

Title	レシピ文入力からの3DCG調理画像生成システムの開発
Author(s)	植松, 秀樹
Citation	
Issue Date	2001-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1434
Rights	
Description	Supervisor: 島津 明, 情報科学研究科, 修士

修 士 論 文

レシピ文入力からの3DCG調理画像
生成システムの開発

指導教官 島津 明 教授

北陸先端科学技術大学院大学
情報科学研究科情報処理学専攻

植松秀樹

2001年2月15日

要旨

言語によって伝えられる情報を、コンピュータ・グラフィックスにより可視化することができれば、受け手によりよく情報を理解させることが期待される。

そこで、本研究では、言語表現を動画に変換する問題を取り上げ、「料理レシピ」を対象に、レシピ文を入力として与えると、調理画像を生成するシステムの開発を目指す。

本研究では、まず、画像化することを念頭に置いて、レシピ文の分析を行ない、画像化すべき「画像要素」の特定と、時間関係を明確にする手がかりとなる「調理動作」の分類を行なった。またレシピ文から抽出すべき情報に加えて、知識として予めシステムに保持させる情報についても特定した。画像化に必要な上記の情報を含んだ中間表現を生成し、自動的に調理画像を生成する実験システムを構築した。

目次

1	はじめに	1
1.1	研究の背景と目的	1
1.2	本論文の構成	2
2	関連研究	3
3	レシピ文の言語表現の分析	5
3.1	分析の目的	5
3.2	調理画像要素	6
3.3	調理動作	7
3.3.1	調理動作の終点について分類	7
3.3.2	副詞句の分類	8
3.3.3	連続した調理動作の分類	9
4	画像化に関する検討	11
4.1	画像要素「材料」の形状変化と内部構造	11
4.2	画像要素間の作用	13
5	実験システム	14
5.1	システムの処理の流れ	14
5.2	中間表現	15
5.3	レシピ文から中間表現への処理内容 (1)	16
5.4	レシピ文から中間表現への処理内容 (2)	18
5.4.1	全ての調理動作の特定	18

5.4.2	材料の形状の変化	19
5.4.3	材料の場所	19
5.5	レシピ文から中間表現への処理内容 まとめ	20
5.6	調理画像	22
5.6.1	画像要素のモデリングとアニメーションの設定	22
5.6.2	画像要素の作用スクリプトと画像再生スクリプト	22
6	レシピ文から画像を生成する動作例	26
6.1	画像化するレシピ文	26
6.2	中間表現の生成	27
6.2.1	レシピ文から読み取れる画像要素	27
6.2.2	中間表現を生成する処理の流れ	27
6.2.3	レシピ文から生成された中間表現	32
6.3	調理画像の生成	33
6.3.1	画像要素	33
6.3.2	アニメーション画像とスクリプト	34
6.3.3	画像再生スクリプト	34
6.3.4	レシピ「アスパラガスとウインナのじか炒め」の調理画像	35
6.3.5	レシピ文を変えた場合の動作	37
7	考察	40
7.1	調理動作	40
7.1.1	パターン 1・パターン 2	40
7.1.2	パターン 3	42
7.1.3	パターン 4	42
7.2	調理画像	42
7.2.1	材料	42
7.2.2	動作	43
8	おわりに	44
8.1	まとめ	44
8.2	今後の課題	45

目 次

4.1	魚の内部構造	12
5.1	システムの処理の流れ	14
5.2	有限数のステップから構成される調理手順	15
5.3	中間表現	16
5.4	知識：3つの辞書	17
5.5	動作辞書における「切る」の表現	18
5.6	調理動作「(材料を場所Aから場所Bへ)移す」の表現	18
5.7	内部構造に分割する形状変化	20
5.8	Maya 2.0 でアニメーションを生成する場合	23
5.9	Digital Loca でアニメーションを生成する場合	23
5.10	画像要素の作用	24
5.11	「中間表現」 「画像生成」	25
6.1	材料・調味料の動作ステップ毎の場所の変化	31
6.2	材料オブジェクトの内部構造	33
6.3	作用スクリプトによる再生画像の切替え	34
6.4	画像再生スクリプト	35
6.5	初期画面	35
6.6	動作ステップ2の調理画像	36
6.7	動作ステップ11の調理画像	36
6.8	動作ステップ12の調理画像	37
6.9	動作ステップ3の調理画像	39
6.10	動作ステップ10の調理画像	39

7.1 終点が特定できない動作の表現	43
------------------------------	----

表 目 次

6.1 対象レシピにおける手順毎の「材料」と「動作」	27
6.2 中間表現	32

第 1 章

はじめに

1.1 研究の背景と目的

近年、ハードウェアの高性能化によって、コンピュータ上での3次元グラフィックス画像技術の進歩は目覚ましいものがある。そして、それらの画像は我々の普段の生活の中にも広く浸透しているが、多くの場合、映画やゲームなどの娯楽の中に使用されるものであり、必ずしも普段の生活の中で必要なものではなかった。

しかし、最近になって言語処理と3次元グラフィックス画像を結び付けようという研究がいくつか行なわれている。ある議題に関してCGで描かれた仮想エージェントにプレゼンテーションを行なわせる研究や、オフィス業務におけるノウハウを持った人間に代わって、擬人化エージェントが問い合わせに答えるシステムの開発、またVR空間に描かれた人物や風景を様々な角度から見るために、話し言葉でカメラに対し位置に関する指示を与える研究などがある。

このように、言語処理とグラフィックス画像を結び付ける研究が行なわれているのは、言語情報を可視化することで、受け手に情報をよりよく理解させることが期待されるからである。

このような言語情報から画像を生成するシステムに関して、考慮すべき点として以下のものが挙げられる。

- 言語情報から抽出すべき情報は何か
- 言語情報に加えて補完すべき情報は何か

- どれだけの画像データを用意すべきか

以上のような背景から、本研究では言語表現を動画に変換する問題を取り上げ、「料理レシピ」を対象にする。

普段料理をすることのない人達にとって、レシピ本の文章を読んだだけで手際良く調理をすることは困難である。これはレシピの文章だけでは効率的な調理手順を正しく理解できなかったり、レシピ文に含まれる調理分野特有の表現や調理技法を知らないという点が原因として挙げられる。

本研究では多くの料理レシピに対応できる知識ベースの構築と、調理の時間関係をレシピ文から明らかにして画像生成することを目的とする。料理レシピ文から調理画像へ変換する際に、どのように言語表現を取り扱う必要があるかについての分析を示し、その分析に基づいて調理画像生成システムを構築する。

1.2 本論文の構成

本論文は本章を含め8章で構成される。第2章では、関連研究について述べる。第3章では、画像化に必要な情報を特定するために行なったレシピ文における、言語表現の分析について述べる。第4章では、分析に基づいて行なった、画像化に関する検討について述べる。第5章では、本研究で構築した実験システムについてアルゴリズム等の説明を行なう。第6章では、構築した実験システムによる、レシピ文から調理画像生成する動作例について述べる。第7章では、考察を述べ、そして最後に第8章では、本研究のまとめと今後の課題について述べる。

第 2 章

関連研究

近年, 言語処理と画像生成を結び付けた研究が数多くなされている.

[Strassmann 94] では, 自然言語入力により仮想エージェントを半自動的に動かすことを試みている. コンピュータによって生成されたキャラクタの振舞いは, おおよそユーザによって与えられる外的なソース(台本・指導)によって制限される. 同様に仮想エージェントを動かす研究として, [Loyal 他 93] では更にキャラクタに感情を持たせ, その振舞いに自主性を与えている.

また, [Cassell 他 94] では, 複数のエージェント間の会話を自動的に生成・アニメ化するシステムを構築した. この研究では, 発話に加え, 発話に意味的にリンクしているジェスチャーをエージェントに行なわせることで, より自然な会話のアニメ化に成功している.

また, [Townes 他 98] では, 高校の物理分野において, ユーザである学生に複雑な空間的概念を 3DCG で分かりやすく, 動的に説明する教育システムを構築している. このシステムでは, CG による説明と発話による説明の 2 つがシンクロして行なわれる. 3Dアニメーションプランナは, 幾何学的な形式での空間知識を必要とし, また自然言語ジェネレータは言語的な形式での空間知識を必要とするため, リアルタイム・マルチメディア・プランナが mediator (仲裁人) として処理を整合している. 同様に, プレゼンテーションを目的として, CG と発話を結び付けた研究に [Andre 96] がある. アニメ化されたユーザ・インタフェース・エージェントは, モデムの回路のようなオブジェクトに対して, 箇所の名前や構造についての説明を発話により行なう. エージェントの動作と発話の時間的なコーディネートのために, システムは時間軸に沿って表現された temporal structure の処理を行ない, 効果的なダイナミックプレゼンテーションを実現している.

