CAVIAR

Context Aware Vision using Image-Based Active Recognition

Robert B. Fisher, Edinburgh University

James L. Crowley, INRIA, INPG, UJF, CNRS

José Santos Victor, IST

CAVIAR
Context Aware Vision using Image-based Active Recognition

Presentation Plan:

Project Overview	(R. B. Fisher)
WP1 - Active image acquisition	(J. Santos-Victor)
WP2 - Control Architectures	(R. B. Fisher)
WP3 - Selective Attention	(R. B. Fisher)
WP4 - Recognition Processes	(R. B. Fisher, J. L Crowley)
WP5 - Model Learning	(J. L Crowley)
WP6 - Performance Evaluation	(R. B. Fisher)
WP7 - Dissemination	(R. B. Fisher)

WP 4: Recognition processes

Objective:

Define, implement and compare attention driven methods for detection and observation of entities and properties

Tasks:

T4.1 Define and implement visual processes

T4.2 Define representations for object, situation and context

T4.3 Develop processes of object recognition under variation of viewpoint

T4.4 Develop processes for diagnosing and recovering from errors

Deliverables

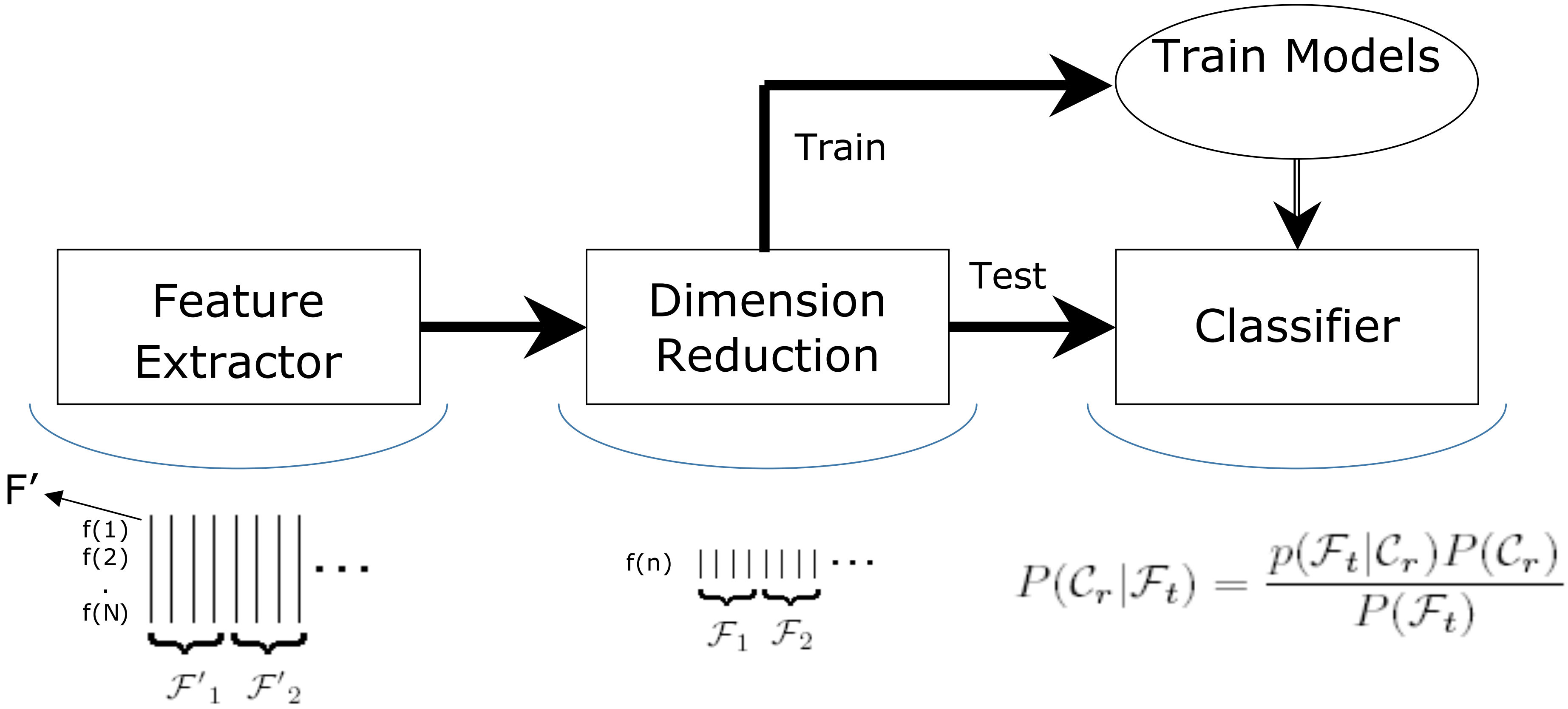
D22/24 - Report and Software on Property Observation

D29/34 - Report and Software for Error Recovery

D31/36 - Report and Software for Object Recognition

Activity Recognition

framework



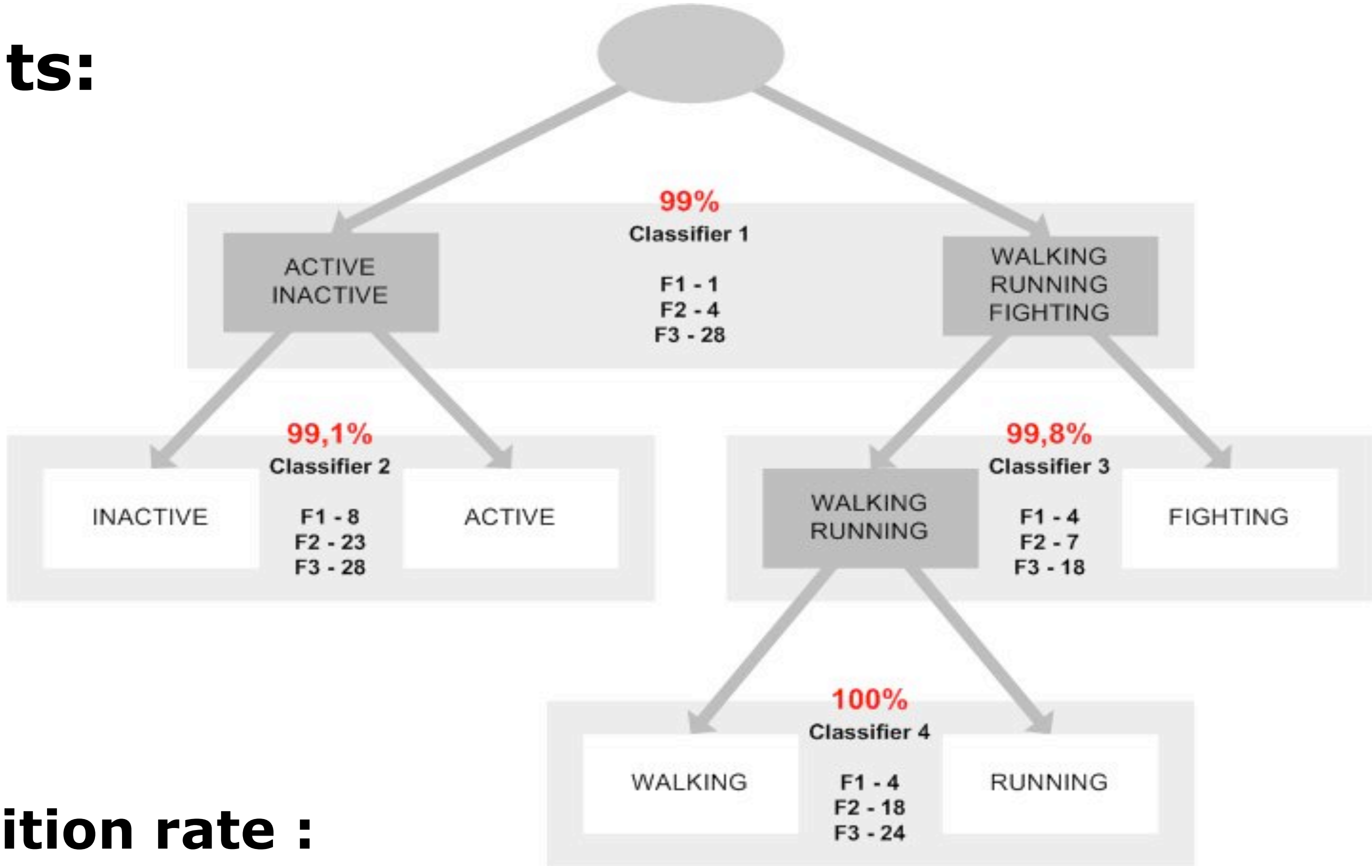
Features set

Selection
Compression

Bayesian
framework

4. classifier structure

Results:



Recognition rate :
98,8%

WP 4: Recognition processes

Objective:

Define, implement and compare attention driven methods for detection and observation of entities and properties

Tasks:

T4.1 Define and implement visual processes

T4.2 Define representations for object, situation and context

T4.3 Develop processes of object recognition under variation of viewpoint

T4.4 Develop processes for diagnosing and recovering from errors

Deliverables

D22/24 - Report and Software on Property Observation

D29/34 - Report and Software for Error Recovery

D31/36 - Report and Software for Object Recognition

D22/24
Report and Software on Property Observation

Problem:

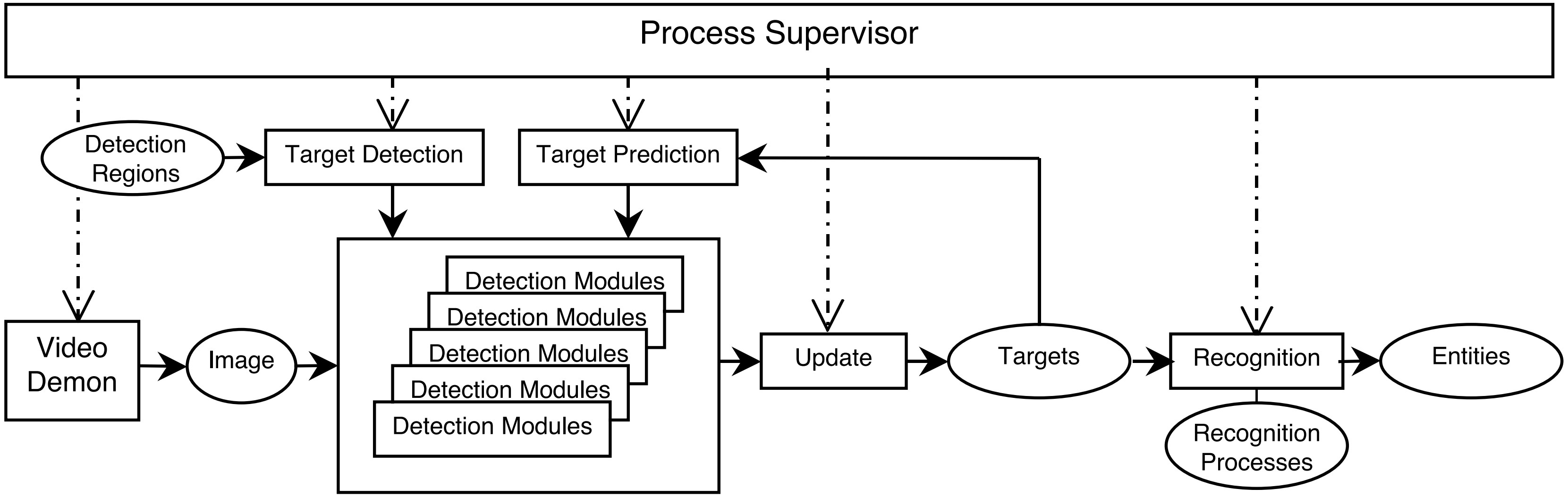
Measure properties for detecting entities and recognizing entities as objects (playing roles in situations).

Approach:

1) Attributes of tracked entities (position, size, velocity, etc)

2) View invariant descriptions of Appearance (structural and statistical)

D22/24
Report and Software on Property Observation



Attributes for tracked entities: Position, Speed, Direction, Activity Level

<u>Experiment</u>:

Compare observed values with ground truth in hand-labeled benchmark sequences.

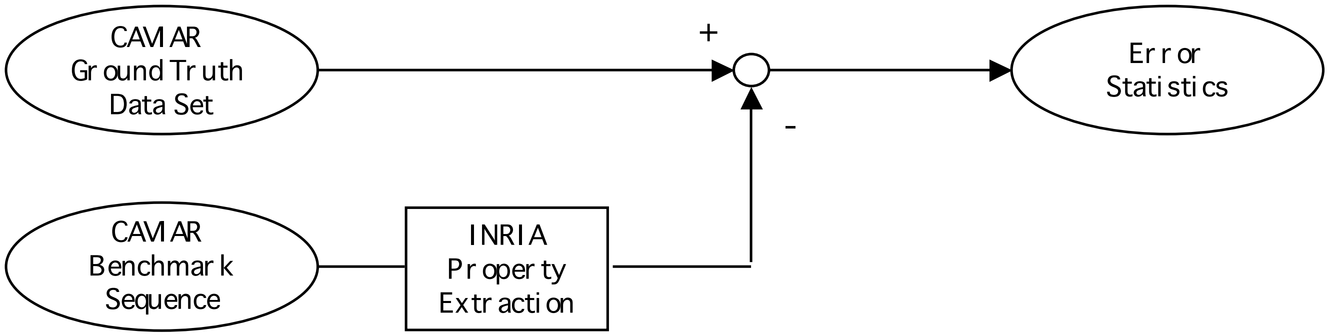
For each target, robust tracking provides

- Image Position, Height and Width
- Velocity (Speed, Direction).
- Position of Entry (target creation) into image
- Motion energy (or activity)

Motion energy is computed as SAD within ROI's

$$En(ROI) = \frac{1}{3NM} \sum_{x,y \in ROI} \sum_{c \in r,g,b} | I^t(x,y,c) - I^{t-1}(x,y,c) |$$

D22/24
Report and Software on Property Observation

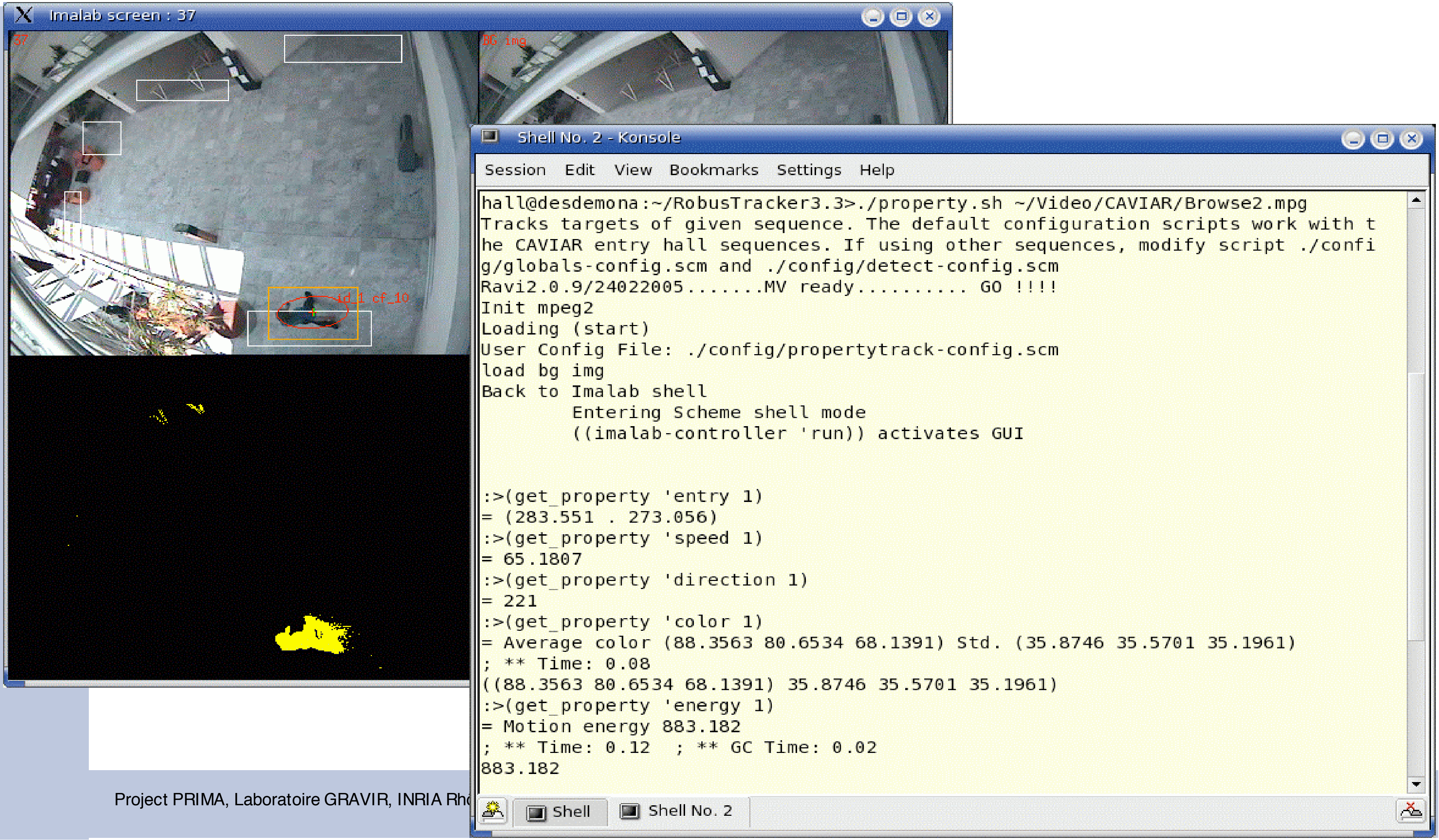


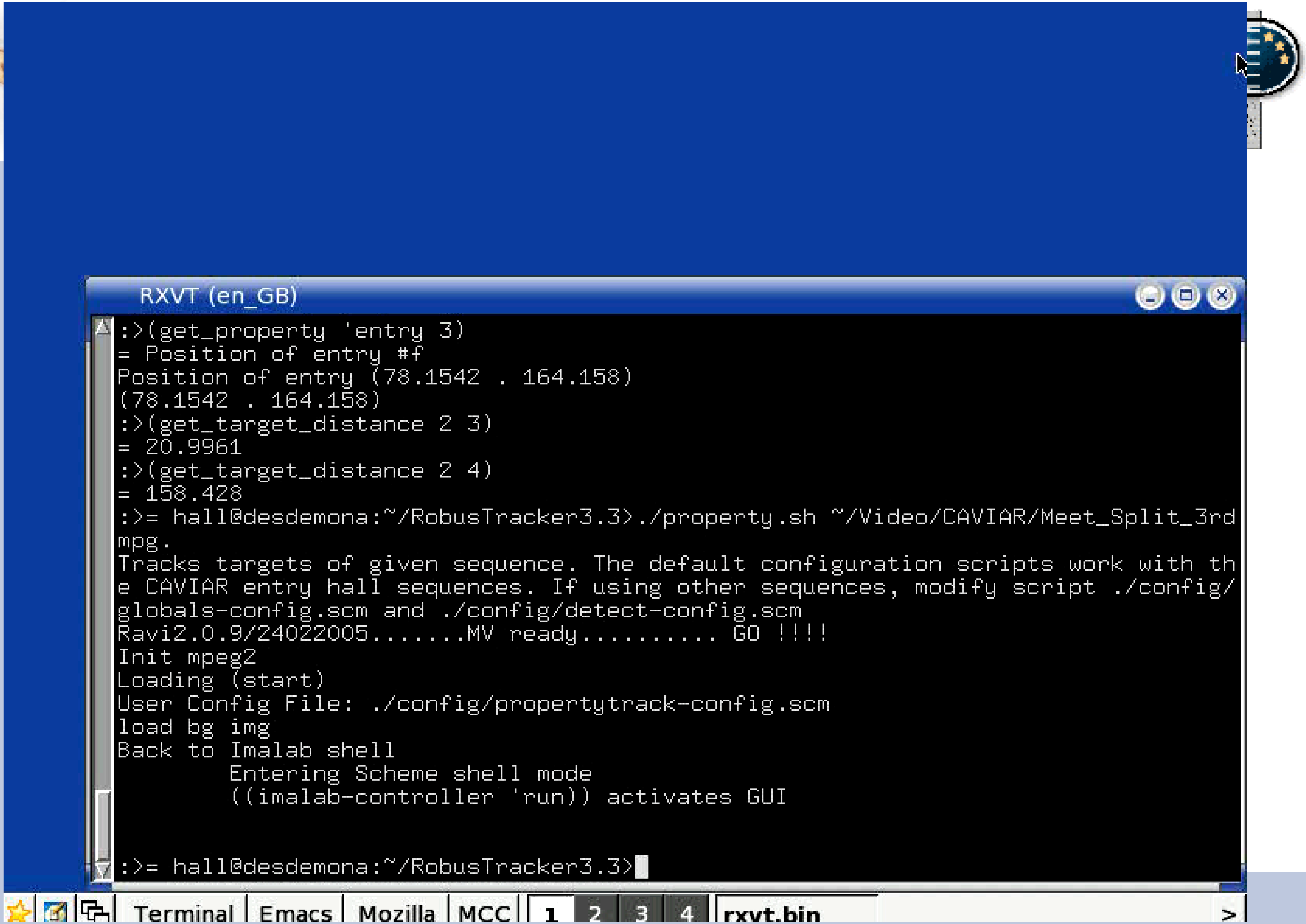
Attributes for tracked entities: Position, Speed, Direction, Activity Level

<u>Experiment</u>:

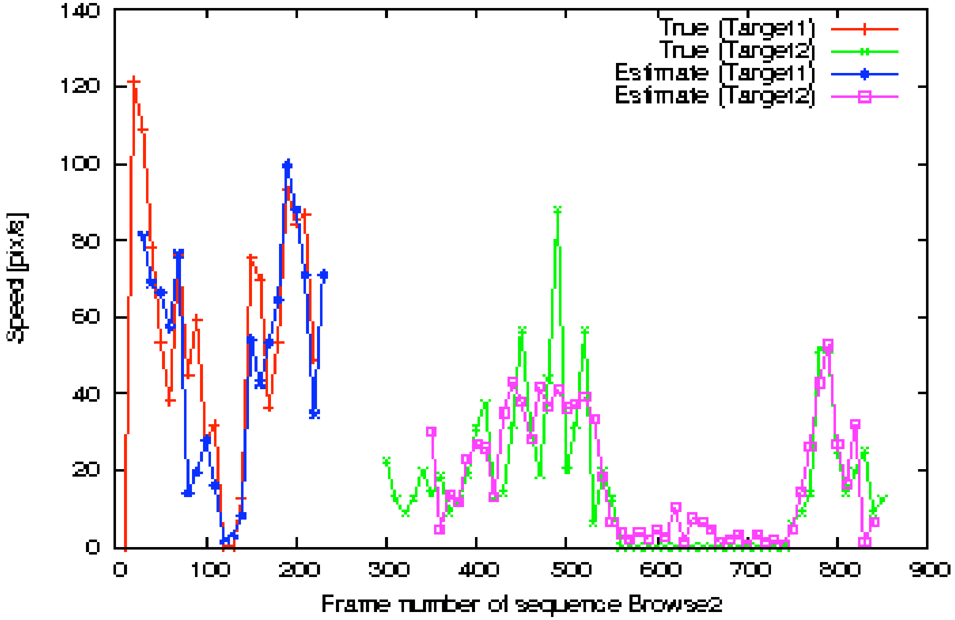
Compare observed values with ground truth in hand-labeled benchmark sequences.

