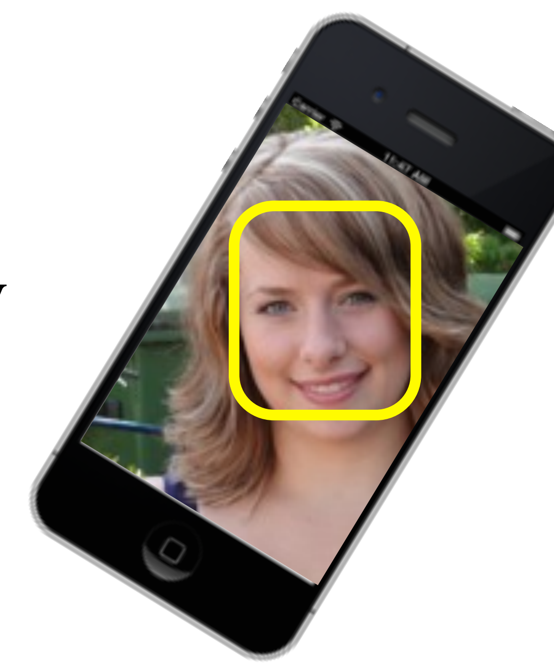
Ego-Motion Compensated Face Detection on a Mobile Device

Björn Scheuermann, Arne Ehlers, Hamon Riazy, Florian Baumann,
Bodo Rosenhahn

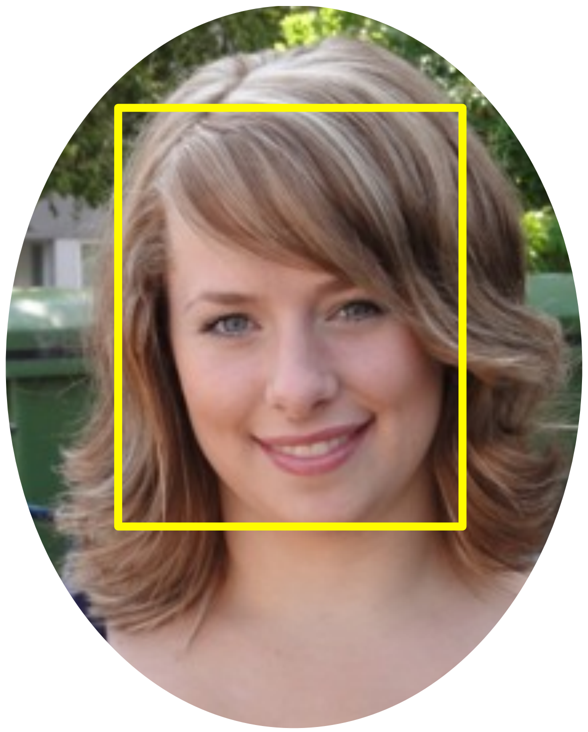
Institut für Informationsverarbeitung
Leibniz Universität Hannover, Germany



Outline

- Motivation
- Boosting Based Face Detection
- Inertial Measurement Unit
- Proposed Method
- Experimental Results
- Conclusion

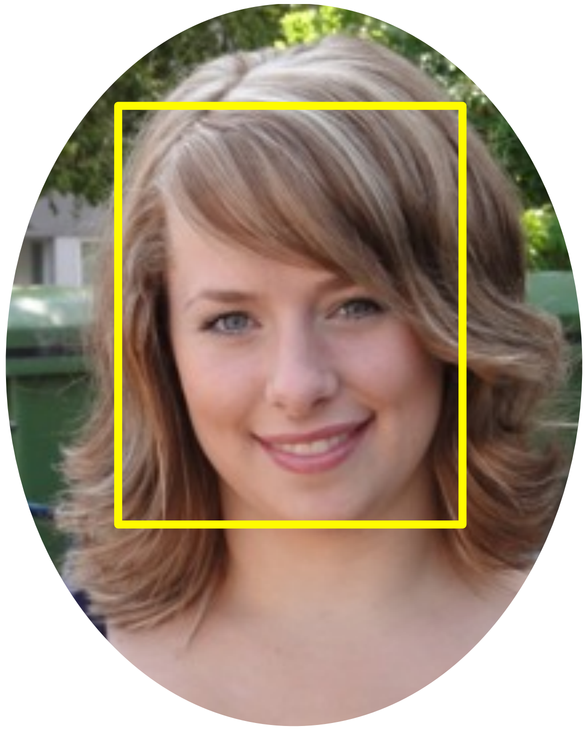
Viola and Jones Face Detection



P.Viola and M.Jones. Rapid object detection using a boosted cascade of simple features. CVPR, 2001

Viola and Jones Face Detection

- well known appearance-based approach
- utilizes Adaboost in the classifier training
  - weak classifiers based on Haar-like features
- integral image for efficient feature computation
- standard in current commercial cameras

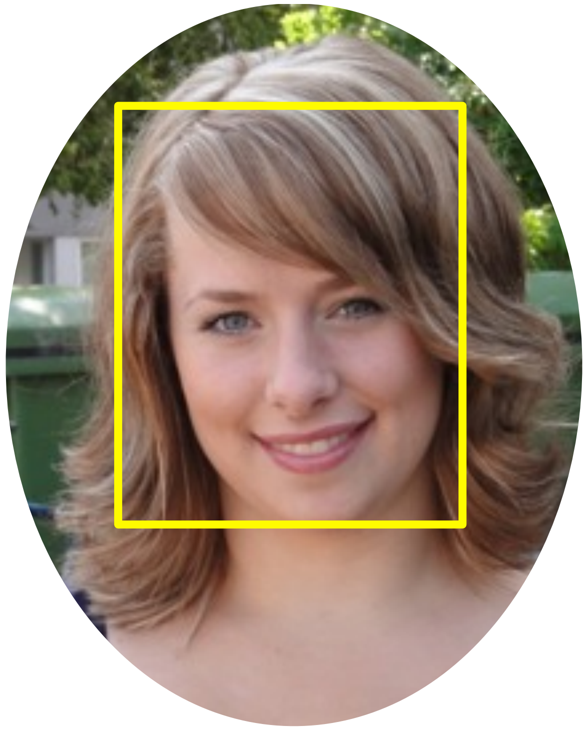


P.Viola and M.Jones. Rapid object detection using a boosted cascade of simple features. CVPR, 2001

Viola and Jones Face Detection

- well known appearance-based approach
- utilizes Adaboost in the classifier training
  - weak classifiers based on Haar-like features
- integral image for efficient feature computation
- standard in current commercial cameras

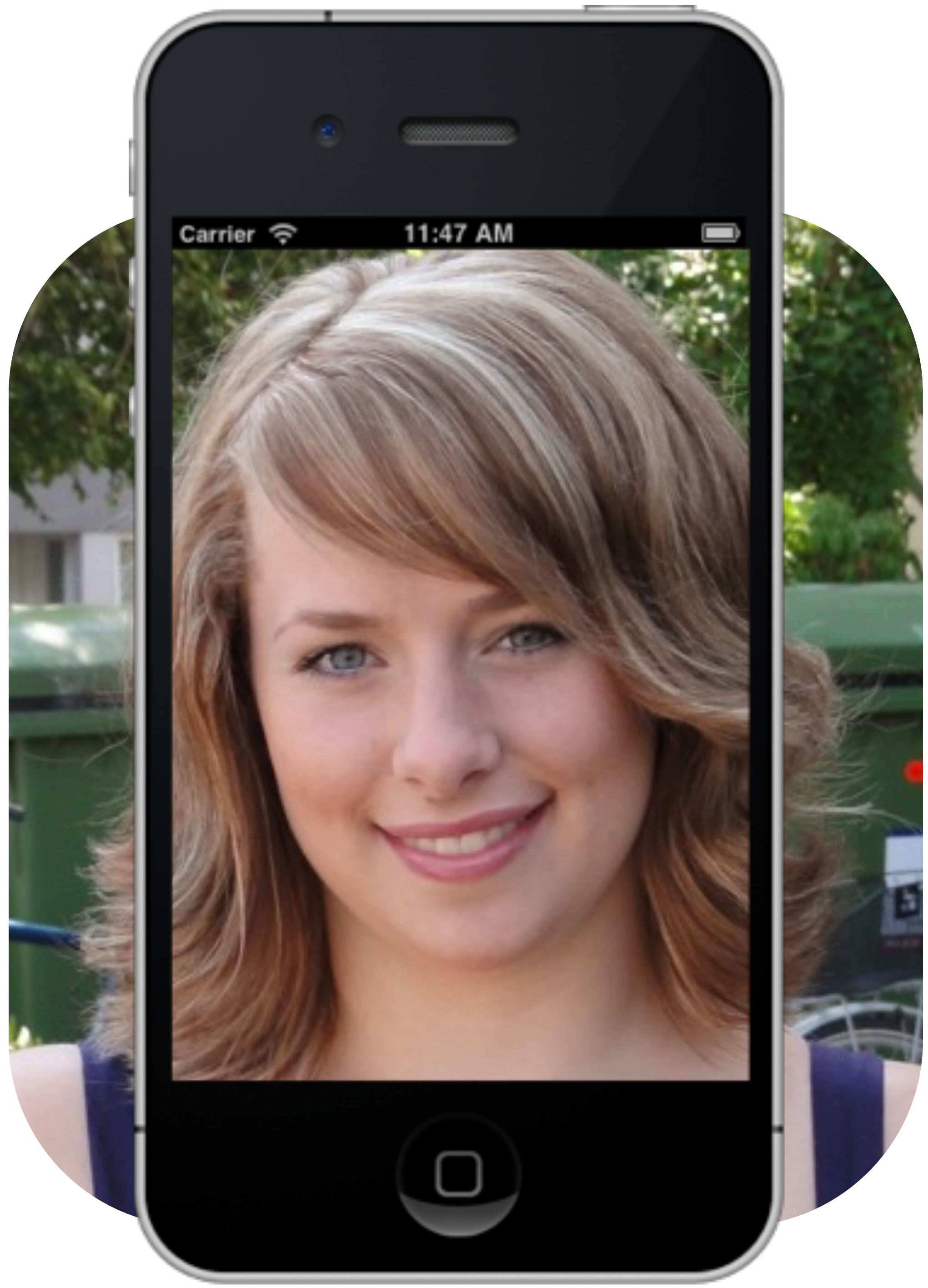
- pros:
  - general object detection framework
  - good detection rates
  - fast detection
- drawbacks:
  - slow training
  - highly rotational variant



