IS WHAT YOU SEE WHAT YOU GET?

Testing Computer-Vision applications

Michael Stahl, SW Validation Architect
Intel
Nov 2015

Agenda

- Intel RealSense® technology
- Demos
- Testing challenges
- Some answers
- Summary

INTEL® REALSENSE™ TECHNOLOGY

Human Senses. Natural, Intuitive and Immersive

Intel® RealSense™ Devices

intel.com/RealSense/devices

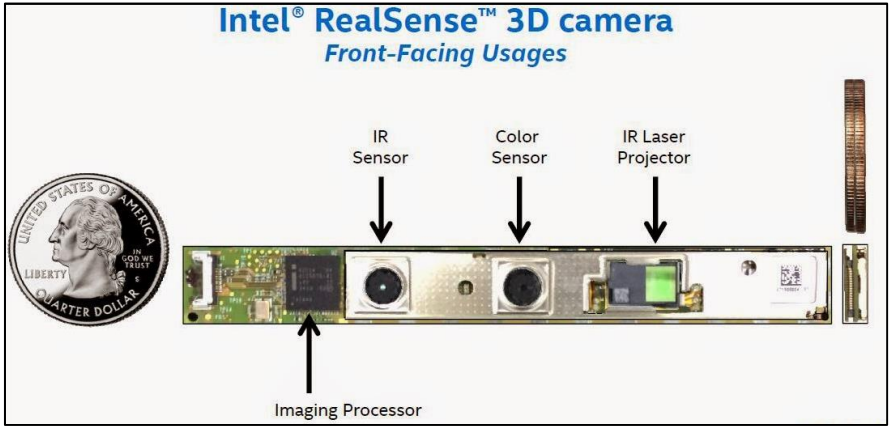
Intel® RealSense™ Cameras - 2015

	User-facing (F200)	World-facing (R200)
devkits available today		
products		

acer ASUS DELL FUJITSU

hp lenovo NEC

Front-facing 3D camera (F200)



© Michael Stahl, 2015, All Rights Reserved

Front-facing 3D camera (F200)

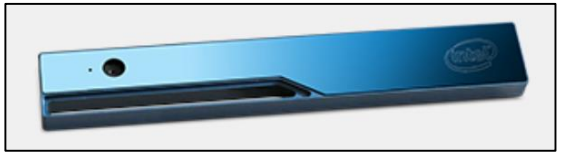
F200

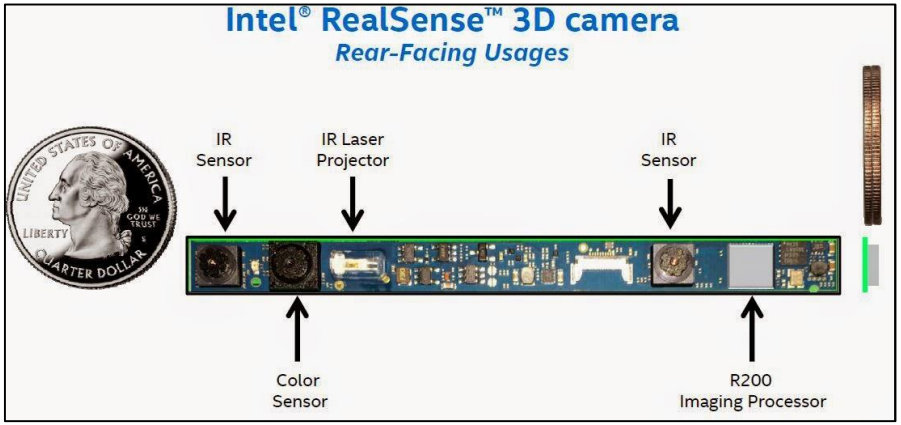
- Features available with the Intel® RealSense™ SDK
  - Hand tracking and gestures
  - Face recognition and animation
  - Background replacement
  - 3D scanning
  - Speech

Source: Intel Developers Forum 2015 presentation

© Michael Stahl, 2015, All Rights Reserved

Rear-facing 3D camera (R200)

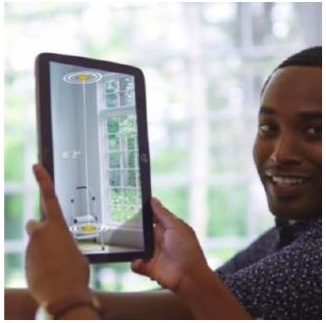


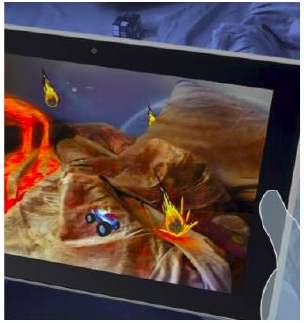


© Michael Stahl, 2015, All Rights Reserved

Rear-facing 3D camera (R200)

- Features available with the Intel® RealSense™ SDK
  - 3D scanning
  - Depth-enabled photo and video
  - Measurement
  - Gaming and play with mixed and augmented reality
    - Scene perception module
    - Camera tracking and localization
    - Mesh reconstruction
  - Face detection and tracking
  - Speech (Windows® SDK only)



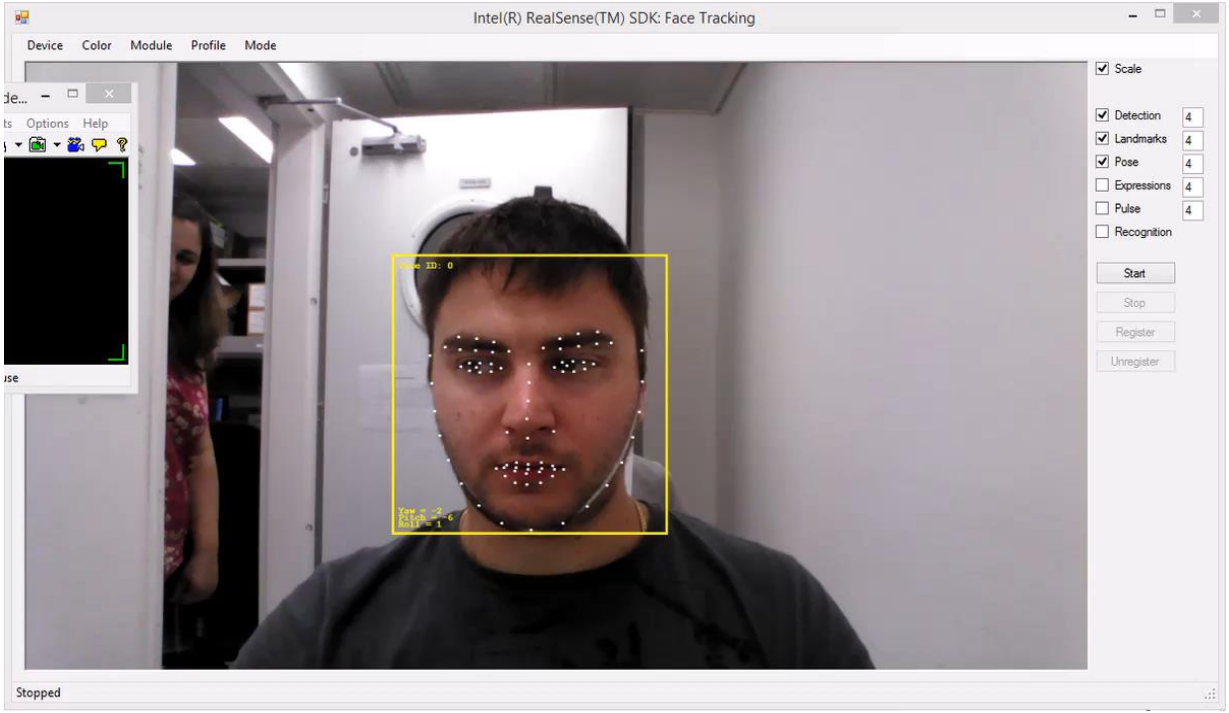


Source: Intel Developers Forum 2015 presentation

© Michael Stahl, 2015, All Rights Reserved

DEMOS

Face



© Michael Stahl, 2015, All Rights Reserved

Face

© Michael Stahl, 2015, All Rights Reserved

Face

© Michael Stahl, 2015, All Rights Reserved

Enhanced Photography

© Michael Stahl, 2015, All Rights Reserved

Enhanced Photography

© Michael Stahl, 2015, All Rights Reserved

Enhanced Photography

© Michael Stahl, 2015, All Rights Reserved

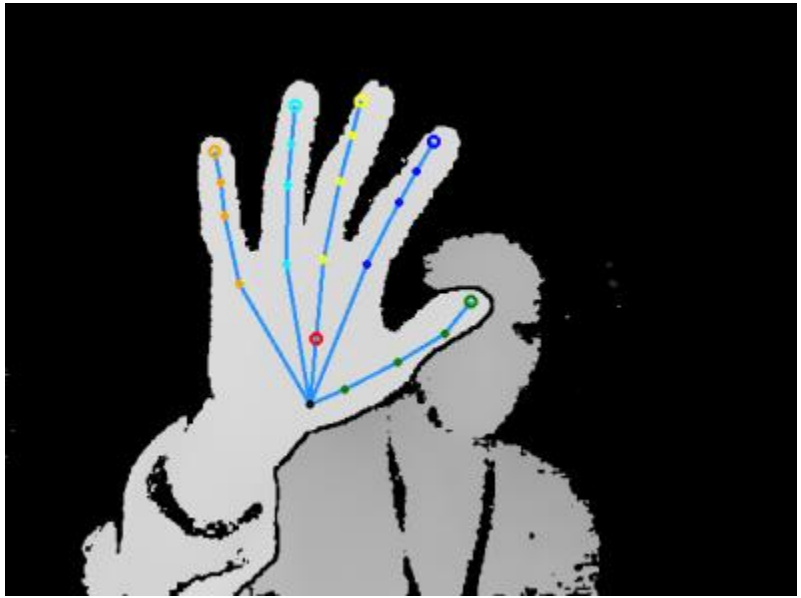
Enhanced Photography

© Michael Stahl, 2015, All Rights Reserved

Background Removal

© Michael Stahl, 2015, All Rights Reserved

Gesture



© Michael Stahl, 2015, All Rights Reserved

Augmented Reality

© Michael Stahl, 2015, All Rights Reserved

TESTING CHALLENGES

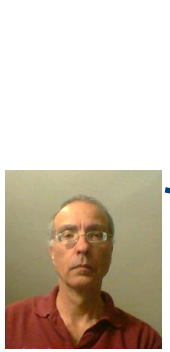
Face Detection

- Pass?
- Fail?
- Good Enough?

https://software.intel.com/en-us/articles/intel-realsense-depth-camera-r200-code-sample-face-tracking

© Michael Stahl, 2015, All Rights Reserved

First test



HW
FW

SW (driver,
MW, apps)

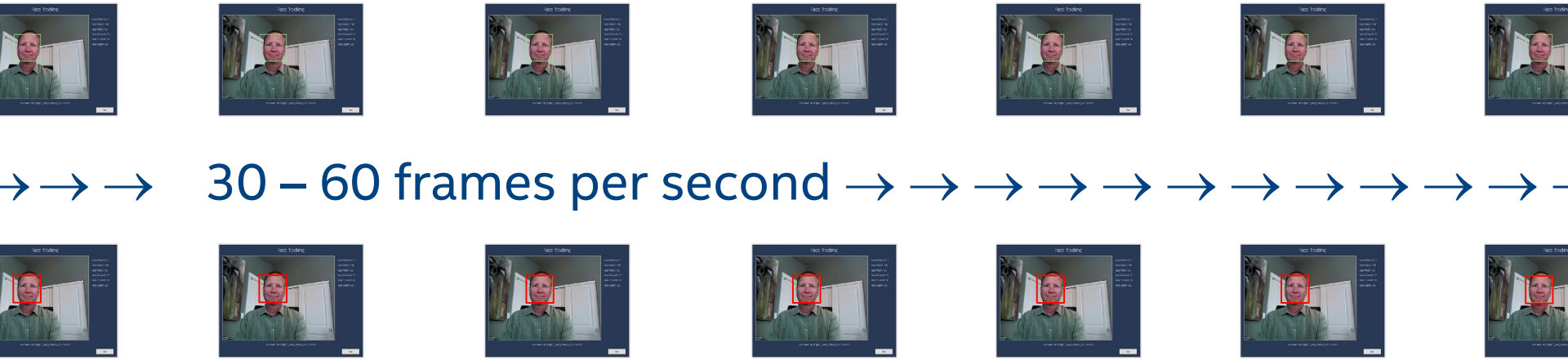
Compare

Pass /
Fail

"Ground Truth"
(G.T.)

