

## Topolo Surface: A 2D Fiducial Tracking System Based on Topological Region Adjacency and Angle Information

HIROKI NISHINO<sup>†1</sup>

In this paper, we describe Topolo Surface, a 2D fiducial tracking system we developed. Topolo Surface is a prototype system that implements a novel fiducial tracking method based on the combination of topological region adjacency and angle information. Existing systems based only on topological region adjacency information, such as D-Touch and ReactIVision, have several desirable features including fast processing speed and robustness against false positive detection. Yet, the method used in these systems also has several deficits. The unique ID range in existing topology-based methods is very narrow and the cost to generate the set of such unique fiducial markers can be computationally very expensive, especially when compared to existing matrix-based systems. Also, several useful techniques to improve robustness, such as CRC or hamming distance, cannot be applied to existing topology-based systems. Our novel fiducial tracking method utilizes the combination of topological region adjacency and angle information. By using topological information together with geometrical information, our prototype system achieved much larger unique ID range at very cheap computational cost to generate its fiducial markers. This is achieved while maintaining the desirable features of fast processing speed and robustness against false positives in a topology-based method. Also, CRC or hamming distance can be applied to our method to improve the robustness, if necessary.

### 1. Introduction

Fiducial tracking is one of the most prevalent methods used to design and implement an interactive environment today. Thanks to the recent improvement in the processing power and the cost of computers/video devices, camera-based user interfaces have become a significant technique in the field of tangible user interfaces and interactive media artworks.

The requirements for a fiducial tracking system can vary considerably between

target application fields. While those mixed/augmented reality applications mostly require 6DoF information or even marker-less tracking, it usually suffices for tabletop tangible user interfaces to provide much simpler 2D information, such as locations, rotation angles and movement vectors of the detected fiducial markers on the tabletop; Instead, a tabletop tangible user interface may require much faster processing speed, the robustness against false detection and a wide unique ID range.

Frequent false detection damages the stability of a system and its user interaction. The processing speed of fiducial tracking can seriously influence the quality of user experience, especially when it is necessary to detect quick gestures performed by a user. To implement a tabletop tangible collaborative environment, it is desirable to have a wide unique ID range to distinguish a large number of personalized objects owned by many users.

Among the available fiducial tracking systems, *reactIVision*<sup>2)</sup> and *D-Touch*<sup>3)</sup> are rapidly gaining considerable popularity in the community of tangible interactions and media art. To name a few major projects, there are *reactTable*<sup>10)</sup> (tangible interactive live computer music performance), *Turtan*<sup>8)</sup> (a LOGO-like tangible programming language) and *Physical Sequencer* (for live computer music performance)<sup>4)</sup>. **Figure 1** shows a picture of *reactTable* performance taken from their paper<sup>10)</sup>.

Both *reactIVision* and *D-Touch* are based on topological region adjacency for fiducial detection, while there are a number of matrix-based or pattern-matching-based fiducial tracking systems/libraries, such as *ARToolkit*<sup>11)</sup>, *ARToolkit Plus*<sup>14)</sup>, *ARTag*<sup>6)</sup>, *CyberCode*<sup>13)</sup>.

Topology-based approach has several desirable features for tangible tabletop interaction, such as fast processing speed and robustness against false detection, and how it achieves such features significantly differs from matrix-based method or pattern-matching-based method as described in the later sections.

However, in those existing systems based only on topological region adjacency, the number of unique fiducial markers is much smaller than most of the matrix-based systems and the extension of unique ID range can increase the actual size of fiducials. It is also computationally expensive to generate a large set of unique fiducial markers that can be stably detected. Furthermore, several

<sup>†1</sup> Graduate School for Integrative Sciences & Engineering, National University of Singapore

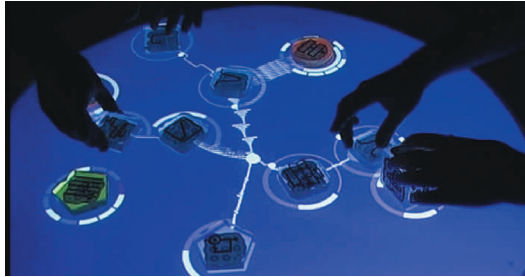


Fig. 1 *reacTable* performance<sup>10)</sup>.



