

# かわいいカフェの検索システム「CafePalette」の研究

橋口恭子<sup>†1</sup> 木川真里<sup>†2</sup> 小川克彦<sup>†3</sup>

CafePalette は、かわいい雰囲気のカフェを探す事のできる検索サービスである。近年ブログ記事に絵文字や顔文字、ひらがなを多用し行間を大きくあけ、文章と共に写真で記事を飾るかわいいブログ記事が数多くみられるようになった。女性は顔に化粧をする行為と同様にブログにも「絵文字」「顔文字」「行間」「ひらがな」などの化粧を施し互いにコミュニケーションを図っている。それらの記事の書き方の違いは記事の書き手の雰囲気や年齢を分ける指標の一つである。本研究では SVM を用いてブログ記事の見た目の雰囲気から「かわいい」「きれい」「おもしろ」に3種類に判定する。そして、雰囲気の違う3種類のブロガーが紹介するかわいいカフェを探す CafePalette を提案し実装および実験を行った結果を報告する。

## CafePalette: A System For Finding Kawaii Cafes

KYOKO HASHIGUCHI<sup>†1</sup> MARI KIKAWA<sup>†2</sup>  
KATSUHIKO OGAWA<sup>†3</sup>

CafePalette is a search system based on three types of kawaii feeling. “Kawaii blog articles”, decorated with emoji and smileys, lots of hiragana characters and whitespace, and rich photos, have recently been growing in number. Similar to when women apply makeup to their face, they apply makeup, such as “emoji”, “smileys”, “whitespace” and “hiragana”, to their blogs, through which they then participate in communication. Differences in these characteristics or writing style differentiates feeling of the blogs and age of the bloggers. In this paper we propose CafePalette, which can find three types of kawaii introduced by bloggers, implement the system and report the result of an experimental evaluation.

### 1. はじめに

近年ブログ記事に絵文字や顔文字、ひらがなを多用し、行間を大きく空け文章と共に写真で記事を飾るかわいいブログ記事が数多く見られるようになった。女性は顔に化粧をする行為と同様にブログにも「絵文字」「顔文字」「行間」「ひらがな」などの化粧を施し、互いにコミュニケーションを図っている。それらの記事の書き方の違いは記事の書き手の雰囲気や年齢を分ける指標の一つである。これらのブログ記事の書き方の違いに着目し、2011年6月にかわいい検索[1][2]を構築した。

かわいい検索とは、ブログ記事の見た目の情報の元に所望のキーワードのブログ記事を検索する事のできる検索サービスである。2011年6月から NTT レゾナント株式会社の運営するポータルサイト goo ラボ上で実験を行っている[3]。検索対象のブログ記事として読者モデルなどのブロガー1500人の約100万件のブログ記事が登録されている。かわいい検索では図1のようにブログ記事の見た目の違いから「ゆるかわ」「キュート」「きれい」「おもしろ」「まじめ」の5種類のかわいいタイプに分類し、見た目から所望のキーワードを含むブログ記事を検索する事ができる。ユーザ行動の実験の結果から、かわいい検索は特に「食べ物」

「ファッション」「人名」などで利用される事が多かった。またブログ記事上で掲載された飲食店に行ったところオシャレで値段が手頃なかわいいお店が多かった。

そこで、本研究ではリアルな世界と結びつきやすい「飲食店」に特化し、特にかわいいカフェにたどりつく事のできる CafePalette を提案し実装と実験を行った結果を報告する。



図1 かわいい検索のインターフェース

Figure 1 Interface of KawaiiSearch

<sup>†1</sup> 慶應義塾大学 政策・メディア研究科  
Keio University Graduate School of Media and Governance  
<sup>†2</sup> 慶應義塾大学 総合政策学部  
Keio University Faculty of Policy Management  
<sup>†3</sup> 慶應義塾大学 環境情報学部  
Keio University Faculty of Environment and Information Studies

## 2. CafePalette のコンセプト

CafePalette とは、かわいいカフェを探す事を目的とした検索サービスである。これまでかわいいカフェを探す場合カフェを掲載していることりっふ[4]などの雑誌や食べログ、ぐるなびのような WEB サービスが利用されてきた。これらのサービスではエリアや金額などから人気のお店を探す事ができる。しかし、女子会やデートなどで利用する事を目的としておらず女性が好むかわいいカフェを探したいという要求に対して適合していない。また、女子会プランのような広告も掲載されているが、店側からの意図的な情報提供であり、リアルな女性同士の口コミなどが反映されていない。

