Geometrically-Correct Projection-Based Texture Mapping onto a Deformable Object

Yuichiro Fujimoto, Ross T. Smith, Takafumi Taketomi, Goshiro Yamamoto,
Jun Miyazaki, Hirokazu Kato, and Bruce H. Thomas

Abstract—Projection-based Augmented Reality commonly employs a rigid substrate as the projection surface and does not support scenarios where the substrate can be reshaped. This investigation presents a projection-based AR system that supports deformable substrates that can be bent, twisted or folded. We demonstrate a new invisible marker embedded into a deformable substrate and an algorithm that identifies deformations to project geometrically correct textures onto the deformable object. The geometrically correct projection-based texture mapping onto a deformable marker is conducted using the measurement of the 3D shape through the detection of the retro-reflective marker on the surface. In order to achieve accurate texture mapping, we propose a marker pattern that can be partially recognized and can be registered to an object's surface. The outcome of this work addresses a fundamental vision recognition challenge that allows the underlying material to change shape and be recognized by the system. Our evaluation demonstrated the system achieved geometrically correct projection under extreme deformation conditions. We envisage the techniques presented are useful for domains including prototype development, design, entertainment and information based AR systems.

Index Terms—Projection-based augmented reality, deformable marker, product design support

◆

1 INTRODUCTION

This paper 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 recognizing the new form. A substrate material constructed from silicon rubber and aluminum mesh with embedded retro-reflective dots is presented to demonstrate the re-construction 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 [1]. 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 [8, 32]. 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 recalculated whenever new deformations are made. Previous tracking

- *Yuichiro Fujimoto is with Nara Institute of Science and Technology. E-mail: yuichiro-f@is.naist.jp*
- *Ross T. Smith is with University of South Australia. E-mail: Ross.Smith@unisa.edu.au*
- *Takafumi Taketomi is with Nara Institute of Science and Technology. E-mail: takafumi-t@is.naist.jp*
- *Goshiro Yamamoto is with Nara Institute of Science and Technology. E-mail: goshiro@is.naist.jp*
- *Jun Miyazaki is with Tokyo Institute of Technology. E-mail: miyazaki@cs.titech.ac.jp*
- *Hirokazu Kato is with Nara Institute of Science and Technology. E-mail: kato@is.naist.jp*
- *Bruce H Thomas is with University of South Australia. E-mail: Bruce.Thomas@unisa.edu.au*

algorithms [17, 28] 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 paper 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) [5]. 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 [9]. 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 [11, 18, 27], entertainment [26] and training [21]. 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 [2, 19]. 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 [15], 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 paper 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 Szentandrasi et al. marker system [25], that allows partially visible markers to be recognized. As an object is deformed, a 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 the object is bent in half. Not only does the digital paint maintain its registration, but the object may have additional digital paint added or removed repeatedly.

The paper begins with an overview of the important related work in both 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, and this is followed with the required extensions to a multi-pro-cam arrangement. A set of evaluation findings is reported to demonstrate the feasibility of our new approach. Finally a description of our new technology in design is discussed as a motivating example, and this is followed by some concluding remarks.

2 RELATED WORK

Three areas of related work are discussed that are relevant to this paper. 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.

Existing technologies such as the Microsoft Kinect and similar methods [31] 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.

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. [14] 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. [7] 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. [10] 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. [12] 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 [29] developed a marker tracking system based on random dots. 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) [16], 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 [28] 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. Mcilroy et al. [13] 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 [17] developed an early real-time method for detecting deformations on fabric material surfaces. 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.

Registering an organic shaped material has also been achieved using multiple markers. Mark Fiala [4] 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.

