POZNAŃ UNIVERSITY OF TECHNOLOGY

FACULTY OF COMPUTING AND TELECOMMUNICATIONS

INSTITUTE OF MULTIMEDIA TELECOMMUNICATIONS

Doctoral dissertation

PREDICTION TECHNIQUES FOR COMPRESSION OF MULTIVIEW VIDEO ACQUIRED USING SYSTEMS WITH VARIOUS CAMERA ARRANGEMENTS

Jarosław Samelak

Supervisor: prof. dr hab. inż. Marek Domański

Auxiliary supervisor: dr inż. Damian Karwowski

Poznań 2023

POLITECHNIKA POZNAŃSKA

WYDZIAŁ INFORMATYKI I TELEKOMUNIKACJI

INSTYTUT TELEKOMUNIKACJI MULTIMEDIALNEJ

Rozprawa doktorska

TECHNIKI PREDYKCJI DLA KOMPRESJI WIZJI WIELOWIDOKOWEJ REJESTROWANEJ ZA POMOCĄ SYSTEMÓW O RÓŻNYM ROZMIESZCZENIU KAMER

Jarosław Samelak

Promotor: prof. dr hab. inż. Marek Domański

Promotor pomocniczy: dr inż. Damian Karwowski

Poznań 2023

TABLE OF CONTENTS

ABSTRACT

The dissertation presents the author's research on novel prediction techniques for multiview video compression. The author identifies the limitations of the state-of-the-art techniques and then proposes two original solutions. The main goal is to improve the inter-view prediction in compression of multiview video acquired using systems with various camera arrangements. The proposed techniques are dedicated to modern applications of multiview video, such as virtual reality (VR), augmented reality (AR), or immersive video systems.

In the first part of the dissertation, the author adapts the state-of-the-art compression technique, 3D-HEVC, to the aforementioned applications. To achieve that, the author first proposes a novel, original rectification method for multiview video acquired by cameras distributed roughly on a circle. Then, the author modifies the inter-view prediction of 3D-HEVC to efficiently compress such rectified video. The modified codec, in the dissertation referred to as ARC-HEVC, is evaluated in terms of rate-distortion (RD) compression efficiency and coding time, and the results are compared to the state-of-the-art 3D-HEVC. According to the results, the author's proposal is both faster and more efficient.

In the second part of the dissertation, the author proposes a novel idea of using Screen Content Coding (SCC) for compression of frame-compatible multiview video. The main idea is to utilize Intra Block Copy as an inter-view prediction tool. The advantage of such a solution is that the codec does not require a complex multi-layer structure dedicated exclusively to the processing of multiview video, contrary to the state-of-the-art Multiview HEVC (MV-HEVC). Additionally, the author proposes a set of original modifications to improve the efficiency of Screen Content Coding in compression of multiview (including stereoscopic) video. Experimental evaluation shows that the author's novel approach, in the dissertation called Advanced SCC or ASCC, provides virtually the same RD compression efficiency and encoding time as MV-HEVC, both in the coding of stereoscopic and multiview video.

Both SCC and ASCC codecs are also employed by the author as the inner codecs in MPEG Immersive Video, replacing commonly used HEVC codec. Experimental evaluation of the proposed change in compression of immersive video shows a significant gain in rate-distortion compression efficiency and the quality of virtual views, at the cost of increased encoding time.

STRESZCZENIE

Niniejsza rozprawa prezentuje przeprowadzone badania w kierunku oryginalnych technik predykcji dla kompresji wizji wielowidokowej. W rozprawie autor identyfikuje ograniczenia aktualnych technik oraz proponuje dwa oryginalne rozwiązania. Głównym celem jest poprawa predykcji międzywidokowej w kompresji wizji wielowidokowej zarejestrowanej systemami o różnym rozmieszczeniu kamer. Zaproponowane techniki są przeznaczone dla nowoczesnych zastosowań wizji wielowidokowej, takich jak wirtualna rzeczywistość (VR), rzeczywistość rozszerzona (AR), czy systemy wizji wszechogarniającej.

W pierwszej części rozprawy autor dostosowuje technikę kompresji 3D-HEVC do wspomnianych zastosowań. W pierwszej kolejności autor proponuje nowatorską, oryginalną metodę rektyfikacji wizji wielowidokowej rejestrowanej za pomocą kamer rozmieszczonych w przybliżeniu na okręgu. Następnie autor modyfikuje predykcję międzywidokową w 3D-HEVC dla efektywnej kompresji tak zrektyfikowanej wizji. Zmodyfikowany kodek, w rozprawie określany jako ARC-HEVC, jest oceniony pod kątem efektywności kompresji i czasu kodowania, a wyniki są porównane z oryginalnym 3D-HEVC. Wyniki pokazują, że proponowana metoda jest zarówno szybsza, jak i bardziej efektywna.

W drugiej części rozprawy autor proponuje oryginalną metodę polegającą na zastosowaniu techniki Screen Content Coding (SCC) w kompresji wizji wielowidokowej zgodnej ramkowo (ang. *frame-compatible*). Główną ideą jest zastosowanie Intra Block Copy jako narzędzia predykcji międzywidokowej. Zaletą takiego rozwiązania jest to, że kodek nie wymaga złożonej, wielowarstwowej struktury dedykowanej wyłącznie kodowaniu wizji wielowidokowej, w przeciwieństwie do aktualnej techniki Multiview HEVC (MV-HEVC). Dodatkowo, autor proponuje szereg modyfikacji poprawiających efektywność Screen Content Coding w kompresji wizji wielowidokowej (także stereoskopowej). Ocena eksperymentalna pokazuje, że metoda zaproponowana przez autora, w rozprawie określana jako Advanced SCC lub ASCC, jest równie efektywna jak MV-HEVC z punktu widzenia poziomu kompresji i czasu kodowania, zarówno dla kodowania wizji stereoskopowej, jak i wielowidokowej.

Kodeki SCC i ASCC zostały także zastosowane przez autora jako wewnętrzne kodeki w technice MPEG Immersive Video, zastępując powszechnie stosowany kodek HEVC. Ocena eksperymentalna dla kompresji wizji wszechogarniającej pokazuje, że zaproponowane przez autora rozwiązania są znacząco lepsze pod względem efektywności kompresji i jakości widoków wirtualnych, kosztem zwiększonego czasu kodowania.

LIST OF SYMBOLS AND ABBREVIATIONS

α_i	angle between the direction of the *i*-th camera optical axis and Z axis, derived in circular rectification
b	bit depth of the depth sample value
c	skew factor (intrinsic camera parameter)
d	sample value from depth map
d_x, d_y	horizontal and vertical component of disparity vector
f_x, f_y	horizontal focal length, vertical focal length
$\mathbb{K}$	[3×3] intrinsic parameter matrix
o_x, o_y	coordinates of the optical center
o'_x	modified horizontal coordinate of optical center
$\mathbb{P}$	[4×4] projection matrix
r	radius of the circle in a circular camera arrangement
$\mathbb{R}$	[3×3] rotation matrix
$\mathbb{T}$	3-component translation vector
x, y	positions of a point in a view
X_{cen}, Z_{cen}	position of the circle center on XZ plane (parallel to the ground)
X_i, Z_i	position of *i*-th camera on XZ plane
X'_i, Z'_i	modified position on the circle of *i*-th camera on XZ plane
z	distance between camera plane and acquired point in 3D space
Z_{near}, Z_{far}	depth maps normalization parameters

2D	two-dimensional
3D	three-dimensional
3DTV	three-dimensional television
ANY-HEVC	HEVC adapted to arbitrary camera arrangements
AR	augmented reality
ARC-HEVC	HEVC adapted to circular camera arrangements
ASCC	Advanced Screen Content Coding
AVC	Advanced Video Coding
BD-rate	Bjøntegaard delta rate
CTC	common test conditions
DCP	disparity-compensated prediction
FTV	free-viewpoint television
HEVC	High Efficiency Video Coding
HM	test model for HEVC
HTM	test model for MV-HEVC and 3D-HEVC
IBC	Intra Block Copy
MIV	MPEG Immersive Video
MPEG	Moving Picture Experts Group
MVD	multiview video plus depth
MV-HEVC	Multiview High Efficiency Video Coding
PSNR	Peak Signal-to-Noise Ratio
QP	Quantization Parameter
RD	rate-distortion
SCC	Screen Content Coding

TMIV	test model for MPEG Immersive Video
VPS	Video Parameter Set
VR	Virtual Reality
VVC	Versatile Video Coding

1. INTRODUCTION

1.1. SCOPE OF THE DISSERTATION

Multiview video [Ho'07, Vetro'11B] is a set of video sequences acquired synchronously by multiple cameras. The number of cameras and their locations vary, depending on the application. Multiview video can be enriched with depth data representing the distance between the camera plane and a given point in the acquired scene [Sullivan'09]. Such representation is called Multiview Video plus Depth (MVD) [Müller'11]. Figure 1.1 presents an example of an MVD frame composed of 3 views and corresponding depth maps obtained from one of the test multiview sequences.

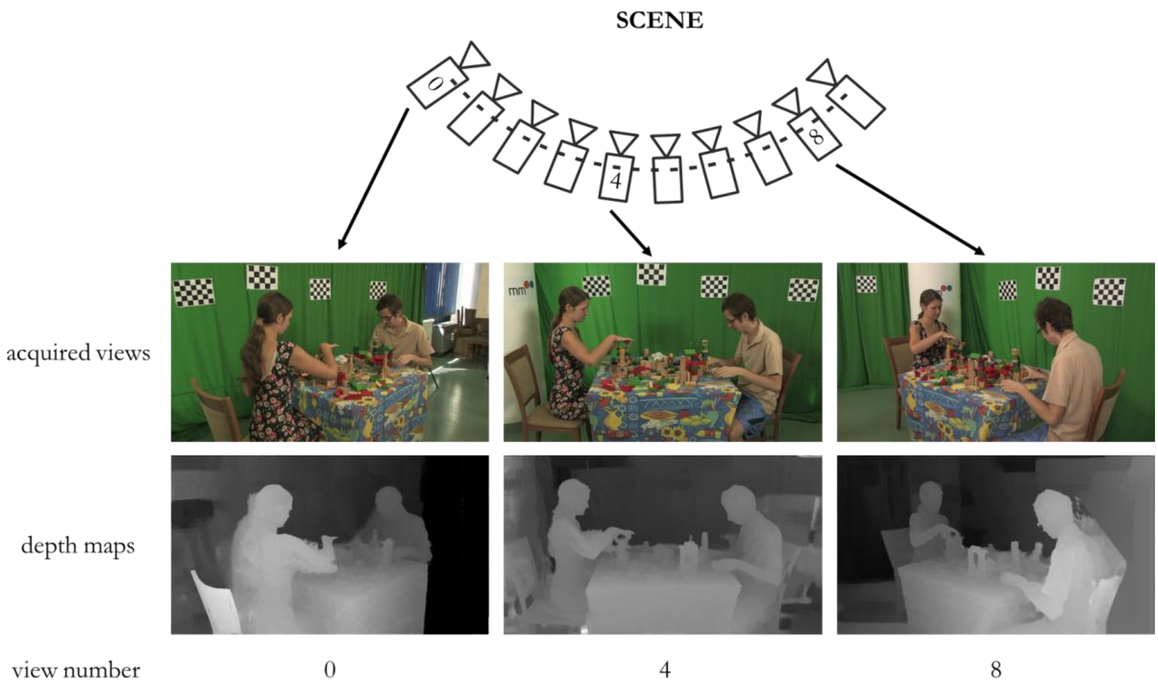

Figure 1.1. Example of a multi-camera setup and an MVD frame.

In recent years, many new applications utilizing multiview video have been developed, e.g., advanced three-dimensional television (3DTV), free-viewpoint television (FTV), virtual reality (VR), augmented reality (AR), and immersive video. Their goal is to satisfy the demand for more realistic and engaging multimedia, compared to the standard two-dimensional video.

A major challenge related to the abovementioned applications is the compression of multiview video [Domański'19]. It is estimated that video data already accounts for roughly 80% of global Internet traffic, and its share continues to grow due to (among other reasons) emerging video applications such as virtual reality or immersive video [Cisco'18, Cisco'20]. Therefore, a vast amount of video data produced by multi-camera setups has to be efficiently compressed, preserving a high quality of the content at the same time. **This dissertation focuses on the efficient compression of multiview video.**

A straightforward way to compress multiview video is to process each view separately (simulcast encoding) using one of the existing compression techniques for monoscopic video, such as High Efficiency Video Coding (HEVC) [ISO'21, Sullivan'12] or Versatile Video Coding (VVC) [ISO'22, Bross'21]. However, this approach is inefficient as it does not utilize the similarities between the views. To address this issue, dedicated multiview video compression techniques have been developed in recent years. The most recent techniques in multiview video coding are listed below.

- Multiview HEVC (MV-HEVC) [Tech'16, Hannuksela'15]. An extension developed on top of HEVC that adds multi-layer coding with inter-view prediction, together with the required signalization.
- 3D-HEVC [Tech'16, Sullivan'13, Müller'13]. This technique is dedicated to the joint coding of camera views and depth maps. It further extends MV-HEVC, primarily by utilizing information about depth, which allows to, e.g., better predict the disparity between the views.
- MPEG Immersive Video (MIV) Coding [Boyce'21]. Dedicated to immersive video that provides playback with six degrees of freedom (6DoF) for end users. Inter-view redundancy is reduced through view synthesis and packing of the occluded parts of each view into atlases. The output images are then compressed using one of the standard compression techniques, usually HEVC.
- Versatile Video Coding (VVC) [Bross'21, ISO'22]. Although it is not its main application, the multi-layer coding functionality in VVC allows it to perform inter-view prediction in a similar manner as MV-HEVC. However, VVC lacks the capability of using depth maps to improve the compression efficiency, known from 3D-HEVC. Moreover, during the author's research, VVC was still under extensive development. **Therefore, this dissertation focuses on HEVC-based coding techniques as the state-of-the-art in multiview video compression.** Nonetheless, in the dissertation, the author takes into account future development and, whenever suitable, provides comments on the possibility of applying proposed solutions to the VVC-based codecs.

The abovementioned coding tools stem from monoscopic video compression techniques but are equipped with additional coding tools for reducing inter-view redundancy.

The coding efficiency of those techniques varies, depending on the application and the content being encoded. For instance, in the compression of video composed of 3 views, MV-HEVC is reported to provide roughly a 30% bitrate reduction compared to simulcast encoding, and 3D-HEVC reduces the bitrate by additional 20% [Tech'16]. However, this applies only when the encoded views are coplanar and arranged on a 1D line. For other view arrangements, the compression efficiency of the multiview extensions of HEVC decreases significantly, which is a major limitation of those techniques. In many modern applications, such as free-viewpoint television or immersive video, the 1D linear camera arrangement does not provide enough information about the scene to create a realistic user experience. On the other hand, non-linear camera arrangements are much more challenging to set up, and, as mentioned before, current compression techniques are not prepared to efficiently compress video data produced by them. **In this dissertation, one of the goals is to improve the rate-distortion compression efficiency for circular camera arrangement.** It is a special case of a non-linear arrangement where the cameras are distributed precisely on a circle with their optical axes directed towards its center.

Another drawback of the state-of-the-art multiview video coding techniques is that they are built on top of monoscopic (single-layer) video codecs, as special profiles dedicated exclusively to the purpose of multiview video compression. Unfortunately, such multi-layer design requires much more complex decoders, which limits their practical applications. Moreover, the development and standardization of a multiview profile require additional work, and therefore such a profile is usually prepared after the standardization of a monoscopic codec. An alternative approach to multiview video compression is to pack several views of a multiview sequence into a single frame and use a standard monoscopic video encoder for compression, with additional signalization to inform the decoder how to separate the views. Such a solution is called frame-compatible coding and is commonly used e.g. in stereoscopic video broadcasting (3D television) [Vetro'10]. The advantage of frame-compatible coding is that the decoding device does not have to be equipped with a codec supporting a dedicated multiview profile, and therefore the 3D television broadcasting signal is more accessible to the customers. On the other hand, a monoscopic video encoder is not able to utilize the similarities between the views packed into a single frame, meaning that there is no inter-view prediction. As a result, the bitstream in such cases is higher than when using dedicated multiview video encoders and comparable to simulcast encoding [Samelak'17A]. **In this dissertation, the author proposes a novel approach that introduces inter-view prediction to frame-compatible multiview video coding by the use of Screen Content Coding** [Xu'16A]. Such a combination is unusual because SCC was developed for entirely different purposes. Moreover, the idea is also applied to the compression of immersive video.

1.2. GOALS AND THESES OF THE DISSERTATION

There are two main goals of this dissertation. The first one is to improve the compression efficiency of rectified multiview video acquired by cameras located on a circle. In this topic, the rectification process is also proposed by the author. The second goal is to prove that Screen Content Coding, which is an extension of HEVC designed for compression of video containing a significant amount of rendered graphics, can be surprisingly used for compression of multiview video.

The theses of this dissertation are as follows:

1. **It is possible to reduce both bitrate and encoding time of 3D-HEVC encoder in compression of rectified multiview video acquired with cameras located on a circle, compared to the state-of-the-art 3D-HEVC encoder, through adaptation of inter-view prediction to circular camera arrangements.**
2. **It is possible to use standard-compliant HEVC Screen Content Coding for compression of stereoscopic video, frame-compatible multiview video, and immersive video. With additional improvements, the rate-distortion compression efficiency of such an approach can be comparable or even higher than the state-of-the-art dedicated techniques.**

In the dissertation, the author proposes novel, original approaches to multiview video compression. The proposed solutions are implemented on top of publicly-available implementations of HEVC encoder and its extensions: Multiview HEVC (MV-HEVC), 3D-HEVC, and Screen Content Coding (SCC). The proposed methods are evaluated experimentally through the compression of test sequences and comparing results with unmodified encoders. The adaptation of 3D-HEVC to circular camera arrangements is additionally compared with 3D-HEVC adapted to any camera arrangement, which is another proposal co-authored by the author of this dissertation [Domański'15A, Domański'16A, Samelak'16, Stankowski'15, Stankiewicz'18].

The author's idea of using Screen Content Coding as an alternative to the state-of-the-art multiview, stereoscopic, and immersive video compression techniques was published in [Samelak'17A-D, Samelak'20A-B, Samelak'21A-B] and [Samelak'22]. The paper presenting the idea of optimizing 3D-HEVC for circular camera arrangements is to be published.

1.3. OVERVIEW OF THE DISSERTATION

This dissertation is divided into seven chapters. Chapter 1 provides information about the scope of the dissertation, as well as its goals and theses.

Chapter 2 contains selected topics in the state-of-the-art of video coding. First, the author briefly describes the principles of video compression. Then, the author presents popular techniques for multiview, 3D, and immersive video coding, which are the dedicated solutions for compression of video acquired by multiple cameras. Finally, the author provides information about Screen Content Coding, which was designed for different applications, but in this dissertation, the author proposes to reuse it also for multiview video compression.

In Chapter 3, the methodology of conducted experiments is presented. The author provides information about test models of video encoders used for the assessment of the proposals. Moreover, the author lists test sequences and test conditions for conducted experiments. The author describes the methods used to evaluate the compression efficiency and the encoding times of the proposed solutions.

Chapter 4 presents the author's idea for adapting 3D-HEVC to circular camera arrangement. The author shows how sequences acquired with cameras located roughly on a circle can be rectified. Next, the author derives the equations for inter-view prediction between views located on a circle. Then, the test model for 3D-HEVC is adapted to make use of the new equations. Finally, the proposal is experimentally evaluated.

In Chapter 5, the author proposes new applications for Screen Content Coding. After explaining the idea, the author shows how to apply SCC to stereoscopic, multiview, and immersive video coding. The goal of this work is to keep the codec standard-compliant, meaning that the bitstream produced by the encoder could be decoded by the state-of-the-art HEVC SCC decoder. The resulting encoders are compared experimentally with the state-of-the-art solutions.

In Chapter 6, the author introduces several customized improvements to the codecs from Chapter 5 to make them more efficient for the new applications. The description of the author's improvements is followed by more experimental results.

In Chapter 7, the author comments on his achievements and summarizes the results with regard to the goals of the dissertation.

2. SELECTED TOPICS IN THE STATE-OF-THE-ART OF VIDEO CODING

2.1. INTRODUCTION

In this chapter, the author briefly describes some aspects of the state-of-the-art of video coding. The main focus is put on multiview, 3D, and immersive video compression. The author also describes selected techniques of Screen Content Coding related to the author's proposal of using SCC for multiview video compression.

As mentioned in Chapter 1, the author's research is based on HEVC and its extensions, nonetheless it should be noted that multiview profiles are also present in previous generations of video coding standards, i.e., multiview profile in MPEG-2 [Chen'97, Ohm'99, ISO'12] and Multiview Video Coding (multiview extension of AVC) [ISO'14, Vetro'11A]. Regardless of the standard, the principle in multiview video compression is to exploit the similarities between the coded views.

2.2. MULTIVIEW AND 3D-VIDEO CODING

Multiview HEVC (MV-HEVC) and 3D-HEVC extensions were developed by the ITU-T/ISO/IEC Joint Collaborative Team on 3D Video Coding Extension Development (JCT-3V) and included in the second and third edition of the HEVC standard [ISO'21, Sullivan'12, Sullivan'13], respectively. They introduce a multilayer coding design, i.e., joint coding of multiple views. Compared to simulcast coding, the main benefit of such design is the encoder's ability to perform inter-view prediction that exploits the similarities between the views. In MV-HEVC and 3D-HEVC, inter-view prediction is performed through block-based matching between the picture being encoded and the previously encoded reference picture of another view and within the same time instance. This approach is analogous to the inter-frame prediction, where the reference picture belongs to the same view but a different time instance.

One of the coding tools that utilize inter-view prediction is Disparity Compensated Prediction (DCP), which estimates disparity vectors and calculates prediction residuals [Müller'13]. A disparity vector points to the best-matching block of samples in the reference view. In 3D-HEVC, disparity vector (d_x, d_y) is also predicted by the use of depth maps, according to Equation (2.1):

$$d_x = \frac{f_{x1}(t_2 - t_1)}{z} - o_{x2} + o_{x1} \quad , \qquad (2.1)$$
$$d_y = 0$$

where:

d_x, d_y – horizontal and vertical component of disparity vector,

f_{x1} – horizontal focal length of the source camera,

t_1, t_2 – positions of the source and target camera along the horizontal axis,

o_{x1}, o_{x2} – horizontal optical centers of the source and target camera,

z – distance between the image plane and acquired point in 3D space, i.e., depth.

In Equation (2.1), all the components except for z are camera parameters, which are assumed to be constant for a given pair of source and target cameras. The distance z is calculated from the depth map sample value using the following Formula (2.2):

$$z = \left(\frac{d}{2^b}\left(\frac{1}{Z_{near}} - \frac{1}{Z_{far}}\right) + \frac{1}{Z_{far}}\right)^{-1}, \qquad (2.2)$$

where:

Z_{near}, Z_{far} – the smallest and the biggest value of z,

b – bit depth of sample values of the depth map,

d – normalized disparity (sample value in depth map).

To avoid confusion, it should be stressed that depth maps are usually represented as greyscale images, where bigger sample values indicate lower depth (a point is closer to the image plane), and sample values closer to 0 indicate greater depth (Figure 1.1). Therefore, the horizontal component of disparity vector (d_x) in MV-HEVC and 3D-HEVC is inversely proportional to depth z and proportional to the value of depth map sample d (i.e., normalized disparity). The bigger the value of the depth map sample (the object is closer to the camera), the bigger the horizontal disparity. Regarding the vertical component of the disparity vector, 3D-HEVC assumes it is always equal to 0, i.e., the views are vertically aligned. On the one hand, 3D-HEVC encoder benefits from those assumptions through significant simplification of the disparity-compensated prediction (DCP). For example, for each pair of cameras, 3D-HEVC prepares dedicated look-up tables that map each possible depth sample value onto disparity; for a typical case of 8-bit depth, such a table contains 256 values. This allows the encoder to significantly reduce the time required to perform inter-view prediction. On the other hand, the aforementioned assumptions limit the use of 3D-HEVC to multiview video acquired by cameras located on a line, aligned vertically and with optical axes in parallel. Even though a great deal of effort was put into building multi-camera systems that would meet those requirements, it is not possible to position the cameras ideally. Therefore, before compression, multiview video is usually rectified, which corresponds to correcting the positions of cameras and suppressing the results of differences in their properties (Figure 2.1) [Hartley'99, Kang'08, Stankowski'10].

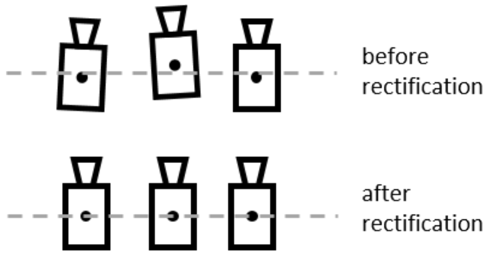

Figure 2.1. Linear camera setup before and after rectification.

It should be stressed that rectification does not correct the real positions of cameras but transforms the views obtained from the cameras into video virtually obtained from an ideally positioned set of cameras [Ho'12, Choi'12]. This improves the accuracy of the inter-view prediction, thus also the overall compression efficiency.

2.3. Screen Content Coding

Screen Content Coding (SCC) is an extension added to the fourth version of the HEVC standard [ISO'21, Xu'16A]. Contrary to the main profile of HEVC dedicated to camera-captured content, the SCC was developed to improve the compression efficiency of video containing a significant portion of rendered graphics, text, or animations (Figure 2.2).

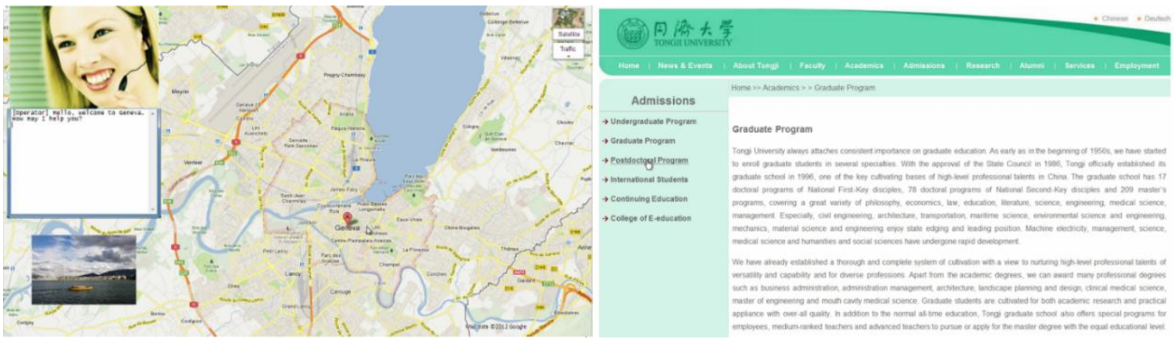

Figure 2.2. Examples of frames from test sequences containing screen content:
sc_map (left) and sc_web_browsing (right).

Applications of SCC include but are not limited to: screen sharing, wireless displays, remote desktop access, streaming e-sports, or cloud gaming. Screen Content Coding profile improves the compression efficiency of computer-generated content by leveraging its characteristics, such as the presence of repeatable patterns, sharp edges, plain monochromatic areas, limited palette of colors, or lack of noise. It is done through a set of dedicated coding tools, from which the most important ones are listed below.

Intra Block Copy (IBC) [Xu'16B], also known as Current Picture Referencing [Xu'19], predicts the content by searching for the most similar block of samples in the currently processed picture. The position of the best matching block of samples is indicated by the IBC vector, similarly to the standard inter-frame prediction. However, since the reference picture is also the current picture, the search area for IBC is restricted to the previously encoded part of the picture (Figure 2.3). Moreover, the search in IBC is performed at full-pel accuracy, contrary to the standard sub-pel inter-frame prediction. Intra Block Copy is a very effective technique for compression of fonts and other repetitive patterns.

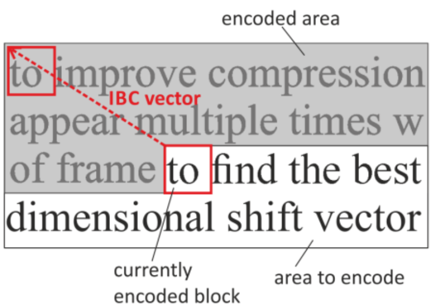

Figure 2.3. Example of using Intra Block Copy.

Palette Mode [Pu'16, Sun'19] – screen content is often composed of a small number of distinct colors, especially at the level of a single coding unit (CU). SCC encoder can decide to enumerate them, prepare palettes and transmit indices to colors from palettes rather than applying standard prediction.

Adaptive Motion Vector Resolution [Xu'16A] allows for dynamic modification of the accuracy of motion-compensated prediction. In the compression of screen content, it is often beneficial to represent motion vectors at full-pel resolution because the cost of spending more bits on sub-pel motion vectors can exceed the profit of more accurate motion estimation.

Adaptive Color Transform [Zhang'15, Jhu'20] allows the encoder to choose between RGB and YCbCr color space for each CU.

It should be noted that Screen Content Coding is not related to multiview video coding, especially when the multiview video is captured with cameras. Nevertheless, in the dissertation, the author proposes the usage of SCC as a novel approach to the compression of camera-captured multiview video and immersive video. Chapter 5 describes the process of adaptation of the unmodified SCC to the new application, whereas in Chapter 6 the author presents how SCC can be modified to achieve rate-distortion compression efficiency similar to that of the dedicated multiview compression techniques. It should also be mentioned that Screen Content Coding was included in the most recent video coding standard – Versatile Video Coding [Xu'22] – and therefore, the author's propositions remain valid.

2.4. IMMERSIVE VIDEO CODING

In recent years, virtual reality (VR), augmented reality (AR), and mixed reality systems are gaining importance. They provide a new type of experience where the user is immersed in the scene and, depending on the number of degrees of freedom provided by the system, can control the viewing position and orientation of the content [Domański'17]. Video content processed by such systems is often referred to as immersive video. It may be computer-generated or captured by a set of cameras, planar (2D), or omnidirectional (360-degree). Immersive video is usually represented as Multiview Video plus Depth (MVD), however other representations, such as point clouds, are also being researched [Hu'23, Li'20, Schwarz'19, Cui'19]. Nevertheless, in the dissertation, the author focuses only on the MVD representation.

One of the critical factors that influence the overall quality of a VR system is the quality of virtual views presented to the user. Therefore, in recent years, a lot of effort has been put into improving virtual view synthesis [Bonatto'21, Ceulemans'18, Fachada'18, Rahaman'18, Dziembowski'17A]. It has been proven that the efficiency of view synthesis strongly depends on the quality of depth maps, hence the research on depth map estimation also attains a lot of interest [Mieloch'21, Rogge'19, Dziembowski'17B, Mieloch'17A, Mieloch'17B, Mieloch'17C]. Another technical challenge for the development of immersive video is the amount of data required to fully represent a scene. On the one hand, the more input views are used, the better the quality of the virtual view synthesis can be achieved. On the other hand, such a vast amount of data may be impractical for transmission and storage, even when compressed using modern multiview video encoders. Therefore, in order to cope with the task of efficient representation of multiview video data, a new type of video codec called MPEG Immersive Video (MIV) was recently developed by the ISO/IEC MPEG group [Boyce'21].

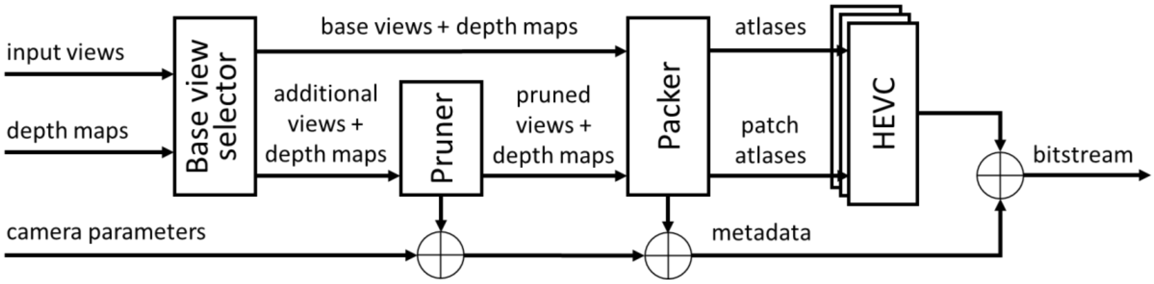

Figure 2.4. Block diagram of MIV encoder.

The goal of a MIV encoder is to reduce the spatial redundancy from the input multiview video, which is composed of multiple input views, corresponding depth maps, and parameters of corresponding cameras (Figure 2.4). In the first step, the encoder chooses a subset of input views (base views) and corresponding depth maps. The base views and their depth maps are merged into frame-compatible structures called atlases. Then, using view synthesis, the areas of additional views that can be synthesized from base views are removed in a process called pruning. The remaining parts (patches), i.e., areas occluded in the base views and present only in the additional views, are packed into several views called patch atlases (Figure 2.5). Pruning also applies to depth maps. An example of the output data after MIV pruning and packing, composed of atlases of base views and patches, is presented in Figure 2.6.

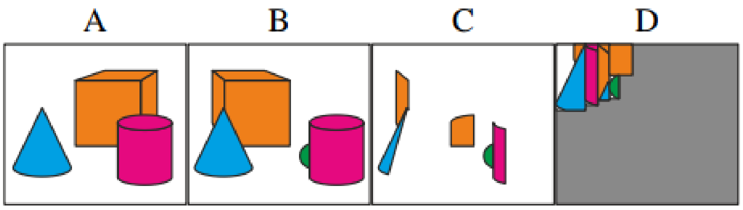

Figure 2.5. The process of preparing a patch atlas.
A – base view, B – additional view, C – pruned additional view, D – patch atlas.

Apart from atlases containing video data and depth maps, a MIV encoder outputs metadata about pruning and packing. This allows MIV decoder to properly unpack and position the patches in the reconstructed views. The metadata is put directly into the bitstream, while the output video and depth atlases are compressed using simulcast HEVC encoders. The choice of HEVC as the internal codec was motivated by HEVC being commonly used in products. Nevertheless, MIV provides flexibility in choosing the video codec for compression of atlases to ensure the possibility of replacing it with VVC and other video coding techniques [Boyce'21].

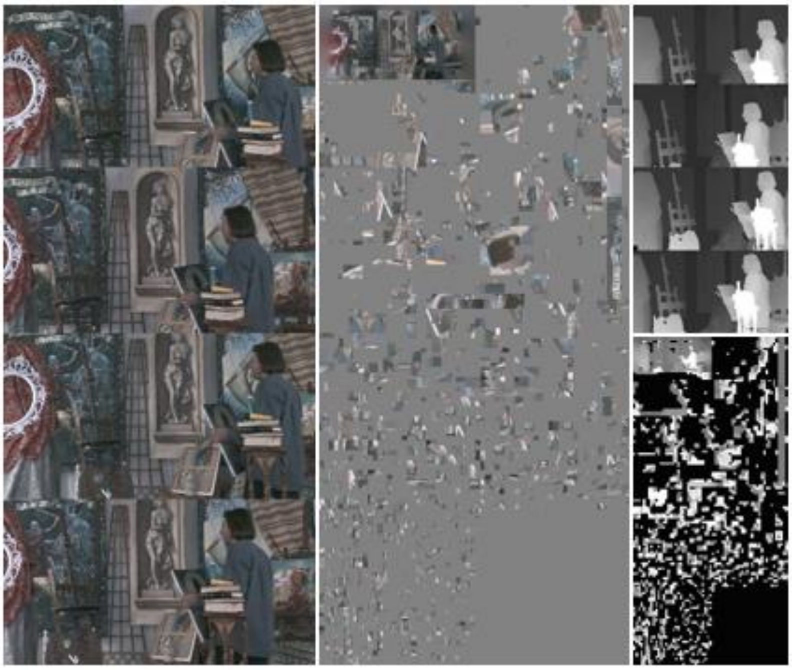

Figure 2.6. Example of MIV output atlases: atlas with packed input base views (left), patch atlas for views (middle), atlas with packed depth maps of corresponding base views (top-right), patch atlas for depth maps (bottom-right).

3. METHODOLOGY OF THE EXPERIMENTS

3.1. INTRODUCTION

In the dissertation, the author presents several proposals for inter-view prediction and multiview video compression, dedicated to various applications. The author implemented the proposed techniques on top of appropriate test models and then evaluated them experimentally through the compression of multiple commonly used test sequences. The results were compared to the state-of-the-art solutions.

In this chapter, the common part of the methodology of performed experiments is described, which includes information about test models and their configurations, test sequences, and metrics used to compare the results. Details that are specific to a given evaluation are described separately in sections dedicated to each of the experiments.

3.2. VIDEO CODEC TEST MODELS

In the previous chapter (Chapter 2), multiple HEVC-based video codecs are described: HEVC Screen Content Coding, Multiview HEVC, 3D-HEVC, and MPEG Immersive Video with HEVC as the inner codec. Each of these codecs is provided with the respective test model, which is a publicly available software implementation commonly used as a reference in experimental research. Versions of test models for each video codec used in the author's research are presented in Table 3.1. All the test models are based on the same HEVC core, HM-16.9. This assures fair comparison in the evaluation because the results are not affected by the changes between different versions of test models. Moreover, each codec is provided with a dedicated Common Test Conditions (CTC) – a document that specifies, e.g., default configurations of test model software to be used in the experiments. Those recommendations are meant to ensure reliable comparison between tested codecs. In the conducted experiments, unless stated otherwise, the configuration of the codecs follows appropriate CTC.

