

画像データからの自動領域抽出手法の考案

Proposing automated region extraction techniques from image data.

小林 秀^{*1} 野村 眞平^{*2} 李 石映雪^{*2} 菊田 遥平^{*1} 神津 友武^{*1}
 Shu Kobayashi Shinpei Nomura Shiyngxue Li Yohei Kikuta Tomotake Koza

^{*1} 有限責任監査法人トーマツ デロイトアナリティクス^{**}
 Deloitte Analytics, Deloitte Touche Tohmatsu LLC

^{*2} 株式会社リクルート住まいカンパニー
 Recruit Sumai Company Ltd.

Image data is widely used for web based services but few service uses effectively. To implement service, not only detecting the object but also extracting the object are important because information on object is more worthful. Many research focused on object detection in foreground but there is also worthful information in background. In this paper, we propose automated region extract process based on GMM and EM algorithm. Specifically we focused not on specific object such as human or car in foreground area but specific object such as floor or wall in background area. As an experimental data set, we use the image data of SUUMO, the largest housing information site. The proposed approach achieves high precision for detecting wood floors.

1. はじめに

本稿では、GMM(GaussianMixtureModel)を利用した画像データ中の領域分割手法をもとに、その短所である分割領域の手作業による指定を行わずに画像から特定の背景部分を抽出するための手法を提案し、不動産ポータルサイト SUUMO で実際にサービスで用いられている画像データを用いた実験を行い、その有用性を検討する。

近年、DeepLearning に対する関心の高まりと共に、画像データの活用に対する注目が高まっている。[Jagadeesh 2014]、[Jing 2015]らのように、Web サービスを中心に、画像データを活用したリコメンドに代表されるアルゴリズムが提供されている。

画像データに対して DeepLearning を用いる場合、対象画像中に映った前景となる特定の物体を判別、抽出することに利用されることが多い。一方で、画像データ中から一般には背景に相当する領域を、可能な限り抽出したい物体のみを対象として、自動的に抽出したい場面が存在する。例えば、本稿で使用する SUUMO の画像を対象とした場合、リビングの天井、壁、床、ドアといったものを切り出し、この色やパターンを物件探しの検索軸とすることが考えられる。この例を図 1 に示す。

一般に画像を領域分割する際には、矩形切出により対象物体を検知するアプローチと画像中の領域そのものを分割するアプローチが存在する。矩形切出の場合には対象物体の形状が矩形に収まっている必要があり、対象物が矩形に収まらないような画像中の広範囲に広がる場面には必ずしも適切な手法とはいえない。一方、矩形によらない領域分割手法のひとつに GMM が存在するが、手作業で抽出したい物体の領域を指定する必要があり、自動的な実行は難しい。

そこで、本研究ではサービスに用いられている実データを対象に、GMM による領域分割を行なったうえで、抽出したい物体の特徴の判別機を構築し適用することで、画像データ中から抽

出した対象物が存在する領域抽出の自動化アルゴリズムを考案した。この結果を用いることで、画像中の意図した領域の色や模様などの画像的な特徴を切り口とした検索機能へ利用可能となる。



図 1. 抽出対象とした領域の例。画像はウェブから収集した著作権フリーのものであるが、実際にサービスで使われる画像に近いものである。

2. 従来研究

画像からの領域分割に関するアプローチには大きく矩形により対象物を特定のうえ分割を実施する手法と矩形によらない手法が存在する。

2.1 矩形による領域分割手法

矩形による手法では、適切な大きさ矩形を設定し、画像データ中を走査のうえ分割したい対象物を特定する。このためには、テンプレート画像を利用したパターンマッチングや[Girshick 2014] らによる DeepLearning を利用した対象物の切り出しなどが存在する。

一方で、パターンマッチングによる分割には、難型となるパターンデータの作成に大きな工数が発生するため、作業工数面

連絡先: 小林秀, 有限責任監査法人トーマツ デロイトアナリティクス, shu.kobayashi@tohmatu.co.jp

** 本研究の内容は有限責任監査法人トーマツの公式見解を示すものではありません。

で実施が困難である。また、パターンマッチングでは切り出した対象の領域形状に沿って物体を切り出すことは難しい。

他方、DeepLearning を用いたアプローチの場合、対象物を矩形で切り出すことが一般的であり、切り出したい対象のみを切り出すことは難しい。また、学習データの作成には画像中から対象物の存在する座標を指定する必要がある、データ作成の工数や、モデル構築のための学習時間など、運用のためには大きな工数が必要となることが多い。

2.2 矩形によらない領域分割手法

矩形によらない領域分割手法として動的輪郭ベースのセグメント化[Chan 2001]、測地的距離ベースのカラーセグメンテーション [Protiere 2007]などの手法が存在する。一方でこれらの手法は前景と背景の分割が主要な目的であり、分割領域数が 2,3 程度を想定しているため今回の目的には合致しない。

