Video Encryption Exploiting Non-Standard 3D Data Arrangements

Stefan A. Kramatsch, Herbert Stögner, and Andreas Uhl
uhl@cosy.sbg.ac.at

UNIVERSITÄT SALZBURG

FACHHOCHSCHULE TECHNIKUM KÄRNTEN

Outline

Basic Idea: Arrange Visual Data in Non-standard Way and Apply Classical Techniques (Compression & Encryption) to improve results.

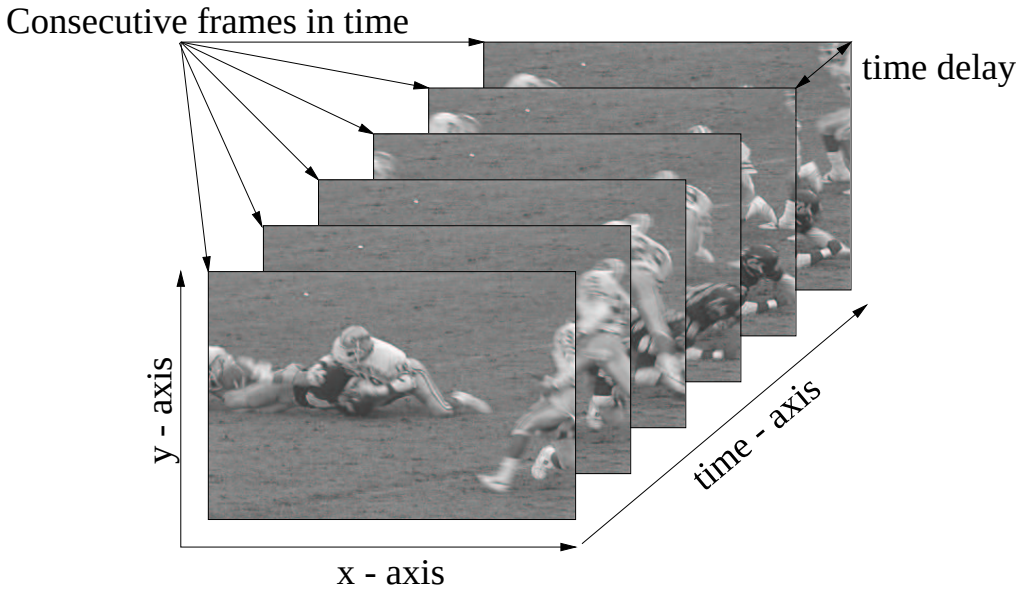
- Introduction - Video Coding
- Non-Standard Frame Structure
- Video Encryption
- The Proposed Encryption Technique
- Experimental Settings
- Experimental Results
- Conclusions

Introduction

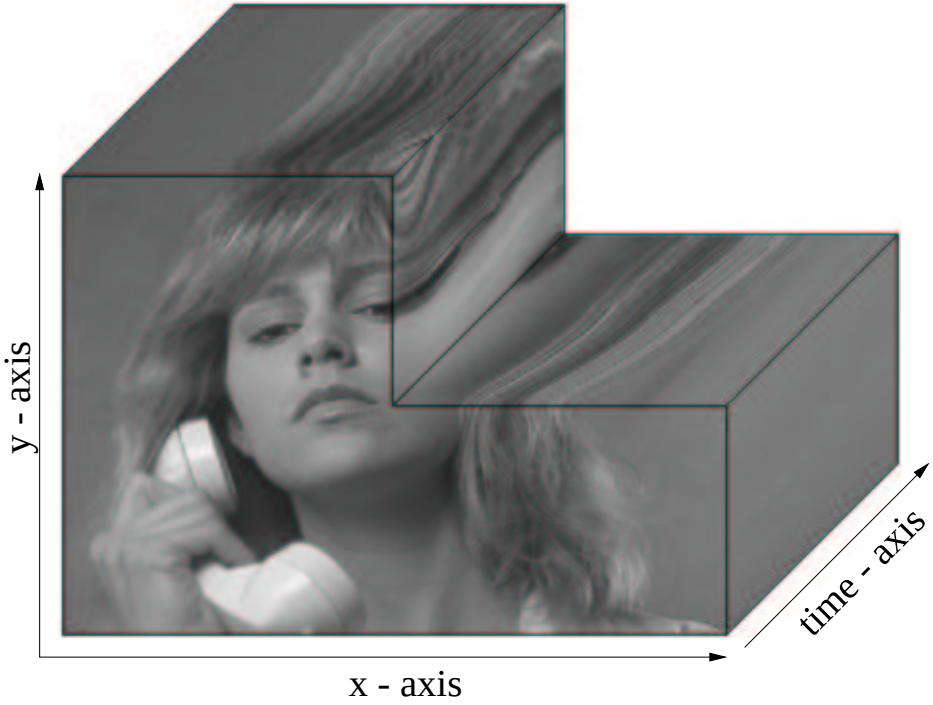
The majority of lossy video coding algorithms is based on motion compensated hybrid coding techniques like MPEG-1, MPEG-2, MPEG-4, and H.261 – H.264. These video codecs exploit the temporal redundancies present in video data (i.e. the similarity of frames over time) by applying algorithms for motion estimation and compensation.

