

1. はじめに

屋内外の風景写真などの画像からシーンを理解するには、色だけでなく、意味を捉えて領域を分割する必要がある。これをセマンティックセグメンテーションと呼ぶ。従来手法では、コンテキストを捉える工夫により、高精度なセマンティックセグメンテーションを実現している。しかし、分割した領域に対応するラベルがひとつであり、複数のラベルが複合したラベル (マルチラベル) に対応できない問題がある。そこで、シングルラベルを扱う従来法 Semantic Texton Forests[1] をベースとしてマルチラベルに拡張を行うことで、マルチラベルに対応したセマンティックセグメンテーション手法を提案する

2. Semantic Texton Forests[1]

Semantic Texton Forests(以下 STF) は画像パッチを Randomized Trees(以下 RT) により学習することで構築する決定木の集合である。図 1 に構築した STF の例を示す。STF の分岐ノードは、以下の 4 つのオペレータ

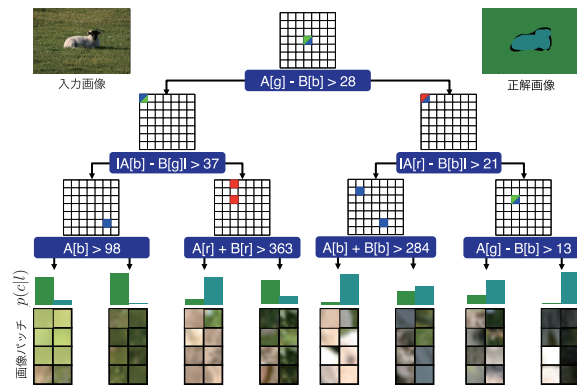


図 1 : Semantic Texton Forests

(i) $p_{x_1,y,b}$, (ii) $p_{x_1,y_1,b_1} + p_{x_2,y_2,b_2}$, (iii) $p_{x_1,y_1,b_1} - p_{x_2,y_2,b_2}$, (iv) $|p_{x_1,y_1,b_1} - p_{x_2,y_2,b_2}|$, から選択した特徴量ときき値により作成する。また、末端ノードには学習サンプルにより算出した確率 $p(c|l)$ を作成する。

STF はこの木構造を利用して特徴抽出を行うための手法であり、ある領域から得られる複数の画像パッチを STF に入力し、各ノードを通過した頻度を算出することで、ある領域に対してヒストグラム特徴 Semantic Texton Histogram(STH) を算出することが可能である。この特徴量はセマンティックセグメンテーションに有効であることが知られており、本研究ではこの特徴量を利用してマルチラベルに対応したセマンティックセグメンテーションを行う。しかし、STF はマルチラベルの学習に対応していないという問題があり、学習法の改良が必要である。

3. 提案手法

提案手法は、従来法である STF をベースに、マルチラベルに対応したセマンティックセグメンテーションを実現するために 2 点の拡張を行う。一つ目は、STF に Label Ranking を用いて LR-STF(Label Ranking-STF) を構築する。これにより、マルチラベルのサンプルを学習可能となる。二つ目は、Label Powerset[2] を用いて LP-RTs(Label Powerset-RTs) を構築する。これにより、対象物体にマルチラベルを考慮することが可能となる。

3.1 LR-STF(Label Ranking-STF)

Label Ranking とは、マルチラベルの学習データを用いてシングルラベルの順位付けを行う識別器を作成するアプローチである。提案手法では、STF に Label Ranking を取り入れることで、マルチラベルを考慮した特徴抽出が可能である LR-STF を構築する。

Label Ranking を行うためには、マルチラベルをシングルラベルに分解する処理を何らかの基準により行う必要がある。そこで、RTs の学習法において、最適な分岐ノードの決定を情報利得を基準に行う点に着目し、マルチラベ

ルのシングルラベルへの分解も同様に情報利得を基準に行う。RTs の学習法では分岐ノードを決定するために、特徴量ときき値の候補をランダムで用意するように、マルチラベルサンプルに付けられているラベルの中から分解する候補のラベルをランダムで用意する。ある分岐ノード n を構築するために利用するサンプル $i \in I_n$ にマルチラベル $C_i = \{c_1, c_2, \dots, c_n\}$ が付けられているとし、ラベルの候補を C_i とすることで、式 (1) により候補を選択する。

$$C_i = \begin{cases} \text{rand}(C_i) & |C_i| > 2 \\ c_1 & \text{otherwise} \end{cases} \quad (C_i \neq \phi) \quad (1)$$

ここで、関数 rand は集合 C_i 中の要素からひとつをランダムで選択する。また c_1 は、シングルラベルを意味し、要素が一つの集合 C_i の要素を意味する。この候補選択を L 回を行い、 L 個の候補の中から、最も情報利得が高くなるラベルの候補をその分岐におけるサンプル i のラベルとする。