There are a number of systems that have explored techniques to capture deformable substrates with sensors embedded into the material. Michael Reed [20] explored developing 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

---

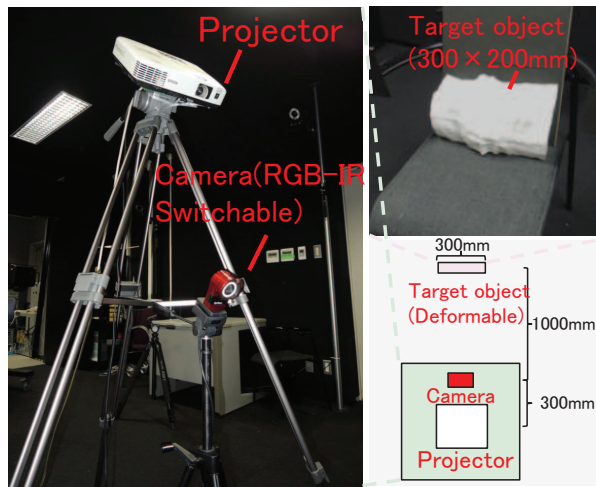[1] http://www.naturalpoint.com/optitrack/

Fig. 1. The major system components: Our system consists of projectors, cameras which are switchable from IR to RGB function, and deformable object.

(a) Image captured by RGB-camera   (b) Metal mesh

(c) Image captured by IR-camera   (d) Back side

(e) Examples of marker shapes

Fig. 2. Overview of deformable substrates: Our new SAR enabled deformable substrate

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. [23, 24] 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.

The work presented in this paper draws upon this previous research and aims to develop a technology that is capable of maintaining its shape after deformations, can capture the shape of the material and continues to operate when partial occlusions are present.

3 SYSTEM COMPONENTS

Our system is composed of three major components that are illustrated in Figure 1. The first sub-section describes the projector camera (procam) system. This is followed by a description of our new deformable object that provides the projection substrate. Finally the new marker pattern is described.

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 procam 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. The proposed 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 1 shows the component of only one pro-cam system, it is also possible to use more than one pro-cam system simultaneously (see Section 5).

3.2 Proposed 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 material to meet these goals shown in Figure 2. The deformable object consists of a metal mesh and a acrylic
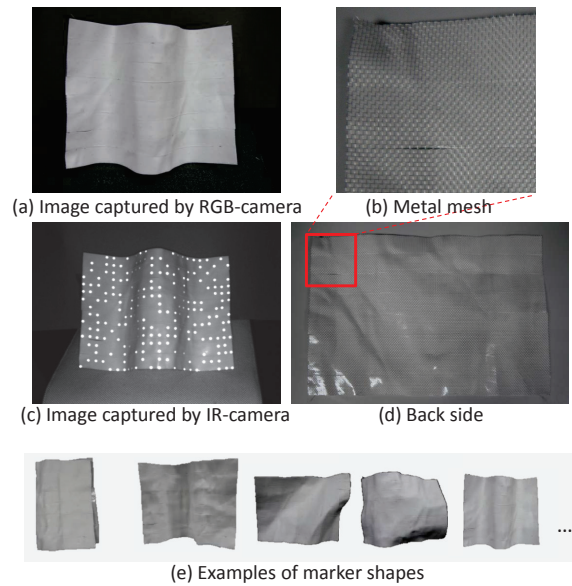
foam sheet. The metal mesh is made of aluminum, and the thickness is 0.2 mm. There are 18 holes per inch of its width and 16 holes per inch of its height. The acrylic foam sheeting has a thickness of 1.1 mm and is attached to the metal mesh, which makes it possible to easily reshape by hand. The metal 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 $\times$ 200 mm sheet in this investigation. However a larger sheet or a smaller sheet can be used to meet different requirements.

3.3 Proposed 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.

Szentandrasi et al. [25] 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 the reference pattern as long as any 4×4 rectangle sub-pattern within the marker is observed by the camera. In our investigation, we adopt a marker that has a similar property. Since the marker-pattern is assumed to be attached to a plane surface in the Szentandrasi 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 3 illustrates the difference between the Szentandrasi et al.'s marker and our new marker. We developed a fundamentally different 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 2(a)), the surface is observed
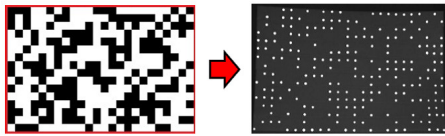
Fig. 3. Szentandrasi et al.'s marker on the left and our new marker system on the right : A minimum unit of Szentandrasi et al.'s marker is a rectangle. A minimum unit for our marker is a point.

as one with many retro-reflective points as shown in Figure 2(c).

