Scalable Recognition with a Vocabulary Tree

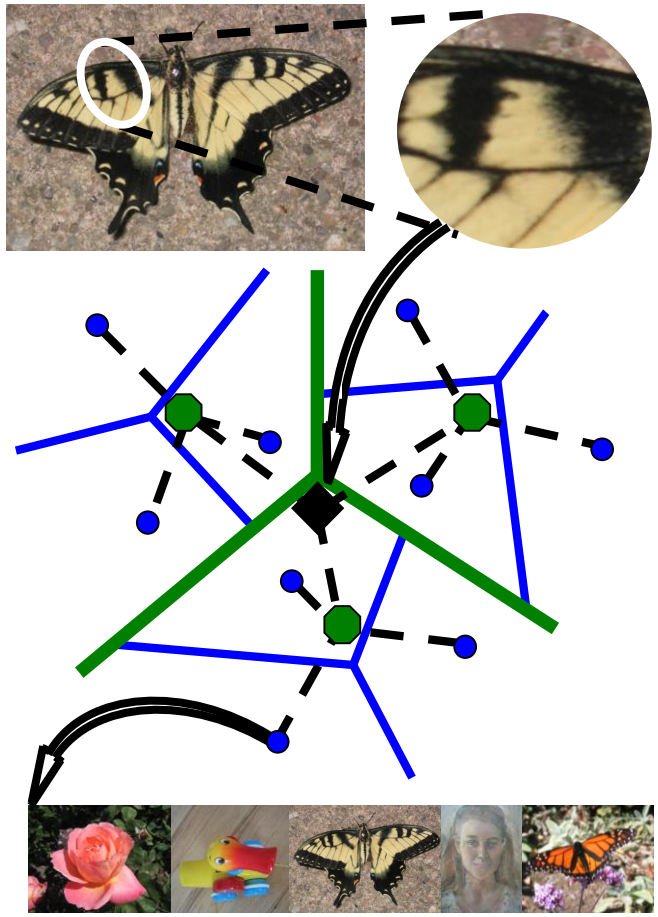
David Nistér, Henrik Stewénius

UK
Center for Visualization & Virtual Environments

New York City
CVPR
2006

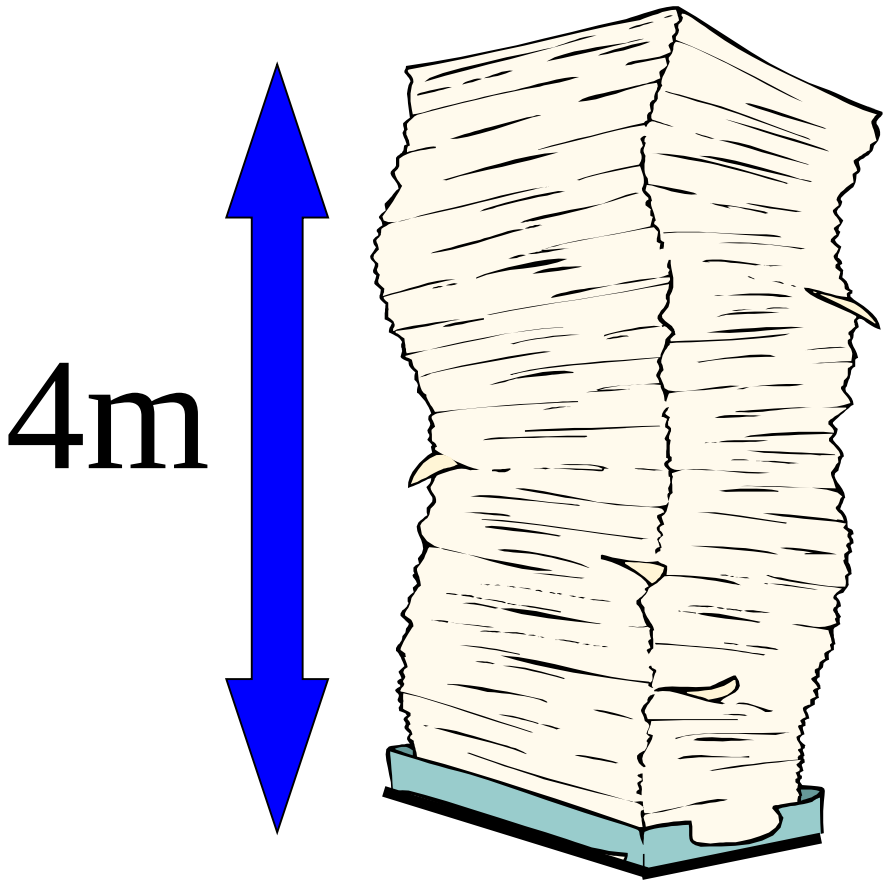
Scalable Recognition with a Vocabulary Tree

David Nistér, Henrik Stewénius



UK
Center for Visualization & Virtual Environments
New York City CVPR 2006

50 Thousand Images



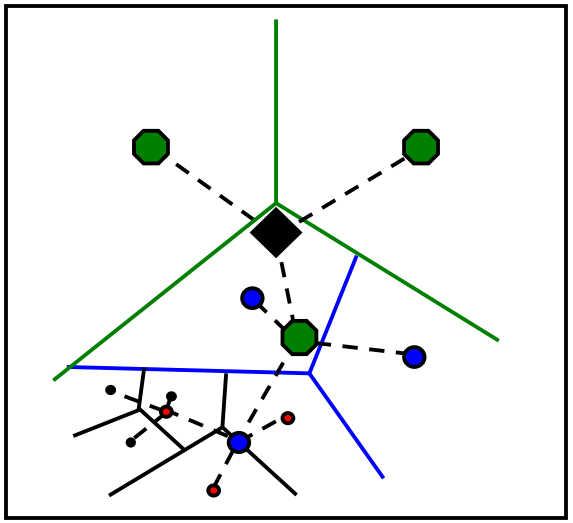
110,000,000 Images in 5.8 Seconds

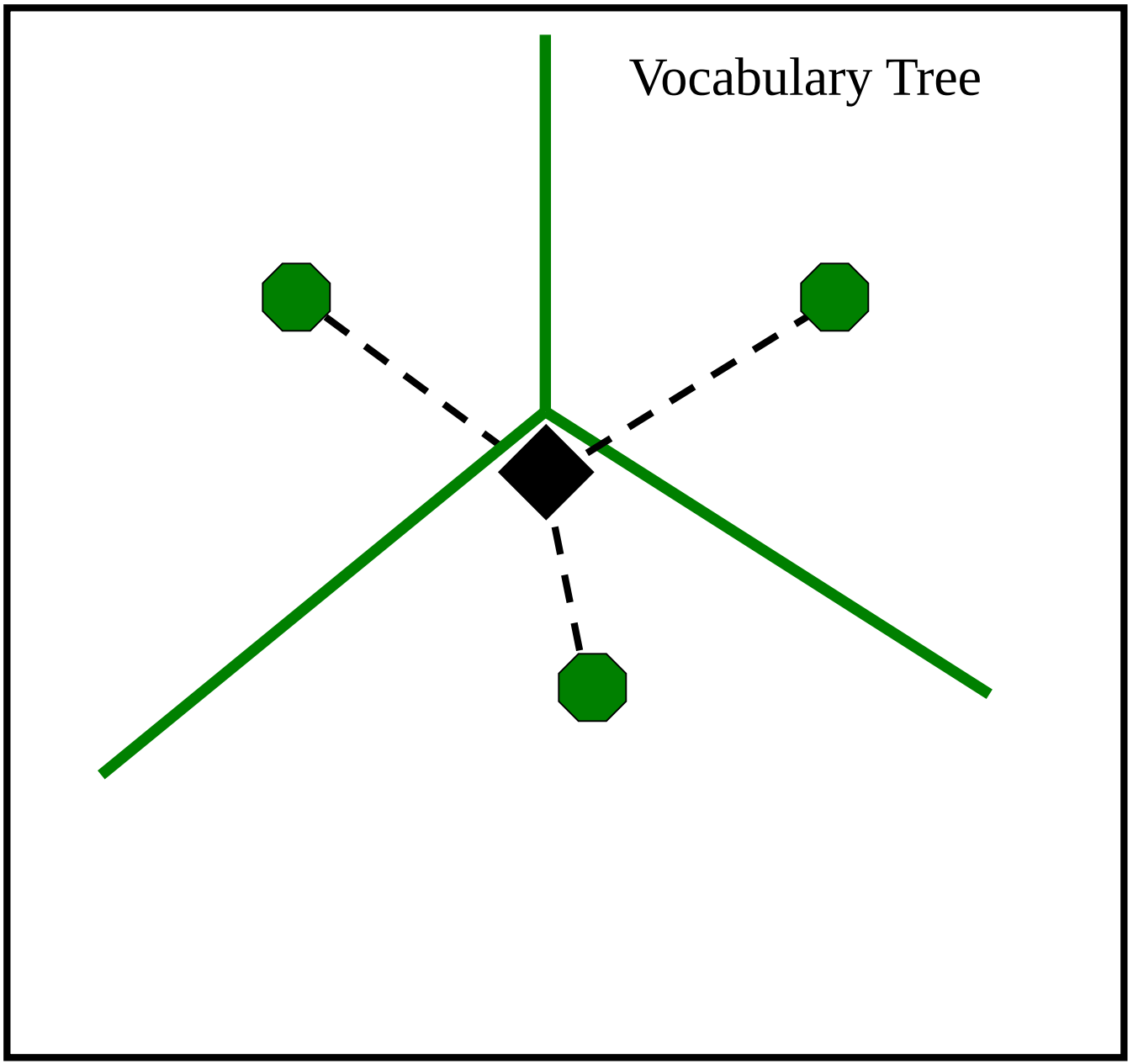
MetLife
Kodak
EASYSHARE Digital Camera and Dock
Shoot.
Touch.
Share.
Ask Merrill
Music
PLANET HOLLYWOOD
New York Sightseeing
GRAY LINE
71317

