

修士論文

Game with a Purposeによる映像エフェクト辞書
構築手法の研究

公立はこだて未来大学大学院 システム情報科学研究科
情報アーキテクチャ領域

平井 彰悟

指導教員 角 康之 教授, 角 薫 教授
提出日 2019年2月18日

Master's Thesis

**Study on Visual Effect Dictionary Construction Method
Using Game with a Purpose**

by

Shogo Hirai

MSc Thesis at Future University Hakodate

Supervisor Prof. Yasuyuki SUMI, Prof. Kaoru SUMI

Media Architecture Field
Future University Hakodate
Submitted on February 18, 2019

Abstract– There are various images in the modifier depending on the target noun. However, due to its versatile image, the image that I wants to convey and the image you thinks can be different. therefore, the contents of conversation may mistake each other. In this research, we developed a system that acquires structured data for visualizing modifiers, and registers the obtained structured data as a visual effect dictionary. So, We tackled the visualization of modifiers with different images depending on the object.

In this research, we propose a visual effect dictionary system corresponding to modifiers such as adjectives and onomatopoeia. Undergraduate research suggests that the modifier is recalled by the visual effect and common sense knowledge of the target object. Therefore, we aimed to express modifiers with multiple images visually by combining common sense knowledge and visual effect. In the proposed system, it is possible to acquire the structured data of the visual effect matching the image of the object for a specific modifier using Game With a Purpose (GWAP). "Emoji" is used for the object, the image of the modifier is displayed with the visual effect, and the game is developed in which the user gains data while enjoying the game by user's vote. The system can learn the relationship between visual effect and common sense knowledge, common knowledge and objects by the user. As a result, each time the game is played, the system dynamically changes the contents of the question and the visual effect to be presented based on the user's answer.

In this paper, We investigate the connection between visual effects and common sense knowledge of objects. Through this experiment, we were able to acquire the visual effect and common sense knowledge, the connection between common sense knowledge and objects as structured data. However, because bias was seen such that one visual effect and multiple common sense knowledge were linked, we also investigated the change when increasing the amount of data. As a result, as the number of times of game play increases, the score increases with respect to the proper connection, and the linkage between multiple visual effects and common sense knowledge was observed. Finally, when adapting the proposed system to modifiers other than "delicious", we investigated what kind of structured data can be obtained. In the network diagram of the same modifier, a combination tied independently to the visual effect was observed by common knowledge knowledge. Finally, we introduce examples of combining the proposed system with the data when applied to modifiers.

Keywords: Adjective, Noun, Commonsense knowledge, Visual effect, Game with a Purpose, Emoji

概要: 修飾語には対象の名詞によって様々なイメージが存在する。しかし、その思っているイメージと相手が考えるイメージが違い、会話がすれ違うことがある。本研究では、修飾語を視覚化するための構造化データを獲得するシステムを開発し、得られた構造化データを映像エフェクト辞書として登録することで対象のオブジェクトによって異なるイメージを持つ修飾語の視覚化に取り組んだ。

修飾語は同じ言葉であっても複数のイメージを持つため視覚化が難しい。本研究では、形容詞やオノマトペなどの修飾語に対応する映像エフェクト辞書システムを提案する。学部の研究により、修飾語は映像エフェクトと対象のオブジェクトの常識知識によって想起されることが示唆された。従って、複数のイメージを持つ修飾語を常識知識と映像エフェクトを組み合わせることで視覚的に表現することを目的とした。しかし、学部の研究では、常識知識とオブジェクトの組み合わせを調査するとき、多くの動画を繰り返し見てもらうだけで被験者への負担が多かった。そこで提案システムでは、Game With a Purpose(GWAP)を用いて特定の修飾語に対してオブジェクトのイメージに合った映像エフェクトの構造化データを獲得することができる。オブジェクトに絵文字を用い、修飾語のイメージを映像エフェクトで表示し、ユーザの投票によってユーザがゲームを楽しみながらデータが獲得されるゲームを開発した。提案システムではユーザが利用すればするほど修飾語とオブジェクトと映像エフェクトとの関係を学習できる。これにより、ゲームをプレイするたびにシステムがユーザの回答をもとに質問内容や提示する映像エフェクトを動的に変更する。

本論文では、同じ修飾語において対象の名詞によって異なる映像エフェクトとの結びつきが得られるかを検証するため、被験者に開発したシステムを利用してもらった。この実験により、映像エフェクトと常識知識、常識知識と名詞の結びつきを構造化データとして獲得できた。しかし、一つの映像エフェクトと複数の常識知識が結びつくなど偏りが見られたため、データ量を増やした時の変化を調査した。その結果、ゲームプレイ回数が増えることで偏りが見られていたデータも適切な結びつきに対してスコアが増えていき、複数の映像エフェクトと常識知識の結びつきが観察された。最後に、提案システムを修飾語に対しても適応した時のデータと組み合わせの事例を紹介する。

キーワード: 擬態語, 形容詞, 名詞, 常識知識, 映像エフェクト, Game with a Purpose, 絵文字

目次

第1章	序論	1
1.1	背景	1
1.2	本論文の構成	1
第2章	関連研究	2
2.1	研究目的	2
2.2	テキストをビジュアルオブジェクトに変える研究	2
2.3	Game with a Purposeに関する研究	3
2.4	常識知識を扱った研究	3
第3章	常識知識と映像エフェクトの結びつきを獲得するシステムの提案	4
3.1	システムを用いた常識知識獲得手法	4
3.2	得られた構造化データの利用方法	4
3.3	GWAPを用いた構造化データ収集システム	4
第4章	構造化データの獲得を目的とした実験	10
4.1	実験概要	10
4.2	結果	10
4.3	考察	11
第5章	構造化データの変化を調査することを目的とした実験	12
5.1	実験概要	12
5.2	結果	12
5.3	比較	14
5.4	アンケートに基づくユーザの反応	14
5.5	考察	15
第6章	本システムを修飾語で利用した時の構造化データの獲得を目的とした実験	16
6.1	実験概要	16
6.2	結果(常識知識について)	16
6.3	結果(映像エフェクトについて)	18
6.4	考察	23
第7章	映像エフェクト辞書システムと評価実験	25
7.1	映像エフェクト辞書システム	25
7.2	評価実験	27

7.3	実験結果の処理方法	28
7.4	結果	28
7.5	考察	30
第8章	おわりに	32
8.1	まとめ	32
8.2	今後の展望	33

第1章 序論

本研究は、開発した GWAP から得られたデータを利用して、抽象的な表現である形容詞やオノマトペといった修飾語を分類することができる新たな映像エフェクトと言葉とのリンクの構造を作成することを目的としている。GWAP とは、ゲームプレイの副産物として何らかの目的を達成しようとするゲームのことである。本システムの目的は、修飾語と名詞の組み合わせから想起される映像エフェクトのイメージを構造化するためのデータを GWAP によりユーザに楽しんでもらいながら収集する。

1.1 背景

人間同士のコミュニケーションにおいて使われる形容詞は、対象とする物によって言葉のイメージが異なる。例えば、「美味しい」という形容詞がある。「美味しいりんご」であれば艶のある果物を想起するが、「美味しいラーメン」となれば、湯気がモヤモヤしているラーメンを想起させるかもしれない。このように 同じ形容詞であっても対象物により想起させるイメージは異なる。学部で行った印象評価実験において、映像エフェクト「湯気」と「温かい食べ物」の組み合わせは「おいしい」という形容詞を想起させることが示唆された [1]。同様に、映像エフェクト「光」と「人工物」は、「新しい」という形容詞を想起させることに相関があった。「温かい」や「人工物」といった常識知識を利用することで、修飾語が表現するオブジェクトの状態を視覚的に表現することができる。そこで本研究では、「湯気、温かい食べ物」、「光、人工物」の組み合わせのように修飾語と常識知識の関係に着目した。しかし、常識知識の獲得には、多くのアンケートを実施する必要があった。そこで、GWAP による簡単なゲームインターフェースで常識知識を収集する「Effect Game」システムを開発した。

1.2 本論文の構成

本稿では、今まで獲得してきた構造化データと常識知識の紹介と、より多くのユーザが利用でき、多くの修飾語で利用できるよう拡張したシステムについて紹介する。本論文の構成は全 8 章で構成する。第 2 章で関連研究、第 3 章で GWAP を用いた構造化データ収集システムの提案と実装、第 4 章で構造化データの獲得を目的とした実験、第 5 章では構造化データの変化を調査することを目的とした実験、第 6 章で本システムを修飾語で利用した時の構造化データの獲得を目的とした実験、第 7 章で映像エフェクト辞書、第 8 章でまとめと今後の展望を述べる。

第2章 関連研究

2.1 研究目的

本研究は、開発した GWAP(Game with a purpose) から得られたデータを利用して、抽象的な表現である形容詞やオノマトペといった修飾語を分類することができる新たな映像エフェクトと言葉とのリンクの構造を作成することを目的としている。

2.2 テキストをビジュアルオブジェクトに変える研究

単語をビジュアルオブジェクトに変換する視覚化研究として、Text to Pictures, Text to Scenes, Text to Animations に変換する方法が開発されてきた。本研究は、映像エフェクトを用いた Text to Pictures に変換するためのデータベースを構築した。Text to Picture の研究として、Story Picturing Engine[2] [3] がある。Story Picturing Engine は、書かれたストーリーを画像や写真に変換し、名詞、形容詞、副詞、動詞に関連するランク付けされたイメージをキーワードとして表示する。また、TTP[4] はテキストから頻出かつ適切な名詞や形容詞を抽出し、複数の画像を結合して新しい画像を合成するシステムである。

