

Doctoral Dissertation

**Evaluation of Augmented Reality as a Memorization
Aid and Projection Mapping onto Deformable Objects**

Yuichiro Fujimoto

February 5, 2015

Department of Information Science
Graduate School of Information Science
Nara Institute of Science and Technology

A Doctoral Dissertation
submitted to Graduate School of Information Science,
Nara Institute of Science and Technology
in partial fulfillment of the requirements for the degree of
Doctor of ENGINEERING

Yuichiro Fujimoto

Thesis Committee:

Professor Hirokazu Kato	(Supervisor)
Professor Naokazu Yokoya	(Co-Supervisor)
Associate Professor Christian Sandor	(Co-Supervisor)
Associate Professor Daisuke Iwai	(Osaka University)

Evaluation of Augmented Reality as a Memorization Aid and Projection Mapping onto Deformable Objects*

Yuichiro Fujimoto

Abstract

Augmented reality (AR) is a technology that overlays information generated by a computer onto the real world and that makes a user perceive the information as if it existed in the real world. Although AR has been gradually applied to entertainment applications such as gaming, advertising and some other fields for general users, AR is also intended to help workers. In particular, one of the most promising fields with active research is industrial task support. However, actual industrial tasks where support systems have been already employed are quite limited.

Objective of this study is to solve several problems which inhibit the application of AR and to expand applicability of AR. To investigate the reason for limited use, we categorized AR task support systems into two types in terms of information characteristics.

The first type of information in AR enables a user to intuitively understand information by performing the association between a virtual object and a real object. The second type of information in AR provides a high realism by the replacement of parts of a real object with a virtual object (e.g., a texture).

The first feature works well for assembly, inspection, maintenance, and some training tasks of these. For the training process, a lot of studies aims at enabling users to intuitively and promptly understand task process with the spatial cognition. Another objective of this process is to make users memorize task flow and information required for the task. In order to find effectiveness of AR for this aspect, I focus on the compatibility between AR and the memorization. AR is a technology that overlays information about a specific location on the user's real-world view. Users can perceive, not only

*Doctoral Dissertation, Department of Information Science, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-DD1261012, February 5, 2015.

the information itself, but also its location. It is also known that humans can easily memorize and retain information if this information is associated with some real-world locations. From these two facts, I hypothesize that “If information is displayed in relation to specific locations on the real world by AR, then the users can have better memorization task results than when the information is displayed in an unrelated location.” If the effectiveness of AR in enhancing a user’s memory skill can be proven, I could argue that AR support systems are effective for not only providing information related to tasks but also facilitating the memorization of tasks. Through the several experiments, I found several significant differences among the situation with AR and without AR which supports this hypothesis.

The second feature works well for the design support task, however, the technique is not as mature compared to the first one. Particularly, one of most significant examples of design process support is rapid-prototyping with AR. AR enables users to check a prototype of a product with various appearances within a short period of time by overlaying or projecting a texture onto it. One of the current problems in this area is the difficulty of using deformable products such as clothes which are easily deformed in a short time. This example shows that current technologies have a big limitation in rapid-prototyping system. In this thesis, I developed a design support system for deformable objects and technologies on projection-based AR required for achieving it.

Keywords:

Augmented Reality, Task Support, User Study, Memorization, Projection, Deformable Object

拡張現実感の記憶行為に対する有効性の検証と柔軟物体 へのプロジェクションマッピング*

藤本 雄一郎

内容梗概

拡張現実感の実世界上の特定の位置にあたかも存在するかのように仮想物体を重畳表示させる技術である。この技術は近年、ゲーム等のアミューズメント分野、商品の広告などに広く利用されるようになり、一般普及しつつある。一方、古くから特に活発に研究されているこの技術の他の応用分野として、工業に関する作業支援が挙げられるが、実際の工業作業の支援に拡張現実感が用いられているケースは現状極めて限られている。

本研究では、拡張現実感の工業作業支援への応用を妨げる具体的問題の解決に取り組む。この原因を探るため、拡張現実感の本質に立ち返ると、この技術を作業支援に用いることの特徴は以下の二つに大別できる。一つ目は仮想物体と実物体の関連付けが視覚的かつ直接的に行えるため、分かりやすい情報提示を可能とすることである。二つ目は実物体の一部を仮想物体で置き換えることで、実物感を損なわない、リアルな知覚を可能とすることである。

一つ目は工業製品の組み立て、廃棄、点検、およびそれらの作業のトレーニング工程に対し有効に働く特徴である。これらトレーニング工程では3次元的な作業をユーザに素早く、直感的に理解させることを主たる目的として今まで研究がなされているが、その他に作業に必要な情報や作業工程の流れをユーザに記憶させる、という目的も存在する。そこで本研究では、トレーニング工程の后者の目的に関する拡張現実感技術の新たな有効性を発見するため、「提示情報が本質的に実世界上の位置という情報を含んでいる」という拡張現実感の特性と「位置に関連付けられた情報は記憶されやすく、また想起されやすい」という人間の記憶メカニズムの特性に着目した。以上の二つの特性を基に、拡張現実感による注釈表示において「対象物体の位置に関連付けて情報を表示した場合、無関係な位置に

*奈良先端科学技術大学院大学 情報科学研究科 情報科学専攻 博士論文, NAIST-IS-DD1261012, 2015年2月5日。

表示した場合と比較して、それを見たユーザの記憶に特定のポジティブな影響を及ぼす」という仮説を立てた。この仮説を示すことができれば、拡張現実感を用いた作業支援システムは、不慣れな作業者が作業工程を覚えるトレーニング工程においても有効であると示すことができる。複数の被験者実験を通して、対象物体の位置に関連付けた表示を行った場合とそうでない場合の記憶結果の間に上記仮説を立証するいくつかの有意な差が見られた。

二つ目は工業製品のデザイン作業等に対し有効に働く特徴であり、前者の特徴と比較するとこれを利用した研究例が少ない。このデザイン作業支援の重要な応用例の一つとして、無地白色の実物体を対象とし、拡張現実感を用いてテクスチャの重畳表示を行い短時間にその見た目を変化させることで、様々なデザインを試すことができるラピッドプロトタイピングが挙げられる。この応用の現状の大きな問題は、多くの既存技術が対象の形状変化を許容しないため、短時間での対象物体の形状変化が行えないことである。それゆえデザイン中に形状が容易に変化する衣類等の工業製品に対し、拡張現実感を用いた既存のラピッドプロトタイピングを導入することは困難であった。そこで本研究では衣類のような柔軟物体を対象とした形状の認識手法の提案、及び得られた形状情報を基にしたプロジェクタを用いたテクスチャ投影により、対象物体表面の見た目を変化させるシステムの構築を行った。

キーワード

拡張現実感, 作業支援, 被験者実験, 記憶, プロジェクション, 柔軟物体

Contents

1	Introduction	1
1.1.	Background and Motivation	1
1.1.1	AR for General Users	1
1.1.2	AR for Workers	4
1.1.3	Motivation	5
1.2.	Classification of Industrial Task Support Systems by AR (IAR)	5
1.2.1	Classification by Product Life-Cycle (Classification A)	6
1.2.2	Classification by Information Characteristics (Classification B)	7
1.3.	Factors Relevant to IAR	9
1.3.1	Human Factors for IAR When a User Perceives Information Regarding a Task (B-1)	9
1.3.2	Technologies for IAR When a User Perceives Information Re- garding a Task (B-1)	10
1.3.3	Human Factors Relevant to IAR When Perceiving Information Itself is a Task (B-2)	11
1.3.4	Technologies for IAR When Perceiving Information Itself is a Task (B-2)	13
1.4.	Scope and Research Goal	13
1.4.1	Verifying Effectiveness of AR for Memorization (Related to B-1)	15

Contents

1.4.2	Developing Tracking Technique for Deformable Objects (Related to B-2)	15
1.5.	Contributions	16
1.6.	Outline of Thesis	19
2	Background and Related Work	21
2.1.	Classification of IAR by Product Life-Cycle	21
2.1.1	Product Design	21
2.1.2	Manufacturing	24
2.1.3	Commissioning: Validation and Documentation	25
2.1.4	Inspection and Maintenance	26
2.1.5	Decommissioning	29
2.2.	Classification of IAR by Information Characteristics	29
2.2.1	When a User Perceives Relevant Information about a Task	30
2.2.2	When Perceiving Information Itself is a Task	31
2.3.	Summary	32
3	Verification of Effectiveness of AR for Memorization	33
3.1.	Introduction	33
3.2.	Related Work and Research Scope	36
3.2.1	Human Memorization	36
3.2.2	Features of AR Relevant to Human Memory and Cognition	39
3.2.3	Objective of Study	40
3.2.4	Method to Verify Hypotheses	40
3.3.	Experiment 1: Memorization with a Location on Visual Perception for Individual Factor	41
3.3.1	Experimental System	41
3.3.2	Experimental Design	45
3.3.3	Experimental Procedure	46
3.3.4	Experimental Results and Consideration	47
3.4.	Experiment 2: Memorization with a Location on Visual Perception for Continuous Factor	51
3.4.1	Experimental System	51
3.4.2	Experimental Design	53
3.4.3	Experiment procedure	54

3.4.4	Experimental Results and Consideration	55
3.5.	Experiment 3: Memorization with a Location on Spatial Cognition for Continuous Factor	58
3.5.1	Experiment System	58
3.5.2	Experimental Design	60
3.5.3	Experimental Results and Consideration	61
3.6.	Summary and Future Work	64
3.6.1	Insight Regarding Information Display Methods Using AR for Effective Memorization	65
3.6.2	Future Work	66
4	Design Support System for Deformable Objects Using Projection-based AR	67
4.1.	Introduction	67
4.2.	Related Work	71
4.2.1	Structured Light Approach	71
4.2.2	Vision Based Recognition of Deformable Substrates	72
4.2.3	Embedded Sensor Based Recognition of Deformable Substrates	74
4.2.4	Clothing Design Simulation System	75
4.3.	System Components	76
4.3.1	Camera and Projector	77
4.3.2	Our Deformable Object	77
4.3.3	Our Marker Pattern	78
4.4.	Algorithm	80
4.4.1	Marker Recognition	83
4.4.2	Point Detection	83
4.4.3	Region Division and Grid Calculation	85
4.4.4	Matching with Reference Pattern	85
4.4.5	Error Removal with Voting	86
4.4.6	Interpolation	88
4.4.7	Projection	89
4.4.8	Discussion on Processing Speed	91
4.4.9	Extension To Multi-Pro-Cams	91
4.5.	Recognition and Evaluation	93

Contents

4.5.1	Projection Evaluation for Single Pro-Cam	94
4.5.2	Projection Evaluation for Multi-Pro-Cams	94
4.6.	Projection-based AR for Design Support and Its Advantages	97
4.6.1	Current Design Prototyping Practice	97
4.6.2	Current Use of SAR for Design Prototyping	98
4.6.3	Current Restrictions Using SAR for Design Support	98
4.6.4	Our New Design Support	99
4.6.5	Limitations	100
4.6.6	Application: Clothing Design Simulation System	101
4.7.	Improving Processing Speed	102
4.7.1	For 3D shape measurement	103
4.7.2	For Marker Recognition and Interpolation	105
4.7.3	Application to Clothing Design Simulation System	105
4.8.	Summary and Future Work	106
5	Conclusion	108
5.1.	Contributions and Summary	108
5.1.1	Verification of Effectiveness of AR for Memorization (Chapter 3)	108
5.1.2	Development of a Design Support System for Deformable Ob- jects and a Technique for Identifying Each Location on a De- formable Object (Chapter 4)	109
5.2.	Future Research	110
5.2.1	Verification of Effectiveness of AR for Memorization	110
5.2.2	Development of a Design Support System for Deformable Ob- jects	111
	Publication List	113
	Acknowledgments	116
	Bibliography	120

List of Figures

1.1	Examples of AR applications for general users: (a) The Eye of Judgment, a video game with AR by Sony, (b) Wikitude, a browser with AR by Wikitude, (c) furniture catalog with AR by IKEA	2
1.2	Task support systems with AR: (a) a maintenance support system for a car [7], (b) a maintenance support system for a military automobile [26]	5
1.3	Bar plot showing the number of AR systems which have been employed in industrial tasks until 2011 [26]	6
1.4	Classification of AR industrial task support systems	8
1.5	Composing factors of IAR when a user perceives relevant information about a task: (a) human factors (b) technologies	12
1.6	Composing factors of IAR when perceiving information itself is a task: (a) human factors (b) technologies	14
1.7	Overview of our research: (a) an example of experiments for evaluating relationship between AR and users' memorization (b) an example of our design support system for deformable object (in this case, clothes) .	17
2.1	Examples of product design support systems by AR (a) a car design support with a HMD [44], (b) a car design support with a projector [103]	23
2.2	Examples of manufacturing support systems by AR (a) a welding task support with AR [13], (b) a picking-up task support with AR [24] . . .	25

List of Figures

2.3	Examples of commissioning: validation and documentation support systems by AR (a) a validation support with AR for a power plant [21], (b) a support for creating documentation with AR [67]	26
2.4	Examples of inspection and maintenance support systems by AR (a) a maintenance support with AR for a printer [17], (b) a maintenance support in the psychomotor phase in a procedural task by AR [27] . . .	28
2.5	Examples of decommissioning support systems by AR for a nuclear power plant [35]	29
3.1	Example of “object in place” experiment [28]	38
3.2	Example of research on the spatial cognition and AR: (a) Tang et al. [95] (b) Goto et al. [22]	39
3.3	Overview of an AR annotation system	42
3.4	User’s view through HMD.	43
3.5	Display methods in the experiment 1: (a) Type 1-1: The information is displayed near the location of each CG point on the paper, (b) Type 1-2: The information is displayed at random locations on the paper, (c) Type 1-3: The information is displayed at the same location on the 2D display	44
3.6	View of two type tests	47
3.7	Rate of question answered correctly with each display methods in experiment 1 (mean with \pm standard deviation)	48
3.8	Number of each score which participants gave for each display method in experiment 1	50
3.9	Display methods in the user study 2.	52
3.10	Overview of the experimental steps: (Left) The step in which a participant memorizes each task process. (Right) The step in which a participant recalls each process and picks some blocks from each drawer in the correct order.	54
3.11	Rate of question answered correctly by display methods in experiment 2 (mean with \pm standard deviation)	56
3.12	Number of each score which participants gave for each display method in experiment 2	57
3.13	Display methods in the experiment 3.	59

3.14	Assembly targets of the experiment 3.	60
3.15	Rate of question answered correctly by display methods in shape A of experiment 3 (mean with \pm standard deviation)	62
3.16	Rate of question answered correctly by display methods in shape B of experiment 3 (mean with \pm standard deviation)	63
4.1	Related work on tracking for a deformable object: (a) random dot pat- tern for tracking of a deformable object [102], (b) tracking of a de- formable object with texture information [69]	74
4.2	Related work on reconstruction for a deformable object and projection for a deformable object: (a) real-time reconstruction for a deformable object [112], (b) real-time projection for a deformable object [73]	75
4.3	Difference between a virtual try-on system with a display and a clothes design support system with a projector (a) a virtual try-on system with a display (e.g., [69]), (b) a clothes design support system with a projector (a example application of our technique)	77
4.4	The major system components: Our system consists of projectors, cam- eras which are switchable from IR to RGB function, and deformable object.	78
4.5	Overview of deformable substrates: Our new SAR enabled deformable substrate	79
4.6	Szentandrasei et al.'s marker on the left and our new marker system on the right : A minimum unit of Szentandrasei et al.'s marker is a rectangle. A minimum unit for our marker is a point.	80
4.7	Overview of marker detection algorithm: (a) a pro-cam which optical axes of a camera and a projector are corresponded in advance (b) a pro-cam when optical axes of a camera and a projector separated.	81
4.8	Overview of system algorithm: Processing consists of four factors: 3D measurements, marker recognition, interpolation and projection	83
4.9	Example of first region division. Allowing duplicates, nine rectangles with a same size are used to divide the whole region (for descriptive purpose, two colored rectangles are used).	84

List of Figures

4.10	Example of grid making: (a) By 3D information, each two consecutive points are detected in each small region. (b) Two salient directions are decided. (c) Grid lines are calculated through as many as possible of the detected points.	84
4.11	Small region integration and error removal: In order to integrate small regions, IDs are assigned to each point for voting. The ID with a maximum number of votes is assigned as the final ID of each point on the image	86
4.12	Interpolation: (a) When # of recognized points is one, a location of a point which should be interpolated cannot be predicted. (b) When # of recognized points is three and all points is in a row, two candidates exist. When all points are not in a row, a candidate can be determined uniquely.	89
4.13	Results of projecting texture image, the same image is deformed and projected onto differently deformed surfaces	90
4.14	Results of projecting texture image when heavy occlusions occur	90
4.15	Example of an two blur effects [59] (1) defocus blur (b) oblique blur . .	93
4.16	Projection evaluation for single pro-cam: Each color represents the error value for each point. Double circles denote points decided by the marker recognition and single circle denote points decide by the interpolation.	95
4.17	Projection evaluation for multi pro-cams: Light blue and pink regions represents projected region by each projector	96
4.18	Example of an apparent plane projection onto an arbitrary curved surface by considering geometry information of the projection surface from [91] (a) Without geometrical consideration (b) With geometrical consideration	100
4.19	Example of current problems: (a) a situation without any large deformations (b) a situation when the object is folded	101
4.20	Projection target: a neutral colored piece of cloth	102
4.21	Application: Design support system for clothes	103
4.22	Application: Design support system (with kinect v2) for clothes	106
4.23	Examples of unnatural artifacts of the projected image derived from the temporal discontinuity and the interpolation inaccuracy	107

List of Tables

4.1	Projection accuracy by the single pro-cam: These value illustrates mean distances between each grid point on the deformable object and each projected point by a projector	94
4.2	Projection accuracy by the multi pro-cams: These value illustrates mean distances between each <i>grid point</i> on the deformable object and each projected point by each projector	97
4.3	Processing time of each process	105

CHAPTER 1

Introduction

A lot of research works have applied augmented reality (AR) on industrial task support. However, few systems have been employed in actual use cases. This introductory chapter starts with current AR systems which have been already used by many people in Section 1.1. In Section 1.2, a classification of existing industrial task supports by AR is conducted in terms of two points of view in order to clarify problems. Next, composing factors of industrial task supports by AR are addressed in Section 1.3. Then, specific scopes to be addressed in this thesis are described for each problem in Section 1.4. We gave an overview of our research contributions in Section 1.5. Finally, the outline of this thesis is described in Section 1.6.

1.1. Background and Motivation

In this section, some AR systems are described in terms of two types of users: general users and workers.

1.1.1 AR for General Users

Augmented reality (AR) is the technology that overlays information generated by a computer onto a real world and makes a user perceive it as if the information existed

Chapter 1. Introduction

in the real world. One of the most important features of this technology is derived from its intervention between the real world and the human perceptual system. This intervention can affect this relationship. AR technology is expected to be applicable to various fields. In fact, this technology has made great strides during the last few decades and has come to influence our lives by taking the form of various applications.

Initially, some entertainment companies paid attention to the effectiveness of this technology. Popular examples in the entertainment industry are EyePet [89] (Figure 1.1(a)), The Eye of Judgment [90] and some other games. These games can recognize customized cards by a web-cam and render virtual characters on top of the display as if they really existed in the real world. Some interactions with these characters make a user feel as if he or she really interact with these.

Layer (Layer [47], Figure 1.1(b)), Metaio GmbH (Junaio [54]), Wikitude GmbH (ex-Mobilizy GmbH, Wikitude [105]), Tab Inc. (ex-Tonchi dot Inc. Sekai Camera, [93]) and some other companies focused attention on the applicability of AR to the locative services. This service uses a mobile device such as a cell-phone with a GPS and an electronic compass, thus putting AR into practical use. These companies provided an application which visualizes information about various shops (e.g., restaurant, clothes, etc.) and allow users to locate (e.g., a text or an image) information associated with specific locations on the real world. These functions are achieved by using built-in GPS and compass to overlay information on the image captured by the camera.



Figure 1.1. Examples of AR applications for general users: (a) The Eye of Judgement, a video game with AR by Sony, (b) Wikitude, a browser with AR by Wikitude, (c) furniture catalog with AR by IKEA

1.1. Background and Motivation

Not only companies focused on the high level amusement of AR technology. It's doubtless that the tremendous amount of videos uploaded onto websites (e.g., Youtube) by programmers, designers, and other general users all over the world facilitated the awareness of AR. In order to make a marker-based AR system, a relative location between a camera and a marker should be calculated. This requires the strong mathematical knowledge and the programing skill. Previously, people had to implement it by themselves, which prevented people from making an AR system. The introduction of ARToolKit [3] removed this barrier and made it possible to use AR as a tool to make interesting applications by programmers who have standard programming skills, and not only an AR researcher who has specialized programming background. Programmers all over the world have combined AR with skills or ideas they had and uploaded a lot of videos about their systems on some website which familiarized the general public with AR.

Recently, we can often see an AR marker which assumes the role of an entrance to an AR experience along the street. One of popular examples is the paper poster with AR marker for an advertisement of a movie. If a user installs a specific application to his or her smart phone and look at this poster through the phone, he or she can enjoy a 3D model animation related to a movie and get information about that movie. The same concepts can be seen in magazines, brochures and so on. These are used for marketing and sales.

As an example of services using an AR feature to make users understand the size of the object intuitively, I explain about IKEA's application [30]. IKEA provides the service for users to check what IKEA's furniture looks like in his or her room directly by placing a 3D model of the furniture with the actual size of furniture onto an image of his or her room (Figure 1.1(c)). The United States Postal Service (USPS) provided a web service to answer the following users' question: "What size of a box should we choose to send this object?" By printing out an AR marker with a specific size and capturing it by a web-cam, a 3D model of the box with the actual size of the box is displayed onto the camera image so that a user can check the size of boxes intuitively in his or her home.

1.1.2 AR for Workers

All of the examples mentioned above are for general users' personal use. On the other hand, AR is also intended to help workers for industrial tasks. We call industrial task support systems using AR "IAR" in this thesis.

In this thesis, we define the industrial task as the process to be performed by workers of companies or some other organizations in order to achieve a specific goal (e.g., making a product). If we think about an example of assembly task in industry, some of tasks are automated and performed by a robot. But other tasks which require complicated procedures or flexible behaviors should still be performed by a worker. In the last several decades, the industrial development had been providing a very significant contribution to the development of Japan itself. However, recently, by the conspicuousness of other developing countries, the speed of the development is decreasing and the good time is coming to an end. In order to acquire continuous industrial developments, it is desirable that each industrial process is performed effectively without any error (of course, the degree depends on each process). Among other problems, it can be exemplified that because of the aging of skilled workers, the training for apprentices is an urgent matter in some industrial fields. In addition, it is also desirable to add additional value to each product by the sophistication of the product design. In order to solve these problems and support workers effectively, AR has been studied by a lot of researchers for a few decades.

One of example of industrial task support by AR is a concept proposed by BMW [7] (Figure 1.2(a)). A car maintenance worker wears a head mounted display (hereinafter HMD) and can confirm procedural maintenance tasks intuitively by overlaying a 3D model of the inside of a bonnet and information related with the task on a camera image attached to an HMD. In addition, this system allows users to do the hands-free manipulation by the voice recognition, which enables users to focus on the maintenance task itself. For another example, Henderson et al. [26] (Figure 1.2(b)) proposed an AR system to assist military mechanics in conducting routine maintenance tasks by enhancing localization with overlaid labels and context-setting 2D and 3D graphics. This system enables users to understand where and how to use maintenance tools intuitively.

1.2. Classification of Industrial Task Support Systems by AR (IAR)

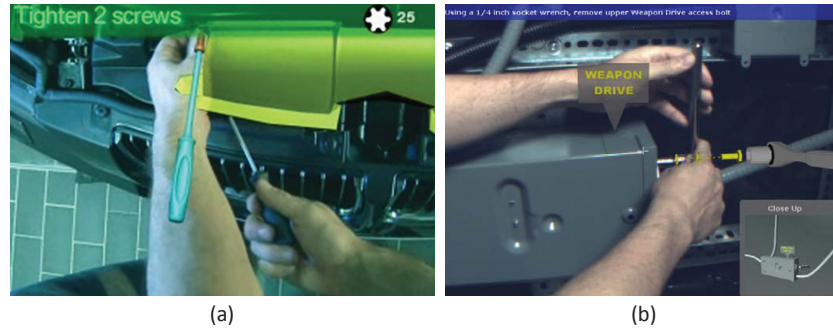


Figure 1.2. Task support systems with AR: (a) a maintenance support system for a car [7], (b) a maintenance support system for a military automobile [26]

1.1.3 Motivation

A lot of studies on similar concepts have been conducted for a long time. However, few systems were used outside of laboratories and employed to actual industrial task support. According to the survey conducted by Fite-Georgel [19] in 2011, only two AR systems (two items of “in the fields” in Figure 1.3) have been employed in the industry. As of 2014, a significant change in the situation has not happened yet. Moreover, in most areas of industrial tasks, no AR system has not been employed yet. Considering AR features and advantages that users can check task procedure in situ in the real world, the use of such kind of systems would be quite effective and promising for various task support. Objective of this research is to solve several specific problems in problems which inhibits the application of AR and expand the applicability of AR. In the next section, We make a classification and a systematization and consider about the remaining technical problems and human factors for the industrial task support by AR.

1.2. Classification of Industrial Task Support Systems by AR (IAR)

In this section, at first, we classify existing IAR into five types in terms of the product life-cycle. Second, we re-classify these systems into two types in terms of feature of information characteristics. Figure 1.4 illustrates these classifications.

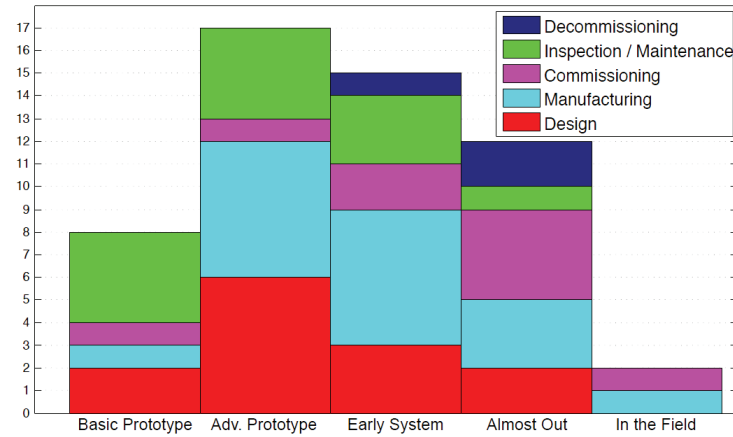


Figure 1.3. Bar plot showing the number of AR systems which have been employed in industrial tasks until 2011 [26]

1.2.1 Classification by Product Life-Cycle (Classification A)

In terms of a product life-cycle, IAR can be divided into five types [19].

Product Design (A-1)

“Product Design” focuses on generating ideas to be conceptualized to a tangible object. This process often requires communication between designer, manufacturer (all stakeholder) and final customers to evaluate ideas and prototypes.

Manufacturing (A-2)

“Manufacturing” is the act of transforming goods (raw material or manufactured) into a more complex product that is ultimately delivered to an end-user. This process mainly focuses on assembling objects. This task also requires the training of inexperienced workers. So this task can be divided into two parts again: (a) Assembly guidance, (b) Training.

1.2. Classification of Industrial Task Support Systems by AR (IAR)

Commissioning: Validation and Documentation (A-3)

“Commissioning: Validation and Documentation” is the process of verifying that all systems and components of a product are installed and functioning as required by the client. This process includes verification of the product against the plan, testing its functionality and document discrepancy when required.

Inspection and Maintenance (A-4)

“Inspection and Maintenance” is the action, sometimes regulated, to verify the condition of the product and, if required, the repair or replacement of faulty components. In order for the product to function properly, Inspection and Maintenance correspond usually apply codified and standardized.

Decommissioning (A-5)

“Decommissioning” is the act of retiring a product when it reaches the end of its life. It may include dismantlement, decontamination and recycling.

1.2.2 Classification by Information Characteristics (Classification B)

If we divide each type of classification A in terms of information characteristics, this can be divided into following two types again: “Case when a user perceives relevant information about a task” and “Case when perceiving information itself is a task”. These two have a tendency to be discussed collectively. However, those actually have completely different objectives and features. In this thesis, we consider those separately.

Case when a user perceives relevant information about a task (B-1)

We define that task support systems displaying information such as a textual annotation and a 3D model onto the appropriate location of the real world are included in this classification. Allowing a few exceptions and categorizing generally, this classification includes A-2, 3, 4 and 5 in section 1.2.1. For instance of A-4: “Inspection and Maintenance”, we exemplify Henderson et al.’s research [26]. Each task of the procedural

Chapter 1. Introduction

maintenance is displayed as an overlaying 3D model and a textual annotation onto the real world image through an HMD for a worker. This system enables users to understand where and how to use maintenance tools intuitively. In that case, displaying a textual annotation and a 3D model are information related to a task to be performed by a worker, and a worker actually performs the task by perceiving displayed information and recognizing a task process.

Case when perceiving information itself is a task (B-2)

On the other hand, there is another case when perceiving information itself is a task process (overlaying a virtual object as an alternative to a real object). Generally categorizing, A-1 (and a few exceptions of A-2, 3, 4 and 5) in section 1.2.1 is included in this classification. For instance of A-1: “Product Design”, we exemplify Verlinden et al.’s research [103]. At first, an automobile-shaped mock-up is made from a white material. The system consists of a camera and a projector which projects a texture image onto it correctly, which enables a user to feel as if it was physically painted by ink or some other ways. With this system, when a user wants to check a different appearance of a mock-up, all he or she has to do is change the projected texture image, and not remake a mock-up. It would be quite useful in order to reduce the cost of the design process. In that case, the displayed information (a projected texture image) is an alternative to physical painting such as using ink. It can be interpreted that perceiving a projected image and a real object and evaluating these themselves are task processes.

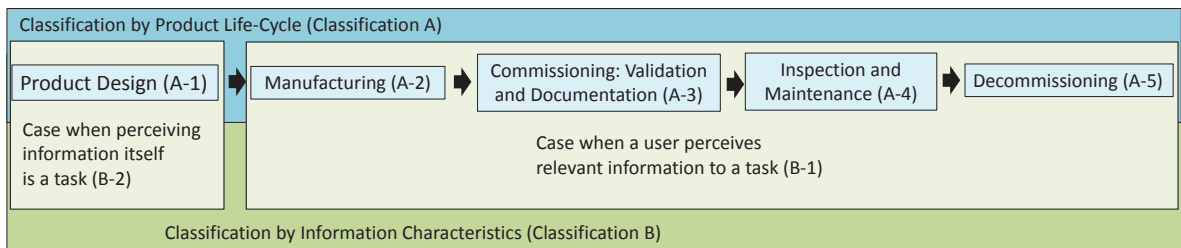


Figure 1.4. Classification of AR industrial task support systems

1.3. Factors Relevant to IAR

We now classify the factors relevant to IAR. In order to distinguish between factors that have already technically matured (or have been verified adequately) from those that have not (or have not yet been adequately verified). We can divide the factors related to IAR into two broad categories: required technologies and human factors. Required technologies denote elemental technologies that required for IAR systems, and define the systems that can be constructed at any given time. Human factors represent the effects of AR features on users in order to justify the use of AR to support industrial tasks. Note that the factors mentioned here are the main factors important in existing IAR systems (see Chapter 2).

1.3.1 Human Factors for IAR When a User Perceives Information Regarding a Task (B-1)

Figure 1.5(a) shows human factors in IAR system when a user perceives relevant information regarding a task (B-1). For descriptive purposes, we separately consider the effects of IAR systems on perception, on the task itself, and on training. Note that these cleavages are not entirely independent and have mutual influences. Effects on perception represent the effects of information displayed using AR on a user's perception. Examples of such effects that have been studied include visibility (e.g., [22, 83, 46]), spatial cognition (e.g., [22, 26, 83, 95]), and mental load (e.g., [85] [100]). In particular, the effect of IAR systems on spatial cognition is intimately related to a main feature of AR, whereby information can be displayed at a specific location with a specific posture in the real world. The effectiveness of AR is derived from this effect. With regard to the effects of IAR systems on the task itself, research involving a comparison of a situation where AR was used to one where it was not has shown that AR improves the speed of task performing (e.g., [27, 26, 24, 85, 84, 13]) as well as task accuracy (fewer mistakes) (e.g., [27, 85, 84]). In the classification of B-1: a "case when a user perceives information relevant to a task (overlaying a virtual object to indicate the task procedure)" the effectiveness of AR in task support can be evidenced by a worker being able to intuitively understand the task procedure as well as information related to the task by observing information related to the task, such as text or a 3D model in situ. Compared to existing methods, such as paper documentation used in many current sce-

narios, the overlaying of information onto the real world in order to enable a worker to more easily and intuitively understand a task process sounds plausible. In fact, several experiments have been conducted in order to verify this concept the world over (e.g., [27, 24]). However, the manner in which AR improves task performance varies with the task and the field of application. If AR was used for every case, the degree of its benefit would remain unclear. Following this, we consider the effects of IAR on training. Although the idea that AR can positively affect a training situation for apprentices has been around since early research on AR [61], concrete and quantitative considerations and verification of these hypotheses have not yet been conducted. Training of the B-1 category consists of two steps. In the first step, a worker memorizes the required information in order to perform tasks and for the task procedure itself. In the second step, a worker repeatedly performs a task in order to improve his or her ability with regard to it. In particular, quantitative evaluations of the effect of IAR on memorization in the first step have not yet been conducted to the best of our knowledge, despite the obvious importance of this effect. In this study, we address this problem and verify the effects of AR on memorization as the first research objective.

