

## ハイパーメディアオーサリングツールにおける

2C-3

## 大規模データへの対応 ～検索支援～

柴田清己 紀伊隆弘 宇山政志 馬場ひとみ  
(株)富士通研究所

## 1. はじめに

我々はハイパーメディアオーサリングツールの研究開発を行ない、実験的に利用を続けている。利用を重ねるにしたがいハイパーメディア文書の量が増え、過去に作成したハイパーメディア文書の一部を再利用することも多くなった。このようなとき、利用者は再利用したいハイパーメディア文書の内容は覚えていても、名称や保存場所を覚えていないことが多い。そのため検索が非常に困難であり不便を感じている。そこで、検索のために補助的なツールを作成した。

## 2. 従来のツールの問題点

大量のハイパーメディア文書の中から欲しいものを探そうとする場合、従来の方法ではおおよそ次のような手順となる。

- i) ハイパーメディア文書の名前の一覧を見て、名前から見当をつけて一つ選ぶ。
- ii) オーサリングツールで内容を確認する。
- iii) 欲しいものでなかったら i) に戻る。

すなわち、適切な名前が付いていなければ、iからiiiを何度も繰り返すことになる。しかし、実際にはハイパーメディア文書が増えるにしたがって紛らわしい名前を付けることが多くなるため、この方法は破綻する。

これと同じことはテキストデータだけからなる文書の場合でもいえるが、テキストデータの場合にはキーワードを与えて計算機に検索させるという強力な手段が存在している。しかし、ハイパーメディア文書はテキストデータに加えて画像データや音声データからなり、さらにリンク構造などを含んでおり、これに対して上のようなキーワードによる検索を実現するのは困難である。そこで、我々はハイパーメディア文書を画像データとして検索するツールを実現した。また、実現にあたって簡単な予備実験を行なった。本稿では、ツールの概要と予備実験について述べる。

## 3. 検索ツールの概要

## 3.1 想定する検索作業

ハイパーメディア文書の検索では、画像データや画面のレイアウトなど視覚的なイメージを頼りに探すことが多い。そこで、ハイパーメディア文書の視覚イメージを使って検索することを想定した。

さらに、作業の手順としては次のようなものを想定した。

- i) たくさんのハイパーメディア文書をざっと見て、目標に近いものを選び出す。
- ii) 数の減った中から、さらに、目標に近いものを選び出す。これを何回か繰り返し、少数の候補にしぼる。
- iii) 少数の候補に絞ったところで、それぞれの候補についてオーサリングツールで細かな内容まで確かめる。

## 3.2 対象データ

我々が開発中のオーサリングツールで作成したハイパーメディア文書を対象とする。一つのハイパーメディア文書は、複数のページから構成される。ページにはテキスト、静止画像、動画像などの素材データが配置されている。

さらに検索のために、ページのある瞬間の状態をページイメージと呼ぶ1枚の静止画に記録するものとする。したがって、各ハイパーメディア文書は検索時には複数の静止画の列で表現することができる。

## 3.3 特徴

本ツールの主な特徴をあげる。また、図1に外観を示す。

- ハイパーメディア文書の一覧として、名称だけでなく、先頭ページのページイメージを表示する

経験的には、ほとんどのハイパーメディア文書で先頭ページがタイトルのページとなっており、ハイパーメディア文書の内容をよく表している。

- 段階的に候補を減らしながら検索を進める

ハイパーメディア文書やページにマークを付けて、候補から落したり、拾い上げて候補にするなどして、一覧するハイパーメディア文書の数を減らしてゆく。

- 候補数に応じて適当な縮小率を選ぶ

一覧性を高めるために候補が多い時には縮小率を上げて、できるだけすべてのページイメージが画面に収まるようにする。また、このとき、候補の多い段階で縮小率が高いために細部を識別しづらくなることはあまり問題にならない。候補の多い段階では、細部を識別することなしに、目標とは明らかに違うと分かるようなものが多数あるためである。

- モード切替の必要性を減らす

ハイパーメディア文書一覧モードからページ一覧モードに移ることなく、その場でハイパーメディア文書一覧モードのままハイパーメディア文書内のページをブ

Authoring Tool for large scale Hypermedia documents

- Visual information retrieval -

Seiki Shibata, Takahiro Kii, Masashi Uyama, Hitomi Baba  
FUJITSU Laboratories LTD.

