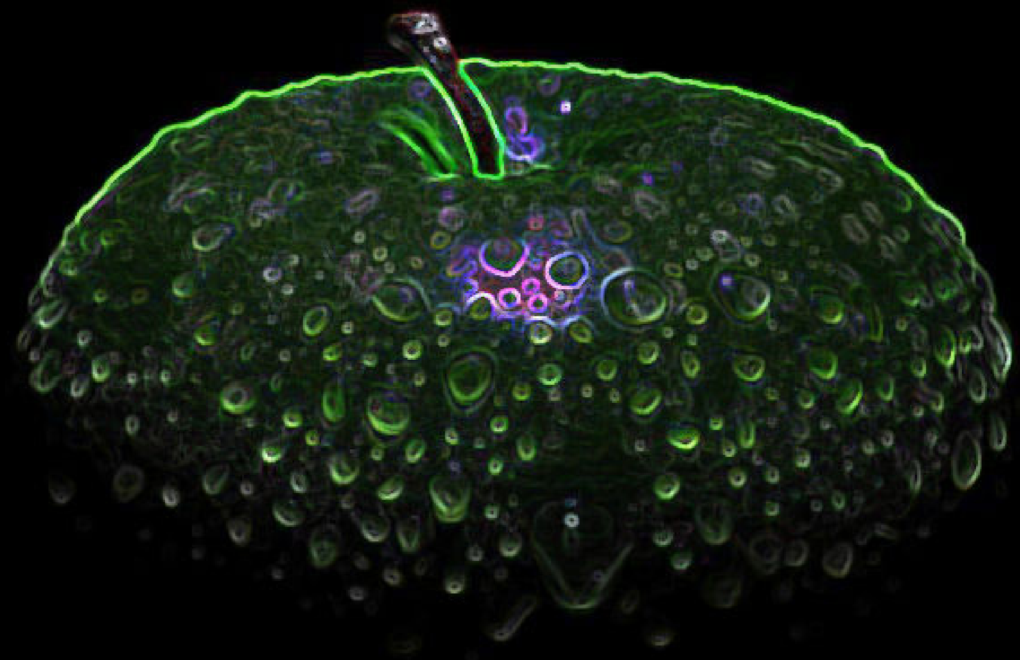
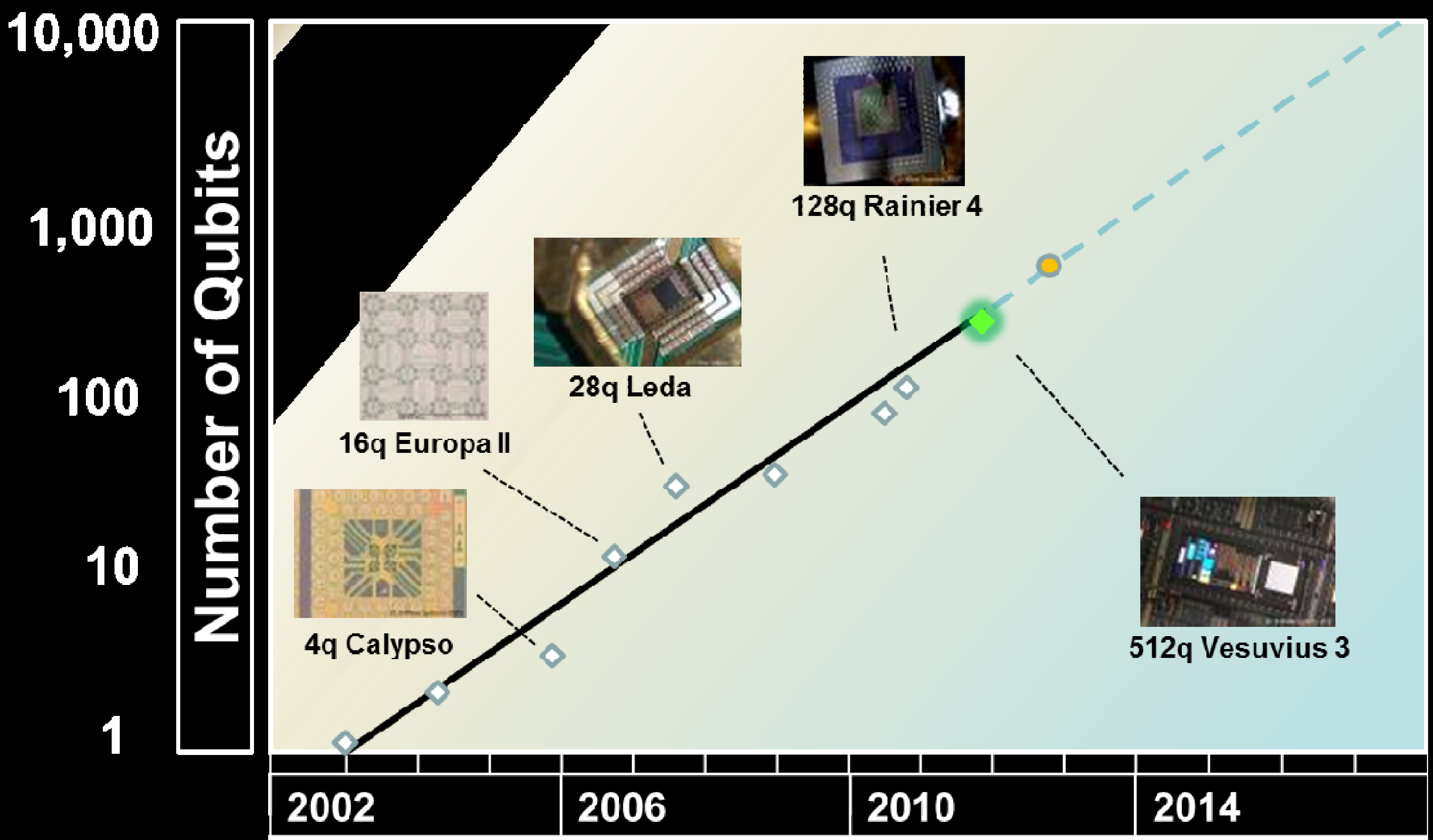
Compressive sensing and semi-supervised feature learning using a D-Wave One