D22/24
Report and Software on Property Observation




RXVT (en_GB)
:>(get_property 'entry 3)
= Position of entry #f
Position of entry (78.1542 . 164.158)
(78.1542 . 164.158)
:>(get_target_distance 2 3)
= 20.9961
:>(get_target_distance 2 4)
= 158.428
:>= hall@desdemona:~/RobusTracker3.3>./property.sh ~/Video/CAVIAR/Meet_Split_3rd
mpg.
Tracks targets of given sequence. The default configuration scripts work with th
e CAVIAR entry hall sequences. If using other sequences, modify script ./config/
globals-config.scm and ./config/detect-config.scm
Ravi2.0.9/24022005.......MV ready.......... GO !!!!
Init mpeg2
Loading (start)
User Config File: ./config/propertytrack-config.scm
load bg img
Back to Imalab shell
Entering Scheme shell mode
((imalab-controller 'run)) activates GUI
:>= hall@desdemona:~/RobusTracker3.3>
Terminal
Emacs
Mozilla
MCC
1
2
3
4
rxvt.bin

Evaluation of property "speed"



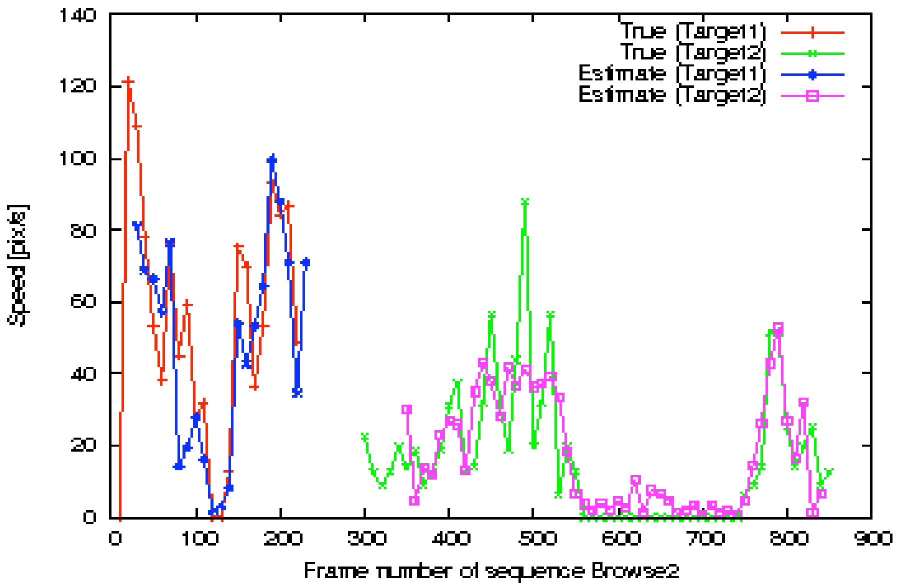
Benchmark sequence "Browse2"

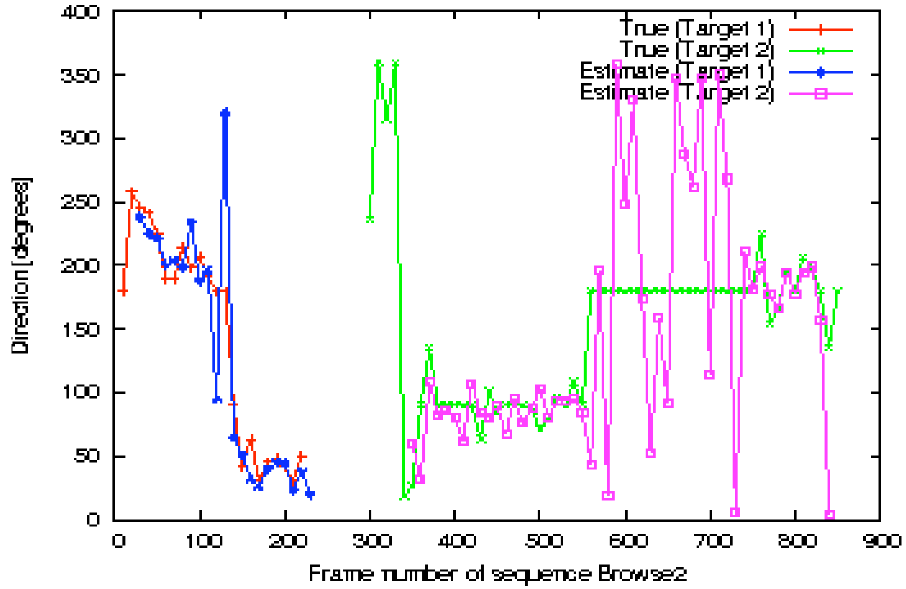
- Ground truth (red, green) is computed from a temporal window of 4 frames
- Estimated speed (blue, mauve) is smoothed by tracking.

Conclusion:

- Estimation is reliable,
- Lag is not a problem.

Evaluation of property "direction"





Benchmark sequence "Browse2"

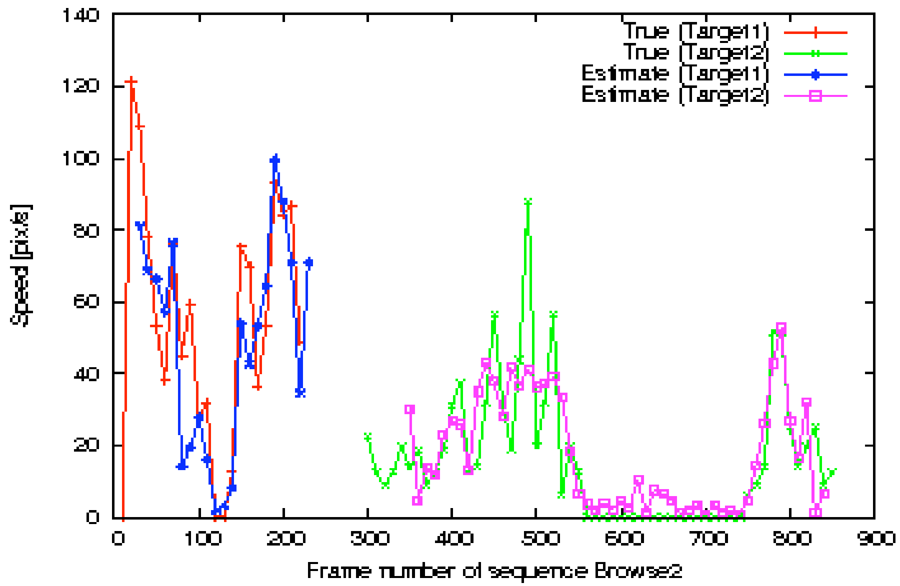
- Ground truth (red, green) is computed from a temporal window of 4 frames

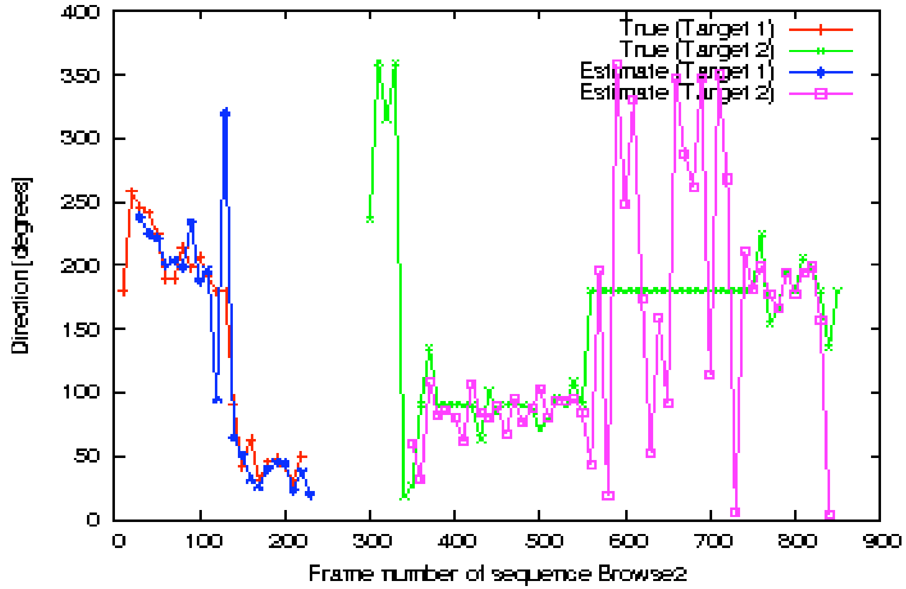
Conclusion:

- Estimated Direction (blue mauve) accurate only when speed is non-zero.

(hand labeled bounding boxes do not move for stationary target because tool for hand labeling allows copy of results from previous frame.

Evaluation of property "direction"

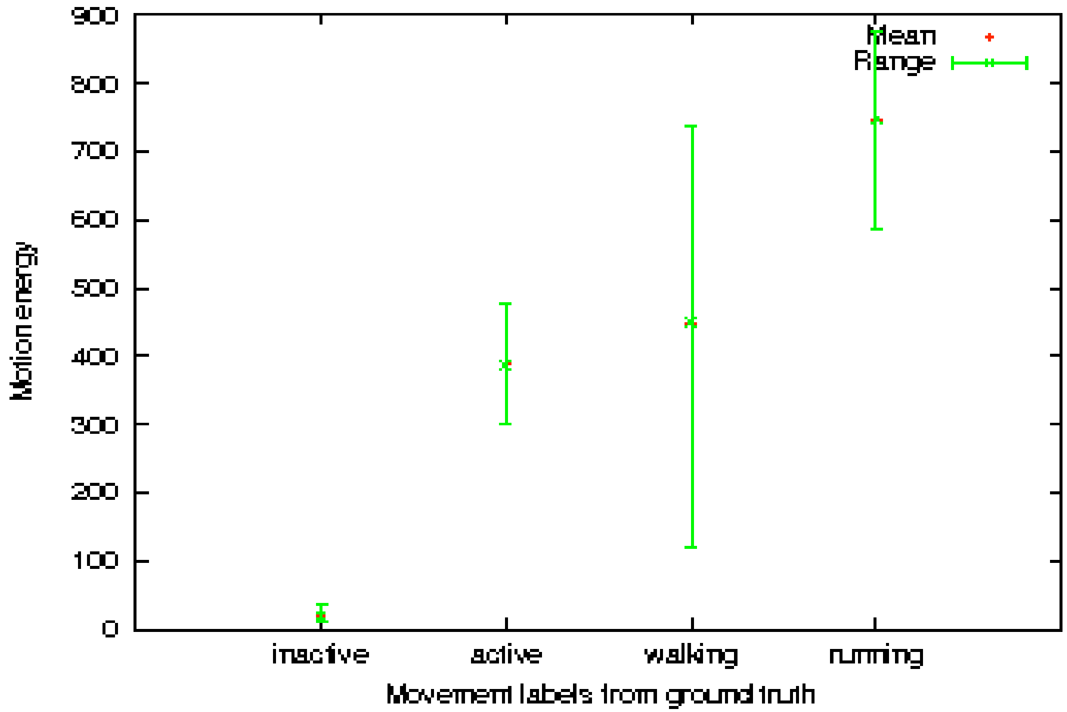




Motion direction estimate is precise when speed is larger than some threshold.

For very little speed, the motion direction is arbitrary.

Detecting Activity from Motion Energy



Labels for activity: inactive, active, walking, running

- Motion energy correlates with inactive, active and running.
- Poor Correlation for Walking (large variance in energy)

WP 4: Recognition processes

Objective:

Define, implement and compare attention driven methods for detection and observation of entities and properties

Tasks:

T4.1 Define and implement visual processes

T4.2 Define representations for object, situation and context

T4.3 Develop processes of object recognition under variation of viewpoint

T4.4 Develop processes for diagnosing and recovering from errors

Deliverables

D22/24 - Report and Software on Property Observation

D29/34 - Report and Software for Error Recovery

D31/36 - Report and Software for Object Recognition

D29/34
Report and Software for Error Recovery

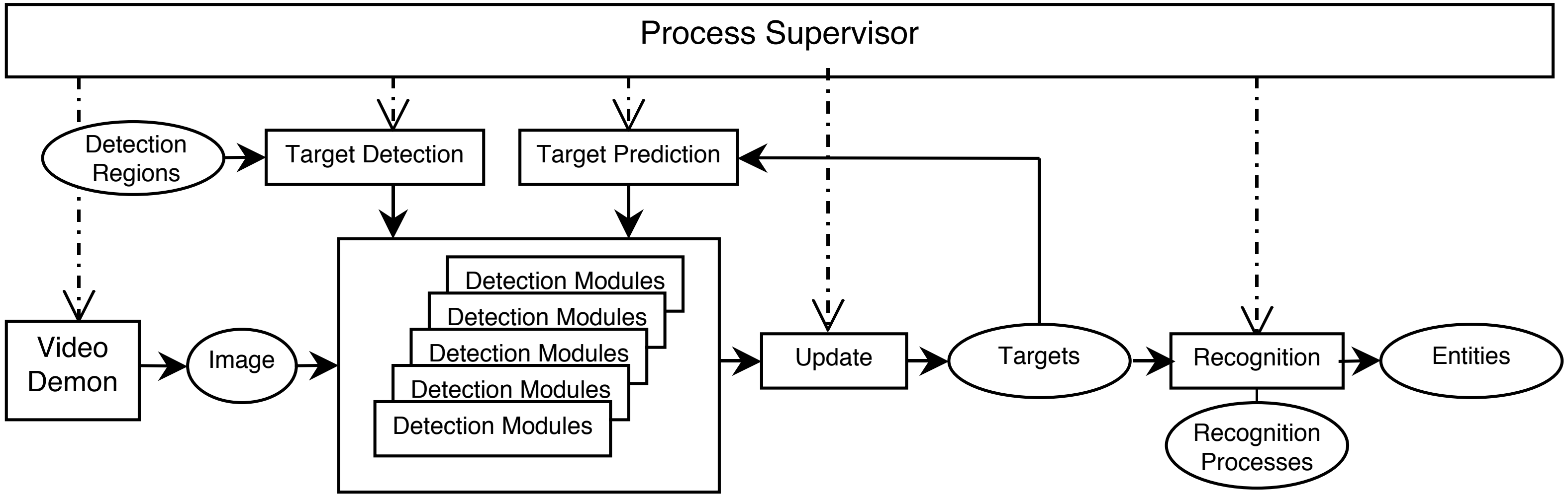
Objective:

Detect, diagnose and repair when system performance degrades.

Approach:

- Learn a model of correct system behaviour
- Monitor system output and detect errors
- Classify errors
- Repair system according to error class

Observation Processes



Pipeline of modules controlled by a supervisor

Supervisor Provides:

- Execution Scheduler
- Parameter Regulation
- Error Handling
- Command Interpreter
- Reflexive Description

Autonomic Properties

(Provided by process supervisor)

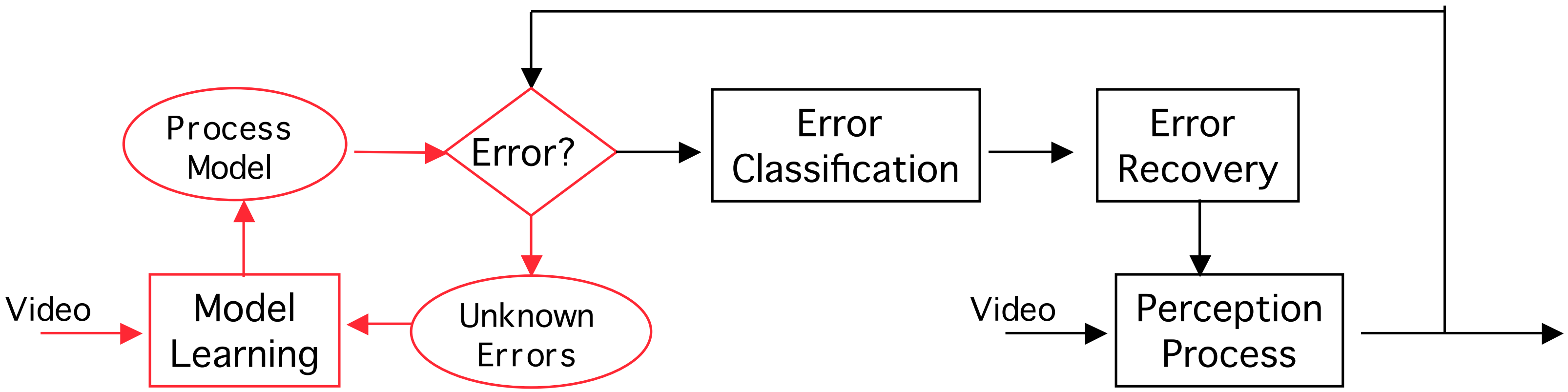
Auto-descriptive: The process controller provides descriptions of the capabilities and the current state of the process.

Auto-regulatory: The process controller can adapt parameters to maintain a desired process state.

Auto-critical: Process estimates confidence for all properties and events.

Self Monitoring: Maintaining a description of process state and quality of service

Training the Process Model



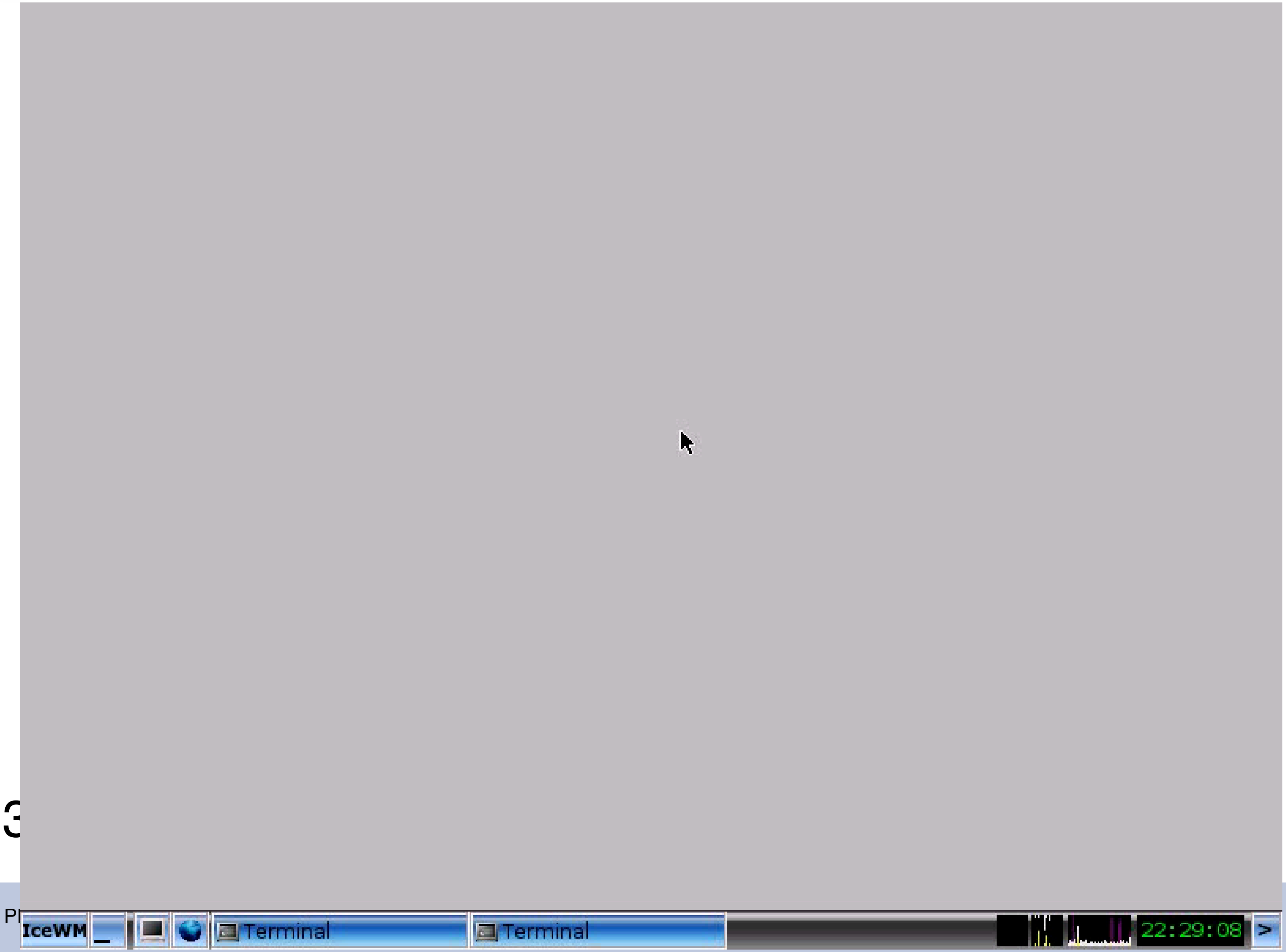
Process Model: histogram for process outputs

Semi-Supervised learning:

- User configures and launches a process.
- System classifies each frame as valid, known error, unknown.
- User validates classification, sequence stored.
- Model updated after each validation.
- Process converges after a few minutes (using CAVIAR indoor testbed).

