

画像データベースにおけるメタデータ付与システムの構築と評価

福本 徹*, 赤堀侃二*

The Development and Evaluation of Metadata Attaching System with Image Database

Toru FUKUMOTO*, Kanji AKAHORI*

A large number of digital images are stored on the Internet. From an educational perspective the fact that a vast number of images on a number of topics are readily available and are in some cases free is a very helpful thing. The aim of this study is to develop and evaluate a system attaching metadata to images. The user can attach a metadata looking the images on Web, the same metadata of similar images with similarity engine. This system also has three modes: Package-images mode, One-image mode and Modification mode. And we evaluate and inspect this system with subjects.

キーワード：知的学習支援, WBT, 画像検索, データベースキーワード：

1. はじめに

近年、教育用デジタルコンテンツ環境の整備が進んでおり、動画・静止画など質の高いコンテンツが数多く開発されている。また学校現場では、これまで写真等の静止画やビデオやテレビ番組といった動画が多く使われており、教育効果をあげてきた。しかし多くのデジタルコンテンツが開発されたが、教育現場において教師が授業で使用する際に使いたいというコンテンツが不足していることが指摘されている⁽¹⁾。今後インターネットの教育利用がますます進んで、どのように学習に活用するかに関する方法論がある程度浸透すれば、教育・学習素材情報に対する要求がますます高まっていくことも想定される。

画像等のデジタルコンテンツに対して検索を行なう

には2つの手法が考えられる。1つは当該のコンテンツデータそのものを対象とする方法、もう1つはそのコンテンツデータを説明するデータ、つまりメタデータを利用する方法である。検索精度やWeb検索との親和性からの点ではユーザにとっては後者が使いやすい。

メタデータ型画像データベースでは検索者はキーワードを入力し、検索システムは入力されたキーワードと画像に付与されているメタデータを照合することで検索を行なう。つまり動作は以下ようになる。まず、システム提供者がメタデータそのものの枠組みを設計する。次に設計した枠組みの下で、付与者が画像にメタデータの内容を記述する。そして検索者はテキストなどを検索のためのキーとして入力し、システムが入力されたテキストと画像に付けられているメタデータの構造と内容とを照合することで検索を行なう。

ここで問題となるのは、実際に1つ1つの動画あるいは静止画に対してメタデータをどのように付与するかという点である。これまでは、コンテンツ提供者が

* 東京工業大学大学院社会理工学研究科
Graduate School of Decision Science and Technology,
Tokyo Institute of Technology

人手に頼ってメタデータを付与しているのが現状であった⁽²⁾。またNTT 東日本による教育用学習素材検索システム⁽³⁾では、画像などの学習素材そのものとメタデータは検索システムへの登録時には分離されており、メタデータはExcelのファイル形式で登録する仕組みになっている。メタデータ入力者にとって使いやすく手軽である点、コンテンツ提供者にとって特別なシステムが不要である点から、例えばInternet Explorerなどの画像が表示できるソフトとExcelによるメタデータ入力が一般的であると思われる。

このようなメタデータを付与する手間を軽減するために、画像を解析しキーワード等のメタデータを自動的に付与するという研究が見られる。静止画を対象としたもの⁽⁴⁾⁽⁵⁾、動画を対象としたもの⁽⁶⁾が提案されている。両方式とも以下のような短所が指摘できる⁽⁷⁾。まず、人間による知識の記述が必要であることとしかもその知識はヒューリスティックなものであり一般性に欠けること、例えば前者であれば画像セグメント構成、後者であればスポーツにおける試合構造といったものである。次に、対象となる画像が極端に限定されること、前者であれば風景画などセグメントが容易に定義できる画像、後者であればスポーツなどの構造を持った画像といったものである。

そこで本論では上述したようなメタデータを付与する手間を軽減し、かつ、一貫性のあるメタデータを付与するシステムの開発と評価について述べる。

2. システムの設計

本章ではメタデータ付与システムの設計方針について述べる。

まず、対象となる画像の扱い方について、画像を見ながら付与することとした。画像とメタデータとを分離して扱うのではなく、一体として扱うことを目指した。また、動画と静止画とをシームレスに扱うために、静止画と動画クリップとを扱うこととした。ここで動画クリップは1シーンのみのものとし、連続したシーンを扱うことは今後の課題である。

