Abstract

For a computer vision system, the task of recognizing human faces from single pictures is made difficult by **variations** in face position, size, expression, and pose (front, profile, ...). We present an automatic system able to recognize human faces on the basis of single grey-level mug-shots matched against a large data set including one image per person (100-250 persons). The system performance and robustness is assessed, and is compared with other systems.

2D views of faces are represented by **labeled graphs**. Graph nodes are labeled with **jets** and graph edges are labeled with distance vectors. Jets are 80-dimensional vectors based on a 2D **Gabor-wavelet transform**. We use 40 complex Gabor-wavelets, which are localized filters, each of a certain spatial frequency and orientation (8 orientations and 5 spatial frequencies). Jets are a concise and robust representation of local grey-level value regions of the image. Since we want to compare faces across pose, we define a set of about 45 facial (or *fiducial*) points at which nodes are positioned. These points are identical in different poses (to the extent they are visible). Points such as the tip of the nose, the corners of the mouth, the pupils, the tip of the chin are included. The graphs thus have an object adapted grid and compatible nodes in different views can be compared with each other.

Abstract

To be able to process a large number of new faces, we have introduced a graph structure, the **face bunch graph**. It is constructed from a representative set (typically 70) of model graphs having the same pose, (e.g. frontal), hence the same graph structure. Each node of the bunch graph is then labeled with all the jets taken from the models at the same fiducial point. For instance, in frontal view, the left-eye node of the face bunch graph has attached different left-eye jets taken from the frontal view model graphs. New faces can be encoded by taking jets from different models at each node, e.g. left-eye jet taken from model 3 while nose jet taken from model 25. This takes advantage of the full combinatorial power of the bunch graph and makes it possible to represent and process faces not seen before.

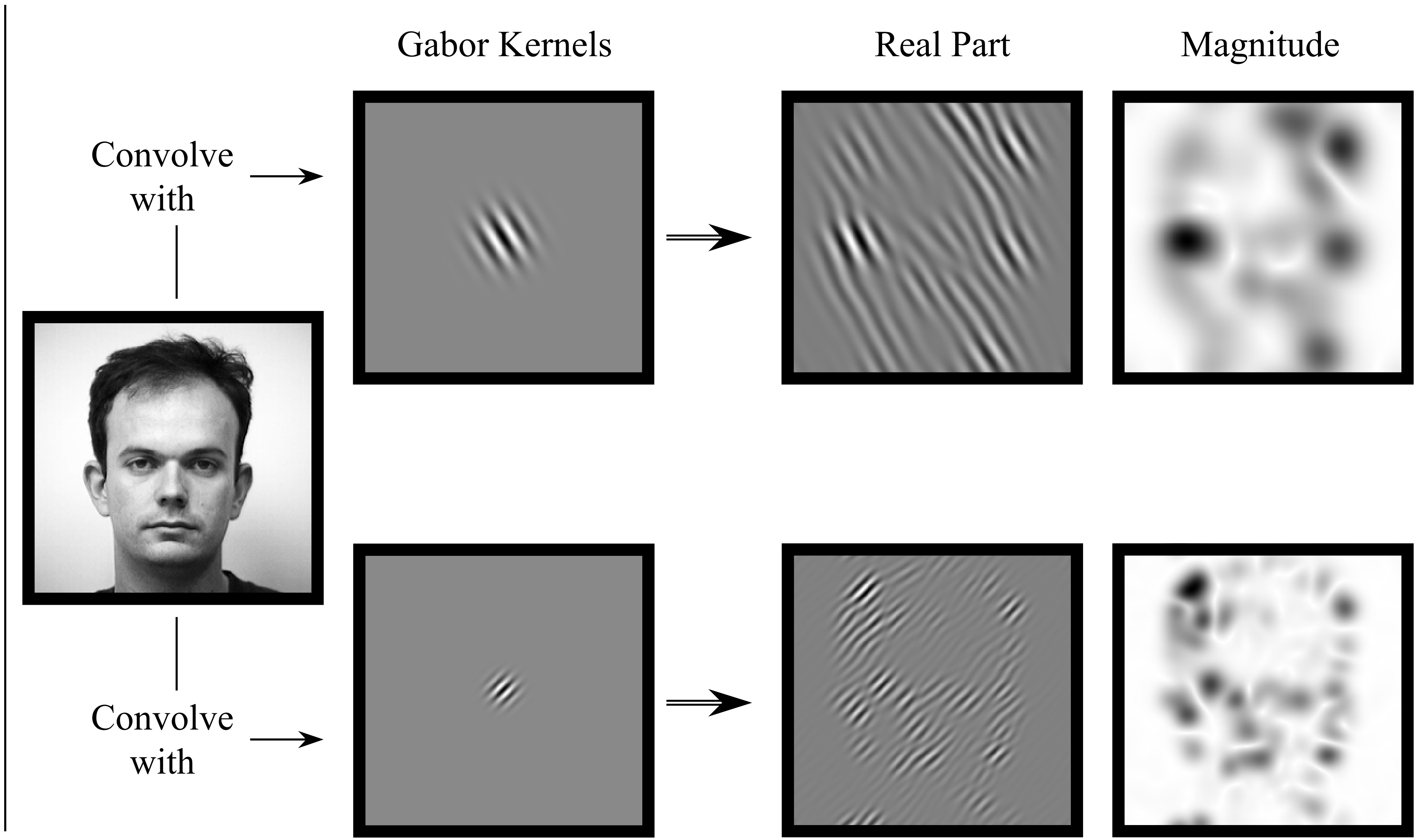
For new pictures, new graphs are automatically generated using **elastic bunch-graph matching**, which is guided by a **graph similarity function** defined between a face bunch graph and a new image graph. This similarity function accounts for spatial distortion and is based on the **jet similarities** between image jets and the best fitting jets in each bunch. Finally, once generated, the new graph is compared with stored graphs and the example with the highest average similarity is taken as the recognized person. This procedure is also possible between different views, because jets representing the same fiducial points across views can be associated. (i.e. recognizing a half-profile from a database of frontals).

