

MPEG-7 記述子と画像検索インターフェースの検討

青木秀一[†] 青木輝勝[†] 安田浩[†]

[†]東京大学大学院工学系研究科

〒153-8904 東京都目黒区駒場 4-6-1

TEL:03-5452-5277 FAX:03-5452-5278

s-aoki@mpeg.rcast.u-tokyo.ac.jp

近年、マルチメディアコンテンツの普及により画像検索の重要性が高まっている。筆者らは、画像検索においてユーザーの意図を適切にシステム側に伝達するインターフェースが重要であり、また検索効率を大きく左右するものと考えている。検索したい目標画像を一意のテキストで表現しにくい場合などは、略画による検索が適していると考えている。画像検索を行う際のユーザーのクエリー入力として略画を検討した際、ユーザーがシンボリックに用いられている、しかし実際のオブジェクトとは必ずしも形の上では似ていない略画を描くことがしばしばある。本稿ではメタデータとして2001年秋に標準化予定のMPEG-7を用いた際に、このようなシンボリックな略画をどのように画像検索に応用するかについて検討を行った。

キーワード：画像検索、略画、MPEG-7

A study on MPEG-7 descriptors and human interface for picture retrieval **Shuichi AOKI, Terumasa AOKI and Hiroshi YASUDA**

Graduate school of engineering, the University of Tokyo

4-6-1 Komaba Meguro-ku, Tokyo, 153-8904 Japan

TEL: +81 3 5452 5277 FAX: +81 3 5452 5278

The importance of technology for picture retrieval is increasing by reason of pervasion of multimedia content, these days. Human interface, which can transmit our intention to system, is important and it affects the efficiency of retrieval very much. We think that retrieval with rough sketch is best suited when we don't express our target picture uniquely with a word. Some people often sketch objects, which don't resemble their target picture as rough sketch for query. In this paper, if we use MPEG-7 that will be standardized in this fall we study appropriate process of symbolic rough sketch for picture retrieval.

Key words: picture retrieval, rough sketch, MPEG-7

1.はじめに

画像検索の要素技術としては、インデキシングとクエリー入力インターフェースに大きく分けることができ、前者はどのような特徴量を抽出するか、また、いかに省力化するかという観点から研究が進められている。一方、ユーザーインターフェースの観点からは大きく次のように分類できる。

a:テキスト(名詞句や動作語)を入力するもの

b:ラフスケッチや簡単な図形を配置し、ユーザーのイメージを構成するもの

aの場合は、「富士山」や「走っている人」など具体的な物を示す場合に加え、「明るい空」や「丸い果物」などのような主観を含む場合もある。一方、bのスケッチによるクエリー入力は厳密な検索に用いることは困難であるが、対象物を一意のテキストで表現しにくい場合や対象物についての記憶があいまいな場合には有用であると考えられる。しかし、ユーザーが画像検索におけるクエリーを入力する際、丁寧に細部までスケッチを描くことは期待できない。本稿では、メタデータとして2001年秋に標準化予定のMPEG-7を用いる場合に、どのようにクエリーを処理するかを検討した。本稿では次の2において既存の画像検索の問題点を提示し、3で略画を処理する方式を提案し、この提案方式を実装し実験したのでその結果について報告する。

2.従来技術の問題点

画像検索の研究は長年行われており、略画を検索クエリーに用いた研究例として、例えば[3]がある。[3]ではシルエット画像を対象とした画像検索プロトタイプシステムを構築している。ここでは、シルエット画像を凸分解し、それらの部分領域の属性及び領域間の関係を対象画像の特徴量とすることで、類似画像検索を実現し

ている。一方、[4]ではウェーブレット変換による多重解像度解析を用いることで、民族資料の画像データベースに対して類似画像検索を行ったことを報告している。しかし、これらはいずれも独自の特徴量を用いるという点で、汎用性に乏しい。また[3]ではユーザーの描いた折れ線に対し、属性を指示することができるが、これは人間の直感的な情報ではないと考えられる。

ここで筆者らは、例えば



図1：標準案内図用図記号(国土交通省)