Fig. 2 Examples of the markers from *CyberCode*<sup>1)</sup>, *ARToolkit*<sup>11)</sup>, *ARToolkit Plus*<sup>14)</sup> and *ARTag*<sup>7)</sup> (L to R).

digital techniques widely used in matrix-based systems, such as CRC or Hamming distance, are not applicable to topology-based method.

Such shortcomings of the existing topology-based approaches present a serious obstacle to implementing an interactive system that needs to handle a large number of objects and users.

To overcome such a deficit of a topology-based approach while maintaining its merits, we developed *Topolo Surface*, a prototype system that implements a novel fiducial tracking method based on the combination of topological region adjacency and angle information.

By such a combination of topology information and geometric information, our method largely extends the unique ID range at almost no computational cost for fiducial generation, without much increase in the actual size of fiducials. Our method also makes it possible to apply several beneficial techniques like CRC or hamming distance to improve the robustness.

These features of topology-based approach, such as fast processing speed and the robustness against false positives are still maintained without any significant damage.

## 2. Related Works

In this section, we briefly review existing matrix-based systems and pattern-matching-based systems to contrast them to topology-based systems. Then, we describe existing topology-based systems in more detail, to clarify their deficits when compared to matrix-based systems.

We also describe the techniques and strategies to achieve sufficient robustness against false detection in both matrix-based and topology-based approaches, since there is a significant difference between them.

### 2.1 Matrix-Pattern and Pattern-matching

The use of 2D matrix or pattern-matching for fiducial markers can be frequently seen in existing fiducial tracking systems. There is already a substantial body of the previous research of this type. **Figure 2** shows the examples of the fiducial markers in this category.

Generally speaking, a fiducial marker using a 2D matrix usually encodes 1 bit for each cell of its matrix. Most of those matrix-based fiducial markers have a larger data capacity or a wider unique ID range than do existing topology-based fiducial markers. *CyberCode* has 24-bit data capacity<sup>13)</sup>. *ARTag* has a set of 2002 predefined fiducial markers<sup>6)</sup> and *ARToolkit Plus* has 4096 fiducial markers<sup>14)</sup>.

However, such fiducial markers with matrix patterns usually require some techniques to improve robustness against false detection, as described later.

*ARToolkit* uses pattern-matching for fiducial detection and provides a *confidence value* as the probability to be any particular marker, to distinguish each fiducial from the others. In such a pattern-matching method, it is desirable to use those patterns that are different from the other patterns as possible, to avoid inter-marker confusion.

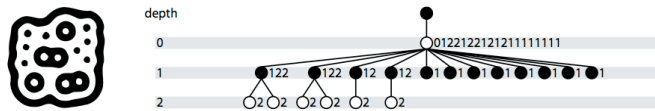
### 2.2 Topological Region Adjacency

There are several existing systems that use topological region adjacency information for fiducial detection, such as *RAG Target* by Johnston<sup>9)</sup>, *D-Touch* by Costanza<sup>3)</sup>, and *reacTIVision* by Kaltenbrunner and his colleagues<sup>2)</sup>. **Figure 3** shows several examples from these existing systems.

Fiducial detection methods in such topology-based systems significantly differ



**Fig. 3** The fiducial markers of *RAGTarget*<sup>9)</sup> (left), *reacTIVision*<sup>2)</sup> (2<sup>nd</sup> from left) and *D-Touch*<sup>5)</sup> (right 4).



**Fig. 4** A *reacTIVision* fiducial marker (left) and its topological region adjacency tree (right)<sup>2)</sup>.

from those of matrix-based systems. We describe the basic method, taking *reacTIVision* as an example.

**Figure 4** shows an example of a *reacTIVision* fiducial marker on the left. In a binarized image, the region adjacency information can be expressed as containment information of the black/white regions, taken from Ref. 2). A *reacTIVision* fiducial marker on the left can be expressed as the topological region adjacency graph on the right.

In those systems based on topological region adjacency information, a fiducial marker is designed to have its own unique topological region adjacency structure, which is distinct from any other fiducial markers in the system. Such a unique topological structure is mapped to its own unique ID.

