Vetronics Winter Workshop
Warren, MI, USA
January 11, 2005

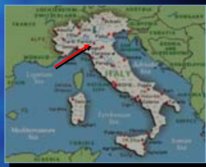
Research on Human Shape Localization using Daylight and FIR Stereo Images

Alberto Broggi, Paolo Grisleri
broggi@ce.unipr.it grisleri@ce.unipr.it
University of Parma, Italy

Artificial Vision and Intelligent Systems Lab

About the speakers

- Artificial Vision and Intelligent Systems Lab, University of Parma, Italy
- The research team:
  - 1 professor
  - 1 full time researcher
  - 3 temporary researchers
  - 4 PhD students
  - 4 external collaborators



Alberto Broggi - University of Parma, Italy

The Initial Projects
in the automotive field

- Prometheus (late 80s – early 90s)
  - Visual sensing in automotive environments
- ARGO (1997 – 2001)
  - Automatic steering in highways (including lane, obstacle, vehicle, and pedestrian detection)
  - \`MilleMiglia in Automatico\` test tour

Alberto Broggi - University of Parma, Italy

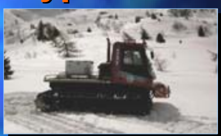
Current Projects
in the automotive field

- Volkswagen (covered by NDA)
  - Pedestrian detection
- FIAT Research Center (covered by NDA)
  - Vehicle detection and headlights detection
- US Army (Tacom, Detroit, USA)
  - Pedestrian detection in unstructured environments
- APALACI (PReVENT European Project – VI Fw) with CRF, Bosch, DC, VTech, IBEO
  - Obstacle detection and classification
- DARPA Grand Challenge (with Oshkosh, USA)
  - Automatic off-road driving

Alberto Broggi - University of Parma, Italy

Old Prototypes


MOB-LAB: 1992-1994


RAS: 1998-2000


ARGO: 1997-2001


WOB1: 2001-2003

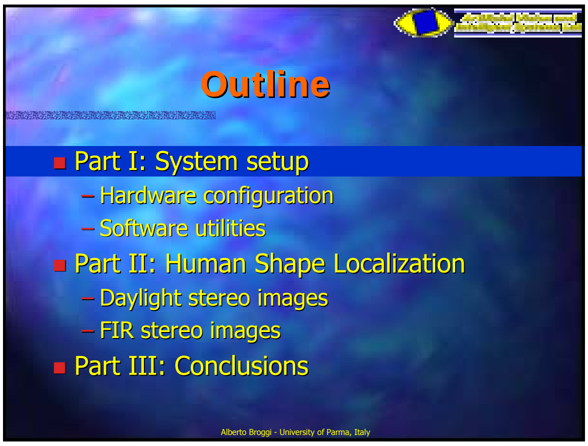
Alberto Broggi - University of Parma, Italy

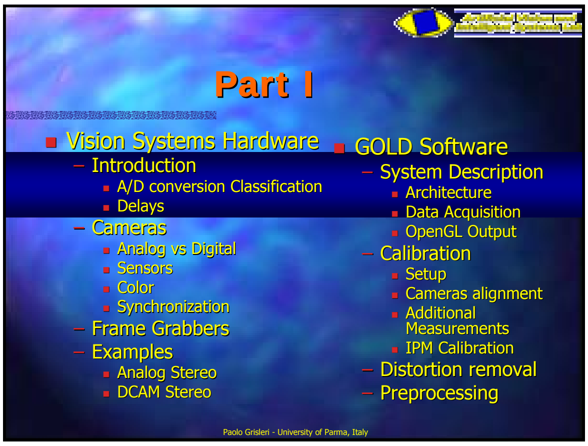
Current Prototypes

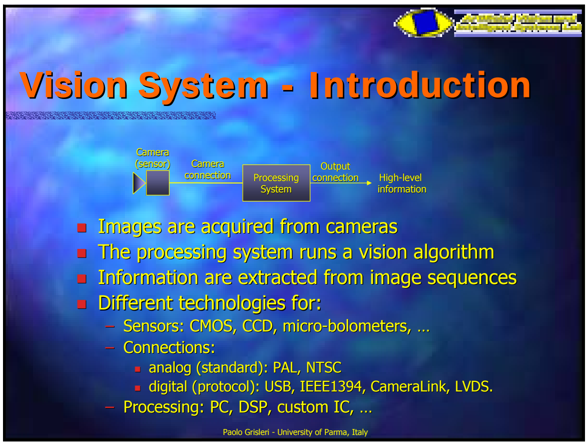

TerraMax: 2004

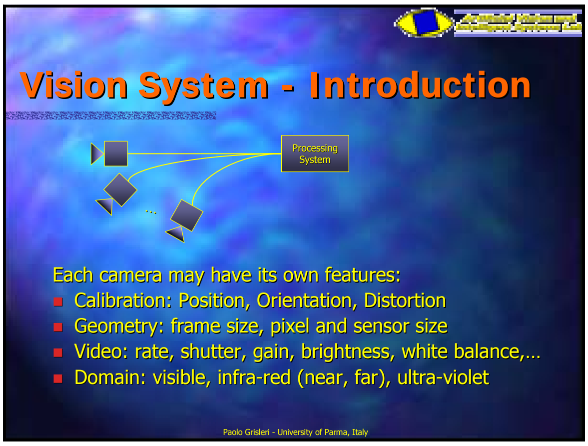

VW Touareg: 2004

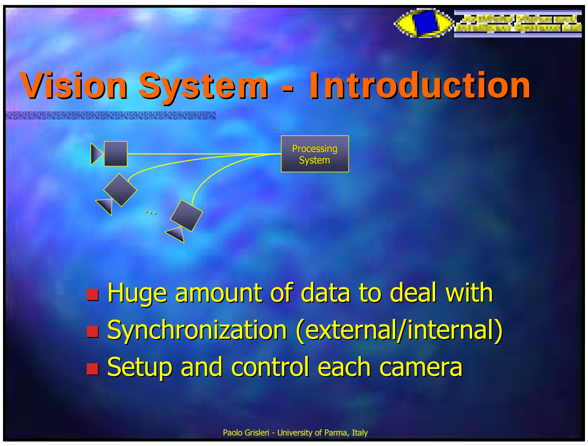
Alberto Broggi - University of Parma, Italy

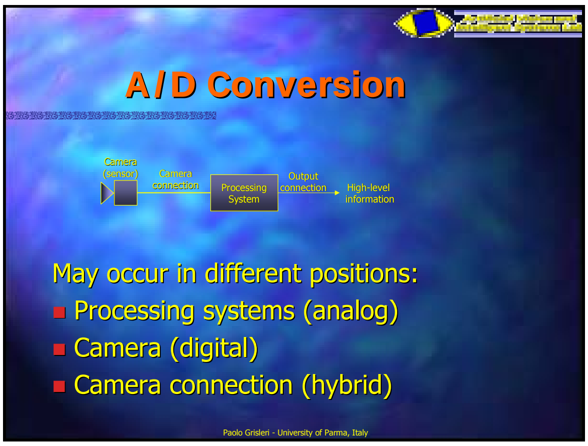

Outline
Part I: System setup
– Hardware configuration
– Software utilities
Part II: Human Shape Localization
– Daylight stereo images
– FIR stereo images
Part III: Conclusions
Alberto Broggi - University of Parma, Italy

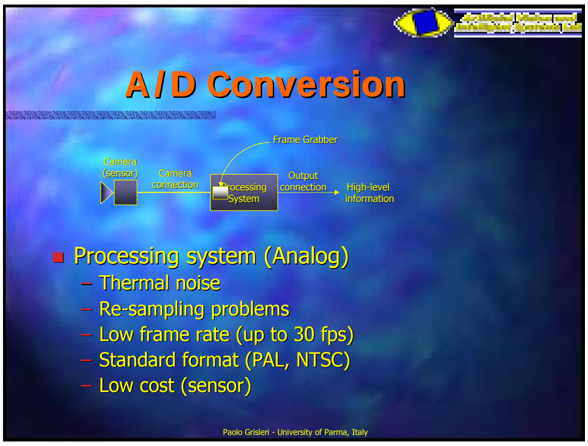

Part I
Vision Systems Hardware
– Introduction
A/D conversion Classification
Delays
– Cameras
Analog vs Digital
Sensors
Color
Synchronization
– Frame Grabbers
– Examples
Analog Stereo
DCAM Stereo
GOLD Software
– System Description
Architecture
Data Acquisition
OpenGL Output
– Calibration
Setup
Cameras alignment
Additional Measurements
IPM Calibration
– Distortion removal
– Preprocessing
Paolo Grisleri - University of Parma, Italy

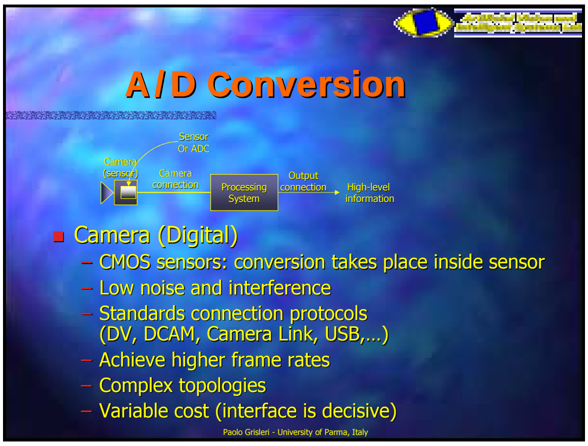

Vision System - Introduction
Camera (sensor)
Camera connection
Processing System
Output connection
High-level information
Images are acquired from cameras
The processing system runs a vision algorithm
Information are extracted from image sequences
Different technologies for:
– Sensors: CMOS, CCD, micro-bolometers, …
– Connections:
analog (standard): PAL, NTSC
digital (protocol): USB, IEEE1394, CameraLink, LVDS.
– Processing: PC, DSP, custom IC, …
Paolo Grisleri - University of Parma, Italy

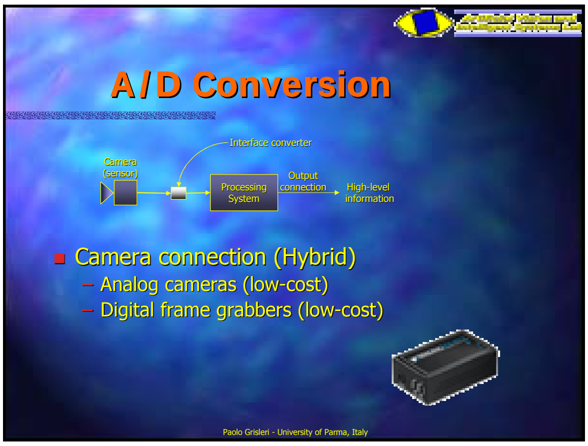

Vision System - Introduction
Processing System
Each camera may have its own features:
Calibration: Position, Orientation, Distortion
Geometry: frame size, pixel and sensor size
Video: rate, shutter, gain, brightness, white balance,…
Domain: visible, infra-red (near, far), ultra-violet
Paolo Grisleri - University of Parma, Italy

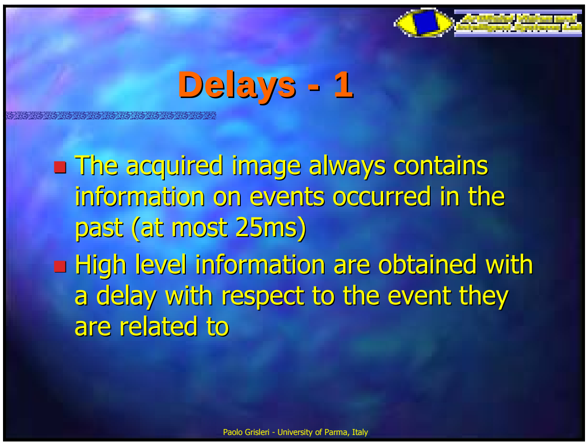

Vision System - Introduction
Processing System
Huge amount of data to deal with
Synchronization (external/internal)
Setup and control each camera
Paolo Grisleri - University of Parma, Italy

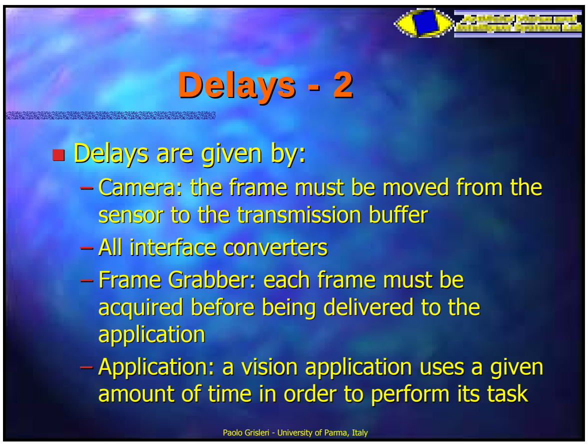

A/D Conversion
Camera (sensor)
Camera connection
Processing System
Output connection
High-level information
May occur in different positions:
Processing systems (analog)
Camera (digital)
Camera connection (hybrid)
Paolo Grisleri - University of Parma, Italy