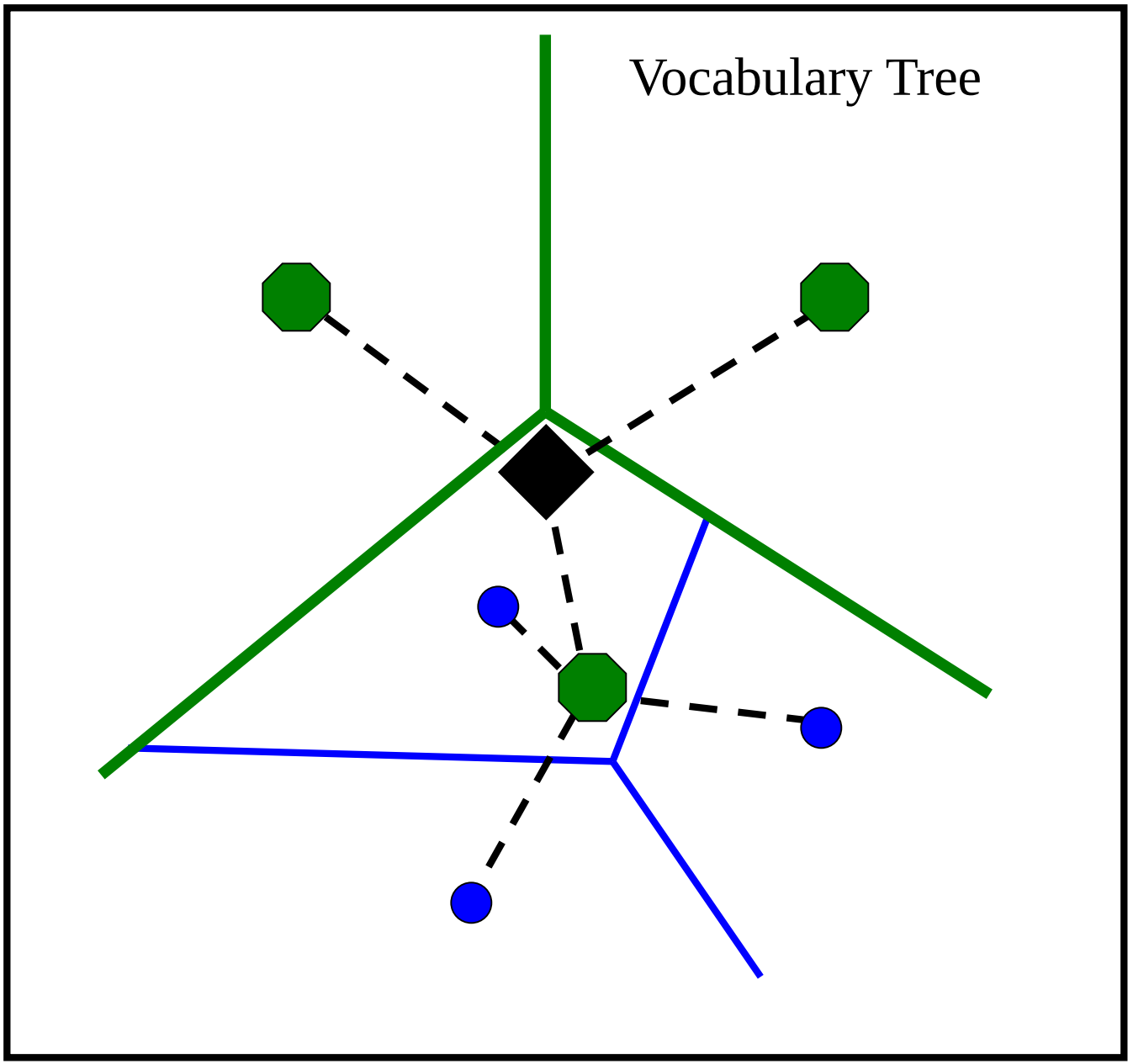
Take-Home Message

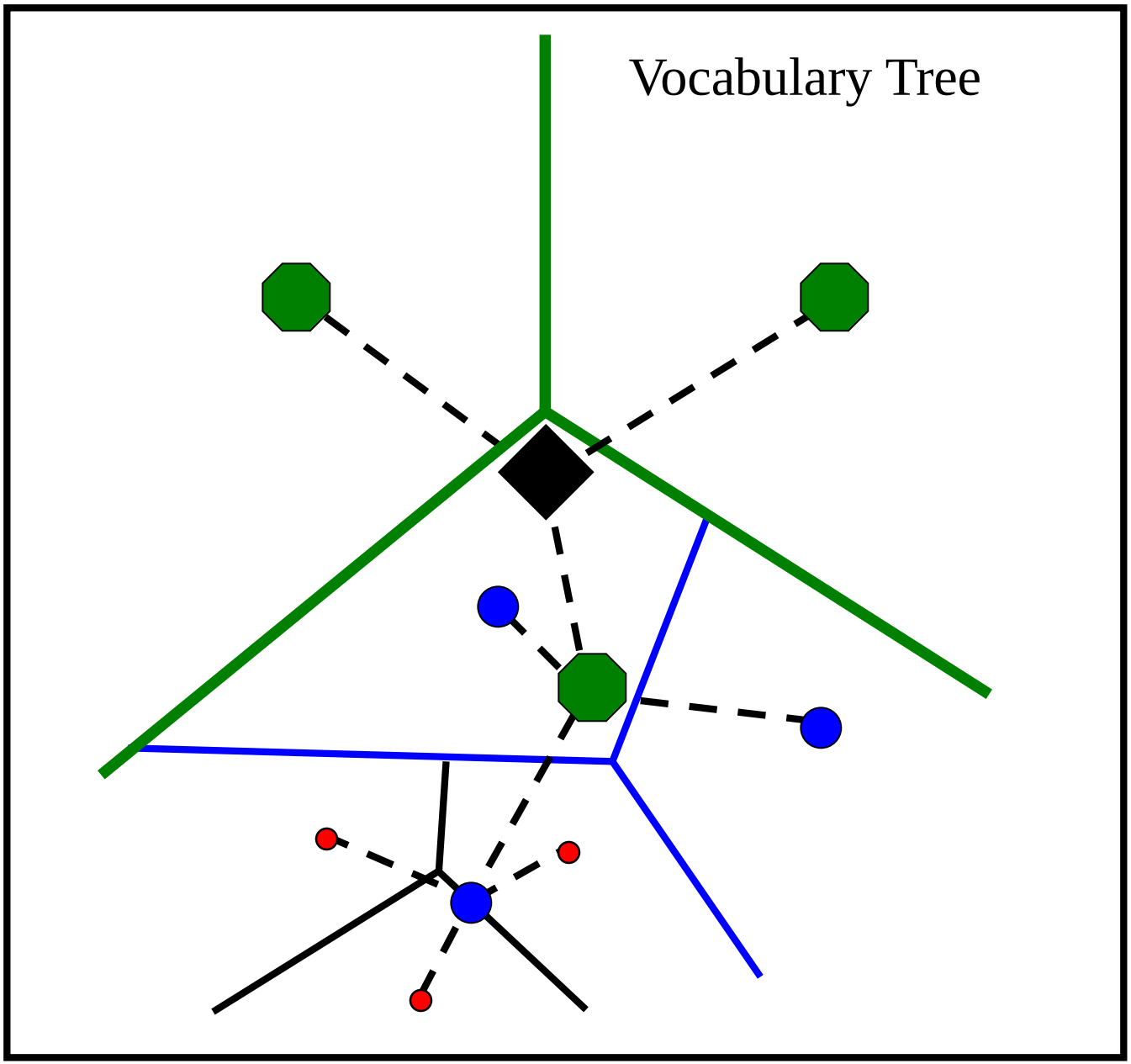
If we can get repeatable, discriminative features,

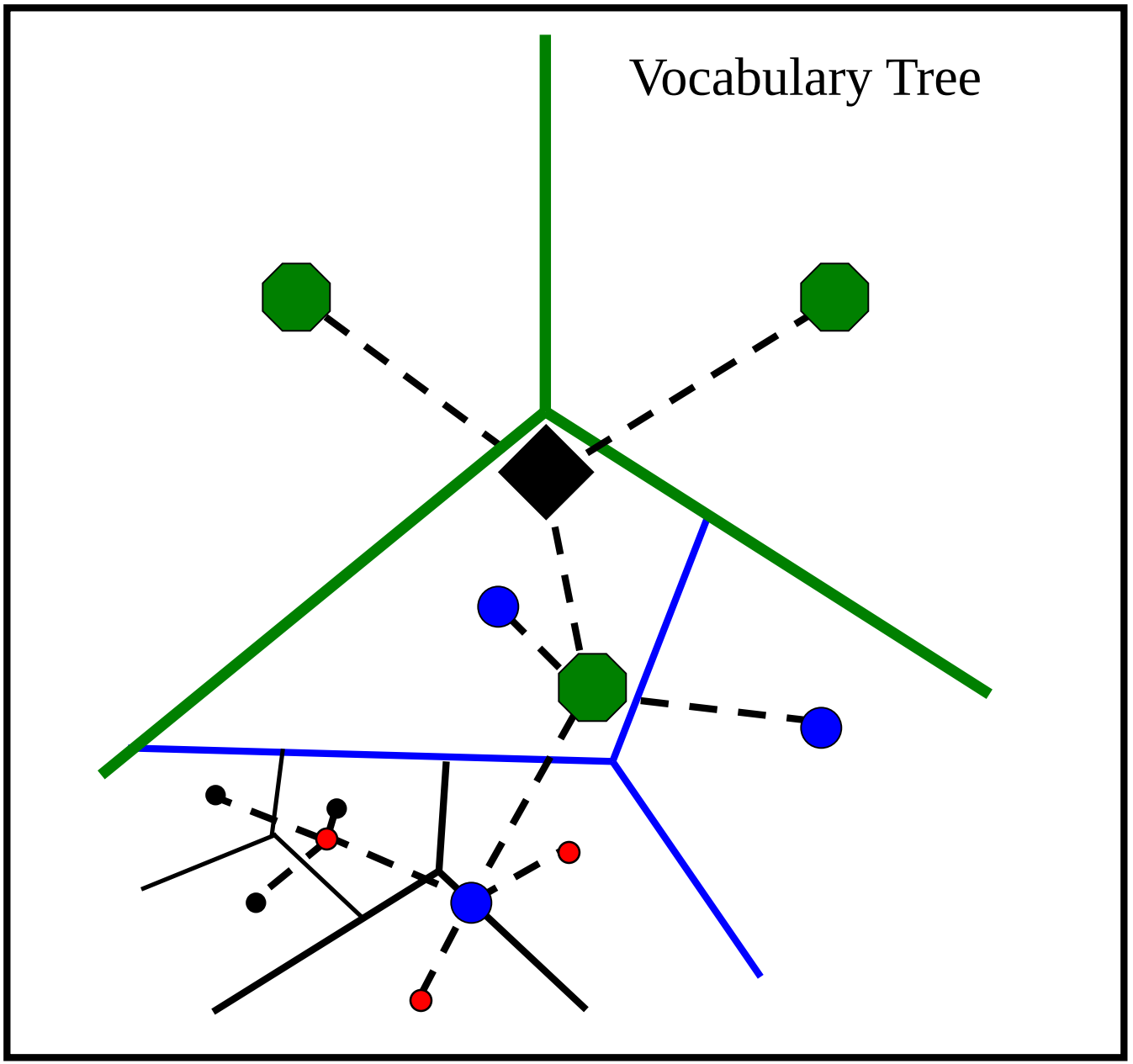
then recognition can scale to very large databases using the vocabulary tree and indexing approach described in Nistér & Stewénius CVPR 2006.



