折り図の自動生成に向けた折り紙画像からの折り線推定

加藤　眸[†]　加藤　大貴[†]　平山　高嗣[††]　駒水　孝裕[†]　井手　一郎[†]

† 名古屋大学　〒464–8601 愛知県名古屋市千種区不老町
†† 人間環境大学　〒444–3505 愛知県岡崎市本宿町上三本松 6–2
E-mail: †{katoht,katoh}@cs.is.i.nagoya-u.ac.jp, ††t-hirayama@uhe.ac.jp,†taka-coma@acm.org,†ide@i.nagoya-u.ac.jp

あらまし　折り紙の折り方を説明する方法として一般に折り図と折り方動画がある．折り図の制作には多くの手間と時間を要すること，スマートフォンなどの普及により折り方動画の撮影が容易になりつつあることを踏まえ，筆者らは任意の折り紙の折り方動画から折り図を自動的に生成するシステムの制作を目指している．本研究では，このようなシステムを実現する上で解くべき部分問題を整理し，その中の 1 つの特別な場合である，直線折り操作における折り線推定に取り組んだ．そのための手法として，本報告では折り操作前後の画像（以下，折り状態画像対）と，それらの間で行なわれた折り操作を表す線分との関係を学習する折り線推定手法を提案する．また，折り線推定器の学習と評価のためのデータセットを作成した．評価実験では，折り線推定器として畳み込みニューラルネットワーク（Convolutional Neural Network; CNN）を用いたモデルを実装し，折り線の推定性能を評価した．その結果，データ拡張によりデータ数を増やすことで，推定された線分の距離の誤差および傾きの誤差が小さくなること，直線折りの中にも，推定が比較的容易なものと難しいものがあることを確認した．
キーワード　折り紙，折り図，直線検出

Crease Estimation from Origami Pictures for Automatic Folding Diagrams Generation

Hitomi KATO[†], Hirotaka KATO[†], Takatsugu HIRAYAMA[††], Takahiro KOMAMIZU[†], and Ichiro IDE[†]

† Nagoya University　Furo-cho, Chikusa-ku, Nagoya, Aichi, 464–8601 Japan
†† University of Human Environments　6–2 Kami Sanbonmatsu, Motojuku-cho, Okazaki, Aichi,444–3505 Japan
E-mail: †{katoht,katoh}@cs.is.i.nagoya-u.ac.jp, ††t-hirayama@uhe.ac.jp,†taka-coma@acm.org,†ide@i.nagoya-u.ac.jp

Abstract　In general, folding diagrams and videos are provided for explaining how to fold origami. However, creating folding diagrams requires a lot of time and effort. Considering that it is becoming easier to take videos of how to fold origami with the spread of smartphones and other devices, we aim to create a system that automatically generates folding diagrams from a video of how to fold any origami. First, we organized the subproblems that need to be solved in order to realize such a system. In this report, we focus on one special case of crease estimation in linear valley fold. For this, we propose a crease estimation method that learns the relationship between images before and after a folding operation (hereinafter, folded state image pairs) and the line segment that represents the folding operation performed between them. We also created a dataset for training and evaluating the fold line estimator. In the evaluation experiment, we implemented a model using a Convolutional Neural Network (CNN) as a fold line estimator and evaluated the fold line estimation performance. As a result, we confirmed that by expanding the data, the error in the distance and inclination of the estimated line segments was reduced, and that even among linear folds, there are some that are relatively easy and difficult to estimate.
Key words　Origami, folding diagrams, line detection

1. はじめに

1.1 研究の背景

折り紙によって様々な教育効果が得られることが期待され，難解な折り紙の折り方を説明した書籍が多く出版されるなど，折り紙を学ぶことに対する需要が高まってきている．折り紙の折り方を伝える方法として一般に折り図と折り方動画がある．両者にはそれぞれ長所と短所があり，双方の欠点を補い合うこ