Medical imaging is the main application field for lossless video coding. In these applications, most techniques employ lossless image coding techniques like lossless JPEG, JPEG-LS, or lossless JPEG2000 on a per-frame basis. Of course, temporal redundancy is ignored in such schemes which results in limited compression performance. On the other hand, three dimensional transforms (e.g. 3D DWT) are used to decorrelate the video data but these techniques suffer from high complexity and high memory requirements. Motion compensation techniques as employed in lossy schemes are seldom used in lossless environments.

Video Data: Classical View

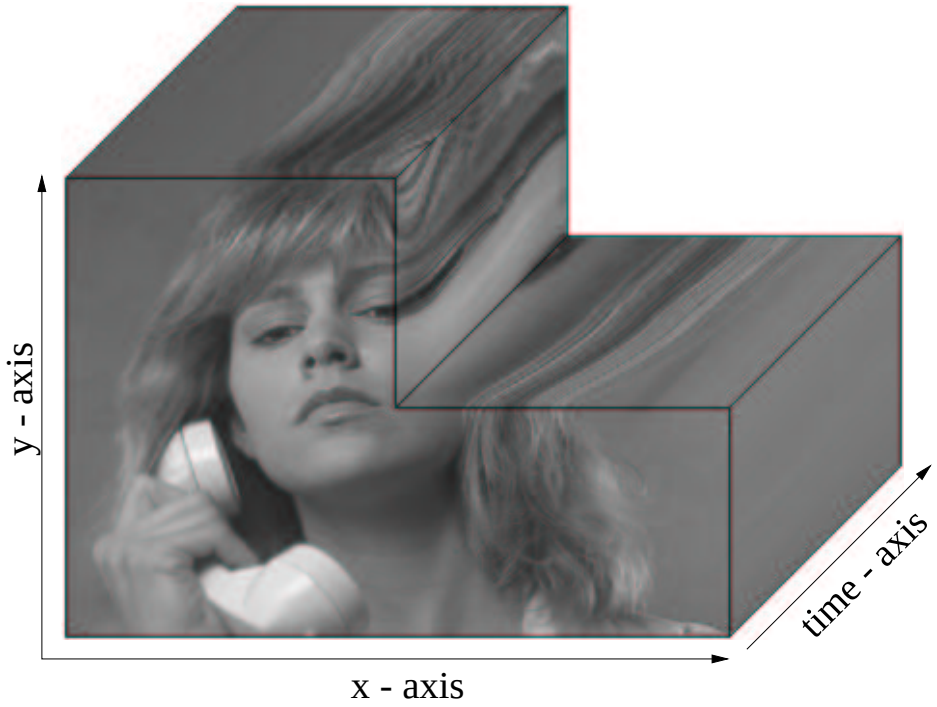


Sequence of Frames

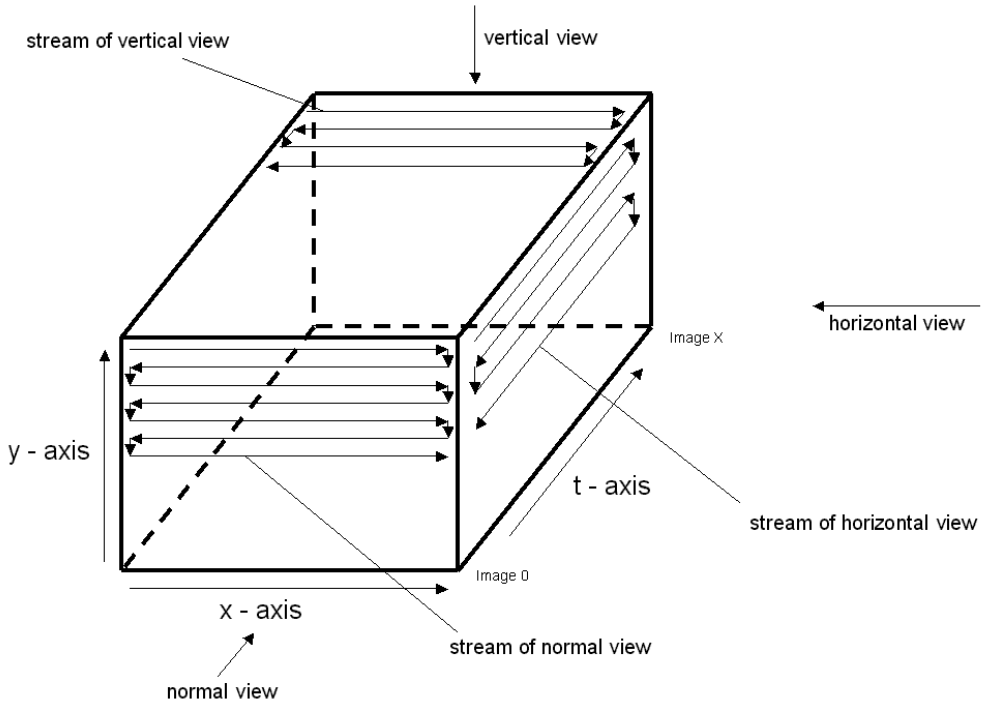


3D Block

Non Standard Frame Structure

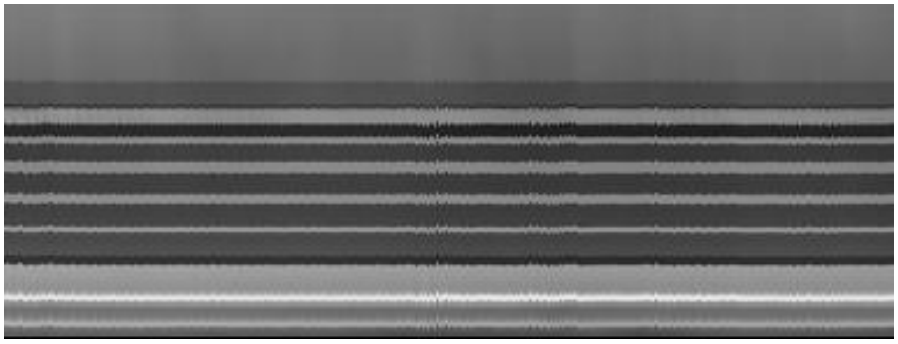


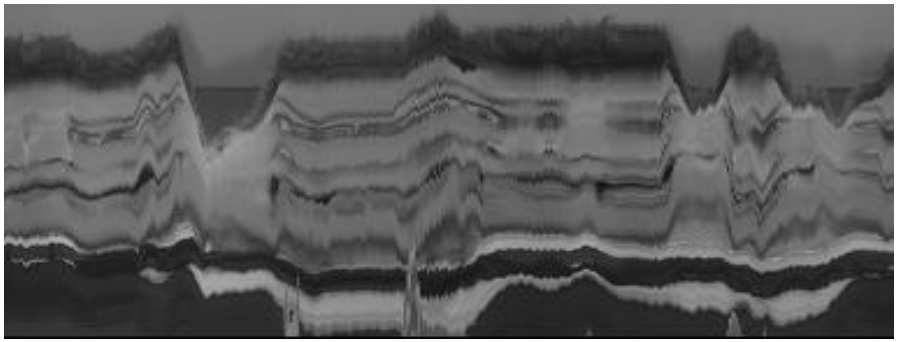
3D Block



Different Scans

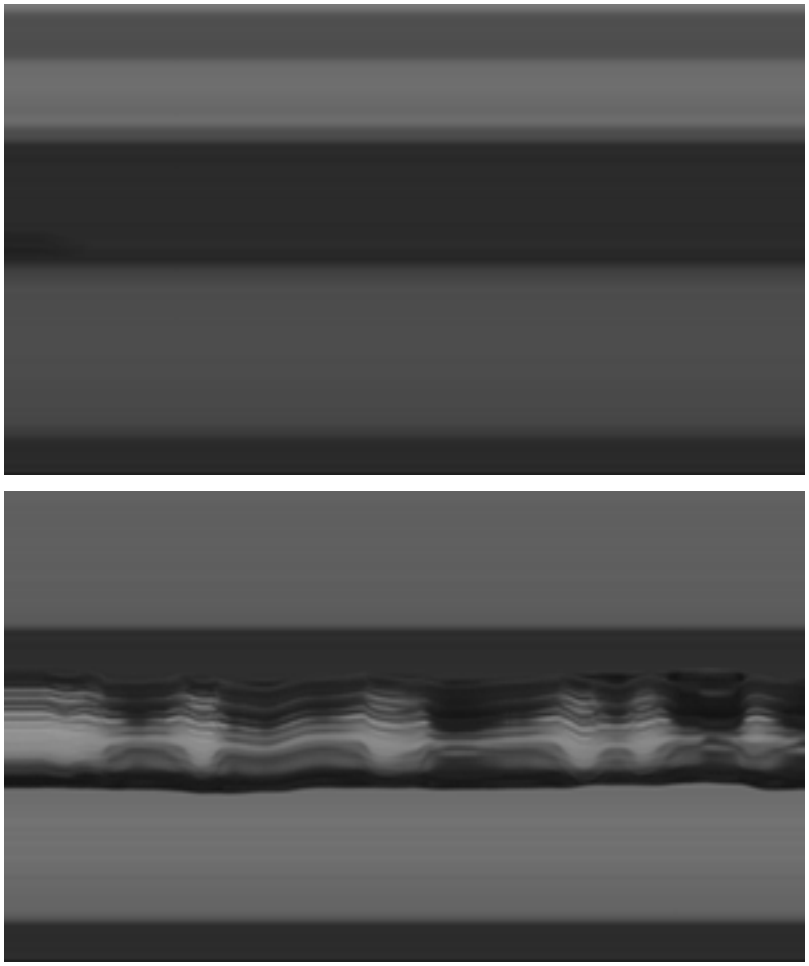
Example: Horizontal Frames - Carphone





- Horizontal frames provide a "side-view" of the video
