The University of Osaka
Institutional Knowledge Archive

Title	事前知識に基づく画像領域分割の高度化に関する研究
Author(s)	二神, 拓也
Citation	大阪大学, 2021, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/82291
rights	© 2020 The Author(s). Published with license by Taylor & Francis Group, LLC. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

事前知識に基づく画像領域分割の
高度化に関する研究

提出先　　大阪大学大学院情報科学研究科

提出年月　2021年1月

二 神 拓 也

研究業績目録

I. 学術論文

1. T. Futagami, N. Hayasaka, I. Taniguchi and T. Onoye. Experimental evaluation of reduction technique for time delay of turning of light to occupancy lighting control. *IEEJ Transactions on Electrical and Electronic Engineering*, Vol. 16, No. 2, Feb. 2021 (to appear).
2. I. Masakazu, T. Futagami, N. Hayasaka and T. Onoye. Acceleration of automatic building extraction via color-clustering analysis. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Science*, Vol. 103, No. 12, pp. 1599–1602, Dec. 2020.
3. T. Futagami and N. Hayasaka. Automatic product region extraction based on analysis of images uploaded to C2C online market. *Journal of Organizational Computing and Electronic Commerce*, Vol. 30, No. 4, pp. 323–334, Jul. 2020.
4. 二神拓也, 早坂昇, 尾上孝雄. 画像センサを用いた照明在不在制御の有効性評価. システム制御情報学会論文誌, Vol. 33, No. 5, pp. 139–148, May 2020.
5. T. Futagami, N. Hayasaka and T. Onoye. Fast and robust building extraction based on HSV color analysis using color segmentation and GrabCut. *SICE Journal of Control, Measurement, and System Integration*, Vol. 13, No. 3, pp. 97–106, May 2020.
6. T. Futagami, T. Yano, C. Huang and T. Enohara. Experimental evaluation for daylight-linked gradation lighting control using image-based motion sensors. *IEEJ Transactions on Electrical and Electronic Engineering*, Vol. 15, No. 5, pp. 723–732, May 2020.
7. 二神拓也, 早坂昇. 色クラスタ解析に基づく建造物領域の自動抽出. システム制御情報学会論文誌, Vol. 32, No. 9, pp. 349–356, Sep. 2019.

8. 二神拓也, 矢野亨, 黄静君, 榎原孝明. 画像センサによる照明個別制御の有効性検証. 電気学会論文誌 C, Vol. 138, No. 11, pp. 1362–1374, Nov. 2018.

II. 国際会議

1. T. Futagami, N. Hayasaka and T. Onoye. Performance comparison of saliency detection methods for food region extraction. *Proceedings of 4th International Conference on Graphics and Signal Processing*, pp. 1–4, Jun. 2020.
2. T. Futagami and N. Hayasaka. Automatic extraction of building regions by using color clustering. *Proceedings of SICE Annual Conference 2019*, pp. 415–419, Sep. 2019.
3. T. Futagami and N. Hayasaka. Automatic Extraction of Product Regions from Images on C2C Online Market. *Proceedings of 2019 2nd International Conference on Communication Engineering and Technology (ICCET)*, pp. 148–151, Apr. 2019.
4. T. Futagami, N. Hayasaka and Youji Iiguni. Speaker adaptation using improved MAP estimation with small amount of adaptation data. *Proceedings of 2013 IEEE Region 10 Conference (TENCON)*, Paper ID 135, Oct. 2013.

III. 国内会議

1. 二神拓也, 北田絢子, 早坂昇. 誘目度評価モデルを応用した食事領域抽出. 第 64 回システム制御情報学会研究発表講演会講演論文集, pp. 57–62, May 2020.
2. 二神拓也, 中島裕哉, 早坂昇. 照明在不在制御における消灯遅延時間適応の有効性検証. 第 64 回システム制御情報学会研究発表講演会講演論文集, pp. 98–103, May 2020.
3. 二神拓也, 岩井将一, 早坂昇. 夜間情景画像に対する建造物領域抽出の性能比較. 第 25 回画像センシングシンポジウム講演論文集, p. IS1-30, Jun. 2019.
4. 高橋雄介, 小林広幸, 二神拓也, 鴨雄史, 服部陽平. 走行安全支援向け前方障害物検知技術の開発 -安全な運転支援を目指して-. 第 54 回鉄道サイバネ・シンポジウム論文集, p. 530, Nov. 2017.
5. 二神拓也, 早坂昇. 画像センサによる照明制御の有効性評価. 第 63 回システム制御

情報学会研究発表講演会講演論文集, pp. 1104–1111, May 2019.

6. 二神拓也, 黄静君, 朝倉啓, 榎原孝明, 西村信孝. 画像センサを用いた照明制御に関する有効性検証. 平成 29 年電気学会産業応用部門大会, pp. V-223-V–226, Aug. 2017.
7. 二神拓也, 黄静君, 榎原孝明, 西村信孝. 画像センサを用いたグラデーション制御の有効性検証. 第 23 回画像センシングシンポジウム講演論文集, p. IS1-30, Jun. 2017.
8. Zhuokun Qi, 二神拓也, 早坂昇, 飯國洋二. 信頼度に基づいた調波構造強調による音声認識手法. 2014 電子情報通信学会総合大会講演論文集, p. A-20-6, Mar. 2014.
9. 二神拓也, 早坂昇, 飯國洋二. 少量の適応データで未観測問題に対応した MAP 推定による話者適応化手法. 日本音響学会春季研究発表会講演論文集, pp. 173–174, Mar. 2013.

IV. 解説記事

1. 二神拓也, 早坂昇. 色クラスタ解析を用いた建造物領域の自動抽出手法の解説. 画像ラボ, Vol. 31, No. 1, pp. 56–60, Jan. 2020.

内容梗概

可視光を検出して画像信号を取得するイメージセンサは, 物理限界に近い性能が得られており, 防災, 防犯, 計測, 車載, 医療, 生物, 農業, 食品, 材料, 宇宙, 環境の幅広い分野で利用されている. イメージセンサは, 静止画や動画の鑑賞目的のみに留まらず, 画像に写る内容をコンピュータが理解する画像認識技術へ応用されている. 画像認識技術には, 画像や矩形単位で物体の概念種別を予測するクラス認識や物体検出が存在するが, クラス認識や物体検出は物体の形状や輪郭線を抽出できない. 一方, 画素単位で領域種別を予測する画像領域分割は物体の輪郭線を抽出可能なため, クラス認識や物体検出の前処理としても用いられる重要な要素技術である.

幅広い分野で研究が進められている画像領域分割には様々な従来手法が存在する. 例として, 学習データを用いる局所特徴に基づく手法, ユーザからの手動入力を用いるインタラクティブ手法, 入力画像と領域種別に関する知見に基づきアルゴリズムが設計される事前知識に基づく手法が挙げられる. 局所特徴に基づく手法は高精度化に向けて学習データ量が増加するため導入コストが増大する課題を有する. また, インタラクティブ手法はユーザからの手動入力を要するためユーザビリティが低下する課題を有する.

一方, 事前知識に基づく手法は学習データやユーザからの手動入力が不要なため, 導入コスト増大とユーザビリティ低下の課題を回避できる. しかし, 事前知識に基づく手法の処理アルゴリズムは入力画像や領域種別に関する知見に基づいて設計されるため, 導入先の処理目的に依存する. そこで, 本論文では画像認識分野に貢献することを目的に, 導入先となる応用システムの特性を加味して事前知識に基づく手法を高度化する研究成果を述べる.

まず, ユーザが周囲を撮影した情景画像を建造物の領域と建造物以外の背景の領域に分割する手法を提案し, その有効性を領域分割精度と計算時間の観点から評価する. 提案手法は, 背景領域が画像の上部や下部に分布する事前知識に基づき, グラフ理論を用いて画像

領域を分割する GrabCut を適用する. GrabCut の初期化に必要な背景領域の候補を求めるために, 画像の上部や下部に分布する傾向のあるクラスタを抽出する処理を提案手法に導入する. 情景画像を用いた評価実験に基づき, 提案手法が建造物領域分割精度の向上と計算時間の短縮を実現することを確認する. また, 領域分割精度の低下を引き起こす従来手法の課題を提案手法が回避することも確認する. 提案手法は, 建造物領域の輪郭線情報を特徴とする建造物クラス認識の高性能化に寄与できる. 建造物クラス認識は情景画像を撮影したユーザの現在位置を測位する技術に応用されるため, 現在位置測位の性能向上に提案手法が貢献できる可能性がある.

続いて, インターネットオークション等で用いられる商品画像を商品の領域と商品以外の背景の領域に分割する手法を提案し, その有効性を領域分割精度と計算時間の観点から評価する. 提案手法は, 商品領域が画像の中央部, 背景が画像の外縁部に分布する事前知識に基づいて GrabCut を適用する. GrabCut の初期化に必要な背景領域の候補を求めるために, 商品領域分布の分析結果に基づいて定めた画像の外縁部を周辺の画素を併合しながら拡大させる処理を提案手法に導入する. 商品画像を用いた評価実験に基づき, 提案手法が商品領域分割精度の向上と計算時間の短縮を実現することを確認する. また, 領域分割精度の低下を引き起こす従来手法の課題を提案手法が回避することも確認する. 提案手法は, 商品領域の輪郭線情報を特徴とする商品クラス認識の高性能化に寄与できる. 商品クラス認識は商品画像に含まれる商品の名称や関連情報を自動取得するインターネットオークションの機能に応用されるため, インターネットオークションの効率化に提案手法が貢献できる可能性がある.

次に, 監視カメラが現在と過去に撮影した画像を比較して, 画素値の差分が大きい領域（以降, 差分領域と呼称する）と差分領域以外の背景の領域を分類する手法を応用した人感センサ（以降, 差分領域に基づく人感センサと呼称する）の有効性を評価する. 人感センサは在室者の検知結果を用いて電気設備を自動制御する技術（以降, 在不在制御と呼称する）に用いられるため, 差分領域分割に基づく人感センサを用いた在不在制御の省エネルギー効果や消費電力量を評価項目に加える. 一般家庭で収集した実験データに基づき, 差分領域分割に基づく人感センサの在室者検知精度と省エネルギー効果が既存の人感センサと比べて向上することを確認する. また, 差分領域分割に基づく人感センサを用いた在不在制御の消費電力量を評価し, 商用化されている差分領域分割に基づく人感センサの高度化に向けた考察を行う. 本研究は, 関連研究で明らかにされていない差分領域分割に基づく人

感センサの利点や課題を述べた点で有用で, 差分領域分割に基づく人感センサの普及を促す可能性がある. 差分領域分割に基づく人感センサは照明や空調設備の在不在制御に用いられるため, 本研究で得られる知見は一般家庭やオフィスにおける電気設備のエネルギー消費量の抑制に貢献できる可能性がある.

本論文は, 全 6 章で構成される. 第 1 章にて序論を述べ, 第 2 章では, 関連研究に基づいて画像領域分割の従来手法を説明する. まず, 一般的な多クラスの画像領域分割を定式化し, 本研究に関連する建造物領域分割, 商品領域分割, 差分領域分割を例示する. 続いて, 画像領域分割の従来手法を局所特徴に基づく手法, インタラクティブ手法, 事前知識に基づく手法に分類し, 各手法を代表例に基づき説明する. 次に, 各手法の利点と課題を整理して, 研究対象とする事前知識に基づく手法とその他の手法との差異を述べる.

第 3 章では, 情景画像の建造物領域を分割する手法を提案する. まず, 事前知識に基づいて建造物領域を分割する従来手法として, 上野らの手法の処理概要をフローチャートに基づき詳細に述べる. 続いて, 領域分割精度の低下を引き起こす上野らの手法の課題を説明する. 次に, 画像を構成する各画素を色情報に基づいてクラスタリングし, 生成されたクラスタを解析する提案手法の処理概要をフローチャートに基づき詳細に述べる. 更に, 106 枚の情景画像を用いた比較実験に基づいて領域分割精度と計算時間の観点から提案手法の有効性を示す. また, 提案手法が領域分割精度の低下を引き起こす上野らの手法の課題を回避することも確認する.

第 4 章では, 商品画像の商品領域を分割する手法を提案する. まず, 事前知識に基づいて商品領域を分割する従来手法として, Cheng らの手法の処理概要をフローチャートに基づき詳細に述べる. 続いて, 領域分割精度の低下を引き起こす Cheng らの手法の課題を説明する. 次に, 商品領域分布の分析結果に基づいて定めた画像の外縁部を周辺画素を併合しながら拡大させる提案手法の処理概要をフローチャートに基づき詳細に述べる. 更に, 341 枚の商品画像を用いた比較実験に基づいて領域分割精度と計算時間の観点から提案手法の有効性を示す. また, 提案手法が領域分割精度の低下を引き起こす Cheng らの手法の課題を回避することも確認する.

第 5 章では, 関連研究で十分に評価されていない差分領域分割に基づく人感センサの有効性を確認し, その利点や課題を明らかにする. 人感センサの在室者検知精度は在不在制御の省エネルギー効果や消費電力量を左右するため, 一般家庭やオフィスの消費エネルギー内訳に占める割合の高い照明設備の在不在制御を再現し, 省エネルギー効果と消費電

力量を評価する. まず, 人感センサの関連技術として, 差分領域分割に基づく人感センサについて説明する. 続いて, 照明在不在制御の処理概要をフローチャートに基づき詳細に述べる. 次に, 人感センサの有効性を確認する実験の実験条件を説明し, 在室者検知精度, 省エネルギー効果, 消費電力量の観点で差分領域分割に基づく人感センサの有効性を評価する. 更に, 実験結果に基づいて差分領域分割に基づく人感センサの利点と課題を整理し, 差分領域分割に基づく人感センサの改良課題を考察する. 最後に, 第 6 章で本研究で得られた成果を結論として要約する.

謝辞

本研究をまとめるにあたり, 研究に対する考え方, 研究者としての姿勢などについて懇切なる御指導, 御支援を賜わりました大阪大学大学院情報科学研究科情報システム工学専攻 尾上 孝雄 教授に深く感謝申し上げます.

本研究を推進するにあたり, 研究方針に関して御助言頂き, 様々な場面で惜しみない御指導, 御支援を頂きました大学大学院情報科学研究科情報システム工学専攻 竹村 治雄 教授, 谷口 一徹 准教授, 大阪電気通信大学 早坂 昇 准教授に心より感謝申し上げます.

本研究を推進するにあたり, 惜しみない御支援を頂きました大阪大学クリエイティブユニット 伊藤 雄一 准教授に心より感謝申し上げます.

共同研究者として, 有益な御討論を頂き, 多大なる御支援を頂きました株式会社東芝 矢野 亨 氏, 東芝インフラシステムズ株式会社 榎原 孝明 氏, 朝倉 啓 氏, 黄 静君 氏, 岩井 将一 氏, 大阪電気通信大学情報通信工学部情報工学科 中島 裕哉 氏に心より感謝申し上げます.

会社生活を通じて画像認識技術について多大なる御指導, 御支援を頂きました東芝インフラシステムズ株式会社 高橋 雄介 氏, 足利大学工学部創生工学科 松木 洋 助教に心から感謝と御礼を申し上げます.

ソフトウェア設計と実装に関して, 多数の貴重な御助言, 御支援を頂いた, 東芝インフラシステムズ株式会社 伊藤 博章 氏, 服部 陽平 氏, 鴨 雄史 氏, 藤浦 一夫 氏, 中村 幸太 氏, 安田 幸司 氏, 高橋 功 氏, 佐口 太一 氏に深く感謝の意を表します.

会社生活, 私生活を通じ, 多数の貴重な御助言, 御支援を頂いた, 東芝インフラシステムズ株式会社 三木 武郎 氏, 猪狩 精司 氏, 助川 寛 氏, 長谷部 光威 氏, 小林 広幸 氏, 鈴木 美彦 氏, 浅野 渉 氏, 山崎 世支明 氏, 大嶽 達哉 氏, 瀬戸 直人 氏, 堀江 勝大 氏, 間嶋 義喜 氏, 大阪電気通信大学情報通信工学部情報工学科 北田 絢子 氏に深く感謝の意を表します.

研究生活を通じて, 多数の貴重な御助言, 御支援を頂いた, 大阪大学大学院基礎工学研究科学研究科システム創成専攻 飯國 洋二 教授, 京都産業大学情報理工学部情報理工学科 川

村 新 教授, 近畿大学工学部電子情報工学科 吉田 大海 講師に厚く御礼申し上げます.

多くのご支援を頂いた尾上研究室秘書 吉田 友紀 氏に厚く御礼申し上げます.

研究生活, 私生活を献身的に支えて下さり, 常に寛大な心で接してくださいました株式会社リゾーム 生田目 真誇 氏に深い敬意と感謝を示します.

最後に, 研究生活を送る上で, 暖かい御支援と多大なる御理解を頂いた, 両親を始めとする家族に心から感謝と御礼を申し上げます.

目次

第 1 章　序論　1

1.1　研究背景 1

1.2　本論文の構成 7

第 2 章　画像領域分割の研究動向　11

2.1　画像領域分割の定式化 11

2.2　従来手法 12

2.2.1　局所特徴に基づく手法 12

2.2.2　インタラクティブ手法 18

2.2.3　事前知識に基づく手法 23

2.2.4　利点と課題 25

2.3　研究方針 26

2.3.1　建造物領域分割 27

2.3.2　商品領域分割 28

2.3.3　差分領域分割 28

2.4　まとめ 31

第 3 章　色クラスタ解析に基づく建造物領域分割手法　33

3.1　はじめに 33

3.2　上野らの手法 34

3.2.1　処理概要 34

3.2.2　課題 38

3.3　提案手法 40

3.3.1 処理概要 40
3.4 評価実験 45
3.4.1 実験条件 46
3.4.2 実験結果 50
3.4.3 考察 52
3.5 まとめ 54

第 4 章 画像外縁部に領域成長法を用いる商品領域分割手法 55
4.1 はじめに 55
4.2 Cheng らの手法 56
4.2.1 処理概要 56
4.2.2 課題 59
4.3 提案手法 59
4.3.1 商品領域分布の分析 60
4.3.2 処理概要 61
4.4 評価実験 64
4.4.1 実験条件 64
4.4.2 実験結果 66
4.4.3 考察 68
4.5 まとめ 69

第 5 章 差分領域分割に基づく人感センサの有効性評価 71
5.1 はじめに 71
5.2 Sarker らの手法 72
5.2.1 処理概要 73
5.3 照明在不在制御 74
5.3.1 処理概要 74
5.3.2 消灯遅延時間 75
5.4 評価実験 77
5.4.1 実験条件 78
5.4.2 実験結果 84

5.4.3 考察 89
5.5 まとめ 92

第 6 章 結論 **95**

参考文献 **99**

表目次

3.1 各手法の建造物領域分割精度 50
3.2 上野らの手法と提案手法の計算時間 52

4.1 各手法による商品領域分割精度 66
4.2 Cheng らの手法と提案手法の計算時間 68
4.3 各手法の改善率 69

5.1 実験スケジュール 79
5.2 リビングルームで観察された挙動例 (2019/1/6) 80
5.3 人感センサや LED 照明の最大・最小の定格電力 82
5.4 在室者検知精度 85
5.5 各日程の在率と省エネルギー効果 86

図目次

1.1 画像認識の例 2
1.2 画像領域分割結果の比較 4
1.3 建造物領域分割 5
1.4 商品領域分割 5
1.5 差分領域分割 6

2.1 画像領域分割の概要図 11
2.2 NN の例 13
2.3 NN の構造 14
2.4 畳み込み層の例 15
2.5 最大値プーリングの例 15
2.6 平均値プーリングの例 16
2.7 畳み込み層の適用例 17
2.8 FCN のネットワーク構造 17
2.9 逆畳み込み層の例 18
2.10 ドロップアウトの例 18
2.11 SegNet のネットワーク構造 19
2.12 逆プーリング層の例 19
2.13 GrabCut による画像領域分割例 20
2.14 GrabCut で生成するグラフ 21
2.15 GrowCut による画像領域分割例 23
2.16 CT 画像に対する肺領域分割 24
2.17 皮膚画像に対する黒子領域分割 24

2.18 食事画像に対する食事領域分割 25
2.19 GrabCut の適用事例（左：入力画像, 右：領域分割結果） 26
2.20 情景画像の例 27
2.21 情景画像と領域種別 28
2.22 商品画像の例 29
2.23 監視カメラ画像の例 29
2.24 レンズの断面図 30
2.25 赤外線センサの検知範囲（PIR sensor: 赤外線センサ） 31

3.1 上野らの手法のフローチャート 34
3.2 2 値化画像 35
3.3 平均隣接数と平均隣接数の増加率 36
3.4 画像中心部 36
3.5 上野らの手法による建造物領域分割過程 37
3.6 背景領域の確信度 37
3.7 上野らの手法による建造物領域分割精度低下の例 39
3.8 提案手法のフローチャート 40
3.9 提案手法による建造物領域分割過程 42
3.10 提案手法による色クラスタ解析の概要 43
3.11 各クラスタの含有率 44
3.12 ZuBuD の画像例 46
3.13 建造物と背景領域を示す正解データ画像の例 47
3.14 建造物領域の分布マップ 47
3.15 CamVid の画像例 48
3.16 半自動化手法による建造物領域の分割結果例 49
3.17 各手法による建造物領域分割結果例 51
3.18 上野らの手法で精度低下が発生していた画像に対する提案手法の適用結果 51
3.19 提案手法による建造物領域分割結果例 53

4.1 Cheng らの手法のフローチャート 57
4.2 図 2.22 の各画像に対するスーパーピクセル分割結果 57

4.3 Cheng らの手法による商品領域分割過程 58
4.4 誘目度マップの例（左：入力画像, 右：誘目度マップ） 58
4.5 Cheng らの手法による商品領域分割精度低下の例 60
4.6 商品と背景領域を示す正解データ画像の例 60
4.7 商品領域の分布マップ 61
4.8 提案手法のフローチャート 62
4.9 背景領域候補とする画像外縁部 62
4.10 提案手法による商品領域分割過程 63
4.11 半自動化手法による商品領域の分割結果例 65
4.12 各手法による商品領域分割結果例 67
4.13 Cheng らの手法で精度低下が発生していた画像に対する提案手法の適用結果 67

5.1 差分領域分割の例 73
5.2 照明在不在制御のフローチャート 75
5.3 照明在不在制御の例 76
5.4 不在判定時間と消灯遅延時間の例 76
5.5 リビングルームのレイアウト（PIR sensor: 赤外線センサ, Camera: カメラモジュール） 78
5.6 利用モジュール 79
5.7 誤不在判定時間のヒストグラム 84
5.8 在室者検知結果例（2019/1/19 10:30–11:00） 87
5.9 (5.17) 式の可視化（灰色領域: 領域 A, 白色領域: 領域 B） 89
5.10 差分領域分割に基づく人感センサの誤検知例 91

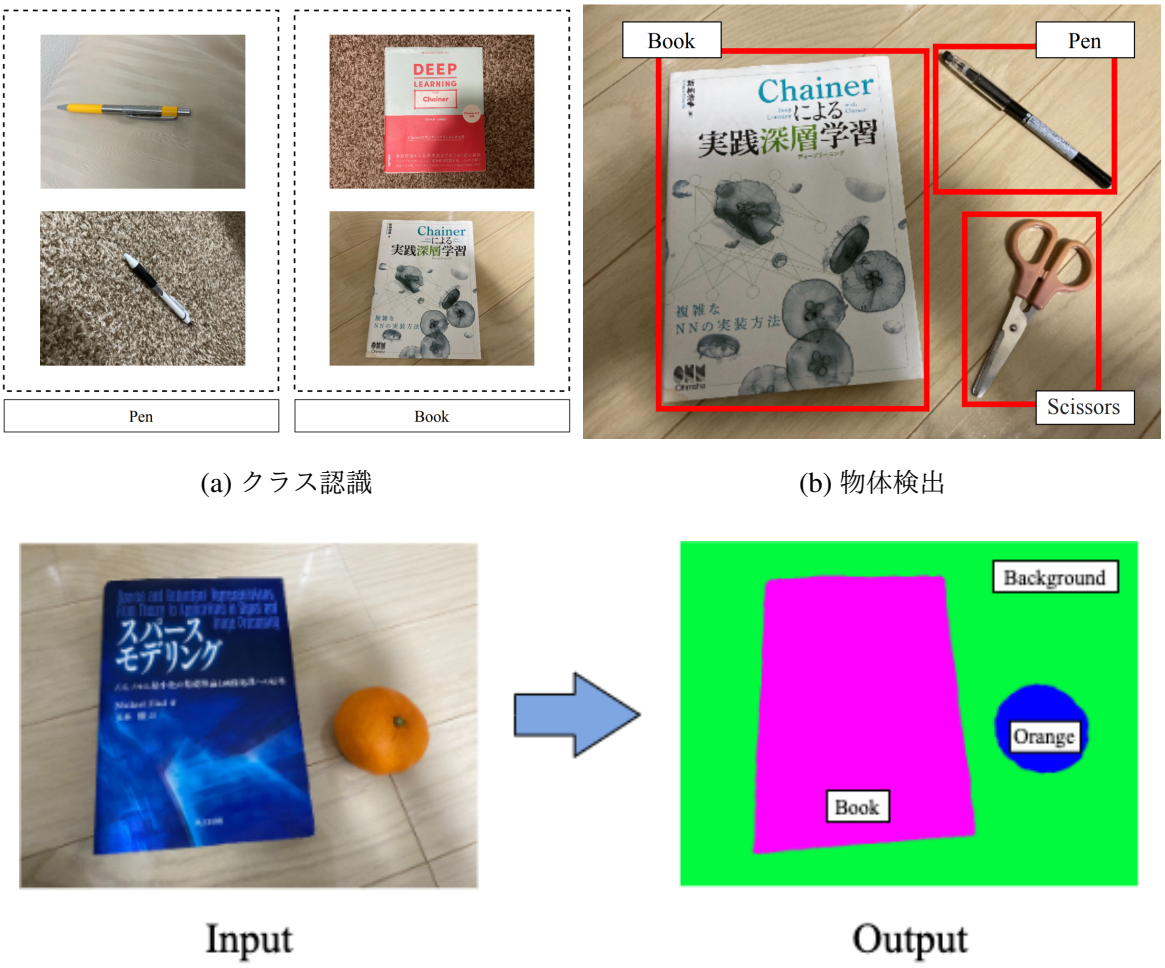
第 1 章

序論

1.1 研究背景

可視光を検出して画像信号を取得するイメージセンサは, スマートフォンやデジタルカメラなどに代表される民生機器のみでなく, 防災, 防犯, 計測, 車載, 医療, 生物, 農業, 食品, 材料, 宇宙, 環境の幅広い分野で利用されている [1–4]. 人間は主に視覚, 聴覚, 触覚, 味覚, 嗅覚の 5 つの感覚器を通じて外界から情報を収集するが, イメージセンサは知覚する情報量全体の 80% 以上を占める視覚を代替可能で [5], 重要なセンサとして高性能化に向けた研究が進められてきた. 現在, イメージセンサは物理限界に近い性能を得られており, 静止画や動画の鑑賞目的に留まらず, 画像に写る内容をコンピュータが自動で理解する画像認識技術への応用が拡大している [3].

画像認識は 1960 年代に研究が始まったと言われており [6], 1970 年代から 1980 年代には研究者の知識に基づいて認識対象の特徴を記述するアルゴリズムが設計されてきた [7]. 1990 年代から 2000 年代には, 汎用コンピュータの進化に伴って, 研究者が設計した局所的な画像特徴に基づく画像認識が盛んに研究されてきた [8]. 例えば, Scale Invariant Feature Transform (SIFT) 特徴量 [9] と Support Vector Machine (SVM) [10, 11] を組み合わせた画像認識は, 高精度化や高速化に向けた改良が重ねられて現在でも幅広い分野で実用化されている. 2010 年代以降には, 研究者の知識に基づいて設計されてきた画像特徴を自動で獲得する深層学習を応用した画像認識が研究されている. 1940 年代から研究されてきた Neural Network (NN) [12] を深層学習で多層化した Deep Neural Network (DNN) [13] は, コンペティションで既存の画像認識に対して大幅な精度向上を達成した [14]. 現在, DNN



(a) クラス認識　(b) 物体検出

(c) 画像領域分割

図 1.1: 画像認識の例

に基づく画像認識は様々な分野に向けて研究や実用化が進められている.

画像認識には, 図 1.1 のように画像, 矩形, 画素単位で物体の属する概念種別を予測するクラス認識, 物体検出, 画像領域分割が存在する [15]. クラス認識は, 画像全体に含まれる物体の概念種別を予測する. 図 1.1 (a) に例示するクラス認識は, 画像全体にペンと本のいずれかが含まれることを予測する. 物体検出は, 矩形の領域単位で物体の概念種別を予測する. 図 1.1 (b) に例示する物体検出は, 画像上に赤色の枠で示すペン, 本, はさみが含まれる矩形の領域を予測する. 画像領域分割は, 画素単位で物体の概念種別を予測する. 図 1.1 (c) に例示する画像領域分割は, 画像の各画素を桃色で示す本, 青色で示すみかん, 緑色で示す背景に分類する.

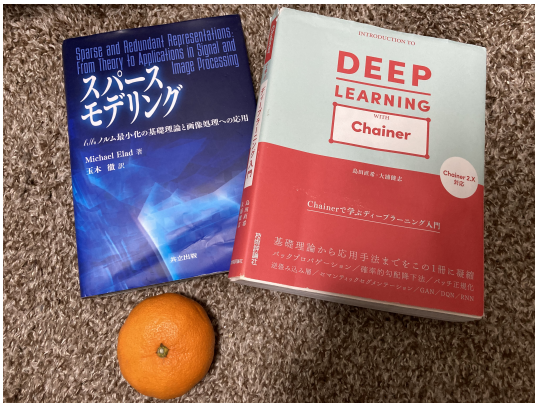
図 1.1 (a) に例示したクラス認識として, 入力画像の代表的な点をサンプリングして, 代

表点の周辺領域から画素間の共起を表現する局所特徴を抽出する手法が存在する [16]. DNN に基づいて局所特徴を抽出する手法も近年盛んに研究されている. しかし, クラス認識は物体が存在する画像上の位置を明らかにすることができない.

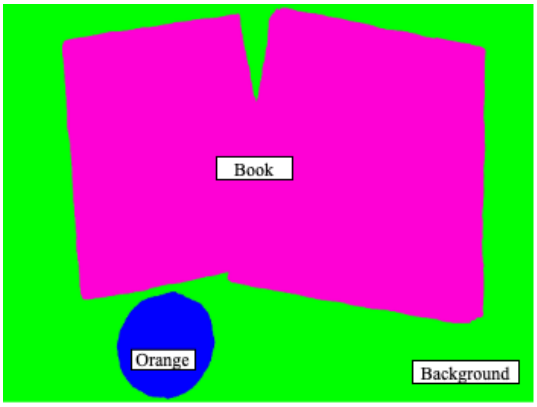
図 1.1 (b) に例示した物体検出として, 入力画像中に物体が存在する可能性の高い領域を物体候補として抽出して, 各物体候補にクラス認識を適用する手法が存在する [17]. 物体検出の計算時間を低減するためには, クラス認識が適用される物体候補を予め絞り込むことが重要である. 例えば, 物体が存在する度合を評価することで物体候補を絞り込む手法が存在する [18]. 物体検出は物体が存在する画像上の大まかな位置が分かるが, 物体が含まれる矩形の内部には物体の存在しない画素が含まれるため, 物体の形状や輪郭線を明らかにすることができない.

図 1.1 (c) に例示した画像領域分割として, 局所特徴を深層学習に基づき抽出する手法（以降, 局所特徴に基づく手法と呼称する）が主流である. 画像領域分割は, 医療, 車載, 計測, 防災, 食品の広い分野で応用が検討されている [19–23]. 認識単位の小さい画像領域分割は画像認識の中で最も困難な技術の 1 つとして捉えられているが [8], クラス認識や物体検出で求められなかった物体の形状や輪郭線を抽出可能な利点を有する. そのため, 画像領域分割はクラス認識や物体検出の前処理として用いられる重要な要素技術である. 例えば, クラス認識には物体の輪郭線に基づいて局所特徴を抽出する手法が存在し [24], 物体検出には物体候補の絞り込みに画像領域分割を用いる手法が存在する [25, 26]. 更に, 画像中に存在する特定の物体を 3 次元画像化するためにも画像領域分割が用いられる [27]. 本論文では, 画像領域分割に関する研究成果を述べる. 以降, 画像領域で各画素に割り当てる概念種別を領域種別と呼称する.

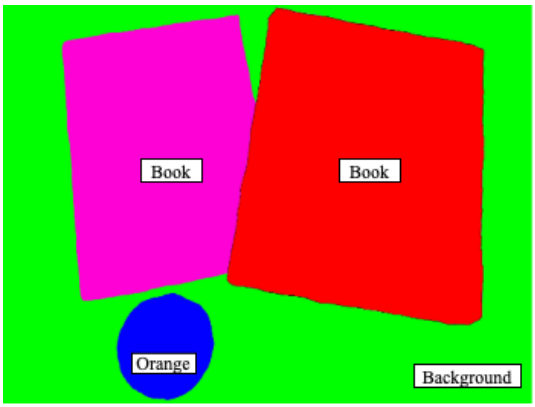
画像領域分割には, 領域種別単位で領域分割するセマンティックセグメンテーション [28] と個体単位で領域分割するインスタンスセグメンテーション [29] が存在する. セマンティックセグメンテーションとインスタンスセグメンテーションによる画像領域分割結果を図 1.2 を用いて説明する. 図 1.2 (a) は入力画像, 図 1.2 (b), (c) はそれぞれセマンティックセグメンテーションとインスタンスセグメンテーションを適用した結果である. 図 1.2 (b), (c) を確認すると, セマンティックセグメンテーションと比べて, インスタンスセグメンテーションは 2 冊の隣接する本を異なる本として区別できる利点を有することが分かる. この利点はインスタンスセグメンテーションが物体検出とセマンティックセグメンテーションの組み合わせで実現されることに起因するが, インスタンスセグメンテー



(a) 入力画像



(b) セマンティックセグメンテーション



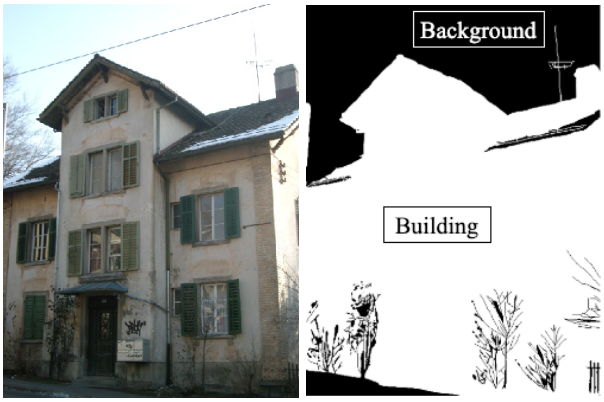
(c) インスタンスセグメンテーション

図 1.2: 画像領域分割結果の比較

ションはセマンティックセグメンテーションと比較して計算コストが高い点で実用化に向けて課題がある [8]. そこで, 本論文では計算コストが低く, 実用化が進められているセマンティックセグメンテーションに関する研究成果を示す.