In some of the conducted experiments, the author considers two coding scenarios defined in CTCs: All Intra and Random Access. All Intra disables temporal prediction, i.e., all the frames are coded as I-frames [Sullivan'12].

Table 3.1. Test Models used in the experiments.

Video coding technique	**Test Model**	**CTC**
HEVC Main	HM-16.9 [HM]	[Bossen'13]
HEVC Screen Content Coding	HM-16.9+SCM-8.0 [HM+SCM]	[Yu'15]
Multiview HEVC	HTM-16.2 [HTM]	[Müller'14]
3D-HEVC	HTM-16.2 [HTM]	[Müller'14]
MPEG Immersive Video	TMIV v.8 [TMIV] + HM-16.9 [HM]	[CTC MIV]

The compression techniques proposed in the dissertation are implemented by the author on top of the aforementioned test models. It should be noted that software implementation of a test model for a video codec is very complex, e.g., the source code of HTM (test model for 3D-HEVC) contains roughly 120 thousand lines of code in C++. Such source code of the test model is called the reference software. It is a result of the collaboration of many research teams and software engineers, and it requires many man-years of work to develop such an advanced video codec. As the evolution of video codecs progresses, achieving significant improvement in compression efficiency becomes more difficult and usually is attained at the cost of increased complexity and longer encoding times.

3.3. Methodology of compression efficiency assessment

This section is dedicated to the description of the methodology used for the evaluation of the author's ideas for improving the compression efficiency of multiview video. The term "rate-distortion compression efficiency" (or simply "compression efficiency") used in the dissertation should be understood as the relation between the quality of the reconstructed video (after compression and decompression of the original video) and the bitrate of the compressed video stream produced by the encoder. Better compression efficiency indicates that the tested codec, compared to the reference one, provides better quality of reconstructed video at a given bitrate or lower bitrate at a given quality level.

In the dissertation, the quality is usually measured as PSNR (Peek Signal-to-Noise Ratio) of luma samples between original and reconstructed video, which is a common approach in the assessment of lossy compression [Wang'04] and is calculated according to Equation (3.1):

$$PSNR = 10\log_{10}\left(\frac{(2^b - 1)^2}{MSE}\right) \ , \tag{3.1}$$

where:

b – bit depth of sample values of the depth map,

MSE – mean squared error, calculated according to Equation (3.2):

$$MSE = \frac{1}{H \cdot W}\sum_{y=0}^{H-1}\sum_{x=0}^{W-1}[V_D(x,y) - V_R(x,y)]^2 \ , \tag{3.2}$$

where:

W, H – width and height of the coded view,

$V_D(x,y), V_R(x,y)$ – luma sample value at position (x,y) in the decoded and reference (before compression) views, respectively.

It should be stressed that PSNR (expressed in dB), when used as a quality metric of video compression techniques, is usually calculated only for the luma component of the input views, ignoring chroma components as well as depth maps. In the dissertation, unless stated otherwise, the author follows the commonly used approach for quality assessment.

As mentioned in Section 2.4, in immersive video applications, the overall quality of a system mostly depends on the quality of synthesized virtual views rather than decoded views. Some artifacts of view synthesis that do not influence the subjective quality, such as slightly shifted edges or minor changes in luminance sample values, may cause a significant decrease in PSNR. Other artifacts may not impact PSNR, but they significantly deteriorate virtual view quality when assessed by end users. Therefore, in the experiments related to immersive video (Section 5.7.4 and Section 6.4.4), the quality is additionally measured for the virtual views using other metrics:

- WS-PSNR – Weighted-to-Spherically-Uniform PSNR [Sun'17],
- VIF – Visual Information Fidelity [Sheikh'06],
- VMAF – Video Multimethod Assessment Fusion [Li'16],
- MS-SSIM – Multi-Scale SSIM [Wang'04],
- IV-PSNR – ISO/IEC MPEG metric for immersive video [Dziembowski'22, ISO'19].

The abovementioned alternative metrics are more resilient to view synthesis artifacts and, therefore, better correlated with the subjective quality assessment.

According to Common Test Conditions, the evaluation of video encoders should be performed at multiple rate points. The parameter that is responsible for controlling the bitrate of a modern video encoder, such as AVC, HEVC (including MV-HEVC and 3D-HEVC), or VVC, is called Quantization Parameter (QP) [Ma'05, Sullivan'12, Bross'21]. Encoding at small values of QP results in a higher bitrate but better quality. Higher QP values reduce the bitrate at the cost of lower quality of reconstructed video and visible compression artifacts. The experiments presented in the dissertation are performed at 4 different values of QP (defined separately in the description of each experiment). If an experiment includes compression of depth map images, their QP sets may differ from the values used for input views. Experimental compression results, i.e., 4 bitrate and quality pairs for each tested encoder, can be used to interpolate the rate-distortion curves using a third-order polynomial. Averaged bitrate reduction at a given quality (alternatively, averaged quality improvement at a given bitrate) is often calculated by MPEG research teams using the Bjøntegaard metric [Bjøntegaard'01]. That metric is also used in the experiments conducted by the author for the compression efficiency assessment.

In the experiments, the author of the dissertation also compares the encoding time. It is calculated as the mean coding time of 4 compression cycles, one per each QP. In order to ensure a fair comparison of the results, all the processing is performed on the same desktop computer equipped with an Intel Core i7 3.4 GHz CPU unit, 64 GB RAM, and Microsoft Windows 10 Pro operating system.

3.4. TEST SEQUENCES

In the evaluation of compression techniques proposed in the dissertation, multiple test sequences are used. The test set contains the following types of sequences:

- Multiview sequences acquired with dense linear camera setups (Table 3.2). The cameras are coplanar, located on a line with their optical axes in parallel. The provided content is already rectified to an ideal linear camera arrangement (Figure 2.1).
- Multiview sequences acquired with nearly-circular camera setups (Table 3.3). The cameras are located roughly in a circle. Acquired sequences are not rectified.
- Multiview sequences acquired with two-dimensional linear camera setups – cameras are coplanar and distributed in the form of two-dimensional matrices (Table 3.4)
- Multiview sequences acquired with omnidirectional cameras (Table 3.5). Such sequences are dedicated to immersive video applications.

- Single-view sequences containing screen content (Table 3.6). These are recommended for the assessment of compression techniques related to screen content.

Table 3.2. Multiview sequences acquired with dense linear camera setups.

Test sequence name	**Luma frame size**	**Frame rate**	**Content type**	**Sequence source**
Balloons	1024×768	30 fps	NC	[Tanimoto'09]
Kendo	1024×768	30 fps	NC	[Tanimoto'09]
Newspaper	1024×768	30 fps	NC	[Ho'08]
Poznan_Hall2	1920×1088	25 fps	NC	[Domański'09]
Poznan_Street	1920×1088	25 fps	NC	[Domański'09]
Poznan_Carpark	1920×1088	25 fps	NC	[Domański'09]
IntelFrog	1920×1080	30 fps	NC	[Salahieh'19]
GT_Fly	1920×1088	25 fps	CG	[Zhang'11]
Dancer	1920×1080	25 fps	CG	[Rusanovskyy'11]
Shark	1920×1080	60 fps	CG	[Senoh'14]

Table 3.3. multiview sequences acquired with nearly-circular camera setups.

Test sequence name	**Luma frame size**	**Frame rate**	**Content type**	**Sequence source**
Ballet	1024×768	25 fps	NC	[Zitnick'04]
Breakdancers	1024×768	25 fps	NC	[Zitnick'04]
BBB_Flowers	1280×768	24 fps	CG	[Kovacs'15]
Poznan_Blocks	1920×1080	25 fps	NC	[Wegner'14]
BBB_Butterfly	1280×768	24 fps	CG	[Kovacs'15]
Poznan_Fencing2	1920×1080	25 fps	NC	[Domański'16B]

Table 3.4. Multiview sequences acquired with omnidirectional cameras.

Test sequence name	**Luma frame size**	**Frame rate**	**Content type**	**Sequence source**
ClassroomVideo	4096×2048	30 fps	CG	[Kroon'18]
TechnicolorMuseum	2048×2048	30 fps	CG	[Doré'18]
TechnicolorHijack	4096×4096	30 fps	CG	[Doré'18]
Chess	2048×2048	30 fps	CG	[Ilola'19]
Group	1920×1080	30 fps	CG	[Doré'20B]
ChessPieces	1920×1080	30 fps	NC	[Guillo'18]

Table 3.5. Multiview sequences acquired with cameras located on a 2D matrix.

Test sequence name	**Luma frame size**	**Frame rate**	**Content type**	**Sequence source**
OrangeKitchen	1920×1080	30 fps	CG	[Boissonade'18]
TechnicolorPainter	2048×1088	30 fps	NC	[Doyen'17]
Fan	1920×1080	30 fps	CG	[Doré'20A]
Mirror	1920×1080	30 fps	CG	[Doré'21]

Table 3.6. Single-view sequences containing screen content.

Test sequence name	Luma frame size	Frame rate	Sequence source
Basketball_Screen	2560×1440	60 fps	[Suzuki'14]
ChinaSpeed	1024×768	30 fps	
ChineseEditing	1920×1080	60 fps	
MissionControlClip2	2560×1440	60 fps	
MissionControlClip3	1920×1080	60 fps	
sc_console	1920×1080	60 fps	
sc_desktop	1920×1080	60 fps	
sc_flyingGraphics	1920×1080	60 fps	
sc_map	1280×720	60 fps	
sc_programming	1280×720	60 fps	
sc_robot	1280×720	30 fps	
sc_web_browsing	1280×720	30 fps	
SlideShow	1280×720	20 fps	

All of the sequences are publicly available and recommended by the appropriate common test conditions for the evaluation of video coding techniques. They all have the same bit depth of samples, equal to 8 bits, and the chroma format is 4:2:0. It is assumed that geometric distortions of camera lenses are corrected, i.e., the cameras are calibrated [Collins'99, Lucchese'03].

The content of used test sequences is very diverse to ensure that the results of experiments reflect the compression efficiency of video encoders when used in practice. Test sequences can be divided by content into two groups: computer-generated (CG) and natural content (NC). Examples of frames from test sequences are presented in Figures 3.1–3.5.

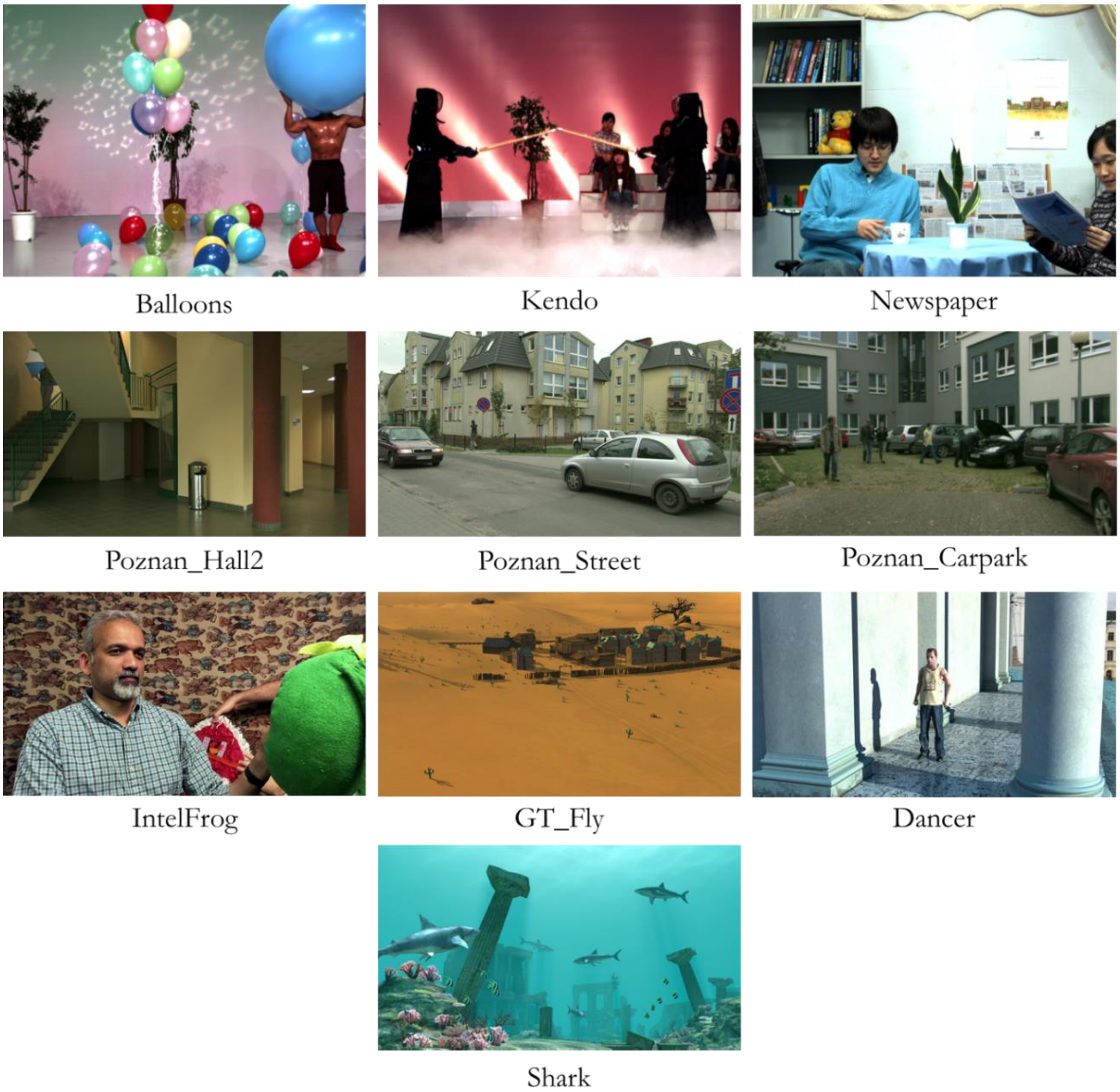

Figure 3.1. Examples of frames from multiview test sequences acquired with dense linear camera setups.

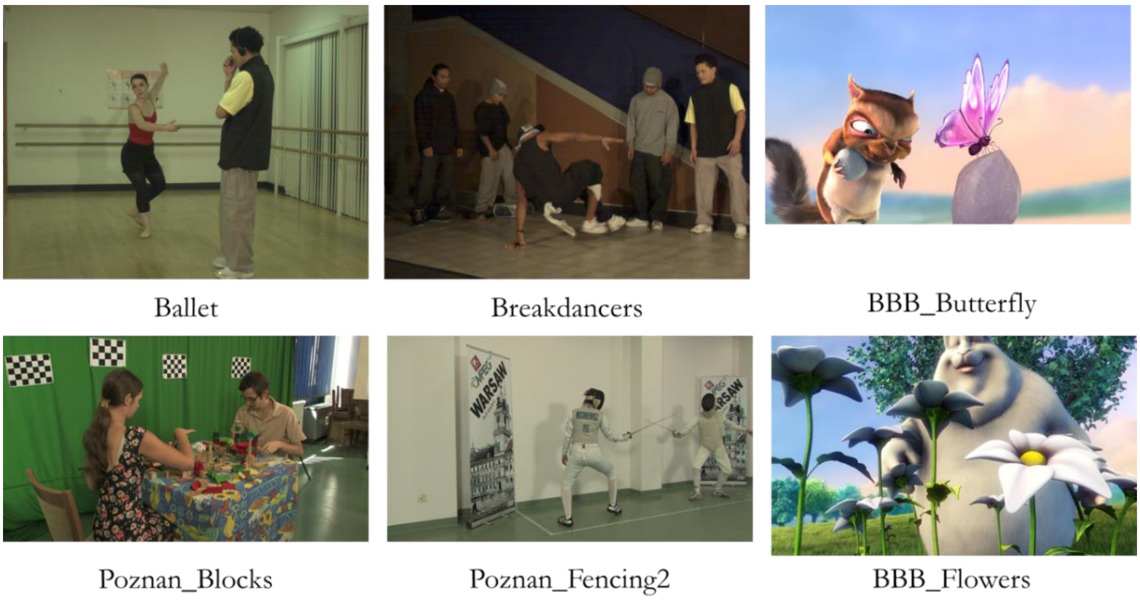

Figure 3.2. Examples of frames from multiview test sequences acquired with nearly-circular camera setups.

Figure 3.3. Examples of frames from multiview test sequences acquired with omnidirectional cameras.

Figure 3.4. Examples of frames from multiview test sequences acquired with rectangular camera setups.

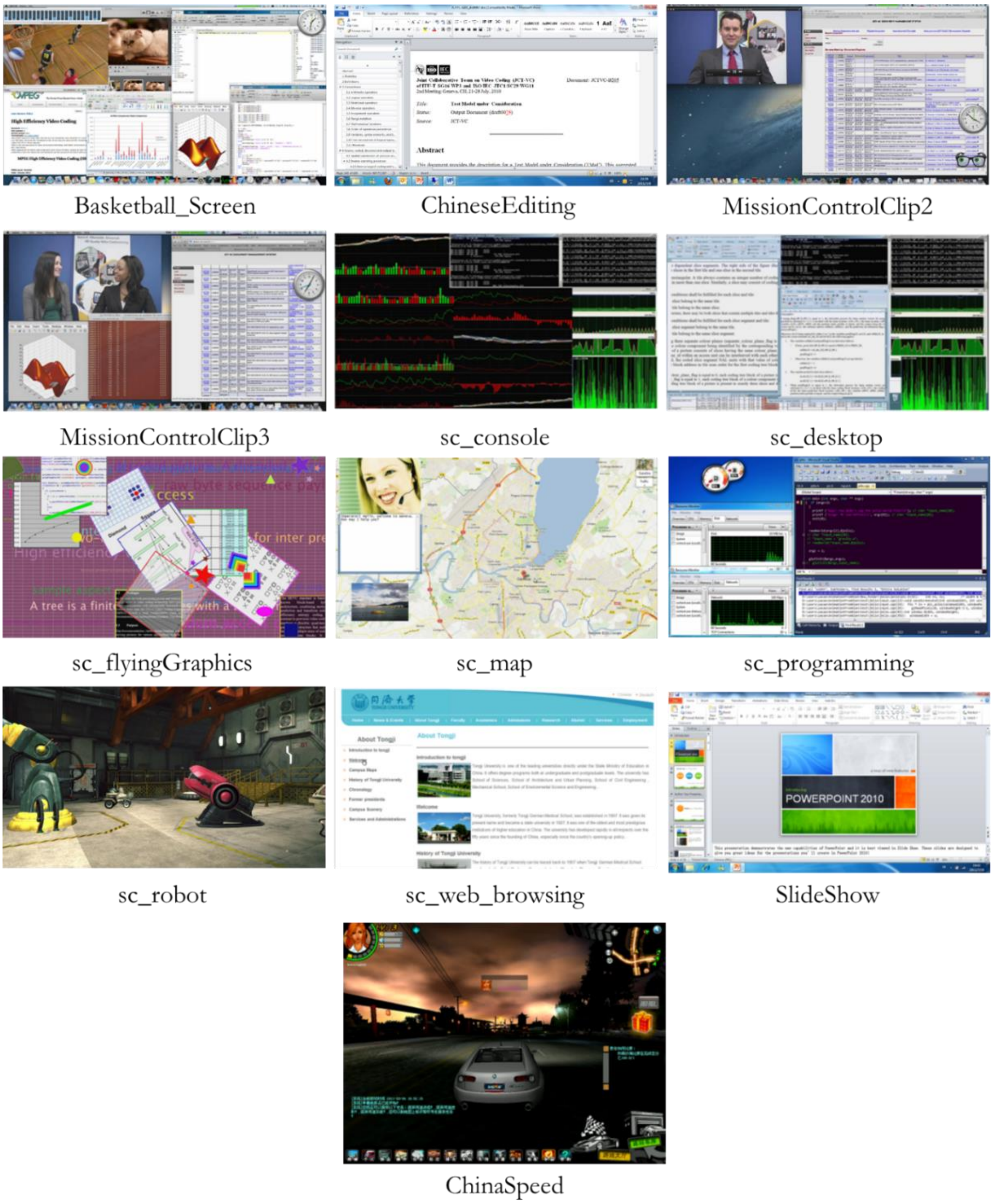

Figure 3.5. Examples of frames from single-view sequences containing screen content.

4. ADAPTING 3D-HEVC TO CAMERA ARRANGEMENTS OTHER THAN LINEAR

4.1. INTRODUCTION

As mentioned in Section 2.2, both MV-HEVC and 3D-HEVC assume that the input video is captured by multiple cameras distributed densely on a line. Such camera arrangement is used mainly for the purpose of displaying multiview content on autostereoscopic displays, which allow users to watch a stereoscopic video without wearing special glasses or other headgear, and even shift the viewer's position horizontally in a limited range. This technology, however, still has not become popular. On the contrary, other multiview applications such as free-viewpoint television (FTV), virtual reality, or immersive video systems are recently gaining more attention. Unfortunately, the aforementioned applications require video content that is captured from different perspectives using sparsely distributed cameras. Even though it is possible to use 3D-HEVC for compression of multiview video obtained from camera arrangements other than linear, the compression gain is significantly smaller than for linear camera arrangements [Stankowski'15, Samelak'16, Tech'16] due to inaccurate inter-view prediction. For that reason, 3D-HEVC and other state-of-the-art multiview profiles are not suitable for modern multiview applications.

The aforementioned problem was addressed by adapting the inter-view prediction of 3D-HEVC to arbitrary camera arrangements by the use of point mapping in 3D space. **The author of the dissertation participated in this research, implementation, and evaluation of the proposed modifications in the reference software, and co-authored several papers on that topic** [Stankowski'15, Domański'16A, Samelak'16, Stankiewicz'18]. A detailed description of this technique is presented in Section 4.2. In the dissertation, this solution is referred to as **ANY-HEVC.**

The main drawback of 3D point mapping is that it significantly increases the complexity of the inter-view prediction compared to standard 3D-HEVC. For instance, it is not possible to prepare look-up tables mapping depth to disparity (as described in Section 2.2) because the position of corresponding points in two views depends on multiple variables. Therefore, even though the proposed technique improves the compression efficiency in the case of camera arrangements other than linear, it also significantly increases the encoding time.

Among other multiview camera arrangements, much attention is paid to 2D arrays of cameras and circular arrangements [Tanimoto'12, Domański'15C, Cserkaszky'18]. In the case of circular arrangements, the cameras surround a scene and acquire it from different perspectives. The advantage of such an approach is that a few sparsely distributed cameras allow capturing the scene from various angles, which is an essential feature in the context of, e.g., FTV or immersive video applications [Domański'15A, Boyce'21].

In this chapter, the author of the dissertation presents **a novel idea of adapting 3D-HEVC for the compression of multiview video captured with cameras located roughly on a circle** and rectified to the views located ideally on a circle with optical axes directed towards its center. Such an approach is dual to 3D-HEVC, where video rectification and the assumption of linear camera arrangement allow to simplify inter-view prediction of the codec. The proposal, in the dissertation called **ARC-HEVC**, is a third way of dealing with the compression of MVD, and it should be considered as a trade-off between simple but with limited applications standard 3D-HEVC and flexible but more complex ANY-HEVC (Figure 4.1).

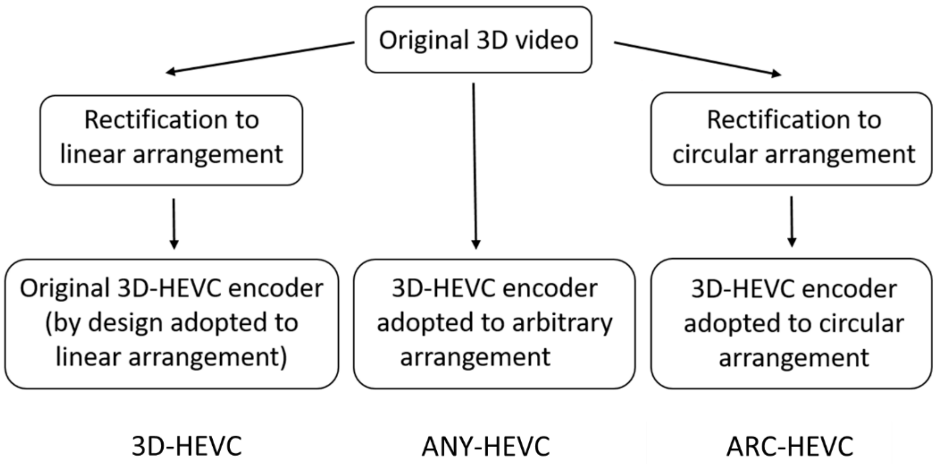

Figure 4.1. Three approaches to compression of 3D video: state-of-the-art 3D-HEVC dedicated for linear camera arrangements (3D-HEVC), modified 3D-HEVC for compression of arbitrary camera arrangements (ANY-HEVC, Sections 4.2-4.3), and the proposed codec for circularly rectified views (ARC-HEVC, Sections 4.4-4.6).

Along with the new codec, the author proposes **a novel procedure for circular rectification of a multiview video acquired by nearly-circular camera arrangements**. This rectification process and the adaptation of 3D-HEVC inter-view prediction are presented in Sections 4.4 and 4.5, respectively.

4.2. 3D-HEVC ADAPTED TO ANY CAMERA ARRANGEMENT (ANY-HEVC)

4.2.1. INTER-VIEW PREDICTION FOR ANY CAMERA ARRANGEMENT

This section describes the modified inter-view prediction by the use of 3D point mapping, which is the core of the ANY-HEVC codec. It is assumed that all camera parameters are known and represented by intrinsic parameters gathered in matrix $\mathbb{K}$ (4.1), and extrinsic parameters, i.e., rotation matrix [3×3] $\mathbb{R}$ and 3-component translation vector [3×1] $\mathbb{T}$. The derivation of the camera parameters is out of the scope of the dissertation and can be found in [Hartley'03, Cyganek'09, LaValle'20].

$$\mathbb{K} = \begin{bmatrix} f_x & c & o_x \\ 0 & f_y & o_y \\ 0 & 0 & 1 \end{bmatrix}, \tag{4.1}$$

where:

f_x, f_y – focal lengths,

o_x, o_y – coordinates of the optical center,

c – skew factor.

The abovementioned intrinsic and extrinsic camera parameters can be used to calculate the projection matrix [4×4] $\mathbb{P}$ for each camera using Equation (4.2) [Hartley'03]:

$$\mathbb{P} = \begin{bmatrix} \mathbb{K} & 0 \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} \mathbb{R} & -\mathbb{R} \cdot \mathbb{T} \\ 0^T & 1 \end{bmatrix}. \tag{4.2}$$

Then, the positions of the corresponding points in two camera views (denoted by indices 1 and 2) can be derived according to Formula (4.3) [Hartley'03]

$$\begin{bmatrix} z_2 \cdot x_2 \\ z_2 \cdot y_2 \\ z_2 \\ 1 \end{bmatrix} = \mathbb{P}_2 \cdot \mathbb{P}_1^{-1} \begin{bmatrix} z_1 \cdot x_1 \\ z_1 \cdot y_1 \\ z_1 \\ 1 \end{bmatrix}, \tag{4.3}$$

where:

(x_1, y_1), (x_2, y_2) – positions of corresponding points in Views 1 and 2,

z_1, z_2 – distances between the acquired point in 3D space and the planes of Camera 1 and 2, respectively, calculated from depth map sample values using Formula (2.2),

$\mathbb{P}_1$, $\mathbb{P}_2$ – projection matrices for Views 1 and 2.

The Formula (4.3) allows to project any point (x_1, y_1, z_1) from the plane of Camera 1 directly onto the plane of Camera 2, resulting in point (x_2, y_2, z_2). This operation, in the dissertation referred to as 3D point mapping, is used in the ANY-HEVC codec instead of simple depth-to-disparity mapping from 3D-HEVC. The description of the modifications is presented in Sections 4.2.2 – 4.2.4, and their evaluation is presented in Section 4.3.

4.2.2. Modified Disparity Compensated Prediction

This section describes the modification of Disparity Compensated Prediction (DCP) in ANY-HEVC. As mentioned in Section 2.2, in 3D-HEVC, a disparity vector has only one (horizontal) component derived directly from camera parameters and a given depth map sample through Equation (2.1). In ANY-HEVC, modified DCP uses 3D point mapping to project the position (x_1, y_1) of the coded block of samples onto the reference view, resulting in a new position (x_2, y_2), as presented in Figure 4.2.

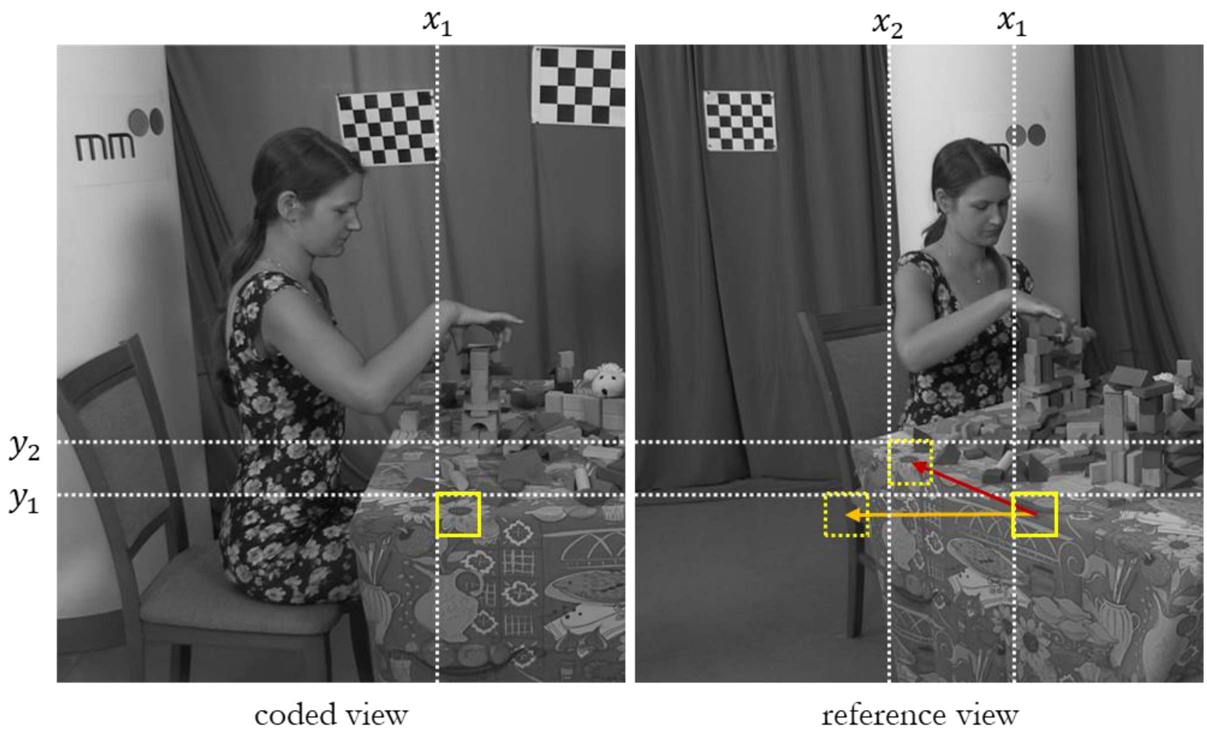

Figure 4.2. Visualization of disparity vectors in 3D-HEVC (orange) and ANY-HEVC (red) and the corresponding predicted blocks of samples (yellow dotted squares) in a multiview video acquired by nearly-circular camera arrangement, Poznan_Blocks sequence.

The difference in the position of the corresponding points in Views 1 and 2 is then used to find the 2-dimensional disparity vector (d_x, d_y) according to Equation (4.4):

$$\begin{aligned} d_x &= x_2 - x_1\,, \\ d_y &= y_2 - y_1\,. \end{aligned} \tag{4.4}$$

The example presented in Figure 4.2 shows that the unmodified DCP from 3D-HEVC can be highly inaccurate when applied to nearly-circular multiview video. First, it is caused by the lack of a vertical component of disparity vector. Secondly, the 3D-HEVC codec does not take into account camera rotation, so the length of the horizontal component of disparity vector can also be incorrect.

On the other hand, adding the vertical component of disparity vector, as proposed in ANY-HEVC, requires more bits to be transmitted in the bitstream. Therefore, to be beneficial, the more accurate DCP should compensate for the overhead caused by 2D disparity vectors. Naturally, this may decrease the rate-distortion compression efficiency when using ANY-HEVC to compress multiview video acquired by linear camera arrangements. The evaluation of such a case is presented in Section 4.3.

4.2.3. MODIFIED INTER-VIEW PREDICTION TOOLS

As a consequence of introducing in ANY-HEVC a non-zero vertical component of disparity vector, several inter-view prediction tools have to be adapted to make use of 2D disparity vectors.

One of the tools to be updated is Inter-View Motion Prediction. In 3D-HEVC, motion vectors can be predicted from the reference view, from the position of the collocated block shifted by disparity. In the proposal, the disparity is a two-dimensional vector, therefore the derivation of the position of the collocated block needs to be changed. Moreover, 3D-HEVC assumes that the cameras capturing the scene are densely distributed on a line with parallel optical axes, hence the motion appears to be the same in each view. However, when the cameras are distributed arbitrarily, the motion direction may be different for each view (Figure 4.3).

In order to correctly predict motion in the case of non-linear camera arrangements, motion vectors predicted from the reference view have to be distorted according to the change of perspective between the cameras. In the proposed solution, both points a_1 and $b_1$that indicate the motion vector are projected onto corresponding points a_2 and b_2 in the target view, and the difference $b_2 - a_2$ results in the new motion vector.

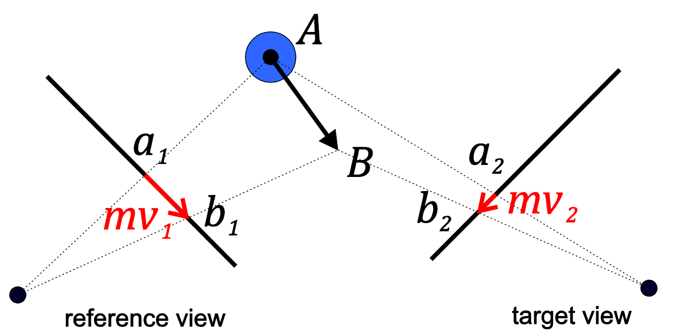

Figure 4.3. Different motion vectors representing the same motion in the reference and target views.

Other tools that are updated to make use of two-dimensional disparity vectors are View Synthesis Prediction, Neighboring Block Disparity Vector, Depth-oriented Neighboring Block Disparity Vector, and Illumination Compensation. In all of the aforementioned tools, the standard approach of mapping depth to disparity is replaced with the projection of depth sample value at a given position in the view, using the 3D point mapping presented in Section 4.2.1.

4.2.4. MODIFIED BITSTREAM SYNTAX

As mentioned in Section 4.2.1, inter-view prediction through 3D point mapping requires all camera parameters, and they also have to be available for the decoder. Therefore, ANY-HEVC encoder has to include them in the bitstream. However, floating point numbers have to be transmitted in the bitstream at finite precision. Therefore, each camera parameter would suffer from rounding errors that would accumulate when calculating projection matrices for 3D point mapping. To avoid that, ANY-HEVC codec transmits the components of projection matrices for each view instead of raw camera parameters [Samelak'16]. Additionally, the precision of the aforementioned components is adjusted dynamically in such a way that the rounding errors do not exceed 0.05%. Precision coefficients are also added to the bitstream.

Depending on whether the parameters change in time or not, ANY-HEVC transmits the projection matrices either in Video Parameter Set or Slice Header [ISO'21]. The updated syntax for both cases is presented in Table 4.1 and Table 4.2.

The presence of additional parameters in the bitstream results in an increased bitrate after ANY-HEVC compression compared to the state-of-the-art 3D-HEVC. In the case of camera arrangements other than linear, this overhead may be reduced by better prediction. However, for linear camera arrangements, such representation is redundant and leads to decreased RD compression efficiency (see Section 4.3).

