

アニメーション生成を目的とした料理ドメイン動作辞書について

大川 寛志 白井 清昭

北陸先端科学技術大学院大学 情報科学研究科

{h-ookawa, kshirai}@jaist.ac.jp

1 はじめに

日常生活において、レシピを見ながら料理をする機会は多い。しかし、特に料理の経験が浅い人は、料理レシピに書かれた手順をなかなか理解できない場合がある。本研究では、人間の料理レシピの理解を助けることを目的に、料理レシピからアニメーションを生成し、ユーザーに提示するシステムを構築することを目的とする。

料理レシピからアニメーションを生成する研究として、植松らや Karlin によるものがある。植松らは、日本語の料理レシピを対象に、レシピに記述されている一連の手順をアニメーションで表現するシステムを構築している [6]。Karlin は、動詞の修飾句を解析し、料理動作の aspekt、繰り返し動作回数、動作時間、スピードなどを特定し、アニメーションを生成する手法を提案している [3]。料理レシピには多種多様な表現が出現するが、これらの先行研究ではスケラビリティの問題、すなわちどれだけの表現をアニメーションに変換できるかについては論じられていない。これに対し、本研究では、様々な料理レシピに対してアニメーションを生成することのできる頑健なシステムの構築を目指している。しかしながら、材料、器具、動作など、アニメーション生成の対象となる要素は数多くあり、多様なレシピに対応するのは決して容易ではない。そこで、本研究では、レシピ文中の動作表現に着目し、動作表現をアニメーションに変換することを第一の目的とする。すなわち、料理レシピで説明されている一連の手続きを1つのアニメーションとしてユーザーに提示するのではなく、レシピの中にユーザーがよく理解できない動作表現があれば、それに対応したアニメーションを提示することを目指す。具体的には、レシピ文書をユーザーに提示し、ユーザーが理解しにくい動作表現をマウスで指定してもらい、それに対応するアニメーションを提示するシステムを構築する。レシピには「そぎ切りにする」や「三枚におろす」などのように料理特有の動作表現が多く含まれる。料理に慣れていない人にとっては、レシピ全体の流れをアニメーションで見ることができなくても、このような料理特有の動作表現だけでもアニメーションで見ることができれば、レシピを理解する助けになると考えられる。

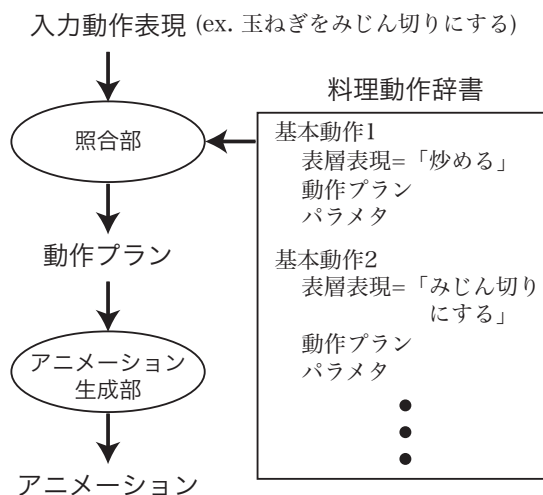


図 1: システム概要

2 提案システムの概要

1 節で述べたように、本研究は、料理レシピ中の1つの動作表現を入力とし、その動作表現に対応したアニメーションを生成するシステムの構築を目的とする。提案システムの基本的な処理の流れを図 1 に示す。

まず、料理動作辞書を用意する。料理動作辞書とは、「炒める」「みじん切りにする」といった料理に関する基本的な動作の集合であり、各動作についてアニメーションの生成に必要な知識が記述されている。ここでは料理動作辞書の1つのエントリを基本動作と呼ぶ。各基本動作には、表層表現、動作プラン、パラメタを記述する。表層表現は、基本動作に対応する言語表現である。動作プラン、パラメタはアニメーションの生成に必要な知識である。詳細は 3.2 項で述べる。以上のように、料理動作辞書は自然言語で書かれた動作表現とアニメーションを対応付ける役割を持つ。

アニメーションを生成したい動作表現が入力されたとき、システムは料理動作辞書の基本動作に記載されている「表層表現」との照合を行い、入力に対応する基本動作を選択する。次に、選択された基本動作に記載されている動作プランとパラメタをアニメーション生成部に渡す。最後に、アニメーション生成部は動作プランを解釈し、対応するアニメーションを生成する。

多様なレシピに対応するために、本研究では特に以下の2点に重点を置く。

1. 大規模な料理動作辞書の構築 (3節)

多様な動作表現に対して適切なアニメーションを生成するためには、多くの料理の基本動作を辞書に登録する必要がある。ここでは、料理レシピに現われる大部分の基本動作を含むような比較的大規模な料理動作辞書を構築する。

