

# 序文

本書を執筆するきっかけになったのは、2001年にワシントン大学で Steve Seitz とともに講義した「コンピュータグラフィックスのためのコンピュータビジョン」という授業である。当時は、コンピュータビジョンの技術が映画などのコンピュータグラフィックス (CG) に用いられ始めた時期だった。実世界の物体のモデルを画像から作成したり、特殊視覚効果映像を作成したり、コンピュータシミュレーションフォトグラフィ技術を用いて実世界の画像と CG を融合したりすることに、コンピュータビジョンが使われていた。そこで、授業ではコンピュータビジョンのおもしろい応用を紹介することに焦点を当てた。パノラマ画像を合成する方法や、自分が撮影した写真から 3 次元の CG モデルを作成する方法など、学生が興味を持ちそうなトピックは多数あった。

それ以降、ワシントン大学とスタンフォード大学では、同じ内容でプロジェクト課題も含んだ、コンピュータビジョンの一般的な授業が行われている (スタンフォードでは 2003 年に David Fleet とともに講義した)。似たような内容の授業は、そのほかの数多くの大学でも取り入れられ、コンピュータシミュレーションフォトグラフィについてのより専門的な授業にも組み込まれている (本書を授業でどのように利用すればよいのかについては、1.4 節の表 1.1 (p.23) を参照)。

本書は、私の 20 年にわたるコンピュータビジョンに関する研究を反映している。その多くは Digital Equipment Corporation 社の Cambridge Research Lab と、Microsoft Research での研究である。これまでの研究において、私が常に取り組んできたのは、問題とそれを解決する技術 (アルゴリズム) である。それらは実世界において現実的な応用場面が必ずあり、しかも実際にうまく動作しなければならない。そのため、本書では実世界の環境下で動作する基本的な技術に焦点を当てて解説する。数学的にエレガントであっても、あまり実用的でないものは割愛する。

本書は、コンピュータサイエンス学科や電子工学科の大学 3 年生に向けた、コンピュータビジョンの授業に適した構成になっている。画像処理や CG の授業の受講経験や基礎知識がすでにあれば、一般的な数学的知識の吸収に時間をあまり割かず、コンピュータビジョン技術そのものを理解することに集中できるだろう。また本書は、大学院レベルのコンピュータビジョンの講義にも適しており (より実際的な応用を追及し、アルゴリズム的な要素を深く掘り下げることができる)、基礎技術や最先端の研究を参照するためのハンドブックとしても使える。そのため、各分野における最新の文献を最低一つは紹介している (ただし、その内容は本書で扱うには複雑すぎるかもしれない)。

これまで授業を行って来てわかったことは、実際にアルゴリズムを実装する課題をいくつも学生に行わせることの有効性である。時には別の課題として実装したものを利用することもある。現実世界の実際の画像を扱い、実際に存在する問題設定をクリアするようなものを作らせるのである。学生をいくつかのグループに分けた後、最終課題の内容は自分たちで選べた (これが時には国際学会の論文にもなった!)。本書の各章末の演習は、中期的なプロジェクト課題の宝庫であり、もちろん未だ多くの研究がなされている未解決問題も含んでいる。可能な限り、学生に

は自分で撮影した写真に対して自分のアルゴリズムを試すことを勧めてきた。そうすることで、学生の動機付けになり、今までになかった新しい問題を発見し、実世界の画像がどんなに多様で複雑なのかを学生自身が身をもって体験できるからである。

コンピュータビジョンの問題を定式化し解決する過程において、私は三つの高い視点からヒントを得ることが多かった。

- 科学的 (scientific) アプローチ：画像生成過程の詳細なモデルを構築し、数式を導出する。そしてそれを逆にして、測定量として得たい値を復元する（必要ならば、簡単化のための仮定を導入し、数式を単純化する）。
- 統計的 (statistical) アプローチ：未知変数の事前尤度確率と入力画像を生成するノイズを含んだ測定の過程を統計的にモデル化し、測定量として得たい値の最も良い推定値を求め、それに含まれる不確実性を解析する。ここで使用される推定アルゴリズムは最適化手法に密接に関連しており、(科学的アプローチにおける) 画像生成過程を逆転するために用いるものと同じものである。
- 工学的 (engineering) アプローチ：対象を単純に記述し、実現可能な技術を開発する。もちろん実用的でなければならない。その技術をテストして、限界と失敗例を理解し、(実行速度としての) 計算コストも見積もる。