In this research, the proposed 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 ALGORITHM FOR SINGLE PRO-CAM

Figure 4 illustrates an overview of our proposed 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 (a) depicts the case of a pro-cam that has the light 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 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 light axes of a camera and a projector are separated (normal pro-cams) are employed in this research, as shown in Figure 4 (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 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 dense a 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). 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 5 illustrates each step of the proposed 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.
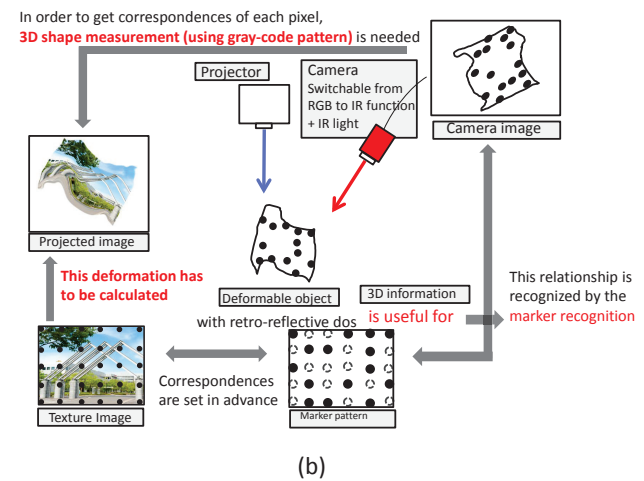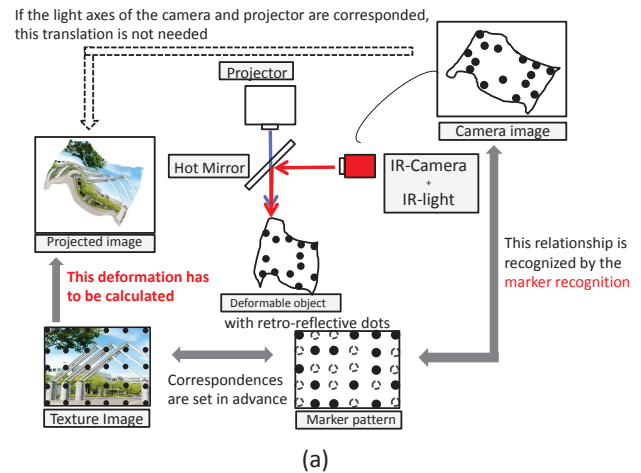

(a)

(b)

Fig. 4. Overview of marker detection algorithm: (a) a pro-cam which light axes of a camera and a projector are corresponded in advance (b) a pro-cam when light axes of a camera and a projector separated.

4.1 Marker Recognition

The steps for recognizing the described 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.

4.1.1 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 light 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.

4.1.2 Region division and grid calculation

This step aims to recognize a relationship between each of the points on the non-planar deformable surface. In order to deal with the chang-
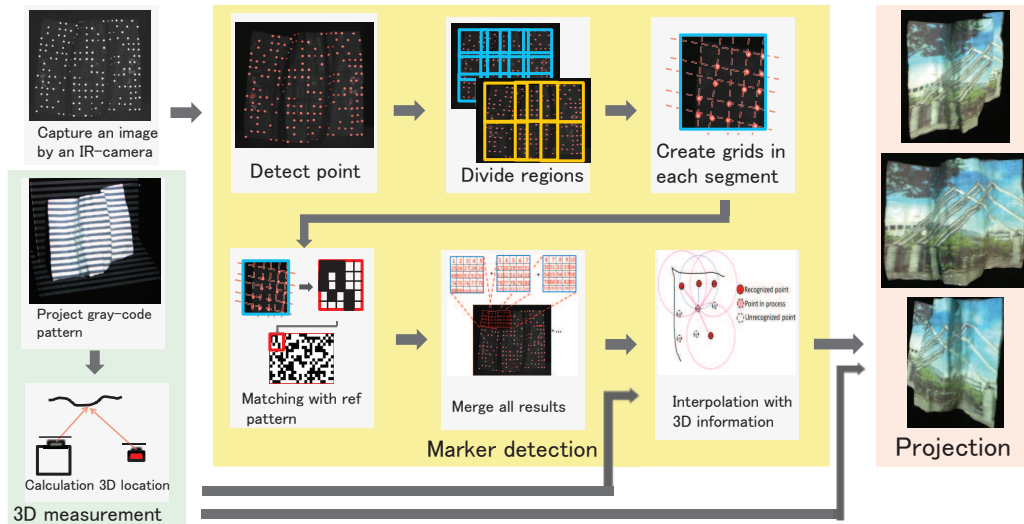
Fig. 5. Overview of system algorithm: Processing consists of four factors; 3D measurements, marker recognition, interpolation and projection
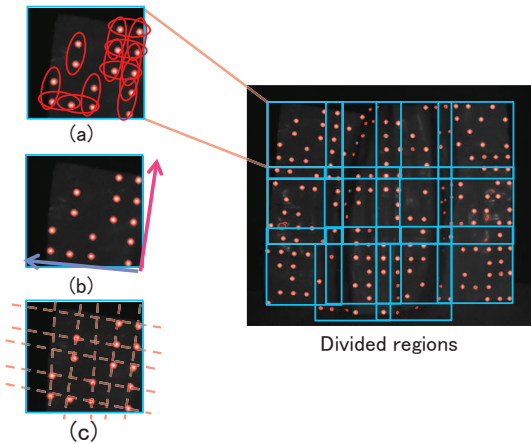

