

Real-time Multiple Object Instances Detection

Chengli Xie, Jinqiao Wang, Yifan Zhang, Hanqing Lu

National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences.
Beijing, China

{clxie, jqwang, yfzhang, luhq}@nlpr.ia.ac.cn

ABSTRACT

In this paper, we present a novel, real-time multiple object instance detection system via template matching and pairwise classification. Instance detection aims to find and locate exactly the same object instances as specified. Our system is composed of two heterogeneous stages. The first stage adopts instance-specific detection to generate candidates. And the second stage makes use of a pairwise-based classifier across instance categories to test and verify these candidates with respect to templates. Experiments show the superiority of our approach.

Categories and Subject Descriptors

I.4.9 [Image Processing and Computer Vision]: Application; J.7.

General Terms

Algorithms Design Experimentation

Keywords

Instance detection, pairwise classification, multiple instances

1. INTRODUCTION

Scientists and researchers have studied object detection for several decades. However, nowadays with the advance in smartphone and mobile computing, people often would like to go one step further. That is to say, they not only want to detect objects in images they took, but also they want to know what it is specifically of the detected objects. This is what we so called object instance detection. Take a flower image for example, object detection tells us whether there is a flower and where is it. While instance detection goes beyond to tell whether there is a peony and where is the peony.

Instance detection is a variation of object detection. However, there are several differences between object detection and instance detection technically. First, the difference between instances is much more subtle to be captured than that of object classes. Another difference between object detection and instance detection lies in the fact that the latter usually has much fewer (positive) samples for training compared to that of the former.

Instance detection is also quite different to other existing techniques such as fine-grained object recognition and image retrieval. Given an example of handbag recognition, these three techniques have different goals. Instance detection attempts to find and locate a specific handbag (for instance a specific type of Louis Vuitton handbag) in images. Object fine-grainedrecognition

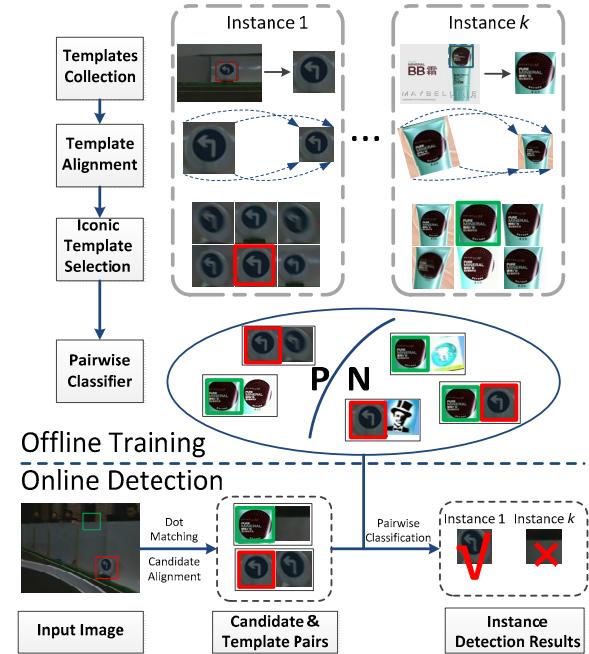


Figure 1. Overview of our method.

just wants to know whether the image is a handbag or not. And in image retrieval, people just find images in a large database which has similar global appearance to the given query handbag image.

We present a template-based real-time multiple object instance detection system (IDS). The IDS system utilizes an online detector for each instance and an overall offline trained cross-instances classifier validator. The detector generates candidates for each instance and the validator test whether the candidates are real instances with respect to the template.

2. SYSTEM OVERVIEW

In this paper, we propose a coarse to fine cascade framework of two stages to combine template matching and pairwise classification for multiple object instances detection, for the purpose of robust and rapidly discovering specified instances of object classes (Figure 1). Our two heterogeneous stages IDS system combines template matching and pairwise classification for multiple instances detection. Figure 1 shows the corresponding system diagram. Our detector and validator take gradient information, texture information and appearance features into consideration.

Offline training: At the beginning, users choose an instance template by drawing a box. And to reduce the instability by initial template selection, the train examples are aligned and one iconic template is selected which could minimize the overall distance with others. Followed this, a pairwise classifier across instance categories is trained to predict whether the two elements in a pair belong to the same instance.