In *reacTIVision*, a canonical sequence called *left heavy depth sequence* is used to express such a tree of topological region adjacency. For instance, the *reacTIVision* fiducial markers on the left can be expressed as *left heavy depth sequence*, 012212212121111111. Then it finds a unique ID that is mapped to such a *left heavy depth sequence* from the predefined database.

Such topological region adjacency information can easily be obtained just by segmenting the binarized image and the time cost to such a fiducial candidate is not so expensive. Also, given a sufficiently complicated topological structure to a fiducial marker, it can easily achieve robustness against false positive detection

even without digital techniques, unlike those matrix-based fiducial markers.

However, the number of unique fiducial markers in such a topology-based system is considerably smaller than that of matrix-based systems. For instance, Johnston's *RAG Target* has only 449 unique fiducial markers, and *reacTIVision* has only 216. *D-Touch* has several different types of fiducial markers, as shown in Fig. 3. The numbers of unique fiducials in *D-Touch* are 1, 6, 24 and 120 for each type (from left 3<sup>rd</sup> to the right most in Fig. 3). Such narrow unique ID ranges in a topology-based system mainly derive from its fiducial detection method.

Since the uniqueness of each fiducial marker depends directly on the uniqueness of its topological region adjacency structure, the ID range is limited by the possible combination of available sub graphs. To extend its unique ID range wider, the number of subgraphs in a fiducial may need to be increased. However, the addition of such subgraphs can also enlarge the size of a fiducial marker.

Such addition of more fiducial markers may also increase inter-marker confusion as there are more chances to have similar topological region adjacency structures among its fiducial markers. Also the generation cost of the large set of fiducial markers is another problem. We describe these issues in the next section.

### 2.3 Techniques to Improve Robustness

There are significant differences in methods to achieve the required robustness against false detection between those fiducial tracking systems based on matrix patterns and those based on topological region adjacency.

Generally speaking, fiducial markers that encode data into matrix patterns have a trade-off relationship between the complexity of matrix patterns and the robustness against false detection. The false detection rate can considerably increase as matrix patterns become more complicated and dense; for instance, a white cell in the matrix can be misinterpreted as a black cell when the quality of input image is not sufficiently good. Such misinterpretation can frequently occur, especially when the cells needs to be placed densely to encode more bit data inside a matrix pattern.

To improve the robustness against such false detection, most fiducial markers with matrix patterns employ digital techniques. Check-sum or CRC are frequently applied to validate the data decoded from matrix-patterns, as seen in *CyberCode*<sup>13)</sup>. *ARTag* also applies hamming distance to the data to be encoded

into its fiducial markers to improve the robustness against inter-marker confusion rate<sup>6)</sup>.

While such digital techniques play a significant role in the matrix-pattern based fiducial markers, a topology-based method takes very different strategies to achieve robustness in fiducial detection.

In existing methods based only on topological region adjacency information, the robustness against false positive detection largely depends by the rarity of the topological structures of the fiducial markers in input.

Unlike those fiducial markers with matrix patterns, topology-based approach can achieve significant robustness against false positives, just by giving enough complicated structures to its fiducial markers.

As seen in *reactIVision*, a topological region adjacency structure with a depth of at least 3 and 19 nodes is sufficiently rare to avoid most of false positive detections<sup>2)</sup>. Such a relationship between the complexity of topological structures of fiducials and the robustness against false positives and false negatives is also described in Ref. 3) in detail.

However, unlike matrix-based systems, such a topology-based method cannot utilize check-sum or CRC to validate detected fiducial markers. The uniqueness of each topology-based fiducial marker depends solely on the uniqueness of its own topological region adjacency structure. So there could be more chances for inter-marker confusion if there are any topological structures alike among the other fiducial markers in the same system. For instance, in the case of *reactIVision*, a topological structure with *left heavy depth sequence* of 012212212121111111 can easily be misinterpreted as 01221221212111111 if one of the nodes in the depth level 1 happened to absorb another node in the same level. (Notice that there ten '1' digits in the later *left heavy depth sequence*, while there are eleven in the former sequence).

To avoid such false detection in a topology-based approach, it is very desirable to avoid the use of those topological region adjacency structures alike to any of the others. False detection rate in a topology-based approach also depends on the actual placement of each region in a fiducial, and it is important to find a good fiducial marker design for a given topological structure. It is desirable to find a design that can be stably detected and at the same time has a similar size to

the other fiducial markers, not to let the sizes of the fiducial markers in the same system vary too much.

