

# 確率潜在コンポーネント木による物体カテゴリ構成の学習

Learning Categorical Object Composition based on Probabilistic Latent Component Trees

渥美雅保

Masayasu Atsumi

創価大学工学部情報システム工学科

Dept. of Information Systems Sci., Faculty of Eng., Soka University

This paper proposes a learning method of categorical object clustering and composition analysis based on a probabilistic latent component tree in which probabilistic latent component classes are clustered according to their probabilistic similarity and category labels are assigned to classes in the manner of incremental semi-supervised learning. Through experiments by using images of plural categories in an image database, it is shown that the method works effectively in categorical object clustering, labeling and composition analysis.

## 1. はじめに

人の認識機構はカテゴリ化の能力を有し、情景に含まれる多くの物体のカテゴリを識別することができる。本論では、複数のカテゴリの物体を含む情景画像の集まりから、物体カテゴリの外見の分類体系、及び情景の物体カテゴリ構成を学習する問題を考える。具体的には、あるカテゴリの物体を他の物体と共に含む情景画像の集合が複数のカテゴリに関して与えられて、それらに含まれる物体カテゴリの外見上の分類体系、及び各情景の物体カテゴリの構成を、各情景画像の集合につけられた事前情景カテゴリラベル、及び物体カテゴリ分類後のカテゴリのいくつかについてそれらの代表的な物体セグメントに与えられる事後物体カテゴリラベルに基づき学習する問題を設定する。ここでの課題は、第1に、各情景カテゴリの画像集合に含まれる物体セグメント集合を、それらの外見の特徴に基づきクラスタリングしてその物体カテゴリ構成を求めること、そして第2に、全情景カテゴリの画像集合から得られた物体カテゴリ集合をそれらの間の類似関係に基づき分類体系化するとともに、与えられた限定的な事後物体カテゴリラベルを教師信号として用いて効果的に物体カテゴリを定めること、である。

本研究では、これら課題のために、1) 各情景カテゴリの物体セグメント集合を外見的局所特徴に基づきクラスタリングするための、局所特徴の BoF (Bag of Feature) [Csurka 04] を用いた潜在クラス数可変の確率潜在コンポーネント解析 (Probabilistic Latent Component Analysis, PLCA) [Hofmann 01]、2) 物体カテゴリを表す潜在クラスの全集合の分類体系を与える確率潜在コンポーネント木 (Probabilistic Latent Component Tree, PLCT) の生成、3) 確率潜在コンポーネント木 (PLCT) の葉に割り当てられた潜在クラスの物体カテゴリを、それらの代表的物体セグメント画像に与えられる物体カテゴリラベルを用いてインクリメンタルに更新する半教師付き学習、からなる手法を提案する。

画像に含まれる物体のカテゴリを識別する問題は、一般物体認識の分野で多くの研究がなされている [柳井 10]。また、情景のコンテキストが物体の認識を促進する効果 [Bar 04] に着目し、物体のカテゴリ識別にコンテキストや他の物体との共起を用いる手法 [Galleguillos 08] も提案されている。本研究のアプローチの既存手法との主な違いは、情景のカテゴリ構成を与える共起クラス集合の PLCA による学習と物体カテゴリの分

類木である PLCT の学習を同時に行う点、及びそれらの物体カテゴリラベルを半教師付き学習によりインクリメンタルに更新する点で、人やロボットがある情景の元で他者を教師としたインタラクションを通じて物体カテゴリを学習していく過程を扱える点である。

以下、2章で問題の形式的定義、3章で提案手法の詳細、4章で実験結果について述べる。

## 2. 情景からの物体カテゴリの学習

情景カテゴリの集合を  $C$ 、情景カテゴリの数を  $N_c$  とする。情景カテゴリ  $c \in C$  は、そのカテゴリの物体を他のカテゴリの物体と共に含む情景画像の集合である。情景カテゴリ  $c$  の情景画像  $i$  から分節された  $j$  番目の物体セグメントを  $s_{c,i,j}$ 、情景カテゴリ  $c$  の任意の情景画像から分節された物体セグメントの集合を  $S_c$ 、 $N_{c_s}$  を  $S_c$  の物体セグメントの数とする。

