

フラクタル符号化による動画像領域分割に関する研究

†上繁 義史 †黒木 祥光 ‡太田 諦二

†鹿児島工業高等専門学校 ‡九州工業大学 工学部

あらまし

井田らは、IFSの応用として、縮小変換の逆過程により、離散的なダイナミカルシステムを定義した。この軌道がトラッピング領域に引き込まれる性質を利用して、Basin図を求め、領域分割に適用した。著者らは、これを動画像に適用し、トラッピング領域とフラクタル次元を用いた方法を提案した。この手法では、フラクタル次元を用いない場合に比べて、粒状雑音を削減し、領域の形状が確認できるようになったが、十分な精度とは言えなかった。そこで、本研究では、ラベルの粒状雑音の分布に注目し、ラベルの更新を行う手法を提案する。計算機シミュレーションにより、粒状雑音を低減させ、分割精度が向上することを示す。

A Study on Video Segmentation using Fractal Coding

†Yoshifumi UESHIGE Yoshimitsu KUROKI ‡Teiji OHTA

†Kagoshima National College of Technology

‡Faculty of Engineering, Kyushu Institute of Technology

Abstract

In this paper, we propose a modified scheme of video segmentation using IFS coding and fractal dimension. Ida and Sambonsugi have proposed a new segmentation scheme as the application of IFS coding. They applied the feature that orbits of the inverse process of the IFS attract to trapping regions. We have proposed a method of video segmentation using IFS coding and fractal dimension. This method can reduce granular error, and shape of regions can be distinguished compared to the segmentation without fractal dimension. However, the accuracy of the segmentation results is not enough. Therefore, we propose a scheme of reduction for granular error. The simulation results show its validity.

1 はじめに

Idaらは、IFS [1] に保存されている、幾何構造を応用した、領域分割法を提案した [2]。この手法では、[3]の方法により、画像についてIFSを求め、この逆写像を用いて、ダイナミカルシステムを定義している。このダイナミカルシステムの軌道を計算し、収束したトラッピング領域によりBasin図を求めている。

筆者らは、3D-IFS [4]を用いた動画像への拡張を検討している。3D-IFSのみを用いて、分割を行った

場合、主にテクスチャ領域において、粒状の雑音が発生し、領域の境界を十分に識別することが出来なかった [5]。そこで、3D-IFSとフラクタル次元を用いる手法を提案した [6]。この手法では、[5]の手法に比べて、領域の境界をある程度分離することができたが、粒状雑音が十分に削減できなかった。

そこで本論文では、この改良として、各画素について、近傍画素のラベルを参照して、ラベルの更新を行う方法を提案する。計算機シミュレーションにより、その有効性について検討を行ったので報告する。

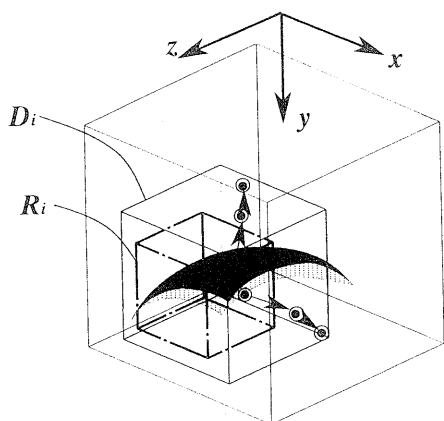


図 1: ダイナミカルシステムの軌道

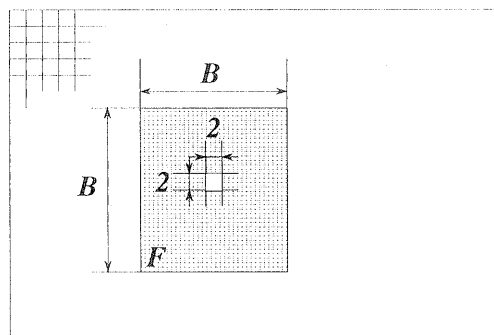


図 2: 局所的フラクタル次元の計算

2 IFS を用いた動画像の領域分割法

領域分割は、3D-IFS により、トラッピング領域の計算、局所的なフラクタル次元の計算、これらによるベクトルを構成してのクラスタリングの3段階で行われる。

トラッピング領域のみで領域分割を行った場合、画像を近似する IFS の精度に依存することになる。だが、IFS の探索範囲は後述するように、大きく取ることが出来ないため、画像を十分に近似できるとは限らない。

