Origami Guru: An Augmented Reality Application to Assist Paper Folding

Nuwee Wiwatwattana, Chayangkul Laphom, Sarocha Aggaitchaya and Sudarat Chattanon

Abstract Origami folders often find themselves unable to recall the entire sequences of foldings, and origami drill books have been the only resources for them to learn to fold. However, many people have difficulties understanding symbols and operations in drill books. The purpose of this research is to design and develop a mobile augmented reality application that acts as a technological aid to origami folders. Extensive user studies have been employed at every step of the implementation. The application recognizes the current state/shape of the paper and instantly overlays the diagrams and instructions of the next step(s) on the real paper through the smartphone's camera view. Evaluation results show that the application can successfully assist folders most of the time, and the qualitative mean score is about the same level of most mobile applications. This work will benefit a lot of origami practitioners such as those in geometry classes and intellectual development classes.

Keywords Augmented reality · Memory aid · Interaction design · Paper folding · Origami

1 Introduction and Motivation

The traditional art of folding flat papers into things is usually associated with Japanese culture and is generally known as Origami. Because of its entertaining and educational benefits, paper folding is often taught in early school years or young at home. In higher education, algorithms and data structures of folding and crease patterns are studied in mathematics and computer science. The principles of origami have also been used to solve real-world problems and inspired cutting-edge developments in science and engineering, such as origami stent grafts [1], nanoscale DNA origami [2],

N. Wiwatwattana(✉) · C. Laphom · S. Aggaitchaya · S. Chattanon
Computer Science Program, Department of Mathematics, Faculty of Science,
Srinakharinwirot University, Bangkok 10110, Thailand
e-mail: {nuwee,sc541010324,sc541011150,sc541010375}@swu.ac.th

S. Latifi (ed.), *Information Technology New Generations*,
Advances in Intelligent Systems and Computing 448,
DOI: 10.1007/978-3-319-32467-8_95

Table 1 Summary of the preliminary user study

Observation/Question	User1	User2	User3	User4	User5
Have you ever folded a crane?	Yes	No	Yes	Yes	Yes
Did he/she consult the book?	Always	Always	No	Sometimes	No
Did he/she fold through trial and error?	No	No	Yes	Yes	No
Have you thought about the next fold before consulting the book?	No	No	N/A	Yes	N/A
Did he/she finish the crane?	Yes	Yes	Yes	Yes	Yes

folded automotive airbags, or folding solar panels for space satellites [3], to name a few.

There are an innumerable number of origami models foldable by hands ranging from basic and popular models (e.g., crane) to advanced and unimaginable models. Instruction books described in Yozhizawa-Randlett system [4] (also called origami drill books) have been the only resources to learn folding sequences. Until recently, the internet has become so widespread and origami instructions have been published on the web. At present, we are in the world where smartphones are becoming part of our organs. As a result, many drill books are built in the form of mobile applications. In the Apple's App Store and the Google's Play Store, there are over three hundred of such applications available for download (at the time of the preparation of this manuscript), which indicates significant interest in paper folding. Most of the applications are free; some of them are gamified; some of them are video-based; some are in 3D, while others are traditional 2D drawings.

It is not possible for everyone to be able to recall the entire sequences of particular origami models as folding papers are not everyday activities. To study how people fold a simple model and the extent to which the drill book aids the folding, five undergraduate students, two men and three women, were asked to fold a typical "crane" in our preliminary study. An origami drill book diagramming the crane model was provided from the beginning. Table 1 summarizes questions that were asked and observations that were made. Results from the preliminary study show that the drill book plays an important role in finishing the model regardless of folders' familiarity with the model. However, there are arguments we would like to note.

- It is not always plausible for folders to follow the instructions diagrammed in the books or applications. Although Yozhizawa-Randlett system provides a standard set of folds, operations, and symbols, not everyone has studied the system before.
- Some folders can not relate 2D portions in the books with the real objects they are folding, especially if the models are complex, as argued in [5].
- Instruction books and their mobile application counterparts are too similar to cheat sheets, and that would defeat the pleasure of paper folding.
- A paper drill book does not come in handy when it is needed.

