HOGgles: Visualizing Object Detection Features




Carl Vondrick Aditya Khosla Tomasz Malisiewicz Antonio Torralba

Massachusetts Institute of Technology
Computer Science and Artificial Intelligence Laboratory

Chair

Car


U.S. NAVY
U.S. NAVY
U.S. NAVY
51

Aeroplane

Aeroplane

Aeroplane

Person

Chair

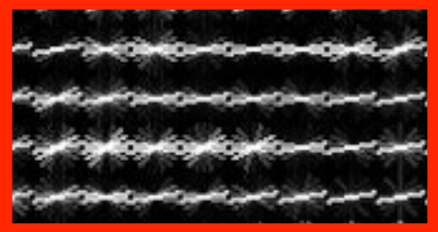
Chair

Chair


Car

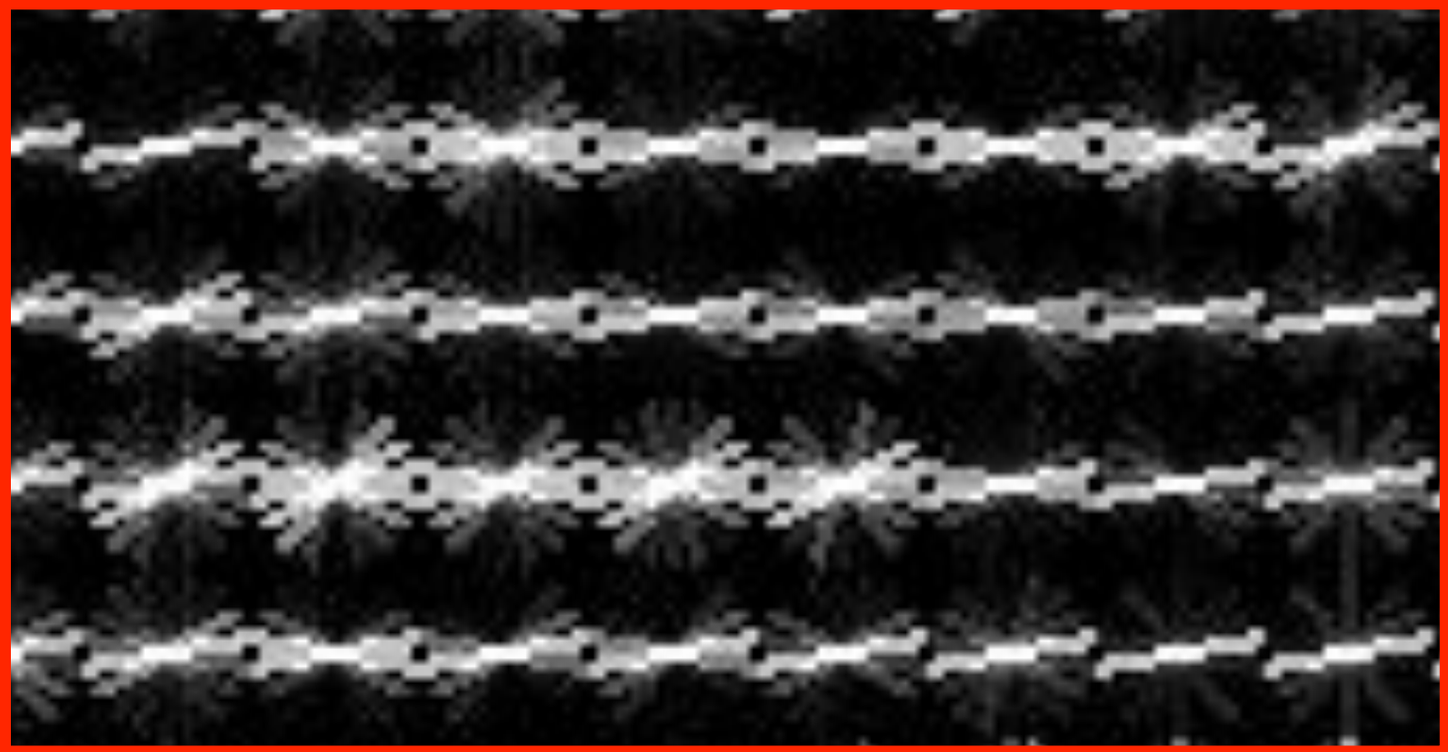

Car

Car

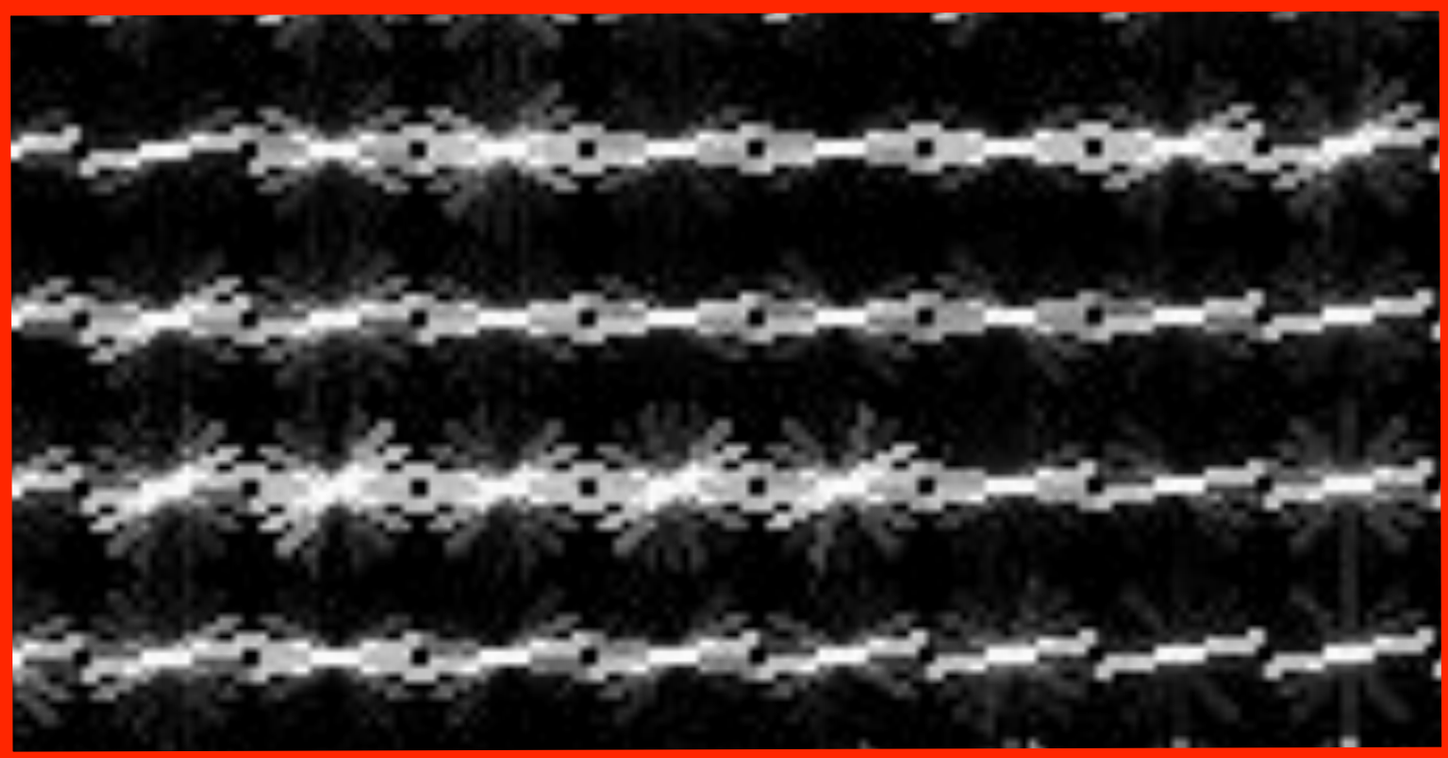



Car

Car



Car



What information does HOG have?

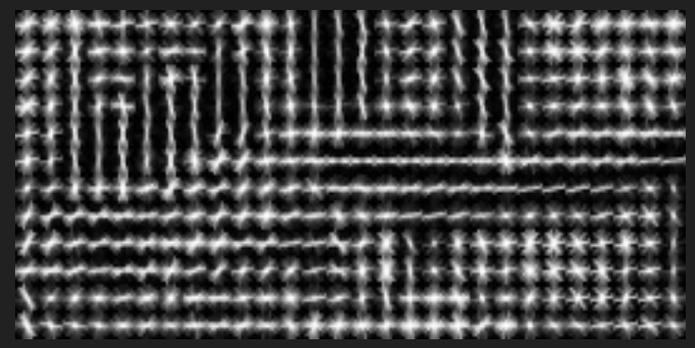
What information does HOG have?

Image

What information does HOG have?

Image

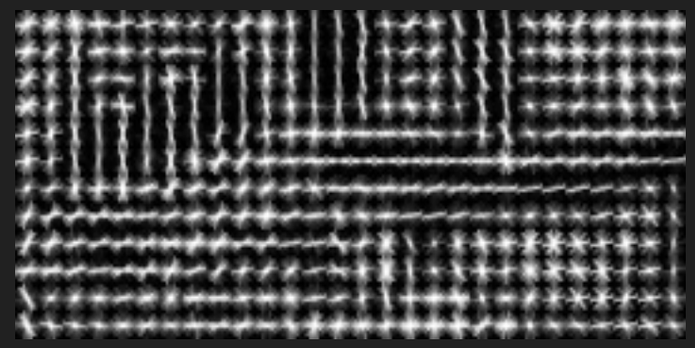
HOG

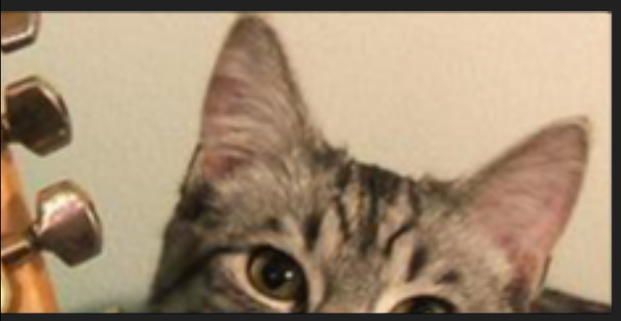


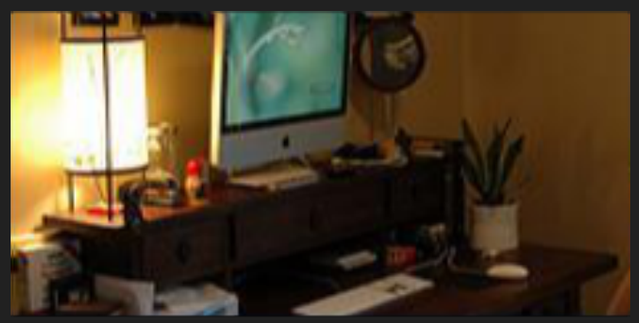
What information does HOG have?

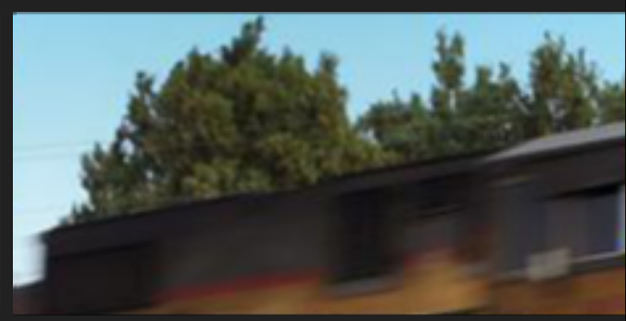
Image

HOG







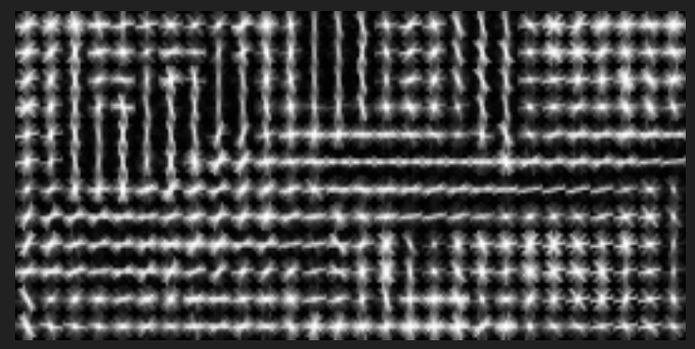


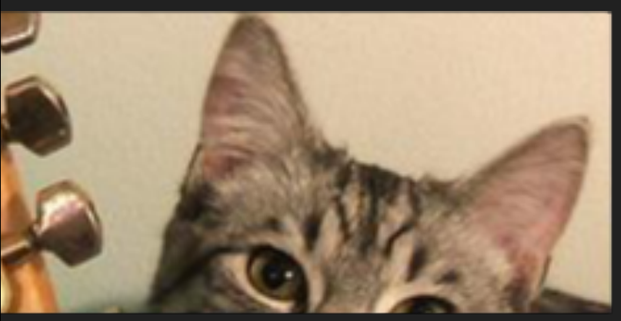
Nearest Neighbors

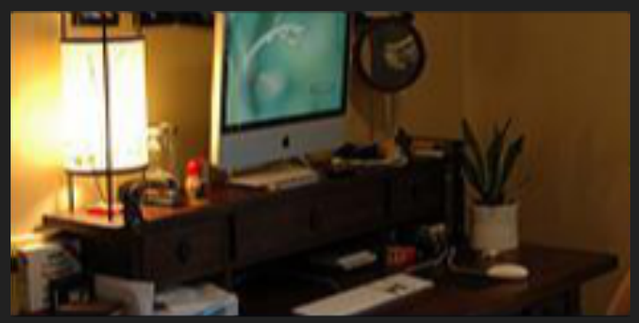
What information does HOG have?

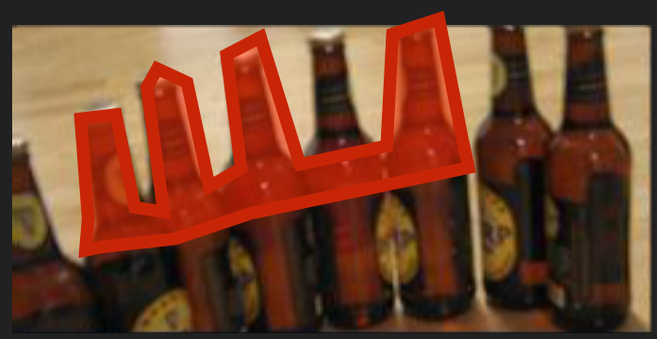
Image

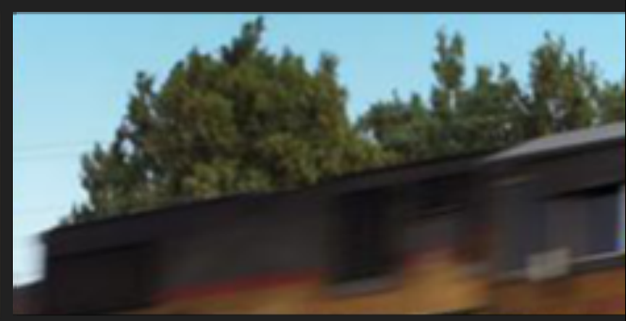
HOG









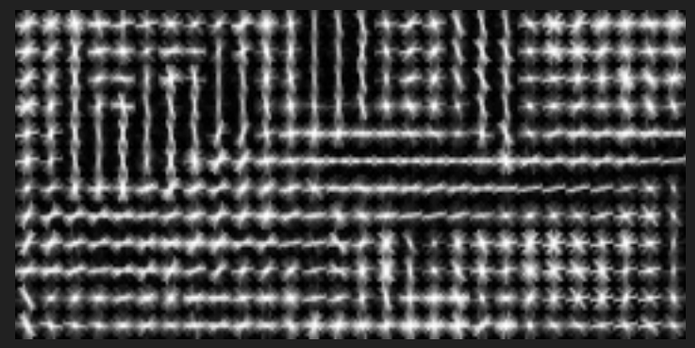


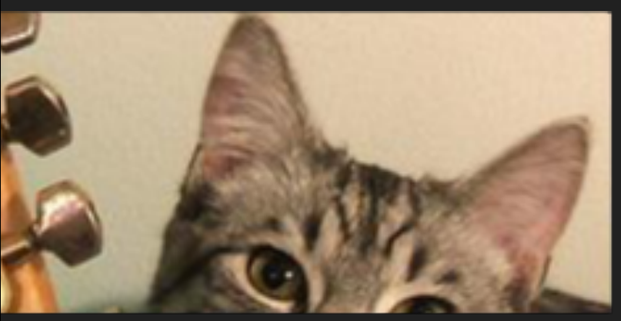
Nearest Neighbors

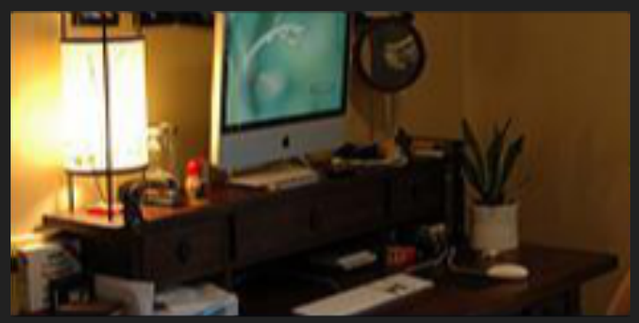
What information does HOG have?

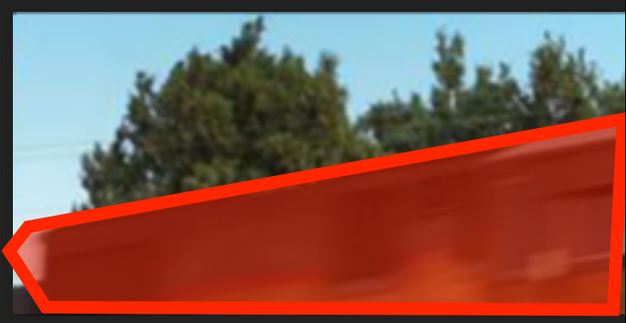
Image

HOG







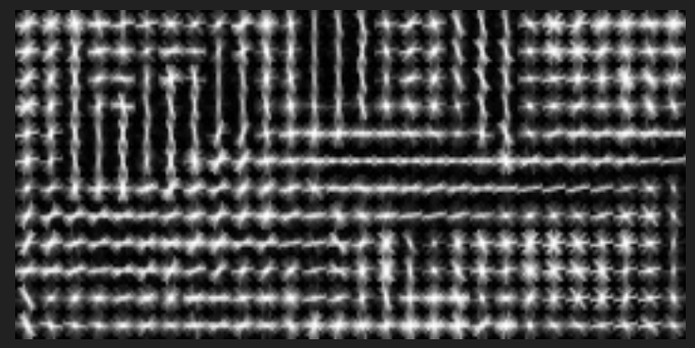


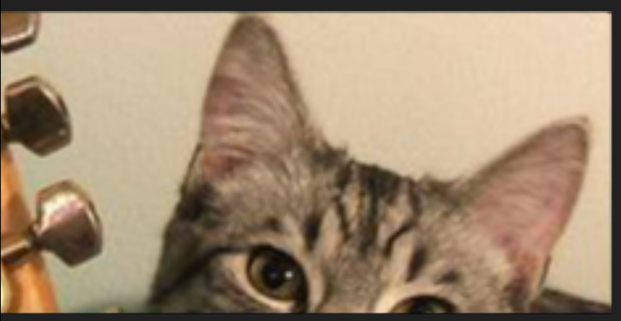
Nearest Neighbors

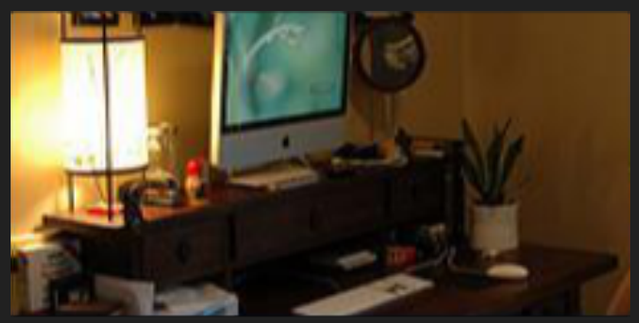
What information does HOG have?

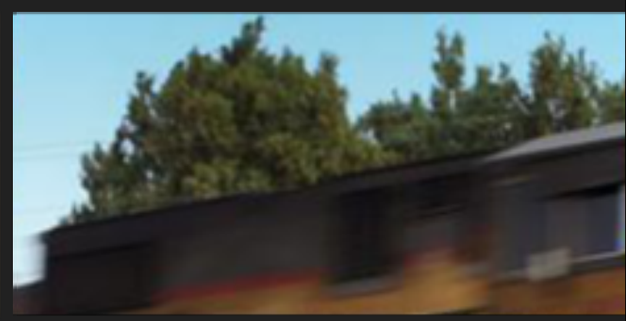
Image

HOG





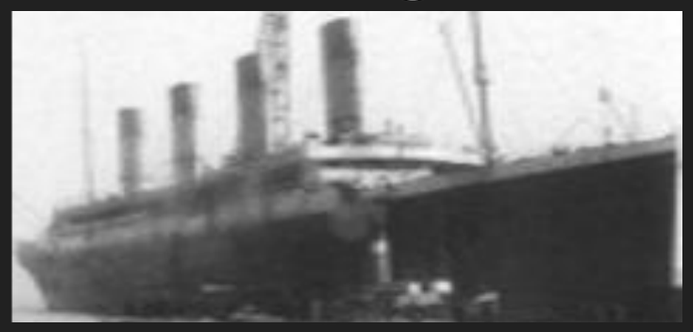


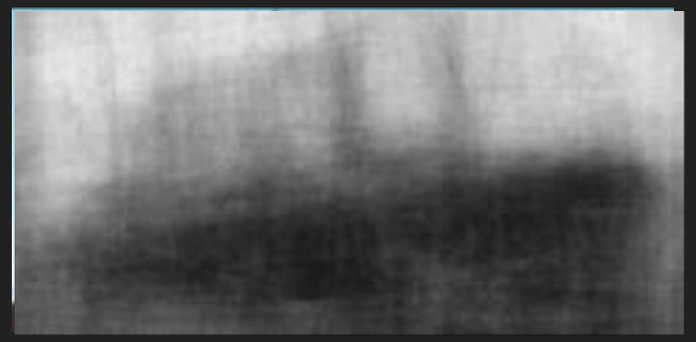


Nearest Neighbors

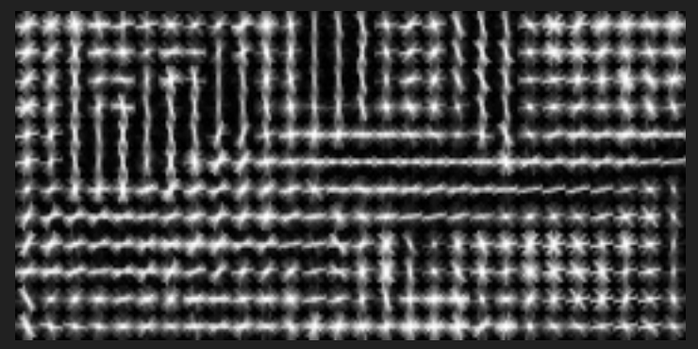
What information does HOG have?

