2

Large-Scale Face Image Retrieval: *A Wyner-Ziv Coding Approach*

Jean-Paul Kouma and Haibo Li
Digital Media Lab
Department of Applied Physics and Electronics,
Umeå University
Sweden

1. Introduction

Great progress in face recognition technology has been made recently. Since the first face recognition vendor test (FRVT)Phillips et al. (2007) in 1993, face recognition performance has been improved by two orders of magnitude in thirteen years. Notably, in the FRVT 2006 it is the first time that algorithms are capable of human performance levels, and at false acceptance rates in the range of 0.05, machines can outperform humans Phillips et al. (2007). The advances are promising for face verification applications where a typical one-to-one match is performed. It is still a grand challenge to power large-scale face image retrieval. Large-scale face image retrieval is the enabling technology behind the next generation search engines (search beyond words), by which web users can do social search with personal photos. High performance face identification algorithms are needed to support large scale face image retrieval. Compared with face verification, face identification is believed **N** times harder than face verification due to its nature of **1:N** problems. The number of individuals **N** in the database has a great impact on both the *effectiveness* and *efficiency* of face identification algorithms.

With the state of the art face identification algorithms, the identification rate is only around 70% (rank = 1) for the FERET database, a gallery of ten thousands individuals. When to serve for large-scale face image retrieval applications, the identification rate will further decrease as the gallery size increase (fortunately not linearly but logarithmically).The computing complexity of face identification is linearly related to the number of individuals N. For large-scale face image retrieval the efficiency of face identification is a key issue. In this paper we focus on the efficiency aspects of face identification.

Technically, it is very challenging to find a person from a very large or extremely large database which might hold face images of millions or hundred millions people. A highly efficient image retrieval technology is needed. Indexing technology based on tree structures has been widely used in commercial search engines. These structures are quite efficient for small dimensions (of the order of 1-10). However, as the data dimensionality increases, the query performance of these structures degrades rapidly. For instance, White and Jain report that as the dimensionality increases from 5 to 10, the performance of a nearest-neighbor query

in multi-dimensional structures such as the SS-tree and the R-tree, degrades by a factor of 12. This phenomenon, known as the *dimensionality curse* is a common characteristic of all multi-dimensional index structures. In spite of the progress in the design and analysis of multi-dimensional structures such as the TV-trees, the X-trees, and the SR-trees, the dimensionality curse persists.

A very efficient approach to large-scale image retrieval is to use an approximate similarity searching strategy Tuncel et al. (2004). Without building an indexing mechanism, the search engine simply accesses *partial* information about *all* the feature vectors. Popular examples of this approach are the VA-file algorithm Weber & Blott (1997), and the dimensionality reduction techniques. Feature vectors are approximated using the accessed partial information. In the state of the art face recognition techniques, all face images in a gallery are transferred into lower resolution used for feature vectors (called face signatures). Two examples are face signatures (images) of 21-by-12 pixels used in the statistical subspace methodShakhnarovich & Moghaddam (2004) and face signatures of 28-by-23 pixels used in our 1D HMM methodLe (2008). The problem is that due to huge number of individuals in a larger gallery all signatures are too big to fit into a single server's memory. The signatures have to be stored in hard disks. The number of disk I/O operations will be the bottleneck for query processing. The processing time is approximately proportional to the size of a signature image. Therefore, compression of face signatures will play a central role in large-scale face image retrieval. High compression will lead to a fast retrieval but the distortion due to compression will affect the retrieval quality. Therefore, two types of scientific challenges are:

- *How to characterize the trade-off between retrieval quality and speed.*
- *How to efficiently compress face signatures under the fixed distortion.*

This chapter brings together our earlier works in a more detailed and coherent whole. A shorter version can be found in Kouma & Li (2009). Our contributions are:

1. *To treat the image retrieval problem as a source coding problem and the rate-distortion theory* ***R(D)*** *is used to characterize retrieval quality* (***D****: distortion of coding) and retrieval speed* (***R****: rate of coding)*
2. *To view compression of signature images it as a typical "Wyner-Ziv Coding" problem, which circumvents the problem that the query images are not available until we decompress the signature image*
3. *To develop a distributed coding scheme based on LDPC codes to compress face signature images*

2. Image retrieval as a Wyner-Ziv coding problem

In the 1970s Slepian and Wolf had already proved that efficient compression can also be achieved by exploring source statistics partially or wholly at the decoder only. This was known as distributed lossless coding Slepian & Wolf (1973). This is illustrated in Figure 1 by an example of compressing an information source **X** in the presence of side information **Y** when **X** and **Y** are two correlated sources. For the simplest case when **Y** is not known at all, **X** can be compressed at the rate larger than or equal to **H(X)**. When **Y** is known at both the encoder and decoder, the problem of compressing **X** is also well understood: one can compress **X** at the theoretical rate of its conditional entropy **H(X|Y)**. But what if **Y** is known