P.Viola and M.Jones. Rapid object detection using a boosted cascade of simple features. CVPR, 2001

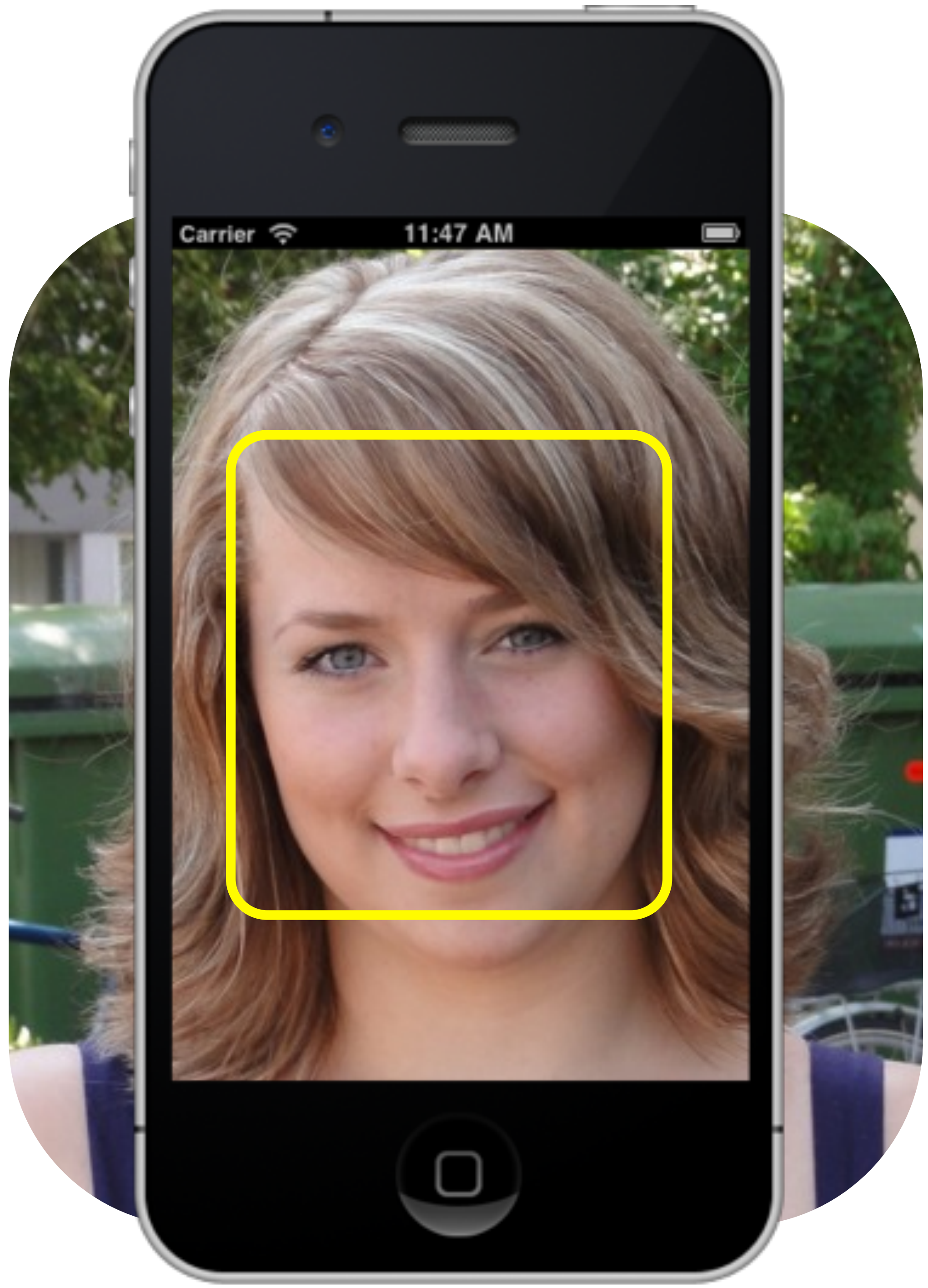
Motivation

Viola and Jones
Face detection:



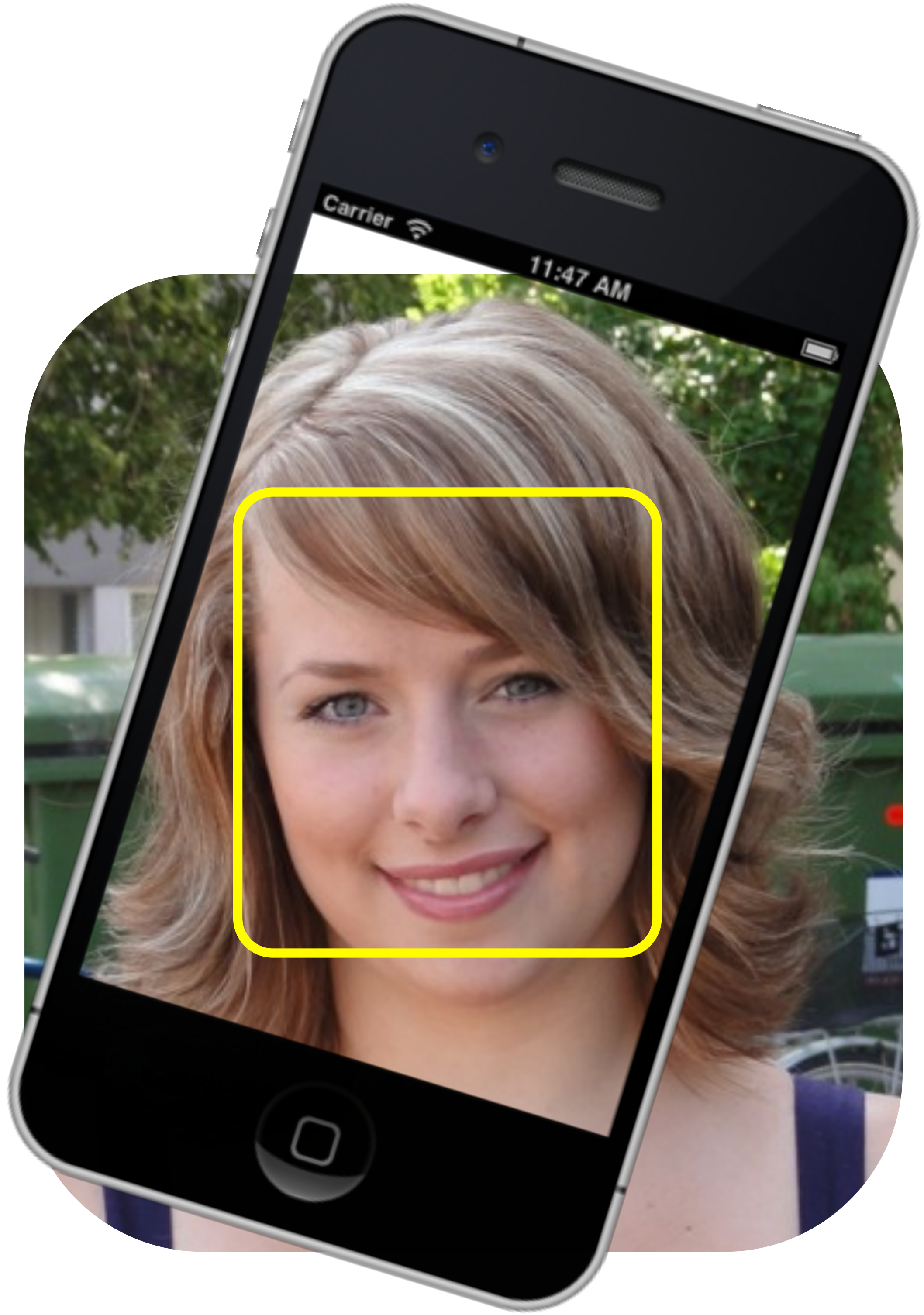
Motivation

Viola and Jones
Face detection:



Motivation

Viola and Jones
Face detection:



Motivation

Viola and Jones
Face detection:

Motivation

Viola and Jones
Face detection:

Motivation

Rotation invariant detection based on Viola and Jones?

- Du et al.: set of rotated Haar like features
- Wu et al.: divide human faces according to different view points

Motivation

Rotation invariant detection based on Viola and Jones?

- Du et al.: set of rotated Haar like features
- Wu et al.: divide human faces according to different view points

Computational too expensive for mobile devices (8-12 times slower)

Motivation

Rotation invariant detection based on Viola and Jones?

