FACE VERIFICATION:
USING LOCAL FREQUENCY BANDS

Chris McCool (Idiap)

An Overview

- Face Recognition and Verification: a brief overview....
- Face Verification using GMM and a Parts-Based approach
- Extending the GMM Parts-Based approach: by applying spatial and frequency decomposition
- Where to from here?

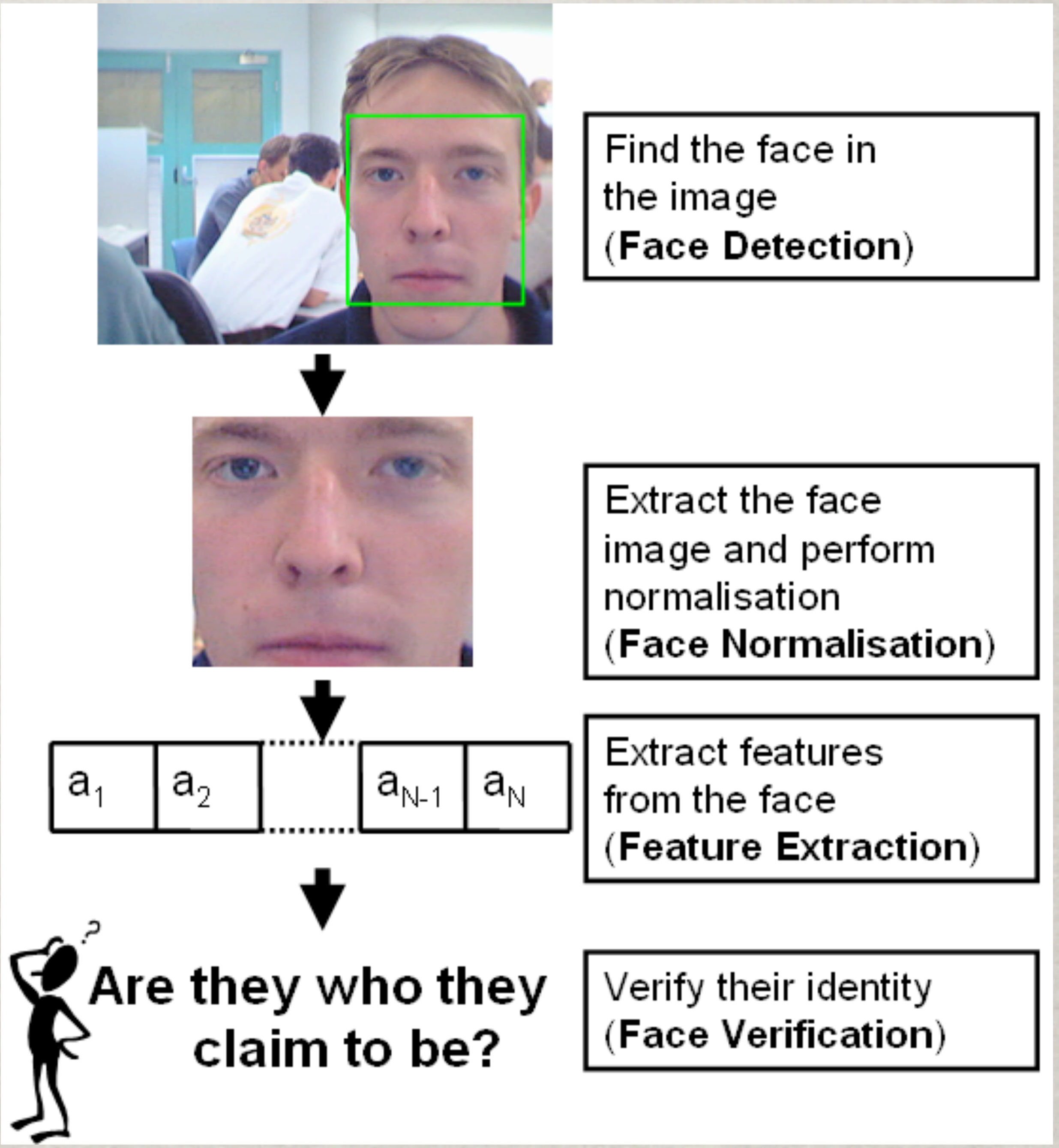
Face Recognition
what is it?

Face Recognition

- When I think of face recognition I think of security guards looking at my driver's licence
- Or immigration officials looking at my passport photo
- So we know that humans can do this but how can we get a computer to automatically recognise someone's face?