そこで本研究では、読者モデルなどの女性たちがブログ記事に掲載した口コミからかわいいカフェを探す事のできる CafePalette を提案する。具体的な利用シーンは、例えば女子会やかわいいカフェ巡り、デートなどで利用するカフェを探す時に CafePalette を利用すると、オシャレでリーズナブルなカフェが見つかる。

## 3. CafePalette のアルゴリズム

### 3.1 かわいいの指標

かわいい検索では「絵文字」「顔文字」「文字数」「単語数」「行間」「画数」「ひらがな」「漢字」「ギザギザ」「画像」「ゆる絵文字」「記号」の見た目の要素となりうる 12 の指標をかわいい特徴点とした。本研究でもこれらの 12 のかわいい特徴点を用いる。実験の結果からかわいいタイプが図 3 のような「かわいい」「きれい」「おもしろ」の 3 分類が適切であると推奨された[5][6]。そこで、かわいい検索の 5 つの分類から代表的な 3 つの分類に絞った。



図 2 3 分類のかわいい視覚特定

Figure 2 Visual Traits of the Three Classes

### 3.2 CafePalette の検索アルゴリズム

CafePalette の見た目の分類は、サポートベクターマシン[7](SVM)を用いて行った。SVM とは、与えられたデータを超平面で正例集合へと分離する際にマージンを最大化する事によって最適な分離超平面を得る学習手法である。

SVM は手書き文字認識、人の顔認識といった高次元の入力属性が必要であるような多くの分類問題において優れた汎化能力を持つ事が明らかになっている。

かわいい見た目の判定で利用する場合、「絵文字」「ギザギザ」「行間」などの特徴点は全てのかawaiiタイプに含まれており線形的に分類する事が難しい。このような特徴点の重なりがある場合、非線形の分類に対応した SVM を用いる方が適切であると考えられる。

### 3.3 教師データの作成と評価結果

SVMの判別性能を検証するための評価実験を行った。

まず、教師データと評価データの作成を行った。実験用のwebサイト上で1400記事を見た目を元に「かわいい」「きれい」「おもしろ」の3分類を行った。

これらの実験で得られた1400記事のデータのうち分類できなかった記事を除き、500記事を教師データとし600記事の評価データとした。評価データに対して正解率を出したところ表1のようになり、かわいいときれいに対する正解率が高い事がわかった。おもしろについては、12のかawaii特徴点では捉えられず正解率が低かったが、再現率がある程度の値だったのでそのまま用いた。

表 1 サポートベクターマシンの評価結果

Figure 1 Evaluation Tool For Support Vector Machine

|      | 正解率 | 再現率 |
|------|-----|-----|
| かわいい | 89% | 77% |
| きれい  | 97% | 79% |
| おもしろ | 2%  | 40% |
| 全体   | 78% | 78% |

## 4. CafePalette のシステム

### 4.1 インタフェース

CafePalette のインタフェースは図 3 の通りである。かわいい検索の多くのユーザがスマートフォンを利用していた事と、現在地から近いカフェを探す検索要求を考慮しスマートフォンで利用する事を前提とした。下記にインタフェースの詳細を述べる。

#### (1) TOP ページ

「渋谷」「原宿」「代官山」「下北沢」「六本木」の 5 箇所を左右にフリックする事で場所を選択する事ができる。かわいい検索で用いられた読者モデルなどの記事を解析したところ上記の 5 箇所の口コミが多かった。

場所を選択したら、図 3 に記載されているように「MAP から探す」と「写真から探す」のどちらかを選択する。

#### (2) MAP から探す

Café を探す際は、例えば現在地から近い店を探す事もあるため MAP から探す機能を実装した。「MAP から探す」

を選択した場合図3のようにかわいいカフェの場所を探すピンが地図上にマッピングされカフェを選択する事ができる。また、「かわいい」「きれい」「おもしろ」の3種類の雰囲気から選択する事ができる。3種類の雰囲気は図2のブログ記事の見た目の雰囲気である。

### (3) 写真から探す

「写真から探す」では、カフェの写真を選べる事ができる。かわいい検索では、文字や内容よりもビジュアル重視の検索行動が多く見られた。そこでCafePaletteでは検索キーワードから探すのではなく写真で情報提示をする。検索上位に同じカフェの情報が提示される事を回避する為に表示順はランダムとした。

