in future workspaces.

Amazon Ex. 1012.0234
Authorized licensed use limited to: Purdue University. Downloaded on April 25,2023 at 14:08:57 UTC from IEEE Xplore. Restrictions apply.



(a) (b) (c) (d)

Fig. 6. Images of our 3×3 array of nine projectors. (a,c) The geometrically registered display without any photometric registration. (b,d) Photometric registration achieved by using an uncalibrated camera.

REFERENCES

[1] Mitsubishi pocketprojector led dlp projector. Technical report, http://www.mitsubishi-presentations.com/proj_pocket.asp.

[2] E. S. Bhasker, P. Sinha, and A. Majumder. Asynchronous distributed calibration for scalable reconfigurable multi-projector displays. *IEEE Transactions of Visualization and Computer Graphics (TVCG)*, 12(5):1101–1108, 2006.

[3] J. Y. Bouguet. Camera calibration toolbox for matlab. Technical report, http://www.vision.caltech.edu/bouguetj/calib_doc/index.html.

[4] D. Brown. Decentering distortion of lenses. *Photometric Engineering*, 32(3):444 – 462, 1966.

[5] H. Chen, R. Sukthankar, G. Wallace, and K. Li. Scalable alignment of large-format multi-projector displays using camera homography trees. *Proceedings of IEEE Visualization*, 2002.

[6] Y. Chen, D. W. Clark, A. Finkelstein, T. Housel, and K. Li. Automatic alignment of high-resolution multi-projector displays using an uncalibrated camera. *Proceedings of IEEE Visualization*, 2000.

[7] K. Cornelis, M. Pollefeys, and L. V. Gool. Lens distortion recovery for accurate sequential structure and motion recovery. *European Conference in Computer Vision*, pages 186 – 200, 2002.

[8] P. E. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. *Proceedings of ACM Siggraph*, pages 369–378, 1997.

[9] T. Disz, M. Papka, and R. Stevens. Ubiworld: An environment integrating virtual reality. *Heterogeneous Computing Workshop, Geneva, Switzerland*, 1997.

[10] J. Fryer and D. Brown. Lens distortion for close-range photogrammetry. *Photogrammetric Engineering and Remote Sensing*, 52(1):51 – 58, 1986.

[11] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Addison Wesley, 1992.

[12] M. Hereld, I. Judson, and R. Stevens. Introduction to building projection-based tiled display systems. *IEEE Computer Graphics and Applications*, 2000.

[13] M. Hereld, I. R. Judson, and R. Stevens. Dottytoto: A measurement engine for aligning multi-projector display systems. *Argonne National Laboratory preprint ANL/MCS-P958-0502*, 2002.

[14] G. Humphreys, M. Houston, R. Ng, R. Frank, S. Ahem, P. Kirchner, and J. Klosowski. Chromium : A stream processing framework for interactive rendering on clusters. *ACM Transactions of Graphics*, 2002.

[15] H. G. Jung, Y. H. Lee, P. J. Yoon, and J. Kim. Radial distortion refinement by inverse mapping-based extrapolation. *Proceedings to 18th International Conference on Pattern Recognition(ICPR)*, 2006.

[16] J. Lai, A. Levas, P. Chou, C. Pinhanez, and M. Viveros. Bluespace: Personalizing workspace through awareness and adaptability. *International Journal of Human Computer Studies*, 57(5):415 – 428, 2002.

[17] K. Li, H. Chen, Y. Chen, D. W. Clark, P. Cook, S. Damianakis, G. Essl, A. Finkelstein, T. Funkhouser, A. Klein, Z. Liu, E. Praun, R. Samanta, B. Shedd, J. P. Singh, G. Tzanetakis, and J. Zheng. Early experiences and challenges in building and using a scalable display wall system. *IEEE Computer Graphics and Applications*, 20(4):671–680, 2000.

[18] A. Majumder and M. Gopi. Modeling color properties of tiled displays. *Computer Graphics Forum*, June 2005.

[19] A. Majumder, Z. He, H. Towles, and G. Welch. Achieving color uniformity across multi-projector displays. *Proceedings of IEEE Visualization*, 2000.

[20] A. Majumder and R. Stevens. Color nonuniformity in projection-based displays: Analysis and solutions. *IEEE Transactions on Visualization and Computer Graphics*, 10(2), March–April 2003.

[21] A. Majumder and R. Stevens. Perceptual photometric seamlessness in tiled projection-based displays. *ACM Transactions on Graphics*, 24(1), January 2005.

[22] T. Mitsunaga and S. Nayar. Radiometric self calibration. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1999.

[23] S. K. Nayar, H. Peri, M. D. Grossberg, and P. N. Belhumeur. A projection system with radiometric compensation for screen imperfections. *Proceedings of IEEE International Workshop on Projector-Camera Systems*, 2003.

[24] C. Pinhanez. The everywhere displays projector: A device to create ubiquitous graphical interfaces. *Proceedings of Ubiquitous Computing, Atlanta, Georgia*, 2001.

[25] C. S. Pinhanez, F. C. Kjeldsen, A. Levas, G. S. Pingali, M. E. Podlaseck, and P. B. Chou. Ubiquitous interactive graphics. Technical Report RC22495 (W0205-143), IBM Research, 2002.

[26] A. Raij, G. Gill, A. Majumder, H. Towles, and H. Fuchs. PixelFlex2: A Comprehensive, Automatic, Casually-Aligned Multi-Projector Display. *IEEE International Workshop on Projector-Camera Systems*, 2003.

[27] R. Raskar. Immersive planar displays using roughly aligned projectors. In *Proceedings of IEEE Virtual Reality 2000*, 1999.

[28] R. Raskar, M. Brown, R. Yang, W. Chen, H. Towles, B. Seales, and H. Fuchs. Multi projector displays using camera based registration. *Proceedings of IEEE Visualization*, 1999.

[29] R. Raskar, J. van Baar, P. Beardsley, T. Willwacher, S. Rao, and C. Forlines. ilamps: Geometrically aware and self-configuring projectors. *ACM Transactions on Graphics*, 22(3), 2003.

[30] 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 display. In *Proceedings of ACM Siggraph*, pages 168–176, 1998.

[31] S. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. Prentice-Hall, 1995.

[32] J. Salvi, X. Armangu, and J. Batlle. A comparative review of camera calibrating methods with accuracy evaluation. *Pattern Recognition*, 35:1617 – 1635, 2002.

[33] P. Steurer and M. B. Srivastava. System design of smart table. *Proceedings of the First IEEE International Conference on Pervasive Computing and Communications*, 2003.

[34] R. Sukthankar, R. Stockton, and M. Mullin. Smarter presentations: Exploiting homography in cameraprojector systems. *Proceedings of International Conference on Computer Vision (ICCV)*, 2001.

[35] R. Swaminathan and S. K. Nayar. Nonmetric calibration of wide-angle lenses and polycameras. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1172 – 1178, 2000.

[36] J. Underkoffler, B. Ullmer, and H. Ishii. Emancipated pixels: Real-world graphics in the luminous room. *Proceedings of ACM Siggraph*, 1999.

[37] R. Yang, D. Gotz, J. Hensley, H. Towles, and M. S. Brown. Pixelflex: A reconfigurable multi-projector display system. *Proceedings of IEEE Visualization*, 2001.

[38] R. Yang, A. Majumder, and M. Brown. Camera based calibration techniques for seamless multi-projector displays. *IEEE Transactions on Visualization and Computer Graphics*, 11(2), March-April 2005.

[39] Z. Zhang. Flexible camera calibration by viewing a plane from unknown orientation. *IEEE International Conference on Computer Vision*, pages 666 – 673, 1999.

Amazon Ex. 1012.0235

Authorized licensed use limited to: Purdue University. Downloaded on April 25,2023 at 14:08:57 UTC from IEEE Xplore. Restrictions apply.

Appendix B-8

User-Centric Design of a Vision System for Interactive Applications

Stanislaw Borkowski[†], Julien Letessier[†§], François Bérard[§], and James L. Crowley[†]
†INRIA Rhône-alpes
GRAVIR laboratory – PRIMA group
655 avenue de l'Europe – Montbonnot
38334 Saint Ismier Cedex – France
{stan.borkowski,julien.letessier,james.crowley}@inrialpes.fr

§CLIPS – IMAG
HCI group
BP 53
38041 Grenoble Cedex 9 – France
francois.berard@imag.fr

Abstract

Despite great promise of vision-based user interfaces, commercial employment of such systems remains marginal. Most vision-based interactive systems are one-time, "proof of concept" prototypes that demonstrate the interest of a particular image treatment applied to interaction. In general, vision systems require parameter tuning, both during setup and at runtime, and are thus difficult to handle by non-experts in computer vision. In this paper, we present a pragmatic, developer-centric, service-oriented framework for the construction of vision-based interactive systems. Our framework is designed to allow developers unfamiliar with vision to use computer vision as an interaction modality. To achieve this goal, we address specific developer- and interaction-centric requirements during the design of our system. We validate our approach with an implementation of standard GUI widgets (buttons and sliders) based on computer vision.

1. Introduction

Computer vision has the potential to provide a rich modality for user input in interactive systems. Numerous examples from the literature illustrate the diversity of application domains and interaction styles created using vision as input modality [7, 11, 9, 14, 2, 12, 10]. Yet, there are surprisingly few examples of commercially available vision-based interactive systems.

Human-computer interface (HCI) developers are used to hardware-based input devices that provide data "out of the box" with little or no off-line or runtime tuning. In contrast, vision-based input systems can only be integrated by HCI developers with great difficulty: they require vision expertise for setup, as well as execution conditions incompatible with typical HCI requirements. In this sense, most current computer vision systems do not *just work*.

We believe that to popularize vision-based interfaces, computer vision systems must be made more developer-friendly as well as more user-friendly. To accomplish this, it is necessary that vision system developers adopt a different perspective. In this paper, we present an approach to vision system design that aims at minimizing the difficulties related to the deployment of vision-based interactive systems by: *(a)* encapsulating vision components in isolated services, *(b)* imposing these services to meet specific usability requirements, and *(c)* limiting communications between the services and the interactive applications to a minimum. We describe our approach in section 3.

In section 4, we design a system for vision-based interactive widgets for augmented surfaces, and propose a breakdown into atomic services. In this context, we identify the necessary information exchange between services and with the application. Meeting the requirements detailed in section 3 influences the choices of the image processing algorithms used; in particular, they require little or no setup and tuning and feature usability-grade latency. Our implementation is robust to light changes and is moderate in CPU usage. Finally, in section 5 we detail an implementation of a simple vision-based calculator built using the described vision system.

In this paper, we consider the requirements of developers and end-users. By "developers", we refer to the developers of the interactive systems aimed at "end-users". In particular, these are not the developers of the vision system and we do not expect them to have computer vision expertise. By "end-users", we refer to the users of the final system, who are assumed to have neither developer nor computer vision expertise.

2. Related work

Much work has been done in vision-based gesture recognition and novel interaction styles research, but the problem

Proceedings of the Fourth IEEE International Conference on Computer Vision Systems (ICVS 2006)
0-7695-2506-7/06 $20.00 © 2006 IEEE

IEEE COMPUTER SOCIETY
Amazon Ex. 1012.0237
Authorized licensed use limited to: Purdue University. Downloaded on April 26,2023 at 14:07:55 UTC from IEEE Xplore. Restrictions apply.

of integrating vision to standard development has only recently received attention. Kjeldsen et al.[5] present an architecture for dynamically reconfigurable vision-based user interfaces (VB-UIs). They propose to separate the application from the vision engine. The core application controls the vision engine by sending XML messages specifying the current interface configuration. The interface is composed of interaction widgets defined at the level of functional core by their function and position in the interface. As end-users actuate widgets, the image processing engine sends messages that correspond to interaction events to the application's functional core.

The modular decomposition of vision components, as well as the separation of image treatment from the functional core of applications are important steps toward "developer-usable" VB-UI components. However, as presented, the VB-UI approach requires its developer to define sets of parameters and calibration data for the vision treatment engine for each interface configuration and location. Setting parameters for particular conditions allows fine tuning of the vision system. On the other hand, it requires the user of the vision system (the developer) to understand the underlying image processing algorithms, which is contradictory with the idea of VB-UI deployment by developers unfamiliar with vision.

In [13] Ye et al. present a framework for vision-based interactive components design. Both Kjeldsen et al. and Ye et al. propose to analyze camera input only in regions of interest around interactive "zones" of the interface, thus making the vision engine more economic in CPU usage. However Ye et al. go a step further, and instead of applying one vision algorithm per widget type, they propose to use a set of different detection techniques to obtain visual events called Visual Interface Cues (VIC). Each VIC detector called also *selector* examines a small part of the camera image for visual events that can be of different nature, like color, texture, motion or object geometry. Selectors are structured into hierarchies, and each selector can trigger one or more other selectors. Interaction events are detected based on sequences of selectors output. For instance, a touch sensitive button uses only motion detection at first. Once motion is detected, a VIC-based button switches to color-based image segmentation and to segmentation-based gesture recognition.

