# A SURF-based Natural Feature Tracking System for Origami Recognition

Kening Zhu*
Keio-NUS CUTE Center
NUS Graduate School for Integrative Sciences and Engineering, Singapore

Owen Noel Newton Fernando†
Keio-NUS CUTE Center, Singapore

Adrian David Cheok‡
Keio-NUS CUTE Center, Singapore

Mark Fiala§
Ryerson University, Toronto, Canada

Theam Wei Yang¶
Keio-NUS CUTE Center, Singapore

Hooman Aghaebrahimi Samani‖
Keio-NUS CUTE Center
NUS Graduate School for Integrative Sciences and Engineering, Singapore

## Abstract

In this paper we introduce a system that can recognize different types of origami performed by users in real time. This system allows users to register and use their desired paper for the interaction, and recognize the folding by using Speed Up Robust Feature (SURF) alogrithm with proper parameters. The paper also describes a paper-based game which has been developed as a proof of concept of our method. This method can be considered as an initial step for seamlessly migrating meaningful traditional art of origami into the digital world and as a part of the interactive media.

**Index Terms:** H.5 [Information Interface and Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities; H.5 [Information Interface and Presentation]: User Interfaces—Interaction styles; I.4 [Image Processing and Computer Vision]: Scene Analysis—Tracking; I.5 [Pattern Recognition]: Implementation—Interactive systems

## 1 Introduction

For a long history, origami has been appreciated by people from different cultures all over the world, and used in various areas, such as cultural ceremony, social communication, education and medical therapy. People appreciate origami as not only a kind of delicate artwork but also an indirect method of communication. In Japan, origami was initially used in cultural ceremonies, such as Shinto weddings, as a symbolically respectful practice to show the value of pureness [12]. As a tool for social communication, origami was also widely used in the ancient eastern Asia. During Heian period of Japan, many samurais were keen to exchange their blades adorned by a special kind of paper flower as a sign of friendship [12].

Today, many other applications of origami have been developed in different areas. Research in education shows the origami training affects brain development [15]. By playing origami, children can improve their creativity, spatial reasoning skills and performing ability. In addition, origami can also be used to enhance the communication among teachers and children [2]. Additionally, origami is playing an important role in today's family communication. In Japan, children can gain knowledge about the past from traditional crafts, such as origami, made by their grandparents [4].

On the other hand, people, especially young generation, find it more familiar with modern entertainment and communication methods such as virtual reality, online social network and so on. They preferred to manipulate the digital contents with mouse and keyboard. However, research showed that manipulating virtual reality with tangible and meaningful objects could provide more advantages in user experience [7]. With its various forms in manipulation and shapes and its rich historical and cultural legacy, origami could be a good candidate for tangible interface in mixed reality interaction. Taking these discussions into consideration, we developed a method to recognize users' origami in real time. By comparing with the existing related works, we believe our method could provide more freedom for users to perform origami during the interaction with virtual reality.

In the following of the paper, we will describe more details about this research. Section 2 will introduces the basic types of origami that can be recognized by our current method, and the related works in origami recognition and interaction will also be discussed and compared with our research. The detailed description of the method will be presented in Section 3. In section 4, application of our method will be discussed. Finally we conclude with the future works in this research.

## 2 Background & Related Works

Many origami books begin with a description of basic origami techniques which are used to construct the models. There are five standard basic origami techniques [16], as shown in Figure 1, including book fold, kite fold, cupboard fold, shawl fold and cushion fold.

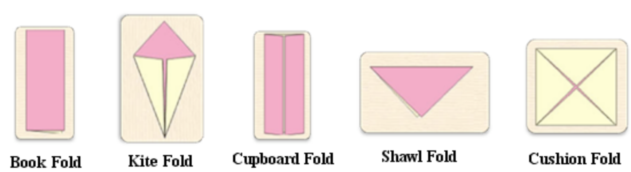



Figure 1: Five Basic Origami Techniques

Figure 2 shows that many complex form of origami can be derived from these basic foldings. Hence, these bases can be viewed as an introduction to origami and their simplicity is suitable for beginners. Our system is able to recognize all the fundamental base folding.

