

画像処理と画像認識

— AI時代の画像処理入門 —

工学博士 山田 宏尚

工学博士 末松 良一

共 著

コロナ社

□□□ ま え が き □□□

前著『画像処理工学』（メカトロニクス教科書シリーズ）は、画像処理を初歩から学習しようとするさまざまな分野の方を対象として、2000年に出版されました。そして2014年には、その後のデジタル化の流れに対応した改訂を行いました。幸いなことに、前著は多くの大学・高専で教科書として採用され、講義等で使用されてきました。

周知のように、この20年で特に深層学習（ディープラーニング）をはじめとするAI（人工知能）関連の技術が大きく発展し、スマートフォンや自動車の安全運転支援をはじめ、多くの分野で画像処理・認識技術が活用されるようになりました。そして、この分野への期待は以前にも増して高まっているといえます。このような背景を踏まえて、本書ではおもに画像認識に関する内容を大幅に拡充し、新たに『画像処理と画像認識—AI時代の画像処理入門—』として出版することになりました。

画像処理で扱われる領域はたいへん広く、信号処理技術、画像処理手法、画像計測やグラフィックスなどさまざまな分野に及びます。それに加えて画像認識の領域も機械学習や深層学習を中心とした広範な分野があり、学ぶべき項目も膨大になります。このため、既存の教科書では、おかれている重点分野もさまざまに異なっています。そこで本書では、前著と同様に幅広いトピックを入門的に扱いながら、これらの分野をバランス良く配分させ、基礎事項を体系的に学べるように構成しました。

本書を画像処理の教科書として用いる場合は、8章以降は発展的内容あるいは自習用の参考という扱いにされてもよいと思います。また、画像認識やコンピュータビジョンの教科書として用いる場合は、2～7章の基礎的内容を必要に応じてスキップし、8～14章を中心に講義で使うことができるものと思います。

近年、インターネットよりOpenCVなどのフリーのライブラリが簡単に手に入り、初学者でも画像処理や認識のプログラムを実行できるようになりました。また、深層学習を含む最先端の研究ソースも大学や研究機関などにより公開されており、誰もが容易に利用できるようになっています。それらを導入する際に必要な理論的な基礎知識を、本書により得ることができるものと思います。一方で、各項目の本質をわかりやすく説明するために、詳細なアルゴリズムや数学的手法の多くは、あえて省略せざるを得ませんでした。そこで、本書によって基礎的知識を習得後、読者の皆さんがさらに高度な知識を得たい場合には、巻末の引用・参考文献やインターネット上の文献等を参照いただければ幸いです。

ii ま え が き

本書の執筆にあたっては、多くの文献を参考にさせていただきました。これら著者の方々には心から感謝申し上げます。

最後に、本書を出版するにあたってたいへんお世話になったコロナ社の関係各位に深くお礼を申し上げます。

2022年8月

著 者

注) 本文中に記載している会社名、製品名は、それぞれ各社の商標または登録商標です。本書では®やTMは省略しています。

□□□ 目 次 □□□

1. 序 論

1.1	人間の視覚機能	1
1.2	画像と画像処理	2
1.3	ビジョンシステム	3
1.4	AIと画像処理	4
	演習問題	5

2. 画像の表現

2.1	アナログ画像とデジタル画像	6
2.2	画像のA-D変換	7
2.2.1	標本化	7
2.2.2	量子化	8
2.3	A-D変換と画質との関係	10
2.4	カラー画像の表現	12
2.4.1	人間の視覚とカラー画像	12
2.4.2	色の表現法	14
2.4.3	カラー画像のデジタル化	16
2.5	画像データの表現	17
2.5.1	画像データの表現方式	17
2.5.2	インデックス方式による画像表現	18
2.5.3	画像のファイル形式	19
	演習問題	20

3. 画像処理システム

3.1	画像処理システムの構成例	21
3.1.1	コンピュータを用いた画像処理システム	21
3.1.2	画像処理プログラムにおける高速化のための留意点	23
3.1.3	専用ハードウェアによる画像処理装置	24

3.2 画像の入出力装置	27
3.2.1 画像の入力装置	27
3.2.2 画像の出力装置	28
演習問題	30

4. 画像情報処理

4.1 画像のフーリエ変換	31
4.1.1 空間周波数	32
4.1.2 複素正弦波	33
4.1.3 フーリエ変換	34
4.1.4 離散フーリエ変換	35
4.1.5 高速フーリエ変換	38
4.1.6 直交変換	38
コラム 1: 直交変換と直交基底	40
コラム 2: 直交変換と DFT	41
4.1.7 その他の直交変換	42
4.2 標本化定理	43
4.3 フィルタ処理	47
4.3.1 周波数領域でのフィルタ処理	47
コラム 3: FIR フィルタと IIR フィルタ	48
4.3.2 空間領域でのフィルタ処理	50
コラム 4: 畳み込みの意味と線形時不変フィルタ	52
4.3.3 フィルタの設計	52
4.3.4 オペレータによるフィルタ処理	55
コラム 5: 局所処理と大局処理	55
4.4 画像データの圧縮	56
4.4.1 可逆符号化	56
4.4.2 不可逆符号化	59
4.4.3 動画の符号化	61
4.4.4 標準化された符号化技術	62
コラム 6: JPEG 非可逆圧縮の特徴	63
演習問題	66

