

# 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

平井彰悟<sup>†1</sup> 角薫<sup>†2</sup>

**概要**：本研究では形容詞に対応した映像エフェクトの辞書を構築するために、構造化されたデータを収集する GWAP(Game with a Purpose)を開発した。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するゲームのことであり、本システムでは、ユーザに名詞(例:りんご)を入力してもらい、それが持つ特徴を答えてもらい、形容詞(例:おいしい)を示す映像エフェクトと組み合わせたオブジェクトを見てもらい、それが形容詞に合った適切な映像エフェクトであるかをユーザに回答してもらうことでデータを更新する。同じ形容詞の表現であってもオブジェクトによって表現する映像エフェクトは異なる。この構造化されたデータは、形容詞と映像エフェクトとを対象の物体の種類に応じて紐づけるものとして利用される。構造化された当初のデータと本システム利用後のデータ構造を比較することで検証した。

**キーワード**：形容詞, 名詞, 常識知識, 映像エフェクト

## GWAP for Constructing Visual Effect Dictionary Linked with Adjective

SHOGO HIRAI<sup>†1</sup> KAORU SUMI<sup>†2</sup>

**Abstract**: In this research, we developed GWAP (Game with a Purpose) which collects structured data in order to construct a dictionary of image effects corresponding to adjectives. GWAP is a game to achieve some purpose as a byproduct of game play. The system prompts the user to input a noun (e.g. apple), ask the user to answer the characteristic that it has, and watch the object combined with the visual effect showing the adjective (e.g. delicious) to the user. The data is updated by having the user answer whether it is an appropriate visual effect matching the adjective. Even with the same adjective expression, the image effect expressed by the object is different. This structured data is used to link adjectives and visual effects according to the type of target object. It was verified by comparing the structured initial data with the data structure after using this system.

**Keywords**: Adjective, Noun, Commonsense knowledge, Visual effect

### 1. はじめに

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、形容詞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。

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、対象とする名詞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。

このように 同じ形容詞であっても、想起させるイメージが異なる。平井らが行った印象評価実験により、映像エフェクト「湯気」と「温かい食べ物」の組み合わせに相関があることが示唆された[1]。図1は映像エフェクトと常識知識の適切な組み合わせを示した図である。本研究は、「湯気, 温かい食べ物」, 「光, 人工物」の組み合わせのように形容詞と常識知識の関係に着目した。そのため、コンピュータが学習する際、人間が暗黙的に持つ常識知識と呼ばれる知識を収集して、知的に学習するアルゴリズムが必要である。

常識知識を蓄積したデータベースに ConceptNet[2]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が、ConceptNet を利用したゲームはいくつか存在する。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」を開発した[3]。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

<sup>†1</sup> 公立はこだて未来大学 システム情報科学研究科  
Future University Hakodate

<sup>†2</sup> 公立はこだて未来大学  
Future University Hakodate

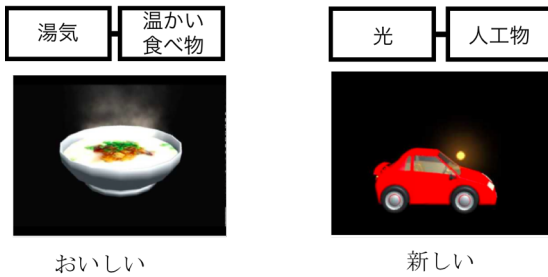


図 1: 常識知識と映像エフェクトの適切な組み合わせ

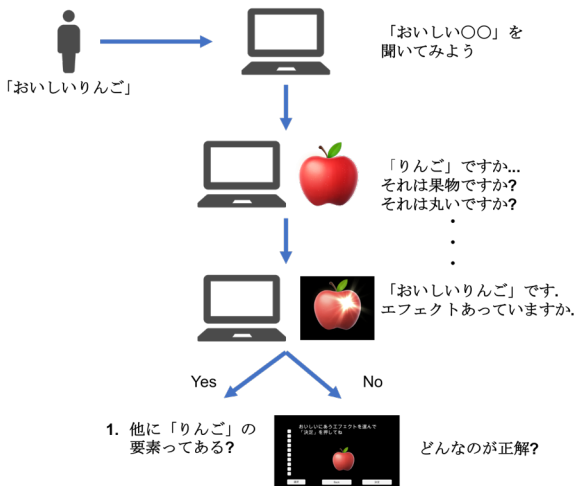


図 2: システムの流れ

ことで、ConceptNet を拡張することが目的だが、本研究では、形容詞とリンクする映像エフェクトを収集する点が異なる。我々は、未知の常識知識を獲得するため、映像エフェクトを用いた GWAP を開発した。この構造化されたデータを用いることで、抽象的な表現を持つ日本語をより明確に視覚化することが可能になると考える。