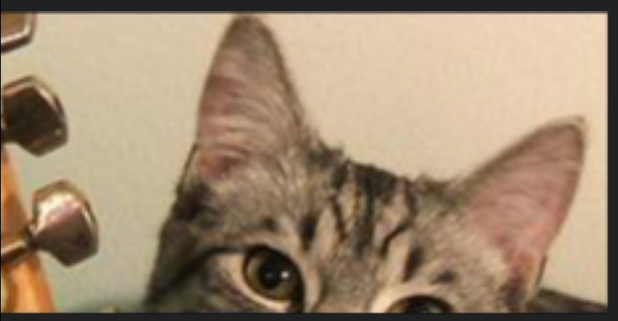
Image

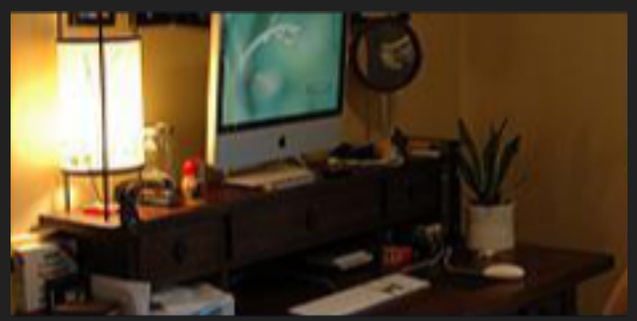


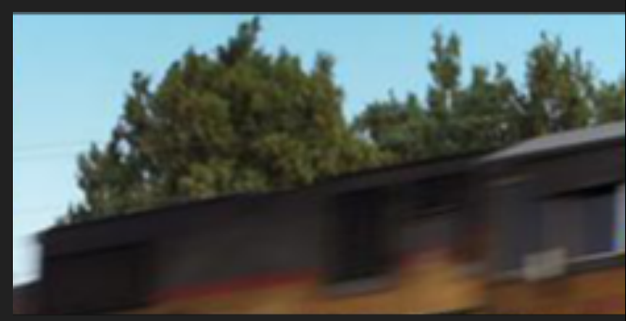


HOG



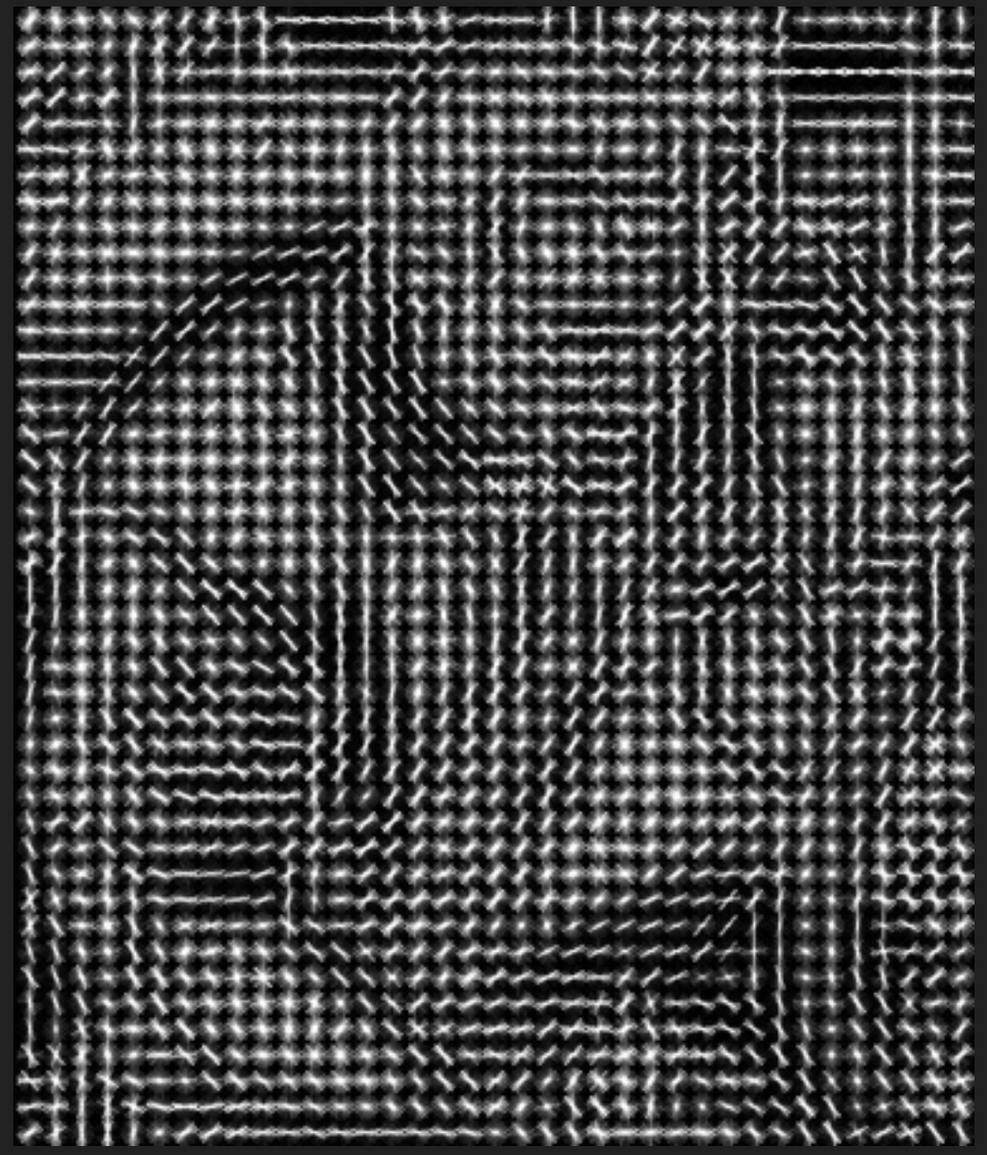




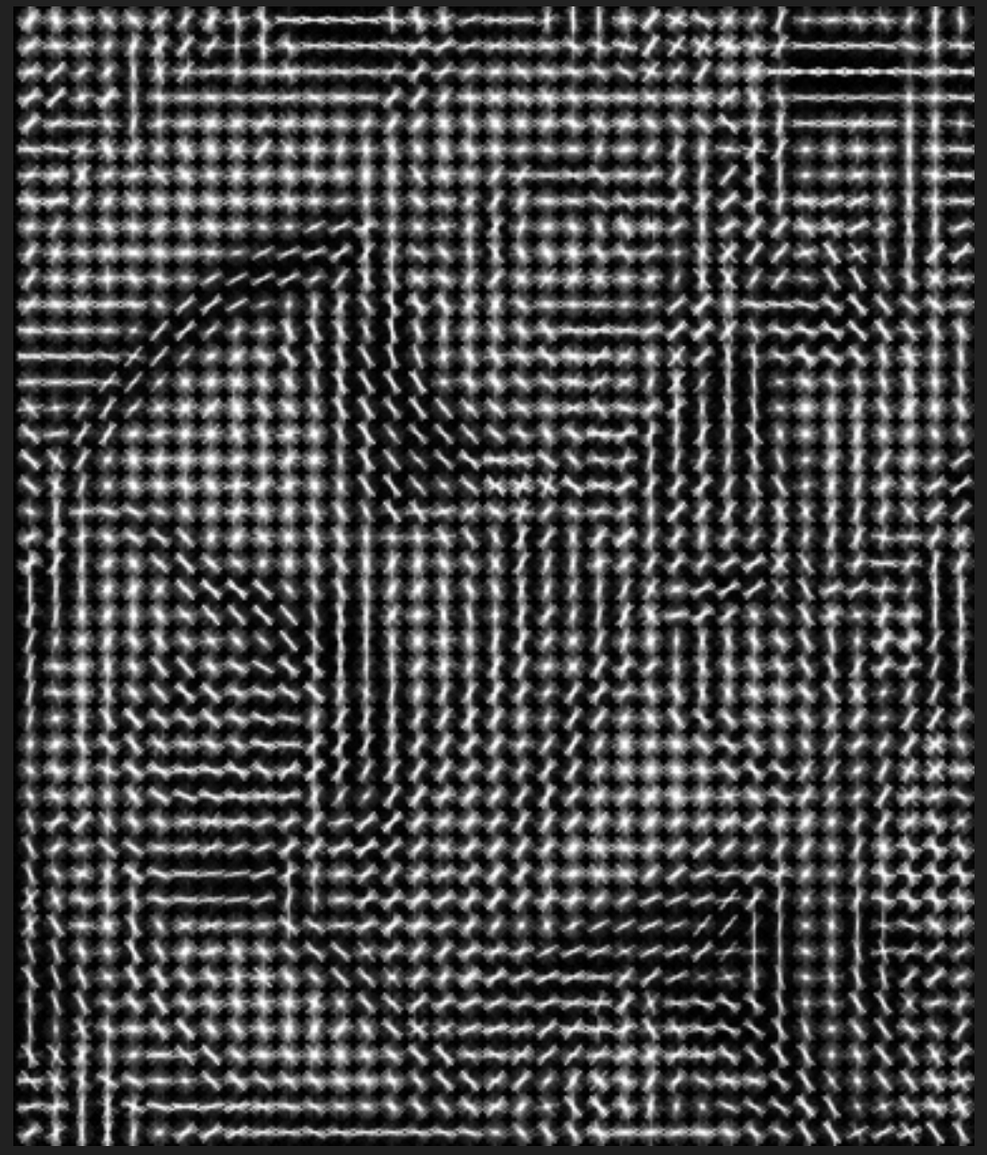


Nearest Neighbors

What information is lost?

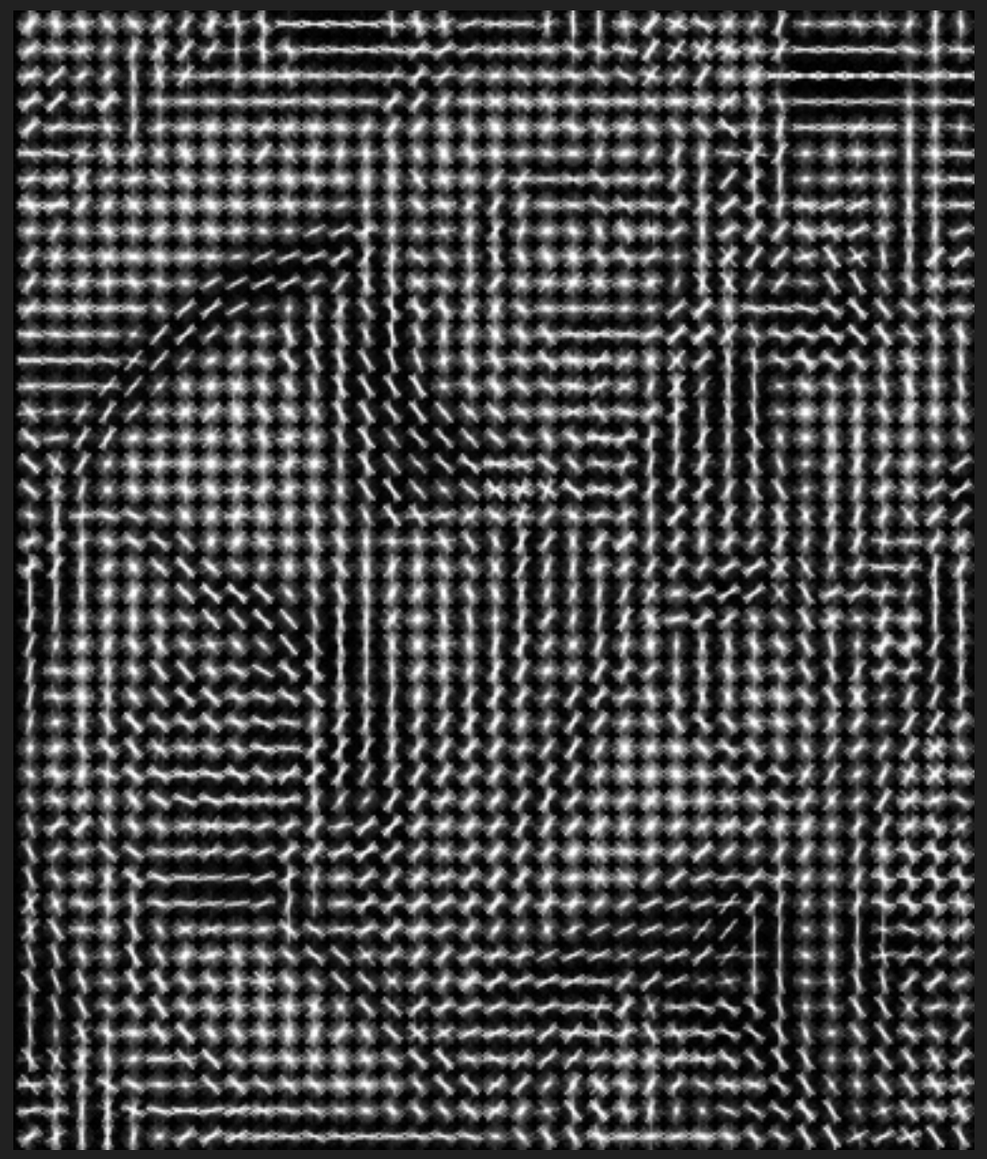


What information is lost?

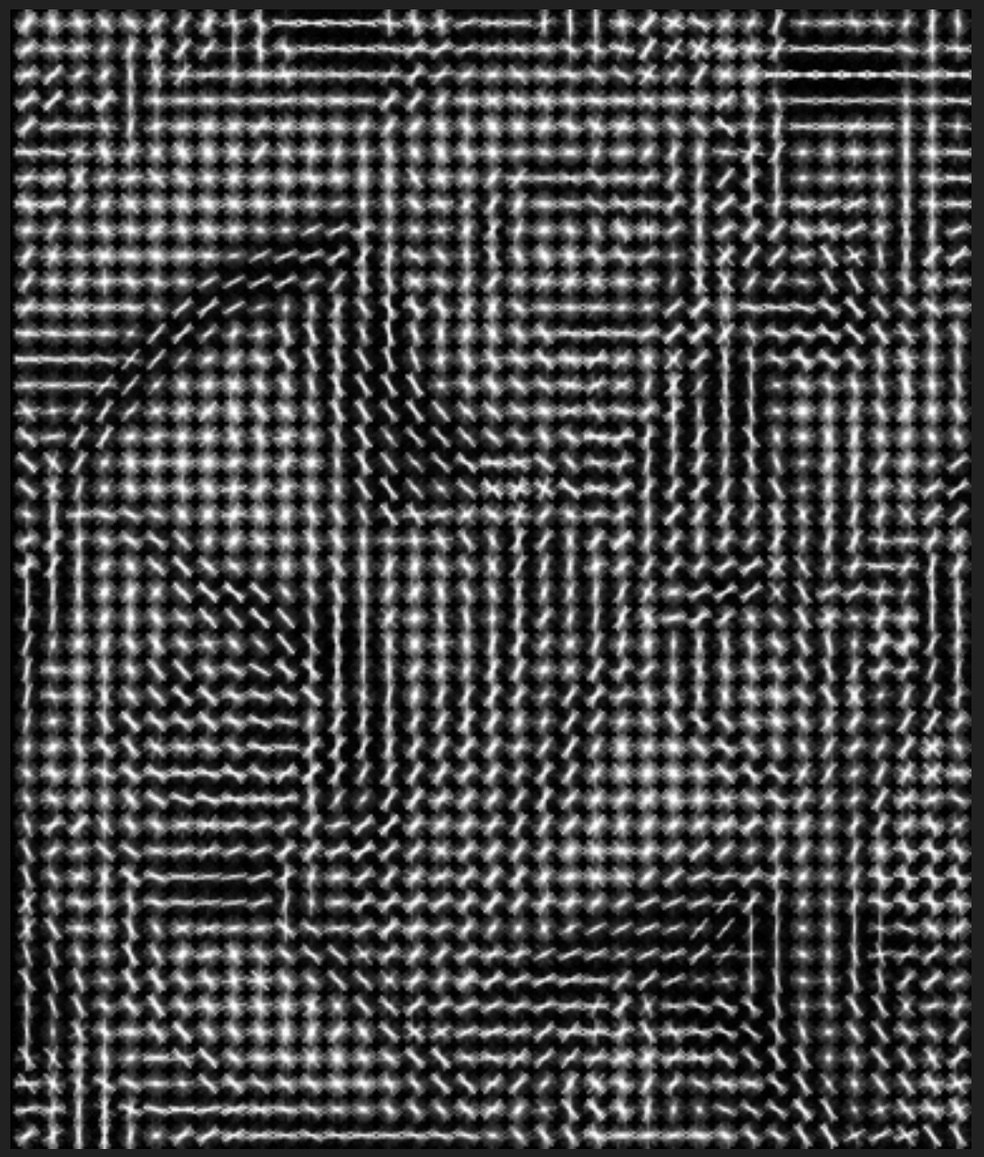


What information is lost?

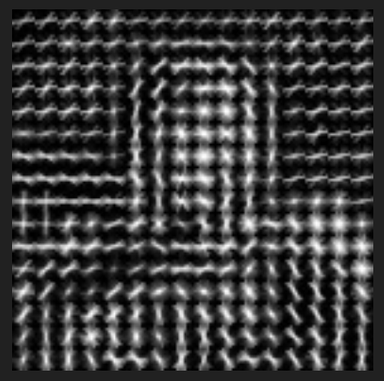
$$\min_{x \in \mathbb{R}^d} ||\phi(x) - y||_2^2$$



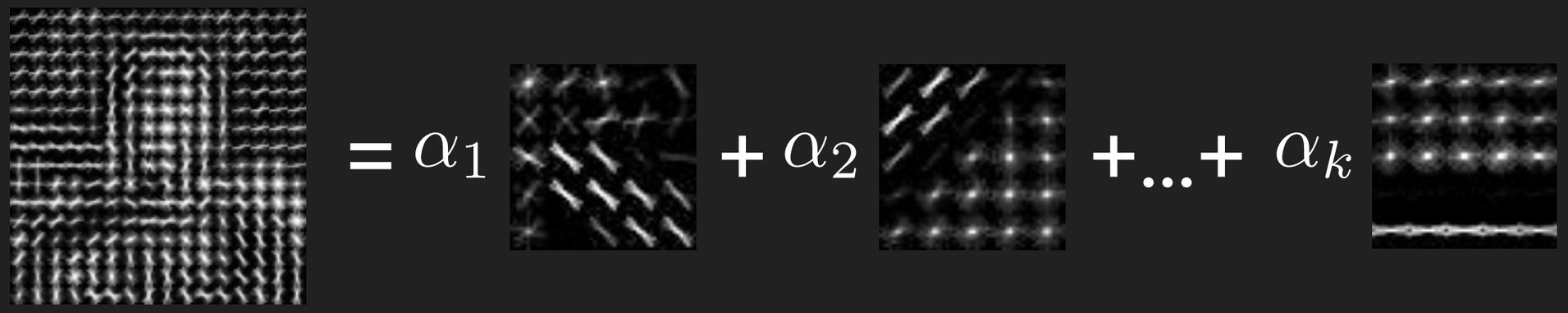
What information is lost?