Table 4.1. ANY-HEVC modified syntax for transmission of camera parameters in Video Parameter Set (VPS).

vps_3d_extension() {	Value
cp_precision	ue(v)
for (n = 0; n < NumViews; n++) {	
i = ViewOIdxList[n]	
cp_in_slice_segment_header_flag[i]	u(1)
if (!cp_in_slice_segment_header_flag[i]) {	
vps_cp_znear[i]	se(v)
vps_cp_zfar[i]	se(v)
for (j=0; j<12; j++)	
vps_cp_projection_matrix[i][j]	se(v)
for (j=0; j<12; j++)	
vps_cp_projection_matrix_prec[i][j]	ue(v)
}	
}	
}	

Table 4.2. ANY-HEVC modified syntax for transmission of camera parameters in Slice Header.

if(cp_in_slice_segment_header_flag[ViewIdx]) {	Value
for (i=0; i<num_cp[ViewIdx]; i++) {	
cp_znear[i]	se(v)
cp_zfar[i]	se(v)
for (j=0; j<12; j++)	
cp_projection_matrix[i][j]	se(v)
for (j=0; j<12; j++)	
cp_projection_matrix_prec[i][j]	ue(v)
}	
}	

4.3. Evaluation of ANY-HEVC codec in compression of multiview video acquired by various camera arrangements

In the previous section, it is stated that using full 3D point mapping in ANY-HEVC provides an accurate inter-view prediction, regardless of the arrangement of the cameras used to capture the scene. On the other hand, the disparity vector in DCP becomes a two-dimensional vector, which requires transmission of the additional component in the bitstream. Moreover, such complex inter-view prediction requires additional camera parameters to be also added to the bitstream. This section presents an experimental comparison of rate-distortion compression efficiency between ANY-HEVC and the standard techniques: HEVC simulcast, MV-HEVC, and 3D-HEVC. The goal is to verify how the camera arrangement and the number of views influence the bitrate at a given quality.

The experiments are performed on 7 linear and 4 nearly-circular sequences. Details for the used sequences can be found in Chapter 3. It should be stressed that nearly-circular sequences are not rectified to an ideal circle – that case is considered later in the chapter in Sections 4.4 – 4.6. The experiments are performed in Random Access configuration, according to the appropriate Common Test Conditions, as described in Chapter 3. Detailed configuration is presented in Table 4.3.

Table 4.3. Configuration details used for the experiments.

Parameter	**Value**
Number of coded frames	50
QP for views	25, 30, 35, 40
QP for depth maps	34, 39, 42, 45
Sample Adaptive Offset	on
View Synthesis Prediction	on (if applicable)
View Synthesis Optimization	off
Inter-view Motion Prediction	on (if applicable)
Neighboring Block Disparity Vector	on (if applicable)
Depth-oriented Neighboring Block Disparity Vector	on (if applicable)

The results for linear camera arrangements are presented in Table 4.4, whereas Table 4.5 gathers the results for nearly-circular camera arrangements. In both cases, HEVC simulcast encoding is used as the reference. Bitrate reduction is expressed as BD-rate for luma samples, as described in Chapter 3.

Table 4.4. Bitrate reduction for compression of 3 views acquired by linear camera arrangements, compared to simulcast HEVC. Negative values indicate lower bitrates at the same quality.

Sequence	**MV-HEVC**	**3D-HEVC**	**ANY-HEVC**
Poznan_Hall2	-30.60%	-36.44%	-36.36%
Poznan_Street	-43.24%	-46.20%	-46.19%
Dancer	-49.27%	-52.20%	-52.19%
Kendo	-32.99%	-41.75%	-41.68%
Ballons	-32.19%	-39.30%	-39.26%
Newspaper	-35.12%	-38.66%	-38.59%
Average	**-37.23%**	**-42.43%**	**-42.38%**

Table 4.5. Bitrate reduction for nearly-circular camera arrangements, compared to simulcast HEVC. Negative values indicate lower bitrates at the same quality.

	3 views			5 views			7 views		
Sequence	**MV-HEVC**	**3D-HEVC**	**ANY-HEVC**	**MV-HEVC**	**3D-HEVC**	**ANY-HEVC**	**MV-HEVC**	**3D-HEVC**	**ANY-HEVC**
Ballet	-19.72%	-21.78%	-26.35%	-23.64%	-26.16%	-31.80%	-24.32%	-27.58%	-33.52%
Breakdancers	-21.84%	-25.87%	-30.00%	-29.72%	-33.15%	-38.65%	-31.21%	-34.43%	-40.36%
BBB_Flowers	-5.59%	-5.69%	-8.00%	-7.33%	-7.20%	-9.79%	-8.47%	-8.22%	-10.74%
Poznan_Blocks	-13.65%	-12.77%	-16.72%	-15.95%	-14.53%	-19.23%	-16.45%	-14.51%	-19.64%
Average	**-15.20%**	**-16.53%**	**-20.27%**	**-19.16%**	**-20.26%**	**-24.86%**	**-20.11%**	**-21.19%**	**-26.06%**

The results for linear camera arrangements (Table 4.4) prove that the dedicated techniques for multiview video compression provide significant gain compared to independent compression of each view. As expected, the smallest gain of roughly 37% is reported for MV-HEVC as it does not utilize the information about depth to improve inter-view prediction. The most significant reduction in the bitrate is achieved for the state-of-the-art 3D-HEVC, which is dedicated to linear camera arrangements. In the case of ANY-HEVC, it is nearly as effective as the standard 3D-HEVC. The slight difference in favor of 3D-HEVC is caused by the increased number of camera parameters used in ANY-HEVC.

Regarding nearly-circular camera arrangements (Table 4.5), the bitrate reduction is noticeably smaller than in the case of linear arrangements. This is caused by the sparse distribution of cameras and the relatively significant perspective difference between the views, which leads to less accurate inter-view prediction. Similarly to the compression of linear sequences, MV-HEVC reports the smallest bitrate reduction compared to the HEVC simulcast. 3D-HEVC is a more efficient technique, however, the difference is not outstanding. This indicates that due to the assumption of linear camera arrangement, 3D-HEVC is inefficient for compression of multiview video captured with camera arrangements other than linear. A more significant gain is observed for ANY-HEVC at the cost of complex inter-view prediction using 3D point mapping.

Another conclusion from the results presented in Table 4.5 is that the more views are coded, the bigger the bitrate reduction is. This is caused by the increasing contribution of inter-view prediction in the coding process. Depending on the sequence, the gain can be as high as 40% for the compression of 7 views using ANY-HEVC. Therefore, multiview video acquired using camera arrangements other than linear can also be efficiently compressed using advanced inter-view prediction techniques.

Table 4.6 presents the comparison of the encoding times for different multiview codecs in reference to the HEVC simulcast. The encoding in all cases was performed on the same number of views, i.e., the depth maps were also encoded using HEVC simulcast and MV-HEVC, even though they are not required for compression of views acquired by cameras.

Table 4.6. Encoding time change for compression of 7 views acquired by nearly-circular camera arrangements, compared to simulcast HEVC. Positive values indicate longer encoding.

Sequence	**MV-HEVC**	**3D-HEVC**	**ANY-HEVC**
Ballet	113.51%	168.27%	191.52%
Breakdancers	85.95%	209.89%	216.17%
BBB_Flowers	98.35%	190.97%	225.61%
Poznan_Blocks	59.69%	95.10%	106.22%
Average	**89.38%**	**166.06%**	**184.88%**

The results in Table 4.6 show that the processing time of the dedicated multiview video encoders is much longer than when using simulcast encoding. This is caused by the inter-view prediction techniques that incorporate complex algorithms to improve the rate-distortion compression efficiency at the cost of longer computational time. For example, MV-HEVC provides roughly 20% bitrate reduction (Table 4.5) at the cost of encoding time increased by, on average, 90%.

In the case of nearly-circular camera arrangements, ANY-HEVC provides the biggest bitrate reduction, but on the other hand, this encoder is also the slowest due to the complex 3D point mapping. However, the author observed that if the cameras were distributed ideally on a circle, i.e., the acquired views were circularly rectified, the inter-view prediction could be simplified, and thus the encoding time could be reduced. This idea is presented in the following sections of this chapter.

4.4. CIRCULAR RECTIFICATION

4.4.1. INTRODUCTION

In the dissertation, the author proposes an efficient modification of the 3D-HEVC codec for processing circularly rectified 3D video (multiview video with depth, MVD) and the procedure for circular rectification. Acquisition systems with multiple cameras sparsely distributed around the scene have already been successfully set up and used to record some of the publicly available 3D video test sequences [Suzuki'09, Stankiewicz'18, Domański'15B, Domański'16B]. Setting up such a multi-camera system is even more challenging than in the case of linear camera arrangements. In practice, the cameras are never located ideally in a circle. The goal of the author's circular rectification is to correct the imperfections of camera positioning in a nearly-circular camera arrangement. This allows preparing an original codec (ARC-HEVC) with simplified inter-view prediction (compared to ANY-HEVC), dedicated to compression of circularly rectified multiview video (Figure 4.4). Such an approach is dual to 3D-HEVC, however the assumption of circular camera arrangement instead of linear is more useful in various applications, as described in Section 4.1.

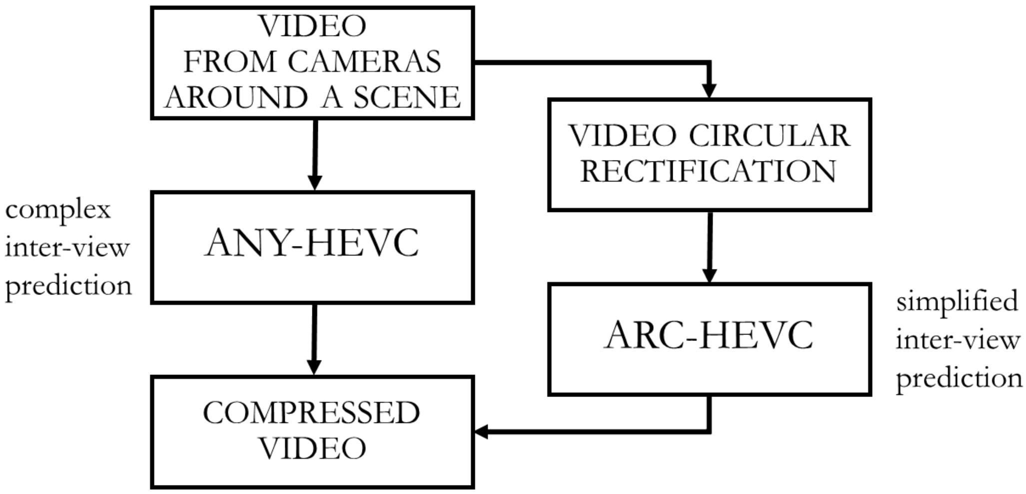

Figure 4.4. Two approaches to compression of multiview video acquired using systems with nearly-circular camera arrangements.

In the dissertation, circular rectification is not considered as a part of ARC-HEVC, but rather as a post-processing phase after capturing multiview video. It is in line with 3D-HEVC where the input video is expected to already be rectified. Therefore, the performance of the proposed circular rectification is not taken into account when comparing ARC-HEVC with other compression techniques.

The proposed circular rectification transforms the original MVD data and results in the multiview video with depth maps virtually obtained from cameras located on an ideal circle and with optical axes collocated on a single plane and intersecting in the center of the circle (Figure 4.5).

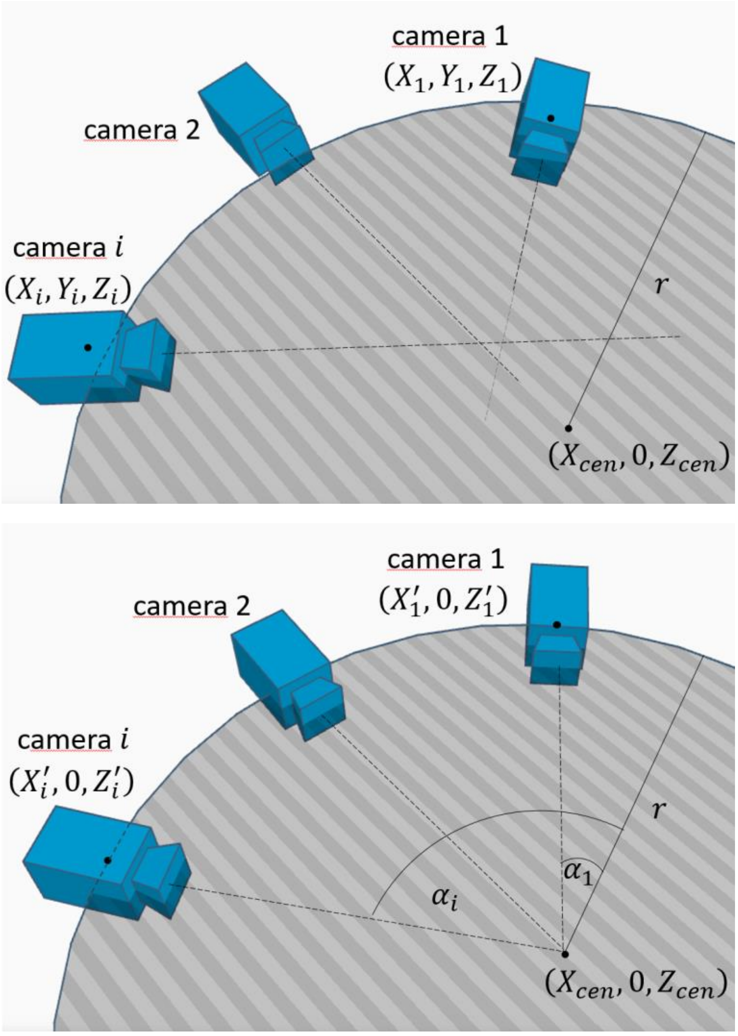

Figure 4.5. Circular camera setup before (top) and after (bottom) the proposed rectification.

The procedure for the proposed circular rectification of the multiview video is based on the equation for 3D point mapping (Equation (4.3)) between the original and the rectified view and is defined as follows:

$$\forall \begin{Bmatrix} x \in \mathbb{Z} : 0 \leq x < W, \\ y \in \mathbb{Z} : 0 \leq y < H \end{Bmatrix}, \qquad \begin{bmatrix} z_{x',y'} \cdot x' \\ z_{x',y'} \cdot y' \\ z_{x',y'} \\ 1 \end{bmatrix} = \mathbb{P}_{circle} \cdot \mathbb{P}_{org}^{-1} \begin{bmatrix} z_{x,y} \cdot x \\ z_{x,y} \cdot y \\ z_{x,y} \\ 1 \end{bmatrix}, \tag{4.5}$$

where:

(x, y), (x', y') – positions of corresponding points in the original and rectified views,

$z_{x,y}$, $z_{x',y'}$ – distances between the point in 3D space and the planes of the original and the rectified view, respectively, calculated from depth map sample values using Formula (2.2),

$\mathbb{P}_{org}$, $\mathbb{P}_{circle}$ – projection matrices for the original and the rectified view, respectively,

W, H – width and height of the original view,

$\mathbb{Z}$ – set of integer numbers.

For each position of a sample in the original image, the corresponding position in the rectified image is derived along with the distance to the rectified image plane, and the view sample is copied. It is possible that multiple positions from the original view are projected onto the same position in the rectified view – in such a case, the point closest to the image plane is picked as the rectified sample. If multiple samples are equally distant, they are averaged. If the rectified view contains empty samples after processing the whole original view, missing samples are interpolated from the surrounding content.

The projection matrix of the original view ($\mathbb{P}_{org}$) is known and can be derived from the input camera parameters using Equation (4.2). The derivation of the projection matrix of the rectified view ($\mathbb{P}_{circle}$) is presented in Sections 4.4.2 – 4.4.4.

4.4.2. Derivation of new camera positions

This section presents the derivation of the parameters of the circle that best fits the original positions of the cameras, as well as the positions of the virtual cameras capturing the rectified views. The circle is represented by the position of its center $(X_{cen}, 0, Z_{cen})$ and radius r. In order to find the aforementioned parameters, the author uses circle Equation (4.6) with positions $(X_i, 0, Z_i)$ of each of the N cameras, and performs non-linear regression by minimizing the sum of squares S according to Equation (4.7).

$$(X_i - X_{cen})^2 + (Z_i - Z_{cen})^2 = r^2 \tag{4.6}$$

$$\mathrm{S} = \sum_{i=1}^{N} (\sqrt{(X_i - X_{cen})^2 + (Z_i - Z_{cen})^2} - r)^2 \tag{4.7}$$

It should be noted that vertical positions are ignored $(Y_{cen} = 0,\ Y_i = 0)$ because the proposed rectification assumes that all cameras, as well as the center of the circle, are located at the same height.

After the derivation of circle parameters, i.e., the position of its center $(X_{cen}, 0, Z_{cen})$ and radius r, the next step is to find for each camera its modified position $(X_i', 0, Z_i')$ on the circle, the closest to the original location. That position lies on a line appointed by the circle center and the original position of a camera and can be derived using Thales theorem:

$$\frac{r}{\sqrt{(X_i - X_{cen})^2 + (Z_i - Z_{cen})^2}} = \frac{X_i' - X_{cen}}{X_i - X_{cen}} \ . \tag{4.8}$$

Therefore:

$$X_i' = X_{cen} + \frac{X_i - X_{cen}}{\sqrt{(X_i - X_{cen})^2 + (Z_i - Z_{cen})^2}} \cdot r \ . \tag{4.9}$$

The rectified position along axis Z can then be derived from the circle equation (Equation (4.6)):

$$Z_i' = Z_{cen} + \sqrt{r^2 - (X_i' - X_{cen})^2} \ . \tag{4.10}$$

EXAMPLE:

Figure 4.6 presents an example of the correction of camera positions of one of the real 3D video test sequences - *Breakdancers*. It can be noticed that the differences are not significant; the cameras before rectification are located roughly on a circle. Table 4.7 presents a statistical analysis of the correction of camera positions. It can be seen that the positions of the *BBB_Flowers* sequence are not corrected at all because that sequence is rendered, and therefore the original camera positions are already positioned ideally on a circle. For the remaining sequences, the difference is still not major and does not exceed 1.5% of the circle radius. Therefore, it is a sensible approach to rectify those viewports to a circle and adapt 3D-HEVC to circular camera arrangements.

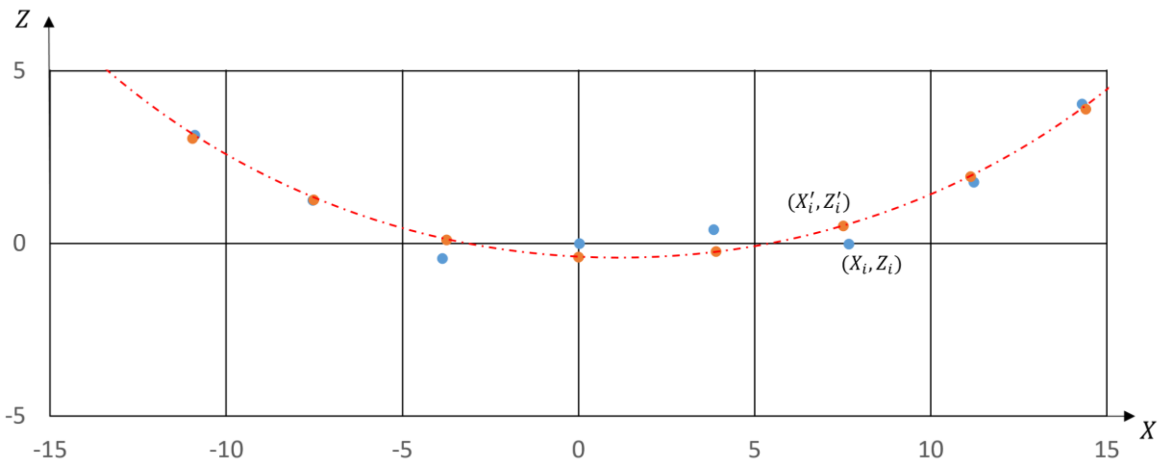

Figure 4.6. Top view of a multi-camera system with original camera positions (blue dots) and corrected to an ideal circle (orange dots). Breakdancers test sequence.

Table 4.7. The average difference between the camera positions before and after rectification, standard deviation, and the ratio of the average difference to the circle radius.

Sequence	**average difference between original and rectified camera positions Δ(x,z)**	**standard deviation of Δ(x,z)**	**average difference to radius ratio Δ(x,z) / r**
Ballet	0.2458	0.2142	0.93%
Breakdancers	0.3282	0.2209	1.45%
BBB_Flowers	0.0000	0.0000	0.00%
Poznan_Blocks	0.1177	0.0992	0.61%

4.4.3. Camera rotation in the ideally circular arrangement

In the proposed process of circular rectification, the goal is not only to correct the locations of the cameras but also to direct their optical axes precisely toward the center of a circle derived in the previous subsection. To achieve that, modification of rotation matrices is necessary.

The rotation matrix $\mathbb{R}$ represents the combined rotation of a camera around three orthogonal axes. In the ideally circular arrangement, optical axes are assumed to be on a single plane. Such camera rotation can be represented by the following matrix (4.11):

$$\mathbb{R}'_i = \begin{bmatrix} \cos\alpha_i & 0 & -\sin\alpha_i \\ 0 & 1 & 0 \\ \sin\alpha_i & 0 & \cos\alpha_i \end{bmatrix}, \tag{4.11}$$

where α_i is the angle between the direction of the *i*-th camera optical axis and circle radius along Z axis (Figure 4.5), therefore:

$$\cos\alpha_i = \frac{Z_{cen} - Z_i'}{r}, \qquad \sin\alpha_i = \frac{X_{cen} - X_i'}{r}. \tag{4.12}$$

All the necessary parameters: circle center position (X_{cen}, Z_{cen}) and its radius r, as well as the modified *i*-th camera position (X_i', Z_i') are already derived, thus there is no need to transmit more camera parameters in the bitstream to find the rotation matrix of rectified cameras.

4.4.4. MODIFICATION OF THE INTRINSIC CAMERA PARAMETERS

Previous subsections 4.4.2 – 4.4.3 present the process of deriving extrinsic parameters of cameras located on an ideal circle. This subsection describes how the internal camera parameters are evaluated in the proposed circular rectification process.

First, the skew factor is set to $c = 0$. As mentioned in Section 3.4, it is assumed that the cameras are calibrated, which includes correction of the skew factor. The same assumption is made in the state-of-the-art 3D-HEVC [Domański'13, Müller'13]. Next, the focal lengths f_x, f_y and the vertical component of the principal point o_y are averaged and set equal for every camera, resulting in f_x', f_y' and o_y'. The same approximation is done in 3D-HEVC. However, since 3D-HEVC assumes that the views are coplanar and vertically aligned, the values of f_y and o_y are not used at all (there is no vertical component in the inter-view prediction). A more sophisticated approach is required to derive the horizontal component of the principal point o_x'. The cameras in a nearly-circular arrangement are usually directed toward the center of the recorded scene, which can be (and often is) much closer to the cameras than the center of a circle. Therefore, due to the modification of rotation matrices towards the center of a circle, the field of view of a camera can be significantly misaligned with the field of view of the original camera. This can result in only a small portion of the original field of view being covered by a given camera after circular rectification (Figure 4.7). In such a case, rectified views would contain only a small part of the original content, which is highly unwanted (Figure 4.8).

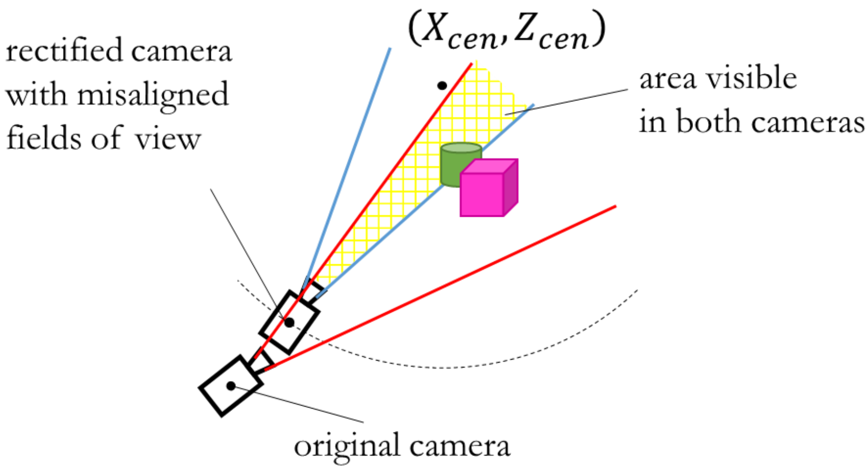

Figure 4.7. The problem with the misaligned fields of view after circular rectification.

Figure 4.8. A real example of the problem with the misaligned fields of view after circular rectification, Breakdancers test sequence.
View number 4 (left) is slightly misaligned, and view number 6 (right) is significantly rotated towards the center of a circle but outside the center of the acquired scene.

Cancellation of the misalignment of the field of view can be achieved not only by rotating the camera itself but also by changing its principal point (Figure 4.9) [Hou'20, Hartley'03]. In the proposed circular rectification technique, the component o'_x of the rectified principal is calculated for each camera to assert roughly the coverage of the recorded scene before rectification. It is done by projecting a point equal to the original optical center (o_x, o_y) from the original view onto rectified camera plane. If the projected point is not located at the optical center, the difference between o_x and its projection is added to o'_x of rectified camera parameters. Therefore, the optical axes of original and projected cameras are directed at the same point in 3D space and rectification of the rotation matrix is compensated.

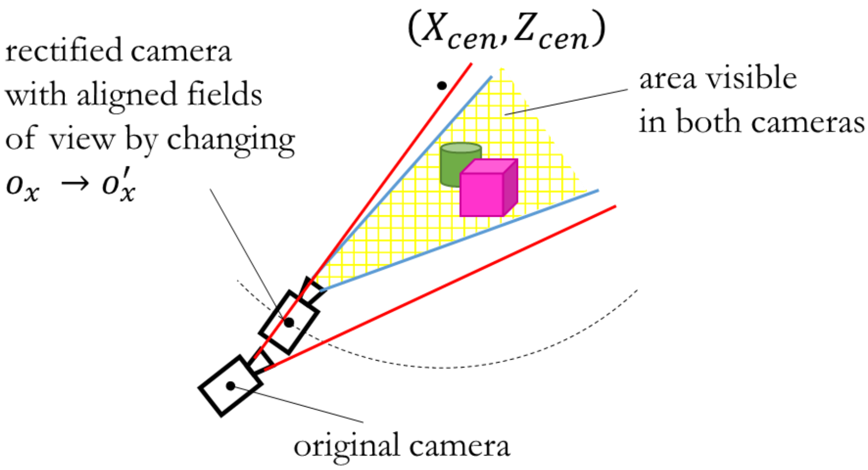

Figure 4.9. Rectified camera directed towards circle center
and with the modified optical center o'_x to align its field of view with the original camera.

4.4.5. Rectification of video and depth

After deriving circular camera parameters, multiview video can be rectified by performing 3D point mapping for every view of a sequence using Equation (4.5). This requires the calculation of projection matrices $\mathbb{P}_{org}$ (for original parameters) and $\mathbb{P}_{circle}$ (for circular parameters). Projection matrix for the original camera ($\mathbb{P}_{org}$) is calculated using all camera parameters according to Formula (4.2). Projection matrix for the rectified camera ($\mathbb{P}_{circle}$) is calculated using rectified camera parameters derived in Sections 4.4.2 – 4.4.4 and, for each view, is represented by Formula (4.13):

$$\mathbb{P}_{circle} = \begin{bmatrix} f'_x & 0 & o'_x & 0 \\ 0 & f'_y & o'_y & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\alpha & 0 & -\sin\alpha & -X'\cos\alpha + Z'\sin\alpha \\ 0 & 1 & 0 & 0 \\ \sin\alpha & 0 & \cos\alpha & -X'\sin\alpha - Z'\cos\alpha \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (4.13)$$

As mentioned before, a rectified view may contain some unfilled areas – these are interpolated from surrounding content. The author of the dissertation prepared **an original software for circular rectification**, which calculates rectified camera parameters from original parameters and transforms original MVD data into rectified multiview video with depth. Figure 4.10 presents an example of a frame of a test sequence *Poznan_Blocks* rectified with the author's software, and the difference image between the original and rectified view. All the views and the corresponding depth maps were rectified and used as input data for the experiments described in Section 4.6.

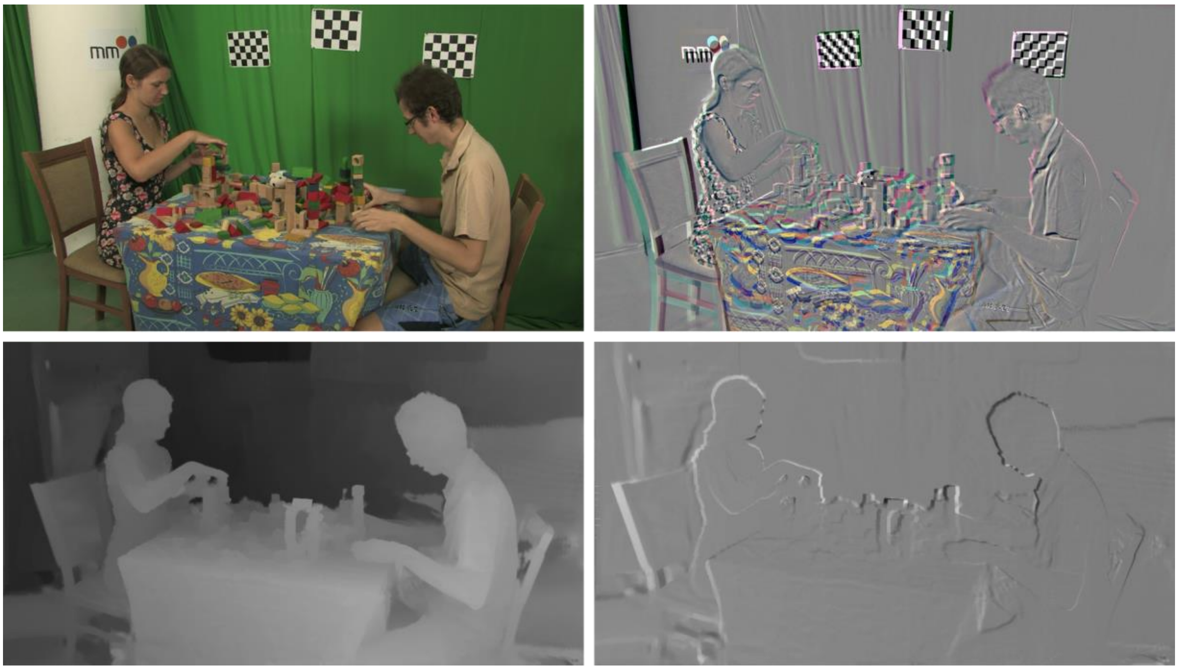

Figure 4.10. Rectified view (top-left), depth map (bottom-left), and the difference between the original and rectified view (top-right) and depth map (bottom-right), Poznan_Blocks test sequence.

4.5. 3D-HEVC ADAPTED TO RECTIFIED CIRCULAR CAMERA ARRANGEMENT (ARC-HEVC)

4.5.1. INTER-VIEW PREDICTION FOR RECTIFIED CIRCULAR CAMERA ARRANGEMENT

In the previous section, 3D point mapping was used to project the views from the original camera planes onto planes of virtual cameras positioned ideally on a circle. Such an operation is referred to as circular rectification. In this section, the author proposes to utilize 3D point mapping for inter-view prediction, which is similar to the approach presented in ANY-HEVC. In this case, however, both reference and coded views are rectified, hence their projection matrices (Equation (4.13)) are less complex than the ones in ANY-HEVC (Equation (4.2)). In general, inter-view prediction for rectified circular camera arrangement is defined as follows:

$$\begin{bmatrix} z_2 \cdot x_2 \\ z_2 \cdot y_2 \\ z_2 \\ 1 \end{bmatrix} = \mathbb{P}_{circle\_2} \cdot \mathbb{P}_{circle\_1}^{-1} \begin{bmatrix} z_1 \cdot x_1 \\ z_1 \cdot y_1 \\ z_1 \\ 1 \end{bmatrix}, \tag{4.14}$$

where:

$$\mathbb{P}_{circle\_1} = \begin{bmatrix} f_x' & 0 & o_{x1}' & 0 \\ 0 & f_y' & o_y' & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\alpha_1 & 0 & -\sin\alpha_1 & -X_1'\cos\alpha_1 + Z_1'\sin\alpha_1 \\ 0 & 1 & 0 & 0 \\ \sin\alpha_1 & 0 & \cos\alpha_1 & -X_1'\sin\alpha_1 - Z_1'\cos\alpha_1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (4.15)$$

$$\mathbb{P}_{circle\_2} = \begin{bmatrix} f_x' & 0 & o_{x2}' & 0 \\ 0 & f_y' & o_y' & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\alpha_2 & 0 & -\sin\alpha_2 & -X_2'\cos\alpha_2 + Z_2'\sin\alpha_2 \\ 0 & 1 & 0 & 0 \\ \sin\alpha_2 & 0 & \cos\alpha_2 & -X_2'\sin\alpha_2 - Z_2'\cos\alpha_2 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (4.16)$$

The multiplication of projection matrices $\mathbb{P}_{circle\_2} \cdot \mathbb{P}_{circle\_1}^{-1}$ results in the following matrix:

$$\mathbb{P}_{circle_2} \cdot \mathbb{P}_{circle_1}^{-1} = \quad (4.17)$$

$$= \begin{bmatrix} \cos\Delta\alpha + \frac{o_{x2}'}{f_x'}\sin\Delta\alpha & 0 & (o_{x2}' - o_{x1}')\cos\Delta\alpha - \left(f_x' + \frac{o_{x1}'o_{x2}'}{f_x'}\right)\sin\Delta\alpha & \begin{matrix} f_x'\left(\Delta X'\cos\alpha_2 - \Delta Z'\sin\alpha_2\right) + \\ +o_{x2}'\left(\Delta X'\sin\alpha_2 + \Delta Z'\cos\alpha_2\right) \end{matrix} \\ \frac{o_y'}{f_x'}\sin\Delta\alpha & 1 & o_y'\left(-\frac{o_{x1}'}{f_x'}\sin\Delta\alpha + \cos\Delta\alpha - 1\right) & o_y'\left(\Delta X'\sin\alpha_2 + \Delta Z'\cos\alpha_2\right) \\ \frac{1}{f_x'}\sin\Delta\alpha & 0 & -\frac{o_{x1}'}{f_x'}\sin\Delta\alpha + \cos\Delta\alpha & \Delta X'\sin\alpha_2 + \Delta Z'\cos\alpha_2 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

where:

$\Delta\alpha = \alpha_2 - \alpha_1$ (difference between the angles of optical axes of source and target camera, see Figure 4.4),

$\Delta X' = X_1' - X_2'$, the difference between camera positions along axis X,

$\Delta Z' = Z_1' - Z_2'$, the difference between camera positions along axis Z.

From Formula (4.12), differences $\Delta X'$ and $\Delta Z'$ can be defined as follows:

$$\Delta X' = X_1' - X_2' = X_{cen} - r\sin\alpha_1 - X_{cen} + r\sin\alpha_2 = r(\sin\alpha_2 - \sin\alpha_1), \quad (4.18)$$

$$\Delta Z' = Z_1' - Z_2' = Z_{cen} - r\cos\alpha_1 - Z_{cen} + r\cos\alpha_2 = r(\cos\alpha_2 - \cos\alpha_1). \quad (4.19)$$

Therefore, Formula (4.17) can be further simplified:

$$\mathbb{P}_{circle_2} \cdot \mathbb{P}_{circle_1}^{-1} = \tag{4.20}$$

$$= \begin{bmatrix} \cos\Delta\alpha + \frac{o'_{x2}}{f'_x}\sin\Delta\alpha & 0 & (o'_{x2} - o'_{x1})\cos\Delta\alpha - \left(f'_x + \frac{o'_{x1}o'_{x2}}{f'_x}\right)\sin\Delta\alpha & f'_x r\sin\Delta\alpha + o'_{x2}r(1-\cos\Delta\alpha) \\ \frac{o'_y}{f'_x}\sin\Delta\alpha & 1 & o'_y\left(-\frac{o'_{x1}}{f'_x}\sin\Delta\alpha + \cos\Delta\alpha - 1\right) & o'_y r(1-\cos\Delta\alpha) \\ \frac{1}{f'_x}\sin\Delta\alpha & 0 & -\frac{o'_{x1}}{f'_x}\sin\Delta\alpha + \cos\Delta\alpha & r(1-\cos\Delta\alpha) \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

Equation (4.14) for point mapping between circularly rectified views (View 1 and View 2) may be represented as the following Equations (4.21 – 4.23):

$$z_2 = (x_1 - o'_{x1})\frac{z_1}{f'_x}\sin\Delta\alpha + (z_1 - r)\cos\Delta\alpha + r\,, \tag{4.21}$$

$$y_2 = o'_y + \frac{z_1}{z_2}(y_1 - o'_y), \tag{4.22}$$

$$x_2 = o'_{x2} + \frac{1}{z_2}[(x_1 - o'_{x1})z_1\cos\Delta\alpha - (z_1 - r)f'_x\sin\Delta\alpha]\,, \tag{4.23}$$

where:

$\Delta\alpha = \alpha_2 - \alpha_1$ (difference between the angles of optical axes of source and target camera, see Figure 4.4),

(x_1, y_1, z_1) – point coordinates in rectified View 1 (reference view),

(x_2, y_2, z_2) – point coordinates in rectified View 2 (target view),

o'_{x1}, o'_{x2} – horizontal components of principal points of rectified Views 1 and 2,

o'_y – averaged vertical component of principal points,

f'_x – averaged horizontal focal length,

r – circle radius.

The above formulas allow to predict the position of a point in View 2 from its position in View 1 and circular camera parameters, provided that both views are circularly rectified. The author of the dissertation modifies the inter-view prediction in 3D-HEVC by replacing standard disparity derivation with point projection that uses the above equations. The description of modifications is presented in the following sections.

