

[ポスター講演] 動画像を用いた超低解像度文字認識

鳥羽 修平[†] 工藤 裕貴[†] 宮崎 智[†] 菅谷 至寛[†] 大町 真一郎[†]

[†] 東北大学工学部 〒980-8579 宮城県 仙台市青葉区 荒巻字青葉 6-6-05

E-mail: [†] shuhei@iic.ecei.tohoku.ac.jp

あらまし 近年の文字認識技術の向上により、携帯カメラを用いた様々な文字認識アプリケーションが登場してきている。しかし、携帯カメラで撮影した画像中の文字は撮影環境や手ぶれの影響により劣化している場合がある。そこで本研究では、20*20pixel以下の超低解像度な文字画像に対する文字認識システムの開発に取り組んでいる。提案するシステムは、生成型学習による学習サンプル数の増加、ウィナーフィルタと位置合わせ処理による文字画像のぼけ除去及び高解像度化、さらには高速化した相互部分空間法による認識の3段階で構成される。文字認識実験により、各段階のアプローチによる認識率の向上と提案する認識手法による処理時間の短縮を確認した。

キーワード 文字認識, ぼけ除去, 位置合わせ処理, 部分空間法

[Poster Presentation]

Ultra-low Resolution Character Recognition by Utilizing Multiple Frames

Shuhei TOBA[†] Hiroataka KUDO[†] Tomo MIYAZAKI[†]

Yoshihiro SUGAYA[†] and Shinichiro OMACHI[†]

[†] Faculty of Engineering, Tohoku University 6-6-05 Aramaki-aza-aoba, Sendai-shi, Aoba-ku, Miyagi, 980-8579 Japan

E-mail: [†] shuhei@iic.ecei.tohoku.ac.jp

Abstract Improvement of character recognition technology brings us various character recognition applications for mobile camera. However, many low-resolution and poor-quality character images exist, due to the influence of environment or motion blur. Therefore, we develop a character recognition system for ultra-low resolution character images less than 20*20 pixels. The proposed system consists of three phases of increased training data with a generative learning method, creating a deblurred high-resolution image with Wiener filter and image alignment, and recognition by efficient Mutual Subspace Method. We examined our system and confirmed the improvement of accuracy and efficiency of the recognition process.

Keywords Character Recognition, Deblurring, Image Alignment, Subspace Method

1. はじめに

現在、様々なアプリケーションにおいて OCR 技術が利用されてきている。書類のデータ化、画像内の単語の検索や翻訳などその用途も様々である。しかし、現在のアプリケーションは低解像、低品質な文字画像では正確な認識が困難である。加えて、動画像を用いる場合は手振れの影響により文字はさらに劣化し、認識の精度は著しく低下する。このため、動画像を用いたリアルタイムな文字認識アプリケーションは未だに普及していない。文字認識技術の汎用性を高めるためには劣化した文字への対処が必要不可欠である。

現在までに低解像な文字画像に対する様々なアプローチが行われている。石井らは推定した点広がり関数を利用して学習データを生成することでカメラの焦点ぼけに対応した学習を行っている[1]。これにより撮影された文字画像の劣化を再現した学習データを効率よく作成することはできるが、手振れによるぼけにま

で対応することは困難である。一方で低解像な画像を高解像化する超解像処理に関する研究も行われており、文字画像に超解像処理を施すことによる認識率の向上も報告されている[2]。超解像処理は低解像な画像中の文字認識のために有益な処理ではあるが、複雑な処理を行うため、多くの処理時間が必要となる。また、画像の位置ずれやぼけの影響を緩和する認識手法として相互部分空間法が提案されている[3]。この手法は複数の画像を入力として用いるため、動画像に適した認識手法だと言えるが、基となる部分空間法に比べて多くの計算時間が必要となってしまう。

本稿で提案する超低解像度文字認識システムは学習と前処理、認識の3段階で構成される。各段階において低解像、低品質な文字画像に対処することにより認識率の向上を図る。

2. 提案手法

学習段階においては、ガウシアンフィルタによる生

成型学習法を用いる。ガウシアンフィルタの標準偏差を変化させ原画像に適用することにより、様々なぼけの強さを表現した学習データを生成する。

前処理段階は位置合わせ処理とぼけ除去の2つの処理で構成される。位置合わせ処理は超解像処理内で用いられる技術である。本稿で提案する位置合わせ処理はモーション推定と画素位置変換の2つのステップから成る。モーション推定では複数の画像間に生じる微小な位置ずれを位相限定相関法[4]により推定する。画素位置変換処理において、モーション推定の結果を基に低解像な各画像の画素を高解像な画素空間に埋め込んでいくことにより高解像な画像が生成される。続いて、生成した高解像な画像に対しウィナーフィルタを用いてぼけ除去を行う。ウィナーフィルタで除去するぼけのモデルとしてガウス関数を用いることとした。

