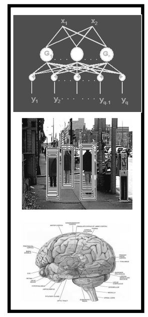
Learning:
Theory, Engineering Applications and Brains

brains and machines

Multidisciplinary Approach to Learning



LEARNING THEORY AND ALGORITHMS

ENGINEERING APPLICATIONS, PLAUSIBILITY PROOFS

NEUROSCIENCE: MODELS AND EXPERIMENTS

Classification and search in images, video, bioinformatics, time series, text, sound,..

Plan

- **Theory: Statistical Learning Theory, Regularization Networks and SVMs**
- **Applications:**
  - **Datamining applications**
  - **Graphics: Face synthesis, TTVS**
  - **Vision: object detection, face estimation**
  - **Bioinformatics: DNA arrays**
  - **Neuroscience: IT model, action recognition, circuits**

Learning from Examples

INPUT → f → OUTPUT

EXAMPLES

$INPUT_1$ → $OUTPUT_1$

$INPUT_2$ → $OUTPUT_2$

....

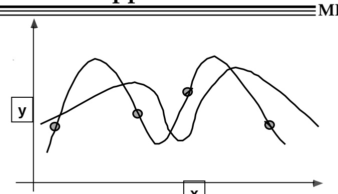
$INPUT_n$ → $OUTPUT_n$

Supervised Learning as Multivariate Function Approximation

= data from f

= function f

= approximation of f



Generalization: **estimating value of function where there are no data**

Regression: **function is real valued**

Classification: **function is binary**

General Statistical interpretation...

Learning from examples needs prior constraints= regularization=capacity control

Generalization problem is ill-posed:

Generalization requires a priori assumption/capacity control such as *smoothness*

"Classical" example: Regularization Networks

- **The function f that minimizes**

$$H[f] = \sum_{i=1}^{N} (y_i - f(\mathbf{x}_i))^2 + \gamma \|f\|_K^2$$

has the form

$$f(\mathbf{x}) = \sum_{i=1}^{N} c_i K(\mathbf{x}, \mathbf{x}_i) + p(\mathbf{x})$$

where K is the basis function associated with the RKHS norm (the regularizer term)

A unified theory

- **Regularization networks -- such as Gaussian Radial Basis Functions -- and also Support Vector Machines for regression (SVMR) and for classification (SVMC) can be justified in terms of a new, fundamental statistical theory of learning.**
- **The theory (Vapnik…) deals with the problem of learning from finite and small sample sizes.**
- **Its focus is on the problem of capacity control, i.e. the learning and statistical version of regularization.**
- **Its key motivation is to do classification and regression without density estimation**

Unified framework: RN, SVMR and SVMC

$$\min_{f \in \mathcal{H}} H[f] = \sum_{i=1}^{l} V(y_i - f(\mathbf{x}_i)) + \gamma \|f\|_K^2$$

Very close relationship: RN, SVMR and SVMC

The three techniques admit the same solution

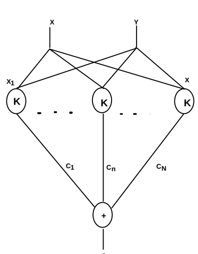
$$f(x) = \sum_{i=1}^{l} c_i K(\mathbf{x}_i, \mathbf{x})$$

(…modulo an additive p(x) term)

Equivalence to Networks

$$f(x) = \sum_{i=1}^{l} c_i K(\mathbf{x}_i, \mathbf{x})$$

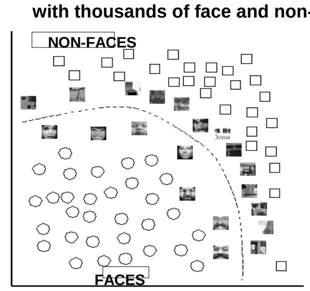
can be "written" as a network….



Example 1: Gaussian Radial Basis Functions with K a Gaussian
Example 2: Linear case with K(x,y) = $x \circ y$

Support Vectors are Sparse!

