Human Activity and Vision Summer School

Estimating Aspects of Social Behaviour with Non-verbal Cues

Hayley Hung, University of Amsterdam

What is social behaviour?

When it's good...

When it's bad...

Why should we care about social behaviour?

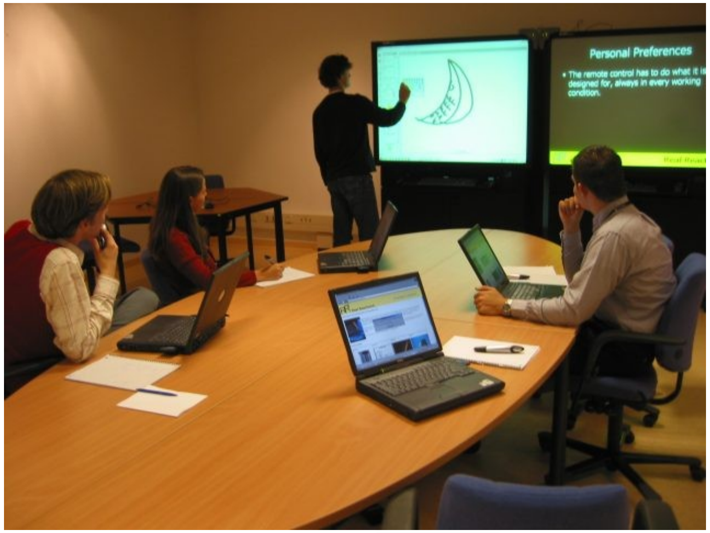
- Human-Robot Interaction

Why should we care about social behaviour?

- Helps for designing man-machine interfaces that are effortless to interact with.
- e.g. Human-Avatar Interaction

Why should we care about social behaviour?

- Face-to-face contact helps to **establish trust**, **friendship**, paves the way for **future relationships** and potentially **influence**.
- e.g. Automated meeting analysis
  - Dominance detection
  - Interest detection
  - (Dis)agreement detection
  - Role Recognition



Why should we care about social behaviour?

- Public Space monitoring

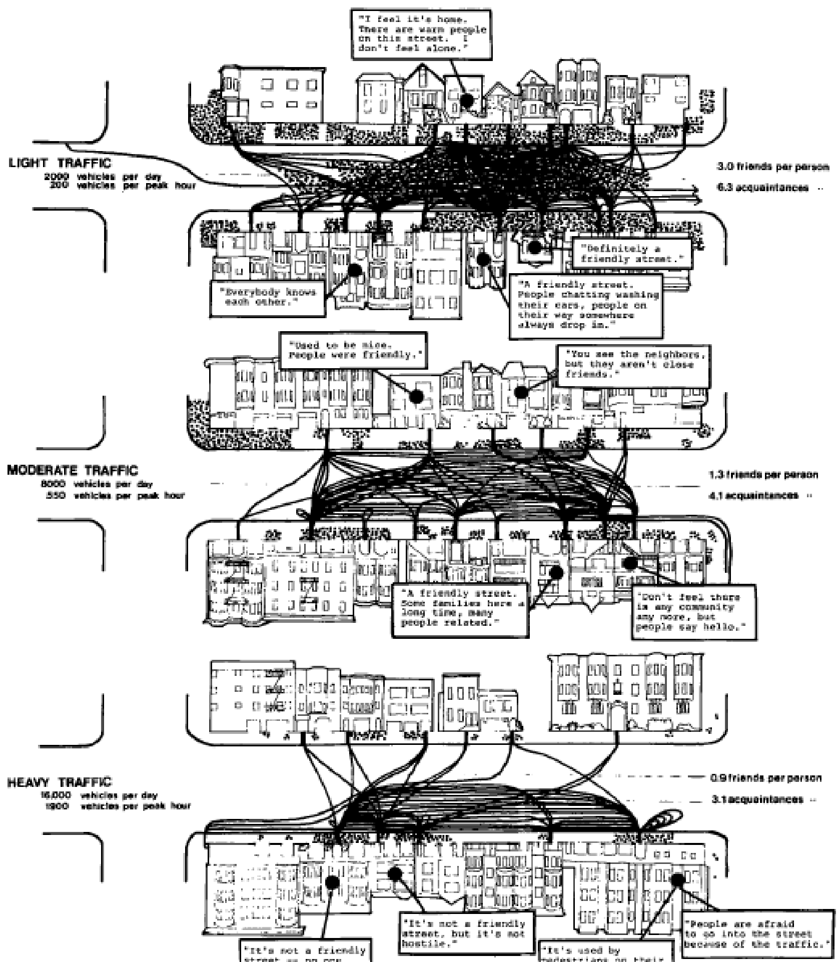
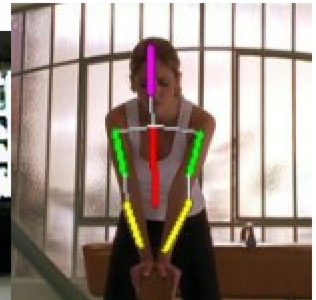


FIGURE 6 *Social Interaction*
Lines show where people said they had friends or acquaintances. Dots show where people are said to gather.

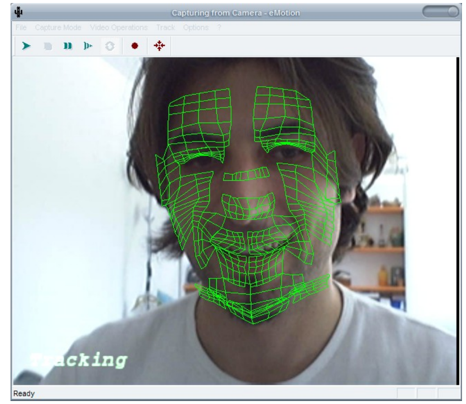
The social semantic gap

We can extract behaviour:

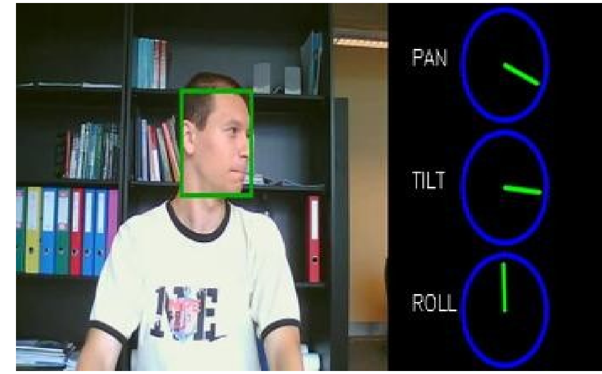
Bazzani et al. 2012



Ferrari et al. 2008

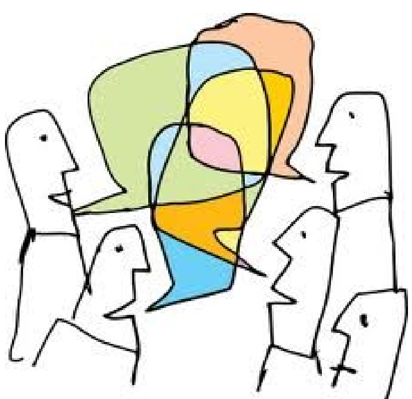


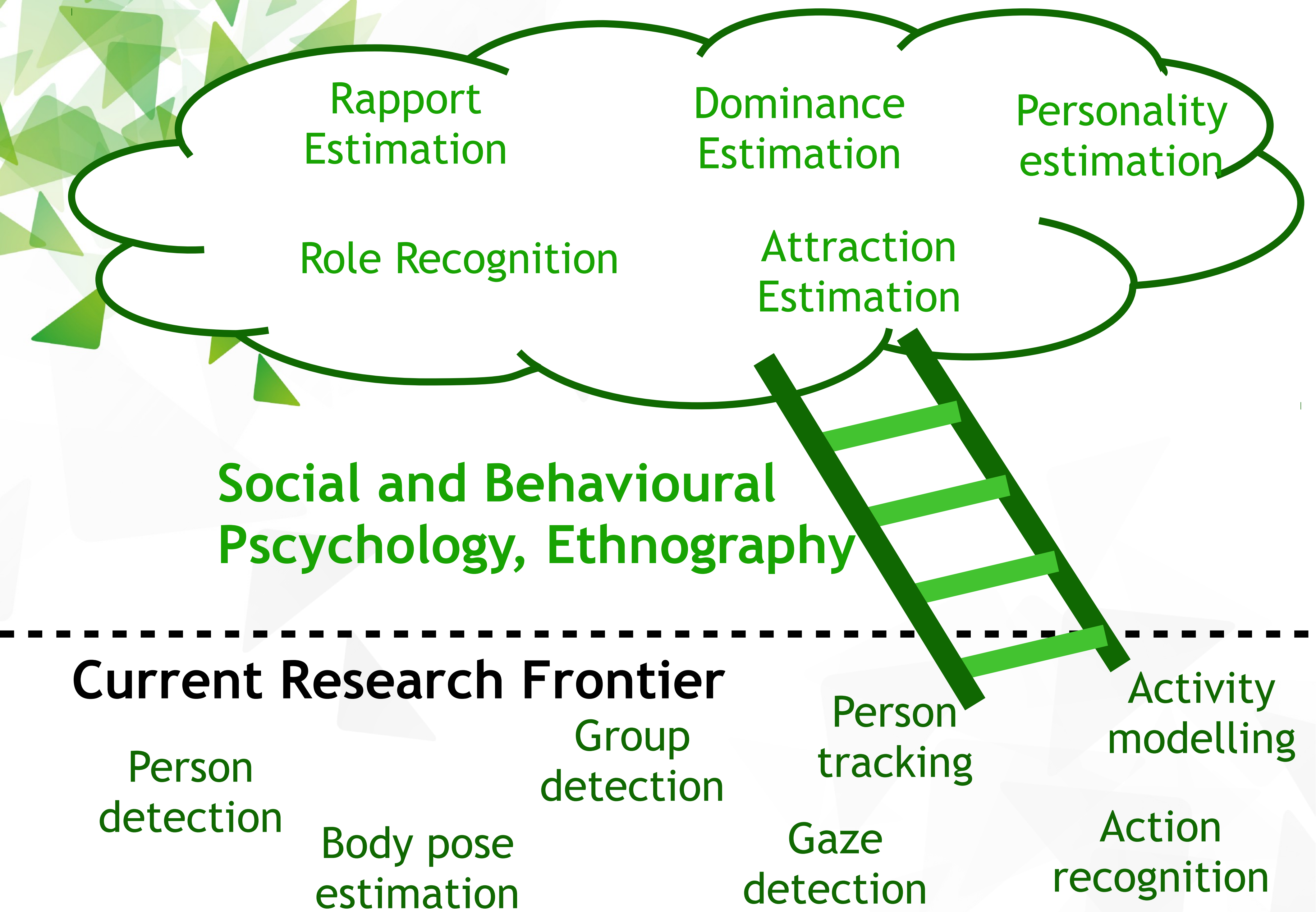
Valenti et al. 2007



Ricci & Odobez 2009

The Social Semantic Gap ?????????????????