While the VIC paradigm presented in [13] is appealing for widget-oriented interface design, it lacks the analysis of VIC selectors suitability for interactive systems. Indeed, the VICs framework is presented from the perspective of an expert in computer vision. It does not consider setup and maintenance issues related to each selector such as threshold setting or lighting requirements. In section 4 we present a VIC-based implementation of basic interactive widgets that was conceived to respect the requirements identified in section 3.

Finally, the Papier-Mâché [6] toolkit features a framework for the creation of multimodal interactive applications, in particular using computer vision. The toolkit design is based on a detailed user study, and offers an abstract, event-based model to work with multiple modalities. The vision part allows object recognition and tracking to be used as application input. While the design seems to be well-thought from the HCI perspective, it has several architectural shortcomings; in particular typical problems linked to object recognition (such as the aspect variations due to lighting changes) are not dealt with, and several thresholds must be set manually, on-line, by the developer using the toolkit.

3. Service-oriented design

Vision-based user interfaces rely on vision processes that extract high-level, abstract information from streams of images. This information is what is relevant for the interaction task. Except for "proof of concept" demonstrations created by computer vision researchers, this information is to be used by interactive application developers who have little to no expertise in the computer vision. This implies that a vision system that aims to be used outside the laboratory must be designed from a developer-centric point of view. In addition, all interactive systems, including vision based once, must consider end-user requirements. In this section, we describe how user-centric and developer-centric requirements influence both the structure of a VB system and its application programming interface (API).

Our goal is to provide a set of computer vision services to developers of interactive systems. We aim at offering these services at different levels of abstraction. At the highest level, a service is an autonomous "black-box" application that, once launched, provides abstract information through some communication channel (a network socket, or a shared memory for examples). Developers should not need any knowledge of the vision processes. They simply connect to the black-box application and receive abstract events such as the position of the fingers in the case of a finger tracker service, or activation of striplets in the case of the SPOD service that we describe later. We facilitate the communication between the vision application and the developer's application by using a very simple communication protocol. This approach is developer-centric in the sense that we only expose and make easy to access what matters to the developer: the high-level events.

3.1. Non-functional requirements

Typical user-centric criteria against which an interactive system is evaluated include overall latency, reliability and

IEEE COMPUTER SOCIETY
Amazon Ex. 1012.0238
Authorized licensed use limited to: Purdue University. Downloaded on April 26,2023 at 14:07:55 UTC from IEEE Xplore. Restrictions apply.

autonomy [7]. Here, we shortly describe these criteria and how they constrain the architecture of vision systems.

Latency. Interactive systems generally require a constraint on latency. For user input systems latency is measured as the time between the user action and the notification of the application. Typically, when using a vision-based finger tracker or a mouse to drag projected objects on a surface, the latency must be under 50 ms to optimize usability [8]. On the other hand, for a system that monitors the number of persons present in a room, a latency of 1 second may be acceptable. In consequence, a general architecture for vision services must be able to respect strict latency limits in communicating information to the application.

Autonomy. Vision-based systems usually require operator intervention for initial setup and maintenance. While acceptable in a laboratory setting, such intervention is poorly suited for deployment. Our "black-box" approach reduces the possibility of operator maintenance by reducing communication between the vision service and the developer's application to a minimum.

Ideally, vision system developers should devise automatic maintenance solutions without offering a maintenance API. However, we acknowledge that designing fully autonomous vision system is a difficult goal that will not be reached in the short run for many services. When operator intervention is mandatory, one must take into account that the calibration task will be performed by a non-expert. Therefore, the maintenance procedure must be designed appropriately to minimize disrupting the end-user from her/his primary tasks.

Reliability. In most cases, information needs to be reliably conveyed between the input system and the application. For instance, an application that receives events when a person enters or exits an augmented room cannot afford to "miss" an event because its state would become incoherent. Losing some events from a tracking system, on the other hand, would not necessarily break the interaction, but would deteriorate the end-user experience. Therefore, our architecture must include reliable transport for interactive events.

3.2 Functional requirements

We propose a number of developer-centric requirements for our system in order to successfully address the target audience. These requirements include abstraction, isolation, and contract. Together they form the notion of a service.

Abstraction. To HCI developers, "abstracting the input [is] the most time consuming and challenging piece of application development" [6]. Since we assume the user of our vision system (the application developer) has no vision expertise, vision-specific information should not be made visible.

Even though coupling processing results with a confidence factor might be richer, this information is, in most cases, of no interest to developers. For instance, optical mice use normalized cross-correlation to determine the direction and speed of the movement. As long as the correlation coefficient is above a threshold they send positional data to the computer, otherwise they remain silent. The correlation coefficient is not sent to the system. The services internal data, such as the confidence factor, should not be considered as first-class output of services and should not be made visible by default. However, in cases where developers need to access this information, it should be available by explicit request from the developer.

End-user input should be abstracted to interaction events that are the same for a given class of input. For example, a finger tracker should produce "motion" events compatible with mouse "motion" events so that developers can, as a first step, quickly substitute vision-based input to existing event processing code. In other words, the abstracted data from the vision service should be in a form that is easily connectable to existing graphical user interface code. Consequently, the API of a VB input system for interaction should *(a)* render the vision aspects invisible, and *(b)* generalize the input used for a given task.

Isolation. Input subsystems like a VB-UI may need to be used by multiple applications, possibly running on different machines – for performance or geographic reasons. For instance, a video capture service might be used concurrently by a surveillance system located on a remote server, by a video-conference application, and by a motion capture service, both co-located to the camera. Since "a particular piece of input can be used for many different types of output" [6], information generated by a vision system must be shareable, and both remotely and locally accessible. Moreover, any input system for interaction should be easily extensible (for instance using output adapters, aggregators, or supervisor patterns) and embeddable in other services or in applications. Designing a VB input system as a federation of "black box" components therefore appears to be adequate.

Contract. Developers of toolkits or input devices usually establish a contract with the UI developer, in terms of HCI-centric, non-functional criteria. For instance, for positional or tracking input (mice, trackpads, laser trackers, etc.), the

Proceedings of the Fourth IEEE International Conference on Computer Vision Systems (ICVS 2006)
0-7695-2506-7/06 $20.00 © 2006 IEEE
IEEE COMPUTER SOCIETY
Amazon Ex. 1012.0239
Authorized licensed use limited to: Purdue University. Downloaded on April 26,2023 at 14:07:55 UTC from IEEE Xplore. Restrictions apply.

relevant criteria are latency, precision, robustness, and autonomy.

- *latency* is implicitly under a usability threshold for standard tracking input devices (the textbook value [8] is around 50 ms);
- *precision* is also implicitly defined for a mouse/trackpad (under 1 display pixel, scaled with respect to the device gain)
- *robustness* is generally "absolute". In other words the device or system either works or doesn't. For instance an optical mouse "just works" as long as used on an adequate surface, and stops emitting positional events as soon as the contact stops.
- *autonomy*, or the lack of setup and maintenance, is also tacit for traditional devices.

In the case of VB input systems, these criteria are difficult to meet. Therefore, it is the role of the vision system designer to explicitly state in contract form the limitations with respect to these criteria. For instance, the system should notify the client application about robustness-related failures rather than provide incorrect or distorted information along with a (low) confidence factor. This corresponds to a binary quality of service evaluation and requires the system to perform introspection.

3.3. A pragmatic approach

Our approach is to generally isolate as much as possible the VB input system from the application that uses it, and minimize the communications between them. We propose to encapsulate the relevant services into independent, "black-box" processes that use only serialized communications.

To allow for low latency while preserving reliability, we propose to use traditional socket-based (TCP and UDP) communications. We use the TCP link to guarantee the connection between services or with the application, and to allow for reliable, high-latency communications, and the UDP link for low-latency communications. This constitutes the base of the "BIP/1.0" protocol we use for communications between services (draft specification at `http://www-prima.inrialpes.fr/prima/pub/Publications/index.php`).

System autonomy cannot be enforced at the architectural level. Nevertheless, not providing support for synchronous communications (thus making it difficult to implement transactions between the application and a service) limits the possibility of implementing setup mechanisms where automatic setup can be devised.

Finally, since the *contract* offered by a VB input system cannot be expressed simply as a part of an API, we propose to document it explicitly. This means that the developer of the system must evaluate it against the criteria presented above in order to circumscribe conditions of use for the vision system.

4. Basic services for vision-based UI design

To illustrate our service-oriented approach with an implementation, we choose to study vision systems applied to traditional interfaces. Typically, we investigate the use of graphical WIMP-like interfaces projected on surfaces of mundane objects. We assume that end-users are free to interact with projected images without wearing markers or actuating any hardware.

We break down our VB input system into three services: *(1)* the widget service, which the client will interoperate with; *(2)* the image capture service, which abstracts the camera, allowing it to be used by other systems eventually; and *(3)* the calibration service, which establishes the geometrical mapping between the camera view and the display.

4.1. Simple Pattern Occlusion Detectors

In [3] we presented an appearance-based implementation of touch sensitive projected buttons which we called "Sensitive widgets". The presence of an object over a button on the interaction surface is detected by observing the change of perceived luminance over the button center area with respect to a reference area. By defining the reference area around the central-one, the button is made robust to complete occlusion, and sensitive to appearance changes made by elongated objects. Very simple image processing allows this system to maintain several dozens of sensitive widgets at camera frame rate (PAL-size images at 25 Hz) on a typical desktop computer. Moreover, this approach is robust to lighting changes, and thus suitable for configurations in which the interface is projected frontally from a video projector.

Implementation and evaluation of several interface prototypes based on sensitive widgets demonstrated that, from the end-user's perspective, robustness to partial occlusions is also necessary. Indeed, a user pointing at a far-part of the interface would likely hover her/his arm over projected buttons, thus triggering partially occluded widgets. A partial, unsatisfactory solution was obtained by deactivating partly occluded widgets based on the input from widgets placed further away from the user. The idea of combining inputs from several sensitive widgets led us to re-think the touch detection approach.

We now choose to assemble atomic occlusion detectors, which are to be placed within and around widgets, in a way

Proceedings of the Fourth IEEE International Conference on Computer Vision Systems (ICVS 2006)
0-7695-2506-7/06 $20.00 © 2006 IEEE
IEEE COMPUTER SOCIETY
Amazon Ex. 1012.0240
Authorized licensed use limited to: Purdue University. Downloaded on April 26,2023 at 14:07:55 UTC from IEEE Xplore. Restrictions apply.

that allows the system to distinguish some simple occlusion patterns. The geometry of the detectors (called "striplets") is simplified to a rectangular strip. These detectors are assembled into a federation to provide the SPOD (Simple Pattern Occlusion Detectors) service. The SPOD service is internally divided into two separated layers: the image processing layer called the Striplet Engine (SE), in charge of image processing, and the Vision Events Interpretation Layer (VEIL), in charge of input abstraction. Both of these layers were designed to meet the requirements described in section 3.

Striplets are defined as sensitive patches on the interface. Their response is calculated as the integral of the perceived luminance multiplied by a gain function over the surface of the striplet. The gain function has to be chosen so that the integral equals zero when the luminance over the whole striplet is constant. In current implementation the gain function is a symmetric step function with positive value over the central part of the striplet and negative value at both ends of the striplet (Figure 1).

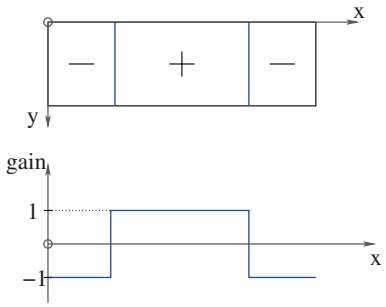


Figure 1. Striplet geometry and the gain function.

Striplets are designed to detect occlusion by elongated objects. Each striplet is 40 millimeters long and 10 millimeters wide on the projected interface. These dimensions are chosen to ensure a maximal response to occlusions made by finger-sized objects occluding the striplet's central area. Occlusions of any extremity of a striplet are intentionally ignored.

The camera coordinates of a striplet are calculated based on its position in the interface as given by the VEIL service and the camera-interface mapping. The event trigger threshold, on the other hand, is estimated individually for each striplet by the SE without any control from the client application. This threshold is automatically set to half of the maximal positive response of a striplet during interaction. The only assumption is that fingers contrast with the interface, which is true in most setups.

The VEIL is the "brain" of the SPOD service. The VEIL *(a)* translates widgets coordinates defined by the client application to a set of striplets coordinates, and *(b)* analyses the occlusion events generated by the SE and issues interaction events when appropriate.

Currently two types of interactive widgets are implemented: touch buttons and sliders. This allows a designer to build simple WIMP-like user interfaces. The button widget is composed of six striplets: two crossed in the center and four other surrounding the button center (Figure 2). Touch events are issued only if occlusion is detected by at least one of the center striplets and no more than one surrounding striplets. Sliders are simply obtained by assembling multiple partially overlapping buttons.

