

映像エフェクト辞書システムによる修飾語のイメージの視覚化

平井彰悟¹ 角薫¹

概要：本研究では、形容詞・擬音語・擬態語などの修飾語としての言葉のイメージを対象の物体である 3D オブジェクトに対して視覚的に映像エフェクトとして表現することのできる映像エフェクト辞書システムを構築した。この映像エフェクト辞書システムは言葉と映像エフェクトとを対象の物体に応じて紐づけた辞書システムである。同じ修飾語でも対象の物体の種類ごとに異なる意味があるため、映像エフェクトの選択の知的なアルゴリズムが必要となる。この辞書は形容詞・擬音語・擬態語などの修飾語だけでイメージを表すことが難しいイメージについてのコミュニケーション支援として利用できると考えている。被験者 30 名の印象評定により名詞や修飾語の分類付けを検討して辞書を拡張し、IBM Bluemix と Microsoft Azure の K-mean 法によるクラスタリングにより修飾語の分類を試みた。

Visualization of Modifiers via a Visual Effect Dictionary System

SHOGO HIRAI¹ KAORU SUMI¹

1. はじめに

本研究では、形容詞・擬音語・擬態語などの修飾語としての言葉のイメージを対象の物体である 3D オブジェクトに対して視覚的に映像エフェクトとして表現することのできる映像エフェクト辞書を構築した。この映像エフェクト辞書は言葉と映像エフェクトとを対象の物体に応じて紐づけた辞書である。同じ修飾語でも対象の物体の種類ごとに異なる意味があるため、映像エフェクトの選択の知的なアルゴリズムが必要となる。この辞書は形容詞・擬音語・擬態語などの修飾語だけでイメージを表すことが難しい言葉を可視化することができる。ここでグラフィックスとは、ものを 3D オブジェクト、形容詞や生成されたものの表面などをテクスチャや映像エフェクト、そしてアニメーションで可視化することであると定義する。本研究では、被験者を用いたアンケートによる印象評定を分析しそのアルゴリズムを検討する。

2. 関連研究

3D オブジェクトを生成するシステムでは、WordsEye がある[1]。WordsEye は、「Yellow elephant」などをテキスト入力すると「黄色い象」が 3D オブジェクトで表示されるシステムである。本研究では 3D オブジェクトに動作やエフェクトを付けることで相手が伝えたい情報をより詳しく表現する点異なる。そして、テキストではなく、音声認

識を利用して音声で入力することでリアルタイムなコミュニケーションに対応させる。

また、田中ら[2]は 3D オブジェクトによって音声情報をアニメーションで表現する Kairai を構築した。Kairai は 3D 空間上の視点や「ちょっと」「かなり」などの話し言葉を重視した自然言語理解システムである。そのため、Kairai では形容詞表現については「遠い」や「青い」などの距離や色彩について表現されている。本研究も同様に形容詞を対象としてさらに擬音語、擬態語を表現するが、生成されるオブジェクトに動作やエフェクトをつける点異なる。例えば、「りんご」という物体に対して、「ざらざらした」という表現を加えることでりんごが変化し、会話を支援すると考えられる。

文字入力することで 3D アニメが表示される Anime de Blog がある[3][4][5]。Anime de Blog は主語であるキャラクターの動作を動詞に対応させて 3D アニメを生成している。本研究では、形容詞および擬音語、擬態語を研究対象とし、動的なエフェクトを生成し、名詞の具体的な特徴を表現する点異なる。

このように、文章から画像、アニメーションを生成するシステムは存在しているが、それを実際のコミュニケーションに利用するものはほとんどない。そこで本研究では、状態を表現した言葉である形容詞、擬音語、擬態語に着目し、動的なエフェクトを付与する。それにより、名詞から生成されるオブジェクトの特徴をイメージによって伝えることでコミュニケーション支援を行うシステムを提案する。

¹ 公立はこだて未来大学

3. 映像エフェクト辞書システム

本システムでは、ユーザはシステムに音声情報を与える。その与えられた音声情報をシステムが映像に変換する。システムの流れは以下のようになっている。

- 音声認識
- 認識した音声情報を名詞と修飾語に分ける
- 名詞を 3D オブジェクト、修飾語を映像エフェクトに変換

システム上ではまず Unity で作成されたメイン画面がある(図 1)。

メイン画面上にある「音声」というボタンをユーザに押しってもらう。押されてから 3 秒間音声認識を行うことができる。音声認識には IBM Watson の Speech to Text を用いる。Speech to Text は認識した音声を即座にテキストに反映することができる。認識した音声情報をシステムが形態素解析 NMeCab を通して「修飾語」と「名詞」に分類する。分類されたそれぞれの品詞を IBM Watson の Natural Language Classifier の分類器にかける。それぞれ分類された情報をもとに「修飾語」は映像エフェクト、「名詞」は 3D オブジェクトに変換し、ユーザに表示することができる。

