

デジタル画像処理の基本的なトピックと手法の紹介

鈴木 将之*

Introduction of digital image processing topics and methods

Masayuki Suzuki*

近年では容易に高解像度のデジタル画像を取得することができ、またストレージ容量の増加により大量のデジタル画像を保存することが可能となった。取得したデジタル画像をビジネスに活用すべく、デジタル画像処理の需要が高まっている。本稿では画像処理における基本的なトピックと、それに対する数理的なデジタル画像処理手法を紹介する。

Keywords: 画像処理、デジタル画像処理、OpenCV

1. 概要

デジタル画像は画素で構成される画像であり、一般に横が N 、縦が M 個の、格子状に $N \times M$ 個の画素が並んだデータである。画素には画素値が設定され、一般にグレースケール画像であれば 8 ビット、カラー画像であれば 8×3 ビットの RGB(赤、緑と青)色空間で表現される色が設定される。図 1 にデジタル RGB 画像のイメージを示す。

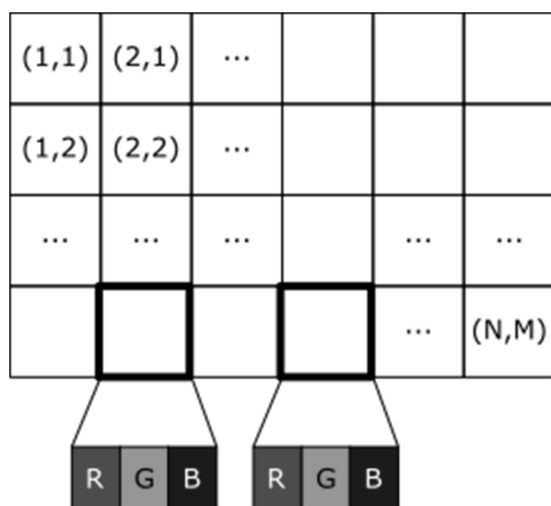


図 1 デジタル RGB 画像

安価なウェブカメラや携帯端末付属のカメラの普及により、いつでも容易にデジタル画像を取得することが可能となった。さらに、ストレージ

容量の増加やクラウド上のストレージサービスの登場により、高解像度の、大量のデジタル画像を保存することが可能となった。また、OpenCV [1]や `pix2pix` [2]といった優れたデジタル画像処理ライブラリがオープンソースとして公開されるようになり、無料または安価かつ高信頼の実装の利用を行うことができる。以上のことより、より高精度の画像処理や、従来画像処理を利用しなかった分野で自動化や追加情報の取得のためにデジタル画像処理を導入することが、今日では可能となった。

本稿では画像処理の基本的な個別トピックと、それに対する数理的な手法の紹介を行う。本稿では 2 章にエッジなどの特徴抽出について、3 章に画像認識について、4 章にロバスト性を高めるための前処理について、0 章にまとめを述べる。

デジタル画像処理において深層学習が利用されることも多いが、本稿では深層学習は簡単に触れるに留める。本稿で挙げるものは汎用な画像処理におけるトピックであり、深層学習を利用する上でも考慮しなければならない事項である。

2. 特徴抽出

画像処理では画像に映る物を定量的に評価するために、エッジや領域といった特徴を抽出することがある。本稿では代表的な特徴抽出としてエッジ抽出、線分抽出と領域分割を紹介する。

*アドバンスソフト株式会社 第 1 事業部

1st Computational Science and Engineering
Group, AdvanceSoft Corporation

2.1. エッジ抽出

最も単純な特徴抽出として、画像中の輪郭部分を 1、それ以外を 0 として抽出するエッジ抽出が挙げられる。代表的なエッジ抽出手法としてケニーのエッジ検出 [3]を挙げる。ケニーのエッジ検出ではまずソーベルフィルタ [4]などから周辺画素値の変化が大きい画素を検出する。画素値の変化は連続的であるため、ソーベルフィルタの出力の画素値は実数かつ太さがあるものとなる。太さが 1 画素の 2 値エッジ画像として抽出するために細線化を行う必要がある。ケニーのエッジ検出では画素値変化が最大の箇所のみをエッジとし、またエッジを途切れにくくすべく閾値を満たす画素もエッジとして追加する。図 2 にケニーのエッジ検出適用例を示す。図 2 上図に示す画像に対し、中図はケニーのエッジ検出適用結果を下図は中図の一部拡大図を示す。図 2 下図より、上図の輪郭をエッジとして抽出することができることと、1 画素のエッジとして抽出することができることが分かる。