以上の三つのアプローチは互いに密接に関連しており、本書のあちらこちらで見かけるだろう。

私自身の研究開発のポリシーは（そしてもちろん本書の演習の主眼は）アルゴリズムを実際に「試すこと」である。コンピュータビジョンでは、いくつかの画像では「うまくいきそうな」アルゴリズムを作ることは非常に簡単である。しかし、それは「正しく動作する」ということではない。自分のアルゴリズムを評価する最も良い方法は、次の3段階の手順を踏むことである。

一つ目は、正確な結果がわかっている、人工的に作成したデータでアルゴリズムを試すことである。二つ目は、そのデータにノイズを加えて、ノイズレベルの増加とともに性能がどの程度落ちるのかを評価することである。三つ目は、実世界のデータでアルゴリズムを試すことである。できれば、インターネットで収集した大量の画像のように、出所がきわめて多様であることが望ましい。そうして初めて、自分のアルゴリズムが現実世界の複雑さに耐えられるかどうかを知ることができる。つまり、単純化されたモデルや仮定に従わないような画像に対しても、うまく動作するかどうか分かることになる。

このやり方を学生が実践する手助けをするために、本書は数多くの補助資料を用意している。それらの資料は本書のウェブサイト <http://szeliski.org/Book/> から入手可能であり<sup>\*1</sup>、付録Cでその説明をする。資料には次のものが含まれる。

- それぞれの問題に対して一般的に利用されており、インターネットで入手できる評価用データセットへのリンク
- 学生が課題に取り組む際に助けとなる、画像入力・出力や画像処理などのソフトウェアライブラリへのリンク
- 本書に対応した講義スライド
- 本書で引用した文献リスト (BibTeX 形式)

<sup>\*1</sup>【訳注】2013年1月現在、まだ入手できない。著者によれば、多忙のため、残念ながら今後もアップロードされる予定はないそうである。

最後の二つの資料は、授業を行う教員や論文を執筆する研究者にとって有用であると思われるが、もちろん通常の学生にとっても役に立つものである。ソフトウェアライブラリの中には、コンピュータビジョンのアルゴリズムを幅広く実装しているものもあり、それらを利用することで（教員の指導があれば）より意欲的なプロジェクトに取り組めるだろう。

## 謝辞

研究と探求への並々ならぬ情熱を持ち、執筆を励ましてくれた、全ての人たちに感謝を申し上げたい。

マギル大学の Steve Zucker は、最初にコンピュータビジョンというものを私に紹介してくれた。そして彼の全ての学生に、研究で使われている技術とその結果について、疑問を持ち討論することを指導してくれた。さらに、この分野での学位を取ることを勧めてくれた。

金出武雄と Geoff Hinton は、カーネギーメロン大学における私の博士学位論文指導教員として、良質な研究の進め方、文章の書き方、発表の仕方の基礎を教えてくれた。彼らは、映像情報処理や3次元モデリング、統計的手法に対する私の興味を掻き立ててくれた。Larry Matthies はカルマンフィルタとステレオマッチングの手ほどきをしてくれた。

私の最初の企業研究職における指導者であった Demetri Terzopoulos は、上手な論文執筆の方法を教えてくれた。短い間ではあったが SRI International でともに働いた Yvan Leclerc と Pascal Fua は、コンピュータビジョンとはまったく異なった新しい物の見方を見せてくれた。

Digital Equipment Corporation 社の Cambridge Research Lab における6年間の研究生活では、素晴らしい仲間たち — Ingrid Carlbom, Gudrun Klinker, Keith Waters, Richard Weiss, Stéphane Lavallée, Sing Bing Kang — と働くという幸運に恵まれた。さらに、何人もの傑出した夏季インターンたち — David Tonnesen, Sing Bing Kang, James Coughlan, Harry Shum — にも恵まれた。ここはまた、ミドルベリー大学に今は移った Daniel Scharstein との長きにわたる共同研究が始まった場所でもあった。

