

Hill-Climbing を用いたイメージセグメンテーション

大橋 巧[†] Zaher AGHBARI[‡] 牧之内 顕文[‡]

[†] [‡]九州大学システム情報科学研究所 〒812-8581 福岡市東区箱崎 6-10-1

E-mail: [†] ohashi@db.is.kyushu-u.ac.jp, [‡] {zaher, akifumi}@is.kyushu-u.ac.jp

あらまし 本論文では、画像を知覚的に重要な領域に分割する新しい手法を提案する。この手法は hill-climbing 法をベースに、大きく 2 つのタスクに分けられる。初めに、hill-climbing アルゴリズムによって画像の 3 次元カラーヒストグラムにおける局所的な最大値を検出する。次に、検出された最大値に関してピクセル同士を連結する。その結果、視覚的に同一な領域を生成する。この領域分割アルゴリズムは、単純かつ高速であり、またノンパラメトリックである。その為、領域分割の際にパラメータを手動で設定する必要がない。さらにこの手法では、クラスタの数または画像の種類によらない領域分割を行うことができる。

キーワード 画像の領域分割, Hill-Climbing

Hill-Climbing Based Image Segmentation

Takumi OHASHI[†] Zaher AGHBARI[‡] and Akifumi MAKINOUCHI[‡]

[†] [‡] Graduate School of Information Science and Electrical Engineering, Kyushu University 6-10-1 Hakozaki, Higashi-ku, Fukuoka, 812-8581 Japan

E-mail: [†] ohashi@db.is.kyushu-u.ac.jp, [‡] {zaher, akifumi}@is.kyushu-u.ac.jp

Abstract In this paper we present a novel image segmentation method that produces a set of perceptually meaningful regions. The method is based on a hill-climbing approach and achieves the segmentation by performing two main tasks. First, the hill-climbing algorithm detects local maxima of clusters in the global three-dimensional color histogram of an image. Then, the algorithm associates the pixels of an image with the detected local maxima; as a result, several visually coherent segments (small regions) are generated. The segmentation algorithm is simple, fast and nonparametric. The whole segmentation process is performed without any hand-tuning of parameters. Furthermore, this method does not assume any a priori knowledge on the number of clusters or the content of an image.

Keyword Image Segmentation, Hill-Climbing

1. はじめに

近年、インターネットで利用可能な画像や、様々なデジタルマルチメディアライブラリの急増に伴い、内容に基づく効果的な画像検索(CBIR)アルゴリズムが必要とされ始めている。多くの研究者が、画像の全体的な特徴、例えばカラーヒストグラムなどを用いて、画像の表現や検索をするためのアルゴリズムを提案している。しかし、画像の全体的な特徴を用いて検索を行っても満足な結果を得ることはできない[1]。その為、画像全体から意味のある領域を見つけ出し、その領域の部分的な特徴によって画像全体を表現することが必要不可欠である。そうすることで、莫大な画像からのCBIRをより効果的に行うことができる[2]。

画像の重要な領域を特定するプロセスは、よく知られている問題で“セグメンテーション”と呼ばれる。

これらの領域は、視覚的または知覚的な基本的性質を表す特徴(色、テクスチャ、形など)に関して定義される。画像のセグメンテーションアルゴリズムは大きく 2 つのカテゴリに分類することができる[3]:(1)領域に基づく手法、(2)輪郭に基づく手法。

領域に基づく手法[4][5][6]は、画像のピクセルを領域に区分しようとするものである。それぞれの領域のピクセルは、色やテクスチャといったある特徴に関して視覚的に均質で、“フラットゾーン”と呼ばれる。一般にこのような手法では、特徴空間でのあらかじめ決めておいた属性のみを使ってセグメンテーションを行う。そのため、画像のピクセルは複数の属性の中の 1 つの属性によって分類される。最後に、特徴が類似しており、空間的に閉じたピクセルの集合は、領域としてグループ化される。