There are quite a few related researches which have focused on tracking or recognizing origami, using either software or hardware. Jun Mitani et al. [11] developed a system that uses a digital camera to read QR-codes preprinted on the folding paper and then employs this information to infer the three-dimensional form as the user folds the paper. Shimanuki et al. [14] proposed a vision-based algorithm to recoginze the folding process of an origami procedure from origami drill books. Yasuhiro Kinoshita [9] developed a silhouette-based method estimating the state of origami in a cam-

*e-mail: ken@cutecenter.org
†e-mail:newtonfernando@cutecenter.org
‡e-mail:adriancheok@cutecenter.org
§e-mail:mark.fiala@ryerson.ca
¶e-mail:u0606604@nus.edu.sg
‖e-mail:hooman@nus.edu.sg

20th International Conference on Artificial Reality and Telexistence (ICAT2010), 1-3 December 2010, Adelaide, Australia
ISBN: 978-4-904490-03-7 C3450 ©2010 VRSJ

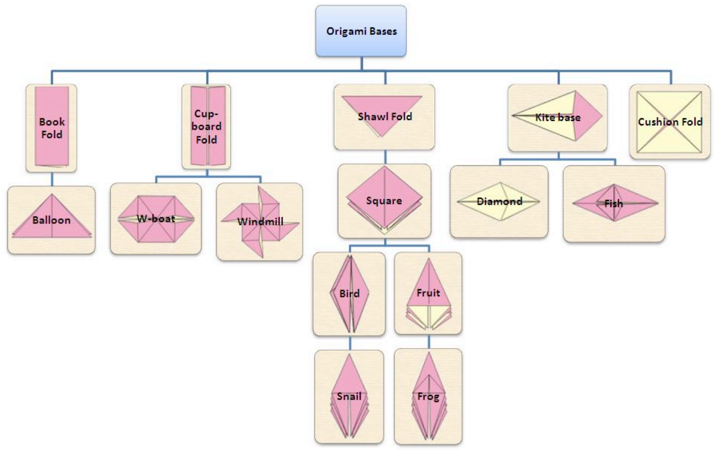



Figure 2: Tree Hiarachy of Origami Creation

era image. David Gallant et al. [3] presented Foldable User Interface (FUI) which are sheets of paper that are augmented with IR reflectors tracked by camera. Yingdan Huang and her colleagues developed Easigami [6], a novel tangible user interface which embeds sensors on the edge of paper, so that user can construct different shapes of model by combining papers and the model will be reflected in 3D virtual representation. MIT Media Lab developed Origami Desk [8], using projected animations to directly map instructions onto the users' paper, electric field sensing to detect touch inputs on the desk surface, and swept-frequency sensors to detect the papers folds. Compared to these related works, our method with a simple webcam enables users to use their desired origami paper without requiring any other special hardware equipment. This means users can even just cut out a square piece from old newspaper as the tool for interacting with our system, hence it provides greater ease of use and more freedom.

## 3 System Description

The whole origami interaction system consists of two main modules: paper registration and folding recognition, which are built on top of Speed Up Robust Feature (SURF) algorithm [5]. As shown in Figure 3, user places a square piece of his/her desired paper under the camera which is set in the top of the table. The paper registration module analyses the paper and extract the natural feature points on the paper based on SURF descriptors. In this step, the system reports to the user whether there are sufficient features to utilize for tracking. Therefore, user should supply sufficiently textured paper to the system in this step. The feature points will be used in the step of origami recognition, where SURF is used to track the visibility of different regions of the origami paper, further to recognized different types of folding performed by the users. The rest of this section will explain why we choose SURF as our fundamental algorithm, and then discuss the features of our method in details.

### 3.1 Experiment for Algorithm Selection

Based on a series of image training and experiments done on different types of detectors and descriptors, SURF with Fast Hessian detector is selected as the natural feature tracking algorithm for our origami recognition system. As shown in Figure 4, we performed the matching of four different types of origami paper with simple folding using SURF algorithm. The results shows that SURF could

