Space Speaks – Towards Socially and Personality Aware Visual Surveillance

G. Zen, B. Lepri, E. Ricci, O. Lanz

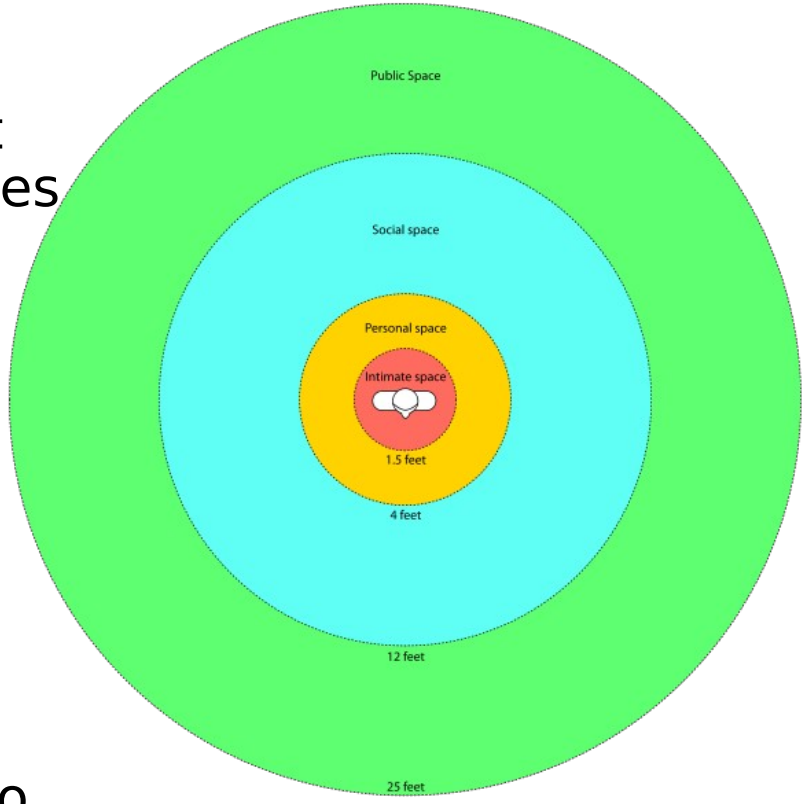
FBK-irst, Trento (Italy)

Workshop on *Multimodal Pervasive Video Analysis* (MPVA)

In conjunction with **ACM-Multimedia**, Florence. October 29th, **2010**

“Space Speaks” - - Proxemics

- Proxemics = the study of personal space [Hall, 1966]
  - Personal space: the space that people ‘form’ around their bodies
- Human perceptions of space are modulated by
  - culture
  - gender
  - social status
  - other individual characteristics e.g. **personality traits** (DeJulio and Duffy, 1977; Williams, 1971)

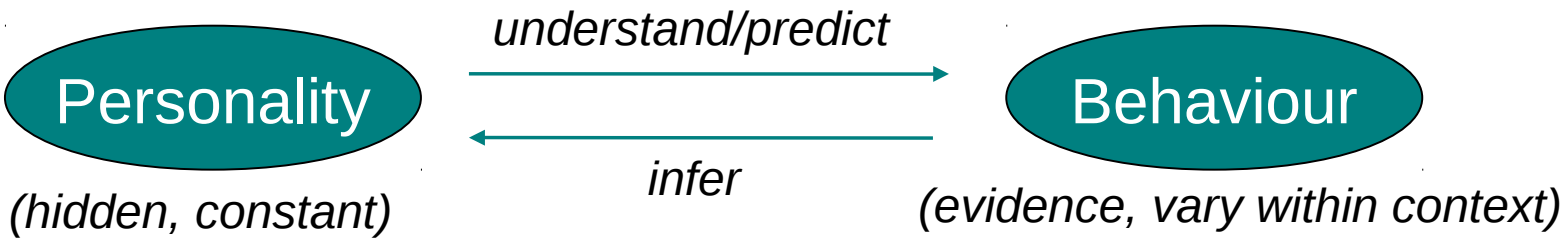


Proxemics – Types of Space

- Public distance
  - the area of space between an audience and a well-known speaker (interactions usually not personal, anonymous)
- Social or formal distance
  - the spaces in which people feel comfortable during business interactions and social interactions with acquaintances and strangers
- Personal or informal distance
  - used among friends and family members
- Intimate distance
  - acceptable only among closest friends and intimates

What is personality?