The (pre-processed) Database

Frontals | Half Profiles (right/left) | Profiles (right/left)

The Database typically includes 250 entries. Pictures are 128x128 pixels, with 255 grey levels.

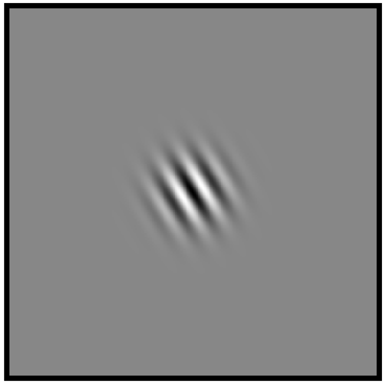
Gabor Wavelet Decomposition



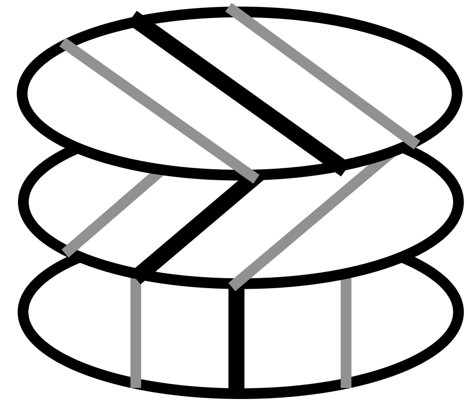
Jet

Real Parts

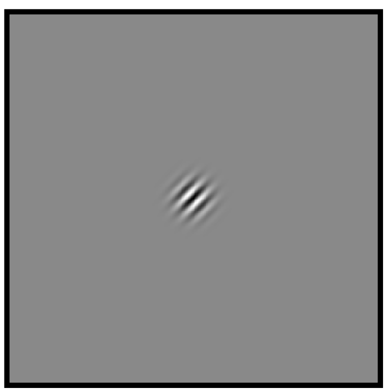
Low Frequency



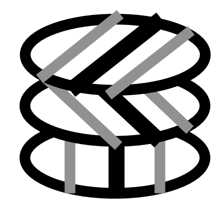