のような一般化している記号を人間にとって直感的と考えた。

特徴量の汎用性については、様々なマルチメディアコンテンツに付与することができるメタデータとしてMPEG-7(ISO/IEC 15938)が2001年秋に標準化の予定である。ISO標準としてのMPEG-7を用いることで次のような点が改善されることが期待できる。

- メタデータ抽出の汎用ツールの開発が進むことで、インデキシングコストが下がる。
- メタデータが標準化されることで、広い互換性をもったマルチメディアデータベースを構築できる。

以上のような理由により、本稿ではMPEG-7をメタデータとして用いた場合に、人間にとって直感的であろう略画からの画像検索を実現するためのプロセスについて提案する。

3.本稿における提案方式

本稿では図2に描かれているような手書き略画をクエリーとして画像検索に用いることを目標としている。図2に描かれているオブジェク

トは本物の人間の形とは大きく異なっているが、この略画が人を表していることは多くの場合間違いないであろう。本稿ではこのような略画をシンボル略画と呼ぶことにする。シンボル略画を画像検索に応用する場合には次のようなアプローチが考えられる。

- a. シンボル略画に最適なメタデータをあらかじめ付与する
- b. シンボル略画に最適な類似度計算方法を用いる
- c. シンボル略画を適当に処理することで従来のメタデータと同様に扱う（シンボル略画に対しプリプロセスを導入しマッチングを行う）

a は汎用性に乏しく実現したとしてもローカルなマルチメディアデータベースで終わってしまうおそれがある。また b のシンボル画像に最適な類似度計算式を検討することは、最終的にシンボル画像の特徴を検討することであり、結果として c と同様にシンボル画像に対しプリプロセスを施すことにつながる。そのため今回は c のアプローチをとり、次のようなプリプロセスのアルゴリズムを検討した。

○閉曲線部分はそのままもしくは大きさを小さくする。

○線分を太くする。

どの程度大きくするか、また太くするかについては次の 3 以降で実験を通し明らかにする。

3.1 実験の方法

検索したいオブジェクトのおおまかな形を描き検索する際に有用である記述子として MPEG-7 には Part3 で RegionShapeDS と ContourShapeDS とが定義されている。輪郭図形のような略画の場合、ContourShapeDS を用いるのが簡便である。そこで本提案の有効性を確認するために、MPEG-7 Visual part FCD [1]

に従って、輪郭線から ContourShapeDS の記述子を生成するプログラムを C++を用いて実装した。ContourShapeDS は Eccentricity、Circularity と輪郭線の曲率から計算される xpeak、ypeak の値をメタデータとする記述子である。

また類似度計算には MPEG-7 Visual part of experimentation Model Version 9.0[2]で提示されている計算式を用いた。実際には、クエリー画像と比較画像の Eccentricity と Circularity の値を用いて global parameter を計算し、これらがあらかじめ定めた閾値以下であった場合には、ContourShapeDS の xpeak と ypeak の値を用いて、類似度 M を次のように計算している。

$$\frac{|c_q[0] - c_r[0]|}{\text{MAX}(c_q[0], c_r[0])} \leq Th_e$$

$$\frac{|c_q[1] - c_r[1]|}{\text{MAX}(c_q[1], c_r[1])} \leq Th_c$$

(今回は[2]を参考にして、

Th_e=0.6、Th_c=1.0 とした)

$$M = 0.4 \times \frac{|c_q[0] - c_r[0]|}{\text{MAX}(c_q[0], c_r[0])} + 0.3 \times \frac{|c_q[1] - c_r[1]|}{\text{MAX}(c_q[1], c_r[1])} + Mcss$$

$$Mcss = \sum_1 ((xpeak[i] - xpeak[j])^2 + (ypeak[i] - ypeak[j])^2) + \sum_2 (ypeak[i])^2$$

ここで $c_q[0]$ 、 $c_r[0]$ はそれぞれクエリー画像、比較画像の eccentricity の値、 $c_q[1]$ 、 $c_r[1]$ はそれぞれクエリー画像、比較画像の circularity の値である。

