

# SynCook:動画メタデータと加速度センサを用いたレシピ動画進行度自動同期システム

加藤 岳大\* 横窪 安奈\* ロペズ ギヨーム\*

**概要.** 加速度センサと動画メタデータを用いたレシピ動画進行度自動同期システム SynCook (シンクック) を提案する. 通常レシピ動画を見ながら同時に料理をする場合, 動画を一時停止する動作や巻戻す動作が発生し料理時間が長くなる. 加えて料理中に機器を触ることは衛生上のリスクもある. SynCook は包丁の柄先に装着可能な小型の加速度センサを用いてユーザの調理行動を認識し, 動画に付与されているメタデータを利用して動画を制御することでレシピ動画とユーザが実際に行っている料理の進行度を自動的に同期させることができる. SynCook を構成している機器は Android 搭載のスマートフォンおよび加速度センサのみである. さらにレシピ動画は YouTube に投稿されている料理動画を使用するため, 日常生活で手軽に SynCook を用いることが可能である.

## 1 はじめに

昨今, インターネットの普及に伴い, レシピサイトの需要が高まり, レシピサイトの利用者が増加している [1]. レシピ動画は文章のみのレシピに比べ, 実際に料理しているイメージを掴みやすいことや, 文章だけではわかりにくい食材の様々な切り方を実際に見て確認できるため料理経験が少ない人でも利用しやすいといった利点がある. しかしレシピ動画は編集された動画であるため, 動画を見ながらの同時作業は困難である. そのため通常ユーザはレシピ動画を一時停止する動作や動画を巻戻す動作が発生するために料理をする際の負荷が大きくなることや料理時間が長くなるという問題がある. さらに料理中にスマートフォンのような機器を触ることは衛生上のリスクがあり, あまり好ましくない. そこで本研究では上記のような問題を解決するためにレシピ動画とユーザが行う実際の料理工程を自動的に同期させるシステム SynCook (シンクック) を提案する.

## 2 関連研究

昨今, キッチンに様々な装置を取り付ける事でユーザの料理を支援するシステムが数多く提案されている. 森岡ら [2] はカメラとプロジェクタと対話ロボットが連携して料理を支援するシステムを開発した. この研究では2台のカメラと3台のプロジェクタを使用し, 食材の置き方や切り方を食材やまな板に重畳表示することでユーザの料理を支援している. さらに対話ロボットも導入しており, ユーザの質問に対してシステムが詳しく答えるという機能も実装

している. Sato ら [3] もプロジェクタと深度カメラを用いて, 遠隔地にいる料理のアドバイザーがボイスチャットおよび手書きメッセージの注釈を用いてユーザの料理を支援するシステムを提案している. また Hasada ら [4] はプロジェクタやカメラを用いず AR ゴーグルを使用することでユーザの調理を支援する研究を行った. 使い方が不明な調理器具の使い方を 3D アニメーションを用いて教示することでビデオを使用した場合よりもタスク遂行までの時間を大幅に短縮できた. しかしいずれの研究もプロジェクタ, カメラ, AR ゴーグルという高価な機材を使用しているため日常生活に導入コストが大きい.

料理は様々な工程からなるため, 料理を支援するにはユーザの調理行動を認識し, 料理レシピを適切に提示する必要がある. そのためユーザの調理行動を認識する研究が数多く行われている. Korematsu ら [5] は連続した調理行動を音を用いて推定する研究を行った. 「カット」, 「グリル」, 「その他」をフレーム毎の推定で 70% の精度で推定できた. また画像認識の技術を用いて調理行動を認識する研究 [6][7][8] も行われている. しかし実際の料理環境では様々な音がノイズになる事に加え, 動画を見ながらの料理では画像認識の精度が落ちてしまう恐れがある.

