

ネットワーク型マルチメディア情報の媒介を目指した 公開型画像データベース GIRLS

谷田部 智之 高羽 洋樹[†] 佐藤 隆^{††} 坂内 正夫

東京大学生産技術研究所

〒106 港区六本木7-22-1

Tel: 03-3402-6231 ex.2640 Fax: 03-3402-5078

E-mail: {yatabe, takashi, takaha, sakauchi}@sak.iis.u-tokyo.ac.jp

[†] 現在 シャープ株式会社

^{††} 現在 NTT ヒューマンインタフェース研究所

ここ2年程の間に World-Wide Web (WWW) は目覚しく発展し、多くの情報が WWW 上で発信されている。今後、さらに情報量の増加、多様化が進むと予想される。しかしながら情報が無秩序に存在するため、現状では、ユーザが必要な情報に到達するという目的は、“サーチエンジン”を使うことである程度満たされている。我々は、情報探索問題への新たな手法として、ホームページに多数含まれる画像を手掛かりにアクセスするプロトタイプシステムである画像データベース GIRLS を開発し、そこで得られたデータについての考察を行った。

The open Image database GIRLS aiming at mediating Multimedia data on the Network

Tomoyuki YATABE Hiroki TAKAHA[†] Takashi SATOU^{††} Masao SAKAUCHI

Institute of Industrial Science, The University of TOKYO

7-22-1 Roppongi, Minatoku, Tokyo 106, JAPAN

Tel: 03-3402-6231 ex.2640 Fax: 03-3402-5078

E-mail: {yatabe, takaha, takashi, sakauchi}@sak.iis.u-tokyo.ac.jp

[†] SHARP Corporation

^{††} NTT Human Interface Laboratories

Recent progresses of the World-Wide Web (WWW) are remarkable. Already, huge amount of information is available of the WWW, and it is expected to grow further. However, information isn't managed on the WWW, so use of "Search Engine" meets the needs of users' purpose of information to want to a certain degree. We develop a prototype system, Image database GIRL, to access the information by indexing images in homepages, as a new solution of the resource discovery problem on the WWW.

1 はじめに

World-Wide Web (WWW) [1] は目覚しく発展し、多くの情報が WWW 上で発信されている。また、これまでネットワーク上で提供されてきた文字情報にとどまらず、画像・映像や音声などの新たなメディアを利用した情報が提供されている。今後はさらに情報量の増加、多様化が進むと予想される。

現状では、ユーザが必要とする情報に到達することは、容易ではないと言えよう。情報が管理されていないために、ユーザはハイパーテキストのリンクを“手探り”でたどって欲しい情報を探索しなければならない。ユーザが必要な情報に到達するという目的は、各種の“サーチエンジン”を使うことである程度満たされている。現在のところ WWW 上のデータを集める手段として、主として以下の方法があげられることができ、これらの手法を利用してデータベースが作られている。それぞれに長短所があり、主に文字情報を利用した検索が行われている。

- 自動的に収集
- 情報発信者が登録
- 第三者が登録

ハイパーメディアと利用可能な WWW を文字情報のみを利用したナビゲーションだけでなく、画像情報を含めたマルチメディア情報を利用したナビゲーションが考えられる。

本稿では、まず、マルチメディア情報を扱う上でのマルチメディアデータベースに関する整理し、マルチメディア情報を探索する新たな手法として、ホームページに多数含まれる画像・映像を自動的に集め、それらのデータを手掛かりにアクセスするプロトタイプシステムを実装した。また、実験に使用したデータについての考察を行った。これらは、閉じたシステムではなく、実際にユーザが使えるように、公開型の画像データベースを目指している。

2 マルチメディア情報

今後のマルチメディアシステムには、一般に3つの要素が必要とされる。映像や画像、図形、音声等の「マルチメディアコンテンツ」、それを伝達するための「通信システム」、ユーザに対してアクセス、検索方法を提供するインタフェースを含む「プラットフォーム」である。この中でマルチメディアコンテンツは、マルチメディアソフト、デジタル放送、映像ソフトに代表されるように、利用者に対して直接付加価値を生み出していくもので、今後のマルチメディアシステム構成の上で重要な要素である。

コンテンツがどこにどのような形で存在するかということによりマルチメディアデータベースを以下のように分類することができる。

- ネットワーク型データベース

1989年から始まったWWWがネットワーク上で膨大なデータリソースを提供している。現在数万のサーバがホームページを持ち、映像・画像を含むマルチメディア情報の提供を行っている。またこれらは、極めて分散した形で存在するため、データベース化と言う点ではこれから解決すべき課題が多い。