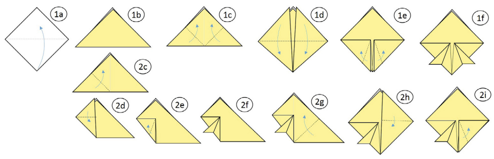


Fig. 1 Two possible paths of folding a samurai helmet: (1) 1a-1b-1c-1d-1e-1h-1i-1f (2) 1a-1b-2c-2d-2e-2f-2g-2h-2i-1f. The rest of the steps are omitted.

Fig. 2 Cue cards used in the second user study

Nobody has ever denied the fact that modern technology has greatly and positively assisted people in learning. Lately, a growing number of handheld augmented reality (HAR) applications are developed to improve the way a human learns. In 2015 NMC Horizon Reports, the renowned five-year horizon reports from an international not-for-profit consortium of learning-focused organizations, the expert panel listed augmented reality as one of the visualization technologies to help process information, identify patterns, and sense order in complex situations [6]. As a matter of fact, we thought that augmented reality technology could aid folders to learn and construct origami in a way.

In this research, we aim to develop an intelligent application utilizing the augmented reality to help people fold origami models. This paper is organized as follows. Starting from the ground up in Section 2, we establish requirements and develop designs from the mind process of the folders and folding sequences modeling. Section 3 addresses technical details of the implementation of the research prototype, and Section 4 evaluates the prototype. Subsequently, works related to our research are discussed in Section 5. Finally, Section 6 concludes our work and discusses future work.

2 Establishing Requirements and Developing Designs

We define our project as a mobile application that can aid origami folders, both novices and experts, to finish the models with the help of augmented reality technology. Our preliminary study has already addressed the needs for such application. Establishing requirements includes understanding the nature of the folding, i.e., the analytic process of the folder. In order to finish a paper model, the folder needs to know the sequence of the folding.

Let us first examine the sequences of folding several models. The technique of going from one (initial/immediate) shape to another (intermediate/final) shape might be a standard one—like a mountain fold, a valley fold, or a reverse fold—or complex one—like a squash fold or a rabbit ear fold. A folding operation consists of both the folding technique and the folded location. As seen in drill books, the folding sequences are often captured in one fix path of sequences. However, many models can be accomplished in different paths. One example of these is folding a samurai helmet, shown partly in Fig. 1. There are two possible paths to fold the samurai helmet. The first path is to fold both the left and the right flap at the same time. The second path is to get an antler-like kuwagata, one flap at a time. Several models also share a number of standard bases as the first few steps in construction. All in all, a set of origami models folding sequences can be described by a state transition diagram, originally stated in [5].

2.1 State Transition Diagram (STD)

We formally define a state transition diagram of origami models' folding sequences as a quintuple:

$$STD(Q, \sum, \delta, q_0, F) \tag{1}$$

where:

- Q is a finite set of states, each state is a unique shape of (finished/unfinished) origami.
- $\sum$ is a finite set of folding operations.
- δ is the folding function, represented as arrows between states labeled by the folding technique and the folded location, $\delta \subseteq Q \times \sum \times Q$.
- q_0 is the start state, a.k.a initial shape, usually a flat rectangular paper, $q_0 \in Q$.
- F is a finite set of end states, a.k.a. completed origami models, $F \in Q$.

For folders who have never folded the model before, they generally imitate the instructions in drill books step by step. For folders who are accustomed to fold the model, they will use their memories to recall the folding operations—a conscious process of problem solving. We already knew from the preliminary study that when folding some people tried to fold through trial and error and some people thought about the next fold before consulting the book. We decided to continue studying the nature of folders' mind process by direct observation. We would like to know the stepwise thinking of folders and apply that thinking into the design. This leads to another round of user study.

