Classifying an image with accuracy and speed

The value of parts-based representations

Naveen Agnihotri, PhD

Milabra

New York, NY

Milabra

The Problem

- The amount of image and video data on the Internet is exploding:
  - Imageshack has 20 billion images in its inventory
  - Facebook has 15 billion, growing at 850 million per month
- The most prevalent method to index images is social tagging, which is unreliable and not comprehensive.
- Images are expensive to maintain, and expensive to serve
- They represent a potential liability to content owners.
- Without indexing, there is no easy way to monetize the images by placing relevant ads

Milabra

How images are stored and processed

Computer

- Serial storage and processing
- Data stored by pixels and frames
- Retrieval is based on index lookup
- Easy to answer storage queries (“Where is this file located?”)
- Difficult to answer pattern queries (“Is this a motorcycle or a boat?”)

Brain

- Parallel storage and processing
- Data stored by features and patterns
- Retrieval is based on pattern reconstruction
- Difficult to answer storage queries (“where is this memory located?”)
- Easy to answer pattern queries: (“Is this is motorcycle or a boat?”)

Milabra

Learning to find patterns in images

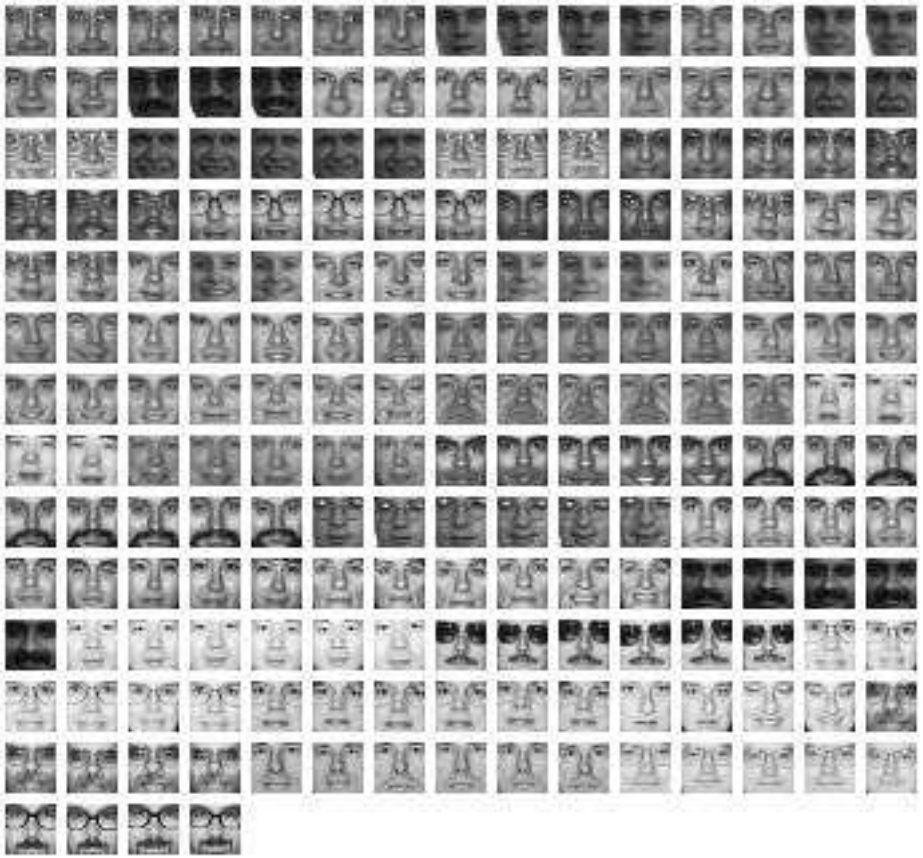
- Feature extraction using “basis images”
  - The input image is represented by a linear combination of basis images
  - Each basis image is the same size
  - Each basis image provides a “feature” (or a set of features)

- Operating procedure:
  - Start with a set of training images
  - The basis images are of the same size as the input images
  - Use a learning heuristic (analogous to PCA) to figure out the optimal basis images for the input image set
  - Assign constraints based on observations from neuroscience: firing rates of neurons are never negative

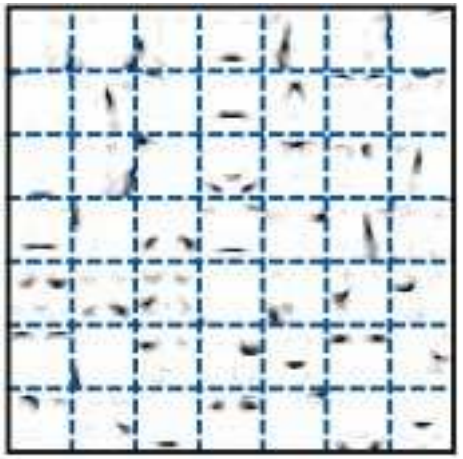
Milabra

Basis image calculation example

Input face images (3431 total)