4.5.2. Modified Disparity Compensated Prediction

The disparity vector is calculated similarly as in ANY-HEVC (Section 4.2.2) as a difference between the position of the source sample and its projection in the reference view, according to Equation (4.4). This time, however, the projection is performed using the equations derived in Section 4.5.1. According to Formula (4.22), the vertical component of disparity vector (i.e., $d_y = y_2 - y_1$) is equal to zero only when the distance between the acquired point and the view plane is equal for both views, i.e., $z_2 = z_1$ (Figure 4.11). Otherwise, the DCP in ARC-HEVC works on two-dimensional disparity vectors, similarly to ANY-HEVC.

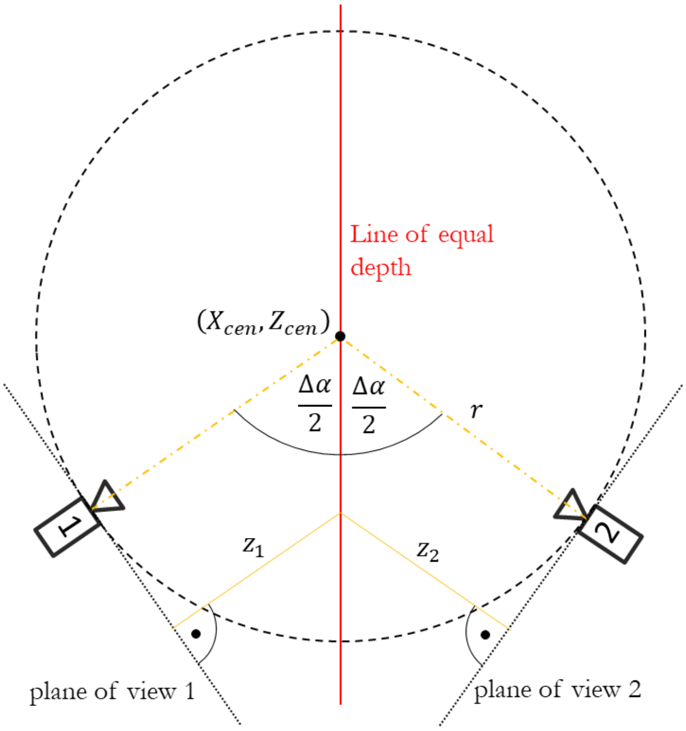

Figure 4.11. Visualization of the line of equal depth for a pair of cameras in a circular arrangement.

4.5.3. Modified inter-view prediction tools

Analogously to ANY-HEVC, the adaptation of 3D-HEVC to circular camera arrangements requires modification of inter-view coding tools that depend on disparity vectors or prediction of depth. The author of the dissertation applied the 2-dimensional disparity vector and the proposed equations for inter-view prediction (Section 4.5.1) to the following coding tools:

- Inter-View Motion Prediction,
- View Synthesis Prediction,
- Neighboring Block Disparity Vector,
- Depth-oriented Neighboring Block Disparity Vector,
- Illumination Compensation.

In particular, the most significant change compared to ANY-HEVC is in the prediction of depth samples. In ANY-HEVC, the derivation of depth sample values in the coded view requires complex calculations, and it depends on the depth sample value (z_1) in the reference depth map and sample position (x_1, y_1) in the reference view. In ARC-HEVC, the prediction formula is much simpler (Formula 4.21), and for a given pair of cameras (with constant parameters), it depends on the depth value from the reference view (z_1) and only one, horizontal component of the sample position (x_1). The author introduced a dedicated caching algorithm that stores in the encoder's memory the value of the predicted depth z_2 for each processed pair of (x_1, z_1) for future use. It is a similar approach to the look-up tables from 3D-HEVC for mapping depth to disparity. The difference is that in 3D-HEVC, look-up tables are prepared before encoding a view, while in the proposal, they are continuously updated whenever a unique pair of values for (x_1, z_1) occurs. Nonetheless, such an approach reduces the number of 3D point mappings and therefore, the encoding time.

4.5.4. MODIFIED BITSTREAM SYNTAX

As mentioned before, the proposed solution for efficient compression of circularly rectified video requires more information about the intrinsic and extrinsic camera parameters than 3D-HEVC. These parameters have to be transmitted in the bitstream because the derivation of disparity vectors has to be performed analogously by the decoder. The number of parameters and their representation in the bitstream varies depending on the compression technique.

Table 4.8 compares the parameters used by 3D-HEVC and modified encoders for compression of multiview video acquired with circular (ARC-HEVC) and arbitrary (ANY-HEVC) camera arrangements.

Table 4.8. Camera parameters required by different 3D video coding techniques.

<table>
<tr><th>Parameter name</th><th>3D-HEVC
(linear camera setup)</th><th>ARC-HEVC
(circular camera setup)</th><th>ANY-HEVC
(arbitrary camera setup)</th></tr>
<tr><td>Horizontal focal length</td><td>f_x (1 for all views)</td><td>f_x (1 for all views)</td><td>f_x</td></tr>
<tr><td>Vertical focal length</td><td>-</td><td>-</td><td>f_y</td></tr>
<tr><td>Horizontal optical center</td><td>o_x</td><td>o_x</td><td>o_x</td></tr>
<tr><td>Vertical optical center</td><td>-</td><td>o_y (1 for all views)</td><td>o_y</td></tr>
<tr><td>Skew factor</td><td>-</td><td>-</td><td>c</td></tr>
<tr><td>Translation</td><td>t_x</td><td rowspan="2">α,
r (1 for all views)</td><td>$\mathbb{T} = [t_x, t_y, t_z]$</td></tr>
<tr><td>Rotation</td><td>-</td><td>$\mathbb{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$</td></tr>
</table>

It should be noted that rectified circular camera setup requires much fewer parameters than arbitrary and only two more values than unmodified 3D-HEVC. However, neither 3D-HEVC nor ANY-HEVC transmits the parameters directly. In 3D-HEVC, they are represented as “scale” and “offset” [ISO’21], while ANY-HEVC transmits calculated projection matrices (Section 4.2.4) [Samelak’16]. In the proposed ARC-HEVC, the parameters are included in the bitstream directly and transmitted in the dedicated 3D extension of the Video Parameter Set (Table 4.9) or in the Slice Header (Table 4.10), depending on whether the parameters are constant or change in time. The parameters for Z_{near}, Z_{far}, as well as metadata parameters (**cp_precision, cp_in_slice_segment_header_flag**) are transmitted in every technique.

Table 4.9. Proposed syntax of VPS extension for ARC-HEVC.

vps_3d_extension() {	Value
cp_precision	ue(v)
vps_cp_focal_length_x	se(v)
vps_cp_principal_point_y	se(v)
vps_cp_radius	se(v)
for (n = 0; n < NumViews; n++) {	
i = ViewOIdxList[n]	
cp_in_slice_segment_header_flag[i]	u(1)
if (!cp_in_slice_segment_header_flag[i]) {	
vps_cp_znear[i]	se(v)
vps_cp_zfar[i]	se(v)
vps_cp_principal_point_x[i]	se(v)
vps_cp_angle[i]	se(v)
}	
}	
}	

Table 4.10. Proposed syntax in Slice Header for ARC-HEVC.

if(cp_in_slice_segment_header_flag[ViewIdx]) {	Value
cp_focal_length_x	se(v)
cp_principal_point_y	se(v)
cp_radius	se(v)
for (i=0; i<num_cp[ViewIdx]; i++) {	
cp_znear[i]	se(v)
cp_zfar[i]	se(v)
cp_principal_point_x[i]	se(v)
cp_angle[i]	se(v)
}	
}	

It should be mentioned that if the camera parameters vary in time, the circular rectification has to be updated accordingly. However, such a case is out of the scope of the dissertation; the parameters of all test sequences used in the evaluation are constant.

The proposed changes in the bitstream result in ARC-HEVC not being compliant with the 3D-HEVC standard. This means that 3D-HEVC is not able to decode the bitstream produced by ARC-HEVC and vice versa. Nevertheless, the author of the dissertation proves that support for rectified circular 3D video compression could be added with only minor changes in the bitstream syntax.

4.6. EVALUATION OF THE PROPOSED ARC-HEVC

4.6.1. INTRODUCTION

In this section, ARC-HEVC, i.e., modified 3D-HEVC codec for efficient compression of circularly rectified video, is experimentally evaluated and compared to the state-of-the-art 3D-HEVC (designed for linear camera arrangement) and ANY-HEVC, which is the modified 3D-HEVC for arbitrary camera arrangement.

The goal of the experiments is to assess the rate-distortion compression efficiency and encoding time using the aforementioned codecs. Additionally, the author compares the encoding time of only intra-view prediction for both modified 3D-HEVC encoders (ARC-HEVC and ANY-HEVC). The RD compression efficiency is compared by measuring average bitrate reduction for the luma component of input views, using Bjøntegaard metric as described in Chapter 3. The experiments are conducted by encoding 7 rectified views of 4 commonly-used multiview test sequences, listed in Section 3.4. All the input views and the corresponding depth maps are rectified and used as the input for tested codecs (Figure 4.12). Rectified camera parameters and complete results of the conducted experiments can be found in the Appendix.

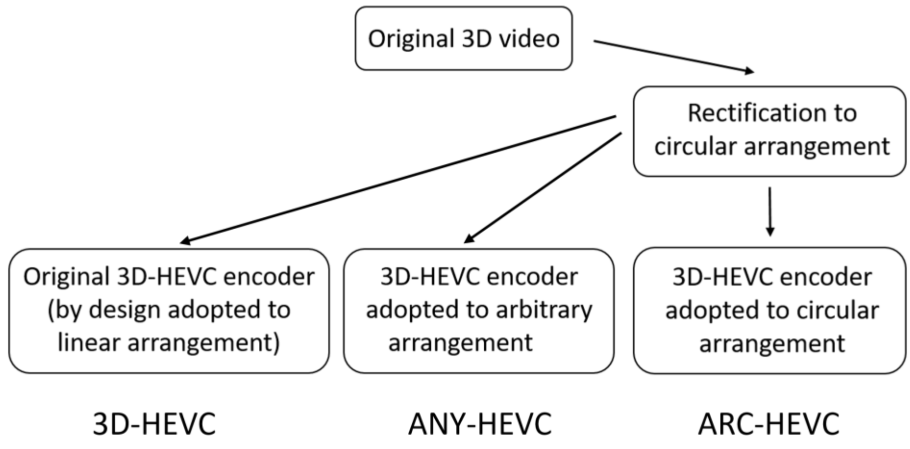

Figure 4.12. Block diagram of the evaluation of tested encoders in compression of circularly rectified video.

The configuration of all three encoders is the same as in the experiment presented in Section 4.3, except that this time the number of encoded frames is equal to 100. The remaining configuration follows Common Test Conditions for 3D video experiments [Müller'14] for Random Access coding scenario. The camera parameters are prepared according to the requirements of each of the encoders (Table 4.8). The rectification of test sequences is done in the pre-processing phase, so it does not affect encoding time results. Moreover, as described in Section 3.2, all three encoders are based on the same version of the 3D-HEVC publicly available test model [HTM].

4.6.2. EXPERIMENTAL RESULTS

This section contains the results of conducted experiments. Table 4.11 shows the comparison of the three in terms of bitrate reduction. **For compression of 7 views, the proposed ARC-HEVC reduces bitrate on average by 6%** compared to the state-of-the-art technique. This is because 3D-HEVC does not perform accurate inter-view prediction if the video was acquired by camera arrangements other than linear. The more views are encoded, the bigger the difference is in favor of ARC-HEVC.

Compared to the encoder that supports arbitrary camera setup, i.e., ANY-HEVC, the solution adapted to circular arrangements provides very similar results. This is expected as both techniques perform an accurate inter-view prediction when the views are circularly rectified. It can be noticed that for a low number of views, ANY-HEVC is slightly better, but for more views, it is the opposite. The difference is caused by a lower number of camera parameters required by the proposed ARC-HEVC and simplified inter-view prediction, which results in a reduced number of rounding errors. Nevertheless, the difference is not significant, therefore **both encoders can be considered equally efficient in terms of compression rate**.

Table 4.11. Bitrate reduction comparison between tested encoders, only for views.
Negative values indicate lower bitrates at the same quality.
ANY=ANY-HEVC, ARC=ARC-HEVC, 3D=3D-HEVC

	3 views			5 views			7 views		
Sequence	**ANY vs. 3D**	**ARC vs. 3D**	**ARC vs. ANY**	**ANY vs. 3D**	**ARC vs. 3D**	**ARC vs. ANY**	**ANY vs. 3D**	**ARC vs. 3D**	**ARC vs. ANY**
Poznan_Blocks	-6.31%	-6.15%	0.18%	-7.76%	-8.05%	-0.31%	-7.93%	-8.14%	-0.23%
BBB_Flowers	-6.87%	-6.08%	0.86%	-8.62%	-7.86%	0.83%	-8.96%	-8.22%	0.82%
Ballet	-2.31%	-2.30%	0.00%	-2.67%	-2.94%	-0.27%	-2.63%	-3.15%	-0.54%
Breakdancers	-3.70%	-3.79%	-0.09%	-4.32%	-4.52%	-0.21%	-4.39%	-4.61%	-0.23%
Average	**-4.80%**	**-4.58%**	**0.24%**	**-5.84%**	**-5.84%**	**0.01%**	**-5.98%**	**-6.03%**	**-0.05%**

Table 4.12 presents the results for the views and depth maps combined. The relation between the modified encoders, i.e., ARC-HEVC and ANY-HEVC, is not changed. However, when compared to 3D-HEVC, both encoders report slightly lower gains.

Table 4.12. Bitrate reduction comparison between tested encoders, views + depth maps.
Negative values indicate lower bitrates at the same quality.
ANY=ANY-HEVC, ARC=ARC-HEVC, 3D=3D-HEVC

	3 views			5 views			7 views		
Sequence	**ANY vs. 3D**	**ARC vs. 3D**	**ARC vs. ANY**	**ANY vs. 3D**	**ARC vs. 3D**	**ARC vs. ANY**	**ANY vs. 3D**	**ARC vs. 3D**	**ARC vs. ANY**
Poznan_Blocks	-5.50%	-5.38%	0.17%	-6.55%	-6.76%	-0.22%	-6.62%	-6.81%	-0.17%
BBB_Flowers	-6.23%	-5.57%	0.76%	-7.70%	-7.10%	0.74%	-7.90%	-7.32%	0.69%
Ballet	-1.99%	-2.01%	-0.04%	-2.30%	-2.57%	-0.29%	-2.28%	-2.78%	-0.51%
Breakdancers	-3.59%	-3.68%	-0.09%	-4.18%	-4.37%	-0.21%	-4.23%	-4.44%	-0.23%
Average	**-4.33%**	**-4.16%**	**0.20%**	**-5.18%**	**-5.20%**	**0.01%**	**-5.26%**	**-5.34%**	**-0.06%**

Table 4.13 presents the reduction of total encoding time, while Table 4.14 compares the inter-view prediction time between two modified encoders. It should be noted that **the proposed encoder is up to 10% faster than the encoder with full 3D point mapping** (ARC-HEVC vs. ANY-HEVC). At the same time, its **inter-view prediction is, on average, 44 times faster**, due to much simpler equations for point mapping, optimized for circularly rectified 3D video. Surprisingly, **the proposed encoder is also faster than 3D-HEVC by more than 4%**, even though the inter-view prediction of the former is more complex than the

state-of-the-art technique. The reason for such a phenomenon is related to the results for compression efficiency (Table 4.11). As mentioned before, ARC-HEVC is more accurate in predicting the content of the circularly rectified video. Therefore, it is able to find the best match for the currently encoded unit faster, compensating for the overhead resulting from more complex calculations.

Table 4.13. Encoding time comparison between tested encoders for compression of 7 views. Negative values indicate faster encoding.

Sequence	ANY-HEVC vs. 3D-HEVC	ARC-HEVC vs. 3D-HEVC	ARC-HEVC vs. ANY-HEVC
Ballet	9.27%	-1.89%	-10.21%
Breakdancers	-1.12%	-8.50%	-7.46%
BBB_Flowers	6.37%	-1.78%	-7.66%
Poznan_Blocks	5.98%	-4.50%	-9.89%
Average	**5.09%**	**-4.17%**	**-8.81%**

Table 4.14. Inter-view prediction time change between ARC-HEVC and ANY-HEVC. Negative values indicate faster inter-view prediction in ARC-HEVC.

Sequence	ARC-HEVC vs. ANY-HEVC
Ballet	-97.96% / 49x
Breakdancers	-97.79% / 45x
BBB_Flowers	-97.77% / 44x
Poznan_Blocks	-97.43% / 39x
Average	**-97.74% / 44x**

4.6.3. CONCLUSIONS

In this chapter, the author of the dissertation proposes a novel approach to the compression of multiview video with depth (3D video). Video acquired by cameras located roughly on a circle, as well as corresponding depth maps, is proposed to undergo circular rectification as presented and discussed in Section 4.4. The author of the dissertation develops a process for correcting camera parameters to an ideal circle together with circular video rectification. The state-of-the-art 3D-HEVC technique is proposed to be modified for efficient compression of such video. Moreover, point mapping equations for simplified inter-view prediction of circularly rectified 3D video are derived and implemented on top of the 3D-HEVC reference test model. The proposed modifications are evaluated experimentally and compared to the unmodified 3D-HEVC and the 3D-HEVC codec adapted to the compression of video acquired by cameras at arbitrary locations (ANY-HEVC), as described in Section 4.2.

The results show that the proposed technique is better than the other 2 solutions, both in terms of compression efficiency and encoding time. The state-of-the-art 3D-HEVC is 6% less efficient due to its restriction to linear camera arrangement. On the other hand, ARC-HEVC is nearly as efficient as the proposal, however, its inter-view prediction is 44 times slower. The author states that circular rectification combined with an adapted encoder allows to exploit a more useful camera arrangement than a linear one without compromising encoder complexity or compression efficiency. Therefore, the technique seems to be an interesting proposal for modern multiview applications such as free-viewpoint television, virtual reality, or immersive video.

5. INTER-VIEW PREDICTION WITH SCREEN CONTENT CODING

5.1. INTRODUCTION

This chapter presents a novel approach to multiview video coding with the use of HEVC Screen Content Coding (SCC). In Sections 5.2 – 5.6, the author of the dissertation presents the original idea and methodology for adapting Screen Content Coding to compression of stereoscopic video, frame-compatible multiview video, and immersive video. Section 5.7 contains experimental results and the evaluation of the proposal.

It should be stressed that the application of SCC to compression of camera-captured multiview video is a novel approach. Given the initial purpose of the SCC, i.e., the compression of screen content, the proposed solution may seem counterintuitive. Nevertheless, the author of the dissertation provides a detailed explanation of this idea, supported by a series of experiments. In this chapter, the author focuses on adapting SCC to the new applications without making any modifications to the encoder other than changes in its configuration, to prove the usability of standard-compliant SCC. Modifications of the SCC encoder that further increase the compression efficiency at the cost of breaking its compatibility with the state-of-the-art standard are presented in Chapter 6.

5.2. SCREEN CONTENT CODING IN MULTIVIEW CODING

As described in Section 2.3, Screen Content Coding was designed for the efficient compression of screencasts, rendered graphics, and other computer-generated video content [Xu'16A]. One of the tools introduced in the SCC extension is Intra Block Copy (IBC), which performs an intra-frame prediction by searching for the most similar block of samples within the previously encoded area of the processed frame. The result of IBC prediction is a 2-dimensional vector that indicates the best matching block of samples, even if it is located in a distant part of the frame. This type of prediction is highly effective for frames containing fonts and other repeatable patterns.

The author of the dissertation observed that **the Intra Block Copy technique could be used for inter-view prediction if all the views from multiview video composed a single frame**. Figure 5.1 presents the block diagram of the **proposed multiview codec**.

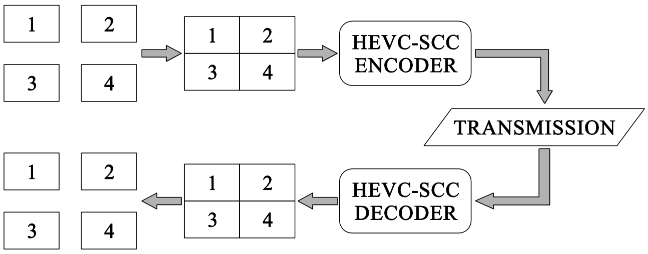

Figure 5.1. Block diagram of the proposed multiview video codec using HEVC Screen Content Coding. The numbers 1-4 denote subsequent views.

In the first step of the encoding process, all the views of a multiview sequence are joined into a single, frame-compatible video. Then, the resulting sequence is compressed with a video encoder that supports the Intra Block Copy technique. In the dissertation, the author uses the state-of-the-art HEVC Screen Content Coding. Obviously, depending on the resolution of the input sequences, the resolution of the frame-compatible video can be very high, and it has to be assured that the encoder is able to process this amount of data. Modern encoders support video compression at resolutions as high as “8K” (7680×4320), which accommodates a maximum of 16 views in the HD 1920×1080 format.

For the decoding, a state-of-the-art HEVC Screen Content Coding decoder is proposed to be used. The output is a reconstructed frame-compatible multiview video, which can then be split into separate views.

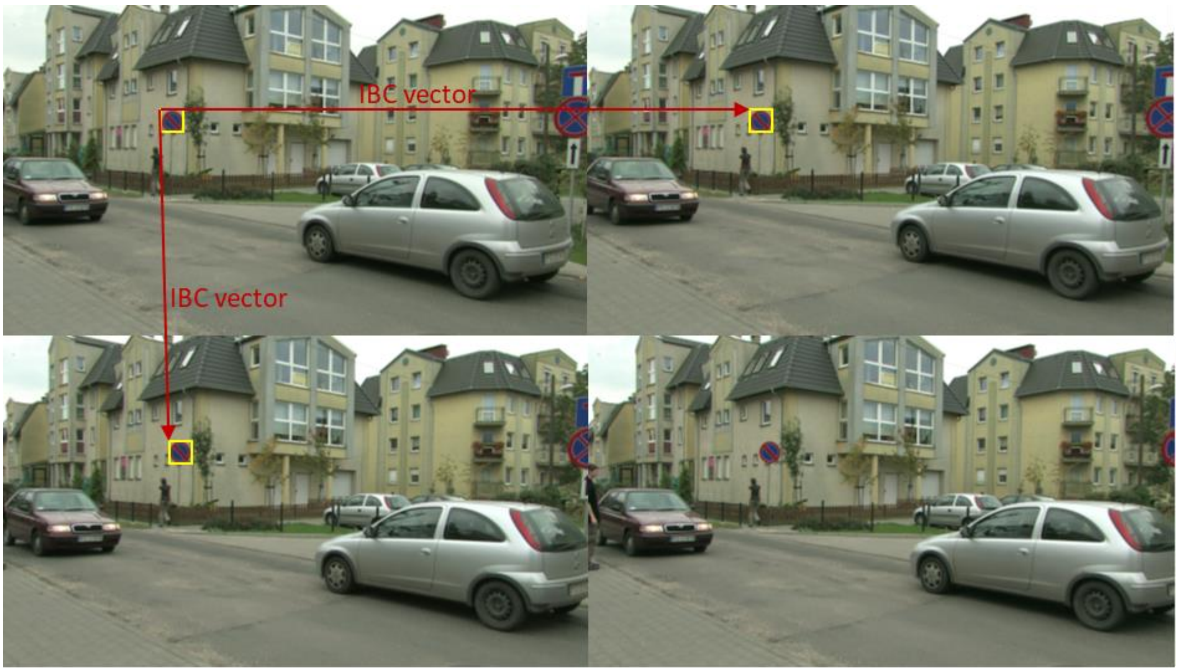

Figure 5.2. Intra Block Copy used on a frame composed of 4 views from a multiview sequence.

Figure 5.2 presents an example of a frame composed of 4 views from a multiview sequence. By the use of Intra Block Copy, the inter-view similarities can be predicted and compressed efficiently. It should be noted that the author of the dissertation proposes using IBC for the **compression of camera-captured content**, which is a **novel approach** and opposite to the original purpose of this technique – the compression of computer-generated video.

5.3. THE CHOICE OF VIEW ALIGNMENT FOR FRAME-COMPATIBLE MULTIVIEW VIDEO COMPRESSION

The first step of the proposed technique is to combine all the views of the input multiview video into one frame-compatible sequence. Figures 5.1 and 5.2 suggest joining the encoded views both horizontally and vertically. Such a solution seems natural because it preserves the original aspect ratio. On the other hand, it cannot be applied in some cases (e.g., 3 or 5 views) because the input sequence has to form a rectangular shape. For a four views coding scenario, the possible alignments are presented in Figure 5.3.

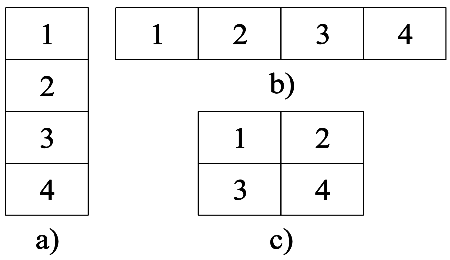

Figure 5.3. Different view alignments: a) 1×4, b) 4×1, c) 2×2.

In this section, the author of the dissertation compares experimentally the encoding time and the efficiency of the HEVC Screen Content Coding encoder in compression of the aforementioned three different view alignments. For the comparison, six commonly used multiview sequences are used. The 2×2 view alignment (Figure 5.3c) is used as a reference. Table 5.1 presents the results of bitrate and time reduction.

Both compression efficiency and encoding time are the best for horizontal (4×1) alignment. In this case, the vector derived by the Intra Block Copy has only a horizontal component because the input views are vertically aligned. Thus, these vectors are encoded more efficiently compared to the remaining test cases. The 4×1 view alignment is therefore used as the input for the remaining experiments presented in Chapter 5 and Chapter 6, as it is proven best for vertically aligned views.

Table 5.1. Bitrate and time reduction against 2×2 view alignment.
Negative values indicate lower bitrates at the same quality or faster encoding.

Sequence	Bitrate reduction [%]		Encoding time reduction [%]	
	1×4	4×1	1×4	4×1
Balloons	1.68	-5.57	-0.22	-8.11
BBB_Butterfly	-0.62	-6.40	5.77	1.44
Kendo	-1.75	-4.33	6.33	2.80
Newspaper	-2.31	-4.13	0.60	-3.19
Poznan_Hall2	3.89	-5.39	5.46	-4.04
Poznan_Street	-0.34	-4.02	9.38	3.47
Average	**-0.04**	**-4.58**	**5.64**	**-0.69**

5.4. THE CHOICE OF SCC CONFIGURATION FOR FRAME-COMPATIBLE MULTIVIEW VIDEO COMPRESSION

Apart from Intra Block Copy, the Screen Content Coding extension contains several other compression tools dedicated to computer-generated visual content. The author of the dissertation proposes applying SCC for camera-captured content, for which some of the SCC coding techniques may not be beneficial or even decrease the overall performance. This section contains the evaluation of selected SCC tools in terms of speed and efficiency in the compression of camera-captured content.

From the improvements introduced in Screen Content Coding, three techniques were selected for evaluation in the compression of camera-captured video:

- Intra Boundary Filter (disabled in Common Test Conditions) [Xu'16A]
- Hash-Based Intra Block Copy Search (enabled in CTC)
- Palette Mode (enabled in CTC)

Next, six commonly used multiview sequences are encoded using the HEVC SCC test model [HM+SCM] with a toggled configuration of the abovementioned tools, and compared to the results of compression with configuration compliant with Common Test Conditions for SCC [Yu'15]. Comparison results are presented in Table 5.2 and Table 5.3.

Table 5.2. Bitrate reduction against default Screen Content Coding configuration. Negative values indicate lower bitrates at the same quality.

Sequence	Intra Boundary Filter enabled	Hash-Based IBC Search disabled	Palette Mode disabled	all improvements
Balloons	-0.30%	0.00%	-0.07%	-0.38%
BBB_Butterfly	-0.19%	0.00%	0.04%	-0.09%
Kendo	-0.17%	0.00%	0.00%	-0.23%
Newspaper	-0.44%	0.00%	-0.05%	-0.46%
Poznan_Hall2	-0.31%	0.00%	-0.07%	-0.30%
Poznan_Street	-0.31%	0.00%	-0.01%	-0.34%
Average	**-0.30%**	**0.00%**	**-0.02%**	**-0.32%**

Table 5.3. Encoding time reduction against default Screen Content Coding configuration. Negative values indicate faster encoding.

Sequence	Intra Boundary Filter enabled	Hash-Based IBC Search disabled	Palette Mode disabled	all improvements
Balloons	1.80%	-1.61%	-16.53%	-20.28%
BBB_Butterfly	0.84%	0.13%	-19.58%	-19.46%
Kendo	-2.95%	-3.06%	-21.55%	-20.54%
Newspaper	1.34%	-2.96%	-26.14%	-25.33%
Poznan_Hall2	-1.59%	3.21%	-18.15%	-19.64%
Poznan_Street	0.80%	-1.09%	-15.76%	-17.73%
Average	**0.12%**	**-0.58%**	**-18.55%**	**-19.78%**

The results show that enabling Intra Boundary Filter slightly reduces the bitrate without a negative impact on the encoding time. The influence of Hash-Based IBC Search and Palette Mode on compression efficiency is negligible, which proves that those techniques are not beneficial for the camera-captured content. However, disabling them improves the encoding time significantly. In total, adapting Screen Content Coding configuration to camera-captured content provides roughly 0.3% of bitrate reduction and 20% faster encoding time. In the further evaluation of applying SCC to compression of camera-captured video, the author of the dissertation configures the HEVC SCC test model according to the modifications proposed in this section: enabled Intra Boundary Filter, disabled Hash-Based IBC Search and Palette Mode.

5.5. Screen Content Coding in stereoscopic video coding

Stereoscopic video is a particular case of multiview video with a variety of applications, such as entertainment (3D television), surveillance, depth estimation, or navigation assistance for the visually impaired [Strumiłło'18, Skulimowski'07, Müller'11, Ratajczak'12]. It comprises only two views, one for each of the spectator's eyes [Fujikawa'10]. The views are shifted horizontally by a distance that corresponds to the distance between humans' eyes. Therefore the majority of the content of the left view is also visible in the right view. A multiview encoder would exploit the inter-view similarities and provide the highest compression efficiency. In practice, however, the use of a dedicated multiview codec for stereoscopic video coding is very limited. Usually, in order to use the existing broadcasting infrastructure and receivers, both views of a stereoscopic video are accommodated into a single frame and compressed using a single-layer encoder with additional signalization information provided in SEI (Supplemental Enhancement Information) [ISO'21]. In such a case, the benefit of the inter-view prediction is sacrificed at the cost of simpler implementation.

In this section, the author of the dissertation proposes to apply HEVC Screen Content Coding to the compression of frame-compatible stereoscopic video as an alternative to the commonly used HEVC Main profile encoder. There are two main ways of packing left and right view into a single frame: Side-by-Side and Top-and-Bottom [Vetro'10]. Given the results presented in Section 5.3, the Side-by-Side solution is considered. Contrary to the approach for multiview video (Section 5.2), both views of the stereoscopic video are decimated in the horizontal direction by a factor of 2, as it is done in practice. The proposed block diagram is presented in Figure 5.4.

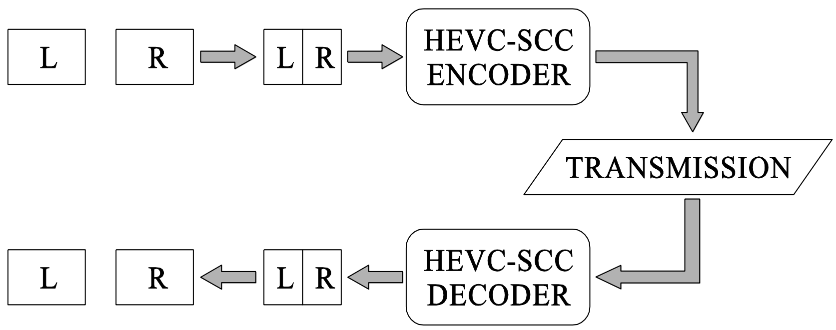

Figure 5.4. The proposed solution for stereoscopic video coding with Screen Content Coding.

5.6. Screen Content Coding in immersive video compression

The state-of-the-art in immersive video coding, as described in Section 2.4, is MPEG Immersive Video (MIV). The encoding process in MIV can be divided into three parts:

1. Preparation of atlases. This part includes dividing input views into base views and additional views, removing inter-view redundancies in the process of pruning, and creating atlases with patches and stacked base views during packing.
2. Compression of atlases. The output sequences from the previous step are fed into a general video encoder, such as HEVC or VVC. Each of the atlases is encoded individually with a separate instance of video encoder.
3. Preparation of MIV bitstream. The bitstreams produced by general video encoders in step 2 are combined into a single bitstream, with additional MIV metadata included.

In this section, the author proposes to use Screen Content Coding as the video encoder in step 2 **instead of the general video encoder**. In the dissertation, replacing HEVC with HEVC SCC is considered. The block diagram of the proposed solution is presented in Figure 5.5.

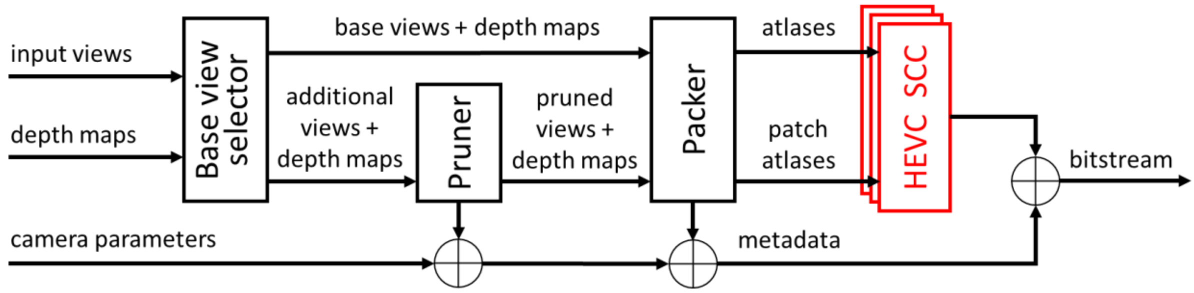

Figure 5.5. MPEG Immersive Video codec with HEVC Screen Content Coding as the internal video encoder.

The rationale for the above change is as follows:

- **Intra Block Copy can be beneficial for the compression of atlases containing base views**. Packing base views into atlases by MIV encoder is principally the same as creating frame-compatible multiview video, proposed in Section 5.2. It is expected that the Intra Block Copy technique would make use of the inter-view similarities between the views that compose an atlas. The difference is that the MIV encoder joins the base views vertically (Figure 2.6), while it was experimentally checked in Section 5.3 that Intra Block Copy is more efficient for horizontally aligned views. However, modifications of the MIV codec are out of the scope of this dissertation, therefore in this case, the vertical accommodation of base views is preserved.

- **Intra Block Copy can be beneficial for the compression of atlases containing patches from additional views**. Such atlases can contain similar patches located far from each other. Obviously, their number depends on the effectiveness of MIV in reducing the inter-view redundancy. Such repeated patches would be efficiently predicted with IBC, similarly to fonts in screen content.
- **Palette Mode can be beneficial for the compression of depth atlases**. As mentioned in Section 2.3, Palette Mode increases the overall compression efficiency if the encoded content has locally only a few colors separated by sharp edges [Pu'16]. It is often true for depth maps, where, e.g., the depth of an object is constant, and there is a sharp edge between an object and the background. In the case of atlases with camera-captured content, using Palette Mode would not provide such gain – it was experimentally checked in Section 5.4.

To summarize, using Screen Content Coding as the internal encoder can be beneficial due to the characteristics of atlases prepared in the initial phase of MIV encoding. The following section contains the results of the experiments conducted to verify the above thesis.

5.7. EVALUATION OF THE PROPOSAL

5.7.1. GOALS OF THE EXPERIMENTS

The goal of the experiments presented in this chapter is to evaluate the use of Screen Content Coding in compression of stereoscopic (side-by-side), frame-compatible multiview, and immersive video. The solutions proposed by the author of the dissertation are compared with the state-of-the-art techniques in terms of compression efficiency and encoding time (in the case of multiview and stereoscopic video) or the quality of synthesized views (in the case of immersive video). The author also performs experiments to determine if using SCC has to be combined with frame-compatibility to provide bitrate reduction.

The general information about the methodology of the experiments (e.g., versions of encoders used or definitions of metrics) is provided in Chapter 3. Sections 5.7.2 - 5.7.4 contain details specific to each experiment and the evaluation results. The complete results of the conducted experiments can be found in the Appendix.

5.7.2. Evaluation of SCC in stereoscopic video coding

In the first experiment, the application of HEVC SCC to stereoscopic video compression is compared with the common approach, which is using the HEVC Main profile. The encoders are obtained from publicly available test models as described in Chapter 3. The comparison is made by coding 100 frames of 2 arbitrarily chosen views obtained from 6 test sequences. The views are chosen according to Table 5.4. The experiments are conducted in All Intra and Random Access configurations according to appropriate Common Test Conditions [Bossen'13, Yu'15]. The configuration of tools specific to Screen Content Coding is prepared as proposed in Section 5.4.