Some of the support vectors found by SVM training with thousands of face and non-face examples



…that is the coeffcients in

$$f(x) = \sum_{i=1}^{l} c_i K(\mathbf{x}_i, \mathbf{x})$$

are sparse!

Learning to do computer graphics

CBCL+AI MIT

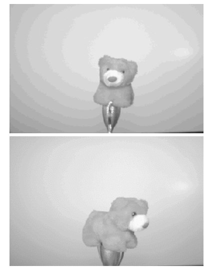
- Learning to draw cartoons from an artist
- Learning to synthesize a specific face
- Making a talking face (TTVS)

IAP, 2000

Image Analysis

CBCL+AI MIT

IMAGE ANALYSIS: OBJECT RECOGNITION AND POSE ESTIMATION



⇒ Bear (0° view)

⇒ Bear (45° view)

IAP, 2000

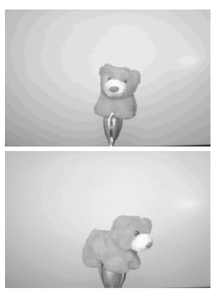
Image Synthesis

CBCL+AI MIT

UNCONVENTIONAL GRAPHICS

Θ = 0° view ⇒

Θ = 45° view ⇒



IAP, 2000

nFX Interactive demo

CBCL+AI MIT

IAP, 2000

3D Morphable Model for 3D Shape-from-Single-Image

CBCL+AI MIT

FOR LINEAR CLASSES ONE 2D IMAGE IS SUFFICIENT TO ESTIMATE 3D SHAPE *BECAUSE (assuming correspondence!)*

$$\mathbf{x}_0^f = \sum_{i=1}^{q} \alpha_i \mathbf{x}_i^f$$

implies

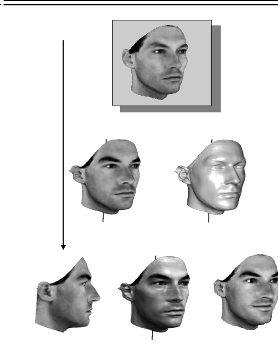
$$X_0 = \sum_{i=1}^{q} \alpha_i X_i$$

THUS 3D SHAPE CAN BE ESTIMATED FROM A SINGLE IMAGE, GIVEN A FLEXIBLE 3D MODEL BUILT FROM A SET OF 3D PROTOTYPES (Poggio and Vetter, 1992)

IAP, 2000

3D from one 2D Image

CBCL+AI MIT



2D Input Image

3D Reconstruction

New Views

Vetter & Blanz ECCV'98

IAP, 2000

Learning, Approximation and Networks

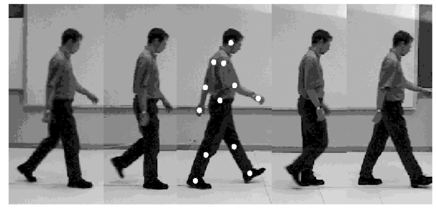
CBCL+AI MIT

Video 1: Thomas Vetter and Volker Blanz

IAP, 2000

Dynamic morphable models (M. Giese)

CBCL+AI MIT

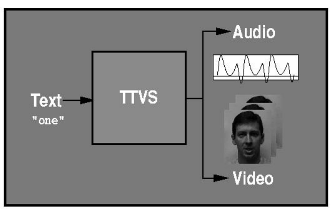


IAP, 2000

Another example of graphics: TTVS (T. Ezzat)

CBCL+AI MIT

- **Input: Text**
- **Output: photorealistic talking face uttering text**

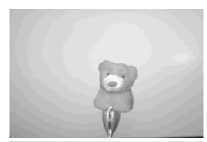


IAP, 2000

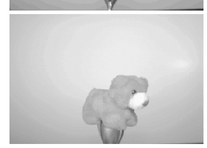
Image Analysis

CBCL+AI MIT

IMAGE ANALYSIS: OBJECT RECOGNITION AND POSE ESTIMATION



⇒ Bear (0° view)



⇒ Bear (45° view)

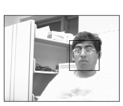
IAP, 2000

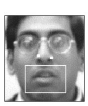
An example of vision: Face analysis system (V. Kumar)

CBCL+AI MIT

T*rainable* system capable of estimating basic mouth features such as degree of openness in *real time*.