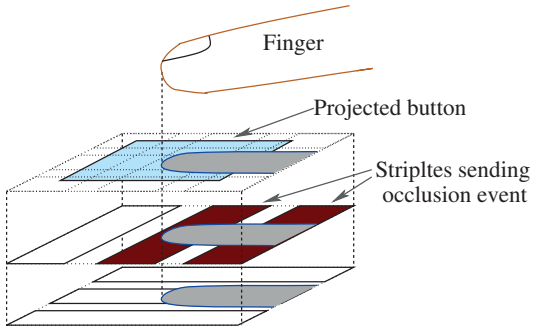


Figure 2. Button widget made of six striplets.

Since monocular vision systems cannot detect when a finger actually touches the interface, interaction events are generated only after detecting a short pause of the finger over a widget. The dwelling period, set to 250 ms, corresponds to the button-down event for mouse-based interaction. To move a slider, the end-user has to first "press" it and then drag it. This is coherent with existing WIMP interface behavior. In contrast, the dragging task requires from the end-user 2D movement coordination. If user's finger exits the SPOD slider area, dragging is stopped.

By assembling striplet detectors in more sophisticated ways, it is also possible to develop different types of interactive widgets, for instance crossing-based menus [1]. An extreme case would cover the whole interface surface with SPOD-button-like structures, thus making a SPOD-based finger tracker.

Service API. The SPOD service requires the client application to specify the position of each interactive widget in a normalized coordinate frame of the interface. Additionally, the SPOD service needs to know the mapping between the camera view and the interface, as well as a rough estimate of the number of pixels per unit length of the interface in the scene. Both the mapping and the scale, are provided by a

Proceedings of the Fourth IEEE International Conference on Computer Vision Systems (ICVS 2006)
0-7695-2506-7/06 $20.00 © 2006 IEEE
Amazon Ex. 1012.0241
Authorized licensed use limited to: Purdue University. Downloaded on April 26,2023 at 14:07:55 UTC from IEEE Xplore. Restrictions apply.

calibration service (discussed below). The communication between SPOD service and calibration service is invisible to the client application. The SPOD service exclusively sends to the client application a stream of interaction events. All communication occurs via TCP/IP connections, using the BIP protocol described in section 3.3.

Inter-layer API. Both the SE and the VEIL are implemented as independent services, running in independent processes. Their communication also is asynchronous and event-driven, using BIP. For a given interface configuration, the VEIL sends to the SE coordinates of all striplets together with the interface-camera mapping matrix. The SE layer sends back to VEIL striplets state-change events that result from end-user interaction with the system.

Contract for the SPOD service. The initial SPOD service implementation can handle up to 300 striplets at camera frame rate (30 Hz) with images of 320x240 pixels size on a 2.8 GHz Pentium IV processor. In terms of widgets, this means the system can handle roughly 50 SPOD buttons simultaneously.

Because actuating buttons is not close coupled interaction, the latency is less of an issue. In fact, the VEIL makes the distinction between accidental occlusions and intentional actions based on a dwell time. On the other hand, the SE service is implemented to minimize latency.

Striplets only provide coarse resolution for finger positions. The resolution can be enhanced by averaging the position of several striplets detecting the same finger. Our implementation of a slider widget achieves a resolution of about 5 millimeters.

The SPOD service is made autonomous (i.e. exept for the UI-camera mapping and scale there are no parameters to set), at the expense of robustness to certain condition changes. In particular, the SE layer would fail to detect occlusion from a finger on a dark background, it would also fail if the image contrast decreases due to a change in camera setting.

While the SPOD service was designed to respect the developer-centric requirements, it does not fully meet user-centric requirements. In particular, the simplistic automatic threshold estimation results in occasional false positive touch detections. However, the SPOD service can be described within the Visual Interface Cues (VICs) framework [13], with local luminance changes as the only visual cue. Using a similar approach for spacio-temporal gesture recognition like in [13], we can hope to alleviate the need of setting an occlusion threshold. Instead, striplets responses would be fed to a neural network based VEIL.

4.2. Support Services

Calibration consists in the geometrical coupling of the camera view (what the vision system percieves) and the displayed interface (what the application developer controls). The calibration service clients need to access the mapping information to transform vision information (e.g. positions in camera coordinates) into application-relevant data (e.g. positions in interface coordinates). This is achieved by providing the associated projective transformation in matrix form. Because the calibration service needs information from both the camera and the application, two approaches are possible: *(a)* If it controls the graphical output, it can work without interaction with the application. This is the case in the PDS example [3], where the interactive surface itself is tracked by the service. *(b)* In the general case, it must negotiate with the client application the display of a calibration grid [7].

Image acquisition service creates an abstraction of the camera. It allows concurrent access to the camera by multiple services. In our case, both the calibration service and the SE require access to the image stream. Low latency video sharing is implemented using shared memory buffers.

5. Application

Using the widget implementation described above, we have implemented a simple calculator application. The calculator interface can be projected and manipulated directly with fingers on the top of a desk. This interface allows a user to perform basic calculations such as addition, subtraction, division and multiplication. Numbers can be either typed on the calculator keyboard or chosen from the history buffer containing results of previous operations. The history buffer can be browsed using a slider on the left side of the calculator.

An informal evaluation of the calculator application, made by volunteers from our laboratory, showed that SPOD-widgets are easy to use and allow fast interface prototyping. A video showing the application working is available in the demo section of http://www-prima.inrialpes.fr/

6. Conclusions

This paper presents a developer oriented design approach for vision-based interactive systems. Inspired by [5], we decompose the vision-based applications to isolated processes of vision components and functional core of the application. The implementation of the vision-components draws on the VICs framework presented by Ye et. al in [13].

Proceedings of the Fourth IEEE International Conference on Computer Vision Systems (ICVS 2006)
0-7695-2506-7/06 $20.00 © 2006 IEEE
Amazon Ex. 1012.0242
Authorized licensed use limited to: Purdue University. Downloaded on April 26,2023 at 14:07:55 UTC from IEEE Xplore. Restrictions apply.

We believe that extending these two design approaches by an HCI-centric requirements analysis allows to build vision systems that can be used for interactive systems designed by developers unfamiliar with vision. Following the guidelines of the developer-centric and end-user-centric requirements analysis we implemented vision-based interactive widgets: buttons and sliders. We illustrate the feasibility of our approach with an implementation of a simple calculator for projection-augmented surfaces.

Acknowledgments

This work has been partially funded by the European project FAME (IST-2000-28323), the FGnet working group (IST-2000-26434), and the RNTL/Proact ContAct project.

References

[1] G. Apitz and F. Guimbretière. Crossy: a crossing-based drawing application. In *UIST '04: Proceedings of the 17th annual ACM symposium on User interface software and technology*, pages 3–12, 2004.

[2] F. Bérard. The magic table: Computer-vision based augmentation of a whiteboard for creative meetings. In *Proceedings of the ICCV Workshop on Projector-Camera Systems*. IEEE Computer Society Press, 2003.

[3] S. Borkowski, J. Letessier, and J. L. Crowley. Spatial control of interactive surfaces in an augmented environment. In R. Bastide, P. A. Palanque, and J. Roth, editors, *EHCI/DSVIS*, volume 3425 of *Lecture Notes in Computer Science*, pages 228–244. Springer, 2004.

[4] J. L. Crowley, J. H. Piater, M. Vincze, and L. Paletta, editors. *Computer Vision Systems, Third International Conference, ICVS 2003, Graz, Austria, April 1-3, 2003, Proceedings*, volume 2626 of *Lecture Notes in Computer Science*. Springer, 2003.

[5] R. Kjeldsen, A. Levas, and C. S. Pinhanez. Dynamically reconfigurable vision-based user interfaces. In Crowley et al. [4], pages 323–332.

[6] S. R. Klemmer, J. Li, J. Lin, and J. A. Landay. Papier-mâché: toolkit support for tangible input. In *Proceedings of the 2004 conference on Human factors in computing systems (CHI'04)*, pages 399–406. ACM Press, 2004.

[7] J. Letessier and F. Bérard. Visual tracking of bare fingers for interactive surfaces. In *UIST '04: Proceedings of the 17th annual ACM symposium on User interface software and technology*, pages 119–122. ACM Press, 2004.

[8] I. S. MacKenzie and C. Ware. Lag as a determinant of human performance in interactive systems. In *Proceedings of the conference on Human factors in computing systems*, pages 488–493. Addison-Wesley Longman Publishing Co., Inc., 1993.

[9] C. Pinhanez, R. Kjeldsen, T. Levas, G. Pingali, M. Podlaseck, and P. Chou. Ubiquitous interactive graphics. In *IBM Research Report RC22495 (W0205-143)*, May 2002.

[10] R. S. Rao, K. Conn, S. H. Jung, J. Katupitiya, T. Kientz, V. Kumar, J. Ostrowski, S. Patel, and C. J. Taylor. Human robot interaction: Application to smart wheelchairs. In *Proceedings of the 2002 IEEE International Conference on Robotics and Automation (ICRA' 02)*, pages 3583–3589, May 2002.

[11] R. Raskar, J. van Baar, P. Beardsley, T. Willwacher, S. Rao, and C. Forlines. iLamps: geometrically aware and self-configuring projectors. *ACM Trans. Graph.*, 22(3):809–818, 2003.

[12] P. Wellner. The digitaldesk calculator: Tactile manipulation on a desk top display. In *ACM Symposium on User Interface Software and Technology*, pages 27–33, 1991.

[13] G. Ye, J. J. Corso, D. Burschka, and G. D. Hager. Vics: A modular vision-based hci framework. In Crowley et al. [4], pages 257–267.

[14] Z. Zhang, Y. Wu, Y. Shan, and S. Shafer. Visual panel: virtual mouse, keyboard and 3d controller with an ordinary piece of paper. In *Proceedings of the 2001 workshop on Percetive user interfaces*, pages 1–8. ACM Press, 2001.

IEEE COMPUTER SOCIETY

Amazon Ex. 1012.0243

Authorized licensed use limited to: Purdue University. Downloaded on April 26,2023 at 14:07:55 UTC from IEEE Xplore. Restrictions apply.

Appendix C-2



Database Processing

Fundamentals, Design, and Implementation

David M. Kroenke and **David J. Auer**

Edition 11

Amazon Ex. 1012.0245

SWAIN HALL LIBRARY

Amazon Ex. 1012.0246

Edition 11

Database Processing

Fundamentals, Design, and Implementation

David M. Kroenke

David J. Auer
Western Washington University

INDIANA UNIVERSITY
LIBRARIES
BLOOMINGTON

Prentice Hall

Boston Columbus Indianapolis New York San Francisco Upper Saddle River
Amsterdam Cape Town Dubai London Madrid Milan Munich Paris Montreal Toronto
Delhi Mexico City Sao Paulo Sydney Hong Kong Seoul Singapore Taipei Tokyo

Amazon Ex. 1012.0247

Executive Editor: Bob Horan
Editorial Director: Sally Yagan
AVP/Editor-in-Chief: Eric Svendsen
Assistant Editor: Kelly Loftus
Director of Marketing: Patrice Lumumba Jones
Senior Marketing Manager: Anne Fahlgren
Marketing Assistant: Susan Osterlitz
Senior Managing Editor: Judy Leale
Production Project Manager: Kelly Warsak
Senior Operations Supervisor: Arnold Vila
Operations Specialist: Ilene Kahn
Senior Art Director: Janet Slowik

Design Director: Christy Mahon
Cover and Interior Design: Karen Quigley
Manager, Rights and Permissions: Charles Morris
Manager, Cover Visual Research & Permissions: Karen Sanatar
Cover Illustration: Stockbyte/Veer/Corbis
Media Director: Lisa Rinaldi
Lead Media Project Manager: Denise Vaughn
Full-Service Project Management: Jennifer Welsch/BookMasters, Inc.
Composition: Integra Software Services
Printer/Binder: Quebecor World-Versailles
Cover Printer: Lehigh-Phoenix Color/Hagerstown
Text Font: 10/12 Kepler MM

Credits and acknowledgments borrowed from other sources and reproduced, with permission, in this textbook appear on appropriate page within text.

Microsoft® and Windows® are registered trademarks of the Microsoft Corporation in the U.S.A. and other countries. Screen shots and icons reprinted with permission from the Microsoft Corporation. This book is not sponsored or endorsed by or affiliated with the Microsoft Corporation.

MySQL®, the MySQL GUI Tools® (MySQL Query Brower® and MySQL Administrator®), the MySQL Command Line Client®, and MySQL Workbench® are registered trademarks of Sun Microsystems, Inc. in the U.S.A and other countries. Screenshots and icons reprinted with permission of Sun Microsystems, Inc. This book is not sponsored or endorsed by or affiliated with Sun Microsystems.

Eclipse® and The Eclipse PHP Development Tools (PDT) Project® are trademarks of the Eclipse Foundation, Inc. The Eclipse platform is copyright Eclipse copyright holders and others, 2000, 2007. Screenshots reprinted under the terms of the Eclipse Public License v1.0 available at www.eclipse.org/legal/epl-v10.html. This book is not sponsored or endorsed by or affiliated with the Eclipse Foundation, Inc.

PHP is copyright The PHP Group 1999–2008, and is used under the terms of the PHP Public License v3.01 available at http://www.php.net/license/3_01.txt. This book is not sponsored or endorsed by or affiliated with The PHP Group.