Face Recognition

- Two basic steps:
  - **Face Detection**: find the face (or faces) in an image
  - **Face Verification**: match the features to the model of the ID they are claiming to be



Face Verification

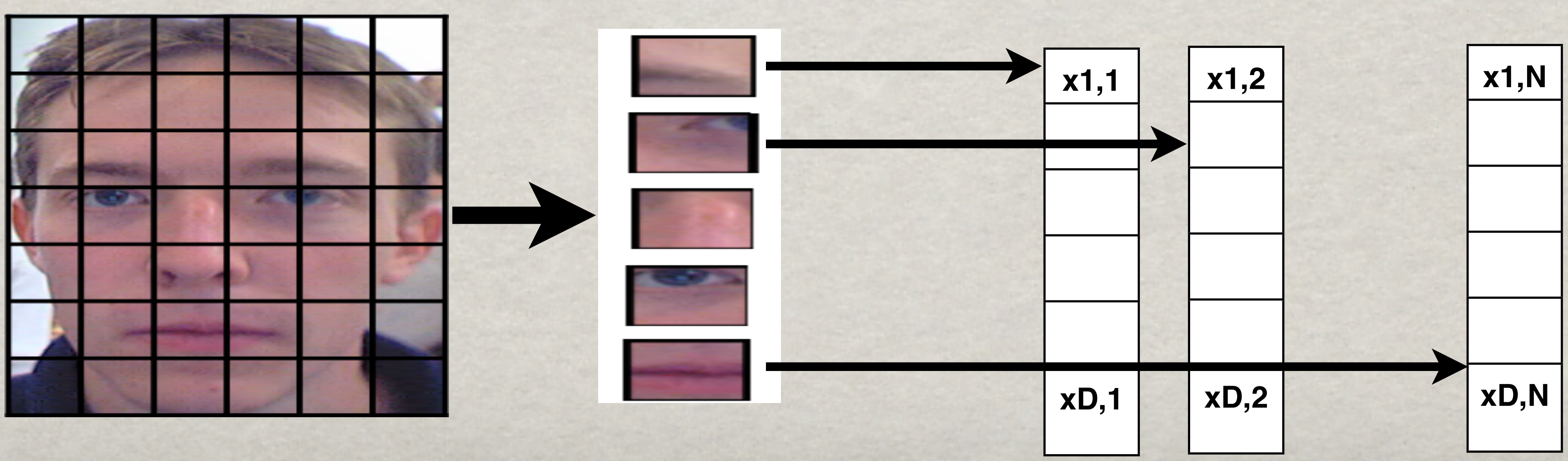
- Face Verification:
  - An input image is supplied along with who they claim to be (a claimed ID)
  - We then match the input to the template or model of this claimed ID
    - This is a 1-1 match
  - We then compare the match (a score) against a threshold to accept or reject

Face Verification

- Many methods have been proposed to perform face verification:
    - For obtaining **features** from a face people have proposed techniques such as:
        - Principal Component Analysis, Linear Discriminant Analysis, Independent Component Analysis, Discrete Cosine Transform (DCT), Gabor Wavelets
    - For **classifying** these features people have proposed many techniques including:
        - Distance or Similarity Measures, Support Vector Machines, Neural Networks, Gaussian Mixture Models, Hidden Markov Models
- Out of all the possibilities there is an interesting paradigm (that is also quite successful) called the GMM Parts-Based approach