この類似度計算式ではクエリー画像と比較画像が完全一致の場合に 0 になり、値が大きくなるほど形状が似ていないことを意味する。

3.2 実験対象の画像

まず人間を表す手書き画像として次の二枚を用意した。



図2：略画1



図3：略画2

また比較画像として、自然画から実際の人間の画像、またデフォルメされているアニメーション画像、比較的人間に近い形をしているアニメーション画像の合計6枚を用意した。(図4～図9) これらの比較画像についてはそれぞれの画像の右側に示すように対象オブジェクトの切り出しを行いその輪郭線から MPEG-7 ContourShapeDS を抽出した。



図4 sizen1 (目標とするオブジェクトは左側を歩いている人間)



図5 sizen2 (目標とするオブジェクトは左側にいる後ろ向きの子供)



(自然画像)



図6 an1



図7 an2

(デフォルメされているアニメーション画像)



図8 cha1



図9 cha2

(実際の人間に近いアニメーション画像)

なお略画から画像検索を行う場合には一般的にはオブジェクトの位置合わせや大きさを合わせた上で類似度計算を行う必要があるが、ContourShapeDS を用いたために、画像内でのオブジェクトの位置の調整やオブジェクトの大きさをあわせるといった作業は不必要なため行っていない。

3.3 プリプロセスを行わない場合の実験(実験1)

3.2 で述べた略画1・2と各比較画像との類似度を表1-1～1-3に示す。実験1では類似度計算の前に略画を加工していない。

	sizen1	sizen2
略画1	1.856	0.982
略画2	2.700	1.515

表1-1

	an1	an2
略画1	6.321	1.266
略画2	9.418	3.073

表 1-2

	cha1	cha2
略画1	2.390	1.108
略画2	3.763	1.746

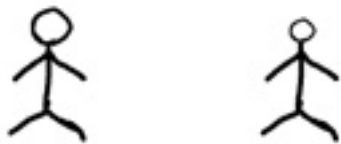
表 1-3

3.4 頭と腕に対し別々にプリプロセスを行った場合の実験（実験2）

次に、手書きの略画1・2に対して、頭の大きさを（縦横ともに）80%、50%にしたもの、また腕の太さ（幅）を120%、140%、160%に拡張したものと比較画像との類似度を計算した。この処理を施した後の略画は次の通りである。



略画1の頭を80%、50%に縮小したもの



略画2の頭を80%、50%に縮小したもの



略画1の腕を120%、140%、160%に拡張したもの



略画2の腕を120%、140%、160%に拡張したもの
これらの処理を施した略画と各比較画像との類似度を表2-1～2-3に示す。

	sizen1	sizen2
略画1	1.856	0.982
略画1 頭0.8	1.511	0.838
略画1 頭0.5	1.340	0.923
略画1 腕1.2	1.895	1.014
略画1 腕1.4	2.017	1.100
略画1 腕1.6	1.118	0.413
略画2	2.700	1.515
略画2 頭0.8	2.302	1.536
略画2 頭0.5	1.336	0.869
略画2 腕1.2	0.977	0.526
略画2 腕1.4	1.812	0.967
略画2 腕1.6	2.340	1.439

表 2-1

	an1	an2
略画1	6.321	1.266
略画1 頭0.8	8.027	3.006
略画1 頭0.5	7.380	0.890
略画1 腕1.2	6.311	1.310
略画1 腕1.4	6.357	1.415
略画1 腕1.6	1.036	0.365
略画2	9.418	3.073
略画2 頭0.8	13.05	1.123
略画2 頭0.5	3.549	0.873
略画2 腕1.2	0.923	0.826
略画2 腕1.4	3.901	1.164
略画2 腕1.6	11.83	2.410

表 2-2

	cha1	cha2
略画1	2.390	1.108
略画1 頭0.8	1.961	0.763
略画1 頭0.5	2.285	0.763
略画1 腕1.2	2.425	1.096
略画1 腕1.4	2.490	1.084
略画1 腕1.6	1.340	1.378
略画2	3.763	1.746
略画2 頭0.8	3.204	1.659
略画2 頭0.5	1.663	0.752
略画2 腕1.2	1.213	0.989
略画2 腕1.4	2.566	0.841
略画2 腕1.6	5.631	3.505

