

# GWAP による映像エフェクト辞書構築手法

平井彰悟<sup>†1</sup> 角薫<sup>†2</sup>

**概要:** データベースを用いることにより言葉を 3DCG に変換し視覚化するシステムの研究が行われている。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の辞書システムを構築した。本システムでは、GWAP(Game With A Purpose)を用いて特定の修飾語におけるオブジェクトに対応する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。過去の研究により、この構造化データは対象のオブジェクトの常識知識に関連していることが示された。GWAP ゲームでは、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。本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。実験により、獲得した構造化データの有用性とユーザが本ゲームを楽しみながら自然に利用できたかについて評価した。

**キーワード:** 擬態語, 形容詞, 名詞, 常識知識, 映像エフェクト

## Visual Effect Dictionary Construction Method using GWAP

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

**Abstract:** Research on a system for converting words into 3DCG and visualizing it by using a database has been conducted. In this research, we constructed a dictionary system of visual effects corresponding to modifiers such as adjectives and onomatopoeia. In this system, structured data of visual effect corresponding to an object in a specific modifier can be acquired using GWAP (Game With A Purpose). Previous studies have shown that this structured data is related to common sense knowledge of objects to be visualized. In GWAP games, pictograms are used for objects, images of modifiers are displayed with visual effects, and data is acquired while users enjoy games by voting by users. In this system, the more the user uses it, the more you can learn the relationship between modifier, object and visual effect. Experiments evaluated the usefulness of the acquired structured data and whether the user was able to use it naturally while enjoying the game.

**Keywords:** Onomatopoeia, 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。

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。著者らが行った印象評

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。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。「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

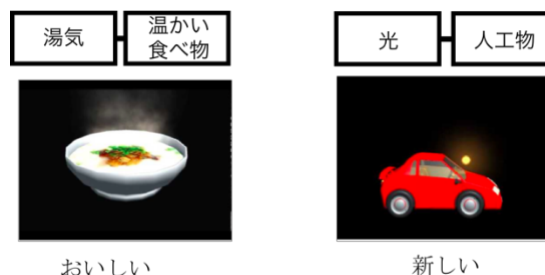


図1 常識知識と映像エフェクトの適切な組み合わせ

<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

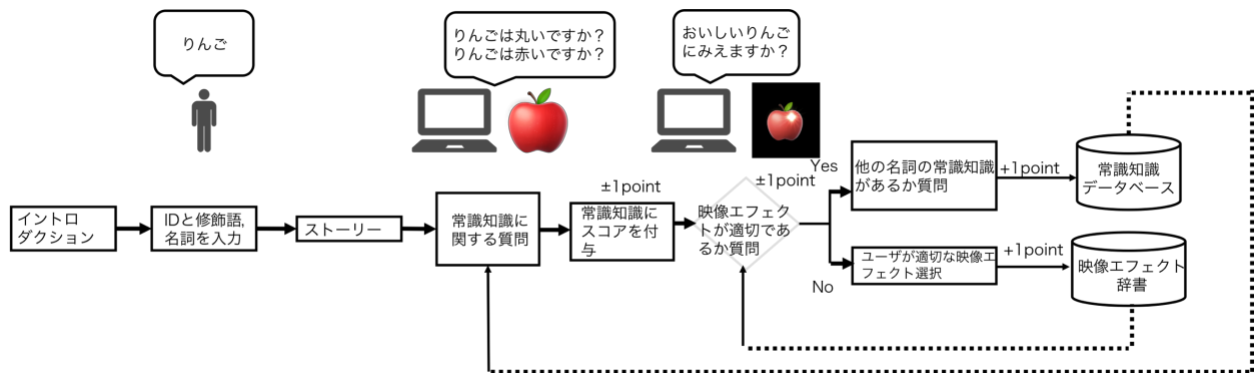


図 2 システムの流れ

あった。そこで、GWAP による簡単なゲームインターフェイスで常識知識を収集する「Effect Game」システムを開発した。「Effect Game」は、ユーザからオブジェクトに関する新たな常識知識を獲得し、映像エフェクトとの結びつきを獲得することができる。今までの研究では、「おいしい」という形容詞を対象に本システムを利用した実験を行ない、常識知識と映像エフェクトで構造化データを獲得することができた[2]。そして、実験によりユーザに複数回試してもらったデータでは、同じ形容詞でも対象が変わることで異なる映像エフェクトと結びつくことが示された[3]。これらの実験から修飾語と常識知識の結びつきを構造化されたデータとして獲得することで、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ことがわかった。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。