5. 濃淡画像処理

5.1 濃度変換	67
5.1.1 コントラスト変換関数を用いた濃度変換	68

コラム 7: ガンマ補正による濃度変換	68
5.1.2 ヒストグラム変換	69
5.1.3 隣接する画素との濃度差の強調	70
5.2 平滑化	70
5.2.1 線形フィルタによる平滑化	71
5.2.2 エッジを保存した平滑化	72
5.3 鮮鋭化	73
コラム 8: 画像超解像	76
5.4 エッジ・線の検出	76
5.4.1 差分型によるエッジ検出	76
5.4.2 零交差法によるエッジ検出	79
5.4.3 レンジフィルタ	80
5.4.4 Canny エッジ検出器	80
5.5 画像表示のための処理	80
5.5.1 限定色表示	80
5.5.2 疑似濃淡表示	81
5.6 幾何学的変換	83
5.6.1 アフィン変換	84
5.6.2 疑似アフィン変換	84
5.6.3 画像の再配列と補間	85
5.6.4 圧縮された画像データからの拡大・縮小	86
演習問題	87

6. 2 値画像処理

6.1 2 値化処理	88
6.2 連結性と幾何学的性質	90
6.2.1 連結と近傍	90
6.2.2 幾何学的性質	91
6.3 2 値画像に対する処理	93
6.3.1 ラベリング	93
6.3.2 膨張・収縮処理	94
6.3.3 線・点図形化処理	95
6.4 図形の形状特徴	98
演習問題	101

7. コンピュータグラフィックス

7.1	2次元グラフィックス	102
7.1.1	線図形の描画	102
7.1.2	面の生成と処理	103
7.1.3	アンチエイリアシング	104
7.2	3次元グラフィックス	105
7.2.1	3次元図形の幾何学的変換	105
7.2.2	モデリング	107
7.2.3	レンダリング	109
	コラム9: CGと画像処理技術の融合	114
	演習問題	115

8. 領域分割

8.1	原画像中のエッジを用いる方法	117
8.2	領域拡張法	117
8.3	特徴空間におけるクラスタリングを用いた方法	117
8.4	テクスチャ解析	118
8.5	グラフカット法	121
8.6	深層学習を用いた領域分割	121
	演習問題	121

9. 特徴・パターンの検出

9.1	テンプレートマッチング	122
9.2	局所特徴	124
9.2.1	Harrisのコーナの検出	124
9.2.2	FASTによるコーナ検出	125
9.2.3	プロブの検出	126
9.2.4	SIFT	127
9.2.5	HOG	128
	コラム10: さまざまな記述子	129
9.2.6	bag-of-visual words	129
	演習問題	130

10. 画像認識

10.1 画像認識のタスクとプロセス	131
10.2 ルールベースに基づく画像認識	132
10.3 統計的パターン認識	133
10.3.1 特徴空間	134
10.3.2 パターンの識別	134
コラム 11: 訓練データの前処理	138
10.3.3 クラスタリング	139
コラム 12: 遺伝的アルゴリズム	141
演習問題	141

11. ニューラルネットワークと深層学習

11.1 ニューラルネットワーク	142
11.2 深層学習	144
11.2.1 活性化関数の改良	145
11.2.2 ドロップアウト	146
11.2.3 訓練データと学習手順	146
コラム 13: 訓練データの作成	147
11.2.4 転移学習とファインチューニング	147
11.3 畳み込みニューラルネットワーク	148
11.3.1 畳み込み層	148
11.3.2 プーリング層	148
11.3.3 出力層	149
11.3.4 CNN の応用	149
コラム 14: CNN のネットワークモデル	151
11.4 オートエンコーダ	152
11.5 敵対的生成ネットワーク	152
演習問題	153

12. 3次元画像処理

12.1 3次元空間の計測と認識	154
12.2 1枚の画像を用いた3次元認識	155
12.2.1 テクスチャによる方法	155
12.2.2 陰影による方法	155

12.2.3	その他の方法	156
12.3	複数の画像を用いた3次元認識	156
12.3.1	カメラモデルとキャリブレーション	156
12.3.2	ステレオ画像処理	158
12.4	レンジファインダ	159
	コラム 15: 3次元データ処理を行うソフトウェア	160
12.5	3次元形状の復元	161
	演習問題	161

13. 動画画像処理

13.1	オプティカルフローの抽出	162
13.1.1	照合法	163
13.1.2	時空間勾配法	163
13.2	差分画像を利用する方法	164
13.3	動画画像を用いた行動認識	166
13.3.1	従来の機械学習を用いた方法	166
13.3.2	深層学習を用いた方法	166
13.3.3	RNN と LSTM	166
13.4	動画画像を用いた3次元復元	167
	演習問題	168

14. 画像処理の応用

14.1	産業応用	169
14.1.1	組立工程	169
14.1.2	検査工程	171
	コラム 16: 従来の機械学習と深層学習	172
14.2	医療用画像処理	173
14.2.1	CT	174
14.2.2	細胞診の自動化	175
14.2.3	深層学習による画像診断支援	176
	コラム 17: 説明可能な AI とネットワークの可視化	176
14.3	文字認識	177
14.3.1	印刷文字認識	177
14.3.2	手書き文字認識	178
14.3.3	深層学習を用いた文字認識	178

14.4 顔 の 認 証	179
コラム 18：オープンソースの活用	180
演 習 問 題	181
引用・参考文献	182
演習問題の解答	183
索 引	187

コンピュータビジョンは、一般的な情景の理解・解析や3次元情報の復元など、人間の視覚と同等の機能をコンピュータ上に実現することを目的とする分野である。ロボットビジョンにおいては、おもに実用的見地から、処理速度・処理精度・価格等が重要であり、高速処理が可能な2値画像を扱うことも多い。また、背景や照明、あるいは画像入力方法などの工夫も重要となる。一方、コンピュータビジョンは、3次元空間における対象やシーンを認識させようとする**物体認識** (object recognition) や**画像理解** (image understanding) などの**人工知能** (artificial intelligence : AI) の分野 (次節参照) と関連の深い研究分野であり、方法論を主体とするものである。

