

Evaluating Performance for Computer Vision Systems

James L. Crowley

Professor I.N.P. Grenoble

Directeur Laboratoire GRAVIR

(UMR CNRS 5527)

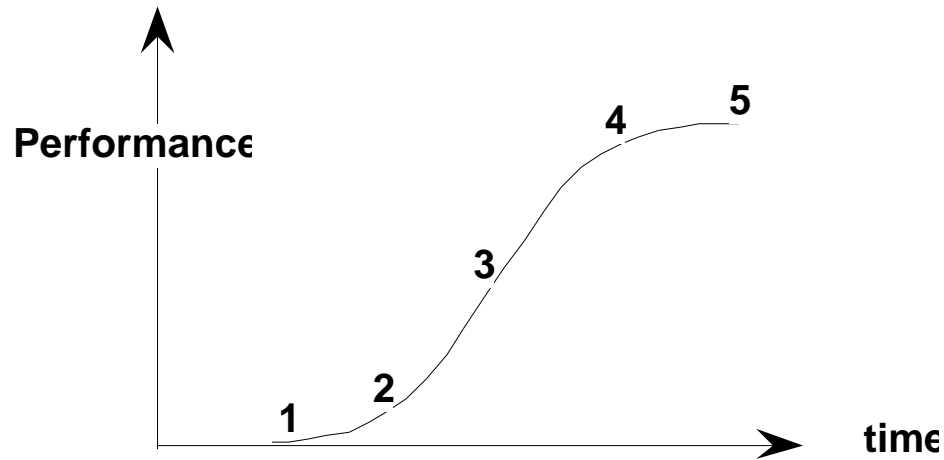
Evaluating Performance of Computer Vision Systems

Plan

- Econometrics: Performance Evaluation in Technological Innovation
- Recent examples of performance evaluation in computer Vision
- Advice for TechnoVision.

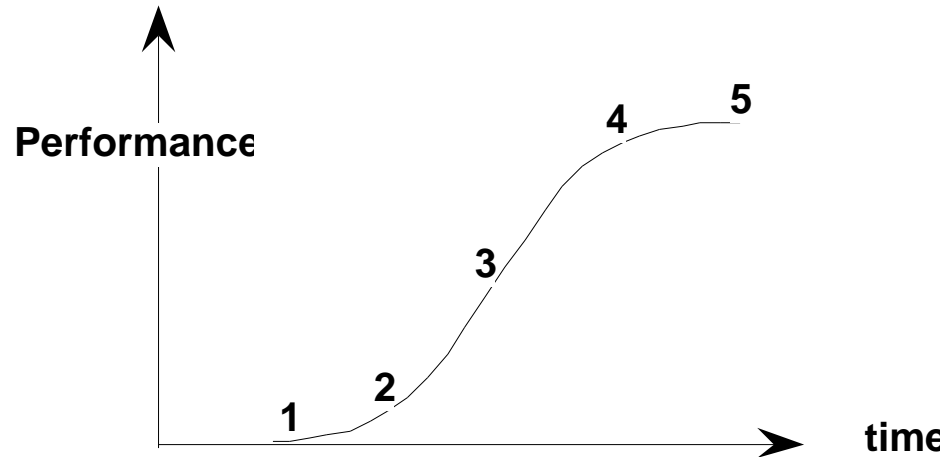
Performance Evaluation in Technological Innovation

The S Curve Model of Technology Development



As measured by any performance metric, technology matures as a cumulative Gaussian distribution.

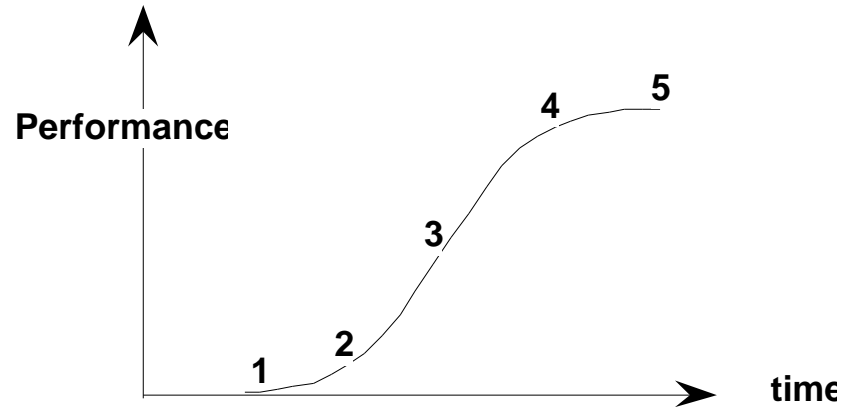
The S Curve Model of Technology Development



Phases of Technology Development:

- 1) Innovation (New idea triggers development)
- 2) Early Development (new prototypes, low volume)
- 3) Expansion (rapid growth, growing volume)
- 4) Maturation (slowing growth rate)
- 5) Saturation (Technology reaches a physical barrier).