### (4) ショップ詳細画面

「写真から探す」画面の写真か、「MAPから探す」画面のピンをタッチすると詳細画面を見る事ができる。詳細画面には、カフェの画像、金額、ブログのリンク、マップ詳細画面への遷移などのアイコンが掲載されている。金額は、ビジュアル重視にするため財布アイコンで表した。財布1つは、平均予算1000円以下であり2つは1000～3000円3つは3000円以上の店である。また、ブログ記事に記載されている口コミをカフェ選択の参考にするためにブログ記事のリンクを貼った。

### (5) マップ詳細画面

マップ詳細画面ではカフェの位置にピンが表示される。地図は手書きの地図を利用した。

## 4.2 CafePalette のシステム構成

CafePalette のシステム構成は図4の通りである。このよ

うな手法は相良ら[8]の研究でもなされている。相良らは、タウンページを検索キーワードとして、web上から施設に関する情報を取得し地図上にマッピングしている。しかし、これらの手法はカフェに特化していない。特にタウンページを利用すると最新のカフェのデータベースとGPS情報を同時取得する事が難しい。本研究ではGooglePlacesAPIを用いて最新のカフェのデータベースを利用しカフェの判定を行った。その際マッシュアップ方式でブログ記事からカフェ名を抽出し、地図上にマッピングする。詳細を下記に述べる。構成は雰囲気抽出と、カフェ情報抽出の2つにわかれている。

### (1) 雰囲気抽出

まず読者モデルなどの1500ブロガーのブログ記事をクロールする。

次に、雰囲気検索する事を目的として、「かわいい」「きれい」「おもしろ」の3つのタイプにSVMを利用して分類する。

### (2) カフェ情報抽出

読者モデルのブログ記事からカフェ名とカフェの画像、値段、カフェのGPS情報を取得する事を目的としている。

まず、クロールしたブログ記事から自動でカフェ名を抽出する。そのため、下記の①～③の手段でカフェでありそうな言葉と場所が含まれる記事を集める。

①「カフェ、雰囲気、ディナー、ランチ」などカフェと結びつきの大きいと考えられる言葉が含まれている辞書を作る。また、「東京、渋谷、原宿」など場所についての辞書も作成する。場所についての辞書はぐるなびのエリア検索



図3 CafePalette のインターフェース

Figure 3 Interface of CafePalette.



を参考にして作成した。

②辞書に含まれている「カフェ+場所」「雰囲気+場所」などで検索した記事を収集する。それらの記事のうち、「英語文字列」「カタカナ」「カフェ」「cafe」などが含まれる言葉を取得する。「カフェラミル」「monsooncafe」などのカフェ名でありそうな言葉が抽出される。

③「live」や「love」などの言葉が抽出され、カフェとは関係のない記事が取得される事が多かったのでそれらのワードを「ダメワードフィルター」として、検索されないようにフィルターをかける。

次にカフェの判定を下記のように行う。

①カフェ名抽出で抽出したカフェが実際のカフェか判定するために、取得したカフェ名と場所を GoogleplacesAPI に投げる。GoogleplacesAPI は、「カフェ名と場所」を元に店の情報を取得できる WebAPI である。例えば、カフェの名前と場所を入力として利用すると、その場所の「カフェの名前」と「緯度経度」、「写真の URL」などを返す。入力として利用したカフェ名と、GoogleplacesAPI が返すカフェ名が一致すればカフェであると判定を行う。

②しかし、判定したブログで書かれていたカフェ名と GoogleplacesAPI で返すカフェ名が英語と日本語で表記のずれがある。例えば、ブログ記事から取得したカフェでありそな言葉が「スターバックスカフェ」であり、GoogleplacesAPI では「StarbucksCafe」と返ってきた場合カフェ名と判定されない。そこでカタカナとローマ字表記を統一する為に MicrosoftTranslateAPI を用いて表記を合わせた。

③CafePalette で利用する値段の情報は、GooglePlacesAPI では取得できなかったため、食べログの API を利用して値段を取得した。

④最後に CafePalette に必要なデータがそろったので、「カフェ名」「写真 URL」「位置情報」「かわいいタイプ」「値段」「ブログ URL」をデータベースに保存する。