Oracle is a registered trademark of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners.

Copyright © 2010, 2006, 2004, 2000, 1998 Pearson Education, Inc., publishing as Prentice Hall, One Lake Street, Upper Saddle River, New Jersey 07458. All rights reserved. Manufactured in the United States of America. This publication is protected by Copyright, and permission should be obtained from the publisher prior to any prohibited reproduction, storage in a retrieval system, or transmission in any form or by any means, electronic, mechanical, photocopying, recording, or likewise. To obtain permission(s) to use material from this work, please submit a written request to Pearson Education, Inc., Permissions Department, One Lake Street, Upper Saddle River, New Jersey 07458.

Many of the designations by manufacturers and seller to distinguish their products are claimed as trademarks. Where those designations appear in this book, and the publisher was aware of a trademark claim, the designations have been printed in initial caps or all caps.

Library of Congress Cataloging-in-Publication Data

Kroenke, David.
Database processing: fundamentals, design, and implementation/David M. Kroenke, David J. Auer.—11th ed.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-13-230267-8 (casebound : alk. paper)
1. Database management. I. Auer, David J. II. Title.

QA76.9.D3K7365 2009
005.74—dc22

2009011144

10 9 8 7 6 5 4 3 2 1

Prentice Hall
is an imprint of

PEARSON

www.pearsonhighered.com

ISBN 10: 0-13-230267-5
ISBN 13: 978-0-13-230267-8

Amazon Ex. 1012.0248

Brief Contents

Preface *xvii*

Part 1 • Getting Started **1**

Chapter 1 Introduction 2
Chapter 2 Introduction to Structured Query Language 31

Part 2 • Database Design **93**

Chapter 3 The Relational Model and Normalization 94
Chapter 4 Database Design Using Normalization 127
Chapter 5 Data Modeling and the Entity-Relationship Model 144
Chapter 6 Transforming Data Models in Database Designs 192

Part 3 • Database Implementation **235**

Chapter 7 SQL for Database Construction and Application Processing 236
Chapter 8 Database Redesign 295

Part 4 • Multiuser Database Processing **319**

Chapter 9 Managing Multiuser Databases 320
Chapter 10 Managing Databases with SQL Server 2008 351

ONLINE CHAPTER: SEE PAGE 423 FOR INSTRUCTIONS
Chapter 10A Managing Databases with Oracle 11g 10A-1

ONLINE CHAPTER: SEE PAGE 425 FOR INSTRUCTIONS
Chapter 10B Managing Databases with MySQL 5.1 10B-1

Part 5 • Database Access Standards **427**

Chapter 11 The Web Server Environment 428
Chapter 12 Database Processing with XML 485
Chapter 13 Database Processing for Business Intelligence Systems 525

ONLINE APPENDICES: SEE PAGE 565 FOR INSTRUCTIONS
Appendix A Getting Started with Microsoft Access 2007 A-1
Appendix B The IDEF1X Standard B-1
Appendix C UML-Style Entity-Relationship Diagrams C-1
Appendix D Data Structures for Database Processing D-1
Appendix E The Semantic Object Model E-1

Amazon Ex. 1012.0249

Contents

Preface xvii

PART 1 ● GETTING STARTED 1

Chapter 1: Introduction 2

Chapter Objectives 2
The Characteristics of Databases 3
A Note on Naming Conventions 3
A Database Has Data and Relationships 4
Databases Create Information 5
Database Examples 6
Single-User Database Applications 6
Multiuser Database Applications 6
E-Commerce Database Applications 7
Reporting and Data Mining Database Applications 7
The Components of a Database System 8
Database Applications and SQL 9
The DBMS 11
The Database 12
Personal Versus Enterprise-Class Database Systems 13
What Is Microsoft Access? 13
What Is an Enterprise-Class Database System? 15
Database Design 16
Database Design from Existing Data 16
Database Design for New Systems Development 17
Database Redesign 18
What You Need to Learn 19
A Brief History of Database Processing 20
The Early Years 20
The Emergence and Dominance of the Relational Model 22
Post-Relational Developments 23
Summary 24
Key Terms 25
Review Questions 26
Project Questions 28

Chapter 2: Introduction to Structured Query Language 31

Chapter Objectives 31
Cape Codd Outdoor Sports 32
The Retail Sales Data Extraction 33
RETAIL_ORDER Data 33
ORDER_ITEM Data 34
SKU_DATA Table 35
Data Extracts Are Common 35

Amazon Ex. 1012.0250

SQL Background 35
The SQL SELECT/FROM/WHERE Framework 36
Reading Specified Columns from a Single Table 36
Reading Specified Rows from a Single Table 38
Reading Specified Columns and Rows from a Single Table 39
Submitting SQL Statements to the DBMS 40
Using SQL in Microsoft Access 2007 40
Using SQL in Microsoft SQL Server 2008 46
Using SQL in Oracle Database 11g 48
Using SQL in Sun Microsystems MySQL 5.1 50
SQL Enhancements for Querying a Single Table 52
Sorting the Results 53
SQL WHERE Clause Options 55
Combing the SQL WHERE Clause and the SQL ORDER BY Clause 59
Performing Calculations in SQL Queries 59
Using SQL Built-in Functions 59
SQL Expressions in SQL SELECT Statements 62
Grouping in SQL SELECT Statements 64
Looking for Patterns in NASDAQ Trading 67
Investigating the Characteristics of the Data 68
Searching for Patterns in Trading by Day of Week 69
Querying Two or More Tables with SQL 70
Querying Multiple Tables with Subqueries 70
Querying Multiple Tables with Joins 73
Comparing Subqueries and Joins 76
Summary 77
Key Terms 77
Review Questions 78
Project Questions 82
Marcia's Dry Cleaning 86
Morgan Importing 89

PART 2 ● DATABASE DESIGN **93**

Chapter 3: The Relational Model and Normalization 94

Chapter Objectives 94
Relational Model Terminology 96
Relations 96
Characteristics of Relations 97
Alternative Terminology 98
Functional Dependencies 100
Finding Functional Dependencies 101
Keys 104
Normal Forms 106
Modification Anomalies 106
A Short History of Normal Forms 107
Normalization Categories 108
From First Normal Form to Boyce-Codd Normal Form 108
Eliminating Anomalies from Functional Dependencies 110
Eliminating Anomalies from Multivalued Dependencies 117
Fifth Normal Form 120
Domain/Key Normal Form 120
Summary 121
Key Terms 121
Review Questions 122

Amazon Ex. 1012.0251

Project Questions 124
Marcia's Dry Cleaning 125
Morgan Importing 126

Chapter 4: Database Design Using Normalization 127

Chapter Objectives 127
Assess Table Structure 128
Designing Updatable Databases 129
Advantages and Disadvantages of Normalization 129
Functional Dependencies 129
Normalizing with SQL 130
Choosing Not to Use BCNF 131
Multivalued Dependencies 132
Designing Read-Only Databases 132
Denormalization 132
Customized Duplicated Tables 134
Common Design Problems 135
The Multivalue, Multicolumn Problem 135
Inconsistent Values 136
Missing Values 137
The General-Purpose Remarks Column 138
Summary 139
Key Terms 139
Review Questions 139
Project Questions 141
Marcia's Dry Cleaning 142
Morgan Importing 143

Chapter 5: Data Modeling with the Entity-Relationship Model 144

Chapter Objectives 144
The Purpose of a Data Model 145
The Entity-Relationship Model 145
Entities 145
Attributes 146
Identifiers 146
Relationships 147
Maximum Cardinality 149
Minimum Cardinality 150
Entity-Relationship Diagrams and Their Versions 151
Variations of the E-R Model 151
E-R Diagrams Using the IE Crow's Foot Model 152
Strong Entities and Weak Entities 153
ID-Dependent Entities 153
Non-ID-Dependent Weak Entities 154
Subtype Entities 156
Patterns in Forms, Reports, and E-R Models 157
Strong Entity Patterns 158
ID-Dependent Relationships 162
Mixed Identifying and Nonidentifying Patterns 168
The For-Use-By Pattern 171
Recursive Patterns 172
The Data Modeling Process 174
The College Report 175
The Department Report 175
The Department/Major Report 176
The Student Acceptance Letter 178

Amazon Ex. 1012.0252

Summary 180
Key Terms 181
Review Questions 182
Project Questions 184
Marcia's Dry Cleaning 191
Morgan Importing 191

Chapter 6: Transforming Data Models into Database Designs 192

Chapter Objectives 192
Create a Table for Each Entity 193
Selecting the Primary Key 194
Specifying Candidate (Alternate) Keys 195
Specify Column Properties 195
Verify Normalization 197
Create Relationships 198
Relationships Between Strong Entities 198
Relationships Using ID-Dependent Entities 201
Relationships with a Weak Non-ID-Dependent Entity 206
Relationships in Mixed Entity Designs 206
Relationships Between Supertype and Subtype Entities 208
Recursive Relationships 208
Representing Ternary and Higher-Order Relationships 210
Relational Representation of the Highline University Data Model 213
Design for Minimum Cardinality 214
Actions When the Parent Is Required 216
Actions When the Child Is Required 217
Implementing Actions for M-O Relationships 218
Implementing Actions for O-M Relationships 219
Implementing Actions for M-M Relationships 219
Designing Special Case M-M Relationships 219
Documenting the Minimum Cardinality Design 220
An Additional Complication 222
Summary of Minimum Cardinality Design 222
The View Ridge Gallery Database 222
Summary of Requirements 222
The View Ridge Data Model 223
Database Design with Data Keys 224
Minimum Cardinality Enforcement for Required Parents 225
Minimum Cardinality Enforcement for the Required Child 227
Column Properties for the View Ridge Database Design Tables 228
Summary 230
Key Terms 231
Review Questions 231
Project Questions 233
Marcia's Dry Cleaning 234
Morgan Importing 234

PART 3 ◎ DATABASE IMPLEMENTATION 235

Chapter 7: SQL for Database Construction and Application Processing 236

Chapter Objectives 236
The View Ridge Gallery Database 237
SQL DDL, DML, and a New Type of Join 237

Amazon Ex. 1012.0253

Managing Table Structure with SQL DDL 238
Creating the View Ridge Database 238
Using the SQL CREATE TABLE Statement 239
Variations in SQL Data Types 239
Creating the ARTIST Table 240
Creating the WORK Table and the 1:N ARTIST-to-WORK Relationship 244
Implementing Required Parent Rows 245
Implementing 1:1 Relationships 245
Casual Relationships 245
Creating Default Values and Data Constraints with SQL 246
Creating the View Ridge Database Tables 248
The SQL ALTER Statement 251
The SQL DROP TABLE Statement 252
SQL DML Statements 252
The SQL INSERT Statement 252
Populating the View Ridge Database Tables 253
The SQL UPDATE Statement 259
The SQL DELETE Statement 260
New Forms of Join 260
The SQL JOIN ON Syntax 260
Outer Joins 262
Using SQL Views 266
Using SQL Views to Hide Columns and Rows 268
Using SQL Views to Display Results of Computed Columns 269
Using SQL Views to Hide Complicated SQL Syntax 270
Layering Built-in Functions 270
Using SQL Views for Isolation, Multiple Permissions, and Multiple Triggers 273
Updating SQL Views 273
Embedding SQL in Program Code 274
Using SQL Triggers 275
Using Triggers to Provide Default Values 276
Using Triggers to Enforce Data Constraints 278
Using Triggers to Update Views 279
Using Triggers to Implement Referential Integrity Actions 279
Using Stored Procedures 281
Advantages of Stored Procedures 281
The WORK_AddWorkTransaction Stored Procedure 283
Summary 283
Key Terms 285
Review Questions 286
Project Questions 290
Marcia's Dry Cleaning 293
Morgan Importing 294

Chapter 8: Database Redesign 295

Chapter Objectives 295
The Need for Database Redesign 296
SQL Statements for Checking Functional Dependencies 296
Correlated Subqueries 297
Analyzing the Existing Database 302
Reverse Engineering 302
Dependency Graphs 303
Database Backup and Test Databases 304
Changing Table Names and Table Columns 305
Changing Table Names 305
Adding and Dropping Columns 306

Amazon Ex. 1012.0254

Changing a Column Data Type or Column Constraints 307
Adding and Dropping Constraints 308
Changing Relationship Cardinalities and Properties 308
Changing Minimum Cardinalities 308
Changing Maximum Cardinalities 309
Adding and Deleting Tables and Relationships 312
Forward Engineering(?) 312
Summary 313
Key Terms 314
Review Questions 314
Project Questions 316
Marcia's Dry Cleaning 317
Morgan Importing 318

PART 4 • MULTIUSER DATABASE PROCESSING 319

Chapter 9: Managing Multiuser Databases 320

Chapter Objectives 320
Database Administration 321
Managing the Database Structure 322
Concurrency Control 323
The Need for Atomic Transactions 324
Resource Locking 328
Optimistic Versus Pessimistic Locking 330
Declaring Lock Characteristics 331
Consistent Transactions 332
Transaction Isolation Level 333
Cursor Type 334
Database Security 335
Processing Rights and Responsibilities 335
DBMS Security 337
DBMS Security Guidelines 337
Application Security 339
SQL Injection Attack 340
Database Backup and Recovery 340
Recovery via Reprocessing 341
Recovery via Rollback/Rollforward 341
Managing the DMBS 343
Maintaining the Data Repository 344
Summary 345
Key Terms 346
Review Questions 347
Project Questions 348
Marcia's Dry Cleaning 349
Morgan Importing 350