- ストリーム型データベース

現在マルチメディアコンテンツの供給源として主役となっているのは、放送である。現在でも通信衛星を通じて150チャンネル以上の番組をアメリカのDirect TVが提供している。また、日本でも今年10月から通信衛星を利用してPerfec TVが70チャンネル程度の放送を始める予定になっている。CATVや地上波のデジタル化が進むにつれて、さらに多くのチャンネルが提供されるようになることが予想される。そこでは、ユーザは映像を見せてもらうというのではなく、映像を中心としたリアルタイムな情報ストリームがあり、ユーザの興味や目的に合わせて、その中から情報を選択することで、データベースを構築することができる [2]。

- ストア型データベース

これまでデータベースと言えば、ディスク等に蓄積されたデータベースであった。マルチメディアデータベースとしてCD-ROMを利用して画像や音声等のデータが提供されている。電子図鑑や電子図書館等のデータベースはここに入れられるだろう。

本稿では、これらマルチメディアデータベースのうちネットワーク型データベースであるWWWを利用したマルチメディア情報を媒介するシステムについて述べる。

3 公開型画像データベース GIRLS

WWWでは、マルチメディア情報の一つである画像・映像を容易に発信でき、ユーザが使用するMosaicやNetscape Navigator等のブラウザが画像を表示できるように対応しているため、各ページの多くに画像が含まれている。文字情報を利用した分類・検索を行う“サーチエンジン”によるサービスはすでに行われているが、WWW上で提供されている画像・映像情報に対して、これまで分類や検索できるようなサービスは行われていない。これは、画像の内容に応じたインデックス付けに高いコストがかかることやあるいは画像情報に対する多数のユーザの多様な検索要求に答える事が、一般的には難しいことなどに起因している。

そこで、本システムではWWW上の情報である文字情報だけではなく、マルチメディア情報として画像・映像情報の検索を行えるようにした。このような検索手段を提供することはユーザにとって、画像を探す事ができるだけでなく、文字情報との連係により検索効率が向上させることができるという利点があると言える。

3.1 システム構成

本システムは図 1に示すように、画像探索ロボット BOYS、データベース部、画像検索部 GIRLS およびユーザインターフェース部に大きく分けることができる。それぞれは独立したものとなっているため、今後ユーザが自由なものを使う事ができるようになっている。

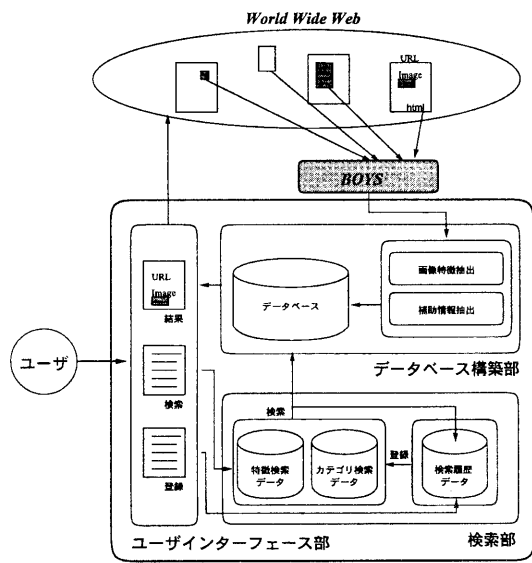


図 1: システム構成

3.1.1 画像探索ロボット BOYS

これまでに WWW 上の情報資源について様々なデータベースが構築されているが、大規模なデータベースを構築するための手段として、“ロボット”あるいは、“スパイダー”と呼ばれる自動的にデータを集めるプログラムが使われている。在世界中で 100 近くの“ロボット”が確認されている [3,4]。

それぞれの“ロボット”の行動の目的は異なっているが、基本的な動作としては、ある URL (Universal Resource Location) から順にリンクをたどりながら、情報 (HTML 文書) を集めて、データベースを構築していく。一種のエージェントとも言えよう。原理的には人間が実際に“ネットサーフ”する代わりに自動的に行ってくれるものとも言える。

実際のプログラムは、perl 言語を用いて実装されたものが多く、これは HTTP プロトコル [5] を実装する上で、簡単に記述できること、HTML 文書を解析しやすいためであると考えられる。また、ライブラリが用意されていることも一因となっている [6]。

我々が開発した“BOYS” [7] は、主としてページ上存在する画像・映像情報を集める事を目的とした“ロボット”