Rapport Estimation
Dominance Estimation
Personality estimation
Role Recognition
Attraction Estimation
Social and Behavioural Pscychology, Ethnography
Current Research Frontier
Person detection
Body pose estimation
Group detection
Person tracking
Gaze detection
Activity modelling
Action recognition

Social Signal Processing

- Social behaviour analysis, interpretation, and synthesis.
- **"A social signal** is a communicative or informative signal that, either directly or indirectly, provides information concerning social interactions, social emotions, social attitudes or social relations."
- "Social signals are manifested through a multiplicity of **non-verbal behavioural cues** including facial expressions, body postures, gestures, vocal outbursts, etc."
- Sspnet.eu

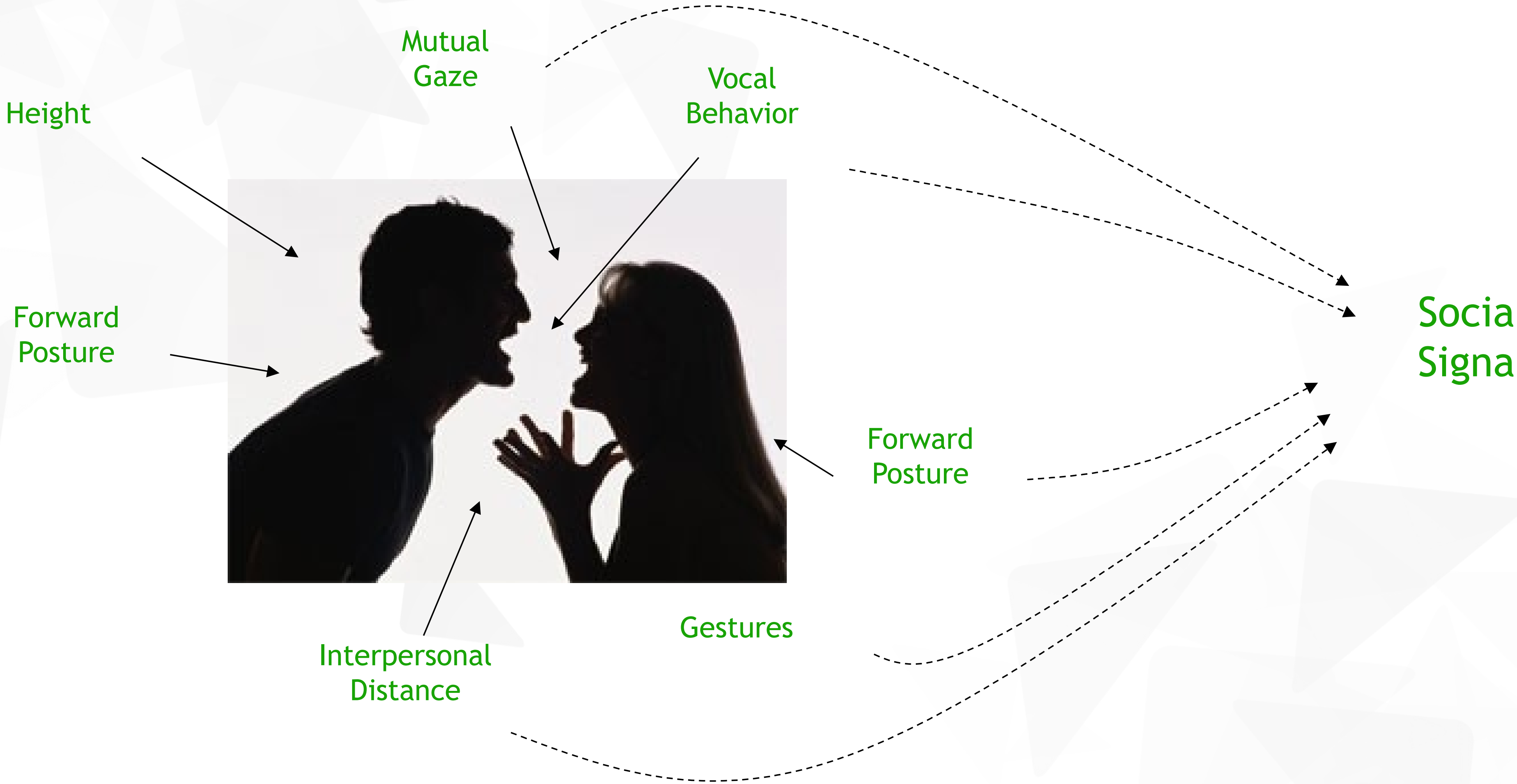
Social Signal Processing Network

A.Vinciarelli et al., 'Bridging the Gap Between Social Animal and Unsocial Machine: A Survey of Social Signal Processing' IEEE Transactions on Affective Computing, 2012

D. Gatica-Perez, 'Automatic Nonverbal Analysis of Social Interaction in Small Groups: a Review' Image and Vision Computing, 2009

S. Pentland,'Honest Signals: How they shape our world', 2008

What are Non-verbal Cues?

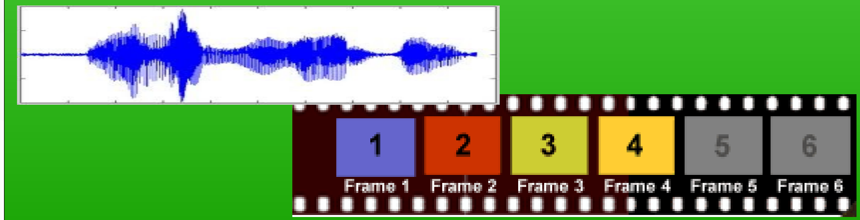


Our focus today

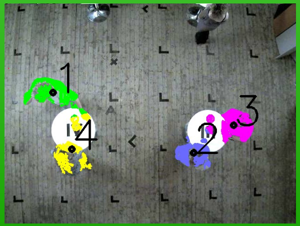
- How can we model and automatically understand social behaviour?
- To what extent can findings from social and behavioural psychology help to **inspire** automated models to understand social behaviour?
- 4 tasks:
  - Dominance Estimation
  - Personality Estimation
  - Attraction Estimation
  - Social Group Estimation

Typical Social Behaviour Estimation Flow Diagram

Data: Audio and/or Video

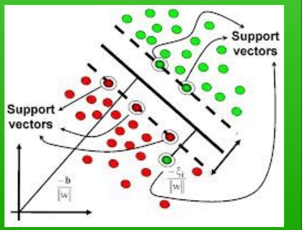


Feature and Cue Extraction



Data Annotation

Social Behaviour Modelling



Model Evaluation

1. Estimating Dominance

Jayagopi et al. "Modeling Dominance in Group Conversations using Non-verbal Activity Cues" IEEE Transactions on Acoustics, Speech and Language Processing, 2009
Hung et al. "Estimating Dominance In Multi-Party Meetings Using Speaker Diarization" , IEEE Transactions on Multimedia, 2011

What's so Interesting About Dominance?

- Fundamental construct in social interaction
- Related to power and status (social verticality)
- Profound effects on brief encounters, relationships & organizations
- It's not always easy to actually tell someone they are dominating...
- ...Having an impartial/neutral judgement from a machine may be more constructive.

Social Psychology: Non-verbal Expressions of Dominance

- Talking louder (Tusing, 2000)
- **Talking longer** (Schmid Mast, 2002)
- Speaking first or respond quickly (Leffler,1982)
- Attempting more **interruptions** (Smith-Lovin,1989)
- **More kinesically expressive** (Dunbar 2005)
- Accompanying their speech with gestures (Dunbar 2005)
- Receiving **more visual attention** (Efran,1968)
- Exhibiting a high **looking-while-speaking to looking-while-listening** ratio (Exline, 1975)

Research Question:

Can dominant people be automatically identified using only Non-verbal Cues?

task-oriented meetings

7 cameras, 24 microphones

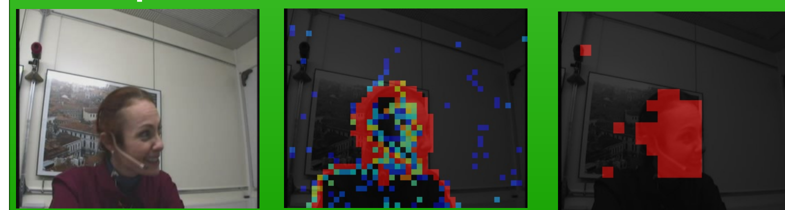
Is it possible to do it from relatively brief observations?

Audio-Visual Dominance Estimation System

Cue Extraction

Audio cues:
Binary Speaking Activity
Prosodic cues

Video cues:
Compressed domain features



Estimate Most Dominant Person

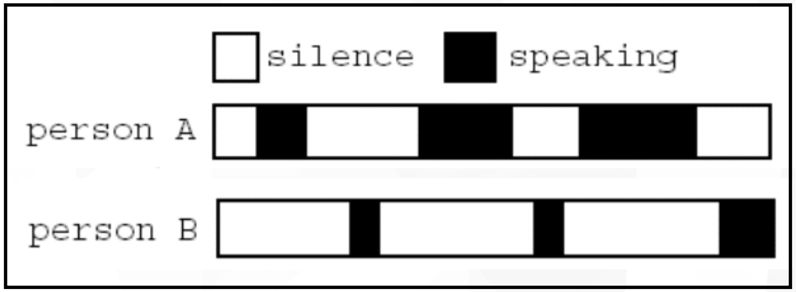
Dominance Annotations

Evaluation:

Jayagopi et al. "Modeling Dominance in Group Conversations using Non-verbal Activity Cues" 2009

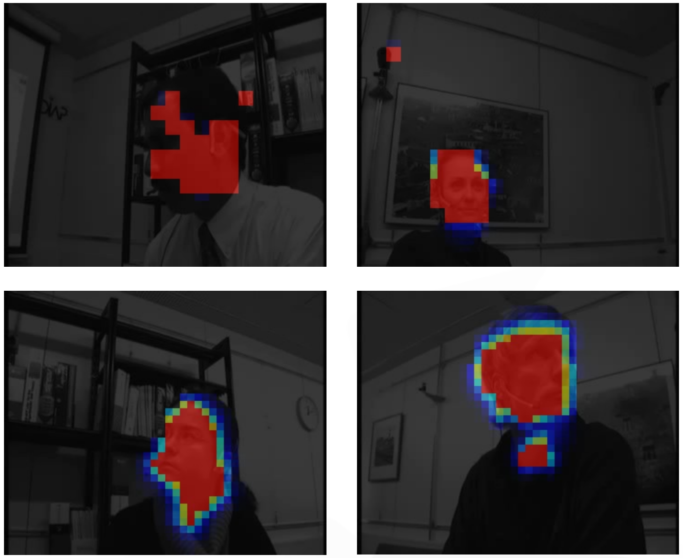
Extracting Audio cues

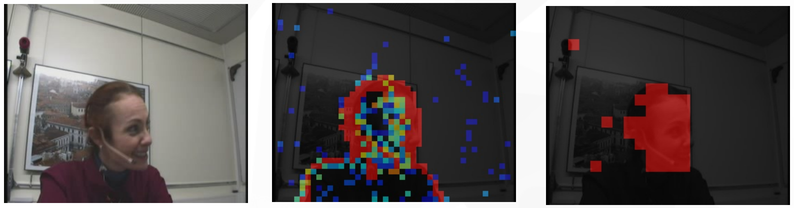
- From head set microphones:
  - Speaker energy
  - speaker-turn segmentation
    - speaking length (TSL)
    - number of turns (TST)
    - number of successful interruptions (TSI)
    - number of times being interrupted (TBI)
    - number of ‘speaking first’ times (TSF)



Extracting Visual Activity Cues

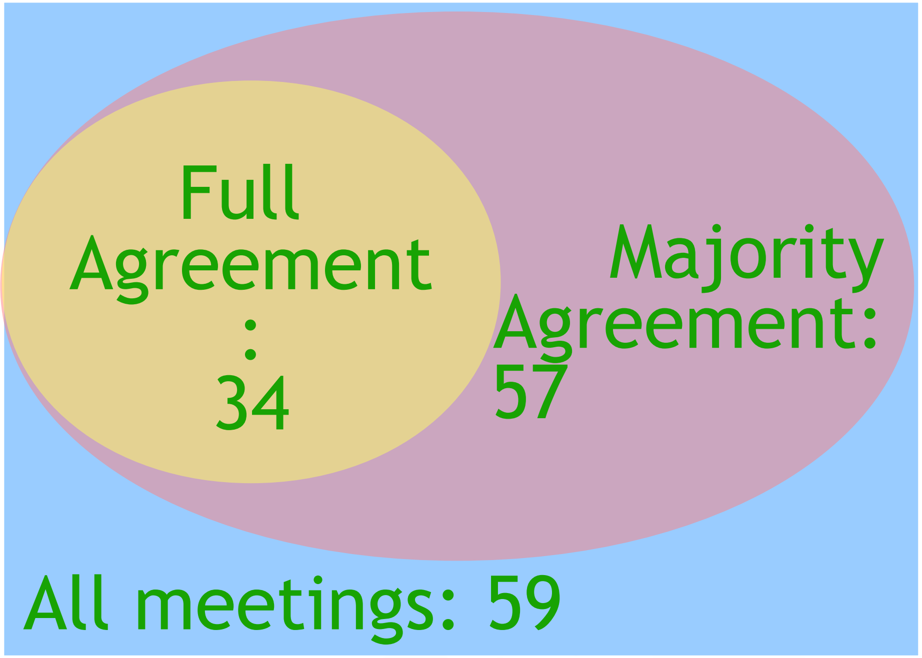
- Inexpensive features computed in compressed-domain
  - DCT coefficients
  - motion vectors
  - residual coding bit-rate
- Used for efficient
  - skin blob detection
  - activity level modeling (high / low)
- Extracted cues (similar to audio cues):
  - visual activity length (TVL)
  - visual activity turns (TVT)





Annotating for Dominance

- **Data**
  - meetings divided into 5-min
  - non-overlapping segments
- **Annotation set up**
  - 21 annotators in total
  - 3 annotators per meeting
- **Annotation procedure**
  - No prior definition of dominance given
  - Absolute rankings: 1 (most dominant) to 4.

