

Origami Foldaway Support for Beginners using Image Processing

L. Gayathri^{1*}, Praveen Kumar²

Department of Computer Science, Christ University, Bangalore.

*Corresponding author E-mail: gayathri.l@cs.christuniversity.in

Abstract

Various origami works are distributed as origami books in which a succession of collapsing operations with basic outlines is portrayed. In any case, the origami book reader often will give up the instruction of the book middle, because it is too difficult to understand a way to fold in accordance with the diagrams. This paper proposes an approach to find the next step how to do the folding operation, especially for origami beginners. First, a method of detecting the folding operation based on camera images is been detected by canny edge detection. Then, in order to get the next operation camera image is been compare with the database images with the help of Bag of a visual word and Speed up robust features(SURF) detector to detect the key points by finding out the nearest neighboring distance ration (NNDR) measures to find out the similarities.

Keywords: Origami, Folding Operation, Canny Edge Detection and Bags Of Features.

1. Introduction

Origami can be split into two part: 'Ori' means folding and 'kami' means paper so all together it describes paper folding which is the most conventional culture of Japanese. Origami can be an open door for Japanese to take in the youth into it and a few people are taught by their parents and grandparents, whereas others are educated by their teachers at a school. Then again the origami workshops are introduced as one of the global trade. Origami makes imitate long back works and makes new works. These days various works are distributed as an origami instruction books by the users. In the book, origami works are spoken as guideline and outline which clarify the instruction steps from a sheet of square paper to the finalized origami models. But still, it is a bit very hard to understand the diagrams for many people, especially for beginners.

The difficulty can be due to the following three reasons. First, normally a folding process is explained by a series of origami instructional steps in the diagrams which cannot show all the portions of the origami model.

Second, a few operations are regularly instructed as one representation due to space constraint. In spite of the fact that specialists can work out regardless of the possibility that some progress is skipped, for beginners, one illustration should be represented briefly.

Finally, since the learners are untrained in folding operations, the folded sheet of origami paper turns out to be poorly molded and well ordered. If the sheet is folded incorrectly at individual step then the further steps will become more difficult.

The works related to the propose method is explained in Section II, The proposed method of the this paper is explained in Section III, the results obtain from the proposed method is shown in Section IV and the overall conclusion is explained in section V.

2. Related work

In many origami books the basics steps have been described as shown in figure 1, some of the basic folding operations are Mountain folding, valley folding, kite folding, cupboard folding, shawl folding and cushion folding [10]. This basic folding is very useful for construct origami models like table and cup. Until the origami beginners understand these basis folding operations they find more difficult to construct the origami model. Therefore, origami experts have developed many frameworks in order to match user origami paper with the diagram,

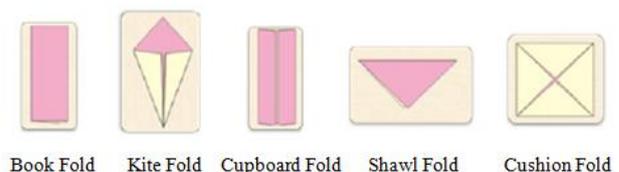


Fig 1: Basic folding [10].

Hiroshi Shimanuki et al.[1] proposes an approach to correction of mistakes on the folding operations especially for origami beginners. First, a method of detecting user's mistakes based on camera images is described. Then, in order to constitute intuitively understandable diagrams, the contents overlaid onto camera images are considered [1].

Hiroshi Shimanuki et al.[1][2] has designed a framework system for identification of folding process of origami instruction book which gives a sequence of folding origami illustration and the illustration is been converted into 3-dimentional animation automatically. So that origami beginners will be able to view the folding steps form different viewpoint. The instruction model will

maintain the change of origami state if there is any explanation of the folding process and it plays an important role to estimate the origami folding phase the model is useful for reconstructing the recognized folding process with the help of CG simulator [2].

Yasuhiro Kinoshita et al.[3] have proposed a method for origami supporting process for origami work.

The main focus on folding operation which represents a view to understand the state of origami and estimate the folding operations by using single camera image.

The folding operation is been estimated by comparing the origami shape in the book for the shapes. A silhouette model was introduced to grasp the association of state of an origami with a single camera image and the folding operation can be estimated by the silhouette model[3]. Toyohide Watanabe[4] proposed a framework which aims to teach the folding operation for the beginners in an understandable manner. Toyohide Watanabe[4] have mainly introduced a model called instruction model which represents a sequential process of origami work. In this model, it consists of a sequence of steps in the process of origami instruction states and each step can be differed from each other states.

