On Compression Encrypted Data – part 2

Prof. Ja-Ling Wu

The Graduate Institute of Networking and Multimedia

National Taiwan University

Brief Summary of Information-theoretic Prescription

- At a functional level, Wyner-Ziv compression of a source X in the presence of a correlated source Y at the decoder proceeds according to a two-step design process:

  (1) a rate-distortion source codebook is designed for quantizing X into its quantized description U

  (2) this source codebook is then "randomly" partitioned into bins or cosets, where the number of cosets depends on the correlation structure between X and Y and the targeted encoding bit rate.

Encoding consists of

(a) quantizing X to U using the source rate-distortion codebook

(b) syndrome-encoding U , i.e., specifying the index of the coset into which U falls.

Decoding consists of

(a) decoding U from its associated coset by using Y to find the "closest" entry in the list of codewords contained in the coset representation for X.

(b) Finding the best (MMSE) estimate for X using both U and Y, which are akin to "noisy observations" of X.

From information theory to video codec practice

- First while the theory assume knowledge of the correlation structure between X and Y, in practice, this structure is not known precisely and needs to be estimated. Þ

one needs a classification module aiming to classify the statistical nature of the correlated side-information into a prescribed set of classes (from zero to maximum correlation).

Each video block will be accordingly classified into a prescribed discrete set of correlation noise classes, which in turn will influence the syndrome trellis (convolutional) channel code choice.

- Secondly, while information-theory dictates for the optimal decorrelating transform in the side-information coding case to be the KL transform of the innovations noise vector process, in practical version, one will invoke the ubiquitous DCT.

- Thirdly, the rate-distortion source codebook prescribed in theory will be approximated with a scalar quantizer. The "random" partitioning operation dictated by information-theory will be approximated by trellis-based coset codes. Thus, while theory calls for a random codebook to be partitioned into random cosets, the practical construction will be based on a scalar quantizer lattice partitioned into trellis-coded cosets.

- The maximum-likelihood syndrome decoding will be done through a Viterbi decoder. Further, unlike the classical Wyner-Ziv coding case, where there is a single known side-information Y, in video case, there will be several side-information candidates Yi corresponding to various motion predictor “matched” to the trellis code choices.
- Finally, the syndrome-decoded output and the side-information source will be optimally “fused” linearly to form an MMSE estimate of the source reconstruction.

Encoding

- The video frame to be encoded is divided into non-overlapping spatial blocks (we choose blocks of size 16x16 or 8x8).There are four main aspects of the encoding process:
  1. classification
  2. Decorrelating transform
  3. Quantization
  4. Syndrome Encoding

- Real video sequences exhibit spatial- temporal correlation noise structures whose statistics are highly spatially varying.
- With the same sequence, some spatial blocks that are part of the scene background do not change much with time. That is , they are highly correlated with their temporal predictors (small N). On the other hand, there are some blocks that are a part of a scene change or occlusion. Such blocks have little correlation with the previous frame (large N). Thus, within the same frame, different blocks exhibits different degrees of correlation with the previous frame. This motivates the modeling of the video source as a composite or a mixture source where the different components of the mixture correspond to sources with different correlation (innovation) noise.

- The classification step therefore aims at classifying these correlation noise structure at a fine-grained block level into a prescribed set of pre-designed classes from zero to maximum correlation.
- These classes correspond to syndrome channel code choices of varying error-correcting capabilities. That is, once a block is classified, it will be tagged with the appropriate channel code for side information coding of that block.

- The “square error difference” between the block to be encoded and the co-located block in the previous frame is used to model the correlation noise N.
- This squared error difference is thresholded and the block is classified into one of many classes enabling the use of the appropriately matched channel code.
- At one extreme is the SKIP mode, where the frame difference is so small that the block is not encoded at all , and at the other extreme is the INTRA mode, where the frame difference is very large suggesting poor correlation, so that intra-coding is appropriate. There are various different syndrome coding modes in between these two extremes. After classification, the classification label is included as part of the header information for use by the decoder.

2. Decorrelating Transform:

We apply a DCT on the source vector to approximate the KL transform of the correlation noise innovation process between the source vector and its side-information counterpart.

This is akin (similar) to the classical operation of a DCT on the motion-compensated Displace Frame Difference (DFD) process in standard coders.

