

# アフィン変換を用いた全方位カメラ画像中の活字文字認識に関する研究

## A Study on Recognition Method of Printed Characters in Omnidirectional Camera Image Using Affine Transformation

高野征直

指導教員 鈴木雅人, 北越大輔, 西村亮

東京工業高等専門学校 情報工学科 知識情報研究室

キーワード: 局所特徴量, アフィン変換, 全方位カメラ, 文字認識

### 1. はじめに

視覚障害者の支援[1]等で広範囲から画像情報を得たい場合, 全方位カメラが用いられることがある. しかし全方位カメラは通常のカメラよりも画像の歪みが大きいため, 全方位カメラで撮影した画像(全方位画像)を用いて画像認識や文字認識を行うと認識精度が低下する傾向がある. そこで本研究では, 全方位画像中の活字文字に対する認識精度改善に関する手法を検討する. 文字位置によって歪み方が異なる全方位画像を小領域に分割し, それぞれの領域の歪みをアフィン変換で補正することで, 歪みのない文字画像に近い局所特徴量を抽出して認識精度を改善する.

### 2. 局所特徴量のマッチング

局所特徴量は, 文字を構成する線分のコーナーや境界を特徴点として検出し, その周辺画像情報をベクトルで表したものである. 局所特徴量のマッチングでは, テンプレート画像と撮影した画像の局所特徴量とを比較し, 類似度の高い特徴点が集中している場所を探すことで文字の場所や種類を認識している. 本研究では局所特徴量の1つであるAKAZE特徴量[2]を用いて文字画像のマッチングを行う.

### 3. 立方体マッピング画像の歪みの補正

全方位画像は真上と真下付近ほど歪みが大きく,

それを歪みが無い状態に戻すのは困難であるため, 立方体マッピングを用いる. 立方体マッピングとは, 球面に撮影した画像を立方体に投影して, 図1のような6枚の正方形画像から構成される画像(キューブマップ画像)に展開することである. しかし立方体マッピングした画像では, 文字が複数画像に分割されたり歪んだりする問題が発生する. 図1は, 格子状の模様の画像を入力し, 立方体マッピングによってどのような歪みが発生するのかを検証したものである. A~Dの画像は四隅に近づくほど歪みが大きくなり, E, Fの画像は全体的に円形の歪みが発生している.

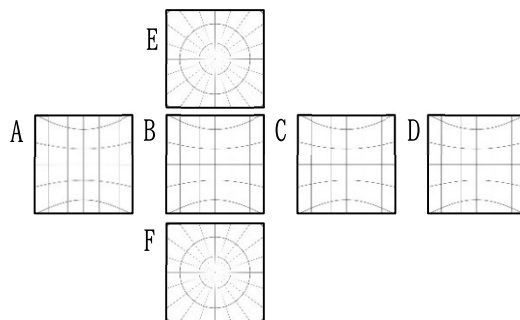


図1 格子模様のキューブマップ画像

### 4. アフィン変換を用いた歪みの補正

本研究では, 図1のA~Dの画像の歪みをアフィン変換で補正する手法を検討した. アフィン変換とは, 点  $(x, y)$  を以下の式(1)の線形変換により

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} e \\ f \end{bmatrix} \quad (1)$$

点 (X, Y) に移す変換であり, 画像の拡大と縮小, 回転, 平行移動, せん断を行うことができる. 図 1 の A~D の画像において, 垂直方向の線の歪みはなく, 水平方向の線のみが歪んでいる. そこで今回の提案手法では, 水平方向の線の歪みをせん断により補正した. 補正では, A~D の画像を縦に 2 分割し, それぞれに対してせん断を加える. 図 2 は左半分にせん断を加えた様子を表している. 右半分についても同様の処理を行い, 2 つの画像を結合して最終的な歪み補正画像を得る (図 3).

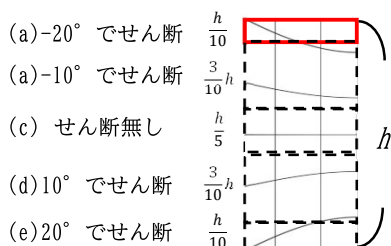


図 2 アフィン変換の過程

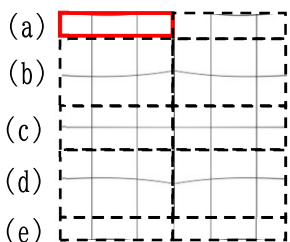


図 3 アフィン変換により得られる最終的な画像

5. 実験と結果

提案手法の精度を評価するために, 全方位画像を用いて実験を行った. カメラの正面を 0 度とし, 水平方向にカメラを 30 度ずつ回転させて平仮名の「あ」を撮影し, A~D のいずれかに「あ」が写る

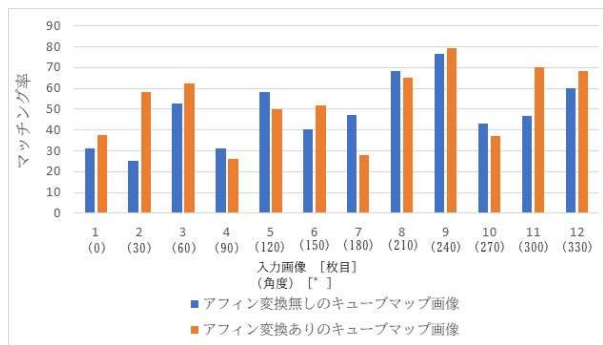


図 4 アフィン変換の効果を確認するための実験の結果

ような 12 枚の画像を生成した. 提案するアフィン変換の有無と局所特徴量のマッチングとの関係を調査したところ, 図 4 のような結果が得られた. マッチング率とは, 予め算出された「あ」の局所特徴量とキューブマップ画像の局所特徴量を比較し, 文字部分と対応が取れた特徴点の数を画像全体で対応が取れた特徴点の数で割ったものである. 図 4 より, 入力した 12 枚の画像のうち 7 枚が, アフィン変換前よりもアフィン変換後のマッチング率が高くなったため, アフィン変換により歪みを補正する効果がある可能性があることが確認できた. 一方今回の提案手法では, 図 3 のように画像を長方形に切り出して一律にアフィン変換を行った. そのため, 図 1 の A~D の画像の中心部分のように, 本来であれば歪みが少ない場所に文字が写る場合 (カメラを回転させた角が 90 の倍数の場合) は, それ以外の場合よりもマッチング率が 10% 以上低下した.

6. まとめ

本研究では, 全方位画像を小領域に分割してアフィン変換を行い, 全方位画像に映り込んだ文字の認識精度を改善することを提案した. 提案手法の有効性を確認する実験では, 全方位カメラを用いた文字認識においても, 前後左右の部分ではアフィン変換を用いることで全方位カメラの歪みを補正し, 認識精度を改善できる可能性があることが分かった. 今後は, 図 1 の A~D の正面付近のアフィン変換の改善や図 1 の E, F の画像の補正手法の提案を行う予定である.

参考文献

[1] 千葉 穂乃美, 宮治 裕, “視覚障害者向け 360 度周辺情報提示システム” 情報システム学会誌 17 巻, 2 号, pp.35-48, Mar. 2022.  
 [2] P. F. Alcantarilla, et al., “Fast Explicit Diffusion for Accelerated Features in Nonlinear Scale Spaces,” Proc. of British Machine Vision Conference 2013, pp.13.1-13.11, Sept. 2013.