- Du et al.: set of rotated Haar like features
- Wu et al.: divide human faces according to different view points

Computational too expensive for mobile devices (8-12 times slower)

Inertial sensors getting major attention

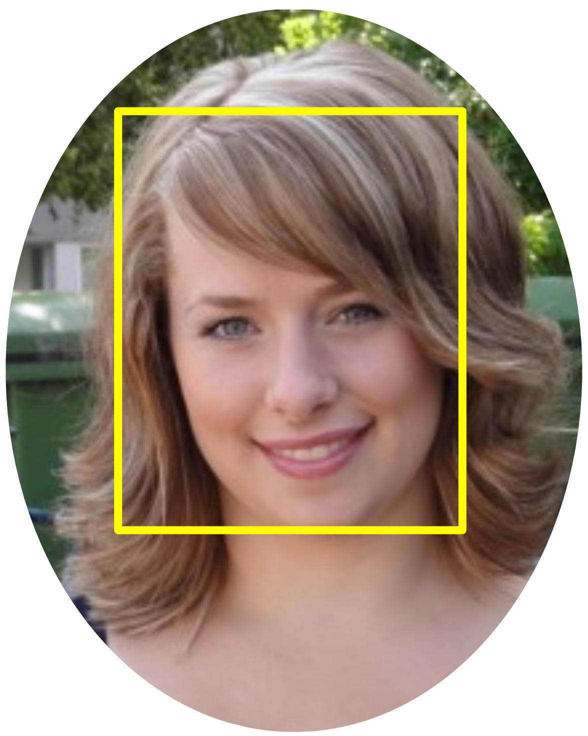
- markerless human motion capturing
- structure from motion
- ...

Inertial sensors are available on most current mobile devices

- mobile phones (iPhone, HTC Desire, ...)
- iPod touch, iPad, ...

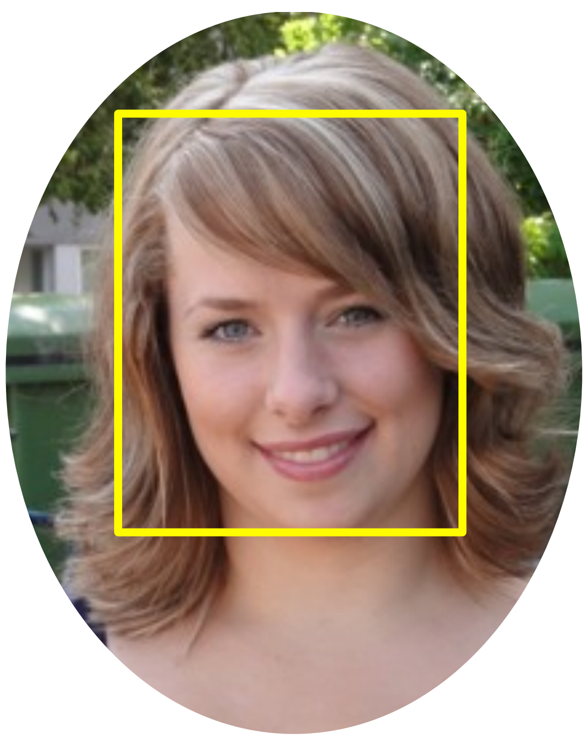
Motivation

Viola and Jones
face detection
framework
(rotational variant)



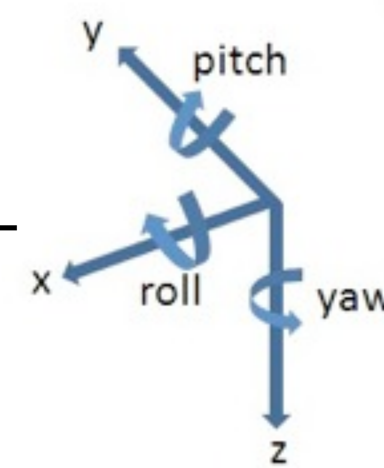
Motivation

Viola and Jones face detection framework (rotational variant)



mobile devices with an IMU (inertial measurement unit)

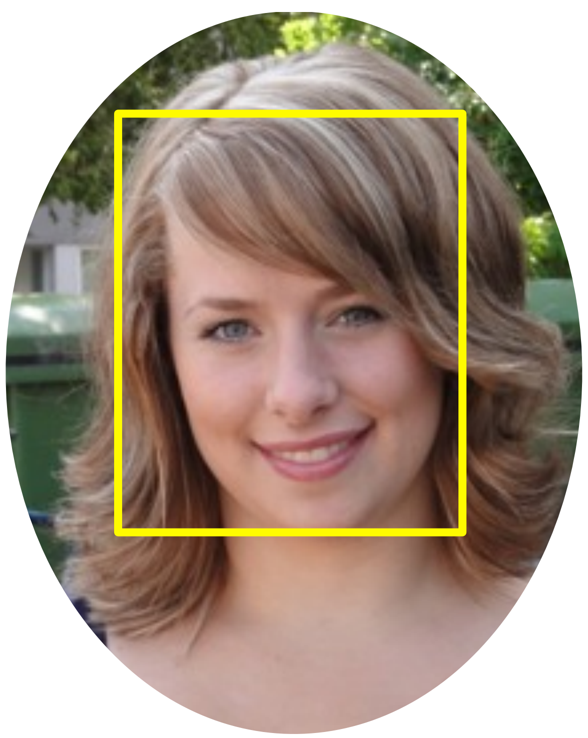
+



Motivation

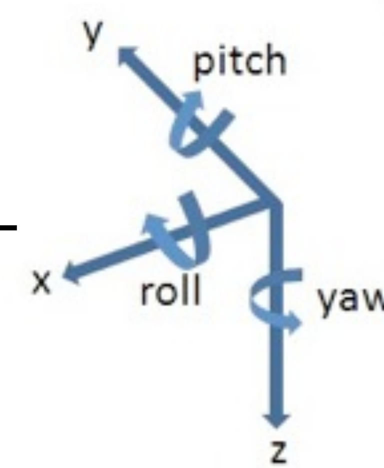
Viola and Jones face detection framework (rotational variant)

mobile devices with an IMU (inertial measurement unit)



fusion?!

+



Boosting Based Face Detection

Adaboost is an algorithm for constructing a „strong“ classifier as a linear combination of simple „weak“ classifiers $h_t(x)$:

$$f(x) = \sum_{t=1}^{T} \alpha_t h_t(x)$$

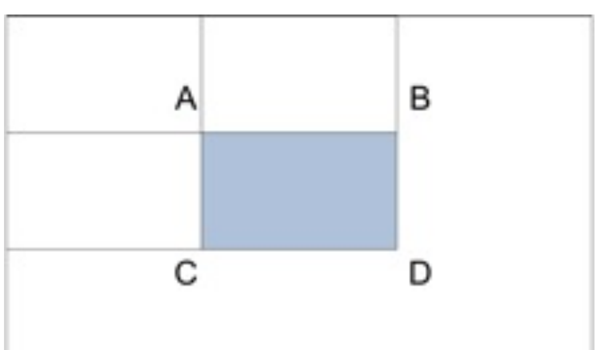
final „strong“ classifier/hypothesis: $H(x) = \text{sign}(f(x))$

Haar-like Features

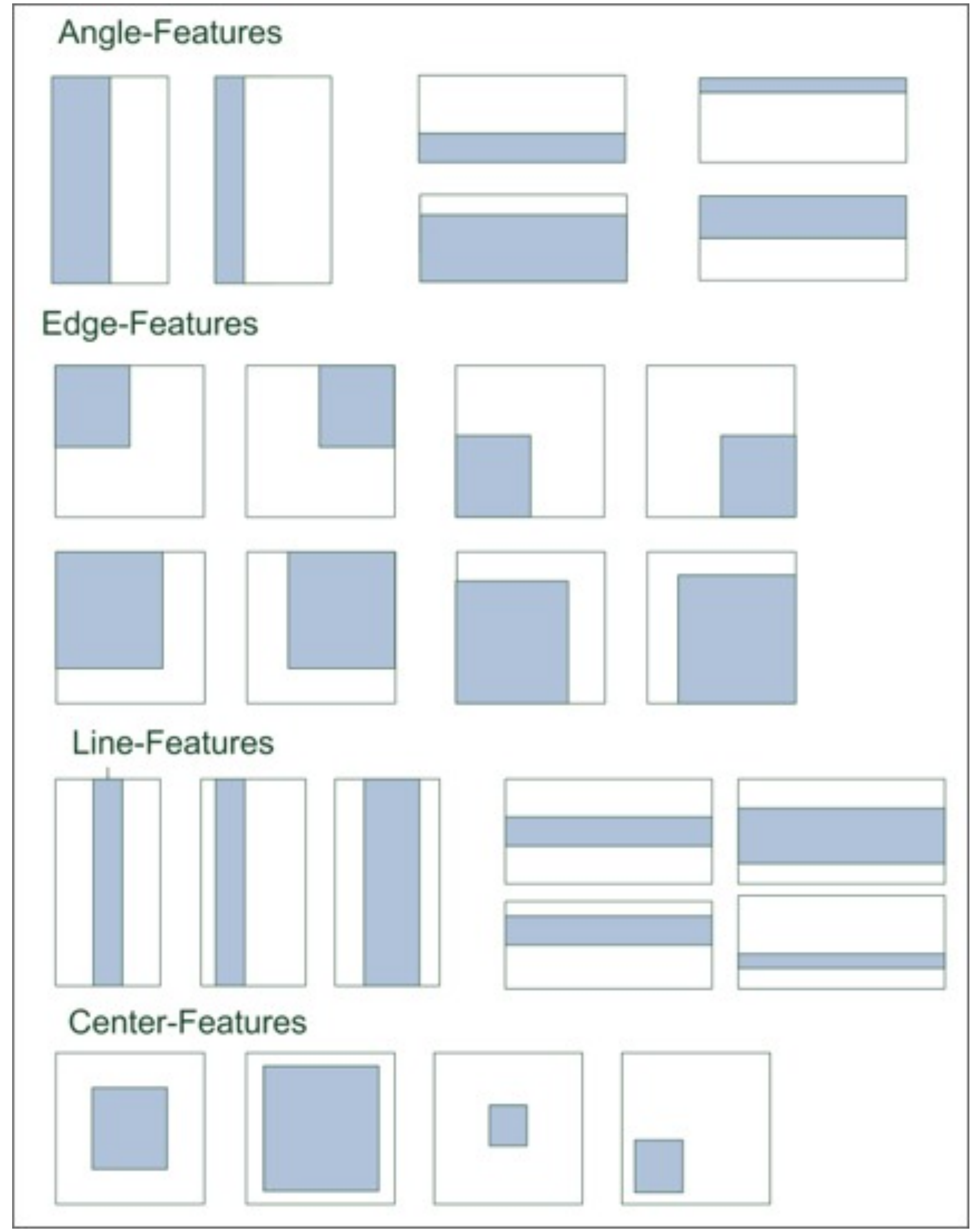
- Easy computation (scalable and changeable in size and position)