3. Quantization

This step involves scalar quantization of the DCT coefficients of the source vector (akin to the source rate-distortion codebook in Wynes-Ziv theory). The information-theoretic bit allocation for independent correlated vectors has to drive the design choice of the quantizer resolution (i.e., quantization factor) for each DCT coefficient according to the bit allocation algorithm applied to the innovations (DFD) coefficients.

4. Syndrome Encoding:

This step represents the constructive counterpart to the random coset partitioning operation, and involves the syndrome encoding of the quantized block codewords.

This step partitions the codeword space of quantized codewords into cosets. Each containing a collection (or uncertainty list) of codewords. A syndrome label is associated with each such coset.

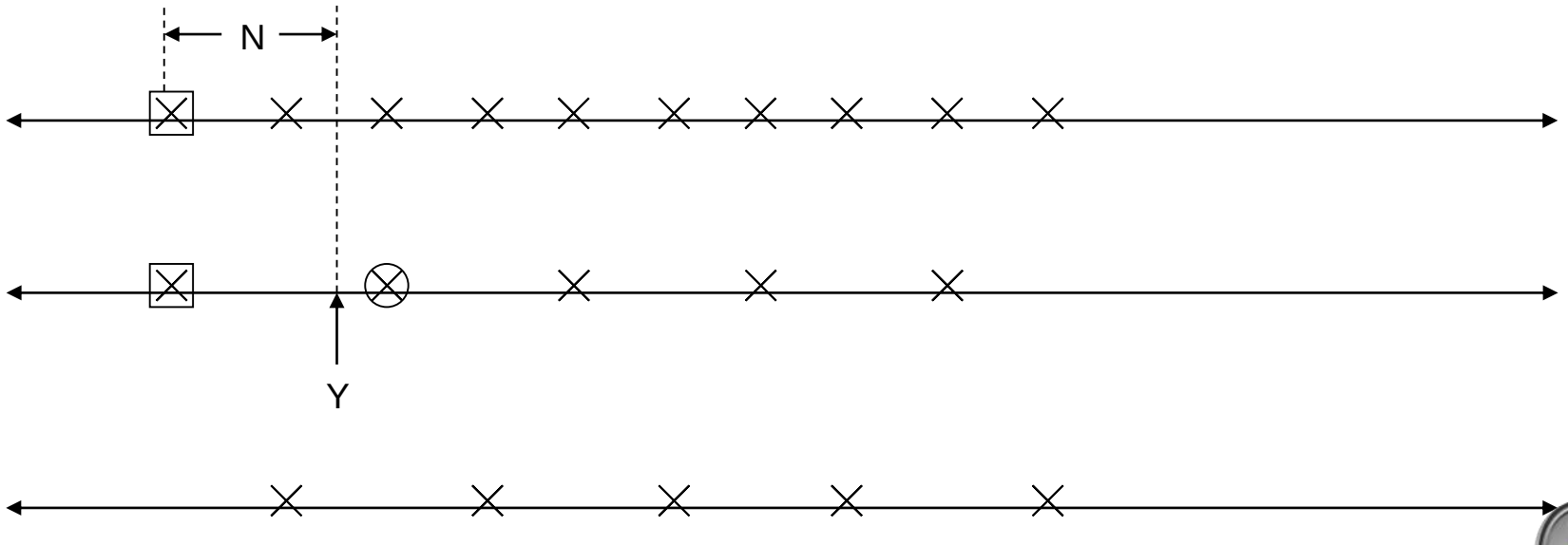
- Compression is obtained here since instead of transmitting each individual codeword index, only the index of the set containing the codeword index is transmitted.
- If syndrome encoding is done using a code matched to the correlation noise (eg. a Gaussian distribution), one can then expect to achieve performance comparable with predictive coding.

The implementation of the above two steps:

(a) Base Scalar Quantization:

The transform coefficients (which are real numbers when DCT is used), are first quantized before encoding. When the total number of coset partitions is fixed, the choice of the quantization step size is limited by the statistics of N.

- If a very fine step size is chosen to encode X, then there can be decoding errors, since the codewords will be too “close” so that the side-information Y cannot disambiguate them correctly (such as the case in topline of the following figure).



- The topline shows the quantized codeword set for X, and the two bottom lines show the partition of the space of quantized codewords (two partitions imply a rate of 1 bit).
- The rectangular box shows the observed codeword which lies in the first partition.
- Since the magnitude of N is large than the quantization step size, the decoder can use the side information Y to decode (reconstruct) the better approximated (circled) codeword.