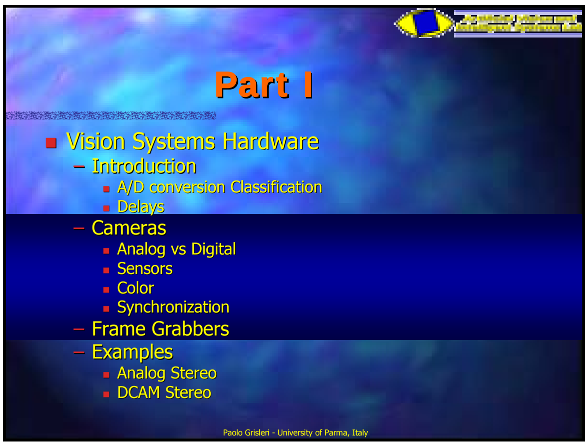

A/D Conversion
Frame Grabber
Camera (sensor)
Camera connection
Processing System
Output connection
High-level information
Processing system (Analog)
– Thermal noise
– Re-sampling problems
– Low frame rate (up to 30 fps)
– Standard format (PAL, NTSC)
– Low cost (sensor)
Paolo Grisleri - University of Parma, Italy


A/D Conversion
Sensor Or ADC
Camera (sensor)
Camera connection
Processing System
Output connection
High-level information
Camera (Digital)
– CMOS sensors: conversion takes place inside sensor
– Low noise and interference
– Standards connection protocols (DV, DCAM, Camera Link, USB,…)
– Achieve higher frame rates
– Complex topologies
– Variable cost (interface is decisive)
Paolo Grisleri - University of Parma, Italy


A/D Conversion
Interface converter
Camera (sensor)
Processing System
Output connection
High-level information
Camera connection (Hybrid)
– Analog cameras (low-cost)
– Digital frame grabbers (low-cost)
Paolo Grisleri - University of Parma, Italy


Delays - 1
The acquired image always contains information on events occurred in the past (at most 25ms)
High level information are obtained with a delay with respect to the event they are related to
Paolo Grisleri - University of Parma, Italy


Delays - 2
Delays are given by:
– Camera: the frame must be moved from the sensor to the transmission buffer
– All interface converters
– Frame Grabber: each frame must be acquired before being delivered to the application
– Application: a vision application uses a given amount of time in order to perform its task
Paolo Grisleri - University of Parma, Italy


Part I
Vision Systems Hardware
– Introduction
A/D conversion Classification
Delays
– Cameras
Analog vs Digital
Sensors
Color
Synchronization
– Frame Grabbers
– Examples
Analog Stereo
DCAM Stereo
Paolo Grisleri - University of Parma, Italy

Cameras

- Input
  - Power
  - Trigger/synchronization
  - Framed scene
- Output
  - Video (signal/stream)
  - Synchronization
- Parameters
  - Adjustable physically, electrically, software
  - AGC (On/Off, offset), shutter, color temp.
  - Frame rate, frame size

Video	→	Photo
	Lens	
	Iris	
Shutter: Electr.	→	Mechanical
Sensor	→	Film

Analog Cameras

- Sensors:
  - visible (gray and color), NIR, FIR
  - technology CCD, CMOS, microbolometers
- Format: fixed
  - depends on video norm (PAL, NTSC, …)
  - may be acquired at different resolutions or cropped
- Output: BNC, Composite, S-Video
- Synchronization: internal, external

Paolo Grisleri - University of Parma, Italy

Digital Cameras - 1

- Sensors:
  - Visible (Grey and Color), NIR, FIR
  - Technology CCD, CMOS, Micro-Bolometers
- Format: variable
  - Software Adjustable Resolution: Region of Interest (ROI) even at run-time
  - Sensor dependent: beyond 4096x4096

Paolo Grisleri - University of Parma, Italy

Digital Cameras - 2

- Output:
  - Physical interface: USB, FW, RS422, LVDS(RS644), CameraLink, Ethernet
  - Some interfaces may use different protocols (FW: DV, DCAM)
  - Compressed format to reduce bandwidth (JPEG,YUV…)
- Synchronization: external, software, or auto

Paolo Grisleri - University of Parma, Italy

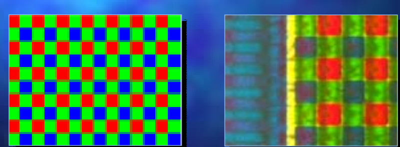
Color

- Color may be used as an additional source of information
- Color does not have the same relevance it has for human vision
- Technologies:
  - Color Filter Mosaic Array (Bayer pattern)
  - Spectral separation

Paolo Grisleri - University of Parma, Italy

Color filter mosaic array (CFA)

- A color filter mosaic array (CFA) is applied on top of a black and white imager



Paolo Grisleri - University of Parma, Italy

Color filter mosaic array (CFA)

- Additional steps are required to reconstruct the color image
- Can be performed:
  - by the camera: requires processing power and a large bandwidth for data transfer
  - by the host: cheap cameras; the processing power is provided by the host, so an additional delay is introduced

Paolo Grisleri - University of Parma, Italy

CFA: Bayer Pattern

- Several CFA patterns can be used
  - Pseudo-random
  - Fixed number of neighbors
- Bayer Pattern:
  - 2x2 kernel (e.g. RGGB) preference for green as luminance channel
  - artifacts for vertical and horizontal lines

Paolo Grisleri - University of Parma, Italy

Synchronization - 1

- Multiple views vision systems (stereo, trinocular,…) need synchronized frames
- Synchronized cameras have:
  - the same integration time (ε_I)
  - the same start acquisition time (ε_S)

Paolo Grisleri - University of Parma, Italy

Synchronization - 2

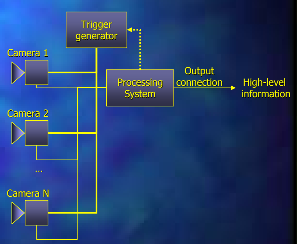
- Synchronization needs depends on the specific application:
  - for slow moving scenes (terrain): ε~s
  - for fast moving scenes (crash-test): ε~µs
- Synchronization can be obtained:
  - Hardware (analog and digital)
  - Software (digital)

Paolo Grisleri - University of Parma, Italy

Hardware Synchronization

- Master/slave cameras
- Trigger generator

+ Very precise
- Camera must have compatible trigger input
- Additional hardware required to change trigger via software



Paolo Grisleri - University of Parma, Italy

Software Synchronization

Digital cameras receive commands via bus

- Trigger command:
  - shared through the bus: generated by an external device or a master camera
  - sent on the bus by control software
- Example: MicroPix C-640:
  - synchronized below 125 µs
  - when connected on the same bus with other C-640

Paolo Grisleri - University of Parma, Italy

Analog Frame Grabbers

- Board hosted on the PC bus
- One ore more video input
- Conversion chip: BT8x8, Philips SAA 7146, …
- Specific device driver needed
- Low cost hardware synchronization
- 3 grey synchronized cameras signals can be acquired using 1 composite input
- 2 grey synchronized interlaced camera signals can be acquired using 1 interlaced gray input
- High level library mask hardware peculiarity

Paolo Grisleri - University of Parma, Italy

Low-cost digital frame grabbers

- Focus on Firewire or USB connection Controllers are integrated on motherboards
- High number of devices
- Easy synchronization
- Highly configurable bus topology
- Bandwidth limitation (FW: 400Mbps → 5 grey640x480@30fps)

Paolo Grisleri - University of Parma, Italy

Part I

- Vision Systems Hardware
  - Introduction
    - A/D conversion Classification
    - Delays
  - Cameras
    - Analog vs Digital
    - Sensors
    - Color
    - Synchronization
  - Frame Grabbers
  - Examples
    - Analog Stereo
    - DCAM Stereo

Paolo Grisleri - University of Parma, Italy

Example: Analog Stereo System

- Cameras: 2 x TELI
- Frame Grabber: 2x BT878 boards
- Additional Hardware: SONY DC-700: Power supply and synchronization
- API: Video for Linux 2

Paolo Grisleri - University of Parma, Italy

Example: Digital Stereo System

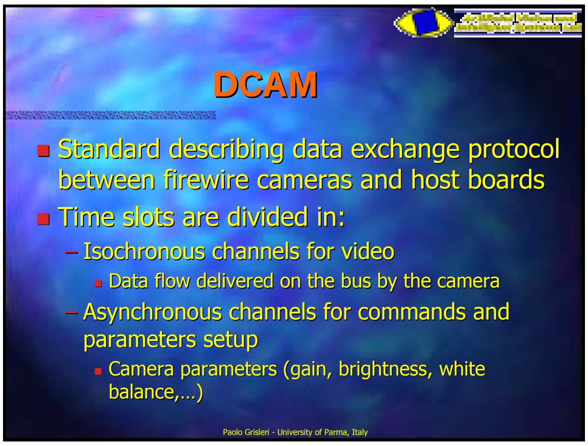
- Cameras: MicroPix C640
- Frame Grabber: OHCI IEEE 1394 board
- API: libdc1394

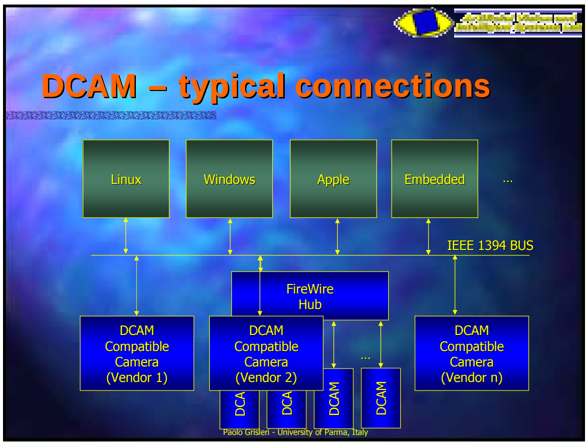
Paolo Grisleri - University of Parma, Italy

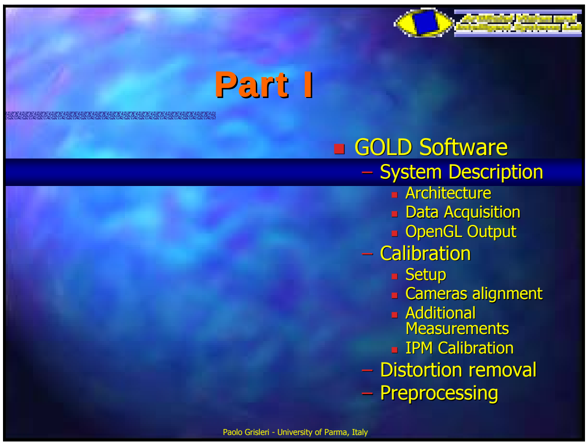
API: Video for Linux 2

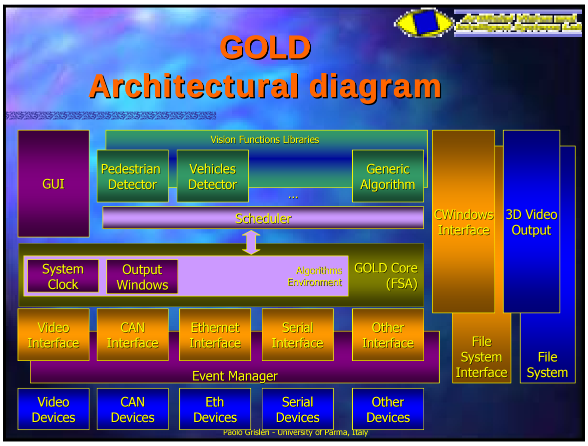
- Video acquisition API integrated in 2.6.x linux kernel:
  - Acquire frames in a customizable format
  - Set up frame grabber properties (brightness, contrast, …)
  - Widely used and stable
- Hardware support:
  - Framegrabbers: BT8xx, Zoran, Philips,…

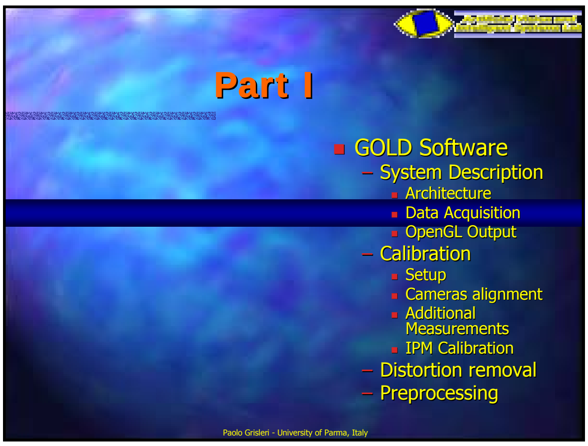
Paolo Grisleri - University of Parma, Italy

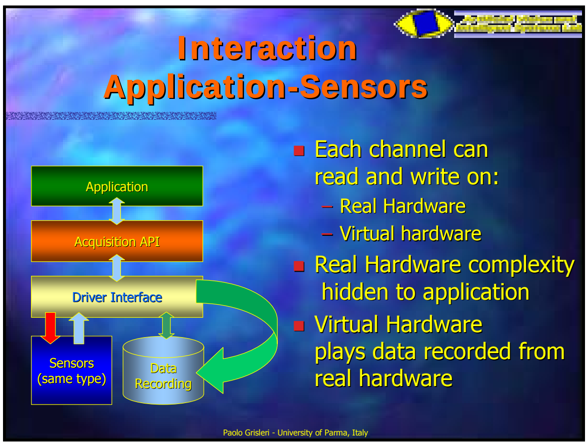

DCAM
Standard describing data exchange protocol between firewire cameras and host boards
Time slots are divided in:
Isochronous channels for video
Data flow delivered on the bus by the camera
Asynchronous channels for commands and parameters setup
Camera parameters (gain, brightness, white balance,…)
Paolo Grisleri - University of Parma, Italy