X 3 Orientations



High Frequency

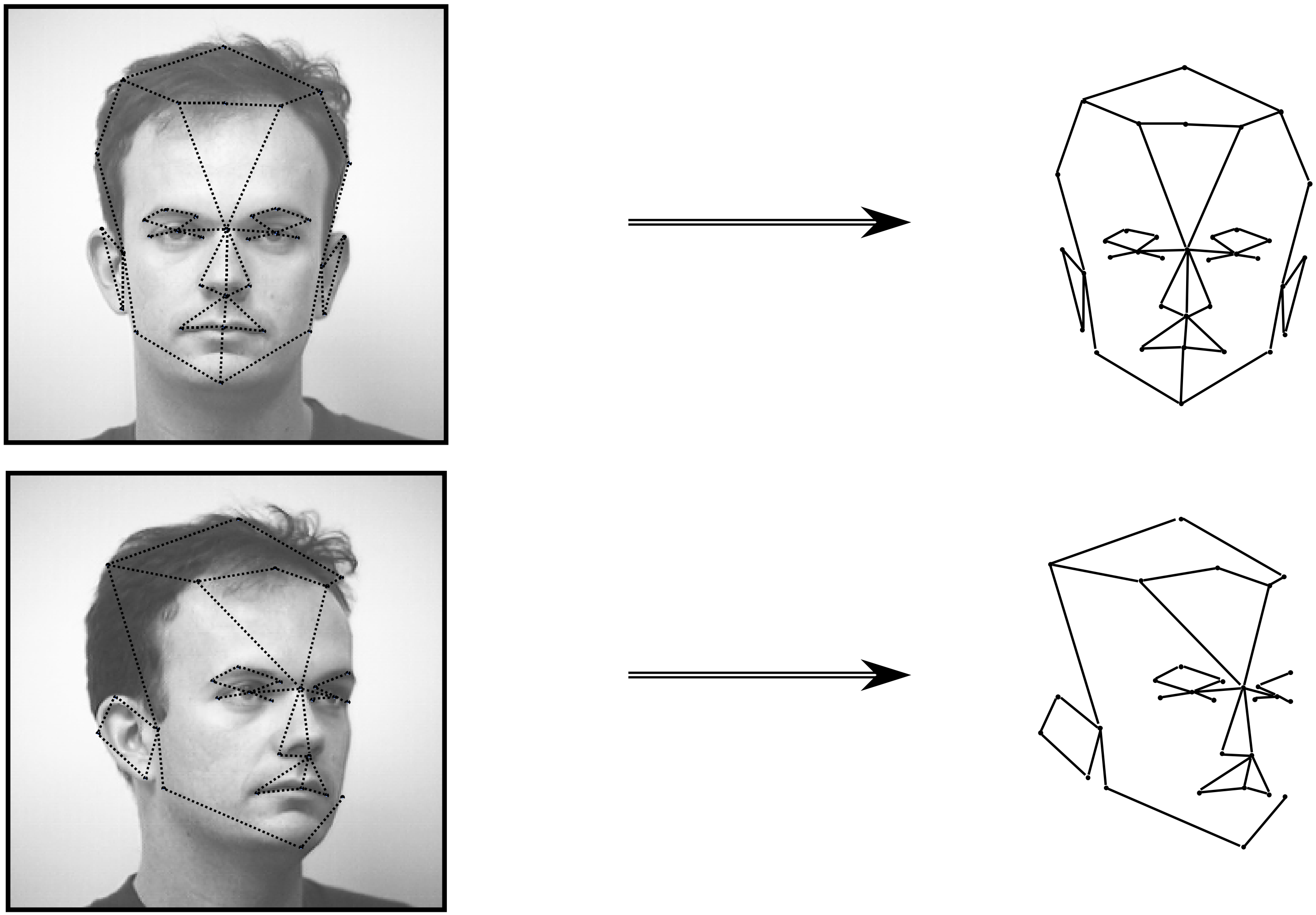


X 3 Orientations



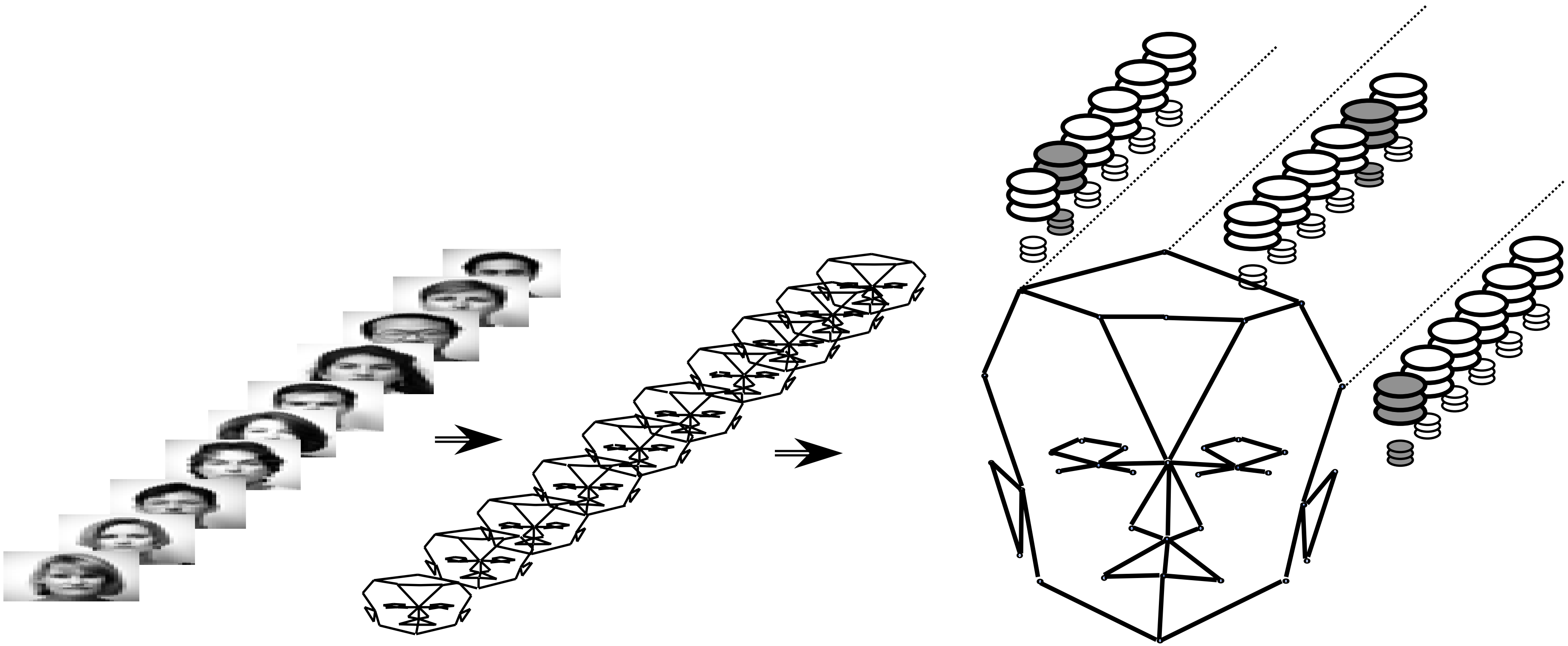
Note: We use real and imaginary parts, at 5 spatial frequencies, and 8 different orientations.

Labeled Fiducial Graph



Jets are extracted at the nodes of the fiducial graphs only. We use about 45 nodes per pose.

The Face Bunch Graph (FBG)