この処理を利用して分岐を繰り返すことにより、サンプル i は分岐した回数 S だけラベルが決定され、次のような履歴が作成される $\{C_i^1, C_i^2, \dots, C_i^S\}$ 。これを用いてサンプル i の最終ラベル C_i^{max} を算出する。まず、履歴を用いて選択したクラスの頻度により確率 $p(c|i)$ を求める。そして、以下の式により最終ラベル C_i^{max} を求める。

$$C_i^{max} = \arg \max_c p(c|i) \quad (2)$$

これらの処理により、サンプル i につけられたマルチラベル C_i は C_i^{max} に収束し、シングルラベルへの分解を行うことができた。この最終ラベル C_i^{max} を用いて末端ノード l の確率 $p(c|l)$ を算出することで、Label Ranking が可能となる。

3.2 LP-RTs(Label Powerset-RTs)

LR-STF により、マルチラベルを考慮したヒストグラム特徴 STH の算出が可能となった。そこで、STH を用いてマルチラベルに対応した識別器を構築する。マルチラベルに対応した識別器を構築するために、RTs に Label Powerset 法を利用した LP-RTs 識別器を構築する。

Read 等の提案した Label Powerset 法 [2] は、従来のマルチクラス識別器を構築する手法と組み合わせることで、マルチラベルに対応することが可能となる手法である。この手法は、マルチラベルを符号化することにより新しいラベルを作成し、マルチラベルをシングルラベルとして学習する。表 1 に処理例を示す。このように変換した教師信号を用いることでマルチラベルを考慮した学習が可能となる。

表 1 : Label Powerset の利用例

サンプル	ラベル	処理後のラベル
1	$\{c_1, c_4\}$	$\{c_{1,4}\}$
2	$\{c_3, c_4\}$	$\{c_{3,4}\}$
3	$\{c_1\}$	$\{c_1\}$
4	$\{c_2, c_3, c_4\}$	$\{c_{2,3,4}\}$

3.3 セマンティックセグメンテーション

3.1.3.2 節で述べた LR-STF と LP-RTs を用いて、マルチラベルに対応したセマンティックセグメンテーションを行う。処理の流れを図 2 と以下に示す。

Step1 注目画素を中心とした $n \times n$ の画像を切り出す

Step2 切り出した画像の全画素を LR-STF に入力

Step3 末端ノードの到達頻度を数えて STH を算出

Step4 算出した STH を LP-RTs により識別

Step5 Step1 に戻る

以上の処理により、マルチラベルに対応したセマンティックセグメンテーションを実現する。

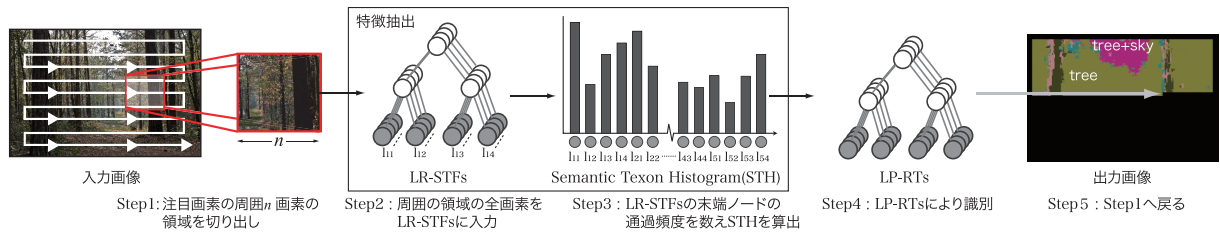


図 2: マルチラベルに対応したセマンティックセグメンテーション

4. 評価実験

4.1 実験データと評価方法

評価実験には, Flickr(<http://www.flickr.com/>) よりダウンロードした画像をラベリングした実験データを用いる. 対象カテゴリは, Face, Body, Tree, Grass, Sky, Flower, Neon の 7 カテゴリ, Grass + Flower, Sky + Tree のマルチラベル 2 カテゴリである. 実験データは, すべて合わせて 1168 枚の画像があり, 学習用に 587 枚の画像, 評価用に 581 枚の画像を利用する.

評価には, 予め作成した正解データを用いて, 以下に定義する Recall をカテゴリ毎に算出して平均した平均再現率を用いる.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

ここで, TP はセグメンテーションが正解した領域の面積, FN は他のカテゴリと間違えた領域の面積である. また, 特徴抽出に用いる画像領域サイズ, LR-STFs と LP-RTs の学習パラメータは実験的に求めた最適なものを利用する.

4.2 マルチラベルの有効性の評価

実験データを用いて, 提案手法の有効性を確認する. 実験データを用いた Confusion Matrix を図 3 に示す. 対角成分が白いほど正しくセグメンテーションできていることを示し, 平均再現率は 70.6% である. 図 3 から, マルチラベルは複数のカテゴリが複合したものであるため, その複合している要素カテゴリに混同されやすいことがわかる. しかし, すべてのカテゴリにおいて対角成分が大きくなりマルチラベルに対応したセマンティックセグメンテーションが可能であるといえる.