- Frames close to the edge of the video consist mostly of background pixels which hardly change over time (left frame)
- The right frame consists of columns at the center of the original frames which change frequently over time.

Example: Vertical Frames – Akiyo



- Vertical frames provide a "top-view" of the video
- Frames close to the top of the video consist mostly of background pixels which hardly change over time (left frame)
- The right frame consists of columns at the center of the original frames which change frequently over time.

Example Application: JPEG2000 Compression

mode \\ video	Carphone	Claire
normal	1,877	2,727
vertical	2,122	3,908
horizontal	2,046	4,080

Compression is significantly improved, especially for low-motion video !

Image & Video Encryption

Image and video encryption schemes have been mostly discussed in the context of digital rights management (DRM) systems where the emphasis lies on lossy compression schemes. In this field, different strategies apply partial encryption either on a per frame basis (where selected coefficients or VLC codewords are protected only) or on a per group of picture (GOP) basis (where selected frames – I-frames – or selected macroblocks – I-blocks are protected only).

Obviously, medical imaging is also an application field where privacy and confidentiality are important aims. We discuss privacy schemes for lossless video. Since our underlying coding scheme does not use GOPs, we employ a frame based encryption scheme in our approach.

Our Approach

The aim is to exploit the alternative perspectives of video data as discussed for lossless video compression and encryption. The classical technique ("normal view") which we compare our results to corresponds to Motion JPEG2000 (MJPEG2000) encryption on a per frame basis (which is in fact JPEG2000 encryption).

We apply classical JPEG2000 encryption on a per frame basis to frames in horizontal and vertical perspective, respectively. We encrypt a certain percentage of the original amount packet data, reconstruct the video data and measure the visual quality in terms of PSNR and edge similarity score (ESS) besides subjective visual inspection.