1.3.2 Technologies for IAR When a User Perceives Information Regarding a Task (B-1)

Figure 1.5(b) shows the required technologies for an IAR system when a user perceives information relevant to a task (B-1). These technologies can be categorized into three parts: tracking, displaying, and human interface technologies. In the B-1 category, tracking technology is necessary to display information, such as textual annotation and 3D models, with specific postures at specific locations. For example, a 3D arrow that indicates a specific location for a user should be displayed at the exact location. We can further categorize tracking into three targets: specific objects, such as cars, and home electronics, users, and environments such as a factory. Although we can consider such a tracking methodology as vision-based (marker-based, natural feature-based and others), and sensor-based, we will not consider technical methodology in detail here. We can divide an object into two types according to rigidity: rigid (e.g., [1]) and non-rigid (deformable object). Environments can be also divided into static environments without moving objects (e.g., [42]) and dynamic environments (e.g., [94]). Among these, tracking dynamic environments and deformable objects has not yet technically

matured. However, considering practical applications in the field of B-1, few target applications exist in any case (see Chapter 2). Therefore, we do not address these problems here. In B-1, examples of displaying technologies are visual effect technologies (such as X-ray AR [81]), which process afford easier perception for users, and view management technologies (e.g., [23]), which decide where information is displayed in order to make it highly visible to users. For human interfaces in IAR systems in B-1, various input methods ensure that the user can make inquiries regarding the necessary information in a given situation. Gesture interfaces that track users' hand movements (e.g., [70]), voice interfaces that use voice command recognition (e.g., [70, 110]), and tangible interfaces that employ an ordinal keyboard and a touch panel, have been actively researched in the last two decades. An example of other required technologies in B-1 is a technology for predicting a user's current situation in a stepwise task procedure (e.g., [68]). This technology is for appropriately supporting a user in accordance with his or her current situation. Moreover, technologies for creating information displayed by AR, such as 3D model reconstruction, are needed. Technologies mentioned here with a high-priority have been researched for a long time, and are reaching relative maturity.

1.3.3 Human Factors Relevant to IAR When Perceiving Information Itself is a Task (B-2)

Figure 1.6 (a) illustrates human factors for the IAR system when perceiving that information itself is a task (B-2). As shown in Figure 1.6(a), the number of human factors of B-2 is fewer than those of B-1. As mentioned in the previous section, the most beneficial merit of AR for this category is that AR can reduce the design costs of products (for both time and money) by replacing expensive repetitive remaking of a physical 3D prototype with overlaying various texture images onto a physical object [49, 78]. Therefore, because the importance of human factors in this category is relatively low, we do not address these in this research.

Human factors and technologies of IAR when a user perceives relevant information about a task

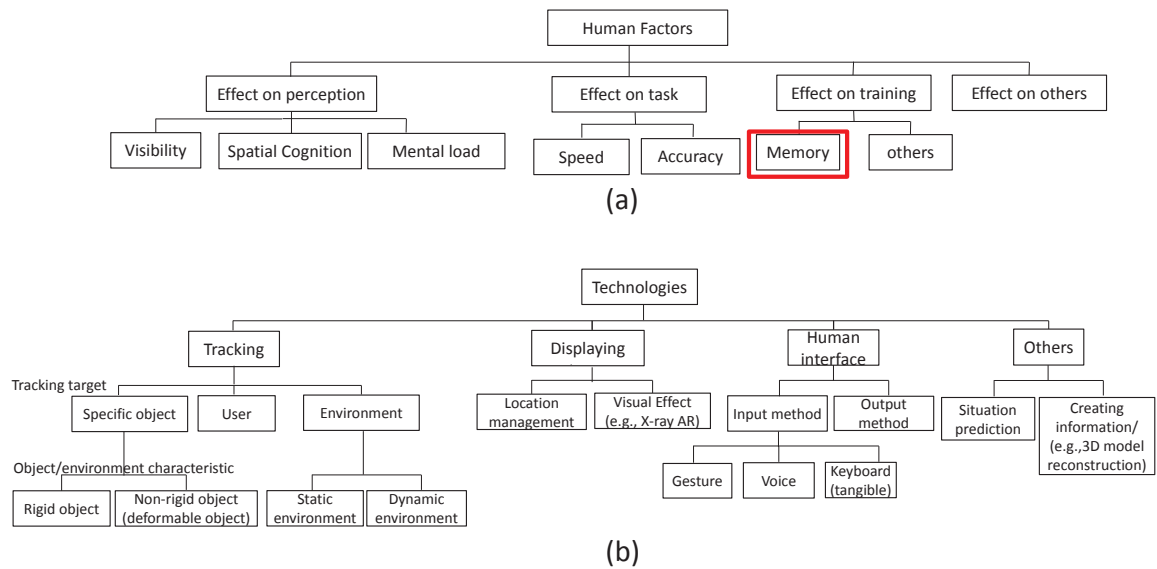


Figure 1.5. Composing factors of IAR when a user perceives relevant information about a task: (a) human factors (b) technologies

1.3.4 Technologies for IAR When Perceiving Information Itself is a Task (B-2)

Figure 1.6(b) shows the required technologies for an IAR system when perceiving that information itself is a task (B-2). As mentioned earlier, these can be categorized into three parts: tracking, displaying, and human interface technologies. In this category, since most target objects are specific, such as cars and home electronics, tracking technologies of users and the environment are not needed in most cases. The tracked objects are divided into rigid and non-rigid objects (deformable). Most existing design support systems are intended for rigid objects, such as cars, home electronics, and tableware [78, 49, 103]. However, design support for a deformable object, such as fabric (e.g., clothes and furniture) has not yet been implemented. Although there are several tracking technologies for deformable objects, applying these technologies to design support leads to limitations and problems. We will address these problems in detail in Section 1.4 and Chapter 4. We develop a tracking technology for deformable objects that is appropriate for design support applications as our second research topic.

Another important technology in this category is high-realistic rendering, which includes technologies that allow users to view an object overlaid or a projected texture image as if it were physically painted by considering complex light modulations (such as inter-reflection, and specular reflection), environmental illumination, and the reflectance properties of the target surface. This is because the design process places a high value premium on careful examination of the texture and the functionality of a product. A large number of studies have been conducted on such high-realistic renderings (e.g., for display situations [38] and for projection scenarios [113]).

1.4. Scope and Research Goal

In the previous section, we described several problems in the use of AR in an industrial task support system. In this research, we aim to address a few specific problems to encourage the application of AR to industrial task support. By summarizing the current challenges facing IAR, described in the previous section, we focus on two specific research themes: “verifying effectiveness of AR for memorization” and “tracking for deformable object” in the context of industrial task support.

Chapter 1. Introduction

Human factors and technologies for IAR when perceiving information itself is a task

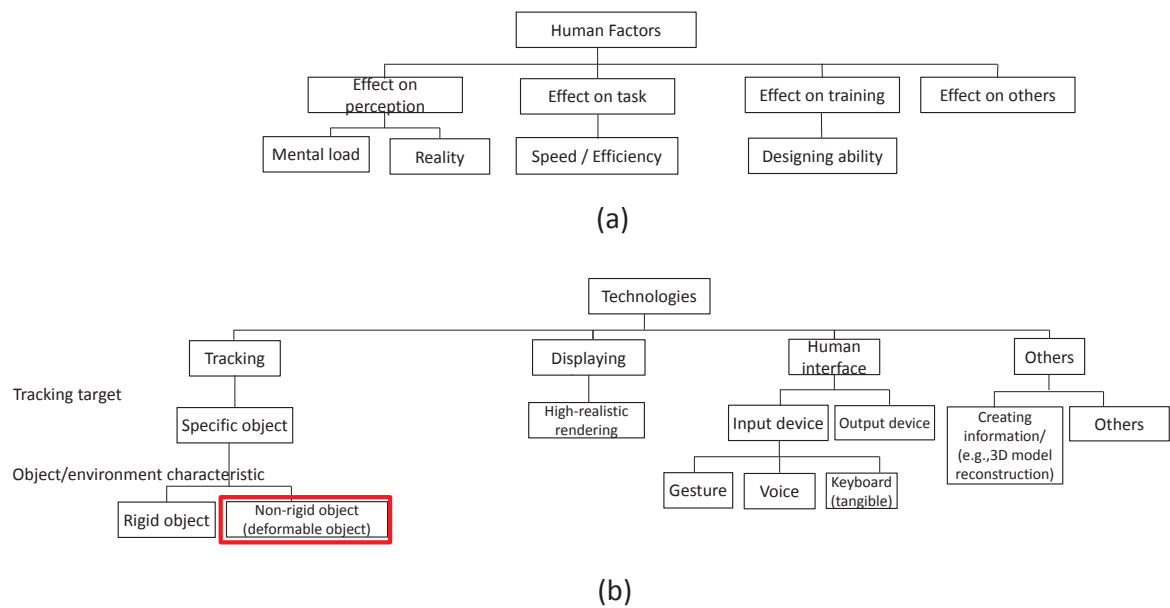


Figure 1.6. Composing factors of IAR when perceiving information itself is a task: (a) human factors (b) technologies

1.4.1 Verifying Effectiveness of AR for Memorization (Related to B-1)

Most research on the effectiveness of AR for task support has focused on how humans perceive information displayed by AR technologies. By contrast, we focus on the next step in said perception: ‘memorization’. As mentioned above, AR is a technology that overlays information about a specific location onto the user’s real-world view. Users may perceive not only the information, itself but also the relevant location. This contributes to the benefits of AR, which have been confirmed by several researchers. We know that humans can easily memorize and retain information if it is associated with real-world locations [32, 15]. Therefore, we focus on available relationship between the AR-mediated method of information display and users who perceive this information. We can take advantage of this relationship to help users memorize information more easily than in other information display methods. If we can find evidence supporting the effectiveness of AR in enhancing a user’s ability to memorize relevant information, we can claim that AR support systems are effective in providing relevant information to users regarding tasks as well as facilitating the memorization of these tasks. This would render IAR efficient not only for A-2-(b): “Manufacturing - Training” in Section 1.2.1 but also for most training tasks, and should have a significant positive influence on a number of tasks. Therefore, we seek to confirm this claim as the first objective of our research.

1.4.2 Developing Tracking Technique for Deformable Objects (Related to B-2)

In existing systems, a prototype of a product (e.g., a mobile phone, or a car) made from the white material is projected by a projector or overlaid through an HMD in order to change its appearance. There are two main types of systems in this design: projector-based and display-based (including HMDs and hand-held devices). In the design process conducted by an industrial designer, a projector-based system would be more suitable because such a design process places a high premium on careful examination of the texture and the functionality of a product, and designers would prefer to directly check the appearance of the products without using a display. This enables a user to check various appearances in a manner similar to observing the actual finished

product. By using such a system, even if various colors need to be checked in the design process, it is also possible to check various appearances of a prototype simply by changing the projected texture of the image without the need to remake it. However, there is considerable demand for the capability to quickly change not only the color or texture, but also the shape of a prototype in a short time. For example, if a user slightly changes the design of the prototype of a car, it is difficult to virtually repaint depending on the deformation in its shape, using existing technologies. In order to attain design support through AR, the target object should be observed through a camera or some other similar device. The position of the target object relative to the camera (or a projector) and the target object need to be calculated. Most existing technologies are intended for rigid objects, rather than deformable ones. The system obtain information in advance regarding the shape of the target object. Therefore, if the shape of the object changes in real-time, it becomes difficult to recognize its position and angle of rotation relative to the camera. Furthermore, the deformation in the object itself should also be calculated. If the target object is a piece of cloth, a sofa, or a cushion, a change in the the shape of the target object in the design process is inevitable. This makes it difficult to deal with such objects for design support using existing technologies. Although several studies focused on tracking deformable objects, their results are not directly applicable to our situation because of the following reasons. First, in a projection situation, because a target surface has to have an almost no-texture in the range of visible light, a tracking method that requires that objects to be tracked have rich textures cannot be used. Second, we assume that the complicated deformations, including large self-occlusions and foldings occur in the design process (Of course, the degree of deformation depends on the application). To the best of our knowledge, no study which fulfills these two properties exists. Therefore, we propose a design support system using projection-based AR for deformable objects, and develop technologies suited to realizing such a system as the second objective of this research.

1.5. Contributions

Thus far, we have classified AR task support systems into two types, described current problems for each type and clarified the scope of this thesis. The contributions of our research here are as follows:

1.5. Contributions

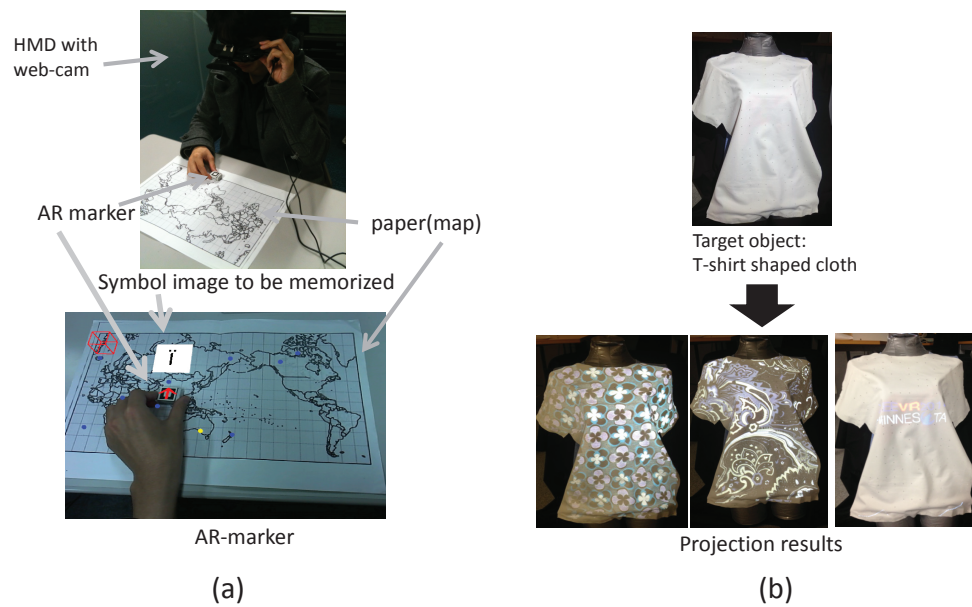


Figure 1.7. Overview of our research: (a) an example of experiments for evaluating relationship between AR and users' memorization (b) an example of our design support system for deformable object (in this case, clothes)

(1) Verification of effectiveness of AR for memorization

We verify a new benefit of AR in task support. We show that rendering AR features efficient allows a user to memorize the displayed information more effectively than otherwise. Three experiments show this benefit. In the first experiment, we acquired basic knowledge regarding the effectiveness of AR for memorization when subjects were instructed to memorize abstract images with and without the aid of AR features (Figure 1.7(a)). The two subsequent experiments targeted more practical situations: pick-up and assembly. To the best of our knowledge, no other study in the area has quantitatively confirmed the effectiveness of AR in this manner. Moreover, on the basis of experimental results, we provide some insights into the guidelines for the design of the method to display information obtained using AR for effective memorization. We think that these insights will help apprentices familiarize themselves with and memorize the task process more easily and intuitively in industrial tasks training.

(2) Construction of a design support system for a deformable object and development of techniques for identifying each location on the deformable object suitable for projection-based AR

We address a technical problem that inhibits the application of AR to task support. A projection-based texture augmentation of a 3D physical model is a promising method for design task support because users can intuitively check the various appearances of prototypes in an intuitive way. However, a deformable object, such as clothes, is not applicable as a target object of such a system because most existing technologies are intended for rigid objects. Moreover, deformable objects such as clothes may not only be deformed, but also even folded or cut, which makes it difficult to manage such objects in existing design support systems. In order to solve this problem, we develop a special marker pattern that can be partially recognized even if it is on a deformed surface, and developed a recognition method for it. The marker pattern and recognition method have the following three features.

- 1) Our marker is almost imperceptible to users, which is a crucial feature for projection scenarios. Our marker consists of various small points on a grid, and the ratio of the region occupied by points to the whole surface of the target is very small. Further, because it consists of only binary information (in other words, gradual information,

such as intensity and color, are not used), a half-transparent retro-reflective material, which makes it difficult to represent gradual information such as the intensity can be used to create this marker. These features render this marker almost imperceptible to users.

2) Because this marker can be partially recognized (which means that marker recognition does not require the optimization of the entire surface of the target object), unlike most existing methods, ours inherently allows complicated deformations that may occur in assumed applications, such as an object that is folded, cut, occluded by other objects, or self-occluded (such deformations do not affect the recognition result of other parts of the object).

3) Even if some parts of the pattern are not recognized correctly due to occlusions, it is comparatively easy to interpolate these by using 3D information because each dot is situated on a grid, and the useful equidistance property of two consecutive points can be used for the interpolation.

As far as we know, no other existing method boasts all three of the above mentioned features. Although none of the above ideas is novel, our system as a whole, including the marker, the recognition method, and the interpolation method is applicable to various applications, and is robust in projection scenarios involving deformable objects. By attaching this marker pattern to a deformable object, we can calculate correspondences between each location on the deformable object and each location on the texture image to be projected for the correct projection. We develop a system that quickly changes the appearance of deformable objects (Figure 1.7(b)) based on 3D shape measurement and marker pattern recognition.

1.6. Outline of Thesis

This thesis consists of five chapters.

Chapter 2 provides more detailed classifications by discussing related work in order to clarify the research scope we focus on.

Chapter 3 describes the first research objective: verification of effectiveness of AR for memorization through several user studies.

Chapter 4 focuses on the second research objective: construction of the design support system for a deformable object by projection-based AR and technical development to achieve it.

Chapter 1. Introduction

Chapter 5 summarizes this thesis, with discussions about the contributions, limitations and feasible future work.

CHAPTER 2

Background and Related Work

In this chapter, we discuss related research and provide detailed classifications of IAR systems in order to clarify the scope of our research. Based on the product life-cycles of industrial tasks mentioned in Section 1.2.1, we describe related research in Section 2.1. According to the classification of information characteristics regarding the tasks discussed in Section 1.2.2, we examine related work and clarify research objectives for our study in Section 2.2.

2.1. Classification of IAR by Product Life-Cycle

In this section, we classify industrial task supporting systems with AR in terms of product life-cycle. We refer to Fite-Georgel’s definition of industrial AR [19], whereby it consists of five parts, “Product Design”, “Manufacturing”, “Commissioning: Validation and Documentation”, “Inspection and Maintenance”, and “Decommissioning”.

2.1.1 Product Design

“Product Design” focuses on generating ideas to conceptualize to with regard to a tangible object. This process often requires communication among a designer, a manufacturer (all stakeholders), and the end users to evaluate ideas and prototypes. The target

Chapter 2. Background and Related Work

products include tableware, cars, home electronics, architectural materials, and factory.

It is common in the design process of many artifacts (for example, automobiles, and home electronics) to start with rough sketches and move toward some form of virtual representation, such as rendering in PhotoShop ¹. During this phase, a large number of ideas are generated and evaluated because these concepts are very quick to instantiate. The designers want to move as quickly as possible to “the 3D,” i.e., some form of physical representation. Prototype physical mock-ups are developed at this stage. These may start as quite simple shapes in order for the designer to gain an understanding of size and form. They also can become quite complex, such as a clay model of the dashboard of an automobile. They lack color or details outside their physical shape.

Surface appearance is an important aspect of the physical mock-up. Designers use paint and inks to color and texture their mock-ups. A drawback of this tactic is that in order to alter a design, either a separate mock-up must be constructed or the original mock-up must be re-painted. Several AR applications have been developed in order to solve this problem. Klinker et al. augmented a car mock-up with different optics to assess its appearance in situ [44] (Figure 2.1 (a)). This offered the possibility of navigating around an augmented mock-up. They emphasized the need to integrate AR into the design process.

Verlinden et al. [103] (Figure 2.1 (b)) developed the idea of *Augmented Prototyping* (AP). They employed *Spatial Augmented Reality* (SAR) to project images onto objects manufactured by standard rapid prototyping techniques. They found that SAR offered a tangible and social interface for designers. Their platform Workbench for AP focused on the early phase of the design process, where the promptness with which an impression of the product is generated is critical. Using this kind of system, designers and developers can change the appearance of the target object by . changing the projected light. This is expected to help reduce the design cost.

Most previously mentioned approaches are designed to work in a prepared environment in advance. Thomas et al. [96] lifted this constraint by proposing an HMD combined with a wearable computer to visualize design data on-site by aligning CAD data to the real world. Users could freely enter CAD data into the real-world and check whether the two matched.

A considerable amount of literature regarding factory planning with AR also exists.

¹<http://www.photoshop.com/>

2.1. Classification of IAR by Product Life-Cycle

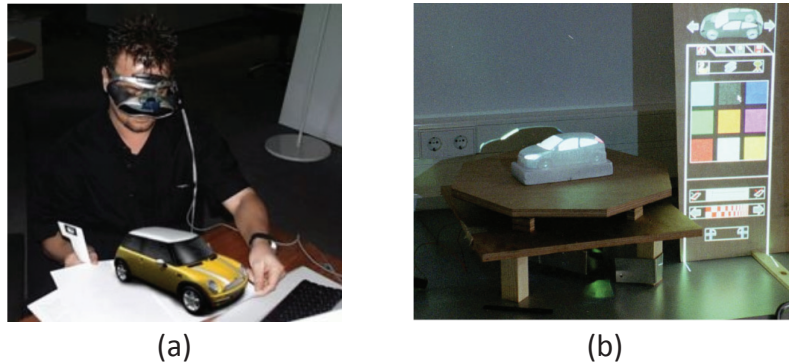


Figure 2.1. Examples of product design support systems by AR (a) a car design support with a HMD [44], (b) a car design support with a projector [103]

In industrial processes, considerable care is afforded to the factory setup when a new item needs to be manufactured. Such factory design can occur in new compounds or already existing production lines that need to be evaluated for their competence to produce the relevant new item. This process is called factory planning [19]. For example, to position components and systems in a new factory, Siltanen et al. proposed an iterative process where a plant operation requested an alteration [86]. This request was generated from the factory floor for the designer team. Using an augmented view of a current plant and a proposed design alteration, the operator could evaluate the proposal. This method allowed the plant operator to communicate from the factory floor. He or she could directly explain his or her requests to the designer and clearly describe problems found on site.

3D printers have lately become cheaper, and are employed in various fields. Due to their ease of use, they are also being applied to manufacturing prototypes of products. However, for the time being, even high-end models of 3D printers do not provide perfect coloring functions. Users thus have to color and texture products after printing them. Resch et al. proposed a system to augment a target object printed by a 3D printer through projected virtual replications using a projector-camera system [78]. They also proposed a tracking method for a target object printed by a 3D printer (this denotes that the system has complete information regarding a target shape) without the user's effort.

2.1.2 Manufacturing

“Manufacturing” is the act of transforming goods (raw material or manufactured) into a more complex product that is ultimately delivered to an end-user. This process mainly focuses on assembling objects. This task requires training of inexperienced workers.

Echtler et al. [13] (Figure 2.2 (a)) with their Intelligent Welding Gun introduced a new product and a related work-flow for the industry. The aim was to help welders shoot studs with high precision for an experimental prototype of cars for which robots could not be programmed because this would have required too much time. These prototypes were mainly hand built. A regular welding gun was tracked using external sensing devices, and is augmented with a display that helped the worker find designated stud locations. This new work-flow replaced a cumbersome procedure that required that the worker manipulate a probe sensor to find stud locations as he or she read from an instruction sheet, while a second worker marked the position and shot the stud. An AR-based setup is clearly more efficient because it only requires one operator. Echtler et al. showed that their setup was up to four times faster than usual work time while sustaining the same precision (This is one of the rare cases where the AR system has an obvious advantage). This AR project is one of the publicly known ones, and was deployed and used by a car manufacturer (BMW). And this system was improved and becoming a product [16]

In order to assemble a complicated object, workers first gather the specific parts to be used. This process is called the “picking-up task”. Grubert created an AR system supporting the picking-up task [24] (Figure 2.2(b)) . A worker intended to pick up specific objects from one of many drawers or shelves. He or she wore an HMD and checked a virtual funnel that showed a location of a drawer in the process through the HMD. They reported that compared with an existing task situation that used paper documentation, an AR-supported setup led to significantly more items being picked up and fewer erroneous pickups.

A few researchers constructed systems to train novice workers. “Training” here denotes a process to improve workers’ skills in manufacturing new items. For example, AR is used to support welders as well as train new welders, which is a complex procedure to learn [45]. Kobayashi et al. proposed an AR simulator that enabled learning to weld, which is a complex process that can be dangerous, in a safe and efficient environment. Moreover, the number of qualified teachers for such trades is limited. Their system used a display system and offered additional haptic and sound feedback.

2.1. Classification of IAR by Product Life-Cycle

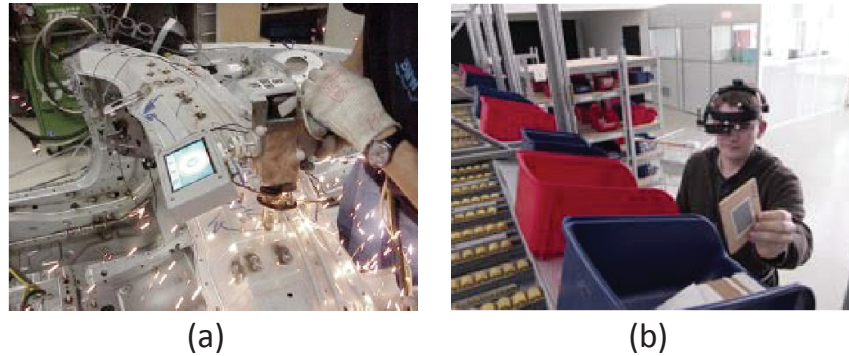


Figure 2.2. Examples of manufacturing support systems by AR (a) a welding task support with AR [13], (b) a picking-up task support with AR [24]

2.1.3 Commissioning: Validation and Documentation

“Commissioning” is the process of verifying that all systems and components of a product are installed, and are functioning as required by the client. This process consists of verifications of the product in light of the plan, the testing functionality, and document discrepancy, when required.

Georgel et al. proposed using AR to perform a discrepancy check assess the construction of a power plant [21] (Figure 2.3(a)). The idea underlying their work was to find differences between 3D CAD models and the corresponding plant that was built. In their system, high-resolution images were aligned with the CAD model of the plant. Workers could intuitively compare the CAD model with images of the actual site in order to recognize discrepancies. This concept has been commercialized by Siemens CT and tested by Areva NP.

The same concepts can be applied to ship building. Olbrich et al. proposed a system for to plan and install pipes in large ships [64]. In the manufacturing process, a wire model was produced according to ship geometry, and was used to check accurate fits for the pipeline. An operator manually adjusted this wire by bending it until it fitted. This modified wire was digitalized with a measurement system, and the resulting geometry data was used to bend a real pipe with a CNC bending machine. To speed up this process, they developed an AR system using which pipes could be virtually visualized, modified, and tested for fitting accuracy prior to the actual installation. The

Chapter 2. Background and Related Work

important features of their system was a measurement tool that could be used to automatically align a pipe along two 3D points. This could be used for the visualization of the differences between the virtual and real models as well as the modification and precise adaptation of virtual content into an existing, partially manufactured industrial object.

Petersen et al. proposed a robust real-time framework for segmenting video sequences and live-streams of manual workflows into the small tasks [67] (Figure 2.3(b)). Using classifiers trained on these segments, a worker could mimic an other worker performing the workflow in real-time and learn task variants from additional video examples. A novel application of this method is the automatic creation of step-by-step task documentation from a video demonstration. They concluded that their system can be used for various tasks to automatically create intuitive documentation.



Figure 2.3. Examples of commissioning: validation and documentation support systems by AR (a) a validation support with AR for a power plant [21], (b) a support for creating documentation with AR [67]

2.1.4 Inspection and Maintenance

“Inspection and Maintenance” is the action, sometimes regulated, to verify the condition of the product and if required, repair or replace fault components. Inspection and Maintenance typically correspond to codified and standardized procedures.

Considerable research has been conducted on the inspection and maintenance of systems using AR.

In case of a complex object, the user takes considerable time to familiarize himself or herself with the spatial and mechanical aspects of various components, their

2.1. Classification of IAR by Product Life-Cycle

operations, and the procedures. The practice of reading paper-based documentation to perform a complex maintenance task has a long tradition in the industry, even if it is not the most productive method. For example, aircraft maintenance evokes images of repair actions on actual hardware, but an airline spokesman reported that 45 % of every technician's shift was spent on finding and reading procedural and other information related to repairs [66].

One of the most famous studies among AR researchers working on maintenance task support is Feiner's example of a printer maintenance task, called "Knowledge-based AR for Maintenance Assistance" (KARMA) [17] (Figure 2.4(a)). Feiner et al. described the first maintenance and repair task supported by AR. The KARMA system graphically guided the user through the repair printer. The system automated the design of augmentations that explain how to perform a 3D task with a set of methods. Augmentation displayed through an HMD helped the worker localize and identify the action to be performed using highlights, labels, and animation.

Henderson et al. [26] proposed an AR system to assist military mechanics conduct routine maintenance tasks by enhancing localization through overlaid labels and context-setting 2D and 3D graphics. This system helped users understand where and how to intuitively use maintenance tools. This system has been shown to be very useful in complex and cramped environments, even for trained repairpersons, by limiting head movement and facilitating a quick repair. Following this study, they conducted another user study involving the psychomotor phase of a procedural task [27] (Figure 2.4(b)). They confirmed that AR is most effective for "align tasks," where an object needs to be correctly aligned to another.

The periodic maintenance of nuclear power plants is not a common task. Therefore, a worker may not remember all maintenance target locations and routes at a workplace. Moreover, nuclear power plants contain dangerous areas that workers should not approach without care. Ishii et al. proposed an AR-based system that avoids undesirable worker behavior and intuitively navigates workers to workplaces consisting of large spaces [34].

Zhu et al. proposed the "AR-Mentor," wearable real-time AR mentoring configured to assist maintenance and repair tasks on complex machinery, such as vehicles, appliances, and industrial machinery [110]. The system consisted of an Optical-see-through display device with high-precision pose tracking and a virtual personal assistant with natural language. The system facilitated verbal conversational interaction, and pro-

Chapter 2. Background and Related Work

vided guidance to the user in the form of visual, audio, and locational cues. Authors mentioned that the system acted as a personal mentor to users, providing human-like understanding and guidance.

In past research, a user was supported by information stored in the system. Another popular approach was to support the user by providing him or her access to a remote expert. The worker in charge of maintenance can generally handle repairs, but is sometimes unable to locate the problem, and can benefit from expert knowledge. In this case, the expert can find the problem and inform the worker of the solution. AR provides a promising technology for such applications. Reitmayr et al. proposed a system to allow an expert to annotate the 3D world [77]. A worker wore an HMD, and camera images seen by the worker were always transmitted to an expert in a remote location in real time in order for the latter to check the situation. They showed that their system could support the maintenance of a computer. The local geometry (e.g. the circle part) was estimated based on simultaneous localization and mapping (SLAM), and the annotations sketched by the expert were snapped on the geometry, thus allowing a precise description of the task to be performed.