2. 類義の動作表現への対応 (4節)

料理レシピに含まれる動作表現は多様であり、その全てを辞書に記述するというアプローチには限界がある。そこで、料理動作辞書に含まれない未知の動作表現が入力されたとき、辞書の中から全く同じ、あるいは類似している基本動作を選択し、アニメーションを生成する。例えば、「刻む」が未知の動作表現であったとする。このとき、システムは「刻む」と類似する基本動作「みじん切りにする」を選択し、それに対応したアニメーションを生成する。

3 料理動作辞書の構築

3.1 基本動作の定義

料理動作辞書を構築する最初のステップは、辞書のエントリとなる基本動作を定義することである。ここで問題となるのは、どのような基準で基本動作を定義するかである。本研究では、料理の基礎について解説している本を参照し、そこに記述されている動作を基本動作と定義する。料理の本を基本動作の定義の基準とした理由については文献 [7] を参照していただきたい。

本研究では、三冊の料理の本 [1, 2, 5] を参照した。これらは、材料別に切り方、焼き方、皮・種の取り方など、料理の基本的な操作を項目として取り上げ、それぞれの項目を写真入りで解説している。これら3冊の本に項目として挙がっている動作を収集し、基本動作とした。

2節で述べたように、本研究の目的の一つは大規模な動作辞書を構築することである。そこで、以上のようにして作成した動作辞書が、料理レシピ文書に出現する動作をどれだけ含むかを調べる予備実験を行った。その結果、レシピに出現する動作の約50%しか動作辞書に含まれないことがわかった。この原因を調べたところ、「加える」「入れる」「のせる」など、レシピ以外にも現われる一般的な動詞で表現される動作であることがわかった。料理の本ではこのような基本的な動作は解説されていないため、辞書に含まれなかったと考えられる。そこで、200個のレシピ文書コーパスを用意し、辞書に該当する

表層表現	4つ割りにする
動作プラン	cut(材料, 器具, 場所, 1) rotate_90(材料, 場所) cut(材料, 器具, 場所, 1)
パラメタ	材料=[じゃがいも] (ヲ格, ハ格) 器具=包丁 場所=まな板

図2: 動作プランとパラメタ

動作がなく、かつ10回以上出現する動作表現を基本動作として新たに辞書に追加した。コーパスの高頻度語を基本動作とする場合、料理動作辞書が参照したコーパスに依存するために好ましくない。しかし、コーパスから追加される基本動作は極めて一般的なものが多く、ある程度の汎用性は認められると考える。

最終的に、料理の本から265、コーパスから32、合計297個の基本動作を収集し、初期の動作辞書を作成した。

3.2 動作プランとパラメタの記述

個々の基本動作に対して動作プランとパラメタを人手で記述する。「動作プラン」とは、基本動作のアニメーションを生成するために必要な個々の動作(以下、動作プリミティブと呼ぶ)の列である。動作プリミティブはアニメーションを生成する際の基本となる動作であり、システムによってユニークなアニメーションに変換される動作の最小単位である。「パラメタ」は、材料、器具、場所など、動作プランに渡すべきパラメタを表わす。パラメタは入力レシピ文によって変化する場合もある。例えば、「材料」は入力される動作表現の内容によって変化する。この場合、動作辞書にはデフォルトのパラメタを記載する。また、入力文におけるパラメタの出現位置に関する情報も記載する。

例として、「4つ割りにする」という基本動作の動作プランとパラメタを図2に示す。図2の動作プランは、この基本動作が、材料を1回切る、材料を90度回転する、材料を1回切る、という3つの動作プリミティブから構成されることを示している。一方、この基本動作のパラメタは材料、器具、場所である。特に材料については、デフォルトの値が「じゃがいも」であり、レシピ文中ではヲ格またはハ格の格要素として出現することを表わしている。

現在、80個の基本動作に対して動作プランとパラメタを記述した。ただし、パラメタの入力文での出現位置に関する情報は記述していない。残りの情報については

表 1: 料理動作辞書の評価

動作表現数	1929
(a) 基本動作	1551 (80.4%)
(a1) 表層表現一致	1088 (56.4%) [65.0%]
(a2) 表層表現不一致	463 (24.0%) [27.6%]
(b) 基本動作以外	124 (6.4%) [7.5%]
(c) 解析誤り	91 (4.7%)
(d) 動作表現でないもの	163 (8.5%)