© Michael Stahl, 2015, All Rights Reserved

The Ground Truth challenge



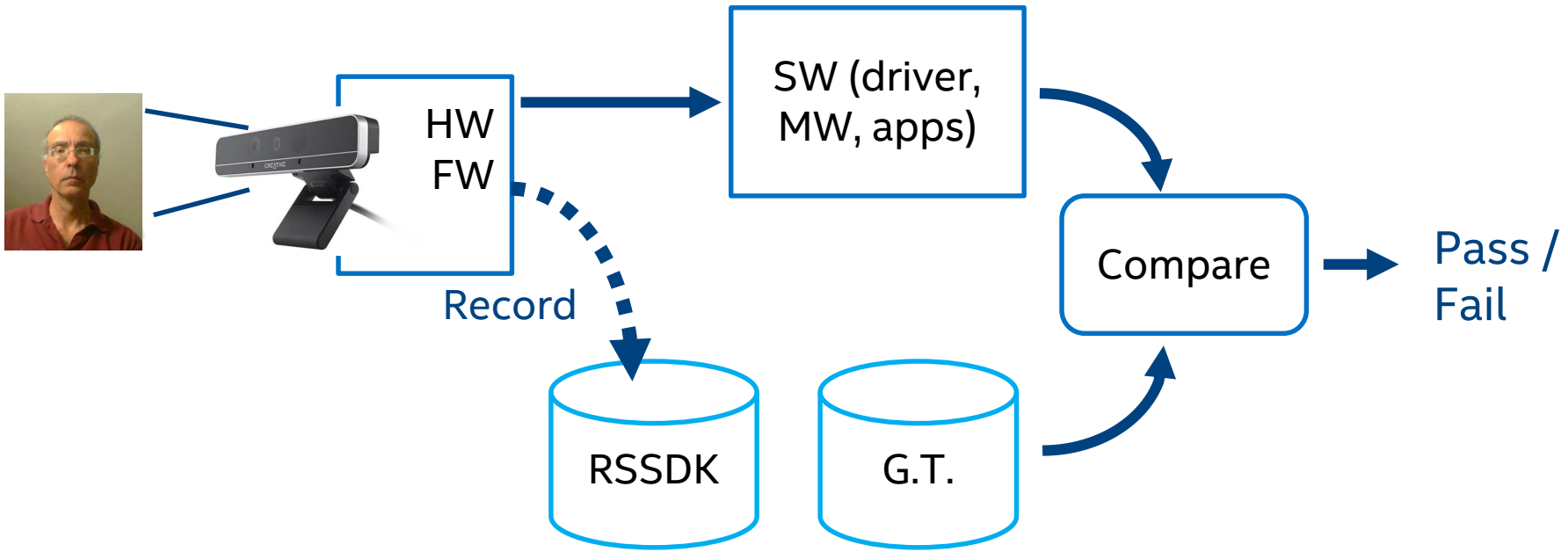
Quis custodiet ipsos custodes?

© Michael Stahl, 2015, All Rights Reserved

The Ground Truth challenge

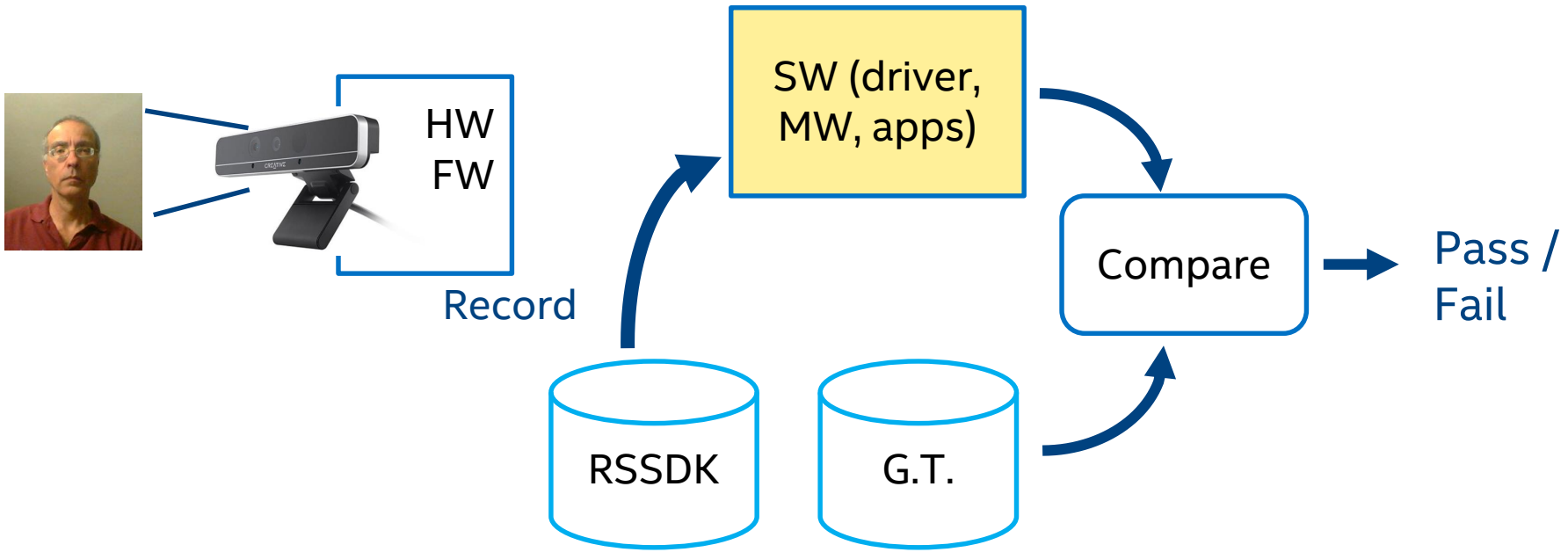
© Michael Stahl, 2015, All Rights Reserved

The Ground Truth challenge: First test



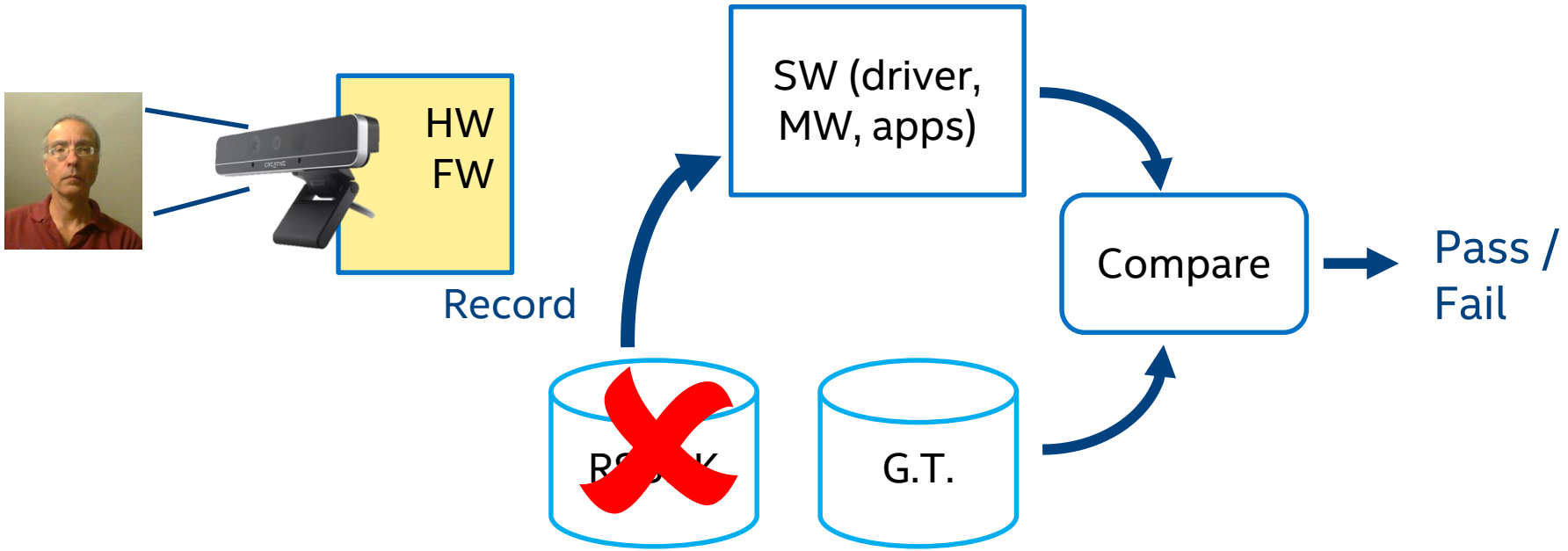
© Michael Stahl, 2015, All Rights Reserved