以上から, 音や画像を使った調理行動認識はユーザの行動を制限しないという利点がある反面, 環境のノイズによる障壁が大きく, プロジェクタなどの機器を導入する事は導入への敷居が高いという問題があった. そこで我々は, 日常生活に比較的簡易に導入可能かつ環境によるノイズが少ない加速度のみを用いて初心者の包丁技術の向上を支援する研究 [9] を行ってきた. 包丁の柄先に装着した加速度センサのみで包丁の動きを計測し, 料理熟練者の包丁技術を判定する指標を明らかにした. この結果から加速度センサのみでも料理の動きの計測が可能であるこ

Copyright is held by the author(s). This paper is non-refereed and non-archival. Hence it may later appear in any journals, conferences, symposia, etc.

\* 青山学院大学

とが明らかになったため、本研究でも加速度センサを用いて料理行動の認識を行う。

### 3 SynCook

#### 3.1 レシピ紹介動画の問題点

1章でも述べたように、昨今のレシピサイトでは動画を使ったレシピ紹介が主流となっているだけでなく、YouTubeのような動画配信サービスに料理動画を投稿する人も多くなり、今後ますます需要が増えていくことが予想される。動画を使うことで文章だけでは伝わらない細かな作業や実際の料理風景を視聴できるため、完成までのイメージがしやすいという利点がある。さらにYouTubeに投稿される動画は投稿者自らがテロップの挿入や効果音の付与といった編集を施し、視聴者にとってわかりやすく見やすいコンテンツとなっている。しかしこれらの動画は視聴用に作られているため、ユーザの視聴に不必要な食材の移動や長い時間を要する料理工程はカットされ短く編集されている。通常料理は完成までに複数の作業をこなす必要がある上に分量や切り方の細かい指示もあるため、一度動画を見ただけでそのレシピを再現する事は非常に困難である。そのため、ユーザは動画を見ながら同時に料理をすることを強いられるが、料理を中断し、一時停止や巻戻すという動作を行う必要があるため料理中の負荷が大きくなり、ユーザはレシピ動画で紹介されている料理を手軽に作るができない。

#### 3.2 SynCook に求められる機能

3.1節で述べたレシピ動画の問題点を解決するために、SynCookに求められる機能は以下の2つである。

- レシピ動画の料理工程を認識し、ユーザがその料理工程を開始したこと、および終了したことを判定し、ユーザの現在の料理工程を認識する。
- ユーザが機器に触れることなく、レシピ動画をユーザが現在行っている料理工程と同期するように制御する

料理工程を認識するためにはその工程が開始されたと判別される料理行動及び終了したと判定される料理行動を認識する必要がある。ユーザの料理行動認識には加速度センサを用いる。またレシピ動画としてYouTubeに投稿されている動画を用いる。そのためレシピ動画と実際の料理工程との同期させるための動画制御にはYouTubeのAPIを使用する。

以上の機能を有するSynCookを用いれば加速度センサで認識した料理行動を元に動画を自動制御するため、巻き戻しや一時停止を行う必要がなく、ユーザはレシピ動画を見ながら手軽に料理をすることが可能となる。

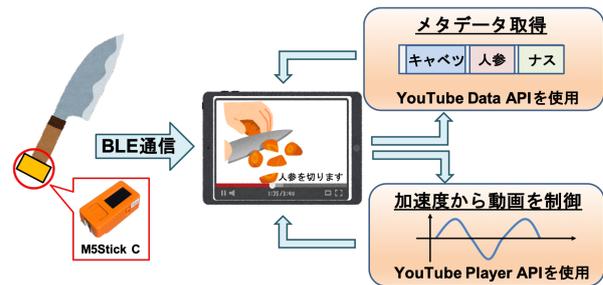


図 1. SynCook のシステム構成

### 4 SynCook の実装

#### 4.1 システム構成

図1にSynCookの構成を示す。SynCookはandroid搭載のスマートフォン（motorola社製のmoto e5）、包丁の柄先に加速度センサ（SWITCH-SCIENCE社製のM5StickC）を装着した包丁型デバイスで構成される。加速度センサとスマートフォンはBluetooth Low Energy(BLE)で通信されており、加速度センサからユーザの切断動作を加速度として取得しスマートフォン上で料理行動・料理工程の認識を行う。認識された情報を元にYouTubeのAPIを使用して動画の自動的な制御を行う。他の研究と比べ安価で導入しやすい小型の加速度を包丁に取り付けるのみで、包丁が単なる食材を切断するツールとしてだけではなく、スマートフォン上の動画とのインタラクションを可能にするデバイスとすることができる。