認識段階では部分空間法の拡張である相互部分空間法による認識を行う。部分空間法では学習データから各クラスの部分空間を生成し、入力ベクトルがどのクラスの部分空間で最もよく近似できるかにより識別を行う。対して相互部分空間法は、複数の入力から部分空間を生成し、学習データとの部分空間同士の類似度を比較することにより識別を行う。部分空間は式(1)の行列 R に対し上位任意数の固有値に対応する固有ベクトルにより構成される空間として計算される。

$$R = \frac{1}{N} \sum_{i=1}^N (x_i x_i^t) \quad (1)$$

ここでの $x_i (i = 1, 2, \dots, N)$ はラスタスキャンによりベクトル化された画像、 N は画像の枚数である。相互部分空間法における類似度は d 次元の部分空間同士の成す正準角 $\theta_j (j = 1, 2, \dots, d)$ から式(3)のように求められる。

$$S_{MSM} = \frac{1}{d} \sum_{j=1}^d \cos^2 \theta_j \quad (2)$$

また、相互部分空間法の処理速度を向上させる方法として、類似候補を抽出した後に認識を行う手法を提案する。まず入力と辞書の部分空間の第1固有ベクトル同士の成す角 θ から $\cos^2 \theta$ を類似度として計算する。続いて閾値以上の類似度を持つクラスのみを類似候補として抽出し、相互部分空間法による認識を行う。

3. 実験

入力用のテストデータは50~100cmまで10cm刻みの距離より撮影した動画から切り取った文字画像(英数字62文字、各文字50枚)である。機材にはソニータブレット(SGPT1211)を使用した。テストデータの一例を図1に示す。学習データは標準偏差1~10のガウシアンフィルタを各文字の原画像(20種類のフォント)に適用したものであり、1文字につき220枚を学習する。

提案する文字認識システムの有効性を示すために、学習、前処理、認識の各段階での手法を変更して実験を行った。結果を表1に示す。前処理と認識手法に関しては、括弧内に1回の処理にかかる時間も示している。提案手法が各組み合わせの中でもっともよい結果であり、類似候補の抽出によって認識率を維持しつつ処理速度を向上できていることが確認できた。



(a).50cm(16*16) (b).70cm(12*12) (c).100cm(9*9)

図1. テストデータの一例

学習	前処理 (5.9ms)	認識手法		
		SM (0.5ms)	MSM (9.5ms)	候補抽出 +MSM (4.6ms)
原画像のみ	無し	41.8%	45.0%	45.1%
生成型学習	無し	55.9%	60.5%	60.4%
	有り	56.8%	63.1%	63.5%

表1. 各段階、各手法の組み合わせ結果。SMは部分空間法(Subspace Method)、MSMは相互部分空間法(Mutual Subspace Method)であり、提案手法の結果を太字で示している。

4. 終わりに

本稿ではガウシアンフィルタを利用した学習データの生成、位置合わせ処理とウィナーフィルタの併用による前処理、類似候補抽出と相互部分空間法による認識の3段階で構成される超低解像度文字認識システムを提案した。各段階において低解像、低品質な文字画像に対処し、認識実験により各段階での手法の組み合わせの中でもっとも高い認識率が得られることと、相互部分空間法に対する処理速度の向上を確認できた。さらなる処理速度の向上が今後の課題である。

謝辞 本研究の成果の一部は、科学技術振興機構CRESTの補助による。

文献

- [1] 石井皓之, 柳詰進介, 目加田慶人, 井出一郎, 村瀬洋, “部分空間法による低解像度文字認識のための生成型学習法,” 電子情報通信学会技術研究報告, WIT, 福祉情報工学, Vol.104(92), pp.37-42, May 2004.
- [2] A. Ohkura, T. Takahashi, I. Ide, H. Murase, “Low-Resolution Character Recognition by Video-Based Super-Resolution,” ICDAR’09, pp.191-195, July 2009.
- [3] 前田賢一, 渡辺貞一, “局所的構造を導入したパターンマッチング法,” 電子情報通信学会論文誌, J68-D(3), pp. 345-352, May 1985.
- [4] 小林孝次, 中島寛, 青木孝文, 川又政征, 樋口龍雄, “位相限定相関法の原理とその応用,” テレビジョン学会技術報告, Vol.20(41), pp.1-6, July 1996.