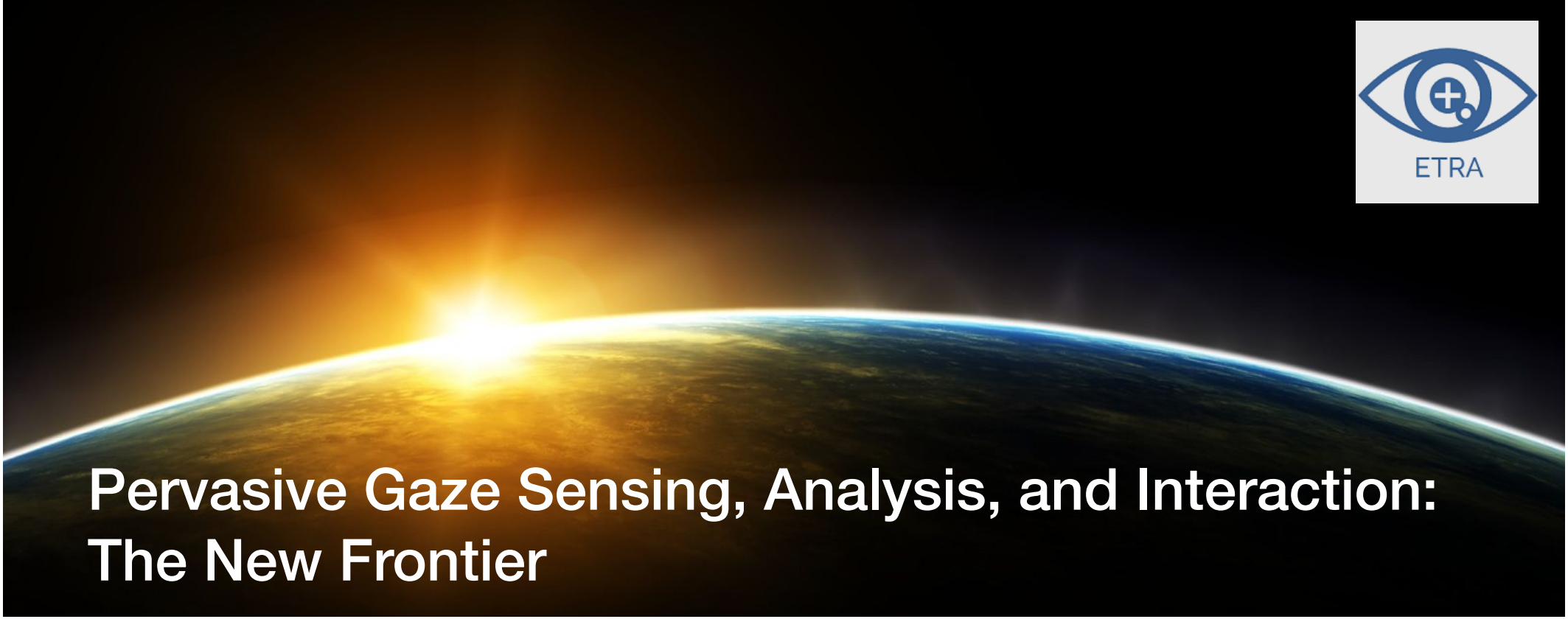


Pervasive Gaze Sensing, Analysis, and Interaction: The New Frontier

Andreas Bulling

Max Planck Institute for Informatics & Saarland University

bulling@mpi-inf.mpg.de

Pervasive eye tracking

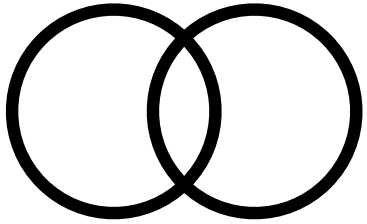
Rich user models

Multimodal sensing

24/7

Everyday

Natural interactions



Stationary/Mobile

Controlled settings

Calibration-free

Accurate

Robust

Pervasive eye tracking

Next generation human-machine systems that **sense, analyse, and adapt to users' gaze** in **all interactions** that users perform **in everyday life**.

Bulling & Gellersen: Toward mobile eye-based human-computer interaction, IEEE Pervasive Computing 2010

1st International Workshop on Pervasive Eye Tracking and Mobile Eye-Based Interaction

PETMEI 2011 – in conjunction with UbiComp 2011

PETMEI 2011 | Topics | Submissions | Organisers | Keynote | Program | Venue | Sponsors | Call for Papers

PETMEI 2011

Recent developments in mobile eye tracking equipment and automated eye movement analysis point the way toward unobtrusive eye-based human-computer interfaces that are pervasively usable in everyday life. We call this new paradigm *pervasive eye tracking* – continuous eye monitoring and analysis 24/7. The potential applications for the ability to track and analyze eye movements anywhere and anytime call for new research to further develop and understand visual behaviour and eye-based interaction in daily life settings.

PETMEI 2011 will focus on pervasive eye tracking as a trailblazer for mobile eye-based interaction and eye-based context-awareness. We provide a forum for researchers from human-computer interaction, context-aware computing, and eye tracking to discuss techniques and applications that go beyond classical eye tracking and stationary eye-based interaction. We want to stimulate and explore the creativity of these communities with respect to the implications, key research challenges, and new applications for

Important Dates

Paper Submission:
June 6, 2011 (closed)

Notification of Acceptance:
June 30, 2011

Camera-ready due:
July 18, 2011

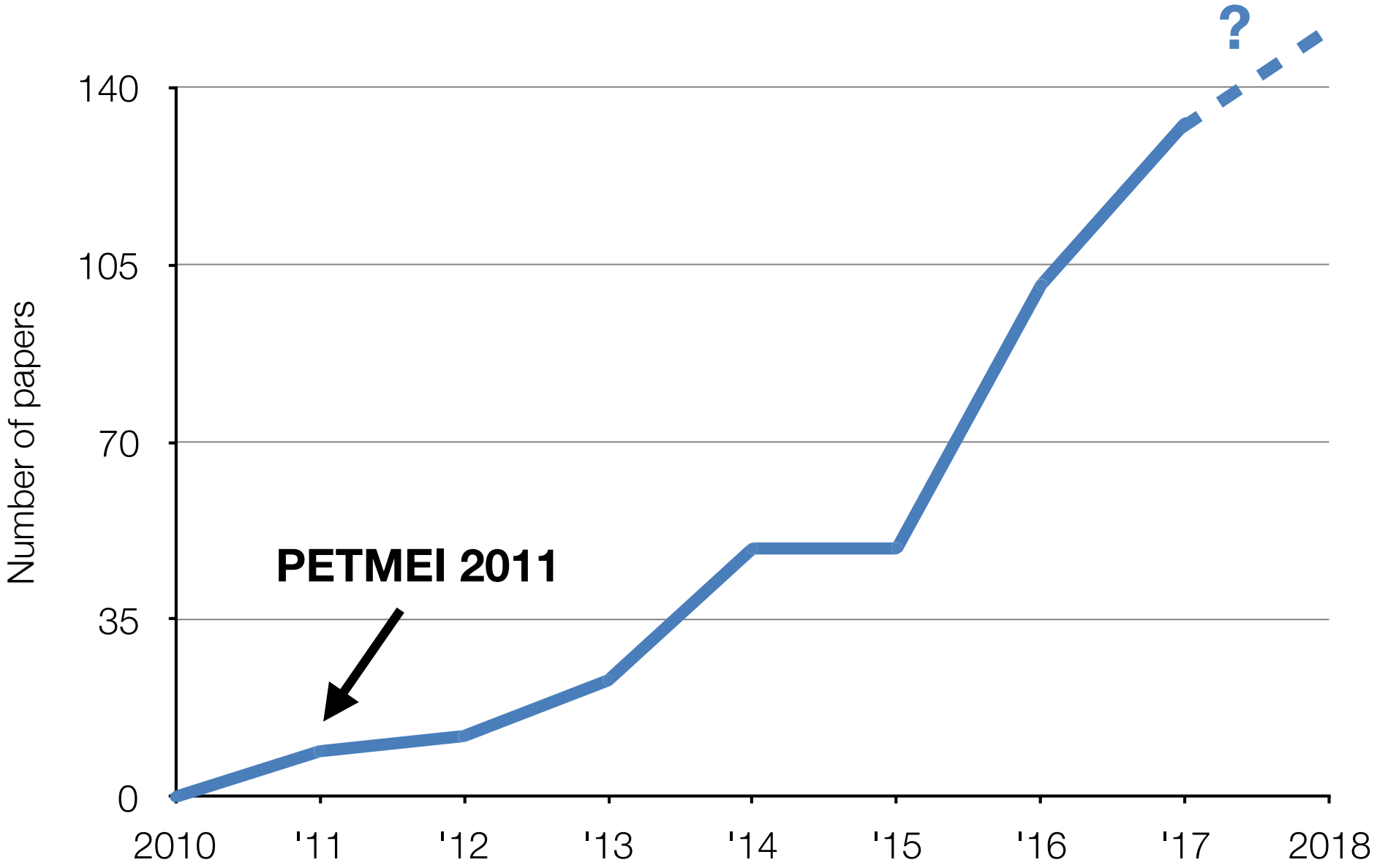
Workshop:
September 18, 2011

Contact

Andreas Bulling
Andrew T. Duchowski
Päivi Majaranta

www.petmei.org

Pervasive eye tracking: A growing topic



source: GoogleScholar

Goals

1. Exciting **new applications impossible before**
2. **Recent advances** in pervasive gaze sensing, analysis, and interaction
3. **Data-driven methods** are instrumental
4. **A lot still remains to be done** to fully realise the vision of pervasive eye tracking

Living in a pervasive eye tracking world

Living in a pervasive eye tracking world

Living in a pervasive eye tracking world

Living in a pervasive eye tracking world

Anticipatory systems

Living in a pervasive eye tracking world

Preference learning

Visual search support

Living in a pervasive eye tracking world

Autism diagnosis and monitoring

Quantified self

There is still a long way to go...


Pervasive eye tracking
Sensing
Analysis
Interaction

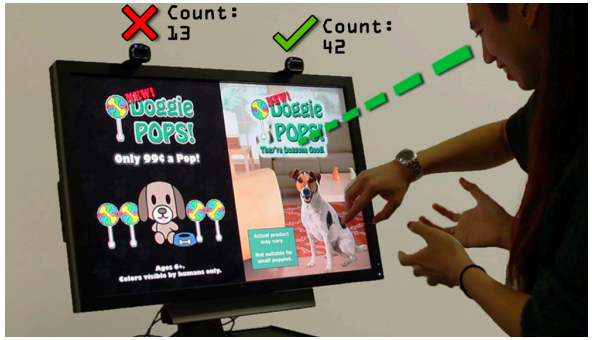
Goals

1. Exciting **new applications impossible before**
2. **Recent advances** in pervasive gaze sensing, analysis, and interaction
3. **Data-driven methods** are instrumental
4. **A lot still remains to be done** to fully realise the vision of pervasive eye tracking

Motivation

Human-human

Joint attention



Human-machine

- Eye contact is pervasive in social communication
  - Can signal attention, interest, attraction, ...
- Eye contact detection has potential for a range of applications
  - Personalised health (autism), intelligent tutoring systems, social robotics, ...