Thus, each of the elements of X is quantized with a step size
Proportional to the standard deviation of the corresponding element in N

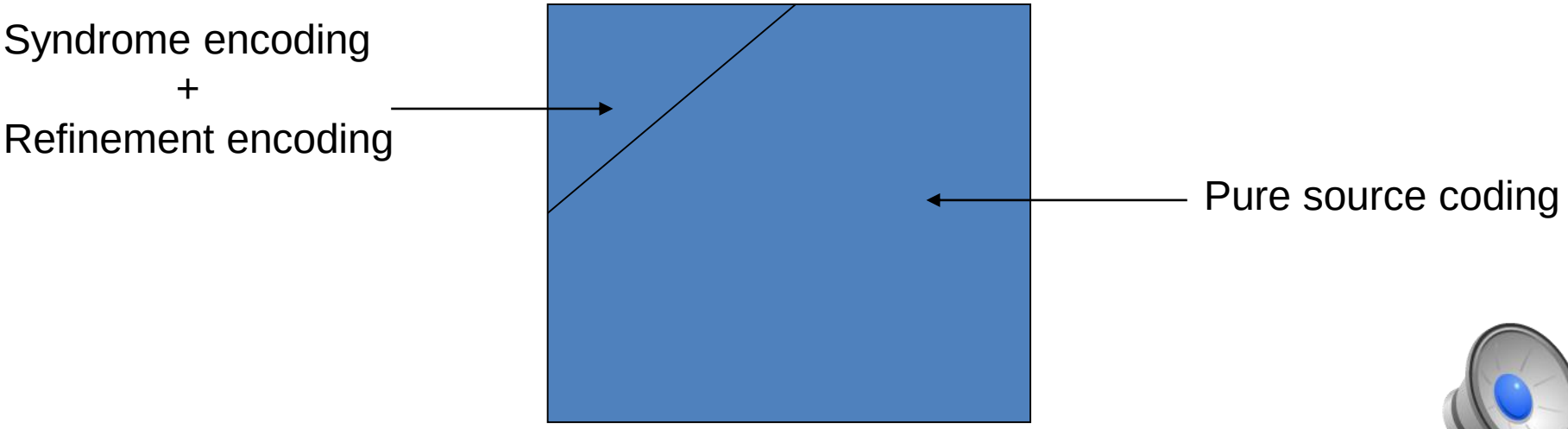
(b) Zig-Zag scan :

- The quantized coefficients are arranged in a 1-dimensional order (size 256 or 64) by doing a Zig-Zag scan on the 2-D block (size 16x16 or 8x8).
- It has been observed in general that arranging 2-D coefficients in a 1-D order using a Zig-Zag scan pattern tends to organize them in decreasing order of energies (importance).

(c) Syndrome Encoding :

- Now the space of quantized codewords which has been appropriately generated using the statistics of N is partitioned using a Euclidean space trellis channel code (i.e., an Error Control Code).
- The generation of the coset index (syndrome) associated with each codeword can be accomplished in a Computational efficient manner through a simple convolution operation (multiplication) between the quantized codeword and the parity check matrix of the trellis code.

- Further, in each block of each class, only the first fraction of the (zig-zag) scanned coefficients are syndrome encoded. The remaining coefficients are purely intra-coded.
- This is based on the observation that for typical natural images, the first few transform coefficients contain most of the information about the block. We thus expect them to exhibit significant correlation with the corresponding predictor blocks. (Empirically, for both 8x8 and 16x16 blocks, only about 20% of the coefficients need to be syndrome encoded).



(d) ”Pure” Source coding :

- The remaining coefficients which comprise about 80% of the total coefficients are intra-coded in the conventional way.
- The coefficients are first quantized, then Zig-Zag scanned and finally are entropy coded using run-length Huffman coding.

(e) Refinement Quantization :

- A target reconstruction quality (e.g. PSNR) corresponds to a particular quantization step size. (Higher desired quality corresponds to a finer quantization step size and lower corresponds to a coarser one).
- The coefficients that are purely intra-coded are quantized with a step size corresponding to the target quality.

- For the coefficients that are syndrome encoded, the choice of the base quantization step size is limited by N. This is done so as to minimize the probability of decoding error of the trellis codes.
- Hence, assuming that the base quantization interval can be conveyed correctly with high fidelity to the decoder, we refine it further to the target quantization step size.