- Fast and efficient computation with integral-image

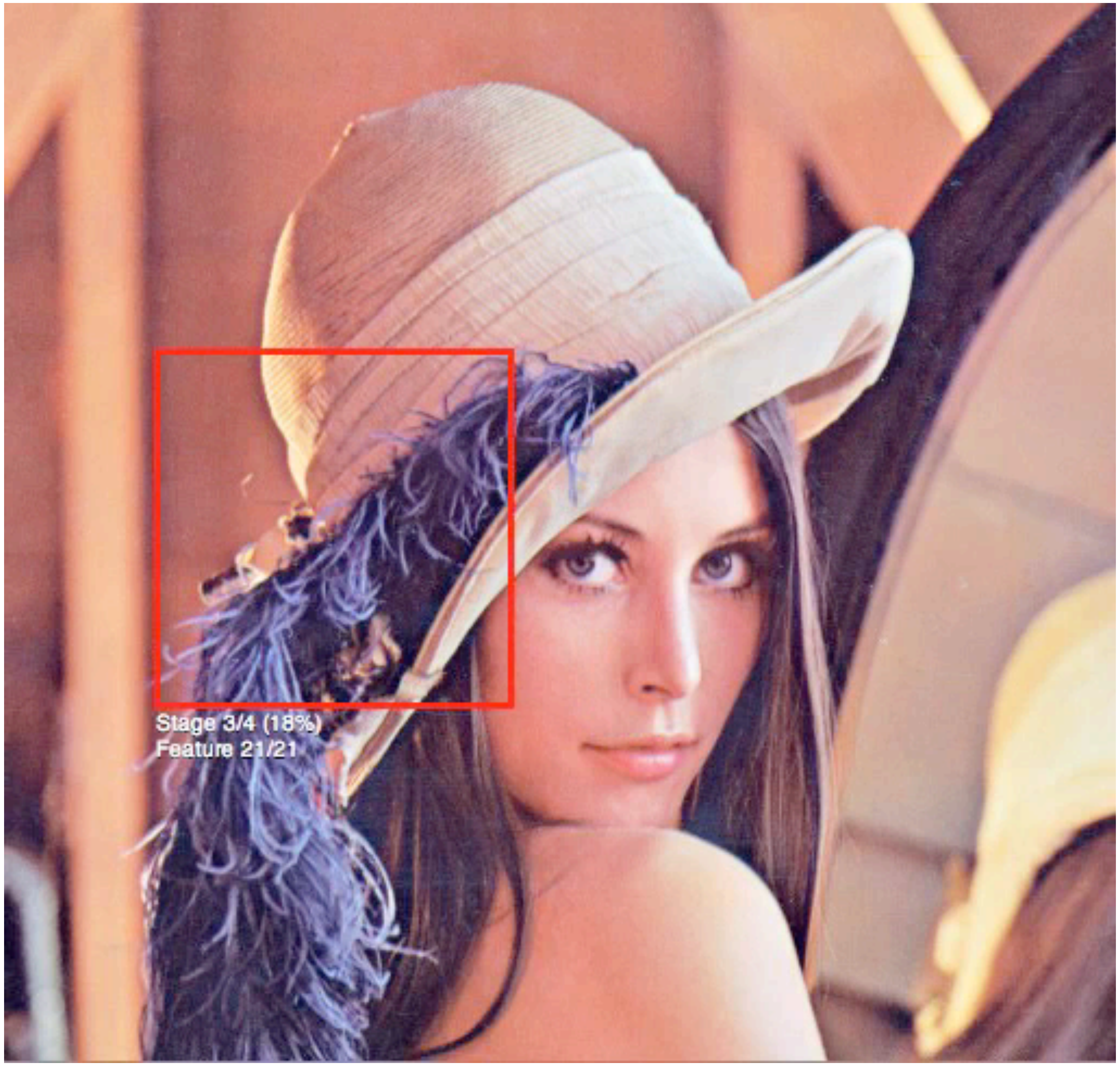
$$I_\Sigma(x,y) = \sum_{i=0}^{i\leq x} \sum_{j=0}^{j\leq y} I(i,j)$$



$$A + D - (C + B)$$



Example



Inertial Measurement Unit

iPhone 4, 4G iPod Touch, iPad 2

two hardware sources:
- three axis accelerometer
- three axis gyroscope

Inertial Measurement Unit

World-coordinate System (defined by the IMU)

spanned by the axes: $(x_{\text{world}}, y_{\text{world}}, z_{\text{world}})$

$$z_{\text{world}} := -\frac{a_{\text{gravitation}}}{||a_{\text{gravitation}}||}$$

$x_{\text{world}}, y_{\text{world}}$ defined orthogonal to gravity (right-hand space)

Inertial Measurement Unit

World-coordinate System (defined by the IMU)

spanned by the axes: $(x_{\text{world}}, y_{\text{world}}, z_{\text{world}})$

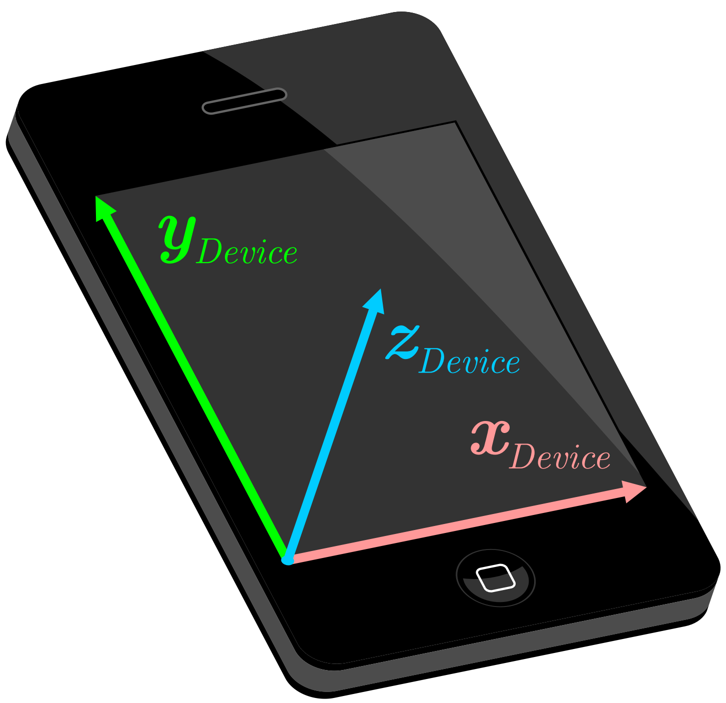
$$z_{\text{world}} := -\frac{a_{\text{gravitation}}}{||a_{\text{gravitation}}||}$$

$x_{\text{world}}, y_{\text{world}}$ defined orthogonal to gravity (right-hand space)

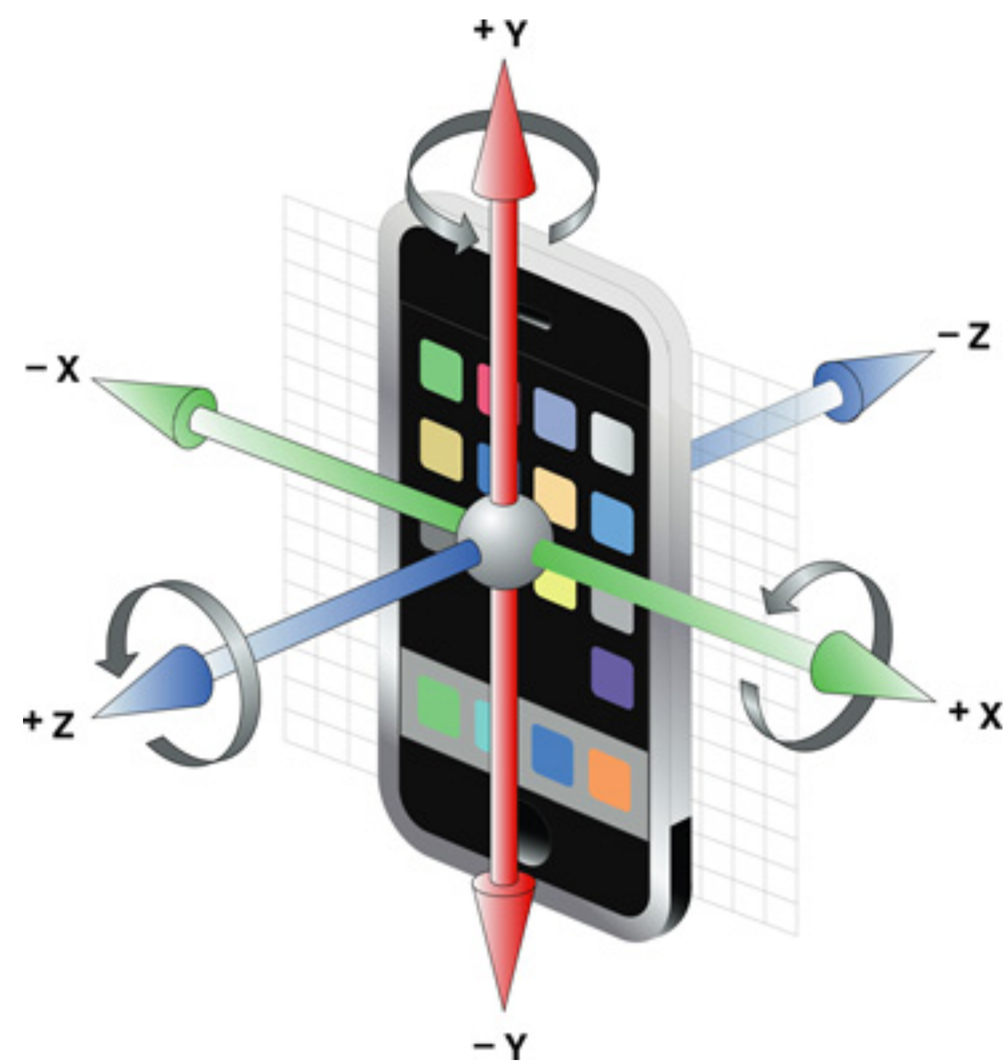
Apple provides a proprietary algorithm to calculate absolute orientations relative to the static world-coordinate system.