The Ground Truth challenge: Regression



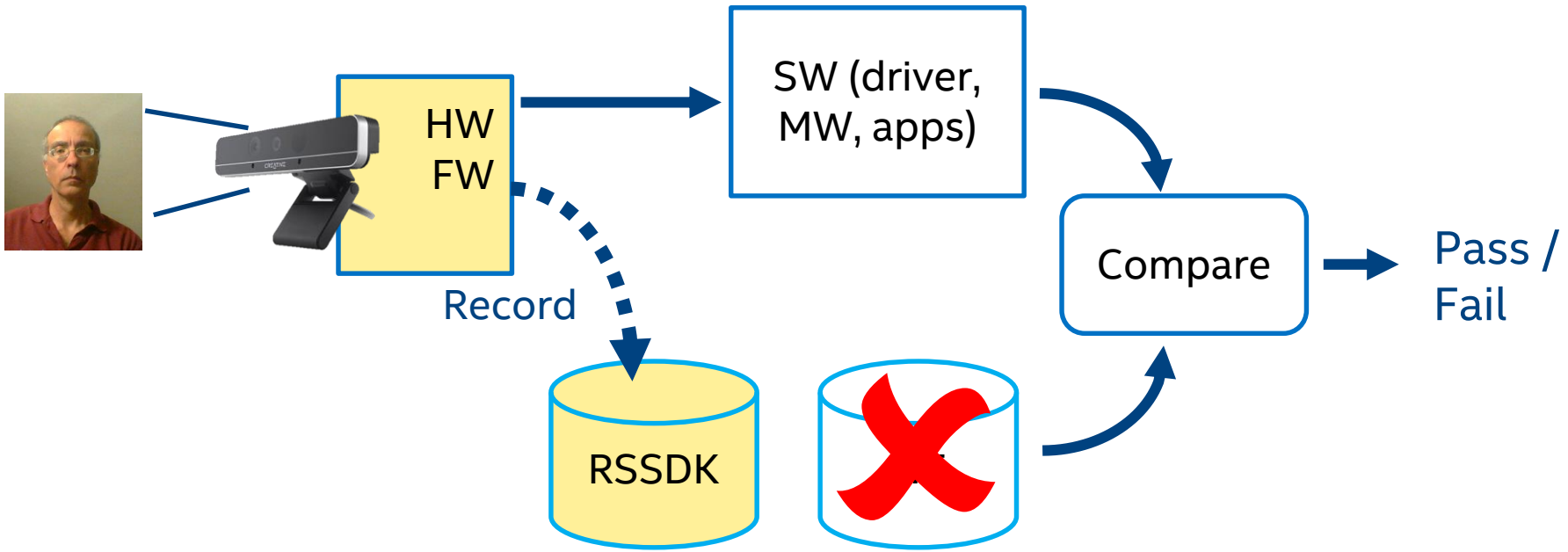
© Michael Stahl, 2015, All Rights Reserved

The Ground Truth challenge: HW, FW change



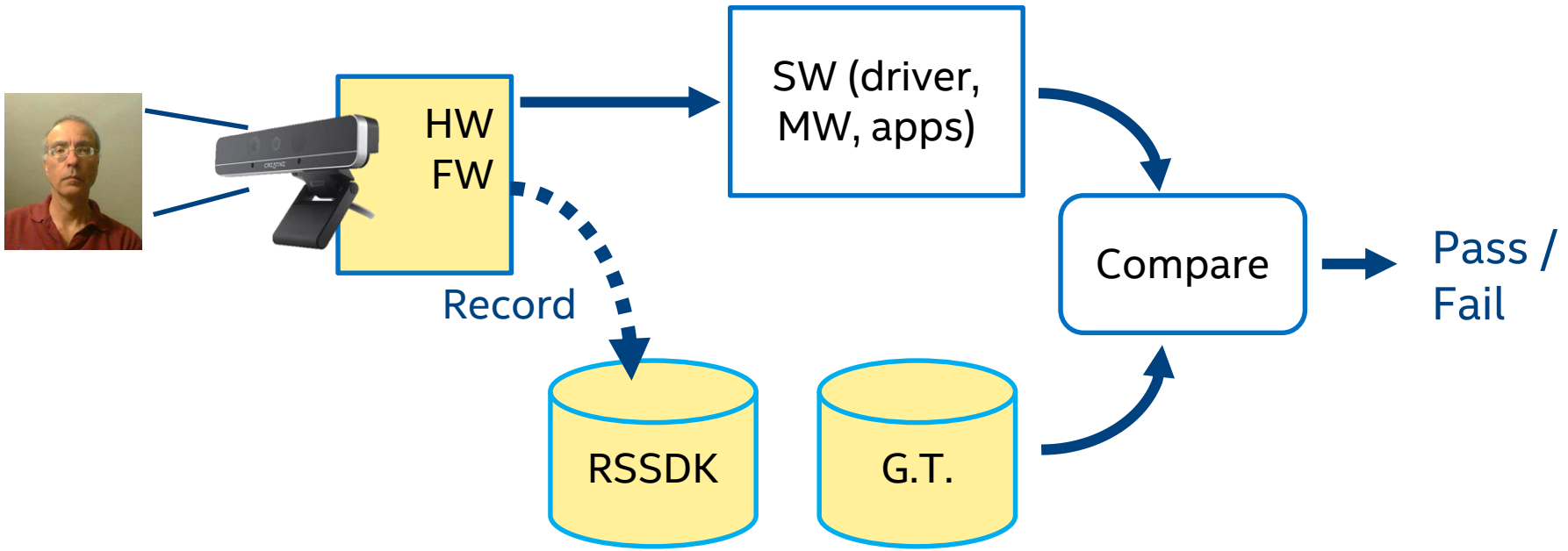
© Michael Stahl, 2015, All Rights Reserved

The Ground Truth challenge: HW, FW change



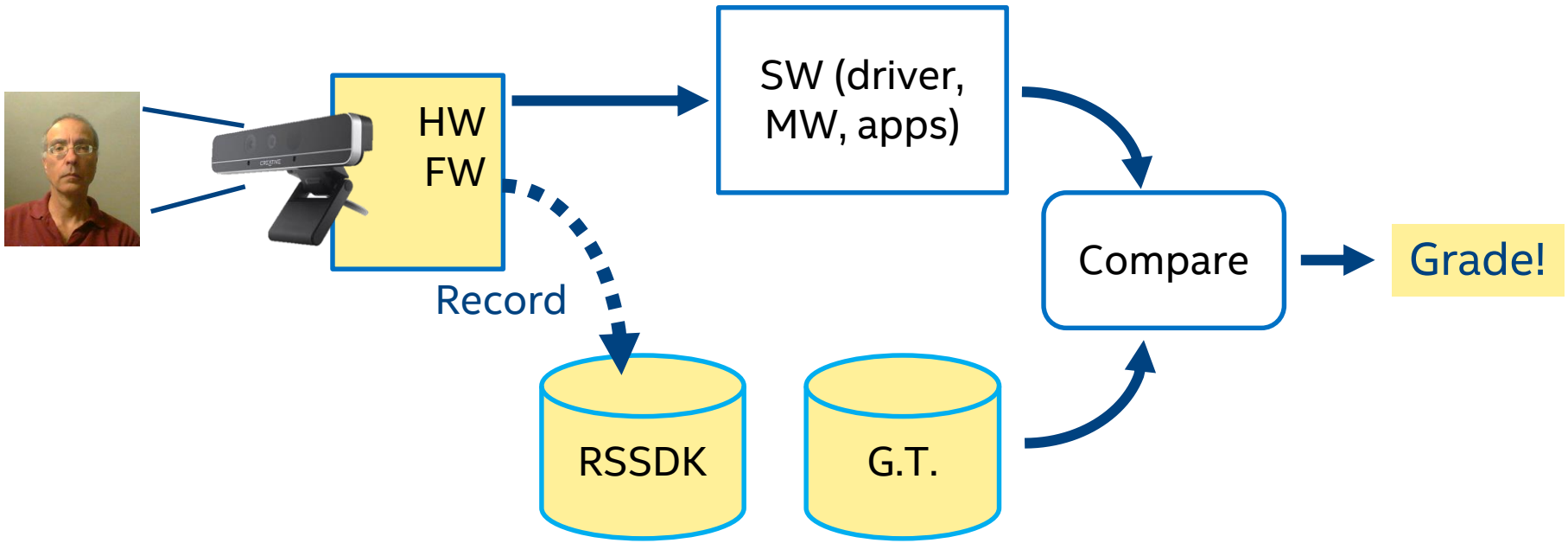
© Michael Stahl, 2015, All Rights Reserved

The Ground Truth challenge: HW, FW change



© Michael Stahl, 2015, All Rights Reserved

The Ground Truth challenge: HW, FW change



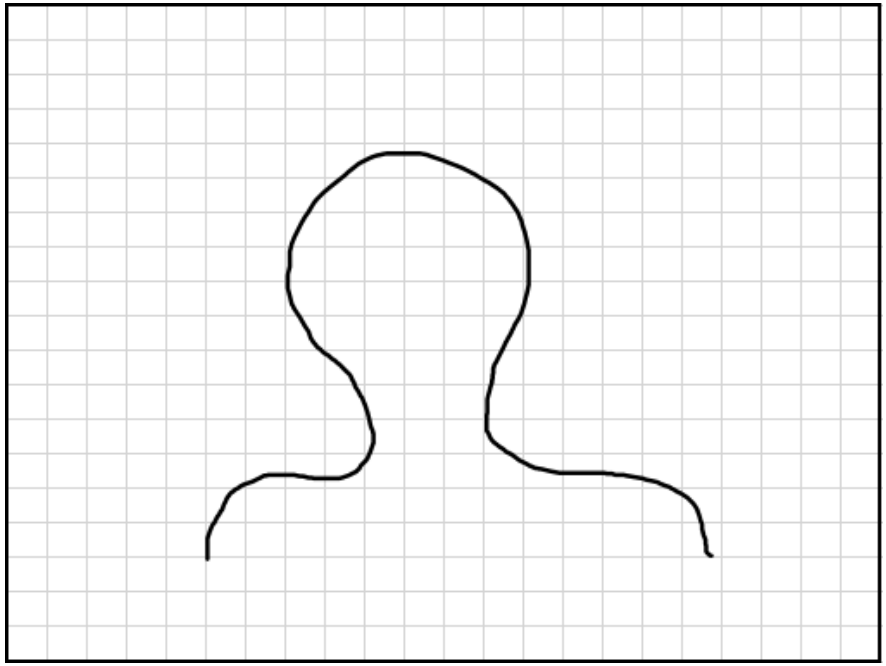
© Michael Stahl, 2015, All Rights Reserved

The Test Data challenge

1 MIN = 1GB

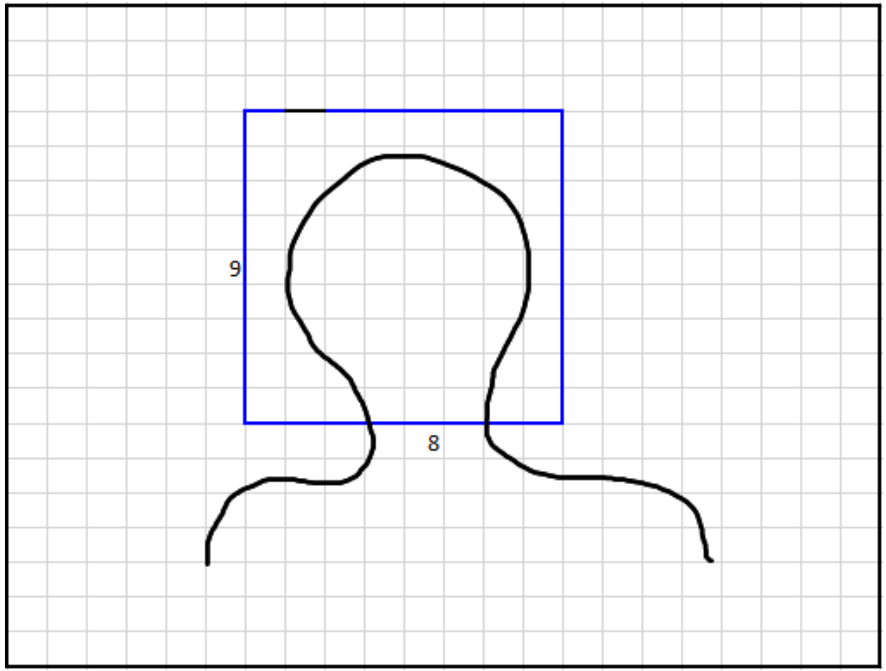
© Michael Stahl, 2015, All Rights Reserved