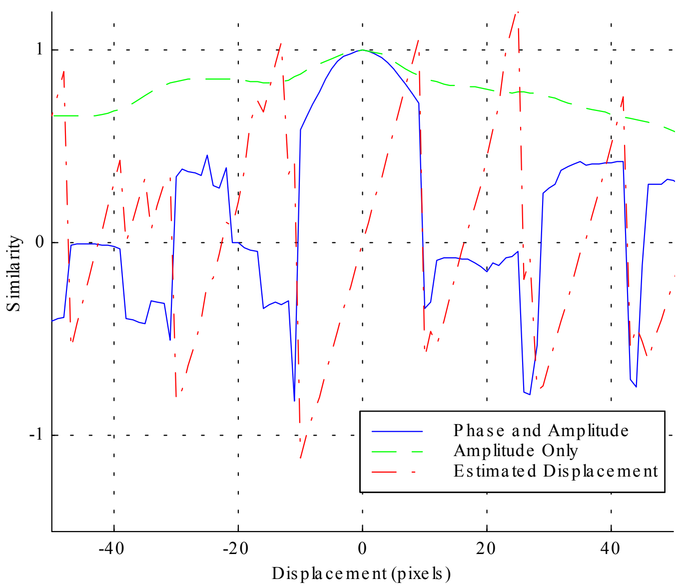


The Bunch Graph is built on the basis of a limited set of examples, of a given pose (frontal, half-profile, ...). We use about 70 examplars per FBG.

The Jet Similarity Functions

Jet

$$J = \begin{pmatrix} a_1 e^{i\phi_1} \\ \cdots \\ a_j e^{i\phi_j} \\ \cdots \\ a_{40} e^{i\phi_{40}} \end{pmatrix}$$



We use 2 measures of Jet-to-Jet comparisons:

Amplitude Only

$$S_a(J, J') = \frac{\sum_j a_j a'_j}{\sqrt{\sum_j a_j^2 \sum_j a'^2_j}}$$

Amplitude and Phase

$$S_\phi(J, J') = \frac{\sum_j a_j a'_j \cos(\phi_j - \phi'_j - \vec{d}.\vec{k}_j)}{\sqrt{\sum_j a_j^2 \sum_j a'^2_j}}$$

$$\vec{d} = f(J, J')$$

The Graph Similarity Function

The Bunch Graph (*B*) contains ***M*** examples. Each graph has ***N*** nodes and ***E*** edges. The similarity between fiducial Graph G, and Bunch Graph B is given by:

$$S(G,B) = \frac{1}{N}\sum_{n}\max_{m}(S_{\phi}(J_n^G, J_n^{B_m})) - \frac{\lambda}{E}\sum_{e}\frac{\left(\Delta\vec{x}_e^{\,G} - \Delta\vec{x}_e^{\,B}\right)^2}{\left(\Delta\vec{x}_e^{\,B}\right)^2}$$

Feature (Jets) comparison term.

Metric (distances) comparison term.
(Distortions)

Overall Procedure

Pre-processing (database-dependent).

Cropped, padded, resized.

Construction of the FBGs (one per pose).

70 ‘charasterisc’ examples - Manual fiducial graph construction.

Constitution of the databases (one per pose).

Use FBG to locate fiducial points automatically.

Store Fiducial graphs (~45 nodes).

Identification of new incoming face.