2.2 The Second User Study

In the second user study, ten undergraduate students were recruited to fold the "crane" again. We prepared instructions in the form of stepwise cue cards as shown in Fig. 2. Unlike the preliminary study, we handed the next-step card to folders only when it is requested. The hypothesis is that, at any step, the folder would like to know one step

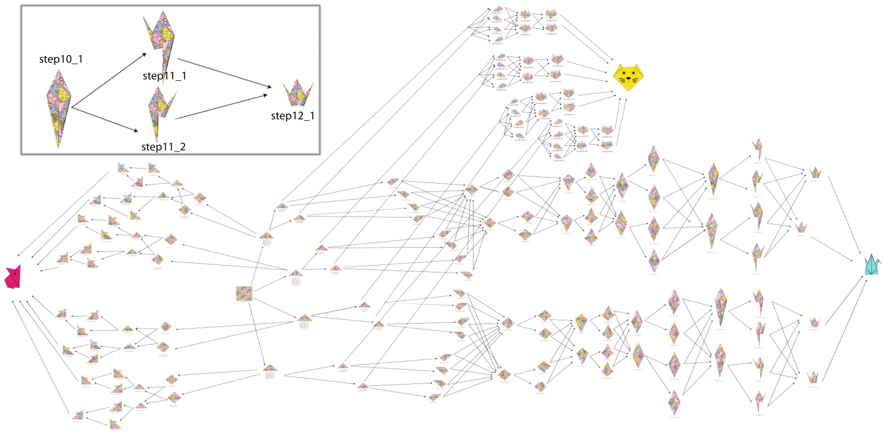


Fig. 3 The actual state transition diagram of three models: crane, fox , and cat face. The insert in the top left shows part of the three states of the crane. One sub-state can be folded into multiple sub-states, and multiple sub-states can be folded into one sub-state.

ahead. It turned out that nine students asked for the cards at least once (One person was very professional!). Five of them, who declared themselves as having folded the crane many times, asked to see the next step and the step after the next before continuing to fold. The rest of them asked for the next three steps. In addition, some of the students, especially those who rarely fold the crane, had difficulties in understanding the diagrams and had to ask for help sometimes.

The results from the second study confirmed the findings that parts of human memories are short-lived and that humans tend to use heuristics. They tend to carry out the folding by finding a short term solution, step by step. For this reason, our application is in a better form to support just enough information for technological intervention. This could be done with augmented reality to superimpose the additional information upon real-world objects in situ. To assist folders relate the real-world 3D papers to the diagrams, we can project the next-step diagram over the paper using the mobile screen and camera. Thinking about the situation when a folder is folding an origami, but finds that he cannot recall the next step. He can take out his mobile phone and points the camera to the currently unfinished paper. Instantly, the application suggests him what he should do next with the diagram and instructions floating right on top of the paper. Folders can readily map the current face, the current crease, or the current edge of the paper to the diagram and easily understand the instant direction of the next folding operation(s). Not only the next step, but we also added two more next operations to the application. Another consideration about perceptive cognition suggested that we should not use the same color/pattern of the diagram to that of the paper, otherwise folders may not be able to distinguish between reality and virtuality.

3 Implementation

We developed a prototype application on the Android platform using Unity [7] and Vuforia's development platform [8]. Incorporated the requirements and designs discussed earlier, the prototype has two main sections: the real-time model recognition and the static diagrams. The latter is an electronic version of drill books for beginners or for studying the entire folding sequences. The former is the augmented reality module, which we shall give more details in the following sub-sections.

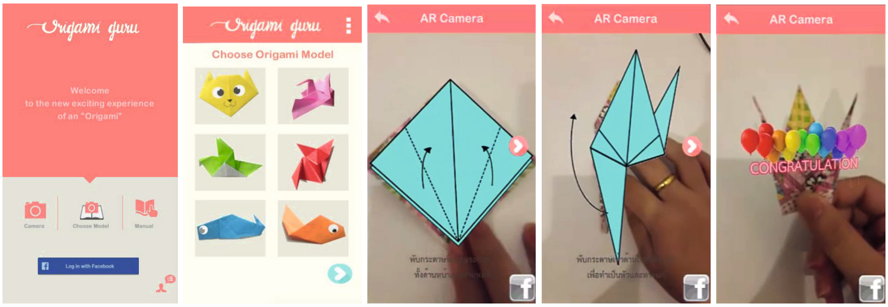
3.1 Architecture

The system consists of three layers of software performing all tasks in an Android application unit, which are made possible by Unity and Vuforia. The bottom layer is a driver to control and send real-time feed from the camera. The middle layer is an augmented reality (AR) engine. The AR engine performs a live image analysis of the camera feed, recognizes the paper model, estimates the 3D position of the paper (pose estimation), and renders the virtual diagrams and instructions. The top layer is an interactive interface and live screen that users see. Users can interact with the application such as rotating the camera and zooming the camera.