図 5 はこれらのシステムを用いてカフェ判定された渋谷のカフェの一例である。ピンクが「かわいい」黄色が「きれい」虹色が「おもしろ」を表している。渋谷駅周辺から渋谷センター街周辺にかけてかわいいカフェが多い事が見てとれる。またかわいいタイプの特徴としては全体的に大きく変わらないが、特に「かわいい」は駅に近いマークシティ周辺に多く、「きれい」「おもしろ」は青山通り周辺のカフェが多い事が見てとれる。

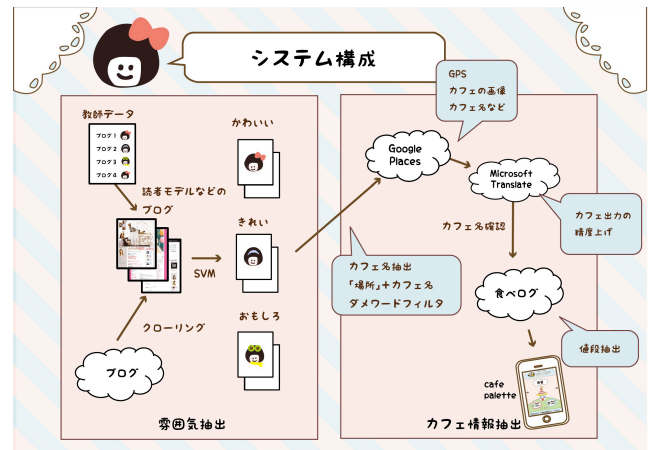


図 4 CafePalette のシステム構成

Figure 4 Overview of CafePalette System.

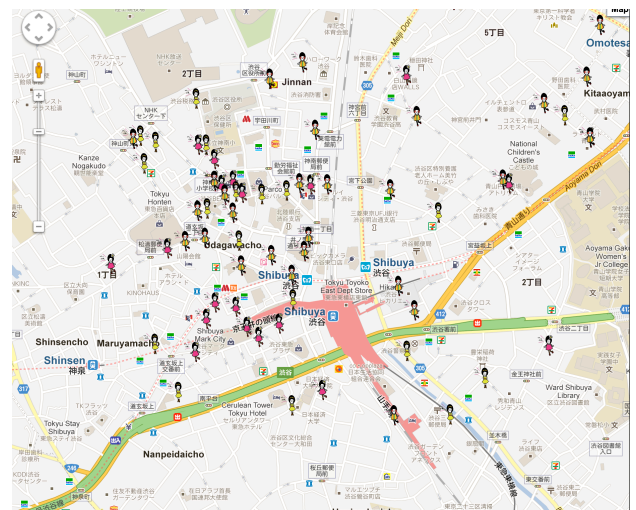


図 5 渋谷のカフェ

Figure 5 Cafés in Shibuya

## 5. CafePalette の評価

### 5.1 評価方法

CafePalette に表示されたリアル店舗と、ブログ記事の見た目の「かわいい」「きれい」「おもしろ」タイプとの整合性を評価することを目的としてカフェ店舗を分析した。

リアルなカフェの視覚特性の研究は大塚らの研究[9]が挙げられる。大塚らは、昭和初期のカフェの建築の内部意匠の中で大きな視覚特性を生み出しているグラフィックな模様についてその組み合わせパターンを分析した。しかし、それらの研究はリアルなカフェの視覚特性の分析のみを対象としており、検索の評価としてネットとリアルのかわいいカフェの視覚特性を考察した文献は見当たらない。本研究では、ブログ記事のかわいいタイプの違いと実世界のかわいいカフェについて考察を行う事で「CafePalette」のサービスとアルゴリズムの有用性を検証する。方法は下記の通りである。まず CafePalette で表示された店舗のうちかわいいタイプごとにランダムに選び出された 10 店舗、合計



30 店舗について調査を行う。目立つ箇所を記録におさめ、それらの中で特徴や関連性を見つけ出し分類する。この分類で見つかった特徴をリアルのかawaiiい視覚特性とする。

## 5.2 評価結果

評価の結果、CafePalette で表示されたカフェは、店の客の男女比が 8:2 と女性の方が高く、使用用途としてはデートや女子会が開かれている事が多かった。これらの事から当初の目的だった女性が好むカフェが検索されたと言える。