テキストをシーンに変換する初期の研究では、NALIG[5] と PUT[6] は、対象、動詞、オブジェクトなどに基づいて 2 次元 (2D) オブジェクトを配置していた。WordsEye[7] は、3 次元 (3D) モデルのリポジトリから共通名詞と固有名詞に関連するオブジェクト、文字、および場所を検索して出力する。このシステムでは、属性 (例えば、色) が形容詞によって修正され、3D モデルの動作および姿勢が動詞から決定される。AVDT[8] は、前置詞や幾つかのオブジェクトを含む詳細な空間説明を用いて、自然言語のテキストから 3D シーンを生成するシステムである。空間関係の視覚化は、位置を表すメタデータを前置詞に割り当てることによって実現されている。暗黙的に関連する情報を視覚化するシステムもある [9][10][11]。例えば、「部屋にはコンピュータと椅子がある」という表現から机を出力することができる。テキストをアニメーションに変換する初期の研究では、SHRDLU[12] は自然言語を使用してスクリーン上でブロックを移動させた。傀儡 [13] は、より複雑な動きや言語表現を対象としたシステムである。音声情報を 3D オブジェクトとして表現している。傀儡は、3D 空間上の「小さな」や「もうちょっと」などの話された言葉に焦点を当てた自然言語理解システムである。形容詞の表現については、「遠い」や「青い」などの距離や色に関する表現も可能である。文字入力することで 3D アニメが表示される Anime de Blog[14][15][16] がある。Anime de Blog はキャラクターの動作を動詞に対応させることで、3D アニメを生成している。

このように、これまでの研究では、名詞や動詞、対象の色に関する形容詞表現が扱われていたが、修飾語を視覚化することは難しいため、修飾語を主な対象とする研究はなかつ

た。本研究は、今まで視覚化されてこなかった修飾語を常識知識と結びつけることで視覚化する。

2.3 Game with a Purpose に関する研究

Vonらは、ユーザが多く時間を費やすゲームへのエネルギーや時間を大規模な問題に対する解決策として利用できると述べており、そのようなゲーム設計を Games with a Purpose と呼んでいる [17][18]。本研究でも同様にユーザにはゲームを楽しんでもらうだけで無意識のうちに多くのデータが集まるよう設計を行った。中原らは、日本の常識知識を獲得する目的で「ナージャとなぞなぞ [19]」を開発した。このゲームでは、日本語の単語が持つ常識知識を獲得することで、ConceptNet[20] を拡張することが目的である。本研究では、修飾語を視覚化するための映像エフェクトとオブジェクトが持つ常識知識との結びつきを獲得する点が異なる。

2.4 常識知識を扱った研究

Cyc[21] では、専門家によって手作業でコモンセンス知識を登録されていた。知識は正確だが、限られた人にだけしか登録できず、獲得が難しかった。GWAPのようなユーザが意識せずに構造化データを収集するような試みがいくつかある。Open Mind Common Sense[22] は、Web を介して何千人もの人々から収集した常識知識データベースを開発した。同様に、常識知識を蓄積したデータベースに ConceptNet がある。ConceptNet 自体もインターネット上で複数の人から収集した常識知識データベースである。本研究では、ConceptNet などですでに獲得されたものではなく、ユーザから得られる常識知識を用いることが異なる。すでに獲得されたものでは、排除されてしまうような常識知識であっても修飾語という観点から強く結びつく常識知識が得られる点などが異なる。ConceptNet で獲得された膨大なデータを利用したアプリケーションも存在する。MAKEBELIEVE[23] は、ユーザが提供したストーリーからシステムと対話することで短いストーリーを展開する。Commonsense Predictive Text Entry[24] では、常識知識を用いた推論と統計的アプローチを組み合わせることで携帯電話の予測変換の方法を提案している。本研究で開発した映像エフェクト辞書システムでは、常識知識を使い、単語やストーリーを作り出すのではなく、視覚的に表現を行える映像エフェクトを利用している点で異なる。

第3章 常識知識と映像エフェクトの結びつきを 獲得するシステムの提案

3.1 システムを用いた常識知識獲得手法

ユーザから擬態語を視覚化するための構造化データを獲得する。本システムでは、常識知識と名詞、常識知識と映像エフェクトの組み合わせを獲得できる、ユーザにはゲームを楽しんでもらいながらシステムでは、データを獲得し、得られたデータをまたシステムでも利用するといった形式をとっている。そのため、ユーザから得られた常識知識が適切であれば、スコアは増加する。これにより、名詞が持つ常識知識を獲得することができる。また、適切な映像エフェクトについても獲得することができるため、擬態語と名詞の組み合わせのイメージに合った映像エフェクトとの繋がりを獲得することができる。

3.2 得られた構造化データの利用方法

擬態語と名詞の組み合わせによって異なる映像エフェクトとの結びつきを構造化データとして獲得することで、一つの擬態語に対してもそのイメージにあった映像エフェクトを表示することができる。例えば、「おいしい」という形容詞に対して、「りんご」や「ラーメン」といった異なる名詞に合わせて、映像エフェクトを表示することが可能になる。「りんご」であれば、「果物」という常識知識において適切な映像エフェクトをデータベースから検索し、表示できる。

3.3 GWAPを用いた構造化データ収集システム

本システムでは、映像エフェクトを通じた言葉のイメージの意見を収集する「Effect Game」を開発した。図 3.1 はシステムの流れである。

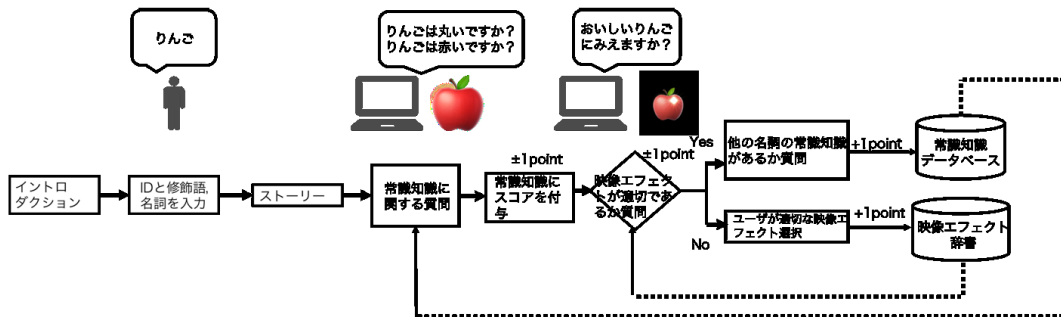


図 3.1: システムの流れ

ネット上を通してユーザにゲームをプレイしてもらうことを前提としているため、最初にイントロダクションの画面でこういったゲームかを説明し、IDを登録することでデータがセーブできるようにした。図 3.2 は擬態語と名詞を入力する画面である。



図 3.2: 擬態語と名詞を入力する画面

この画面からユーザは自分が表現したい好きな修飾語と名詞を選択し、ゲームをプレイすることができる。ここで選択する修飾語は修飾語登録画面から自由に登録することができる。入力された名詞に対して、システムは、ユーザが入力した名詞の絵文字を映しながら、○×の質問を提示する。名詞の絵文字は、絵文字の中から食物を用いている。図 3.3 では、○×のゲーム画面を示す。このゲーム中にゲーム画面の上の部分で敵がバトルを行う。

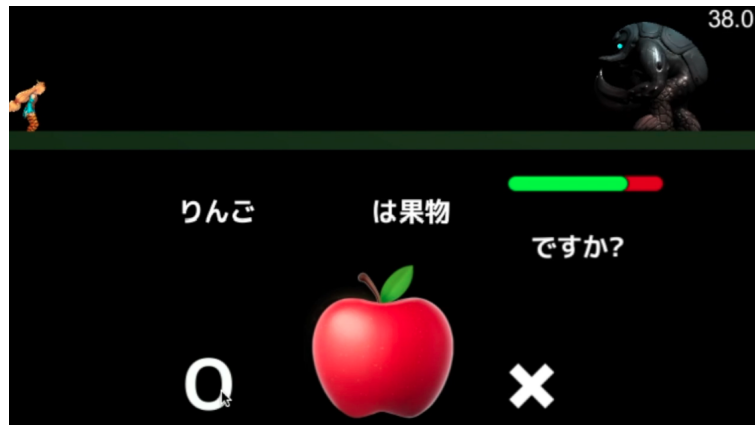


図 3.3: ○×ゲームの画面

現在名詞は 77 種類を登録している。質問内容は、その名詞に属する常識知識についてユーザに質問を行う。例えば、「それは果物ですか?」といった質問である。質問に用いる常識知識は以下の 2 種類である。(1) WordNet で一階層上の上位語 (2) ユーザから得られる新しい常識知識 WordNet の上位語とは、名詞が持っている広義の意味のことである。WordNet とは、意味辞書のことである [25]。今回は、登録数が多い英語の WordNet を用いる。例えば、アイスクリームであれば frozen desert, おかゆであれば dish といったものである。(2) では、ユーザから得られた常識知識を利用する。○×の質問が終了した後、システムは、ユーザから得られた常識知識をもとに映像エフェクトを選択する。映像エフェクトは、「おいしい」という形容詞のイメージに近い映像エフェクト 4 種類と「Unity ゲームエフェクト入門」という本から 36 種類の合計 40 種類を登録している。ユーザに 40 種類から適切な映像エフェクトを 1 つ選択してもらうため、映像エフェクトを 6 種類に分類した。映像エフェクトの例は図 3.4, 図 3.5, 図 3.6, 図 3.7, 図 3.8, 図 3.9 である。それぞれの分類と種類は以下である。

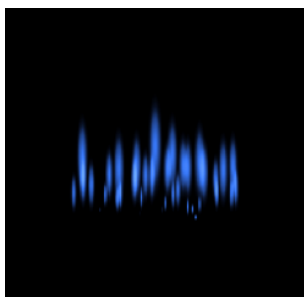


図 3.4: ゲームエフェクト



図 3.5: 光グループ

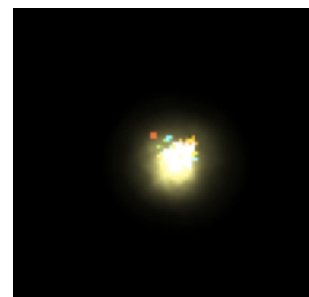


図 3.6: 光グループ 2

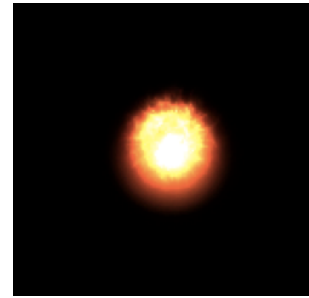
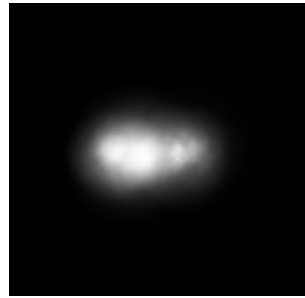
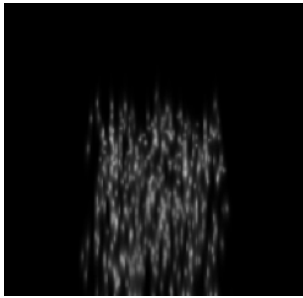