In Vuforia, each real object that the developer would like to recognize is called a target (a marker in a broad sense). The image targets are kept in the Vuforia's device database. To be a valid target, and to be able to perform a high quality pattern recognition, an ample amount of feature points has to be extracted. We found that plain color papers do not produce a high enough number of feature points. Thus, we needed to use a sophisticated pattern of papers with a lot of feature points. The pattern is shown in the middle of Fig. 3. A target is associated with one or more virtual diagram to be superimposed on. All targets and virtual diagrams are to be registered with the AR engine.

3.2 Data Structures

A state in the state transition diagram is, however, not a one-to-one mapping to an AR target. This is because of two main reasons. (1) An origami model is a 3D statue that has an unlimited number of perspectives. There is no automatic way, as of yet, to directly recognize an arbitrary 3D object in Vuforia. (2) An origami paper, says a rectangle with 4 sides: a-b-c-d, can be folded in different directions. For example, a valley folding from side a and c is the same as a valley folding from side b and d. If the paper pattern is not symmetric, like ours, there will be multiple image patterns for a given shape. Therefore, we had to create multiple targets for each state of the origami model. We call each of them a sub-state. A sub-state is folded to one or more



(a) Home page (b) Choose page (c) Step 6 (d) Step 11 (e) Last step

Fig. 4 Screenshots of the mobile application before and during the folding of a crane

sub-states of the next state. Fig. 3 depicts an actual state transition diagram that we built. Out of 8 models available in the application, only 3 models—the crane, the fox, and the cat face—are shown here due to space limitation. The insert on the top left is a close-up of the crane model. A sub-state of state no. 10 (step10_1) can be reverse-folded on the left leg to the first sub-state of state no.11 (step11_1), or can be reverse-folded on the right leg to the second sub-state of state no.11 (step11_2).

Each image target has a corresponding identifier. Each sub-state in the state transition diagram is a JavaScript Object Notation (JSON) object. Each JSON object consists of the model identifier, the image target identifier, the state identifier, the next-state virtual diagram object identifier, the second-next-state virtual diagram object identifier, and the third-next-state virtual diagram object identifier. Three next states are maintained because, according to the user study, this is the number of steps folders generally requested at a time. At present, although a sub-state can be folded into several next sub-states, we only encoded one path.

3.3 User Interface and Functions

Our application is called "Origami Guru". The user interface is designed using the material design principle. On the homepage, users can login with their Facebook accounts. Users can choose between three modes of functions, shown in Fig. 4a. The first mode opens the camera view and let users point the camera to the paper model they are folding. If there are multiple models that belong to the shape that is being recognized, as are the cases for the first few folds, the application asks users to choose the model they are folding. The virtual diagram (with instructions) of the next step is superimposed on the paper model with an arrow button to superimpose the step after the next, and so on.

The second mode lets users choose the model they are folding first before opening up the camera view. Choosing the model page is shown in Fig. 4b. A partial screen flow of recognizing the crane model is shown in Fig. 4c, Fig. 4d, and Fig. 4e. Users can also share their model with the Facebook community via the share button. The third mode is the static diagram page.