N.J. Emery, Neuroscience & Biobehavioral Reviews 2000

Everyday eye contact detection

- Previous methods required special-purpose eye tracking equipment and/or assumed controlled settings
- First method for eye contact detection in daily-life settings that only requires a single off-the-shelf camera

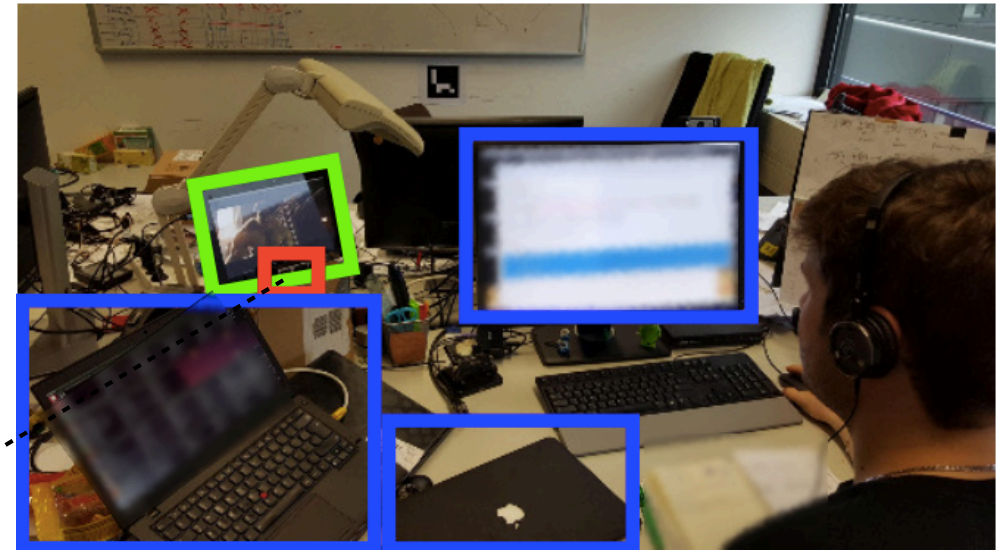
Zhang et al., Everyday Eye Contact Detection Using Unsupervised Gaze Target Discovery, *Proc. UIST 2017*
best paper honourable mention award

Everyday eye contact detection

https://www.youtube.com/watch?v=ccrS5XuhQpk

Unsupervised gaze target discovery

Training images

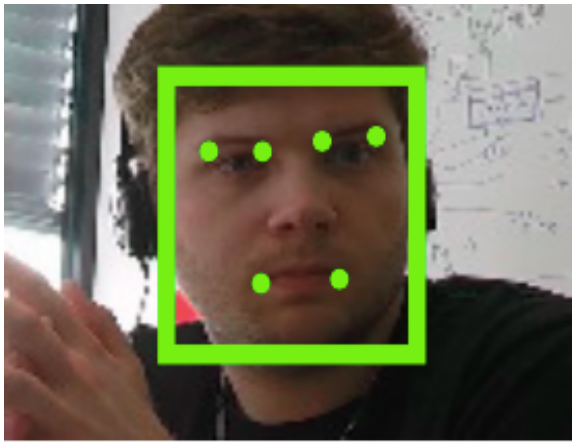


Target object

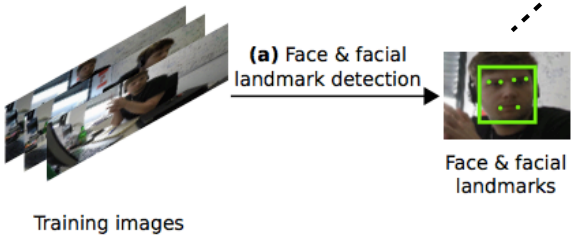
Camera

Other distractive objects

Unsupervised gaze target discovery

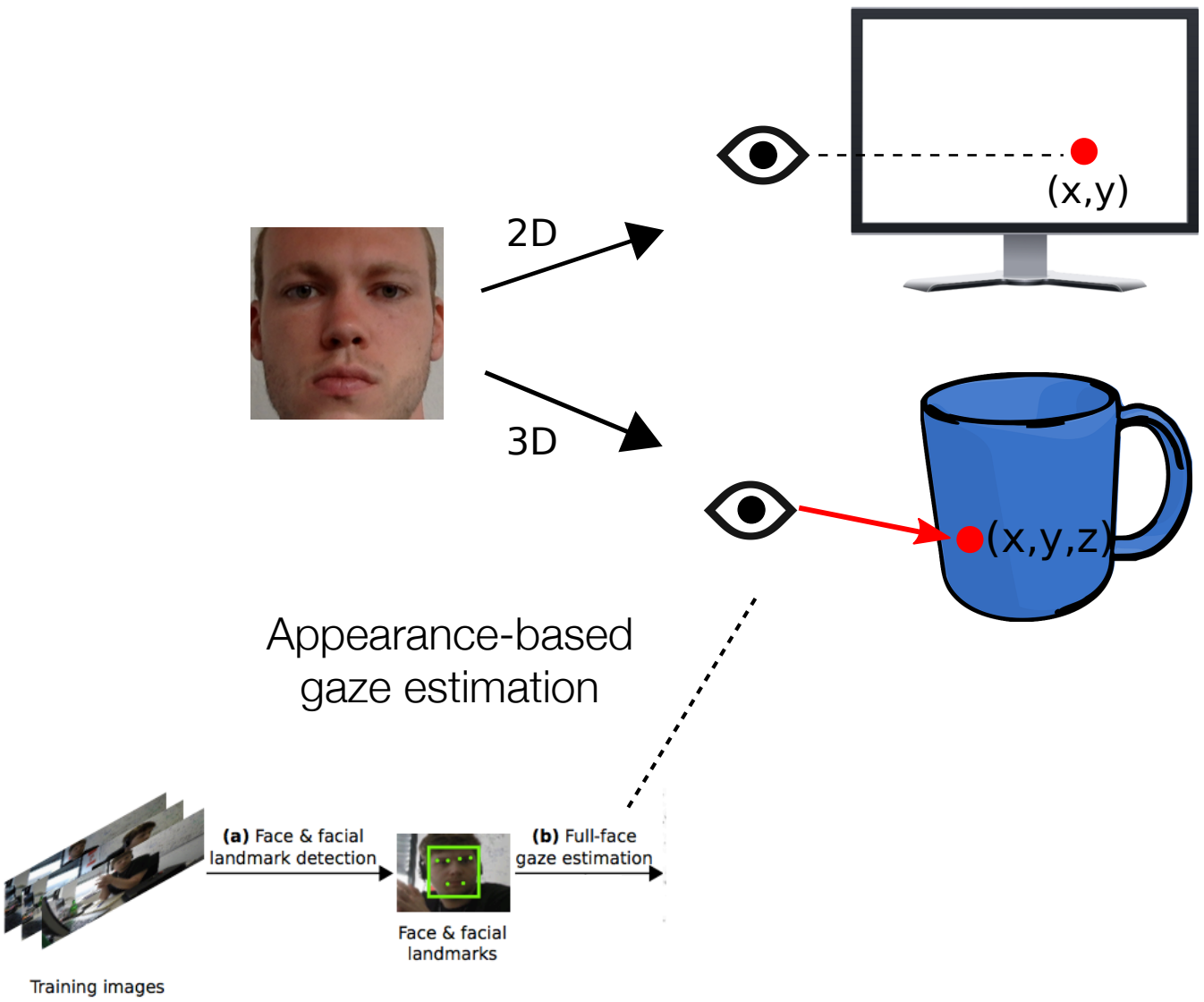


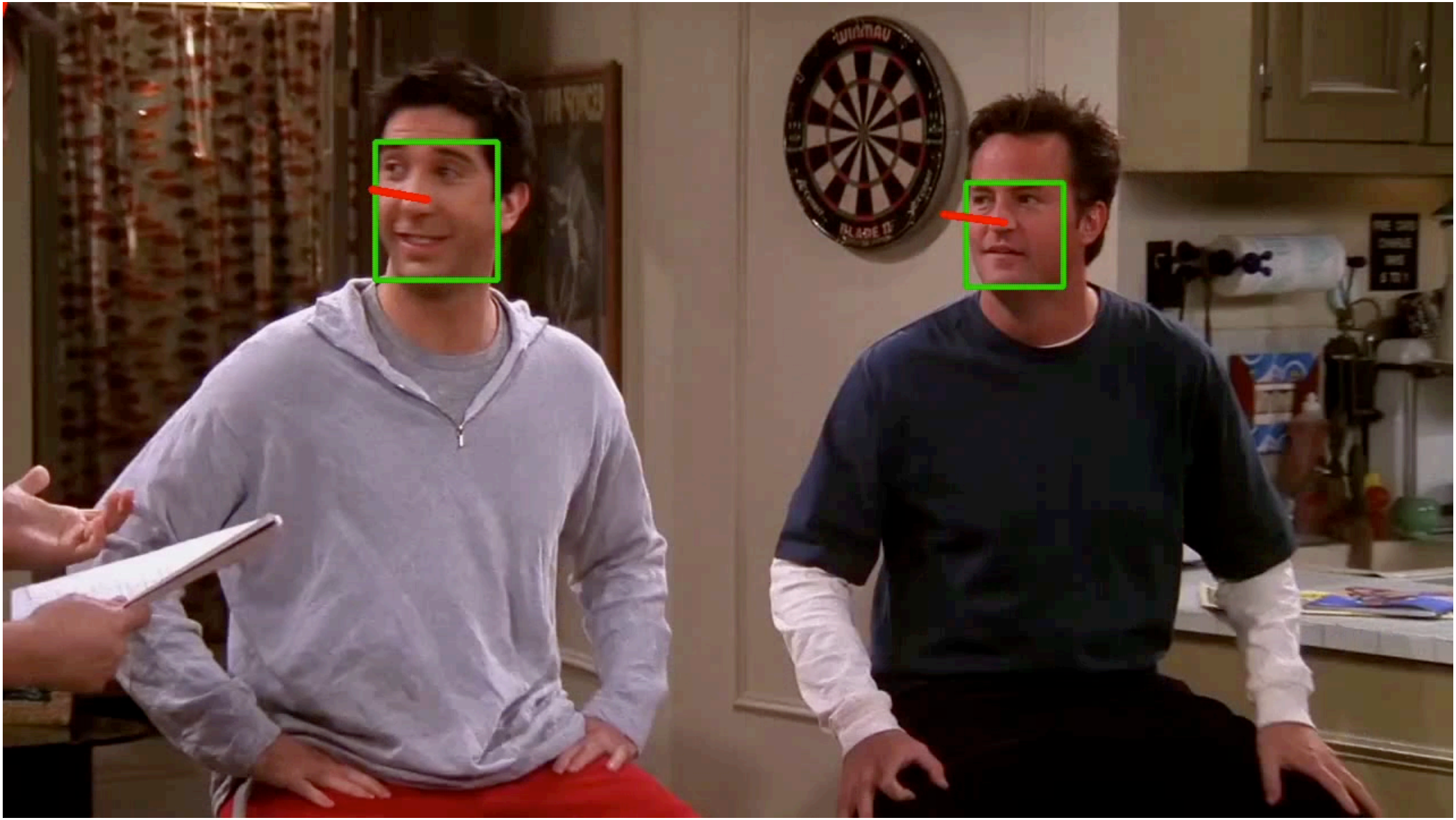
Face and facial landmark detection



J. Li and Y. Zhang, *Proc. CVPR 2013*
T. Baltrušaitis et al., *Proc. ICCVW 2013*

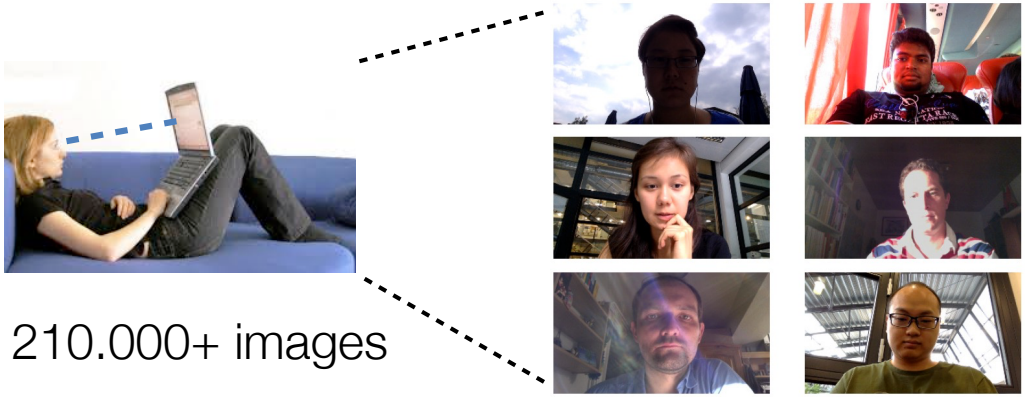
Unsupervised gaze target discovery





Appearance-based gaze estimation

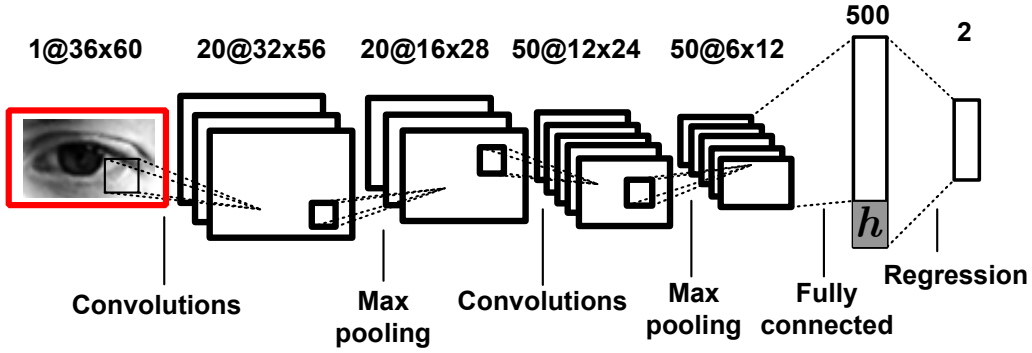
Real-world data (MPIIGaze)