また、料理分野を扱ったコンピュータアニメーション生成の研究として、[Karlin 88] がある。英語の部分集合を使うことで、「調理アニメーション生成のための意味解析 SEAFAC-T(Semantic Analysis For the Animation of Cooking Tasks)」は、ユーザに料理定義を提示することが出来る。英語入力を分析し、動作手順に沿って行なわれるタスクの表現を生成する。この研究では [Moens 87] の定義に従って、イベントをアスペクトに関して、「最高点」「プロセス」「最高点に達するプロセス」「ポイント」の4つに分類し、料理領域に見られるイベントの多くは「最高点に達するプロセス」であるとしている。また、英語入力を処理する上でイベントの特定のために、動詞の修飾語の意味解析について分類も行なっている。

[Adachi 97] では、「知識獲得、情報抽出」と「要約」に焦点を絞り、領域特定知識を得る方法を示し、料理領域に関して「レシピ」を取り上げている。日本語で書かれたレシピ本の中に見られる多くの料理名は、「AのB」のように名詞の混合によって記述されている。ここで、Bは「基準となる料理の名前」を示し、Aは「その料理を特徴付けるメイン原料の名前」を示している。この研究では、類似した2つの料理レシピ間の共通部分を抽出し、料理定義を生成している。

[Webber 90] では、「Animation from Instructions」システム構築のために、自然言語命令文における Free Adjuncts(自由付加句) について分析を行なっている。自由付加句とは副詞従属節の機能を持った述語として定義できる。ここでは、命令文における主節と付加句の関係を、「Augmentation」「Generation」「Simultaneity」3つのクラスに分類している。

第 3 章

レシピ文の言語表現の分析

3.1 分析の目的

本研究では、先ず、料理レシピ文における言語表現の分析を行なった。分析の目的は、言語表現から画像を生成するまでの処理を行なうために必要な情報が何であるか、を明確にすることである。

通常、動作に関する言語表現は、動作全体を包括して指し示すことが多く、その動作を行なうために必要な具体的な細かい動作の記述は省略されている。「車を運転する」という動作を例に挙げると、この動作は実際には、「車のドアを開ける」「運転席に座る」「キーを差し込む」などの一連のサブ動作から成り立っている。

料理領域における動作についても同様のことが言える。レシピ文の言語表現から調理画像を生成するには、以下の 3 点を明確にする必要がある。

- 言語情報から抽出すべき情報は何か
- 言語情報に加えて補完すべき情報は何か
- どれだけの画像データを用意すべきか

上記の 3 点について明らかにするために行なった、料理レシピ文の分析の結果を次頁より示す。

分析の結果、料理レシピを動画に変換するためには、「調理画像要素」と「調理動作」を明確にする必要があることが分かった。以下、それぞれについて説明する。

3.2 調理画像要素

調理画像生成のためには、実際に画面上に表示させるグラフィック画像要素を明らかにしなければならない。レシピ文の分析の結果、画像要素として「材料」「調理動作」「調理器具」があることが分かった。また、動作がどこで行なわれるか、あるいはどこに向けて行なわれるかといった「調理場所」の特定も必要となる。画像要素として、

- 材料
- 調理動作
- 調理器具
- 調理場所

について、それぞれ明らかにする必要がある。

しかし、料理レシピ文によっては、これらの画像要素に関する記述が全てについてなされていないというわけではないため、全てを料理レシピ文から特定することは困難であることが分かった。

例えば、

(例)「いんげんは、斜め切りにします」

のような手順では、調理動作「切る」ために必要な調理器具「包丁」についての記述がない。調理場所についても同様のことが言える。

しかし、多くの場合、調理動作が何であるかによって、調理器具と調理場所を特定することができる。

以上のことから、調理器具と調理場所に関する情報を保持した調理動作の辞書をあらかじめ用意し、システムがレシピ文の解析を行なう際、動作辞書を参照しながら、各画像要素を特定していくことにする。

詳細については、5章「システムの構成」で述べる。

3.3 調理動作

本節では、材料に与える調理動作に関して行なった分析について述べる。調理動作に関して、3つの点に着目して分析を行なった。

- 調理動作の終点についての分類
- 副詞句の分類
- 連続した調理動作の分類

3.3.1 調理動作の終点について分類

料理は限られたリソースで実行される有限数のステップから成り立っている。そのため、料理領域における全てのステップ、つまり調理動作は必ず「終点」を持ち、それにより材料に新しい状態への変化をもたらす。

調理動作はこの「終点」に関して、2つに分類できる。

- 「終点」を持つ動作

「終点」を持つ動作とは、その動詞単独ですでにその終点を特定できるような意味を含んでいる。

(例)「切る」「加える」「むく」「取る」など

- 「終点」を持たない動作

「終点」をもたない動作とは、その動詞単独では終点を特定することができない動作である。

(例)「焼く」「煮る」「炒める」「かき混ぜる」など

- 調理分野特有表現によって表される動作

状態や時間に関する情報がある程度保持していて、調理者の知識に依存する動作。通常、動詞を修飾する副詞を伴うことはない。

「万能ねぎは小口切りにする」

「大根は皮をむき、薄いいちょう切りにします」

3.3.2 副詞句の分類

動作は、動詞を修飾する副詞句によって、より具体的に特定される。また、副詞句は動詞の意味と目的において、本来固有のものではない動作に関する情報を特定する助けとなる。これらの副詞句についても分類を行なった。

- 「状態」に関する副詞句

- 数値を含んだ具体的な「状態」に関する修飾語を伴う場合

「玉ねぎは 1センチ角に、ベーコンは 1センチ幅に 切ります」

「にんにくの茎は水で洗い、3～4センチの長さに切ります。豚肉は 2～3ミリ幅に 切りにします」

- 抽象的に表現された「状態」に関する修飾語を伴う場合

「小さじ1ほどの塩を 平均に ふって」

「いんげんはゆで、食べやすい長さに斜め切りにします」

- 「期間」に関する副詞句

前述したような「終点」が特定できな動作を修飾し、その動作をいつまで行なえばよいかの判断の手がかりとなる。

- 数値を含んだ具体的な「期間」に関する修飾語を伴う場合

「牛乳を加えて 2～3分 煮て」

「1～2時間 おいてなじませます」

- 抽象的に表現された「期間」に関する修飾語を伴う場合

「熱湯で さっと ゆでます」

「170度Cの油に入れ、ゆっくりと 揚げます」

- 材料の形状を表す表現を用いて「期間」を特定する場合

「豚肉を加えて 色が変わるまで 炒め」

「汁気がなくなるまで 煮て」

3.3.3 連続した調理動作の分類

料理タスクのそれぞれのステップは有限シーケンスである。本節では、それぞれの動作ステップがどのタイミングで行なわれるべきか、言い替えると、どのタイミングで画像化すべきか（またはすべきでないか）について、言語表現の分析から明らかになった点について述べる。

ここでは、レシピ文中に頻繁に出現する、動詞の連用形で接続した文（「A、B」）を例に挙げる。動作が2つ連続した場合、その調理の画像化に関して4つのパターンが存在する。

- パターン1 「Aの画像 + Bの画像」

レシピの手順通りに、Aの画像を再生した後、Bの画像を再生する場合。

（例）「なすはへたを取り、1 cmの厚さに輪切りにします」

「なすのへたを取る」動画のあとに「1 cmの厚さに輪切りにする」動画を順番に流す。

- パターン2 「Aの画像中にBのサブ画像を重ねる」

Aの画像再生中に、Bの画像を重ねて同時に再生する場合。

（例）「玉ねぎを炒め、小麦粉を加える」

「玉ねぎを炒める」画像のある時点で、「小麦粉を加える」画像を重ね合わせる。

- パターン3 「Aだけ画像化（Bは目的）」

Aの画像のみを再生すればよい場合。

（例）「なめこはざるに水で洗い、ぬめりを落す」

Bが「ぬめりを落す」「香りをつける」などの場合には、「BはAの目的」であり、この例のように水で洗うだけでよい時には、画像化する必要はない。

- パターン4 「AとBの間に画像Cをはさむ」

Aの画像とBの画像の間に新たに画像Cをはさむ必要がある場合。

(例)「大根を水でさっと洗い, 鍋で煮る」

AとBの間に大根を鍋に移す動作画像Cを挿入する.