Chapter 10: Managing Databases with SQL Server 2008 351

Chapter Objectives 351
Installing SQL Server 2008 352
The Microsoft SQL Server 2008 Management Studio 353
Creating an SQL Server 2008 Database 354
SQL Server 2008 Utilities 356
SQL CMD and Microsoft PowerShell 356
Microsoft SQL CLR 356

Amazon Ex. 1012.0255

SQL Server 2008 GUI Displays 357
SQL Server 2008 SQL Statements and SQL Scripts 357
Creating and Populating the View Ridge Database Tables 360
Creating the View Ridge Database Table Structure 360
Reviewing Database Structures in the SQL Server GUI Display 364
Indexes 368
Populating the VRG Tables with Data 369
Creating Views 379
SQL Server 2008 Application Logic 381
Transact-SQL 382
Stored Procedures 385
Triggers 393
Concurrency Control 407
Transaction Isolation Level 408
Cursor Concurrency 408
Locking Hints 409
SQL Server 2008 Security 410
SQL Server Database Security Settings 412
SQL Server 2008 Backup and Recovery 414
Backing Up a Database 414
SQL Server Recovery Models 415
Restoring a Database 416
Database Maintenance Plans 416
Topics Not Discussed in This Chapter 417
Summary 417
Key Terms 417
Review Questions 418
Project Questions 420
Marcia's Dry Cleaning 421
Morgan Importing 421

ONLINE CHAPTER: SEE PAGE 423 FOR INSTRUCTIONS

Chapter 10A: Managing Databases with Oracle Database 11g 10A-1

Chapter Objectives 10A-1
Installing Oracle Database 11g 10A-2
Oracle Database 11g Documentation 10A-4
Oracle and Java 10A-5
Oracle Database 11g Administration and Development Tools 10A-5
The Oracle Database Configuration Assistant 10A-6
The Oracle Enterprise Manager 11g Database Control 10A-6
Oracle Tablespaces 10A-8
Oracle Security 10A-11
User Privileges 10A-11
Creating a User Account 10A-11
Creating a Role 10A-13
Oracle Application Development Tools 10A-15
Oracle SQL*Plus 10A-15
Oracle SQL Developer 10A-16
Oracle Schemas 10A-17
Oracle Database 11g 2008 SQL Statements and SQL Scripts 10A-18
Creating and Populating the View Ridge Database Tables 10A-20
Creating the View Ridge Database Table Structure 10A-20
Reviewing Database Structures in the SQL Developer GUI Display 10A-23

Amazon Ex. 1012.0256

Indexes 10A-26
Populating the VRG Tables 10A-28
Creating Views 10A-36
Application Logic 10A-39
Oracle PL/SQL 10A-40
Stored Procedures 10A-42
Triggers 10A-50
Concurrency Control 10A-61
Read Committed Transaction Isolation Level 10A-62
Serializable Transaction Isolation Level 10A-63
Read-Only Transaction Isolation 10A-63
Additional Locking Comments 10A-63
Oracle Backup and Recovery 10A-64
Oracle Recovery Facilities 10A-64
Types of Failure 10A-64
Topics Not Discussed in This Chapter 10A-65
Summary 10A-65
Key Terms 10A-66
Review Questions 10A-67
Project Questions 10A-69
Marcia's Dry Cleaning 10A-70
Morgan Importing 10A-70

ONLINE CHAPTER: SEE PAGE 425 FOR INSTRUCTIONS

Chapter 10B: Managing Databases with MySQL 5.1 10B-1

Chapter Objectives 10B-1
The MySQL 5.1 DBMS 10B-2
Installing and Updating MySQL 10B-2
Configuring MySQL 10B-2
MySQL Storage Engines 10B-3
The MySQL GUI Utilities 10B-4
Creating and Using a MySQL Database 10B-6
MySQL Utilities 10B-7
MySQL Command Line Client 10B-8
MySQL GUI Displays 10B-8
MySQL SQL Statements and SQL Scripts 10B-10
Creating and Populating the View Ridge Database Tables 10B-13
Creating the View Ridge Database Table Structure 10B-13
Reviewing Database Structures in the MySQL GUI Display 10B-16
Indexes 10B-17
Populating the VRG Tables with Data 10B-22
Creating Views 10B-30
MySQL Application Logic 10B-33
MySQL Procedural Statements 10B-34
Stored Procedures 10B-36
Triggers 10B-43
A Last Word on MySQL Stored Procedures and Triggers 10B-53
Concurrency Control 10B-54
MySQL 5.1 Security 10B-55
MySQL Database Security Settings 10B-56
MySQL 5.1 Server Backup and Recovery 10B-58
Backing Up a MySQL Database 10B-58
Restoring a MySQL Database 10B-59
Topics Not Discussed in This Chapter 10B-61

Amazon Ex. 1012.0257

Summary 10B-61
Key Terms 10B-62
Review Questions 10B-62
Project Questions 10B-63
Marcia's Dry Cleaning 10B-65
Morgan Importing 10B-65

PART 5 ● DATABASE ACCESS STANDARDS 427

Chapter 11: The Web Server Environment 428

Chapter Objectives 428
The Web Database Processing Environment 429
The ODBC Standard 431
ODBC Architecture 431
Conformance Levels 432
Creating an ODBC Data Source Name 434
The Mircosoft .NET Framework and ADO.NET 437
OLE DB 438
ADO and ADO.NET 442
The ADO.NET Object Model 442
The JAVA Platform 446
JDBC 446
Java Server Pages (JSP) and Servlets 448
Apache Tomcat 449
Web Database Processing with PHP 449
Web Database Processing with PHP and Eclipse 451
Getting Started with HTML Web Pages 453
The index.html Web Page 453
Creating the index.html Web Page 453
Using PHP 456
Challenges for Web Database Processing 463
Web Page Examples with PHP 463
Example 1: Updating a Table 464
Example 2: Using PHP Data Objects (PDO) 470
Example 3: Invoking a Stored Procedure 470
Summmary 478
Key Terms 480
Review Questions 481
Project Questions 483
Marcia's Dry Cleaning 485
Morgan Importing 485

Chapter 12: Database Processing with XML 486

Chapter Objectives 486
The Importance of XML 487
XML as a Markup Language 488
XML Document Type Declarations 489
Materializing XML Documents with XSLT 489
XML Schema 494
XML Schema Validation 494
Elements and Attributes 495
Flat Versus Structured Schemas 496
Global Elements 500

Amazon Ex. 1012.0258

Creating XML Documents from Database Data 502
Using the SQL SELECT . . . FOR XML Statement 502
Multitable SELECT with FOR XML 506
An XML Schema for All CUSTOMER Purchases 510
A Schema with Two Multivalue Paths 514
Why Is XML Important? 514
Additional XML Standards 520
Summary 522
Key Terms 523
Review Questions 523
Project Questions 524
Marcia's Dry Cleaning 525
Morgan Importing 525

Chapter 13: Database Processing for Business Intelligence Systems 526

Chapter Objectives 526
Business Intelligence Systems 526
The Relationship Between Operational and BI Systems 526
Reporting Systems and Data Mining Applications 527
Reporting Systems 528
Data Mining Applications 528
Data Warehouses and Data Marts 528
Components of a Data Warehouse 528
Data Warehouses Versus Data Marts 531
Dimensional Databases 532
Reporting Systems 539
RFM Analysis 540
Producing the RFM Report 541
Reporting System Components 544
Report Types 545
Report Media 545
Report Modes 546
Report System Functions 546
OLAP 549
Data Mining 552
Unsupervised Data Mining 554
Supervised Data Mining 555
Three Popular Data Mining Techniques 555
Market Basket Analysis 555
Using SQL for Market Basket Analysis 555
Summary 557
Key Terms 558
Review Questions 559
Project Questions 561
Marcia's Dry Cleaning 563
Morgan Importing 564

APPENDICES

ONLINE APPENDICES: SEE PAGE 565 FOR INSTRUCTIONS

Appendix A: Getting Started with Microsoft Access 2007 A-1
Chapter Objectives A-1
Table Keys A-2
Relationships Among Tables A-2

Amazon Ex. 1012.0259

Creating an Access Database A-3
The Microsoft Office Fluent User Interface A-5
The Ribbon and Command Tabs A-5
Contextual Command Tabs A-5
Database Objects and the Navigation Pane A-6
Closing a Database and Exiting Microsoft Access A-7
Opening an Existing Access Database A-8
Creating Access Database Tables A-9
Inserting Data into Tables—The Datasheet View A-17
Modifying and Deleting Data in Tables in the Datasheet View A-19
Creating Relationships Between Tables A-20
Working with Microsoft Access Queries A-23
Microsoft Access Forms and Reports A-26
Creating Access 2007 Forms A-26
Creating Access 2007 Reports A-27
Closing a Database and Exiting Access 2007 A-27
Key Terms A-28
Review Questions A-28

Appendix B: The IDEF1X Standard B-1
Chapter Objectives B-1
IDEF1X Entities B-1
IDEF1X Relationships B-2
Nonidentifying Connection Relationships B-2
Identifying Connection Relationships B-3
Nonspecific Relationships B-4
Categorization Relationships B-4
Domains B-6
Domains Reduce Ambiguity B-6
Domains Are Useful B-7
Base Domains and Typed Domains B-7
Key Terms B-8
Review Questions B-8

Appendix C: UML-Style Entity-Relationship Diagrams C-1
Chapter Objectives C-1
UML Entities and Relationships C-1
Representation of Weak Entities C-2
Representation of Subtypes C-2
OOP Constructs Introduced by UML C-3
The Role of UML in Database Processing Today C-4
Key Terms C-5
Review Questions C-5

Appendix D: Data Structures for Database Processing D-1
Chapter Objectives D-1
Flat Files D-1
Processing Flat Files in Multiple Orders D-2
A Note on Record Addressing D-2
Maintaining Order with Linked Lists D-3
Maintaining Order with Indexes D-5
B-Trees D-5
Summary of Data Structures D-7
Representing Binary Relationships D-8
Review of Record Relationships D-8
Representing Trees D-10

Amazon Ex. 1012.0260

Representing Simple Networks D-12
Representing Complex Networks D-13
Summary of Relationship Representations D-14
Secondary-Key Representations D-16
Linked-List Representation of Secondary Keys D-17
Index Representation of Secondary Keys D-18
Key Terms D-18
Review Questions D-19

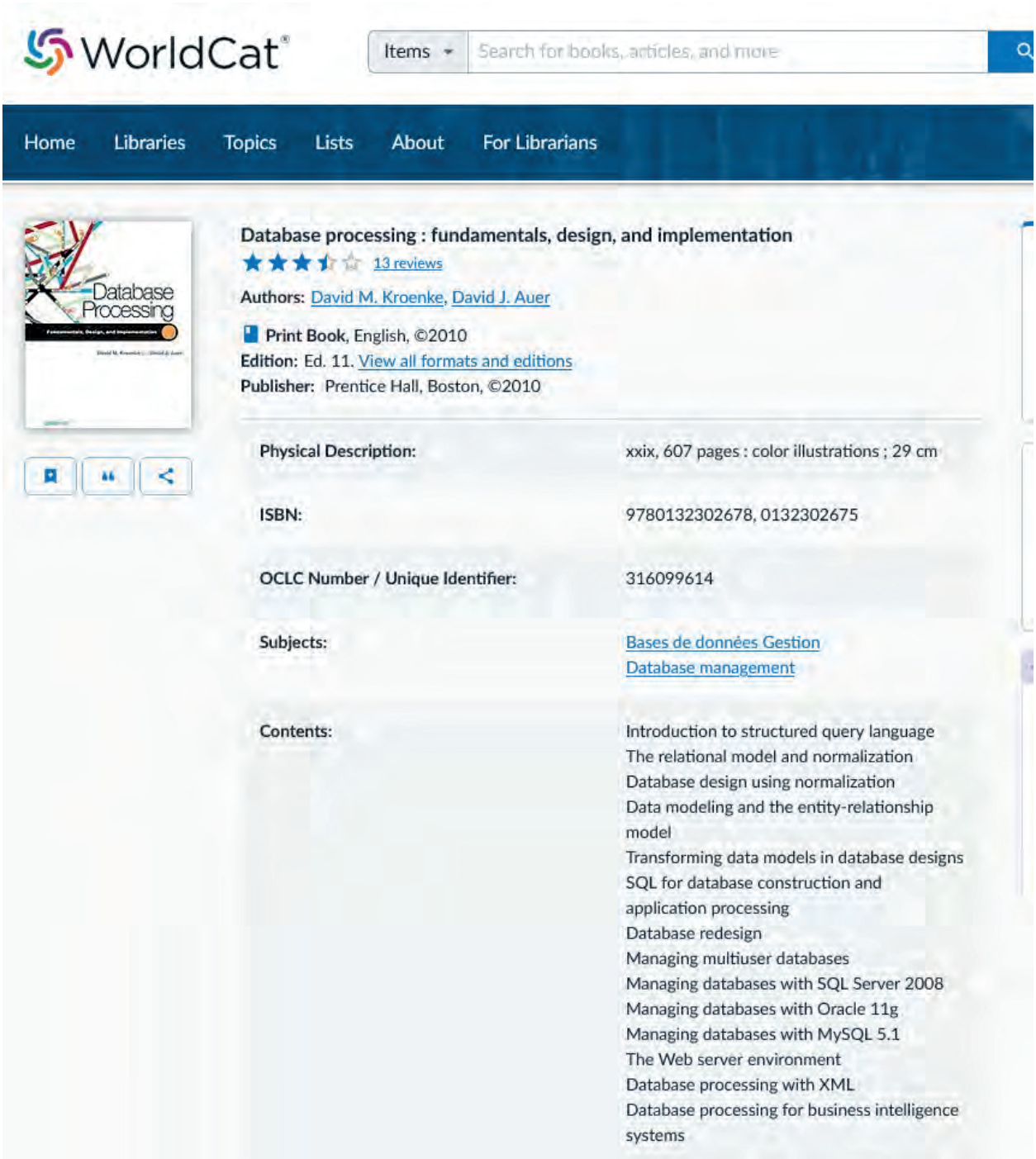
Appendix E: The Semantic Object Model E-1
Chapter Objectives E-1
Semantic Objects E-2
Defining Semantic Objects E-2
Attributes E-3
Object Identifiers E-5
Attribute Domains E-6
Semantic Object Views E-6
Types of Objects E-7
Simple Objects E-7
Composite Objects E-8
Compound Objects E-11
Representing One-to-One Compound Objects E-12
Representing One-to-Many and Many-to-One Relationships E-14
Representing Many-to-Many Relationships E-14
Hybrid Objects E-17
Representing Hybrid Relationships E-20
Association Objects E-22
Parent/Subtype Objects E-24
Archetype/Version Objects E-26
Comparing the Semantic Object and the E-R Models E-29
Key Terms E-31
Review Questions E-32