一方、輪郭に基づく手法[7][8][9]は、まず領域のエッジを検出し、断続するエッジ群を連続した曲線に繋ぎ合わせることで領域を特定する。この方法の欠点は、エッジかどうかを局所的にかつ初期段階で決定する必要があるため、エッジの連結が不可能になってしまう場合があることである。この問題を解決するためには、エッジ検出の際に小さなしきい値を用いる必要がある。しかし、数多くの不規則なエッジが至るところで検出された場合、これを繋ぎ合わせることは非常に困難である[7]。

本論文では、HSV 色空間において視覚的に均質な領域分割を行う新しい手法を提案する。この手法は単純かつ高速、またノンパラメトリックであり、“Hill-Climbing”と呼ばれるアルゴリズムを利用している[10]。このアルゴリズムはAIの分野などで利用されており、ヒューリスティック探索アルゴリズムの一種である。セグメンテーションは2つのステップで行われる：(1)Hill-Climbing 法を用いて、画像の3次元カラーヒストグラムからクラスタの局所的最大値を求める、(2)検出された局所的最大値に関して画像のピクセルを連結する。その結果、いくつかの視覚的に均質な特徴をもつ領域を生成する。セグメンテーションアルゴリズムがクラスタを特定する際に一般的な色特徴を考慮していることから、出力される領域はユーザにとって知覚的に意味のあるものとなる。

2章では特徴空間の解析と Hill-Climbing を用いたセグメンテーションについて論じる。そして3章では、本手法を用いた結果について考察する。最後に、本論文の結論について述べる。

2. セグメンテーション

画像のセグメンテーションは、画像を均質な領域に分割することである。均質な領域とは、それに含まれるピクセルの値が類似していることとして定義される。セグメンテーションの目指す最も重要な目的は、(1)視覚的に均質な領域を抽出する、(2)一般的な画像データから満足できる結果を出す、つまり画像に関するなんらかの予備知識を必要としない、そして、(3)教師無しで行うことができることである。今回提案する手法は、色に基づいたセグメンテーションであり、前述の目的をすべて満たしている。セグメンテーションアルゴリズムを述べる前に、今回使用した HSV 表色系について説明する。

2.1. HSV 表色系

セグメンテーションにとって都合の良い色空間とは、視覚的な色の違いが色空間におけるそのユークリッド距離となっているものである。HSV 表色系はこの

条件を満たしている[11]。さらに、HSV 表色系は人間の知覚特性と非常に似ている。Fig. 1は、HSV 色空間が持つ3つの構成要素を表している：(1)Hは、その円盤における角度によって色相を表し、(2)Sは、円盤の中心からの距離で彩度を表し、(3)Vは、円盤に垂直な方向への距離で色の明るさ、明度を表している。これに従って、RGB 表色系から HSV 表色系へ色空間を変換している。

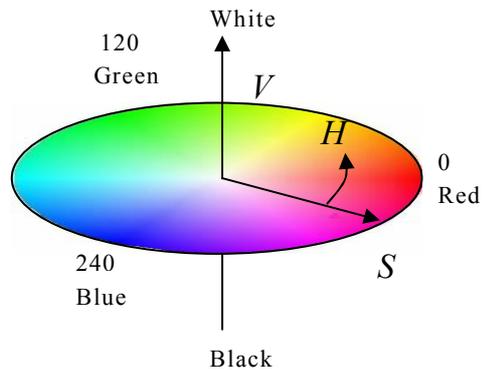


Fig. 1. HSV 表色系。Hはカラー円盤上の位置、Sはカラー円盤の中心からの距離、Vはカラー円盤と垂直な軸上での位置。この3つの値で色を表現する。

2.2.3 次元カラーヒストグラム

それぞれ画像について、HSV ヒストグラムを作成する為に3つの構成要素(H,S,V)の最大値と最小値を見つける。ヒストグラムのそれぞれの軸は、この最小値から最大値までの大きさを持つ。このとき、それぞれの構成要素の値は0から1に正規化を行っている。H, S, そしてVは、ヒストグラムのそれぞれの軸で16, 8, 8の分割数で量子化を行う。HがS, Vに比べて倍の量子化を行う理由は、画像データベースにおける色の多様性に対応するためである。