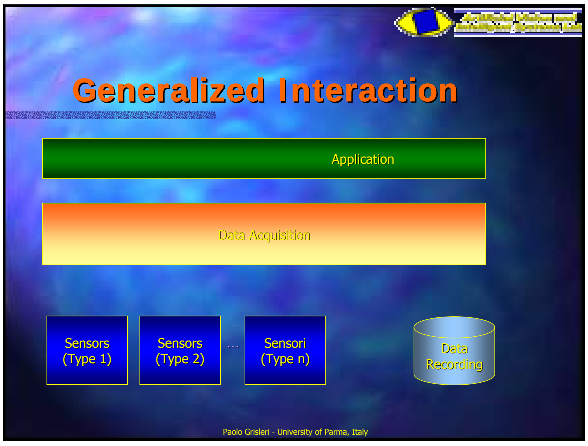

DCAM – typical connections
Linux
Windows
Apple
Embedded
…
IEEE 1394 BUS
FireWire Hub
DCAM Compatible Camera (Vendor 1)
DCAM Compatible Camera (Vendor 2)
DCAM Compatible Camera (Vendor n)
DCAM
Paolo Grisleri - University of Parma, Italy


Part I
GOLD Software
System Description
Architecture
Data Acquisition
OpenGL Output
Calibration
Setup
Cameras alignment
Additional Measurements
IPM Calibration
Distortion removal
Preprocessing
Paolo Grisleri - University of Parma, Italy


GOLD
Architectural diagram
Vision Functions Libraries
GUI
Pedestrian Detector
Vehicles Detector
Generic Algorithm
Scheduler
CWindows Interface
3D Video Output
System Clock
Output Windows
GOLD Core (FSA)
Video Interface
CAN Interface
Ethernet Interface
Serial Interface
Other Interface
File System Interface
File System
Event Manager
Video Devices
CAN Devices
Eth Devices
Serial Devices
Other Devices
Paolo Grisleri - University of Parma, Italy

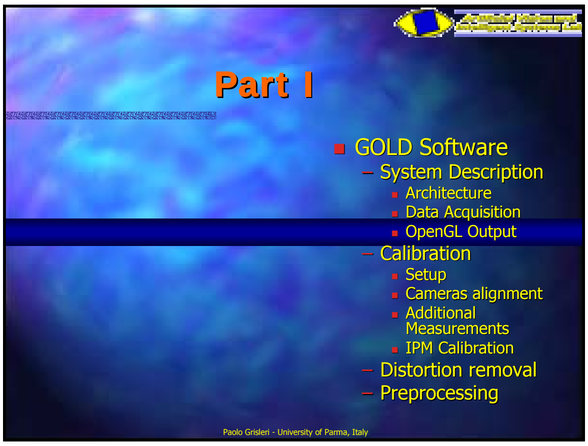

Part I
GOLD Software
System Description
Architecture
Data Acquisition
OpenGL Output
Calibration
Setup
Cameras alignment
Additional Measurements
IPM Calibration
Distortion removal
Preprocessing
Paolo Grisleri - University of Parma, Italy

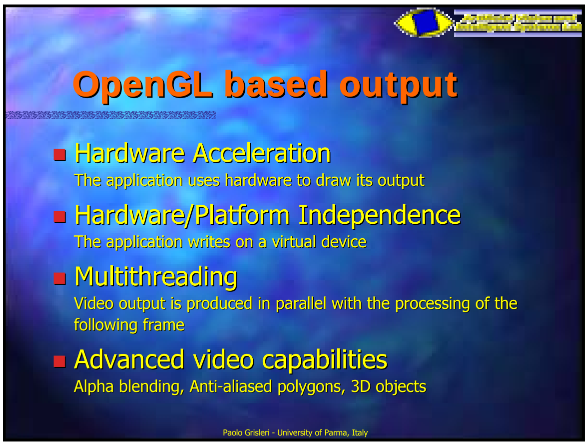

Interaction Application-Sensors
Application
Acquisition API
Driver Interface
Sensors (same type)
Data Recording
Each channel can read and write on:
Real Hardware
Virtual hardware
Real Hardware complexity hidden to application
Virtual Hardware plays data recorded from real hardware
Paolo Grisleri - University of Parma, Italy

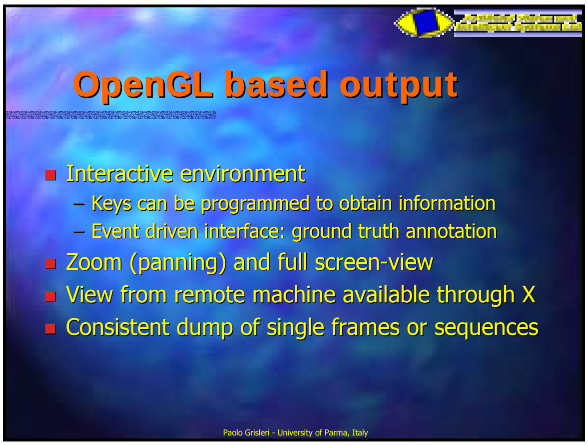

Generalized Interaction
Application
Data Acquisition
Sensors (Type 1)
Sensors (Type 2)
...
Sensori (Type n)
Data Recording
Paolo Grisleri - University of Parma, Italy

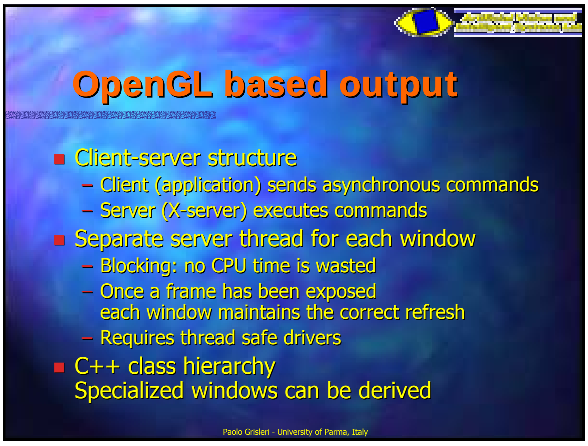

Video Acquisition
Video Acquisition API
Virtual Device
Acquisition Driver Interface
Source 1
Source 2
...
Source n
Events File
Images
Paolo Grisleri - University of Parma, Italy

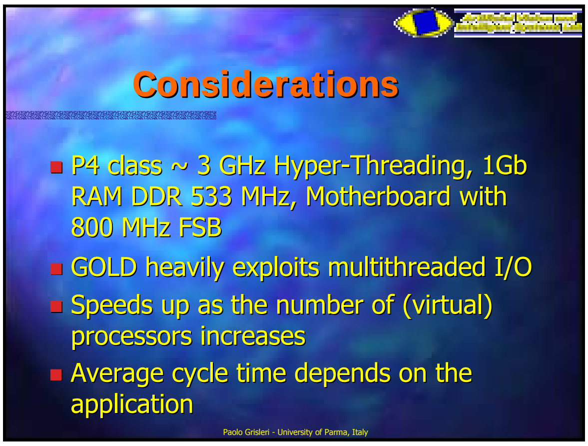

Video Acquisition
Phisical devices can be
– Groups of cameras (analog or digital)
– Image servers (even remote)
– Network file systems
Specific connections interfaces can be developed
– Ethernet
– Firewire
– CameraLink
– ...
Paolo Grisleri - University of Parma, Italy

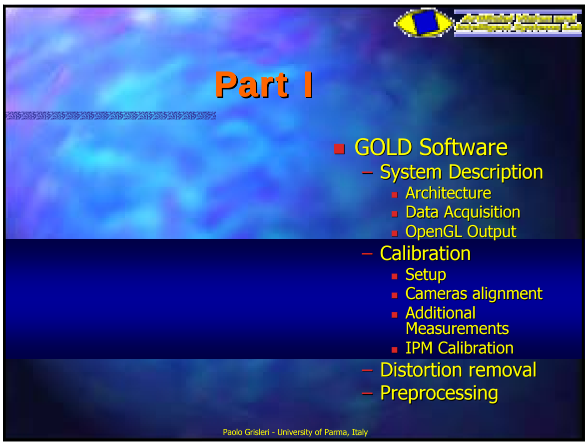

Part I
GOLD Software
– System Description
Architecture
Data Acquisition
OpenGL Output
– Calibration
Setup
Cameras alignment
Additional Measurements
IPM Calibration
– Distortion removal
– Preprocessing
Paolo Grisleri - University of Parma, Italy


OpenGL based output
Hardware Acceleration
The application uses hardware to draw its output
Hardware/Platform Independence
The application writes on a virtual device
Multithreading
Video output is produced in parallel with the processing of the following frame
Advanced video capabilities
Alpha blending, Anti-aliased polygons, 3D objects
Paolo Grisleri - University of Parma, Italy


OpenGL based output
Interactive environment
– Keys can be programmed to obtain information
– Event driven interface: ground truth annotation
Zoom (panning) and full screen-view
View from remote machine available through X
Consistent dump of single frames or sequences
Paolo Grisleri - University of Parma, Italy

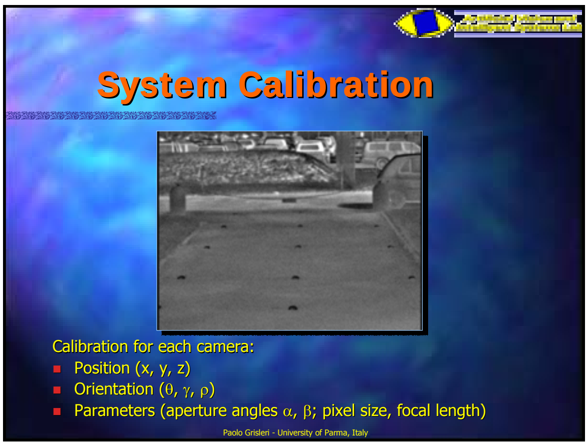

OpenGL based output
Client-server structure
– Client (application) sends asynchronous commands
– Server (X-server) executes commands
Separate server thread for each window
– Blocking: no CPU time is wasted
– Once a frame has been exposed each window maintains the correct refresh
– Requires thread safe drivers
C++ class hierarchy
Specialized windows can be derived
Paolo Grisleri - University of Parma, Italy

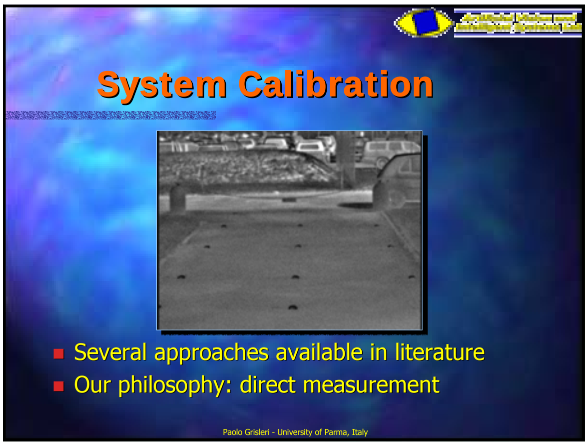

Considerations
P4 class ~ 3 GHz Hyper-Threading, 1Gb RAM DDR 533 MHz, Motherboard with 800 MHz FSB
GOLD heavily exploits multithreaded I/O
Speeds up as the number of (virtual) processors increases
Average cycle time depends on the application
Paolo Grisleri - University of Parma, Italy

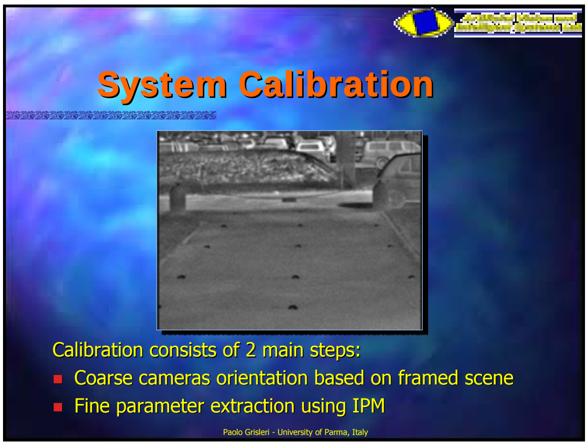

Part I
GOLD Software
– System Description
Architecture
Data Acquisition
OpenGL Output
– Calibration
Setup
Cameras alignment
Additional Measurements
IPM Calibration
– Distortion removal
– Preprocessing
Paolo Grisleri - University of Parma, Italy

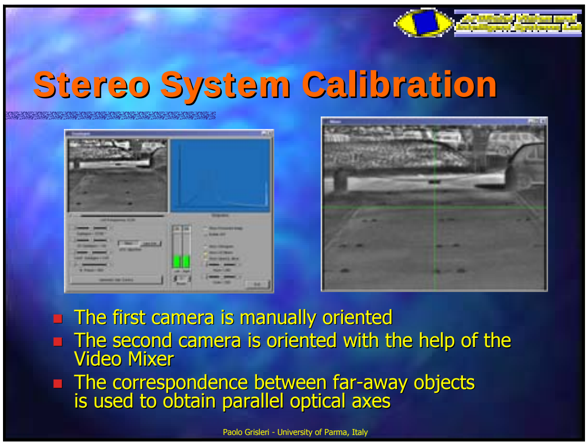

System Calibration
VIS, NIR, FIR
Calibration setup
Paolo Grisleri - University of Parma, Italy