Proposed Method

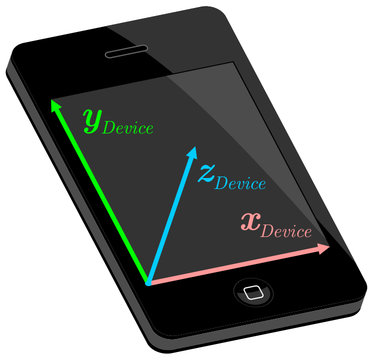
Camera-coordinate system:

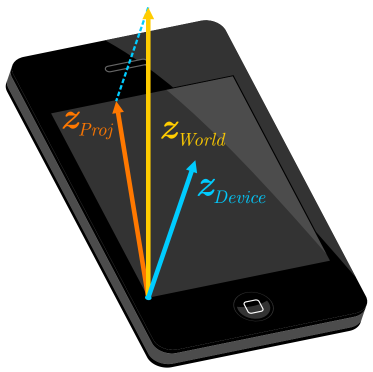


We are interested in the rotation around the z-axis (yaw-angle)



Proposed Method

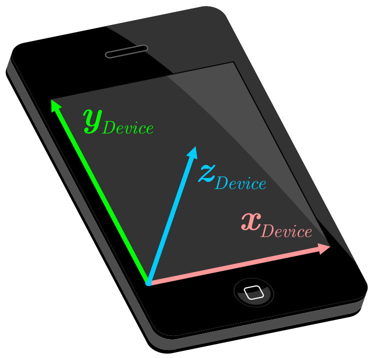


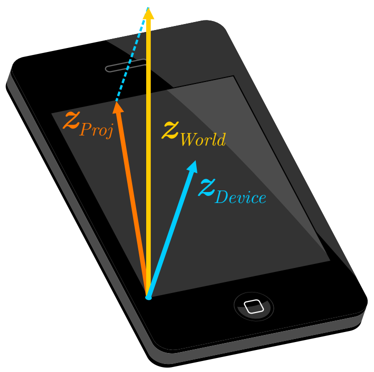


Proposed Method

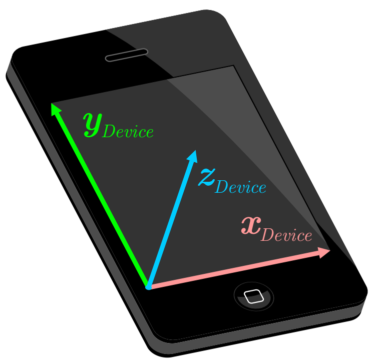
Apples proprietary algorithm gives:

$$(0, i_{\text{device},t}) = q(0, i_{\text{world}})q^{-1} \qquad (i \in \mathbb{R}^3)$$





Proposed Method



Apples proprietary algorithm gives:

$$(0, i_{\text{device},t}) = q(0, i_{\text{world}})q^{-1} \qquad (i \in \mathbb{R}^3)$$

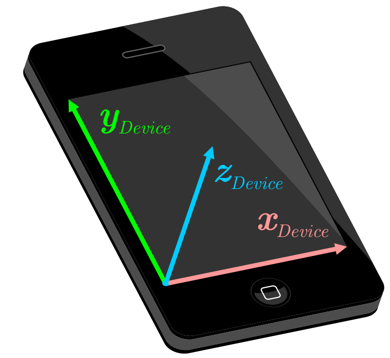
project z_{world} on the devices screen surface:

$$z_{\text{proj},t} = z_{\text{world}} - \langle z_{\text{world}} | z_{\text{device},t} \rangle \cdot z_{\text{device},t}$$

Proposed Method

Apples proprietary algorithm gives:

$$(0, i_{\text{device},t}) = q(0, i_{\text{world}})q^{-1} \qquad (i \in \mathbb{R}^3)$$



project z_{world} on the devices screen surface:

$$z_{\text{proj},t} = z_{\text{world}} - \langle z_{\text{world}} | z_{\text{device},t} \rangle \cdot z_{\text{device},t}$$

compute the yaw-angle:

$$\cos(\varphi_{\text{yaw}}) = \pm \langle y_{\text{device},t} | z_{\text{proj},t} / || z_{\text{proj},t} || \rangle$$

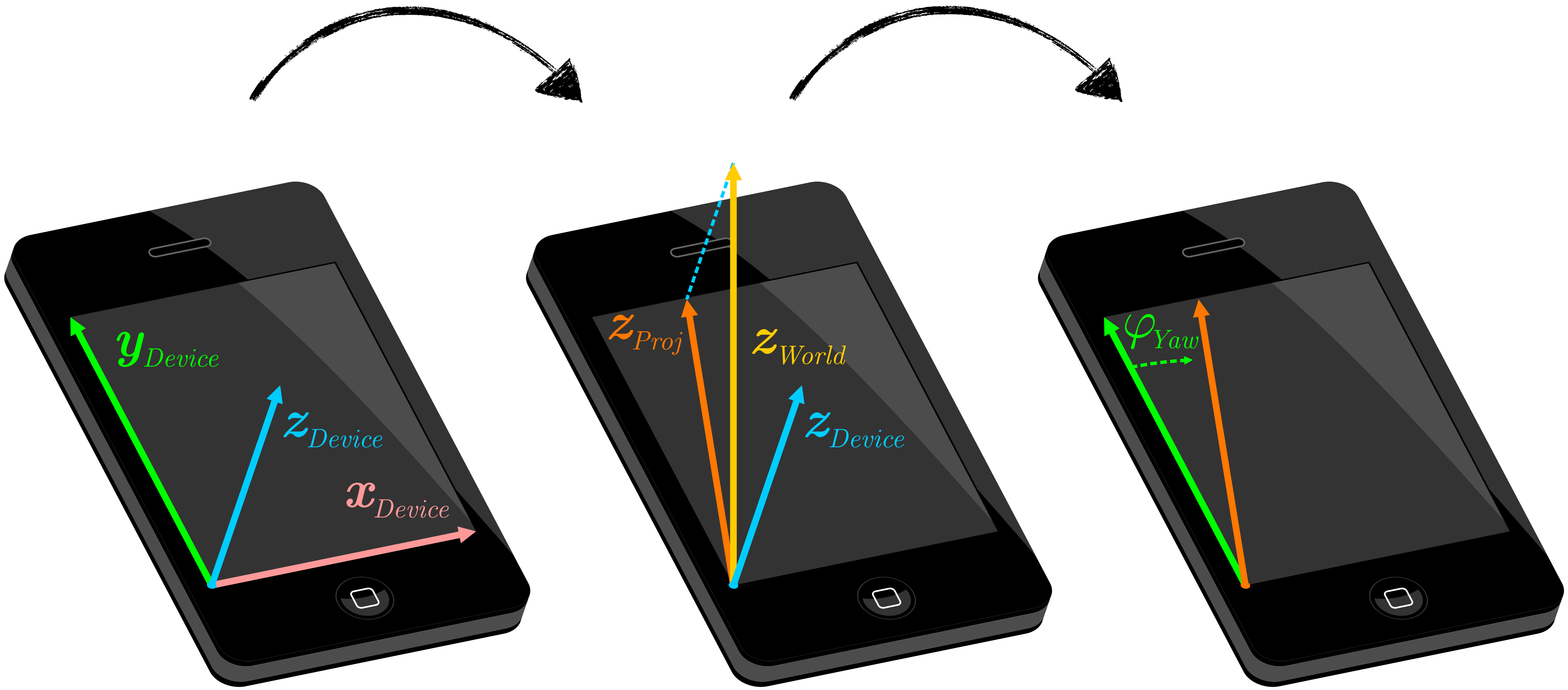
having the yaw-angle, we rotate the camera image
to produce a virtual upright image