- A straightforward implementation of the refinement operation is just a progressive sub-dividing of the base quantization interval into intervals of size equal to the target quantization step size.
- The index of the refinement interval inside the base interval is transmitted to the decoder.

- Coefficients with significant correlation (small N) have a fine base quantization step size and hence the refinement stage results in fewer refinement intervals and hence fewer refinement bits.
- Thus, the bit allocation is proportion to correlation – fewer bits are required if correlation is high and more bits when the correlation is weak.

(f) Cyclic Redundancy Check (CRC) :

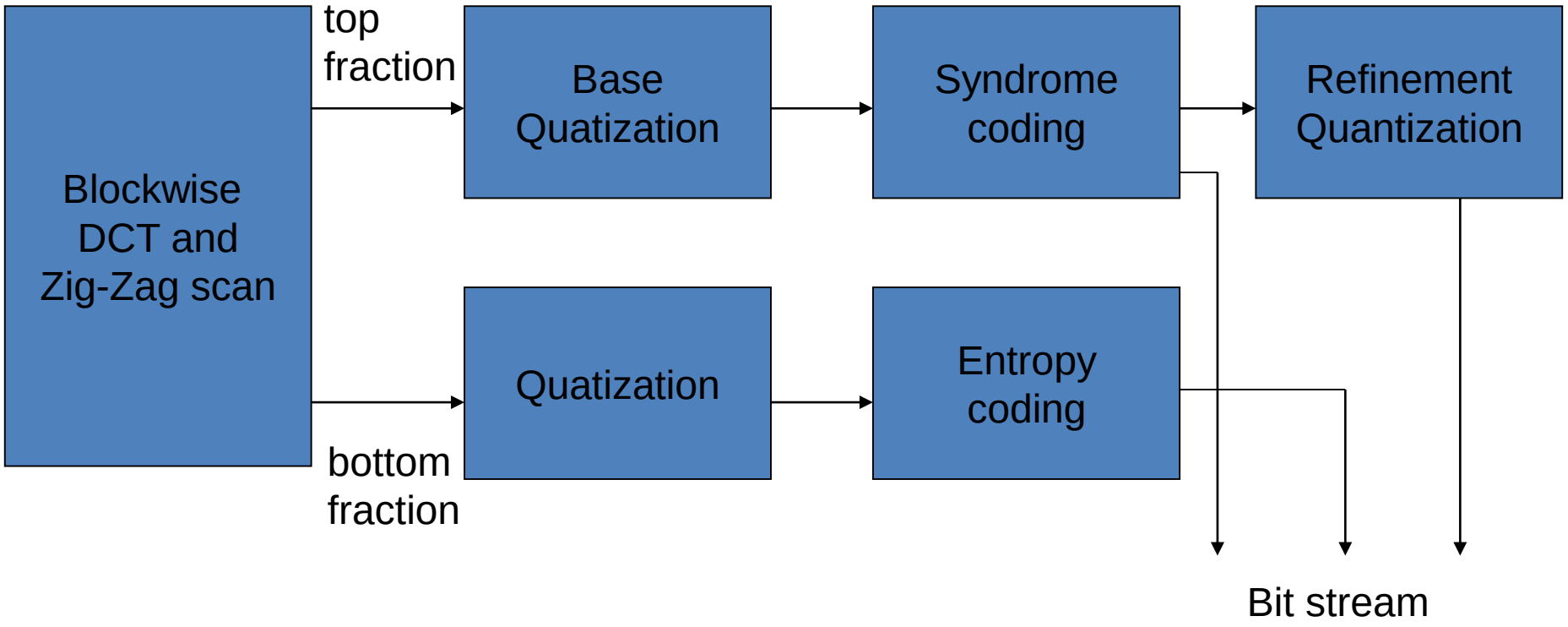
- Note that at the encoder, side-information encoding is done in principle with respect to the statistics of the motion compensated prediction error between the block x that is to be encoded and the “best” predictor Y for this block in the frame memory.
- At the decoder, all that is available in the frame memory. The decoder does not know the “best” predictor for the block X.

- The encoder transmits not only the syndrome for the side-information encoded coefficients but also a CRC check (of sufficient strength) of the quantized sequence. This CRC check serves as a “signature” of the quantized codeword sequence.
- In contrast to the conventional paradigm, it is the decoder’s task to do motion search here, and it searches over the space of candidate predictors one-by-one to decode a sequence from the set labeled by the syndrome.