System Calibration
Paolo Grisleri - University of Parma, Italy

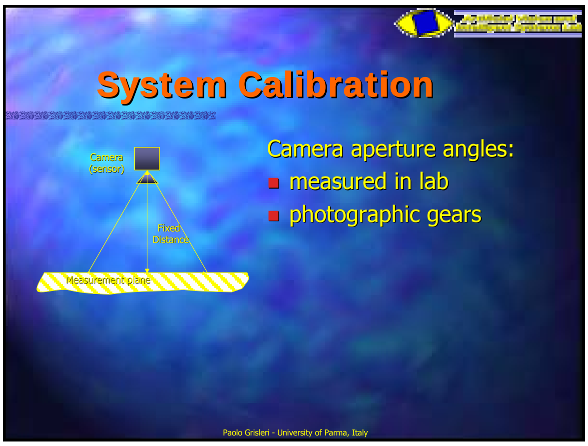

System Calibration
Calibration for each camera:
Position (x, y, z)
Orientation (θ, γ, ρ)
Parameters (aperture angles α, β; pixel size, focal length)
Paolo Grisleri - University of Parma, Italy

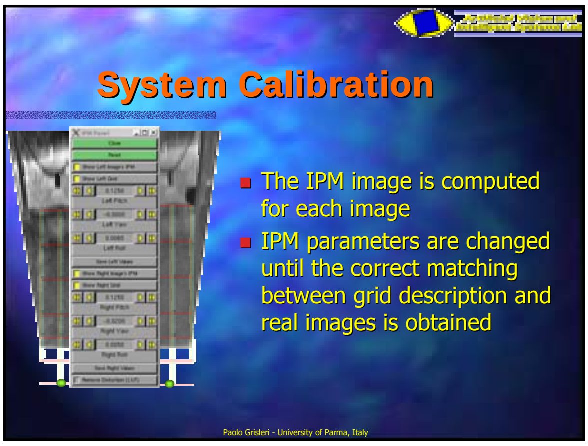

System Calibration
Several approaches available in literature
Our philosophy: direct measurement
Paolo Grisleri - University of Parma, Italy

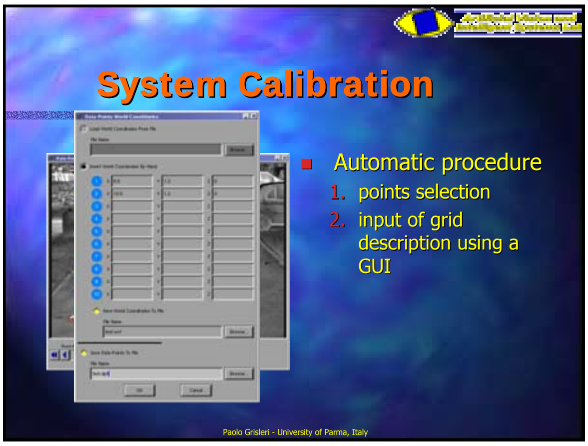

System Calibration
Calibration consists of 2 main steps:
Coarse cameras orientation based on framed scene
Fine parameter extraction using IPM
Paolo Grisleri - University of Parma, Italy

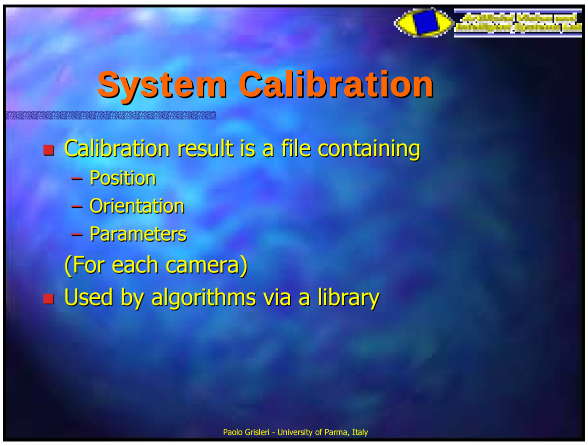

Stereo System Calibration
The first camera is manually oriented
The second camera is oriented with the help of the Video Mixer
The correspondence between far-away objects is used to obtain parallel optical axes
Paolo Grisleri - University of Parma, Italy

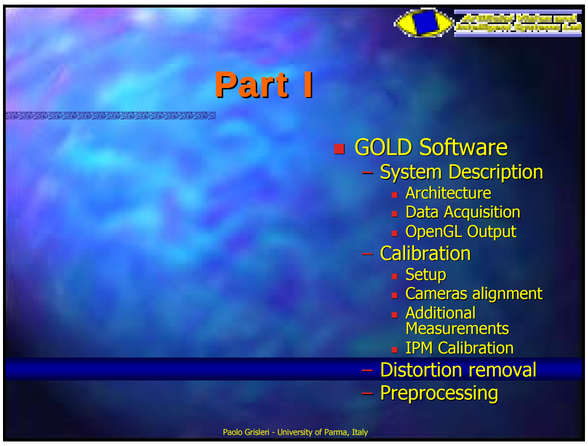

Video Stereo Calibration
Each orientation angle can be adjusted separately

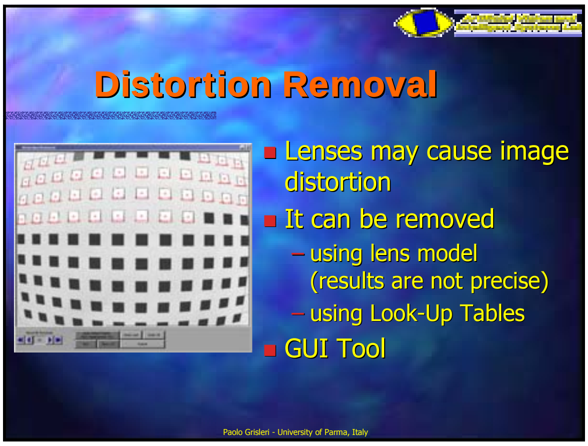

System Calibration
Camera (sensor)
Fixed Distance
Measurement plane
Camera aperture angles:
measured in lab
photographic gears
Paolo Grisleri - University of Parma, Italy

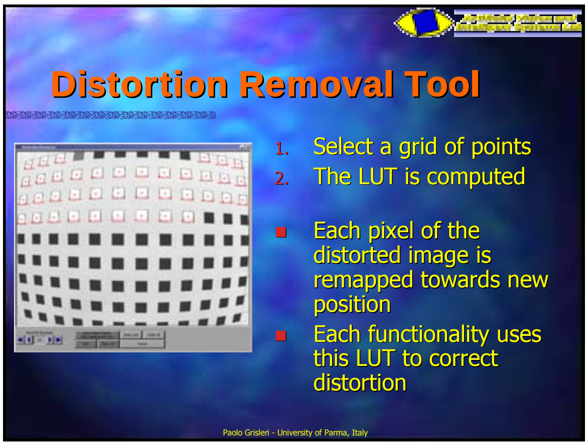

System Calibration
The IPM image is computed for each image
IPM parameters are changed until the correct matching between grid description and real images is obtained
Paolo Grisleri - University of Parma, Italy

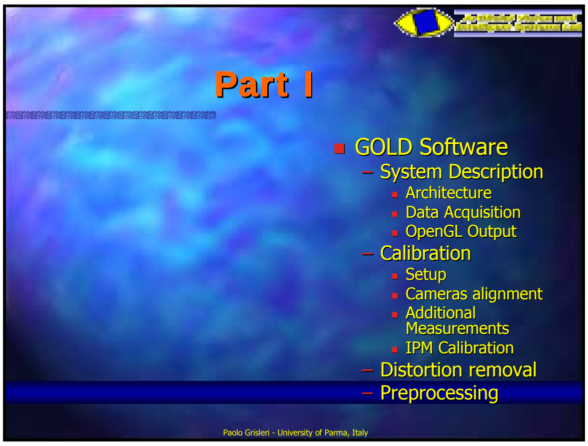

System Calibration
Automatic procedure
1. points selection
2. input of grid description using a GUI
Paolo Grisleri - University of Parma, Italy

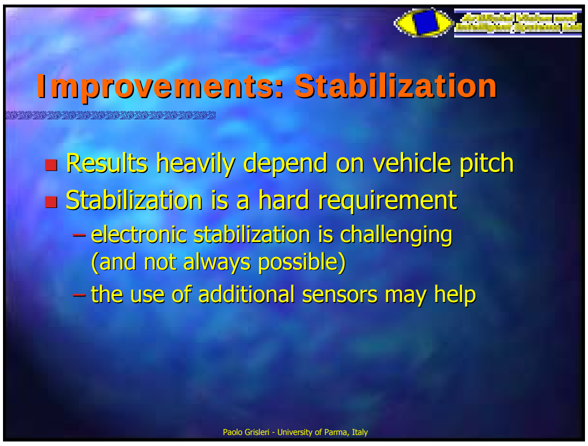

System Calibration
Calibration result is a file containing
– Position
– Orientation
– Parameters
(For each camera)
Used by algorithms via a library
Paolo Grisleri - University of Parma, Italy

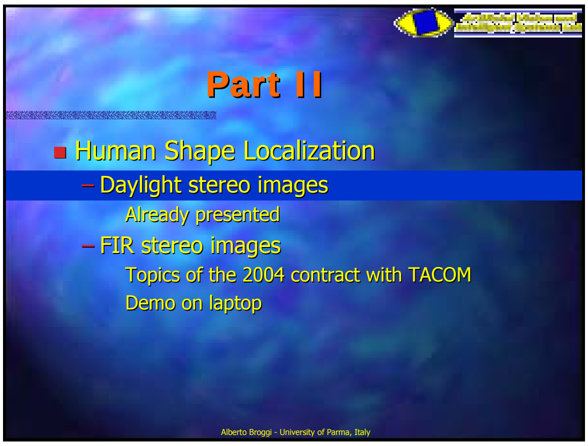

Part I
GOLD Software
– System Description
Architecture
Data Acquisition
OpenGL Output
– Calibration
Setup
Cameras alignment
Additional Measurements
IPM Calibration
– Distortion removal
– Preprocessing
Paolo Grisleri - University of Parma, Italy

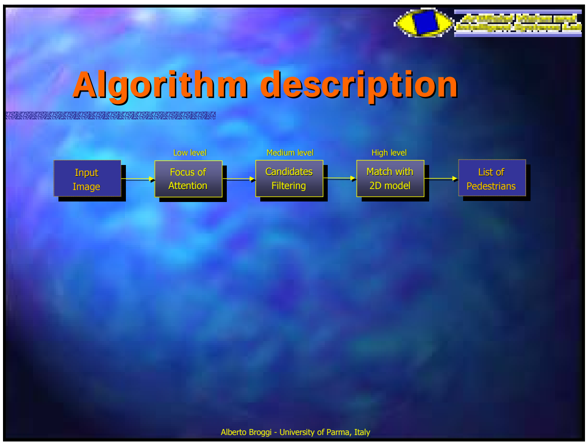

Distortion Removal
Lenses may cause image distortion
It can be removed
– using lens model (results are not precise)
– using Look-Up Tables
GUI Tool
Paolo Grisleri - University of Parma, Italy

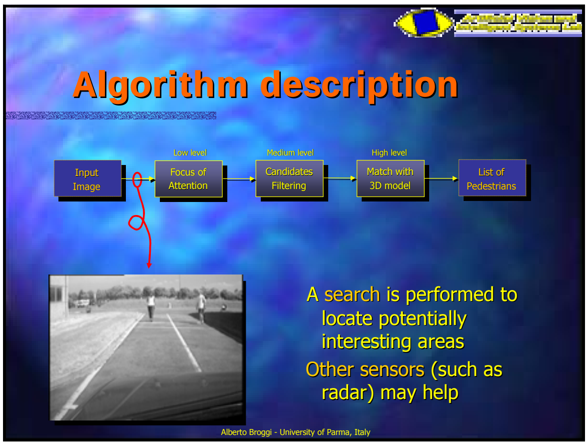

Distortion Removal Tool
1. Select a grid of points
2. The LUT is computed
Each pixel of the distorted image is remapped towards new position
Each functionality uses this LUT to correct distortion
Paolo Grisleri - University of Parma, Italy

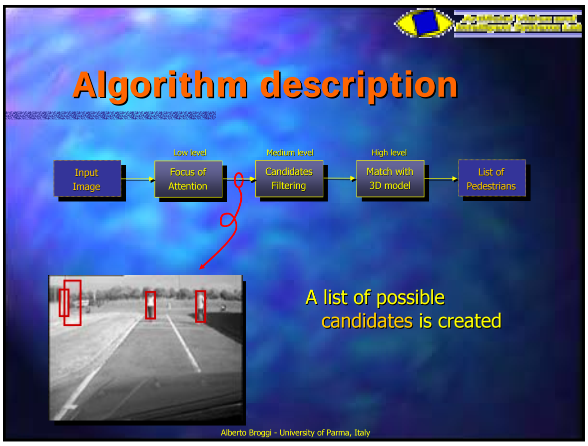

Part I
GOLD Software
– System Description
Architecture
Data Acquisition
OpenGL Output
– Calibration
Setup
Cameras alignment
Additional Measurements
IPM Calibration
– Distortion removal
– Preprocessing
Paolo Grisleri - University of Parma, Italy

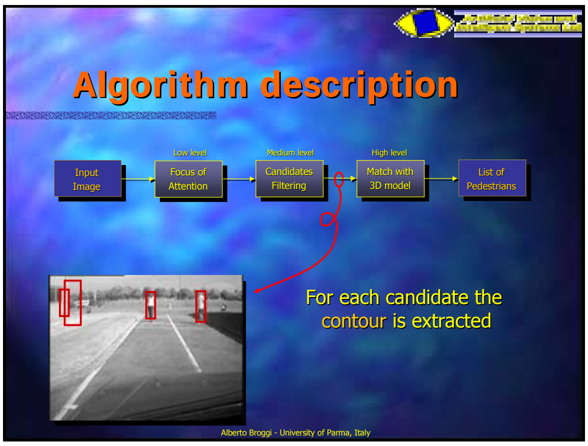

Improvements: Stabilization
Results heavily depend on vehicle pitch
Stabilization is a hard requirement
– electronic stabilization is challenging (and not always possible)
– the use of additional sensors may help
Paolo Grisleri - University of Parma, Italy