図 3.7: 水・冷気グループ 図 3.8: 湯気グループ 図 3.9: 炎・爆発グループ

それぞれのエフェクトの数はゲームエフェクトが6種類、光グループが8種類、光グループ2が7種類、水・冷気グループが8種類、湯気グループが6種類、爆発・炎グループが5種類である。本システムでは、システムが選択した映像エフェクトと名詞の組み合わせをユーザに提示し、システムが「エフェクトが適切だと思いますか」とユーザに Yes か No で問う。図 3.10 に YesNo ゲームの画面を示す。



図 3.10: YesNo ゲームの画面

Yes を選択した場合、映像エフェクトと名詞の組み合わせを適切だと判断する。その後、システムは「他にりんごの特徴はありますか?」といった質問をして、ゲームを終了する。No であった場合、システムは「この組み合わせの時に適切な映像エフェクトを選んでください」とユーザに質問を行う。映像エフェクト選択画面では、ユーザが適切だと思う映像エフェクトを選択してもらう。このゲームで、獲得されるデータは、○×ゲームが名詞に対して、常識知識が適切であるかを問う。Yes/No ゲームでは映像エフェクトと名詞の常識知識の組み合わせが適切であるかを問う。Yes を解答した場合、名詞が持つ未知の常識知識を獲得する常識知識記入画面に移動する。No の場合はユーザが考える新たな映像

エフェクトと名詞の組み合わせを取得する映像エフェクト選択画面に移動する。図 3.11 に常識知識記入画面を、映像エフェクト選択画面を図 3.12 に示す。



図 3.11: 常識知識記入画面

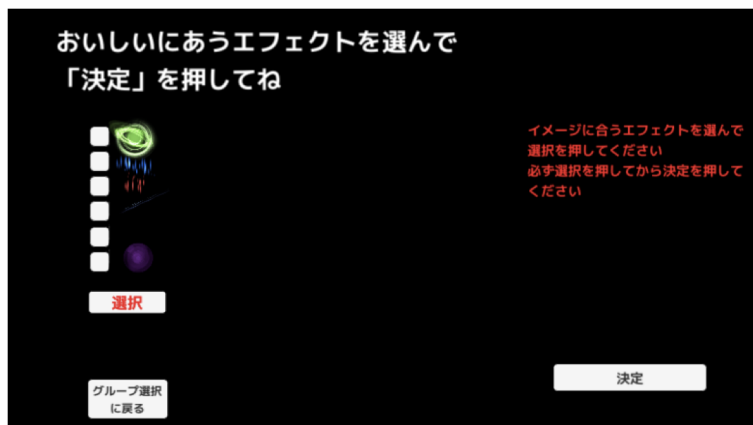


図 3.12: 映像エフェクト選択画面

これらの獲得されたデータはユーザが解答するごとに、それぞれの組み合わせに対して Score が 1 ずつ加算される。○×ゲームでは、次のユーザが解答する際、入力された名詞から Score の高いデータを○×の質問として更新していくことができる。次に、Yes/No 形式で獲得できるデータについて説明する。Yes/No では、食べ物と映像エフェクトが適切であるかを Score の点数で獲得することができる。Score が高いほど名詞と映像エフェクトが適切であることがわかる。ゲームから取得されるデータの保存は、ニフクラ Mobile Backend を用いて行った。ニフクラ mobile backend とは、mBaaS (mobile Backend as a Service) と呼ばれるスマートフォンアプリに利用される機能を提供するサービスの一つである。今回はニフクラ mobile backend のデータストアを利用して、取得される常識知識を保存、更新を行った。本ゲームでは、ユーザのモチベーションを保つ為、「Effect Game」にゲームストーリーを設定した。ゲームストーリーとして、主人公がモンスターに負けそ

うな場面を設定した。図 3.13 にストーリー画面を示す。

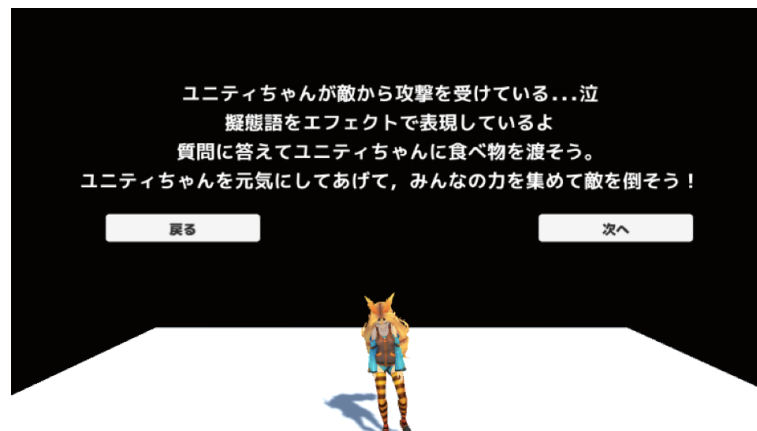


図 3.13: ストーリー画面

ユーザが「食べ物」を答え、食べ物に関する質問を○×で答える。その後、映像エフェクトと食べ物の組み合わせが表示され、適切な組み合わせであるかを Yes か No で答える。この Yes/No ゲームで同じ解答をした他のユーザが多いほど、主人公の攻撃する力が上昇する。一致する割合が一定数を越えた場合、モンスターを倒すことができる。

第4章 構造化データの獲得を目的とした実験

本章では、開発したシステムで生成された構造化データについて記述する。4.1では、開発したシステムを用いた実験の概要を説明する。4.2で実験結果について説明する。4.3では、結果から考えられる考察について説明する。

4.1 実験概要

本実験では、はこだて未来大学の学生6人を対象に実験を行った。用いた名詞は「りんご」、「オレンジ」、「ケーキ」、「餃子」、「ピザ」、「カレーライス」、「アイスクリーム」、「ソフトクリーム」の8種類を用いた。形容詞は「おいしい」のみを用いた。映像エフェクトに関しては、被験者一人と協力しておいしくみえる映像エフェクトを10種類選択した。

4.2 結果

被験者から得られたデータの一部を構造化した。構造化データを図4.1に示す。

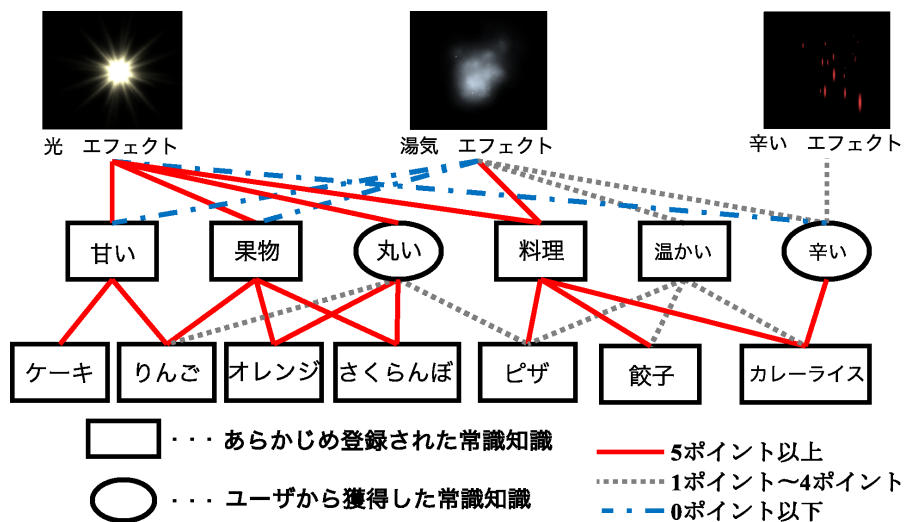


図 4.1: 構造化データの一部

長い赤色の線は5点以上を示す。灰色の点線は1~4点を示す。青い点線と短い線はマイナス点を示す。四角は、事前に登録されていた常識知識を表している。丸い形は、ユーザーから新たに獲得した常識知識を示している。○/×ゲームでは、「果物」と「りんご」、

「果物」と「オレンジ」,「丸い」と「オレンジ」はスコアが5点以上だった。また,「料理」に関しては,「料理」と「ピザ」,「料理」と「餃子」,「料理」と「カレーライス」が5点以上のつながりが獲得された。ユーザから獲得された常識知識「丸い」は,「オレンジ」と,常識知識「辛い」と「カレーライス」はそれぞれ5ポイント以上で獲得された。Yes/Noゲームでは,「料理」と「光」エフェクト,「果物」と「光」エフェクト,「料理」と「湯気」エフェクトが5ポイント以上で結びつきが獲得された。「デザート」と「湯気」エフェクト,「甘い」と「湯気」エフェクト,「果物」と「湯気」エフェクト,「温かい」と「光」エフェクトは, マイナスのスコアであった。1点から4点の結びつきは図 4.1 に示す。

4.3 考察

本章では, 映像エフェクトを利用して構造化データと未知の常識知識を獲得する GWAP を紹介した。実験結果では,「光」エフェクトに複数の常識知識と結びつきが示唆された。WordNet と同様に「料理」との結びつきとして「カレー」や「ピザ」との結びつきが獲得されました。提案したシステムでは,「温かい」と「湯気」エフェクトのように映像エフェクトとの結びつきも獲得された。さらに本システムでは,「辛い」や「丸い」などの新たな常識知識を獲得した。次に, 得られた結果の常識知識と映像エフェクトの結びつきに注目した。「温かい」は「湯気」エフェクトが適切であり,「辛い」の場合は「辛い」エフェクトが適切であった。「果物」については,「光」エフェクト,「甘い」および「丸い」は「光」エフェクトが示唆された。これらの結果から形容詞「おいしい」を表現する際, 常識知識によって異なる映像エフェクトとの結びつきが観測された。つまり, 異なるイメージを持つ形容詞を映像エフェクトと常識知識を用いて視覚化できることが示唆された。