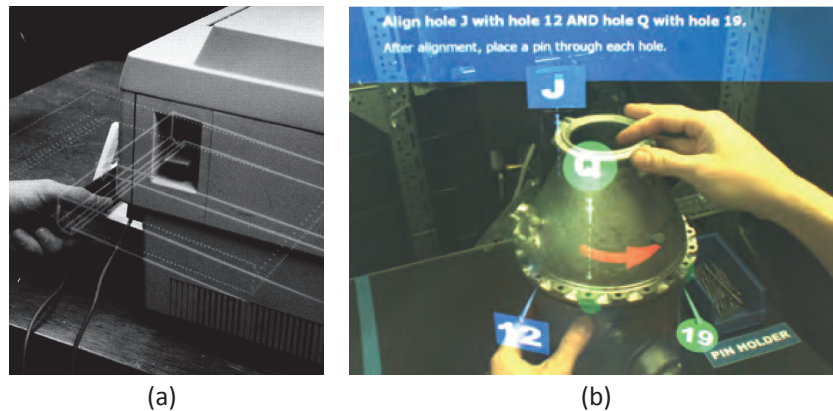


Figure 2.4. Examples of inspection and maintenance support systems by AR (a) a maintenance support with AR for a printer [17], (b) a maintenance support in the psychomotor phase in a procedural task by AR [27]

2.1.5 Decommissioning

“Decommissioning” is the act of retiring a product when it reaches the end of its life. This can include dismantlement, decontamination, and recycling.

Most important examples that require this process are from nuclear power plants. Such tasks are heavily regulated for obvious security and safety reasons. It first needs to be planned, and the feasibility of the process needs to be verified. Ishii et al. proposed a Dismantling Planning Support System (DPSS) based on AR [35] (Figure 2.5). DPSS operators could simulate the method of locating a scaffolding and temporary enclosures in a real dismantling field in order to determine their layout and predict the number of required parts. Its effectiveness and applicability to a real working field were evaluated using a survey involving staff of the Fugen Decommissioning Engineering Center.

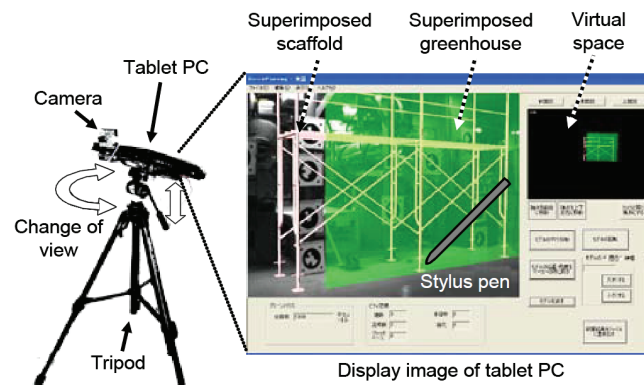


Figure 2.5. Examples of decommissioning support systems by AR for a nuclear power plant [35]

2.2. Classification of IAR by Information Characteristics

In the previous section, we described a lot of existing IAR systems based on product life-cycle. The detailed applications are diversified. However, if we classify the systems in terms of AR characteristics (in other words, how to use AR for task), they can be

classified into two main types: 1) cases when a user perceives relevant information to a task (when a virtual object is overlaid to indicate the task procedure), and 2) cases when perceiving information itself is a task (when a virtual object is overlaid as an alternative to a real object). Based on these two types, we describe current concerns and clarify the specific issues to be addressed in this study.

2.2.1 When a User Perceives Relevant Information about a Task

Most research on “Manufacturing”, “Commissioning: Validation and Documentation”, “Inspection and Maintenance”, “Decommissioning” as described in the previous section, fall in the first category, namely when user perceives information about a task. Much research has been conducted during the last few decades. Although these applications are diverse, their basic framework and objectives are similar: that is, AR is designed to replace existing ineffective methods such as paper documents and allow workers to understand tasks easily and intuitively. This requires that workers possess spatial cognitive abilities necessary to perform a task. AR has been shown to improve the task efficiency and reduce task performance time, mistakes, and worker cognitive load based on a worker’s intuitive understanding of task procedure. For example, Henderson et al. showed that workers can perform a procedural task faster and more accurately in situations employing AR than in those not employing AR (instead, an instruction is displayed on a standard LCD display) [27]. In addition, they also evaluated user subjective factors such as ease of system use, user satisfaction, and user intuitiveness. The study revealed that workers prefer to the situation employing AR from most of the point of view. Another example of this category involves the picking-up task when workers pick up specific items from one of many drawers. Grubert et al. reported that, compared with an existing task situation based on paper-documents, more items were picked up and few picking errors occurred when AR support was provided [24].

Most studies on the effectiveness of AR for task support focus on human perception of information displayed by AR. However, in order to increase AR effectiveness, we must attend to the next step in perception, ‘memorization’. As previously discussed, AR is a technology that overlays information about a specific location on a user’s real-world view. Users may perceive not only the information itself but also its location in the real world. Humans are known to memorize and retain information easily if the information is associated with real-world locations [15]. From these two facts, we hy-

2.2. Classification of IAR by Information Characteristics

pothesized that if we display information associated with specific real-world locations by means of AR, a user's memorization task results are improved compared to situations in which information is displayed at unrelated locations. If the effectiveness of AR in enhancing a user's memory can be proven, we can argue that AR support systems are effective for not only providing task-specific information but also facilitating the memorization of tasks. Helping an apprentice become familiar with and memorize a task procedure by actually performing the task in site is crucial to the performance of industrial tasks. Providing the previously stated hypothesis, one that has received little attention to date, is crucial to reducing the training cost of an apprentice and can lead to the practical application of AR in industrial work support. Therefore, we attempt to verify this point as a first objective in this study. Related studies are listed in Section 3.2.

2.2.2 When Perceiving Information Itself is a Task

Most studies discussed in the previous section concerning "Product Design" fall in the second category, namely when perceiving information itself is a task. These studies have a general framework as follows. When a company develop a product, the design process at first is mostly performed on a computer. However, evaluating the concept designs by means of a paper document or a computer display is not accurate enough in most cases [8]. This is because the process places a high value on careful examination of the texture and functionality of a product. In addition, the more the product development process advances, the higher is the cost of the initial design change. Confirming the design of the product during the early stages of development is crucial. Producing a prototype is often necessary for certain product designs in order to gauge the appearance of the product when compared to physical 3D model. However, this process is too expensive to be performed repeatedly. Therefore, to reduce the cost of producing prototype, employing AR and replacing physical 3D models are necessary. AR can enable a worker to conceptualize and understand a virtual 3D model in a manner similar to that of perceiving the actual physical product. No other technologies exist for this goal at the present. Therefore, we believe that applying AR to the design process would yield advantages and the priority in this application compared with other methods. Considering this point, target objects of current design support systems are limited to automobiles, home electronics, architectures, and others. A common feature

Chapter 2. Background and Related Work

of these conventional targets is that all can be regarded as rigid objects. This means that the shape of the target object is known in advance. Research on tracking with previously-computed information for a rigid object has shown that overlaying a texture image onto a real object with a high precision is possible. For instance, Resch et al. proposed a system to augment a target object printed by a 3D printer with projected virtual replications using a projector-camera system [78]. With this method, users can analyze the appearance of a target object having various designs in a short time. However, examples of industrial products that are easily deformed are seat cushions, beds and clothes. Compared with the use of a rigid object, using an easily deformed product as a target object in the design process is difficult. This is because the system should calculate not only its translation and rotation but also its shape every time it deforms during the design process. Most existing tracking technologies are designed for rigid objects. Although several recent studies attempted to solve this problem, a technology for uniquely identifying each location on the deformable object does not yet exist. Therefore, we conceived a design support system using AR for such deformable objects. Developed a practical technology is a second objective of this study. Related studies are listed in Section 4.2.

2.3. Summary

This chapter presented a summary of existing industrial task support systems that employ AR. Classification based on analysis of information characteristics was performed.

Although many studies have evaluated the effectiveness of AR systems, most those focus on the perception of information displayed by AR, task efficiency and others. We focused on the relationship between the memorization and AR to clarify a novel benefit for the training task. Several experiments conducted for this study are described in Chapter 3.

In addition, we examined on the lack of design support systems for deformable objects in certain industrial fields. To manage such target objects for projection, a technique to calculate their shape and identifying each location on it is essential. In Chapter 4, our technique for projecting deformable objects is presented.

CHAPTER 3

Verification of Effectiveness of AR for Memorization

This chapter presents research related to the learning and training of an industrial task through AR. In particular, we aim to verify the hypothesis: “If information is associated with the relevant locations in being projected onto the real world by AR, it would positively affect the user’s memorization task results over the case where the information is displayed at an unrelated location.” This hypothesis is derived from two features. The first is a feature of AR: AR can provide information associated with specific locations in the real world. The second is a feature of human memory, whereby humans can easily memorize information if the information in question is associated with specific locations. In Section 3.1, we describe the background of research in the area. Section 3.2 is devoted to an explanation of related studies on human memorization, and features of AR on human memory and cognition. We describe three experiments intended to verify the above hypothesis in Sections 3.3-3.5. In Section 3.6, we summarize the results of all the experiments and offer concluding remarks.

3.1. Introduction

AR has undergone continual development over the past decade and its technical foundation has hence improved to create practical systems that support user tasks. The

Chapter 3. Verification of Effectiveness of AR for Memorization

usefulness of these systems has been widely researched. [26, 77].

One of the most important merits of using AR is to be able to see information as virtual images associated with specific locations in the real world. This allows users to perceive an image as if it were a real object. Due to this feature, AR systems can assist users effectively understand the size, location, and rotation of objects in the real world. Henderson et al. [26] proposed an AR system to assist military mechanics conduct routine maintenance tasks by enhancing localization through overlaid labels, and context-setting 2D and 3D graphics. This system enabled users to understand where and how to intuitively use maintenance tools. Reitmayr et al. [77] proposed an AR collaborative task-supporting system for two users at distinct locations. In this system, the two users shared views using HMDs. User A could dynamically create CG arrows to point out the locations of specific objects to User B. Thus, User B could easily recognize the locations of objects.

These studies placed a high premium on the user's visibility, but considerable uncertainty still exists regarding the possibility of other fields. In this research, we focus on the effectiveness of AR in assisting the user's memory. "Memorization" means an action performed by the user to retain information having acquired information (by looking at objects, in this case). By using AR task-supporting systems, users can visually refer to the information required to perform tasks at any time. However, such support might induce user dependence on external memory (stored in a PC). In general, humans perform memory rehearsals for information required in tasks by autonomously recalling it during tasks. Thus, such information is gradually translated into long-term memory. Using AR task-supporting systems and passively referring to the required information may result in memory degradation.

At the same time, the features of AR may have a positive effect on the user's memorization ability. AR overlays information regarding a specific location on the user's real-world view. Users may perceive not only the information itself but also the relevant location. Thus, information displayed by AR naturally contains real-world locations. It is also known that humans can easily memorize and retain information if it is associated with real-world locations [15, 32].

Therefore, we focus on the possible relationship between the information, display method in AR and the users who perceive it. We might be able to take advantage of this relationship to help users memorize information more easily than through other display methods. We provide the following three hypotheses related to the effectiveness of the

features of AR to user memorization:

Hypothesis X

Displaying the relevant information near the location of a target object by AR enhances the efficiency of users' ability to memorize information in comparison with when it is shown at an unrelated location.

Hypothesis Y

Displaying the related information near the location of a target object through AR reduces the users' cognitive load on users in comparison with when it is shown at an unrelated location.

Hypothesis Z

In a task requiring spatial cognition, such as a 3D assembly tasks, performing observation task procedures from the first-person-view using AR increases users' memorization efficiency compared with situations where AR is not used.

Many systems use AR to display annotations associated with specific locations to provide context-sensitive assistance in tasks such as machine maintenance and object assembly. Such support is mainly intended for users performing unfamiliar tasks. Users can complete these tasks if the AR support system can be used at all times. However, it is desirable for users to be able to completely memorize and understand all tasks in a given process. As described in Chapter 2, industrial tasks include training to familiarize apprentices with work processes and help them memorize. If the effectiveness of AR in enhancing a user's memory can be shown, we can claim that AR support systems are effective not only in providing information related to tasks but also in facilitating the memorization of tasks.

3.2. Related Work and Research Scope

3.2.1 Human Memorization

In this section, we give basic knowledge about the human memorization and clarify the scope of this research. Human memory process consists of three main phases. The phase when human tries to encode and memorize information is called “Memorizing”. The phase when human retains information she or he memorized is called “Retention” and the phase when human tries to bring back information she or he retains is called “Recall”. Remembrance can be further divided into recognition and recall. Recognition involves judging only whether he or she knows the observed information. Recall involves remembering the information itself or related information with or without clues.

Human memory storage mainly consists of two parts, short-term memory (STM: working memory) and long-term memory (LTM). Information is initially acquired by humans from the external world via sensory organs. Most of the information is cleared out immediately, but some information that attracts attention is stored in the STM. The storage limit of the STM is 7 ± 2 items and the duration of memorizing is strictly limited, which means that information is not retained indefinitely. Some of it is transferred to the LTM when elaborative rehearsals such as repetitions are conducted. Compared with the STM, the LTM can store much larger quantities of information for a potentially unlimited duration.

The feature of human memory about both STM and LTM that information associated with specific location is easily memorized, is well known [15, 32].

Feature of the Memory Associated with Locations

We consider this feature from two points of view.

First is the existence of “place cell,” “cognitive map,” and “spatial memory.” The existence of these perspective was confirmed by experiments in cognitive science involving rodents. A place cell is a kind of pyramidal neuron within the hippocampus that becomes active when an animal enters a particular place in the environment, called the “place field.” A given place cell will only have one or a few place fields in a small environment but more in a larger region. There is no apparent topography of the pattern of place fields, unlike other brain areas, such as the visual cortex, - where a place cell is as likely to have distant fields as its neighbors. In this manner, as different combinations

3.2. Related Work and Research Scope

of pyramid cells are activated, depending on their location in the environment, numerous locations can be encoded by the combination of more than 100,000 pyramidal cells. The place cells are considered collectively, and provide the cognitive representation of a specific location in a space known as a “cognitive map”[63, 43]. A cognitive map retained in the hippocampus is called “spatial memory.” Although the existence of these perspectives has been confirmed in rodents, it is predicted that humans have the same ability because we have the same cells in our hippocampuses [14, 25].

Spatial memory works at the same time as other memories and reinforces the user’s memorization. Therefore, information at a specific location can to be memorized easily. Further, it is known that imaginary locations can be used as locations applied to spatial memory instead of actual locations [10]. The usefulness of this feature is evident from by the fact that “the method of loci” is widely used as a memorization technique [62]. This is a memorization enhancement technique where a person imagines a place with which he or she is familiar, and creates an imaginary space called a “mental map” or “memory palace.” If this feature is applied to AR, we can assume that a user can recall more easily the displayed objects (information) associated with specific locations in the real world than ones that are not.

The second factor is that when humans perceive their surroundings visually and memorize specific objects, the surroundings of that object may serve as a clue when they attempt to recall the objects later. When humans try to memorize an object, its surroundings are also memorized subconsciously. If its effect is sufficiently strong, such a memory is retained not as semantic memory but a kind of episodic memory. Semantic memory refers to the memory of meaning, understanding, and other concept-based knowledge unrelated to specific experiences. Episodic memory is the memory of events (times, places, associated emotions, and other contextual knowledge) that can be explicitly stated. Episodic memories can be memorized and recalled more easily than semantic memories [99]. This is one of other main reasons for why information associated with a specific location is easily memorized. In cognitive science, this feature is verified by an experiment called “object in place.” In an object in place test, a participant is instructed to look at and memorize a test image consisting of complex surroundings featuring a few objects. The subject is then instructed to look at two answer images containing a specific object from the test image, and to identify the image that contains the same object. By comparing the subject’s answer for images with and without (only an object) the surrounding image, Hollingworth showed that participants

Chapter 3. Verification of Effectiveness of AR for Memorization

could choose images more accurately when they saw the answer image in the same surroundings as the test image [28] (Figure 3.1).

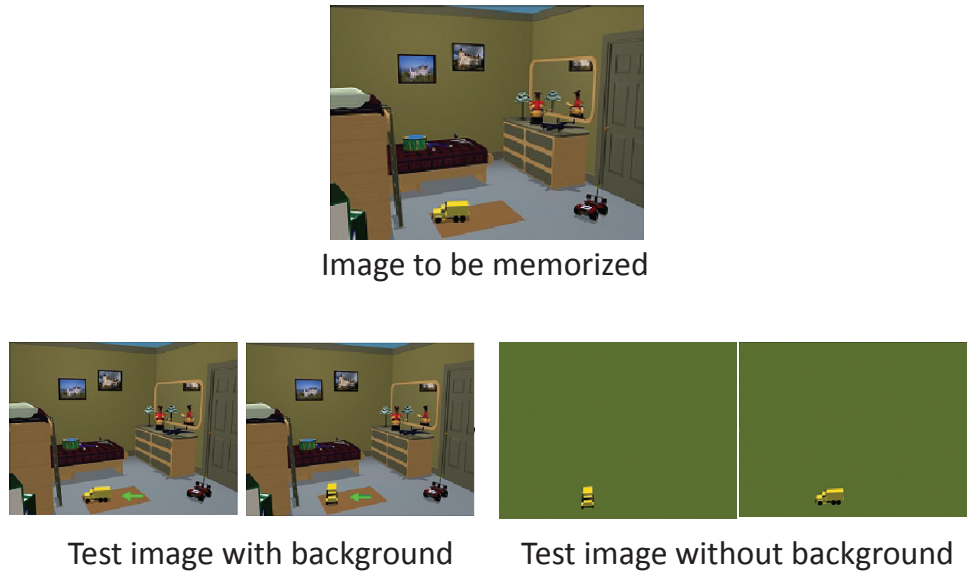


Figure 3.1. Example of “object in place” experiment [28]

This was also confirmed in another test, where the user was instructed to look at different images, with meaningless symbols in the background, in to search for a specific symbol. It was found that users’ responses were quicker if the same background was used [12]. This was not a memory test, however, as it was similar to the “object in place” experiment with regard to the fact that users subconsciously memorized the background, which serves as a clue to recall memories. Since background information differs depending on locations in the real world, memorizing background information results in the ease of recalling locations. If this feature can be applied to AR, we can assume that when objects (information) are shown at specific locations containing some background images, users might be able to easily recall the objects previously displayed by merely observing the same background on some level.

3.2.2 Features of AR Relevant to Human Memory and Cognition

In the fields of neuroscience and cognitive science, it has been found that spatial position is strongly associated with human working memory. On the other hand, one of the most important objectives of technologies such as AR or other wearable systems is to provide users with appropriate support in certain contexts from time to time. Hence, by definition, such technologies are positively compatible with users' tasks in a task space.

Tang et al. claims that features of AR have positive effects on block assembly tasks with regards to supporting human spatial cognitive ability [95] (Figure 3.2(a)). Similarly, Goto et al. studied how different display methods enhance users' spatial cognition in order to enable them to follow the working process when AR is used to support assembly tasks [22] (Figure 3.2(b)). Kirth et al. claimed that humans have their own imaginary maps called “mental maps,” and concluded that if information to be used often was located at the specific points around the body using AR, users could easily refer to the information [6].

Ikei et al. showed that overlaying symbols on pictures of specific locations in the real world through a wearable system enabled users to memorize these pictures and retain the information longer than usual [31]. Kawamura et al. proposed a method for associating information with real target objects and internalizing such information by touching the object in the real world using a wearable system [41].

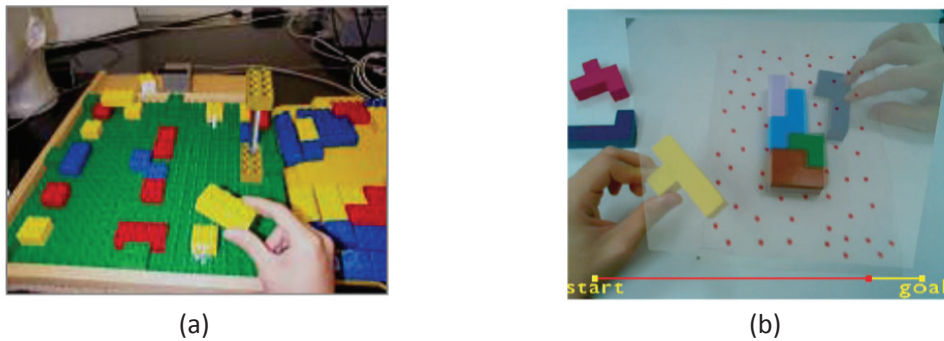


Figure 3.2. Example of research on the spatial cognition and AR: (a) Tang et al. [95]
(b) Goto et al. [22]

3.2.3 Objective of Study

A few studies have associated the fact that information displayed by AR includes spatial information with human spatial cognition or memorization. Others have connected target information with specific objects, thereby enabling users to memorize information as episodic memory by using wearable systems. However, the manner in which how such AR display methods affect user memorization, when the latter is instructed to memorize and recall information displayed by AR, remains to be analyzed. Therefore, we aim to investigate the relation between features of AR and human memorization. As mentioned in the previous section, human memory storage consists of two main parts: STM and LTM. By the way of a first step toward our objective, we focus on the verification for STM, which is relatively easy to be verified in this research.

3.2.4 Method to Verify Hypotheses

In Section 3.2.1, we described memorization using location association between two points of view. However, since each is not completely independent of the other and the two interact, this memorization cannot be clearly divided using these points of view. In this chapter, we classify this memorization to two memorizations for descriptive purposes.

One is the “Memorization with a location on the visual perception”. This is the procedure whereby when humans see at an object in the environment and perceive these as an image, they memorize the object and its location by taking advantage of the location onto the image (which is defined by humans’ field of view). The other is the “Memorization with a location on the spatial cognition.” This is the procedure whereby humans recognize one another’s relational positions using spatial cognition and an object in the environment by memorizing the object. It is assumed that “Memorization with a location on the visual perception” is mainly related to Hypotheses X and Y described in Section 3.1, and “memorization with a location on the spatial cognition” is related to Hypothesis Z.

Furthermore, we use two types of information as memorization targets with regard to Hypotheses X and Y. One is a character and image that is independent of each other, hereinafter called it “Individual factor.” This is defined as a memorization target that is easily recalled without information of the order when memorized. The other is a series of task processes with continuity, hereafter called “Continuous factor.” We assume

3.3. Experiment 1: Memorization with a Location on Visual Perception for Individual Factor

that it is difficult to recall this target at random, and that it is intended to be recalled in the same order as in the memorization step. A difference between these can be defined according to whether the information is dependent on the order of memorization. Considering the possibility that this difference might affect memorization results, we separately evaluated these for descriptive purposes.

With regard to “memorization with a location on the spatial cognition” concerning “Individual factor,” we do not consider this here because it is not related to the three hypotheses described in Section 3.1. Ikei et al.’s research [32] related to this classification. Note that these separations are not completely independent of each other and have mutual influences.

Using this classification, we verified each hypothesis through experiments in the following three sections. Section 3.3 describes an experiment related to “Memorization with a location on the visual perception” for “Individual factor” as verification of Hypotheses X and Y. Section 3.4 describes an experiment related to “Memorization with a location on the visual perception” for “Continuous factor” as verification of Hypotheses X and Y. Section 3.5 describes an experiment related to “Memorization with a location on the spatial cognition” for “Continuous factor” as verification of Hypothesis Z.

3.3. Experiment 1: Memorization with a Location on Visual Perception for Individual Factor

3.3.1 Experimental System

An experimental system consisted of a Head Mounted Display (HMD, video-see-through type, Vuzix VR920, resolution: 640×480) equipped with a single web-camera (Logitech HD Pro Webcam C910, resolution: 640×480), A3 sized paper (297 mm \times 420 mm) as a target object and an AR marker as an input device. The HMD provided 30 frames per second and 32° -field of view. Figure 3.3 shows the overview of the system for the experiment.

Users wore the HMD and all of the information was displayed through it. To conduct an experiment, we created a system that displays specific information associated with each location when users do the following simple operation in each of the locations of the paper. A3 sized map was used as the target on which information was put.

Chapter 3. Verification of Effectiveness of AR for Memorization

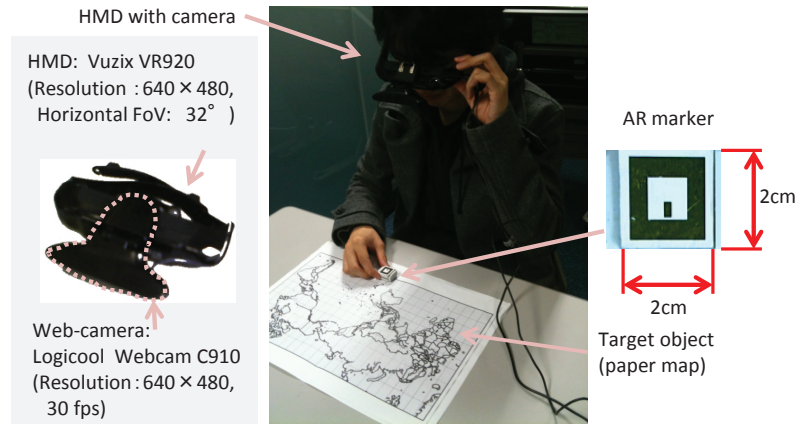


Figure 3.3. Overview of an AR annotation system

Since feature points of the paper were extracted and learned in advance [40], it enabled participants to recognize the position and rotation from the web-camera to the paper by capturing any part of the paper. In this experiment, since participants were intended to do the operations on top of the paper, part of the paper was assumed to become frequently occluded from the web-camera attached to the HMD. Because a prior learning method could maintain robust tracking in such a situation, it was decided to be used in this experiment. In the pilot study, five participants used the system with this tracking method and all commented that there was no tracking loss and they could concentrate on the tasks of the experiment.

Figure 3.4 shows view for participants. The red cube on the left upper side of the image is displayed to show participants that the tracking of the system works with enough accuracy.

CG points were displayed at 10 locations randomly chosen on the paper. When participants put an AR marker (20 mm×20 mm) near the location of each CG point, information (symbol image) is displayed on the specific location for 10 seconds. After each symbol image is displayed at the specific period, it is disappeared and can never be checked again. Participants were instructed to memorize “the symbol image itself” and the “location of the CG point associated with it” each the symbol image was displayed. When each point was selected once by participants, the color of that point was changed from blue to yellow. With it, the participants could distinguish CG points which had

3.3. Experiment 1: Memorization with a Location on Visual Perception for Individual Factor

not been selected yet from CG points which already had been selected.

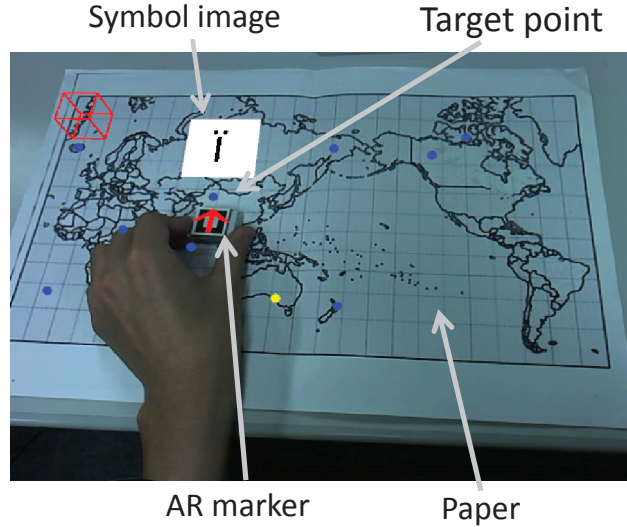


Figure 3.4. User's view through HMD.

This system enables displaying of information with three type of methods. Figure 3.5 illustrates example of three displaying methods. These images illustrate an approximate range participants can perceive through the HMD when they look at the center of the paper during the experiment. We evaluated how differences among those three displaying methods affected user's memory skills.

Type 1-1: The information is displayed near the location of each CG point on the paper

In this display method, the symbol images are displayed near the location of each associated CG point on the paper. This display method takes advantage of the AR feature which can associate information to a specific location on the real world.

Type 1-2: The information is displayed at random locations on the paper

In this display method, the symbol images are displayed on locations randomly chosen from 50 locations which decided in advance without relation to locations of CG points.

Chapter 3. Verification of Effectiveness of AR for Memorization

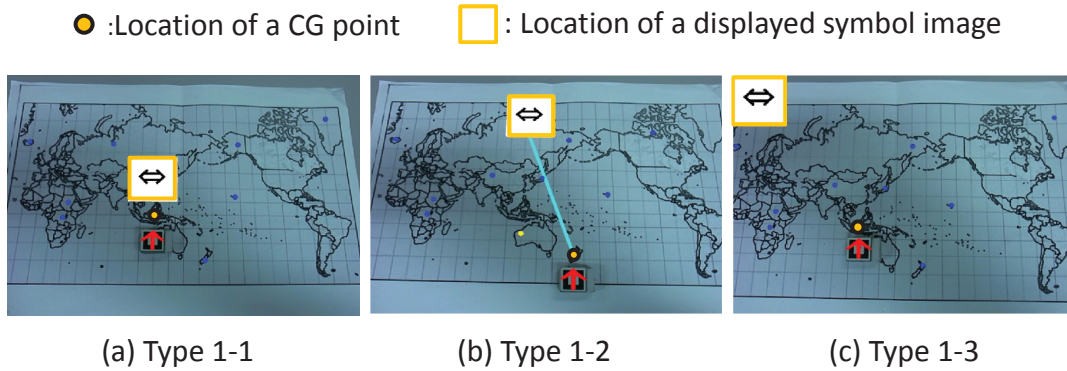


Figure 3.5. Display methods in the experiment 1: (a) Type 1-1: The information is displayed near the location of each CG point on the paper, (b) Type 1-2: The information is displayed at random locations on the paper, (c) Type 1-3: The information is displayed at the same location on the 2D display

If the display locations chosen randomly were too near to the selected CG points (this threshold value was set 120 mm from the result of the pilot user study), reselecting the location was conducted to distinguish between Type 2 and Type 1 methods. It was not intended that users had to take time to look for the location on which symbol images were displayed each time. Thus, a CG line connecting the symbol image and each CG point was also drawn to enable a participant to recognize the position of each symbol as soon as possible.

Type 1-3: The information is displayed at the same location on the 2D display

In this display method, AR is not used as the technology overlaying information on the real world. All symbol image is displayed at the same location on the 2D display without relation to each location of CG point.

3.3. Experiment 1: Memorization with a Location on Visual Perception for Individual Factor

3.3.2 Experimental Design

We had a total of 35 participants in the study, 21 to 35 years old (average 24.3), 31 males and four females. All participants had regular memory skill. 18 reported they had the experience of using similar AR annotation systems and 17 reported they had no previous experience. The research was piloted with five participants and was later conducted with 30 participants.

We used a within-subject design with a single variable (display methods). We recorded the accuracy rate in each test. In addition, participants were instructed to evaluate the degree of memorization difficulty by keeping a score. The order of the display methods was completely balanced for all possible orderings, with each possible ordering traversed to remove the order effect.

Our hypotheses about the experiment 1's outcome were as follows. Since Type 1-1 displays symbol images near the location associated with each CG point, it was assumed participants could memorize them as a 'single location-based image' including the background map image. We assumed that participants can easily memorize and recall information in particular about the its location with Type 1-1. Therefore we hypothesized X-1 X-2, and Y-1.

Hypothesis X-1

In the test of symbol images, a participant memorizes information more accurately with display method Type 1-1 than Type 1-2 and 1-3.

Hypothesis X-2

In the test of locations of symbol images, a participant memorizes information much more accurately with display method Type 1-1 than Type 1-2 and 1-3.

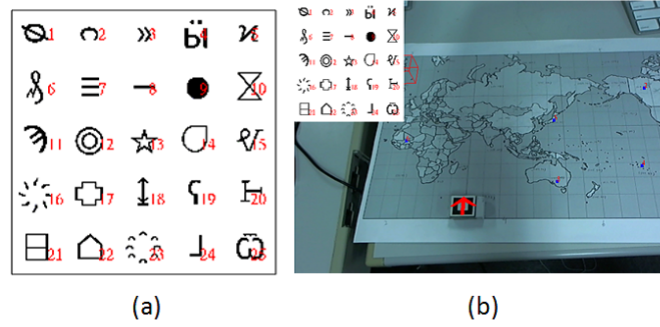
Hypothesis Y-1

A participant's cognitive load is lower with display method Type 1-1 than Type 1-2 and 1-3.