2.2. 線分抽出

エッジ検出に対し線分抽出は線分のみを検出し、その始点と終点を得ることである。線分抽出ではより機械的に利用しやすい始点と終点座標を得ることができることと、画像内でより特徴的といえる直線部分のみを抽出することができる。線分抽出手法として、Fast Line Detector (FSD) [5]を挙げる。FSD はケニーのエッジ検出により得られたエッジを元に、しきい値以下の曲率を持つエッジに対し直線フィッティングを行う。また線分同士の結合を行う。線分結合時に結合方向へ小領域を作成し、小領域内のヒストグラムの比較を行うことで大きく異なる画素値から得られたエッジ、すなわち異なる対象同士のエッジの結合を防ぐことを期待することができる。図 3 に FSD 適用例を示す。図 3 上図に示す画像に対し、中図に FSD により得られた線分を赤線で示した図を、下図に中図の一部拡大図を示す。図 3 下図において、白い石柱部分とレンガ部分同士で線分が生成されていることが分かる。単純な近傍同士のエッ

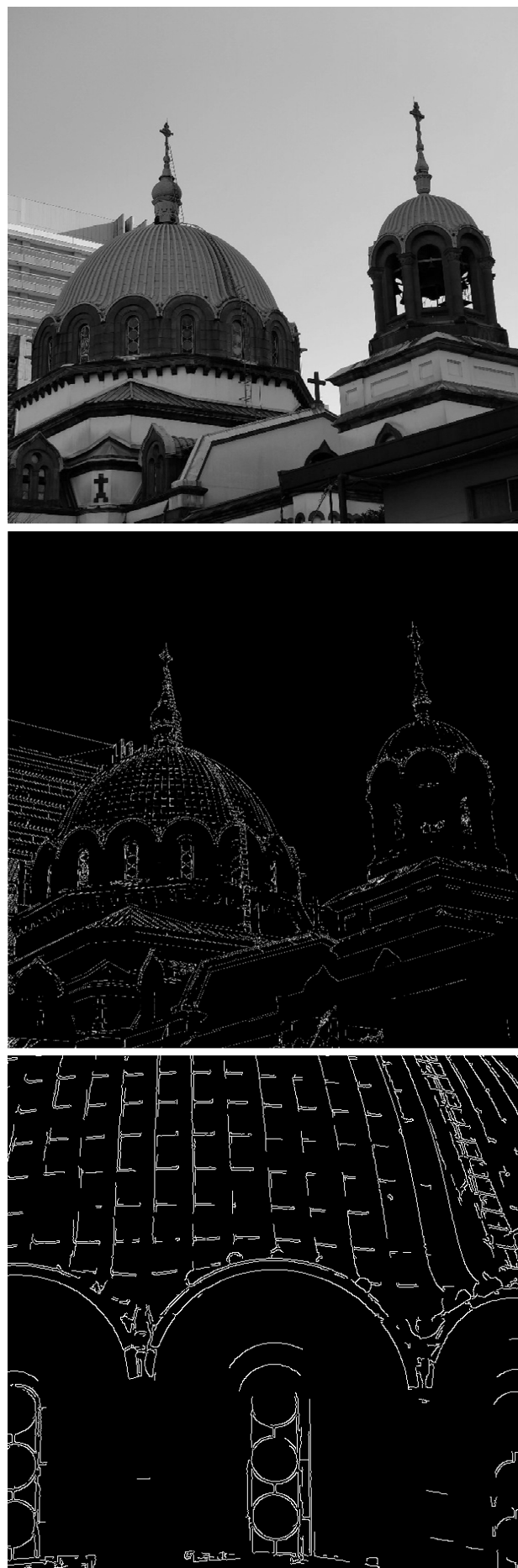


図 2 ケニーのエッジ検出適用例



図 3 FSD 適用例

ジ結合ではなく、FSD が画素値の分布を考慮してエッジを結合していることが分かる。

2.3. 領域分割

画像より対象物のみを取り出したい場合、領域分割を行い、前景と背景や対象と対象以外のものを切り分ける。