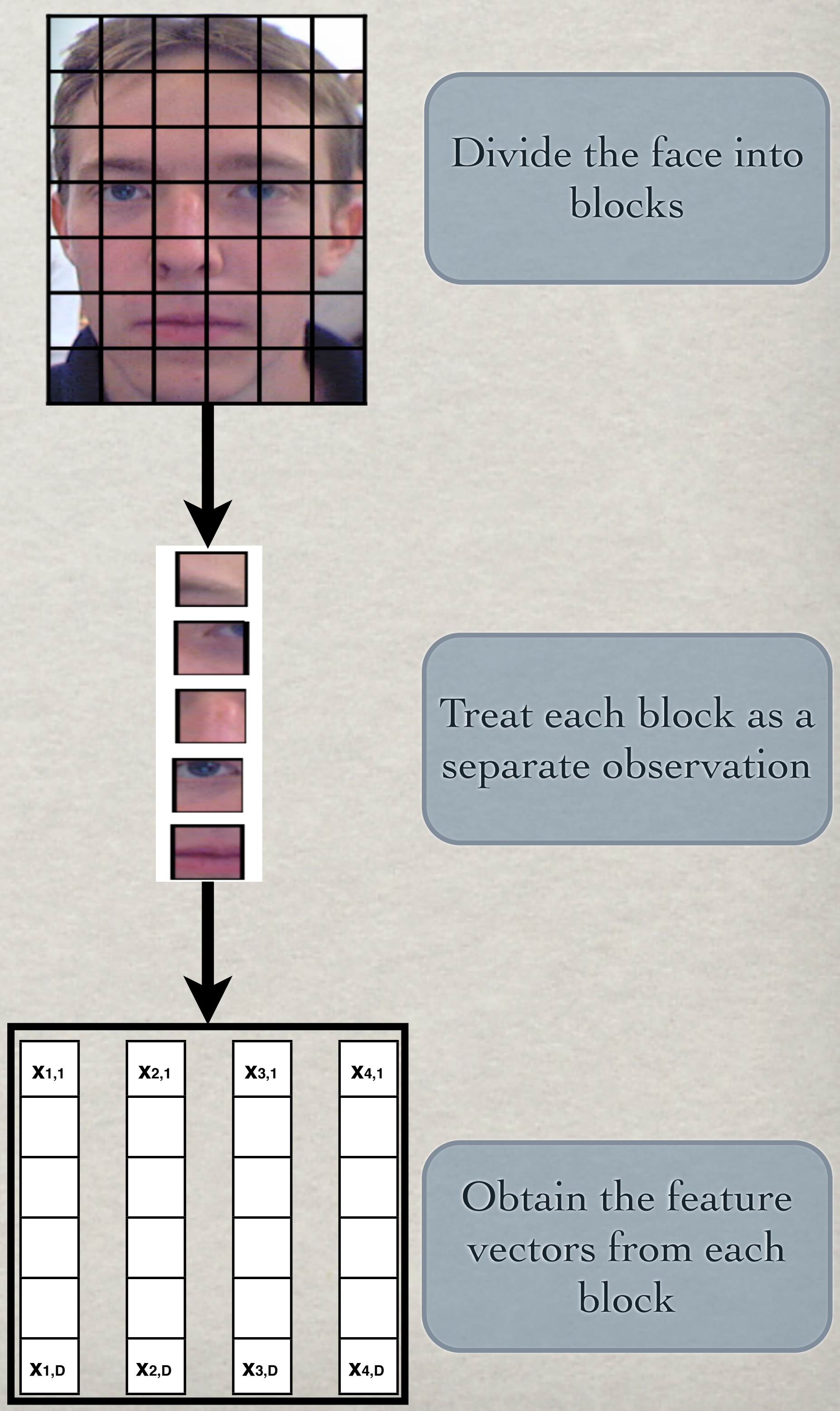
GMM Parts-Based Approach

- Interesting: multiple feature vectors are obtained from a single face image
- The face is divided into blocks: DCT feature vectors are obtained from each block and treated independently
  - Getting local frequency information
- The feature vectors are then modelled with a Gaussian Mixture Model
  - Trying to describe the probability density function (pdf) of these features



Face Verification

- **Important Aspects**
  - Each block is treated as an independent observations of the same signal/object
  - This gives us many observations from a single image
  - This method is performing a spatial decomposition of the face