3.3.3 Experimental Procedure

At the beginning of the experiment session, the study administrator explained how to use the system, experimental procedure and how to response in the test sessions to the participant. The administrator was careful to only explain the individual features about each display method and not to recommended any particular strategy. Next, the participant was given the HMD and a three minute training session to get accustomed to the interface. When the administrator determined that the participant was familiar enough with the system, he or she was instructed to remove the HMD and take a break to concentrate on the experimental objectives. After this introduction part, the memorization test session started. 10 CG points were displayed on 10 locations of an A3 sized map of the real world. When the participant put the AR marker near each of the CG points, the symbol images associated with each CG point in advance, were displayed. After 10 seconds, symbol images disappeared. During the displaying of symbol images, the participant tried to memorize each symbol image and the location of the CG points associated with them as correctly as possible. In this study, the administrator did not give any instructions on how to memorize to the participant, instead let each participant find his or her own way of memorization. After the memorization session, the participant was instructed to remove the HMD and wait for 30 seconds for the short-term memory of the session to fade a bit. After the short clearing session, the participant was instructed to look at 25 numbered symbol image chart displayed on a PC monitor and was asked to answer which images were 10 images they had memorized (This is the ‘Symbol Image Test’). Next, the participant looked at the same 25 numbered symbol image chart, the 10 numbered CG points and the tests’ map image displayed on the PC monitor and was instructed to associate each symbol image with each CG point in the correct order (This is the ‘Location Test’). Figure 3.6 illustrates displays participants’ see in two tests. After each test, the participant filled out a brief questionnaire asking about his or her memorizing methods and ease of memorization. These procedures were repeated for all of the three different display methods. After the last display method was completed, the participant was asked to fill out a questionnaire asking which display method would be the best and participated in a brief interview about his or her impressions on the experiment.

3.3. Experiment 1: Memorization with a Location on Visual Perception for Individual Factor



(a)Symbol image test chart (b)View of location test

Figure 3.6. View of two type tests

3.3.4 Experimental Results and Consideration

We did not find any significant difference in learning or in display order effects, which implies that the training session was adequate. Figure 3.7(a) shows the accuracy rate of participants' answer in the Symbol Image Test and Figure 3.7(b) shows the accuracy rate of participants' answer in the Location Test with each display method. With an ANOVA, we found significant effect between display methods on the accuracy rate of the answer in the Symbol Image Test ($F(2,58)=4.82$, $p=0.0116$). The effect size (η^2) was 0.0973. Using Holm method to evaluate each displaying method, it was found that participants memorized symbol images more accurately with Type 1-1 than Type 1-2 and 1-3 (Type 1-1 ($M=0.990$, $SD=0.040$) > 1-2 ($M=0.950$, $SD=0.057$) = 1-3 ($M=0.957$, $SD=0.063$), $MSe=0.0029$, 5% level of significance). With a ANOVA, we found significant effect between display methods on the accuracy rate of the answer in the Location Test ($F(2,58)=6.02$, $p=0.0192$). The effect size (η^2) was 0.0869. Using Holm method to evaluate each displaying method, it was found that participants memorized the locations significantly more accurately with Type 1-1 than Type 1-2 and 1-3 (Type 1-1 ($M=0.873$, $SD=0.198$) > 1-2 ($M=0.703$, $SD=0.269$) = 1-3 ($M=0.750$, $SD=0.236$), $MSe=0.0385$, 5% level of significance).

Since the Symbol Image Test was the test where participants freely chose 10 images which they memorized from 25 images, this test was similar to the recognition memory test where participants were instructed to answer only whether they know specific objects or not [2]. This test was easier than other tests, so the results of each test in

Chapter 3. Verification of Effectiveness of AR for Memorization

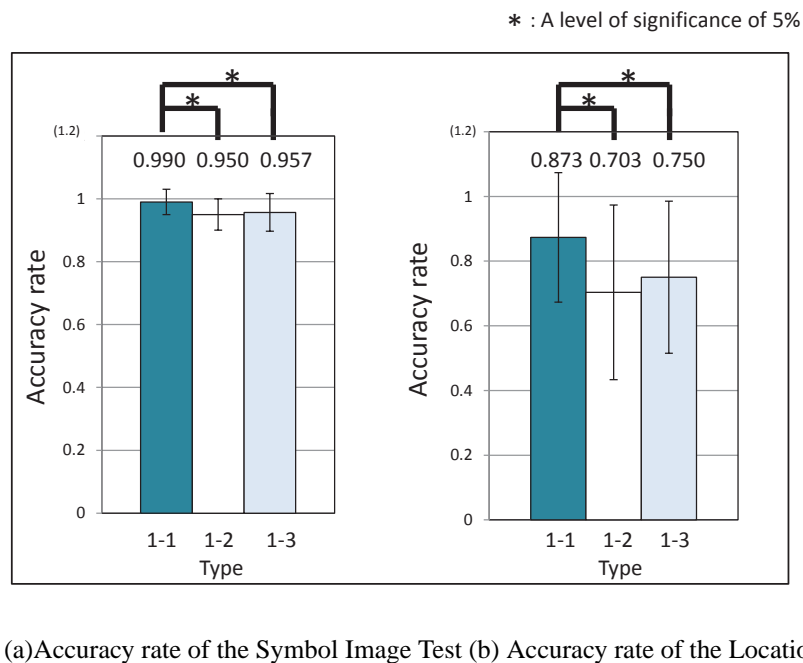


Figure 3.7. Rate of question answered correctly with each display methods in experiment 1 (mean with \pm standard deviation)

3.3. Experiment 1: Memorization with a Location on Visual Perception for Individual Factor

all display methods were very accurate and we found only slight differences with Type 1-1 and 1-2, and also between Type 1-1 and 1-3.

In the post-study interview, 10 participants confirmed their preference to Type 1 method with the comment that they could easily concentrate on the memorization task because they did not need to shift their attention away from the location of where the AR marker was already put. It is known that when a human memorize a object put in a specific location and its location in the environment, forcing him or her shifting his or her attention to non-relevant locations disturbs the memorization procedure [111]. It is assumed that effect of such disturbing is smaller in Type 1-1 than other displaying methods. Therefore, it can be considered that in the memorization task associating location of information with the target object in the real world would be useful, because this would save the participants' trouble of shifting their focus away from the object to be memorized. Thus Hypothesis X-1 was confirmed.

In the Location Test, the answering order was randomly set and the accuracy rates of this test were much lower than in the Symbol Image Test. Type 1-1 method achieved a much better score in the accuracy rate than other display methods. In the post-study questionnaire, memorization methods mainly seemed to consist of two types. One method was as follows. First, participants decided a starting point from the 10 CG points and the created a path for memorizing the other 10 symbol images associated with each CG point. After that, they memorized each symbol image using this predetermined path. When they tried to recall this information, they usually recalled each symbol images starting from the first one until they got to the requested symbol image. The other memorization method was in which participants visualized a square of space in their mind and allocated each symbol image to each location of that square. In addition, some participants commented that they used both methods combined.

With all memorization methods, the display method that showed information near the location of CG points, like in Type 1-1, achieved better scores than other display methods that showed information further away. As mentioned above, information displayed near the specific locations like Type 1-1 saved participants' trouble of shifting their attention. In addition to that, because participants could memorize their target CG points and the symbol images as one, the accompanying background map image could be the clue which reminded participants of the symbol images as mentioned in section 3.2.4. This is a possible explanation for why Type 1-1 achieved the best score. Thus, Hypothesis X-2 was confirmed. In the post-study questionnaires and interviews, seven

Chapter 3. Verification of Effectiveness of AR for Memorization

participants commented that they especially had difficulties memorizing the locations with Type 1-2. The reason for this might be that participants had to have an ‘extra step’ to recognize where to next shift their focus from their current location, as they could not know the location of the next symbol image in advance. This makes us believe that displaying information on unrelated locations partly disturbs humans’ memorization task and enlarge humans’ cognitive load.

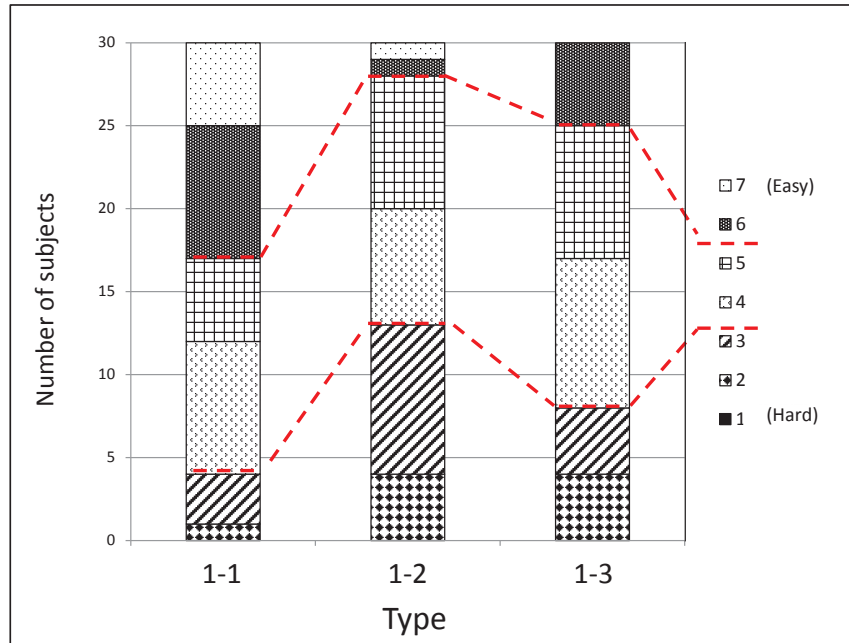


Figure 3.8. Number of each score which participants gave for each display method in experiment 1

Finally, in the subjective results, Figure 3.8 shows the scoring results about the easiness for each displaying method (1: Difficult, 7: Easy) by participants. Regarding scores as ranked values, we found significant effect between display methods with a Friedman test ($\chi^2 = 15.883, p < .01$). Using Scheffe method, results of multiple test were Type 1-1 > 1-2 ($S = 10.4727$, 5% level of significance), Type 1-1 > 1-3 ($S = 8.792$, 5% level of significance), Type 1-2 = 1-3 ($S = 0.07346$, n.s.). Thus Hypothesis Y-1 was also confirmed. Many participants ranked Type 1-2 lower than Type 1-1 and 1-3. One participant commented on being distracted by the display feature of Type 1-1: “because symbol images were displayed near each of the CG points in Type 1-1, I could

3.4. Experiment 2: Memorization with a Location on Visual Perception for Continuous Factor

not see the part of the background map during displaying of symbol information”. As a whole, however, the results indicated that the participants could memorize information most easily with Type 1-1.

As a result, we can conclude that displaying information associated with locations of specific objects by AR can increase the amount of information which users can memorize within a certain period of time. For that, we indicate the following conclusion. When a location on which information has to be associated, is important, displaying information associated to that specific location on the real world by AR is useful in terms of memorizing it.

3.4. Experiment 2: Memorization with a Location on Visual Perception for Continuous Factor

We confirm the relationship between the memorization with a location on the visual perception and features of AR in the case where the target information to be memorized is a continuous task procedure. In this experiment, participants were instructed to memorize a task like picking some blocks from each of nine drawers in the correct order.

3.4.1 Experimental System

The experimental device consists of an HMD (video-see-through type) equipped with a single web-camera. The HMD provides 640×480 images at 30 fps and 32° field of view. Participants wore the HMD and all the information was displayed using AR. All hardware used in this experiment is same as these used in the experiment 1.

Each of the nine drawers contains 10 blocks, as shown in Figure 3.9, and participants are instructed to pick those from the drawers in the correct order. The total size of all drawers is $330\text{mm} \times 330\text{mm} \times 140\text{mm}$ so that participants sitting on a chair can reach and open every drawers effortlessly. The drawers are labeled with AR markers and seals that distinguish each location, such as ‘L1’ (first one in the left column) or ‘M2’ (second one in the middle column). When the system recognizes one of these at least, a task process annotation is displayed for participants. 12 task processes are continuously displayed for participants in turn as one task process annotation, which indicates “how many blocks” are to be picked and “from which drawers” these are to

Chapter 3. Verification of Effectiveness of AR for Memorization

be picked. After displaying for 10 seconds, each step is automatically switched to the next step. Participants are instructed to memorize “how many blocks” are picked and “from which drawers” those are picked while each step annotation is displayed.



(a) Type 2-1: The information is displayed near the location of each drawer. (b) Type 2-2: The information is displayed at the same location at any time.

Figure 3.9. Display methods in the user study 2.

The information is displayed in the system by the following two types of display methods.

Type 2-1: The information is displayed near the location of each drawer.

Figure 3.9(a) shows this display method. In this display method, a CG drawer, an arrow, and an annotation that indicates the number of blocks to be picked are displayed near the location of each drawer. This display method exploits the AR feature, which can associate information with a specific location in the real world.

Type 2-2: The information is displayed at the same location at any time.

Figure 3.9(b) shows this display method. In this display method, an annotation is displayed at the same location without the relation to the location of each drawer. The location of each drawer is displayed as the labeling number marked on the drawers, e.g., ‘L1’ or ‘M2’. This display method simulates the situation in which target objects

3.4. Experiment 2: Memorization with a Location on Visual Perception for Continuous Factor

and information are memorized separately without using AR, such as the existing situation where a user memorizes the task process by referring to a paper list of the task process.

These two display methods have several different properties. As mentioned earlier, Type 2-1 takes advantages of AR features which can associate information with a specific location in the real world. On the other hand, Type 2-2 simulates the existing situation where a participant memorizes the task process by referring labeling numbers in a paper list. This experiment aims at investigating the effectiveness of AR for continuous task memorization by comparing situation with two different display methods.

3.4.2 Experimental Design

19 males and two females, 23 to 27 years old (average 24.8), participated in the experiment. The pilot experiment involved three participants, and was later conducted using 18 participants. All participants were experienced in the use of AR systems but no one was experienced in this experimental system. We used a within-subjects design for arranging the display methods (Type 2-1 and 2-2) within each participant. We recorded the accuracy rates of “drawer locations” and “number of blocks to be picked” and the ease of memorization for subjective evaluation. The order of the display methods was completely balanced for all possible orderings, with each possible ordering traversed to eliminate the order effect.

To calculate the accuracy rate of the drawer locations, true-false judgments are conducted for each step of 12 steps separately. The accuracy rate of the drawer locations is the value of the number of process that participants picked blocks from the correct drawer divided by the total number of tasks. In the same manner, the accuracy rate of the number of blocks is the value of the number of process that participants picked the correct number of blocks from the drawer divided by the total number of tasks.

Our hypotheses on experiment 2 outcome were as follows. Since Type 2-1 displays information near the location associated with each drawer, it was assumed that participants could memorize these as a “single image” including actual drawers as a background. In addition, since participants do not need to remember the locations of drawers in the form of cumbersome characters like L1 or M2, nor to recognize the relation of pairs between each character and the actual location of the drawer, it was assumed that they could remember the task process with a smaller number of memory

Chapter 3. Verification of Effectiveness of AR for Memorization

processes in Type 2-1. Then, we predicted that participants would memorize drawer locations and the number of blocks more accurately with the display method Type 2-1 than with Type 2-2, similar to the results of experiment 1 for the memorization target of “individual factor”. Therefore, we hypothesized Hypothesis X-3, X-4 and Y-2.

Hypothesis X-3

A participant memorizes drawer locations more accurately with display method Type 2-1 than with Type 2-2.

Hypothesis X-4

A participant memorizes numbers of blocks to be picked more accurately with display method Type 2-1 than with Type 2-2.

Hypothesis Y-2

A participant’s cognitive load is lower with display method Type 2-1 than Type 2-2.

3.4.3 Experiment procedure

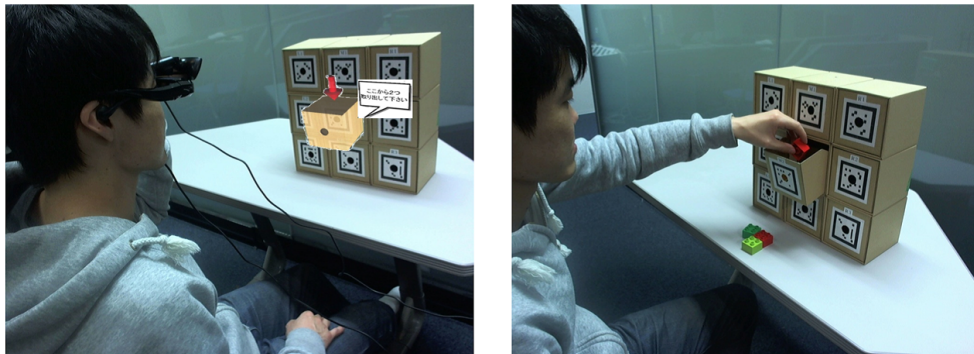


Figure 3.10. Overview of the experimental steps: (Left) The step in which a participant memorizes each task process. (Right) The step in which a participant recalls each process and picks some blocks from each drawer in the correct order.

3.4. Experiment 2: Memorization with a Location on Visual Perception for Continuous Factor

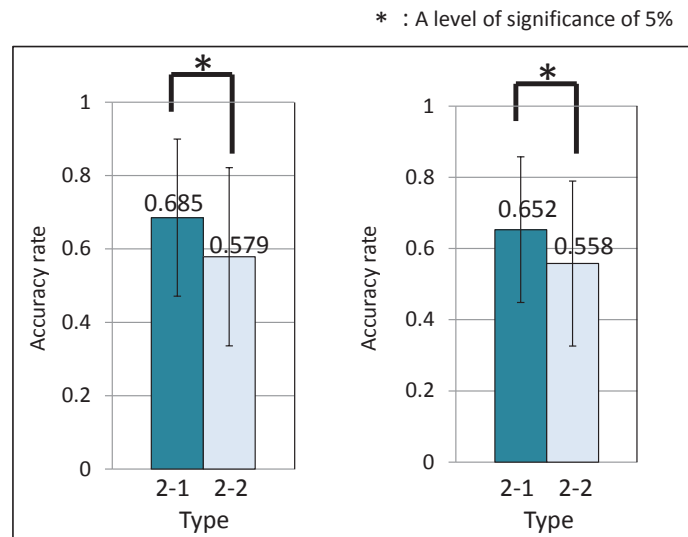
Figure 3.10 illustrates the experimental steps. The figure on the left shows the step in which a participant memorizes each task process in the correct order. The figure on the right shows the step in which a participant recalls each process and picks some blocks from each drawer in the correct order. At the beginning of the experimental session, the study administrator explained how to use the system, the experimental procedure, and how to respond in the test sessions. The administrator was careful to only explain the individual features of each display method and not to recommend any particular strategy. This was because an individual participant was intended to memorize information by the method that he or she thought to be effective. Participants were instructed to wear a HMD and to look at the drawers from the front. 12 task processes were displayed to the participants in turn continuously as one step that includes information indicating how many blocks are to be picked and from which drawers they are to be picked. After displaying for 10 s, each step automatically switched to the next step. Participants were instructed to memorize these task processes in order. After the memorization session, they were instructed to remove the HMD and to wait for 30 seconds to facilitate fading of the short-term memory of the session. After this clearing session, the participants were instructed to recall the task process and to pick blocks from each drawer in the correct order. In this experiment, the colors of blocks were not designated. These procedures were repeated for the two different display methods. As mentioned earlier, the order of the display methods was changed for each participant. After the last display method was completed, the participants filled out a brief questionnaire about their memorizing methods and ease of memorization.

3.4.4 Experimental Results and Consideration

Figure 3.11 shows the accuracy rate of the test answer for each display method. Figure(a) shows the accuracy of “from which drawers” and figure(b) shows the accuracy of “how many blocks”. With a paired t-test, it was found that the participants memorized “from which drawers” more accurately with Type 2-1 ($M=0.685$, $SD=0.215$) than with Type 2-2 ($M=0.579$, $SD=0.243$) (two-tailed test, $t(17)=2.29$, $p=0.0349$). The effect size (Cohen’s d) was 0.478. It was also found that the participants memorized “how many blocks” more accurately with Type 2-1 ($M=0.652$, $SD=0.205$) than with Type 2-2 ($M=0.558$, $SD=0.232$) (two-tailed test, $t(17)=2.82$, $p=0.0118$). The effect size (Cohen’s d) was 0.447. Thus, hypotheses X-3 and X-4 were confirmed.

Chapter 3. Verification of Effectiveness of AR for Memorization

The following comments were elicited from a participant: “In Type 2-1, I could intuitively memorize the task process without the slightly cumbersome transformation from characters of CG annotation to the actual location of the drawers”. In Type 2-2, some participants memorized the location of the drawers as character strings such as “M13L21...”. However, they memorized less accurately than participants who visually memorized the locations. From these results, it might be useful for memorization to overlay information on the target location intuitively, even if information can be expressed as character strings. This usefulness could be derived from the feature of AR directly overlaying information on the real world. However, a few participants commented that “CG such as drawers, arrows, and annotations, overlaid on the actual locations of drawers to be memorized, disturbs the memorization task”, which indicates the possibility that displaying a combination of many elements is visually intrusive and sometimes disturb the user memorization task on some level. Thus, it is more effective to display only simple annotations, which can be easily recognized visually, than to displaying a combination of many elements, in terms of a memorizing task process such as this experiment.



(a) The accuracy of “from which drawers” (b) The accuracy of “how many blocks”.

Figure 3.11. Rate of question answered correctly by display methods in experiment 2 (mean with \pm standard deviation)

3.4. Experiment 2: Memorization with a Location on Visual Perception for Continuous Factor

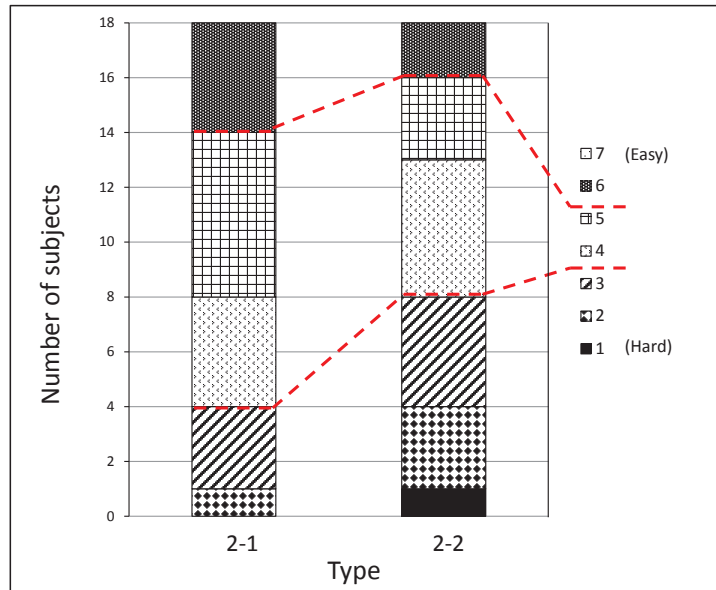


Figure 3.12. Number of each score which participants gave for each display method in experiment 2

Figure 3.12 shows the scores for ease of memorization of each display method that the participants gave (with easy=7, difficult=1). Many participants gave higher scores for ease of memorization with Type 2-1 than with Type 2-2, which indicates that they could memorize information more easily with Type 2-1. Regarding scores as ranked values, we found a significant effect between two display methods with a Wilcoxon rank-sum test ($T = 39.5, p < .05$). Thus, hypothesis Y-2 was confirmed.

3.5. Experiment 3: Memorization with a Location on Spatial Cognition for Continuous Factor

We evaluated the relationship between memorization given a location using spatial cognition and features of AR when the target information to be memorized was a Continuous Factor procedure. In this experiment, participants were instructed to memorize a block-assembling task.

3.5.1 Experiment System

Participants wore HMDs and all information was displayed using AR. All hardware used in this experiment was the same as that used in Experiments 1 and 2.

Ten AR markers were attached to a table on which participants were asked to assemble blocks. As long as the system recognized at least one of these markers, the information regarding the task procedure was continuously displayed. This information consisted of an animated 3D model of the blocks to illustrate what the function of each block in the step in question. A step consisted of an action where participants fit a specific block at a specific location. Information for 12 steps was continuously shown to the participants in order. The information for each step was shown for 10 seconds, following which information for the next step was displayed.

In this experiment, two displaying methods were used.

Type 3-1: 3D models are displayed fixed with the real world coordinate

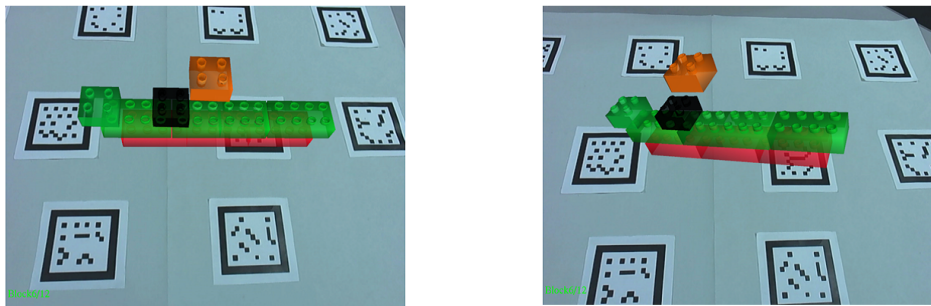
Figure 3.13(a) illustrates this displaying method. In this method, 3D models of blocks that had been assembled until a current step were displayed as fixed with real-world

3.5. Experiment 3: Memorization with a Location on Spatial Cognition for Continuous Factor

coordinates (markers coordinate onto the table). A 3D model of block that was used in a current step and its animation to illustrate the location where it should be assembled were also displayed. Participants could check blocks from various point of views by physically moving their body or head and changing their view. We assumed that this displaying method was similar to the situation where real block was put onto the table and participants observe it.

Type 3-2: 3D models are displayed fixed with the HMD display coordinate

Figure 3.13(b) illustrates this displaying method. In this method, 3D models of blocks that had been already assembled until a current step is displayed fixed onto the HMD display coordinate regardless of the participants' point of view. Animation that illustrates the location where the current block should be assembled was also displayed in the same manner as Type 3-1. Participants could freely rotate 3D model of blocks by dragging a computer mouse and could check it from arbitrary point of views.



(a) Type 3-1: 3D models are displayed fixed with the real world coordinate
(b) Type 3-2: 3D models are displayed fixed with the HMD display coordinate

Figure 3.13. Display methods in the experiment 3.

Two types of target assembling objects (Shapes A and B) were used. Shape A was ship-shaped (as shown in Figure 3.14(a)) and Shape B was a no-meaning-shape (as shown in Figure 3.14(b)). Both shapes consisted of the same number of blocks. We assumed that it was easier to assemble Shape A because it consisted of one-lined blocks, and hence participants did not need to consider depth perception during assembly. However, Shape B was complicated, and consisted of blocks of varying depth.

Chapter 3. Verification of Effectiveness of AR for Memorization

Moreover, in Shape B, following a few steps, it became difficult to recheck blocks in the preceding layers and it was difficult to imagine the completed shape of Shape B. For these reasons, Shape B was assumed to be a more difficult memorization target, than Shape A.

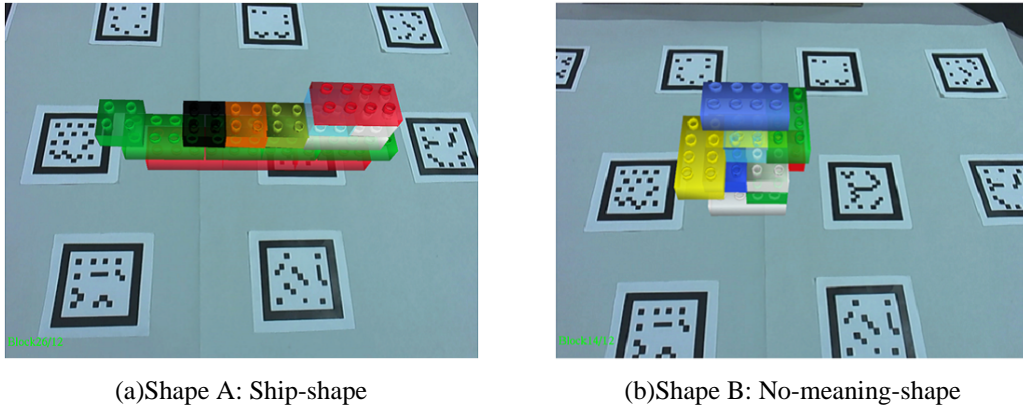


Figure 3.14. Assembly targets of the experiment 3.

3.5.2 Experimental Design

A total of 21 subjects, 19 males and two females, aged 23 to 27 years (average age, 24.8) participated in our study. The pilot study had involved three participants, and was later conducted using 18 more participants. Each participant performed a memorization test twice in two different situations that combined one of two display methods with the assembly of one of the two target shapes. For the two experiments, two different situations were assigned to each participant, without duplication, regarding to the display methods and the assembly of targets. Because we assumed that the differences in difficulty of memorization between the targets to be assembled was significant, we regarded this experiment as a between-subject-experiment in order to separately evaluate each situation involving different assembly targets. There were nine participants for each situation (in Type 3-X and Shape X). We recorded the accuracy of the calculation of “block location” and “the types of block (color and size)” when participants performed an assembling task by relying on their memory.

Our hypotheses regarding the outcome of Experiment 3 were as follows. In the situation where the 3D model was shown on fixed real-world coordinates such as Type

3.5. Experiment 3: Memorization with a Location on Spatial Cognition for Continuous Factor

3-1, participants had to move their head or body in order to check the blocks from different positions. Such observation requires that the participants possess the same spatial cognition that they would need while observing real blocks. We thus expected that the cognitive load on participants would be comparatively lower using this display method and it would be useful for the memorization. However, when the target object did not include blocks of varying depths, such as Shape A, such spatial cognition was not needed. There was no significant difference between display methods with Shape A. Therefore, we hypothesized Z-1 and Z-2:

Hypothesis Z-1

A participant memorizes block locations more accurately with display method Type 3-1 than with Type 3-2 in case an the target of assembly is Shape B.

Hypothesis Z-2

A participant memorizes the types of blocks to be assembled more accurately with display method Type 3-1 than with Type 3-2 in case the target of assembly is Shape B.

The experimental procedure and method to calculate the accuracy rate was identical to Experiment 2.

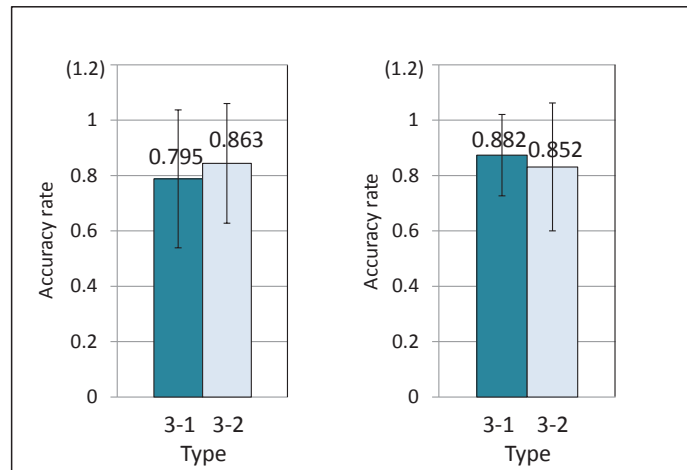
3.5.3 Experimental Results and Consideration

Figure 3.15 shows the accuracy rate of the test answer for each display method with Shape A. Figure 3.16 shows the accuracy rate with Shape B. Each figure (a) shows the accuracy rate of “block locations”. Each figure (b) shows the accuracy rate of “types of blocks (color and size).” We conducted the F-test for the results for “block locations” in situations involving each display method with Shape A, and confirmed that the two variances were statistically identical. With the t-test, there was no significant difference between the results of Type 3-1 ($M=0.795$, $SD=0.249$) and those of Type 3-2 ($M=0.864$, $SD=0.216$) (two-tailed test, $t(16)=0.611$, $p=0.549$). The effect size (Cohen’s d) was 0.307. We conducted the F-test for the results for “types of blocks (a color and a size)” in situations involving each display method and Shape A but could not confirm that the two variance value were is statistically identical. Then with a Welch’s test, there was no significant difference between the results with Type 3-1 ($M=0.882$, $SD=0.161$) and

Chapter 3. Verification of Effectiveness of AR for Memorization

those of Type 3-2 ($M=0.852$, $SD=0.231$) (two-tailed test, $t(16)=0.306$, $p=0.764$). The effect size (Cohen's d) was 0.161.

In the same manner, we conducted the F-test for the results for “block locations” in situations involving each display method and Shape B, and confirmed that the two variance values were statistically identical. With a t-test, there was no significant difference between results of Type 3-1 ($M=0.585$, $SD=0.270$) and those of Type 3-2 ($M=0.468$, $SD=0.297$) (two-tailed test, $t(16)=0.864$, $p=0.400$). The effect size (Cohen's d) was 0.435. We conducted the F-test for the results for “block locations” in situations with involving each display method with Shape B, but could not confirm that the two variance values were statistically identical. Then with a Welch's test, there was no significant difference between results of Type 3-1 ($M=0.693$, $SD=0.292$) and those of Type 3-2 ($M=0.491$, $SD=0.227$) (two-tailed test, $t(16)=1.612$, $p=0.131$). The effect size (Cohen's d) was 0.835. From these results, we could not verify Hypotheses Z-1 and Z-2.