表 2-3

3.5 プリプロセスを行った場合の実験(実験3)
 実験1、2を参考にして頭の拡大縮小と腕の幅を同時に変化させた場合の類似度を表3-1～3-3に示す。

	size1	size2
略画1 頭0.5 腕1.6	1.150	0.517
略画2 頭0.5 腕1.6	2.395	1.579

表 3-1

	an1	an2
略画1 頭1.0 腕1.3	2.274	1.660
略画2 頭1.0 腕1.3	3.699	0.905

表 3-2

	cha1	cha2
略画1 頭0.5 腕1.5	1.475	0.739
略画2 頭0.5 腕1.5	1.196	1.037

表 3-3

3.6.考察

まず実験1の結果より、いずれの比較画像に

おいても略画2よりも略画1の方が類似度Mの値が小さく、より比較画像に近いことが分かる。略画1と略画2の違いは両腕が一つの線分で表されるか、二つの線分で表されるかの違いであるが、ContourShapeDSのnumber of peaks記述子の値を参照すると略画2では25であったのに対し、略画1では19であった。一方、比較画像の方は、11～19までの間に分布していた。そのため、略画2の類似度Mの値が大きくなったものと考えられる。

次に、略画にプリプロセスを行った実験2の結果より、処理を加えることによりすべての場合で類似度の値を小さくすることができることが分かる。しかし、その方法は一定ではなく、逆に類似度が大きくなってしまう場合があることも同時に分かる。その中で、頭の部分を縮小することはデフォルメされたアニメーション画像(an1、an2)以外ではおおむね有効であった。腕の幅については、120%に拡幅することで多くの場合の妥協点を見いだすことができた。

また実験3では頭と腕の幅の両方を同時に変化させているが、これらの結果と実験1、実験2との結果から今回のサンプル例においては次のことが言える。

- 自然画に対しては腕幅をそのままに頭の大きさだけを50%にすることが最適である
- デフォルメされたアニメーション画像に対しては頭の大きさをそのまま腕幅を130%にすることが最適である。
- 比較的人間の形に近いアニメーション画像に対しては頭の大きさを50%、腕幅を150%にすることが最適である。

以上より、シンボル略画のプリプロセスとして、データベースに蓄積されている画像の種類によってプリプロセスのパラメータを変えることで

最適な結果を得られることが分かる。また、このようなプリプロセスを導入することによりシンボル略画を画像検索のクエリーとして用いることが可能になる。

4.まとめ

本稿ではシンボル略画を画像検索のクエリーとして用いるためのプリプロセスについて、特に人を表すシンボル略画に対し、どのようなプリプロセスが適当であるかを検討した。その結果、プリプロセスを行うことによりシンボル略画の類似度の値を小さくすることが可能なことが明らかになった。

5.今後の課題

シンボル略画に対するプリプロセスは、本来示しているオブジェクトに対してはより類似度の値が小さくなるようにすることが必要であるのはもちろんだが、同時に本来意図していないオブジェクトに対しては類似度の値を大きくできることが望ましい。また、このようなシンボリックな意味を持つ略画は人を表す物だけではない。例えば、次のような略画についても同様の処理が検討されるであろう。



これらは左側の図は花を描いたものであろうし、右側は車を描いたものであろう。しかし、実際のオブジェクトとは似てないことが多いことが予想される。今後は MPEG-7 の IS に沿った形でメタデータを生成した上で実用的なマルチメディアデータベースを構築し、どの程度の略画であれば実用的な検索になるのか検証を進めたい。また検索の際の候補画像が増加してしまうことが予想されるが、候補画像の絞り込みの方

法についても検討を行う必要がある。

参考文献

- [1]ISO/IEC JTC1/SC29/WG11/N4062 “Text of ISO/IEC 15938-3/FCD Information technology Part 3 Visual”
- [2]ISO/IEC JTC1/SC29/WG11/N3914 “MPEG-7 Visual part of experimentation Model Version 9.0”
- [3]金原史和、佐藤真一、濱田喬：形状分解によるユーザーの視点に基づいたシルエット画像検索、情報処理学会論文誌 Vol36 No.12 pp.2800-2810(1995)
- [4]小早川倫広、星守、大森匡、照井武彦：ウェーブレット変換を用いた対話的類似画像検索と民族資料データベースへの適用、情報処理学会論文誌 Vol40 No3 pp.899-911(1999)