Perceptive Context for Pervasive Computing

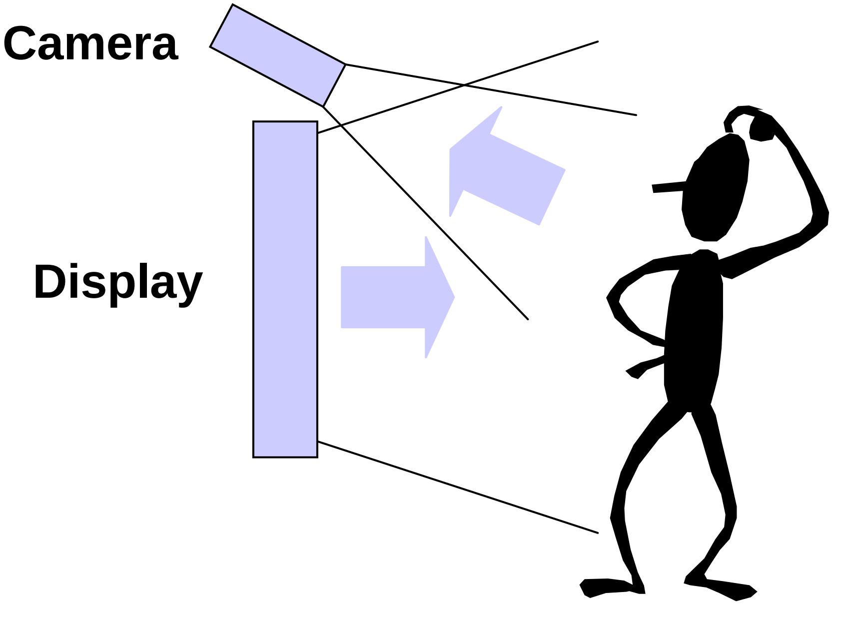
Trevor Darrell
Vision Interface Group
MIT AI Lab

Perceptually Aware Displays

Camera associated with display

Display should respond to user

- font size
- attentional load
- passive acknowledgement



e.g., “Magic Mirror”, Interval

Compaq’s Smart Kiosk

ALIVE, MIT Media Lab

Example: A Face Responsive Display

- Faces are natural interfaces!
  - Ubiquitous, fast, expressive, general.
  - Want machines to generate and perceive faces.
- A Face Responsive Display...
  - Knows when it's being observed
  - Recognizes returning observers
  - Tracks head pose
  - Robust to changing lighting, moving backgrounds…

A Face Responsive Display

Tasks

- Detection
- Identification
- Tracking

How? Exploit multiple visual modalities:

- Shape
- Color
- Pattern

Tasks and Visual Modalities

	shape	***color***	***pattern***
detection	silhouette classifier	skin classifier	face detection
identification	biometrics	flesh hue	face recognition
tracking	coarse motion estimation	clothing histogram	fine motion estimation / pose tracking

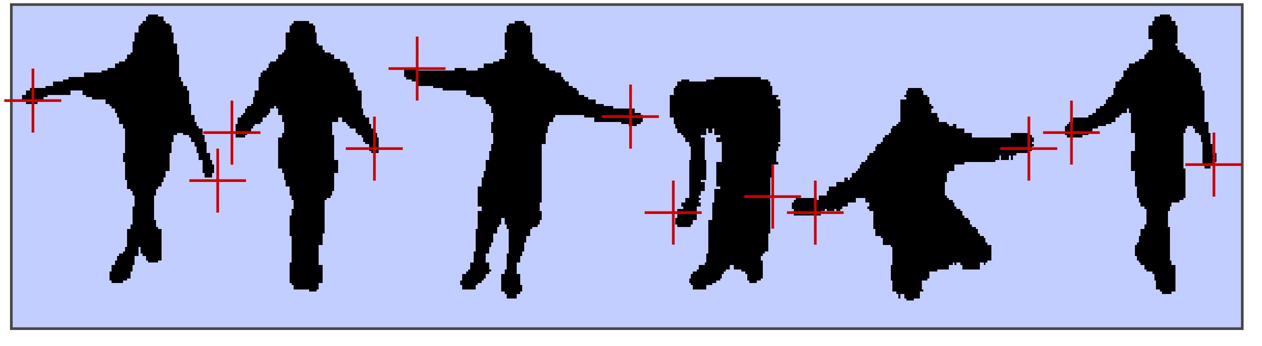
Mode and Task Matrix

	shape	***color***	***pattern***
detection	silhouette classifier	skin classifier	face detection
identification	biometrics	flesh hue	face recognition
tracking	Shape change	clothing histogram	Appearance change