## 2. 関連研究

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はいくつかある。Open Mind Common Sense[4]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[5]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[6]。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 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。

言葉を視覚化する研究として、WordsEye がある[7]。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。例えば、「黄色い象が家の上にいる」といった文章が入力されると、家の 3D オブジェクトの上に黄色い象が出力される。また、話し言葉を視覚化する研究として、

傀儡がある[8]。傀儡では、「その」や「もうちょっと」といった話し言葉を視覚化していた。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[9][10][11]。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った。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。

## 3. Effect Game

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。今回のシステムでは、自分自身が説明せずにユーザー一人一人が説明を見てプレイできるようシステムを修正した。図 2 はシステムの流れである。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ど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 は形容詞と名詞を入力する画面である。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、



図 3 形容詞と名詞入力画面

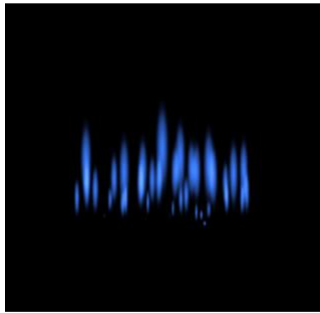


図 4 ゲームエフェクトの例



図 5 光グループの例

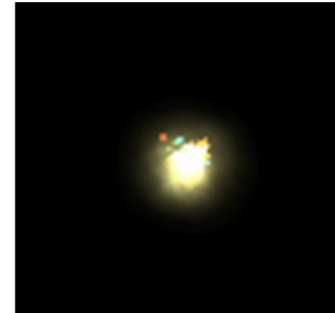


図 6 光グループ 2 の例

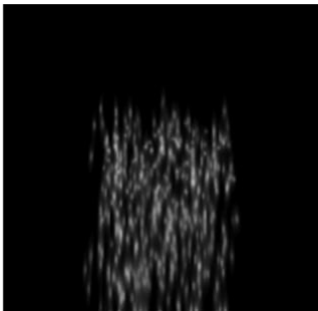


図 7 水・冷氣グループの例

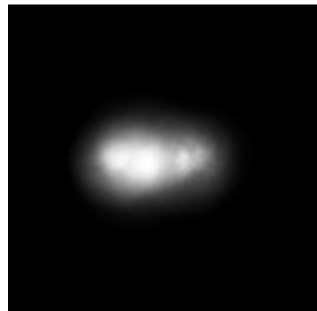


図 8 湯気グループの例

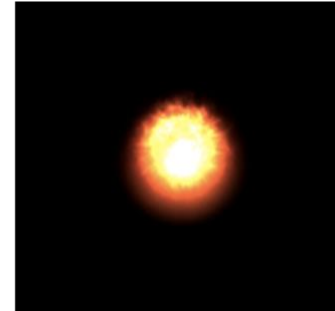


図 9 爆発・炎グループの例

絵文字の中から食べ物を用いている。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[12]。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 4, 図 5, 図 6, 図 7, 図 8, 図 9 である。それぞれの分類と種類は以下である。

- ・ゲームエフェクト : 6 種類
- ・光グループ : 8 種類
- ・光グループ 2 : 7 種類
- ・水・冷氣グループ : 8 種類
- ・湯気グループ : 6 種類
- ・爆発・炎グループ : 5 種類

映像エフェクト選択画面を図 10 に示す。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。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、以

