

調理作業スペースの概念モデルの可視化の検討

Visualization of conceptual model of cooking workspace

濱 龍太郎^{*1}

Ryutaro Hama

森 梓^{*1}

Azusa Mori

原田 篤^{*1}

Atsushi Harada

高久 由香里^{*1}

Yukari Takaku

橋本 敦史^{*2}

Atsushi Hashimoto

^{*1}株式会社 LIXIL

LIXIL Corporation

^{*2}京都大学教育学研究科

Kyoto University

This paper proposes a method for visualizing and analyzing conceptual models of cooking space in kitchen. We hypothesized that users recognize the space as a set of separated areas to which the users allocate cooking items (tools and foods) along with the concepts in their mind. To model such concepts, we separated cooking areas according to the detected items' location, then characterized these areas by the located items' categories; we use Single shot multibox detector (SSD) for item detection, the EM algorithm with Gaussian Mixture Model (GMM) for dividing the space into areas based on the detected items, and characterize each area with term frequency - inverse document frequency (tf-idf), where we regard each area as a document and items as terms. In conclusion, we visualized their conceptual models of relationships of kitchen items and the areas, and their shape patterns.

1. 緒言

キッチン住生活において食を支える設備であり、何年にも渡り日々使われる設備であるため、そのユーザビリティは極めて重要である。プロダクトデザインのユーザビリティにおいて、ユーザーの持つ、行為の対象物に対する概念モデルと、その対象のシステムイメージとが一致しないときにユーザビリティが下がると言われている [Norman 13]。キッチン設備においても調理に関するユーザーの概念モデルを把握できれば、概念モデルに則した機能提案や評価ができる可能性がある。

キッチン設備は図1の様にシンク、コンロ、作業スペース、収納などの要素からなるが、多くの作業は作業スペースを中心に行われるため、作業スペースの概念モデルを把握することはその他の機能提案にも重要であると考え、本研究では作業スペースに注目した。

作業スペースに対する概念モデルは、作業スペースが基本的には平面からなるため、自由度が高いが、用いる器具や食材の特性、水気や火気、身体からの距離などの影響を受けていると考えられる。これらの影響からユーザーは作業スペースを空間的に分割して捉え、分割したエリアに対して物体を配置しているのではないかと仮説を立てた。

本研究では、食材や調理器具が設備中のどこに配置されていたかは調理の概念モデルに基づいた行動の結果が表れていると考えられることから、食材・調理器具（以下アイテムとする）の種類、作業スペースへの配置に基づいた分析を行うこととした。

本研究では分析データの取得において深層学習による物体検出技術に着目した。キッチンワークトップ全体を天井カメラで撮影し、深層学習の物体検出フレームワーク Single Shot MultiBox Detector (SSD) [Liu 16] でワークトップ上での物体検出を行うことにより、目的とするアイテムの種類ならびに配置座標を推定できる。

この物体検出結果を元に、作業スペースに対するユーザーの概念モデルの解析手法を提案し、その可視化結果を示した。

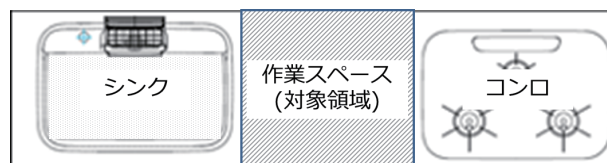


図 1: キッチン定義

2. 問題設定

2.1 数学的記述

本研究において、調理中に登場しうるアイテムを既知として、その集合を O とし、物体検出の出力は検出された物体 $o \in O$ 、時刻 t 、および矩形座標 (x_1, y_1, x_2, y_2) からなる。対象領域とする作業スペースの左手前を原点とした。アイテムの検出座標の代表値を矩形中心 (cog, center of graph) を $cog = (x_1 + x_2)/2 + (y_1 + y_2)/2$ 。また、ワークトップのエリア分割 S 、および各エリア $s \in S$ とし、検出アイテムのエリアに対する滞留時間を $\tau_{o,s}$ とする。

3. 提案手法

本論文では、ユーザーは作業領域を空間的広さや用いるアイテムの特性に基づいて、大きく分割して認識し、その領域に用途や形状などの特性に基づいてアイテムを配置していると仮定した。以下の手順で配置とアイテムの関連性や空間の分割のモデルを分析する手法を提案する。まず、調理中のアイテムを物体検出アルゴリズムを用いて検出し、配置されたアイテムを抽出する。次に、空間的な分割の概念モデルを実際のアイテム配置位置に基づいて分析する。さらに、分割されたエリアに配置されたアイテムを特徴としてエリアをクラスタリングし、意味づけを行う。

