

## 複数特徴に基づく自動領域分割

矢野 雅之 勝間 大輔 清水 大輔 渡邊 睦

鹿児島大学理工学研究科 〒890-0065 鹿児島県鹿児島市郡元 1-21-40

E-mail: sc100062@ics.kagoshima-u.ac.jp, mutty@ics.kagoshima-u.ac.jp

**あらまし** 画像認識を行う際、まず領域分割を行って認識対象の候補となる領域を設定することが行われる。しかし、領域数は一般に未知であり、又、明度/色相値の類似性のみで領域を生成した場合には影や模様の影響により適切に設定できないという問題がある。そこで本稿では、明度による自動領域分割に対してDFT処理した領域をクラスタリングした分割結果を統合することにより、領域分割結果の精度向上を行う方式について提案する。屋内シーン、屋外シーンにおいて本手法の有効性を確認した。

**キーワード** 領域分割, 離散フーリエ変換, k-means クラスタリング, 明度ヒストグラム

### Research of automatic image segmentation based on multiple image features

Masayuki YANO Daisuke KATSUMA Daisuke SHIMIZU Mutsumi WATANABE

Graduate School of Science and Engineering, Kagoshima University

1-21-40 Korimoto, Kagoshima, 890-0065 Japan

E-mail: sc100062@ics.kagoshima-u.ac.jp, mutty@ics.kagoshima-u.ac.jp

**Abstract** The first stage of object recognition is generally the region(description) creation to identify candidates of recognized objects by means of segmentation. However, several difficulties exist in the segmentation, such as, unknown region number and miss-segmentation due to shadows or irregular texture patterns. We propose a unified method of both intensity-based segmentation result and texture-based one to improve the accuracy. The texture-based segmentation is executed by clustering DFT outputs. Experimental result have shown the effectiveness of the proposed method

**Keyword** image segmentation Discrete Fourier Transform, k-means clustering, brightness histogram

#### 1. はじめに

コンピュータビジョンにおける画像認識は一般に認識対象の分割、モデルによる認識、シーンの記述の順に行われる。領域分割はこの認識対象の分割における基本的な処理であり、認識性能を大きく左右する[1]。

領域を分割するための画像特徴としては、明度、色などのスペクトル特徴、テクスチャなどの空間的特徴、形状特徴などの幾何学的な特徴がある。これらを単一的に用いるだけでは、分割できない領域に対して、結果を改善するための画像特徴情報を得られないとい

う問題がある。そこで明度ヒストグラムと2次元離散フーリエ変換を用いた空間周波数の複数特徴を用いる分割手法を考案し、実験により有効性を確認した。

#### 2. 従来研究

画像中に撮影された各対象物に対応する各領域を求める問題は、領域分割(region segmentation)または領域生成(region growing)と呼ばれ、1970年代初期から活発に研究が行われてきた。

これまで提案されてきた手法は、クラスタリング処

理を行う対象領域の観点からは、局所的手法、大局的手法、分割と併合の組み合わせ手法に分類される。ここで、局所的手法とは各画素に対して近傍の画外の画像特徴の類似性評価に基づき領域を併合生成していく手法であり、大局的手法は、画像全体に分布する画素群の大局的な性質に基づいて領域を分割生成する方式である。また、分割と併合の組み合わせ手法においては、局所領域への分割と複数の局所領域の併合を繰り返すことにより、目的とする領域生成を行う方式である。

クラスタリング処理を行う特徴空間選択の観点からは、画像平面でのクラスタ化、低次元特徴空間でのクラスタ化、高次元特徴空間でのクラスタ化に分類される。

クラスタリング処理を行う対象領域の観点からの分類において、局所的手法および分割と併合の組み合わせ手法は、処理開始時点または処理途中での多数の処理領域を対象とするため処理時間が多くかかる、という問題点を有する。局所的手法においてはさらに、画像特徴の選択及び類似性の判定基準の設定法が、分割と併合の組み合わせ手法においては、分割と併合のどちらを実施するか判断基準の設定及び収束の判定が、技術課題として存在する。

これに対し大局的手法では、画像全体を扱うため、多数の局所的領域を扱うことによる処理時間の増大は問題とならず、高速な初期記述領域を行うことができる。

低次元特徴空間における大局的領域分割手法として最も一般的なものは、「閾値処理」である。実際のシーンにおいては、シーンを構成する対象の明度範囲が未知であるために、閾値を固定的に設定することはできない。そこで1次元特徴量である明度ヒストグラムを利用する、p-タイル法、モード法[2]、判別閾値選定法[3]が提案されている。