Face Detection → Localization of Facial Features → Analysis of Facial parts



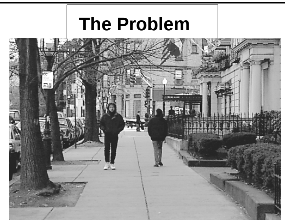


IAP, 2000

A Trainable Object Detection System

CBCL+AI MIT

The Problem

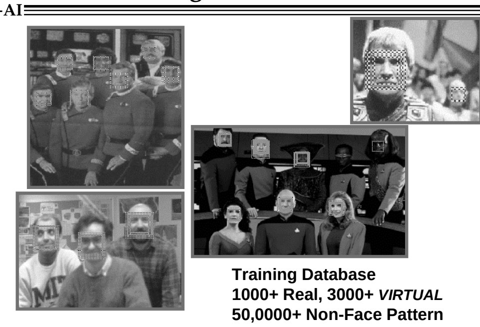


- **Developing a *general* paradigm for object detection in cluttered scenes**
- **Applications: target detection, visual data base search...**
- **Trainable system…for “any” desired object class**

IAP, 2000

Learning Object Detection: Finding Frontal Faces ...

CBCL+AI MIT



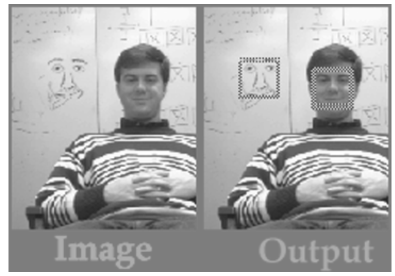
Training Database
1000+ Real, 3000+ *VIRTUAL*
50,0000+ Non-Face Pattern

Sung, Poggio 1995

IAP, 2000

Learning Face Detection

CBCL+AI MIT

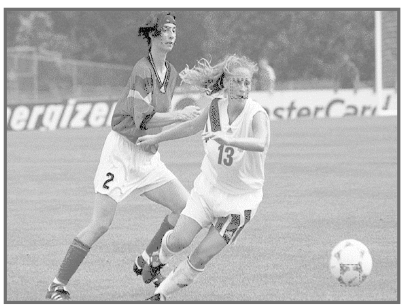


Sung, Poggio 1994

IAP, 2000

Face detection:…

CBCL+AI MIT



IAP, 2000

A trainable system for object detection

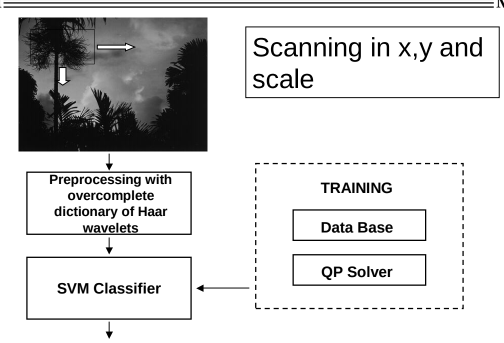
CBCL+AI MIT

Video 2: K.K. Sung

IAP, 2000

System Architecture (K.K. Sung, C. Papageorgiou + A. Mohan)

CBCL+AI MIT



IAP, 2000

Combining component classifier, i.e. "boosting" for pedestrian detection

CBCL+AI MIT

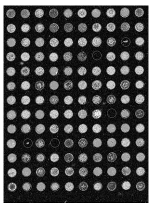
IAP, 2000

BIOINFORMATICS

CBCL+AI MIT

Support Vector Machine Classification of Microarray Data

S. Mukherjee, P. Tamayo, J.P. Mesirov, D. Slonim, A. Verri, T. Poggio



Collaboration with Whitehead Institute, MIT

IAP, 2000

DNA arrays: high dim (7100 genes expression data)

CBCL+AI MIT

Gene expression data: 38 training, 34 test points

Tamayo et al (Science Sept 99), form of linear classifier with feature selection: 100% correct, 5 rejections, 2 errors

Sayan Mukherejee et al., linear SVM, no feature selection: 100% correct, 1 rejection, 1 error

More difficult lymphoma data soon!