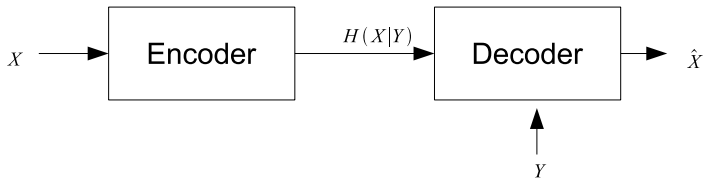

Fig. 1

only at the decoder for **X** and not at the encoder? The surprising answer from the Slepian-Wolf coding theorem Slepian & Wolf (1973) is that one can still compress **X** using only **H(X | Y)** bits, the same bits as the case where the encoder does know **Y**! This was extended to the lossy encoding by Wyner and Ziv Wyner & Ziv (1976) and yielded a similar result: *Under certain conditions, as when* **X** *and* **Y** *are jointly Gaussian with the MSE measure, when the decoder knows* **Y**, *then whether or not the encoder knows* **Y**, *the rate-distortion performance for coding* **X** *is identical*.
In face retrieval applications, for a given signature **X**, **R** $\geq$ **H(X)** bits are needed to represent it. If the query image **Y** is from the same individual, then **Y** will be highly related to **X**. If we know **Y** in advance then we don't need to store the whole information of **X**, instead just the conditional information, **R** $\geq$ **H(X | Y)**. Obviously, **H(X | Y)** $\leq$ **H(X).** As mentioned, retrieval speed is determined by (linearly proportional to) the rate **R**, so it makes large sense to reduce the rate from **H(X)** to **H(X | Y)**. The challenge is that in practice, **Y** is not known in advance and we can't directly make use of the knowledge of **Y** to help the compression of **X**. The solution is to treat it as the Wyner-Ziv coding problem: *take the query image* **Y** *as the side information*. There will be no rate loss as long as **X** and **Y** are jointly Gaussian. In fact, in the state of art face recognition techniques Gaussian modeling of human faces are commonly used. For example, in statistical subspace methods Shakhnarovich & Moghaddam (2004), it is assumed that Δ**=X-Y** is a Gaussian distribution if **X** and **Y** are from the same individual. The Gaussian distribution is used to characterize intra-personal variations Ω, caused by different facial expressions, lighting, and poses of the same individual. According to Wyner-Ziv coding theory there is no rate loss when **Y** is available only at the decoder since it is the quadratic Gaussian case. The rate will be:

$$R_{WZ}(D) = R_{X|Y}(D) = \frac{1}{2}\sum_{i} \log\left(\frac{\text{œ}_{Z_i}}{D_i}\right) \tag{1}$$

where

$$D = \sum_{i} D_i \tag{2}$$

Note that **Y**can be arbitrarily distributed. The rate-distortion $R_{WZ}(D)$, put it in the image retrieval language, says the minimum time complexity (**R**) achieving retrieval distortion **D**. It governs the tradeoff between the retrieval quality (**D**) and retrieval speed **R** in practice.
Just as the information theory, the Wyner-Ziv theorem only tells us a theoretical bound on information rate but not how to reach the bound in practice. The image coding community has high interest in exploring how to design practical Slepian-Wolf and Wyner-Ziv codecs. With Low-density parity-check codes, when the code performance approaches the capacity

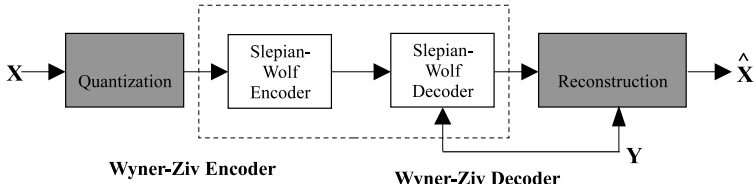

Fig. 2

of the correlation channel, the compression performance approaches the Slepian-Wolf bound (see referenced listed in Varodayan et al. (2005)). In contrast, efforts toward practical Wyner-Ziv coding have been undertaken only recently. Zamir and Shamai proved that linear codes and nested lattices might approach the Wyner-Ziv rate-distortion function if the source data and side information are jointly Gaussian Zamir et al. (n.d.). Xiong et al implemented a Wyner-Ziv encoder as a nested lattice quantizer followed by a Slepian-Wolf coder Xiong et al. (2003). In general, a practical Wyner-Ziv coder can be thought to consist of a quantizer followed by a Slepian-Wolf encoder, as illustrated in figure 2. This makes it possible for us to focus on two basic components: quantization and reconstruction. As an example of practical codec a Wyner-Ziv video coding system is reported to perform 10-12 dB better than H.263+ intra-fame codingVarodayan et al. (2005). In the face image retrieval, compression of face signature images are much more challenging than distributed video coding due to rather large variations between face images of the same person, which may be taken at different time, by using different cameras. In this paper we focus on using Slepian-Wolf coding to compress face signature images.

3. Low Density Parity-Check (LDPC) codes as Slepian-Wolf coder

Low Density Parity-Check Codes are intensively studied in other literatures, But for the sake of completeness we briefly review it here.
LDPC codes are a class of linear block codes. They were invented by Gallager in the early 60's. But due the computational complexity (at that time), LDPC codes were largely forgotten until the early 90's.
LDPC codes are specified by a sparse *parity-check* matrix $\mathbb{H}$, as well as a bipartite graph, introduced by Tanner Tanner (1981). Equation 3 and figure 3 show a parity-check matrix and its graphical representation, respectively. an LDPC code consists of N *variable nodes* (number of bits in a codeword) and M *check nodes* (number of parity bits). A check node c_m is *connected* to a variable node v_n if the element h_{ij} in $\mathbb{H}$ is 1.

$$A = \begin{bmatrix} 0\,1\,0\,1\,1\,0\,0\,1 \\ 1\,1\,1\,0\,0\,1\,0\,0 \\ 0\,0\,1\,0\,0\,1\,1\,1 \\ 1\,0\,0\,1\,1\,0\,1\,0 \end{bmatrix} \tag{3}$$

Typically a parity-check matrix is very big - of size over 2000 entries - and very sparse.