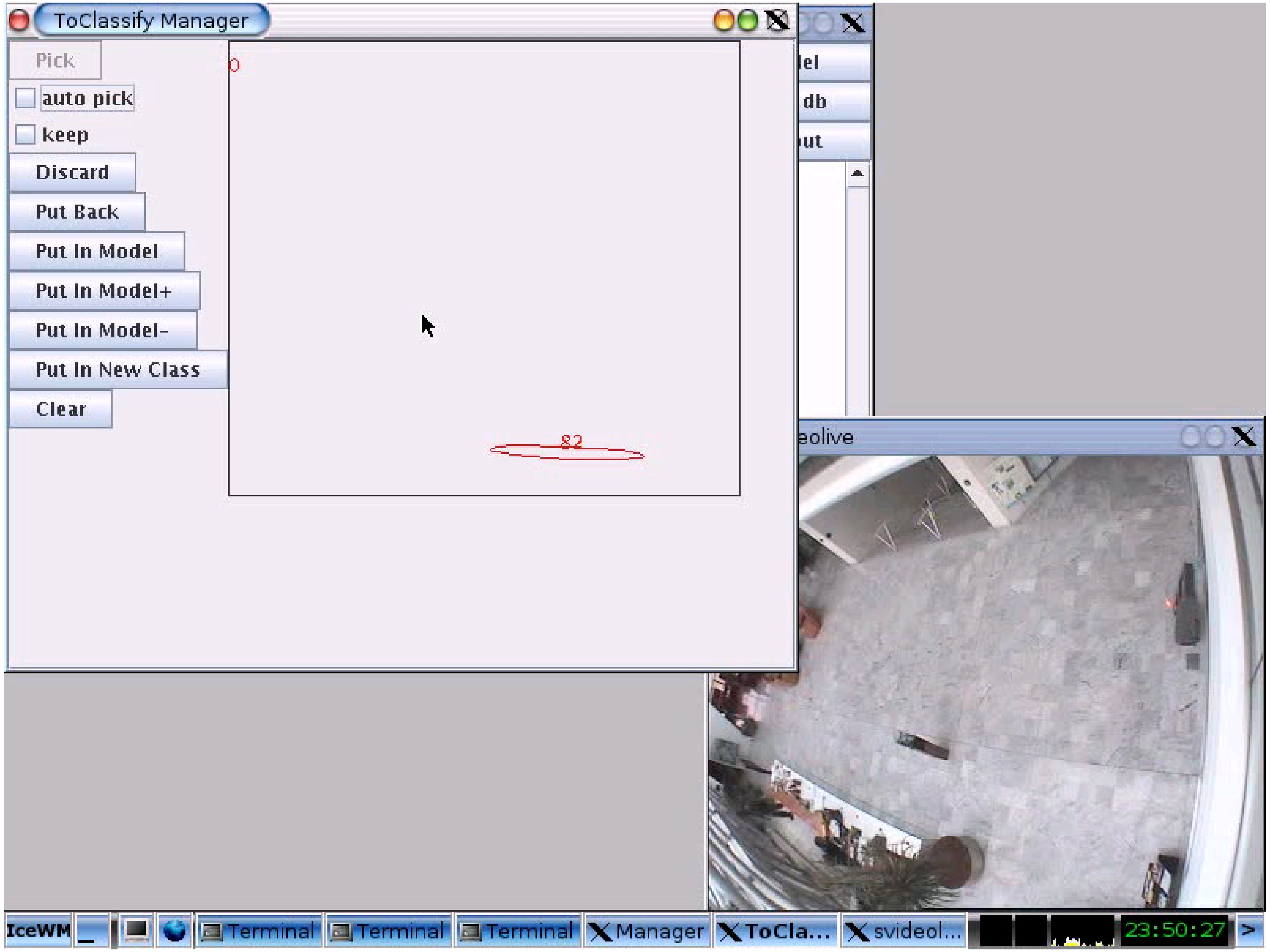
CAVIAR

Training the Process Model

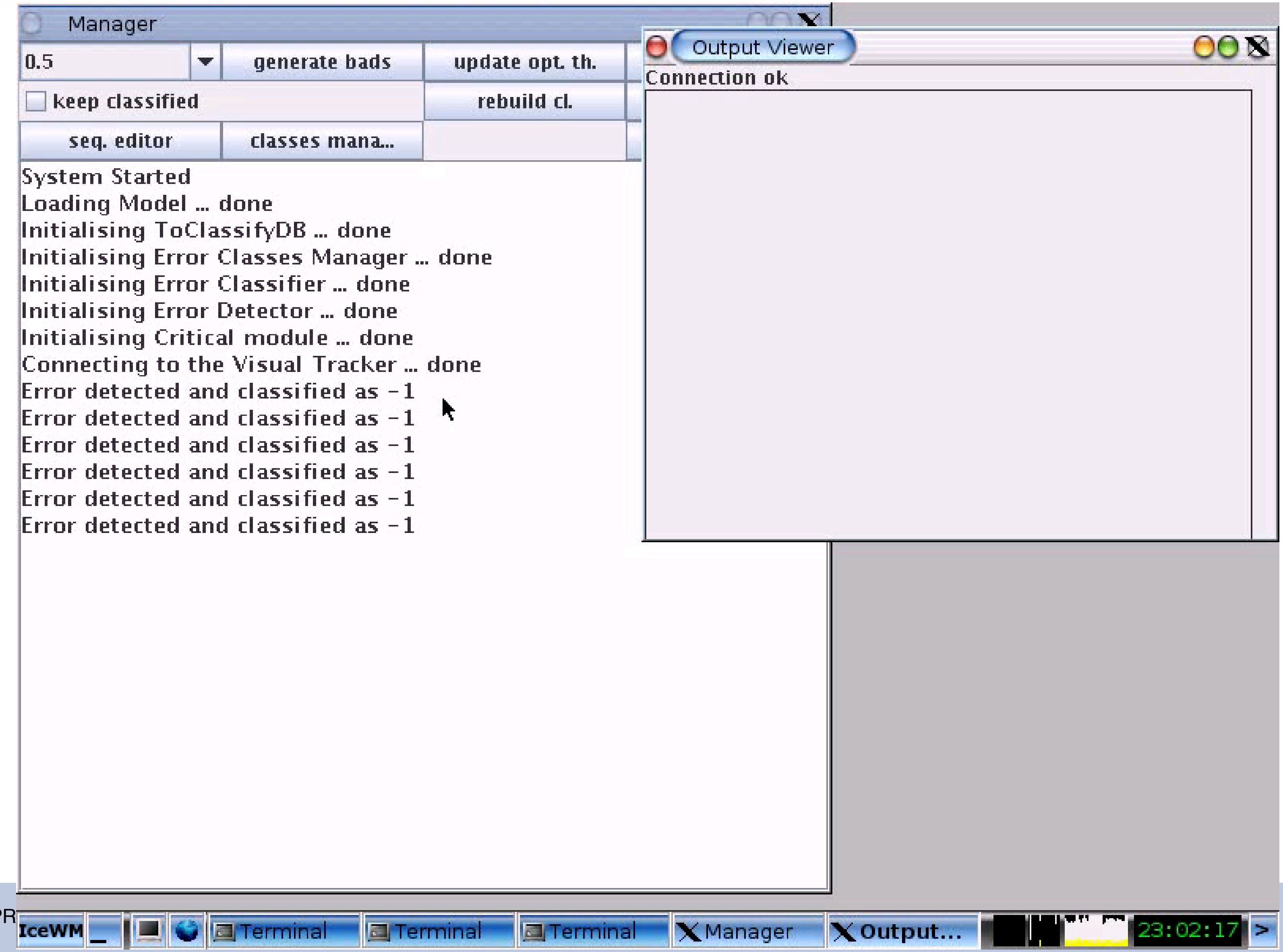


D3

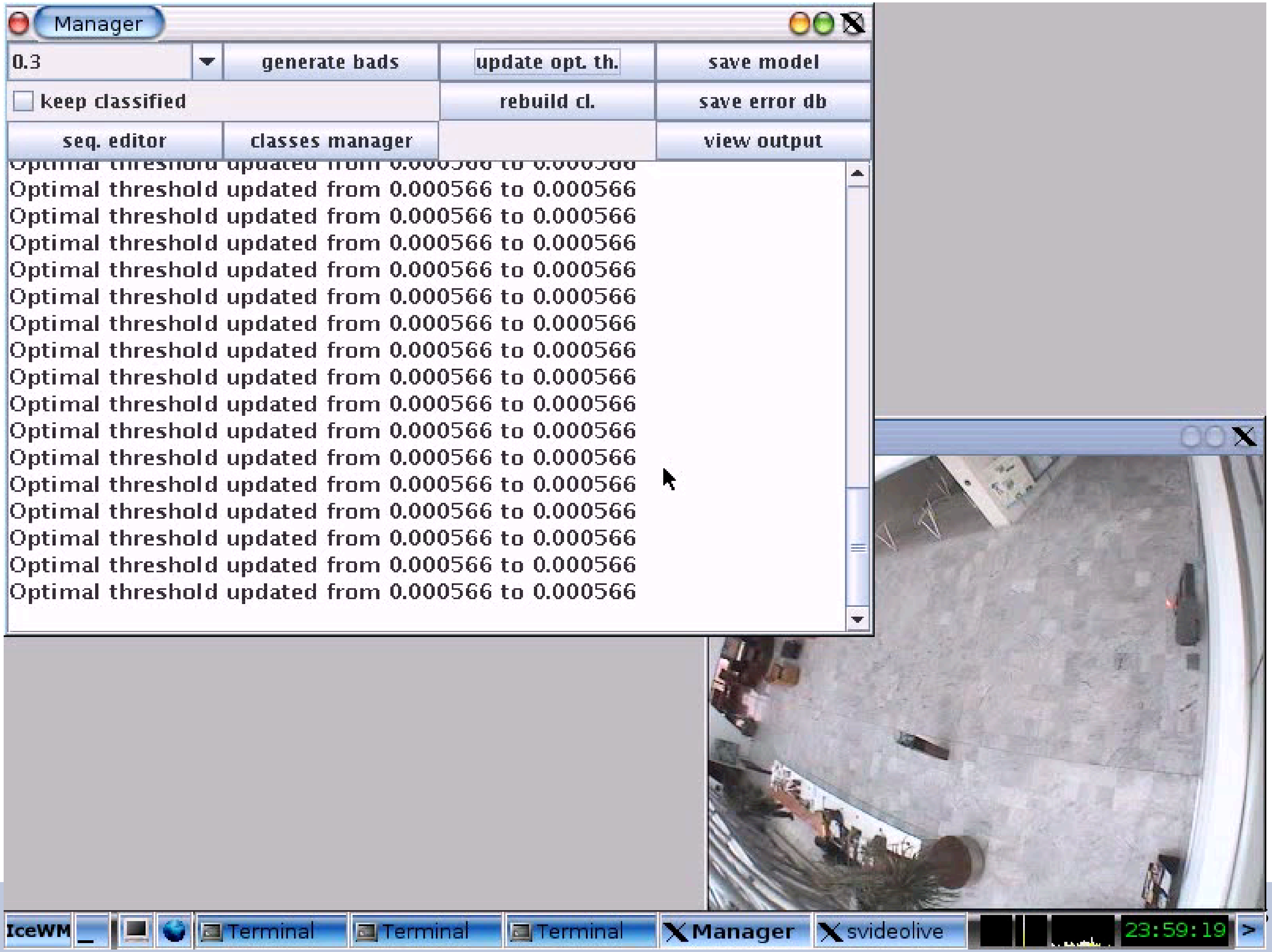
Training the Process Model



Training the Process Model

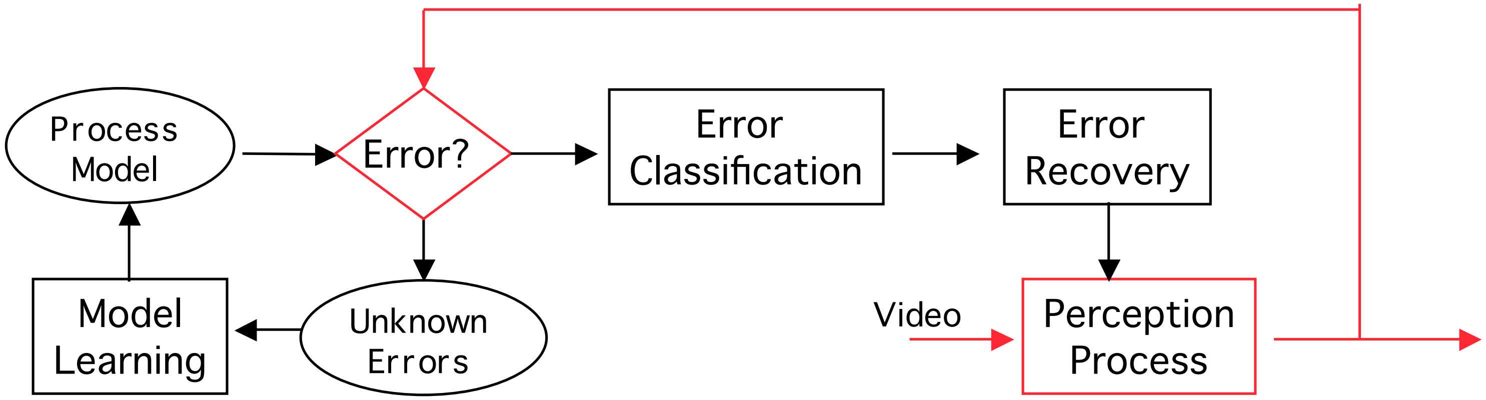


Training the Process Model



CAVIAR

Error Monitoring and Recovery



Error Detection:

Calculates likelihood from output of N most recent frames

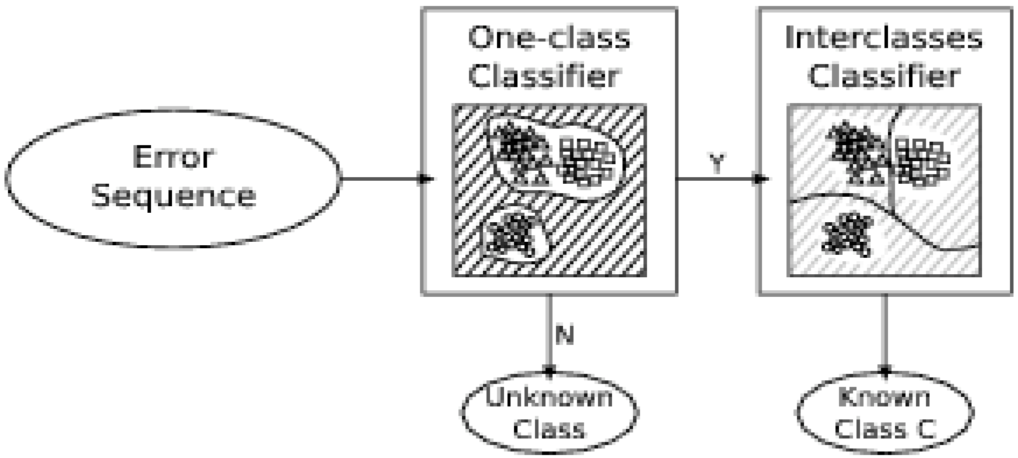
Raises event if likelihood falls below threshold for N frames

Error Classification

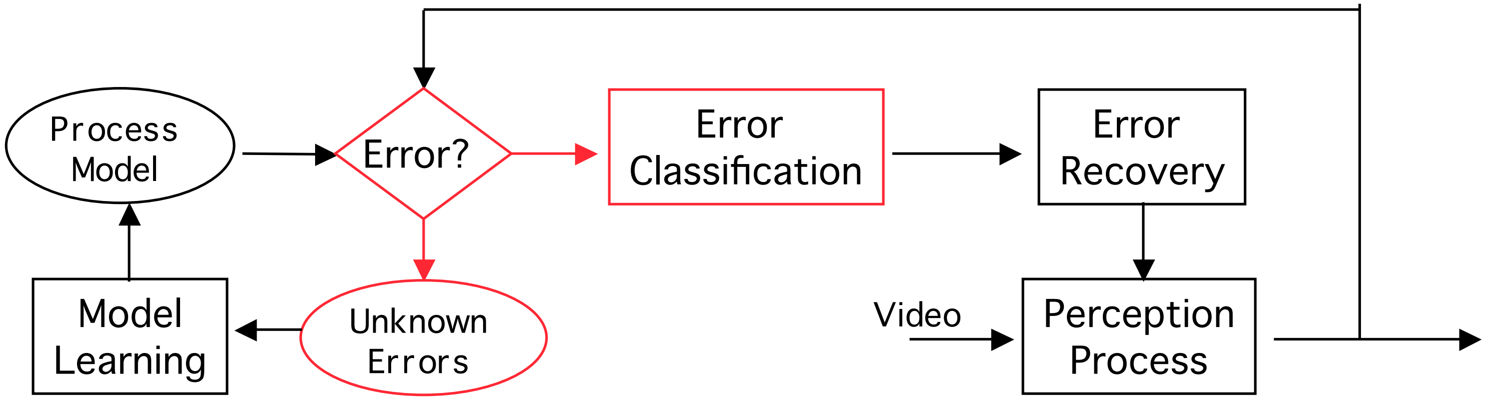
Method: Cascaded Support Vector Machines

1) Classify error as known/unknown
2) If known, assign error to class.

Output: The error sequence and class.



Error Monitoring and Recovery



Two Cases:

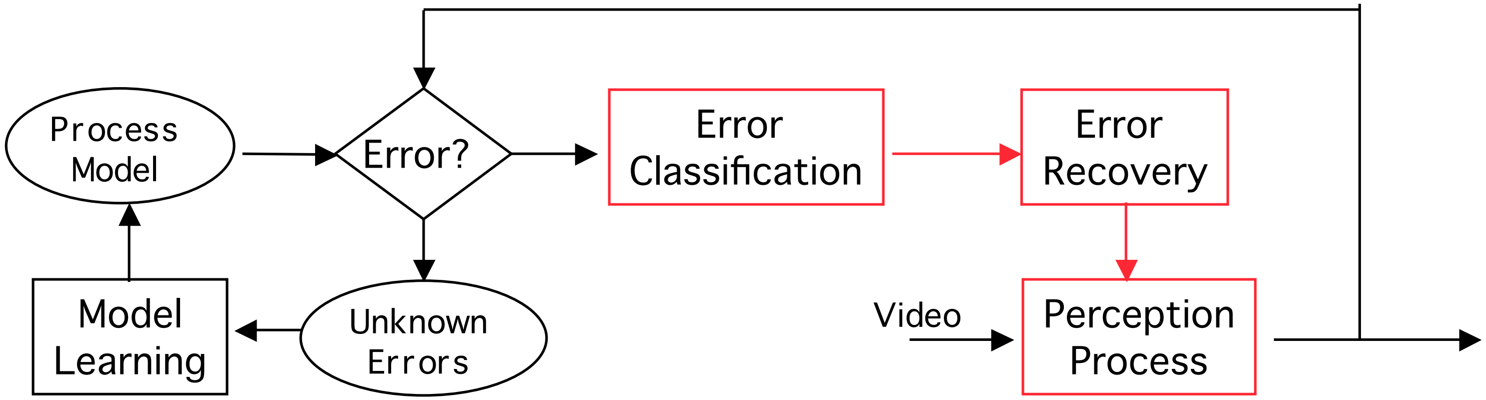
If Error labeled as a known class

- Use repair code for class to reconfigure process.

if Error labeled as *Unknown*

- Store data sequence in data base for off line learning.

Error Monitoring and Recovery

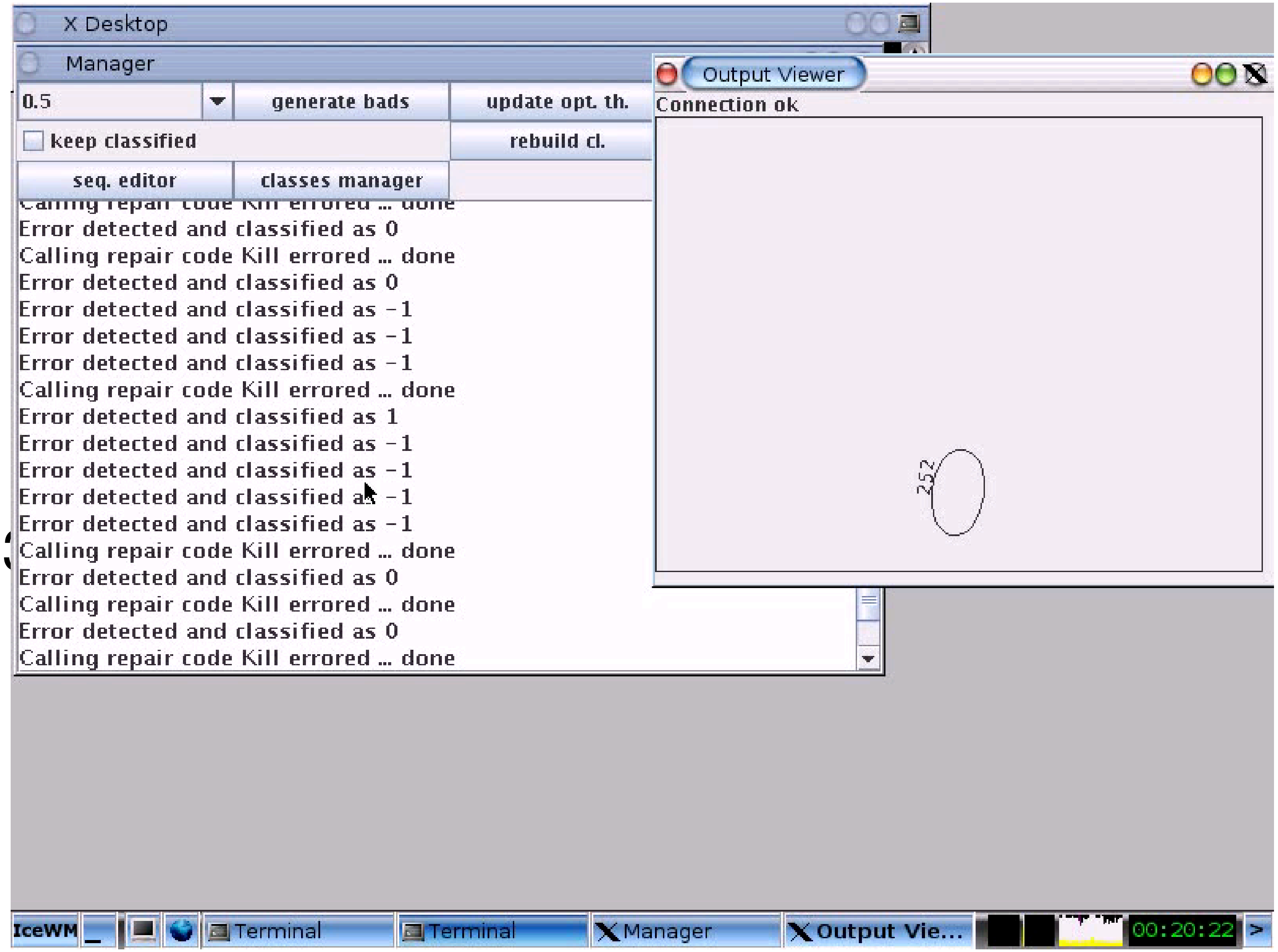


Error Recovery:

Error Class is used to select recovery procedure

(hand coded)

Error Monitoring and Recovery



D29/34
Report and Software for Error Recovery

Conclusion:

Self Monitoring allows error recovery and auto-regulation

Error Recovery:

- Provides robust system behaviour under real world conditions.

Auto-regulation (Y2):

- Provides robust, real time quality of service under variations of illumination and operating Conditions.
- Allows system installation by unskilled personnel.

WP 4: Recognition processes

Objective:

Define, implement and compare attention driven methods for detection and observation of entities and properties

Tasks:

T4.1 Define and implement visual processes

T4.2 Define representations for object, situation and context

T4.3 Develop processes of object recognition under variation of viewpoint

T4.4 Develop processes for diagnosing and recovering from errors

Deliverables

D22/24 - Report and Software on Property Observation

D29/34 - Report and Software for Error Recovery

D31/36 - Report and Software for Object Recognition

D31/36 - Report and Software for Object Recognition under Variations of Viewpoint

Problem:

Recognize target under variations of view direction and illumination conditions.

Approach:

- Scale and orientation invariance from normalised receptive fields
- Comparison of Statistical and Structural approachs

Structural Approach:

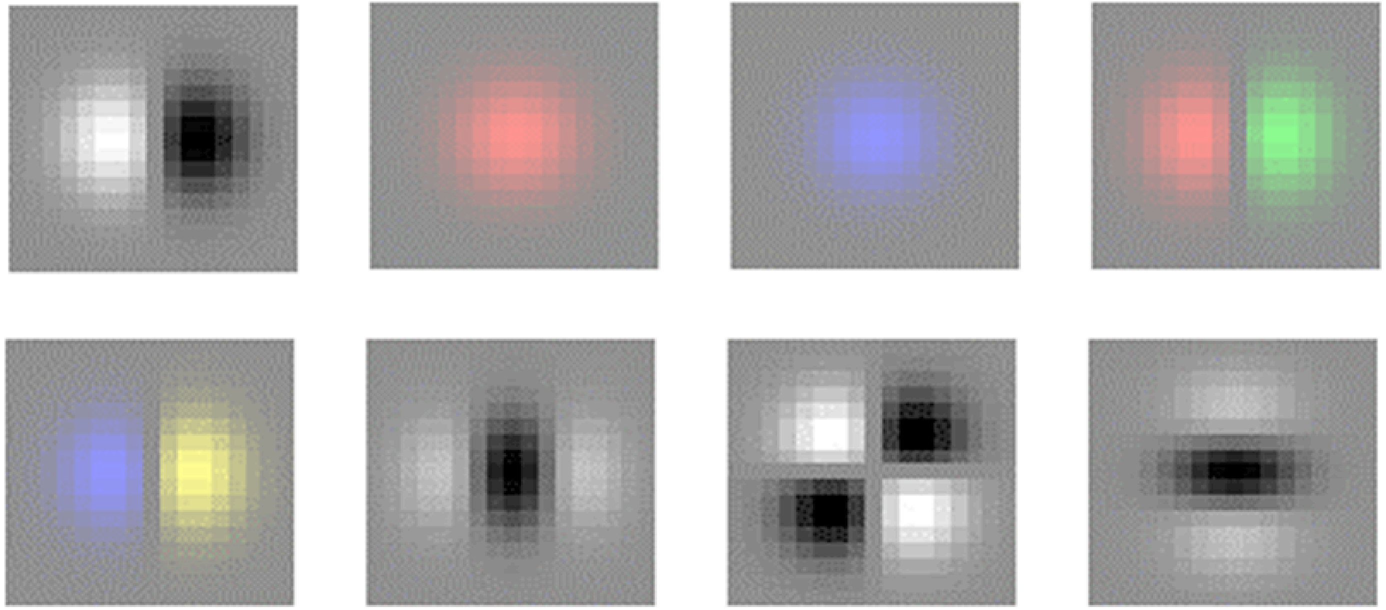
- Natural interest points and Natural interest "ridges".