The metrics challenge



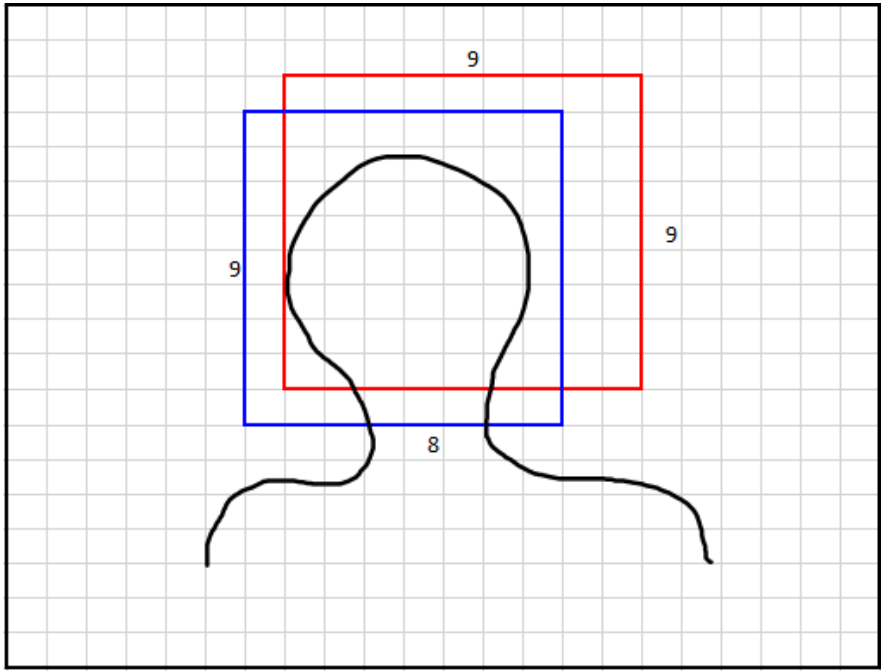
© Michael Stahl, 2015, All Rights Reserved

The metrics challenge



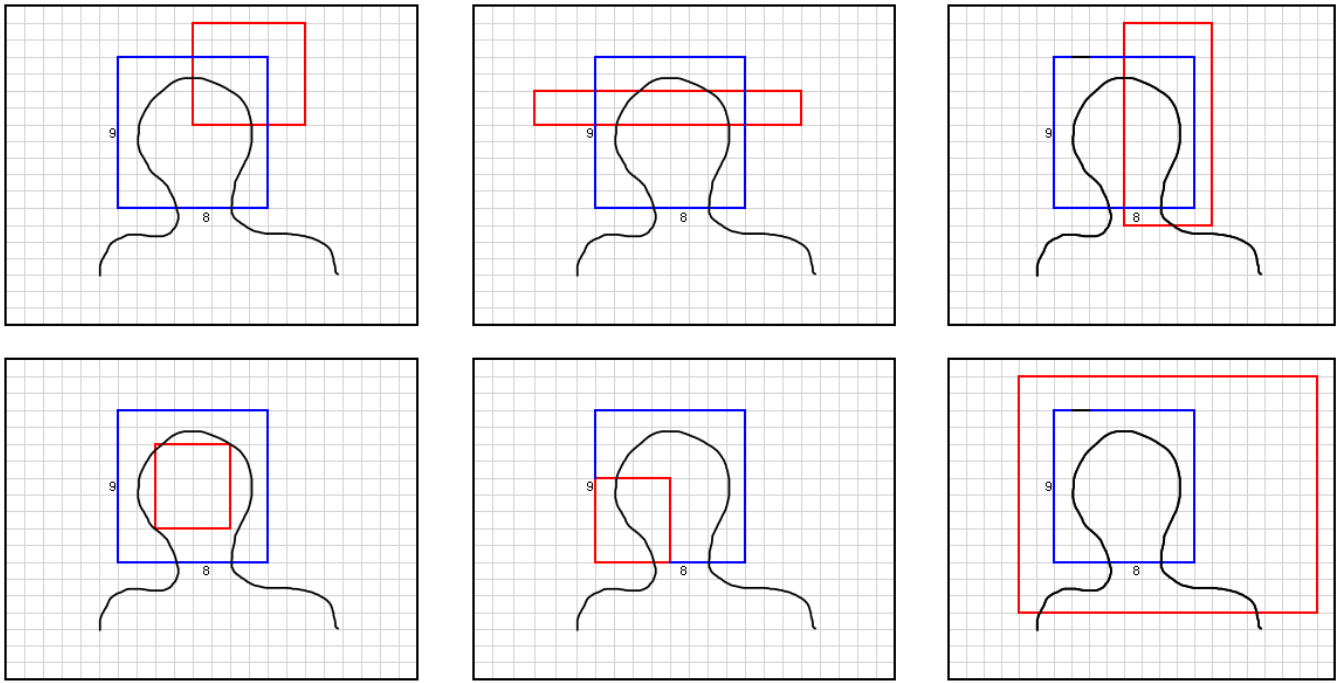
© Michael Stahl, 2015, All Rights Reserved

The metrics challenge



© Michael Stahl, 2015, All Rights Reserved

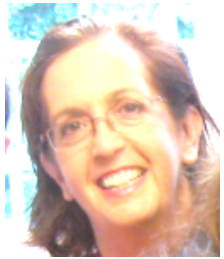
The metrics challenge – What's better?



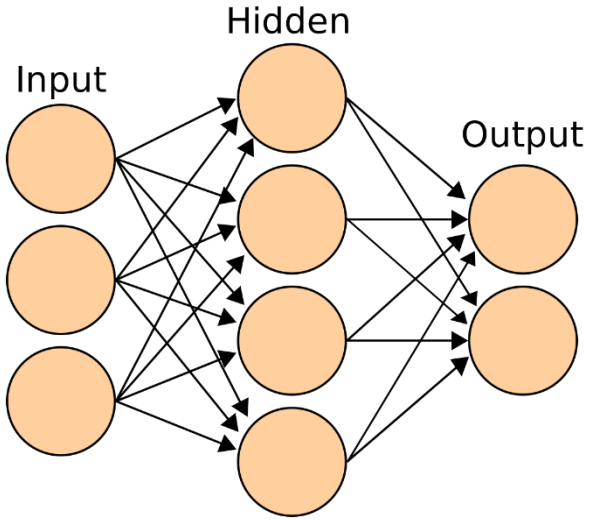
© Michael Stahl, 2015, All Rights Reserved

The test technique challenge

Valid EC



Invalid EC



The end of EC testing?

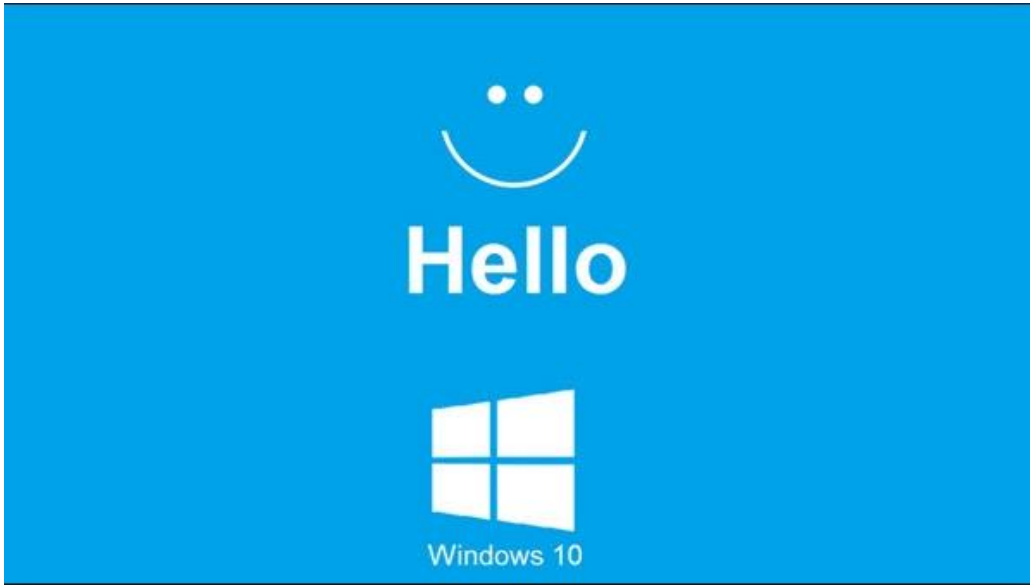
© Michael Stahl, 2015, All Rights Reserved

The test technique challenge

What's the Boundary Value?...

© Michael Stahl, 2015, All Rights Reserved

The combinatorial coverage challenge



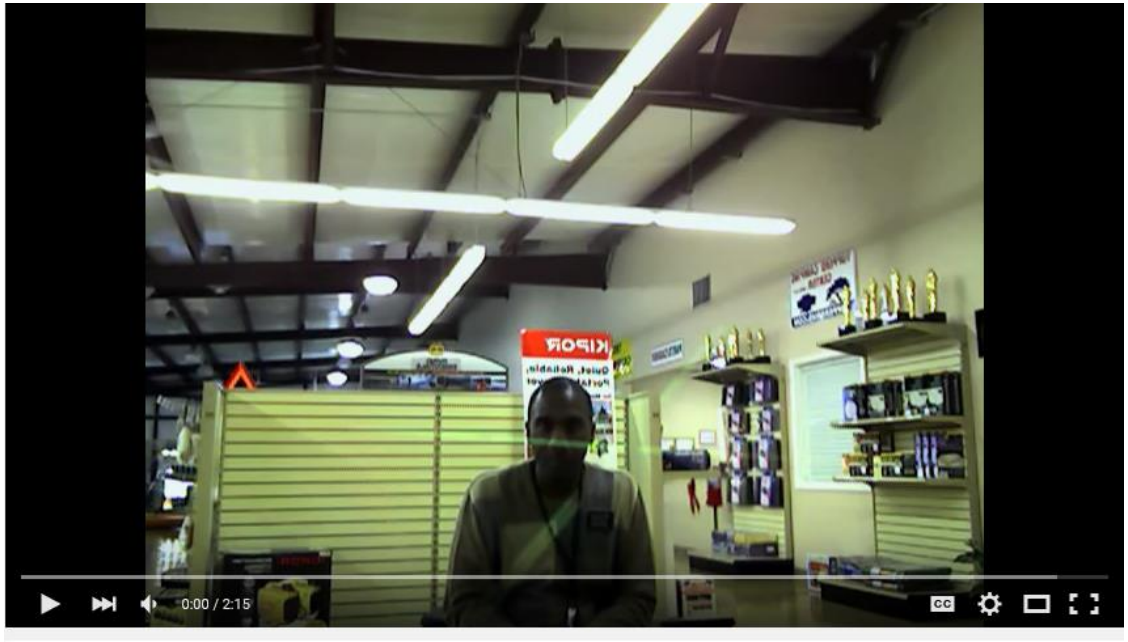
© Michael Stahl, 2015, All Rights Reserved