Bibliography 567
Glossary 569
Index 591

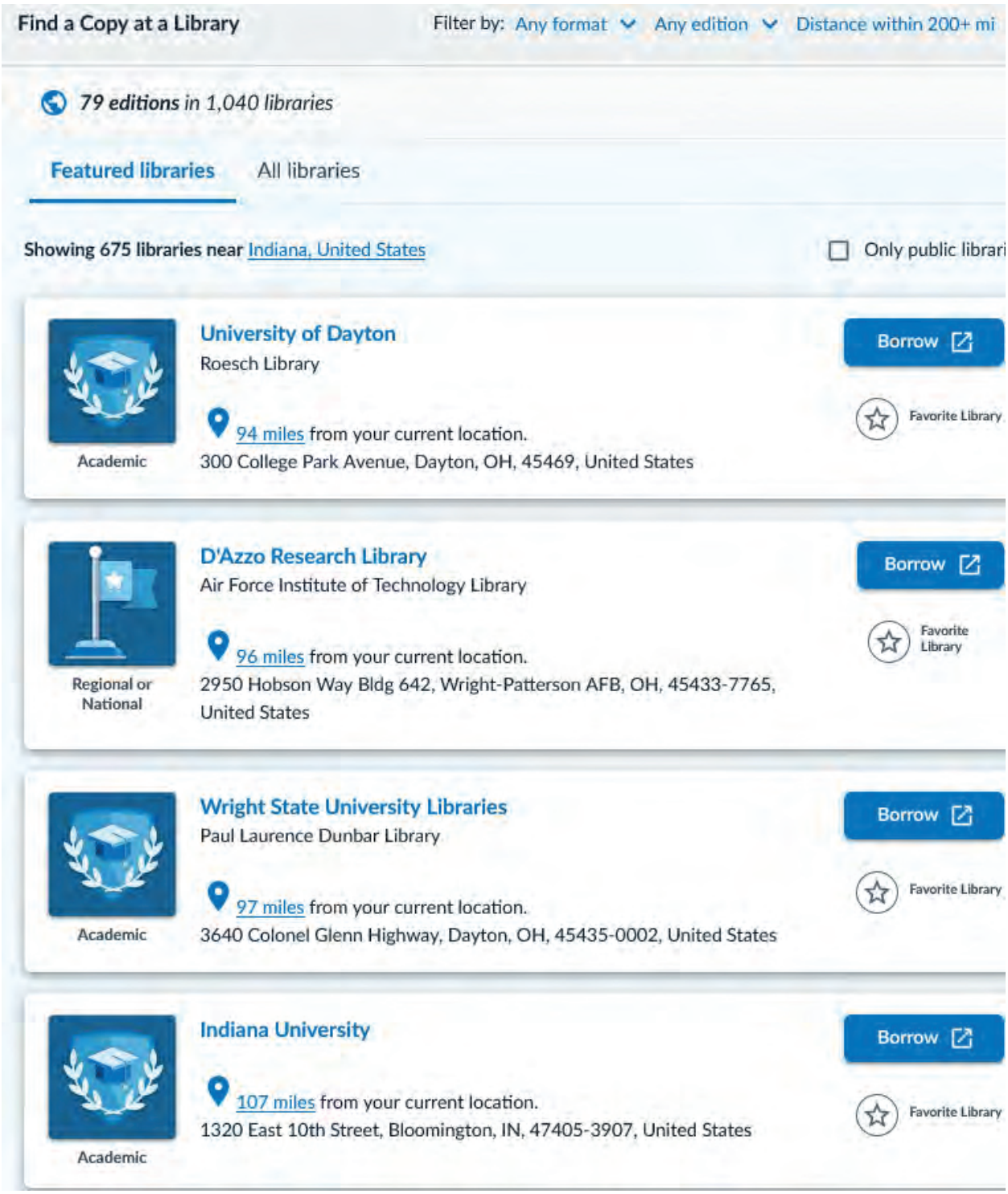
Amazon Ex. 1012.0261

Appendix C-3

Screen capture completed by J L Mullins from *WorldCat* on October 26, 2023 at URL:
https://www.worldcat.org/title/316099614?oclcNum=316099614



Screen capture completed by J L Mullins from *WorldCat* on October 26, 2023 at URL:
https://www.worldcat.org/title/316099614?oclcNum=316099614



Appendix C-4

All Fields

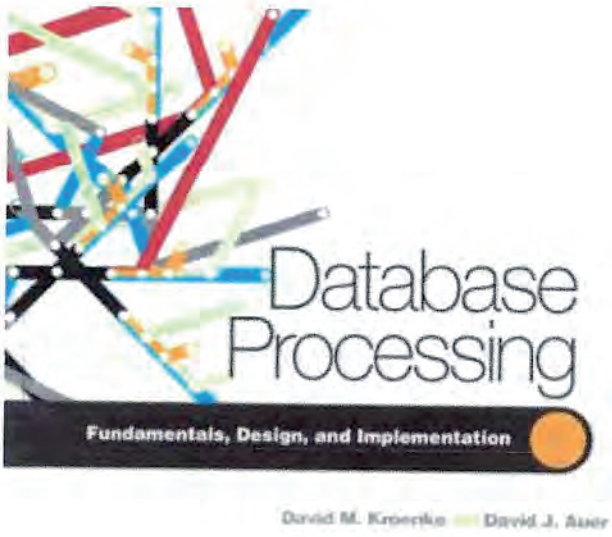
Online Only

Due to unforeseen courier issues, we anticipate delays in fulfilling Request Delivery requests from other campuses through October. In the meantime, you can still use Interlibrary Loan to request materials. Please contact your campus library with any questions.

Search Results

Back to Search Start Over

View Tools



Edition 11

Amazon Ex. 1012.0266

Database processing : fundamentals, design, and implementation

(8370410)

Author

Kroenke, David M., 1948-

Title

Database processing : fundamentals, design, and implementation / David M. Kroenke, David J. Auer.

Format

Book

Edition

Ed. 11.

Published

Boston : Prentice Hall, c2010.

Description

xxix, 607 p. : col. ill. ; 29 cm.

Other contributors

Auer, David J.

Notes

Includes bibliographical references and index.

Subject headings

Database management.

ISBN

9780132302678 (casebound : alk. paper)
0132302675 (casebound : alk. paper)

Holdings

Library

BImgtn - Sciences Library

Call Number

QA76.9.D3 K7365 2010

Location

Stacks

Amazon Ex. 1012.0267

Browse by call number

Database concepts
Kroenke, David M., 1948-
QA76.9.D3 K736 2008

Database processing : fundamentals, design & implementation
Kroenke, David M., 1948-
QA76.9.D3 K7365 2004

Database processing : fundamentals, design, and implementation
Kroenke, David M., 1948-
QA76.9.D3 K7365 2010

Database processing : fundamentals, design, & implementation
Kroenke, David M., 1948-
QA76.9.D3 K737 2000 Text

Database processing : fundamentals, design, implementation
Kroenke, David M., 1948-
QA76.9.D3 K76 1983

‹ Previous

Next ›

Amazon Ex. 1012.0268

Class Reserves
My Account
Feedback
IUCAT Help
Privacy Policy
IU Home

Powered by Blacklight

Amazon Ex. 1012.0269

Appendix C-5

Screen capture by JL Mullins from Indiana University Libraries Online Catalog on October 23, 2023 at: https://iucat.iu.edu/catalog/8370410

Librarian View

LEADER 00000cam a22000004a 45 0
001 8370410
003 OCoLC
005 20090906000000.0
008 090313s2010 maua b 001 0 eng||
010 a| 2009011144
035 a| (Sirsi)8370410
040 a| DLC c| DLC d| C#P d| BWX
020 a| 9780132302678 (casebound : alk. paper)
020 a| 0132302675 (casebound : alk. paper)
035 a| (OCoLC)ocn316099614
050 0 0 a| QA76.9.D3 b| K7365 2010
082 0 0 a| 005.74 2| 22
082 4 a| 004
099 a| QA76.9.D3 b| K7365 2010
049 a| IUN4
100 1 a| Kroenke, David M., d| 1948- ?| UNAUTHORIZED
245 1 0 a| Database processing : b| fundamentals, design, and implementation / c| David M. Kroenke, David J. Auer.
250 a| Ed. 11.
260 a| Boston : b| Prentice Hall, c| c2010.
300 a| xxix, 607 p. : b| col. ill. ; c| 29 cm.
504 a| Includes bibliographical references and index.
650 0 a| Database management. =| ^A1116188
700 1 a| Auer, David J. =| ^A2231874
596 a| 7
999 a| QA76.9.D3 K7365 2010 i| 30000124581368 j| QA 000076.9 .D3 K7365 2010 l| STACKS m| B-CHEM n| STACKS w| LC z| CAT:09-15-2009:AP537

Appendix C-6

Screen capture completed by JL Mullins from the Copyright Public Records System on October 26, 2023 at URL: https://publicrecords.copyright.gov/detailed-record/23325118

Copyright | Public Records System Home Recent Records Recent Searches Name D

Home / Search / Detailed Record View

Title | entals, design and implemer | Search Advanced Search

Detailed Record View

Registration record TX0007071581

Copyright Catalog

Search Results: Displaying 7 of 1446 entries

Previous Next

Database Processing: Fundamentals, Design, and Implementation, 11/E.

Registration Number / Date	TX0007071581 / 2009-10-28
Previous Registration	2005, TX6-221-618.
Type of Work	Text
Title	Database Processing: Fundamentals, Design, and Implementation, 11/E.
Application Title	Database Processing: Fundamentals, Design, and Implementation, 11/E.
Date of Creation	2008
Date of Publication	2009-07-28
Copyright Claimant	Pearson Education, Inc., Transfer: By written agreement. Address: One Lake Street, U 07458, United States.
Authorship on Application	David M. Kroenke; Domicile: United States; Citizenship: United States. Authorship: text.
	David J. Auer; Domicile: United States; Citizenship: United States. Authorship: text.
Basis of Claim	Revised and updated.
Rights and Permissions	Pearson Education, Inc., Pearson Education, Inc., One Lake Street, Upper Saddle River, NJ
Pre-existing Material	Previously copyrighted material.
ISBN	0-13-230267-5
Description	Book, 607 p.
Nation of First Publication	United States
Names	Kroenke, David M.
	Auer, David J.
	Pearson Education, Inc.

Appendix D

JAMES L. MULLINS, PhD

Prior Art Documentation Librarian Services, LLC

205 Saint Cuthbert, Williamsburg, VA 23188

jlmullins@priorartdoclib.com

ph. 765 479 4956

Prior Art Documentation Librarian Services, LLC. (PADLS). Founded January 2018. As of July, 2023, 1 have completed 123 declarations for 39 law firms; five depositions were scheduled, three cancelled a few days prior, one was held in February, 2019 and a second was held in June 2023.