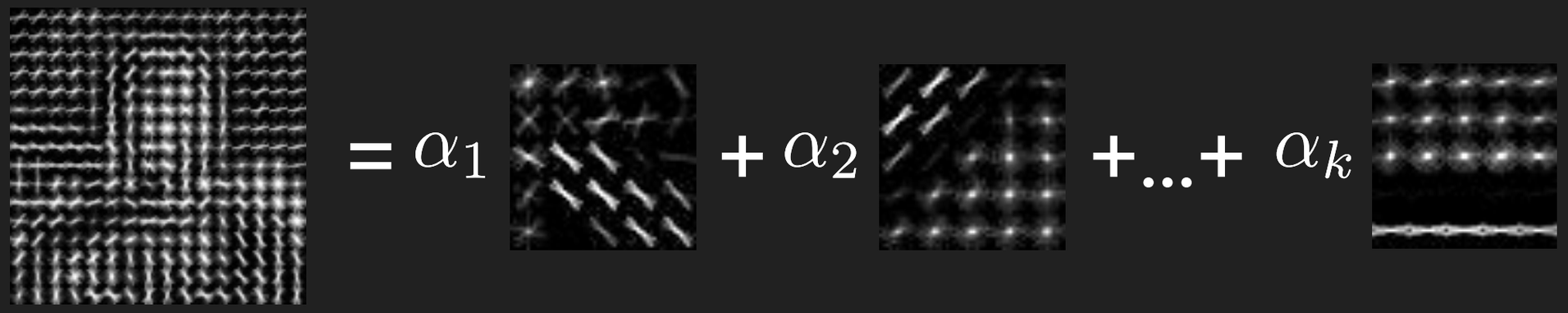
Method: Paired Dictionary

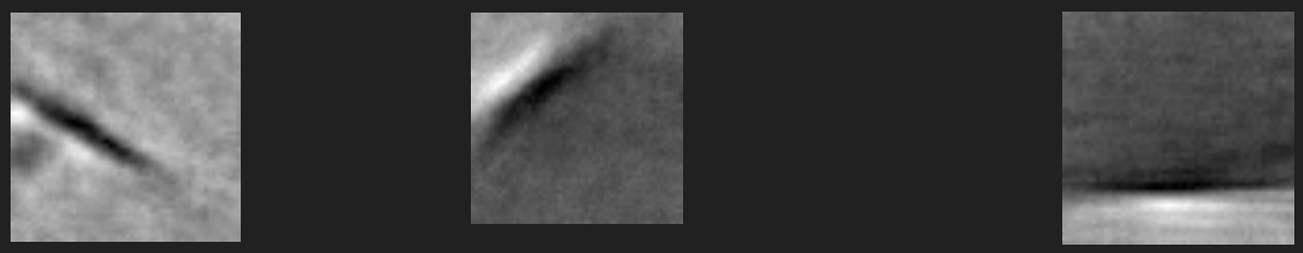


Method: Paired Dictionary

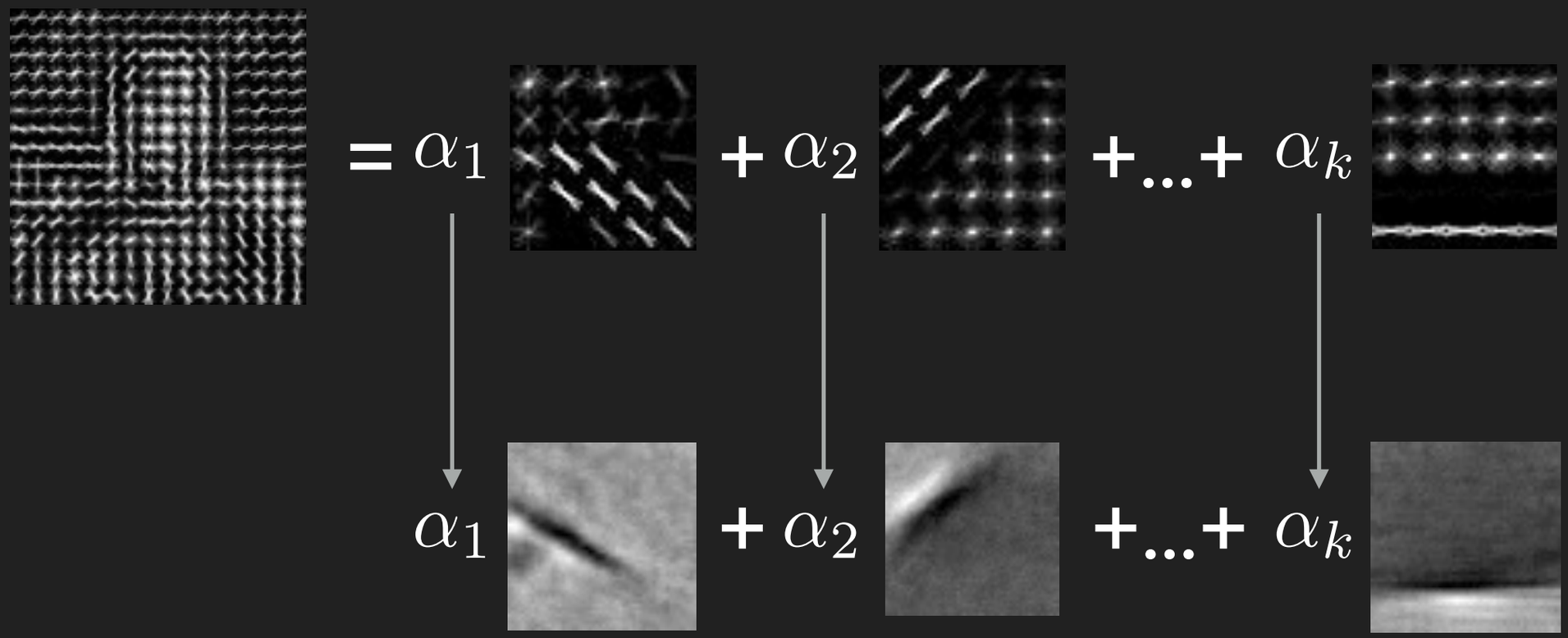


Method: Paired Dictionary

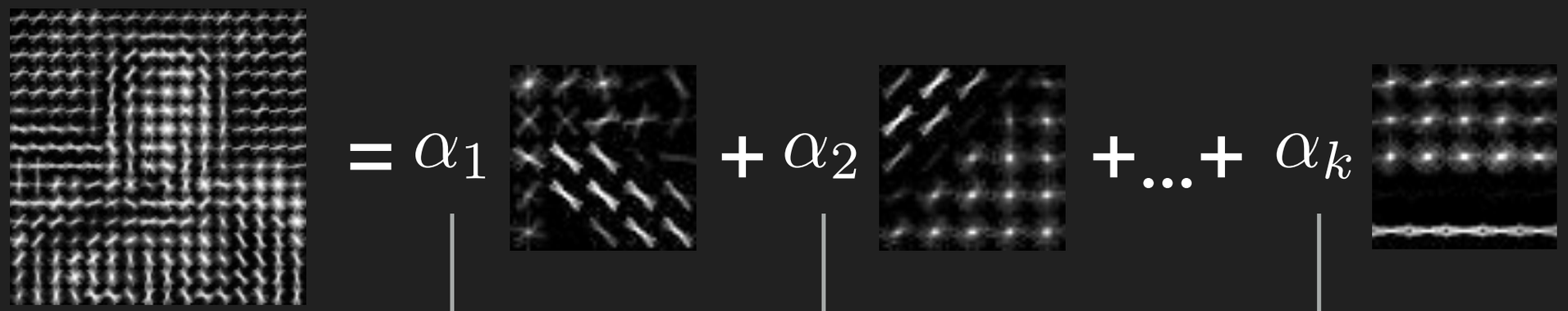


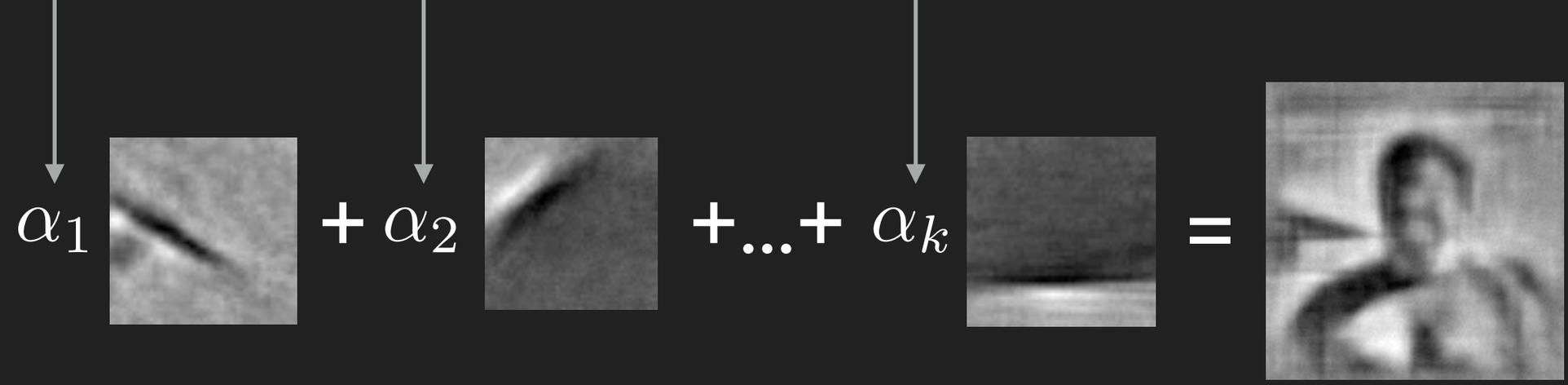


Method: Paired Dictionary

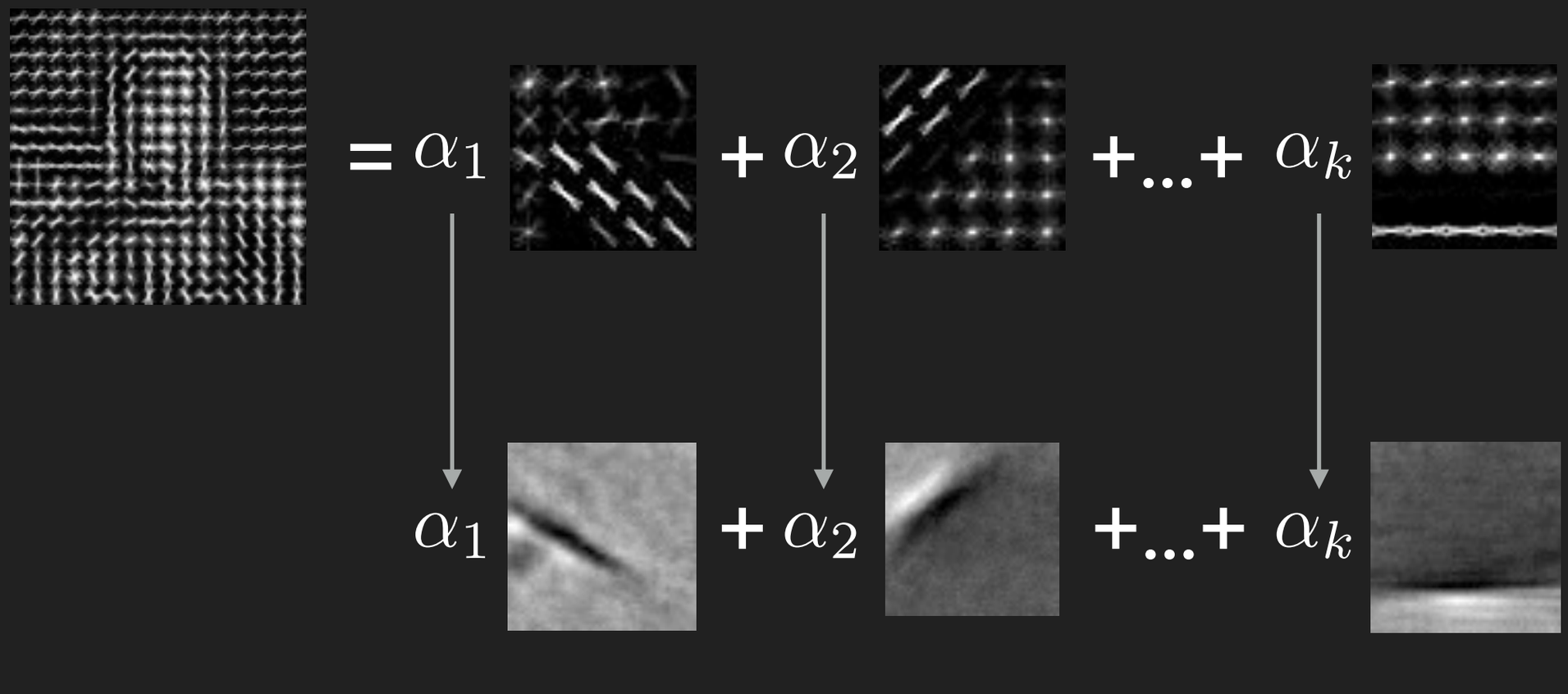


Method: Paired Dictionary



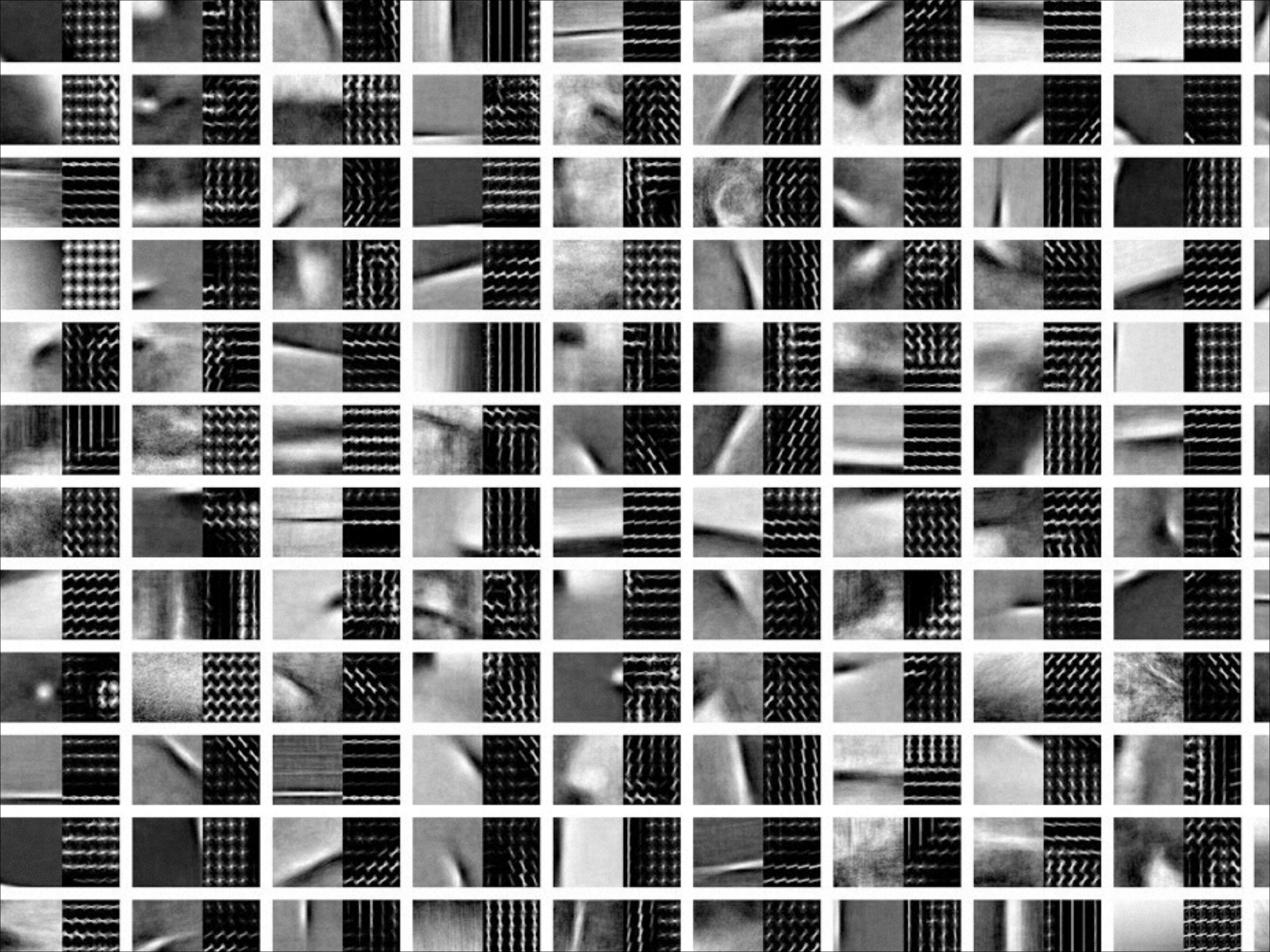


Method: Paired Dictionary



$$\hat{y} = f(x) = V\hat{\alpha}$$

$$\text{where} \quad \hat{\alpha} = \arg\min_{\alpha} ||x - U\alpha||_2^2 \quad \text{s.t.} \quad ||\alpha||_1 \leq \lambda$$



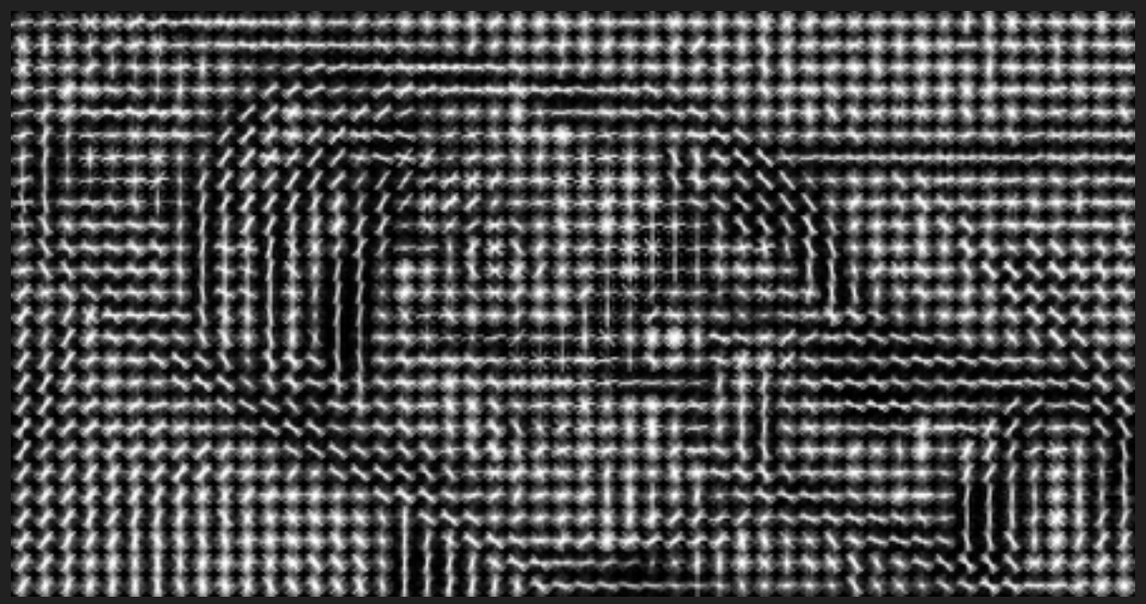
Paired Dictionary Learning

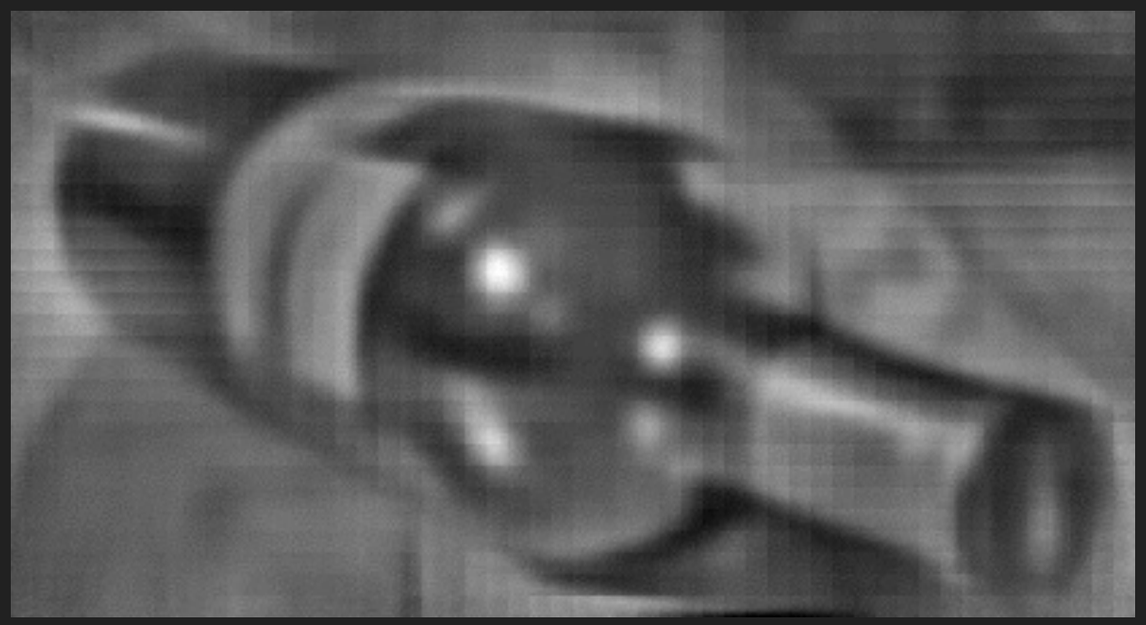
$$\arg\min_{U,V,\alpha} \sum_{i=1}^{N} ||x_i - U\alpha_i||_2^2 + ||y_i - V\alpha_i||_2^2$$

$$\text{s.t.} \quad ||\alpha_i||_1 \leq \lambda,\ ||U||_2^2 \leq \gamma_1,\ ||V||_2^2 \leq \gamma_2$$

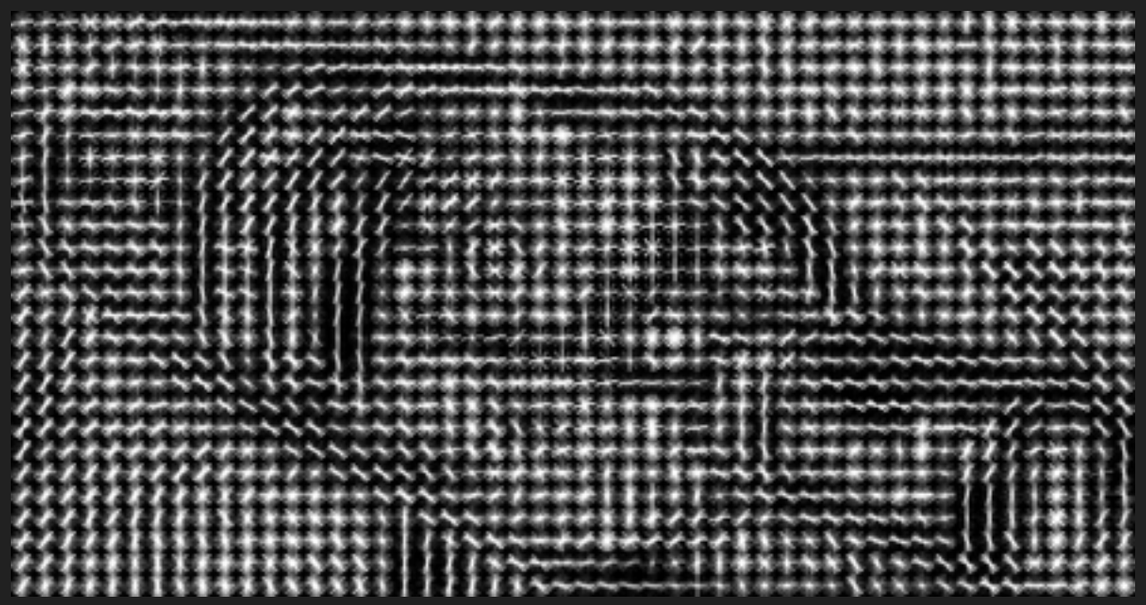
Just sparse coding!