任意の領域数を指定する際に、[Fu 2012]らによる色情報利用した GMM と EM アルゴリズムによるアプローチが存在する。GMM と EM アルゴリズムによる領域分割手法では色情報を入力とし、パラメータとして分割領域数を指定する。したがって、適切な領域分割数は試行錯誤のうえで設定する必要がある。また、領域分割の精度は、画像中に存在する切り出した対象物の情報を認識するわけではなく、与えられた画像中の色情報に基づいた分割を行なうため、必ずしも切り出された対象の精度が高いとは限らない。そこで、一般的には、領域分割の精度向上のために、切り出したい対象物の大まかな領域を手作業で指定するが、この操作を実施する場合には自動化は困難である。

以上のような問題点は存在するものの、本稿で対象とする問題は後述の通り、抽出したい物体の対象の色特性が類似した領域であること、領域分割数が任意に設定できること、データ作成工数や計算時間が一定以下に収まることが重要であるため、この手法を元にしたアプローチを検討することとした。

3. 分析のアプローチ

3.1 分析データの概要

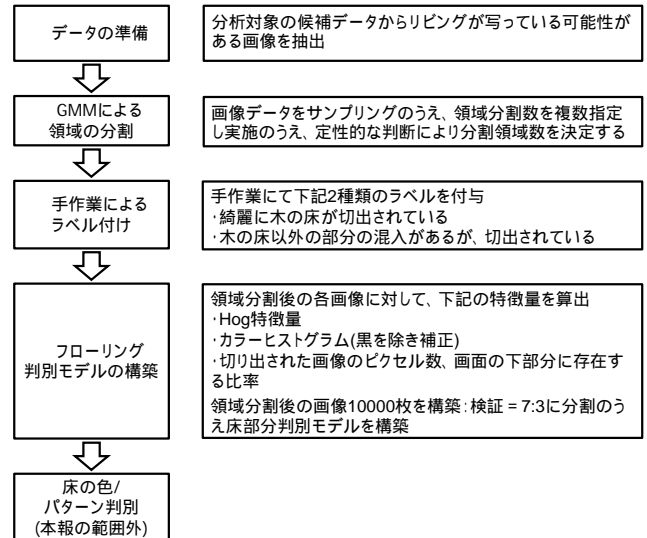
本研究では、リクルート住まいカンパニーが運営する不動産に関する情報サイトである「SUUMO」(<http://suumo.co.jp>)のデータを利用した。この概要を表 1 に示す。尚、本画像中には机、ソファ、イス、のような家具や照明器具などの物体が多いことが特徴であると考えられる。したがって、これらの混入を防ぎながら、背景部分を切り出すことが技術的な課題となる。背景部分の切り出しにあたり、実現した際の有用性が高いと考えられる木の床を抽出することを目標とした。

表 1. 分析データの内容

対象データ	注文住宅領域
分析対象画像データの件数	注文住宅領域の画像 31,870 枚より、リビングが映っていると考えられる 25,574 枚の画像を抽出。さらに、ランダムに 1,000 枚を抽出して使用
画像データの大きさ	(幅,高さ)=(350,350)

3.2 分析アプローチの全体像

本稿で提案するアプローチでは大きく 4 つのステップが存在する。この概要を図 2 に示す。分析にあたり、データ処理やモデル構築には Matlab2015a を使用した。



4. GMM による領域の分割

4.1 データの準備

本稿で用いるデータには画像の撮影対象に関する情報が整備されていなかった。そこで、画像の説明テキストから下記単語を含むを含む 25,574 枚の画像を選定した。

{リビング, キッチン, 階段, ダイニング, 天井, 和室, カウンター, 寝室, 棚, 家具, 畳, インテリア, 書斎, テーブル}

4.2 分割領域数の検討

GMM を用いて、パラメータとして分割領域数のみを与え領域分割を行なった。GMM とは画像内の色分布を Gauss 関数の線形結合で表現したモデルである。モデルのパラメータは EM アルゴリズムにより決定する。分割数は、5, 10, 15 の 3 パターンを比較し、今回の分析では、定性的な判断から 10 に設定した。この実行結果の比較例の一部を図 3 に示す。

