Employing Vector Quantization on Detected Facial Parts for Face Recognition

Ahmed Aldhahab, Taif Alobaidi, and Wasfy B. Mikhael
Department of Electrical and Computer Engineering , University of Central Florida, Orlando, FL USA
Aldhahab2012@knights.ucf.edu, taif.alobaidi@knights.ucf.edu, and Wasfy.Mikhael@ucf.edu

Abstract

Facial Parts Detection (FPD) approach in conjunction with Vector Quantization (VQ) algorithm are proposed for face recognition. Detecting facial parts, which are nose, both eyes, and mouth, and choosing appropriate dimensions for each part are done in the preprocessing phase. In the feature extraction phase, four groups for each person, one group for each detected part, are constructed for dimensionality reduction and feature discrimination by considering all parts of all training poses. For further data compression, VQ algorithm is applied to each of the four groups. Finally, Euclidean distance criterion is used to obtain the recognition rates. Four databases, namely, ORL, YALE, FERET, and FEI are used to evaluate the proposed system. Then K-Fold Cross Validation (CV) is used to analyze the results. The proposed system consistently improved the recognition rates as well as the storage requirements. Sample results are given.

Related Work and Applications

Introduction

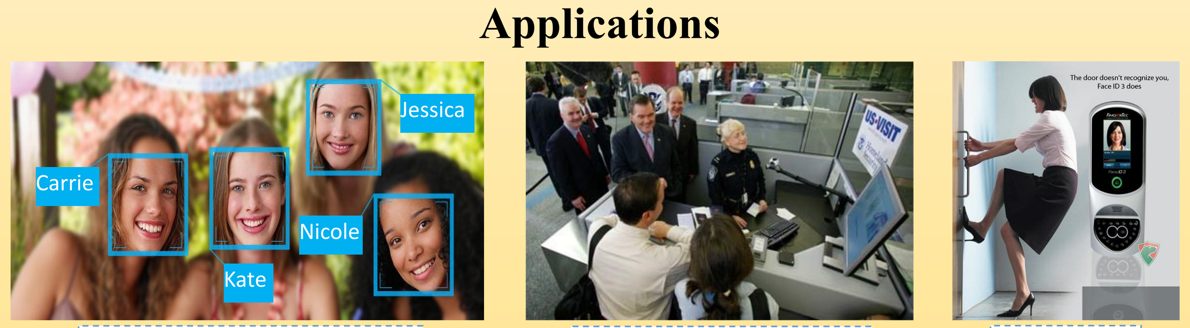
⇒ Facial recognition is an important task in computer vision, pattern recognition, and image processing, which has received renewed attention in recent years due to its wide applicability in **security**, **control**, and **personal identification, etc**.

⇒ The **efficiency** and **reliability** of the different recognition systems depend on several factors, including the **computational complexity**, **storage requirements**, and **recognition rates**.

Related Work

⇒ PCA and Wavelet Transform [**Mukhedka et al' 2015**].
⇒ Independent Component Analysis and DWT [**Kinage et al' 2010**].
⇒ Stationary DMWT [**Tarik et al' 2013**]
⇒ Discrete Cosine Transform &PCA [**Chelali et al' 2015**], and etc.

Applications



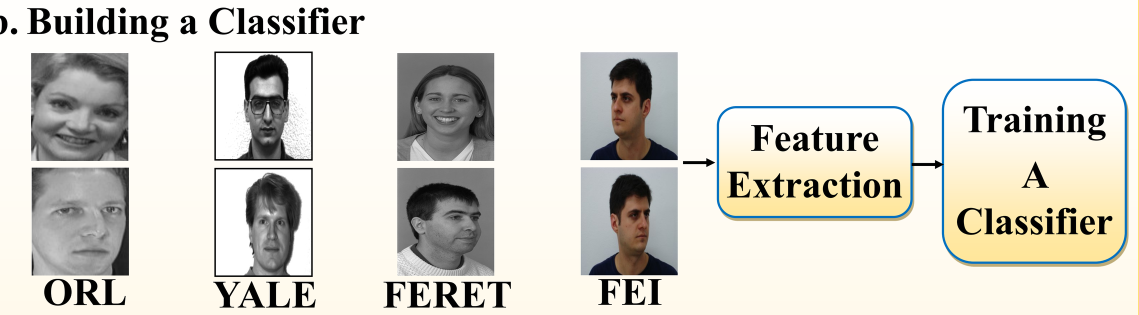
Personal Identification | US-VISIT (Security) | Control