1.4 AI と画像処理

AI はコンピュータにより推論、認識、判断などの人間と同様の知能を人工的に実現させようという技術である。今のところ人間と同等のレベルで認識・理解し、自ら判断ができるようなAI (これを**強い AI** という) は実現できていない。現実には、あらかじめプログラムされた範囲内で処理を行う限定的な人工知能 (これを**弱い AI** という) の実用化が進んでいる。ただし、弱いAIであってもある特定の範囲 (画像処理、言語処理、将棋や囲碁など) では人間のレベルを超えたAIを実現できているものもある。

AIの初期の研究では人間の思考過程を記号で表現し実行する**推論** (inference) やその処理として**探索** (search) により問題を解決する研究が行われていた。それにより**トイプロブレム** (Toy problem) と呼ばれるチェスやオセロなどの比較的簡単な問題解決や、コンピュータによる特定の定理の証明などが行われていた。その後、専門家の知識やノウハウなどの**知識ベース** (knowledge base) を人間がルールとして記述し、そのルールに従ってコンピュータに処理させようという**ルールベース** (rule base) による研究 (例えば専門家のような知識を持ち、推論や判断ができる**エキスパートシステム**など) も活発に行われた。しかし、人間の膨大な知識をどう定式化して表現するのか、最適な答えを得るために矛盾のないルールをどのように決めるかなどの問題があり、この方法では汎用的なAIの実装が難しい。

これに対して、現在では人間が行っている学習能力と同様の機能をコンピュータで実現しようとする**機械学習** (machine learning) を用いた画像認識が主流となっている。その背景には、コンピュータ性能の向上およびインターネット上から学習に必要な大量の情報 (ビッグデータ) が比較的容易に手に入るようになったことがある。人間はいろいろな経験をして、そこから学習したことを一般化できる。これと同じように、機械学習ではコンピュータに多くのデータを入力して学習させた後、未知の例について正確に判断させることを目標とする。機械学習では、確率や統計などの手法を使って問題を解決する手法が一般的だが、その中でも**深層学習** (deep learning, **ディープラーニング**) が重要な手法となっている。深層学習は、ニューラルネットワークの一種で、その詳細については11章で述べる。AIの手法に関する分類方法

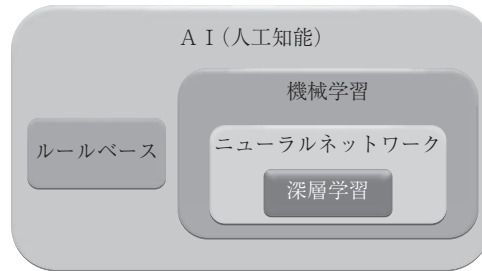


図 1.3 AI の分類

はさまざまなものがあるが、上記で説明した手法をまとめると図 1.3 のようになる。

また、機械学習には**教師あり学習** (supervised learning) と**教師なし学習** (unsupervised learning) がある。教師あり学習は、入力データとその正解を与えた**訓練データ** (training data) (**学習データ**もしくは**教師信号**ともいう) に基づいて学習させるもので、おもに**回帰** (regression) と**分類** (classification) の問題を対象とする。ここで、回帰は学習データに基づいて未知のデータ (数値) を予測するものである。これに対して分類は、あるデータがどのクラス (カテゴリ) に属するかを予測するもの (例えば画像の分類など) である。一方、教師なし学習は学習データに正解を与えない状態で学習をさせる手法である。なお、教師あり学習と教師なし学習の中間的手法として、学習データに正解を与える代わりに、目的として設定された報酬 (スコア) を最大化するための行動を学習する**強化学習** (reinforcement learning) という手法もある。

演 習 問 題

- 【1】 人間が持つ視覚と画像処理の能力について、その特徴をまとめよ。
- 【2】 画像処理が応用されている身近な例をあげよ。
- 【3】 ロボットビジョンとコンピュータビジョンの違いについて述べよ。
- 【4】 AI (人工知能) について、とくに機械学習、ニューラルネットワーク、深層学習 (ディープラーニング) の関係性を中心にまとめよ。

数（サブルーチン）を用いる必要がある。一般に、このような関数は、画像入力インターフェースに付属するソフトウェアや OpenCV（コラム 18 参照）などの画像処理ソフトウェアと同梱されていることが多い。

3.1.3 専用ハードウェアによる画像処理装置

画像処理のための専用ハードウェアは、リアルタイム処理あるいは、それに準ずる高速での画像処理を行うことを目的として開発されている。例えば、デジタルカメラにおいて画像の補正処理やノイズ除去などの処理（5 章参照）を行う画像信号処理（image signal processing：ISP）用のプロセッサや、車載機器等で用いられる人物などの認識処理（10 章参照）、ビデオレコーダなどで動画像の符号化・復号（4.4.3 項参照）などを行うための専用ハードウェアが開発されている。また、工場における組立・検査工程などに用いられる画像処理システムでは、高速な処理が必要であるため、画像処理専用プロセッサおよびメモリ、画像入出力部を一体化した画像処理ハードウェアや、それをコントロールするためのコンピュータを組み合わせたシステムなどがある。このようなシステムでは、画像処理に要する必要最小限の機器構成および高速性・信頼性が要求される。