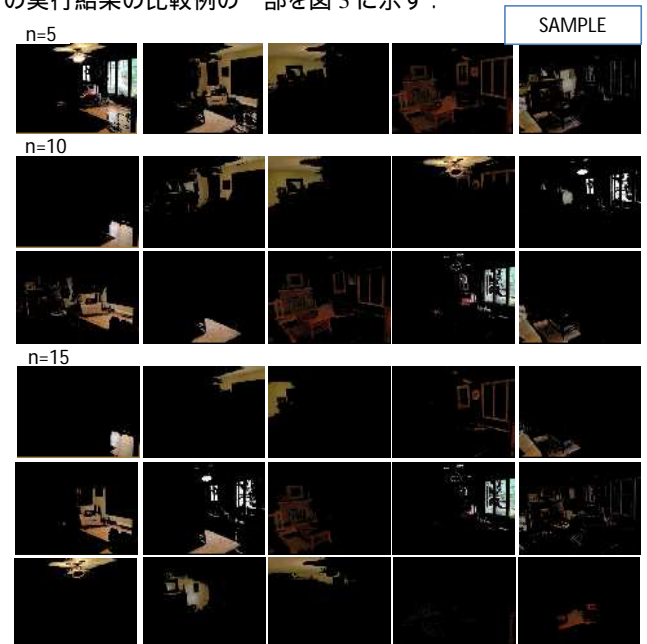


図 3. 領域分割数ごとの GMM による領域分割の適用結果例

この結果を元にランダムサンプリングした 1,000 枚の画像データを対象に、GMM による領域分割を適用した。この結果の例を図 4 に示す



図 4. GMM による領域分割の適用結果例

4.3 手作業によるラベル付け

領域分割後の 10,000 枚に対して、手作業にて下記 3 種類のラベルを付与した。この例を図 5 に示す。

- 1) 綺麗に木の床が切り出されている
- 2) 木の床以外の部分の混入があるが切り出されている
- 3) 領域分割の失敗



図 5. “綺麗に木の床が切り出されている”と、“木の床以外の部分の混入があるが切り出されている”の例

「2)木の床以外の部分の混入があるが切り出されている」とは木の床を含む領域へ分割が出来ているものの、ソファなどの物体も同一の領域とされている場合である。領域分割の精度向上のために必要な改善点を得ると共に、後述する木の床判別モデルの構築にあたり、全く木の床が含まれない画像の誤判別するよりはこちらを判別したほうが良いためこのような区分を設けた。

「3)領域分割の失敗」とは、床と判別できる画像が存在しなかったような場合である。

手作業に手分類を行なった件数の内訳を表 2 に示す。分割元の画像で 662 枚中 427 枚の 64.9%は綺麗に木の床が切り出されており、期待した精度が得られている。また、領域分割に失敗した画像は 662 枚中の 5 枚であり、99.2%の画像は GMM に

より何らかの形で木の床が含まれる領域の分割が実現できている。

表 2. フラグの件数内訳

	元の枚数	分割後枚数
(1)綺麗に木の床が切り出されている	427	567
(2)木の床以外の部分の混入があるが切り出されている	230	260
(3)領域分割の失敗	5	50
木の床が写っている画像 (1)+(2)+(3)	662	877
全体	1,000	10,000

5. 木の床判別モデルの作成

5.1 木の床判別モデルの作成

本稿では、木の床判別モデルの構築にあたり、説明変数として 5 種類の特徴量を作成した。それぞれの概要を表 3 に示す。

表 3. 各特徴量と概要

特徴量	概要
A:カラーヒストグラム	木の色には特徴的な色があることを想定し、画像ファイル中の RGB の組み合わせごとにピクセルの構成比を現したもの
B:黒領域	領域分割後の画像中、分割対象以外の画像は黒調整後カラー(RGB=0,0,0)へ置換しているため、黒色のピクセルを除外したうえで構成比を算出したもの
C:Hog特徴量	木の床が切り出しがうまく行っている際には形状的なパターンがあることを想定し、画像ファイル中の輝度勾配を算出し、物体の存在を数値的に表現したもの
D:パターンマッチングによるスコア	木の床と層でないエリアを対象に 10 ピクセル×10 ピクセルのパッチを作成し、各ピクセルの RGB 値からなる 300 の変数を元にスコア判定を行なう
E:分割後画像の存在位置	木の床は画像中の下半分に存在することが多いため、領域分割後の画像が、画像中の下 n%に存在するピクセルの比率を算出したもの

5.2 木の床判別モデルの精度

木の床判別モデルの構築にあたり、表 2 に示すデータを構築:検証 = 7:3 に分割した。

表 4. 構築, 検証データの内訳

区分	構築	検証
(1)綺麗に木の床が切り出されている	396	171
(2)木の床以外の部分の混入があるが切り出されている	194	66
全体	7,000	3,000

モデル構築にあたり、アルゴリズムは MATLAB 中の fitcsvm, fitensemble, TreeBagger 関数を使用し、アルゴリズムは SVM, Bagging, AdaBoost, RUSBoost を実施のうえ精度を比較した。目的変数は“(1)綺麗に木の床が切り出されている”とした。精度指標に関して、precision に関しては、Precision1:“(1)綺麗に木