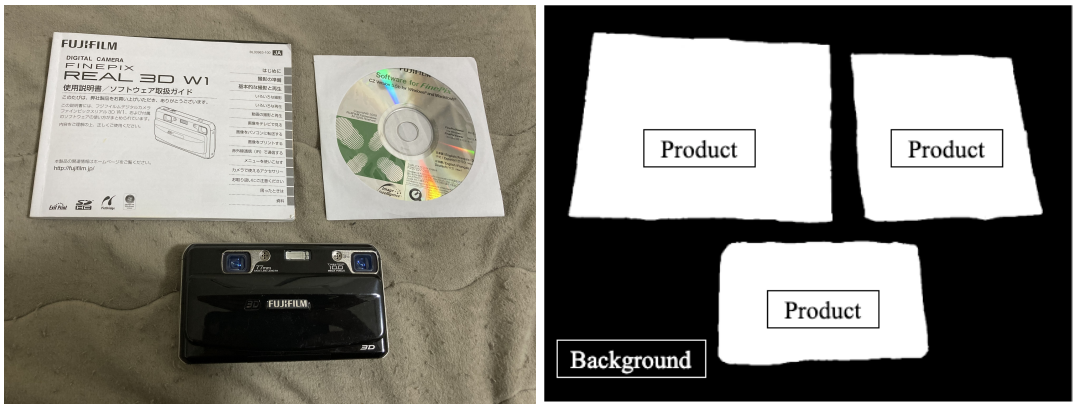
ここで画像領域分割の例を示す. 図 1.3 は, ユーザが周囲を撮影した情景画像の各画素を建造物領域と建造物以外の背景領域に分類する建造物領域分割の例である. 図 1.3 (a) は入力画像で, 図 1.3 (b) は建造物領域分割を適用した結果である. 図 1.3 (b) の白色と黒色に示す領域はそれぞれ建造物領域と背景領域をである. 建造物領域分割を高度化することで, 建造物の輪郭線を特徴とする建造物クラス認識の高性能化に寄与できる. 更に, 建造物クラス認識は情景画像を撮影したユーザの位置を測位する技術に応用されているため, 建造物領域分割を高度化することで現在位置測位技術の高性能化に貢献できる可能性がある.

図 1.4 は, インターネットオークション等で用いられる商品画像の各画素を商品領域



(a) 情景画像　(b) 領域分割結果

図 1.3: 建造物領域分割



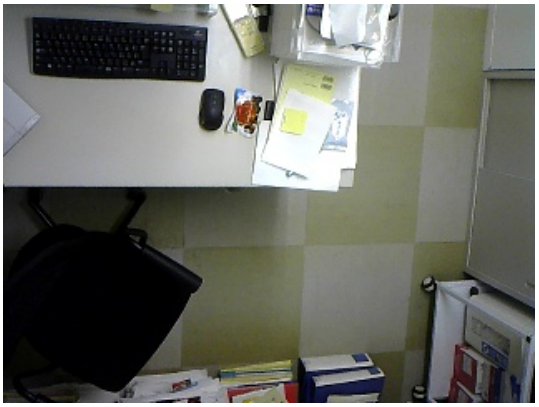
(a) 商品画像　(b) 領域分割結果

図 1.4: 商品領域分割

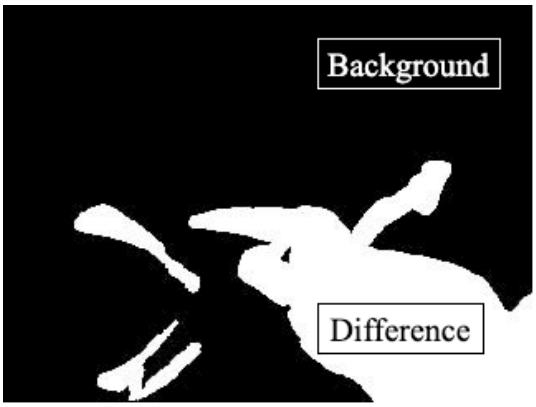
と商品以外の背景領域に分類する商品領域分割の例である．図 1.4 (a) は入力画像で, 図 1.4 (b) は商品領域分割を適用した結果である．図 1.4 (b) の白色と黒色に示す領域はそれぞれ商品領域と背景領域である．商品領域分割を高度化することで, 商品の輪郭線を特徴とする商品クラス認識の高性能化に寄与できる．更に, 商品クラス認識は商品画像に含まれる商品の名称や関連情報を自動で取得するインターネットオークションの機能に応用されているため, 商品領域分割を高度化することでインターネットオークションの効率化に貢献できる可能性がある．

図 1.5 は, 監視カメラで過去と現在に撮影した画像を比較して, 差分が発生する領域（以降, 差分領域と呼称する）と差分領域以外の背景領域に各画素を分類する差分領域分割で

(a) 過去の監視カメラ画像



(b) 現在の監視カメラ画像



(c) 領域分割結果

図 1.5: 差分領域分割

ある. 図 1.5 (a),(b) は過去と現在に撮影した入力画像で, 図 1.5 (c) は差分領域分割を適用した結果である. 図 1.5 (c) の白色と黒色に示す領域はそれぞれ差分領域と背景領域である. 近年, 監視カメラを内蔵し, 撮影した画像に差分領域分割を適用することで在室者の検知結果を出力する人感センサ（以降, 差分領域分割に基づく人感センサと呼称する）が商用化されているが, 差分領域分割に基づく人感センサの有効性が関連研究で十分に評価されていない. 差分領域分割に基づく人感センサの利点や高度化に向けた課題を明らかにすることで, 差分領域分割に基づく人感センサの普及を促す可能性がある. また, 人感センサの在室者検知結果に基づいて電気設備を自動制御する技術（以降, 在不在制御と呼称する）の高性能化に寄与し, 一般家庭やオフィスのエネルギー消費量を抑制できる可能性がある.

上述した画像領域分割は, 入力画像から分離したい建造物領域, 商品領域, 差分領域を前景領域として背景領域と区別する. 局所特徴に基づく手法を用いて色や形状の異なる前景

領域と背景領域を分割するには, 前景領域や背景領域に応じて学習データを事前に準備する必要がある. 特に, 現在主流の深層学習に基づいて局所特徴を抽出する手法はその他の機械学習と比較して大量かつバリエーション豊富な学習データを要するため, 導入コストが増大する [30]. 他にも, 入力画像を目視確認したユーザからの手動入力に基づいて画像領域を分割する手法（以降, インタラクティブ手法と呼称する）も存在するが, ユーザからの手動入力を要するためユーザビリティが低下する課題がある.

一方, 入力画像や領域種別に関する事前知識に基づいてアルゴリズムを手動で設計する手法（以降, 事前知識に基づく手法と呼称する）が存在する. 図 1.3 の建造物領域分割は, 建造物が画像の中心部に分布することや, 上空や道路などの背景は画像の上半分と下半分にそれぞれ分布することを事前知識として用いる. 図 1.4 の商品領域分割は, 商品が背景領域と比較して人の注目を集めるように画像の中央部に撮影されることを事前知識として用いる. 図 1.5 の差分領域分割は, 現在と過去に撮影した画像間で画素値の差分が大きい領域に差分領域が含まれることを事前知識として用いる.

事前知識に基づく手法は, 局所特徴に基づく手法のように学習データを必要としないため導入コストを低減可能である. 更に, インタラクティブ手法のようにユーザからの手動入力を必要としないためユーザビリティの低下が存在しない. そこで, 本論文は導入コスト増加やユーザビリティ低下を回避して実用化できる事前知識に基づく手法を研究対象とする. 入力画像や領域種別に関する知見に基づいて設計される事前知識に基づく手法のアルゴリズムは, 処理目的に強く依存する点に課題がある. そのため, 導入先となる応用システムの特性を加味して事前知識に基づく手法を高度化する研究成果を述べる.

事前知識に基づく手法の領域分割精度を向上するには, 入力画像や領域種別に関する特徴をアルゴリズムに適切に反映することが重要である. そこで, 本研究では既存の事前知識に基づく手法の課題を分析し, 課題を回避可能なアルゴリズムを提案することで事前知識に基づく手法の高度化を実現する. また, 事前知識に基づく手法に関する有効性評価を実施し, 高度化に向けた改良課題を得る.

1.2　本論文の構成

本論文は, 全 6 章で構成される. 第 2 章では, 画像領域分割の従来手法について詳細に説明する. まず, 画像領域分割を定式化する. 続いて, 画像領域分割の従来手法を局所特徴に基づく手法, インタラクティブ手法, 事前知識に基づく手法に分類し, 各手法の代表例を詳

細に説明する. 次に, 各手法の利点や課題を整理し, 研究対象とする事前知識に基づく手法とその他の手法の差異を明確化する. 更に, 研究方針を述べる.

第 3 章では, 情景画像から建造物領域を分割する手法を提案し, その有効性を示す. まず, 建造物領域を分割する従来手法の処理アルゴリズムをフローチャートに基づき詳細に述べ, 領域分割精度が低下する従来手法の課題を整理する. 続いて, 提案手法の処理アルゴリズムをフローチャートに基づき詳細に詳べる. 画像の上部や下部に背景領域が分布することを事前知識とする提案手法には, 入力画像の各画素を色情報でクラスタリングし, 背景領域を含む割合の高いクラスタを判定する処理を導入する. 次に, 建造物領域分割精度と計算時間の観点から提案手法の有効性を確認する実験の実験条件を説明し, 実験結果に基づき提案手法が 86.4% の平均計算時間低減と 10.1% の領域分割精度向上を実現したことを述べる. 領域分割精度の低下を引き起こす従来手法の課題を提案手法が回避できることも確認する. 他にも, 更なる精度向上に向けて提案手法で発生した課題を整理する.

第 4 章では, 商品画像から商品領域を分割する手法を提案し, その有効性を示す. まず, 商品領域を分割する従来手法の処理アルゴリズムをフローチャートに基づき詳細に述べ, 領域分割精度が低下する従来手法の課題を説明する. 続いて, 提案手法の処理アルゴリズムをフローチャートに基づき詳細に詳べる. 画像の外縁部に背景領域が分布することを事前知識とする提案手法には, 周辺画素を併合しながら画像の外縁部を拡大させることで背景領域の候補を求める処理を導入する. 次に, 商品領域分割精度と計算時間の観点から提案手法の有効性を確認する実験の実験条件を説明し, 実験結果に基づき提案手法が 76.7% の平均計算時間低減と 3.2% の領域分割精度向上を実現したことを述べる. 更に, 領域分割精度の低下を引き起こす従来手法の課題を提案手法が回避できることを確認し, 提案手法を構成する各処理の有効性を考察する.

第 5 章は, 関連研究で十分に評価されていない差分領域分割に基づく人感センサの有効性を実環境下で収集したデータを用いて確認し, その利点や課題を考察する. 人感センサは, 在室者の検知結果に基づいて電気設備を自動制御する技術（以降, 在不在制御と呼称する）の省エネルギー効果や消費電力量を左右する. そこで, 一般家庭やオフィスの消費エネルギー内訳に占める割合の高い照明設備の在不在制御を再現し, 省エネルギー効果と消費電力量を評価する. まず, 現在商用化が進められている差分領域分割に基づく人感センサを関連研究に基づいて詳細に説明する. 続いて, 照明在不在制御の処理概要をフローチャートに基づき説明する. 次に, 人感センサの有効性を評価する実験の実験条件を説明

し, 在室者検知精度, 省エネルギー効果, 消費電力量の観点で差分領域分割に基づく人感センサの有効性を考察する. 差分領域分割に基づく人感センサは安価で普及している赤外線センサに対して 21.7% の在室者検知精度改善を実現し, 照明に対する在不在制御の省エネルギー効果を赤外線センサと比べて 9.68% 向上したことを説明する. 更に, 差分領域分割に基づく人感センサの利点と課題を考察し, 差分領域分割に基づく人感センサの改良課題について述べる. 具体的には, 差分領域分割に基づく人感センサ自体の消費電力が赤外線センサと比べて高く, 照明在不在制御の消費電力量削減を実現できない結果が得られたことを述べる.

第 6 章は, 本研究で得られた成果を結論として要約し, 研究成果に基づいて今後の展望を示す. なお, 第 3 章は文献 [31–34], 第 4 章は文献 [35], 第 5 章は文献 [36–38] で公表した内容に関連する.

第 2 章

画像領域分割の研究動向

本章の内容について説明する．2.1 節で研究対象とする画像領域分割を定式化する．2.2 節で画像領域分割の従来手法を局所特徴に基づく手法, インタラクティブ手法, 事前知識に基づく手法に分類し, 関連研究に基づいて詳細を述べる．各手法の利点や課題を整理し, 研究対象とする事前知識に基づく手法とその他の手法との差異を明確化する．2.3 節で研究方針を述べ, 本章の内容を 2.4 節でまとめる．

2.1 画像領域分割の定式化

本節では, 幅と高さの画素数がそれぞれ W, H で総画素数が $P(= H \cdot W)$ の画像 $X = \{x_0, \ldots, x_{P-1}\}$ を入力とする画像領域分割を定式化する．画像領域分割は, 入力画像の各画

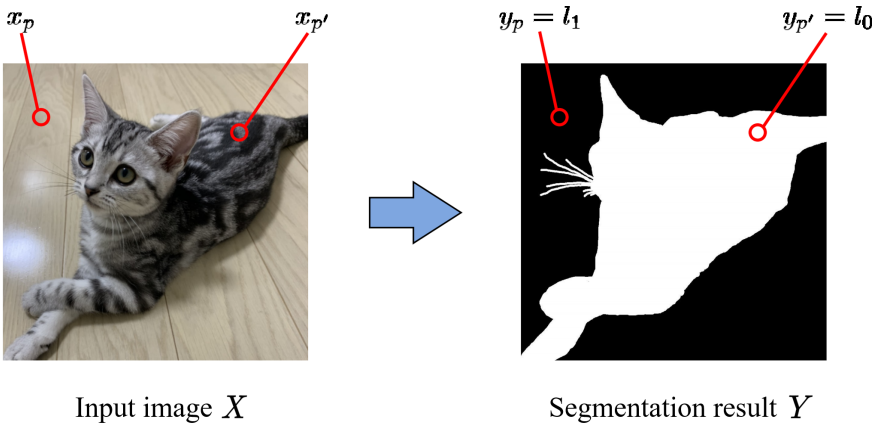


図 2.1: 画像領域分割の概要図

素に事前に定めた N_L 種類の領域種別 $L = \{l_0, \ldots, l_{N_L-1}\}$ のいずれかを割り当て, 領域分割結果 $Y = \{y_0, \ldots, y_{P-1}\}$ を出力する処理を意味する. 図 2.1 に例示する画像領域分割は, 画像の各画素に 2 種類の領域種別 $L = \{l_0, l_1\}$ のいずれかを割り当てる. l_0 と l_1 は, それぞれ猫の領域とそれ以外の背景領域を示す.

以降の本論文は, ユーザが入力画像から分離したい前景領域 l_{fgd} と, それ以外の背景領域 l_{bgd} のいずれかを入力画像の各画素に割り当てる画像領域分割に関して述べる. ここで, 本論文に関連する画像領域分割の具体例を示す.

まず, 図 1.3 の画像領域分割は建造物が撮影された情景画像の各画素に領域種別 $L = \{l_{\mathrm{bld}}, l_{\mathrm{bkg}}\}$ のいずれかを割り当てる建造物領域分割である. l_{bld} と l_{bkg} は, それぞれ前景領域に相当する建造物領域と背景領域を示す.

続いて, 図 1.4 の画像領域分割はインターネットオークション等で用いられる商品画像の各画素に領域種別 $L = \{l_{\mathrm{pdc}}, l_{\mathrm{bkg}}\}$ のいずれかを割り当てる商品領域分割である. l_{pdc} と l_{bkg} は, それぞれ前景領域に相当する商品領域と背景領域を示す.

次に, 図 1.5 の画像領域分割は現在と過去に撮影された監視カメラ画像去に撮影された監視カメラ画像（それぞれ X^{t-1}, X^t とする）を比較して, X^t の各画素に領域種別 $L = \{l_{\mathrm{dif}}, l_{\mathrm{bkg}}\}$ のいずれかを割り当てる差分領域分割である. l_{dif} と l_{bkg} は, それぞれ前景領域に相当する差分領域と背景領域を示す.

2.2　従来手法

本節では, 前節で定式化した画像領域分割の従来手法を示す. 従来手法は, 学習データを用いる局所特徴に基づく手法 [39,40], ユーザからの手動入力を用いるインタラクティブ手法 [41,42], 入力画像や領域種別に関する知見に基づいてアルゴリズムが設計される事前知識に基づく手法 [43–49] に分類される. 2.2.1〜2.2.3 節で各手法の代表例を挙げて従来手法を詳細に説明する. 2.2.4 節で各手法の利点と課題を整理して, 研究対象とする事前知識に基づく手法とその他の手法との差異を明確化する.

2.2.1　局所特徴に基づく手法

局所特徴に基づく方式として, 人間の脳の神経回路を模擬した NN を応用する手法が提案されている. 入力層に 4 個のユニット, 隠れ層に 2 個のユニット, 出力層に 3 個のユニッ

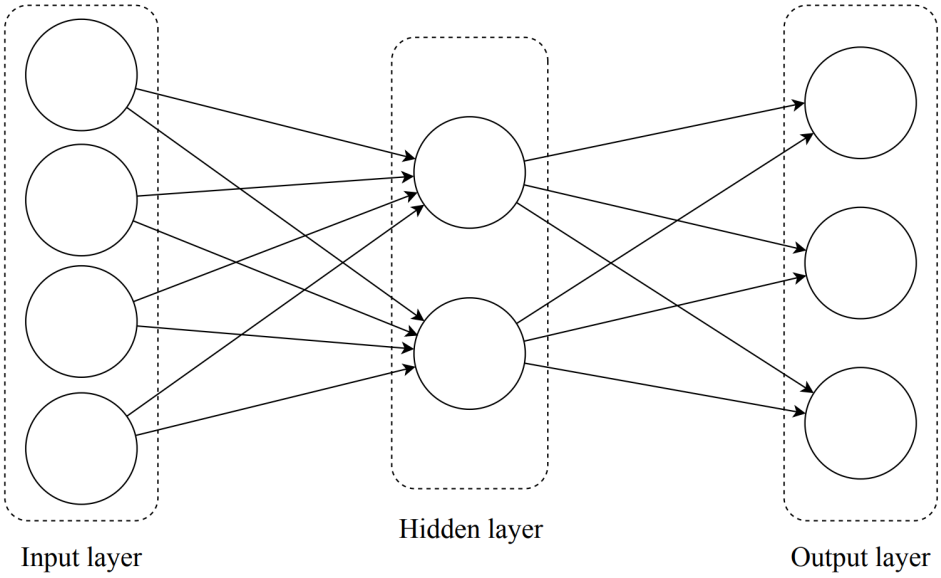


図 2.2: NN の例

トが存在する NN を図 2.2 に例示する．なお, 図中の矢印は前方のユニットから後方のユニットに対する接続を示している．

入力層と隠れ層のユニット間における入出力を図 2.3 に例示する．図 2.3 における隠れ層への入力値 z は, 入力層の出力値 s を用いて以下で示される．

$$z_1 = w_{11}s_1 + w_{12}s_2 + w_{13}s_3 + w_{14}s_4 + b_1 \tag{2.1}$$

$$z_2 = w_{21}s_1 + w_{22}s_2 + w_{23}s_3 + w_{24}s_4 + b_2 \tag{2.2}$$

w と b はそれぞれユニット間のつながりの強さを表す重みとバイアスを示すパラメータである．隠れ層の出力値 z' は, 脳のシナプスがある閾値を超えると発火することを模倣して設計されている活性化関数 σ_f を用いて入力値 z を変換することで求められる [50]．画像認識分野における活性化関数の例として, Rectified Linear Unit (ReLU) 関数やソフトマックス関数が挙げられる．ReLU 関数は入力値 z を以下のように変換し, 隠れ層の出力値 z' を求める [51]．

$$z' = \max(z, 0) \tag{2.3}$$

ここで, max 関数は引数の中で最大値を返す関数を意味する．ソフトマックス関数は, 各ユ

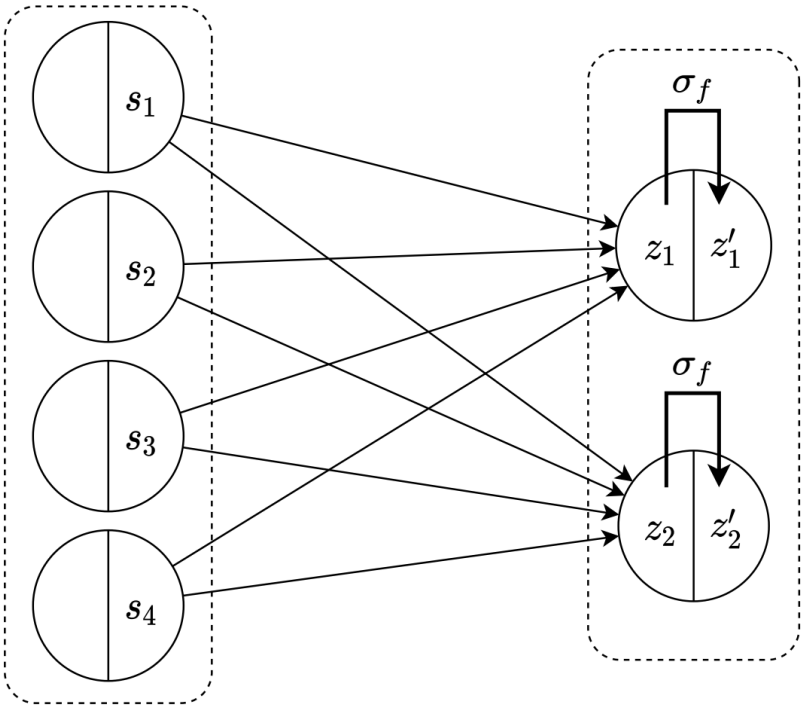


図 2.3: NN の構造

ニットからの出力値の総和が 1 となるような以下の補正をする [52].

$$z'_k = \frac{\exp(z_k)}{\sum_{k'=1}^{K} \exp(z_{k'})} \tag{2.4}$$

ここで, K はユニットの総数を示す.

局所特徴に基づく手法として, 近年, Convolutional Neural Network (CNN) が広く研究されている. CNN は, 入力画像の各画素値をユニットに対応させた入力層に対して, 関係性の強い隣接する画素間の共起を表現する局所特徴を畳み込み層, プーリング層を用いて抽出する [53].

畳み込み層は, 入力データに同一の係数を有するフィルタを重畳し, 局所特徴として特徴マップを出力する. 図 2.4 に畳み込み層を例示する. プーリング層は, 特徴マップのサイズを低下させるために特徴マップの局所領域内における代表値を抽出する. 入力画像の僅かな差異や変化が吸収される効果を期待できる. 2×2 領域における最大値を代表値とする最大値プーリングを図 2.5 に例示する. 他にも, 局所領域内の平均値を代表値とする平均値プーリングも存在するが, 画像認識分野では最大値プーリングが一般的に用いられる [54].

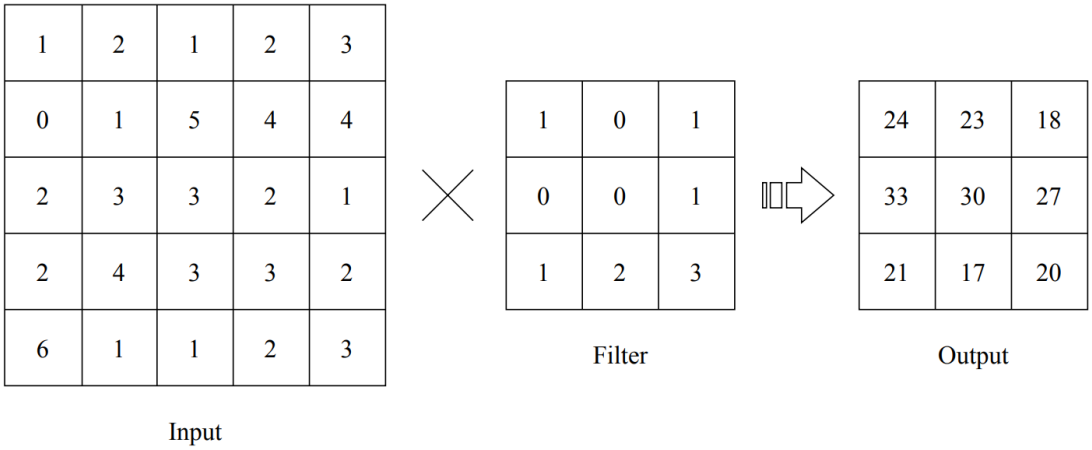


図 2.4: 畳み込み層の例

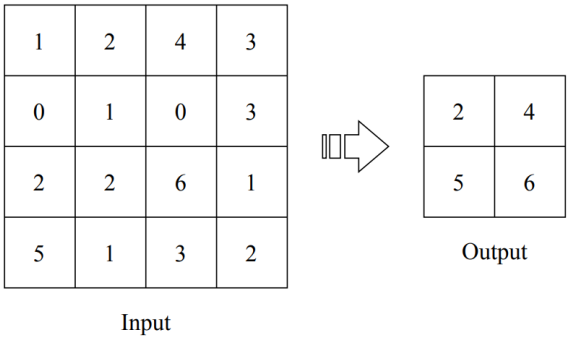


図 2.5: 最大値プーリングの例

2×2 領域における平均値を代表値とする平均値プーリングを図 2.6 に例示する. CNN は畳み込み層とプーリング層を繰り返し用いることで, 入力画像のサイズを少しずつ低下させながら異なるスケールの局所特徴を抽出する.

畳み込み層を入力画像の RGB 成分に適用し, 特徴マップを出力する例を図 2.7 に示す. 図 2.7 では, 幅と高さの画素数がそれぞれ W, H で, RGB の 3 チャンネルを有する入力画像の各画素をユニットとした入力層に, $W_{\mathrm{f}} \times H_{\mathrm{f}} \times 3$ のフィルタを重畳している. 合計 Q 種類のフィルタを重畳することで, 幅と高さの画素数がそれぞれ W, H で, Q チャンネルを有する特徴マップが出力される.

CNN の出力と学習について説明する. CNN は, 入力画像の画素 p が領域種別 l に所属する確率 $s'_{p,l}$ を出力するため, 以下の式で領域種別を決定する.

$$y_p = \arg\max_{l \in L} s'_{p,l} \tag{2.5}$$

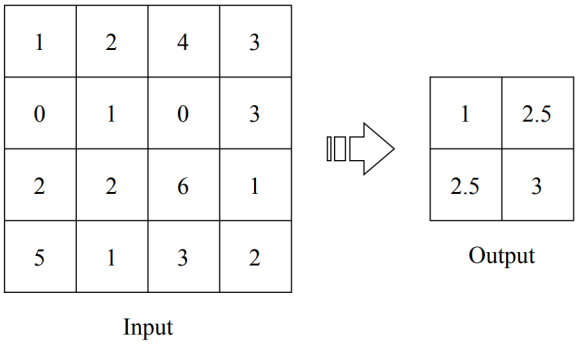


図 2.6: 平均値プーリングの例

CNN を学習する際には, 画像の各画素に領域種別が正しく割り当てられた正解データ画像を用いる. 以下の交差エントロピー誤差 E_{N} を最小化するようにネットワークを学習する.

$$E_{\mathrm{N}} = \sum_{\tilde{n}=0}^{\tilde{N}-1} \sum_{p=0}^{P-1} s^{\mathrm{gt}}_{p,l,\tilde{n}} \log(s'_{p,l,\tilde{n}}) \tag{2.6}$$

ここで, $s'_{p,l,\tilde{n}}$ は $\tilde{n}$ 枚目の画像の画素 p が領域種別 l に所属する確率を意味する. $s^{\mathrm{gt}}_{p,l,\tilde{n}}$ は, 正解データ画像に基づいて以下で求められる.

$$s^{\mathrm{gt}}_{p,l,\tilde{n}} = \begin{cases} 1 & (u_{p,\tilde{n}} = l) \\ 0 & (\text{otherwise}) \end{cases} \tag{2.7}$$

上式の $u_{p,\tilde{n}}$ は, $\tilde{n}$ 枚目の正解データ画像の画素 p における領域種別を意味する. 以降, 局所特徴に基づく手法の代表例として様々な分野に応用されている Fully Convolutional Network (FCN) [55] と SegNet [56] を説明する.

FCN

図 2.8 に FCN のネットワーク構造を示す. FCN は, 畳み込み層, 最大値プーリング層, 逆畳み込み層によって構成されている. まず, FCN は畳み込み層とプーリング層を組み合わせて, 入力画像を低解像度の特徴マップに変換する. 続いて, 逆畳み込み層で特徴マップを入力画像の解像度まで拡大し, 入力画像の各画素に領域種別を割り当てる. 特徴マップをアップサンプリングさせる逆畳み込み層を図 2.9 に例示する. 逆畳み込み層は, 畳み込み層のフィルタ処理の前に出力解像度を増大させる処理を有する [57].

過学習の防止に向けて, FCN にはバッチ正規化やドロップアウトが導入されている. バッチ正規化は, 入力データ分布の変化を抑えるために各層に入力する値の平均値と分

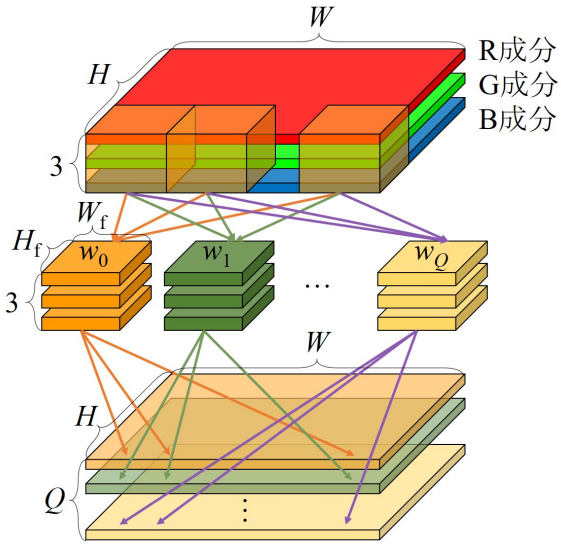


図 2.7: 畳み込み層の適用例

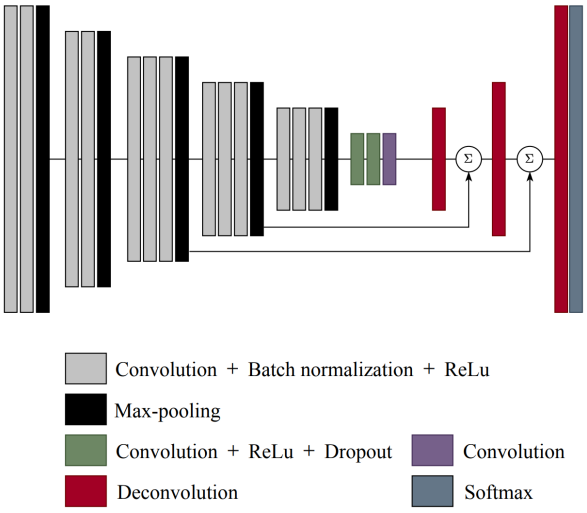


図 2.8: FCN のネットワーク構造

散値を正規化する処理を意味する [58]. ドロップアウトは, ランダムに選択した一部のユニットとそのユニットに対する接続を取り除いて学習する [59]. ドロップアウトを図 2.10 に例示する.

SegNet

図 2.11 に SegNet のネットワーク構造を示す. 中間層の特徴マップを記憶する必要がある FCN はメモリ使用量が大きい課題を有するが, SegNet は中間層の特徴マップを記憶す

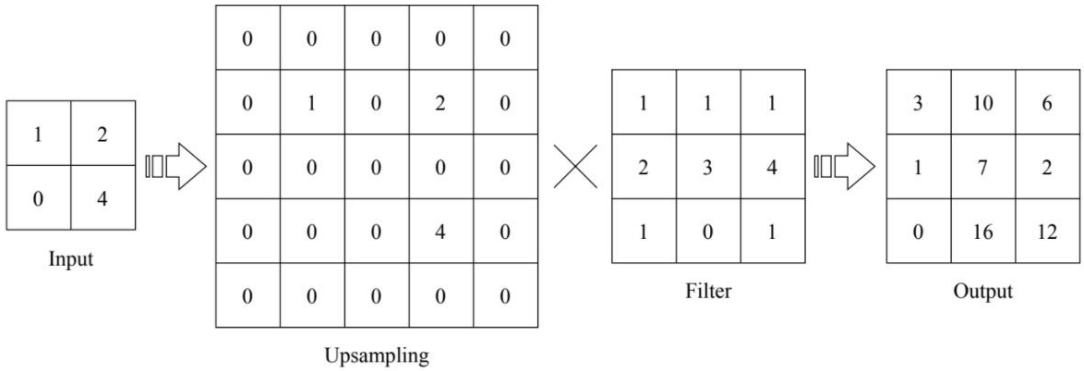


図 2.9: 逆畳み込み層の例

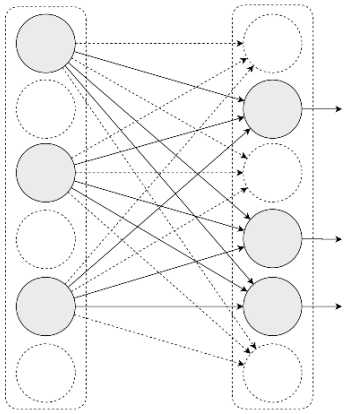


図 2.10: ドロップアウトの例

る必要がないためメモリ使用量を低減できる. SegNet のエンコーダ部は VGG-16 [60] のエンコーダ部にバッチ正規化を加えたものと等しく, FCN と同様に畳み込み層とプーリング層を用いて特徴マップの解像度を低下させる. 一方, デコーダ部はプーリング層で最大値を出力したユニットに基づいてアップサンプリングする逆プーリング層を用いて特徴マップの解像度を増大させる. 逆プーリング層を図 2.12 に例示する.

2.2.2　インタラクティブ手法

インタラクティブ手法は, 入力画像を目視確認したユーザからの手動入力に基づいて画像領域を分割する. ユーザからの手動入力の例として, 図 2.13 (a) に緑色で示す領域種別の境界を表す矩形情報が挙げられる.

