\section{Introduction}

\textsf{Immersive virtual reality (VR) and its applications are growing very fast in the recent years. Hence, the need for wide field of view contents that can cover the whole 360-degree scene and make the interaction in this virtual environment feasible is an important factor. Panoramic content can be represented by a sphere that has been mapped to a two dimensional (2D) image plane using cylindrical or pseudo-cylindrical projections. In the case of cylindrical projections, where the spherical coordinates are mapped to the full rectangular 2D coordinates, the resulted image suffers from over stretching specially in the polar areas. Although cylindrical projections maintain the rectangular image format which is suitable for standard video codecs such as High Efficiency Video Coding (H.265/HEVC) and Advanced Video Coding (H.264/AVC), but projected images contain redundant information due to the over stretching. }\\

\textsf{A family of pseudo-cylindrical projections attempts to minimize the distortion of the polar areas of the cylindrical projections such as equirectangular projections, by bending the meridians toward the center of the map as a function of longitude while maintaining the cylindrical characteristics by preserving the parallel latitude lines, parallel. These projections approximate equidistant sampling of 360-degree scene (which can be represented by 3D spherical coordinates). Hence, the pixel density is roughly equal regardless of the position on the sphere, providing spatially stable quality without the need of processing an excessive amount of pixels in compression.}\\

\textsf{The benefits of pseudo-cylindrical projections include that they preserve the image content locally and avoid over stretching of polar areas. Moreover, images are represented by fewer pixels compared to respective cylindrically projected images (e.g. equirectangular panorama images) due to the fact that polar areas are not stretched. Due to fewer pixels, they may also compress better and are good candidates for panoramic image projection formats. Pseudo-cylindrical projections may be characterized based upon the shape of the meridians to sinusoidal, elliptical parabolic, hyperbolic, rectilinear and miscellaneous pseudo-cylindrical projections. Examples of these pseudo-cylindrical panoramas are shown in Figure 1. Two types of pseudo-cylindrical projections are shown in the figure: sinusoidal (Figure 1.a) and miscellaneous (Figure 1.b). Figure 2 represents a model of projected pseudo-cylindrical panorama. The effective picture area which contains the 360-degree panoramic data indicated by the solid line and the rectangular block grid is depicted with a dashed line.}\\

\section{PROBLEMS IN CODING OF PSEUDO-CYLINDRICAL PANORAMAS}

The boundary of the effective picture areas of pseudo-cylindrically projected spherical images is not rectangular and hence it is not aligned with the block partitioning grid used in the image and video encoding and decoding process. Hence, the blocks that includes the boundary of the effective picture area contain sharp edges which are not favorable for image/video coding standards. The mentioned non-rectangular content format affects both intra-frame and inter-frame coding process. This section analyzes this problem in intra-frame and inter-frame coding and later for each problem proposes solutions in order to improve the RD performance of these contents.

\subsection{Intra-Frame Coding Problem}

In intra-frame coding, sharp edges in the boundary areas of pseudo-cylindrical panoramas create a blocks with non-homogeneous texture which contain both actual picture content and pixels that are outside the effective picture area of the image. These non-homogeneous blocks will produce many high-frequency components after Discrete Cosine Transform (DCT) and quantization processes in the block compared to blocks with homogeneous texture, which typically have very few high-frequency values after DCT and quantization. The main problems that occur with the blocks containing sharp edges are as below:

\begin{itemize}

\item Intra prediction signal is typically not able to reproduce the sharp edge, causing the prediction error signal to be substantial and comprise a sharp edge too.

\item The high-frequency components cause an increase in bitrate. However, many coding schemes, such as zig-zag scan of DCT coefficients, have been tuned with the expectation that the high-frequency components are less likely and/or with a smaller magnitude than the low-frequency components.

\item The quantization of high-frequency components causes visible artifacts, such as ringing effect, for the entire decoded block particularly in the proximity of the sharp edges.

\end{itemize}

\subsection{Inter-Frame Coding Problem}

The inter-frame prediction of pseudo-cylindrical panoramas in the boundary areas is not efficient due to the fact that the samples in the reference frame are not available in the non-effective areas close to the boundaries of the effective picture. The reconstructed reference pictures that have non-rectangular effective picture area cause a sub-optimal inter prediction performance when:

\begin{itemize}

\item The prediction block or block to be encoded is in the boundary areas of the image and hence partially filled with non-effective picture area samples.

\item Both the prediction block and the block to be encoded cover a boundary of the effective picture area, therefore both include some data from the non-effective picture area.