Table 5.4. Chosen views of sequences used in the experiment.

Sequence	**Views (left, right)**
Balloons	3, 4
BBB_Butterfly	45, 50
Kendo	3, 4
Newspaper	4, 6
Poznan_Hall2	6, 5
Poznan_Street	4, 3

In order to fully evaluate the proposed solution, the following coding scenarios are applied:

- HEVC Main simulcast – the views are not joined into a frame-compatible sequence, only decimated horizontally and encoded independently with the HEVC Main profile. This approach is used as a reference as it is the most straightforward solution.
- HEVC Main side-by-side – both views decimated and packed next to each other and encoded with HEVC Main profile. This approach is commonly used in systems broadcasting stereoscopic video [Vetro'10]. Comparing it to simulcast would show if the HEVC encoder would benefit from frame-compatibility itself.
- HEVC SCC simulcast – both views are decimated and encoded individually with HEVC SCC to check if Screen Content Coding provides compression gain if the views are not accommodated into a single frame.
- HEVC SCC side-by-side – **the proposed solution**. Combines frame-compatibility and using Screen Content Coding.

All four encoders are compared in terms of encoding time and compression efficiency, represented as bitrate reduction at a constant quality of the reconstructed video (Chapter 3). The results are presented in Table 5.5 and Table 5.6.

Table 5.5. Bitrate reduction compared to HEVC Main simulcast for All Intra (AI) and Random Access (RA) coding scenarios. Negative values indicate lower bitrates at the same quality.

Sequence	Main side-by-side		SCC simulcast		SCC side-by-side	
	AI	RA	AI	RA	AI	RA
Balloons	0.03%	0.32%	0.27%	0.45%	-21.95%	-13.65%
BBB_Butterfly	-0.05%	-0.62%	0.28%	-0.13%	-25.70%	-19.92%
Kendo	0.07%	0.37%	0.23%	-0.35%	-23.36%	-16.05%
Newspaper	0.04%	-0.30%	0.13%	0.20%	-17.75%	-13.97%
Poznan_Hall2	0.04%	0.70%	-0.60%	-0.45%	-14.01%	-8.65%
Poznan_Street	0.06%	-0.11%	-0.89%	-0.59%	-20.49%	-16.23%
Average	**0.09%**	**0.12%**	**-0.33%**	**-0.15%**	**-20.07%**	**-14.70%**

Table 5.6. Encoding time reduction compared to HEVC Main simulcast for All Intra (AI) and Random Access (RA) coding scenarios. Negative values indicate faster encoding.

Sequence	Main side-by-side		SCC simulcast		SCC side-by-side	
	AI	RA	AI	RA	AI	RA
Balloons	2.82%	-4.94%	84.65%	1.74%	84.32%	-1.45%
BBB_Butterfly	-0.41%	6.20%	42.61%	-1.10%	15.20%	-6.53%
Kendo	6.12%	6.91%	68.17%	17.70%	55.26%	11.51%
Newspaper	3.93%	7.88%	125.22%	-3.66%	113.25%	-3.80%
Poznan_Hall2	2.07%	-5.90%	39.07%	-26.34%	30.37%	-30.11%
Poznan_Street	-1.27%	-3.68%	113.24%	-17.07%	103.75%	-20.92%
Average	**1.41%**	**-0.68%**	**78.48%**	**-9.60%**	**67.51%**	**-13.45%**

The conclusions from the results presented in Tables 5.5 and 5.6 are as follows:

- **The influence of frame-compatibility is negligible when using the HEVC Main encoder**, both in terms of bitrate and encoding time. This confirms that the inter-view similarities between the left and right views are not exploited by the standard HEVC encoder.
- **Using HEVC SCC for simulcast compression of camera-captured content does not affect the efficiency** and has a negative impact on encoding time. Such a result is expected because Screen Content Coding tools were designed for the compression of computer-generated content.

- **Application of HEVC SCC for compression of frame-compatible stereoscopic video provides a significant gain** of roughly 20% bitrate reduction for All Intra and 15% for Random Access, compared to other compression scenarios. This result indicates that Intra Block Copy can be successfully used to predict the inter-view similarities. In the case of All Intra, bitrate reduction is achieved at the cost of considerably longer encoding time, which is expected since the IBC and other SCC tools are additionally applied to every frame. On the other hand, for Random Access, the encoding time is reduced on average by 14%. In this case, intra frames are encoded longer due to the Intra Block Copy, but thanks to the higher quality of the reconstructed I-frame, the remaining B-frames are encoded much faster, compensating for the time increase of the I-frame.

5.7.3. Evaluation of SCC in multiview video coding

The experiment presented in this section aims to measure the impact of Screen Content Coding in the compression of multiview video and compare it to the state-of-the-art Multiview HEVC. The methodology of this experiment is similar to the previous one (Section 5.7.2). This time, however, 4 views of each sequence are chosen according to Table 5.7, and they are not decimated. The frame-compatible sequences are prepared by joining all 4 views horizontally, as described in Section 5.3. In addition, encoding with Multiview HEVC is performed.

Table 5.7. Chosen views of sequences used in the experiment.

Sequence	Views (from left to right)
Balloons	2, 3, 4, 5
BBB_Butterfly	40, 45, 50, 55
Kendo	2, 3, 4, 5
Newspaper	2, 4, 6, 8
Poznan_Hall2	7, 6, 5, 4
Poznan_Street	5, 4, 3, 2

Table 5.8 presents the results of bitrate reduction, compared to simulcast compression of all views using the HEVC Main profile, both for the All Intra (AI) and Random Access (RA) configuration. The results of encoding time reduction are gathered in Table 5.9.

Table 5.8. Bitrate reduction compared to HEVC Main simulcast for All Intra (AI) and Random Access (RA) coding scenarios. Negative values indicate lower bitrates at the same quality.

Sequence	**Main side-by-side**		**SCC simulcast**		**SCC side-by-side**		**Multiview**	
	AI	**RA**	**AI**	**RA**	**AI**	**RA**	**AI**	**RA**
Balloons	0.02%	0.29%	0.27%	0.27%	-32.35%	-20.88%	-42.59%	-36.66%
BBB_Butterfly	-0.09%	-0.84%	0.10%	-0.25%	-38.93%	-29.21%	-45.17%	-41.02%
Kendo	0.04%	0.16%	0.19%	-0.19%	-33.19%	-22.71%	-44.97%	-41.07%
Newspaper	0.35%	0.17%	0.06%	0.29%	-23.98%	-18.06%	-31.13%	-30.79%
Poznan_Hall2	0.00%	0.75%	-0.66%	-0.60%	-15.16%	-9.98%	-26.70%	-22.18%
Poznan_Street	0.06%	-0.06%	-0.73%	-0.55%	-23.49%	-19.27%	-34.70%	-40.19%
Average	**0.06%**	**0.08%**	**-0.13%**	**-0.17%**	**-27.85%**	**-20.02%**	**-37.54%**	**-35.32%**

Table 5.9. Encoding time reduction compared to HEVC Main simulcast for All Intra (AI) and Random Access (RA) coding scenarios. Negative values indicate faster encoding.

Sequence	**Main side-by-side**		**SCC simulcast**		**SCC side-by-side**		**Multiview**	
	AI	**RA**	**AI**	**RA**	**AI**	**RA**	**AI**	**RA**
Balloons	0.00%	-8.43%	42.75%	-9.48%	54.93%	-9.72%	120.80%	-2.76%
BBB_Butterfly	-4.77%	-2.44%	40.77%	-7.72%	11.85%	-5.08%	51.80%	10.99%
Kendo	3.68%	0.33%	58.40%	0.47%	74.57%	4.04%	152.19%	9.59%
Newspaper	-10.30%	0.61%	88.88%	-13.04%	93.24%	-9.03%	157.34%	13.56%
Poznan_Hall2	-3.24%	5.60%	34.42%	-24.10%	38.58%	-24.39%	119.57%	8.52%
Poznan_Street	-9.94%	5.81%	79.67%	-14.60%	110.56%	-14.40%	104.26%	5.18%
Average	**-4.09%**	**0.25%**	**57.48%**	**-11.41%**	**63.96%**	**-9.76%**	**117.66%**	**7.51%**

The conclusions for using frame-compatible HEVC Main and simulcast HEVC SCC are the same as for stereoscopic video compression: these encoders provide roughly the same efficiency as simulcast HEVC Main. Regarding frame-compatible compression with HEVC SCC, the bitrate reduction is even bigger than for stereoscopic video - on average, 28% for All Intra and 20% for Random Access. This is caused by the bigger impact of inter-view prediction due to more views being compressed. Nevertheless, the best results in terms of compression efficiency are achieved by Multiview HEVC. Bitrate reduction for this encoder is as high as 37% for All Intra and 35% for Random Access. On the other hand, the encoding time is much longer compared to frame-compatible HEVC SCC, which means that the proposed solution is less efficient in terms of compression efficiency, but much faster than the state-of-the-art Multiview HEVC.

5.7.4. EVALUATION OF SCC IN IMMERSIVE VIDEO CODING

This section evaluates the efficiency of Screen Content Coding applied to the compression of immersive video. The experiments are conducted using the publicly available MPEG Immersive Video test model [TMIV]. In the assessment, 97 frames of 7 immersive video test sequences are used. The atlases generated by the TMIV encoder are first compressed with the HEVC Main profile, which is the state-of-the-art approach in MIV coding. Then, the compression of atlases is repeated using HEVC SCC as proposed by the author of the dissertation in Section 5.6. The configuration of coding parameters is in line with MPEG Common Test Conditions for Immersive Video [CTC MIV], and it is the same for both encoders. Obviously, in the case of HEVC SCC, the Screen Content Coding tools are additionally configured according to MPEG Common Test Conditions for Screen Content Coding [Yu'15]. It should be stressed that Palette Mode is enabled, contrary to the experiments for stereoscopic and multiview video. As explained in Section 5.6, Palette Mode should be beneficial for the compression of depth maps; therefore, it is not disabled this time.

The results of the experiments are divided into two subsections. In the first subsection, the proposal is evaluated by comparing it to HEVC Main in terms of bitrate reduction and quality of reconstructed atlases after the coding cycle. Moreover, the results of the compression of views and depth atlases are presented and discussed separately. The second subsection focuses on the quality of virtual view synthesis performed using the reconstructed data. In this comparison, five commonly used objective quality metrics were calculated: Weighted-to-Spherically-Uniform PSNR (WS-PSNR) [Sun'17], Multi-Scale SSIM (MS-SSIM) [Wang'04], Visual Information Fidelity (VIF) [Sheikh'06], Video Multimethod Assessment Fusion (VMAF) [Li'16] and ISO/IEC MPEG's metric for immersive video: IV-PSNR [Dziembowski'22]. The results of such a comparison are more consistent with the subjective quality assessment of an end user of an immersive video system.

A. COMPRESSION EFFICIENCY OF ATLASES

Experimental results in this subsection are presented separately for views and depth data. First, the outcome of the comparison between HEVC Main and HEVC SCC in compression of views data is presented in Table 5.10. The results are divided based on the input data type (base views or patch atlases) and the content: computer-generated (CG) or natural content (NC). Figures 5.6 and 5.7 present rate-distortion curves of compression of views data for computer-generated and natural content, respectively.

Table 5.10. Bitrate reduction and quality improvement of HEVC SCC for views data, compared to HEVC Main. A positive number indicates a lower bitrate or better quality.

Sequence	Bitrate reduction [%]		Quality improvement [dB]	
	Base views	Patch atlases	Base views	Patch atlases
ClassroomVideo	0.78	1.74	0.02	0.04
TechnicolorMuseum	1.16	4.00	0.01	-0.01
TechnicolorHijack	0.00	6.55	0.03	0.22
OrangeKitchen	0.49	7.44	0.05	0.11
CG average	**0.61**	**4.93**	**0.03**	**0.09**
TechnicolorPainter	-0.09	1.53	-0.01	22.86
IntelFrog	0.44	1.39	0.00	33.00
Poznan_Fencing2	0.30	0.69	0.00	0.02
NC average	**0.22**	**1.20**	**0.00**	**18.63**
Average	**0.44**	**3.33**	**0.01**	**8.03**

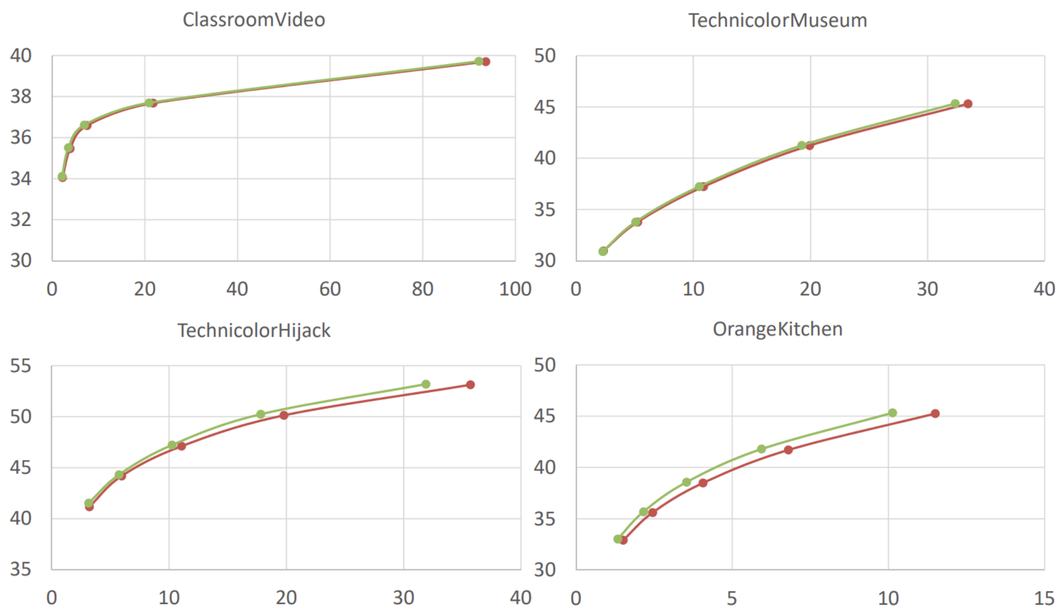

Figure 5.6. RD curves for compression of computer-generated sequences with HEVC (red) and HEVC SCC (green), views data. Horizontal axis – bitrate [Mbps], vertical axis – PSNR [dB].

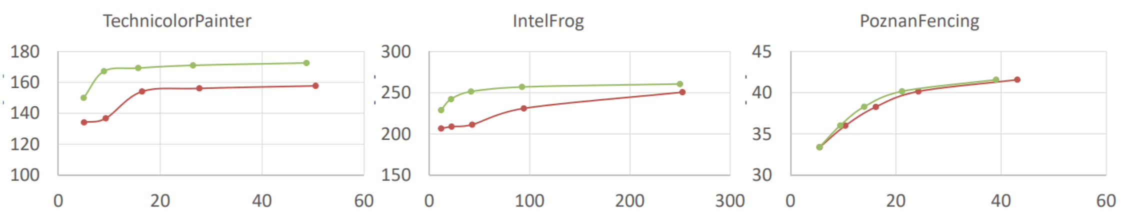

Figure 5.7. RD curves for compression of natural sequences with HEVC (red) and HEVC SCC (green), views data. Horizontal axis – bitrate [Mbps], vertical axis – PSNR [dB].

The results show that using HEVC SCC for compression of views data is more efficient than the state-of-the-art solution with HEVC Main profile. The difference in bitrate is more significant for patch atlases. In the case of quality, the difference is usually negligible, however, for patch atlases of two sequences, *TechnicolorPainter* and *IntelFrog*, the difference in quality is surprisingly high. Detailed investigation shows that such difference is caused by frames with no patches within sequences produced by the TMIV encoder. When using HEVC SCC, reconstruction images of such empty frames are ideal, which results in a very high PSNR value (99.99 dB) and strongly affects mean PSNR.

An analogous comparison of results is made for depth data. Table 5.11 presents bitrate reduction and quality improvement of base views and atlases encoded with HEVC SCC, compared to using HEVC Main. Rate-distortion curves for compression of depth data per sequence are presented in Figure 5.8 (computer-generated content) and Figure 5.9 (natural content). The numbers and plots clearly show that the proposed solution outperforms the state-of-the-art, both in terms of bitrate and quality of encoded video data. The gain is much higher than for views, which is expected due to the characteristics of depth maps. First of all, depth maps often contain large and smooth areas, as well as repeatable patterns, which can be efficiently predicted by the Intra Block Copy technique. Secondly, the values in fragments of depth maps (especially computer-generated) often belong to a very limited set, which can be utilized by Palette Mode.

Table 5.11. Bitrate reduction and quality improvement of HEVC SCC for depth data, compared to HEVC Main. A positive number indicates a lower bitrate or better quality.

Sequence	**Bitrate reduction [%]**		**Quality improvement [dB]**	
	Base views	**Patch atlases**	**Base views**	**Patch atlases**
ClassroomVideo	11.76	18.38	1.85	3.14
TechnicolorMuseum	8.52	13.12	0.56	0.62
TechnicolorHijack	7.89	9.25	1.09	1.43
OrangeKitchen	15.34	25.10	1.28	1.80
CG average	**10.88**	**16.47**	**1.20**	**1.75**
TechnicolorPainter	4.60	4.27	0.27	8.77
IntelFrog	2.01	3.16	0.12	32.32
Poznan_Fencing2	14.23	12.22	0.90	0.87
NC average	**6.95**	**6.55**	**0.43**	**13.99**
Average	**9.19**	**12.22**	**0.87**	**6.99**

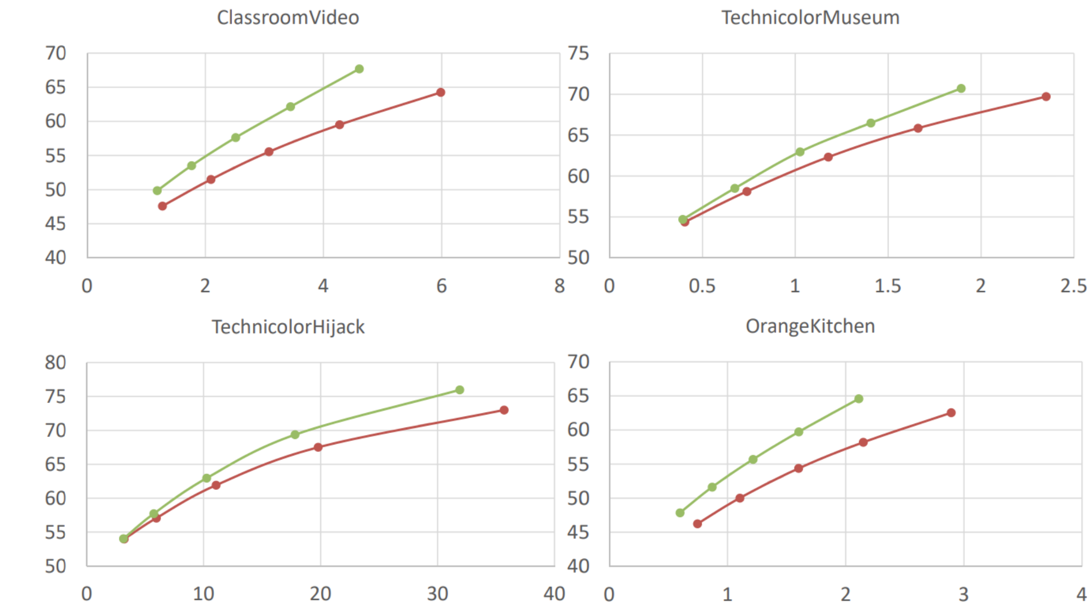

Figure 5.8. RD curves for compression of computer-generated sequences with HEVC (red) and HEVC SCC (green), depth data. Horizontal axis – bitrate [Mbps], vertical axis – PSNR [dB].

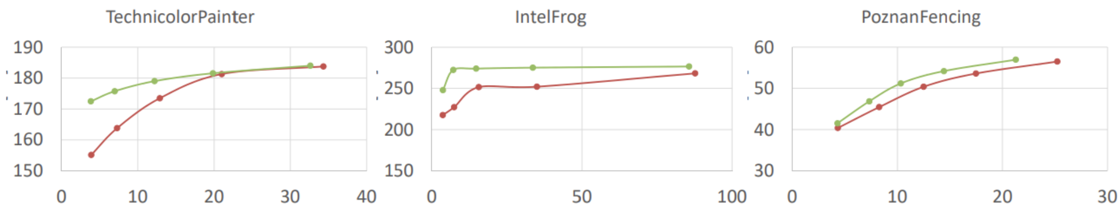

Figure 5.9. RD curves for compression of natural sequences with HEVC (red) and HEVC SCC (green), depth data. Horizontal axis – bitrate [Mbps], vertical axis – PSNR [dB].

Regarding differences between results for base views and patch atlases, it is observed that for computer-generated video, bitrate reduction is noticeably higher for patch atlases than for base views. In the case of natural content, however, the gain is comparable or even worse for patch atlases. Such an effect is a result of differences in creating depth maps. For computer-generated video, depth maps are usually rendered, and therefore they are smooth across whole objects and have sharp edges between objects and the background. On the contrary, depth maps for camera-captured video are usually estimated algorithmically from the input views, which causes some artifacts, e.g., grained objects or the background, blurred edges. These issues negatively affect the efficiency of both Intra Block Copy and Palette Mode. Nonetheless, the overall performance of standard-compliant HEVC Screen Content Coding in compression of atlases with depth maps and patches is much better than the commonly used HEVC Main profile.

B. View synthesis quality comparison

In the previous subsection, it was proven that HEVC SCC improves the compression efficiency and the quality of immersive video patch atlases, especially those containing depth data. However, the quality of depth does not concern end users of immersive video systems, but they care about the quality of synthesized video data that is generated from base views and depth maps and presented to end users. This subsection evaluates the quality of synthesized virtual views when using the proposed compression technique – HEVC SCC. It is compared to the synthesis quality achieved with video data compressed using the HEVC Main profile.

Figures 5.10 and 5.11 present rate-distortion plots for computer-generated and natural sequences, respectively. Bitrates (horizontal axis) are calculated as the sum of all atlases, while the quality (vertical axis) is represented by averaged WS-PSNR metric for the luma component of synthesized views. The results show that when using HEVC SCC, as proposed by the author of the dissertation, the quality of virtual views is higher at a given bitrate than when using standard HEVC.

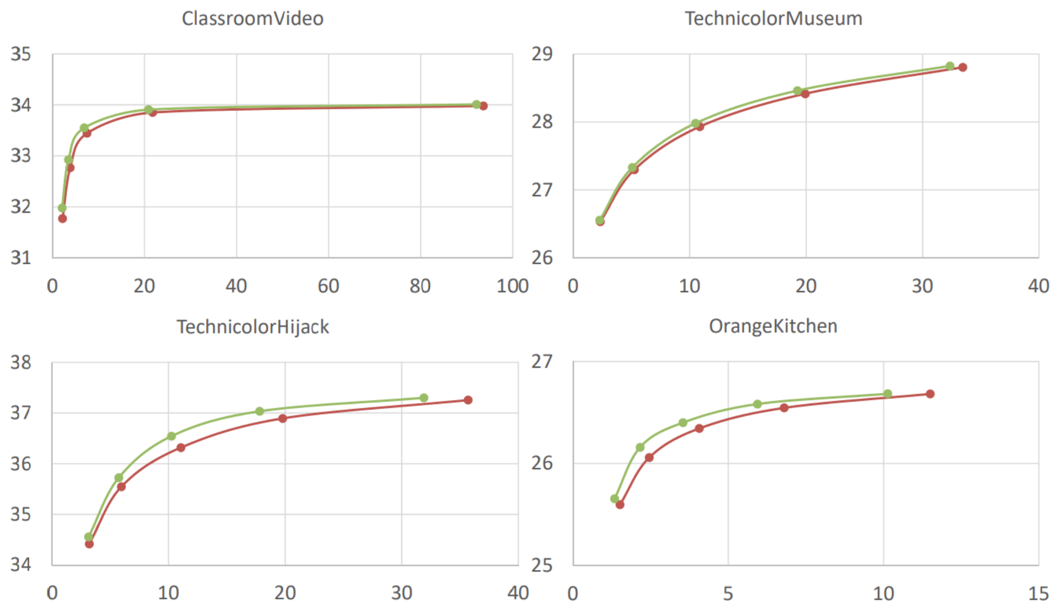

Figure 5.10. RD curves for view synthesis of computer-generated sequences with HEVC (red) and HEVC SCC (green). Horizontal axis – bitrate [Mbps], vertical axis – WS-PSNR [dB].

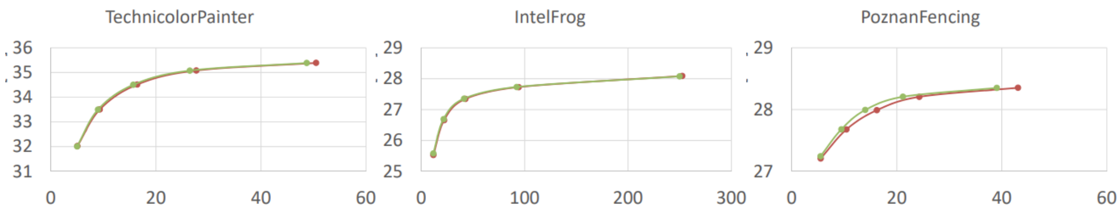

Figure 5.11. RD curves for view synthesis of natural sequences with HEVC (red) and HEVC SCC (green). Horizontal axis – bitrate [Mbps], vertical axis – WS-PSNR [dB].

Table 5.12 presents the average bitrate reduction calculated as Bjøntegaard Delta for WS-PSNR and four other commonly used objective quality metrics: VIF, VMAF, SSIM, and IV-PSNR. It can be noticed that the proposal is better than HEVC Main in all cases, and the gain can be as high as 30% (VMAF, sequence *OrangeKitchen*). Another observation is that the performance is better for computer-generated sequences than for natural content. This is in line with the compression efficiency results, where computer-generated depth maps were encoded more efficiently. As mentioned before, depth maps of higher quality allow to synthesize virtual views with fewer artifacts, thus higher quality.

Table 5.12. Bitrate reduction for different quality metrics.
Positive values indicate bitrate reduction for a given quality metric.

Sequence	**WSPSNR**	**VIF**	**VMAF**	**SSIM**	**IVPSNR**
ClassroomVideo	22.36	10.87	16.42	10.28	10.24
TechnicolorMuseum	7.82	4.07	8.77	4.48	4.78
TechnicolorHijack	19.83	14.91	20.06	15.88	9.20
OrangeKitchen	22.64	16.35	30.19	16.30	5.65
CG average	**18.16**	**11.55**	**18.86**	**11.74**	**7.47**
TechnicolorPainter	3.37	3.75	2.92	3.37	3.33
IntelFrog	3.85	2.70	5.04	4.14	1.48
Poznan_Fencing2	11.41	11.18	10.23	10.31	9.46
NC average	**6.21**	**5.88**	**6.06**	**5.94**	**4.76**
Average	**13.04**	**9.12**	**13.38**	**9.25**	**6.31**

Figure 5.12 presents the comparison of fragments of virtual views synthesized with original data (left), reconstructed data after compression with HEVC Main (middle), and reconstructed data after compression with HEVC SCC (right). It can be observed that the proposed solution results in fewer artifacts in view synthesis, especially at the edges between objects, which has a positive influence on the subjective quality when viewed by the end user of an immersive video system.

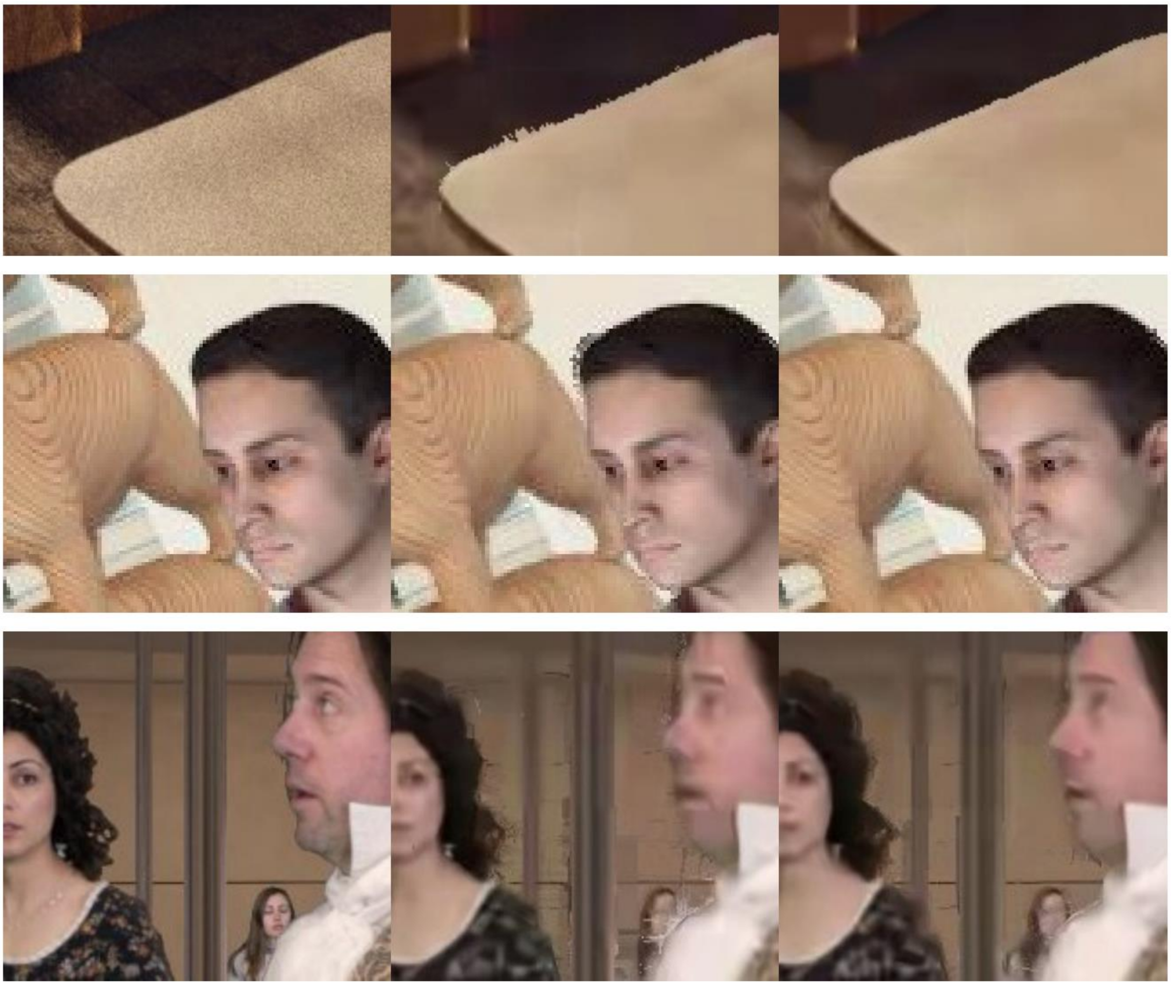

Figure 5.12. Fragments of: input views (left), views synthesized using data encoded with HEVC Main (middle), and views synthesized using data encoded with HEVC SCC (right).

5.8. CONCLUSIONS

In this chapter, the author of the dissertation considers using Screen Content Coding as an inter-view compression technique and proposes methods for using it in several new applications: stereoscopic, multiview, and immersive video coding. The proposal is a novel approach, and using it to compress camera-captured content may appear unintuitive because Screen Content Coding was developed to compress computer-generated content. Although SCC does not improve the compression efficiency of a single view containing camera-captured content, combining Intra Block Copy technique with frame-compatibility turns out to be an efficient way to exploit the inter-view similarities. Moreover, the author evaluates the influence of view alignment in frame-compatible sequence, as well as the impact of SCC coding tools other than IBC, on the overall performance of Screen Content Coding in the compression of natural content. The results help in preparing optimal configurations of the proposal and applying it in experiments presented in Section 5.7.

A complete proposal assessment is performed by conducting several experiments, comparing HEVC Screen Content Coding with state-of-the-art solutions for stereoscopic, multiview, and immersive video compression.

In the case of stereoscopic video coding, the proposed solution is compared to the HEVC Main profile in compression of decimated views joined side-by-side, which is a common approach in broadcasting so-called "3D video" content. The results show that **HEVC SCC provides, on average, 15% lower bitrate than HEVC Main profile**. Taking into account that HEVC SCC can be efficiently applied to screen content, and by design is less complex than Multiview HEVC (which is a multi-layer codec), using SCC for stereoscopic video compression in practical systems is possible. Moreover, support for Screen Content Coding was already provided in the recently developed video coding standard, Versatile Video Coding, therefore the proposed approach could be easily applied in the new generation of codecs.

Regarding multiview video, **using HEVC SCC in compression of 4 views provides roughly 20% bitrate reduction compared to simulcast encoding**. However, the gain is significantly lower than when using the state-of-the-art Multiview HEVC encoder. It should be noted that this chapter only describes encoding with standard-compliant HEVC SCC, which is not optimized for efficient compression of camera-captured multiview video content. In the next chapter (Chapter 6), the author of the dissertation proposes several modifications of standard HEVC SCC to adapt it to the new applications and achieve as high bitrate savings as the dedicated Multiview HEVC extension.

Finally, the proposed HEVC SCC method is applied to immersive video coding. MPEG Immersive Video (MIV) standard uses a general video encoder internally to compress atlases. MIV is, however, codec-agnostic, so replacing HEVC with HEVC SCC could be done effortlessly and without violating the standard. The evaluation of such change also proves the superiority of the proposal over the HEVC Main profile, especially in the compression of computer-generated content and depth maps. Further assessment shows that the improved quality of depth maps results in **a significantly better quality of synthesized virtual views** – the proposal is compared to HEVC Main with 5 different quality metrics, and it turns out to be **noticeably more efficient in all cases**.

The idea of using SCC as an internal video encoder in MIV was proposed by the author during one of MPEG meetings [Samelak'20A-B]. As a result, MPEG group investigated such a possibility by evaluating using VVC with Screen Content Coding as an internal MIV encoder, compared to plain VVC. Experimental results confirmed that the application of Screen Content

Coding is beneficial in the compression of immersive video data [Vadakital'20]. Therefore, **the novel approach proposed by the author of the dissertation remains valid for the forthcoming generation of video coding standards**.

6. ADVANCED SCREEN CONTENT CODING

6.1. INTRODUCTION

In the previous chapter (Chapter 5), the author of the dissertation proposes using Screen Content Coding for efficient compression of frame-compatible multiview and immersive video. The main idea for such change is to exploit similarities between the views with Intra Block Copy. The experiments presented in Section 5.7 prove that the proposed approach significantly improves the compression efficiency of frame-compatible video compared to the HEVC Main encoder. Nevertheless, the proposal is less efficient than the state-of-the-art dedicated solution for multiview compression – Multiview HEVC (Section 5.7.3). Despite that, HEVC SCC has a few important advantages over Multiview HEVC, such as shorter encoding time, less complex implementation, and broader versatility (one encoder for multiple applications).

In this chapter, the author of the dissertation focuses on **improving the efficiency of Screen Content Coding in the compression of multiview and immersive video**. The author proposes a novel approach to multiview video compression, including **adapting the configuration of HEVC SCC** to the new application and **a set of improvements** aimed at increasing the compression efficiency of frame-compatible multiview video and immersive video. The goal of the work is to prepare a **competitive alternative for the state-of-the-art Multiview HEVC** with comparable compression efficiency and better usability. In the dissertation, the encoder with the author's improvements is called **Advanced Screen Content Coding (ASCC)**. The ASCC encoder is also applied to immersive video coding, and the author proposes a MIV metadata parser for controlling tools, depending on the type of input video, to achieve the highest compression efficiency.

6.2. IMPROVEMENTS OF SCC FOR MULTIVIEW VIDEO COMPRESSION

6.2.1. FRAME-COMPATIBILITY WITH SPECIFIC ORDERING

As explained in Section 5.2, the views of a multiview video have to compose a single frame to allow the Intra Block Copy technique to take advantage of the similarities between them. In Section 5.3, different view alignments are compared, which results in choosing horizontal view alignment as the most efficient. In this section, the author of the dissertation proposes to change the order of the views that compose a frame-compatible sequence. The preferred order for compression of 3 views is middle-leftmost-rightmost, as presented in

Figure 6.1. Such an ordering reflects the most common approach in Multiview HEVC, where the middle view is encoded as the first and then becomes a reference for inter-view prediction of the remaining views [Müller'14]. Obviously, the information about the order of encoded views has to be included in the bitstream to allow the decoder to properly organize the output after decoding and splitting the frame-compatible video into separate views. The author includes such information in the bitstream as an extension of the Video Parameter Set (VPS) [ISO'21].

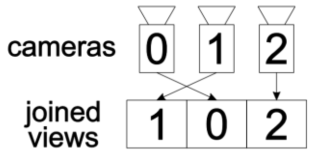

Figure 6.1. Order of positioning 3 cameras in a scene and joining acquired views.

