Continuous Learning for Vision?

Horst Bischof

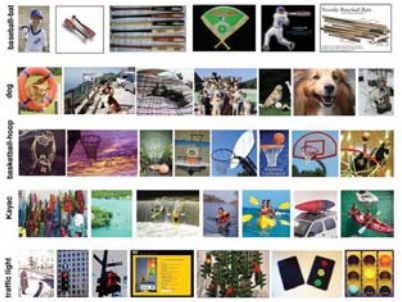
Inst. for Computer Graphics and Vision
Graz University of Technology

Recent Progress in Vision

1. Robust Features (for a certain class of objects)
   - SIFT
   - SURF
   - HOG
   - ….
2. Lots of learning algorithms
   - SVM
   - Boosting
   - MKL
   - ….

Typical Vision Tasks/Trends

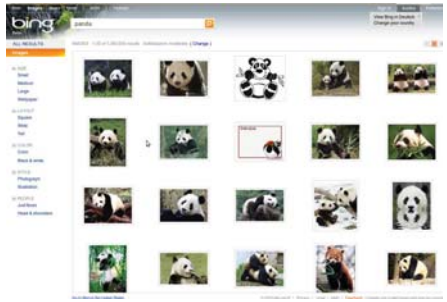
Object Categorization



Huge Amounts of data

Typical Vision Tasks/Trends

Internet: Image Search/Classification Categorization



Huge Amounts of data, data from different sources

Typical Vision Tasks/Trends

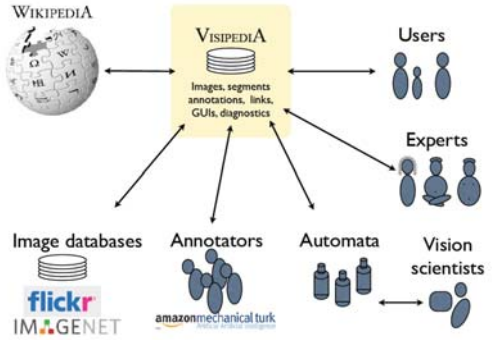
Internet: Various Video and Image Databases

Picasa
flickr
ImageShack
YouTube
Broadcast Yourself
facebook
myspace
a place for friends

Huge Amounts of data, multimodal data

Perona 2009

Visipedia



Typical Vision Tasks/Trends

Surveillance: On-line data/Detection-Tracking

Huge Amounts of data, On-line processing, Scene adaptation

Typical Vision Tasks/Trends

Robot Vision

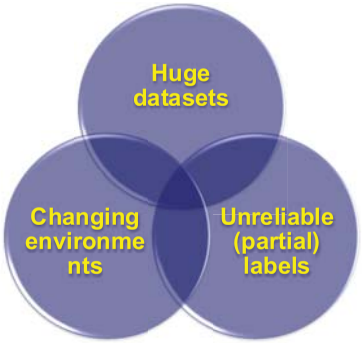
Multiple sensors, On-line adaptation, …

Typical Vision Tasks/Trends

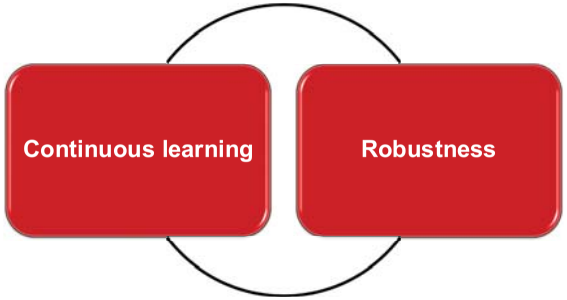
Many other tasks:

- Tracking: Adaptation to object
- Interactive segmentation: Changing model on the fly
- Interactive labeling: Suggestions as you label
- …..

Requirements



Requirements on Learner



Continuous Learning

Vision system deployed that:

1. Continuously operates in its environment
2. While in operation learns/adapts continuously

Continuous Learning

2 Problems

1. On-line learning (constant adaptation etc.)
2. Generate Learning signals/labels/teacher

Why on-line learning?

Too much training data to fit in memory
- Internet!!!