The combinatorial coverage challenge

amazon fire

© Michael Stahl, 2015, All Rights Reserved

The PC challenge



HP computers are racist

https://www.youtube.com/watch?v=t4DT3tQqgRM

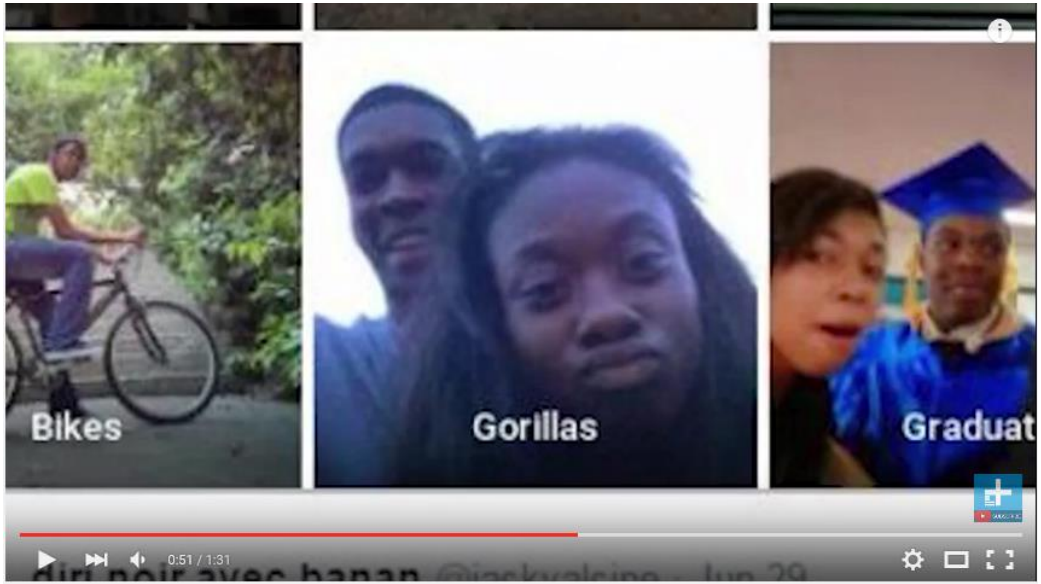
© Michael Stahl, 2015, All Rights Reserved

The PC challenge

https://www.youtube.com/watch?v=t4DT3tQqgRM

© Michael Stahl, 2015, All Rights Reserved

The PC challenge

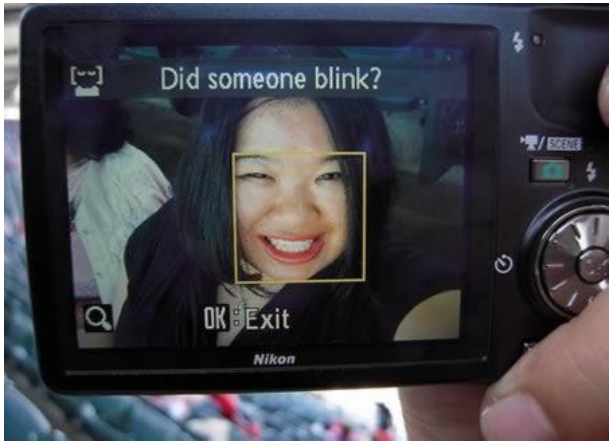


Google's Photos app labels black couple as gorillas: DT Daily

https://www.youtube.com/watch?v=MsR7GPnpzMY

© Michael Stahl, 2015, All Rights Reserved

The PC challenge



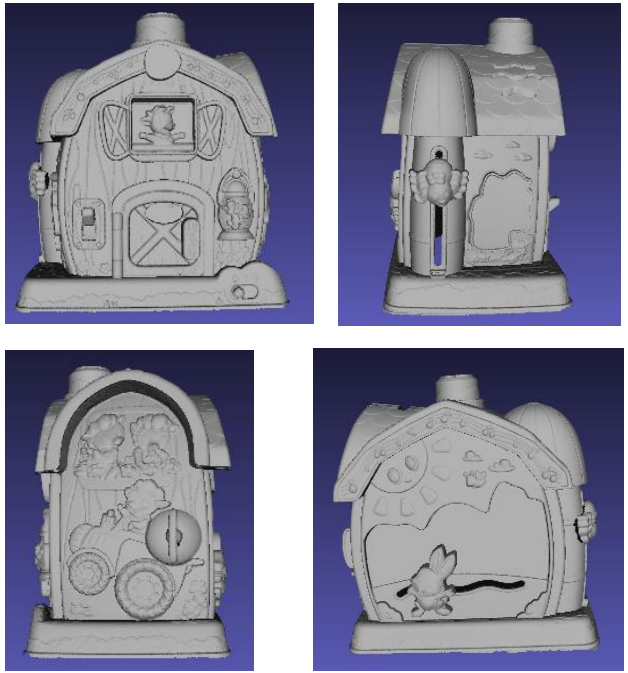
© Michael Stahl, 2015, All Rights Reserved

SOME ANSWERS

The Ground Truth challenge

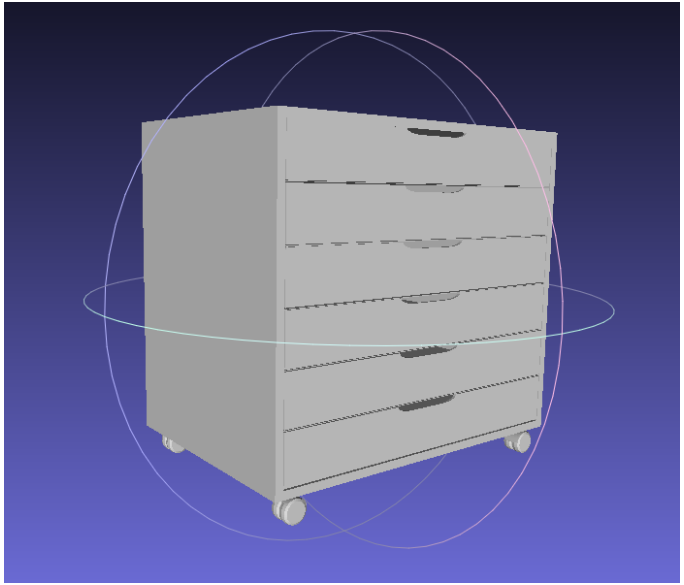
© Michael Stahl, 2015, All Rights Reserved

Ground Truth from competing products



© Michael Stahl, 2015, All Rights Reserved

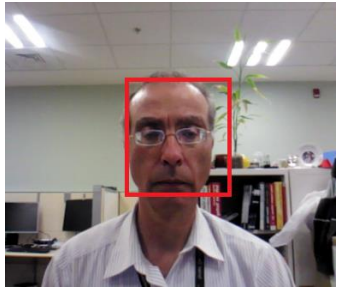
Ground Truth from CAD models



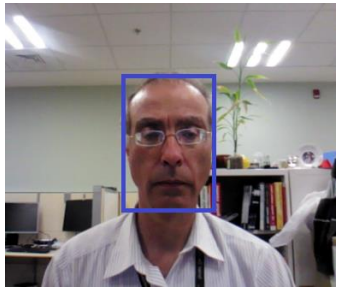
© Michael Stahl, 2015, All Rights Reserved

Ground Truth by other algorithms

Algorithm under Test

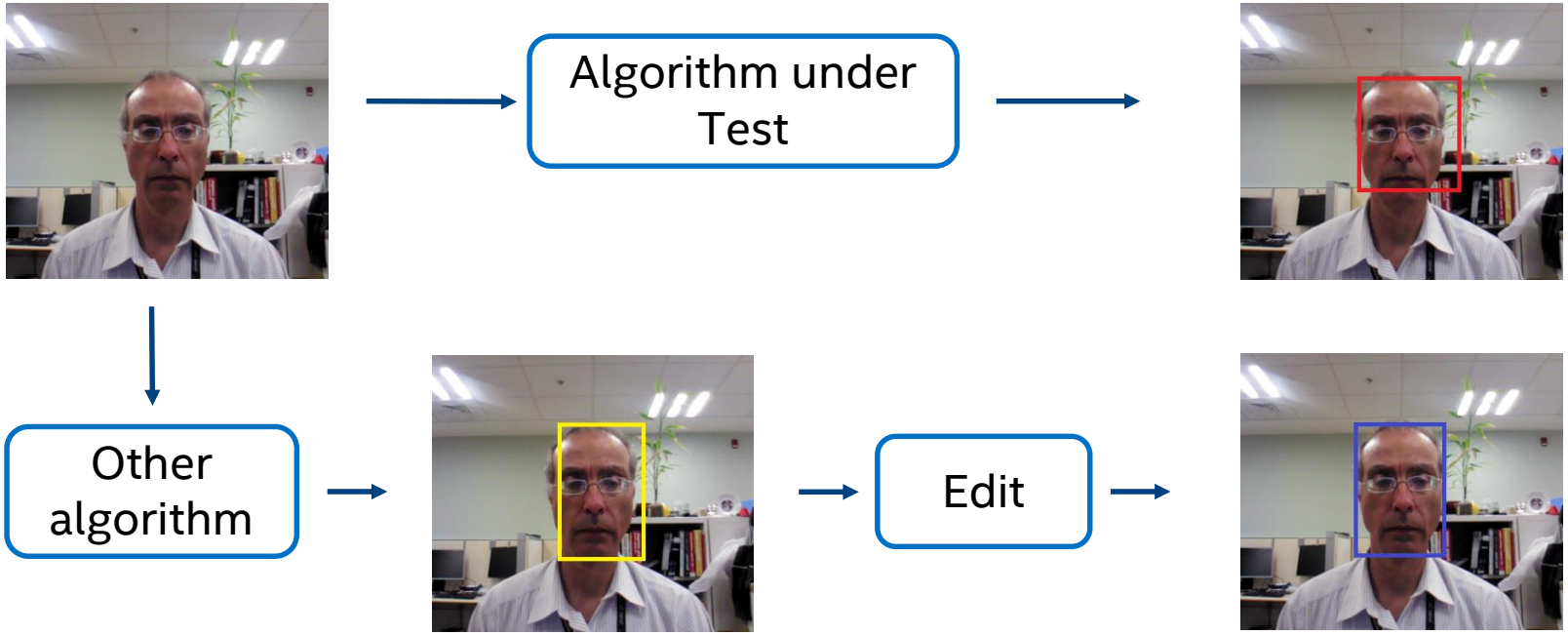


Other algorithm (e.g. slow)



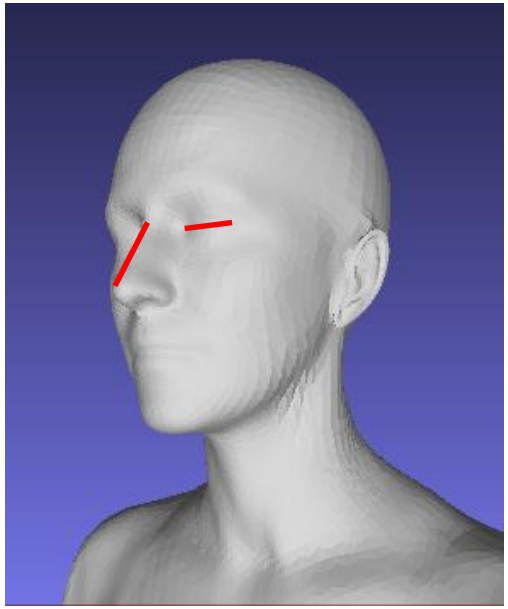
© Michael Stahl, 2015, All Rights Reserved