物体セグメントの特徴を局所特徴の BoF により表現する。そのために、まず、各物体セグメントのキーポイントを抽出し、それらキーポイントの 128 次元の SIFT 特徴 [Lowe 04] を計算する。次に、全ての物体セグメントの SIFT 特徴の集合を K-tree 手法 [Shlomo 00] によりクラスタリングしてキー特徴の集合を求める。そして、最後に、このキー特徴の集合をコードブックとして用いて、物体セグメントの BoF を計算し物体セグメントの特徴表現とする。キー特徴の集合を  $F$ 、キー特徴の数を  $N_f$ 、 $F$  の  $n$  番目のキー特徴を  $f_n$  とする。このとき、物体セグメント  $s_{c,i,j}$  の BoF を  $H(s_{c,i,j}) = [h_{c,i,j}(f_1), \dots, h_{c,i,j}(f_{N_f})]$  で表す。  $H(\cdot)$  はキー特徴のヒストグラムである。

問題は、情景カテゴリの集合  $C$  の各情景カテゴリ  $c \in C$  に対して、その情景画像の集合から求められた BoF の集合  $H_c = \{H(s_{c,i,j}) | s_{c,i,j} \in S_c\}$  が与えられて、 $H_c$  から物体カテゴリを表すクラスの集合を、また、全情景カテゴリから求められたクラス的全集合から物体カテゴリを表すクラスを葉ノードに持つ分類木を求めて、さらに、各クラスの物体カテゴリを定めることである。この問題に対して、本研究では、潜在クラス数可変の PLCA 手法、PLCT 生成手法、及び物体カテゴリのインクリメンタルな設定のための半教師付き学習手法を提案し、これら手法を用いて次のようにアプローチする。

第1に、各情景カテゴリ  $c \in C$  の BoF 集合  $H_c$  に対して、潜在クラス数可変の PLCA 手法を適用することにより、情景カテゴリ  $c$  の情景画像を構成する物体カテゴリを表すクラス集

合  $Q_c = \{q_{c,r} | r = 1, \dots, N_{c_q}\}$  を、それらクラスのクラス確率、クラスのキー特徴確率分布、物体セグメントに対するインスタンス確率付きで求める。ここで、クラス確率は物体カテゴリの存在比率、キー特徴確率分布はクラスの表すカテゴリの特徴、インスタンス確率は物体セグメントがそのクラスのインスタンスである程度を与える。また、 $N_{c_q}$  は  $Q_c$  のクラス数である。

第2に、全情景カテゴリから求められるクラスの全集合  $Q = \cup_{c \in C} Q_c$  に対して PLCT 生成手法を適用することにより、クラスを葉に持つ PLCT と呼ぶ2分木を生成する。PLCT においてはクラスのキー特徴確率分布が類似のクラスが近くの葉ノードに配置される。また、ブランチノードにも、そのブランチをルートとする部分木を特徴付けるキー特徴確率分布が求められる。これより、PLCT は物体カテゴリの外見的特徴の分類体系を与える一種のソーラスとみなすことができる。

第3に、PLCT の葉ノードに割り当てられたクラスの物体カテゴリを定めるために、PLCT のブランチノード、及び葉ノードが表す物体カテゴリの仮説規則を定義する。そして、情景カテゴリ名として与えられる事前情景カテゴリラベル、及び葉ノードのクラスのインスタンス確率最大の物体セグメントに対して与えられる事後物体カテゴリラベルを半教師信号として、物体カテゴリ仮説規則に基づき効果的に周辺の葉ノードに物体カテゴリ教師信号を伝播させることにより、葉ノードのクラスの物体カテゴリをインクリメンタルに定める。