Fig. 6. Example of grid making: (a) By 3D information, each two consecutive points are detected in each small region. (b) Two salient directions are decided. (c) Grid lines are calculated through as many as possible of the detected points.

ing 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, see Figure 6. At first, the whole region is divided to rectangles. 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 × 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 6). The grid is determined from two salient directions from all possible directions of lines through pairs of points in the divided region. The process first creates pairs both consisting of 4-connective-points (shown in Figure 6(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 6(b)). Based on these two directions, grid lines are calculated through as many as possible detected

points in the region (shown in Figure 6(c)) for use with matching the reference pattern.

4.1.3 Matching with reference pattern

We reffer 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 $2^{16}/4(Up, down, left, right) = 16384$ (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 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.

4.1.4 Error Removal with Voting

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 7. 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. 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.

4.2 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 light axis of the IR-camera. Considering these factors, there are some cases the process has difficultly 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.
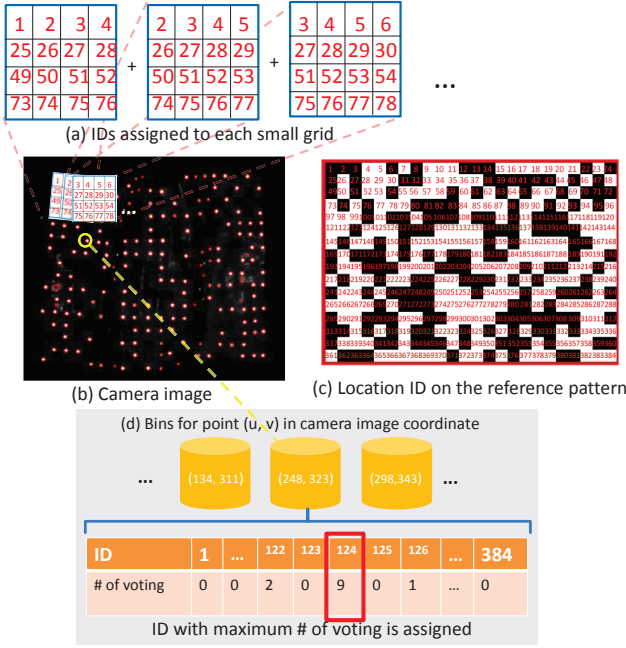
Fig. 7. Small grid integration and error removal: In order to integrate small grids, IDs are assigned to each point for voting. The ID with a maximum number of votes is assigned as the final ID of each point on the image

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 $13 \times \sqrt{2} \fallingdotseq 18.4$mm). It is possible to decide a required location on the surface of the object if some 8-connective-IDs has been assigned. Figure 8 (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 can not be decided.

# of recognized points: 2-3 If all points are in a row (vertical or horizontal), two candidates exist. A location can not 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.

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.