以降, インタラクティブ手法として広く用いられる GrabCut [61], GrowCut [62] につい

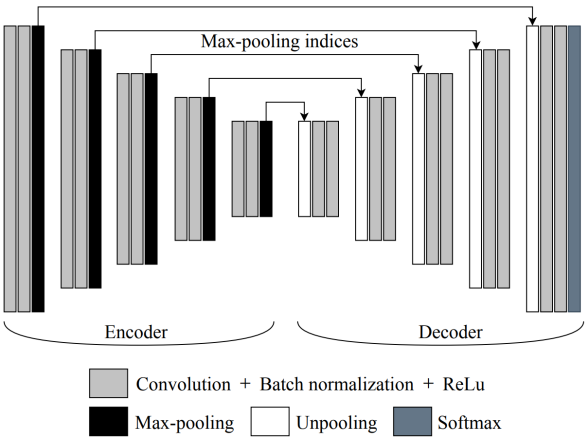


図 2.11: SegNet のネットワーク構造

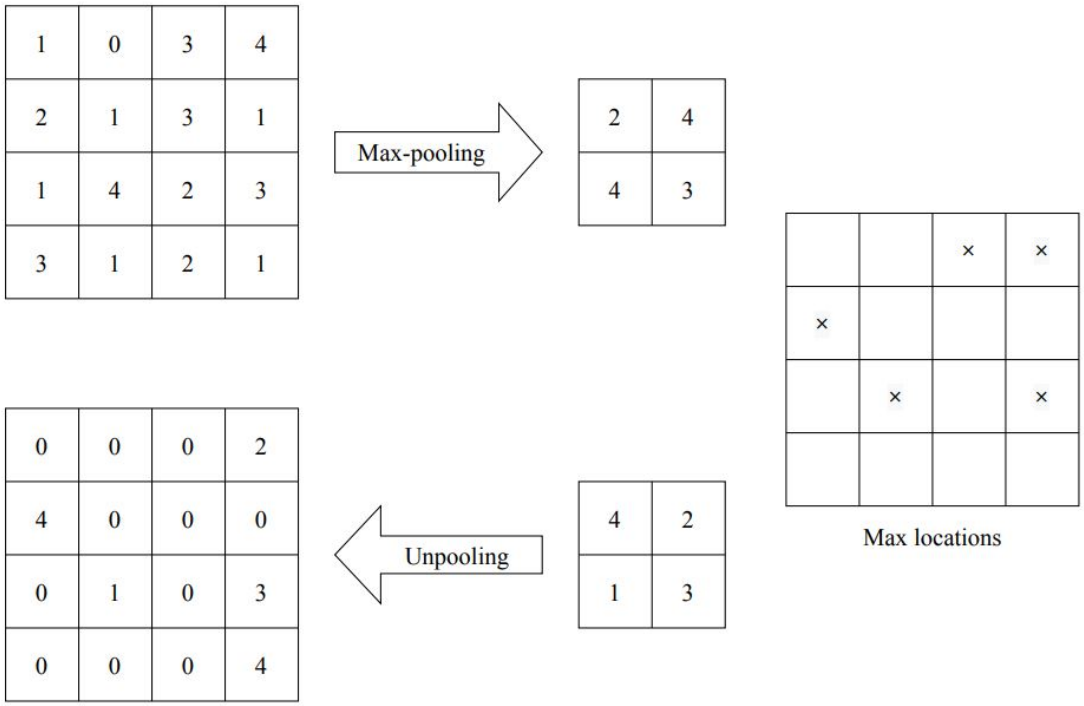


図 2.12: 逆プーリング層の例

て詳細を述べる．なお, GrabCut と GrowCut は次節で説明する事前知識に基づく手法でも用いられる．

(a) 入力データ　　(b) 領域分割結果

図 2.13: GrabCut による画像領域分割例

GrabCut

GrabCut は, ユーザの手動入力に基づいて初期化された画素値の分布と各画素をノードとしたグラフに基づいて各画素に領域種別を割り当てる. 文献 [61] では, 図 2.13 (a) に示すユーザが手動入力した矩形情報を基準に定めた前景と背景領域の候補に基づきグラフが初期化されている. 矩形の内側と外側がそれぞれ前景と背景領域の候補に分類されている.

まず, ユーザからの手動入力に基づいて入力画像の画素 p に領域種別の候補 l'_p を割り当て, 以下のエネルギー関数 E に基づいてグラフを生成する. なお, l'_p には $l_{\mathrm{fgd}}, l_{\mathrm{bgd}}$ のいずれかが代入される.

$$E = \sum_{p \in [0, P-1]} D(p) + \sum_{(p,q) \in C} S(p,q) \tag{2.8}$$

ここで, C は画素 p と画素 p に隣接する画素 q のペアを示す. 前項の $D(p)$ がデータ項, 後項の $S(p,q)$ が平滑化項を表している.

データ項の $D(p)$ は, 前景と背景領域候補の画素値分布を Gaussian Mixture Model (GMM) [63] でモデル化することで求められる. 前景と背景領域候補の画素値分布をモデル化する GMM のパラメータ群をそれぞれ $\Theta_{\mathrm{fgd}}, \Theta_{\mathrm{bgd}}$ とすると, 以下の対数尤度を最大化するパラメータ群 $\hat{\Theta}_{\mathrm{fgd}}, \hat{\Theta}_{\mathrm{bgd}}$ を求めることで前景と背景領域候補の画素値分布

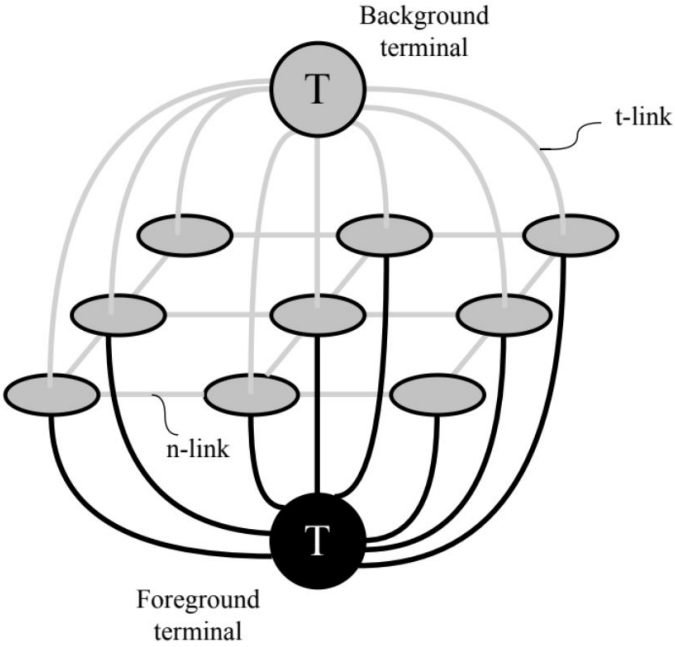


図 2.14: GrabCut で生成するグラフ

がモデル化される.

$$L(\Theta_{\mathrm{fgd}}) = \sum_{p=0}^{P-1} \delta(l'_p, l_{\mathrm{fgd}}) \cdot \log(\psi(x_p|\Theta_{\mathrm{fgd}})) \tag{2.9}$$

$$L(\Theta_{\mathrm{bgd}}) = \sum_{p=0}^{P-1} \delta(l'_p, l_{\mathrm{bgd}}) \cdot \log(\psi(x_p|\Theta_{\mathrm{bgd}})) \tag{2.10}$$

なお, δ はクロネッカーのデルタ, $\psi(x_p|\Theta_{\mathrm{fgd}}), \psi(x_p|\Theta_{\mathrm{bgd}})$ はそれぞれパラメータ群 $\Theta_{\mathrm{fgd}}, \Theta_{\mathrm{bgd}}$ を有する GMM を表す. 前景と背景領域候補の画素値分布をモデル化した GMM をそれぞれ $f_{\mathrm{fgd}} = \psi(x_p|\hat{\Theta}_{\mathrm{fgd}}), f_{\mathrm{bgd}} = \psi(x_p|\hat{\Theta}_{\mathrm{bgd}})$ とすると, データ項は以下で定義される.

$$D(p) = \begin{cases} -\log(f_{\mathrm{fgd}}(x_p)) & (l'_p = l_{\mathrm{fgd}}) \\ -\log(f_{\mathrm{bgd}}(x_p)) & (\text{otherwise}) \end{cases} \tag{2.11}$$

次に, 平滑化項の $S(p,q)$ は以下で定義される.

$$S(p,q) = \begin{cases} 0 & (l'_p = l'_q) \\ \gamma \exp(-\kappa \|x_p - x_q\|) & (\text{otherwise}) \end{cases} \tag{2.12}$$

ここで, γ, κ は定数, $\|\cdot\|$ はユークリッド距離を意味する.

GrabCut で生成するグラフを図 2.14 に例示する. 図 2.14 に示すように, GrabCut で生成するグラフは各画素を表すノードと, 前景と背景領域の頂点に相当するターミナルから構

成される．ノード間は n-link と呼ばれるエッジ, ターミナルとノード間は t-link と呼ばれるエッジで接続される．n-link と t-link の切断コストは, それぞれ (2.11) 式と (2.12) 式に示すデータ項と平滑化項に相当する．生成されたグラフは, エッジの切断コストを最小化する最小カットアルゴリズム [64] に基づいて, 前景と背景領域に分割される．

続いて, 分割された各領域の画素値分布を GMM で再度モデル化し, 再生成したグラフに最小カットアルゴリズムを繰り返し適用する．繰り返し回数が一定以上, もしくは領域分割結果が収束した場合に処理を終了する．図 2.13 (a) の矩形情報を基準に定めた前景領域と背景領域の候補に基づき初期化された GrabCut は, 図 2.13 (b) のように画像領域を分割する．

GrowCut

GrowCut は, ユーザからの手動入力に基づいて初期化されたセルオートマトンを用いて各画素に領域種別を割り当てる．規則性を持つセルから構成されるセルオートマトンは, 現在と近傍の状態から次の状態を求める計算モデルを意味する．GrowCut の初期化には, セルに相当する画素 p に対して領域種別候補 l'_p と, 領域種別の確実さを表す確信度 $g_p(\in [0,1])$ を割り当てる必要がある．

各画素 p は隣接画素 q を侵食し, 隣接画素 q は画素 p に侵食されないように以下で定義した防御力 $d(p,q)$ を用いて防御する．

$$d(p,q) = \frac{\|x_p - x_q\|}{255\sqrt{3}} \tag{2.13}$$

なお, 上式の分母は防御率の値域を [0,1] に正規化させる役割を持つ．以下を満たす場合, 画素 p が隣接画素 q の侵食に成功したとし, 隣接画素 q の領域種別を画素 p の領域種別に置き換える．

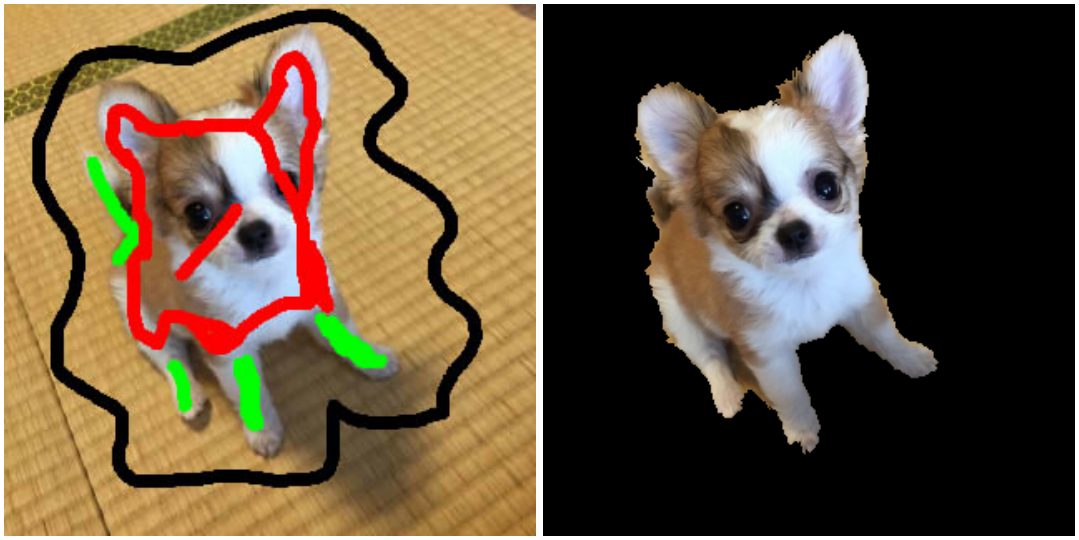
$$(1 - d(p,q))g_p > g_q \tag{2.14}$$

更に, 隣接画素 q の確信度を以下で更新する．

$$g_q = (1 - d(p,q))g_p \tag{2.15}$$

上記の侵食処理を画像全体にラスタ走査し, 領域分割結果が収束するまで繰り返す．

図 2.15 に GrowCut の適用例を示す．図 2.15 (a) は, 入力画像に対して手動で領域種別候補と確信度を割り当てた例である．赤色と緑色で示す領域を前景領域候補, それ以外の領域を背景領域候補とし, 赤色と黒色で示す領域の確信度を 1, 緑色で示す領域の確信度を



(a) 入力データ　　(b) 領域分割結果

図 2.15: GrowCut による画像領域分割例

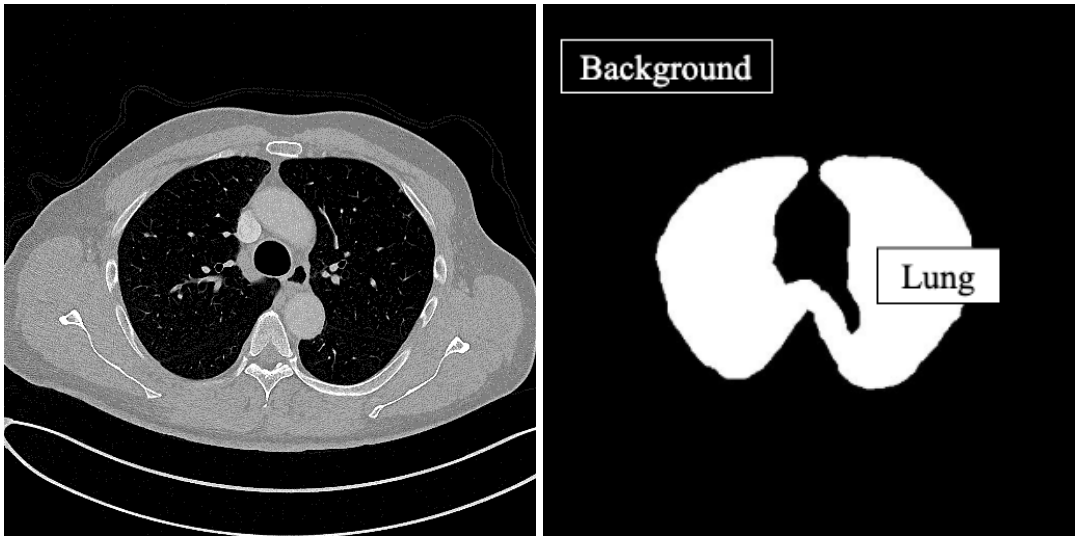
0.5, それ以外の領域の確信度を 0 とした. 図 2.15 (b) は, 図 2.15 (a) に基づいて初期化された GrowCut を適用した結果を示す.

2.2.3　事前知識に基づく手法

事前知識に基づく手法は, 入力画像や領域種別に関する知見に基づいて設計されたアルゴリズムを用いて画像領域を分割する. 事前知識に基づいて前景と背景領域の候補を求めて, 前節で説明した GrabCut や GrowCut を適用する手法が主流である. 本節では, これまで提案されてきた事前知識に基づく手法の例を挙げて概要を説明する.

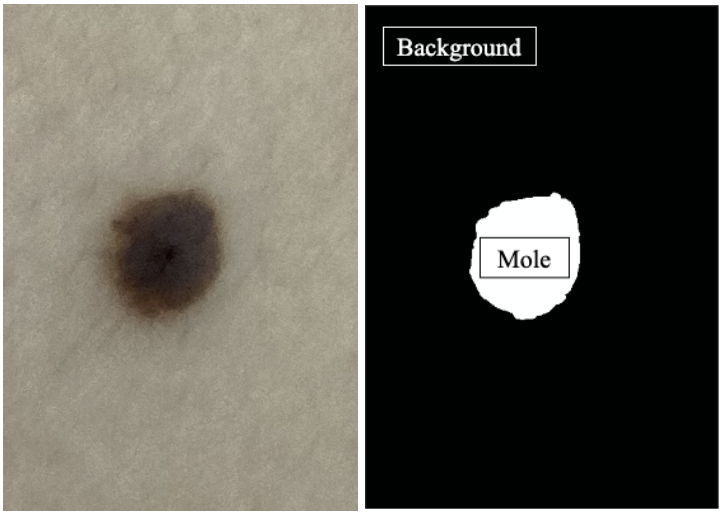
例えば, 医療の効率化を実現するために, CT 画像に対して肺の領域とそれ以外の背景領域を分割する手法を Zhang らが提案した [47]. 図 2.16 (a) の CT 画像に対して, 肺領域を分割した結果を図 2.16 (b) に示す. 図 2.16 (b) の白色と黒色に示す領域はそれぞれ肺領域と背景領域である. CT 画像において, 肺領域は左右で 2 つ存在することや, 肺領域の輝度は背景領域と比べて低いことが事前知識として挙げられる. まず, Zhang らの手法はこの事前知識に基づいて入力画像に大津の 2 値化 [65] を適用し, 2 つの連結要素を抽出する. 続いて, 抽出された 2 つの連結要素を基準に定めた肺と背景領域の候補で初期化した GrabCut を適用する.

また, 人間の肌を撮影した皮膚画像に対して, 黒子の領域とそれ以外の背景領域を分割す



(a) CT 画像　　(b) 肺領域分割結果

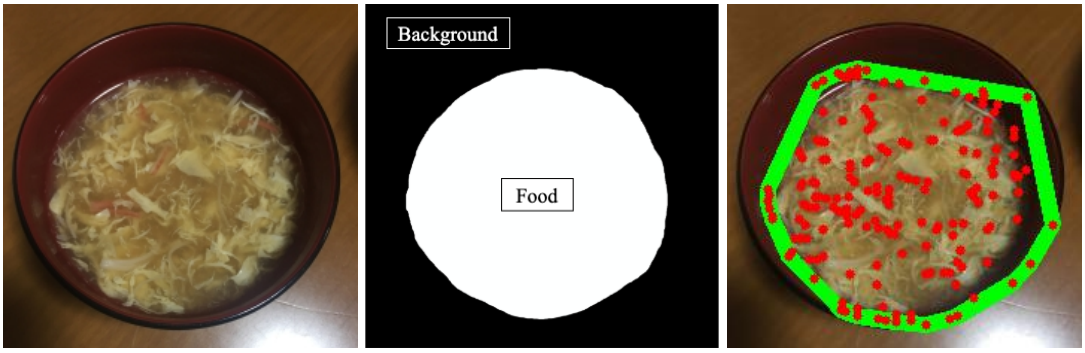
図 2.16: CT 画像に対する肺領域分割



(a) 皮膚画像　　(b) 黒子領域分割結果

図 2.17: 皮膚画像に対する黒子領域分割

る手法を Tushar らが提案した [48]. 図 2.17 (a) の皮膚画像に対して, 黒子領域を分割した結果を図 2.17 (b) に示す. 図 2.17 (b) の白色と黒色に示す領域はそれぞれ黒子領域と背景領域である. 皮膚画像において, 黒子領域の輝度値は背景領域と比べて低いことが事前知識として挙げられる. まず, Tushar らの手法はこの事前知識に基づいてヒストグラムが正規化された入力画像の輝度値を 2 値化し, 輝度値の低い領域を求める. 続いて, 2 値化の結



(a) 食事画像　(b) 食事領域分割結果　(c) 輝度極値点と凸包

図 2.18: 食事画像に対する食事領域分割

果を基準に定めた黒子と背景領域の候補で初期化した GrabCut を適用する.

他にも, 食事を効率的に記録することを目的に撮影した食事画像に対して, 食事の領域とそれ以外の背景領域を分割する手法を杉山らが提案した [49]. 図 2.18 (a) の食事画像に対して, 食事領域を分割した結果を図 2.18 (b) に示す. 図 2.18 (b) の白色と黒色に示す領域はそれぞれ食事領域と背景領域である. 食事画像において, 被写体の食事領域は背景領域と比べて注目を集めるように撮影されることが事前知識として挙げられる. まず, 杉山らの手法はこの事前知識に基づいて人の視線を集める度合（誘目度） [66] が高い領域を, 輝度値の極値点を全て含む面積が最小の多角形（凸包）として求める. 図 2.18 (c) の赤点が輝度の極値点で, 全ての赤点を囲う緑線が凸包である. 続いて, 凸包の内部と外部をそれぞれ食事と背景領域の候補とし, GrabCut を適用する.

2.2.4　利点と課題

本節では, 2.2.1～2.2.3 節で説明した従来手法の利点と課題を整理する. 2.2.1 節で説明した局所特徴に基づく手法は, 処理対象となる入力画像や領域種別に応じた学習データを用意することで様々な分野に適用できる利点を有する. 例えば, SegNet は導入先に応じた学習データを用意することで医療, 車載, 防災, 食品の分野に応用されている [67–70]. しかし, 領域分割精度を向上させるには学習データ量を増加させる必要がある. 特に, FCN や SegNet のようなネットワークは画素単位で領域種別が付与された学習データが必要なため, 導入コストの増大が課題として挙げられている.

図 2.19: GrabCut の適用事例（左：入力画像, 右：領域分割結果）

2.2.2 節で説明したインタラクティブ手法は, 入力画像を目視確認したユーザからの手動入力に基づいて領域種別を分類する. インタラクティブ手法はユーザからの適切な手動入力が得られた場合に, 図 2.19 に示すように様々な入力画像に対して前景と背景領域を高精度に分割できる利点を有する. しかし, インタラクティブ手法にはユーザからの手動入力が必要なため, ユーザビリティが低下する点に課題がある.

2.2.3 節で説明した事前知識に基づく手法は, 局所特徴に基づく手法やインタラクティブ手法で発生した導入コスト増大やユーザビリティ低下を回避可能な利点を有する. 入力画像や領域種別に関する知見に基づいてアルゴリズムが設計されるため, アルゴリズムが処理目的に強く依存する点に課題がある.

2.3　研究方針

前節では, 画像領域分割の従来手法に関する概要を代表例に基づいて説明し, 各手法の利点と課題を整理した. 本論文では, 導入コスト増加やユーザビリティ低下を回避して実用化できる事前知識に基づく手法の高度化が研究目的である. 入力画像や領域種別に関する知見に基づいて設計される事前知識に基づく手法は, アルゴリズムが処理目的に依存する. そこで, 次章以降は導入先となる応用システムの特性を加味して事前知識に基づく手法を高度化する研究成果を述べる. 以降, 詳細な説明は各章に委ねるが, 次章以降の内容に関連する建造物領域分割, 商品領域分割, 差分領域分割の研究動向を述べ, 研究方針の概要を説明する.

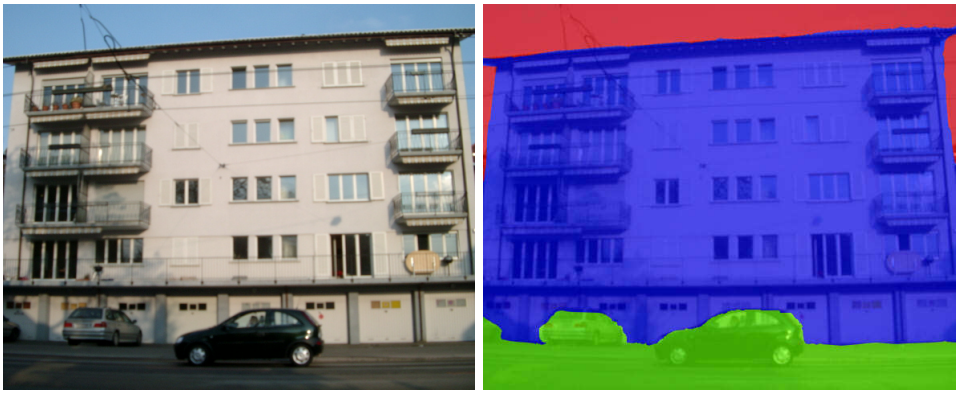
図 2.20: 情景画像の例

2.3.1　建造物領域分割

本研究では, ユーザが周辺を撮影した情景画像の建造物領域を事前知識に基づいて分割する手法の高度化を目的とする. 図 2.20 に情景画像を例示する. 情景画像に関する事前知識の例として, 被写体の建造物は画像の中心部に分布することや, 上空や道路などの背景は画像の上半分と下半分にそれぞれ分布することが挙げられる. この事前知識と一致する例を図 2.21 に示す. 図 2.21 (b) は, 図 2.21 (a) の建造物領域と背景領域をそれぞれ青色と青色以外で示した画像である. 背景領域に相当する上空と道路をそれぞれ赤色と緑色で示す. 建造物領域が画像の中心部に, 背景領域の上空と道路がそれぞれ画像の上半分と下半分に分布している.

Åhlen らは, 上空が青いことや, 建造物の下に道路が存在することを事前知識として建造物領域を分割する手法を提案した [71]. Åhlen らの手法では, まず, 青色成分の弱い画素を抽出することで上空の可能性が低い領域（以降, 非上空領域と呼称する）を求める. 次に, 画像の下半分に含まれる道路の可能性が高い非上空領域を背景領域と仮定して建造物領域を分割する. 一方, 文献 [72] で上野らは GrowCut を応用した手法を提案し, Åhlen らの手法と比べて 40% 以上の領域分割精度向上を達成した. 上野らの手法では, 情景画像の中心部に建造物領域, 上半分に上空の領域が分布することや, 被写体の建造物は比較的大きく撮影されることを事前知識とする.

しかし, 上野らの手法には領域分割精度の低下を引き起こす複数の課題が存在することが分かった. そこで, 本研究では上野らの手法で発生した課題を回避可能な建造物領域分割手法を提案し, その有効性を示す.



(a) 情景画像　　(b) 領域種別（青: 建造物, 赤・緑: 背景）

図 2.21: 情景画像と領域種別

2.3.2　商品領域分割

本研究では, インターネットオークションやフリーマーケットで用いられる商品画像の商品領域を事前知識に基づいて分割する手法の高度化を目的とする. 図 2.22 に商品画像を例示する. 商品画像に関する事前知識の例として, 被写体となる商品は背景と比較して人の注目を集めるように撮影されることや, 画像の中心部に撮影されることが挙げられる.

商品領域を事前知識に基づいて分割する手法を提案し, その精度を評価した関連研究は存在しない. そこで, 商品クラス認識で用いられている商品領域分割に着目すると, 誘目度が高い領域を分割する Cheng らの手法 [73] が文献 [74] で応用されている. これは人の注目を集めるように撮影される商品領域の誘目度が背景領域と比較して高いことを事前知識としている.

しかし, Cheng らの手法には領域分割精度の低下を引き起こす課題が存在することが分かった. そこで, 本研究では Cheng らの手法で発生した課題を回避可能な商品領域分割手法を提案し, その有効性を示す.

2.3.3　差分領域分割

本研究では, 監視カメラで撮影した画像の差分領域を事前知識に基づいて分割する手法の高度化を目的とする. 図 2.23 に監視カメラ画像を例示する. 監視カメラ画像に関する事前知識の例として, 現在と過去に撮影した画像間で画素値の差分が大きい領域に差分領域

図 2.22: 商品画像の例

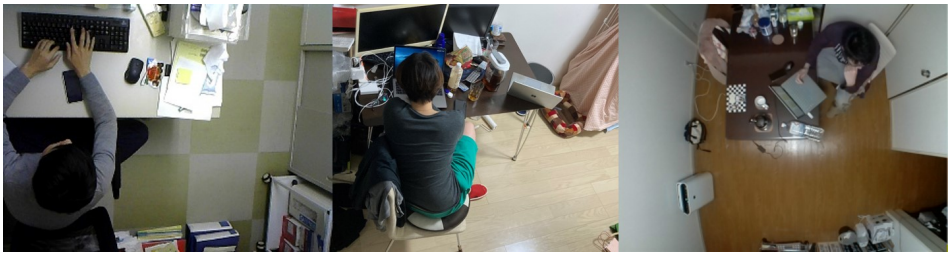


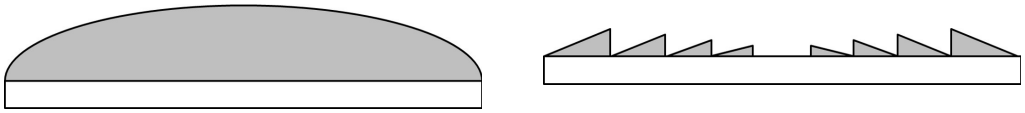
図 2.23: 監視カメラ画像の例

が含まれることが挙げられる.

近年, 監視カメラを内蔵し, 撮影した画像に画像認識技術を駆使することで在室者を検知する人感センサが商用化されている [75, 76]. 例えば, CNN に基づく物体検出を応用することで在室者の検知精度向上を見込めるが, 商用化が進められている人感センサは計算コストや導入コストの低い差分領域分割をベースとしている [75]. 在室者が動作することで現在と過去に撮影した画像の画素値間に差分が発生することを前提としている.

現在と過去に撮影した画像間で差分が閾値以上の画素を, 在室者による動きを含む差分領域に分類し, 差分領域が占める割合が一定以上の場合に在判定する手法を Sarker らが提案している [77]. Sarker らの手法は, 文献 [78] で低消費電力で安価なシングルボードコンピュータに実装されているため, 一般家庭向けにも広く導入が可能と考えられる. そこで, 商用化されている差分領域分割に基づく人感センサのアルゴリズムとして Sarker らの手法を想定し, 従来の人感センサ技術との比較実験に基づきその有効性を評価する. 比較実験の結果に基づいて差分領域分割に基づく人感センサの利点と課題を整理し, 差分領域分割に基づく人感センサの高度化に向けた改良課題を考察する.

在室者を高精度に検知できる人感センサ技術として, ユーザが携行する ID カードや専用アプリケーションを導入したスマートフォンを利用する手法が存在する [79, 80]. しかし, ユーザは常に ID カードやスマートフォンを携行する必要がある点で制約が生じる. そ



(a) 凸レンズ　(b) フレネルレンズ

図 2.24: レンズの断面図

こで, 本研究ではユーザの携行品に制約を与えることなく, いずれの在室者も検知することが可能な人感センサ技術を差分領域分割に基づく人感センサの比較対象とする.

いずれの在室者も検知することが可能な人感センサ技術として, 室内の二酸化炭素濃度に基づく CO_2 センサや, 受信した超音波の波長に基づく超音波センサが存在する [81]. しかし, 赤外線センサが安価で最も普及しているため [82], 差分領域分割に基づく人感センサの比較対象とする.

赤外線センサに関する関連研究に着目すると, 赤外線センサを 2 次元状に並べた赤外線センサアレイに関する研究が進められていることが分かる [83, 84]. 赤外線センサアレイは, 単画素の赤外線センサと比較して在室者の検知精度が高く, 在室者が存在する位置を求めることが可能だが, 現在普及している赤外線センサは安価な単画素式が主流である. そこで, 本研究では単画素の赤外線センサを差分領域分割に基づく人感センサの比較対象とする. 以降, 単画素の赤外線センサを説明する.

赤外線センサと凸レンズの断面図を図 2.24 で比較する. 赤外線センサは, 軽量化や低コスト化を目的に凸レンズ (図 2.24 (a)) の一部を切り出したフレネルレンズ (図 2.24 (b)) を用いる. フレネルレンズによって広げられた視野内の人体が発する赤外線量の変化に基づき在室者を検知する.

フレネルレンズを用いることで赤外線センサの低コスト化を実現できるが, 検知範囲内に赤外線センサが在室者の動作を検知できない隙間（以降, Gap と呼称する）が発生する. 図 2.25 に赤外線センサの検知範囲と Gap を例示する. 文献 [85] では, Gap が原因で赤外線センサは在室者の細かな動作に対して正しく在判定できないことが指摘されている. 例えば, 複数の赤外線センサに対して検知精度を調査した文献 [86] において, 在室者の腕の動作をロボットアームで模擬した結果, 正しく在判定できた割合の中央値は 35% であった.

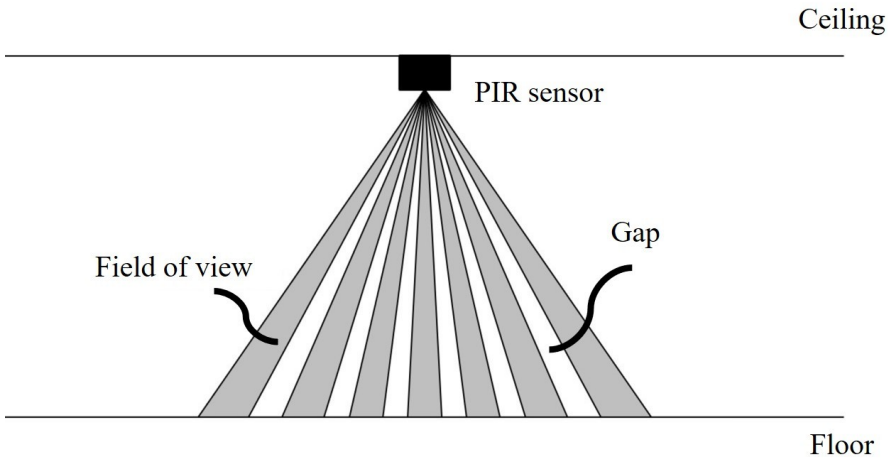


図 2.25: 赤外線センサの検知範囲（PIR sensor: 赤外線センサ）

2.4　まとめ

本章では, 入力画像の各画素に領域種別を割り当てる画像領域分割の従来手法を関連研究に基づいて説明した. まず, 次章以降で研究成果を述べる建造物領域分割, 商品領域分割, 差分領域分割を例に画像領域分割を定式化した. 続いて, 画像領域分割の従来手法を, 局所特徴に基づく手法, インタラクティブ手法, 事前知識に基づく手法に分類し, 各手法を代表例に基づき説明した. 局所特徴に基づく手法の代表例として, CNN に基づく FCN と SegNet を挙げた. インタラクティブ手法の代表例として, グラフ理論に基づく GrabCut とセルオートマトンに基づく GrowCut を挙げた. 事前知識に基づく手法の代表例として, 肺領域を分割する Zhang らの手法, 黒子領域を分割する Tushar らの手法, 食事領域を分割する杉山らの手法を挙げた. 次に, 各手法の利点と課題を整理して, 研究対象とする事前知識に基づく手法とその他の手法の差異を明確に説明した. 更に, 本論文に関連する建造物領域分割, 商品領域分割, 差分領域分割の研究動向を述べ, 次章以降で述べる研究の方針を説明した.

第 3 章

色クラスタ解析に基づく建造物領域分割手法

3.1 はじめに

近年, スマートフォンの普及とともに Global Navigation Satellite System (GNSS) による現在位置の測位が比較的容易に行えるようになった. しかし, 電波遮蔽物の多い都市部において測位誤差が増大する課題があり [87, 88], GNSS を利用することなく都市部における現在位置を高精度に測位できるシステムが求められている. そこで, 文献 [89, 90] で現在位置の測位に応用が可能な情景画像中に存在する建造物をクラス認識する技術が研究されている.

Åhlen らは建造物クラス認識の前処理として, 2.1 節で例示した建造物領域分割を適用し, 建造物の輪郭線を抽出する処理を導入している [71]. 文献 [71] において, 実験的に建造物領域の分割精度が認識性能に影響することが示唆されている. 以上より建造物クラス認識の性能改善に向けて, 情景画像の建造物領域を分割する技術の高精度化が求められていることが分かる. また, 建造物領域を分割する技術は建造物の外観検査にも応用が可能で, 外観検査の効率・省人化を実現するためにも高精度が求められる.