#### 4.2 料理行動認識

##### 4.2.1 加速度センサの装着位置

3.2節で述べたようにSynCookではユーザが行っている料理工程を認識するためにユーザの料理行動を認識する必要がある。料理行動を認識する研究は数多く行われているが、本研究では加速度センサを用いて認識を行う。図2に加速度センサを装着した包丁型デバイスを示す。加速度センサはユーザが食材を切断する際の動作の妨げにならないように包丁の後ろ部分に装着した。実際に包丁型デバイスを複数人に使用してもったが、全員が装着感は全く気にならないという意見であった。加速度センサから得られる加速度の方向は図2に示すようにx軸が上下方向、y軸が左右方向、z方向が前後方向である。

##### 4.2.2 切断工程

料理の切断工程は包丁を動かし食材を切っているだけではない。様々な動作が切断工程には伴い、それぞれの動作における包丁型デバイスの状態を適切に認識することでユーザの料理行動を把握し、レシピ動画と同期することができる。食材の違いより想

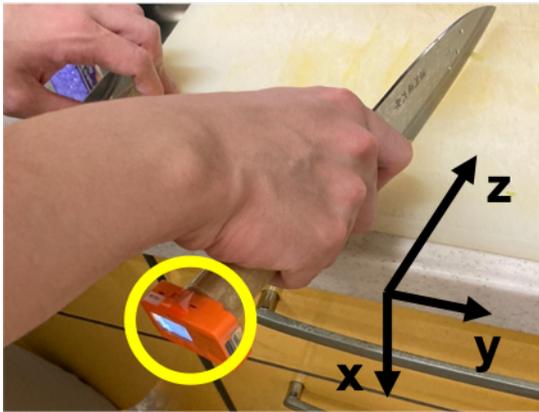


図 2. 加速度センサの装着位置

きゃべつなど下処理の不要な食材の場合

まな板に載せる	切断する	まな板から退ける
---------	------	----------

にんじんなど皮むきが必要な食材の場合

まな板に載せる	皮むき	切断する	まな板から退ける
---------	-----	------	----------

ピーマンなど切断動作が伴う下処理が必要な食材の場合

まな板に載せる	切断する	たねを除去する	切断する	まな板から退ける
---------	------	---------	------	----------

図 3. 食材の違いにより想定される切断工程

定される切断工程を図3に示す。きゃべつなどの食材は切断工程に切断する以外の下処理を必要としない。この場合は「まな板に載せる」、「切断する」、「まな板から退ける」の3工程から切断工程が構成されると考える事ができる。一方、にんじんや大根といった切断する前に皮を剥く必要がある食材も多く存在する。その場合は先程の3工程に「皮むきをする」工程を加えた4工程と考える事ができる。さらにピーマンなど内部に種を有する食材では下処理の前に1度切断動作が伴う。そのため下処理のための切断が加わった5工程から成ると考える事ができる。

#### 4.2.3 料理行動認識手法

4.2.2項で想定した切断工程を認識する手法を述べる。SynCookではレシピ動画と実際の料理工程が同期するため、切断工程の開始と終了を正しく判定する手法が必要である。しかし包丁型デバイスで取得できるデータは包丁の動きのみであるため、「まな板に載せる」工程や「皮むき」、「種を除去する」といった包丁を伴わない工程を動きとして計測することができない。そのため、本システムでは包丁を伴う動作である「切断する」工程および「まな板から退ける」工程を開始と終了のトリガーとして使用することで料理工程を認識する。