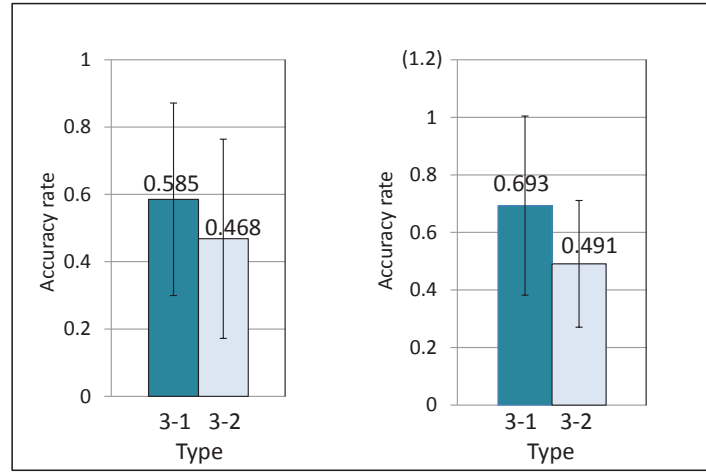


(a)The accuracy rate of block locations (b)The accuracy rate of the types of blocks

Figure 3.15. Rate of question answered correctly by display methods in shape A of experiment 3 (mean with \pm standard deviation)

As expected, there was no difference between the two display methods for Shape A. Furthermore, there was no difference between the two display methods for Shape B, despite the fact that the average value of accuracy rate for Type 3-1 was greater than that for Type 3-2. This was because the variance among each participants was quite

3.5. Experiment 3: Memorization with a Location on Spatial Cognition for Continuous Factor



(a)The accuracy rate of block locations (b)The accuracy rate of the types of blocks

Figure 3.16. Rate of question answered correctly by display methods in shape B of experiment 3 (mean with \pm standard deviation)

large. As mentioned above, Shape A was easier to assemble because it consisted of one-lined blocks and participants do not need consider depth perception. By contrast, Shape B consisted of blocks of varying depth. Shape B can be considered more difficult to memorize due to the complexity of the spatial cognitive reasoning involved. As we expected, when participants memorized each step while assembling Shape B, the number of times they moved their bodies or heads to see the information displayed was much greater than in situations involving the assembly of Shape A. Two participants commented that such movements were cumbersome and their results for Shape B were below average. Thus, the display method with fixed real-world coordinates was not effective for all participants.

However, five participants positively commented on Type 3-1, saying as if they could intuitively observe the blocks from a first-person-view of the other expert who assembled blocks perfectly. Although statistically significant results were not obtained, they did suggest that displaying the 3D model fixed with real-world coordinates might be useful for memorization using spatial cognition because users can perceive the target object in a similar manner as when they perceive an object in the real world. However,

the results also indicated that some users feel burdened by cognitive burdens similar to those incurred by the real-world spatial reasoning of this sort because it required frequent head and body movements. Therefore, we conclude that it is desirable to combine Type 3-1 and Type 3-2 and switch between them as needed in actual task supporting systems.

3.6. Summary and Future Work

In this chapter, we classified memorization into two types for descriptive purposes: “Memorization with a location on the visual perception” and “Memorization with a location on the spatial cognition.” We verified three hypotheses regarding situations where humans actively memorize information displayed using AR through three experiments. The results were as follows:

Hypothesis X

Displaying the relevant information near the location of a target object by AR enhances the efficiency of users’ ability to memorize information in comparison with when it is shown at an unrelated location.

- For Hypotheses X-1 and X-2, significant differences were found in the case involving individual factors.
- For Hypotheses X-3 and X-4, significant differences were found in the case involving continuous task procedures.

Hypothesis Y

Displaying the related information near the location of a target object through AR reduces the users’ cognitive load on users in comparison with when it is shown at an unrelated location.

- For Hypothesis Y-1, a significant difference was found in the case for individual factors.
- For Hypothesis Y-2, a significant difference was found in the case for continuous task procedures.

Hypothesis Z

In a task requiring spatial cognition, such as a 3D assembly tasks, performing observation task procedures from the first-person-view using AR increases users' memorization efficiency compared with situations where AR is not used.

- For Hypotheses Z-1 and Z-2, significant differences were not found.
- In parts of the questionnaires for Experiment 3, a tendency that would indicate the correctness of this hypothesis was found among the responses.

3.6.1 Insight Regarding Information Display Methods Using AR for Effective Memorization

It can be concluded that the experimental results indicate a design guidelines for information display using AR to enable users to effectively memorize information. The examples are as follows:

Display information relevant to an object near it (so that a user can see both the information and the object without significantly moving of his or her eyes)

Experiments 1 and 2 indicated this guideline. If information relevant to an object (e.g., text information regarding how to operate a system) is displayed using AR, the information should be displayed near the object by considering the observing point and field of view of users. This is because a user can then memorize relevant information and the location of the object as one image visually and effectively by looking at simultaneously at both.

Do not display information at random locations (ones a user cannot predict)

Experiment 1 indicated this guideline. Even if the information is not related to any objects that a user sees, it should not be displayed at random locations. This is because this would disturb user memorization if he or she cannot predict the location at which information is shown.

Display minimum required information to explain task process

Experiment 2 indicated this guideline. The displayed information should have sufficient detail to clarify the task process for a user, but this should not be excessive. If the displayed information consists of unnecessary elements (e.g., visual effects, text, 3D models, animations, movies, etc.), it fatigues a user and affects memorization.

Provide two types of display methods and a function to switch between them at any time

Experiment 3 indicated this guideline. In particular, in memorization tasks requiring spatial cognition, each user's method of memorizing information varies. A suitable display method for each user would thus be different. Moreover, even if a user felt that learning using the first-person view was helpful, frequent movements would distress him or her. Therefore, providing two display methods, such as Method 3-1 and 3-2, and a function that enables the user to switch between these at any time is recommended.

Moreover, as described in Section 3.1, using AR task-supporting systems and referring passively to the required information may result in memory degradation. In order to prevent this, it is thus necessary to devise a method to encourage users to actively memorize the displayed information when AR is used for task support.

3.6.2 Future Work

In this study, all our experiments involved situations where participants sat on chairs and performed tasks at tables without producing expansive bodily movements. This indicates that the participants did not walk freely in the experimental environment and observe the target object from the back. This is limitation considering the various tasks that workers need to move to carry out in the real world. Future studies involving the same type of experiments should be conducted in larger settings where participants can freely move. This will require that participants possess a different degree of spatial cognition which, we believe, will allow us to further clarify the relationship between AR features and human memorization.

CHAPTER 4

Design Support System for Deformable Objects Using Projection-based AR

This chapter provides an overview of research related to the design of task supporting systems using AR. Section 2.2.2 describes issues with existing technologies and systems. We address one of these: that “existing systems and technologies cannot deal with a deformable object.” In Section 4.1, we provide the background. Section 4.2 is devoted to an overview of related research in the area. In Sections 4.3 and 4.4, we detail the system components and the algorithm for our proposed method. Section 4.5 provides an assessment of the system. In Section 4.6, we explain feasible applications of this technology, followed by an explanation of a constructed application. Section 4.7 is devoted to the the explanation of improvement of the processing speed of our system. Finally, we discuss the design of our new technology, outstanding problems with it, and our plans for future research in the area.

4.1. Introduction

This chapter presents a new optical processing technique that registers digital projected imagery onto deformable substrates. The novel contribution is the ability to change the physical shape of the deformable substrate through twisting, bending, folding and

Chapter 4. Design Support System for Deformable Objects Using Projection-based AR

recognizing the new form. A substrate material constructed from silicon rubber and aluminum mesh with embedded retro-reflective dots is presented to demonstrate the reconstruction algorithm. The material can also be wrapped around a variety of organic shapes such as, spheres, cubes and complex geometric shapes. The wrapped surface may then be augmented with registered digital imagery.

A fundamental challenge of Augmented Reality (AR) research is the registration of digital information onto physical objects [4]. The majority of surface capture techniques and tracking systems assume the physical objects have a static shape and size. Deformable objects represent a new challenge of how to capture a changing form to support matching digital imagery on the deformable substrate.

Computer vision techniques have made considerable progress in helping the registration problem in AR [39, 108]. Deformable objects are particularly challenging for computer vision, as a non-planar substrate requires more information to be gathered to understand the shape. Additionally the feature points do not remain in a fixed correspondence with other feature points, instead they need to be re-calculated whenever new deformations are made. Previous tracking algorithms [69, 101] employ time-series data, and the processing for each camera frame is related with previous frames. In these previous algorithms, curve-fitting calculations are made to determine the shape of the deformable objects. A major contribution of this work is our solution that directly measures and/or interpolates the individual points on the surface of the deformable object, thus allowing the determination of the local position of a point on the target object for the projection of digital information. We refer to this as a “local position estimation algorithm”. Our solution additionally determines the position and orientation of the target object.

Once the technical challenges of implementing a local position estimation algorithm are solved, a number of novel virtual environment interactions are possible. Deformable materials that are recognized by a system can be employed for creating Deformable User Interfaces (DUI) [29]. DUIs provide a unique set of interactive capabilities; they can intentionally be squashed, twisted, bent and distorted to provide a fulfilling and powerful human-computer interaction experience [48]. The feel of the device and its dynamic shape allow physical world interactions such as fingertip sculpting and gestures to be used for system design. Rather than using materials with a sterile hard plastic feel, DUIs use soft smart materials with the goal to enhance the interface functionality and improve the user experience.

We believe there are uses in domains such as design [51, 71, 98], entertainment [97] and training [76]. The design domain demonstrates some compelling uses. For example, the iterative process employed when constructing a new car body form. Currently a combination of CAD modeling and physical clay prototypes are used to demonstrate new design concepts. This is a tedious and costly process that restricts the number of designs that can be explored. Once the clay models have been created, details such as color or textures may be added via painting. If the shape of the clay model is modified (a change in shape or size), the clay model has to be repainted. Our new material makes it possible for substrate alterations to be easily made and enhanced with projected imagery, and when the model is modified, the details remain attached to the model. This is particularly well suited to Spatial Augmented Reality applications to support new forms of interactions through physical touch.

Spatial Augmented Reality (SAR) is a specialized form of Augmented Reality that uses projectors as the primary display device [5, 74]. Physical models (props) are illuminated by projectors allowing the system to simulate different surface properties. In our system we will project against deformable objects with a neutral surface color. SAR offers an intuitive approach to AR for application domains such as design. An elegant feature of SAR is it does not require the user to wear or operate any form of display device, such as a head-mounted display or a hand-held device. Additionally, this AR approach scales well to multiple users [57], such as a meeting, and offers affordance with the physical presence of the props further increasing the realism of the virtual display.

The system presented in this work employs both a projector and camera (pro-cam) that simultaneously estimate a deformable objects pose (position and orientation) and changes in shape. The current method of tracking the pose in SAR is performed by a standard tracking system, such as an OptiTrack by NaturalPoint Inc.¹, and our new method incorporates this ability without the need of specialized 6DOF tracking technology. The system employs a natural light projector and an IR based computer vision fiducial marker tracking system. This IR tracking system allows for imperceptible features to be added to the projector substrate. We extend the concepts of the Szentandrási et al. marker system [92], that allows partially visible markers to be recognized.

The contributions of our research here are as follows:

¹<http://www.naturalpoint.com/optitrack/>

1. Construction of a system to correctly project a texture image onto a deformable object intended for use in design support

As an object is deformed, the mapping of the digital textures to exact positions on the physical object is re-calculated. For example, an object may be digitally painted, and then bent in half. Not only does the digital paint maintain its registration, but the object may have additional digital paint repeatedly added or removed. For example, clothing design can be considered an application of the system. Most existing systems, such as virtual try-on systems, use a display to show a user clothes. However, our system allows a user to perceive the target object with various appearances directly by adapting a projector instead of the display. It is suitable for a situation where designer cuts or attaches several pieces of cloth with each other in the design process.

2. Development of a partially recognizable, nearly invisible marker pattern suitable for projection scenarios

In order to implement a system of the sort mentioned above, we developed a novel marker pattern with the following features:

1) Our marker is almost imperceptible to users, which is a crucial feature for the projection scenarios. The marker consists of various small points on grid, and the region occupied by them is a small part of the entire surface of the target. Furthermore, because it consists of only binary information (in other words, gradual information, such as that regarding intensity and color is not used), a half-transparent retro-reflective material (on which it is difficult to represent gradual information) can be used to create this marker. These features render this marker almost imperceptible to users.

2) Because our marker is partially recognizable (this means that the marker recognition algorithm does not require the optimization such as a single cost function representing the entire object), it inherently allows complicated deformations that may occur in assumed applications, such as objects that are folded, cut, occluded by other objects, and self-occluded.

3) Even if some parts of the pattern are not recognized correctly due to occlusions, it is easy to interpolate these by using 3D information because each dot is placed on the grid, and the equidistance property of two consecutive points can be used.

To the best of our knowledge, no other application-centric technology exist with

these three features. This chapter begins with an overview of important related work in deformable surfaces and computer vision techniques for deformable surfaces. The major system components are then described to provide context. The algorithmic details of a single pro-cam arrangement are presented, followed by a description of the required extensions to a multi-pro-cam arrangement. A set of evaluation findings is then reported to show the feasibility of our new approach. Finally, we discuss our new technology in design as a motivating example, followed by some concluding remarks.

4.2. Related Work

Three areas of related work are discussed that are relevant to this study. Firstly we discuss structured light systems that capture physical object data. Following this we discuss vision based systems that have been developed to capture deformable substrates. Finally deformable materials that provide real-time surface geometric information through sensors embedded in the substrate material are summarized.

4.2.1 Structured Light Approach

Existing technologies such as the Microsoft Kinect and similar methods [72, 58, 104, 79] use structured light to capture depth information. Microsoft Kinect employs multiple techniques including parallax and multi-focal lenses across the X and Y planes to calculate the depth array in conjunction with a traditional RGB camera for vision techniques. While these methods can make correspondences between the light emitter and a camera to calculate 3D information, they do not recognize specific locations on the target object. Post processing of the depth information can identify 3D shape features however specific details are still unknown. For example, with a white cube such an RGBD camera could not uniquely identify each of the six faces. A major aim of our new technique is to project a texture onto specific locations of a deformable object where each part of the texture is associated with the physical object. The projected image should appear to be printed on the surface of the deformable in any orientation or deformed state. Our new approach is to identify each local location on the deformable object with dot pattern markers that provide unique identification of regions on the surface.

4.2.2 Vision Based Recognition of Deformable Substrates

A number of systems have successfully identified changes in deformable substrates using vision based methods with substrate materials including silicon rubber and paper. Milczynski et al. [56] developed a deformable silicon surface with black dots that is stretched over a circular hole in a planar surface with a camera under the silicon surface to capture how the dots change when deformed. The surface reconstruction is described as two components, firstly a two-dimensional vector is calculated based on the marker's image position and secondly a depth vector is calculated for each marker using the corresponding Voronoi cell area. When the force deforming the surface is released, the silicon springs back to its original shape.

Kamiyama et al. [37] developed a novel surface sensor called Gelforce that measures 3D force vector field of a deformable silicon surface. Force distributions are calculated by recognizing, with a CCD camera, how red and blue markers deform. This technique calculates only magnitude and does not provide direction vectors. The authors proposed this method might be employed to capture a sense of touch on robotic platforms. Again this technique springs back to the original form when the force is removed and does not maintain its shape.

Mass et al. [50] developed a silicon rubber material with embedded IR ARToolkit markers. Quimo is a white malleable sheet-material that can be molded with bare hands to produce low-fidelity physical prototypes. The sheet form allows hollow physical models to be constructed by cutting and bending the material into shape. Employing SAR to project imagery onto these low-fidelity mock-ups allows for complex surface appearances to be presented. Quimo supports maintaining a deformed shape but lacked fine grain re-construction resolution due to the minimal size of ARToolkit markers that can be employed.

Researchers have also explored how to capture deformations on paper with embedded markers. Martedi et al. [52] develop a method for folded surface detection and tracking for augmented maps. Plane detection is iteratively applied to 2D correspondences between an input image and a reference plane to determine where the folded surface is composed of multiple planes. Their algorithm computes the folding line from the intersection line of the folded planes of their positional relationship. Each plane can be individually tracked after this line is detected. They applied an overlay of virtual geographic data on each of the detected plane.

Uchiyama and Saito [102] developed a marker tracking system based on random

dots (Figure 4.1(a)). These markers are composed of randomly placed dots on a planar surface. The dot patterns are utilized for marker retrieval and tracking. A major feature is the marker does not require a square black frame. Using Local Likely Arrangement Hashing (LLAH) [60], they utilize the local patterns of the dots for descriptors in keypoint matching and tracking. Leveraging the fact LLAH works well on the unique local patterns of keypoints, they make the assumption the randomness of the distribution will lead to the unique patterns in the dots. Uchiyama and Marchand [101] extended the random dots marker system to non-rigidly deformable markers. They take advantage of the keypoint matching recognition and tracking of random dot markers. Their algorithm performs an estimation of the deformation of the markers with non-rigid surface detection from keypoint correspondences. The initial pose of the markers is determined. An estimation of the deformations is calculated from the minimization of a cost function for deformable surface fitting. Their algorithm inherently does not allow a large occlusion or self-occlusion.

McIlroy et al. [53] developed a fast dot pattern matching algorithm for the 6DOF tracking of a moving IR emitter. These dots are projected onto the surface by the moving IR emitter and detected by a fixed IR camera. In this system the surface is assumed to be a planar surface.

Pilet, Lepetit, and Fua [69] developed an early real-time method for detecting deformations on fabric material surfaces (Figure 4.1(b)). To detect a potentially deformable object, the process employs a strategy of building correspondences between a model image that contains small deformations and an input image in which there could be large deformations. A fast wide-baseline matching algorithm is employed to this end. As they mentioned in the paper, their method also does not allow a large occlusion or self-occlusion because their algorithm is based on 2D meshes and an optimization of one cost function without a care for severe occlusions.

Registering an organic shaped material has also been achieved using multiple markers. Mark Fiala [18] developed “The SQUASH 1000” using the ARTag fiducial marker system where objects of an arbitrary shape can be used as an input device. This system uses a non-planar substrate but does not use a deformable substrate, instead the ARTag markers are rigidly attached to the surface of a cucurbit.

Recently, a few studies succeeded the reconstruction of the deformable object in real-time. Zollhöfer et al. [112] combined hardware and software solution for marker-less reconstruction of non-rigidly deforming physical objects with arbitrary shape in

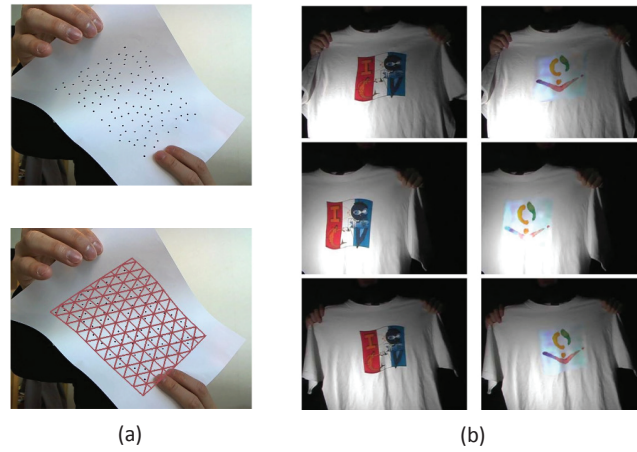


Figure 4.1. Related work on tracking for a deformable object: (a) random dot pattern for tracking of a deformable object [102], (b) tracking of a deformable object with texture information [69]

real-time (Figure 4.2(a)). With a single-contained stereo camera, a stereo matching algorithm estimates real-time RGB-D data. At first by scanning a smooth template model of the subject as the move rigidly with a geometric surface prior. Next, a GPU pipeline performs non-rigid registration of live RGB-D data to the smooth template using an extended non linear as-rigid-as-possible. However, this method is also not applicable to our objective because this method does not provide the unique correspondences between each location of the target object and each location of the texture to be projected.

Punpongsanon et al. developed a projection system for a deformable object (Figure 4.2(b), [73]). The motions of the future points are measured between two successive frames. Users can interact in real-time with the deformable object, while the realistic projected graphics deform according to the deformation of the surface. However, this system is also not applicable to our objective because it allows only restricted deformations such as the tangential deformation.

4.2.3 Embedded Sensor Based Recognition of Deformable Substrates

There are a number of systems that have explored techniques to capture deformable substrates with sensors embedded into the material. Michael Reed [75] explored de-

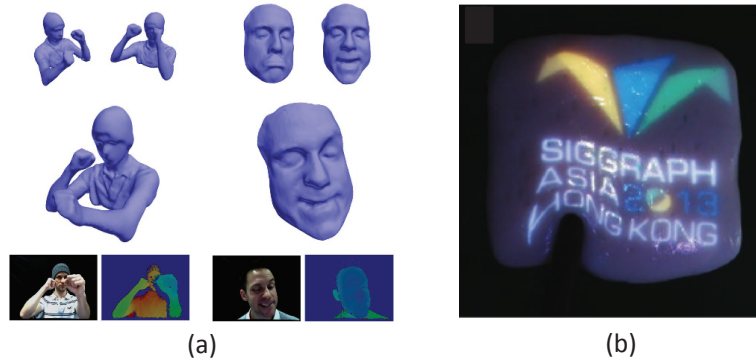


Figure 4.2. Related work on reconstruction for a deformable object and projection for a deformable object: (a) real-time reconstruction for a deformable object [112], (b) real-time projection for a deformable object [73]

veloping a digital clay material by embedding 6DOF wireless tracking sensors into a modeling clay material. He describes the ideal implementation with many thousands of miniaturized tracking units would be incorporated into the clay material. In his prototype a proof-of-concept with six wireless Polhemus sensors were used. A unique aspect of this system is the clay can be divided into multiple parts that are detected by the system. The biggest limitation of this system is the low count of sensors in the current implementation.

Smith et al. [87, 88] developed a deformable foam material, Digital Foam, which can approximate the surface shape when deformed. Digital Foam employs tubes of conductive foam that change in resistance when compressed. A 2D grid of foam tubes is insulated from each other with polyurethane to create an interactive surface. A planar and spherical form was used to perform sculpting inspired operations on digital models. One limitation of this system is the foam material springs back to its original shape when the users stops compressing the surfaces.

4.2.4 Clothing Design Simulation System

An example of design support targeting deformable objects that is also a fertile topic for research is design support for clothes. Of these, a popular example is the virtual try-on system that enables users to virtually “try-on” different clothes without having

Chapter 4. Design Support System for Deformable Objects Using Projection-based AR

to change their real clothes. The general idea underlying try-on systems is to track the user's motion in either 2D or 3D and synthesize a clothing image that can be overlaid onto the user's image [20, 107, 109]. In real try-on situations in general, people use a mirror to check their appearance with clothes on. Thus, this kind of system, which uses a display as if it was a mirror in a real situation, is highly compatible with a real situation (as shown in Figure 4.3(a)).

However, our intended application focuses on the design and development support for designers. In general, a designer cuts several pieces of clothing and attaches these to a clothing form to design clothes. In this case, the designer can directly observe clothes without a mirror. This is because the design process highly values careful examination of the texture and the functionality of a product. Therefore, projection-based AR, which projects images onto real clothes, is more suitable for this purpose than display-based AR (as shown in Figure 4.3(b)) because a designer would prefer to directly observe clothes. With our technique, a designer can check various appearances of clothes within a short time and without repeating the same cuttings, as if he or she had created different items of clothing with different pieces of cloth. For this purpose, we need to address a situation involving cutting (the entire region of a piece of cloth is cut and separated into several smaller regions) or one where self-occlusion occurs. With existing techniques such as Pilet et al.'s proposal [69] and Uchiyama et al.'s work [102], it is difficult to deal with such situations because these techniques recognize the shape of a target object as a monolith by optimizing a cost function that represents the entire region occupied by the object by disregarding large occlusions. By contrast, described in Section 4.4, our marker pattern is separately recognized as several very small regions, and is not inherently affected so much by cuttings or self-occlusions as a whole.

4.3. System Components

Our system is composed of three major components that are illustrated in Figure 4.4. The first sub-section describes the projector camera (pro-cam) system. This is followed by a description of our new deformable object that provides the projection substrate. Finally the new marker pattern is described.

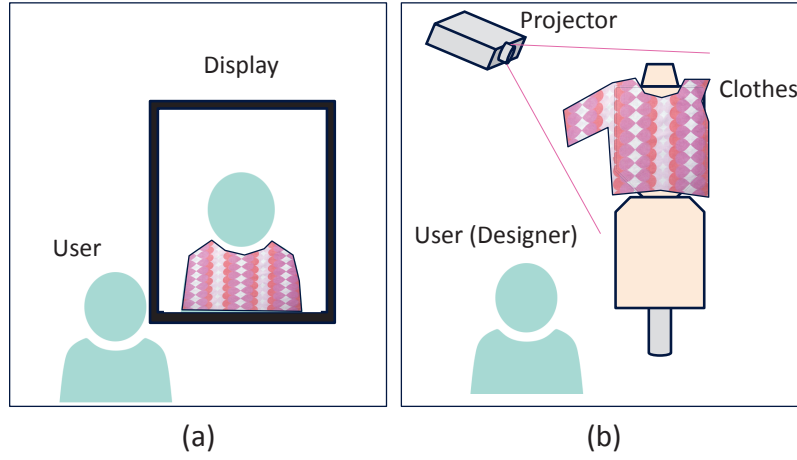


Figure 4.3. Difference between a virtual try-on system with a display and a clothes design support system with a projector (a) a virtual try-on system with a display (e.g., [69]), (b) a clothes design support system with a projector (a example application of our technique)

4.3.1 Camera and Projector

We employ a pro-cam based system to register the computer generated digital information onto the physical object. As with many pro-cam systems, the camera captures features of target objects such as shape, color and the projector displays texture images onto a target, in our case a deformable object. We employ a camera that captures both visible and infrared light under computer control. Each projector and camera is calibrated in advance, and the relative position and orientation between each pair is known during the operation of the system. Our system operates on a Window 7 PC with a Core i7-3930K 3.2 GHz, 16.0 GB of memory, NVIDIA GeForce GTX 560 Ti. A DLP projector with 1024×768 resolution and a camera with 1024×768 resolution are employed. While Figure 4.4 shows the component of only one pro-cam system, it is also possible to use more than one pro-cam system simultaneously (see Section 4.4.9).

4.3.2 Our Deformable Object

The goal of this investigation is to develop a tracked material that can be easily reshaped by hand and are suitable for natural light projection. We developed a prototype

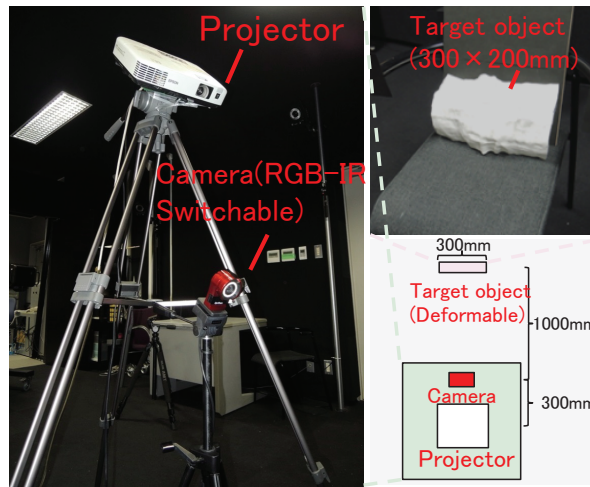


Figure 4.4. The major system components: Our system consists of projectors, cameras which are switchable from IR to RGB function, and deformable object.

material to meet these goals shown in Figure 4.5. The deformable object consists of a metal mesh and a acrylic foam sheet. The metal mesh is made of aluminum, and the thickness is 0.2 mm. There are 18 holes per inch of its width and 16 holes per inch of its height. The acrylic foam sheeting has a thickness of 1.1 mm and is attached to the metal mesh, which makes it possible to easily reshape by hand. The metal mesh ensures the deformable object retains its shape after the user manipulation. This object exhibits stretch properties because as it consists of a acrylic foam. However, we make the assumption that the small amount of stretch does not disturb the marker recognition process. We employ a 300 mm × 200 mm sheet in this investigation. However, a larger sheet or a smaller sheet can be used to meet different requirements.

4.3.3 Our Marker Pattern

In order to project texture images onto the deformable object, local positions of the object need to be recognized in a camera image to make correspondences between each location of the object and each location of the texture. To achieve this goal, we have developed a marker that can be partially recognized.

Szentandrási et al. [92] developed a marker system which can be partially recognized. Their marker is a grid pattern consisting of black and white rectangles developed for AR tracking on planer paper. This grid can recognize corresponding locations on

4.3. System Components

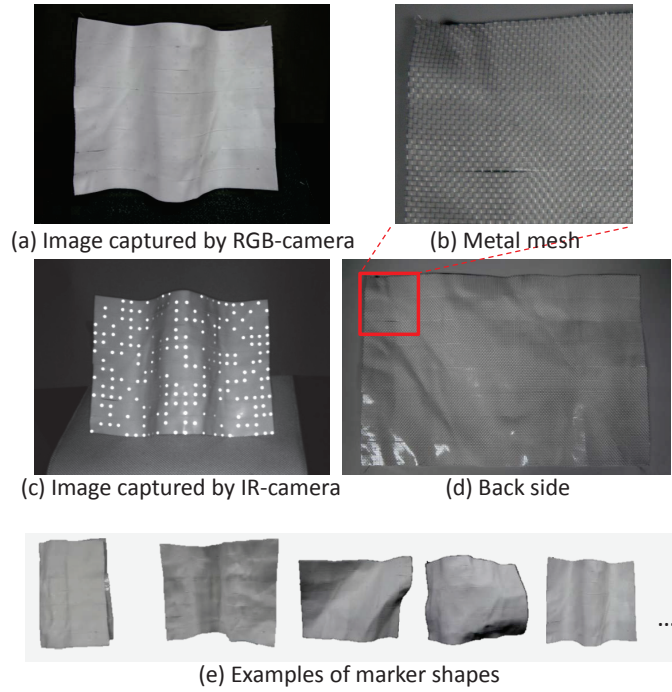


Figure 4.5. Overview of deformable substrates: Our new SAR enabled deformable substrate

the reference pattern as long as any 4×4 rectangle sub-pattern within the marker is observed by the camera. Their marker is based on aperiodic 4-orientable n^2 -window arrays (De Bruijn tori) [9] - two-dimensional arrays of binary values, where each $n \times n$ sub-window is presented only once, including all four possible rotations. In our investigation, we adopt a marker that has a similar property. We used the patterns with the 4×4 sub-window. In this case, the theoretical upper bounds of sizes for arrays are 127×127 . Instead of this pattern, we can easily use a more larger pattern if it is required. Since the marker-pattern is assumed to be attached to a plane surface in the Szentandrasei et al.'s research, 'rectangles' which can be easily processed in the separation process by a simple line detection algorithm, are employed as the minimum-unit. This unit has the characteristic that adjacencies between minimum-units are easily detected. Our marker-pattern is assumed to be attached to a non-planar deformable object. Therefore, we employ 'points' as the minimum-unit that can be easily and separately detected even if the surface is bent. Figure 4.6 illustrates the difference between the Szentandrasei et al.'s marker and our new marker. We developed a fundamentally diffe-

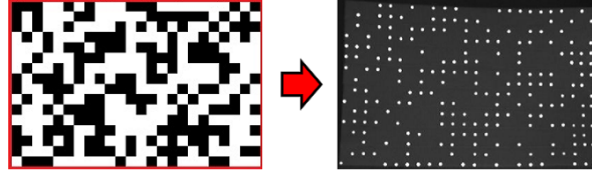


Figure 4.6. Szentandrasei et al.'s marker on the left and our new marker system on the right : A minimum unit of Szentandrasei et al.'s marker is a rectangle. A minimum unit for our marker is a point.

rent algorithm to recognize this new form of marker and the algorithm is described in the next section.

A desirable feature of the marker is a neutral color to act as a projection surface while simultaneously function for surface re-construction. We employ a half-transparent retro-reflective material for the construction of the marker-pattern. While the surface with this marker made of the retro-reflective material is observed as an almost white surface under the visible light region (see Figure 4.5(a)), the surface is observed as one with many retro-reflective points as shown in Figure 4.5(c).

In this research, our marker is attached to the sheet-shaped deformable object. The marker may be attached to a surface of a material with a higher ability of deformation (e.g. clay).