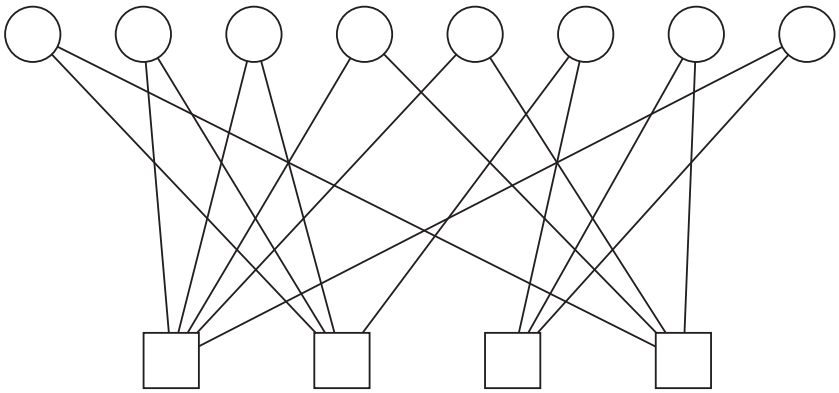

Fig. 3. Equivalent Tanner graph of parity-check matrix in equation (3)

An LDPC code is called *(ir)regular* if the total number of 1's in every column of the matrix is (not) the same as well as the total number of 1's in every row. Equivalently if all the check nodes have (not) the same number of connections to the variable nodes as well as the variable nodes to the check nodes.
LDPC codes are randomly constructed subject to these (ir)regularity constraints Gallager (1962).

3.0.1 Encoding

Given a binary source $X \in \{0,1\}^{1\times n}$ and an LDPC code $\mathbb{H} \in \{0,1\}^{k\times n}$ - $k < n$ - we multiply X with $\mathbb{H}$ and find the corresponding syndrome[1] $Z = \mathbb{H}^T X$, $Z \in \{0,1\}^{1\times n}$. Equivalently in the tanner graph we add all the variable nodes connected to the same check node. All operations are performed in modulo 2. The corresponding syndrome Z will be the compressed version of X.

3.0.2 Decoding

The decoder must estimate X, say $\hat{X}$, from Z, given $\mathbb{H}$ and $Y \in \{0,1\}^{1,n}$, known to be correlated to X. That is $Pr(X_i \neq Y_i) < 0.5$, $i = 1,2,...,n$.
As in Liveris et al. (2002), the conventional message passing[2] the LDPC decoder Casado et al. (2007); Leiner (2005); Shokrollahi (2003) is modified for the syndrome information to be taken into account. This yields to the following syndrome decoding algorithm:

- $\{x_i, y_i\} \in \{0,1\}$, $i = 1,2,...,n$ are the values in X and Y, respectively
- $s_i \in \{0,1\}$, $i = 1,2,...,k$ are the values in Z
- q_{ij} is the message passed variable node v_i to a check node c_j
- r_{ji} is the message passed from a check node c_j to a variable node v_i

[1] Actually the concept of compressing a binary source to its syndrome was first introduced by S. Pradhan et al. Pradhan & Ramch (1999). But that concept was rather an inspiration to constructive frameworks

[2] The message passing algorithm itself, even called *Belief propagation* in some literatures, is intensively studied in Bishop (2006); Kschischang et al. (2001)

- Q_i is the set of connected check nodes to the $i:th$ variable node.
- $Q_{i\setminus j}$ is the set of connected check nodes, excluding the $i:th$ check node, to the $j:th$ variable node.
- R_j is the set of connected variable nodes to the $j:th$ check node.
- $R_{j\setminus i}$ is the set of connected variable nodes, excluding the $i:th$ variable node, to the $j:th$ check node.

1. initialize; Prior Log Likelihood Ratios (LLR)s of x_i:

$$q_i^0 = \log \frac{Pr[x_i = 0|y_i]}{Pr[x_i = 1|y_i]} = (1 - 2y_i) \log \frac{1-p}{p} \tag{4}$$

2. Message (or LLR) sent from $i:th$ variable node to $j:th$ check node:

$$r_{ji} = 2 \operatorname{arctanh} \left((1 - 2s_j) \prod_{i' \in R_{j\setminus i}} \tanh \left(\frac{q_{i'j}}{2} \right) \right) \tag{5}$$

3. Message sent from $j:th$ check node to $i:th$ variable node:

$$q_{ij} = q_i^0 + \sum_{j' \in Q_{i\setminus j}} r_{j'i} \tag{6}$$

4. Hard decision:

$$\hat{x}_i = \begin{cases} 0, & \text{if } q_i^0 + \sum_{j=1}^{k} q_{ij} \geq 0 \\ 1, & \text{otherwise} \end{cases} \tag{7}$$

5. If $\mathbb{H}^T \hat{X} \doteq Z$, stop. Else goto 3

3.1 Potentials and limitations of LDPC codes

We carried out approximatively the same simulations as in Liveris et al. (2002). We compressed a (randomly generated) binary source X with codeword length $n = 16384$ bits to different compression ratios. The side information was generated with different crossover probabilities. The rates were increased until lossless compression was achieved. The results are presented in table 1

Although experimental results showed that LDPC-based compression of binary sources provides rates very close to Slepian-Wolf bound, it is important to mention a few caveats:

- No convergence at all is observed. Probable causes:
  - The **real** crossover probability is higher than required
  - The source signals are "far"[3] [4] from **random**

[3] Experimentally, source signals do not need to be *strictly* random for the decoder to work

[4] There are ongoing research especially dealing with non-random sources Garcia-Frias & Zhong (2003)

- The maximum number of iterations is reached[5], but convergence is not.
- Convergence is reached but the decoded codeword is the wrong one.

To generate parity-check matrix we used the implementation from Avudainayagam (2002). Additionally Liveris et al. showed that the performance achieved by LDPC codes is seen to be better than recently published results using Turbo codes. LDPC codes seem therefore to be more attractive solution to our "Wyner-Ziv Coding of Face Images" problem.