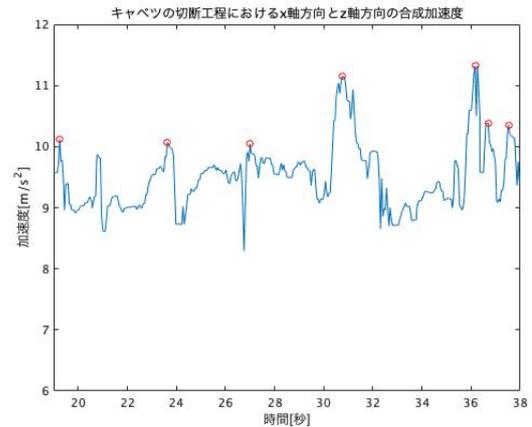


図 4. キャベツの切断工程における x 軸方向と z 軸方向の合成加速度

#### 4.2.4 データ収集

本手法を実装するために提案システムと同様の包丁型デバイスおよびスマートフォンを使用して切断工程のデータを収集した。食材には想定した切断工程を網羅できるようにきゃべつ、白菜、にんじん、大根、ピーマン、こんにやくを用意し、すべての切断工程を一連のデータとして収集した。なお加速度センサのサンプリング周波数はおおよそ20Hzであり、食材の皮むきにはピーラーを使用した。

#### 4.2.5 切断判定

キャベツの切断工程における x 軸方向の加速度と z 軸方向の合計加速度を図4に示す。包丁動作はユーザからみて下方向に押し切り、上方向に引いて戻す動作であるため、x-z平面で考えた時に特徴的な動作を行うことになる。そのため、x軸とz軸の合成加速度には切断を行った時に特徴的なピークが現れる。図4中の赤丸はこのピークを閾値を用いて検出し、プロットしたものであり、この時点で切断動作が行われたと考える事ができる。よって x 軸方向と z 軸方向の合成加速度を切断を判定する特徴量として用いる。

#### 4.2.6 料理工程の終了判定

料理工程の終了は、切断が行われてかつ食材を「まな板から退ける」動作をユーザが行い、終了した場合にその料理工程が終了したと判定する。「まな板から退ける」動作の判定を行うため、ユーザの”包丁のせ”を包丁型デバイスで検出する。包丁のせとは図5に示すように包丁の側面に食材を載せて食材を移動させる方法である。多くの食材を切断した際に一度に多量の食材を移動できることや包丁を持っている手が食材に触れないため方手を汚さずに済むという利点があり、料理の切断工程でよく見られる



図 5. 包丁のせ

動作である。包丁のせを用いて食材を「まな板から退ける」動作を行っているときの 3 軸加速度を図 6 に示す。この図からわかるように、包丁のせを用いて食材を「まな板から退ける」動作を行っている時は x 軸加速度と y 軸加速度が相反する増減を繰り返すという特徴が現れる。通常の切断動作では包丁を立てて扱うが、包丁のせは包丁を寝かせて包丁を扱う。これは切断動作の中でも特異的な動作であるためにこのような特徴が現れたと考察できる。よってこの特徴を食材を「まな板から退ける」動作を検出する特徴量として用いる。さらに「まな板から退ける」動作の終了判定について述べる。収集したデータの中には「切断する」、「まな板から退ける」、「切断する」、「まな板から退ける」といったように細かく「まな板から退ける」という動作を繰り返す場合もあった。しかし食材が完全にまな板から退けられ終わった場合、ユーザは包丁を持って静止あるいは包丁を置くという行動を取っていた。そのため、包丁が静止している状態である加速度に増減が現れないという特徴が現れた。よって「まな板から退ける」動作が判定され、直後に包丁が動いていない動作が判定された時「まな板から退ける」動作が終了しその料理工程が終了したと判定する。