Finding Features

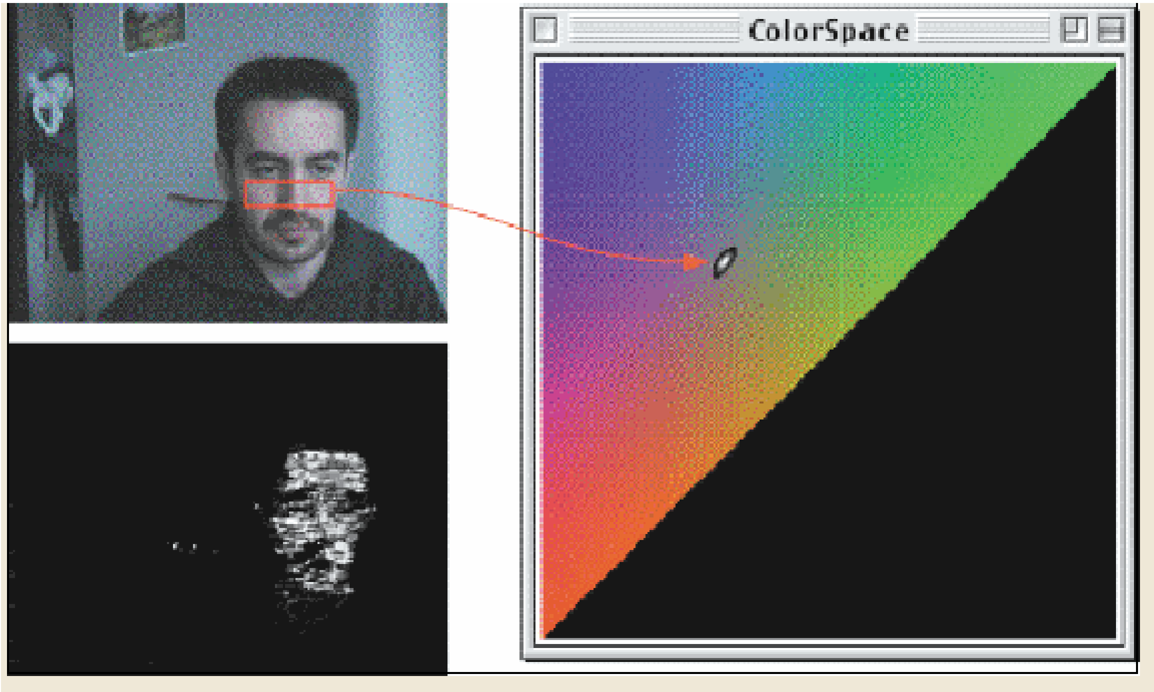
2D Head / hands localization

- contour analysis: mark extremal points (highest curvature or distance from center of body) as hand features
- use skin color model when region of hand or face is found (color model is independent of flesh tone intensity)



Flesh color tracking

- Often the simplest, fastest face detector!
- Initialize region of hue space



[Crowley, Coutaz, Berard, INRIA]

Color Processing

- Train two-class classifier with examples of skin and not skin
- Typical approaches: Gaussian, Neural Net, Nearest Neighbor
- Use features invariant to intensity

  Log color-opponent [Fleck et al.]

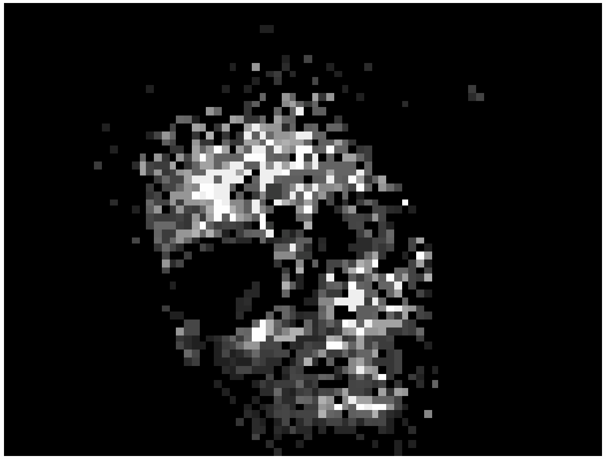
  $(\log(r) - \log(g), \log(b) - \log((r+g)/2))$

  Hue & Saturation

Flesh color tracking

Can use Intel OpenCV lib's CAMSHIFT algorithm for robust real-time tracking.

(open source impl. avail.!)

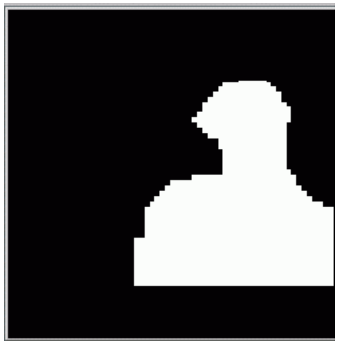


[Bradsky, Intel]

Intel’s computer vision library

Detection with multiple visual modes

Shape



Find head sized peaks in 2-D or 3-D.

Flesh Color Detection

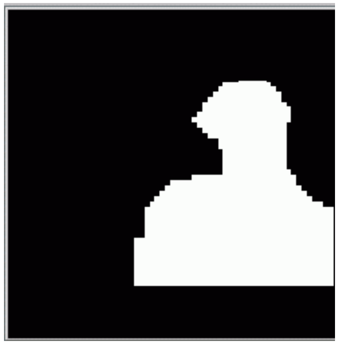
Detect skin pigment in hue-based color space

Face Pattern Detection

Classify intensity vector corresponding to face class

Common Detection Failure Modes

Shape



Fooled by head shaped peaks

Flesh Color Detection

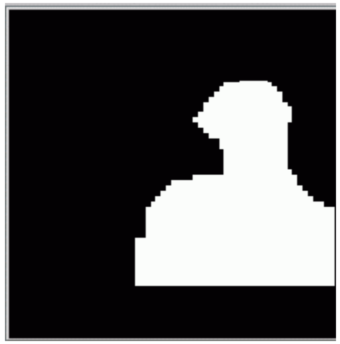
Fooled by flesh colored objects

Face Pattern Detection

Misses out of plane rotation or expression

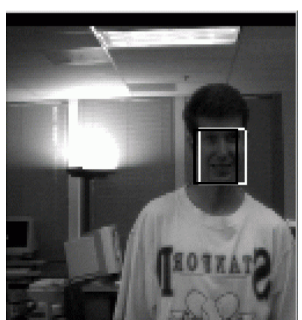
Robust real-time performance

Shape



Flesh Color Detection

Face Pattern Detection



Integrated Face Detection Algorithm
(temporally asynch. voting scheme)

