

調理行動理解に向けた視線情報分析方法の検討

井上 裕哉^{1,a)} 平山 高嗣¹ 道満 恵介¹ 川西 康友¹ 井手 一郎¹ 出口 大輔¹ 村瀬 洋¹

1. はじめに

近年、料理教室が人気を博しており、調理技術の向上を目指す人が増えている。情報システムによる調理者の支援を考える際には、調理者の調理行動を理解する必要がある。従来研究では、台所上の固定カメラで撮影された調理シーンの映像特徴に基づいて調理動作が認識されている [1]。本研究では、行動が認知・判断・動作の過程から生成されると考え、調理状況の視覚的な注意や認知を重要視し、調理者の視行動を分析対象とする。視線運動は人間の内部状態を反映し [2]、その視野の周辺には行動に関連する情報が含まれる [3]。視線情報の分析に基づいて調理者の行動を理解できれば、調理者の支援のほか、たとえば料理レシピへの掲載や調理映像の要約のために見どころを抽出できる。また、視行動の個人差を抽出することで、調理動作のコツだけではなく、判断のコツを形式化することができる可能性もある。

そこで、本研究ではそのような調理行動理解の前段階として、視線運動と調理動作の関係性を検証するために、視線情報に基づいて調理の基本動作を識別することを目標とする。関連研究としては Bulling らが視線運動を利用して数種類のデスクワークの識別を行っている [4]。また大垣らは注意を向ける方向の小さな変化に関する視線運動と、その大きな変化に関する頭部運動を反映した 1 人称視点映像の Optical flow とを組み合わせて更に高精度にデスクワークを識別できることを示している [5]。しかし、本研究で対象とする調理行動とデスクワークでは動作の性質が異なる。そこで本研究では、これらの従来手法を調理動作識別に適した特徴量を導入して改良することで、視線情報の分析に基づいて調理の基本動作の識別を行う。

2. 視線情報の分析に基づく調理動作識別手法

図 1 に本手法の処理手順を示す。事前処理で視線情報を記号化し、そこから得られる N -gram のヒストグラムを特徴量とし、学習・識別に用いる。図 2 に事前処理の手順を示す。計測した視線座標系列はノイズを含むため、まず x 座標系列、 y 座標系列それぞれに対してメディアンフィル

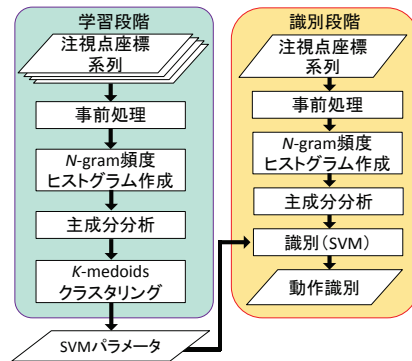


図 1 調理動作識別処理の手順

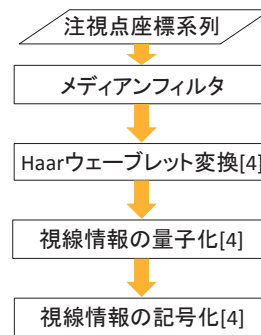


図 2 事前処理手順

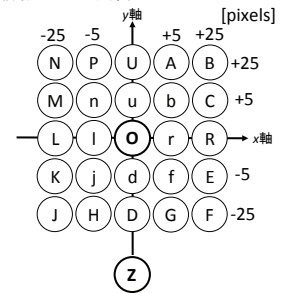


図 3 注視点の移動に応じて付与する記号

タを適用する。次に文献 [4] と同様に Haar ウェーブレットによる連続ウェーブレット変換を行う。更に大小 2 種類 (負の閾値を含めれば 4 種類) の閾値を設けて 5 段階に量子化を行ったうえで、量子化した x 座標系列、 y 座標系列を統合して記号化する。現フレームに付与された記号は前フレームからの相対的な注視点移動を表す。注視点の移動に応じて付与する記号を図 3 に示す。文献 [4] における記号に加えて、調理行動には調理対象への注視を継続させる振る舞いが頻繁に現れるため、視線移動のない原点 (記号「O」) を設ける。さらに注視点が計測されなかった区間は瞬きを行っていると仮定し、瞬き記号「Z」を設け、合計 26 通りの記号のいずれかに変換する。瞬き区間においては、その前後の座標で線形補間した座標に事前処理を適用する。以上の事前処理によって作成した記号列を視線情報として利用する。本研究ではこの記号列をある一定の時間幅の解析窓で区切り、各々に 1 つの調理動作が含まれるものとする。解析窓はその幅を 900 フレーム (15 秒) とし、