Statistical:

- SIFT, Ridge normalised SIFT
- Auto-Associative Classifiers

CAVIAR

Chromatic Gaussian Receptive Fields



$$\vec{\varphi} = \begin{bmatrix} G_x^L \\ G^{C_1} \\ G^{C_2} \\ G_x^{C_1} \\ G_x^{C_2} \\ G_{xx}^L \\ G_{xy}^L \\ G_{yy}^L \end{bmatrix}$$

Normalized in scale and orientation of local neighborhood
Real Time calculation with binomial pyramid.

CAVIAR

Scale Normalised Gaussian RF's

The normalisation to intrinsic scale provides scale invariant description of appearance

$$\sigma_i(i,j) = \underset{\sigma}{\mathrm{Arg-Max}}\{ < \nabla^2 G(\sigma) \cdot A(i,j) > \}$$

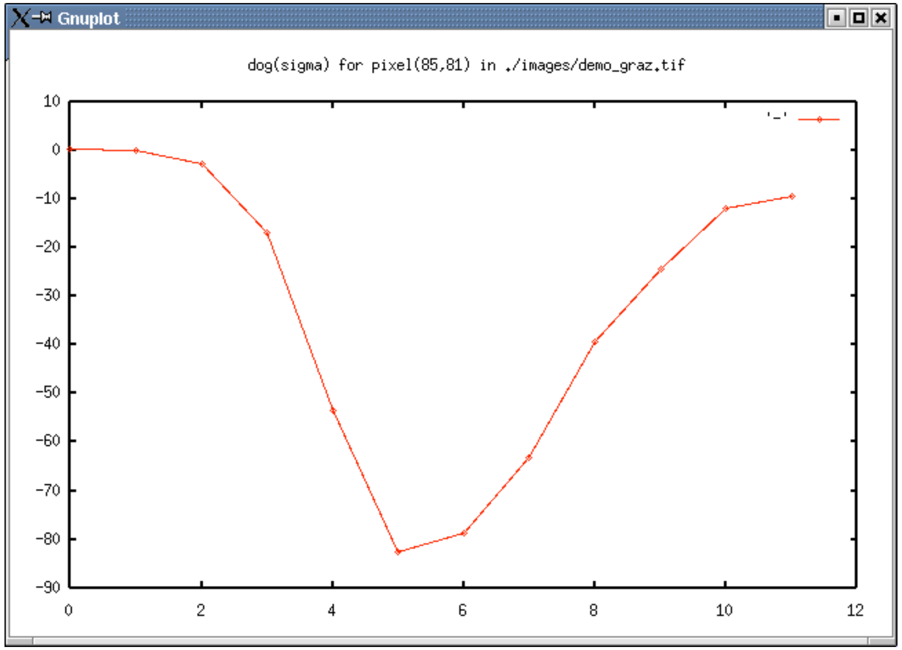
Natural Interest Points

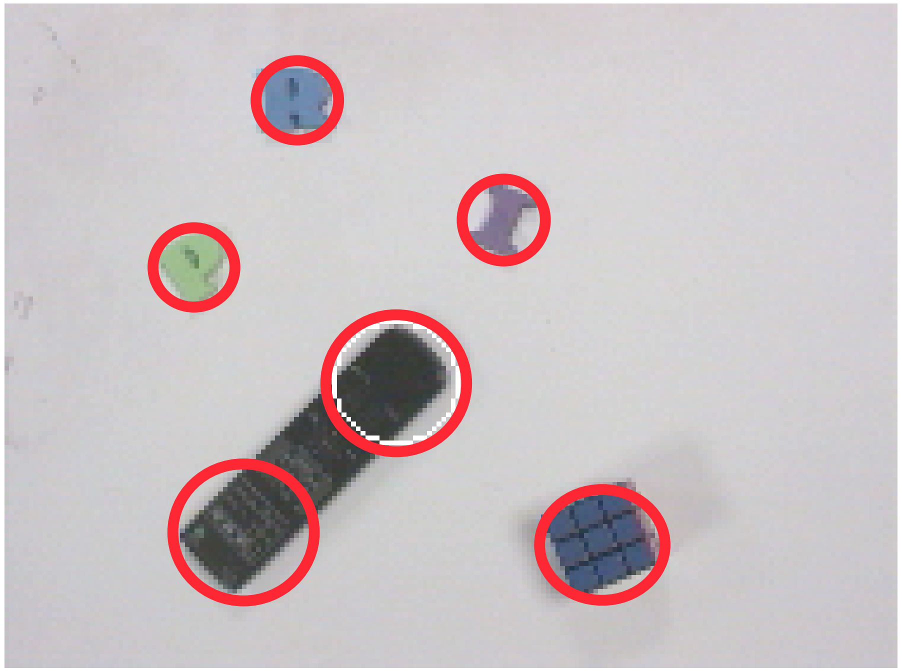
(Scale Invariant "Salient" image features)

Natural Interest points:

Local extrema of $<\nabla^2 G(i,j,s) \cdot A(i,j)>$ over i, j, s

(Current state of the art in "salient image features")





Natural Interest Points

(Scale Invariant "Salient" image features)

Local extrema of $<\nabla^2 G(i,j,s) \bullet A(i,j)>$ over i, j, s

Problems with Points

- Elongated shapes
- Lack of discrimination power
- No orientation information

Proposal: Natural Interest Ridges

Maximal ridges in Laplacian Scale Space:

$<\nabla^2 G(i,j,s) \bullet A(i,j)>$ over i, j, s

Natural Ridge Detection [Tran04]

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

Laplacian

$$\mathcal{H} = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial xy} \\ \frac{\partial^2 f}{\partial xy} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

Hessian

Compute Derivatives at different Scales.

For each point (x,y,scale)

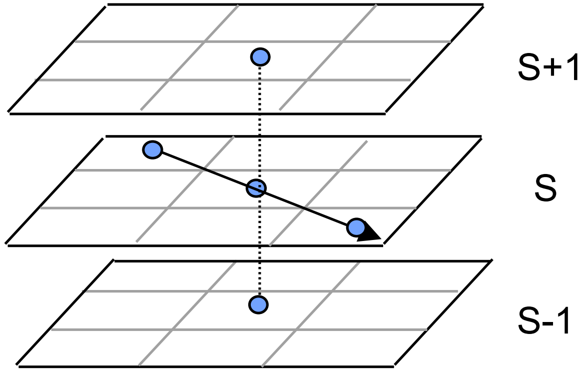
- Compute second derivatives: f_{xx}, f_{yy}, f_{xy}
- Compute eigenvalues and eigenvectors of Hessian matrice:
- Detect local extremum in the direction corresponding to the largest eigenvalue.
- Assemble ridge points,

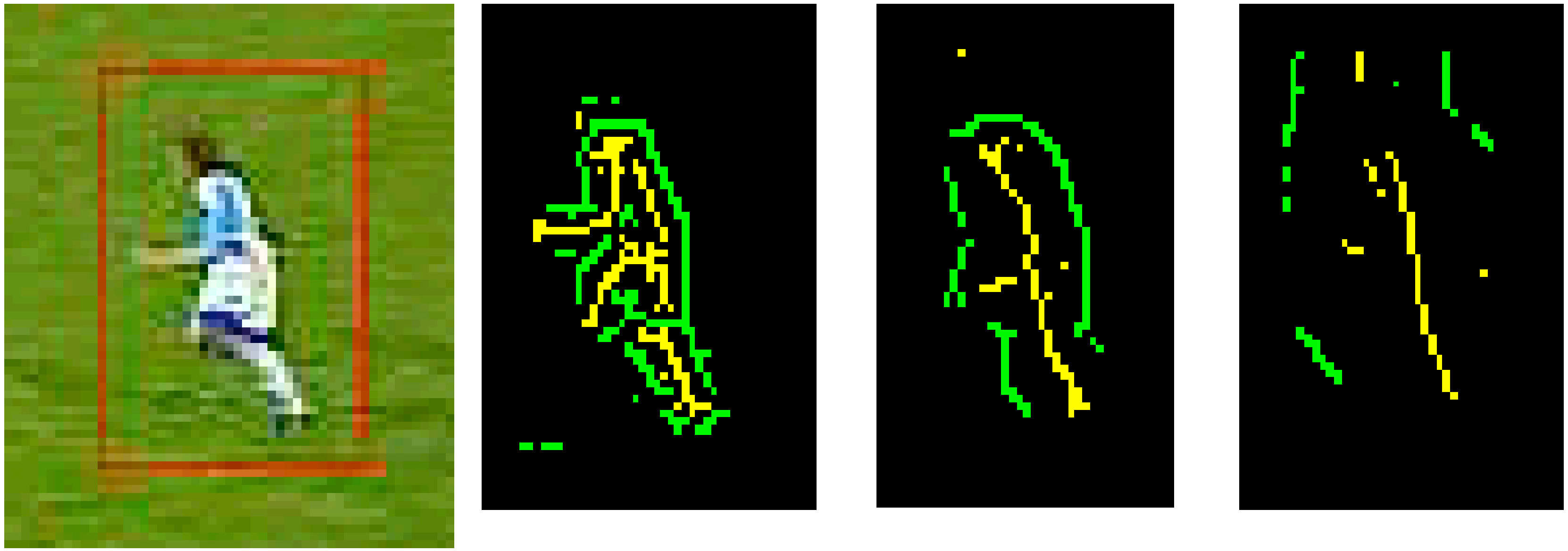
Ridge Detection and Linking

- For each ridge point (x,y,s) at scale s:
- If: Lap(x,y,s) > threshold

  Lap(x,y,s-1)< Lap(x,y,s) > Lap(x,y,s+1)

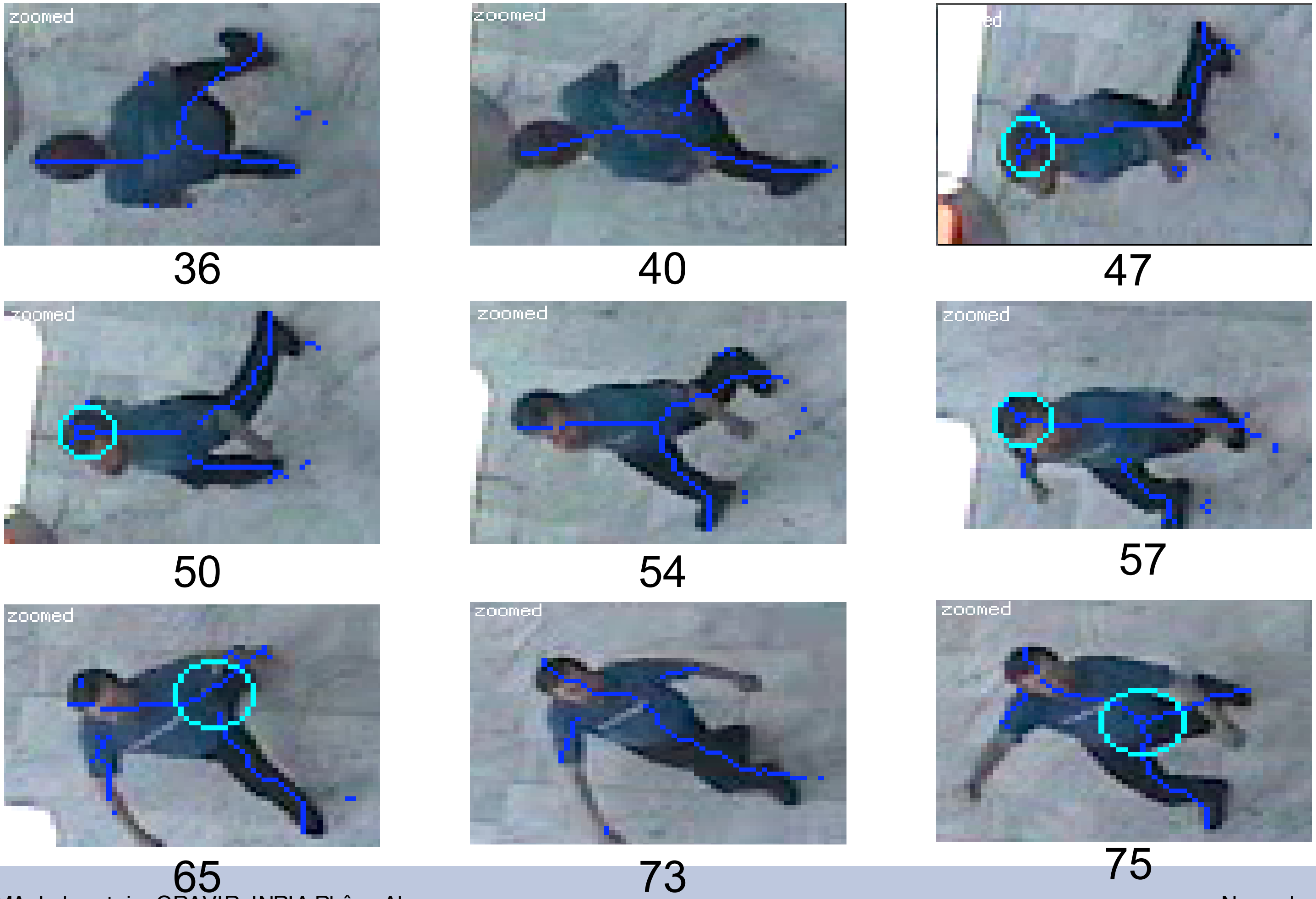
Then note as ridge point





Structural Modelling of Humans using Ridges

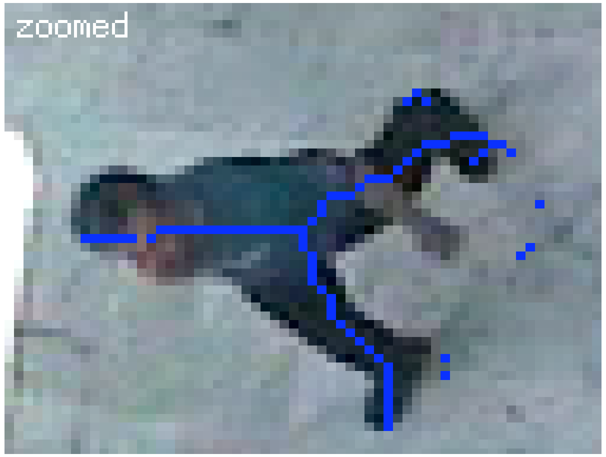
Human form can be represented by three dominant ridges:

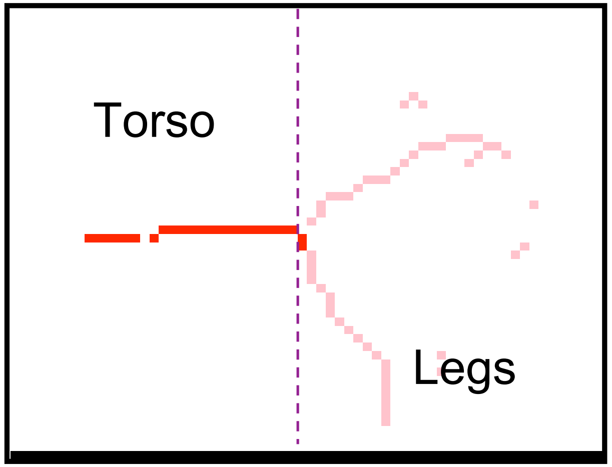


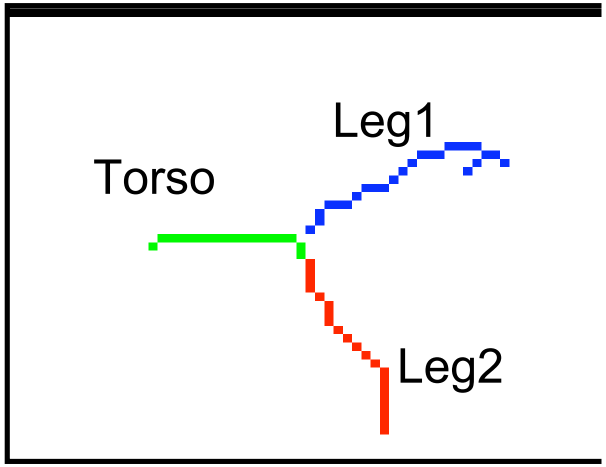
Determining ridges: Torso, Legs

Dominant Ridge: Torso ridge

- Ridge at largest scale.
- Tests : Length to scale ratio, direction (scene vertical)
  - (Scene vertical observed at set-up)





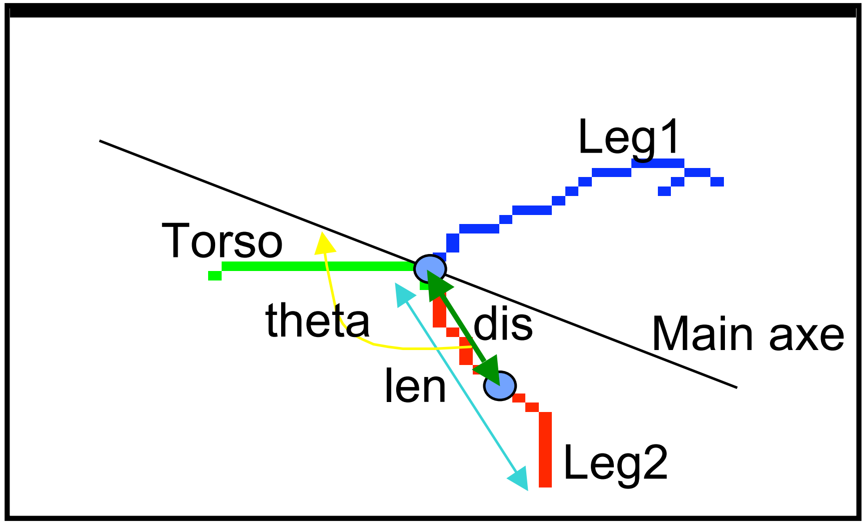


Structural model for Human Form

A human: Feature vector of 10 dimensions

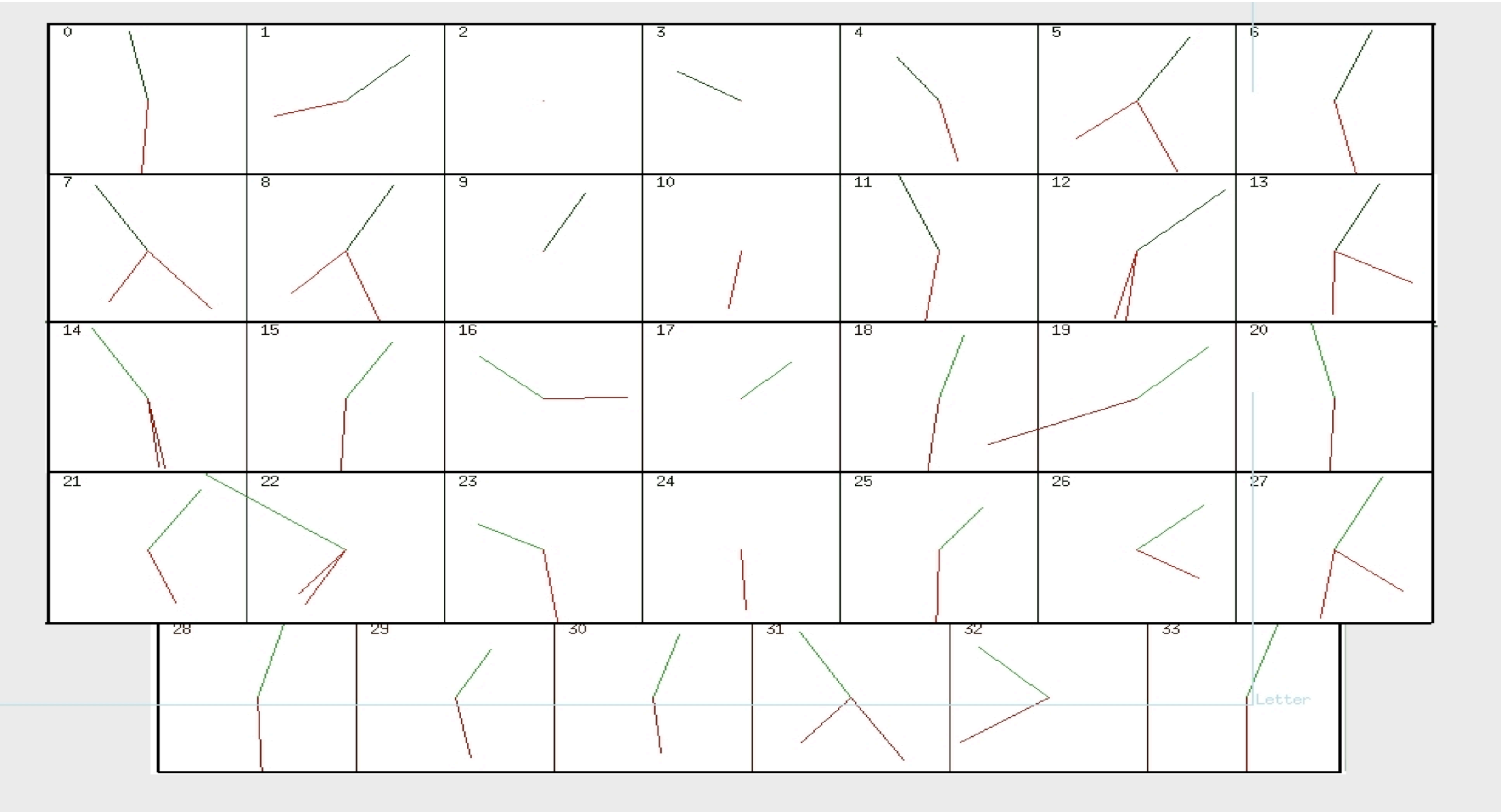
(N_{ridges},theta1,len1,dis1,theta2,len2,dis2,theta3,len3,dis3)

Features are normalized by scale of human



Learning articulated models

Human models are learned from 12 video sequences in the hall of INRIA by using K-Means (34 clusters)



Example Results: Browsing

Task: Classify Imagette
Class 1: Contains one or more Human (Green)
Class 2: Does not contain a human (Red)

Example Results: Browsing

Task: Classify Imagette
Class 1: Contains one or more Human (Green)
Class 2: Does not contain a human (Red)

Example Results: Browsing

Task: Classify Imagette

Class 1: Contains one or more Human (Green)