### 4.3 動画の自動制御

加速度センサで認識したユーザの行動を元に YouTube 動画を制御する方法を述べる。

#### 4.3.1 動画メタデータの取得

動画を制御するためのメタデータとして YouTube の動画投稿者がチャプター機能を使うために記述するタイムスタンプを使用する。YouTube では図 7 のように動画投稿者が動画説明欄にタイムスタンプとタイトルを記述することでチャプター機能を使う事がができる。チャプター機能を用いることで視聴者が動画内の見たいチャプターに素早くアクセスすることができる。動画説明欄に記載されている情報

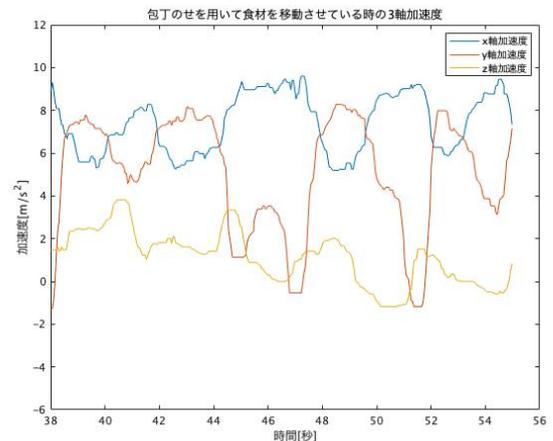


図 6. 包丁のせを用いて食材を移動させてる時の 3 軸加速度

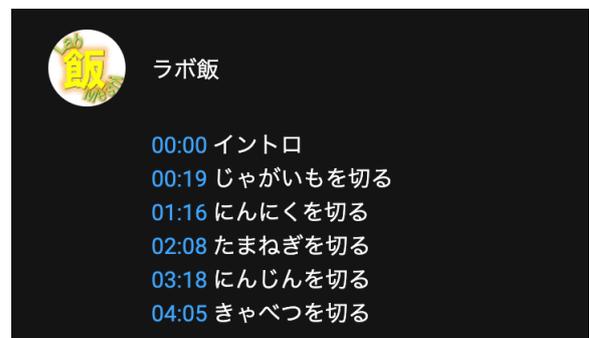


図 7. チャプター機能を使用するために記述するタイムスタンプ [10] より引用

は YouTube Data API と YouTube の全動画に割り当てられている ID を用いることで誰でも取得することが可能である。SynCook ではこの情報を動画を制御するためのメタデータとして取得し、使用する。

#### 4.3.2 動画の制御

加速度センサで得られた情報と取得したメタデータを用いて動画の制御を行う。制御には YouTube Player API を使用する。4.3.1 項で述べたように動画のメタデータから現在動画で行っている料理工程を SynCook が認識する。本来であれば停止しない限り動画は進んでしまうが、SynCook ではユーザの料理行動が終了しない限り動画はメタデータによりセグメントされた該当の料理工程をループする仕様になっている。動画の料理工程に合わせた料理工程をユーザが加速度センサ搭載の包丁型デバイスで行い、終了と判定された場合に次の料理工程に動画が遷移する。このようにしてレシピ動画と現在ユーザが行っている料理工程が同期する。