6.2.2. TILE ENCODING

When processing a video using the default configuration, the HEVC encoder divides each frame into so-called Coding Tree Units (CTUs) and then compresses them in rows from left to right, starting from the top left CTU. Considering a frame-compatible format, the top rows of each view are encoded first, then the second rows, etc. As explained in Section 2.3, the reference area for Intra Block Copy has to be restricted to the part of the frame that was already compressed, otherwise, it would not be possible to reproduce the prediction at the decoder side. Therefore, in the unmodified SCC compression of frame-compatible multiview video, Intra Block Copy cannot perform inter-view prediction by matching the area of another view below the vertical position of the compressed unit. Due to such limitations, the efficiency of IBC as an inter-view prediction technique may be deteriorated.

As a solution for this issue, the author of the dissertation proposes to apply compression in tiles, as presented in Figure 6.2. In the proposal, the division of a frame into tiles is analogous to the accommodation of views within a frame-compatible video, which means that every view corresponds to a single tile. In such a configuration, the leftmost tile (which corresponds to the middle view according to the proposed accommodation of views presented in Section 6.2.1) is compressed in whole before the compression of the remaining views begins. Therefore, the restriction for IBC on vertical position does not apply if the block matching is performed within the area of the previously encoded tile. This way, Intra Block Copy can search for the best matching block of samples within a bigger area, even below the vertical position of the currently processed unit. However, it should be mentioned that the application of the proposed solution introduces a dependency between tiles, which disables the capability of parallel processing of tiles.

Figure 6.2. HEVC coding order without (above) and with (below) tile encoding.

6.2.3. Intra Block Copy vectors precision

As described in Section 2.3, Intra Block Copy produces a block vector that points to the best matching block of samples within the reconstructed part of the current picture. The resulting vector's precision is full-pel because, in the case of computer-generated images, the benefit from using sub-pel precision is usually negligible or even negative, while the processing time is noticeably longer. The author of the dissertation proposes to apply Intra Block Copy as an inter-view prediction technique for multiview camera-captured content, for which performing a sub-pel block matching could be beneficial. The state-of-the-art dedicated multiview video coding technique, Multiview HEVC, supports inter-view prediction at a quarter-pel level. Therefore, for a fair comparison, the author of the dissertation improves the Intra Block Copy by implementing a quarter-pel block matching. Obviously, such a change can affect the efficiency of Advanced SCC in the compression of screen content. The impact of the proposed modification on screen content compression is evaluated in Section 6.4.3.

6.2.4. Starting point for block matching in Intra Block Copy

Application of Screen Content Coding for compression of frame-compatible multiview video aims at utilizing Intra Block Copy as an inter-view prediction technique, and tile encoding allows the IBC to use the whole previously encoded tile as a reference. The expected result of Intra Block Copy search is a prediction vector that points to the best matching block of samples in a different tile (Figure 6.3). Such a vector is relatively long, which is not desired due to the relatively high cost of representing it in the bitstream. Moreover, Intra Block Copy has to search distant areas to find the best matching block of samples, which can deteriorate the processing time.

Figure 6.3. Long inter-view prediction vector derived by Intra Block Copy.

In the ASCC, the author of the dissertation proposes to change the starting point for the Intra Block Copy search to the position of the collocated unit in the leftmost tile. This way, the probability of finding the expected best-matching block of samples in the reference view is higher, and the IBC processing time is shorter. Additionally, the distance between the processed unit and the collocated one in the reference tile is subtracted from the horizontal component of the block vector returned by Intra Block Copy. This means that the block vector (0, 0) indicates the collocated position in the reference view instead of the current position. Such an approach reduces the average length of vectors found by the Intra Block Copy algorithm, which is beneficial for the overall compression efficiency.

6.2.5. IN-LOOP FILTERING PER TILE

As mentioned, Intra Block Copy searches for the best matching block of samples in the previously encoded area of the same picture. Therefore, the prediction is made on a reconstructed part of a frame before the in-loop filtering. Loop filters are executed at the end of the compression of each slice to reduce the encoding artifacts and thus improve the quality of the reconstructed picture, which can then be used as a reference for the compression of the following frames [Sullivan'12].

In the proposed Advanced SCC, the author of the dissertation applies the in-loop filtering after the compression of each tile. This improves the quality of the reference view (which is represented by a single tile) and allows Intra Block Copy to predict the content of the side views more accurately. Moreover, HEVC allows choosing whether the in-loop filters should be applied at the boundaries between tiles [ISO'21]. In the frame-compatible multiview video, the sharp edges between the views are intended, therefore the filtering at the boundaries of tiles is disabled to preserve them.

6.2.6. Different Quantization Parameter for side views

The compression level in HEVC is mainly controlled by Quantization Parameter (QP). The higher the QP, the lower the bitrate, but at the same time, the quality of decoded video becomes worse. In Multiview HEVC, the QP can be set individually for each view, which is the case in the default configuration in Common Test Conditions for Multiview HEVC [Müller'14]. In order to prepare a fair comparison, the author of the dissertation adds to the ASCC a possibility to specify Quantization Parameter per tile. It is a non-standard information that must be provided in the bitstream to allow the decoder to decompress it correctly. In the proposal, the difference between the QP of the leftmost tile and another one is included in the Video Parameter Set extension.

6.2.7. Reference tile border extension

When a Prediction Unit (PU) is close to the right border of a tile, the search area of Intra Block Copy can cover the border between tiles, as well as a part of a tile next to the reference one (Figure 6.4). Therefore, the prediction error of matching a PU with a block of samples that contains parts of two tiles is very high.

In the Advanced SCC, the author of the dissertation implements an algorithm that extracts the reconstructed reference tile and extends its borders by interpolating the values from the edge of a tile. Such a solution results in better prediction than matching the boundaries of two tiles and is in line with inter-frame prediction, where the borders of reference views are also extended in a similar manner.

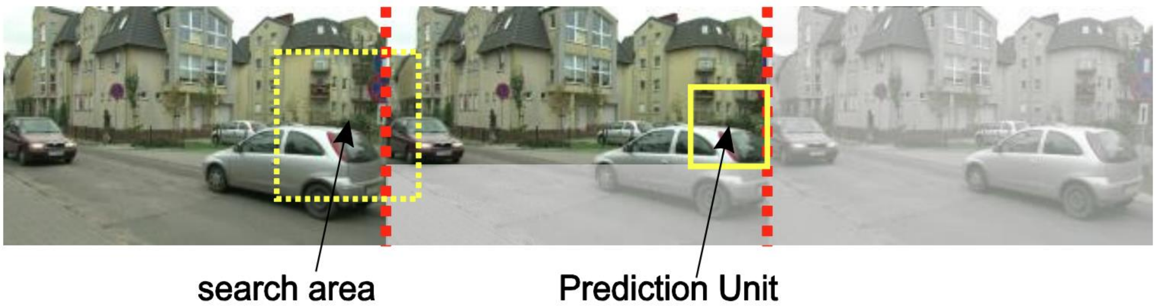

Figure 6.4. Intra Block Copy search at the border of two tiles.

6.3. Improvements of SCC for immersive video compression

6.3.1. Adaptation of Advanced SCC to immersive video

The improvements of Screen Content Coding (collectively called Advanced SCC) for compression of frame-compatible multiview video, presented in Section 6.2, are also applied to immersive video coding. However, some of the proposed improvements cannot be applied in the same manner as for multiview video, while others cannot be applied at all. The list of changes compared to ASCC for multiview video is as follows:

- Ordering of view accommodation in frame-compatible input video (Section 6.2.1) is not changed. In the case of immersive video, the views are stacked into atlases directly by the TMIV encoder. Modifications of TMIV are out of the scope of the dissertation. The video data produced by TMIV is encoded with ASCC "as is", therefore no manipulation with the ordering of views is done.
- Starting point for IBC block matching (Section 6.2.4) is shifted vertically instead of horizontally. Since the views in atlases are stacked vertically, the Intra Block Copy search should start from the position of a collocated block in the topmost tile instead of the leftmost. Also, in this case, the vertical distance between the current and collocated position is subtracted from the vertical component of the Intra Block Copy vector.
- Different Quantization Parameters for side views (QPs) are not applied. All tiles composing a frame are compressed with the same QP. This is to assert that the view synthesis is performed using base views of similar quality.
- Enabled Palette Mode. As presented in Chapter 5, Palette Mode can be beneficial for the compression of depth maps, therefore it was enabled for immersive video coding using Advanced Screen Content Coding.

The remaining improvements prepared for the frame-compatible multiview video are applied to the compression of immersive video with no changes.

6.3.2. Using MIV metadata to control the modifications

As described in Section 2.4, MPEG Immersive Video encoder reduces the inter-view similarities and forms the remaining video data into so-called patch atlases, which are then compressed individually with general video encoders such as HEVC or, as proposed in Chapter 5, HEVC SCC. The atlases, however, have different characteristics, and therefore the efficiency of the encoder may differ depending on the type of atlas provided at the input. This observation was experimentally proven in Section 5.7.4, where the benefit from using HEVC SCC was, e.g., greater for atlases with depth patches than views patches.

In this section, on top of the ASCC modifications, the author proposes to incorporate a dedicated metadata parser into the MIV coding scheme for controlling the configuration of video encoders, depending on the type of input atlas. Figure 6.5 presents a block diagram of the proposed modified MIV encoder. In the state-of-the-art MIV encoder, metadata is directly included in the resulting bitstream. At the same time, in the proposal, it is additionally parsed and can be used to configure such options as the precision of Intra Block Copy vectors (full-pel or quarter-pel), configuration for tile encoding, enabling Palette Mode for depth atlases, etc. Another advantage of such a solution is that it doesn't require any additional signaling because metadata is already present in the MIV bitstream.

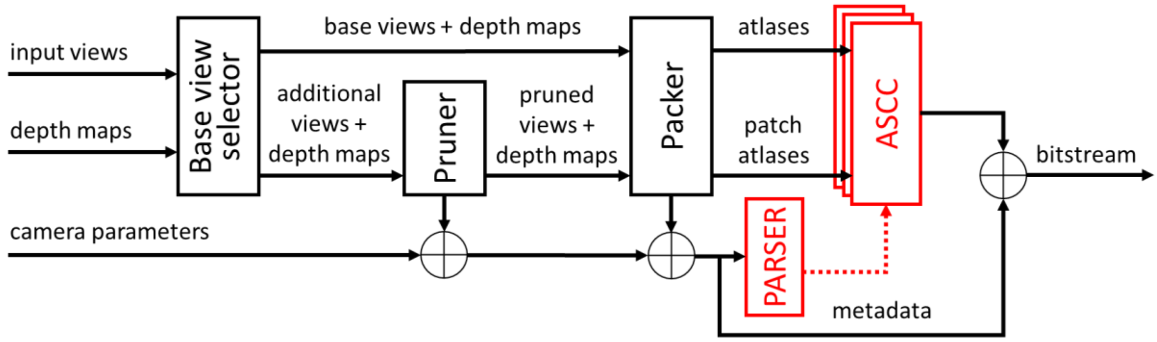

Figure 6.5. MPEG Immersive Video coding scheme with ASCC encoder controlled by MIV metadata.

The evaluation of Advanced SCC modifications in the compression of different types of atlases is presented in Section 6.4, along with comments on how the encoders should be set up for each atlas to achieve the highest compression efficiency.

6.4. Evaluation of the proposal

6.4.1. Goals of the experiments

In this chapter, the author's several modifications of Screen Content Coding are described. The goal of these changes is to better adapt SCC to new applications, namely multiview and immersive video compression. The proposed improvements are implemented in a publicly available test model for the HEVC SCC codec (details in Chapter 3). In the case of multiview video coding, the goal of the author's proposal is to achieve compression efficiency comparable to Multiview HEVC, which is the dedicated coding technique for multiview video content. A comparison of the proposal with the state-of-the-art is presented in Section 6.4.2. Additionally, the author of the dissertation conducts an experiment that evaluates the influence of the proposed modifications in the compression of screen content (Section 6.4.3).

Regarding the application of ASCC to immersive video coding, the main goal of the experiments (presented in Section 6.4.4) is to assess the compression efficiency of different types of video data generated by the MIV encoder. Based on the results, the author proposes a novel approach to immersive video coding that would adjust the coding tools to the input video data to increase the compression efficiency and the quality of synthesized video presented to end users.

The complete results of the conducted experiments can be found in the Appendix.

6.4.2. Evaluation of Advanced SCC in multiview video coding

In this section, the proposed ASCC is compared to Multiview HEVC in terms of encoding time and compression efficiency. The base configuration of both encoders is set up according to the appropriate Common Test Conditions as described in Chapter 3, with some additional changes. Regarding the configuration of ASCC, tools specific to Screen Content Coding are configured as proposed in Section 5.4. Moreover, the configuration is adapted to tile encoding and modified in-loop filtering as proposed in Sections 6.2.2 and 6.2.5. Regarding Multiview HEVC configuration, the vertical range limit for disparity search is set to 64 (from default 128), and Early Skip Detection is enabled to make the configuration fair and consistent with HEVC SCC. The experiments are conducted using 3 views of 8 multiview sequences, chosen and ordered according to Common Test Conditions for multiview video coding. The evaluation is performed in All Intra and Random Access coding scenarios.

Experimental results of the comparison are presented in Tables 6.1 – 6.4. Negative numbers indicate better compression efficiency (lower bitrate at constant quality) or shorter encoding time.

Table 6.1. Bitrate reduction [%] against HEVC simulcast and MV-HEVC; All Intra.
Negative values indicate lower bitrates.

Sequence	HEVC simulcast				MV-HEVC			
	2 views		3 views		2 views		3 views	
	SCC	**ASCC**	**SCC**	**ASCC**	**SCC**	**ASCC**	**SCC**	**ASCC**
Poznan_Hall2	-16.79	-28.01	-21.08	-40.57	17.24	2.93	27.36	-1.67
Poznan_Street	-20.66	-31.57	-27.53	-47.35	15.30	2.69	29.05	-0.55
Kendo	-20.25	-29.19	-26.33	-40.76	14.10	2.19	22.34	0.06
Balloons	-21.49	-29.29	-30.08	-41.90	12.01	1.80	18.27	0.00
Newspaper	-18.32	-25.37	-25.51	-39.06	9.75	1.87	18.56	-0.35
Dancer	-35.14	-39.98	-47.16	-55.71	8.86	1.25	17.12	-0.79
GT_Fly	-38.56	-42.27	-50.71	-58.22	6.85	0.62	15.36	-1.54
Shark	-35.39	-40.50	-49.78	-57.29	8.89	0.64	15.68	-0.97
Average	**-26.28**	**-34.21**	**-34.56**	**-48.52**	**12.47**	**1.94**	**22.25**	**-0.90**

Table 6.2. Encoding time reduction [%] against HEVC simulcast and MV-HEVC; All Intra.
Negative values indicate faster encoding.

Sequence	HEVC simulcast				MV-HEVC			
	2 views		3 views		2 views		3 views	
	SCC	**ASCC**	**SCC**	**ASCC**	**SCC**	**ASCC**	**SCC**	**ASCC**
Poznan_Hall2	+56	+8	+63	+22	+43	-1	+32	-5
Poznan_Street	+170	+38	+183	+48	+79	-9	+67	-12
Kendo	+169	+117	+183	+147	+25	+1	+11	-4
Balloons	+195	+121	+212	+148	+35	+1	+24	-1
Newspaper	+242	+131	+259	+151	+51	+2	+37	-3
Dancer	+330	+30	+319	+31	+224	-2	+204	-5
GT_Fly	+150	+11	+126	+4	+124	+0	+111	-3
Shark	+180	+35	+142	+27	+111	+2	+87	-2
Average	**+175**	**+41**	**+175**	**+50**	**+99**	**-2**	**+85**	**-5**

Table 6.3. Bitrate reduction [%] against HEVC simulcast and MV-HEVC; Random Access. Negative values indicate lower bitrates.

Sequence	HEVC simulcast				MV-HEVC			
	2 views		3 views		2 views		3 views	
	SCC	ASCC	SCC	ASCC	SCC	ASCC	SCC	ASCC
Poznan_Hall2	-11.30	-23.55	-13.82	-32.09	13.47	-1.92	25.43	-0.63
Poznan_Street	-13.88	-25.78	-19.29	-38.91	16.73	1.43	30.82	0.51
Kendo	-10.30	-19.45	-13.80	-27.52	12.37	1.45	18.39	0.37
Balloons	-13.26	-21.38	-18.44	-30.11	11.28	1.34	16.19	0.33
Newspaper	-16.00	-23.55	-20.46	-33.04	8.93	-0.03	17.74	0.46
Dancer	-25.82	-35.02	-34.32	-47.59	11.89	-1.69	24.67	0.05
GT_Fly	-21.74	-31.09	-31.67	-46.23	13.87	0.21	27.18	0.05
Shark	-29.50	-37.76	-40.39	-52.22	12.82	-0.10	23.93	-0.18
Average	**-16.61**	**-26.98**	**-22.58**	**-38.47**	**13.67**	**-0.11**	**25.30**	**0.07**

Table 6.4. Encoding time reduction [%] against HEVC simulcast and MV-HEVC; Random Access. Negative values indicate faster encoding.

Sequence	HEVC simulcast				MV-HEVC			
	2 views		3 views		2 views		3 views	
	SCC	ASCC	SCC	ASCC	SCC	ASCC	SCC	ASCC
Poznan_Hall2	+80	-5	+77	-2	+88	-1	+80	+0
Poznan_Street	+141	+22	+140	+23	+102	+2	+99	+2
Kendo	+169	+45	+127	+47	+88	+1	+56	+1
Balloons	+190	+45	+183	+45	+100	+0	+92	-1
Newspaper	+184	+47	+189	+50	+97	+2	+94	+0
Dancer	+175	+17	+176	+17	+137	+1	+138	+1
GT_Fly	+162	+10	+142	+7	+134	-2	+123	-1
Shark	+164	+30	+123	+28	+104	+0	+74	+0
Average	**+145**	**+18**	**+132**	**+18**	**+110**	**+0**	**+99**	**+0**

Experimental results confirm that the proposed modifications significantly improve both compression efficiency and encoding time compared to the unmodified HEVC SCC. For Random Access, **encoding 3 views with ASCC results in a nearly 40% smaller bitrate** compared to simulcast, at the cost of around 20% longer encoding time. Compared to Multiview HEVC, the proposed solution is **as efficient as a dedicated technique with no significant difference in encoding time**. The above results prove that such a novel approach to multiview compression could substitute the state-of-the-art complex codec.

6.4.3. Evaluation of Advanced SCC in screen content coding

As mentioned earlier in this chapter, the author's modifications to SCC are supposed to improve the compression efficiency of camera-captured multiview video. Therefore, some of the proposed changes may negatively impact the compression of screen content. In this section, this influence is evaluated. Advanced SCC is used for the compression of 13 test sequences commonly used in the evaluation of HEVC SCC. Since only one view is encoded, some of the ASCC modifications (e.g., frame-compatibility, tile encoding) do not apply.

As in previous experiments, tests are performed in All Intra and Random Access coding scenarios. The configuration is in line with Common Test Conditions for screen content coding, including enabled Palette Mode. The results of the comparison between ASCC and unmodified SCC are presented in Table 6.5.

Table 6.5. Experimental results for ASCC against HEVC-SCC in compression of single view screen content.

Sequence	All Intra		Random Access	
	Bitrate [%]	Encoding time [%]	Bitrate [%]	Encoding time [%]
Basketball_Screen	3.90	+22	2.60	+10
ChinaSpeed	0.58	+14	0.32	+23
ChineseEditing	3.95	+13	3.50	+12
MissionControlClip2	0.56	+19	0.42	+10
MissionControlClip3	3.05	+20	2.38	+10
sc_console	15.07	+30	8.34	+15
sc_desktop	13.40	+23	8.82	+13
sc_flyingGraphics	9.07	+20	5.38	+31
sc_map	2.35	+16	1.42	+10
sc_programming	4.61	+18	1.89	+13
sc_robot	0.28	+13	0.10	+20
sc_web_browsing	15.24	+26	11.27	+18
SlideShow	-0.98	+23	-0.77	+33
Average	**5.47**	**+20**	**3.51**	**+17**

The results indicate that ASCC is slower by roughly 20% than the unmodified SCC. That is because ASCC performs Intra Block Copy search at quarter-pel precision, while originally it is limited to full-pel. Obviously, this introduces overhead in encoding time. Regarding compression efficiency, the sub-pel accuracy of IBC is not beneficial for computer-generated content. However, the results strongly depend on the content of the video. As observed, the compression efficiency of ASCC is close to the unmodified SCC for sequences that contain a lot of fluent motion, gradients, parts with camera-captured content, or computer-generated images that are supposed to imitate natural images. On the other hand, the bitrate is significantly increased for the compression of mostly computer-generated content with simple graphics and motion. In this case, full-pel precision from the original SCC seems enough, and further improving it to quarter-pel does not improve the IBC prediction but generates a significant amount of redundant bits. Therefore, if the ASCC would be applied both to compression of multiview video and screen content, the accuracy of Intra Block Copy search would need to be dynamically adjusted by the encoder to optimize the efficiency for various types of input data.

6.4.4. Evaluation of Advanced SCC in immersive video coding

This section presents the evaluation of ASCC adapted to immersive video coding as described in Section 6.3. The tests are performed according to Common Test Conditions for immersive video [CTC MIV] on 17 frames of 6 camera-captured and 9 computer-generated sequences. The state-of-the-art Test Model for MIV [TMIV] was used in the experiments. It should be noted that the process of preparing atlases is not changed; only the internal video encoder is replaced with ASCC and compared to HEVC Main in terms of encoding time and quality of video synthesized from reconstructed data after the encoding and decoding cycle. The assessment of quality is performed using two full-reference metrics: WS-PSNR and IV-PSNR, described in Chapter 3. Then, bitrate reduction represented as Bjøntegaard delta is calculated for both metrics.

The experiments are divided into two sections. Section A presents experimental results of using only quarter-pel IBC precision in compression of all types of atlases, while Section B additionally evaluates the efficiency of tile encoding.

A. QUARTER-PEL IBC ACCURACY

In order to fully evaluate the influence of IBC accuracy on immersive video coding, three experiments are performed. In the first experiment, quarter-pel accuracy is compared to full-pel used by unmodified Screen Content Coding. The results of such a comparison are presented in Tables 6.6 – 6.7. Table 6.8 gathers encoding results of the remaining experiments: in one of them, quarter-pel accuracy is applied only to depth data, while in the other, only to atlases of views.

Table 6.6. Results for compression of SCC with quarter-pel IBC accuracy compared to basic SCC. Negative values indicate lower bitrates at the same quality or faster encoding.

Sequence	BD-rate WS-PSNR$_Y$	BD-rate IV-PSNR	Δ Encoding time
ClassroomVideo	– 0.3%	+0.7%	+7.2%
TechnicolorMuseum	+0.6%	+0.9%	+8.7%
Fan	+0.3%	+0.5%	+3.7%
OrangeKitchen	+0.1%	– 0.2%	+6.3%
Chess	– 0.5%	– 1.3%	+4.3%
Group	+0.6%	+0.5%	+8.0%
ChessPieces	– 1.0%	– 3.0%	+8.3%
TechnicolorHijack	– 0.1%	– 0.5%	+13.2%
Mirror	+0.5%	+0.0%	+1.5%
CG average	**+0.0%**	**– 0.3%**	**+6.8%**
TechnicolorPainter	– 1.6%	– 1.9%	+10.4%
IntelFrog	+0.1%	– 0.0%	+7.5%
Poznan_Carpark	– 1.3%	– 1.4%	+8.7%
Poznan_Fencing2	– 2.0%	– 2.4%	+7.2%
Poznan_Hall2	– 6.0%	– 7.3%	+6.2%
Poznan_Street	– 2.9%	– 3.1%	+8.6%
NC average	**– 2.3%**	**– 2.7%**	**+8.1%**
Average	**– 0.9%**	**– 1.2%**	**+7.3%**

The results presented in Table 6.6 show that using quarter-pel IBC search accuracy instead of full-pel is more efficient when encoding camera-captured content. In the case of computer-generated content, the influence on the efficiency is much smaller for IV-PSNR and negligible for WS-PSNR. In terms of encoding time, more accurate IBC block matching results in longer processing, with an average increase of 7% compared to unmodified Screen Content Coding.

Table 6.7 presents the results separately for views and depth data and shows an average bitrate reduction at each of the four rate points. Additionally, the values of IV-PSNR quality for synthesized views are compared. As observed, a more accurate IBC vector search does not impact the quality of synthesized views nor the bitrates of compressed computer-generated content. However, a significant gain was observed for the compression of depth data estimated for camera-captured sequences.

Table 6.7. Results for compression of SCC with quarter-pel IBC accuracy compared to basic SCC, bitrate, and quality change for different rate points. Rate points are defined in CTC separately for each sequence, therefore they are denoted as R1-R4.

Content type	**Rate**	**Views bitrate [Mbps]**			**Depth bitrate [Mbps]**			**IV-PSNR [dB]**	
		FPel	**QPel**	**Δ**	**FPel**	**QPel**	**Δ**	**FPel**	**QPel**
CG	R1	58.95	58.38	– 1%	7.38	7.35	0%	44.31	44.31
	R2	15.81	15.72	– 1%	5.39	5.37	0%	43.04	43.05
	R3	5.01	5.07	1%	3.72	3.71	0%	40.89	40.90
	R4	1.65	1.66	1%	2.78	2.78	0%	38.14	38.14
NC	R1	79.46	79.02	– 1%	31.07	28.95	– 7%	45.41	45.42
	R2	16.59	16.53	0%	13.81	12.74	– 8%	43.92	43.93
	R3	5.42	5.51	2%	6.19	5.90	– 5%	41.76	41.75
	R4	2.02	2.04	1%	3.76	3.69	– 2%	38.46	38.47

Since the results in Table 6.7 indicate a significant difference in bitrate reduction between depth and views atlases, another experiment is conducted to assess the performance of mixed full-pel and quarter-pel solutions. Table 6.8 gathers the results of two tests: in the first one, the quarter-pel accuracy is applied only to the compression of depth atlases, while in the second one, it is applied only to views atlases.

According to the results, the optimal configuration of IBC search precision is quarter-pel for depth atlases and full-pel for views atlases. Although the achieved gain is small for computer-generated content, bitrate reduction can be as high as 7% for natural content. Moreover, using the proposed mixed precision does not negatively impact encoding time – the results indicate roughly the same encoding time compared to the full-pel only scenario.

Table 6.8. Results for compression of SCC with mixed full-pel and quarter-pel IBC accuracy, compared to SCC. Negative values indicate lower bitrates at the same quality or faster encoding.

Sequence	QPel accuracy for depth, FPel for views			FPel accuracy for depth, QPel for views		
	BD-rate WS-PSNR$_Y$	BD-rate IV-PSNR	Δ Encoding time	BD-rate WS-PSNR$_Y$	BD-rate IV-PSNR	Δ Encoding time
ClassroomVideo	– 0.2%	– 0.2%	+ 0.8%	– 0.1%	+ 0.9%	+ 7.0%
TechnicolorMuseum	– 0.0%	– 0.1%	+ 0.3%	+ 0.6%	+ 1.0%	+ 8.3%
Fan	+ 0.3%	+ 0.2%	– 0.9%	– 0.0%	+ 0.3%	+ 2.7%
OrangeKitchen	– 0.1%	– 0.3%	– 2.8%	+ 0.2%	+ 0.0%	+ 5.9%
Chess	– 0.1%	– 0.1%	– 2.4%	– 0.4%	– 1.1%	+ 3.7%
Group	+ 0.4%	+ 0.2%	+ 1.9%	+ 0.3%	+ 0.3%	+ 7.4%
ChessPieces	– 1.0%	– 1.5%	– 0.2%	– 0.0%	– 1.5%	+ 7.8%
TechnicolorHijack	+ 0.1%	– 0.2%	+ 0.9%	– 0.2%	– 0.4%	+12.0%
Mirror	– 0.2%	– 0.2%	– 0.6%	+ 0.7%	+ 0.3%	+ 1.1%
CG average	**– 0.1%**	**– 0.2%**	**– 0.3%**	**+ 0.1%**	**– 0.0%**	**+ 6.2%**
TechnicolorPainter	– 1.3%	– 1.3%	+ 0.4%	– 0.3%	– 0.6%	+ 8.9%
IntelFrog	– 0.5%	– 1.0%	+ 1.1%	+ 0.6%	+ 1.0%	+ 6.9%
Poznan_Carpark	– 1.5%	– 1.7%	+ 3.2%	+ 0.2%	+ 0.3%	+ 7.8%
Poznan_Fencing2	– 1.9%	– 2.5%	+ 0.8%	– 0.0%	+ 0.0%	+ 6.3%
Poznan_Hall2	– 6.3%	– 7.1%	+ 1.3%	+ 0.4%	– 0.2%	+ 5.9%
Poznan_Street	– 4.5%	– 4.6%	+ 0.6%	+ 1.5%	+ 1.5%	+ 8.2%
NC average	**– 2.7%**	**– 3.0%**	**+ 1.2%**	**+ 0.4%**	**+ 0.3%**	**+ 7.3%**
Average	**– 1.1%**	**– 1.4%**	**+ 0.3%**	**+ 0.2%**	**+ 0.1%**	**+ 6.6%**

B. QPEL + TILE-BASED IBC ANALYSIS

In this section, the performance of tile-based IBC analysis is evaluated. The reference for the experiments is HEVC SCC with quarter-pel accuracy of IBC vector search. In the first experiment, tile-based encoding is enabled only for atlas with base views. The reason for that is because such type of an atlas is somewhat a frame-compatible multiview video, where the views are stacked vertically. Therefore, tile encoding is set up in such a way that each tile contains exactly one of the views. The results of the abovementioned experiments are presented in Table 6.9.

Table 6.9. Results for compression of SCC with quarter-pel IBC accuracy and tile-based IBC analysis for the first atlas of views, compared to basic SCC. Negative values indicate lower bitrates at the same quality or faster encoding.

Sequence	**BD-rate WS-PSNR$_Y$**	**BD-rate IV-PSNR**	**Δ Encoding time**
ClassroomVideo	+ 2.6%	– 5.2%	– 45.6%
TechnicolorMuseum	+ 1.3%	+ 2.1%	– 21.5%
Fan	+ 0.7%	+ 1.4%	– 19.3%
OrangeKitchen	+ 1.6%	+ 2.8%	– 9.6%
Chess	+ 0.3%	– 0.5%	– 9.4%
Group	+ 0.8%	+ 0.3%	– 22.9%
ChessPieces	– 0.6%	– 2.4%	– 6.5%
TechnicolorHijack	+ 0.6%	+ 1.4%	+13.0%
Mirror	+ 2.4%	+ 2.8%	– 8.6%
CG average	**+ 1.1%**	**+ 0.3%**	**– 15.5%**
TechnicolorPainter	– 2.3%	– 4.2%	– 4.3%
IntelFrog	– 0.7%	– 5.6%	– 24.2%
Poznan_Carpark	– 6.1%	– 8.2%	– 24.0%
Poznan_Fencing2	– 1.5%	– 3.5%	– 17.5%
Poznan_Hall2	– 6.7%	– 7.7%	– 9.3%
Poznan_Street	– 10.5%	– 12.3%	– 23.2%
NC average	**– 4.6%**	**– 6.9%**	**– 17.1%**
Average	**– 1.2%**	**– 2.6%**	**– 15.5%**

In the remaining experiments, the tile-based IBC is enabled either for both views atlases or for all atlases. The gathered results are collectively presented in Table 6.10.

The results of the first experiment (Table 6.9) show that enabling tile-encoding for atlases containing base views with natural content significantly improves compression efficiency. At constant IV-PSNR, the **bitrate is reduced by nearly 7%**. Moreover, the **encoding time is noticeably shorter (15.5% on average)**. That is due to tile-based encoding, in which the IBC search is performed only on the referenced tile, much faster than analyzing the whole previously encoded area.

Regarding computer-generated content, enabling tile-based encoding slightly decreases the overall compression efficiency, nevertheless, the encoding time is reduced at a similar level as for natural content. Therefore, the results clearly indicate that tile encoding should be enabled for the first views atlas with natural content. It is possible to automate the configuration setup based on the atlas type by adding a MIV metadata parser, as described in Section 6.3.2.

Table 6.10 presents the influence of tile-based IBC enabled for different sets of atlases. The left part of the table contains experimental results of immersive video compression with tile-based encoding enabled for atlases with views data. The right part presents the results achieved when all atlases (views + depth) are compressed with tile-based encoding.

Table 6.10. Results for compression of SCC with Quarter-Pel IBC Accuracy and Tile-Based IBC Analysis for different sets of Atlases, compared to basic SCC. Negative values indicate lower bitrates at the same quality or faster encoding.

Sequence	**Tile-based IBC analysis for views atlases**			**Tile-based IBC analysis for all atlases (views + depth)**		
	BD-rate WS-PSNR$_Y$	**BD-rate IV-PSNR**	**Δ Encoding time**	**BD-rate WS-PSNR$_Y$**	**BD-rate IV-PSNR**	**Δ Encoding time**
ClassroomVideo	3.7%	– 4.5%	– 60.3%	9.0%	– 1.8%	– 58.4%
TechnicolorMuseum	2.2%	2.9%	– 52.7%	4.4%	4.7%	– 49.1%
Fan	1.6%	2.5%	– 48.2%	6.4%	6.8%	– 56.4%
OrangeKitchen	3.8%	4.9%	– 42.4%	8.4%	8.7%	– 36.6%
Chess	2.1%	0.9%	– 43.4%	7.4%	5.3%	– 39.6%
Group	1.7%	1.1%	– 54.0%	6.9%	4.5%	– 54.8%
ChessPieces	1.4%	0.4%	– 43.9%	7.7%	4.9%	– 40.3%
TechnicolorHijack	1.8%	2.8%	– 42.0%	8.5%	8.9%	– 46.3%
Mirror	4.3%	4.6%	– 44.2%	7.0%	7.1%	– 46.0%
CG average	**2.5%**	**1.7%**	– 47.9%	**7.3%**	**5.5%**	– 47.5%
TechnicolorPainter	– 1.9%	– 4.0%	– 47.3%	2.0%	– 0.2%	– 61.0%
IntelFrog	– 0.1%	– 5.6%	– 55.6%	3.3%	– 2.6%	– 66.3%
Poznan_Carpark	– 5.6%	– 7.9%	– 36.2%	– 2.6%	– 5.1%	– 44.3%
Poznan_Fencing2	– 0.7%	– 3.0%	– 46.0%	4.6%	1.9%	– 55.2%
Poznan_Hall2	– 6.3%	– 7.2%	– 28.2%	3.5%	2.5%	– 46.0%
Poznan_Street	– 10 %	– 12 %	– 33.7%	– 6.2%	– 8.0%	– 41.6%
NC average	**– 4.1%**	**– 6.6%**	– 41.2%	**0.8%**	**– 1.9%**	– 52.4%
Average	**– 0.1%**	**– 1.6%**	– 45.2%	**4.7%**	**2.5%**	– 49.5%

When comparing Tables 6.10 and 6.9, it can be observed that performing tile-based encoding on patch atlases for views decreases the overall compression efficiency. Such a result is expected because patches in atlases are packed pseudo-randomly, therefore, it is not guaranteed that similar patches will be placed in different tiles.

Enabling tile-based encoding in compression of depth atlases significantly deteriorates the BD-rates for both quality metrics. Such results can be explained by different characteristics of views and depth maps. The latter often contain sharp edges between objects or between objects and the background, as well as large, smooth, self-similar areas. Such characteristics make standard IBC more efficient as it takes advantage of self-similarities, opposite to the proposed tile-encoding. On the other hand, tile-based encoding has a positive impact on encoding time, which is almost halved compared to HEVC SCC encoding with only quarter-pel IBC search enabled.

6.5. CONCLUSIONS

In this chapter, the author of the dissertation proposes several modifications of Screen Content Coding aimed at improving the compression efficiency of multiview and immersive video. The proposal includes changes that are compatible with HEVC SCC, as well as modifications that require modifying the standard. The improvements are implemented on top of the test model for HEVC Screen Content Coding and adapted to the new applications – multiview and immersive video compression.

The evaluation of the author's novel compression technique, Advanced SCC, includes a series of experiments and a comparison of the results with the state-of-the-art multiview and immersive video compression techniques. The results show that **ASCC is significantly more efficient than standard-compliant SCC**. In the case of multiview encoding, **the proposal is as efficient as the dedicated solution – Multiview HEVC**, without a negative impact on encoding time. When applied to camera-captured immersive video, the author's improvements **significantly decrease the output bitrate and the encoding time**, compared to using state-of-the-art SCC as a video encoder in TMIV. Due to the different characteristics of atlases generated by MIV and differences between computer-generated and camera-captured content, the author also proposes including a MIV metadata parser as a controller of internal video encoder to adjust the configuration of ASCC to the input data and thus optimizing the overall performance.

The influence of the proposed modifications is also evaluated for the compression of screen content (Section 6.4.3). The results highly depend on the content of the test sequence, but usually, the proposal is less efficient than the original SCC. Therefore, modifications of Screen Content Coding should be adaptively toggled based on the input video data, similarly to the depth flag in 3D-HEVC that enables additional coding techniques dedicated to depth maps [ISO'21].

