

超解像処理を用いた低解像度文字の認識 Recognition of Low-resolution Characters Using Super Resolution

大倉 直[†] 高橋 友和[‡] 井手 一郎[†] 村瀬 洋[†]
Ataru Ohkura Tomokazu Takahashi Ichiro Ide Hiroshi Murase

1. はじめに

近年、デジタルカメラや携帯電話などの携帯デジタル撮影機器が低価格で広く普及しており、撮影した文字を認識する機能も備わり始めている。しかし、このような機器で文書を一度に撮影し、認識する場合、各文字が低解像度となってしまいうため、高精度な認識をすることは難しい。

そこで本講演では、超解像処理を用いた低解像度文字の認識手法を提案する。まず、デジタルカメラで撮影した動画像に対して超解像処理をすることで、高解像度文字を再構成する。次に、再構成した文字を入力として文字認識をする。

2. 超解像処理を用いた低解像度文字の認識手法

本研究では、まず、動画像中の連続したフレームと画像を用いて超解像処理を行い、これにより再構成された高解像度画像を認識に用いることで、認識率の向上を図る。

2.1 超解像処理

一般に、カメラで撮影された画像は、撮影時の振動によるぶれやレンズの光学的ぼけ、さらに低解像度化やノイズなどの様々な影響を受ける。これらの影響は、一般にPSF (Point Spread Function) によりモデル化される。

超解像処理では、まず、微小な位置のずれを持つ複数枚の低解像度画像から1枚の高解像度画像を再構成する[1]。目的とする解像度に合わせて、サブピクセル単位で位置合わせした各低解像度画像の画素値の集合からなる画像をレジストレーション画像と呼ぶ。

次に、このレジストレーション画像を用いての式(1)の評価関数 I が小さくなるように目的画像の更新を繰り返すことにより、PSFを取り除いた高解像度画像を求める。

$$I = \sum_{y=0}^{H-1} \sum_{x=0}^{W-1} [\mathbf{b}(x, y)^T \mathbf{h} - f(x, y)]^2 \quad (1)$$

ここで、 \mathbf{h} は更新していく目的画像のベクトル表現を、 $\mathbf{b}(x, y)$ は位置 (x, y) に対応するPSFカーネルのベクトル表現を、 $f(x, y)$ は位置 (x, y) のレジストレーション画像の画素値を、それぞれ表している。また、 H 、 W はレジストレーション画像及び目的画像の縦、横のサイズを表す。目的画像 \mathbf{h} を変数として、評価関数 I を Fletcher-Reeves の共役勾配法[2]を用いて最小化する。

実際に超解像処理により再構成された高解像度文字画像の例を図1に示す。



図1. 超解像処理の例

2.2 部分空間法による認識

文字認識において一般的に用いられる部分空間法によって学習及び認識を行う。

あるカテゴリ c の超解像処理された画像 N 枚を学習に用いるとする。 n 番目の画像から各画素値を要素とするベクトルを作成する。このとき、各ベクトルの要素の平均が0で、ノルムが1になるように正規化したベクトルを $\mathbf{x}_n^{(c)}$ とおく。 N 個のベクトル $\mathbf{x}_n^{(c)}$ を列ベクトルとして並べた行列の自己相関行列 $\mathbf{X}^{(c)}$ を以下の式で求める。

$$\mathbf{X}^{(c)} = [\mathbf{x}_1^{(c)} \cdots \mathbf{x}_N^{(c)}][\mathbf{x}_1^{(c)} \cdots \mathbf{x}_N^{(c)}]^T \quad (2)$$

この行列を固有値展開し、固有値の大きい方から R ($< N$) 個に対応する固有ベクトル $\mathbf{e}_r^{(c)}$ を認識に用いる。実際に作成された固有ベクトルを画像化した例を図2に示す。

超解像処理した入力画像 \mathbf{y} に対し、学習時に求めた固有ベクトル $\mathbf{e}_r^{(c)}$ と \mathbf{y} の類似度 $L^{(c)}(\mathbf{y})$ を以下の式により定義する。

$$L^{(c)}(\mathbf{y}) = \sum_{r=1}^R (\mathbf{e}_r^{(c)} \cdot \mathbf{y})^2 \quad (3)$$

\mathbf{y} に対する類似度 $L^{(c)}(\mathbf{y})$ が最大となるカテゴリを認識結果とする。



図2. 作成された固有ベクトル画像の例

3. 評価実験

3.1 実験条件

各実験で学習及び認識に使用する画像は、表1に仕様を示すデジタルカメラを用い、手ぶれ補正機能を解除して接写モードで撮影した。対象文字のフォントは Century とし、文字種は数字とアルファベットの大文字と小文字 (計 62 カテゴリ) である。文字画像は、屋内の壁面に文字を印刷した紙 (図3) を貼り、約 40 cm 離れた位置からカメラを手で持って撮影した。超解像処理におけるPSFとしてガウシアンフィルタ ($\sigma = 0.3$) を想定し、縦横 4

[†] 名古屋大学 大学院情報科学研究科 Graduate School of Information Science, Nagoya University

[‡] 岐阜聖徳学園大学 経済情報学部 Faculty of Economics and Information, Gifu Shotoku Gakuen University