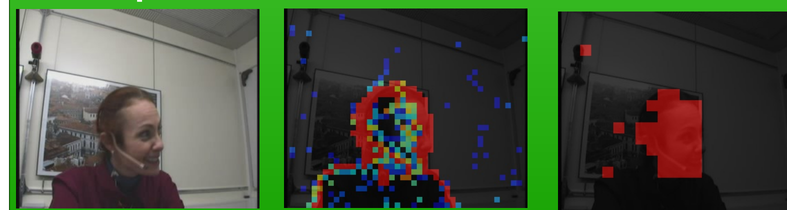


Audio-Visual Dominance Estimation Results

Cue Extraction

Audio cues:
Binary Speaking Activity
Prosodic cues

Video cues:
Compressed domain features



Estimate Most Dominant Person

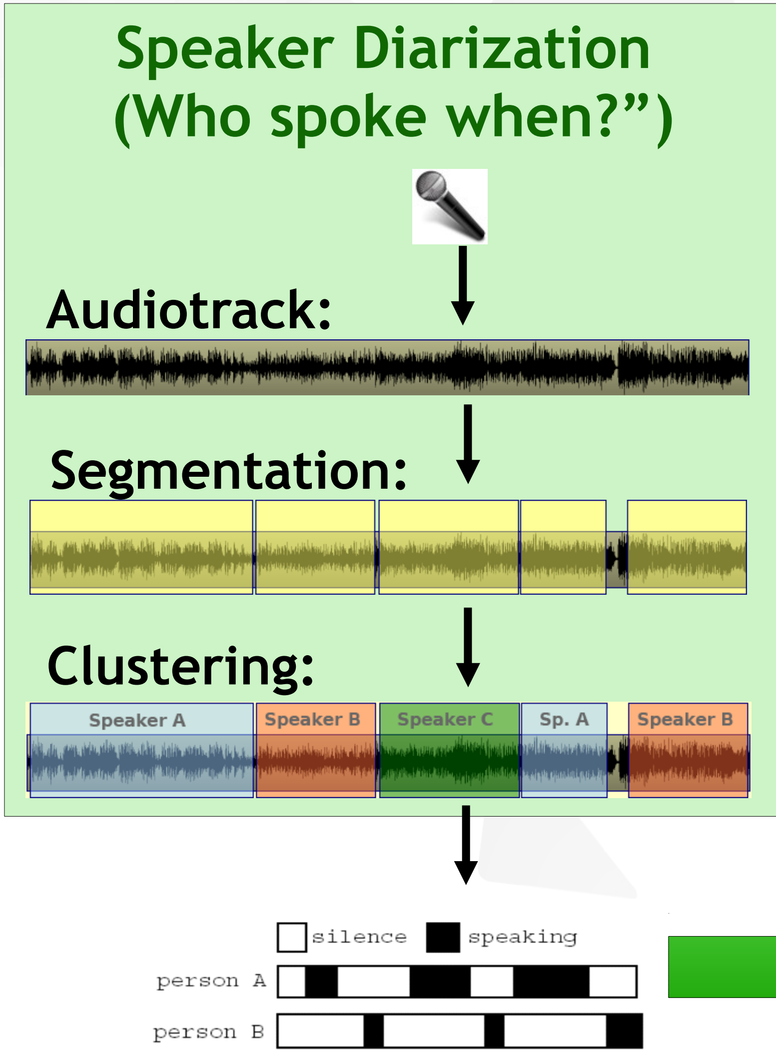
Dominance Annotations:
21 annotators
3 per 5-minute meeting
59 data points
34 full agreement
57 majority agreement

Evaluation:
Best single feature 85.3%
Best feature fusion 91.2%
Class. Acc.

'Dominance as Expressed and Inferred through Speaking Time: A Meta-Analysis', Schmid-Mast 2002

Jayagopi et al. "Modeling Dominance in Group Conversations using Non-verbal Activity Cues" 2009

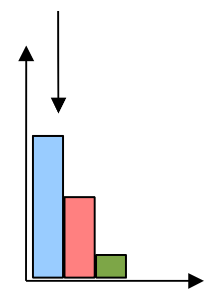
Dominance Estimation using a Single Microphone



Cue Extraction

Dominance Annotations

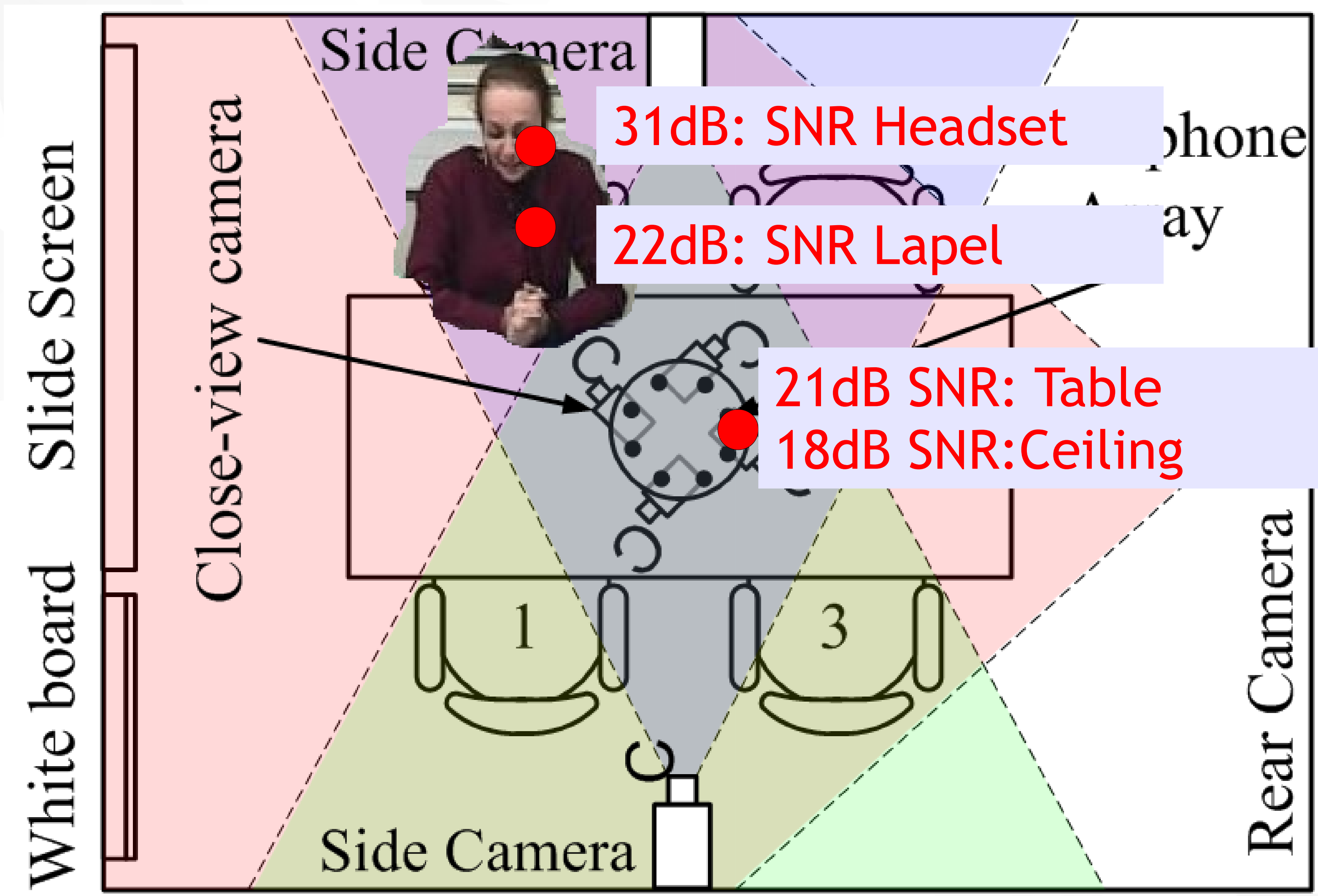
Who is dominant?



Evaluation

Hung et al. "Estimating Dominance In Multi-Party Meetings Using Speaker Diarization" 2011

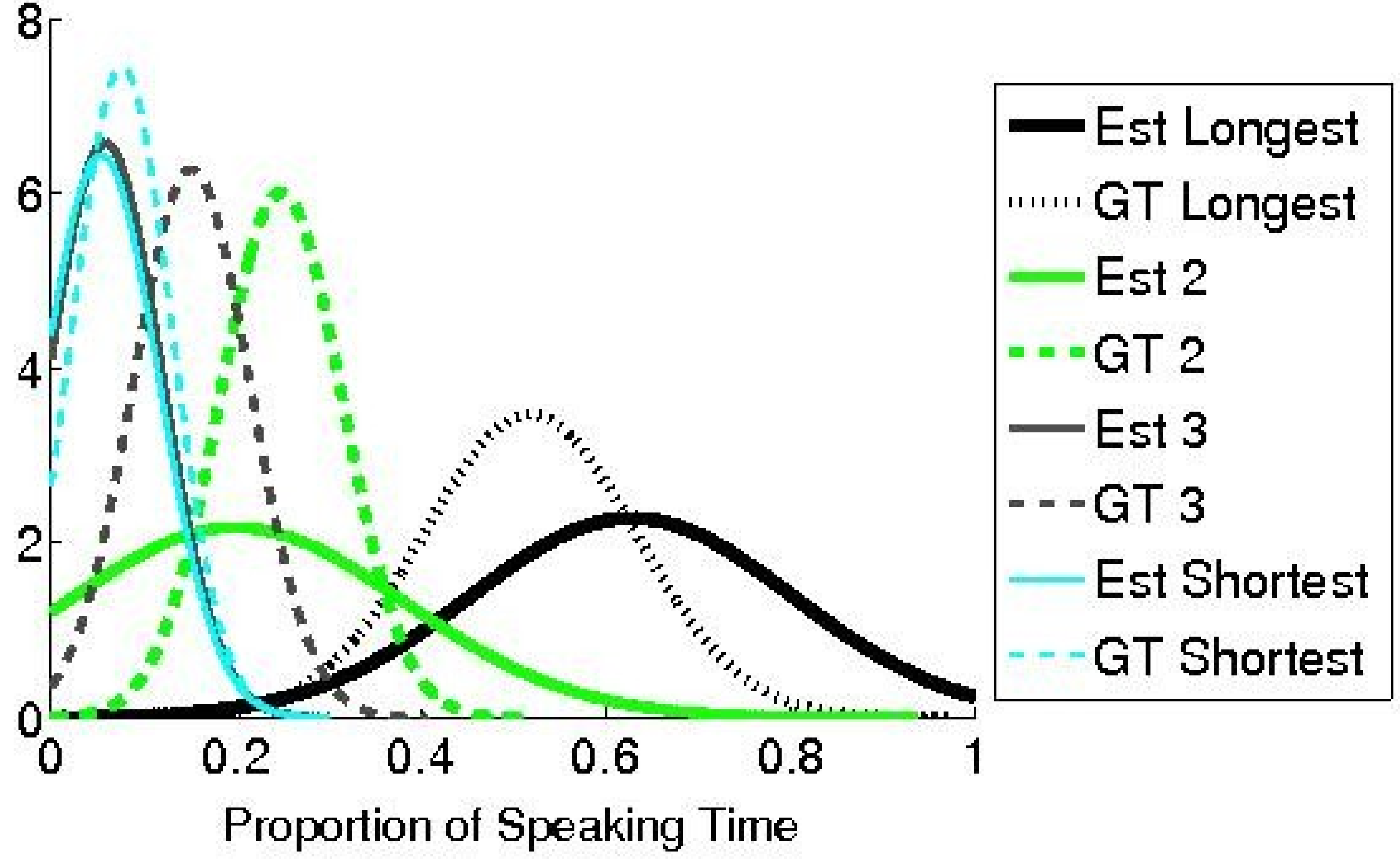
Microphone Experimental Conditions



Audio Dominance Estimation Experiments and Results

- **Experiments:**
  - Different speaker diarization strategies:
    - Thorough and accurate method (Slow), Approximated clustering method (Fast).
  - Different experimental conditions
    - Distance->Signal to Noise Ratio
- **Experimental results:**
  - Speaker Diarization Error:
    - Error reduces as the SNR increases.
  - Highest dominance classification accuracy (74%)
    - One of the worst SNR.
    - fastest clustering approximation.
    - One of the worst diarization errors.
    - Baseline using individual headset microphones was 85%

Diarization Accuracy, Dominance and Speaking Length



Hung et al. "Estimating Dominance In Multi-Party Meetings Using Speaker Diarization" 2011

Dominance Estimation Conclusion

- SNR did not appear to be correlated with dominance estimation performance.
- Making 'shortcuts' with the diarization algorithm did not appear to affect the dominance estimation performance.
- Using a single microphone to estimate dominance is not as good as using multiple microphones.
  - Best result using Speaker Diarization: 74%
  - Using Headset microphones : 85%
- Group self-regulation of behaviour ensures certain interaction rules are maintained.
  - Dominance estimation performance not correlated with speaker diarization performance.

