

## オブジェクト抽出・領域分割の比較手法に関する検討 A Study of Comparison Scheme of Object Extraction / Region Segmentation for Object Based Coding

高木 幸一  
Koichi Takagi

小池 淳  
Atsushi Koike

松本 修一  
Shuichi Matsumoto

### 1. はじめに

MPEG-4 の Visual Part[1]でオブジェクト符号化に関する枠組みが採用されたことにより、それを行うためのオブジェクト抽出、あるいは領域分割技術の発展が望まれている。そこで、筆者らは、2次元映像信号を入力としたオブジェクト符号化を行うための、汎用的なオブジェクト抽出および領域分割手法に関する検討を行っている。

一方、コンピュータビジョンの分野では、オブジェクト抽出・領域分割に関する検討が数多くされてきている。また、対象・用途が限られたものに関しては特に深く検討が進められており、それに関連する製品も、極めて精練されたものが市場に存在する。

そこで、これらの技術を応用し、すなわち、複数のオブジェクト抽出・領域分割手法（のエッセンスとなる部分）を統合することで、精度の観点で最適な汎用オブジェクト抽出・領域分割手法を構築することが可能であると考えられる。

ところで、これを行うためには、各方式によるオブジェクト抽出・領域分割結果がどのくらい類似しているのかを測る指標が必要となる。しかし、これらの抽出・分類結果

の類似度を比較する方式については、あまり検討がされていない。

そこで、本稿ではこの類似度の指標に関し検討を行う。

### 2. 従来手法とその課題

まず、本検討の概略を示す。入力として2次元映像信号を与えたとき、それに複数のオブジェクト抽出・領域分割手法を適用することにより、図1のようにそれぞれ異なるオブジェクト抽出・領域分割結果が得られたとする（ここでは、簡単のため2種類の手法の比較とする）。その結果がどのくらい類似しているのかを測定するための手法の検討が本稿の目的である。

最も容易な方法として、任意の2画像から抽出された形状信号の L1(L2)ノルムを取る方法が考えられる。しかし、この方法は画像の特徴（形状全体のバランス）等を全く考慮していないため、これをもって類似度と呼ぶにはふさわしくない。

一方、MPEG-7 の Visual Part[2]では、映像信号のさまざまな特性を効率的に記述するツールが規定されている。これにより、主に大量のデータベースから意にかなった映像を検索したり、映像を識別したりするなどの機能を実現することが可能となる。また、複数の映像信号の類似度の指標をこの中で使われている記述子の情報から導出するための手法が非標準情報として報告されている[3]。本標準は、従来画像検索・識別の検討で培われてきた技術をふんだんに活用しているため、キーとなる情報を映像信号のみから抽出し類似度を計測するためには極めて効率的なツールであると言える。そこで、本検討ではこのツールを可能な限り有効に活用することとする。

ところで、本検討はあくまで抽出・分割された領域の形状の類似度を検証することが目的であるため、[2][3]の中でも特に形状情報を記述するツール (Shape description tools) ,および位置情報を記述するツール (Localization tools) に着目すればよい。

Shape description tools は、比較的簡単な閉曲線の輪郭線形状情報を記述するための Counter Shape Descriptor と一般的なオブジェクト形状を記述するための Region Shape Descriptor, および3次元形状に対応した 3D Shape Descriptor の3つからなる。

また、Localization tools は1画面（静止画）のみを考慮したものと、時空間を考慮したものが存在する。

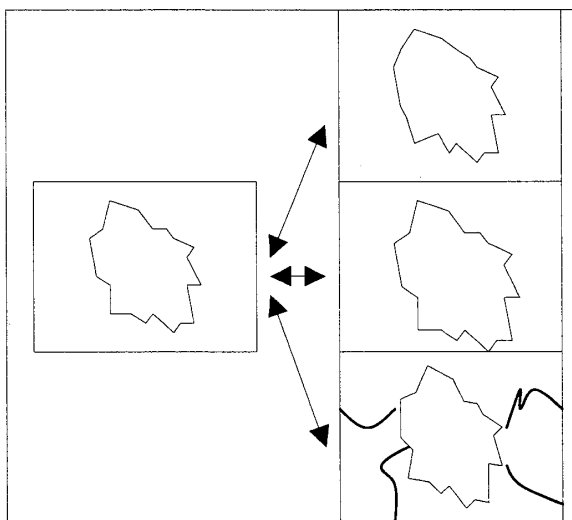


図1. どれが一番類似度が高い？

† (株) KDDI 研究所, KDDI R&D Labs. Inc.

この両者を用いることで、本稿における課題をほぼ達成できると考えられるが、以下の点で問題が残る。

- ・ オブジェクト符号化への適用を考える必要があるが、これらの記述子のみでは符号化に対する考慮が一切なされていない。
- ・ これらの類似度計測手法は抽出されたオブジェクトの比較を対象とするものであるため、領域分割の結果を比較するためには何らかの工夫が必要となる。(例えば、画像全体をクラスタリングした結果とオブジェクト抽出結果を直接比較することができない。)

### 3. 提案手法

前節の問題点を踏まえ、本稿における課題を解決するためのポイントを整理すると、以下のようになる。

1. 領域の位置は極めて重要な因子である。それに基づきスケール等も関与すると考えられる。
2. 形状による符号化難易度に依存する指標が必要である。
3. 前段が領域分割・オブジェクト抽出のいずれであるかにより、比較の方法を別に検討しないとイケない。これらの条件を満足すべく、以下の手法を用いて類似度を決定することとする。