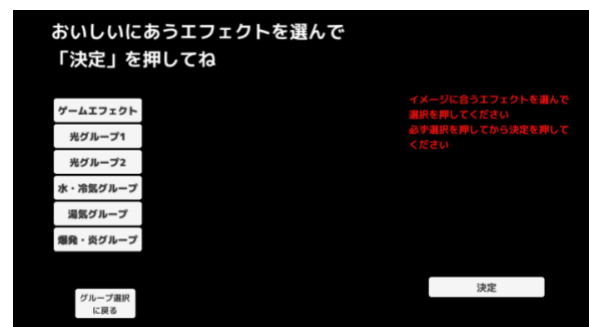


図 10 映像エフェクト選択画面

表 1 獲得されるデータ

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

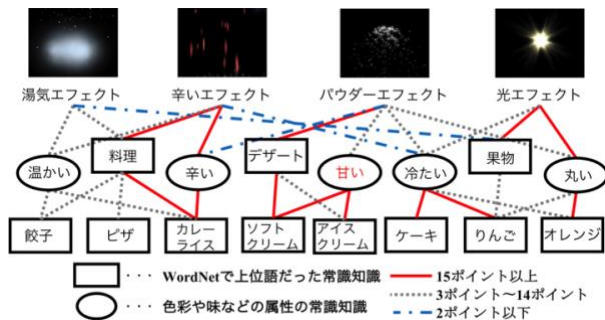


図 11：形容詞「おいしい」の時の構造化データ

下である。

- ○×ゲーム：名詞に対して、常識知識が適切であるか
- Yes/No ゲーム：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
- Yes の場合：名詞が持つ未知の常識知識を獲得
- No の場合：ユーザが考える新たな映像エフェクトと名詞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。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、主人公がボスに負けそうな場面を設定した。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超えた場合、ボスを倒すことができる。

#### 4. 実験「形容詞」

映像エフェクトと名詞の常識知識との関係の変化を調査することを目的として、本システムを 8 人の被験者に利用してもらった[3]。この実験では、形容詞「おいしい」の

表 2 前回の実験で得られた常識知識

獲得した常識知識
三角
甘い
軽い
茶色
熱い
柔らかい
酸っぱい
安い
なめらか

みを対象に 8 つの名詞と組み合わせて 1 日 8 回ゲームをプレイしてもらった。被験者には、8 回のゲームを 3 日間プレイしてもらった。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」である。3 日間プレイして得られた構造化データの一部を図 11 に示す。図中のスコアは長い線が 15 点以上、点線が 3~14 点、点線と単線の繰り返しは 2 点以下を表している。また、映像エフェクトにそれぞれ名称が付いているが図の表記をわかりやすくするためであり、被験者には提示していない。Yes/No ゲームで得られた結果として、「辛い」エフェクトと常識知識「料理」、「辛い」エフェクトと常識知識「辛い」、「パウダー」エフェクトと常識知識「デザート」、「光」エフェクトと常識知識「果物」、「光」エフェクトと常識知識「丸い」のスコアが 15 点以上だった。○×ゲームでは、ユーザから獲得した「甘い」という常識知識がソフトクリームやアイスクリームと 15 点以上の結びつきを獲得した。

#### 5. 考察

本稿では、構造化データと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「Effect Game」について議論した。形容詞「おいしい」を対象とした実験では、「辛い」エフェクトと常識知識「料理」、「パウダー」エフェクトと常識知識「デザート」、「光」エフェクトと常識知識「果物」のように WordNet から獲得した上位語の意味を持つ常識知識とそれぞれ 15 点以上のスコアを獲得していた。従って、異なるイメージを持つ形容詞を対象の物体が持っている常識知識と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。また、ユーザから得られた常識知識として「甘い」があった。この常識知識の結びつきは 2 点以下の結びつきが存在しなかった。よって、ユーザから得られたデータであっても適切な組み合わせであれば、スコアが上がっていく傾向にあることも示唆された。

