Selective and Scalable Encryption of Enhancement Layers for Dyadic Scalable H.264/AVC by Scrambling of Scan Patterns

Zafar SHAHID, Marc CHAUMONT and William PUECH
LIRMM, UMR CNRS 5506, Université Montpellier II
zafar.shahid@lirmm.fr, marc.chaumont@lirmm.fr, william.puech@lirmm.fr

Problem Statement

THE PROBLEM:

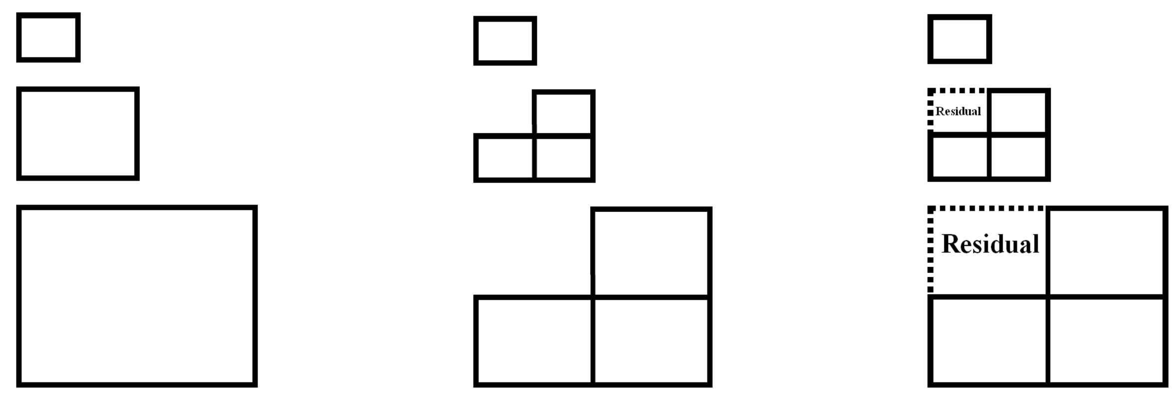
- To devise a method for selective and scalable encryption (SSE) of *intra* dyadic scalable coding framework based on wavelet/subband (DWTSB) for H.264/AVC.

OUR APPROACH:

- Separate secret key has been used for each layer. The key has been embedded in LL subband using fragile watermarking.
- Scalable encryption has been introduced, in accordance with the energy distribution of quantized transformed coefficients (QTCs) in the specific subband.
- SVC offers a new profile for *intra* video coding named Profile B Intra only.

H.264/AVC

H.264/AVC is the state of the art video codec and performs better than previous standards.
Scalable video coding (SVC) is based on H.264/AVC.
Different SVC approaches have been used in literature: a) Pyramid coding used in JSVM, b) Wavelet subband coding used in JPEG2000, c) DWTSB for dyadic scalable intra frame of JSVM.



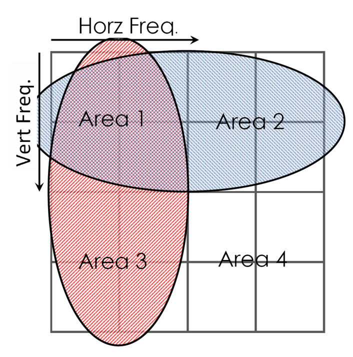
(a) (b) (c)

In this work, we have used DWTSB approach.

The Proposed Method

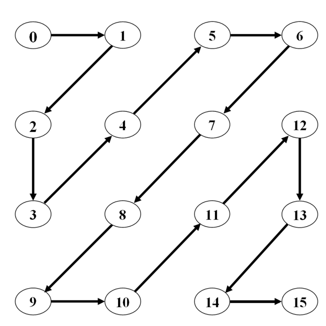
Scanning is a bijective function from $P_{m \times n}$ to Q_{mn}. Let the QTCs be a 2-dimensional array given as:
$P_{m \times n} = \{p(i,j) : 1 \leq i \leq m, i \leq j \leq n\}$.
After scanning the 2-dimensional array, we get a set:
$Q_{mn} = \{1, ..., mn\}$.
Every subband has different energy distribution. Hence, separate scan should be used for each scan. For example, in LL subband:

Energy distribution in QTCs of LL subband.