Once the beginners finish a folding operation system will estimate his/her origami work state by comparing the shape of origami work with the model.

The next folding operational state will be decided based on his/her origami folding state. For estimating the location, the system uses a silhouette model which represents an instruction state of origami and in order to detect the location of origami, the silhouette model which extract origami sheet leaving the user's hand region is used as a processed image[4].

Takashi Terashima et al.[5] approach a method for recognizing instruction of origami from an origami book with a viewpoint to understand the folding process of origami to overcome the problem Takashi Terashima et al.[5] has purposed a method for comparing and matching two kinds of graphic. First approach is to represent and define a graph which can generate an origami instruction images by applying image process techniques.UGS can also consist of line segments equivalent to edges of origami and their intersections which is also called as features points[4][5]. The other is defined as a graph which is generated by projecting the internal model and a data structure which maintains 3D information about changes of origami states step by step and on the 2D plane which from the same viewpoint and tells that as the next illustration. T. Terashima [4][5] In this paper the system compare and match the USG with one or more possible ISGs to determine the correct folding operation at each step. These methods are based on Hough transformation which takes into an account the characteristic of an origami state instruction image.

Kenta Tanaka et al.[12] have proposed an origami folding model as a human hand. Though it does not look like a human hand its works are the same as the works which a human hand can do. They aimed mainly in understanding the movements of human hands and develop an automated system for hands. For this purpose they have taken origami folding as a component and based on that they have analyzed the dexterity of the hands of humans. Finally they developed an automated system to fold a tadpole.

Though it involved human intervention in the process of image segmentation the folding was perfect as if a human has folded. This paper followed a bottom up approach to build an automated system[12].

The instruction model plays important role in our system, because both estimation and instruction of folding processes are based on this instruction model it maintains the state transition of origami work step by step[1]. In order to represent the folding processes, a data structure defined by Miyazaki et al. [6] is used. The node represents each state of origami paper, and the arrow means state transition by a folding operation (OP)[1].

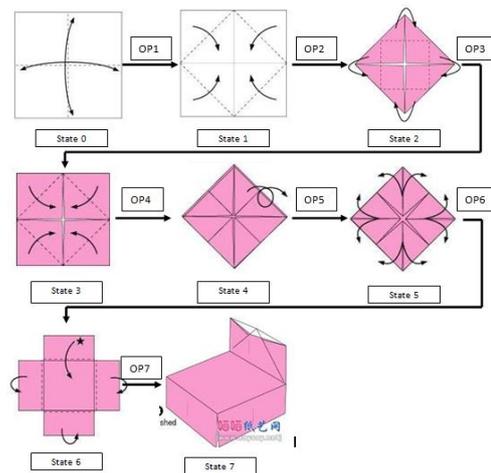


Fig 2: A instruction model representing the folding process of "Chair".

3. Proposed approach

The proposed method adopts the following techniques, where each technique can be done once after finishing the previous technique since the previous technique output is given as input for the next technique. The proposed method involves scanning or capturing of the origami image and it converts the image into standard formats like jpeg, png etc. once the image is inputted in to system, median filtering, as the first step is processed since median filtering removes the noise such as impulsive noise, salt and pepper noise by preparing the image for further processes such as Canny edge detection to extract the edge and to get the accurate edge threshold value of 0.1 and 0.9 is been given and in order to get the connected edges a morphological process is involved. Once after a morphological process is done processed image is been taken into content-based image retrieval to retrieve the image from the database.

3.1. Median filter

Median filtering is a type of filtering method which is a nonlinear method used to remove noise from the image. The median filters can user impulsive, salt and pepper noise. But in the proposed model impulsive noise is used in order to get the smoothen image.

$$M(x, y) = \text{median} \{n[i, j], (i, j) \in \omega\} \quad (1)$$

$M(x, y)$ represents center of the matrix with row x and column y . $n[i, j]$ are elements which belongs to ω . Here, ' ω ' is set of window elements of $M(x, y)$.