次にシステムの使い方として以下の3通りが考えられる。大量の画像を対象として一気に付与する、手持ちの1枚の画像に対して付与する、データベース中に

既に存在する画像に対して付与されたメタデータの修正を行なう、である。大量の画像を対象とするのはコンテンツビジネス等を対象としたものと想定でき、1枚の画像を対象とするのは個人による登録型や個人の使用と想定できる。本システムでは両方の用途に対応するため、一気に付与するモード、画像のアップロードによる1枚ずつの付与、キーワード検索による画像指定を経たメタデータ修正モード、の3通りを実装することとした。以下ではそれぞれのモードを、一括付与モード、個別付与モード、修正モードと呼ぶ。

そして、メタデータを付与する際に類似した画像間での付与の手間を軽減することと、統一したメタデータを付与できることを目指す。具体的には、既にシステム中に存在している画像と類似している画像に対してメタデータを付与する場合には、類似した画像からメタデータをコピーできるようにする。なお、本機能を実現するには類似画像検索エンジンを利用する。学校現場での使用を想定すると、イラストやグラフ等の様々な画像をハンドリングする必要がある。スケッチ(線画や略画)を入力してシステム側でスケッチとデータベース中の肖像画や風景画を検索し、55%の被験者が満足した結果を得たとの報告⁽⁹⁾があり、イラストやグラフ等であっても類似画像検索は有効性はあると考えられる。

そして、特にコンテンツビジネスを想定すると、複数人で本システムを利用することが考えられる。そのため、システムを利用するには認証することとし、メタデータを書き込み等で扱う際には排他制御を備えることとした。すなわち、メタデータ付与者に対してシステムが入力画面を表示する直前に当該メタデータについてWriteロックし、入力終了した時点でロックを解除する。

最後に、メタデータの一貫性であるが、次の2通りが考えられる。

- ①新規に画像にメタデータを付与する際の、独立したメタデータ同士
- ②既にデータベース中に存在する画像に付与されたメタデータと、メタデータが付与されていない画像に対して新たに付与されるメタデータとの間

①については先行研究⁽⁸⁾によって示されているため、本研究では②を取り扱うこととした。つまりデー

データベース内での一貫性ということになる。

3. システムの構成とインターフェイス

3.1 システム構成

本システムの構成図を図1に示す。クライアントはサーバの入出力部より PHP3によって動的に Web ページを作成し表示する。サーバの OS は RedHat Linux 6.2であり、システム全体は C 言語で記述し、Web サーバには Apache を使い、データベースには PostgreSQL 6.5を使用して、全てをフリーソフトで構築している。サーバの構成は、ユーザ認証部・画像蓄積部・メタデータ検索エンジン・類似画像検索エンジン・入出力部の5つから成る。メタデータの構造としては先行研究⁽⁸⁾に基づくものを採用した。

以下ではシステムの主な動作を説明する。画像そのものは静止画であれば JPEG 形式、動画であれば MPEG-1形式のものを扱い画像データベースに、画像に付与されたメタデータはメタデータデータベースに、画像から抽出した画像特徴量は画像特徴量データベースに、それぞれ保持されている。なお動画像に対する類似画像の扱いは、当該動画の先頭フレームを対象としている。まず、クライアントでは最初にユーザ ID とパスワードによって利用者の認証を行なう。次に、一括付与・個別付与・修正の各モードの選択を行なう。一括付与の場合は以下のように動作する。まず、クライアントコンピュータのハードディスクに存在する画像ファイル名を複数指定する。指定された画像のうち1つに対して画像特徴量を求め、画像特徴量検索エンジンはこれを元に画像特徴量データベースに対して検索を行なう。検索結果は画像 ID として得られ、画像データベースより当該画像そのものを得て、入出力部では PHP3によって画面の整形を行ってクライアント部に結果を返し、3.2で述べる入力画面の類似画像表示部分に表示する。そして、3.2で示す入力画面によってユーザは各メタデータを入力する。これを最初に指定した画像分だけ繰り返す。