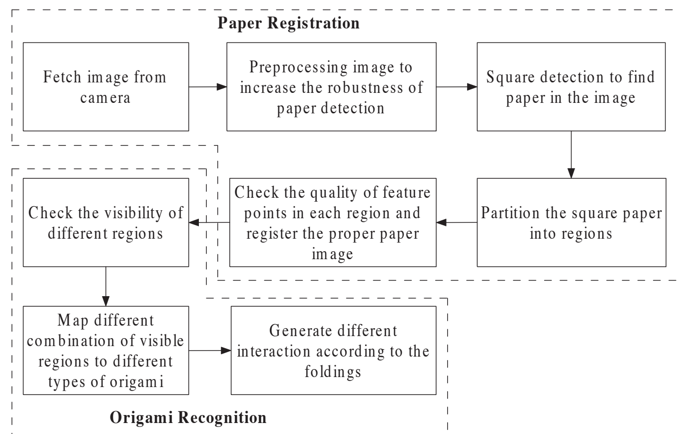



Figure 3: Overview of the system flow

provide good feature tracking even with the presence of human hands. Another main reason is that SURF descriptor requires less computational time compared to others, resulting in faster tracking of object. Although SIFT [10] is more robust in term of correct matching, affine distortion, and illumination changes, our origami recognition system uses only static camera and nearly constant lightning condition which gives less distortion and less illumination changes. The computation speed is considered more important as origami recognition system is required to run in real time. On the other hand, SURF descriptor is better for 3D object recognition while Ferns [13] is only good for estimating planr homographies, since we are planning to recognize origami in a 3D form. In addition, compare to Ferns which requires to process the image in a complex classifier at the beginning, we will use a relatively simpler approach to pre-process the registered paper image.

As the SURF algorithm consists of three main parameters: Fast Hessian Threshold, number of octaves and number of layers, it is necessary to obtain the optimum values that fulfill our requirement of efficiency and correctness in the origami recognition system. Figure 5 shows the testing image used for experiment.

As shown in Figure 6, with the increase of feature points and the

(a) Postcard (b) Newspaper

Figure 4: Test of Different Paper Tracking using SURF

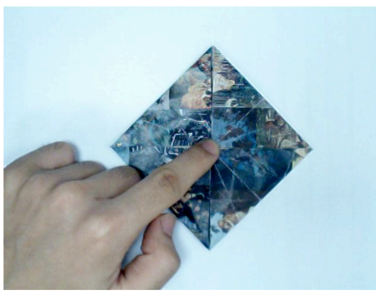


Figure 5: Tesing Image for SURF Parameter Selection

robustness, the computation time is increased in the SURF processing. Since the origami recognition system only performs detection every 1000 millisecond, the highest number of interest points with acceptable long computational time is selected.

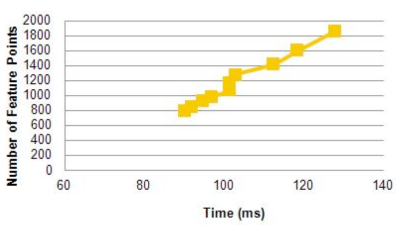



Figure 6: The relation of number of feature points and computational time in SURF

After experimenting in different combination of the three main parameters in SURF as shown in Figure 7, 8 and 9, we decide the values as Fast-Hessian threshold = 100, Number of octaves = 5 and Number of layers = 4. In addition, the 128-element descriptor is chosen.

## 3.2 Detailed Description of Our Method

### 3.2.1 Square Paper Detection

Since the most common origami paper is approximately in a square shape, our system performs a process of square detection, as shown in Figure 10, when the user places his/her origami paper under the camera in the beginning.

Firstly, to reduce the noise and peak intensity in pixels, image smoothing is executed by using pyramid downsize and upsize of the image captured by the camera. Here 5x5 Gaussian kernel, as shown in the following equation , is used for the smoothing.

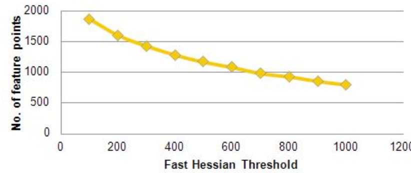



Figure 7: The relation of number of feature points and Fast-Hessian threshold in SURF