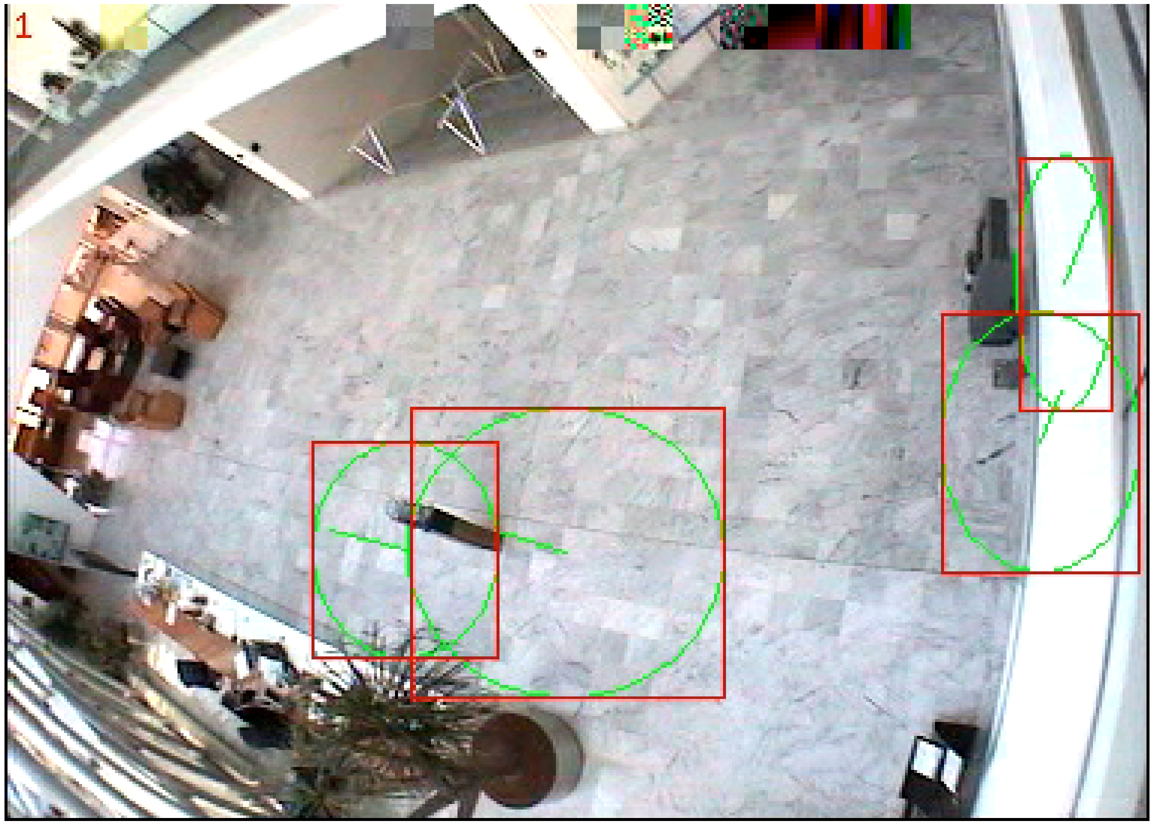
Class 2: Does not contain a human (Red)

Example Results: Random ROIs

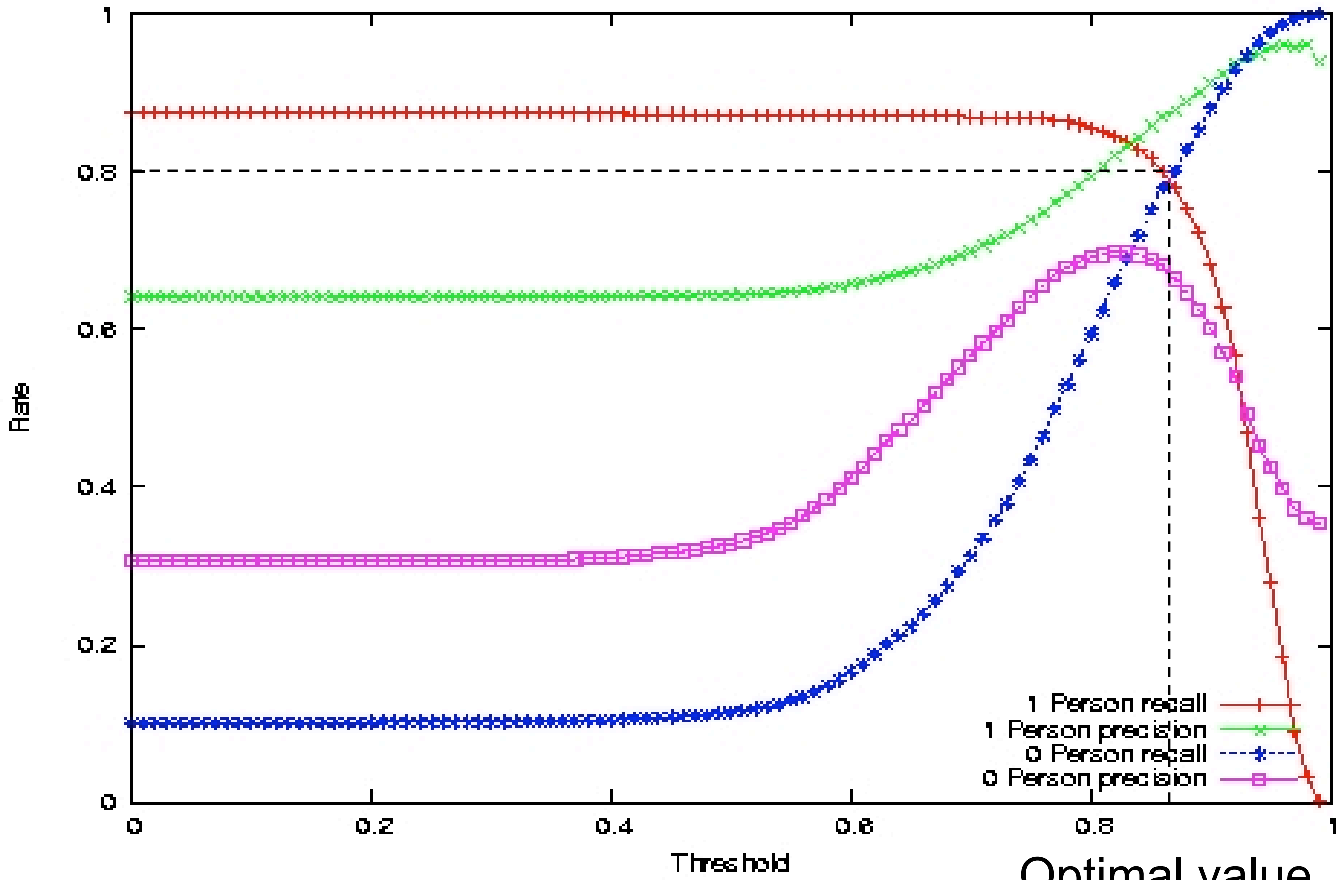
Task: Classify Imagette

Class 1: Contains one or more Human (Green)

Class 2: Does not contain a human (Red)



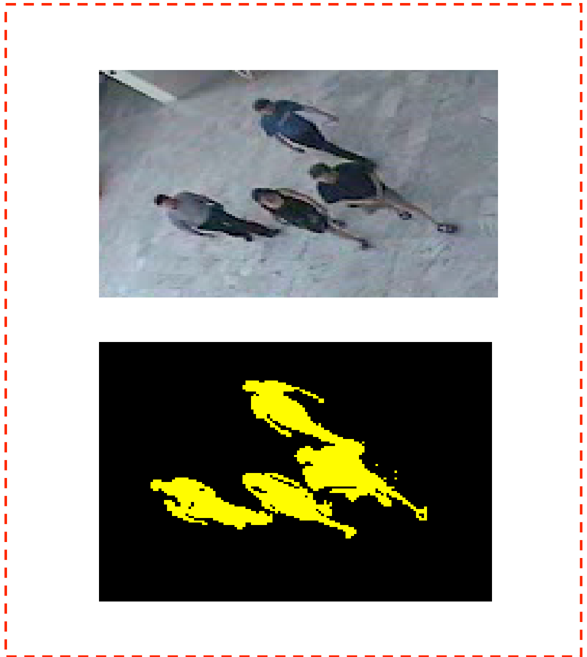
Recognition rates for human class vs matching threshold

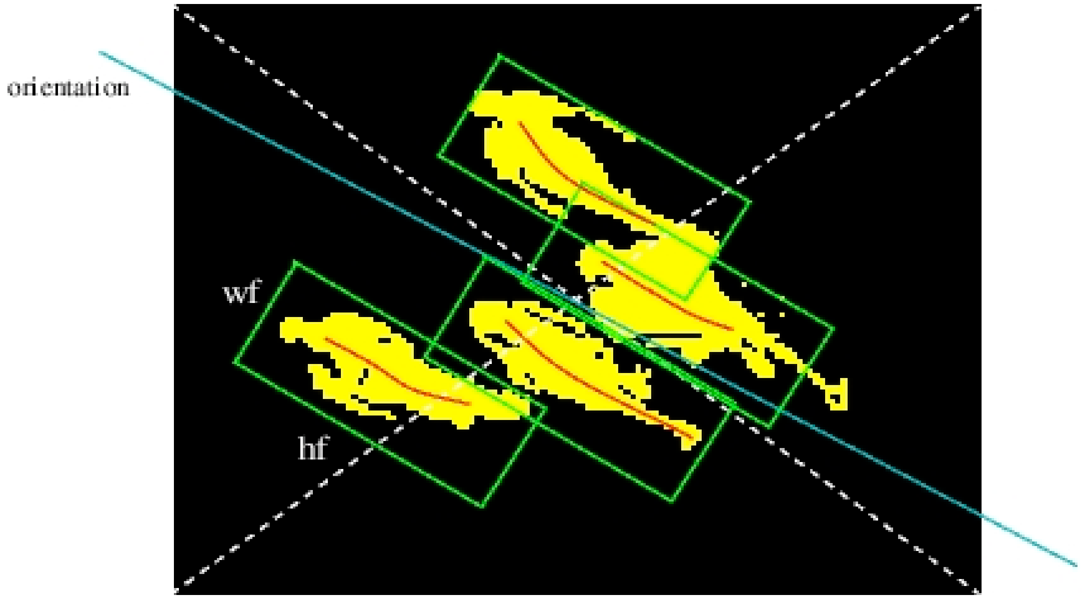


Optimal value

Estimating number of people in groups by using majors ridges

Counting the number of people in a group using number of major ridges.





Early results: only 51.3 % of Imagettes correctly classified in first experiment

Statistical Recognition of Imagettes

Objectives :

- Classification of imagettes based on appearance.
- Invariant to viewpoint (affine transformations)
- Robust to illumination changes

Approach:

- Gradient histograms normalized by ridge parameters

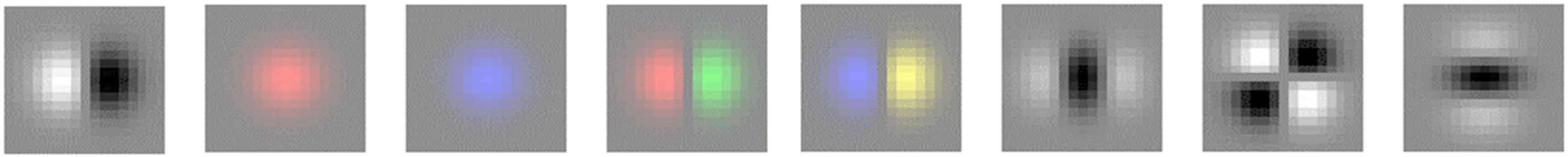
Inspired by:

- SIFT descriptor [Lowe04] and
- Gaussian Receptive Field Histograms [Schiele00]

Method:

- Estimation of a local reference using ridges
- Normalize scale and orientation of gradient
- Compute histograms for 4 quadrants
- Compare to model for human(s)

Gaussian RF's are Steerable in Orientation



$$< G_x^{\theta} \cdot A(i,j) > = < G_x \cdot A(i,j) > \mathrm{Cos}(\theta) + < G_y \cdot A(i,j) > \mathrm{Sin}(\theta)$$

Problem: How to determine orientation
Classic Approach:

$$\theta_i(i,j) = \mathrm{Tan}^{-1}\left(\frac{< G_y \cdot A(i,j) >}{< G_x \cdot A(i,j) >}\right)$$

Alternative: Direction of dominant ridge

Ridge Description

Ridge Linking:

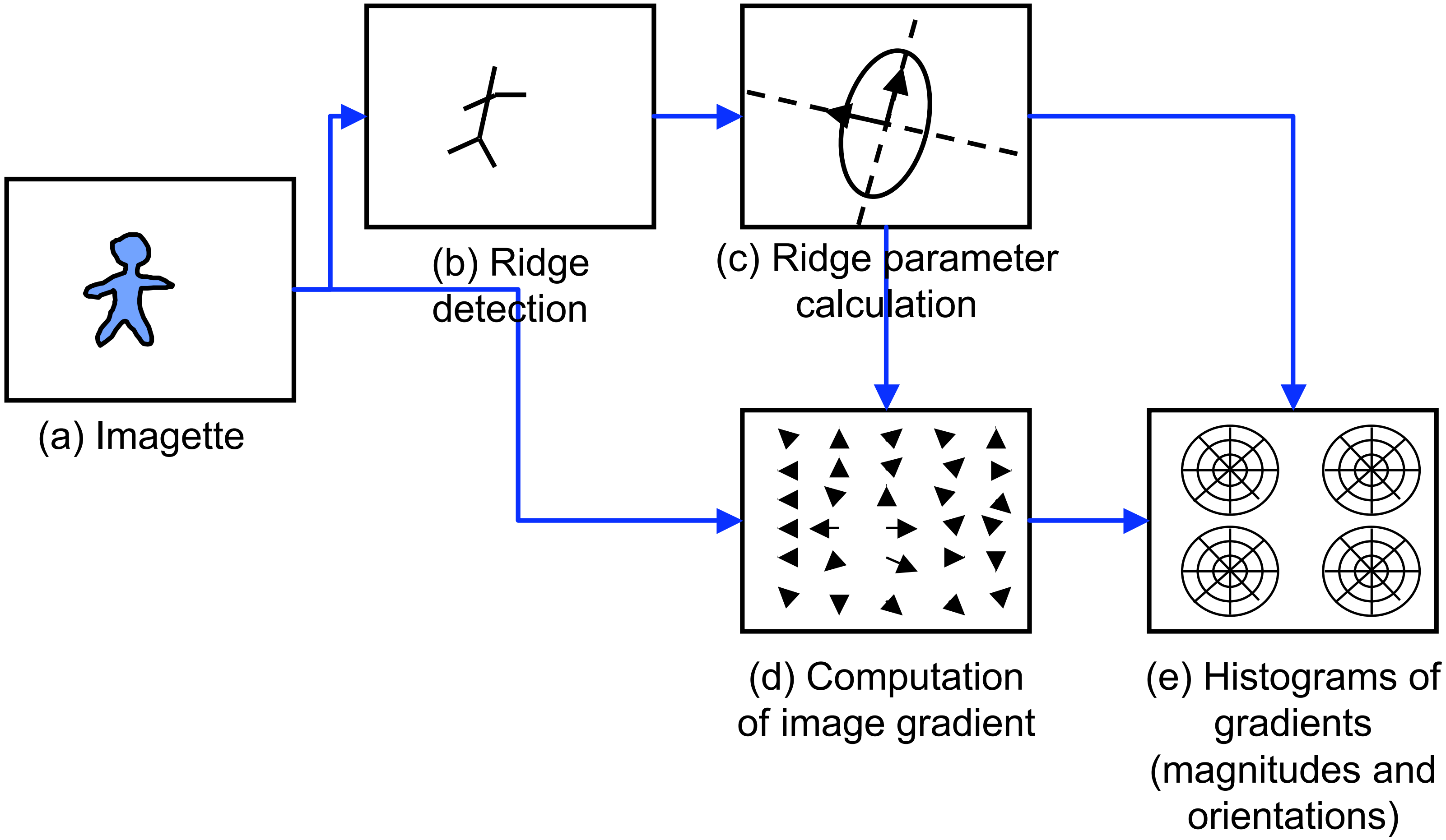
- Ridge detection at individual scales
- Group overlapping ridges across scales to form 3d blob in (i, j, s).

Ridge feature estimation

- Position, Scale: first moments (C.o.G.) of ridge blob
- Length: Major axis of second moment
- Orientation: Angle of second moment.

Ridge parameter estimation

SIFT Statistical Classification



Experimental results

Two classes ("1person" and "0person")

Training database:

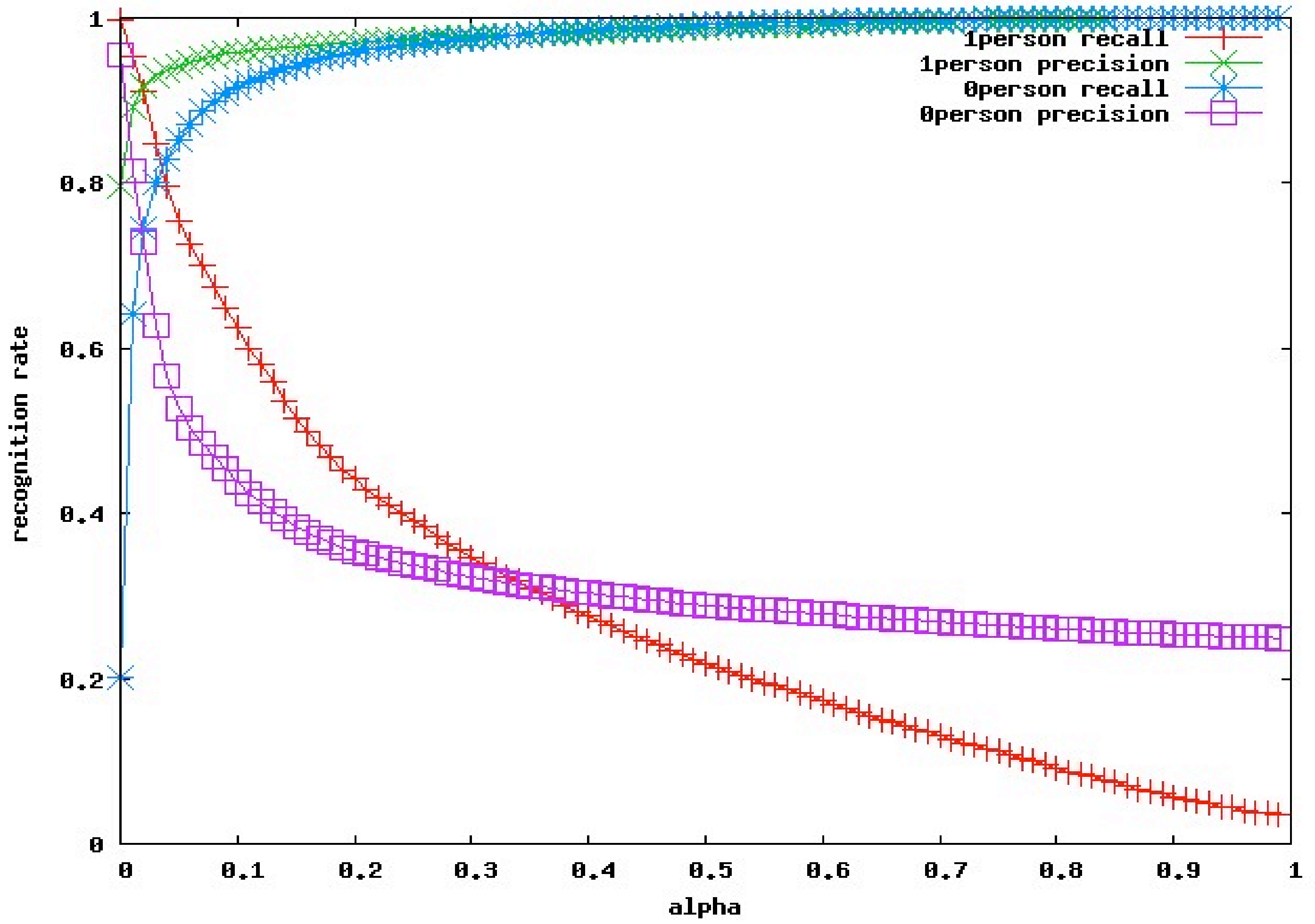
12 video sequences: 20000 imagettes with one person.

Test database:

12 video sequences: 9500 imagettes with one person

2 video sequences: 5000 imagettes without person.

Recognition Rate vs Precision



Recognition rate and precision for two classes

Experimental results: Browsing

Task: Classify imagettes
Class 1: Contains one or more human (Green)
Class 2: Does not contain a human (Red)

Experimental results: Meeting

Task: Classify imagettes
Class 1: Contains one or more human (Green)
Class 2: Does not contain a human (Red)

Experimental results: Empty

Task: Classify Imagettes
Class 1: Contains one or more human (Green)
Class 2: Does not contain a human (Red)

Conclusion

Ridge-based normalisation for statistical classification is

- Affine Invariant
- Robust to Occlusions
- Useful at Video rate.
- Useful for any object with a dominant ridge.
- Easily trained to new classes,

But also:

- Sensitive to local illumination (shadows)
- Less reliable than structural approach because of ridge detection errors (can be improved).

Recognition with Linear Auto-Associative Networks

Objectives :

- Classification of imagettes based on appearance.

Approach:

- Use Imagette ROI to normalize scale and orientation.
- Classify imagette based on Appearance

Method:

- Auto-associative Memories [Abdi94]
- Widrow-Hoff learning rule to learn a linear subspace

Recognition with Linear Auto-Associative Networks

Widrow-Hoff learning rule to learn a linear subspace

Learning: For M normalised imagettes $\{x_m\}$

$$W^{(t+1)} = W^{(t)} + h(X_m - W^{(t)} X_m) X_m^T$$

Recognition: $S = \| X - W X \|$

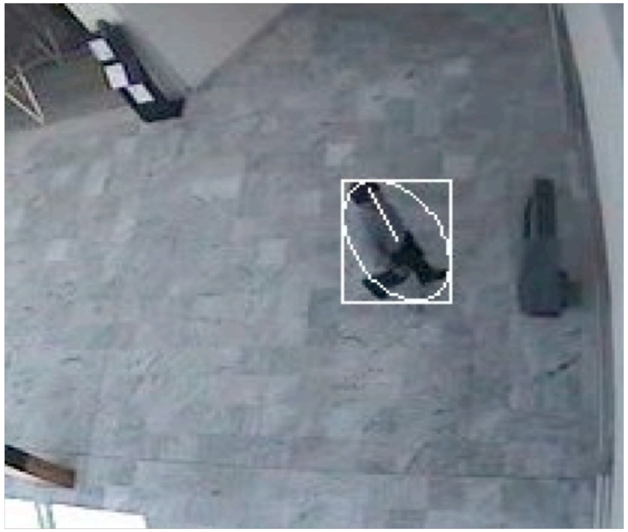
S is a similarity score $0 \leq S \leq 1$

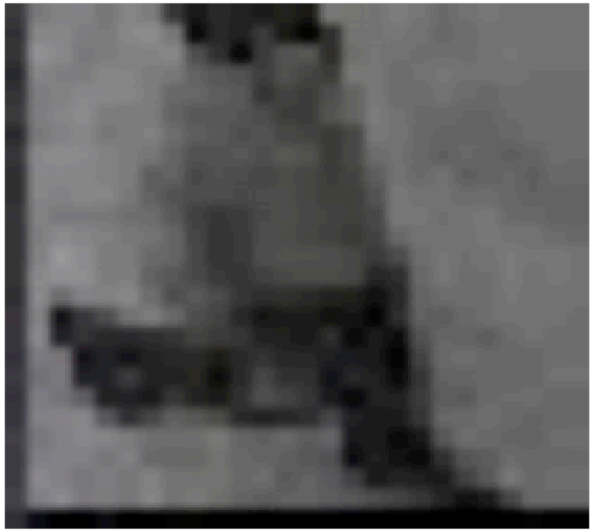
Normalisation

Normalise scale and orientation using ROI parameters

Affine transformation of imagette to 25 x 25 pixels.