本システムでは、「絵で学ぶ擬音語・擬態語カード(著: 富川和代)」を参考に映像エフェクトを「光る」、「乱れる」、「湿気」、「動的」、「浮遊」、「音」、「匂い」の 7 種類に分類した。そのうち、動きのある映像エフェクトで表現可能なものとして、「光る」、「乱れる」、「湿気」、「動的」、「浮遊」を選択し、それぞれ映像エフェクトを作成した。映像エフェクト、3D オブジェクトの分類方法を説明する。映像エフェクトはある特定の 3D オブジェクトに対して、

1. 表面のテクスチャを変更すること
2. 周辺に動きのある映像エフェクトを付与すること
3. オブジェクト自体に動きを与えること

の 3 つで表現する。上記のうち 2 つの例を図 2、図 3 に示す。例えば、1 に対しては、「おいしい」「綺麗な」「汚れた」といった抽象的な言葉を表現する。2 は、「赤

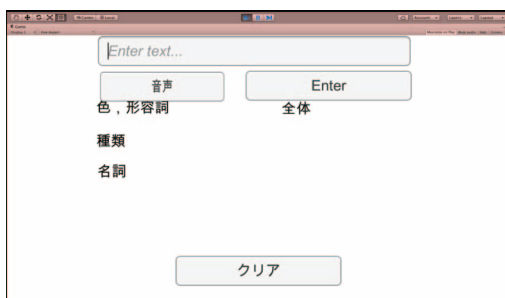


図 1: メイン画面

い」、「青い」などの色彩表現をする。3 では「速い」、「遅い」などの緩急のある表現に適応させるために 3D オブジェクトの座標を変更する。制作工程として、「光る」という映像エフェクトを例にあげる。まず光のテクスチャを用意し、それを点滅させるスピード、明度などを指定し、ループさせる。これにより、「光る」という分類を映像エフェクトで表現している。

3D オブジェクトは、「果物」、「温かい食べ物」、「冷たい食べ物」、「ボール」、「車」を用意した。分類された形容詞、擬音語、擬態語を関連付けられた映像エフェクトに変換する。本システムの開発環境並びに動作環境について、OS は Mac を使用し、ゲーム開発エンジンとして Unity 5 を使用し、開発言語として C# を使用した。また、機械学習を行う環境として IBM Watson を使用した。動的な映像エフェクトは Unity の Particle System で作成する。

4. 実験

4.1 実験の目的

本実験の目的は、映像エフェクトと 3D オブジェクトによって言葉のイメージを正しく想起できているか調査する。本研究では、映像エフェクトと 3D オブジェクトの想起を調査することで、映像エフェクトからわかる新しい分類方法を模索できると仮定した。

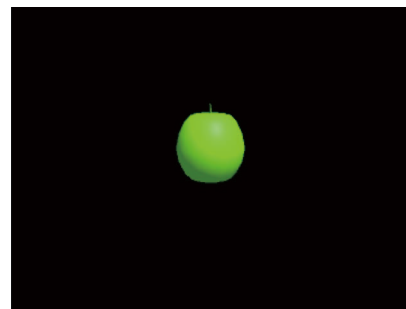


図 2: テクスチャを変更した映像エフェクト

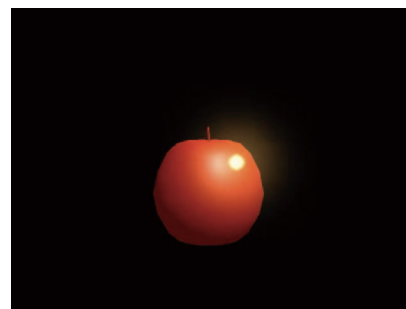


図 3: 周辺に動きのある映像エフェクト

表 2: 果物のアンケート内容

設問内容 1	設問内容 2	映像エフェクト	オブジェクト
新しい	おいしい	光る	りんご
			バナナ
			桃
新しい	おいしい	湯気	りんご
			バナナ
			桃
ぐらぐら	—	揺れる	りんご
速い	—	横移動	りんご
速い	—	動く線	りんご
まずい	汚い	黒い光	りんご
			バナナ
			桃
まずい	こげた	黒い湯気	りんご
			バナナ
			桃
ふわふわ	—	揺れる	りんご
青い	緑の	青色	りんご
			バナナ
			桃
青い	緑の	緑色	りんご
			バナナ
			桃
じめじめした	—	湿気	りんご
			バナナ
			桃