Library Experience:

2018-present Dean Emeritus, Libraries & Norton Professor Emeritus, Purdue University

2011 - 2017 Dean of Libraries & Esther Ellis Norton Professor, Purdue University

2004 - 2011 Dean of Libraries & Professor, Purdue University, West Lafayette, IN

2000-2004 Assistant/Associate Director for Administration, MIT Libraries, Massachusetts Institute of Technology, Cambridge, MA.

1996-2000 University Librarian & Director, Falvey Memorial Library. Villanova University, Villanova, PA.

1978-1996 Director of Library Services, Indiana University South Bend.

1974-1978 Associate Librarian, Indiana University Bloomington, School of Law.

1973-1974 Instructor/Catalog Librarian. Georgia Southern College (now University).

Teaching Experience:

1977-1996 Associate Professor (part-time), School of Library and Information Science, Indiana University. Subjects taught: Cataloging, Management, and Academic Librarianship.

Education:

1972 - University of Iowa. Honors Bachelor of Arts, History, Religion & Political Science.

1973 - University of Iowa. Master of Arts in Library Science.

1984 - Indiana University. Doctor of Philosophy. Concentration: Academic Library Administration. Emphasis: Law Librarianship.

Awards and Recognition:

2017 Wilmeth Active Learning Center/Library of Engineering and Science, Grand Reading Room, was announced by President Mitch Daniels, Purdue University, that it would be re-named the James L. Mullins Reading Room to honor his leadership and reputation in the academic library profession. September 2017. Portrait unveiled December 2017.

2017 Distinguished Alumnus Award by the School of Informatics and Computing, Indiana University, Bloomington. Given June 25, 2017.

2016 Hugh C. Atkinson Memorial Award, jointly sponsored by the four divisions of the American Library Association (ALA), June 27, 2016.

2015 ACRL Excellence in University Libraries Award, April 23, 2015.

Named Esther Ellis Norton Professor of Library Science by Purdue Trustees, December 11, 2011.

International Review Panel to evaluate the University of Pretoria Library, February 20 – 24, 2011. Pretoria, South Africa.

Publications: (selected)

Designing the Academic Library, Chapter Contributor. Association of College and Research Libraries, American Library Association, 2022.

A Purdue Icon: Creation, Life, and Legacy, edited by James L. Mullins, Founder's Series, Purdue University Press, 138pp., August 2017.

"The policy and institutional framework." In *Research Data Management, Practical Strategies for Information Professionals,* edited by Ray, J M. Purdue University Press, pp.25-44, 2014.

"DataCite: linking research to data sets and content." In Benson, P and Silver, S. *What Editors Want: An Author's Guide to Scientific Journal Publishing*. University of Chicago Press, pp. 21-23, December 2012.

"Library Publishing Services: Strategies for Success," with R. Crow, O. Ivins, A. Mower, C. Murray-Rust, J. Ogburn, D Nesdill, M. Newton, J. Speer, C. Watkinson. *Scholarly Publishing and Academic Resources Coalition (SPARC),* version 2.0, March 2012.

"The Changing Definition and Role of Collections and Services in the University Research Library." *Indiana Libraries,* Vol 31, Number 1 (2012), pp.18-24.

"Are MLS Graduates Being Prepared for the Changing and Emerging Roles that Librarians must now assume within Research Libraries?" *Journal of Library Administration.* Volume 52, Issue 1, 2012, p. 124-132

Baykoucheva, Svetla. What Do Libraries Have to Do with e-Science?: An Interview with James L. Mullins, Dean of Purdue University Libraries. Chem. Inf. Bull. [Online] 2011, 63 (1), 45-49. http://www.acscinf.org/publications/bulletin/63-1/mullins.php (accessed Mar 16, 2011).

"The Challenges of e-Science Data-set Management and Scholarly Communication for Domain Sciences and Technology: a Role for Academic Libraries and Librarians," chapter in, *The Digital Deluge: Can Libraries Cope with e-Science?"* Deanna B. Marcum and Gerald George, editors, Libraries Unlimited/Teacher Ideas Press, 2009. (a monograph publication of the combined proceedings of the KIT/CLIR proceedings).

"Bringing Librarianship to e-Science," *College and Research Libraries.* vol. 70, no. 3, May 2009, editorial.

"The Librarian's Role in e-Science" *Joho Kanri (Journal on Information Processing and Management),* Japan Science and Technology Agency (formerly Japan Information Center of Science and Technology), Tokyo, Japan. Translated into Japanese by Taeko Kato. March 2008.

The Challenge of e-Science Data-set Management to Domain Sciences and Engineering: a Role for Academic Libraries and Librarians," KIT (Kanazawa Institute of Technology)/CLIR (Council of Library and Information Resources) International Roundtable for Library and Information Science, July 5-6, 2007. Developments in e-science status quo and the challenge, The Japan Foundation, 2007.

"An Administrative Perspective," Chapter 14, *Proven Strategies for Building an Information Literacy Program*, Susan Curzon and Lynn Lampert, editors, Neal-Schuman Publishers, Inc., New York, 2007. pp. 229-237.

Library Management and Marketing in a Multicultural World, proceedings of the IFLA Management and Marketing (M&M) Section, Shanghai, China, August 16-17, 2006, edited. K.G. Saur, Munchen, Germany, June 2007. 390 pp.

Top Ten Assumptions for the Future of Academic Libraries and Librarians: a report from the ACRL Research Committee, with Frank R. Allen and Jon R. Hufford. *College & Research Libraries,* April 2007, vol.68, no.4. pp.240-241, 246.

To Stand the Test of Time: Long-term Stewardship of Digital Data Sets in Science and Engineering. A report to the National Science Foundation from the ARL Workshop on New Collaborative Relationships: the Role of Academic Libraries in the Digital Data Universe. September 26-27, 2006, Arlington, VA. p.141. http://www.arl.org/bm~doc/digdatarpt.pdf

"Enabling Interaction and Quality in a Distributed Data DRIS," *Enabling Interaction and Quality: Beyond the Hanseatic League*. 8th International Conference on Current Research Information Systems, with D. Scott Brandt and Michael Witt. Promoted by euro CRIS. Leuven University Press, 2006. pp.55-62. Editors: Anne Garns Steine Asserson and Eduard J. Simons.

"Standards for College Libraries, the final version approved January 2000," prepared by the ACRL College Libraries Standards Committee (member), *C&RL News*, March 2000, p.175-182.

"Standards for College Libraries: a draft," prepared by the ACRL College Libraries Section, Standards Committee (member), *C&RL News*, May 1999, p. 375-381.

"Statistical Measures of Usage of Web-based Resources," *The Serials Librarian,* vol. 36, no. 1-2 (1999) p. 207-10.

"An Opportunity: Cooperation between the Library and Computer Services," in *Building Partnerships: Computing and Library Professionals.* Edited by Anne G. Lipow and Sheila D. Creth. Berkeley and San Carlos, CA, Library Solutions Press, 1995. p. 69-70.

"Faculty Status of Librarians: A Comparative Study of Two Universities in the United Kingdom and How They Compare to the Association of College and Research Libraries Standards, " in *Academic Librarianship, Past, Present, and Future: a Festschrift in Honor of David Kaser.* Englewood, Colorado; Libraries Unlimited, 1989. p. 67-78. Review in: *College & Research Libraries,* vol. 51, no. 6. November 1990, p. 573-574.

Presentations: (Representative)

"How Long the Odyssey? Transitioning the Library and Librarians to Meet the Needs and Expectations of the 21st Century University," David Kaser Lecture, School of Informatics & Computing, Indiana University, Bloomington, IN, November 16, 2015.

Presentation at University of Cape Town, Cape Town, South Africa, August 20, 2015.

"The Challenge of Discovering Science and Technology Information," Moderator, International Federation of Library Associations (IFLA) Science and Technological Libraries Section Program, Cape Town, South Africa, August 18, 2015.

"An Odyssey in Data Management: Purdue University," International Federation of Library Associations (IFLA) Research Data Management: Finding Our Role – A program of the Research Data Alliance, Cape Town, South Africa, August 17, 2015.

Presentation at University of Pretoria, Pretoria, South Africa, August 11, 2015.

Co-Convener with Sarah Thomas, Harvard University, at the Harvard Purdue Symposium on Data Management, Harvard University, Cambridge, MA, June 15-18, 2015.

"Strategic Communication," panel discussion on the Director's role and perspective on library communications at Committee on Institutional Cooperation (CIC) Center for Library Initiatives (CLI) Annual Conference, University of Illinois Urbana-Champaign, May 20, 2015.

"Issues in Data Management," panel discussion moderated by Catherine Woteki, United States Undersecretary for Research, Education & Economics at 20th Agriculture Network Information Collaborative (AgNIC) Annual Meeting in the National Agricultural Library, Beltsville, MD, May 6, 2015.

"Active learning/IMPACT & the Active Learning Center at Purdue University," Florida Institute of Technology, Melbourne, FL, February 11, 2015.

"Science+art=creativity: libraries and the new collaborative thinking," panel moderator, International Federation of Library Associations (IFLA) 80th General Conference and Assembly, Lyon, France, August 19, 2014.

"Purdue University The Active Learning Center—A new concept for a library," Association of University Architects 59th Annual National Conference, University of Notre Dame, South Bend, IN, June 23, 2014.

"Big Data & Implications for Academic Libraries," keynote speaker, Greater Western Library Alliance (GWLA) Cyber-infrastructure Conference, Kansas City, MO, May 28, 2014.

"Research Infrastructure," panel moderator, Association of Research Libraries (ARL) 164th Membership Meeting, Ohio State University, Columbus, OH, May 7, 2014.

"An Eight Year Odyssey in Data Management: Purdue University," International Association of Scientific and Technological University Libraries (IATUL) 2013 Workshop Research Data Management: Finding Our Role, University of Oxford, UK, December 2013.

"Purdue University Libraries & Press: from collaboration to integration," Ithaka Sustainable Scholarship, The Evolving Digital Landscape: New Roles and Responsibilities in Higher Education, libraries as publishers, New York, New York, October 2013.

"Tsinghua and Purdue: Research Libraries for the 21st Century," Tsinghua University, Tsinghua, China, August 2013.

"Purdue Publishing Experience in the Libraries Publishing Coalition," Association of American University Presses Annual Meeting, Press-Library Coalition Panel, Boston, Massachusetts, June 21, 2013.

"Indiana University Librarians Day: Purdue University Libraries Ready for the 21st Century," Indiana University Purdue University Indianapolis (IUPUI), June 7, 2013.

"Purdue University Libraries and Open Access; CNI Project Update," Coalition for Networked Information, San Antonio, TX, April 5, 2013.

Memorial Resolution, honoring Joseph Brannon, to the Board of the Association of College & Research Libraries, Seattle, WA, January 2013.

"An overview of sustaining e-Science collaboration in an Academic Research Library—the Purdue experience," Duraspace e-Science Institute webcast, October 17, 2012.

"The Role of Libraries in Data Curation, Access, and Preservation: an International Perspective, " Panel Moderator, 78th General Conference and Assembly, International Federation of Library Associations, Helsinki, Finland, August 15, 2012.

"21st Century Libraries," moderator of First Plenary Session, International Association of Technological University Libraries 33rd Annual Conference, Singapore, June 4, 2012.

"Planning for New Buildings on Campus," panel presenter, University of Calgary Building Symposium on Designing Libraries for the 21st Century, Calgary, Alberta, Canada, May 17, 2012.

"Data Management and e-Science, the Purdue Response." Wiley-Blackwell Executive Seminar-2012, Washington, DC, March 23, 2012.

"An overview of Sustaining e-Science Collaboration in Academic Research Libraries and the Purdue Experience." Leadership & Career Development Program Institute, Association of Research Libraries (ARL). Houston, TX, March 21, 2012.

"An overview of Data Activities at Purdue University in response to Data Management Requirements." Coalition for Academic Scientific Computation (CASC). Arlington, VA, September 8, 2011.

"Getting on Track with Tenure," Association of College and Research Libraries (ACRL) Research Program Committee. Washington, DC, June 26, 2011.

"Integration of the Press and Libraries Collaboration to Promote Scholarly Communication," Association of Library Collections & Technical Services (ALCTS) Scholarly Communication Interest Group – American Library Association, New Orleans, Louisiana, June 25, 2011.

"Cooperation for improving access to scholarly communication," with N. Lossau (Germany), C. Mazurek (Poland), J. Stokker (Australia), panel moderator and presenter, Second Plenary Session, International Association of Scientific and Technological University Libraries (IATUL) 32nd Conference 2011, Warsaw, Poland. May 29-June 2, 2011.

"Riding the Wave of Data," STM Annual Spring Conference 2011. Trailblazing & transforming scholarly publishing **2011**. Washington, D.C., April 28, 2011.

"Confronting old assumptions to assume new roles: physical and operational integration of the Press and Libraries at Purdue University," keynote speaker, 2011 BioOne Publishers & Partners Meeting. Washington, D. C., April 22, 2011.

"Are MLS Graduates Being Prepared for the Changing and Emerging Roles that Librarians must now assume within Research Libraries?" University of Oklahoma Libraries Seminar, March 4, 2011, Oklahoma City, Oklahoma.