4.4. Algorithm

Figure 4.7 illustrates an overview of our algorithm. With the goal of geometrically correct projection of a texture image onto an arbitrary surface shaped object. The algorithm is required to recognize local locations on the surface of the target object and deform a texture image to fit the surface. The projection employs small patches surrounded by four points. We assume that the correspondences between each point of a texture image and the reference marker pattern are predefined. This is a common assumption used with standard graphics texturing algorithms.

Figure 4.7(a) depicts the case of a pro-cam that has the optical axes of a camera and a projector corresponded with each other. The first step is to determine the relationship between the points on the reference marker pattern and the corresponding points in the captured image by the camera. The marker recognition algorithm described in the next

4.4. Algorithm

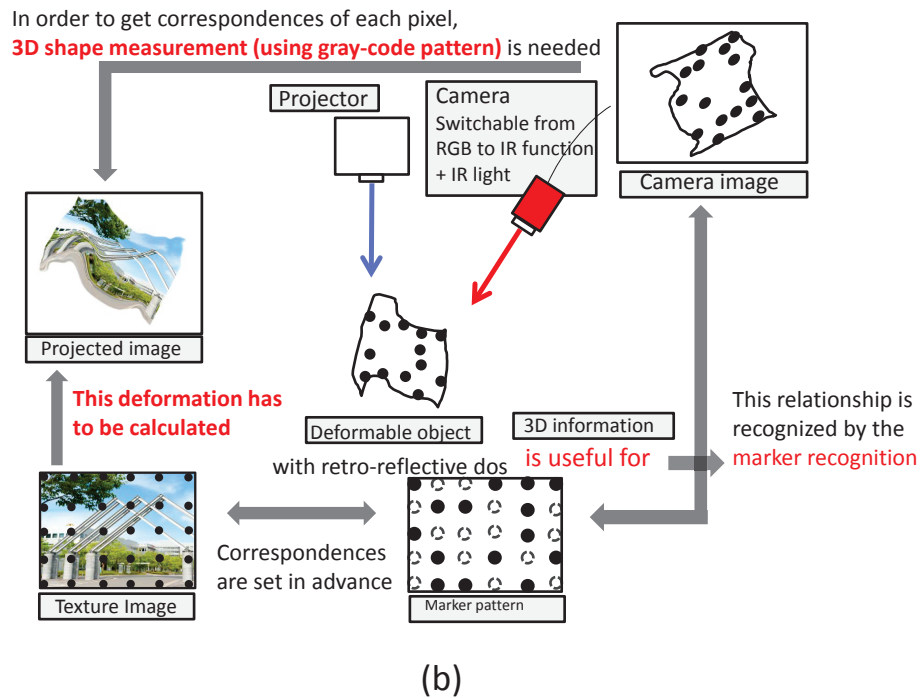
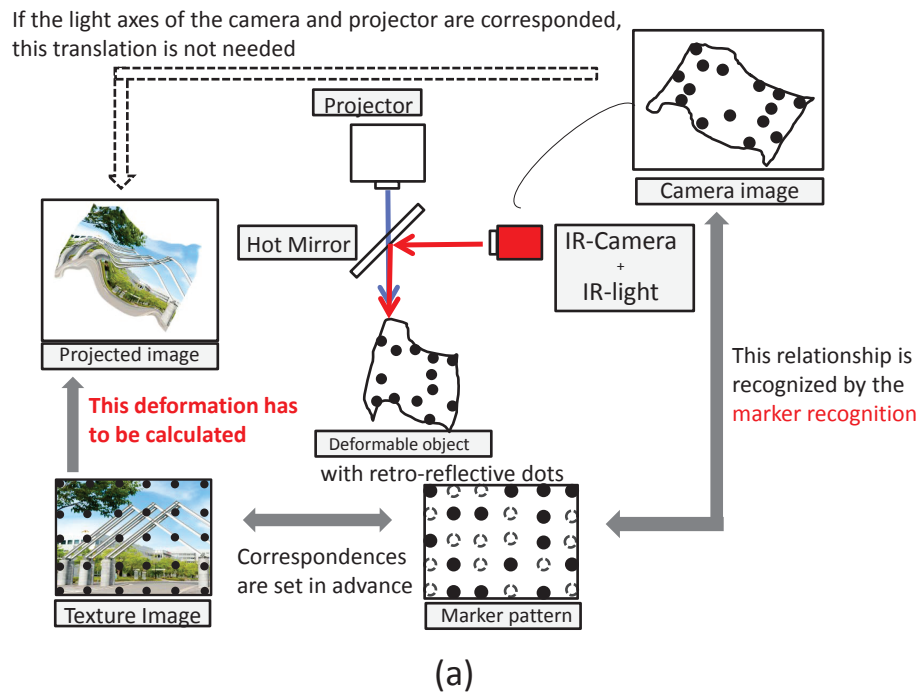


Figure 4.7. Overview of marker detection algorithm: (a) a pro-cam which optical axes of a camera and a projector are corresponded in advance (b) a pro-cam when optical axes of a camera and a projector separated.⁸¹

Chapter 4. Design Support System for Deformable Objects Using Projection-based AR

section performs this. After that, making correspondences between a camera image and a projector image per each pixel is required. This process is not required for the pro-cam case in sub-figure (a). In order to detect retro-reflective dots on the deformable object, an IR-camera (with an IR-light) is needed. Because RGB-light by a projector and IR-light by an IR-camera have to be separated in such systems, a hot mirror (which reflects infrared light back into a light source, while allowing visible light to pass) and its subtle arrangement are required. However, such pro-cams are expensive and not very accessible.

In order to make more portable and inexpensive systems, pro-cam systems in which the optical axes of a camera and a projector are separated (normal pro-cams) are employed in this research, as shown in Figure 4.7(b). In this case, because the projection has to be changed depending on distances between the projector and local locations of the object, making correspondences is required as mentioned earlier. Then we employ a gray-code pattern method [33] to get correspondences between each pixel of cameras and projectors. A gray-code pattern is projected onto the object and observed by an RGB-camera. This is an encoding method which projects several different patterns made of each bit of the gray-codes in time series. While the method takes time to project each gray-code pattern, it is possible to calculate a dense 3D point cloud that corresponds to each camera pixel. The dense 3D information calculated in this process will be employed in the marker recognition and the interpolation process. In order to achieve this construction, both an RGB-camera and an IR-camera (with an IR-light) is required for capturing both the gray-code pattern (the RGB-light) and the retro-reflective dots (which reflects the IR-light). We adopt a camera that is switchable from the RGB to the IR function by a PC command (e.g. OptiTrack by NaturalPoint [65]). This feature of the camera makes for a straightforward construction of the system hardware and makes an expansion for multi-pro-cams easier.

In this section, processing with only one pro-cam is described. Figure 4.8 illustrates each step of the algorithm with a single pro-cam. Processing consists of four factors; 3D measurement, marker recognition, interpolation and projection. 3D measurement is performed by a standard gray-code pattern method, and each of last three steps is described in detail below.

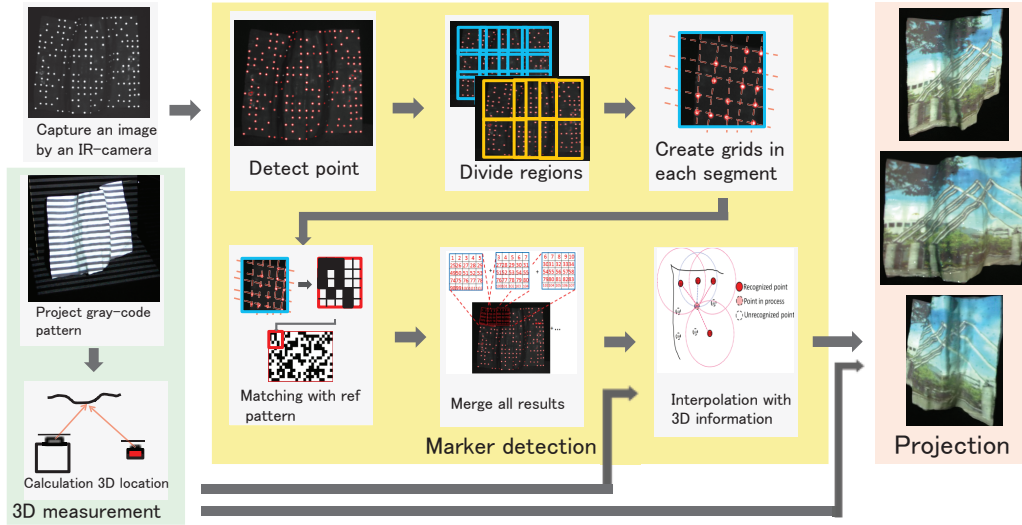


Figure 4.8. Overview of system algorithm: Processing consists of four factors: 3D measurements, marker recognition, interpolation and projection

4.4.1 Marker Recognition

The steps for recognizing the marker are described in this section. The process starts with the point detection algorithm. The process moves onto region division (sub areas that are visible and detectable on the marker) and aligns these regions to a smaller grid pattern. These smaller regions are then matched against the full marker. Because the regions may be detected when overlapped with other regions, a decision process for determining the best match is required.

4.4.2 Point Detection

Initially each retro-reflective point is detected in an image captured by the camera in IR mode. We refer to these points as *detected points*. One of the principal issues of this process is detecting false points caused by specular highlights on parts of the surface of the deformable object. It is also important to correctly detect dimly lit points due to the small amount of reflected light that occur due to regions where the angle between a surface normal vector and a camera optical vector are large. Deletion is conducted towards wrongly detected points in an integration step, and an interpolation process is performed for undetected points in the subsequent steps.

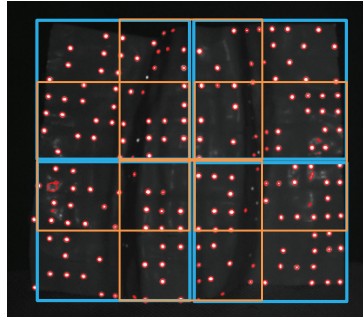


Figure 4.9. Example of first region division. Allowing duplicates, nine rectangles with a same size are used to divide the whole region (for descriptive purpose, two colored rectangles are used).

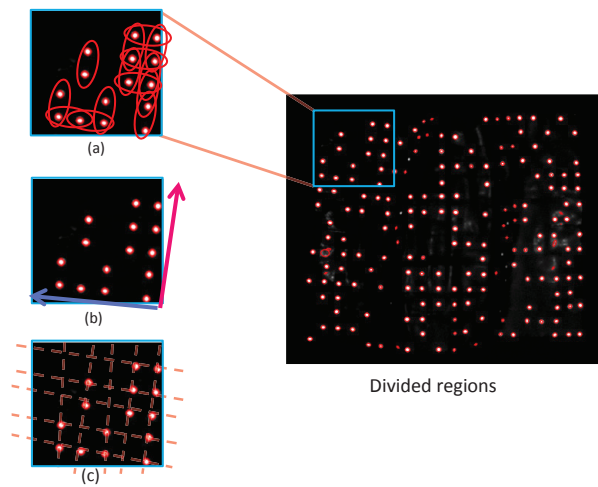


Figure 4.10. Example of grid making: (a) By 3D information, each two consecutive points are detected in each small region. (b) Two salient directions are decided. (c) Grid lines are calculated through as many as possible of the detected points.

4.4.3 Region Division and Grid Calculation

This step aims at recognizing a relationship between each of the points on the non-planar deformable surface. In order to deal with the changing form factor, we decided to use several sized regions and recognize them separately. If a region is small enough as compared to the curvature of the surface, it can be expected that configuration of points in this region are similar with each other. The system divides the entire surface into smaller overlapping regions in a hierarchical manner. At first, a large rectangle including all *detected points* is detected. The region of the rectangle is divided to smaller rectangles. A half length of each small rectangle is simply overlapped with neighboring rectangles along both horizontal and vertical axes as shown in Figure 4.9. If the number of points included in each divided region is larger than a preset threshold value, the rectangle is subdivided to a $3/4 \times 3/4$ size of the previous rectangles and the division of the region is conducted again. This processes is repeated until the number of points in each sub-region is small enough.

Grids are then created through all the *detected points* in each divided region (shown in Figure 4.10). The grid is determined from two salient directions from all possible directions of lines through pairs of points in the divided region. With 3D information, the process first creates pairs both consisting of 4-connective-points (shown in Figure 4.10(a)) that are a predefined euclidean distance apart and within a margin of error. Next, directions of lines through the pairs are calculated. Two salient directions are decided by counting the number of each line directions. These two directions can be regarded as base directions of a grid in this small region (shown in Figure 4.10(b)). Based on these two directions, grid lines are calculated through as many as possible *detected points* in the region (shown in Figure 4.10(c)) for use with matching the reference pattern.

4.4.4 Matching with Reference Pattern

We refer to each intersection on the grid as “*grid points*” that can be represented in binary and can be used as a small marker pattern. We search for the reference pattern to find the most similar region. A location on the reference pattern can be uniquely recognized among $2^{16}/4(Up, down, left, right) = 16,384$ (types of patterns) as long as any 4×4 rectangle pattern is observed by a camera. The matching with the reference pattern is conducted for every 4×4 units among the created grids. As a result of this

Chapter 4. Design Support System for Deformable Objects Using Projection-based AR

matching process, IDs of locations on the reference pattern (in our case, 1-384, these are predetermined) are temporarily assigned to each *grid point* in each small pattern.

4.4.5 Error Removal with Voting

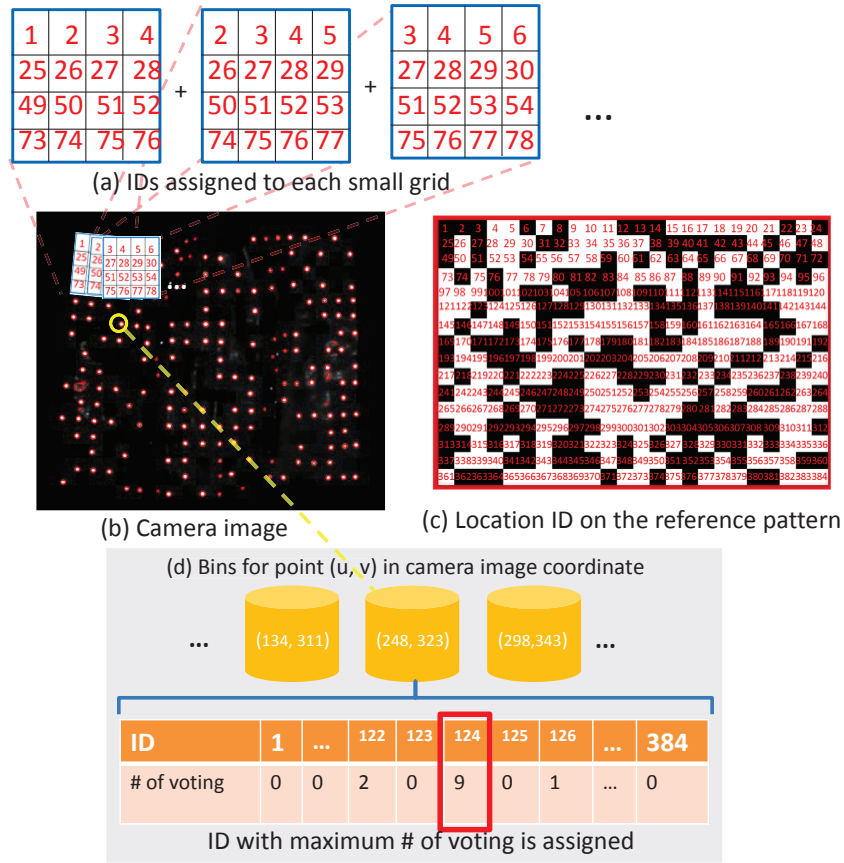


Figure 4.11. Small region integration and error removal: In order to integrate small regions, IDs are assigned to each point for voting. The ID with a maximum number of votes is assigned as the final ID of each point on the image

After matching each grid with the reference pattern in the small region, these are merged again to make a whole marker pattern. Then the IDs given to each *grid point* in the matching phase are assigned to bins of corresponding points in the whole image as shown in Figure 4.11. Each smaller region is allowed to overlap so that points on the image obtains several IDs' votes. The process can be regarded as a method of voting.

4.4. Algorithm

The ID with a maximum number of votes is assigned as the final ID of each point on the entire image. ID points with a few number of votes are rejected because they can be regarded to having wrong assignments. We employ this voting mechanism to retain only reliable points.

Algorithm 1 Interpolation for unrecognized points

```

1:  $N \leftarrow 4$ 
2: while  $N > 1$  do
3:   for all unrecognized points ( $X_{unrec}$ ) do
4:      $k \leftarrow \#$  of recognized 8-connective-points of  $X_{unrec}$  ( $X_i$ )
5:      $n \leftarrow \#$  of neighboring points of  $X_{unrec}$  ( $X_j$ ) in an image
6:      $const_i \leftarrow 13.0$  (if  $X_i$  is a horizontal or vertical adjacent point),  $13.0 * \sqrt{2}$  (if  $X_i$ 
       is a diagonal adjacent point)
7:     if  $k \geq N$  then
8:       if  $k = 2$  or  $3$  and all  $X_i$  is in a row then
9:         continue
10:      end if
11:      for all  $X_i$  do
12:        for all  $X_j$  do
13:           $D_{ij} \leftarrow$  3D distance between  $X_i$  and  $X_j$ 
14:        end for
15:      end for
16:       $j_{inter} \leftarrow \operatorname{argmin}_j \sum_{i=1}^k (D_{ij} - const_i)$ 
17:       $X_{j_{inter}}$  is a interpolated location
18:    end if
19:  end for
20:  if all unrecognized point does not have  $N$  or more recognized 8-connective-
    points then
21:     $N \leftarrow N - 1$ 
22:  end if
23: end while

```

4.4.6 Interpolation

At this point in the process, IDs have been assigned to each of *detected points*, but some grids may not be correctly created in the regions of high curvature. In such regions the maximum number of votes are deemed to be not large enough to be classified; therefore the *detected points* remain in an ID-unassigned state. An additional contributing factor is that amount of reflected light decreases dramatically in a region where surface normals are highly oblique to the optical axis of the IR-camera. Considering these factors, there are some cases the process has difficulty observing and processing all the points on the surface of the deformable object. In these cases we attempt to interpolate the location of the IDs which have not been associated with a detected point.

The interpolation is conducted using point locations which have been associated with IDs (hereinafter called “recognized points”) among 8-connective-IDs on the reference pattern of an ID in process. Since the ideal 3D distances between a point in process and recognized points of its 8-connective-IDs are known (In our situation, a distance between horizontal or vertical adjacent points is 13mm and a distance between diagonal adjacent points is $13 \times \sqrt{2} \doteq 18.4\text{mm}$). It is possible to decide a required location on the surface of the object if some 8-connective-IDs has been assigned. Figure 4.12 (a) depicts potential locations of a point, the sub-figure (b) shows a simplified process calculating intersecting spheres to determine potential interpolated locations. The distance used in the interpolation process is along the surface of the deformable object. So an equidistant curve of each detected point does not draw a sphere. Error is accumulated when interpolated points are used to interpolate additional points in the process of this method. In order to reduce accumulation error, we preferentially interpolate high reliability points. Here the high reliability denotes that the number of recognized 8-connective-points is large. The following lists the selection criteria for the determination if a location may be decided.

of recognized points: 0-1 A location of a point cannot be decided.

of recognized points: 2-3 If all points are in a row (vertical or horizontal), two candidates exist. A location cannot be decided uniquely. Otherwise, a location can be decided uniquely.

of recognized points: 4- A location can be decided uniquely.

Based on this ability to uniquely decide a candidate, we conduct the interpolation

process preferentially from points which have more than three recognized 8-connective-points.

Algorithm 1 shows the whole procedure of the interpolation.

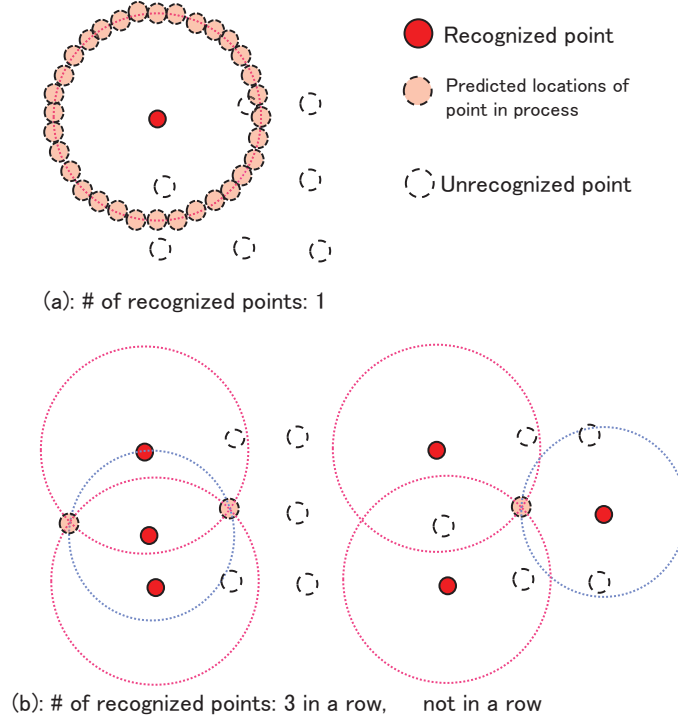


Figure 4.12. Interpolation: (a) When # of recognized points is one, a location of a point which should be interpolated cannot be predicted. (b) When # of recognized points is three and all points is in a row, two candidates exist. When all points are not in a row, a candidate can be determined uniquely.

In addition, due to the 3D information and an RGB image (with simple thresholds), we can easily detect the deformable object region, which makes it possible for a correct interpolation under situations when the occlusion occurs and when the whole deformable object is not captured by cameras.

4.4.7 Projection

The textured image has to be mapped onto the deformed object. Since it is predicted that each small patch surrounded by four points is small enough, the projection is conducted based on each patch. By translating locations in the camera image coordinate

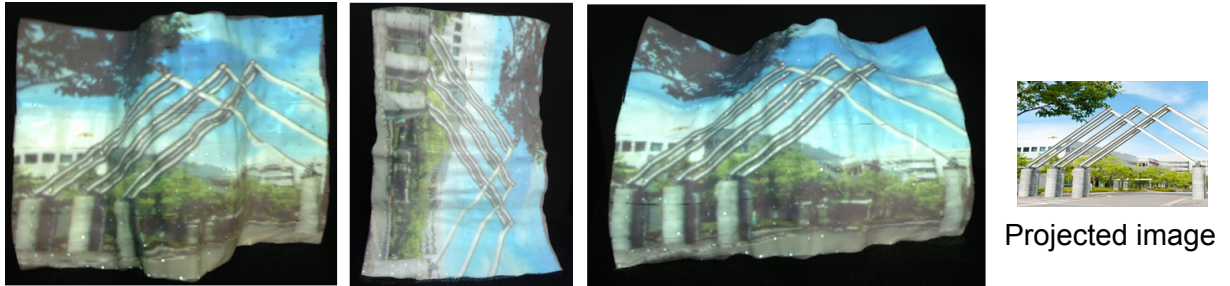


Figure 4.13. Results of projecting texture image, the same image is deformed and projected onto differently deformed surfaces

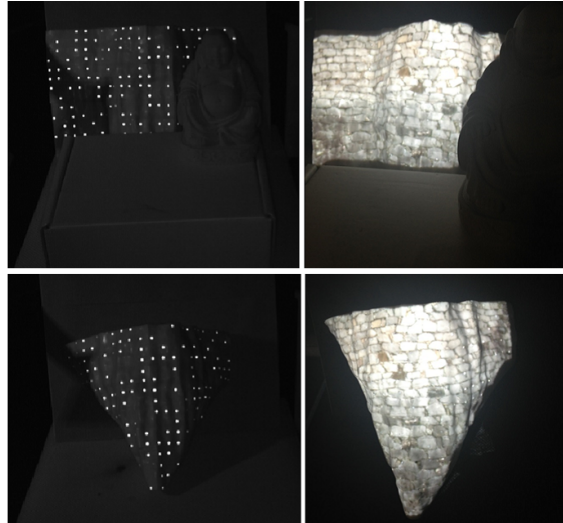


Figure 4.14. Results of projecting texture image when heavy occlusions occur

of the IDs calculated in the previous step to those in the projector image coordinate, a small part of the texture image is deformed in accordance with the shape of target deformable object. Figures 4.13 and 4.14 illustrates projection results of the same texture images onto several differently shaped deformable objects. You can see that the texture image is projected geometrically-correct onto several shaped-objects. While in this example, correspondences between each local location on the deformable object and the texture image is set in advance, it is also possible to make correspondences dynamically by user's interaction.

The time efficiencies of the process are important for interactive interactions. While the current implementation does not provide the required time efficiencies, the process

can be optimized in the future. We measured the average processing time of the system operating 50 times. We found each process takes the following time: 3D shape measurement - 5.04 secs, marker recognition - 0.42 secs, interpolation - 0.21 secs, and ready for projection - 0.05 secs.

4.4.8 Discussion on Processing Speed

Currently, the most time consuming aspect is the 3D shape measurement using gray-code patterns. A significant speed improvement can be achieved by synchronizing the camera and project frames. With a 10 bit gray-code pattern (each pattern has a negative and a positive) the minimum number of required gray-code images is 20. Employing each of this optimization aspects it is possible to shorten the processing time to approximately one second.

To achieve a faster processing speed we can stop using gray-code patterns. One option is the use of active depth sensors that work in a very short time instead of the gray-code patterns. To retain the same projection accuracy, a sensor that has the same 3D measurement accuracy of the gray-code patterns (in the current construction, the Euclidean distance error is about 1 mm) is required. This approach for a faster processing speed is described again in Section 4.7. Another solution is the use of a corresponding optical axes for the camera and projector. The advantage of this approach is that correspondences are not changed and gray-code patterns are not required which leads to the improvement of the processing speed. However, some modifications are needed to make these algorithms work without the 3D information for the grid making process.

4.4.9 Extension To Multi-Pro-Cams

A practical system requires more than one pro-cam system to increase the range of the projection area. Of course, if several cameras are placed surrounding a target object, a larger marker region of the object can be recognized correctly by at least one camera. In this section, the algorithm for a multi-pro-cams system is described. Here, we consider the situation where several paired projector-cameras are used for the sake of convenience. However, an algorithm described here can be easily adopted to the situation where some projectors are not paired with any cameras and vice versa.

Chapter 4. Design Support System for Deformable Objects Using Projection-based AR

Highly accurate locations must be determined by the marker recognition process, and these are deemed to be better than those determined by the interpolation process. This is the case when a region (surrounded by four locations of *grid points*) is recognized by only one camera and is interpolated by several other cameras. When this occurs, locations determined by the marker recognition are used preferentially.

Next, our algorithm chooses which projector should be employed on each region. If several projectors project textured images onto the same regions of an object simultaneously, overlapped regions get an increased brightness and disrupts the blending with other projected regions. The angle between a optical axis of each projector and normal vectors of the target object's surface is adopted as the selection criteria. The chosen projector has the minimum angle for the projection, as the projected image has the best resolution for users. We regard each region (surrounded by four locations of *grid points*) as a patch and decide which projector should be used for each patch.

We have developed a straightforward algorithm for choosing the best projectors. At first, all the 3D information in a region is gathered. A covariance matrix C which elements are distances between a centroid and each point are calculated as shown in Equation 4.1.

$$C = \frac{1}{n} \sum_{j=1}^n (p_i - \bar{p})(p_i - \bar{p})^T \quad (4.1)$$

Here \bar{p} is a centroid vector in the region. By conducting principal component analysis to the matrix, vectors corresponding with first and second principal components are obtained. They are tangent vectors to the surface in the region. A normal vector can be obtained by calculating cross product of them. We use a value calculated by the inner product of the normal and a projector's optical vectors as the reliability. Comparing this value of each projector in each patch, a projector which has maximum value is adopted for the projection to each patch.

Here, we consider only the oblique blur effect (loss of high-spatial-frequency components according to the incidence angle of the projection light) [59] as shown in Figure 4.15 since we can assume that the distance variance on the object is relatively small. The defocus blur effect can be negligible in our current situation. However, if a larger object (with various depths) is adopted, the defocus blur derived from the variance of the distance between a projector and an object also has to be considered. In this case, the projector selection method proposed by Nagase et al. [59] is promising. This method chooses an appropriate projector among multi projector for each region in in-

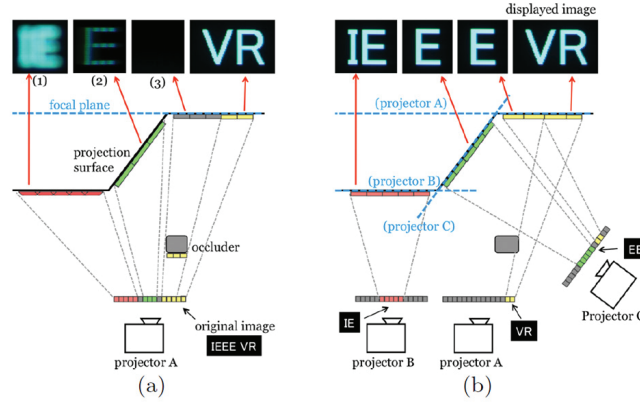


Figure 4.15. Example of an two blur effects [59] (1) defocus blur (b) oblique blur

teractive systems and can treat both the defocus blur effect and the oblique blur effect explicitly.

4.5. Recognition and Evaluation

In order to confirm the ability of the algorithm, several evaluations of the projection process were conducted. First, a projection evaluation for a single projector camera is described in the situation where the whole region is captured by one camera in Section 4.5.1. Second, an evaluation is described in the situation where more than one camera is required since the whole region cannot be capture by only one camera in Section 4.5.2.

The local projection accuracy is evaluated for a single pro-cam system. The evaluation process starts with the determination of each grid point location (without retro-reflective points) on the deformable object. These *grid points* are marked in advance. After deforming the deformable object to a specific shape, a grid point texture image is projected onto it. If the recognition and the projection grid patterns align perfectly correct, each grid point in the texture image is projected onto the marked grid point locations of each point on the deformable object. In order to evaluate accuracy of recognized points and interpolated points separately, the recognized points and interpolated points have different colors in the texture image. We then visually confirm the locations of projected points and marked them on the surface. For repetitive evaluations, We used an ink which has the feature that it becomes invisible gradually over time.

After marking at all projected points, the shape of the surface is reinstated to the almost plane surface. We measure distances between each actual point on the surface and each marked point (projected points) in the previous process. Figure 4.16 illustrates results for two shaped object. Double circles denote points decided by the marker recognition and single circle denote points decide by the interpolation.

4.5.1 Projection Evaluation for Single Pro-Cam

Table 4.1 illustrates the mean error values for each shape. The mean error is 2.04 mm for shape 1 and 2.47 mm for shape 2. As shown in Figure 4.16, the projection accuracy is high in regions which has little shape change because there were recognized points via the marker recognition. By contrast, the projection accuracy decreases slightly in regions which has a large shape change because locations of each *grid points* are decided by the interpolation process. It is believed that the improvement of the 3D measurement accuracy leads to the projection accuracy. In addition, if several pro-cams are used, the number of recognized points increases, which results in the improvement of the projection accuracy.