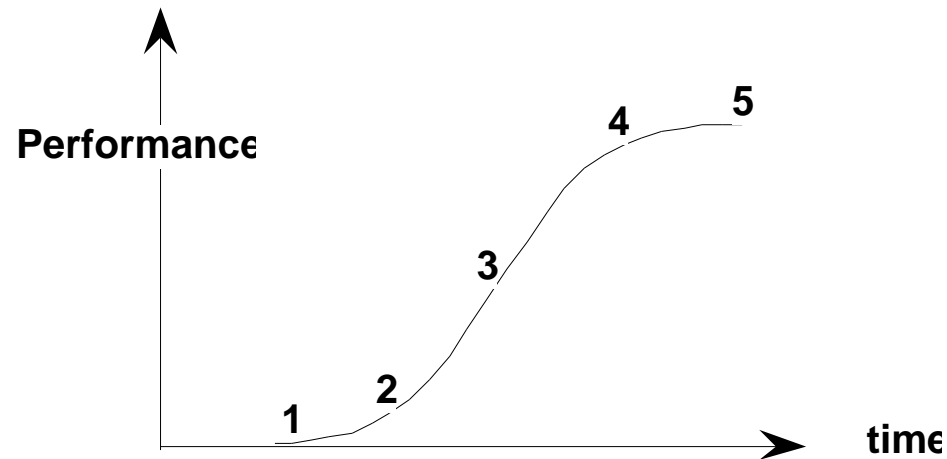
Examples



Some technologies that have matured in the 20th Century

- Electricity (1870 - 1930)
- Telephone (land lines) (1875 - 1935)
- Aviation (1900 - 1950)
- Telephones (Land Lines) (1870 - 1930)
- Radio-Communications (1895 - 1945)

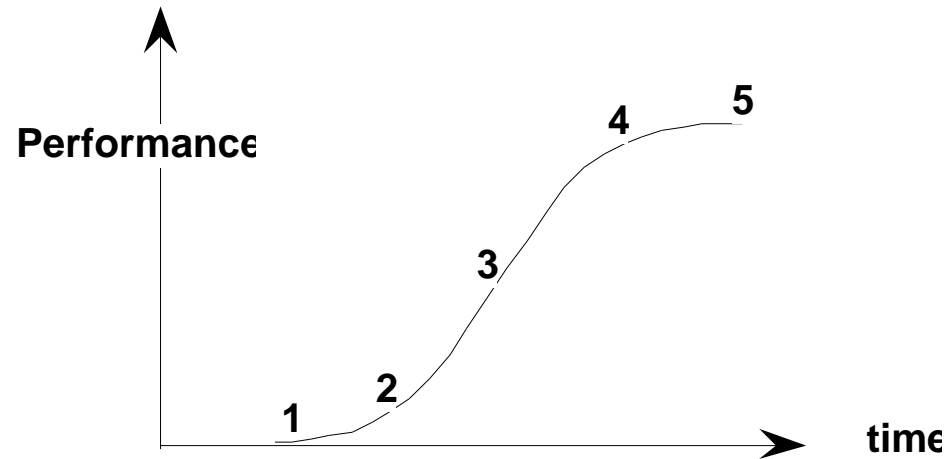
The S Curve Model of Technology Development



S-Curve measures a Virtuous Cycle:

- Performance growth creates market growth.
- Market growth generates resources for innovation
- Resources stimulate performance growth (through research).

The S Curve Model of Technology Development



Performance Measures:

Define the cycle: Make competition possible

Maintain the cycle: By documenting effects of investment

Performance Evaluation in Computer Vision

Some Recent Examples:

- The PETS Series
- Pointing '04
- Performance Evaluation in CHIL

The PETS Workshop Series

Goal: Foster the emergence of methods and metrics for performance evaluation.