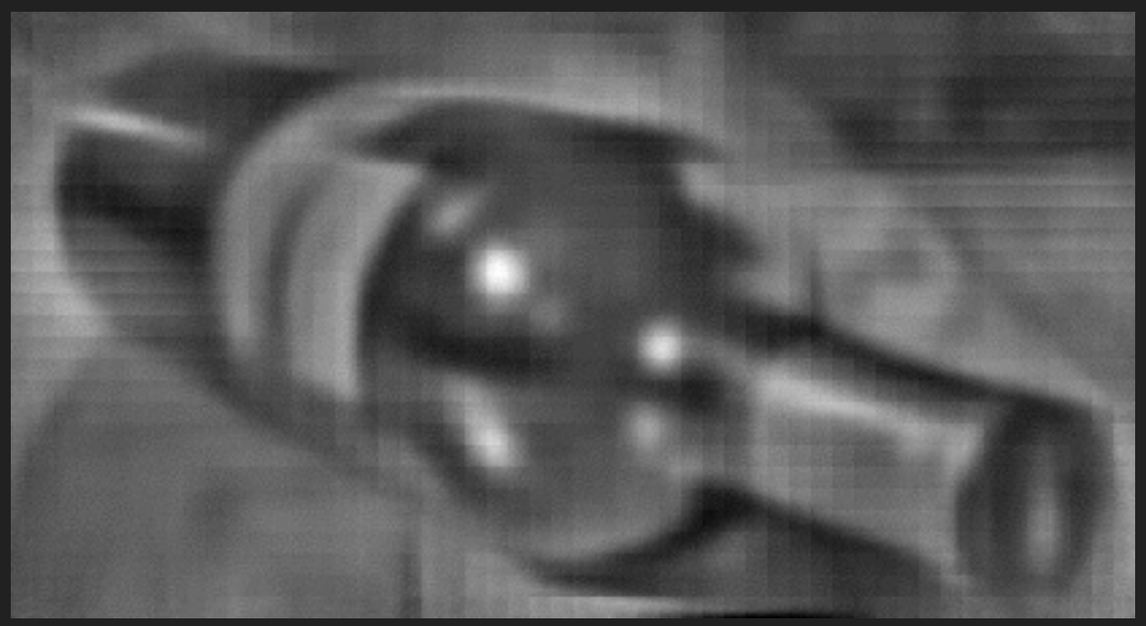
(Optimize using off the shelf solvers)