これらのヒストグラム処理による分割結果では、閾値付近のデータが過分割されることがあり、エッジ付近などではその影響が大きくなる。

統計判別手法による分割として、ヒストグラムを混合正規分布とみなし、隠れ変数を有するモデルパラメータの最尤推定手法であるEMアルゴリズムを用いて領域分割を行う手法が提案されている[4][5]。この手法は事前知識を必要とせずに分割数を推定することができる。しかし、屋内シーンのような環境変動が小さい場合でのヒストグラムの分布が混合正規分布では無いときに、分布推定が正しく行われないう問題がある。また、複数の分布が近い位置に推定されることがあり、その分布が重なる明度範囲の領域に対して分割の選択が不安定なため、領域が過分割される傾向にある。

一方、テクスチャ特徴を用いた手法としてウェーブレットを用いて分割を行う手法が提案されている[6]。この手法は、木や草のようなテクスチャ変化の異なる領域の分割に有効な結果を示している。しかし明度変化が緩やかな領域に対しては周波数の変化の特徴を得ることが難しく、分割が困難になる。

以上で述べたエッジ付近での誤分割結果を改善するために、本稿では、明度による自動領域分割に対してDFT処理した領域をクラスタリングした分割結果を統合することにより、領域分割結果の精度向上を行う方式について提案する。

### 3. 領域分割処理の流れ

#### 3.1. 領域併合指標

領域を併合するためのデータとして明度ヒストグラムによる分割結果と、空間周波数成分による分割結果の2つが存在するが、本手法での領域分割のベースは明度ヒストグラムの領域分割結果とする。

#### 3.2. 処理の流れ

本手法の処理の流れを図1に示す。一枚の画像に対して、明度ヒストグラムと画像中の空間周波数成分の2種類の特徴を独立に処理することにより各々2つの領域分割画像を作成する(a)。2つの画像についてのラベリングを施し(b)、隣接ラベルにおいて併合条件を用いて領域の併合を行う(c)。図1(a)の明度ヒストグラムによる領域分割の手法についての説明を3節で、空間周波数成分による分割手法を4節で述べ、(b)(c)についての説明を5節で行う。

#### 3.3. 明度による領域分割

本研究では明度ヒストグラムによる大局的手法を用いた。明度ヒストグラムを分割する手法としてモード法がある。モード法はヒストグラムに双峰性があるとき、分布の極小値を探索することで分割を行う手法である。本手法ではこれを応用し、ヒストグラム分布中で極大と思われる箇所を全て探し出し、その間の極小値を取得することで、分布の山の数に対応した領域分割数を得るようにした。これにより、ヒストグラムの分布の違いに応じて、クラスタを自動的に決定することが可能となる。ヒストグラムに対し、尖度評価を行い、混合正規分布に近いと判断されたときにはEMアルゴリズムを用いた自動領域分割を行い、そうでない分布の時は分布の極小値を利用し分割を行うようにした。