そこで, 本研究では情景画像の建造物領域を分割する手法の高性能化を目的とする. この目的を実現するために, 入力画像の各画素を色情報でクラスタリングし, 背景領域を含む割合の高いクラスタを判定する処理を導入した建造物領域分割手法を提案する. 続いて, 2.3.1 節で述べた上野らの手法との比較実験に基づいて領域分割精度と計算時間の観点か

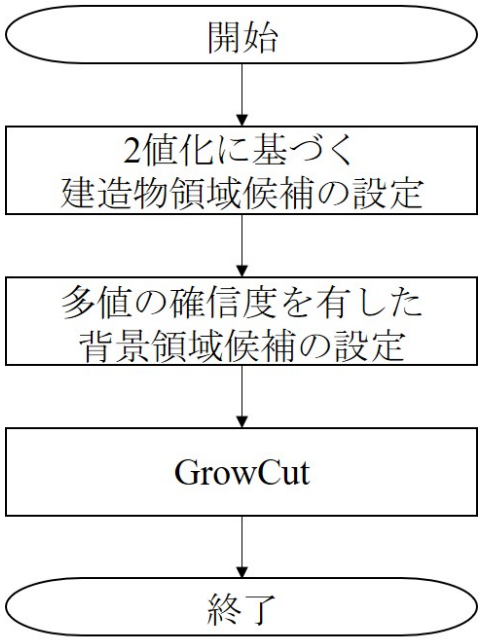


図 3.1: 上野らの手法のフローチャート

ら提案手法の有効性を示す．なお, 本節では建造物が被写体として正立して撮影された情景画像が入力されることを前提とする．

以降の内容について説明する．3.2 節で上野らの手法の各処理を詳細に述べ, 領域分割精度低下を引き起こす上野らの手法の課題をまとめる．3.3 節で建造物領域を分割する提案手法についてフローチャートを用いて各処理を詳細に述べる．3.4 節で情景画像を用いた比較実験に基づいて提案手法の有効性や課題を考察し, 本節の内容を 3.5 節でまとめる．

3.2　上野らの手法

本節では, 図 2.20 に例示する情景画像の建造物領域を分割する上野らの手法について説明する．上野らの手法は, 情景画像の中心部に建造物領域, 上半分に上空の領域が分布することや, 被写体の建造物は比較的大きく撮影されることを事前知識としている．

3.2.1　処理概要

上野らの手法は, 情景画像に対して建造物と背景領域の候補や確信度を求めて, GrowCut を適用する．上野らの手法のフローチャートを図 3.1 に示す．以降, 図 3.1 の各項目を詳細に述べる．

(a) 入力画像　(b) 2 値化画像 ($\theta_b = 105$)　(c) 2 値化画像 ($\theta_b = 158$)

図 3.2: 2 値化画像

2 値化に基づく建造物領域候補の設定

GrowCut を初期化するために必要な建造物領域候補を求める．建造物領域候補は, 平均隣接数 [91, 92] の増加率に基づいて輝度値を 2 値化した画像を基準に設定する．

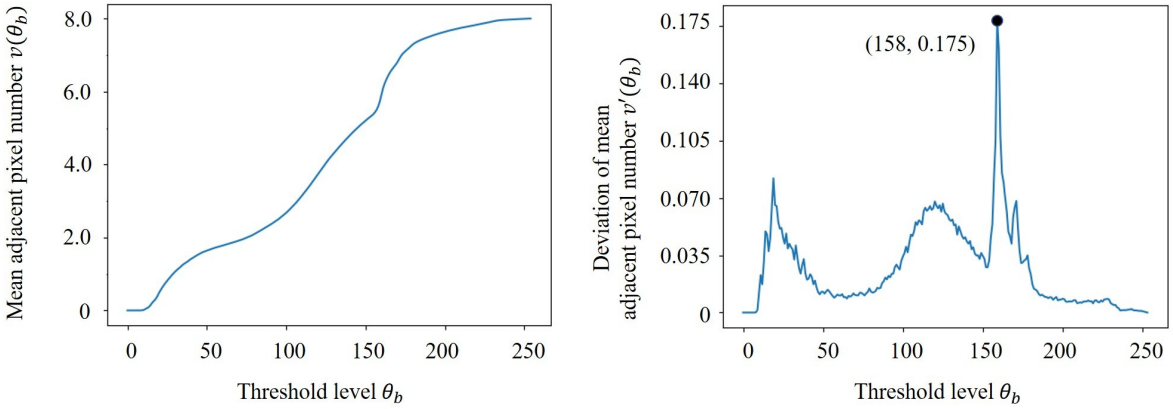
まず, 入力画像を平均隣接数に基づいて 2 値化する．平均隣接数とは, 2 値化画像における黒画素の 8 近傍に存在する黒画素数の平均値を意味し, 連結成分のまとまりの良さを意味する．図 3.2 (a) に示す入力画像を閾値 105, 158 で 2 値化した結果を図 3.2 (b), (c) に示す．図 3.2 (b), (c) の平均隣接数はそれぞれ 2.89, 5.27 で, 平均隣接数の多い図 3.2 (c) は図 3.2 (b) と比べて建造物がまとまって黒画素に分類されていることが分かる．

建造物が黒画素でまとまって得られる閾値付近では, 以下で定義する平均隣接数の増加率が増大する傾向がある．

$$v'_{\theta_\mathrm{b}} = v_{\theta_\mathrm{b}+1} - v_{\theta_\mathrm{b}} \tag{3.1}$$

なお, v_{θ_b} は閾値 θ_b で 2 値化された画像の平均隣接数を示す．この傾向と被写体の建造物は比較的大きく撮像される事前知識に基づいて, 上野らの手法は全画素に対して黒画素が占める割合が 1/3 以上で平均隣接数の増加率が極大となる閾値を採用する．なお, 文献 [72] において平均隣接数の増加率が極大となる閾値が複数存在する場合の選択方法について言及されていないため, 本研究では平均隣接数の増加率が最大となる閾値を採用する．

図 3.2 (a) に示す入力画像に対して, 平均隣接数と平均隣接数の増加率を求めた結果をそれぞれ図 3.3 (a), (b) に示す．なお, 全画素数に対して 2 値化画像の黒画素数の占める割合が 1/3 以上となる閾値は 105 以上であった．図 3.3 (b) を確認すると, $\theta_b \geq 105$ で平均隣接数の増加率が極大値をとる閾値は複数存在することが分かる．平均隣接数の増加率が最大となる閾値を選択すると閾値 158 が採用される．



(a) 平均隣接数　　(b) 平均隣接数の増加率

図 3.3: 平均隣接数と平均隣接数の増加率

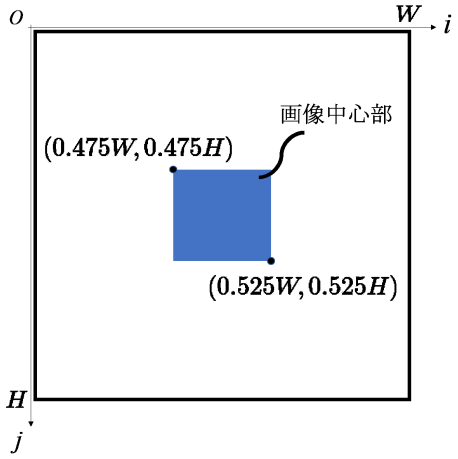
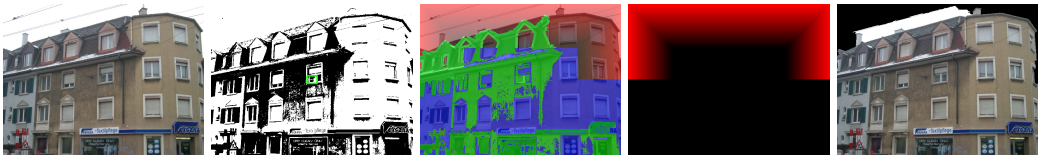


図 3.4: 画像中心部

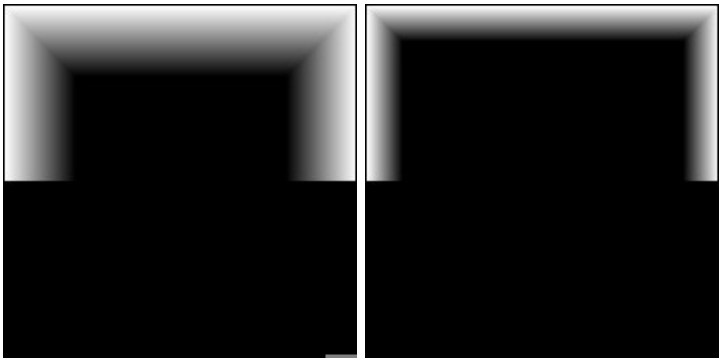
続いて, 画像の中心部に建造物領域が分布する事前知識に基づいて, 幅と高さの画素数がそれぞれ $0.05W, 0.05H$ の画像中央部に連結する黒画素成分を建造物領域候補とする. 画像中央部を図 3.4 に例示する. 次に, 建造物領域を背景領域に誤分類する割合を低減するために, 建造物領域候補にモルフォロジー演算の一種である膨張処理を適用する.

図 3.5 (a) の入力画像に平均隣接数に基づく 2 値化を適用することで, 図 3.5 (b) に示す 2 値化画像が得られる. 図 3.5 (b) の画像中心部に連結する黒画素成分を膨張させることで, 図 3.5 (c) に緑色で示すような建造物領域候補が求められる.



(a) (b) (c) (d) (e)

図 3.5: 上野らの手法による建造物領域分割過程



(a) $\alpha = 10$ (b) $\alpha = 20$

図 3.6: 背景領域の確信度

多値の確信度を有した背景領域候補の設定

GrowCut を初期化するために必要な背景領域候補と確信度を求める．まず, 背景領域候補の確信度を求める．画素位置 (i, j) における確信度 $g_{i,j}$ は以下で設定される．

$$g_{i,j} = \begin{cases} \max\left(1 - \frac{\alpha}{W+H} i, 1 - \frac{\alpha}{W+H}(W - i), 1 - \frac{\alpha}{W+H} j\right) & (j < H/2) \\ 0 & (\text{otherwise}) \end{cases} \tag{3.2}$$

なお, α は確信度を多値化するときの量子化レベルを決定するパラメータである．

図 3.6 に (3.2) 式で求められる背景領域候補の確信度を可視化する．なお, 確信度の値の大きさを白色成分の強度で示している．図 3.6 を確認すると, (3.2) 式は被写体の建造物が画像中心部に分布する事前知識に基づいて画像の外縁部から中心部に向かって背景領域候補の確信度の値を低下させていることが分かる．図 3.6 (a), (b) を比較すると, 図 3.6 (a) は画像の中心部に向かって確信度の値が滑らかに低下している．(3.2) 式の α は確信度の値を低下させる割合を調整するために用いられることが分かる．

次に, 背景領域候補を設定する. 建造物領域候補に設定されていない $g_{i,j} > 0$ となる画素を背景領域候補に設定する.

以上の処理で, 図 3.5 (c) のように建造物と背景領域の候補が求められる. なお, 図 3.5 (c) の緑色は建造物領域候補, 赤色は背景領域候補, 青色はいずれにも該当しない未設定の領域を示している. 図 3.5 (d) は, (3.2) 式で求められる背景領域候補の確信度を示す. なお, 確信度の値の大きさを赤色成分の強度で示している. 図 3.5 (c) で背景領域候補に設定された画素にのみ背景領域候補の確信度を設定する.

GrowCut

前節の処理で求められた建造物と背景領域の候補, 確信度に基づいて初期化された GrowCut を適用する. 図 3.5 (c), (d) に示す建造物と背景領域の候補, 背景領域候補の確信度で初期化された GrowCut を適用することで, 図 3.5 (e) のように建造物領域が分割される.

3.2.2　課題

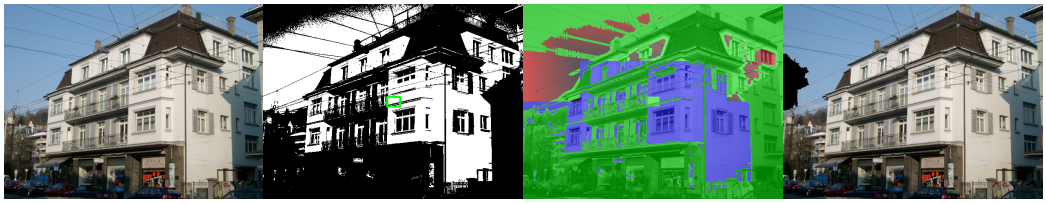
上野らの手法を ZuBuD [93] に収録されている 5 つの視点から撮像した 201 種類の情景画像に適用したところ, 図 3.7 に示す領域分割精度が低下する事例を確認した. 図 3.7 では, 各事例に対して左から 1 番目が入力画像, 2 番目が 2 値化画像, 3 番目が建造物と背景領域の候補設定, 4 番目が建造物領域分割結果を示している.

まず, 図 3.7 (a) の事例 1 では, 木の領域を建造物領域候補に誤分類する閾値が得られていることが原因で, 背景領域候補に建造物領域が広範囲に含まれている. その結果, GrowCut の適用後に建造物領域を背景領域に誤分類する割合が高くなる.

次に, 図 3.7 (b) の事例 2 では, 背景に相当する上空の領域が画像中心部に連結している. そのため, 建造物領域候補に上空の領域が広範囲に含まれており, GrowCut の適用後に上空の領域を建造物領域に誤分類する割合が増加した. この課題は, 空領域の輝度が建造物領域より低い傾向にある場合に頻出し, 天候や時間帯に大きく依存すると考えられる.

続いて, 図 3.7 (c) の事例 3 では, 2 値化で建造物領域を建造物領域候補に正しく分類できているものの, 建造物領域候補の設定時に行われる膨張処理が原因で, 上空の領域が建造物領域候補に誤分類されている. 上野らの手法では建造物領域候補の確信度を最大値に設定するため, GrowCut によって建造物領域候補が背景領域に分類されることはなく, 上空

(a) 事例 1



(b) 事例 2

(c) 事例 3

図 3.7: 上野らの手法による建造物領域分割精度低下の例

の領域を建造物領域に広範囲に誤分類した結果が得られている.

上記で述べた領域分割精度が低下する事例において, 建造物と背景領域の候補を設定する処理が有する課題を以下にまとめる.

(i)　2 値化後の黒画素が画像中心部に連結していないことや建造物領域が黒画素として得られないことが原因で, 背景領域候補に建造物領域を誤分類する割合が高くなる.

(ii)　建造物領域と比較して背景領域の輝度が低いことが原因で, 建造物領域候補に背景領域を高確信度で誤分類する割合が高くなる.

(iii)　建造物領域候補の膨張処理が原因で, 背景領域を建造物領域候補に高確信度で誤分類する割合が高くなる.

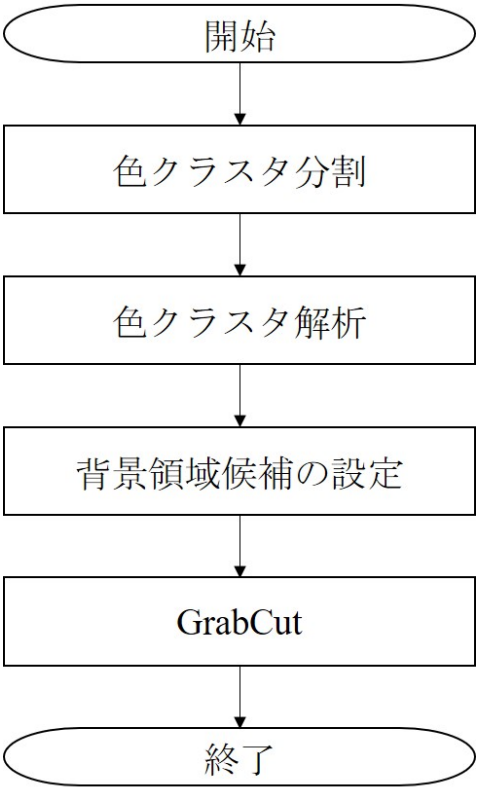


図 3.8: 提案手法のフローチャート

3.3　提案手法

本節では, 情景画像の建造物領域を分割する手法を提案する. 提案手法も, 上野らの手法と同様に画像の中心部に建造物, 上部や下部にそれぞれ空や道路などの背景が分布することを事前知識として用いる. 提案手法は, 入力画像の各画素を色情報に基づいてクラスタリングし, 色クラスタを解析することで画像の上部や下部に分布する傾向のあるクラスタを抽出する. 次に, 色クラスタの解析結果に基づいて建造物と背景領域の候補を求めて GrabCut を適用する. 以降, 提案手法の処理概要を説明する.

3.3.1　処理概要

提案手法のフローチャートを図 3.8 に示す. 以降, 図 3.8 の各項目を詳細に述べる.

色クラスタ分割

入力画像を構成する各画素を画素値に基づいてクラスタリングする．まず, 入力画像の解像度を $\rho(\in [0,1])$ 倍に低下させ, 幅と高さの画素数がそれぞれ ρW, ρH で総画素数が $P'(=\rho^2 WH)$ の縮小画像 $X' = \left\{x'_0, \ldots, x'_{P'-1}\right\}$ を生成する．なお, 提案手法では画素値として RGB 値を用いる．

続いて, Variational Bayesian Gaussian Mixture (VBGMM) [94, 95] に基づくクラスタリングを適用し, 縮小画像の画素値を混合数 m_{G} の GMM$(=\psi(x|\Theta_{m_{\mathrm{G}}}))$ でモデル化する．なお, $\Theta_{m_{\mathrm{G}}}$ は GMM のパラメータ群を意味する．

VBGMM に基づくクラスタリングは, まず, 画素値の分布を GMM でモデル化する．縮小画像の画素値 x'_p を用いて, 以下の対数尤度 $L(\Theta_{m_{\mathrm{G}}})$ を最大化するパラメータ群 $\hat{\Theta}_{m_{\mathrm{G}}}$ を求める．

$$L(\Theta_{m_{\mathrm{G}}}) = \sum_{p=0}^{P'-1} \log(\psi(x'_p|\Theta_{m_{\mathrm{G}}})) \tag{3.3}$$

続いて, 入力画像に応じて適切な混合数 m_{G} を自動設定するために, 以下のベイズ情報量規準 (BIC(m_{G})) [96] を最小化する混合数 $\hat{m}$ を採用する．

$$\mathrm{BIC}(m_{\mathrm{G}}) = -L(\hat{\Theta}_{m_{\mathrm{G}}}) + \frac{9m_{\mathrm{G}} - 1}{2}\log(P') \tag{3.4}$$

(3.3) 式の対数尤度は, 混合数 m_{G} の増加とともに増大する傾向がある．しかし, ベイズ情報量規準は (3.4) 式の 2 項目のように混合数 m_{G} に応じた罰則を加えて画素値の分布をモデル化した GMM を評価できるため, 入力画像に応じて混合数を自動で設定することができる．なお, $9m_{\mathrm{G}} - 1$ は RGB 空間を混合数 m_{G} の GMM でモデル化した場合の自由パラメータ数を表す．

上記処理に基づいて, VBGMM は画素値分布を混合数 $\hat{m}$ の GMM でモデル化する．次に, 入力画像の各画素を GMM に入力し, 出力確率が最大となるガウス分布を求めることで, 各画素にクラスタ番号を割り当てたクラスタリング結果 $M = \{m_0, \ldots, m_{P-1}\}$ $(m_p \in [0, \hat{m} - 1])$ を求める．

GMM に基づくクラスタリングは混合数をパラメータとして入力する必要があるが, 建造物が塗装されている色の種類や時刻で背景が変動する入力画像に対して, 適した混合数を入力することは難しい．一方, VBGMM は (3.4) 式に基づいて GMM の混合数を自動で設定できる利点がある．VBGMM は画素値の分布を GMM でモデル化するため, ユーク

(a) (b) (c) (d) (e)

図 3.9: 提案手法による建造物領域分割過程

リッド距離を使用する k-means 法 [97] とは異なり, 楕円の長軸方向に長いガウス分布が隣接しているような観測データに対しても良好なクラスタリング結果を期待できる [98].

図 3.9 (a) の入力画像に以上の処理を適用することで, 図 3.9 (b) のように各クラスタを異なる色で示した色クラスタリング結果が得られる.

色クラスタ解析

建造物が被写体の情景画像は地上から見上げて撮影されるため, 画像の上部に空, 下部に歩行者や道路などの背景領域が存在する傾向がある. 前節において得られたクラスタリング結果に着目しても, 背景領域が占める割合が比較的高いクラスタは画像の上部や下部に分布する傾向が確認できる.

上記の傾向に基づいて画像の上部に分布するクラスタを抽出するために, 画像上端から中心に向かって図 3.10 のように一定の画素幅で移動する矩形 $R_{n'}$ $(0 \leq n' \leq N' - 1)$ における, クラスタ m' の含有率 $(\zeta_{n'}^{m'})$ を以下の式で求める.

$$\zeta_{n'}^{m'} = \frac{\sum_{p \in R_{n'}} \delta(m_p, m')}{W_R \cdot H_R} \tag{3.5}$$

なお, δ はクロネッカーのデルタ, W_R と H_R は矩形 $R_{n'}$ の幅と高さ, m_p は前節で得られた画素 p におけるクラスタ番号を意味する. 続いて, 画像の下部に分布するクラスタを抽出するために画像を 180° 回転させて同様に含有率を求める.

ここで背景領域を含む割合の高いクラスタの含有率を分析し, 次節で背景領域候補を決定する際の条件に反映させる. 図 3.11 (a), (b) に画像の上半分, 下半分を対象に求めた各クラスタの含有率 $\zeta_{n'}^{m'}$ を示す. なお, 図 3.11 におけるグラフの色は図 3.9 (b) における色と対

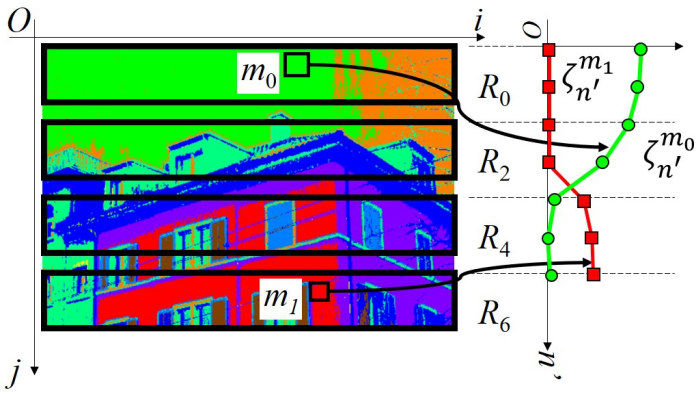


図 3.10: 提案手法による色クラスタ解析の概要

応している．

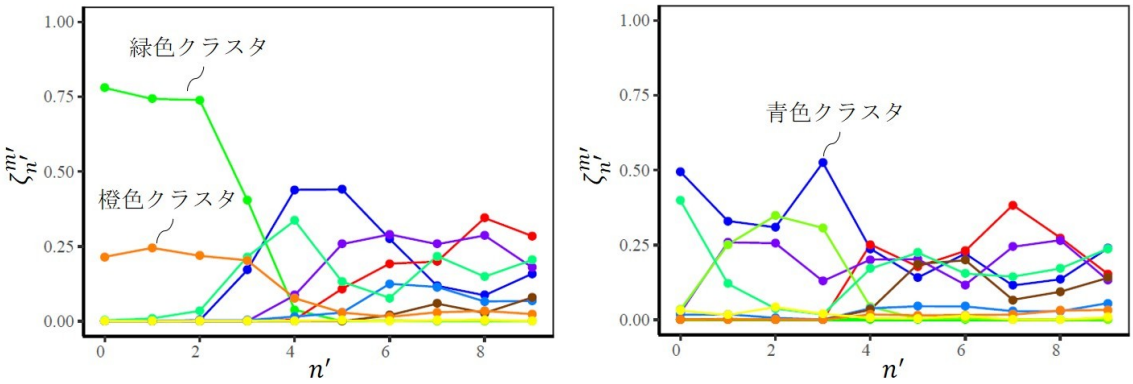
まず, 図 3.11 (a) に着目する．背景領域を含む割合の高いクラスタとして, 上空や木を多く含む図 3.9 (b) における緑, 橙色クラスタを分析する．橙色クラスタの含有率は $n' = 1$ で僅かに増加する傾向があったが, 緑, 橙色クラスタの含有率は n' が増加するにつれて単調減少する傾向がある．その他の画像における背景領域を含む割合の高いクラスタにも, 同様の傾向が確認された．そこで, 次節で背景領域を含む割合の高いクラスタを決定する際の条件に上述した傾向を反映させる．

次に, 図 3.11 (b) に着目する．背景領域を含む割合の高いクラスタとして, 道路や車を多く含む図 3.9 (b) における青色クラスタを分析する．青色クラスタは図 3.11 (a) の緑, 橙色クラスタのように単調減少する傾向はなく, $n' = 2$ で極小値を取る．これは青色クラスタが背景領域と建造物領域の一部から構成されていることに起因している．その他の画像においても同様の傾向が散見されたため, 次節で背景領域を含む割合の高いクラスタを決定する際の条件に上述した傾向を反映させる．

背景領域候補の設定

前節で分析したように, 背景領域を含む割合の高いクラスタの含有率は n' が増加するにつれて減少する傾向がある．そこで, 前節の分析結果に基づき, 以下の条件を満たしたクラスタを背景領域を含む割合の高いクラスタと判定する．

(I)　含有率が単調減少傾向にある．すなわち $\zeta_{n'}^{m'}$ の増加率 $\dot{\zeta}_{n'}^{m'} = \zeta_{n'+1}^{m'} - \zeta_{n'}^{m'}$ が $0 \leq n' \leq N' - 2$ で常に閾値 T_α 以下である．



(a) 画像上半分　　(b) 画像下半分

図 3.11: 各クラスタの含有率

(II)　含有率が極小値を取る最小の $n_{\min}$ まで, 含有率の増加率が常に閾値 T_α 以下である.

条件 (I) は, 前節で図 3.11 (a) を分析した際に着目した緑, 橙色クラスタで確認された傾向に基づき定めた. 背景領域を含む割合が高く, 含有率が単調減少する傾向にあるクラスタには, 橙色クラスタのように含有率の増加率が必ずしも負にならないものが含まれる. そこで, 閾値 T_α が非負値を取ることを許容する. 条件 (II) は, 前節で図 3.11 (b) を分析した際に着目した青色クラスタで確認された傾向に基づき定めた.

続いて, 条件 (I), (II) のいずれかを満たすクラスタを基準に背景領域候補を求める. 以降では, 画像の上半分のみを対象に背景領域候補を設定するが, 同様の処理を画像下半分にも適用する. 条件 (I) を満たすクラスタは, $0 \leq n' \leq N' - 1$ の矩形 $R_{n'}$ で背景領域を含む割合が高いと考える. そこで, 条件 (I) を満たすクラスタに所属する画像上半分の画素全てを背景領域候補とする. 条件 (II) を満たすクラスタは, $0 \leq n' \leq n_{\min}$ の矩形 $R_{n'}$ で背景領域を含む割合が高いと考える. そこで, $0 \leq n' \leq n_{\min}$ の矩形 $R_{n'}$ に含まれる, 条件 (II) を満たすクラスタに所属する画素全てを背景領域候補とする.

図 3.9 (b) に示すクラスタリング結果に以上の処理を適用することで, 図 3.9 (c) の非黒色に示すように背景領域候補が求められる.

次に, 背景領域候補にモルフォロジー演算の一種の膨張処理を適用し, 建造物領域が背景領域候補に誤分類される割合を低減する. 背景領域候補に設定されていない残りの画素を建造物領域候補とする.

図 3.9 (c) に示す背景領域候補を膨張させることで, 図 3.9 (d) のように建造物と背景領域の候補が求められる. なお, 図 3.9 (d) において桃色と青色は背景領域候補, 緑色は建造物領域候補を示す. 桃色領域は画像の外縁部に接する背景領域候補のため, 以降の処理で建造物領域に分類されない背景領域の可能性が高い領域とする. 一方, 青色領域は画像の外縁部に接さない背景領域候補のため, 以降の繰り返し処理で建造物領域に分類される背景領域の可能性が低い領域とする.

GrabCut

前節までの処理で求められた建造物と背景領域の候補に基づいて初期化された GrabCut を適用する.

上野らの手法では多値の確信度を有した背景領域候補を設定するため, 確信度に基づく初期化が可能な GrowCut が採用されていた. 一方, 確信度を設定しない提案手法は GrowCut と GrabCut のいずれも採用することが可能である.

提案手法が GrowCut ではなく GrabCut を採用したのは, 多値の確信度に基づく初期化を行わない場合に, GrabCut が GrowCut と比較して高精度かつ高速に画像領域を分割できることが文献 [99–103] で示唆されているためである. 文献 [99, 100] では, GrabCut が GrowCut と比べて画像領域の分割精度を低下させることなく, 5〜30 倍の計算速度で画像領域を分割したことが複数のデータセットを用いた実験結果から示されている. この結果は, 文献 [99, 100] で詳細が分析されていないものの, GrowCut が収束までに要する繰り返し回数が多いことが原因だと考えられる. 文献 [101–103] では, GrowCut が 700 回以上繰り返しを要した入力画像に GrabCut が 10 回以下の繰り返しで収束したことが示されている.

図 3.9 (d) に示す建造物と背景領域の候補で初期化された GrabCut を適用することで, 図 3.9 (e) のように建造物領域が分割される.

3.4　評価実験

本節では, 建造物領域を分割する提案手法の有効性を確認するために従来手法との比較実験を行う. 3.4.1 節で実験条件を示した上で, 3.4.2 節で実験結果に基づき提案手法の有効性を確認する. 3.4.3 節で提案手法の改良点を考察する.

図 3.12: ZuBuD の画像例

3.4.1　実験条件

提案手法の有効性を示すために, 建造物が撮像された情景画像で構成される ZuBuD に存在する 106 枚の画像を用いて, 領域分割精度と計算時間を評価する. ZuBuD には, 5 つの視点から撮像した 201 種類の情景画像が収録されており, その中から正解データ画像作成が困難でない 106 枚の画像を選択した. なお視点はランダムに選択し, 建造物が横向きに撮影されている画像に対しては, 3.1 節で説明した入力画像に関する前提条件に基づいて建造物が正立するように画像を 90° 回転させた. 図 3.12 に ZuBuD の画像例を示す.

建造物領域 (l_{bld}) と背景領域 (l_{bkg}) が正しく分類された建造物領域の正解データ画像は, 文献 [72] と同様に被写体である建造物の領域を建造物領域として, 人手で教示することで作成した. 本研究において, 建造物領域分割は建造物認識の前処理と位置付けているため, 認識に不要と考えられる草木や雪などの領域は背景領域とした. 図 3.13 (b) は, 図 3.13 (a) に対して作成した正解データ画像である. 建造物領域と背景領域がそれぞれ白色と黒色で示されている.

図 3.14 に, 評価画像において建造物領域が分布する傾向のある画素位置を示すマップ（以降, 建造物領域の分布マップと呼称する）を可視化する. なお, 建造物が存在する割合の高い領域を白色で示している. 建造物領域を被写体とした評価画像は, 画像の中心部に建造物領域, 上部や下部に背景領域が分布する傾向が確認できる. 図 3.14 に示す建造物領域の分布マップ $\tilde{U} = \{\tilde{u}_0, \ldots, \tilde{u}_{P'-1}\}$ は, 正解データ画像の高さと幅の画素数を揃えるようにリサイズした画像を用いて以下の式で算出した.

$$\tilde{u}_{p'} = \frac{\sum_{\tilde{n}=0}^{\tilde{N}-1} \delta(u'_{\tilde{n},p'}, l_{\mathrm{bld}})}{\tilde{N}} \tag{3.6}$$

(a) 入力画像　　(b) 正解データ画像

図 3.13: 建造物と背景領域を示す正解データ画像の例

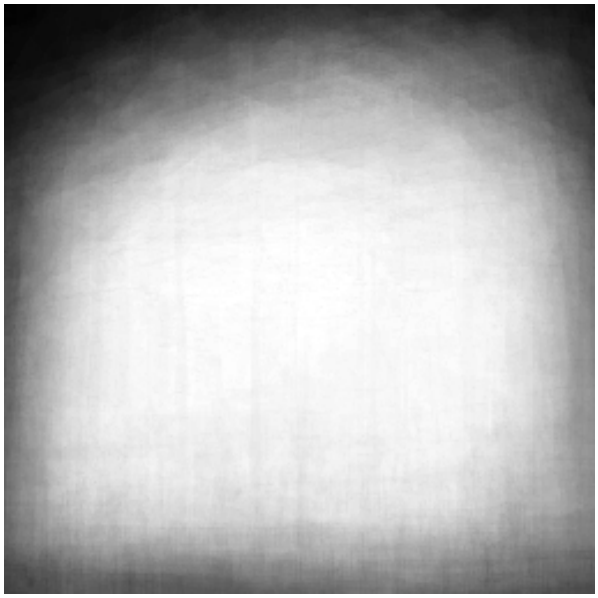


図 3.14: 建造物領域の分布マップ

ここで, $\tilde{N}$ は評価画像の枚数 (=106) を示し, $u'_{\tilde{n},p'}$ は $\tilde{n}$ 番目のリサイズした正解データ画像 U' の画素 p' における領域種別を示す.

評価対象となる建造物領域の分割手法は, 上野らの手法（3.2.1 節）, 提案手法（3.3 節）, SegNet（2.2.1 節）, 半自動化手法である. 上野らの手法は, 提案手法と同様に事前知識に基づく手法に分類されている. 提案手法が有効であるためには, 全ての指標で上野らの手法に対して高性能化を実現する必要がある. SegNet と半自動化手法は, それぞれ局所特徴に基づく手法とインタラクティブ手法に分類される.

図 3.15: CamVid の画像例

上野らの手法は文献 [72] で本実験と同様に ZuBuD を評価対象としており, 実験的に求めたパラメータが記載されているため, その値を流用した. 提案手法におけるパラメータは実験的に最適な値を求めた.