さて、画像処理に特化した専用ハードウェアの場合、用途に応じて高速化のためのさまざまな工夫がなされている。汎用コンピュータのための CPU は（高速化のために一度に数個の命令を同時に実行できる機能もあるが）、通常は逐次処理であり基本的には一度に一つの演算しかできない。このため、画素を大量に持つ画像データに対して高速な処理を行うことが難しい。そこで、画像処理に特化した高速演算プロセッサ等が提案されている。処理速度の面では、すべての画素に対して並列に処理モジュールを並べて画像処理を行う**完全並列型プロセッサ**が理想的であるが、複雑な処理には向いておらず、並列処理のためのプログラミングが難しいため実用面で問題がある。このため、一般的には画像の局所的なデータをプロセッサに読み込んで並列的に処理を行う手法が使われる。プロセッサの処理では、できるだけ並列的に処理を進めた方が一度に多くの処理ができる。この並列処理のレベルには、実行単位の小さい方から順に（a）データレベル、（b）命令レベル、（c）スレッドレベルがあり、それぞれにおける並列処理について以下で詳しく述べる。

（a）データレベルの並列処理

画像処理の中で 4 章や 5 章で述べるようなさまざまな処理の多くでは、大量のデータに対し並列に同一の演算を行って処理を進めることが可能である。これをデータレベルの並列性が高いという。この場合、複数の画素値をまとめてベクトルデータ（1 次元配列）で表し、それを一度に計算することが可能である。このような計算処理が可能なプロセッサを**ベクトルプロセッサ**（アレイプロセッサともいう）という。ベクトルの演算では、一つの命令で同時に複数のデータに対して演算処理を適用できるため、**SIMD**（single instruction multiple data）とも呼ばれる。

(b) 命令レベルの並列処理

1枚の画像の連結領域に対するラベリング処理(6.3.1項参照)などでは、隣接する領域のデータ処理を並列に同時実行できないためデータレベルの並列処理は難しい。そこで、命令レベルの並列処理を利用して処理速度の向上をはかる必要がある。命令レベルで並列な処理を実現するには、**スーパースカラ**を用いる方法と、**VLIW**(very long instruction word)による方法がある。**表3.1**にスーパースカラとVLIWの特徴の比較をまとめる。

表3.1 スーパースカラとVLIWの特徴の比較

	命令実行方法	手 法	長 所	短 所
スーパースカラ	命令処理を実行する回路を複数用意し、複数の命令処理を同時に実行	プロセッサがプログラムの並列性を判別して同時処理を行う	実行プログラムに変更を加えなくても高速で実行できる	プロセッサのハードウェア構造が複雑になる
VLIW	複数の命令を一つの命令としてまとめて投入し複数の実行ユニットで並列に実行	プログラム言語のソースコードを機械語に変換する際に並列に実行できる処理をまとめ、一つの命令としてプロセッサに与える	プロセッサ構造をシミュレーションにでき製造コストを下げられる	本来の性能を十分に発揮させるためのプログラミングが難しい

なお、命令レベルの並列処理性能を向上させるために**命令パイプライン処理**が採用されることも多い。これは、一つの命令を複数の段階に分割してそれぞれをパイプライン状に接続した別々の回路モジュールで実行するものである。これにより、複数の処理をベルトコンベア式にオーバーラップさせて実行できる。

(c) スレッドレベルの並列処理

一つの画像処理アプリケーションソフトウェアは一つのプロセスを実行するが、その中には**スレッド**と呼ばれるプロセスよりも小さな複数の処理単位が実行される。このスレッドを複数同時並行に実行することを**マルチスレッド**という。一方、プロセッサの内部で独立して機能する演算・制御装置のことを**コア**(もしくは**プロセッサコア**)といい、二つ以上のコアを有するプロセッサ製品を**マルチコアプロセッサ**という。マルチコアプロセッサではそれぞれのコアでスレッドを並列に実行できる(一つのコアで複数のスレッドを実行できるプロセッサもある)。多くのコアを持つプロセッサを使えば同時並行で多くのスレッドを実行できるので、アプリケーションの処理速度が速くなる。

上記の(a)~(c)はプロセッサに組み込む際に用途に応じてカスタマイズされる。ここでは、画像処理の用途に応じて、(1)カメラ用ISP、(2)ビデオやテレビにおける動画像処理用プロセッサ、(3)コンピュータグラフィックス用プロセッサ(GPU)、(4)画像認識用プロセッサを例にあげ、専用ハードウェアの構成概要について述べる。

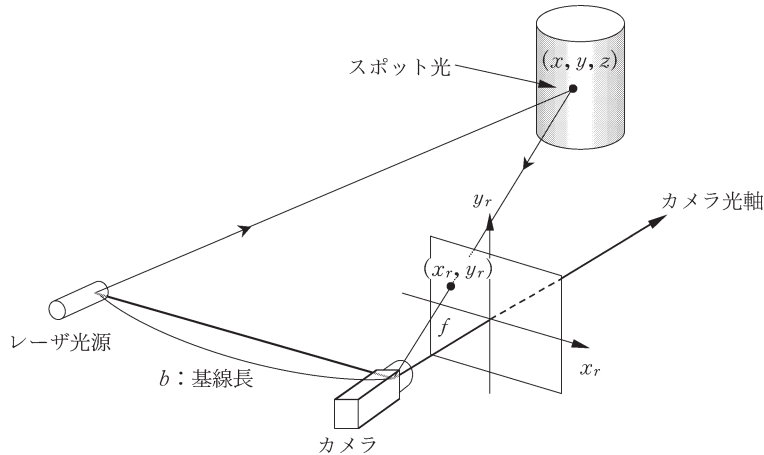


図 12.7 レンジファインダによる距離計測の原理

帰ってくるまでの時間により対象物までの距離を計測する **ToF** (time of flight) (光レーダ法ともいう) がある。

ToF には、光パルスを利用する方法と、光の位相差を利用する方法がある。ToF を利用した代表的なレンジファインダに **LiDAR** (light detection and ranging) がある。LiDAR は、レーザレーダとも呼ばれ、電波を使うレーダよりも検出精度が高い。また、レンジファインダにより得られた 3 次元の位置情報と同時に RGB のカラー画像を取得できる **RGB-D カメラ** の普及も進んでいる。