Proposed Method

Camera coordinate system

project z_{world} on the device screen surface

compute the yaw-angle



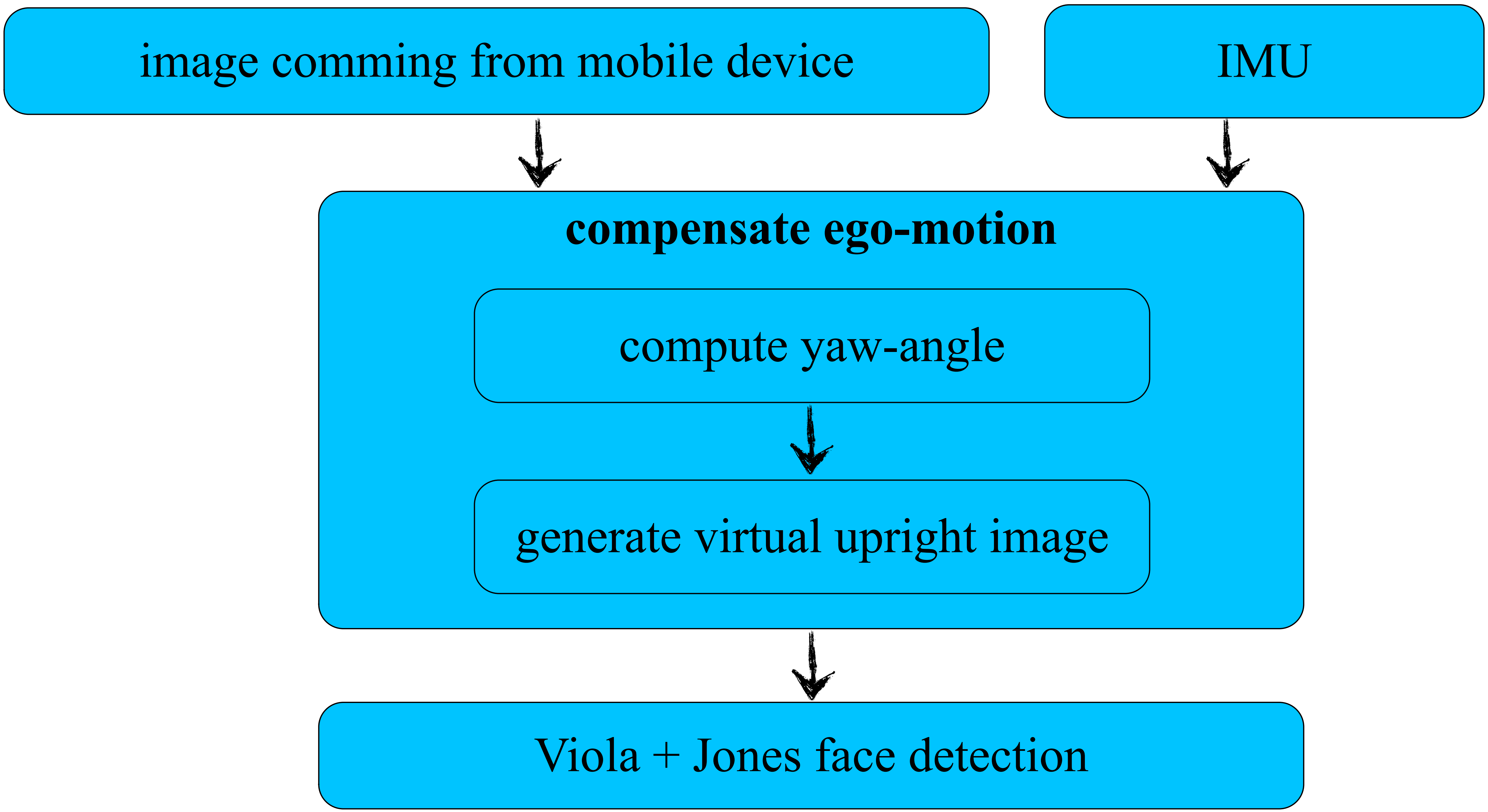
Standard Approach

image comming from mobile device

↓

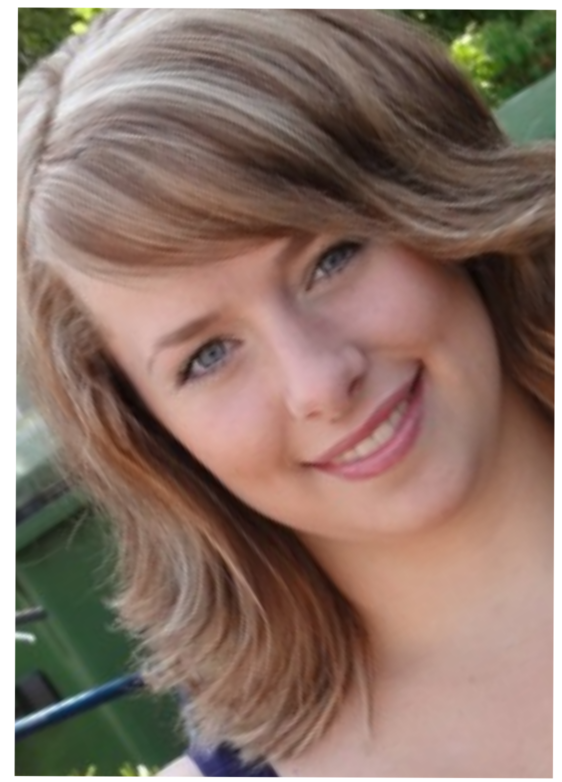
Viola + Jones face detection

Proposed Method

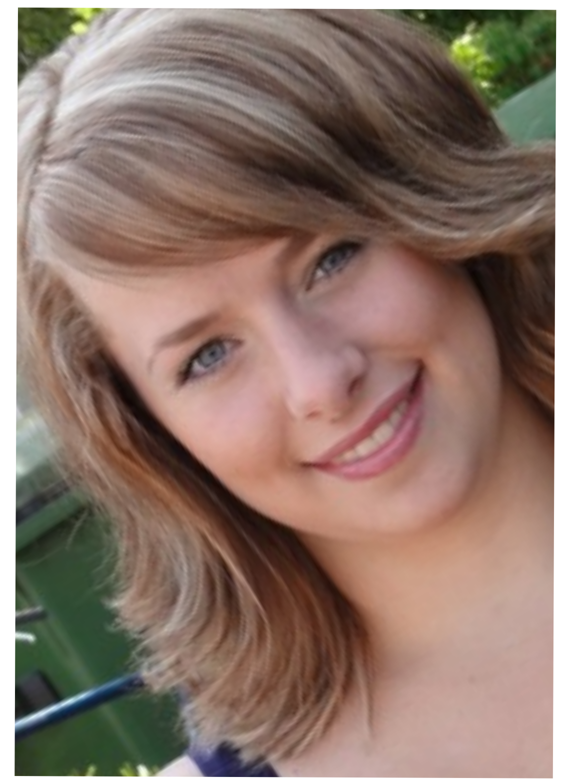


Proposed Method

Proposed Method

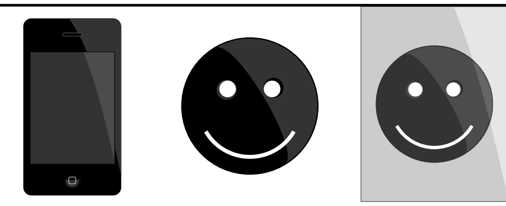
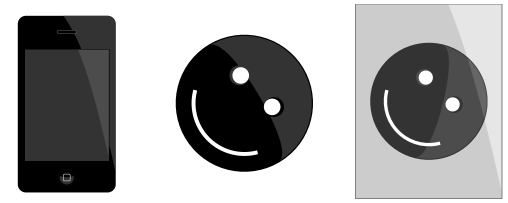
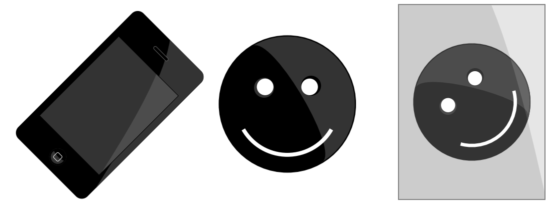


Proposed Method

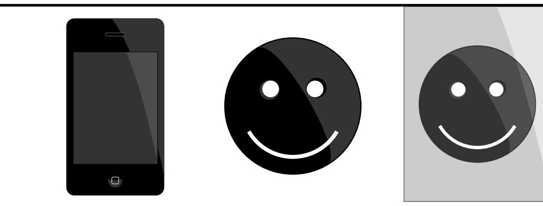
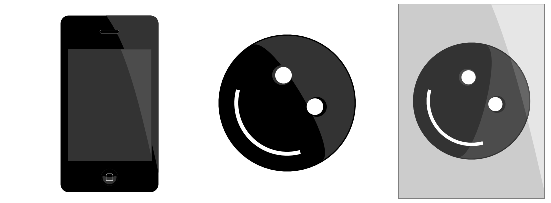
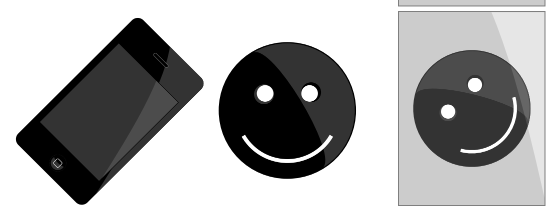