3.2. Canny edge detection

Canny edge detector is an edge detection operator which is used to extract useful structural information such as detection of edges with low error rate. The accurate edge can be obtained by varying the two threshold values. The process of detecting the edges using canny edge detection can be done in the following steps. First Gaussian filter is applied to smoothen the image and to remove noise in it. Then the intensity gradients are taken from that image. Then a non-maximum suppression matrix method is applied to remove the weak edges in the image. A strong threshold value is applied twice and a number of edges are taken. Finally, the edges which are not weak and which are not connected to strong edges are suppressed and the rest all are taken to be the edge detected[8][11].

The smoothing can be computed and denoted by $I[x, y]$ to an image. $G[x, y, \sigma]$ where G is Gaussian filtering and σ is the degree for smoothening the image. The result can be represented in array form

$$S[x, y] = G[x, y, \sigma] * I[x, y] \quad (2)$$

The equation for a Gaussian filter kernel of size $(2z+1)*(2z+1)$ can be written as:

$$G_{xy} = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(x-(z+1))^2+(y-(z+1))^2}{2\sigma^2}\right); 1 \leq x, y \leq (2z + 1) \tag{3}$$

To find the intensity gradient, In image after detecting the edge, the edges may vary the directions, so the canny edge detector use four filters to detect horizontal, vertical and diagonal edges in the image. The edge detector returns a value for the first derivative in the horizontal direction and vertical direction which is represented as G_x and G_y [8].

$$G = \sqrt{G_x^2 + G_y^2} \tag{4}$$

$$\theta = \text{atan2}(G_x, G_y) \tag{5}$$

3.3. Morphological processing

The morphological process is a collection of nonlinear operation related to the shape features of the image with the binary values and it is a process used to make the thicker edges thin. The morphological process consists erosion and dilation. In proposed method erosion operation is used.

Erosion of binary image 'a' by a structuring element 'b' produces a new binary image $h = a \ominus b$ with all ones locations (x,y) of a structuring elements which fits the input image 'a' i.e. $h(x,y) = 1$ is 'b' fits 'a' and 0 otherwise[7]. Erosion with small square structuring elements of size $2*2$ or $5*5$ which shrinks an image by stripping away a layer of pixels for both inner and outer boundaries of regions[7].

The gaps and holes between different regions become larger and small details are eliminated so the larger structuring elements have a more pronounced effect which are obtained by iterating erosion using a smaller structuring element of the same shape. If a_1 and a_2 are structuring element identical in shape, with a_2 twice the size of a_1 , then

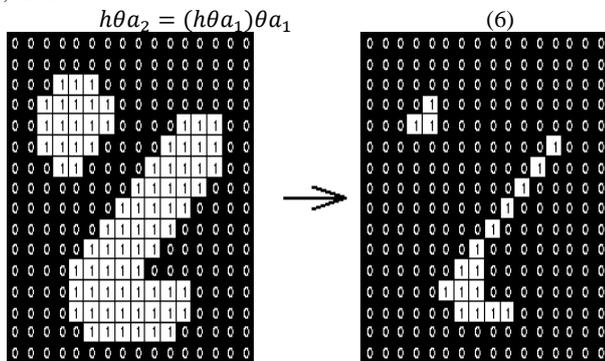


Fig 3: Erosion operation of an image [7]

3.4. Content based Image Retrieval

Content-based image retrieval is an approach which is used for retrieval of an image from the database. Retrieval of an image can be based on color, shape, and texture within a large database [9]. The features used to retrieve the image can be primitive or semantic.

Basically, content-based image retrieval has two technique: bags of features and k-means classification. Bags of visual words approach a method for image retrieval, text retrieval of visual classification and *k-means* which are alone designed for classification with the help of SURF detector and nearest neighboring distance ration (NNDR) measures to find out the similarities and primary application of k means is clustering or unsupervised classification which can also adapt to for supervised classification.

Steps for Image Retrieval

Step1: Create image Set

Step2: Extract the bags of visual words using SURF detector

Step3: Attain image index

Step 4: Search the image

Step5: Obtain the image from the image Set.

3.5. Flow Diagram

In this proposed Method Figure 4 shows the overall process model for origami folding operational support for the beginners with the help of image processing,

Where training and testing represent the process state of the image and the database represents the image storage directory such that testing image can be compared with the database and retrieves the current folding operational image and with the current folding operation image next folding operation is obtained and display as output image.