### 3. 確率潜在コンポーネント木の学習に基づく物体カテゴリの分類と解析

#### 3.1 確率潜在コンポーネント解析による情景解析

各情景カテゴリ  $c \in C$  の物体カテゴリ構成の学習は、情景カテゴリ  $c \in C$  の BoF 集合  $H_c = \{H(s_{c,i_j}) | s_{c,i_j} \in S_c\}$  に対して、対数尤度

$$L_c = \sum_{i_j} \sum_n h_{c,i_j}(f_n) \log(p(s_{c,i_j}, f_n)) \quad (1)$$

を最大にする  $p(s_{c,i_j}, f_n) = \sum_r p(q_{c,r}) p(s_{c,i_j} | q_{c,r}) p(f_n | q_{c,r})$ 、即ち  $\{p(q_{c,r}) | q_{c,r} \in Q_c\}$ 、 $\{p(s_{c,i_j} | q_{c,r}) | s_{c,i_j} \in S_c, q_{c,r} \in Q_c\}$ 、 $\{p(f_n | q_{c,r}) | f_n \in F, q_{c,r} \in Q_c\}$ 、及び潜在クラスの数  $N_{c_q}$  を推定することによりなされる。与えられた潜在クラスの数のもとでのこれら確率は EM アルゴリズムにより、

[E-step]

$$p(q_{c,r} | s_{c,i_j}, f_n) = \frac{[p(q_{c,r}) p(s_{c,i_j} | q_{c,r}) p(f_n | q_{c,r})]^\beta}{\sum_{q_{c,r'}} [p(q_{c,r'}) p(s_{c,i_j} | q_{c,r'}) p(f_n | q_{c,r'})]^\beta} \quad (2)$$

[M-step]

$$p(f_n | q_{c,r}) = \frac{\sum_{s_{c,i_j}} h_{c,i_j}(f_n) p(q_{c,r} | s_{c,i_j}, f_n)}{\sum_{f_n'} \sum_{s_{c,i_j}} h_{c,i_j}(f_n') p(q_{c,r} | s_{c,i_j}, f_n')} \quad (3)$$

$$p(s_{c,i_j} | q_{c,r}) = \frac{\sum_{f_n} h_{c,i_j}(f_n) p(q_{c,r} | s_{c,i_j}, f_n)}{\sum_{s_{c,i_j}'} \sum_{f_n} h_{c,i_j'}(f_n) p(q_{c,r} | s_{c,i_j'}, f_n)} \quad (4)$$

$$p(q_{c,r}) = \frac{\sum_{s_{c,i_j}} \sum_{f_n} h_{c,i_j}(f_n) p(q_{c,r} | s_{c,i_j}, f_n)}{\sum_{s_{c,i_j}} \sum_{f_n} h_{c,i_j}(f_n)} \quad (5)$$

を繰り返すことにより推定される。ここで、 $\beta$  は温度係数である。

潜在クラスの数の決定は、1つまたはある少数の潜在クラスから初めて、一定回数を上限とする EM の繰返し毎に各クラス内の散らばり度

$$\delta_{c,r} = \sum_{s_{c,i_j}} \left( \sum_{f_n} |p(f_n | q_{c,r}) - D(s_{c,i_j}, f_n)| \right) \times p(s_{c,i_j} | q_{c,r}) \quad (6)$$

ここで、

$$D(s_{c,i_j}, f_n) = \frac{h_{c,i_j}(f_n)}{\sum_{f_n'} h_{c,i_j}(f_n')} \quad (7)$$

を求め、それら値がある閾値より大きい、かつクラス確率がある閾値より大きい潜在クラスのうち散らばり度が最大値をとるクラスを2分割して EM 手続きを続けることを、すべてのクラスの散らばり度が閾値以下、またはクラス確率が閾値以下という停止条件を満たすまで繰り返すことによりなされる。潜在クラスの分割は、元のクラスを  $q_{c,r_0}$ 、分割後のクラスを  $q_{c,r_1}$ 、 $q_{c,r_2}$  とするとき次のようになされる。

