

# 물체-행동 컨텍스트를 이용하는 확률 그래프 기반 물체 범주 인식

윤성백\*, 배세호\*, 박한재\*, 이준호<sup>o</sup>

## Probabilistic Graph Based Object Category Recognition Using the Context of Object-Action Interaction

Sung-baek Yoon\*, Se-ho Bae\*, Han-je Park\*, June-ho Yi<sup>o</sup>

### 요약

다양한 외형 변화를 가지는 물체의 범주 인식성능을 향상 시키는데 있어서 사람의 행동은 매우 효과적인 컨텍스트 정보이다. 본 연구에서는 Bayesian 접근법을 기반으로 하는 간단한 확률 그래프 모델을 통해 사람의 행동을 물체 범주 인식을 위한 컨텍스트 정보로 활용하였다. 다양한 외형의 컵, 전화기, 가위 그리고 스프레이 물체에 대해 실험을 수행한 결과 물체의 용도에 대한 사람의 행동을 인식함으로써 물체 인식 성능을 8%~28% 개선할 수 있었다.

**Key Words** : object recognition, object-action context, object-human interaction

### ABSTRACT

The use of human actions as context for object class recognition is quite effective in enhancing the recognition performance despite the large variation in the appearance of objects. We propose an efficient method that integrates human action information into object class recognition using a Bayesian approach based on a simple probabilistic graph model. The experiment shows that by using human actions as context information we can improve the performance of the object class recognition from 8% to 28%.

### 1. 서론

다양한 외형을 가지는 물체의 범주(object class)를 인식하는 문제는 매우 어려운 문제로 인식 되어왔다. 사람이 사용하는 물체는 그 용도가 정해져 있고 그 용도에 대응하는 물체에 대한 사람의 행동들을 연계할 수 있으므로 사람의 행동을 물체의 용도를 나타내는 컨텍스트 정보로서 활용하여 물체의 범주 인식 성능

향상이 가능하다. 이에 본 연구는 베이지안(Bayesian) 방식을 통해 사람의 행동을 컨텍스트 정보로 사용하여 물체 범주 인식을 하고자 한다.

사람의 행동 및 물체 외형을 표현하는 방법과 이를 활용하는 인식 알고리즘은 상이하나 본 연구와 유사한 아이디어를 채택한 연구가 있었다. 사람의 포즈(pose)를 물체 검출을 위한 컨텍스트 정보로 활용하는 연구로는 베이지안 방식을 사용하여 사람의 포즈 추

\* 본 연구는 한국연구재단 논문연구과제(NRF-0213R1A1A2006164) 지원 및 성균관 대학교 산학 협력단의 관리로 수행되었습니다.

• First Author: Electrical and Computer Engineering, Sungkyunkwan University, beagii@skku.edu, 정희원

o Corresponding Author: Electronic and Electrical Engineering, Sungkyunkwan University/ Information and Communication Engineering, North university of China, jhyi@skku.edu, 정희원

\* Electrical and Computer Engineering, Sungkyunkwan University, pr0411@skku.edu, kanje0602@skku.edu

논문번호 : KICS2015-07-235, Received July 23, 2015; Revised October 13, 2015; Accepted October 13, 2015

정 및 물체를 검출하는 연구<sup>[1]</sup>와 바디 파트와 물체간의 연결 관계를 모델링하는 random field를 이용하는 연구<sup>[2]</sup>가 있다. 물체를 사용하는 일상적인 행동에서 상호간에 겹치는 포즈들로 인해 하나의 포즈가 어떤 행동의 일부인지 명확하지 않기 때문에 하나의 포즈만 사용하는 것은 오히려 물체의 범주를 모호하게 할 수 있다. 물체를 컨텍스트 정보로 활용하여 사람의 행동을 인식하고 사람의 행동과 물체의 관계를 통해 물체 검출 및 인식을 하는 연구<sup>[3]</sup>은 손의 경로로 사람의 행동을 표현하고 HMM(Hidden Markov model)을 활용하여 인식한다. 행동을 정확하게 인식하기 위해서는 각 개인의 행동 속도로 인해 발생하는 서로 다른 길이의 손의 경로를 정확하게 비교할 수 있는 방법이 요구된다. 물체와 사람의 행동을 상호간의 컨텍스트 정보로서 사용하는 연구<sup>[4]</sup>은 물체 인식과 행동 인식을 결합하여 사람과 물체의 상호 작용(interaction)을 인식한다. 손의 경로를 추적하여 HMM을 통해 몇 개의 atomic action으로 나누어 사람의 행동을 표현한다. 하지만 각각의 행동을 여러 행동 간에 공유되는atomic action으로 정확하게 분할하는 것이 어렵고 유사한 외형을 갖는 물체로 인식 대상이 한정되어있기 때문에 물체의 외형변화가 다양한 경우에는 적용하기 어렵다.