2. Personality Analysis of Vlogs

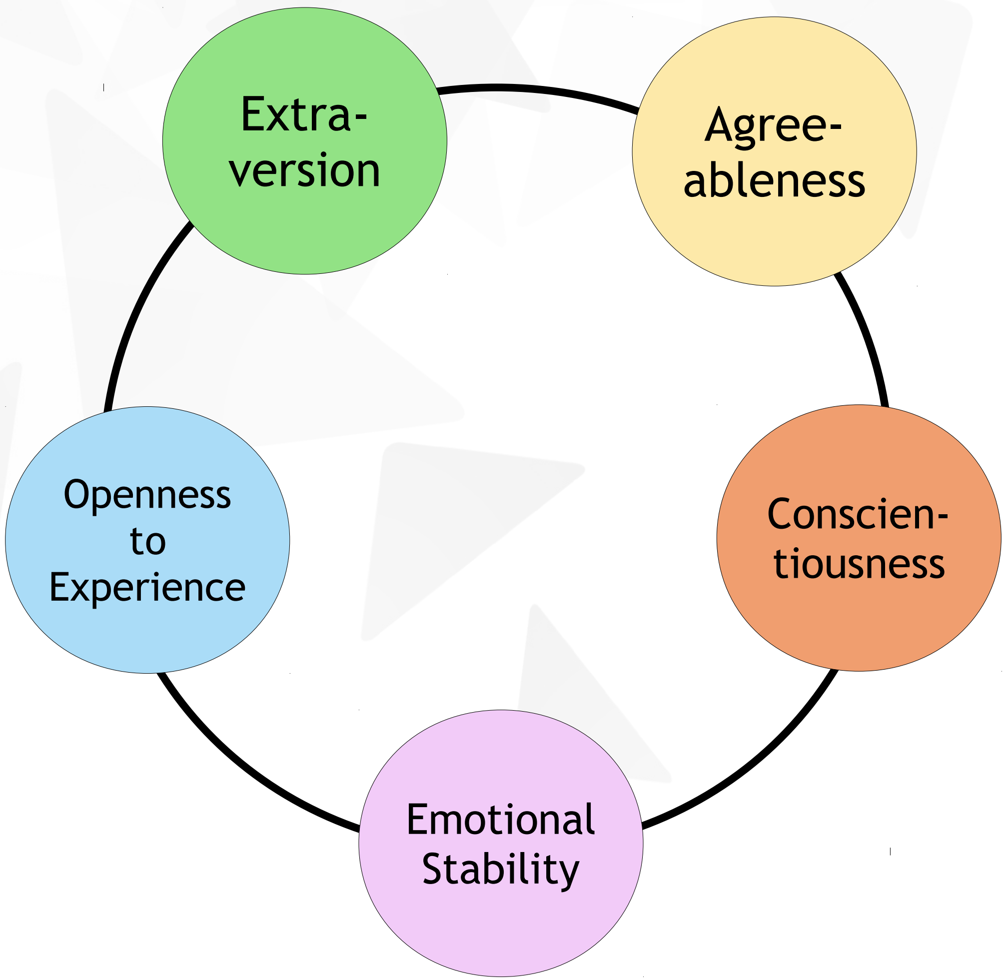
J.-I. Biel, O. Aran, and D. Gatica-Perez, “You Are Known by How You Vlog: Personality Impressions and Nonverbal Behavior in YouTube” in Int. Conf. on Weblogs and Social Media (ICWSM), 2011
J.-I. Biel and D. Gatica-Perez, „The YouTube Lense: Crowdsourced Personality Impressions and Audiovisual Analysis of Vlogs“, Transactions on Multimedia, 2012

Source: http://images4.fanpop.com/image/photos/18000000/Personality-Test-personality-test-18054186-400-327.jpg

Social Psychology: The big-five Personality Traits

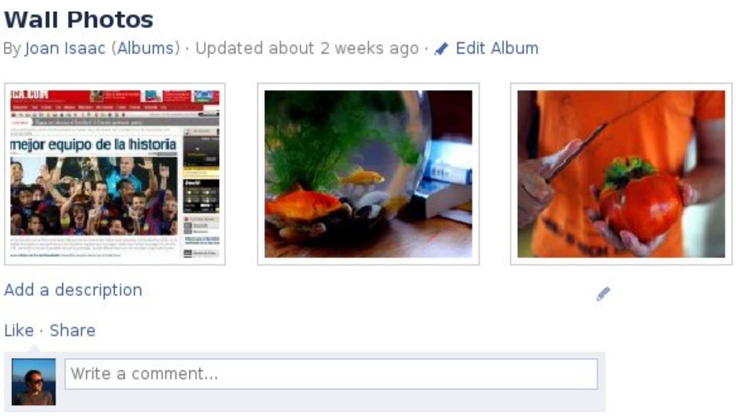
"the Big-Five traits have been broadly accepted as a way of presenting all the major traits of a person at the highest level of abstraction"

Gosling et al., 2003



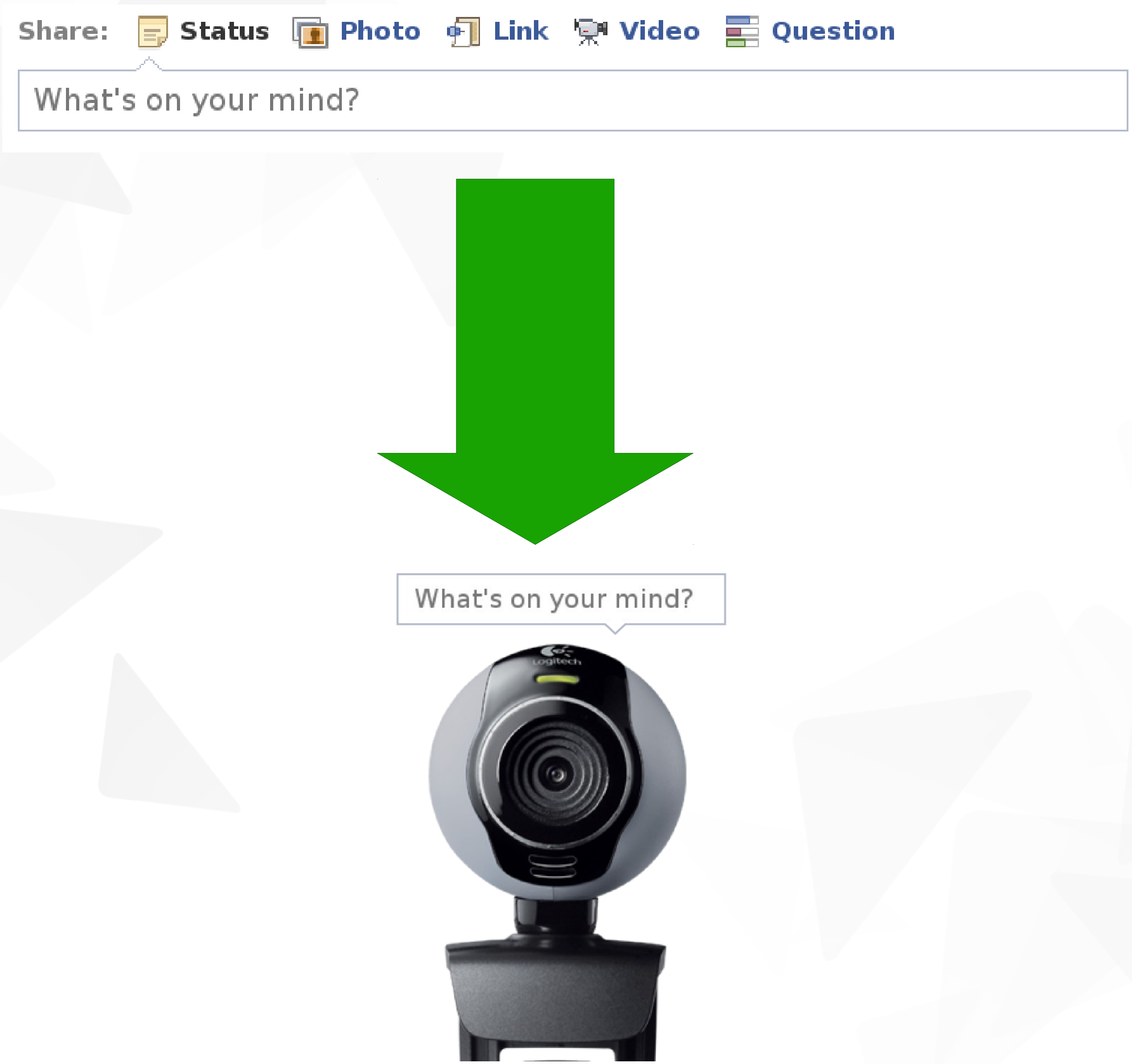
- **Extraversion**: talkative, assertive, energetic
- **Agreeableness**: good-natured, co-operative, trustful
- **Conscientiousness**: orderly, responsible, dependable
- **Emotional Stability**: calm, not neurotic, not easily upset
- **Openness to Experience:** intellectual, imaginative, independent-minded.

Social Media and Social Behaviour



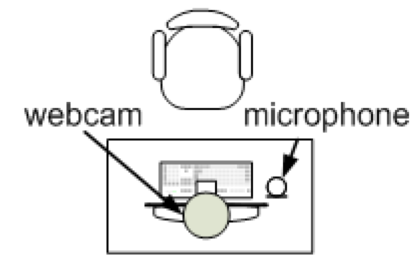
Most behavioral research in social media has sensed text interactions

From text to video:
The video blogging revolution…



vlogs=video blogs

- Multimedia life documentary and communication tool
- Rich social media behavior: verbal and nonverbal
- Resembles face to face interaction and skype
- Huge variety of content



How does Non-verbal behaviour in vlogs relate to personality impressions?

http://www.rickey.org/anthony-fedorov-video-blog/

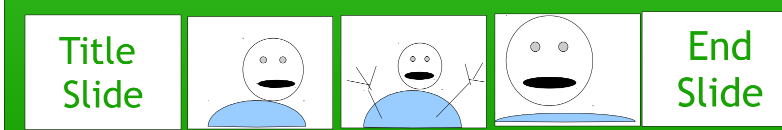
http://static.onemansblog.com/wp-content/uploads/2007/09/Robyn_Tippins.jpg