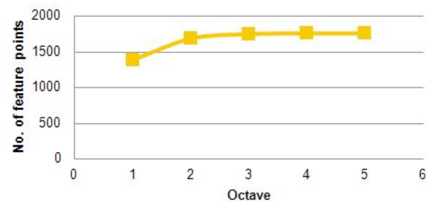



Figure 8: The relation of number of feature points and number of octaves in SURF

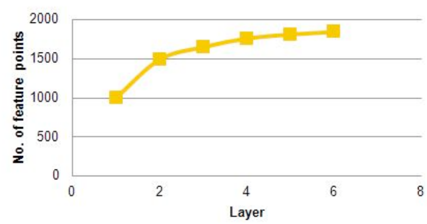



Figure 9: The relation of number of feature points and number of layers in SURF

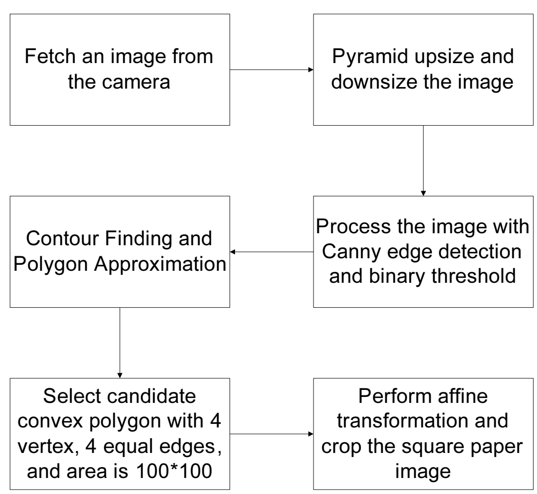



Figure 10: Flowchart of Square Paper Detection

$$5x5-Gaussian-kernel = 1/16 * \begin{pmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 1 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 1 \\ 1 & 4 & 6 & 4 & 1 \end{pmatrix}$$

After smoothing, Canny edge detection [1] and image dilation are performed with a 3x3 rectangular structuring element, and potential holes between edge segments from Canny detection are removed. Then different levels of binary threshold setting are done in each of the three color channels, to search possible polygons in different channels, as shown in Figure 11.

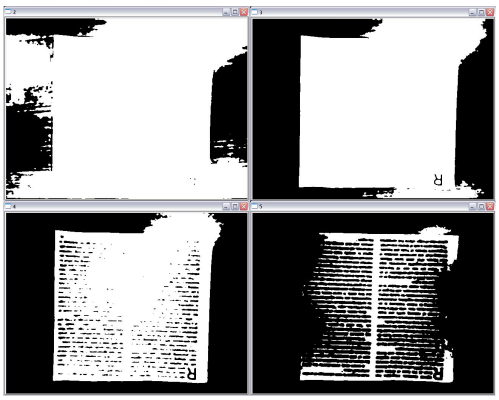


Figure 11: Binary Result of Paper Image

Finally, a series of filters are set to obtain the proper square. The possible candidates have to first pass through a corner angle test that requires the angle to be approximately 90 degree and an edge length test that requires 4 edges are approximately equal in length. The final filter for the square detection is to remove the adjacent squares which are only five pixels deviated from others. The main purpose of this final filter is to remove squares overlapping in the similar location which result from image smoothing and dilation. After the filtering, rotation and cropping of the image are executed, due to the requirement of exact coordinates of feature points during the origami recognition. The final result of origami paper detection is shown in Figure 12, and this process will be performed in twice for both front side and back side of the paper.

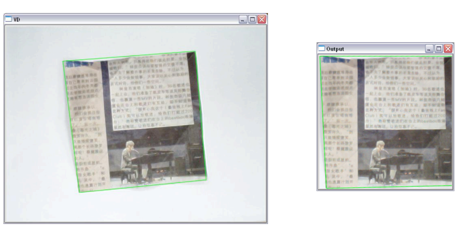


Figure 12: Result of Square Paper Detection

### 3.2.2 SURF-based Paper Analysis

The images of the square paper are passed to the module of SURF-based paper analysis, as shown in Figure 13, which determines whether the paper is textured enough for the camera tracking and extracts necessary feature points for the origami recognition.

Based on the five basic types of origami, the system first divides both the front square and the back square into regions as shown in Figure 14. These regions are defined mathematically using the combination of the equations of straight lines. Then the regions will be checked one by one for the sufficient number of feature points.

