Face Recognition II

COMP 388-002/488-002 Biometrics

Daniel Moreira

Fall 2023

LOYOLA UNIVERSITY CHICAGO

Today you will...

Get to know
Face acquisition and enhancement.

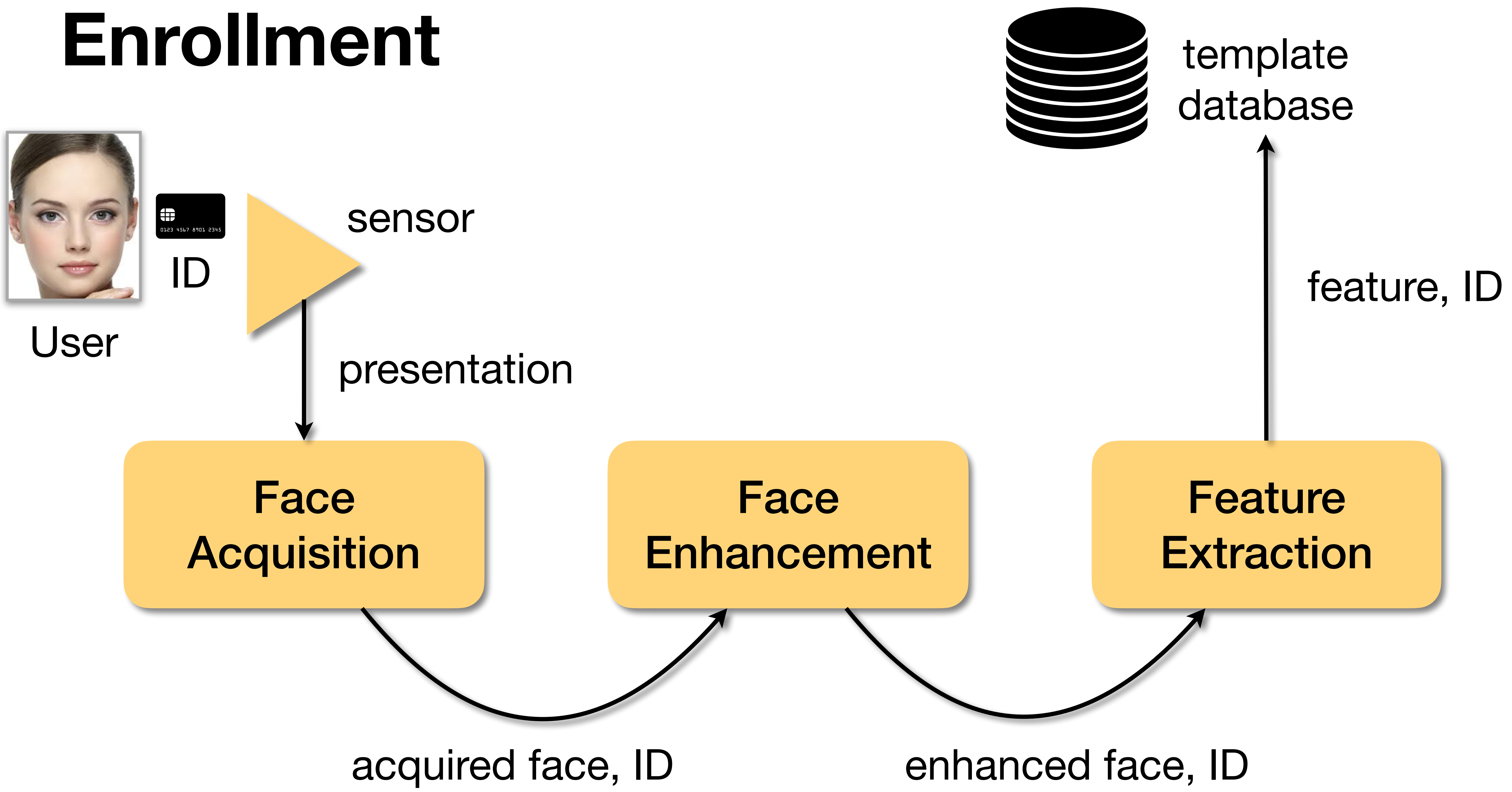
Today's attendance

Please fill out the form

https://forms.gle/zQF51qPNPc4gVdfi8

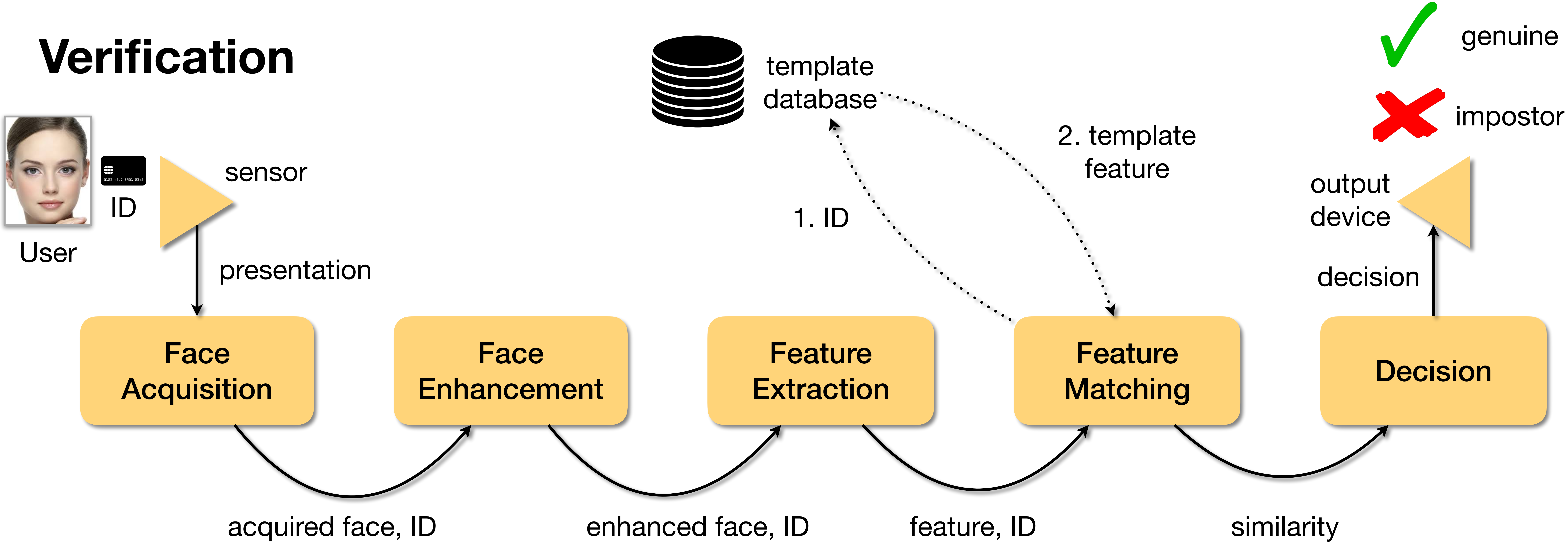
Face Recognition

Enrollment



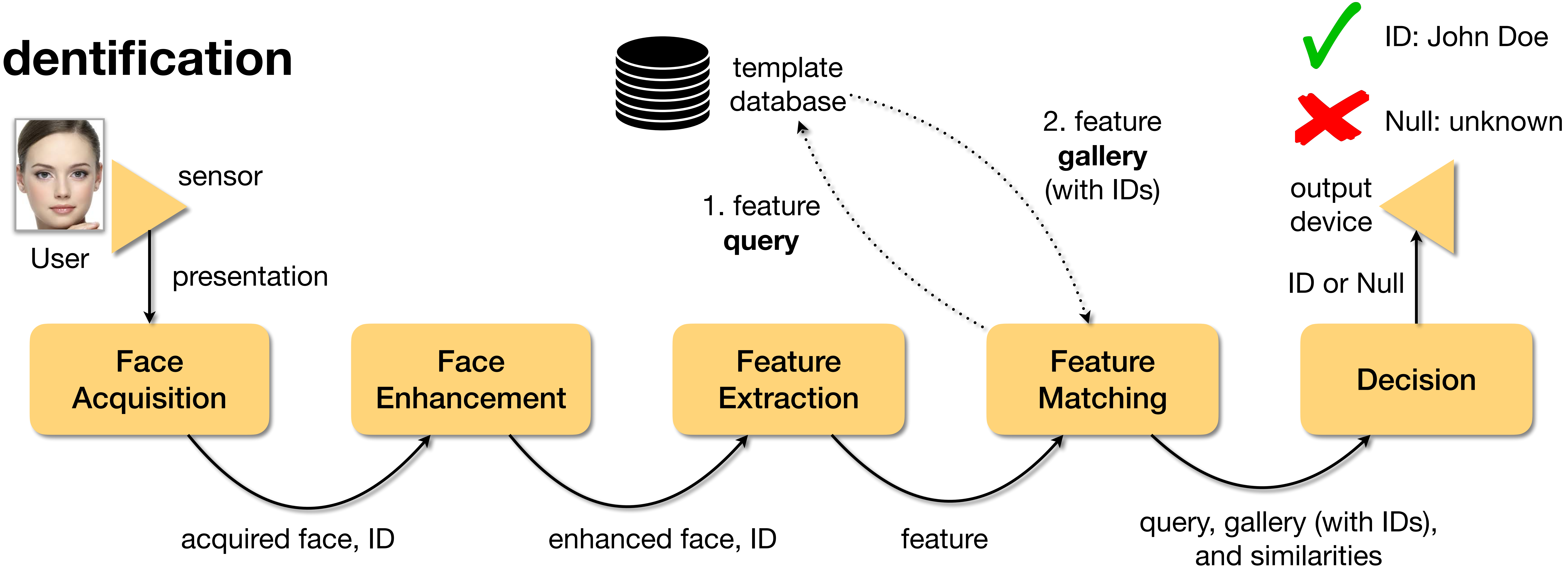
Face Recognition

Verification



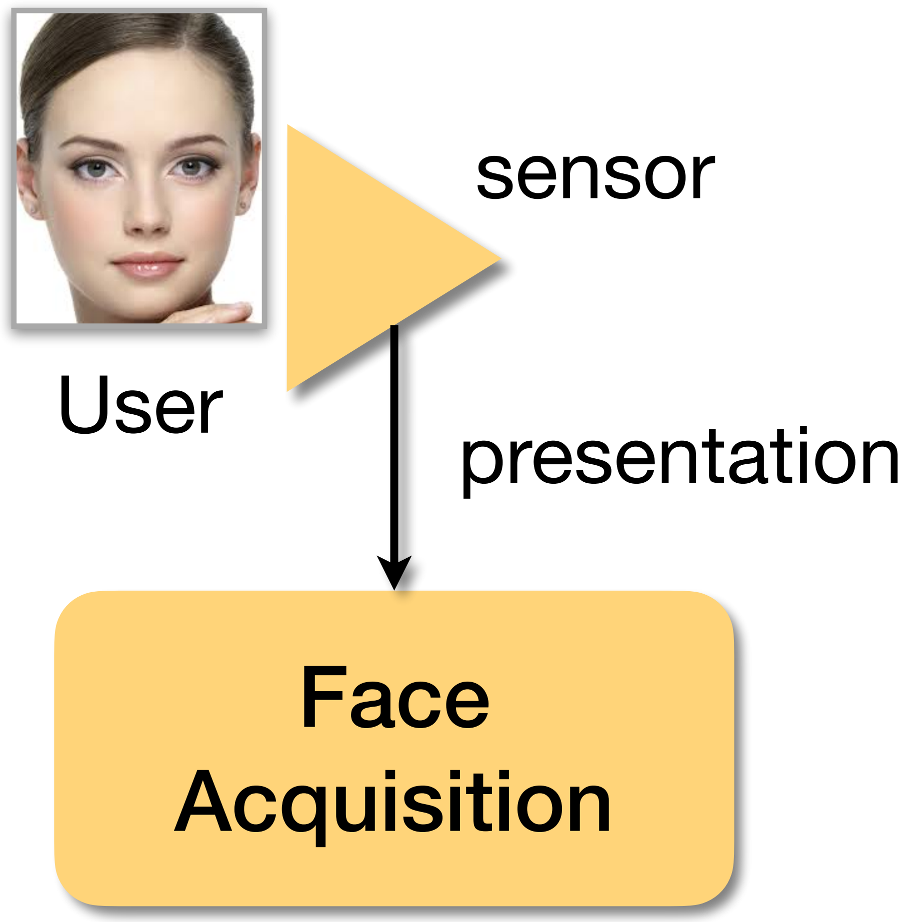
Face Recognition

Identification



LOYOLA UNIVERSITY CHICAGO

Face Recognition



LOYOLA UNIVERSITY CHICAGO

Acquisition

On-line versus Off-line

https://www.youtube.com/watch?v=BYN4oF_bi4c

Acquisition

Controlled Acquisition
Right pose, distance and illumination.

https://www.youtube.com/watch?v=BYN4oF_bi4c

https://www.youtube.com/watch?v=-cjoJR3oWcQ

LOYOLA UNIVERSITY CHICAGO

Acquisition

Controlled Acquisition
Different light wavelengths.

Jain, Ross, and Nadakumar
Introduction to Biometrics
Springer Books, 2011

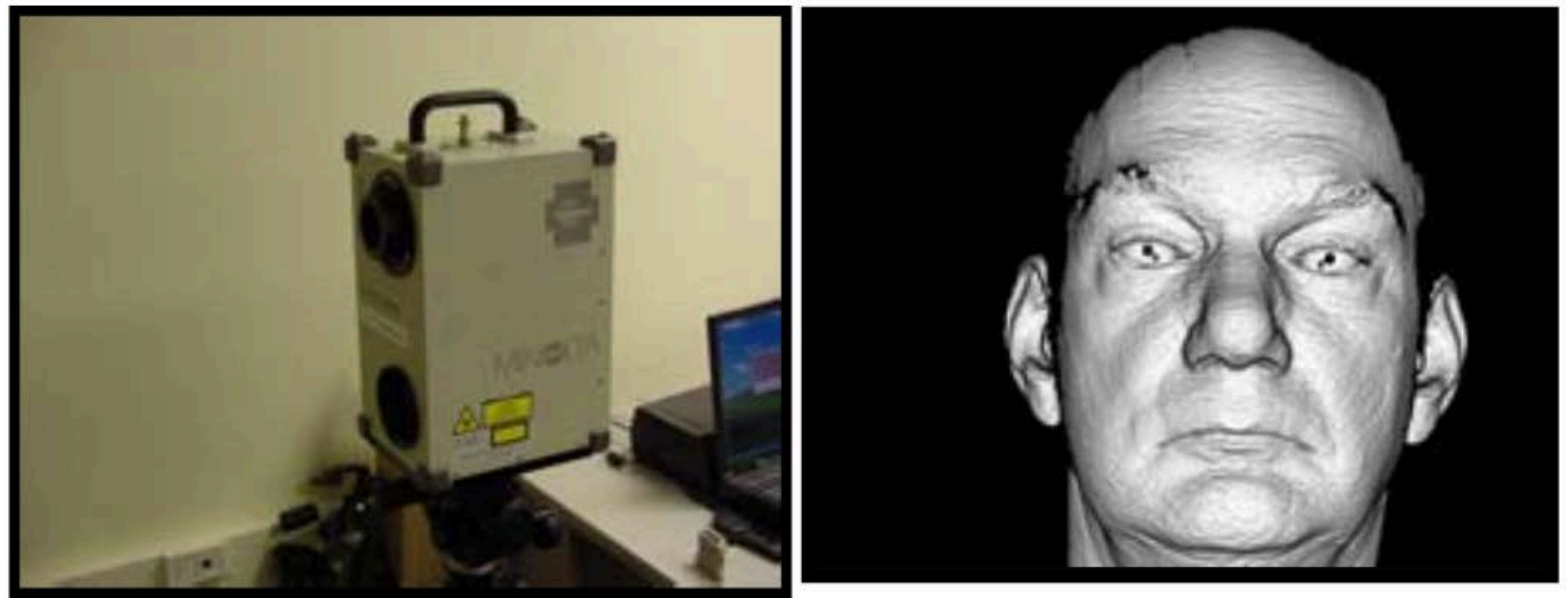
Captures at visible and near-infrared spectra.