"The Future Role of University Librarians," the University of Cape Town, South Africa, February 25, 2011.

"New Roles for Librarians: the Application of Library Science to Scientific/Technical Research – Purdue University – a case study. International Council for Science and Technology (ICSTI); Ottawa, Canada. June 9, 2009.

"Reinventing Science Librarianship: Models for the Future," Association of Research Libraries / Coalition for Networked Information. October 16-17th, 2008, Arlington, VA. Moderator and convener of Data Curation: Issues and Challenges.

"Practical Implementation and Opportunities Created at Purdue University," African Digital Curation Conference, Pretoria, South Africa, (live video transmission), February 12, 2008.

Keynote speaker. "*Scholarly Communication & Academe: The Winter of Our Discontent,"* XXVII Charleston Conference on Issues in Book and Serial Acquisition, Charleston, South Carolina. November 8, 2007.

Keynote speaker. "*Enabling Access to Scientific & Technical Data-sets in e-Science: a role for Library and Archival Sciences,"* Greater Western Library Alliance (GWLA), Tucson, Arizona. September 17, 2007. A meeting of library directors and vice presidents for research of member institutions.

"*The Challenge of e-Science Data-set Management to Domain Sciences and Engineering: a Role for Academic Libraries and Librarians,"* KIT (Kanazawa Institute of Technology)/CLIR (Council of

Library and Information Resources) International Roundtable for Library and Information Science, July 5-6, 2007. Invited to participate by the Deputy Librarian of Congress.

International Association of Technological University Libraries (IATUL), Stockholm, Sweden. June 8, 2007**.** Invited paper**,** *Enabling International Access to Scientific Data-sets: creation of the Distributed Data Curation Center (D2C2).*

"A New Collaboration for Librarians: The Principles of Library and Archival Sciences Applied to the Curation of Datasets," Symposium of the Libraries and the College of Engineering, University of Louisville, April 6, 2007.

"Purdue University Libraries: Through Pre-eminent Innovation and Creativity, Meeting the Challenges of the Information Age," Board of Trustees, Purdue University, February 15, 2007.

ARL Workshop on New Collaborative Relationships: The Role of Academic Libraries in the Digital Data Universe, September 26-27, 2006, Arlington, VA. Invited participant.

NARA and SDSC: A partnership. A panel before the National Science Foundation, June 27, 2006. Arlington, VA. Invited participant.

"Kaleidoscope of Scientific Literacy: fusing new connections," with Diane Rein, American Library Association, Association of College and Research Libraries, Science & Technology Section, Annual Conference, New Orleans, June 26th, 2006.

"Leadership for Learning: Building a Culture of Teaching in Academic Libraries – an administrative perspective," American Library Association, Association of College and Research Libraries, Instruction Section, Annual Conference, New Orleans, June 25th, 2006.

"Building an interdisciplinary research program in an academic library:
how the Libraries'associate dean for research makes a difference at Purdue University," International Association of Technological University Libraries (IATUL), Porto, Portugal, May 23rd, 2006.

"Enabling Interaction and Quality in a Distributed Data DRIS," *Enabling Interaction and Quality: Beyond the Hanseatic League*. 8th International Conference on Current Research Information Systems, with D. Scott Brandt and Michael Witt. Promoted by euro CRIS, Bergen, Norway, May 12th, 2006, Brandt, and Witt presented in person

"Interdisciplinary Research," with D. Scott Brandt, Coalition for Networked Information (CNI) Spring Meeting: Project Briefing, Washington, D.C., April 3rd, 2006.

"An Interview with Purdue's James Mullins," a podcast submitted by Matt Pasiewicz, on *Educause Connect,* http://connect.educause.edu/James_L_Mullins_Interview_CNI_2005

"Managing Long-Lived Digital Data-sets and their Curation: Interdisciplinary Policy Issues," Managing Digital Assets Forum, Association of Research Libraries (ARL), Washington, D.C., October 28th, 2005.
"The Odyssey of a Librarian." Indiana Library Federation (ILF), District 2 Meeting, South Bend, Indiana. October 4th, 2005.
"New College Library Standards," Standards Committee Presentation, ALA, Chicago, July 7, 2000.
SUNY Library Directors, Lake George, New York. "*The College Library Standards: a Tool for Assessment."* April 5, 2000.
Tri-State College Library Association, *Finding You Have Talents You Never Knew You Had,* Penn State Great Valley, March 25, 2000.
Using Web Statistics, American Library Association, New Orleans, June 24, 1999.
Keynote speaker at the JSTOR Workshop, January 29, 30, 1999. University of Pennsylvania, Philadelphia, PA.
"The New Standards for Electronic Resources Statistics," Society of Scholarly Publishers, Washington, D.C., September 17, 1998.
"Evaluating Online Resources: Now that you've got them what do you do?," joint presenter with Chuck Hamaker, LSU, at the NASIG Conference, Boulder, Colorado. June 1998.
"What Employers Are Looking for in New Librarians?" Pennsylvania Library Association, Philadelphia. September 26, 1997.
"The Theory of Matrix Management" panel presentation of the Comparative Library Organization Committee of the Library Organization and Management Section of the Library Administration and Management Association, a division of the American Library Association, Annual Meeting, Chicago, June 24, 1990.

Professional Involvement: (summary of recent emphasis)

The focus for my professional involvement and research has moved recently toward managing massive data-sets. This has resulted in working with faculty in the sciences and technology to determine how librarians can collaborate in managing, curating, and preserving data-sets for future access and documentation. This has included various speaking opportunities as well as participation in planning with the National Science Foundation (NSF) on ways in which librarians can be integrated more completely into the funded research process. Participation in the Kanazawa Institute of Technology/Council of Library Resources Roundtable was particularly rewarding and provided new opportunities to share with international colleagues the issues surrounding data-set management. I was

the champion for the creation of the Distributed Data Curation Center (D2C2) at Purdue University (http://d2c2.lib.purdue.edu/)

Throughout my career, beginning with my dissertation, I have been actively involved with assessing and evaluating libraries. In the fall of 1999, I contacted twenty-two academic library directors to determine whether the need was also felt by others. The response was overwhelmingly affirmative. This resulted in a meeting at ALA Midwinter, January 2000. A formal meeting followed at Villanova University in April 2000. As convener, I helped to form the University Libraries Group (ULG), modeled after the Oberlin Group for college libraries. The ULG is made up of university libraries that support diverse wide-ranging programs through doctoral level and have a level of support that places them in the top tier of academic institutions. A few of the member libraries, along with Villanova, are William and Mary, Wake Forest, Lehigh, Carnegie-Mellon, Tufts, Marquette, Miami of Ohio, and Southern Methodist.

In 1994 appointed to the Standards Committee, College Section, Association of College and Research Libraries. During the next six years, the Committee concentrated on changing the focus of the standards from quantitative analysis of input and output factors to emphasis on assessment of the outcome. Culmination of the work was a re-issue of the *Standards for College Libraries* in 2000. The knowledge gained through my work experience enabled me to formulate the changes needed in the standards. This work allowed for close collaboration with accrediting agencies, both professional and regional.

During this same time another focus emerged, the impact of digital resources. Through my work on the JSTOR Statistics Task Force, standards were developed on the collection of use of electronic databases. This Standard was later adopted in 1998 by the International Consortium of Library Consortia (ICOLC).

In 2002, the American Library Association appointed me to serve as the liaison to the Marketing and Management Section of the International Federation of Library Associations (IFLA).

Professional Service: (representative list)

Nominations Committee, Association of Research Libraries (ARL), 2016.

Steering Committee, Scholarly Publishing and Academic Resources Coalition (SPARC), 2016 – 2017.

"Excellence in Library Services," Chair, Review Team, University of Hong Kong, Hong Kong, August 24-27, 2015.

Chair, Management Advisory Board, 2015-2017; Member, Scientific Advisory Board, arXiv, Cornell University, 1/1/2013 – present.

Advisory Board for the Wayne State University School of Library and Information Science, July 2012 – present.

Advisory Board for Microsoft Academic Search, 2012 – 2015. Redmond, WA.
Transforming Research Libraries, a Strategic Direction Steering Committee of the Association of Research Libraries (ARL), 2012-2015.
Science and Technology section, representing ARL, International Federation of Library Associations (IFLA), Chair, 2013 – 2017; Member, 2011 to present.
Member of University of Pretoria, South Africa, Library Review Committee. August 2013.
Co-chair, Local Arrangements Planning Committee for 2013 Conference, Association of College and Research Libraries (ACRL), a division of the American Library Association (ALA).
Association of Research Libraries Leadership & Career Development Program Mentor, 2011-2017.
e-Science Task Force, Association of Research Libraries. July 2006 – present. Chair, October 2011 – October 2012.
Board of Directors, International Association of Technological University Libraries (IATUL). January 2008 – December 2014.
Midwest Collaborative for Library Services (MCLS); Board Member, October 2010 – December 2012.
Chair, Library Directors, Committee on Institutional Cooperation (CIC), July 2010 – June 2012.
Board of Directors, Association of Research Libraries (ARL); October 2008 – October 2011.
Scholarly Communication Steering Committee, Association of Research Libraries (ARL) 2008-2011.
Editorial Board, *College and Research Libraries*, Association of College and Research Libraries, American Library Association. January 2008 – December 2014.
Chair, Organizing Committee for IATUL Conference 2010, June 21-24, 2010, Purdue University, West Lafayette, Indiana/Chicago, Illinois.
Conference Planning Committee for National Conference of the Association of College and Research Libraries, 2009, Seattle, Washington.
Research Committee, Association of College and Research Libraries, ACRL, division of ALA. 2002-2007, chair, 2005-2007.
Association of Research Libraries, Search and Screen Committee, Executive Director. March – January 2008.
Center for Research Libraries, Board of Directors. April 2006 – April 2012.
Academic Libraries of Indiana, Board of Directors, 2004 – present. Vice-president, 2005-2007. President, 2007- 2009.

ALA Representative to the International Federation of Library Associations (IFLA), Marketing and Management (M&M) Section, initial term 2003-2007, re-appointed for second term, 2007-2011.
Invited to represent Research Libraries at the ACRL/3M Wonewok Retreat to assess Marketing of Academic Libraries, October 2002.
Hugh A. Atkinson Award Committee, LAMA Representative, ALA, 2001-2005.
Program Committee, Library Administrators and Management Association (LAMA), a division of ALA. 1996-2001.
ACRL, Standards and Accreditation Committee, a division of ALA. Liaison to RBMS Section of ACRL. 1997-2002.
Elected to the Executive Committee of LAMA, LOMS, a division of the American Library Association, 1998-2000. Nominated as Chair/Elect for 2003 – 2005.
Columbia University Press Advisory Committee. 1996 - 2000.
LITA/LAMA Conference Evaluation Committee, Pittsburgh, Pennsylvania, October 1996.
"New Learning Communities," Coalition for Networked Information, Indianapolis. November 19-21, 1995. Facilitator for invitational, national conference committed to developing collaborative learning and teaching techniques, involving librarians.
Planning Committee-Evaluation. LITA/LAMA 1996 Conference, Pittsburgh. This first conference, to be held jointly between two divisions of ALA, will focus on new technologies within libraries.
Indiana Cooperative Library Services Authority (InCoLSA), elected to Executive Committee, April 1991, served as President in 1993-94. InCoLSA is a statewide network of academic, public, school, and special libraries that supports library cooperation for cataloging, interlibrary loan, collection development and application of new technologies.
Governor’s Conference on Libraries and Information Services. Served on Planning Committee, Academic Libraries Representative, appointed by the Governor to represent academic libraries in Indiana, Chair, Finance Committee, April 1989-July 1991.
Indiana Library Endowment Foundation Board, 1984-92. Charter Member, 1984, President, 1988-1992. 2004-2005.

University Service: (Summary)

Served on search and screen committees for senior positions including chancellor, dean, and directors; most recently I have been asked to serve on the search committee for the provost of Purdue University. At MIT service included the Library Council & appointment to the Administrative Council by President

Vest, 2001-2003 & Member of the Faculty Committee on the Library System. At Purdue appointed by the President to the Search Committee for the Provost, October 2007 to May 2008; member of the Capital Projects Committee, and IT Operational Oversight Committee as senior academic dean, 2008-2014.

Global Council, Global Policy Institute, 2012 – 2016.

Academic Program Excellence and Rankings (APER) project team, 2014.

Representative of the Academic Deans on the Re-engineering Business Operations, Purdue University, 2016 –

Academic Deans Council chaired by Provost – 2004 – 2017.

University Promotion and Tenure Committee – 2006 – 2017.

"Outstanding Team Award, Electronic Reserve Project," served as Chair, recognition awarded by the President of Villanova University to one team who made an outstanding contribution to the operations of the University, selected by a committee of administrators, faculty, and staff. Awarded September 9, 1999.

Nominated for the IUSB Lundquist Award, 1995 & 1996. The Lundquist award is given to faculty who have "exhibited excellence in teaching, scholarly or artistic achievement, and diversified relevant service..."