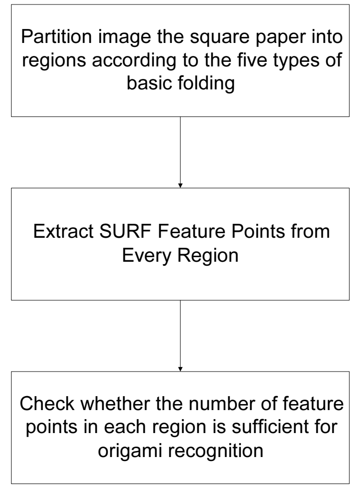
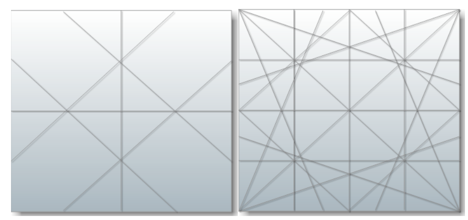



Figure 13: Flow Chart of SURF based Paper Analysis



(a) front partition (b) back partition

Figure 14: Partition of Paper

The division straight lines, in Figure 14, will be defined in equations as follow:

Given that w = the width of image and h = the height of image, and define $\alpha = 0.4142$ and $\beta = 0.5858$, and $x_i$ and $y_i$ are the coordinates horizontally and vertically with the center of the paper as the original point.

Front side line equation for 6 straight lines:

$$x_i = \begin{cases} \frac{w}{2} & y = 0 \\ \frac{(y_i - hb_i)w}{m_i h} & otherwise \end{cases} \quad i = 1...6 \tag{1}$$

$$m_i = \begin{bmatrix} 0 & -1 & 1 & -1 & 1 \end{bmatrix}$$
$$b_i = \begin{bmatrix} 0 & 1+\alpha & \alpha & \beta & -\alpha \end{bmatrix}$$

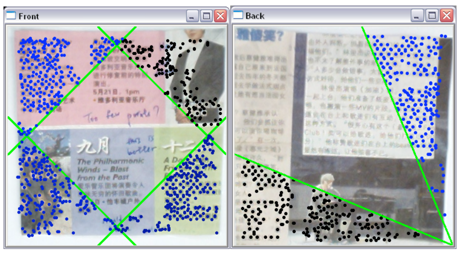
Back side line equation for 20 straight lines:

$$\begin{cases} x_j = \frac{w}{4}, \frac{2w}{4}, \frac{3w}{4} & y = 0 \\ y_j = \frac{h}{4}, \frac{2h}{4}, \frac{3h}{4} & x = 0 \\ y_j = m_j \frac{h}{w} x_j + b_i h & x \neq 0, y \neq 0 \end{cases} \quad j = 1...20 \tag{2}$$

$$m_j = \begin{bmatrix} -1 & 1 & -1 & 1 & 1 & -1 & -\alpha \\ (-\frac{1}{\alpha}) & \alpha & (\frac{1}{\alpha}) & (-\frac{1}{\alpha}) & -\alpha & (\frac{1}{\alpha}) & \alpha \end{bmatrix}$$
$$b_j = \begin{bmatrix} \frac{3}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & 0 & 1 & \alpha \\ \frac{1}{\alpha} & 0 & 0 & 1 & 1 & -(1+\alpha) & \beta \end{bmatrix}$$

The procedure of the paper analysis is based on SURF algorithm. With the proper parameters, our system will check every region in the square paper based on a set of criteria, such as the total number of feature points and the density and the distribution of the feature points in a particular region. Figure 15 shows the results of feature points in the front and back of the origami.

After this analysis, our system will report the user whether his/her current origami paper is textured enough for the origami tracking interaction. The system will register the final well-textured paper for the origami recognition module.




(a) front image (b) back image

Figure 15: Feature Points Checking in Different Regions of the Paper

### 3.2.3 Origami Recognition

In the module of origami recognition, as shown in Figure 16, based on SURF matching method, the paper image in each frame is compared with the square images of the square images stored by the previous steps.

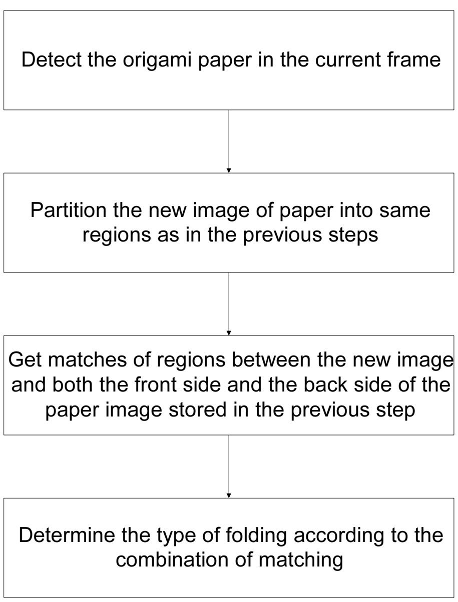



Figure 16: Flow Chart of Origami Recognition

The folding procedure does not have the high speed, as a person is limited in how fast they can fold, therefore we can implement more accurate tracking. We decided 1000 milliseconds is an acceptable update rate for origami. This allows us to use more interest points and a full 128-element description is chosen for the matching between the paper image captured real time and the pre-stored template images of the paper.