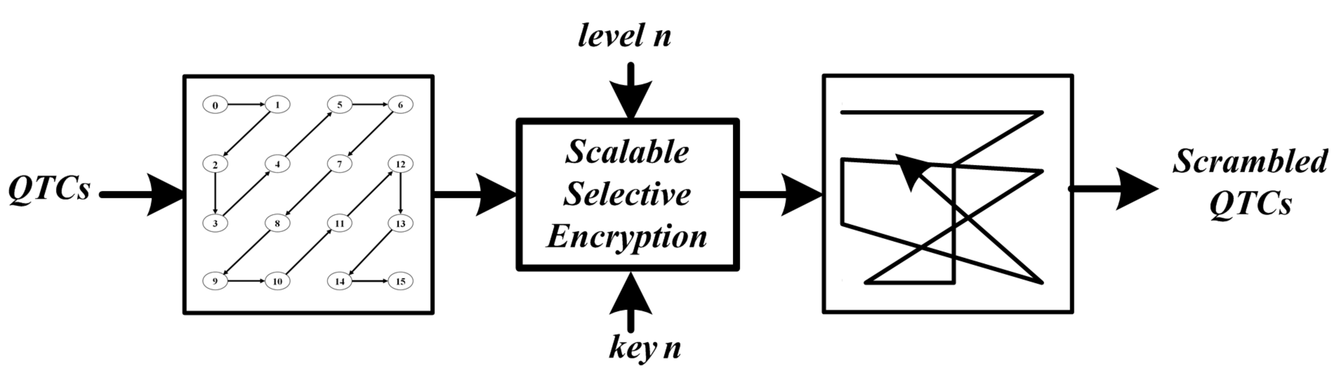
(a)

Scan used for this frequency distribution.

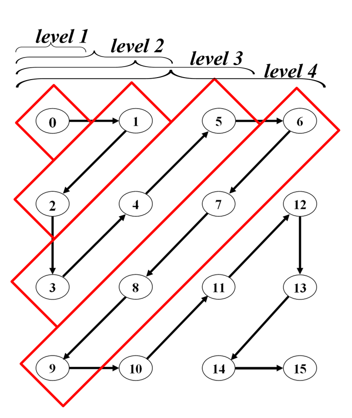


(b)

Encryption is performed by scrambling of scan pattern of transformed coefficients as shown below.