一方、フラクタル次元は、テクスチャの不連続性の程度に関する情報をもつ。すなわち、次元が低ければ、滑らかな曲面であり、逆に、次元が高ければ、複雑な形状をもつ曲面といえる。

そこで、フラクタル次元を用いて、テクスチャに関する特徴を補い領域分割を行う [6]。

2.1 トラッピング領域の計算

まず、動画像を近似する IFS を求める。IFS は Lazer の 3D-IFS [4] を用いる。ただし、符号化において、ドメインブロックの探索範囲をレンジブロックの近傍に制限し、輝度のスケール係数とオフセットの値の範囲を制限する。

3D-IFS において、位置に関するアフィン変換は

以下のようになる。

$$f_i \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r_i \cos \theta_i & -r_i \sin \theta_i & 0 \\ r_i \sin \theta_i & r_i \cos \theta_i & 0 \\ 0 & 0 & r_i (-1)^{\tau_i} \end{pmatrix} \times \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} e_i \\ f_i \\ g_i \end{pmatrix} \quad (1)$$

式 (1) の逆変換を用いて、ダイナミカルシステムを以下のように定義する。

$$M(x, y, z)^T = f_i^{-1}(x, y, z)^T, (x, y, z) \in R_i \quad (2)$$

式中の R_i はレンジブロックである。初期点として、IFS のアトラクタ上の点を用いる。この場合は、画像上の各点がそれに相当する。初期点を (x_0, y_0, z_0) として、

$$(x_n, y_n, z_n)^T = M^{on}(x_0, y_0, z_0)^T \quad (3)$$

により変換する。これを画像上の各点について行い、トラッピング領域を求める。エッジに近い初期点の変換点は、式 (3) により、元のエッジから遠ざかる性質をもつ。また、 $\theta = 0$ とすれば、写像された点は、エッジを越えない。したがって、閉曲線の内部の点は、同一のトラッピング領域に引き込まれる (図 1 参照) [2]。

2.2 フラクタル次元の計算

局所的なフラクタル次元の値を求める。画像を 2×2 画素のブロックに分割し、それぞれのブロック

クについてフラクタル次元の値を求める。ただし、ブロック内で次元の値は一定とする。図 2 に示すように、それぞれのブロックを中心として、 $B \times B$ の領域 F に対して計算を行う。なお、フラクタル次元は、Peleg らの方法 [7] により求める。

2.3 セグメンテーション

上で求めた、各座標に対するフラクタル次元の値 $FD(x_0, y_0, z_0)$ と、その点を変換して得られたトラッピング領域内の点 (x_t, y_t, z_t) を成分として、ベクトル $(x_t, y_t, z_t, FD(x_0, y_0, z_0))$ を構成する。このベクトルの集合を LBG アルゴリズムにより、 γ 個のクラスタに分類する。ただし、ベクトル $\mathbf{A} = (x_{nA}, y_{nA}, z_{nA}, FD(x_{0A}, y_{0A}, z_{0A}))$ 、 $\mathbf{B} = (x_{nB}, y_{nB}, z_{nB}, FD(x_{0B}, y_{0B}, z_{0B}))$ 間の距離として、

$$d(\mathbf{A}, \mathbf{B}) = \sqrt{(x_{nA} - x_{nB})^2 + (y_{nA} - y_{nB})^2 + (z_{nA} - z_{nB})^2} + \lambda |FD(x_{0A}, y_{0A}, z_{0A}) - FD(x_{0B}, y_{0B}, z_{0B})| \quad (4)$$

を定義する。式 (4) は、トラッピングされた点の平方自乗距離と、フラクタル次元の絶対差分との重み付き線形和となっている。定数 λ は、フラクタル次元に対する重み係数である。