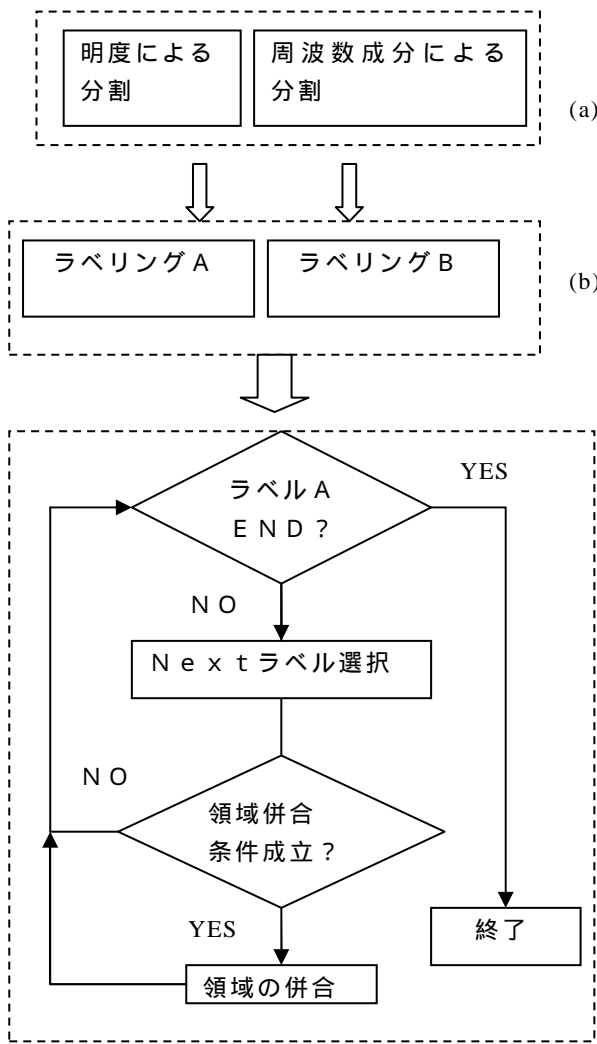


図 1 本研究の処理の流れ

$$w_1 = e^{-j\frac{2\pi}{M}}, \quad w_2 = e^{-j\frac{2\pi}{N}} \quad (2)$$

### 3.4.2. クラスタリング

上記の処理で出力した周波数成分を画素のもつ特徴次元として画像領域を分割するために、k-means クラスタリング[7]を使用した。k-means は非階層のクラスタリング手法で n 個のデータをクラスタの平均を用い、指定した k 個のクラスタに分類する手法である。K-means アルゴリズムの基本的な流れは以下のとおりである。

1. K 個の初期クラスタの中心をランダムで与える。
2. 全てのデータと K 個のクラスと中心との距離を求め、個々のデータを最短距離のクラスタに分類する。
3. 新たにできたクラスタの中心をクラスタの中心とする。
4. クラスタの中心が全て前の結果と同じになる。または、指定繰り返し回数に達するまで、2、3を繰り返す

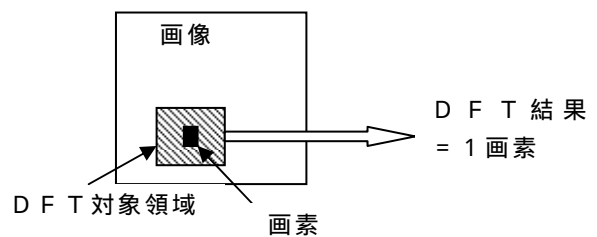


図 2 画素と周波数領域

## 3.4. 空間周波数による領域分割

### 3.4.1. 離散フーリエ変換

明度情報とは異なる画像情報として、空間周波数を利用して領域分割を行った。空間周波数とは、単位長さあたりの正弦波の濃淡変化の繰り返しを示すものである。空間周波数成分の取得法として、2次元離散フーリエ変換 (DFT) を使用した。2次元画像信号を  $f[m,n]$ 、画素数を  $M \times N$ 、フーリエ係数を  $F[k,l]$  で表現すると2次元DFTの式は式(1)のようになる。2次元DFTを行うことで、ある画像領域での周波数成分の分布を知ることができる。

$$F[k,l] = \frac{1}{MN \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} f[m,n] w_1^{km} w_2^{ln}} \quad (1)$$

本研究では、明度が1画素単位で領域分割結果を得ていることから、周波数成分でも1画素単位で領域分割結果を得た方が良く考えたため、図2のように画像中の各画素を中心とした  $32 \times 32$  画素に対してDFT処理を施した結果を対象画素が持つデータとするようにした。そのため  $1024$  次元の情報に対して k-means 処理を行うことになる。

k-means クラスタリングを行うときにはクラスタ数 K を設定する必要がある。そこで各々の K に対してクラスタリングされた領域画像に対しラベリングを施し、ラベリング総数が最も小さくなるときの K を領域分割に最適なクラスタ数であるとし、分割を行う。

### 3.5. 領域併合条件

領域を併合するために2つの領域分割データに対してラベル付けを行い、ラベル画像を作成する。ここ

で、明度ヒストグラムによる分割結果に対してラベル付けを行ったものをメインラベル、空間周波数による分割結果に対してラベル付けを行ったものをサブラベルとして説明する。メインラベルは以下の情報を持つようにする。