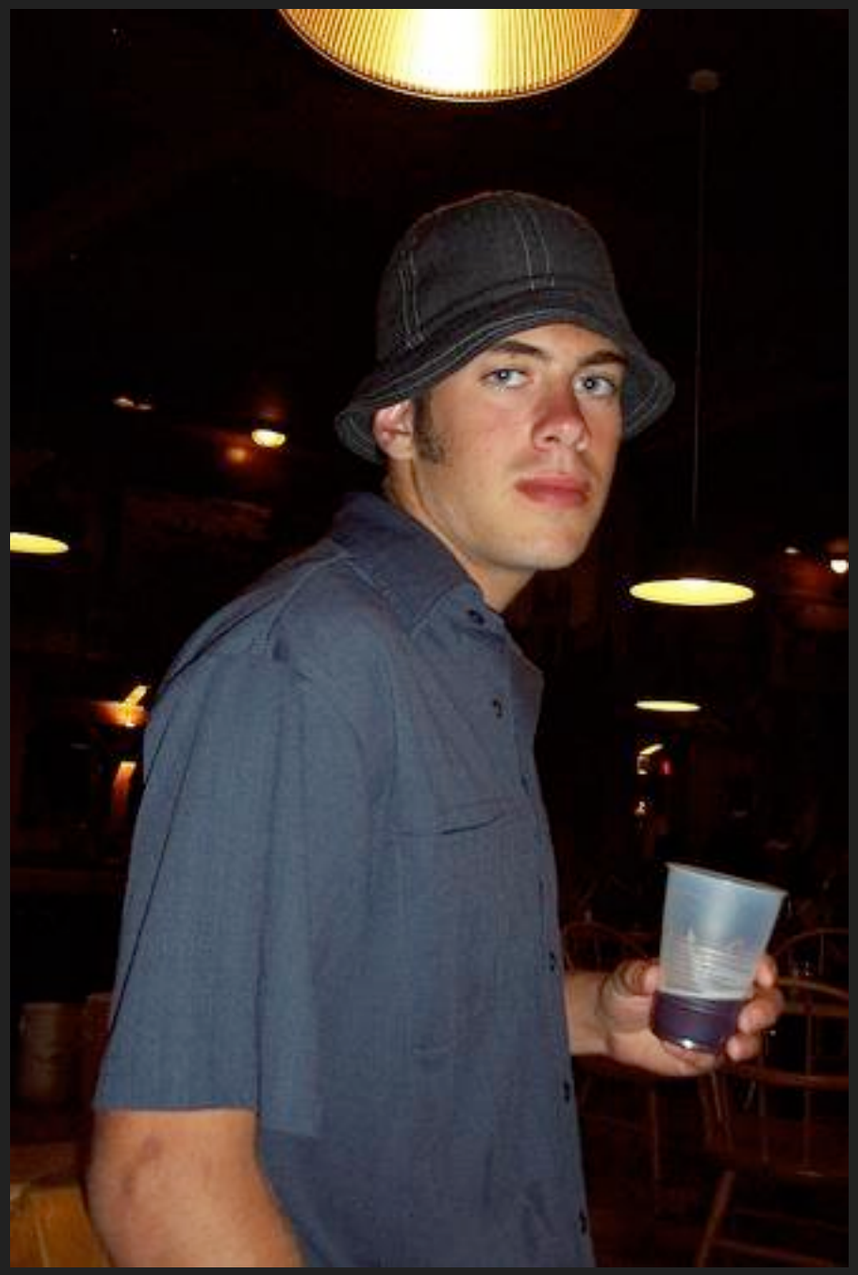
A microscope to view HOG





A microscope to view HOG

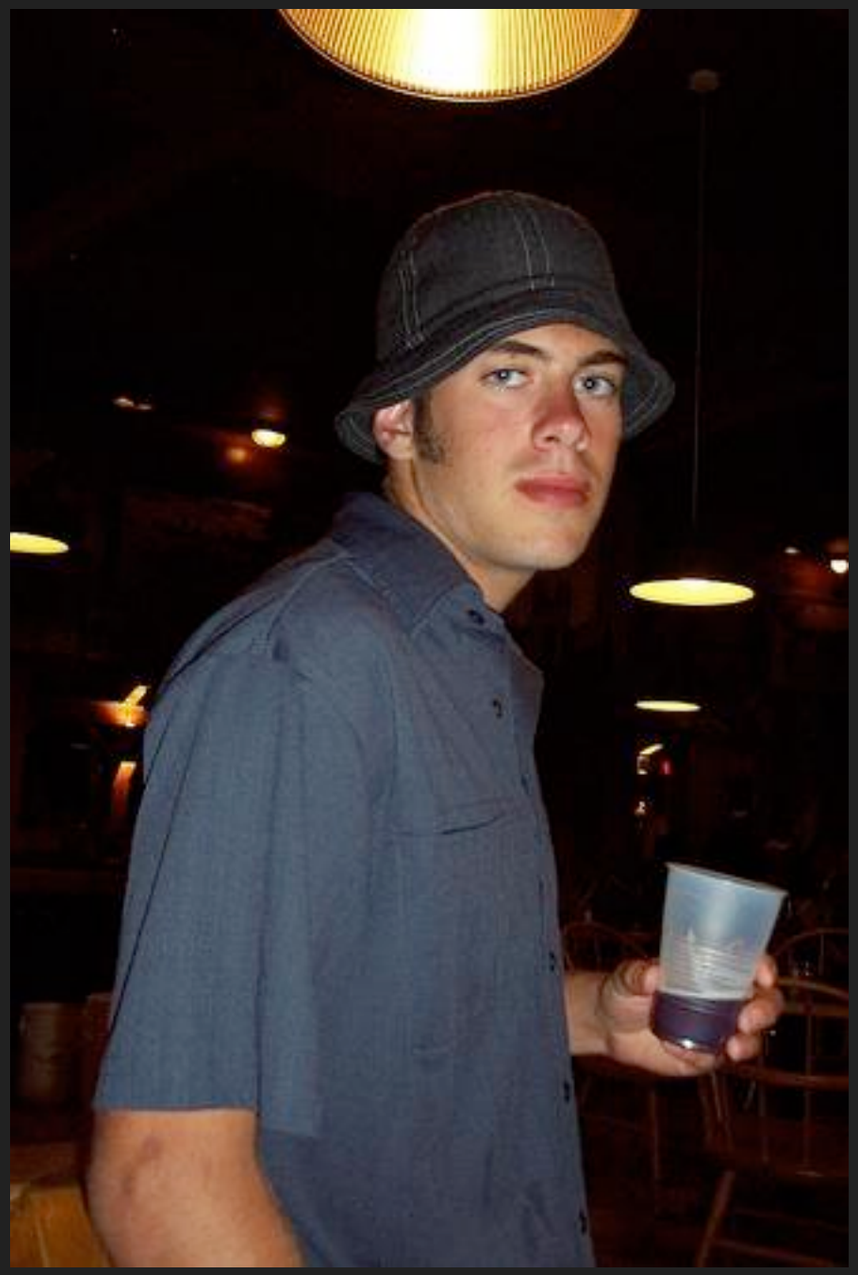
2x more intuitive



vs

Human Vision

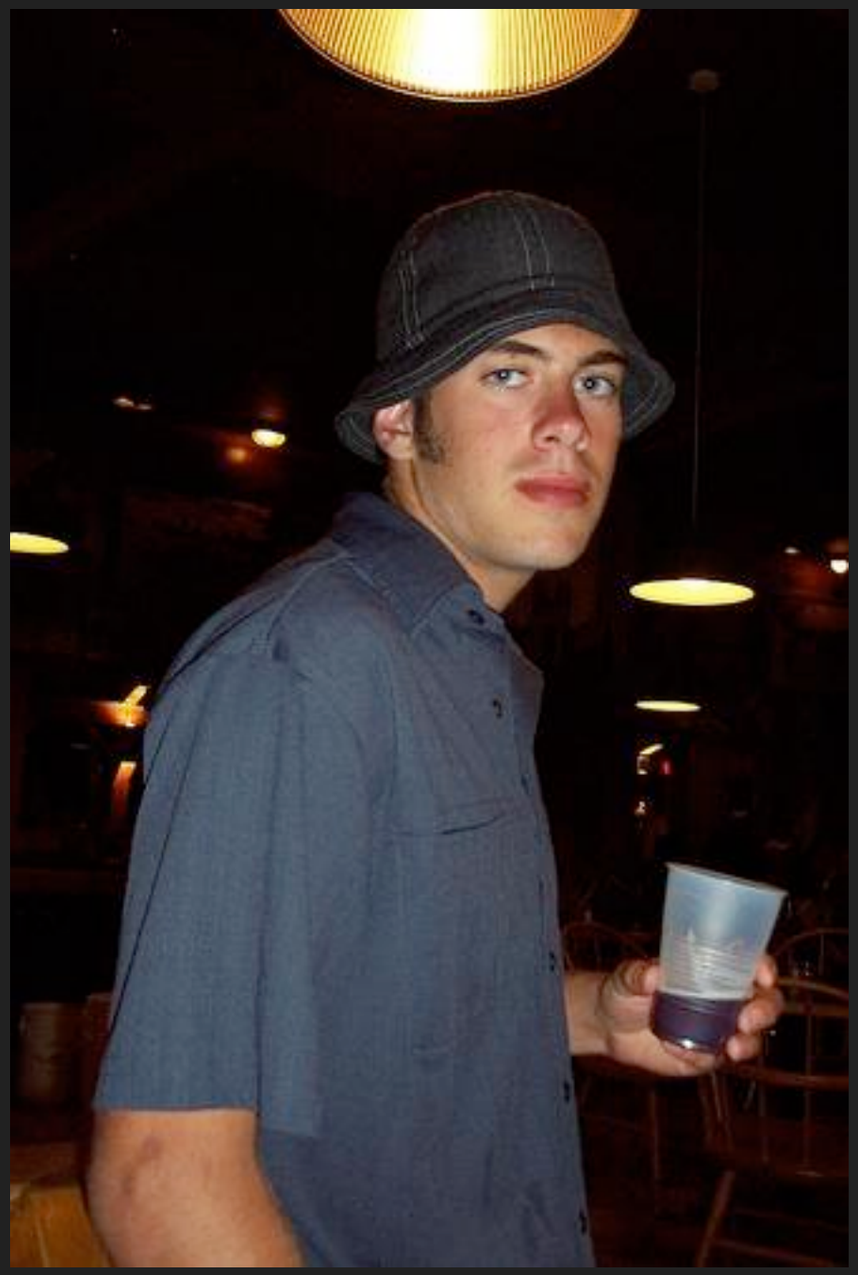
HOG Vision



Human Vision

vs

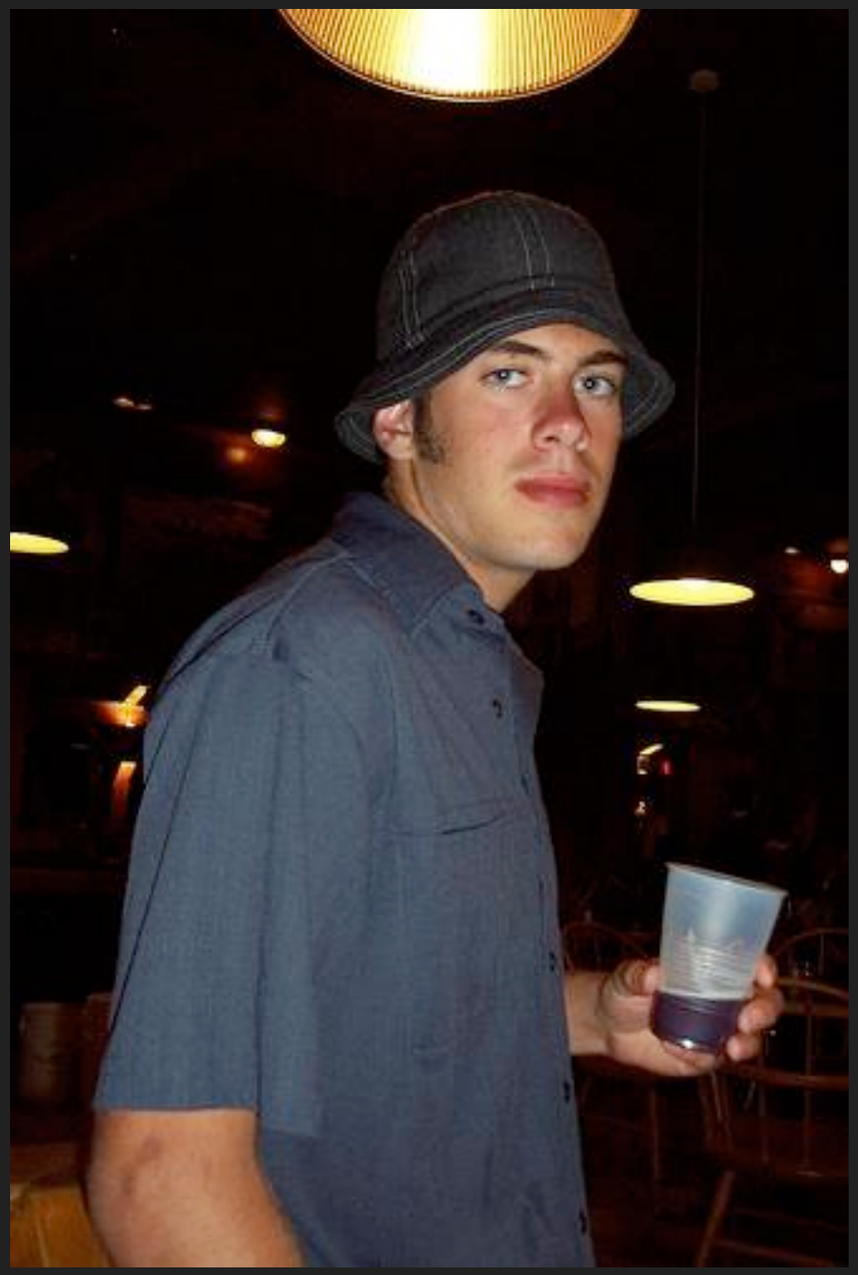
HOG Vision



vs

Human Vision

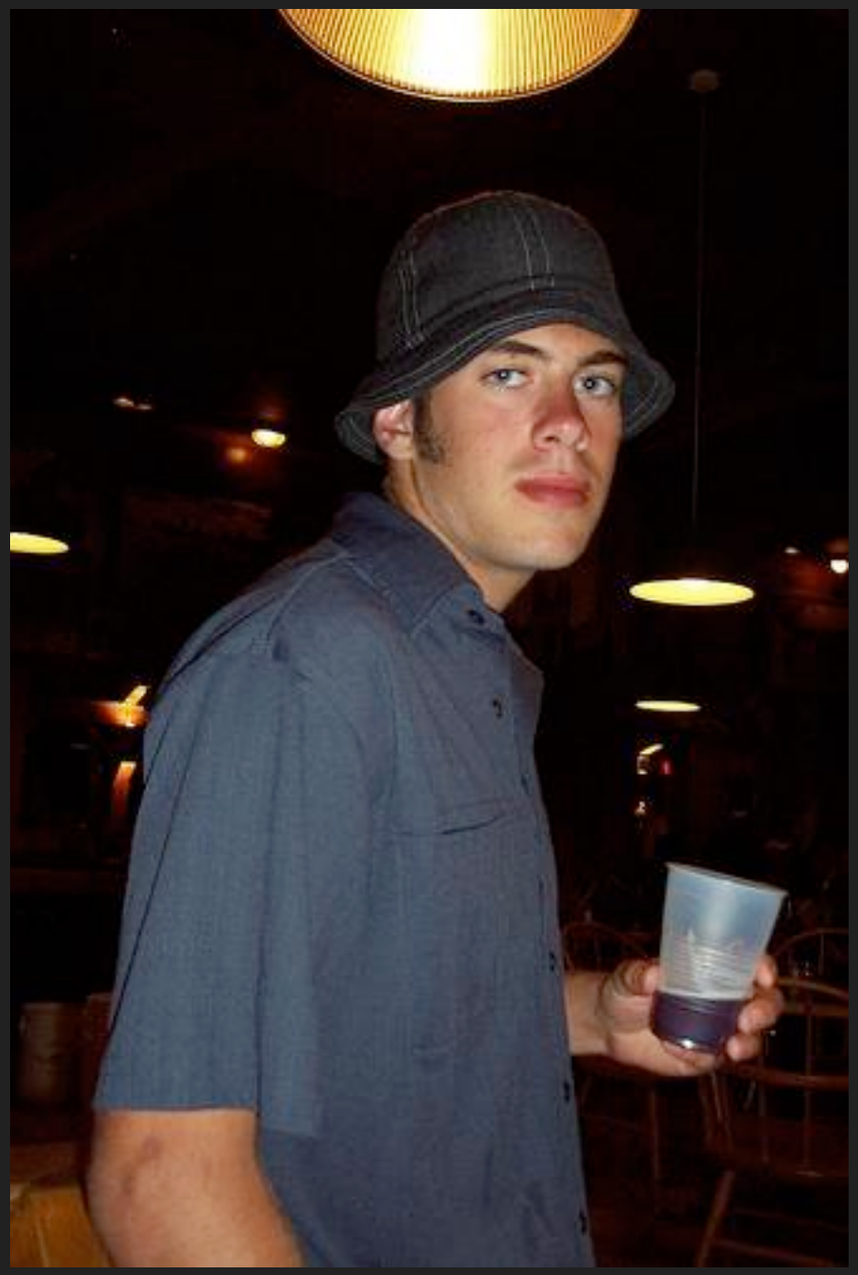
HOG Vision



vs

Human Vision

HOG Vision



vs

Human Vision

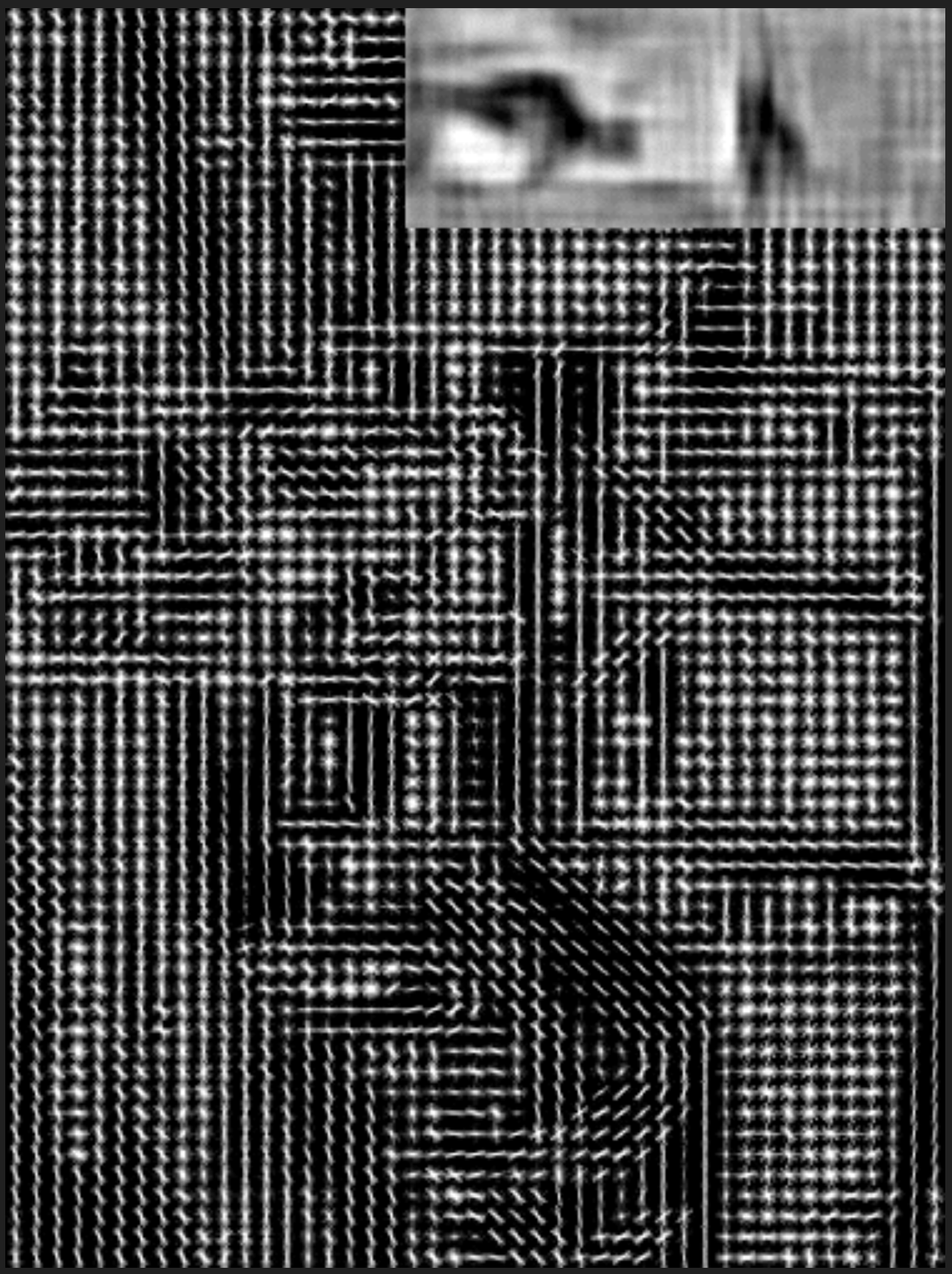
HOG Vision

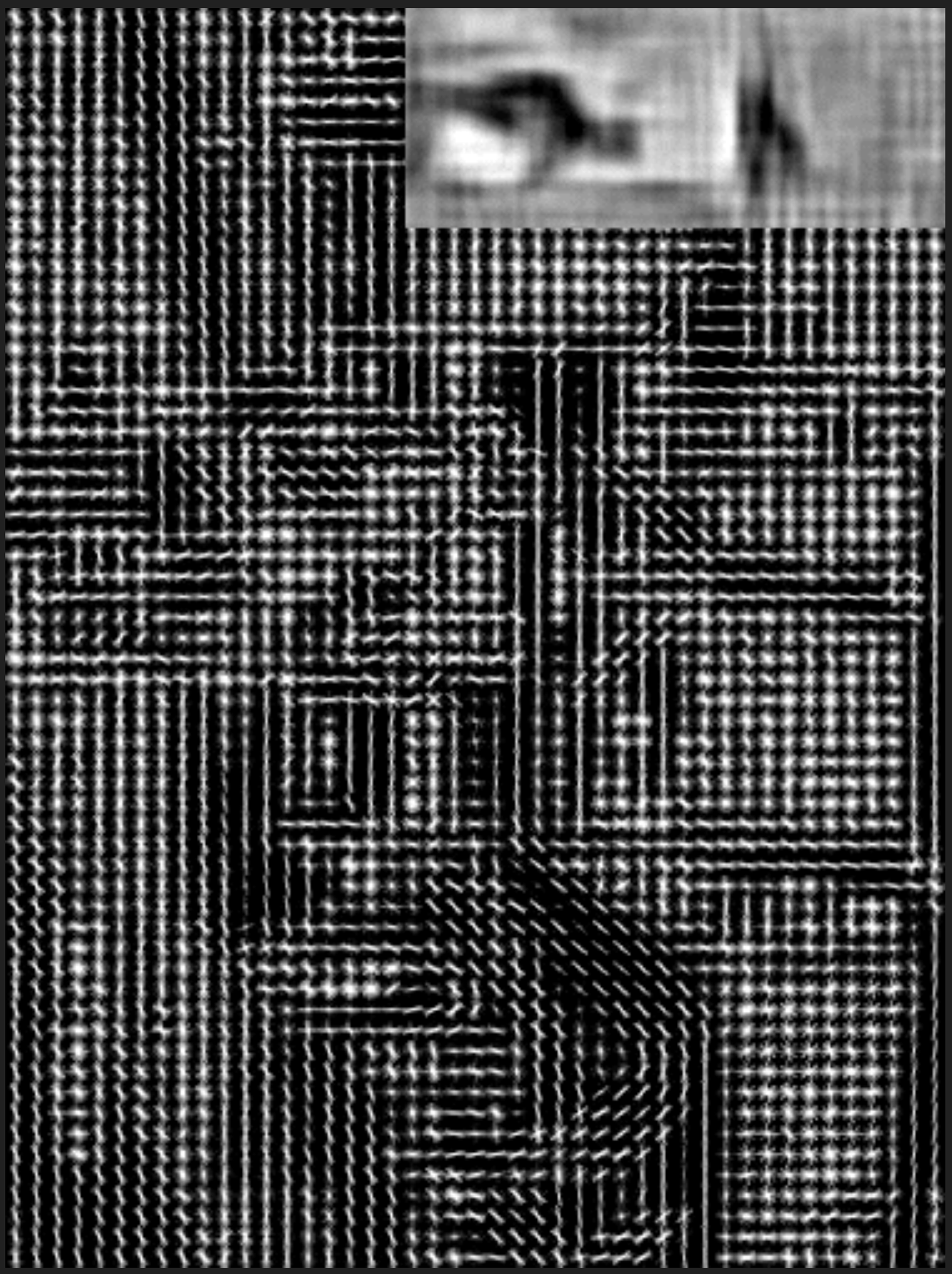

Car

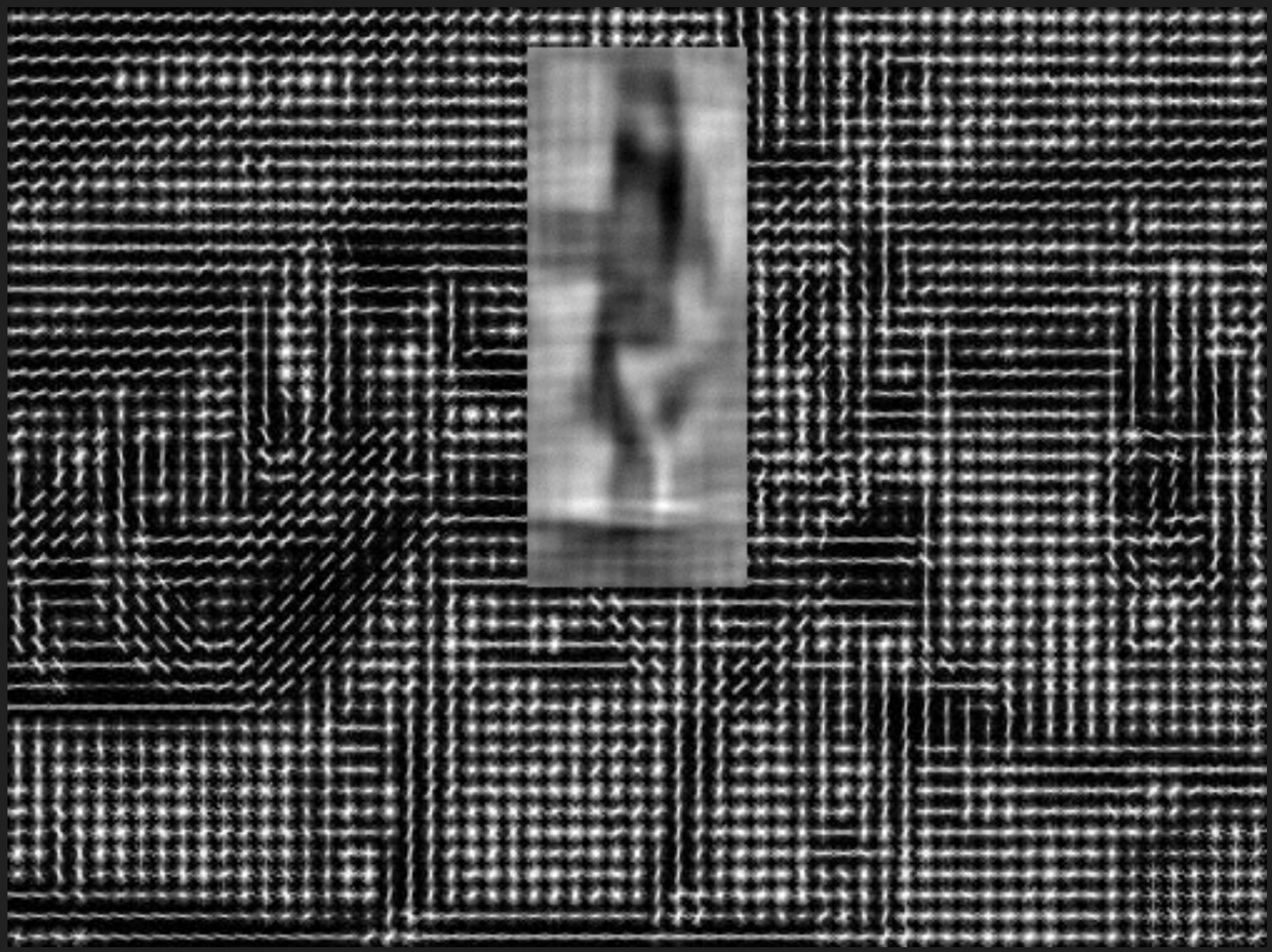

Image
HOG
SVM


Image
HOG
SVM


Image
HOG
SVM
Human





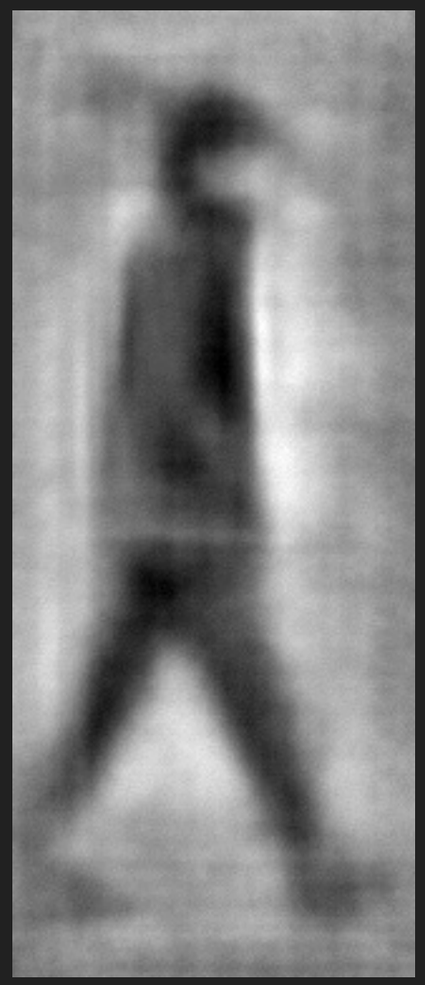


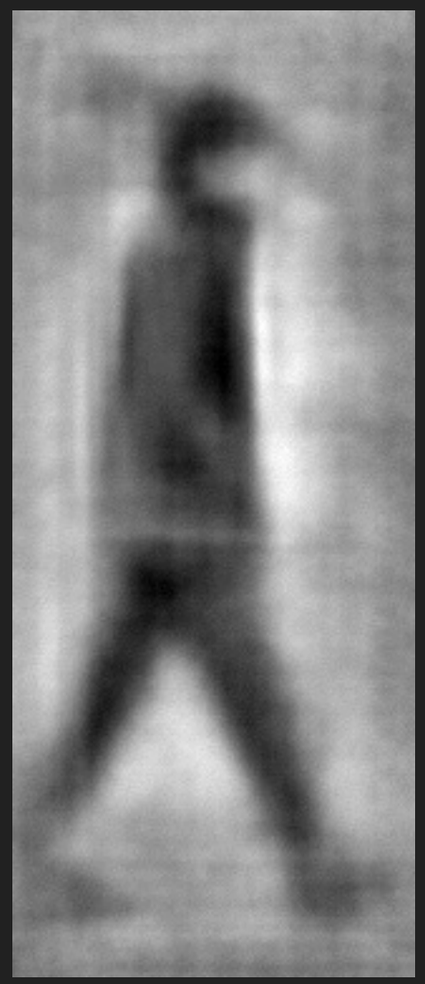

9
DEC 15 2006

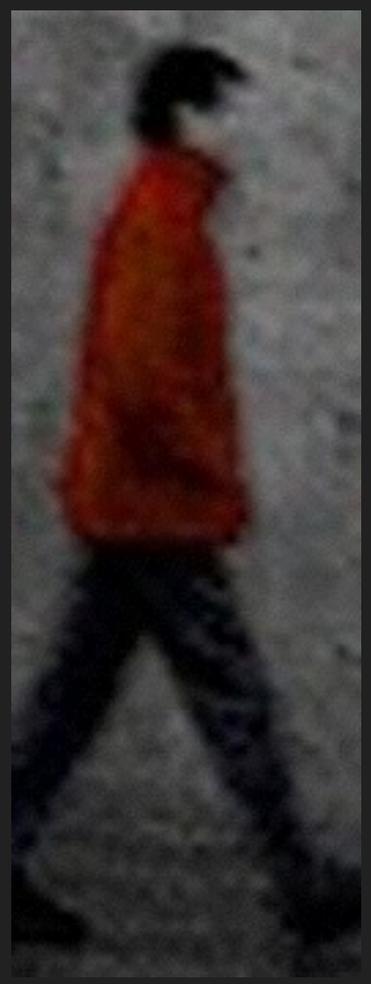
The HOGgles Challenge

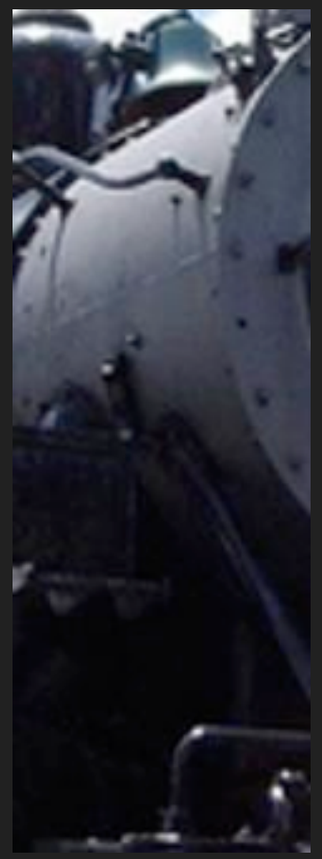
The HOGgles Challenge