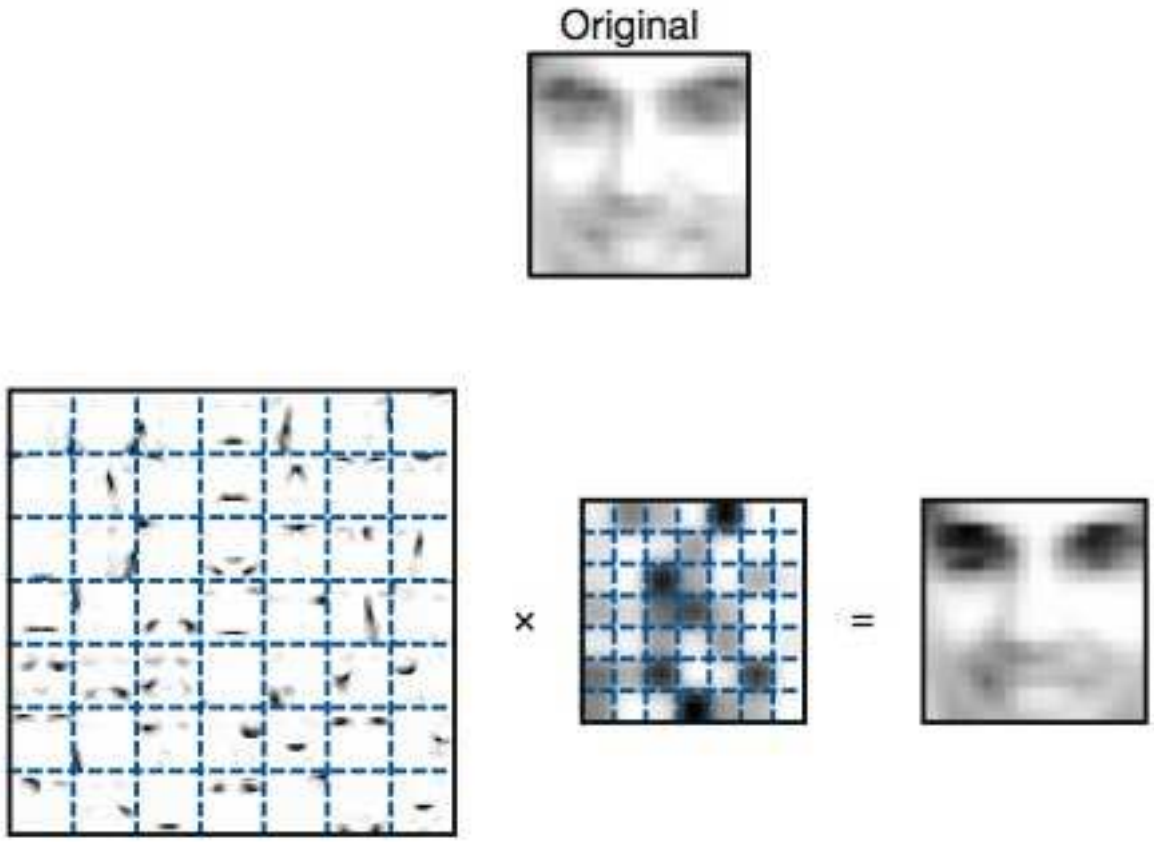


Basis images (49)



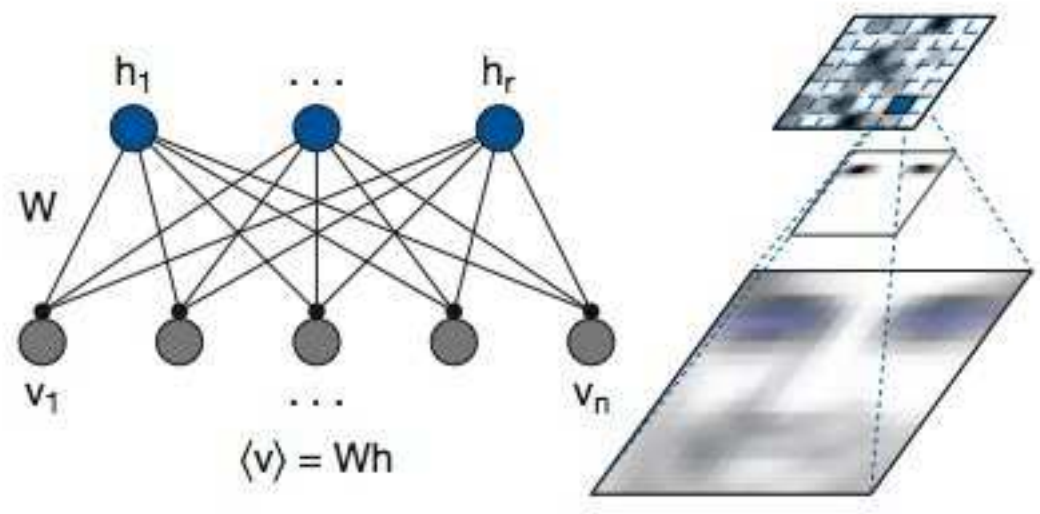
Milabra

How the process works (reconstruction)



Milabra

Hidden variables model (neural network)



- $h_1 \ldots h_r$ = hidden layer neuron strengths
- W = weight matrix (basis images)
- $v_1 \ldots v_n$ = image pixel values

Advantages of parts-based feature extraction

- Searches use feature vectors based on patterns in actual image data, instead of associated text.
- Helps with database and file-based storage: instead of storing the complete image, we can store the feature vector for the image.
- Pattern-based searches are much more accurate than using pixel-based searches.
- Heavy computation is restricted to the point of database creation, and can be accelerated with parallelization.
- Dramatically reduces the dimensionality of search (25-50 features vs. 100K pixels), so searches are faster and less expensive.

Milabra


Milabra