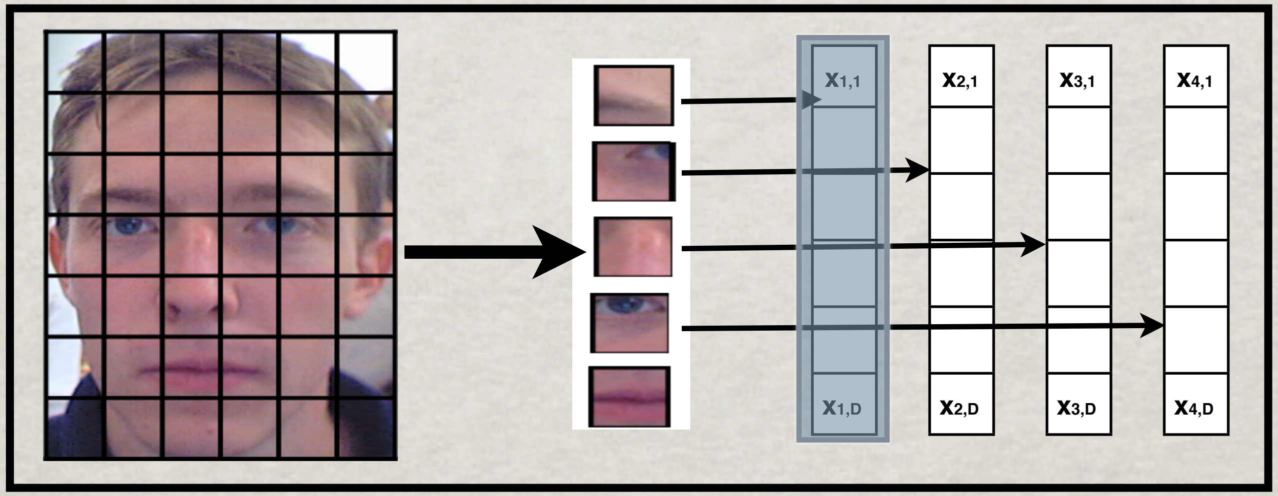


Extending the Parts-Based Approach

Spatial and Frequency Decomposition

Local Frequency band Approach

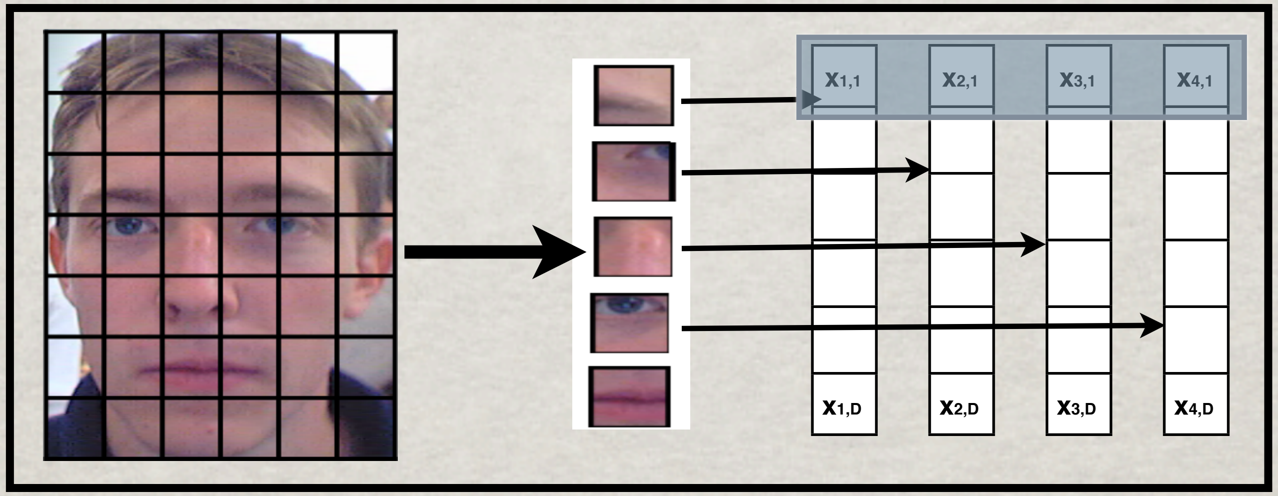
- The original story was that we obtained a feature vector from each block



- Each feature vector is a frequency response
  - Obtained using the Discrete Cosine Transform (DCT)

Local Frequency band Approach

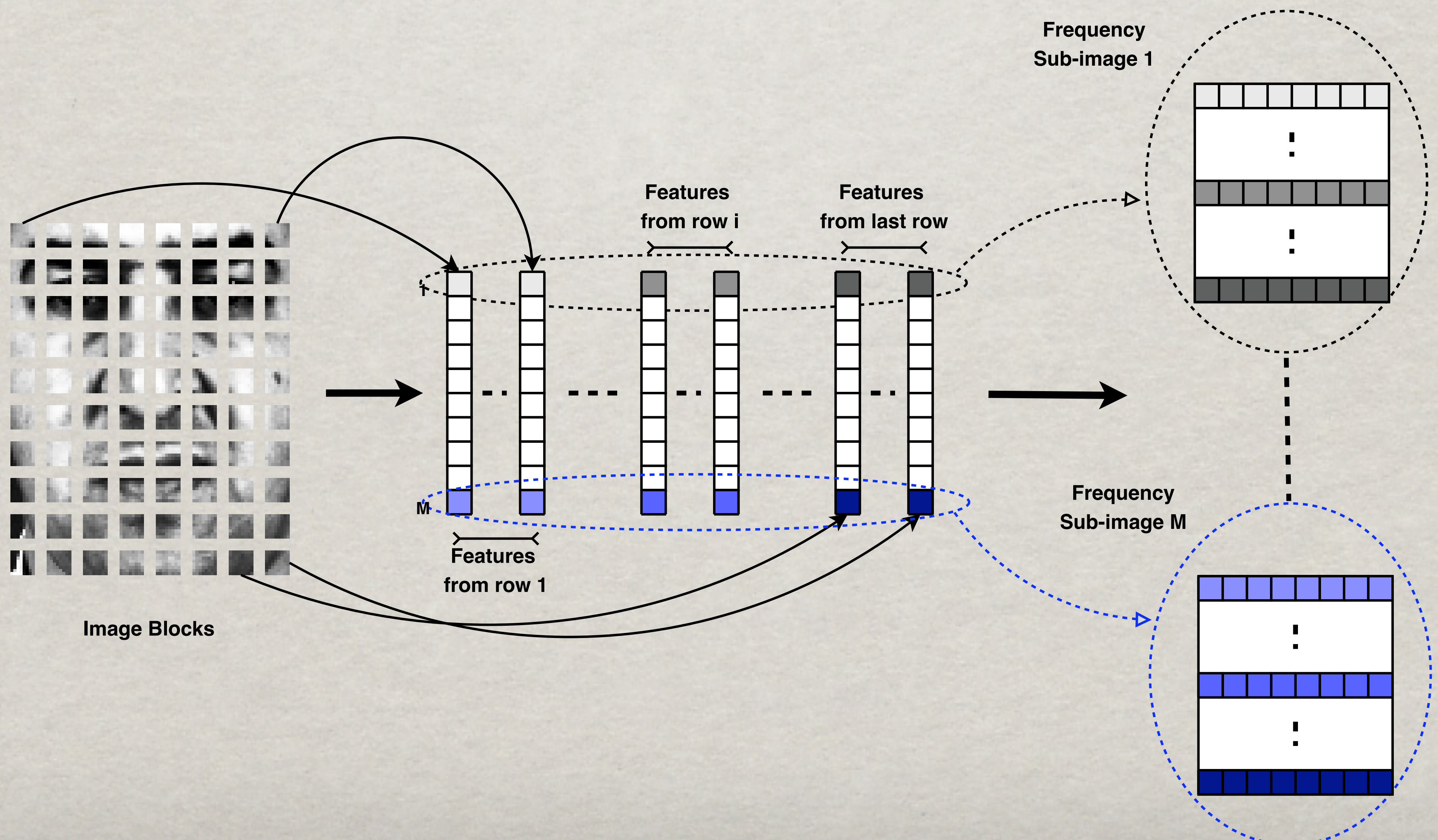
- What happens if we treat the frequency response separately?



