

## 利用者モデルを用いたインタラクティブ画像検索の評価 Evaluation of Interactive Image Retrieval Using User Models

井上 雅史<sup>†#</sup> ゲン ホン<sup>‡</sup>  
Masashi Inoue Hong Ngyuen Manh

### 1. はじめに

検索システムを利用する際に、必要とする情報が明確であり、その情報がどこで手に入るかが正確に分かっている場合はまれである。特に、画像などのテキスト以外のメディアによって情報が表現されている場合、利用者が自身の情報要求を適切に表現することに不慣れであること、検索システムが要求された情報をテキストほどには適切に回収できないことなどから、システムとの一度の操作で情報探索課題を完遂することは難しい。このため、システムとの反復的なやり取りが必要とされる。本研究では、一例として、インタラクティブな画像検索システムに焦点を当てる。

効果的な情報検索システムの実現上の主要な課題の一つが、効率的かつ公平な評価の実施手法の開発である。大量の被験者を雇い、長時間を掛けたテストによって、システムの有効性を比較的正確に計測することは可能であるが、コストが高くなりすぎる。一方で、開発者の主観のみでシステムの良し悪しを判定すると、一般の利用者にとって適切なシステムが開発されているという保証がない。そこで、本研究では開発者以外の利用者の行動様式を導入しつつも、評価コストを削減する試みとして、利用者モデルを構築・利用する枠組みを提案する。

### 2. 関連研究

システムに対する問い合わせとその結果提示で検索が完結する静的なシステムにおいては、共通のテストコレクションを用いて、異なるシステムをある程度公平かつ自動的に評価可能である。テストコレクションの適切さや、評価尺度の特性等が問題になることはあるが、人手による作業は、テストコレクション構築段階でのみ必要と考えられている。さらに、この人手によるコレクション作成のコストを自動化により削減しようとする試みも存在する[1][2]。

例えば著名なウェブ検索エンジンなど、既に実用に供されていて、多数の利用者を獲得しているサービスの場合には、機能やデザインの変更を、既存の利用者を実験被験者とみなして評価することができる。一方、革新的な新規検索サービス等を提供しようとする場合、サービスを成熟させるための評価は、自前で行う必要がある。特にランキング結果の評価に加えてインタフェース部分や多様な検索機能の提供にも関心がある場合、実際の人間による操作が円滑に行われているかを観察することが欠かせない。その際の行動履歴のデータが、システム評価の基盤となる。行動履歴に加えて、例えば高久らは、検索の際の利用者の視線を計測することで、利用者属性や検索課題と、検索中の注視点の特徴付ける方法を提案している[3]。このほかにも実際の利用者から様々な情報を収集して評価の質を高めることができるが、収集のコストはさらに上昇してしまう。

機能やインタフェースの変更・評価を迅速に行うためには、利用者実験のコストを抑える必要がある。そのための方法として、ユーザのモデルを作成し、行動を予想すると

いう方法がある。「ペルソナ」という名称で、仮想的なユーザ像をいくつか用意する方法が試みられている[4]。しかし、それらのモデルは設計者の主観に基づいていることが多く、定性的な性質のみを対象とするという限界がある。検索ではなく情報機器のインタフェースに対しては、ボタン操作等の遷移を数理的なモデルで表現・評価しようとする試みがある[5]。我々は、同様の発想でインタラクティブ画像検索システム利用者のモデル化を試みる。

### 3. データ収集環境

#### 3.1 検索対象画像と検索課題

検索対象として、ImageCLEF2008 のテストコレクション（画像集合及び検索課題）を使用した。画像総数は 40,000 で、検索課題数は 39 であり、その内容を付録の表 1 に示した。各画像にはテキストによる注釈が付けられており、検索は、表 1 に挙げられている検索課題の内容文を用いたテキスト検索により開始される。各検索課題には、適合画像の集合が定められているため、ある参加者が、検索のある時点でどの程度適合画像を集めることができているかが、定量的に把握できる。

#### 3.2 実験参加者

実験には、24 名が参加した。参加者は 19 歳から 32 歳までの、計算機科学の素養がある技術者、研究者、学生、教師であった。システム利用シナリオとして、各課題に対して適合する画像をなるべく多く集めるという目的を設定した。つまり、再現率を重視することになる。参加者は、収集結果に満足するか、それ以上検索作業を継続することを望まないとき、いつでも検索を終了することができる。従って、同一の検索課題に対しての、インタラクションのラウンド数及び検索の結果収集できた適合画像数は、参加者ごとにまちまちである。参加者のうち、検索課題をこなした数が上位であった 10 名を真剣に実験に取り組んだと考え、その操作履歴をデータとした。

#### 3.3 プロトタイプ検索システム

検索行動履歴を利用者情報及びシステムの内部状態の情報と併せて収集するために、我々独自の簡易なインタラクティブ画像検索システムを作成し、インターネット経由で利用可能とした。実験参加者は、ユーザ登録後にインターネットブラウザ (Firefox) を用いて、画像を検索することができる。ただし、初期クエリは、前述の 40 課題に固定されている。このシステムでは、以下の 3 種類の行動が記録される。