前述したように, 調理動作には動作の「終点」が判断出来る動作と, 判断できない動作が存在する. それらの動作を表す動詞と, レシピ文中の動作期間を表す副詞などの情報から, 時系列に関して明らかにし, 適切な順序で画像化を行なうこととする.

第 4 章

画像化に関する検討

3章において、レシピ文における言語表現の分析を示した。言語情報から調理画像を生成するためには、言語表現の分析結果に加えて、画像化に関する検討を行なう必要がある。本章では、画像要素に着目し、「画像要素「材料」の形状変化と内部構造」「画像要素間の作用」に関する行なった検討について述べる。

4.1 画像要素「材料」の形状変化と内部構造

料理分野において調理する材料の種類は多々あるが、本研究においては画像化することを考慮した上で画像要素である「材料」の分類を行なう。材料は、その形状に関して以下のように分類できる。

- 形状が固形のもの
- 形状が液状のもの

また、多くの場合、調理を進めるにしたがって、材料はその形状を著しく変化させる。例えば、

(例)「ピーマンはへたと種を除いてせん切りにする」

のようなケースでは、調理後のピーマンは「へたと種を除かれ」、「せん切り」にされたピーマンであり、調理前の形状とは大きく異なる。このように、調理の最終状態において、材料が調理前の原形をとどめていない場合が多い。本研究においては、調理画像を生

成する際に、調理に必要な材料のCG画像を予めモデリングしておくが、上記のことから、ピーマンの原形のモデルだけを用意しても、それだけでは調理画像を生成することはできない。従って、モデリングを行なう際に、調理が進むにつれて材料が変化していくことを考慮しなければならない。

上記のレシピにおいて、調理動作によって材料に与えられる形状変化は、内部構造に分けられていく変化である。材料が内部構造を持ち、調理において材料の構造を考慮しなければならないことは、ピーマンのような野菜類に限らず、魚類(図 4.1)、肉類など多くの材料に言えることである。

魚	頭
	身
	内蔵
	うろこ
	ひれ

図 4.1: 魚の内部構造

上記の動作は、調理前も調理後も材料の形状は固形であったが、調理によっては、固形の材料を「かき混ぜる」「すりおろす」などの動作によって液状、またはほぼ液状に変化する場合もある。また逆に、液状の材料を冷凍庫などで凍らせるなどの動作によって、材料が固形に形状を変化させる場合もある。当然、調理の前と後で液状のまま変わらない材料も存在する。

これらのことから、調理動作によって変化する材料の形状に関してまとめると、以下のようになる。

材料の形状変化

- 固形 固形
- 固形 液状
- 液状 固形
- 液状 液状

4.2 画像要素間の作用

4.1 節において、材料は調理が進むにつれて形状を変化させる場合があることを述べた。画像要素「材料」がその形状を変化させるのは、ある画像要素「調理動作」が行なわれた場合である。本研究では、システムに汎用性を持たせるため、「材料」と「調理動作」のグラフィックス画像をそれぞれ別に用意し、料理レシピに応じて必要な画像を組み合わせて、調理画像を生成する。

料理レシピ中に「アスパラガスを切る」という調理があった場合に、必要となる画像要素は「アスパラガス(材料)」と「切る(調理動作)」である。ここで、自然な調理画像を生成するためには、「切る(調理動作)」が再生中に、「包丁(調理器具)」が「アスパラガス(材料)」と接触したら、切られていく画像を再生する必要がある。このように、各画像要素をあらかじめ生成するだけでなく、画像要素同士の接触による作用についても考慮しなければならない。

第 5 章

実験システム

本章では、3章で示したレシピ文の言語表現の分析結果と、4章で示した画像化に関する検討に基づいて設計・構築した調理画像生成システムについて述べる。

5.1 システムの処理の流れ

構築したシステムの、レシピ文の入力から調理画像を出力するまでの処理の流れを示す。入力であるレシピ文と知識から画像化に必要な情報を含んだ中間表現を生成する。そして、中間表現に記述された情報に従い、あらかじめ作成した画像要素を呼び出し、再生する(図 5.1)。

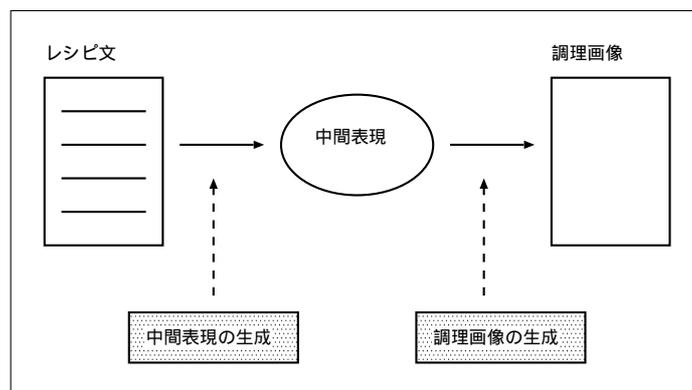


図 5.1: システムの処理の流れ

- 中間表現の生成

レシピ文分析から得られた、画像を生成するために必要と思われる情報を含んだ中間表現を生成する。画像化に必要な情報とは、3章で述べた「画像要素」と「調理動作」に関する情報である。

- 調理画像の生成

中間表現から特定される調理画像（キーフレームアニメーション）を呼びだし、正しい順序で再生を行なう。

5.2 中間表現

本節では、調理画像生成のために必要な中間表現について述べる。

料理レシピ中には有限数の調理手順が存在するが、必ずしも1つの調理手順に1つの動作というわけではなく、1つの調理手順中に複数の動作が行なわれる場合も多く見られる(図 5.2)。料理レシピ文を画像化するためには、すべての動作を特定し画像化しなければならない。

本研究において、調理手順よりさらに詳細な区切りを「ステップ」と定義する。

$$\text{調理手順 } 1 = \text{ステップ } 1, \dots, \text{ステップ } n \\ (1 \leq n)$$

図 5.2: 有限数のステップから構成される調理手順

これにより、中間表現は、料理レシピ中の全てのステップについて、調理画像を生成するための情報を保持したものでなければならない。

レシピ文の分析より、調理画像を生成するために必要な情報としてフレームを以下のように定義した。

- 「材料」
- 「調理動作」
- 「調理器具」

- 「調理場所」
- 「副詞句」

料理レシピ文における言語情報と、システムが保持する辞書等の知識を用いて、各ステップに関してこれらのフレームを埋めていく(図 5.3)。また、調理器具を使わない「洗う」など、動作によってはこれらの全ての特定が必ずしも必要ではないため、動作はそれぞれに固有の格フレームで動詞辞書において表現される。

```

STEP数 = X

[STEP1]
材料 = ~
調理動作 = ~
調理器具 = ~
調理場所 = ~
副詞句 = ~

[STEP2]
材料 =
調理動作 =
調理器具 =
調理場所 =
副詞句 =

.
.
.
.

[STEPX]
材料 =
調理動作 =
調理器具 =
調理場所 =
副詞句 =

```

図 5.3: 中間表現

5.3 レシピ文から中間表現への処理内容 (1)

中間表現を生成するために、以下の4点をシステムが特定する。

1. レシピ文中の画像要素
2. 補完すべき調理動作
3. 調理が進むにつれて変化する材料の形状

4. 調理が進むにつれて変化する材料の場所

上記の4点は、1. が「言語表現と知識の参照により得られる動作ステップの要素」であり、2.3.4. は、「ルールを用いて情報を補完し得られる動作ステップの要素」である。本節では、1. の特定について述べる。

調理で行なわれる全ての動作をステップとする中間表現を生成するため、入力となる料理レシピ文に対して形態素解析・構文解析を行ない、動詞で区切る。次に、区切られたそれぞれのステップについて、システムが保持する知識(図 5.3) を用いて画像要素である「材料」「調理器具」「調理場所」を特定していく。