To sum up, the Advanced SCC technique proposed by the author of the dissertation appears to be an efficient, versatile solution that can provide high performance in multiple applications. The need for such versatile video codecs can be observed during the development of modern standards such as MPEG Immersive Video, which aims at efficient compression of both natural and computer-generated content acquired with different camera setups (omnidirectional, multi-camera systems, etc.) and is codec-agnostic to assure that the internal video codec can be easily changed when more efficient solutions appear. Moreover, the practical use of video extensions dedicated to one type of application is very limited, while the development of such extensions requires a lot of effort.

7. SUMMARY OF THE DISSERTATION

7.1. RECAPITULATION

This dissertation focuses on inter-view prediction techniques for the compression of multiview video acquired using systems with various camera arrangements. In Chapter 2, the author describes the state-of-the-art of multiview video coding and explains the motivation for the research. As stated, the dedicated codec for compression of such video, 3D-HEVC, is not efficient in the compression of video acquired using multi-camera systems with camera arrangements other than linear. Moreover, dedicated multiview profiles built on top of single-layer codecs, significantly increase the complexity of the codec, are not reusable in other applications, and require a lot of additional research to be developed. In order to address those issues, the author of the dissertation formulated two theses in Section 1.2:

1. **It is possible to reduce both bitrate and encoding time of 3D-HEVC encoder in compression of rectified multiview video acquired with cameras located on a circle, compared to the state-of-the-art 3D-HEVC encoder, through adaptation of inter-view prediction to circular camera arrangements.**
2. **It is possible to use standard-compliant HEVC Screen Content Coding for compression of stereoscopic video, frame-compatible multiview video, and immersive video. With additional improvements, the rate-distortion compression efficiency of such an approach can be comparable or even higher than the state-of-the-art dedicated techniques.**

Chapter 4 presents the verification of the first thesis. The author of the dissertation proposes a process for rectification of multiview video acquired roughly on a circle, using full 3D point mapping. Then, the author introduces a novel method for efficient inter-view prediction in compression of circularly rectified video. The proposal is implemented on top of 3D-HEVC through modification of Disparity Compensated Prediction, Inter-View Motion Prediction, View Synthesis Prediction, and other inter-view prediction tools. The new approach entails changes in the representation of camera parameters in 3D-HEVC bitstream syntax. In Section 4.6, the author presents experimental results of comparison between state-of-the-art and ARC-HEVC, in compression of circularly rectified video. The results are additionally compared to 3D-HEVC adapted to arbitrary camera arrangement (ANY-HEVC), which is also a non-standard codec co-authored by the author of this dissertation.

The verification of the second thesis is presented in Chapter 5 and Chapter 6. First, the author of the dissertation presents the idea of adapting standard-compliant Screen Content Coding for the compression of frame-compatible multiview video. In that idea, Intra Block Copy works as an inter-view prediction, even though its original purpose was different. The author describes the process of choosing the optimal view alignment, as well as the best configuration of other SCC tools. Then, the novel approach is used in 3 applications: stereoscopic, multiview, and immersive video compression. Section 5.7 presents the experimental evaluation of the proposal. In Chapter 6, the author's idea is further developed by several improvements. The goal of the modifications is to improve the inter-view prediction accuracy of SCC for multiview and immersive video coding. The proposal is experimentally evaluated, and the results are compared to the state-of-the-art dedicated solutions.

7.2. RESEARCH WORK DONE

During the research, the author created an original implementation of the following software:

- ANY-HEVC – part of modifications added on top of 3D-HEVC; ~2000 lines of code in C++,
- ARC-HEVC – all modifications added on top of 3D-HEVC; ~3500 lines of code in C++,
- Circular rectifier – software for deriving rectified camera parameters and performing video rectification; ~1000 lines of code in C++ and Python,
- Advanced SCC – modifications of HEVC SCC; ~2500 lines of code in C++.

Additionally, the author created software for preparing and running the experiments, and for processing their results. In total, the author prepared roughly 10000 lines of code. It should also be noted that the author's improvements were added on top of test models for 3D-HEVC and HEVC SCC, which are very complex software implementations of state-of-the-art video codecs. For example, as mentioned in Chapter 3, the test model for 3D-HEVC contains roughly 120 thousand lines of code. Introducing modifications to the core algorithms of such advanced software requires many hours of code analysis and debugging.

To assess the solutions presented in the dissertation, the author conducted multiple time-consuming experiments. If they were performed on a single core of a CPU, the processing would take roughly 180 days. The amount of video data used in the experiments was close to

300 GB. The aforementioned numbers show that the research on multiview video compression is a challenging task that requires a lot of effort and computational power.

7.3. ORIGINAL ACHIEVEMENTS OF THE AUTHOR AND CONCLUSIONS

In the dissertation, several original achievements of the author can be found. The most important ones are summarized below.

- **Development of the concept and the procedure for circular rectification.** The author proposes the process of derivation of circle parameters and virtual camera positions that best fit the real positions of the cameras. Both intrinsic and extrinsic camera parameters are rectified to an ideal circle with optical axes collocated on a single plane and intersecting in the center of the circle. The author also proposes a solution for the problem of misalignment of the field of view if the optical axis of a camera before rectification is far from the center of the circle.
- **Efficient modification of 3D-HEVC codec for processing of the circularly rectified 3D video (ARC-HEVC).** The author proposes formulas for mapping points between circularly rectified views. They are used in the author's modification of 3D-HEVC that includes adapting Disparity Compensated Prediction and other inter-view prediction tools to circular camera arrangements. The author also proposes modifications of 3D-HEVC syntax. Experimental results show that the proposed codec **reduces bitrate on average by 6% when compared to standard 3D-HEVC.** At the same time, the average encoding time is **reduced by more than 4%.** Therefore, for compression of multiview video with depth acquired with cameras sparsely distributed around the scene, **the proposed codec is objectively better than the dedicated state-of-the-art solution**, both in terms of compression efficiency and encoding time.
- **Adaptation of standard-compliant HEVC Screen Content Coding to efficient stereoscopic, multiview, and immersive video compression.** The author presents a novel and unexpected use of Intra Block Copy as an inter-view prediction tool. Even though SCC is designed to compress computer-generated video, the author proves that it can be successfully reused as a multiview codec for camera-captured content. This idea requires the preparation of a multiview video in a frame-compatible structure. The author experimentally found the best view alignment within a frame and the optimal configuration of SCC tools. The evaluation of the proposal proves that for stereoscopic

video compression, **SCC reduces the bitrate roughly by 20% for All-Intra and 15% for Random Access** when compared to the HEVC Main profile. In the case of multiview video coding (4 views), the bitrate reduction is even higher – **roughly 28% for All-Intra and 20% for Random Access**. Nevertheless, without further modifications, the proposed method is not as efficient as the dedicated MV-HEVC profile. When HEVC SCC is used as an internal MIV codec instead of HEVC Main, **the bitrate is also reduced, especially in the compression of computer-generated content (up to more than 7%) and depth maps (up to more than 15% for base views and 25% for patch atlases)**, however, the results depend on the content. Increased quality of depth maps results in a significantly better quality of synthesized virtual views, which is a crucial factor for immersive video applications. The idea of using **SCC as an internal MIV codec has been independently evaluated by the MPEG** group responsible for the development of MIV, and **it has been confirmed to be beneficial**.

- **Efficient modifications of HEVC SCC for compression of camera-captured multiview and immersive video.** The author proposes a set of modifications to SCC, aimed at increasing its compression efficiency. Experimental results show that **with the proposed modifications, SCC can be as efficient as MV-HEVC**, and slightly faster (on average 5% for 3 views) at the same time. Therefore, modified SCC could successfully replace MV-HEVC as a more versatile solution. Regarding immersive video coding, the efficiency of the proposed modifications strongly depends on the content, however, for camera-captured video they usually provide a significant gain both in terms of bitrate and encoding time reduction.

To sum up, in the dissertation, the author proves both theses to be valid. For compression of circularly rectified multiview video, the proposed modifications of inter-view prediction in 3D-HEVC reduce both bitrate and encoding time, compared to the state-of-the-art 3D-HEVC, even though ARC-HEVC is more complex and introduces additional parameters.

Regarding the second thesis, the author proposes a novel approach that reuses existing Screen Content Coding technique and improves it to provide equally efficient inter-view prediction as in the dedicated Multiview HEVC. This achievement shows that the development of future video coding should be directed toward the unification of coding techniques rather than creating dedicated extensions. Research presented in the dissertation can be the starting point for adapting the emerging Versatile Video Coding to frame-compatible multiview video coding and immersive video coding.

REFERENCES

[Bjøntegaard'01] G. Bjøntegaard, *Calculation of Average PSNR Differences between RD-curves*, ITU-T SG16, Doc. VCEG-M33, Austin, USA, April 2001.

[Boissonade'18] P. Boissonade and J. Jung, *[MPEG-I Visual] Proposition of new sequences for Windowed6DoF experiments on compression, synthesis, and depth estimation*, ISO/IEC JTC1/SC29/WG11 MPEG/M43318, Ljubljana, Slovenia, 2018.

[Bonatto'21] D. Bonatto, S. Fachada, S. Rogge, A. Munteanu and G. Lafruit, *Real-Time Depth Video-Based Rendering for 6-DoF HMD Navigation and Light Field Displays*, in IEEE Access, vol. 9, pp. 146868-146887, 2021, doi: 10.1109/ACCESS.2021.3123529.

[Bossen'13] F. Bossen, *Common Test Conditions and software reference configurations*, JCT-VC of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 12th Meeting: Doc: JCTVC-L1100, Geneva, CH, January 2013.

[Boyce'21] J. Boyce, R. Doré, A. Dziembowski, J. Fleureau, J. Jung, B. Kroon, B. Salahieh, V. K. Vadakital, Lu Yu, *MPEG Immersive Video Coding Standard*, Proceedings of the IEEE, vol. 109, no. 9, pp. 1521-1536, 2021.

[Bross'21] B. Bross et al., *Overview of the Versatile Video Coding (VVC) Standard and its Applications*, IEEE Transactions on Circuits and Systems for Video Technology, vol. 31, no. 10, pp. 3736-3764, October 2021, doi: 10.1109/TCSVT.2021.3101953.

[Ceulemans'18] B. Ceulemans et al., *Robust Multiview Synthesis for Wide-Baseline Camera Arrays*, IEEE Tr. on Multimedia, 2018.

[Chen'97] X. Chen and A. Luthra, *MPEG-2 multi-view profile and its application in 3DTV*, Proc. SPIE IS&T Multimedia Hardware Architectures, San Diego, USA, Vol. 3021, pp. 212-223, February 1997.

[Choi'12] M. Choi, J. Kim, W. -K. Cho and Y. Chung, *Low complexity image rectification for multi-view video coding, 2012 IEEE International Symposium on Circuits and Systems (ISCAS),* 2012, pp. 381-384.

[Cisco'18] *VNI Complete Forecast Highlights* https://www.cisco.com/c/dam/m/en_us/solutions/service-provider/vni-forecast-highlights/pdf/Global_Device_Growth_Traffic_Profiles.pdf, Accessed 8 May 2023

[Cisco'20] *Cisco Annual Internet Report (2018-2023)* https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.pdf, Accessed 8 May 2023

[Collins'99] R. T. Collins and Y. Tsin, *Calibration of an Outdoor Active Camera System,* Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition 1999, vol. 1, pp. 528-534.

[Cserkaszky'18] A. Cserkaszky, P. A. Kara, A. Barsi, M. G. Martini and T. Balogh, *Light - fields of Circular Camera Arrays,* 2018 26th European Signal Processing Conference (EUSIPCO), Rome, Italy, 2018, pp. 241-245, doi: 10.23919/EUSIPCO.2018.8552998.

[CTC MIV] *Common Test Conditions for MPEG Immersive Video. Standardization document*, ISO/IEC JTC1/SC29/WG04 MPEG VC N0051, Online, January 2021.

[Cui'19] L. Cui et al., *Point-Cloud Compression: Moving Picture Experts Group's New Standard in 2020*, IEEE Consumer Electronics Magazine, 2019

[Cyganek'09] B. Cyganek, J. Siebert, *An Introduction to 3D Computer Vision Techniques and Algorithms*, Wiley, 2009.

[Domański'09] M. Domański, T. Grajek, K. Klimaszewski, M. Kurc, O. Stankiewicz, J. Stankowski, K. Wegner, *Poznań multiview video test sequences and camera parameters*, ISO/IEC JTC1/SC29/WG11 MPEG Doc. M17050, Xian, China, October 2009.

[Domański'13] M. Domański, O. Stankiewicz, K. Wegner, M. Kurc, J. Konieczny, J. Siast, J. Stankowski, R. Ratajczak, *High efficiency 3D video coding using new tools based on view synthesis*, in IEEE Transactions on Image Processing, vol. 22, no. 9, pp. 3517-3527, September 2013

[Domański'15A] M. Domański, A. Dziembowski, T. Grajek, A. Grzelka, Ł. Kowalski, M. Kurc, A. Łuczak, D. Mieloch, R. Ratajczak, J. Samelak, O. Stankiewicz, J. Stankowski, K. Wegner, *Methods of high efficiency compression for transmission of spatial representation of motion scenes*, IEEE International Conference on Multimedia and Expo, ICME 2015, Turyn, Włochy, 29.06-3.07, 2015

[Domański'15B] M. Domański, A. Dziembowski, M. Kurc, A. Łuczak, D. Mieloch, J. Siast, O. Stankiewicz, K. Wegner, *Poznan University of Technology test multiview video sequences acquired with circular camera arrangement – "Poznan Team" and "Poznan Blocks" sequences,* ISO/IEC JTC1/SC29/WG11 MPEG Doc. M35846, Geneva, Switzerland, February 2015.

[Domański'15C] M. Domański, A. Dziembowski, D. Mieloch, A. Łuczak, O. Stankiewicz, K. Wegner, *A Practical Approach to Acquisition and Processing of Free Viewpoint Video*, 31st Picture Coding Symposium PCS 2015, Cairns, Australia, 31 May - 3 June 2015, pp. 10-14,

[Domański'16A] M. Domański, M. Bartkowiak, A. Dziembowski, T. Grajek, A. Grzelka, A. Łuczak, D. Mieloch, J. Samelak, O. Stankiewicz, J. Stankowski, K. Wegner, *New results in free-viewpoint television systems for horizontal virtual navigation*, IEEE International Conference on Multimedia and Expo, ICME 2016, Seattle, Stany Zjednoczone, 11-15.07.2016

[Domański'16B] M. Domanski et al., *Multiview test video sequences for free navigation exploration obtained using pairs of cameras*, ISO/IEC JTC1/SC29/WG11 MPEG/M38247, Geneva, Switzerland, 2016.

[Domański'17] M. Domański, O. Stankiewicz, K. Wegner and T. Grajek, *Immersive visual media — MPEG-I: 360 video, virtual navigation and beyond*, 2017 International Conference on Systems, Signals and Image Processing (IWSSIP), Poznan, Poland, 2017, pp. 1-9, doi: 10.1109/IWSSIP.2017.7965623.

[Domański'19] M. Domański, T. Grajek, C. Conti, C. Debono, S. de Faria, P. Kovacs, L. Lucas, P. Nunes, C. Perra, N. Rodrigues, M. Sjöström, L. Soares, O. Stankiewicz, *Emerging Imaging Technologies: Trends and Challenges*, 3D Visual Content Creation, Coding and Delivery, Signals and Communication Technology: Springer International, 2019, pp. 5-39, ISBN: 978-3-319-77841-9, ISSN: 1860-4862.

[Doré'18] R. Doré, *Technicolor 3DoF+ test materials*, ISO/IEC JTC1/SC29/WG11 MPEG/M42349, San Diego, USA, 2018.

[Doré'20A] R. Doré, G. Briand, F. Thudor, *[MPEG-I Visual] InterdigitalFan0 content proposal for MIV*, ISO/IEC JTC1/SC29/WG11 MPEG/M54732, online, July 2020.

[Doré'20B] R. Doré, G. Briand, F. Thudor, *[MPEG-I Visual] InterdigitalGroup content proposal*, ISO/IEC JTC1/SC29/WG11 MPEG/M54731, online, July 2020.

[Doré'21] R. Doré, G. Briand, *[MPEG-I] Interdigital Mirror Content Proposal for advanced MIV investigations on reflection*, ISO/IEC JTC1/SC29/WG11 MPEG/M55710, online, January 2021.

[Doyen'17] D. Doyen et al., *Light field content from 16- camera rig*, ISO/IEC JTC1/SC29/WG11 MPEG, M40010, Geneva, Switzerland, 2017.

[Dziembowski'17A] A. Dziembowski, A. Grzelka, D. Mieloch, O. Stankiewicz, M. Domański, *Enhancing view synthesis with image and depth map upsampling*, International Conference on Systems, Signals and Image Processing, IWSSIP 2017, Poznań, Polska, 22-24.05.2017

[Dziembowski'17B] A. Dziembowski, A. Grzelka, D. Mieloch, *Zwiększanie rozdzielczości obrazu i mapy głębi w celu poprawy jakości syntezy widoków wirtualnych*, Przegląd Telekomunikacyjny, tom 86, nr 6, s. 405-408, 2017

[Dziembowski'22] A. Dziembowski, D. Mieloch, J. Stankowski, A. Grzelka, *IV-PSNR – the objective quality metric for immersive video applications*, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 32, 1 June 2022, pp. 7575 - 7591, ISSN: 1558-2205

[Fachada'18] S. Fachada, D. Bonatto, A. Schenkel and G. Lafruit, *Depth image based view synthesis with multiple reference views for virtual reality*, 2018 - 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), Helsinki, Finland, 2018, pp. 1-4, doi: 10.1109/3DTV.2018.8478484.

[Fujikawa'10] S. Fujikawa, *Three-dimensional imaging technology: A revolution in the world of imaging*, 2010 Symposium on VLSI Circuits, Honolulu, HI, USA, 2010, pp. 9-12, doi: 10.1109/VLSIC.2010.5560254.

[Guillo'18] L. Guillo, X. Jiang, G. Lafruit, C. Guillemot, *[MPEG-I Visual] Light field video dataset captured by a R8 Raytrix camera (with disparity maps)*, ISO/IEC JTC1/SC29/WG11 MPEG Doc. M42468, San Diego, USA, April 2018.

[Hannuksela'15] M. M. Hannuksela, Y. Yan, X. Huang and H. Li, *Overview of the multiview high efficiency video coding (MV-HEVC) standard*, 2015 IEEE International Conference on Image Processing (ICIP), Quebec City, QC, Canada, 2015, pp. 2154-2158, doi: 10.1109/ICIP.2015.7351182.

[Hartley'03] R. Hartley, A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge: Cambridge Univ. Press, 2nd ed., 2003.

[Hartley'99] Hartley, R.I., *Theory and Practice of Projective Rectification*, International Journal of Computer Vision 35, 115–127 (1999). https://doi.org/10.1023/A:1008115206617

[HM] JCT-VC, *HEVC reference software repository*, https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.9. Accessed 8 May 2023

[HM+SCM] JCT-VC, *HEVC Screen Content Coding reference software repository*, https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftw are/tags/HM-16.9+SCM-8.0. Accessed 8 May 2023

[Ho'07] Ho, Yo-Sung & Oh, Kwan-Jung, *Overview of Multi-view Video Coding*, 5 - 12. 10.1109/IWSSIP.2007.4381085, 2007

[Ho'08] Y.S. Ho, E.K. Lee, C. Lee, *Multiview video test sequence and camera parameters*, ISO/IEC JTC1/SC29/ WG11 MPEG Doc. M15419, Archamps, France, April 2008.

[Ho'12] Y. Ho, *Video processing techniques for 3D television*, *2012 Visual Communications and Image Processing,* 2012, pp. 1-13.

[Hou'20] Y. Hou, X. Su, W. Chen, *Alignment Method of an Axis Based on Camera Calibration in a Rotating Optical Measurement System,* Applied Sciences. 2020; 10(19):6962. https://doi.org/10.3390/app10196962

[HTM] JCT-3V, *Multiview HEVC reference software repository*, https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSoftware/tags/HTM-16.2. Accessed 8 May 2023

[Hu'23] P. Hu, E. S. L. Ho and A. Munteanu, *AlignBodyNet: Deep Learning-Based Alignment of Non-Overlapping Partial Body Point Clouds From a Single Depth Camera*, in IEEE Transactions on Instrumentation and Measurement, vol. 72, pp. 1-9, 2023, Art no. 2502609, doi: 10.1109/TIM.2022.3222501.

[Ilola'19] L. Ilola, V.K.M. Vadakital, K. Roimela, J. Keränen, *New test content for Immersive Video – Nokia Chess*, ISO/IEC JTC1/SC29/WG11 MPEG/M50787, Geneva, Switzerland, 2019.

[ISO'12] ISO/IEC IS 13818-2: 2013 and ITU-T Rec. H.262 (V3.1) (2012), *Generic coding of moving pictures and associated audio information – Part 2: Video*,

[ISO'14] ISO/IEC IS 14496-10: 2014, *Coding of audio-visual objects - Part 10: Advanced Video Coding*, and ITU-T Rec. H.264 (V9) (2014), *Advanced video coding for generic audiovisual services*

[ISO'19] ISO/IEC MPEG. *Software manual of IV-PSNR for Immersive Video*. ISO/IEC JTC1/SC29/WG11 MPEG/N18709, Goeteborg, Sweden, 2019

[ISO'21] ISO/IEC Int. Standard 23008-2: 2020, *High efficiency coding and media delivery in heterogeneous environment – Part 2: High efficiency video coding*, and ITU-T Rec. H.265 (V8) (2021), *High efficiency video coding*.

[ISO'22] ISO/IEC Int. Standard 23090-3: 2022, *Information technology — Coded representation of immersive media — Part 3: Versatile video coding*, and ITU-T Rec. H.266 (V2) (2022), *Versatile video coding*.

[Jhu'20] H. -J. Jhu, X. Xiu, Y. -W. Chen, T. -C. Ma and X. Wang, *Adaptive Color Transform in VVC Standard*, 2020 IEEE International Conference on Visual Communications and Image Processing (VCIP), Macau, China, 2020, pp. 314-317, doi: 10.1109/VCIP49819.2020.9301798.

[Kang'08] Y. -S. Kang, C. Lee and Y. -S. Ho, *An Efficient Rectification Algorithm for Multi-View Images in Parallel Camera Array*, 2008 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video, Istanbul, Turkey, 2008, pp. 61-64, doi: 10.1109/3DTV.2008.4547808.

[Kovacs'15] P. T. Kovacs, *[FTV AHG] Big Buck Bunny light-field test sequences*, MPEG M35721, Geneva, February 2015.

[Kroon'18] B. Kroon, *3DoF+ test sequence ClassroomVideo*, ISO/IEC JTC1/SC29/WG11 MPEG/M42415, San Diego, USA, 2018.

[LaValle'20] S. LaValle, *Virtual Reality*, Cambridge Univ. Press, 2020.

[Li'16] Z. Li et al., *Toward a practical perceptual video quality metric*, Netflix Technology Blog, 2016.

[Li'20] L. Li et al., *Advanced 3D Motion Prediction for Video-Based Dynamic Point Cloud Compression*, IEEE Tr. on Image Processing, 2020.

[Lucchese'03] L. Lucchese and S. K. Mitra, *Correction of geometric lens distortion through image warping,* 3rd International Symposium on Image and Signal Processing and Analysis, 2003. ISPA 2003. Proceedings of the, Rome, Italy, 2003, pp. 516-521 Vol.1, doi: 10.1109/ISPA.2003.1296951.

[Ma'05] S. Ma, W. Gao and Yan Lu, *Rate-distortion analysis for H.264/AVC video coding and its application to rate control*, in IEEE Transactions on Circuits and Systems for Video Technology, vol. 15, no. 12, pp. 1533-1544, December 2005, doi: 10.1109/TCSVT.2005.857300.

[Mieloch'17A] D. Mieloch, A. Dziembowski, A. Grzelka, O. Stankiewicz, M. Domański, *Graph-based multiview depth estimation using segmentation*, IEEE International Conference on Multimedia and Expo, ICME 2017, Hongkong, Chiny, 10-14.07.2017

[Mieloch'17B] D. Mieloch, A. Dziembowski, A. Grzelka, O. Stankiewicz, M. Domański, *Temporal enhancement of graph-based depth estimation method*, International Conference on Systems, Signals and Image Processing, IWSSIP 2017, Poznań, Polska, 22-24.05.2017

[Mieloch'17C] D. Mieloch, A. Dziembowski, A. Grzelka, *Estymacja głębi dla systemów wielowidokowych*, Przegląd Telekomunikacyjny, tom 86, nr 6, s. 479-482, 2017

[Mieloch'21] D. Mieloch, A. Dziembowski and M. Domański, *Depth Map Refinement for Immersive Video*, in IEEE Access, vol. 9, pp. 10778-10788, 2021, doi: 10.1109/ACCESS.2021.3050554.

[Müller'11] K. Müller, P. Merkle, T. Wiegand, *3D video representation using depth maps*, Proceedings of the IEEE, tom 99, nr 4, s. 643-656, 04.2011

[Müller'13] K. Müller, H. Schwarz, D. Marpe, Ch. Bartnik, S. Bosse, H. Brust, T. Hinz, H. Lakshman, Ph. Merkle, F. Hunn Rhee, G. Tech, M. Winken, Th. Wiegand, *3D High Efficiency Video Coding for multi-view video and depth data*, IEEE Trans. Image Processing, IEEE Trans. Image Processing, vol. 22, pp. 3366-3378, 2013.

[Müller'14] K. Müller, A. Vetro, *Common Test Conditions of 3DV Core Experiments*, JCT-3V of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 7th Meeting: Doc. JCT3V-G1100, San José, US, January 2014.

[Ohm'99] J.-R. Ohm, *Stereo/Multiview Video Encoding Using the MPEG Family of Standards*, Proc. SPIE Conf. Stereoscopic Displays and Virtual Reality Systems VI, San Jose, CA, January 1999.

[Pu'16] W. Pu et al., *Palette Mode Coding in HEVC Screen Content Coding Extension*, in IEEE Journal on Emerging and Selected Topics in Circuits and Systems, vol. 6, no. 4, pp. 420-432, December 2016, doi: 10.1109/JETCAS.2016.2605661.

[Rahaman'18] D. Rahaman and M. Paul, *Virtual view synthesis for free viewpoint video and multiview video compression using Gaussian mixture modelling*, IEEE Tr. on Image Processing, 2018.

[Ratajczak'12] R. Ratajczak, M. Domański and K. Wegner, *Vehicle size estimation from stereoscopic video*, 2012 19th International Conference on Systems, Signals and Image Processing (IWSSIP), Vienna, Austria, 2012, pp. 405-408.

[Rogge'19] S. Rogge, D. Bonatto, J. Sancho, R. Salvador, E. Juarez, A. Munteanu, G. Lafruit, *MPEG-I Depth Estimation Reference Software*, 2019 International Conference on 3D Immersion (IC3D), Brussels, Belgium, 2019, pp. 1-6, doi: 10.1109/IC3D48390.2019.8975995.

[Rusanovskyy'11] D. Rusanovskyy, P. Aflaki, and M. M. Hannuksela, *Undo dancer 3DV sequence for purposes of 3DV standardization*, ISO/IEC JTC1/SC29/WG11, Geneva, Switzerland, Tech. Rec. M20028, March 2011.

[Salahieh'19] B. Salahieh et al., *Kermit test sequence for Windowed 6DoF Activities*, ISO/IEC JTC1/SC29/WG11/M43748, Ljubljana, 2019.

[Samelak'16] J. Samelak, J. Stankowski, M. Domański, *Adaptation of the 3D-HEVC coding tools to arbitrary locations of cameras*, International Conference on Signals and Electronic Systems, ICSES 2016, Kraków, Poland, September 5-7 2016.

[Samelak'17A] J. Samelak, J. Stankowski, M. Domański, *Efficient frame-compatible stereoscopic video coding using HEVC Screen Content Coding*, IEEE International Conference on Systems, Signals and Image Processing IWSSIP 2017, Poznań, Poland, 22-24 May 2017.

[Samelak'17B] J. Samelak, O. Stankiewicz, M. Domański, *Do we need multiview profiles for future video coding generations?*, ISO/IEC JTC1/SC29/WG11 MPEG2017, M41699, Macao, China, October 2017

[Samelak'17C] J. Samelak, O. Stankiewicz, M. Domański, *Do we need multiview profiles for future video coding generations?*, ISO/IEC JTC1/SC29/WG11 Joint Video Exploration Team (JVET), JVET_H0044, Macao, China, October 2017

[Samelak'17D] J. Samelak, J. Stankowski, M. Domański, *Experimental results for frame – compatible multiview video coding using HEVC SCC*, ISO/IEC JTC1/SC29/WG11 Joint Collaborative Team on Video Coding (JCT-VC), JCTVC-Z0041, Geneva, Switzerland, January 2017

[Samelak'20A] J. Samelak, A. Dziembowski, D. Mieloch, M. Domański, *[MPEG-I Visual] Impact of HEVC profile on TMIV performance*, MPEG130/m53910, Alpbah, Austria, April 2020

[Samelak'20B] J. Samelak, A. Dziembowski, D. Mieloch, M. Domański, *[MPEG-I Visual] HEVC-SCC in TMIV*, MPEG130/m53427, Alpbah, Austria, April 2020

[Samelak'21A] J. Samelak, A. Dziembowski, D. Mieloch, M. Domański, M. Wawrzyniak, *Efficient Immersive Video Compression using Screen Content Coding*, 29th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision : WSCG 2021, WSCG 2021, Prague, Czech Republic, 17-20 May 2021

[Samelak'21B] J. Samelak, M. Domański, *Multiview Video Compression Using Advanced HEVC Screen Content Coding*, https://arxiv.org/abs/2106.13574

[Samelak'22] J. Samelak, A. Dziembowski, D. Mieloch, *Advanced HEVC Screen Content Coding for MPEG Immersive Video*, Electronics, Vol. 11, No. 23, December 2022, ISSN: 2079-9292

[Schwarz'19] S. Schwarz et al., *Emerging MPEG Standards for Point Cloud Compression.* IEEE J. on Emerging and Sel. Topics in Circuits and Systems, 2019.

[Senoh'14] T. Senoh, A. Ishikawa, M. Okui, K. Yamamoto, N. Inoue, *FTV AHG: Super-Multiview Sequences of NICT*, ISO/IEC JTC1/SC29/WG11 MPEG Doc. M32201, San Jose, USA, January 2014.

[Sheikh'06] H. Sheikh and A. Bovik, *Image information and visual quality*, IEEE Tr. on Image Proc., 2006.

[Skulimowski'07] P. Skulimowski, P. Strumiłło, *Obstacle localization in 3D scenes from stereoscopic sequences*, 2007 15th European Signal Processing Conference, Poznan, Poland, 2007, pp. 2095-2099.

[Stankiewicz'18] O. Stankiewicz, M. Domański, A. Dziembowski, A. Grzelka, D. Mieloch, J. Samelak, *A free-viewpoint television system for horizontal virtual navigation*, IEEE Transactions on Multimedia, tom PP, nr 99, s. 1-1, 2018

[Stankowski'10] J. Stankowski, K. Klimaszewski, *Application of epipolar rectification algorithm in 3D Television*, Image Processing and Communications Challenges 2, Advances in Intelligent and Soft Computing: Vol. 84, Springer-Verlag, Berlin, 2010, pp. 345-352, ISBN: 978-3-642-16294-7, ISSN: 1867-5662

[Stankowski'15] J. Stankowski, Ł. Kowalski, J. Samelak, M. Domański, T. Grajek, K. Wegner, *3D-HEVC extension for circular camera arrangements*, 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), Lizbona, Portugalia, 8-10.07.2015

[Strumiłło'18] Strumiłło, P. et al. (2018). *Different Approaches to Aiding Blind Persons in Mobility and Navigation in the "Naviton" and "Sound of Vision" Projects*. In: Pissaloux, E., Velazquez, R. (eds) Mobility of Visually Impaired People. Springer, Cham.

[Sullivan'09] G. J. Sullivan, *Standards-based approaches to 3D and multiview video coding*, Proc. SPIE Conf. Applications of Digital Image Processing XXXII, San Diego, CA, August 2009.

[Sullivan'12] G. J. Sullivan, J. Ohm, W. J. Han, and T. Wiegand, *Overview of the High Efficiency Video Coding (HEVC) Standard*, in IEEE Transactions on Circuits Systems for Video Technology, vol. 22, no. 12, pp. 1649-1668, December 2012.

[Sullivan'13] G.J. Sullivan, J.M. Boyce, Y. Chen, J.R. Ohm, C.A. Segall, A. Vetro, *Standarized extensions of High Efficiency Video Coding (HEVC)*, IEEE Journal of Selected Topics in Signal Processing, tom 7, nr 6, s. 1001-1016, 12.2013

[Sun'17] Y. Sun et al., *Weighted-to-Spherically-Uniform Quality Evaluation for Omnidirectional Video*, *IEEE Sig. Proc. Lett.*, vol. 24, no. 9, September 2017.

[Sun'19] Y. -C. Sun, J. Lou, Y. -H. Chao, H. Wang, V. Seregin and M. Karczewicz, *Analysis of Palette Mode on Versatile Video Coding*, 2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), San Jose, CA, USA, 2019, pp. 455-458, doi: 10.1109/MIPR.2019.00091.

[Suzuki'09] K. Suzuki, M. Tanimoto, *Multiview Video Test Sequence in Circular Camera Arrangement for 3DV/FTV*, ISO/IEC JTC1/SC29/WG11, MPEG Doc. M24846, Geneva, Switzerland, May 2012.

[Suzuki'14] T. Suzuki, R. Choen, T. K. Tan, S. Wenger, *JCT-VC AHG report: Test sequence material (AHG16)*, ISO/IEC JTC1/SC29/WG11 MPEG Doc. M33409, Valencia, Spain, March 2014.

[Tanimoto'09] M. Tanimoto, T. Fujii, M.P. Tehrani, M. Wildeboer, N. Fukushima, H. Furihata, *Moving Multiview Camera Test Sequences for MPEG-FTV*, ISO/IEC JTC1/SC29/WG11, MPEG Doc. M16922, Xian, China, October 2009.

[Tanimoto'12] M. Tanimoto, M. Panahpour Tehrani, T. Fujii and T. Yendo, *FTV for 3-D Spatial Communication*, in Proceedings of the IEEE, vol. 100, no. 4, pp. 905-917, April 2012, doi: 10.1109/JPROC.2011.2182101.

[Tech'16] G. Tech, Y. Chen, K. Müller, J. R. Ohm, A. Vetro and Y. K. Wang, *Overview of the Multiview and 3D Extensions of High Efficiency Video Coding*, in IEEE Transactions on Circuits and Systems for Video Technology, vol. 26, no. 1, pp. 35-49, September 2015.

[TMIV] *Test Model 8 for MPEG Immersive Video*, Standardization document: ISO/IEC JTC1/SC29/WG04 MPEG VC N0050, Online, January 2021

[Vadakital'20] V. K. M. Vadakital, A. Dziembowski, D. Mieloch, *[MPEG-I Video][MIV] Using VVC Class-F configuration improves the performance of MIV*, MPEG131/m54390, online, June 2020

[Vetro'10] A. Vetro, *Frame compatible formats for 3D video distribution*, 2010 IEEE International Conference on Image Processing, 2010, pp. 2405-2408, doi: 10.1109/ICIP.2010.5651071.

[Vetro'11A] A. Vetro, T. Wiegand and G. J. Sullivan, *Overview of the Stereo and Multiview Video Coding Extensions of the H.264/MPEG-4 AVC Standard*, Proceedings of the IEEE, vol. 99, no. 4, pp. 626-642, April 2011, doi: 10.1109/JPROC.2010.2098830.

[Vetro'11B] A. Vetro, A. M. Tourapis, K. Muller and T. Chen, *3D-TV Content Storage and Transmission*, in IEEE Transactions on Broadcasting, vol. 57, no. 2, pp. 384-394, June 2011, doi: 10.1109/TBC.2010.2102950.

[Wang'04] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, *Image quality assessment: from error visibility to structural similarity*, IEEE Transactions on Image Processing, tom 13, nr 4, s. 600-612, 04.2004

[Wegner'14] K. Wegner, O. Stankiewicz, K. Klimaszewski, M. Domański, *Poznan Blocks – a multiview video test sequence and camera parameters for FTV*, ISO/IEC JTC1/SC29/WG11, Doc. MPEG M32243, San Jose, January 2014.

[Xu'16A] J. Xu, R. Joshi, and R. A. Cohen, *Overview of the Emerging HEVC Screen Content Coding Extension*, in IEEE Transactions on Circuits and Systems for Video Technology, vol. 26, no. 1, pp. 50-62, January 2016.

[Xu'16B] X. Xu et al., *Intra Block Copy in HEVC Screen Content Coding Extensions*, in IEEE Journal on Emerging and Selected Topics in Circuits and Systems, vol. 6, no. 4, pp. 409-419, December 2016.

[Xu'19] X. Xu, X. Li and S. Liu, *Current Picture Referencing in Versatile Video Coding*, 2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), San Jose, CA, USA, 2019, pp. 26-31, doi: 10.1109/MIPR.2019.00013.