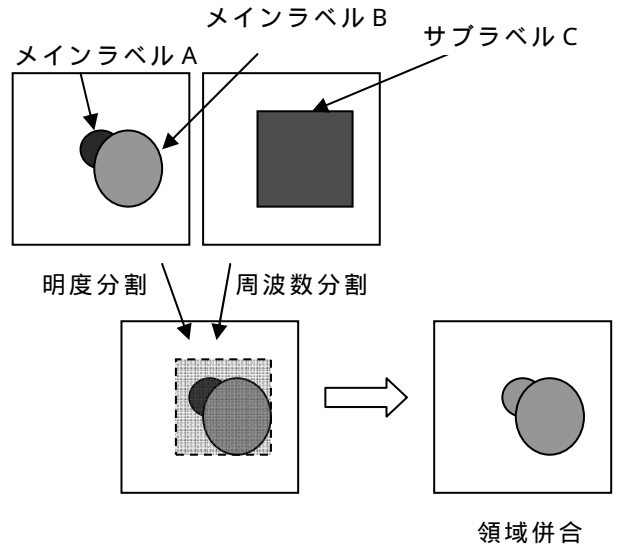
- 隣接メインラベル
- 画素数
- ラベル値

このラベルを用い、以下の領域併合条件に照らし合わせて領域を併合する。図3を元に説明を行う

1. メインラベルAの画像領域と同じ領域におけるサブラベル画像の分布を調べる。この時、一つのサブラベルCの占める領域が閾値  $Th1$  以上であるならばこのメインラベルは次のステップへ進む。閾値以下ならば、次のメインラベルへ移動し、1の処理を行う。
2. 隣接メインラベル情報を調べ、Aの画素数より大きく、その中で画素数が一番高いメインラベルBを併合対象に指定する。1と同様にメインラベルB領域でのサブラベル画像の分布を調べる。Aと同じサブラベルCの占める領域が閾値  $Th2$ 、または規定数以上の場合、領域の併合を行う。
3. 全てのメインラベルについても同様に、1、2の処理を行う。

1の条件が満たされることは、メインラベルAは1つの空間周波数分割領域に位置していることになる。同様に2の条件が満たされるときメインラベルBは、Aと同じ空間周波数分割領域を保持していることになる。つまりメインラベルA、B両者は同じ空間周波数特徴を持っていることになる。この時、両者は同じ認識物体の領域の一部だと判断され、領域が併合される。

併合処理が行われるとラベルA領域のラベル値をラベルBのラベル値に変更する。この手法では、併合処理済という条件を考慮にいれていないため、併合した後のラベルに対して併合処理を行った場合、処理結果がおかしくなる。しかし、この結果に対して再ラベリングを施しもう一度併合処理を行うことで結果は改善される。



A,B が同じ周波数分割領域 C 中に存在するならば、A,B の領域を併合する

図3 領域併合イメージ

## 4. 実験

### 4.1. 実験環境

前章で述べた方式に基づく自動領域分割システムを、汎用パソコン上のソフトウェアとして試作し、提案手法の妥当性を確認する実験を行った。CPU:3.2GHzのデスクトップPC, OS:Windows XP Professional 開発言語に Microsoft Visual C++6.0を使用した。640 \* 480ピクセルのモノクロ画像を対象とした。

### 4.2. 実験条件

今回の実験では明度ヒストグラムの分割は極小値による分割手法を用いて行うようにした。3章4節において k-means のクラスタリングに用いる特徴次元を1024次元としたが、次元が増えてしまうと処理時間が大きくなり、また、初期点の設定とクラスタ数Kによって大きく異なる分割結果が得られてしまうことが確認された。そこで、得られたDFTのうちの8分の1のデータ、つまり128次元に対してクラスタリングを行った。また、ラベル数削減の為に、領域分割画像に対して近傍処理を施した。これは、対象となる画素及び8近傍の値を調べ、最頻値に置き換える処理である。これによりドットのような領域を、周辺の領域に置き換え、ラベル数を減らす。併合条件で使用する閾値  $Th1$  を80%、 $Th2$  を40%に設定した。

また、本研究との比較の為にEMアルゴリズムを用

いた自動領域分割手法[5]との比較を行った。

### 4.3. 実験結果

最初に屋内シーンでの実験結果について述べる。これは天井カメラを用いてとられた机を中心とした画像である。机上中心付近に3つの箱、右部にノートと手帳、左部に白紙が存在している。