These deficits in a topology-based approach also can largely increase the generation costs of the set of fiducial markers. To generate the set of 180 fiducial markers, *reactIVision* involved several techniques, including genetic algorithm, to design the topological structures and their actual placements of each node that can be stably detected with the similar marker size as possible. As described in Ref. 2), such generation can be computationally very expensive, and even required 12 hours for only 128 fiducial markers on 11 PCs, each with dual Pentium 3 CPUs.

Such a generation cost can be even more expensive as the number of unique fiducial markers increase, to find those topological region adjacency structures distinct from the others and that can be placed in a space as small as possible. Even though, such a topology-based method still lacks any validation techniques that can detect falsely-detected fiducial markers.

### 3. Description of Our Algorithm

In this section, we describe our fiducial detection method based on the combination of topological region adjacency information and angle information.

By a combination of topological information and geometrical information, our method can implement several desirable features lacking in existing methods based only on topological region adjacency information, while maintaining its fast processing speed and the robustness against false positive detection.

Such desirable features include a wider unique ID range, constant fiducial marker size, cheap generation cost of fiducial markers, and applicability of digital techniques to improve the robustness. The lack of such features has been the deficits in the existing topology-based systems.

#### 3.1 Fiducial Marker Design

**Figure 5** is a picture of three examples of our fiducials (From left to right: 16-bit, 12-bit, 8-bit). Our fiducial markers have a circular shape as shown.

Our fiducial consists of two main components, *centre area* and *data area*. The first component is *centre area*, which is the black circle in the centre that contains a white circle within. This part is used to estimate the approximate rotation



Fig. 5 Three examples of our fiducials.



Fig. 6 The images from each phase in our system; input image, gray-scaling, binarization and detection (from L to R).

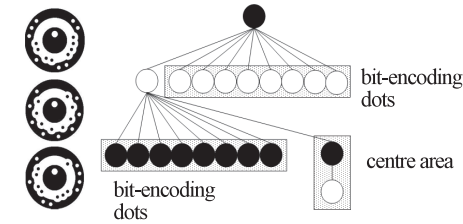


Fig. 7 An example of 3 different fiducial markers that share the same topological region adjacency tree.

angle, which is required in the decoding phase.

The second component is *data area*, which surrounds *centre area*. This *data area* is composed of an inner white ring-like region that contains black dots and an outer black ring-like region that contains white dots. These black and white dots are *bit-encoding dots* and each dot encodes 1 bit of data, 0 for black and 1 for white in clockwise order. Applying this decoding method, those fiducial in Fig. 5 can be decoded as 43690, 1017, 115 respectively from left to right.

Thus, our method directly encodes a unique ID into each fiducial, whereas *reactIVision* and *D-Touch* both require mapping from the topological region adjacency structure of a fiducial to its unique ID.

### 3.2 Fiducial Detection

Figure 6 shows images from each phase from our fiducial tracking method.

Basically speaking, our method uses the topological region adjacency information to find fiducial candidates in input image and uses the angle information to decode their unique IDs.

First, we binarize a frame of input image and segmentize it to obtain a region adjacency information. Then, we seek for a topological region adjacency structure that is likely to be a fiducial. In case of 16 bit setting, there can be 17 different topological structures as mentioned.

After finding a fiducial candidate, *centre area* is used to obtain the approximate rotation angle, using the vector from the centre of this area to the centre of the white circle within it.

Then, the coordinates of those *bit-encoding dots* in *data area* are obtained and sorted in clockwise order, using the approximate rotation angle as an origin. Then the data inside a fiducial can be recovered translating each *bit-encoding dot* to 0 and 1, according to its color. Any fiducial candidate with invalid angles between *bit-encoding dots* is rejected.

All the information required for this decoding can be obtained when we extract topological region adjacency information in the segmentation phase.

One of the merits in our method is that it can encode different unique IDs even into the same topological structure. Figure 7 shows three 16 bit fiducial markers, each with a different unique ID (58563, 43690, and 37662). Notice all these fiducials have exactly the same topological region adjacency graph as described in Fig. 7.