Sony infrared camera.

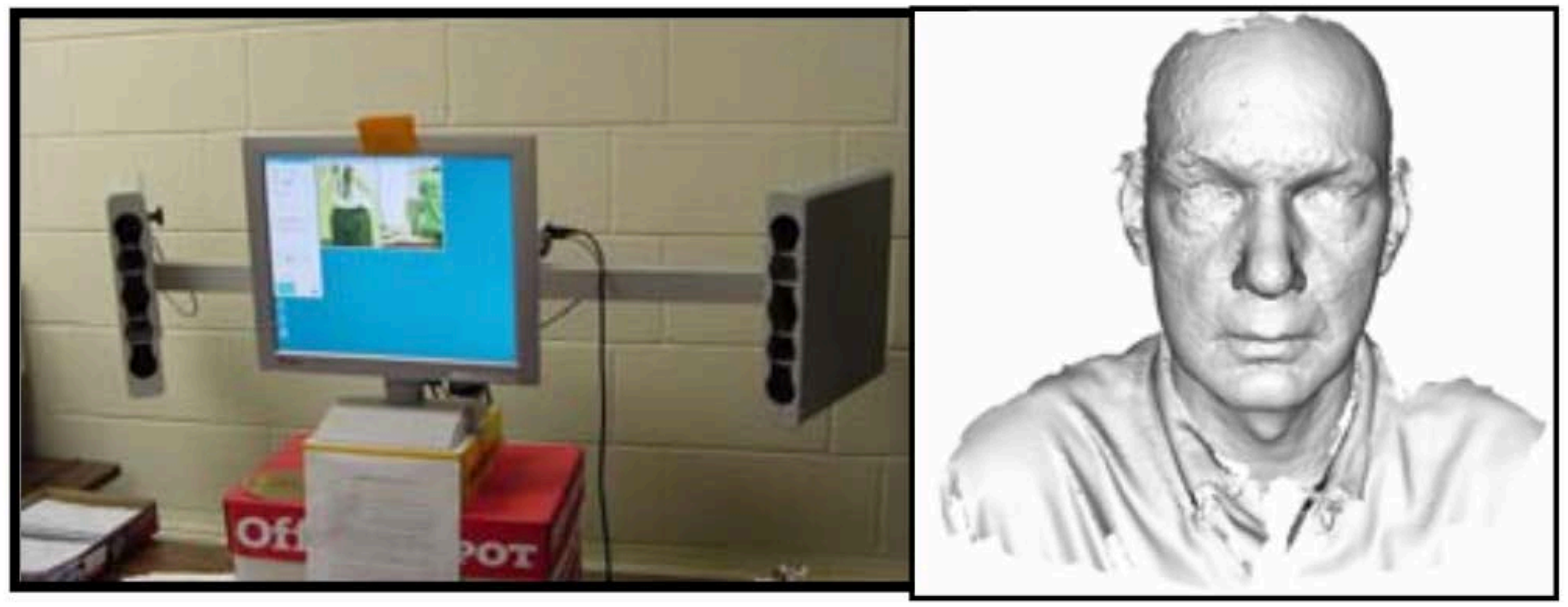
Acquisition

Controlled Acquisition
3D Information

Source:
Dr. Walter Scheirer



Minolta Vivid 900/910



3DMD “Qlonerator”

Acquisition

Unconstrained Acquisition

No illumination control.

https://www.nist.gov/system/files/documents/itl/iad/ig/05771424.pdf

Acquisition

Unconstrained Acquisition

No distance control.

Jain, Ross, and Nadakumar
Introduction to Biometrics
Springer Books, 2011

1m

3m

5m

Acquisition

Unconstrained Acquisition

No pose control.

Hsu
Face detection and modeling for recognition
PhD Thesis, MSU, 2002.

Acquisition

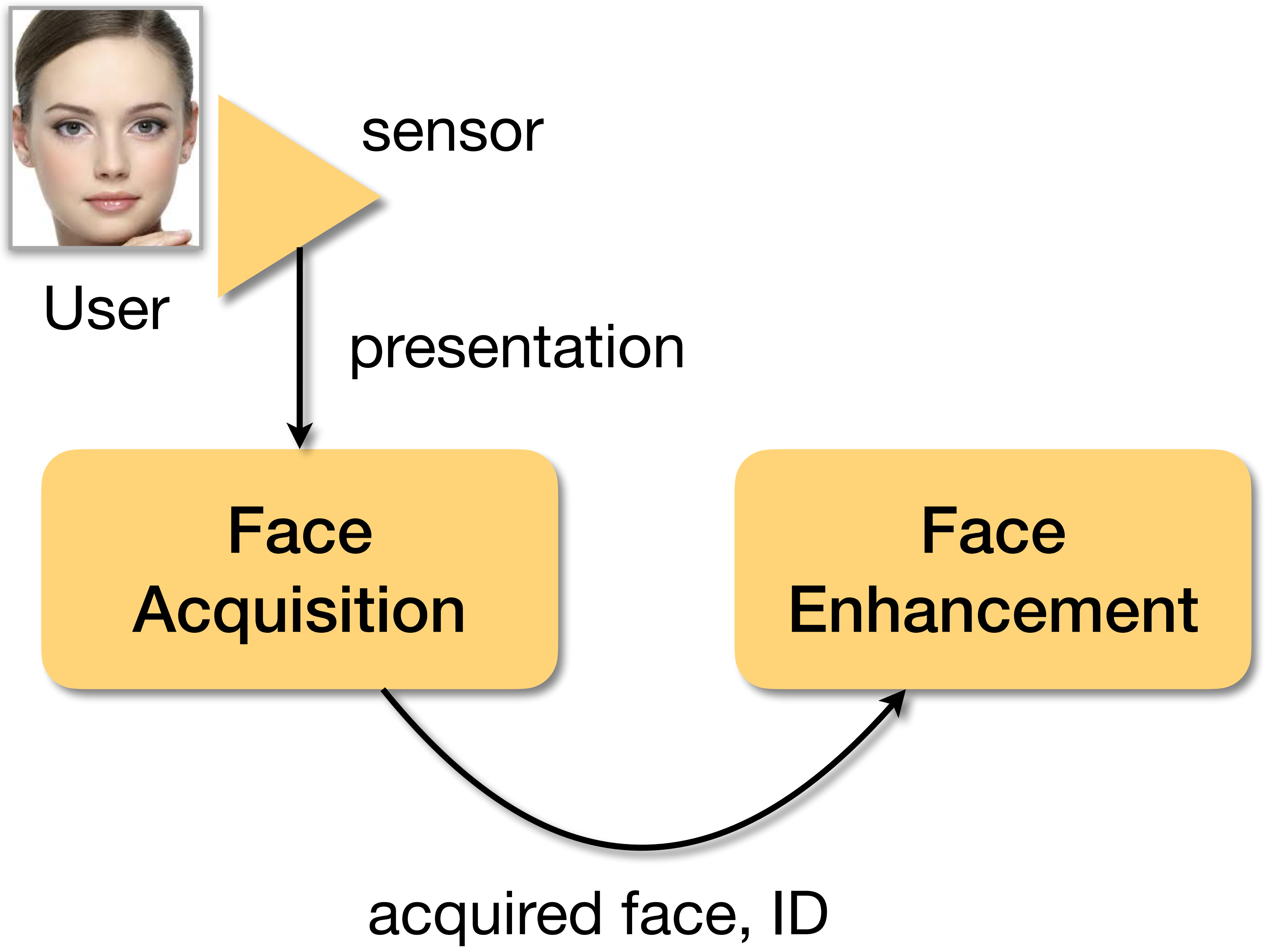
Problems

Presentation Attack

https://www.youtube.com/watch?v=BGgQ9woZQOg

LOYOLA UNIVERSITY CHICAGO

Face Recognition

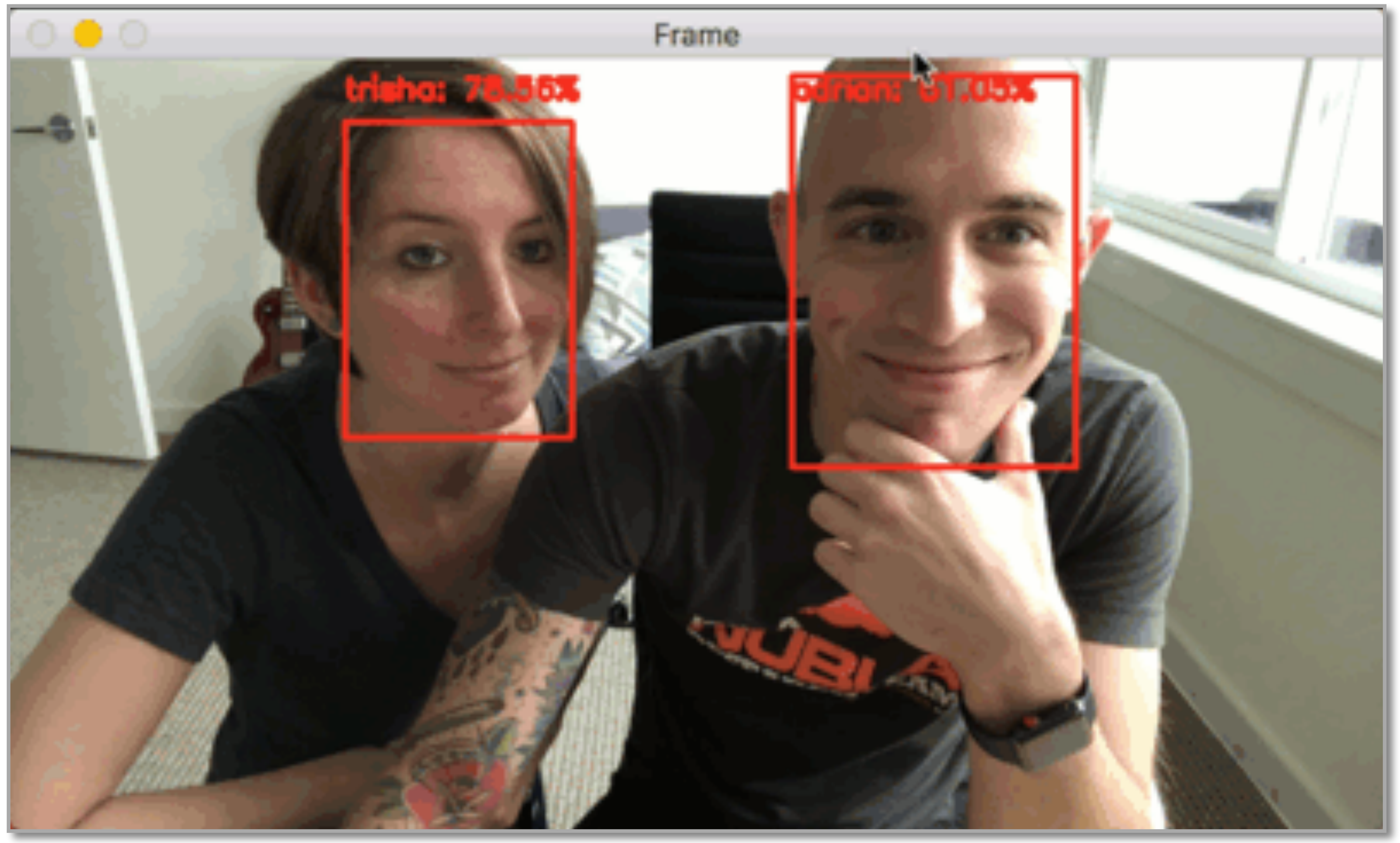


Enhancement

Face Detection

Goal

Localize faces for segmentation and further recognition.



https://www.pyimagesearch.com/2018/09/24/opencv-face-recognition/

Enhancement

Face Detection

Challenges
Megapixel image
Nearly millions of possible locations and scales combined.

False positives should be below 1 in 1 million.

Source: Hu et al., *Finding Tiny Faces*, 2016 (https://arxiv.org/abs/1612.04402)

LOYOLA UNIVERSITY CHICAGO

Enhancement

Face Detection

State of the Art
Megapixel image
Nearly millions of possible locations, scales,
and poses combined.
Detection and pose estimation.

Source: Albiero et al.
img2pose: Face Alignment and Detection via 6DoF, Face Pose Estimation
2021 (https://arxiv.org/abs/2012.07791)

Available at
https://github.com/vitoralbiero/img2pose

LOYOLA UNIVERSITY CHICAGO

Enhancement

Face Detection

Methods
Either based on *sliding windows* or on *regions of interest*.

LOYOLA UNIVERSITY CHICAGO

Enhancement

Face Detection

Sliding Windows
Scans of the image with windows of different scales.