## 6. 考察

ブログ記事のかawaiiいタイプとリアル店舗のマッチング状況を考察する。

店の特徴を抜き出したところ、図6の「キャンドル」「黒板」「シャンデリア」などの30の要素が見つかった。図6は、リアルのかawaiiい視覚特性に対して、「かawaiiい」「きれい」「おもしろ」が含まれる店舗数を表している。「テラス」「黒板」「オリジナル食器」は全てのタイプに存在する視覚特性であり、食べ物による装飾よりも家具などの装飾が多い事が見てとれる。これらの3要素がCafePaletteにおけるかawaiiいカフェの特徴であると考えられる。次にそれぞれのかawaiiいタイプについて比較を行う。

### (1) かawaiiいタイプの特徴

かawaiiいタイプについて、最も多いリアルのかawaiiい視覚特性は「テラス」と「ウッド調の家具」であった。次いで「黒板」「シャンデリア」「パンケーキ」「ネオン」「ランプ」「白壁」である。図7はかawaiiいタイプに分類されたカフェの一例であるが、上記に挙げたウッド調の家具や黒板がみとれる。

記事の書き手であるかawaiiいタイプのブロガーの特徴として、図2のように絵文字や顔文字行間が多い記事を書き装飾を好む事ときれいタイプよりも年齢層が低く学生が多いため低価格なブランドを身につける事が挙げられる[6]。かawaiiいタイプのcaféは「テラス」や「ウッド調の家具」など自然と調和した優しい雰囲気のものも多く「黒板」「ランプ」など小さく細かい装飾が見られた事が特徴的である。また、きれいに特徴的な「皮」「ガラス」の素材を用いる店舗よりも低価格なものが多い。ブロガーの年齢による経済状況の違いと、ブログ記事の装飾の好みからリアルでもそのようなカフェを好んだのではないかと考えられる。また「かawaiiい」という言葉は「まるい」「優しい」「親しみやすい」「幼稚な」「小さい」「隙のある」などを意味する事が多い[10][11][12]のでかawaiiいタイプで検索されたカフェは「かawaiiい」で形容されやすいカフェであると言える。

### (2) きれいタイプの特徴

きれいタイプでは「ガラス張り」の視覚特性が最も多く、図8で表すような外観がガラス張りで中の様子がよく見えるカフェが多かった。また、「皮等のイス」のようにかawaiiいタイプの「ウッド調の家具」とは素材が異なり、形が揃

えられた家具が多く見られた。これらの事から、かawaiiいタイプときれいタイプは構成される素材が違い、きれいタイプでは皮やガラスのように高級感のある素材が用いられたカフェが多い事が見て取れる。

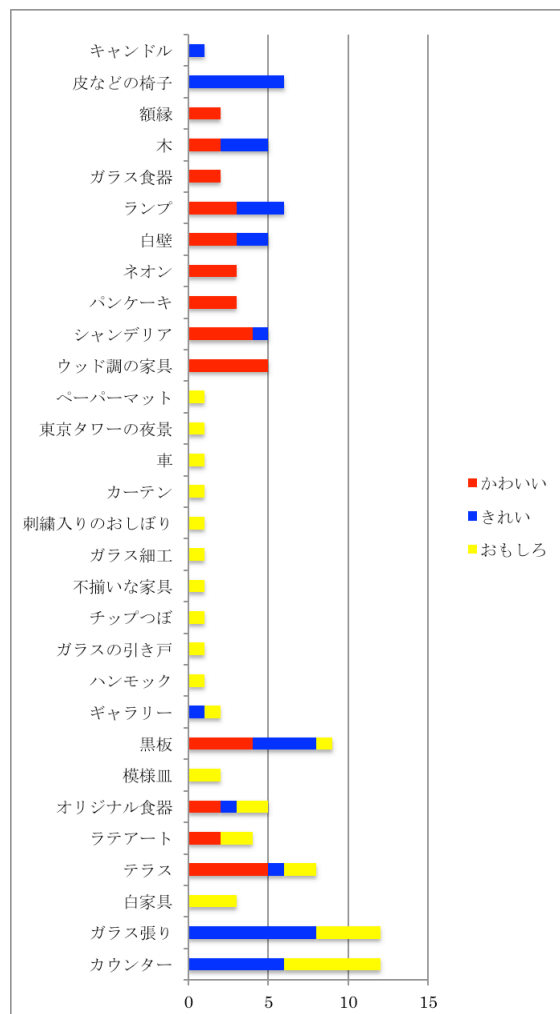


図 6 リアルのかawaiiい視覚特性とかawaiiいタイプ

Figure 6 Visual Trait and kawaii Type



図 7 代官山モンスーンカフェと下北沢の CAFE VIET ARCO

Figure 7 Daikanyama's MonsoonCafe and Shimokitazawa's CAFE VIET ARCO

図2のようなきれいタイプのブロガーは、ブログ記事の装飾が少なく年齢層が高いブランド嗜好な特徴がある。ブログ記事の見た目と同様、細かい装飾よりも高級感のある素材を用いたカフェを好む傾向にあり、かawaiiいタイプとの比較において高級嗜好な傾向があると考えられる。