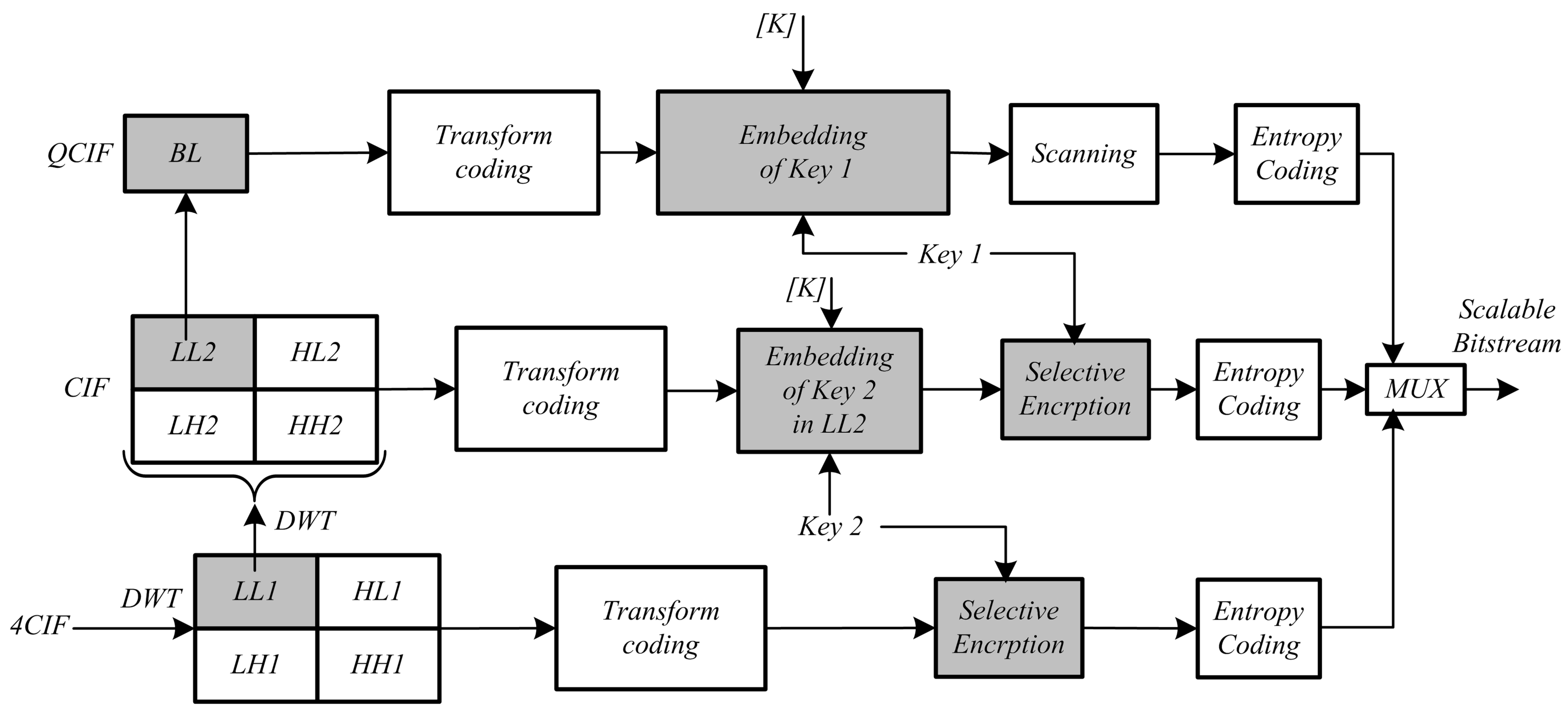


Scalable SE has been used in which we have a compromise among required computational power, bitrate and PSNR of encrypted bitstream.
In each encryption level of SSE, we leave a subgroup of QTCs unscrambled, which depends on the DWTSB subband. Since in each subband QTCs have different energy distribution.
Subgroup of QTCs which is not scrambled in each encryption level is shown in figure below:



Selective and Scalable Encryption

Block diagram showing SSE of ELs in scalable bitstream is:



Experimental Results

- We have used 'JVT-W097' which is referenced H.264 JSVM 8.9 with wavelet framework integrated.
- Daubechies 9/7 wavelet filter set has been used to transform the image to the wavelet subbands.
- CAVLC entropy coding mode has been used.
  - Analysis of trade off among bitrate, PSNR and SE *level* for *city* over whole range of QP values.