これらの 3 次元計測によって求めた距離データは、図 12.8 のように物体までの距離を濃淡画像で表示する **距離画像** (もしくは **深度画像**) や **点群** (7.2.2 項 [7]), **メッシュ** (7.2.2 項 [2]) などで表現できる。

コラム 15 : 3次元データ処理を行うソフトウェア

3次元データ処理を行う際に、得られた点群の処理が容易にできる点群処理ライブラリや可視化ツールなどがオープンソース(コラム 18 参照)で公開されている。点群の基本的な処理には、ノイズ処理フィルタリング、法線推定、近傍点探索などがある。代表的な点群処理ライブラリには Point Cloud Library (**PCL**) や **Open3D** が、可視化ツールでは、**MeshLab** や **Cloud Compare** などがある。PCL はリアルタイム処理が可能で、多くの OS に対応している。Open3D は PCL と比較して容易に導入ができるなどの特長がある。Cloud Compare は、多機能な点群ビューワで、フィルタリング処理、複数点群の位置合わせなど、さまざまな機能が実装されている。

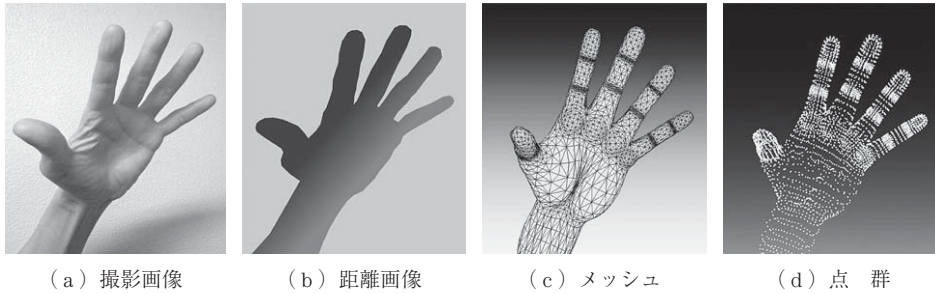


図 12.8 距離データの画像表現

12.5 3次元形状の復元

画像センサからの情報を用いて物体の3次元の形状を復元する多くの手法が提案されている。初期に提案された手法に、撮影された積み木のシーンから個々の積み木の位置関係等を理解させる**積み木の世界** (blocks world) と呼ばれる研究がある。これは積み木のような多面体の各辺の位置関係に基づく手法であるが、ノイズに弱く一般的なシーンへの適用が難しいなどの問題がある。そこで、さまざまな手がかりから3次元の形状を復元する **Shape from X** を利用する手法が提案されている。ここで、Xには、テクスチャ、陰影、輪郭、自動焦点、動きなどがある。例えば、12.2.1項で述べた面に描かれたテクスチャの傾き情報を用いて面の向きを求める **Shape from Texture** や、12.2.2項で述べた陰影などの面の明るさの情報を利用する **Shape from Shading**、複数の視点で撮影した画像中の物体のシルエットから3次元形状を復元する **Shape from Silhouette**、カメラの自動焦点機能を利用した **Shape from Defocus** (もしくは **Depth from Defocus** 法ともいう) などがある。また、異なる視点の複数画像を入力として対象物の3次元形状を復元する **Shape from Motion** もある。これは **Structure from Motion (SfM)** とも呼ばれる (13.4節参照)。

なお、3次元形状を用いた物体認識の方法として、計測された3次元データから3次元特徴量を求め、あらかじめ用意した立体形状パターンとのマッチングを行う方法がある。これは2次元画像におけるSIFTのようなキーポイントと同様の考え方であり、3次元のキーポイントを求め、想定される3次元モデルのパターンとの対応付けを行う方法である (14.1.1項 [2])。

演習問題

- 【1】 カメラキャリブレーションの目的と方法について説明せよ。
- 【2】 3次元計測の手法を、三角測量の原理を用いるものと用いないものに分類せよ。
- 【3】 ステレオ画像処理においてエピポーラ線はどのような役割を持つのか説明せよ。
- 【4】 Shape from X の代表的手法を説明せよ。

【あ】

アンシャープマスキング	75
アクティブステレオ法	159
アダプスト	139, 179
アダマール変換	42
アナログ量	6
アノテーション	147
アフィン変換	84
アンサンブル学習	139
アンダーシュート	73
アンチエリアシング	104

【い】

位相スペクトル	35
一次視覚野	2
一般化ハフ変換	98
一般物体認識	131
遺伝的アルゴリズム	141
移動平均フィルタ	71
移動ベクトル	62, 162
イメージスキャナ	27
イメージベーストモデリング	114
イメージベーストライティング	114
イメージベーストレンダリング	114
印刷文字認識	177
インスタンスセグメンテーション	121
インスタンス認識	131
隠線消去	110
インタレース方式	28
インデックス方式	18
インパルス応答	50
隠面消去	110

【う】

ウェーブレット変換	61
動き補償フレーム間予測符号化	62

【え】

液晶ディスプレイ	28
エキスパートシステム	4

エッジ	76
エビポーラ線	159
エポック	147
エリアシング	10, 44, 104
エリアシング誤差	45
円形度	99
エントロピー符号化	59
エンボス効果	78

【お】

オイラー数	91
大津の2値化	89
オートエンコーダ	152
オーバーシュート	73
オブティカルフロー	162
オープニング	94
オフライン文字認識	178
オープンソースソフトウェア	180
オペレータ	55
重みフィルタ	148
オリエンテーション	127
折り返し誤差	45
オンライン学習	146
オンライン文字認識	178