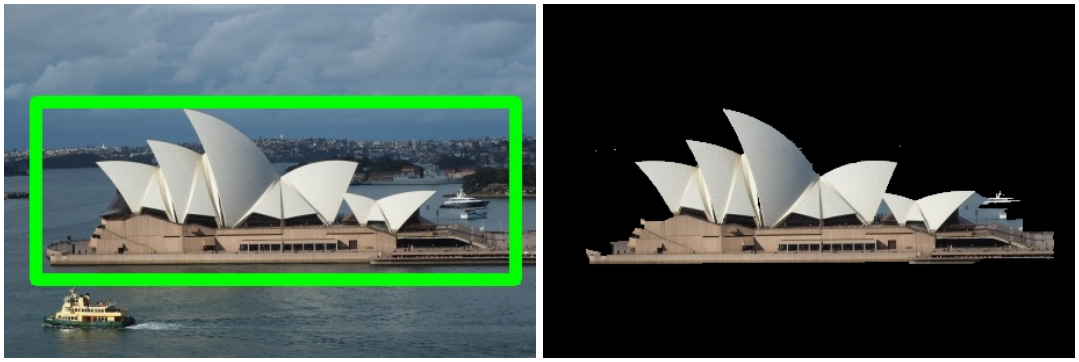
本研究では 2.3 節で述べているように, 導入コストの低い事前知識に基づく手法の高性能化を目的としているが, 近年では画素単位で領域種別が付与された学習データがインターネット上で公開されることもあり, それらを利用することで導入コストを抑えて局所特徴に基づく手法を構築することが可能である. そこで, 様々な画像領域分割問題に応用されている SegNet も比較対象とした. 本実験で利用する SegNet は, 情景画像で構成された CamVid [104] の学習データ 367 枚を利用して学習されたものである. 図 3.15 に CamVid の画像例を示す. 導入コストを抑えて構築した SegNet と比べて, 提案手法の建造物領域分割精度が向上していることを確認する.

半自動化手法は, ユーザから理想的な手動入力に基づいて建造物と背景領域の候補を設定し, GrabCut を適用する. 本実験では, 正解データ画像に基づいて求めた建造物領域の外接矩形の内側と外側をそれぞれ建造物と背景領域の候補とする. 図 3.16 (a) に緑色で示す矩形情報に基づいて分割された建造物領域を図 3.16 (b) に示す. なお, GrabCut や GrowCut を用いる上野らの手法, 提案手法, 半自動化手法において, 画像全体が建造物領域候補, もしくは背景領域候補になる場合は GrabCut や GrowCut を適用せず, 領域分割結果の全てを建造物領域, もしくは背景領域とした.

建造物領域の分割精度を示す指標として, 領域分割結果 Y と正解データ U の比較に基づく, 以下の Precision, Recall, F-measure を用いる.

- True Positive (TP): 建造物領域に正しく分類した画素数

$$\mathrm{TP} = \sum_{p=0}^{P-1} \delta(y_p, l_{\mathrm{bld}}) \cdot \delta(u_p, l_{\mathrm{bld}}) \tag{3.7}$$



(a) 入力画像　　(b) 領域分割結果

図 3.16: 半自動化手法による建造物領域の分割結果例

- False Positive (FP): 建造物領域に誤分類した画素数

$$\mathrm{FP} = \sum_{p=0}^{P-1} \delta(y_p, l_{\mathrm{bld}}) \cdot \delta(u_p, l_{\mathrm{bkg}}) \tag{3.8}$$

- True Negative (TN): 背景領域に正しく分類した画素数

$$\mathrm{TN} = \sum_{p=0}^{P-1} \delta(y_p, l_{\mathrm{bkg}}) \cdot \delta(u_p, l_{\mathrm{bkg}}) \tag{3.9}$$

- False Negative (FN): 背景領域に誤分類した画素数

$$\mathrm{TN} = \sum_{p=0}^{P-1} \delta(y_p, l_{\mathrm{bkg}}) \cdot \delta(u_p, l_{\mathrm{bld}}) \tag{3.10}$$

- Precision: 建造物領域に分類した画素数に対する建造物領域に正しく分類した画素数の割合

$$\mathrm{Precision} = \mathrm{TP}/(\mathrm{TP} + \mathrm{FP}) \tag{3.11}$$

- Recall: 建造物領域の画素数に対する建造物領域に正しく分類した画素数の割合

$$\mathrm{Recall} = \mathrm{TP}/(\mathrm{TP} + \mathrm{FN}) \tag{3.12}$$

- F-measure: Precision と Recall の調和平均

$$\text{F-measure} = \frac{2(\mathrm{Precision} \times \mathrm{Recall})}{\mathrm{Precision} + \mathrm{Recall}} \tag{3.13}$$

なお, Precision と Recall にはトレードオフの関係があるため, それらの調和平均を用いた F-measure を総合的な評価指標とする. 提案手法のような GrabCut を用いる手法において,

表 3.1: 各手法の建造物領域分割精度

	上野らの手法 [%]	提案手法 [%]	SegNet [%]	半自動化 [%]
Precision	75.5	85.6	90.8	87.5
Recall	89.7	93.8	84.0	96.8
F-measure	78.5	88.6	86.9	91.6

建造物と背景領域の候補を高精度に求めることができれば, 半自動化手法と同等の建造物領域分割精度が得られる可能性が高い. そこで, 以下の式に基づいて上野らの手法に対する提案手法の改善度も評価する.

$$\text{改善度} = \frac{\text{提案手法の F-measure} - \text{上野らの手法の F-measure}}{\text{半自動化手法の F-measure} - \text{上野らの手法の F-measure}} \tag{3.14}$$

計算時間は各画像に対する処理時間の平均値と中央値で評価した. 事前知識に基づく手法（上野らの手法と提案手法）間で計算時間を比較する. 上野らの手法と提案手法は, Numpy, scikit-image, OpenCV をライブラリとして用いた Python で実装し, Core i5-4460 (3.20 GHz) CPU, 16-GB RAM の PC 上で動作させた.

3.4.2　実験結果

領域分割精度

各手法による建造物領域の分割精度を表 3.1 に示す. 表 3.1 を見る限り, 提案手法は上野らの手法に対して, Precision, Recall, F-measure のいずれの評価値においても性能改善を実現しており, 総合的な評価指標である F-measure は 10.1% 向上した. (3.14) 式に基づいて求められる改善度は 77.1% だった. 続いて, 提案手法は SegNet と比べて F-measure が 1.7% 高く, インターネットで公開される学習データを用いて構築したモデルに対して高精度化を実現した. 一方, 上野らの手法は SegNet と比べて F-measure が 8.4% 低下した. しかし, 提案手法は半自動化手法と比べて全ての評価値で劣っていた. 特に F-measure は 3.0% 低く, さらなる改善の余地があると考えられる. それぞれの評価画像に対する F-measure をノンパラメトリック, パラメトリックな多重比較検定である Tukey-Kramer 法, Steel-Dwass 法を用いて有意水準 5% で検定したところ, 全手法間において有意差が確認された（p 値は 1% 以下）.

各手法による建造物領域の分割結果を図 3.17 に示す. 3 行目では画像下部の建造物領域

図 3.17: 各手法による建造物領域分割結果例

(a) 事例 1　(b) 事例 2　(c) 事例 3

図 3.18: 上野らの手法で精度低下が発生していた画像に対する提案手法の適用結果

を背景領域に誤分類している点に留意すべきだが, 提案手法は上野らの手法や SegNet と比較して全体的に分割精度が高く, 背景領域と建造物領域との境界が正しく得られている点で有効であると考えられる. そのため, 提案手法は輪郭線の類似度を利用することで建造物をクラス認識する Åhlen らの手法 [71] への応用が期待できる. 一方, 図 3.17 を見る限り, 提案手法は半自動化手法や SegNet と比較して草木などの背景領域を建造物領域に誤分類する傾向がある点で課題があると考えられる.

上野らの手法で性能低下が発生していた図 3.7 の事例 1〜3 の画像に対して, 提案手法を適用した結果を図 3.18 (a)〜(c) で確認する. 図 3.18 を見る限り, 提案手法はいずれの画像に対しても性能を改善したことが確認できる.

表 3.2: 上野らの手法と提案手法の計算時間

	上野らの手法 [s]	提案手法 [s]
平均値	27.2	3.7
中央値	11.9	3.2

まず, 提案手法では特定の領域との連結性を加味した建造物領域候補の設定を行っていないため, 上野らの手法のように画像中心部に連結しないことに起因して背景領域候補に建造物領域を誤分類する割合が高くなる可能性がない. 続いて, 提案手法は上野らの手法のように平均隣接数に基づく 2 値化で画像を 2 クラスタに分割するのではなく, RGB 空間において VBGMM による多クラスタ分割を行っているため, 輝度値の低い背景領域を建造物領域候補に誤分類する可能性が低下する. また, 提案手法も上野らの手法と同様に膨張処理に基づき建造物領域候補を拡張しているものの, 提案手法の建造物領域候補は GrabCut の繰り返し処理で背景領域に分類されることを許容するため, 膨張処理が原因で分割精度が低下する可能性は低くくなる.

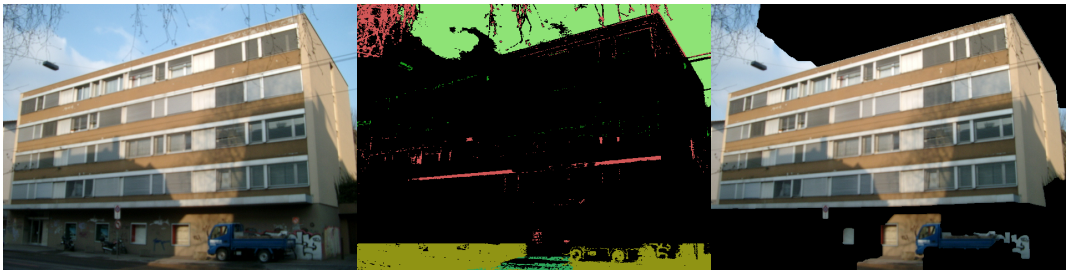
計算時間

上野らの手法と提案手法の各入力画像に対する計算時間を表 3.2 で比較する. 提案手法の計算時間は, 上野らの手法と比較して平均値と中央値でそれぞれ 23.5, 8.7 s 低く, 計算時間の 86.4%, 73.1% を低減したことが分かった.

以上, 本節では提案手法の有効性を確認するために建造物領域の分割精度と計算時間を比較評価した. 事前知識に基づく手法に分類される上野らの手法と提案手法の比較結果に着目すると, 提案手法は平均計算時間を 86.4% 低減し, F-measure を 10.1%（改善率 77.1%）向上した点で有効であった.

3.4.3　考察

本節では更なる性能向上を目的に提案手法で発生した課題を考察する. 図 3.19 に提案手法で性能改善が必要だと考える 2 つの事例を示す. それぞれの事例に対して, 左から 1 番目の画像が入力画像, 2 番目の画像が背景領域を含む割合の高いクラスタの判定結果（図



(a) 事例 A

(b) 事例 B

図 3.19: 提案手法による建造物領域分割結果例

3.9 (c) に対応）, 3 番目の画像が建造物領域の分割結果である.

図 3.19 (a), (b) に示す事例 A, B に着目すると, 建造物領域の一部が背景に誤分類されている点で精度低下が確認できる. 図 3.19 (a) の事例 A では, 画像下部の影で輝度が低下した領域が背景領域候補に誤分類されていることが原因で, GrabCut の適用後に建造物領域が背景領域に誤分類された. また図 3.19 (b) の事例 B では, 窓の色が背景領域と類似していることが原因で窓が背景領域に誤分類された. 他にも事例 A では, 雲の領域が建造物領域に誤分類される精度低下を確認した. これは, 雲の領域が背景領域候補に設定されていないことが原因である.

上記の課題を解決するために, 建造物領域の輪郭線は直線成分を多く含むことを利用した事後処理の導入が有効であると考えられる. しかし事後処理の導入によって, 一部の画像に対しては精度が低下する可能性がある点に留意する必要がある.

3.5 まとめ

本研究では, 建造物を地上で撮像した情景画像の建造物領域を分割する手法を提案した. 提案手法は, 背景領域が画像の上部や下部に分布する事前知識に基づいて, GrabCut を適用することで建造物領域を分割する. GrabCut の入力に必要な建造物と背景領域の候補は, RGB 空間を VBGMM によってクラスタリングした各クラスタのうち画像の上部もしくは下部に分布する傾向があるものを基準に定める.

提案手法の有効性を示すために, ZuBuD に収録されている 106 枚の情景画像を用いて比較評価した. まず, 提案手法は, 事前知識に基づく手法に分類される上野らの手法と比べて 86.4% の平均計算時間低減と 10.1% の平均精度改善（改善率は 77.1%）を実現した. また, 上野らの手法で性能低下が発生していた画像に対しても, 建造物領域の分割精度が向上していることを確認した. 続いて, 提案手法は, 情景画像 367 枚で学習された, 局所特徴に基づく手法に分類される SegNet と比べて, 1.7% の平均精度改善を実現した. しかし, 提案手法はインタラクティブ手法に分類される半自動化手法と比べて平均精度が 3.0% が低く, 更なる改良の余地があると考えられる.

提案手法による建造物領域の分割結果を確認したところ, 背景領域に類似した建造物領域を背景領域に誤分類する課題があることが判明した. この課題に対しては, 建造物領域と背景領域の境界を調査することが有効だと考えられるため, 今後の性能改善に向けて検証を進めていく.

第 4 章

画像外縁部に領域成長法を用いる商品領域分割手法

4.1 はじめに

近年, スマートフォンの普及とともにインターネットを介したフリーマーケットやオークションの市場規模が増大傾向にある [105, 106]. インターネットオークションやフリーマーケットにおいて, 販売者は商品情報を提示するために商品の外観を撮像した画像をアップロードし, 商品の名称や種別（ファッション製品, 美容製品, スポーツ製品等）を手動入力する. 購入者は販売者に提示された商品の情報に基づいて購入意思を決定する.

販売者がアップロードする商品画像は商品価格に影響を与えることが考察されており, インターネットオークションやフリーマーケットにおいて重要である [107]. また, 商品画像は利用者の作業負荷を低減することにも用いられている. 例えば, 商品画像の商品をクラス認識することで商品の名称を特定する機能や, 関連した商品情報を自動入力する機能が存在する.

商品クラス認識として, 2.1 節で例示した商品領域分割を適用することで得られる商品領域の輪郭線情報を局所特徴とする手法が存在する [74, 108–113]. 上述した商品クラス認識において輪郭線情報の抽出精度が認識性能を左右するため, 商品クラス認識の性能改善に向けて商品領域を分割する技術の高精度化が求められている. また, 商品領域を分割する技術は商品画像に不要な背景を除去して, 背景を商品に合わせて変更する画像加工技術にも応用可能である.

そこで, 本研究では商品画像の商品領域を分割する手法の高性能化を目的とする. この目的を実現するために, 周辺画素を併合しながら画像の外縁部を拡大させることで背景領域の候補を求める処理を導入した商品領域分割手法を提案する. 続いて, 2.3.2 節で述べた Cheng らの手法との比較実験に基づいて領域分割精度と計算時間の観点から提案手法の有効性を示す.

以降の内容について説明する. 4.2 節で Cheng らの手法の各処理を詳細に述べ, 領域分割精度低下を引き起こす Cheng らの手法の課題をまとめる. 4.3 節でインターネットオークションで実際に使用されている商品画像を分析し, 分析結果に基づいて構築した提案手法についてフローチャートを用いて各処理を詳細に述べる. 4.4 節で商品画像を用いた比較実験に基づいて提案手法の有効性を考察し, 本章の内容を 4.5 節でまとめる.

4.2 Cheng らの手法

本節では, 図 2.22 に例示する商品画像の商品領域を分割する Cheng らの手法について説明する. Cheng らの手法は, 人の注目を集めるように撮影される商品領域の誘目度が背景領域と比較して高いことを事前知識としている.

4.2.1 処理概要

Cheng らの手法は, 商品画像に対して商品と背景領域の候補を求めて, GrabCut を適用する. Cheng らの手法のフローチャートを図 4.1 に示す. 以降, 図 4.1 の各項目を詳細に述べる.

誘目度マップ生成

まず, 入力画像を, 明度を表す L*信号と色度を表す a*, b*信号から構成される L*a*b*空間に変換する. 続いて, L*a*b*空間に変換された画像に Felzenszwalb らの手法 [114] を適用し, 色が類似した近傍の画素集合（以降, スーパーピクセルと呼称する）を求める. Felzenszwalb らの手法では, L*a*b*成分の値と画素の位置で表される 5 次元の特徴量を用いて各画素をクラスタリングする. 図 2.22 の各画像に対してスーパーピクセルを求めた結果を図 4.2 に示す. 図 4.2 の黄線は, スーパーピクセル間の境界を示す. 色と位置の近い画素が同一のスーパーピクセルに分類されていることが確認できる.

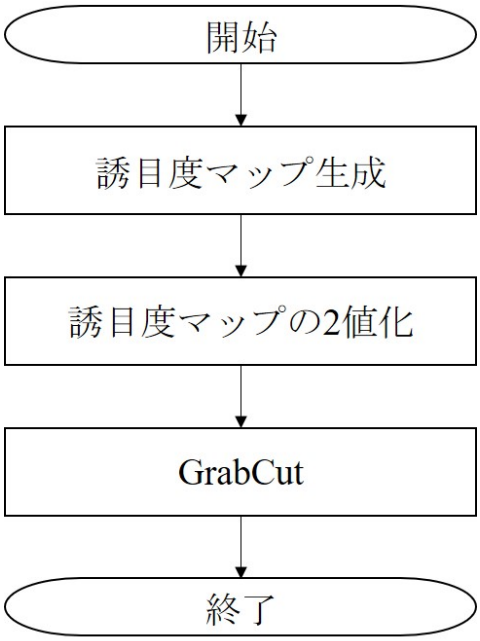


図 4.1: Cheng らの手法のフローチャート

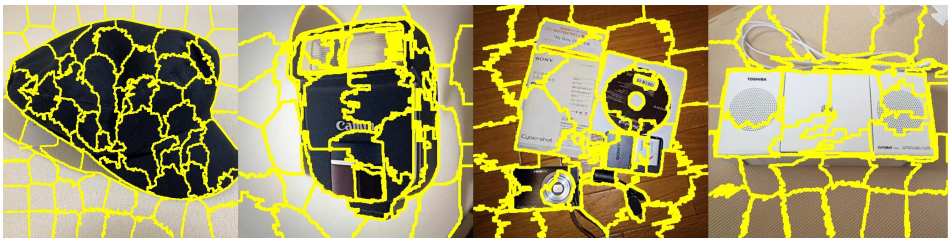


図 4.2: 図 2.22 の各画像に対するスーパーピクセル分割結果

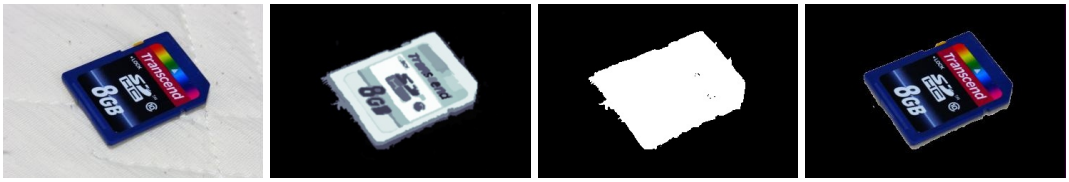
次に, スーパーピクセル $\tilde{p}$ における誘目度 $\tilde{S}$ を以下で求める.

$$\tilde{S}(\tilde{p}) = \sum_{\tilde{p} \neq \tilde{p}'} \exp\left(-\tilde{D}_s(\tilde{p}, \tilde{p}')/\tilde{\sigma}_s^2\right) \tilde{w}(\tilde{p}) \tilde{D}_r(\tilde{p}, \tilde{p}') \tag{4.1}$$

ここで, $\tilde{D}_s(\tilde{p}, \tilde{p}')$ はスーパーピクセル $\tilde{p}$ と $\tilde{p}'$ の重心座標間のユークリッド距離を意味し, $\tilde{\sigma}_s$ はスーパーピクセル間のユークリッド距離が誘目度に与える影響を調整するパラメータを意味する. $\tilde{w}(\tilde{p})$ はスーパーピクセル $\tilde{p}$ に含まれる画素数を示す. $\tilde{D}_r(\tilde{p}, \tilde{p}')$ は, 以下で定義するスーパーピクセル $\tilde{p}$ と $\tilde{p}'$ に含まれる色に関する距離を示す.

$$\tilde{D}_r(\tilde{p}, \tilde{p}') = \sum_{\tilde{i}=0}^{\tilde{N}_{\tilde{p}}-1} \sum_{\tilde{i}'=0}^{\tilde{N}_{\tilde{p}'}-1} \tilde{f}(\tilde{p}, \tilde{i}) \tilde{f}(\tilde{p}', \tilde{i}') \| v_{\tilde{p},\tilde{i}} - v_{\tilde{p},\tilde{i}'} \| \tag{4.2}$$

ここで, $v_{\tilde{p},\tilde{i}}$ はスーパーピクセル $\tilde{p}$ に含まれる $\tilde{i}$ $(\in [0, \tilde{N}_{\tilde{p}}])$ 番目の色情報を L*a*b*空間で



(a) (b) (c) (d)

図 4.3: Cheng らの手法による商品領域分割過程

図 4.4: 誘目度マップの例（左：入力画像, 右：誘目度マップ）

表す. また, $\tilde{f}(\tilde{p},\tilde{i})$ はスーパーピクセル $\tilde{p}$ に $\tilde{i}$ 番目の色が含まれる割合を意味する.

以上の処理を図 4.3 (a) に適用することで, 図 4.3 (b) に示す誘目度マップが求められる. なお, 誘目度の高さを白色成分の強度で示している. また, 図 4.4 に以上の処理で求めた他の誘目度マップを参考として示す.

誘目度マップの 2 値化

前節で求めた誘目度マップを事前に設定した閾値で 2 値化する. 誘目度が閾値以上の画素を商品領域候補, 残る画素を背景領域候補とする.

図 4.3 (b) に示す誘目度マップに以上の処理を適用することで, 図 4.3 (c) のように誘目度マップが 2 値化される. 白色と黒色がそれぞれ商品と背景領域の候補を示している.

GrabCut

前節までの処理で求められた商品と背景領域の候補に基づいて初期化された GrabCut を適用する. GrabCut の繰り返し処理で, 最小カットアルゴリズムの適用後に得られる商品と背景領域の候補にモルフォロジー演算の一種である収縮処理を用いる. 収縮処理を適

用した後の商品と背景領域の候補は, 以降の最小カットアルゴリズムで異なる領域に分類されない.

図 4.3 (c) に示す商品と背景領域の候補に以上の処理を適用することで, 図 4.3 (d) のように商品領域が分割される.

4.2.2　課題

Cheng らの手法を, 無作為抽出した実際に商品画像として使用されている 300 枚の商品画像（以降, 分析データセットと呼称する）に適用したところ, 図 4.5 のように領域分割精度が低下する事例を確認した. 図 4.5 では, 各事例に対して左から 1 番目が入力画像, 2 番目が誘目度マップ, 3 番目が商品領域分割結果を示している. なお, 誘目度マップは図 4.3 (b) と同様に誘目度の高さを白色成分の強度で示している.

まず, 図 4.5 (a) の事例 1 は, 反射で白く撮影された領域を含む茶色のフローリングが背景で, ロボットの玩具が商品の画像である. 背景領域と色が類似していることが原因で, 玩具の白色や茶色領域の誘目度が低下している. その結果, GrabCut の適用後に玩具に含まれる誘目度が低下した領域の一部が背景領域に誤分類されている.

次に, 図 4.5 (b) の事例 2 は, 白色と茶色のタイルマットが背景で, 灰色のボイスレコーダが商品の画像である. 背景領域と色が類似していることが原因で, ボイスレコーダの白色成分の強い一部の領域における誘目度が低下している. その結果, GrabCut の適用後にボイスレコーダに含まれる誘目度が低下した領域の一部が背景領域に誤分類されている.

上記で述べた領域分割精度が低下する課題は, 背景の色と類似する一部の商品領域の誘目度が低下することに起因している.

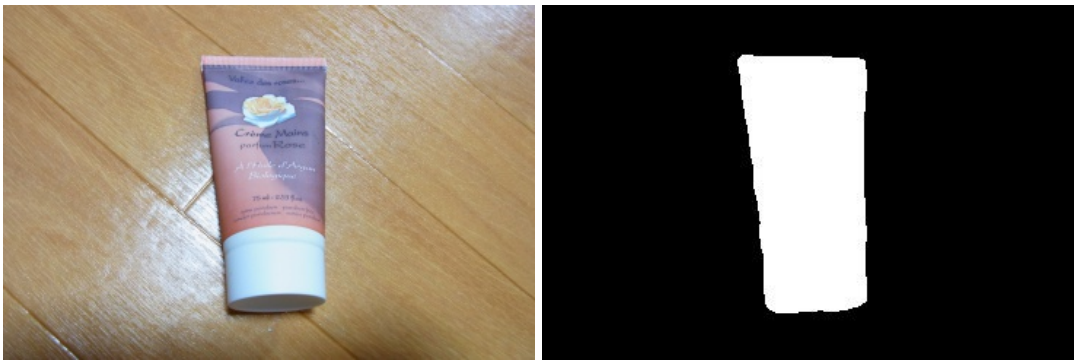
4.3　提案手法

商品画像には被写体の商品が画像の中心部に分布する傾向があると考えられる. そこで, 4.3.1 節で商品領域の分布傾向を分析した上で, 分析結果に基づいて構築した商品領域分割の処理概要を 4.3.2 節で説明する.

(a) 事例 1

(b) 事例 2

図 4.5: Cheng らの手法による商品領域分割精度低下の例



(a) 入力画像　　(b) 正解データ画像

図 4.6: 商品と背景領域を示す正解データ画像の例

4.3.1　商品領域分布の分析

インターネットオークション等で用いられる商品画像を対象に商品領域の分布を分析する．まず, 4.2.2 節で用いた分析データセットに含まれる各画像に対して, 3.4.1 節と同様に商品と背景領域が正しく分類された正解データ画像を作成した．図 4.6 (b) は, 図 4.6 (a) に

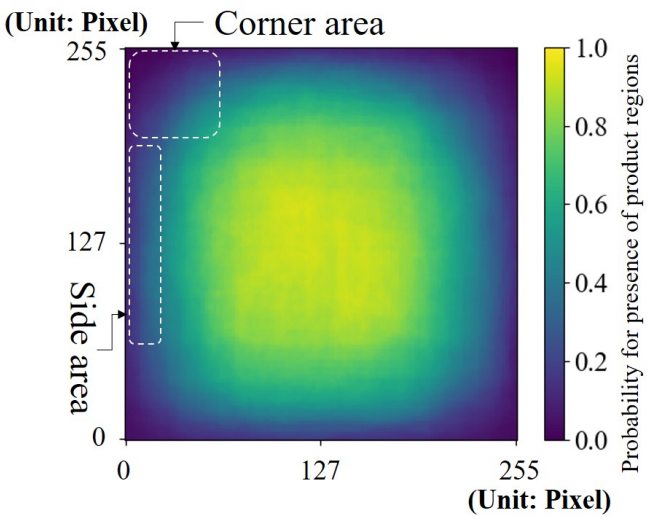


図 4.7: 商品領域の分布マップ

対して作成した正解データ画像である．続いて, 3.4.1 節と同様に正解データ画像の高さと幅の画素数を揃えたリサイズ画像を用いて, 商品領域が分布する傾向のある画素位置を示すマップ（以降, 商品領域の分布マップと呼称する）を (3.6) 式のように求めた．

図 4.7 に商品領域の分布マップを可視化する．可視化されたマップを見る限り, 商品領域は画像の外縁部と比較して中央部に分布する傾向があることが確認できる．また, 画像の 4 隅周辺部 (corner areas と図中に記載) は 4 辺周辺部 (side areas と図中に記載) と比較して, 商品が低確率で分布する領域の面積が広い．

4.3.2　処理概要

提案手法は, 4.3.1 節の分析結果に基づいて商品と背景領域の候補を求めて, GrabCut を適用する．提案手法のフローチャートを図 4.8 に示す．以降, 図 4.8 の各項目を詳細に述べる．

外縁部に対する背景領域候補の設定

4.3.1 節の分析結果によると, 商品画像の 4 隅, 4 辺周辺部において商品領域が分布する可能性が低下する．そこで, 図 4.9 に示す A_c, A_s から構成される外縁部を背景領域候補とする．なお, $\tilde{r}_c, \tilde{r}_s$ ($\in [0\%, 50\%]$) は, A_c, A_s の大きさを決定するパラメータを意味する．4.3.1 節における分析結果によると, 商品画像の 4 隅周辺部は 4 辺周辺部と比べて商品が低

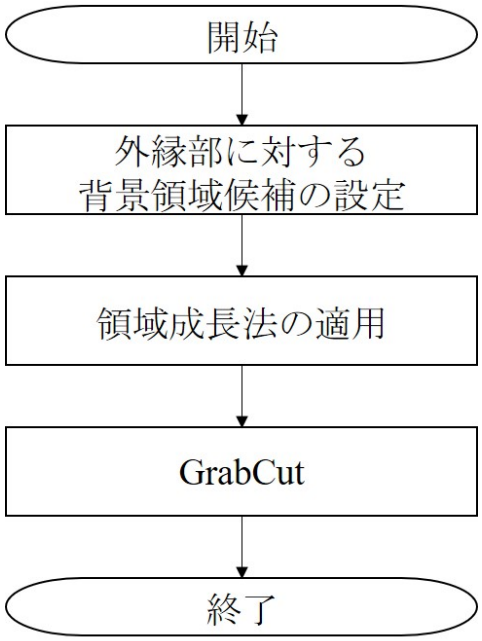


図 4.8: 提案手法のフローチャート

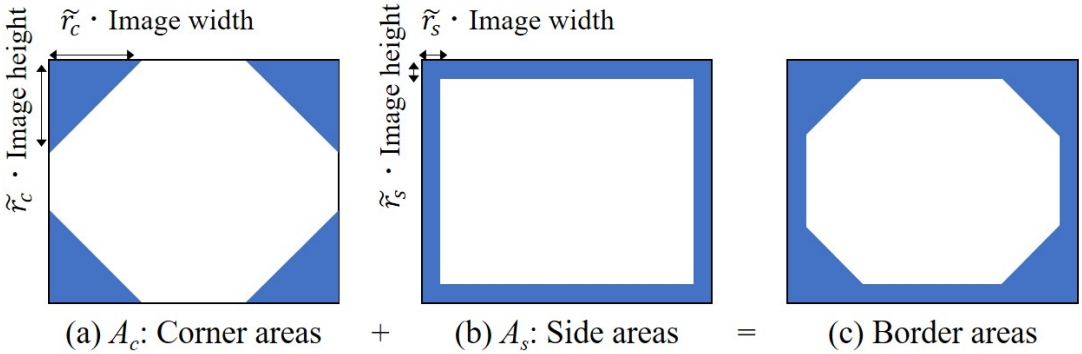
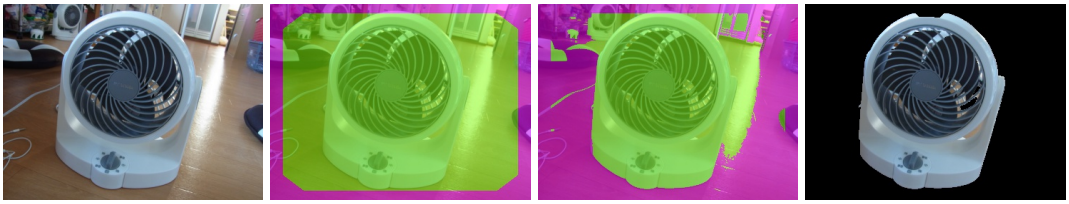


図 4.9: 背景領域候補とする画像外縁部

確率で分布する領域の面積が広い．そこで, 分析結果に基づいて外縁部を設定するために $\tilde{r}_s < \tilde{r}_c$ とする.

図 4.10 (a) に示す入力画像に対して, 以上の処理で求めた背景領域候補を図 4.10 (b) に桃色で示す.



(a)　(b)　(c)　(d)

図 4.10: 提案手法による商品領域分割過程

領域成長法の適用

前節で設定された背景領域候補に領域成長法 [115] を適用し, 背景領域候補を拡張する. 領域成長法とは, 指定された領域に対して連結条件を満たす周辺の画素を併合しながら領域を拡張する処理を意味する. 指定された領域に隣接する画素間の距離を計算し, 閾値以下の場合に連結条件を満たしたとし, その画素を併合する. 画素 p, q 間の距離として, 画素値間のユークリッド距離 $(= \|x_p - x_q\|)$ を用いる. なお, 提案手法では画素値として RGB 値を用いる.

4.3.1 節において, 商品画像の 4 隅周辺部は 4 辺周辺部と比べて商品が低確率で分布する領域の面積が広いことが分かった. そこで, 提案手法は 4 隅周辺部と 4 辺周辺部に対して異なる閾値（それぞれ T'_c, T'_s とする）で領域成長法を適用する. 分析結果に基づいて 4 隅周辺部により多くの周辺画素を併合させるために $T'_c > T'_s$ とする. 背景領域候補に分類されなかった残りの画素は商品領域候補に分類する. なお, 以上の処理で求めた背景領域候補は, 以降の処理で商品領域に分類されない背景領域の可能性が高い領域とする.

図 4.10 (b) に桃色で示す背景領域候補に以上の処理を適用することで, 図 4.10 (c) に桃色で示すように背景領域候補が拡張される. 緑色領域は商品領域候補を示す.

GrabCut

前節までの処理で求められた商品と背景領域の候補に基づいて初期化された GrabCut を適用する. 図 4.10 (c) に示す商品と背景領域の候補で初期化された GrabCut を適用することで, 図 4.10 (d) のように商品領域が分割される.

4.4 評価実験

本節では, 商品領域を分割する提案手法の有効性を確認するために従来手法との比較実験を行う. まず, 4.4.1 節で実験条件を示した上で, 4.4.2 節で実験結果を確認する. 続いて, 4.4.3 節で提案手法を構成する各処理の有効性を考察する.

4.4.1 実験条件

提案手法の有効性を示すために, インターネットオークションで実際に使用された商品画像 341 枚を用いて領域分割精度と計算時間を評価する. なお, 評価画像には 4.2.2 節の分析データセットに含まれる画像が存在しない. 商品領域 (l_{pdc}) と背景領域 (l_{bkg}) が正しく分類された商品領域の正解データ画像は, 3.4.1 節と同様に人手で教示することで作成した.

評価対象となる商品領域の分割手法は, Cheng らの手法（4.2.1 節）, 提案手法（4.3.2 節）, SegNet（2.2.1 節）, 半自動化手法である. Cheng らの手法は, 提案手法と同様に事前知識に基づく手法に分類されている. 提案手法が有効であるためには, 全ての指標で Cheng らの手法に対して高性能化を実現する必要がある. SegNet と半自動化手法は, 2.2 節で示した通りそれぞれ局所特徴に基づく手法とインタラクティブ手法に分類される.

本実験で利用する SegNet は, 4.3.1 節で提案手法の構築に用いた分析データセットを用いて学習させる. 提案手法と同等の導入コストで構築した SegNet と比べて, 提案手法の商品領域分割精度が向上していることを確認する.

半自動化手法は, 3.4.1 節で説明した通り, 正解データ画像に基づいて商品と背景領域の候補を設定し, GrabCut を適用する. 図 4.11 (a) に緑色で示す矩形情報に基づいて分割された商品領域を図 4.11 (b) に示す.

商品領域の分割精度を示す指標として, 3.4.1 節で定義されている Precision((3.11) 式), Recall((3.12) 式), F-measure((3.13) 式) を用いる. 総合的な評価指標として, トレードオフの関係がある Precision と Recall ではなく, F-measure を用いる. なお, TP, FP, TN, FN は分割結果 Y と正解データ画像 U の比較に基づいて以下のように定義される.