得られた各クラスタのラベルを初期点 (x_0, y_0, z_0) に割り当てることにより、Basin 図が求まり、これが分類結果となる。

3 ラベルの更新

上述の手法により、領域分割を行うと、図 3 (b) のようになる。ただし、原画像は QCIF の動画画像 Susie (図 3 (a)) を用いている。また、式 (4) において、 $\lambda = 7$ としている。図 3 (b) から、粒状雑音の発生している領域において、それぞれの領域に、特定の数種のラベルが分布していることが分かる。これは、テキスト領域を初期点とする軌道は必ずしも、2.1 で述べた仮定を満たしていないためと思われる。

ここで、各画素の縦、横 ± 2 画素、前後 ± 1 フレームの範囲で、分布しているラベルの個数と、その中で最大の画素数をもつラベルの発生確率の関係を求めたところ、図 3 のようになった。この図から、全画素の 78.6% の各近傍において、支配的なラベル



(a)



(b)

図 3: (a) 原画像 Susie (第 1 フレーム), (b) その分割例

が存在することが分かる。そこで、各領域において、最大個数を取るラベルを用いて更新することにより、粒状の雑音を削減できると考えられる。

本手法では、各画素において、近傍画素のラベルの分布を求め、その中で、最大となるものを用いて、更新を行う。今、画素の座標 (x_0, y_0, z_0) と、その近傍領域を

$$P(p, q) = \{(x, y, z) | x_0 - p \leq x \leq x_0 + p, y_0 - p \leq y \leq y_0 + p, z_0 - q \leq z \leq z_0 + q\} \quad (5)$$

とする。 p, q の値により、近傍領域の大きさを決定することが出来る。 P に含まれる画素のラベルを $\{label_i | i = 1, 2, \dots, l\}$ とする。 P の部分集合で、ラベル $label_i$ をもつ画素全体の集合を P_i とし、集合の要

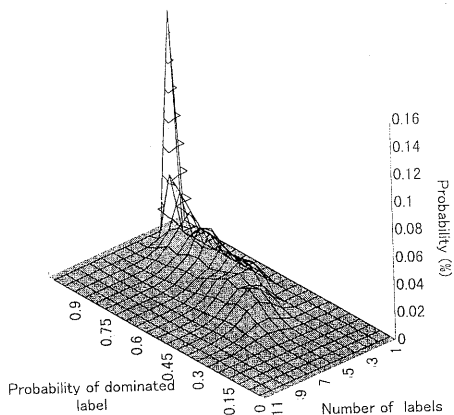


図 4: 支配的なラベルの発生確率

素の個数を $|\cdot|$ と表すと、求めるラベル $label(x_0, y_0, z_0)$ は

$$label(x_0, y_0, z_0) = \max_i \{P_i | P_i \subseteq P(p, q)\} \quad (6)$$

となる。この処理を全画素に対して行う。

4 シミュレーション

原画像として、画像 Susie の 1-32 フレームを用いる。

3D-IFS の計算において、ブロックサイズを 8 の固定サイズとし、縮小写像はフレーム内、フレーム間方向の縮小、および輝度値方向の縮小、オフセットで構成し、フレーム内、フレーム間方向の縮小率は $r_i = 0.5$ 、輝度方向の縮小率は 0.9 とする。また、輝度のオフセット値は 0-30 とする。ドメインセルの探索範囲は、レンジセルを中心として、 ± 4 の範囲とした。フラクタル次元の計算において、 $B = 8$ とした。LBG アルゴリズムにおいて、 $\gamma = 32$ とする。なお、定数 λ の値は、0, 5.0, 7.0, 10.0 の 4 通りとした。ラベルの更新において、 $p = 2, q = 1$ とした。

シミュレーションの結果を図 4 に示す。この図において、図 3 (b) で見られた、粒状雑音が大きく削減されていることが分かる。

図 4(a) はフラクタル次元を用いない場合である。女性の髪の毛の領域形状がほぼ失われており、背景と同じラベルとなっていることが分かる。また、受話器の部分でも、同様の傾向が見られる。これは、3D-IFS の近似精度に依存するためと考えられる。ラベルの