図 8 渋谷の Legato と六本木 MADURO

Figure 8 Shibuya's Legato and Roppongi's MADURO

### (3) おもしろタイプの特徴

おもしろタイプのカフェは、図9で示す「ハンモック」や「チップ瓶」「ふぞろいな家具」「車」など他のタイプには見られない特徴を持つ店舗が多かった。かわいい検索の調査ではおもしろタイプの記事の書き方には特徴が出なかったが、12の視覚特性で捉える事のできない個性的な書き方の中に特徴のある店を見つける事ができたのではないかと考えられる。



図 9 渋谷カフェマンドゥーカと下北沢 BEAR POND

ESPRESSO

Figure 9 Shibuya's Café Manduka and Shimokitazawa's BEAR

POND ESPRESSO

## 7. おわりに

本研究では、かわいいカフェを探す為の CafePalette の提案と実装および実験を行った。まず雰囲気抽出する事を目的に SVM でかわいい視覚特性の3分類を行った。次に、分類されたブログ記事からカフェ名を抽出し CafePalette を実装した。CafePalette で表示された店舗を分析したところ下記の事が分かった。

①全てのかわいいタイプのカフェで女子会やデートなどに利用される事が多かった。

②かわいいタイプときれいタイプの違いは素材であり、かわいいタイプでは「木」などの親しみやすく、きれいタイプでは「皮」などの高級感のある素材を利用する事が多かった。

③おもしろタイプは特に特徴的であり他のカフェには存在しないアイテムが多かった。

これらの事からまず、CafePalette は当初の目的だった女性の好むかわいいカフェが検索できると言える。また、ブログ記事の見た目がリアルな場所の好みに少なからず反映されているという事が見てとれた。

今後の課題として CafePalette を実用化する事を考えている。実用化の過程で下記の課題が挙げられる。

①おもしろの特徴点の改善

おもしろの正解率が低かったため、おもしろの見た目を表す特徴点を増やす必要があると考えられる。

②ユーザ投稿

CafePalette を用いてかわいいカフェに足を運んだユーザがカフェの写真とコメントを投稿する機能を実装しユーザ参加型にする。

③かわいいカフェ紹介プラットフォーム

現在ブログ記事からカフェ名を抽出している。今後は、かわいいカフェ紹介プラットフォームを作り、読者モデルなどがカフェを紹介しやすいインタフェースを作成する。

このような機能を実装しサービス化する事で、かわいいカフェ文化を発展させたいと考えている。

## 参考文献

- 1) かわいい検索  
<http://kawaii-search.jp>
- 2) 橋口恭子, 松野香織, 若林里奈, 小川克彦, : 見目から判定するブログの「かわいい検索」エンジンの提案, ヒューマンインタフェースシンポジウム(2011).
- 3) Goo ラボ  
<http://labs.goo.ne.jp>
- 4) ことりつぷ 東京, 昭文社(2010).
- 5) 橋口恭子, 小川克彦: 「かわいい検索」における検索行動の分析と評価, インタクション(2012).
- 6) 橋口恭子, 見目から判定するブログのかわいい検索の研究, (2013).
- 7) I. Steinwart and A. Christmann, *Support Vector Machines*. Springer(2008).
- 8) 相良毅, 有川正俊: ジオパースによる Web からの空間コンテンツ獲得, 電子情報通信学会第15回データ工学ワークショップ(2004).
- 9) A. Ohtsuka and T. Hatsuda, "9203 A Study on interior graphical patterns in cafe and tea room," *Summaries of technical papers of Annual Meeting Architectural Institute of Japan. F-2, History and theory of architecture*, vol. 2004, pp. 405-406 (2004).
- 10) 四方田犬彦, : 「かわいい」論, 筑摩書房(2006).
- 11) 真壁智治: カワイイパラダイムデザイン研究, 平凡社, (2009).
- 12) 読売新聞新日本語企画班, 新日本語の現場. 中公新書クラレ, (2003).