In order to additionally assess the security of the scheme, we exploit a built-in error resilience functionality in JJ2000 in order to conduct an "error concealment attack".

Experimental Settings

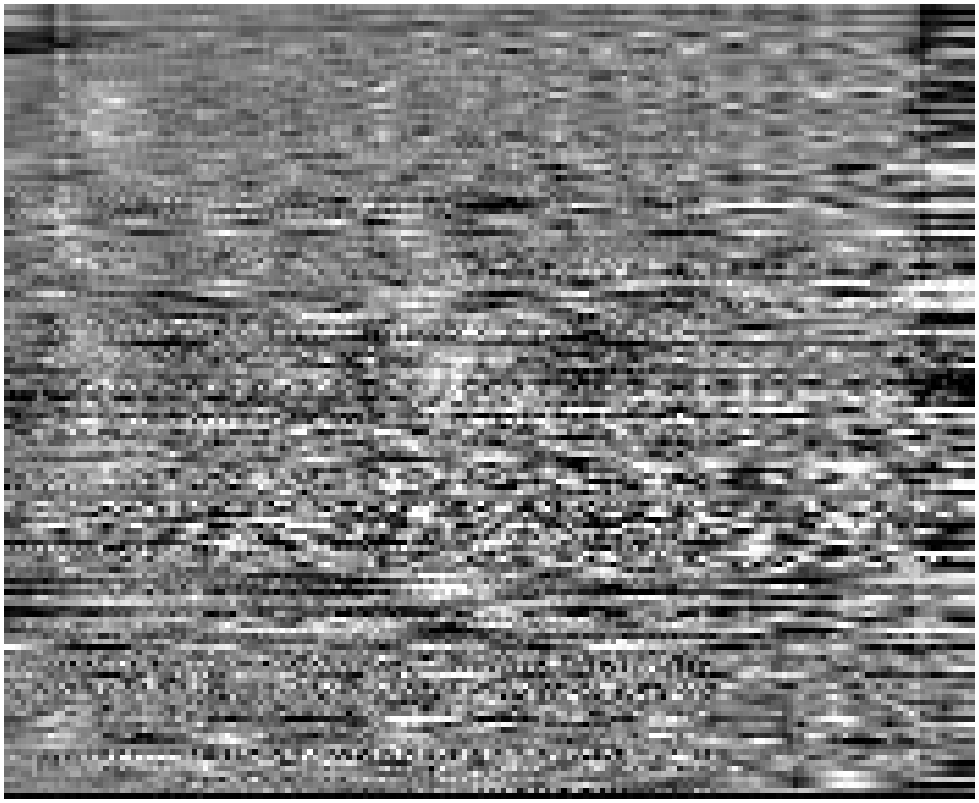
Carphone (176 x 144 x 383) and Claire (176 x 144 x 494) are used as testvideos, which represent sequences with high motion content and low motion content, respectively.

All displayed result show the 18th ("normal") frame of the corresponding test videos, numerical values represent average ESS and/or PSNR values for the entire sequence (in normal view) considered.

Error concealment attack: An error resilience segmentation symbol is inserted at the end of each bit-plane. Decoders can use this information to detect and conceal errors – encrypted data is treated as being errorness. This method is invoked in JJ2000 encoding using the option `-Cseg_symbol on`.