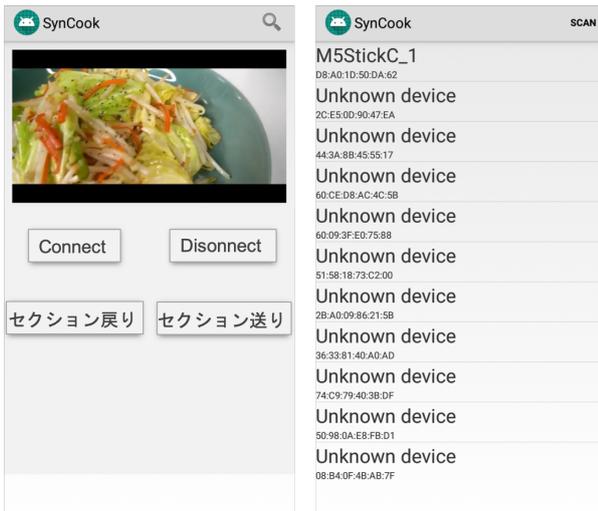


図 8. SynCook のインターフェース

#### 4.4 利用方法

SynCook は Android アプリケーションであり、アプリを立ち上げると図 8 の左側の画面が起動する。起動した後、画面右上にある虫眼鏡のアイコンを押下すると図 8 の右側の図に画面が遷移する。遷移先の画面では BLE デバイスの検出が行われ、検出されたデバイスの名前がリスト表示される。もし仮に検出されなかった場合も画面右上の scan ボタンを押下すると再度 BLE デバイスを検出する。包丁に取り付けられた加速度センサのデバイス名を押下すると図 8 の左側の図に戻り Connect ボタンを押下すると包丁型デバイスと接続され、動画が再生される。あとはレシピ動画通りに食材を切断していただくだけであるが、もし仮に SynCook の誤判定によりレシピ動画との進行度がずれてしまった場合に備え、セクション戻りボタン、セクション送りボタンも用意しており、それぞれ押下することでレシピ動画のセクションを移動する事ができる。

### 5 今後の展望

実装した SynCook を用いた評価実験を行い、システムの有用性を検証していく。評価軸として得られるデータは以下の 3 つである。

- 料理の切断工程が完了しすまでに要する時間
- 料理中にスマートフォンの画面をタッチした回数
- アンケートでの結果

SynCook を使用することで料理中に一時停止や巻戻しを行う必要がないため切断工程が完了するまでの時間および画面をタッチする回数が減少することが期待できる。またアンケートでは評定尺度法によってユーザの負荷を定量的に算出し、SynCook の

有無でユーザが感じる負荷に差がでるかを確認していく。

#### 参考文献

- [1] 株式会社ドゥ・ハウス. 「レシピサイト」に関する調査結果. <https://www.dohouse.co.jp/datacolle/rs20200218/>. (Accessed on 09/30/2020).
- [2] 森岡俊介, 上田博唯. カメラとプロジェクタと対話ロボットが連携する調理支援システム. HAI シンポジウム 2011, 2011.
- [3] A.Sato, K.Watanabe, and J.Rekimoto. Shadow cooking: Situated guidance for a fluid cooking experience. *UAHCI2014*, 2014.
- [4] H.Hasada, J.Zhang, K.Yamamoto, B.Ryskeldiev, and Y.Ochiai. Ar cooking: Comparing display methods for the instructions of cookwares on ar goggles. In *International Conference on Human-Computer Interaction*, pp. 127–140. Springer, 2019.
- [5] Y.Korematsu, S.Daisuke, and M.Nobuaki. Cooking state recognition based on acoustic event detection. *CEA2019*, 2019.
- [6] 大井翔, 池ヶ谷剛, 佐野睦夫. Dp マッチングと移動方向ヒストグラムに基づく一人称視点調理行動認識. 画像電子学会誌, Vol. 46, No. 4, pp. 570–578, 2017.
- [7] S.Ooi, T.Ikegaya, and M.Sano. Cooking behavior recognition using egocentric vision for cooking navigation. *Journal of Robotics and Mechatronics*, Vol. 29, No. 4, pp. 728–736, 2017.
- [8] 宮澤飛鳥, 中村和晃, 橋本敦史, 船富卓哉, 美濃導彦. 調理者の手と容器の位置関係を利用した「かき混ぜる」行動の認識. 電子情報通信学会技術研究報告. DE, データ工学, Vol. 112, No. 75, pp. 25–30, 2012.
- [9] T.kato, H.Kobayashi, A.Yokokubo, and G.Lopez. Proposing a system to supporting for kitchen knife skill improvement using acceleration sensor. *CogInfoCom2020*, 2020.
- [10] YouTube. ラボ飯, ボルシチ風スープ (切断工程のみ). <https://www.youtube.com/watch?v=7dtv1kxseCBg>. (Accessed on 11/18/2020).