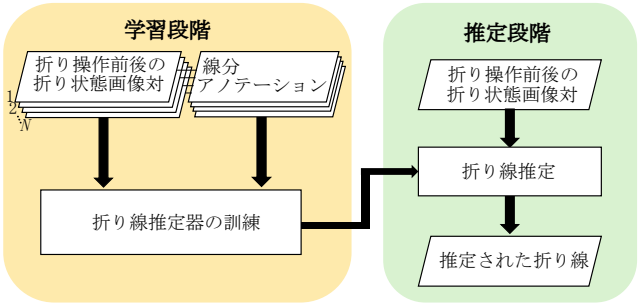

図 2　提案手法の概要．

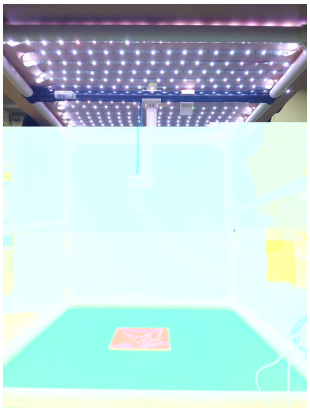

図 3　撮影環境．天井部に白色 LED テープライトを設置し，装置全体を暗幕で覆うことで光源を管理している．

2.1　学 習 段 階

学習段階では，折り操作前後の折り状態画像対と，教師データにあたる折り線の線分データを組として学習する．以下，学習段階全体を定式化する．

学習および推定に用いる折り状態画像対は，任意の状態の折り紙に対して行なわれる任意の 1 つの直線折り操作について，その操作前の画像と操作後の画像である．以後これらの折り状態画像対をそれぞれ I_{i1}, I_{i2}（$i = 1, 2, \cdots, N$）と表す．ただし N は学習データの総数を表す．

線分データは，折り操作によって加えられた折り線について，画像上の両端点の座標を記述したものであり，折り状態画像対に対して線分データがアノテーションされる．以後この線分アノテーションを l_i と表す．線分データの具体的な表現方法については 3.2 節で述べる．折り状態画像対と折り線との関係を学習させることにより，折り線推定器を得る．折り線推定器は折り状態画像対 I_{i1}, I_{i2} を入力とし，折り線を出力する．折り線推定器を f，そのパラメータを Θ と表し，折り線推定器が出力する折り線を $\hat{l}_i$ と表すと，折り線推定は式 1 のように定式化される．

$$\hat{l}_i = f(I_{i1}, I_{i2}; \Theta) \tag{1}$$

また，2 つの線分 l, l' の違いを表す関数を $E(l, l')$ とすると，折り線推定器の訓練は，式 2 に示す Θ の最適化問題として定式化される．折り線推定器には画像を扱える任意の機械学習手法を使用できる．そのため，具体的な実装や構造については 4.1.1 で述べる．

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmin}} \sum_{i=1}^{N} E\left(l_i, f(I_{i1}, I_{i2}; \Theta)\right) \tag{2}$$

2.2　推 定 段 階

推定段階では学習段階で得られた折り線推定器に対し，折り状態画像対を入力することで折り線推定を行なう．すなわち，訓練された折り線推定器 f に対し，未知の折り状態画像対 I_{i1}, I_{i2} を入力し，折り線 $\hat{l}_i$ を出力させることになるため，この推定は式 3 のように表される．

$$\hat{l}_i = f(I_{i1}, I_{i2}; \hat{\Theta}) \tag{3}$$

3.　データセット

本研究では，49 種類の折り方について学習用データとテスト用データからなるデータセットを作成した．学習用とテスト用のデータに含まれる折り方は全く同じ種類のものとし，それぞれの折り方で折り始める度に新しい折り紙を使用した．また，撮影後には折り操作前後の折り状態画像対に対して折り線の線分アノテーションを付与し，学習用とテスト用で各 49 組のデータを得た．

3.1　折り状態画像の撮影

本研究では，折り紙制作者の手元を真上から撮影した動画像を用いて折り操作の結果加わる折り線を推定することを想定する．そのため図 3 に示す撮影環境を用意して折り状態画像を撮影した．撮影面にはグリーンバックを設置し，折り紙が画面中央に位置するように置いて撮影した．また，線分データをアノテーションしやすくするために，折り操作前後で折り紙の位置を極力維持するように配慮した．画像は RGB 各色 8 bit，解像度 1,024 × 1,024 画素の正方形で撮影した．カメラは撮影面から高さ 35 cm の地点に固定した Point Grey 社製の Flea3 を使用し，外光の影響を受けないように装置全体を暗幕で覆い，装置の天井部には SwitchBot 社製の白色 LED テープライトを 2 本（全長合計 10 m）設置した．また，折り紙用紙は学習用，テスト用データともに市販されている一般的なもの（トーヨー社製の単色おりがみだいだい色 15.0 cm）を使用した．

また，撮影後に学習のための事前処理として背景除去と縮小を行なった．具体的には，グリーンバックの上に折り紙を置いて撮影したことを利用し，各画素の RGB 値の R，B がともに G の $\frac{2}{3}$ より小さい場合にはその画素の RGB 値をすべて 0 にするという閾値処理によって背景を除去した．この除去の閾値 $\frac{2}{3}$ は実験的に定めた．続いて，Python の OpenCV モジュールの `cv2.resize` 関数を使用し，バイリニア補完によって画像を 1,024 × 1,024 画素から 224 × 224 画素に縮小した．

3.2　線分データのアノテーション

線分データは折り紙に対して直線折り操作が行なわれた際に加えられた折り線を表す 1 本の線分であり，その両端点の