- They could be used to rebuild a set of images which now represent the local frequency response

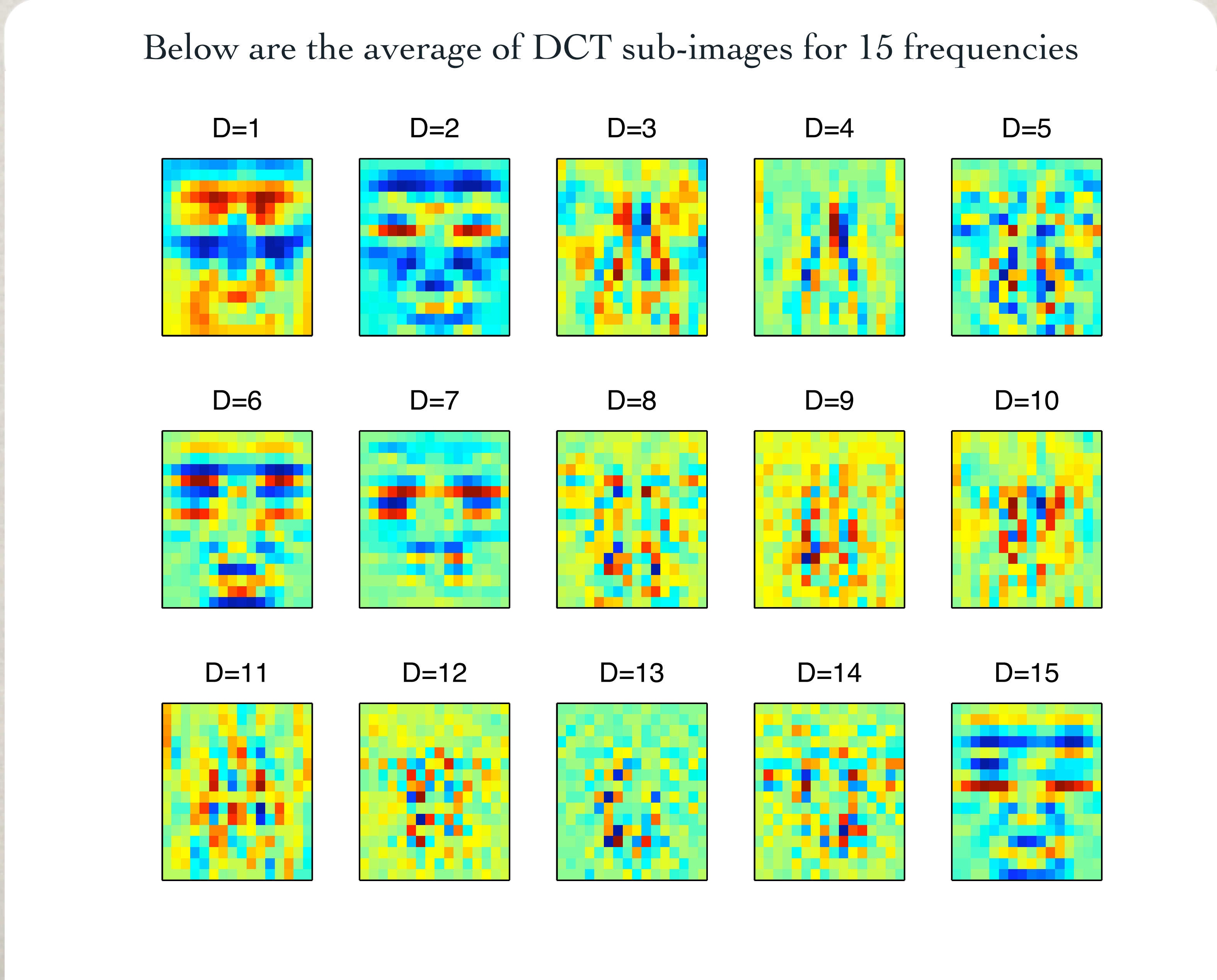
Local Frequency band Approach

- Building a set of frequency images
    - From a set of local frequency responses