Sample generation process
- Tracking, Co-Training

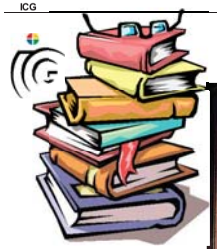
Changing processes
- Changing Environment

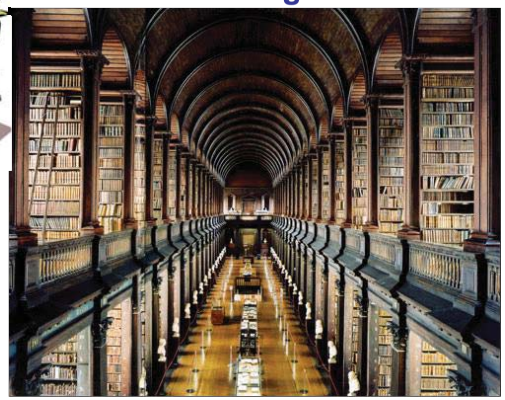
Why on-line learning?

Specializing
- Forget irrelevant information
- Specialize to current scene

Interactive Applications
- Data labeling
- Classifier Training
- Specializing (Human in the loop)
- Interactive Training (Segmentation)

Off-line learning





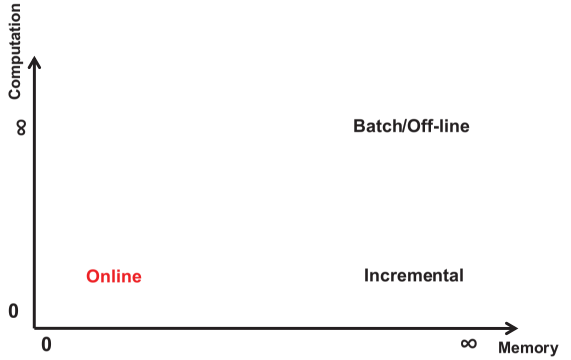
On-line learning

Learning Algorithm

Labeled Information

Teacher

Online Learning in Perspective



On-line learning algorithm

Hypothesis h_t, and new sample (x,y) output a new hypothesis h_t+1

slightly weaker version you are allowed to keep k samples and use them to update

Goal: Be as good as batch learning that has access to all samples

On-line learning algorithms

Neural Networks
- Stochastic Gradient Descent $w = w + \alpha \Delta w; \quad \Delta w = -\frac{\partial E}{\partial w}$

Winnow (Littlestone 1988)

Decision Trees (Utgoff, Berkman, & Clouse, 1997)

Universal linear prediction (Merhav& Feder, 1998; Singer & Feder, 1999)

Online Ensemble Learning (Oza 2001)

Online Adaboost (Grabner & Bischof 2006)

Online Random Forests (Leistner, Saffari, Bischof 2010)

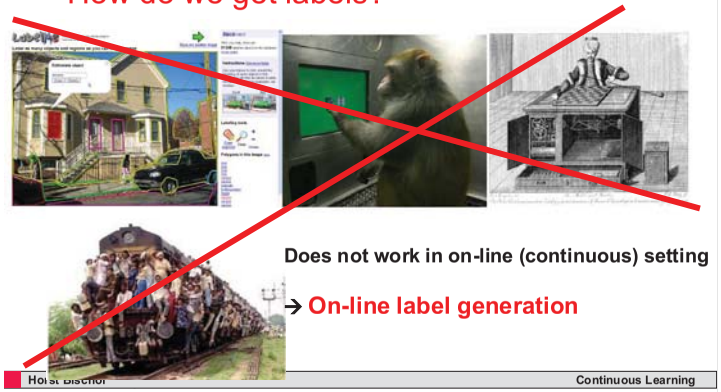
Why is it hard?

- No control over samples
- Know nothing about current sample
- Additionally in Vision:
  - **Labels** needs to be generated on-line →

    Noisy labels

Labels

How do we get labels?



Does not work in on-line (continuous) setting

→ **On-line label generation**

Tracking

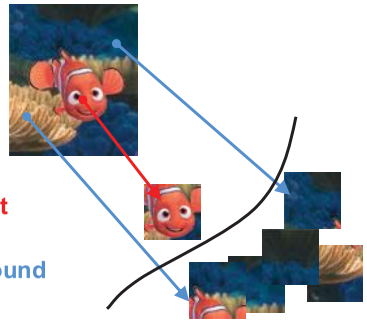
A simple example of continuous learning:

Keep Track of an unknown object

Tracking as Classification

S. Avidan. **Ensemble tracking**. CVPR 2005.
J.Wang, et al. **Online selecting discriminative tracking features using particle filter**. CVPR 2005.