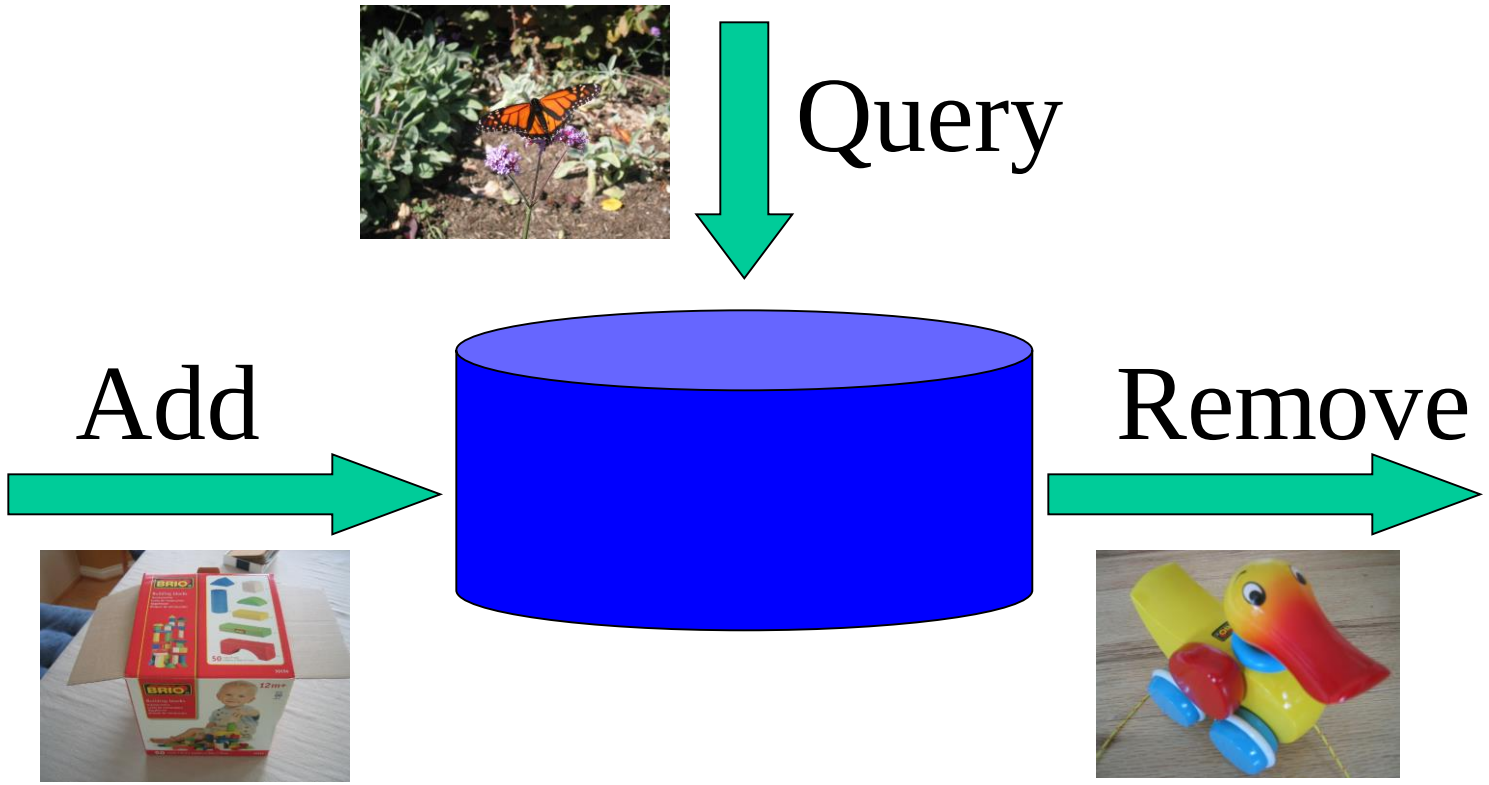

Vocabulary Tree


Vocabulary Tree


Vocabulary Tree

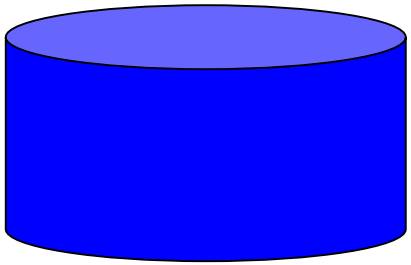

Vocabulary Tree

Adding, Querying and Removing Images at full speed

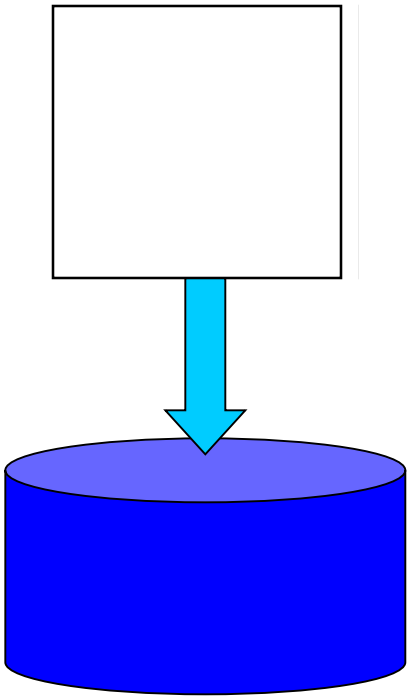


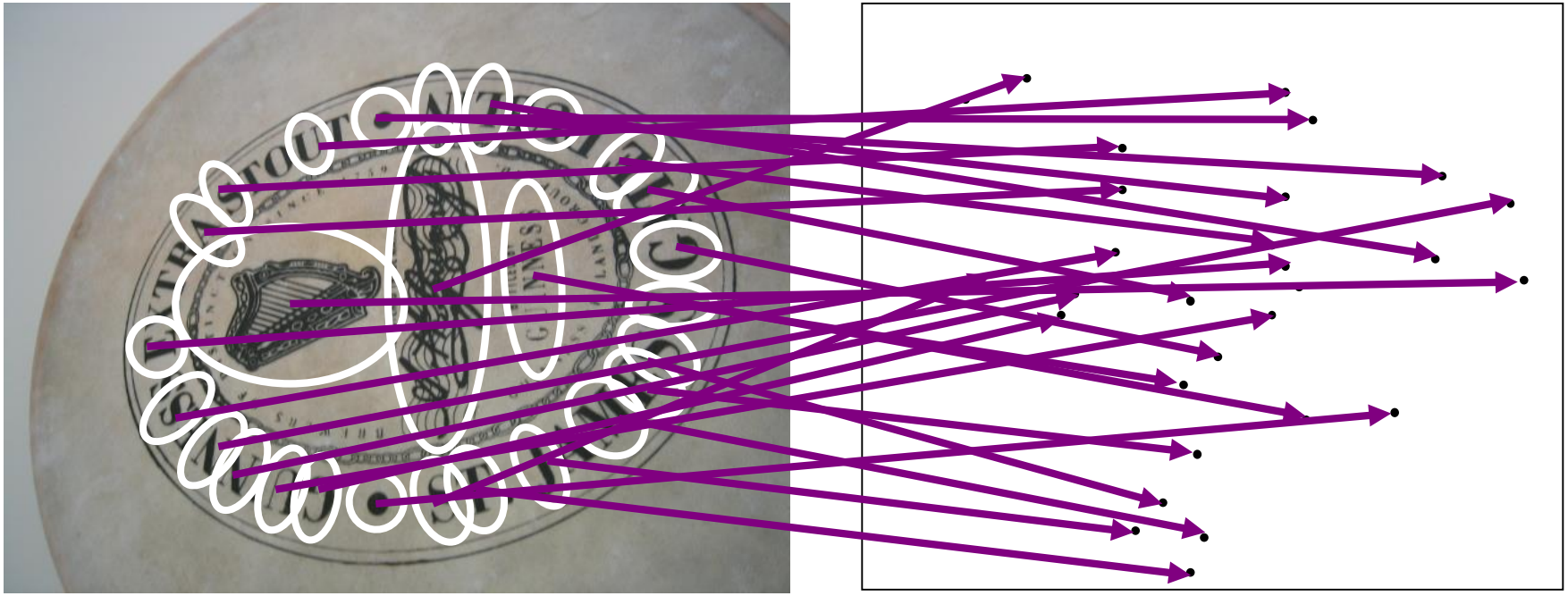
Training and Addition are Separate

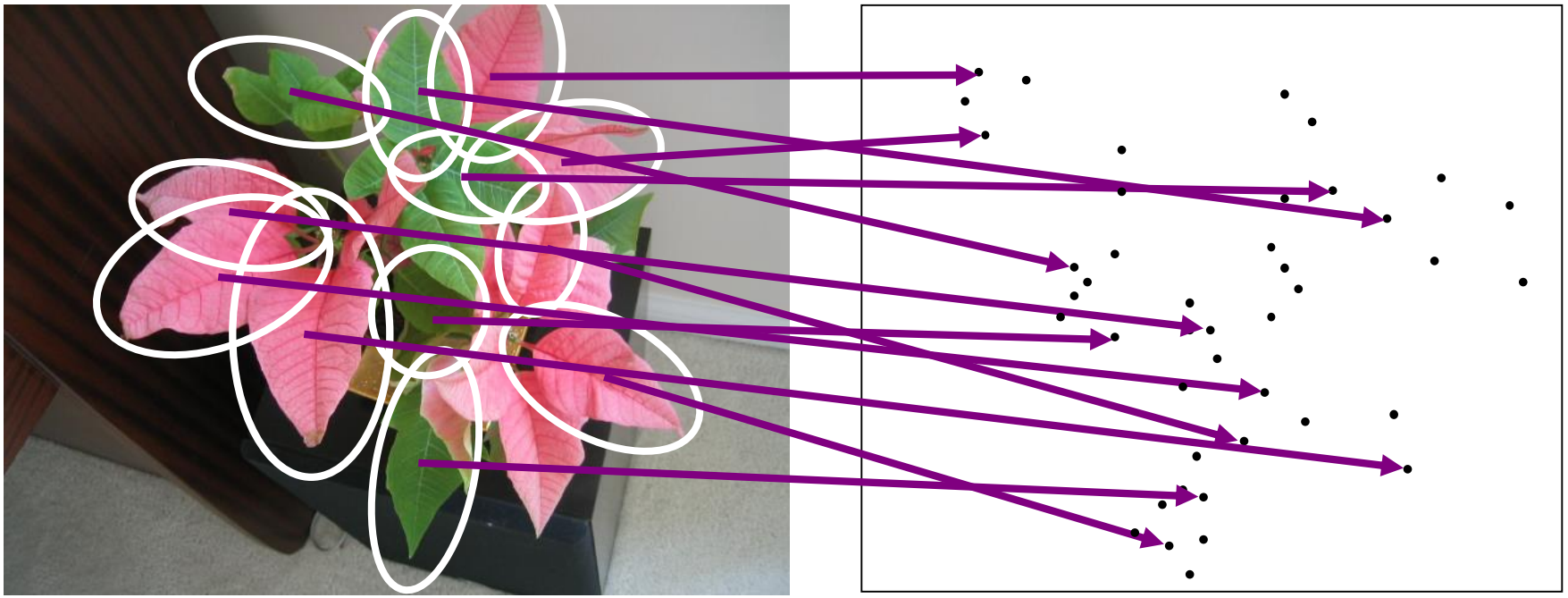
Common Approach

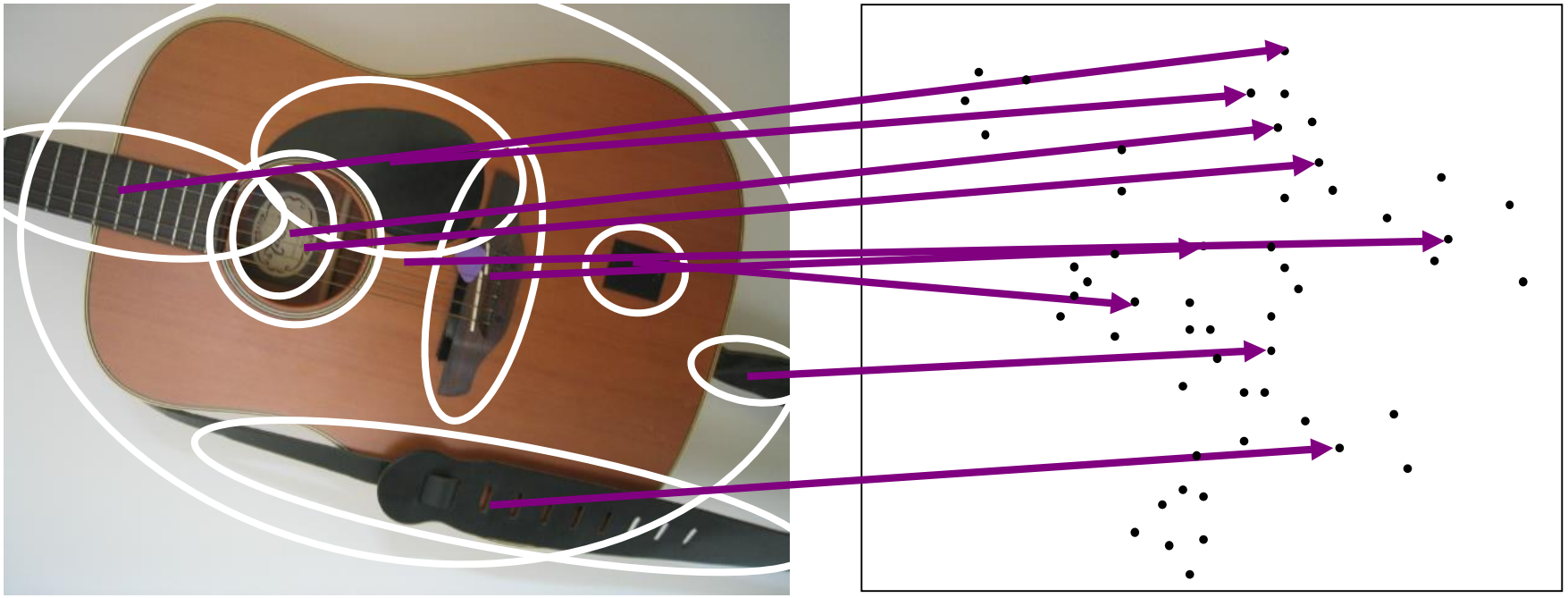


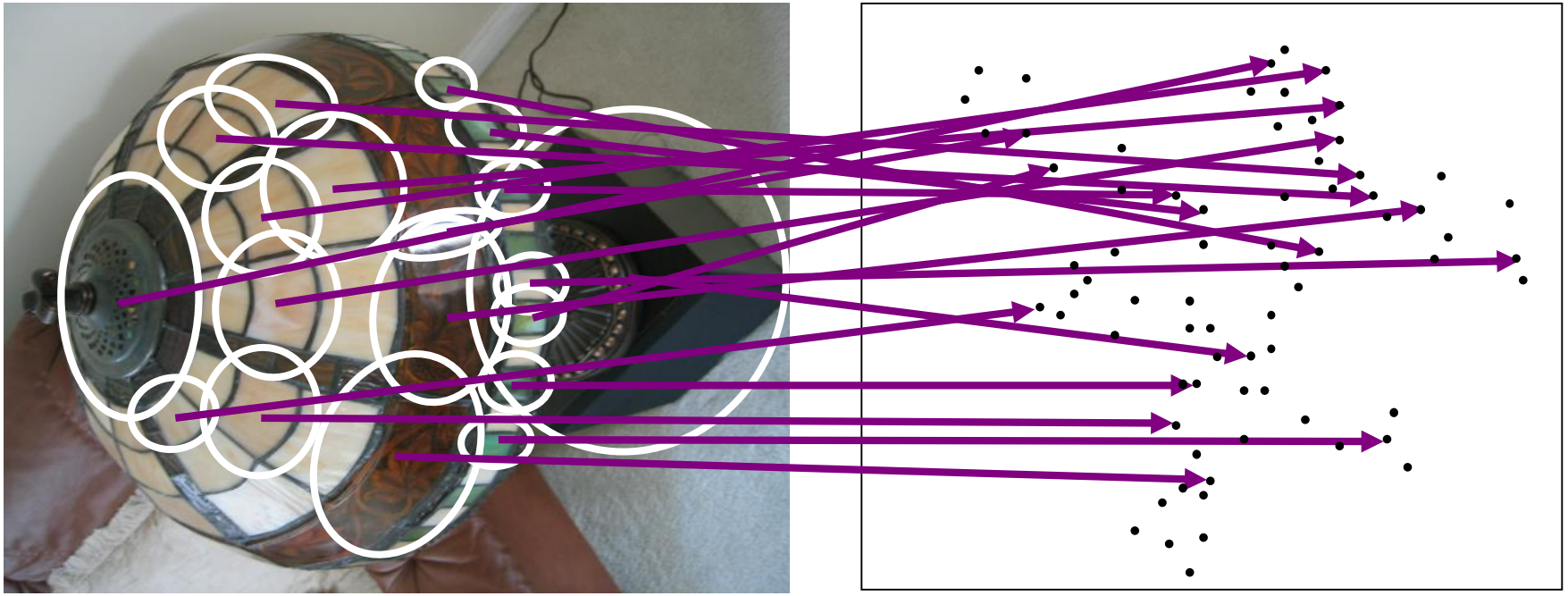
Our approach

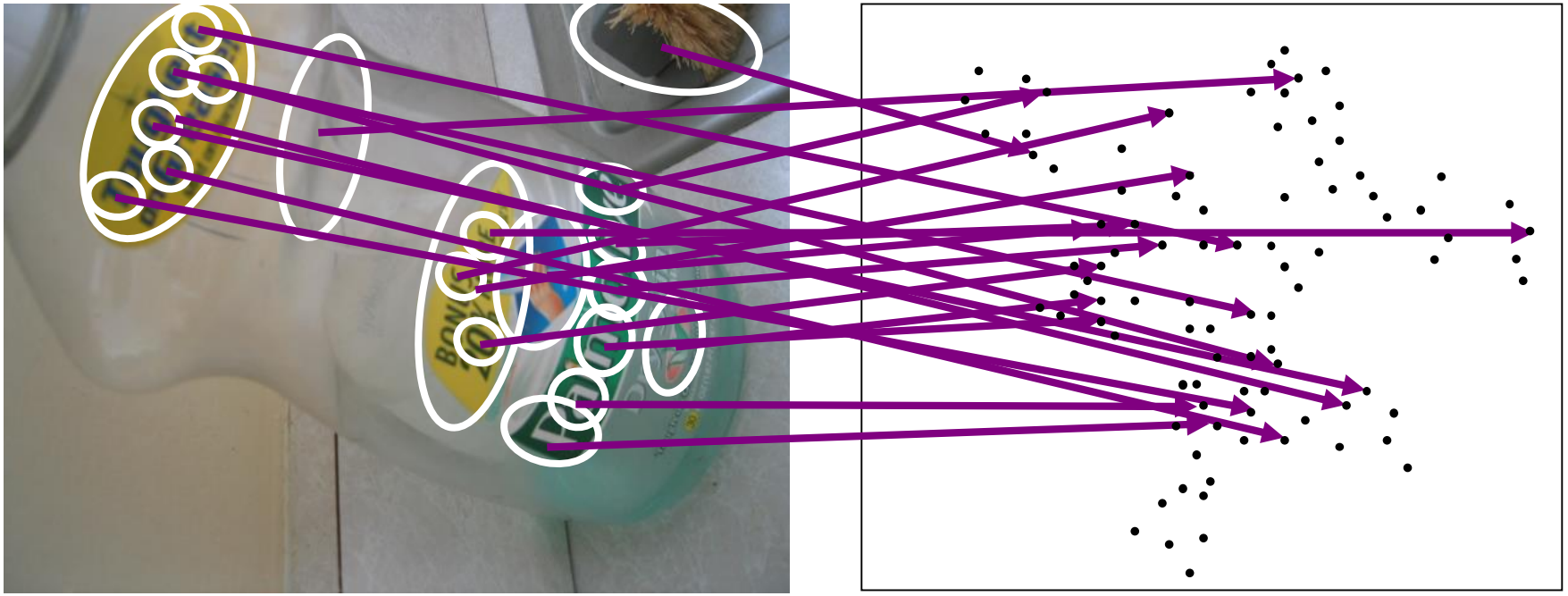


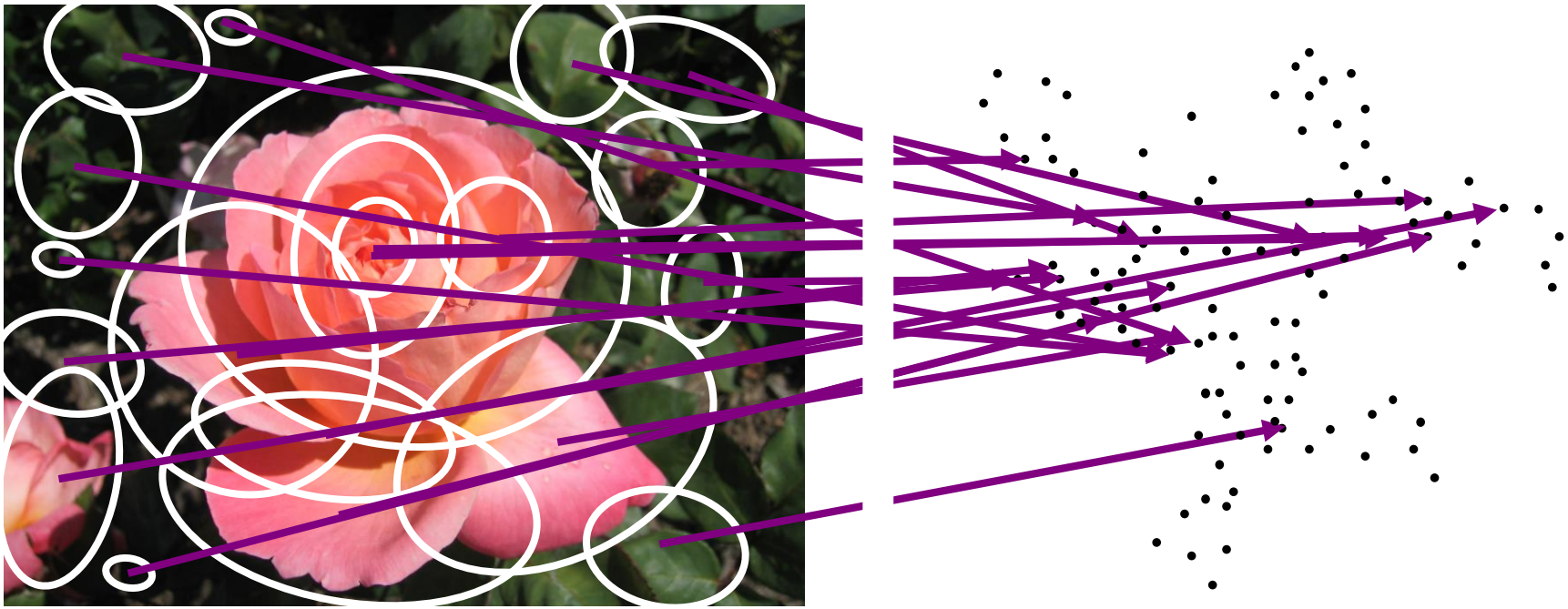


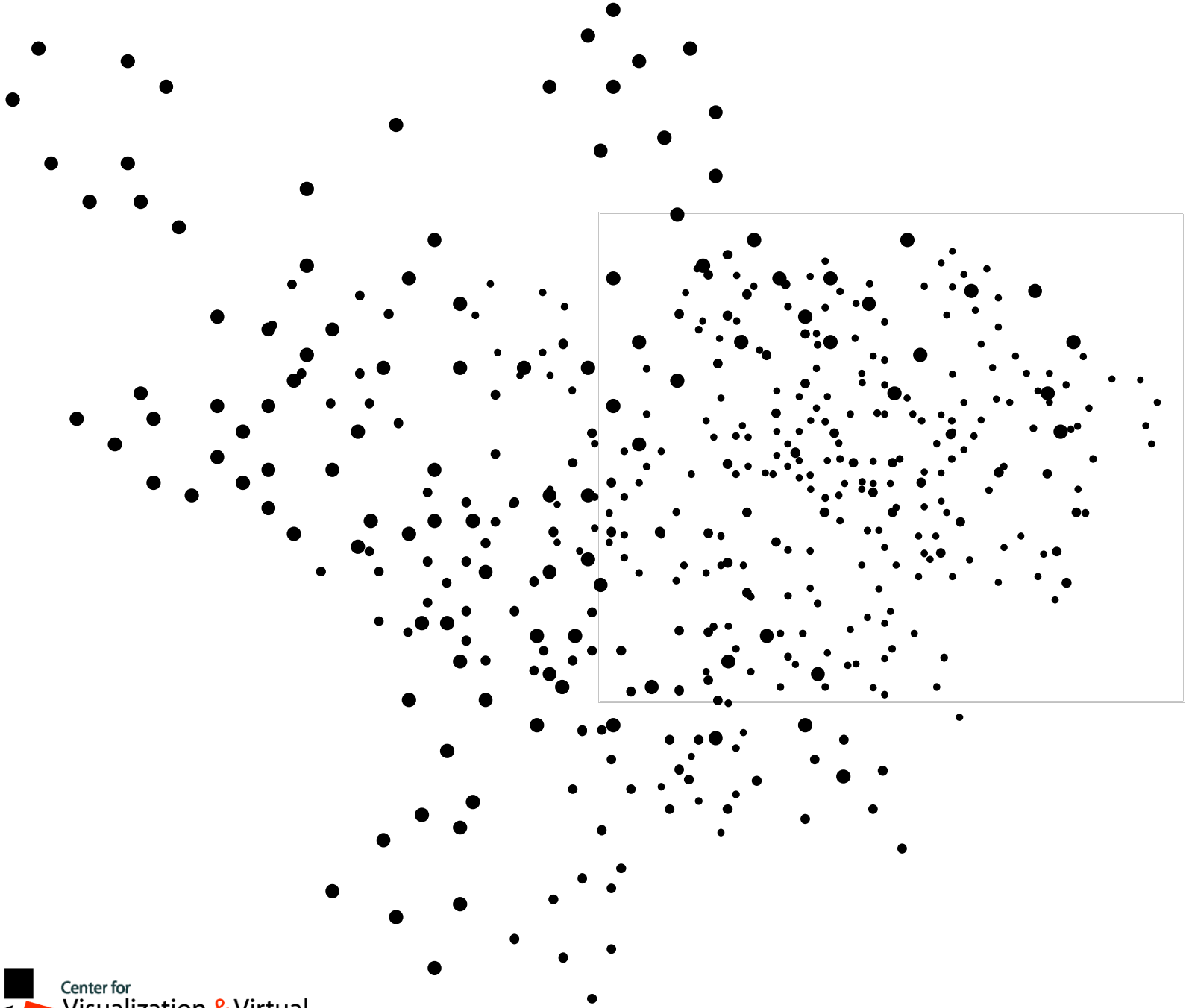