Mode and Task Matrix

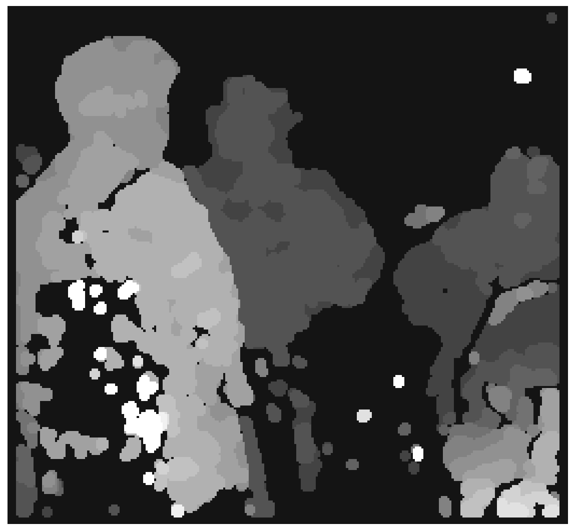
	shape	*color*	*pattern*
detection	silhouette classifier	skin classifier	face detection
identification	biometrics	flesh hue	face recognition
tracking	Shape change	clothing histogram	Appearance change

A Key Technology: Video-Rate Stereo

- Two cameras –> stereo range estimation; disparity proportional to depth
- Depth makes tracking people easy
  - segmentation
  - shape characterization
  - pose tracking
- Real-time implementations becoming commercially available.