【か】

回帰	5
階層型ネットワーク	142
階層的クラスタリング	140
解像度	10
階調	8
階調値	8
外部パラメータ	157
ガウシアンフィルタ	71
顔検出	179
顔照合	180
過学習	137
可逆符号化	56
拡散反射	111
学習データ	5, 139
学習モデル	151
拡張	94
拡張現実	167
隠れ層	143
陰	111
影付け	111

可視化	102
加重平均フィルタ	71
カスケード型識別器	179
画素	7
画像	2
画像位置合わせ	128
画像照合	131
画像処理	1, 2
画像処理プロセッサ	21
画像診断支援	176
画像超解像	76
画像認識	131
画像ピラミッド	123
画像理解	4
活性化関数	143
カテゴリ	133
カーネル	50, 55, 148
カーネルトリック	138
カーネル法	138
可変しきい値法	90
可変長符号	57
加法混色	13
ガボールフィルタ	120
カメラキャリブレーション	158
カメラ座標系	157
カラーテーブル	18
カラーマップ方式	18
環境マッピング	113
完全拡散反射面	155
完全並列型プロセッサ	24
ガンマ値	68
ガンマ特性	68
ガンマ補正	68

【き】

機械学習	4
幾何学的変換	83
疑似アフィン変換	84
疑似エッジ	12
疑似カラー	70
記述子	127
疑似輪郭	12
基線長	158
基底	39
輝度信号	17
輝度値	8

相互相関係数 123
 速度ベクトル 162
 阻止域 47
 組織的ディザ法 83
 ソフトマージン SVM 137
 ソフトマックス関数 149
 ソリッドテキストチャ 113
 ソリッドモデル 108

【た】

大局処理 55
 対数量子化法 12
 ダイナミックレンジ 9
 多重解像度表現 123
 多重連結成分 91
 畳み込み 50
 畳み込み層 148
 畳み込みニューラルネット
 ワーク 148
 多値画像 9
 探索 4
 探索問題 141
 単色光 12
 単連結成分 91

【ち】

チェインコード 100
 チェイン符号化 100
 チェス盤距離 92
 知識ベース 4
 チップセット 22
 中央値フィルタ 72
 中間層 143
 直線位相 48
 直線の検出 97
 直流成分 32
 直交基底 40
 直交行列 39
 直交変換 38
 直交変換符号化 59

【つ】

通過域 47
 積み木の世界 161
 強い AI 4

【て】

ディザ法 82
 デジタル量 6
 ディープニューラルネット
 ワーク 144
 ディープラーニング 4
 手書き文字認識 177
 敵対的サンプル 176

敵対的生成ネットワーク 152
 テクスチャ 118
 テクスチャ解析 118
 テクスチャマッピング 113
 データ拡張 147
 データバス 22
 デプスバッファ 110
 転移学習 147
 点群 109,160
 伝播 94
 テンプレート 122
 テンプレートマッチング 122

【と】

トイプロブレム 4
 投影 106
 投影変換 106
 統計的背景差分法 165
 統計的パターン認識 133
 等高線計測法 159
 同次座標 84
 透視投影 107,157
 動的しきい値法 90
 特徴空間 117,134
 特徴選択 134
 特徴抽出 98
 特徴ベクトル 135
 特徴マップ 148
 特定物体認識 131
 ドロップアウト 146

【な】

ナイキスト周波数 44
 内部パラメータ 157

【に】

入力層 143
 ニューラルネットワーク 142
 ニューロン 142

【ぬ】

塗りつぶし 103

【ね】

ネオコグニトロン 148
 ネットワークの可視化 176

【の】

濃淡画像 9
 濃淡レベル 8
 能動的計測法 155,159
 濃度共起行列 119
 濃度値 8
 濃度パターン法 82

濃度ヒストグラム 69
 濃度変換 67
 ノード 143
 ノーフリーランチ定理 139
 ノンインタレース方式 28

【は】

バイオメトリクス 181
 バイキュービック補間法 86
 背景差分法 165
 ハイパスフィルタ 47
 ハイパーパラメータ 138,172
 ハイライト 111
 バイラテラルフィルタ 72
 バイリニア補間法 85
 バウンディングボックス 150
 バギング 139
 白色化 138
 白色光 12
 パーセプトロン 136
 パターン認識 131
 パターンの識別 134
 パターンマッチング 122
 パーチャル内視鏡 175
 パーチャルリアリティ 102
 バックプロパゲーション 142
 ハッシング 128
 バッチ 109
 バッチ学習 146
 バッチサイズ 147
 ハフ変換 97
 ハフマン符号 58
 パラメータ空間 98
 パレット 18
 パワースベクトル 35
 汎化能力 135,144
 半教師あり学習 147
 ハンドアイシステム 171
 バンドパスフィルタ 47
 バンプマッピング 113
 判別分析法 89

【ひ】

非可逆符号化 56
 光切断法 159
 光レーダ法 160
 ビジョンシステム 3
 ヒストグラム 57
 ——の均一化 69
 ヒストグラム変換 69
 非線形量子化 10,12
 非直線位相 48
 ビットマップデータ 17
 微分ヒストグラム法 89