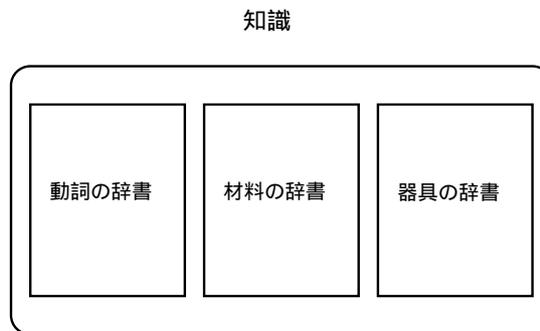


図 5.4: 知識：3つの辞書

まず、材料の辞書を参照して、レシピ文中で「材料」を特定する。「調理器具」と「調理場所」に関して、これらは動作に固有の情報として考えることができ、あらかじめ動作に対応した「調理器具」と「調理場所」の値を動詞辞書に持たせておく。ただし、「洗う」のように「調理器具」を必要としなかったり、「加える」のように「調理場所」をあらかじめ決めにくいような動作に対しては、これらの値は与えない。動作毎にこのような点を考慮しなければならない。

具体的に、動詞辞書には各動詞を先頭に、動作を行なう「場所」と動作の対象となる「材料」、必要な「器具」、そして「副詞句」という4つのパラメータが基本的に備わっている。以下は調理動作「切る」の辞書中の表現形式である(図 5.5)。調理場所はデフォルトとして「まな板」に、調理器具には「包丁」という値が入っている。

ただし、「調理器具」について、レシピ中に動詞辞書の器具とは異なる記述があった場合には、これを優先する。異なる「調理器具」を判断するために、材料の辞書と同様に調

切る_<場所>まな板</場所>_<材料></材料>_<器具>包丁</器具>_<切り方></切り方>

図 5.5: 動作辞書における「切る」の表現

理器具の辞書を用意する.

5.4 レシピ文から中間表現への処理内容 (2)

本節では, 5.2.1 節の 2.3.4. の特定について述べる.

5.4.1 全ての調理動作の特定

レシピ文中に出現する多くの調理動作は, 最小限の情報しか含まれていない. 正しい調理動作の画像化を行なうためには, 前提となる条件が満たされていなければならない. 例えば, 「材料を切る」場合に「材料がまな板の上にある」ということなど. しかし, レシピ中には, この「材料をまな板の上につす」などの動作は常識として省略されるため, これらの前提条件を満たすための動作 (サブ動作) の画像化も行なうこととする.

本研究では, 上述したような「調理動作」を行なうための「調理場所」が満たされているかに着目し, 満たされていない場合は新たに調理動作「移す」(図 5.6) を動作ステップの中に加えることにした.

移す_<場所>from A to B</場所>_<材料></材料>

図 5.6: 調理動作「(材料を場所 A から場所 B へ) 移す」の表現

具体的には, 料理レシピ中に,

1. 玉ねぎはざく切りにします

という手順が存在した場合, この動作を行なうために「玉ねぎ」は調理場所「まな板」になければならない. そこで現在, 「玉ねぎ」がどの場所にあるかをチェックする. 「まな板」の上にある場合は問題ないが, 「まな板」の上でない場合, 現在の場所から「まな板」へ「移す」動作を新たに動作ステップに加えなければならない.

上の例では、手順1の動作である。初期状態として「材料」は全て調理場所「皿」にあるから、「皿」から「まな板」へ移す動作が、「ざく切りにする」動作の前に入る。

5.4.2 材料の形状の変化

通常、調理が進むに従って材料の形状は変化していく。ある動作が行なわれた場合に、材料はどのように変化するのかといった知識が必要である。材料の形状変化に関しては、4.1節で示したように、以下の4つのパターンがある。

材料の形状変化

- 固形 固形
- 固形 液状
- 液状 固形
- 液状 液状

本研究では、画像化するうえでの難しさを考慮し、「固形 固形」の変化を扱うことにする。これは、多くの場合、材料が「内部構造」に分割される変化である。また、材料を「内部構造」に分割する調理動作は、「むく」「除く」などあるが、例えば「アスパラガスを切る」調理動作では、調理前と後において、レシピ文中での言語表現は同じ「アスパラガス」であるが、調理前の「アスパラガス」と調理後の「アスパラガス」では、その形状に関して大きく異なる。調理前では1本の「アスパラガス」であるが、調理動作「切る」を行なった後では、複数の断片に分割される。そして、その後の手順において、それぞれの断片が別々に調理される。このことから、画像化を考慮すると、調理前と調理後では同じアスパラでも別のオブジェクトとして扱う必要がある。それに対応して、中間表現を生成する際にも、材料を内部構造に分割する動作が行なわれた場合に、中間表現において別材料として扱う必要がある(図5.7)。

5.4.3 材料の場所

5.4.1節で述べた材料を「移す」動作を加えることに関連して、材料の形状変化と同様に、調理が進むに従って材料が現在置かれている場所をシステムは管理するようにしなけ

アスパラガス - - - - アスパラガス 1_アスパラガス 2
調理動作「切る」

図 5.7: 内部構造に分割する形状変化

ればならない。本研究では、材料と調味料がそれぞれ置かれている場所を記述したリストを用意する。

初期状態における材料と調味料の場所

- 材料
キッチンに置かれた皿の上。
- 調味料
キッチンのテーブルの上。

これらの場所は、調理ステップの順に調理動作によって随時書き換えられていく。

5.5 レシピ文から中間表現への処理内容 まとめ

本節では、5.3 節と 5.4 節で示した内容に基づいた、入力である料理レシピ文から中間表現を生成する処理の流れを示す。以下の手順で中間表現となるステップ毎の「材料」、「調理動作」、「調理器具」、「調理場所」、「副詞句」を特定していく。

中間表現の生成

1. レシピ文を形態素解析・構文解析した結果と動詞、材料、器具の各辞書から、レシピ文中の画像要素にタグを付与する
2. タグが付いたレシピ文を動作で区切り、対応する動詞辞書内の格フレームに変換する
3. レシピ中の「材料」「調理器具」「調理場所」タグが付与されている名詞でフレームのブランクを埋める
4. タグが付与されずに残った部分を「副詞句」としてフレームを埋める

5. 各動作のフレーム表現において、空白があった場合にそれ以前の動作の値を与える
6. 各動作ステップにおいて、調理手順の順に、「材料」リストとそれに対応した「場所」のリストを参照して、場所が変化している場合に「材料を移す動作」を動作ステップに加える
7. 各動作ステップの中に、材料を「内部構造に分ける動作」が存在した場合に、そのステップ以降で、「材料」を分割された表現に書き換える
8. 「材料」「調理動作」「調理器具」「調理場所」等について、全てのステップで特定し、完全な中間表現を生成する

5.6 調理画像

本節では、本システムにおける調理画像の扱いについて述べる。

調理画像生成に必要な画像要素（CGオブジェクト）を3DCG専用のソフトウェアを用いてモデリングを行なった。さらに、各画像要素に対してアニメーションを設定した。これらの調理画像の作成には以下の2つのソフトウェアを使用した。2つのソフトウェアの用途は以下の通りである。

MaYa 2.0 CGモデリングソフト。材料など画像要素のモデリング

Digital Loca 2.0 3Dオーサリングツール。調理動作・材料の形状変化を表現するアニメーションと、スクリプトによるアニメーションフレーム制御

5.6.1 画像要素のモデリングとアニメーションの設定

MaYa 2.0 を用いてモデリングした画像オブジェクト（調理画像要素）を、Digital Loca で読み込んで、それぞれにキーフレームアニメーションを設定した。

キーフレームアニメーションを作る際の節目のコマのこと。これを数フレームおきに正確に作成し、その間を補間して動画を作成する一般的なアニメーションの制作技法。

Maya 2.0 でもアニメーションは生成は可能だが、新たに画像要素が加えられた場合の汎用性が得られない。これは、Maya 2.0 のアニメーション方法が、全てのフレームに対し、2次元の画像ファイルを生成し、秒速数フレームずつに表示されるというものだからである。これにより、「材料Aを切る」画像と「材料Bを切る」画像をそれぞれ別に用意しなければならない(図 5.8)。材料が増えれば増えるほど画像データ量と作成の時間が増えてしまう。