- TP: 商品領域に正しく分類した画素数

$$\mathrm{TP} = \sum_{p=0}^{P-1} \delta(y_p, l_{\mathrm{pdc}}) \cdot \delta(u_p, l_{\mathrm{pdc}}) \tag{4.3}$$

(a) 入力画像　　(b) 領域分割結果

図 4.11: 半自動化手法による商品領域の分割結果例

- FP: 商品領域に誤分類した画素数

$$\mathrm{FP} = \sum_{p=0}^{P-1} \delta(y_p, l_{\mathrm{pdc}}) \cdot \delta(u_p, l_{\mathrm{bkg}}) \tag{4.4}$$

- TN: 商品領域に正しく分類した画素数

$$\mathrm{TN} = \sum_{p=0}^{P-1} \delta(y_p, l_{\mathrm{pdc}}) \cdot \delta(u_p, l_{\mathrm{bkg}}) \tag{4.5}$$

- FN: 商品領域に誤分類した画素数

$$\mathrm{TN} = \sum_{p=0}^{P-1} \delta(y_p, l_{\mathrm{bkg}}) \cdot \delta(u_p, l_{\mathrm{pdc}}) \tag{4.6}$$

提案手法のように GrabCut を用いる手法において, 商品と背景領域の候補を高精度に求めることができれば, 半自動化手法と同等の商品領域分割精度が得られる可能性が高い. そこで, 以下の式に基づいて Cheng らの手法に対する提案手法の改善度を (3.14) 式と同様に評価する.

$$\text{改善度} = \frac{\text{提案手法の F-measure} - \text{Cheng らの手法の F-measure}}{\text{半自動化手法の F-measure} - \text{Cheng らの手法の F-measure}} \tag{4.7}$$

計算時間は 3.4.1 節で説明したように各画像に対する処理時間の平均値と中央値で評価した. 3.4.2 節における実験結果と同様に, 事前知識に基づく手法（Cheng らの手法と提案手法）間で計算時間を比較する. Cheng らの手法と提案手法は, Numpy, scikit-image, OpenCV をライブラリとして用いた Python で実装した.

表 4.1: 各手法による商品領域分割精度

	Cheng らの手法 [%]	提案手法 [%]	SegNet [%]	半自動化 [%]
Precision	88.0	92.2	81.2	90.7
Recall	87.0	89.2	77.2	96.2
F-measure	87.5	90.7	79.2	93.4

4.4.2　実験結果

領域分割精度

各手法による商品領域の分割精度を表 4.1 に示す. 表 4.1 を見る限り, 提案手法は Cheng らの手法に対していずれの評価値においても性能改善を実現しており, 総合的な評価指標である F-measure は 3.2% 向上した. (4.7) 式に基づいて求められる改善度は 54.2% だった. 続いて, 提案手法は SegNet と比べて F-measure が 11.5% 向上した. 分析データセットに含まれる 300 枚の学習データで, 多種多様な商品画像に対して SegNet を高精度化するのは困難なことが分かった. しかし, 提案手法は半自動化手法と比べて Precision は向上したものの, Recall と F-measure が低下した. 特に F-measure は 5.9% 低く, さらなる改善の余地があると考えられる. 3.4.2 節と同様に, F-measure に対してノンパラメトリック, パラメトリックな多重検定を適用したところ, 全手法間において有意差が確認された（p 値は 1% 以下）.

各手法による商品領域の分割結果を図 4.12 に示す. 1 行目では, 全ての手法で商品領域を高精度に分割できていることが確認できるが, 2 行目では, 鞄の紐で囲われている壁を商品領域に誤分類する傾向がある. 3 行目では, 提案手法のみが外縁部に位置する電源ケーブルを背景領域に誤分類していることが確認できる. これは, 外縁部を背景領域候補に設定する提案手法の処理に起因している.

Cheng らの手法で性能低下が発生していた図 4.5 の事例 1, 2 の画像に対して, 提案手法を適用した結果をそれぞれ図 4.13 (a), (b) で確認する. 図 4.13 を見る限り, 提案手法はいずれの画像に対しても性能を改善したことが確認できる. これは提案手法が画像外縁部の周辺画素を背景領域候補として併合するため, 商品領域の内部に誤って背景領域候補に設定された画素を含む割合が低下するためである.

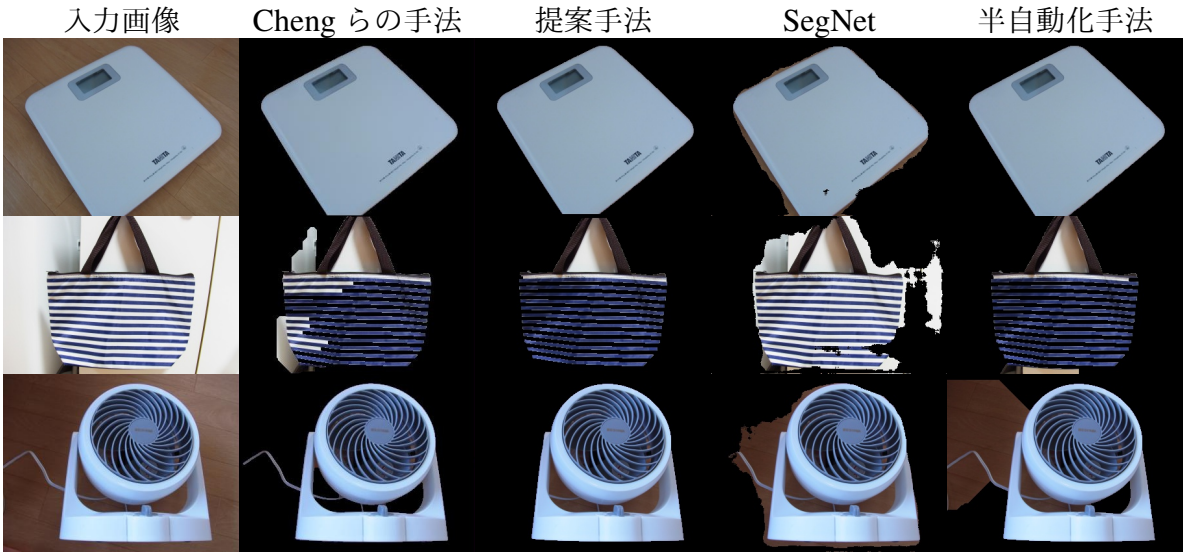
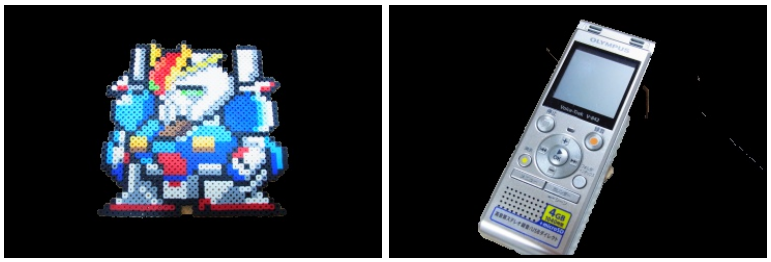


図 4.12: 各手法による商品領域分割結果例



(a) 事例 1　　(b) 事例 2

図 4.13: Cheng らの手法で精度低下が発生していた画像に対する提案手法の適用結果

計算時間

Cheng らの手法と提案手法の各入力画像に対する計算時間を表 4.2 で比較する．提案手法の計算時間は, Cheng らの手法と比較して平均値と中央値でそれぞれ 12.1, 9.7 s 低く, 計算時間の 76.7%, 74.5% を低減したことが分かった．Cheng らの手法における誘目度マップを生成する処理の計算時間は, 平均値と中央値でそれぞれ 14.0, 11.2 s 以上で全体処理の 85% 以上を占めていた．一方, 提案手法は Cheng らの手法のように誘目度マップを生成する処理を含まないため, 計算時間の低減を実現できた．

以上, 本節では提案手法の有効性を確認するために商品領域の分割精度と計算時間を比較した．事前知識に基づく手法に分類される Cheng らの手法と提案手法の比較結果に着目すると, 提案手法は平均計算時間を 76.7% 低減し, F-measure を 3.2%（改善率 54.2%）向

表 4.2: Cheng らの手法と提案手法の計算時間

	Cheng らの手法 [s]	提案手法 [s]
平均値	15.8	3.7
中央値	13.0	3.3

上した点で有効であった.

4.4.3　考察

本節では, 4.3.2 節で説明した提案手法を構成する各処理の有効性を考察する. 以下に示す提案手法の一部処理のみを用いた 3 種類の手法（SBG, SBG + RG, SBG + GC）の改善率を表 4.3 にまとめる.

- SBG: 画像の外縁部を背景領域とし, 残る領域を商品領域に分割する手法
- SBG + RG: 画像の外縁部に領域成長法の適用し, 拡大された背景領域候補を背景領域, 残る領域を商品領域に分割する手法
- SBG + GC: 画像の外縁部を背景領域候補, 残る領域を商品領域候補として初期化された GrabCut によって商品領域を分割する手法

まず, GrabCut を用いない SBG と SBG + RG は提案手法と比べて改善率が 56.1% 以上低下している. そのため, GrabCut は商品領域の分割精度向上に必要な処理であることが分かる. 続いて, 領域成長法を用いない SBG + GC は提案手法と比べて改善率が 15.2% 低下したため, 領域成長法も GrabCut と同様に商品領域の分割精度向上に向けて必要なことが分かる. 次に, 外縁部に基づいて初期化された GrabCut を適用する SBG + GC の改善率は 39.0% で Cheng らの手法に対して商品領域分割の精度向上を実現した. この結果は, 外縁部を背景領域候補に設定する提案手法の処理が, 商品領域の分割精度向上に向けて重要であることを示している.

表 4.3: 各手法の改善率

手法	外縁部に対する背景領域候補の設定	領域成長法の適用	GrabCut	改善率 [%]
SBG	✓			-4.0
SBG + RG	✓	✓		-1.9
SBG + GC	✓		✓	39.0
提案手法	✓	✓	✓	54.2

4.5　まとめ

本研究では, インターネットオークション等で用いられる商品画像の商品領域を分割する新規手法を商品領域分布に関する分析結果に基づいて提案した. 提案手法は, 商品領域が画像の中央部に分布する事前知識に基づいて GrabCut を適用することで商品領域を分割する. GrabCut の入力に必要な商品と背景領域の候補は, 画像の外縁部に領域成長法を適用することで定める.

提案手法の有効性を示すために, インターネットオークションで実際に使用された 341 枚の商品画像を用いて比較評価した. まず, 提案手法は事前知識に基づく手法に分類される Cheng らの手法に対して 76.7% の平均計算時間低減と 3.2% の平均精度改善（改善率は 54.2%）を実現した. また, 商品領域の一部で誘目度が低下したことに起因した, Cheng らの手法で発生していた精度低下も, 提案手法を用いることで回避できる傾向が確認された. 続いて, 提案手法は, 提案手法を構築するための分析に用いた 300 枚の商品画像を利用して学習した, 局所特徴に基づく手法に分類される SegNet と比べて, 11.5% の平均精度改善を実現した. しかし, 提案手法はインタラクティブ手法に分類される半自動化手法と比べて平均精度が 5.9% 低く, 更なる改良の余地があると考えられる.

また, 提案手法の各処理が Cheng らの手法に対して精度向上を実現するために必要な処理であることを考察した. 領域成長法を用いない提案手法は, Cheng らの手法に対して 39.0% の改善率を実現できることが分かった. この結果は, 誘目度ではなく, 外縁部を基準に商品と背景領域の候補を定めることの有効性を示している.

第 5 章

差分領域分割に基づく人感センサの有効性評価

5.1 はじめに

近年, 低炭素社会の実現に向けて電気設備のエネルギー消費量を抑制することが求められている. 電気設備の利用者に過度な省エネルギーを強いることは利便性の低下に繋がるため, 利用者が意識することなく省エネルギーを実現できる自動制御技術の高度化が求められている. 電気設備の自動制御技術として, 人感センサの在室者検知結果を用いる在不在制御が存在する [116–119].

照明や空調を例に在不在制御を説明する. 在不在制御では, 人感センサが在室者を検知した（在判定した）場合に照明を点灯, 空調能力を向上させ, 不在判定した場合に照明を消灯, 空調能力を低下させる. 詳細は後述するが, 在不在制御の省エネルギー効果は人感センサの在室者検知精度に左右されるため, 在室者検知精度を向上できる人感センサの普及が求められている.

在室者検知精度を向上できる可能性がある人感センサ技術として, 2.1 節で例示した差分領域分割を用いた在室者検知が挙げられる. 近年, 監視カメラを内蔵し, 撮影した画像に差分領域分割を適用することで在室者を検知する人感センサが商用化されている [75]. しかし, その他の人感センサに対する差分領域分割に基づく人感センサの有効性が関連研究で十分に明らかになっていない. 例えば, 差分領域分割に基づく人感センサはその他の人感センサと比較して在室者の細かい動作を正しく在判定できることが示唆されている

が [75, 76, 120], それを実環境下で検証した関連研究は存在しない. また, 差分領域分割に基づく人感センサを用いた在不在制御の省エネルギー効果や消費電力量を実環境下で評価した関連研究は存在しない. 差分領域分割に基づく人感センサの在室者検知精度, 省エネルギー効果, 消費電力量を評価し, その有効性を考察することは人感センサの今後の改良方針を定める上で重要である.

物体検出やクラス認識を応用した在室者検知も広く研究されているが, 実環境下における有効性をその他の人感センサと比較した関連研究は存在しない. 画像認識技術の実用化や普及に向けた研究を進める上でも, 差分領域分割に基づく人感センサをその他の人感センサと比較評価することが求められている.

そこで, 本章は差分領域分割に基づく人感センサの有効性を明らかにすることを目的とする. 具体的には, 差分領域分割に基づく人感センサを在室者検知精度, 省エネルギー効果, 消費電力量の観点から評価し, 2.3.3 節で説明した赤外線センサに対する利点や課題を整理する. なお, 省エネルギー効果や消費電力量を評価する際に, 本研究では照明設備との在不在制御（以降, 照明在不在制御と呼称する）を対象とする. これは, 一般家庭やオフィスの消費エネルギー内訳において照明設備が占める割合が高く [121, 122], 低炭素社会実現に向けて省エネルギー化が必要不可欠なためである.

以降の内容について説明する. 5.2 節で現在商用化が進められている差分領域分割に基づく人感センサの処理を詳細に述べる. 5.3 節で照明設備に対する照明在不在制御をフローチャートを用いて説明する. 5.4 節で人感センサを設置する環境, 実験条件, 評価方法について説明し, 差分領域分割に基づく人感センサの有効性を実験結果に基づき考察する. 5.5 節で本章の内容をまとめる.

5.2 Sarker らの手法

本節では, 2.3.3 節で述べたように差分領域分割に基づく人感センサの在室者検知アルゴリズムとして想定する Sarker らの手法の詳細を述べる. 人感センサは, 在室状態を表す種別 $\Xi = \{\xi_{\mathrm{o}}, \xi_{\mathrm{f}}\}$ のいずれかを時刻 t における予測結果 y_t^{ξ} として出力する. ここで, ξ_{o} と ξ_{f} はそれぞれ在状態と不在状態を示す.

(a) 過去画像 X^{t-1}

(b) 現在画像 X^{t}

(c) 差分領域分割結果 Y

図 5.1: 差分領域分割の例

5.2.1　処理概要

Sarker らの手法は, 図 5.1 (a) に示す過去画像 (X^{t-1}) と図 5.1 (b) に示す現在画像 (X^{t}) を比較して, 差分領域 (l_{dif}) と背景領域 (l_{bkg}) を以下で分割する.

$$y_p = \begin{cases} l_{\mathrm{dif}} & (|x_p^t - x_p^{t-1}| > T_f) \\ l_{\mathrm{bkg}} & (\text{otherwise}) \end{cases} \tag{5.1}$$

なお, T_f は画素値間の差分に関する閾値を意味する.

続いて, 上式で求めた図 5.1 (c) の差分領域分割結果 Y に基づき, 以下のように在室者の

検知結果を出力する.

$$y_t^{\xi} = \begin{cases} \xi_{\mathrm{o}} & (r_y > T_r) \\ \xi_{\mathrm{f}} & (\text{otherwise}) \end{cases} \tag{5.2}$$

$$r_y = \frac{\sum_{p=0}^{P-1} \delta(y_p, l_{\mathrm{dif}})}{P} \tag{5.3}$$

なお, T_r は差分領域が占める割合に関する閾値を意味する.

5.3　照明在不在制御

本節では, 前節で説明した人感センサの在室者検知結果に基づいて照明を制御する照明在不在制御について説明する. 照明在不在制御では, 照明を誤って消灯すること（誤消灯）を防止するために不在判定が一定時間連続した場合に照明を消灯する. この消灯までの遅延時間は照明在不在制御の省エネルギー効果を左右する. 5.3.1 節で照明在不在制御の処理概要を述べ, 5.3.2 節で消灯遅延時間について説明する.

5.3.1　処理概要

照明在不在制御は, 人感センサが在判定した場合に照明を点灯する制御手法でオフィスや一般家庭に広く導入されている. 照明在不在制御は常時点灯方式に対して 20%〜46% の省エネルギー効果があることが述べられており [119, 123–125], 更なる省エネルギーの実現に向けて研究が進められている.

図 5.2 に照明在不在制御の処理フローチャートを示す. 照明在不在制御では, 人感センサの在室者検知結果と人感センサが最後に不在判定してから経過した時間（以降, 不在判定経過時間と呼称する）に基づいて照明を点灯制御する. 人感センサが在判定した場合には, 照明を点灯, もしくは点灯維持する. 人感センサが不在判定した場合には, 不在判定経過時間に基づいて照明を点灯制御する. 不在判定経過時間が消灯遅延時間以上のときに照明を消灯, もしくは消灯維持し, 消灯遅延時間未満のときに照明を点灯維持する. なお, 照明在不在制御開始時の不在判定経過時間は消灯遅延時間の値と等しく設定する.

消灯遅延時間が照明の省エネルギー効果を左右する重要なパラメータであることを図 5.3 を用いて説明する. 図 5.3 (a) は実際の在室状態, 図 5.3 (b), (c) は消灯遅延時間をそれぞれ 2, 5 min に設定した照明在不在制御の点灯状態を例示する. 消灯遅延時間が 2 min の

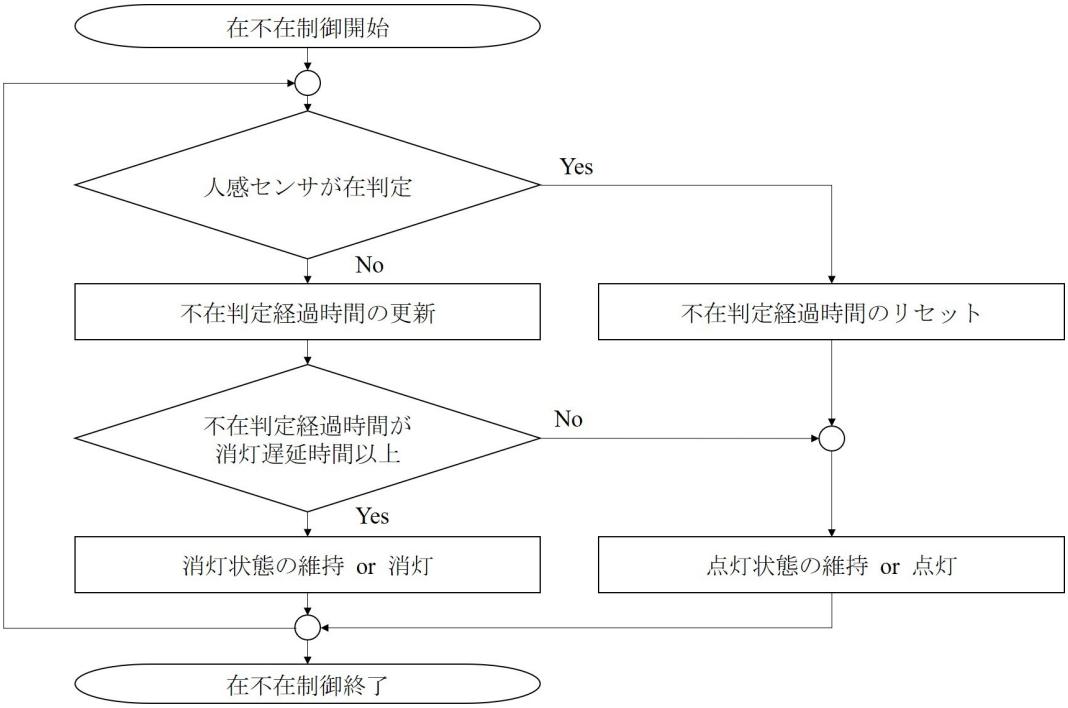


図 5.2: 照明在不在制御のフローチャート

照明在不在制御（図 5.3 (b)）は, 消灯遅延時間が 5 min の照明在不在制御（図 5.3 (c)）と比較して, 誤消灯の発生を抑えた上で照明を小まめに消灯できていることが分かる. 消灯遅延時間が 5 min の照明在不在制御（図 5.3 (c)）は, 在室者が不在にも関わらず照明を不必要に点灯する時間が長い. 在室者の快適性を損なうことなく照明在不在制御の省エネルギー効果を向上させるために, 誤消灯の発生を抑えた上で消灯遅延時間を短縮できる技術が求められている.

5.3.2　消灯遅延時間

文献 [126] で消灯遅延時間を 20 min から 5 min まで減少させることで省エネルギー効果が最大 13% 増加する事例が確認されているように, 消灯遅延時間は重要なパラメータである. そこで, 消灯遅延時間の設定方法を調査すると, 文献 [118, 119] では図 5.4 (1) の上側に細矢印に示す, 人感センサが連続して不在判定した時間（以降, 不在判定時間と呼称する）が利用されている. 不在判定時間のみを用いて消灯遅延時間を設定することで, 在室者の動作の大きさや頻度に合わせて消灯遅延時間を自動で設定できる. しかし, 図 5.4 (2) に示す実際の在室状態が加味されていないため, 誤消灯の発生しない消灯遅延時間を厳密に

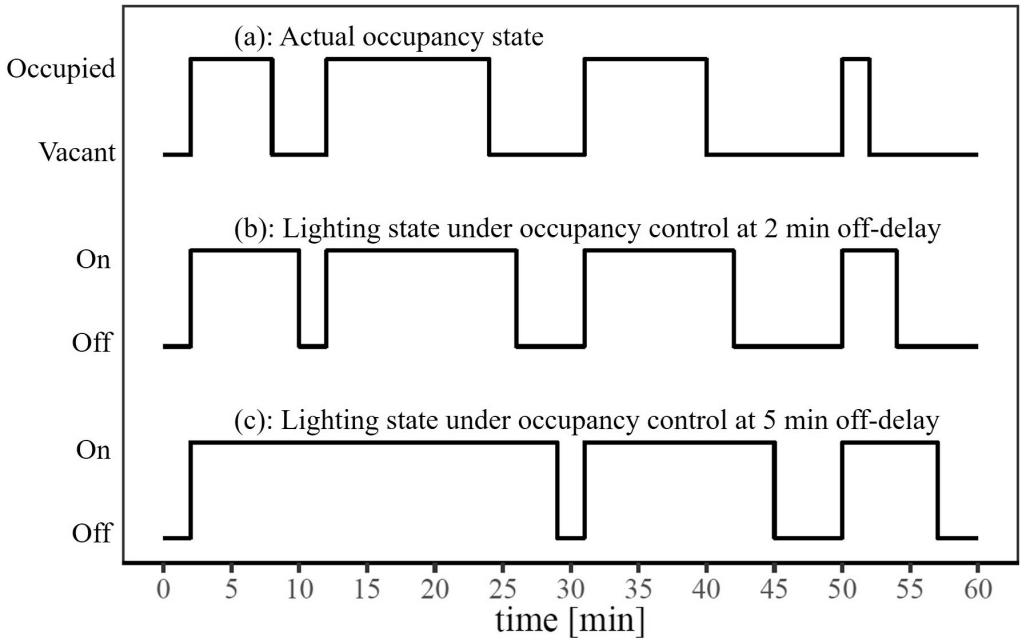


図 5.3: 照明在不在制御の例

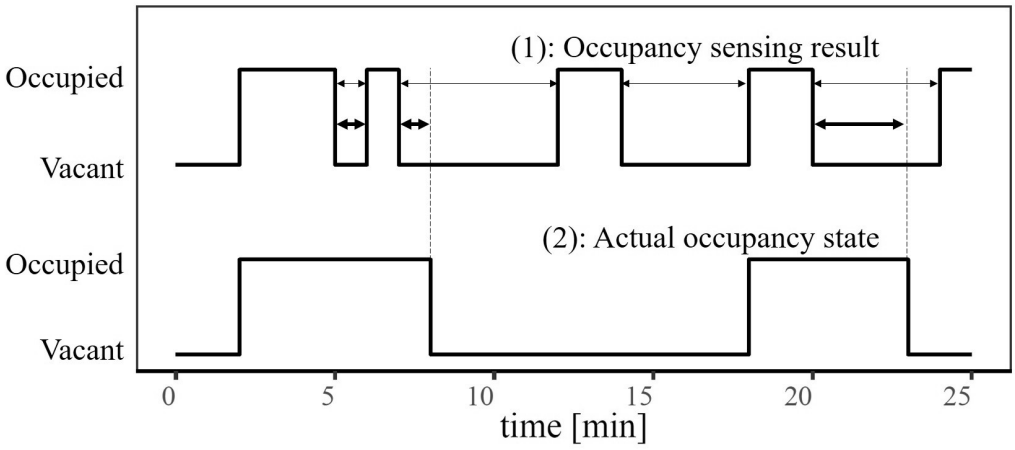


図 5.4: 不在判定時間と消灯遅延時間の例

求めることができない.

一方, 図 5.4 (1) の下側に太矢印で示す在室者が存在するにも関わらず誤って不在判定した時間（以降, 誤不在判定時間と呼称する）を長期に渡って観察し, 発生した誤不在判定時間の最大値を求めることで誤消灯の発生を抑えられる消灯遅延時間を評価することができる. 文献 [118, 119] と異なり, 在室者の検知結果と実際の在室状態との比較が必要で, 消灯遅延時間を自動で設定することができないが, 誤消灯の回避に必要な消灯遅延時間を厳密に評価できる.

消灯遅延時間の自動設定手法

本節では, 文献 [119] に基づいて提案された文献 [118] による消灯遅延時間の自動設定手順を示す. 観測された不在判定時間に基づいて消灯遅延時間を自動設定するための各ステップを以下に記す.

1. ビン幅 1 min で観測された 1 min 以上の不在判定時間の相対度数を生成して, 以下の確率密度分布（一般化パレート分布）で近似する.

$$\phi_\vartheta(v') = \begin{cases} (\vartheta v' + 1)^{-\frac{\vartheta+1}{\vartheta}} & (\vartheta \neq 0) \\ \exp(-v') & (\text{otherwise}) \end{cases} \tag{5.4}$$

2. 前ステップで近似した確立密度分布 $\phi_\vartheta(v')$ の累積密度分布を以下の式で求める.

$$\Phi_\vartheta(t') = \int_{-\infty}^{t'} \phi_\vartheta(v')dv' \tag{5.5}$$

$$= \begin{cases} 1 - (\vartheta t' + 1)^{-1/\vartheta} & (\vartheta \neq 0) \\ 1 - \exp(-t') & (\text{otherwise}) \end{cases} \tag{5.6}$$

3. 適応パラメータ $\kappa'(\in [0\%, 100\%])$ を用いて以下の式で消灯遅延時間を適応する.

$$\Phi_\vartheta(t'_{\text{ad}}) = \kappa' \tag{5.7}$$

$$t'_{\text{ad}} = \Phi_\vartheta^{-1}(\kappa') \tag{5.8}$$

なお, t'_{ad} は自動設定された消灯遅延時間, Φ_ϑ^{-1} は累積密度分布の逆関数を示す.

5.4　評価実験

本節では, 差分領域分割に基づく人感センサの有効性を確認するために赤外線センサとの比較実験を行う. まず, 5.4.1 節で人感センサを設置する環境や評価項目を実験条件として示し, 5.4.2 節で実験結果に基づき差分領域分割に基づく人感センサの有効性を確認する. 続いて, 5.4.3 節で赤外線センサや差分領域分割に基づく人感センサの利点や課題を考察する.

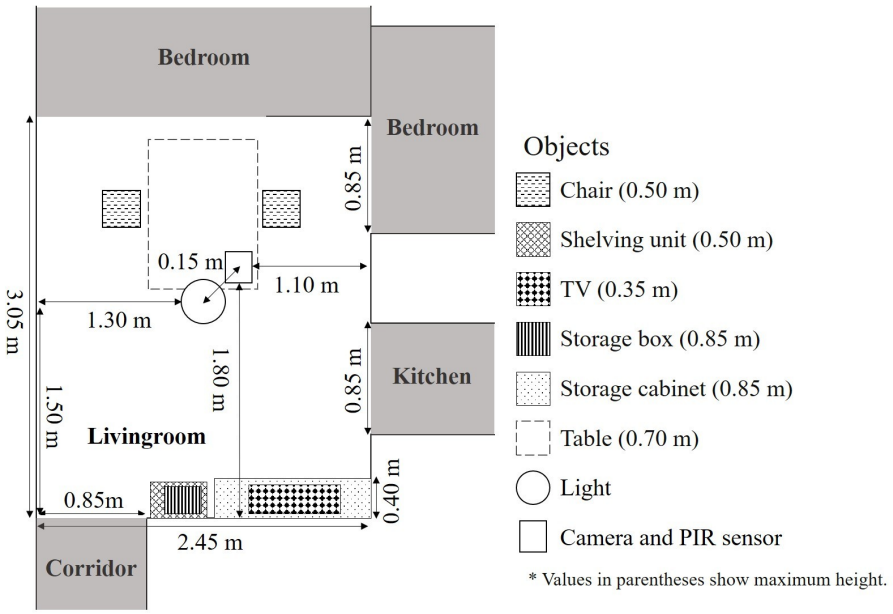


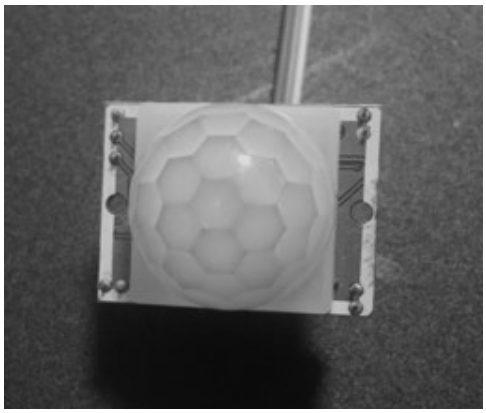
図 5.5: リビングルームのレイアウト（PIR sensor: 赤外線センサ, Camera: カメラモジュール）

5.4.1　実験条件

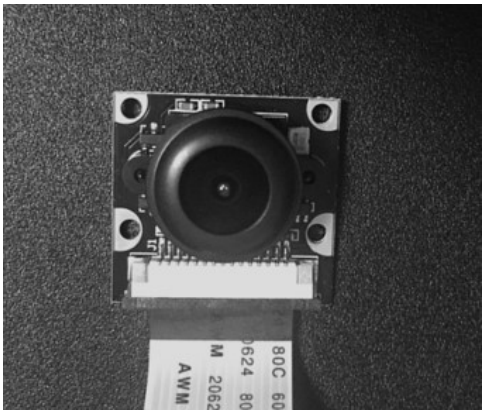
実験環境

本実験では, 2 名の 20 代社会人が居住する住宅におけるリビングルームの天井に監視カメラと赤外線センサを設置する. 図 5.5 にリビングルームのレイアウトを示す. 監視カメラと赤外線センサはリビングルームの中央部に位置する照明から 0.15 m 離れた位置に設置した. リビングルームにはベッドルーム, キッチン, 廊下が隣接する. リビングルーム内に在室する被験者は椅子に座って行動する傾向がある.

赤外線センサは文献 [127] で使用されている画角 120 度の SainSmart 製 HC-SR501 を使用した. リビングルーム内における実際の在室状態を表す正解データを作成するために, 文献 [128] で使用されている OmniVision 製カメラモジュール OV5647 に画角 130 度のレンズを装着し, 320×240 画素の監視カメラ画像を撮影した. 赤外線センサとカメラモジュールの外観をそれぞれ図 5.6 (a), (b) に示す. 赤外線センサとカメラモジュールの詳細な仕様は文献 [129, 130] に示されている. 赤外線センサの在室者検知結果と監視カメラ画



(a) HC-SR501



(b) OV5647 とカメラレンズ

図 5.6: 利用モジュール

表 5.1: 実験スケジュール

日付	開始時刻	終了時刻	実験時間 [h]
2019/1/6	08:00	22:30	13.00
2019/1/12	13:00	18:00	5.00
2019/1/14	18:00	23:15	5.25
2019/1/19	08:00	23:00	14.50
2019/1/20	12:30	22:15	9.50
2019/1/26	13:00	19:30	6.50
2019/1/27	09:00	21:30	12.00
2019/2/2	10:00	20:45	10.75

像はいずれも 100 ms 周期で実験データとして保存される.

差分領域分割に基づく人感センサの処理アルゴリズムとして, 5.2.1 節で説明した Sarker らの手法を用いる. 実験期間終了後に保存された監視カメラ画像を用いて, PC 上で Sarker らの手法を動作させることで差分領域分割に基づく人感センサを再現した. 赤外線センサと差分領域分割に基づく人感センサのパラメータは実験的に最適値を求める.

実験データは 2019/1/6〜2019/2/2 の期間を対象に収集した. 被験者が長時間在室する休日を中心に, 表 5.1 に示すスケジュールで実験を行った. 被験者の都合上, 一部の実験データを除外して 4 列目に示す合計 76.50 h (= T とする) 分の実験データを解析対象とした.

表 5.2: リビングルームで観察された挙動例 (2019/1/6)

時間	行動	在率 [%]	在室状態変化数
08:00-08:59	食事, TV 視聴	37.7	19
09:00-09:59	食事, デバイス操作, TV 視聴	86.5	14
10:00-10:59	デバイス操作, 歯磨き	58.1	5
11:00-11:59	デバイス操作, 飲水	26.5	10
12:00-12:29	デバイス操作, 飲水	81.5	7
13:00-13:59	食事, デバイス操作, 歯磨き, TV 視聴	94.9	9
14:00-14:59	デバイス操作	36.5	12
15:00-15:59	PC・デバイス操作	92.2	4
16:00-16:59	PC・デバイス操作	93.6	3
18:00-18:59	食事, デバイス操作, TV 視聴	87.6	22
19:00-19:59	食事, TV 視聴	88.2	9
20:00-20:59	スキンケア, デバイス操作, TV 視聴	86.3	11
21:00-21:59	食事, TV 視聴	59.6	4
22:00-22:29	歯磨き, TV 視聴	94.9	1

表 5.1 に示す実験データにリビングルームにおける普段の行動を全て網羅していることを被験者に確認済みである. 参考として, リビングルーム内で観察された被験者 A と B の挙動を表 5.2 に示す. 本実験環境下で観察された被験者の挙動は, 一般的な住宅のリビングルーム内に存在する在室者の挙動から逸脱しないことを関係者間で確認済みである. 表中の在率は在室者が存在した時間の割合を示し, 在室状態変化数は不在状態から在状態に推移した回数を示す. またデバイス操作とは, タブレット/スマートフォン/パソコンのいずれかを操作したことを意味する. なお, 被験者からの希望で実験から除外した 17:00-17:59 の実験データで観測された被験者の挙動例は表 5.2 から省いた.