(Note: Problems encountered with hand labeled data)





Recognition with Linear Auto-Associative Networks

Three Experiments:

1) Train single class for imagettes with 1 person.
   (no separate Class for 0 persons or N>1 persons)
2) Train 2 Classes: 0-persons, and 1 person.
3) Train 2 Classes: 0-persons and N>0 persons

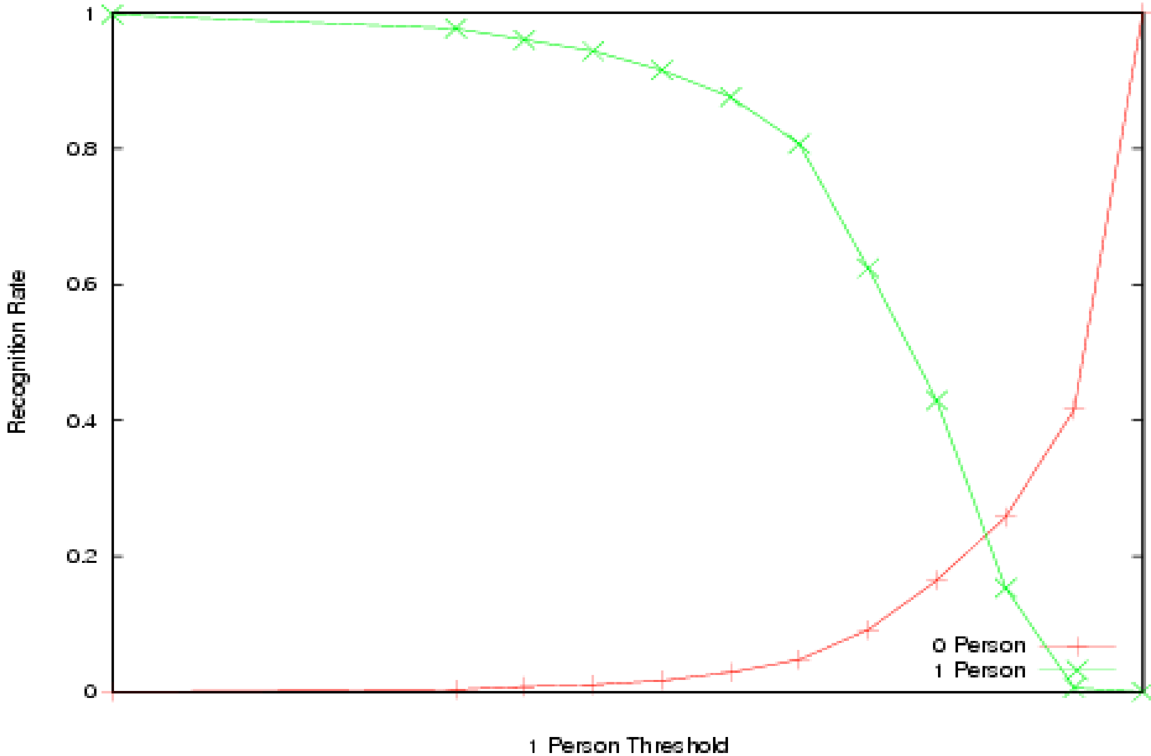
Experiment 1

Train single class for imagettes with 1 person.

Interest: Can be learned in lab and applied to all sites.

Conclusion: 0.2 recognition rate

Not reliable. Need to learn Background (imagettes with 0 person).

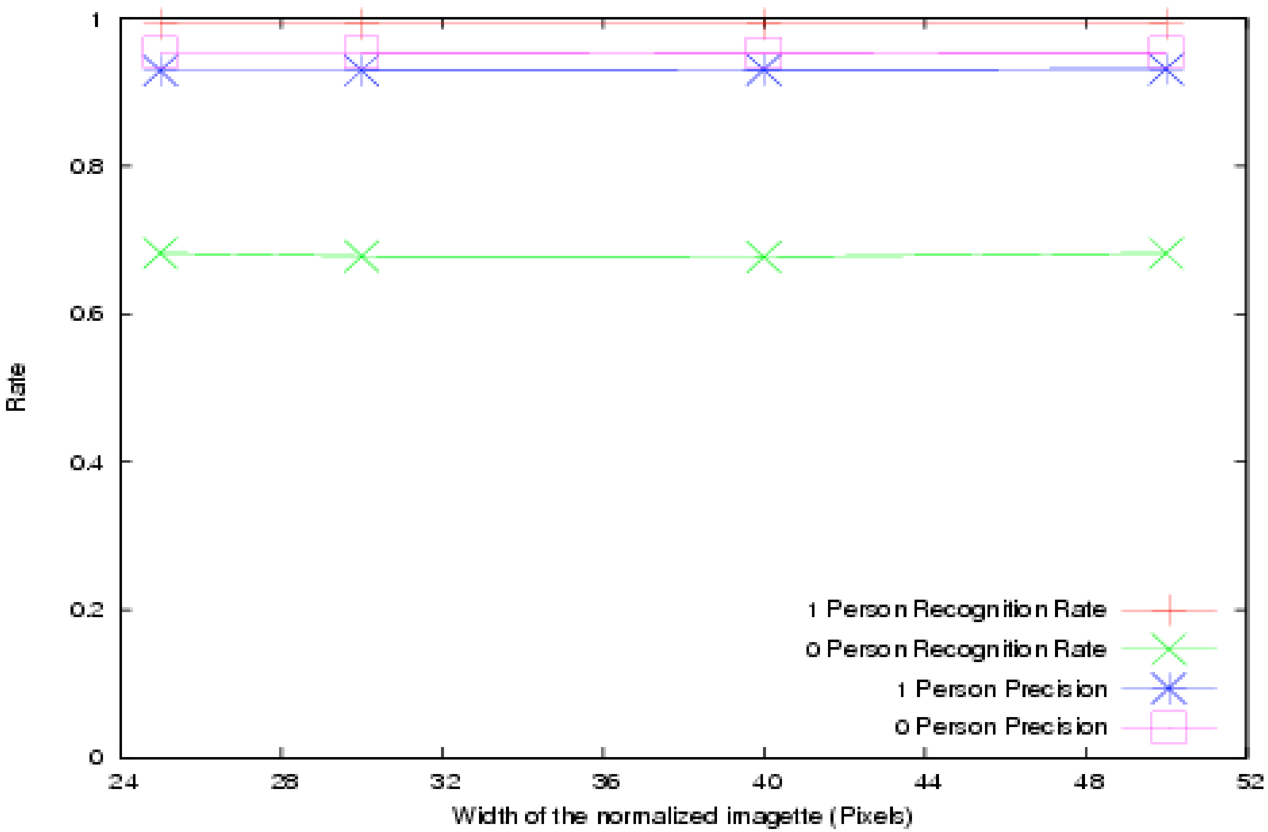


Experiment 2

Train 2 Classes: 0-persons, and 1 person.

Inconvenience: Requires learning in-situ.

Conclusion: Reliable discrimination for 0/1 person class imagettes

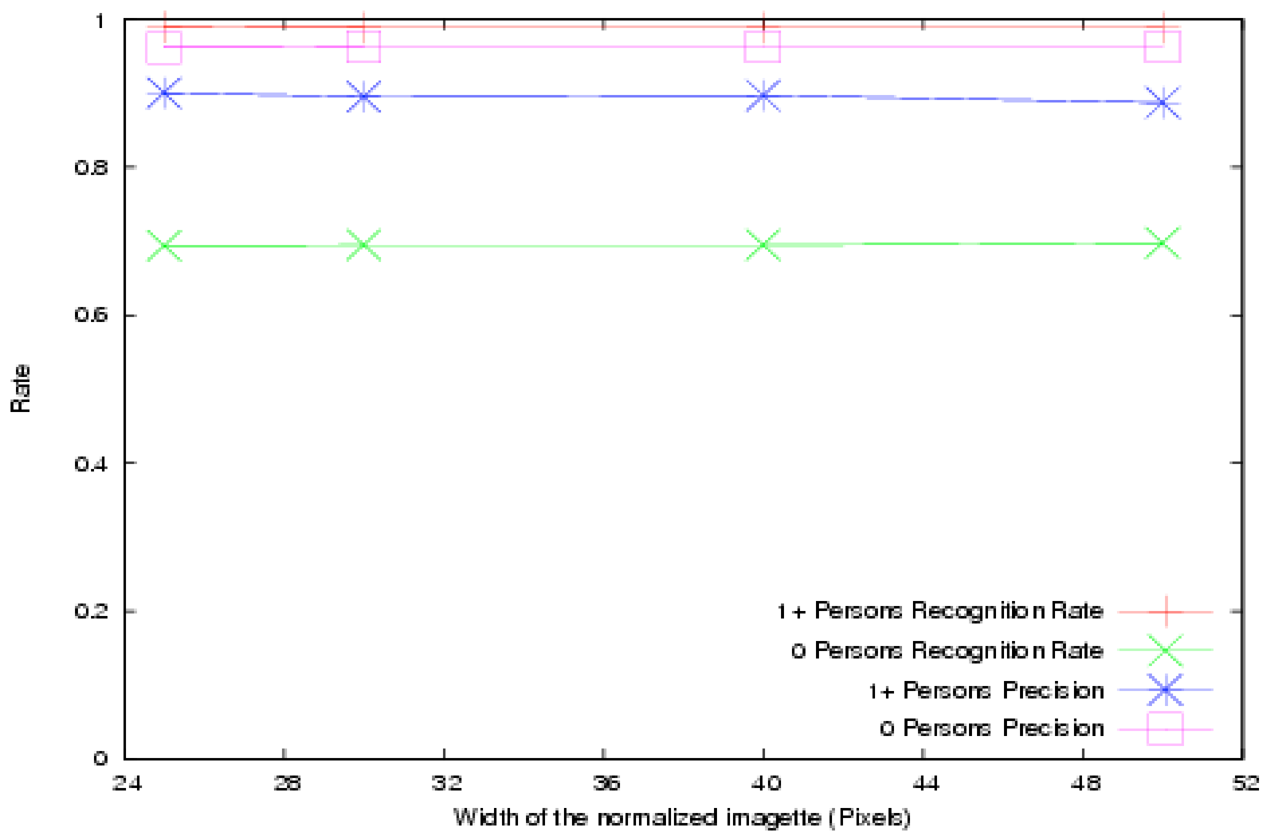


Experiment 3

Train 2 Classes: 0-persons, and N person.

Inconvenience: Requires learning in-situ.

Conclusion: Works almost as well as 0/1 person classification



Precision and Recall

Experiment Class	1 1/0 Person	2 1/0 Person	3 N/0 Persons
1st class recall	-	99 %	99 %
2nd class recall	-	68 %	70 %
1st class precision	-	95 %	93 %
2nd class precision	-	93 %	90 %

Demo: Browse Sequence

Green: N persons in Imagette
Red: No Person in Imagette

Demo: Meeting Sequence

Green: N persons in Imagette
Red: No Person in Imagette

Demo: Empty Image

Green: N persons in Imagette
Red: No Person in Imagette

Conclusions:

1) Natural Interest ridges can provide affine invariant normalisation when reliably detected.

2) Structural description is reliable, but requires hand crafted model.

3) Statisical (SIFT) description:

- Easily implemented
- Problems when prinicipal ridge is not properly detected

4) Linear Auto-Associative Networks

- Reliable and easily used for imagette classification.
- Usable for target “recognition” for reacquisition.

WP 5: Acquiring models for contexts, situations and control of perception

Objective:

Provide tools for manual and automatic acquisition of models for context, situation and control of perception.

Tasks:

- T5.1 Develop a tool for specifying process federations
- T5.2 Determine methods for automatic recognition process optimisation
- T5.3 Develop tools for modeling situations and contexts
- T5.4 Determine methods for automatic federation assembly
- T5.5 Determine methods for learning context model

Deliverables

- D27/32 Report / Software on Context Modeling
- D30/37 Report / Software on Federation Assembly

D27/32
Report / Software on Context Modeling

<u>Objective</u>:

- Provide tools for automatic acquisition of situation networks.

<u>Approach</u>:

- Statistical learning using Hidden Markov Models.

<u>Learning</u>:

- Use BaumWelch (EM) to learn models with different numbers of states
- Select model with best results

<u>Recognition</u>:

- Use viterbi algorithm

Training Data:

28 CAVIAR Benchmark Sequences with Ground truth.

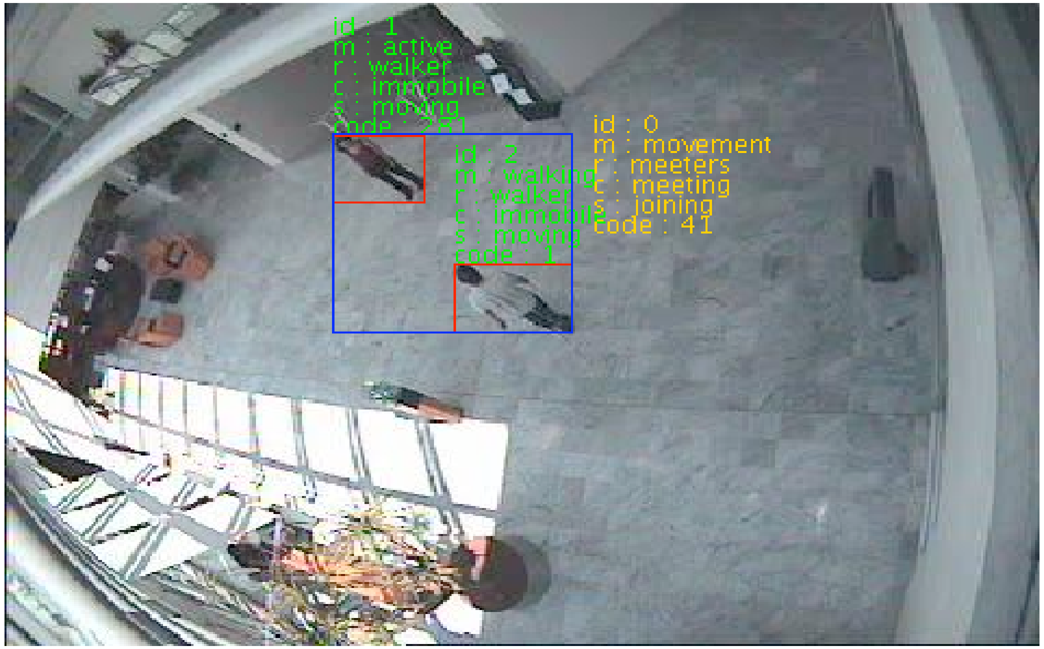

```
<object id="3">
   <orientation>101</orientation>
   <box xc="303" yc="98" w="23" h="45"/>
   <appearance>visible</appearance>
   <hypothesislist>
     <hypothesis id="1" prev="1.0"
     evaluation="1.0">
         <movement
     evaluation="1.0">active</movement>
         <role evaluation="1.0">browser</role>
         <context
     evaluation="1.0">browsing</context>
         <situation
     evaluation="1.0">browsing</situation>
     </hypothesis>
   </hypothesislist>
</object>
```

CAVIAR

D27/32
Report / Software on Context Modeling

Groups:

Use ground truth data about groups




```
<group id="0">
  <orientation>103</orientation>
  <box xc="228" yc="110" w="55" h="126"/>
  <members>1,2</members>
  <appearance>appear</appearance>
  <hypothesislist>
    <hypothesis id="1" prev="0.0"
    evaluation="1.0">
      <movement
    evaluation="1.0">movement</movement>
      <role evaluation="1.0">meeters</role>
      <context
    evaluation="1.0">meeting</context>
      <situation
    evaluation="1.0">joining</situation>
    </hypothesis>
  </hypothesislist>
</group>
```

Converting Ground Truth to Labels

<movement evaluation="1.0">active</movement>

<role evaluation="1.0">browser</role>

<situation evaluation="1.0">browsing</situation>

[active, browser, browsing]

[3, 2, 2]

62

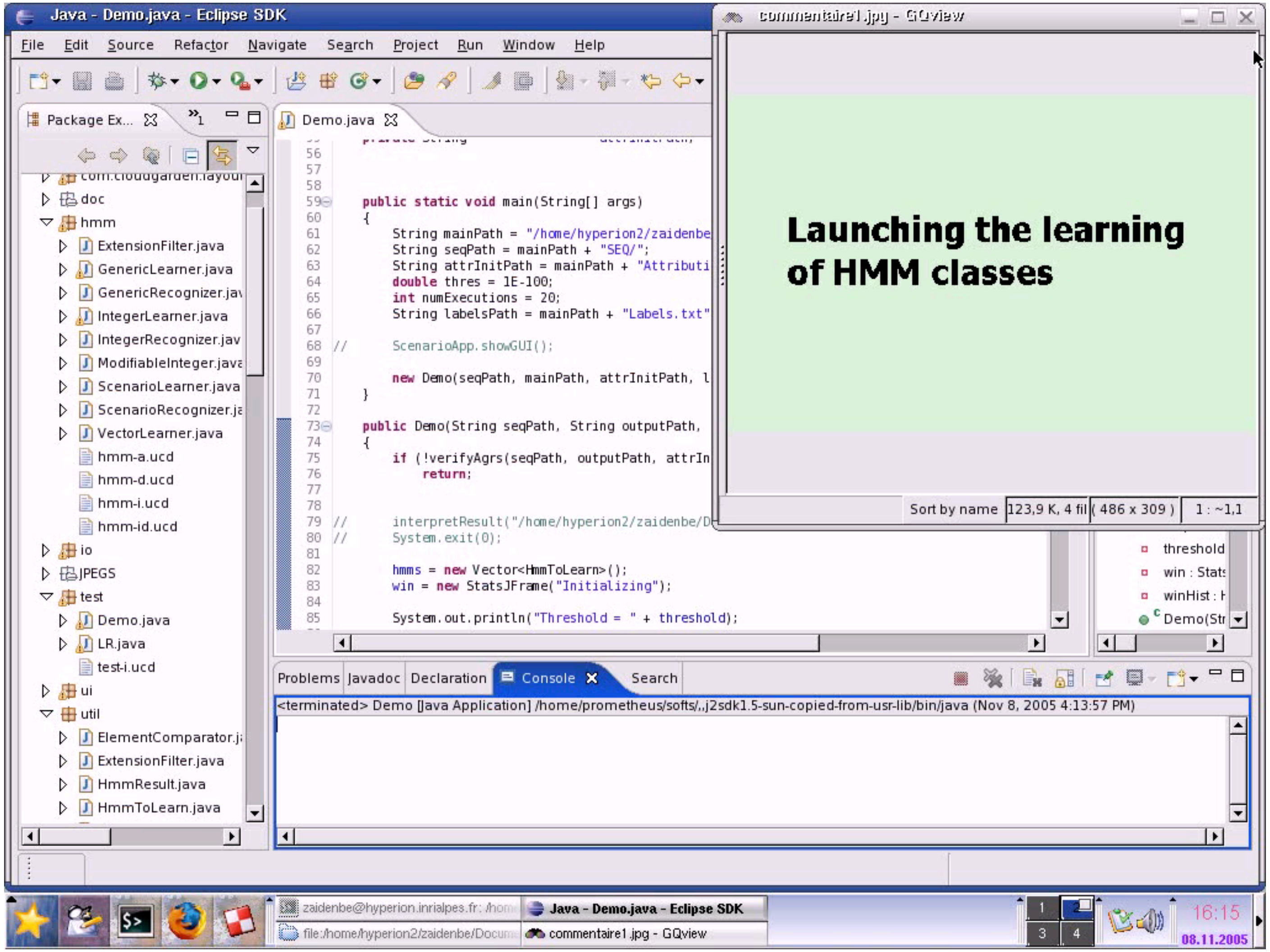
Label Sequences normalised in length to 100 labels per benchmark.

Example of sequences

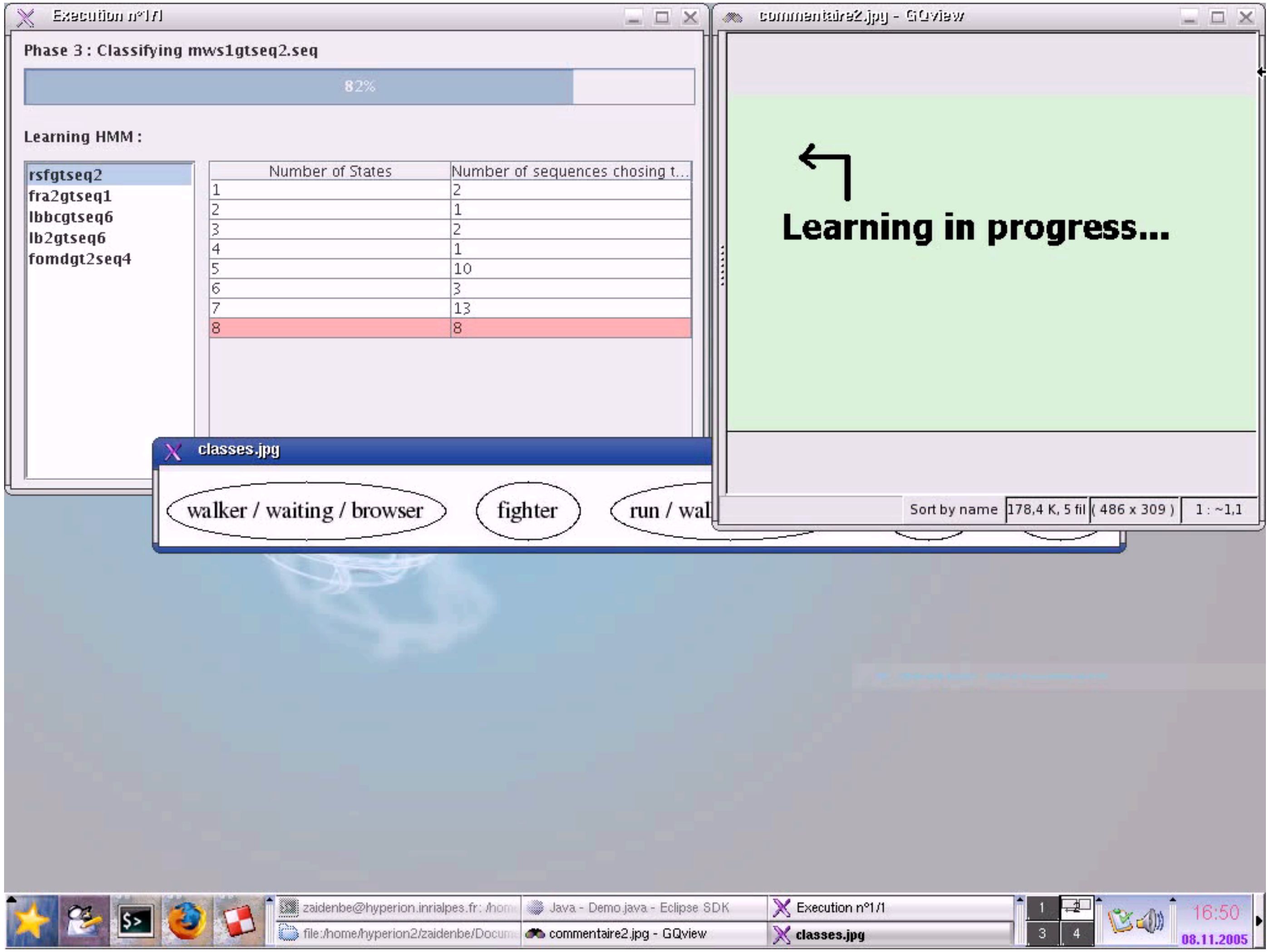
Example (“browsing 2”):

```
# object id = 1
1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1;
1; 1; 1; 1; 1; 281; 281; 281; 281; 281; 281; 281; 281; 281; 281; 281; 281; 143; 143; 143;
143; 143; 143; 143; 143; 143; 143; 143; 143; 143; 281; 281; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1;
1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1; 1;
# object id = 2
285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285;
285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285;
285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285;
285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285; 285;
285; 285; 285; 285; 285; 285;
# object id = 3
9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9;
9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 310; 310; 310; 310; 310; 310; 310; 310; 310; 310;
310; 310; 310; 310; 310; 310; 310; 310; 310; 310; 310; 310; 310; 310; 310; 310; 310; 310;
310; 310; 310; 310; 310; 310; 310; 310; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9; 9;
9; 9;
```

Demo: Learning Context Model



Demo: Recognizing Situations



Evaluation

- Error metric: number of misclassified validation sequences.
- Need of a label for each sequence and one for each class to compare.
- Sequence labeling done by hand.