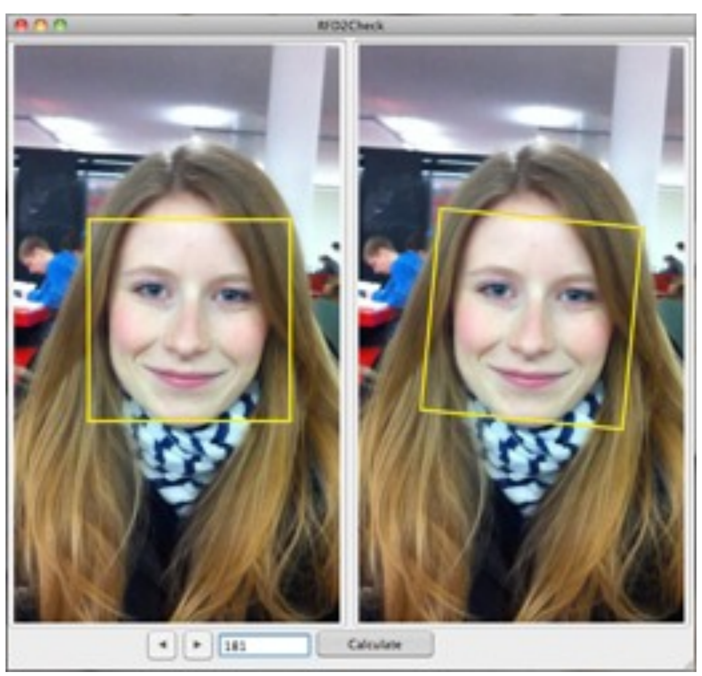
Proposed Method

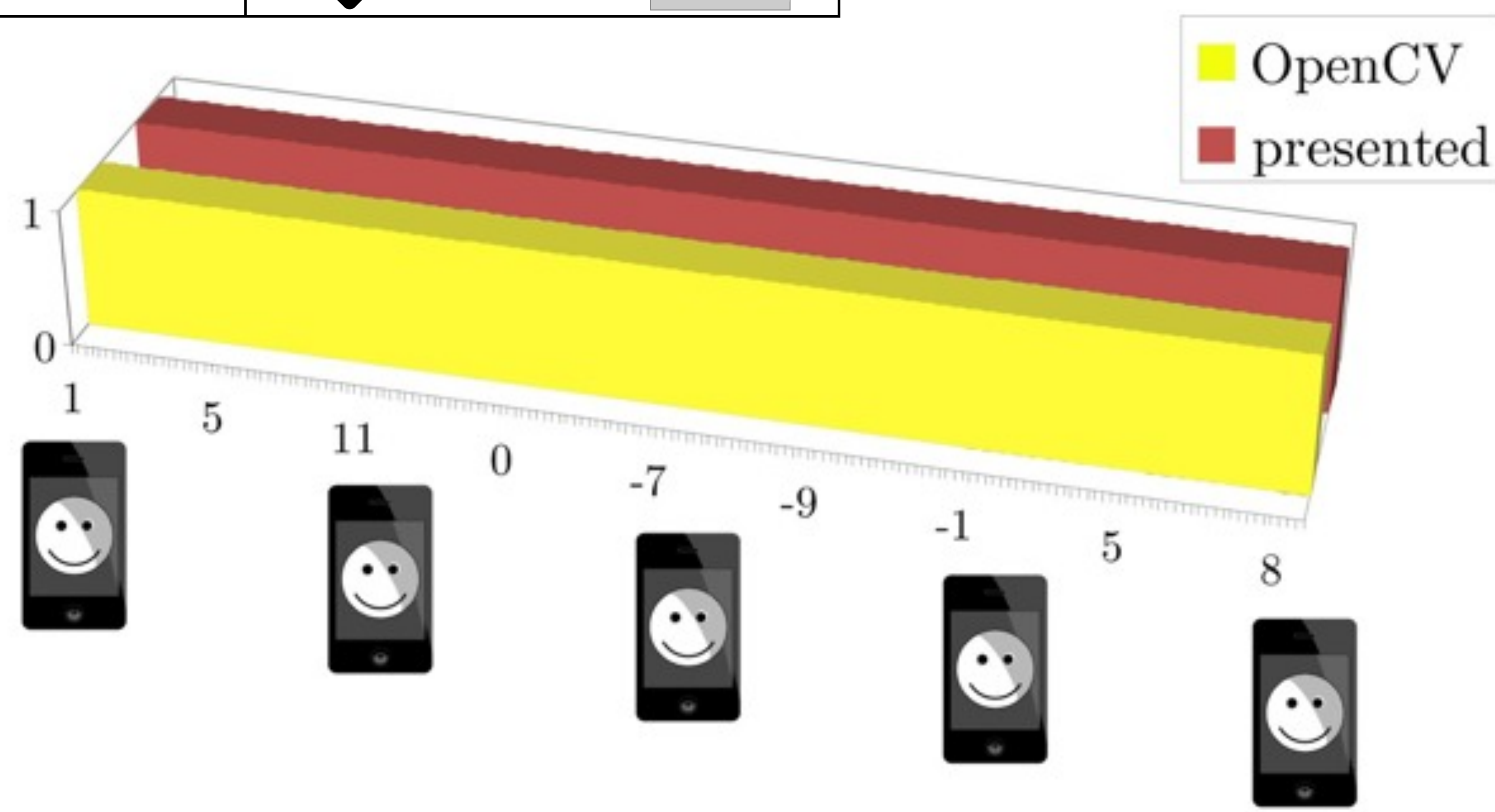
Experimental Results

	CAMERA	FACE	SYMBOLIC
Sequence 1	straight	straight	
Sequence 2	straight	rotated	
Sequence 3	rotated	straight	

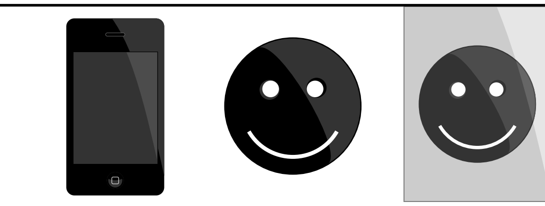
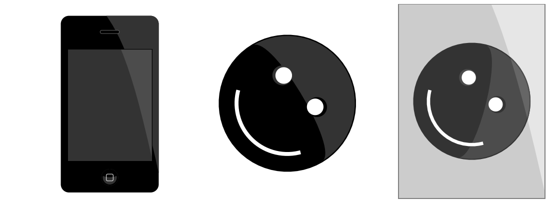
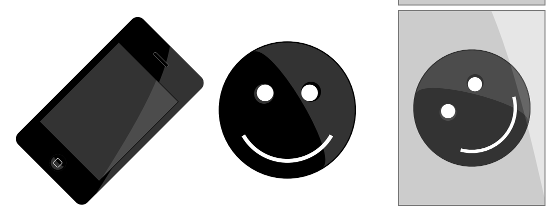
Experimental Results

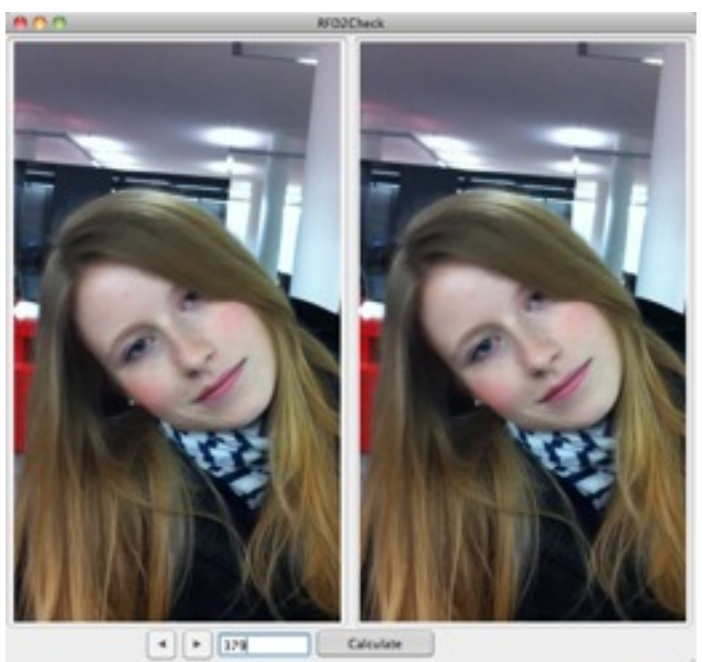
	CAMERA	FACE	SYMBOLIC
Sequence 1	straight	straight	
Sequence 2	straight	rotated	
Sequence 3	rotated	straight	