LOYOLA UNIVERSITY CHICAGO

Enhancement

Face Detection

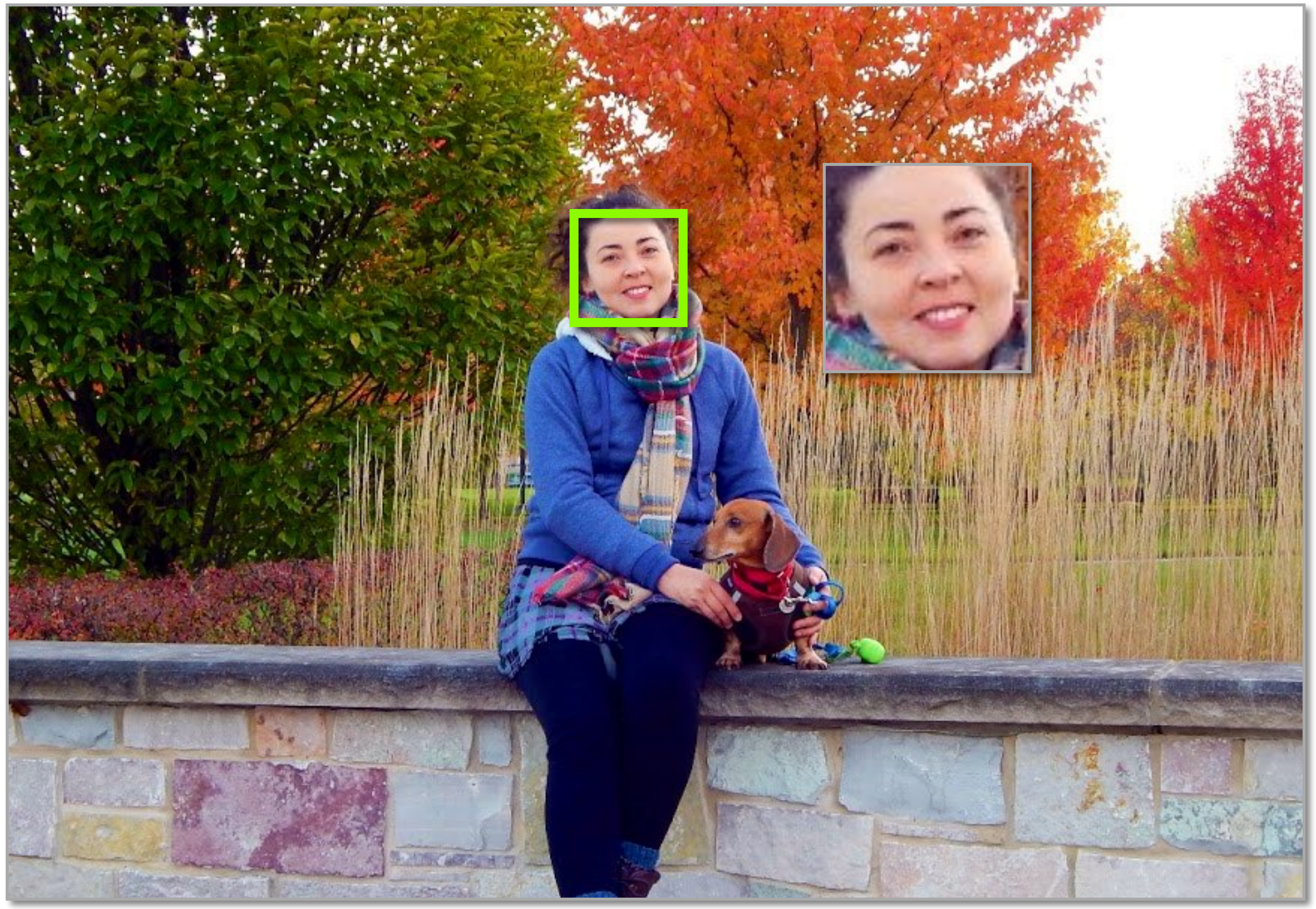
Sliding Windows
Scans of the image with windows of different scales.

LOYOLA UNIVERSITY CHICAGO

Enhancement

Face Detection

Sliding Windows
Scans of the image with windows of different scales.



Enhancement

Face Detection

Regions of Interest
Techniques from Computer Vision or Machine Learning to segment regions.

E.g., Maximally Stable Extremal Regions (MSER[1]) or Deep Local Features (DELF[2]).

1. Matas et al. *Robust Wide Baseline Stereo from Maximally Stable Extremal Regions*. BMVC 2002.
2. Noh et al. *Large-Scale Image Retrieval with Attentive Deep Local Features*. ICCV 2017.

LOYOLA UNIVERSITY CHICAGO

Enhancement

Face Detection

Regions of Interest
Techniques from Machine Learning to classify each region as *face or non-face.*

E.g., Support Vector Machines (SVM).

LOYOLA
UNIVERSITY CHICAGO

Enhancement

Face Detection

Viola-Jones Detector
First real-time face detector.
Based on sliding windows.

Key Ideas (4)
Haar-like features.
Integral image.
Boosting for feature selection.
Attentional Cascade to reject non-faces.

SECOND INTERNATIONAL WORKSHOP ON STATISTICAL AND COMPUTATIONAL THEORIES OF VISION – MODELING, LEARNING, COMPUTING, AND SAMPLING

VANCOUVER, CANADA, JULY 13, 2001.

Robust Real-time Object Detection

Paul Viola
viola@merl.com
Mitsubishi Electric Research Labs
201 Broadway, 8th FL
Cambridge, MA 02139

Michael Jones
mjones@crl.dec.com
Compaq CRL
One Cambridge Center
Cambridge, MA 02142

Abstract

This paper describes a visual object detection framework that is capable of processing images extremely rapidly while achieving high detection rates. There are three key contributions. The first is the introduction of a new image representation called the "Integral Image" which allows the features used by our detector to be computed very quickly. The second is a learning algorithm, based on AdaBoost, which selects a small number of critical visual features and yields extremely efficient classifiers [6]. The third contribution is a method for combining classifiers in a "cascade" which allows background regions of the image to be quickly discarded while spending more computation on promising object-like regions. A set of experiments in the domain of face detection are presented. The system yields face detection performace comparable to the best previous systems [18, 13, 16, 12, 1]. Implemented on a conventional desktop, face detection proceeds at 15 frames per second.

1. Introduction

This paper brings together new algorithms and insights to construct a framework for robust and extremely rapid object detection. This framework is demonstrated on, and in part motivated by, the task of face detection. Toward this end we have constructed a frontal face detection system which achieves detection and false positive rates which are equivalent to the best published results [18, 13, 16, 12, 1]. This face detection system is most clearly distinguished from previous approaches in its ability to detect faces extremely rapidly. Operating on 384 by 288 pixel images, faces are detected at 15 frames per second on a conventional 700

1

LOYOLA UNIVERSITY CHICAGO

Enhancement

Face Detection

Viola-Jones Detector
First real-time face detector.
Based on sliding windows.

Key Ideas (4)
Haar-like features.
Integral image.
Boosting for feature selection.
Attentional Cascade to reject non-faces.

SECOND INTERNATIONAL WORKSHOP ON STATISTICAL AND COMPUTATIONAL THEORIES OF VISION – MODELING, LEARNING, COMPUTING, AND SAMPLING

VANCOUVER, CANADA, JULY 13, 2001.

Robust Real-time Object Detection

Paul Viola
viola@merl.com
Mitsubishi Electric Research Labs
201 Broadway, 8th FL
Cambridge, MA 02139

Michael Jones
mjones@crl.dec.com
Compaq CRL
One Cambridge Center
Cambridge, MA 02142

Abstract

This paper describes a visual object detection framework that is capable of processing images extremely rapidly while achieving high detection rates. There are three key contributions. The first is the introduction of a new image representation called the "Integral Image" which allows the features used by our detector to be computed very quickly. The second is a learning algorithm, based on AdaBoost, which selects a small number of critical visual features and yields extremely efficient classifiers [6]. The third contribution is a method for combining classifiers in a "cascade" which allows background regions of the image to be quickly discarded while spending more computation on promising object-like regions. A set of experiments in the domain of face detection are presented. The system yields face detection performace comparable to the best previous systems [18, 13, 16, 12, 1]. Implemented on a conventional desktop, face detection proceeds at 15 frames per second.

1. Introduction

This paper brings together new algorithms and insights to construct a framework for robust and extremely rapid object detection. This framework is demonstrated on, and in part motivated by, the task of face detection. Toward this end we have constructed a frontal face detection system which achieves detection and false positive rates which are equivalent to the best published results [18, 13, 16, 12, 1]. This face detection system is most clearly distinguished from previous approaches in its ability to detect faces extremely rapidly. Operating on 384 by 288 pixel images, faces are detected at 15 frames per second on a conventional 700

1

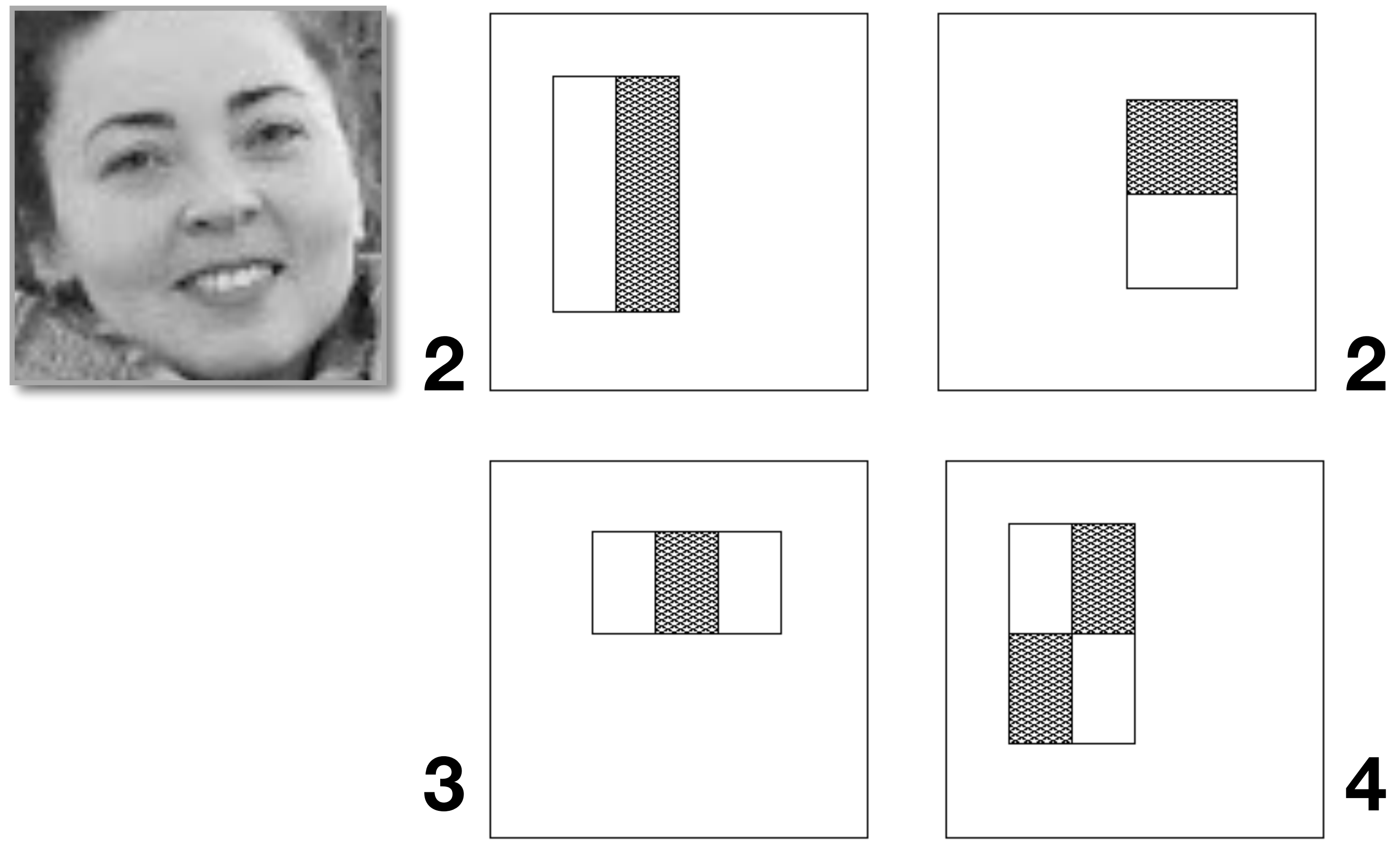
Enhancement

Viola-Jones Detector

Haar-Like Features (1/4)
Binary rectangle filters used to extract features from the sliding window.

$$value = \sum pixels\ in\ white\ area - \sum pixels\ in\ black\ area$$

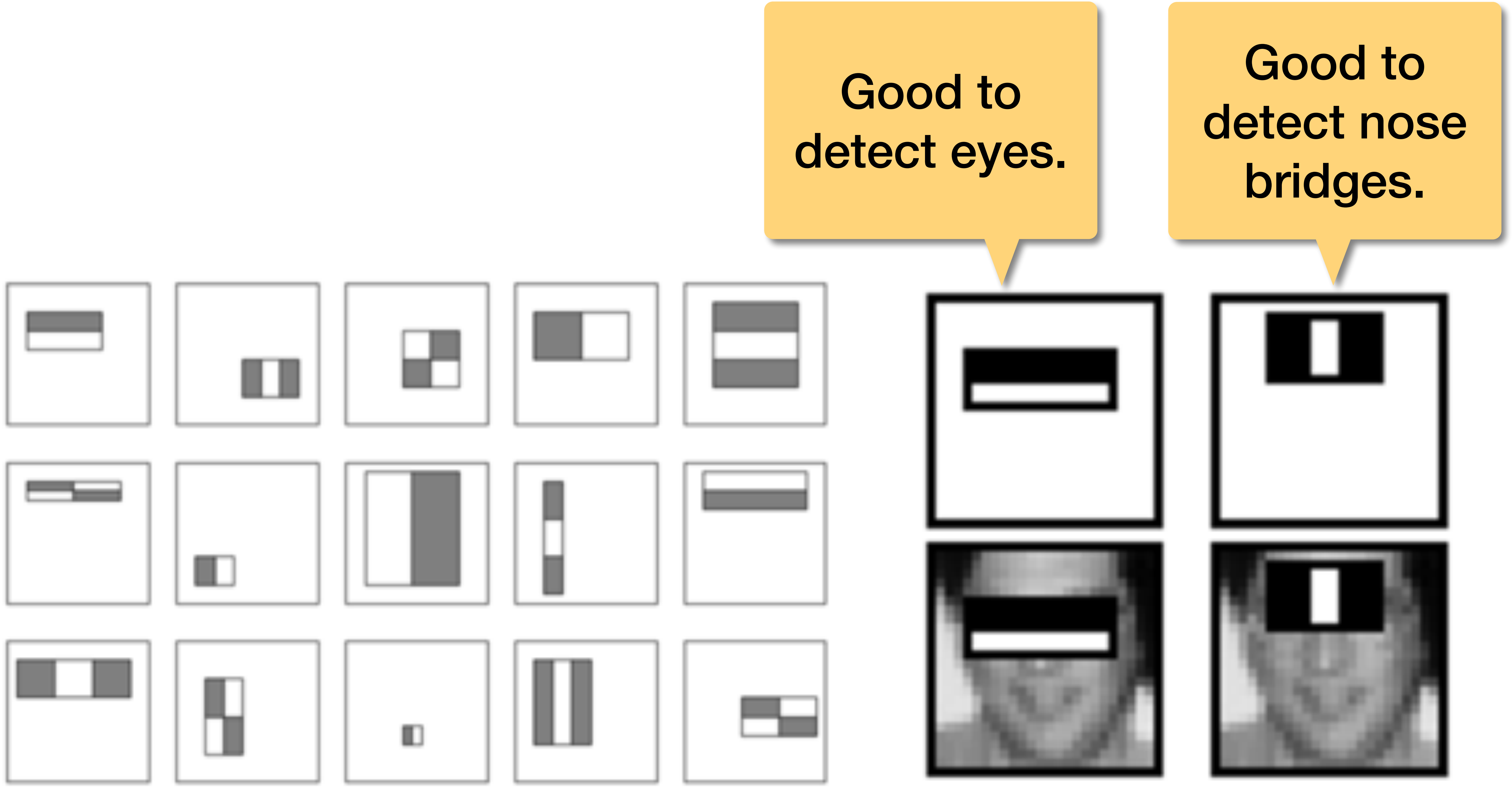
Filter types
2, 3, and 4 rectangles.