With the proper parameters, the tracking algorithm recognizes different types of origami folding according to the visibility of the feature points in each region in the paper. The regions represented in the front side and back side of the origami paper are different due to the fact that most folding types cover the entire front side of the paper. Before determining the appearance of a particular region, the threshold for each region has to be calculated. The threshold refers to the number of feature points so that when detected points are higher than this threshold, the particular region is considered shown and not blocked. As the total number of feature points for different origami paper is not the same, a dynamic threshold has to be implemented in order to determine the appearance of a particular region. Different types of folding are mapped to different combinations of the appearance of regions. For example, the disappearance of the top right corner of front side and the appearance of the top left corner of the back side can be recognized as the folding shown in Figure 17. Finally, our system is able to recognize 24 types of simple folding as shown in Figure 18.

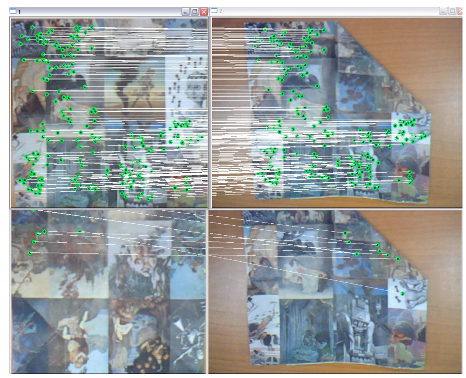


Figure 17: Result of Square Paper Detection

## 4 Application

This Origami recognition system provides a platform for a wide range of application in education, entertainment and social communication, including storytelling, gaming, designing etc. As a proof of concept, a flash game application is developed to realize our origami recognition method. As shown in Figure 19, the main objective of the game is to construct towers to defeat the virtual monsters. The towers in the game are represented uniquely by different kind of fundamental origami folding. Hence, in order to win the game, players are required to successfully construct one or more Origami basic folding to build defence towers.

In addition, this game application also presented a way to replace or substitute keyboard, mouse, or joystick as the input to the computer. This game only requires one paper from the users, and then he/she is able to navigate the virtual world using the origami paper. For instance, by constructing the arrow shape pointing upwards, which is shown in Figure 19, the navigation box will move to the up direction accordingly.

During the interaction, a real-time video is streamed and shown in the flash game. The origami recognition algorithm is running at the background and is providing real-time update to the gaming interaction. 30 frames per second is achieved by using video frame splitter. As each folding requires some time to complete, it is unnecessary to waste resources to do folding recognition every frame. Hence, the video frame is split and one of them is used for updating the screen for real time streaming and another is used to detecting folding for every 1 second which is approximately 30 frames apart. Although there may result in some delay for the folding recognition, it is hard for users to notice in this small interval of time according to our primitive observation.