評価項目

差分領域分割に基づく人感センサの有効性を確認するための評価項目 **(I)**〜**(IV)** を以下に列挙する.

(I) 在室者検知精度

時刻 t における在状態 (ξ_{o}) と不在状態 (ξ_{f}) が正しく分類された在室状態の正解データ u'_t を監視カメラ画像に基づいて人手で作成した. 正解データと人感センサの在室者検知結果を比較することで在室者検知精度を評価する.

赤外線センサの検知範囲や, 監視カメラ画像の撮影範囲には, リビングルームと接するベッドルーム, キッチンを僅かに含むが, 本実験では在室者がリビングルームに存在する状態を在状態と定義した. 赤外線センサの検知範囲や, 監視カメラ画像の撮影範囲は概ね等しく, 範囲内に僅かに含まれるリビングルーム以外の領域は全て部屋間の移動のために使用されていた. したがって, リビングルーム以外の部屋が在室者検知精度に与える影響は小さい.

在室者検知精度を示す指標として, 3.4.1 節で定義されている Precision((3.11) 式), Recall((3.12) 式), F-measure((3.13) 式) を用いる. 3.4 節の評価実験と同様に F-measure を総合的な評価指標とする. なお, TP, FP, TN, FN は, 時刻 t ($\in [0,T]$) における在室者検知結果 y_t^{ξ} と正解データ u'_t の比較に基づいて以下のように定義される.

- TP: 正しく在判定した回数

$$\mathrm{TP} = \sum_{t=0}^{T-1} \delta(y_t^{\xi}, \xi_{\mathrm{o}}) \cdot \delta(u'_t, \xi_{\mathrm{o}}) \tag{5.9}$$

- FP: 誤って在判定した回数

$$\mathrm{FP} = \sum_{t=0}^{T-1} \delta(y_t^{\xi}, \xi_{\mathrm{o}}) \cdot \delta(u'_t, \xi_{\mathrm{f}}) \tag{5.10}$$

- TN: 正しく不在判定した回数

$$\mathrm{TN} = \sum_{t=0}^{T-1} \delta(y_t^{\xi}, \xi_{\mathrm{f}}) \cdot \delta(u'_t, \xi_{\mathrm{f}}) \tag{5.11}$$

- FN: 誤って不在判定した回数

$$\mathrm{TN} = \sum_{t=0}^{T-1} \delta(y_t^{\xi}, \xi_{\mathrm{f}}) \cdot \delta(u'_t, \xi_{\mathrm{o}}) \tag{5.12}$$

表 5.3: 人感センサや LED 照明の最大・最小の定格電力

パラメータ	最小値 [W]	最大値 [W]
$\varepsilon_{\mathrm{pir}}$	0.013	2.0
$\varepsilon_{\mathrm{img}}$	1.0	4.0
$\tilde{\varepsilon}$	21.3	46.0

(II) 誤不在判定時間

誤消灯が発生しない消灯遅延時間を求めるために, 5.3.2 節で図 5.4 を用いて定義した誤不在判定時間を評価する. 収集した実験データには被験者の普段の基本的な行動が全て網羅されているため, 誤不在判定時間の最大値を消灯遅延時間として設定することで誤消灯を回避できる. そこで, 本実験においては誤不在判定時間の最大値を消灯遅延時間とし, 在室時の誤消灯を回避する.

(III) 省エネルギー効果

評価項目 **(II)** で求めた消灯遅延時間を用いて再現した照明在不在制御の点灯時間を算出し, 常時点灯方式に対する省エネルギー効果 Ψ を以下で評価する.

$$\Psi(\varrho, \tilde{\varrho}) = \frac{\tilde{\varrho} - \varrho}{\tilde{\varrho}} \tag{5.13}$$

なお, ϱ と $\tilde{\varrho}$ は照明在不在制御と常時点灯方式による照明点灯時間を意味する.

(IV) 消費電力量

評価項目 **(III)** の省エネルギー効果は照明点灯時間に基づいて評価されており, 照明と人感センサ自体が消費する電力量の合算値に基づく評価がされていない. そこで, 人感センサを用いた照明在不在制御の消費電力量を以下で評価する.

$$E_{\mathrm{con}}(\varepsilon, \tilde{\varepsilon}, \varrho, \tilde{\varrho}) = \tilde{\varepsilon}\varrho + \varepsilon\tilde{\varrho} \tag{5.14}$$

なお, ε と $\tilde{\varepsilon}$ はそれぞれ人感センサと照明の消費電力を示す.

本研究では商用化されている人感センサと照明の定格消費電力を参照して照明在不在制御の消費電力量を評価する. 表 5.3 に天井設置可能な赤外線センサと差分領域分割に基づく人感センサの定格消費電力の最小値 ($\varepsilon_{\mathrm{pir}}^{\mathrm{min}}$, $\varepsilon_{\mathrm{img}}^{\mathrm{min}}$) と最大値 ($\varepsilon_{\mathrm{pir}}^{\mathrm{max}}$, $\varepsilon_{\mathrm{img}}^{\mathrm{max}}$) を示す. また, 図

5.5 のリビングルームで照明として使用されているアイリスオーヤマ製 CL6D-5.0 と同等な, 6 畳向け LED 照明の定格消費電力の最小値と最大値 ($\tilde{\varepsilon}^{\min}$, $\tilde{\varepsilon}^{\max}$) も示す. 表 5.3 を得るために 7 社の赤外線センサ, 4 社の差分領域に基づく人感センサ, 15 社の照明を調査した.

差分領域分割に基づく人感センサを用いた照明在不在制御が消費電力量の観点で有効であるには以下を満たす必要がある.

$$E_{\mathrm{con}}(\varepsilon_{\mathrm{img}}, \tilde{\varepsilon}, \varrho_{\mathrm{img}}, \tilde{\varrho}) - E_{\mathrm{con}}(\varepsilon_{\mathrm{pir}}, \tilde{\varepsilon}, \varrho_{\mathrm{pir}}, \tilde{\varrho}) < 0 \tag{5.15}$$

なお, $\varepsilon_{\mathrm{pir}}$ と $\varepsilon_{\mathrm{img}}$ は, それぞれ赤外線センサと差分領域分割に基づく人感センサの消費電力を示す. ϱ_{pir} と ϱ_{img} は, それぞれ赤外線センサと差分領域分割に基づく人感センサを用いた照明在不在制御の点灯時間を示す.

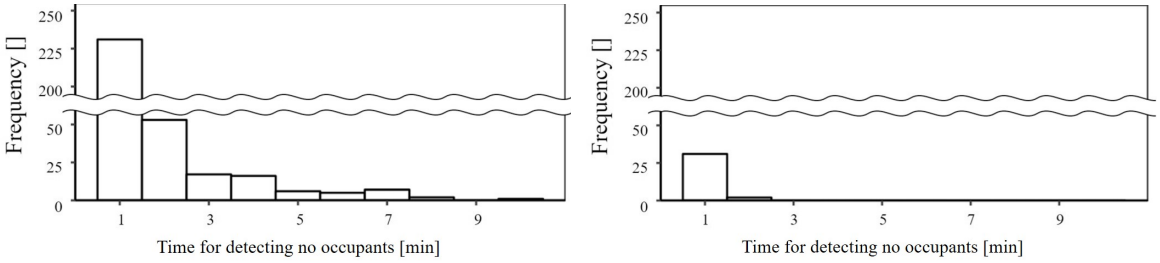
(5.15) 式に基づいて照明在不在制御の消費電力量を比較評価した文献 [36, 38] について解説する. (5.15) 式の左辺は以下のように展開される.

$$E_{\mathrm{con}}(\varepsilon_{\mathrm{img}}, \tilde{\varepsilon}, \varrho_{\mathrm{img}}, \tilde{\varrho}) - E_{\mathrm{con}}(\varepsilon_{\mathrm{pir}}, \tilde{\varepsilon}, \varrho_{\mathrm{pir}}, \tilde{\varrho}) = \tilde{\varepsilon}(\varrho_{\mathrm{img}} - \varrho_{\mathrm{pir}}) + \tilde{\varrho}(\varepsilon_{\mathrm{img}} - \varepsilon_{\mathrm{pir}}) \tag{5.16}$$

上式の前項は照明の消費電力量の差分値を表し, 後項は人感センサの消費電力量の差分値を表す.

文献 [36] は, (5.16) 式の前項を Building Energy Management System (BEMS) で計測された実測値に基づいて評価し, 後項を定格消費電力に基づく計算値を用いて評価した. 具体的には, $\tilde{\varrho}$ は実験時間 (=T), $\varepsilon_{\mathrm{img}}$ は商用化されている差分領域分割に基づく人感センサの定格消費電力 (=[$\varepsilon_{\mathrm{img}}^{\min}\varepsilon_{\mathrm{img}}^{\max}$]), $\varepsilon_{\mathrm{pir}}$ は実験を実施したオフィスに設置された赤外線センサの定格消費電力 (= 0.25 W) を用いて評価した. 文献 [36] の評価方法は, 実測値と計算値が含まれている点で改善の余地があった. また, 実験を実施したオフィスに設置された赤外線センサの定格消費電力のみを用いて (5.16) 式の後項を評価した点で実験環境に依存した結果が得られる可能性があった.

一方, 文献 [38] は全て計算値に基づいて (5.16) 式を評価した. 前項の $\tilde{\varepsilon}$ は商用化されている照明の定格消費電力 (=[$\tilde{\varepsilon}^{\min}$, $\tilde{\varepsilon}^{\max}$]), $\varrho_{\mathrm{img}}, \varrho_{\mathrm{pir}}$ は在室者検知結果に基づいて再現した照明在不在制御の点灯時間（評価項目 **(III)** と同様）を用いて評価した. 後項の $\tilde{\varrho}$ は実験時間 (=T), $\varepsilon_{\mathrm{img}}, \varepsilon_{\mathrm{pir}}$ は商用化されている人感センサの定格消費電力 (=[$\varepsilon_{\mathrm{img}}^{\min}$, $\varepsilon_{\mathrm{img}}^{\max}$], [$\varepsilon_{\mathrm{pir}}^{\min}$, $\varepsilon_{\mathrm{pir}}^{\max}$]) を用いて評価した. 文献 [38] では, 実験環境に依存しない結果を得るために商用化されている人感センサと照明の定格消費電力に基づき (5.15) 式が評価されている. 本論文も, 実験環境に依存しない評価結果を得ることを目的に文献 [38] と同様の方法で (5.15) 式



(a) 赤外線センサ　　(b) 差分領域分割に基づく人感センサ

図 5.7: 誤不在判定時間のヒストグラム

を評価する.

5.4.2　実験結果

本節では, 5.4.1 節で述べた評価項目 **(I)**〜**(IV)** に対する評価結果を示す. まず, 評価項目 **(II)** に基づいて誤不在判定時間の最大値を求める. 続いて, 評価項目 **(I)** に基づいて在室者検知精度を評価する. 次に, 評価項目 **(III)** に基づいて省エネルギー効果を考察する. 更に, 評価項目 **(IV)** に基づいて消費電力量を評価する.

誤不在判定時間

赤外線センサと差分領域分割に基づく人感センサで発生した誤不在判定時間のヒストグラムをそれぞれ図 5.7 (a), (b) に示す. なお, 発生頻度が高い 1 min 以下の誤不在判定時間は図 5.7 から省略した. 図 5.7 を見る限り, 赤外線センサと比べて差分領域分割に基づく人感センサで発生する誤不在判定時間は全体的に短い傾向がある. 実験期間中に発生した誤不在判定時間の最大値は, 赤外線センサと差分領域分割に基づく人感センサにおいてそれぞれ 10.58 min, 2.66 min で, 差分領域分割に基づく人感センサは赤外線センサに対して誤不在判定時間の最大値を 7.92 min (74.8%) 短縮した.

ここで赤外線センサにおいて比較的長い誤不在判定時間が発生した時間帯に着目すると, 在室者はタブレット/スマートフォン/パソコン操作, 食事, 筆記作業に代表される腕や指レベルの細かい動作を行っており, それらを赤外線センサが見逃し, 在判定できなかったことに起因して誤不在判定時間が増大することが分かった. 一方, 差分領域分割に基づく人感センサは在室者の細かい動作を見逃す割合が低く, 誤不在判定時間を短縮できた.

表 5.4: 在室者検知精度

(a) 消灯遅延時間導入前

指標	赤外線センサ	差分領域分割に基づく人感センサ
Precision [%]	88.8	96.3
Recall [%]	28.1	48.4
F-measure [%]	42.7	64.4

(b) 消灯遅延時間導入後

指標	赤外線センサ	差分領域分割に基づく人感センサ
Precision [%]	66.9	73.7
Recall [%]	100.0	100.0
F-measure [%]	80.2	84.8

以降, 5.4.1 節の評価項目 **(II)** で述べた通り, 消灯遅延時間の設定値として上記で求めた誤不在判定時間の最大値を用いる. 消灯遅延時間の設定値として誤不在判定時間の最大値を用いることで, 実験期間中に誤消灯が発生することを回避できる.

本実験データには被験者の行動が網羅されているため, 本実験で求めた誤不在判定時間の最大値を消灯遅延時間に設定することで実験期間外においても本リビングルームにおける誤消灯を回避できる可能性がある. しかし, 在室者の動作の大きさや頻度は曜日や季節によって変動するため, 実験期間外においても誤消灯が発生しない消灯遅延時間を求めるには更なる実験を要すると考えられる.

在室者検知精度

表 5.4 に各人感センサの在室者検知精度を示す. なお, 消灯遅延時間導入後の評価結果を示した表 5.4 (b) では, 人感センサが不在判定した場合においても消灯遅延時間以内であれば在判定したものと見なして精度を算出した.

まず, 消灯遅延時間導入前の在室者検知精度を示した表 5.4 (a) に着目すると, 差分領域分割に基づく人感センサは全ての評価指標において赤外線センサを上回る結果を得た. 特に総合的な評価指標である F-measure は 21.7% 高く, 差分領域分割に基づく人感センサの

表 5.5: 各日程の在率と省エネルギー効果

Date	在率 [%]	Ψ_{pir} [%]	Ψ_{img} [%]	$\Psi_{img} - \Psi_{pir}$ [%]
1/6	72.0	9.96	15.99	6.03
1/12	88.9	2.44	4.17	1.73
1/14	46.9	12.87	24.9	12.08
1/19	44.1	31.46	41.89	10.43
1/20	70.9	4.78	8.05	3.27
1/26	22.3	70.19	73.56	3.37
1/27	49.4	7.95	13.16	5.21
2/2	33.8	32.39	45.15	12.76
平均	52.8	18.62	28.30	9.68

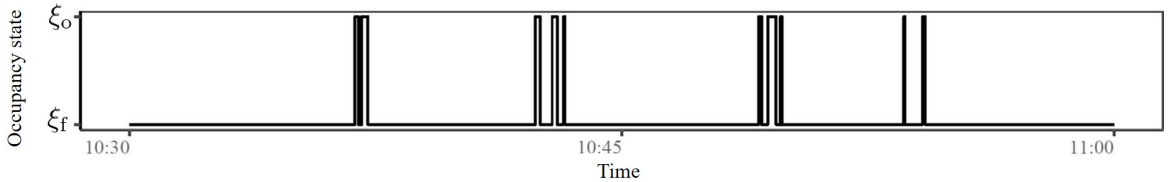
在室者検知精度面における有効性を確認できた.

次に, 消灯遅延時間導入後の在室者検知精度を示す表 5.4 (b) に着目すると, 消灯遅延時間導入前と同様に差分領域分割に基づく人感センサは全ての評価指標において赤外線センサを上回る結果を得た. 消灯遅延時間導入後には人感センサ間で性能差が小さくなったものの, 差分領域分割に基づく人感センサの F-measure は赤外線センサと比較して 4.6% 高く, 差分領域分割に基づく人感センサは消灯遅延時間導入前と同様に赤外線センサに対して在室者検知精度面で有効性を維持していた.

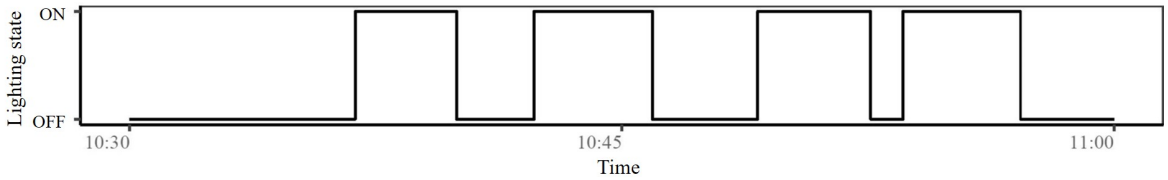
省エネルギー効果

各人感センサを用いた照明在不在制御の省エネルギー効果 ((5.13) 式) を表 5.5 に示す. なお, 表中の在率は実験時間中に在室者が存在した割合を示し, Ψ_{pir} と Ψ_{img} は赤外線センサと差分領域分割に基づく人感センサの省エネルギー効果を示す. 差分領域分割に基づく人感センサは赤外線センサと比べて省エネルギー効果を平均 9.68% 向上できることが分かった. なお, この平均値は日程単位の平均値ではなく, 全日程（76.50 h 分）の実験データを対象に評価した値である.

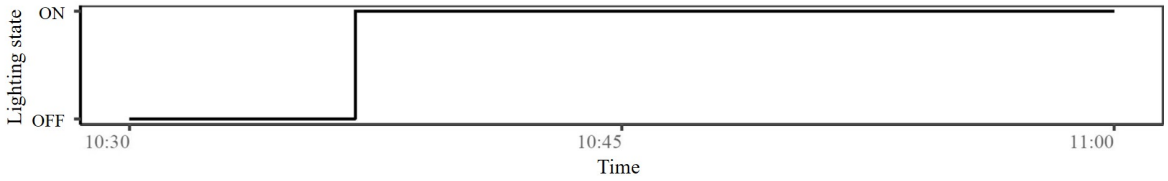
図 5.8 (a) に在室者が在と不在を繰り返した 2019/1/19 の 10:30〜11:00 における実際の在室状態を示す. 図 5.8 (b), (c) に各人感センサを用いた照明在不在制御の点灯状態を示す.



(a) 実際の在室状態



(b) 赤外線センサを用いた照明在不在制御



(c) 差分領域に基づく人感センサを用いた照明在不在制御

図 5.8: 在室者検知結果例（2019/1/19 10:30–11:00）

差分領域分割に基づく人感センサは, 消灯遅延時間の短縮効果に起因して照明点灯時間を低減できていることが分かる. 図 5.8 のような傾向は異なる日程, 時間帯においても確認されており, 差分領域分割に基づく人感センサによる省エネルギー効果向上は広く得られると考えられる.

差分領域分割に基づく人感センサによる省エネルギー効果の向上率 ($\Psi_{\mathrm{img}} - \Psi_{\mathrm{pir}}$) は, 1.73% から 12.76% まで変動した. 在率が 70% を超える照明を消灯できる機会が少ない日程は省エネルギー効果の向上率が平均値を下回った. 1/26, 27 の在率はそれぞれ 23.3%, 49.4% で照明を消灯できる機会が比較的多く存在したが, 差分領域分割に基づく人感センサによる省エネルギー効果の向上率が平均値以下であった. これは, 扇風機によって乾燥されている洗濯物を差分領域分割に基づく人感センサが在室者として誤検知し, 照明を不要に点灯することに起因した.

一方, 差分領域分割に基づく人感センサによる省エネルギー効果の向上率が平均値を上回る日程の在率は 33.8%〜46.9% で, 在室者が断続的にリビングルームを利用する場面が

散見され, 照明を消灯できる機会が多く存在した. 差分領域分割に基づく人感センサを導入するコストに見合った省エネルギー効果を得るためにも在室状況や部屋の用途を事前に調査する必要がある.

また, 全実験時間中に在室者が存在した時間の総和は 40.42 h で, 差分領域分割に基づく人感センサを用いた照明在不在制御の点灯時間の総和は 54.85 h であった. 差分領域分割に基づく人感センサを用いた照明在不在制御は, 在室者が存在しないのにも関わらず照明を不要に点灯した時間が点灯時間の 26.3% を占めていることが分かった. これは差分領域分割に基づく人感センサが誤って在判定したことや, 誤消灯を回避するために消灯遅延時間を 2.66 min に設定したことが原因で, 省エネルギー効果の更なる向上を実現するためにも, 差分領域分割に基づく人感センサの高精度化が必要である.

消費電力量

全実験時間中 ($T = 76.50$ h) における照明在不在制御の点灯時間は $\varrho_{\mathrm{pir}} = 60.22$ h, $\varrho_{\mathrm{img}} = 54.85$ h だった. 表 5.3 を参照すると, 商用化されている赤外線センサ, 差分領域分割に基づく人感センサ, 照明の定格消費電力の値域はそれぞれ [0.013 W, 2.0 W], [1.0 W, 4.0 W], [21.3 W, 46.0 W] である.

(5.14) 式に上記の値を代入し, 赤外線センサと差分領域分割に基づく人感センサを用いた照明在不在制御が実験期間中に消費した電力量を求めた結果, それぞれ 1.28〜2.92 kWh と 1.24〜2.82 kWh だった. 差分領域分割に基づく人感センサは, センサ自体の消費電力が高いことに起因して赤外線センサと比べて消費電力量を低減できない可能性が示唆された.

そこで, (5.15) 式を用いて差分領域分割に基づく人感センサの消費電力について考察する. (5.15) 式に上述した値を代入し, 差分領域分割に基づく人感センサが消費電力量の観点から有効である条件式を以下で導出する.

$$\begin{aligned} &E_{\mathrm{con}}(\varepsilon_{\mathrm{img}}, \tilde{\varepsilon}, \varrho_{\mathrm{img}} = 54.85\ \mathrm{h}, \tilde{\varrho} = 76.50\ \mathrm{h}) \\ &\qquad - E_{\mathrm{con}}(\varepsilon_{\mathrm{pir}} = 0.013\ \mathrm{W}, \tilde{\varepsilon}, \varrho_{\mathrm{pir}} = 60.22\ \mathrm{h}, \tilde{\varrho} = 76.50\ \mathrm{h}) < 0 \end{aligned} \tag{5.17}$$

図 5.9 に (5.17) 式の境界条件を黒色実線で示す. 照明と差分領域分割に基づく人感センサの消費電力を示す $\tilde{\varepsilon}$ と $\varepsilon_{\mathrm{img}}$ で構成される平面空間は 2 つの領域 A, B に分割される. 図中に灰色で示す領域 A は (5.17) 式を満たし, 領域 B は (5.17) 式を満たさない.

図 5.9 に点線で囲われる定格消費電力の範囲内 ($[\varepsilon_{\mathrm{img}}^{\mathrm{min}}, \varepsilon_{\mathrm{img}}^{\mathrm{max}}]$, $[\tilde{\varepsilon}^{\mathrm{min}}, \tilde{\varepsilon}^{\mathrm{max}}]$) には上述した領域 B が含まれる. そのため, 差分領域分割に基づく人感センサを用いた照明在不在制御

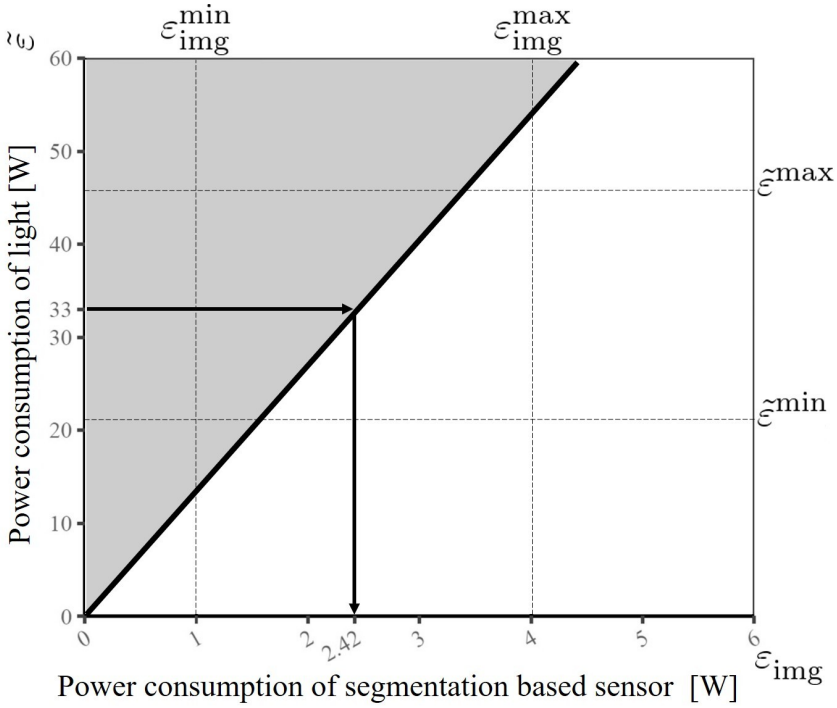


図 5.9: (5.17) 式の可視化（灰色領域: 領域 A, 白色領域: 領域 B）

が赤外線センサを用いた照明在不在制御と比べて消費電力量の観点で有効であるには, 照明の消費電力に合わせて消費電力の低い差分領域分割に基づく人感センサを選定する必要がある. 例えば, 本実験環境下で用いた照明 ($\tilde{\varepsilon}$ = 33.0 W) を使用した際に, 差分領域分割に基づく人感センサが消費電力量の観点から有効であるには $\varepsilon_{\mathrm{img}} < 2.42$ W を満たす必要があった.

5.4.3　考察

本節では, まず, 評価実験で確認された差分領域分割に基づく人感センサの利点と課題を考察する. 続いて, 赤外線センサと差分領域分割に基づく人感センサに共通して確認された課題を述べる.

差分領域分割に基づく人感センサの利点

前節の評価実験を通じて, まず, 差分領域分割に基づく人感センサは在室者検知精度を 21.7% 以上向上できることを確認した. 続いて, 腕や指レベルの細かい動作を検知できることに起因して誤不在判定時間を 74.8% 短縮した. 更に, 差分領域分割に基づく人感セン

サを用いた照明在不在制御は誤消灯の発生を抑えた上で省エネルギー効果を 9.68% 向上させた. 他にも, 差分領域分割に基づく人感センサは検知範囲を任意に区切ったエリア単位で在室者を検知できるため, 設置台数を抑えつつ在室者直上の照明のみを点灯できる利点を有している [38].

差分領域分割に基づく人感センサの課題

差分領域分割に基づく人感センサは, 在室者が不在の場合でも自動で動作する物体を在室者として誤検知する課題があり, この課題によって照明在不在制御の省エネルギー効果が低下する. 例えば, 文献 [131] で首振り扇風機や PC のスクリーンセーバーを誤検知することが示唆されている. 本実験で発生した誤検知を図 5.10 に示す. 図 5.10 (a) に示す首振り扇風機によって乾燥されている洗濯物や, 図 5.10 (b) に示すディスプレイが自動オフされたタブレット端末が誤検知された. 一方, 赤外線センサの在室者検知結果に着目したところ, 図 5.10 に起因した誤検知は確認されなかった. 以上より, 上記の誤検知は差分領域分割に基づく人感センサ特有の課題であると考えられる. 近年では CNN を用いたクラス認識や物体検出が広く研究されており, 学習データの整理に要するコストや処理時間に関する課題があるものの, それらを応用することで図 5.10 に示す誤検知を回避できる可能性がある. 他にも, 5.4.2 節における消費電力量の評価で述べたように, 差分領域分割に基づく人感センサ自体の消費電力が高い課題があるため, 低消費電力化を実現可能な回路設計技術の導入を検討する必要がある.

共通課題

本実験環境下では, 図 5.5 の中央に位置する椅子に在室者が座った際に, 赤外線センサと差分領域分割に基づく人感センサに対して在室者の下半身がテーブルで隠れるオクルージョンが発生して上半身の動作のみで在室者が検知された. オクルージョンが発生しているため着座時に下半身が動作する頻度を確認できないが, 上半身に加えて下半身の動作を捉えることができれば, スマートフォンやタブレット操作に代表される腕や指レベルの細かい動作が高頻度で発生する着座時の誤不在判定を低減できる可能性がある. その他の在室者検知精度低下が懸念されるオクルージョンは未確認だが, 赤外線センサや差分領域分割に基づく人感センサを設置する際には設置位置や什器配置を工夫してオクルージョンを回避すべきである.

1/6/2019 16:08:37.04　　1/6/2019 16:08:41.14

(a) 洗濯物

1/20/2019 15:57:53.46　　1/20/2019 15:57:53.56

(b) タブレット端末

図 5.10: 差分領域分割に基づく人感センサの誤検知例

また 5.4.2 節における誤不在判定時間の評価では, 誤消灯が発生しない消灯遅延時間の最小値を正解データに基づいて求めたが, 実運用下で正解データを作成するのは困難である. そこで通常, 省エネルギー効果が低減するが, 誤消灯を回避できる十分に長い消灯遅延時間を設定することで実運用される. 差分領域分割に基づく人感センサの省エネルギー効果を実運用下で向上させるためにも, 在室者検知結果のみを利用して誤消灯を回避可能な消灯遅延時間の最小値を自動推定する手法 [118, 119] の導入が求められている.

本実験で得られた在室者検知結果に Nagy らの手法 [118] を適用することで, 赤外線センサと差分領域分割に基づく人感センサに対して誤消灯の発生しない消灯遅延時間の最小値を推定した結果, それぞれ 3.04〜6.46 min と 3.51〜8.31 min が得られた. なお, 適応パラメータ κ' の値は文献中に記載されている推奨値を用いた.

Nagy らの手法で推定した誤消灯の発生しない消灯遅延時間の最小値と, 実験期間中に観察された誤消灯の発生しない消灯遅延時間を比較する. まず, 赤外線センサに誤消灯を発生させないためには消灯遅延時間を 10.58 min 以上にする必要があり, Nagy らの手法で求めた値は誤消灯の発生を回避できない. 次に, 差分領域分割に基づく人感センサに誤消灯を発生させないためには消灯遅延時間 2.66 min 以上にする必要があり, Nagy らの手法で求めた値は在室者が不在にも関わらず照明を不要に点灯する時間を増大させる.

Nagy らの手法はオフィスに導入された赤外線センサ向けに適応パラメータ κ' の値が設

定されているため, 本実験のように住宅のリビングルームに設置された人感センサに応用する際には適応パラメータの値を見直す必要があると考えられる. 文献 [118] に記載された適応パラメータ κ' の推奨値は [85%, 95%] であるが, 本実験環境下で誤消灯が発生しない消灯遅延時間の最小値が得られる適応パラメータ κ' は赤外線センサで 97.8%, 差分領域分割に基づく人感センサで 75.5% だった.

5.5　まとめ

本研究では, 関連研究で十分に評価されていない差分領域分割に基づく人感センサの有効性を確認し, その利点や課題を明らかにした. 差分領域分割に基づく人感センサの有効性を確認するために, 一般家庭で収集した 8 日間の実験データを用いて安価で普及している赤外線センサと比較評価した.

まず, 差分領域分割に基づく人感センサは赤外線センサに対して 21.7% の在室者検知精度改善を実現し, 人感センサが誤って不在判定し続ける時間（誤不在判定時間）を 7.92 min (74.8%) 短縮できることが分かった. 差分領域分割に基づく人感センサは, スマートフォンやパソコン操作に代表される腕や指レベルの細かい動作を見逃す割合が低かったため赤外線センサと比較して誤不在判定時間を短縮できた.

次に, 各人感センサに誤消灯が発生しない消灯遅延時間を設定し, 実験データを基に照明在不在制御を再現することで照明点灯時間を評価した. その結果, 差分領域分割に基づく人感センサを用いた照明在不在制御は赤外線センサを用いた照明在不在制御に対して省エネルギー効果を 9.68% 向上できることが分かった.

続いて, 商用化されている人感センサや照明の定格消費電力に基づいて実験期間中に照明と人感センサで消費した電力量の合算値を評価した. その結果, 赤外線センサによる照明在不在制御の消費電力量は 1.28〜2.92 kWh, 差分領域分割に基づく人感センサによる照明在不在制御の消費電力量は 1.24〜2.82 kWh だった. 差分領域分割に基づく人感センサ自体の消費電力が赤外線センサと比べて高い傾向にあるため, 消費電力量を低減できない可能性があることが本実験結果より示唆された. 本実験環境下において差分領域分割に基づく人感センサが赤外線センサと比べて照明在不在制御の消費電力量を低減するには, 差分領域分割に基づく人感センサ自体の定格消費電力を 2.42 W 以下に抑える必要があることが明らかになった.

更に, 差分領域分割に基づく人感センサの改良課題について考察した. 差分領域分割に

基づく人感センサには自動で動作する物体（例えば, 首振り扇風機）を誤って在判定する課題が存在するため, 誤検知の発生頻度低減に向けてクラス認識や物体検出の導入を検討する必要がある. 他にも, 誤消灯が発生しない最短の消灯遅延時間を自動推定する手法の開発も今後の課題として挙げられる.

本研究で得られた差分領域分割に基づく人感センサの有効性に関する知見は, 1 戸の住宅のリビングルームで 8 日間に亘って実施した実験結果に基づいている点に留意する必要がある. 用途の異なる部屋や家族構成の異なる住宅に差分領域分割に基づく人感センサを導入し, 長期に亘る評価実験を行うことで本研究で得られた実験結果とは異なるものが得られる可能性がある.

第 6 章

結論

本論文では, 画像を構成する各画素に領域種別を割り当てる画像領域分割に関する研究成果を述べた. 幅広い分野で研究や実用化が進められている画像領域分割には様々な従来手法が存在するが, 本研究では導入コスト増大とユーザビリティ低下の課題を回避可能な事前知識に基づく手法を研究対象とした. 事前知識に基づく手法のアルゴリズムは入力画像や領域種別に関する知見に基づいて設計されるため, 導入先の処理目的に依存する. そこで, 導入先となる応用システムの特性を加味して事前知識に基づく手法を高度化する研究成果を述べた.

以下, 本論文で述べた研究成果を要約する. まず, ユーザが周囲を撮影した情景画像の建造物領域を分割する手法を提案した. 背景領域が画像の上部や下部に分布する傾向があることを事前知識とする提案手法には, 入力画像の各画素を色情報で分類したクラスタの解析処理を導入した. ZuBuD に収録されている 106 枚の情景画像を用いた比較実験の結果, 同じく事前知識に基づく手法に分類される上野らの手法に対して, 提案手法は 86.4% の平均計算時間低減と 10.1% の領域分割精度向上を実現した. 領域分割精度低下を引き起こす上野らの手法の課題を提案手法は全て回避した.