Microsoft Research においては、世界のトップを走るコンピュータビジョンとコンピュータグラフィックスの研究者たち — Michael Cohen, Hugues Hoppe, Stephen Gortler, Steve Shafer, Matthew Turk, Harry Shum, Anandan, Phil Torr, Antonio Criminisi, Georg Petschnigg, Kentaro Toyama, Ramin Zabih, Shai Avidan, Sing Bing Kang, Matt Uyttendaele, Patrice Simard, Larry Zitnick, Richard Hartley, Simon Winder, Drew Steedly, Chris Pal, Nebojsa Jojic, Patrick Baudisch, Dani Lischinski, Matthew Brown, Simon Baker, Michael Goesele, Eric Stollnitz, David Nistér, Blaise Aguera y Arcas, Sudipta Sinha, Johannes Kopf, Neel Joshi, Krishnan Ramnath — とともに働くという素晴らしい恩恵に浴している。また、優秀な学生インターンたち — Polina Golland, Simon Baker, Mei Han, Arno Schödl, Ron Dror, Ashley Eden, Jinxiang Chai, Rahul Swaminathan, Yanghai Tsin, Sam Hasinoff, Anat Levin, Matthew Brown, Eric Bennett, Vaibhav Vaish, Jan-Michael Frahm, James Diebel, Ce Liu, Josef Sivic, Grant Schindler, Colin Zheng, Neel Joshi, Sudipta Sinha, Zeev Farbman, Rahul Garg, Tim Cho, Yekeun Jeong, Richard Roberts, Varsha Hedau, Dilip Krishnan — にも恵まれた。

Microsoft で働きながら、（私が客員教授のポストを持つ）ワシントン大学の素晴らしい仲間たちと共同研究する機会も得た。Tony DeRose と David Salesin には、ワシントン大学で進められている研究に加わるように勧めてくれたことに感謝しなければならない。共同研究者として

付き合いの長い Brian Curless, Steve Seitz, Maneesh Agrawala, Sameer Agarwal, 古川泰隆にも感謝する。そして、指導教員として責任を持って指導した学生たち — Fr deric Pighin, Yung-Yu Chuang, Doug Zongker, Colin Zheng, Aseem Agarwala, Dan Goldman, Noah Snavely, Rahul Garg, Ryan Kaminsky — にも感謝する。冒頭にも述べたように、本書の執筆のきっかけになったのは、Steve Seitz が一緒に講義をしようと誘ってくれたコンピュータビジョンの授業であり、Steve の励ましであり、授業ノートであり、コメントである。

本書について多大なる建設的コメントをいただいた多くのコンピュータビジョン研究者たちにも、お礼を申し上げる。本書の編集者の役割を果たしてくれた Sing Bing Kang, MRF 推論のための線形計画法に関する付録 B.5.5 項を寄稿してくれた Vladimir Kolmogorov, そのほかにも Daniel Scharstein, Richard Hartley, Simon Baker, Noah Snavely, Bill Freeman, Svetlana Lazebnik, Matthew Turk, Jitendra Malik, Alyosha Efros, Michael Black, Brian Curless, Sameer Agarwal, Li Zhang, Deva Ramanan, Olga Veksler, Yuri Boykov, Carsten Rother, Phil Torr, Bill Triggs, Bruce Maxwell, Jana Ko eck, Eero Simoncelli, Aaron Hertzmann, Antonio Torralba, Tomaso Poggio, Theo Pavlidis, Baba Vemuri, Nando de Freitas, Chuck Dyer, Song Yi, Falk Schubert, Roman Pflugfelder, Marshall Tappen, James Coughlan, Sammy Rogmans, Klaus Strobel, Shanmuganathan, Andreas Siebert, Yongjun Wu, Fred Pighin, Juan Cockburn, Ronald Mallet, Tim Soper, Georgios Evangelidis, Dwight Fowler, Itzik Bayaz, Daniel O'Connor, Srikrishna Bhat に感謝する。Shena Deuchers は本書の原稿を整理するという素晴らしい仕事をこなし、さらに良いものに仕上げるための意見をいただいた。Springer 社の Wayne Wheeler と Simon Rees には、本書を出版する全ての段階でお世話になった。Keith Price の Annotated Computer Vision Bibliography は、関連研究を見つけ、文献を探し出すために、何物にも代えがたく有益であった。