In case of 16 bit fiducial markers, there are only 17 different topological structures for 65536 unique fiducials, because the number of 0 (or 1) can vary only between 0–16 in 16 bit binary data; contrastingly, *reactIVision* requires 108 topological structures to express 216 unique fiducials ( $= 108 \times 2$  different coloring of black/white).

Because of such a data-encoding method, our fiducial can have a constant size for each fiducial marker. There is also almost no computational cost to generate a large number of fiducial markers; it only requires the very simple translation and the placement of bit data into black/white bit-encoding dots in clockwise order.

This is a considerable merit compared to the existing systems based only on topological region adjacency, in which generation of the large set of fiducial markers can be very computationally expensive.

### 3.3 Robustness in Our Algorithms

Since our method finds fiducial candidates by topological region adjacency information, the merit of the robustness against false positive detection in the topology-based fiducial tracking is still maintained to some degree.

The topological region adjacency tree of our fiducial markers has the depth of 4 and 20 nodes in 16-bit data capacity setting. This topological structure is likely complicated enough to avoid most false positives, since such a complex topological structure can hardly exist except in an actual fiducial marker. (In *reactIVision*, the topological region adjacency tree for each fiducial marker has at least with a depth of 3 and 19 nodes.)

However, in our marker design, *bit-encoding dots* are actually the leaf nodes in the upper level of the topological region adjacency tree as shown in Fig. 7. This may result in less robustness. As the result of Costanza's experiment in Ref. 3) suggests, the presence of empty nodes in the upper level of the topological region adjacency tree can increase the number of false positives and unbalanced tree structures can also increase the false positives.

Since our method also utilizes the angles between the *bit-encoding dots* to filter out invalid fiducial candidates, it can be significantly more robust against false positive detection than if it used topological region adjacency alone without angle information.

As described in the following evaluation section, even when a false positive with a valid topological structure is found in input image, such a false positive is mostly rejected because of invalid angles between *bit-encoding dots*.

Another significant merit in our method is applicability of digital techniques to improve the robustness against inter-marker confusion. Since we encode an ID directly to a fiducial, we can apply such digital techniques as CRC, parity-bit, check-sum and hamming distance as can be already done with matrix-based markers.

It may narrow the unique ID range in our fiducial to apply such digital techniques, yet there is still a much wider unique ID range compared to existing topology-based systems; Even if we sacrifice 4 bit for CRC in 16 bit fiducial markers, there are still  $2^{12} = 4096$  unique fiducials.

As described in the evaluation section, our method can achieve sufficient

robustness for tangible interaction, as in those existing topology-based systems and it is usually not necessary to apply such digital techniques.

However, such applicability of the existing digital techniques is a significant merit in our method, compared to the *reactIVision* or *D-Touch*. In both systems, such digital techniques to validate detected ID cannot be applied to avoid inter-marker confusion or other false detections.

As above, the combination of topology information and geometrical information in our method can improve the robustness in a topology-based approach.

## 4. Evaluation of Our Method

We evaluated our new fiducial tracking technique by a series of experiments. We also compare our prototype system to *reactIVision*, since it is one of the most widely used fiducial tracking systems for tangible interaction, based on the topological region adjacency information. We used 16 bit setting for all these experiments. The experiments in this section were done in the environments shown in **Table 1**, in  $640 \times 480$  resolutions.

### 4.1 Unique ID Range

One of the significant merits in our method is much wider unique ID range than those existing topology-based systems. It is also wider than most of those matrix-based systems, as shown in **Table 2**.

### 4.2 Processing Speed

First, we measured the performance of fiducial recognition of our system. We processed 1000 frames and calculated the average in millisecond. We also

**Table 1** Testing Environments.

Computer (Desktop PC)	
CPU	Intel Core 2 Quad CPU 2.5Ghz
MEMORY	4GB
OS	Windows VISTA Home Premium
Video Capture Device	
Capture (a)	Logicool QCam Pro web camera USB 2.0
Capture (b)	Alphawireless EZCAP video capturing device USB 2.0 & Sony DVD Player: DVP-PR50P

**Table 2** The number of unique fiducials in each system.

	System	The number of unique fiducials
Topology	<i>Topolo Surface</i> (our prototype system)	65,536 (in 16bit setting)
	<i>reactIVision</i>	216
	<i>D-Touch</i>	151 (total of 4 different types)
Matrix	<i>ARTag</i>	2,002
	<i>ARToolkit Plus</i>	4,096
	<i>CyberCode</i>	16,777,216 (24bit data capacity)

