

解説 : Mean Shift Segmentation の OpenCV 実装

2016/03/27

ビジョン&IT ラボ 皆川卓也

1. 概要

OpenCV の `imgproc` モジュールには `cv::pyrMeanShiftFiltering()` という Mean Shift 法を用いた画像の領域分割アルゴリズムが実装されています。本書では OpenCV 3.1 の `cv::pyrMeanShiftFiltering()` 関数について、そのアルゴリズムを解説します。

Mean Shift を用いた領域分割の方法については、以下で提案されています。

Comaniciu, D., & Meer, P. (2002). Mean Shift: A Robust Approach Toward Feature Space Analysis. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 603–619.

この手法は、各画素の周辺の局所的な色分布のピークを探索し、共通のピークに達した画素を同じ領域内に存在するとみなすことで行います。

ピークの探索は、各画素を起点とし、その画像二次元空間および色三次元空間における周辺について、その局所範囲内の平均へ注目画素を順次移動することで行います。この周辺範囲の平均への移動を繰り返すと、局所的なピークへ収束します。

実装では画像ピラミッドを用いて、**Coarse-to-Fine** に領域分割を行います。すなわち、まず最も解像度の小さな画像で領域分割を行った後、徐々に解像度を上げていき、領域の境界付近のみ計算することで、計算の高速化を図っています。

Mean Shift Segmentation の例を図 1 に示しました。

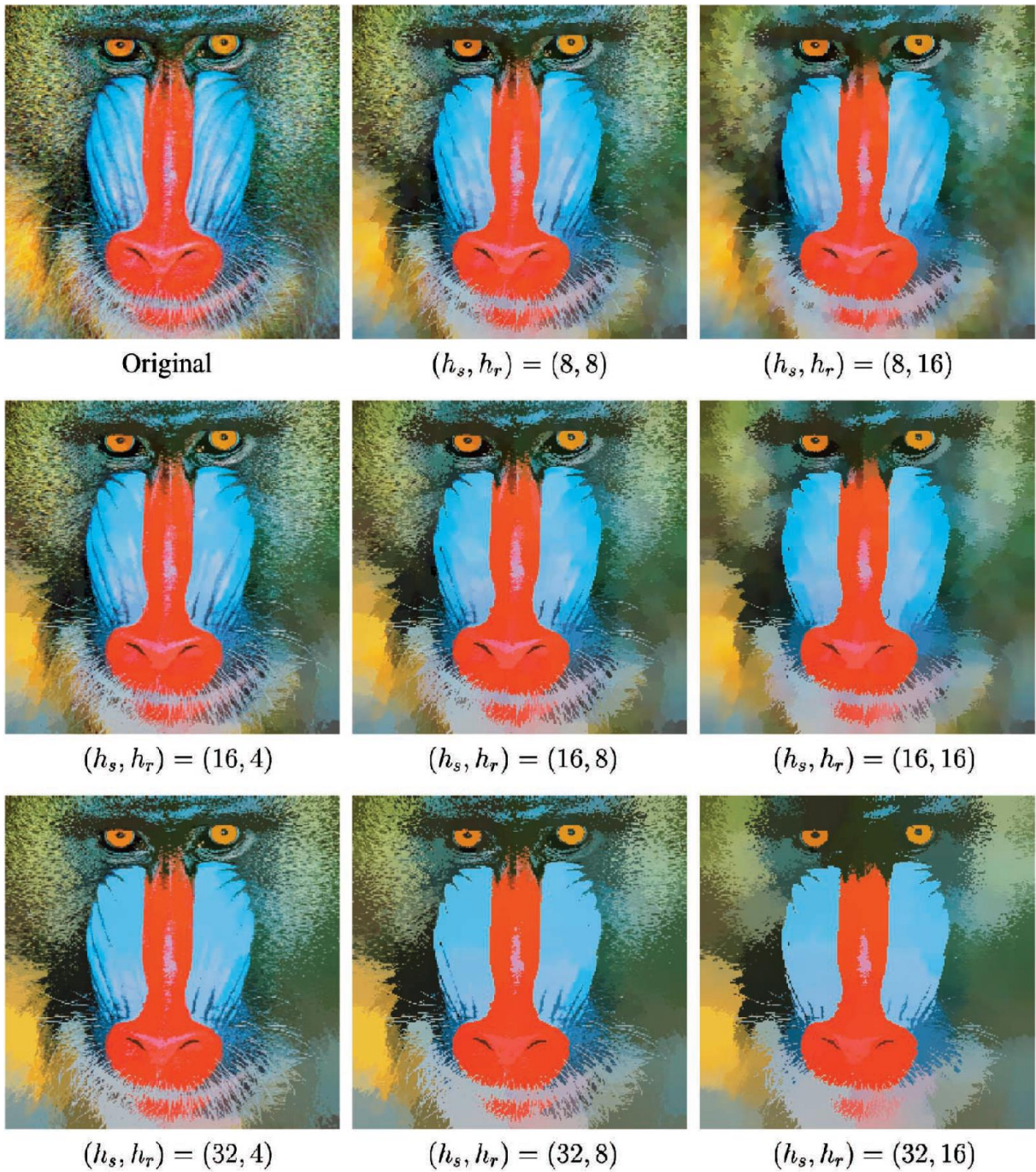


図1 Mean Shift Segmentation の例 (Comaniciu & Meer, 2002)