一方のクラス  $q_{c,r_1}$  は、元のクラスの最大インスタンス確率の物体セグメント  $s_{c,i_j^*} = \arg \max_{i_j} \{p(s_{c,i_j} | q_{c,r_0})\}$  とその  $\text{BoFH}(s_{c,i_j^*}) = [h_{c,i_j^*}(f_1), \dots, h_{c,i_j^*}(f_{N_f})]$  に対して、キー特徴確率分布、物体セグメントのインスタンス確率、クラス確率を

$$p(f_n | q_{c,r_1}) = \frac{h_{c,i_j^*}(f_n) + \alpha}{\sum_{f_n'} (h_{c,i_j^*}(f_n') + \alpha)} \quad \forall f_n \in F \quad (8)$$

$$p(s_{c,i_j} | q_{c,r_1}) = \begin{cases} p(s_{c,i_j^*} | q_{c,r_0}) & (i_j = i_j^*) \\ \frac{1 - p(s_{c,i_j^*} | q_{c,r_0})}{N_{c_s} - 1} & (\forall i_j (i_j \neq i_j^*) \in S_c) \end{cases} \quad (9)$$

$$p(q_{c,r_1}) = \frac{p(q_{c,r_0})}{2} \quad (10)$$

とすることで定める。ここで、 $\alpha$  は正值の補正係数である。もう一方のクラス  $q_{c,r_2}$  は、キー特徴確率分布  $\{p(f_n | q_{c,r_2}) | f_n \in F\}$  をランダムに、物体セグメントのインスタンス確率  $\{p(s_{c,i_j} | q_{c,r_2}) | i_j \in S_c\}$  を物体セグメント  $s_{c,i_j^*}$  に対して0、それ以外の物体セグメント  $s_{c,i_j} (i_j \neq i_j^*)$  に対して等確率  $\frac{1}{N_{c_s} - 1}$  に、そして、クラス確率を  $p(q_{c,r_2}) = \frac{p(q_{c,r_0})}{2}$  とすることで定める。

すべての潜在クラスが分割停止条件を満たしクラス数が決定するまでは温度  $\beta$  を1.0に固定し、その後、1.0から与えられたスケジュールの元で段階的に減らして Tempered EM 手続きを収束するまで実行する。

#### 3.2 確率潜在コンポーネント木による物体カテゴリ分類

物体カテゴリの外見の分類木 PLCT の学習は、全情景カテゴリから求められたクラスの全集合  $Q^* = \cup_{c \in C} Q_c$  に対して、それらクラスのキー特徴確率分布とクラス確率を用いて、次の手続きにより2分木を生成することによりなされる。

いま、あるブランチノードをルートとする部分木の葉ノードに割り当てられるクラスの集合を  $Q^0 (\subseteq Q^*)$  とし、このブランチノードを  $B(Q^0)$  と表す。ここで、PLCT のルートノードに対しては  $Q^0 = Q^*$  である。このとき、まず、各キー特徴  $f_n \in F$  について、 $Q^0$  をそれに含まれるクラスのキー特徴  $f_n$  の確率値に応じて、2つの部分集合  $Q_{f_n}^1 = \{q_{c,r} | p(f_n | q_{c,r}) \leq \epsilon, q_{c,r} \in Q^0\}$  と  $Q_{f_n}^2 = \{q_{c,r} | p(f_n | q_{c,r}) > \epsilon, q_{c,r} \in Q^0\}$  に分割する。ここで、 $\epsilon$  は0または非常に小さい正の整数で、デフォルトでは0とする。次に、 $Q_{f_n}^1$  と  $Q_{f_n}^2$  について、それぞれに含まれるク