Online detection: In this section, DOT matching is applied firstly, where the matched image patches are called candidates. Then similarity between a candidate and its relative template is measured. Finally, the candidate is validated by classifying the similarity measurement via the learned pairwise classifier.

3. ALGORITHM DESCRIPTION

In the IDS system, first stage is a detector stage, which makes use of multi-angle multi-scale gradient template to extract candidates based on a variation of DOT [1][3]. Following, pairwise classification is done in the second stage, a validator stage. A candidate and its corresponding template compose a pair. For each pair, this stage extracts its similarity feature and classifies them to further filter false positive candidates by judge whether the two elements in a pair are the same object instance.

Note that the proposed method is not restricted to specific descriptors or classifiers in each layer. Taking the 2nd stage as an example, MLP can be replaced by SVM or other classifiers.

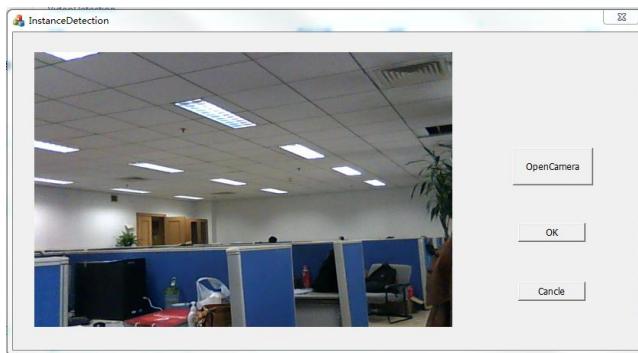


Figure 2. System Interface

3.1 Detector Stage

A variation [3] of Dominant Orientation Template matching (DOT) [1] is the algorithm we apply. In our application, a merge and non-max suppress manipulator is embedded. And template is updated by iconic image in initialization.

3.2 Validator Stage

This stage contains two steps. The first step calculates the similarity feature of a pair. And then the similarity feature is classified to validate the candidate in this pair in the second step.

Specifically, in the first step a local texture descriptor, histogram of pyramid LTP (pLTP), is computed for each candidate and its principal component (PCA) is calculated to reduce the dimension of the pLTP descriptor. Then a similarity measurement is extracted from the PCA components of a pair. While in the second step, machine learning method (currently used multi-layer perception neuron network) is employed to validate whether the

two elements in a pair are the same by classifying the similarity measure got above so as to further filter false positive candidates.

In offline training procedure, all the similarity features across pairs composed by all the training samples are used to learn a multilayer perception neuron network. In this way, samples number increases relative to the traditional 1-vs-all method and classifiers number reduce to 1, which is good for training.

Our pairwise method provides a way to solve few positive training samples problem among multiple object instances detection. Another benefit for the pairwise classifier is that we can generate to unlearned instance.

4. EXPERIMENT

Several instance categories are tested in our experiments, including book titles, notebook, pictures and different logos. We don't train all the instance categories for our pairwise classifiers. And you can find that the untrained instance can also be detected well in video.

Please note that in the whole video, no tracking technology is applied, because we would like to generalize this work to data forms where tracking may not be applicable, such as images.

The implement use Microsoft Visual C++ and MFC with a Logitech web camera on an Intel Core2Duo computer with 2.8GHz CPU and 4 GB RAM. Our approach runs at 18 frame/s for detecting 5 object instances simultaneously.

5. CONCLUSION

In this paper, we study the problem of multiple object instances detection and propose a cascade framework to deal with it. Iconic template selection with template image alignment improves the stability of template matching. While pairwise classification compensates the scarcity of positive samples and could generate to unlearned instances to a certain extent. Finally, accurate detection results (category instances) are found effectively.

With the help of OpenCV on Android [2], this system can be transplant to Smartphone technically. That is one of our future works.

6. ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China 973 Program (Project No. 2010CB327905) and the National Natural Science Foundation of China (Grant No. 60905008, 61070104)

7. REFERENCES

- [1] S. Hinterstoisser, V. Lepetit, S. Ilic, P. Fua and N. Navab. 2010. Dominant Orientation Templates for Real-Time Detection of Texture-Less Objects. In Proceedings of the CVPR, 2010.
- [2] <http://sourceforge.net/projects/opencvlibrary/files/opencv-android/>
- [3] C. Xie, J. Li, T. Wang, J. Wang and H. Lu. Real time cascade template matching for object instance detection. PCM 2011.