Personality Estimation from Vlogs : Flow Diagram

Cue Extraction

Audio cues:
Binary Speaking Activity
Prosodic cues

Video cues:
Looking, Camera proximity, Motion

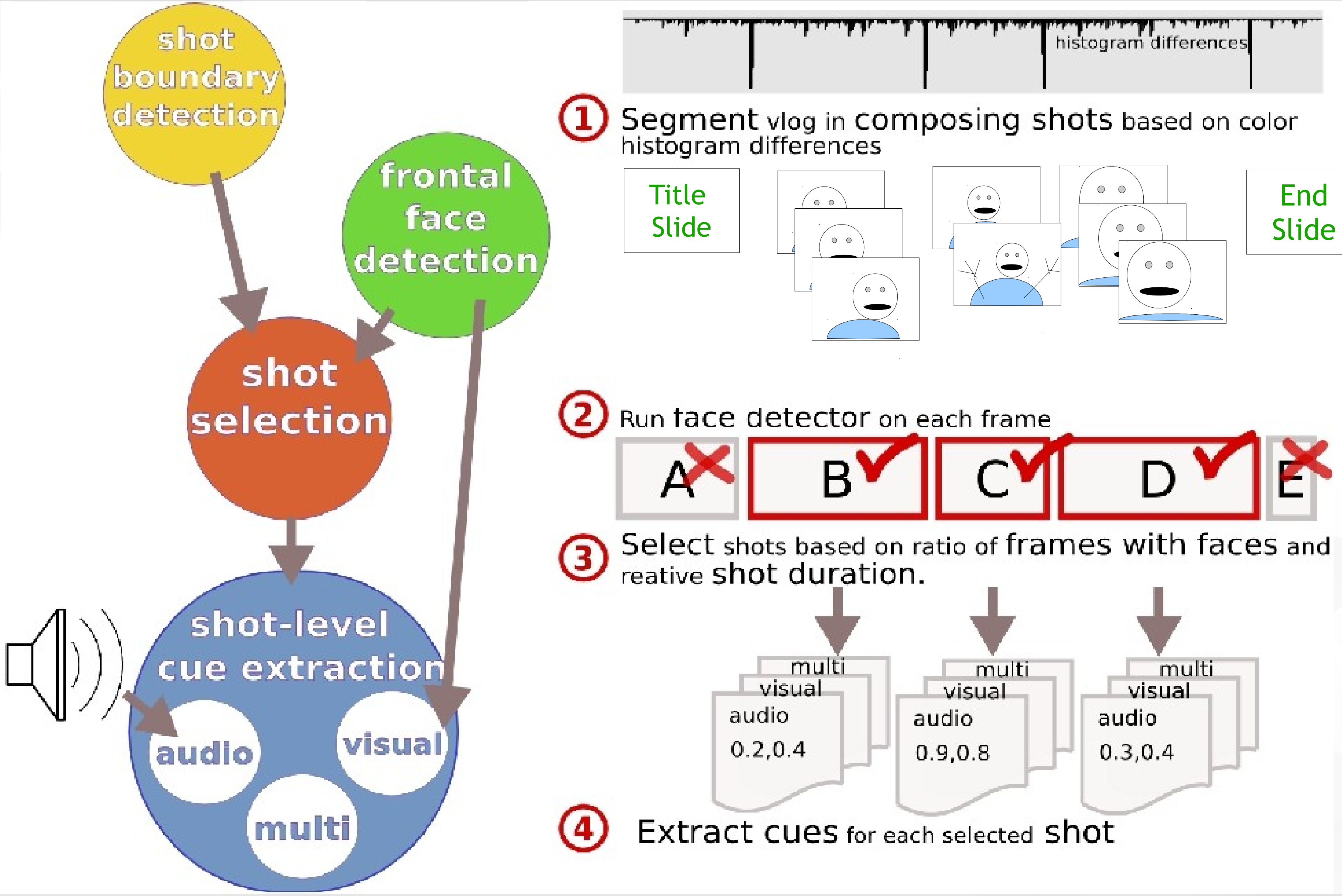


Estimate Personality

Personality Annotations:
Mechanical Turk

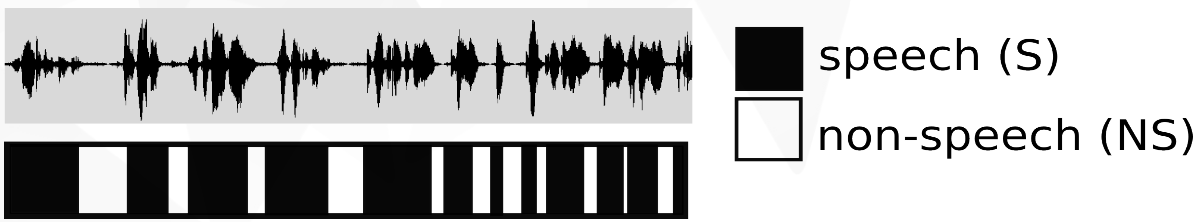
Evaluation

Automatic Vlog Processing



Audio cues

- speaking activity features are computed from speech/non speech segmentations



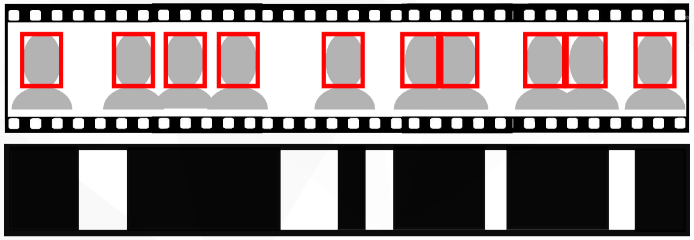
Speaking Time
Length of Speaking Seg
Speak Rate
Speech Turns

- voice quality measures are extracted from speech segments

Energy
Pitch
Autocorrelaton peak

Visual Cues

- Looking/non-looking segmentations from face detection



frontal face detection

looking (L)

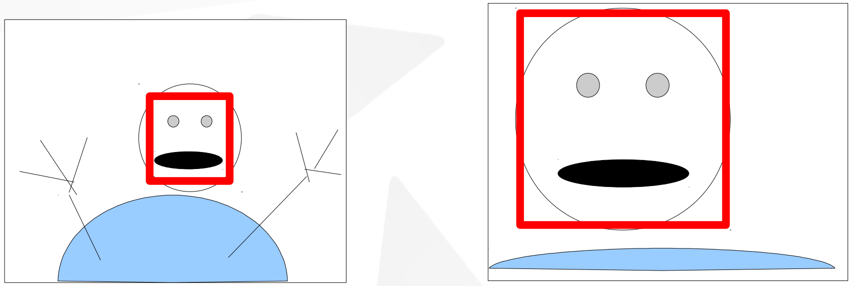
not-looking (NL)

looking time

length of looking segments

looking turns

- face detection bounding box

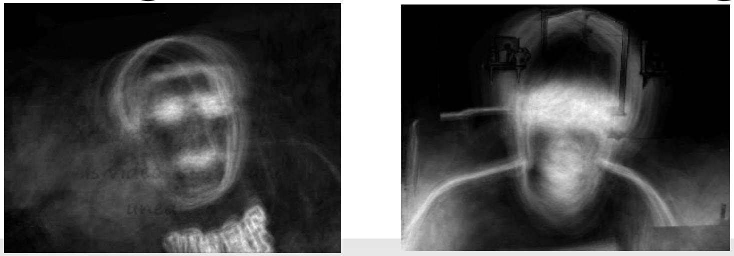


proximity to camera

framing

head motion

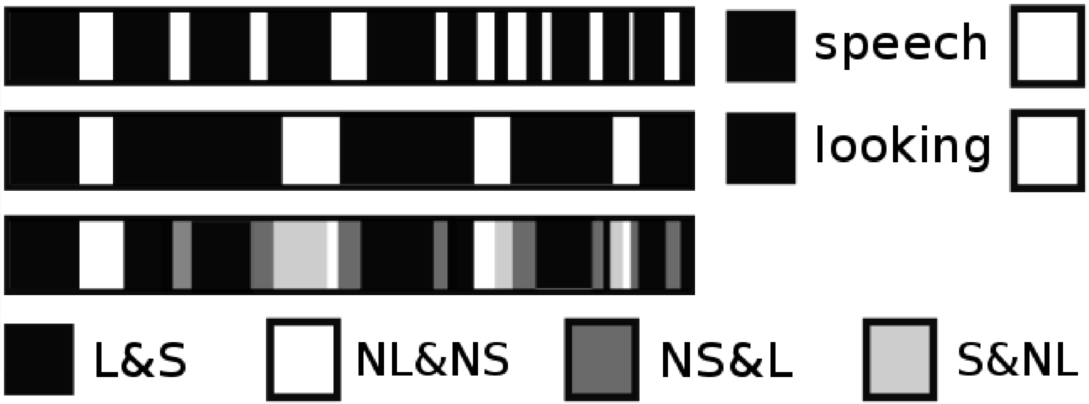
- weighted motion energy images



amount of motion throughout video

Multimodal Cues

- multimodal segmentations from audio and video segmentations



L = looking
NL = not-looking
NS = not-speaking
S = speaking

Time Look & Speak

Time Look & Not speak

Dominance ratio:
L&S/L&NS
(Dovido and Ellyson,1982)

Dovidio and Ellyson. 1982. Decoding visual dominance: Attributions of power based on relative percentages of looking while speaking and looking while listening. Journal of Social and Personal Relationships 45, 2, 106–113

Crowdsourced Personality Annotations

- Data:442 vlogs
- Annotation:
  - First minute of conversational video
  - Questionnaire based on zero-acquaintance judgments of personality
  - 2210 HITs (5x442)
  - 113 workers from US & India
  - Highest annotator agreement for extraversion trait.

Annotating the big-five: Ten-Item Personality Instrument (Gosling, 2003)

Please, INDICATE HOW MUCH YOU AGREE OR DISAGREE with each one of the following STATEMENTS about the person in the video.

(!) Rate the extent to which the pair of the trait applies to the person, even if one characteristic applies more stronlgy than the other.

STATEMENTS:

You see the person in the video as...

P1. Extraverted, enthusiastic	1-Disagree strongly	2	3	4	5	6	7-Agree strongly
P2. Cristical, quarrelsome	1-Disagree strongly	2	3	4	5	6	7-Agree strongly
P3. Dependable, self-disciplined	1-Disagree strongly	2	3	4	5	6	7-Agree strongly
P4. Anxious, easily upset	1-Disagree strongly	2	3	4	5	6	7-Agree strongly
P5. Open to new experiences, complex	1-Disagree strongly	2	3	4	5	6	7-Agree strongly
P6. Reserved, quiet	1-Disagree strongly	2	3	4	5	6	7-Agree strongly
P7. Sympathetic, warm	1-Disagree strongly	2	3	4	5	6	7-Agree strongly
P8. Disorganized, careless	1-Disagree strongly	2	3	4	5	6	7-Agree strongly
P9. Calm, emotionally stable.	1-Disagree strongly	2	3	4	5	6	7-Agree strongly
P10. Conventional, uncreative	1-Disagree strongly	2	3	4	5	6	7-Agree strongly

Vlog Personality Prediction Results

- Support Vector Machine Regression
  - A. Smola and B. Schlkopf, “A Tutorial on Support Vector Regression,”Royal Holloway College, University of London, Tech. Rep., 1998.
- standard prediction figure

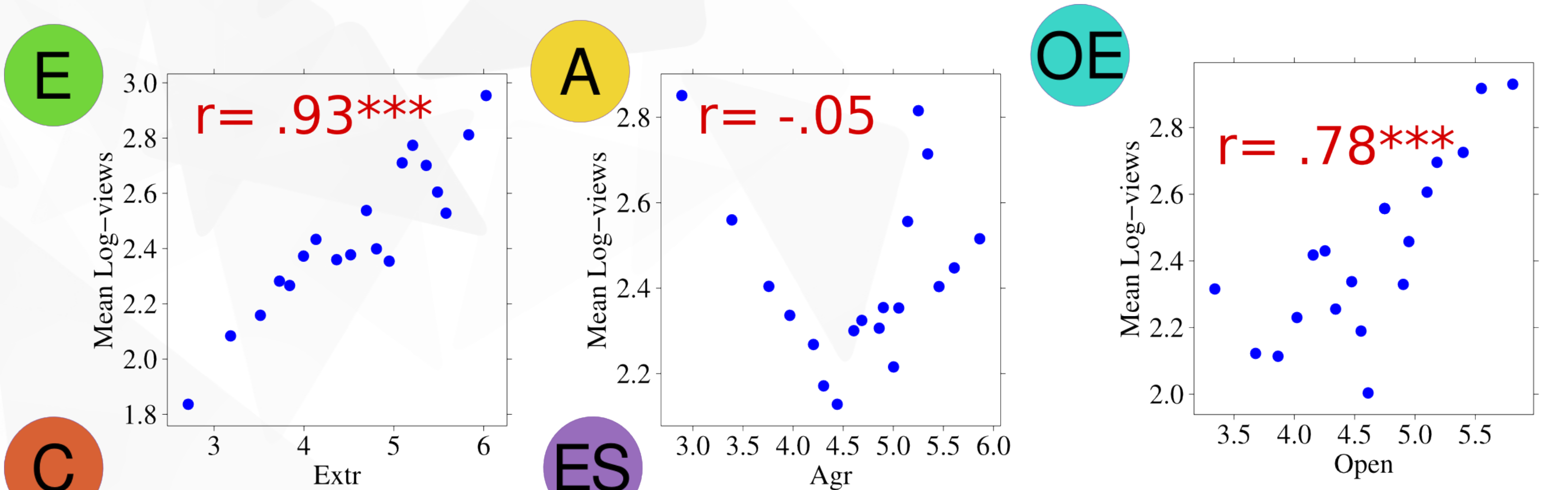
$$R^2 = 100 \times \left(1 - \frac{\sum (y_{obs} - y_{pred})^2}{\sum (y_{obs} - \bar{y}_{obs})^2}\right)$$

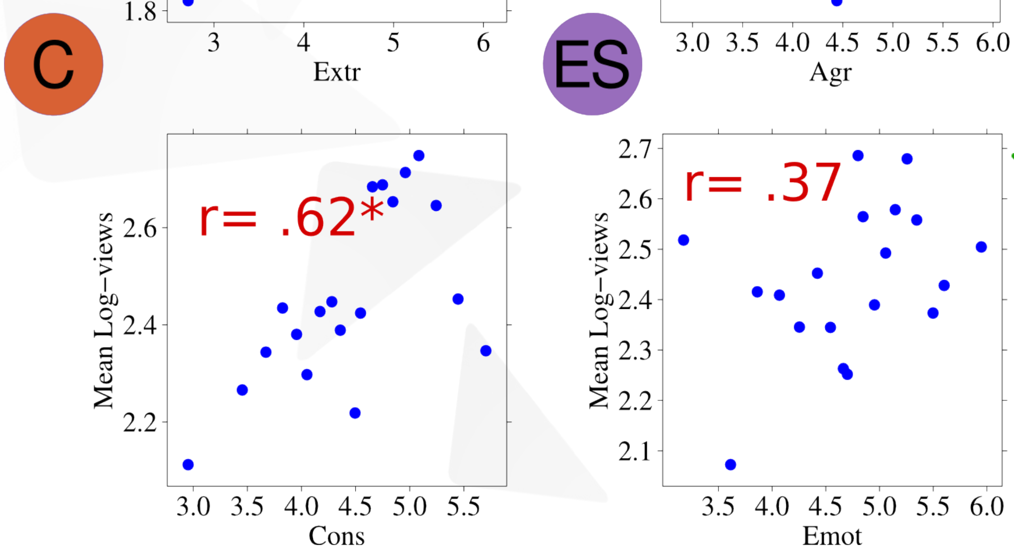
Baseline performance

R-squared measures how much better than the baseline are we?