これに対し、Digital Loca ではアニメーションを予め設定しておく、それぞれの画像をレイヤーのように重ね合わせ、再生することができる(図 5.9)。

5.6.2 画像要素の作用スクリプトと画像再生スクリプト

自然な調理画像を生成するため、調理画像に関して2種類のスクリプトを作成した。本節では、それぞれのスクリプトについて述べる。

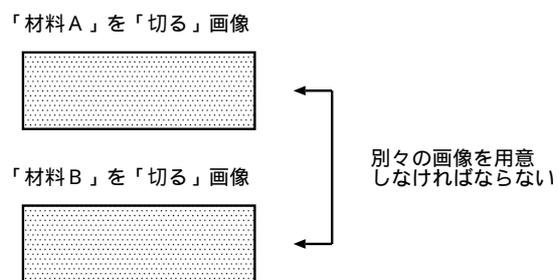


図 5.8: Maya 2.0 でアニメーションを生成する場合

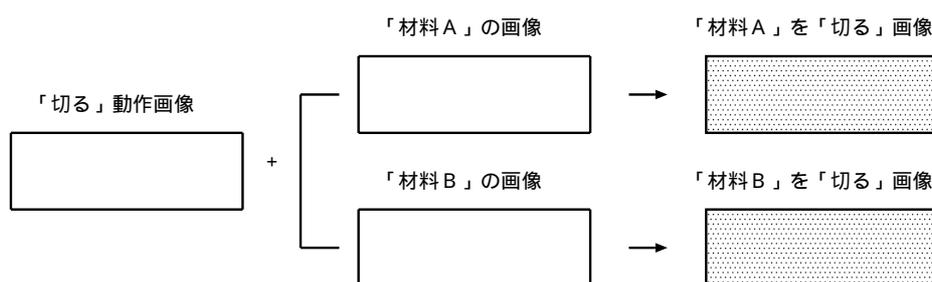


図 5.9: Digital Loca でアニメーションを生成する場合

画像要素の作用スクリプト

4.2 節において、材料が調理動作によって形状を変化させることから、画像要素間の相互の作用について考慮しなければならないことを述べた。また、5.4.2 節では材料が形状を変化させる場合、調理前のオブジェクトと調理後のオブジェクトは別のオブジェクトとして扱うことを述べた。そこで、画像要素間の作用に関してスクリプトを作成した。これにより、材料に形状を変化させるような動作が与えられた場合に、変化する過程を表現したアニメーションを再生させるようにした。例として、材料「ウインナ」を「切る」場合には、画像要素である材料「ウインナ」と動作「切る」が再生中に、調理器具「包丁」が「ウインナ」に接触したら、ウインナのデフォルトの画像から「切られる」アニメーション画像に切替えて、再生するようにした (図 5.10)。

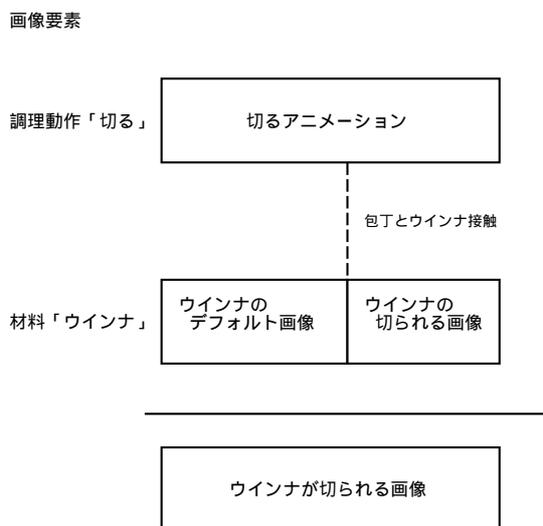


図 5.10: 画像要素の作用

画像再生スクリプト 各画像要素が記述された中間表現を入力として与えると、各画像要素の指定された画像フレームを再生していくスクリプト (図 5.11) を作成した。

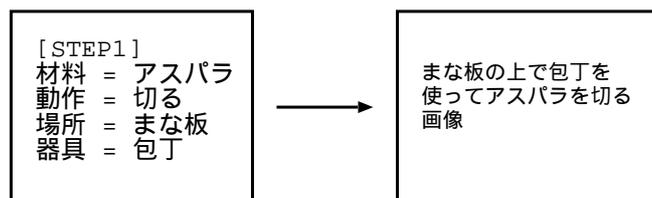


図 5.11: 「中間表現」 「画像生成」

第 6 章

レシピ文から画像を生成する動作例

用意できる画像要素には限りがあるため、本研究では実際に調理画像を生成する対象レシピを選定した。本章では、対象レシピ文についての説明と対象レシピを構築したシステムで処理を行なわせた時の動作に関して述べる。

6.1 画像化するレシピ文

画像化するレシピとして「アスパラガスとウインナのじか炒め」を選んだ。

作り方

「アスパラガスとウインナのじか炒め」

材料	
グリーンアスパラガス	2束
ウインナソーセージ	6本
バター	大さじ1
塩・こしょう	各少々

1. アスパラガスは3～4センチの長さに切ります。
2. ウインナは斜め切りにします。
3. フライパンにバターを溶かし、アスパラガスを中火で、焦がさないように炒めます。六分どおり火が通ったらウインナを加えてさっと炒め、塩、こしょうをふって、味をととのえます。

6.2 中間表現の生成

本節では、対象レシピの入力から画像生成に必要な情報を含んだ中間表現を生成するまでの処理の流れを示す。

6.2.1 レシピ文から読み取れる画像要素

中間表現を生成する処理の説明の前に、対象レシピ文について、手順毎に「材料」に対して行なわれる「調理動作」は以下に挙げる(表 6.1)。

これらは対象レシピ文から容易に読み取れる画像化すべき要素である。中間表現を生成するために、知識の参照や言語表現の分析により設定した規則を使って、さらに画像化に必要な画像要素を特定し、調理動作を具体化する。

表 6.1: 対象レシピにおける手順毎の「材料」と「動作」

レシピ手順	動作	材料
1	切る	アスパラ
2	切る	ウインナ
3	溶かす	バター
	炒める	アスパラ
	加える	ウインナ
	炒める	ウインナ
	ふる	塩・こしょう

6.2.2 中間表現を生成する処理の流れ

入力の対象レシピ文から得られる情報に加えて、以下に挙げた規則を設けて各動作ステップを具体化していく。

規則

- 場所に関する初期状態
 - － 「材料」は「皿」の上にある

- 「調味料」は「テーブル」の上にある
- 動作の前提条件
 - 「切る」動作を行なう時, 材料は「まな板」の上
 - 「炒める」動作を行なう時, 材料は「フライパン」の上
- 材料の形状変化
 - 「材料」は「切る」と内部構造に分かれる
- 場所の制約
 - 「まな板」の上に置ける「材料」は一つ
 - 「フライパン」の上に「材料」は複数置くことができる

レシピ文から中間表現への処理内容

1. 対象レシピ文を形態素解析・構文解析した結果と, 動詞, 材料, 器具の各辞書から, レシピ文中の画像要素にタグを付与する

手順 1 : アスパラガスは 3 ~ 4 センチの長さに切ります.

<材料>アスパラガス</材料>は 3 ~ 4 センチの長さに<動作>切り</動作>ます.

2. タグが付いたレシピ文を動作で区切り, 対応する動詞辞書内の格フレームに変換する

ステップ 1 : 切る_<場所>まな板</場所>_<材料></材料>_<器具>包丁</器具>
_<切り方></切り方>

3. レシピ文中の「材料」「調理器具」「調理場所」タグが付与されている名詞でフレームの値が入れられていないブランクを埋める

ステップ 1 : 切る_<場所>まな板</場所>_<材料>アスパラガス</材料>
<器具>包丁</器具><切り方></切り方>

4. タグが付与されずに残った部分を, 動作の具体化の手がかりとなる「副詞句」としてフレームのブランクを埋める

ステップ 1 : 切る_<場所>まな板</場所>_<材料>アスパラガス</材料>
<器具>包丁</器具><切り方>3 ~ 4 センチの長さ</切り方>

5. 各動作のフレーム表現において、ブランクがあった場合にそれ以前の値を入れる
6. 各動作ステップにおいて、調理手順の順に、「材料」リストとそれに対応した「場所」のリストを参照する (図 6.1). 場所が変化している場合には材料を「移す」動作を動作ステップに加える

ステップ 1 : 移す_<場所>from 皿 to まな板</場所>_<材料>アスパラガス</材料>

ステップ 2 : 切る_<場所>まな板</場所>_<材料>アスパラガス</材料>
<器具>包丁</器具><切り方>3 ~ 4 センチの長さ</切り方>

ステップ 3 : 移す_<場所>from まな板 to 皿</場所>
_<材料>アスパラガス</材料>

7. 各動作ステップの中に、材料を「内部構造に分ける動作」が存在した場合に、そのステップ以降、「材料」を分割された表現に書き換える

ステップ 2 : 切る_<場所>まな板</場所>_<材料>アスパラガス</材料>
<器具>包丁</器具><切り方>3 ~ 4 センチの長さ</切り方>

ステップ 3 : 移す_<場所>from まな板 to 皿</場所>
_<材料>アスパラガス 1_アスパラガス 2</材料>

8. 「材料」「調理動作」「調理器具」「調理場所」等について、全ての動作ステップで特定し、中間表現を生成する.