赤い線（長い線）は5点以上を示している。灰色の線（点線）は1~4点を示している。青い線（点線と短い線の繰り返し）は0点未満を示している。図中の四角は、WordNetから登録した常識知識を示している。丸い形は、ユーザーから得られた常識知識や色彩や味などの常識知識を示している。○×形式では、「料理」と「餃子」、「料理」と「ピザ」、「料理」と「カレーライス」、「辛い」と「カレーライス」、「デザート」と「アイスクリーム」、「デザート」と「ソフトクリーム」、「果物」と「りんご」、「果物」と「オレンジ」が5点以上であった。また、新しくユーザーから得られた常識知識として「甘い」が存在した。「甘い」に関しては、「甘い」と「ケーキ」、「甘い」と「りんご」で5点以上の結びつきであった。Yes/Noゲームでは、「光」エフェクトと「料理」、「光」エフェクトと「甘い」、「光」エフェクトと「冷たい」、「光」エフェクトと「丸い」が5点以上であった。「パウダー」エフェクトと「料理」、「パウダー」エフェクトと「果物」、「光」エフェクトと「温かい」、「光」エフェクトと「辛い」は否定的なつながりであった。次に3日目ゲームをプレイしてもらったときにできた構造化データの結果が図5.2である。

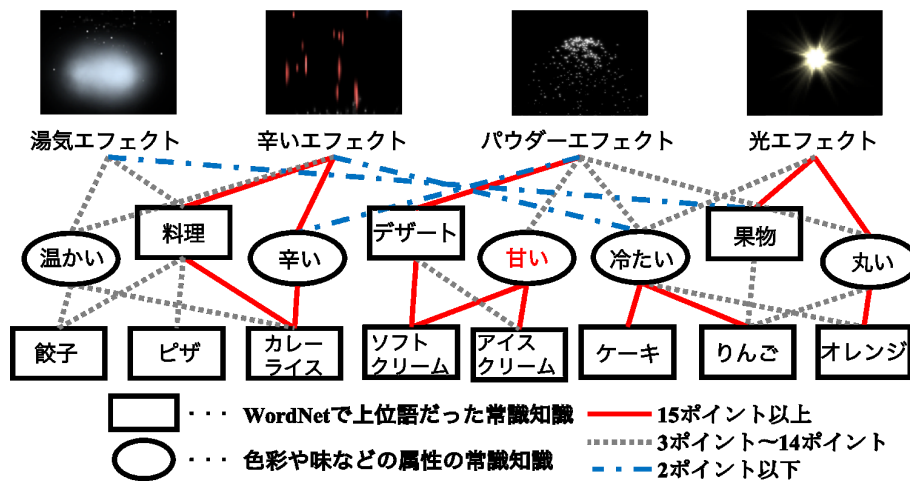


図 5.2: 3 周目の構造化データ

結びつきのスコアに関しては、3 回分プレイしてもらったため、3 倍に直して表記する。○×形式では、「料理」と「カレーライス」、「辛い」と「カレーライス」、「デザート」と「ソフトクリーム」、「冷たい」と「ソフトクリーム」、「冷たい」と「アイスクリーム」、「甘い」と「ケーキ」、「甘い」と「りんご」、「丸い」と「オレンジ」が5点以上であった。Yes/Noゲームでは、「辛い」エフェクトと「料理」、「辛い」エフェクトと「辛い」、「パウダー」エフェクトと「デザート」、「光」エフェクトと「果物」、「光」エフェクトと「丸い」が5点以上であった。「湯気」エフェクトと「果物」、「辛い」エフェクトと「甘い」、「パウダー」エフェクトと「辛い」、「光」エフェクトと「料理」は否定的なつながりであった。

5.3 比較

複数回データを取得した時の構造化されたデータの変化を調査するため、1回目と3回目の構造化されたデータを比較した。1回目のデータでは、名詞と常識知識の結びつきが適切であった。例えば、「料理」に属しているものとして、「餃子」や「ピザ」、「カレーライス」が強く結びついていた。これは WordNet の構造化データと同様の結びつきとなっているため、適切であると考えられる。同様に WordNet で上位語であると述べられている「デザート」や「果物」にも同様の名詞が結びついている。しかし、映像エフェクトと名詞の常識知識の結びつきに着目すると、「光」エフェクトに偏っていることがわかった。そして、他のエフェクトで5点以上の映像エフェクトが存在しなかった。一方で、3回やってもらったゲームの構造化データでは、複数の映像エフェクトと15点以上の結びつきが存在することがわかった。

5.4 アンケートに基づくユーザの反応

3回のゲームをしてもらった後、アンケートを行った。本アンケートの目的は、ユーザが楽しめるゲームとなっているかを調査することである。質問項目は以下である。

- 攻撃力がどのように上がっていたか理解できましたか
- おいしいを適切に表現した映像エフェクトはありましたか
- ゲームは使いやすかったですか (操作性)
- もう一度ゲームをしたいと思いましたか

「おいしいを適切に表現した映像エフェクトはありましたか」という質問項目では、「複数あった、一つだけあった、わからない、なかった」の4段階評価であった。結果としては、「複数あった」が87.5%、「わからない」が12.5%であった。「もう一度ゲームをしたいと思いましたか」という質問項目では、「また利用したい、利用したい、わからない、利用したくない、もう利用したくない」の5段階評価であった。結果として、「利用したい」が62.5%、「わからない」が37.5%であった。「ゲームの設定は面白かったですか」という質問項目では、「とても面白かった、面白かった、わからない、面白くない、とても面白くない」の5段階評価であった。結果としては、とても面白かったが25.0%、面白かったが37.5%であった。自由記述で回答してもらった意見が表 5.1 である。

表 5.1: 自由記述で得られた回答

面白くするために増やして欲しい機能はありますか
名詞の種類を8単語だけでなく、もっと増やして欲しい
攻撃パターンの種類を増やして欲しい
映像エフェクトを複数選べるようにして欲しい

5.5 考察

抽象的な表現である形容詞を分類することができる新たな映像エフェクトと言葉とのリンクの構造を作成することを目的として、ユーザに楽しんでもらいながら構造化データを獲得する手法を提案した。実験結果では、一つ一つの映像エフェクトには複数の常識知識と適切な組み合わせがあることが示唆された。さらに、WordNetに登録されている「料理」や「果物」、「デザート」によって、適切である映像エフェクトが存在することがわかった。今回の実験では、ユーザが入力された「甘い」が質問に使われていた。ユーザから得られたデータであっても高いスコアを獲得することがわかった。また、1回ゲームをしてもらった時の構造化データと3回ゲームをしてもらった時の構造化データを比較した結果、3回ゲームをしてもらった構造化データの方がより多くの映像エフェクトと名詞の常識知識の適切な組み合わせを獲得することができることが示唆された。これにより、本システムは一つの映像エフェクトに偏らず、適切な組み合わせをユーザから獲得することができると思われる。アンケート結果では、「おいしいを適切に表現した映像エフェクトはありましたか」という質問項目に対して、半数以上が複数あったと回答していた。よって、映像エフェクトで「おいしい」という形容詞を適切に可視化できることが示唆された。また、「もう一度ゲームをしたいと思いませんか」という質問項目では、「利用したい」が半数以上を占めていた。そして、「ゲームの設定は面白かったですか」という質問では、「とても面白い」が25%で「面白い」が37.5%という結果であった。この2つのアンケート結果から利用者が自主的にこのゲームを繰り返し行えるような面白いゲーム設定であると示唆された。一方で、「ゲームは使いやすかったですか(操作性)」という質問項目では、「使いづらい」が半数以上を占めていたため、操作性を向上して行く必要がある。

第6章 本システムを修飾語で利用した時の構造化データの獲得を目的とした実験

6.1 実験概要

本章では、常識知識を獲得する Web ゲーム「Effect Game」について述べる。「Effect Game」を用いて、擬態語に適応した時に常識知識によって映像エフェクトの結びつきが変化するかを調査した。被験者は事前調査で 3277 名が回答した。事前調査では Web ゲームの特性上、動作が被験者のパソコンに依存するため、パソコンのスペックについて質問を行なった。事前調査で条件が合った被験者 87 名を対象に本システムを利用してもらった。被験者 87 名は男性が 58 名、女性が 29 名だった。年齢は 10 代が 1 名 20 代が 13 名、30 代が 28 名、40 代が 45 名であった。今回の実験では、スコアがプラスのものをランダムに取得する方式を採用した。開発したゲームを各プレイヤーが 10 回名詞と擬態語を組み合わせてもらい、プレイしてもらった。用いた擬態語は、「みずみずしい」、「ほくほく」、「ふわふわした」、「ほっこりな」、「しゅわしゅわ」、「あつあつな」、「こんがりした」、「ひんやりした」、「しっとりした」、「ジューシーな」の 10 種類である。選定方法は日本語テクスチャー用語体系 [26] から映像エフェクトで表現できそうな修飾語を 6 種類選択した。そして、映像エフェクトを用いるため湯気や炎、水っぽさを表現している修飾語「あつあつ」、「こんがりした」、「ひんやりした」、「しっとりした」の 4 種類を選択した。名詞は「じゃがいも」、「くり」、「ケーキ」、「パン」、「餃子」、「ビール」、「オレンジ」、「りんご」、「ぶどう」、「カレーライス」、「クッキー」、「ソフトクリーム」、「にんじん」、「お茶」、「ジュース」の 15 種類である。名詞は絵文字で用意できるものの中から、WordNet の常識知識「果物」、「料理」、「デザート」、「飲み物」、「オーブンで焼かれた食品」、「野菜」をもとに 15 種類を選択した。

6.2 結果 (常識知識について)

本実験でユーザから新たに 332 種類の常識知識を獲得した。「餃子」と結びついた常識知識は図 6.1, 2 点以上のスコアの組み合わせを表 6.1 である。

結びついていた。

6.3 結果 (映像エフェクトについて)