4.3 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

---

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:       $j_{inter} \Leftarrow \arg \min_j \sum_{i=1}^{k} (D_{ij} - const_i)$
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**

enough, the projection is conducted based on each patch. By translating locations in the camera image coordinate 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 9 and 10 illustrates projection results of the same texture image 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 - 8.04 secs, marker recognition - 0.41 secs, interpolation - 0.043 secs, and ready for projection - 0.040 secs.

4.4 Discussion

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. Another area of optimization is to use only one, either horizontal or vertical, patterns. While both horizontal and vertical code patters are employed in order to deal with various shapes and locations of the target object only one is needed to calculate 3D information. Additionally 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 these 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. Another solution is the use of a corresponding light 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
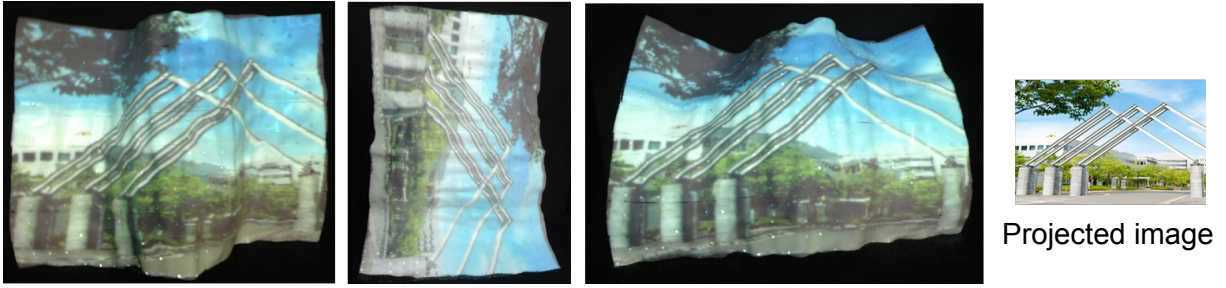
Projected image

Fig. 9. Results of projecting texture image, the same image is deformed and projected onto differently deformed surfaces
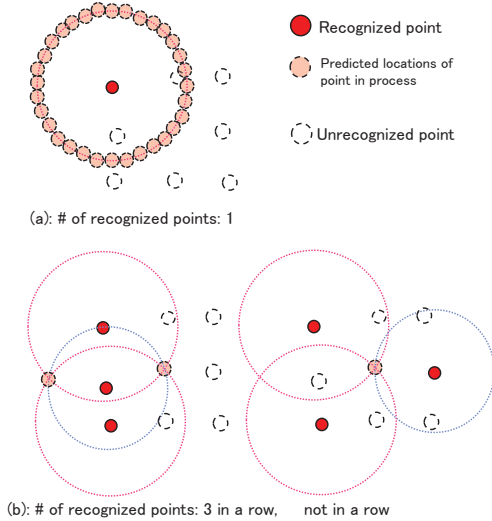

Fig. 8. Interpolation: (a) When # of recognized points is one, a location of a point which should be interpolated can not be predicted. (b) When # of recognized points is three and all points is in a row, two candidates exist. When all points are not in a row, a candidate can be determined uniquely.

Fig. 10. Results of projecting texture image when heavy occlusions occur

improvement of the processing speed. However some modifications are needed to make these algorithms work without the 3D information for the grid making process.

5 EXTENSION TO MULTI-PRO-CAMS

A practical system requires more than one pro-cam system to increase the range of the projection area. In this section, the algorithm for a multi-pro-cam system is described.