Dr. Geordie Rose

Founder and CTO, D Wave
10:15AM Friday January 20th 2012
@ NASA Ames



The evolution of an idea



The USC – Lockheed Martin Quantum Computing Center

"… the possibility of solving some of the world's most complex optimization and machine learning problems."

USC Viterbi Dean Yannis C. Yortsos

Image and text from
http://dailytrojan.com/2011/10/30/usc-receives-first-quantum-computer/

Quantum computation …will be the first technology that allows useful tasks to be performed in collaboration between parallel universes.

David Deutsch @ TED 2005

Image from
http://conferences.ted.com/TEDGlobal2010/program/

… quantum computers … can solve problems whose solution will never be feasible on a conventional computer.

Quantum computing for everyone

Michael Nielsen (2008)

http://michaelnielsen.org/blog/quantum-computing-for-everyone/

Image from http://www.longexposure.ca/2011/03/tedx-waterloo-the-uncharted/

Someday, perhaps soon, we will build a machine that will be able to perform the functions of a human mind, a thinking machine.

The Connection Machine
Danny Hillis (1985)

Image from http://compass-summit.com/danny-hillis/

… if you were to have a working quantum computer today, the business of doing machine learning would entirely change… quantum computing might be the missing link that brings true human level intelligence to machines.

Hartmut Neven (2007)

http://www.youtube.com/watch?v=I56UugZ_8DI

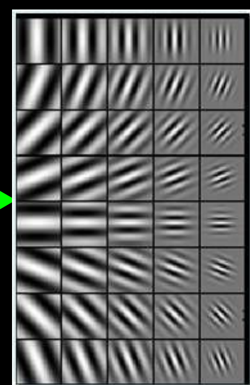
Image from http://www.flickr.com/photos/ecommconf/4540760079/

There's a fascinating hypothesis that a lot of human perception … can be explained by a single learning algorithm.

Unsupervised Feature Learning and Deep Learning
Andrew Ng (2011)
http://www.youtube.com/watch?v=I56UugZ_8DI

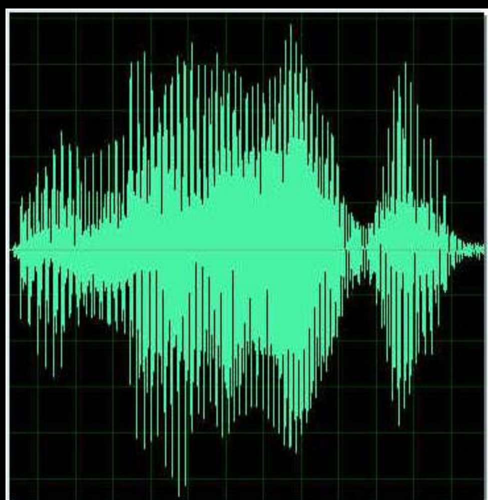
Image from http://www.biotricks.net/2011/08/stanford-machine-learning-classes-in.html

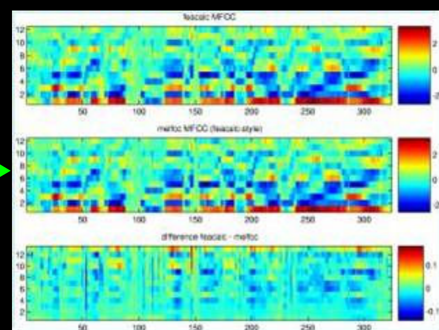
Input → Low-level features → Learning algorithm



SIFT, Spin image, HoG,
RIFT, Textons, GLOH,
Gabor Wavelets

Input → Low-level features → Learning algorithm





Spectrogram, MFCC, Flux, ZCR, Rolloff

http://news.bbc.co.uk/2/hi/technology/7487645.stm

Input → Low-level features → Learning algorithm

Finance
Business
Sports
Music ✓
Realty
Eldritch horrors ✓

Bag of words, Parser features, NER/SRL, Stemming, Anaphora, POS tagging, WordNet features

Learning features: images

Warm-up: how may bits does it take to download this highly compelling movie from Netflix?

Option 1.

Send all the bits for all eight images –
80x112x3x8 x 8 = 1,720,320 bits

Option 2.

Send one picture, plus instructions that there are eight –
80x112x3x8 + 8 = 215,048 bits

+ [1, 1, 1, 1, 1, 1, 1, 1]

Option 2.

Send one picture, plus instructions that there are eight –
80x112x3x8 + 8 = 215,048 bits

+ [1, 1, 1, 1, 1, 1, 1, 1]

Feature or dictionary atom

Question:

Is the equality below:

- ☐ Obvious
- ☐ Deep

=  + [1, 1, 1, 1, 1, 1, 1, 1]

Question:

Is the equality below:

☒ **Obvious**
☒ **Deep**

=  + [1, 1, 1, 1, 1, 1, 1, 1]

What if our 'video' is more interesting?