- | | | | | | |
|--------------------|-------|----------------|--------|--------------|-----|
| 標準化 | 138 | 平行投影 | 107 | 面積 | 98 |
| 表色系 | 14 | バイズの識別規則 | 138 | | |
| 標本化 | 7 | バイズの定理 | 138 | 【も】 | |
| 標本化定理 | 43 | ベクトルデータ | 17 | モスキート雑音 | 60 |
| 頻度法 | 81 | ベクトルプロセッサ | 24 | モデラ | 105 |
| ピンピッキング | 170 | ベクトル量子化 | 61,130 | モデリング | 105 |
| | | ベジェ曲線 | 103 | モード法 | 88 |
| 【ふ】 | | ペンタプレット | 27 | モーメント特徴 | 99 |
| ファインチューニング | 147 | | | モルフォロジ演算 | 94 |
| フィルタ処理 | 47 | 【ほ】 | | | |
| フォールスカラー | 70 | ポイントベーストモデリング | | 【ゆ】 | |
| フォンノイマン型コンピュータ | | | 109 | 有彩色 | 13 |
| | 22 | 方向コード | 100 | ユークリッド距離 | 92 |
| フォンノイマンボトルネック | | 膨張 | 94 | ユニタリ行列 | 39 |
| | 22 | ボクセル | 109 | ユニット | 143 |
| フォンのスムーズシェーディング | 112 | 補色 | 13 | | |
| フォンのモデル | 111 | ホップフィールドネットワーク | | 【よ】 | |
| 復号 | 8 | | 142 | 予測符号化 | 56 |
| 複雑度 | 99 | ポリゴン | 107 | 弱い AI | 4 |
| 符号化 | 8 | ポリウムレンダリング | 114 | | |
| 符号数 | 8 | ポルツマンマシン | 142 | 【ら】 | |
| ブースティング | 139 | | | ライトフィールドカメラ | 114 |
| 物体検出 | 131 | 【ま】 | | ライン CCD | 27 |
| 物体座標系 | 157 | マシンビジョン | 3 | ラジオシティ法 | 115 |
| 物体認識 | 4,131 | マスク | 55 | ラスト型データ | 17 |
| フラクタル | 109 | マックスプーリング | 149 | ラプラシアンフィルタ | 73 |
| ブラシ処理 | 104 | マッピング | 113 | ラベリング | 93 |
| フーリエ記述子 | 99 | 窓関数 | 52 | ラン | 58 |
| フーリエ逆変換 | 32 | 窓関数法 | 53 | ランダムディザ法 | 82 |
| フーリエ変換 | 31 | マトリックス CCD | 27 | ランダムフォレスト | 139 |
| プリミティブ | 108 | マハラノビス距離 | 136 | ランレングス | 58 |
| プーリング層 | 148 | マルチコアプロセッサ | 25 | ランレングス符号化 | 58 |
| プリンタ | 29 | マルチスライス CT | 175 | | |
| フルカラー画像 | 16 | マルチスレッド | 25 | 【り】 | |
| フレーム間差分法 | 164 | マルチメディア | 102 | 離散コサイン変換 | 42 |
| フレーム間予測符号化 | 62 | マンセル表色系 | 14 | 離散フーリエ逆変換 | 36 |
| プログラマブルシェーダ | 26 | | | 離散フーリエ変換 | 35 |
| プログレッシブ走査 | 28 | 【み】 | | リージョンプロポーザル | 150 |
| プロジェクト | 29 | ミニバッチ学習 | 146 | リフラクションマッピング | 113 |
| ブロックノイズ | 60 | ミーンシフト法 | 141 | リフレクションマッピング | 113 |
| ブロックマッチング法 | 163 | | | 領域拡張法 | 117 |
| プロトタイプ法 | 136 | 【む】 | | 領域分割 | 116 |
| プロブ (コンピュータグラフィクス) | 108 | 無彩色 | 13 | 量子化 | 7,8 |
| プロブ (局所特徴) | 124 | 無相関化 | 138 | 量子化間隔 | 9 |
| 分割統合法 | 117 | | | 量子化誤差 | 11 |
| 分割法 | 117 | 【め】 | | 量子化雑音 | 11 |
| 分離型フィルタ | 54 | 明度 | 14 | 量子化レベル | 8 |
| 分類 | 5 | 命令パイプライン処理 | 25 | 輪郭線 | 91 |
| | | メインメモリ | 21 | 輪郭線追跡 | 96 |
| 【へ】 | | メタボール | 108 | 隣接 | 90 |
| 平滑化 | 70 | メッシュ | 160 | | |
| 平均値プーリング | 149 | メディアンカット量子化法 | | 【る】 | |
| | | | 12,81 | ルックアップテーブル方式 | 18 |
| | | メディアンフィルタ | 72 | ルールベース | 4 |