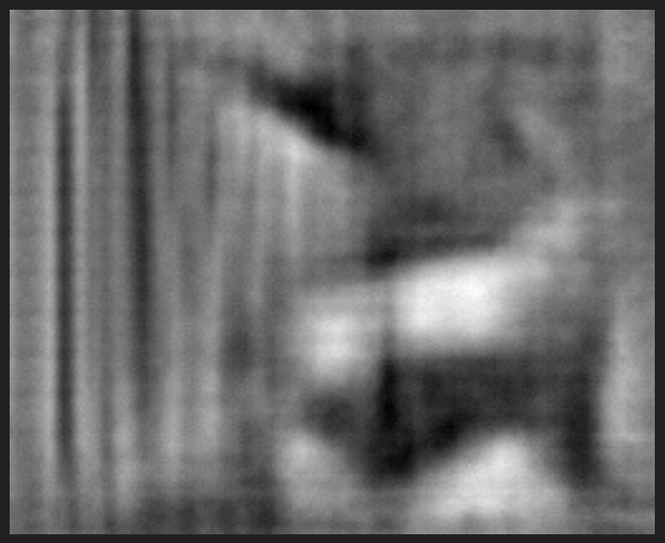
Clap your hands when you see a person

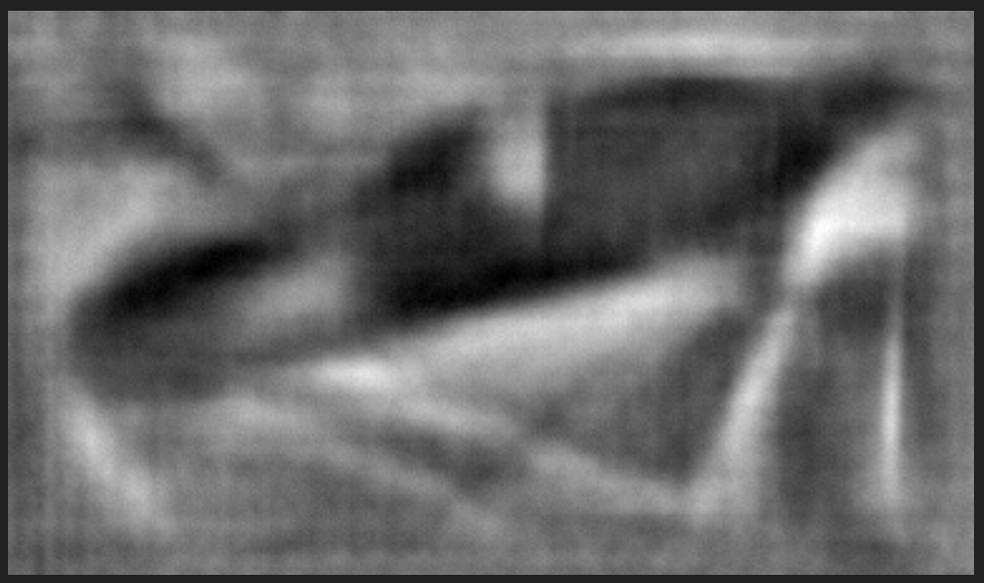






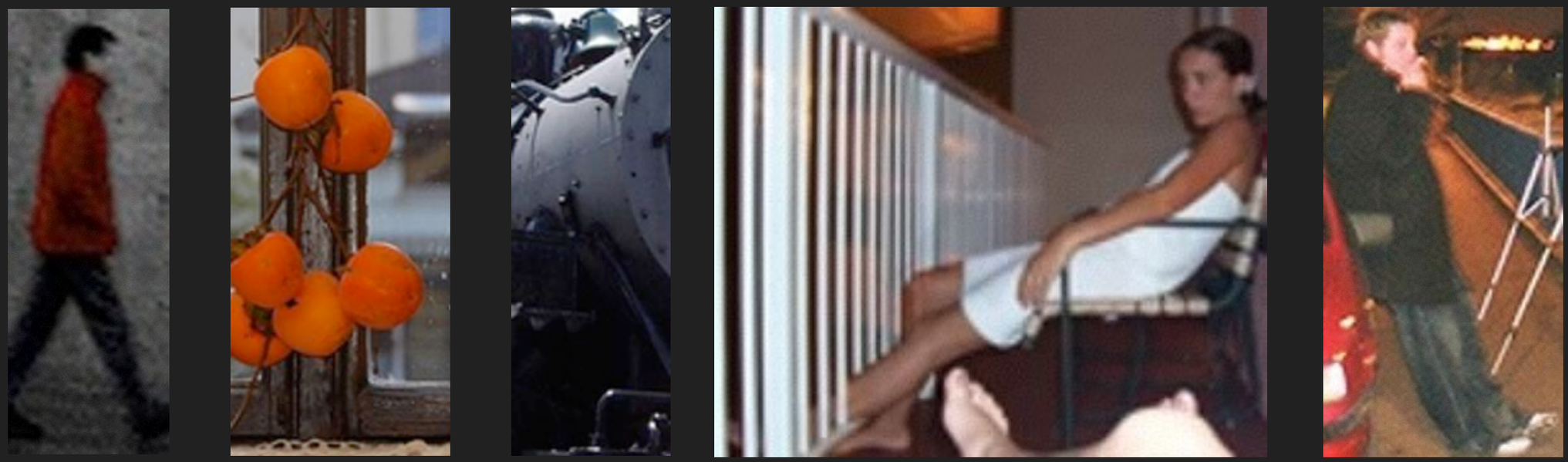




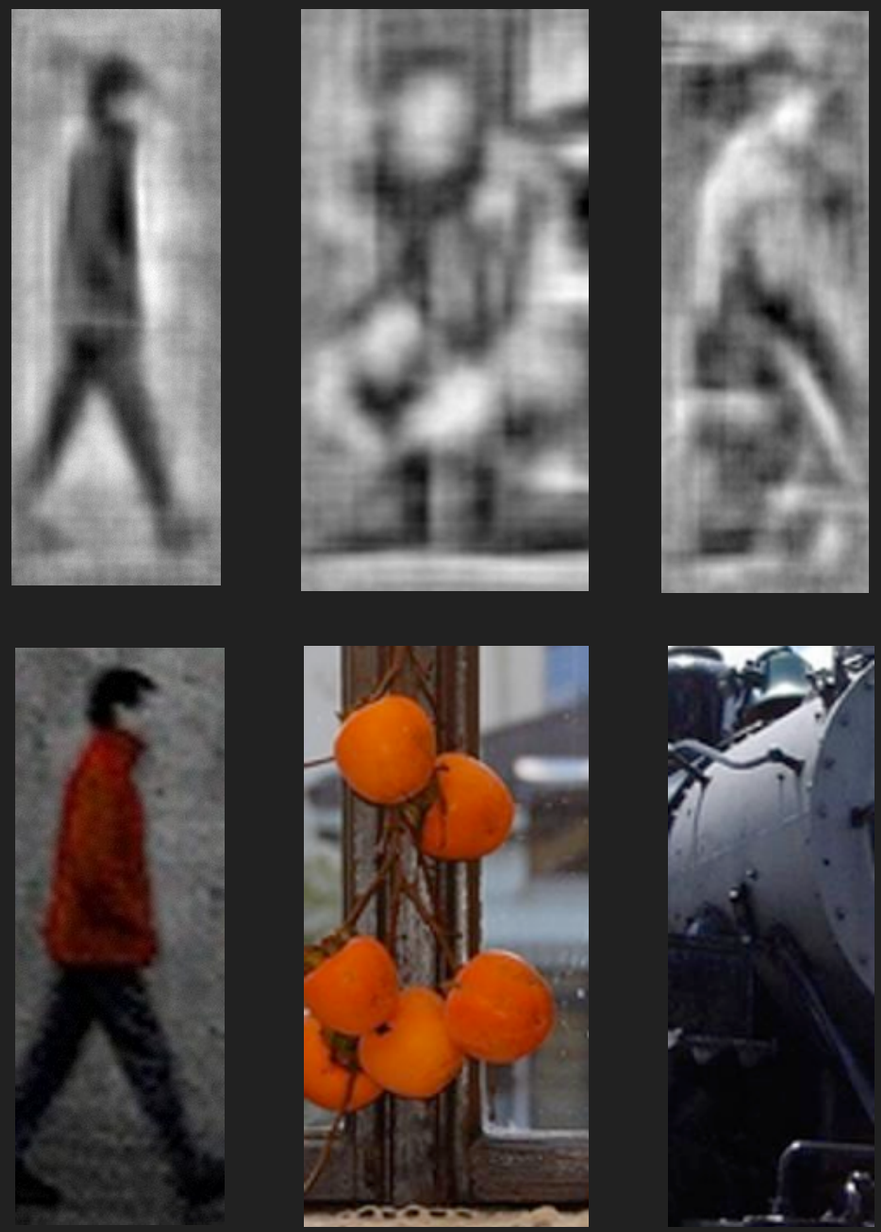


The HOGgles Challenge

The HOGgles Challenge



The HOGgles Challenge



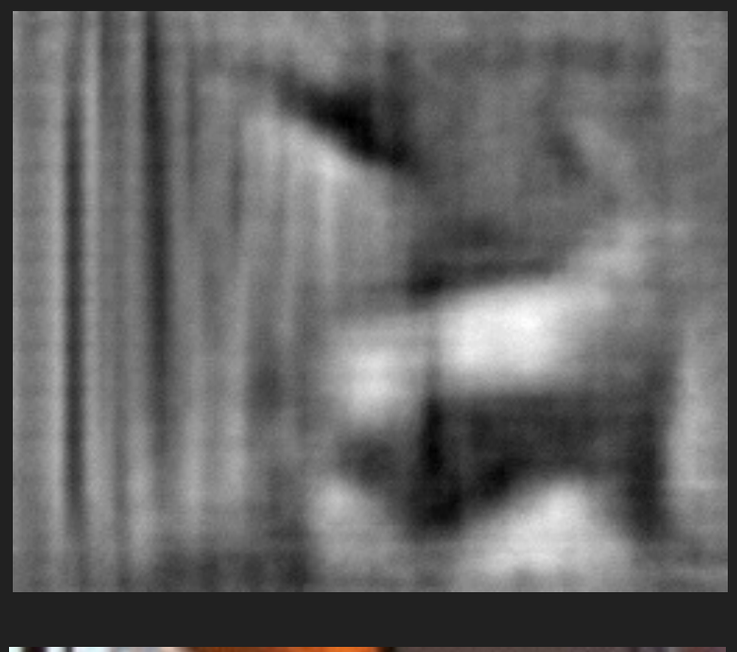
Humans detect

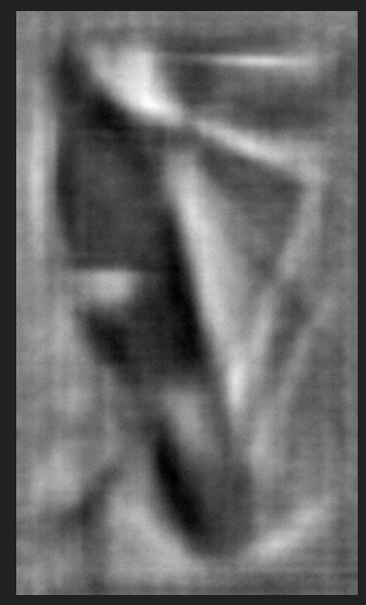
&

DPMs detect

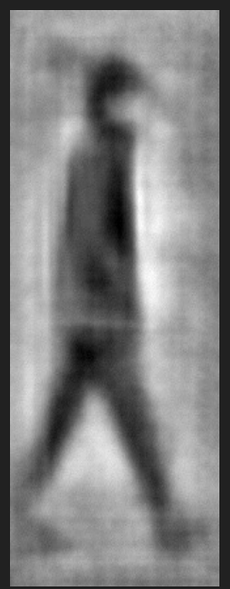
The HOGgles Challenge

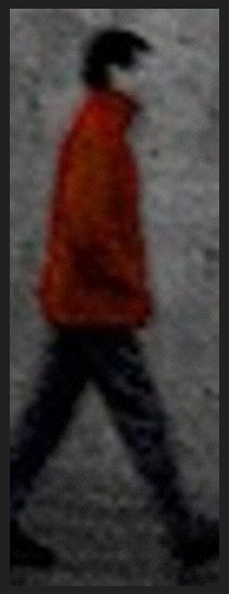
Humans miss
&
DPM miss

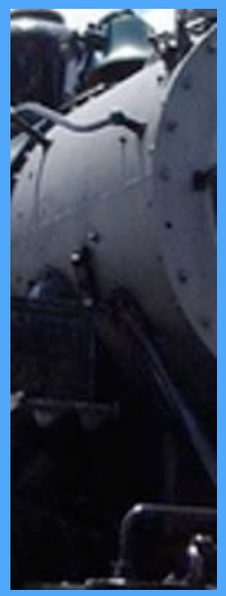




The HOGgles Challenge



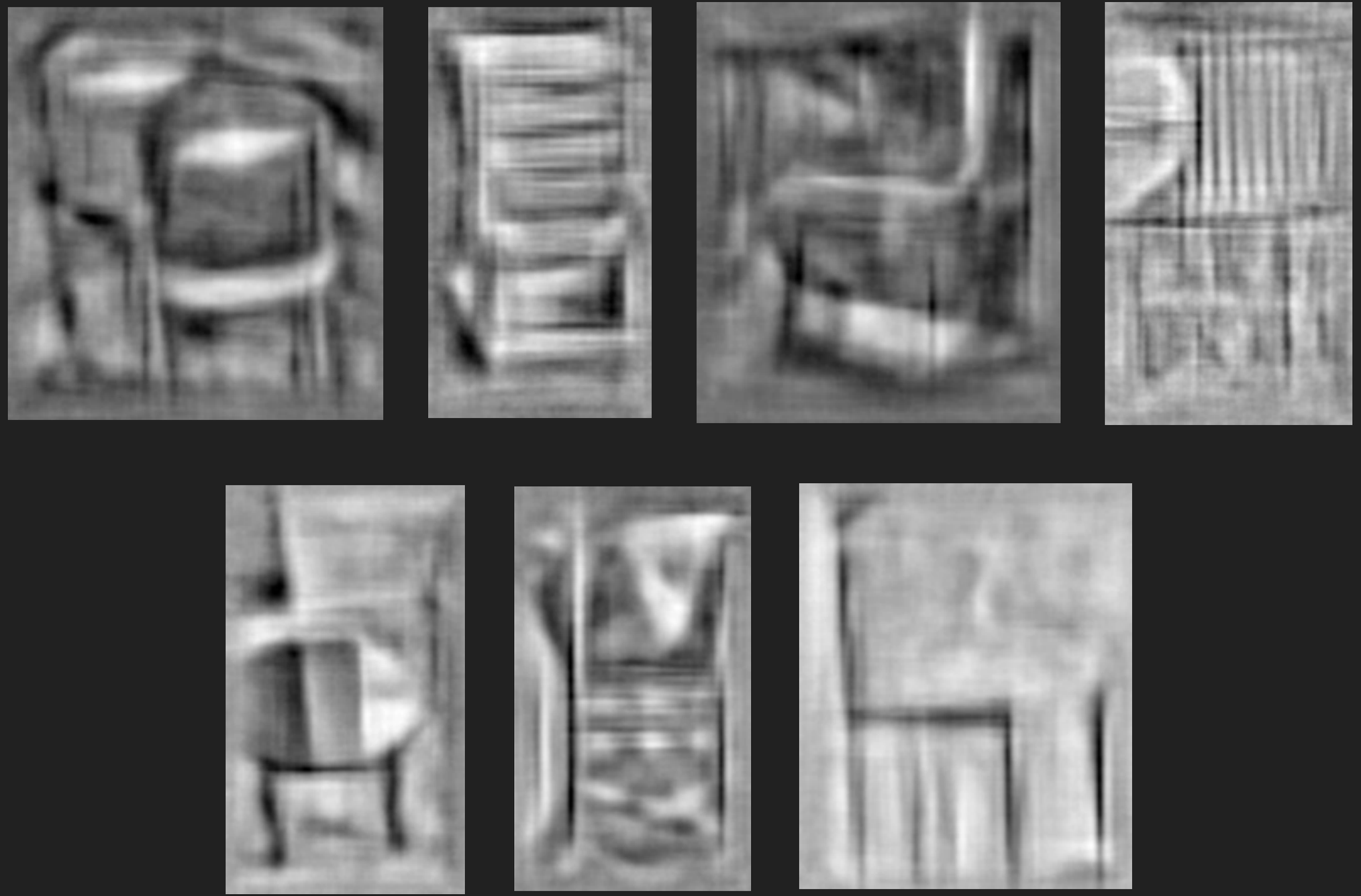




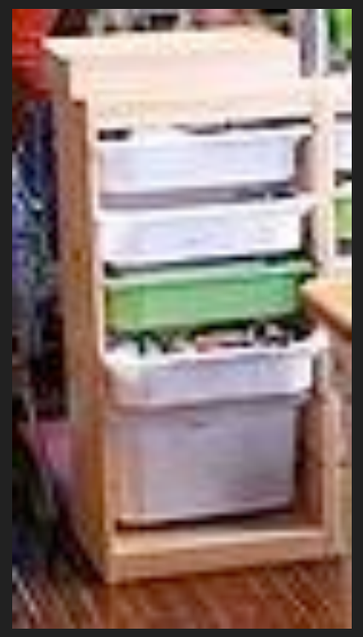
Common False Alarms

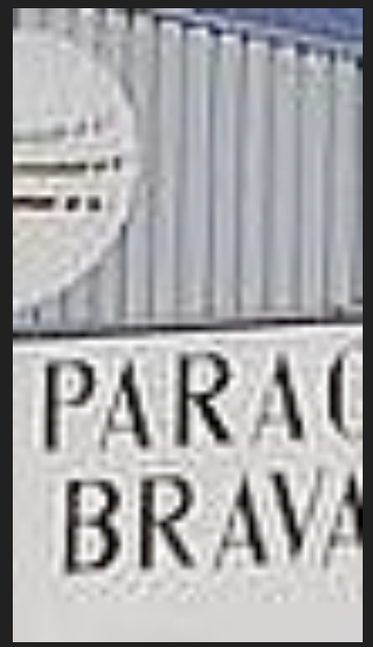
Common Missed Detections

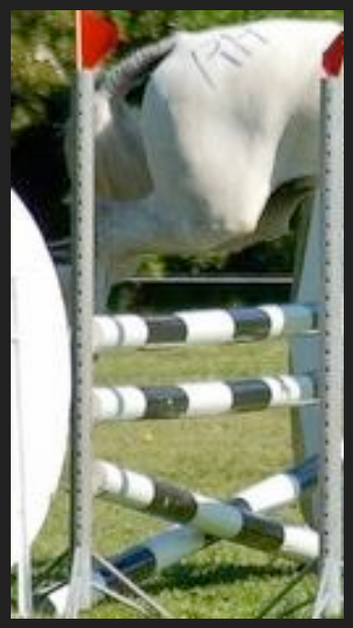
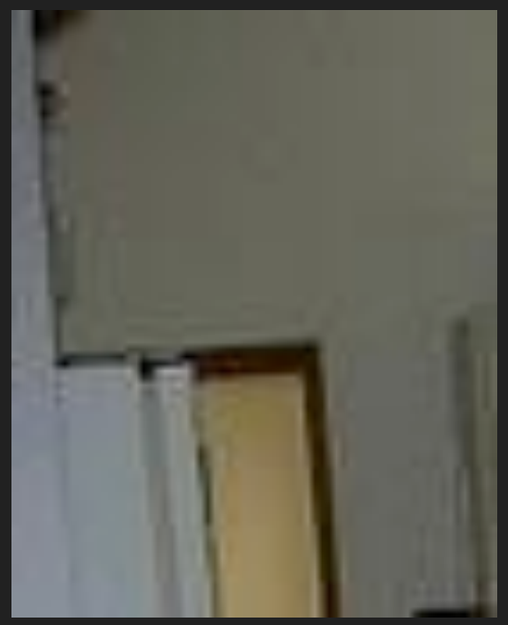
Chair Detections