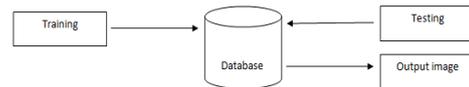


Fig 4: A overall process model of proposed method.

Training Phase

The proposed algorithm for training the origami steps of an image. In this process, user gives an origami folding operation image as an input such that the input goes all the process steps shown in figure 5 and the image is been saved to the database in .jpg format. So that it can be considered as a training set of images.

Step 1: Read Training image data

Step 2: Perform median filtering to smoothen

Step 3: Convert to Grayscale

Step 4: Detect edge using canny edge detector

Step 5: Apply threshold value of 0.1 and 0.9 to get the accurate edges.

Step 6: Convert the detected edge to binary form and reduce the thickness of edge with Morphological

Step 7: Store the image in the database as .jpg format.

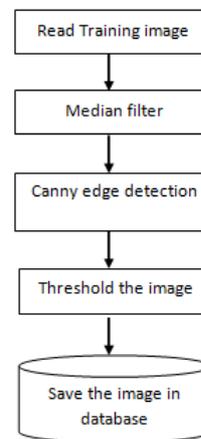


Fig 5: Process 1 for training image of an origami shape

Testing phase

The proposed algorithm for testing the origami steps of an image. In this process, origami beginners give an origami folding operation image as an input such that the input goes all the process steps shown in figure 6. So that the processed image can be considered for the image comparison with the trained database of images in order to retrieve the current folding operation image and from the current folding operational image the next folding operation can be obtained.

Step 1: Read testing image data

Step 2: Perform median filtering to smoothen

Step 3: Convert to Grayscale

Step 4: Detect edge using canny edge detector

- Step 5: Apply threshold value of 0.1 and 0.9 to get the accurate edges.
- Step 6: Convert the detected edge to binary form and reduce the thickness of edge with Morphological process
- Step 7: Extract bag of visual words from morphological image.
- Step 8: Search and compare the image in database using SURF detector and nearest neighboring distance ration with *k-means* clustering.
- Step 9: Display image of the next step in building the origami

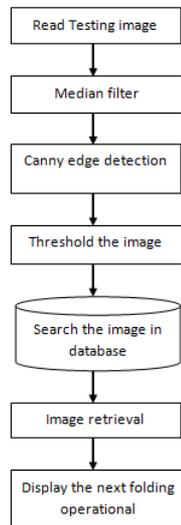
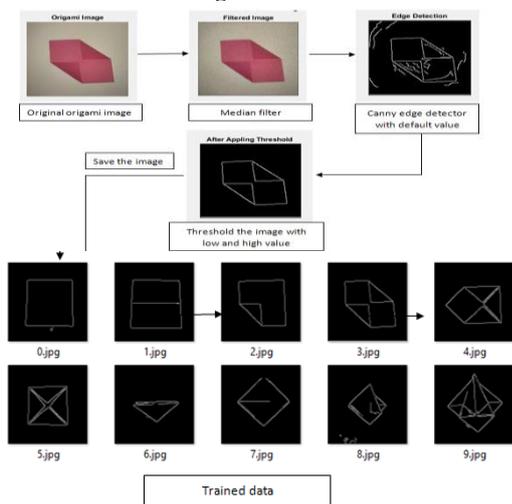


Fig 6: Process 2 for testing image of an origami shape

4. Experimental result

The proposed approach is implemented in MatLab version R2015a with system specification of windows 10, RAM 4 GB and CPU Intel core i5 processor speed. Since various origami model has different folding steps in this paper, for the proposed method homemade dataset such as each folding operation is been captured using a camera and it is been saved as a .jpg format in a directory. So that the folding operational image is been taken into system as shown in figure 7 to extract the edge using the following sequential steps like median filtering which helps to smoothen the image, canny edge detector with default is used to detect the edge form origami folding operational image, canny edge detector using threshold value of 0.1 and 0.9 to get the accurate edges and content-based image retrieval is used to retrieve an image from the database with the help of visual words and SURF detection to get the nearest neighboring distance ration (NNDR) measures. As explained in section D for each folding operational image. The following steps show the implementation which gives the exact result as shown in figure 8.



As shown in the above figure 7 the initial step for the training process is to apply a median filter for smoothening the inputted origami image then the default canny edge detector is been applied on the median filtered image in order to extract the edges of the origami image. 0.1 and 0.9 threshold value is been applied to obtain the accurate edge. The resultant image of the mentioned process is stored in the database as trained data for the following process.