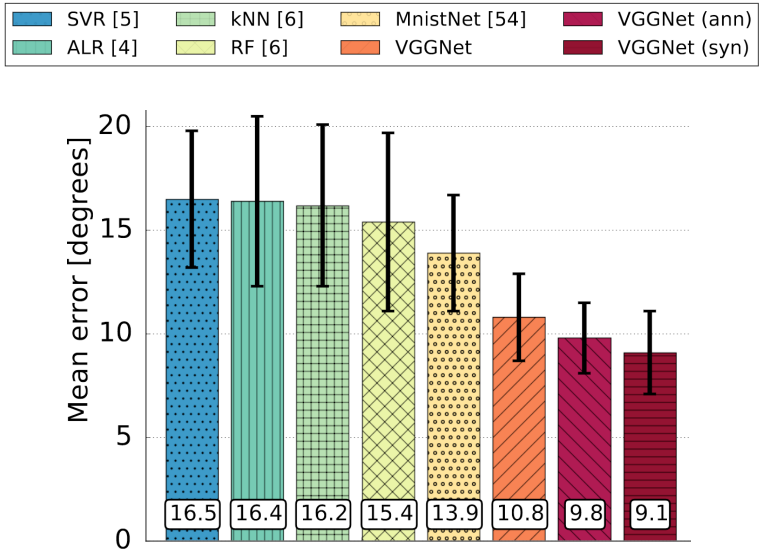
210.000+ images

+

Deep learning



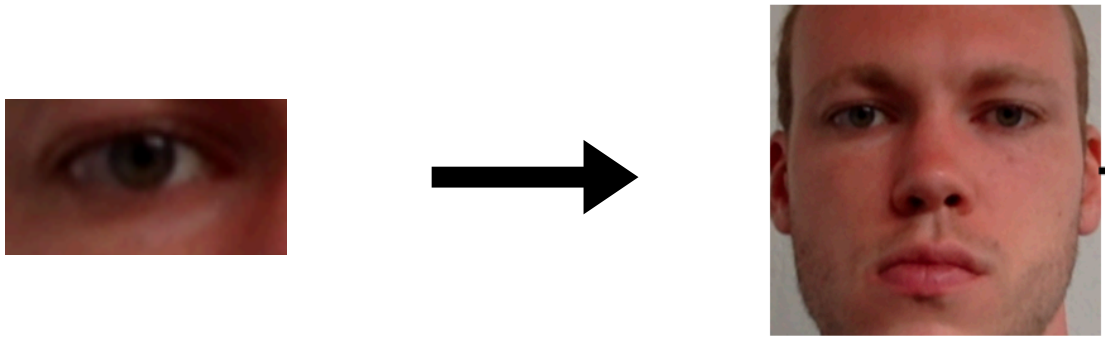
=

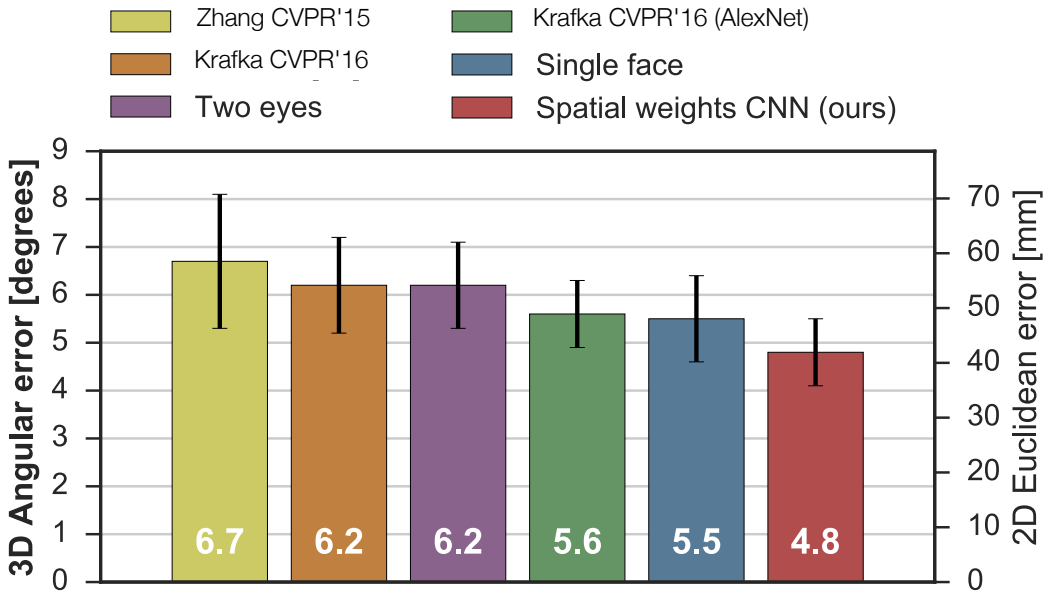


Cross-dataset evaluation: **+41%**

Zhang et al., *Proc. CVPR 2015 / IEEE TPAMI 2018*

Full-face appearance-based gaze estimation



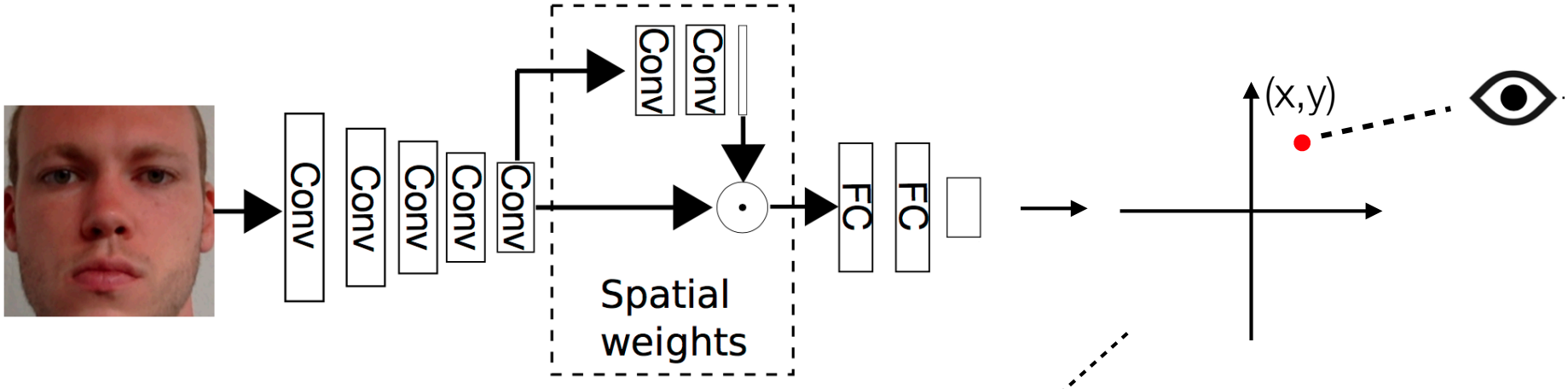


Within-dataset evaluation (MPIIGaze):

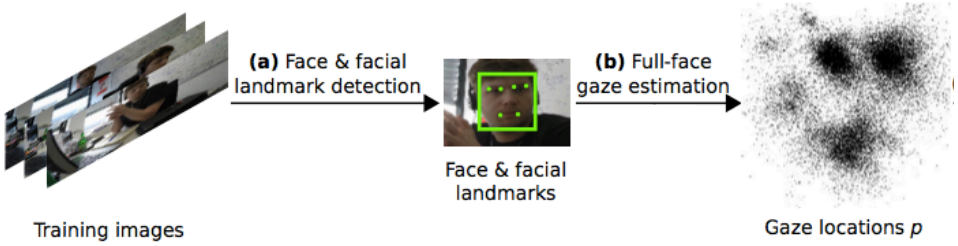
↑14.3%

Zhang et al., It's Written All Over Your Face: Full-Face Appearance-Based Gaze Estimation, *Proc. CVPRW 2017*

