北京郵電大學
BEIJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS

BUPT-MCPRL at Trecvid2015 Instance Search Task

Wenhui Jiang (jiang1st@bupt.edu.cn)
Zhicheng Zhao, Fei Su, Mei Liu, Shanwei Zhao, Anni Cai

MCPR Lab

Beijing University of Posts and Telecommunications

Brief Overview

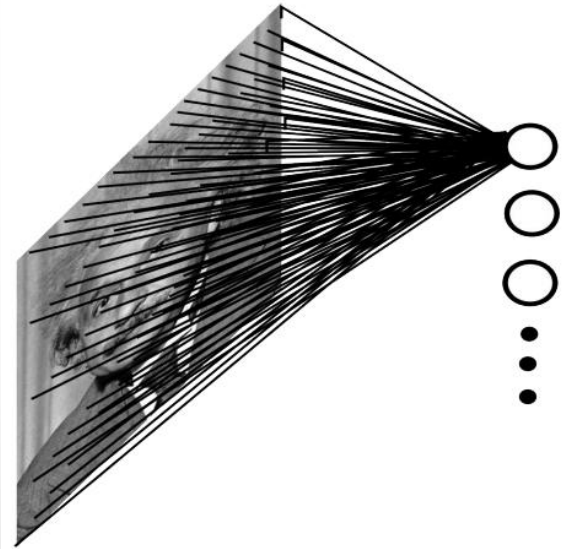
- Three local features
  - MSER + RootSIFT
  - Hessian Affine + RootSIFT
  - Deep Conv5
- One global feature
  - Deep FC6
- Feature fusion
  - Manual tuned
  - Query adaptive
- Trial feature
  - Hessian Affine + Deep Conv

Brief Overview

Features	mAP (2013)	mAP (2014)	mAP (2015)
MSER + RootSIFT	15.86	13.00	
Hessian Affine + RootSIFT	21.59	17.03	
Deep Conv5	16.58	18.37	
Deep Fc6	4.52	4.03	

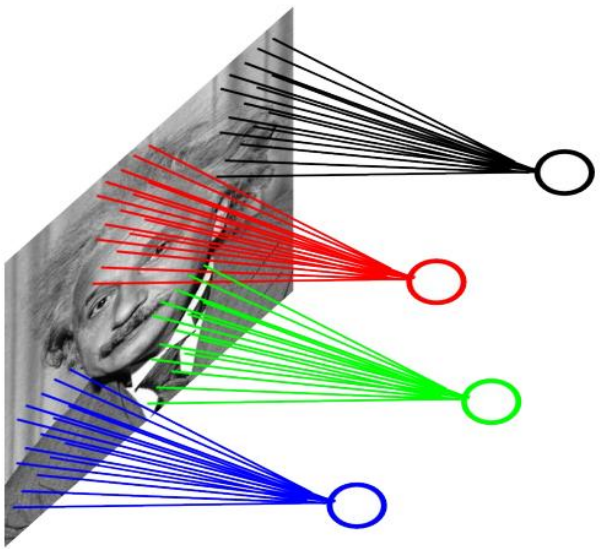
Deep Conv Feature

Fully connected layer



Deep FC

Locally connected layer



Deep Conv

Deep Conv Feature

Center point

Receptive field for conv1

Receptive field for conv5

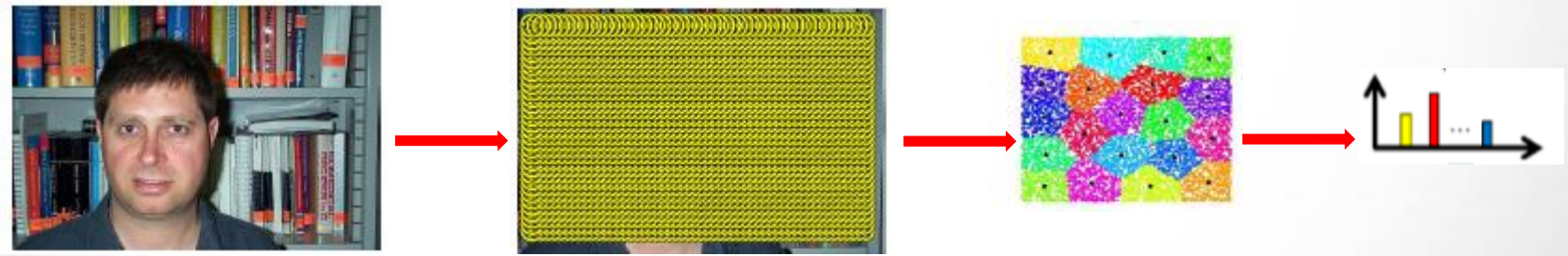
Receptive field sizes and strides for AlexNet