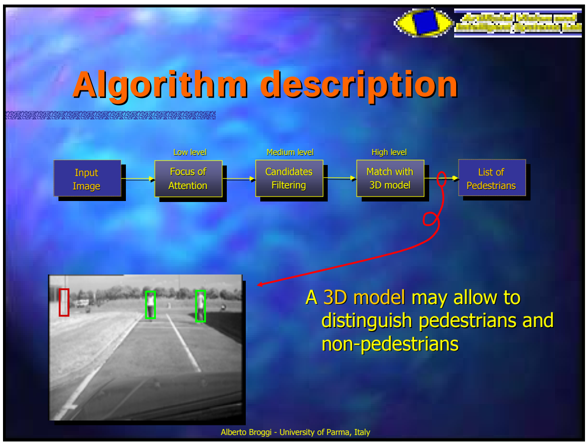

Part II
Human Shape Localization
– Daylight stereo images
Already presented
– FIR stereo images
Topics of the 2004 contract with TACOM
Demo on laptop
Alberto Broggi - University of Parma, Italy


Algorithm description
Input Image
Low level
Focus of Attention
Medium level
Candidates Filtering
High level
Match with 2D model
List of Pedestrians
Alberto Broggi - University of Parma, Italy


Algorithm description
Input Image
Low level
Focus of Attention
Medium level
Candidates Filtering
High level
Match with 3D model
List of Pedestrians
A search is performed to locate potentially interesting areas
Other sensors (such as radar) may help
Alberto Broggi - University of Parma, Italy


Algorithm description
Input Image
Low level
Focus of Attention
Medium level
Candidates Filtering
High level
Match with 3D model
List of Pedestrians
A list of possible candidates is created
Alberto Broggi - University of Parma, Italy

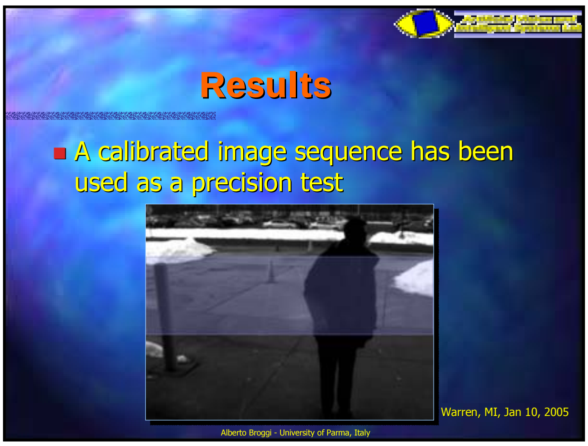

Algorithm description
Input Image
Low level
Focus of Attention
Medium level
Candidates Filtering
High level
Match with 3D model
List of Pedestrians
For each candidate the contour is extracted
Alberto Broggi - University of Parma, Italy


Algorithm description
Low level
Medium level
High level
Input Image
Focus of Attention
Candidates Filtering
Match with 3D model
List of Pedestrians
A 3D model may allow to distinguish pedestrians and non-pedestrians
Alberto Broggi - University of Parma, Italy

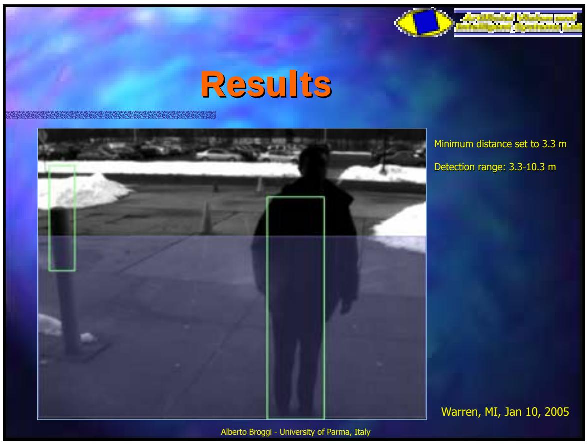

The Low Level
Input Image
Grey level Symmetry
Vertical Edge Symmetry
Vertical Edge Extraction
Preattentive Filter
Vertical Edge Density
Linear Combination
Maxima Detection
List of BBs
Alberto Broggi - University of Parma, Italy

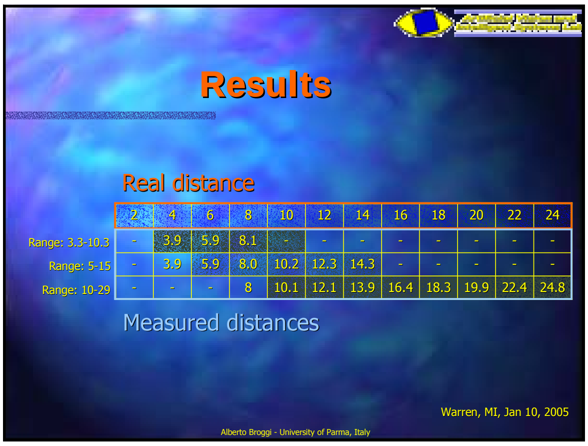

TARDEC Installation
24 meters-long grid
Warren, MI, Jan 10, 2005
Alberto Broggi - University of Parma, Italy


TARDEC Installation
Cameras installed on a HMMWV
Warren, MI, Jan 10, 2005
Alberto Broggi - University of Parma, Italy

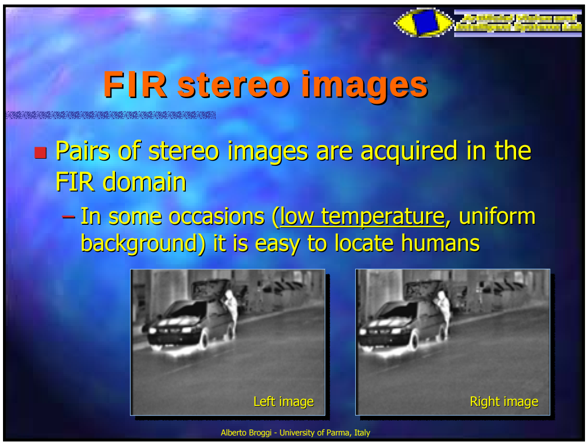

Results
A calibrated image sequence has been used as a precision test
Warren, MI, Jan 10, 2005
Alberto Broggi - University of Parma, Italy

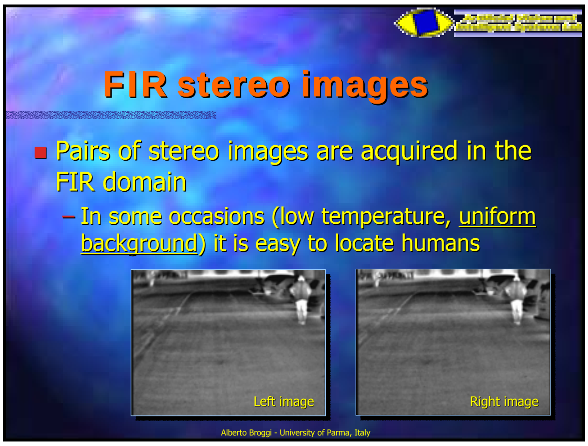

Results
Minimum distance set to 10 m
Detection range: 10-29 m
Warren, MI, Jan 10, 2005
Alberto Broggi - University of Parma, Italy

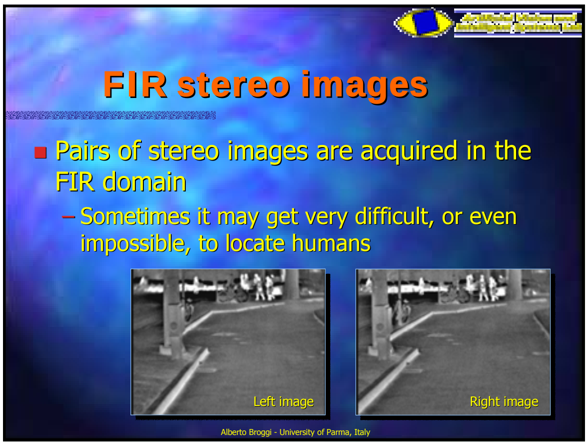


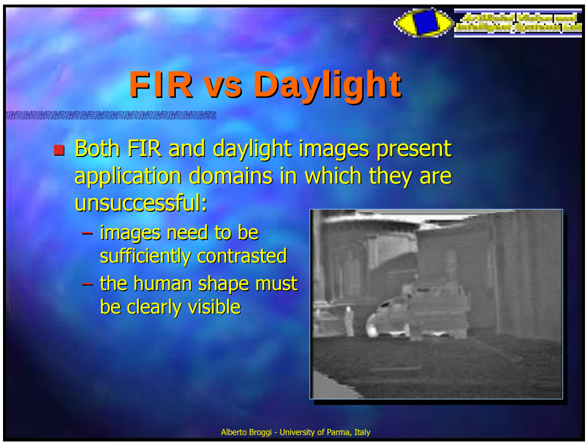


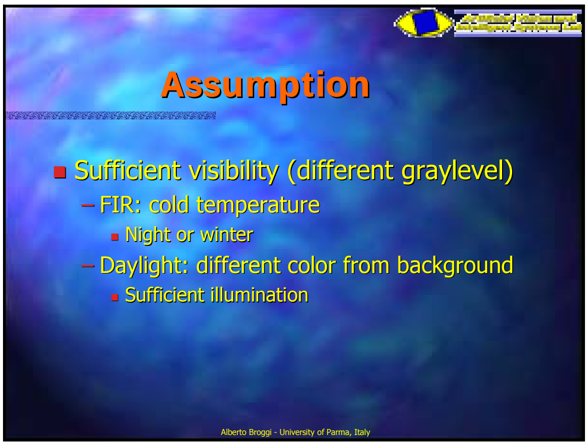
	2	4	6	8	10	12	14	16	18	20	22	24
Range: 3.3-10.3	-	3.9	5.9	8.1	-	-	-	-	-	-	-	-
Range: 5-15	-	3.9	5.9	8.0	10.2	12.3	14.3	-	-	-	-	-
Range: 10-29	-	-	-	8	10.1	12.1	13.9	16.4	18.3	19.9	22.4	24.8

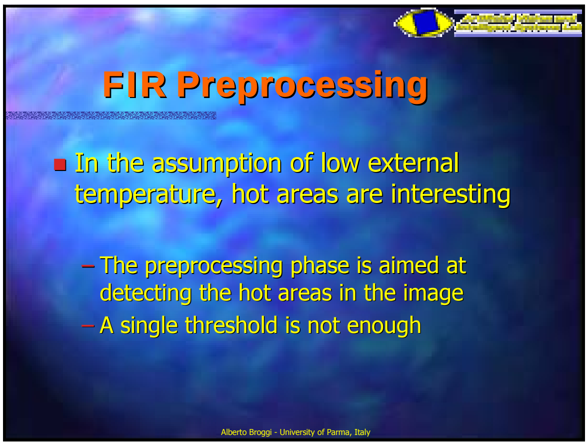


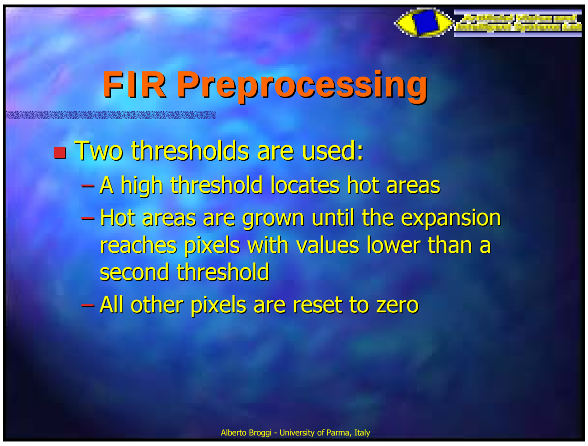


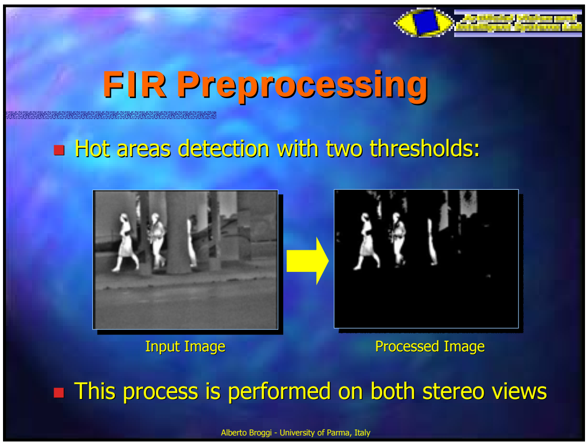

FIR stereo images
Pairs of stereo images are acquired in the FIR domain
– Sometimes it may get very difficult, or even impossible, to locate humans
Left image
Right image
Alberto Broggi - University of Parma, Italy

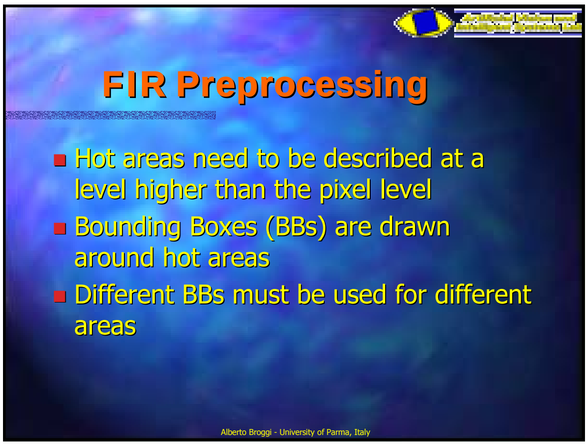

FIR vs Daylight
Both FIR and daylight images present application domains in which they are unsuccessful:
– images need to be sufficiently contrasted
– the human shape must be clearly visible
Alberto Broggi - University of Parma, Italy