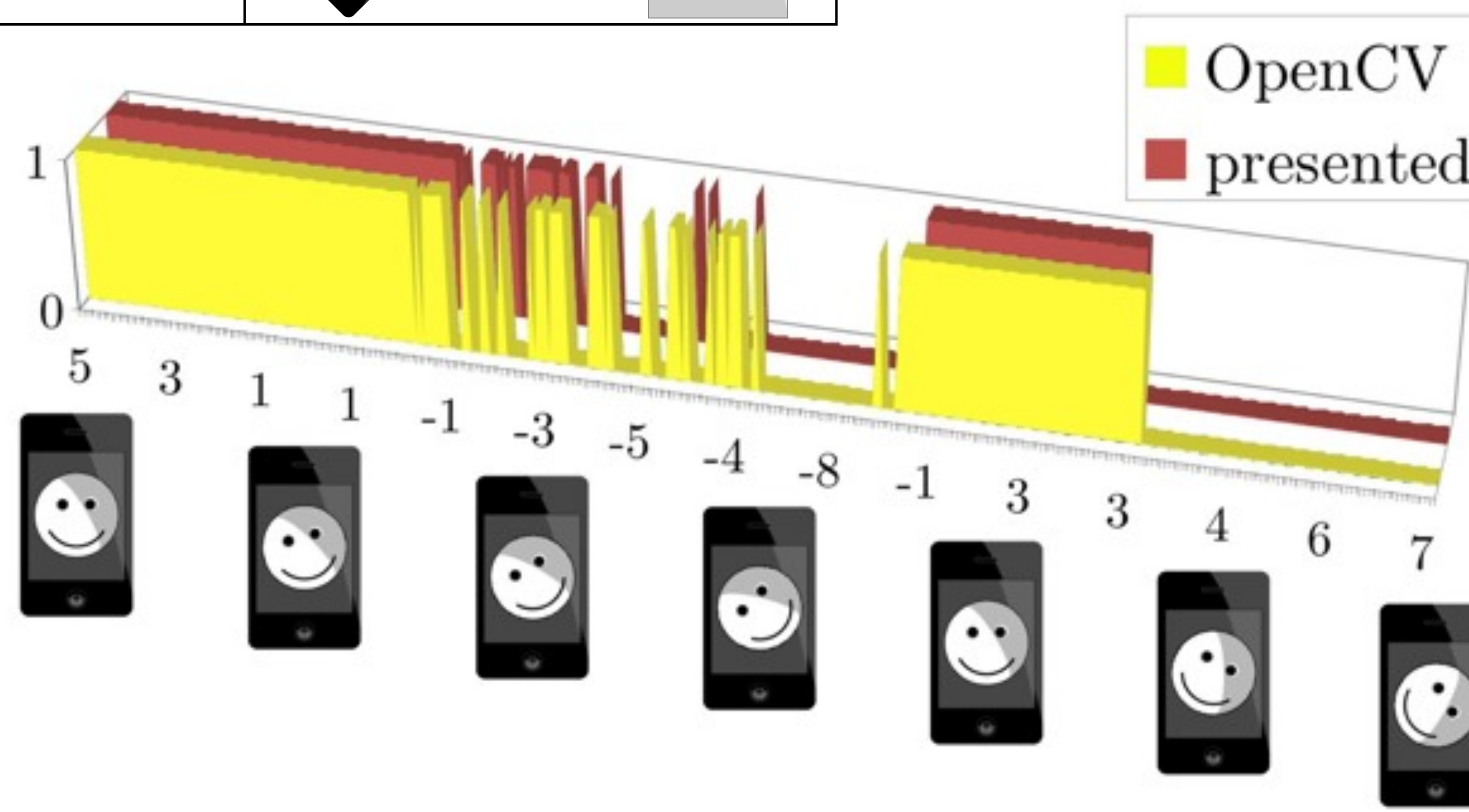




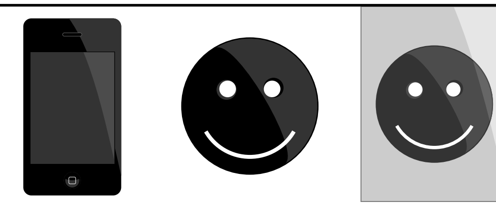
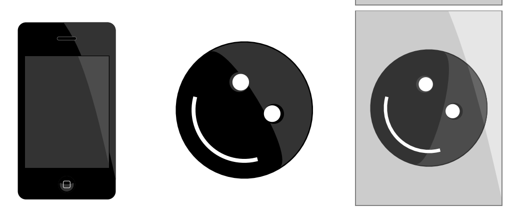
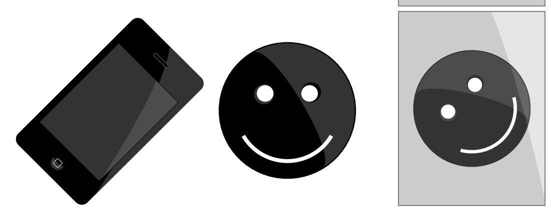
Experimental Results

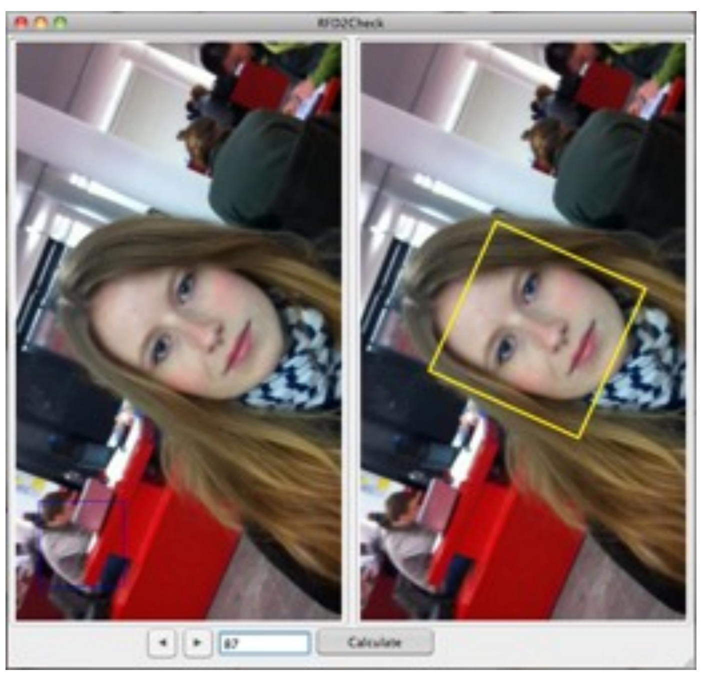
	CAMERA	FACE	SYMBOLIC
Sequence 1	straight	straight	
Sequence 2	straight	rotated	
Sequence 3	rotated	straight	

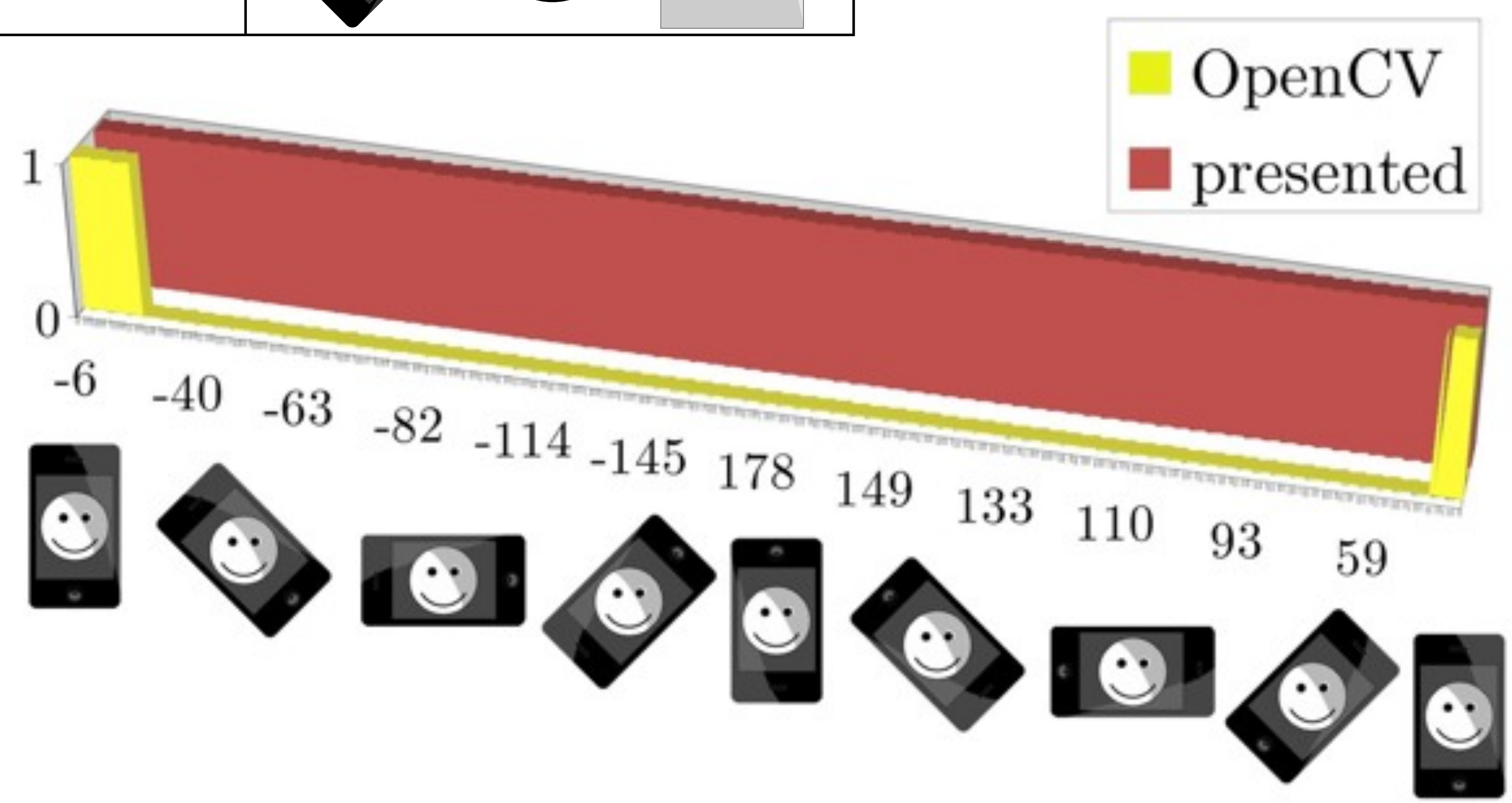




Experimental Results

	CAMERA	FACE	SYMBOLIC
Sequence 1	straight	straight	
Sequence 2	straight	rotated	
Sequence 3	rotated	straight	





Conclusion

- proposed an efficient fusion of IMU and boosting based face detection by Viola and Jones

Conclusion

• proposed an efficient fusion of IMU and boosting based face detection by Viola and Jones

• compared to other (rotation invariant) boosting based method our method is fast enough for a mobile device

Conclusion

• proposed an efficient fusion of IMU and boosting based face detection by Viola and Jones

• compared to other (rotation invariant) boosting based method our method is fast enough for a mobile device

• video sequences + motion data will be available

Conclusion

• proposed an efficient fusion of IMU and boosting based face detection by Viola and Jones

• compared to other (rotation invariant) boosting based method our method is fast enough for a mobile device

• video sequences + motion data will be available

• online demo is available (just ask)

Conclusion

- proposed an efficient fusion of IMU and boosting based face detection by Viola and Jones

- compared to other (rotation invariant) boosting based method our method is fast enough for a mobile device

- video sequences + motion data will be available

- online demo is available (just ask)

- compensate face motion is an ongoing project

Conclusion

• proposed an efficient fusion of IMU and boosting based face detection by Viola and Jones

• compared to other (rotation invariant) boosting based method our method is fast enough for a mobile device

• video sequences + motion data will be available

• online demo is available (just ask)

• compensate face motion is an ongoing project

Thanks for your attention!!!