Enhancement

Viola-Jones Detector

Haar-Like Features (1/4)
Take a 24-by-24-pixels window.
The number of possible features is nearly 160,000.



How to apply and how to select features fast?

Enhancement

Face Detection

Viola-Jones Detector
First real-time face detector.
Based on sliding windows.

Key Ideas (4)
Haar-like features.
Integral image.
Boosting for feature selection.
Attentional Cascade to reject non-faces.


SECOND INTERNATIONAL WORKSHOP ON STATISTICAL AND COMPUTATIONAL THEORIES OF VISION – MODELING, LEARNING, COMPUTING, AND SAMPLING

VANCOUVER, CANADA, JULY 13, 2001.

Robust Real-time Object Detection

Paul Viola
viola@merl.com
Mitsubishi Electric Research Labs
201 Broadway, 8th FL
Cambridge, MA 02139

Michael Jones
mjones@crl.dec.com
Compaq CRL
One Cambridge Center
Cambridge, MA 02142

Abstract

This paper describes a visual object detection framework that is capable of processing images extremely rapidly while achieving high detection rates. There are three key contributions. The first is the introduction of a new image representation called the "Integral Image" which allows the features used by our detector to be computed very quickly. The second is a learning algorithm, based on AdaBoost, which selects a small number of critical visual features and yields extremely efficient classifiers [6]. The third contribution is a method for combining classifiers in a "cascade" which allows background regions of the image to be quickly discarded while spending more computation on promising object-like regions. A set of experiments in the domain of face detection are presented. The system yields face detection performace comparable to the best previous systems [18, 13, 16, 12, 1]. Implemented on a conventional desktop, face detection proceeds at 15 frames per second.

1. Introduction

This paper brings together new algorithms and insights to construct a framework for robust and extremely rapid object detection. This framework is demonstrated on, and in part motivated by, the task of face detection. Toward this end we have constructed a frontal face detection system which achieves detection and false positive rates which are equivalent to the best published results [18, 13, 16, 12, 1]. This face detection system is most clearly distinguished from previous approaches in its ability to detect faces extremely rapidly. Operating on 384 by 288 pixel images, faces are detected at 15 frames per second on a conventional 700

1

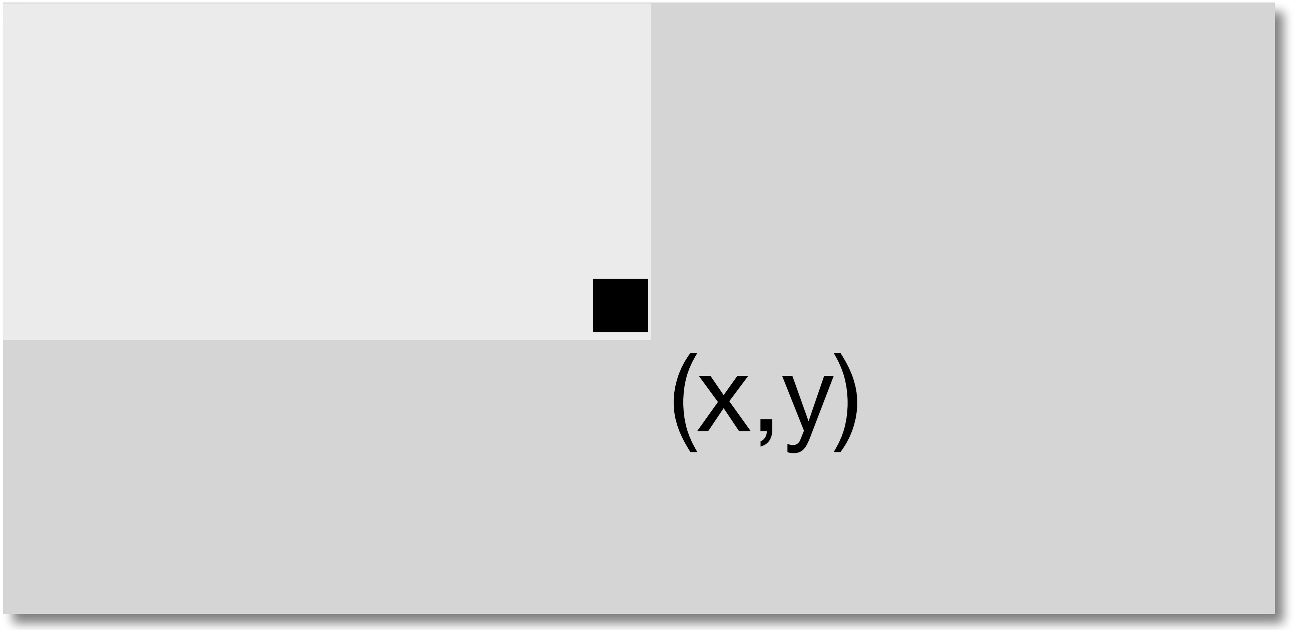
Enhancement

Viola-Jones Detector

Integral Image (2/4)
Solution to apply Haar-like features fast.

Precomputed data structure with the same dimensions of the target image.

Target Image



Integral Image

Enhancement

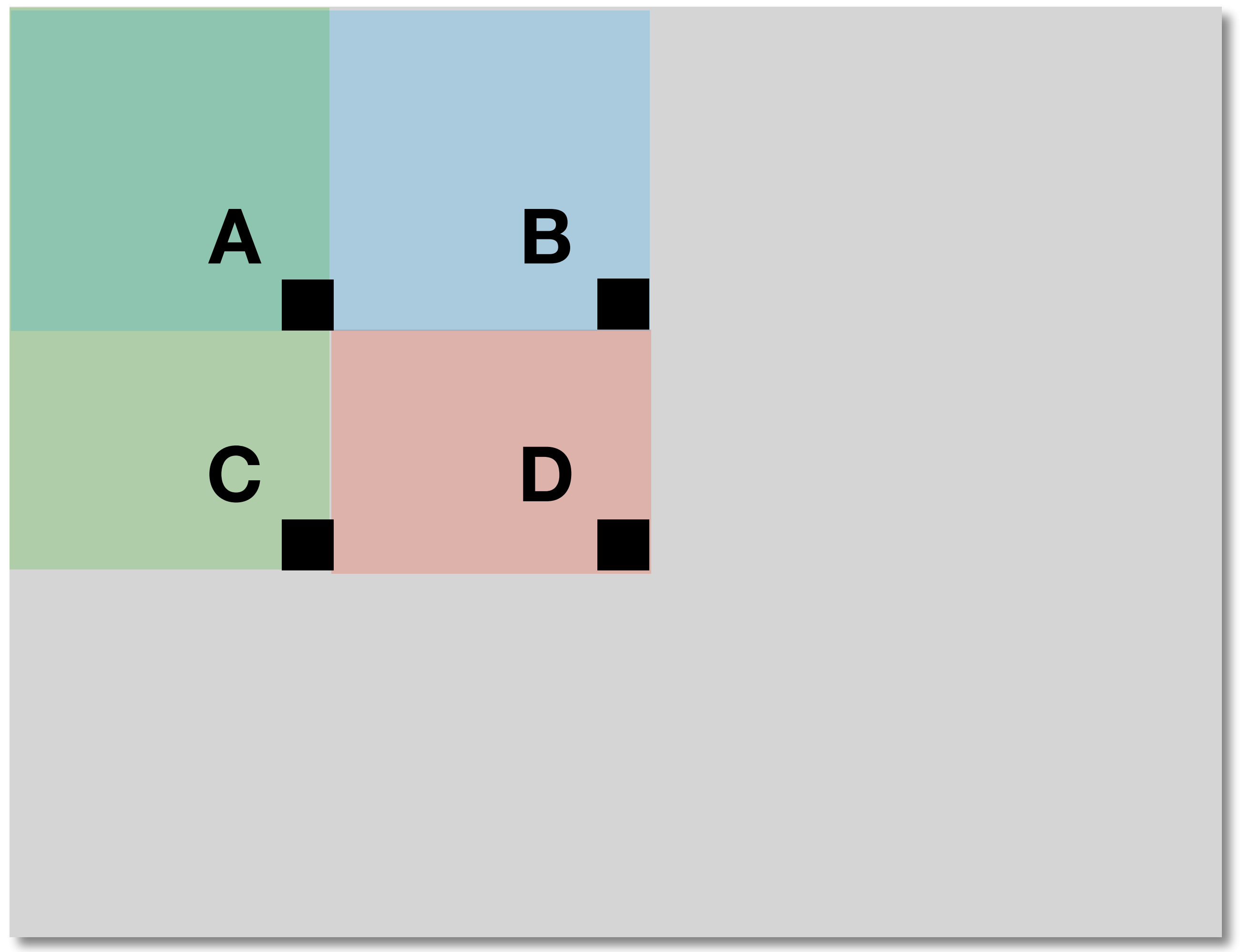
Viola-Jones Detector

Integral Image (2/4)

Remember Haar feature *value*:

$$value = \sum pixels\ in\ white\ area - \sum pixels\ in\ black\ area$$

Integral images allow the computation of the sum of pixel values in any target area in constant time, regardless of the size of the area.



Sum of pixels in red area

$$content = D - B - C + A$$

Only and always 4 accesses.

Enhancement

Face Detection

Viola-Jones Detector
First real-time face detector.
Based on sliding windows.

Key Ideas (4)
Haar-like features.
Integral image.
Boosting for feature selection.
Attentional Cascade to reject non-faces.

SECOND INTERNATIONAL WORKSHOP ON STATISTICAL AND COMPUTATIONAL THEORIES OF VISION – MODELING, LEARNING, COMPUTING, AND SAMPLING

VANCOUVER, CANADA, JULY 13, 2001.

Robust Real-time Object Detection

Paul Viola	Michael Jones
viola@merl.com	mjones@crl.dec.com
Mitsubishi Electric Research Labs	Compaq CRL
201 Broadway, 8th FL	One Cambridge Center
Cambridge, MA 02139	Cambridge, MA 02142

Abstract

This paper describes a visual object detection framework that is capable of processing images extremely rapidly while achieving high detection rates. There are three key contributions. The first is the introduction of a new image representation called the "Integral Image" which allows the features used by our detector to be computed very quickly. The second is a learning algorithm, based on AdaBoost, which selects a small number of critical visual features and yields extremely efficient classifiers [6]. The third contribution is a method for combining classifiers in a "cascade" which allows background regions of the image to be quickly discarded while spending more computation on promising object-like regions. A set of experiments in the domain of face detection are presented. The system yields face detection performace comparable to the best previous systems [18, 13, 16, 12, 1]. Implemented on a conventional desktop, face detection proceeds at 15 frames per second.

1. Introduction

This paper brings together new algorithms and insights to construct a framework for robust and extremely rapid object detection. This framework is demonstrated on, and in part motivated by, the task of face detection. Toward this end we have constructed a frontal face detection system which achieves detection and false positive rates which are equivalent to the best published results [18, 13, 16, 12, 1]. This face detection system is most clearly distinguished from previous approaches in its ability to detect faces extremely rapidly. Operating on 384 by 288 pixel images, faces are detected at 15 frames per second on a conventional 700