4.2 実験の方法

本実験では、2016年12月に18歳以上の男女30人を対象として行った。

被験者が3Dオブジェクトと映像エフェクトから修飾語を想起できるか調査するため「映像エフェクト+3Dオブジェクト」の組み合わせ動画から「修飾語+名詞」の文節を想起できるかを質問するアンケートを用意した。アンケートは紙ではなくGoogleフォームを用いて制作した。果物、食べ物、その他でアンケートを制作した。果物が動画25種類、質問43種類、食べ物が動画23種類、質問37種類、その他が動画28種類、質問42種類のアンケート構成になっている。アンケートの例として、果物のアンケートをまとめたものが表1である。それぞれの動画をみてもらい質問に答えてもらう。例えば、光っているエフェクトとりんごの3Dオブジェクトを組み合わせた動画をみってもらう。これに対して、動画は「おいしいりんご」を想起できるか、といった質問を提示する。被験者には「想起できる」、「やや想起できる」、「わからない」、「あまり想起できない」、「想起できない」の5択から適切なものを選んでもらう。

4.3 実験の結果

アンケートによると、被験者は男性が29人、女性が1人であった。また、被験者の平均年齢は21.5歳であった。アンケートで「光る」エフェクトと「湯気」エフェクトを反

表 1: 光る, 湯気エフェクトの結果

	光るエフェクト	湯気エフェクト
おいしいりんご	60.5%	16.7%
新しいりんご	60.0%	6.6%
おいしいバナナ	30.5%	24.1%
新しいバナナ	56.7%	16.6%
おいしい桃	36.6%	13.7%
新しい桃	56.6%	10.3%
おいしい温かい食べ物	60.0%	93.3%
新しい温かい食べ物	33.3%	66.7%
おいしい温かい食べ物 2	53.4%	100.0%
新しい温かい食べ物 2	40.0%	63.3%
おいしい冷たい食べ物	40.0%	23.3%
新しい冷たい食べ物	36.7%	26.6%
新しいボール	83.3%	3.3%
壊れたボール	0.0%	6.7%
新しい車	56.7%	3.3%
壊れた車	6.6%	86.7%
新しいバス	76.6%	3.3%
壊れたバス	6.7%	83.3%

映させた結果が表2である。「光る」エフェクトは、「光る」、「湯気」エフェクトが「湿気」を表している。各割合は被験者全体のうち想起できるとやや想起できるを選択した人数の割合である。例えば、「おいしいりんご」と「光る」エフェクトの組み合わせであれば、「光る」エフェクトと「りんご」の組み合わせが「おいしいりんご」を想起できるかという質問に対して60.50%の人が「想起できる」か「やや想起できる」を選択したことになる。

「光る」エフェクトでは、「おいしいりんご、新しいりんご」の想起できる割合がともに過半数を占めていた。また、バナナ、桃に関しても「新しい」を想起できる割合が過半数を占めていた。一方、温かい食べ物のオブジェクト2つは「おいしい」、「温かい」とともに「湯気」エフェクトで想起できる割合が過半数を占めていた。冷たい食べ物に関しては、どちらのエフェクトも「想起できない」の割合の方が多かった。ボールに対する印象は「光る」に対して「新しいボール」が想起できる割合が過半数を占めていた。乗り物である車、バスに関しては光るエフェクトで「新しい」を想起できる割合が過半数を占めていた。「壊れた」に対しては「湯気」エフェクトが過半数を占めていた。「黒い光」、「黒い湯気」エフェクトの結果が表3である。「黒い光」と「黒い湯気」エフェクトでは、3Dオブジェクトと提示した修飾語の全ての組み合わせで想起できると答えた割合が過半数を占めていた。揺れているエフェクトの結果が表4である。「ぐらぐら」という表現を提示した全ての3Dオブジェクトで想起できる結果となった。また、食べ物以外では「ぼろい」という修飾語もアンケートをとった。結果としてはどの3Dオブジェクトでも揺れるエフェクトでは「ぼろい」を想起できなかった。

表 3: 黒い光, 黒い湯気エフェクトの結果

	黒い光	黒い湯気
まずいりんご	86.6%	70.0%
汚いりんご	90.0%	/
こげたりんご	/	90.0%
まずいバナナ	46.6%	73.4%
汚いバナナ	56.7%	/
こげたバナナ	/	87.0%
まずい桃	83.3%	76.7%
汚い桃	96.7%	/
こげた桃	/	90.0%
まずい温かい食べ物	63.3%	90.0%
汚い温かい食べ物	76.7%	/
こげた温かい食べ物	/	90.0%
まずい温かい食べ物 2	73.3%	73.4%
汚い温かい食べ物 2	96.6%	/
こげた温かい食べ物 2	/	80.0%
まずい温かい食べ物	76.7%	70.0%
汚い冷たい食べ物	90.0%	/
こげた冷たい食べ物	/	33.3%
汚いボール	96.7%	36.6%
こげたボール	/	93.3%
汚い車	86.7%	50.0%
こげた車	/	90.0%
汚いバス	89.7%	53.3%
こげたバス	/	90.0%