- A set of stable properties of people



- Multifactorial models / traits:
  - **the Big Five**
    - Extraversion vs. Introversion
    - Emotional stability vs. Neuroticism
    - Agreeableness vs. Disagreeableness
    - Conscientiousness vs. Un-conscientiousness
    - Openness to experience vs. non-open-minded

Why personality?

Personality-awareness is gaining importance in many fields.

- **Human-Computer Interaction.** Towards user centered design [Goren-Bar et al. 2007], [Conati et al. 2003].
- **Robotics.** Robots showing human behaviour enhance people's confort and trustness [Brooks and Arkin 2007].
- **Computer Vision.** Stable properties of human behaviour (not specifically related to personality) encoded as an 'a priori'-knowledge to enhance the people tracking systems accuracy.

Goals

- Exploit **only** significant **proxemics cues** for the **automatic recognition of two personality traits**, Extraversion and Neuroticism

- Build subject-specific behavioral models and use them as “a priori” knowledge to **improve** the estimates of a particle filter based **tracking system**

- Use the people spatial behaviors, recognized online, to control **active cameras** in order to collect near-field data (e.g., data for head pose and facial expressions recognition)

Goals

- Exploit **only** significant **proxemics cues** for the **automatic recognition of two personality traits**, Extraversion and Neuroticism

- Build subject-specific behavioral models and use them as “a priori” knowledge to **improve** the estimates of a particle filter based **tracking system**

- Use the people spatial behaviors, recognized online, to control **active cameras** in order to collect near-field data (e.g., data for head pose and facial expressions recognition)

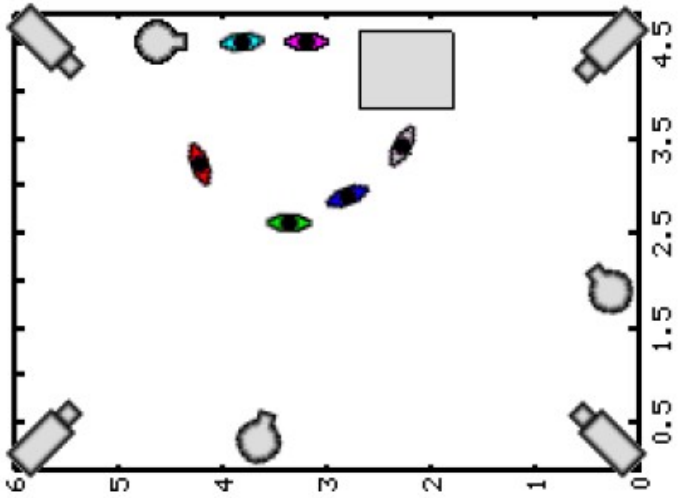
Automatic Personality Recognition

Assumptions

- **personality** shows up in a clear form in the course of social interaction
- **thin slices** of social behavior are enough for a reliable classification of personality traits
- **social signals** (e.g., proxemics cues) can be used for thin slice-driven classification

Experimental Setup

- Scenario: **cocktail party** where a group of people (6-7 subjects) interact spontaneously
- **People moving freely** rather than in a seating arrangement as in previous works.
- Data: 2 sessions of 30 min each, for a total of 13 targets
- **Personality test**: Big Five Marker Scale
- Video sensors: 4 fixed and 3 PTZ cameras.

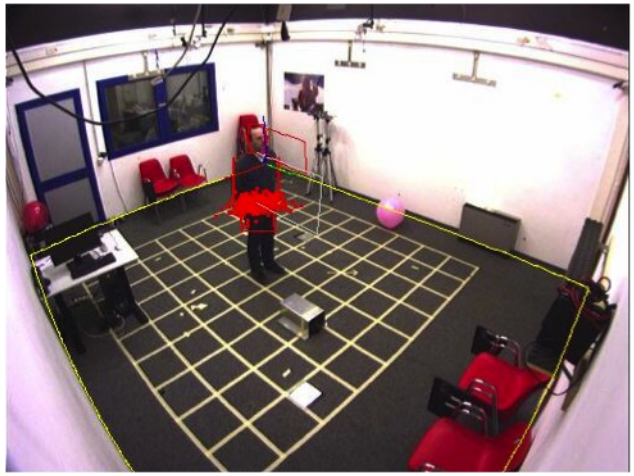


Joint Tracking and Pose Estimation

Multiple people tracking system

Joint (**position+pose**) particle filter tracking

- 3D appearance model (shape+color)
- Effective occlusion handling (HJS algorithm) [Lanz, 2006] [Lanz&Brunelli,2006]
- Spatial exclusion principle implemented through a MRF