Ground Truth by other algorithms



© Michael Stahl, 2015, All Rights Reserved

Ground Truth by Invariants



© Michael Stahl, 2015, All Rights Reserved

Computer Vision Test Images

https://www.cs.cmu.edu/~cil/v-images.html

Ground Truth by Databases

- **AT&T Labo**
- **AVHRR Pat**
- **Air Freight** - The Air Freight data set is a
- **Amsterdam Library of Object Images** -
- **Annotated face, hand, cardiac & meat i**
- **Brown University Stimuli** - A variety of
- **CAVIAR video sequences of mall and p**
- **CCITT Fax standa** - 8 i
- **CMU CIL's Stereo Data**
- **CMU PIE Database** - A database
- **CMU VASC Image Database** - Im
- **Caltech Image Database** - about 20
- **Columbia-Utrecht Reflectance an**
- **Computational Colour Constancy**
- **Content-based image retrieval dat**
- **Densely Sampled View Spheres** -
- **Digital Embryos** - Digital embryos
- **El Salvador Atlas of Gastrointestinal**
- **FG-NET Facial Aging**
- **FVC2000**
- **Face a**
- **Germa**

Aachen Uni

- **Groning**
- **ICG Test**
- **IEN Imag** images, mos
- **INRIA's S** images database** - 15 c
- **INRIA's Syntim stereo databases** - 34 ca
- **Image Analysis Laboratory** - Images ob
- **Image Database** - An image database incl
- **JAFFE Facial Expression Image Datab**

Grayscale images.) *(ATR Research, Kyoto, J*

- **JISCT Stereo Evaluation** - 44 image pai

(SRI)

- **MIT Vision Texture** - Image archive (100+ images) (Formats: ppm)
- **MIT face images and more** - hundreds of images (Formats: homebrew) *(MIT)*
- **Machine Vision** - Images from the textbook by Jain, Kasturi, Schunck (20+ images) (Formats: GIF TIFF)
- **Mammography Image Databases** - 100 or more images of mammograms with ground truth. Additional images available by request, and links to several other mammography databases are provided. (Formats: homebrew)
- **Michigan State images** - many images (Formats: unknown)
- **Middlebury Stereo Data Sets with Ground Truth** - Six multi-frame stereo data sets of scenes containing planar regions. Each data set contains 9 color images and subpixel-accuracy ground-truth data. (Formats: ppm) *(Middlebury Stereo Vi*

...pixels). (Formats: PNG)

...variation in object recordings, we systematically varied viewing angle,

...Formats: png) *(ISIS, Intelligent Sensory Information Systems / Universi*

(Image Analysis and Computer Graphics / Technical Universit De

...behavior classification (Formats: MP

...mats: bmp)

...(Formats: tiff) *(Computation*

Content-based Retrieval Gro

of Edinburgh) *Dortmund)*

erity of Minnesota Vision La

...2000 benchmark (3520 fingerprints in all). *(Biometric Systems Lab / U*

...non-uniform daylight lighting conditions. (Formats: mpg,jpg) *(Langua*

hics and Vision / Graz University of Technology)

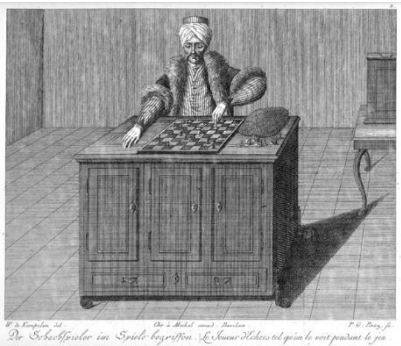
...ew) *(Image Analysis Laboratory / North Carolina State University)*

...al pose. Ratings on emotion adjectives are also available, free of charge

...rkshop paper ``The JISCT Stereo Evaluation'' by R.C.Bolles, H.H.Bake

Ground Truth by Databases

Manual creation of Ground Truth



- Crowd Sourcing
- Outsourcing
- Just Do it…

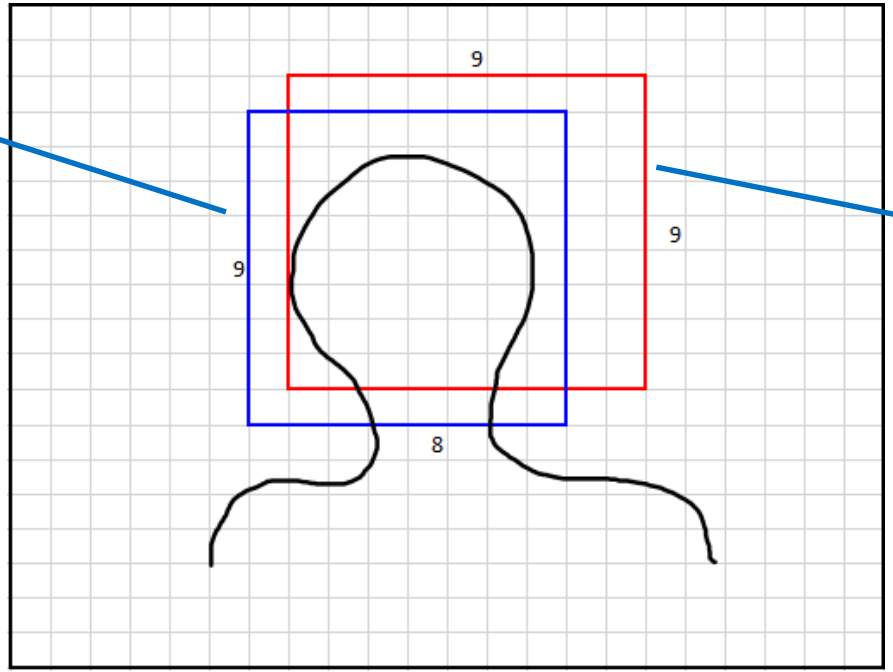
© Michael Stahl, 2015, All Rights Reserved

The metrics challenge

© Michael Stahl, 2015, All Rights Reserved

The metrics challenge

Ground Truth
(G)



Detected
(D)

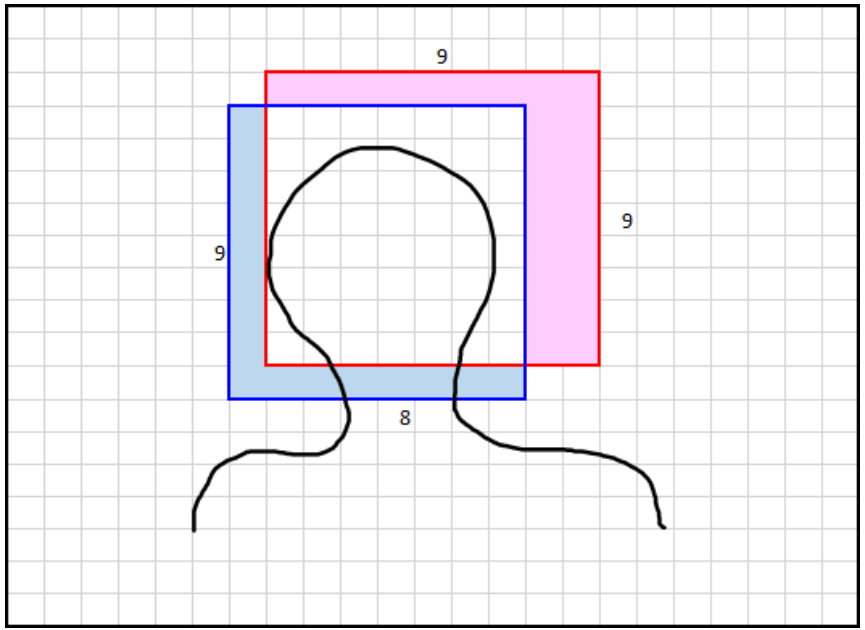
© Michael Stahl, 2015, All Rights Reserved

The metrics challenge

$G \cap D = 56$

$G \cup D =$

$8 + 8 + 2*9 + 7 + 56 = 97$



© Michael Stahl, 2015, All Rights Reserved

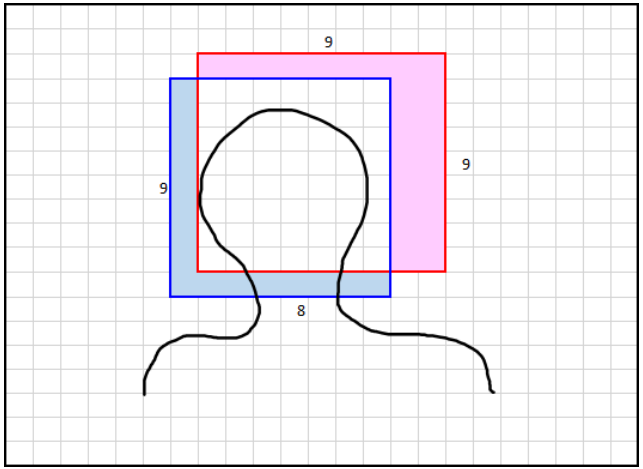
Overlap metrics

What part of D area is correct?

$$\text{Precision} = \frac{G \cap D}{D} = \frac{56}{97} = 0.58$$

What part of G area was detected?

$$\text{Recall} = \frac{G \cap D}{G} = \frac{56}{72} = 0.78$$

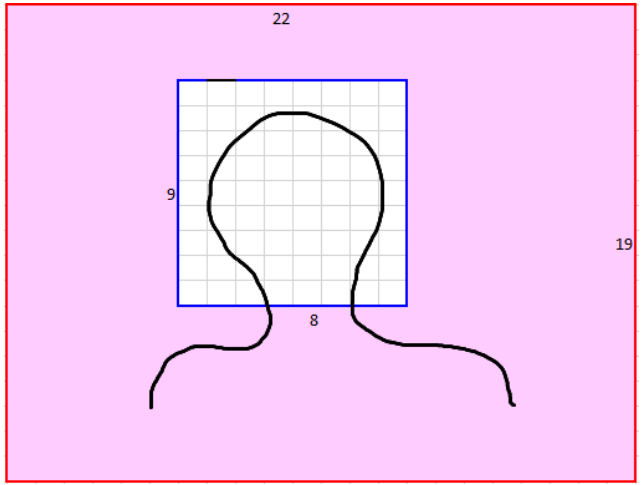


© Michael Stahl, 2015, All Rights Reserved

Extreme cases

$$\text{Precision} = \frac{G \cap D}{D} = \frac{72}{418} = 0.17$$

$$\text{Recall} = \frac{G \cap D}{G} = \frac{72}{72} = 1.0$$

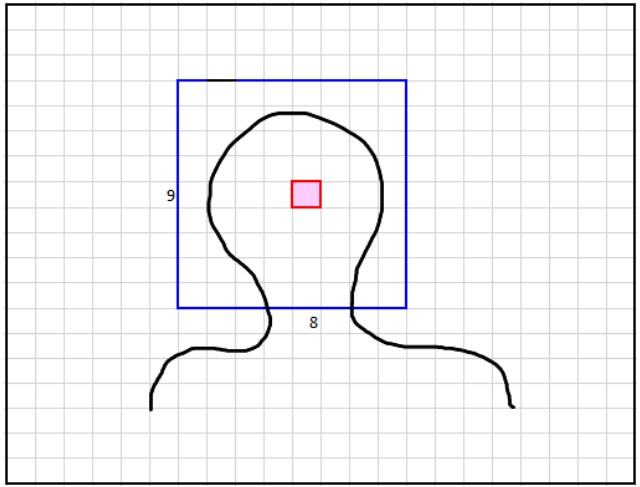


© Michael Stahl, 2015, All Rights Reserved

Extreme cases

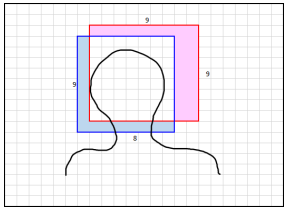
$$\text{Precision} = \frac{G \cap D}{D} = \frac{1}{1} = 1.0$$