個別付与の場合は以下のように動作する。まず、クライアントコンピュータのハードディスクに存在する画像ファイル名を指定する。その後は一括付与と同様に動作する。

修正モードではまず、ユーザがキーワードを入力すると、サーバの入出力部がこれを受け取ってメタデータ検索エンジンに渡す。次にメタデータ検索エンジンはこのキーワードを元にメタデータデータベースに対して検索を行なう。検索結果は画像 ID の列として得られ、画像データベースより画像そのものを得て、画像の列として入力画面に表示する。表示順序はスコアの高いものから順に低いものへと表示する。メタデータ検索エンジンは一般的なもので、入力されたキーワード間で and/or/not の指定が可能である。利用者が表示された画像のうち1つをクリックすると、当該画像に対応するメタデータに対してメタデータ検索エンジンは排他制御を行ない、3.2で示す入力画面へと遷移する。この際にも他のモードと同じく、類似画像検索エンジンを用いて似た画像を入力画面の類似画像表示部分に表示する。その後は個別付与モードと同じ動作をする。入力が終了した際には、当該画像に対応するメタデータに対してメタデータ検索エンジンは排他制御を外す。

メタデータデータベースには、画像 ID に対応したキーワードが保存されている。類似画像検索エンジンによって、ユーザが指定した画像を似た画像を抽出し表示することができ、後述するように似た画像の間でメタデータをコピーすることが可能になる。

類似画像検索エンジンは、最も単純であるが基本的なアルゴリズムを使用していることと、色空間に RGB ではなく HSI を利用している点を考慮に入れ、一般的な画像特徴量検索エンジン⁽⁹⁾に基づくものを利用した。一般に類似画像検索は1つの画像に対してデータベース中の全ての画像との照合を行なうため、計算コストを考慮する必要がある、精度と計算コストはトレードオフとなる。本研究においてはユーザとの対話型であることに鑑み、多少の精度を犠牲にして計算時間を優先させて、単純なアルゴリズムを選択した。また、RGB 色空間は明暗変化に弱い⁽¹⁰⁾ことと、また別の色空間である $L^*a^*b^*$ では RGB 色空間からの変換に計算コストがかかる⁽¹⁰⁾ため、HSI 色空間を選択した。このエンジンは画像を縦横ともにそれぞれ等間隔で8点を選び、各点の色値を後述する画像特徴量データベースに保存する。画像間の類似度を求める際には、この値をそれぞれの点で比較する。比較の際に求める

類似度をスコアと呼ぶ。

スコアの求め方は以下のとおりである。いま、入力画像を P、データベース内のある画像を Q とする。

画像 P, Q の HSI 空間内での色値はそれぞれ、

$$Pxy = (phxy, psxy, piry)$$

$$Qxy = (qhxy, qsxy, qiry)$$

($1 \leq x, y \leq 8$) となる。

P, Q の点 (x, y) における距離は

$$Dxy(P, Q) =$$

$$\sqrt[3]{phxy \times qhxy + psxy \times qsxy + piry \times qiry} \dots\dots (1)$$

と定義される。

$$\text{スコアは } S(P, Q) = \sum_{1 \leq x, y \leq 8} Dxy(P, Q) \dots\dots\dots (2)$$

と求められる。

ある入力画像 P に対して、データベース内の画像すべてにスコアが計算される。このスコアが高いほど、類似度が高い、つまり 2 つの画像はより似ているということになる。例えば 2 つの画像 Q と R が存在している場合に、入力画像 P に対してスコア S (P, Q) がスコア S (P, R) よりも大きい場合、画像 Q が画像 R よりも画像 P に似ているということになる。

次に、図 1 中の各データベースの構成を表 1 と表 2 に示す。

表 1 に示したメタデータデータベースには、画像 ID と当該画像に付加されたメタデータが格納されている。第 1 列には画像 ID、第 2 列には第 3 列のメタデータに対応する属性、第 3 列にはメタデータが一語ずつ格納されている。ひとつの画像に対し一般に複

