

# ラドン・ヒストグラム変換を用いたモバイル文字認識の検討

伊藤 光彦<sup>†</sup> 長谷川 誠<sup>‡</sup>

東京電機大学大学院工学研究科<sup>†</sup>

東京電機大学工学部情報通信工学科<sup>‡</sup>

## 1. はじめに

カメラで撮影した情景画像中の文字を認識する方法について検討する。ここでは、情景画像(図1)中の文字を抽出し、これをラドン・ヒストグラム変換[1, 2]する。ラドン・ヒストグラム変換とは、原画像をラドン変換し、各偏角成分における変換係数のヒストグラムを算出する変換であり、変換結果は、文字の配転、拡大縮小、平行移動に頑健となる。文字のデータセットをラドン・ヒストグラム変換し、変換結果をテンプレートとして予め記録しておく。文字認識する際には、検索対象の文字(クエリ)をラドン・ヒストグラム変換し、テンプレートの変換結果と照合する。

文字認識方法を提案する。カタカナ文字および数字からなるデータセットを作成してシミュレーションする。フーリエ変換を用いる従来方法(フーリエ・メリン変換)よりも認識率が高いことを示す。

## 2. ラドン変換

画像が与えられている二次元領域  $R^2$  の座標を位置ベクトル  $\mathbf{X} = (x, y)$  で表し、 $\mathbf{X}$  における輝度値を  $f(\mathbf{X})$  とする。濃淡画像  $f(\mathbf{X})$  のラドン変換は、デルタ関数  $\delta(t)$  を用いて

$$R_f(\theta, \rho) = \int f(\mathbf{X}) \delta(\mathbf{X} \cdot \xi - \rho) d\mathbf{X} \quad (1)$$

と定義される。なお、 $\xi = (\cos \theta, \sin \theta)^T$  とする。すなわち、原点からの距離が  $\rho$ 、 $y$  軸とのなす角を  $\theta$  とする直線

$$L_{\rho} = \{\mathbf{X} \in R^2 \mid \mathbf{X} \cdot \xi = \rho\} \quad (2)$$

を積分路とした  $f(\mathbf{X})$  の線積分がラドン変換  $R_f(\theta, \rho)$  である。

濃淡画像  $f(\mathbf{X})$  における被写体が回転すると、ラドン変換結果  $R_f(\theta, \rho)$  は横方向(偏角  $\theta$  方向)に平行移動する。被写体が拡大縮小すると、変換結果は縦方向(動径  $\rho$  方向)のみに拡大縮小する。また、被写体が平行移動すると、変換結果は上下方向に波を打つように変形する。



図1 情景文字

## 3. ラドン・ヒストグラム変換

一次元信号  $f(x)$  のヒストグラムを

$$H(f)(v_i) = \frac{\#\{u \in \mathbf{U} \mid |f(u) - v_i| < \frac{\Delta}{2}\}}{|\mathbf{U}|} \quad (3)$$

と定義する。なお、ここでの  $\Delta$  は、量子化幅

$$\Delta = \frac{\max f(u) - \min f(u)}{N} \quad (4)$$

であり、 $N$  はビン数である。

ラドン変換の結果  $R_f(\theta, \rho)$  をその最大値  $\max_{\theta, \rho} (R_f(\theta, \rho))$  をもちいて、

$$R_f(\theta, \rho) \rightarrow \frac{R_f(\theta, \rho)}{\max_{\theta, \rho} (R_f(\theta, \rho))} \quad (5)$$

と正規化し、列ごとのヒストグラム

$$D_f(\theta, v) = H(R_f(\theta, \cdot))(v) \quad (6)$$

を算出する変換が、ラドン・ヒストグラム変換である。なお、ヒストグラムを作成する際には、

$$\mathbf{U}_{\theta} = \{\rho \mid R_f(\theta, \rho) > 0\} \quad (7)$$

のように、 $R_f(\theta, \rho)$  がゼロ以上の値を用いる。

濃淡画像  $f(\mathbf{X})$  における被写体が拡大縮小しても、平行移動してもラドン・ヒストグラム変換の結果  $D_f(\theta, v)$  は不変である。被写体が回転すると、ラドン・ヒストグラム変換の結果  $D_f(\theta, v)$  は横方向(偏角  $\theta$  方向)に平行移動する。ラドン・ヒストグラム変換の結果を図2に