まず初めに、修飾語「ジューシーな」の結びつきを図 6.2 に名詞とエフェクトの組み合わせのスコアを表 6.2 に示す。結びつきの図と表は 1 点のものを含むと数が増加し、図に示しづらいため、スコアが 2 点以上のものを示す。スコアが 5 点以上の結びつきを赤線で表している。修飾語「ジューシーな」において、「エフェクト 37」と「料理」、「エフェクト 37」と「元気がでそう」、「エフェクト 7」と「果物」、「エフェクト 7」と「丸い」、「エフェクト 7」と「野菜」、「エフェクト 10」と「丸い」、「エフェクト 10」と「果物」、「エフェクト 14」と「野菜」が高いスコアで結びついていた。

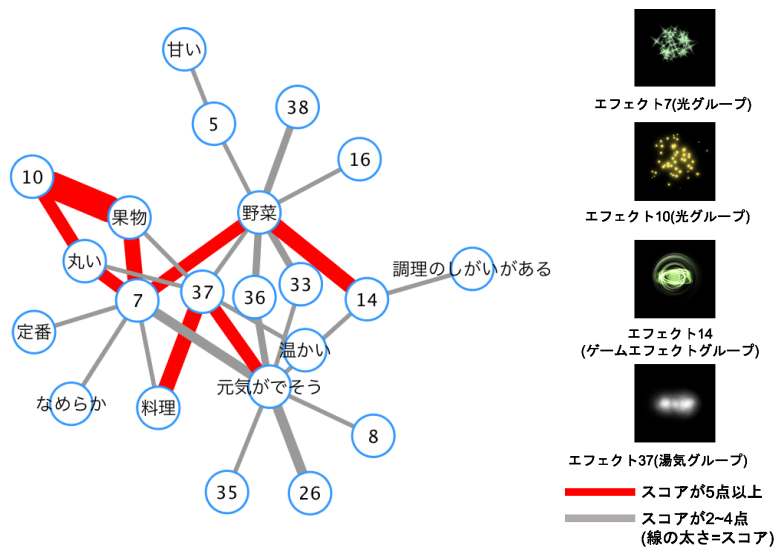


図 6.2: 修飾語「ジューシーな」と常識知識の結びつき

表 6.2: 「ジューシーな」の時の常識知識と映像エフェクトのスコア

常識知識	エフェクト	スコア
果物	10	10
果物	7	6
料理	37	6
野菜	14	6
野菜	7	5
丸い	7	5
丸い	10	5
元気がでそう	37	5
元気がでそう	26	4
元気がでそう	7	4
野菜	33	3
野菜	38	3
野菜	36	3
元気がでそう	36	3
野菜	5	2
元気がでそう	8	2
料理	7	2
野菜	37	2
温かい	37	2
元気がでそう	33	2
元気がでそう	35	2
果物	37	2
定番	7	2
なめらか	7	2
丸い	37	2
調理のしがいがある	14	2
野菜	16	2
甘い	5	2
元気がでそう	14	2

次に「しゅわしゅわした」の結びつきを図 6.3 に、常識知識と映像エフェクトのスコアを表 6.3 示す。「しゅわしゅわした」では、「エフェクト 32」と「飲み物」,「エフェクト 43」と「飲み物」,「エフェクト 41」と「元気が出そう」が高いスコアで結びついた。

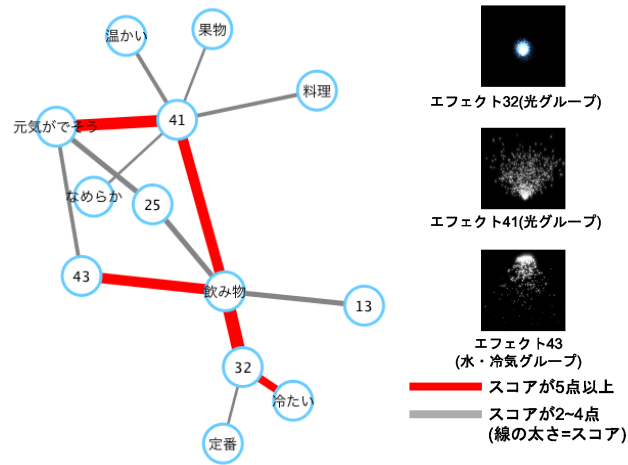


図 6.3: 修飾語「しゅわしゅわした」と常識知識の結びつき

表 6.3: 「しゅわしゅわした」ときの常識知識と映像エフェクトのスコア

常識知識	エフェクト	スコア
飲み物	32	10
元気がでそう	41	9
飲み物	41	8
飲み物	43	8
冷たい	32	6
飲み物	13	4
元気がでそう	25	4
飲み物	25	4
温かい	41	3
元気がでそう	43	3
料理	41	3
果物	41	2
定番	32	2
なめらか	41	2

「しっとりした」の結びつきを図 6.4 と常識知識と映像エフェクトのスコアを表 6.4 に示す。「しっとりした」において、「元気がでそう」と「エフェクト 7」が 5 点以上の結びつきを示した。

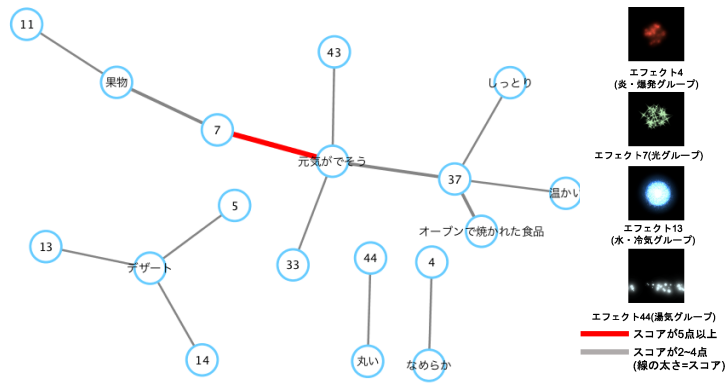


図 6.4: 修飾語「しっとりした」と常識知識の結びつき

表 6.4: 「しっとりした」ときの常識知識と映像エフェクトのスコア

常識知識	エフェクト	スコア
元気がでそう	7	5
果物	7	3
オープンで焼かれた食品	37	3
元気がでそう	37	3
果物	11	2
温かい	37	2
丸い	44	2
元気がでそう	33	2
元気がでそう	43	2
しっとり	37	2
なめらか	4	2
デザート	13	2
デザート	14	2
デザート	5	2

「みずみずしい」の結びつきを図 6.5 と常識知識と映像エフェクトのスコアを表 6.5 に示す。「みずみずしい」では、「果物」と「エフェクト 18」,「果物」と「エフェクト 26」が5点以上で結びついた。

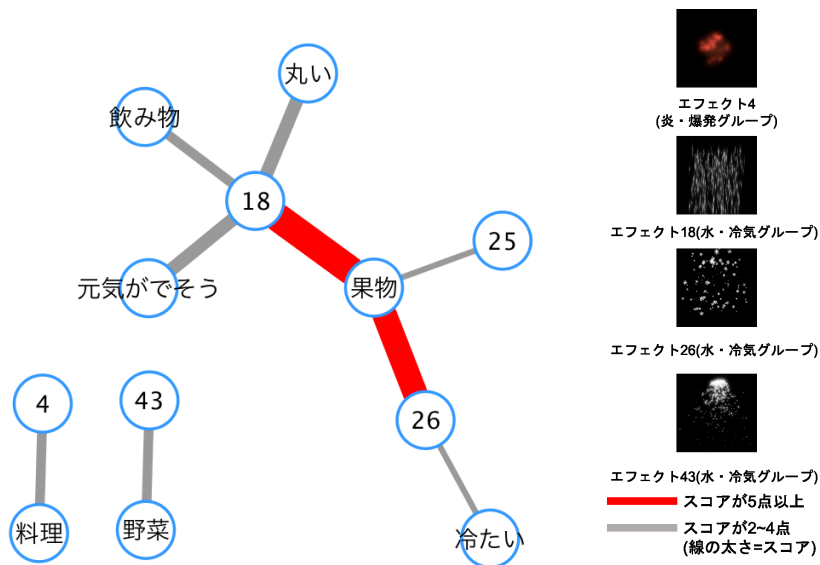


図 6.5: 修飾語「みずみずしい」と常識知識の結びつき

表 6.5: 「みずみずしい」のときの常識知識と映像エフェクトのスコア

常識知識	映像エフェクト	スコア
果物	18	8
果物	26	7
丸い	18	4
元気がでそう	18	4
飲み物	18	3
料理	4	3
野菜	43	3
果物	25	2
冷たい	26	2

「ほっこりした」の結びつきを図 6.6 と常識知識と映像エフェクトのスコアを表 6.6 に示す。「ほっこりした」では、「野菜」と「エフェクト 6」が 5 点以上で結びついた。

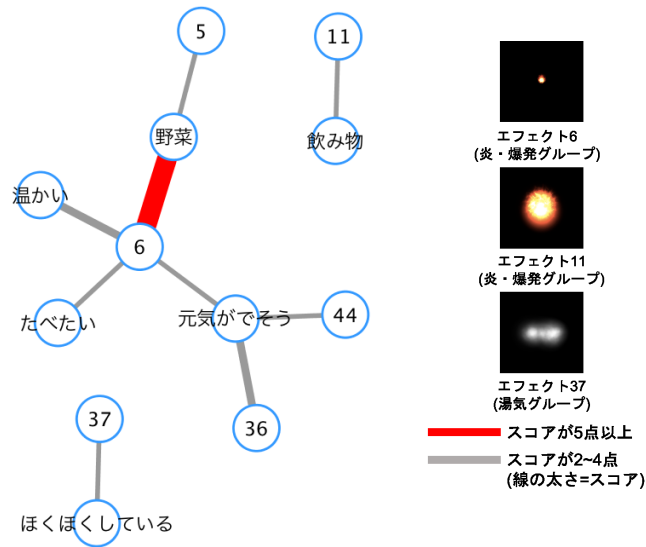


図 6.6: 修飾語「ほっこりした」と常識知識の結びつき

表 6.6: 「ほっこりした」のときの常識知識と映像エフェクトのスコア

常識知識	映像エフェクト	スコア
野菜	6	7
温かい	6	3
元気がでそう	36	3
元気がでそう	6	2
元気がでそう	44	2
飲み物	11	2
野菜	5	2
たべたい	6	2
ほくほくしている	37	2