- How many features do we need to represent images from the world around us?

- How do we find them?

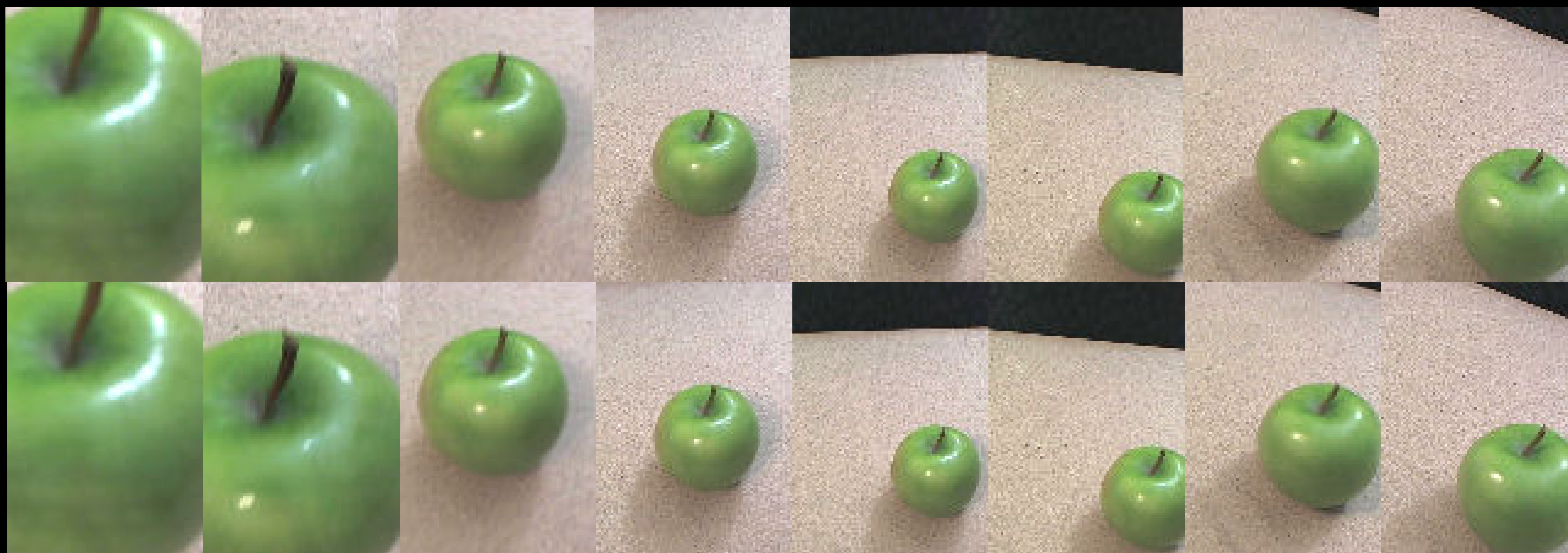
One feature

Like an "average"

Feature Dictionary

One feature per image

Guarantee of perfect reconstruction



Feature Dictionary →

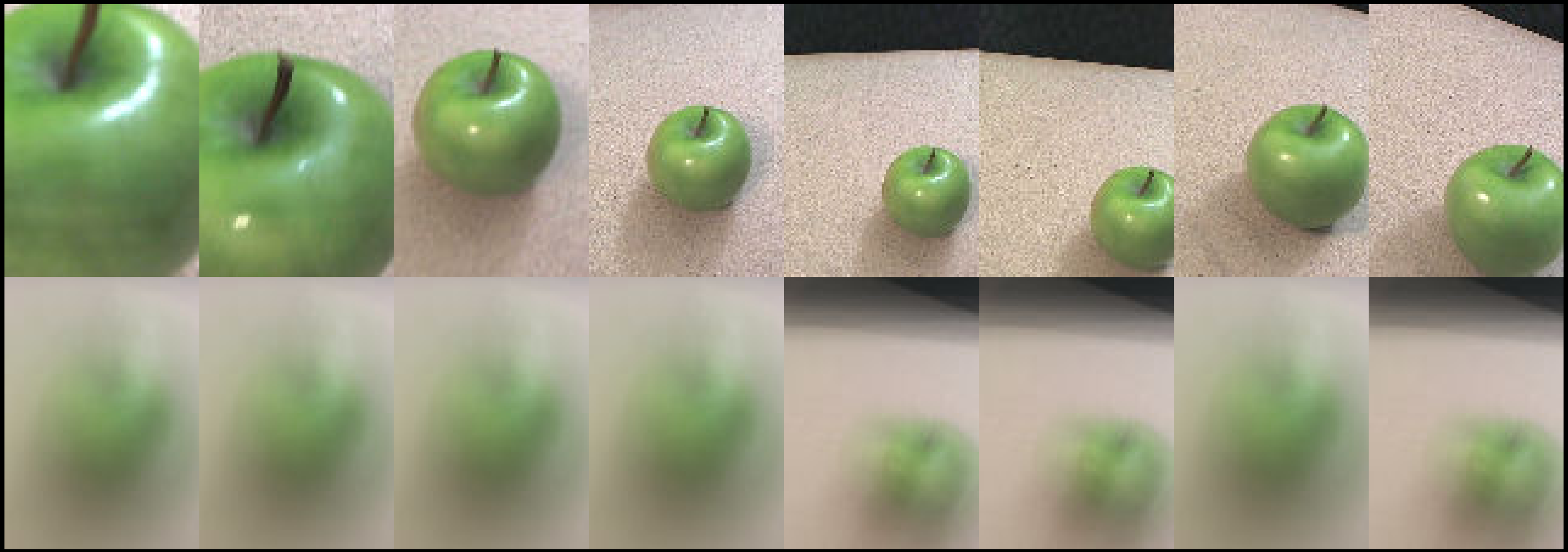
[MANY NATURAL SIGNALS ARE SPARSE OR COMPRESSIBLE IN THE SENSE THAT THEY HAVE CONCISE REPRESENTATIONS WHEN EXPRESSED IN THE PROPER BASIS.]