数のメタデータが付加され、表 1 に示した例では、画像 ID 1 の画像に物体として「ビル」、背景として「青空」というメタデータが、画像 ID2 の画像には物体として「自動車」「ビジネスマン」というメタデータが、画像 ID3 の画像には物体として「ビル」、背景として「空」「曇り」というメタデータが付加されていることを示している。また、メタデータが付加されていない画像 ID については、当該画像 ID はこのデータベースの第 1 列には現れない。

表 2 に示した画像特徴量データベースには、画像 ID と画像のファイル名、および画像より予め抽出した画像特徴量が格納されている。画像 ID は本装置の中で一意に決定し、画像 ID と画像ファイル名との間是一对一に対応する。また、画像特徴量はバイナリ形式で格納される。表 2 に示した例では、画像 ID1 の画像のファイル名は D001_001.jpg であることを示している。なおこのデータベースは類似画像検索で使用するものである。

3.2 インターフェイス

本システムのインターフェイスを図 2 に示す。Web の画面により表示を行なう。画面の左側にメタデータを付与する対象となる画像が表示される。右側にキー

表 1 メタデータデータベース

画像 ID	属性	メタデータ
1	物体	ビル
1	背景	青空
2	物体	自動車
2	物体	ビジネスマン
3	物体	ビル
3	背景	空
3	背景	曇り
⋮	⋮	⋮

表 2 画像特徴量データベース

画像 ID	ファイル名	特徴量
1	D001_001.jpg	5f3da3...
2	D001_002.jpg	006cd4...
3	D001_003.jpg	0c0d0c...
⋮	⋮	⋮

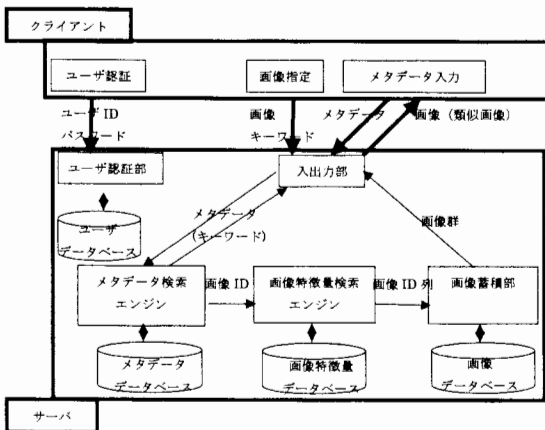


図 1 システム構成

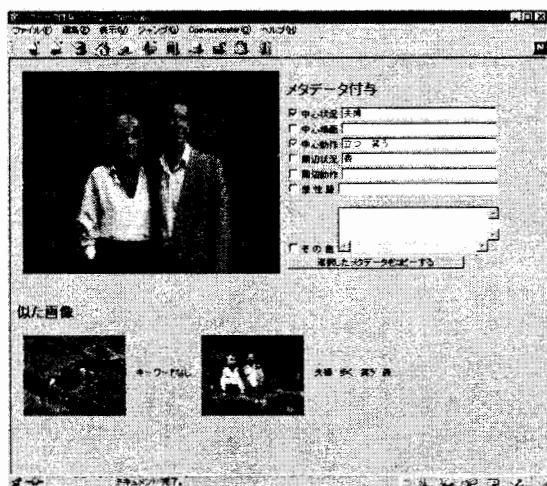


図2 システムインターフェイス

ワード入力部があり、ユーザはキーワードを入力する。また、類似画像検索ボタンにより、システムが当該画像に似た画像を検索しそして表示され、それぞれの画像にキーワードが存在する場合にはキーワードが表示される。チェックボックスによりキーワードを指定することで、該当のキーワードが付与対象である画像に対してコピーされる。これにより、似た画像の間でキーワードを参考にしたたり、入力する手間を軽減したりすることができる。

4. 評価実験

4.1 実験概要

本システムの有用性を確かめるために評価実験を行った。

被験者は成人男女合わせて10名（男性7名女性3名）であり、Web等の検索エンジンの使い方には習熟している。類似画像検索部における画像データとして、テクスチャパターン（幾何学模様）、会社の社内風景、山や海といった自然の風景、といった画像1000枚を用意した。それぞれの画像には統一した基準⁽⁸⁾に従ってメタデータを予め付与してある。