Feature Extraction

- **Minimum Distance**: distance among a given target and the closest one
  - Introverts tend to stay at a more intimate distance to few people
  - Extroverts tend to stay at a less intimate distance but to interact with more people
- **Velocity**: variation of the target position every 2 sec.
  - Introverts tend to isolate from the group or get closer to person of confidence
  - Extroverts tend to move freely and to interact with different subjects
- **N° of intimate relationships (INT)**
  - distance < 0.46 meters
- **N° of personal relationships (PER)**
  - Close phase distance < 0.76 meters
  - Far phase distance < 1.2 meters
- **N° of social relationships (SOC)**
  - Close phase distance < 1.2 meters
  - Far phase distance < 2.10 meters
- **N° of public relationships (PUB)**
  - Close phase distance > 2.10 meters

The features vector is computed for each target, for each frame

SVM classification

- Goal: indentify the **proxemics cues** significant for the automatic recognition of Extraversion and Neuroticism
- 2 binary classification tasks (Extroversion vs Introversion / Neuroticism vs Emotional Stability)
- Support Vector Classification
  - 2-order polynomial kernel
- Testing: leave one subject out
- Performance evaluation: Accuracy and F-score

SVM Results

	Extroversion		Neuroticism	
	Accuracy	F-measure	Accuracy	F-measure
All features	**0.66**	**0.66**	**0.75**	**0.69**
Distance	0.54	0.38	**0.69**	**0.52**
Relationships	**0.63**	**0.62**	**0.69**	**0.52**
Intimate	0.53	0.37	0.67	0.50
Personal-Close	0.56	0.46	0.56	0.51
Personal-Far	0.56	0.48	0.64	0.39
Social	0.59	0.59	0.64	0.39
Velocity	0.52	0.46	0.64	0.39

Goals

- Exploit **only** significant **proxemics cues** for the **automatic recognition of two personality traits**, Extraversion and Neuroticism

- Build subject-specific behavioral models and use them as “a priori” knowledge to **improve** the estimates of a particle filter based **tracking system**

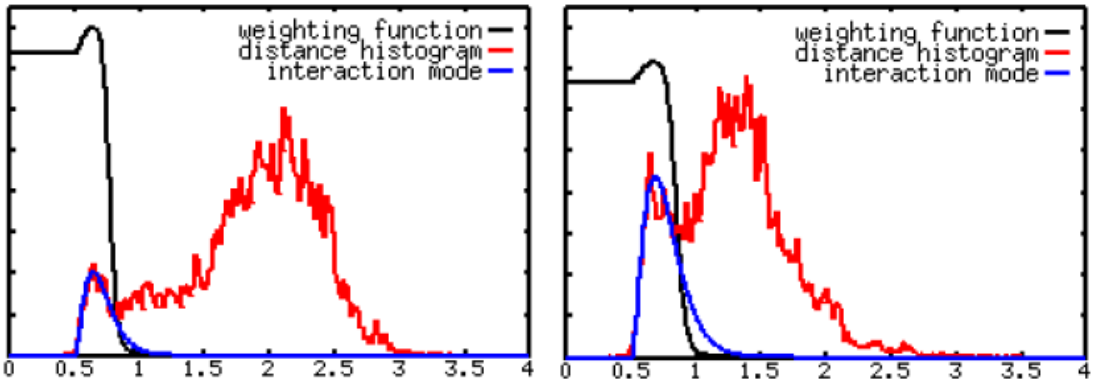
- Use the people spatial behaviors, recognized online, to control **active cameras** in order to collect near-field data (e.g., data for head pose and facial expressions recognition)

Modeling Interaction Patterns

- Embed a **model of interactive behavior** into a Bayesian tracking framework
- Calibrate it to each individual to capture inter-personal variations that **correlate with personality traits**.
- Show that such model, once calibrated to each subject, **improves tracking performance**.