図4に屋内シーンの結果を示す。(b)は明度ヒストグラムを3.3で紹介した手法で分割した結果である。この画像でヒストグラムは7つの領域に分けられた。エッジ付近では明度変化が大きい事があり、細かい領域を多く抽出してしまうことが確認できる。(c)は3.4で紹介した手法で分割した結果である。この時のk-meansのクラスタ数Kは4であった。1画素が周辺領域の特徴も持っているため、箱や白紙のエッジ付近での大きな周波数特徴周りの画素に含まれてしまう。それ故、エッジ周辺で領域がまとまるという結果になった。(d)は領域併合の結果である。壁、ティッシュ箱、ノート、白紙が代表的な領域として抽出された。椅子は(b)の明度ヒストグラムの分割結果時点で机と重なる領域となっていた。併合処理前のラベル数は243であったが、併合処理後には54に減少した。画面中央のティッシュ箱は(b)では多数の領域が存在していたが、(d)では2つの領域に減少した。

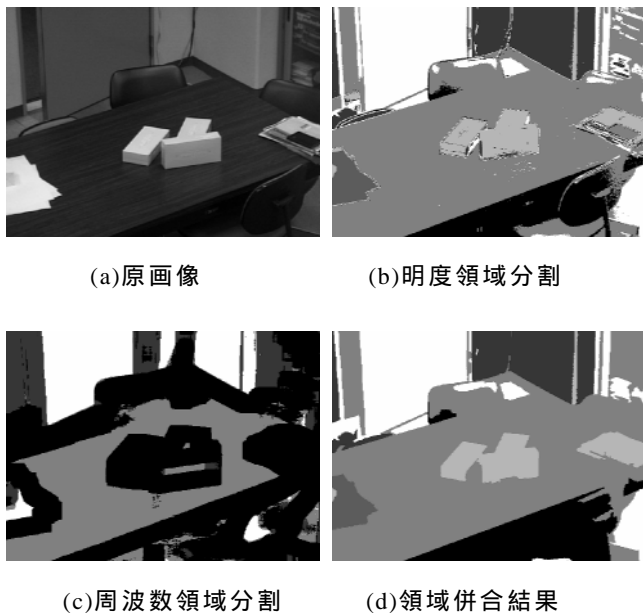


図4 屋内シーン処理結果

	E M 法利用	明度のみ	併合後
机	43%	85%	87%
箱	33%	58%	59%

表1 屋内シーン精度評価

同様に屋外シーンとして山の風景に対して実験を行った。画像中には雲、空、山(森)が存在している。図5が屋外シーンでの処理結果となる。(a)の画像に対して、明度ヒストグラムを用いた自動領域分割での分割数は3、空間周波数を用いた領域の分割結果は3であった。この場合の(c)の結果を見ても山の輪郭付近や雲の輪郭で領域がまとまっていた。明度ヒストグラムみの領域分割結果のラベル数は81であったが、領域併合結果後のラベル数は8に減少された。空の領域は一つに、山の領域が大きく2つの領域に分割された。

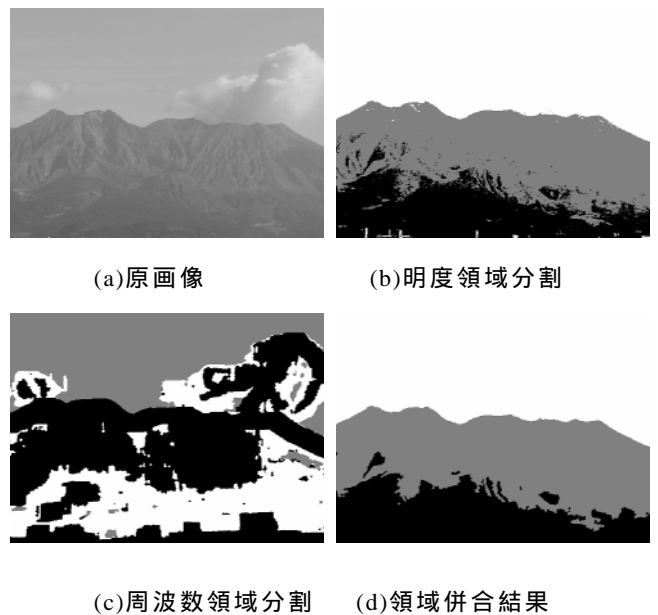


図5 屋外シーン処理結果

	E M 法利用	明度のみ	併合後
山	85%	59%	60%
空	50%	93%	93%

表2 屋外シーン精度評価

## 5. 考察

表 1 は物体認識に有効な領域分割の実現という観点から、対象がどの程度正確に抽出されているかを定量的に示すべく、人の手で領域を抽出した結果と、自動抽出した領域分割結果の面積値との比を正答率で示した。単位はパーセンテージで示す。評価対象として、画像中央のティッシュ箱と机を用いた。