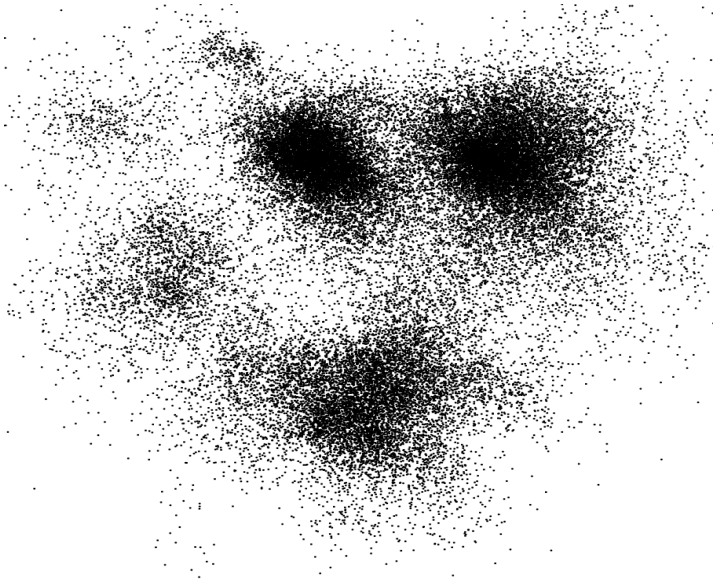
Unsupervised gaze target discovery



Estimated 3D gaze direction projected to camera image plane

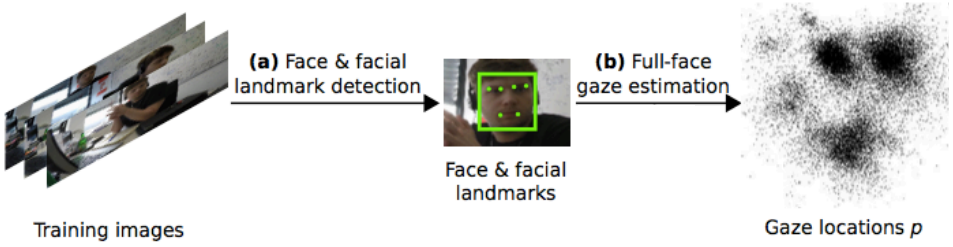


Unsupervised gaze target discovery

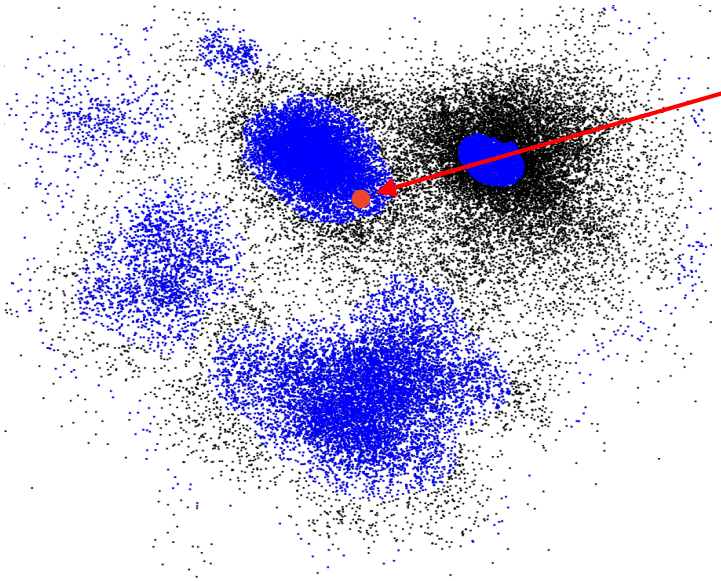


Gaze locations $\boldsymbol{p}$ in the camera image plane

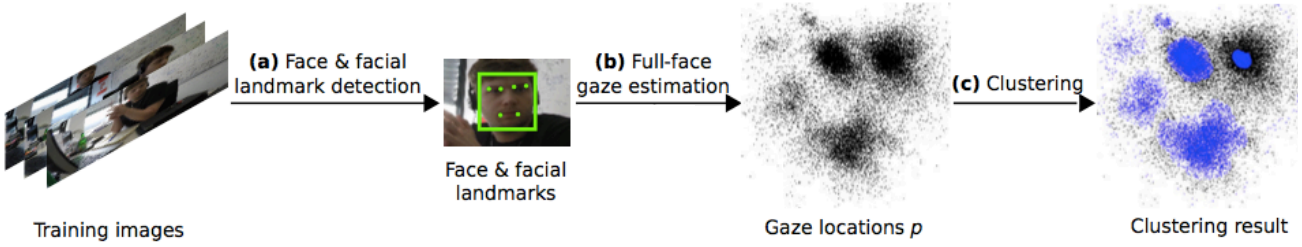
Density-based clustering



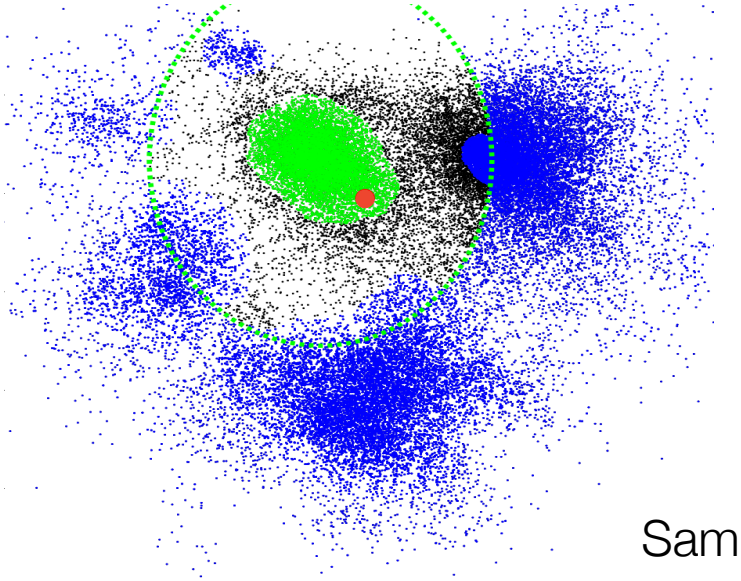
Unsupervised gaze target discovery



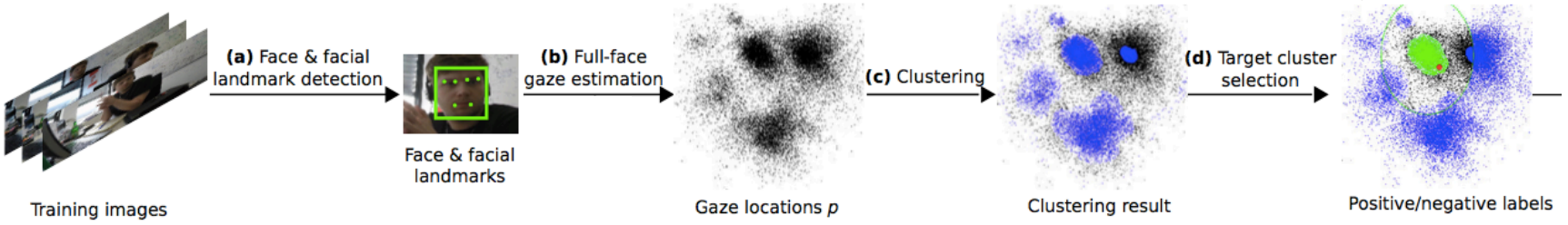
The target is assumed to be a single salient object closest to the camera



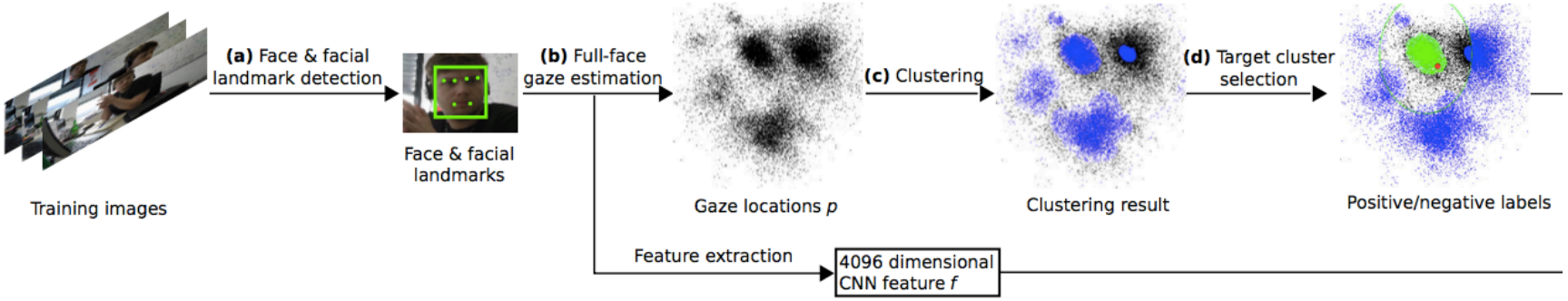
Unsupervised gaze target discovery

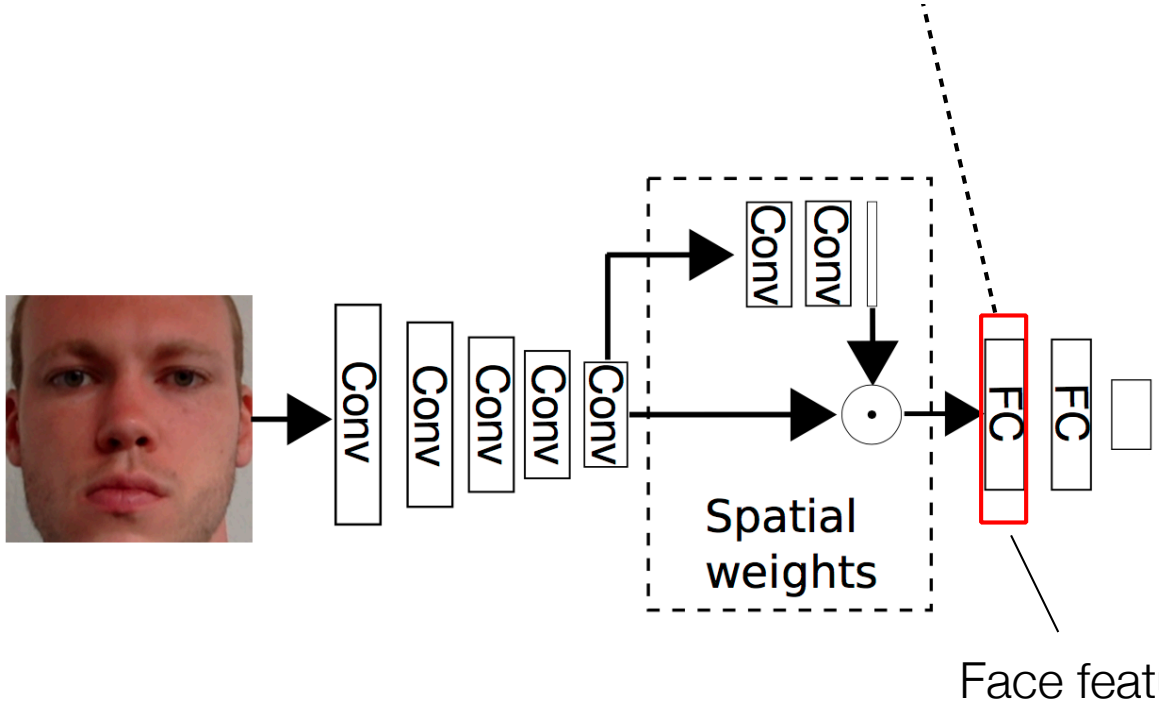


Samples in the cluster closest to the camera are labelled as positive, others as negative



Unsupervised gaze target discovery

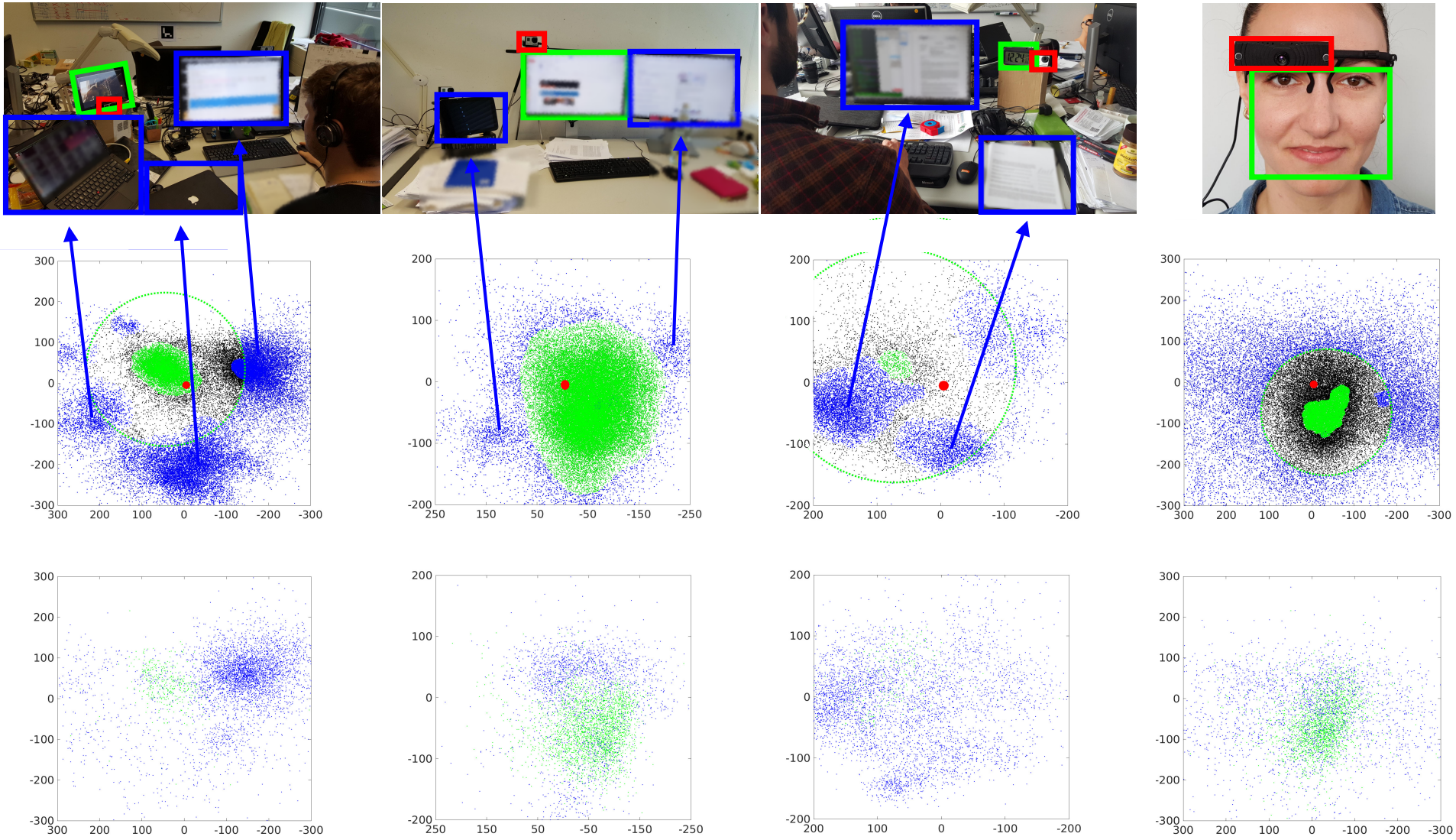




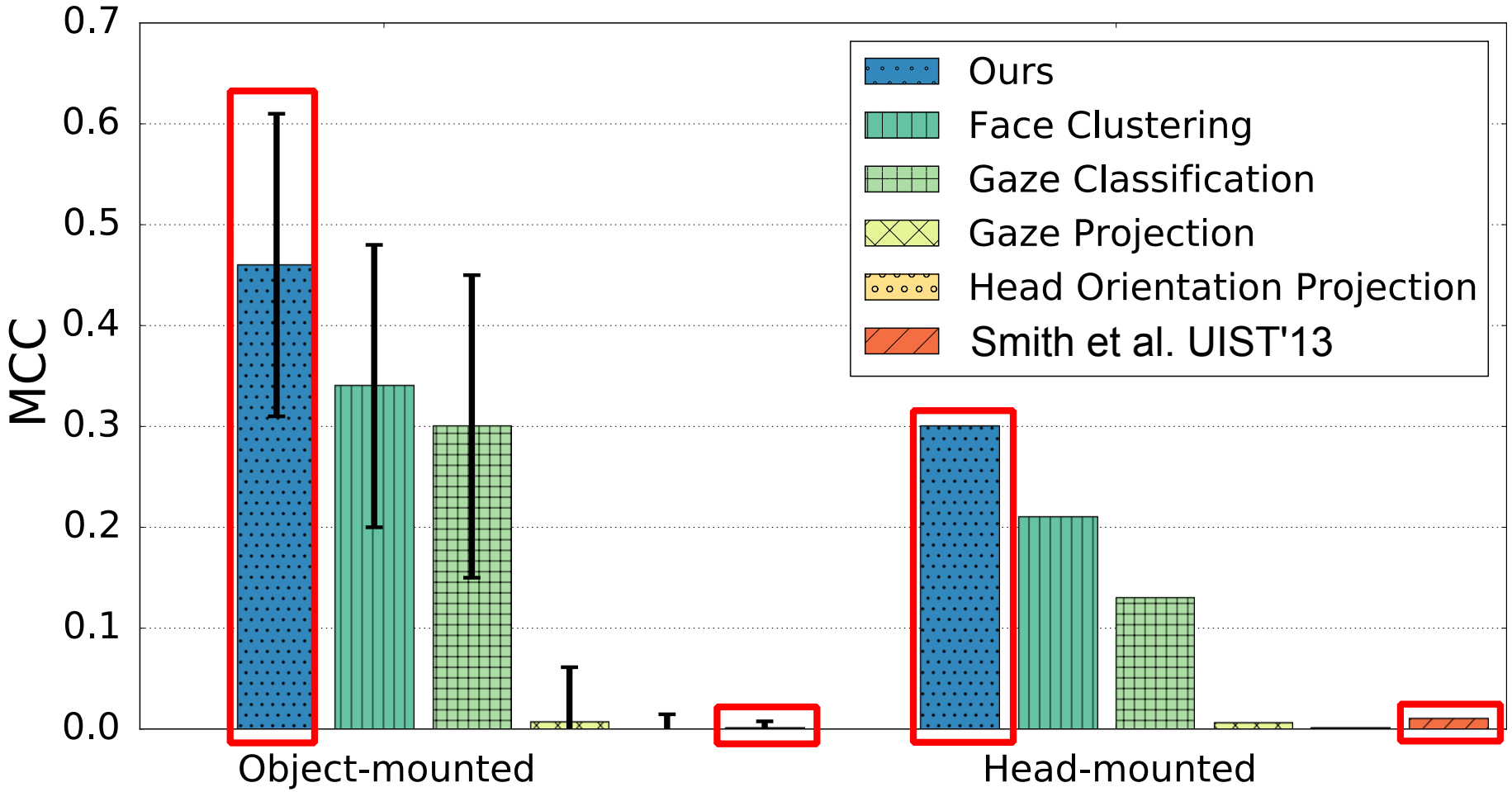
Eye contact detector training using labelled samples and face feature **after deployment**

