

超解像手法を利用した画像認識の改善

Improvement of Image Recognition Utilizing Super-Resolution

村上 隆昭† 竹内 泰斗† 後藤 富朗† 平野 智† 桜井 優†
Takaaki Murakami Taito Takeuchi Tomio Goto Satoshi Hirano Masaru Sakurai

1. はじめに

画像認識において、特徴量は画像の中から顔や人などの特定の物体を認識するために重要な要素の 1 つである。特徴量は学習手法と密接な関係があり、新たな学習手法の提案に伴い、学習手法を有効に活用する特徴量が提案されている。そこで本稿では、あらかじめ設定した特定の物体を認識する手法を提案する。しかし現在この技術は、周囲の状況変化で認識精度が低下してしまうという問題点がある。また低解像度カメラで撮影した画像は認識率が低く、近づかないと認識できない。これらの解決策として、認識対象の画像を明るさや角度など様々な状況を想定することに加えて、超解像手法を利用して低解像度画像から高解像度画像を生成することで、認識精度の向上を図った。

2. 超解像

提案手法のブロック図を図 1 に示す。提案手法ではまず、入力画像を Total Variation 正則化によって骨格成分とテクスチャ成分に分離する。その後、骨格成分は Shock Filter、テクスチャ成分は Pulse Enhancement Filter によって強調し、両者を合成することで鮮明な拡大画像を得る。

2.1. Total Variation 正則化

Total Variation 正則化は入力画像をエッジおよび滑らかな成分で構成される骨格成分と細かな振動成分およびノイズで構成されるテクスチャ成分の分離に利用される。式 (1) で示される評価関数 $F(u)$ の最小化問題を解くことで画像分離が実現できる。

$$F(u) = \sum_{i,j} |\nabla u_{i,j}| + \lambda \sum_{i,j} |f_{i,j} - u_{i,j}|^2 \quad (1)$$

ここで f は入力画像、 u は出力画像、 λ は正の定数である。この最小化問題を求める高速解法として式 (2) に示す Chambolle の射影法 [1] を用いる。

$$P_{i,j}^{(n+1)} = \frac{P_{i,j}^{(n)} + \left(\frac{\tau}{\lambda}\right) \nabla (f + \lambda \operatorname{div} P_{i,j}^{(n)})}{\max \left\{ 1, \left| P_{i,j}^{(n)} + \left(\frac{\tau}{\lambda}\right) \nabla (f + \lambda \operatorname{div} P_{i,j}^{(n)}) \right| \right\}} \quad (2)$$

$$v = \lambda \operatorname{div} P \quad u = f - v \quad (3)$$

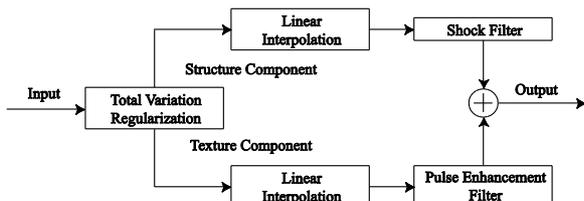


Figure 1. 提案手法のブロック図

式 (3) のように原画像 f と骨格画像 u の差分からテクスチャ画像 v を導出できる。図 2 に Total Variation 正則化分離の例を示す。

2.2. Shock Filter

Shock Filter は Osher と Rudin によって提案された非線形エッジ先鋭化フィルタ [2] であり、式 (4) の繰り返し演算によって与えられる。

$$u_{i,j}^{(n+1)} = u_{i,j}^{(n)} - \operatorname{sign} \left(\Delta \left(K_{\sigma} * \Delta u_{i,j}^{(n)} \right) \right) \left| \nabla u_{i,j}^{(n)} \right| dt \quad (4)$$

ここで u は骨格成分であり、簡単な演算処理で急峻なエッジを再構成することが可能である。また、エッジ強調処理時に発生するいくつかのアーティファクトを正常に制御することができる。図 3 に入力画像と Shock Filter によって処理された出力画像の例を示す。

2.3. Pulse Enhancement Filter

図 2 (c) に示されるように、テクスチャ成分はエッジを除く高周波成分のみで構成され、パルスの集合として扱うことができる。このテクスチャ成分の強調処理の基本的な考え方としては、テクスチャ成分の各パルスを鋭くすることである。その処理は、各パルスの検出、 N 乗の計算、パルスの高さの正規化の順に行う。式 (5) に入力信号 x と出力信号 y の関係を示す。

$$y = A|x|^N \operatorname{sign}(x) \quad (5)$$

ここで、 A は元のパルスと比較してパルスの高さを強調するように設定された正規化係数であり、 N は自然画像で 1.5 ~ 3 に設定される強調係数である。図 4 にテクスチャ強調の一例を示す。