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. Our algorithm chooses which projector should be employed on each region observed by several cameras. 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) can be recognized by only one camera and the region is interpolated by several other cameras. When this occurs, the projector paired with the camera which recognized the region is chosen. A preferred projector must be chosen in cases where more than one option is available, such as when a region can be recognized by several cameras or when the region can only be interpolated by several cameras. A selection of projectors is conducted among projectors corresponded to each image which is used in the interpolation process. The angle between a light axis of each projector and normal vectors of the target object's surface is adopted as the criteria. The chosen projector has the minimum angle for the projection, as the projected image has the best resolution
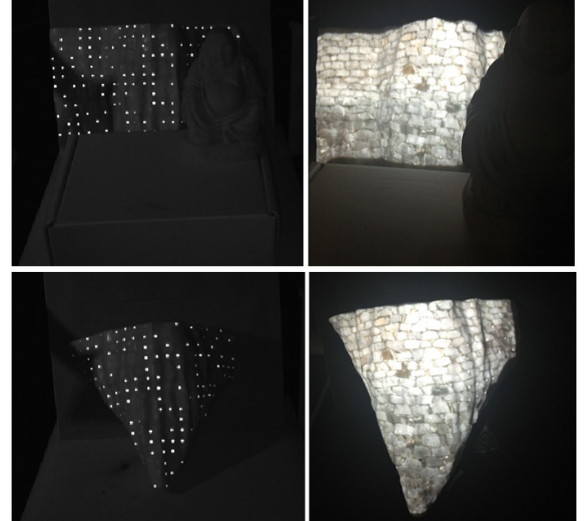
for users. We regard each region (surrounded by four locations of grid points) as a patch and decide which projector should be used for each patch.

We have developed an 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 1.

$$C = \frac{1}{n} \sum_{j=1}^{n} (p_i - \overline{p})(p_i - \overline{p})^T \qquad (1)$$

Here $\overline{p}$ 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 projectors 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.

6 RECOGNITION AND EVALUATION

In order to confirm the ability of the proposed 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. 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.

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 11 illustrates results for two shaped object. Double circles denote points decided by the marker recognition and single circle denote points decide by the interpolation.

6.1 Projection evaluation for single pro-cam

Table 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 11, 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.

6.2 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 12 (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 observed by one camera.

The procedure of the evaluation is the same to one for the single pro-cam and Figure 12 (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 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 of the pro-cams relative to the tracked object. However, it is predicted that using several pro-cams increases the whole projection accuracy.
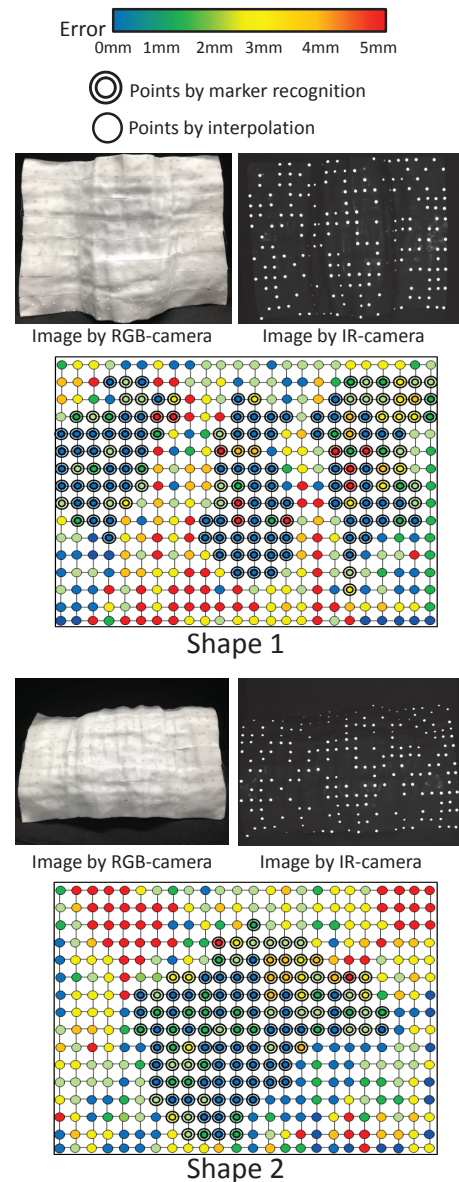

Fig. 11. Projection evaluation for single pro-cam: Each color represents the error value for each point. Double circles denote points decided by the marker recognition and single circle denote points decide by the interpolation.