Sample gaze distributions

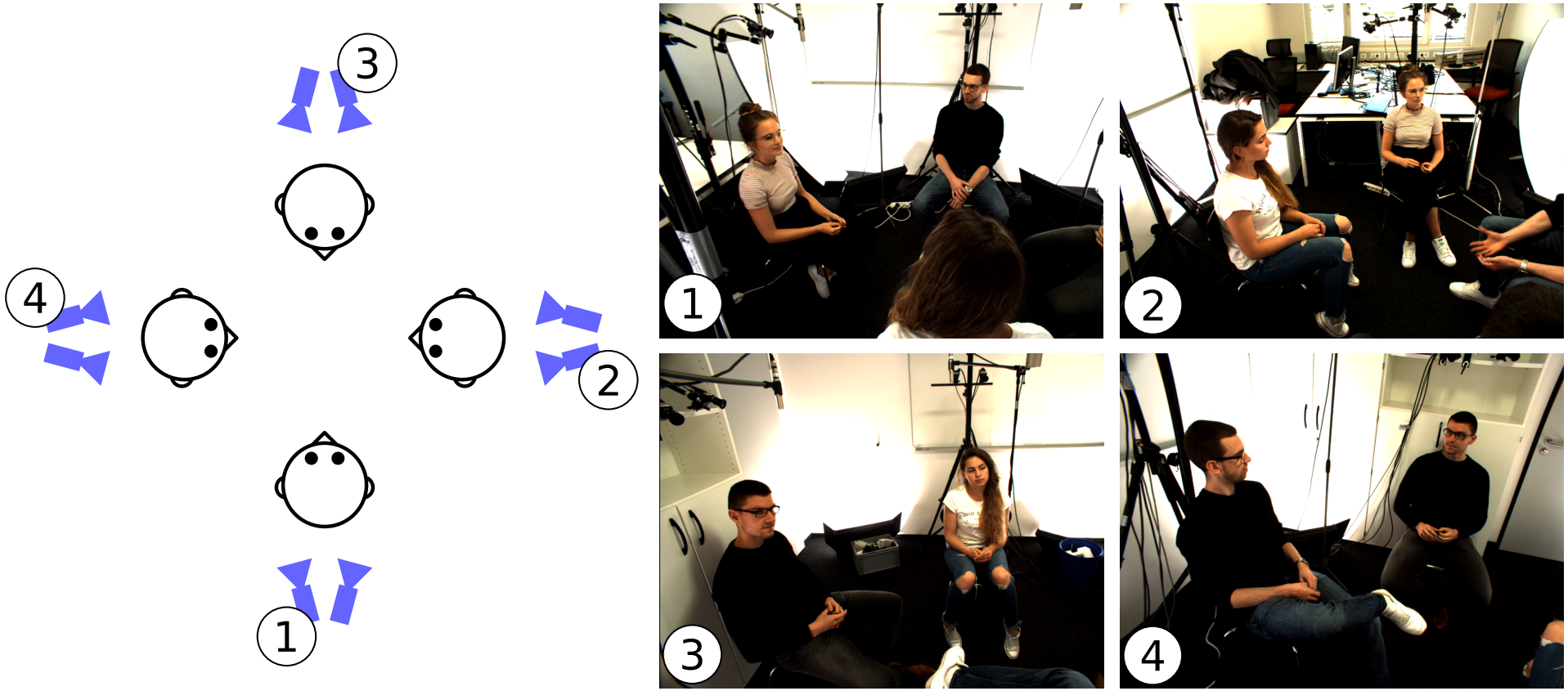
target object camera other distractive objects



Performance evaluation



Multi-target eye contact detection?

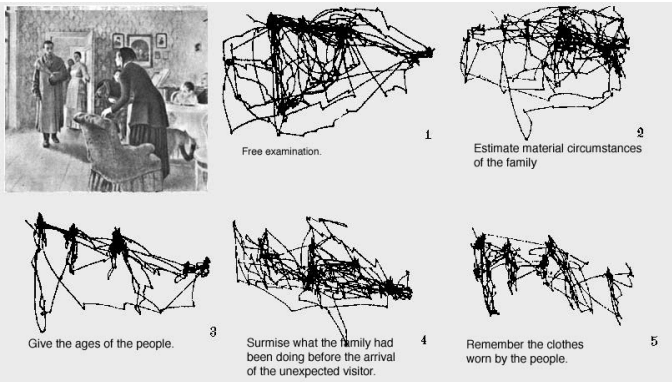


Müller et al., Robust Eye Contact Detection in Natural Multi-Person Interactions Using Gaze and Speaking Behaviour, *Proc. ETRA 2018*


Pervasive eye tracking
Sensing
Analysis
Interaction

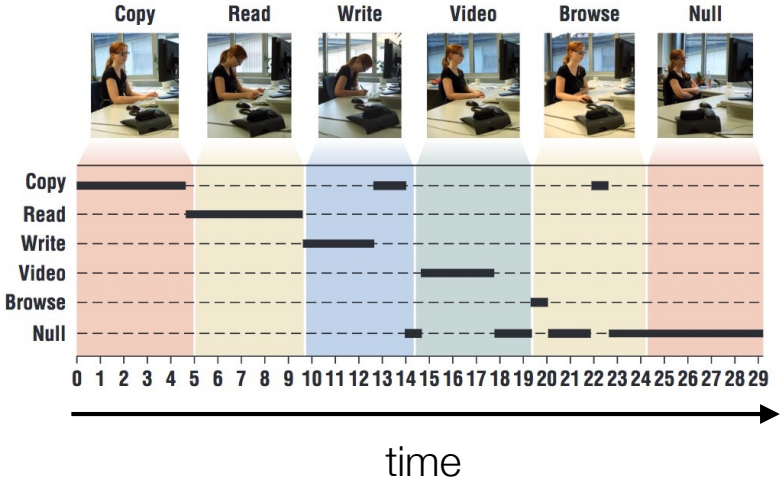
Computational gaze behaviour analysis

Task → Gaze behaviour



Yarbus, Eye Movements and Vision (1967)

Gaze behaviour → User model

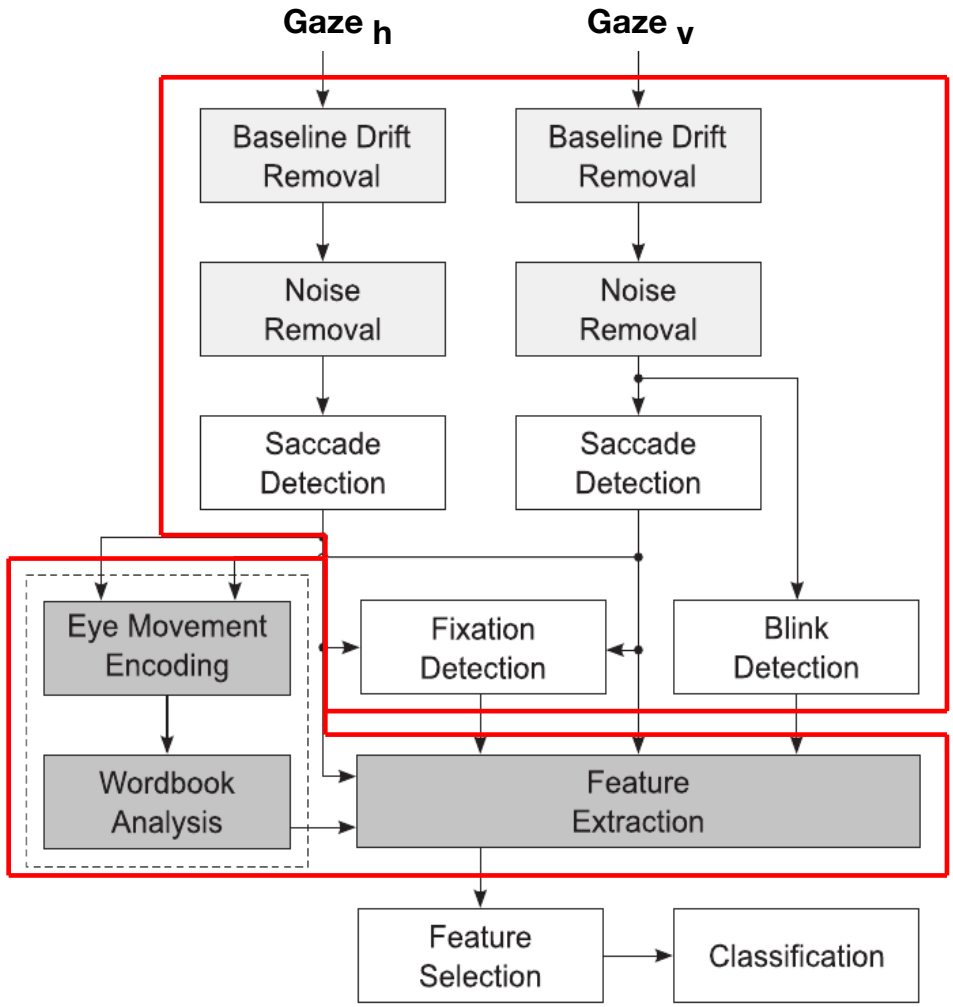


- x Specific eye characteristics
- x Controlled stimuli / tasks
- x Short-term
- x Laboratory settings
- x Gaze-only

- ✓ Full gaze behaviour
- ✓ Everyday activities
- ✓ Long-term
- ✓ Daily life
- ✓ Multimodal

Bulling et al., *IEEE Pervasive Computing 2011*

Computational gaze behaviour analysis



Rich eye movement features

- Saccades
  Mean/variance of saccade amplitude in different directions, saccade rates, ...
- Fixations
  Fixation rate, mean/variance of fixation duration, ...
- Blinks
  Blink rate, mean/variance of blink duration, ...
- Eye movement sequences

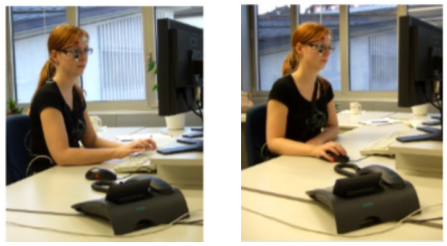
Bulling et al., *Proc. ACM UbiComp 2009 / IEEE TPAMI 2011*

Computational gaze behaviour analysis

Recognition of reading activity in transit

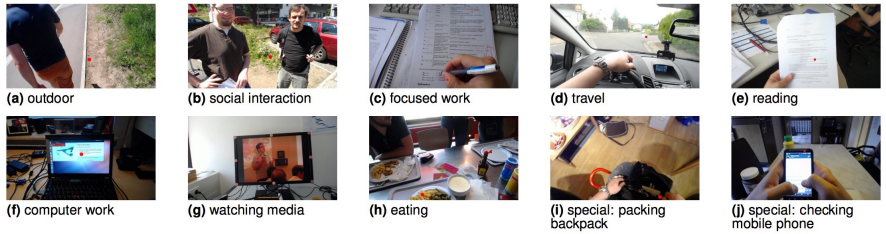
Recognition of document types

Recognition of office activities

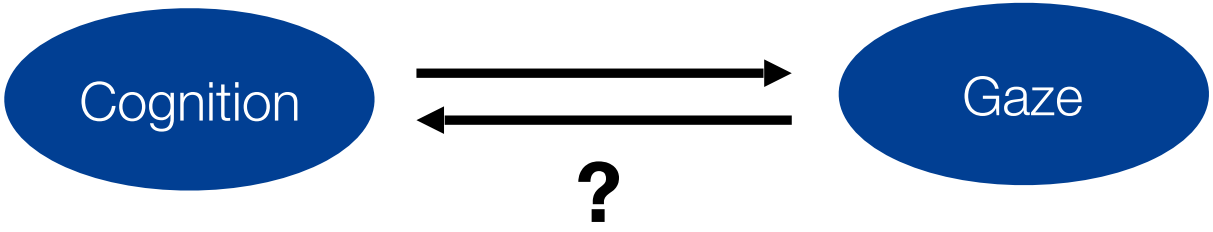


Recognition of high-level user context

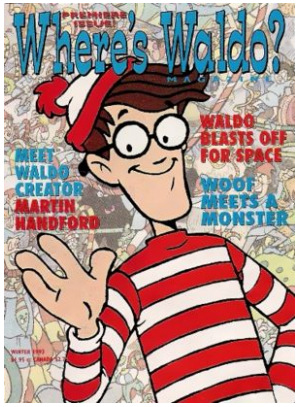
Long-term activity discovery



The eyes: A window into the mind



- Attention
- Cognitive load
- Visual memory / visual search
- Learning / experience