である。これは、perl 言語を用いて実装されており、基本的な動作は図 2 の通りである。

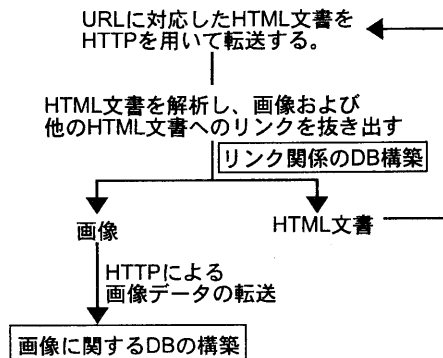


図 2: BOYS の基本動作

動作を大きく分けると、以下の3つの機能に分けられる。

1. HTTP (Hyper Text Transfer Protocol) に基づくデータの転送

データの転送には、それぞれの HTML 文書および画像の URL に対応するファイルを HTTP を用いて転送を行う。この時、ネットワークの負荷を軽くするために Proxy サーバを通じて、データの転送を行う。また一つのサーバにアクセスが集中しないようにアクセスする順にも注意した。各サーバごとに/robots.txt を見て、“ロボット”によるデータの取得を拒否していないかどうか調べる。

2. HTML 文書の解析

 や 等で示された新たなハイパーリンク先である HTML 文書や画像等を抽出する。それぞれのハイパーリンク先の URL について、相対的に書かれた場合も含め、絶対的な URL に直す¹。

3. 取得したデータのデータベース化

1 で集めたデータから 以下の手順により URL や画像情報に関するデータベースを構築する。

- (a) 画像処理により画像特徴抽出
- (b) 画像の URL が記述されていた HTML 文書から補助情報抽出
- (c) 画像データの縮小画像を生成
- (d) 以上のデータをデータベースに保存、画像データを破棄。

¹ただし、http://www.nnt.jp/foo.html という書き方ではなく、http://www.nnt.jp/foo のようにディレクトリを示すのか、ファイルを示すのか分からないものもあるが、それらについてはある程度の推定を行った。

画像特徴は主に検索時に使用される。補助情報は、例えば、その一つである画像の URL を検索結果に添付する事により、WWW 上のリソースにアクセスする事ができる。特徴の種類などの詳細については後述する。保存するデータ量を削減するため、画像データを縮小化した後破棄する。ここで、作られた縮小画像は検索時の表示用として使用する。

3.1.2 画像検索システム GIRLS

ここでは検索手法として、以下の 2 種類の手法を提供している。

- 類似検索

各画像を 27 色に量子化した各色の頻度ヒストグラムを用い、ユーザから指定されたサンプル画像に近いものを提示する(図 3 参照)。具体的な手法として、27 色の色を用いてそれぞれの頻度をベクトルとみなし、画像間の“距離”により検索を行う。また、画像の大きさを指定できる。

HTML を構造を記述するためだけの言語ではなく、プレゼンテーション用の言語であるということもできるので²、画像そのものだけでなく、ページ全体を一つの画像とした場合にも検索を行えるようにした(図 4 参照)。この場合は、ページレイアウトを含め

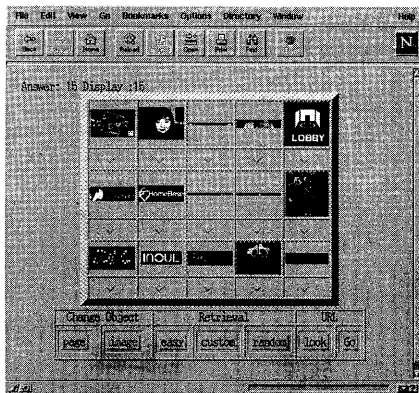


図 3: 類似検索例

- カスタム検索

ユーザが複数のパラメータを指定し、それに合致した画像を表示する。画像の面積、縦横比などや色番号やその占有面積、画像カテゴリなどを指定することができる。画像カテゴリは意味レベルで分類された画像の集合を指し、ここでは、表 1 に従って、写真やイラスト画、顔に分類した。ここでも、前述の類似検索同様ホームページ自体の検索も可能になっている。