今後記述する予定である。

3.3 評価実験

作成した料理動作辞書が実際の料理レシピに現われる動作表現をどれくらい含んでいるかを調べるために、簡単な評価実験を行った。まず、インターネットのレシピ紹介サイト¹から100個の料理レシピを獲得し、テストコーパスとした。これは辞書に高頻度語を基本動作として追加したときに用いたコーパス(3.1項参照)とは異なる。レシピ文の形態素解析、文節の係り受け解析を行い、動作表現を抽出した。形態素解析にはJUMAN²を、文節の係り受け解析にはKNP³を用いた。ここで抽出した動作表現とは、用言とそれに係る格および格要素の組である。これらの動作表現について、対応する基本動作が料理動作辞書に含まれているかどうかを手で調査した。この際、解析して取り出された動作表現とともに元のレシピ文も参照した。結果を表1に示す。

表1の数値は各項目に該当する動作表現の数であり、()内は全体に対する割合である。(c)は解析の誤りによって正しく抽出することができなかった動作表現の数を、(d)は「(次にソースを)作る」「(出来たら熱いうちに)食べる」のような調理動作を表わさない動作表現の数を表わす。これらを除いた動作表現が、本研究でアニメーションを生成する対象となる。(a)は料理動作辞書のいずれかの基本動作に該当する動作表現の数であり、(a1)は辞書に記載されている表層表現と完全に一致するもの、(a2)は一致しないものを表わす。一方、(b)は辞書中に該当する基本動作が存在しない動作表現の数である。また、[]内の数値は、アニメーションを生成すべき動作表現全体(a1+a2+b)に対する各項目の割合を示している。この結果から、アニメーション生成の対象となる動作表現のうち、約92%は料理動作辞書に含まれることがわかる。この結果から、レシピ文書に現われる動作表現の大部分を包含する比較的規模の大きい料理動

作辞書を構築できたといえる。

4 照合部

4.1 動作表現の柔軟な照合

照合部では、レシピ文中の動作表現 ve_{rec} と辞書に記載されている基本動作の表層表現 ve_{dic} を照合し、対応する基本動作を1つ選択する。ここで、 ve_{rec} 、 ve_{dic} はともに動詞、格、格要素で構成される格構造である。ここで問題となるのは、「刻む」と「みじん切りにする」のように、 ve_{rec} と ve_{dic} の表層表現が必ずしも一致しないことである。多くの動作表現に対して適切なアニメーションを生成するためには、表層表現が一致しなくても適切な基本動作を選択できるような柔軟な照合が必要となる。

照合部の処理の流れは以下の通りである。まず、以下の方針にしたがって ve_{rec} と ve_{dic} の照合を行う。

- ve_{rec} と ve_{dic} の動詞を比較し、一致するかどうかを調べる。 ve_{dic} が格要素を持つ場合は、その格要素についても比較を行う。
- 表記での照合に失敗したときは読みでの照合を試みる。 ve_{rec} の読みは形態素解析の結果を利用し、 ve_{dic} の読みはあらかじめ人手で記述する。
- ve_{rec} の動詞を基準として、動詞の前の1~3単語を連結した表現を取り出し、 ve_{dic} との照合を行う。同様に、動詞の後の1~3単語を連結した表現についても照合を試みる。これは、 ve_{rec} が「かつら/むき/に/する」で ve_{dic} が「かつらむき/に/する」のように、両者の形態素区切りが異なる場合を考慮したためである⁴。
- ve_{rec} の動詞とその2つ前にある形態素を連結した表現を取り出し、 ve_{dic} との照合を行う。これは、 ve_{rec} が「酒蒸し/に/する」で ve_{dic} が「酒蒸し/する」のように、助詞が省略されている場合を考慮したためである。
- ve_{dic} にある格要素が ve_{rec} に存在しなくても、動詞が一致していれば照合に成功したとみなす。これは、レシピ側で格要素が省略されている可能性があるためである。
- 上記で考慮した読みや形態素区切りの違いは比較的照合が容易であるが、中には表現そのものが全く異なる場合もある。例えば、 ve_{rec} が「砂出し/を/する」で、 ve_{dic} が「塩水/に/ひたす」のとき、両者の

¹<http://www.bob-an.com/>

²<http://www.kc.t.u-tokyo.ac.jp/nl-resource/juman.html>

³<http://www.kc.t.u-tokyo.ac.jp/nl-resource/knp.html>

⁴/ は形態素区切りを表わす。

表層表現は全く異なるが同じ動作を表わしている
 とみなせる。本論文では、辞書中の各基本動作につ
 いて、このような等価な表現をコーパスから収集し、
 辞書に登録した。等価表現の収集に用いたコーパス
 は、3.1 項で基本動作の定義に用いた 200 レシピと
 同じものを用いた。137 個の基本動作に対してのべ
 438 個の等価表現を登録した。

一般に、上記の照合によって、1 つの ve_{rec} に対し
 て複数の基本動作の ve_{dic} が照合に成功する。しかし、ア
 ニメーションを生成するためには基本動作を 1 つに決め
 る必要がある。そこで、照合によって複数の基本動作が
 得られたとき、以下の基準に基づいて基本動作の数を絞
 り込む。