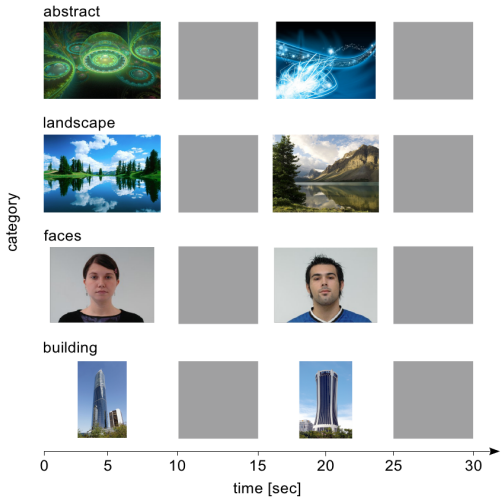


Cognition-aware computing
Computing systems that sense and adapt to users' cognitive states/processes

Bulling, PhD Thesis, ETH Zurich 2010 / Bulling and Zander, *IEEE Pervasive Computing* 2014

Cognition-aware computing

Visual memory recall

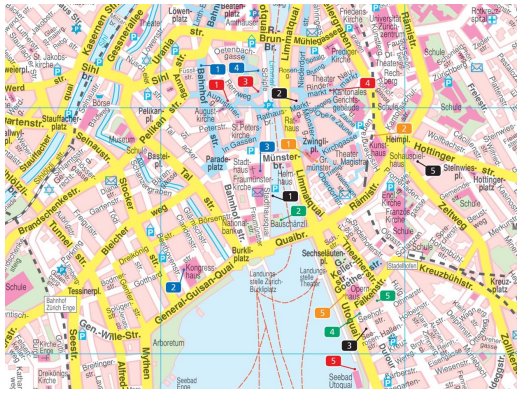


Cognitive load

Personality traits

Big Five

Visual search

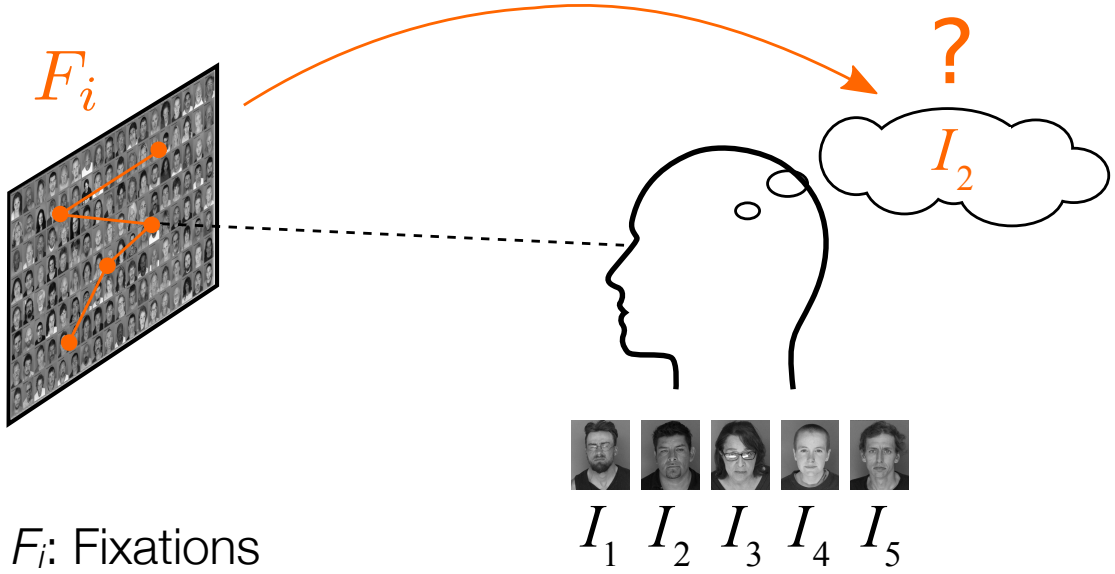


- Pervasive activity
  - Significant potential for developing assistive search interfaces
- Observation: Particular search intent results in specific gaze behaviour

Can we predict users' search intents from their gaze behaviour?

Prediction of search intents

- Prior works focused on the closed-world setting



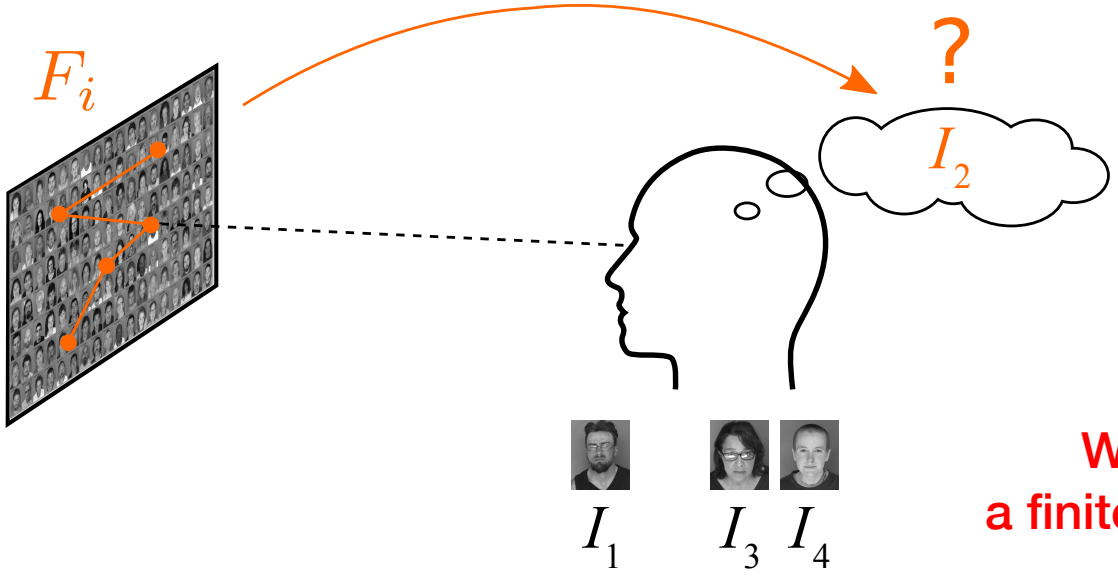
$$F \mapsto I \in \mathbf{I}$$

F_i: Fixations
I_i: Search intents

Zelinsky et al., *JoV 2013* / Borji et al., *Neurocomputing 2015*

Prediction of search intents

- Closed-world prediction is severely limited
  - All potential search intents and corresponding fixations are required at training time
- We instead study <u>open-world</u> prediction



We still predict over a finite set of search intents!

Sattar et al., Prediction of Search Targets From Fixations in Open-World Settings, *Proc. CVPR 2015*

Prediction of search intents

- We require a learning mechanism that can predict over a set of intents that are unknown at training time
  - Standard supervised learning not possible any more
- Idea: Learn compatibilities between fixations and search intents that (hopefully) generalise to other fixations and intents

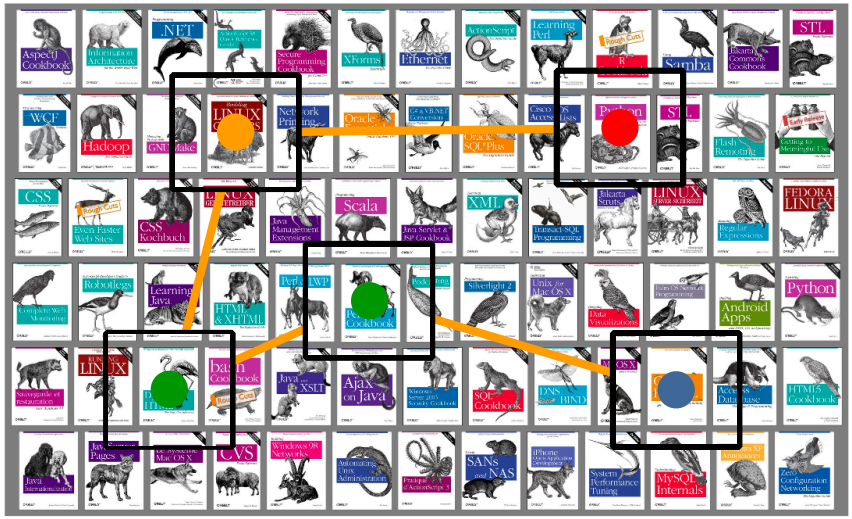
$$F \mapsto I \in \mathrm{I} \quad \Rightarrow \quad (F, I) \mapsto Y \in \{0,1\}$$

