

## CookSum : 動画レシピ作成支援システム

橋本 遼太郎 塚田 浩二\* 栗原 一貴†

**概要.** 現在、料理のレシピは数多く存在するが、その多くは文字や画像で説明されたものが多数を占めている。しかし、素人が文字や画像のみのレシピを参考にしてその通りに料理を再現することは難しい。そこで、本研究では、手軽に動画レシピを作成し、短時間で効率的に閲覧するための動画レシピ作成支援システムを提案する。本システムは、まず調理している様子を撮影し、その中で特徴的なタスクを検出し字幕を生成する。さらに、動画の字幕付きの部分とそうでない部分の再生速度を変換して動画を圧縮する。このように、動画を圧縮することで短時間でレシピの概要を効率的に理解することができる。

### 1 はじめに

レシピを参考にして調理をしたものの、上手くいかなかったことは多くの人を経験しているだろう。こうしたケースでは、多く人は、料理本やクックパッド<sup>1</sup>のような文字や画像のみのレシピを用いている。しかし、料理を普段やらない初心者が文字や画像のみのレシピを参考にしてその通りに料理を再現することは難しい。なぜならば、こうした既存のレシピは、暗黙的に一定の料理スキルを前提としているからである。すなわち、材料と手順のみが規定されており、詳細なタイミングは個人の経験に任されていることが多い。初心者にはこのタイミングが分からないため、調理を再現できない可能性が高まってしまふ。そこで、我々は、初心者が料理を再現するには、実際に調理している様子を動画で撮影した「動画レシピ」が有効であると考えた。本研究では、短時間で効率的に閲覧可能な「高速動画レシピ」を手軽に生成するシステム「CookSum」を提案する。



図 1. 調理時の動画撮影の様子。自立型フレームに取り付けたカメラで上部から撮影する。

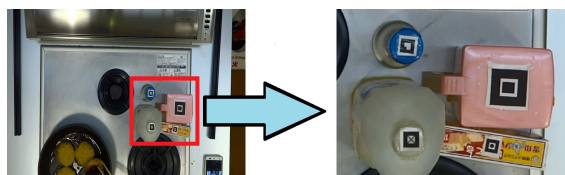


図 2. 撮影動画と調味料等に取り付けた二次元マーカーの一例

### 2 CookSum

CookSum は、手軽に閲覧可能な「動画レシピ」を作成するシステムである。まず、調理過程を動画で記録し、その中での特徴的なタスクを二次元マーカーや音声認識を用いて検出する。次に、検出されたタスクに基づいて自動で字幕を付けて、字幕ファイルを作成する。そして、作成された字幕ファイルを基に撮影した動画を圧縮する。

次に、実装について説明する。まず、独自のフレームにビデオカメラ (SONY 製 DSC-QX10) を取り付け、キッチン上部に設置した上で動画撮影を行う (図 1)。その中で、特徴的なタスク (例: 調味料を入れる、

味見をする) を検出して、動画につける字幕ファイルを自動的に作成する。タスクの検出方式は、二次元マーカーを用いた方式 (図 2) と音声認識を用いた方式の実装を進めている。二次元マーカーの登録/認識には QPToolkit<sup>2</sup> を利用した。登録したマーカーを調味料に付け、キッチン上部から動画撮影した時、常にマーカーが撮影されている状態にする。また、調理過程を撮影後、動画を一定時間 (例: 1 秒) 毎に JPG 形式の静止画に変換する。そこで、各静止画からマーカーを検出していき、一定時間連続で検出されないマーカーがあったときに、「調味料を利用した」と判定する。音声認識を用いた方式では、撮影動画内の発話部分を撮影動画内の発話部分を、MFCC

Copyright is held by the author(s).

\* Ryotaro Hashimoto and Koji Tsukada 公立はこだて未来大学 システム情報科学部 情報アーキテクチャ学科

† Kazutaka Kurihara 津田塾大学 学芸学部 情報科学科

<sup>1</sup> <http://coocpad.com>

<sup>2</sup> QPToolkit <http://kougaku-navi.net/QPToolkit/>

(Mel-Frequency Cepstral Coefficients) を特徴量とする GMM (Gaussian Mixture Model) に基づく音声区間識別 [1] により検出して srt 形式の字幕ファイルの作成を行う。なお、ここでは発話期間のタイミング情報のみを srt 形式で保存している。次に、作成した字幕ファイルを動画に埋め込み、字幕部分とそうでない部分の動画の再生速度を AviSynth スクリプト<sup>3</sup>を用いて変換を行うことで、字幕がある部分は等倍～2 倍程度の遅い速度で再生し、字幕がない部分は 16～32 倍程度の速い速度で再生することで動画の圧縮を行う。そして、Youtube 等の動画サイトにアップロードして、字幕の編集を行う (図 3)。