1. 動詞のみが一致する基本動作よりも、動詞と全ての
 格要素が一致する基本動作を優先する。
2. 形態素を連結して照合を行ったとき、多くの形態素
 を連結した場合を優先する。
3. 読みで一致した基本動作よりも表記で一致した基本
 動作を優先する。

以上の基準に基づいて基本動作の数を絞り込んでも、同
 じ基準を満たす基本動作が複数あり、1 つに絞り切れな
 い場合がある。本来は基本動作は 1 つだけ選択するべき
 であるが、本論文ではその全てを選択することにした。

4.2 予備実験

照合部を評価する予備実験を行った。表 1 の (a1),(a2)
 に該当する 1,551 個の動作表現を ve_{rec} とし、照合部によ
 って適切な基本動作を選択できたかどうかを調べた。
 結果を表 2 に示す。(c) は適切な基本動作の選択に失敗
 した場合を表わす。一方、(c) 以外の約 87% の動作表現
 については適切な基本動作を選択することができた。こ
 のうち、(a) は適切な基本動作のみを選択することがで
 きた場合を、(b) は適切な基本動作を含む複数の基本動
 作が選択された場合を表わす。照合部の目的は、 ve_{rec}
 が与えられたとき、アニメーションを生成する基本動作
 を 1 つ選択することであるので、(b) に該当する動作表
 現が 23.8% も存在することは望ましくなく、改善の余地
 がある。

現在の照合部の問題は大きく分けて 2 つある。ひとつ
 は、 ve_{rec} と ve_{dic} の表層表現が完全に異なる場合への
 対応である。現在は、200 レシピから得られた等価表現
 を人手で辞書に登録しているが、等価表現を人手で収集
 するアプローチには限界がある。したがって、柴田らの
 研究 [4] のように、大量のコーパスから等価表現 (類義

表 2: 照合部の評価

	動作表現数
(a) 成功 (1 つ抽出)	975 (62.9%)
(b) 成功 (複数抽出)	369 (23.8%)
(c) 失敗	207 (13.3%)

表現) を自動的に獲得する手法を検討すべきである。ま
 た、現在の料理動作辞書では、表層表現が同じでも材料
 によって異なる動作をする場合は、それぞれ別の基本動
 作として定義している。(b) のように適切な基本動作は
 選択できても他の不適切な基本動作を排除できない理由
 は、表層表現が同じで材料が異なる基本動作が多く存在
 するからである。現在の照合部では、基本動作を 1 つに
 絞り込む際に、材料については全く考慮していない。し
 たがって、 ve_{rec} の材料と ve_{dic} の材料との類似度を計算
 するなど、基本動作が対象とする材料を考慮に入れる必
 須がある。このとき、材料間の類似性は料理分野に特化
 していると考えられるので、既存のシソーラスを利用す
 る代わりに料理分野に特化したシソーラスをコーパスか
 ら自動的に獲得する必要があると考えられる。

5 おわりに

本論文では、料理レシピからアニメーションを生成す
 るシステムの概要について述べ、レシピ中の動作表現と
 アニメーションを対応付けるための料理動作辞書を構築
 した。また、レシピ中の動作表現と辞書の基本動作の表
 層表現の照合について述べた。今後は、3.2 項で述べた
 ように辞書の動作プランとパラメタを記述するとともに、
 4.2 項での考察に基づき照合部の改良を行いたい。

参考文献

- [1] 熱田陽子. 野菜の切り方 BOOK. 集英社, 2004.
- [2] 藤野嘉子. 新・料理の基本. SS コミュニケーションズ, 2003.
- [3] Robin F. Karlin. Defining the semantics of verbal mod-
 ifiers in the domain of cooking tasks. In *Proceedings
 of the Annual Meeting of the Association for Computa-
 tional Linguistics*, pp. 61–67, 1988.
- [4] 柴田知秀, 黒橋禎夫. 料理教示発話の理解と作業構造の自
 動抽出. 情報処理学会情報処理学会自然言語処理研究会,
 pp. 117–122, 2004. NL-164-20.
- [5] 高城順子, 検見崎聡美. スタンダードクッキング 料理の基
 礎. 小学館, 2004.
- [6] Hideki Uematsu, Akira Shimazu, and Manabu Oku-
 mura. Generation of 3D CG animations from recipe
 sentences. In *Proceedings of the Natural Language Pro-
 cessing Pacific Rim Symposium*, pp. 461–466, 2001.
- [7] 白井清昭, 大川寛志. アニメーション生成のための料理動
 作辞書の構築. 情報処理学会情報処理学会自然言語処理研
 究会, pp. 123–128, 2004. NL-164-21.