- Best performances: Extraversion (36%), Conscientiousness (9%), Openness to Experience (10%) using audio and video features combined
  - Surprisingly, agreeableness not so easy, though social psychology studies tend to have high agreement.
    - Extracted cues were not informative.

Can your personality affect the popularity of your vlog?





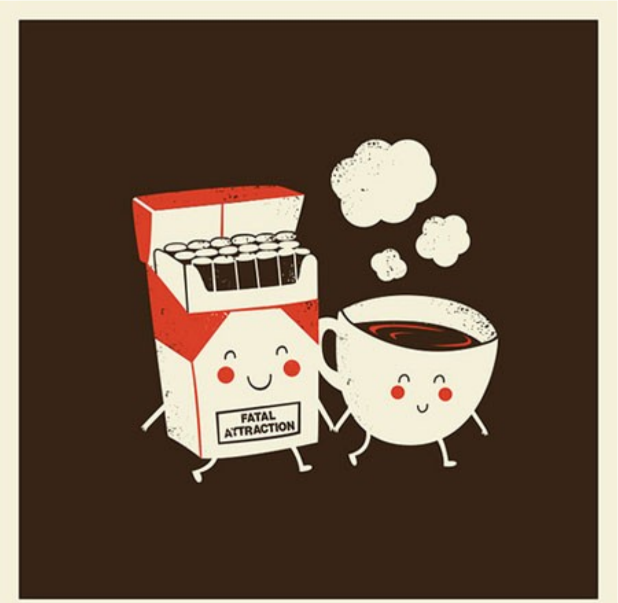
correlations with K = 50 bins
*p < .01, **p < .001, *** p < 0.001

- Higher average levels of attention in YouTube:
  - more Extraverted, Open to experience, and Conscientious vloggers
  - “nasty” & “pleasant” vloggers

Vlog Personality Prediction: Conclusion

- We can extract non-verbal cues from vlogs
  - simple yet robust method
  - need techniques for subtle cues
  - But does not account for personal/cultural differences
- Personality Prediction with vlogs:
  - some nonverbal cues are correlated with personality judgments.
  - Extraversion is easiest to predict, followed by OE and C.
  - Agreeableness: smiling is generally reported to be useful and could be tried in the future.

3. Estimating Attraction



Veenstra and Hung, “Do They Like Me? Using Video Cues to Predict Desires during Speed-dates” in ICCV Workshops 2011

Source: http://catinbag.blogspot.nl/2010/07/fatal-attraction.html

Finding a mate...

- Finding a partner can be difficult, dating services/sites are abundant
- Interpersonal communication in general and dating in particular is often guided by misperception and misinterpretation (Ranganath et al. 2009).
- Speed-Dating
- Support human-human interaction by analysing behavior and giving feedback

Chris Hondros/Getty Images

Speed Dating, Non-verbal cues and Attraction

- Can proximity-related video cues be used to automatically predict attraction in speed-dates?

Predicting Attraction: Related work

- Audio cues to predict outcomes in speed dates:
  - Madan, Caneel and Pentland, “Voices of attraction”, Aug Cog,2005.
  - Cues used: Activity, engagement, emphasis and mirroring
- Audio and Linguistic differences between intention and perception during speed dates
  - Ranganath, Jurafsky and McFarland,“It's Not You, it's Me: Detecting Flirting and its Misperception in Speed-Dates”,EMNLP, 2009
- Visual Motion energy for observing courtship communication:
  - Grammer et al. (1999) Fuzziness of Nonverbal Courtship Communication Unblurred by Motion Energy Detection

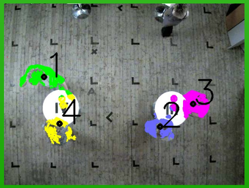
Non-verbal Cues of Attraction

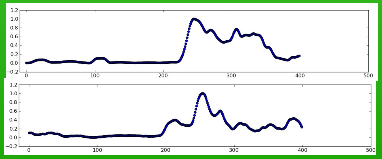
- Behavioral synchrony or mimicry
  - indicates affiliation, attraction, rapport.
  - “The chameleon effect: The perception-behavior link and social interaction.”, Chartrand and Bargh, Journal of Personality and Social Psychology, 1999
- Closer **proximity**, more direct **orientation**, more **gaze**, more mutual gaze, more **smiling**, more **head nods**, lively **movement**, open arms stretched towards other, more personal touching, higher pitch...etc.
  - (Argyle, “Bodily Communication”,1988)
- Our focus today is proximity and movement.

Speed Dating Analysis:Flow Diagram

Cue Extraction

Video cues:
Proximity, Motion, Synchrony





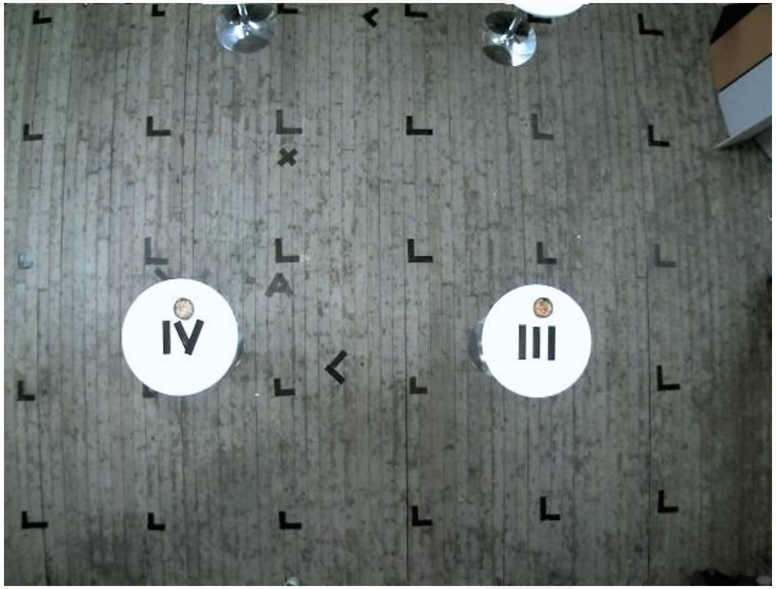
Estimation
Attraction
Exchanging Contact Information

Personality Annotations:
Self, Other perceptions

Evaluation

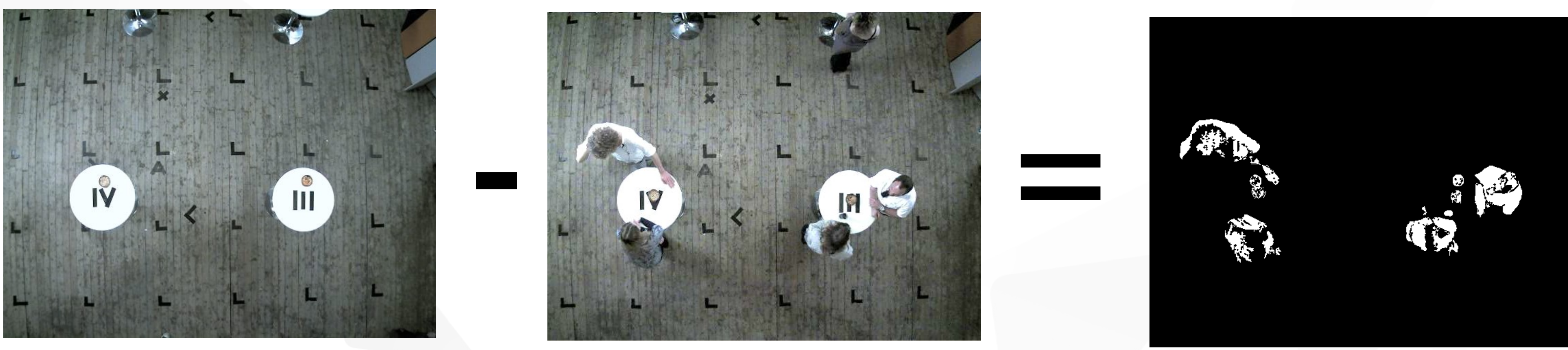
Automated Position Extraction

- Construct eigenbackground



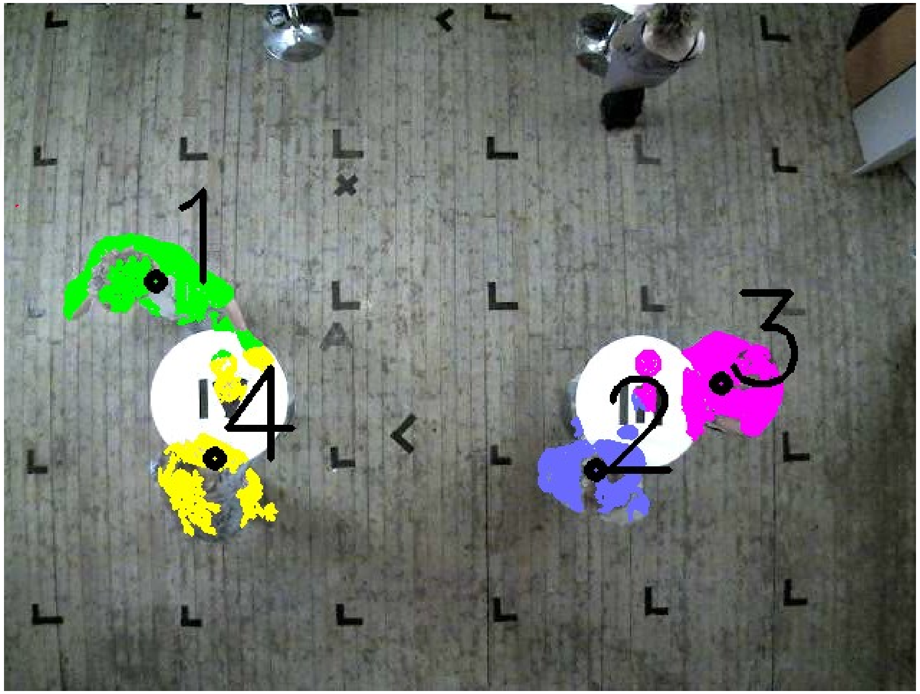
Automated Position Extraction

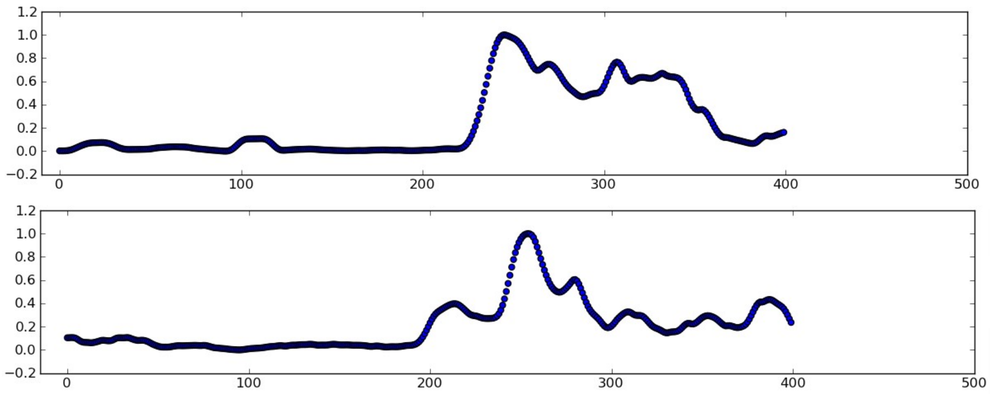
- Construct eigenbackground
- Subtract eigenbackground