続いて, インターネットオークション等で用いられる商品画像の商品領域を分割する手法を提案した. 背景領域が画像の外縁部に分布する傾向があることを事前知識とする提案手法には, 周辺画素を併合しながら画像の外縁部を拡大させる処理を導入した. インターネットオークションで実際に使用されている 341 枚の商品画像を用いた比較実験の結果, 同じく事前知識に基づく手法に分類される Cheng らの手法に対して, 提案手法は 76.7% の平均計算時間低減と 3.2% の領域分割精度向上を実現した. 領域分割精度低下を引き起

こす Cheng らの手法の課題を提案手法は回避した.

次に, 監視カメラ画像の差分領域を分割することで在室者を検知する人感センサの有効性を実環境下で収集したデータを用いて評価し, 関連研究で明らかにされていない差分領域分割に基づく人感センサの利点と課題を整理した. 一般家庭で収集した 8 日間の実験データを用いて比較評価した結果, 差分領域分割に基づく人感センサは安価で普及している赤外線センサに対して 21.7% の在室者検知精度改善を実現した. また, 差分領域分割に基づく人感センサは, 照明に対する在不在制御の省エネルギー効果を赤外線センサと比べて 9.68% 向上できることが分かった. しかし, 差分領域分割に基づく人感センサ自体の消費電力が赤外線センサと比べて高く, 照明在不在制御の消費電力量削減を実現できない可能性が示唆された.

以下, 本論文の内容を章単位で要約する. 第 1 章では, まず, 画像認識技術に関する研究動向を述べた. 次に, クラス認識, 物体検出, 画像領域分割を画像認識技術の代表例として説明し, 研究対象とする画像領域分割とその他の画像認識技術の差異を述べた. 続いて, 画像領域分割には領域種別単位で領域分割するセマンティックセグメンテーションと個体単位で領域分割するインスタンスセグメンテーションが存在することを述べた. 本論文では, 隣接した領域種別が同じ物体を区別できないが計算コストの低いセマンティックセグメンテーションを研究対象とすることを説明した. 更に, 本論文の研究目的を述べた.

第 2 章では, 関連研究に基づいて画像領域分割の従来手法を説明した. まず, 画像領域分割を定式化し, 本研究に関連する建造物領域分割, 商品領域分割, 差分領域分割を例示した. 次に, 画像領域分割の従来手法を局所特徴に基づく手法, インタラクティブ手法, 事前知識に基づく手法に分類し, 各手法を代表例に基づき説明した. 局所特徴に基づく手法は, CNN に基づく FCN と SegNet を代表例として挙げた. インタラクティブ手法は, グラフ理論に基づく GrabCut とセルオートマトンに基づく GrowCut を代表例として挙げた. 事前知識に基づく手法は, 肺領域を分割する Zhang らの手法, 黒子領域を分割する Tushar らの手法, 食事領域を分割する杉山らの手法を代表例として挙げた. 続いて, 各手法の利点と課題を整理して, 研究対象とする事前知識に基づく手法とその他の手法との差異を述べた. 本論文では導入コスト増加やユーザビリティ低下を回避可能な事前知識に基づく手法を研究対象とすることを説明し, 第 3 章以降の研究方針を述べた.

第 3 章では, 情景画像から建造物領域を分割する手法を提案し, その有効性を示した. まず, 事前知識に基づいて建造物領域を分割する上野らの手法を説明した. 上野らの手法の

処理概要をフローチャートに基づき詳細に述べた上で, 領域分割精度が低下する上野らの手法の課題を説明した. 続いて, 背景領域が画像の上部や下部に分布することを事前知識とする提案手法の処理概要をフローチャートに基づき詳細に述べた. 次に, ZuBuD に収録されている 106 枚の情景画像を用いた比較実験に基づき提案手法の有効性を確認した. 提案手法は, 上野らの手法と比べて領域分割精度を 10.1% 向上し, 平均計算時間を 86.4% 低減した. 他にも, 更なる精度向上に向けて提案手法で発生した課題を整理した.

第 4 章では, 商品画像から商品領域を分割する手法を提案し, その有効性を示した. まず, 商品領域の分割に応用されている Cheng らの手法を説明した. Cheng らの手法の処理概要をフローチャートに基づき詳細に述べた上で, 領域分割精度が低下する Cheng らの手法の課題を説明した. 続いて, 商品領域が画像の中央部に分布することを事前知識とする提案手法の処理概要をフローチャートに基づき詳細に述べた. 次に, インターネットオークションで実際に使用されている 341 枚の商品画像を用いた比較実験に基づき提案手法の有効性を確認した. 提案手法は, Cheng らの手法と比べて領域分割精度を 3.2% 向上し, 平均計算時間を 76.7% 低減した. 他にも, 提案手法の各処理が領域分割精度向上の実現に向けて必要な処理であることを実験的に考察した.

第 5 章では, 関連研究で十分に評価されていない差分領域分割に基づく人感センサの利点や課題を一般家庭で収集した実験データを用いて明らかにした. 人感センサの在室者検知精度は, 電気設備に対する在不在制御の省エネルギー効果や消費電力量を左右するため, 本研究では在室者検知精度, 省エネルギー効果, 消費電力量の観点から人感センサを評価した. なお, 省エネルギー効果や消費電力量を評価する際には, 一般家庭やオフィスの消費エネルギー内訳に占める割合が高い照明設備の在不在制御を対象とした. まず, 商用化が進められている差分領域分割に基づく人感センサを関連研究に基づいて詳細に説明し, 照明在不在制御の処理概要をフローチャートに基づき詳細に述べた. 続いて, 差分領域分割に基づく人感センサの有効性を確認する実験の実験条件を説明した. 次に, 一般家庭で収集した実験データに基づき差分領域分割に基づく人感センサの利点と課題を整理した. 差分領域分割に基づく人感センサは在室者検知精度を 21.7%, 省エネルギー効果を 9.68% 向上した. しかし, 差分領域分割に基づく人感センサ自体の消費電力が比較的高く, 照明在不在制御の消費電力量を削減できない可能性が示唆された.

以下, 今後の展望を示す. 第 3 章では, 背景領域が画像の上部や下部に分布する傾向があることを事前知識とする建造物領域分割を提案し, 領域分割精度と計算時間の観点から有

効性を示した．提案手法は, 建造物領域の輪郭線情報を特徴とする建造物クラス認識や物体検出の高性能化に寄与できる．また, 建造物クラス認識と物体検出は情景画像を撮影したユーザの現在位置を測位する技術に応用されるため, 現在位置測位の性能向上に提案手法が寄与できる可能性がある．他にも, 提案手法は建造物領域分割に基づいて 3 次元モデルを生成する技術や, 外壁を検査する技術の高性能化を実現できる可能性がある．

第 4 章では, 背景領域が画像の外縁部に分布する傾向があることを事前知識とする商品領域分割を提案し, 領域分割精度と計算時間の観点から有効性を示した．提案手法は, 商品領域の輪郭線情報を特徴とする商品クラス認識や物体検出の高性能化に寄与できる．また, 商品クラス認識や物体検出は商品画像に含まれる商品の名称や関連情報を自動取得するインターネットオークションの機能に応用されるため, インターネットオークションの効率化に提案手法が寄与できる可能性がある．他にも, 提案手法は商品画像に不要な背景を除去し, 商品に合わせた背景に変更する画像加工技術の高性能化を実現できる．商品に合わせて適切な背景を用いることで購入者の購買意欲を高められる可能性がある．

第 5 章は, 関連研究で明らかにされていない差分領域分割に基づく人感センサの利点と課題を整理した点で有用で, 差分領域分割に基づく人感センサの普及に寄与できる可能性がある．また, 差分領域分割に基づく人感センサは照明や空調設備の在不在制御に用いられるため, 本章で得られた知見は一般家庭やオフィスで消費するエネルギー量の抑制に貢献できる．他にも, 差分領域分割に基づく人感センサの更なる改良に貢献できる．例えば, 本章では差分領域分割に基づく人感センサの消費電力が高いことを指摘したため, 差分領域分割に基づく人感センサに低消費電力化を実現する回路設計技術が今後導入される可能性がある．

参考文献

[1] R. J. Gove. CMOS image sensor technology advances for mobile devices. *High Performance Silicon Imaging*, pp. 185–240, Oct. 2019.

[2] D. J. Duke, T. Knast, B. Thethy, L. Gisler and D. Edgington-Mitchell. CMOS image sensor technology advances for mobile devices. *Measurement Science and Technology*, Vol. 30, No. 7, p. 075403, Jun. 2019.

[3] 須川成利. イメージセンサ技術. 電子情報通信学会論文誌 C, Vol. 100, No. 10, pp. 474–482, Oct. 2017.

[4] P. Jerram and K. Stefanov. CMOS and CCD image sensors for space applications. *High Performance Silicon Imaging*, pp. 255–287, Oct. 2019.

[5] 教育機器編集委員会. 産業教育機器システム便覧. 日科技連出版社, 1972.

[6] M. Sanchez-Garcia, R. Martinez-Cantin and J. J. Guerrero. Semantic and structural image segmentation for prosthetic vision. *Plos One*, Vol. 15, No. 1, p. e0227677, Jan. 2020.

[7] 柳井啓司. 一般物体認識の現状と今後. 情報処理学会論文誌コンピュータビジョンとイメージメディア, Vol. 48, No. SIG16 (CVIM19), pp. 1–24, Nov. 2017.

[8] 藤吉弘亘, 山下隆義. 深層学習による画像認識. 日本ロボット学会誌, Vol. 35, No. 3, pp. 180–185, May 2017.

[9] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, Vol. 60, No. 2, pp. 91–110, Nov. 2004.

[10] J. Cervantes, F. Garcia-Lamont, L. Rodriguez-Mazahua and A. Lopez. A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing*, Vol. 408, No. 30, pp. 189–215, May 2020.

[11] V. N. Vapnik. An overview of statistical learning theory. *IEEE Transactions on Neural*

Networks, Vol. 10, No. 5, pp. 988–999, Sep. 1999.

[12] W. S. McCulloch and W. Pitts. A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, Vol. 5, No. 4, pp. 115–133, Dec. 1943.

[13] A. Shrestha and A. Mahmood. Review of deep learning algorithms and architectures. *IEEE Access*, Vol. 7, pp. 53040–53065, Apr. 2019.

[14] M. Pak and S. Kim. A review of deep learning in image recognition. *Proceedings of 2017 4th International Conference on Computer Applications and Information Processing Technology (CAIPT)*, pp. 1–3, Aug. 2017.

[15] A. Dhillon and G. K. Verma. Convolutional neural network: A review of models, methodologies and applications to object detection. *Progress in Artificial Intelligence*, Vol. 9, No. 2, pp. 85–112, Dec. 2020.

[16] C. Leng, H. Zhang, B. Li, G. Cai, Z. Pei and L. He. Local feature descriptor for image matching: A survey. *IEEE Access*, Vol. 7, pp. 6424–6434, Dec. 2018.

[17] L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu and M. Pietikainen. Deep learning for generic object detection: A survey. *International Journal of Computer Vision*, Vol. 128, No. 2, pp. 261–318, Oct. 2019.

[18] B. Alexe, T. Deselaers and V. Ferrari. Measuring the objectness of image windows. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 34, No. 2, pp.2189–2202, Jan. 2012.

[19] M. H. Hesamian, W. Jia, X. He and P. Kennedy. Deep learning techniques for medical image segmentation: Achievements and challenges. *Journal of Digital Imaging*, Vol. 32, No. 4, pp. 582–596, May 2019.

[20] A. Jahedsaravani, M. Massinaei and M. H. Marhaban. An image segmentation algorithm for measurement of flotation froth bubble size distributions. *Measurement*, Vol. 111, pp. 29–37, Dec. 2017.

[21] N. A. Muhadi, A. F. Abdullah, S. K. Bejo, M. R. Mahadi and A. Mijic. Image segmentation methods for flood monitoring system. *Water*, Vol. 12, No. 6, p. 1825, Apr. 2020.

[22] M. Siam, M. Gamal, M. Abdel-Razek, S. Yogamani, M. Jagersand and H. Zhang. A comparative study of real-time semantic segmentation for autonomous driving. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 587–

597, Jun. 2018.

[23] J. Dehais, M. Anthimopoulos and S. Mougiakakou. Food image segmentation for dietary assessment. *Proceedings of 2nd International Workshop on Multimedia Assisted Dietary Management*, pp. 23–28, Oct. 2016.

[24] M. Caballo, J. M. Boone, R. Mann and I. Sechopoulos An unsupervised automatic segmentation algorithm for breast tissue classification of dedicated breast computed tomography images. *Medical Physics*, Vol. 45, No. 6, pp. 2542–2559, Apr. 2020.

[25] S. N. Aslan, A. Uçar and C. Guzeliş. Semantic segmentation for object detection and grasping with humanoid robots. *Proceedings of 2020 Innovations in Intelligent Systems and Applications Conference (ASYU)*, pp. 1–6, Oct. 2020.

[26] T. Cane and J. Ferryman. Evaluating deep semantic segmentation networks for object detection in maritime surveillance. *Proceedings of 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pp. 1–6, Nov. 2018.

[27] 北村嘉郎, 石川博. 高階エネルギー最小化による医用画像セグメンテーション. 電子情報通信学会論文誌 D, Vol. 101, No. 1, pp. 3–26, Jan. 2020.

[28] F. Lateef and Y. Ruichek. Survey on semantic segmentation using deep learning techniques. *Neurocomputing*, Vol. 338, No. 21, pp. 321–348, Apr. 2019.

[29] A. M. Hafiz and G. M. Bhat. A survey on instance segmentation: State of the art. *International Journal of Multimedia Information Retrieval*, Vol. 9, pp. 171–189, Jul. 2020.

[30] A. Khamparia and K. M. Singh. A systematic review on deep learning architectures and applications. *Expert Systems*, Vol. 36, No. 3, p. e12400, Jan. 2019.

[31] 二神拓也, 早坂昇. 色クラスタ解析に基づく建造物領域の自動抽出. システム制御情報学会論文誌, Vol. 32, No. 9, pp. 349–356, Sep. 2019.

[32] T. Futagami, N. Hayasaka and T. Onoye. Fast and robust building extraction based on HSV color analysis using color segmentation and GrabCut. *SICE Journal of Control, Measurement, and System Integration*, Vol. 13, No. 3, pp. 97–106, May 2020.

[33] T. Futagami and N. Hayasaka. Automatic extraction of building regions by using color clustering. *Proceedings of SICE Annual Conference 2019*, pp. 415–419, Sep. 2019.

[34] I. Masakazu, T. Futagami, N. Hayasaka and T. Onoye. Acceleration of automatic building extraction via color-clustering analysis. *IEICE Transactions on Fundamentals*

of Electronics, Communications and Computer Science, Vol. 103, No. 12, pp. 1599–1602, Dec. 2020.

[35] T. Futagami and N. Hayasaka. Automatic product region extraction based on analysis of images uploaded to C2C online market. *Journal of Organizational Computing and Electronic Commerce*, Vol. 30, No. 4, pp. 323–334, Jul. 2020.

[36] 二神拓也, 早坂昇, 尾上孝雄. 画像センサを用いた照明在不在制御の有効性評価. システム制御情報学会論文誌, Vol. 33, No. 5, pp. 139–148, May 2020.

[37] T. Futagami, T. Yano, C. Huang and T. Enohara. Experimental evaluation for daylight-linked gradation lighting control using image-based motion sensors. *IEEJ Transactions on Electrical and Electronic Engineering*, Vol. 15, No. 5, pp. 723–732, May 2020.

[38] 二神拓也, 矢野亨, 黄静君, 榎原孝明. 画像センサによる照明個別制御の有効性検証. 電気学会論文誌 C, Vol. 138, No. 11, pp. 1362–1374, Nov. 2018.

[39] H. Yu, Z. Yang, L. Tan, Y. Wang, W. Sun, M. Sun and Y. Tang. Methods and datasets on semantic segmentation: A review. *Neurocomputing*, Vol. 304, No. 23, pp. 82–103, Aug. 2018.

[40] F. Cao and Q. Bao. A survey on image semantic segmentation methods with convolutional neural network. *Proceedings of 2020 International Conference on Communications, Information System and Computer Engineering (CISCE)*, pp. 458–462, Jul. 2020.

[41] J. He, C. S. Kim and C. C. J. Kuo. Interactive segmentation techniques: Algorithms and performance evaluation. *Springer Science and Business Media*, Aug. 2013.

[42] H. Ramadan, C. Lachqar and H. Tairi. A survey of recent interactive image segmentation methods. *Computational Visual Media*, Vol. 6, No. 4, pp. 1–30, Dec. 2020.

[43] J. Dogra, S. Jain and M. Sood. Gradient-based kernel selection technique for tumour detection and extraction of medical images using graph cut. *IET Image Processing*, Vol. 14, No. 1, pp. 84–93, Jan. 2019.

[44] Z. Yanqing, X. Heru, J. Xinhua, G. Xiaojing, Z. Lina, D. Yajuan and W. Siyu. Sheep image segmentation and contour extraction of multi-scale watershed and fuzzy c-means based on graph cuts. *Acta Technica*, Vol. 62, No. 6A, pp. 145–156, Dec. 2019.

[45] 堀江勝大, 鹿嶋雅之, 佐藤公則, 渡邊睦. 動的輪郭モデルによる心臓内壁輪郭自動抽出

法. 映像情報メディア学会誌, Vol. 63, No. 11, pp. 1675–1679, Nov. 2009.

[46] 堀江勝大, 鹿嶋雅之, 佐藤公則, 渡邊睦. 標準左心室形状モデルの適用による心臓内壁輪郭自動抽出法. 電気学会論文誌 C, Vol. 130, No. 9, pp. 1537–1545, Sep. 2010.

[47] S. Zhang, Y. Zhao and P. Bai. Object localization improved GrabCut for lung parenchyma segmentation. *Procedia Computer Science*, Vol. 131, pp. 1311–1317, May 2018.

[48] F. I. Tushar. Automatic skin lesion segmentation using GrabCut in HSV colour space. *ArXiv Preprint ArXiv:1810.00871*, Sep. 2015.

[49] 杉山春樹, 森川茶民, 相澤清晴. 極値点情報と GrabCut による食事画像領域分割. 映像情報メディア学会誌, Vol. 66, No. 5, pp. J179–J181, May 2012.

[50] A. Apicella, F. Donnarumma, F. Isgro and R. Prevete. A survey on modern trainable activation functions. *ArXiv Preprint ArXiv:2005.00817*, May. 2020.

[51] V. Nair and G. E. Hinton. Rectified linear units improve restricted boltzmann machines. *Proceedings of 27th International Conference on Machine Learning*, pp.807–814, Jun. 2010.

[52] B. Yuan. Efficient hardware architecture of softmax layer in deep neural network. *Proceedings of 2016 29th IEEE International System-on-Chip Conference (SOCC)*, pp. 323–326, Sep. 2016.

[53] S. Albawi, T. A. Mohammed and S. Al-Zawi. Understanding of a convolutionalneural network. *Proceedings of International Conference on Engineering and Technology (ICET)*, pp. 1–6, Aug. 2017.

[54] 岡谷貴之. ディープラーニング. 映像情報メディア学会誌, Vol. 68, No. 6, pp. 139–148, Jul. 2014.

[55] E. Shelhamer, J. Long and T. Darrell. Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39, No. 4, pp. 640–651, Apr. 2017.

[56] V. Badrinarayanan, A. Kendall and R. Cipolla. SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39, No. 12, pp. 2481–2495, Dec. 2017.

[57] H. Noh, S. Hong and B. Han. Learning deconvolution network for semantic segmen-

tation. *Proceedings of IEEE International Conference on Computer Vision*, pp. 1520–1528, Dec. 2009.

[58] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *ArXiv Preprint ArXiv:1502.03167*, Feb. 2015.

[59] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, Vol. 15, No. 1, pp. 1929–1958, Apr. 2017.

[60] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *ArXiv Preprint ArXiv:1409.1556*, Sep. 2014.

[61] C. Rother, V. Kolmogorov and A. Blake. GrabCut: interactive foreground extraction using iterated graph cuts. *ACM Transactions on Graphics*, Vol. 23, No. 3, pp. 309–314, Aug. 2004.

[62] V. Vezhnevets and V. Konouchine. GrowCut: interactive multi-label N-D image segmentation by cellular automata; *Proceedings of Graphicon*, pp. 150–156, Jun. 2005.

[63] D. A. Reynolds. Gaussian mixture models. *Encyclopedia of Biometrics*, pp. 659–663, 2009.

[64] Y. BoyKov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transaction of Pattern Analysis and Machine Intelligence*, Vol. 26, No. 9, pp. 1124–1137, Jul. 2004.

[65] 大津展之. 判別および最小 2 乗規準に基づく自動しきい値選定法. 電子情報通信学会論文誌 D, Vol. 63, No. 4, pp. 349–356, Apr. 1980.

[66] L. Itti, C. Koch and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20, No. 11, pp. 1254–1259, Nov. 1998.

[67] S. Alqazzaz, X. Sun, X. Yang and L. Nokes. Automated brain tumor segmentation on multi-modal MR image using SegNet. *Computational Visual Media*, Vol. 5, No. 2, pp. 209–219, Apr. 2019.

[68] R. Manickam, S. K. Rajan, C. Subramanian, A. Xavi, G. J. Eanoch and H. R. Yesudhas. Person identification with aerial imaginary using SegNet based semantic segmentation. *Earth Science Informatics*, Vol. 13, No. 4, pp. 1293–1304, Sep. 2020.

[69] S. Aslan, G. Ciocca and R. Schettini. Semantic segmentation of food images for automatic dietary monitoring. *Proceedings of 2018 26th Signal Processing and Communications Applications Conference (SIU)*, pp. 1–4, May. 2018.

[70] N. Deepika and V. S. Variyar. Obstacle classification and detection for vision based navigation for autonomous driving. *Proceedings of 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 2092–2097, Sep. 2017.

[71] J. Åhlen and S. Seipei. Knowledge based single building extraction and recognition. *Recent Advances in Computer Engineering, Communications and Information Technology*, pp. 29–35, Dec. 2014.

[72] 上野大, 吉田大海, 飯國洋二. 多値の初期確信度を用いた建造物画像の自動 GrowCut. システム制御情報学会論文誌, Vol. 29, No. 6, pp. 266–274, Jun. 2016.

[73] M. M. Cheng, G. X. Zhang, N. J. Mitra, X. Huang and S. M. Hu. Global contrast based salient region detection. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 409–416, Mar. 2011.

[74] J. He, J. Feng, X. Liu, T. Cheng, T. H. Lin, H. Chung and S. F. Chang. Mobile product search with bag of hash bits and boundary reranking. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3005–3012, Jun. 2011.

[75] 上田泰佑. 製造者から見た照明制御の現状と展望. 電気設備学会誌, Vol. 38, No. 1, pp. 16–19, Jan. 2018.

[76] 渡邉泰則. 照明制御製造者が考える照明設備, 制御手法. 電気設備学会誌, Vol. 38, No. 1, pp. 12–15, Jan. 2018.

[77] A. Sarkar, M. Fairchild and C. Salvaggio. Integrated daylight harvesting and occupancy detection using digital imaging. *Sensors, Cameras and Systems for Industrial/Scientific Applications IX*, Vol. 6816, Feb. 2008.

[78] M. Antonino, M. Nicola, D. M. Claudio, B. Luciano and R. C. Fulvio. Office building occupancy monitoring through image recognition sensors. *International Journal of Safety and Security Engineering*, Vol. 9, No. 4, pp. 371–380, Nov. 2019.

[79] F. Manzoor, D. Linton and M. Loughlin. Occupancy monitoring using passive RFID technology for effcient building lighting control. *Proceeding of Fourth International*

EURASIP Workshop RFID Technology, pp. 83–88, Sep. 2012.

[80] F. Kaku, M. Miki, T. Hiroyasu, M. Yoshimi, S. Tanaka, T. Nishida, N. Kida, M. Akita, J. Tanisawa and T. Nishimoto. Construction of intelligent lighting system providing desired illuminance distributions in actual office environment. *Proceedings of International Conference on Artiffcial Intelligence and Soft Computing*, Jun. 2010.

[81] Z. Chen, C. Jiang and L. Xie. Building occupancy estimation and detection: A review. *Building and Environment*, Vol. 169, No. 15, pp. 260–270, Jun. 2018.

[82] C. de Bakker, M. Aries, H. Kort and A. Rosemann. Occupancy-based lighting control in open-plan office spaces: A state-of-the-art review. *Building and Environment*, Vol. 112, pp. 308–321, Feb. 2017.

[83] N. Gu, B. Yang and T. Zhang. S. Parnin and M. M. Rahman. Dynamic fuzzy background removal for indoor human target perception based on thermopile array sensor. *IEEE Sensors Journal*, Vol. 20, No. 1, pp. 67–76, Jan. 2020.

[84] S. Parnin and M. M. Rahman. Human location estimation using thermopile array sensor. *Proceeding of 6th International Conference on Mechatronics*, Vol. 260, pp. 1–6, Nov. 2017.

[85] X. Guo, D. K. Tiller, G. P. Henze and C. E. Waters. The performance of occupancy-based lighting control systems: A review. *Lighting Research and Technology*, Vol. 42, No. 4, pp. 415–431, Aug. 2010.

[86] National Lighting Product Information Program (NLPIP). Occupancy sensors: Motion sensors for lighting control. Troy, NY. Lighting Research Center, Oct. 1998.

[87] P. D. Groves. Shadow Matching: A New GNSS positioning technique for urban canyons. *Journal of Navigation*, Vol. 64, No. 3, pp. 417–430, Jul. 2011.

[88] Z. Jiang and P. D. Groves. GNSS NLOS and multipath error mitigation using advanced multi-constellation consistency checking with height aiding. *Proceedings of 25th International Technical Meeting of the Satellite Division of the Institute of Navigation*, pp. 79–88, Sep. 2012.

[89] W. Zhang and J. Kosecka. Localization based on building recognition. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Jun. 2005.

[90] Y. C. Chung, H. X. Tony and Z. He. Building recognition using sketch-based representations and spectral graph matching. *Proceedings of IEEE International Conference on Computer Vision*, pp. 2014–2020, Sep. 2009.

[91] 笹川耕一, 黒田伸一, 池端重樹. 平均隣接数に着目したしきい値決定法. 電子情報通信学会論文誌 D, Vol. 73, No. 3, pp. 360–366, Mar. 1990.

[92] 松尾賢一, 上田勝彦, 梅田三千雄. 適応しきい値法を用いた情景画像からの看板文字列領域の抽出. 電子情報通信学会論文誌 D, Vol. 80, No. 6, pp. 1617–1626, Jun. 1997.

[93] H. Shao, T. Svoboda and L. Van Gool. Zubud - Zurich building database for image based recognition. *ETH Z̈urich, Technical Report 260*, Jan. 2003.

[94] C. M. Bishop. Pattern recognition and machine learning. Springer, pp. 461–517, Aug. 2006.

[95] H. Attias. A variational bayesian framework for graphical models. *Advances in Neural Information Processing Systems 12*, Jun. 2000.

[96] D. L. Weakliem. A critique of the Bayesian information criterion for model selection. *Sociological Methods and Research*, Vol. 27, No. 3, pp. 359–397, Feb. 1994.

[97] J. B. MacQueen. Some methods for classification and analysis of multivariate observations. *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*, pp. 281–291, 1967.

[98] 西谷卓史. 変分混合ガウス分布を用いた視差のクラスタリングによる連続立体視画像中の移動体追跡. 画像の認識・理解シンポジウム (MIRU2011) 論文集, pp. 162–169, Jul. 2011.

[99] X. Li, K. Liu and Y. Dong. Superpixel-based foreground extraction with fast adaptive trimaps. *IEEE Transactions on Cybernetics*, Vol. 48, No. 9, pp. 2609–2619, Sep. 2017.

[100] X. Song, L. Zhou, Z. Li, J. Chen, B. Yan and L. Zeng. Interactive image segmentation based on hierarchical superpixels initialization and region merging. *Proceedings of 2014 7th International Congress on Image and Signal Processing*, pp. 410–414, Oct. 2014.

[101] S. Bhowal and B. C. Dhara. Segmentation of similar color objects using modified GrowCut method. *International Journal of Advanced Computer Research*, Vol. 3, No. 3, pp. 36–41, Sep. 2013.

[102] D. Corrigan, S. Robinson and A. Kokaram. Video matting using motion extended GrabCut. *Proceedings of IET 5th European Conference on Visual Media Production*, pp. 1–9, Nov. 2008.

[103] P. Shanmugavadivu and G. Thenmozhi. Detection of microcalcification in mammogram images using semi-automated texture based GrabCut segmentation. *Proceedings of 2012 International Conference on Emerging Trends in Science, Engineering and Technology (INCOSET)*, pp. 198–203, Dec. 2012.

[104] G. Brostow, J. Fauqueur and R. Cipolla. Semantic object classes in video: A high-definition ground truth database. *Pattern Recognition Letters*, Vol. 30, No. 2, pp. 88–97, Jan. 2009.

[105] W. Chin-Shan, C. Fei-Fei and C. Y. David. The influence of seller, auctioneer, and bidder factors on trust in online auctions. *Journal of Organizational Computing and Electronic Commerce*, Vol. 24, No. 1, pp. 36–57, Jan. 2014.

[106] H. Yamamoto and H. Ohshima. Proactive or reactive? Platform governance strategy in C2C marketplace. *Proceedings of 21th Pacific Asia Conference in Information Systems (PACIS 2017)*, p. 116, Sep. 2017.

[107] T. T. Ow, B. I. Spaid, C. A. Wood and S. Ba. Trust and experience in online auctions. *Journal of Organizational Computing and Electronic Commerce*, Vol. 28, No. 4, pp. 294–314, Oct. 2018.

[108] G. A. Cushen and M. S. Nixon. Mobile visual clothing search. *Proceedings of IEEE Conference on Multimedia and Expo Workshops*, pp. 1–6, Jul. 2013.

[109] J. He, T. H. Lin, J. Feng and S. F. Chang. Mobile product search with bag of hash bits. *Proceedings of ACM Multimedia 2011*, pp. 839–840, Nov. 2011.

[110] S. S. Tsai, D. Chen, V. Chandrasekhar, G. Takacs, N. M. Cheung, R. Vedantham, R. Grzeszczuk and B. Girod. Mobile product recognition. *Proceedings of ACM Multimedia 2010*, pp. 1587–1590, Oct. 2010.

[111] X. Shen, Z. Lin, J. Brandt and Y. Wu. Mobile product image search by automatic query object extraction. *Proceedings of European Conference on Computer Vision*, pp. 114–127, Oct. 2012.

[112] S. A. Oyewole and O. O. Olugbara. Product image classification using eigen colour

feature with ensemble machine learning. *Egyptian Informatics Journal*, Vol. 19, No. 2, pp. 83–100, Jul. 2018.

[113] S. Jia, X. Kong and G. Jin. Automatic fast classification of product-images with class-specific descriptor. *Journal of Electronics (China)*, Vol. 27, No. 6, pp. 808–814, Jul. 2010.

[114] P. Felzenszwalb and D. Huttenlocher. Efficient graph-based image segmentation. *International Journal of Computer Vision*, Vol. 59, No. 2, pp. 167–181, Sep. 2004.

[115] R. Adams and L. Bischof. Seeded region growing. *IEEE Transaction of Pattern Analysis and Machine Intelligence*, Vol. 16, No. 6, pp. 641–647, Jun. 1994.

[116] Z. Nagy, F. Y. Yong, M. Frei and A. Schlueter. Occupant centered lighting control for comfort and energy efficient building operation. *Energy and Buildings*, Vol. 94, No. 1, pp. 100–108, May 2015.

[117] L. L. Fernandes, E. S. Lee, D. L. DiBartolomeo and A. McNeil. Monitored lighting energy savings from dimmable lighting controls in The New York Times Headquarters Building. *Energy and Buildings*, Vol. 68, No. A, pp. 498–514, Jan. 2014.

[118] Z. Nagy, M. Hazas, M. Frei, D. Rossi and A. Schlueter. Illuminating adaptive comfort: dynamic lighting for the active occupant. *Proceedings of 8th Windsor Conference: Counting the Cost of Comfort in a Changing World*, Apr. 2014.

[119] V. Garg and N. K. Bansal. Smart occupancy sensors to reduce energy consumption. *Energy and Buildings*, Vol. 32, No. 1, pp. 81–87, Jun. 2000.

[120] J. Takahashi, S. Nishikado, H. Yamamoto, Y. Morimoto, M. Nagata and T. Enohara. Development of motion sensor with image sensing devices. *Proceeding of 6th Lighting Conference of China, Japan and Korea*, pp.289-290, Aug. 2013.

[121] U.S Energy Information Administration. 2012 Commercial Buildings Energy Consumption Survey: Energy Usage Summary. `https://www.eia.gov/consumption/commercial/reports/2012/energyusage/`（最終アクセス日: 2021 年 1 月 9 日）.

[122] 省エネルギーセンター. オフィスビルのエネルギー消費量の特徴. `https://www.eccj.or.jp/office_bldg/01.html`（最終アクセス日: 2021 年 1 月 9 日）.

[123] D. Maniccia, B. Rutledge, M. S. Rea and W. Morrow. Occupant use of manual lighting controls in private offices. *Journal of the Illuminating Engineering Society*, Vol. 28,

No. 2, pp. 42–56, Sep. 1991.

[124] T. Lashina, S. Vleuten-Chraibi, M. Despenic, P. Shrubsole, A. Rosemann and E. Loenen. A comparison of lighting control strategies for open offices. *Building and Environment*, Vol. 149, pp. 68–78, Feb. 2019.

[125] A. D. Galasiu, G. R. Newsham, C. Suvagau and D. M.Sander. Energy saving lighting control systems for open-plan offices: A field study. *Leukos*, Vol. 4, No. 1, pp. 7–29, Sep. 2007.

[126] B. V. Neida, D. Manicria and A. Tweed. An analysis of the energy and cost savings potential of occupancy sensors for commercial lighting systems. *Journal of the Illuminating Engineering Society*, Vol. 30, No. 2, pp. 111–125, Sep. 2001.

[127] J. Alvarez, A. Acero, S. Gutierrez, P. M. Rodrigo and A. Lay-Ekuakille. A low cost presence detection system for smart homes. *Proceeding of 2018 International Conference on Research in Intelligent and Computing in Engineering (RICE)*, pp. 1–6, Aug. 2018.

[128] A. R. Mead and K. M. Mosalam. Ubiquitous luminance sensing using the Raspberry Pi and camera module system. *Lighting Research Technology*, Vol. 49, No. 7, pp. 904–921, Nov. 2017.

[129] PerkinElmer Inc. LHI778 Datasheet. `https://html.alldatasheet.com/html-pdf/14901/PERKINELMER/LHI778/605/1/LHI778.html`（最終アクセス日: 2021 年 1 月 9 日）.

[130] OmniVision Inc. OV5647 Datasheet. `http://www.dragonwake.com/download/camera/ov5647/ov5647_full.pdf`（最終アクセス日: 2021 年 1 月 9 日）.

[131] 東芝インフラシステムズ株式会社. SMART EYE SENSOR MULTI (TSBHS01). `http://www.toshiba.co.jp/sis/imagesensor/index_j.htm`（最終アクセス日: 2021 年 1 月 9 日）.