の床が切り出されている”に加えて、Precision2:”(2)木の床以外の部分の混入があるが切り出されている”も加えた場合の2種類を算出した。Recall に関しては、“(1)綺麗に木の床が切り出されている”のみを算出した。表3に示す各利用特徴量ごとの精度比較結果を表4に示す。

表5. 各モデルごとの精度

使用特徴量	Precision1	Precision2	Recall	F 値
A	0.451	0.533	0.322	0.375
B	0.407	0.520	0.357	0.380
C	0.407	0.453	0.357	0.380
D	0.647	0.647	0.567	0.604
E	0.367	0.473	0.322	0.343
B,C	0.615	0.725	0.327	0.427
B,C,D	0.733	0.740	0.643	0.685
A,B,C,D,E	0.800	0.800	0.468	0.590

変数単独では D:パターンマッチングによるスコアによる精度が最良となった。B:黒領域調整後カラーヒストグラム, C:Hog 特徴量に関しては単独での精度はさほど良くないが、変数を組み合わせることで精度の向上が見られた。

また、B:黒領域調整後カラーヒストグラム, C:Hog 特徴量, D:パターンマッチングによるスコアの組み合わせで F 値が最大になった。一方、Precision1 と Precision2 に関しては D:パターンマッチングによるスコアを使用した場合に変化が見られないものの、それ以外の場合には値の上昇が見られる結果となった。

6. 考察

本稿で提案したアプローチにより、一定水準の精度により画像から、背景部分に相当し、様々な形状や大きさが存在する木の床部分が切り出せることが分かった。中でも黒領域調整後カラーヒストグラム, Hog 特徴量, パターンマッチングによるスコアの組み合わせで精度が最良となった。これは、木の色の情報、木の床が画像中に出現しやすい場所、木の木目に現れるパターンのくみあわせにより一定水準の精度が得られたと考えられる。本稿で得られた精度の水準は実際のサービスで要求される精度水準を満たしており、運用有用性は高い。また、[Jagadeesh 2014], [Jing 2015]らの他の事例と比較しても遜色のない水準となっている。

一方で、机などの混入のある画像が一定数存在するため、領域分割の精度向上と木の床判別モデルそれぞれについて、精度向上が必要であると考えられる。

領域分割においては、GMM をさらに発展させた GrabCut などの手法の利用や、物体検知をした上で、前景部分の物体を取り除いたうえで領域分割を行なうことが考えられる。

木の床判別モデルの精度向上に関しては、本稿で使用しなかった手法の適用も考えられる。

また、本稿で使用したデータは全 25,574 枚中から 1,000 枚をランダムサンプリングしたうえで分析を実施した。したがって、必ずしも木の床の全パターンが網羅されたとは考えづらい。したがって分析対象のデータ枚数を増やし、モデルの頑健性を担保する必要があると考えられる。

7. まとめと今後の課題

本稿では、GMM を利用した画像データ中の領域分割手法をもとに、その短所である分割領域の手作業による指定を行わずに画像から特定の背景部分を抽出するための手法を提案した。

今後の課題として、第一に本手法の性能向上が挙げられる。なかでも計算時間の短縮、精度のさらなる向上が挙げられる。第二に壁や天井などの他の背景部分についても同様の手法の有効性を検証することが挙げられる。

参考文献

- [Girshick 2014] Ross Girshick, Jeff Donahue, Student Member, Region-based Convolutional Networks for Accurate Object Detection and Segmentation, IEEE, Trevor Darrell, Member, IEEE, IEEE, 2014
- [Chan 2001] T. F. Chan, L. A. Vese, Active contours without edges. IEEE Transactions on Image Processing, Volume 10, Issue 2, pp. 266-277, 2001
- [Protiere 2007] A. Protiere, G. Sapiro, Interactive Image Segmentation via Adaptive Weighted Distances, IEEE Transactions on Image Processing, Volume 16, Issue 4, 2007.
- [Fu 2012] Zhaozia Fu and Liming Wang, Color Image Segmentation Using Gaussian Mixture Model and EM Algorithm, In Proceedings of Second International Conference CMSP 2012, pp.61-66, 2012
- [Jing 2015] Visual Search at Pinterest, In Proceedings of the International Conference on Knowledge Discovery and Data Mining (SIGKDD), 15, pages 1889-1898, 2015
- [Jagadeesh 2014] V. Jagadeesh, R. Piramuthu, A. Bhardwaj, W. Di, and N. Sundaresan. Large scale visual recommendations from street fashion images. In Proceedings of the International Conference on Knowledge Discovery and Data Mining (SIGKDD), 14, pages 1925-1934, 2014.