1. 「類似」ある画像と視覚的に類似した画像を検索
2. 「書換」検索語を書き換えてテキストで再検索
3. 「戻る」前のページに戻る

このうち、「類似」操作では、画像から取り出した色情報に基づいた類似検索を実行している。また、「書換」で利

用可能な検索語は、事前に準備された検索語候補からの利用者の選択によって実行される。

4. 収集データ分析

収集された行動履歴のうち、基本となるのが、システムとやり取りを行った回数である。10名の参加者による合計299回の検索試行について、インタラクションの回数(ラウンド)毎に頻度を示したものが図1である。ラウンドが進むにつれて単調に減少しており、最大のラウンド数である25の頻度は1である。つまり、ある特定の参加者が特定の検索において、25回の操作を行ったことがあるだけであり、平均的には検索は7.28回の操作で終了する。この数字は、今回設定した検索シナリオが、なるべく多くの適合画像を収集するというものだったため、比較的長時間システムと交わっていた可能性がある。

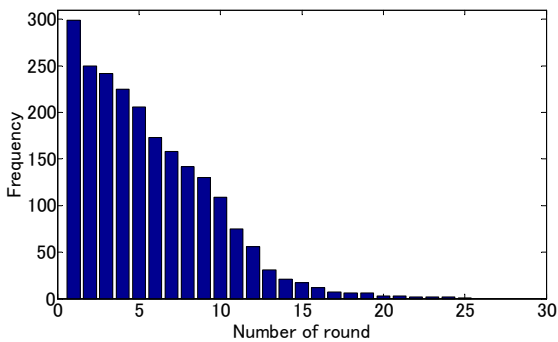


図1 利用者が行ったインタラクションの回数とその頻度

5. ユーザモデル生成

ユーザ履歴に基づいて、次の行動が確率的に決定されるというモデルを推定する。様々な要因が行動決定に影響するが、例えば何回目のラウンド r であるかのみ依存するモデルの場合、

$$P \begin{pmatrix} \text{類似} \\ \text{書換} \\ \text{戻る} \end{pmatrix} | r=2 = \begin{pmatrix} 0.339 \\ 0.533 \\ 0.128 \end{pmatrix} \quad P \begin{pmatrix} \text{類似} \\ \text{書換} \\ \text{戻る} \end{pmatrix} | r=4 = \begin{pmatrix} 0.272 \\ 0.602 \\ 0.126 \end{pmatrix} \quad P \begin{pmatrix} \text{類似} \\ \text{書換} \\ \text{戻る} \end{pmatrix} | r=6 = \begin{pmatrix} 0.253 \\ 0.652 \\ 0.095 \end{pmatrix}$$

となる。この場合には、次第に視覚的に類似した画像の検索よりも、テキストによる問い合わせの書き換えが好まれる傾向が見て取れる。

6. まとめ

本研究では、インタラクティブなマルチメディアシステム、特に画像検索システムの開発を支援するために、モデルを用いた評価の枠組みを提案した。実際の利用者から利用履歴を取得する方法、履歴からモデルを構築する方法をそれぞれ述べた。実際に収集された利用履歴の性質について考察した。

今後は、様々な要因を考慮に入れて、利用履歴から利用者モデルを推定し、推定されたモデルから、システム評価に役立つ擬似行動履歴を生成することができるかを検証する計画である。

参考文献

[1]M. Inoue, N. Ueda, "Retrieving Lightly Annotated Image Using Image Similarities", SAC '05: Proceedings of the 2005 ACM symposium on Applied computing, pp. 1031-1037 (2005).

[2]E. Graf, L. Azzopardi, "A methodology for building a patent test collection for prior art search" EVIA-2008 Workshop (2008).  
 [3]高久 雅生, 江草 由佳, 寺井 仁, 齋藤 ひとみ, 三輪 眞木子, 神門 典子, "サーチエンジン検索結果ページにおける視線情報の分析", 情報知識学会誌, Vol.19 No.2, pp.224-235(2009).  
 [4]A. Cooper, "About Face 3: The Essentials of Interaction Design", Wiley (2007).  
 [5]H. Thimbleby, "User interface design with matrix algebra", ACM Trans. Comput.-Hum. Interact., Vol.11, No. 2, pp. 181-236 (2004).

付録

表1 検索課題と課題毎の実験実施回数(1人の参加者が同一の課題を複数回実施している場合がある)

ID	検索内容	回数
2	church with more than two towers	17
3	religious statue in the foreground	11
4	animal swimming	12
5	straight road in the USA	9
6	destinations in Venezuela	8
7	black and white photos of Russia	12
8	people observing football match	10
9	exterior view of school building	10
10	night shots of cathedrals	7
11	people in San Francisco	8
12	lighthouse at the sea	10
13	sport stadium outside Australia	7
14	exterior view of sport stadium	9
15	close-up photograph of an animal	9
16	accommodation provided by host families	10
17	sport photos from California	8
18	snowcapped building in Europe	8
19	cathedral in Ecuador	7
20	views of Sydney's world-famous landmarks	8
21	volcanoes around Quito	9
22	group picture on a beach	5
23	bird flying	9
24	sights along the Inka-Trail	7
25	people in bad weather	9
26	tourist destinations in bad weather	6
27	winter landscape in South America	6
28	sunset over water	7
29	mountains on mainland Australia	5
30	vehicle in South Korea	5
31	images of typical Australian animals	5
32	indoor photos of a church or cathedral	7
33	sports people with prizes	5
34	views of walls with unsymmetric stones	4
35	famous television (and telecommunication) towers	6
36	drawings in Peruvian deserts	4
37	photos of oxidised vehicles	6
38	seals near water	5
39	creative group pictures in Uyuni	5
40	salt heaps in salt pan	4