An Introduction to compressed sampling

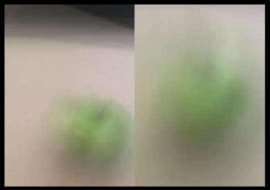
IEEE Signal Processing Magazine 21 March 2008

Two features

A little better!

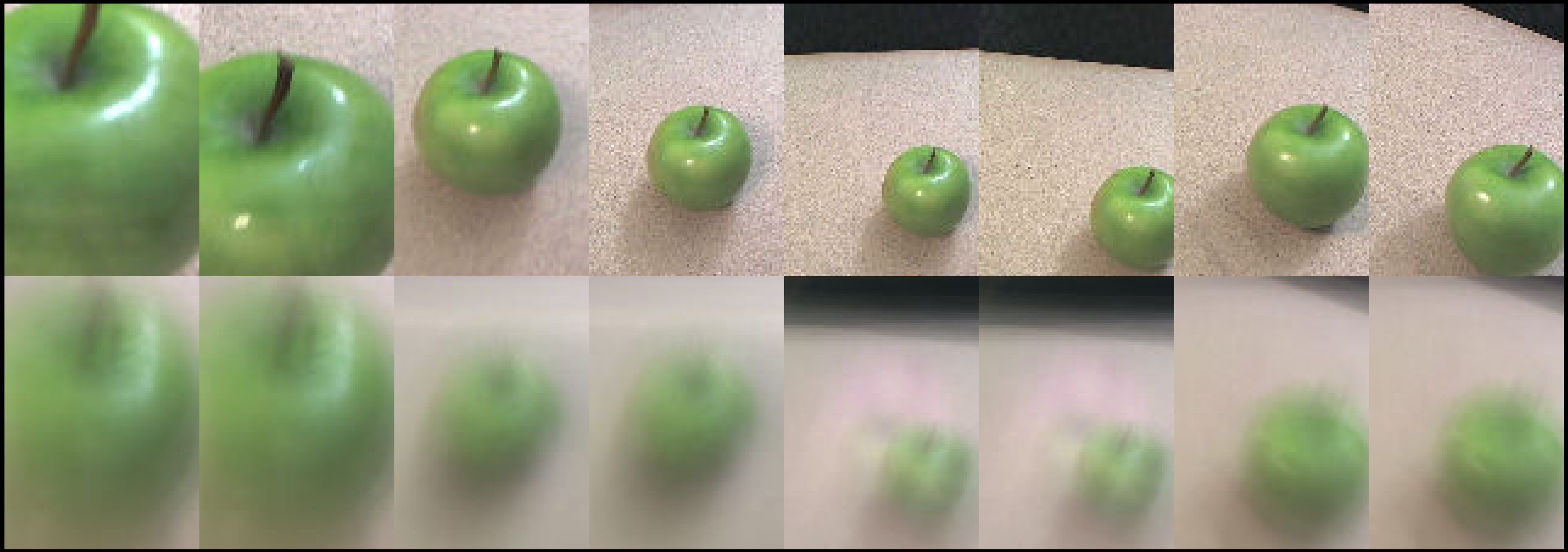


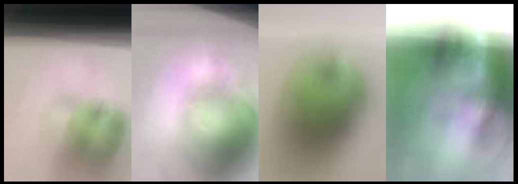
Feature Dictionary →



Four features

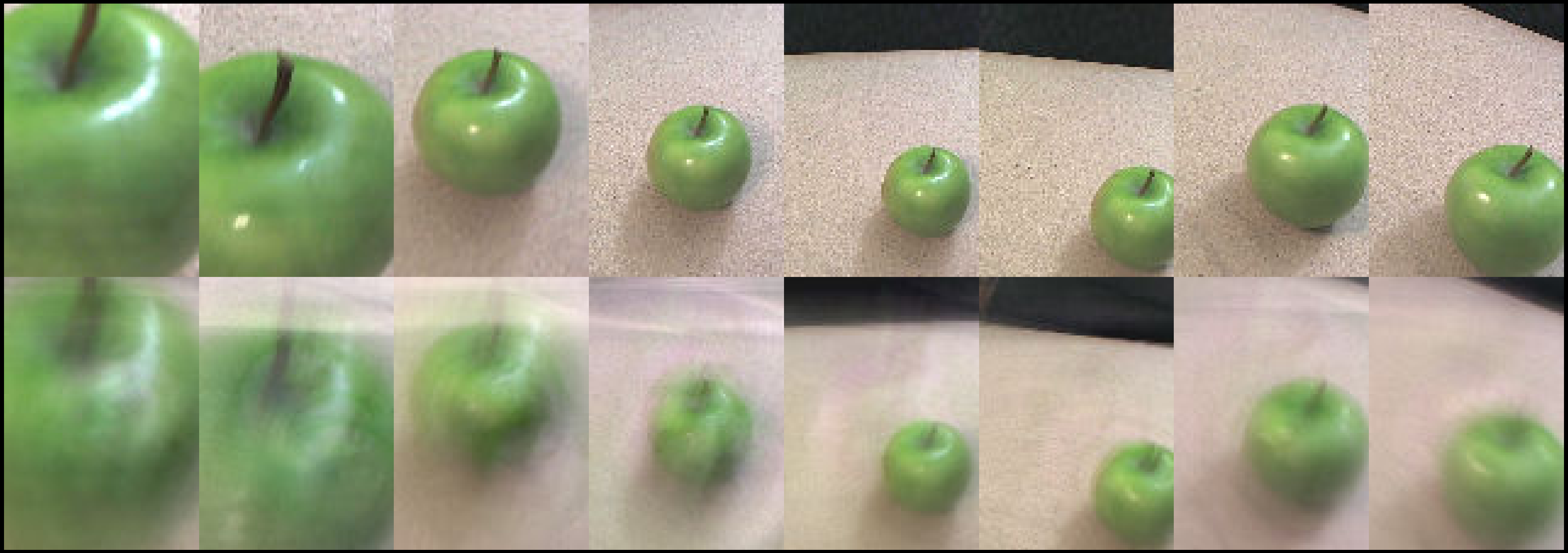
Better still…



Feature Dictionary → 

Twenty features

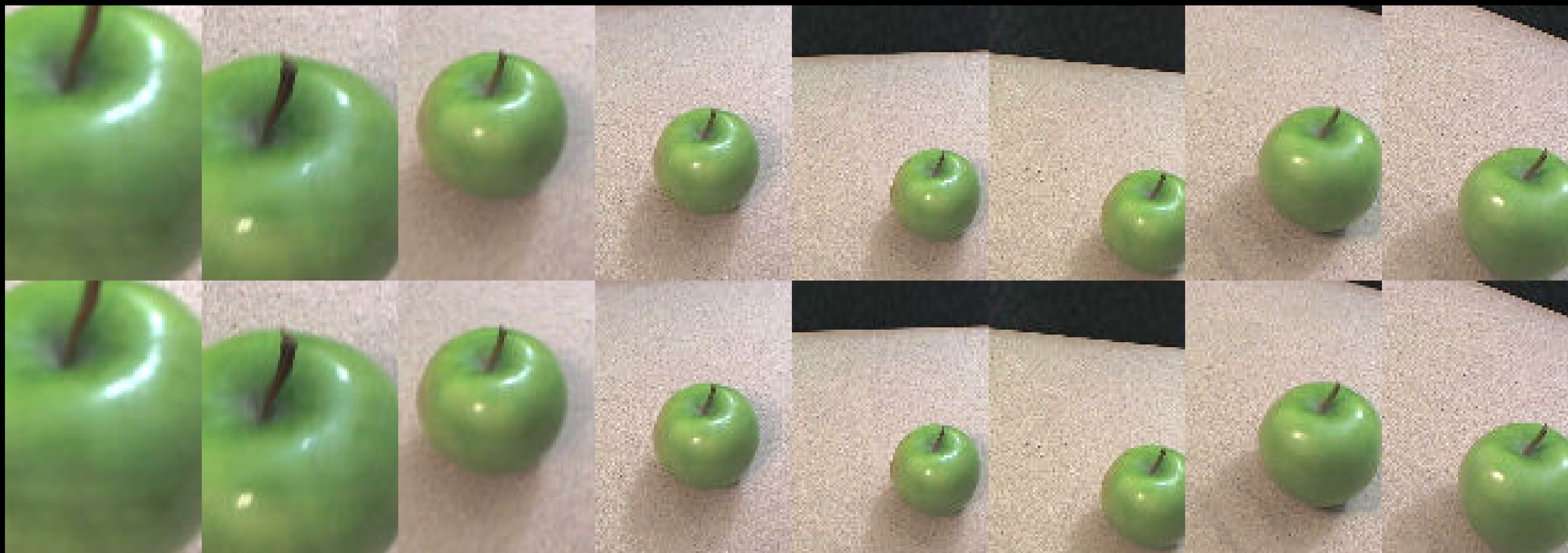