Chair Detections

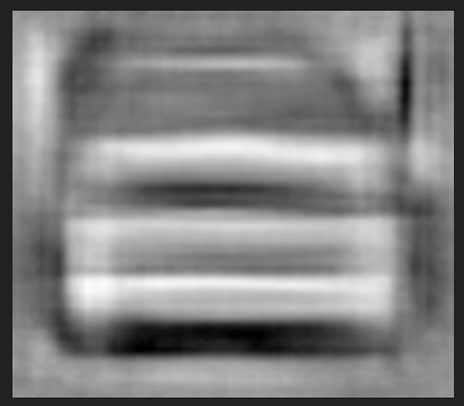





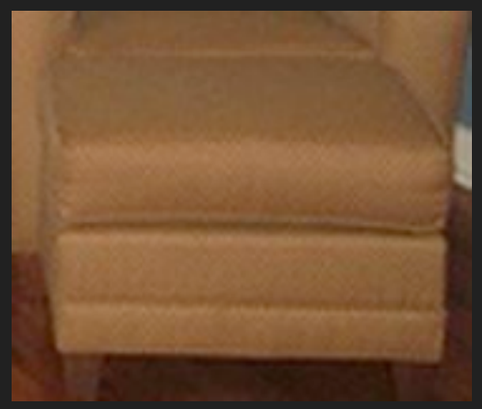
Car Detections



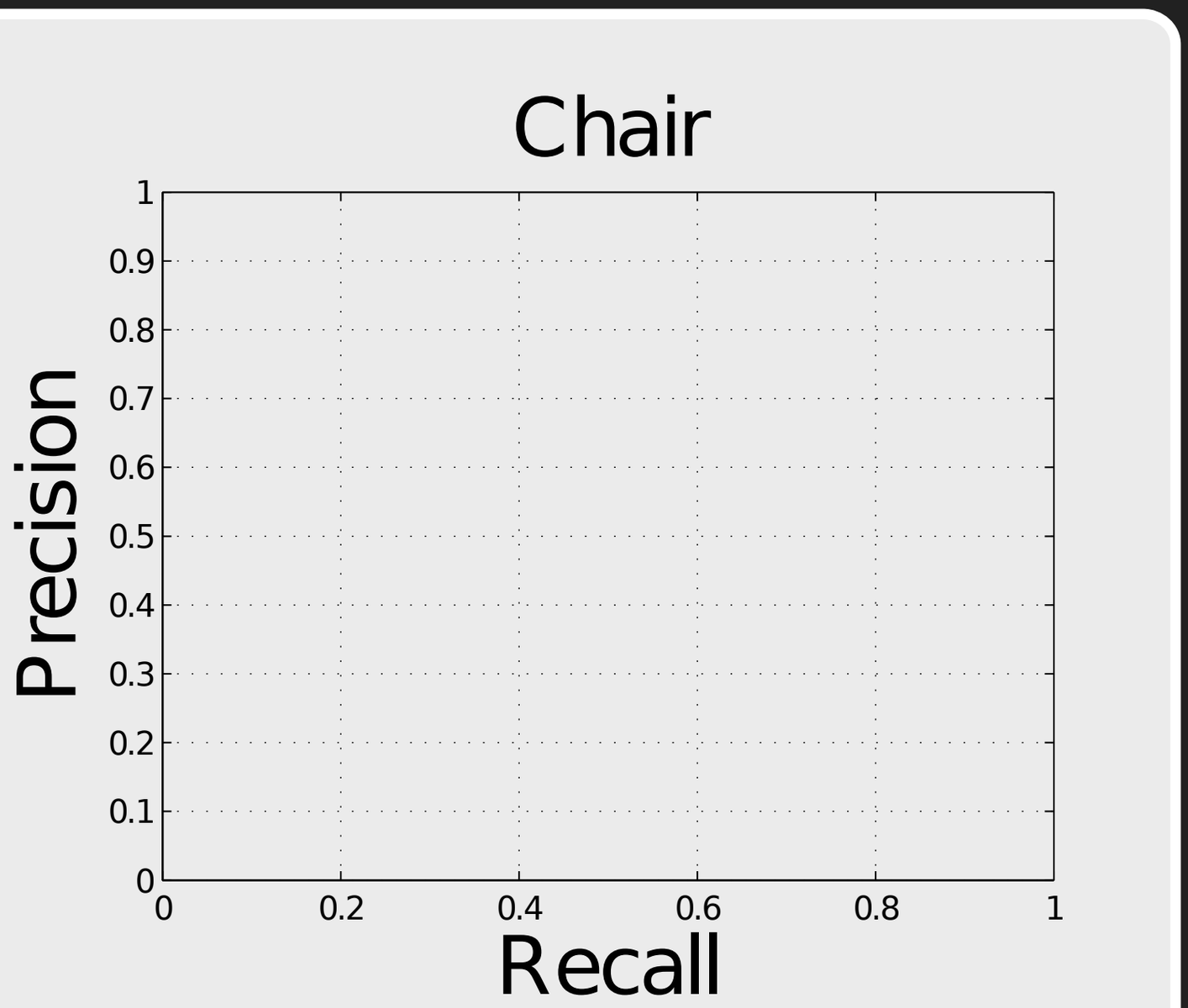




Car Detections

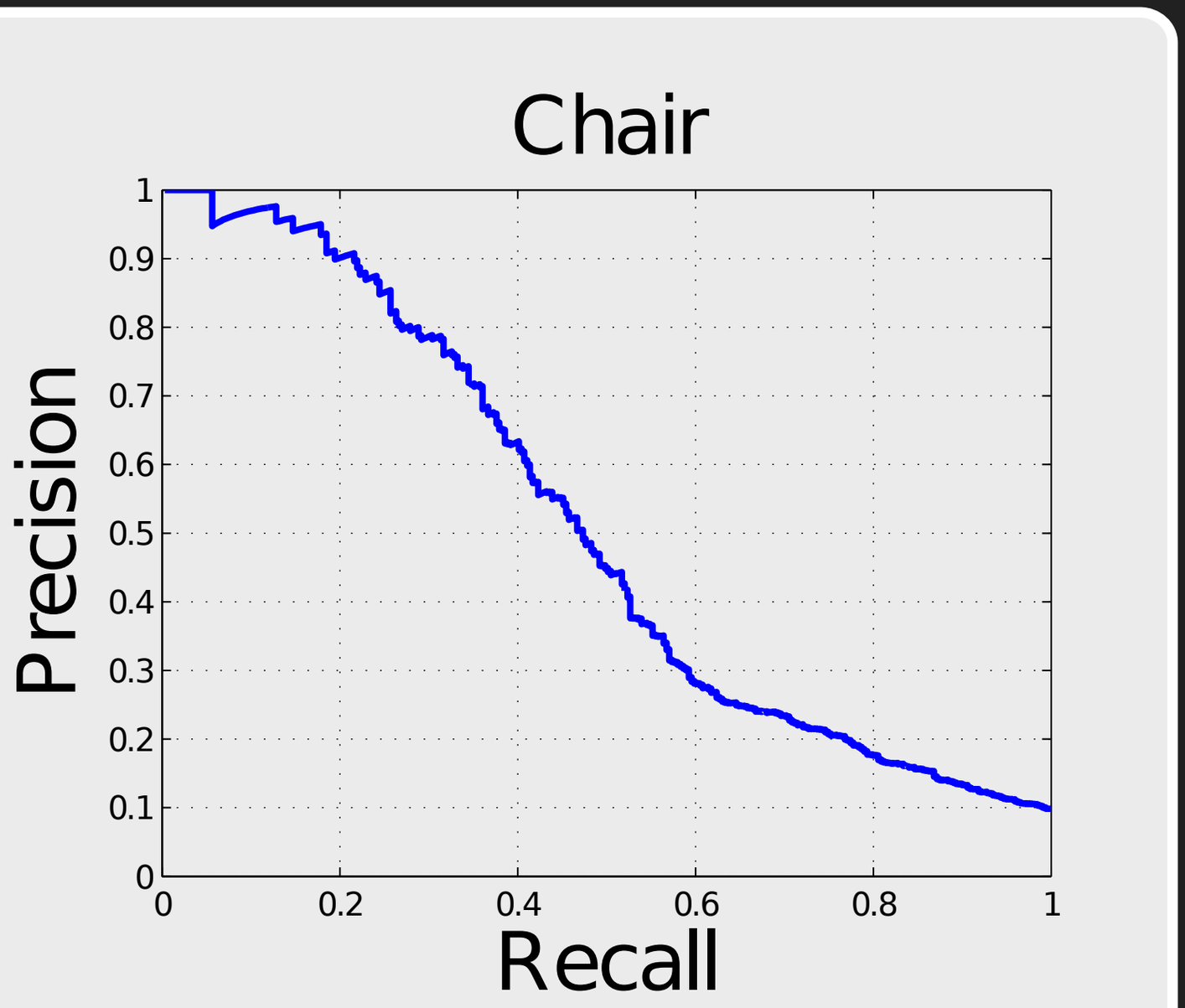



HOG+Human Detector

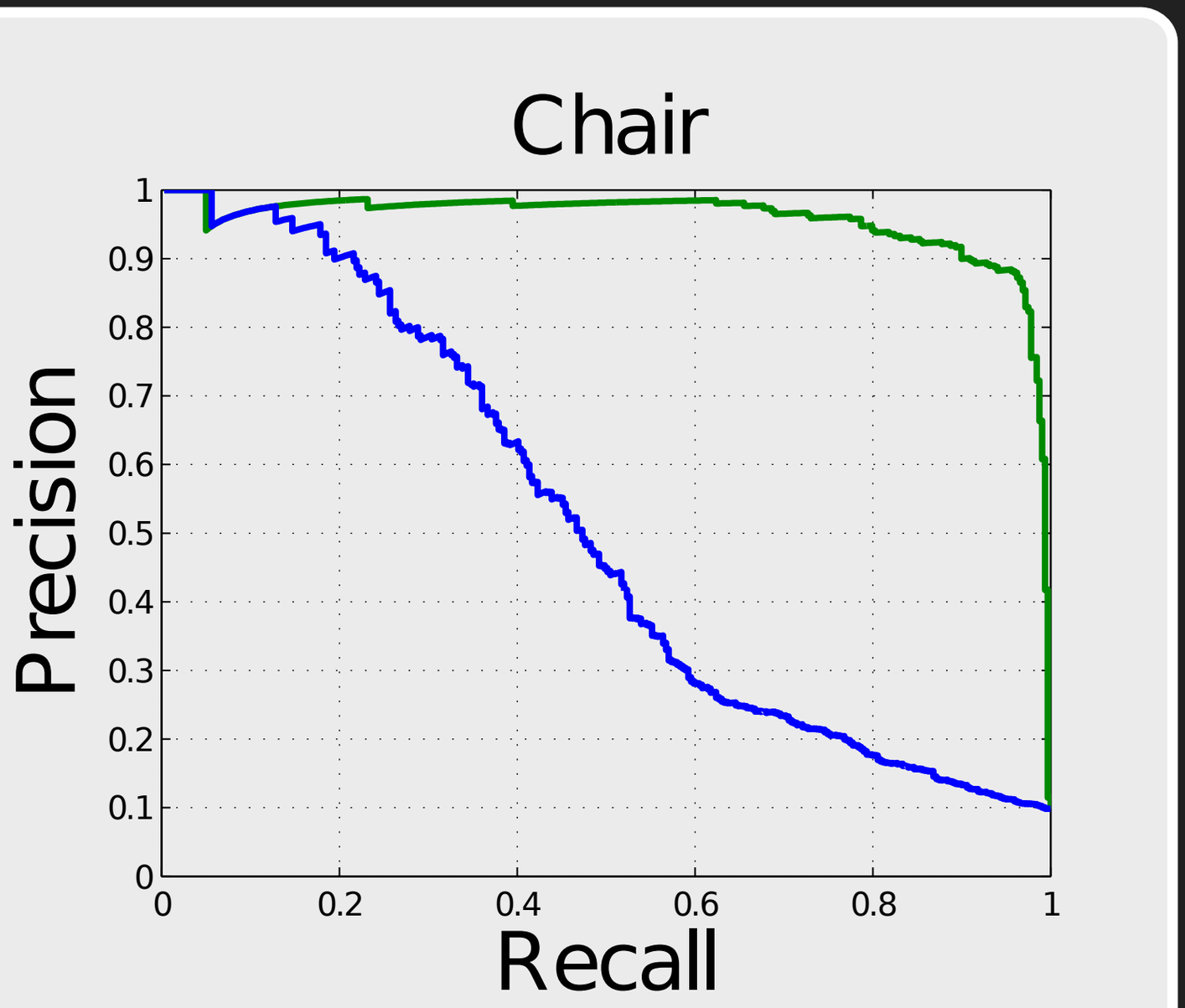


HOG+Human Detector



HOG+DPM

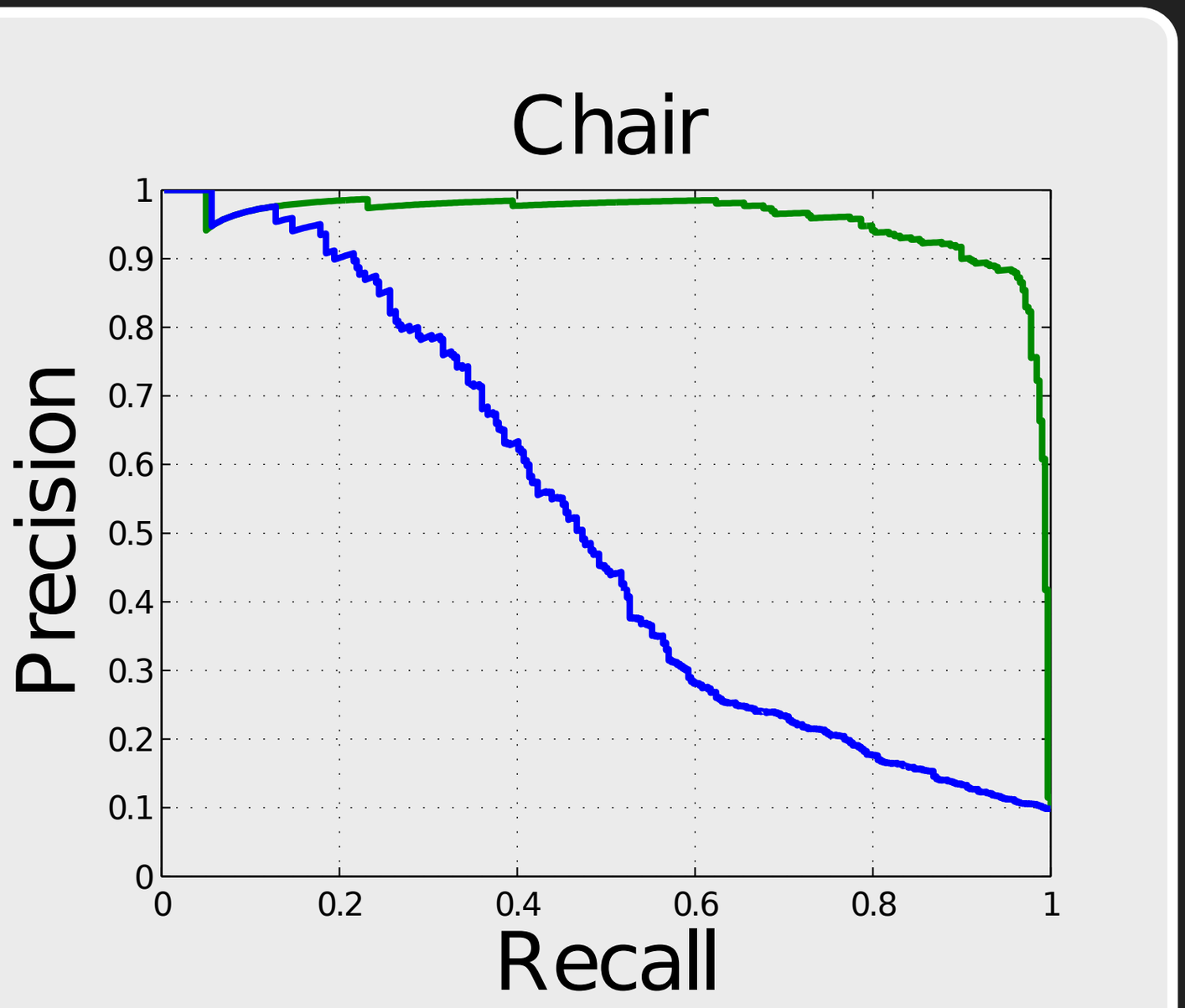
HOG+Human Detector



HOG+DPM

RGB+Human

HOG+Human Detector

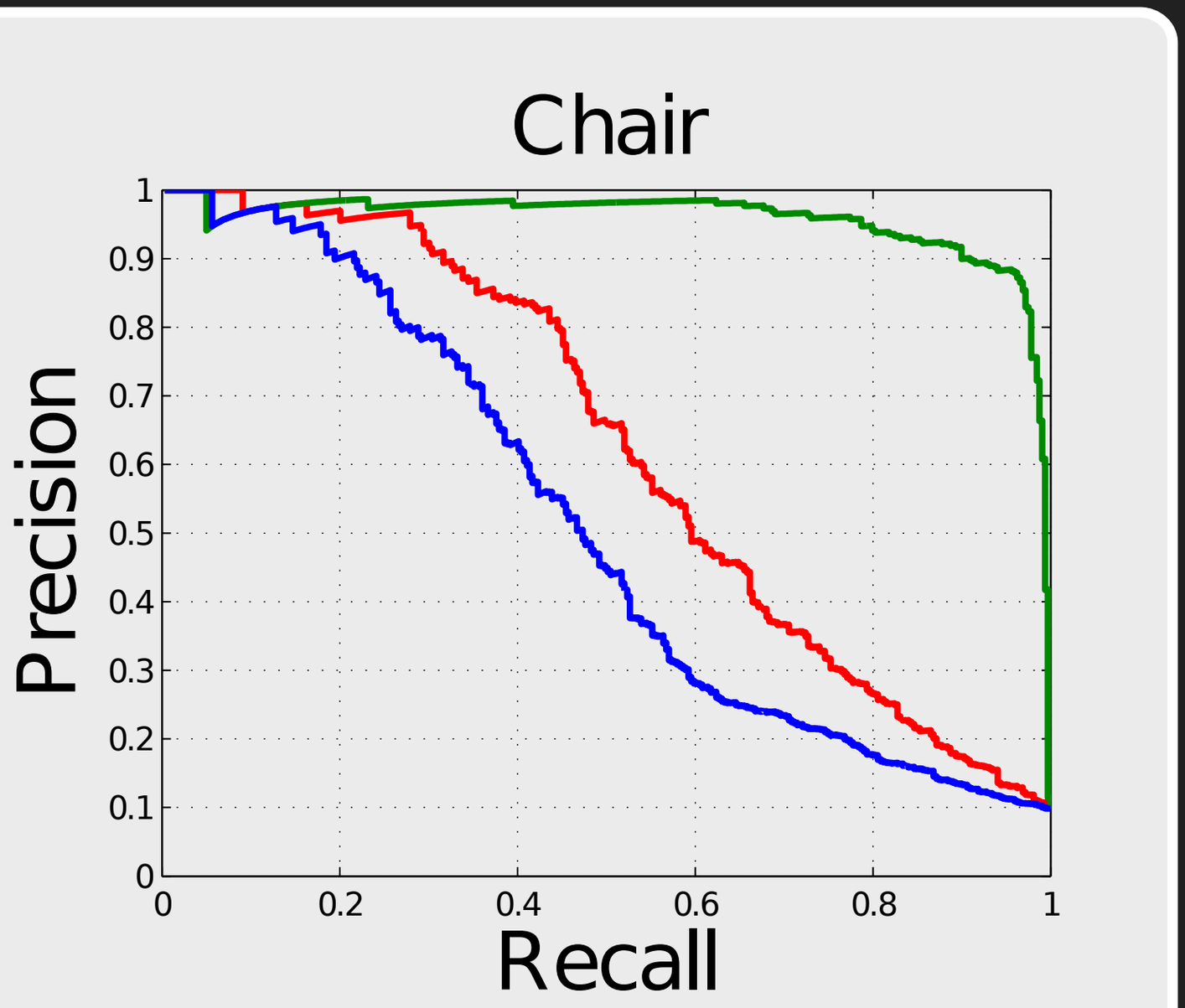


HOG+DPM

RGB+Human

HOG+Human

HOG+Human Detector

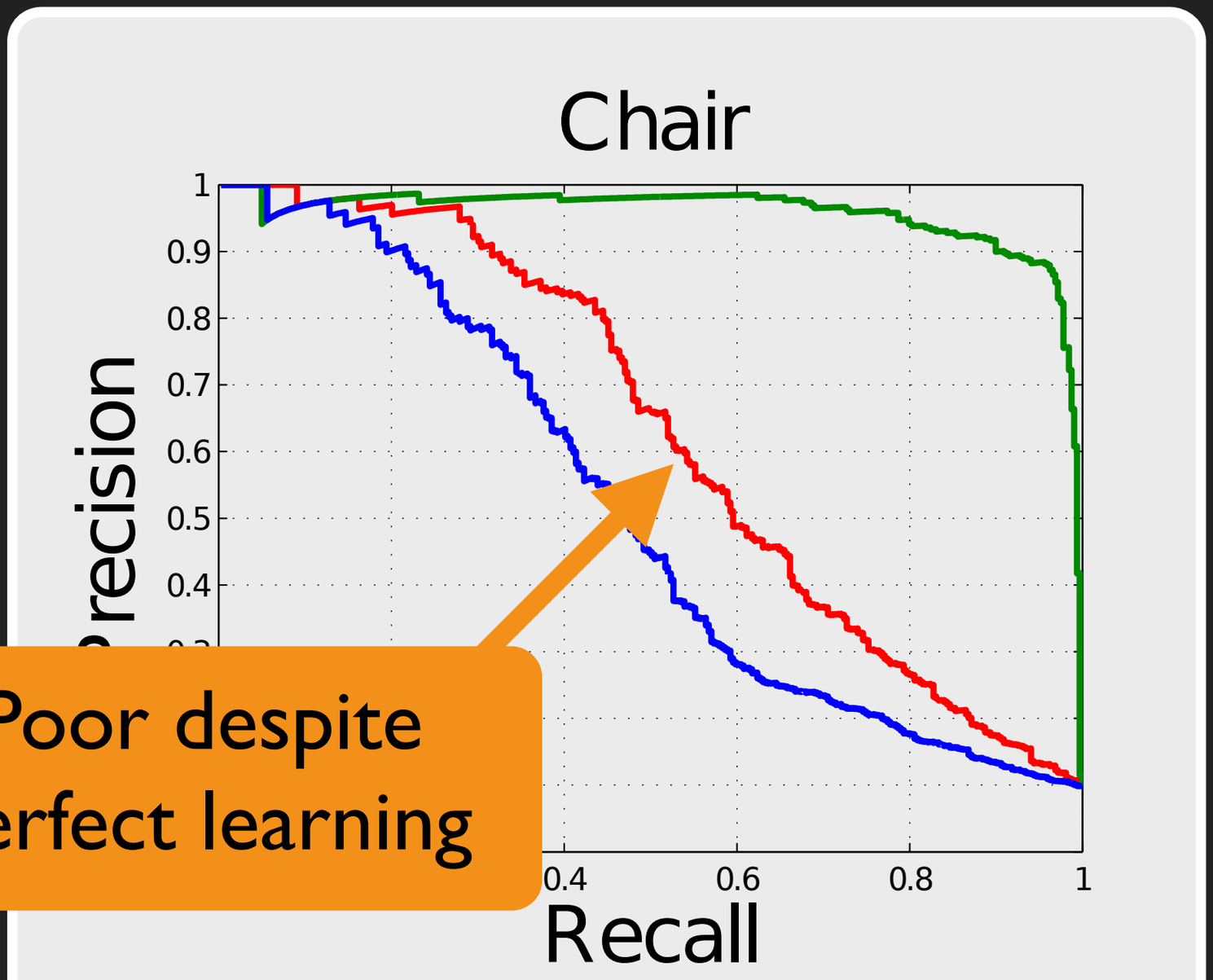


HOG+DPM

RGB+Human

HOG+Human

HOG+Human Detector



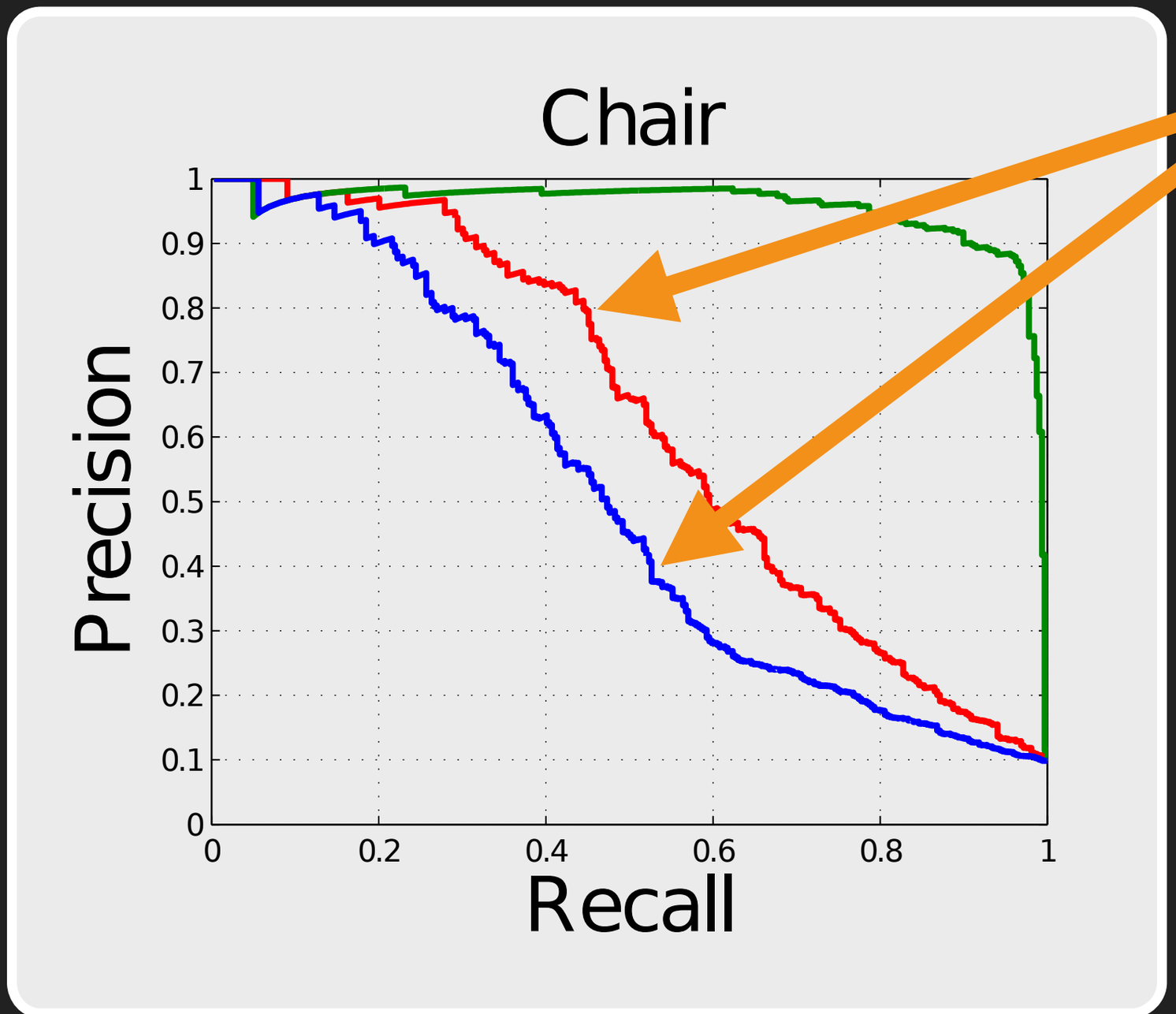
Poor despite perfect learning

HOG+DPM

RGB+Human

HOG+Human

HOG+Human Detector



0.5 correlation coefficient

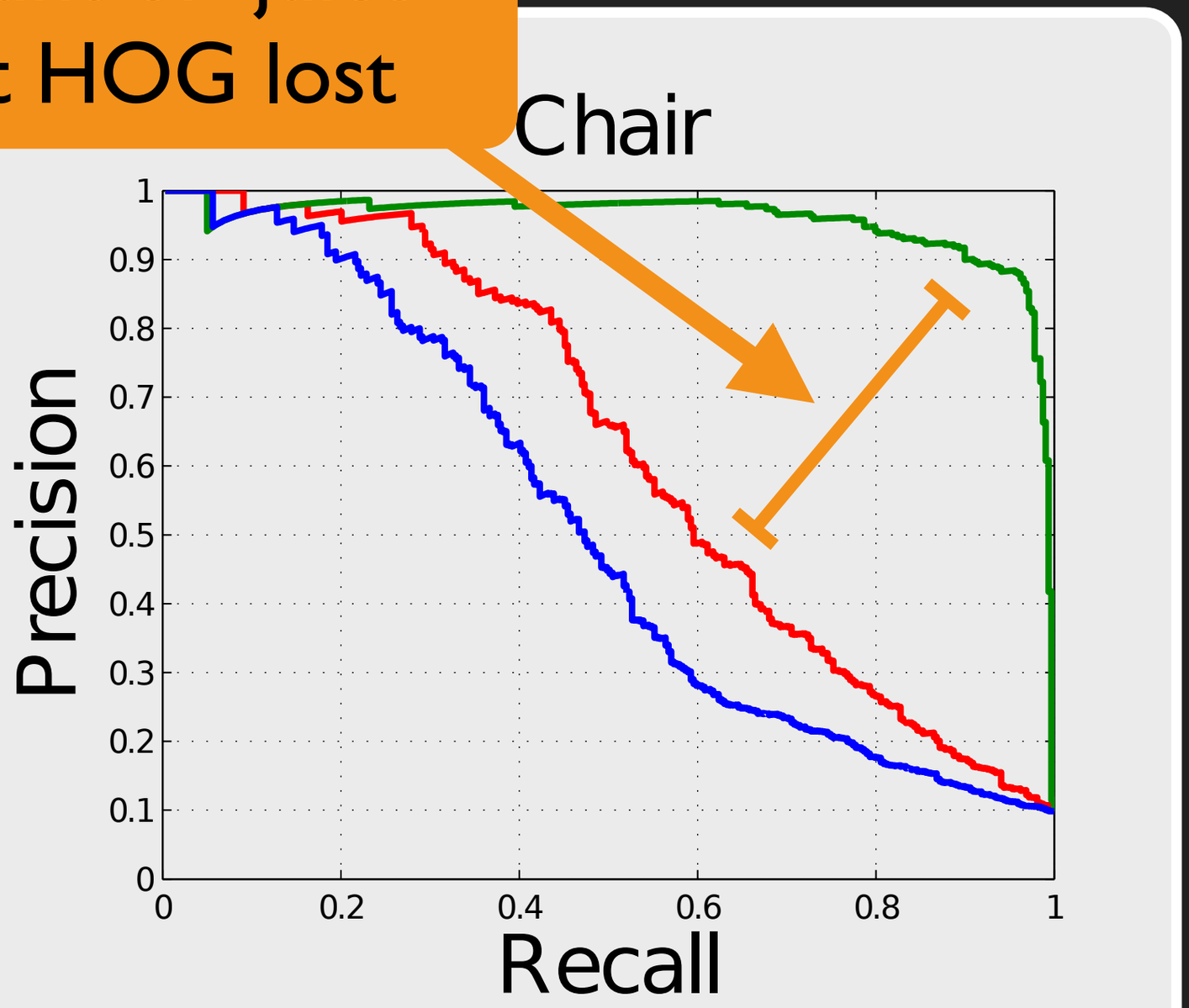
HOG+DPM

RGB+Human

HOG+Human

HOG+Human Detector

Amount of "juice" that HOG lost

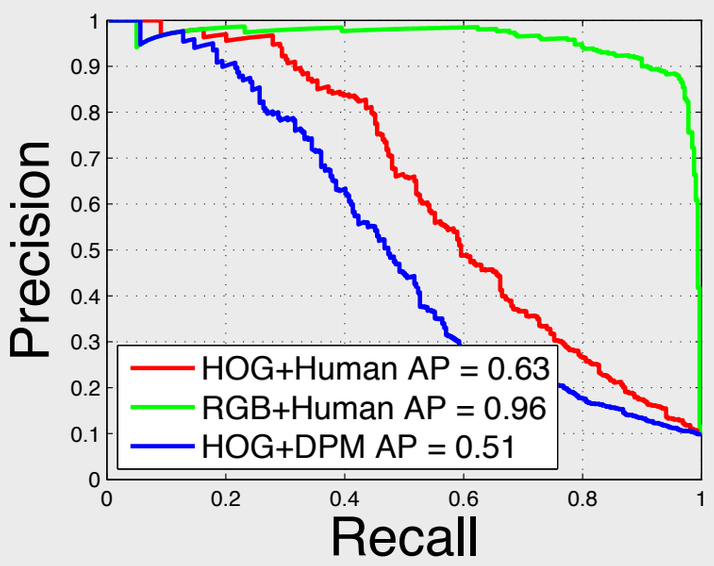


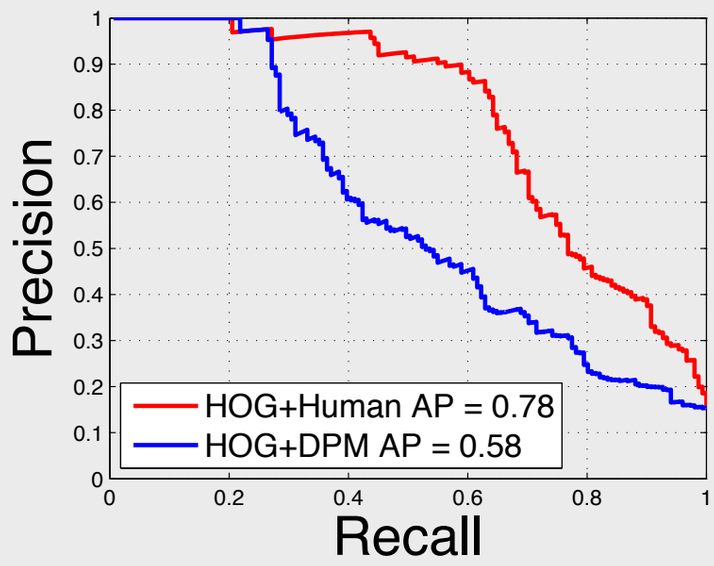
HOG+DPM

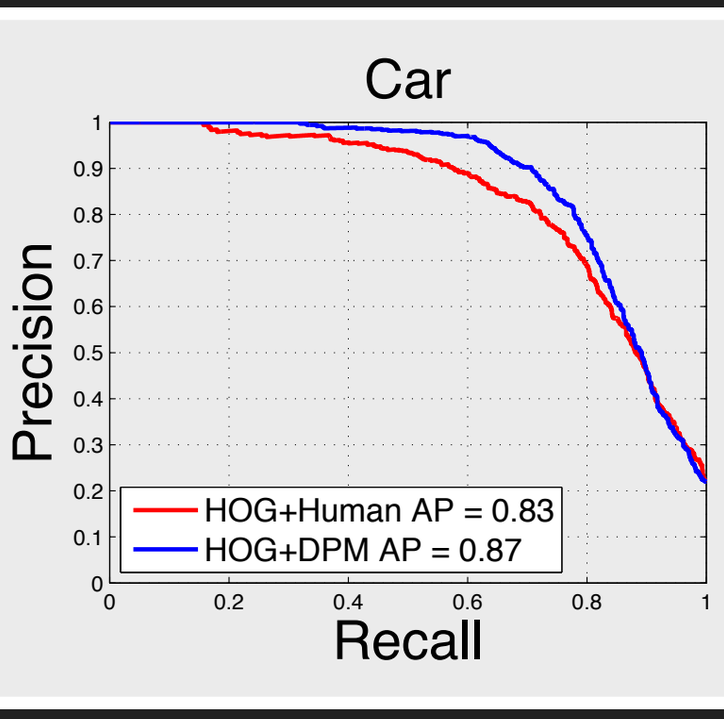
RGB+Human

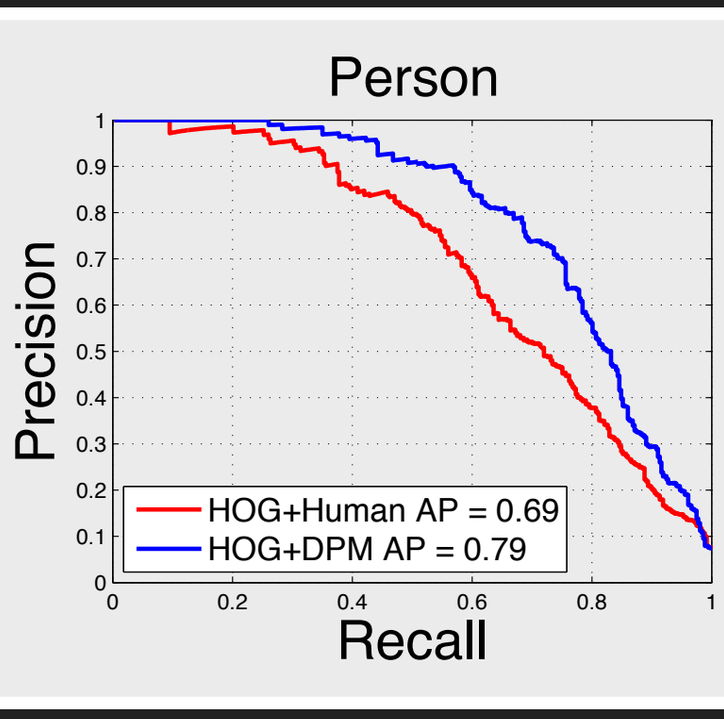
HOG+Human

HOG+Human Detector









Comparing Features

Comparing Features

HOG

Comparing Features

HSC

HOG

Comparing Features

HSC

HOG

Comparing Features

HSC

HOG

What HOG sees,
but HSC does not

Inverting Deep Learning

Inverting Deep Learning

Last convolutional layer of Krizhevsky et al.
Code from DeCAF by Donahue et al.

Experimental

Extracting Color from HOG

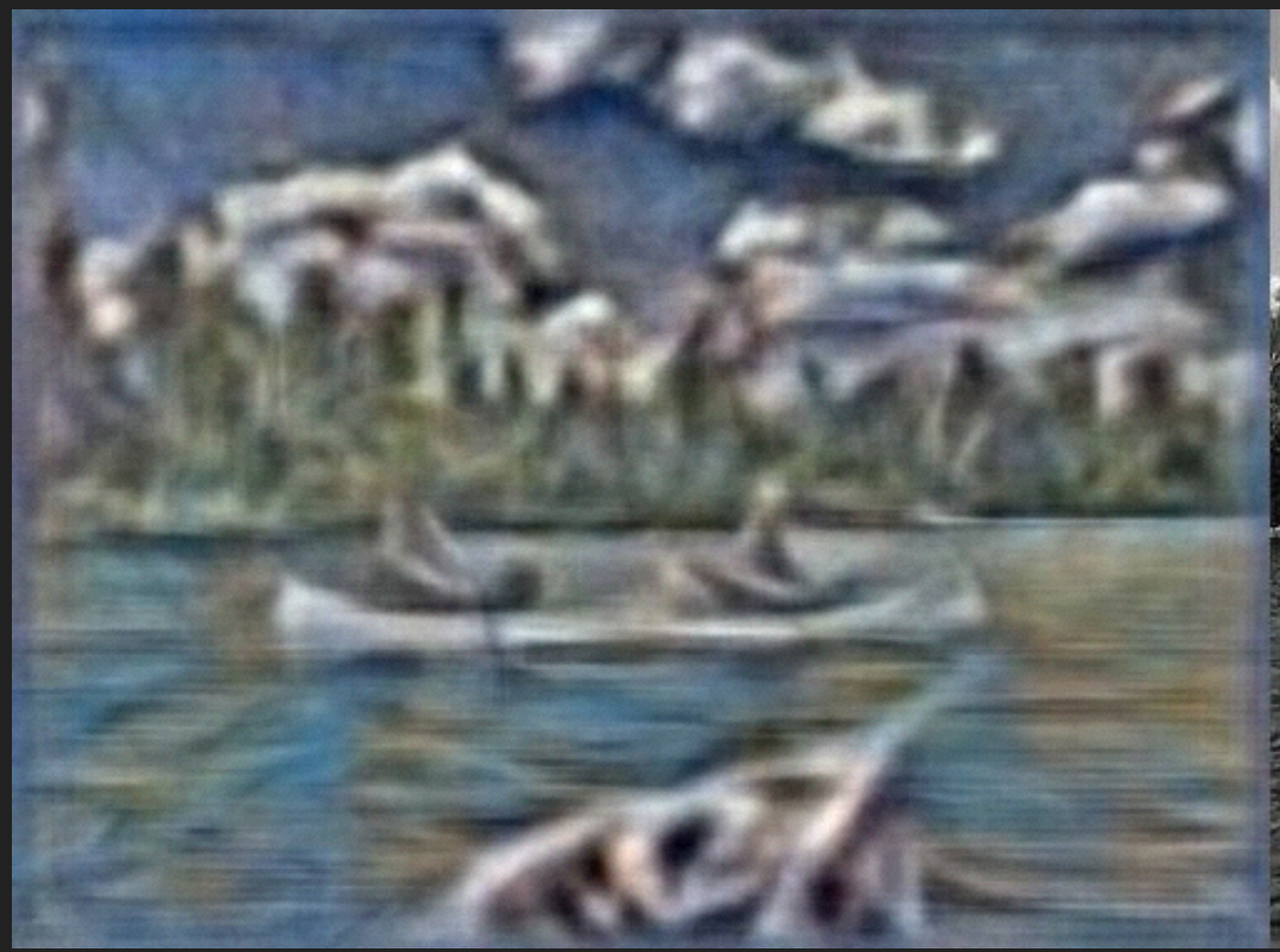
HOG

Image

Extracting Color from HOG

HOG

Image



Visualizing Learned Models

Visualizing Learned Models

Car

Visualizing Learned Models

Car

Person

Visualizing Learned Models

Car

Person

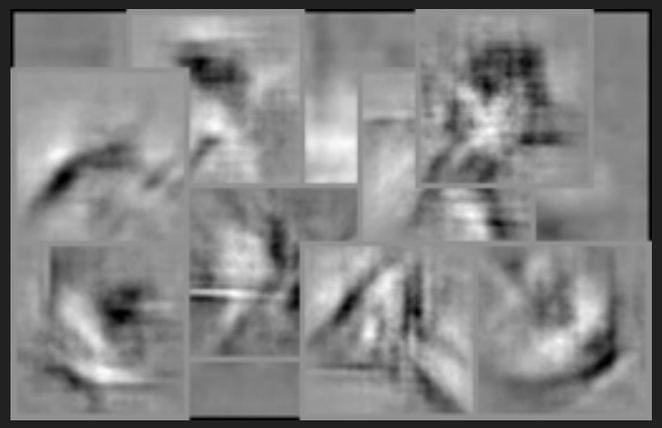
Bottle

Visualizing Learned Models

Car

Person

Bottle