²HTML 3.2 が公開されたが、SGML としての言語よりもプレゼンテーション用の言語としての意味合いが強くなったようである。

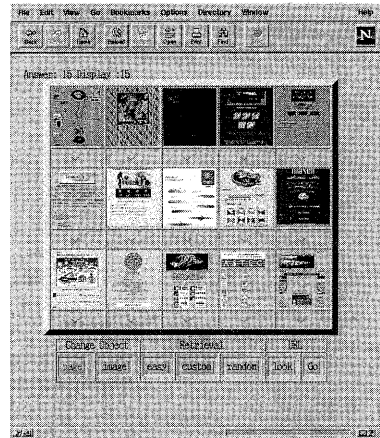


図 4: ホームページの類似検索例

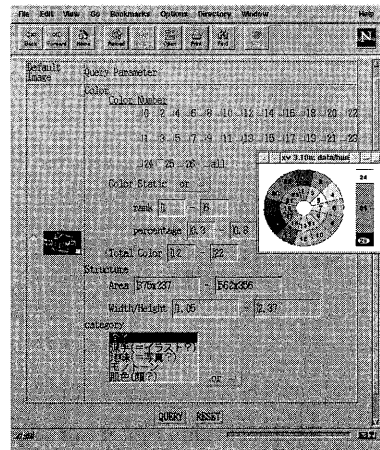


図 5: カスタム検索例

表 1: 画像カテゴリ

カテゴリ	パラメータ
イラスト	$S(\text{高彩度色}) \geq S(I) \times 50\%$ 色数 $> 10 \wedge S(\text{高彩度色}) \geq S(I) \times 10\%$ 最頻度の色が有彩色 $\wedge S(\text{最頻度の色}) \geq S(I) \times 50\%$
写真	イラストでない
顔(肌色)	$S(I) \times 5\% \leq S(\text{色番号 } 1, 2, 3, 5) \leq S(I) \times 50\%$
モノクロ	$S(\text{無彩色}) \geq S(I) \times 90\%$

ここで、 $S(x)$ は x の占める面積、 $S(I)$ は画像の全面積

3.2 ユーザインタフェース部

ユーザインタフェース部はHTMLで記述を行い、Mozilla、Netscape Navigatorなどに代表されるWebブラウザで操作し、検索が行えるように構築している。検索方法の選択、検索結果の表示など、ユーザの要求を入力し、CGIを用いて、検索サーバにデータを渡し、検索結果をHTML文書に変換しブラウザに返す。

3.3 実験の対象としたデータ

日経BP社の企業リスト³ [9]を利用し、各企業のホームページが1503、この中に含まれる7886枚の画像に対して実験を行った。現在の検索用データ量は1画像当たり約20KByte、表示用縮小画像は平均すると1画像当たり15KByte、検索速度は約1000画像/秒程度である。実験に用いたデータの場合全データ量が約300MByteあり、平均検索時間が約8秒ということになる。

また、ページ当たりの画像を含むデータの転送量の統計は図6の通りであり、指数分布状になっている事がわかる。

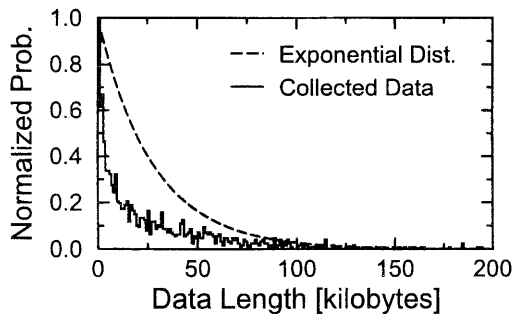


図6: ホームページのデータ長分布

queryの度に、画像から実際に検索するためのデータを作り直す事は時間がかかるため、データベースを生成する段階で、画像処理を行い、検索用のキーを生成している(3.1.1参照)。

また、画像のカテゴリに関する簡易な実験結果を以下に示す。使用したデータは、東大生産研内のサーバが31、画像が約1000個のデータのうちの一部分である。

ここで定めたカテゴリは簡易なもののため、実験結果はあくまで参考程度である。検索キーとしてはより高度なものを考える必要がある。

表1とは別に、表4に沿って分類したデータも参考として図7に示しておく。これは、他の“ロボット”の取得したデータ[10]を利用して、集めたものである。分類は、手作業により行った。ただし、画像の分類自体は客観的なものではないので参考程度であるが、アイコンのような小さな画像、つまり転送するデータ量が少ない画像が多いこと

³現在では構成が代わっているため、<http://www.nikkeibp.co.jp/NDR/SYL/>にあるものを合わせたものとなっている。

表2: 実験結果

	イラスト	写真	モノクロ	顔	合計
イラスト	48	20	6	28	68
写真	4	45	4	36	49
顔	2	25	3	22	27
合計	51	59	10	58	

表3: 再現率と適合率

	再現率	適合率
イラスト	0.706(48/68)	0.941(48/51)
写真	0.918(45/49)	0.763(45/59)
顔	0.815(22/27)	0.379(22/58)

再現率: 検索洩れの少なさ

適合率: 検索の正確さ

が分かる。これはネットワークがそれほど高速でないことが原因であろう。

表4: 画像カテゴリ(参考実験)