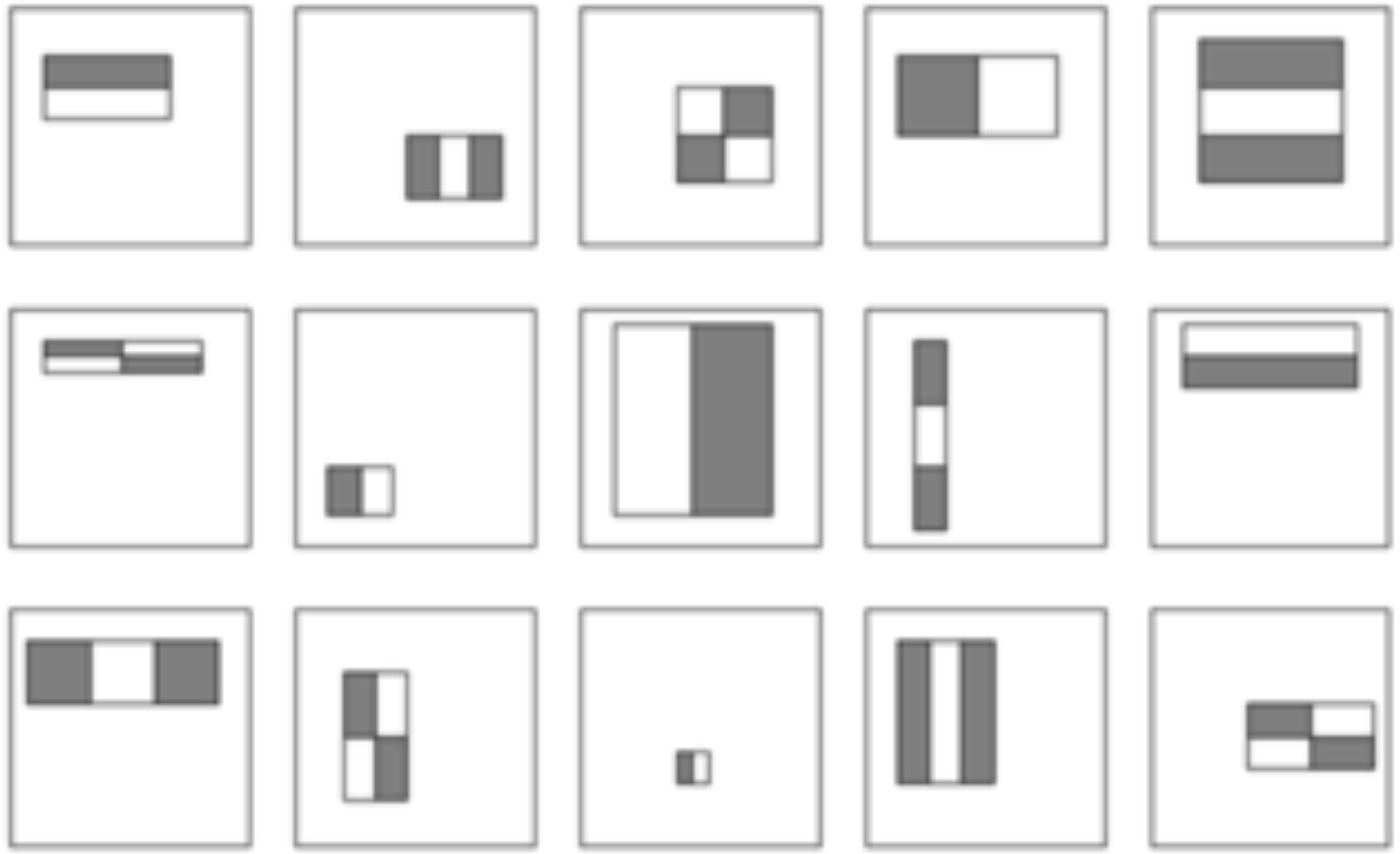
1

LOYOLA UNIVERSITY CHICAGO

Enhancement

Viola-Jones Detector

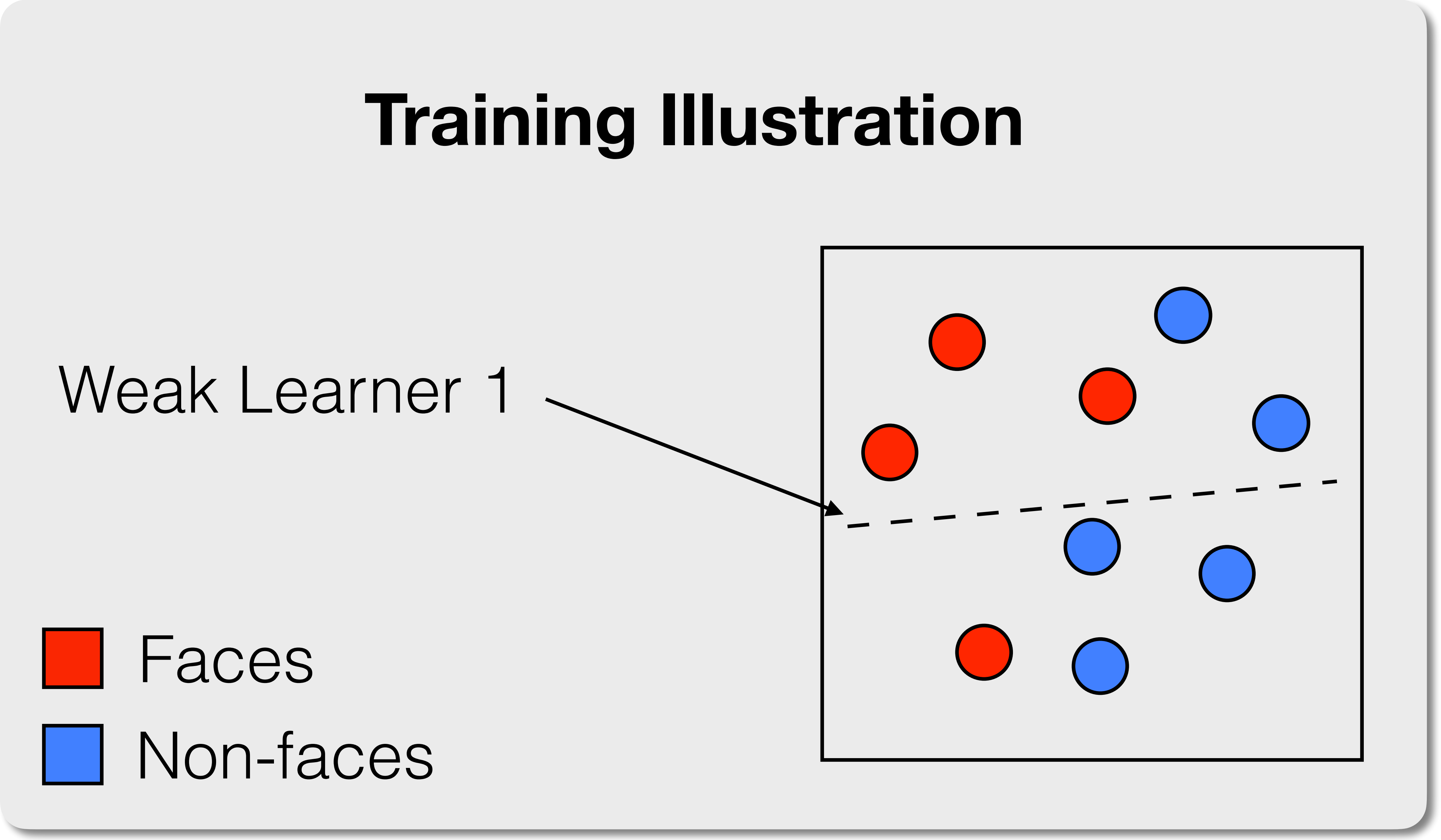
Boosting for
Feature Selection (3/4)
Goal: select combinations of Haar-like features that are useful for face detection.



Enhancement

Viola-Jones Detector

Boosting for
Feature Selection (3/4)
Solution: *boosting*, a combination of weak classifiers that when learned in sequence and applied together, lead to better final classification.



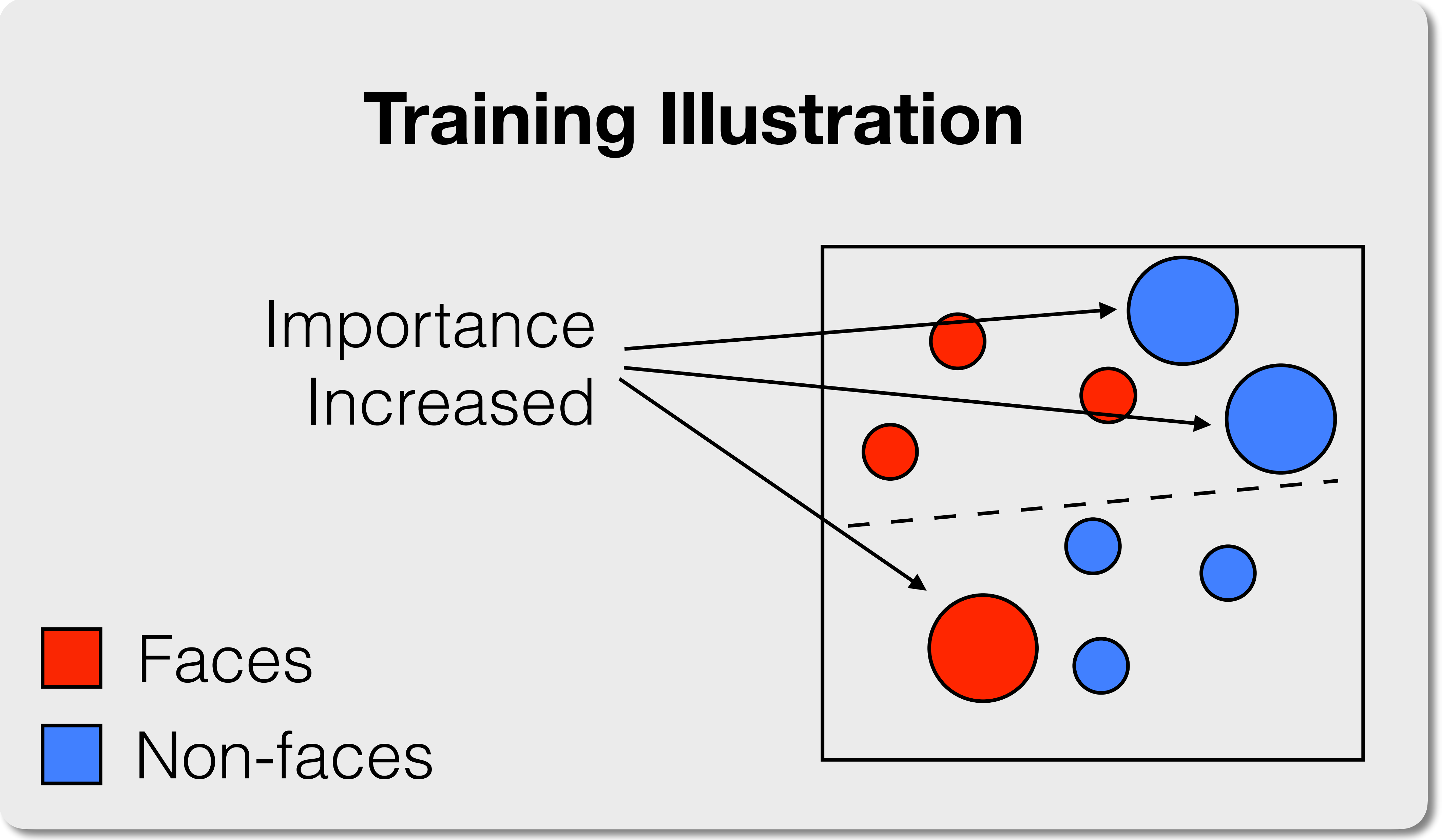
Source: Dr. Walter Scheirer

LOYOLA UNIVERSITY CHICAGO

Enhancement

Viola-Jones Detector

Boosting for
Feature Selection (3/4)
Solution: *boosting*, a combination of weak classifiers that when learned in sequence and applied together, lead to better final classification.



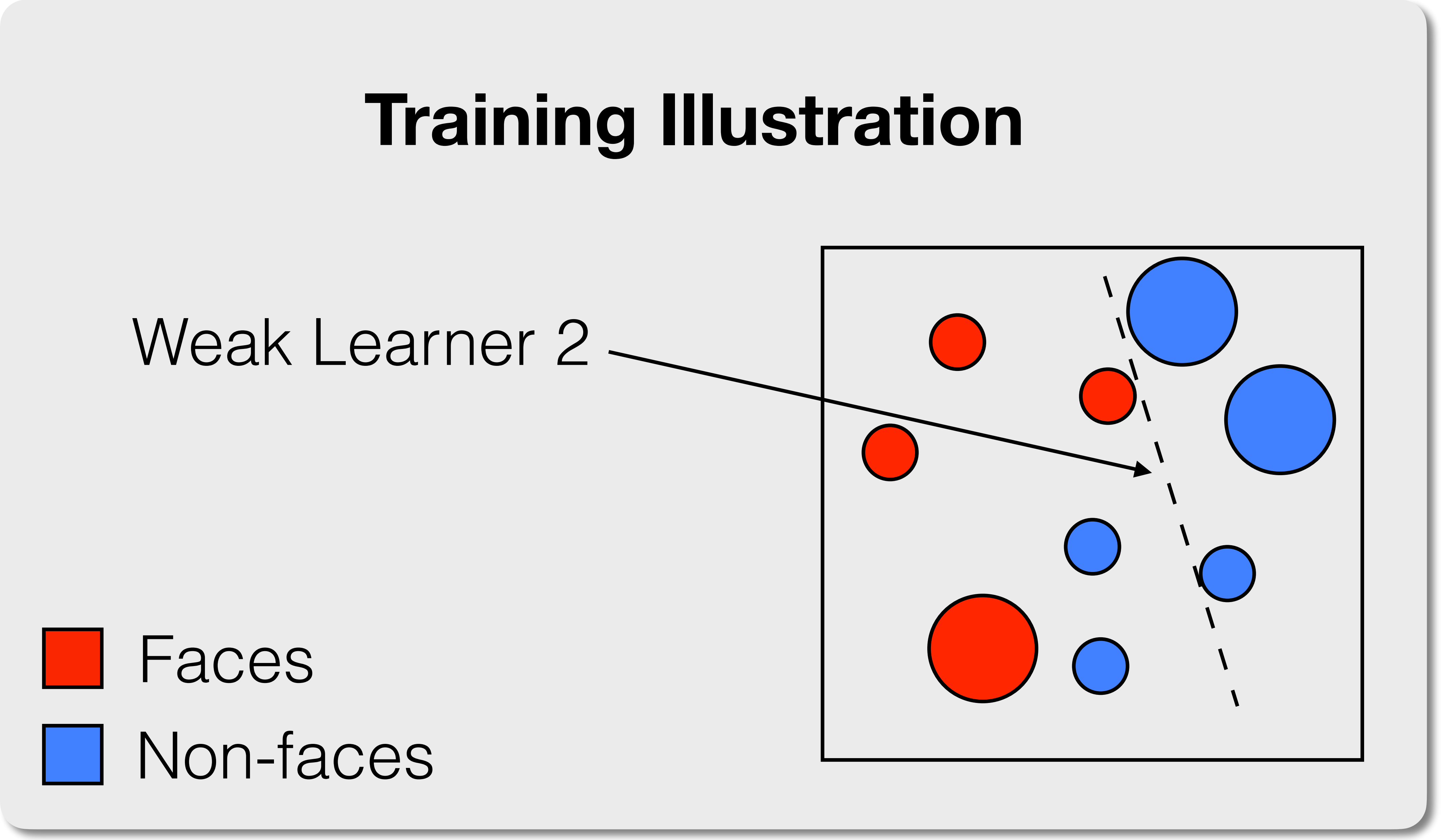
Source: Dr. Walter Scheirer

Enhancement

Viola-Jones Detector

Boosting for
Feature Selection (3/4)

Solution: *boosting*, a combination of weak classifiers that when learned in sequence and applied together, lead to better final classification.