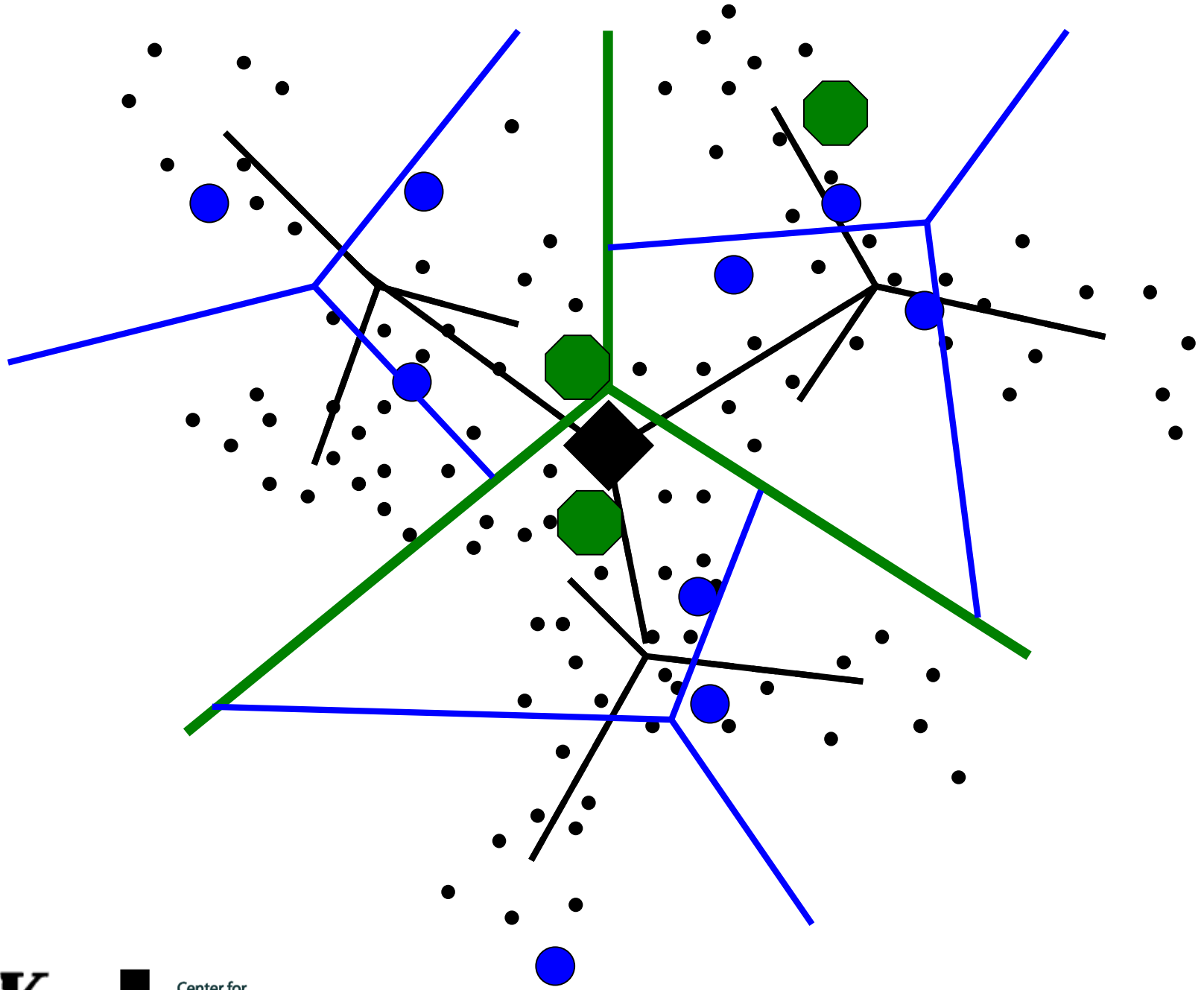


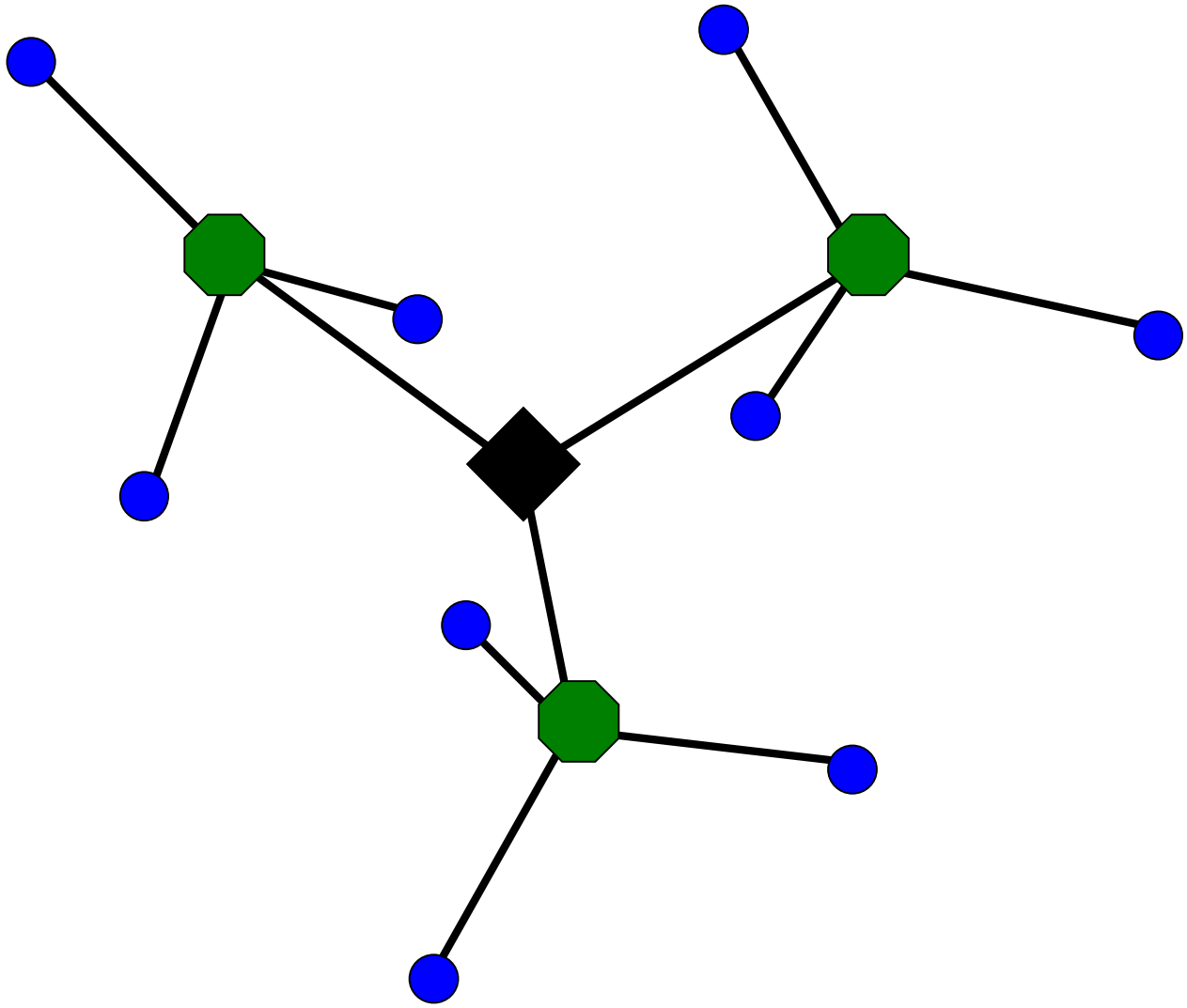


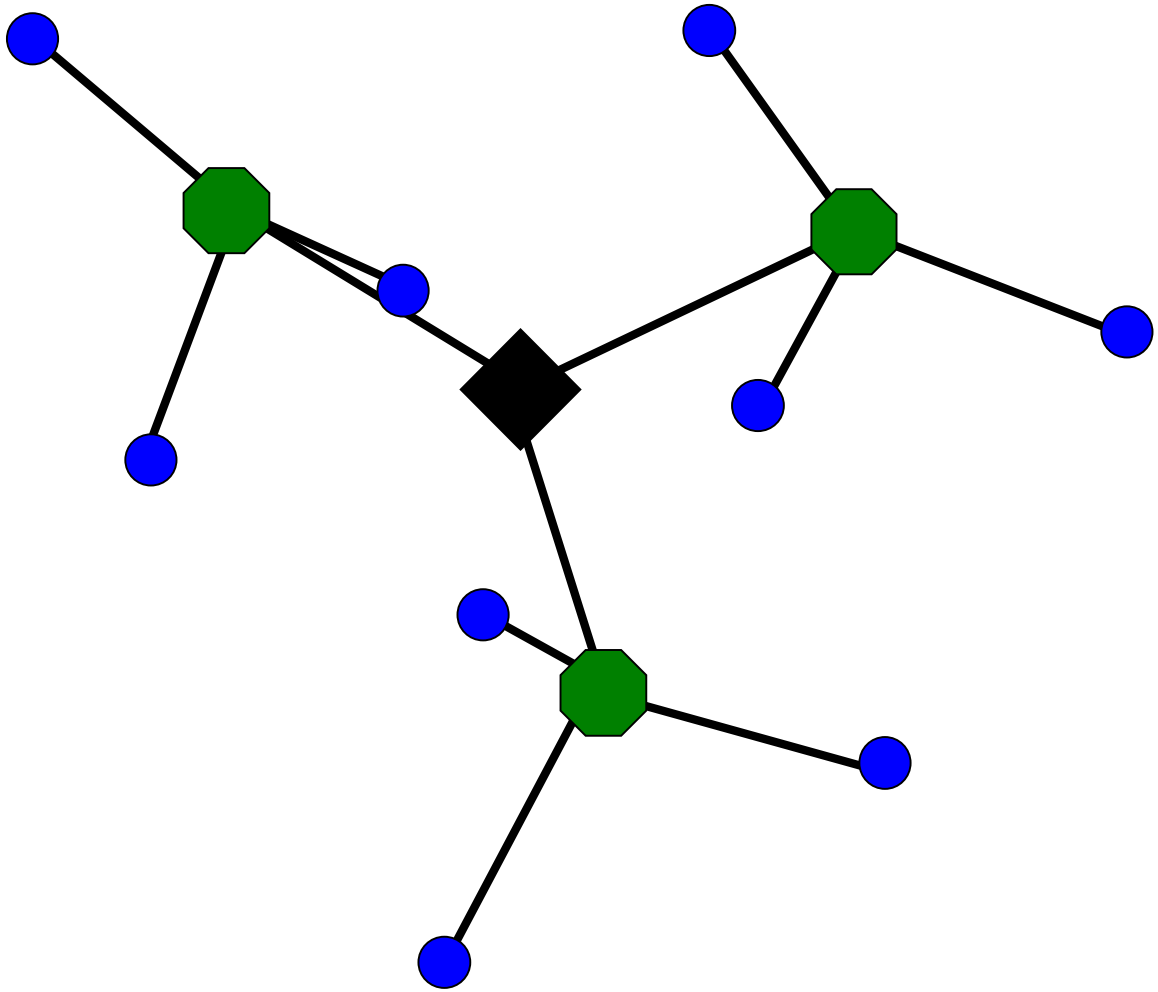


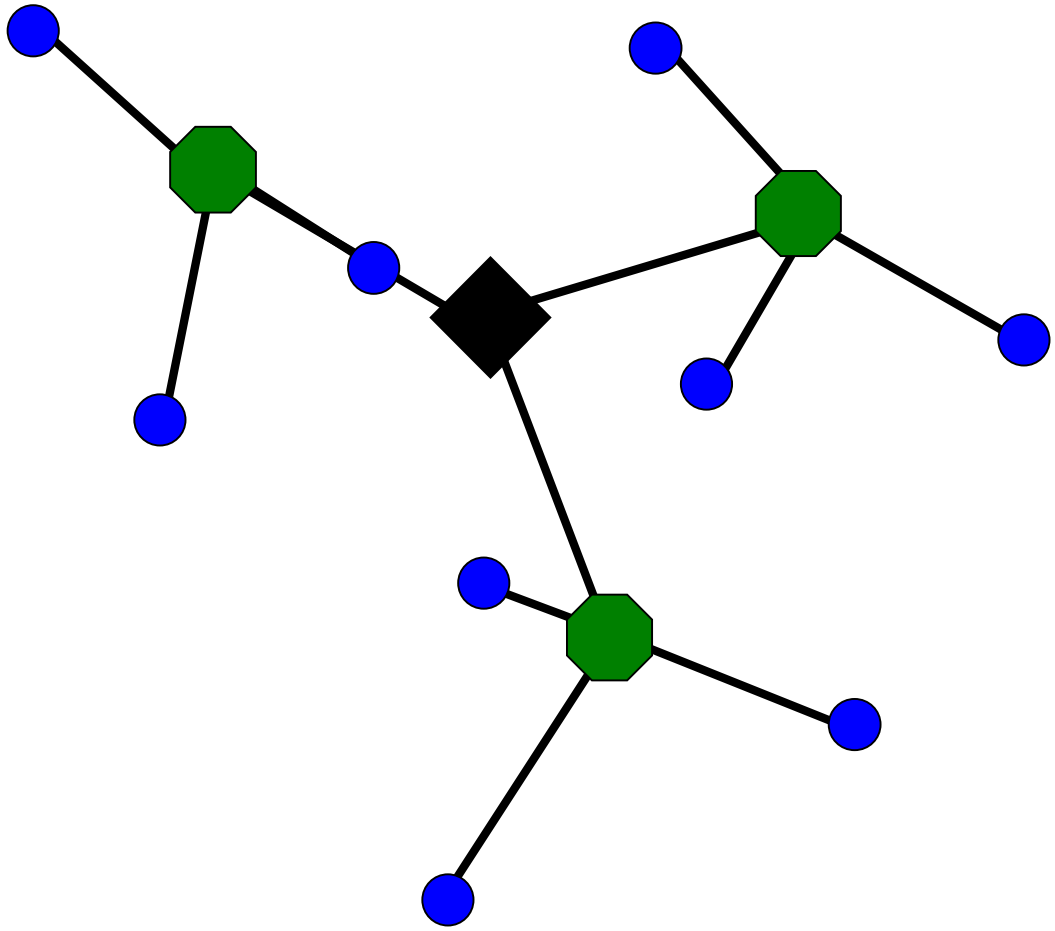


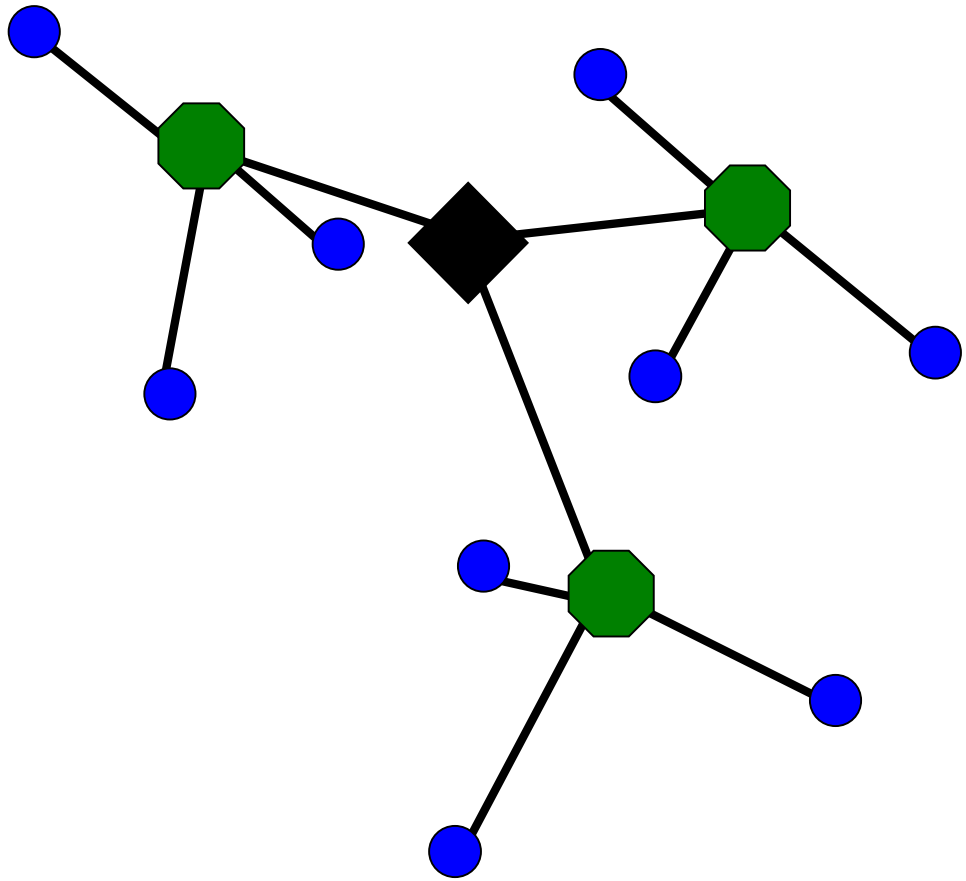


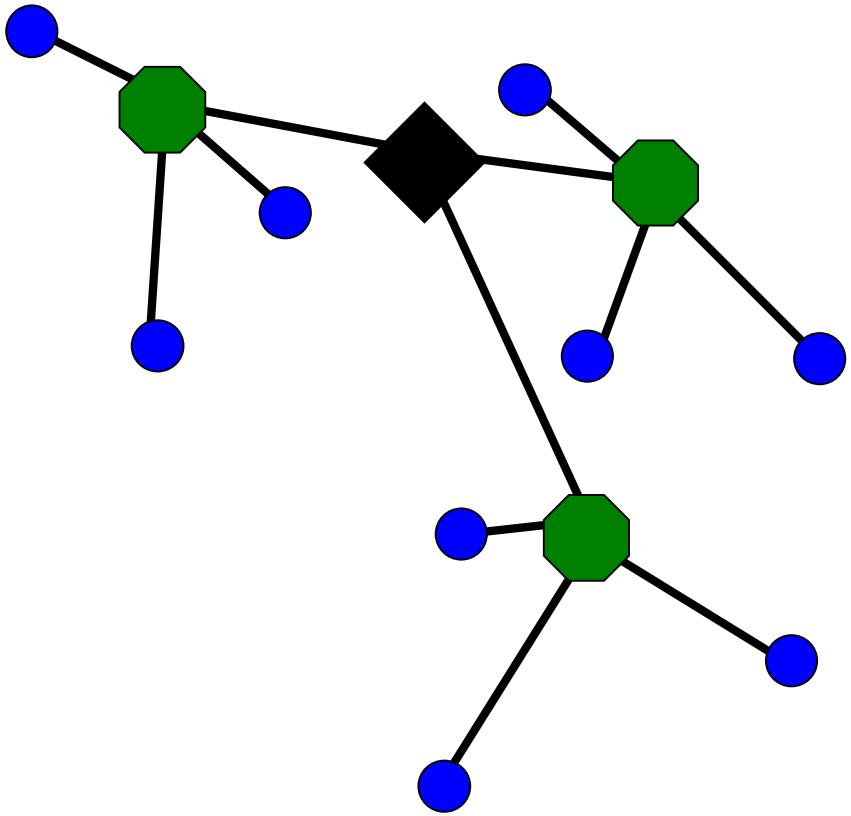


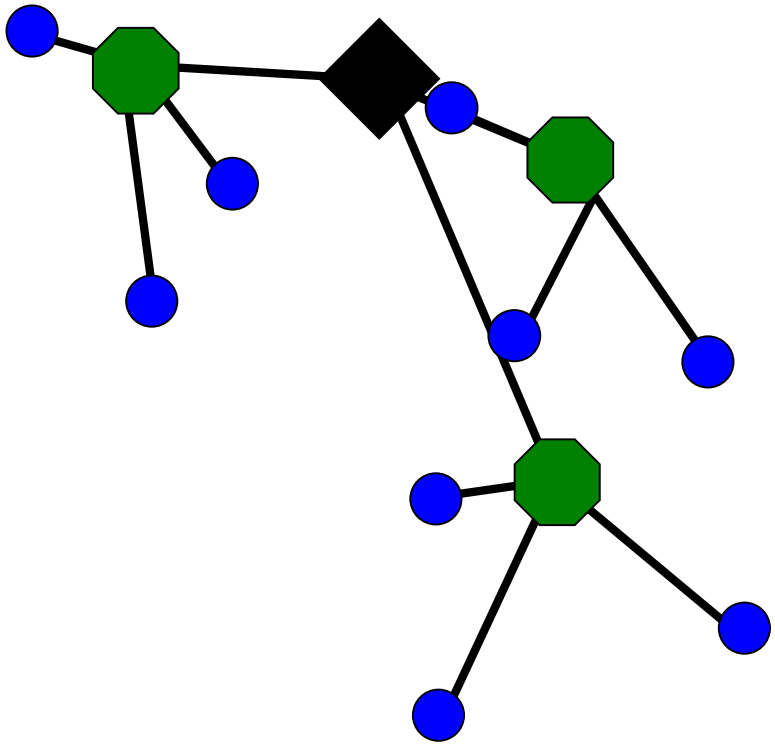


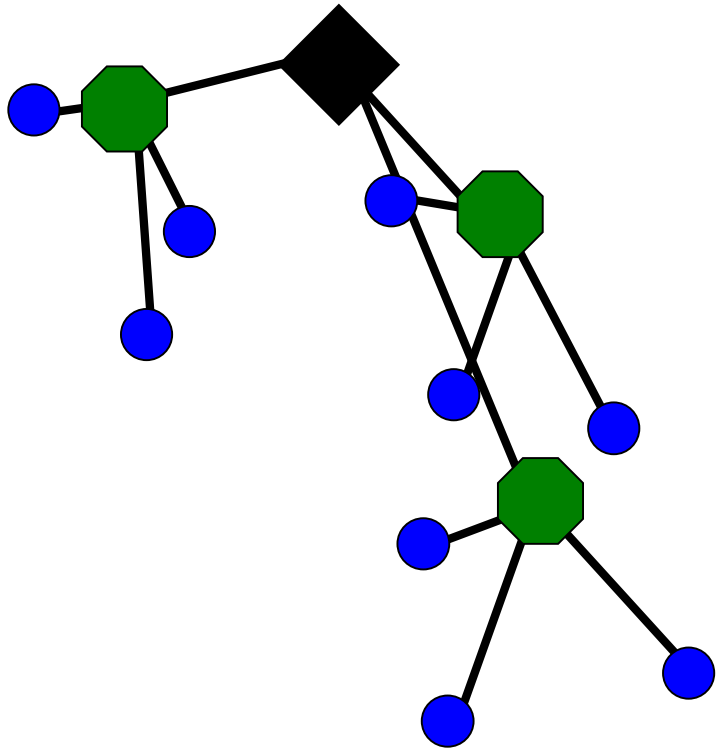


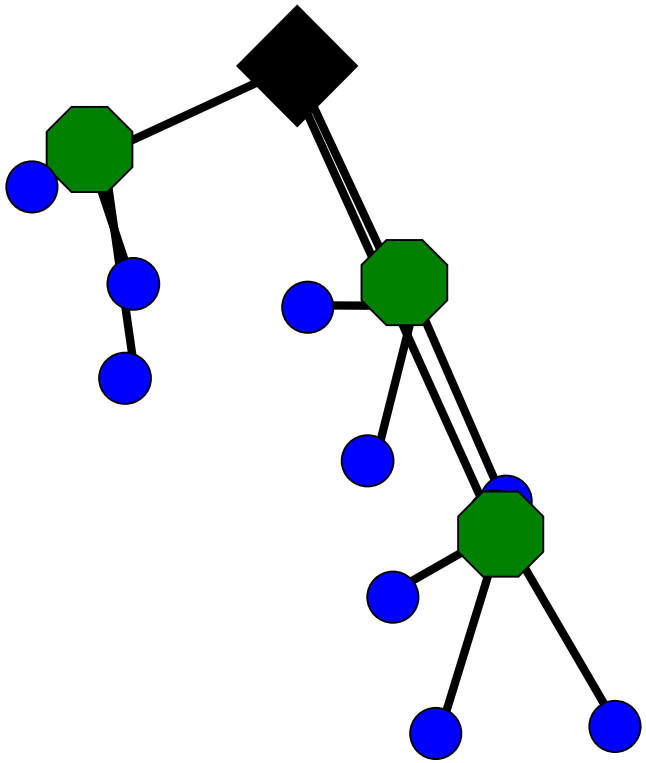


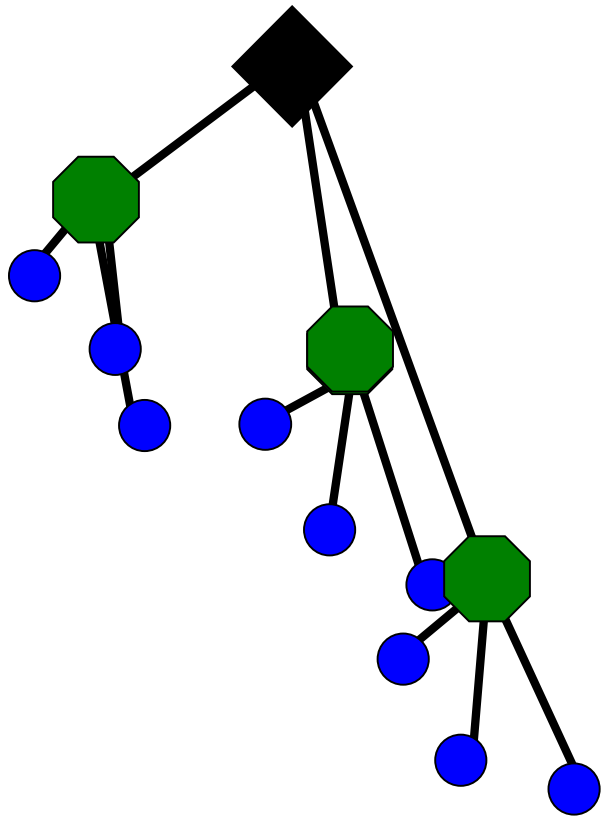


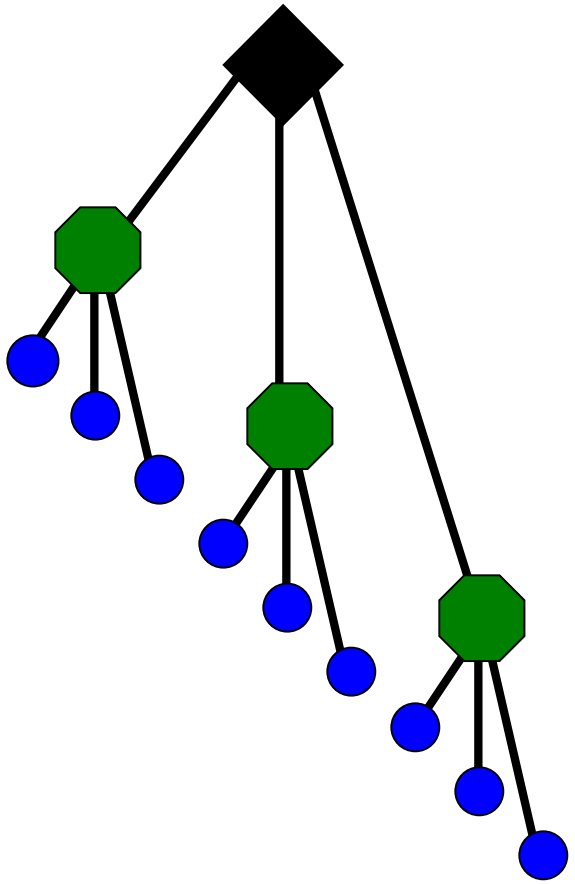


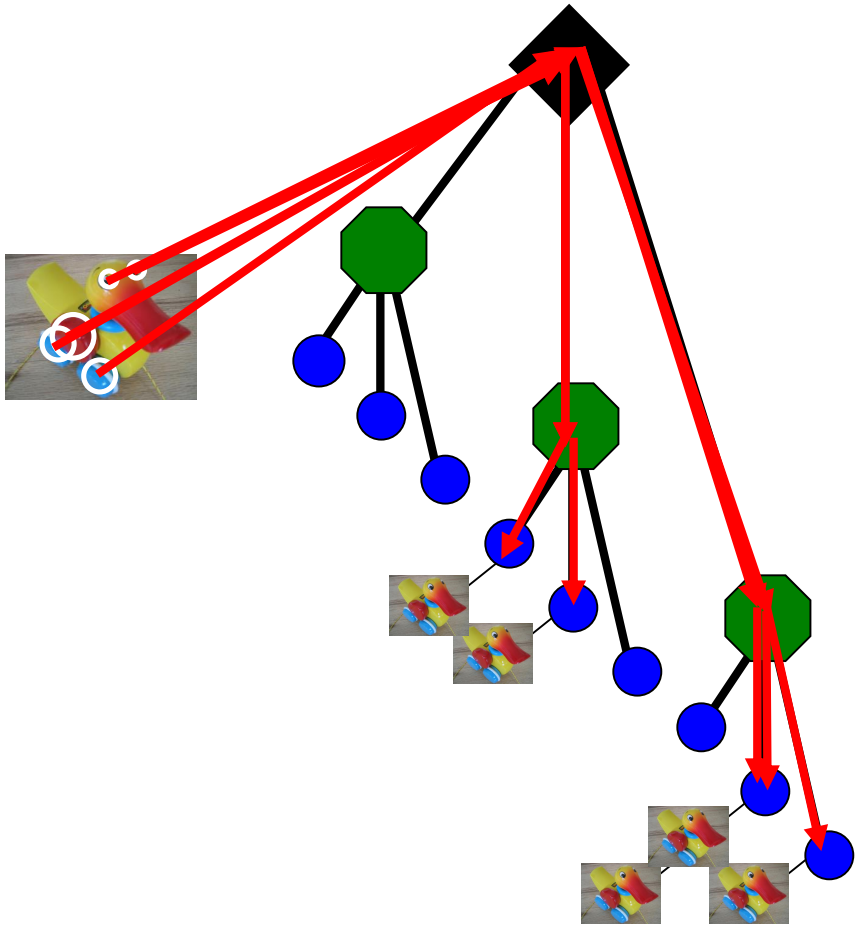


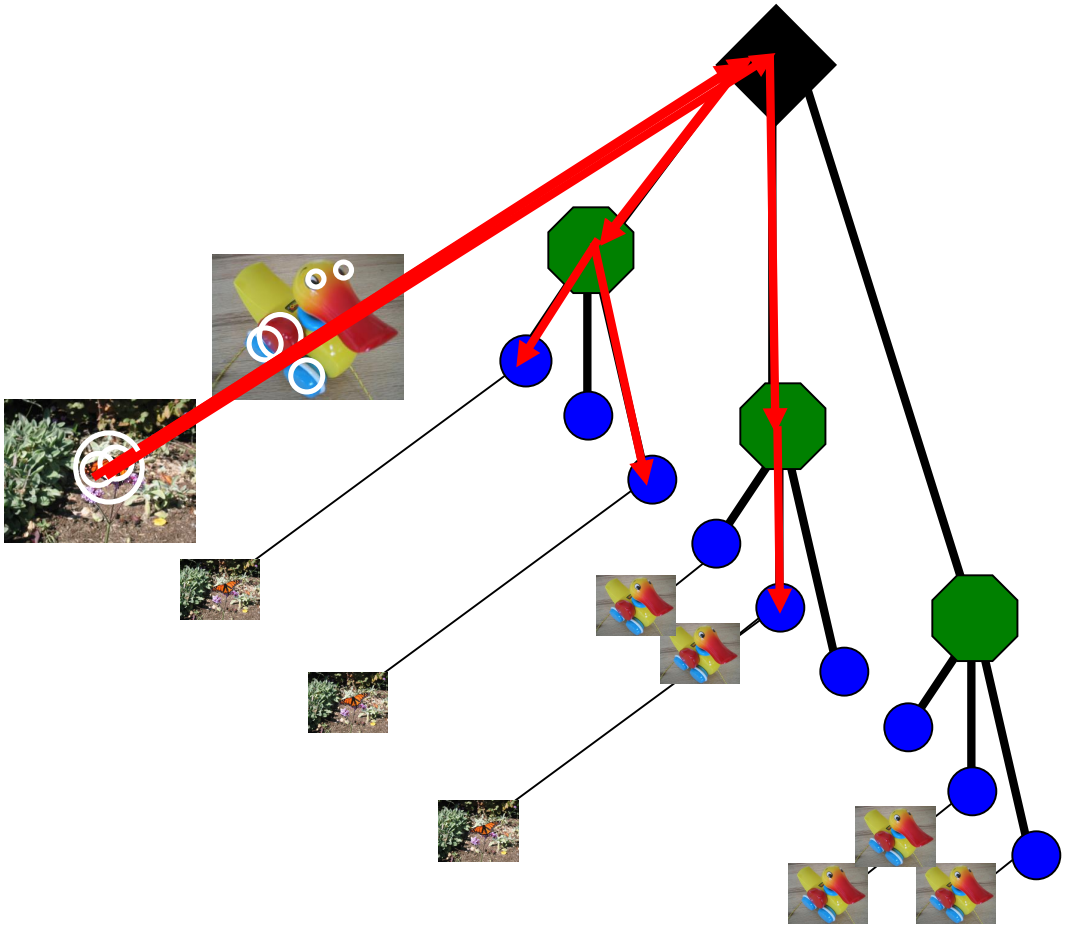


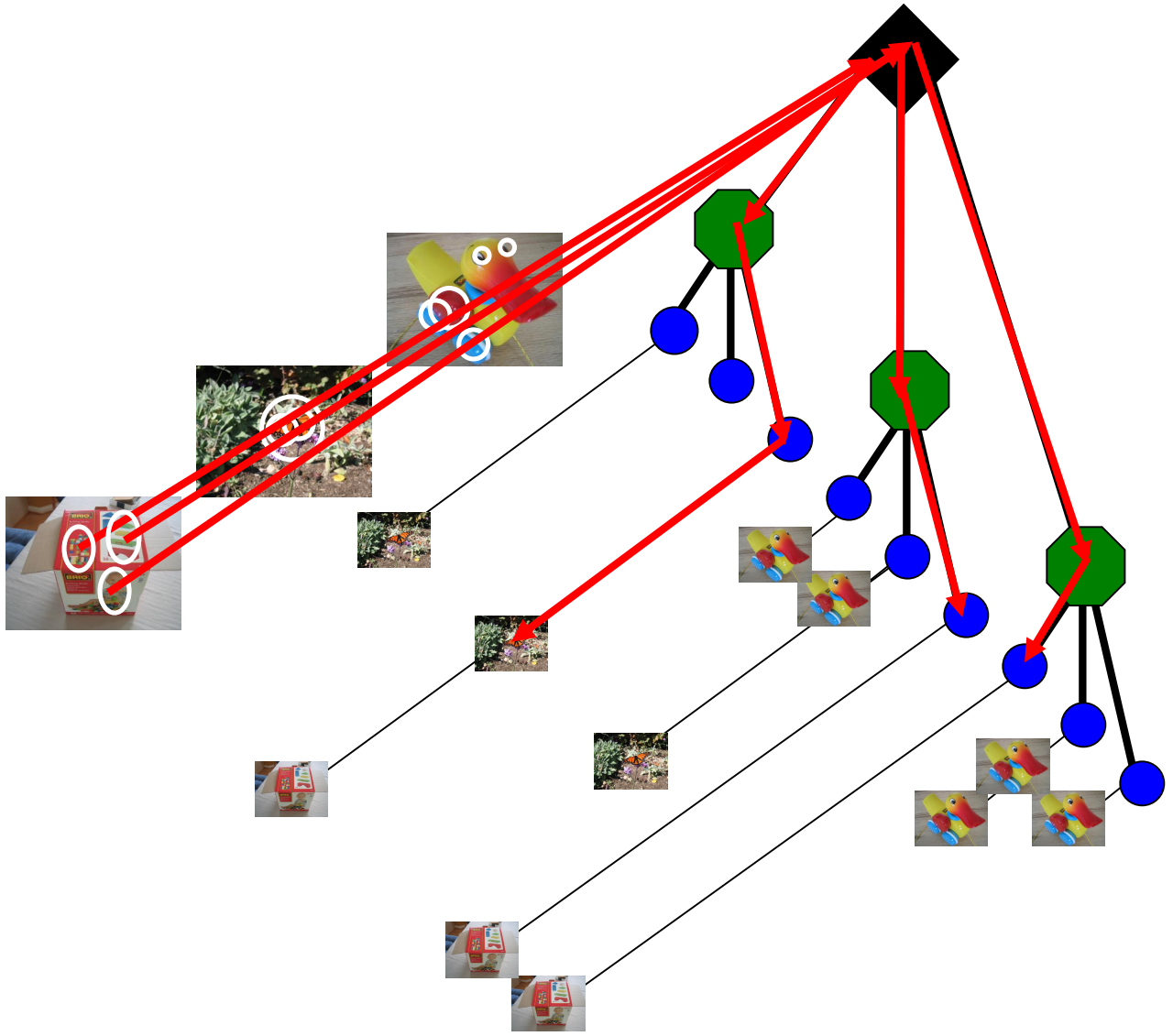