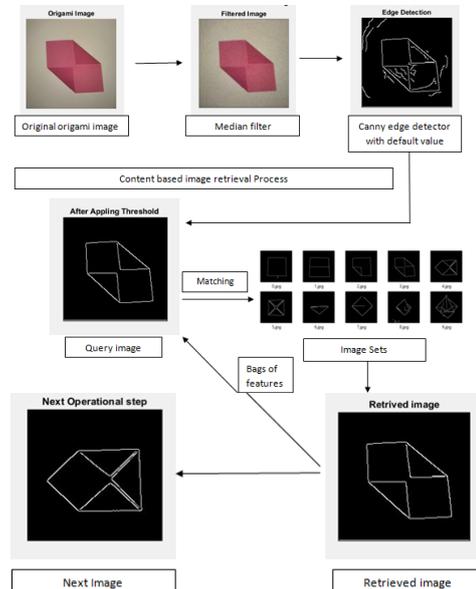


Fig 8: Illustration of the testing process

Figure 8 illustrates the process of testing in which median filter and canny edge detection with default and threshold value of 0.1 and 0.9 have been applied to the input image of the testing process. In the content-based image retrieval process, the output of the threshold image is compared with the image sets in the database. Once the appropriate match is found for the query image, the image is retrieved with the help of SURF detector which is used to detect the key point by finding the NNDR measures to find the similarities. To complete the testing process the succeeding step is obtained with the help of the retrieved image.

5. Conclusion

This paper gives an operational support for the origami beginners to find out the next operational step with the image. The proposed method is used to find the next folding operational step when the user does not know what the next folding operation is. Future analysis, the proposed method can be used for acquiring the time consumed by various origami users and the collected data can be used for further analysis and similar research purposes.

References

- [1] Shimanuki H, Yasuhiro K, Toyohide W, Koichi A & Hideki S, "Operational Support for Origami Beginners by Correcting Mistakes", *In Proceedings of the 10th International Conference on Ubiquitous Information Management and Communication*, (2016).
- [2] Shimanuki H, Jien K & Toyohide W, "A recognition system for folding process of origami drill books", *In International Workshop on Graphics Recognition*, Springer, Berlin, Heidelberg, (2003), pp.244-255.
- [3] Kinoshita Y & Toyohide W, "Estimation of folding operation using silhouette of origami", *IAENG Int. J. Computer Science*, Vol.37, No.2, (2008), pp.1-8.
- [4] Watanabe T & Yasuhiro K, "Folding support for beginners based on state estimation of origami", *In TENCON 2012-2012 IEEE Region 10 Conference*, (2012), pp.1-6.

- [5] Kato J, Toyohide W, Hiroyuki H & Takeshi N, "Understanding illustrations of origami drill books", *Transactions of Information Processing Society of Japan*, Vol.41, No.6, (2000).
- [6] Miyazaki S, Takami Y, Shigeki Y & Jun T, "An origami playing simulator in the virtual space", *Journal of Visualization and Computer Animation*, Vol. 7, No. 1, (1996), pp.25-42.
- [7] <https://www.cs.auckland.ac.nz/courses/compsci773sl/lectures/ImageProcessing-html/topic4.htm#refsl>
- [8] [https://en.wikipedia.org/wiki/canny edge detector](https://en.wikipedia.org/wiki/canny%20edge%20detector)
- [9] [https://en.wikipedia.org/wiki/content-based image retrieval](https://en.wikipedia.org/wiki/content-based_image_retrieval).
- [10] Zhu K, Owen Noel Newton F, Adrian David C, Mark F, Theam Wei Y & Hooman Aghaebrahimi S, "A SURF-based natural feature tracking system for origami recognition", *In Proceedings of 20th International Conference on Artificial Reality and Telexistence, Adelaide, Australia*, (2010).
- [11] Canny J, "A computational approach to edge detection", *IEEE Transactions on pattern analysis and machine intelligence*, Vol.6, (1986), pp.679-698.
- [12] Tanaka K, Yusuke K & Yasuyoshi Y, "Origami folding by a robotic hand", *IEEE/RSJ International Conference on Intelligent Robots and Systems*, (2007), pp. 2540-2547.
- [13] http://fourier.eng.hmc.edu/e161/lectures/smooth_sharpen/node2.html