倍の解像度の画像を再構成するものとした。部分空間法のパラメータは、認識に用いる固有ベクトル数を 5、正規化画像サイズを 32×32 pixels とした。

表 1. デジタルカメラの仕様

撮影機器	Canon PowerShot G9
解像度	360×240 pixels
フレームレート	30 fps

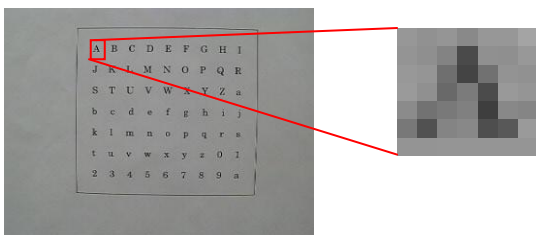


図 3. 実際に撮影された画像 (320×240 pixels)

3.2 撮影文字画像サイズと認識性能の関係

本実験では、撮影された文字画像のサイズが認識性能にどのような影響を及ぼすかを調査するために、30 枚の低解像度画像から超解像処理により作成した超解像文字を入力とした場合と低解像度文字をそのまま入力する場合の認識率を、撮影された文字画像のサイズを変化させながら比較した。1 枚の低解像度画像と 30 枚から作成される超解像画像では情報量に大きな違いがある。単一画像からでは文字の識別に必要な情報量が得られないとき、複数の画像の類似度の平均が最大となるカテゴリに分類することで、認識精度が向上することが報告されている [3]。これらを考慮して、低解像度画像を認識する場合には、部分空間法の類似度を 1 枚の画像から求める手法に加え、30 枚の画像から求める手法も比較に用いた。学習では、各カテゴリに対して 100 枚 (全 6,200 枚) の低解像度画像を用いた。なお、低解像度文字を認識する場合には低解像度文字を学習した識別器を、超解像文字には超解像文字を学習した識別器を用いた。

撮影文字の画像サイズを 6×6 , 7×7 , 8×8 , 9×9 pixels とした (図 4)。各サイズ、各カテゴリについて試行回数 50 回ずつ (全 3,100 回) 行った場合の認識率を図 5 に示す。図 5 より、提案手法は低解像度の文字をそのまま認識する手法よりも認識率が高いことを確認した。

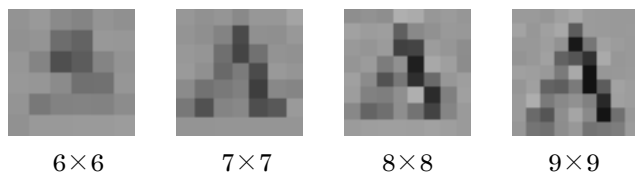


図 4. 各サイズの撮影文字画像例 (pixels)

3.3 超解像処理の使用画像数と認識性能の関係

本実験では、超解像処理に用いる低解像度画像の枚数が認識性能にどのような影響を及ぼすかを調査するために、低解像度画像の枚数を 1, 2, 4, 8, 16, 32 枚と変えた場合の認識率を比較した。学習は先の実験と同じ条件で行った。入力には 8×8 pixels の低解像度文字を縦横 4

倍に超解像処理した超解像文字を用いた。撮影された文字画像の例を図 6 に示す。超解像処理に用いた低解像度画像の枚数それぞれに対し、各カテゴリについて試行回数 50 回ずつ (全 3,100 回) 行った場合の認識率を図 7 に示す。図 7 より、超解像処理に用いる低解像度画像数が多いほど認識率が向上することを確認した。

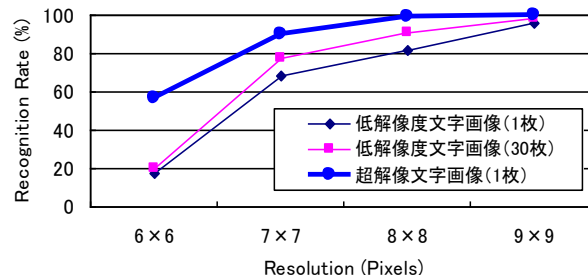


図 5. 撮影文字画像サイズと認識率の関係



図 6. 用いる低解像度画像枚数に対する超解像処理結果例

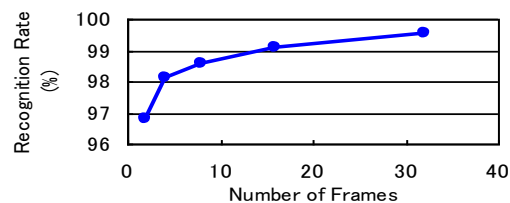


図 7. 用いる低解像度画像枚数と認識率の関係

4. おわりに

動画像に対して超解像処理を用いることで、低解像度の文字を高精度に認識する手法を提案した。また、実際に撮影した動画像を用いて実験を行い、提案手法の有効性を確認した。今後の課題として、文字以外を対象とした認識の検討、超解像処理により適した認識手法の検討などが挙げられる。

謝辞

日頃より熱心に御討論頂く名古屋大学村瀬研究室諸氏に感謝する。本研究の一部は日本学術振興会科学研究費補助金による。本研究では、画像処理に MIST ライブラリを使用した。

参考文献

- [1] 田中正行, 奥富正敏, “再構成型超解像処理の高速化アルゴリズムとその精度評価,” 信学論 (D-II), vol. J88-D-II, no. 11, pp. 2200–2209, Nov. 2005.
- [2] R. Fletcher, C. Reeves, “Function minimization by conjugate gradients,” Computer Journal, vol. 7, pp. 149–154, British Computer Society, 1964.
- [3] S. Yanadume, Y. Mekada, I. Ide, H. Murase, “Recognition of very low-resolution characters from motion images,” Proc. PCM2004, Lecture Notes in Computer Science, vol. 3331, pp. 247–254, Springer Verlag, Dec. 2004.