Figure 2. Total variation 正則化による画像分離



Figure 3. Shock Filter によるエッジ強調

† 名古屋工業大学大学院 工学研究科 情報工学専攻

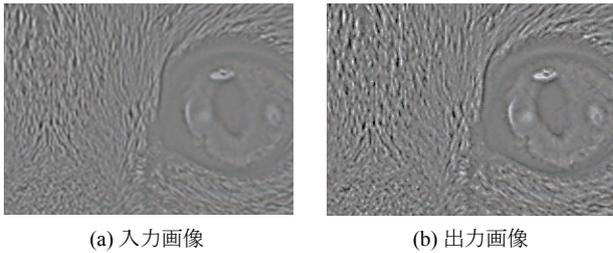


Figure 4. Pulse Enhancement Filter によるテクスチャ強調

3. 画像認識

本稿で用いる画像認識の流れを図 5 に示す。まず、学習フェーズで学習画像を特徴量データに変換し、変換データを機械学習させる。そして、認識フェーズで入力画像を特徴量データに変換し、変換されたデータを機械学習により、抽出されたルールに基づいて判別を行う。

3.1. Haar-like 特徴量

Haar-like 特徴量 [3] は、照明条件の変動やノイズの影響を受けやすい輝度値をそのまま用いるのではなく、近接した 2 つの矩形領域の輝度差により得られる特徴量である。図 6 に示すように、白領域 $r1$ と黒領域 $r2$ の輝度差を式 (6) により算出し、Haar-like 特徴量とする。

$$H(r1, r2) = S(r1) - S(r2) \quad (6)$$

ここで $S(r)$ は、領域 r の輝度和を算出する関数である。図 6 に示すパターンをもとに、位置やスケールを網羅的に変化させることにより膨大な特徴量を生成する。また、Integral image を用いることにより、各領域の合計輝度を領域の大きさに関わらず、4 点の値の加減算で算出できるため高速に特徴量を算出することが可能である。

3.2. AdaBoost

AdaBoost [4] は、二値 (対象クラス/非対象クラス) を出力する弱識別器を複数組み合わせる高精度な強識別器を学習するアルゴリズムである。AdaBoost では、学習サンプルに対して重みを付け、多数の弱識別器候補から誤識別するサンプルの重みの和を、最小とする弱識別器を選択している。その後、選択された弱識別器を用いて学習サンプルを識別し、間違えたサンプルに対して高い重みを与えることにより、識別が困難なサンプルに対しても正しく識別する弱識別器が選択される。

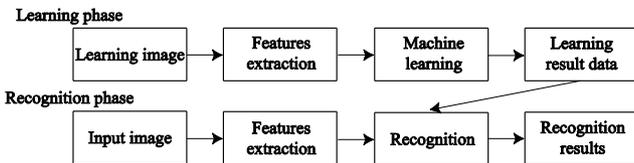


Figure 5. 画像認識の流れ

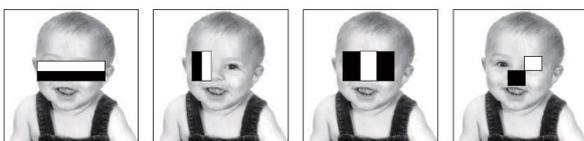


Figure 6. Haar-like 特徴量

4. 実験結果

提案手法の有効性を確認するために認識距離を比較する実験を行った。この実験では、対象物体の大きさを 10×4 cm とし、原画像 (640×480 pixel) と超解像によって 4×4 倍に拡大した画像を比較する。表 1 に実験条件を示し、表 2 および図 7 に実験結果を示す。実験結果より、原画像での認識距離は 90 cm となったのに対し、超解像手法では 140 cm の距離まで正確に認識可能であることが確認でき、超解像手法によって拡大された画像の認識距離は、原画像と比較して約 1.6 倍に増加した。また、図 7 より原画像では文字が認識できないのに対し、超解像後の画像では文字が認識できることも確認した。

5. むすび

本論文では、あらかじめ設定した特定の物体を認識する際の認識精度を向上させるために、画像認識システムに超解像手法を適用する手法を提案した。実験結果より、超解像によって画像が鮮明になること、より遠くの物体を認識可能となることを確認した。

参考文献

- [1] A. Chambolle, "An algorithm for total variation minimization and applications," J. Mathematical Imaging and Vision, vol. 20, no. 1, pp. 89-97, 2004.
- [2] S. J. Osher and L. I. Rudin, "Feature-oriented image enhancement using shock filters," SIAM Journal on Numerical Analysis, vol. 27, pp. 910-940, August 1990.
- [3] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol. 1, pp. 511-518, 2001.
- [4] Y. Freund and R. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," Journal of Computer and System Sciences, vol. 55, no. 1, pp. 119-139, 1997.

TABLE I. 実験条件

Total Variation 正則化	
反復回数	30
λ	0.02
Shock Filter	
ステップ幅	1.2
Pulse Enhancement Filter	
N	2.0

TABLE II. 認識距離の比較

	認識距離
原画像	90 cm
超解像画像	140 cm

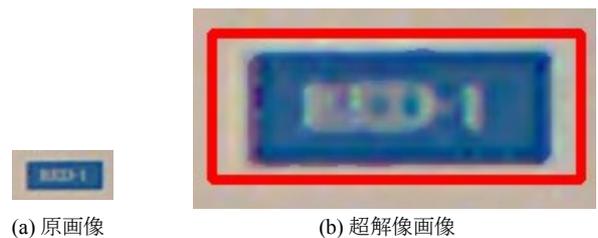


Figure 7. 実験結果