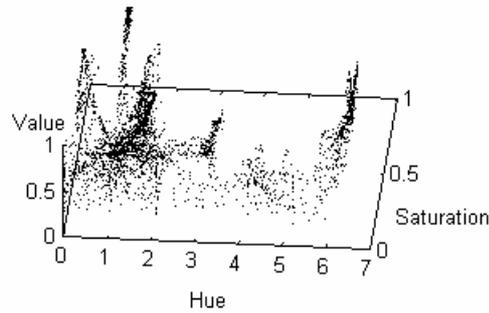
次に、先ほど量子化した画像の値をHSV ヒストグラムにマッピングしていく。Fig. 2に示すように、原画像(Fig. 2(a))は3次元のHSV空間にマッピングされ、いくつかの不規則なクラスタを形成していることがわかる(Fig. 2(b))。クラスタ間は連続して推移している為、パラメトリックな方法で特徴空間をクラスタリングすることは難しい[6]。そのため、本論文では、特徴空間の解析とクラスタの決定に、Hill-Climbing とよばれるノンパラメトリックな方法を利用している。

2.3. Hill-Climbing アルゴリズム

有効なセグメンテーションを行うためには、ノンパラメトリックであり、また局所的な特徴量と全体的な特徴量の分布を考慮した手法でなければならない。ノンパラメトリックなクラスタリング手法には2つのタイ



(a)



(b)

Fig. 2. 特徴空間の例 : (a) 原画像 (b) 3次元特徴空間でのピクセルの分布

プがある．距離に基づくクラスタリング，そして濃度推定に基づくクラスタリングである．距離に基づくクラスタリングは，全体的な特徴量の分布を考慮せず，特徴ベクトルの距離を測ることでピクセルを結合する[12]．全体的な特徴量の分布という重要な情報を切り捨てるため，セグメンテーション結果に誤差が生じることが予想される．一方，濃度推定によるクラスタリング手法は，確率密度関数として特長空間を考え，ここでは局所的最大値がセグメンテーション結果である領域に相当する．一度ピークが決定されると，クラスターは特徴空間の局所的な構造に基づいて決定される[1][4][6]．

本論文で提案する Hill-Climbing を用いたセグメンテーション手法は，画像の3次元カラーヒストグラムにおけるクラスターのピークを高速に，かつノンパラメトリックに探索するアルゴリズムである．クラスターのピークを求める際に，ピクセルそのものを扱うのではなくヒストグラムを扱うことで，効果的なピークの探索が可能である．ピークを求めた後，クラスターの画像上での位置を基にピクセルを連結する．Hill-Climbingを用いたセグメンテーションアルゴリズムの概略を以下に示す．まず初めに，1次元のカラーヒストグラムに対する Hill-Climbing について Fig. 3 に示す．その後，3次元である HSV カラーヒストグラムに対する Hill-Climbing について一般化する．

Algorithm : Hill-Climbing セグメンテーション
 Input : 画像データ
 Output : 知覚的に均質な領域

- 1 画像の1次元カラーヒストグラムを計算する．例えば，画像に含まれる色の色相(hue)とそれに対するピクセル数(Fig. 3(b))．
- 2 カラーヒストグラムのピクセル数がゼロでない