4.5.2 Projection Evaluation for Multi-Pro-Cams

This section presents the evaluation for multi-pro-cam system. A two pro-cam system and the algorithm are used to decide which projector should be employed for each patch. Figure 4.17 (a) illustrates deformable object's images captured by two cameras. Since the center of the deformable object is raised, each side of the uplift can be

Table 4.1. Projection accuracy by the single pro-cam: These value illustrates mean distances between each grid point on the deformable object and each projected point by a projector

Shape	1	2
recognized point	1.40 (SD=1.61) mm	1.25 (SD=1.48) mm
interpolated location	2.92 (SD=1.85) mm	3.29 (SD=2.11) mm
sum	2.04 (SD=1.95) mm	2.47 (SD=2.05) mm

4.5. Recognition and Evaluation

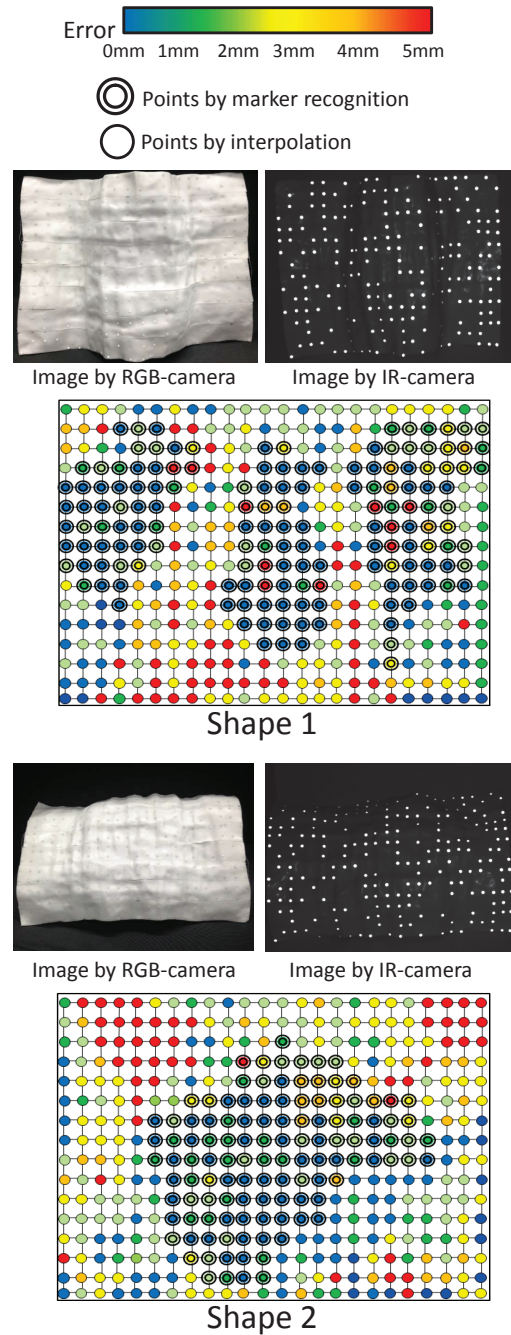


Figure 4.16. Projection evaluation for single pro-cam: Each color represents the error value for each point. Double circles denote points decided by the marker recognition and single circle denote points decided by the interpolation.

Chapter 4. Design Support System for Deformable Objects Using Projection-based AR

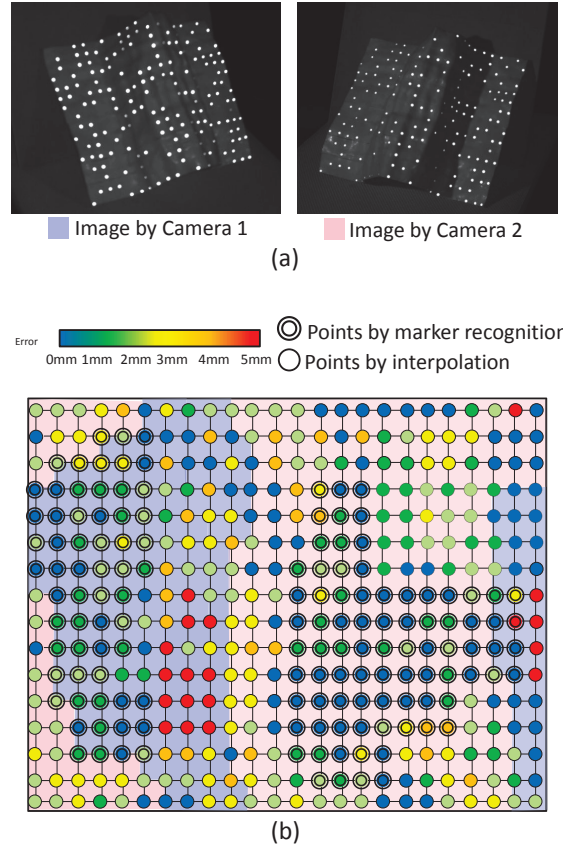


Figure 4.17. Projection evaluation for multi pro-cams: Light blue and pink regions represents projected region by each projector

observed by one camera.

The procedure of the evaluation is the same to one for the single pro-cam and Figure 4.17 (b) illustrates the results. Double circles denote points decided by the marker recognition and single circle denote points decided by the interpolation. Light blue regions represents regions projected by the projector 1 (paired camera 1) and light pink regions represents regions projected by the projector 2 (paired camera 2). In this evaluation, points on boundaries of projection images are projected by a projector which has higher reliability in neighboring regions of boundaries. Each side of the central uplift is appropriately projected by each projector. Table 4.2 denotes the mean error value of projection. These error values are an improvement than our single pro-cam system. These values are affected by the shape of target object and physical placement

4.6. Projection-based AR for Design Support and Its Advantages

of the pro-cams relative to the tracked object. However, it is predicted that using several pro-cams increases the whole projection accuracy.

4.6. Projection-based AR for Design Support and Its Advantages

Our new technology was inspired in part to help provide better prototyping tools for product designers. We are exploring how phases of the modeling methodology used by industrial designers can be enhanced by using SAR technologies. The drive is to enable designers to visualize their concepts with complex detail and be provided with a more flexible modeling environment. This section will outline how this new technology tackled a key problem for applying SAR in the product design prototyping process. The section starts with an overview of current design prototyping practices. The use of SAR in design prototyping is briefly described, and a description of some major limitations in the use of SAR. How our new technology overcomes these problems is then presented.

4.6.1 Current Design Prototyping Practice

Common in the design process of many different artifacts (for example, automobiles, and home electronics) is to start with rough sketches and move towards some form of virtual representation (such as rendering in PhotoShop ²). During this phase a large number of ideas are generated and evaluated, as these concepts are very quick to in-

Table 4.2. Projection accuracy by the multi pro-cams: These value illustrates mean distances between each *grid point* on the deformable object and each projected point by each projector

Projector	1	2
recognized point	0.99 (SD=1.20) mm	1.20 (SD=1.11) mm
interpolated location	2.88 (SD=2.28) mm	2.22 (SD=1.72) mm
sum	2.01 (SD=1.99)mm	1.97 (SD=1.44) mm

²<http://www.photoshop.com/>

stantiate. The designers wish to move as quickly as possible to “the 3D”, i.e. some form of physical representation. Prototype physical mock-ups at this point are developed. These may start as quite simple shapes to gain an understanding of size and form. These also can become quite complex, such as a clay model of the dashboard for an automobile. They lack any color or details outside their physical shape.

The surface appearance is an important aspect of the physical mock-up. Designers use paint and inks to color and texture their mock-ups. A drawback with this tactic is when altering a design, either a separate mock-up requires to be constructed or the original mock-up must be re-painted. Although clay and polymer plasticine are able to have their shape changed unremittingly, this is not probable once painted, as the painted surfaces lose malleability.

4.6.2 Current Use of SAR for Design Prototyping

SAR provides a means for enhancing the details of the physical models in a flexible manner. Current industrial design research is exploring this application of SAR [11, 82]. Verlinden et al. [103] developed the idea of *Augmented Prototyping* (AP). They employed SAR to project onto objects that have been manufactured by standard rapid prototyping techniques. They found SAR offers a tangible and social interface for the designers. Their platform Workbench for Augmented Rapid Prototyping focuses on the early phase of the design process in which the promptness of producing an impression of the product is critical. In general the design process that employs SAR follows these steps: 1) design a physical prototype, 2) build the physical prototype, 3) define a virtual model of a similar shape as the physical prototype, 4) texture the virtual model, and 5) project the virtual model onto the physical prototype via SAR. Changes to the virtual model’s texture may be made at any time. If the physical model is altered, the process has to be repeated.

4.6.3 Current Restrictions Using SAR for Design Support

As previously mentioned, a major restriction with SAR is the shape of the physical models must be known at all times during the projection of SAR design information. Kinetic models may be employed to adjust articulated rigid bodied objects, but highly deformable materials require real-time scanning to changes in shape and size. While 3D scanning technologies allow for the reconstruction of a virtual model of the new

4.6. Projection-based AR for Design Support and Its Advantages

physical shape [36], the correspondence between the projected imagery and the physical shape is lost. This breaks the flow of the designer's process. What we are striving for is a continuous looping between working in the virtual space (textures and 3D graphical objects) and in the physical space (physical prototypes and mockups).

4.6.4 Our New Design Support

The system presented in this chapter solves the problem of losing correspondence between the projected imagery and the physical shape. Using our new system, the designer is allowed to apply virtual detail to the model at anytime in the process, and for that model to be further deformed. When using SAR for development the goal is to allow both the physical model and appearance to be effortlessly altered without demanding a new prototype to be constructed. The rubber and mesh substrate provides a construction material for basic mock-ups used for projection that is very simple and lightweight to use. The exterior of the mock-up is transformed via projected visual information to present all the fine grained details of the design. Our technology overcomes the restriction by allowing numerous appearances to be projected in succession onto one mock-up. The designer may digitally paint straight onto the model by interactively amending the projected texture to alter the appearance. The texture can be saved and recalled in the future for further evaluation.

A second application of our technology is *digital draping*.³ Although we have not constructed this application, we believe it is technically feasible. A neutral colored piece of cloth is embedded with our retro-reflective pattern. The cloth is placed, or draped, over an existing object and the shape of the object is captured. Digital information may now be added and manipulated on this cloth surface as with our previous examples. The digital draping technique could be used for visualization - for example the visual appearance of a bus's interior design may be altered under computer control. To achieve this digital draping material (the deformable cloth substrate) may be placed to cover existing elements of the bus, such as seats and ticket vending machines and the pro-cam system is used to provide virtual illumination and presentation of design concepts.

As a third application, our technology is applicable to non-planar geometrical registration in pro-cam systems as well as design simulations as shown in Figure 4.18.

³The term digital draping was coined by Sean Pickersgill.



Figure 4.18. Example of an apparent plane projection onto an arbitrary curved surface by considering geometry information of the projection surface from [91] (a) Without geometrical consideration (b) With geometrical consideration

Non-planar geometrical registration in pro-cam systems is a technique that provides an apparent plane projection onto an arbitrary curved surface by considering the geometrical information of the projection surface and projecting a texture image with this information. For complex surfaces of unknown geometry, correspondences between a projector image and a camera image (lookup-table) have frequently been used. Yang et al. proposed a method based on feature matching involving a projected image [106]. More simply, structured light techniques [80], such as gray-code patterns, have been frequently used in various studies. For greater robustness robust when forming correspondences, Zollmann et al. used a combined method involving encoded points and simple gray-code patterns [113]. While past methods required more than one pattern, Sun et al. proposed a method based on a robust checker board pattern recognition that used only one pattern image [91]. These methods required re-projections of patterns as often as a target object moved or was deformed. However, our method allows the instant recalculation of correspondences between a projector image and a camera image without any projections, with the limitation that the surface of the target projection is embedded with retro-reflective marker patterns.

4.6.5 Limitations

The current technology has several limitations.

The first is that the retro-reflective dots must maintain their spacing throughout the surface re-construction process. In the future, we would like to investigate accommodating materials that allow for stretching.

4.6. Projection-based AR for Design Support and Its Advantages

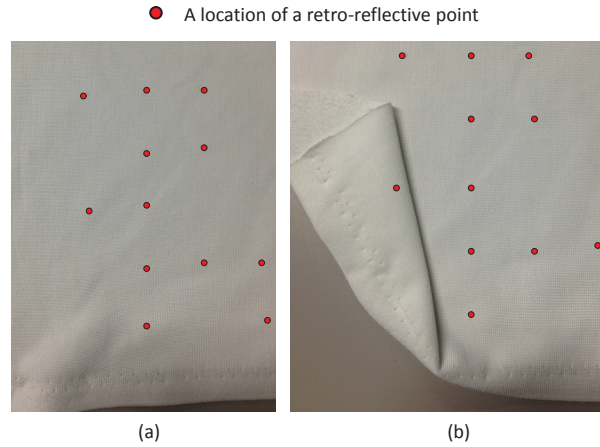


Figure 4.19. Example of current problems: (a) a situation without any large deformations (b) a situation when the object is folded

The second limitation is that the shape of the surface maybe altered by bending. There are other processes the designer uses, such as cutting, scraping, and carving. These pose two particular difficulties: they maybe very fine in detail, and the act may remove the retro-reflective material.

Figure 4.19 shows the third limitation. Figure 4.19(a) shows a situation without any large deformations in the object. Figure 4.19(b) shows a situation where the object is folded. In this case, a retro-reflective point in the folded region is placed at a different point occluded by the folded region. Because these two sub-figures look similar through an image captured by an IR camera, the system cannot recognize any differences between the two images. This will lead to an incorrect projection along the border of the folded area in Figure 4.19(b). We think that this problem can be solved by carrying out surface reconstruction with 3D information and recognizing each border.

4.6.6 Application: Clothing Design Simulation System

We constructed a simple system an example of the application of our algorithm.

As this system was used for clothing design support, the target object was a neutral-colored piece of cloth, as shown in Figure 4.20.

This application showed that our algorithm is applicable not only to acrylic foam sheeting, but also to other materials with little elasticity (less than 5% of the original size), such as a piece of cloth. The cloth had the same dot marker pattern as described



Figure 4.20. Projection target: a neutral colored piece of cloth

in Section 4.3.3. Because the size of the cloth ($100\text{ cm} \times 80\text{ cm}$) was greater than that of the acrylic form, the distance between two consecutive dots was changed to 26 mm. Figure 4.21 shows the target object and some projection results. The piece of cloth was cut in the shape of a t-shirt. Correspondences with each location on the t-shirt-shaped cloth and each location on the textures to be projected were provided in advance. Once the system detected each dot on the cloth, any texture image could be projected onto the surface of the t-shirt-shaped cloth, which allowed users to feel as if the cloth was physically painted with different textures.

4.7. Improving Processing Speed

The implemented system described above can be used for situations where the shape of the target object does not change frequently (e.g., a situation when a deformation occurs once every three minutes). However, it is not applicable to situations where the target object frequently changes. For example, when clothes are the target object and a user actually wears them, he or she has to wait until the re-calculation operation as often as he or she moves, even slightly, in front of the system.

One of the most time consuming aspects of the current system is 3D shape measurement, as shown in Table 4.3. In this section, we describe our approach for improving the processing speed of our system.

4.7. Improving Processing Speed



Figure 4.21. Application: Design support system for clothes

4.7.1 For 3D shape measurement

The problem is the use of the gray-code pattern method for 3D shape measurement. Although the gray-code pattern provides highly accurate 3D information, it is too slow for some applications, such when the real-time deformation occurs frequently. The use of Kinect v2 [55] instead would be promising. Kinect v2 provides both an RGB camera and an IR camera and can capture 3D information in real-time (30 fps).

However, there are the following problems with the application of Kinect v2 to our system, in the stead of the current switchable OptiTrack camera and gray-code patterns:

1. Calibration between the RGB and IR cameras of Kinect v2

Unlike the OptiTrack camera, the optical axes of an RGB camera and an IR camera on Kinect are not correspondent. Hence, the two cameras needed to be calibrated. We used the conventional stereo camera calibration method with a checker-board pattern. However, it is important to be careful regarding the different properties of the RGB and the IR cameras.

2. Reduction in the amount of reflected IR light from retro-reflective dots

The locations of the IR emitter and the IR camera on Kinect are not same. Thus, some IR light reflected from the dots made of the retro-reflective material does not reach the IR camera and, accordingly, the amount of reflected light captured by the IR camera decreases in comparison with the previous setting. It becomes more difficult to capture all dots in the range of the image. In order to mitigate this effect, we applied a different retro-reflective material with different reflectance properties (including more diffuse properties) for dots.

3. 3D information with low accuracy

Compared with Kinect v1, Kinect v2 provides highly accurate 3D information. However, its accuracy is lower still than that obtained from 3D information calculated by the optimized gray-code pattern method, its accuracy. In particular, it leads to a reduction in the accuracy of the interpolation process. In order to deal with this effect, we made some modification to the interpolation algorithm.

In the interpolation process described in Section 4.4.6, we applied a new weight value to each point neighboring one in process as a reliability. The range of a weight value varied from from 0.0 to 1.0, and the maximum number of points neighboring any given point is four. The weight value is 1.0 if the point is a recognized point. Following each interpolation, a weight value is set for each interpolated point. For example, if a point is interpolated by three recognized points, the value for this point is $(1.0 \times 3)/4 = 0.75$. If a point is interpolated by two recognized points and one interpolated point, the weight value of which is 0.5, the value for this point is $(1.0 + 1.0 + 0.5)/4 = 0.625$. This value is used as the weight coefficient in the selection of an appropriate pixel among those neighboring pixels of a point in process. With such a weight value, the reduction in interpolation accuracy due to the cumulative interpolation (the interpolation of interpolated points) can be mitigated to some extent.

We conducted a small test to verify the difference in processing speeds between a system with gray-code patterns (and OptiTrack camera) and Kinect v2. We used a 45×30 dot marker pattern, where the distance between any two neighboring dots was 26 mm. Table 4.3 shows the processing speed of each process with the two systems. The

4.7. Improving Processing Speed

speed of the 3D shape measurement process dramatically improved due to Kinect v2 and some modifications to the algorithm. Please note that because the resolutions of the Kinect cameras were different from the resolution of the OptiTrack camera (kinect v2 RGB camera: 1920×1080 , IR camera: 512×424 , OptiTrack RGB and IR camera: 1024×768), the processing times for marker recognition and interpolation slightly changed between the two systems.

4.7.2 For Marker Recognition and Interpolation

The bottlenecks in the process are marker recognition and interpolation. These processing times account for 84.7 % of the total processing time. Fortunately, our algorithms for marker recognition and interpolation are easy to parallelize because each small region is separately processed. It can be predicted that these processing time can be dramatically improved by using general-purpose computing on graphics processing units (GPGPU) parallelization. We address this problem in future work.

4.7.3 Application to Clothing Design Simulation System

In order to show the effects of the approaches proposed here for faster processing speeds for our system, we applied these to the system described in Section 4.6.6. Figure 4.22 illustrates the projection results. Sub-figures (a) show continuous projection results in case a user flipped up the edge of the clothes. The re-calculation and re-projection are conducted approximately 0.67 seconds. A user can then check various appearances of the target object without waiting long. In the same manner, sub-figure (b) shows results in case a user flipped up a sleeve of the clothes. Sub-figure (c) illustrates results in case a user suddenly rotated the target. We confirmed that the projected texture image also rotated in accordance with the rotation of the target object.

In the current implementation, no consecutive information between two frames was

Table 4.3. Processing time of each process

System - Process	3D Shape Measurement	Marker Recognition	Interpolation	Others	Total
With Gray-code and OptiTrack	5.04 [sec.]	0.42	0.21	0.05	5.72
With Kinect v2	0.033	0.30	0.27	0.070	0.673

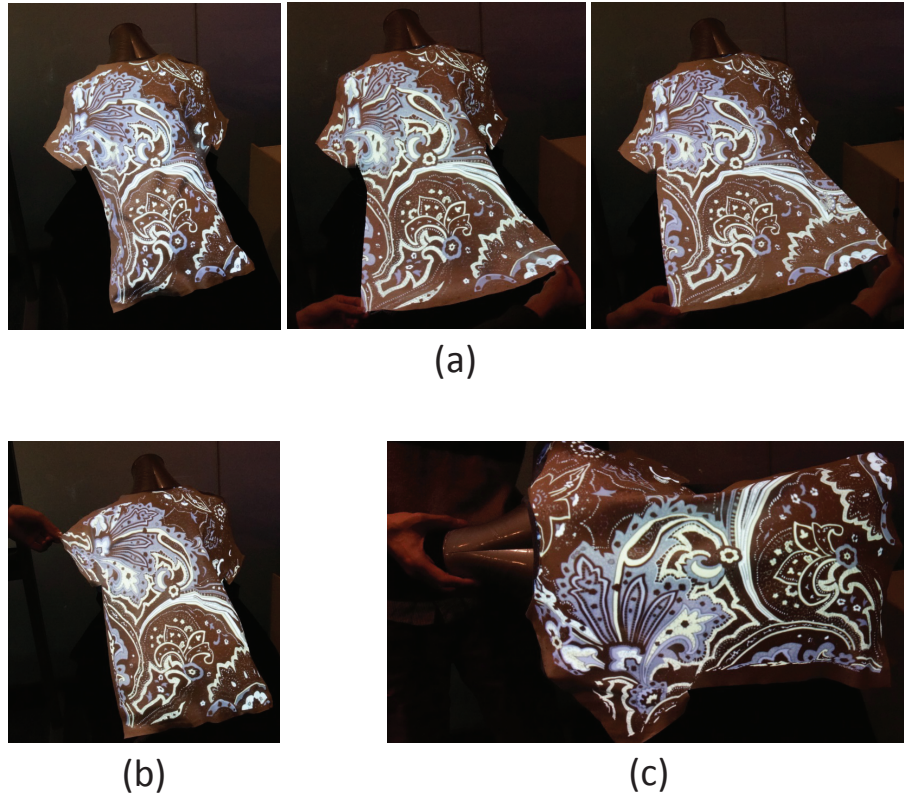


Figure 4.22. Application: Design support system (with kinect v2) for clothes

used. Thus, unnatural artifacts of the projected image were created due to temporal discontinuity and interpolation inaccuracy. Figure 4.23 shows this problem. A white region in a red rectangle waves unnaturally between consecutive temporal frames. This problem significantly detracted from the “experience of reality” of the users’ perceptions. We intend to solve this problem by applying temporal information (e.g., tracking for recognized points) in future work.

4.8. Summary and Future Work

Inspired by projection based-AR support for design as a promising application we found existing techniques are not available where the target objects need to be re-shaped. We proposed a new deformable material with an embedded marker pattern

4.8. Summary and Future Work

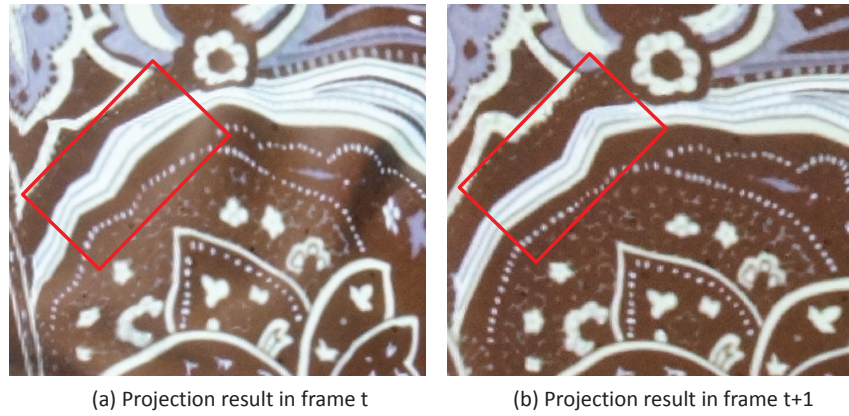


Figure 4.23. Examples of unnatural artifacts of the projected image derived from the temporal discontinuity and the interpolation inaccuracy

and a recognition method for the projection of texture-mapping. We have illustrated that our technique can project a geometrically correct texture onto the deformable material and can be reshaped with the user's hands. We have demonstrated and evaluated the effectiveness of our approach with a single and a dual pro-cam system. While rapid prototyping is presented as an example application, the new technology would be also suitable to fields including entertainment, training, and fashion.

In future work, we will explore improving the processing speed for the marker recognition and the interpolation process with the parallelization, described in Section 4.7.2. Under enough parallelization, we predict that this system will work at around 20 fps.

For the other direction, we will tackle to achieve the more realistic projection. Our objective is to enable a user to perceive the projected target object as if it was physically textured. We have to say the current quality of the projection is not sophisticated enough for this objective. In order to achieve that, applying the color calibration and the color compensation by considering the reflectance property of the target object should be worked as future work.

CHAPTER 5

Conclusion

This concluding chapter summarizes research problems that must be addressed and discuss our approaches. It also considers the contributions of this research to the field of augmented reality. Possible directions for future research are also provided.

5.1. Contributions and Summary

This study investigated the problems of current industrial task support systems based on AR. We focused on two main problems. First, the effectiveness of AR for memorization was verified through experiments. Second, we developed a design support system for deformable objects and a technical means of recognizing each location on the deformable object. Details of our study's contributions to these research topics are provided as follows.

5.1.1 Verification of Effectiveness of AR for Memorization (Chapter 3)

Through several experiments, we verified the effectiveness of AR for user memorization. We initially categorized memorization as it relates to location into two parts: 1)

5.1. Contributions and Summary

memorization with a location in the visual perception, and 2) memorization with a location in spatial cognition. We confirmed that displaying the relevant information near the location of a target object by means of AR increases user memorization more than when the information is displayed at an unrelated location. The confirmed hypotheses are as follows:

- Display information relevant to an object near it (so that a user can see both the information and the object without significantly moving of his or her eyes)
- Displaying the related information near the location of a target object through AR reduces the users' cognitive load on users in comparison with when it is shown at an unrelated location.

To the best of our knowledge, ours is the first study to address the issue of the effectiveness by AR in detail and quantitatively evaluate it. Furthermore, based on experimental results, we provided insights into the system design of information presentation through AR for effective memorization. Our insights are as follows:

- Display relevant information to a specific real object near it (so that a user can see both information and the object without large movement of his or her eyes).
- Do not display information at random locations (ones a user cannot predict)
- Display minimum required information to explain task process
- Provide two types of display methods and a function to switch between them at any time

5.1.2 Development of a Design Support System for Deformable Objects and a Technique for Identifying Each Location on a Deformable Object (Chapter 4)

For a geometrically-correct projection, we developed dot marker patterns that could be partially recognized by means of our recognition method. The marker pattern had suitable features such that each location on the entire marker pattern could be identified, even if it was attached onto the curved surface. Moreover, it allowed complicated

Chapter 5. Conclusion

deformations, such as ones involving folded and self-occluded highly deformable objects, such as pieces of cloth. Moreover, the marker pattern was suitable for projection scenarios because the area of the dot pattern that accounted for the entire surface of the object was small and nearly invisible to users due to the half-transparent retro-reflective material. The features of our technique are as follows:

- Our marker is almost imperceptible to users, which is a crucial feature for projection scenarios because these involve numerous points on a grid. The region occupied by the points on the whole surface of the target is a small portion. Furthermore, the marker can be made of a half-transparent retro-reflective material.
- Due to its parallelism, our algorithm inherently allows complicated deformations, which may occur in assumed applications, such as scenarios involving objects that are folded, cut, occluded by other objects, or self-occluded.
- It is easy to interpolate the unrecognized region of our marker using 3D information because each dot is placed on the grid, and equidistance property of consecutive two points can be used for the interpolation.

Based on this marker recognition method and the conventional 3D shape measurement method, we constructed a design support system for the production of t-shirts was constructed. Compared to previous system that used projectors to change the appearances of clothes, our method had the advantage that it did not require the predetermination of the motion of the target object. By means of this system, a user can freely and quickly change a texture image to project onto a t-shirt as well as the shape of the target object during the design procedure.

5.2. Future Research

This section concludes this study by discussing directions for further research.

5.2.1 Verification of Effectiveness of AR for Memorization

- All experiments conducted for this study involved situations in which participants sat on chairs and performed tasks at a table without producing expensive bodily movements. These experiments were performed relatively small environments.

Future studies involving the same type of experiments should be conducted in larger settings in which participants can walk around freely. This will require that participants possess different degree of spatial cognition which, we believe, will allow us to further clarify the relationship between AR features and human memorization.

- In this study, we analyzed attention to the AR effectiveness for short-term memory. We predict that learning through visualization and association of information using AR will be useful for long-period memory retention. In addition, we hope to determine whether AR support affects memory retention in long-term memory by examining how soon a user learns complex task procedures and then forgets them. We want to accomplish this by examining situations in which AR support is both present and absent.

5.2.2 Development of a Design Support System for Deformable Objects

- The system currently operates at approximately 1.5 fps. Although the size of the marker pattern used affects it, this speed is not acceptable for some real-time applications. To apply the system to situations involving continual movement or deformation of target objects, increasing processing speeds is necessary. A promising direction for future research is the parallelization of each process with a general-purpose computing on graphics processing units (GPGPU) because the small number of dots is processed separately in our algorithm for marker recognition and interpolation. These processes are easy for parallelization. Another direction for future work is to apply a simple tracking methodology for recognized points. In the current system implementation, information in consecutive frames is not used. However, such information would not only increase processing speed by avoiding the repeat recognition of the marker which is already recognized at every frame, but it would also to alleviate unnatural artifacts of the projected image derived from temporal discontinuity.
- In the current system, the color and reflectance properties of the target objects are not considered. In order to improve the quality of the natural appearance of projected target objects, considering these factors is desirable.

Chapter 5. Conclusion

- In the current system, a texture image that is to be projected is prepared in advance. Providing a function which enables a user to produce a texture image to be projected in situ, which can be accomplished immediately and easily, will expand applicability of our system. For example, in the design process for clothes, performed on a computer display, when a user draws a cut line on the thumbnail image of a piece of cloth using such functions, the cut line is projected onto the corresponding location of the actual cloth, thus allowing the user to cut it accurately.
- Applying our technique to other practical applications is interesting. Besides its application to clothing design support, another promising application is digital draping. In digital draping, a user first covers the target object (e.g., a piece of furniture such as a cushion and a sofa) with a neutral-colored piece of cloth. Our marker pattern is attached to the cloth. The system then detects each location on an entire of the target object based on marker recognition and projects any texture images onto the target object so that users can confirm the various appearance of existing target objects and without having to consult a detailed 3D model or digital information about the target object.

Publication List

Journal Papers

1. Yuichiro Fujimoto, Ross T. Smith, Takafumi Taketomi, Goshiro Yamamoto, Jun Miyazaki, Hirokazu Kato, and Bruce H. Thomas. Geometrically-Correct Projection-Based Texture Mapping onto a Deformable Object, *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 20(4): 540-549, 2014 (related to Chapter 4).
2. Yuichiro Fujimoto, Goshiro Yamamoto, Takafumi Taketomi, Jun Miyazaki, and Hirokazu Kato. Relation between Displaying Features of Augmented Reality and User's Memorization, *Transactions on the Virtual Reality Society of Japan*, 18(1): 81-91, 2013 (in Japanese, related to Chapter 3).

International Conferences and Workshops (reviewed)

1. Yuichiro Fujimoto, Ross T. Smith, Takafumi Taketomi, Goshiro Yamamoto, Jun Miyazaki, Hirokazu Kato, and Bruce H. Thomas. Geometrically-Correct Projection-Based Texture Mapping onto a Deformable Object, *IEEE International Conference on Virtual Reality 2014 (VR2014)*, 2014 (Presentation for Journal Paper 1, related to Chapter 4).

Publication List

2. Yuichiro Fujimoto, Goshiro Yamamoto, Takafumi Taketomi, Jun Miyazaki, and Hirokazu Kato. Relationship between Features of Augmented Reality and User Memorization, In *Proceedings of IEEE International Symposium on Mixed and Augmented Reality 2012 (ISMAR2012)*. pages. 279-280, (Poster), 2012 (related to Chapter 3).
3. Yuichiro Fujimoto, Goshiro Yamamoto, Jun Miyazaki, and Hirokazu Kato. Relation between Location of Information Displayed by Augmented Reality and User's Memorization, In *Proceedings of Augmented Human International Conference (AH2012)*, pages. 93-100, 2012 (related to Chapter 3).

International Conferences and Workshops (not reviewed)

1. Yuichiro Fujimoto, Takafumi Taketomi, Goshiro Yamamoto, Jun Miyazaki, and Hirokazu Kato, Ross T. Smith, and Bruce H. Thomas. Geometrically-Correct Projection-Based Texture Mapping onto a Cloth, In *Proceedings of IEEE International Conference on Virtual Reality 2014 (VR2014)*, pages. 169-170, (Demonstration), 2014 (related to Chapter 4).
2. Yuichiro Fujimoto, Tomohisa Yamada, Takafumi Taketomi, Goshiro Yamamoto, Jun Miyazaki and Hirokazu Kato. Construction of a Projection-based Augmented Reality System based on Shape Measurement using Flexible Marker for Design Support, In *Proceedings of Korea-Japan Workshop on Mixed Reality*, 2013 (related to Chapter 4).
3. Yuichiro Fujimoto, Chu Cheng Tse, Makoto Fujisawa, Toshiyuki Amano, Jun Miyazaki, and Hirokazu Kato. AR Dictionary: Dynamic Annotation for English Words in Printed Documents. In *Proceedings of International Conference on Artificial Reality and Telexistence (ICAT2010)*, page. 239, (Demonstration), 2010.