3.1 食材・調理器具の検出

調理における道具や食材（アイテム）のデータの取得方法として、調理作業を上方より撮影した動画像に対し、フレームごとに物体検出モデルを適用した。この検出には SSD [Liu 16]

連絡先: 濱 龍太郎, 株式会社 LIXIL

Technology Research 本部 人間情報科学研究所

〒136-8535 東京都江東区大島 2-1-1 LIXIL WING ビル

E-mail: ryutaro.hama@lixil.com

を用いた。SSD の畳み込み層には ResNet50 を用いた。対象物が画像領域に対して微小なため、検出精度の高い SSD-512 を用いた。

3.2 移動・滞留判定

アイテムの配置位置を評価するにあたり、検出されたアイテムの移動・滞留を判定し、滞留時のデータを用いることとした。

3.1 で得られた検出アイテムに対し、アイテムの cog のフレーム間移動距離を求め、移動平均によって平滑化を行なった。この移動距離の平均・偏差を求め、移動距離が平均+標準偏差の範囲内を滞留・それ以上を移動と判定することとした。平滑化は検出領域の変動による物体の中心位置の移動の影響をなくするためである。

3.3 滞留領域のエリア分割

ユーザーは自身の空間や物体の概念モデルに基づいて、空間を大まかにエリアに分割してアイテムを配置するものとする。アイテムはユーザーの概念モデルに合致するときに長期的に配置位置に滞留し、合致しないとき又は一時置き概念モデルのときには短期的な滞留となると考えられる。このとき配置は、滞留時間を重みとした確率密度が正規分布に近似できると仮定して、EM アルゴリズムを用いた GMM (Gaussian Mixture Model) へのフィッティングを行い、その分布を抽出することとした。この際、前後左右それぞれに 3 分割程度のエリアを想定し、10 を最大分割数として BIC 基準でエリア数を決定することとした。

3.4 エリアの内包アイテムによる意味付け

3.3 にて分割したエリアがユーザーの空間に対する分割意図を反映しているのか、その意図は何か、をエリアの座標情報だけでなく、内包するアイテムにより意味付けすることで評価したい。本研究ではアイテムによるエリアの意味づけにおいて、term frequency - inverse document frequency (tf-idf) を用いたベクトル空間モデルを適用し、余弦類似度でエリア同士の類似度を表現した [奥野 16]。

tf-idf の適用にあたり、エリアに対するアイテムの出現頻度を tf, あるアイテムの存在するエリアの頻度を idf とする。式 1 に計算過程を示す。エリアの特徴をアイテムの出現頻度あるエリア s に対するあるアイテム o の tf の計算において、被験者の違いによる調理時間差や用いたアイテム数の差による影響を少なくするために、アイテム i の滞留時間 $\tau_{o,s}$ を全アイテムの滞留時間 $\sum_k \tau_{k,s}$ で正規化した。ここで、滞留時間は短いもので数秒、長いもので数十分と差が大きいため、tf を対数で正規化した。この際に、tf が負の値を取らない様に 1 を加算した。

また、アイテム o の出現するエリアの数 df_o を、全エリア数を N で除した値を idf とした。

$$tf-idf(o, s) = \log\left(\frac{\tau_{o,s}}{\sum_{o'} \tau_{o',s}} + 1\right) \cdot \frac{\log N}{df_o} \quad (1)$$

この余弦類似度を元に Ward 法で階層的クラスタリングを行い、各クラスタにおけるアイテムの特徴量を元にエリアの意味づけを行う。

4. 実験

4.1 実験条件

4.1.1 実験環境

モニター調査は弊社施設内の実験環境にて行った。実験環境全体図を 2 に示す。LIXIL 製システムキッチン RICHEL

SI(幅:2585 mm, 奥行:650 mm, 左勝手; 左手にシンク)を使用した。また、それぞれの横幅はシンク幅 770 mm, 作業スペース 750 mm, コンロ (LIXIL 製ひろまるコンロ) 815 mm である。カウンタータイプのキッチンであるが、壁付けタイプを想定して奥行き 650 mm 地点に境界を設けてその手前にて作業を行うこととした。撮影は天井設置カメラ (Panasonic 製 AG-HMR10A) を使用し、29.97 fps で撮影後、1920 × 1080 pixel で保存した。