Crossover probabilities	0.01	0.05	0.1	0.2
Theoretical conditional entropies	0.169	0.286	0.469	0.722
Experimental conditional entropies	0.300	0.420	0.600	0.880
Experimental conditional entropies in Liveris et al. (2002)	0.310	0.435	0.630	0.890

Table 1. Lossless compression results using LDPC

4. Compression of face signature images using LDPC codes

Since the query images are not available when face signature images are compressed, a direct solution is to treat face signature images as binary sources X. Since X_i and X_j are two views from the same person and they are highly correlated with $Pr[X_i \neq X_j] = p < 0.5$. To allow the use of distributed coding to compress X the correlation between X_i and X_j can be modeled with a binary symmetric channel (BSC) with crossover probability p as shown in Fig 4 Following Liveris et al. (2002), LDPC codes $\mathbb{H}$ will be used to compress the binary sources with the query image as side information. That is, given a binary source X and a LDPC code $\mathbb{H}$, which is a $k \times n$ parity-check matrix, we multiply X with $\mathbb{H}$ and find the corresponding syndrome $Z = \mathbb{H}X$ with the length $(n-k)$. The LDPC decoder estimates the n-length sequence X from its (n-k) length syndrome X and the side information, query image Y (length n). The system is shown in Fig 5. The compression ratio achieved with this scheme is $\frac{n}{n-k}$. Figure 12 illustrates faces images and their respective syndrome face (signatures).

4.1 Binary coding of face signature images

The simplest way to transfer a grayscale face image into binary sequences is to employ the bit-plane coding to convert each gray level to its binary representation with prefixed resolution and then encode each bit-plane separately. Figure 6 shows the probability distributions for inter-face and intra-face variations over a small-scale face database (containing 40 subjects with 10 photos each). The low correlation is caused by the fact that the bit-plane coding is very sensitive to luminance changes, small changes in gray level can have a significant impact on the complexity of the bit planes. Obviously, it is inadequate to use LDPC to compress bit-planes, directly.

Since our preliminary goal is to investigate how to use LDPC to compress face signature images, here we just select a working binary coding scheme for our experiments. Our choice goes to Expectation Maximization (EM) Dempster et al. (1977) for segmentation Weiss (1997).

[5] For the sake of simplicity, we designed a decoding scheme that runs a predefined number of iterations. Intense studies for convergence rules are carried out in Casado et al. (2007); Daneshgaran et al. (2007); Hou et al. (2001); Matache et al. (2000)

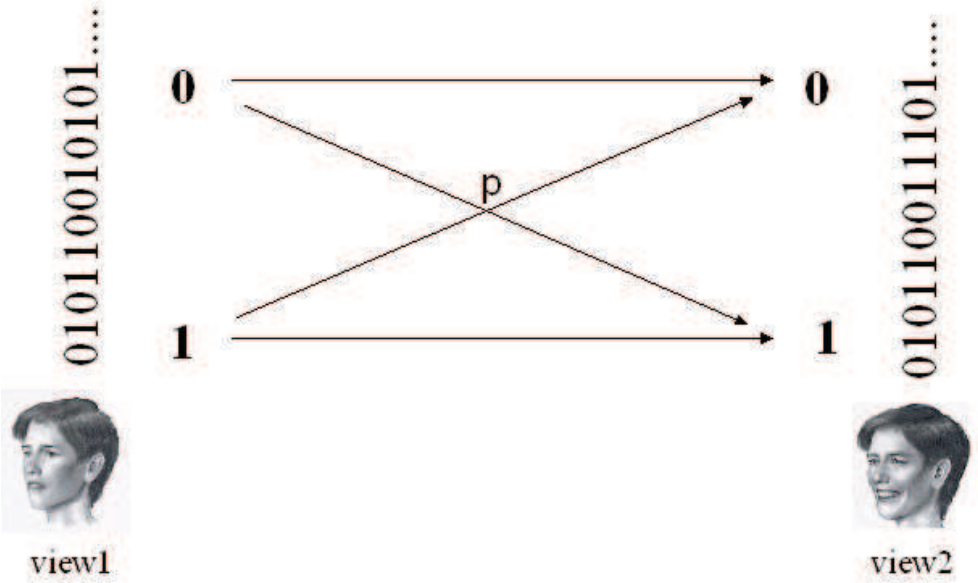

Fig. 4. Intra-face variation example

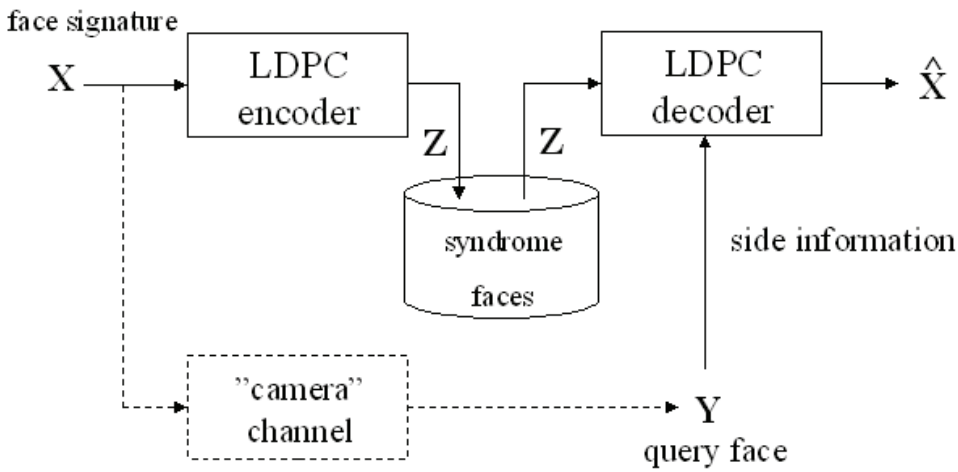

Fig. 5. Overview of system

EM attempts to assign objects a class but in a unsupervised way, tending to maximize the inter-class variation, while keeping the within-class semantic. Figure 7 shows segmentation results using EM. The results in figure 8 show that intra-face and inter-face variations are approximately $p_{intra} \sim N(0.21, 1.4310 \cdot 10^{-2})$ and $p_{inter} \sim N(0.35, 1.3510 \cdot 10^{-2})$, respectively. A certain improvement in correlation over intra-person faces is noticed (as shown in figure 6, yet not sufficient, because to employ LDPC the crossover probability has to be less than