ものからスタートする．以下の手順で，ヒストグラムのピークに到達するまで移動を繰り返す．

- 2.1 ヒストグラムの現在位置と隣り合うヒストグラムとのピクセル数を比較(1次元の場合は左右と比較)する．ここで，色相は Fig. 1 に示したように円形である為，左端と右端はそれぞれ隣り合っていることに注意しなければならない．
- 2.2 ピクセル数を比較した結果，その数の大きい方へ現在位置を移動させる．
- 2.3 隣り合うピクセル数が同じであった場合，さらに次の隣り合うピクセル数との比較を行い，数の大きい方へ現在位置を移動させる．
- 2.4 2.1 から 2.3 を移動が終了するまで行う．この結果，隣り合う位置のピクセル数が現在位置のピクセル数よりも小さいとき，現在位置をピーク(局所的最大値)と定める．
- 3 ステップ2による移動が行われていない位置について，同様にピークを求める．この結果，すべての位置からそれに対応するピークが求まる．
- 4 特定されたピークは，入力された画像の初期クラスターを表している．画像のピクセルとそれに対応したピークを保存する．
- 5 最後に，画像上で隣接するピクセルのもつピークを比較し，同じピークを持つものを結合する．すべてのピクセルに対してこの作業を行うことで，入力画像の領域分割が完了する．

出力される領域の数を，多くても 20 個程度にする為，HSV 空間の量子化数をそれぞれの軸で 16, 8, 8 としている．量子化によるヒストグラム空間の分割数を多くするとオーバーセグメンテーションとなり，また逆に少なくするとアンダーセグメンテーションとな

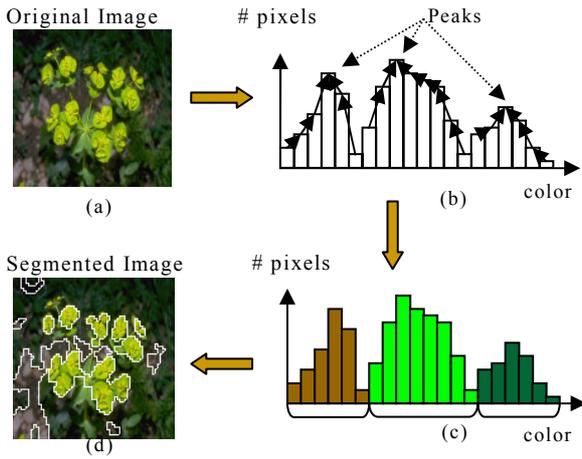


Fig. 3. セグメンテーションプロセスの例: (a) 原画像 (b) 1次元カラーヒストグラム(色相)での Hill-Climbing (c) 3つのクラスターが特定された Hill-Climbing 結果 (d) 領域分割された結果画像

る. この量子化数の変化に伴う結果の違いは, 3章で述べる.

2.4. Hill-Climbing アルゴリズムの一般化

3次元カラーヒストグラム(HSV カラーヒストグラム)への Hill-Climbing アルゴリズムの適用は, 1次元の場合と同様で簡単に行うことができる. ステップ1において, 扱う特徴空間が3次元となるために, ヒストグラムは $3^d - 1 = 26$ の隣接する空間を持つ. ここで d は特徴空間の次元数なので3となる. このため, 移動先を決める際の現在位置のピクセル数との比較が26回必要となる. また, H軸については, 左端と右端は連続している為比較を行うが, S軸とV軸に関しては連続していない為比較は行わない.

HSV空間を用いた場合, 最適なセグメンテーションを行うために, ここで以下のような条件を付加する必要がある: もし S の値がしきい値 τ よりも小さい場合, V 軸方向への移動を禁止する. この条件は, S の値が小さい場合は人間の知覚では色相を判別できない, というに基づいている. ここでしきい値は, 実験結果から0.1が適当であると判断した. Fig. 4に, この制約による効果を示す. しきい値によるV軸への移動禁止を行わない場合, 深い青色と黒との領域分割ができていないことがわかる(Fig. 4(b)).