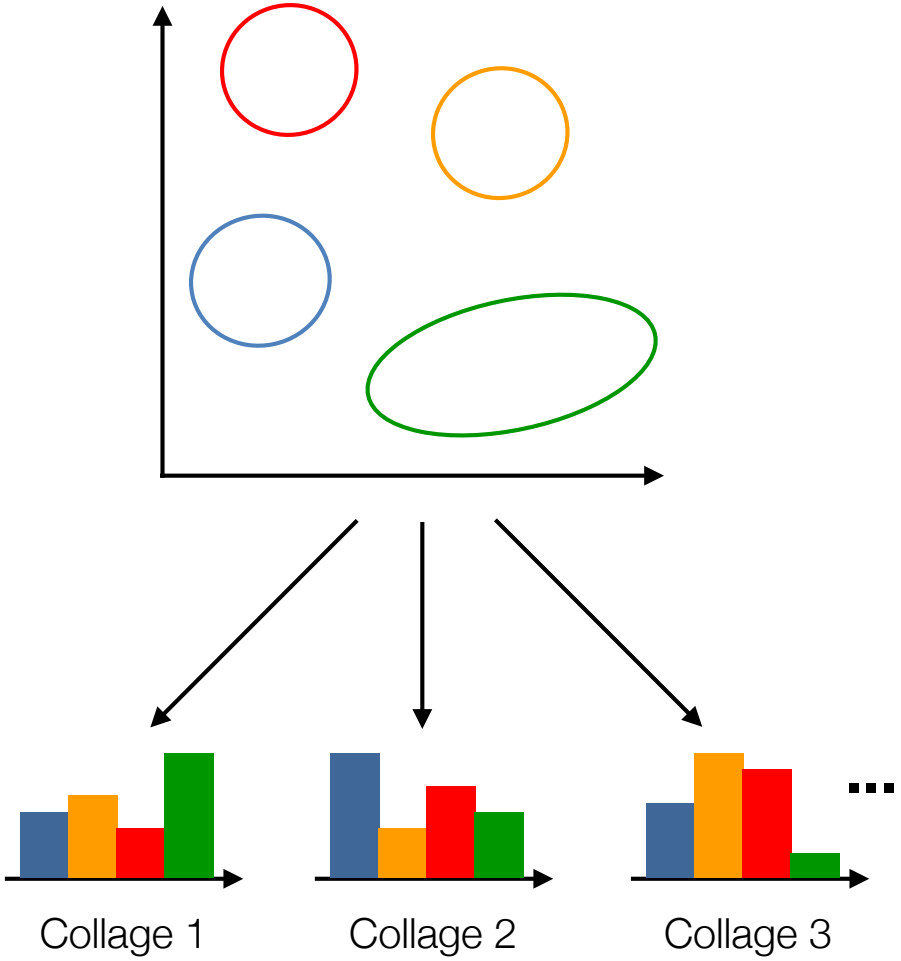
closed-world → open-world

- We propose a featurisation of fixations using visual (image) information

$$(\phi(F), I) \mapsto Y \in \{0,1\}$$

Bag of visual words





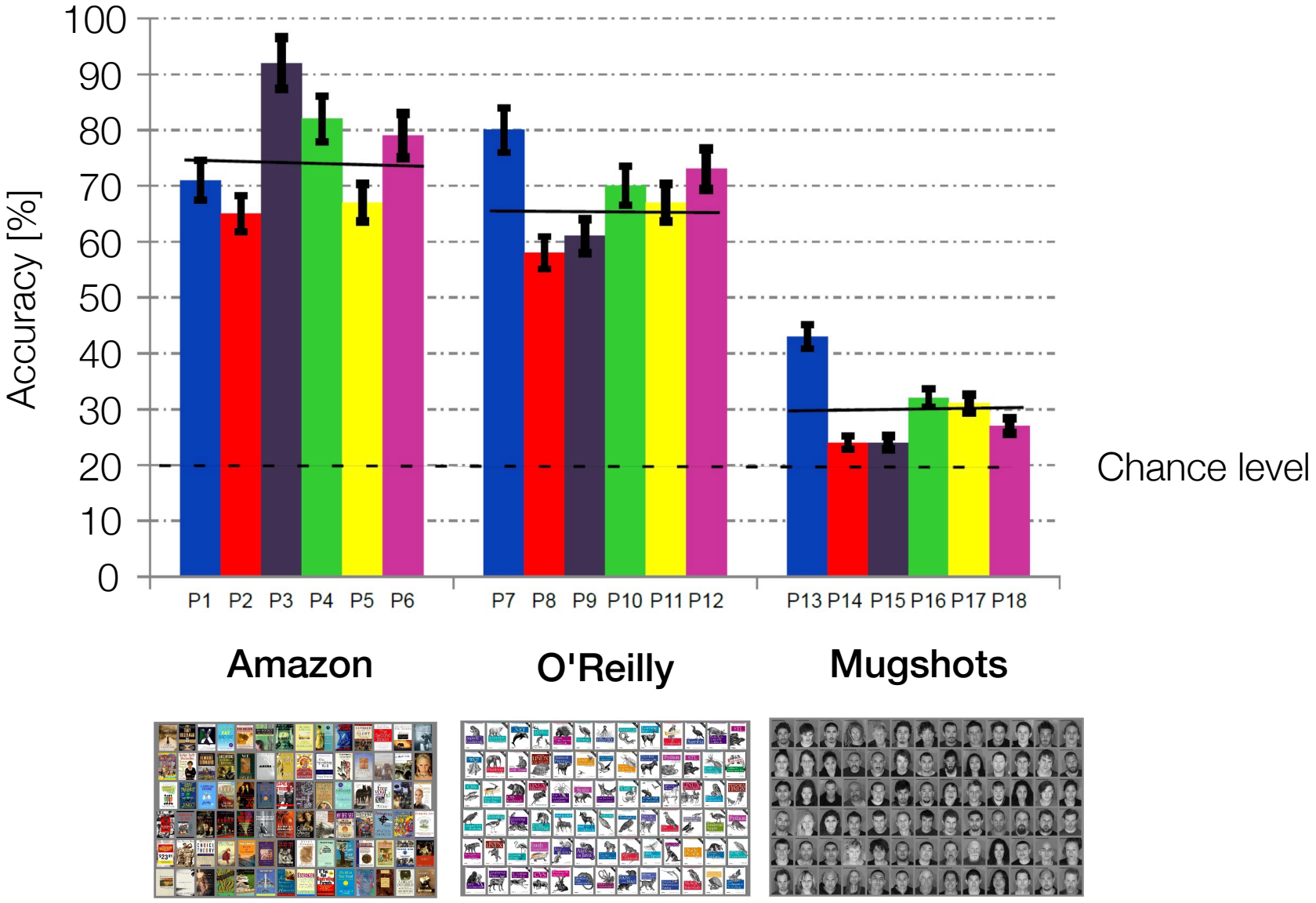
Fixations on collage as key points

Patch extraction

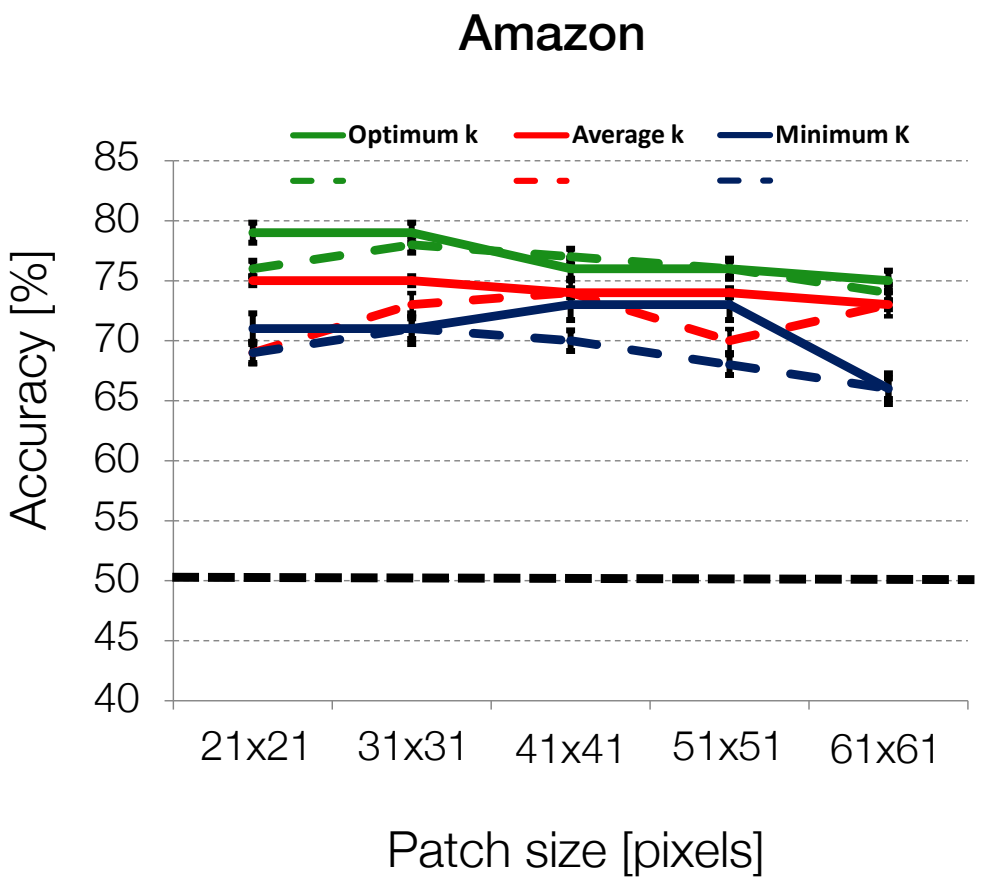
Clustering

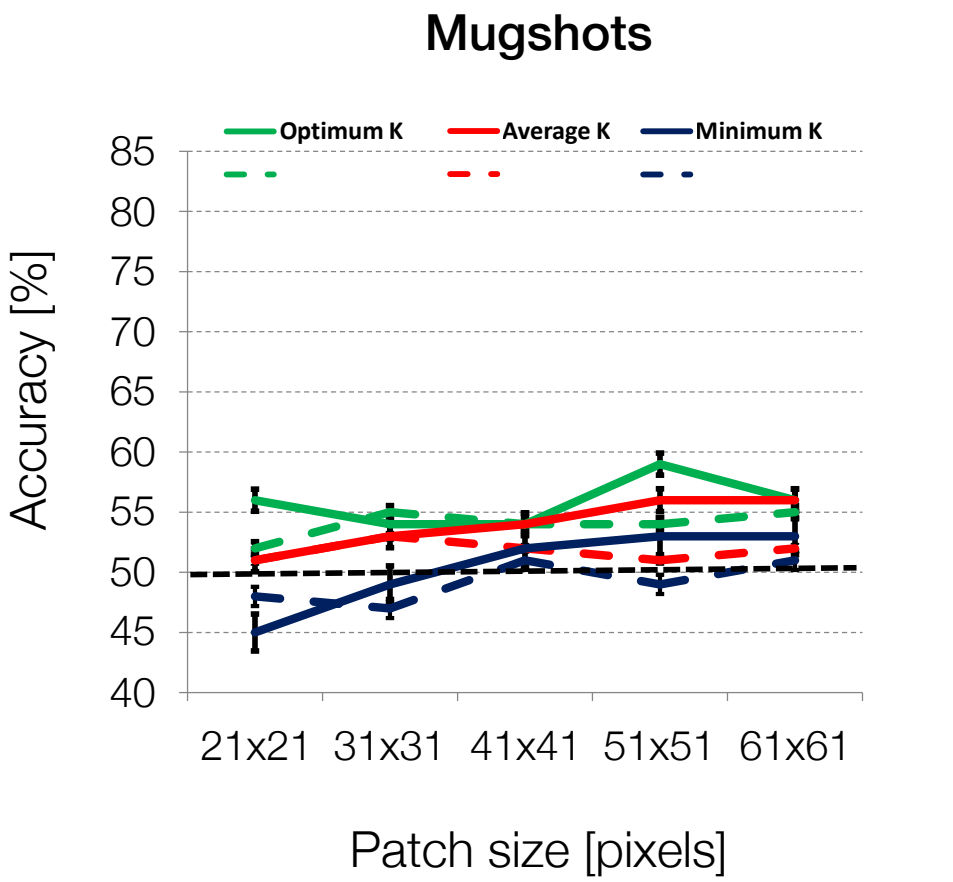
Visual word vectors

Results: Closed-world



Results: Open-world





Motivation

- Previous works assumed a concrete search intent
- In practice, search intents are often abstract, e.g. "*red dress*"

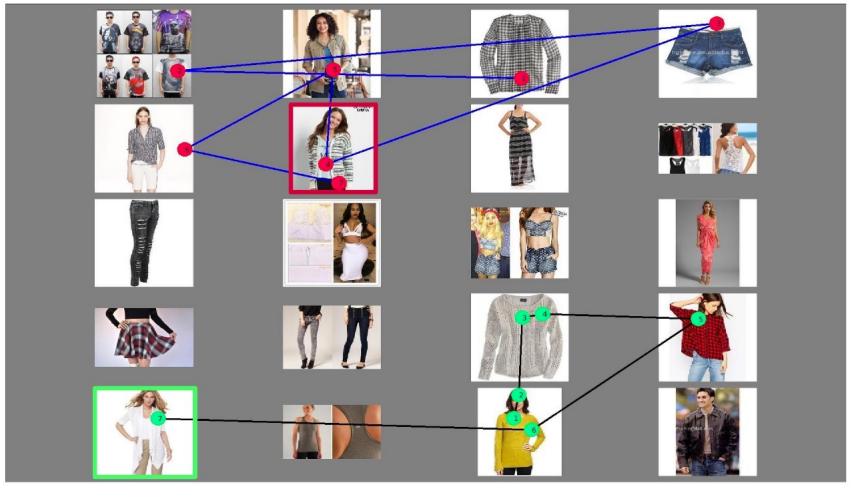
User 1

User 2

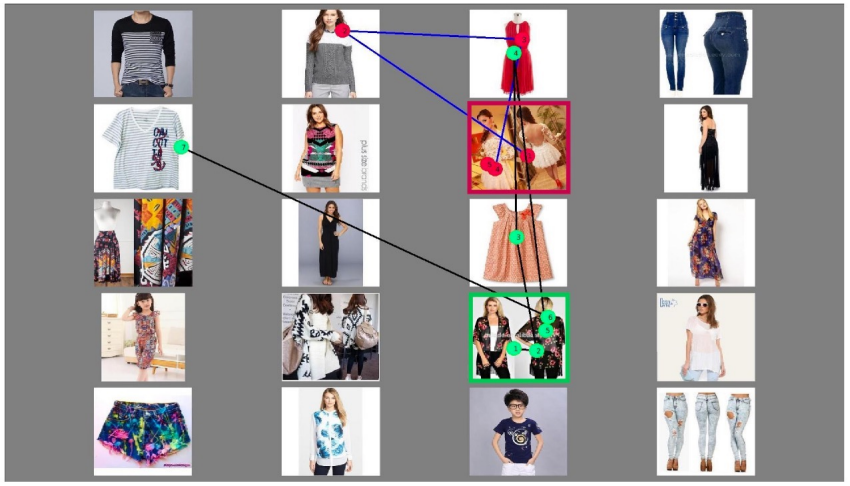
User 3

Mental image search

- Highly challenging because the same mental image can result in drastically different fixation behaviour