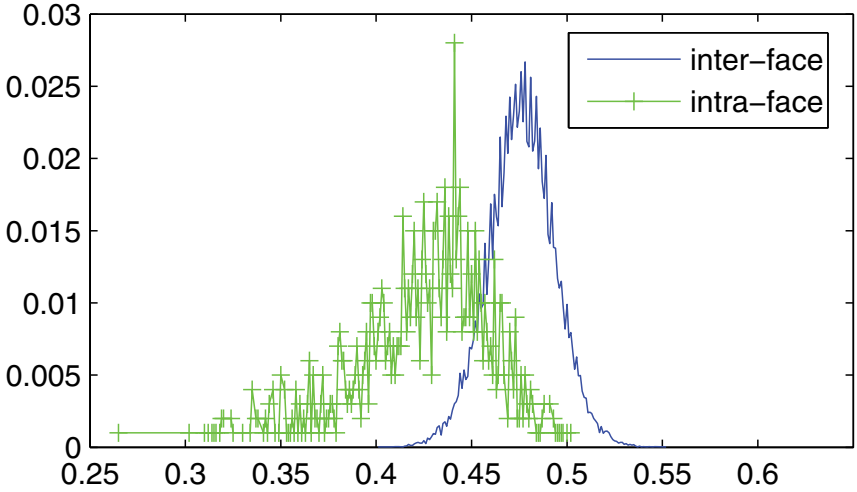

Fig. 6. Distribution of facial variation in grayscale

Fig. 7. Segmentation using EM

0.2 Liveris et al. (2002). We have to go for further processing. After carefully examining all face signatures one will see that intra face variations are mainly caused by the following factors: facial expressions, face poses, illuminance changes, camera factors etc. We have to do alignment between face signatures images.

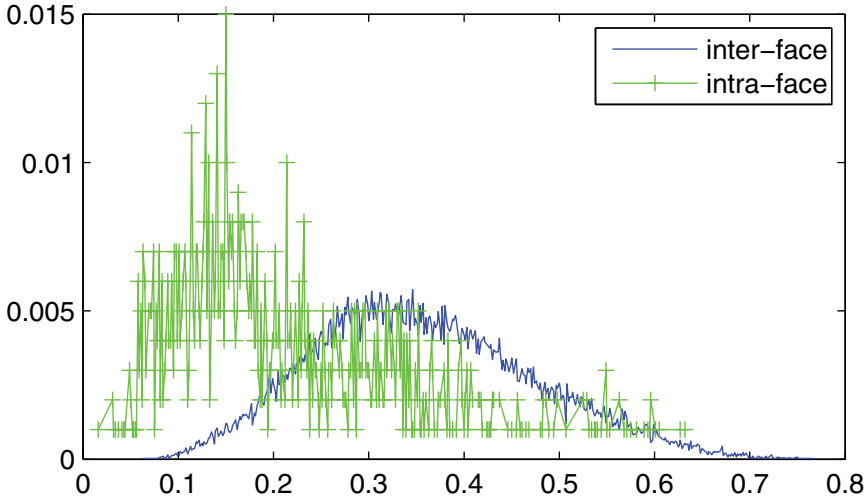

Fig. 8. Distribution of facial variation in binary before alignment

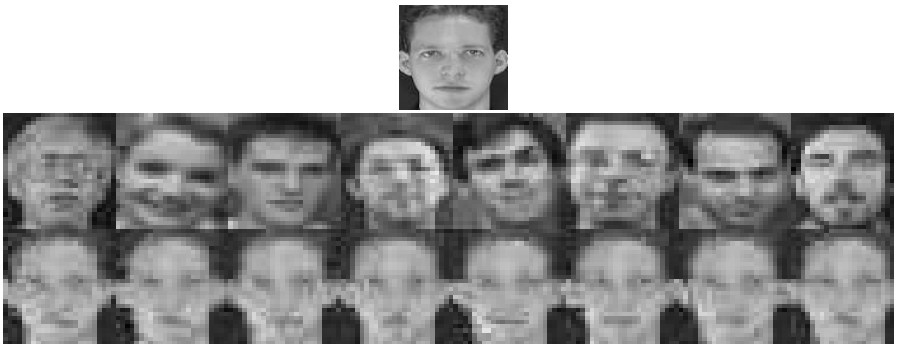

Fig. 9. Alignment result on some input images. First row: reference image, second row: input images, last row: result from alignment