Layer	Rf size	Stride
Conv1	11 X 11	4 X 4
Conv2	51 X 51	8 X 8
Conv3	99 X 99	16 X 16
Conv4	131 X 131	16 X 16
Conv5	163 X 163	16 X 16
Pool5	195 X 195	32 X 32

Reference: Exploiting Local Features from Deep Networks for Image Retrieval, CVPR Workshop 2015

Deep Conv Feature

Feature representation workflow for Deep conv features



1. Input image

2. Dense sampling conv5 activations

3. 1M codebook

4. BoW feature (Deep Conv5)

Deep Conv Feature

Features	mAP (2013)	mAP (2014)
MSER + RootSIFT	15.86	13.00
Hessian Affine + RootSIFT	21.59	17.03
Deep Conv5	16.58	18.37
Deep Fc6	4.52	4.03

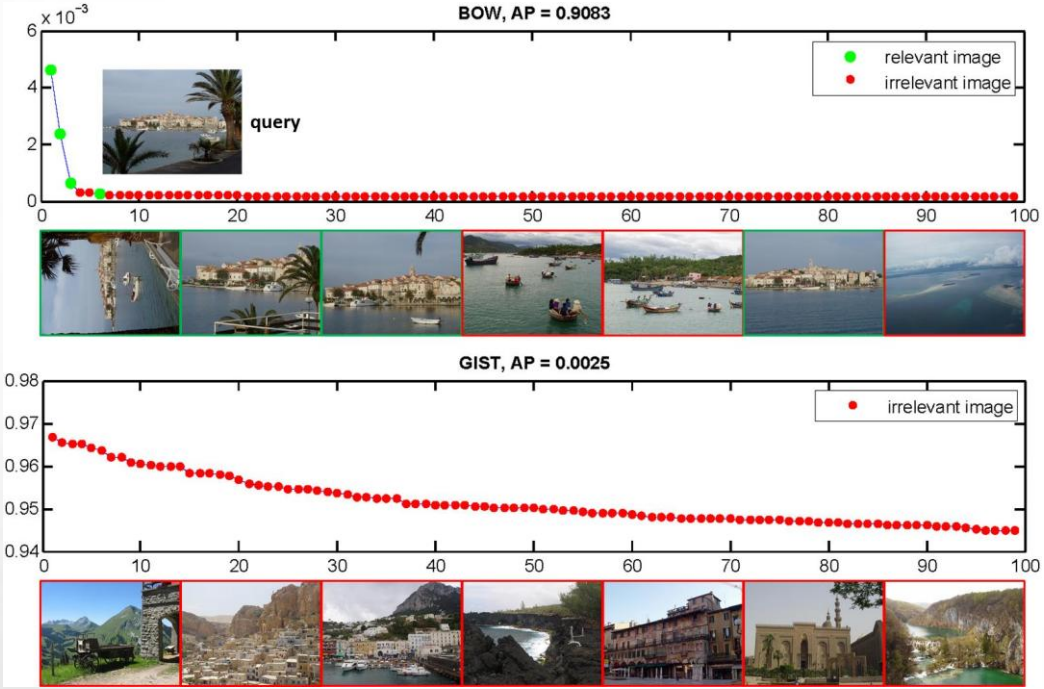
Multiple Features Fusion

Courtesy: Query-Adaptive Late Fusion for Image Search and Person Re-identification, CVPR2015

Multiple Features Fusion

Courtesy: Query-Adaptive Late Fusion for Image Search and Person Re-identification, CVPR2015

Multiple Features Fusion



Good feature:
L-shaped score curve

Bad Feature:
Flat score curve

Courtesy: Query-Adaptive Late Fusion for Image Search and Person Re-identification, CVPR2015