Modeling Interaction Patterns

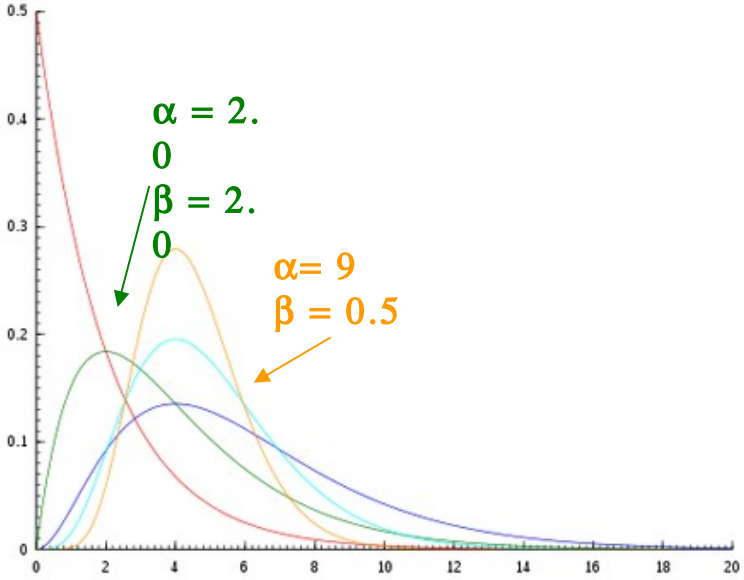
- Compute for each person the **histogram of** its **distance** from the other targets within its visual attention field.
- Histograms with a prominent peak at a distance (<1m) on which people tend to interact.
- Model the **‘’interaction’’ mode** in the histogram as a **Gamma distribution** and fit such model to the data at hand.



Modeling Interaction Patterns

- Different people behaviors correspond to different distributions

$$\Gamma(x;\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta}$$



	Extro-version	Neuro-ticism	mode	spread α	weight β
T1	62	45	0.80	0.15	0.39
T2	56	44	0.66	0.09	0.29
T3	55	57	0.68	0.09	0.25
T4	40	41	0.82	0.08	0.25
T5	37	39	0.64	0.07	0.09
T6	31	43	0.62	0.07	0.25
T7	29	36	0.66	0.09	0.25

Modeling Interaction Patterns

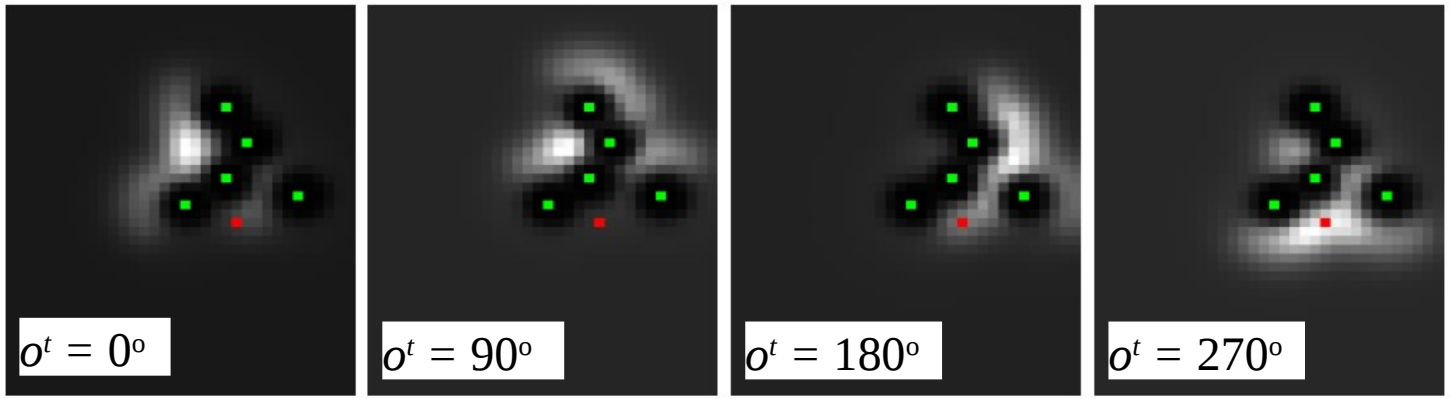
- Before: a MRF potential encodes the spatial exclusion principle.

- Now: the MRF accounts for the interaction prior learnt for target

$$\psi(p_t^k, o_t^k; p_t, o_t) = \Gamma(\|p_t^k - p_t\|; \alpha^k, \beta^k) \cdot N(o_t^k - o_t; \sigma_o)$$

Target position

Head pose



The field perceived by the probe target (the red dot) is maximal when $o^t = 270^o$ where he/she sees a subject at the learnt social distance.

Tracking Results

Real time estimates computed with (top row) and without (bottom row) interaction prior.