図 2: 対象としたキッチン

4.1.2 調理モニター

実験参加者は一般、社員含めて 20 ~ 60 代の男女 24 名を対象とした。調理メニューは調理の基本動作のうち「切る、和える、味付け、焼く、茹でる、盛り付け」を含むように設計し、ご飯 (2 合・電気炊飯器使用)、照り焼きチキン (鶏もも肉 2 枚)、添え野菜 (カットトマト、茹でブロッコリー)、味噌汁 (大根、しめじ、長ネギ)、きゅうりの酢の物 (きゅうり、シラス) とした。食器および食材は 2 名分とした。実験条件の統一のため、調理メニューは事前に告知し、用意した道具の位置や機器の使い方以外の調理作業の動作教示は行わず、それぞれの手順にて調理を行った。

4.2 深層学習による物体検出

4.2.1 データセット

先述の被験者のうち、16 名の動画を 0.5 fps でフレーム化し、33,045 枚の画像を切り出し、出現した 89 種類のアイテム全てにアノテーションをおこなったものをデータセットとした。

4.2.2 SSD の学習・検出

SSD512 に ResNet50 の事前学習済みのモデルを適用し学習を行った。学習用・検出用の画像は全て 512 × 512 pixel に縮小して処理を行った。PASCAL VOC 基準で評価を行うと mAP (mean Average Precision) が 0.62 と算出された。アイテムの検出精度において、出現回数が少ないもの (フォークなど)、形状が大きく変化するもの (食材や料理など) に関しては検出精度が下がってしまっていた。その検出結果の例を図 3 に示す。

本研究では分析対象を配置に関わる調理器具および加工前の食材の 61 種類とした。これにより形状変化の大きい加工された料理を除外した。また、出現回数が少ないものは影響が限定的であると考え、そのまま扱うこととした。データの拡充と変化する物体の検出は今後の課題である。

4.3 エリア分割

手法 3.3 に基づき分割された領域数を BIC 基準で選定するとそれぞれ 8~10 であった。EM アルゴリズムによる GMM

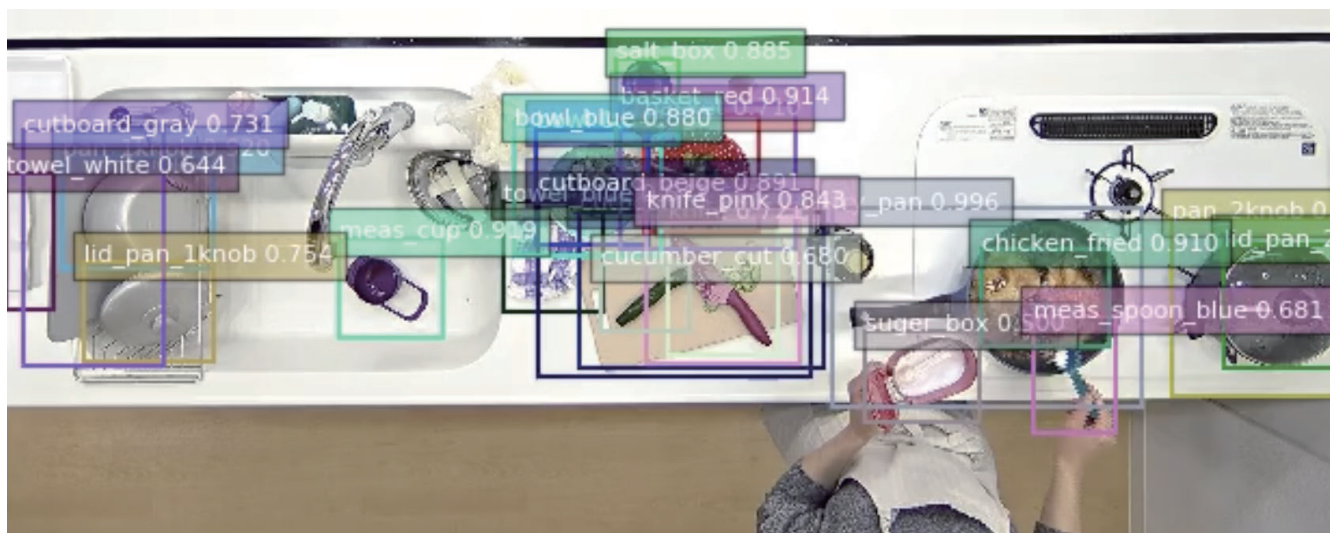


図 3: 深層学習 (SSD) による調理器具および食材の物体検出例

へのフィッティングは scikit-learn v0.19.2 におけるモデルを用いた [Pedregosa 11]. 図 4 にその結果の一部を示す. 灰色の楕円部は GMM で分割された各エリアの平均と分散からなる領域である. 色および記号は 4.4 にて記述する. 図 4(a),(b) では空間の前後, 左右に分割されたエリアが形成されたが, 図 4(c),(d) では広域のエリアと複数の細かいエリアが形成された.