表 4: 揺れるエフェクトの結果

	揺れるエフェクト
ぐらぐらしたりんご	76.7%
ぐらぐらした温かい食べ物	90.0%
ぐらぐらした冷たい食べ物	73.3%
ぐらぐらしたボール	100.0%
ぼろいボール	3.3%
ぐらぐらした車	96.7%
ぼろい車	40.0%
ぐらぐらしたバス	90.0%
ぼろいバス	36.6%

速いエフェクトの結果が表 5 である。3D オブジェクト自体に動きを与える方法では提示した全ての 3D オブジェクトで想起できる割合が過半数を占めていた。一方、周辺に動きのあるエフェクトを付与する方法でとった動く線のエフェクトでは温かい食べ物に対して「速い」と想起できない割合の方が多かった。浮いているエフェクトの結果が表 6 である。「ふわふわ」と想起できる割合が全ての 3D オブジェクトで過半数を占めていた。3D オブジェクトのテクスチャを変化させたエフェクトの結果が表 7 である。この質問はりんごの時のみ「緑色」、「青色」どちらを質問しても想起できると答えた割合が過半数を占めていた。その他は、青色は青色、緑色は緑色という想定内の回答であった。りんごという丸い性質を持つボールに関しても同様の質問をしたが、ボールも青色は青色、緑色は緑色に見えるという結果だった。

表 5: 速いエフェクトの結果

	横移動	動く線
りんご	83.3%	53.4%
温かい食べ物	73.4%	40.0%
ボール	90.0%	63.3%
乗り物	96.6%	90.0%
バス	96.7%	86.2%

表 6: 浮いているエフェクト

	浮いているエフェクト
りんご	83.4%
暖かい食べ物	80.0%
冷たい食べ物	76.7%
ボール	100.0%
乗り物	83.3%

表 7: 表面のテクスチャを変化させているエフェクトの結果

	緑色	青色
青いりんご	66.7%	86.7%
緑りんご	90.0%	26.6%
青いバナナ	43.3%	90.0%
緑バナナ	96.7%	26.6%
青桃	20.0%	80.0%
緑桃	76.6%	31.0%
青色冷たいもの	16.7%	93.3%
緑色冷たいもの	87.0%	10.0%
青色ボール	100.0%	6.6%
緑ボール	13.3%	100.0%

4.4 Microsoft Azure を用いた機械学習

この分析では、アンケート結果で用いた映像エフェクトによって結果にオブジェクトによって差がでた「光」と「湯気」を伴うオブジェクトにクラスタが存在するか調査する。今回の実験で集計したアンケート結果を Microsoft Azure によってクラスタリング分類を行った。Microsoft Azure とは、機械学習を活用したデータ分析ができるクラウドの予測分析サービスである。

図 4 が今回用いた Microsoft Azure のクラスタリングのモジュール構成である。「K-Means Clustering」がクラスタリングのアルゴリズム、「Train Clustering Model」が学習モデル、「Assign Data to Clusters」がモデルの適用を表すモジュールである。「K-Means Clustering」モジュールでは「K-Means++」