7 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

Table 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

Table 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

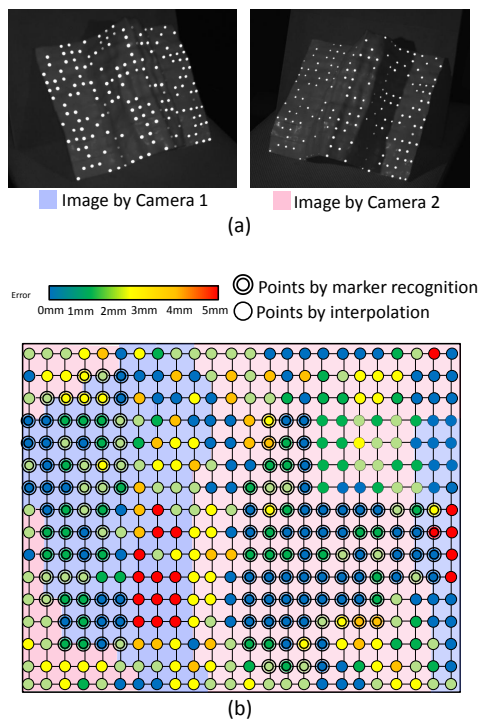| Projector	1	2
recognized point	0.99 (SD=1.20) mm	1.20 (SD=1.11) mm
interpolated location	2.88 (SD=2.28) mm	2.22 (SD=1.72) mm
sum	2.01 (SD=1.99)mm	1.97 (SD=1.44) mm

Fig. 12. Projection evaluation for multi pro-cams: Light blue and pink regions represents projected region by each projector

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.

7.1 Current design prototyping practice

Common in the design process of many different artifacts (for example white goods, 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 instantiate. 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.

7.2 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 [3, 22]. Verlinden et al. [30] 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.

7.3 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 physical shape [6], 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).

7.4 Our new design support

The system presented in this paper 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.

7.5 Digital draping

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.

7.6 Limitations

There are a number of limitations with the current technology. The first is the retro-reflective dots must maintain their spacing throughout the surface re-construction process. In the future we would like to investigate how accommodate materials that allow for stretching. A second limitation is shape of the surface maybe altered through bending. There are other processes the designer employs such as cutting,

---

[2]http://www.photoshop.com/

[3]The term digital draping was coined by Sean Pickersgill.

scraping, and carving. These pose two particular difficulties: firstly they maybe very fine in detail and secondly the act may remove the retro-reflective material.

8 CONCLUSION

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 reshaped. We proposed a new deformable material with an embedded pattern marker 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 with corresponding light axes of cameras and projector, and some modification in the algorithm to remove gray-code patterns that have the most impact on processing time. Even if the gray-code patterns are not needed, the marker recognition process also takes the non-negligible processing time (approximately 0.4 sec). Currently, the marker recognition has to be conducted again when an user moves the target object without any deformation. This operation may occur regularly during design and might adopt a position tracking algorithm that employs the retro-reflective dots without re-calculating the surface shape to improve the system performance.