被験者には本システムによって実際にメタデータを付与してもらった。付与してもらった画像の一例を図3に示す。付与してもらった画像は動画像2種類・静止画5種類であり、各被験者に同じ画像を提示した。な



図3 評価に用いた画像例

おすべてのモードを用いており、被験者は付録に示すような、予め用意されたスクリプトに従って操作を行った。

手順としては、まず次に被験者にデータベース中に含まれる画像を1枚ずつ提示し、付与を行ってもらった。これを一括付与モードで5枚（動画像1枚・静止画4枚）、個別付与モードで2枚（動画像1枚・静止画1枚）の画像に渡って行った。最後に被験者にアンケートに答えてもらった。アンケートの項目としては、類似画像による付与のしやすさ、各モードの使い方、目的の画像に対するメタデータの付与しやすさ、もう少し使いたいのか、必要性を感じるか、総合評価、自由記述である。なお、類似画像による付与と各モードの使いやすさは本システムのための評価、その他の項目は先行研究⁽³⁾や1章で述べたような一般的と思われる入力手段との比べてどうかという意味でExcelによる付与との比較である。なお、モードについては各モードの使いやすさという項目ではなく各モードについて評価してもらうことも考えられる。しかし3.1で述べた一括付与・個別付与・修正という3つの状況に対して本システムでは3つのモードを用意したので、各モードを使い分けることについてどのような評価が得られるかを評価してもらうこととした。

4.2 実験結果

実験結果を表3に示す。各項目についてウィルコクソンの符号付順位検定を用いて実験結果に対する検定を行った。付与のしやすさや類似画像による付与については高い評価が得られた。自由記述からも「画像を見ながらキーワードを付与できるとわかりやすい」というユーザインターフェイスに関する意見が3名から得られた。また「画像の分類に有効であろう」という分類に関する意見が2名から得られた。類似画像によって似た画像に対するメタデータ付与を行なうこと

表3 評価実験の結果

項目	本システム	Excel	有意差
類似画像による付与	4.0	----	----
各モードの使い方	3.2	----	----
付与しやすさ	4.0	2.2	*
もう少し使いたい	3.6	2.0	**
必要性を感じるか	3.2	2.4	+
総合評価	4.0	2.2	*

+ $p<0.1$, * $p<0.05$, ** $p<0.01$

について、本システムは有効であることがわかった。しかし、各モードの使い方は高い評価ではなく、利用対象を絞り込んだ付与システムとする必要がある。付与しやすさと総合評価についてはExcelによる場合と比べて、有意に高い評価が得られた。自由記述にも「手間がかからなくなった」「便利である」との意見が得られた。付与しやすさについても、本システムが有効であることがわかった。

その他に自由記述からは、ユーザインターフェイスや分類に関する意見の他に、「それでも面倒である」「自分の写真なら日付だけで十分」という必要性を感じられないという意見も2名から得られた。評価項目の必要性を感じられるかの得点が他の項目より下がっていることから、必要性を感じない被験者も存在した。

この評価実験により、本システムにおけるメタデータ付与機能の有効性がおおむね確認できた。

5. まとめと今後の課題

本章では、画像データベースを構築するにあたりメタデータを付与する手間を軽減し、かつ、データベース内での一貫性のあるメタデータを付与するシステムを開発することを試みた。

今後の課題としては、複数のシーンを含む動画画像への対応がある。動画画像が長いものになると、メタデータを動画画像の一部に付与したり、MPEG等の動画画像の中間フレームでの画像へ類似画像検索を適用したりする必要がある。また家庭用ビデオカメラ等に付属した動画画像編集ソフトでは一般に⁽¹¹⁾⁽¹²⁾、タイムスケールを表示しながら当該動画画像の一部分を削除したりや

並び替えたりなどの作業を行なう。このようなインターフェイスでは、編集と同時にメタデータを付与することができる。その際に静止画とのシームレスなインターフェイスやデータベースへの格納といった課題を解決する必要がある。

そして、個人向けとしては日付や時刻または場所などといった、様々な観点でのグループ化を行なうことも必要である。特定の行事や旅行の際の動画画像や静止画に対してまとめてメタデータを付与するような機能が求められてくるとと思われる。