PETS '00: Grenoble March 2000 (FG 2000)

PETS '01: Kuai, Dec 2001, (CVPR '01)

PETS '02: Copenhagen, June 2002 (ECCV '02)

PETS-ICVS '03: Graz April 2003 (ICVS '03)

VS-PETS-ICCV '03: Nice August 2003 (ICVS '03)

PETS '04: Prague, May 2004 (ECCV)

PETS '05: Colorado, Jan 2005

PETS 2000 Data Sets

Provided 3 Image Sequences

Task: track cars and people.

Ground truth NOT published

(No sponsor)

Authors invited to propose metrics
for their results.

QuickTime[®] and a
decompressor
are needed to see this picture.

Workshop objective:

Agree on performance metrics!

FG Net

Face and Gesture Recognition Working group
(IST-2000-26434)

Coordinator : Univ of Manchester (P. Courtney)

Duration: Oct 2001 - Sept 2004

Activities:

- Technology Foresight workshops
- Benchmark definition and collection
- Performance Evaluation Workshops

FGNet:

Role of INRIA Rhône Alpes

Web Site Maintenance

Performance Evaluation Workshops

Workshops:

2002 - PETS 2002 - Shop Window Scenario

2003 - PETS 2003 - Meeting Minutes

2004 - Pointing 2004 - Diectic Gestures

PETS 2002 Data Set

Shop Window Scene

Three Scripts (Easy, Challenging, Hard), two sequences per scenario.

Task:

- how many people are passing in front of the shop window
- how many people stop and look into the window
- how many people are looking into the window at each instant (frame) in time
- the trajectories of people passing in front of the store
- the time spent per frame (processing time):

The task is made difficult by:

- the reflections from the ground and from the opposing window
- occlusions from the text on the window
- occluding groups of people walking together
- occlusions among the people standing in front of the store

PETS 2002 Scenario 1

Scenario 1 : Isolated People

QuickTime^a and a
decompressor
are needed to see this picture.

PETS'2002
3rd IEEE International Workshop on
Performance Evaluation of Tracking and Surveillance

Copenhagen, June 1 2002 (In conjunction with ECCV)
Organised by J. Ferryman and J. L. Crowley

11 papers presented.

35 Registered participants

50 copies of proceedings printed

PETS-ICVS'03

Observation of a Meeting

Three views of the smart meeting Ombi-directional and wall mounted.

The Task

Automatically annotate the smart meeting.

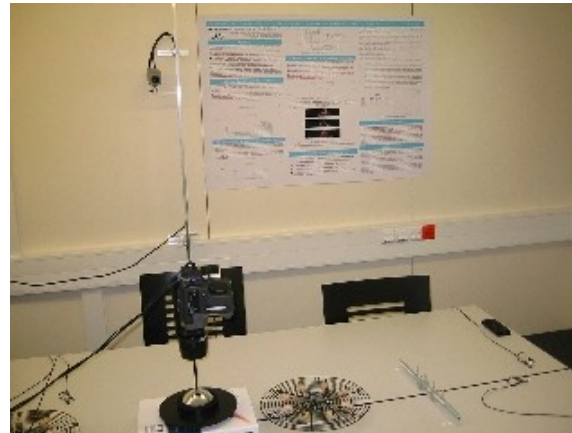
Four scenarios A, B, C and D.

Each scenario consists of a number of separate sub-tasks.

- * face localisation (centre location of eyes)
- * recognition of facial expression
- * recognition of face/hand gesture
- * estimation of face/head direction (gaze)
- * recognition of actions

PETS ICVS Data Sets

Three cameras: One Panoramamic and Two Wall Mounted



PETS-ICVS'03

Observation of a Meeting

31 March 2003 - Graz, Austria - In conjunction with ICVS '03
Organised by J. Ferryman and J. L. Crowley

9 Papers presented.

35 Copies of Proceedings Printed

22 Registered Attendees.

PETS ICVS Data Sets

4 Scenarios A, B, C, D - Scripts and Ground Truth Provided

Scenario A: "Performing distinct Facial Expressions"

Actions: Sitting down, getting up, smile, angry, neutral, looking at other participants

Duration/size: 500 seconds, 1.2 Gb (JPEG)

Scenario B: "Performing face & hand gestures"

Actions: Sitting down, getting up, raising hand, shaking head, nodding head, yawning, laughing

Duration/size: 333 seconds, 1.0 Gb (JPEG)

Scenario C: "Going to the white board"

Actions: Sitting down, getting up, going to the white board

Duration/size: 407 seconds, 1.2 Gb (JPEG)

Scenario D: "Artificial meeting"

Actions: A little bit of everything

Duration/size: 577 seconds, 2.5 Gb (JPEG)

PETS 2004 Data Sets

observing human activity in public places

Six scenarios were recorded at Lobby of INRIA Rhône-Alpes

Activities included

- a person walking in a straight line (3 sequences),
- a person browsing at information displays (5 sequences),
- behaviours while seated in a chair (3 sequences),
- persons abandoning packages (5 sequences),
- groups of people encountering (6 sequences), and
- people fighting (4 sequences).