2.5. Hill-Climbing アルゴリズムの計算量

上述の Hill-Climbing アルゴリズムが, 視覚的に同一な領域分割を効果的に行うことを示した. このアルゴリズムを用いたセグメンテーションは, ピークの特定とピクセルの結合という2つの別個なタスクから成り

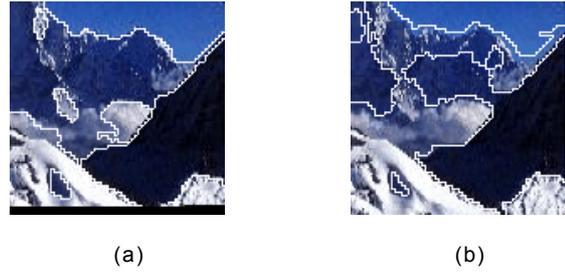


Fig. 4. $S < 0.1$ の制限による効果: (a) 制限を設けた場合の結果画像 (b) 制限を設けなかった場合の結果画像. 暗い青と黒, 白と灰色がうまく分割されていない.

立っている. 1つ目のタスクは, 現在位置と隣接する空間のピクセル数を比較する. 特徴空間が3次元の場合は $3^d - 1$ 個の隣接する空間を持つ. ここで, $d = 3$ である. また量子化された特徴空間は, $N_i: (H, S, V) = 16 \times 8 \times 8$ 個に分割されている. そのため, ピーク特定のタスクは, オーダーが $O((3^d - 1) \times N_i)$ であることがわかる.

一方, ピクセル結合のタスクは, 同じピークを持つピクセルを結合していく. 画像のピクセル数を N_p とすると, ピクセル結合のタスクは, オーダーが $O(N_p)$ であることがわかる. よって, Hill-Climbing アルゴリズムのトータルオーダーは以下で表される.

$$\text{complexity: } O((3^d - 1) \times N_i) + O(N_p)$$

3. 実験結果

実際の画像データを用いて, Hill-Climbing アルゴリズムを用いたセグメンテーションを行った. 特徴空間に HSV 表色系を使用した場合の実験結果をここに示す. このアルゴリズムはノンパラメトリックであり教師無しで実行可能である. よってセグメンテーションにおいてパラメータの手動設定をなんら必要としない. さらに, 生成する領域の個数は, Hill-Climbing アルゴリズムを実行した際の局所的最大値の個数と一致するため, その個数を指定する必要もない.

HSV 特徴空間は, $(H, S, V) = (16, 8, 8)$ に量子化されている. この量子化数が多いと, 例えば $(H, S, V) = (20, 10, 10)$ とした場合, Fig. 5(a) に示すようにオーバーセグメンテーションとなる. 逆に, 量子化数が小さいと, 例えば $(H, S, V) = (8, 4, 4)$ とした場合, Fig. 5(c) に示すようにアンダーセグメンテーションとなる. Fig. 5(b) は, 量子化数を $(H, S, V) = (16, 8, 8)$ とした場合の結果であり, 理想的な領域分割がなされている. 適当な量子化数を決定することで, 意味のない領域やノイズを軽減することができる. Hill-Climbing アルゴリズム

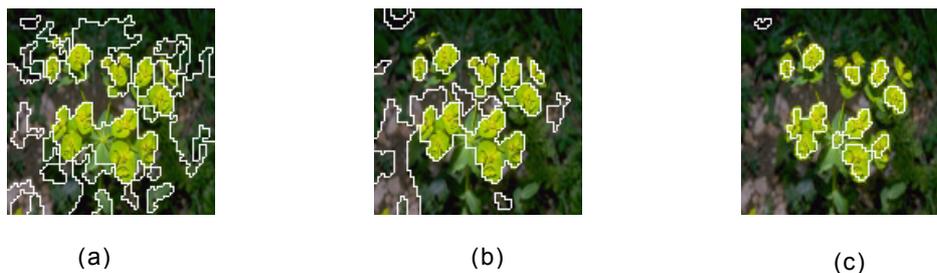


Fig. 5. 量子化数を変えた場合の領域の変化: (a) over-segmentation (b) correct segmentation (c) under-segmentation