Problem Statement

1. Training Phase

a. **Feature Extraction**

b. **Building a Classifier**



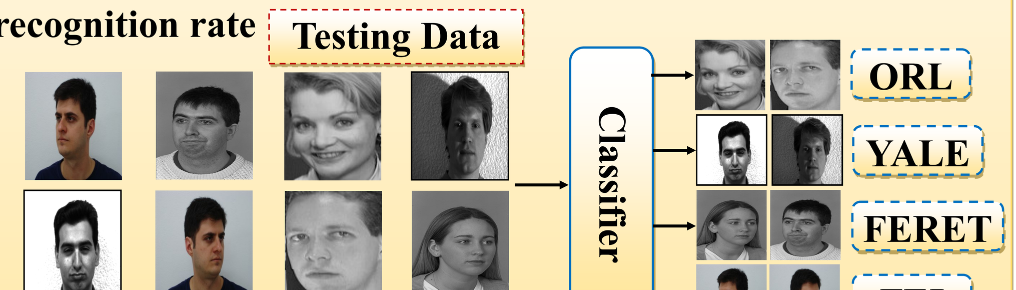
Training Data

2. Testing Phase

a. **Time-sensitive**

b. **High recognition rate**

Testing Data

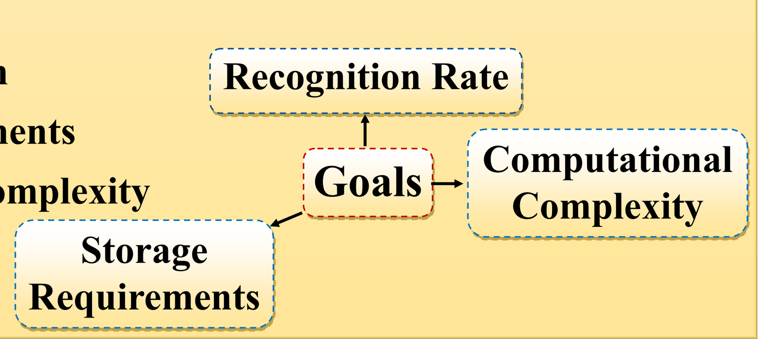


3. Goals

a. **Dimensionality reduction**

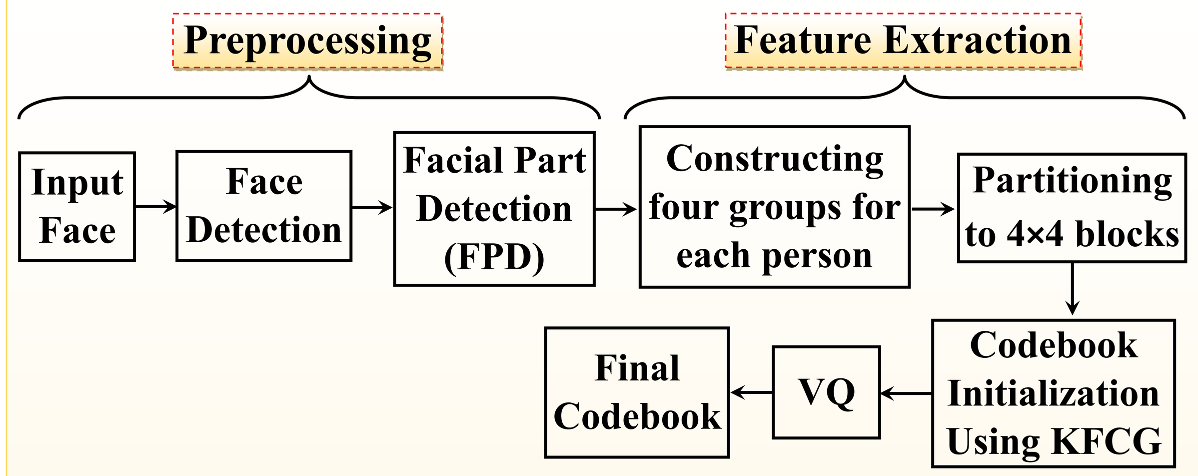
b. **Reduce storage requirements**

c. **Reduce computational complexity**

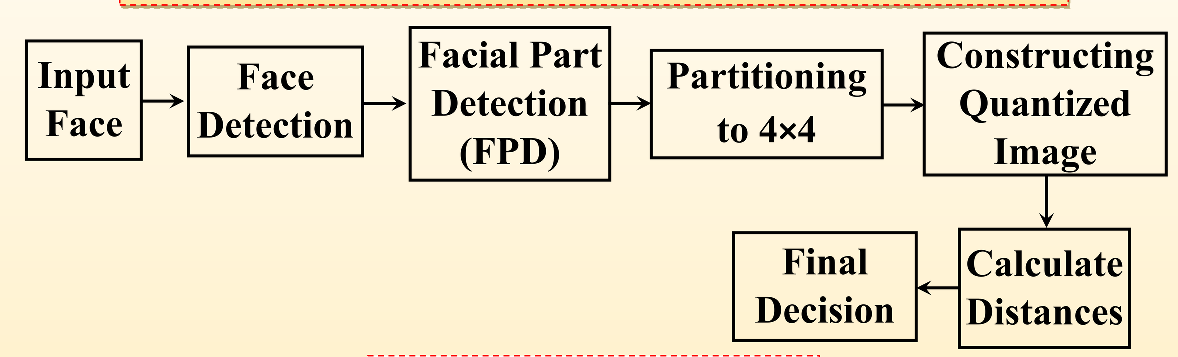


Proposed System

The Proposed System



Preprocessing and Feature Extraction Phase



Classification Phase

A face image may contain information that is not essential for the recognition purpose. Hence, the following effective tools are applied to get an efficient representation of the images:

1. Facial Part Detection is used for:

a. Identifying human in Digital Camera.
b. Locating faces in the crowd.
c. Dimensionality reduction/ Background elimination.

2. The VQ algorithm is used for:

a. Data Compression.
b. Facial image representation by using its codebook as a feature.

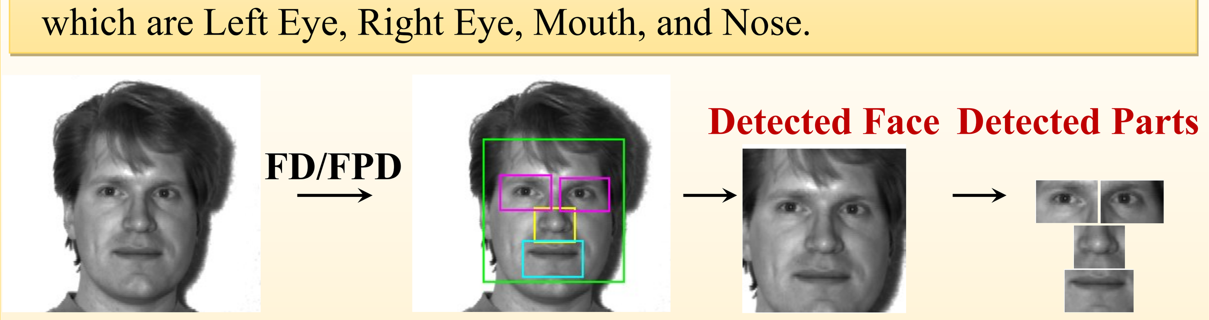
Our approach is based on the FPD and VQ. The VQ algorithm, which uses Kekre Fast Codebook Generation (KFCG), is applied to the resultant FPD features for further feature compression and better facial representation. It is shown to significantly improve the classification rate as well as reduce the storage requirements.

Proposed System

1. Preprocessing

⇒ The first step aims to detect faces from the images using Face Detection Algorithm that employs Viola-Jones algorithm.

⇒ FPD algorithm is applied to the resultant of step 1 to detect facial parts, which are Left Eye, Right Eye, Mouth, and Nose.



Face/facial parts detection of the first Person of YALE database

⇒ Unify the dimensions among all parts. The dimensions of the detected face and parts (Left Eye, Right Eye, Mouth, and Nose) are converted to the appropriate ones as shown in the Table below.