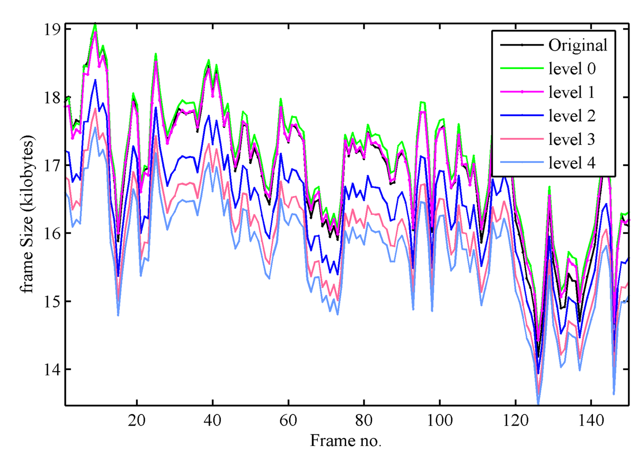
SE level	12.0 (kbps) (dB)	18.0 (kbps) (dB)	24.0 (kbps) (dB)	30.0 (kbps) (dB)	36.0 (kbps) (dB)	42.0 (kbps) (dB)
Orig	5337 48.71	3386 44.20	1999 39.60	1088 35.19	543 31.17	250 27.60
level 0	5327 19.20	3418 19.75	2045 20.61	1144 20.12	595 18.79	288 18.20
level 1	5285 18.48	3394 19.47	2038 20.24	1147 19.88	598 18.65	286 18.12
level 2	5119 20.79	3270 21.75	1955 22.16	1100 21.04	576 19.28	281 18.87
level 3	5012 22.91	3193 23.99	1910 23.92	1082 22.22	573 20.21	283 19.43
level 4	4942 24.82	3146 25.99	1882 25.73	1071 23.40	571 20.82	283 20.06

  - Analysis of change in bitrate and PSNR without encryption and with SSE of benchmark video sequences at QP value '12'.

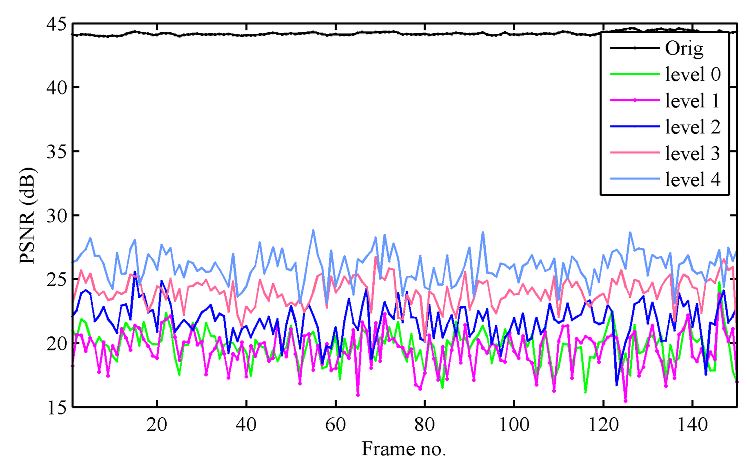
Seq.	Orig (kbps) (dB)	level 0 (kbps) (dB)	level 1 (kbps) (dB)	level 2 (kbps) (dB)	level 3 (kbps) (dB)	level 4 (kbps) (dB)
city	5337 48.71	5327 19.20	5285 18.48	5119 20.79	5012 22.91	4942 24.82
crew	4124 49.08	4100 23.16	4066 22.54	3962 25.12	3883 28.04	3831 28.63
harbour	5260 48.71	5291 19.00	5263 18.57	5080 21.23	4966 23.39	4892 24.67
ice	2729 49.51	2790 24.48	2772 24.34	2715 26.46	2674 29.13	2649 30.23
soccer	2348 49.38	2404 21.66	2350 21.19	2282 23.12	2227 26.09	2198 29.04

Framewise Analysis

- Frame-wise analysis of 2nd EL at 4CIF resolution with QP value '18' of *city*: a) Frame size, b) PSNR.



(a)



(b)

Visual Analysis

- Subframe of 280 × 240 pixels with offset of (400,200) in original frame of 2nd EL (4CIF) from 1st frame of *city* at QP 12: a) Without encryption, b) With SSE.

(a)

(b)

Conclusion

Encouraging results in the following contexts:

- Protection of different layers of spatially scalable bitstream.
- Real-time constraints have been achieved as:
  - No increase in bitrate. Hence it can be used in streaming applications,
  - It can be used for handheld devices as there is negligible increase in processing power.
  - The encrypted bitstream is completely decodable.