본 연구의 목적은 다양한 외형을 갖는 물체 인식 성능을 향상시키기 위해 사람의 행동을 효과적으로 활용하는 것이다. 사람의 행동을 몇 개의 키 포즈(key pose)로 나타내고 간단한 확률 그래프 모델을 통해 물체 인식을 위한 컨텍스트 정보로 활용한다. 본 연구의 기여점은 다음과 같다. 첫째, 물체 범주 인식을 위해 물체에 연계되는 사람의 행동을 컨텍스트로 사용하는 것이다. 둘째, 랜덤 포레스트(random forest)<sup>[5]</sup>와 multi-class Adaboost<sup>[6]</sup>를 활용하여 학습한 물체와 사람의 행동에 대한 확률 분포를 이용하는 효율적인 베이저안 구현이다.

물체-행동(object-action) 컨텍스트를 이용하는 물체 범주 인식은 그림 1과 같다. 먼저 훈련 동영상에서 획득한 행동을 나타내는 특징 벡터와 물체 영상으로부터 획득한 특징 벡터들을 사용하여 각각 독립적으로 행동에 대한 랜덤 포레스트와 물체에 대한 랜덤 포레스트를 훈련한다. 이를 각각 행동 랜덤 포레스트와 물체 랜덤 포레스트라고 한다. 그리고 각 랜덤 포레스트에 속하는 트리를 약한 분류기로 간주하고 multi-class Adaboost<sup>[6]</sup>를 활용하여 트리에 대한 가중치를 학습하여 행동과 물체의 특징 벡터로부터 각각 행동과 물체의 범주에 대한 사후 확률을 계산한다. 입력 데이터에 대한 물체의 범주는 행동과 물체의 특징

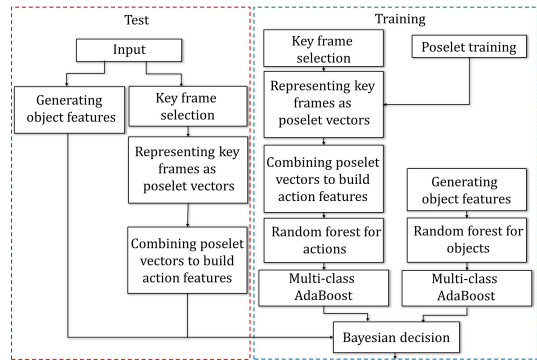


그림 1. 물체-행동 컨텍스트를 활용한 물체 범주 인식 개요  
Fig. 1. Object class recognition using object-action context

벡터로부터 계산한 각각의 확률을 베이저안 방식에 적용하여 결정한다. 행동을 나타내는 특징 벡터는 동영상의 키 프레임(key frame)들을 poselet<sup>[7]</sup> 벡터로 표현하고, 이를 하나의 벡터로 합쳐서 표현한다. Poselet 벡터의 원소들의 값은 키 프레임에서 각각의 poselet에 대한 최대응답 값이며, 이는 SVM(Support Vector Machine)을 통해 획득한다. Poselet은 그림 5에서 보여지듯이 사람의 포즈의 부분을 표현한 것으로 외형 정보를 기준으로 매우 엄밀하게 분류된 특징 벡터이다. 각 범주에 속하는 물체는 HOG(Histogram of Oriented Gradients)<sup>[8]</sup>로 표현한다.

본 논문의 구성은 다음과 같다. 2절에서는 물체 범주 인식을 위한 물체-행동 컨텍스트에 대한 확률그래프 모델을 다룬다. 3절에서는 랜덤 포레스트와 multi-class Adaboost<sup>[6]</sup>를 이용하여 물체와 사람의 행동에 대한 확률 분포 학습 방법에 대해 알아본다. 4절에서는 물체와 사람의 행동 표현 방법을 5절에서는 실험결과를 다룬다.