ステップ 1 : 移す_<場所>from 皿 to まな板</場所>_<材料>アスパラガス</材料>

ステップ 2 : 切る_<場所>まな板</場所>_<材料>アスパラガス</材料>
<器具>包丁</器具><切り方> 3 ~ 4 センチの長さ</切り方>

ステップ 3 : 移す_<場所>from まな板 to 皿</場所>

_<材料>アスパラガス 1_アスパラガス 2</材料>

ステップ 4 : 移す_<場所>from 皿 to まな板</場所>_<材料>ウインナ</材料>

ステップ 5 : 切る_<場所>まな板</場所>_<材料>ウインナ</材料>

<器具>包丁</器具><切り方> 斜め</切り方>

ステップ 6 : 移す_<場所>from 皿 to フライパン</場所>_<材料>バター</材料>

ステップ 7 : 溶かす_<場所>フライパン</場所>_<材料>バター</材料>

ステップ 8 : 移す_<場所>from 皿 to フライパン</場所>

_<材料>アスパラガス 1_アスパラガス 2</材料>

ステップ 9 : 炒める_<場所>フライパン</場所>_<材料>アスパラガス 1

アスパラガス 2</材料><方法>中火で, 焦がさない</方法>

ステップ 10 : 移す_<場所>from まな板 to フライパン</場所>

_<材料>ウインナ 1_ウインナ 2</材料>

ステップ 11 : 炒める_<場所>フライパン</場所>_<材料>ウインナ 1

ウインナ 2</材料><方法>さっと</方法>

ステップ 12 : ふる_<場所>フライパン</場所>_<材料>ウインナ 1

ウインナ 2</材料>(に)<調味料>塩</調味料>

ステップ 13 : ふる_<場所>フライパン</場所>_<材料>ウインナ 1

ウインナ 2</材料>(に)<調味料>こしょう</調味料>

料理手順	材料・調味料の場所
初期状態	(アス、ウイ、バター、塩、こしょう) = (皿、皿、皿、テーブル、テーブル)
1. (切る(アスパラ))	(アス、ウイ、バター、塩、こしょう) = (まな板、皿、皿、テーブル、テーブル)
	(アス1_アス2、ウイ、バター、塩、こしょう) = (まな板_まな板、皿、皿、テーブル、テーブル)
	(アス1_アス2、ウイ、バター、塩、こしょう) = (皿_皿、皿、皿、テーブル、テーブル)
2. (切る(ウインナ))	(アス1_アス2、ウイ、バター、塩、こしょう) = (皿_皿、まな板、皿、テーブル、テーブル)
3. (溶かす(バター))	(アス1_アス2、ウイ1_ウイ2、バター、塩、こしょう) = (皿_皿、まな板_まな板、フラ、テーブル、テーブル)
(炒める(アスパラ))	(アス1_アス2、ウイ1_ウイ2、バター、塩、こしょう) = (フラ_フラ、まな板_まな板、フラ、テーブル、テーブル)
(加える(ウインナ))	(アス1_アス2、ウイ1_ウイ2、バター、塩、こしょう) = (フラ_フラ、フラ_フラ、フラ、テーブル、テーブル)
(炒める(ウインナ))	(アス1_アス2、ウイ1_ウイ2、バター、塩、こしょう) = (フラ_フラ、フラ_フラ、フラ、テーブル、テーブル)
(ふる(塩))	(アス1_アス2、ウイ1_ウイ2、バター、塩、こしょう) = (フラ_フラ、フラ_フラ、フラ、 フラ、テーブル)
(ふる(こしょう))	(アス1_アス2、ウイ1_ウイ2、バター、塩、こしょう) = (フラ_フラ、フラ_フラ、フラ、テーブル、 フラ)

図 6.1: 材料・調味料の動作ステップ毎の場所の変化

6.2.3 レシピ文から生成された中間表現

最終的に生成される中間表現をまとめると以下の表のようになる。動作を具体化する手がかりとなる「副詞句」に関して、ここでは省略している。

表 6.2: 中間表現

ステップ	材料	動作	器具	場所
1	アスパラ	移す		皿 まな板
2	アスパラ	切る	包丁	まな板
3	アスパラ 1, アスパラ 2	移す		まな板 皿
4	ウイナ	移す		皿 まな板
5	ウイナ	切る	包丁	まな板
6	バター	移す		皿 フライパン
7	バター	溶かす		フライパン
8	アスパラ 1-アスパラ 2	移す		皿 フライパン
9	アスパラ 1-アスパラ 2	炒める	フライパン	フライパン
10	ウイナ 1-ウイナ 2	加える		まな板 フライパン
11	ウイナ 1-ウイナ 2	炒める	フライパン	フライパン
12	塩	ふる		フライパン
13	こしょう	ふる		フライパン

6.3 調理画像の生成

本章では、対象レシピ「アスパラガスとウインナのじか炒め」の画像化のために行なった必要な画像要素のモデリングと、アニメーションの生成に関するスクリプトの説明を行ない、最終的にシステムが出力した調理画像について述べる。

6.3.1 画像要素

対象レシピ「アスパラとウインナのじか炒め」画像化のために、「材料」「調理動作」「調理器具」等の画像要素とステージとなるキッチンなど、必要な全ての画像オブジェクトを、Maya2.0 を使用してモデリングを行なった。

画像オブジェクト

ステージ キッチン

両手 右手, 左手

材料 ウインナ, アスパラ, バター

器具 包丁, フライパン, まな板, 皿

調味料 塩(ケース), こしょう(ケース)

4章で述べたように、調理手順が進むにつれて調理を施された材料はその形状を変化させることが多い。この事を考慮して、材料については「内部構造」を持たせてモデリングを行なった(図 6.2)。

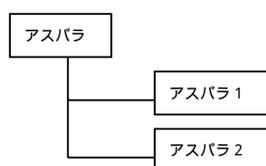


図 6.2: 材料オブジェクトの内部構造

6.3.2 アニメーション画像とスクリプト

それぞれの画像要素に対して、キーフレーム毎にアニメーションを作成した。また、画像要素間の作用について、材料の形状の変化に関するアニメーションを自動的に再生させる「作用スクリプト」を作成した。

画像要素間の作用

例えば「材料を切る」画像を再生すること考える。「材料」が「デフォルト」の状態から「切られる」画像に切り替わるタイミングを「包丁」と接触する時とし、より自然な調理画像を再生させる。この画像要素間の接触における画像の切り替え等をスクリプトとして、アニメーションフレームに保持させる(図 6.3)。

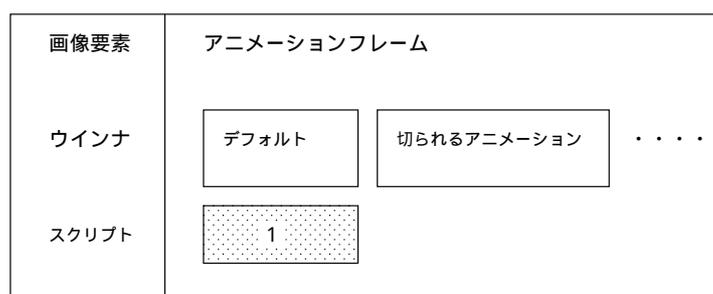


図 6.3: 作用スクリプトによる再生画像の切替え

作用スクリプト 1

「ウインナ」のデフォルトフレーム再生中、「ウインナ」が「包丁」と接触した場合に、「切られるアニメーションフレーム」の先頭へ移動

6.3.3 画像再生スクリプト

6.2.3 節で示した中間表現を入力として与え、予め用意したアニメーションフレーム群の中から、ステップ毎に必要なアニメーションフレームを特定し、ステップが終わるまで順番に再生していく「画像再生スクリプト」を作成した(図 6.4)。

```

begin
  if 再生する必要がある (STEP1 ~最後のSTEPまで全ステップについて) then
    中間表現の読み込み (画像要素の特定 (STEP毎))

    各画像要素の
    アニメーションフレームを再生

  else
    \\ 全て再生し終わった
  end;
end;

```

図 6.4: 画像再生スクリプト

6.3.4 レシピ「アスパラガスとウインナのじか炒め」の調理画像

本節では、構築したシステムに対象レシピ「アスパラガスとウインナのじか炒め」のレシピ文を入力として与えた時に、システムが出力した調理画像を示す。

- 初期画面 (図 6.5)