表 8 : 「2017 新しい」 のデータの例

オブジェクト	オブジェクトの常識	エフェクトの結果	正答
バナナ	果物	0.567	1
おかゆ	温かい食べ物	0.333	0

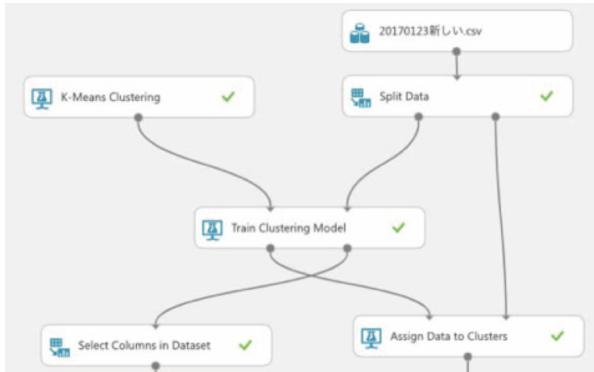


図 4 : Microsoft Azure を用いたクラスタリングのモジュール構成

という手法を用いた。k-means 法とは、非階層型クラスタリングのアルゴリズムである。クラスタの平均を用い、与えられたクラスタ数 k 個に分類する。K-means++では初期値の選択に改良をおこなったものである。クラスタ数は任意で選択できる。システムで分類した 3D オブジェクトが 5 種類、映像エフェクトが 5 種類であるため、今回はクラスタ数も 5 個とした。他の映像エフェクトに比べて、「光」と「湯気」を用いたときに印象が異なっていたため、「新しい」に関するアンケート結果について調査を行った。図 4 の「20170123 新しい」が「新しい」という修飾語に関するアンケート結果である。表 8 は、「20170123 新しい」のデータの例である。この例は、「光る」エフェクトのときに修飾語「新しい」を想起できるかという質問をした時のデータである。「2017 新しい」には、アンケートで用いた「映像エフェクト+オブジェクト」の組み合わせ動画と、その動画を見た際に質問をした修飾語のデータが含まれている。また、その組み合わせ動画を見て修飾語が想起できる割合が含まれている。正答はエフェクトの結果が 0.5 以上の時 1 を、0.5 未満の時 0 としたものである。図 4 のモジュール構成では学習データと評価データに分けているがクラスタリングに明確な精度を図る必要はないため、クラスタの分類だけを見ることにした。なお、表上にはオブジェクトの常識という行が存在するがクラスタの目安として表記したもので、入力データからは除外している。表 8 のデータをクラスタリングした結果が表 9 である。Assignments が分類されたクラスタを示す。

表 9 より、それぞれのクラスタの特徴として、

- クラスタ 0: 「新しい」 + 「湯気」 + 「温かい食べ物」
- クラスタ 1: 「新しい」 + 「光る」 + 「冷たい食べ物」以

表 9: Microsoft Azure のクラスタリング分類の結果

エフェクト	オブジェクト	オブジェクトの常識	Assignments
光る	ボール	丸いもの	1
光る	りんご	果物	1
湯気	バナナ	果物	3
湯気	小籠包	暖かい食べ物	0
湯気	桃	果物	3
光る	桃	果物	1
光る	アイスクリーム	冷たい食べ物	2
湯気	車	乗り物	4
湯気	りんご	果物	4
湯気	おかゆ	暖かい食べ物	0
光る	小籠包	暖かい食べ物	1
光る	バス	乗り物	1
光る	バナナ	果物	1
光る	車	乗り物	1
湯気	バス	乗り物	4
光る	おかゆ	暖かい食べ物	1
湯気	アイスクリーム	冷たい食べ物	4

外の「光る」を用いた群

- クラスタ 2: 「新しい」 + 「光る」 + 「冷たい食べ物」
- クラスタ 3: 「新しい」 + 「湯気」 + 「りんご以外の果物」
- クラスタ 4: 「新しい」 + 「湯気」 + 「上記のその他」という群が得られた。このクラスタからわかる点として、光る映像エフェクトより、湯気の映像エフェクトを用いた動画の方が、入力データにオブジェクトの常識を反映させなくても似たようなグループを形成することがわかる。このグループ例えば、クラスタ 0 においては温かい食べ物として分類される「おかゆ」、「小籠包」のみで群を作っている。また、クラスタ 3 においても「りんご」を除いた果物で群を構成している。クラスタ 4 に関しても「りんご」以外が乗り物という特徴で分類されていることがわかる。つまり、「りんご」を除いて修飾語とオブジェクト、アンケート結果によってオブジェクトの常識で分類することができることがわかった。

4.5 オブジェクトの常識・知識の情報を加えた重回帰分析

「新しい」という修飾語をオブジェクトの常識から調査した。オブジェクトの常識は Wordnet という意味辞書を用いて各名詞を検索して出てきた常識と一般的な項目を挙げた。一般的なオブジェクトの常識に当てはまるものを 1、当てはまらなければ 0 とした。オブジェクトの常識をまとめたものが表 10 である。表 10 と修飾語「新しい」、映像エフェクト「光」の組み合わせのアンケート結果を用いて重回帰分析を行った。表 11 が各項目の標準化係数、有意水準を示したものである。

変数減少法を用いたため、検定の際、不必要なクラスタは除去された。除去して選ばれたモデルがモデル 3 である。

表 10 : オブジェクトの常識

	冷たいもの	温かいもの	果物	丸いもの	人工物	おもちゃ	乗り物	食べ物	結果
りんご	0	0	1	1	0	0	0	1	0.605
車	0	0	0	0	1	0	1	0	0.567
バス	0	0	0	0	1	0	1	0	0.766
ボール	0	0	0	1	1	1	0	0	0.833
桃	0	0	1	1	0	0	0	1	0.566
バナナ	0	0	1	0	0	0	0	1	0.567
おかゆ	0	1	0	0	0	0	0	1	0.333
小籠包	0	1	0	0	0	0	0	1	0.4
アイスクリーム	1	0	0	0	0	0	0	1	0.367