更新は、フラクタル次元を用いない場合には、十分な効果が得られないことを示す。

図 (b) では、図 (a) と比較して、女性の髪の毛の一部が領域として検出できていることが分かる。また、受話器の形状が検出できるようになった。これは、フラクタル次元値の変化が比較的大きい領域であることから、クラスタリングにより分離できていたためと考えられる。また、予備実験により、この領域において、近傍画素のラベルの個数 l が比較的小さいことが分かっており、このような領域では、提案法は正しいラベルに更新するものと考えられる。

図 (c), (d) において、フラクタル次元の影響が図 (b) よりも大きくなっており、画像の上部における、女性の髪の毛の領域と背景が分離できている。特に図 (c) では、画像の右上部に見られる、円形の領域が検出できており、提案法が有効に作用していることが分かる。図 (d) では、女性の顔と首の領域を分離している。また、画像の右上部の女性の髪の毛の領域が図 (c) では 1 領域となっているのに対して、2 領域に分割されている。これは λ の値が大きいため、フラクタル次元値の変化の影響を受けやすくなったことが原因と考えられる。

図 (a),(b),(c),(d) に共通して、女性の右手、および画像の右下部の女性の髪と肩の境界が検出できていないことが分かる。この領域全体では、特定の種類のラベルにより構成されていたが、局所的には、各ラベルの密度が一樣ではなかったために、結果として、別の領域に割り当てられたラベルが選ばれたものと思われる。

5 まとめ

本論文では、IFS 符号化とフラクタル次元を用いた、動画の領域分割法の改良として、粒状の誤差の除去について検討を行った。各画素とその近傍において、最大個数となるラベルを用いて、ラベルの更新を行う手法を提案した。

その結果、粒状雑音が低減され、各領域を分離できることが確認された。しかしながら、一部の領域において、孤立した小領域が発生し、領域形状が消失することが分かった。

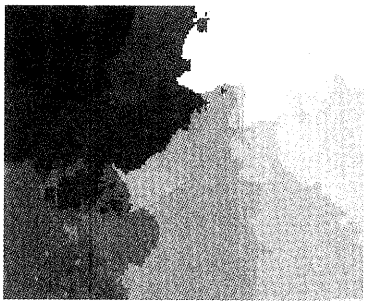
今後は領域形状の消失を防ぐために、ラベルの更新において、近傍画素のラベルを全て参照して判定を行う手法について検討を行う予定である。



(a) $\lambda = 0$ の場合



(b) $\lambda = 5$ の場合



(c) $\lambda = 7$ の場合



(d) $\lambda = 10$ の場合

図 5: シミュレーション結果

参考文献

- [1] M. F. Barnsley, *Fractals Everywhere second edition*. Academic Press, 1993.
- [2] T. Ida and Y. Sambonsugi, "Image segmentation using fractal coding", *IEEE Trans. Circuits Syst. for video technology*, **5**, pp. 567-570, Jun. 1995.
- [3] A. E. Jaquin, "Image coding based on a fractal theory of iterated contractive image transformations," *IEEE Trans. Image Processing*. vol. 1, pp.18-30, Jan. 1992.
- [4] M. S. Lazer and L. T. Bruton, "Fractal block coding of digital video", *IEEE Trans. Circuits Syst. for video technology*, **4**, pp. 297-308, Jun. 1994.
- [5] 上繁義史, 黒木祥光, 太田諦二, "IFS 符号化を用いた動画像の領域分割に関する一検討", 信学技報 (画像工学), Vol. 98, IE-98-33, pp.25-30, Jul. 1998.
- [6] 上繁義史, 黒木祥光, 太田諦二, "フラクタル符号化を用いた動画像の領域分割に関する研究", 画像符号化シンポジウム, PCSJ'98, pp.27-28, Oct. 1998.
- [7] S. Peleg, J. Naor, R. Hartley, D. Avnir, "Multiple resolution texture analysis and classification", *IEEE Trans. Pattern Anal. and Machine Intell.*, **6**, pp. 518-523, Jul. 1984.