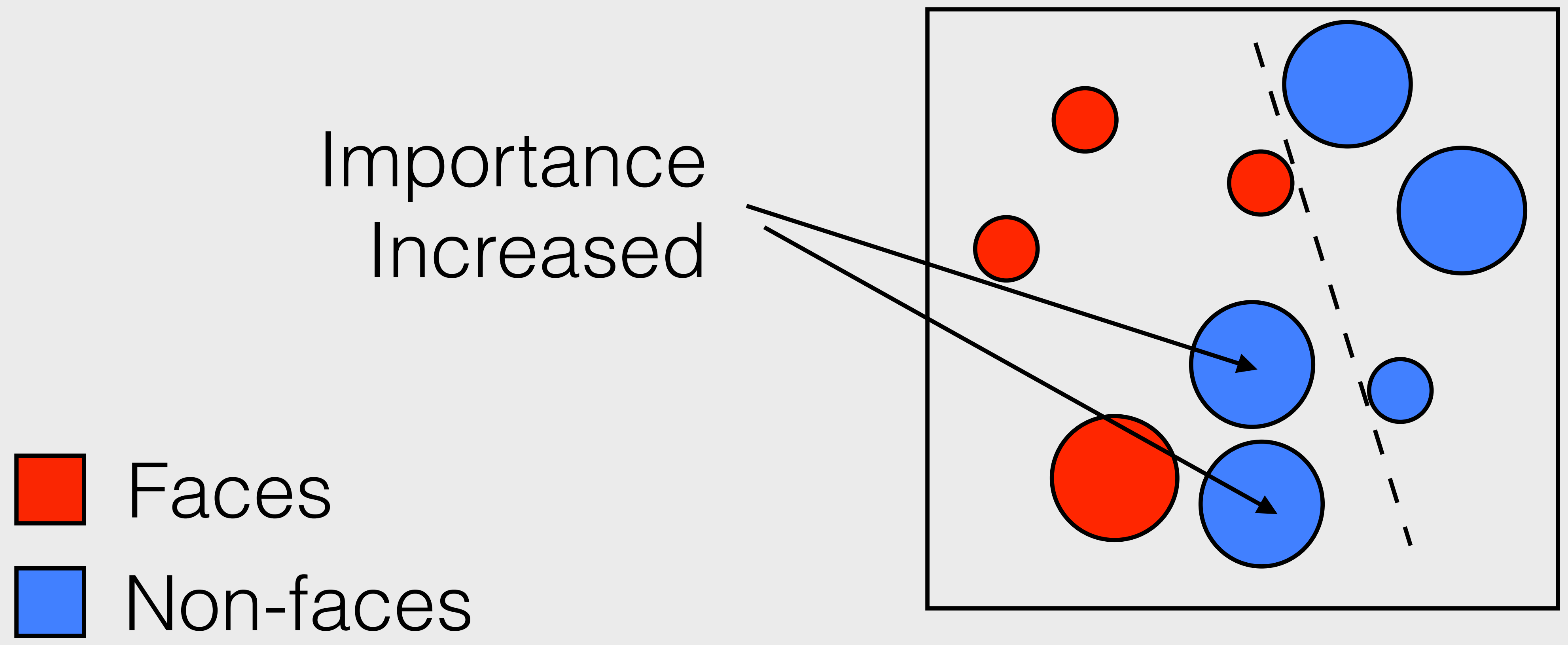
Source: Dr. Walter Scheirer

Enhancement

Viola-Jones Detector

Boosting for
Feature Selection (3/4)
Solution: *boosting*, a combination of weak classifiers that when learned in sequence and applied together, lead to better final classification.

Training Illustration

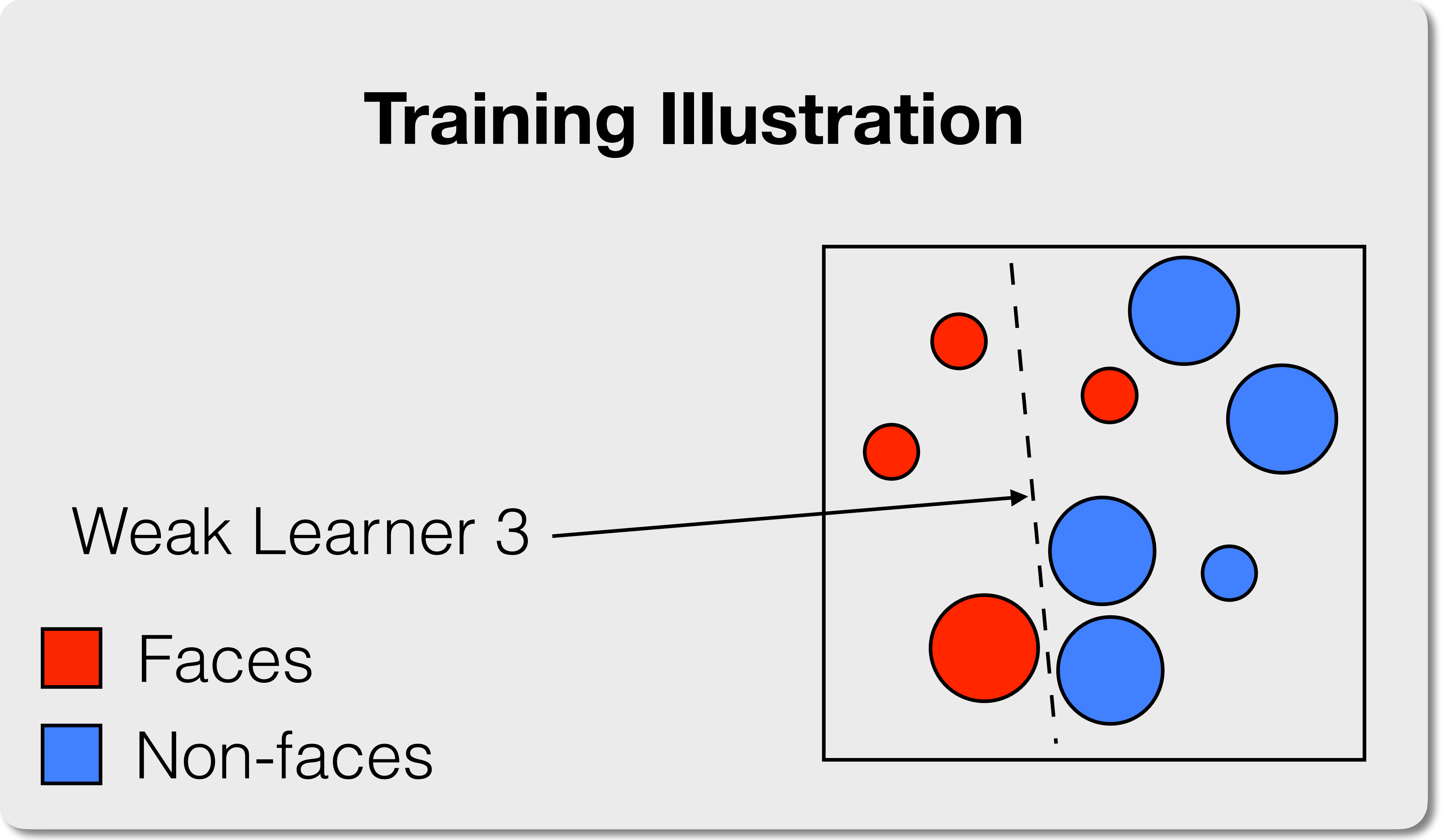


Source: Dr. Walter Scheirer

Enhancement

Viola-Jones Detector

Boosting for
Feature Selection (3/4)
Solution: *boosting*, a combination of weak classifiers that when learned in sequence and applied together, lead to better final classification.



Source: Dr. Walter Scheirer

LOYOLA UNIVERSITY CHICAGO

Enhancement

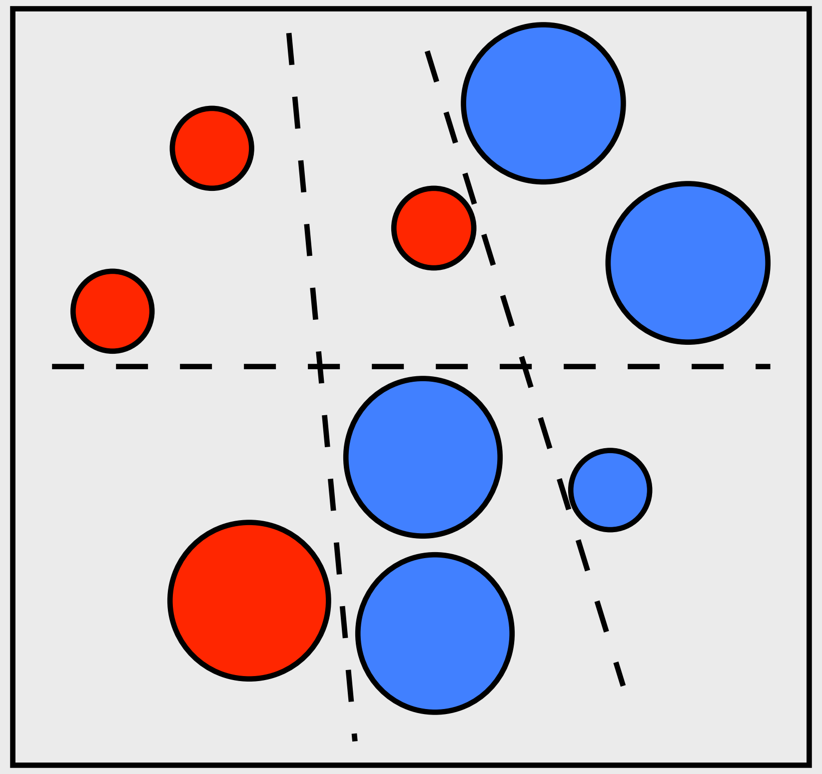
Viola-Jones Detector

Boosting for
Feature Selection (3/4)

Solution: *boosting*, a combination of weak classifiers that when learned in sequence and applied together, lead to better final classification.

Training Illustration

Final classifier
is a combination of 3
weaker classifiers.



Faces

Non-faces

Source: Dr. Walter Scheirer

LOYOLA UNIVERSITY CHICAGO

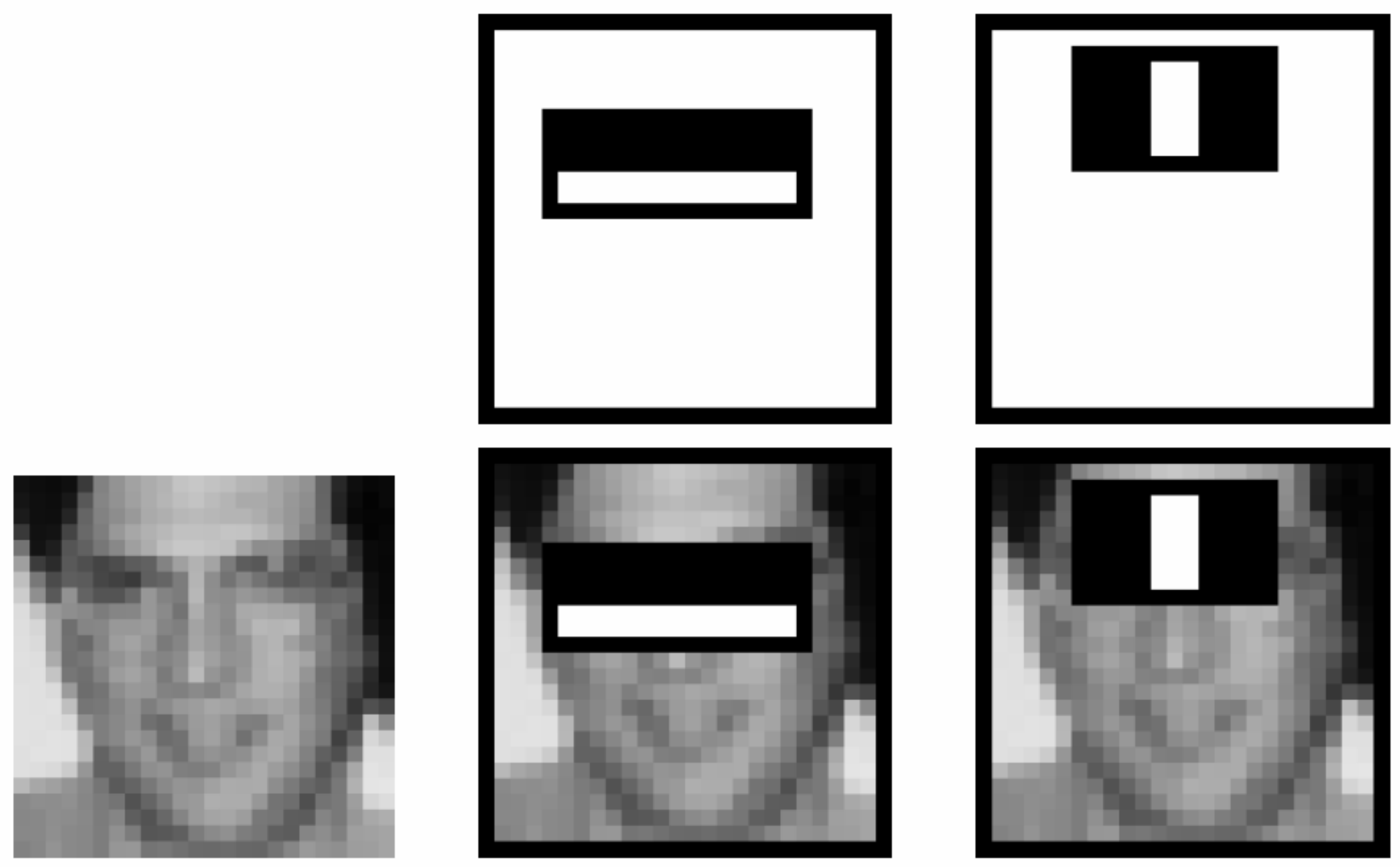
Enhancement

Viola-Jones Detector

Boosting for
Feature Selection (3/4)
Possible outcome.

This combination is enough
to lead to perfect True Positive Rate,
but poor False Positive Rate.

All faces are detected as positive, but many non-faces are detected as positive too.



First two selected features.

Whenever this classifier says an object is not a face (rejection), it is probably right.

LOYOLA UNIVERSITY CHICAGO

Enhancement

Face Detection

Viola-Jones Detector
First real-time face detector.
Based on sliding windows.

Key Ideas (4)
Haar-like features.
Integral image.
Boosting for feature selection.
Attentional Cascade to reject non-faces.

SECOND INTERNATIONAL WORKSHOP ON STATISTICAL AND COMPUTATIONAL THEORIES OF VISION – MODELING, LEARNING, COMPUTING, AND SAMPLING

VANCOUVER, CANADA, JULY 13, 2001.

Robust Real-time Object Detection

Paul Viola	Michael Jones
viola@merl.com	mjones@crl.dec.com
Mitsubishi Electric Research Labs	Compaq CRL
201 Broadway, 8th FL	One Cambridge Center
Cambridge, MA 02139	Cambridge, MA 02142

Abstract

This paper describes a visual object detection framework that is capable of processing images extremely rapidly while achieving high detection rates. There are three key contributions. The first is the introduction of a new image representation called the "Integral Image" which allows the features used by our detector to be computed very quickly. The second is a learning algorithm, based on AdaBoost, which selects a small number of critical visual features and yields extremely efficient classifiers [6]. The third contribution is a method for combining classifiers in a "cascade" which allows background regions of the image to be quickly discarded while spending more computation on promising object-like regions. A set of experiments in the domain of face detection are presented. The system yields face detection performace comparable to the best previous systems [18, 13, 16, 12, 1]. Implemented on a conventional desktop, face detection proceeds at 15 frames per second.