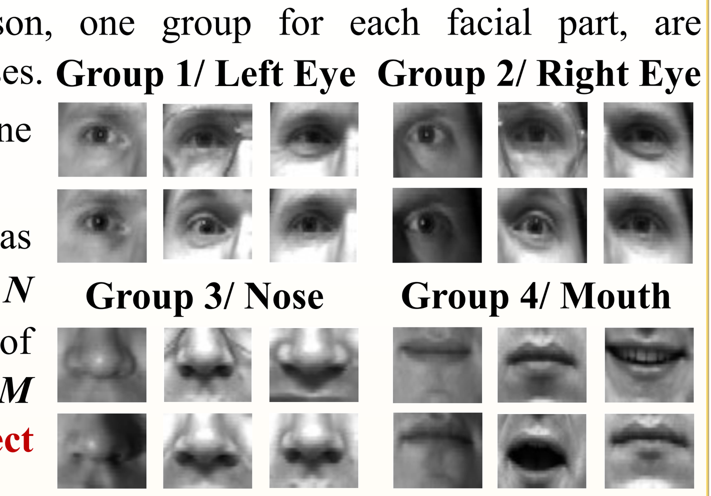
Table I. The Dimension of The Databases

Databases	Actual Size	Faces and Facial Parts Detected				
		Face	Left Eye	Right Eye	Mouth	Nose
ORL	112 × 92	64× 64	32×32	32×32	32×32	32×32
YALE	243 × 320	128×128	32×32	32×32	32×32	32×32
FERET	384 × 256	128 × 128	32×32	32×32	32×32	32×32
FEI	480 × 640	256 × 256	64×64	64×64	64×64	64×64

2. Feature Extraction Step

Training Mode:

1. Four groups for each person, one group for each facial part, are established using different poses.
2. Each group belonging to one detected facial part.
3. In general, each group has dimensions of $N \times M$, where N and M depend on the number of training poses used, and N=M when *the group* is a **perfect square**.



Constructing 4 groups using 6 training poses.

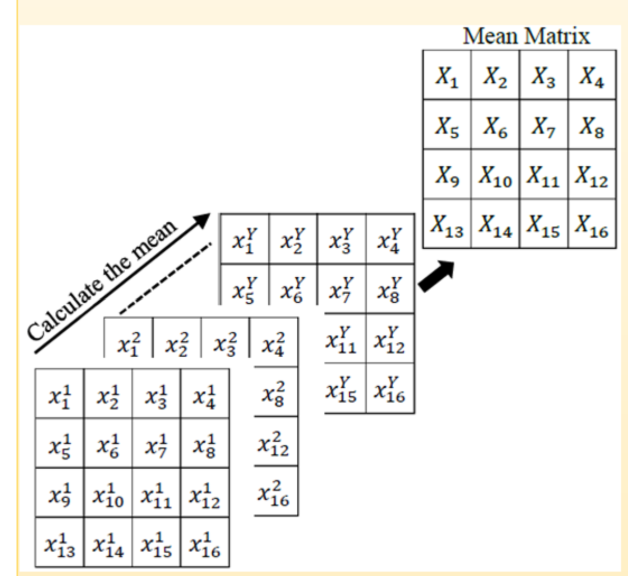
Vector Quantization (VQ)

VQ is a lossy compression technique. Linde-Buzo-Gray (**LBG**) is the well-known algorithm for **VQ**. **The design parameters chosen are:**

a. Centroid (codeword) dimensions are **4 × 4 (q = p = 4)**.
b. Codebook size is **C = 16** (number of centroids required for each group).
c. Mean Square Error (**MSE**) is used for the distortion criterion.

To perform the VQ, the following processes are required:

1. Each group of features is **partitioned into non-overlapping blocks**, each with **4 × 4 (p × q)** dimensions.
2. Calculate the first mean of the matrix, which can be expressed as:



$$Y = \frac{\text{group dimensions}}{\text{codeword dimensions}} \quad \text{--- 1}$$

$$X_i = \frac{x_i^1 + x_i^2 + \cdots + x_i^Y}{Y} \quad \text{--- 2}$$

Where X_i is the i^{th} average of all corresponding elements across all blocks. $i \in Z : \{ i = 1, 2, \ldots , p \times q \}$. x is the i^{th} element in the y^{th} block, $y \in \{ 1, 2, \ldots, Y \}$, Y is the total number of blocks.

3. The **KFCG** method is used to generate the initial **VQ** codebook.
4. Then, the **LBG** algorithm is applied to calculate the new codebook.
5. The final feature extracted for each person has **4 × Centroid (4×4×16)** dimensions **regardless the number of poses used in the training mode**.