Local Frequency band Approach

- To explain this it will help to show graphically what I mean



Local Frequency Subband Approach

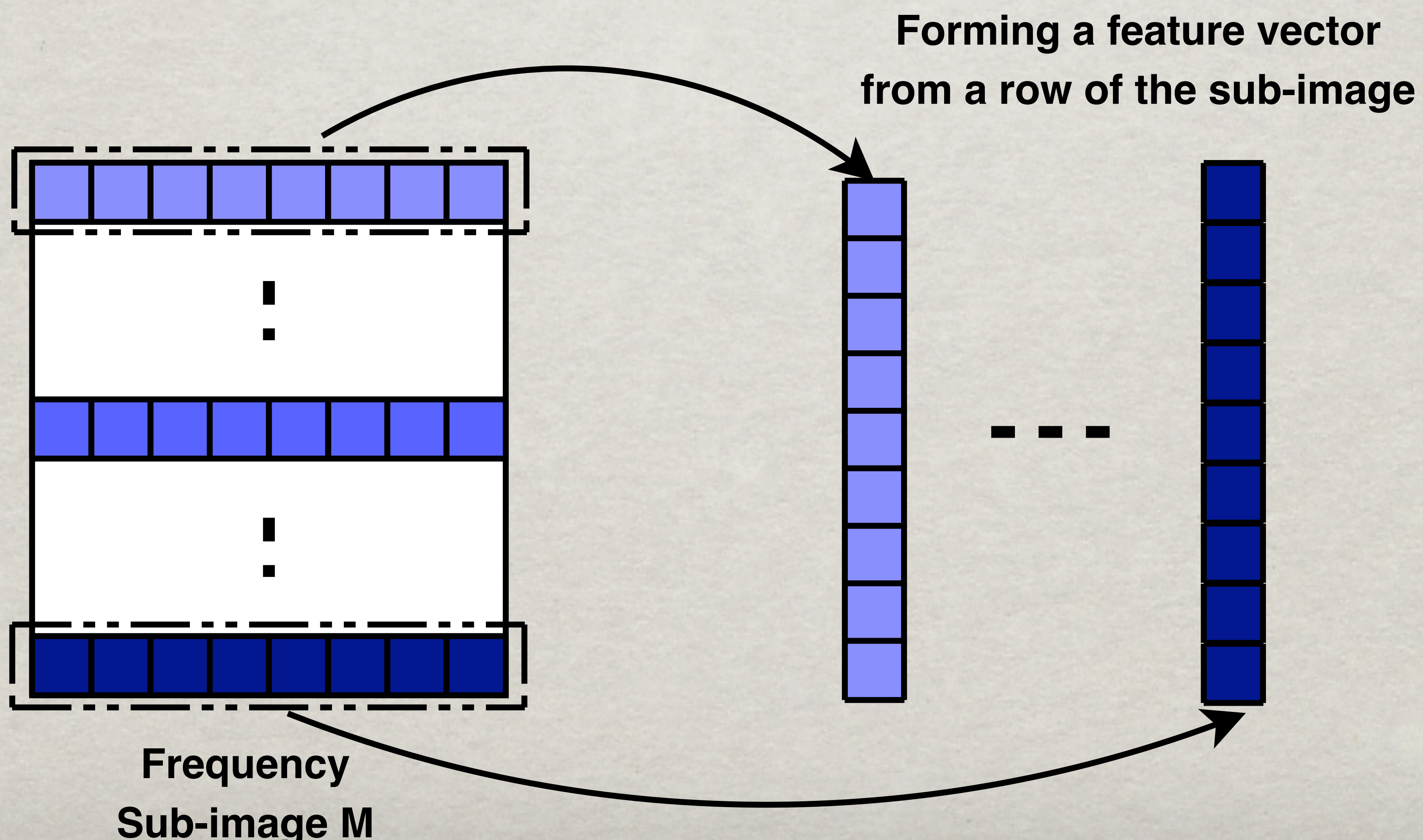
- A similar Parts-Based approach is applied to these frequency sub-images
  - Features are obtained from these sub-images
  - These features are used to derive a GMM classifier for each frequency sub-image
- There are two issues to deal with here:
  - How do we get feature vectors from the frequency sub-image?
  - How do I combine the information from each classifier?