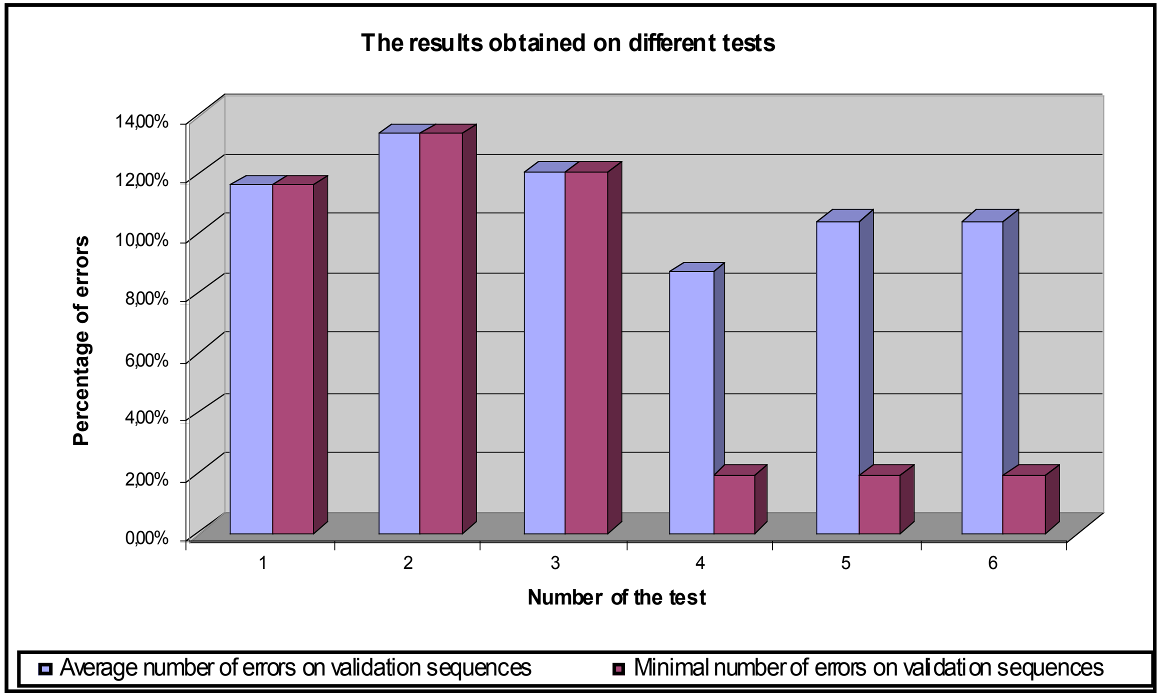
Test:

- Use of Group information for merged imagettes
- Robustness to noise.
- Influence of threshold.

test cases

Num ber	Method to create the classes		Group s		
	Manual	Automatic	Ignored	Boolean attr.	Replace by group attrs.
1	×		×		
2	×			×	
3	×				×
4		×	×		
5		×		×	
6		×			×

D27/32
Test results



Number	Method to create the classes		Groups		
	Manual	Automatic	Ignored	Boolean attr.	Replace by group attrs.
1	×		×		
2	×			×	
3	×				×
4		×	×		
5		×		×	
6		×			×

Simulation of noise

9	9	310	310	380	289	289	289	289	289	15	15	15

Random selection of *n* observations to add noise

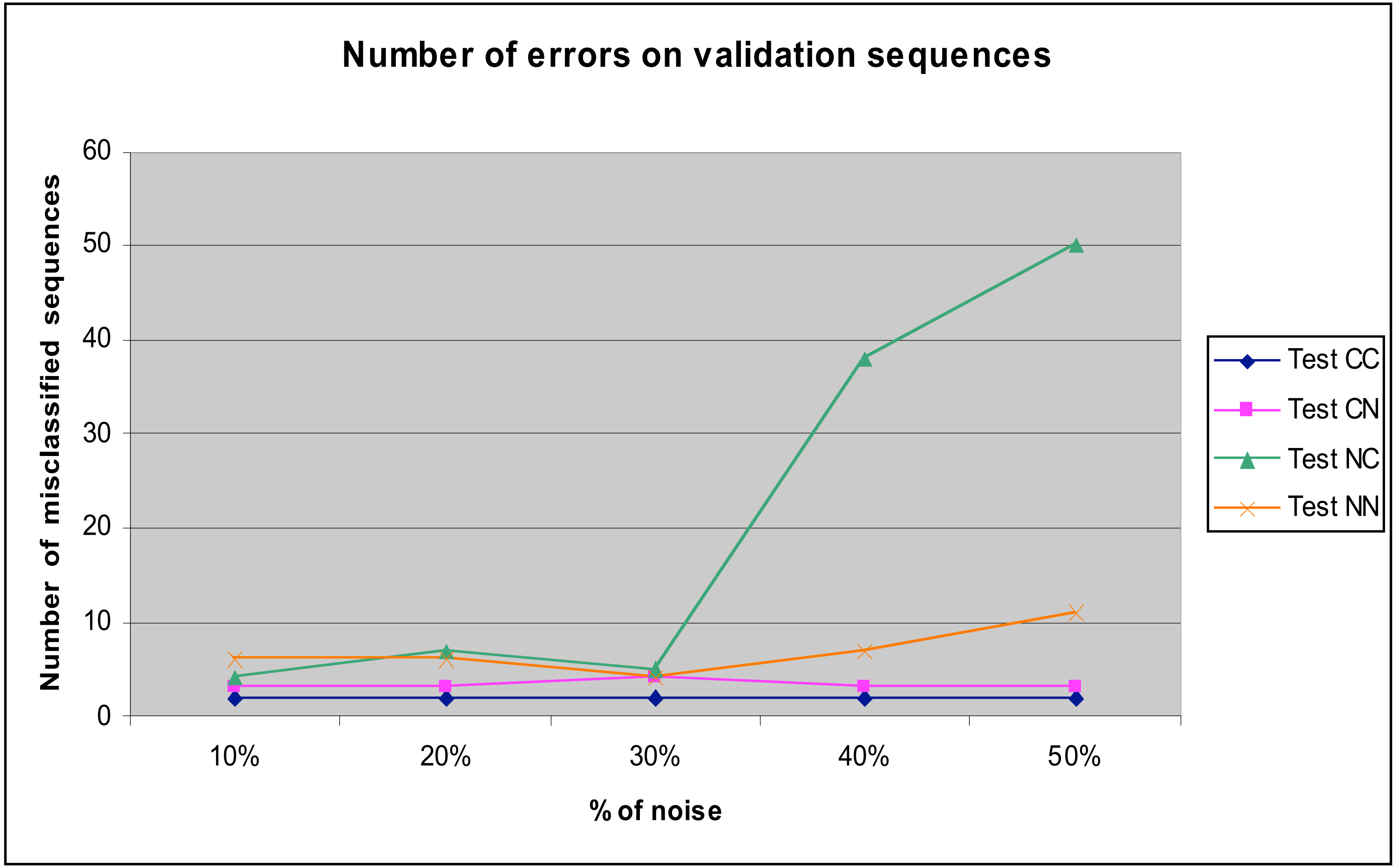
9	351	310	310	380	289	100	289	289	380	207	15	15

For each selected observation, replace code with random number

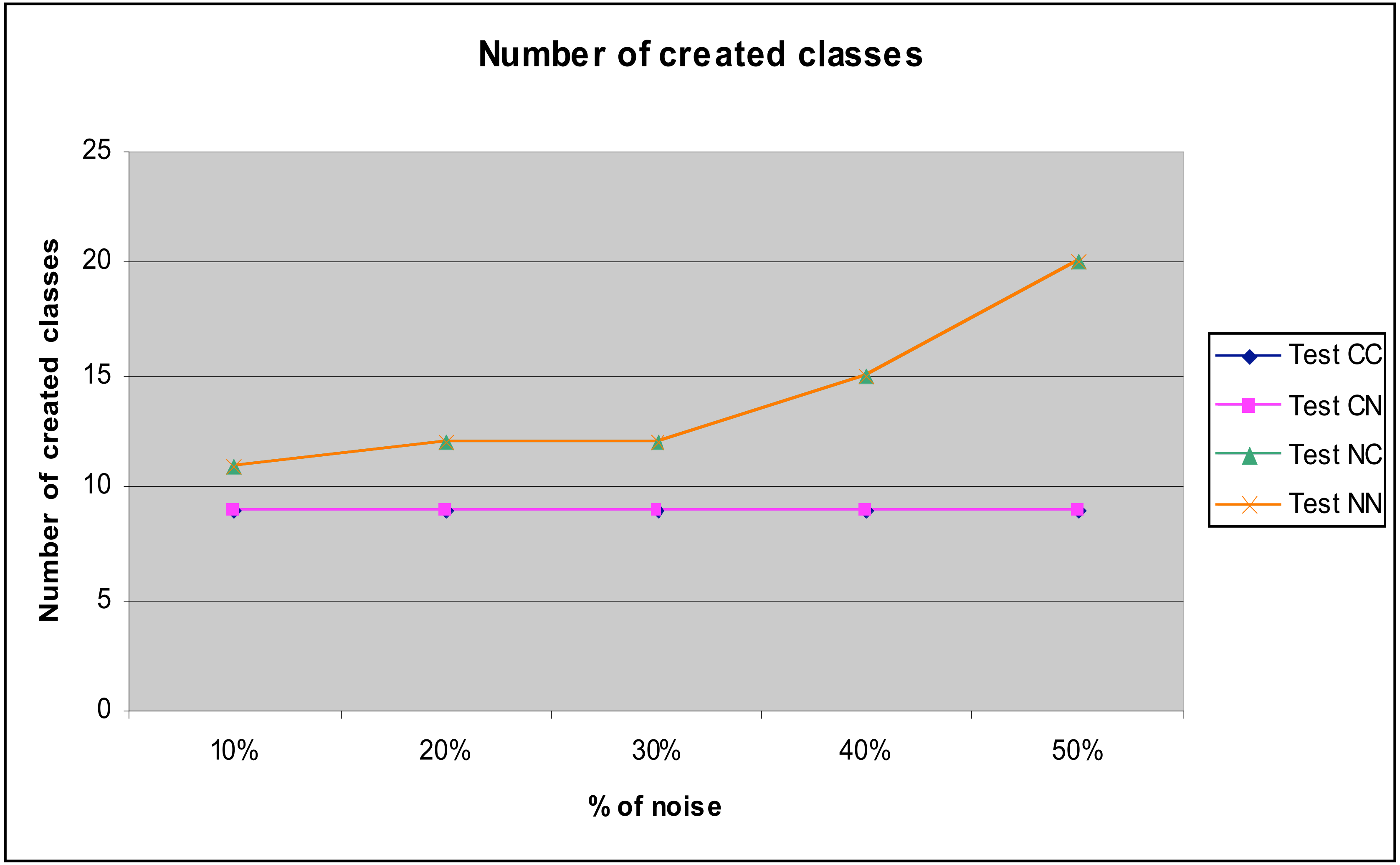
Executed tests:

		Learning	
		clean	noisy
Validation	clean	Test CC	Test NC
	noisy	Test CN	Test NN

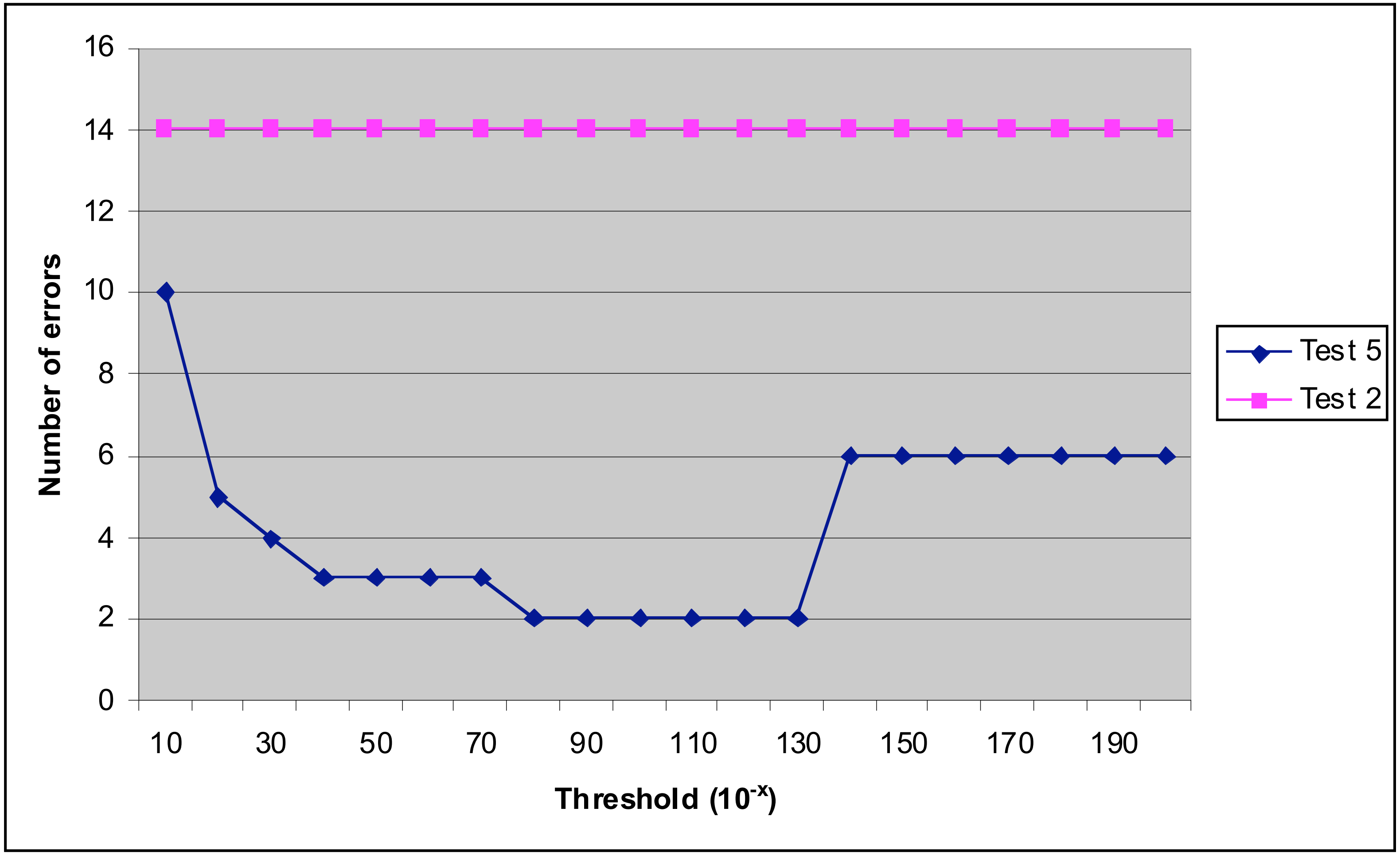
Results: number of errors



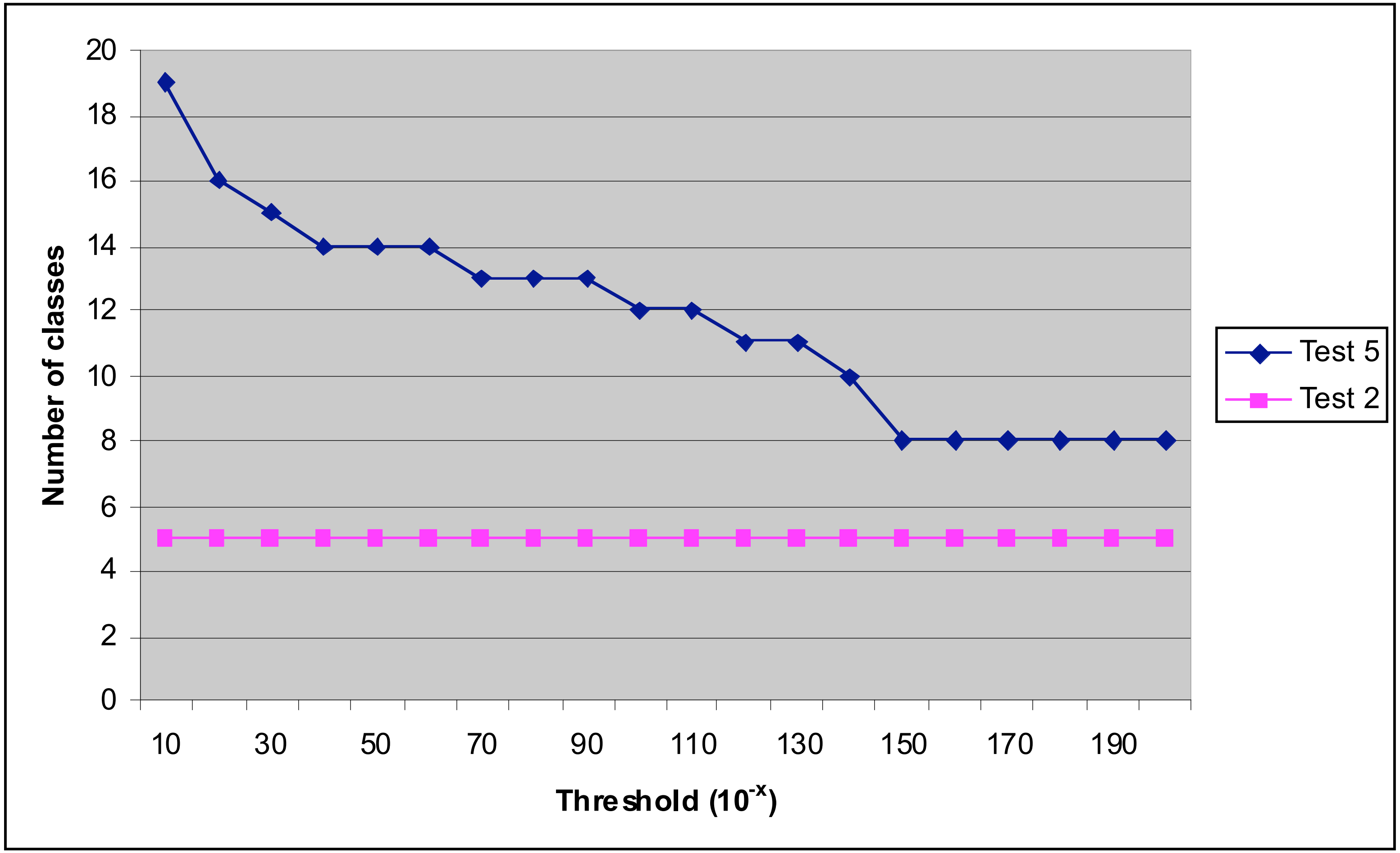
Results: number of classes



Influence of threshold



Influence of threshold



Results on Context learning

Best results when the system organizes itself.

Notion of groups does not improve performance

Good robustness to noise when less than 30% of codes corrupted.

Conclusion

Good recognition rate: 1.96% of errors
(2 errors for 102 sequences).

Strong point of method: Automatic Learning

Opens possibility for interactive model acquisition.

WP 5: Acquiring models for contexts, situations and control of perception

Objective:

- To provide tools for manual and automatic acquisition of models for contexts, situations and for control of perception.

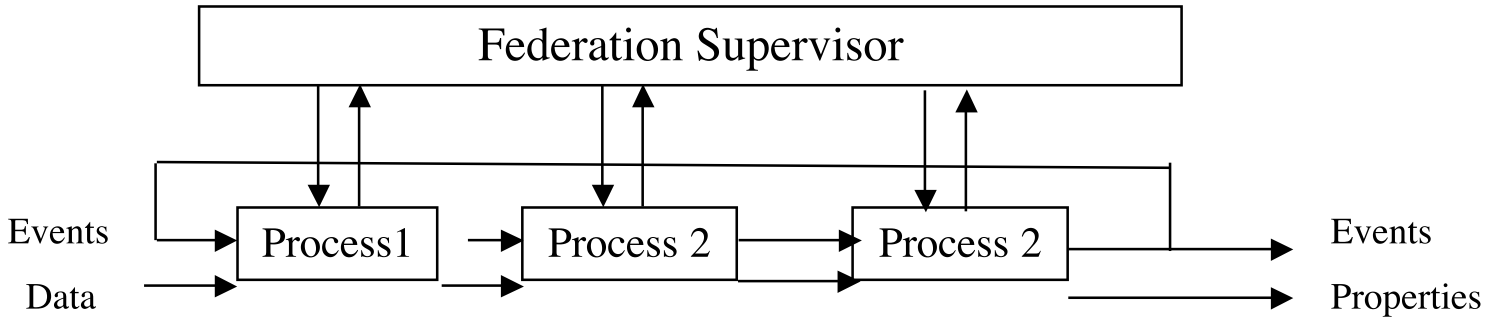
Tasks:

- T5.1 Develop a tool for specifying process federations
- T5.2 Determine methods for automatic recognition process optimisation
- T5.3 Develop tools for modeling situations and contexts
- T5.4 Determine methods for automatic federation assembly
- T5.5 Determine methods for learning context model

Deliverables

- D27/32 Report / Software on Context Modeling
- D30/37 Report / Software on Federation Assembly

Process Federation Assembly

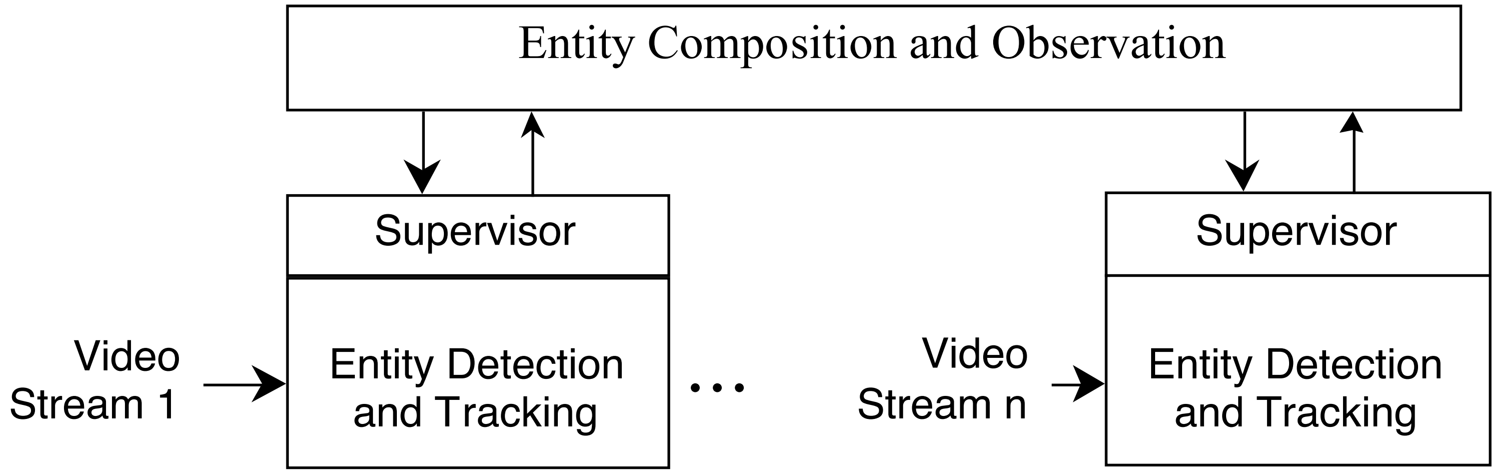


Process Federations:

Assemblies of autonomous processes

Example:
Multi-Camera target tracking

Example: Multi-Camera Target tracking using INRIA ParkingLot TestBed.