Automated Position Extraction

- Construct eigenbackground
- Subtract eigenbackground
- Cluster points (k-means) and find centres
- Sanity check for irregularities (temporal smoothing)



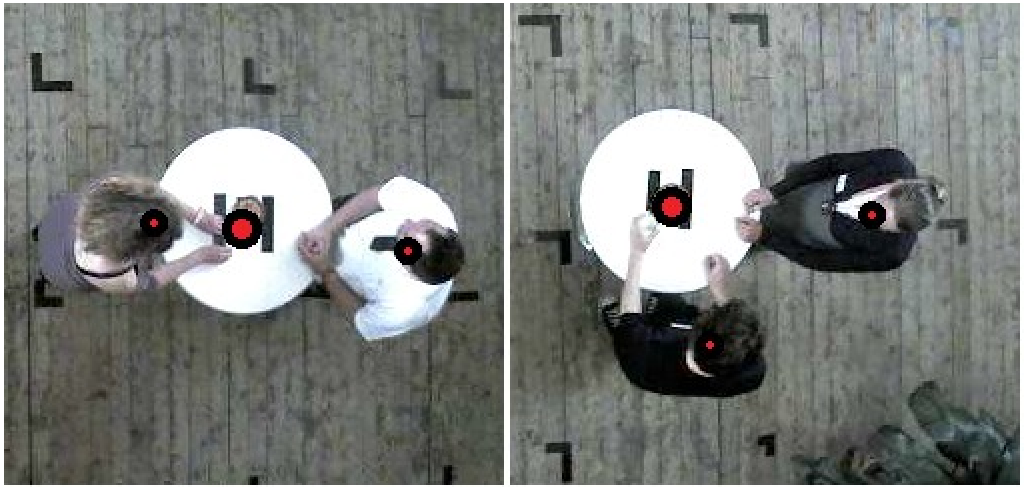


Position-based Behavioural Cues

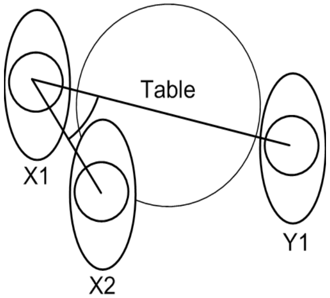
Position	
AVG-DIFANGLE	Average angle between participants with respect to table
Movement	
VARDIS	Variance in distance
VAR-DIFANGLE	Variance in angle between participants with respect to table
VARPOS	Variance in position
VARPOS-OTHER	Variance in position of the other
DECRDIS	Decrease in distance
MOVDISTR	Movement distribution
MOVDISTR-OTHER	Movement distribution of other person
Distance	
AVGDIS	Average distance
Synchrony	
MOTIONSYNC	Synchrony in motion
MOTION-REACTION	Distribution of motion reaction
MOTION-REACTION-OTHER	Distribution of motion reaction of the other

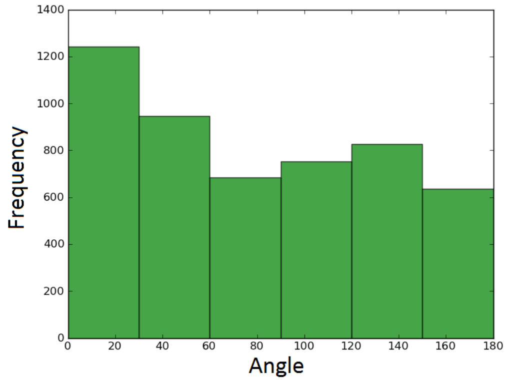
Derive Position Features: Selected Examples

Average angle between participants with respect to table

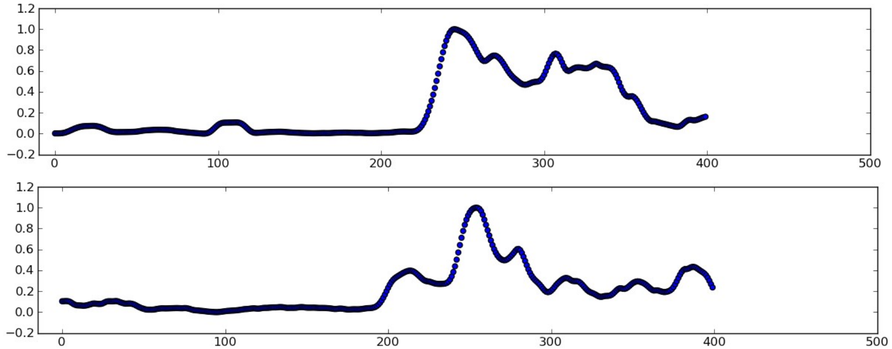


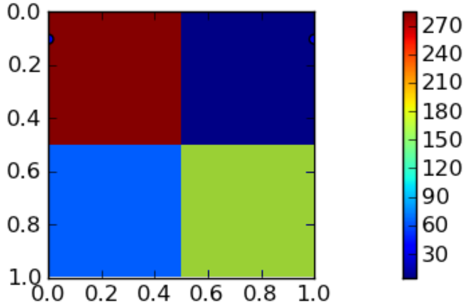
Movement distribution: Angular direction of X relative to Y over the date.





Synchrony: How often does the motion of X and Y, or not match?





Data

- 5 minute speed dates
- Alternated with questionnaire answering
  - Questions from interpersonal attraction scale from McCrosky and McCain (1974)
- 16 participants:8 male 8 female
  - 64 dates for our experiments

Attraction Experiments

- Support Vector Machine used for classification.
- Experiments were split by gender
  - Makes sense from a psychological and biological perspective (e.g. Grammer et al. 1999, Buss and Schmitt 1993)
- Baseline was created by labelling all test items as the most frequent class

Speed Dating Results

- Predicting attraction
  - Variance in position is best feature predictor for women (70%).
  - Variance in position of the women and synchrony both perform well (70%) for men.

Fusion of all movement features

	Male	Female
	SVM	SVM
AVG-DIFANGLE	0.50	**0.66**
Movement	0.55	0.55
VARDIS	0.59	0.39
VARDIFANGLE	0.33	**0.59**
VARPOS	0.59	**0.70****
VARPOS-OTHER	**0.70***	**0.61**
MOVDISTR	0.59	**0.63**
MOVDISTR-OTHER	0.55	0.44
AVGDIS	0.42	0.30
Synchrony	**0.70***	0.31
MOTIONSYNC	0.53	**0.63**
Baseline	0.59	0.55

Fusion of all synchrony features

Speed Date Experiments : Conclusion

- The video channel can indeed be a source of valuable information in speed-dates
- Results differ per gender:
  - Movement synchrony information is more important for males than females.
  - For females, information on the movement of their male counterpart gives good results

4. Identifying Conversing Groups (F-formation Detection)

Source: http://www.aboutleaders.com/bid/141733/Tips-for-Communication-Skills-with-Groups

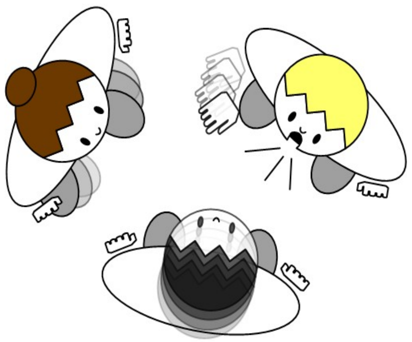
Hung and Krose, “Detecting F-formations as Dominant Sets”, ICMI, 2011

What is an F-formation?

"whenever two or more individuals in close proximity orient their bodies in such a way that each of them has an easy, direct and equal access to every other participant's transactional segment, and when they maintain such an arrangement, they can be said to create an F-formation" (p. 243) Ciolek and Kendon 1980

Why is F-formation detection relevant?

- So far, we have seen scenarios with pre-determined, fixed numbers of participants.
- What if the setting is free? How do we know who has the potential to influence whom?
- Relevant for public space monitoring
- **Mutual** co-operation to ensure equal access to shared space indicative of **relationships**.



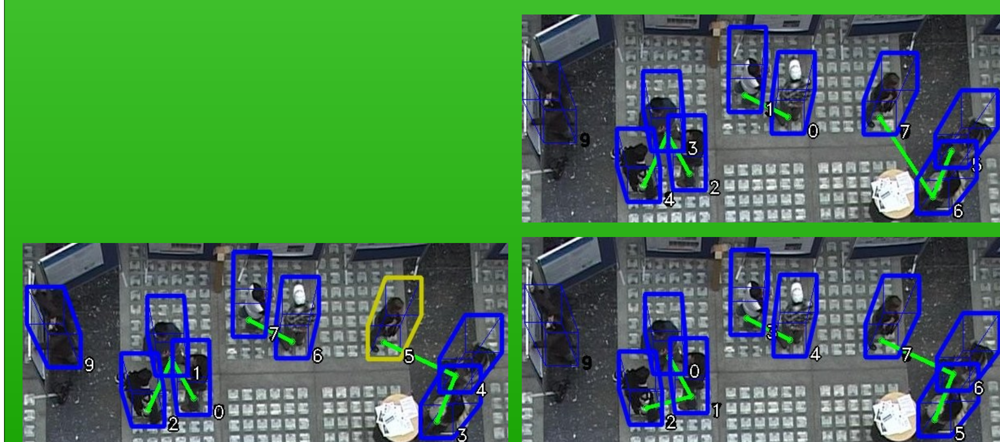
Identifying F-formations : Flow Diagram

Data:
82 still images from poster and coffee session.
~50 people

Behavioural Cues
Proximity
Orientation

F-formation Detection
Dominant set
Modularity cut (baseline)

Annotation:



Evaluation

Behavioural Cues and F-formations

- Position (manual annotation)
- Body Orientation (manual annotation)
- Proximity :

$$A_{ij}^{prox} = -e^{\frac{d_{ij}}{2\sigma^2}}$$

d_{ij} is the distance between person i and j

σ defines width of the Gaussian kernel surrounding each person

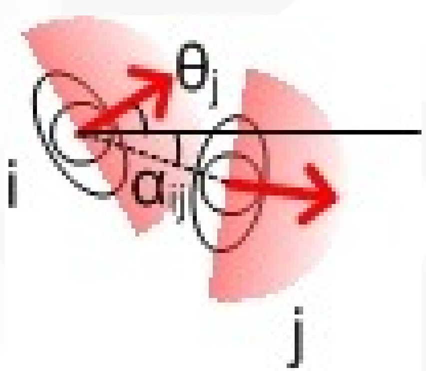
- Proximity and Orientation :

$$A_{ij}^{ori} = \underset{\mathbf{q}}{\operatorname{argmin}}(\begin{cases} e^{-\frac{d_{\mathbf{q}}}{2\sigma^2}} & \text{if } -\frac{\pi}{2} \geq \theta_{\mathbf{q_1}} - \alpha_{\mathbf{q}} \geq \frac{\pi}{2} \\ 0 & \text{otherwise} \end{cases},$$

$\forall \mathbf{q} \in \{(i,j),(j,i)\}$, and $\mathbf{q_1}$ is the first element of $\mathbf{q}$

θ_i body orientation angle of person i

$\alpha_{\mathbf{q}}$ is the angle of the vector from i to j



Behavioural Cues and F-formations

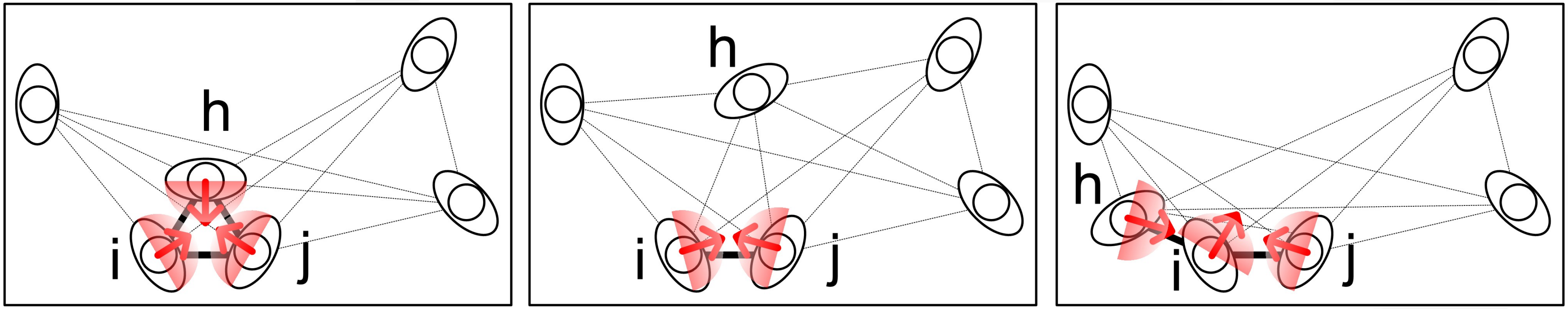
- SMEFO: Socially Motivated Estimate of Visual Focus (Semi-Automatic)
  - Estimate of body orientation from position information only.
  - Centre of visual focus:

$$f_i = \frac{1}{k_i} \sum_j p_j A_{ij}^{prox} \quad , \; k_i = \sum_j A_{ij}^{prox}$$