ブラウジングすることができる。この表示方法は、同じ領域にページイメージを次々と表示するという単純なもので、本のページをバラバラとめくるような感覚を再現しようと試みたものである。

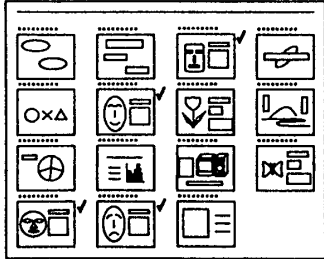


図 1: 検索ツールの外観

#### 4. 予備実験

上で述べた連続ページめくりを実現するにあたり、ページめくりの速さ、すなわち、各ページを呈示する時間についての予備実験をおこなった。

##### 4.1 実験の概要

本実験の目的は、適当なページめくりの速度を調べることにある。特に、実験1では速度の限界を、実験2ではユーザの好む速度を調べた。

世界各国の国旗の絵柄をイメージスキャナで読み込んだ複数枚の画像を用意する。被験者には、あらかじめ呈示された目標の画像を、その中から探し出すという作業をもらった。画像は14インチカラーディスプレイに表示した。縦横とも全画面の1/2の大きさで画面のほぼ中央に表示した。

##### 4.2 実験1

まず、被験者には画像1枚を2秒間呈示する。続いて40枚の画像を連続ページめくりの方式で表示する。最後に、先に呈示された1枚の画像が、後に表示された40枚の画像の中にあっただうかを質問する。

画像はあらかじめ選んだ41枚を用いた。呈示する順序は無作為に決定したが、ただし、先に表示した1枚が後に表示した40枚の中に含まれる確率は、1/2となるように制約を加えた。

1枚あたりの呈示時間を長くしながら、それぞれ5回ずつ試行した。

##### 結果

それぞれの呈示時間における正答率を全被験者について平均をとったものを、図2に示す。

どの被験者についても、呈示時間が80から100msecを越えると、安定して約8割程度の高い正答率を示していた。また、呈示時間が100msecより長くなっても正答率はほとんど上がらなかった。

##### 4.3 実験2

まず、被験者に画像1枚を2秒間呈示する。被験者には最初に呈示した画像を80枚の画像の中から探してもら

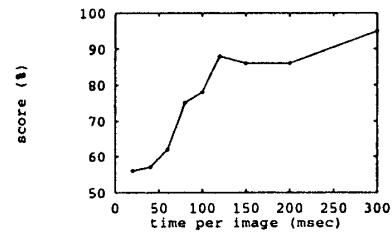


図 2: 実験1の結果

う。被験者はいつでも自由にページめくりの開始と停止ができ、方向も変更でき、また、速度も1msec刻みで変更できる。ただし、初めてキーが押されるまでは20msecの速度でページがめくられる。

80枚の画像はあらかじめ選んだものを用いた。順序やはじめに呈示される1枚の画像は毎回無作為に決定した。

これも5回ずつ試行した。

##### 結果

各被験者について、ページめくりの速度の平均を求めるため、ページめくりの継続した時間を重みとした加重平均を求めた。

被験者	1	2	3	4	5
平均速度 (msec)	97	136	110	141	159

6	7	8	9	10	全被験者の平均
55	62	73	105	68	101

##### 4.4 実験のまとめ

実験1から100msec付近に認識率の下がる点があることが分かり、さらに、実験2から被験者もこの付近の速度を選んでいようである。

#### 5. まとめ

本稿では、大量のハイパーメディア文書を画像イメージとして検索するツールの機能について述べた。このツールは連続ページめくりの機能によって高速ブラウジングを支援する。さらに、この連続ページめくりの適切な速度を求めるための予備実験について述べた。その結果からおおよその目安となりそうな値を求めた。

今後の課題は、ツールの評価実験を行なうことである。特に、連続ページめくりの有効性と実験1で求めた速度が適切であったかどうかを確かめたい。

また、実験2の結果についてさらに詳しく検討を加え、速度以外の要因についても調べたい。

##### 参考文献

- [1] 馬場ひとみ 他: 「オーサリング・ツールの概要」, 情報処理学会第42回全国大会, 1991.
- [2] 紀伊隆弘 他: 「ハイパーメディアオーサリングツールにおけるデータの大規模化への対応~データ管理~」, 情報処理学会第44回全国大会, 1992. (予定)