2. 処理の流れ

パラメータとして、以下を与えます。

- ・ 画像の二次元空間のカーネルサイズ h_s
- ・ 色空間のカーネルサイズ h_r

- ・ 画像ピラミッドのサイズ L
- ・ Mean Shift の繰り返し回数 N
- ・ Mean Shift の収束条件を表す閾値 ε

処理の流れは以下の通りです。

- 1 入力画像から L 枚の画像ピラミッドを作成します。画像ピラミッドのそれぞれの解像度は $1/2, 1/4, \dots, 1/2^{L-1}$ になります。
- 2 以下の処理を、解像度の最も低いものから順に行っていきます。2枚目以降は、領域の境界でのみ行います。

2.1 画像内の入力画素集合 $\{P_s\}$ 、出力画素集合 $\{P_d\}$ とします。

$$\{P_d\} \leftarrow \emptyset$$

2.2 入力画素集合 $\{P_s\}$ から画素 P_i を取得します。

$$\{P_s\} \leftarrow \{P_s\} \setminus P_i$$

2.3 画素 P_i を座標 (x_i, y_i) 、色 (r_i, g_i, b_i) の 5 次元ベクトルとした時、 (x', y', r', g', b') へ代入します。また $n = 1$ とします。

$$x' \leftarrow x_i, \quad y' \leftarrow y_i, \quad r' \leftarrow r_i, \quad g' \leftarrow g_i, \quad b' \leftarrow b_i$$

2.4 画素 P_j の座標 (x_j, y_j) 、色 (r_j, g_j, b_j) とした時、eq.(1) を満たす画素集合 $\{P_j\}$ を選択します。この式は図示すると図 2 のようになります。

$$|x_j - x'| \leq h_s, |y_j - y'| \leq h_s, \sqrt{(r_j - r')^2 + (g_j - g')^2 + (b_j - b')^2} \leq h_r \quad \text{eq. (1)}$$

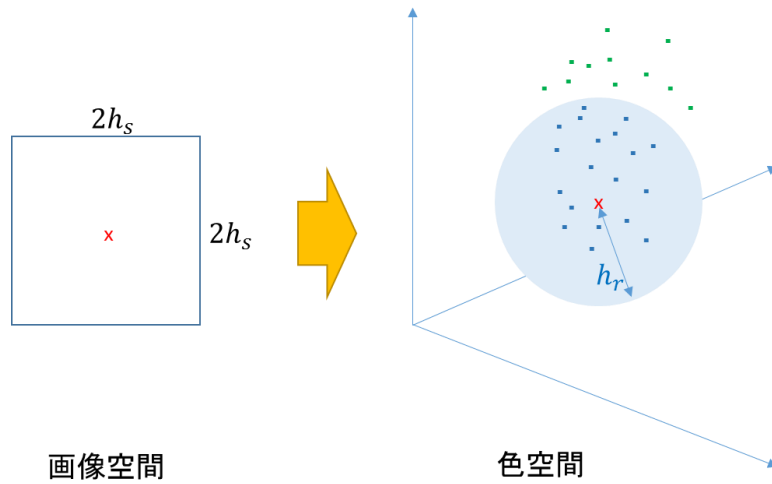


図2 画素の選択

2.5 画素集合 $\{P_j\}$ の位置の重心 (x^*, y^*) 、および色の重心 (r^*, g^*, b^*) を求め、以下のいずれかの終了条件を満たすか判定します。

$$\begin{cases} x' = x^* \cap y' = y^* \\ n \geq N \\ |x' - x^*| + |y' - y^*| + (r' - r^*)^2 + (g' - g^*)^2 + (b' - b^*)^2 \leq \varepsilon \end{cases}$$

2.5.1 終了条件を満たさない場合以下の代入を行って、2.4 から処理を繰り返します。これは図3のように処理中心の移動を意味します。

$$x' \leftarrow x^*, \quad y' \leftarrow y^*, \quad r' \leftarrow r^*, \quad g' \leftarrow g^*, \quad b' \leftarrow b^*, \quad n \leftarrow n + 1$$

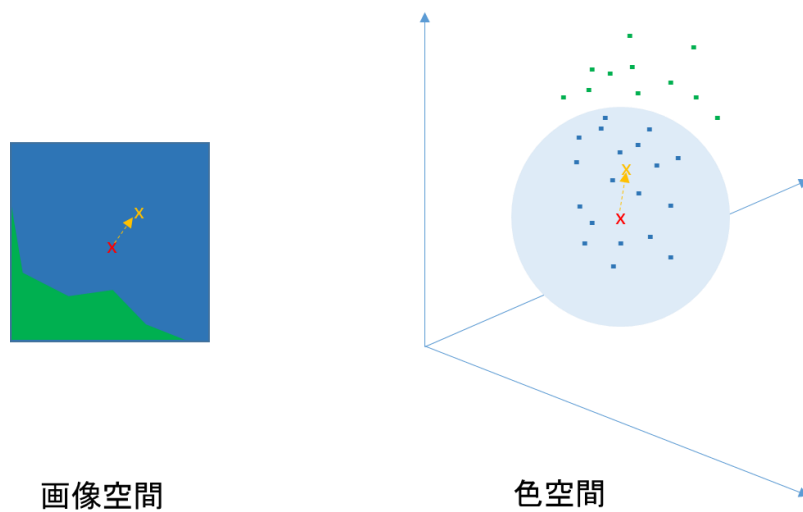


図3 Mean Shift

2.5.2 終了条件を満たす場合、出力画素を $P'_i = (x_i, y_i, r^*, g^*, b^*)$ として出力画素集合 $\{P_d\}$ に格納し、 $\{P_s\}$ が空集合でないなら、2.2 から処理を続けます。

$$\{P_d\} \leftarrow \{P_d\} \cup P'_i$$