- When the decoded sequence matches the CRC check, decoding is declared to be successful. Note that the CRC needs to be sufficiently strong so as to act as a reliable signature for the codeword sequence.

Syndrome bits	CRC bits	Refinement bits	Pure Source coded bits

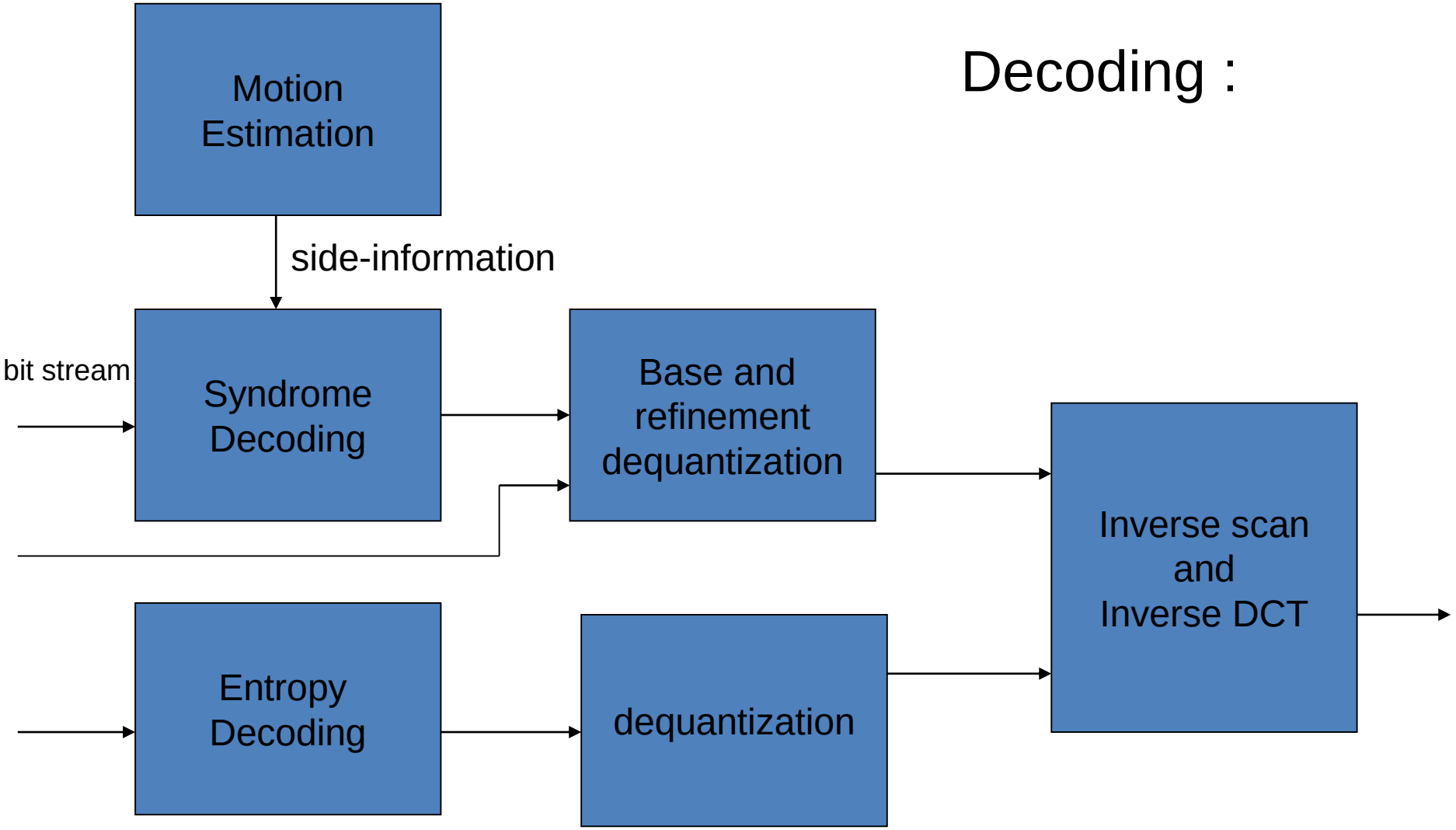
Bit stream syntax associated with a block.



Functional block diagram of the encoder.