Better still…



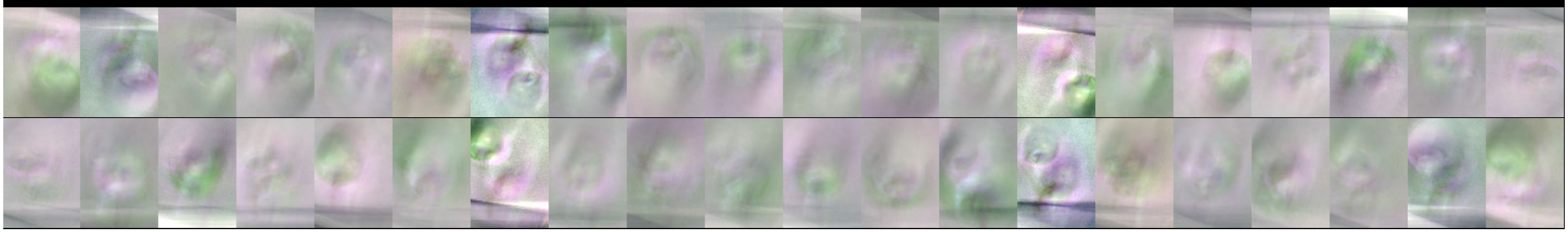
Feature Dictionary →

Forty features

Near perfect reconstruction of a real 256 image movie



Feature Dictionary



Not just apples

Another 20-element dictionary for a 256-image movie

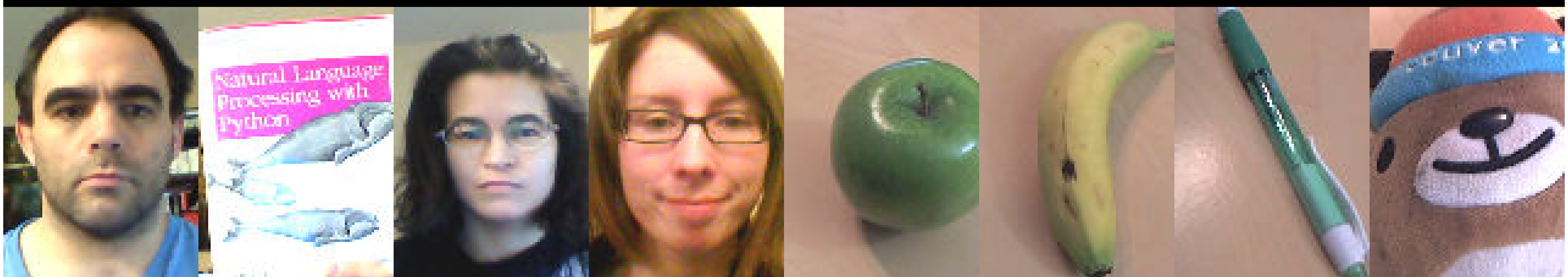
Feature Dictionary →

Not just apples

Another 20-element dictionary for a 256-image movie

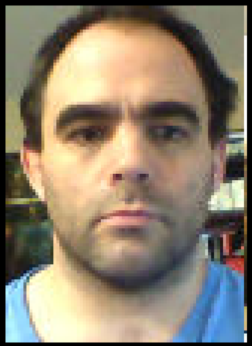