IAP, 2000

Center for Biological and Computational Learning + AI Lab

CBCL+AI MIT

Models of object and action recognition in cortex

(M. Riesenhuber, M. Giese. A. Yu)

IAP, 2000

1. A key problem in (subordinate level) object recognition

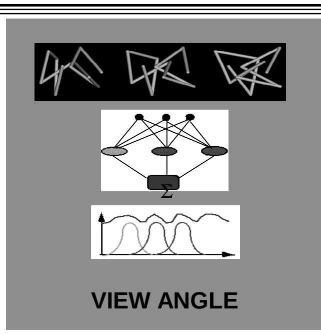
CBCL+AI MIT

How does the brain learn to recognize a specific object -- like a face -- in a view-independent manner?

IAP, 2000

Model of view-invariant recognition: learning from views

CBCL+AI MIT



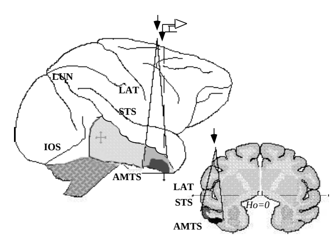
A graphical rewriting of mathematics of regularization (GRBF), a learning technique

Poggio, Edelman *Nature*, 1990.

IAP, 2000

Recording Sites in Anterior IT

CBCL+AI MIT

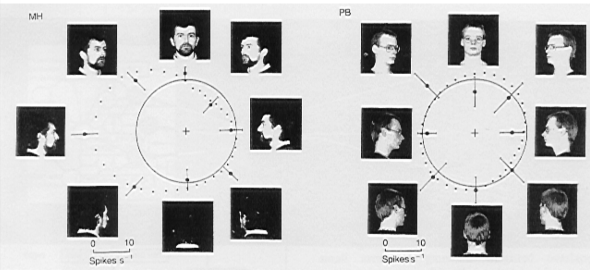


Logothetis, Pauls, and Poggio, 1995; Logothetis, Pauls, 1995

IAP, 2000

Previous glimpses: cells tuned to face identity *and* view

CBCL+AI MIT

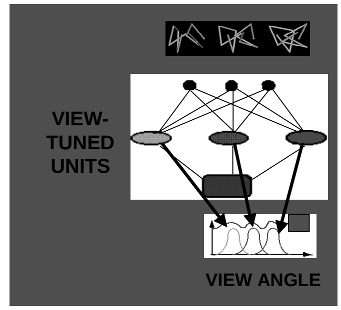


Perrett, 1989

IAP, 2000

Model's predictions: View-tuned Neurons

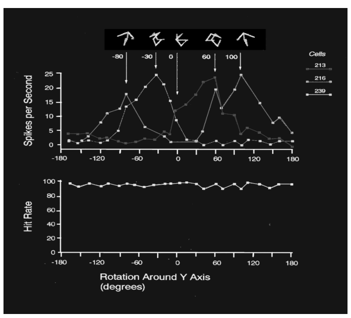
CBCL+AI MIT



IAP, 2000

The Cortex: Neurons Tuned to Object Views

CBCL+AI MIT



Logothetis, Pauls, Poggio 1995

IAP, 2000

Sparse Representations in IT

CBCL+AI MIT

In the recording area in AMTS -- a specialized region for paperclips (!) -- we estimate that there are, after training, *(within an order of magnitude or two)* …

- **About 400 view tuned cells per object**
- **Perhaps 20 view-invariant cells per object**

Logothetis, Pauls, Poggio, 1997

IAP, 2000

2. View-tuned IT neurons

CBCL+AI MIT

View-tuned cells in IT Cortex:

how do they work?

How do they achieve selectivity and invariance?

Max Riesenhuber and T. Poggio, Nature Neuroscience, just published

IAP, 2000

One of the main predicions

CBCL+AI MIT

There should be evidence of softmax at several layers in cortex (not all!). In particular, complex cells in V1/V2 should show softmax properties (under certain conditions)

IAP, 2000