**Table 3** Average processing time cost in milliseconds.

	binarization	RAG building	fiducial detection	total
<i>Topolo Surface</i>	1.621	1.651	0.051	3.324
<i>reactIVision</i>	1.574	1.659	0.093	3.326

compared our prototype system to *reactIVision*. *QueryPerformanceCounter()* Windows API is used for measurement. This experiment used *Capture (a)* as a video input device, in our lab under a normal room light condition. 12 fiducial markers were put in the input image, each with a different ID.

*ReactIVision* has several features to improve its robustness, such as the pre-processing phase of input image before the binarization and the use of previous frame information, which are not implemented in our prototype system. We excluded the time cost for such features so to compare *reactIVision* and our prototype system in the same condition as possible.

The only significant difference in algorithm between our prototype system and *reactIVision* is just the phase of fiducial detection. The other two phases of binarization and region adjacency graph building are using the same algorithms, but using a different implementation.

**Table 3** describes the result of this experiment. The values are rounded off to three decimal place. Both our prototype system and *reactIVision* is fast enough for tangible interaction. The difference in algorithms seem to show some improvement in the fiducial detection phase, however the ratio in the

**Table 4** False positive detection.

	room	outdoor	film	anime
<i>Topolo Surface</i>	0	0	0	0
<i>reactIVision</i>	1	1	127(30)	23

whole time cost is very small and do not result in any significant benefit in its performance.

### 4.3 False Positive Detection

So to measure how robust against false positive detection our new method is, we used four different video inputs, all of which contain no fiducials.

Two of them are video files of real-world environments taken by *Capture (a)*. One of them is 5 min video taken inside our lab (*room*) and the other is 5 min video taken outdoor in our campus (*outdoor*). The other two are an animation (*the whole 25 min of Ghost in the Shell SAC:Ep. No.4/PAL-DVD*) and a film (*the whole 125 min of Matrix/NTSC-DVD*). Both were tested with *Capture (b)*.

As shown in **Table 4**, our system is significantly robust against false positives, especially in larger bit size setting. 16 bit setting showed significant robustness in all the test inputs. In *reactIVision*, the most of the false positives were observed in the very last part of end credits. These false positives were caused by a trademark that happened to have the same topological structure as one of *reactIVision* fiducial markers. The value inside parentheses is the number of the false positives when this trade mark is excluded.

We also measured how much of the fiducial candidates in our system are rejected in the phase that checks if the angles between *bit-encoding dots* are valid. The numbers of fiducial candidates with valid topological structures observed in our prototype system are 60 for *room*, 153 for *outdoor*, 1227 for *anime* and 5512 for *film*. Such numbers of false positive fiducial candidates with valid topological structures are much larger than those of *reactIVision*.

This observation matches the result of Costanza's experiment, which suggests unbalanced topological region adjacency tree structures and the presence of empty nodes in the upper level can increase false positives, as described in the previous section; the tree structure of our fiducial markers are less balanced and *bit-encoding dots* in our markers are actually such empty nodes.



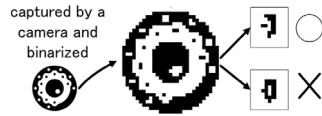


Fig. 8 A possible inter-marker confusion.

However, all of these false positive fiducial candidates with valid topological structures were rejected by the following test by angle information and did not result in any false positive in our prototype system.

Thus, the combination of topology and geometrical information can achieve significant robustness against false positives in our marker design; in *reactIVision*, there is no such validation phase to test those false positives with the valid topological structures; they simply result in the actual false positives as in Table 4.

#### 4.4 Inter-marker Confusion

In our method, inter-marker confusion can occur when any black or white *bit-encoding dots* is falsely recognized to the opposite color. This results in the recognition of a wrong ID to be detected.

**Figure 8** describes one of the situations of such a case. In the upper case on the right, this fiducial is correctly decoded, but if the binarization of the input image causes this black encoding dot to be absorbed by the outer ring, creating a white node inside as in the below case on the right, this causes a change in the topological region adjacency, without altering the number of *bit-encoding dots*. This results in a recognition of the wrong ID.