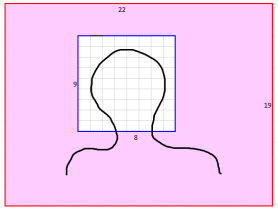
$$\text{Recall} = \frac{G \cap D}{G} = \frac{1}{72} = 0.014$$

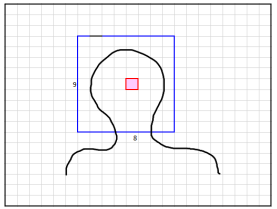


© Michael Stahl, 2015, All Rights Reserved

Combined metrics







Harmonic Mean $= 2 * \frac{\text{Precision} * \text{Recall}}{Precision + Recall}$	0.66	0.29	0.03
$Accuracy = \frac{(G \cap D) + \overline{(G \cap D)}}{Total\ image\ size}$	0.86	0.17	0.83
$\frac{G \cap D}{G \cup D}$	0.58	0.17	0.013
Etc.			

© Michael Stahl, 2015, All Rights Reserved

Choosing metric

8678387

8678387

8678387

© Michael Stahl, 2015, All Rights Reserved

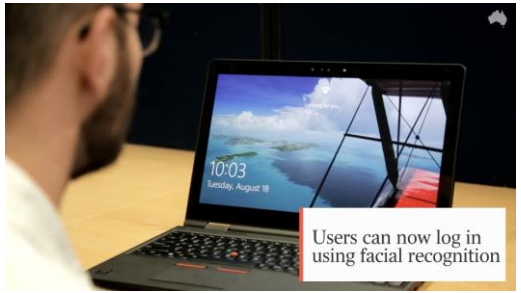
Choosing metric

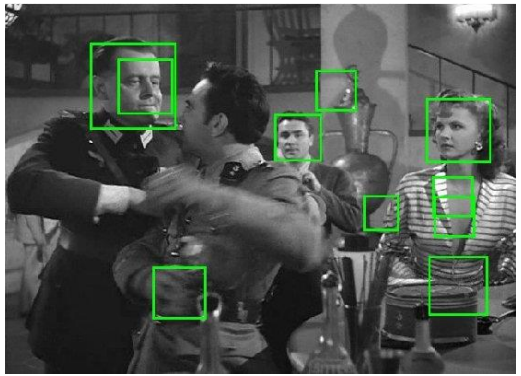
© Michael Stahl, 2015, All Rights Reserved

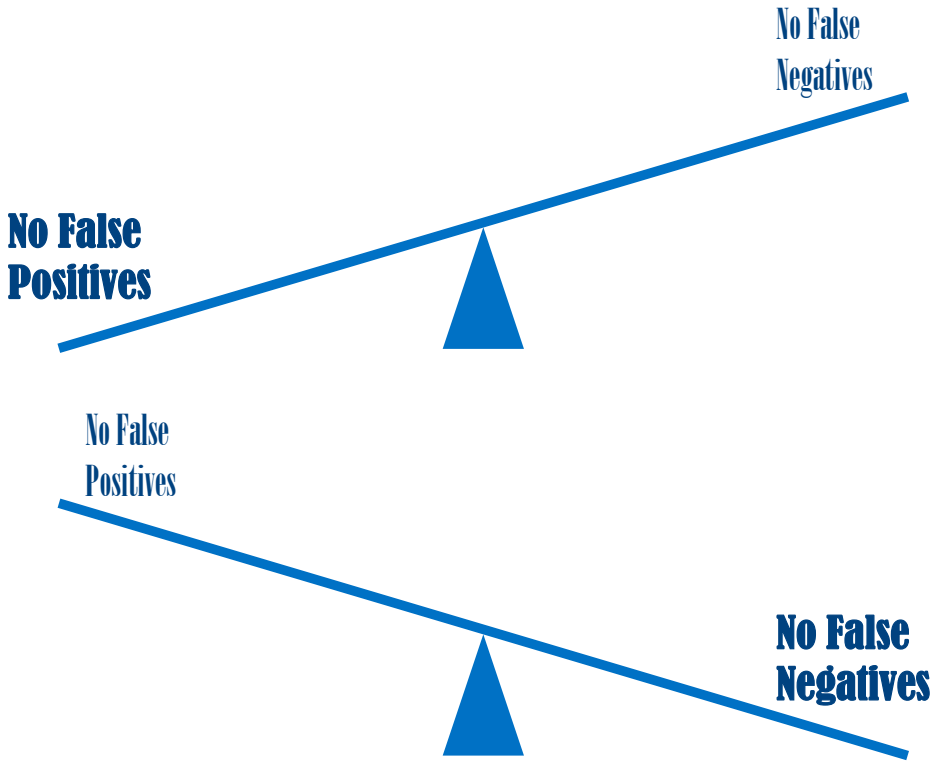
Choosing metric

© Michael Stahl, 2015, All Rights Reserved

Choosing metric







© Michael Stahl, 2015, All Rights Reserved

The metrics challenge

→ → → 30 – 60 frames per second → → → → → → → → → → → →

→ → → → → → → multiple people, multiple conditions → → → →

© Michael Stahl, 2015, All Rights Reserved

The metrics challenge

Distance	Light	Skin Tone	Movement	Angle	Hair color	Accessories	Facial Hair	Score
250cm	500lux	LightBrown	Slow	0	White	Eyeglasses	VeryLong	40
150cm	**200lux**	**Brown**	**None**	**90**	**Red**	**Earphones**	**CleanShaved**	54
50cm	50lux	Black	Medium	60	Black	Sunglasses	Long	96
150cm	**Nolight**	**White**	**Fast**	**60**	**Blond**	**Rings**	**ThinMoustache**	100
120cm	50lux	DarkBrown	Slow	40	Brown	Necklace	Stubble	77
150cm	**1000lux**	**Asian**	**Medium**	**40**	**White**	**Earings**	**ThickMoustache**	60
30cm	50lux	Olive	None	80	Blond	Hat	Short	66
300cm	**5000lux**	**LightBrown**	**Fast**	**10**	**Black**	**Earings**	**Short**	40
100cm	5000lux	Black	None	90	White	Rings	Stubble	47
50cm	**100lux**	**LightBrown**	**Medium**	**80**	**Red**	**Necklace**	**ThinMoustache**	47
100cm	100lux	Olive	Fast	30	Brown	Sunglasses	CleanShaved	84
70cm	**100lux**	**White**	**Slow**	**50**	**Black**	**Earphones**	**ThickMoustache**	74
100cm	Nolight	Asian	Slow	70	Red	Hat	Long	54
50cm	**500lux**	**White**	**Medium**	**90**	**Brown**	**Hat**	**VeryLong**	68
150cm	5000lux	Brown	Slow	10	Blond	Sunglasses	VeryLong	63
100cm	**Nolight**	**Brown**	**Medium**	**50**	**Brown**	**Eyeglasses**	**Short**	90
70cm	1000lux	DarkBrown	None	60	Red	Eyeglasses	CleanShaved	40
120cm	**200lux**	**Brown**	**Fast**	**60**	**White**	**Hat**	**ThickMoustache**	81
70cm	500lux	Black	None	80	Brown	Earings	Long	31
300cm	**200lux**	**Asian**	**Medium**	**20**	**Blond**	**Earphones**	**Stubble**	82
200cm	5000lux	DarkBrown	Fast	0	Brown	Earphones	ThinMoustache	77
250cm	**100lux**	**Asian**	**None**	**20**	**Black**	**Rings**	**CleanShaved**	51
300cm	Nolight	Olive	None	0	Red	Necklace	ThickMoustache	78
70cm	**50lux**	**Olive**	**Medium**	**10**	**White**	**Hat**	**ThinMoustache**	89
30cm	100lux	DarkBrown	Fast	20	White	Eyeglasses	Long	94

Final Score

?

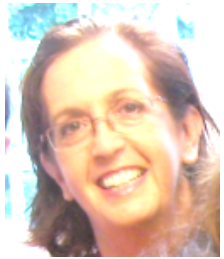
→ → → → → → → multiple people, multiple conditions → → → →

The test technique challenge

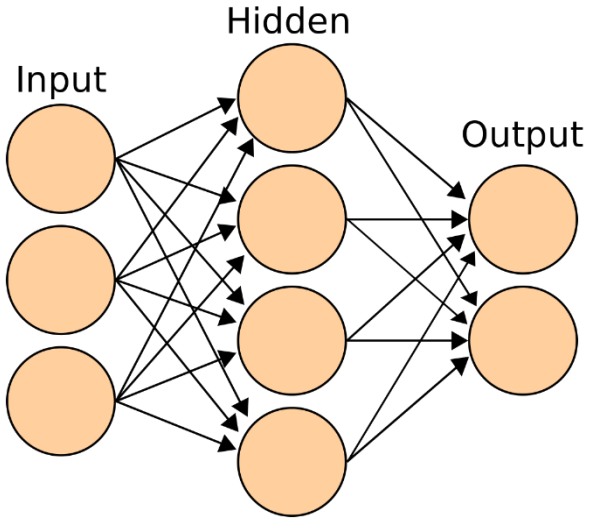
© Michael Stahl, 2015, All Rights Reserved

The test technique challenge

Valid EC



Invalid EC



The end of EC testing?

© Michael Stahl, 2015, All Rights Reserved

“make up” Equivalence Classes

Age

ROI size

Image quality

Color quality

Facial hair
Skin tone

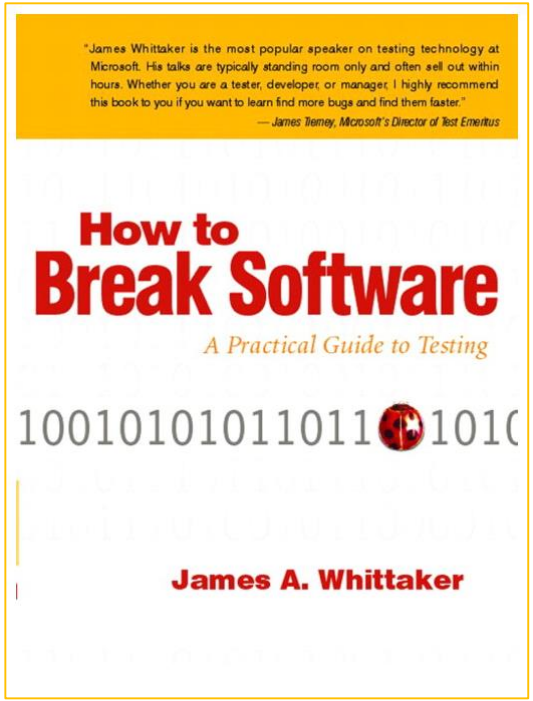
Ethnic groups

Accessories

Etc. etc.