## II. 물체 범주 인식을 위한 확률 모델

본 연구의 목적은 사람의 행동을 컨텍스트 정보로 활용하여 물체의 범주를 인식하는 것이다. 그림 2는 이를 위한 확률 그래프 모델을 보여준다.  $O$ 는 물체의 범주를,  $A$ 는 사람의 행동 범주를,  $x^o$ 는 물체의 외형 특징을 그리고  $x^a$ 는 물체에 대응되는 사람의 행동 특징을 각각 나타낸다.

$x^o$ 와  $x^a$ 가 주어졌을 때 물체의 범주에 대한 확률,  $p(A|x^o, x^a)$ 은 베이저안 룰을 사용하여 식 (1)과 같이 나타낼 수 있다.

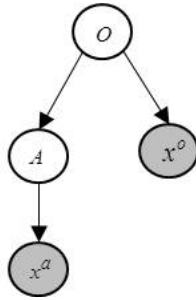


그림 2. 사람의 행동을 컨텍스트 정보로 활용하는 물체 범주 인식을 위한 그래프 모델  
Fig. 2. Our graph model for recognizing object class using human action as context information

$$\begin{aligned}
 p(O|x^o, x^a) &= \frac{p(x^o, x^a, O)}{p(x^o, x^a)} \\
 &= \frac{\sum_A p(x^o, x^a, O, A)}{\sum_O \sum_A p(x^o, x^a, O, A)} \\
 &= \frac{\frac{p(O|x^o)}{p(O)} \sum_A p(O, A) p(A|x^a)}{\sum_O \frac{p(O|x^o)}{p(O)} \sum_A p(O, A) p(A|x^a)}
 \end{aligned} \tag{1}$$

물체의 범주에 대한 확률  $p(O|x^o, x^a)$ 는 식 (2)와 같이 근사화 할 수 있다.

$$p(O|x^o, x^a) \propto p(O|x^o) \sum_A p(O, A) p(A|x^a) \tag{2}$$

물체의 특징  $x^o$ 가 주어졌을 때의 물체 범주에 대한 확률,  $p(O|x^o)$ 와 사람의 행동에 대한 특징  $x^a$ 가 주어졌을 때의 확률,  $p(A|x^a)$  그리고  $p(O, A)$ 를 학습하는 방법은 다음 절에서 알아보기로 한다.

### III. 물체와 사람의 행동에 대한 확률 분포 학습

물체의 범주에 대한 확률은 식 (2)와 같이  $p(O|x^o)$ 와  $\sum_A p(O, A) p(A|x^a)$ 의 곱에 비례하는 양으로 근사화 할 수 있다. 먼저  $p(O|x^o)$ 와  $p(A|x^a)$ 를 학습하는 과정을 알아보고  $p(O, A)$ 를 계산하는 방법을 설명한다.

$p(O|x^o)$ 와  $p(A|x^a)$ 는 랜덤 포레스트와 multi-class Adaboost를 활용하여 학습한다. 그림 3은 물체 범주에 대한 확률 계산 과정을 보여준다. 랜덤 포레스트의 각 트리,  $\theta_i$ 에서 획득한 물체 범주에 대한

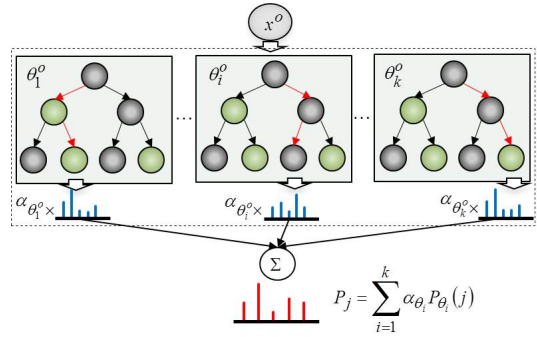


그림 3. 랜덤 포레스트와 multi-class Adaboost를 통한 물체 범주에 대한 확률 계산  
Fig. 3. Computation of the probabilities of object classes using random forest and multi-class Adaboost

확률,  $P_{\theta_i}(j)$ 에 multi-class Adaboost를 사용하여 학습한 트리에 대한 가중치,  $\alpha_i$ 를 곱하여 합한 확률 값,  $P_j$ 가 물체 범주의 확률이 된다.