Dimensionality Reduction:

Compared with the dimensions of final features extracted, the dimensions of the input poses are reduced by:

$$\text{Dimentionality Reduction} = \left(1 - \frac{4 \times \text{Centroids}}{\text{Trained Poses} \times 243 \times 320}\right) \times 100\% \quad \text{--- 3}$$

3. Classification

Testing Mode:

1. The same preprocessing steps are applied to the test face image.
2. All codebooks of all groups of all subjects in the database are used to reconstruct the quantized versions of the detected part.
3. The Euclidean distances, between detected part and the reconstructed ones, are calculated.
4. Then, the detected part will match the person whose codebook has the minimum distance.
5. If **Three** out of **Four** detected facial parts are matched to the same person, that test pose will declare as a correctly matched.
6. The recognition rates are measured as:

$$\text{Recognition Rates} = \frac{\text{Total Number of Poses Correctly Matched}}{\text{Total Number of Poses in The Dtatabase}} \times 100\% \quad \text{--- 4}$$

Experimental Results

Our technique is tested on four databases that have a large number of persons, poses, and lighting conditions. **K-fold Cross Validation** is used to evaluate our techniques.

A. Experimental Results for the ORL Database

There are **40 persons** each with **10 poses**. Table II summarizes the results.

Table II: Experimental Results for ORL Database

K-Fold CV	*Recognition Rates for the Proposed System*		2D DCT [Karhan 2013]	2D DWT/PCA [Mukhedka 2015]	VQ-KFCG [Natu 2010]
	Training Mode	*Testing Mode*			
K=2	100%	95.33%	88%	92.5%	89%
K=5	100%	98.25%	94.25%	96.5%	94.75%

B. Experimental Results for the YALE Database

There are **15 persons** each with **11 poses**. Table III summarizes the results.

Table III: Experimental Results for YALE Database

K-Fold CV	*Recognition Rates for the Proposed System*		2D DCT [Karhan 2013]	2D DWT-PCA [Mukhedka 2015]	VQ-KFCG [Natu 2010]
	Training Mode	*Testing Mode*			
K=2	100%	95.75%	87.33%	92.12%	88.33%
K=5	100%	98.51%	93.1%	95.15%	94.17%

C. Experimental Results for the FERET Database

There are **200 persons** each with 11 poses. Table IV shows the results.

Table IV: Experimental Results for FERET Database

K-Fold CV	*Recognition Rates for the Proposed System*		2D DCT [Karhan 2013]	2D DWT-PCA [Mukhedka 2015]	VQ-KFCG [Natu 2010]
	Training Mode	*Testing Mode*			
K=2	100%	95.01%	87.5%	92.5%	88.21%
K=5	100%	97.98%	92.21%	95.68%	93.57%

D. Experimental Results for the FEI Database

There are **200 persons** each with 14 poses. The facial images here are rotated by 180 degrees. Table V summarizes the results.

Table V: Experimental Results for FEI Database

K-Fold CV	*Recognition Rates for the Proposed System*		2D DCT [Karhan 2013]	2D DWT-PCA [Mukhedka 2015]	VQ-KFCG [Natu 2010]
	Training Mode	*Testing Mode*			
K=2	100%	95.18%	87.32%	92.25%	87.5%
K=5	100%	97.92%	92%	95.58%	93.81%

Conclusion & Acknowledgement

⇒ A facial recognition system based on **Facial Part Detection (FPD)** and **Vector Quantization (VQ)** was proposed.
⇒ Four Groups for each person, one for each detected part, are constructed.
⇒ **VQ**, employing **KFCG** for codebook initialization, was applied to each group to achieve further feature compression.
⇒ Each person was efficiently represented using **four centroids**.
⇒ The extracted features of each person were of size **4 × Centroid**, which is **4×4×4×16 (256),** while the training poses were of size shown in **Table I.**
⇒ The **recognition rates** achieved were improved compared with those reported by **[Karhan 2013, Mukhedka 2015, and Natu 2010].**
⇒ **K-fold CV** was used to analyze the experimental results.
⇒ **Recognition accuracies**, thus realized, were **98.25%, 98.51%, 97.98%, and 97.92%** for **ORL, YALE, FERET, and FEI** databases, respectively.

Acknowledgement

This work was supported by the Iraqi government scholarship (HCED).