4.2 Motion compensated alignment

To take away the displacement between two face signature images, a motion-compensated alignment technique is employed Li & Forchheimer (1995). The idea is first to select a natural face from the database as the reference face. To align a face signature image with respect to the reference image, the reference image has to be divided into blocks, $\{D_k\}$. Given a block D_k in the reference image, the aim is to find the corresponding block B_i in the face image by minimizing the distance $d(D_k, B_i)$. The distance is given by:

$$d(D_k, B_i) = |D_k - (a_0 + a_1 B_i)| \tag{8}$$

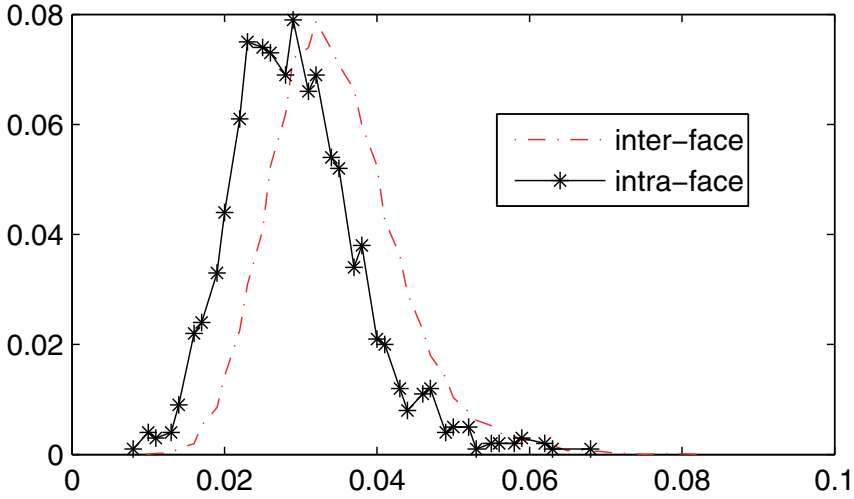

Fig. 10. Face distributions in binary vs. crossover probabilities, after alignment. dash-star line: intra-face. dash-point line: inter-face.

Coefficients a_0 and a_1 are defined by

$$a_0 = u_d - \frac{\sigma_d}{\sigma_b} u_b \tag{9}$$

$$a_1 = \frac{\sigma_d}{\sigma_b} \tag{10}$$

where u_b and σ_b are the mean value and variance of the block B_i and u_d and σ_d are the mean value and variance of the block D_k. The effect of the employed motion compensated alignment on face signature images is shown in figure 9. It is noted here that motion compensated alignment is performed before binary coding via the EM segmentation approach.
We carry out experiments over the face database and computed inter-face and intra-face variations as shown in figure 10. p_{inter} and p_{intra} can be modeled as normal distributions, $p_{inter} \sim N(0.035, 7.04 \cdot 10^{-5})$ and $p_{intra} \sim N(0.029, 7.17 \cdot 10^{-5})$. The small intra face variations make it perfect to use LDPC for compression of face signature images. More important, we know the compression bound: the theoretical limit for lossless compression of face signature images X is

$$nR \geq nH(X|Y) = nH(p) \tag{11}$$

$$H(p) = -plog_2(p) - (1-p)log_2(1-p) \tag{12}$$

Here $p = 0.0298$ corresponding to $H(X|Y) = 0.1934$, that is, a compression ratio of 5 can be achieved.

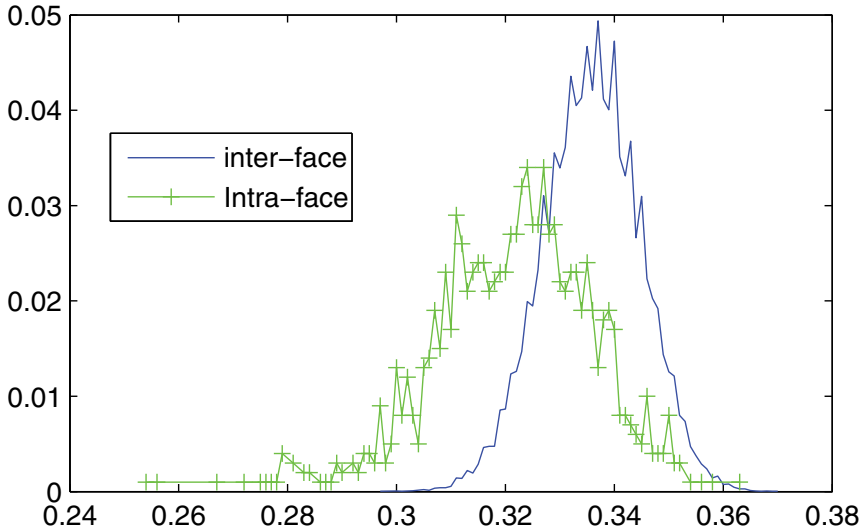

Fig. 11. Facial variation distributions in grayscale after alignment. Solid line: inter-face. dash-plus line: intra-face.

5. Retrieval metric

We introduce a similarity measure where the key is using syndrome decoding Liveris et al. (2002) and normalized Hamming distance.

At the retrieval phase, Given a query Y, we will process Y in the same manner as in the enrollment phase, first motion compensated alignment followed by binary coding. For each syndrome Z_i, $\hat{X}_i$ is estimated with respect to $\mathbb{H}$. This is equivalent to the Slepian-Wolf's insight of *sources coding with side information available only at the decoder* Slepian & Wolf (1973), where Y represents the side information and Z_i, the compressed correlated source. See figures 2 and 5. Normalized Hamming distance is performed between every $(Y, \hat{X}_i)$ pair. The normalized Hamming Distance is given by:

$$D_i = \frac{1}{n} \sum_{j=1}^{n} Y_j \oplus \hat{X}_{ij} \tag{13}$$

The templates are then ranked according to their distance to the query.