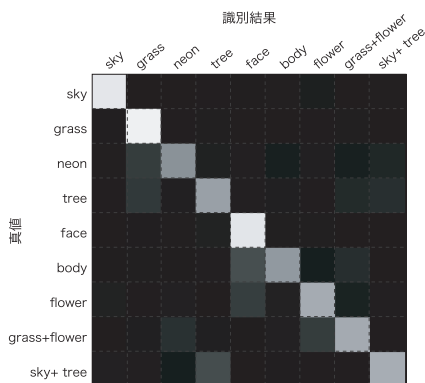


図 3: Confusion Matrix

セグメンテーション例を図 4 に示す. 比較のため, シングルラベルのみを対象とした STF s によるセグメンテーション結果も同時に示す. (a) より, シングルラベルは STF s と同様にセグメンテーション可能であることがわかる. (b) より, STF s は花畑を flower と grass に分割しているのに対し, 提案手法では grass+flower としてそれらの領域を捉えられていることがわかる. (c) より, STF s では木の領域を flower と誤ってセグメンテーションしているのに対し, 提案手法では tree+sky としてセグメンテーション可能であることがわかる. 以上の結果よりマルチラベルを取り入れたことの有効性を確認できた.

4.3 Label Ranking 手法の評価

提案手法では Label Ranking を実現するために STF s の学習アルゴリズムを改良している. 一方, 文献 [3] では, Label Ranking を実現するために, Label Powerset のよ

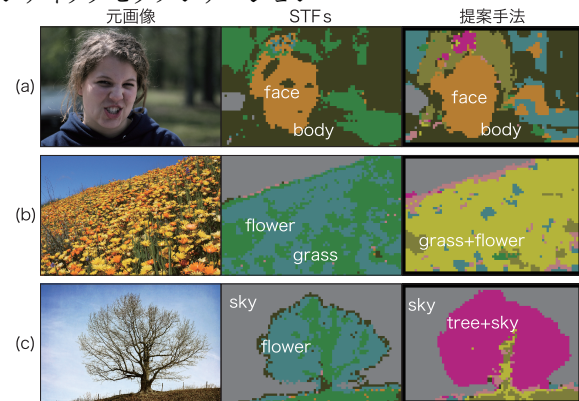


図 4: セグメンテーション結果

うに予め教師信号を変換する手法を用いている. そこで, 提案手法の Label Ranking 手法の有効性を評価するために表 2 に示す文献 [3] の手法と比較を行う.

表 2: 従来の Label Ranking 手法例 [3]

サンプル	ラベル	max	min	Random	Ignore
1	$\{c_1, c_4\}$	c_4	c_1	c_1	
2	$\{c_3, c_4\}$	c_4	c_3	c_4	
3	$\{c_1\}$	c_1	c_1	c_1	c_1
4	$\{c_2, c_3, c_4\}$	c_4	c_2	c_3	

LR-STFs をこれらの手法を用いて構築し, セマンティックセグメンテーションの精度を比較する. 図 5 に比較結果を示す.

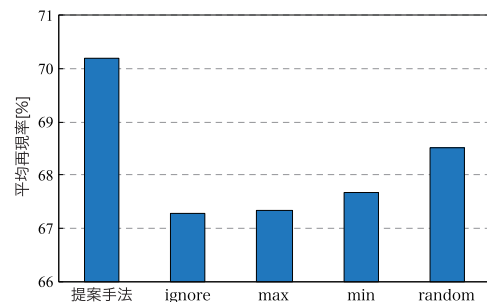


図 5: Label Ranking 手法の比較結果

結果より, 提案手法の平均再現率が最も高くなり提案手法の有効性を確認した.

5. おわりに

本研究では, マルチラベルを考慮した特徴抽出を LR-STFs により実現し, LP-RTs により識別することで, マルチラベルに対応したセマンティックセグメンテーションを実現した. 今後は, 対象カテゴリと, マルチラベルの組み合わせ数を増やしたデータを用いて評価を行う予定である.

参考文献

- [1] J. Shotton, M. Johnson and R. Cipolla, " Semantic Texton Forests for Image Categorization and Segmentation ", Computer Vision and Pattern Recognition, 2008.
- [2] J. Read, " A Pruned Problem Transformation Method for Multi-label classification ", New Zealand Computer Science Research Student Conference ", 2008
- [3] G. Tsoumakas, I. Katakis, I. Vlahavas, " Mining Multi-label Data ", Data Mining and Knowledge Discovery Handbook, 2010.

研究業績

- [1] 福田孝見, 藤吉弘亘, " Label Ranking による Semantic Texton Forests のマルチラベルへの拡張 ", 電子情報通信学会 研究報告 PRMU, 2011.

(他 研究会技術報告 1 件)