GWAP の研究は盛んに研究されており、大谷らは、コモンセンス知識を獲得する連想ゲームを開発した[4]。このゲームでは、既知の知識の一部をプレイヤーに推測してもらうことで新しい知識を獲得する。例えば、システムがキーワード「携帯電話」を正解として、「これは家電の一種です」というヒントを提示した場合、ユーザが「テレビ」と間違えて答えたとする。ゲームでは不正解になるが、これにより、「テレビは家電の一種である」という知識を得ることができる。このゲームの目的は、ConceptNet から文章を生成し、新たな知識を収集することである。GWAPを使ってオントロジを構築する手法として、安永らは、連想ゲームを提案している[5]。連想ゲーム、関係ゲーム、OXゲームをユーザにプレイしてもらうことで、概念、概念間の関係、不適切なものを除去することができる。この手法では、「筑波大といえど?」という名詞のみからオントロジ

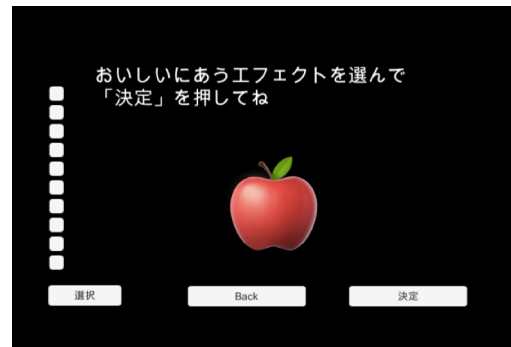


図 3: 映像エフェクト選択画面

を構築している。本研究では形容詞とリンクする映像エフェクトの構造を作成する点と、新たな知識を収集する点が異なる。

また言葉を視覚化する研究として、WordsEyeがある[6]。WordsEyeでは、文章を3Dオブジェクトで可視化することができるが、「象」といった1つの単語に1つの3Dオブジェクトといったように1対1の関係である。文字入力することで3Dアニメが表示されるAnime de Blogがある[7][8][9]。Anime de Blogはキャラクタ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。Anime de Blogも1つの名詞に対して、1つのオブジェクトという1対1の関係である。本研究では、1つの形容詞に対して、名詞が変化することで複数の映像エフェクトで表現する1対多の関係である点が異なる。

## 2. 提案するシステム

図2はシステムの流れである。本システムでは、ユーザに言葉のイメージを教えてくれる「エフェクトゲーム」を行ってもらう。ゲームの初めに、システムが「おいしい○を入力して」と提示する。ユーザは○○に当たる名詞を入力する。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画像を映しながら、○×の質問を提示する。名詞の画像は、絵文字の中から食べ物を用いている。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3種類である。

- (1) WordNetの一階層上の上位語
- (2) 映像エフェクトと適切な組み合わせを示した「温かい」という要素。今回は、対義語である「冷たい」も用いる。
- (3) 形状、色
- (4) ユーザから得られる新しい常識知識

(1)の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[10]。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)は、平井ら[1]が行った印象評価実験により映像エフェクトと関連が示唆された「温かい」という要素を用いる。その対義語である「冷たい」について

表 1: 獲得されるデータ

○×形式	食べ物	常識知識	Score
Yes/No形式	食べ物	映像 エフェクト	Score

も問う。これは ConceptNet や WordNet では得られない概念である。(3)は、人間の持つ視覚的な情報を常識知識として用いる。今回は、形状や色彩を初期データとして利用する。(4)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。映像エフェクトの種類は 10 種類である。選定方法は、「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を 10 種類選び、その映像エフェクトがおいしそうに見えるかを 1 人に見てもらい、判断してもらったデータを用いている。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。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る?」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「どんなのが正解?」とユーザに質問を行う。そして、図 3 で示す映像エフェクト選択画面を表示する。この画面で、ユーザが適切だと思う映像エフェクトを選択してもらい、ゲームを終了する。このゲームで、獲得されるデータは、以下である。

- **○×形式:** 名詞に対して、常識知識が適切であるか
- **Yes/No 形式:** エフェクトと名詞の組み合わせが適切であるか
- **Yes の場合:** 名詞が持つ未知の常識知識と一番影響のあると感じた常識知識を獲得
- **No の場合:** ユーザが考える新たな映像エフェクトの組み合わせを取得

それぞれの形式から獲得されるデータを表 1 で示す。ユーザが解答するごとに、Score が 1 ずつ加算される。○×形式では、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。Yes/No を解答後は、システムと対話しながら未知の常識知識や、新たな映像エフェクトの組み合わせをユーザから調査する。このシステムの開発環境はゲームエンジンである Unity を使用した。動的な映像エフェクトは、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