\end{itemize}

This mismatch between block being encoded and prediction block in the reference picture causes extra error samples in the prediction error block and hence incur some bitrate. Particularly, this happens in the following cases:

\begin{itemize}

\item Figure 3.a and Figure 3.b represent the block to be encoded and the prediction block respectively and the object motion in them. The gray area illustrates the effective picture area and as it can be seen, the motion of the object is toward the inside of the effective picture area. As it can be seen from the prediction block, the object is partially inside the effective area and this missing parts lead to huge residual values. The resulted extra residuals are shown in Figure 3.c

\item The predicted block in the reference picture contains more samples than the block to be encoded in the current picture. The motion in this case is towards outside the effective picture area. Figure 3.d and Figure 3.e represent the situation where the predicted block in the reference picture contains more samples than the block to be encoded in the current picture. The extra prediction error samples will occur in the prediction error block as shown in Figure 3.f.

\item Another problem arises when a block in the current frame is inter predicted from a boundary region with fractional-pixel motion vector, in which case a motion compensation filter is applied to generate the prediction samples. Close to the boundary of the effective picture area, the motion compensation filter may use as input sample values from locations that are outside the effective picture area. The values of sample locations at different sides of the boundary may differ a lot, hence the motion compensation filter generates pixel values that has overshooting and undershooting effect because of the boundary edge. This overshooting and undershooting predicted values increase the values of residuals, which also increase the bitrate or distortion.

\end{itemize}

\section{IMPROVING COMPRESSION PERFORMANCE OF PSEUDO-CYLINDRICAL PANORAMAS}

This section, proposes two methods for intra-frame coding in order to overcome the sharp edge problem in the boundary areas of the pseudo-cylindrical panoramas. Along with the methods for intra-frame coding, a method used for enhancing the performance in inter-frame coding of these panoramas.

\subsection{Intra-frame Coding}

As it has been discussed in section 2.1, the high-frequency components resulted in the boundaries caused by sharp edges are not favorable to the current video coding standards such as HEVC and AVC. In order to avoid these high-frequency components, the boundary blocks which contain samples from the non-effective picture area must be filled with samples that are more correlated with the effective picture area samples. As a result, this correlation of pixel values can be easily handled with DCT transform and quantization in the encoding. process.

\subsubsection{Padding the Boundary Samples}

\textsf{Filling the boundary blocks which are partially containing the samples of non-effective picture area, are done by using the boundary samples of effective picture area. In this method, the first and last pixel of the effective picture boundary are replicated to the boundary blocks in the left and right side of the effective picture area respectively. By padding the boundary pixels row-wise in the neighbor blocks we make the boundary block samples high correlated and this high sample correlation helps the encoder to encode these blocks efficiently.}\\

\textsf{The results of padding method in the boundary block areas are shown if Figure xx. As it can be observed, the texture in boundary areas is uniform and hence this enables the codec to compress these areas efficiently. }\\

\subsubsection{Copying and Padding the Boundary Samples from the Opposite-Side}

\textsf{Since the left-most pixels of the left boundary and the right-most samples of the right boundary of the 360-degree panoramic images are adjacent to each other, the samples from the opposite side of the effective picture can be used in order fill the boundary blocks. This can be effective particularly when there are significant amount of samples within the boundary block that are also within the effective picture area. These empty parts of the block may be filled with data from the other side to make the content of the block smooth to be efficiently compressed. }\\

\textsf{The copying method can be applied easily to the boundary areas; however, the polar areas of the pseudo-cylindrical panoramas are the problematic parts, since there are not many samples in these areas to be copied to boundary blocks, so the boundary blocks in the poles will be partially filled. The partially filled blocks will create high-frequency components in the encoding process, hence it would be more efficient to fill the boundary blocks which are partially filled with samples that are copied from the opposite side with some data that can preserve the correlation of samples inside the block. The padding method which is used in section ‎3.1.1 can be helpful in this situation. After copying samples from the opposite side to fill the boundary blocks, partially filled blocks are detected, and then the rest of the block are filled by replicating the first and the last pixel of each row to the remaining areas inside the block. The resulted images of this method are shown in Figure x. }\\

\subsubsection{Encoding and Decoding Side}

Figure xx illustrates the whole process of encoding and decoding of proposed intra-frame methods. The “padding” or “copying plus padding” methods can be applied either as a pre-processing step or can be implemented as an in-loop process inside the codec. We considered this process as a pre-processing step. The benefit of doing pre-processing is that, it does not require to change the coding standard and the whole process takes only one pre-processing before encoding and one post-processing after decoding. As can be seen from block diagram in the figure, the pre-processed video is fed to the encoder and then the bitstream sent through network for receiver. In the receiver side, the video will be decoded; but the decoded video contains extra information in the boundary areas resulted from pre-processing state. The post-processing step is applied in order to extract the video in original format by detecting the effective picture area using a pre-defined mask.