Assumption
Sufficient visibility (different graylevel)
– FIR: cold temperature
Night or winter
– Daylight: different color from background
Sufficient illumination
Alberto Broggi - University of Parma, Italy

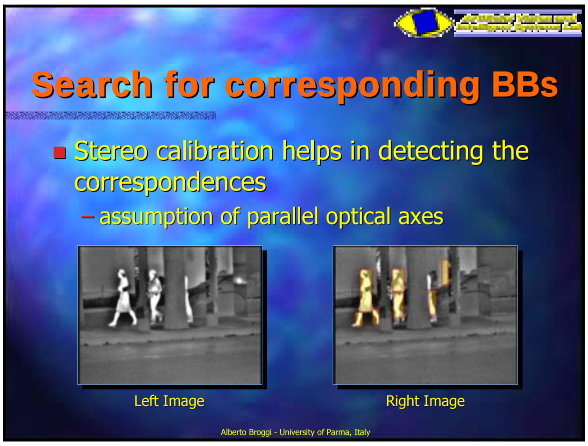

FIR Preprocessing
In the assumption of low external temperature, hot areas are interesting
– The preprocessing phase is aimed at detecting the hot areas in the image
– A single threshold is not enough
Alberto Broggi - University of Parma, Italy

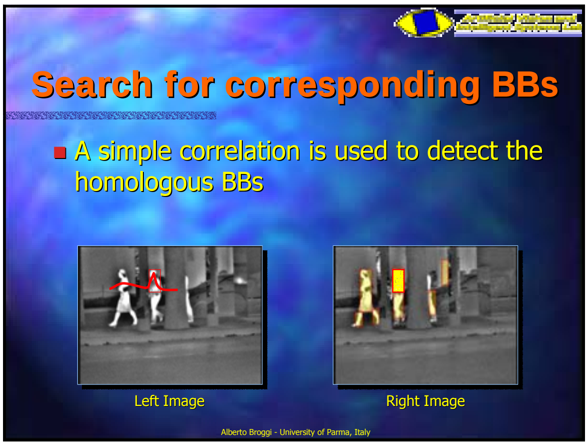

FIR Preprocessing
Two thresholds are used:
– A high threshold locates hot areas
– Hot areas are grown until the expansion reaches pixels with values lower than a second threshold
– All other pixels are reset to zero
Alberto Broggi - University of Parma, Italy

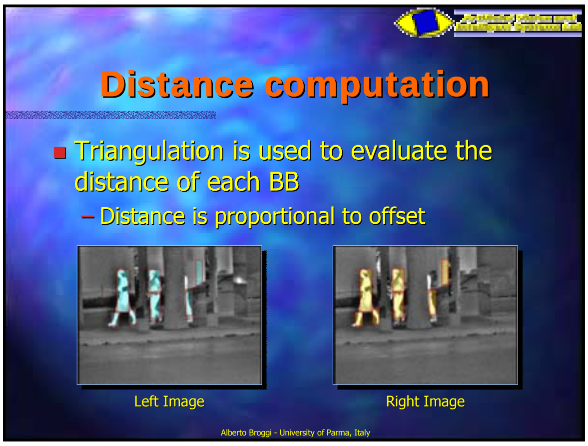

FIR Preprocessing
Hot areas detection with two thresholds:
Input Image
Processed Image
This process is performed on both stereo views
Alberto Broggi - University of Parma, Italy

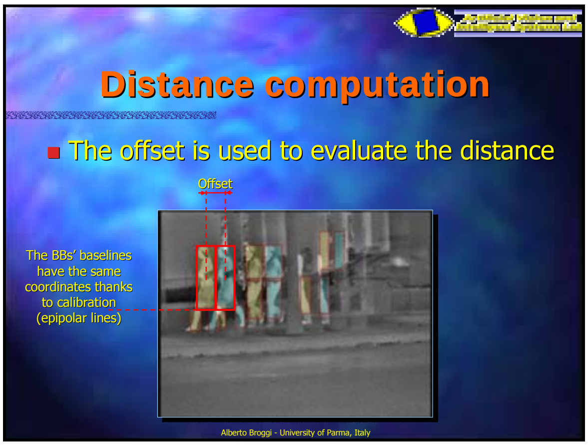

FIR Preprocessing
Hot areas need to be described at a level higher than the pixel level
Bounding Boxes (BBs) are drawn around hot areas
Different BBs must be used for different areas
Alberto Broggi - University of Parma, Italy

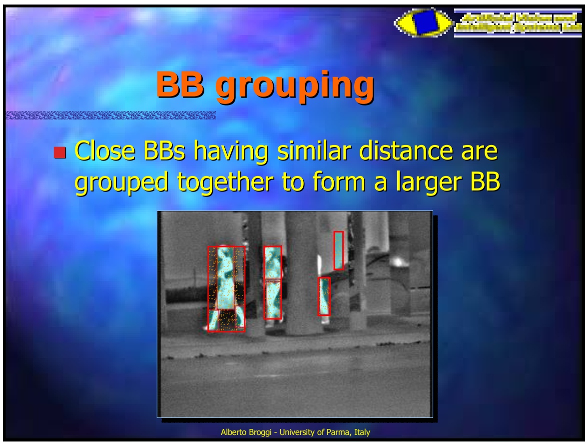

FIR Preprocessing
An iterative process based on histograms is used
Alberto Broggi - University of Parma, Italy

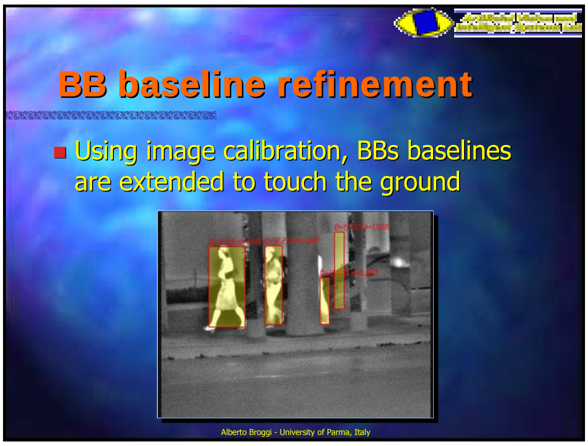

Search for corresponding BBs
Stereo calibration helps in detecting the correspondences
– assumption of parallel optical axes
Left Image
Right Image
Alberto Broggi - University of Parma, Italy

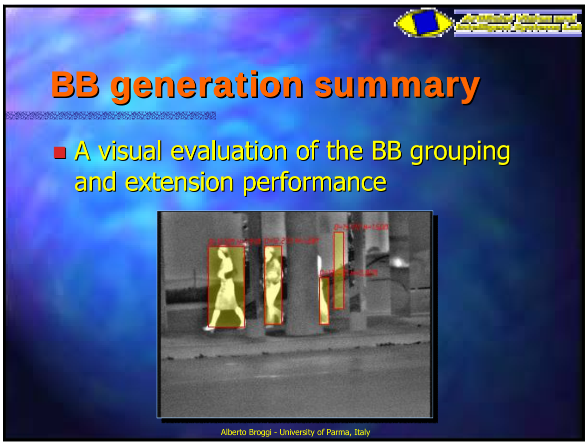

Search for corresponding BBs
A simple correlation is used to detect the homologous BBs
Left Image
Right Image
Alberto Broggi - University of Parma, Italy

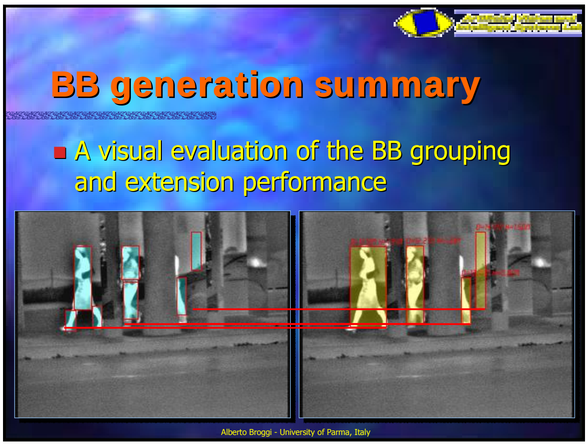

Distance computation
Triangulation is used to evaluate the distance of each BB
– Distance is proportional to offset
Left Image
Right Image
Alberto Broggi - University of Parma, Italy

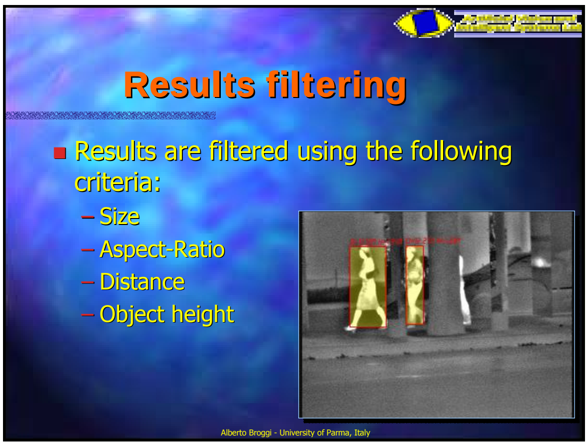

Distance computation
The offset is used to evaluate the distance
Offset
The BBs' baselines have the same coordinates thanks to calibration (epipolar lines)
Alberto Broggi - University of Parma, Italy

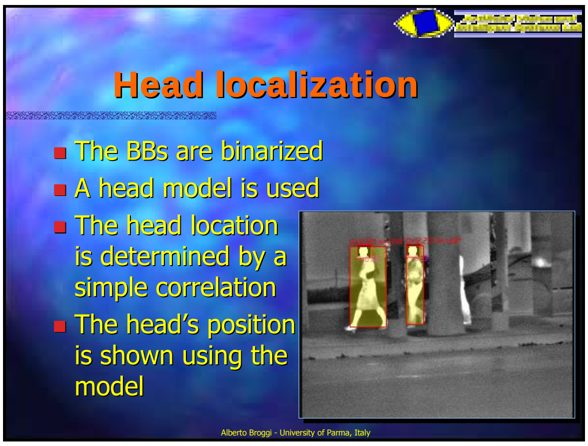

BB grouping
Close BBs having similar distance are grouped together to form a larger BB
Alberto Broggi - University of Parma, Italy

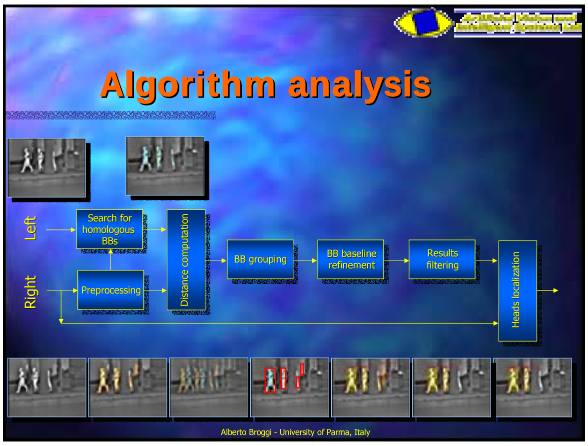

BB baseline refinement
Using image calibration, BBs baselines are extended to touch the ground
Alberto Broggi - University of Parma, Italy


BB generation summary
A visual evaluation of the BB grouping and extension performance
Alberto Broggi - University of Parma, Italy

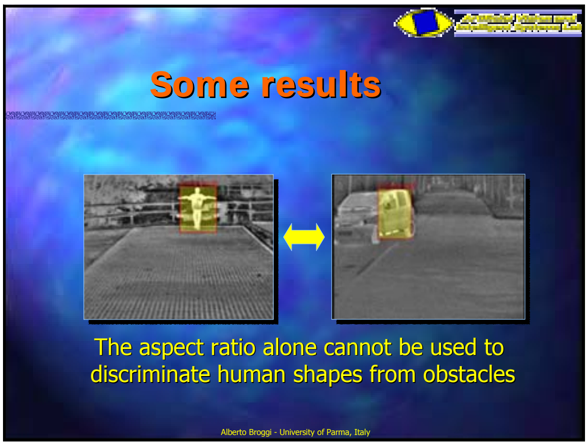

BB generation summary
A visual evaluation of the BB grouping and extension performance
Alberto Broggi - University of Parma, Italy


Results filtering
Results are filtered using the following criteria:
– Size
– Aspect-Ratio
– Distance
– Object height
Alberto Broggi - University of Parma, Italy

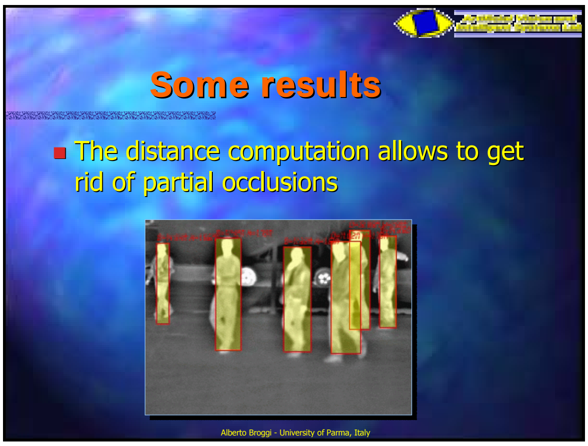

Head localization
The BBs are binarized
A head model is used
The head location is determined by a simple correlation
The head's position is shown using the model
Alberto Broggi - University of Parma, Italy