表 11 : 「新しい」, 「光」の標準化係数

モデル		係数 ^a						
		非標準化係数		標準化係数		有意確率	共線性の統計量	
		B	標準誤差	ベータ	t 値			許容度
1	(定数)	.667	.061		10.854	.002		
	冷たいもの	-.200	.123	-.391	-1.629	.202	.563	1.778
	温かいもの	-.201	.106	-.518	-1.885	.156	.429	2.333
	丸いもの	.016	.106	.047	.150	.890	.333	3.000
	おもちゃ	.151	.150	.294	1.001	.391	.375	2.667
	食べ物	-.100	.106	-.292	-.936	.419	.333	3.000
2	(定数)	.667	.053		12.486	.000		
	冷たいもの	-.211	.087	-.412	-2.417	.073	.844	1.185
	温かいもの	-.211	.069	-.546	-3.064	.037	.771	1.296
	おもちゃ	.167	.092	.325	1.801	.146	.750	1.333
	食べ物	-.089	.069	-.260	-1.289	.267	.600	1.667
3	(定数)	.613	.036		17.069	.000		
	冷たいもの	-.246	.088	-.481	-2.798	.038	.938	1.067
	温かいもの	-.247	.067	-.638	-3.671	.014	.918	1.089
	おもちゃ	.220	.088	.429	2.498	.055	.938	1.067

a. 従属変数 結果

冷たいものが 0.038 で有意確率 5%以下で妥当なことがわかる。また、温かいものについても 0.014 で有意確率 5%以下で有意なことがわかる。しかし、冷たいものと温かいものの標準化係数はそれぞれ -0.481, -0.638 で負の相関を示している。つまり、「新しい」という修飾語において本研究で用いている 3D オブジェクトと映像エフェクトの組み合わせで適切な組み合わせが存在しないことがわかった。

同様に、「新しい」と「湯気」の場合も調査を行った。結果として、決定係数(R²乗)は 0.989 で 10%水準で有意であることがわかる。回帰式全体の有意性は 0.00 で 1%水準で有意であった。表 12 が各項目の標準化係数, 有意水準を示したものである。

各項目の標準化係数は温かい食べ物.000 で有意確率 0.01%水準で有意, 冷たいものが 0.011, 食べ物が 0.034 で有意確率 0.1%水準で有意なことがわかった。この中で、温かいものの標準化係数が 0.932 で正の相関でさらに大きな影響

を与えていることがわかる。つまり、「新しい」という修飾語を表現することに、「湯気」エフェクトと「温かい食べ物」の組み合わせが適切なことがわかった。以下が「新しい」の「光」と「湯気」の映像エフェクトに関するまとめである。

- 光の映像エフェクトは冷たいものと温かいもので負の相関を示した
- 湯気の映像エフェクトに適切なオブジェクトの常識として温かいもの, 冷たいもの, 食べ物で正の相関を示した。その中でも温かいものは特に関係性が強いことがわかった。

また、異なる修飾語として同じ映像エフェクトを用いた「おいしい」にも着目し重回帰分析を行った。「おいしい」と「光」の組み合わせで変数減少法によって重回帰分析を行った結果全てのオブジェクトの常識が除外されてしまい、有意確率を算出できなかった。つまり、重回帰分析をする上で適

表 12 : 「新しい」, 「湯気」の標準化係数

モデル		係数 ^a					共線性の統計量	
		非標準化係数 B	標準誤差	標準化係数 ベータ	t 値	有意確率	許容度	VIF
1	(定数)	.033	.015		2.275	.107		
	冷たいもの	.100	.029	.131	3.447	.041	.563	1.778
	温かいもの	.484	.025	.838	19.264	.000	.429	2.333
	丸いもの	-.082	.025	-.160	-3.244	.048	.333	3.000
	おもちゃ	.079	.036	.103	2.209	.114	.375	2.667
	食べ物	.133	.025	.261	5.294	.013	.333	3.000
2	(定数)	.046	.019		2.479	.068		
	冷たいもの	.126	.037	.165	3.394	.027	.675	1.481
	温かいもの	.510	.031	.883	16.403	.000	.551	1.815
	丸いもの	-.042	.025	-.083	-1.694	.165	.667	1.500
	食べ物	.094	.025	.184	3.759	.020	.667	1.500
	3	(定数)	.032	.019		1.642	.162	
冷たいもの	.154	.039	.202	3.960	.011	.844	1.185	
温かいもの	.538	.031	.932	17.471	.000	.771	1.296	
食べ物	.080	.028	.156	2.891	.034	.750	1.333	