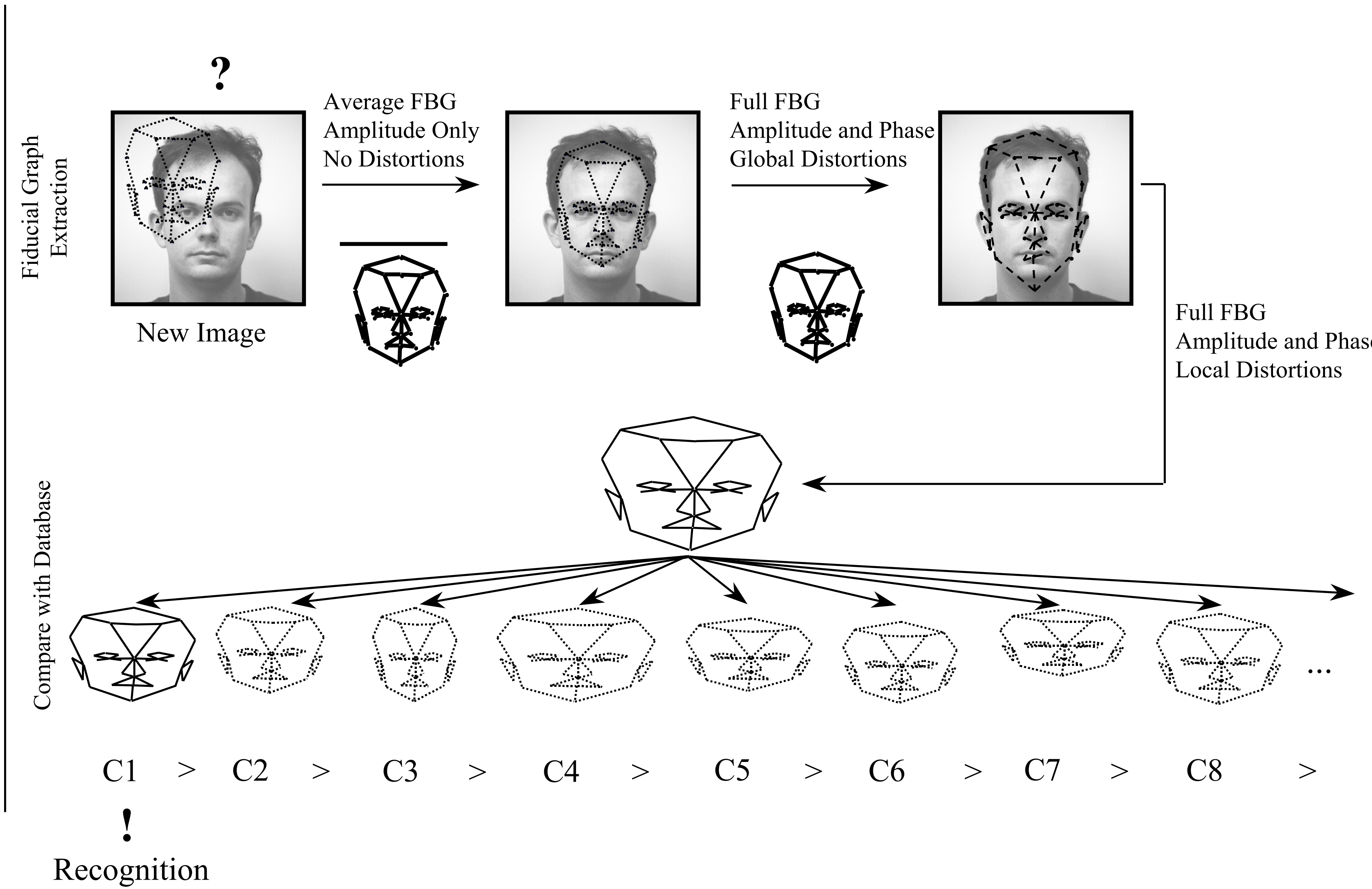
(Use FBGs to determine pose).

Use FBG to locate fiducial points automatically.

Compare with all database entries. Classify results by decreasing order of similarity.

Make decision.

Elastic Bunch Graph Matching



Conclusions

On the ARL/FERET database recognition rates using databases of 250 persons are 98% for frontal poses, 84% for profiles, and 57% for half-profiles. Performance degrades to the range of 10-20% for recognition across poses.

Database	Probe	1st Rank #	1st Rank %	10 Ranks #	10 Ranks %
250 fa	250 fb	245	98	248	99
250 hr	181 hl	103	57	147	81
250 pr	250 pl	210	84	236	94
249 fa + 1 fb	171 hl + 79 hr	44	18	111	44
171 hl + 79 hr	249 fa + 1 fb	42	17	95	38
170 hl + 80 hr	217 pl + 33 pr	22	9	67	27
217 pl + 33 pr	170 hl + 80 hr	31	12	80	32

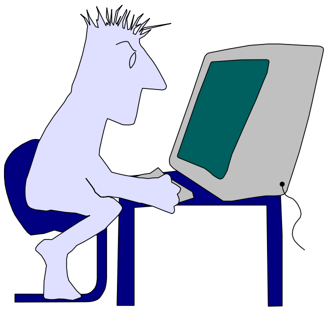
fa, fb: Frontals with different facial expressions.
hr/l: Half-profile right/left.
pr/l: Profile right/left.

Note: Chance level is about .5%

Conclusions

Supported by grants from the German Federal Ministry for Science and Technology (413-5839-01 IN 101 B9) and from ARPA and the U.S. Army Research Laboratory (01/93/K-109).

The Poster is On-line at:



http://emotion.ccs.brandeis.edu/JM/Poster/BU97Poster.htnl