そして、LOMやDublin Coreなど様々なメタデータ規格があるが、それぞれに応じて様々なシステムを使えるようにする必要もある。この機能を実現するには、データベースの設計変更やユーザインターフェイスの変更など課題は多く、データベースに関する研究開発が待たれる。

(2003年3月8日受付)

参考文献

- (1) 永野和男：“教材コンテンツの開発とその普及の戦略”，日本教育工学会シンポジウム（2002）。
- (2) 照井武彦：“CDROMとネットワークによる民俗・歴史・考古画像データベースの提供方式の開発研究”，平成7年度～平成9年度科学研究補助金（基盤研究A-1）研究報告書（1998）
- (3) 篠原正典，徳畑香菜，岡本麻由美，三宅丈夫，永野和男：“小・中学校教育用学習素材検索システムの開発と児童・生徒の検索時における検索過程”教育システム情報学会誌 Vol.18, No.2, pp.200-209（2001）
- (4) 小野敦史，天野督士，斗谷充宏，佐藤隆，坂内正夫：“状態遷移モデルとシーン記述言語による自動キーワード付与機能を持つ画像データベースとその評価”，電子情報通信学会論文誌 D-Ⅱ Vol.J79-D-Ⅱ No.4, pp.476-483（1996）
- (5) Christopher Town, David Sinclair：“Content Based Image Retrieval using Semantic Visual Categories”，Technical Report TR2000-14, AT&T Laboratories（2000）
- (6) 新田直子，馬場口登，北橋忠弘：“放送型スポーツ

映像の構造を考慮した重要シーンへの自動アノテーション付け”, 電子情報学会論文誌 D-Ⅱ Vol.J84-D-Ⅱ No.8, pp.1838-1847 (2001)

- (7) 椋木雅之, 美濃導彦, 池田克夫: “対象物スケッチによる風景画像検索とインデックスの自動生成”, 電子情報通信学会論文誌 D-Ⅱ Vol.J79-D-Ⅱ No.6, pp.1025-1033 (1996)
- (8) 福本徹, 赤堀侃司: “画像データベースに適したメタデータの分析と評価”, 日本教育工学会論文誌 Vol.26 No.4, pp.337-348 (2003)
- (9) 西山晴彦, 松下温: “画像の構図を用いた絵画検索システム”, 情報処理学会論文誌 Vol.37 No.1, pp.101-109 (1996)
- (10) 高木幹雄, 下田陽久: “画像解析ハンドブック”, 東京大学出版会 (1991).
- (11) 上田博唯, 宮武孝文, 炭野重雄, 長坂晃朗: “動画像解析に基づくビデオ構造の視覚化とその応用”, 電子情報通信学会論文誌 D-Ⅱ Vol.J76-D-Ⅱ No.8, pp.1572-1580 (1993)
- (12) 馬場口登: “メディア理解による映像メディアの構造化”, 信学技報 IE99-18, pp.39-46 (1999)
- (13) 福本徹, 赤堀侃司: “画像データベースにおけるメタデータ付与システムの開発”, 第17回日本教育工学会大会講演論文集 (2002)

著 者 略 歴

福本 徹



1971年生まれ, 大阪府豊中市出身。1992年大阪大学基礎工学部情報工学科中退, 1994年大阪大学大学院基礎工学研究科物理系専攻情報工学分野修士課程修了, キヤノン株式会社入社。

2003年東京工業大学大学院社会理工学研究科人間行動システム専攻博士後期課程修了, 博士(工学), 現在に至る。画像検索, 教育工学の研究に従事。情報処理学会, 日本教育工学会, 教育システム情報学会, 各会員。



赤堀 侃司

1944年生まれ, 広島県呉市出身。1969年東京工業大学大学院理工学研究科物理学の修士課程を修了, 静岡県高等学校教諭, 東京学芸大学講師, 助教授, 東京工業大学助教授を経て, 平成

3年3月から現職。現在, 東京工業大学・教育工学開発センター, および大学院社会理工学研究科, 人間行動システム専攻教育工学講座に在籍。教育工学, 教育情報科学の研究に従事。工学博士。教育システム情報学会, 電子情報通信学会, 日本教育工学会各会員。