ラスのキー特徴確率分布の平均分布  $\{\mu_{Q^1}^{f_n'}(f_n')|f_n' \in F\}$  と  $\{\mu_{Q^2}^{f_n'}(f_n')|f_n' \in F\}$  を求める．そして、これら平均キー特徴確率分布に対して、次の KL 情報量に基づく距離指標

$$D_{f_n} = \sum_{q_{c,r} \in Q^1_{f_n}} p(q_{c,r}) \left( \sum_{f_n' \in F} p(f_n'|q_{c,r}) \log \frac{p(f_n'|q_{c,r})}{\mu_{Q^1}^{f_n'}(f_n')} \right) \quad (11)$$

$$+ \sum_{q_{c,r} \in Q^2_{f_n}} p(q_{c,r}) \left( \sum_{f_n' \in F} p(f_n'|q_{c,r}) \log \frac{p(f_n'|q_{c,r})}{\mu_{Q^2}^{f_n'}(f_n')} \right)$$

を計算する．最後に、すべてのキー特徴  $f_n \in F$  に対して、 $D_{f_n}$  の最小値を与えるクラス集合分割  $Q^1, Q^2$  を求める．そして、それら部分クラス集合  $Q^1, Q^2$  の各々に対して、クラス数が 1 より大きい場合はそれぞれブランチノード  $B(Q^1)$  と  $B(Q^2)$  を生成し、ブランチノード  $B(Q^0)$  の子ノードとする．一方、クラス数が 1 の場合は、そのクラスを持つ葉ノードを生成し、ブランチノード  $B(Q^0)$  の子ノードとする．以上の処理を、PLCT のルートノード  $B(Q^*)$  から始めて、生成された子ブランチノードに関して再帰的に繰り返す．ただし、ブランチノード  $B(Q^0)$  に割り当てられたクラス数が 2 の場合には、それらの各々のクラスを持つ 2 つの葉ノードを生成し、それらを  $B(Q^0)$  の子ノードとする．

ところで、上記のクラス集合分割において、部分クラス集合  $Q^1, Q^2$  のいずれか一方のクラス数が 0 になる場合が生じうる．そのような場合は、 $\epsilon$  を平均確率値  $\epsilon = \sum_{q_{c,r} \in Q^0} p(f_n|q_{c,r})/|Q^0|$  としてクラス集合分割を行い、子ノードを求める．ここで、 $|Q^0|$  は  $Q^0$  のクラス数を表す．

葉ノードには 1 つのクラスが割り当てられ、そのクラスのクラス確率、キー特徴確率分布、及び物体セグメントに対するインスタンス確率が保持される．ただし、葉ノードのクラス確率は、クラスのクラス確率を情景カテゴリ数で割った値  $p(q_{c,r})/N_c$  とする．また、ブランチノードにも、クラス確率とキー特徴確率分布が保持される．いま、ブランチノードを  $n_p$ 、その子ノードを  $n_{c1}, n_{c2}$  とし、子ノードのクラス確率をそれぞれ  $p(n_{c1}), p(n_{c2})$ 、子ノードのキー特徴確率分布をそれぞれ  $\{p(f_n|n_{c1})|f_n \in F\}, \{p(f_n|n_{c2})|f_n \in F\}$  とする．このとき、ブランチノード  $n_p$  のクラス確率は  $p(n_p) = p(n_{c1}) + p(n_{c2})$ 、キー特徴確率分布  $\{p(f_n|n_p)|f_n \in F\}$  は、

$$p(f_n|n_p) = \frac{p(n_{c1})}{p(n_p)} \times p(f_n|n_{c1}) + \frac{p(n_{c2})}{p(n_p)} \times p(f_n|n_{c2}) \quad (12)$$

により求められる．

### 3.3 物体カテゴリの設定

PLCT の葉ノードに割り当てられたクラスの物体カテゴリの学習では、情景カテゴリ名として与えられる事前情景カテゴリラベル、及びいくつかのクラスのインスタンス確率最大の物体セグメントに順次与えられる事後物体カテゴリラベルを半教師信号として用いて、クラスの物体カテゴリを定める．葉ノードのクラスの物体カテゴリは次の手順で定められる：