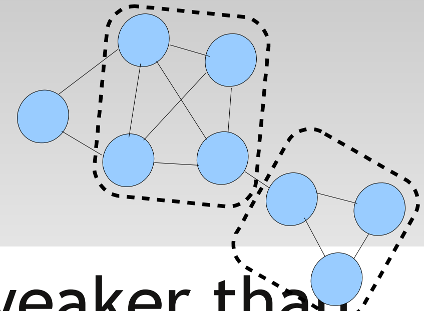
  - Estimated focus angle:

$$\gamma_i = \arccos(\alpha(p_i, f_i))$$

where $\alpha(p_i, f_i)$ is the angle of the vector from person i (p_i) to f_i



F-formation Detection



- **Modularity Cut Clustering:** Cut edges based on weaker than expected connections
  - Yu et al. “Monitoring, recognizing and discovering social networks”, CVPR 2009
  - Grouping based on pairwise connections being **more than expected connection** with the entire network.
  - **Global optimisation** based on recursive bisection.
- F-formation is like a maximal clique
  - Maximal clique: A cluster of nodes in a graph that are **fully connected** and cannot be enlarged.
  - Edge weights measured as inter-personal affinity.
- **Dominant set:** A maximal clique in an edge-weighted graph
  - Pavan and Pelillo, “Dominant sets and pairwise clustering”, IEEE PAMI, 2007.
  - Exploits clique context when grouping people - **local optimization**

Coffee and Poster Session Data and Annotation

- 82 images selected for annotation and evaluation.
  - No consecutive images contained exactly the same F-formations.
  - Tried to maximise crowdedness
  - Tried to maximise ambiguities from associates
  - > 1700 instances of people
- Annotation:
- 24 annotators grouped into 8 triads
  - 3 annotators per image

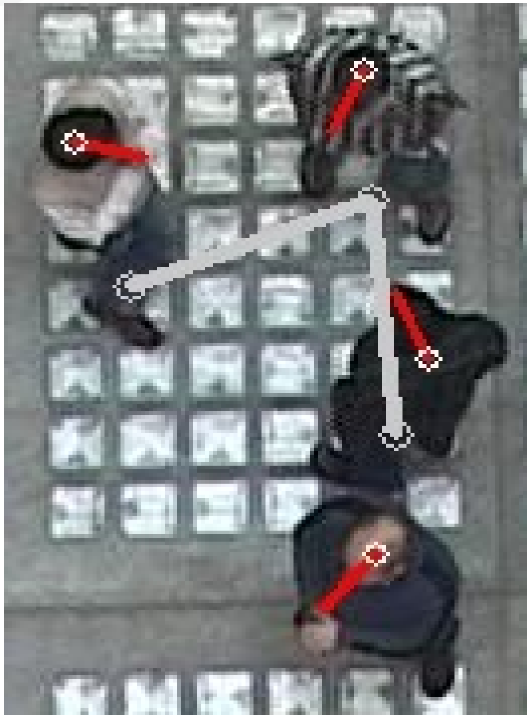
Experimental Results

Cues	Methods	F-Measure
Proximity **Baseline (Yu et al. 2009)**	Dom. Set	**86.83**
	Mod. Cut	76.57
Proximity + Orientation	Dom. Set	**92.24**
	Mod. Cut	92.02
Proximity + SMEFO	Dom. Set	**86.50**
	Mod. Cut	81.44
Label Everyone as Singletons		75.57

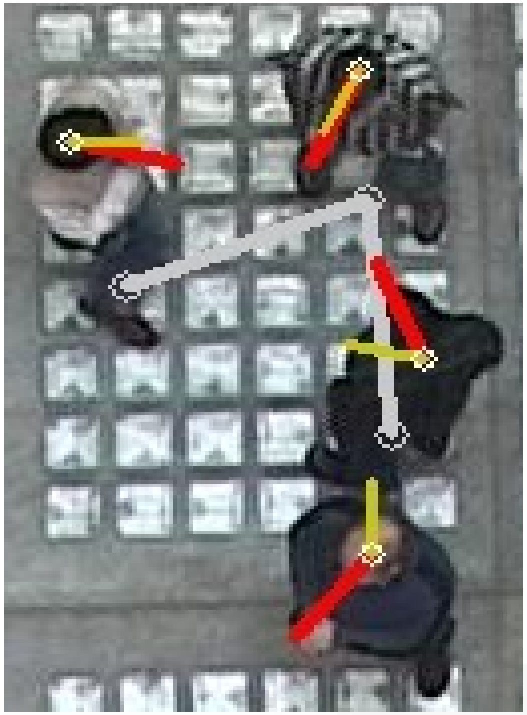
SMEFO Orientation Estimate Error: 12.9° (±14.3)

How the SMEFO Can Help

MC



MC+
Orientation



MC +
SMEFO

Detected F-formation

SMEFO

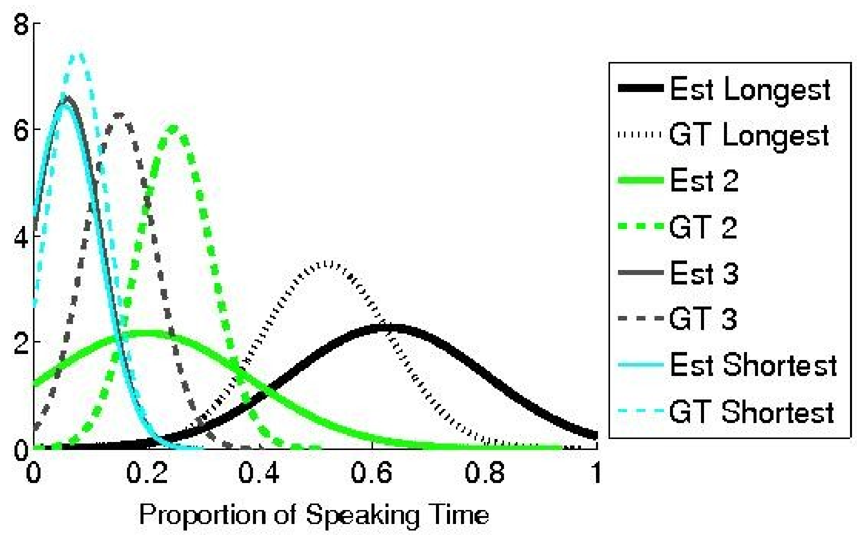
Labeled Body Pose

F-formation Detection: Conclusion

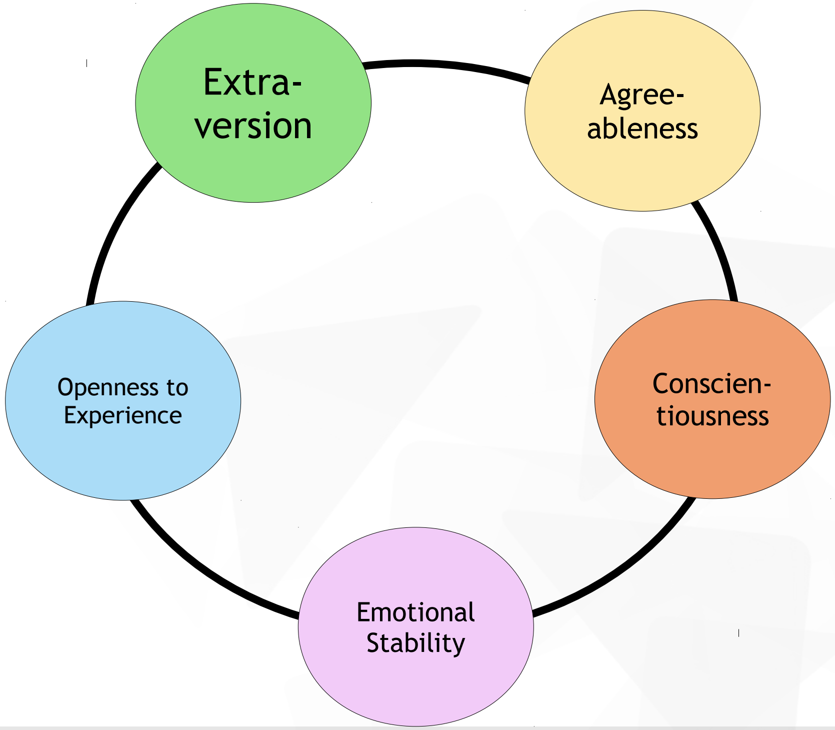
- Detecting F-formations using social context (dominant sets) leads to better performance than using global context (modularity cut).
- SMEFO does well when used with modularity cut but does not improve performance with dominant set method.
- Body orientation helps a lot in estimating F-formations.

Summary

- Dominance:
  - The way turns are regulated during discussions gives strong, automatically extractable patterns

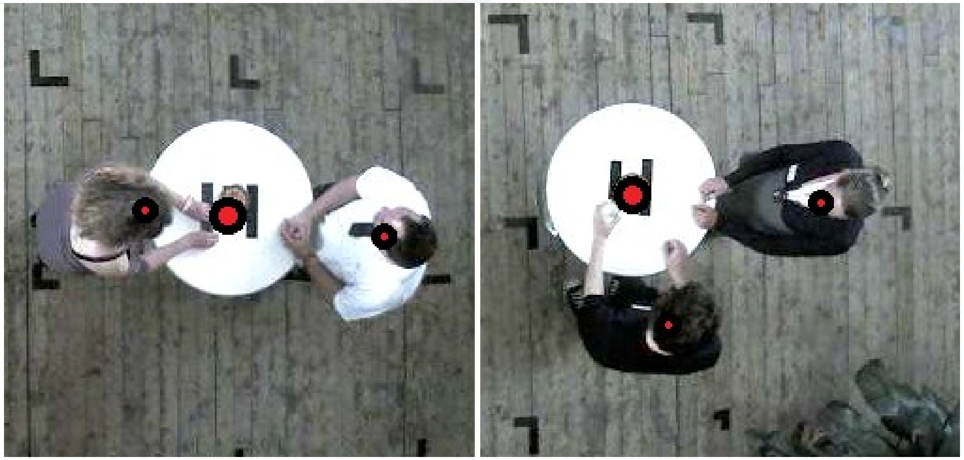


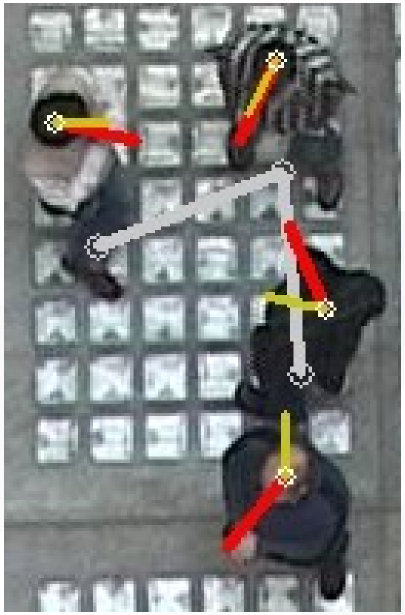
- Personality:
  - Simple behavioural cues extracted during vlogging can be discriminative for E,C, and OE.



Summary

- Attraction:
  - Cues as simple as proximity and movement can indicate a lot about attraction.
- F-formation detection:
  - Social context is a powerful prior for extracting behavioural cues.





Summary and Discussion

- Today we have seen simple cue extraction
  - Sometimes noisy but still robust for the task.
  - Sometimes too much data to validate (e.g. Social media)
  - Exploiting social context can improve cue extraction.
- Open Questions:
  - How can we deal sensibly with multiple annotations?
  - What is the ground truth?
  - Do more complex cues enable better advances in automated social behavior understanding?
  - If noisy feature extraction still works...
    - how far away can the sensors go? e.g. Lower resolution video
    - How simple could the sensor become? e.g. Just a single motion detector

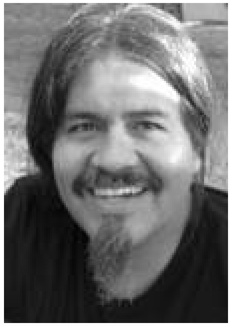
Slide Credits

- Thanks to the following slide contributors

Alessandro
Vinciarelli

Joan-Isaac
Biel

Arno
Veenstra



Daniel
Gatica-Perez