6.4 考察

抽象的な表現である修飾語を分類することができる新たな映像エフェクトと言葉とのリンクの構造を作成することを目的として、ユーザに楽しんでもらいながら構造化データを獲得する Web ゲーム「Effect Game」を開発した。名詞と常識知識の結びつきに関しては、「餃子」が「料理」に属しているといった WordNet で得られた常識的なものが高スコアが見つかることが示唆された。このことから適切な常識知識と名詞の組み合わせに対して、システムを利用することで構造化データとして組み合わせを獲得できると考えられる。次に

常識知識と映像エフェクトの組み合わせについて述べる。「しっとりした」では「丸い」という常識知識の時には湯気グループの「エフェクト44」,「なめらか」という常識知識では炎・爆発グループの「エフェクト4」と結びつきが見られた。「みずみずしい」のときでは、「料理」と炎・爆発グループの「エフェクト4」,「野菜」と水・冷氣グループの「エフェクト43」,「果物」であれば、水・冷氣グループの「エフェクト18,26」と結びつきが見られた。「ほっこりした」では、「ほくほくしている」と湯気グループの「エフェクト37」,「飲み物」と炎・爆発グループの「エフェクト11」と結びつきが見られた。このことから同じ修飾語であっても常識知識によって異なるグループの映像エフェクトと結びつくことが示唆された。従って、同じ修飾語であっても対象の名詞によって異なるイメージが存在し、本システムを利用することで常識知識に合った映像エフェクトの組み合わせを獲得できることが示唆された。しかし、「しっとりした」や「ほっこりした」といった修飾語では、常識知識によって映像エフェクトと結びついているが、スコアが低く必ずしも適切である組み合わせではないものも見られた。また、「しゅわしゅわした」では、「冷たい」と「飲み物」は、「エフェクト32」と結びついているが、「温かい」と「飲み物」では、「エフェクト41」と結びついていた。従って、今回得られたデータでは、「温かい」かつ「飲み物」という組み合わせに対しては、「エフェクト32」のイメージであるが、「冷たい」かつ「飲み物」であれば、「エフェクト41」のイメージといったように常識知識を組み合わせることでより細かいイメージの違いも表現できることが示唆された。しかし、今回の実験では、「ホットティー」など明確に「温かい」といった常識知識を持つ飲み物は含んでおらず、「料理」という常識知識を持つ名詞から「温かい」という常識知識が結びついたと考えられるため、常識知識の組み合わせにも着目する必要がある。

第7章 映像エフェクト辞書システムと評価実験

本章では、実験で得られた構造化データを活用した言葉を視覚化するシステム「映像エフェクト辞書システム」の紹介とそのシステムがユーザに提示する映像エフェクトと名詞の組み合わせの適切さを評価した評価実験について述べる。

7.1 映像エフェクト辞書システム

現在開発した映像エフェクト辞書の画面を図 7.1 に示す。この映像エフェクト辞書では、今まで獲得した修飾語と名詞を入力することで、修飾語と常識知識、常識知識と映像エフェクトの組み合わせから高いスコアのものを検索、表示することができる。これにより、「おいしいりんご」と「おいしいラーメン」のように対象によって異なるイメージをもつ修飾語を映像エフェクトで表示できる。

映像エフェクト辞書

擬態語入力

名詞入力

エフェクトを調べる

みずみずしい、ほくほくな、ふわふわな、ほっこりした
しゅわしゅわした、あつあつな、こんがりした
ひんやりした、しっとりした、ジューシーな、おいしい

図 7.1: 映像エフェクト辞書画面

例えば、6章で得られた修飾語「ジューシーな」に関して映像エフェクト辞書で検索した例が図 7.2, 図 7.3, 図 7.4 に示す。それぞれ、修飾語「ジューシーな」と名詞「餃子」, 「ジューシーな」と「にんじん」, 「ジューシーな」と「りんご」を検索した結果である。こ

これらの結果は、映像エフェクト辞書に格納されているデータの組み合わせでスコアの高い組み合わせをシステムが表示する。図 7.2 では、「餃子」が持つ常識知識「料理」や「元気がでそう」とスコアが高い組み合わせである映像エフェクト 37 をシステムが選択し、表示している。次に、「しゅわしゅわした」の組み合わせの例を図 7.5, 図 7.6, 図 7.7 に示す。これは修飾語「しゅわしゅわした」と名詞「飲み物」、「しゅわしゅわした」と「元気がでそう」を組み合わせた時の結果を示している。最後に、「しっとりした」の組み合わせの例を図 7.8, 図 7.9 に示す。これは修飾語「しっとりした」と「果物」、「しっとりした」と「オープンで焼かれた食品」を組み合わせた時の結果を示している。

この辞書システムを拡張して行くことで、システムには登録されていない名詞が入力されたとしても、常識知識をシステムが質問することで、未知の名詞に対しても適切な映像エフェクトを選択することができる。



図 7.2: エフェクト 37 と「料理」の組み合わせ

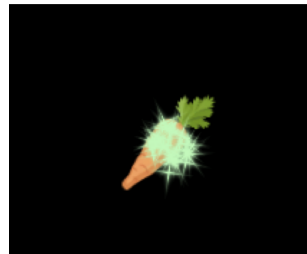


図 7.3: エフェクト 7 と「野菜」の組み合わせ

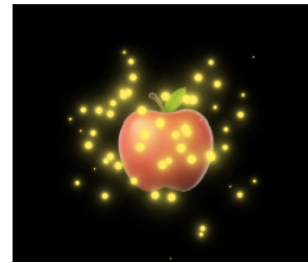


図 7.4: エフェクト 10 と「果物」の組み合わせ

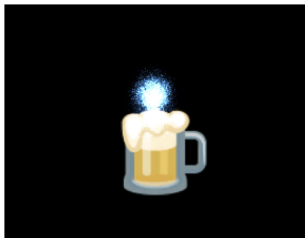


図 7.5: エフェクト 32 と「飲み物」の組み合わせ

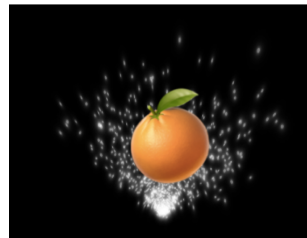


図 7.6: エフェクト 41 と「元気が出そう」の組み合わせ

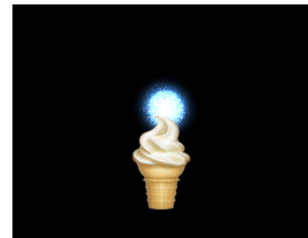


図 7.7: エフェクト 32 と「冷たい」の組み合わせ

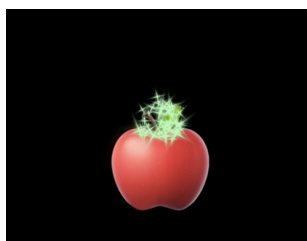


図 7.8: エフェクト 7 と「果物」の組み合わせ

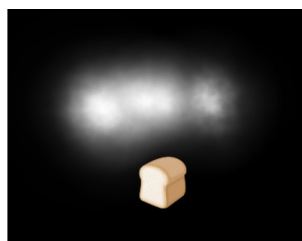


図 7.9: エフェクト 37 と「オープンで焼かれた食品」の組み合わせ

7.2 評価実験

本システムがユーザに提示する映像エフェクトと名詞の組み合わせが適切であるかどうかを調査するため、評価実験を行なった。被験者は公立はこだて未来大学の学生 16 人を対象に行なった。被験者には、40 種類の映像エフェクトと名詞を組み合わせさせた動画をみてもらい、修飾語を想起できるかを 7 つの選択肢から選んでもらった。選択肢は「とても思う」「思う」「やや思う」「やや思わない」「思わない」「とても思わない」「想起できるか判別できない」の 7 つである。40 種類の組み合わせは「Effect Game」で得られた映像エフェクトと名詞の常識知識の組み合わせのスコアがマイナスだった組み合わせから 8 種類、0 点だったものから 8 種類、1 点だったものから 8 種類、2 から 4 点だったものから 8 種類、5 点以上だったものから 8 種類を選択した。組み合わせデータは全部で 661 種類であったがそのうち 358 個の組み合わせが 1 点だったため、一つの群とした。映像エフェクト辞書システムでは高いスコアのもを推薦するシステムのため、5 点以上の群を設けた。プラスのスコアのもが組み合わせとして適切であることを評価するため、比較用にマイナスの群を作成した。また正でも負でもない 0 点のグループとプラスのスコアである 2 から 4 点の群を設けた。表 7.1 にマイナス群と 5 点以上の群の質問内容と質問番号を示す。

今回の評価実験では、それぞれの群内では、有意差がない群で分けられ、本システムがユーザに提示するスコアの高い群とそれ以外の群同士を比較した際に有意差がみられることで本システムがユーザに提示する映像エフェクトと名詞の組み合わせが適切であると仮定した。

表 7.1: 質問内容と質問番号の組み合わせ

群	質問番号	質問内容
マイナス群	1	「ジューシーなじゃがいも」を想起できるか
	2	「ひんやりした餃子」を想起できるか
	3	「ジューシーなにんじん」を想起できるか
	4	「しゅわしゅわしたケーキ」を想起できるか
	5	「ほっこりしたパン」を想起できるか
	6	「あつあつなお茶」を想起できるか
	7	「あつあつなパン」を想起できるか
	8	「ひんやりしたカレーライス」を想起できるか
5点以上の群	33	「ひんやりしたソフトクリーム」を想起できるか
	34	「ジューシーなりんご」を想起できるか
	35	「みずみずしいりんご」を想起できるか
	36	「しゅわしゅわしたビール」を想起できるか
	37	「ジューシーな餃子」を想起できるか
	38	「しゅわしゅわしたお茶」を想起できるか
	39	「しゅわしゅわしたオレンジ」を想起できるか
	40	「みずみずしいぶどう」を想起できるか