Sufficient margin space around a *bit-encoding node* can help improve robustness against such inter-marker confusion. Currently, each black *bit-encoding dot* is given the surrounding margin space with the length of 15% of its diameter. In most practical applications, this margin space achieves the sufficient robustness against inter-marker confusion in our prototype system. However, we should also consider input with undesirable noise, to see if such a margin space is enough to avoid inter-marker confusion.

We measured the robustness against such inter-marker confusion, adding Gaussian noise with the different variance levels (average = 0). We used several  $640 \times 480$  bitmap files as an idealized input, each with a different size in pixels for the markers, so to see how the size of the markers in input image

Table 5 True positive detection/inter-marker confusion.

		<i>variance</i>				
		10	20	30	40	50
		<i>Topolo Surface</i>				
size in pixels	30x30	30000 (0)	30000 (0)	29900 (0)	23991 (23)	5959 (314)
	60x60	30000 (0)	30000 (0)	29424 (0)	7709 (60)	0 (0)
	120x120	30000 (0)	30000 (0)	27180 (4)	49 (21)	0 (0)
		<i>reactIVision</i>				
size in pixels	30x30	30000 (0)	30000 (0)	29918 (0)	26067 (2)	10811 (75)
	60x60	30000 (0)	30000 (0)	29524 (0)	10493 (0)	12 (0)
	120x120	30000 (0)	30000 (0)	27572 (0)	123 (0)	0 (1)

can be affected by given Gaussian noise. This is to simulate the robustness in various distances. Each bmp file contains 3 markers with ID = 0, 65535, 43690 (1010101010101010 in binary) for our prototype system and ID = 0, 90, 107 for *reactIVision*. We processed 10,000 frames for this test.

The result of this test is described in **Table 5**. The values outside parentheses are the numbers of fiducial markers correctly detected, and the values inside parentheses are the numbers of inter-marker confusion. For instance, given the markers of  $30 \times 30$  pixels and the variance of 50, *Topolo Surface* (our prototype) detected 6273 markers in 10,000 frames and detected 5959 correct IDs and 314 wrong IDs (inter-marker confusion).

We also compared our prototype system to *reactIVision*. *ReactIVision* can use the information of the previous frame so to gain more robustness. Since our prototype system still has not implemented such a feature, we counted only those fiducial markers detected without any other information than the current frame.

As shown in Table 5, in our prototype system, additional Gaussian noise results mostly in false negative detection and rarely in inter-marker confusion.

With *variance* of 30 and below, there is no significant difference between the two systems. Above *variance* = 40, *reactIVision* showed better true positive rate and less inter-marker confusion. Most false negatives in our method seem to



be caused when an isolated 1 dot pixel caused by noise appeared inside a marker, resulting in the addition of one more region to its topological structure. Filtering out such an accidental noise regions inside a marker may improve true positive detection rate.

However, such filtering may not be required, since such a high level Gaussian noise input would rarely happen in a practical application.

As for inter-marker confusion with such highly noisy input in our method, there were not significant inter-marker confusions below  $variance = 30$ , and those 4 inter-marker confusions in  $variance = 30$  in  $120 \times 120$  pixels were caused by an error of one bit in 16 bit ID. Since digital techniques such as CRC or hamming distance are applicable to our method, it may be desirable to utilize such techniques in cases where such very occasional inter-marker confusions are curial for application. Yet, since such inter-marker confusion is very occasional, it would be sufficient just to use the previous frame information to correct such falsely-detected IDs, as is done in *reactIVision*.

#### 4.5 Testing in a Practical Situation

We also experimented with several real-world situations to simulate practical applications. We put 3 fiducial markers, each with ID = 0, 65535, 43690, the sizes of which are  $5\text{ cm} \times 5\text{ cm}$ . In our current fiducial marker design, the maximum distance that a fiducial can be detected is about 20 times as long as a fiducial size in  $640 \times 480$  resolution. Actual minimal detection size in pixels is about  $30 \times 30$  pixels. This area is about 0.3% at  $640 \times 480$  resolution.