Feature Dictionary →

Framework easily handles combination of labeled and unlabeled data



{Geordie, NLTK, Mary, Suz, Apple, Banana, Pen, MukMuk}

Framework easily handles combination of labeled and unlabeled data



Just append label data
[+1, -1, -1, -1, -1, -1, -1, -1]
to image data vector!

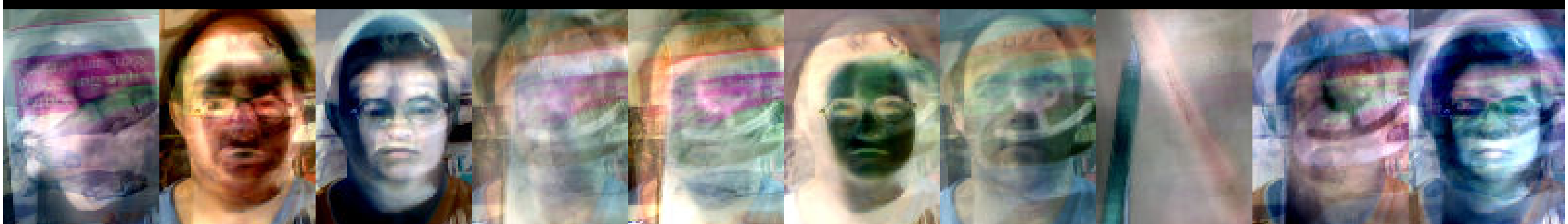
{Geordie, NLTK, Mary, Suz, Apple, Banana, Pen, MukMuk}