図 6.5: 初期画面

- 動作ステップ 2 「アスパラガスは 3 ~ 4 センチの長さに切ります」(図 6.10)

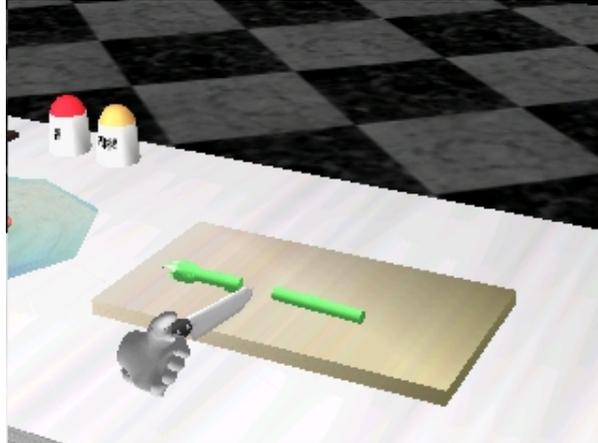


図 6.6: 動作ステップ 2 の調理画像

- 動作ステップ 1 1 「ウイナ 1-ウイナ 2 をさっと炒める」(図 6.7)



図 6.7: 動作ステップ 1 1 の調理画像

- 動作ステップ 1 2 「塩をふる」(図 6.8)

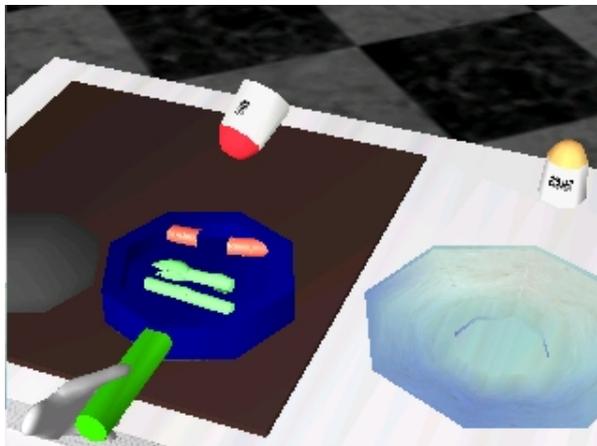


図 6.8: 動作ステップ 1 2 の調理画像

6.3.5 レシピ文を変えた場合の動作

本節では, 対象レシピ文を以下のように変えた時のシステムが出力した調理画像を示す.
調理手順

1. アスパラガスは半分に切って, 塩をふります.
2. フライパンにバターを溶かし, アスパラガスを 5 分間炒めます.
3. こしょうをふります.
4. 六分どおり火が通ったらウインナを加えてさっと炒めます.

中間表現

ステップ 1 : 移す_<場所>from 皿 to まな板</場所>_<材料>アスパラガス</材料>

ステップ 2 : 切る_<場所>まな板</場所>_<材料>アスパラガス</材料>
<器具>包丁</器具><切り方>半分</切り方>

ステップ 3 : ふる_<場所>まな板</場所>_<材料>アスパラガス 1
アスパラガス 2</材料>(に)<調味料>塩</調味料>

ステップ 4 : 移す_<場所>from 皿 to フライパン</場所>_<材料>バター</材料>

ステップ 5 : 溶かす_<場所>フライパン</場所>_<材料>バター</材料>

ステップ 6 : 移す_<場所>from まな板 to フライパン</場所>
_<材料>アスパラガス 1_アスパラガス 2</材料>

ステップ 7 : 炒める_<場所>フライパン</場所>_<材料>アスパラガス 1
アスパラガス 2</材料><方法>5 分間</方法>

ステップ 8 : ふる_<場所>フライパン</場所>_<材料>アスパラガス 1
アスパラガス 2</材料>(に)<調味料>こしょう</調味料>

ステップ 9 : 移す_<場所>from 皿 to フライパン</場所>_<材料>ウインナ</材料>

ステップ 10 : 炒める_<場所>フライパン</場所>_<材料>ウインナ</材料>
_<方法>さっと</方法>

調理画像

- 動作ステップ 3 「塩をふる」(図 6.9)

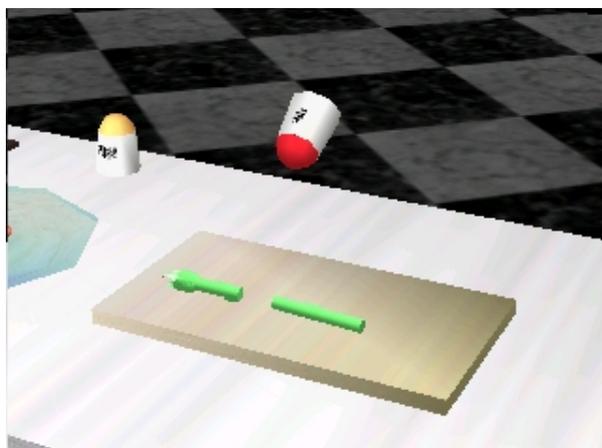


図 6.9: 動作ステップ 3 の調理画像

- 動作ステップ 10 「塩をふる」(図??)

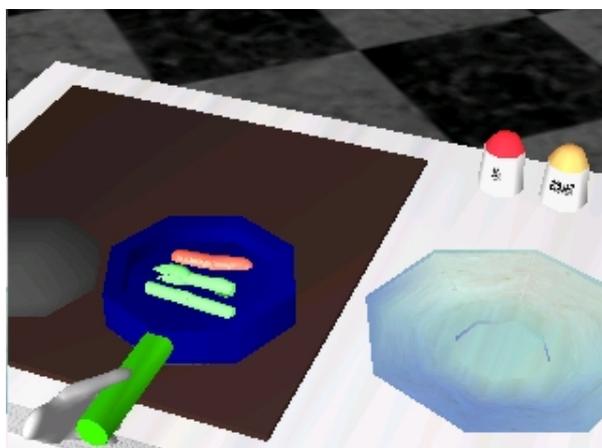


図 6.10: 動作ステップ 10 の調理画像

第 7 章

考察

本章では、作成したシステムが言語表現における「調理動作」の扱いと、「調理画像」について、分析結果をどの程度反映できたかについての考察を示す。

7.1 調理動作

レシピ文は複数の調理動作から成り立つ。レシピ中に頻出する動詞の連用形で接続した文「A, B」について、3.3.3 節で取り上げた 4 つのケースが含まれたレシピを対象レシピとして扱った。

1. A の画像 + B の画像
2. A の画像中に B のサブ画像を重ねる
3. A だけ画像化 (B は目的)
4. A と B の間に画像 C をはさむ

7.1.1 パターン 1 ・パターン 2

パターン 1 とパターン 2 の違いは動作 A の種類の違いである。

動作の種類について

動作はその終点に関して 2 通りある。

- 終点が特定できる動作

終点が特定できる動作に関しては、その動作により材料が変化し終わった時点（切る・移す・加える・溶かす）を終点とした。パターン1の動作Aは、この終点が特定できる動作である。

- 終点が特定できない動作

パターン2の場合、動作Aに含まれる動詞は単独では終点が特定できない。このような動詞には、期間を特定する手がかりとなる副詞句がついている。

- － 期間を特定する手がかりとなる副詞句がついている場合

- * 材料の形状の変化「きつね色になるまで～」
- * 明示的な時間「5分間～」
- * 程度「さっと～」

画像フレームの長さが可変ではないため、今回画像には反映できなかった。しかし、レシピ文中にこれらの副詞句が存在していれば、それが係る動作画像再生中に文字を表示させるようにした。

- － 期間を特定する手がかりとなる副詞句がついていない場合

また、期間が特定できない動作であっても、副詞句が伴っていない場合もある。多くの場合、料理をする人の経験的な知識に頼っている動作である（「冷ます」「洗う」など）。対象レシピで言えば、「塩、こしょう（調味料）をふる」という動作。この場合は「切る」動作と同じように、固定した長さの画像を用意しておく。

レシピ文は通常、手順通りの順序で書かれているが、中には時間関係が前後したものもある。

例: 「あらかじめ切っておいた～を、…」

のような表現が含まれたレシピは、本研究では対象外とした。

7.1.2 パターン3

「～(して), 香りをつける」「～(して), めめりを落す」など, 画像化する必要のない直前の動作の目的を表す部分は, 動詞辞書・材料表から画像化のための情報として必要ではないと判断する. レシピ文の処理の段階で, 中間表現には含まれないようにした.