アイコン・ロゴ	ボタンなどのような比較的画像サイズの小さなもの、企業等のロゴマークなど
写真	スキャナなどで入力されたと思われるもの
グラフ・図	研究用等で使われるグラフや図
絵・イラスト	アイコンとは区別され、コンピュータを利用して描かれた画像
その他	文章や数字だけなどどれも当てはまらないもの

3.4 問題点及び今後の課題

現在、作られたシステムには、以下のような問題点及び課題となる点があげられる。

- データ収集の高速化

これは、ネットワーク全体の高速化が望まれるが、転送を並列に行うこと以外に現状ではこれ以上の高速化は困難であろう。また、“ロボット”によるデータの収集はネットワークに対する負荷が大きいと予想されるので、できるだけ無駄なトラフィックを減らす必要がある。また、

- データベースの構築の高速化

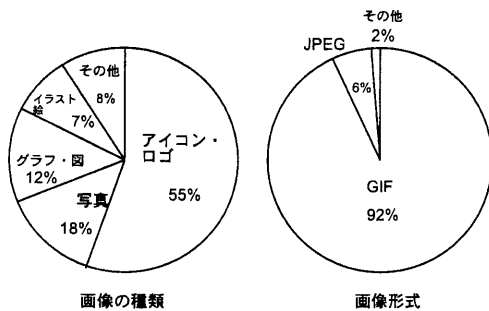


図 7: 画像の分類 (11796 画像)

これに対しては、データベースの分散化にも関係し、並列化により、比較的容易に行えるだろう。これは検索の高速化にもつながる。

● 検索手法の改良

現在では、画像の統計量を利用した検索に対応しているが、さらに高度な条件による検索方法を利用できるようなインタフェースに改良する。表 1、2でも分かる通り、非常に基本的な認識を用いた検索しかできないため、より高度な検索ができるようにデータベース自体の構造も改良の余地がある。これについては、状態遷移型モデル [2] を用いて認識を行うことを現在検討中である。

4 おわりに

WWW 上で提供されている画像データを利用して、ユーザが自由に使うことができる、いわば公開型の画像検索システム GILRS、そのフロントエンドとしてのデータベース構築のための“ロボット”BOYSのプロトタイプについて、実際の動作、集めたデータについて検討を行った。

今後は、さらに多くのデータを用いて、より複雑なカテゴリに対応するための認識手法の検討を行いたい。また、外部からデータベースを使用できるように現在準備中である。

【 参 考 文 献 】

[1] “The World Wide Web Consortium”, <http://www.w3.org/pub/WWW/>

[2] 佐藤 隆, 坂内 正夫, “ライブハイパーメディアにおける映像情報の獲得”, 電子情報通信学会論文誌, Vol.J79-D-II, No.4, pp.559-567, Apr. 1996

[3] Martijn Koster, “World Wide Web Robots, Wanderers, and Spiders”, <http://info.webcrawler.com/mak/projects/robots/robots.html>

[4] 林 良彦, 菊井 玄一郎, 鷲崎 誠司, 砂場 倫太郎, “WWW 情報空間における Resource Discovery と Naviga-

tion 支援”, 電子情報信学会, 人工知能と知識処理研究会, AI-95-31, <http://isserv.tas.nitt.jp/chisho/paper/9511HayashiSIGAI.ps.Z>, Sep. 1995

[5] “Basic HTTP”, <http://www.w3.org/pub/WWW/Protocols/HTTP/HTTP2.html>

[6] Roy Fielding, “libwww-perl: Distribution Information”, <http://www.ics.uci.edu/pub/websoft/libwww-perl/>, May 1995

[7] 谷田部 智之, 高羽 洋樹, 坂内 正夫, “WWW における画像探索ロボット”, 電子情報通信学会 1996 年総大会, D-337, pp.125, Mar. 1996

[8] 高羽 洋樹, 谷田部 智之, 佐藤 隆, 坂内 正夫, “World Wide Web 上の公開型画像検索システム GIRLS”, IE95-76, 信学技報, Vol.95, No.366, pp.1-8, Nov. 1995

[9] “日本企業 URL ディレクトリー”, <http://www.nikkeibp.co.jp/NDR/SYL/index.html>, Mar. 1996

[10] 田村 健人, “千里眼”, <http://www.info.waseda.ac.jp/search.html>

[11] 加藤 俊一, 下垣 弘行, 藤村 是明, “画像対話型商標・意匠データベース TRADEMARK”, 電子情報通信学会論文誌, Vol.J72-D-II, No.4, pp.535-544, Apr. 1989