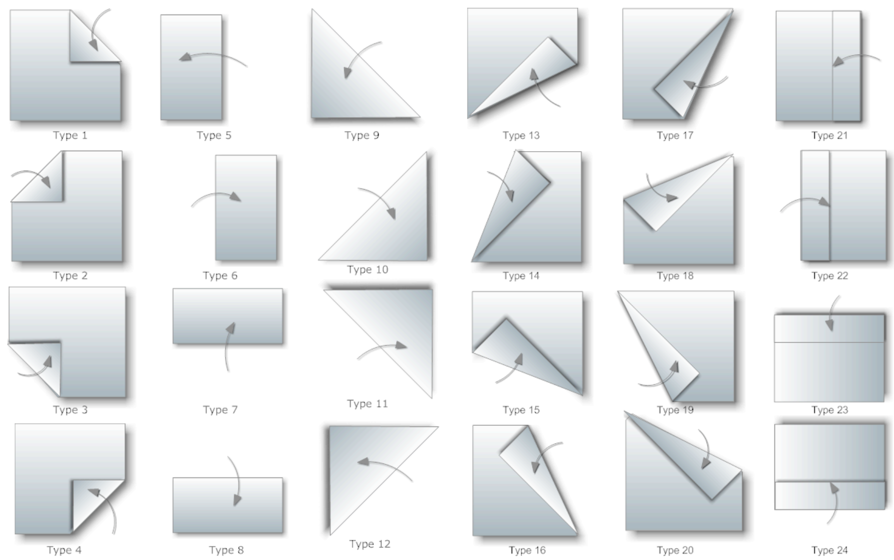



Figure 18: Possible Types of Origami to Recognize


Figure 19: Origami-based Interactive Gaming

## 5 Conclusion & Future Works

This paper describes a paper-based interaction system that recognizes different types of origami based on natural feature tracking. Our algorithm is built on top of SURF algorithm with proper parameters which meet the requirement of a sufficiently real-time recognition for origami. An simple game application has been developed to the robustness and responsiveness of our algorithm. The main purpose of this research is migrating meaningful traditional art of origami into the digital world and as part of the interactive media.

In addition to the current features, our system can be easily expanded to more complex types of origami since different folding are corresponding to different combination of the visiblity of different regions in the paper. Our future work will include machine learning algorithms to increase the robustness of the system. Meanwhile, a design platform and framework are planned to be developed for end-user to design their own origami interaction based on the presented method. We also plan to conduct user study to evaluate our approach in virtual reality interaction.

## References


[1] J. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 8(6):679–698, 1986.

[2] F.-T. Deborah. Origami and communication strategies. *Doshisha studies in language and culture*, 1(2):315–334, 12 1998.

[3] D. T. Gallant, A. G. Seniuk, and R. Vertegaal. Towards more paper-like input: flexible input devices for foldable interaction styles. In *UIST '08: Proceedings of the 21st annual ACM symposium on User interface software and technology*, pages 283–286, New York, NY, USA, 2008. ACM.

[4] R. Goodman. *Family and social policy in Japan*. Cambridge University Press, 2002.

[5] T. T. Herbert Bay, Andreas Ess and L. V. Gool. Surf: Speeded up robust features. *Computer Vision and Image Understanding (CVIU)*, 110(3):346–359, 2008.

[6] Y. Huang, M. D. Gross, E. Y.-L. Do, and M. Eisenberg. Easigami: a reconfigurable folded-sheet tui. In *TEI '09: Proceedings of the 3rd International Conference on Tangible and Embedded Interaction*, pages 107–112, New York, NY, USA, 2009. ACM.

[7] H. Ishii. Tangible bits: beyond pixels. In *TEI '08: Proceedings of the 2nd international conference on Tangible and embedded interaction*, pages xv–xxv, New York, NY, USA, 2008. ACM.

[8] W. Ju, L. Bonanni, R. Fletcher, R. Hurwitz, T. Judd, R. Post, M. Reynolds, and J. Yoon. Origami desk: integrating technological innovation and human-centric design. In *DIS '02: Proceedings of the 4th conference on Designing interactive systems*, pages 399–405, New York, NY, USA, 2002. ACM.

[9] Y. Kinoshita and T. Watanabe. Estimation of folding operation using silhouette of origami. *IAENG International Journal of Computer Science*, 37(2), 2008.

[10] D. Lowe. Object recognition from local scale-invariant features. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 2, pages 1150 –1157 vol.2, 1999.

[11] J. Mitani. Recognition and modeling of paper folding configuration using 2d bar code [in japanese]. *Transactions of Information Processing Society of Japan*, 48(8):2859–2867, 8 2007.

[12] K. Ohashi. The roots of origami and its cultural background. In *Proceedings of the second international meeting of origami science and scientific origami*, pages 503–511, 1997.

[13] M. Ozuysal, M. Calonder, V. Lepetit, and P. Fua. Fast keypoint recognition using random ferns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(3):448 –461, 2010.

[14] H. Shimanuki, J. Kato, and T. Watanabe. Recognition of folding process from origami drill books. In *ICDAR '03: Proceedings of the Seventh International Conference on Document Analysis and Recognition*, page 550, Washington, DC, USA, 2003. IEEE Computer Society.

[15] K. Shumakov and Y. Shumakov. *Functional interhemispheric asymmetry of the brain in dynamics of bimanual activity in children 7-11 year old during origami training*. PhD thesis, Rostov State University, 2000.

[16] Yuri and K. Shumakov. *Origami Chic*. Oriland Publications, 2001.