물체의 범주에 대한 확률  $p(O|x^o)$ 를 훈련하는 과정은 다음과 같다. 먼저, 훈련 데이터  $S = \{(x^o, O_1), \dots, (x^o, O_n)\}$ ,  $x^o = \{x_1^o, \dots, x_M^o\}$ 가 주어졌을 때,  $k$ 개의 트리,  $\theta_1, \dots, \theta_k$ 로 이루어진 랜덤 포레스트를 생성한다.  $x_i^o$ 는 물체 범주  $O_j$ 에 속하는  $i$ 번째의 물체의 특징을 나타낸다. 또한, 랜덤 포레스트에 속하는 각각의 트리,  $\theta_i^o$ 를 약한 분류기로 보고 multi-class Adaboost 알고리즘<sup>[9]</sup>을 활용하여, 각 트리에 대한 가중치를 훈련한다. 물체 범주에 대한 확률,  $p(O|x^o)$ 는 식 (3)과 같다.

$$\begin{aligned}
 p(O_j|x^o) &\equiv p(O_j|x^o, \theta^o) \\
 &= \sum_{i=1}^k p(O_j|x^o, \theta_i^o) p(\theta_i^o|D) = \frac{1}{|\theta^o|} \sum_{i=1}^k \alpha_{\theta_i^o} \frac{n_{O_j, \theta_i^o}}{n_{\theta_i^o}}
 \end{aligned} \tag{3}$$

식 (3)에서  $\theta^o$ 는 물체 외형 특징으로부터 생성된 랜덤 포레스트이고  $\theta_i^o \in \theta^o$ 는  $i$ 번째 결정 트리,  $D$ 는 훈련 데이터 그리고  $p(\theta_i^o|D) = 1/|\theta^o|$ 이다.  $n_{O_j, \theta_i^o}$ 는 트리  $\theta_i^o$ 의 리프(leaf) 노드에서 물체 범주,  $O_j$ 로 분류되는 훈련 데이터의 개수이고  $n_{\theta_i^o}$ 는 트리  $\theta_i^o$ 의 리프 노드에서의 총 훈련 데이터 개수이다.  $|\theta^o|$ 는 랜덤 포레스트에 속하는 트리의 개수를 나타낸다.  $\alpha_{\theta_i^o}$ 는 랜덤 포레스트의 각 트리에 대한 가중치이다. 사람의 행동 특징을 사용하여 훈련하는 확률,  $p(A|x^a)$ 의 생성과

정도 위와 동일하다.

그림 4는 multi-class Adaboost 알고리즘을 통해 랜덤 포레스트의 각 트리에 대한 가중치 학습과정이다. 이는 순차적으로 분류규칙을 생성하고 이전 분류규칙을 사용하여 훈련 데이터의 분포를 재조정하는 이진 분류기인 Adaboost의 약한 분류기에 대한 가중치 학습과정과 동일하다. 학습 데이터가 총  $n$ 개, 분류하고자 하는 범주(class)가  $C$ 개 일 때, 첫 번째 단계에서 훈련 데이터에 대한 초기 분포  $w$ 를 계산한다. 분포  $w$ 는  $k$ 회 반복 되면서 그 값이 업데이트되고, 높은 값일 수록 분류가 잘 되지 않은 데이터이다. 단계 2에서는 약한 분류기  $T^{(m)}(x)$ 의 에러를 계산하고 이를 바탕으로  $w$ 를 갱신한다. 최종적으로 단계 3에서  $k$ 회 반복하면서 구한 가중치  $\alpha$ 와 약한 분류기인 각 트리에서 생성한 확률 분포의 선형 결합으로 강한 분류기인 물체 범주에 대한 확률 분포를 획득한다. 단계 2의 (c)에서  $\log(C-1)$ 이 추가 된 것이 이진 분류기인 Adaboost와 다른 점이다. 이진 분류기의 경우 random guessing의 에러율(error rate)이 1/2인 것과는 다르게 multi-class인 경우에는 에러율이  $(C-1)/C$ 이기 때문에 약한 분류기의 에러율이 1/2보다 작다는 Adaboost의 가정이 만족 되지 않는다. 따라서, 이러한 Adaboost의 단점을 보완하기 위해서  $\log(C-1)$ 이 추가 되었다.