7.3 実験結果の処理方法

アンケート結果の処理方法として、各質問の選択肢「とても思う」、「やや思う」、「思う」、「やや思う」、「やや思わない」、「思わない」、「とても思わない」を3点、2点、1点、-1点、-2点、-3点とした。「想起できるか判別できない」に対しては0点として集計した。点数付けの理由として、システムのスコアを参考に群分けを行なったため、評価された質問の平均値も負から正の点数でプロットされたほうがわかりやすいと感じたため、このような点数付けとした。今回の実験では、あらかじめスコアを基準に5群に分けていたため、3群以上かつ平均値にばらつきがあるため、被験者内要因の一元配置の分散分析による多重比較検定をおこなった。

7.4 結果

図 7.10 はマイナス群と5点以上の群の質問に関する平均点と標準偏差を示している。マイナス群のすべての質問で平均点は負の数を示していた。また、5点以上の群でも全ての質問の平均点が正の数を示していた。表 7.2 はマイナス群だけで比較した有意確率を示した表、表 7.3 は5点以上の群だけで比較した有意確率を示した表である。n.s. は有意差なし、†は10%水準で有意傾向を示す。*は5%水準で有意を示す。どちらの群内でも有意差はみられなかった。

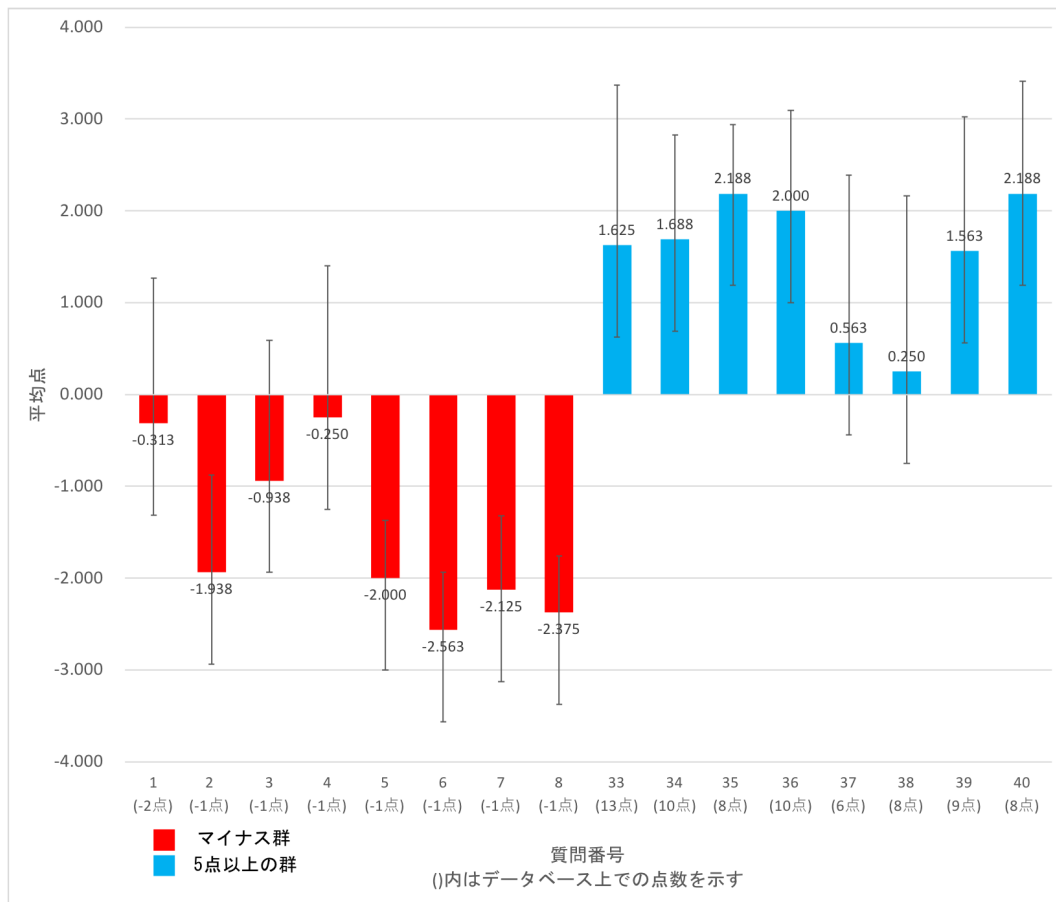


図 7.10: マイナス群と 5 点以上の群の質問の平均点と標準偏差

表 7.2: マイナスの群だけで比較した有意確率を示した表

質問番号	1	2	3	4	5	6	7	8
1		n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
2			n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
3				n.s.	n.s.	n.s.	n.s.	n.s.
4					n.s.	n.s.	n.s.	n.s.
5						n.s.	n.s.	n.s.
6							n.s.	n.s.
7								n.s.
8								

表 7.3: 5 点以上の群だけで比較した有意確率を示した表

質問番号	33	34	35	36	37	38	39	40
33		n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
34			n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
35				n.s.	n.s.	n.s.	n.s.	n.s.
36					n.s.	n.s.	n.s.	n.s.
37						n.s.	n.s.	n.s.
38							n.s.	n.s.
39								n.s.
40								

次に、表 7.4 はマイナス群と 5 点以上の群を比較した時の有意確率を示した表である。マイナス群の質問番号 4 の「しゅわしゅわしたケーキ」では、5 点以上の群と比較した結果、質問番号 33 の「ひんやりしたソフトクリーム」以外とは有意差が見られなかった。また質問番号 38 の「しゅわしゅわしたお茶」が多くのマイナス群と有意差がみられなかった。質問番号 37 の「ジューシーな餃子」は質問番号 1,3,4 と有意差がみられなかった。質問番号 39 の「しゅわしゅわしたオレンジ」では、質問番号 1,4 と有意差がみられなかった。質問番号 40 の「みずみずしいぶどう」と質問番号 4 と有意差がみられなかった。それ以外の組み合わせでは 10%水準の有意傾向または 5%水準で有意であった。

表 7.4: マイナスの群と 5 点以上の群を比較した有意確率を示した表

質問番号	33	34	35	36	37	38	39	40
1	†	*	*	*	n.s.	n.s.	n.s.	*
2	*	*	*	*	†	n.s.	*	*
3	†	*	*	*	n.s.	n.s.	†	*
4	n.s.	†	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.
5	*	*	*	*	*	n.s.	*	*
6	*	*	*	*	*	†	*	*
7	*	*	*	*	*	†	*	*
8	*	*	*	*	*	†	*	*

7.5 考察

被験者内要因の一元配置の分散分析による多重比較検定の結果から、本システムがユーザに提示する映像エフェクトと名詞の組み合わせが適切であることが認められた。なぜならマイナス群と 5 点以上の群を総当たりで比較した結果から、5 点以上のスコアを有している動画を見てもらった時の質問群がマイナスのスコアを有している動画を見てもらった時の質問群よりも有意に想起できていることがわかったためである。しかし、質問番号 4 の「しゅわしゅわしたケーキ」や質問番号 38 の「しゅわしゅわしたお茶」では特に多く

の組み合わせで有意差がみられなかった。このことから、「しゅわしゅわした」という修飾語では、システムが示す映像エフェクトが適切ではない可能性があり、異なる手法が必要になることが考えられる。またマイナス群内の全通りの動画同士の組み合わせにおいて有意差が無いため、想起できるかどうかの認識に明確な相違は見られなかった。また、負の数の群内の平均点は全て負の数であったため、想起できない傾向が強かった。したがって、負の数のグループ内の全ての動画は、想起できないような名詞とエフェクトの組み合わせであり、動画間の認識にも差異はないと考えられる。同様に5点以上の群でも平均点は全て正の数であり、全通りの動画同士の組み合わせで有意差はなかった。つまり、5点以上の群内のすべての動画は、想起できるような名詞と映像エフェクトの組み合わせであり、全通りの動画同士の組み合わせで有意差はなかった。

第8章 おわりに

8.1 まとめ

本論文では、抽象的な表現である形容詞やオノマトペといった修飾語を分類することができる新たな映像エフェクトと言葉とのリンクの構造化データ獲得手法を提案した。そのために、ユーザが楽しみながらシステムがデータを獲得するための Web ゲームである、Effect Game を開発した。

開発したゲームによる構造化データの検証を行うため、形容詞「おいしい」を題材に被験者実験を行なった。1度目の実験では、「おいしい」という同じ形容詞であっても常識知識によって異なる映像エフェクトと高いスコアで結びつく結果であった。このことから複数のイメージを持つ形容詞を映像エフェクトを用いることで視覚的に表現できることが示唆された。しかし、この実験では、一つの映像エフェクトに集中して高いスコアが結びついてしまう問題点があった。

1度目の実験データが人数が少なめであったことから2度目の実験では、同じ実験を3回行い、データ量を増やすことで構造化データの変化を調査することを目的とした。結果としては一回目の構造化データでは、一度目の実験同様に一つの映像エフェクトに対して複数のスコアの高い結びつきが集中していた。しかし、3回ゲームをプレイしてもらった構造化データでは、複数の映像エフェクトと常識知識が高いスコアで結びついていくことが示唆された。

3度目の実験では、本システムを修飾語に適応したときの構造化データを調査することを目的とした。「ジューシーな」や「しゅわしゅわした」といった視覚的にイメージしやすい修飾語に対して本システムを利用することでそのイメージにあった映像エフェクトと常識知識が高いスコアで結びつくことが示唆された。一方で「しっとりした」といった修飾語に対しては映像エフェクトで適切に表現することは難しかった。つまり、修飾語にも本システムを利用する際には視覚で表現される言葉をあらかじめ選ぶことでより適切に言葉を映像エフェクトで表現できると考えられる。表現できない修飾語に対しては、例えば擬音語といったものには音声エフェクトを使って表現を行うなど修飾語にあったエフェクトを拡張する必要がある。最後に実験で得られたデータを活用する映像エフェクト辞書において表示される映像エフェクトと名詞の組み合わせが適切であることを示すため評価実験を行なった。評価実験の結果から本システムがユーザに提示する映像エフェクトと名詞の組み合わせが適切であることが認められた。