6. Preliminary results

In our experiment we use the ORL Database of Faces. In the database there are 10 different images of each of 40 distinct subjects, taken at different times, varying light conditions and facial expressions. For our purpose 5 randomly chosen images out 10 of each 40 subjects are used as training set and 5 for validation and test. It is also important to mention that the images were resized to 28×24 before further processing. The resizing parameters are

mainly motivated by the psychological assumption made in Torralba et al. (2008) and our own research on face recognition Le (2008).
Three LDPC codes are employed corresponding to different compression ratios, $R = 0.31$, $R = 0.50$ and $R = 0.76$. Recall that we found experimentally that $p_{inter} \sim N(0.035, 7.04{\cdot}10^{-5})$ and $p_{intra} \sim N(0.029, 7.17{\cdot}10^{-5})$. Theoretically a compression rate $R = 0.1934$ is thus expected Slepian & Wolf (1973). This makes great sense since a 28×24 grayscale image, when transformed to binary, requires 672 bits to be stored. With The theoretical compression rate 537 bits saved. Using an LDPC code with rate $R = 0.31$, we were able to save 464 bits per template, while achieving comparable results the with the scheme with no compression. Thus for 200 templates, we save 92800 bits! Figure 13 reports performance results, where the retrieval efficiency is plotted against the number of outputs. the line specifications in figure 13 are denoted as follow:

- *Solid-asterisk*: Alignment followed by bit-plane-wise binary representation
- *Solid*: Alignment followed by binarization. No compression
- *Dash-dot*: proposed scheme with rate 0.3
- *Solid-upward triangle*: proposed scheme with rate 0.5
- *Solid-downward triangle*: proposed scheme with rate 0.7

Fig. 12. Face images and respective resulting syndrome

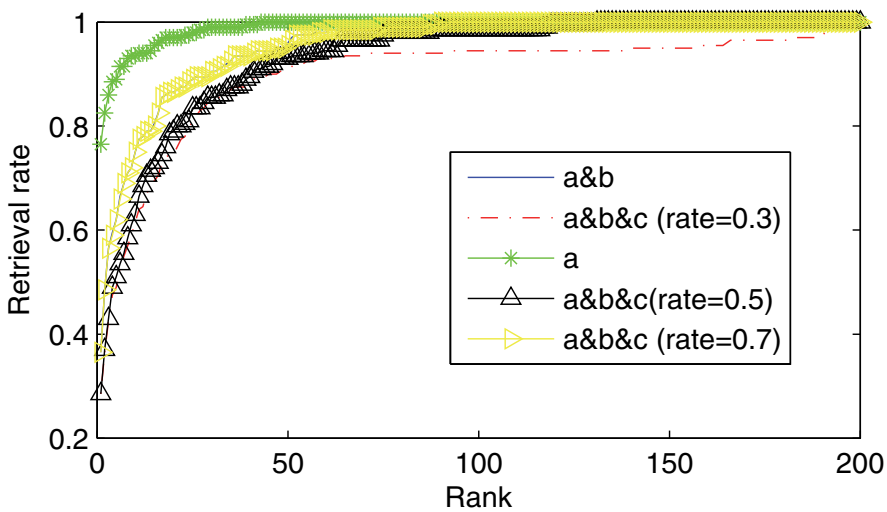

Fig. 13. Recognition Rate vs. Rank

7. Concluding remarks

Wyner-Ziv coding is radically different from conventional image coding. It gives a totally new coding paradigm. Most research efforts are devoted to how image and video compression can be done under the new paradigm. This paper is the debut of our effort to investigate how Wyner-Ziv coding can be used for large-scale image retrieval problem.

Image coding and image retrieval have been conventionally two different disciplines. In this paper image retrieval is considered as an image-coding problem. The powerful rate-distortion theory can be directly used to characterize the tradeoff between retrieval quality and retrieval speed through the crossover probability p.

Wyner-Ziv coding has a great potential to improve the efficiency of large-scale image retrieval. Under the Wyner-Ziv coding framework the query information provided by a huge number of web users can be utilized to reduce the storage and transmission of face images. Considering that Google receives hundreds of millions of queries per day and they use a million servers to run their search service, it is a big impact to our environment if consumption of storage and transmission can be reduced 90% by adopting Wyner-Ziv coding.

The results we reported here are very preliminary. We focus ourselves on how to use LDPC codes to compress binary coding of face signature images X to reach the Slepian-Wolf bound $H(X|Y)$. We haven't addressed at all how to quantize X to achieve Wyner-Ziv coding, $H(Q(X)|Y)$. To focus on LDPC coding of X, we just use an EM-strategy to do binary coding of face signature images and use it for our benchmark. We already see that the binary coding results in a significant loss in the quality of image retrieval. To build an efficient whole system, the study of $Q(X)$ has to be carried out. In addition, motion compensated alignment plays a very important role and has a big impact on both compression efficiency and retrieval quality. How to achieve an optimal alignment is an important topic for future research.

8. References

Avudainayagam, A. (2002). Ldpc toolkit for matlab.
URL: *http://arun-10.tripod.com/ldpc/ldpc.htm*

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*, Springer, chapter 8, pp. 360–418. URL: *http://www.amazon.ca/exec/obidos/redirect?tag=citeulike09-20&path=ASIN/0387310738*

Casado, A. I. V., Griot, M. & Wesel, R. D. (2007). Informed dynamic scheduling for belief-propagation decoding of ldpc codes, *Communications, 2007. ICC '07. IEEE International Conference on*, pp. 932–937.
URL: *http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4288829*

Daneshgaran, F., Laddomada, M. & Mondin, M. (2007). Ldpc-based iterative algorithm for compressionof correlated sources at rates approaching the slepian-wolf bound, *Technical report*.

Dempster, A. P., Laird, N. M. & Rdin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm, *Journal of the Royal Statistical Society, Series B* 39: 1–38.