Video-rate stereo

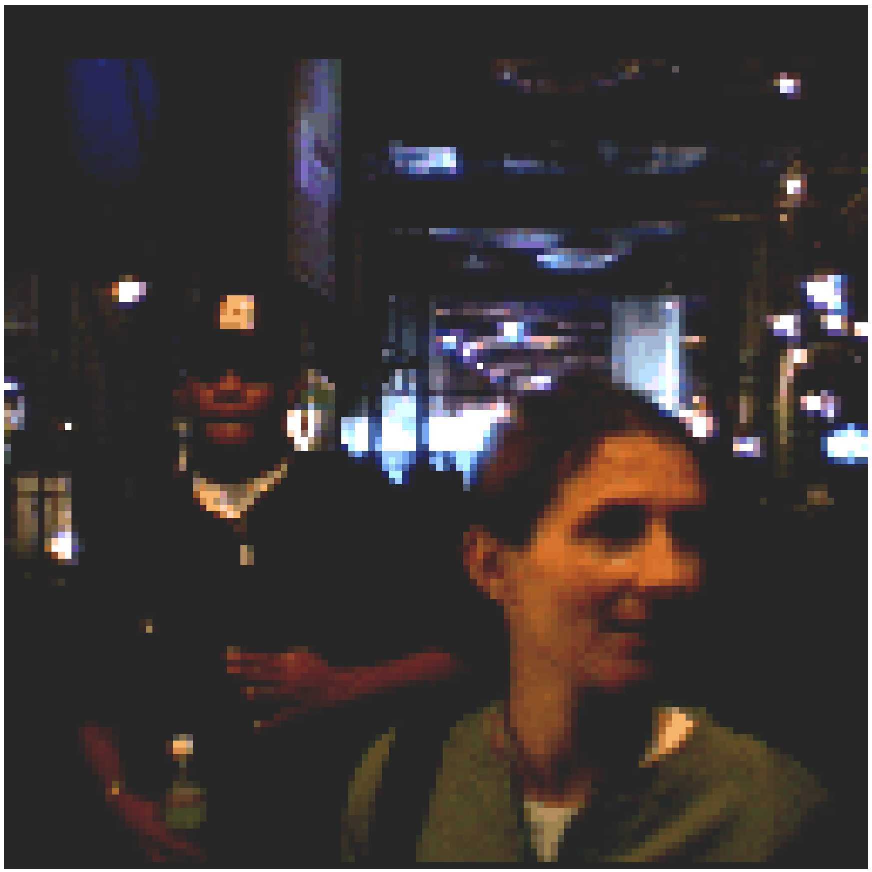
Left and right images



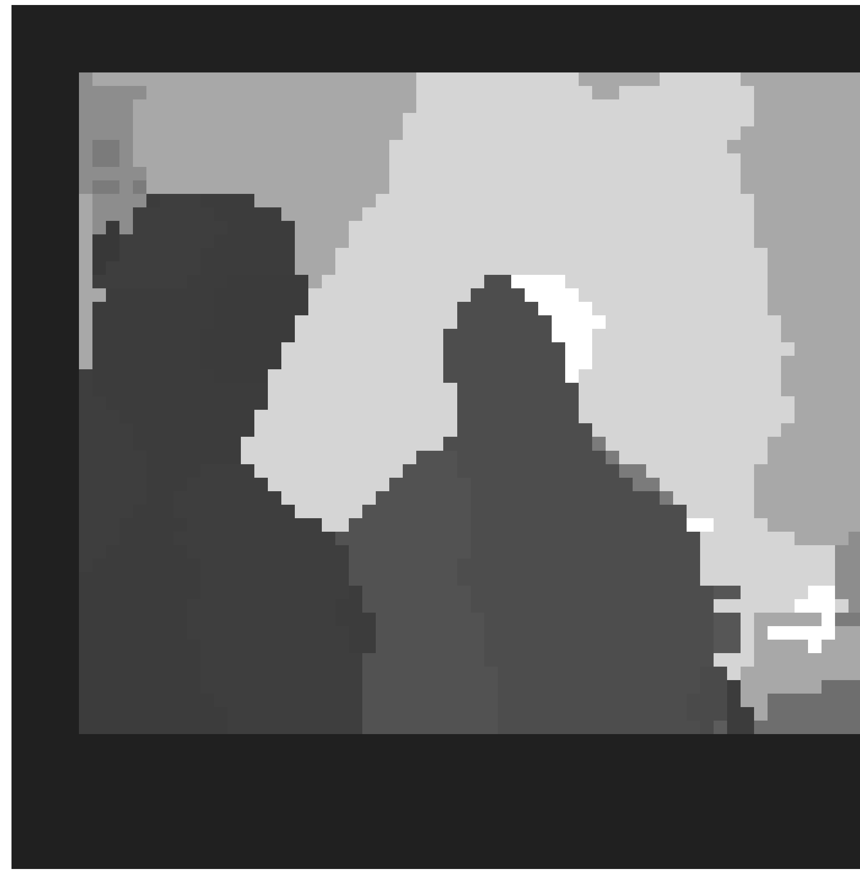
Computed disparity

Foreground pixels; grouped by local connectivity

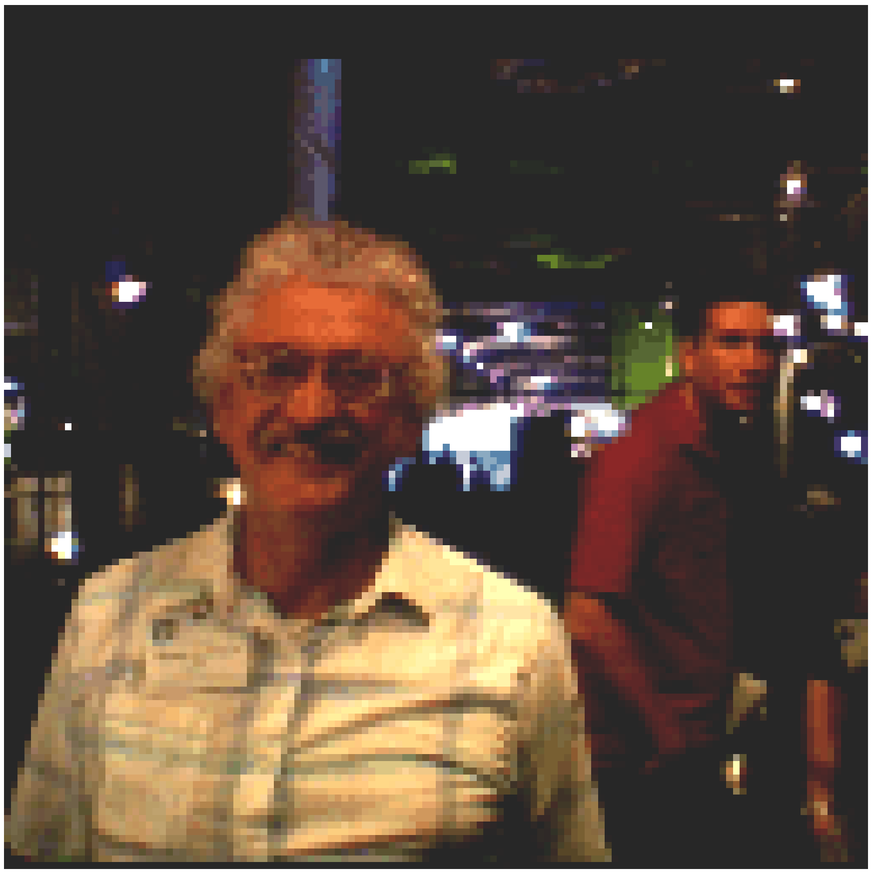
RGBZ input

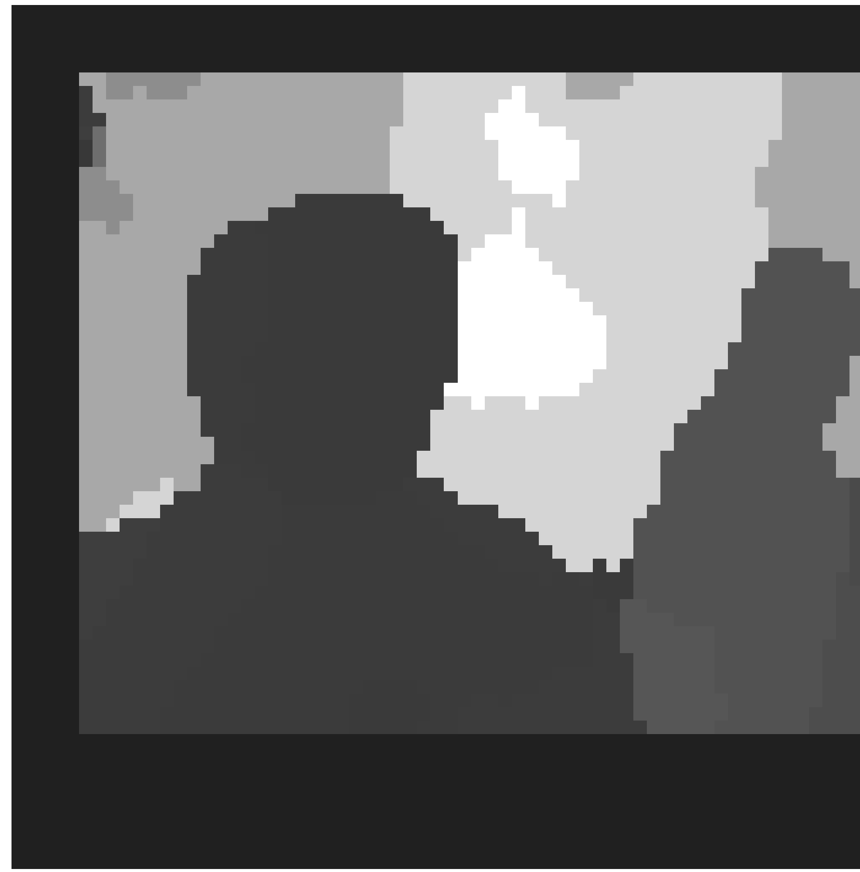


RGBZ input



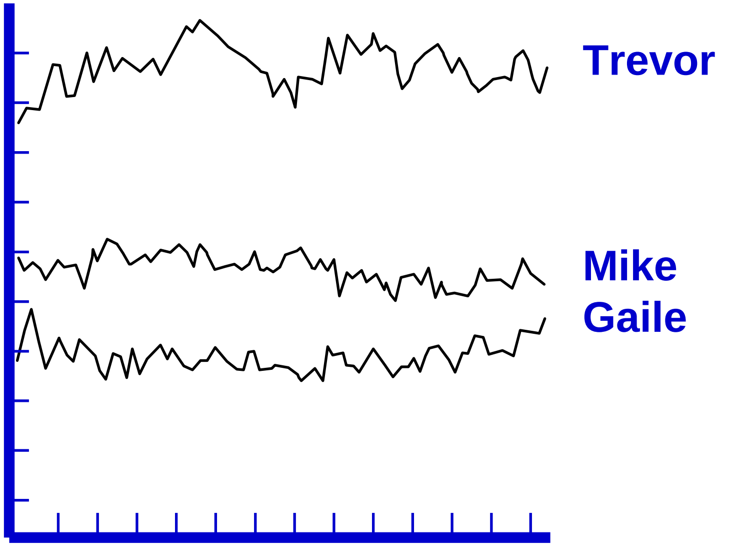
RGBZ input





Range feature for ID!

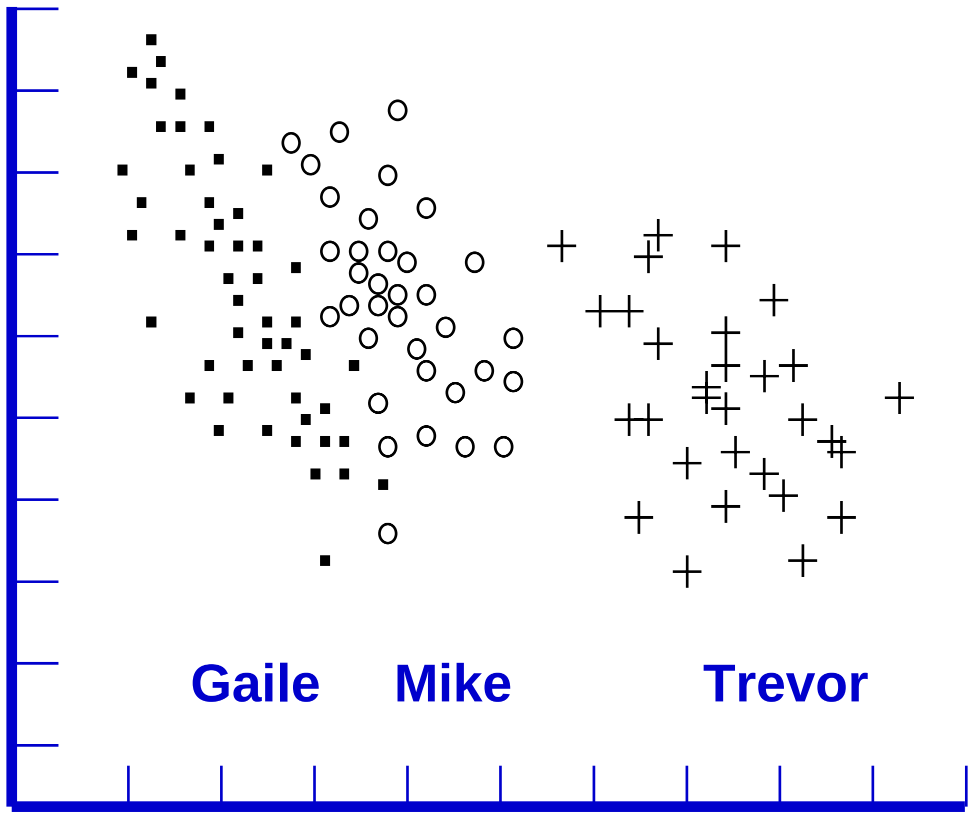
- Body shape characteristics -- e.g., height measure.
- Normalize for motion/pose: median filter over time



- Near future: full vision-based kinematic estimation and tracking-- active research topic in many labs.

Color feature for ID!

For long-term tracking / identification, measure color hue and saturation values of hair and skin….



For same-day ID, use histogram of entire body / clothing

Mode and Task Matrix

	shape	***color***	***pattern***
detection	silhouette classifier	skin classifier	face detection
identification	biometrics	flesh hue	face recognition
tracking	Shape change	clothing histogram	Appearance change

See lectures by Trevor later in the course

Robust, Multi-modal Algorithm

Combine modules for detection:

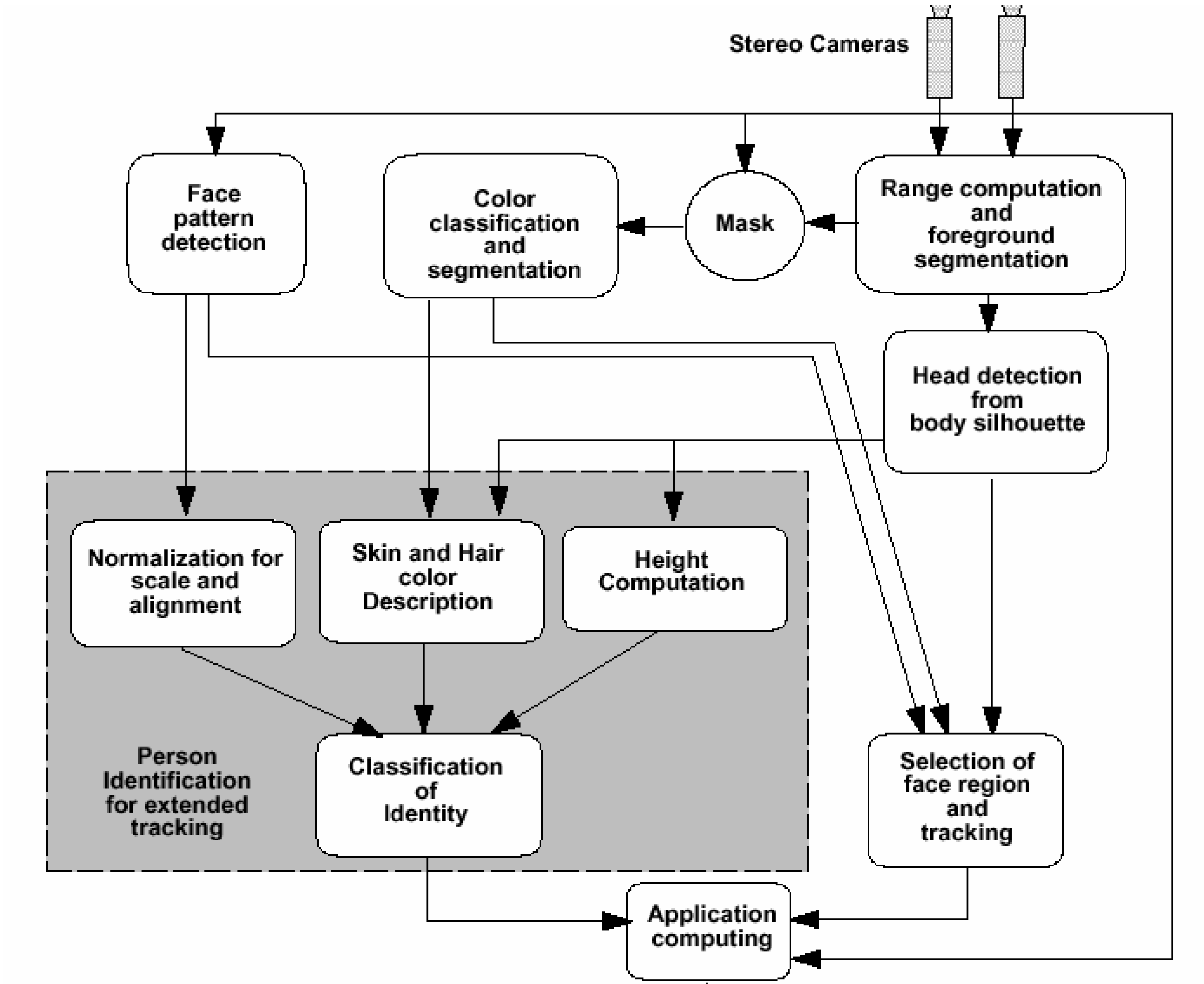
- Silhouette finds body
- Color tracks extremities
- Pattern discriminates head from hands.

Use each also to recognize returning people:

- Face recognition
- Biometrics (skeletal structure)
- Hair and Skin hue
- Clothing (intra-day.)

[CVPR ‘98; T. Darrell, G. Gordon, M. Harville, J. Woodfill]

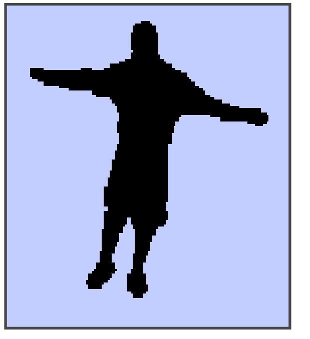
System Overview



Classic Background Subtraction model

- Background is assumed to be mostly static
- Each pixel is modeled as by a gaussian distribution in YUV space
- Model mean is usually updated using a recursive low-pass filter

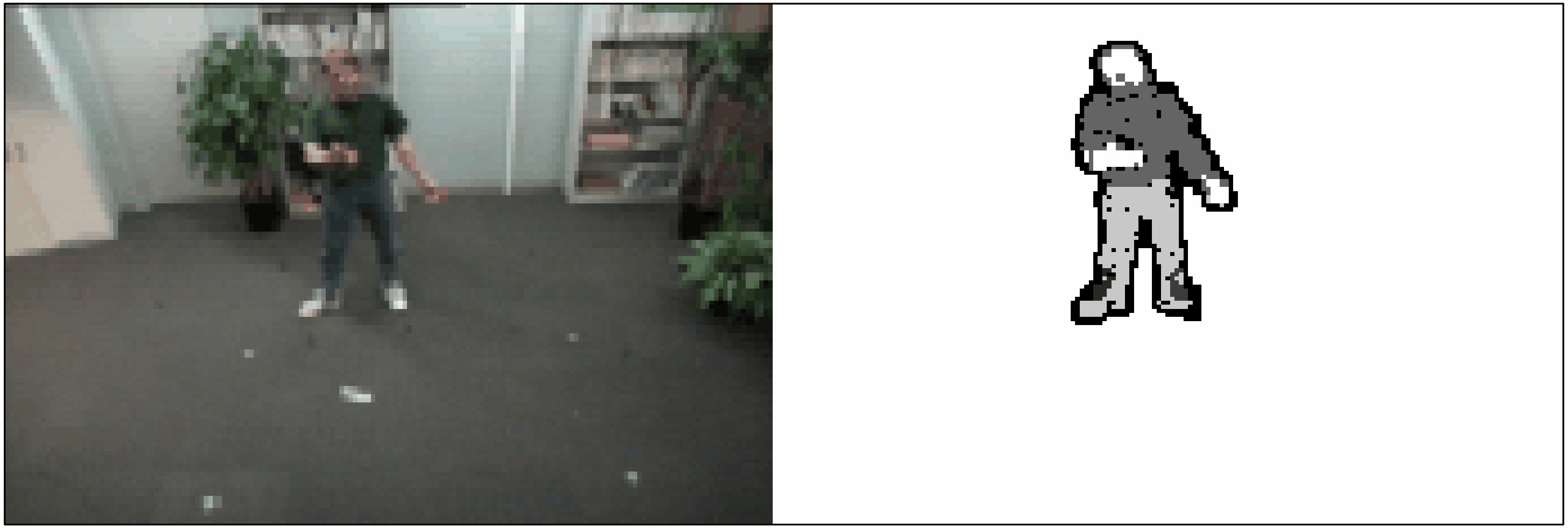
Given new image, generate silhouette
by marking those pixels that are significantly
different from the “background” value.