領域分割手法としてミーンシフト法[6]を挙げる。ミーンシフト法はクラスタリングの一手法であり、ミーンシフトと呼ばれる更新を行い確率密度関数の極大点を求める手法である。画像処理においては、周囲の似たような画素値を持つ画素全ての画素値を同一のものに変更することで、平滑化と領域分割を行うものである。図 4 にミーンシフト法適用例を示す。図 4 上図に示した画像に対し、下図が適用結果を示す。図 4 において、彫刻の凹凸、木のつなぎ目や汚れによる画素値変化が軽減していることが分かる。

3. 画像認識

2 章で述べたような特徴の利用と学習を行い、画像中の物体検出や画像分類など画像認識が行われる。

図 5 に Haar 特徴量とカスケード分類器 [7]による顔認識例 [1]を挙げる。Haar 特徴量は図 6 に示すような矩形領域の白部分の画素値総和から灰部分の画素値総和を引いたものである。図 5 においては複数パターンの Haar 特徴量に重みを付けた総和（強識別器）を直列接続し、全ての強識別器の判定に成功した場合のみを顔と判定している。

現在では画像認識には深層学習がよく使用される。これは深層学習を使用した AlexNet が大規模画像認識のコンペティションにて従来手法を大幅に上回る精度を得たことによる [4][8]。深層学習を用いると ResNet [9]のように人間の能力と同等以上の性能を得ることを期待することができる [4]。ただし、深層学習では学習のために大量のデータを必要とする。また、従来手法では知見に基づき特徴抽出と表現を行うことができるのに対し、深層学習では内部処理がブラックボックスになりやすい。

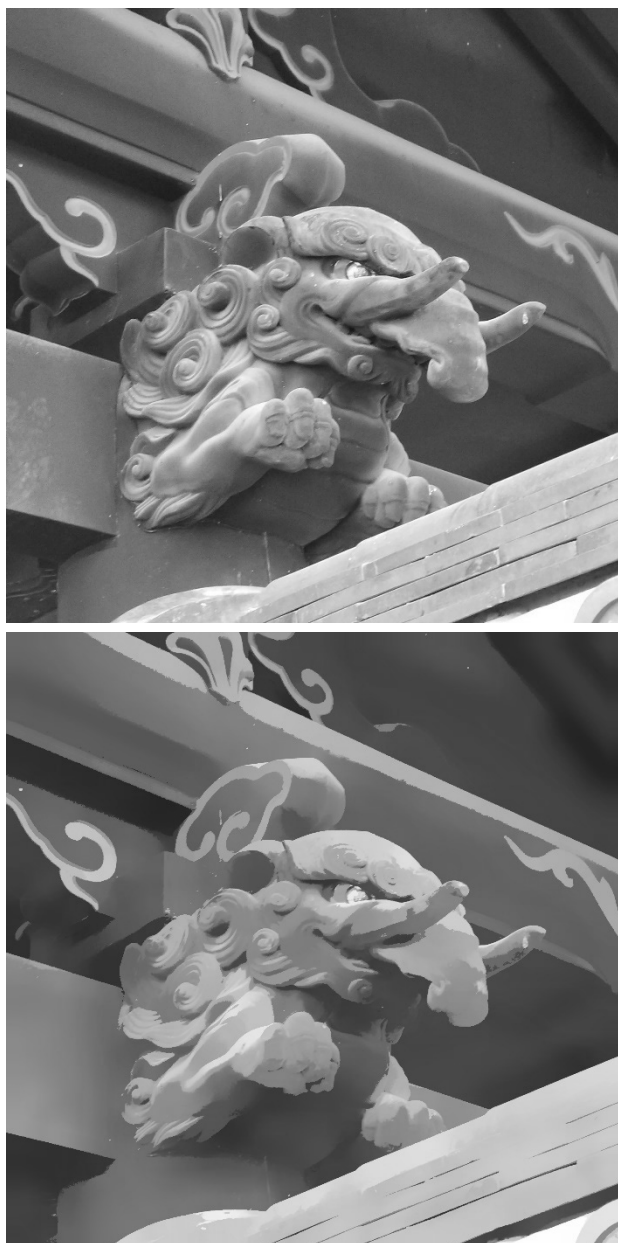
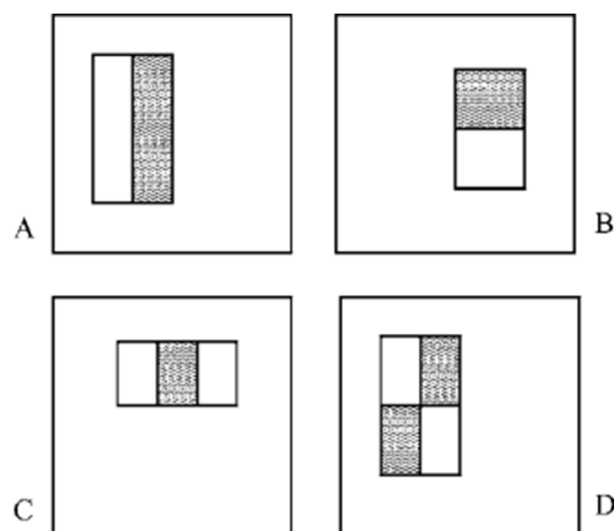


