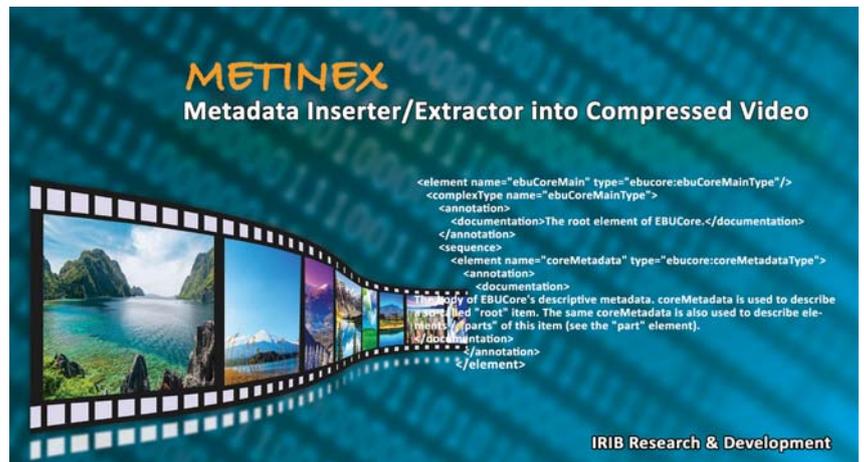# Metadata Inserter/Extractor into Compressed Video

In digital environments, metadata can be effectively hidden into the multimedia itself for general purpose metadata management. One of the most important applications of metadata in broadcast industry is in audio and video archiving. Broadcaster's archives contain thousands or billions of valuable contents so metadata has an essential role in searching and finding the required archived materials. In the archive automation workflow, besides the original video content, a high resolution (Hi Res.) and a low resolution (Low Res.) versions of the original content are generally generated and stored based on H.264/AVC codec. Users could then access the high or low resolution content according to their network access levels. For simple access to the desired archived contents, the search engines of the automation system are usually implemented based

on the Web. Therefore, the content security considerations and metadata confidentiality are challenges of these systems. Applying data hiding in compressed video is one of the agile and reliable solutions that can guarantee content security and metadata confidentiality. The main goal of the project is to insert metadata into compressed video without significantly increasing its bit rate volume and degrading its quality.

METINEX

Metadata Insertion into Compressed Video

The system is developed to insert metadata into archived compressed video and is used in automation system as two API modules: metadata inserter and metadata extractor.
The video files are archived in compressed forms, so it is necessary to insert metadata in the compressed video domain to avoid any decompressing process. The proposed metadata insertion system has two important advantages over the others; the quality transparency and minimum bit rate or capacity change after metadata insertion into compressed video.

Generally, there are two methods for inserting metadata into compressed video:
   (I) metadata insertion during video compression process.
   (II) direct metadata insertion into compressed video without full decompression.
In the first approach, if the video is in compressed form, it must be decompressed, and after metadata insertion then re-compressed.

This implies that in metadata insertion, video decoding and encoding is required and quality loss is inevitable. This undesired outcome limits the possibility of several times metadata insertion into the compressed video. Whereas in the second, there are no video compression/decompression artefacts. Therefore, one of the unique advantages of the proposed approach is its multi times fast metadata insertion and extraction capability with minimum video quality degradation.

Adding any strange data into the video can damage its statistical property, making it less compression efficient. As a consequence, the compressed bit rate is increased, and its quality deteriorates. The art of data hiding technique in video is to insert the data in a way that the bit rate and the quality of compressed video is not significantly altered by the hidden data. To do this, we have developed a novel method for H.264/AVC coded video, where metadata are hidden in the last non zero coefficient (LNZ) level of quantized DCT block in the scanning order. The unique advantage of this method is that, since data is hidden at the position of last non-zero high frequency DCT coefficient, they do not contribute to the video distortion. Moreover, since data is hidden at the highest possible frequency, its visual perception is very minimal. Second, altering the position of the last non-zero coefficients hardly affects the run-length bits and hence the increased bit rate is at its minimal. Experimental results on several test video sequences prove that the proposed approach can realize blind extraction with real-time performance, delivers very high capacity and low distortion and also keeps the increased bit-rate at a negligible level. On the data confidentiality, although for video this is not as important as the increased bit-rate and quality degradation are, but with a simple key, the metadata is first encrypted and then inserted into the LNZ of the H.264/AVC coefficients. This novel method is then adapted for data insertion into the compressed video stream. To do this, the compressed video is partially decoded, and the encrypted metadata is simply inserted in the LNZ. The performance of this method for various types of compressed I, P and B pictures are tested, and it is shown that while several times data insertions is possible, the multiple inserted data do not interfere with each other. For simple real-time operations, the data insertion and extraction parts are designed in two separate modules, each with the following characteristics.

Metadata Insertion Module Metadata inserter is a small part of decoder, where after entropy decoding of the Run-length DCT coefficients, the encrypted metadata is inserted into the LNZ of each block of 4x4 coefficients with a high frequency coefficient above a certain threshold. If entropy is the type of CABAC, some transitional probabilities of the H.264/AVC encoder are required to be added to this inserter module, but for Huffman entropy coding, the decoder itself has the required tables. In both types of entropy coding, since the compressed data is only partially decoded, the video quality is not degraded and more over the system is extremely fast. A large volume of metadata can be easily inserted to hundreds of frames, in a fraction of a minute.
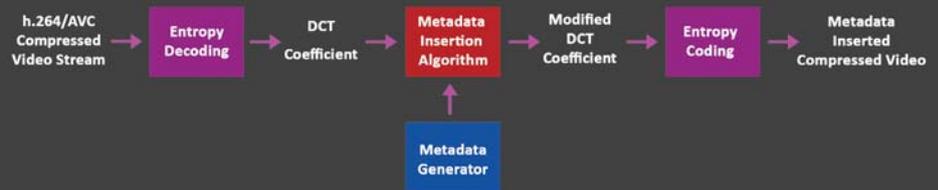
Metadata Extraction Module The inserted metadata bits are extracted in the decoding process of H.264/AVC where the quantized DCT levels for each macro block are entropy decoded. Then, according to metadata insertion algorithm, the inserted bits in macroblocks are identified and extracted to create inserted metadata information.
To obtain the raw metadata stream, the extracted encrypted metadata are decrypted by the encryption key. The data extractor is even faster than the data insertor, since it does not apply any entropy encoder. However, both modules are extremely fast, considering that a second of video comprises of 25 frames, while insertion or extraction of metadata in these frames can be done at a fraction of a video frame.