1. Introduction

This paper brings together new algorithms and insights to construct a framework for robust and extremely rapid object detection. This framework is demonstrated on, and in part motivated by, the task of face detection. Toward this end we have constructed a frontal face detection system which achieves detection and false positive rates which are equivalent to the best published results [18, 13, 16, 12, 1]. This face detection system is most clearly distinguished from previous approaches in its ability to detect faces extremely rapidly. Operating on 384 by 288 pixel images, faces are detected at 15 frames per second on a conventional 700

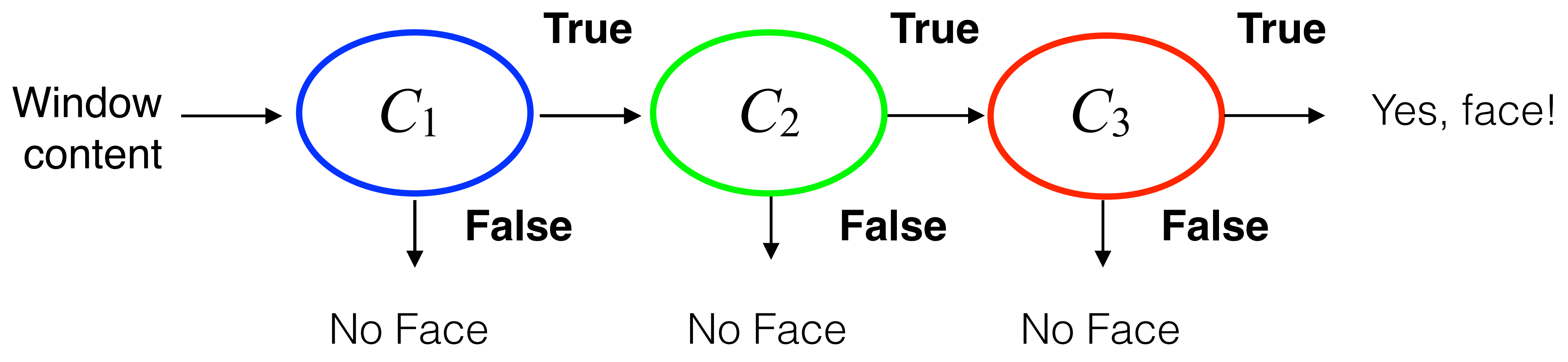
1

Enhancement

Viola-Jones Detector

Attentional Cascade (4/4)
Make a cascade of different classifiers that are good at rejecting faces.

Start with simpler and faster classifiers.



Enhancement

Viola-Jones Detector

Results

Jain, Ross, and Nadakumar
Introduction to Biometrics
Springer Books, 2011

clean background

cluttered background

tilted head

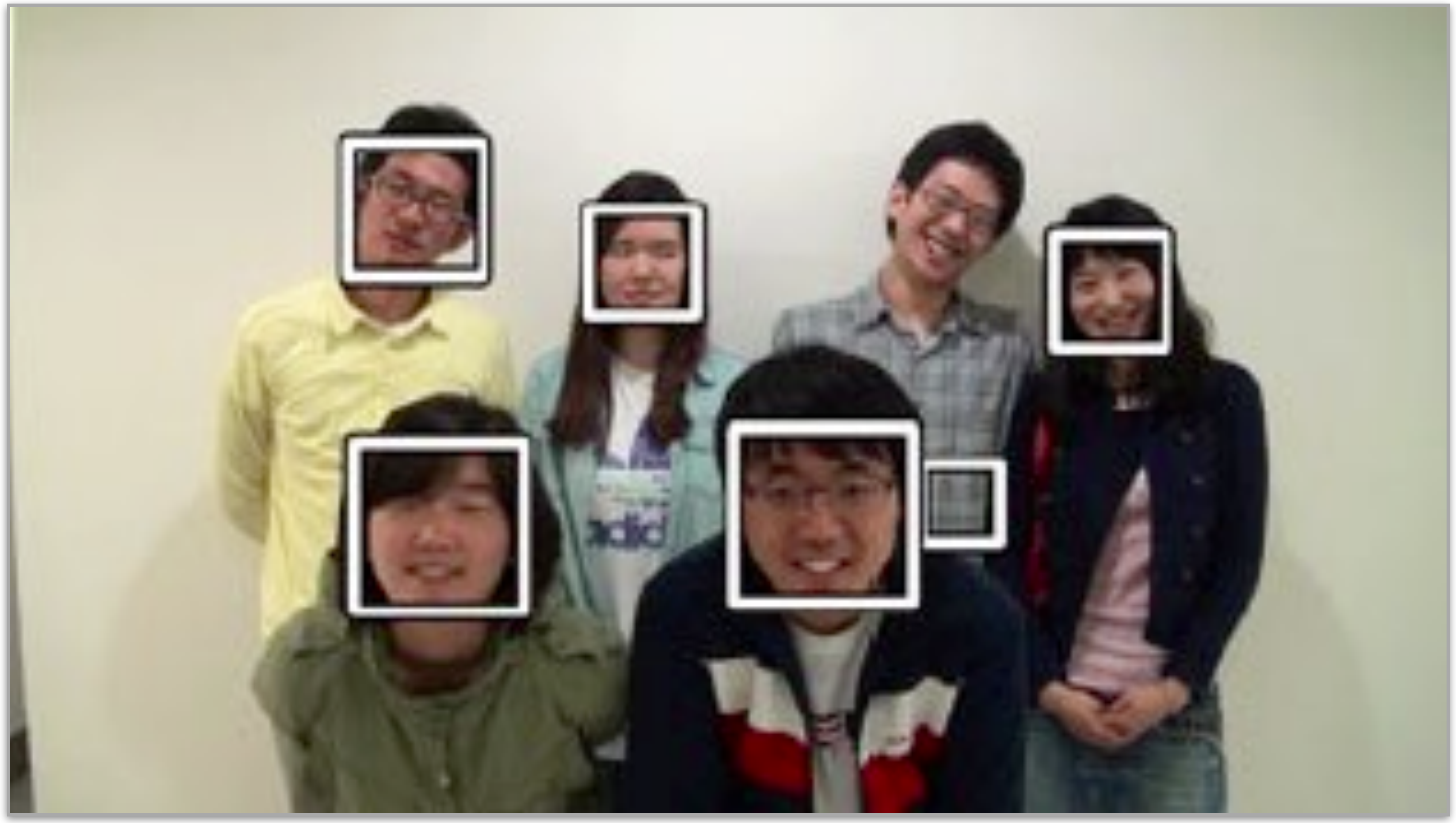
upside down

LOYOLA UNIVERSITY CHICAGO

Enhancement

Viola-Jones Detector

Results



Jain, Ross, and Nadakumar
Introduction to Biometrics
Springer Books, 2011

Enhancement

Face Detection

Attack
Non-live faces and some special patterns may be used to trigger the face detector on purpose.

If it happens too often, it will flood the system.


Support The Guardian
Available for everyone, funded by readers
Contribute → Subscribe →
Sign in
The Guardian
News Opinion Sport Culture Lifestyle
World ▸ Europe US Americas Asia Australia Middle East Africa Inequality More
Surveillance
This article is more than **6 months old**
The fashion line designed to trick surveillance cameras

https://www.theguardian.com/world/2019/aug/13/the-fashion-line-designed-to-trick-surveillance-cameras

Enhancement

Face Detection

https://twitter.com/glichfield/status/925425702194810882

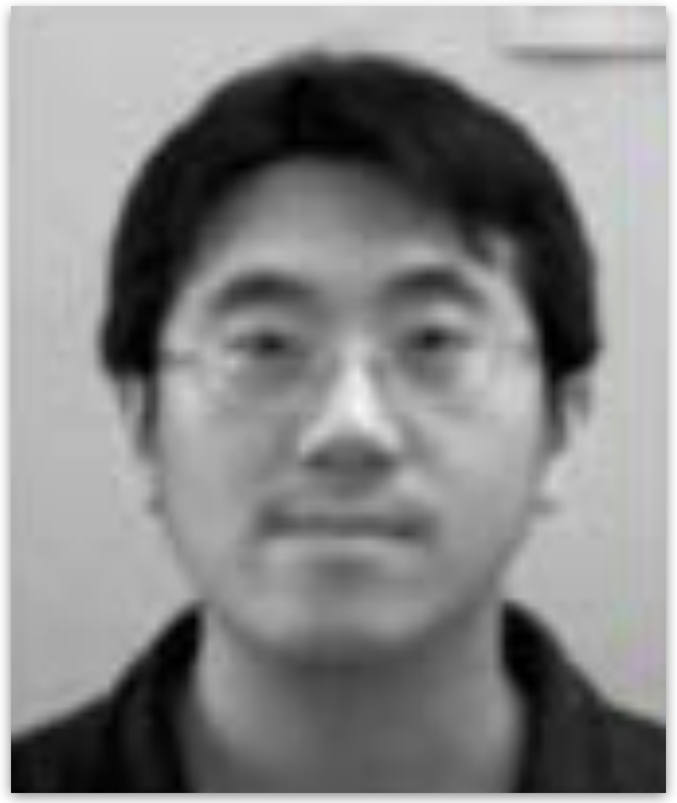
Attack
Make-up can be used to hinder detection.

LOYOLA UNIVERSITY CHICAGO

Enhancement

Face Alignment

Goal
Make template and sample faces be in similar poses, to make further description and matching easier.


template


sample

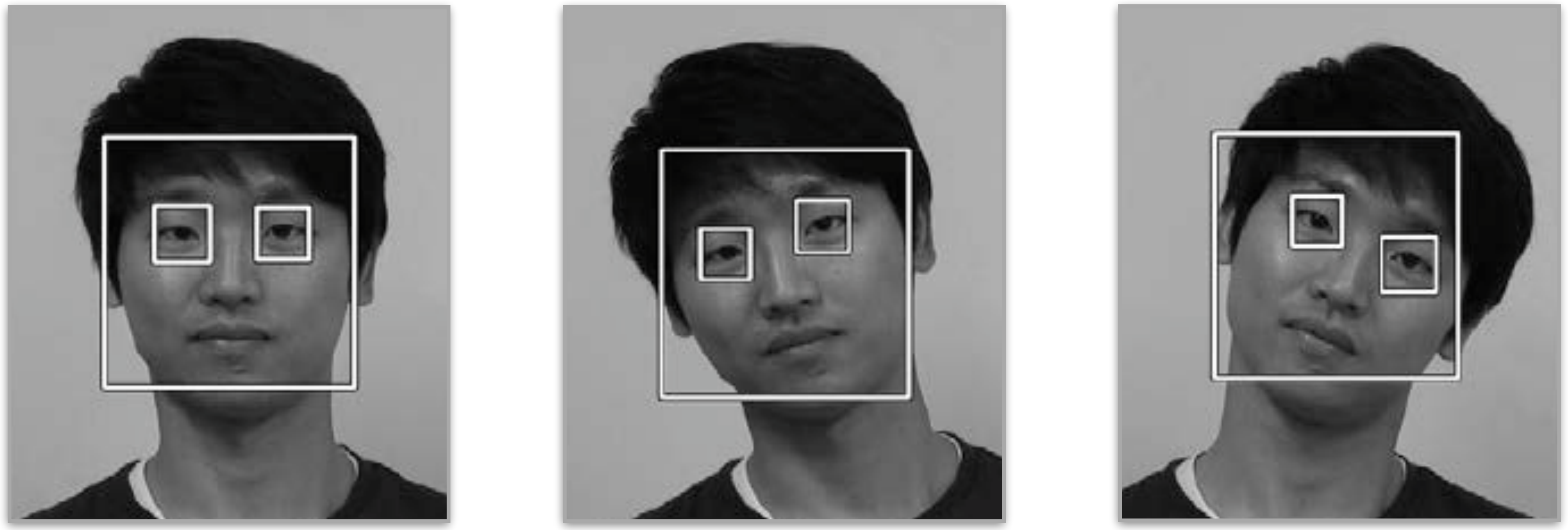
LOYOLA UNIVERSITY CHICAGO

Enhancement

Face Alignment

Detection of
Face Landmarks
E.g., position of eyes.

Jain, Ross, and Nadakumar
Introduction to Biometrics
Springer Books, 2011



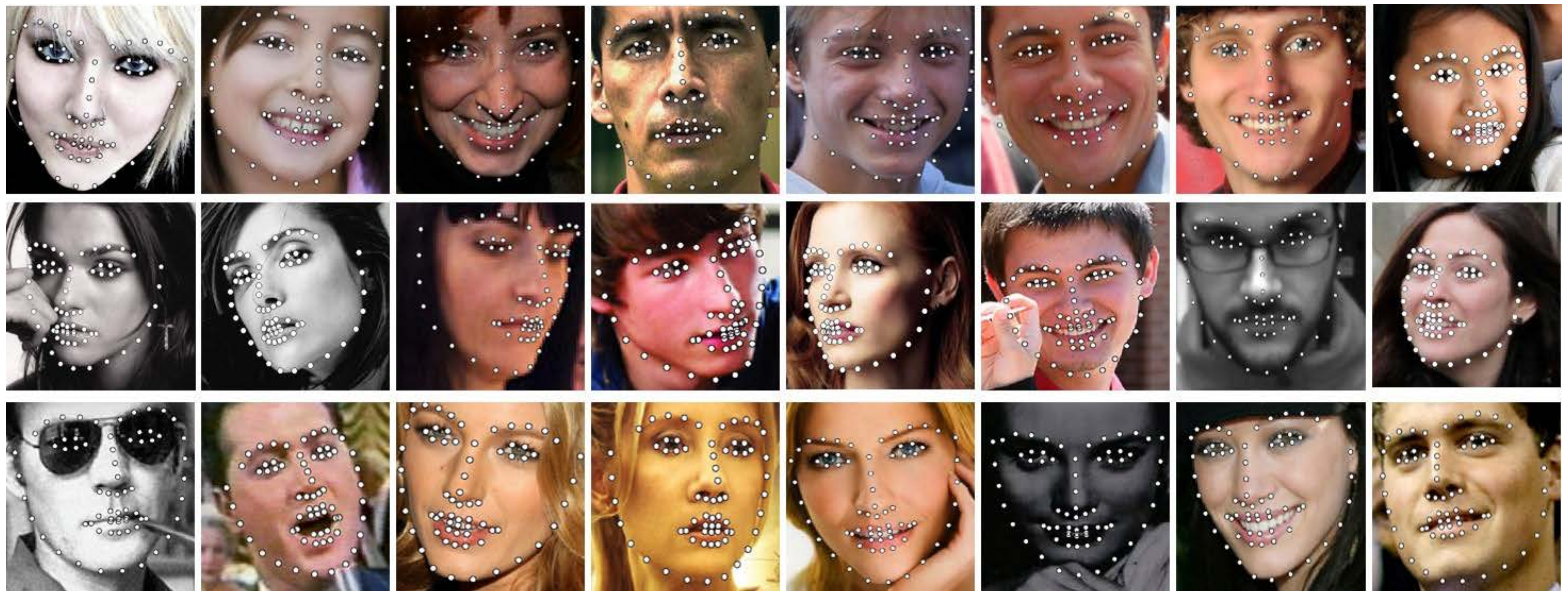
Possible solution: eye detection using Viola-Jones approach.