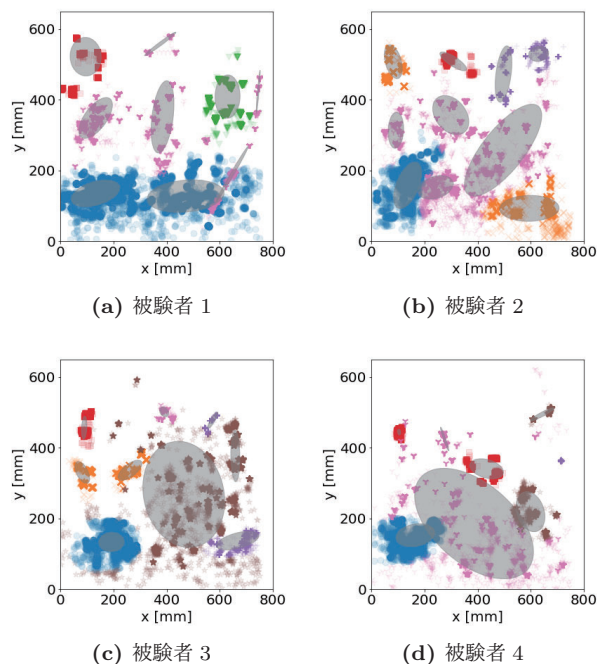


図 4: アイテムの配置に基づくエリア分け

また, 図 5 に示すように全被験者の BIC を比較すると, オーダーが同一であったため, およそ同程度の基準で分割できていた.

以上より, 被験者の作業スペースのエリア分割の概念モデルを明確なエリア分割があるタイプと, エリア分割が曖昧なタイプとで表現することができたと考えられる.

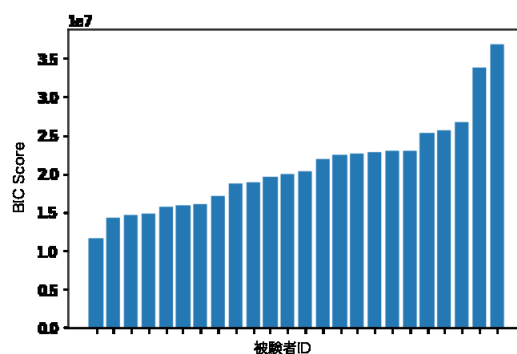


図 5: 被験者間の BIC 比較

4.4 エリアの意味づけ

前述 4.3 により得られたエリア $s \in S$ に対し, 3.4 により類似したエリアを集約した. クラスタ内の tf-idf の平均値と内容を元に 7 クラスタに決定した. 分類のデンドログラムを図 6 に示す. このクラスタごとの tf-idf の値が高いアイテムを表 1 に示す. 図 4 の色および記号が, アイテムの配置と種類を反映した例である. この凡例を表 1 中に示した.

同様にエリアに対し特徴的なアイテムを表 1 の 1~6 に示す. また, お玉や計量カップなどの道具類, 醤油やみりんなどのボトル調味料は作業スペース上のエリアに固定される傾向が薄いアイテムであることがわかった.

この 24 名の作業スペースのエリア分割と, その主要アイテムから以下の結果が得られた.

もっとも特徴として強く出たまな板・包丁からなるエリアは, 図 4(a) のように左右に分布している群 9 名と, 図 4(b)~(d) に示すようにシンク寄りの手前側に配置される群 15 名とに分けることができた.

5. 考察

クラスター配置と種類, 実際の動画観察を元に概念モデルの形状を図 7 のように表現した. 図 7(b)~(d) のうち, 図 7(a),(b) は奥行き方向で分割し, 長時間置く物を奥に置いていた. 両被験者の違いとしてカット位置を固定しているかがあったが, 被

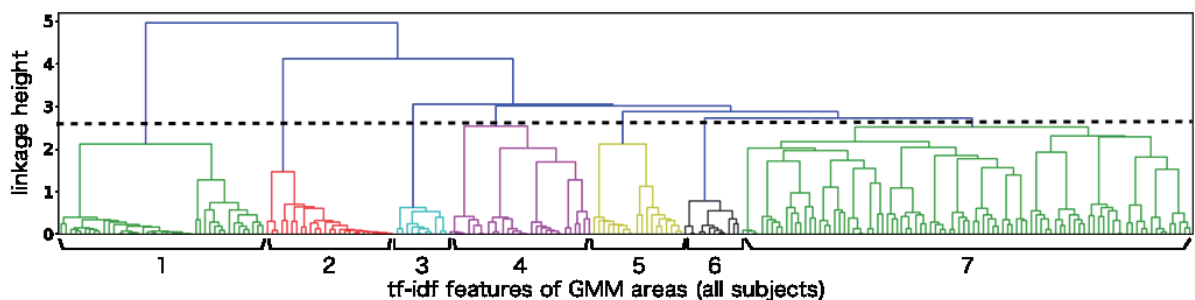


図 6: アイテムを特徴量としたエリアのクラスタリング

表 1: エリアの意味付け.

