CHAPTER 2

MOTIVATION AND OBJECTIVES OF WORK

2.1. Identification of probable issues

Transcoding is the technology which converts one format of compressed bitstream to another format. The compressed bitstream is decoded to spatial domain video frames. These video frames are subjected to screen size resizing and frame rate conversion. The modified video frames, which are in spatial domain, are again coded by different encoder to compressed bitstream. Instead of decoding and encoding in spatial domain, there is a possibility of performing the transcoding operations in compressed domain in order to reduce the computational complexity.

Although the MPEG and ITU have released many video compression standards, the recently released H.264 Video compression standard is substantially different from previous standards (Siwei, M. et al. 2003). Though this standard is highly complex, it gives lower bitrate for same video quality than other standards. The syntax and the algorithms used in H.264 are so different that transcoding a compressed video to H.264 compliant bitstream will face many difficulties.

Compressed domain transcoding involves conversion between transforms used in video compression standards. Gao, C. et al. (2006) performed Fast DCT-to-Integer Transform conversion between MPEG2 and H.264 to improve the computational efficiency. But in H.264, 2D-DCT is used for 4x4 block size and performed in a different way. The DCT operation in H.264 is combined with quantization. The combined transform coefficients are scaled and truncated into integer numbers to perform integer transform. This integer transform process cannot be adopted for transcoding, because it leads to information loss. So there is a need of integer transform which does not involve with quantization.

Moreover, the integer transform should be reversible, so that it can switch between the spatial domain and compressed domain without any information loss. Researchers have
used 2D-DCT with either rounded or truncated coefficients to perform integer transform operations. For both the types of coefficients, the compressed domain results have not been equivalent to spatial domain results. So far, according to literature survey and review made by the author, the integer transform-based compressed domain transcoding has not been reported.

In the transcoding path, the combination of resolution change, bitrate change, frame rate change and codec change is possible. This research work focuses on only the resolution change, called resizing. Different approaches of changing larger resolution to smaller resolution have been studied and it has been found that integer transform-based resizing has not been attempted.

2.2. Motivation

The issues that arose out of the literature review, as listed above, which are considered worthy to be investigated in the H.264 related transcoding are the following.

H.264 employs 4x4 Integer transform and is different from 8x8 DCT used in other standards. The intra prediction is performed only in H.264. Transcoding of intra frame requires techniques for intra prediction in the compressed domain. The prediction block structures and motion estimation algorithms in H.264 are different from the previous standards. The Motion Vectors (MVs) extracted from a source video coded with other standards are not appropriate for the H.264 target video. New algorithms are needed to acquire more adequate MVs of the target video.

The drift error, due to data loss in compressed domain operation, because of its bits precision, is needed to be addressed. All the compressed domain transcoding architectures explained are having the compressed domain processes with reference to spatial domain frames. There is no compressed domain transcoding architecture which uses the compressed domain reference frames. Compressed domain resizing processes are implemented with the help of floating point DCT coefficients. There is no integer transform-based resizing process available. Another method of resizing by truncating high frequency coefficients has lost the edge patterns and blurred the high texture blocks.

Reusing the decoded syntax elements intelligently in the encoder was reviewed by many researchers. As the H.264 Standard is the newly developed video compression
standard, few researchers have addressed the complete utilization of all syntax elements from H.264 bitstream. Some researchers have addressed the modification of existing encoding or decoding algorithms to suit the hardware implementation to meet specific requirement. Also it is opined that a general approach on hardware implementation of compressed domain transcoding processes is required.

All the above issues have motivated to understand the need for compressed domain transcoding and define the problem statements which led to objective of this research work. Further studies made on H.264 Standard have given insight of encoding and decoding processes that could be exploited in the development of an efficient transcoder.

It has been shown that there is a need of defining the integer transform, which is reversible. Based on the integer transform, the processes such as intra prediction, motion compensation and motion estimation in compressed domain decoder and encoder are to be defined. The mode decision has to be re-defined based on compressed domain information. The compressed domain frames are to be resized into different resolution in the transcoding path. The modified syntax elements which are derived from decoder are to be adapted in encoding process. All the processes are to be modified to fit in for hardware implementation.

2.3. Objectives

The objective of the present work is to design a compressed domain homogeneous video transcoding through integer transform in compliance with H.264 Video compression standard. The transcoding techniques are to be modified to fit in hardware implementation. Computer codes are to be developed in MatLab environment to implement this transcoder. The correctness of the code is to be verified and validated with reference to H.264 Standard. The metrics to check the quality of video bitstream and efficacy of transcoder are to be identified. The efficacy of the transcoder algorithm is to be demonstrated with reference to standard software through the identified metrics.
2.4. Scope of Work

In order to meet the objectives of this research work, the scope of the work is defined as follows:

A compressed domain decoder will be designed and implemented with the help of integer transform. In the decoder, the incoming H.264 bitstream will be parsed and segregated into the syntax elements and residue coefficients. The intra prediction and motion compensation will be performed based on extracted syntax elements in the compressed domain. But the MV prediction will be compliant to H.264 Standard. The video frame in compressed domain will be reconstructed after adjusting the drift due to pixel overflow. This frame will be filtered as per H.264 Standard to remove blockiness. This filtered frame will be used as reference for future motion compensation.

The decoded frame will be resized to smaller screen size with the pixel count as multiples of 16, keeping any resizing ratio. This resizing process will be implemented in compressed domain with the help of integer transform. The required resizing filters will be designed based on the resizing ratio. And these filters will be applied on decoded frame to get resized frame in compressed domain. The syntax element reuse engine for the encoder will be designed and implemented to use decoded or parsed syntax elements such as MB type, intra modes, block sizes, MVs, sMB type, QP and SKIP information.

A compressed domain encoder will be designed with the help of integer transform. In this encoder, intra mode decision and motion estimation with reference to reconstructed frame will be performed in compressed domain. Various decisions such as, (i) best intra mode decision based on SATD and RD-Cost, (ii) motion search based on SAD, (iii) best block size decision based on RD-Cost, and (iv) best MB type decision among intra and inter based on RD-Cost, will be implemented in compressed domain. The identified syntax elements will be coded by fixed length or variable length and Exp-Golomb techniques in compliance with H.264 Standard. The residue coefficients will be coded through CAVLC. The output bitstream should comply with the H.264 Standard. All the compressed domain processes will be designed to suit implementation in hardware environment.

In the ensuing chapters detailed reporting of the research work carried out is presented starting with literature review.