Head pose estimation improves significantly!

Tracking Results

Real time estimates computed with (left) and without (right) interaction prior.

With Interaction Prior

Without Interaction Prior

Goals

- Exploit **only** significant **proxemics cues** for the **automatic recognition of two personality traits**, Extraversion and Neuroticism

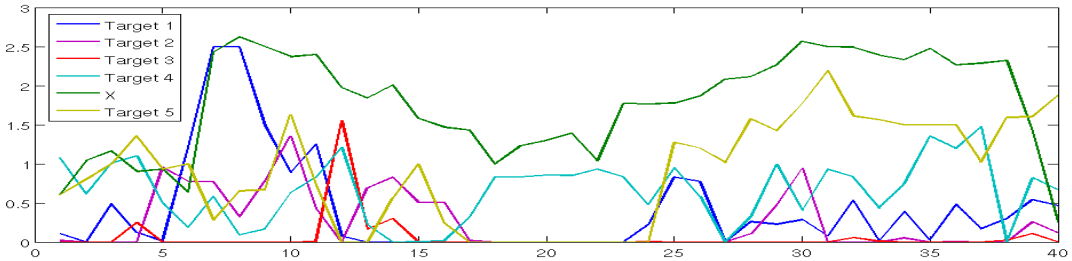
- Build subject-specific behavioral models and use them as “a priori” knowledge to **improve** the estimates of a particle filter based **tracking system**

- Use the people spatial behaviors, recognized online, to control **active cameras** in order to collect near-field data (e.g., data for head pose and facial expressions recognition)

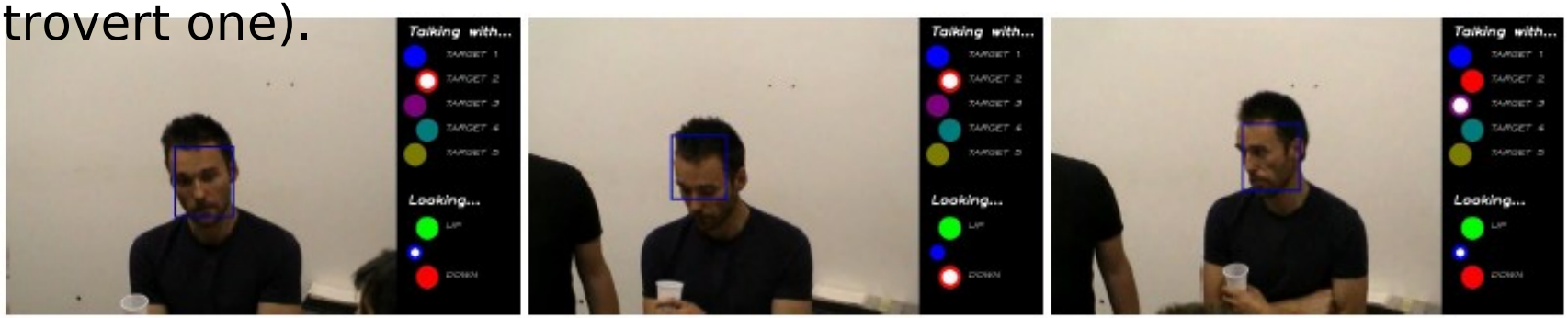
Personality Aware PTZ

A 'social logic' can be used to determine the target of interest to follow with a PTZ cameras.
Near field views can provide other informations than localization (e.g. emotion recognition). Here we refine head pose estimation.

Example: SOC features computed for the first half of the meeting.



PERSONALITY AWARE PTZ: Define a score of people extroversion and follow with PTZ the target with the lowest score (the most introvert one).



Conclusions

Main Contribution:

- Estimation of the level of Extroversion and Neuroticism of people with a SVM classifier using visual features
- Integration, in form of a prior, of a model of interactive behavior into a particle filter tracking algorithm.

Future works:

- validate our approach on more data and different scenarios
- further elaborate on personality aware active camera and better exploit adaptivity aspects.

Conclusions

- Relationship between proxemics, visual attention and personality traits during interaction.
- Estimation of the level of Extroversion and Neuroticism of people with a SVM classifier using visual features
- Integration, in form of a prior, of a model of interactive behavior into a particle filter tracking algorithm.

Future works:

- validate our approach on more data and different scenarios
- tighter integration between low-level (tracking and head pose estimation) and high-level tasks (personality trait classification)
- further elaborate on personality aware active camera and better exploit adaptivity aspects.