本書をより良いものにするためのコメントは、ぜひメールで私宛てに送っていただきたい。できるだけ本書を、正確かつ有益で、タイムリーなものにしたいと思っている。

最後に、本書は、私の家族のひとかたならぬ支えと励ましがなかったならば、実現しなかったであろう。本書を両親と姉妹、そして家族に捧げる。また、私の両親である Zdzisław と Jadwiga に捧げる。両親が示してくれた愛すること、受け入れること、成し遂げたことは常に私に刺激を与え続けてくれた。妹の Basia に捧げる。彼女は生涯を通じた友情を教えてくれた。そして特に、日々出会う全ての物事（本書の執筆も含めて）を素晴らしいものにしてくれた、Lyn, Anne, Stephen に捧げる。

ウェナッチ湖にて  
2010年8月

# 訳者序文

本書は Richard Szeliski 著, *Computer Vision: Algorithms and Applications* (2010 年 11 月, Springer, ISBN 978-1848829343)\*2 の全訳である。本書は、「コンピュータに視覚を持たせる」技術であるコンピュータビジョンのほとんどのトピックを網羅したハンドブック的教科書と言ってよい。2000 年以降に開発された技術も数多く取り入れ、その様々な応用を紹介しながらも、多視点幾何や画像生成過程など、従来から扱われていた基礎的な内容も詳しく記述している。また、これまでは CG に分類されコンピュータビジョンのテキストではあまり扱われてこなかった、イメージベーストレンダリングやコンピュータシミュレーションフォトグラフィ、拡張現実 (AR) などについても詳しく解説し、映画において特殊視覚効果のために用いられているコンピュータビジョン技術も紹介している。パターン認識と機械学習にも章を割り、2005 年頃から普及したデジタルカメラの顔検出機能の原点となった手法も説明している。

800 ページを超える大著であるにもかかわらず、コンピュータビジョンという分野が広すぎるために、一つひとつの手法の記述は比較的簡素なものとなっている。そのため、本書だけでは各手法の詳細を理解することは難しいかもしれない。また、本書の特定の章のトピックを専門にしている研究者にとっては、その章の説明は物足りないだろう。しかし、そのトピックの全体を把握するための手法はほとんど紹介されており、重要な論文はもれなく参考文献に挙げられているはずである。つまり、本書は全容を理解し広範な知識を得ることができる教科書であり、また、主要な文献に到達するためのハンドブックでもある。詳細が知りたいと感じる手法を本書で見つけた場合には、ぜひ原論文に当たってほしい。

著者の Richard Szeliski 氏 (発音は「リチャード・シェリスキ」) は、コンピュータビジョンの分野だけでなく CG 分野でも著名な研究者であり、現在 Microsoft Research に籍を置き、活発な研究活動を進めている (その内容については、著者ウェブサイト <http://research.microsoft.com/en-us/um/people/szeliski/> を参照してほしい)。原著の出版は 2010 年秋であるが、その 1 年以上前からドラフトの PDF をインターネットで公開し、メーリングリストで意見を求めていたため、本書の存在は出版前から海外でも日本でもあちこちですでに話題に上っていた。現在も、原著ドラフトの PDF 版は原著ウェブサイト (<http://szeliski.org/Book/>) において無料で公開されており、内容の検索には便利である。

原著の出版からこの邦訳書の出版までに 2 年以上が経過している。この分野の研究の進展はめざましく、すでに原著には登場していない新しい概念、新しい手法、新しい研究成果が次々に発表されている。しかし、次の引用からもわかるように、この状況は 30 年以上も前から変わっていない。この分野は、萌芽期からその発展のスピードを維持しているようであり、これからもそうであり続けるだろう。

---

\*2 中国語版は「計算機視覚 — 算法与応用」(2011 年 12 月, 清华大学出版社, ISBN 978-7302269151)。

計算機による画像情報の処理についての最初の論文が出てから、もう約 20 年になる。それ以来今日まで、この分野は急速に発展してきており、英語による論文だけを数えても、(この分野で) 年間 500 以上の論文が発表されていると推定される。