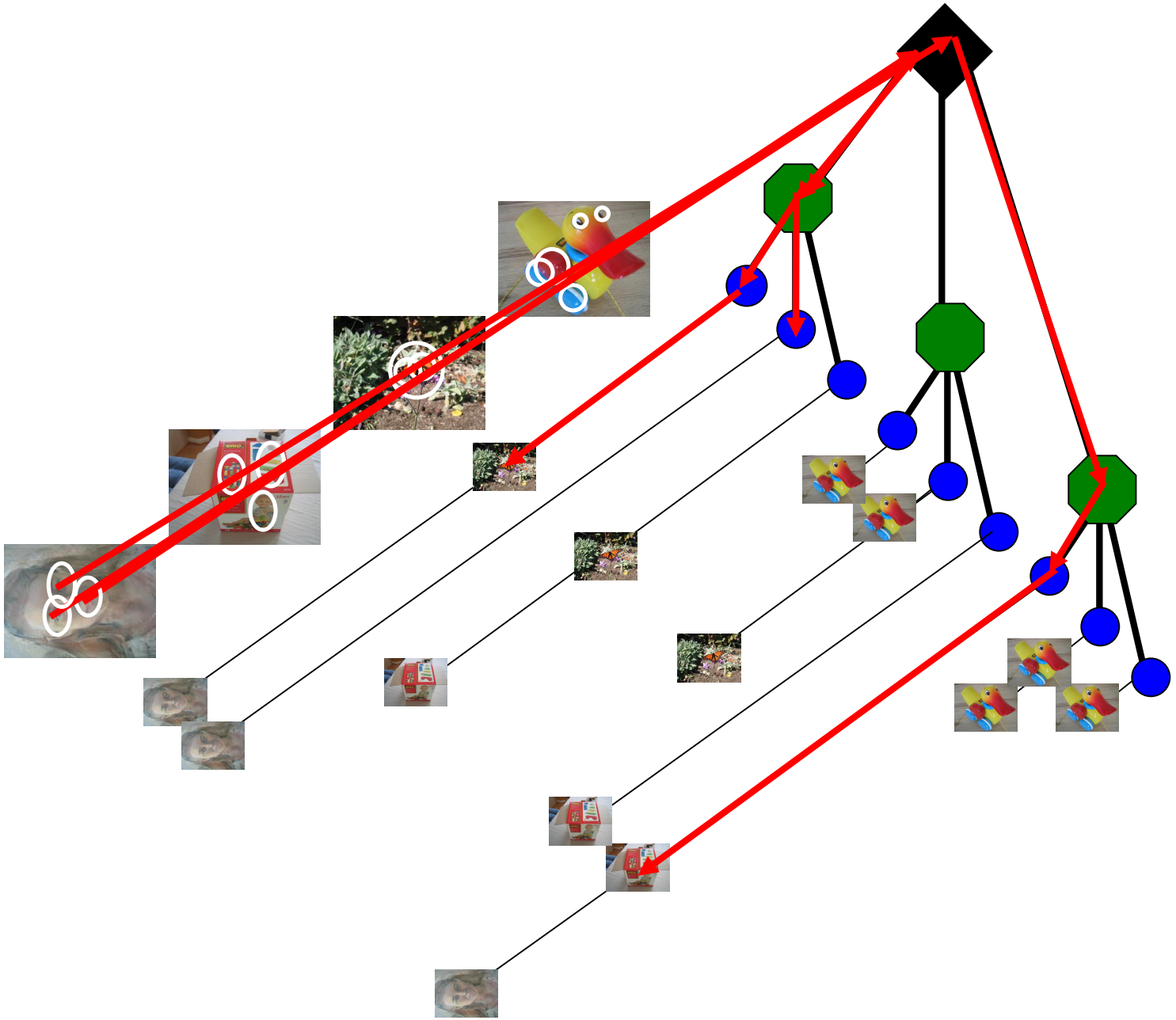


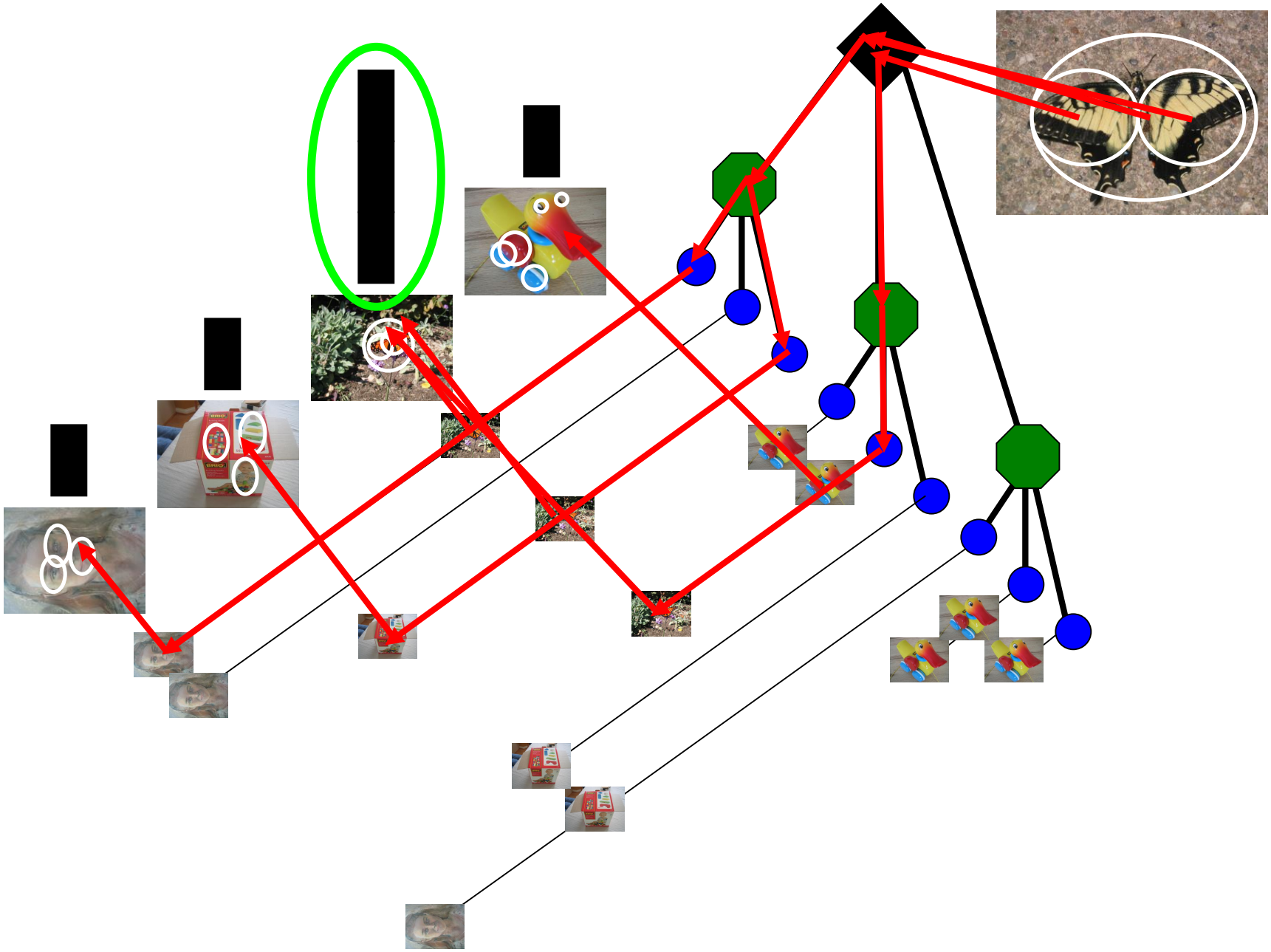




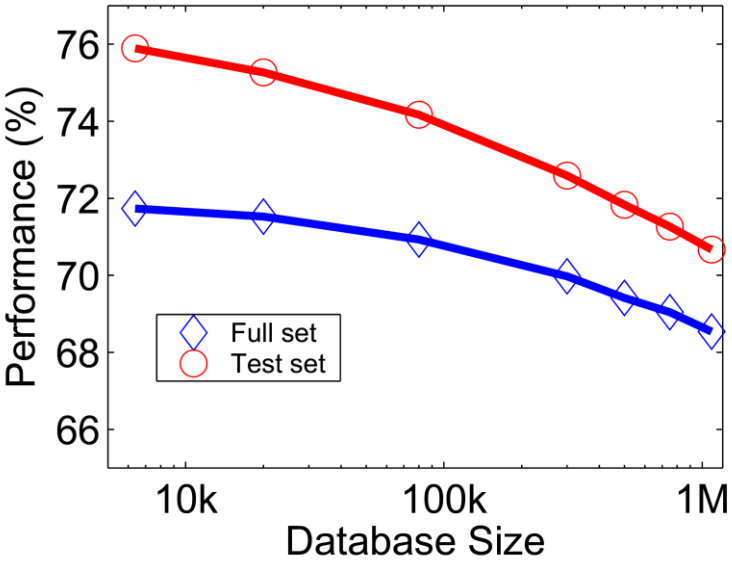








Performance



UK

Center for Visualization & Virtual Environments

Center for Visualization & Virtual Environments

Recognition Benchmark Images

Henrik Stewénius and David Nistér

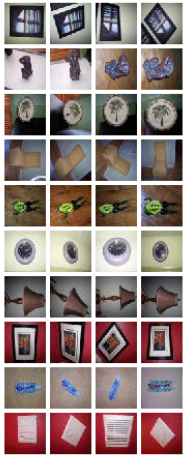
The set consists of 2604 groups of 4 images each for a total of 10416 images. All the images are 640x480.

If you use the dataset, please refer to:
- D. Nistér and H. Stewénius, Scalable Recognition with a Vocabulary Tree, CVPR 2006. PDF

Subsets

For users of subsets of the database please note that the difficulty is dependent on the chosen subset. Important factors are:

1. Difficulty of the objects themselves. CD-covers are much easier than flowers. See performance curve below.
2. Sharpness of the images. Many of the indoor images are somewhat blurry and this can affect some algorithms.
3. Similar or identical objects. All the pictures where taken by CS students/faculty/staff and thus keyboards and computer equipment are popular motives. So is computer vision literature.



Download

Please note BEFORE starting your download that the file is almost **2GB**. Please save a local copy in order to save bandwidth at our server.
- Zipped File.

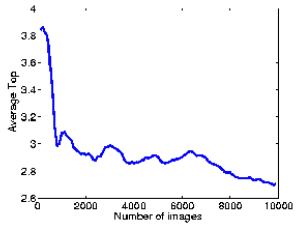
Performance

In the paper we give results either for a subset of 6376 images (all we had at that time) or a smaller subset of 1400 images. The smaller set was used when we did not have an efficient enough implementation in order to handle the larger set.

Performance Measures

- Our simplest measure of performance is to count how many of the 4 images which are top-4 when using a query image from that set of four images.

A matlab implementation which computes this measure: Download.