The main complexity in the encoding process incurred in step 2 (complexity of the DCT) and in the entropy coding stage.
=> The encoding complexity is of the order of standard intra-coding complexity.

Decoding :



Functional block diagram of the decoder.

1.Generation of Side Information (Motion Search):

- The decoder does motion search to generate candidate predicators to decode the sequence of quantized codewords from the set indicated by the received syndrome. For example, half pixel motion search is used to obtain various candidate predictors as is also done at the encoding side in the standard video algorithms.
- The framework is very general so as to accommodate any motion estimation procedures. The choice of a more sophisticated algorithm can only serve to enhance the performance of this coding scheme.

2.Syndrome Decoding:

- Each of the candidate predictors generated by the motion search module is used to decode a sequence of quantized codewords from the set indicated by the syndrome. Since trellis codes are used, this decoding can be accomplished using the Viterbi algorithm.
- Here the set of all sequences labeled by the received syndrome is represented on a trellis. The Viterbi algorithm is then used to identify the sequence in this set that is “closet” to the candidate predictor.

- If the decoded sequence matches the CRC check, then the decoding is declared to be successful. Else using the motion search module, the next candidate predictor is obtained and the whole procedure repeated.

3. Estimation and Reconstruction:

- Once the quantized codeword sequence is recovered , it is used along with the predictor to obtain the best reconstruction of the source.
- Any sophisticated signal processing algorithm(e.g. spatial-temporal interpolation) or post processing mechanism can be deployed in this framework and they serve to improve the overall performance.

4.Pure Source Decoding:

- For the coefficients (about 80%) that have been intra-coded, the decoding action consists of entropy decoding followed by dequantization.

5.Inverse Zig-Zag Scan:

- Once all the transform coefficients have been dequantized, the zig-zag scan operation carried out at the encoder is inverted to obtain a 2-D block of reconstructed coefficients.

6.Inverse Transform:

- The transform coefficients are then inverted using the IDCT to give the reconstructed pixels.

- An Introduction to Coset Decomposition

Corset Decomposition

Consider a finite group G, let H $\{h_1, h_2, \cdots, h_n\}$ be

a subgroup of G,

Perform

h_1	h_2	h_3	$\cdots$	h_n	
$g_1 * h_1$	$g_1 * h_2$	$g_1 * h_3$	$\cdots$	$g_1 * h_n$	$g_1 \in G/H$
$g_2 * h_1$	$g_2 * h_2$	$g_2 * h_3$	$\cdots$	$g_2 * h_n$	g_2 Î $G/H/\{2nd \text{ row}\}$
			⋮		
$g_m * h_1$	$g_m * h_2$	$g_m * h_3$	$\cdots$	$g_m * h_n$	

Each row is called a left coset of H in G.

: consists of the elements generated by performing binary operation * on H with the same element g_k

$$g_k \in G/H/\{2nd \text{ row}\}/\cdots/\{(k-1)th \text{ row}\}$$

Each element in G appears once and only once in a coset decomposition of G

$$|H| \bullet (\#\text{ of coset G w.r.t H}) = G$$

remark : Coset decomposition forms a partition of G.

In ECC applications:

$h_i \rightarrow$ codewords

$g_i \rightarrow$ error patterns

Each column is a decoding sphere

Since each coset corresponds to a specific

error pattern, each row gives a certain syndrome in decoding.

Where syndrome denotes the product of error pattern vector with the parity check matrix.

All vectors in the same coset (the same row of the above table) have the same syndrome unique to that coset.

Given a received codeword $\underline{u}$, compute the syndrome and lookup its associated coset leader (error pattern).

Subtract the coset leader from $\underline{u}$ to get the decoded result.

For a source X, if we can partition
it into n disjoint subsets according
to some specific relation, and within
each subset find the representing codeword
based on certain rules, such as VQ. Associating
these representing codewords with the codewords
of an ECC.
Based on certain similarity of distance measure,
we can partition each subset further into different
M levels according to the distance to the center of
the subset, and associate each level with a codeword so that
the Hamming distance to the center of the subset is proportio
to the prescribed distance.

in other words, the so-obtained codewords of a subset behaves like the error-incurred received codewords of the representing codeword of the subset.

The number of subsets n determines the quality of this construction.

The number of levels M corresponds to the number of syndromes.

The error-tolerance (robustness) of this coding system depends on the chosen ECC.