Ground Truth : University of Edinburgh

PETS 2004 Data Sets

observing human activity in public places

QuickTime^a and a
decompressor
are needed to see this picture.

PETS 2004 Data Sets

observing human activity in public places

The PETS challenge: Demonstrate automatic labeling

6 Scenarios - 26 Sequences

Hand labeled ground truth for all sequences

For each sequence: Half of Ground Truth published.

PETS ECCV 2004

6th International Workshop on
Performance Evaluation for Tracking and
Surveillance

Zofin Palace, Slovansky ostrov
Prague 1, Czech Republic

10 May 2004

POINTING 2004

Visual Observation of Deictic Gestures

International Workshop

ICPR 2004

Sponsored by FGnet

The Pointing 04 Data Sets

Head Pose Image Data Base

Pointing Gesture Data Base

The Pointing '04 Head-Pose Image Database

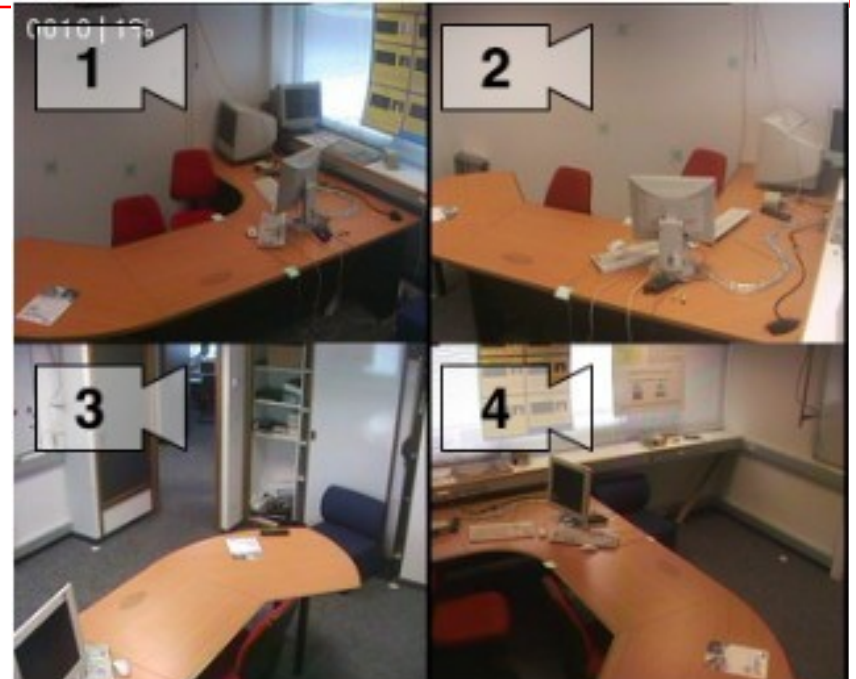
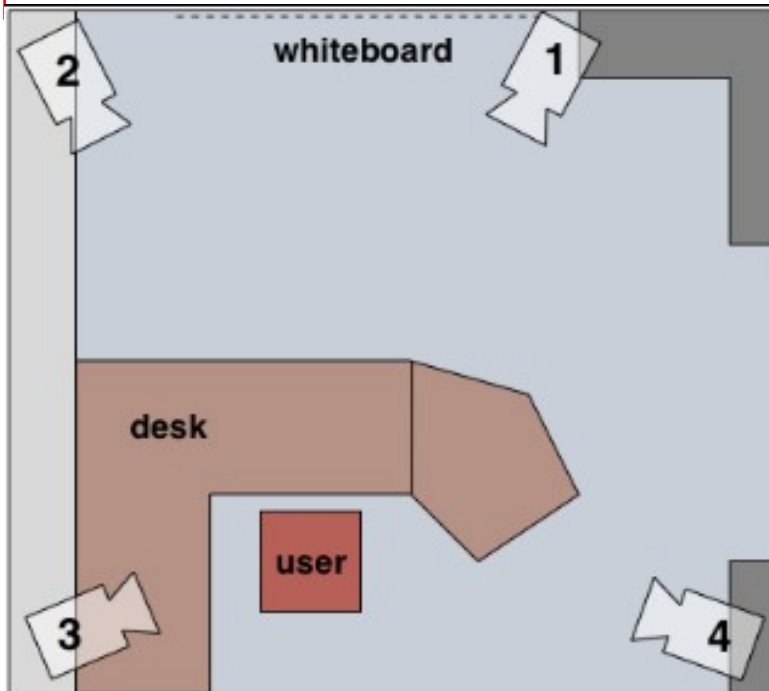


15 Subjects = 15 image sets.

For each subject 2 series of 93 images
first series for training, second for testing.