Static Background Modeling Examples

[MIT Media Lab Pfinder / ALIVE System]

Static Background Modeling Examples

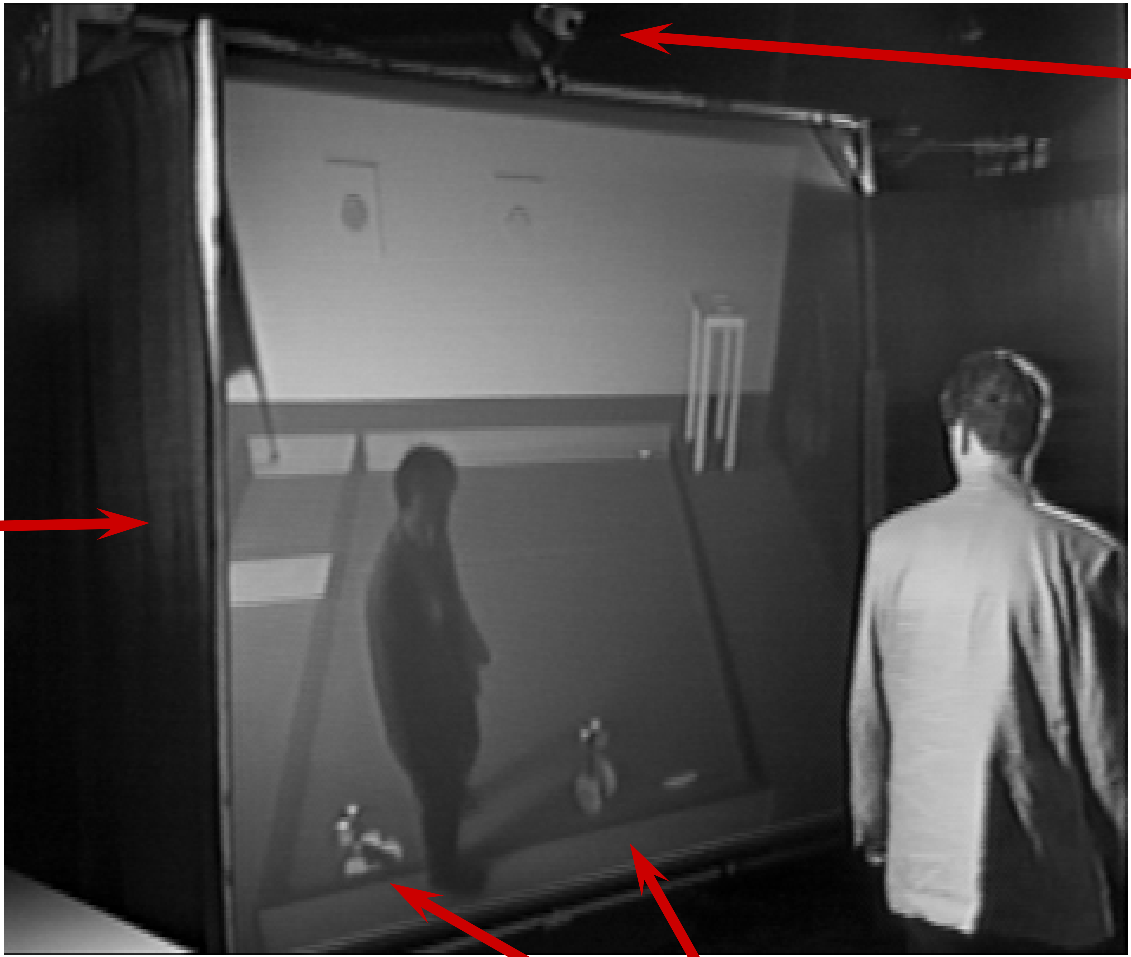


[MIT Media Lab Pfinder / ALIVE System]

Static Background Modeling Examples

[MIT Media Lab Pfinder / ALIVE System]