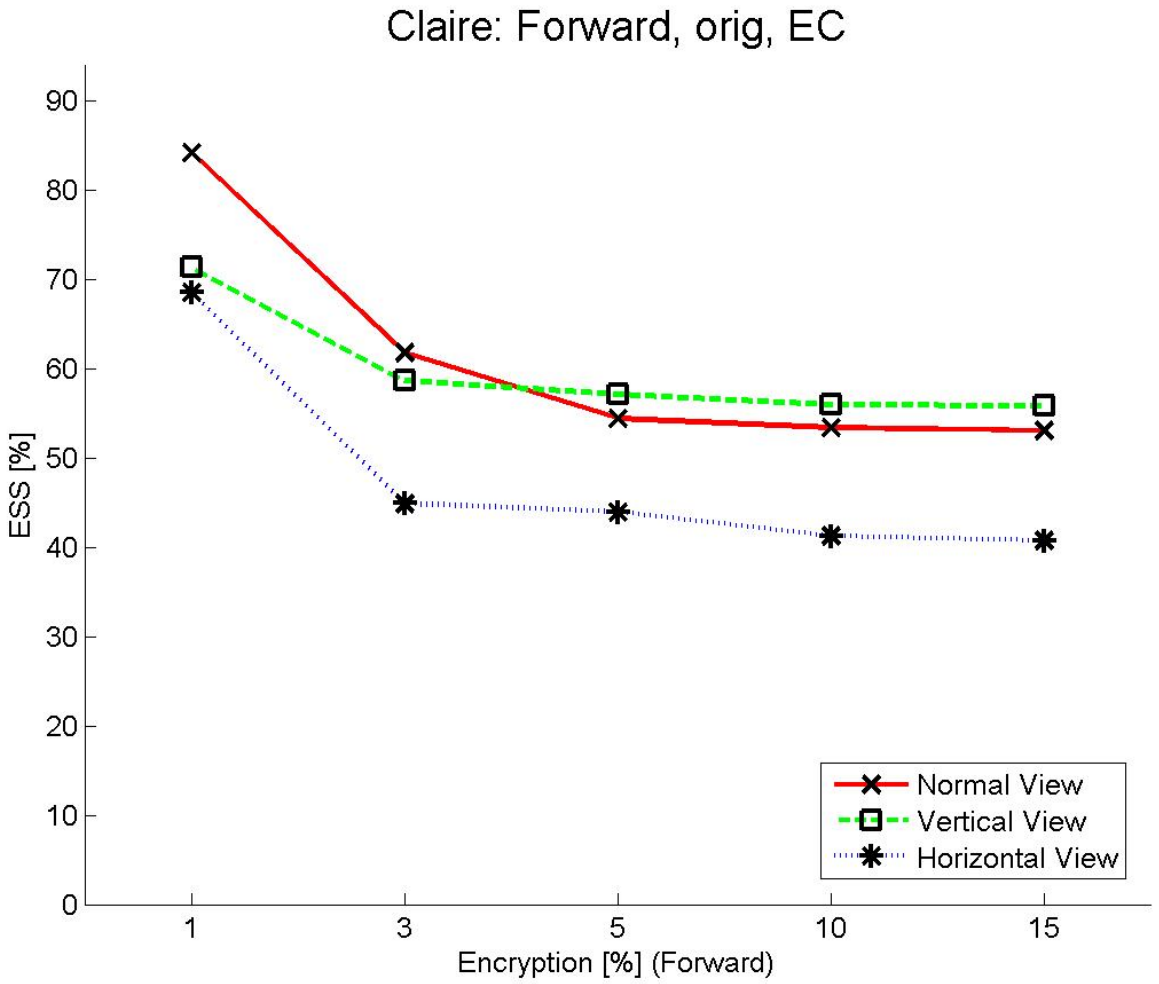
Results: Claire 3% encrypted, no attack

normal: ESS 0.54, 10.63 dB



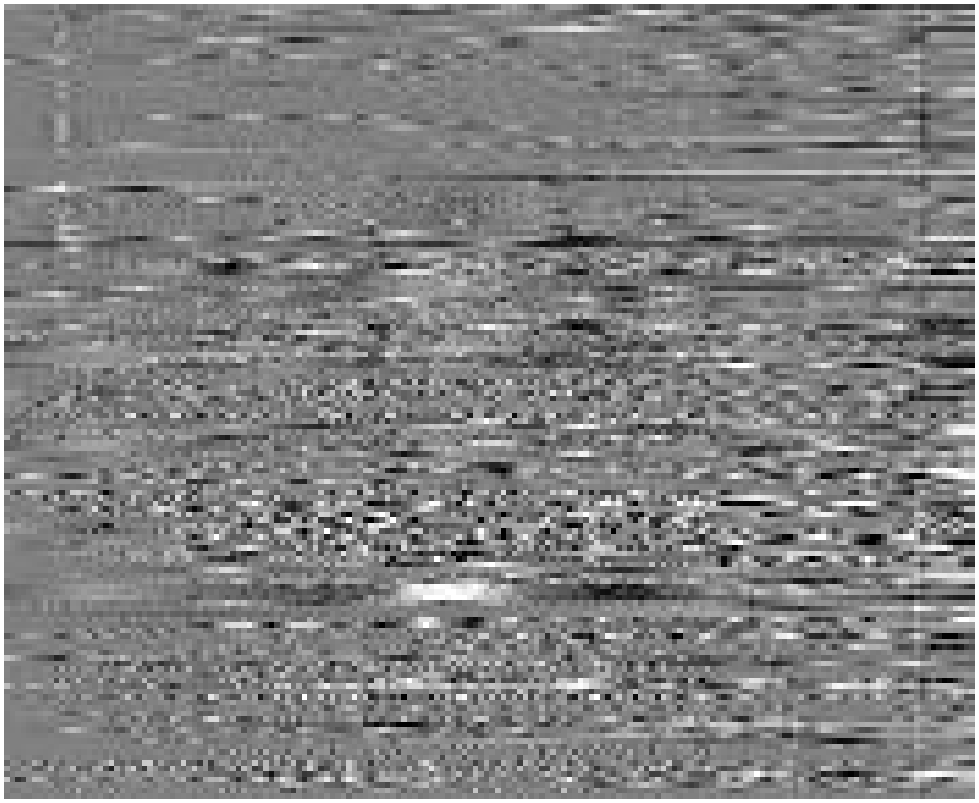
vertical: ESS 0.55, 10.67 dB

Results: Claire ESS vs. Encryption Amount (attacked)



