

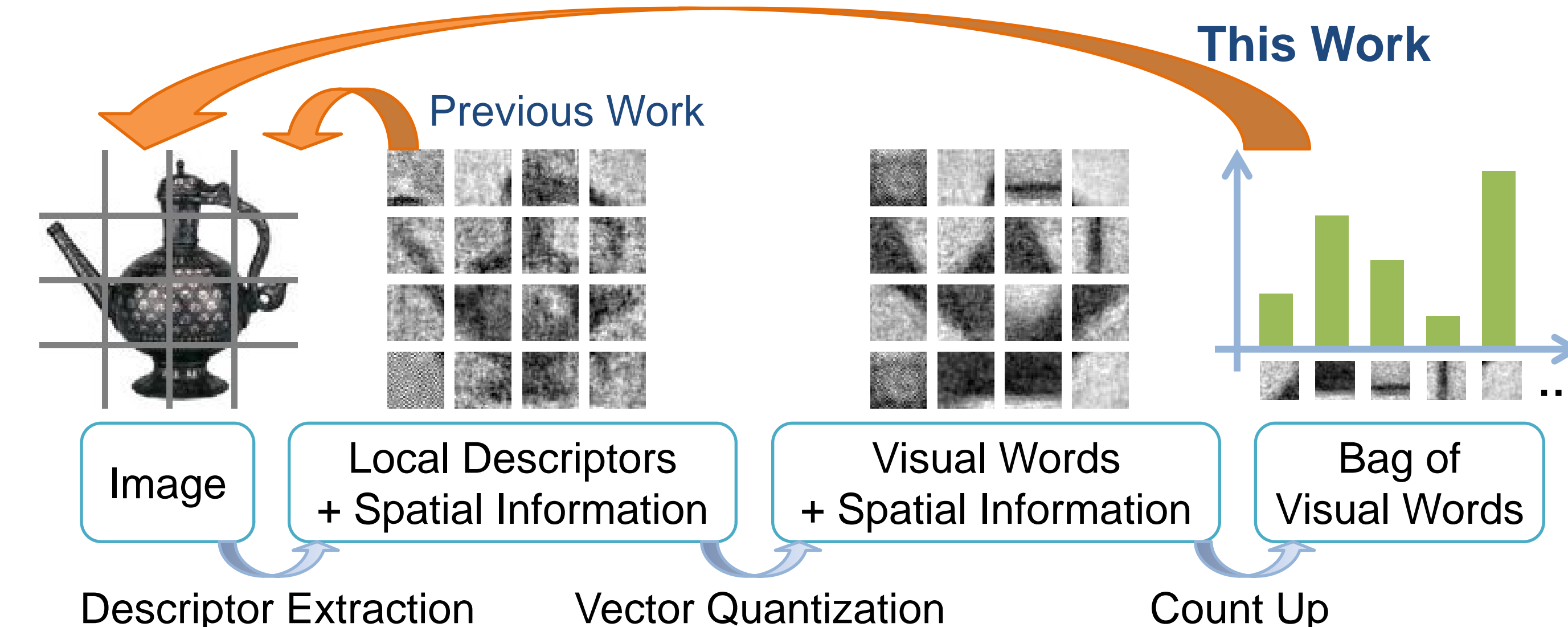


Overview

- **Background:** to reconstruct an image from its image feature is useful for understanding the feature intuitively, however Bag-of-Visual-Words (BoVW) has not been reconverted yet.
- **Objective:** to reconstruct images from BoVW.
- **Main problem:** BoVW lacks spatial information of visual words.
- **Solution:** spatial arrangement of them is estimated like solving a jigsaw puzzle, using statistics of local co-occurrences and absolute positions of visual words in an image database.
- **Contributions:**
 1. It is the first work to reconstruct images from BoVW.
 2. A method to estimate the spatial arrangement of visual words by using an image database is proposed.
 3. Relations among our problem, jigsaw puzzle problem, and the Quadratic Assignment Problem is shown.

Extraction of Bag-of-Visual-Words

- **Assumptions:** dense and single scale sampling.



Reconstruction Method

Our proposed method consists of two steps.

1. Estimation of spatial arrangement of visual words
 2. Generation of an image patch from each visual word image patch
- For the latter, we use HOGgles*¹. For the former, to assign n visual words in an image at n grid points, we solve the following optimization problem.

$$\min \lambda \sum_{i,j,k,l=1}^n C_{ijkl}^a x_{ik} x_{jl} + (1-\lambda) \sum_{i,k=1}^n C_{ik}^l x_{ik}$$

$$\text{s.t.} \quad \sum_{i=1}^n x_{ik} = 1, \quad \sum_{k=1}^n x_{ik} = 1, \quad x_{i,k} \in \{0,1\}$$

- $x_{ik} = 1$ if the i -th visual word in BoVW is assigned to k -th grid point.
- **Adjacency Cost C^a** gives a reconstructed image consistent edges and shapes.



- **Global Location Cost C^l** makes a reconstructed image globally feasible.






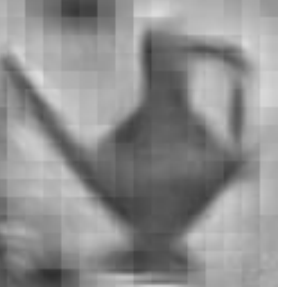




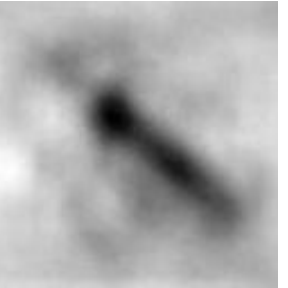
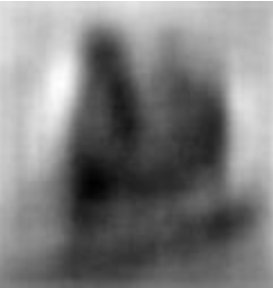
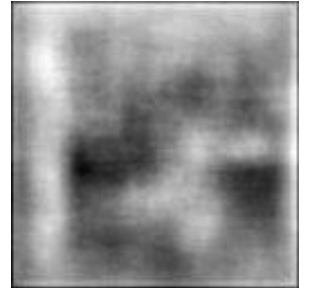

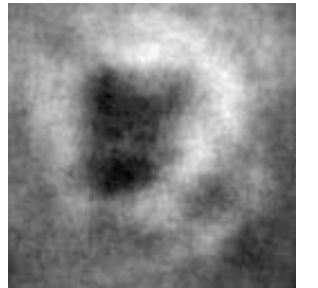







Optimization

- This problem can be result in the Quadratic Assignment Problem.
- Solved by Hybrid algorithm of Genetic Algorithm and Hill Climbing.

Experimental Results

We used 101 object images. Five results of them are shown here.

		Good Results				
Original Images						
Obtained from BoVW	Our method					
	HOGgles* ¹					
	Image Retrieval					

This is for HOG originally

From 1M images

*1 Vondrick et al., ICCV, 2013.

[Settings]

- Image size: 128x128 px
- Descriptor size: 32x32 px
- Local descriptor: SIFT
- Vocabulary size of BoVW: 5000
- Descriptor extraction step: 8px

Discussion and Conclusion

- **Reconstructability:** images of single objects are reconstructed more stably than that of complicated textures.
- **Computational Cost:** About a minute for one image. The bottleneck is optimization.
- **Limitations:** there are several assumptions which are desirable to be relaxed, such as single scale sampling and hard assignment of local descriptors.

All experimental results are available at
<http://www.mi.t.u-tokyo.ac.jp/kato/cvpr2014.html>