Local Frequency Subband Approach

- Extracting feature vectors from the sub-band images:
    - form a feature vector along a row of the frequency sub-image
    - form a feature vector along a column of the frequency sub-image
    - form a feature vector from blocks of the frequency sub-image

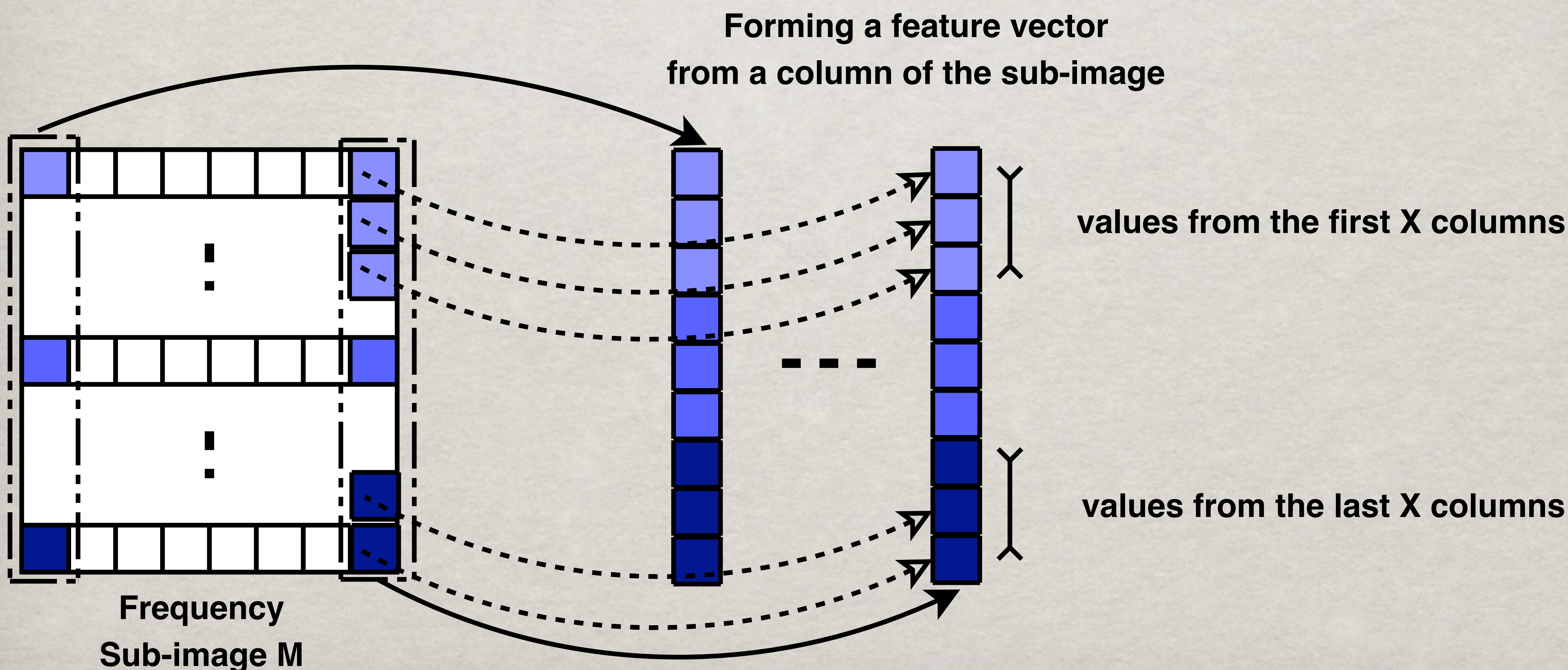
Row-Based Feature Vectors

- Forming a **set of feature vectors** across the row of a frequency sub-image



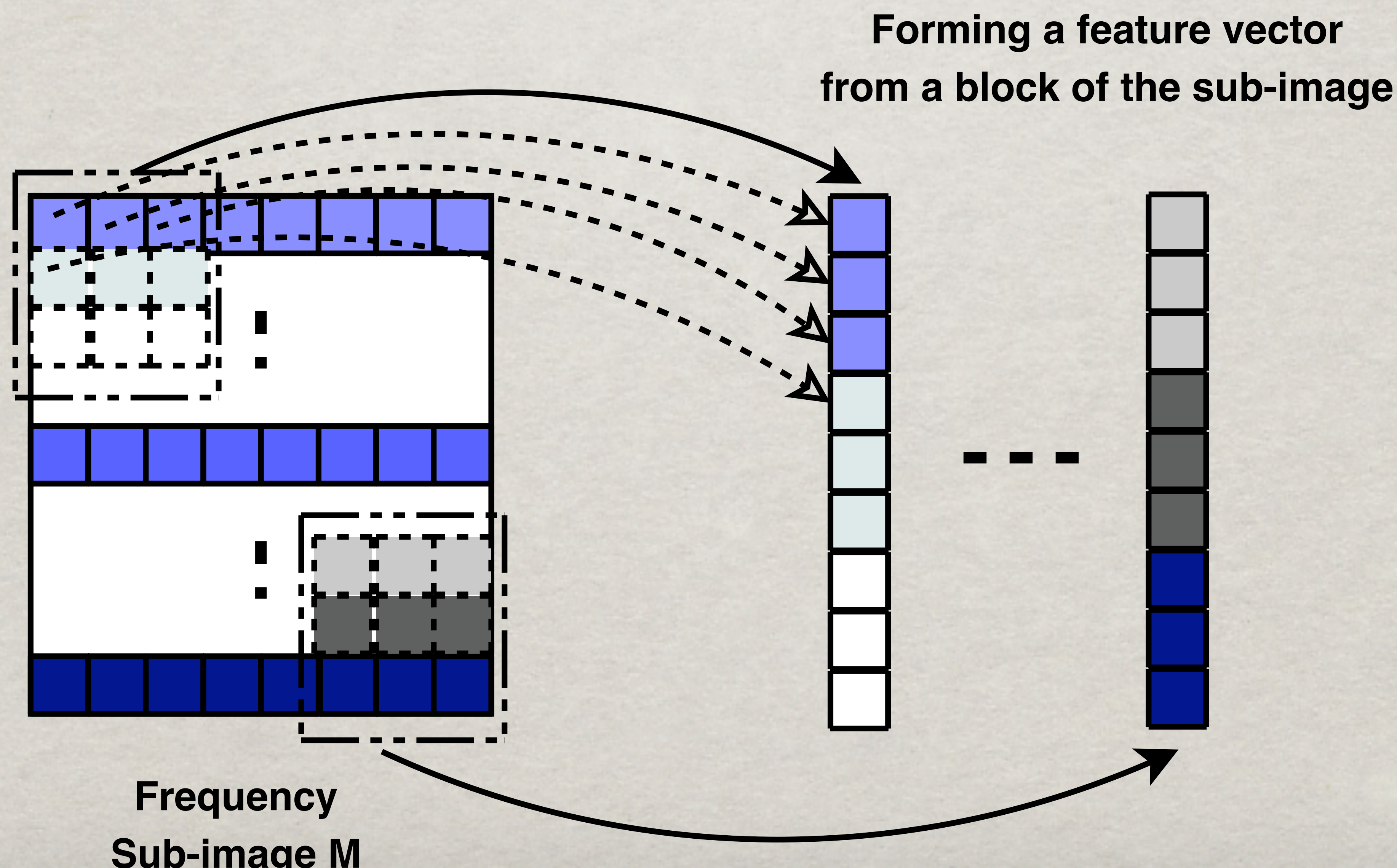
Column-Based Feature Vectors

- Forming a **set of feature vectors** across the column of a frequency sub-image



Block-Based Feature Vectors

- Forming a **set of feature vectors** from blocks of a frequency sub-image



Classifier Fusion

- Fusion was performed using weighted fusion
  - It's robust to estimation errors and is a relatively well used method for fusion
    - The weights are learnt on the tuning data set using linear logistic regression

$$C_{fused} = \sum_{i=1}^{D} \beta_i C_i$$

Results

- This method was tested on BANCA database
    - ~6,500 images, several well defined protocols
- In this presentation we only present results for P protocol
- We compare the performance of manual and automatic eye locations as well
- Performance: Average Half Total Error Rate (HTER)
  (Average False Acceptance Rate + Average False Rejection Rate)/2

Results

- The first few results are comparing the sub-band approaches against one another
  - Manual Eye Locations to get an idea of “optimal” performance
  - Automatic Eye Locations to get an idea of “real” performance

	Manual Eye Locations	Automatic Eye Localtions
Baseline GMM	26.59%	27.84%
Row Features	19.73%	26.58%
Block Features	18.05%	21.57%
Column Features	**14.85%**	**16.62%**