Algorithm analysis
Left
Right
Search for homologous BBs
Preprocessing
Distance computation
BB grouping
BB baseline refinement
Results filtering
Heads localization
Alberto Broggi - University of Parma, Italy


Some results
Alberto Broggi - University of Parma, Italy


Some results
The aspect ratio alone cannot be used to discriminate human shapes from obstacles
Alberto Broggi - University of Parma, Italy

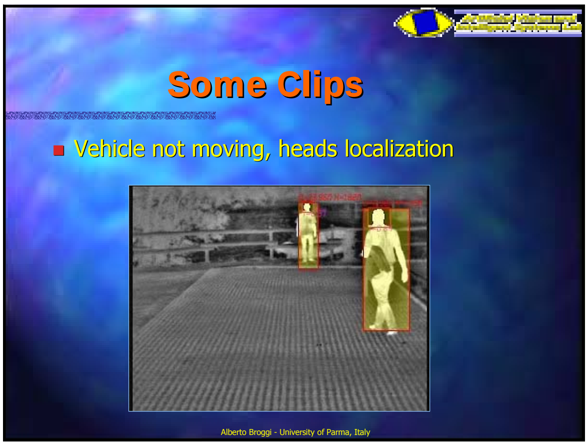

Some Results
Too small Bounding Boxes:
Left Image
Right Image
Alberto Broggi - University of Parma, Italy

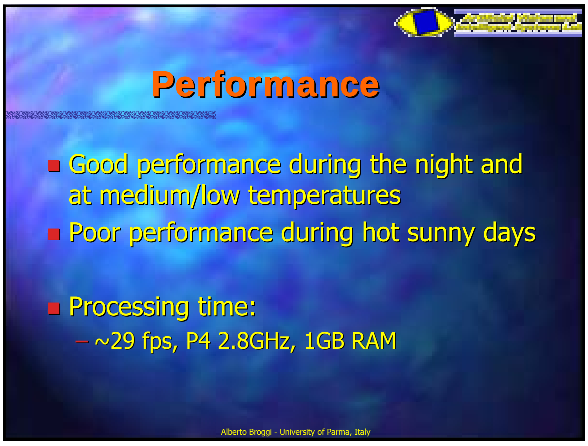

Some results
The distance computation allows to get rid of partial occlusions
Alberto Broggi - University of Parma, Italy

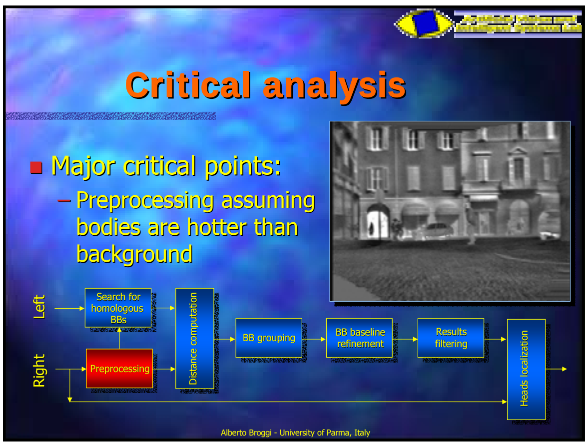

Some Clips
Vehicle not moving, 3 pedestrians
Alberto Broggi - University of Parma, Italy

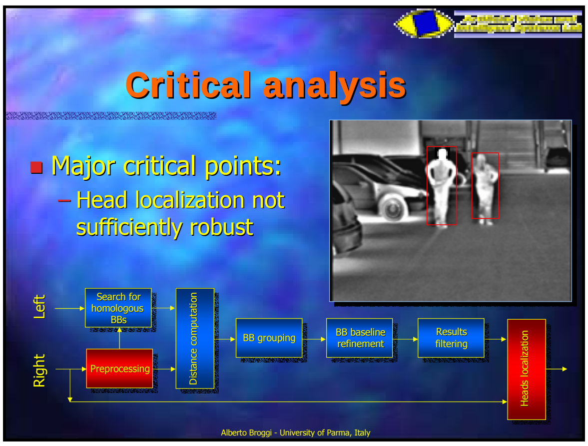

Some Clips
Vehicle moving, many pedestrians
Alberto Broggi - University of Parma, Italy


Some Clips
Vehicle not moving, occlusions
Alberto Broggi - University of Parma, Italy


Some Clips
Vehicle not moving, heads localization
Alberto Broggi - University of Parma, Italy

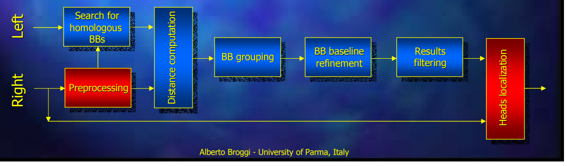

Performance
Good performance during the night and at medium/low temperatures
Poor performance during hot sunny days
Processing time:
– ~29 fps, P4 2.8GHz, 1GB RAM
Alberto Broggi - University of Parma, Italy


Critical analysis
Major critical points:
– Preprocessing assuming bodies are hotter than background
Left
Right
Search for homologous BBs
Preprocessing
Distance computation
BB grouping
BB baseline refinement
Results filtering
Heads localization
Alberto Broggi - University of Parma, Italy


Critical analysis
Major critical points:
– Head localization not sufficiently robust
Left
Right
Search for homologous BBs
Preprocessing
Distance computation
BB grouping
BB baseline refinement
Results filtering
Heads localization
Alberto Broggi - University of Parma, Italy

Critical analysis

- Current research is focused on algorithm improvement on high-temperature images
  - New method to generate BBs
  - New head localization approach



Alberto Broggi - University of Parma, Italy

New methods to generate BBs

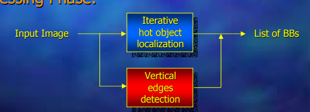
- The following methods are under evaluation:
  - Based on vertical edges (like in daylight image processing)
  - Based on stereo obstacles detection (using the experience developed in the DARPA Grand Challenge '04)

Alberto Broggi - University of Parma, Italy

New methods to generate BBs

- The following methods are under evaluation:
  - Based on vertical edges (like in daylight image processing)

The Preprocessing Phase:



Alberto Broggi - University of Parma, Italy

Method based on Vertical edges

- The approach:
  - Vertical edge detection
  - Morphological dilation
  - BB generation
  - Removal of BBs based on
    - size and
    - average internal temperature

Alberto Broggi - University of Parma, Italy

Method based on Vertical edges

- Intermediate results:

Original — Vertical Edges — Dilated edges + BBs

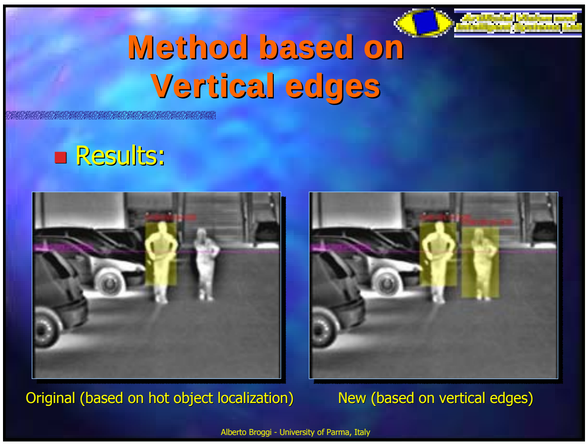
Alberto Broggi - University of Parma, Italy

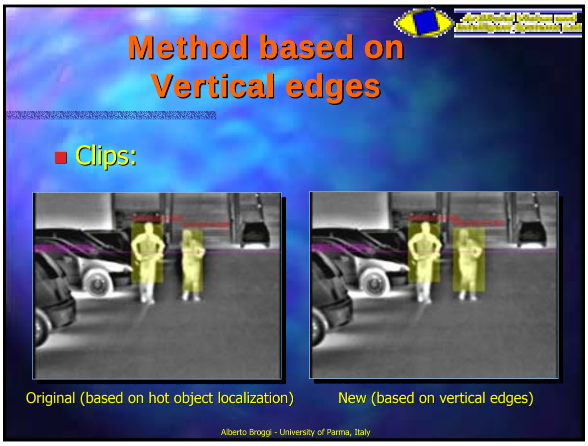
Method based on Vertical edges

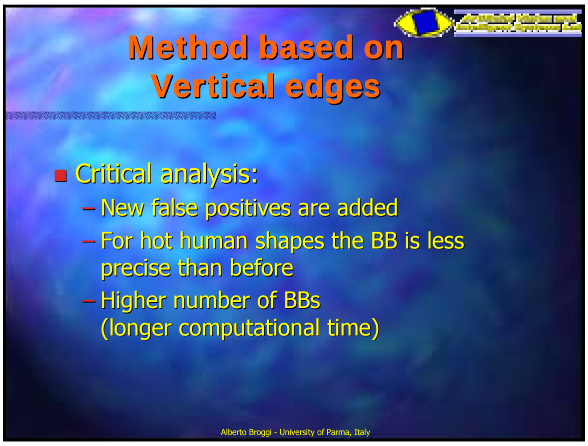
- Intermediate results:

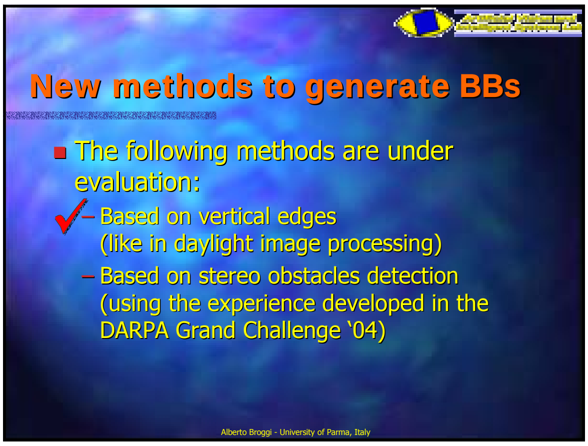
Original (based on hot object localization) — New (based on vertical edges)

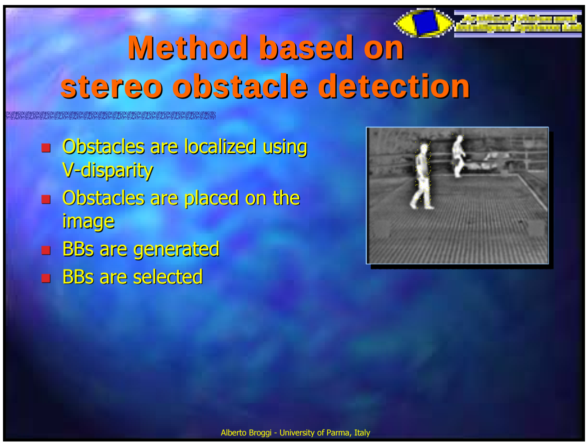
Alberto Broggi - University of Parma, Italy

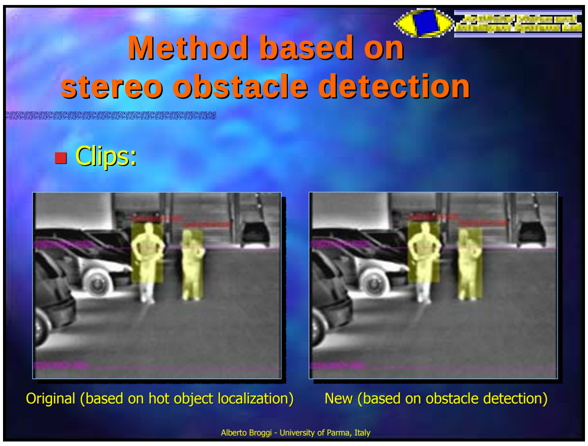

Method based on
Vertical edges
Results:
Original (based on hot object localization)
New (based on vertical edges)
Alberto Broggi - University of Parma, Italy

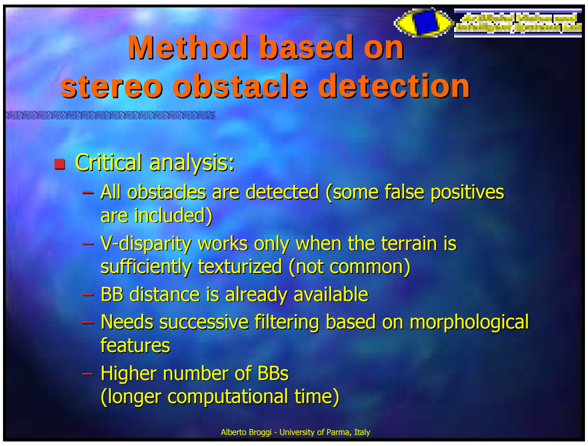

Method based on
Vertical edges
Clips:
Original (based on hot object localization)
New (based on vertical edges)
Alberto Broggi - University of Parma, Italy

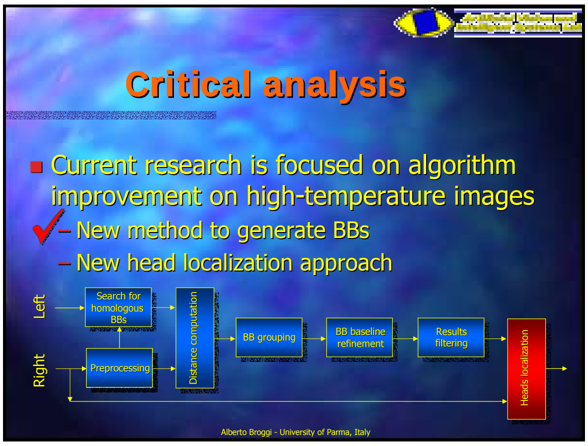

Method based on
Vertical edges
Critical analysis:
– New false positives are added
– For hot human shapes the BB is less precise than before
– Higher number of BBs (longer computational time)
Alberto Broggi - University of Parma, Italy