このシステムを修飾語で利用した際に得られる構造化データについて述べる。このシステムを「おいしい」という

表 3：実験で行う修飾語と名詞の組み合わせ

	みずみずしい	ほくほく	ふわふわ	ほっこり	しゅわしゅわ	あつあつ	こんがりした	ひんやりした	しっとりした	ジューシーな
じゃがいも	/	o	/	o	/	o	o	o	o	/
くり	/	o	/	o	/	o	o	o	o	/
ケーキ	/	/	o	/	/	o	/	o	o	/
パン	/	/	o	/	/	o	o	o	o	/
餃子	/	o	o	/	/	o	o	o	o	o
ビール	/	/	/	/	o	/	/	o	/	/
オレンジ	o	/	/	/	o	/	/	o	o	o
りんご	o	/	/	/	/	o	o	o	o	o
ぶどう	o	/	/	/	o	/	/	o	/	o
カレーライス	/	/	/	/	/	o	o	o	/	/
クッキー	/	/	/	/	/	o	o	/	/	/
ソフトクリーム	/	/	/	/	/	/	/	o	/	/
にんじん	o	/	/	/	/	/	/	/	/	/
お茶	/	/	/	/	/	/	/	o	/	/
ジュース	o	/	/	/	o	/	/	o	/	o

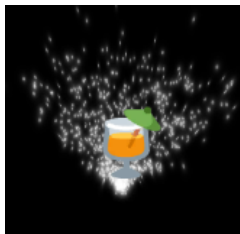


図 12 しゅわしゅわしたジュース

形容詞で利用した時、複数回システムを利用してもらうことで、対象のオブジェクトの常識知識ごとに適切な映像エフェクトとの組み合わせが獲得された。次の実験では、「おいしい」の構造化データを獲得した際の常識知識を適応して実験を開始する。表 2 が前回の実験で得られた常識知識である。このようにこのシステムで獲得されたデータは常に蓄積され続け、○×ゲームの質問に活用することができる。また、修飾語で実験する際に利用する名詞は 15 種類、修飾語は 10 種類である。修飾語の選定方法や組み合わせは、日本語テクスチャー用語体系[13][14]を参考にしたものと視覚的に表現できそうな修飾語を選択した。被験者に行なってもらう組み合わせを表 3 に示す。表 3 において○がついている組み合わせで実験を行い、明らかにあり得ない組み合わせを排除して実験を行う。被験者は 120 人を想定しており、1 つの組み合わせあたり 20 以上のデータを獲得する予定である。この実験で得られたデータを利用することで例えば、図 12 のような「しゅわしゅわしたジュース」といった修飾語も視覚的に表現することが可能になる。

## 6. 結言

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手法を提案した。「おいしい」という形容詞を対象に実験を行なった結果、常識知識によって異なる映

像エフェクトと結びつく構造化データを獲得することができた。今後の課題として修飾語の視覚化があげられる。インターネットを通じてより多くのユーザから本システムを利用してもらい、形容詞と同様に実験を行うことで修飾語を視覚化するための構造化データを獲得していく。

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。

## 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] S. Hirai, K. Sumi, Collecting Visual Effect Linked Data using GWAP, The 12th International Conference on E-learning and Games (Edutainment 2018), (2018, 6). (unpublished)
- [3] 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).
- [4] 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).
- [5] R. Speer and C. Havasi, "Representing General Relational Knowledge in ConceptNet 5," in LREC, pp. 3679-3686, (2012).
- [6] 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.
- [7] 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).

[8] Hozumi T., Takenobu T., Yusuke S.: Animated Agents Capable of Understanding Natural Language and Performing Actions, Life-Like Characters, Helmet Prendinger and Mitsuru Ishizuka Eds, pp.163-187, Springer, (2004).

[9] K. 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).

[10] K. 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).

[11] 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, (2009. 2).

[12] Princeton University, "About WordNet", available from<<https://wordnet.princeton.edu>> (accessed 2017-12-19)

[13] Hayakawa et al., Classification of Japanese texture terms, Journal of Texture Studies, 44, 140-159 (2013).

[14]早川文代：日本語テクスチャー用語体系，入手先  
<[http://www.naro.affrc.go.jp/archive/nfri/introduction/files/yougotaikai\\_1.1.pdf](http://www.naro.affrc.go.jp/archive/nfri/introduction/files/yougotaikai_1.1.pdf)>(accessed 2018-11-29)