a. 従属変数 結果

表 13 : 「おいしい」, 「湯気」の標準化係数

モデル		係数 ^a					共線性の統計量	
		非標準化係数 B	標準誤差	標準化係数 ベータ	t 値	有意確率	許容度	VIF
1	(定数)	.241	.037		6.566	.022		
	冷たいもの	-.008	.052	-.008	-.154	.892	.600	1.667
	温かいもの	.726	.045	.934	16.139	.004	.500	2.000
	丸いもの	-.089	.045	-.115	-1.980	.186	.500	2.000
2	(定数)	.237	.021		11.118	.002		
	温かいもの	.730	.030	.939	24.198	.000	.750	1.333
	丸いもの	-.085	.030	-.109	-2.820	.067	.750	1.333

a. 従属変数 結果

切なデータとは言えず、関係性が見られなかった。最後に「おいしい」と「湯気」の組み合わせをみた。結果として、決定係数(R²乗)は0.997で10%水準で有意であることがわかった。回帰式全体の有意性は0.00で1%水準で有意であった。表13が各項目の標準化係数、有意水準を示したものである。

各項目の標準化係数は温かい食べ物.000で有意確率0.01%水準で有意、標準化係数も0.939で正の相関を示した。つまり、「おいしい」という修飾語を表現する上で、湯気と温かいものの組み合わせが適切であることがわかった。

4.6 分析まとめ

上記の重回帰分析の結果をまとめたものが図5である。「新しい」という修飾語において湯気の映像エフェクトと温かい食べ物の組み合わせが最適であることが示唆された。また、光の映像エフェクトにおいて温かい食べ物と冷たい食

べ物が不適切であることが示唆された。これにより、冷たい食べ物や果物においては他の異なる映像エフェクトを用いることが必要であることがわかった。クラスタリングでは、「映像エフェクト」、「オブジェクト」、「想起できる割合」などを入力することで、人を介さずにオブジェクトをさらに細分化したクラスタを構成していた。

5. 考察

本研究では、クラスタリングと重回帰分析を行うことで映像エフェクトと修飾語、オブジェクトの関係を調査した。その結果、クラスタとして、「新しい」+「湯気」+「温かい食べ物」の群、「新しい」+「光る」+「冷たい食べ物」以外の「光る」エフェクトを用いた群、「新しい」+「光る」+「冷たい食べ物」の群、「新しい」+「湯気」+「りんご以外の果物」の群、「新しい」+「湯気」+「上記その他」という群が



図 5：重回帰分析のまとめ

得られた。この群を構成するのに用いた入力情報は、「修飾語」、「エフェクト」、「オブジェクト」、「修飾語を想起できる割合」である。すなわち、事前に映像エフェクトを用いて修飾語を想起できる割合を入力することで、同じ修飾語であってもシステムが映像エフェクトとオブジェクトを分類できることが示唆された。さらにクラスタリングでは、オブジェクトの常識・知識を入力に用いていない。にもかかわらず、結果は、オブジェクトを「温かい食べ物」や「冷たい食べ物」といったオブジェクトの常識・知識に細分化されていた。すなわち、人が入力せずに、自動的にシステムでオブジェクトの細分化を行うことができる可能性が示唆された。

次にオブジェクトの細分化を行うことが可能になった際にそのオブジェクトの常識から映像エフェクトと修飾語の関係を明確にできるか調査を行った。分析として重回帰分析を用いた。重回帰分析では、「温かい食べ物+湯気」の組み合わせが「新しい」という形容詞を想起させやすいことが示唆された。つまり、温かい食べ物に対して、湯気エフェクトを付与することで新しいイメージが湧きやすかった。また、「おいしい」という修飾語に対しても「温かい食べ物+湯気」の組み合わせが適切なことがわかった。温かい食べ物に対して、湯気のエフェクトを付与することで、美味しそうに見えることがわかった。

今回の分析では、光の映像エフェクトでは負の相関、つまり不適切なものしか抽出することができなかった。しかし、映像エフェクトと 3D オブジェクトには適切または不適切な表現が存在し、アンケートをとることでオブジェクトの常識レベルで修飾語を可視化することができた。この関係をより詳細に調べていくことで、修飾語単体では複数の意味を持ち可視化が難しい問題を、修飾される名詞の常識・知識から意味を限定し、可視化することが可能になると推測される。

このシステムでは課題点が幾つか浮上した。まず、複数の意味を持つ修飾語に対して修飾語と名詞の 1 つ 1 つのクラス分類だけでは名詞に対して適切な修飾語の映像エフェクトを反映させることが困難だった。2 つ目は、アンケート結果では、果物に分類していたものでも各果物によって想定できるものと出来ないものがあった。3 つ目は WordNet