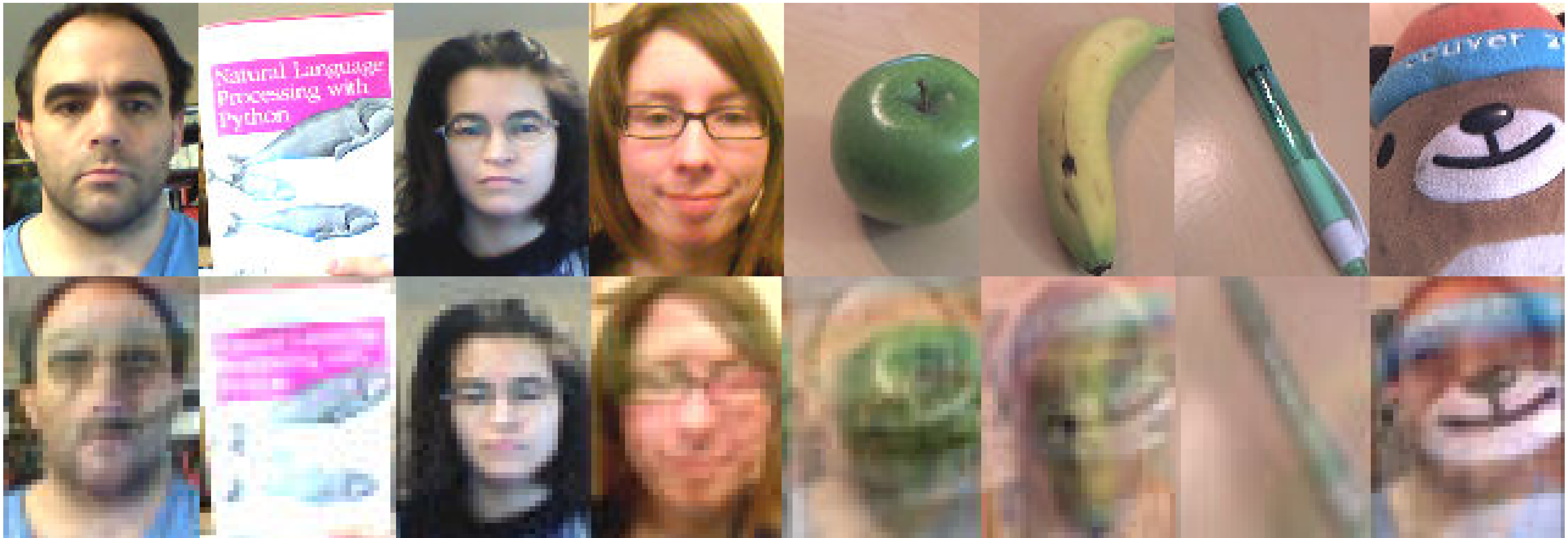
Eight categories, 128 images from each

64 labeled, 64 unlabeled

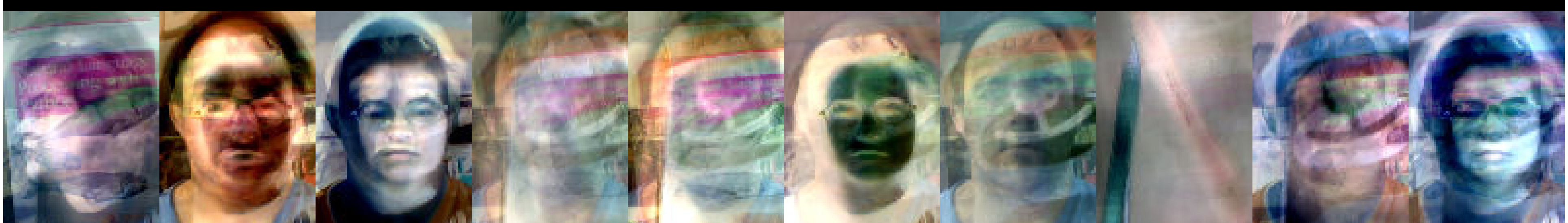
Learn 10 features for a 1,024-image movie

Feature Dictionary →





Feature Dictionary



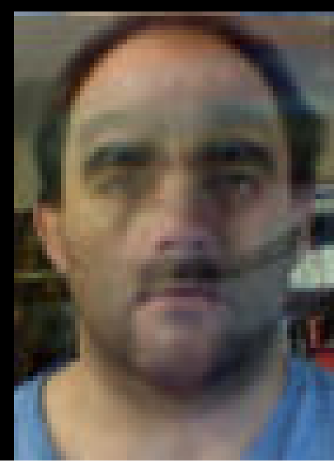
(Extremely hard) optimization problem!

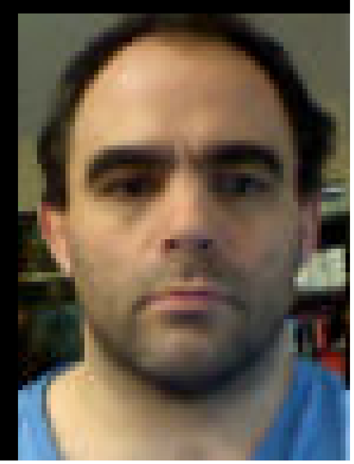
Find $\vec{D}_m$ and $\vec{w}_j$ that minimize the difference between ground truth and reconstructions

$\vec{D}_1$ $\vec{D}_2$ $\vec{D}_3$ $\vec{D}_4$ $\vec{D}_5$ $\vec{D}_6$ $\vec{D}_7$ $\vec{D}_8$ $\vec{D}_9$ $\vec{D}_{10}$