— Rosenfeld & Kak 著 (1976) 「デジタル画像処理」序文より<sup>\*3</sup>

コンピュータ・ビジョンは比較的新しく、かつ急速に成長しつつある分野である。最初の実験は 1950 年代遅くに行われており、数多くの本質的な概念は最近の 5 か年に展開されている。

— Ballard & Brown 著 (1982) 「コンピュータ・ビジョン」序文より<sup>\*4</sup>

マシンビジョンはまだ新しく、急激に変化している分野である。この分野について執筆することは胸おどる仕事であるが、新しい結果がつつぎとあらわれるので、どこで見切りをつけ執筆を終えればよいかを決めるのが難しい。

— Horn 著 (1986) 「ロボットビジョン」まえがきより<sup>\*5</sup>

私 (玉木) が本書の翻訳を思い立ったのは 2011 年に入ってからである。原著を英語で読むことができればそれに越したことはないが、それは研究者や大学院生の場合である。学部生が 1 週間で読み通すためには、やはり日本語でなければならない。原著を翻訳すればそれが可能になり、ひいてはこの分野の発展に役立つという思いで、翻訳プロジェクトを (最初は一人で) 開始した。そうは言うものの、翻訳出版については右も左もわからず、出版業界に知人もいないため、どこにどう問い合わせたらよいのかもわからない。そこで、同じ分野の Forsyth & Ponce 著、*Computer vision: A Modern Approach* の翻訳書 (分量も同程度である) を出版していた共立出版に、2011 年 1 月中旬に問い合わせた (これを「翻訳企画の持ち込み」と言うことを、あとから知った)。そして早速検討するとの回答をいただいたのだが、その後に東日本大震災が発生し、企画内定の連絡をいただいたのは 2011 年 4 月であった。当時の混乱した状況において丁寧に対応していただいた共立出版編集部の石井徹也氏には大変お世話になり (現在も)、本書の出版の機会をいただけたことに、非常に感謝している。

その後、2011 年 4 月中に合計 8 人の翻訳メンバーを集め、担当する章の割り当てを決定し、Google グループと Google Docs (現在は Google Drive) を利用して、連絡手段と情報共有手段を整備した。各担当者の翻訳期間は 2011 年 6 月から 2012 年 3 月末までの 10 か月間と設定し、毎月 1 回は原稿を収集して PDF にまとめ、進捗状況を確認した。その後、2012 年 4 月から 8 月末までの 5 か月間で、用語と文体の統一、自然な日本語文章への修正、訳抜けのチェック、誤植と誤訳がないかどうかの確認を行った。この監訳作業の統一を図るためには一人で作業を行うほうが好ましいため、全てを玉木が行った。この段階で各担当者の翻訳文のほとんど全ての行に何らかの修正を施したため、もしわかりにくい文章や誤訳がある場合には、各担当者と玉木に責がある。そして 2012 年 9 月からの校正作業の終了を待ち、10 月下旬から 11 月末にかけて著者校正 (初校) を行い、さらに 2012 年 12 月末から 2013 年 1 月上旬にかけて再び著者校正 (再校)

<sup>\*3</sup> 長尾 真 監訳, 近代科学社 (1978).

<sup>\*4</sup> 福村晃夫ほか訳, 科学技術出版社 (1987).

<sup>\*5</sup> NTT ヒューマンインタフェース研究所プロジェクト RVT 訳, 朝倉書店 (1993).

を行った。この校正作業では(株)グラベルロードの担当の方々に大変お世話になり、そのおかげで表記の揺れや数式のミスなどはほとんどなくなったと思っている。特に、すべての URL のリンク切れがチェックされていたり、100 ページにもわたる参考文献リスト中に数か所あった原著の誤植までもが修正されていたことには、頭の下がる思いである。

翻訳については、この分野の大著を過去に翻訳した下記の諸先輩方を見習い、翻訳者として最大限の努力を払ったつもりである。しかし、どれだけ翻訳作業に慎重を期したとしても、誰にとっても納得できる訳語や翻訳文はでき上がらない。翻訳書というものは常に議論的になってしまうのである。翻訳書の日本語はわかりにくいから原書を読むべきだという意見や、原書と照らし合わせて翻訳を読んだら誤訳をいくつも見つけたという話を、あちこちで耳にする。本書の原著 PDF はウェブから入手できるので、簡単に訳文と原文とを比較できてしまう。そのため、本書についても翻訳の良し悪しが議論されるであろうことは想像に難くない。しかし、専門的な研究が本業で和英論文は執筆したことがあっても翻訳には携わったことのない者としては、人事を尽くして天命を待つよりほかはない。翻訳の力不足を補うためにも、読者諸氏のご意見やご批判をいただければ幸いである。