1. Initialize the observation weights  $w_i = 1/n \quad i = 1, \dots, n$
2. For  $m = 1$  to  $k$ 
  - (a) Fit a classifier  $T^{(m)}(x) = \arg \max_{\Theta} \frac{1}{|\Theta|} \cdot p(O|x, \theta_m)$  to the training data using weights  $w_i$
  - (b) Compute  $err^{(m)} = \sum_{i=1}^m w_i I(c_i \neq T^{(m)}(x)) / \sum_{i=1}^n w_i$
  - (c) Compute  $\alpha^{(m)} = \log \frac{1 - err^{(m)}}{err^{(m)}} + \log(C-1)$
  - (d) Set  $w_i \leftarrow w_i \cdot \exp(\alpha^{(m)} \cdot I(c_i \neq T^{(m)}(x)))$   
for  $i = 1, \dots, n$
  - (e) Renormalize  $w_i$
3. Outputs  $p(O|x) = \frac{1}{|\Theta|} \sum_{m=1}^M \alpha^{(m)} \cdot p(O|x, \theta_m)$

그림 4. Multi-class Adaboost를 활용하는 랜덤 포레스트의 각 트리에 대한 가중치 훈련  
Fig. 4. Training the weight of each tree in random forest using multi-class Adaboost

$p(OA)$ 는 식 (4)와 같이 베이지안 룰을 사용하여  $p(A|O)$ 를 통해 계산한다.

$$p(OA) = \frac{p(A|O)p(O)}{\sum_O p(A|O)p(O)} \quad (4)$$

$p(A|O)$ 는 동일한 범주에 연계되는 행동 특징들에 대한 행동 범주의 수로 계산되며, 식 (5)와 같다.

$$p(A = a_j | O = o_i) = \frac{n_j}{N_i} \quad (5)$$

식 (5)의  $N_i$ 는 물체 범주  $o_i$ 에 연계되는 행동 특징들의 총 개수이며,  $n_j$ 는 그 중에서  $a_j$ 를 행동 범주로 갖는 행동 특징들의 개수이다. 행동 특징에 대한 행동 범주,  $a_j$ 는  $k$ -NN(Nearest Neighbors)을 활용하여 훈련 집합에 속하는 행동 특징의 이웃들로부터 다수표를 획득한 행동 범주로 결정된다.

#### IV. 물체와 사람의 행동 표현

사람의 행동은 연속되는 포즈들의 집합으로 볼 수 있다. 하지만 인접한 프레임들의 포즈는 매우 유사하기 때문에 동영상의 모든 프레임으로부터 포즈를 추출하여 사용하는 것은 불필요한 중복을 야기 시킨다. 따라서, 행동을 표현 할 수 있는 최소한의 키 프레임을 추출하고 이를 poselet 벡터로 표현한다. 선택된 키 프레임들에 대한 poselet 벡터들을 시간 순서대로 결합하면 행동을 나타내는 특징이 된다.

그림 6은 키 프레임을 poselet 벡터로 표현하고 이를 행동에 대한 특징으로 나타내는 과정을 보여준다. 키 프레임을 poselet 벡터로 표현하기 위해서는 그림 5에서와 같이 주석이 달린 poselet들을 HOG<sup>[8]</sup>를 이용하여 나타내고 각 poselet들에 대한 분류기로 SVM을 사용한다. 동영상의 키 프레임에서 슬라이딩 윈도우 방법을 적용하여 훈련된 모든 poselet들에 대한 최대 응답 값으로 poselet 벡터를 생성한다. 사람의 행동을 세 개의 키 프레임으로부터 획득한 poselet 벡터의 결합으로 나타내었으며, 총 38종류의 poselet을 사용하였기 때문에 각 프레임에 대한 poselet 벡터는 38차원이 되며 행동 특징은 114차원의 크기를 가진다.

입력 동영상에서 키 프레임을 추출하기 위해서 동영상의 각 프레임을 poselet 벡터로 표현하고 훈련 데이터의 키 프레임들과 유사한 시간대에 있는 입력 동