1. 葉ノードに物体カテゴリ観測を割り当てる、
2. ブランチノードの物体カテゴリ仮説を葉ノードの物体カテゴリ観測と葉ノードのクラスのクラス確率に基づき定める、
3. 葉ノードの物体カテゴリ仮説を PLCT 上で定義される物体カテゴリ仮説規則に基づき定める．

葉ノードのクラスの物体カテゴリは、PLCT 生成時に事前情景カテゴリラベルを物体カテゴリ観測として用いて初期設定さ

れ、その後、クラスのインスタンス確率最大の物体セグメントに対して与えられる事後物体カテゴリラベルを物体カテゴリ観測として用いてインクリメンタルに更新される．

上記手順の 2 において、ブランチノードの物体カテゴリ仮説は、葉ノードに割り当てられた物体カテゴリ観測と葉ノードのクラスのクラス確率に基づき次のように定められる．いま、ブランチノードを  $B(Q)$ 、それをルートとする部分木の葉ノードに割り当てられるクラス集合を  $Q = \{q_{c,r}\}$ 、それら葉ノードのクラス確率と物体カテゴリ観測のペアの集合を  $\Gamma_{(B(Q))} = \{(p(q_{c,r}), l_{c,r})|q_{c,r} \in Q, l_{c,r} \in L\}$  とする．ここで、 $L$  はカテゴリラベルの集合である．このとき、ブランチノード  $B(Q)$  について、カテゴリ毎にクラス確率の和の集合  $\{p_{B(Q),l} = \sum_{q_{c,r} \in Q} p(q_{c,r})| (p(q_{c,r}), l_{c,r}) \in \Gamma_{(B(Q))}\}$  を求めると、 $p_{B(Q),l}$  はブランチ  $B(Q)$  が物体カテゴリ  $l$  を表す確からしさの指標を与える．そこで、この指標の最大値を与えるカテゴリラベル  $l^* = \arg \max_l \{p_{B(Q),l}\}$  を求め、それをブランチ  $B(Q)$  の物体カテゴリ仮説とする．

上記手順の 3 における葉ノードの物体カテゴリ仮説を定める物体カテゴリ仮説規則は次のように与えられる．いま、PLCT 上でカテゴリルートノードを次の規則により定める．あるブランチノードが次の 2 つの条件を満たすとき、そのブランチノードをカテゴリルートとする：

[条件 1] 兄弟ノードがブランチノードでその物体カテゴリ仮説が異なるか、兄弟ノードが葉ノードでその物体カテゴリ観測が異なる、

[条件 2] 2 つの子ノードに関して次のいずれかが成り立つ：

- (条件 2-1) 両方が同じ物体カテゴリ仮説を持つブランチノードである、
- (条件 2-2) 両方が同じ物体カテゴリ観測が与えられた葉ノードである、
- (条件 2-3) 一方がブランチノード、もう一方が葉ノードで、かつそれらの物体カテゴリ仮説と物体カテゴリ観測が同じである．

一方、ブランチノードの子である葉ノードについては、(条件 2-2) において、2 つの葉ノードに与えられた物体カテゴリ観測が異なるとき、それら葉ノードをカテゴリルートとする．また、(条件 2-3) において、子のブランチノードと葉ノードの物体カテゴリ仮説と物体カテゴリ観測が異なるとき、その葉ノードをカテゴリルートとする．

このとき、葉ノードの物体カテゴリ仮説を、PLCT 上で定められたカテゴリルートノードに対して、次の規則により設定する：

- 葉ノードがカテゴリルートならばそれに与えられた物体カテゴリ観測を葉ノードの物体カテゴリ仮説とする、
- そうでないならば、葉ノードからルートノードに至るまでの経路の直近上位のカテゴリルートノードの物体カテゴリ仮説を葉ノードの物体カテゴリ仮説とする．