We captured 1,800 frames (15 fps/2 min), moving the camera from about 80 cm above to 20 cm above the fiducial markers on the desk. All the markers were stably detected except one frame that cannot detect ID = 65535, but no inter-marker confusion nor false positive was observed. 5399 out of 5400 fiducial markers in 1,800 frames were correctly detected. The false negative detection rate in this test is  $1/5400 * 100 =$  less than 0.02%.

As for the detection from acute angles, we have tested with various angles, capturing from 30 cm away from the same set of the markers. All the fiducial markers were stably detected without true negatives in the angles between +60 and -60 degrees from the vertical position (or between 30 and 150 degrees, in other words).

However, it is important to consider that, generally speaking, camera-based fiducial tracking can be significantly influenced by real-world parameters such as lighting condition, camera exposure and gain, moving speed of objects, quality of printed markers and the like.

## 5. Conclusion

We developed a 2D fiducial tracking system based on the combination of topological region adjacency and angle information. Our new method can provide a wide unique ID range while maintaining fast processing speed that can tolerate real-time video input. It is also significantly robust against false positives and inter-marker confusion.

Unlike the existing topology-based systems, IDs are directly encoded into markers, and digital techniques as CRC or hamming distance can be applied to the detected fiducial IDs. Such digital techniques can improve the robustness.

Such features of our novel method are beneficial for tangible interaction research and interactive artworks.

**Acknowledgments** The author would like to thank Information-Technology Promotion Agency Japan for supporting part of this software development by Exploratory Software Project Funding and Pola Art Foundation for supporting a one year residency at National University of Singapore by Grants for Overseas Study by Young Artists.

## References

- 1) Ayatsyka, Y. and Rekimoto, J.: Active CyberCode: A directly controllable 2D code, *Proc. CHI'06 Extended Abstracts on Human Factors in Computing Systems* (2006).
- 2) Bencina, R., Kaltenbrunner, M. and Segri, J.: Improved Topological Fiducial Tracking in the reactIVision system, *Proc. CVPR'05* (2005).
- 3) Costanza, E. and Robinson, J.A.: A region adjacency tree approach to the detection and design of fiducials, *Proc. VVG'03* (2003).
- 4) Cotanza, E., et al.: Introducing Audio D-Touch: A Tangible User Interface for Music Composition and Performance, *Proc. DaFX'03* (2003).
- 5) Costanza, E.: Enrico Costanza — Research <http://web.media.mit.edu/~enrico/research/research.php?projectTitle=Older%20Markers> (accessed on 31-Aug-2009).
- 6) Fiala, M.: ARTag, a Fiducial Marker System Using Digital Techniques, *Proc.*

*CVPR'05*, Vol.2, pp.590–596 (2005).

- 7) Fiala, M.: ARTag. <http://www.artag.net/> (accessed on 31-Aug-2009).
- 8) Gallardo, D., et al.: TurTan: A Tangible Programming Language for Creative Exploration, *Proc. TABLETOP'08* (2008).
- 9) Johnston, D.: *Untethered VR*, Ph.D thesis, University of Essex (2001).
- 10) Jordà, S., et al.: The reacTable: exploring the synergy between live music performance and tabletop tangible interfaces, *Proc. 1st Int'l Conf. on Tangible and Embedded Interaction* (2007).
- 11) Kato, H. and Billinghurst, M.: Marker Tracking and HMD Calibration for a video-based Augmented Reality Conferencing System, *Proc. IWAR'99*, pp.85–94 (1999).
- 12) Rekimoto, J. and Ayatuka, Y.: CyberCode: designing augmented reality environments with visual tags, *Proc. DARE 2000* (2000).
- 13) Wagner, D.: ARToolkit Plus. [http://studierstube.icg.tu-graz.ac.at/handheld\\_ar/artoolkitplus.php](http://studierstube.icg.tu-graz.ac.at/handheld_ar/artoolkitplus.php) (accessed on on 31-Aug-2009).
- 14) Wagner, D. and Schmalstieg, D.: ARToolKitPlus for Pose Tracking on Mobile

Devices, *Proc. CVWW'07* (2007).

(Received April 20, 2009)  
(Accepted November 6, 2009)  
(Released February 10, 2010)



**Hiroki Nishino** is currently studying as a Ph.D. student at NUS Graduate School for Integrative Sciences & Engineering, National University of Singapore, with a focus on Computer Music Interaction.