© Michael Stahl, 2015, All Rights Reserved

Boundary Values



© Michael Stahl, 2015, All Rights Reserved

Boundary Value by Facts of Life

זה מה שצה"ל מנפק

Trial & error, common sense, known technology limitations, etc.

© Michael Stahl, 2015, All Rights Reserved

The test technique challenge

© Michael Stahl, 2015, All Rights Reserved

The combinatorial coverage challenge

© Michael Stahl, 2015, All Rights Reserved

The combinatorial coverage challenge

Distance	Light	Skin tone	Movement	Angle	Hair color	Accessories	Facial hair	...
30cm	Nolight	White	None	0	Black	None	Clean shaved	...
50cm	50lux	LightBrown	Slow	10	Brown	Eyeglasses	Stubble	...
70cm	100lux	Olive	Medium	20	Blond	Sunglasses	Short	...
100cm	200lux	Brown	Fast	30	Red	Earings	Long	...
...	500lux	DarkBrown		40	White	Necklace	VeryLong	...
400cm	1000lux	Black		50		Hat	ThinnMoustache	...
	5000lux	Asian		...		Earphones	ThickMoustache	...
				90		...		

© Michael Stahl, 2015, All Rights Reserved

The combinatorial coverage challenge

One Way coverage

Distance	Light	Skin Tone	Movement	Angle	Hair color	Accessories	Facial Hair
250cm	200lux	Asian	Slow	90	Black	Eyeglasses	Stubble
70cm	Nolight	Olive	None	10	White	Earphones	ThinMoustache
150cm	5000lux	DarkBrown	Fast	20	Brown	Hat	ThickMoustache
200cm	500lux	Brown	Medium	30	Red	Rings	CleanShaved
30cm	1000lux	Black	Fast	80	Blond	Necklace	VeryLong
100cm	100lux	LightBrown	Fast	60	Black	Earings	Short
120cm	50lux	White	Slow	0	Black	None	Long
300cm	Nolight	DarkBrown	None	70	Red	Earings	Short
50cm	500lux	White	Slow	50	Black	Earings	VeryLong
150cm	5000lux	Asian	Medium	40	Brown	Necklace	Long

© Michael Stahl, 2015, All Rights Reserved

The combinatorial coverage challenge

All-pairs

Distance	Light	Skin Tone	Movement	Angle	Hair color	Accessories	Facial Hair
250cm	500lux	LightBrown	Slow	0	White	None	VeryLong
150cm	**200lux**	**Brown**	**None**	**90**	**Red**	**Earphones**	**CleanShaved**
50cm	50lux	Black	Medium	60	Black	Eyeglasses	Long
150cm	**Nolight**	**White**	**Fast**	**60**	**Blond**	**Rings**	**ThinMoustache**
120cm	50lux	DarkBrown	Slow	40	Brown	Necklace	Stubble
150cm	**1000lux**	**Asian**	**Medium**	**40**	**White**	**Earings**	**ThickMoustache**
30cm	50lux	Olive	None	80	Blond	Hat	Short
300cm	**5000lux**	**LightBrown**	**Fast**	**10**	**Black**	**Earings**	**Short**
100cm	5000lux	Black	None	90	White	Rings	Stubble
50cm	**100lux**	**LightBrown**	**Medium**	**80**	**Red**	**Necklace**	**ThinMoustache**
100cm	100lux	Olive	Fast	30	Brown	Eyeglasses	CleanShaved
70cm	**100lux**	**White**	**Slow**	**50**	**Black**	**Earphones**	**ThickMoustache**
100cm	Nolight	Asian	Slow	70	Red	Hat	Long
50cm	**500lux**	**White**	**Medium**	**90**	**Brown**	**Hat**	**VeryLong**
150cm	5000lux	Brown	Slow	10	Blond	Eyeglasses	VeryLong
100cm	**Nolight**	**Brown**	**Medium**	**50**	**Brown**	**None**	**Short**
70cm	1000lux	DarkBrown	None	60	Red	None	CleanShaved
120cm	**200lux**	**Brown**	**Fast**	**60**	**White**	**Hat**	**ThickMoustache**

The combinatorial coverage challenge

Base Choice

Distance	Light	Skin tone	Movement	Angle	Hair color	Accessories	Facial hair	...
30cm	Nolight	White	None	0	Black	None	Clean shaved	...
50cm	50lux	LightBrown	Slow	10	Brown	Eyeglasses	Stubble	...
70cm	100lux	Olive	Medium	20	Blond	Sunglasses	Short	...
100cm	200lux	Brown	Fast	30	Red	Earings	Long	...
...	500lux	DarkBrown		40	White	Necklace	VeryLong	...
400cm	1000lux	Black		50		Hat	ThinnMoustache	...
	5000lux	Asian		...		Earphones	ThickMoustache	...
				90		...		

© Michael Stahl, 2015, All Rights Reserved

The combinatorial coverage challenge

Base Choice

Distance
30cm
50cm
70cm
100cm
...
250cm
...
250cm
250cm
250cm
250cm

Light	Skin tone	Movement	Angle	Hair color	Accessories	Facial hair	...
200lux	LightBrown	None	0	Black	None	CleanShaved	...

© Michael Stahl, 2015, All Rights Reserved

The combinatorial coverage challenge

Base Choice

Distance	Light	Skin tone	Movement	Angle	Hair color	Accessories	Facial hair	...
70cm	Nolight	LightBrown	None	0	Black	None	CleanShaved	...
	50lux							
	100lux							
	200lux							
	500lux							
	1000lux							
	5000lux							
	5000lux							
	5000lux							
	5000lux							
	5000lux							
	5000lux							
	5000lux							

© Michael Stahl, 2015, All Rights Reserved

The combinatorial coverage challenge

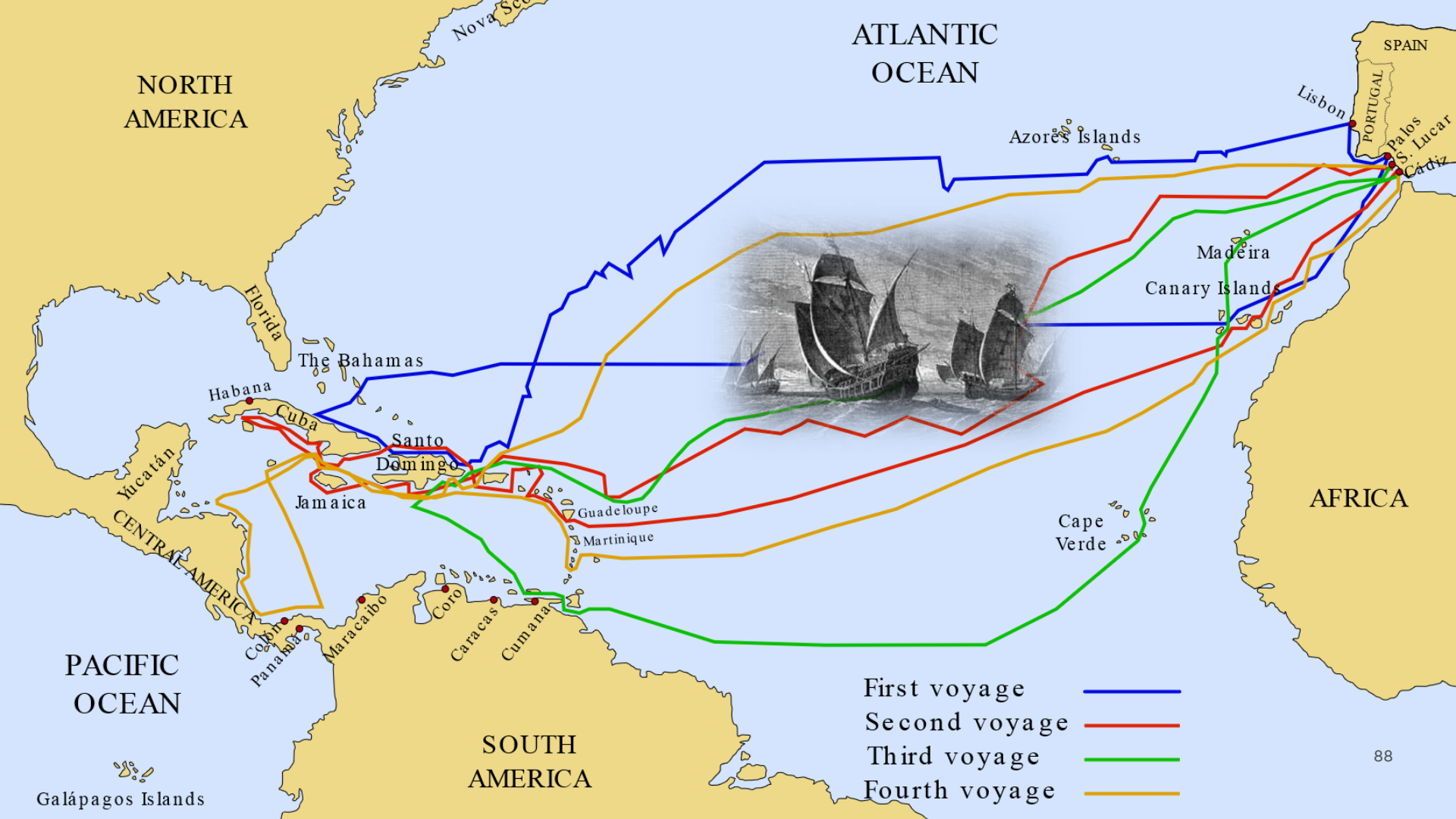
The combinatorial coverage challenge

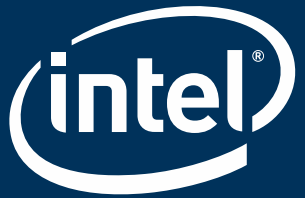
amazon fire

Summary

- Testing CV applications is a new frontier for testers
- G.T. is a huge problem
- Metrics, test techniques:
  - Lots of academic work
  - Little to No industry information
- We figure it out as we go…

intel


ATLANTIC OCEAN
NORTH AMERICA
Nova Scotia
SPAIN
PORTUGAL
Lisbon
Palos
S. Lucar
Cádiz
Azores Islands
Madeira
Canary Islands
Florida
The Bahamas
Habana
Cuba
Santo Domingo
Jamaica
Yucatán
CENTRAL AMERICA
Guadeloupe
Martinique
Cape Verde
AFRICA
Colón
Panamá
Maracaibo
Coro
Caracas
Cumaná
PACIFIC OCEAN
SOUTH AMERICA
Galápagos Islands
First voyage
Second voyage
Third voyage
Fourth voyage



THANK YOU!