\subsection{Inter-Frame Coding}

\textsf{In order to prevent the mentioned problems in section 2.2 for inter-frame coding of pseudo-cylindrical panoramas, the advantage of 360-degree characteristics of pseudo-cylindrical panoramic images which the right-most pixel in each row considered to be adjacent to the left-most sample in the same row inside the effective picture area is considered. In order to improve the inter prediction of these 360-degree panoramic videos, in the reference frames, samples are copied from the opposite-side of the effective picture area in the corresponding pixel row, to fill the non-effective picture area in each side of the cropped image.}\\

\textsf{The row-wise circular copying samples from the opposite-side in the reference frame results for MyShelter stationary camera and Bear Attack sequences are shown in Figure xx respectively. As it can be noticed from the figures, the sample continuity is established in the boundary areas of the manipulated reference frame which enhances the inter prediction of pseudo-cylindrical panoramas. }\\

Expanding the samples from the opposite side to the non-effective picture area helps the prediction of inter frames by filling the prediction blocks in the reference picture with adjacent samples. Two main advantages of this method that improves the coding of inter frames are:

\begin{itemize}

\item The non-effective picture area is filled by samples from the opposite side which provides continuity in the boundary areas of the reference picture. This continuity of samples in the boundary areas helps the better prediction from the manipulated reference picture.

\item Fractional sample interpolation is improved since the boundary areas do not contain edges anymore, as no overshooting or undershooting pixel values is generated when using motion compensation filter.

\end{itemize}

Manipulating the reference frame improves the inter prediction, but on the other hand the bitrate increases due to the following reasons.

\subsubsection{Residual Manipulation}

Manipulating the reference frame by copying the samples from the opposite side creates unwanted extra residuals in the prediction error blocks outside of the effective area. These extra residuals should not be coded into bitstream, otherwise it will increase the bitrate significantly. Hence, in order to avoid such bitrate increment, these motion compensation residuals are replaced by zero values. By replacing the residuals which are located outside of the effective picture area with zero values, the encoder can code these areas with fewer bits compared to non-zero residual values. By replacing these unwanted residuals with zeros we avoid the extra bitrate but on the other hand, the reconstructed image will contain unwanted data in non-effective area which is the result of the copied data in reference frame. The following steps are required for handling this extra data in reconstructed frame:

\begin{itemize}

\item The extra data in non-effective area are replaced by zero values before applying the reference frame manipulation step.

\item The extra data are removed as a post-processing step after decoding.

\end{itemize}

\subsubsection{Manipulating Distortion Calculation Functions}

The extra residuals will affect the distortion cost calculation. During the rate-distortion optimization process at the encoder side, the reconstruction error of the pixels outside of the effective picture area should be excluded from distortion cost (e.g. sum of absolute differences) calculation. Since the residual values outside of the effective picture area of the current frame are replaced by zeros, then the reconstructed picture will contain the copied sample information from the reference frame in the non-effective picture areas. These samples will be omitted in the decoder end.

\subsubsection{SAO Modification}

After prediction process of current frame, HEVC applies some in-loop filtering techniques (e.g. deblocking filter (DBF)) in order to reduce the coding artifacts in the frames ‎[16]. One of the filtering techniques that applied on reconstructed frames is Sample Adaptive Offset (SAO). SAO applies after deblocking filter and it tries to reduce the mean sample distortion between original and the reconstructed image ‎[17]. Since the reconstructed picture includes the extra samples from the reference frame, the SAO process adds huge offset values to the samples outside the effective area in order to compensate this difference with original picture. The added offset values cause very high bitrate in the encoding process. Hence, to avoid this unnecessary offsets, the SAO is disabled in the encoding side.

\subsubsection{Encoding and Decoding Process}

\textsf{The hybrid block diagram of encoder and decoder process in HEVC including the proposed methods for inter prediction in this work is represented in Figure xx and Figure xx, respectively.}\\

\textsf{As Figure xx illustrates, in the encoding process, the reconstructed frame is passed to Reference Frame Manipulation (RFM) unit prior to filtering (F) and storing in Reference Frame Memory (MEM). The reference frame manipulation is applied in RFM unit and then stored in MEM for inter-frame prediction operations. The process of setting the residual values to zeros outside of the effective picture area is applied in SRZ unit before transform (T) unit, this process is applied only in the encoder side.}\\

\textsf{Figure xx demonstrates the decoding process of the proposed algorithm. The similar operations as the encoder side performed in reverse order. The reconstructed picture in pixel prediction operation is passed to RFM unit for reference frame manipulation. The decoded frames later passed to Output Cropping (OC) unit in order to extract the pseudo-cylindrical panorama from the manipulated format. This process includes a predefined mask representing the effective picture area boundaries. The samples outside of the effective picture area will be set to initial background values.}\\

\textsf{}\\