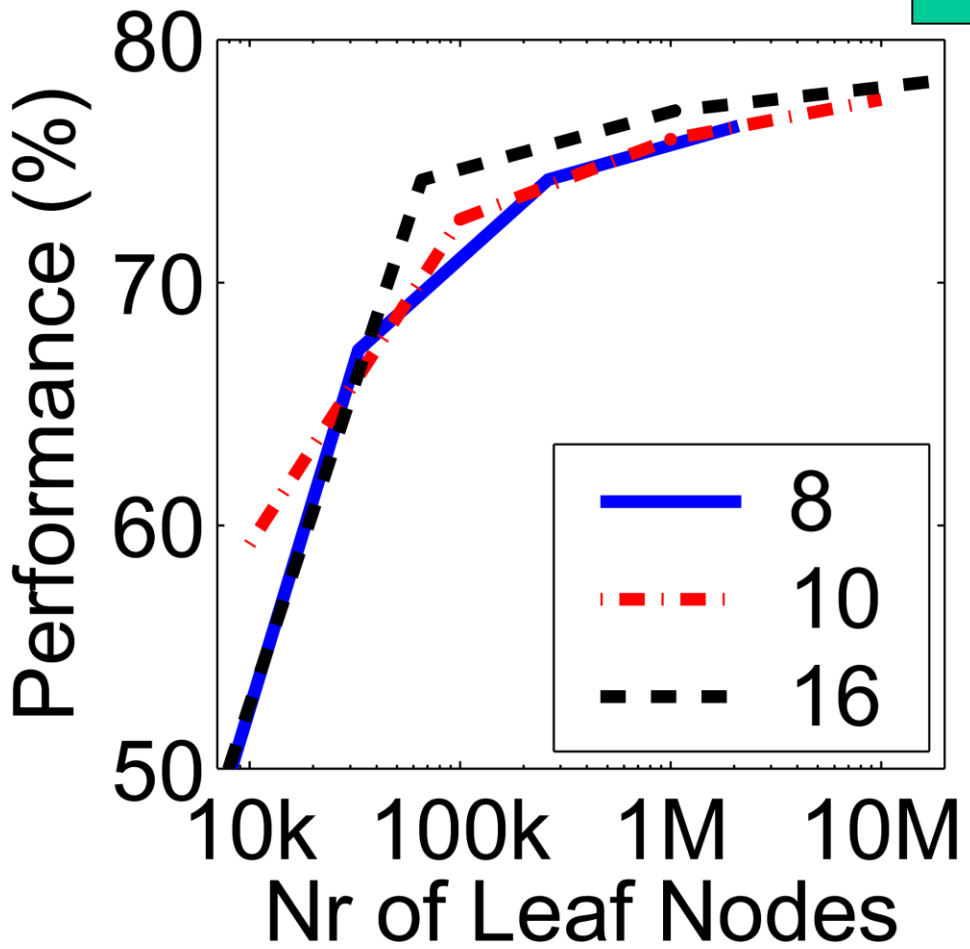


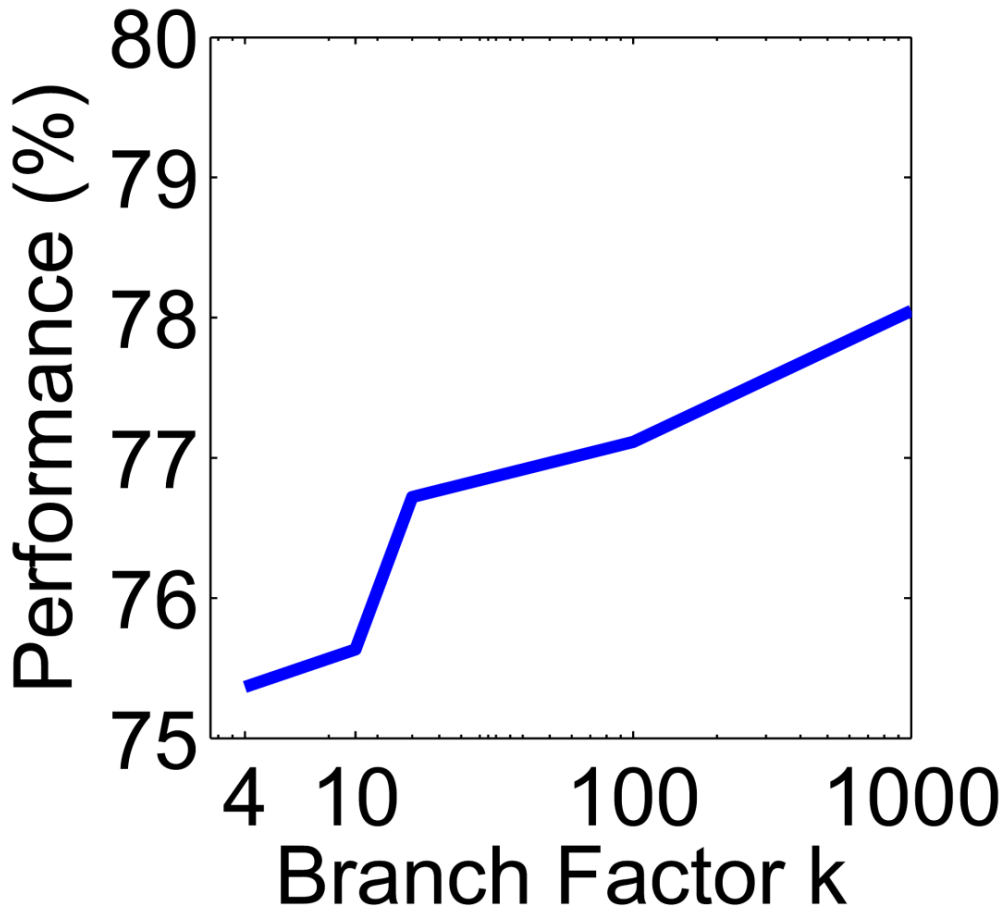
How our performance varies when taking subsets 0:n from the set. These results were run with settings optimized for speed.

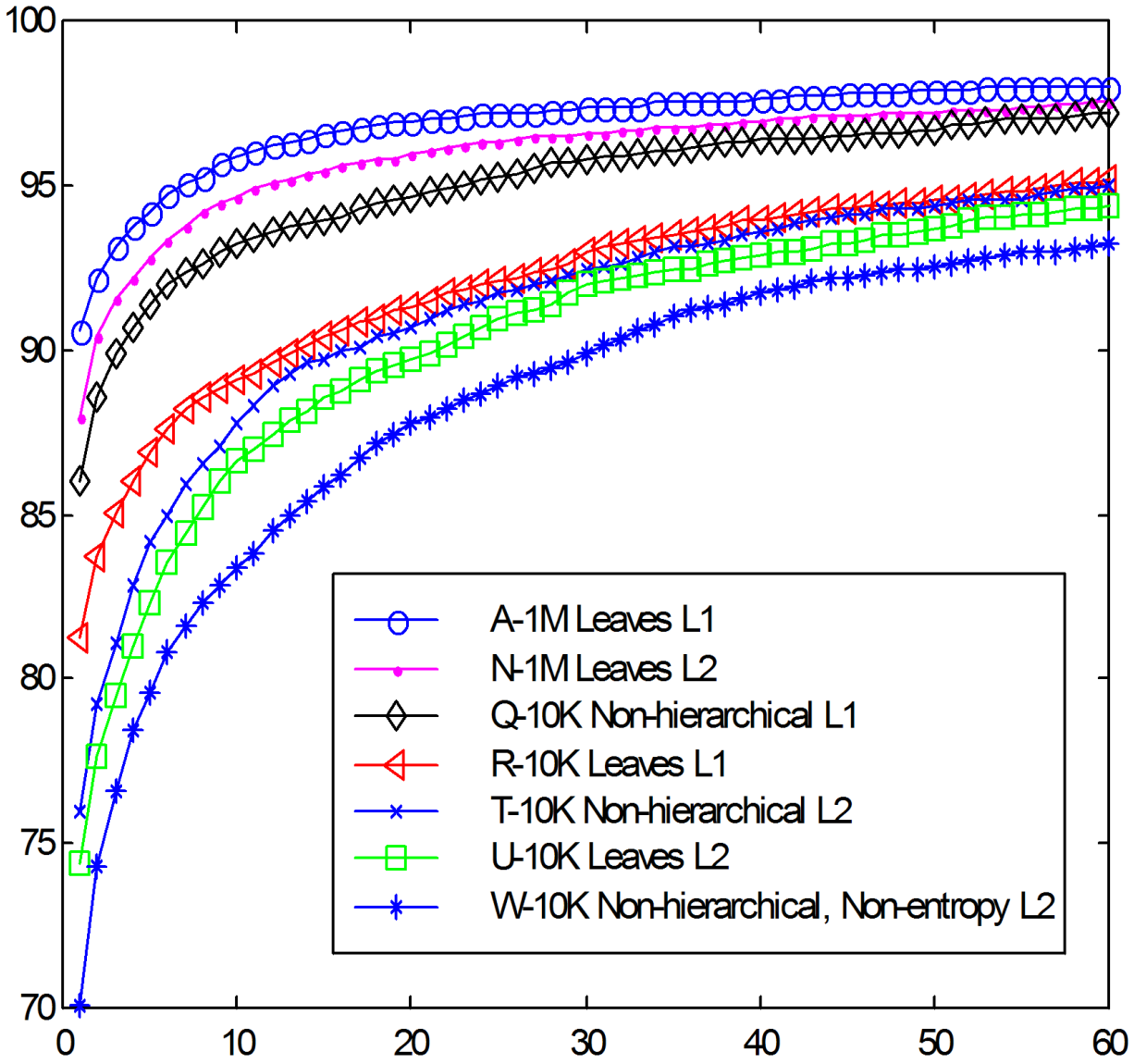
Size Matters

Improves Retrieval

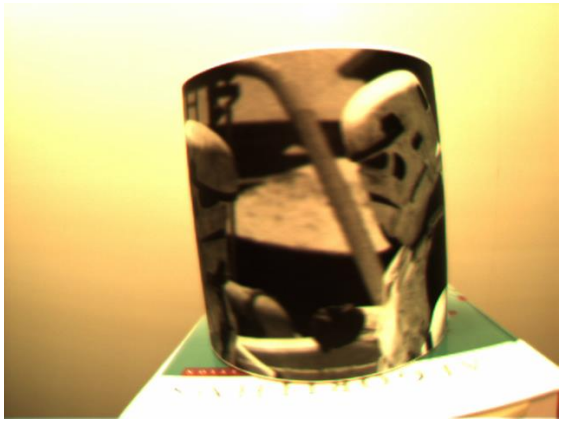
Improves Speed

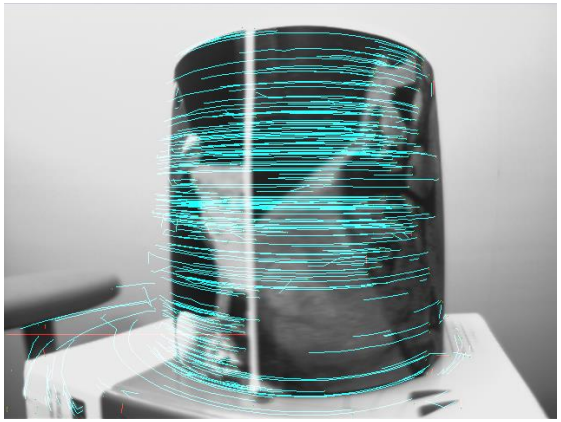


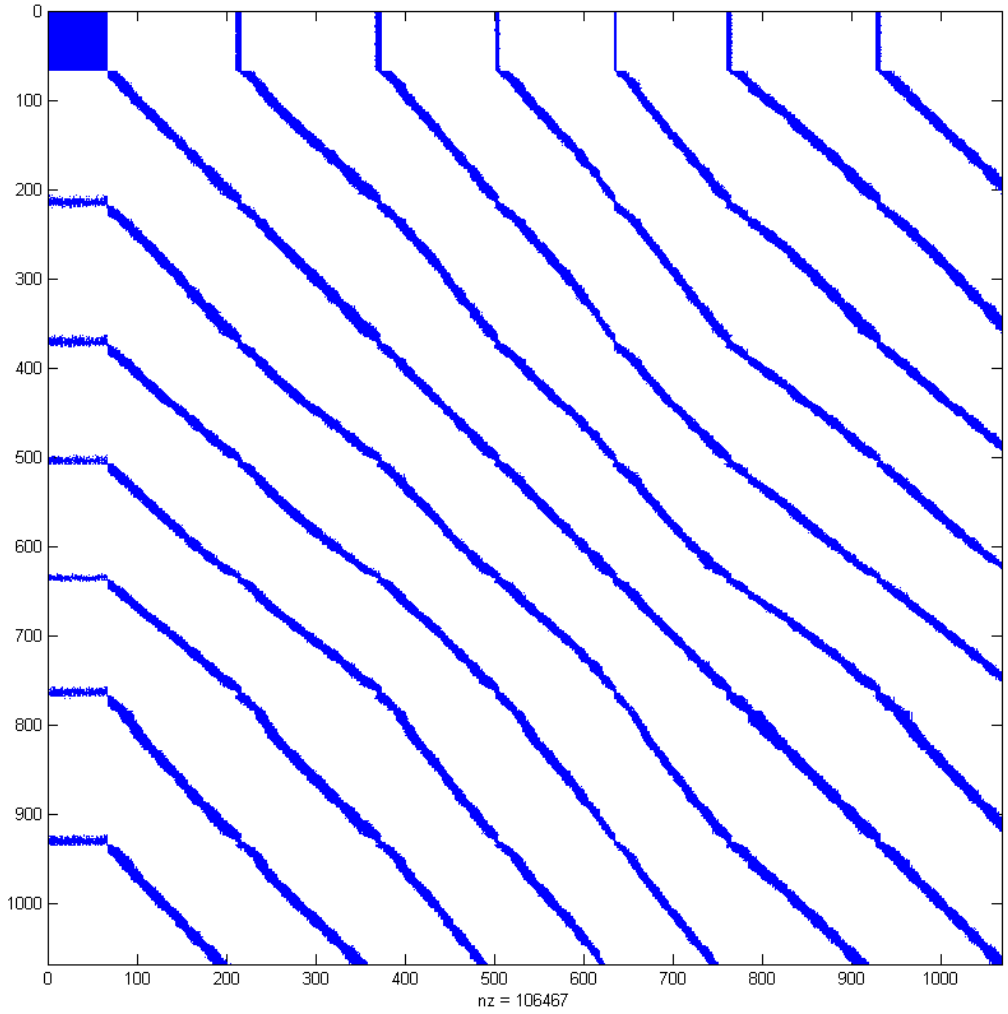

Performance (%)
80
79
78
77
76
75
4
10
100
1000
Branch Factor k

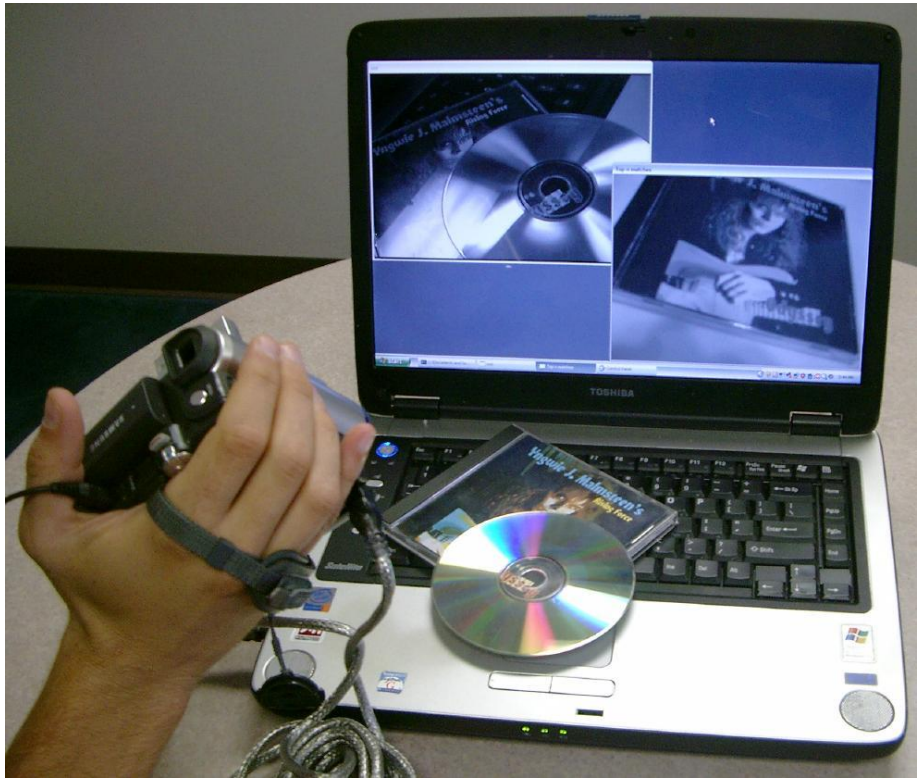

100
95
90
85
80
75
70
0
10
20
30
40
50
60
A-1M Leaves L1
N-1M Leaves L2
Q-10K Non-hierarchical L1
R-10K Leaves L1
T-10K Non-hierarchical L2
U-10K Leaves L2
W-10K Non-hierarchical, Non-entropy L2

Geometric Verification







Robust to Clutter and Occlusion

- Local Regions
- Like Web-search