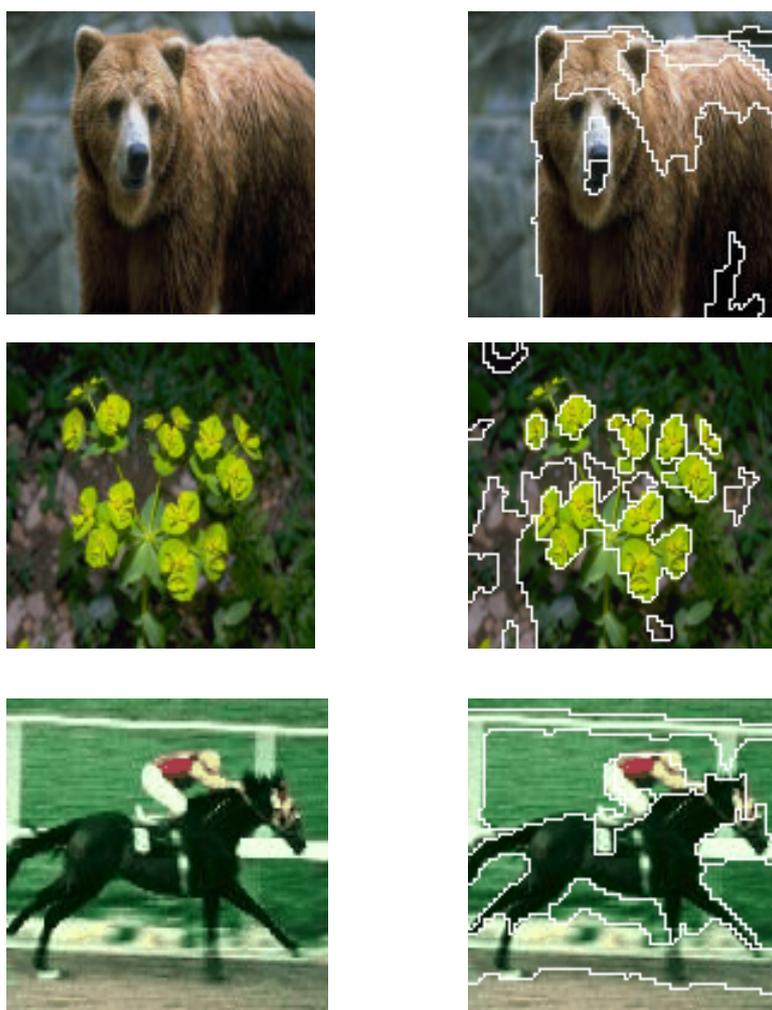


Fig. 6. HSV 表色系を用いた時のセグメンテーション結果 左側: 原画像 右側: 領域分割された画像

ムは、平均して 20 以下の領域に画像を分割する。もし意図的にオーバーセグメンテーション、またはアンダーセグメンテーションを行いたい場合は、ユーザはこの量子化数を任意に設定することで、所望の結果を得ることができる。

Fig. 6 は Hill-Climbing を用いたセグメンテーションアルゴリズムによる一般的な画像のセグメンテーシ

ョン結果である。左側が原画像データ、右側が領域の境界を白でマーキングしたセグメンテーション結果画像である。熊が描かれている上段の画像は、7 つの領域に分割されている。分割された領域の 1 つは、ほぼ背景と一致している。白い鼻の部分と黒い口元の部分は 2 つの領域として分割されている。また、熊の体は光の当たり方の違いにより、明るい茶色、茶色、暗い

茶色の3つの領域に分割されている。

Fig. 6 の中段の画像は、色のはっきりした違いによりいくつかの領域に分割されている。Fig. 5 に示したように、画像の HSV 空間から3つの局所的最大値を特定している。原画像データを念入りに見てみると、(1)花の部分の明るい緑、(2)葉の部分の暗い緑、(3)地面の部分の茶色、の3つの領域を見ることができる。Hill-Climbing アルゴリズムにより、うまく領域分割されていることがわかる。