$$\vec{I}_j = \sum_{m=1}^{K} \vec{D}_m \vec{w}_j$$

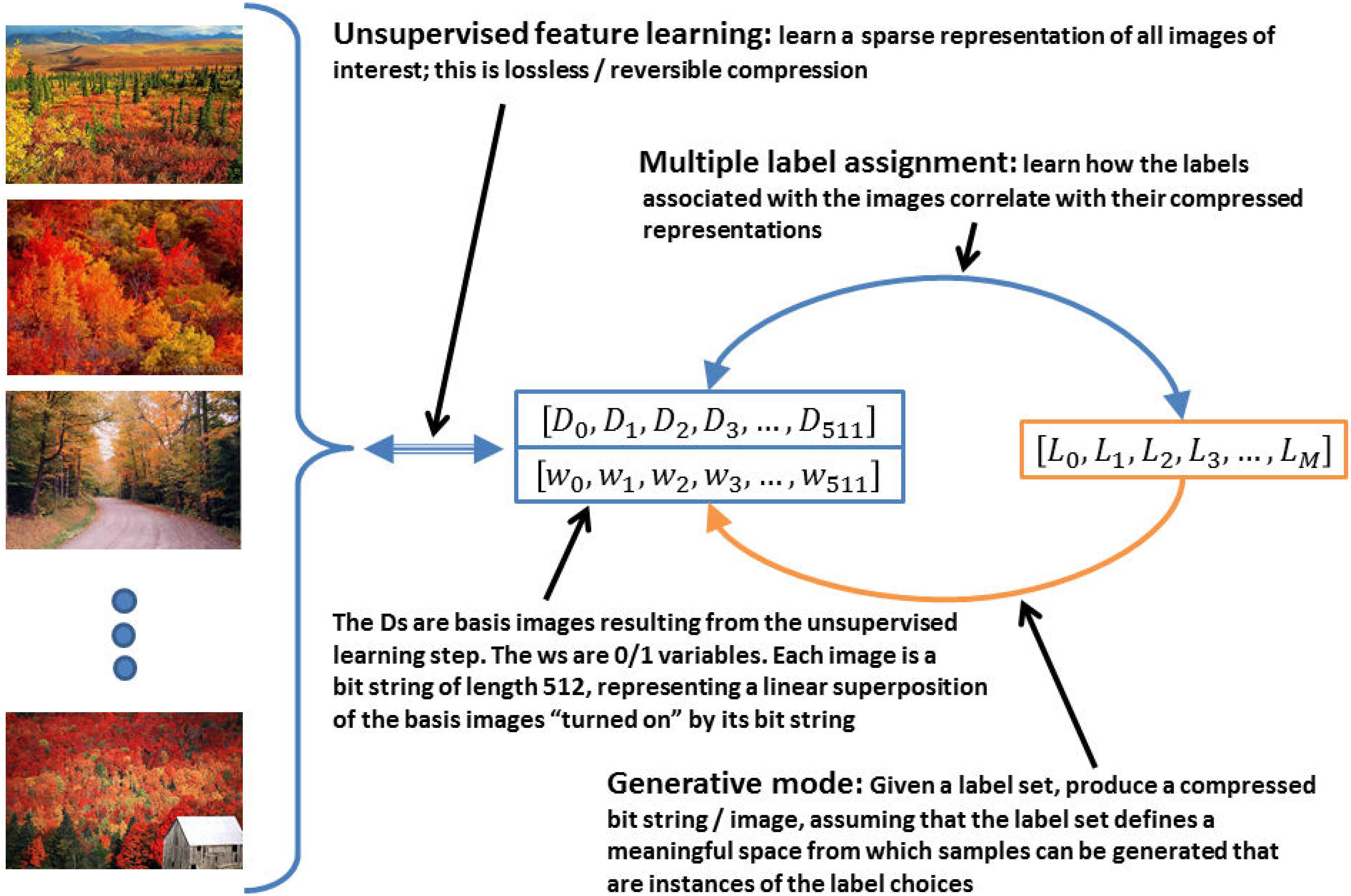
$$\vec{w}_j = [0,1,0,1,0,0,1,0,0,0\ ] \longrightarrow$$





Once you've learned your features…

1. Assign multiple labels to new objects
2. Anomaly detection
3. Generative mode – assign an object to a new label set
4. Use features as inputs to learning algorithms
5. Objects can have multiple data types seamlessly included at the same time – e.g. image + speech + text + category labels

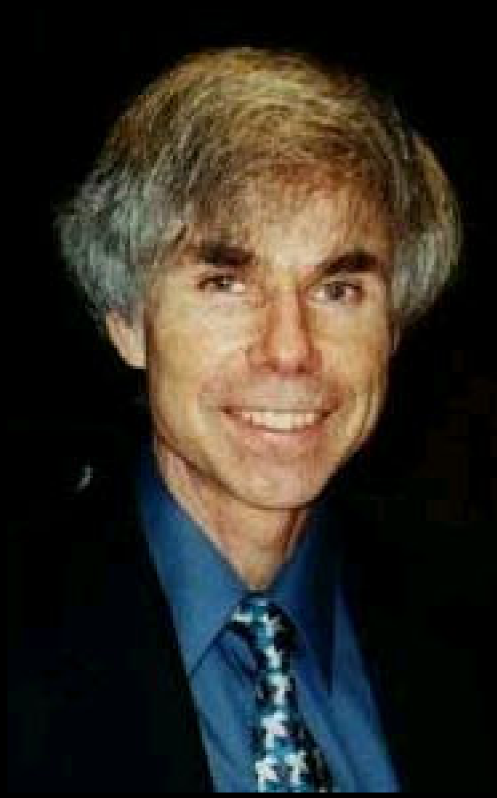


Unsupervised feature learning: learn a sparse representation of all images of interest; this is lossless / reversible compression
Multiple label assignment: learn how the labels associated with the images correlate with their compressed representations
$[D_0, D_1, D_2, D_3, \dots, D_{511}]$
$[w_0, w_1, w_2, w_3, \dots, w_{511}]$
$[L_0, L_1, L_2, L_3, \dots, L_M]$
The Ds are basis images resulting from the unsupervised learning step. The ws are 0/1 variables. Each image is a bit string of length 512, representing a linear superposition of the basis images "turned on" by its bit string
Generative mode: Given a label set, produce a compressed bit string / image, assuming that the label set defines a meaningful space from which samples can be generated that are instances of the label choices

Do androids dream of electric sheep?

Generative mode – assign an object to a new label set

Think of this as "the inverse of classification"



http://www.cogs.indiana.edu/people/homepages/hofstadter.html

Thanks!

rose@dwavesys.com