図 4 ミーンシフト法適用例

図 5 Haar 特徴量とカスケード分類器による顔認識 [1]¹図 6 Haar 特徴量 [7]²

4. 前処理

2 章と 3 章では画像処理の中でも画像をビジネスに利用するための、画像解析と呼ばれるトピックを紹介した。しかし、実画像ではノイズ、明るさの変化やコントラストの不均一といった、画像解析の障害となりうる要素が含まれている。図 7 にノイズが含まれる画像の例を示す。図 7 は高 ISO 感度にて暗い箇所で撮影を行ったものである。図 7 において高 ISO 感度によりノイズが増幅された結果、本来白一色の紙面部分の色が均一ではなくなっている。本稿では代表的な前処理である、平滑化とコントラスト調整を紹介する。

¹ OpenCV: Cascade Classifier. OpenCV documentation index [online]. [viewed 11 January 2022]. Available from: <https://docs.opencv.org/3.4/db/d28/tutorial_cascade_classifier.html>

² Figure 1. VIOLA, Paul; JONES, Michael J. Robust real-time face detection. International journal of computer vision, 2004, 57.2: 137-154.



図 7 ノイズを含む画像

4.1. 平滑化

平滑化は画素変化を滑らかにする処理である。画素変化を滑らかにすることにより、ノイズの軽減と、不要な画素値変化による画像解析への影響を軽減することができる。

代表的な平滑化手法として 2.3 章で示したミーシフト法と、ガウシアンフィルタに画素値変化の重みを加えたバイラテラルフィルタ [10] を挙げる。ガウシアンフィルタは注目画素とその周辺画素にガウス分布に従うように重みを付けて足し合わせるフィルタである。ガウシアンフィルタも平滑化を目的として使用されるがエッジ部分も平滑化され画像がぼやけてしまう。バイラテラルフィルタではガウシアンフィルタに画素値変化の重みを追加し、画素値変化が大きな箇所ではガウシアンフィルタの効果を弱めることにより、エッジを保持したまま平滑化することを期待することができる。図 8 に図 7 に対しバイラテラルフィルタを適用した例を示す。図 8 において、平滑化がなされノイズが軽減されていること、文字が判別可能なことからエッジを保持することができたことが分かる。

4.2. コントラスト変換

明るさやコントラストを事前に調整することは画像処理において重要である。明るさが足りず

不明瞭な画像では特徴を抽出することは難しい。また影のようにコントラストが大きい箇所をエッジや線分として抽出することがある。

コントラストを調整する方法として画素値ヒストグラム平坦化がある。例えば影を考えると、影の箇所と影でない箇所に画素値が分布していることとなり、画素値ヒストグラム上では 2 つの山として現れる。画素値ヒストグラム平坦化は頻度が高い画素値を持つ画素を変更することで、暗い箇所を明るく、逆に明るい箇所を暗くすることが可能となり、コントラストを抑えることができる。代表的な画素値ヒストグラム平坦化手法として Contrast Limited Adaptive Histogram Equalization (CLAHE) [11] を挙げる。CLAHE は画像をタイルに分け、タイルごとにヒストグラム平坦化を行う手法である。タイルごとにヒストグラム平坦化を行うことで、画像内で極端にコントラストが大きい箇所がある画像に対し、コントラストが似たような箇所ごとにコントラストを変換する。そのため画像全体のヒストグラム平坦化に比べ、白飛びと黒飛びの防止を期待することができる。図 9 に CLAHE 適用例を示す。図 9 上図に示す画像に対し、下図が適用結果である。図 9 上図では判別が難しい軒下部分が、下図においては判別が可能である。