PLCT 生成時の葉ノードのクラスの物体カテゴリ仮説の初期設定は、まず、事前情景カテゴリラベルを葉ノードの物体カテゴリ観測として割りあて、次に、それに基づきブランチノードの物体カテゴリ仮説を求め、最後に、葉ノードの物体カテゴリ仮説を物体カテゴリ仮説規則に基づき定めることによりなされる．葉ノードのクラスの物体カテゴリ仮説の更新は、いくつかのクラスのインスタンス確率最大の物体セグメントに対して順次与えられる教師信号としての事後物体カテゴリラベルを用いて、葉ノードに与えられた物体カテゴリ教師信号がその葉ノードに設定された物体カテゴリ仮説と異なるときに次のようになされる．もし、葉ノードに与えられた物体カテゴリ教師

信号がその葉ノードに割り当てられている物体カテゴリ観測，即ち事前情景カテゴリラベルと異なるならば，まず，その物体カテゴリ観測を与えられた物体カテゴリ教師信号，即ち事後物体カテゴリラベルに更新する．次に，その葉ノードからルートノードに至るまでの経路上のノードとその兄弟ノードに関して，ブランチノードの物体カテゴリ仮説の更新とカテゴリルートノードの再判定を行い，最後に，それらの変更がなされたノードをルートとする部分木の葉ノードのすべてに関して，葉ノードの物体カテゴリ仮説を更新する．一方，もし，葉ノードに与えられた物体カテゴリ教師信号がその葉ノードに割り当てられている物体カテゴリ観測と同じならば，その葉ノードの物体カテゴリ仮説を与えられた物体カテゴリ教師信号に変更する．これら更新において，ある葉ノードに与えられた物体カテゴリ教師信号が，ブランチノードの物体カテゴリ仮説の更新と葉ノードの物体カテゴリ仮説規則を介して，その周りの葉ノードにも伝播し，それら周辺の葉ノードの物体カテゴリ仮説も更新される．

#### 4. 実験

##### 4.1 実験枠組

物体カテゴリの分類と構成の学習実験を，MSRC ラベル付き画像 DB \*1 に含まれる 19 個のカテゴリの 269 枚の画像を用いて行った．図 1 に，実験で用いたいくつかのカテゴリの画像と，画像から分節されたセグメントとそのラベルを示す．実験で用いたパラメータは次のとおりである．PLCA のクラス数決定過程における散らばり度とクラス確率の閾値はそれぞれ 1.0 と 0.2，式 (8) の補正係数  $\alpha$  は 1.0 である．クラス数決定後の Tempered EM 手続きにおいて，温度係数  $\beta$  は 20 回の繰返し毎に 0.95 倍ずつ 0.8 まで減少させた．また，K-tree 手法により求められたキー特徴のコードブックのサイズは 533 で，学習にはキーポイント数が 50 以上のセグメントを用いた．

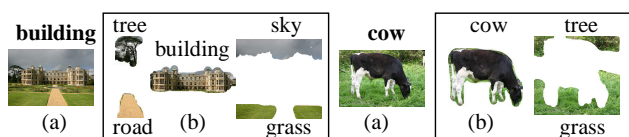


図 1: Examples of (a) images and (b) segments with labels.

##### 4.2 実験結果

情景カテゴリ 19 個の各々から PLCA により生成されたクラスの総数は 114 個で，これらクラスに対して生成された PLCT において，各クラスのインスタンス確率最大の物体セグメントに付けられたラベルを葉ノードのクラスの物体カテゴリの正解とみなしたとき，情景カテゴリラベルを物体カテゴリ観測とした物体カテゴリ設定のもとで 114 個のクラス中 66 個 (58%) に正しい物体カテゴリが設定された．図 2 に生成時の PLCT の一部を示す．葉ノードの物体カテゴリ仮説規則により，情景カテゴリラベル “bird” が割り当てられた葉ノード L3 の物体カテゴリ仮説が正しく “water” に設定されている．