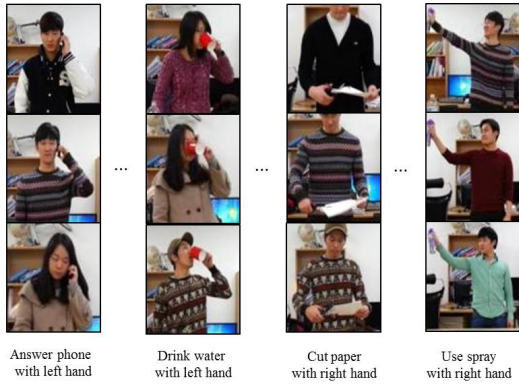


그림 5. Poselet의 예  
Fig. 5. Examples of poselets

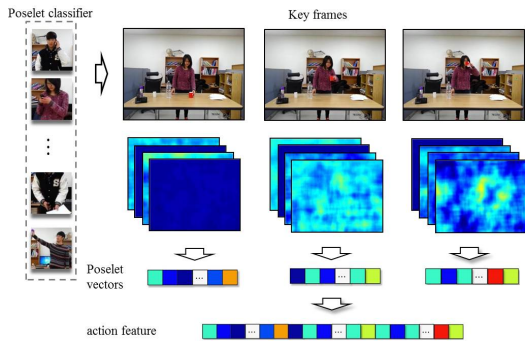


그림 6. Poselet 벡터 및 행동에 대한 특징 생성 과정  
Fig. 6. Creation of poselet vectors and action features

여상의 프레임들의 poselet 벡터와 훈련데이터의 poselet 벡터의 거리를 계산하여 최소 값을 갖는 프레임을 키 프레임으로 선택한다. 이때, poselet 벡터들간의 거리는 유클리디안(Euclidean) 거리를 사용하여 측정한다.

물체를 표현하기 위해 Haar-like 특징<sup>9,10)</sup>과 direct pixel-mapping (DPM)<sup>11)</sup>을 사용하는 방법들이 보고되었으나 본 연구에서는 보행자 및 물체의 인식률이 높다고 알려진 HOG를 사용하여 각 범주에 속하는 물체를 나타낸다. 물체 영상의 크기는 50×50이다.

### V. 실험 결과

물체 범주 인식을 위해 사람의 행동을 컨텍스트 정보로 활용하는 대표적인 연구인 Gupta et al.<sup>4)</sup>의 방법과 성능을 비교하였다. Gupta et al.<sup>4)</sup>은 사람의 포즈를 컨텍스트 정보로 활용하는 다른 연구와는 달리 사람의 행동을 컨텍스트 정보로 활용하고 인식 대상이 되는 물체로 컵, 전화기, 플래쉬 라이트 그리고 스프

레이를 사용하여 본 연구와 유사하다.

본 연구는 실험을 위해 19명을 대상으로 컵, 가위, 전화기 그리고 스프레이 총 네 종류의 물체를 사용하는 동영상들 각각의 물체에 대해 세 번씩 촬영하여 총 228개의 동영상 데이터 베이스를 구축하였다.

Poselet 분류기를 학습하기 위해 38 종류의 포즈 영상 20,308개와 부영상 2,321개를 사용하였으며, 각 poselet을 학습하기 위한 훈련 영상의 크기는 96×94이다. 분류기로는 linear SVM이 사용 되었으며 하나의 poselet 종류를 제외한 다른 모든 부류의 poselet은 부영상으로 활용하여 학습하였다.

행동 랜덤 포레스트를 훈련하기 위해 물 마시는 행동 특징 1,625개, 전화를 거는 행동 7,149개, 종이 자르는 행동 1,674개 그리고 스프레이 뿌리는 행동 678개를 사용 하였으며, multi-class Adaboost를 학습하기 위해 각각 848개, 1,890개, 330개 그리고 658개 사용 되었다.

실험에 사용된 물체 영상을 Google Image Search와 ImageNet으로부터 획득하였으며, 컵 3,120장, 전화기 4,131장, 가위 2,263장 그리고 스프레이 2,006장이다. 이중 각 범주에서 1,200장을 사용하여 물체 랜덤 포레스트를 훈련하고, 600장은 multi-class Adaboost 학습에 사용되었다. 그림 7은 실험에 사용된 물체 영상의 일부분이다. 동일한 범주에 속하는 물체들의 다양한 외형뿐만 아니라 컵과 스프레이와 같이 서로 다른 물체 범주에 속하지만 원통형 구조로 인해 유사한 외형을 가지는 물체 또한 포함하고 있다.