4 Evaluation

To measure the quality of our application, we conducted two evaluation studies and derived two scores: the average accuracy of recognizing the folded paper, and the overall usability score. The LG G2 smartphone with Android 4.4 (KitKat) was used in these studies. The first study measured the average accuracy of recognizing 6 models across all foldable shapes and was conducted by our team of researchers. For each model, we performed ten rounds of foldings. For each shape, we calculated the consistency percentage, which is the percentage the application was able to recognize as the correct model(s) within 5 seconds using the second mode. Fig. 5 shows the comparison between these 6 models. Detecting the pigeon's folded papers yields the highest accuracy of 88 percent, followed by the cat face, the crane, the fox, the fox face and the fish respectively. This is because the pigeon's folded shapes have very rich natural features such as corners and crease patterns. The less, the harder, such as the fox face or the fish. On average, out of 10 times, there are about 7 times that the application were able to detect models correctly. However, these numbers can vary depending on the phone camera, the environment (lighting and contrast), and the augmented reality platform.

We recruited 20 participants, 9 men and 11 women of age between 15 and 25, to evaluate the application. We asked each participant to select one origami model that he/she has folded before, gave him/her the paper, and explained the basis of the application. Due to its evidence as a reliable method for web and mobile usability assessment, the Software Usability Scale (SUS) [9] was employed in our study. The standard version of SUS was translated into Thai and distributed to the participants after they finished the test. As shown in Fig. 6, the mean usability score is 69.72 (SD = 13.14)

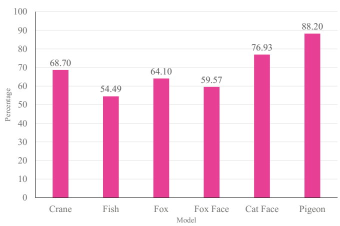


Fig. 5 The average accuracy of recognizing each origami model across all foldable shapes

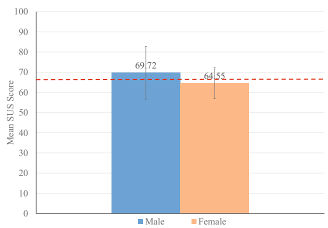


Fig. 6 The mean score for male and female participants according to the Software Usability Scale (SUS)

for men and 64.55 (SD = 7.65) for women. However, this does not show evidence of a statistically significant gender difference (p-value = 0.32, by the two-sample t-test with unequal variance). Results involving all participants are acceptable with an overall mean score of 66.88 (SD = 10.51) in red dashed line, about the level of the mean for web and mobile interface type, a "C" grade, according to [10].

Comments from participants suggested that the application is innovative and surprised them about how augmented reality can be combined with folding origami. One participant even encouraged us to add more models to the application. Another participant suggested superimposing animated instructions in place of the current static picture and suggested that we should opt for hands-free devices.

5 Related Work

Shimanuki et al. [11] argued that some people had difficulties with the origami drill books because paper books could not show all portions of the shape. Their research proposed a vision mechanism to extract diagrams from the drill books and converted them into digital graphics that can be explored in 3D spaces. The most-related researches are the interactive system to support folding operations by Watanabe and Kinoshita [5] and their earlier works [12], and [13]. Nonetheless, their system uses an algorithmic approach to recognize folding processes via a web-camera system while ours is an enhancement in design to incorporate augmented reality technology to a smartphone application.

Nobu Kobayashi [14] introduced Goal Structuring Notation (GSN) to the process of origami folding. His goal was to study the implicit knowledge of folding to "make it beautiful" and to make the knowledge explicit by GSN. We are not concerned about folding beautifully, but rather to finish it, although GSN can be used in the same manner for our work. Furthermore, if folders follow the diagram precisely, shape recognition is easier.

Different works have provided practical examples of technology support for memory capacity limits. Only a few of them are listed here. Cook's Collage [15] provides surrogate memory support for memory-slipped cooking tasks using a display of six most recent actions in a temporal sequence. Geo-reminder, a type of mobile applications to remind its users based on location as well as time, is often shipped with the smartphone platform.

6 Conclusion and Future Work

In this paper, we are the first to propose a mobile application to aid origami folding using the augmented reality technology. We investigated how folders consult the drill books for guidance of steps through a series of user studies. Results from the requirement studies were carefully used in the design of our prototype. The prototype application we developed are able to recognize 8 origami models. Once recognized, the diagrams and instructions of the next folding operations are superimposed on the real paper model in the camera view. Experimental evaluation has proved that the application is able to perform correctly most of the time. The user satisfaction is quite positive. Our application is expected to come in handy for people who have problems recalling origami folding sequences and those who have problems reading diagrams and instructions.