EM アルゴリズムを利用した自動領域分割では、静止画且つ明度ヒストグラムが混合正規分布ではなかったことが影響したのか、最適だと思われる分割結果が得られなかった。その為領域が過分割されてしまい、低い結果になっている。本手法は、正答率に関してテ机は 85% から 87% に、ティッシュ箱は 58% から 59% に上昇し、精度の向上が確認された。

屋外シーンでは分割結果として雲の分割まで期待していたが、明度ヒストグラムの分割手法により、空と雲の明度が近いことから同じ領域に分割された。しかし、空間周波数による分割結果では、雲の特徴を捉えていた。屋外シーンの山の領域において評価結果も同様に精度向上が確認された。空の領域はすでに一つの領域になっていたため併合することはなく精度が向上することはなかった。しかし、山の評価比較において EM アルゴリズムによる自動領域分割の評価が上になる結果となった。これは、目視で設定した山の定義が影響しているものと思われる。今回の実験で使用した画像において筆者は山を、山から下の画像全部（森も含める）を山と定義したため、このような結果となった。森を別物として定義した場合には山の精度評価は変わらない結果となる。

このように明度においてエッジ付近において細かく分割されていた領域に対して、それらを併合するという処理結果が確認できた。

屋内シーンにおけるラベル数は 243 から 53 に、屋外シーンにおけるラベル数は 81 から 8 に減少され、目視による物体の領域に対して、少ないラベルで記述される結果になった。領域記述の点においても性能は向上した。

## 6. 課題

実験から、本手法に対して改善されるべきことが幾つか確認された。一つは明度ヒストグラムによる分割結果を主にし、併合処理のみを行うため、領域分割結果が大きく明度ヒストグラム分割手法に依存されることである。それに対処する分割手法の改良、もしくは再分割処理を啜る必要がある。

また、領域併合に対して以下の条件が不足していた。併合の対象を隣接する一番大きな領域としているため

画像の大半を占める領域に隣接する領域では併合条件に合わないことが多いはずである。その時、処理を終了しているが、それでは他の隣接領域に対して併合条件が合う可能性があっても領域併合を無視していることになる。大きい領域に併合するという考えにも妥当性があまりないため、これらの条件について検討する必要がある。

## 7. 結論

明度ヒストグラムを用いた自動領域分割画像に、空間周波数成分による領域分割結果の条件を組み合わせることにより画像領域を併合する手法を開発し、実画像を用いた実験によりエッジ付近領域の併合により領域分割結果の精度向上と領域数の減少を確認した。今後は明度ヒストグラム以外の低次元特徴を利用して同様の併合処理を行っていく予定である。

## 8. 謝辞

本研究の一部は、日本学術振興会の科学研究補助金（「分散視覚エージェントの統合に基づく親和的情報空間の展開」, 課題番号:15500115）の援助を受ける。

## 文 献

- [1] 白井良明, “ コンピュータビジョンの誕生と成長,” 電子情報通信学会誌, Vol.89, pp.173 - 175, 2006-2.
- [2] A.Rosenfeld and A.C.kak, “ デジタル画像処理”, 近代科学社 (1978)
- [3] 大津展之, “ 判別および最小 2 乗基準に基づく自動閾値選定法”, 信学論, vol.J63-D, no.4, pp.349-356
- [4] 渡邊睦, 東友紀 “ 明度分布パラメータに推定に基づく自動領域分割 ” 情報メディア学会誌, Vol.56, No.2, pp.280-289
- [5] Mutsumi Watanabe, “ Adaptive Decomposition of Dynamic Scene into Object-based Distribution Components based on Mixture Model Framework”, IEICE Trans. Vol.E88-D, No.4, pp.758 - 766
- [6] 狩野芳正, 大野信一郎, 阿曾弘具, “ 特徴選択によるテクスチャ画像の教師なし領域分割 ”, 信学論, D-11, vol.J86-D-11, no.7, pp.988-995
- [7] 高木幹夫, 下田陽久 “ 新編画像解析ハンドブック ”, 東京大学出版会, pp.1576-1579