によって増やした修飾語のテストデータの偏りによって、テストデータの多い分類に振り分けられる割合が高くなった点である。これらの問題のうち複数の意味を持つ修飾語に対してどのようにオブジェクトと修飾語を関連付けるか調査を行った。

アンケート結果では、「りんご」と「バナナ」と「桃」に「光る」エフェクトを用いて、「おいしい」を想起できるかについては「りんご」については過半数が想起できると回答した。しかし、「バナナ」と「桃」については過半数に満たなかった。一方で「新しい」を想起できるかについては、「りんご」、「桃」、「バナナ」で全て過半数を占めていた。しかし、「温かい食べ物」、「冷たい食べ物」に「光る」エフェクトを用いて、「新しい」を想起できるかについては、過半数を占めていたものはなかった。よって今回の実験であげた事例に関しては、食べ物において果物にのみ光るエフェクトが新しいという形容詞を表現していること示唆された。また、「湯気」エフェクトと「りんご」、「桃」、「バナナ」に関しては、「おいしい」も「新しい」も想起できない結果となった。「温かい食べ物」に「湯気」エフェクトを用いて、「おいしい」、「新しい」を想起できるかについては、全ての組み合わせで過半数を占めていた。つまり、「湯気」エフェクトが「温かい食べ物」において、「新しい」、「おいしい」の修飾語を表現することに適していることが示唆された。「揺れる」エフェクト、「ふわふわな」エフェクトについてはどのオブジェクトに対しても「ぐらぐらした」、「ふわふわ」を表現可能なことが示唆された。速い映像エフェクトに関しては、全体的に想起できる結果になっているが、実際に 3D オブジェクトを動かして速さを表現した動画の方が 3D オブジェクトの周辺に線を動かした動画より想起できる割合が多い結果となった。「まずい」や「こげた」、「汚い」といった用途に用いた映像エフェクトはおいしいと新しいで用いた映像エフェクトと真逆の色を採用してアンケートを行ったが「光」や「湯気」の映像エフェクトと違い、どの 3D オブジェクトに対しても想起できる割合が過半数を占める結果となった。「冷たい食べ物」に関しては想起できるものが少ないため、異なる映像エフェクトを用意する必要がある可能性が示唆された。

アンケート結果、重回帰分析などの分析は、実際に準備したオブジェクトのみ、映像エフェクトと修飾語の関連を見ることができなかった。しかし、オブジェクトが変わることで同じ映像エフェクトでも想起する修飾語が変わることがわかった。そのため、オブジェクトの常識や映像エフェクトの想起できる割合などのデータをより多く回収できるシステムにすることで、抽象的な修飾語を明確化する可能性を検討していきたい。

今後の展望としては、外国人が日本語の修飾語を学ぶ際に難題となる繊細な日本語表現をイメージで学ぶようなシステムとしても利用できると考えられる。音声同時通訳は

発展形だが、本システムのような視覚化で分かりやすく伝えられるシステムは有用であると言える。

6. 結言

本研究では、3D オブジェクトに対応した形容詞、擬音語、擬態語の映像エフェクトを返す辞書の構築を紹介した。結果、オブジェクトの常識から調査することで、抽象的な表現である修飾語に対して適切な映像エフェクトを与えることができることがわかった。

参考文献

- [1] Bob Coyne and Richard Sproat: WordsEye: An Automatic Text-to-Scene Conversion System, SIGGRAPH 2001, Proceedings of the 28th Annual Conference on Computer Graphics, Los Angeles, California, USA. ACM, 2001.
- [2] Hozumi Tanaka, Takenobu Tokunaga, and Shinyama Yusuke: Animated Agents Capable of Understanding Natural Language and Performing Actions, Life-Like Characters, Helmet Prendinger and Mitsuru Ishizuka Eds, pp.163-187, Springer, 2004.
- [3] Kaoru Sumi: Anime de Blog: Animation CGM for Content Distribution, Proc. of International Conference on Advances in Computer Entertainment Technology (ACE2008), pp.187-190, SIGCHI ACM, (2008.11)
- [4] Kaoru Sumi: Animation-based Interactive Storytelling System, published in Ulrike Spierling and Nicolas Szilas ed., Interactive Storytelling, LNCS 5334, Springer Lecture Note in Computer Science, pp.48-50, Springer (2008.11).
- [5] Kaoru Sumi: Capturing Common Sense Knowledge via Story Generation, Common Sense and Intelligent User Interfaces 2009: Story Understanding and Generation for Context-Aware Interface Design, 2009 International Conference on Intelligent User Interfaces (IUI2009), SIGCHI ACM, (2009.2).