Results

- The first few results are comparing the sub-band approaches against one another
  - Manual Eye Locations to get an idea of “optimal” performance
  - Automatic Eye Locations to get an idea of “real” performance

	Manual Eye Locations	Automatic Eye Localtions
Baseline GMM	26.59%	27.84%
Row Features	19.73%	26.58%
Block Features	18.05%	21.57%
Column Features	**14.85%**	**16.62%**

The performance degrades quite badly when compared to manual eye locations ~3-7% worse

Results

- The first few results are comparing the sub-band approaches against one another
  - Manual Eye Locations to get an idea of “optimal” performance
  - Automatic Eye Locations to get an idea of “real” performance

	Manual Eye Locations	Automatic Eye Localtions
Baseline GMM	26.59%	27.84%
Row Features	19.73%	26.58%
Block Features	18.05%	21.57%
Column Features	**14.85%**	**16.62%**

Results

- The first few results are comparing the sub-band approaches against one another
  - Manual Eye Locations to get an idea of “optimal” performance
  - Automatic Eye Locations to get an idea of “real” performance

	Manual Eye Locations	Automatic Eye Localtions
Baseline GMM	26.59%	27.84%
Row Features	19.73%	26.58%
Block Features	18.05%	21.57%
Column Features	**14.85%**	**16.62%**

The performance is relatively stable ~1-2% worse

Conclusions and Future Work...

Conclusions

- Extended the GMM Parts-Based approach: by applying spatial and frequency decomposition
    - Obtained significant improvements
    - The work for manual and automatic eye locations
- Where to from here?
    - Perhaps this should be extended to be convolution (a pixel by pixel formation of the sub-band images) and then try to obtain features from these images

?