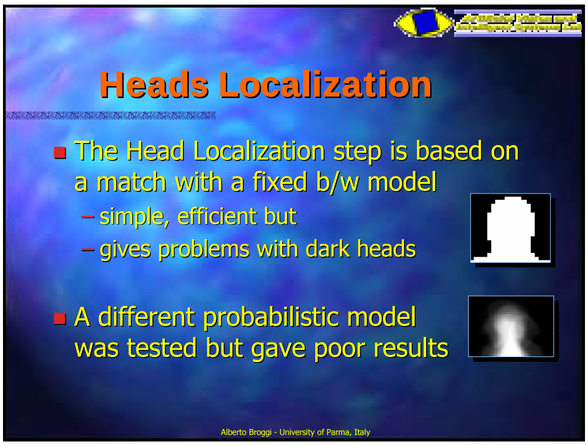

New methods to generate BBs
The following methods are under evaluation:
– Based on vertical edges (like in daylight image processing)
– Based on stereo obstacles detection (using the experience developed in the DARPA Grand Challenge '04)
Alberto Broggi - University of Parma, Italy

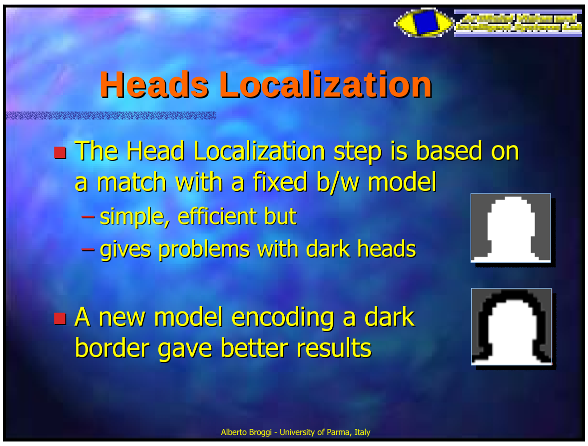

Method based on
stereo obstacle detection
Obstacles are localized using V-disparity
Obstacles are placed on the image
BBs are generated
BBs are selected
Alberto Broggi - University of Parma, Italy

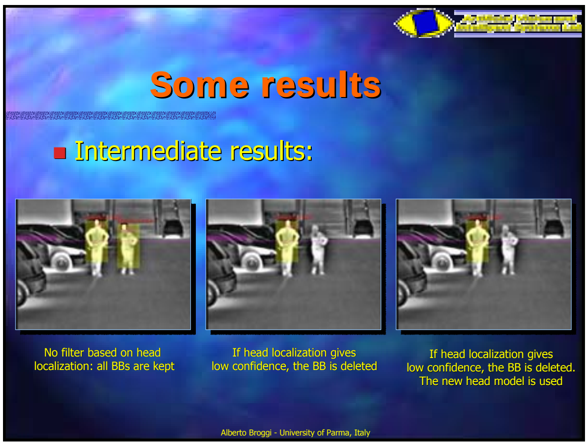

Method based on
stereo obstacle detection
Clips:
Original (based on hot object localization)
New (based on obstacle detection)
Alberto Broggi - University of Parma, Italy

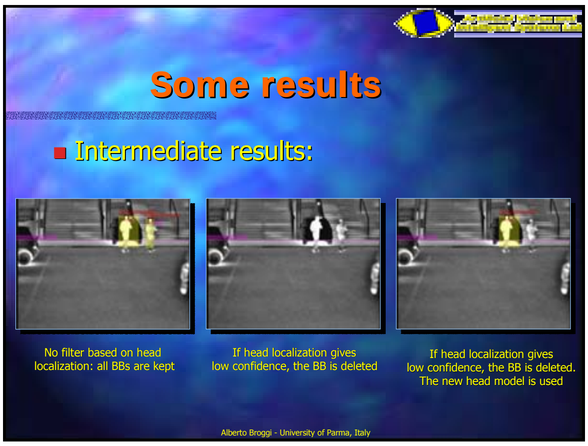

Method based on
stereo obstacle detection
Critical analysis:
– All obstacles are detected (some false positives are included)
– V-disparity works only when the terrain is sufficiently texturized (not common)
– BB distance is already available
– Needs successive filtering based on morphological features
– Higher number of BBs (longer computational time)
Alberto Broggi - University of Parma, Italy

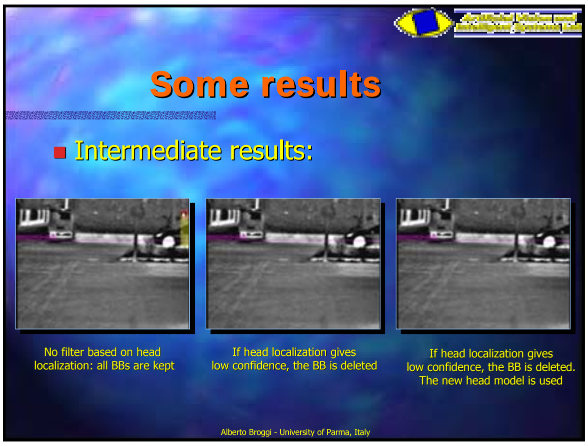

Critical analysis
Current research is focused on algorithm improvement on high-temperature images
– New method to generate BBs
– New head localization approach
Left
Right
Search for homologous BBs
Preprocessing
Distance computation
BB grouping
BB baseline refinement
Results filtering
Heads localization
Alberto Broggi - University of Parma, Italy

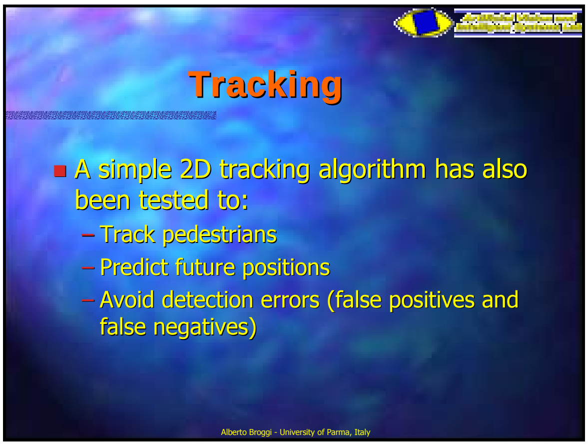

Heads Localization
The Head Localization step is based on a match with a fixed b/w model
– simple, efficient but
– gives problems with dark heads
A different probabilistic model was tested but gave poor results
Alberto Broggi - University of Parma, Italy

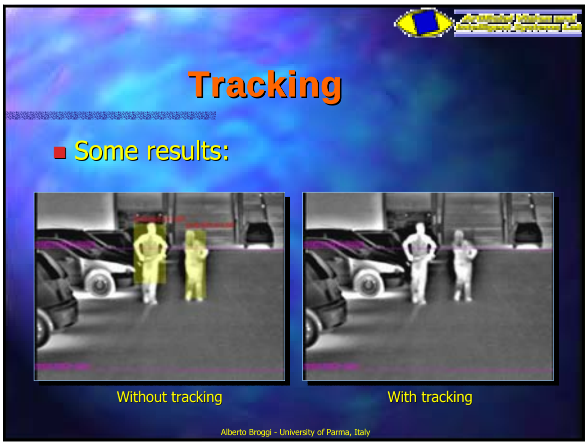

Heads Localization
The Head Localization step is based on a match with a fixed b/w model
– simple, efficient but
– gives problems with dark heads
A new model encoding a dark border gave better results
Alberto Broggi - University of Parma, Italy

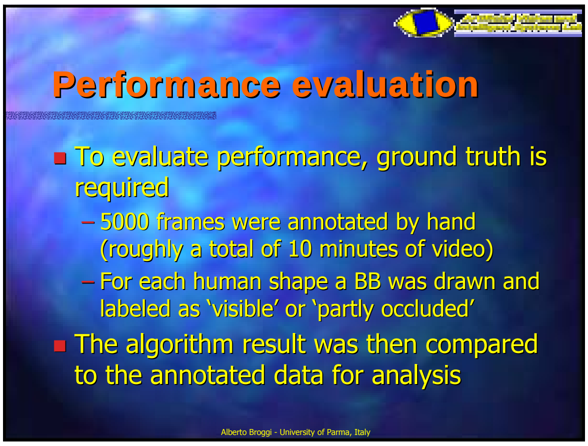

Some results
Intermediate results:
No filter based on head localization: all BBs are kept
If head localization gives low confidence, the BB is deleted
If head localization gives low confidence, the BB is deleted. The new head model is used
Alberto Broggi - University of Parma, Italy

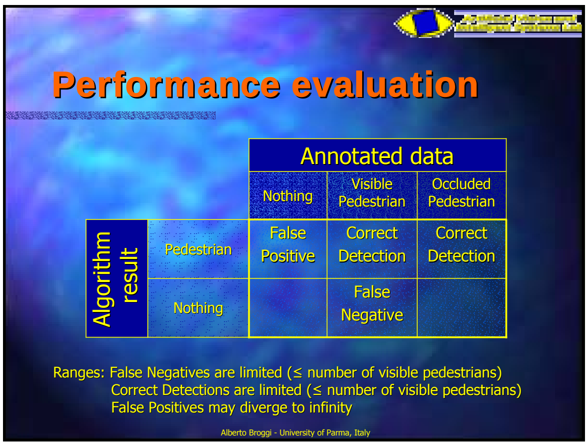

Some results
Intermediate results:
No filter based on head localization: all BBs are kept
If head localization gives low confidence, the BB is deleted
If head localization gives low confidence, the BB is deleted. The new head model is used
Alberto Broggi - University of Parma, Italy

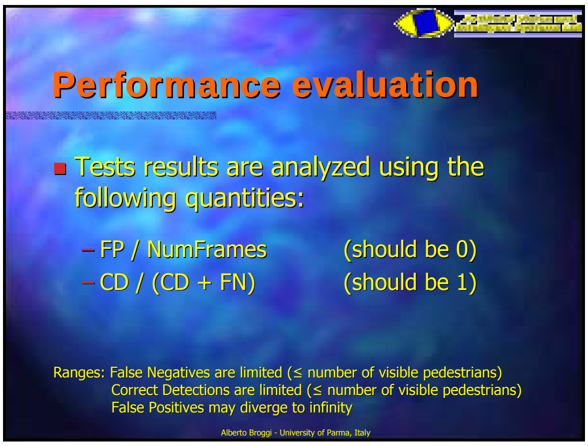

Some results
Intermediate results:
No filter based on head localization: all BBs are kept
If head localization gives low confidence, the BB is deleted
If head localization gives low confidence, the BB is deleted. The new head model is used
Alberto Broggi - University of Parma, Italy


Tracking
A simple 2D tracking algorithm has also been tested to:
– Track pedestrians
– Predict future positions
– Avoid detection errors (false positives and false negatives)
Alberto Broggi - University of Parma, Italy


Tracking
Some results:
Without tracking
With tracking
Alberto Broggi - University of Parma, Italy


Performance evaluation
To evaluate performance, ground truth is required
– 5000 frames were annotated by hand (roughly a total of 10 minutes of video)
– For each human shape a BB was drawn and labeled as 'visible' or 'partly occluded'
The algorithm result was then compared to the annotated data for analysis
Alberto Broggi - University of Parma, Italy


Performance evaluation
Annotated data
Nothing
Visible Pedestrian
Occluded Pedestrian
Algorithm result
Pedestrian
False Positive
Correct Detection
Correct Detection
Nothing
False Negative
Ranges: False Negatives are limited (≤ number of visible pedestrians)
Correct Detections are limited (≤ number of visible pedestrians)
False Positives may diverge to infinity
Alberto Broggi - University of Parma, Italy


Performance evaluation
Tests results are analyzed using the following quantities:
– FP / NumFrames (should be 0)
– CD / (CD + FN) (should be 1)
Ranges: False Negatives are limited (≤ number of visible pedestrians)
Correct Detections are limited (≤ number of visible pedestrians)
False Positives may diverge to infinity
Alberto Broggi - University of Parma, Italy

Performance evaluation

- Tests with the old ( ) and new ( ) head models

Algorithm	FP/NumFrames	CD/(CD+FN)
Original model	0.07	0.69
New model	0.09	0.74

Alberto Broggi - University of Parma, Italy

Performance evaluation

Algorithm Enhancement		FP/NumFrames	CD/(CD+FN)
	Edge	0.08	0.80
	Obstacles	0.11	0.78
	Edge + Tracking	0.09	0.81
	Obstacles + Tracking	0.11	0.80
	Edge + Obstacles + Tracking	0.11	0.81

Alberto Broggi - University of Parma, Italy

Part II

- Human Shape Localization
  - Daylight stereo images
    - Already presented
  - FIR stereo images
    - Topics of the 2004 contract with TACOM
    - Demo on laptop

Alberto Broggi - University of Parma, Italy

Part III

- Conclusions:
  - Two systems have been developed
  - Further work is required to include:
    - Tracking and motion
    - In daylight processing:
      - Color processing
      - 3D or 2D human shape model (high level processing)
    - In FIR processing:
      - Human shape model
  - Fusion

Alberto Broggi - University of Parma, Italy