Fig. 6 の下段の画像は、複雑な構図を持つ画像でのセグメンテーション結果である。この結果は、Hill-Climbing の正確さを証明している。画像データは10個の領域に分割されている。ここで、馬が1つの領域で表されていることが重要である。また、ジョッキーは服の色によって2つの領域に分割されている。Hill-Climbing アルゴリズムの性質に起因して、カラーヒストグラムの同じピークを持つ場合、下段の画像ではジョッキーの服の黄色の部分と赤い部分は画像の下部の砂の色を介して同じピークに移動したために、分割されず1つの領域となっている。

4. おわりに

本論文では、Hill-Climbing を用いたセグメンテーションによって、画像データを知覚的に意味のある領域に分割する新しい手法について述べた。このアルゴリズムは、ノンパラメトリックかつ教師無しで実行可能である。この手法は、実験結果で証明したように、2章で提言した以下の目標に達している：(1)視覚的に均質な領域を抽出する、(2)一般的な画像データから一般的なユーザが満足できる結果を出力する、(3)画像データに関する予備知識を必要としない、(4)セグメンテーションの際のパラメータに関して手動設定を必要としない。現在のところ、Hill-Climbing において、ピークを検出する際の移動距離には制限を設けていない。このため、ある条件下で、例えば赤とオレンジのようなそれほど似ていない色でも、画像の一部に赤みがかかったオレンジが写っているような場合、同じ色にグルーピングしてしまうということが起こり得る。しかし、この問題は移動距離に制限を加えることで簡単に解決できると思われる。今後は、この制限を設けると共に、画像からセマンティックなオブジェクトを生成する為に、階層的な領域のグルーピングアルゴリズムの研究を行う予定である。

文 献

[1] E.J.Pauwels, G.Frederix. "Finding Salient Regions in Images: Non-parametric Clustering for Image Segmentation and Grouping". *Journal of Computer Vision and Understanding*, 75(1,2), pp.73-85, 1999.

- [2] C.H.Kuo, A.Tawfik. "Unsupervised Color Image Segmentation for Content Based Application". *IEEE Int'l Conference on Multimedia and Expo (ICME)*, New York, USA, 2000.
- [3] T.Leung, J.Malik. "Contour Continuity in Region Based Image Segmentation". *5th Euro. Conference on Computer Vision*, Germany, June 1988.
- [4] S.Belongie, C.Carson, H.Greenspan, J.Malik. "Color and Texture Based Image Segmentation Using EM and Its Application to Content-Based Image Retrieval". *Int'l Conf. On Computer Vision (ICCV)*, pp.675-682, 1998.
- [5] L.Garrido, P.Salembier, D.Garcia. "Extensive Operators in Partition Lattices for Image Sequence Analysis". *EURASIP Signal Processing*, 66(2), pp.157-180, Apr. 1998.
- [6] D.Comaniciu, P.Meer. "Mean Shift: A Robust Approach Toward Feature Space Analysis". *IEEE Trans. on Pattern Analysis and Machine Intelligence*. 24(5), pp.1-18, May 2002.
- [7] J.Malik, S.Belongie, T.K.Leung, J.Shi. "Contour and Texture Analysis for Image Segmentation". *Int'l Journal of Computer Vision*, 43(1), pp.7-27, 2001.
- [8] B.R.Lee, A.B.Hamza, H.Krim. "An Active Contour Model for Image Segmentation: A Variational Perspective". *IEEE Int'l Conference on Acoustic Speech and Signal Processing*, May 2002.
- [9] D.Schluter, F.Kummert, G.Sagerer, S.Posch. "Integration of Regions and Contours for Object Recognition". *IEEE 19th Int'l Conference on Pattern Recognition*, Sep. 2000.
- [10] M.Ginsberg, "Essentials of Artificial Intelligence", A Morgan Kaufmann, 1993.
- [11] P.Green. "Understanding Digital Color". Buleprint, an imprint of Chapman & Hall, USA, 1995.
- [12] P.Salembier, F.Marques. "Region-based Representation of Image and Video: Segmentation Tools for Multimedia Services". *IEEE Trans. on Circuits and Systems for Video Technology*. 9(8), pp.1147-1167, Dec. 1999.