REFERENCES

[1] R. T. Azuma. A survey of augmented reality. *Presence*, 6(4):355–385, 1997.
[2] O. Bimber and R. Raskar. *Spatial Augmented Reality - Merging Real and Virtual Worlds*. A K Peters Ltd., 2005.
[3] A. P. Chatzimichali, W. H. Gijselaers, M. S. Segers, P. Van den Bossche, H. van Emmerik, F. E. Smulders, P. P. Jonker, and J. 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.
[4] M. Fiala. The squash 1000 tangible user interface system. In *Proceedings of the 4th IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 180–181. IEEE Computer Society, 2005.
[5] D. Holman and R. Vertegaal. Organic user interfaces: designing computers in any way, shape or form. *Communications of the ACM*, 51(6):48–55, 2008.
[6] S. Izadi, D. Kim, O. Hilliges, D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A. Davison, and A. Fitzgibbon. Kinectfusion: real-time 3D reconstruction and interaction using a moving depth camera. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pages 559–568. ACM, 2011.
[7] K. Kamiyama, T. Mizota, V. Kevin, H. Kajimoto, N. Kawakami, and S. Tachi. Gelforce: Applying traction field sensation to robot finger. In *Proceedings of 1st Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (World Haptics)*, 2005.
[8] H. Kato and M. Billinghurst. Marker tracking and hmd calibration for a video-based augmented reality conferencing system. In *Proceedings of 2nd IEEE and ACM International Workshop on Augmented Reality, (IWAR)*, pages 85–94. IEEE, 1999.
[9] S.-S. Lee, S. Kim, B. Jin, E. Choi, B. Kim, X. Jia, D. Kim, and K.-p. 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.
[10] E. T. A. Maas, M. R. Marner, R. T. Smith, and B. 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, Los Alamitos, CA, USA, 2011. IEEE Computer Society.
[11] M. R. Marner, R. T. Smith, S. R. Porter, M. M. Broecker, B. Close, and B. H. Thomas. *Large Scale Spatial Augmented Reality for Design and Prototyping*. Springer, 2011.
[12] S. Martedi, H. Uchiyama, G. Enriquez, H. Saito, and T. Miyashita. Foldable augmented maps. *IEICE TRANSACTIONS on Information and Systems*, 95(1):256–266, 2012.
[13] P. McIlroy, S. Izadi, and A. Fitzgibbon. Kinectrack: Agile 6-dof tracking using a projected dot pattern. In *Proceedings of the 2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, ISMAR '12, pages 23–29, Washington, DC, USA, 2012. IEEE.
[14] M. Milczynski, T. Hermann, T. Bovermann, and H. Ritter. A malleable device with applications to sonification-based data exploration. In T. Stockman, editor, *Proceedings of the International Conference on Auditory Display*, pages 69–76, 2006.
[15] M. Mine, J. van Baar, A. Grundhofer, D. Rose, and B. Yang. Projection-based augmented reality in disney theme parks. *Computer*, 45(7):32 –40, 2012.
[16] T. Nakai, K. Kise, and M. 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.
[17] J. Pilet, V. Lepetit, and P. Fua. Fast non-rigid surface detection, registration and realistic augmentation. *International Journal of Computer Vision*, 76(2):109–122, 2008.
[18] S. R. Porter, M. R. Marner, R. T. Smith, J. E. Zucco, and B. H. Thomas. Validating spatial augmented reality for interactive rapid prototyping. In *Proceedings of 9th IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 265–266. IEEE, 2010.
[19] R. Raskar and K.-L. 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.
[20] M. Reed. Prototyping digital clay as an active material. In *Proceedings of the 3rd International Conference on Tangible and Embedded Interaction*, pages 339–342, 2009.
[21] H. Regenbrecht. Industrial augmented reality applications. *Emerging technologies of Augmented Reality: Interfaces and Design*, pages 283–304, 2006.
[22] P. Santos, H. Graf, T. Fleisch, and A. Stork. 3d interactive augmented reality in early stages of product design. In *Proceedings of 10th International Conference on Human-Computer Interaction (HCI)*, pages 1203–1207, 2003.
[23] R. T. Smith, B. H. Thomas, and W. Piekarski. Digital foam interaction techniques for 3d modeling. In *Proceedings of the ACM symposium on Virtual reality software and technology*, pages 61–68. ACM, 2008.
[24] R. T. Smith, B. H. Thomas, and W. Piekarski. Tech note: Digital foam. In *Proceedings of IEEE Symposium on 3D User Interfaces (3DUI)*, pages 35–38. IEEE, 2008.
[25] I. Szentandrasi, M. Zacharias, J. Havel, A. Herout, M. Dubska, and R. Kajan. Uniform marker fields: Camera localization by orientable de bruijn tori. In *Proceedings of 11th IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 319–320. IEEE, 2012.
[26] B. H. Thomas. A survey of visual, mixed, and augmented reality gaming. *Computers in Entertainment (CIE)*, 10(3):3, 2012.
[27] B. H. Thomas, G. S. Von Itzstein, R. Vernik, S. Porter, M. R. Marner, R. T. Smith, M. Broecker, B. Close, S. Walker, S. Pickersgill, S. Kelly, and P. 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.
[28] H. Uchiyama and E. Marchand. Deformable random dot markers. In *Proceedings of 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 237–238. IEEE, 2011.
[29] H. Uchiyama and H. Saito. Random dot markers. In *Proceedings of IEEE Virtual Reality Conference (VR)*, pages 35–38. IEEE, 2011.
[30] J. C. Verlinden, A. De Smit, A. W. Peeters, and M. 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.
[31] 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, feb 1990.
[32] F. Zhou, H. B.-L. Duh, and M. Billinghurst. Trends in augmented reality tracking, interaction and display: A review of ten years of ismar. In *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 193–202. IEEE Computer Society, 2008.