Gallager, R. (1962). Low-density parity-check codes, *Information Theory, IEEE Transactions on* 8(1): 21–28.
URL: *http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1057683*

Garcia-Frias, J. & Zhong, W. (2003). Ldpc codes for compression of multi-terminal sources with hidden markov correlation, 7(3): 115–117.

Hou, J., Siegel, P. H. & Milstein, L. B. (2001). Performance analysis and code optimization of low density parity-check codes on rayleigh fading channels, *Selected Areas in Communications, IEEE Journal on* 19(5): 924–934.
URL: *http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=924876*

Kouma, J.-P. & Li, H. (2009). Large-scale face images retrieval: A distribution coding approach, *International Conference on Ultra Modern Telecommunications (ICUMT 2009)*, St. Petersburg, Russia.

Kschischang, F. R., Frey, B. J. & Loeliger, H. A. (2001). Factor graphs and the sum-product algorithm, 47(2): 498–519.

Le, H. S. (2008). *Face Recognition: A Single View Based HMM Approach*, PhD thesis, Umeå University, Faculty of Science and Technology, Applied Physics and Electronics.
URL: *urn:nbn:se:umu:diva-1485*

Leiner, B. M. (2005). Ldpc codes Ű a brief tutorial, *Technical report*.

Li, H. & Forchheimer, R. (1995). A transformed block-based motion compensation technique, *IEEE Transactions on Communications*, Vol. 43, pp. 1673–1676.

Liveris, A. D., Xiong, Z. & Georghiades, C. N. (2002). Compression of binary sources with side information at the decoder using ldpc codes, *Communications Letters, IEEE* 6(10): 440–442.
URL: *http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1042242*

Matache, A., Dolinar, S. & Pollara, F. (2000). Stopping rules for turbo decoders, *JPL TMO Progress Report 42-142, Aug. 15, 2000. 103*, pp. 42–142.

Phillips, J. P., Scruggs, T. W., OŠtoole, A. J., Flynn, P. J., Bowyer, K. W., Schott, C. L. & Sharpe, M. (2007). Frvt 2006 and ice 2006 large-scale results, *Technical report*, National Institute of Standards and Technology.

Pradhan, S. S. & Ramch, K. (1999). Distributed source coding using syndromes (discus): Design and construction, *IEEE Trans. Inform. Theory* 49: 626–643.

Shakhnarovich, G. & Moghaddam, B. (2004). Face recognition in subspaces, *in: S.Z. Li, A.K. Jain (Eds.), Handbook of Face Recognition*, Springer, pp. 141–168.

Shokrollahi, A. (2003). Ldpc codes: An introduction, *Digital Fountain, Inc., Tech. Rep* p. 2.

Slepian, D. & Wolf, J. K. (1973). Noiseless coding of correlated information sources, *Information Theory, IEEE Transactions on* 19(4): 471–480.
URL: *http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1055037*

Tanner, R. (1981). A recursive approach to low complexity codes, *Information Theory, IEEE Transactions on* 27(5): 533–547.
URL: *http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1056404*

Torralba, A., Fergus, R. & Freeman, W. T. (2008). 80 million tiny images: A large data set for nonparametric object and scene recognition, *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 30(11): 1958–1970.
URL: *http://dx.doi.org/10.1109/TPAMI.2008.128*

Tuncel, E., Koulgi, P. & Rose, K. (2004). Rate-distortion approach to databases: storage and content-based retrieval, *Information Theory, IEEE Transactions on* 50(6): 953–967.

Varodayan, D., Aaron, A. & Girod, B. (2005). Rate-adaptive distributed source coding using low-density parity check codes, *In: Proc. Asilomar Conf. on Signals, Syst., Comput*.

Weber, R. & Blott, S. (1997). An approximation-based data structure for similarity search.

Weiss, Y. (1997). Motion segmentation using em - a short tutorial.

Wyner, A. & Ziv, J. (1976). The rate-distortion function for source coding with side information at the decoder, *IEEE Transactions on Information Theory* 22(1): 1–10.
URL: *http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1055508*

Xiong, Z., Liveris, A., Cheng, S. & Liu, Z. (2003). Nested quantization and slepian-wolf coding: a wyner-ziv coding paradigm for i.i.d. sources, *Statistical Signal Processing, 2003 IEEE Workshop on* pp. 399–402.

Zamir, R., Shamai, S. & Erez, U. (n.d.). Nested linear/lattice codes for structured.

New Approaches to Characterization and Recognition of Faces

Edited by Dr. Peter Corcoran

ISBN 978-953-307-515-0

Hard cover, 252 pages

Publisher InTech

Published online 01, August, 2011

Published in print edition August, 2011

As a baby, one of our earliest stimuli is that of human faces. We rapidly learn to identify, characterize and eventually distinguish those who are near and dear to us. We accept face recognition later as an everyday ability. We realize the complexity of the underlying problem only when we attempt to duplicate this skill in a computer vision system. This book is arranged around a number of clustered themes covering different aspects of face recognition. The first section presents an architecture for face recognition based on Hidden Markov Models; it is followed by an article on coding methods. The next section is devoted to 3D methods of face recognition and is followed by a section covering various aspects and techniques in video. Next short section is devoted to the characterization and detection of features in faces. Finally, you can find an article on the human perception of faces and how different neurological or psychological disorders can affect this.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Jean-Paul Kouma and Haibo Li (2011). Large-Scale Face Image Retrieval: A Wyner-Ziv Coding Approach, New Approaches to Characterization and Recognition of Faces, Dr. Peter Corcoran (Ed.), ISBN: 978-953-307-515-0, InTech, Available from: http://www.intechopen.com/books/new-approaches-to-characterization-and-recognition-of-faces/large-scale-face-image-retrieval-a-wyner-ziv-coding-approach

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821