The application is a preliminary prototype and therefore still has many more rooms for improvement in terms of usability and technologies used. In order to make the application more usable, we have to overcome the limitation of the fixed paper pattern, possibly by implementing an origami recognition that is specific to detect creases, faces, corners, and edges; for instance, works by [5] and [16]. An even harder question is how can we identify wrong foldings and suggest the correction. It is interesting to know whether the time-to-completion when using our application is better than searching through a drill book, and if yes, in which circumstances. Yet, less time-to-completion does not always mean more pleasure because sometimes struggling and recalling in between processes are the elements of satisfaction, not only the finished product. In games, we feel fun because of the playing we attempted, not just because the boss are finally killed.

Acknowledgment This research was supported by a grant from Srinakharinwirot University, Thailand, under grant no. 778/2558. The authors express our warm thanks to Dr. Thitima Srivanatakul for her comments and to Srinakharinwirot University for the travel grant.

References

1. Kuribayashi, K., Tsuchiya, K., You, Z., Tomus, D., Umemoto, M., Ito, T., Sasaki, M.: Self-deployable Origami Stent Grafts as a Biomedical Application of Ni-rich TiNi Shape Memory Alloy Foil. Materials Science and Engineering: A **419**(1–2), 131–137 (2006)
2. Smith, L.M.: Nanostructures: The Manifold Faces of DNA. Nature **440**(7082), 283–284 (2006)
3. Nishiyama, Y.: Miura Folding: Applying Origami to Space Exploration. International Journal of Pure and Applied Mathematics **79**(2), 269–279 (2012)
4. Randlett, S.: The Art of Origami; Paper Folding, Traditional and Modern, 1st edn. E P Dutton, June 1961
5. Watanabe, T., Kinoshita, Y.: Folding support for beginners based on state estimation of origami. In: Proceedings of IEEE Region 10 Conference TENCON, pp. 1–6 (2012)
6. Johnson, L., Brown, M., Becker, S.A., Cummins, M., Diaz, V.: NMC Horizon Reports (2015)
7. Unity - Game Engine. https://unity3d.com/
8. Qualcomm Vuforia Developer Portal. https://developer.vuforia.com/
9. Brooke, J.: Usability Evaluation in Industry, chap. SUS: A Quick and Dirty Usability Scale, pp. 189–194. Taylor and Francis, London, June 1996
10. Sauro, J.: A Practical Guide to the System Usability Scale: Background, Benchmarks and Best Practices. CreateSpace Independent Publishing Platform (2011)
11. Shimanuki, H., Kato, J., Watanabe, T.: Recognition of folding process from origami drill books. In: Proceedings of Seventh International Conference on Document Analysis and Recognition, pp. 550–554 (2003)
12. Kinoshita, Y., Watanabe, T.: Estimation of Folding Operation Using Silhouette of Origami. IAENG International Journal of Computer Science **37**(2) (2008)

13. Kato, J., Watanabe, T., Nakayama, T., Guo, L., Kato, H.: A model-based approach for recognizing folding process of origami. In: Proceedings of Fourteenth International Conference on Pattern Recognition (1998)
14. Hiranabe, K.: How to make it better? - Extracting implicit knowledge in Origami (paper folding) via GSN. http://astahblog.com/2014/08/24/
15. Tran, Q.T., Calcaterra, G., Mynatt, E.D.: Cook's collage: deja vu display for a home kitchen. In: Proceedings of Home Oriented Informatics and Telematics (HOIT 2005), York, UK (2005)
16. Zhu, K., Fernando, O.N.N., Cheok, A.D., Fiala, M., Yang, T.W., Samani, H.A.: A SURF-based natural feaure tracking system for origami recognition. In: Proceedings of 20th International Conference on Artificial Reality and Telexistence, Adelaide, Australia (2010)