8.2 今後の展望

本研究の活用する場面として、消費者生成メディアとしての利用が期待される。消費者によって修飾語ごとに映像エフェクトで表現してもらうことで、蓄積したデータをもとに、映像エフェクトからイメージの検索が行える。Googleでも画像検索で、形容詞から絞り込むような検索システムが存在するが、アップロードされている画像のみがリストアップされており、果たして本当にその形容詞にあった画像が表示されているかは曖昧である。本システムでは、実際にユーザによってスコアづけされた修飾語にあった映像エフェクトを表示できる。これにより、母国語が違うコミュニケーションにおいても複数のイメージを持つ修飾語を用いて視覚的に相手に伝えることができる。また、SNSのコミュニケーションの拡張にも利用できると考えている。現在の画像を用いたSNSへの投稿では、人の顔を認識し仮面をつけるなどの加工技術が多く見られる。本研究で得られるデータを利用することでものを加工して投稿するなど今までとは異なるコミュニケーションが可能になると考える。

謝辞

本研究を進めるにあたり，多大なるご指導下さいました指導教員の角康之先生，角薫先生に深く感謝致します．お忙しい中にもかかわらず，本研究について多くの助言を下さいました副査の寺井あすか先生に深く感謝いたします．角康之研究室，角薫研究室の各位には研究生生活にあたり日頃より有益なご討論ご助言いただき，深く感謝いたします．

発表・採録実績

発表

- [I] 平井彰悟, 角薫: 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, マルチメディア, 分散, 協調とモバイル (DICOMO2017) シンポジウム, 情報処理学会 (2017, 06).
- [II] 平井彰悟, 角薫: ことばを映像エフェクトに変換するシステム, 日本デザイン学会 秋季企画大会, 日本デザイン学会 (2017.10). (ポスター発表)
- [III] 平井彰悟, 角薫: 形容詞とリンクする映像エフェクト辞書構築のための GWAP の研究, 人文科学とコンピュータ研究会, 情報処理学会研究報告, 情報処理学会 (2018,01)
- [IV] 平井彰悟, 角薫: 「GWAP による映像エフェクト辞書構築手法」第 50 回情報処理学会エンタテインメントコンピューティング研究会, 情報処理学会研究報告, 情報処理学会 (2018,12)

学術論文等 (査読付き)

- [I] S. Hirai, K. Sumi, Visual-Effect Dictionary for Converting Words into Visual Images, International Conference on Entertainment Computing(ICEC), 2017.09.
- [II] S. Hirai, K. Sumi, Collecting Visual Effect Linked Data using GWAP, The 12th International Conference on E-learning and Games (Edutainment 2018), 2018.06. unpublished.
- [III] S. Hirai, K. Sumi, A Game with a Purpose to Collect Visual Effect Linked Data from Players, IEEE International Conference on Systems, Man, and Cybernetics (SMC 2018), pp.1473-1478,(2018,10).

参考文献

- [1] S. Hirai, K. Sumi, Visual-Effect Dictionary for Converting Words into Visual Images, International Conference on Entertainment Computing(ICEC), 2017.09.
- [2] D. Joshi, J. Z. Wang, and J. Li, “The story picturing engine: finding elite images to illustrate a story using mutual reinforcement,” in Proceedings of the 6th ACM SIGMM international workshop on Multimedia information retrieval, pp. 119126, 2004.
- [3] D. Joshi, J. Z. Wang, and J. Li, “The Story Picturing Engine—a system for automatic text illustration,” ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), vol. 2, no. 1, pp. 6889, 2006, DOI:10.1145/1126004.1126008.
- [4] X. Zhu, A. B. Goldberg, M. Eldawy, C. R. Dyer, and B. Strock, “A text-to-picture synthesis system for augmenting communication,” in AAAI, vol. 7, pp. 15901595, 2007.
- [5] G. Adorni, M. D. Manzo, and G. Ferrari, “Natural language input for scene generation,” in Proceedings of the first conference on European chapter of the Association for Computational Linguistics, pp. 175182, 1983.
- [6] S. R. Clay and J. Wilhelms, Put: Language-based interactive manipulation of objects. IEEE Comput, Graphics vol. 16, no. 2, pp. 3139, 1996, DOI:10.1109/38.486678.
- [7] B. Coyne and R. Sproat, “WordsEye: an automatic text-to-scene conversion system,” in Proceedings of the 28th annual conference on Computer graphics and interactive techniques, pp. 487496, 2001.
- [8] C. Spika, K. Schwarz, H. Dammertz, and H. P. A. Lensch, AVDT-Automatic Visualization of Descriptive Texts, in VMV, pp. 129136, 2011.
- [9] A. Chang, M. Savva, and C. Manning, “Interactive learning of spatial knowledge for text to 3D scene generation,” in Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces, pp. 1421, 2014.
- [10] A. Chang, M. Savva, and C. Manning, “Semantic parsing for text to 3D scene generation,” in Proceedings of the ACL 2014 Workshop on Semantic Parsing, pp. 1721, 2014.

- [11] A. Chang, M. Savva, and C. D. Manning, “ Learning spatial knowledge for text to 3D scene generation,” in Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 2028-2038, 2014.
- [12] Winograd, T.: Understanding Natural Language, Academic Press, 1972.
- [13] H. Tanaka, T. Tokunaga, and Y. Shinyama, Animated agents capable of understanding natural language and performing actions, in Life-Like Characters, Springer, pp. 429-443, 2004.
- [14] K. Sumi, “ Anime de Blog: Animation CGM for Content Distribution ”, Proc. of International Conference on Advances in Computer Entertainment Technology (ACE2008), pp.187-190, 2008.
- [15] K. Sumi, Animation-based Interactive Storytelling System ”, Interactive Storytelling, LNCS 5334, Springer Lecture Note in Computer Science, pp.48-50, 2008.
- [16] K. 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, Feb 2009.
- [17] L. Von Ahn. Games with a purpose. IEEE Computer, 39(6):pp. 92-94, 2006.
- [18] L. Von Ahn, L. Dabbish, ”Designing games with a purpose”, Communications of the ACM, Vol.51, pp.58-67, 2008.
- [19] omcs, “ JAPANESE OPEN MIND COMMON SENSE ”, available from<<http://omcs.jp> >, 2010. (accessed 2018-02-14) (in Japanese).
- [20] R. Speer and C. Havasi, “ Representing General Relational Knowledge in Concept-Net 5, ” in LREC, pp. 3679-3686, 2012.
- [21] Douglas B. Lenat, CYC: a large-scale investment in knowledge infrastructure, Communications of the ACM, Volume 38 Issue 11, ACM, Inc., 32-38, 1995.
- [22] P. Singh, T. Lin, E. T. Mueller, G. Lim, T. Perkins, and W. L. Zhu, “ Open Mind Common Sense: Knowledge acquisition from the general public,” in OTM Confederated International Conferences” On the Move to Meaningful Internet Systems”, pp. 1223-1237, 2002.
- [23] H. Liu and P. Singh, ”MAKEBELIEVE: using commonsense to generate stories”, Proceedings of the Eighteenth National Conference on Artificial Intelligence, AAAI Press, pp 957—958, 2002.
- [24] T. Stocky, A. Faaborg and H. Lieberman, ”A Commonsense approach to predictive text entry”, Proceedings of Conference on Human Factors in Computing Systems, Vienna, Austria 2004.

- [25] Princeton University "About WordNet." WordNet. Princeton University. 2010. available from<<https://wordnet.princeton.edu> >(accessed 2017-12-19)
- [26] Hayakawa et al., Classification of Japanese texture terms, Journal of Texture Studies, 44, 140-159 (2013)

目 次

3.1 システムの流れ	5
3.2 擬態語と名詞を入力する画面	5
3.3 ○×ゲームの画面	6
3.4 ゲームエフェクト	6
3.5 光グループ	6
3.6 光グループ2	6
3.7 水・冷氣グループ	7
3.8 湯気グループ	7
3.9 炎・爆発グループ	7
3.10 YesNo ゲームの画面	7
3.11 常識知識記入画面	8
3.12 映像エフェクト選択画面	8
3.13 ストーリー画面	9
4.1 構造化データの一部	10
5.1 1周目の構造化データ	12
5.2 3周目の構造化データ	13
6.1 餃子と結びついた常識知識	17
6.2 修飾語「ジューシーな」と常識知識の結びつき	18
6.3 修飾語「しゅわしゅわした」と常識知識の結びつき	20
6.4 修飾語「しっとりした」と常識知識の結びつき	21
6.5 修飾語「みずみずしい」と常識知識の結びつき	22
6.6 修飾語「ほっこりした」と常識知識の結びつき	23
7.1 映像エフェクト辞書画面	25
7.2 エフェクト37と「料理」の組み合わせ	26
7.3 エフェクト7と「野菜」の組み合わせ	26
7.4 エフェクト10と「果物」の組み合わせ	26
7.5 エフェクト32と「飲み物」の組み合わせ	26
7.6 エフェクト41と「元気が出そう」の組み合わせ	26
7.7 エフェクト32と「冷たい」の組み合わせ	26
7.8 エフェクト7と「果物」の組み合わせ	27
7.9 エフェクト37と「オープンで焼かれた食品」の組み合わせ	27
7.10 マイナス群と5点以上の群の質問の平均点と標準偏差	29

表 目 次

5.1	自由記述で得られた回答	14
6.1	餃子と常識知識のスコア	17
6.2	「ジューシーな」の時の常識知識と映像エフェクトのスコア	19
6.3	「しゅわしゅわした」ときの常識知識と映像エフェクトのスコア	20
6.4	「しっとりした」ときの常識知識と映像エフェクトのスコア	21
6.5	「みずみずしい」のときの常識知識と映像エフェクトのスコア	22
6.6	「ほっこりした」のときの常識知識と映像エフェクトのスコア	23
7.1	質問内容と質問番号の組み合わせ	28
7.2	マイナスの群だけで比較した有意確率を示した表	29
7.3	5点以上の群だけで比較した有意確率を示した表	30
7.4	マイナスの群と5点以上の群を比較した有意確率を示した表	30