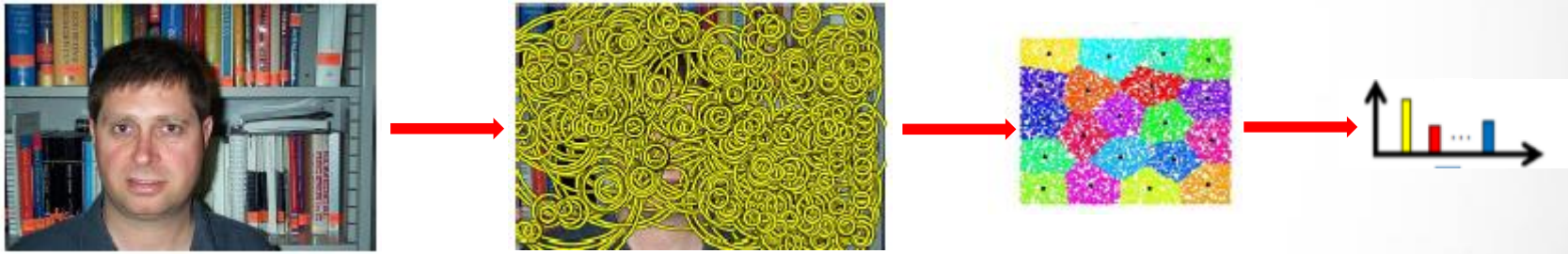
Multiple Samples Fusion

Query Fusion

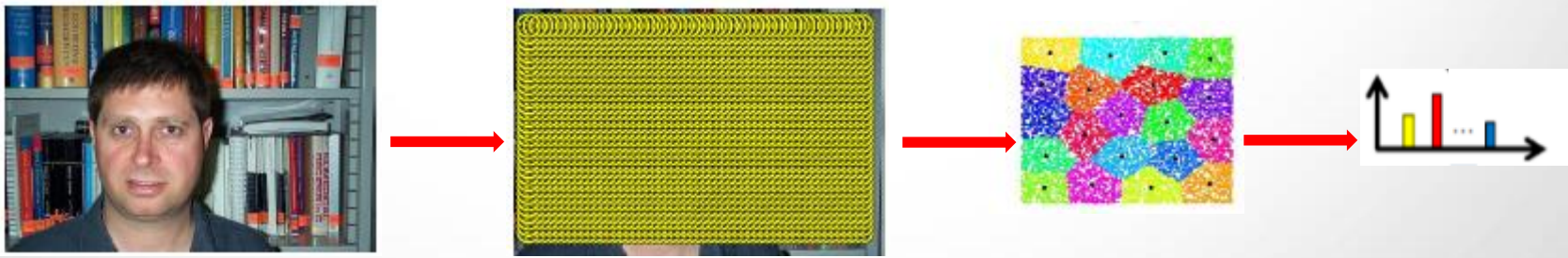
Take four samples as four features

Dense VS Sparse

Feature representation workflow for SIFT baselines

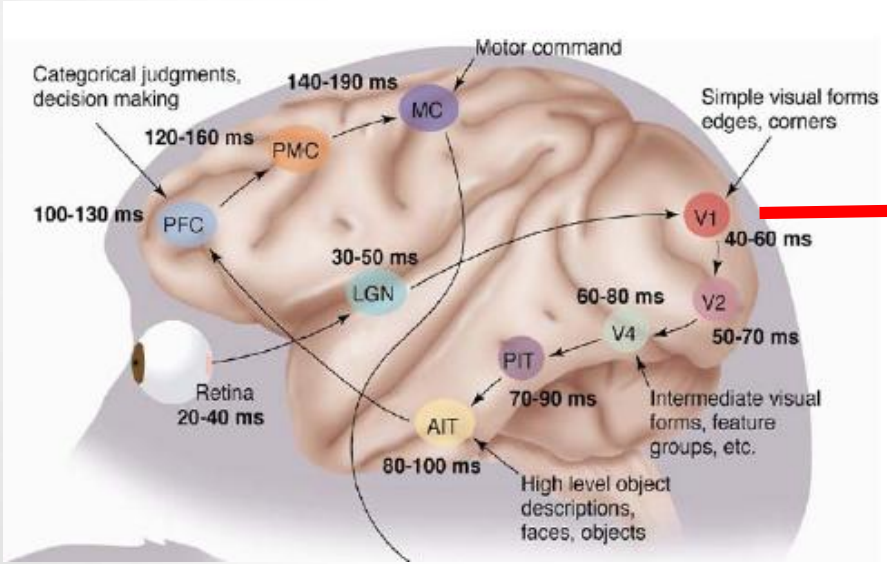


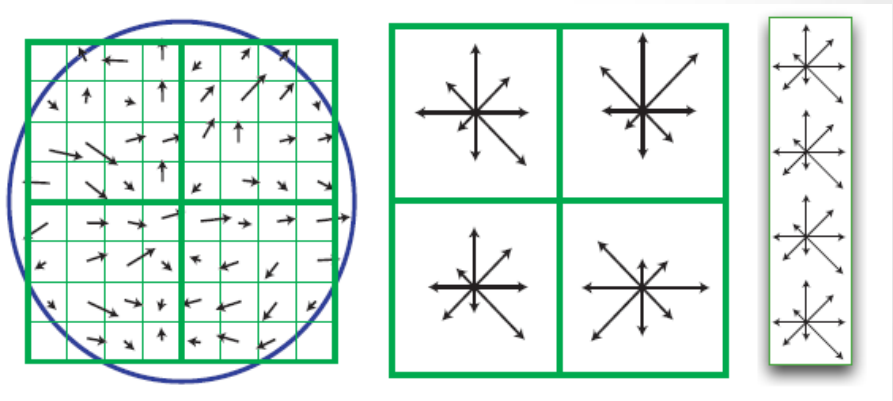
Feature representation workflow for Deep conv features



Courtesy:Dense Interest Points, CVPR2010