<p>【れ】</p> <p>レイトレーシング 113</p> <p>連結 90</p> <p>連結数 91</p> <p>連結成分 90</p> <p>レンジファインダ 159</p>	<p>レンジフィルタ 80</p> <p>レンダラ 105</p> <p>レンダリング 105,109</p> <p>【ろ】</p> <p>ロジスティック回帰 136</p> <p>露出 114</p>	<p>ローパスフィルタ 47</p> <p>ロボットビジョン 3</p> <p>【わ】</p> <p>ワイヤフレームモデル 107</p>
◆ ◆		
<p>【A】</p> <p>A-D 変換 7</p> <p>AI-OCR 178</p> <p>AKAZE 129</p> <p>AlexNet 151</p> <p>AR 167</p> <p>【B】</p> <p>bag-of-features 129</p> <p>bag-of-visual words 129</p> <p>blob 124</p> <p>BMP 20</p> <p>BoVW 129</p> <p>BRIEF 129</p> <p>【C】</p> <p>CAD 3,102,176</p> <p>Canny エッジ検出器 80</p> <p>CCD 27</p> <p>CG 3,102</p> <p>CNN 148</p> <p>Cloud Compare 160</p> <p>Convolutional Pose Machine 151</p> <p>CPU 21</p> <p>CRT 28</p> <p>CSG 表現 108</p> <p>【D】</p> <p>DCT 42</p> <p>DCT 符号化 60</p> <p>Dense Trajectories 166</p> <p>Depth from Defocus 161</p> <p>DFT 35</p> <p>DNN 144</p> <p>DoG 126</p> <p>DPCM 62</p> <p>dpi 11</p> <p>DP マッチング 178</p> <p>【E】</p> <p>End-to-End 学習 149</p> <p>【F】</p> <p>FairMOT 151</p>	<p>FAST 125</p> <p>Fast R-CNN 150</p> <p>FFT 38</p> <p>FIR フィルタ 48</p> <p>【G】</p> <p>GA 141</p> <p>GAN 152</p> <p>GF 20</p> <p>GoogLeNet 151</p> <p>GPGPU 23</p> <p>GPU 23,26</p> <p>【H】</p> <p>Haar-Like 特徴 179</p> <p>Harris のコーナ検出 124</p> <p>HDR 114</p> <p>HOG 128</p> <p>HSI カラーモデル 15</p> <p>HSV 15</p> <p>【I】</p> <p>I3D 166</p> <p>ICP 171</p> <p>IFFT 38</p> <p>IIR フィルタ 48</p> <p>I/O 22</p> <p>ISP 24</p> <p>【J】</p> <p>JPEG 20,62</p> <p>【K】</p> <p>Kd-tree 法 128</p> <p>K-L 変換 42,60</p> <p>K-means 法 140</p> <p>K-NN 法 136</p> <p>K 近傍法 136</p> <p>【L】</p> <p>LCD 28</p> <p>LiDAR 160</p> <p>LoG 79</p> <p>LSTM 167</p>	<p>Lucas-Kanade 法 164</p> <p>LUT 方式 18</p> <p>【M】</p> <p>MeshLab 160</p> <p>MOT 151</p> <p>MPEG 65</p> <p>【N】</p> <p>NN 法 136</p> <p>NTSC 17</p> <p>【O】</p> <p>OCR 177</p> <p>OpenCV 180</p> <p>OpenPose 151</p> <p>Open3D 160</p> <p>ORB 129</p> <p>OSS 180</p> <p>【P】</p> <p>pbm 20</p> <p>PCL 160</p> <p>PE 26</p> <p>pgm 20</p> <p>PNG 20</p> <p>pnm 20</p> <p>PnP 問題 167</p> <p>ppi 11</p> <p>ppm 20</p> <p>Prewitt のエッジ検出オペレータ 78</p> <p>PSPNet 151</p> <p>PyTorch 180</p> <p>P-タイル法 88</p> <p>【R】</p> <p>RAW データ 26</p> <p>ReLU 145</p> <p>ResNet 151</p> <p>RGB 13</p> <p>RGB-D カメラ 109,160</p> <p>RNN 166</p> <p>Roberts のエッジ検出オペレータ 77</p>

	【S】		【T】	YOLO	150
SegNet	151	TensorFlow	180	【Z】	
SfM	161	TIFF	20	Z ソート法	110
Shape from Defocus	161	ToF	160	Z バッファ法	110
Shape from Motion	161			【数字】	
Shape from Shading	161	【V】		2 次元直交変換	39
Shape from Silhouette	161	VGGNet	151	2 次元ユニタリ変換	39
Shape from Texture	161	Visual SLAM	167	2 値画像	9
Shape from X	161	VLIW	25	3D CNN	166
SIMD	24	VRAM	23	4 近傍	90
SLAM	167			4 近傍距離	92
Sobel のエッジ検出オペレータ	78	【X】		4 連結	90
SSD	150	XYZ 表色系	14	8 近傍	90
Structure from Motion	161	X 線 CT	174	8 近傍距離	92
SURF	129			8 連結	90
		【Y】			
		YCbCr	17		
		YIQ	17		

— 著 者 略 歴 —

山田 宏尚 (やまだ ひろなお)

1986年 名古屋大学工学部機械工学科卒業
1991年 名古屋大学大学院博士課程修了(機械工学専攻)
工学博士(名古屋大学)
1991年 名古屋大学助手
1994年 岐阜大学助教授
2007年 岐阜大学教授
現在に至る

末松 良一 (すえまつ よしかず)

1966年 名古屋大学工学部機械工学科卒業
1971年 名古屋大学大学院博士課程修了(機械工学専攻)
1974年 工学博士(名古屋大学)
1977年 名古屋大学助教授
1988年 名古屋大学教授
2005年 名古屋大学名誉教授
2005年 豊田工業高等専門学校長
2011年 愛知工業大学総合技術研究所客員教授
現在に至る

画像処理と画像認識—AI時代の画像処理入門—

Image Processing and Image Recognition — Introduction to Image Processing in AI Era —

© Hironao Yamada, Yoshikazu Suematsu 2022

2022年10月26日 初版第1刷発行

★

検印省略

著 者 山 田 宏 尚
末 松 良 一
発 行 者 株式会社 コロナ社
代 表 者 牛 来 真 也
印 刷 所 壮 光 舎 印 刷 株 式 会 社
製 本 所 株式会社 グリーン

112-0011 東京都文京区千石4-46-10

発 行 所 株式会社 コロナ社
CORONA PUBLISHING CO., LTD.

Tokyo Japan

振替00140-8-14844・電話(03)3941-3131(代)

ホームページ <https://www.coronasha.co.jp>

ISBN 978-4-339-02931-4 C3055 Printed in Japan

(森岡)



JCCOPY < 出版者著作権管理機構 委託出版物 >

本書の無断複製は著作権法上での例外を除き禁じられています。複製される場合は、そのつと事前に、出版者著作権管理機構(電話 03-5244-5088, FAX 03-5244-5089, e-mail: info@jcopy.or.jp)の許諾を得てください。

本書のコピー、スキャン、デジタル化等の無断複製・転載は著作権法上での例外を除き禁じられています。購入者以外の第三者による本書の電子データ化及び電子書籍化は、いかなる場合も認めていません。落丁・乱丁はお取替えいたします。