Tracking as binary classification



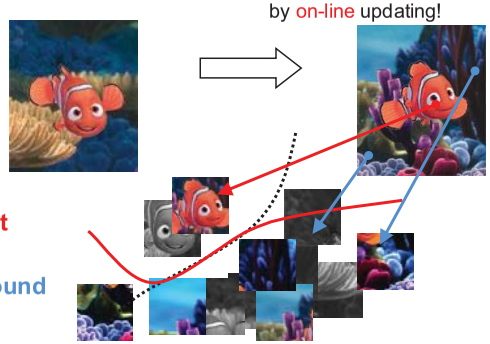
object
vs.
background

Tracking as Classification

S. Avidan. **Ensemble tracking**. CVPR 2005.
J.Wang, et al. **Online selecting discriminative tracking features using particle filter**. CVPR 2005.

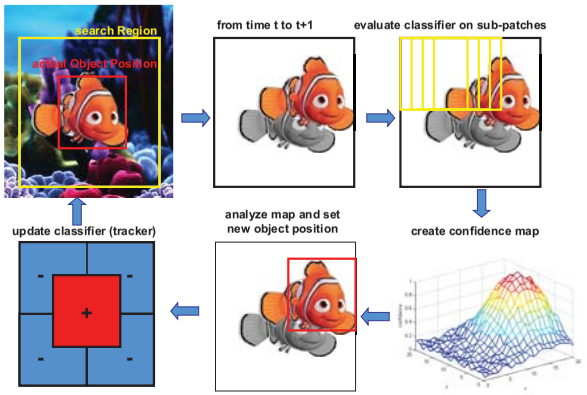
Tracking as binary classification

Object and background changes are robustly handled by on-line updating!



object
vs.
background

Template based Object Tracking



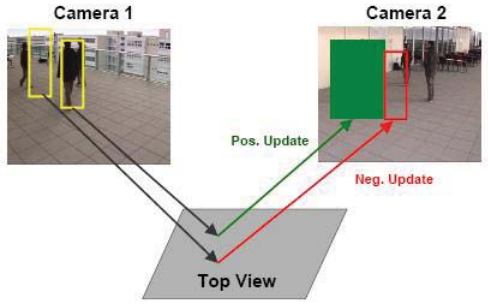
Tracking Examples

Long Term Experiment

Where do we get on-line labels?

- Own predictions (self-learning)
  - Bad idea
- User
  - Only interactive settings
- Other algorithms (Co-training)
  - Are they correct? → Robust Learning
- Other modalities (Co-training)
  - Are they correct? → Robust Learning
- Future (System success)
  - How much memory?
  - Credit assignment problem → Reinforcement Learning

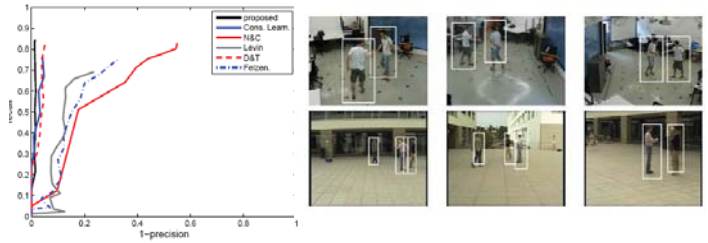
Multiple Camera Co-training



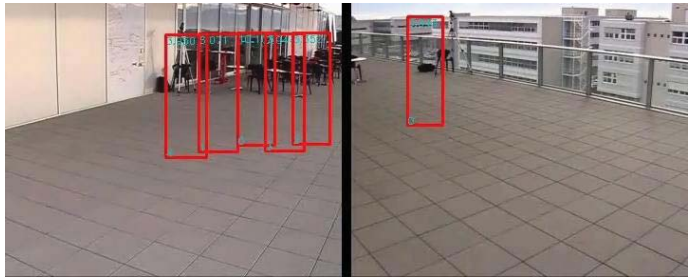
Geometry (Homography) for verification

Co-training

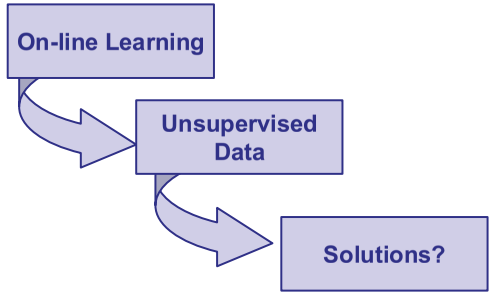
Online Multiple Instance Co-Training (Roth et.al)



Co-training



Summary



Continuous Learning in Vision

- Only plausible learning method for autonomous system
- Label generation harder than learning (stability)
- If we can learn fast/stable we can specialize at the task on the fly

The last paper I want to write

Computer vision solved!

A big step forward

Computer Wischen solved!