Tentative Experiment

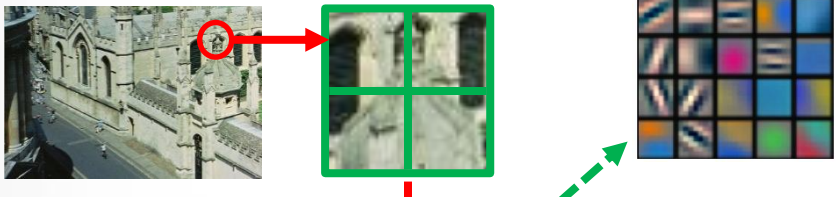




Visual system of human

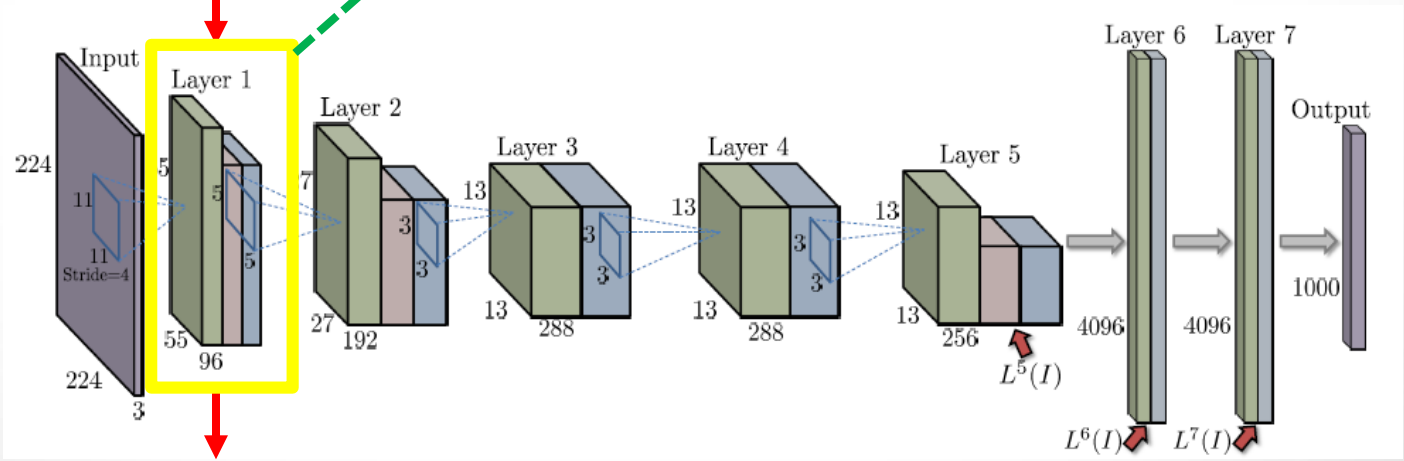
SIFT Descriptor

Tentative Experiment



Layer1 Visualization

AlexNet



384-D feature vector

Tentative Experiment

Features	mAP (2013)	mAP (2014)
MSER + RootSIFT	15.86	13.00
Hessian Affine + RootSIFT	21.59	17.03
Deep Conv5	16.58	18.37
Hessian Affine + Deep Conv1	<1	<1

Thank you