#### 3.1. 領域分割・オブジェクト抽出の比較に関する場合分け

はじめに、3の課題を解決するために以下の場合分けを行う。領域分割をRS, オブジェクト抽出をOEと書く。また、RSの結果を $A_i$ ,  $B_i$  ( $i$ は領域のインデックス)等で表す。

##### (1)OEの結果とOEの結果の比較

何もせず、両者の類似度を3.2節を用いて計算する。また3.2節における $D$ は0とする。

##### (2)OEの結果 $X$ とRSの結果 $A_i$ の比較

$Y = \phi$  (空集合)とする。 $X \cap A_i \neq \phi$ となる $A_i$ に対し、 $X$ と $A_i$ の共通部分が $X$ あるいは $A_i$ の大部分を占める場合を抽出し $Y$ に加える。すなわち、

$$\text{Area}(X \cap A_i) / \text{Area}(A_i) < Th$$

または

$$\text{Area}(X \cap A_i) / \text{Area}(X) < Th$$

(ここで $\text{Area}(K)$ は領域 $K$ の面積、 $Th$ は $Th < 1$ なる閾値である。)となる $A_i$ に対し、 $Y = Y + A_i$ とする。そして、 $X$ と $Y$ の類似度を3.2節を用いて計算する。なお、3.2節の $D$ は( $Y$ に含まれる領域数) / (全領域数(つまり $A_i$ の個数))とする。

##### (3)RSの結果 $A_i$ とRSの結果 $B_j$ の比較

まず、画面全体を $Z$ とする。もし、画面中に注目すべき領域があったら、まず、その部分を囲み(これは適当でよい)、 $Z$ を囲まれた領域と置き換える。

次に、 $A_i \cap Z \neq \phi$ となる $A_i$ の個数 $\alpha$ と $B_j \cap Z \neq \phi$ となる $B_j$ の個数 $\beta$ を比較し、 $\alpha < \beta$ の場合 $A_i$ と $B_j$ をすべて入れ替える(この操作により、 $Z$ と領域を共にする $A_i$ の個数より $B_j$ の個数の方が多くなる。)

そして、 $B_j \cap Z \neq \phi$ となるすべての $B_j$ に対し、 $B_j = X$ と置き、上記(2)と同じ操作を施す。この操作により3.2節を使うことで全 $B_j$ に対し類似度が求められるので、これの平均値を類似度として出力する。

### 3.2. 類似度関数の定義

類似度関数 $E$ は、上記ポイント1,2,3より、以下の式を用いることとする。

$$E = w_l L + w_s S + w_c C + w_d D$$

$L$ は位置情報の類似度、 $S$ は形状表現の類似度、 $C$ は符号化難易度の類似度である。また、 $D$ は、上で求めた領域数に基づく指標である。これらの関数の取りうる値はすべて $[0,1]$ で正規化されているものとし、重み係数 $w_l, w_s, w_c, w_d$ を適宜決定する(ただし $w_l + w_s + w_c + w_d = 1$ )。なお、各指標は、類似しているほど0に近い値を出力することとするが、仮に逆(つまり1に近い値ほど類似度が高い)の場合には1からその値を引いた値を指標とすればよい。この仮定により一般性は失われない。

ここで、各指標の例として $L, S$ は[3]で述べられている類似度指標を、 $C$ は形状情報符号量(もちろん、本来ならば実際に符号量を求めるのが理想であるが、それが困難な場合は、形状の複雑度をフラクタル次元等を用いて求めたり、[4]の形状情報見積もり法を用いるなどの方法で解決される。)を用いることで本課題が解決される。もちろん、他の方法も考えられるが、これらの指標の有効性に関する検討は今後の課題とする。

なお、本検討を深めていくことで、以下の分野への応用が考えられる。

- ・ ある映像信号に対し、それをオブジェクト符号化するための極めて良好なオブジェクト抽出・領域分割結果がわかった場合、別の抽出・分割手法のパラメータ設定のための参考に本指標を使うことができる。
- ・  $w_c = 0$ とおくことにより、符号化を除いた一般的なオブジェクト抽出・領域分割の類似度を評価することもできる。

### 4. おわりに

本稿では、オブジェクト符号化のためのオブジェクト抽出・領域分割結果の類似度計測手法に関する検討を行った。まず、本検討に際し、極めて有効なツールであると考えられるMPEG-7 Visualの2つのツールを取り上げ、その課題を整理し、それを解決するための手段に関する提案を行った。今後の課題として、実験等による本提案の有効性に関する検証が挙げられる。

なお、本研究は、通信・放送機構より委託の「ISDB技術に関する研究開発」に基づき行われた。

#### 参考文献

- [1] ISO/IEC 14496-2:2001 "Coding of audio-video objects -Part 2 Visual".
- [2] ISO/IEC 15938-3 "Multimedia content description interface Part 3: Visual".
- [3] ISO/IEC JTC1/SC29/WG11 N4579, "Text of ISO/IEC 15938-8 DTR (Extraction and Use of MPEG-7 Descriptions)," 2002.3.
- [4] 高木他, "動画像の高効率領域分割符号化方式," 2000 映像全大, 2000.8.