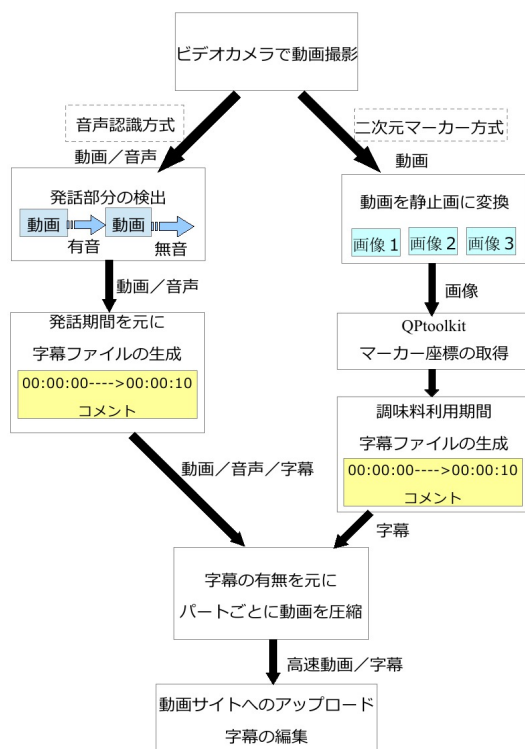


図 3. システム全体の流れ

### 3 関連研究

オリジナルレシピを作成を支援するシステムの研究としては、CookTab[2]がある。CookTabは、任意のレシピを使用した食材と調味料の種類/利用タイミング/使用量を自動的に記録する計量デバイスを用いて、オリジナルレシピの生成を支援するシステムである。また、Kitchen of the Future[3]は、カメラ、ディスプレイ、マイクなどを組み込んだキッチン環境を用いて、調理過程の記録/共有/コミュニケーションを支援するシステムである。MimiCook[4]は、キッチンにデブスカメラ、プロジェクト、重量計等を導入し、調理過程における段階ごとに表

示を行い、ユーザーの調理支援を行うシステムである。CinemaGazer[5]は、字幕の有無を元に、パートごとに映画などの動画の圧縮を行うシステムである。Fabnavi[6]は、ものづくりの過程を写真/動画等で手軽に記録、共有し、再現を支援するシステムである。本研究では、調理を対象として、短時間で手軽に閲覧可能な動画レシピの作成を支援するところが異なる。

### 4 まとめと今後の展望

本研究では、初心者による調理を支援するために、短時間で効率的に閲覧可能な「高速動画レシピ」を手軽に生成するためのシステム「CookSum」を提案、試作した。本システムでは、二次元マーカーと音声認識を用いて調理過程を撮影することで、字幕付きの動画レシピを作成出来る。これらのタスク検出手法はまだ試作段階であり、十分な検証は進めていないが、音声認識については、カメラ内蔵マイクを用いても、発話期間を高い精度で検出することができる感触を得た。二次元マーカーの検出は、若干安定性に欠けるように思われるが、マーカーのサイズや、調味料利用判定のアルゴリズムを工夫することで、安定性を高める工夫を進めていく。さらに、本システムを用いて多くの動画レシピを実際に作成し、作成時の使い勝手を検証すると共に、動画レシピが初心者にとって理解しやすいかどうか併せて検証していきたい。

### 参考文献

- [1] Kazutaka Kurihara, Yoko Sasaki, Jun Ogata, and Masataka Goto, Two-level fast-forwarding using speech detection for rapidly perusing video. In Proceedings of AH '14, Article 19, (2014)
- [2] Ayaka Sato, Koji Tsukada, CookTab: Smart cutting board for creating recipe with real-time feedback Adjunct Proceedings of Ubicomp 2012, pp.543-544, (2012)
- [3] 椎尾 一郎, 浜田 玲子, 美馬 のゆり, Kitchen of the Future: コンピュータ強化キッチンとその応用 コンピュータソフトウェア, Vol.23, No.4, pp.78-83, (2006)
- [4] Ayaka Sato, Keita Watanabe, Jun Rekimoto, MimiCook: A Cooking Assistant System with Situated Guidance In Proceedings of TEI '14, pp.121-124, (2014)
- [5] 栗原一貴, "CinemaGazer: 動画の極限的な高速鑑賞のためのシステムの開発と評価", コンピュータソフトウェア, Vol. 29, No4, pp.293-304, (2012)
- [6] Koji Tsukada, Keita Watanabe, Daisuke Akatsuka, and Maho Oki, FabNavi: Support system to assemble physical objects using visual instructions, Paper presented at Fab10, Barcelona, 2-8 July (2014)

<sup>3</sup> スクリプトベースの動画編集ソフトウェア  
<http://www.avisynth.info/>