The ALIVE System



Camera

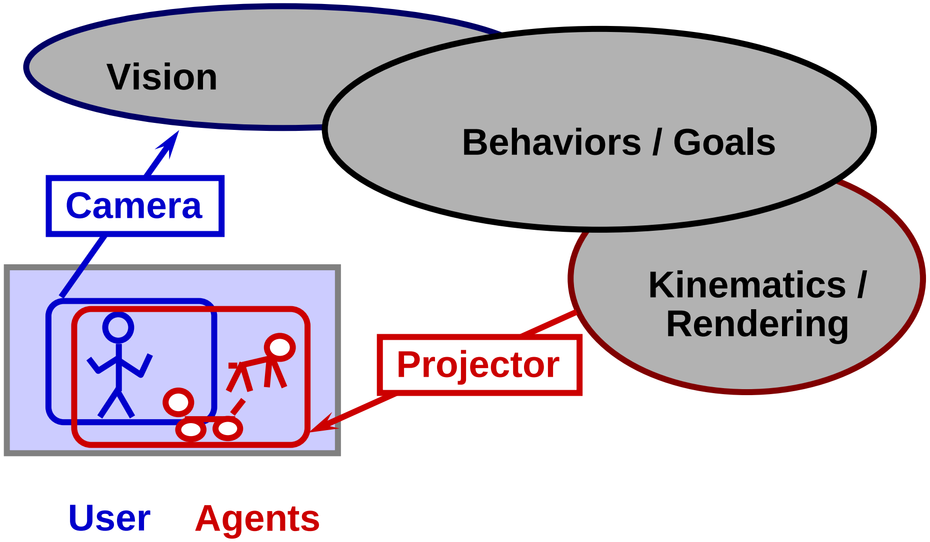
User

Video Screen

Autonomous Agents

ALIVE

- Real sensing for virtual world
- Tightly coupled sensing-behavior-action
- Vision routines: body/head/hand tracking



[Blumberg, Darrell, Maes, Pentland, Wren, … 1995]

ALIVE system, MIT

M.I.T. Media Laboratory Perceptual Computing Technical Report No. 257
(To appear, ACM Multimedia Systems)

The ALIVE System:

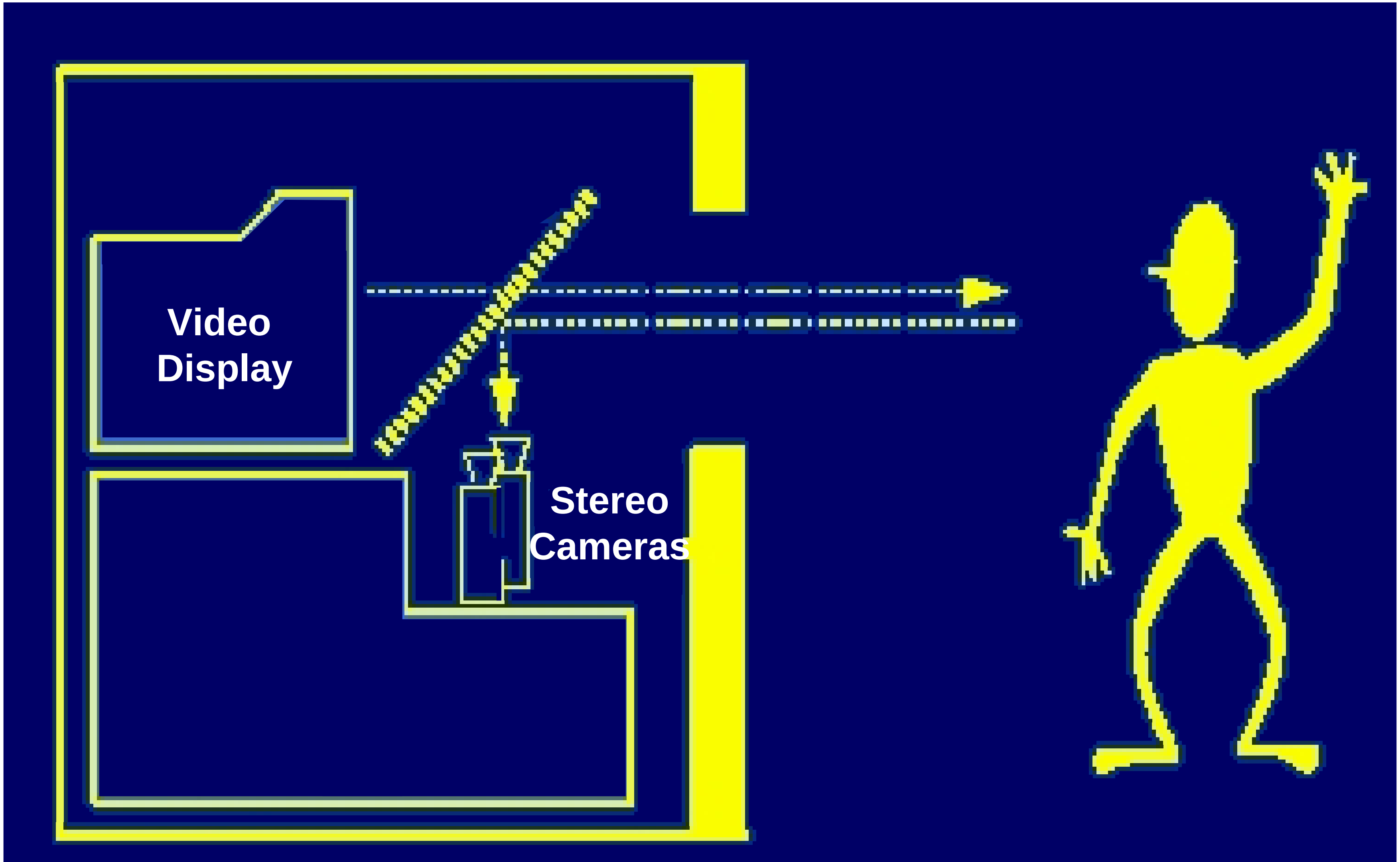
Wireless, Full-body Interaction with Autonomous Agents

Pattie Maes, Trevor Darrell, Bruce Blumberg, Alex Pentland
MIT Media Laboratory

http://vismod.www.media.mit.edu/cgi-bin/tr_pagemaker (TR 257)

http://vismod.www.media.mit.edu/cgi-bin/tr_pagemaker (TR 257)

A Face Responsive Display



Vision-only Application:
Interactive Video Effects

end