BIP : Basic Interaction Protocol for Event Flow Services

Objective:

- Provide Software tools for manual and automatic assembly of process federations

Approach: BIP: A process federation middleware providing

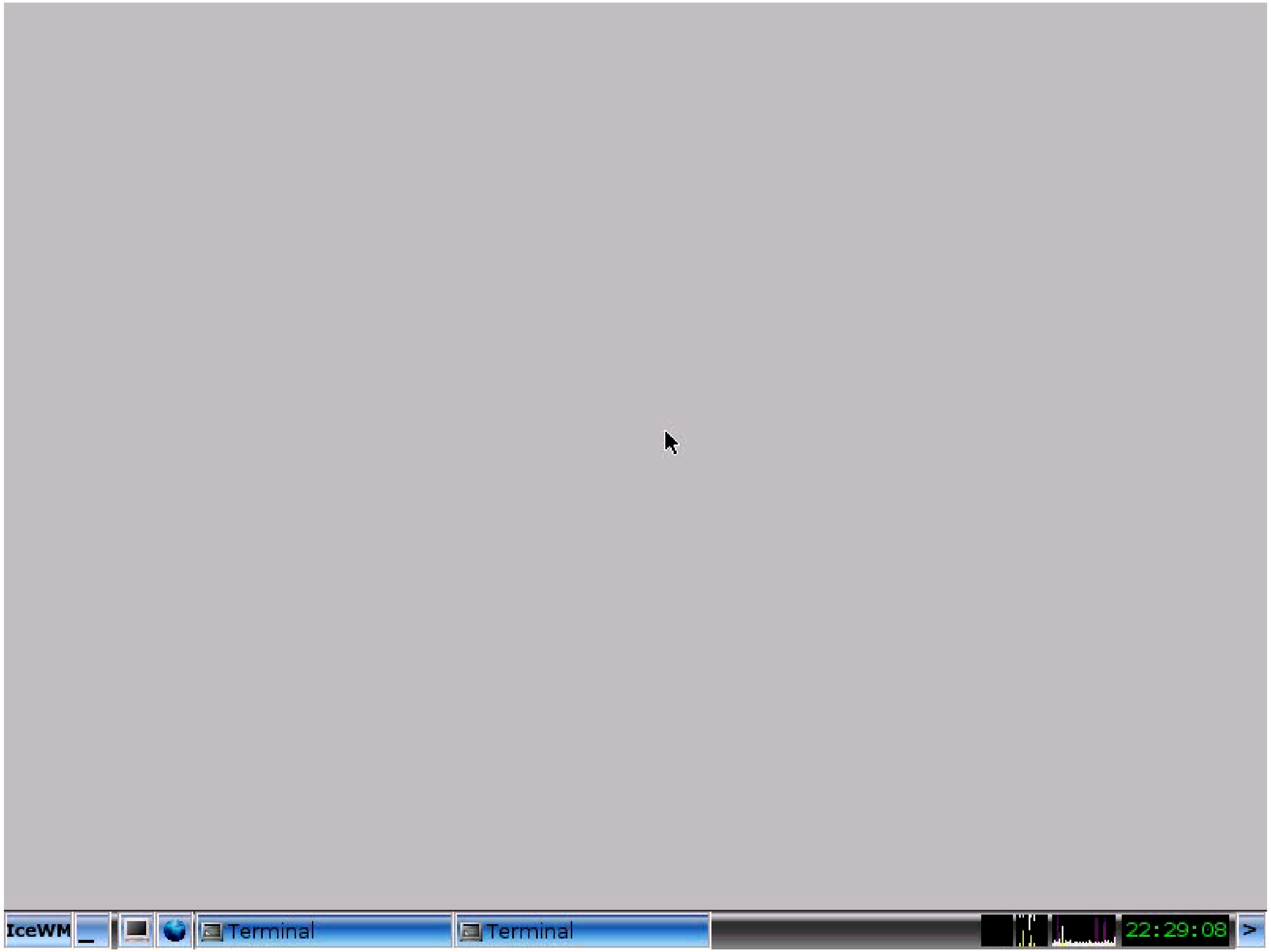
- Event Dispatching
- Establishing Direct Data streams
- Ontology registry for data and processes
- Graphical Interface tools.

BIP : Basic Interaction Protocol for Event Flow Services

(Adapted from Process Federation Tool)

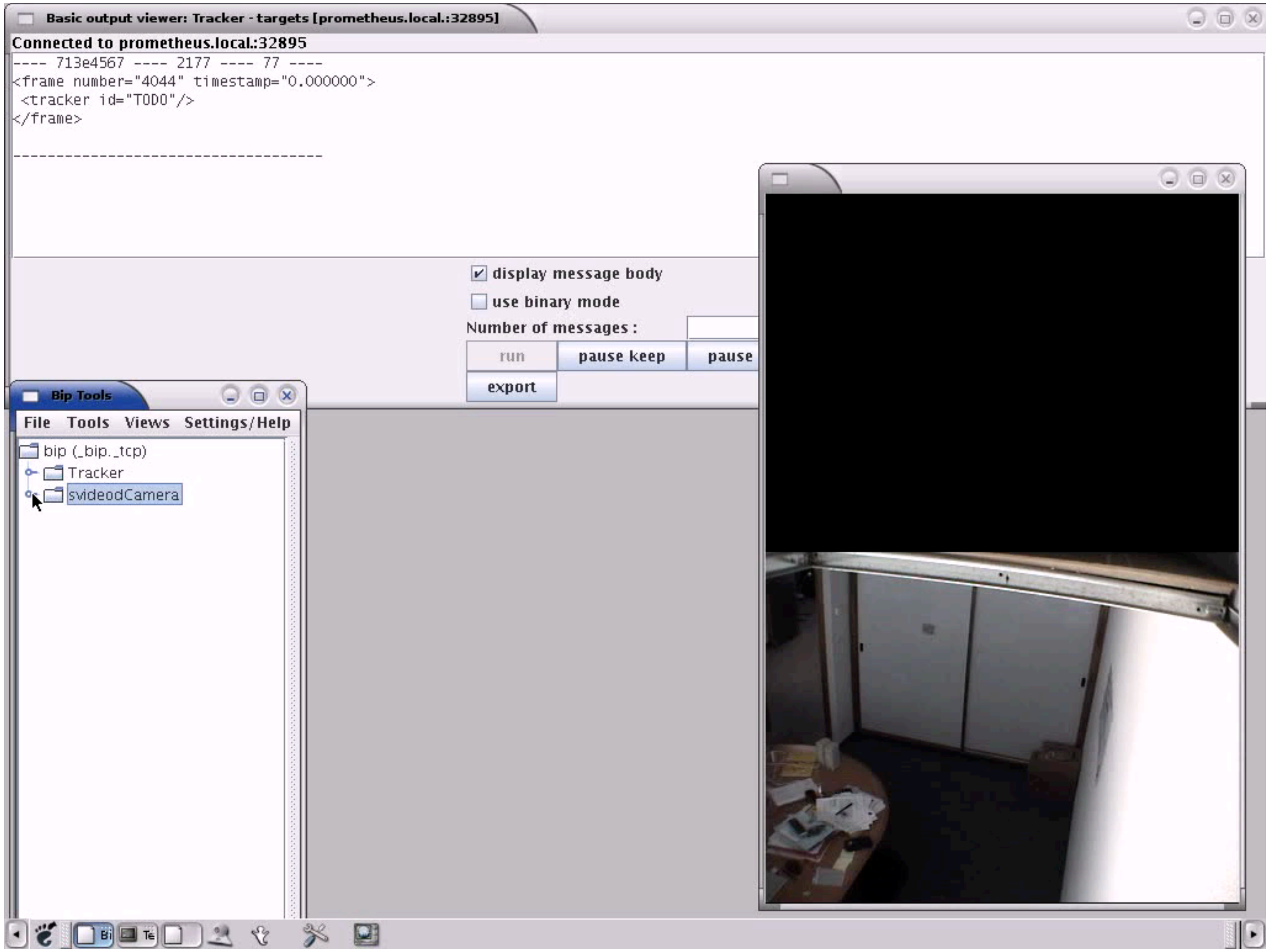
- Module Library Management Tool (in ImaLab)
- Process Configuration Tool
- Federation Configuration tool
- Process and Federation Monitoring Tool

BIP: first version (test)

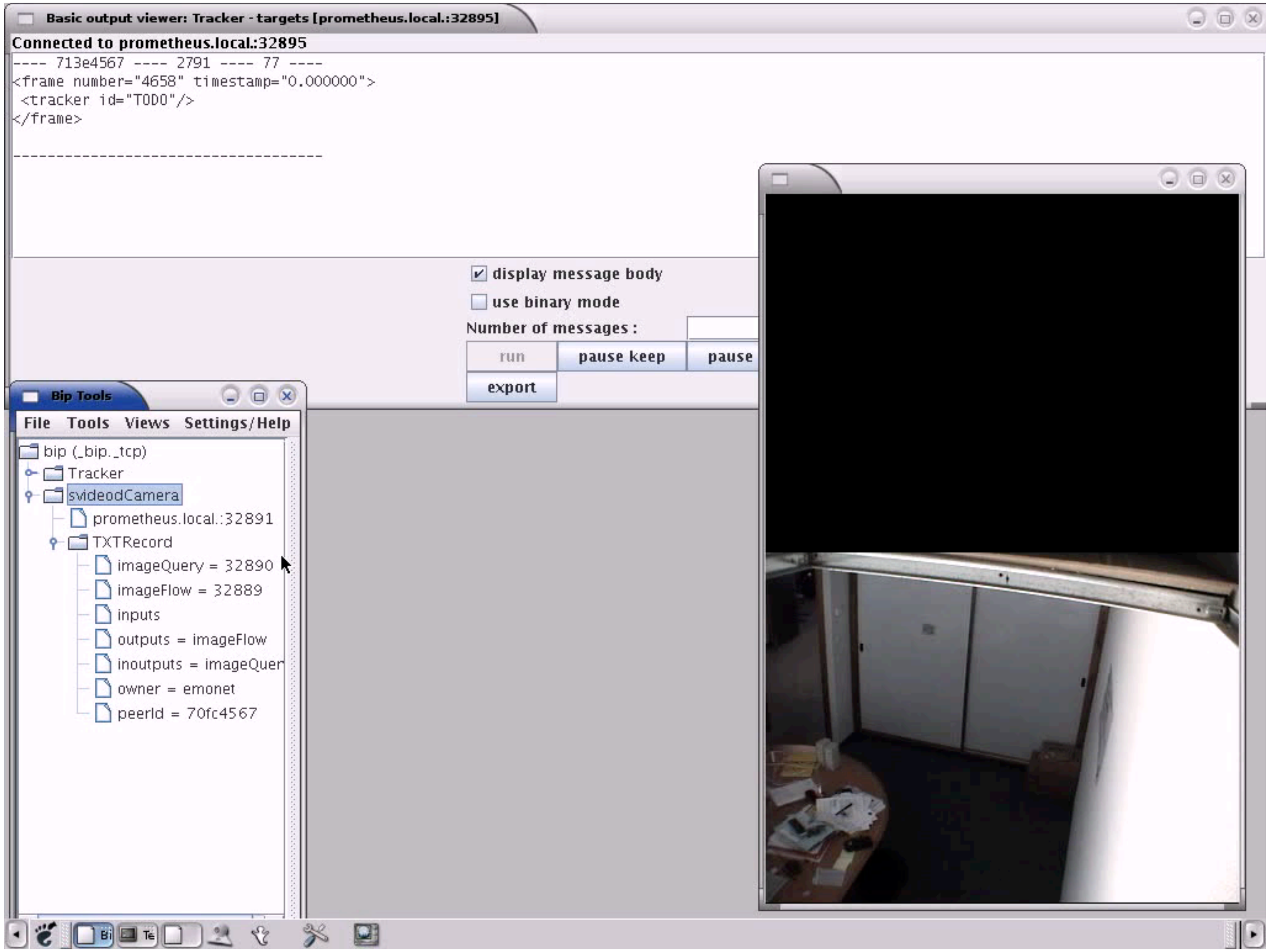


BIP: Federating a Camera and a Tracker

BIP Demo: Federation of image query connected to a tracker debugger

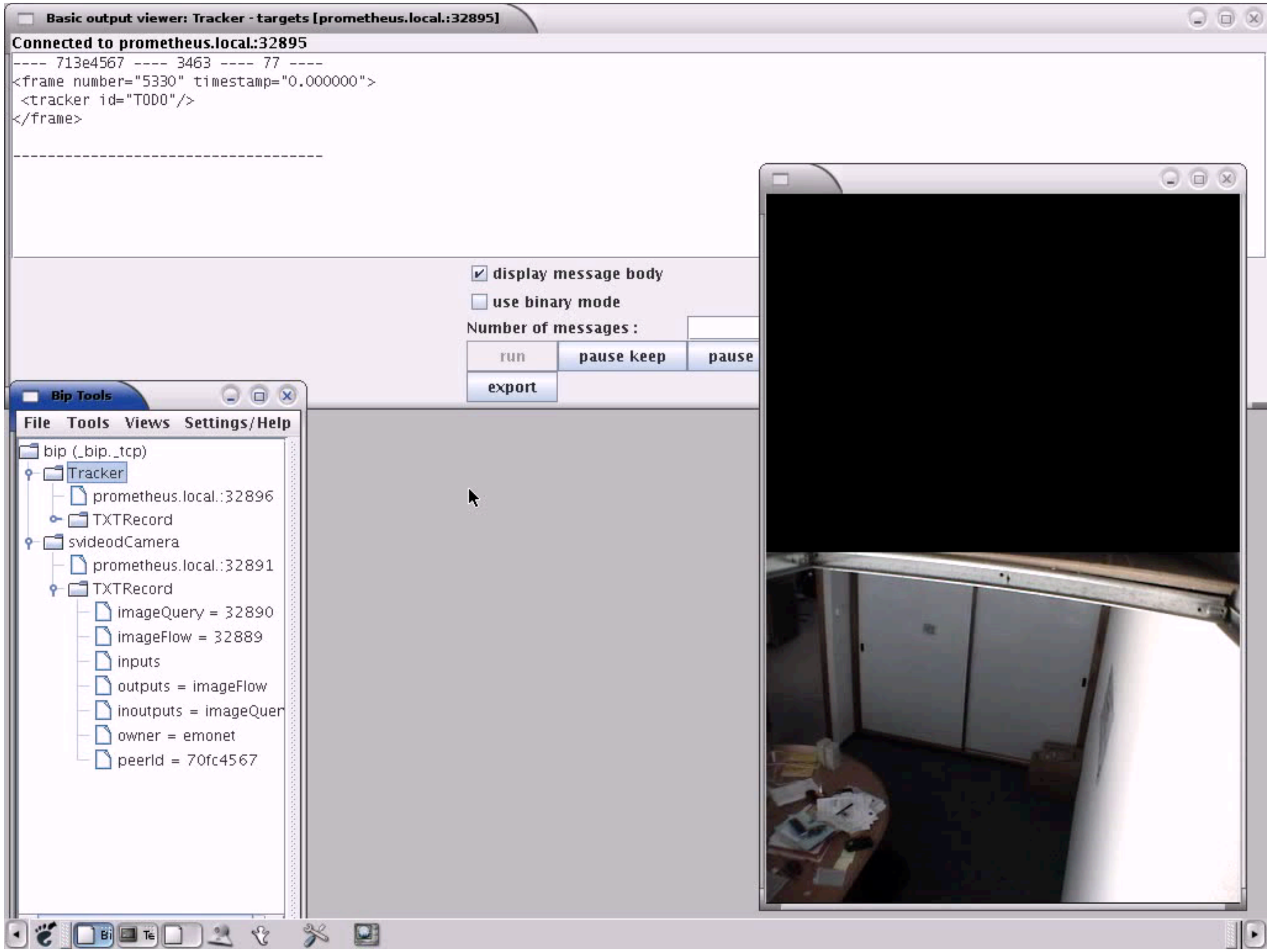


BIP Demo: Federation with tracker connected to a tracker debugger

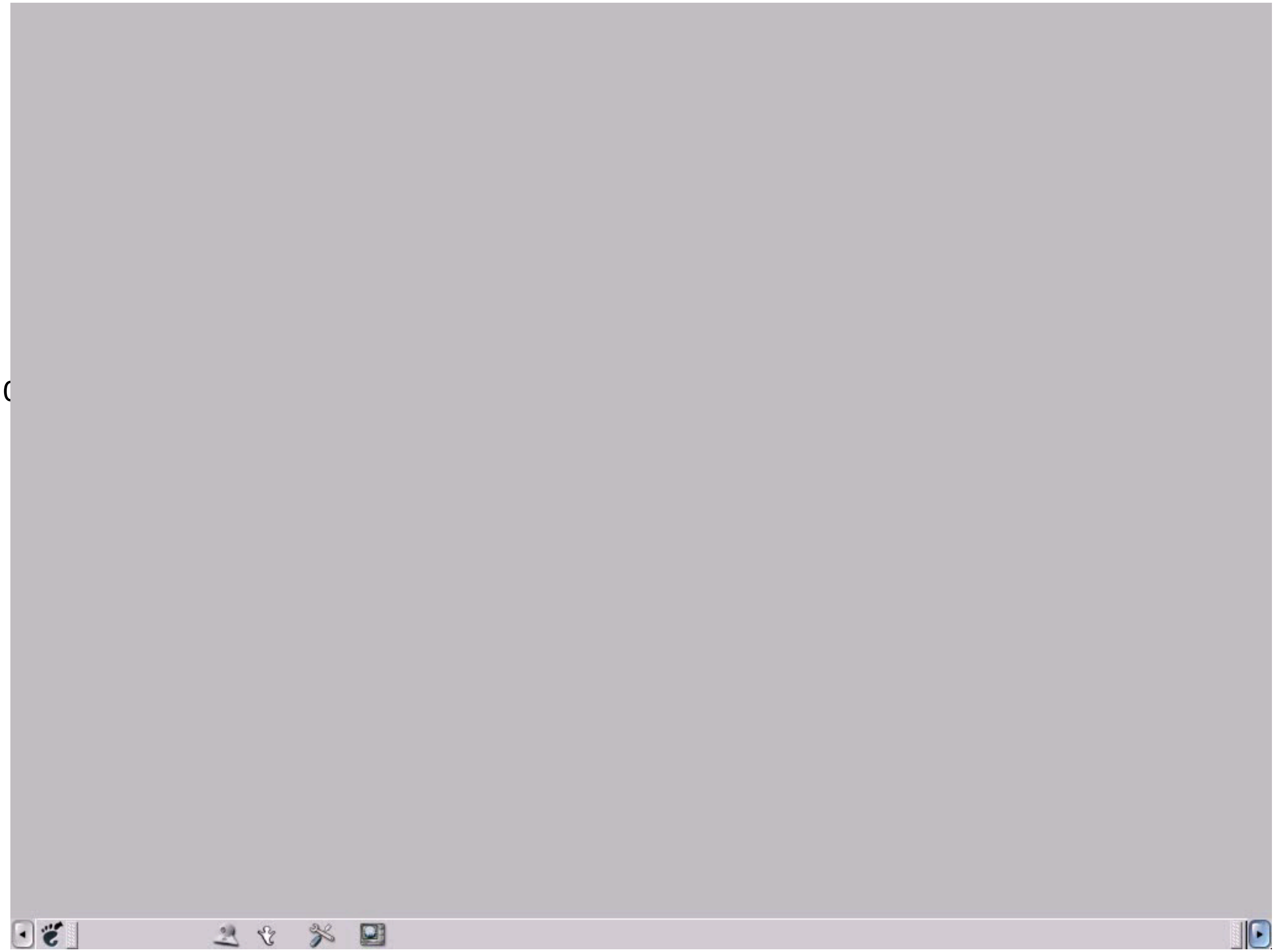


CAVIAR

BIP Demo: Process Shutdown



BIP Demo: distributed service discovery



WP 5: Acquiring models for contexts, situations and control of perception

Objective:

- To provide tools for manual and automatic acquisition of models for contexts, situations and for control of perception.

Tasks:

- T5.1 Develop a tool for specifying process federations
- T5.2 Determine methods for automatic recognition process optimisation
- T5.3 Develop tools for modeling situations and contexts
- T5.4 Determine methods for automatic federation assembly
- T5.5 Determine methods for learning context model

Deliverables

- D27/32 Report / Software on Context Modeling
- D30/37 Report / Software on Federation Assembly

WP 6 Performance Evaluation

INRIA has constructed two test bed environments for acquiring benchmark data sets

Indoor Test-bed: INRIA RA Entrance Hall

- 2 Cameras: one w/wide angle lens, one steerable pan-tilt-zoom

Outdoor Test-bed: INRIA Back Parking lot

- 2 Cameras platforms:
  - 4 Fixed cameras w/wide angle lens
  - 1 steerable surveillance cameras
  - Fixed cameras form wide baseline stereo pairs

WP7 Dissemination

Objectives

- Disseminate the research results of CAVIR
- Disseminate the research experience to colleagues

Tasks

- T7.1 Scientific dissemination
- T7.2 Educational dissemination
- T7.3 Industrial dissemination

Deliverables

- D3: (month 12) ECVision network activity
- D11: (month 24) ECVision network activity
- D26: (month 36) ECVision network activity

CAVIAR

WP7 Dissemination

T7.1 Scientific dissemination

Publications:

2 book chapter, 2 MSc dissertations, 12 conf. papers, 1 journal papers.

T7.2 Educational dissemination

Workshop: Sixth IEEE Int. Work. on Performance Evaluation of Tracking and Surveillance (PETS04), in association with ECCV 2004, Prague, May 2004.

T7.3 Industrial dissemination

- INRIA and Blue Eye Video: commercial Applications
- IST & Observit: access to new shopping centre surveillance data for ground truth
- UEDIN: new EPSRC project that complements CAVIAR, investigating pre-fight detection and dense crowd behaviour.

Project Web site: http://homepages.inf.ed.ac.uk/rbf/CAVIAR/

EXTRA SLIDES