¹ 名古屋大学 〒464-8601 愛知県名古屋市千種区不老町

^{a)} inoueh@murase.m.is.nagoya-u.ac.jp

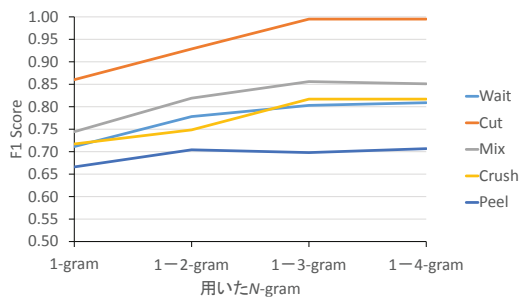


図 4 考慮する上限 N の値と精度の関係

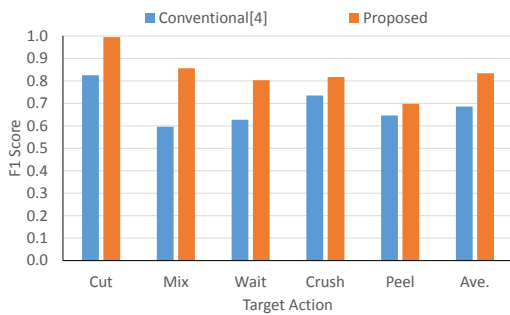


図 5 調理動作識別結果における従来手法との比較

60 フレーム間隔で移動させる。ただし解析対象の時区間に続く次の 60 フレームの間で動作が変わる場合は、その動作が変わる時点までその解析窓を拡げる。そして解析窓ごとに N -gram の頻度ヒストグラムを作成する。頻度ヒストグラムは 1-gram から n -gram までのヒストグラムを結合する。ここで、この頻度ヒストグラムをそのまま特徴ベクトルとすると、次元が 18,278 次元と高次元になってしまうため、主成分分析により次元数を削減する。最後に学習データ全ての解析窓で作成した特徴ベクトルを用いて SVM 識別器を学習する。

3. 実験

本手法で調理の基本動作が識別可能かを検証するために、調理映像 7 本に対して調理動作の識別を行った。これらの調理映像は、4 本がハンバーグの、3 本がポテトサラダの調理過程を撮影したもので、調理動作として「Cut」, 「Mix」, 「Wait」, 「Crush」, 「Peel」の 5 種類を含む。前述の解析窓の幅で視線記号列を分割したところ、サンプル数は 7,880 となった。それぞれに上記 5 つの動作のいずれかが割り当てられる。このサンプルに対して映像単位での交差検定を行った。視線計測には NAC 製 EMR-9 を用いた。評価実験では動作ごとに学習した識別器を用いて、テストデータが対象動作かそれ以外の動作のいずれかという 2 クラス識別を行い、F 値で評価を行った。この時、対象動作を行っているサンプル以外はネガティブサンプルとして扱うため、2 クラスのサンプル数に偏りが生じる。そのため本研究ではポジティブサンプルとネガティブサンプルの比に応じて、学習するネガティブサンプルを k -medoids クラスタリング [5] により間引く。

用いた N -gram と F 値の関係を図 4 に示す。3-gram まででは用いた N -gram の N の値が大きいほど精度が上がった。そして 1-3-gram を用いた場合の精度は平均 83.4% であった。この場合に、従来手法と比較した、調理動作と精度の関係を図 5 に示す。図より全ての動作において比較手法 [4] より高精度の識別を行えたことが分かる。なお最も精度が低かった Peel については、調理者が皮むきを手元に引き付けて行ったため、注視対象が視線計測装置のキャリブレーション平面から離れてしまい、視線を正確に計測できなかったことが原因と考えられる。また Wait の精度が低いのは、被験者によって待機中の動作に個人差があったためであると考えられる。

4. むすび

本報告では視線情報を用いて 5 種類の基本的な調理動作の識別を行った。実験では動作が混ざらない時区間を対象に 2 クラス識別で各動作の識別を行ったため高い精度が得られた。今後は動作区間の切り替わりを推定し、同一動作の区間を自動検出する改良を行う。動作が切り替わる直前に、調理者の意識が次の動作に向かう。その際に、調理者の注意が次の動作に関連する調理器具や食材に移るため、特徴的な視線パターンが表れる可能性がある。また、クラス数を増加させる必要があり、多クラスで識別を行うための改良を検討する。このようにして調理動作が識別可能となれば、これを調理支援に応用するために、視線運動パターンをさらに詳細に解析する必要がある。例えば、調理上級者と初級者では現れる N -gram の出現順序と持続時間が異なる可能性がある。このような特徴に基づいてコツを抽出するなどの応用を目指していく。

5. 謝辞

本研究の一部は、科学研究費補助金による。

参考文献

- [1] Y. Hayashi et al., "Automatic authoring of domestic cooking video based on the description of cooking instructions," Proc. 5th Int. Workshop on Multimedia for Cooking and Eating Activities, pp.21–26, Oct. 2013.
- [2] 平山高嗣, "人間の内部状態を顕在化する視覚的インタラクション," 情報研報, 2013-CVIM-188-27, Aug. 2013.
- [3] Y. Li et al., "Learning to predict gaze in egocentric video," Proc. 2013 IEEE Int. Conf. on Computer Vision (ICCV2013), pp.3216–3223, Dec. 2013.
- [4] A. Bulling et al., "Eye movement analysis for activity recognition using electrooculography," Proc. IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 33, No. 4, pp.741–751, Feb. 2011.
- [5] K. Ogaki et al., "Coupling eye-motion and ego-motion features for first-person activity recognition," Proc. IEEE Workshop on Egocentric Vision in conjunction with CVPR2012, pp.1–7, June 2012.
- [6] D. Vinod et al., "Integer programming and the theory of grouping," journal of american statistical association, Vol. 64, No. 326, pp. 506–519, June. 1969.