Results: Claire 10% encrypted, attacked

normal: ESS 0.56, 9.81 dB



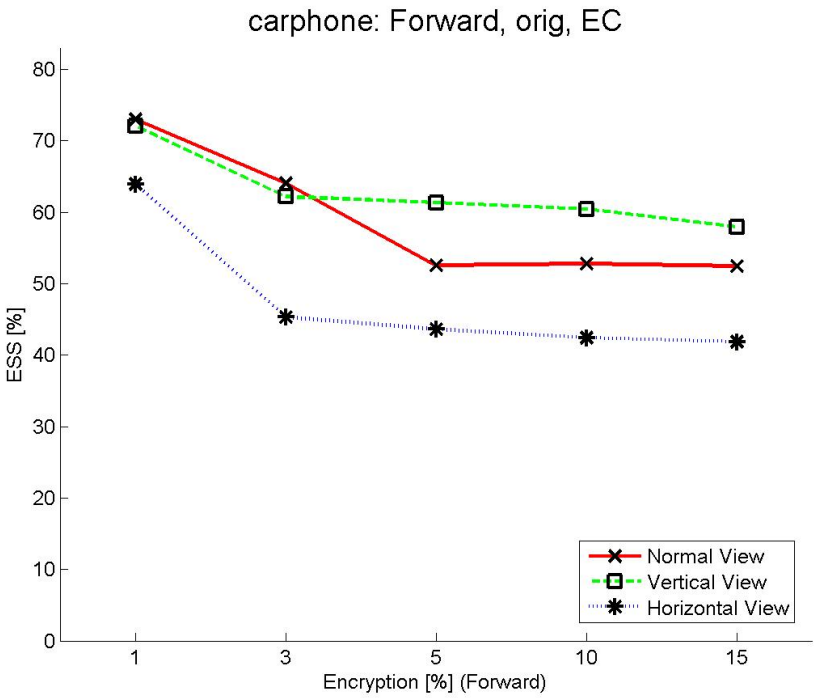
vertical: ESS 0.50, 11.43 dB

Results: Claire 10% transparently encrypted, attacked

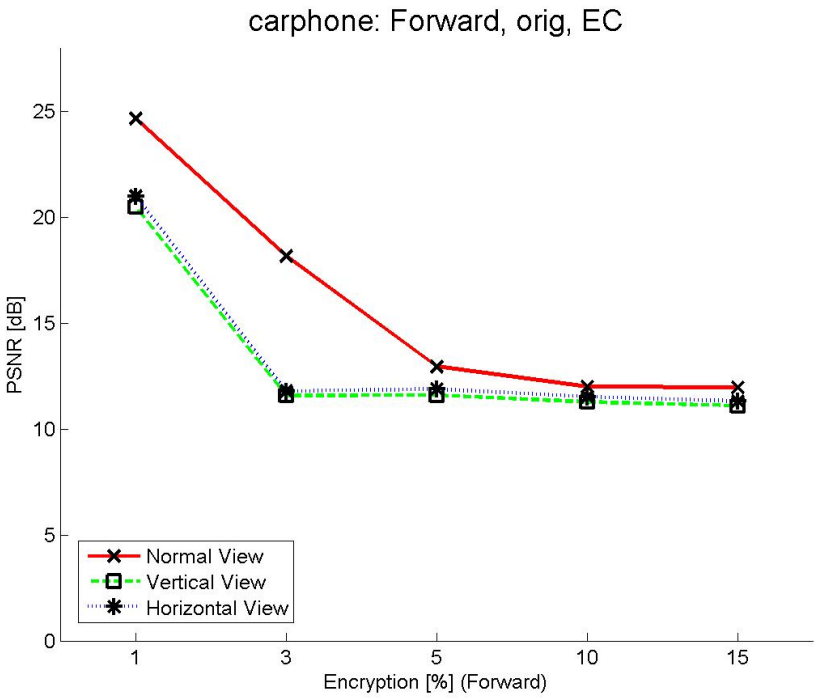
normal: ESS 0.78, 22.28 dB

vertical: ESS 0.80, 22.21 dB

Results: Carphone Quality vs. Encryption Amount



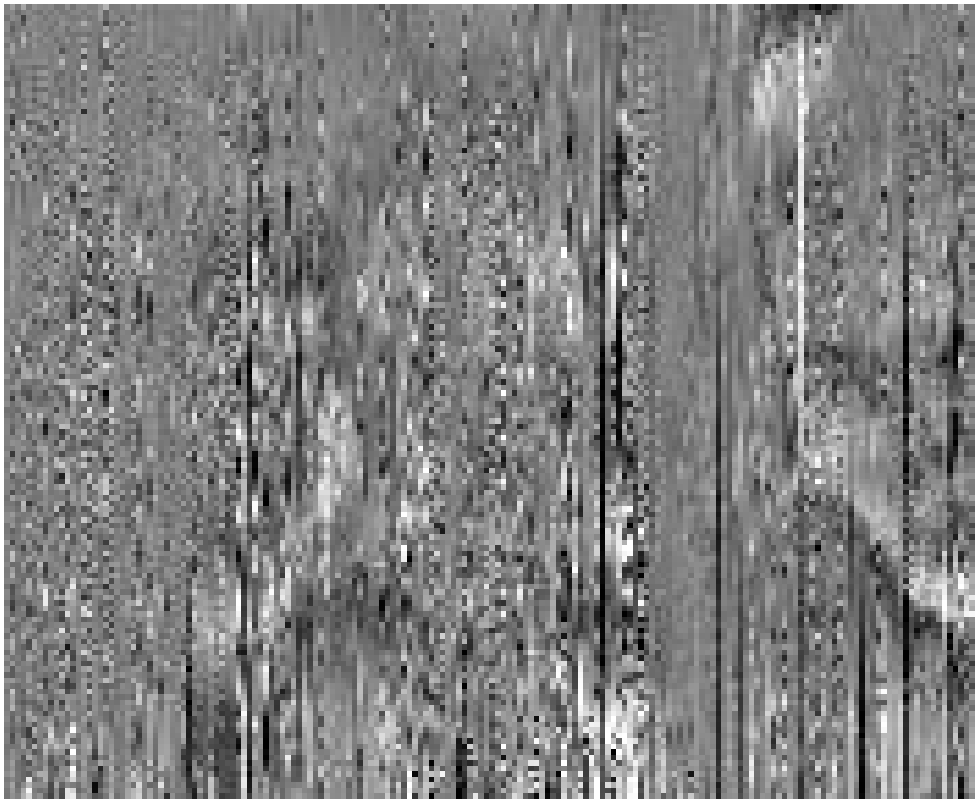
ESS



PSNR

Results: Carphone 15% encrypted, attacked

ESS 0.46, 12.45 dB



ESS 0.45, 11.48 dB

Conclusions

1. We have found that the advances with respect to compression performance caused by alternative interpretation and scan order of video data do carry over to the encryption of JPEG2000 based corresponding frames.

2. Privacy focused applications require a significantly lower amount of encryption effort when applying alternative scan orders as compared to classical MJPEG2000 video.

3. This findings do npt apply for transparent encryption scenarios.

Thank you for your attention !

Questions ?