術語については、新しい研究分野では日本語訳が定着しておらず、研究者間では英語のまま使っていることが多い。慣用的にカタカナのままの方が自然なものを除いて、本訳書では、なるべく適当な訳をつけるようにした。

— (前出)「デジタル画像処理」訳者序文より

原本は解説書ではないため、わかり難い表現も多いが、翻訳に当ってはなるべく簡明な表現に意識した。また必要に応じて訳注をつけて本文を補った。

— Winston 編, 白井良明・杉原厚吉 訳「コンピュータービジョンの心理」訳者序文より<sup>\*6</sup>

訳し方は、表現上の技巧よりは内容に重きを置いて翻訳の経験不足を補うようにした。したがって、いささか大胆と思われるほどの意識をほどこした部分もあるが、意を尽せずに終わった部分も多い。

— (前出)「コンピュータ・ビジョン」訳者序文より

多人数での翻訳のため、用語の統一、および、文体の統一には最大限の努力を払った。

— (前出)「ロボットビジョン」訳者序文より

原著は L<sup>A</sup>T<sub>E</sub>X で執筆されていたが、残念ながらソースファイルは(参考文献 Bib<sub>T</sub>E<sub>X</sub> ファイルも)入手できなかったため、数式や表、参考文献リストなどは全て各担当者が入力した。原著にはない誤植が発生する可能性があるため、細心の注意を払って間違いをなくしたつもりである。また、索引は(ソースがないために)原著と同じものにすることができず、全て監訳作業において(和文と英文の索引を併記して)独自に付与した。図らずも、全て L<sup>A</sup>T<sub>E</sub>X で入力し直したことにより、原著の誤植が多数見つかった。これらの誤植は原著者に連絡済みであり、翻訳書ではそれを修正している。また、リンク先がなくなっていた URL は、可能な限り移転先の URL に修正してある。参考文献のうち、原著出版時に投稿中となっていたものや、ページ番号などが確定

<sup>\*6</sup> Winston (1975), 産業図書 (1979).

していなかったものについても、最新の文献情報に変更してある。これらの点に関しては、原著よりも正しい情報を提供できているだろう。参考文献リストのうち、翻訳書が出ているものについてはその情報を併記した。日本人が執筆した英語文献について、その内容が関連している日本語文献が見つけれられたものについても、関連日本語文献として併記してある。また、各章の内容に対応して、日本語で得られる参考文献情報を、原著にはない付録 D を設けて日本語版補遺という形で掲載した。本文中でも、用語や概念についてわかりにくいと思われるものに対しては、適宜訳注を添えた。主な訳注には索引をつけ、和文索引の「訳注」の項目に掲載してある。これらが読者の参考になれば幸いである。

本書の翻訳にあたって非常に多くの方々の協力を得た。たかだか一文の解釈や一単語の適切な訳語に困ったことも多々あり、その都度様々な方から意見をいただいた。特に早稲田大学の石川博氏、広島市立大学の日浦慎作氏、九州大学の長原一氏には、草稿中の数節に目を通していただいた。東京工業大学の野沢和輝氏には 6 章と 7 章の数式入力に協力していただいた。広島大学図書館副図書館長の甲斐重武氏とスタッフの方々には、推敲のための草稿製本にご協力いただいた。ここに心から感謝する。

2013 年 2 月

訳者代表 玉木 徹

.....

「次はこれを訳してみてもどうか」といって示されたのが本書である。私が、「このような本なら、日本にも書ける人が大ぜいいる」と言ったところ、「書ける人がいることは承知しているが、実際に書く人はいないだろう」という。

— Courant & Robbins 著 (1941), 森口繁一 監訳「数学とは何か」  
岩波書店 (1966), 監訳者あとがきより

.....