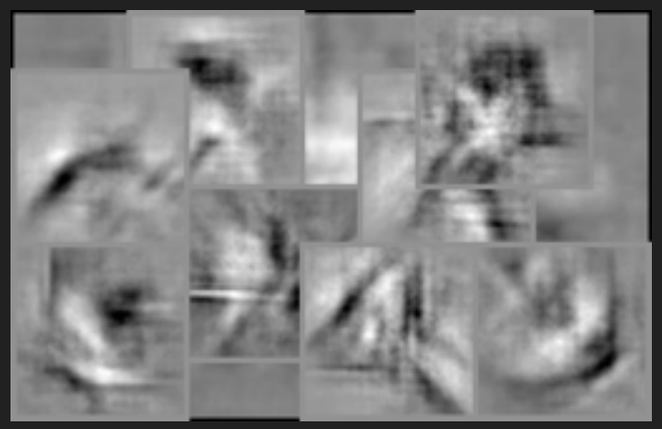
Bicycle

Visualizing Learned Models

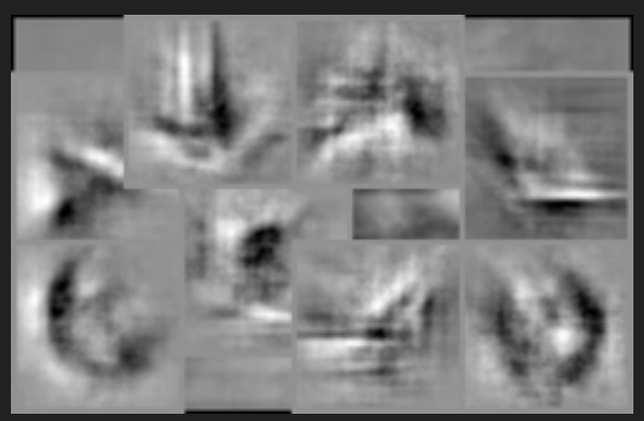
Car

Person

Bottle



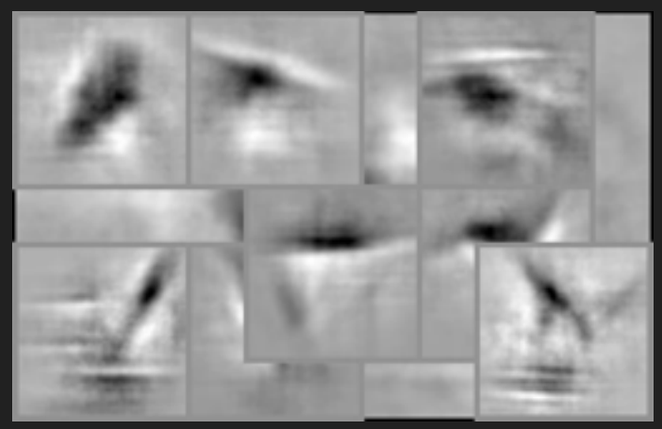
Bicycle



Motorbike

Chair

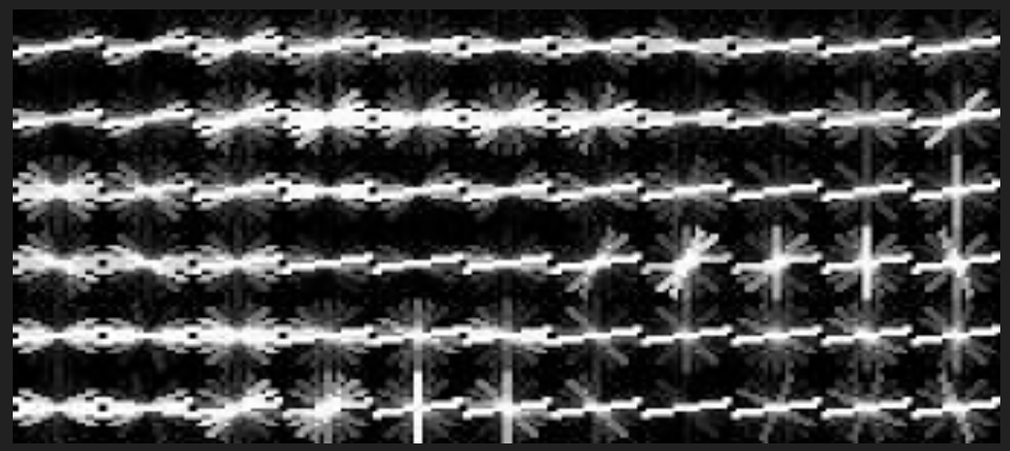
TV



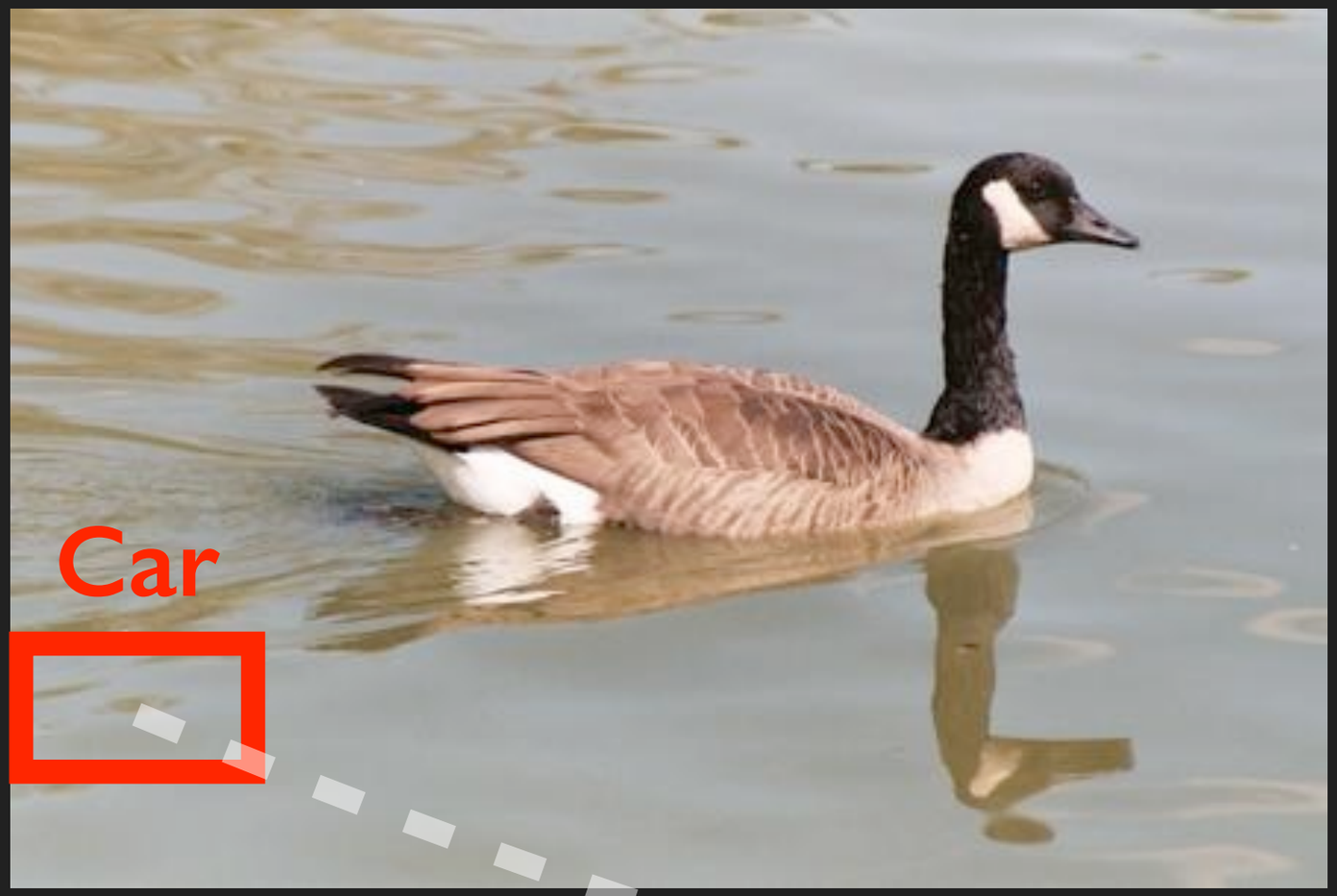
Horse

Why did the detector fail?

Why did the detector fail?



Why did the detector fail?



Code Available

```
ihog = invertHOG(feat);
```

Download at
mit.edu/hoggles

Thanks!

Code + demos + more:

mit.edu/hoggles

ELDA Ridge Direct PairDict

Recovering Pixels

Category	ELDA	Ridge	Direct	PairDict
bicycle	0.452	0.577	0.513	0.561
bottle	0.697	0.683	0.666	0.671
car	0.668	0.677	0.652	0.639
cat	0.749	0.712	0.687	0.705
chair	0.666	0.621	0.604	0.617
table	0.656	0.617	0.582	0.614
horse	0.686	0.633	0.586	0.635
motorbike	0.573	0.617	0.549	0.592
person	0.696	0.667	0.646	0.646
Mean	**0.671**	**0.656**	**0.620**	**0.637**

HOGgles on MTurk

Image Classification

In each image below, indicate whether you think it is a Chair or not. We know the answers to some of the easy images, and we will use this to verify your work. Instructions

Showing image 1 of 10

Is this image a Chair?

○ Image is NOT a Chair (press n) ○ Image is a Chair (press y)

Previous Image Next Image

Submit HIT

Recovering Semantics

Category	ELDA	Ridge	Direct	PairDic	Glyph	Expert
bicycle	0.327	0.127	0.362	0.307	0.405	0.438
bottle	0.269	0.282	0.283	0.446	0.312	0.222
car	0.397	0.457	0.617	0.585	0.359	0.389
cat	0.219	0.178	0.381	0.199	0.139	0.286
chair	0.099	0.239	0.223	0.386	0.119	0.167
table	0.152	0.064	0.162	0.237	0.071	0.125
horse	0.260	0.290	0.354	0.446	0.144	0.15
motorbike	0.221	0.232	0.396	0.224	0.298	0.35
person	0.458	0.546	0.502	0.676	0.301	0.375
Mean	**0.282**	**0.258**	**0.355**	**0.383**	**0.191**	**0.223**