그림 8은 물체 범주 인식 결과를 보여주는 confusion 행렬들이다. 그림 8의 첫 번째 열은 물체의

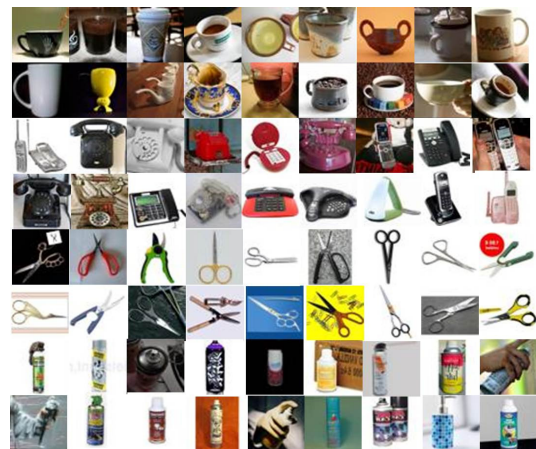


그림 7. 실험에 사용된 물체 영상의 일부  
Fig. 7. The part of object images used in experiments

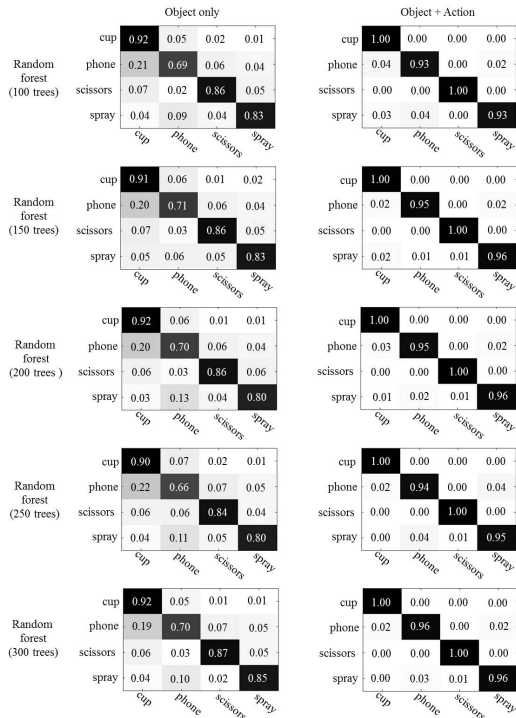


그림 8. 물체 범주 인식의 결과(첫 번째 열은 물체의 외형 정보만 사용한 물체 범주 인식 결과, 두 번째 열은 물체의 외형정보와 사람의 행동정보를 컨텍스트 정보로 사용한 물체 범주 인식 결과)

Fig. 8. The results of object recognition. The left column shows the results using only the object's appearance and the right column represents the results using the object's appearance and human actions

외형 정보만을 사용한 물체 범주 인식 결과이고, 두 번째 열은 물체의 외형 정보와 사람의 행동을 컨텍스트 정보로 활용한 물체 범주 인식의 결과이다. 랜덤 포레스트의 트리 개수를 100, 150, 200, 250 그리고 300으로 변화시켜가며 실험을 진행 하였다. 그림 8에서 보여지듯이 전체적으로 사람의 행동을 활용한 물체 범주 인식의 성능이 물체의 외형 정보만을 사용한 경우보다 월등히 향상되었음을 알 수 있다. 작게는 8%(컵)이상의 인식을 향상을 보이고 많게는 28%(전 화기)의 성능 향상을 보이고 있다.

그림 9는 Gupta et al.<sup>[4]</sup>의 알고리즘을 실험 데이터에 적용한 결과이다. 컵을 제외한 다른 물체들은 본 연구에서 제안한 알고리즘보다 낮은 인식 성능을 보이고 있다. 전화기는 27%~28%, 가위는 33% 그리고 스프레이의 경우 6%~7%의 성능 차이를 보인다. 이는 사람의 행동을 atomic action으로 정확하게 분할하지 못한 것이 낮은 인식 성능의 원인으로 판단된다. 실험

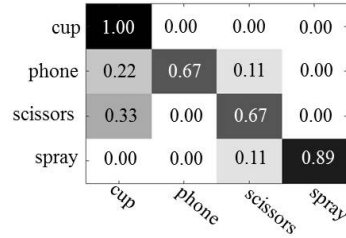


그림 9. 실험 데이터에 Gupta et al.[4] 알고리즘을 적용한 결과

Fig. 9. The result of the Gupta et al.[4] algorithm for our experimental data

결과를 통해 본 연구에서 제안하는 알고리즘이 사람의 행동을 물체 인식을 위한 컨텍스트 정보로 활용하는 것에 있어서 더욱 효과적임을 알 수 있다.