図 8 バイラテラルフィルタ適用例



図 9 CLAHE 適用例

5. 最後に

本稿では画像処理における基本的なトピックと、それに対するデジタル画像処理手法を示した。

本稿で取り上げなかったこととして、精度とリアルタイム性がある。リアルタイム性は放送やライブ配信におけるフレーム間隔や、定時観測における時間間隔内で一処理を終えることである。精度とリアルタイム性はトレードオフである。ミーンスフト法など精度が高いもののリアルタイム性を得ることが難しい手法もある。精度や厳密性を低減して処理を高速化し、リアルタイム性を得るといったことも考えられる。ビジネスの要件を定義した上で手法を選択する必要がある。

本稿で挙げた手法は適切なパラメータを設定する必要がある。例えば 2.3 章に挙げたミーンスフト法では空間の分散と画素値の分散の 2 つのパラメータを設定する。空間の分散は画像大きさや対象の大きさ、画素値の分散は対象の画素値の分布に応じて適切に値を設定する必要がある。しかし、例えば動画で対象が移動し画像内での大きさが変化する、時刻や照明の変化により画素値の分布が変化するなどのことが考えられる。数理的な手法では対象や外環境の変化に対し事前決定したパラメータが不適切であることが多々あり、ロバスト性が不十分となることがある。機械学習及び深層学習では大量の画像から学習することで、より有力な外環境の変化に対する前処理や特徴の抽出手法を構築することができ、ロバストな画像処理を期待することができる。

アドバンスソフト株式会社は本稿で挙げたような数理的な手法及び深層学習による、デジタル画像処理の受託開発の実績を有する。もし画像の機械的な利用を検討されているならば、ぜひ弊社にご相談いただきたい。

参考文献

- [1] Home - OpenCV. OpenCV [online]. [viewed 11 January 2022]. Available from: <https://opencv.org/>
- [2] Image-to-Image Translation with Conditional Adversarial Networks. phillipi.github.io [online]. [viewed 11 January 2022]. Available from: <https://phillipi.github.io/pix2pix/>
- [3] CANNY, John. A computational approach to edge detection. IEEE Transactions on pattern analysis and machine intelligence, 1986, 6: 679-698.
- [4] デジタル画像処理編集委員会. デジタル画像処理[改訂第二版]. 公益財団法人画像情報教育振興協会, 2020. ISBN 9784903474649.
- [5] LEE, Jin Han, et al. Outdoor place recognition in urban environments using

- straight lines. In: 2014 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2014. p. 5550-5557.
- [6] COMANICIU, Dorin; MEER, Peter. Mean shift: A robust approach toward feature space analysis. IEEE Transactions on pattern analysis and machine intelligence, 2002, 24.5: 603-619.
- [7] VIOLA, Paul; JONES, Michael J. Robust real-time face detection. International journal of computer vision, 2004, 57.2: 137-154.
- [8] KRIZHEVSKY, Alex; SUTSKEVER, Ilya; HINTON, Geoffrey E. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 2012, 25: 1097-1105.
- [9] HE, Kaiming, et al. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2016. p. 770-778.
- [10] TOMASI, Carlo; MANDUCHI, Roberto. Bilateral filtering for gray and color images. In: Sixth international conference on computer vision (IEEE Cat. No. 98CH36271). IEEE, 1998. p. 839-846.
- [11] ZUIDERVELD, Karel. Contrast limited adaptive histogram equalization. Graphics gems, 1994, 474-485.

※ 技術情報誌アドバンスシミュレーションは、アドバンスソフト株式会社 ホームページのシミュレーション図書館から、PDF ファイル(カラー版) がダウンロードできます。(ダウンロードしていただくには、アドバンス/シミュレーションフォーラム会員登録が必要です。)