Human-focused Computer Vision

Massimo Piccardi
Faculty of Information Technology,
University of Technology, Sydney (UTS)
e-mail: massimo@it.uts.edu.au

cgiv06, Sydney, 26 July 2006
Computer Vision

- How computers see
- Detect, locate, recognise objects; recognise and understand events from still images and videos
- Geometrical computer vision, image processing and analysis, pattern recognition, perceptual psychology
Applications

- Many "traditional" fields of application:
  - Industrial automation
    - quality inspection
    - pick & place
  - Guidance of mobile robots
  - Analysis of biomedical images
  - Analysis of satellite images

- 40 years or so down the track, in some way computer vision is still in its infancy
Systems: Yesterday

- The Massively Parallel Processor (MPP - Goddard Space Flight Center, 1983; 16384 processors, 250 Mflops)

source: http://sdcd.gsfc.nasa.gov/
Systems: Today

- Small, much more powerful, ubiquitous systems enable new applications

Intel PXA270 card
520 MHz, 128 MB RAM
6.5 cm x 6.0 cm

Prosilica GE 2040
2040 x 2040, Gigabit Ethernet

Wireless Gigabit links
Human-focused Computer Vision

- Humans as subjects of the computer vision process
- Interest is growing rapidly
- Challenges anyone?
Challenges: 1

- Compared to other target objects such as gears, apples, and cars, humans are deformable
  - Articulated motion: chain of non-deformable parts connected by joints; constraints
- Surfaces are deformable and not planar
- Materials are heterogeneous
Challenge 1: Example

Details from CAVIAR Project video clips
Challenges: 2

- Many features of interest are small

From the MMI Face Database
Challenges: 3

- Humans move across very different illumination (non uniform in space and variable in time)
Challenges: 4

- Recognition of activities and behaviours is complex

J. Owens, A. Hunter, VS’2000
Observing humans

Detecting, recognising, interpreting

- Presence and location
- Body parts (face, hands, limbs,...)
- Gestures
- Expressions and emotional states
- Actions and interaction
- Activities, Behaviours
- Biometrics
Human-focused applications

- Visual surveillance
- Ambient intelligence
- Video data mining
- HCI: human to computer
- Affective computing
- Identification through biometrics
Visual Surveillance

- Tracking
- Recognising events of interest
  - Abandoned objects
  - Unauthorised access to places
  - Overcrowded areas
  - Anomalous situations
  - Identifying offenders
Tracking people

Larry Davis et al., VSAM Project
Tracking

- Originates in a radar environment
- Typically based on a coherent model of motion, shape, appearance
- Suffers from occlusions
- Disjoint views?
- Highly crowded scenes?
Classifying activities

- Start, End (temporal segmentation)
- Rule-based classification
- Unsupervised clustering
- Anomaly detection
Interaction

Oliver, Rosario, Pentland, *T-PAMI* 22(8) 2000
Related applications

- Ambient Intelligence – user-centric intelligent environments
- Video Data Mining - extraction of implicit knowledge and patterns not explicitly stored in the video data
- Techniques are similar, but requirements can be very different
HCI – human to computer

- Tracking hands, head, lips, gaze for facilitated input
- Automatic lip reading
- Interpreting sign languages

http://gri.gallaudet.edu/~cvogler/research/
Affective computing

- Recognise affective and emotional states from human display
- Application to human-computer interaction for an improved computer end
- More challenging than recognising explicit gestures
Affective computing

- Ground truth from psychological research
  - Face: Ekman, Unmasking the face, 1975

- Multi-modality helps:
  - face
  - upper body
  - audio
  - other wearable sensors
AC from visual display

Face

Upper body

Gunes, Piccardi, ICPR 2006
AC from visual display

- Each modality displas Action Units (AUs)
- AUs develop along stages: neutral, onset, apex, offset, (neutral, ...)
- Face AUs and Body AUs are not stage-synchronous
  - Synchronization
- Fusion
Biometrics

- Face
- Fingerprint
- Iris
- Other biological: ears, hands, veins
- Behavioural: gait, keystrokes
- Thermograms
Face Recognition

FR: Modes

Is face recognition a mature technology??

- Verification
- Identification
FR: Intrinsic limitations

Distance template-instance

More, dense templates

Aging shifts the templates
FR: Challenges

- In a controlled environment: mutual similarity, aging, cosmetics
- In an unconstrained environment: face location, illumination, pose, expressions, glasses, scarves, ...
- FRVT 2006
  - High resolution still imagery
  - 3D facial scans
  - Multi-sample still facial imagery
  - Pre-processing algorithms that compensate for pose and illumination
Fingerprint

- Highly unique and immutable
- Current sensors attempt to detect whether a finger is live *

Iris

- Highly accurate matching

- Responses of the iris to changes in light prove it is live*

* Delac, Grgic, ELMAR-2004
Ears

- Uniqueness?
- Accuracy?
A viewpoint of difference

PillCam, Given Imaging