クラスのインスタンス確率最大の物体セグメントに付けられたラベルを教師信号として，クラス確率の高い葉ノードから順に与えたとき，物体カテゴリ仮説規則のもとで 114 個のクラス中 107 個 (94%) に正しい物体カテゴリが設定され，さらに続けることによりすべてのクラスに正しい物体カテゴリが設定された．図 3 に更新中の PLCT の一部を示す．葉ノードの物体カテゴリ仮説規則により，葉ノード L106 に教師信号

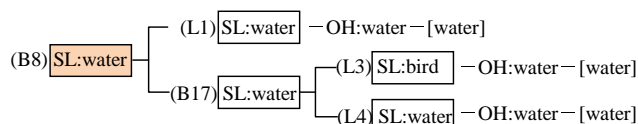


図 2: A part of an initial PLCT. SL: A scene category label, OH: An object category hypothesis, [ ]: A representative segment, and a colored square indicates a categorical root.

“cow” を与えることで，葉ノード L106 と L110 の物体カテゴリ仮説が正しく “cow” に更新されている．

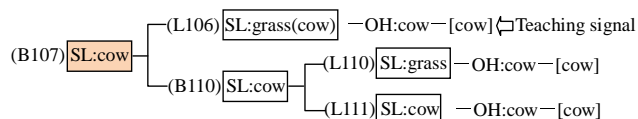


図 3: A part of an updated PLCT.

これら学習による情景カテゴリ “dog” の構成例を，クラスのインスタンス確率最大の物体セグメントを用いて図 4 に示す．物体クラスは情景カテゴリラベルと同じ物体カテゴリ仮説を持つクラスで，それ以外がコンテキストクラスである．

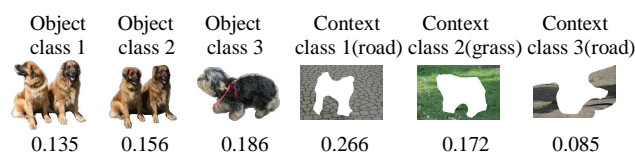


図 4: An example of categorical object composition.

#### 5. おわりに

本論では，情景カテゴリ毎に集められた画像の集合から情景カテゴリの物体カテゴリ構成，及び物体カテゴリの分類体系を学習し，それら物体カテゴリにインクリメンタルにラベル付けをする確率潜在コンポーネント木に基づく手法について述べた．そして，MSRC ラベル付き画像 DB を用いた実験により，本手法の持つ機能的特徴が有効に働くことを確かめた．

#### 参考文献

- [Bar 04] Bar, M.: Visual Objects in Context, Nature Reviews Neuroscience, Vol.5, pp.617-629 (2004)
- [Csurka 04] Csurka, G., Bray, C., Dance, C. and Fan, L.: Visual Categorization with Bags of Keypoints, Proc. of ECCV Workshop on Statistical Learning in Computer Vision, pp.1-22 (2004)
- [Galleguillos 08] Galleguillos, C., Rabinovich, A. and Belongie, S.: Object Categorization using Co-Occurrence, Location and Appearance, Proc. of IEEE CS Conf. on CVPR pp.1-8 (2008)
- [Hofmann 01] Hofmann, T.: Unsupervised Learning by Probabilistic Latent Semantic Analysis, Machine Learning, Vol.42, pp.177-196 (2001)
- [Lowe 04] Lowe, D. G.: Distinctive Image Features from Scale-Invariant Keypoints, International Journal of Computer Vision, Vol.60, No.2, pp.91-110, (2004)
- [Shlomo 00] Shlomo, G.: K-Tree; A Height Balanced Tree Structured Vector Quantizer, Proc. of the 2000 IEEE Signal Processing Society Workshop, Vol.1, pp.271-280 (2000)
- [柳井 10] 柳井啓司: 一般物体認識における機械学習の利用, 信学技報, Vol.110, No.76, IBISML2010-16, pp.103-112 (2010)

\*1 <http://research.microsoft.com/vision/cambridge/recognition/>