clusterNo	主要アイテム	凡例
1	まな板・包丁	●
2	台拭き	×
3	塩・砂糖	▼
4	ボール・ザル	■
5	鍋蓋類	+
6	大皿	★
7	その他 (食材・器具)	Y

験者 2 においてはカット以外の作業を立ち位置を変えて行っていた. (c) は左右で分割され, 右側全体を盛り付けを含むフリーエリアとして利用していた. また, (d) ではカット以外の作業も作業位置を固定し, 極座標系の手の届く距離で配置を行なう傾向が見られた.

このように, 作業スペースの空間認識において, 奥行き方向, 左右方向だけでなく, 極座標系の概念モデルを主要な形状として抽出することができた.

このような方法でタイプを分類すると, 全被験者においては奥行き方向の意識が強い前後分割タイプの被験者が 11 名, 左右分割意識が強い左右分割タイプの被験者が 4 名, 極座標系タイプの被験者が 9 名であった.

また, 多くの被験者において単独のモデルでなく左右や奥行き, 極座標などのモデルを組み合わせることで作業スペースを認識していたと考えられる.

6. まとめ

本研究では物体検出データを用いた調理行動における作業スペースのエリア分割とその意味づけによる概念モデルの可視化の手法を提案した. 実際に被験者の調理モニターを行い, 作業や設備と配置アイテムの関連性を評価することで, ユーザーのキッチンスペースに対する概念モデルとして, 物体配置において特徴的なアイテムと配置の形状パターンを得ることができた. この形状パターンはユーザーによって異なっていたため, 個人の性格や調理に対する意識, 経験が影響を与えていると考えられる.

今後の課題として, 形状パターンの分類が定性的な評価を含むため, 定量的な評価指標の検討を行いたい. また作業スペースに対するコンロやシンクの配置変化や調理のメニューや器具変化により, 配置の傾向がどのように変化するか, その要因は何かを分析すること, 調理中の時系列変化を捉えることが挙げられる.

さらに, この調理行動のパターンを分析することでユーザービリティの高いキッチン設備の提案につなげていきたい.

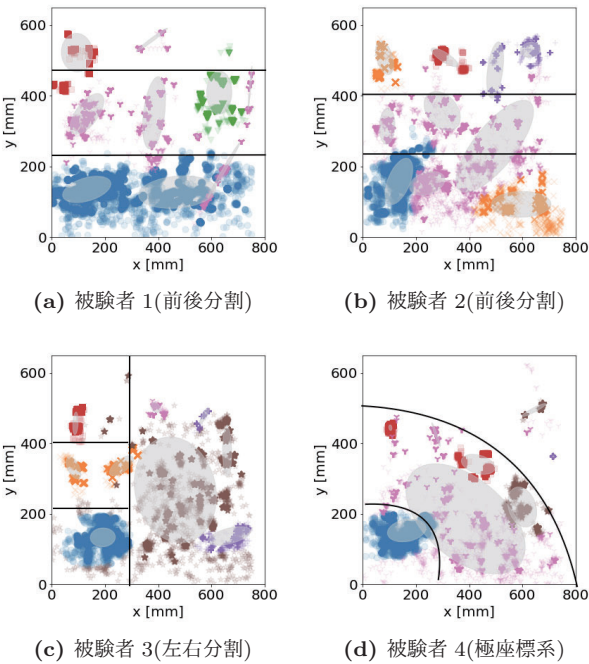


図 7: 作業スペースの意味区分

参考文献

[Norman 13] Norman,D.A.: The design of everyday things ?Revised and Expanded edition, New York:Basic Books(2013)(邦訳 :岡本 明ら, 新曜社 (2015))

[Liu 16] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C., and Berg, A.: SSD: Single shot multi-box detector, *European conference on computer vision*, pp. 21?37, Springer (2016)

[Pedregosa 11] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Pret-tenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. and Duchesnay, E.: Scikit-learn: Machine Learning in Python, *Journal of Machine Learning Research*, Vol. 12, pp. 2825?2830 (2011)

[奥野 16] 自然言語処理の基本と技術 (仕組みが見えるゼロからわかる)・奥野 陽ら, 翔泳社 (2016)