User 1
User 2



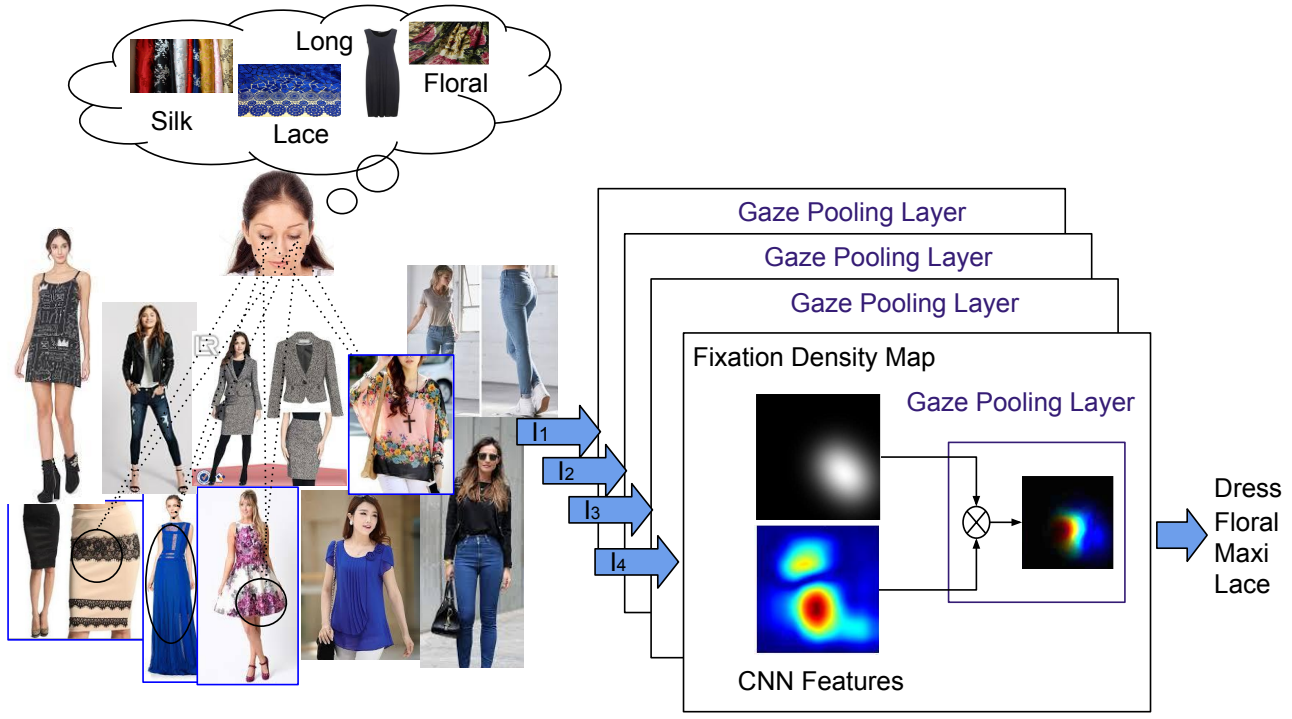
Cardigan

Floral dress

Can we predict the category or attributes of mental images?

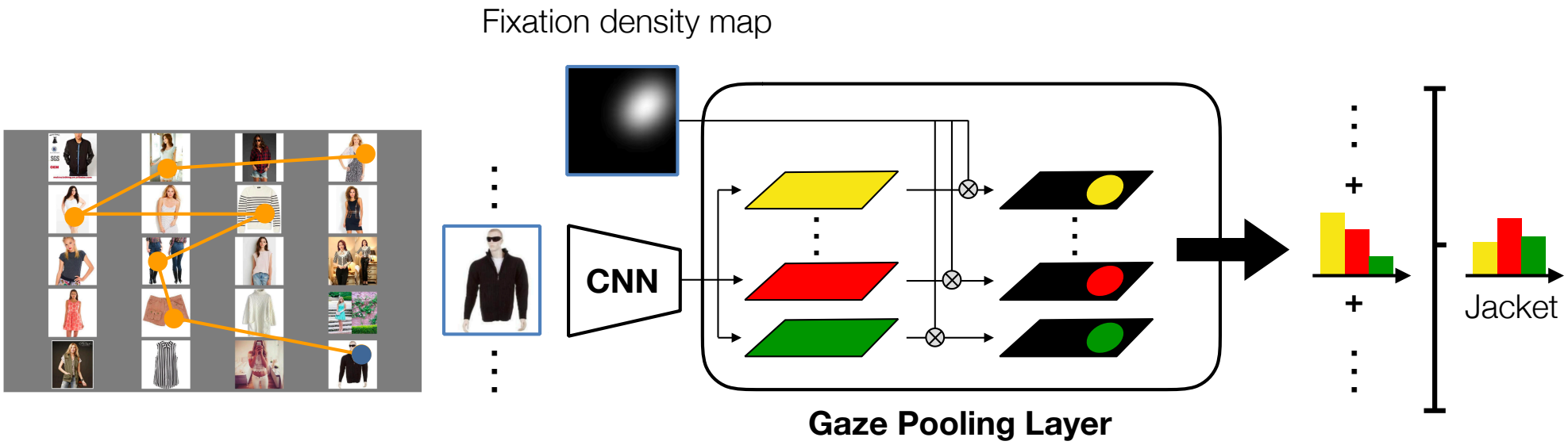
Prediction of mental image search intents

- Collecting large training corpora is prohibitive for human gaze data
- Idea: Use fixations only as an attention mechanism to selectively weight a pre-trained deep image representation → Gaze Pooling Layer



Sattar et al., Predicting the Category and Attributes of Visual Search Targets Using Deep Gaze Pooling, *Proc. ICCWV 2017*

Gaze Pooling Layer



- Each fixated image is encoded using a pre-trained CNN [Liu CVPR'16]
- Image representation and fixation density maps are combined using spatial re-weighting, followed by global average pooling [Zhou CVPR'16]
- Final prediction by integrating class posteriors across all fixated images

Quantitative results

Global vs. Local		Category Top1	Category Top2	Category Top3	Attribute Accuracy
Global		31%±5	48% ±8	62% ±8	20%±1
Local		49%±7	68%±6	78%±6	26 %±1
Global	✓	52%±6	68%±6	78%±6	25%±1
Local	✓	**57%±8**	**74%±7**	**84%±4**	**34%±1**

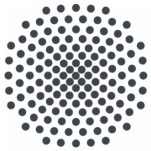
- Global: uniform weight across the whole fixated image
- Local: localised weights for each fixated image
- Weighting of average class posteriors using fixation duration

Take home messages

1. Exciting **new applications impossible before**
2. **Recent advances** in pervasive gaze sensing, analysis, and interaction
3. **Data-driven methods** are instrumental
4. **A lot still remains to be done** to fully realise the vision of pervasive eye tracking

For my new research group

I am looking for highly motivated and skilled



University of Stuttgart
Germany

PhD students and **PostDocs** with a **background in**

i) **Machine learning**, e.g. deep learning, generative models, (inverse) reinforcement learning

ii) **Computer vision or graphics**, e.g. gaze estimation, egocentric vision, scene understanding, object detection/recognition

SimTech
Cluster of Excellence

TRR 161
Transregional Collaborative Research Center
Quantitative Methods for Visual Computing

Strong interest in applying these methods to HCI
e.g. intelligent user interfaces

CyberValley

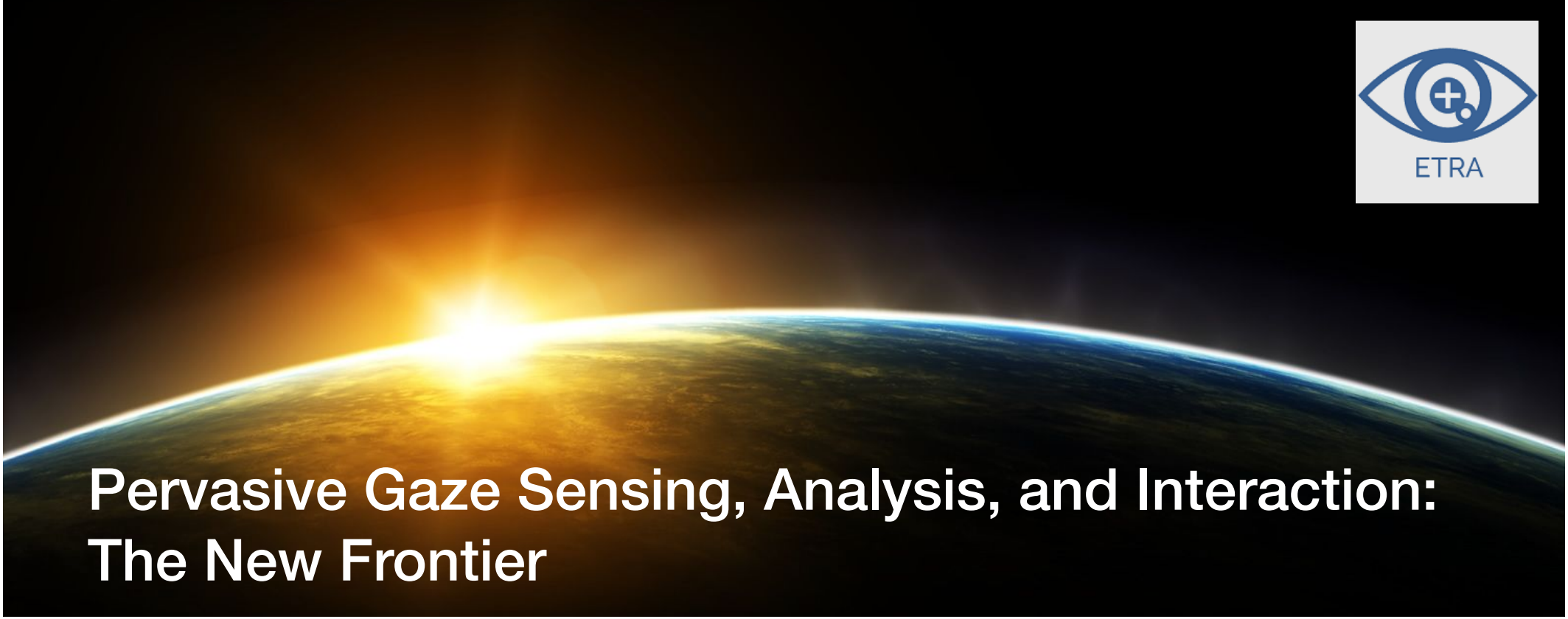
Excellent programming skills in C++ or similar languages are expected; experience with Python, MATLAB, or CUDA

BOSCH AI

Mercedes-Benz

Fluent English written and presentation skills

If you consider yourself to belong to the top 10% of your peer group I'd love to talk to you!



Pervasive Gaze Sensing, Analysis, and Interaction: The New Frontier

Andreas Bulling

Max Planck Institute for Informatics & Saarland University

bulling@mpi-inf.mpg.de

Additional references (slide 36)

Computational gaze behaviour analysis

Julian Steil; Andreas Bulling. Discovery of Everyday Human Activities From Long-Term Visual Behaviour Using Topic Models. Proc. of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp 2015), pp. 75-85, 2015.

Kai Kunze; Andreas Bulling; Yuzuko Utsumi; Shiga Yuki; Koichi Kise. I know what you are reading -- Recognition of document types using mobile eye tracking. Proc. of the 17th International Symposium on Wearable Computers (ISWC 2013), pp. 113-116

Andreas Bulling; Christian Weichel; Hans Gellersen. EyeContext: Recognition of High-level Contextual Cues from Human Visual Behaviour. Proc. of the 31st SIGCHI International Conference on Human Factors in Computing Systems (CHI 2013), pp. 305-308

Andreas Bulling; Jamie A. Ward; Hans Gellersen; Gerhard Tröster. Eye Movement Analysis for Activity Recognition Using Electrooculography. IEEE Transactions on Pattern Analysis and Machine Intelligence, 33 (4), pp. 741-753, 2011

Andreas Bulling; Daniel Roggen; Gerhard Tröster. What's in the Eyes for Context-Awareness? IEEE Pervasive Computing, 10 (2), pp. 48 - 57, 2011.

Andreas Bulling; Jamie A. Ward; Hans Gellersen; Gerhard Tröster. Robust Recognition of Reading Activity in Transit Using Wearable Electrooculography. Proc. of the 6th International Conference on Pervasive Computing (Pervasive 2008), pp. 19–37, 2008

Additional references (slide 38)

Cognition-Aware Computing

Sabrina Hoppe; Tobias Loetscher; Stephanie Morey; Andreas Bulling. Eye Movements During Everyday Behavior Predict Personality Traits. Frontiers in Human Neuroscience, 12 , pp. 105:1-105:8, 2018.

Sabrina Hoppe; Tobias Loetscher; Stephanie Morey; Andreas Bulling. Recognition of Curiosity Using Eye Movement Analysis. Adj. Proc. of the ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp 2015), pp. 185-188, 2015.

Bernd Tessendorf; Andreas Bulling; Daniel Roggen; Thomas Stiefmeier; Manuela Feilner; Peter Derleth; Gerhard Tröster. Recognition of Hearing Needs From Body and Eye Movements to Improve Hearing Instruments Proc. of the 9th International Conference on Pervasive Computing, pp. 314-331, Springer, 2011.

Andreas Bulling; Daniel Roggen. Recognition of Visual Memory Recall Processes Using Eye Movement Analysis Proc. of the 13th International Conference on Ubiquitous Computing (UbiComp 2011), pp. 455-464, 2011.