7.1.3 パターン4

動作Aと動作Bを行なう場所によって判断した. 動作の場所が異なる場合は, 間に材料を移動する動作をはさむようにした.

7.2 調理画像

画像要素に関して, 対象レシピに出現する画像要素についてのみ予め用意した. 今後, 様々なレシピを扱うためには, アニメーションを設定した画像要素(材料・動作)の数を増やしていく必要がある.

7.2.1 材料

調理する「材料」の形状は様々であるが, グラフィック画像で表現する場合に, その形状を表現することが困難な場合がある. 「麺類」や「液状の材料」などがそれである. 本研究では, 「材料」に関して, 調理動作前と後での形状が「固形」であるものに限定した.

またこの事に関連して, 4.1節で述べたように「材料」は調理が進むにつれて, その形状を変化させていくことがある.

材料の形状変化

- 固形 固形
- 固形 液状
- 液状 固形
- 液状 液状

本研究において, 液状のオブジェクトについて画像化が困難なため, 材料が液状の場合, また動作によって材料が液状に変化する場合は対象外とした.

7.2.2 動作

本研究で使用したソフトウェアの機能の制限により、作成したアニメーションの画像フレームの長さを自在に変えることができなかった。そのため、同じ動作でも異なる副詞句がついたものについては、画像化に関してバリエーションを与えることが出来なかった。

しかし、分析の結果、終点を特定できない動作の多くは「炒める」「煮る」など、火を使うものであることが分かっている。この場合、例えば動作「炒める」の調理画像の長さは固定であるが、炒められた材料は場所が変わらない限りフライパンの上にある。表示上、これは材料に対して動作「炒める」が継続して行なわれていることを表現できる(図 7.1)。同様に、「煮る」「焼く」など火にかけることでおこなわれる終点を特定できない動作は、材料の置かれている場所の継続により表すことができると考えられる。



図 7.1: 終点が特定できない動作の表現

第 8 章

おわりに

8.1 まとめ

本研究では、言語表現から画像を生成する問題を取り上げ、「料理レシピ」を対象とした。レシピ文から調理画像を生成する際に、どのように言語表現を取り扱う必要があるかについて分析を行ない、調理画像生成システムの実装を行なった。

まず、料理レシピ文の分析から、レシピ文から「画像要素」と「調理動作」を特定しなければならないことが分かった。「画像要素」として、「材料」「調理動作」「調理器具」「調理場所」があり、これら全てが必ずしもレシピ中に記述されているわけではないため、あらかじめシステムに知識を保持させておく必要があることが分かった。

また、「調理動作」については、その動作をいつまで継続するかという点に関して、「終点を特定できる動作」と「終点を特定できない動作」の2つに分類ができ、「終点を特定できない動作」の多くは動作期間を特定する手がかりとなる副詞句が存在することが分かった。さらに、動作期間を表す副詞句には、「数値を含んだ具体的な副詞句」と「抽象的な程度副詞句」と「材料の形状表現によって表される副詞句」の3種類あることが分かった。次に、料理は有限数の調理動作のシーケンスであることから、レシピ文中に複数の調理動作が接続した場合について分析を行ない、画像化する上で4つのパターンが存在することが分かった。

以上の分析結果に基づいて、調理画像生成システムの実装を行なった。用意できるCG画像には限りがあるため、対象レシピを「アスパラガスとウインナのじか炒め」に設定し、画像要素をモデリングツールにより生成した。画像要素「材料」を生成する際に、材料は

調理が進むにつれて形状を変化させることを考慮して、内部構造をもたせてモデリングを行なった。また、画像要素間の作用による材料の形状を記述したスクリプトを作成し、自然な調理画像を表示するようにした。言語処理に関しては、料理レシピ文を形態素解析・構文解析を行なった結果と材料・器具の辞書により、レシピ文中から必要な画像要素を特定させるようにした。レシピ文から抽出できない情報に関しては、「調理器具」や「調理場所」の知識を保持した動詞辞書を利用して補完する処理をさせ、それぞれの動作ステップについて新たな動作を挿入、または必要のない動作を削除することで、最終的に画像化すべき動作ステップを記述した中間表現を生成するようにした。

生成された中間表現から必要な画像要素を呼びだし、適切な順序で再生を行なう画像再生スクリプトによって、「アスパラガスとウインナのじか炒め」の調理画像を生成することができた。

8.2 今後の課題

本研究では、料理レシピにおける言語表現の処理に関して、レシピ手順をさらに動作ステップに分割して実際に行なわれる動作を特定するようにした。通常、料理レシピの手順に書かれている通りの順序で調理を行なえば、料理を作ることができるが、中には「あらかじめ切っておいた～を、…」のように時間関係が前後したものもあるため、調理時系列の処理に関して更に拡張が必要である。また、「ひと煮立ちさせる」などの料理分野固有の表現に関しても扱えるようにする必要がある。

また、今後、より多くのレシピに対応するには、相応の画像要素をあらかじめ生成しておく必要がある。その際、画像要素「材料」に関して、本研究では取り扱わなかった麺類のような軟らかな材料や液状の材料のような、より複雑な形状のものについても生成しなければならない。このことは、本研究のようなリアルタイムで調理する様を画像で表現する必要がある領域において、非常に困難な問題であるが、今後のグラフィックス技術の進歩によってこの問題は解決することが期待され、より広範なジャンルの料理レシピを扱えるようになるだろう。

謝辞

本研究を進めるにあたり、多くの方々に御指導・御助言を賜りました。

心より感謝の意を表したいと思います。

終始熱心な御指導を賜りました島津 明教授に。

数多くの貴重な御意見を頂いた奥村 学助教授, 望月 源助手, 明星大学 大石亨助教授に。

本研究の節目節目に、貴重な御意見, 御助言を下さいました, 奥村研究室博士後期課程の難波 英嗣氏に。

さらに、常日頃より議論を重ね、研究に関して良きアドバイスを下さった島津・奥村研究室の学生および卒業生の皆様に。

最後に、3年にわたる JAIST での生活を支えてくれた家族, そして友人たちに。

植松 秀樹

2001年2月15日

参考文献

- [Eugenio 92] Barbara Di Eugenio. Understanding Natural Language Instruction: THE CASE OF PURPOSE CLAUSES. ACL Proceedings 30th Annual Meeting, pp.120-127, 1992.
- [Adachi 97] Adachi Hisahiro. A Generation Method of Cooking Definitions Based on Similarity between A Couple of Recipes. NLPRS'97, pp.135-140,1997.
- [Webber 90] Bonnie Lynn Webber, Barbara Di Eugenio. Free Adjuncts In Natural Language Instructions. COLING 90, pp395-400, 1990.
- [Karlin 88] Robin F.Karlin. Defining the Semantics of Verbal Modifiers in the Domain Cooking Tasks. Proceedings of the 26st Annual Meeting of ACL, pp.61-67, 1988.
- [Strassmann 94] Steve Strassmann. Semi-Autonomous Animated Actors. Proceeding of the 12th.National Conference on Artificial Intelligence, Vol.1, pp.128-134, 1994.
- [Loyall 93] A.Bryan Loyall, Joseph Bates. Real-time Control of Animated Broad Agents. Proceedings of the 15th Annual Conference of the Cognitive Science Society, pp.664-669, 1993.
- [Eugenio 他 92] Barbara Di Eugenio, Michael White. On the Interpretation of Natural Language Instructions. COLING 92, pp.1147-1151, 1992.
- [Waltz 81] David L.Waltz. Toward A Detailed Model of Processing for Language Describing the Physical World. IJCAI-1981, pp.1-6, 1981.

- [Cassell 他 94] Cassell, J. ANIMATED CONVERSATION: Rule-based Generation of Facial Expression, Gesture & Spoken Intonation for Multiple Conversation Agents. Proc.SIGGRAPH'94, pp.413-420, 1994.
- [Towns 他 98] Stuart G.Towns and Charles B.Callaway, James C.Lester. Generating Coordinated Natural Language and 3D Animations for Complex Spatial Explanations.AAAI-98/IAAI-98 Proceedings, pp.112-119, 1998.
- [Andre 96] Elisabeth Andre, Thomas Rist. Coping with Temporal Constraints in Multimedia Presentation Planning. Proceedings of the Thirteen National Conference on Artificial Intelligence, pp.142-147,1996.