Domestic Conferences and Workshops (not reviewed)

1. Yuichiro Fujimoto, Takafumi Taketomi, Goshiro Yamamoto, Jun Miyazaki, and Hirokazu Kato. Geometrically-Correct Projection-Based Texture Mapping onto

- a Deformable Object, In *Proceedings of Annual Conference of Virtual Reality Society of Japan*, pages. 238-241, 2013 (in Japanese, related to Chapter 4).
2. Yuichiro Fujimoto, Goshiro Yamamoto, Jun Miyazaki, and Hirokazu Kato. Relation between Location of Information Displayed by Augmented Reality and User's Memorization, In *DVD-ROM Proceedings of Workshop of Human Interface Society*, 2011 (in Japanese, related to Chapter 3).
 3. Yuichiro Fujimoto, Childs Brodrick, Goshiro Yamamoto, Jun Miyazaki, and Hirokazu Kato. The Effects of Location of Information Displayed by Augmented Reality on a User's Memorization, In *Proceedings of Annual Conference of Virtual Reality Society of Japan*, pages. 606-609, 2011 (in Japanese, related to Chapter 3).
 4. Yuichiro Fujimoto, Chu Cheng Tse, Makoto Fujisawa, Toshiyuki Amano, Jun Miyazaki, and Hirokazu Kato. Development of Augmented Reality Dictionary System and Evaluation of Effectiveness, In *Proceedings of Workshop of Pattern Recognition and Media Understanding*, PRMU2010-165, 2011 (in Japanese).

Acknowledgments

Five years ago, an introduction to ARToolKit, a library for Augmented Reality (AR) implemented by Professor Hirokazu Kato, led me to NAIST. It was one of the most impactful experiences of my life; as a result, I chose to abandon the mechanical engineering I had studied until that point, instead devoting my life to research in this field. I would like to give considerable thanks to Professor Kato for creating the key for this turning point in my life. In addition, despite my coming to his laboratory from a different field of research and my lack of knowledge of his field, he always provided me with valuable advice when I encountered difficulty, opening my eyes on how to conduct research, write papers and make presentations. His patient teaching encouraged me to work harder, and I was greatly influenced by his attitude toward other people and research, which affected not only my research but also my life-style. I can never thank him enough, and I will be eternally grateful.

I would also like to thank the members of my thesis committee, my co-supervisors Professor Naokazu Yokoya and Associate Professor Christian Sandor, and Associate Professor Daisuke Iwai. Their work in reviewing my thesis and their insightful comments and suggestions helped me to improve my thesis's overall quality.

Professor Naokazu Yokoya always provided me with insightful and constructive comments in my seminar presentations since I joined NAIST. His comments were truly helpful in improving both my master and doctoral theses. I truly appreciate him for this assistance.

Associate Professor Christian Sandor taught me in the last year of my doctoral

course, the most important year. His teaching and comments were always energetic and effective. Thanks to him, I learned what sort of attitude a top researcher should have. I would like to thank him for guidance and example.

Associate Professor Daisuke Iwai at Osaka University was in charge of my thesis committee. I am certain that the valuable comments I received from him, a researcher working on various excellent studies in the field of projector-based AR, improved the quality of my doctoral thesis. In addition, I acquired a fantastic internship with his recommendation and support. I would like to thank him for his help and for making this opportunity possible.

Assistant Professor Goshiro Yamamoto provided me with basic knowledge on how to live as a researcher, among many other vital things. I hold his attitude toward students and research styles as my ideal. I would like to thank him for this inspiration.

Assistant Professor Takafumi Taketomi always assisted me with concrete advice based on his comprehensive through knowledge on this research field. His advice is and will remain a valuable tool in my research life. I would like to thank him for this guidance.

I would also like to note that Assistant Professor Goshiro Yamamoto and Assistant Professor Takafumi Taketomi took me out for an informal dinner on a daily basis. This was impressive and meaningful to me, as I had never had such an experience with faculty before. The assistant professor's friendly attitudes have consistently motivated me during the last five years. I would like to express my deepest thanks to them again.

I am grateful to former Associate Professor Miyazaki (now Professor at Tokyo Institute of Technology). He always presented me with numerous critical questions and comments at lab meetings, all based on a different research field's point of view. This has surely improved the foundation of my research.

Special thanks are due to former Assistant Professor Toshiyuki Amano (now Associate Professor at Wakayama University) and Makoto Fujisawa (now Assistant Professor at University of Tsukuba) for their patient, thorough teaching when I was a master student first learning how to do research.

Former Assistant Professor Sei Ikeda (now Assistant Professor at Osaka University) gave me advice when I encountered difficulty in my research and he always showed me courtesies when I met with him at conferences. I appreciate his help.

Former Assistant Professor Yuki Uranishi (now Assistant Professor at Kyoto University) provided me with concrete advice on implementations. In addition, he helped

Acknowledgments

me a great deal with while I prepared an application form for a student grant. I still take advantage of what he taught me every time I write applications. I would like to thank him again.

I would also like to thank my collaborative researchers, Professor Bruce H. Thomas and Senior Lecturer Ross T. Smith at the University of South Australia. Our e-mails and face-to-face discussions taught me how to conduct collaborative research. Additionally, they taught me quite a bit about paper writing while we worked together on a conference paper. I am certain this experience improved my skills and will continue to help me in future.

I am grateful to appreciate Professor Petri Pulli at the University of Oulu, who approved my three-month stay at the university. My time at his laboratory, which has a completely different research style from that of Japan, opened my eyes to other ways of thinking.

Principal Research Scientist Paul Beardsley and Senior Research Engineer Anselm Grundhöfer at Disney Research Zürich offered me an internship opportunity. This was a fantastic experience, which allows me to expand my worldview through communication with many smart, capable researchers and designers. I warmly thank them for this experience.

I also would like to thank Dr. Julian Pilet at École Polytechnique Fédérale de Lausanne. He kindly accepted my visit to his laboratory, and answered my questions on a paper of his I referred to.

Of course, sincere thanks to every single person at the Interactive Media Design (IMD) Laboratory at NAIST!

First, I would like to thank the laboratory's secretary, Ms. Makiko Ueno, for her help with official documents and other office procedures, even when I tended to forget to submit documents on time.

Marina Oikawa, Igor Almeida, and Henry Chu were senior students in the laboratory since I joined NAIST. They were always kind and helpful when I struggled with English or had difficulties preparing for presentations and user studies. I do not know what I would have done without them. In addition, Igor and Marina gave me weekly English lectures, which were very interesting and instructive!. I am very thankful to them all.

Atsushi Keyaki (now Assistant Professor at Tokyo Institute of Technology) gave me essential advice on research. He also kindly agreed to proofread when I wrote research

papers and student grant application. His attitude toward his research and his friendly communication with everyone are a standard I now aspire to. I appreciate him.

I cannot thank Arno Lübke enough. He always showed consideration for me. His graciousness always motivated and encouraged me, and his humorous jokes and humour helped me overcome several difficult periods in my research.

I occasionally disturbed Tatsuya Fukui by forcing him to go eat ramen with me at midnight. I would like to apologize to him for that here.

Every time I wrote a research paper or other documents in English, I received help from foreign students. Without their assistance, I would not have been able to finish even a single research paper. In particular, I would like to thank Marc Ericson Santos and Jaakko Hyry.

Many of my fellow students have graduated in the time since I joined NAIST. My thanks extended to all current and former members of IMD lab. Friendship with them made my five years at NAIST much more meaningful.

I also wish to thank the current students who often played Jinro (‘One Night Werewolf’) with me. Spending this time with them encouraged and motivated me.

Finally, I owe my deepest gratitude to my parents, Shuzo Fujimoto and Miyuki Fujimoto who raised me for 27 years and gave me a supportive push when I told them I wanted to study for five years at NAIST.

Bibliography

- [1] Hugo Alvarez, Iker Aguinaga, and Diego Borro. Providing Guidance for Maintenance Operations Using Automic Markerless Augmented Reality System. In *Proceedings of International Symposium on Mixed and Augmented Reality (IS-MAR ' 11)*, pages 181–190. IEEE, 2011.
- [2] Shunichi Amari and Keiji Tanaka. *Brain Science for Cognition and Action (in Japanese)*. Publisher for The University of Tokyo, 2008.
- [3] ARToolKit. <http://www.hitl.washington.edu/artoolkit>. Accessed in November, 2014.
- [4] Ronald T. Azuma. A Survey of Augmented Reality. *Presence*, 6(4):355–385, 1997.
- [5] Oliver Bimber and Ramesh Raskar. *Spatial Augmented Reality - Merging Real and Virtual Worlds*. A K Peters Ltd., 2005.
- [6] Frank Biocca, Arthur Tang, David Lamas, Jehn Gregg, Robert Brady, and Ping Gai. How Do Users Organize Virtual Tools Around Their Body in Immersive Virtual and Augmented Environment?: An Exploratory Study of Egocentric Spatial Mapping of Virtual Tools in the Mobile Infosphere. Technical report, Media Interface and Network Design Labs, Michigan State University, 2001.

- [7] BMW. Research Projects - Virtual World Meets Reality. http://www.bmw.com/com/en/owners/service/augmented_reality_introduction_1.html. Accessed in November, 2014.
- [8] M. Bordegoni and C. Rizzi. *Innovation in Product Design - From CAD to Virtual Prototyping*. Springer, 2011.
- [9] John Burns and Chris J. Mitchell. Coding Schemes for Two-Dimensional Position Sensing. *Cryptography and Coding III*, pages 31–65, 1993.
- [10] Mary J. Carruthers. *The Book of Memory: A Study of Memory in Medieval Culture*. Cambridge University Press, 1997.
- [11] Anna P. Chatzimichali, Wim H. Gijselaers, Mien SR. Segers, Piet Van den Bossche, Hetty van Emmerik, Frido E. Smulders, Pieter P. Jonker, and Jouke C. Verlinden. Bridging the Multiple Reality Gap: Application of Augmented Reality in New Product Development. In *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 1914–1919. IEEE, 2011.
- [12] Marvin M. Chun and Yuhong Jiang. Contextual Cueing: Implicit Learning and Memory of Visual Context Guides Spatial Attention. *Cognitive Psychology*, 36(1):28–71, 1988.
- [13] Florian Echtler, Fabian Sturm, Kay Kindermann, Gudrun Klinker, Joachim Stilla, Joern Trilk, and Hesam Najafi. The Intelligent Welding Gun: Augmented Reality for Experimental Vehicle Construction. *Virtual and Augmented Reality Applications in Manufacturing, Chapter 17*, 2003.
- [14] Arne D. Ekstrom, Michael J. Kahana, Jeremy B. Caplan, Tony A. Fields, Eve A. Isham, Ehren L. Newman, and Itzhak Fried. Cellular Networks Underlying Human Spatial Navigation. *Nature*, 425(6954):184–188, 2003.
- [15] K. Anders Ericsson. Memory Skill. *Canadian Journal of Psychology*, 39(2):188–231, 1985.
- [16] Extend3D. Extend3d. <http://www.extend3d.com/en/>. Accessed in November, 2014.

Bibliography

- [17] Steven Feiner, Blair Macintyre, and Doree Seligmann. Knowledge-based Augmented Reality. *Communications of the ACM*, 36(7):53–62, 1993.
- [18] Mark Fiala. The SQUASH 1000 Tangible User Interface System. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR' 05)*, pages 180–181. IEEE, 2005.
- [19] Pierre Fite-Georgel. Is there a Reality in Industrial Augmented Reality. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR' 11)*, pages 201–210. IEEE, 2011.
- [20] Fitnect. Fitnec. <http://www.fitnec.hu/>. Accessed in January, 2015.
- [21] Pierre Fite Georgel, Pierre Schroeder, and Nassir Navab. Navigation tools for viewing augmented cad models. *Journal of Computer Graphics and Application*, 29(6):65–73, 2009.
- [22] Michihiko Goto, Yuko Uematsu, Hideo Saito, Shuji Senda, and Akihiko Iketani. Task Support System by Displaying Instructional Video onto AR Workspace. In *Proceedings of IEEE International Symposium on Mixed and Augmented Reality (ISMAR' 10)*, pages 83–90. IEEE, 2010.
- [23] Raphaël Grasset, Tobias Langlotz, Denis Kalkofen, Markus Tatzgern, and Dieter Schmalstieg. Image-Driven View Management for Augmented Reality Browsers. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR ' 12)*, pages 177–186. IEEE, 2012.
- [24] Jens Grubert, Daniel Hamacher, Rudiger Mecke, Irina Bockelmann, Lutz Schega, Anke Huckauf, Mario Urbina, Michael Schenk, Fabian Doil, and Johannes Tumler. Extended Investigations of User-Related Issues in Mobile Industrial AR. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR ' 10)*, pages 229–230. IEEE, 2010.
- [25] Damis Hassabis, Carlton Chu, Geraint Rees, Nikolaus Wweiskopf, Peter D. Molyneux, and Eleanor A. Maguire. Decoding Neuronal Ensembles in the Human Hippocampus. *Current Biology*, 19(7):546–554, 2009.

- [26] Steven J. Henderson and Steven K. Feiner. Evaluating the Benefits of Augmented Reality for Task Localization in Maintenance of an Armored Personnel Carrier Turret. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR ' 09)*, pages 135–144. IEEE, 2009.
- [27] Steven J. Henderson and Steven K. Feiner. Augmented Reality in the Psychomotor Phase of a Procedural Task. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR ' 12)*, pages 191–200. IEEE, 2012.
- [28] Andrew Hollingworth. Scene and Position Specificity in Visual Memory for Objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(1):58–69, 2006.
- [29] David Holman and Roel Vertegaal. Organic User Interfaces: Designing Computers in Any Way, Shape or Form. *Communications of the ACM*, 51(6):48–55, 2008.
- [30] IKEA. AR Furniture Catalog. http://www.ikea.com/ca/en/about_ikea/newsitem/2014catalogue. Accessed in November, 2014.
- [31] Yasushi Ikei and Hirofumi Ota. SROM: Spatial Electronic Mnemonics : Basic Concepts and the Characteristics of Registration Process. *Journal of The Virtual Reality Society of Japan (in Japanese)*, 14(2):241–249, 2009.
- [32] Yasushi Ikei, Hirofumi Ota, and Takuro Kayahara. Spatial Electronic Mnemonic: A Virtual Memory Interface. In *Proceedings of Conference on Human Interface*, pages 30–37. ACM, 2007.
- [33] Seiji Inokuchi, Kensuke Sato, and Fumihiko Matsuda. Range Imaging System for 3-D Object Recognition. In *Proceedings of the International Conference on Pattern Recognition*, pages 806–808. IEEE, 1984.
- [34] Hirotake Ishii, Z. Bian, Tomoki Sekiyama, Toshinori Nakai, Hiroshi Shimoda, Masanori Izumi, and Yoshitsugu Morishita. Augmented Reality Applications for Nuclear Power Plant Maintenance Work. In *Proceedings of International Symposium on Symbiotic Nuclear Power Systems for 21st Century (ISSNP)*. IEEE, 2007.

Bibliography

- [35] Hirotake Ishii, Satoshi Oshita, Weida Yan, Hiroshi Shimoda, and Masanori Izumi. Development and Evaluation of a Dismantling Planning Support System Based on Augmented Reality Technology. *Electric Journal of Nuclear Safety and Simulation*, 2(1):52–60, 2011.
- [36] Shahram Izadi, David Kim, Otmar Hilliges, David Molyneaux, Richard Newcombe, Pushmeet Kohli, Jamie Shotton, Steve Hodges, Dustin Freeman, Andrew Davison, and Andrew Fitzgibbon. KinectFusion: Real-time 3D Reconstruction and Interaction Using a Moving Depth Camera. In *Proceedings of Symposium on User Interface Software and Technology*, pages 559–568. ACM, 2011.
- [37] Kazuto Kamiyama, Terukazu Mizota, Vlack. Kevin, Hiroyuki Kajimoto, Naoki Kawakami, and Susumu Tachi. Gelforce: Applying Traction Field Sensation to Robot Finger. In *Proceedings of Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (World Haptics)*, 2005.
- [38] P. Kan and Hannes Kaufmann. High-Quality Reflections, Refractions, and Causatics in Augmented Reality and Their Contribution to Visual Coherence. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR'12)*, pages 99–108. IEEE, 2012.
- [39] Hirokazu Kato and Mark Billinghurst. Marker Tracking and HMD Calibration for a Video-Based Augmented Reality Conferencing System. In *Proceedings of International Workshop on Augmented Reality (IWAR)*, pages 85–94. IEEE, 1999.
- [40] Hirokazu Kato, Keihachiro Tachibana, Mark Billinghurst, and Michael Grafe. A Registration Method based on Texture Tracking using ARToolKit. In *Proceedings of Augmented Reality Toolkit Workshop (ART'03)*, pages 77–85. IEEE, 2003.
- [41] Tatsuyuki Kawamura, Tomohiro Fukuhara, Satoshi Murata, Hideaki Takeda, Yasuyuki Kono, and Masatsugu Kidode. Ubiquitous Memories: Associating Everyday Memory with Real World Objects Using a Touching Operation. *IEICE Transactions on Information and Systems (in Japanese)*, J88-D-I(7):1143–1155, 2005.

- [42] Peter Keitler, Benjamin Becker, and Gudrun Klinker. Management of Tracking for Industrial AR Setups. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR ' 10)*, pages 3–12. IEEE, 2010.
- [43] Robert M. Kitchin. Cognitive Maps: What are they and why study them? *Journal of Environmental Psychology*, 14(1):1–19, 1994.
- [44] Gudrun Klinker, Allen H. Dutoit, and Martin Bauer. Fata Morgana - A Presentation System for Product Design. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR ' 02)*, pages 76–85. IEEE, 2011.
- [45] Kazuhiko Kobayashi, Shinobu Ishigame, and Hideo Kato. Simulator of Manual Metal Arc Welding with Haptic Display. In *Proceedings of International Conference on Artificial Reality and Telexistence*, 2001.
- [46] Ernst Kruijff, J. Edward Swan II, and Steven Feiner. Perceptual Issues in Augmented Reality Revisited. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR ' 10)*, pages 3–12. IEEE, 2010.
- [47] Layar. Layar. <https://www.layar.com/>. Accessed in November, 2014.
- [48] Sang-Su Lee, Sohyun Kim, Bopil Jin, Eunji Choi, Boa Kim, Xu Jia, Daeop Kim, and Kun-pyo Lee. How Users Manipulate Deformable Displays as Input Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1647–1656. ACM, 2010.
- [49] Woohun Lee and Jun Park. Augmented Foam: a Tangible Augmented Reality for Product Design. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR' 04)*, pages 106–109. IEEE, 2005.
- [50] Ewald T. A. Maas, Michael R. Marner, Ross T. Smith, and Bruce H. Thomas. Quimo: A Deformable Material to Support Freeform Modeling in Spatial Augmented Reality Environments. In *Proceedings of IEEE International Symposium on 3D User Interfaces (3DUI)*, pages 111–112. IEEE, 2011.
- [51] Michael R. Marner, Ross T. Smith, Shane R. Porter, Markus M. Broecker, Benjamin Close, and Bruce H. Thomas. *Handbook of Augmented Reality*, chapter 10: Large Scale Spatial Augmented Reality for Design and Prototyping. Springer, 2011.

Bibliography

- [52] Sandy Martedi, Hideaki Uchiyama, Guillermo Enriquez, Hideo Saito, and Tsutomu Miyashita. Foldable Augmented Maps. *IEICE Transactions on Information and Systems*, 95(1):256–266, 2012.
- [53] Paul McIlroy, Shahram Izadi, and Andrew Fitzgibbon. Kinectrack: Agile 6-DoF Tracking Using a Projected Dot Pattern. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR' 12)*, pages 23–29. IEEE, 2012.
- [54] Metaio. Junaio. <http://www.junaio.com/>. Accessed in November, 2014.
- [55] Microsoft. Kinect v2. <http://www.microsoft.com/en-us/kinectforwindows/>. Accessed in November, 2014.
- [56] Matthias Milczynski, Thomas Hermann, Till Bovermann, and Helge Ritter. A Malleable Device with Applications to Sonification-based Data Exploration. In *Proceedings of International Conference on Auditory Display*, pages 69–76, 2006.
- [57] M.R. Mine, J. van Baar, A. Grundhofer, D. Rose, and Bei Yang. Projection-Based Augmented Reality in Disney Theme Parks. *Computer*, 45(7):32–40, 2012.
- [58] Hiroyoshi Morita, Kaanyasn Yajima, and Shojiro Sakata. Reconstruction Of Surfaces Of 3-D Objects By M-array Pattern Projection Method. In *Proceedings of the 2nd International Conference on Computer Vision (ICCV)*, pages 468–473. ACM, 1988.
- [59] Momoyo Nagase, Daisuke Iwai, and Kosuke Sato. Dynamic Defocus and Occlusion Compensation of Projected Imagery by Model-Based Optimal Projector Selection in Multi-projection Environment. *Virtual Reality*, 15(2):119–132, 2011.
- [60] Tomohiro Nakai, Koichi Kise, and Masakazu Iwamura. Use of Affine Invariants in Locally Likely Arrangement Hashing for Camera-Based Document Image Retrieval. In *Proceedings of the 7th international conference on Document Analysis Systems*, pages 541–552. Springer, 2006.

- [61] Ulrich Neumann and Anthony Majoros. Cognitive, Performance, and Systems Issues for Augmented Reality Applications in Manufacturing and Maintenance. In *Proceedings of IEEE International Conference on Virtual Reality (VR ' 98)*, pages 4–11. IEEE, 1998.
- [62] Dominic O’Brien. *Learn to Remember*. Duncan Baird Publisher, 2000.
- [63] John O’Keefe and Lynn Nadal. *The Hippocampus as a Cognitive Map*. Oxford University Press, 1978.
- [64] Manuel Olbrich, Harald Wuest, Patrick Riess, and Ulrich Bockholt. Augmented Reality Pipe Layout Planning in the Shipbuilding Industry. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR ' 11)*, pages 269–270. IEEE, 2011.
- [65] OptiTrack. FLEX 13. <https://www.naturalpoint.com/optitrack/>. Accessed in November, 2014.
- [66] James Ott. Maintenance executives seek greater efficiency. *Aviation Week & Space Technology*, 142(20):43–44, 1995.
- [67] Nils Petersen and Didier Stricker. Learning Task Structure from Video Examples for Workflow Tracking and Authoring. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR ' 12)*, pages 237–246. IEEE, 2012.
- [68] Nils Peterson and Didier Stricker. Learning Task Structure from Video Examples for Workflow Tracking and Authoring. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR ' 12)*, pages 237–246. IEEE, 2012.
- [69] Julien Pilet, Vincent Lepetit, and Pascal Fua. Fast Non-Rigid Surface Detection, Registration and Realistic Augmentation. *International Journal of Computer Vision*, 76(2):109–122, 2008.
- [70] Thammathip Piumsomboon, David Altimira, Hyung Kim, Adrian Clark, Gun Lee, and Mark Billingham. Grasp-Shell vs Gesture-Speech: A Comparison of Direct and Indirect Natural Interaction Techniques in Augmented Reality.

Bibliography

- In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR ' 14)*, pages 73–82. IEEE, 2014.
- [71] Shane R. Porter, Michael R. Marner, Ross T. Smith, Joanne E. Zucco, and Bruce H. Thomas. Validating Spatial Augmented Reality for Interactive Rapid Prototyping. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR' 10)*, pages 265–266. IEEE, 2010.
- [72] Jeffrey L. Posdamer and Martin D. Altschuler. Surface Measurement by Space-Encoded Projected Beam Systems. *Computer Graphics and Image Processing*, 18(1):1–17, 1982.
- [73] Parinya Punpongsanon, Daisuke Iwai, and Kosuke Sato. DeformMe: Projection-based Visualization of Deformable Surfaces using Invisible Textures. In *Proceedings of ACM SIGGRAPH Asia 2013 Emerging Technologies*. ACM, 2013.
- [74] Ramesh Raskar and Kok-Lim Low. Interacting with Spatially Augmented Reality. In *Proceedings of the 1st international conference on Computer graphics, virtual reality and visualisation*, pages 101–108. ACM, 2001.
- [75] Michael Reed. Prototyping Digital Clay as an Active Material. In *Proceedings of the 3rd International Conference on Tangible and Embedded Interaction*, pages 339–342. ACM, 2009.
- [76] Holger Regenbrecht. Industrial Augmented Reality Applications. *Emerging technologies of Augmented Reality: Interfaces and Design*, pages 283–304, 2006.
- [77] Gerhard Reitmayr, Ethan Eade, and Tom W. Drummond. Semiautomatic Annotations in Unknown Environment. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR ' 07)*, pages 67–70. IEEE, 2007.
- [78] Christoph Resch, Peter Keitler, and Gudrun Klinker. Sticky Projections - A New Approach to Interactive Shader Lamp Tracking. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR ' 14)*, pages 151–156. IEEE, 2014.

- [79] Joaquim Salvi, Joan Batlle, and El Mustapha Mouaddib. A Robust-Coded Pattern Projection for Dynamic 3D Scene Measurement. *Pattern Recognition Letters*, 19(11):1055–1065, 1998.
- [80] Joaquim Salvi, Jordi Pagés, and Joan Batlle. Pattern Codification Strategies in Structured Light Systems. *Journal of Pattern Recognition*, 37(4):827–849, 2004.
- [81] Christian Sandor, Andrew Cunningham, Arindam Dey, and Ville-Veikko Mattila. An Augmented Reality X-Ray System Based on Visual Saliency. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR '10)*, pages 27–36. IEEE, 2010.
- [82] Pedro Santos, Holger Graf, Timo Fleisch, and André Stork. 3D Interactive Augmented Reality in Early Stages of Product Design. In *Proceedings of International Conference on Human-Computer Interaction (HCI)*, pages 1203–1207, 2003.
- [83] Ralph Schoenfelder and Dieter Schmalstieg. Augmented Reality for Industrial Building Acceptance. In *Proceedings of IEEE International Conference on Virtual Reality (VR '08)*, pages 83–90. IEEE, 2008.
- [84] Björn Schwerdtfeger and Gudrun Klinker. Supporting Order Picking with Augmented Reality. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR '08)*, pages 91–94. IEEE, 2008.
- [85] Björn Schwerdtfeger, Rupert Reif, Willibald A. Günthner, Gudrun Klinker, Daniel Hamacher, Lutz Schega, Irina Böckelmann, Fabian Doil, and Johannes Tümler. Pick-by-Vision: A First Stress Test. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR '09)*, pages 115–124. IEEE, 2009.
- [86] Pekka Siltanen, Tommi Karhela, Charles Woodward, and Paula Savioja. Augmented Reality for Plant Lifecycle Management. In *Proceedings of International Conference on Concurrent Enterprising*, pages 4–6. IEEE, 2007.
- [87] Ross T. Smith, Bruce H. Thomas, and Wayne Piekarski. Digital Foam Interaction Techniques for 3D Modeling. In *Proceedings of Symposium on Virtual Reality Software and Technology*, pages 61–68. ACM, 2008.

Bibliography

- [88] Ross T. Smith, Bruce H. Thomas, and Wayne Piekarski. Tech Note: Digital Foam. In *Proceedings of Symposium on 3D User Interfaces (3DUI)*, pages 35–38. IEEE, 2008.
- [89] Sony. Eyepet. <http://www.eyepet.com/>. Accessed in November, 2014.
- [90] Sony. The Eye of Judgment. <http://us.playstation.com/games-and-media/games/the-eye-of-judgment-ps3.html>. Accessed in November, 2014.
- [91] Weibin Sun, Xubo Yang, Shuangjiu Xiao, and Wencong Hu. Robust Checkerboard Recognition for Efficient Nonplanar Geometry Registration in Projector-Camera Systems. In *Proceedings of ACM/IEEE International Workshop on Projector Camera Systems (PROCAMS '08)*, pages 2:1–2:7, 2008.
- [92] Istvan Szentandrasei, Michal Zacharias, Jiri Havel, Adam Herout, Marketa Dubska, and Rudolf Kajan. Uniform Marker Fields: Camera Localization by Orientable De Bruijn Tori. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR '12)*, pages 319–320. IEEE, 2012.
- [93] Tab. Tab. <http://corp.tab.do/>. Accessed in November, 2014.
- [94] Wei Tan, Haomin Liu, Zilong Dong, Guofeng Zhang, and Hujun Bao. Grasp. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR '13)*, pages 73–82. IEEE, 2013.
- [95] Arther Tang, Charles Owen, Frank Biocca, and Weimin Mou. Comparative Effectiveness of Augmented Reality in Object Assembly. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 73–80. ACM, 2003.
- [96] Bruce H. Thomas. *Augmented Reality Visualization Facilitating the Architectural Process - Using Outdoor Augmented Reality in Architecture, Design Construction*. Springer Netherlands, 2009.
- [97] Bruce H. Thomas. A survey of Visual, Mixed, and Augmented Reality Gaming. *Computers in Entertainment (CIE)*, 10(3):1–3, 2012.
- [98] Bruce H. Thomas, G. Stewart Von Itzstein, Rudi Vernik, Shane Porter, Michael R. Marner, Ross T. Smith, Markus Broecker, Benjamin Close, Sandy

- Walker, Sean Pickersgill, Steve Kelly, and Peter Schumacher. Spatial Augmented Reality Support for Design of Complex Physical Environments. In *Proceedings of Pervasive Computing and Communications Workshops (PERCOM Workshops)*, pages 588–593. IEEE, 2011.
- [99] Endel Tulving. *Episodic and Semantic Memory, Organization of Memory*. Academic Press, 1974.
- [100] Johannes Tümler, Rüdiger Mecke, Michael Schenk, Anke Huckauf, Fabian Doil, Georg Paul, Eberhard A. Pfister, Irina Böckelmann, and Anja Roggentin. Mobile Augmented Reality in Industrial Applications: Approaches for Solution of User-Related Issues. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR '08)*, pages 87–90. IEEE, 2008.
- [101] Hideaki Uchiyama and Eric Marchand. Deformable Random Dot Markers. In *Proceedings of 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR'11)*, pages 237–238. IEEE, 2011.
- [102] Hideaki Uchiyama and Hideo Saito. Random Dot Markers. In *Proceedings of IEEE Virtual Reality Conference (VR)*, pages 35–38. IEEE, 2011.
- [103] Jouke C. Verlinden, A. De Smit, Aernout WJ. Peeters, and Martijn H. van Gelderen. Development of a Flexible Augmented Prototyping System. In *Proceedings of International Conference on Computer Graphics, Visualization and Computer Vision (WSCG)*, pages 496–503, 2003.
- [104] P. Vuylsteke and A. Oosterlinck. Range Image Acquisition with a Single Binary-Encoded Light Pattern. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(2):148–164, 1990.
- [105] Wikitude. Wikitude. <http://www.wikitude.com/>. Accessed in November, 2014.
- [106] Ruigang Yang and Greg Welch. Automatic and Continuous Projector Display Surface Calibration Using Every-Day Imaging. In *Proceedings of International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG)*, 2001.

Bibliography

- [107] Mao Ye, Huamin Wang, Nianchen Deng, Xubo Yang, and Ruigang Yang. Real-time Human Pose and Shape Estimation for Virtual Try-On Using a Single Commodity Depth Camera. *IEEE Transactions on Visualization and Computer Graphics*, 20(4):550–559, 2014.
- [108] Feng Zhou, Henry Been-Lirn Duh, and Mark Billinghurst. Trends in Augmented Reality Tracking, Interaction and Display: A Review of Ten Years of ISMAR. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR' 08)*, pages 193–202. IEEE, 2008.
- [109] Zhenglong Zhou, Bo Shu, Shaojie Zhuo, Xiaoming Deng, Ping Tan, and Stephen Lin. Image-Based Clothes Animation for Virtual Fitting. In *Proceedings of SIGGRAPH Asia 2012 Technical Briefs*, pages 33:1–33:4, 2012.
- [110] Zhiwei Zhu, Vlad Branzoi, Michael Wolverton, Louise Yarnall, Girish Acharya, Supun Samarasekera, Rakesh Kumar, Glen Murray, and Nicholas Vitovitch. AR-Mentor: Augmented Reality Based Mentoring System. In *Proceedings of International Symposium on Mixed and Augmented Reality (ISMAR ' 14)*, pages 17–22. IEEE, 2014.
- [111] Hubert D. Zimmer, Harry R. Speiser, and Beate Seidler. Spatio-Temporal Working-Memory and Short-Term Object-Location Tasks Use Different Memory Mechanisms. *Acta Psychologica*, 114(1):41–65, 2003.
- [112] Michael Zollhöfer, Matthias Nießner, Shahram Izadi, Christoph Rehmann, Christopher Zach, Matthew Fisher, Chenglei Wu, Andrew Fitzgibbon, Charles Loop, Christian Theobalt, and Marc Stamminger. Real-time Non-rigid Reconstruction using an RGB-D Camera. *ACM Transactions on Graphics (TOG)*, 33(4), 2014.
- [113] Stefanie Zollmann and Oliver Bimber. Imperceptible Calibration for Radiometric Compensation. In *Proceedings of Conference of the European Association for Computer Graphics (EUROGRAPHICS 2007)*, pages 61–64, 2007.