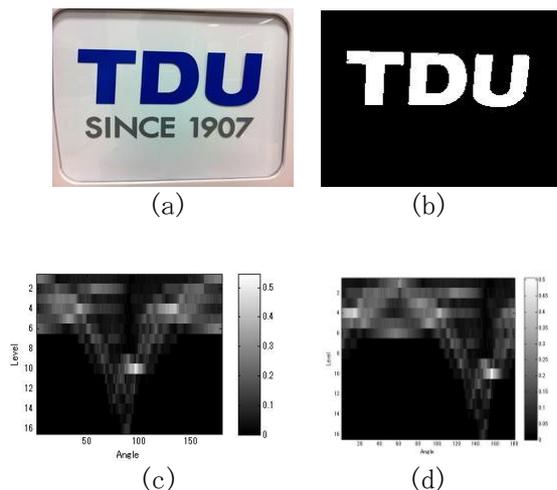


図2 情景画像(a)と文字抽出結果(b) ; (b)のHRT変換結果(c) ; カメラを60°回転させた時のHRT変換結果(d)

示す. 被写体が回転すると, 図2(d)に示すように変換結果は横方向に平行移動する.

#### 4. 文字認識

カメラで撮影した情景画像をHSV変換し, 色相成分(H)の濃淡画像を二値化して文字列を抽出する. 二値化された検索対象画像(クエリ画像)を $f(\mathbf{X})$ とする. ラドン・ヒストグラム変換し, 変換結果 $D_f(\theta, \nu)$ を算出する. 同様に, テンプレートとして $g(\mathbf{X})$ ,  $D_g(\theta, \nu)$ を算出する.

$D_f(\theta, \nu)$ と $D_g(\theta, \nu)$ との誤差

$$E = \min_{\alpha} \iint D_f(\theta, \nu) - D_g(\text{mod}(\theta + \alpha), \nu) d\theta d\nu \quad (8)$$

を算出してマッチングする. なお, ここでの $\text{mod}$ は,  $\theta$ を弧度法で表した場合, 180の剰余である. 閾値 $T$ を設定し,  $E < T$ の場合, クエリ画像の文字はテンプレート画像の文字であると判定する.

#### 5. 実験

文字列を紙に印刷して掲示し, カメラを用いて文字を撮影する. ここでは, カタカナ文字および数字からなる32種類の文字について, 各文字回転, 拡大縮小, 平行移動を含む8枚の画像を撮影し, 計256枚の画像を用いる. 文字どうし総当たりで照合し, 再現率(Recall)と適合率(Precision)を算出する. また, 同じ実験を, フーリエ変換を用いる従来方法(フーリエ・メリ

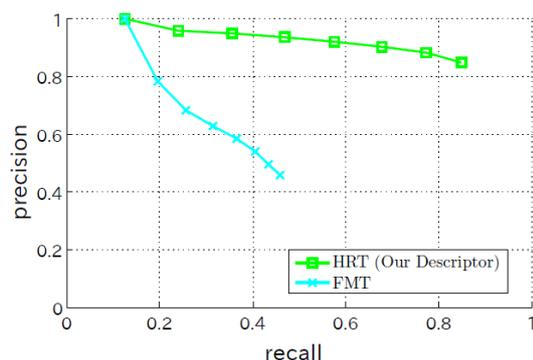


図3 Precision-Recall 特性

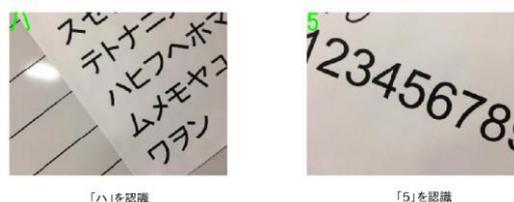


図4 文字認識システム

ン変換)で行った. 実験の結果, 図3に示すように本方法の認識率は, フーリエ・メリン変換よりも高いことが明らかとなった. また, 文字認識システムを構築し, 図4に示すように実現可能性を示した.

#### 6. まとめ

カメラで撮影した情景文字をラドン・ヒストグラム変換して認識する方法を提案した. 本方法の認識率はフーリエ・メリン変換よりも高い.

#### 文献

- [1] S. Tabbone, R. Terrades, and S. Barrat, "Histogram of Radon transform. a useful descriptor for shape retrieval," Proc. ICPR, pp.1-4(2008)
- [2] 長谷川誠, "ラドン変換を用いたアフィン変換に頑健なパターンマッチング," 筑波大学数理物質科学研究科数学専攻解析セミナー(Nov. 2012)
- [3] S. R. Deans, The Radon Transform and Some of Its Applications, Dover Publications (Oct. 2007)

Study on Mobile Character Recognition Using Histogram Radon Transform

† Graduate School of Engineering, Tokyo Denki University

‡ Department of Information and Communication Engineering, School of Engineering, Tokyo Denki University