LOYOLA UNIVERSITY CHICAGO

Enhancement

Face Alignment

Detection of
Face Landmarks
There are better solutions in the literature, using deep neural networks, for instance.



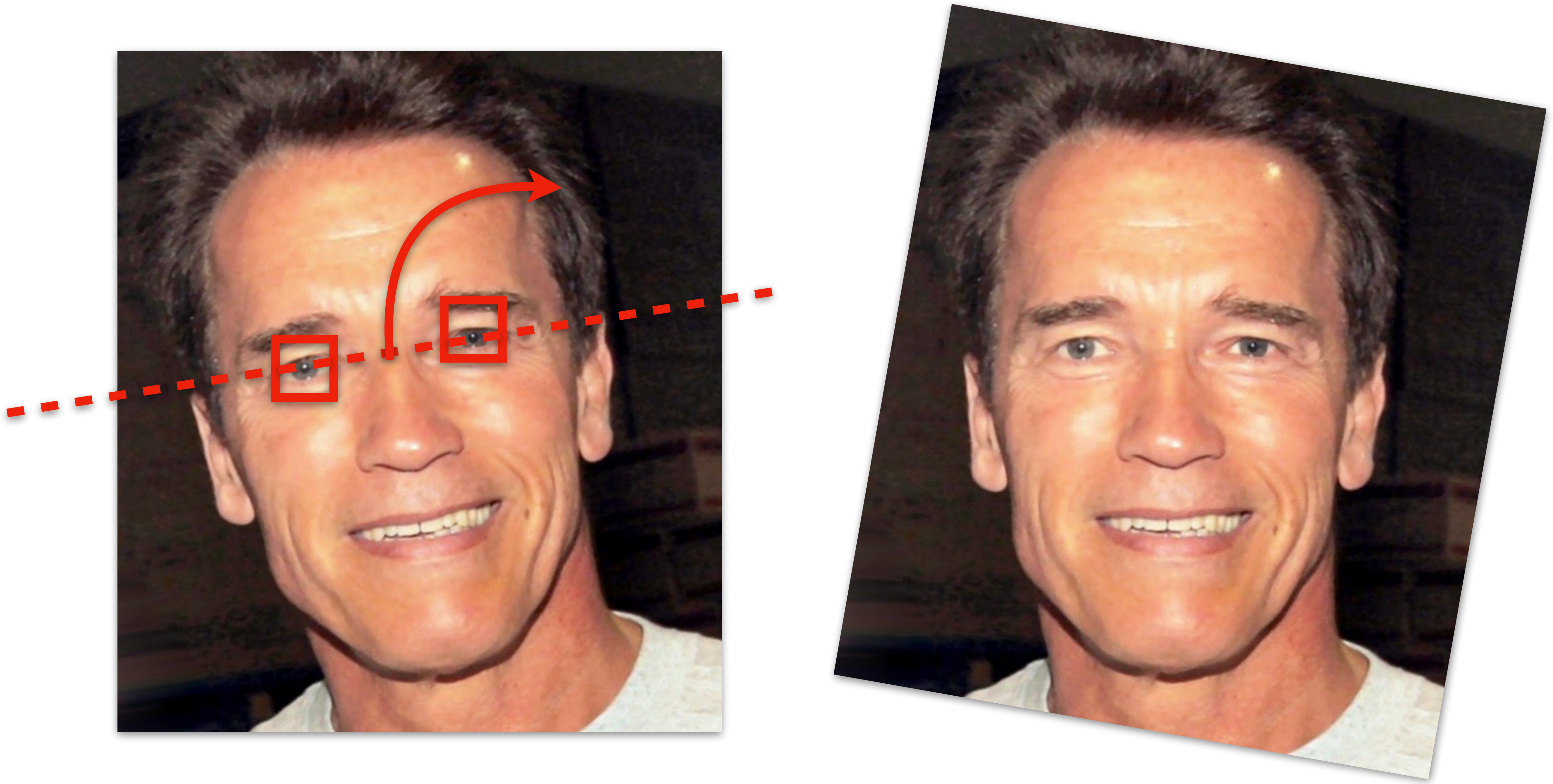
Zhang et al.
Facial Landmark Detection by Deep Multi-task Learning
ECCV 2014

LOYOLA UNIVERSITY CHICAGO

Enhancement

Face Alignment

Landmark Alignment
E.g., make the positions of the eyes horizontally aligned, by rotating the face image.



http://www.bytefish.de/blog/aligning_face_images/

Enhancement

Face Alignment

Cropping
Make a tight crop of the face, to remove background.

Keep eyes, nose, and mouth.

http://www.bytefish.de/blog/aligning_face_images/

LOYOLA UNIVERSITY CHICAGO

Enhancement

Face Alignment

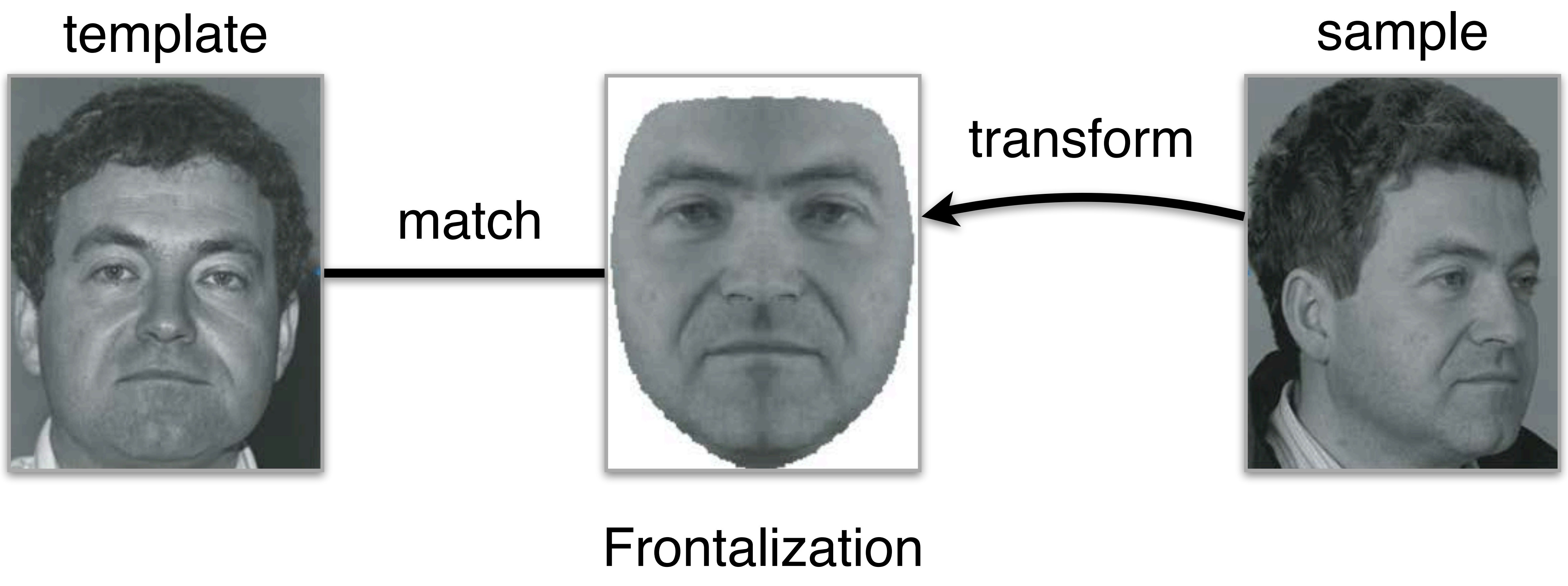
More Severe
Pose Variations
Naïve approach will not work.

Enhancement

Face Alignment

More Severe
Pose Variations
Alternative approaches.
3D information will help
to do frontalization.

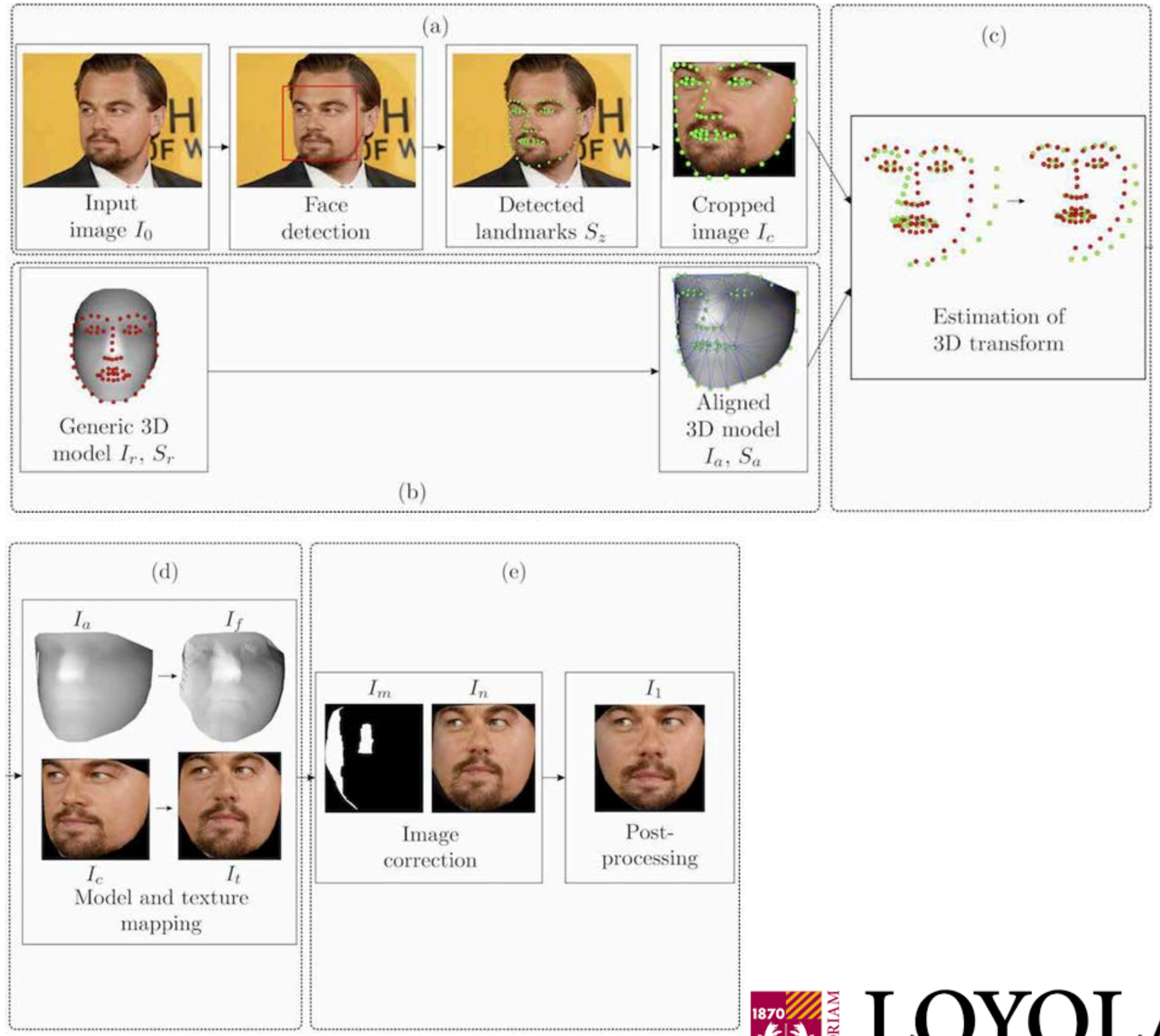
Yi et al.
Towards Pose Robust Face Recognition
CVPR 2013



Enhancement

Face Alignment

More Severe
Pose Variations
Alternative approaches.
3D information will help
to do frontalization.



Banerjee et al.
To frontalize or not to frontalize: Do we really need elaborate pre-processing to improve face recognition?
WACV 2018

LOYOLA UNIVERSITY CHICAGO

Enhancement

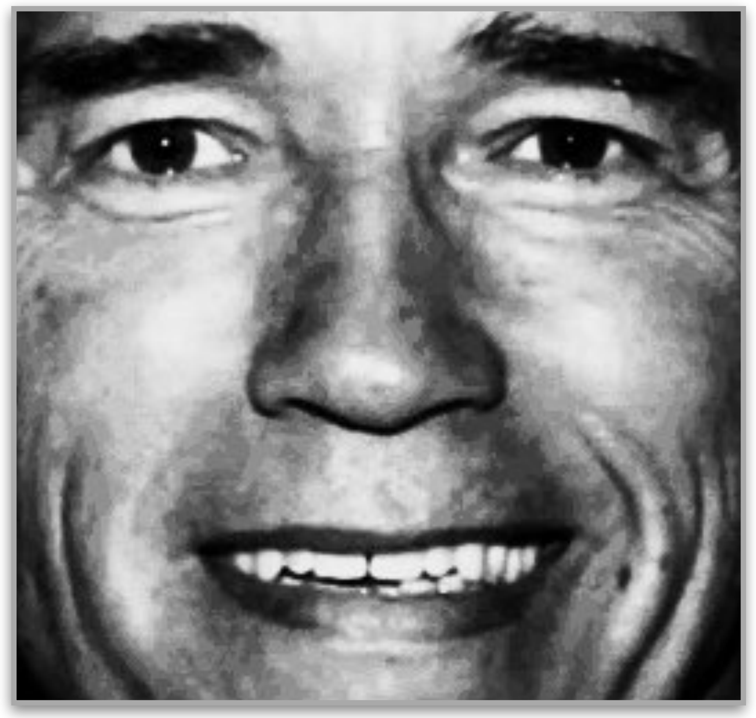
Illumination Correction

Simplest Solution
Color histogram equalization.

Alternatives
Photometric normalization, illumination modeling, etc.

Original

Grayscale



Equalized

What’s Next?

Face Description and Matching

Fill out your
***Today-I-missed* Statement**
Please visit https://sakai.luc.edu/x/PnQvIG.