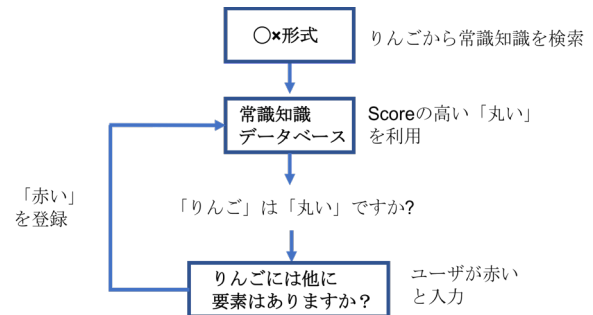


図 4: ○×で問う質問の作成方法

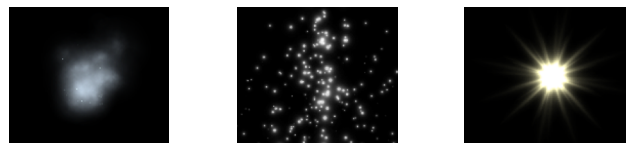


図 5: 映像エフェクトの例

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、上記で述べた獲得するシステム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、主人公がボスに負けそうな場面を設定した。ユーザが「おいしい食べ物」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。多くのユーザと同じ答えを解答すると「おいしい食べ物」が主人公にプレゼントされる。同じ解答をした他のユーザが多いほど、主人公の力が上昇する。一致する割合が一定数を越えた場合、ボスを倒すことができる。

### 3. 適用

○×形式に用いる質問の作成方法を図 4 に示す。このゲームは、ユーザが利用するたびに○×の選択肢が変化していく。ユーザが解答した内容を Score として保存しているため、Score の高いものからユーザに質問をしていくことができる。例えば、ユーザが「りんご」を入力した際、りんごから常識知識を検索する。「丸い」という常識知識が最も Score が高い場合、「りんご」は「丸い」ですか。といった質問を行う。質問の形式は以下になる。

- ・りんごは果物ですか ○ ×
- ・りんごは赤いですか ○ ×
- ・りんごは丸いですか ○ ×

質問に用いられる常識知識のデータベースは、2 章で述べた常識知識の(1)と(2),(3)が初期データとして登録されてい

る。それらを複数回ユーザに○×解答させることで、システムは最も解答が一致する映像エフェクトを選択することができる。ユーザが他の要素として「赤い」と入力した場合、「赤い」というデータを常識知識のデータベースに保存する。このように、ユーザから得た未知の常識知識も獲得されているため、次のユーザに○×の問題として利用することができる。また、映像エフェクトが適切であるかも問うため、名詞によって異なる映像エフェクトを付与することができる。映像エフェクトの例として、図 5 を示す。左から冷気を示す映像エフェクト、塩をイメージした映像エフェクト、光を示す映像エフェクトである。例えば、「おいしいりんご」に対しては、「光」の映像エフェクト、「おいしいアイスクリーム」に対しては、「冷気」の映像エフェクトを付与するなどユーザがそれぞれの名詞に対して、適切だと答えた映像エフェクトを付与することができる。これにより、利用者が増えるほど、同じ形容詞でも名詞によって異なるイメージを持つ形容詞を適切な映像エフェクトで可視化することができる。このように選択された映像エフェクトが適切であるかをユーザに Yes/ No で答えてもらうことで、Yes の選択が多い映像エフェクトが名詞に対して適切な映像エフェクトを示していることがわかる。このゲームを被験者に利用してもらい、通常のゲームと同様にストーリーを楽しんでもらい、ユーザのモチベーションを保ちながら、形容詞を体系化するデータを蓄積していく。蓄積されたデータをもとに、形容詞と名詞を入力するだけで、形容詞を可視化する映像エフェクト辞書として登録する。

#### 4. 考察

本稿では、未知の常識知識を獲得するための映像エフェクトを用いた GWAP を紹介した。今後は、名詞の種類を拡張して行くことで、形容詞によって想起させるイメージと言葉の構造を明確にしていく予定である。

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。消費者によって形容詞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。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伝えたい表現を視覚的に相手に伝えることができる。また、デザインに関わる業界にも利用できると考えている。今までのデザインでは、コーヒーなどを図示化する際、その人が持つ一般常識をもとに、コーヒーに対して、湯気をつけるなど美味しく見せる工夫を施していた。このゲームから得られる未知の常識知識を利用することで、コーヒーやおかゆなど、名詞の分類上、「食べ物」という共通点しか得られなかったものに対して、「温かい」といった直接結びついた常識知識からデザインを支援することができる。

#### 参考文献

- [1] S. Hirai, K. Sumi, Visual-Effect Dictionary for Converting Words into Visual Images, International Conference on Entertainment Computing(ICEC), 2017.9.
- [2] R. Speer and C. Havasi, "Representing General Relational Knowledge in ConceptNet 5," in LREC, 2012, pp. 3679-3686.
- [3] 中原和洋, 山田茂雄, 日本でのコモンセンス知識獲得を目的とした Web ゲームの開発と評価. UNISYS TECHNOLOGY REVIEW, 第 107 号, pp 13-23, 2011.
- [4] 大谷直樹, 河原大輔, 黒橋禎夫ほか, 連想ゲームによるコモンセンス知識の獲得, 言語処理学会, 第 22 回年次大会 発表論文集, pp. 897-900, 2016.
- [5] 安永ゆい, 望月祥司, 森嶋厚行, GWAP によるオントロジ構築手法の提案. 情報処理学会第 73 回, pp.765-766, 2011.
- [6] 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.
- [7] 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).
- [8] 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).
- [9] 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).
- [10] Princeton University, "About WordNet", available from<<https://wordnet.princeton.edu>> (accessed 2017-12-19).