Head orientation (h,v), from -90 degrees to +90 degrees.

Pointing Gestures: video sequence database



4 subjects, sequences per subject = 8 video sequences.

Simultaneous video from four ceiling cameras

User points at target projected on wall by video projector

Ground truth published for 4 of 8 sequences (1 per subject).

Pointing Gestures sequences database

Capture

25Hz, non-interleaved PAL stream (352 x 288 pixels).

Captured by *video4linux* and *ffmpeg*

Four view synchronized *a posteriori*;

Maximum delay between two views is one frame (40 ms).

Script : The subject

Enters the office and sits

Displays a projected pattern on the desk

Points at 8 positions on the whiteboard

Exits the office.

Lighting conditions

Scene illumination 60% natural diffuse light and 40% neon light.

Calibration Data provided for four cameras

Pointing'04

Visual Observation of Deictic Gestures

22 August 2004 - Cambridge, UK - In conjunction with ICPR '04
Organised by J. L. Crowley, J. Letessier and N. Gourier

8 paper presented

29 registered participants

35 copies of proceedings printed.

FGNet Web Site

Working Group on Face and Gesture Recognition

http://mercury.local/FGnet/index.html

Trombinosco...ire GRAVIR Google Amazon.co.uk SIA - La réf...éronautique OLIVIA Météo Fran... - AEROWEB Apple



FGnet - IST-2000-26434

Face and Gesture Recognition Working group

FGnet Home Page

[About FGnet](#)

[Members](#)

[How to Join](#)

[FGnet](#)

[Meetings](#)

[Scientific](#)

[Conferences](#)

[Technology](#)

[Surveys](#)

[Evaluation](#)

[Reports](#)

[Evaluation](#)

[Methods](#)

[Benchmark Data](#)

[Network Intranet](#)

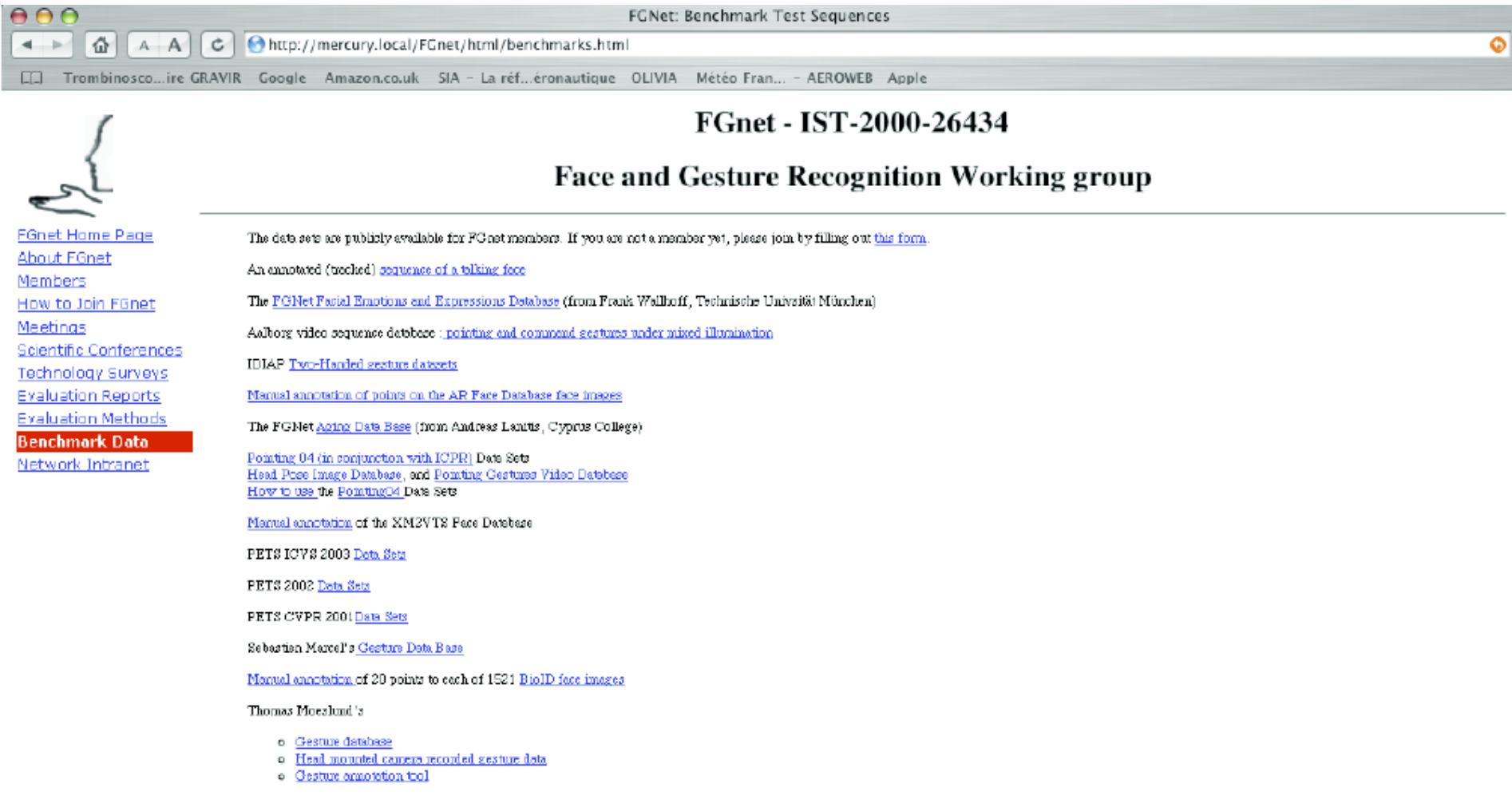
FGnet is the European working group on face and gesture recognition funded by the E.C.IST program. The objectives of FGnet are to encourage development of a technology for face and gesture recognition. The network goals are:

- (1) to assist development face and gesture recognition technology
- (2) to create a set of foresight reports defining development roadmaps and future use scenarios for the technology in the medium (5-7 years) and long (10-20 years) term
- (3) to specify, develop and supply resources (eg image sets) supporting these scenarios

This web site is maintained by FGnet as a tool for internal communication and as a means of providing information and results to the larger scientific community.

The Coordinator of FGnet is [Tim Cootes](#), University of Manchester, UK.


FGNet Web Site



FGNet: Benchmark Test Sequences

http://mercury.local/FGnet/html/benchmarks.html

Trombinosco...ire GRAVIR Google Amazon.co.uk SIA - La réf...éronautique OLIVIA Météo Fran... - AEROWEB Apple



FGnet - IST-2000-26434

Face and Gesture Recognition Working group

The data sets are publicly available for FGnet members. If you are not a member yet, please join by filling out [this form](#).

An annotated (tracked) [sequence of a talking face](#)

The [FGNet Facial Emotions and Expressions Database](#) (from Frank Wallhoff, Technische Universität München)

Aalborg video sequence database: [pointing and command gestures under mixed illumination](#)

IDIAF [Two-Handed gesture datasets](#)

[Manual annotation of points on the AR Face Database face images](#)

The FGNet [Aging Data Base](#) (from Andreas Lantini, Cyprus College)

[Pointing 04 \(in conjunction with ICPR\)](#) Data Sets

[Head Pose Image Database](#), and [Pointing Gestures Video Database](#)

[How to use the Pointing04 Data sets](#)

[Manual annotation](#) of the XM2VIS Face Database

FETS ICVS 2003 [Data Sets](#)

FETS 2002 [Data Sets](#)

FETS CVPR 2001 [Data Sets](#)

Sebastian Marcel's [Gesture Data Base](#)

[Manual annotation](#) of 20 points to each of 1521 [BioID face images](#)

Thomas Muesel's

- o [Gesture database](#)
- o [Head mounted camera recorded gesture data](#)
- o [Gesture annotation tool](#)

Lessons for TechnoVision

Applications drive Markets:

- Define Applications Scenarios
- Define Performance Requirements for Applications
- Define Benchmark examples from Scenarios

Iterative Evaluation (cooperative competition)

- Organise the evaluations in "rounds"
- The first round is for evaluating the evaluation!
- Expect convergence after several rounds

Publish your evaluation metrics!

- Publish (some of) your data
- Industry can not grow without recognized metrics