## VI. 결론

본 연구에서는 물체의 범주 인식 성능 향상을 위해 물체의 용도를 나타내는 사람의 행동을 컨텍스트 정보로 활용하는 방법을 제안하고 실험 하였다. 이를 통해 사람의 행동 정보는 다양한 물체의 외형 변화에도 불구하고 물체의 범주 인식을 효과적으로 향상시킬 수 있는 컨텍스트 정보임을 확인할 수 있었다.

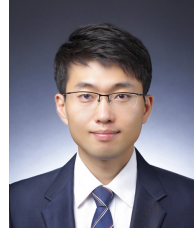
## References

- [1] V. Singh and R. Nevatia, "Multiple pose context trees for estimating human pose in object context," in *CVPR Workshop on Structural Models in Computer Vision*, pp. 17-24, San Francisco, CA, USA, Jun. 2010.
- [2] B. Yao and L. FeiFei, "Recognizing human-object interactions in still images by modeling the mutual context of objects and human poses," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 9, pp. 1691-1703, Sept. 2012.
- [3] D. J. Moore, I. A. Essa, and M. H. Hayes, "Exploiting human actions and object context for recognition tasks," in *Proc. Int. Conf. Computer Vision*, vol. 1, pp. 80-86, Corfu, Greece, Sept. 1999.
- [4] A. Gupta, A. Kembhavi, and L. S. Davis, "Observing human-object interactions: Using spatial and functional compatibility for

recognition,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 10, pp. 1775-1789, Oct. 2009.

- [5] W. Fan, E. Greengrass, and J. McCloskey, “Effective estimation of posterior probabilities: Explaining the accuracy of randomized decision tree approaches,” in *Proc. Int. Conf. Data Mining*, pp. 154-161, Houston, Texas, USA, Nov. 2005.
- [6] J. Zhu, H. Zou, S. Rosset, and T. Hastie, “Multi-class Adaboost,” *Stats. Its Interface*, vol. 2, no. 3, pp. 349-360, 2009.
- [7] L. D. Bourdev and J. Malik, “Poselets: Body part detectors trained using 3d human pose annotations,” in *Proc. Int. Conf. Computer Vision*, pp. 1365-1372, Kyoto, Japan, Sept. 2009.
- [8] N. Dalal and B. Triggs, “Histogram of oriented gradients for human detection,” in *Proc. Computer Vision and Pattern Recognition*, pp. 886-893, San Diego, CA, USA, Jun. 2005.
- [9] B. W. Chung, K. Y. Park, and S. Y. Hwang, “A fast and efficient Haar-like feature selection algorithm for object detection,” *J. KICS*, vol. 38A, no. 6, pp. 486-491, Jun. 2013.
- [10] K. Y. Park and S. Y. Hwang, “An improved normalization method for Haar-like features for real-time object detection,” *J. KICS*, vol. 36, no. 8, pp. 505-515, Aug. 2013.
- [11] H. G. Chung and E. S. Kim, “Improved recognition of far objects by using DPM method in curving-effective integral imaging,” *J. KICS*, vol. 37A, no. 2, pp. 128-134, Feb. 2013.

**윤성백 (Sung-baek Yoon)**



2005년 8월 : 성균관대학교 전자전기컴퓨터공학부 졸업  
 2007년 8월 : 성균관대학교 전자전기컴퓨터공학과 석사  
 2007년 9월~현재 : 성균관대학교 전자전기컴퓨터공학과 박사과정

<관심분야> 컴퓨터 비전

**배세호 (Se-ho Bae)**



2015년 2월 : 성균관대학교 전자전기공학부 졸업  
 2015년 3월~현재 : 성균관대학교 전자전기컴퓨터공학과 석박사 과정

<관심분야> 컴퓨터 비전

**박한재 (Han-je Park)**



2014년 8월 : 성균관대학교 전자전기공학부 졸업  
 2014년 9월~현재 : 성균관대학교 전자전기컴퓨터공학과 석박사 과정

<관심분야> 컴퓨터 비전

**이준호 (June-ho Yi)**



1985년 2월 : 서울대학교 전기공학과 졸업  
 1987년 2월 : Pennsylvania State University Electrical Engineering 석사  
 1994년 2월 : Purdue University Electrical Engineering 박사

1997년 3월~현재 : 성균관대학교 전자전기공학부 교수  
 <관심분야> 컴퓨터 비전