[Xu'22] X. Xu and S. Liu, *Overview of Screen Content Coding in Recently Developed Video Coding Standards*, in IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 2, pp. 839-852, February 2022, doi: 10.1109/TCSVT.2021.3064210.

[Yu'15] H. Yu, R. Cohen, K. Rapaka, J. Xu, *Common Test Conditions for Screen Content Coding*, JCT-VC of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 21st Meeting: Doc. JCTVC-U1015r2, Warsaw, PL, June 2015.

[Zhang'11] J. Zhang, R. Li, H. Li, D. Rusanovskyy, and M. M. Hannuksela, *Ghost town fly 3DV sequence for purposes of 3DV standardization*, ISO/IEC JTC1/SC29/WG11, Geneva, Switzerland, Tech. Rec. M20027, March 2011.

[Zhang'15] L. Zhang, J. Chen, J. Sole, M. Karczewicz, X. Xiu and J. -Z. Xu, *Adaptive Color-Space Transform for HEVC Screen Content Coding*, 2015 Data Compression Conference, Snowbird, UT, USA, 2015, pp. 233-242, doi: 10.1109/DCC.2015.33.

[Zitnick'04] C.L. Zitnick, S. B. Kang, M. Uyttendaele, S. Winder and R. Szeliski, *High-quality video view interpolation using a layered representation*, ACM Transactions on Graphics, vol. 23, pp. 600-608, 2004

PUBLICATIONS OF THE AUTHOR

BOOK CHAPTERS

1. Kukolj, L. Bolecek, L. Polak, T. Kratochvil, O. Zach, J. Kufa, M. Slanina, T. Grajek, J. Samelak, M. Domański, D. Milovanovic, *3D Content Acquisition and Coding*, 3D Visual Content Creation, Coding and Delivery, Signals and Communication Technology: Springer International, 2019, pp. 41-95, ISBN: 978-3-319-77841-9, ISSN: 1860-4862

INTERNATIONAL JOURNALS

1. J. Samelak, A. Dziembowski, D. Mieloch, *Advanced HEVC Screen Content Coding for MPEG Immersive Video*, Electronics, Vol. 11, No. 23, December 2022, ISSN: 2079-9292

2. O. Stankiewicz, M. Domański, A. Dziembowski, A. Grzelka, D. Mieloch, J. Samelak, *A free-viewpoint television system for horizontal virtual navigation*, IEEE Transactions on Multimedia, tom PP, nr 99, s. 1-1, 2018

INTERNATIONAL CONFERENCES

1. D. Klóska, A. Dziembowski, J. Samelak, *Versatile input view selection for efficient immersive video transmission*, 31. International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision, WSCG 2023, Pilsen, Czech Republic, 2023-05-15

2. J. Samelak, A. Dziembowski, D. Mieloch, M. Domański, M. Wawrzyniak, *Efficient Immersive Video Compression using Screen Content Coding*, 29th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision : WSCG 2021, WSCG 2021, Prague, Czech Republic, 17-20 May 2021

3. A. Dziembowski, J. Samelak, M. Domański, *View selection for virtual view synthesis in free navigation systems*, International Conference on Signals and Electronic Systems, ICSES 2018, Kraków, Poland, September 10-12 2018

4. J. Samelak, J. Stankowski, M. Domański, *Efficient frame-compatible stereoscopic video coding using HEVC Screen Content Coding*, IEEE International Conference on Systems, Signals and Image Processing IWSSIP 2017, Poznań, Poland, 22-24 May 2017,

5. M. Domański, M. Bartkowiak, A. Dziembowski, T. Grajek, A. Grzelka, A. Łuczak, D. Mieloch, J. Samelak, O. Stankiewicz, J. Stankowski, K. Wegner, *New results in free-viewpoint television systems for horizontal virtual navigation*, IEEE International Conference on Multimedia and Expo, ICME 2016, Seattle, USA, 11-15.07.2016

6. J. Samelak, J. Stankowski, M. Domański, *Adaptation of the 3D-HEVC coding tools to arbitrary locations of cameras*, International Conference on Signals and Electronic Systems, ICSES 2016, Kraków, Poland, September 5-7 2016,

7. M. Domański, A. Dziembowski, T. Grajek, A. Grzelka, Ł. Kowalski, M. Kurc, A. Łuczak, D. Mieloch, R. Ratajczak, J. Samelak, O. Stankiewicz, J. Stankowski, K. Wegner, *Methods of high efficiency compression for transmission of spatial representation of motion scenes*, IEEE International Conference on Multimedia and Expo, ICME 2015, Turin, Italy, 29.06-3.07.2015

8. J. Stankowski, Ł. Kowalski, J. Samelak, M. Domański, T. Grajek, K. Wegner, *3D-HEVC extension for circular camera arrangements*, 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), Lisboa, Portugal, 8-10.07.2015

INTERNATIONAL CONFERENCES - STANDARDIZATION

1. A. Dziembowski, D. Mieloch, J. Samelak, *3D-HEVC in MIV verification tests*, ISO/IEC JTC 1/SC 29/WG 4 MPEG 136, M57753, October 2021

2. M. Domański, J. Samelak, S. Różek, T. Grajek, S. Maćkowiak, O. Stankiewicz, *[VCM] Stereoscopic and multiview video coding for machines*, MPEG130/m54407, Geneva, Switzerland, June 2020

3. J. Samelak, A. Dziembowski, D. Mieloch, M. Domański, *[MPEG-I Visual] Impact of HEVC profile on TMIV performance*, MPEG130/m53910, Alpbah, Austria, April 2020

4. J. Samelak, A. Dziembowski, D. Mieloch, M. Domański, *[MPEG-I Visual] HEVC-SCC in TMIV*, MPEG130/m53427, Alpbah, Austria, April 2020

5. M. Domański, J. Samelak, *Unified Screen Content and Multiview Video Coding - Experimental results*, Joint Video Exploration Team JVET of ITU-T SG 16 WP 3 and ISO/IEC JTC1/SC29/WG11, JVET-M0765, Marrakech, Morocco, 9-18 January 2019

6. M. Domański, A. Dziembowski, T. Grajek, A. Grzelka, D. Mieloch, R. Ratajczak, J. Samelak, O. Stankiewicz, J. Stankowski, K. Wegner, *Free-viewpoint television demonstration for sports events*, ISO/IEC JTC1/SC29/WG11 MPEG2018, M41994, Gwangju, South Korea, 22-26.01.2018

7. J. Samelak, O. Stankiewicz, M. Domański*, Do we need multiview profiles for future video coding generations?*, ISO/IEC JTC1/SC29/WG11 MPEG2017, M41699, Macao, China, October 2017

8. J. Samelak, O. Stankiewicz, M. Domański, *Do we need multiview profiles for future video coding generations?*, ISO/IEC JTC1/SC29/WG11 Joint Video Exploration Team (JVET), JVET_H0044, Macao, China, October 2017

9. J. Samelak, J. Stankowski, M. Domański, *Experimental results for frame –compatible multiview video coding using HEVC SCC*, ISO/IEC JTC1/SC29/WG11 Joint Collaborative Team on Video Coding (JCT-VC), JCTVC-Z0041, Geneva, Switzerland, January 2017

10. M. Domański, A. Dziembowski, A. Grzelka, Ł. Kowalski, D. Mieloch, J. Samelak, J. Stankowski, O. Stankiewicz, K. Wegner, *Experimental video conding software for free navigation applications*, ISO/IEC JTC1/SC29/WG11 MPEG2016, M39527, Chengdu, China, 17-21.10.2016

11. M. Domański, A. Dziembowski, A. Grzelka, Ł. Kowalski, D. Mieloch, J. Samelak, J. Stankowski, O. Stankiewicz, K. Wegner, *Coding results for Poznan Fencing 2 and Poznan Blocks 2 test sequences in free navigation scenario*, ISO/IEC JTC1/SC29/WG11 MPEG2016, M39215, Chengdu, China, 17-21.10.2016

12. M. Domański, A. Dziembowski, A. Grzelka, Ł. Kowalski, D. Mieloch, J. Samelak, J. Stankowski, O. Stankiewicz, K. Wegner, *[FTV AHG] Extended results of Poznan University of Technology proposal for Call for Evidence on free-viewpoint television*, ISO/IEC JTC1/SC29/WG11 MPEG2016, M38246, Geneva, Switzerland, 30.05-3.06.2016

13. M. Domański, A. Dziembowski, A. Grzelka, Ł. Kowalski, D. Mieloch, J. Samelak, J. Stankowski, O. Stankiewicz, K. Wegner, *[FTV AHG] Technical Description of Poznan University of Technology proposal for Call for Evidence on free-viewpoint television*, ISO/IEC JTC1/SC29/WG11 MPEG2016, M37893, San Diego, USA, 22-26.02.2016

14. M. Domański, J. Samelak, O. Stankiewicz, J. Stankowski, K. Wegner, *[FTV AHG] 3D-HEVC extensions for free navigation*, ISO/IEC JTC1/SC29/WG11 MPEG2016, M38101, San Diego, USA, 22-26 February 2016

NATIONAL JOURNALS

1. J. Samelak, C. Korzeniewski, Ł. Kowalski, J. Stankowski, *Zoptymalizowana implementacja dekodera HEVC*, Przegląd Telekomunikacyjny, No. 6/2014, Warsaw, Poland, June 2014, pp. 583-586

PAPERS IN THE PROCESS OF PUBLICATION

1. J. Samelak, M. Domański, *Multiview Video Compression Using Advanced HEVC Screen Content Coding*, https://arxiv.org/abs/2106.13574

2. J. Samelak, M. Domański, *Circular Rectification of 3D Video and Efficient Modification of 3D-HEVC*, https://arxiv.org/abs/2306.06285

APPENDIX: EXPERIMENTAL RESULTS

Table A.1. Circularly rectified camera parameters for ARC-HEVC (Section 4.4).

Parameter	View	Ballet	Breakdancers	BBB_Flowers	Poznan_Blocks
α	1	-0,39495	-0,45938	0,34034	0,64312
	2	-0,24341	-0,28765	0,22689	0,45021
	3	-0,09630	-0,12535	0,11345	0,26069
	4	0,05464	0,04794	0,00000	0,08178
	5	0,19954	0,21473	-0,11345	-0,12344
	6	0,34641	0,38939	-0,22689	-0,29812
	7	0,48447	0,55978	-0,34034	-0,48627
o_x	1	226,44198	53,14096	645,86711	1133,54227
	2	400,67535	234,18259	631,84719	1117,72118
	3	518,14647	467,25264	631,80032	1111,15753
	4	673,76843	694,48473	640,00000	1090,62062
	5	776,87453	845,11840	648,19968	1080,48483
	6	967,56713	1090,17933	648,15281	1057,95950
	7	1110,30450	1306,68418	634,13289	1039,62190
o_y		356,23763	374,51650	384,00000	562,05012
r		26,28679	22,67017	1,05000	30,18157
f_x		1914,95375	1882,34125	830,45711	1721,66111

Table A.2. PSNR [dB] values for the experiment evaluating ARC-HEVC in compression of multiview video (Section 4.6).

Encoder	QP	Ballet	Breakdancers	BBB_Flowers	Poznan_Blocks	average
3D-HEVC	25	40.43	38.83	40.24	41.94	**40.36**
	30	38.71	37.25	37.36	39.34	**38.17**
	35	36.79	35.52	34.59	36.69	**35.90**
	40	34.70	33.56	31.85	33.99	**33.53**
ANY-HEVC	25	40.49	38.85	40.26	41.99	**40.40**
	30	38.79	37.30	37.39	39.40	**38.22**
	35	36.91	35.61	34.62	36.77	**35.98**
	40	34.86	33.67	31.90	34.07	**33.63**
ARC-HEVC	25	40.49	38.86	40.28	42.00	**40.41**
	30	38.79	37.30	37.41	39.42	**38.23**
	35	36.92	35.60	34.64	36.79	**35.99**
	40	34.87	33.67	31.92	34.09	**33.64**

Table A.3. Bitrate [kbps] values for the experiment evaluating ARC-HEVC in compression of multiview video (Section 4.6).

Encoder	QP	Ballet	Breakdancers	BBB_Flowers	Poznan_Blocks	average
3D-HEVC	25	2758.83	4835.11	6188.50	7844.20	**5406.66**
	30	1188.92	1944.54	3214.43	3911.95	**2564.96**
	35	574.86	972.50	1760.20	2076.74	**1346.08**
	40	304.42	513.94	969.19	1126.31	**728.47**
ANY-HEVC	25	2694.10	4654.14	6099.87	7735.32	**5295.86**
	30	1147.64	1841.11	3161.51	3827.75	**2494.50**
	35	547.06	906.07	1721.46	2011.29	**1296.47**
	40	283.84	466.14	940.47	1075.03	**691.37**
ARC-HEVC	25	2698.13	4683.97	6101.67	7739.86	**5305.91**
	30	1145.83	1857.88	3160.50	3834.07	**2499.57**
	35	547.89	913.64	1719.77	2012.46	**1298.44**
	40	283.07	468.27	937.40	1076.48	**691.31**

Table A.4. PSNR [dB] values for the experiment comparing different view alignments (Section 5.3).

View alignment	QP	Balloons	BBB_Butterfly	Kendo	Newspaper	Poznan_Hall2	Poznan_Street	average
2×2	22	45.04	48.14	45.69	43.55	43.97	43.24	**44.94**
	27	42.84	45.29	43.72	40.50	42.35	39.84	**42.42**
	32	40.33	42.43	41.44	37.68	41.07	37.12	**40.01**
	37	37.42	39.68	38.76	34.95	39.44	34.70	**37.49**
1×4	22	45.04	48.14	45.68	43.54	43.97	43.24	**44.93**
	27	42.85	45.29	43.70	40.48	42.36	39.84	**42.42**
	32	40.31	42.42	41.41	37.66	41.09	37.12	**40.00**
	37	37.40	39.66	38.72	34.93	39.45	34.69	**37.47**
4×1	22	45.02	48.11	45.67	43.53	43.97	43.23	**44.92**
	27	42.80	45.27	43.69	40.48	42.33	39.82	**42.40**
	32	40.29	42.41	41.42	37.66	41.06	37.10	**39.99**
	37	37.39	39.66	38.74	34.94	39.42	34.69	**37.47**

Table A.5. Bitrate [Mbps] values for the experiment comparing different view alignments (Section 5.3).

View alignment	QP	Balloons	BBB_Butterfly	Kendo	Newspaper	Poznan_Hall2	Poznan_Street	average
2×2	22	30.39	17.34	20.23	55.05	50.64	150.11	**53.96**
	27	15.67	9.44	10.21	27.27	17.08	64.43	**24.02**
	32	8.84	5.13	5.65	13.80	8.35	26.81	**11.43**
	37	5.14	2.93	3.25	7.32	4.56	12.20	**5.90**
1×4	22	30.49	17.54	20.09	54.79	50.87	149.38	**53.86**
	27	16.03	9.42	10.08	26.57	17.79	64.73	**24.10**
	32	8.94	5.08	5.50	13.30	8.85	26.59	**11.38**
	37	5.25	2.83	3.10	7.05	4.84	11.77	**5.81**
4×1	22	29.23	16.18	19.60	53.99	50.34	147.72	**52.84**
	27	14.80	8.70	9.73	26.22	16.23	61.88	**22.93**
	32	8.23	4.82	5.35	13.05	7.71	25.14	**10.72**
	37	4.75	2.74	3.08	6.81	4.13	11.42	**5.49**

Table A.6. Encoding time [s] values for the experiment comparing different view alignments (Section 5.3).

View alignment	QP	Balloons	BBB_Butterfly	Kendo	Newspaper	Poznan_Hall2	Poznan_Street	average
2×2	22	7955	5164	6362	11119	18720	30717	**13340**
	27	6103	4273	4819	9060	11983	26219	**10409**
	32	5658	3655	3971	6667	9576	16965	**7749**
	37	4563	3612	3718	4463	7998	9475	**5638**
1×4	22	8892	5977	7473	12332	18178	36773	**14937**
	27	6887	4572	4927	9310	12566	27667	**10988**
	32	5324	3892	4363	6173	10123	18678	**8092**
	37	3752	3318	3629	4720	9292	9971	**5780**
4×1	22	7820	5277	7114	11442	18704	37657	**14669**
	27	6570	4185	5279	8608	11382	26721	**10458**
	32	4557	3909	3823	6407	9000	16750	**7408**
	37	3979	3467	3631	4315	7870	8869	**5355**

Table A.7. PSNR [dB] values for the experiment comparing SCC tools in compression of frame-compatible multiview video (Section 5.4).

SCC configuration	QP	Balloons	BBB_Butterfly	Kendo	Newspaper	Poznan_Hall2	Poznan_Street	average
configuration according to SCC CTC	22	45.02	48.11	45.67	43.53	43.97	43.23	**44.92**
	27	42.80	45.27	43.69	40.48	42.33	39.82	**42.40**
	32	40.29	42.41	41.42	37.66	41.06	37.10	**39.99**
	37	37.39	39.66	38.74	34.94	39.42	34.69	**37.47**
Intra Boundary Filter (enabled)	22	45.02	48.11	45.67	43.53	43.97	43.23	**44.92**
	27	42.81	45.27	43.70	40.48	42.33	39.82	**42.40**
	32	40.30	42.41	41.42	37.67	41.05	37.11	**39.99**
	37	37.40	39.66	38.74	34.94	39.43	34.70	**37.48**
Hash-Based IBC Search (disabled)	22	45.02	48.11	45.67	43.53	43.97	43.23	**44.92**
	27	42.80	45.27	43.69	40.48	42.33	39.82	**42.40**
	32	40.29	42.41	41.42	37.66	41.06	37.10	**39.99**
	37	37.39	39.66	38.74	34.94	39.42	34.69	**37.47**
Palette Mode (disabled)	22	45.02	48.11	45.67	43.52	43.97	43.23	**44.92**
	27	42.80	45.28	43.69	40.48	42.33	39.82	**42.40**
	32	40.30	42.41	41.42	37.66	41.05	37.10	**39.99**
	37	37.39	39.66	38.74	34.93	39.42	34.69	**37.47**
all improvements	22	45.02	48.10	45.67	43.53	43.97	43.23	**44.92**
	27	42.81	45.27	43.69	40.48	42.33	39.82	**42.40**
	32	40.30	42.41	41.42	37.67	41.05	37.11	**39.99**
	37	37.40	39.66	38.74	34.94	39.43	34.70	**37.48**

Table A.8. Bitrate [Mbps] values for the experiment comparing SCC tools in compression of frame-compatible multiview video (Section 5.4).

SCC configuration	QP	Balloons	BBB_Butterfly	Kendo	Newspaper	Poznan_Hall2	Poznan_Street	average
configuration according to SCC CTC	22	29.23	16.18	19.60	53.99	50.34	147.72	**52.84**
	27	14.80	8.70	9.73	26.22	16.23	61.88	**22.93**
	32	8.23	4.82	5.35	13.05	7.71	25.14	**10.72**
	37	4.75	2.74	3.08	6.81	4.13	11.42	**5.49**
Intra Boundary Filter (enabled)	22	29.17	16.16	19.58	53.82	50.29	147.37	**52.73**
	27	14.76	8.70	9.72	26.12	16.18	61.80	**22.88**
	32	8.21	4.81	5.35	13.01	7.68	25.08	**10.69**
	37	4.74	2.73	3.07	6.78	4.12	11.40	**5.48**
Hash-Based IBC Search (disabled)	22	29.23	16.18	19.60	53.99	50.34	147.72	**52.84**
	27	14.80	8.70	9.73	26.22	16.23	61.88	**22.93**
	32	8.23	4.82	5.35	13.05	7.71	25.14	**10.72**
	37	4.75	2.74	3.08	6.81	4.13	11.42	**5.49**
Palette Mode (disabled)	22	29.23	16.18	19.61	53.93	50.32	147.72	**52.83**
	27	14.79	8.70	9.73	26.19	16.22	61.90	**22.92**
	32	8.23	4.82	5.35	13.04	7.69	25.14	**10.71**
	37	4.75	2.76	3.07	6.80	4.13	11.42	**5.49**
all improvements	22	29.17	16.17	19.58	53.77	50.28	147.34	**52.72**
	27	14.75	8.69	9.72	26.12	16.18	61.77	**22.87**
	32	8.21	4.80	5.34	13.01	7.68	25.08	**10.69**
	37	4.74	2.77	3.07	6.78	4.12	11.39	**5.48**

Table A.9. Encoding time [s] values for the experiment comparing SCC tools in compression of frame-compatible multiview video (Section 5.4).

SCC configuration	QP	Balloons	BBB_Butterfly	Kendo	Newspaper	Poznan_Hall2	Poznan_Street	average
configuration according to SCC CTC	22	8085	5420	6309	11897	18940	33813	**14077**
	27	6311	4241	5198	9026	11405	23135	**9886**
	32	4786	3885	4139	6282	8976	14313	**7063**
	37	3644	3475	3361	4608	7634	9387	**5351**
Intra Boundary Filter (enabled)	22	7915	5623	6213	11888	19207	33008	**13976**
	27	6049	4402	5053	9156	10947	23089	**9783**
	32	4961	3798	3994	6420	8644	14725	**7090**
	37	4162	3420	3277	4631	8093	9781	**5561**
Hash-Based IBC Search (disabled)	22	8592	5343	6327	11513	18967	33015	**13959**
	27	6022	4256	4928	8642	11852	22813	**9752**
	32	4691	3901	4035	6221	9309	14290	**7075**
	37	3671	3482	3289	4409	7874	9314	**5340**
Palette Mode (disabled)	22	6181	4033	4815	7706	14475	25425	**10439**
	27	5074	3367	4029	6505	9147	19602	**7954**
	32	3989	3184	3240	4873	7634	12508	**5905**
	37	3541	2976	2776	3695	6404	8221	**4602**
all improvements	22	6271	4015	4950	7812	14846	25316	**10535**
	27	4924	3388	4079	6489	9181	19181	**7874**
	32	3864	3179	3293	5002	7229	12104	**5779**
	37	3010	2981	2762	3734	6224	7988	**4450**

Table A.10. PSNR [dB] values for the experiment evaluating SCC in compression of stereoscopic video (Section 5.7.2).

	Encoder	QP	Balloons	BBB_Butterfly	Kendo	Newspaper	Poznan_Hall2	Poznan_Street	average
All Intra	HEVC Main simulcast	22	45.20	48.41	45.93	43.56	44.92	43.17	**45.20**
		27	42.57	45.24	43.54	40.33	43.34	39.96	**42.50**
		32	39.59	42.17	40.82	37.33	41.44	37.11	**39.74**
		37	36.38	39.21	37.81	34.37	39.22	34.38	**36.90**
	HEVC Main side-by-side	22	45.19	48.40	45.93	43.56	44.92	43.17	**45.19**
		27	42.57	45.23	43.53	40.33	43.33	39.95	**42.49**
		32	39.58	42.15	40.80	37.32	41.44	37.11	**39.73**
		37	36.36	39.20	37.79	34.36	39.21	34.37	**36.88**
	HEVC SCC simulcast	22	45.19	48.40	45.92	43.55	44.91	43.16	**45.19**
		27	42.56	45.24	43.53	40.33	43.33	39.94	**42.49**
		32	39.57	42.16	40.81	37.32	41.44	37.09	**39.73**
		37	36.37	39.20	37.79	34.36	39.22	34.37	**36.88**
	HEVC SCC side-by-side	22	45.01	48.22	45.80	43.43	44.87	43.06	**45.07**
		27	42.32	45.06	43.34	40.20	43.26	39.80	**42.33**
		32	39.29	41.97	40.55	37.17	41.35	36.92	**39.54**
		37	36.03	38.96	37.49	34.17	39.13	34.16	**36.65**
Random Access	HEVC Main simulcast	22	44.06	47.41	44.72	42.99	44.10	41.80	**44.18**
		27	41.70	44.19	42.33	40.24	42.88	39.39	**41.79**
		32	38.84	41.07	39.54	37.40	41.22	36.94	**39.17**
		37	35.88	38.23	36.70	34.58	39.18	34.51	**36.51**
	HEVC Main side-by-side	22	44.05	47.40	44.71	42.99	44.10	41.81	**44.17**
		27	41.68	44.17	42.30	40.24	42.87	39.39	**41.77**
		32	38.82	41.05	39.50	37.39	41.18	36.94	**39.15**
		37	35.84	38.21	36.65	34.57	39.14	34.51	**36.49**
	HEVC SCC simulcast	22	44.03	47.39	44.70	42.97	44.09	41.79	**44.16**
		27	41.66	44.18	42.31	40.22	42.88	39.38	**41.77**
		32	38.81	41.04	39.53	37.38	41.20	36.93	**39.15**
		37	35.84	38.20	36.70	34.56	39.16	34.52	**36.50**
	HEVC SCC side-by-side	22	44.00	47.34	44.66	42.97	44.09	41.79	**44.14**
		27	41.59	44.09	42.23	40.22	42.87	39.38	**41.73**
		32	38.71	40.96	39.37	37.38	41.19	36.89	**39.08**
		37	35.69	38.09	36.48	34.52	39.14	34.45	**36.39**

Table A.11. Bitrate [kbps] values for the experiment evaluating SCC in compression of stereoscopic video (Section 5.7.2).

	Encoder	QP	Balloons	BBB_Butterfly	Kendo	Newspaper	Poznan_Hall2	Poznan_Street	average
All Intra	HEVC Main simulcast	22	12544.53	7964.31	8512.40	18411.66	11750.48	42779.69	**16993.84**
		27	7836.23	4721.33	5227.89	10680.18	6107.82	23191.02	**9627.41**
		32	4963.24	2791.18	3285.59	6304.77	3393.94	12585.00	**5553.95**
		37	3149.78	1699.81	2070.70	3778.06	1960.07	6972.65	**3271.84**
	HEVC Main side-by-side	22	12542.57	7947.78	8503.68	18411.24	11744.76	42775.63	**16987.61**
		27	7832.67	4709.75	5221.36	10673.15	6102.39	23187.73	**9621.17**
		32	4955.36	2782.11	3278.47	6298.27	3388.10	12583.93	**5547.71**
		37	3141.87	1691.49	2063.78	3774.14	1950.69	6971.66	**3265.61**
	HEVC SCC simulcast	22	12540.04	7970.06	8516.49	18406.82	11680.41	42500.74	**16935.76**
		27	7840.10	4727.75	5230.35	10677.97	6033.65	22932.12	**9573.66**
		32	4966.89	2798.03	3289.51	6309.37	3365.31	12405.49	**5522.43**
		37	3155.65	1704.58	2070.48	3779.54	1957.20	6880.40	**3257.97**
	HEVC SCC side-by-side	22	10234.71	5899.91	6936.42	16047.12	10931.90	37107.11	**14526.20**
		27	6006.02	3413.30	3965.15	8747.86	5217.57	18319.68	**7611.60**
		32	3616.03	1991.70	2345.14	4903.66	2776.71	9194.73	**4138.00**
		37	2182.50	1176.29	1403.45	2796.51	1568.48	4733.79	**2310.17**
Random Access	HEVC Main simulcast	22	1763.97	1408.72	1620.03	1763.16	1726.55	5079.73	**2227.03**
		27	950.37	754.06	855.64	949.53	733.34	2219.61	**1077.09**
		32	542.91	413.79	485.13	537.60	382.54	1090.32	**575.38**
		37	323.83	242.77	290.27	316.95	214.71	573.51	**327.00**
	HEVC Main side-by-side	22	1769.75	1398.84	1616.50	1762.79	1730.83	5077.29	**2226.00**
		27	950.84	748.02	854.01	946.64	737.23	2216.78	**1075.59**
		32	541.94	409.15	482.55	534.12	381.78	1088.20	**572.96**
		37	321.57	239.00	288.65	314.34	212.94	571.56	**324.68**
	HEVC SCC simulcast	22	1761.44	1402.42	1609.67	1762.95	1707.61	5054.18	**2216.38**
		27	947.04	750.23	850.81	948.29	727.50	2204.37	**1071.37**
		32	541.91	411.72	481.74	536.33	378.04	1079.25	**571.50**
		37	322.86	242.22	288.95	316.07	213.31	568.23	**325.27**
	HEVC SCC side-by-side	22	1605.55	1146.92	1432.18	1623.63	1677.24	4715.10	**2033.44**
		27	820.00	596.28	717.64	831.35	680.71	1908.83	**925.80**
		32	448.59	320.60	387.24	448.60	341.90	871.08	**469.67**
		37	255.27	185.50	219.52	254.52	186.40	431.39	**255.43**

Table A.12.Encoding time [s] values for the experiment evaluating SCC in compression of stereoscopic video (Section 5.7.2).

	Encoder	QP	Balloons	BBB_Butterfly	Kendo	Newspaper	Poznan_Hall2	Poznan_Street	average
All Intra	**HEVC Main simulcast**	22	805	852	812	837	1772	2203	**1213**
		27	687	828	658	697	1574	1904	**1058**
		32	604	834	569	663	1474	1673	**969**
		37	632	727	582	623	1448	1569	**930**
	HEVC Main side-by-side	22	792	814	764	868	1765	2202	**1201**
		27	710	847	671	765	1590	1847	**1072**
		32	644	837	661	664	1526	1679	**1002**
		37	598	678	579	605	1478	1543	**914**
	HEVC SCC simulcast	22	1537	1325	1108	1811	2732	5043	**2259**
		27	1241	1194	1102	1697	2094	4209	**1923**
		32	1185	1106	1059	1454	2117	3548	**1745**
		37	1001	1135	875	1228	1865	2738	**1474**
	HEVC SCC side-by-side	22	1474	1060	1270	1918	2777	5187	**2281**
		27	1344	918	1075	1627	2186	4230	**1897**
		32	1118	975	868	1331	1813	3124	**1538**
		37	889	806	772	1013	1664	2289	**1239**
Random Access	**HEVC Main simulcast**	22	2297	2242	2110	1859	5117	5037	**3110**
		27	2050	2118	1886	1621	4512	4300	**2748**
		32	1906	2263	1810	1552	4552	4027	**2685**
		37	1678	2199	1755	1530	4262	4080	**2584**
	HEVC Main side-by-side	22	2199	2477	2375	1791	4871	4765	**3080**
		27	1963	2427	1920	1636	4383	4167	**2749**
		32	1771	2316	2049	1836	4184	3919	**2679**
		37	1662	1952	1679	1719	4038	3818	**2478**
	HEVC SCC simulcast	22	3098	2729	3161	2093	5141	5215	**3573**
		27	2308	2308	2424	1754	3781	3843	**2736**
		32	1738	2009	1963	1343	3051	3109	**2202**
		37	1462	1731	1717	1234	2733	2685	**1927**
	HEVC SCC side-by-side	22	2920	2574	3089	2093	4770	5099	**3424**
		27	2204	2223	2382	1711	3615	3665	**2633**
		32	1680	1842	1808	1373	2933	2943	**2096**
		37	1450	1650	1465	1230	2485	2530	**1802**

Table A.13.PSNR [dB] values for the experiment evaluating SCC in compression of multiview video (Section 5.7.3).

	Encoder	QP	Balloons	BBB_Butterfly	Kendo	Newspaper	Poznan_Hall2	Poznan_Street	average
All Intra	HEVC Main simulcast	22	45.20	48.33	45.82	43.68	43.98	43.29	**45.05**
		27	43.14	45.51	43.95	40.70	42.40	39.99	**42.62**
		32	40.70	42.67	41.76	37.93	41.16	37.37	**40.27**
		37	37.83	39.97	39.15	35.24	39.56	35.00	**37.79**
	HEVC Main side-by-side	22	45.20	48.33	45.81	43.67	43.98	43.29	**45.05**
		27	43.13	45.49	43.94	40.68	42.39	39.99	**42.60**
		32	40.69	42.66	41.75	37.91	41.15	37.36	**40.25**
		37	37.82	39.95	39.14	35.20	39.55	34.99	**37.78**
	HEVC SCC simulcast	22	45.20	48.33	45.82	43.68	43.97	43.29	**45.05**
		27	43.13	45.50	43.94	40.69	42.39	39.98	**42.61**
		32	40.68	42.67	41.75	37.93	41.16	37.35	**40.26**
		37	37.82	39.95	39.14	35.23	39.56	34.98	**37.78**
	HEVC SCC side-by-side	22	45.02	48.10	45.67	43.53	43.97	43.23	**44.92**
		27	42.81	45.27	43.69	40.48	42.33	39.82	**42.40**
		32	40.30	42.41	41.42	37.67	41.05	37.11	**39.99**
		37	37.40	39.66	38.74	34.94	39.43	34.70	**37.48**
	Multiview HEVC	22	45.25	48.42	45.89	43.81	44.27	43.52	**45.19**
		27	43.12	45.65	43.97	40.82	42.44	40.13	**42.69**
		32	40.76	42.84	41.82	38.07	41.24	37.47	**40.37**
		37	37.96	40.13	39.25	35.37	39.71	35.14	**37.93**
Random Access	HEVC Main simulcast	22	43.98	47.37	44.67	42.90	42.66	40.85	**43.74**
		27	42.23	44.49	42.87	40.50	41.88	39.02	**41.83**
		32	39.93	41.64	40.60	37.99	40.84	36.99	**39.67**
		37	37.26	39.02	38.06	35.44	39.41	34.97	**37.36**
	HEVC Main side-by-side	22	43.97	47.37	44.66	42.89	42.66	40.85	**43.73**
		27	42.22	44.47	42.85	40.48	41.87	39.02	**41.82**
		32	39.91	41.63	40.57	37.97	40.83	36.98	**39.65**
		37	37.23	39.02	38.02	35.40	39.39	34.96	**37.34**
	HEVC SCC simulcast	22	43.96	47.36	44.66	42.89	42.65	40.84	**43.73**
		27	42.21	44.49	42.87	40.49	41.88	39.01	**41.83**
		32	39.90	41.64	40.59	37.98	40.84	36.98	**39.66**
		37	37.23	39.01	38.05	35.43	39.39	34.96	**37.35**
	HEVC SCC side-by-side	22	43.94	47.29	44.61	42.87	42.65	40.83	**43.70**
		27	42.16	44.39	42.77	40.46	41.88	39.01	**41.78**
		32	39.82	41.53	40.42	37.92	40.84	36.97	**39.58**
		37	37.11	38.87	37.83	35.34	39.38	34.90	**37.24**
	Multiview HEVC	22	43.50	46.70	44.17	42.40	42.39	40.34	**43.25**
		27	41.71	43.89	42.33	40.07	41.66	38.74	**41.40**
		32	39.34	41.13	40.05	37.58	40.57	36.79	**39.24**
		37	36.69	38.54	37.54	35.04	39.07	34.78	**36.94**

Table A.14. Bitrate [Mbps] values for the experiment evaluating SCC in compression of multiview video (Section 5.7.3).

	Encoder	QP	Balloons	BBB_Butterfly	Kendo	Newspaper	Poznan_Hall2	Poznan_Street	average
All Intra	HEVC Main simulcast	22	37.12	26.23	25.17	61.98	51.26	158.91	**60.11**
		27	22.14	14.92	14.66	34.31	18.95	79.07	**30.67**
		32	13.81	8.41	9.07	19.36	10.10	39.79	**16.76**
		37	8.76	4.90	5.71	11.23	5.71	21.25	**9.59**
	HEVC Main side-by-side	22	37.12	26.20	25.14	61.99	51.30	158.94	**60.12**
		27	22.12	14.89	14.63	34.30	18.92	79.07	**30.66**
		32	13.78	8.37	9.04	19.34	10.07	39.77	**16.73**
		37	8.74	4.86	5.68	11.20	5.67	21.22	**9.56**
	HEVC SCC simulcast	22	37.10	26.22	25.17	61.95	51.17	158.29	**59.98**
		27	22.14	14.93	14.66	34.29	18.78	78.34	**30.52**
		32	13.82	8.41	9.07	19.36	9.99	39.17	**16.64**
		37	8.77	4.90	5.71	11.23	5.67	20.95	**9.54**
	HEVC SCC side-by-side	22	29.17	16.17	19.58	53.77	50.28	147.34	**52.72**
		27	14.75	8.69	9.72	26.12	16.18	61.77	**22.87**
		32	8.21	4.80	5.34	13.01	7.68	25.08	**10.69**
		37	4.74	2.77	3.07	6.78	4.12	11.39	**5.48**
	Multiview HEVC	22	28.18	15.69	18.86	53.25	57.20	146.10	**53.21**
		27	13.54	8.56	8.73	25.57	15.34	58.69	**21.74**
		32	7.51	4.68	4.71	12.78	7.02	23.02	**9.95**
		37	4.42	2.63	2.73	6.72	3.73	10.38	**5.10**
Random Access	HEVC Main simulcast	22	5.56	4.63	4.89	6.18	9.56	24.64	**9.24**
		27	2.74	2.37	2.42	3.04	2.43	8.45	**3.57**
		32	1.55	1.26	1.35	1.64	1.16	3.61	**1.76**
		37	0.92	0.72	0.81	0.93	0.63	1.81	**0.97**
	HEVC Main side-by-side	22	5.56	4.61	4.88	6.18	9.58	24.65	**9.24**
		27	2.74	2.35	2.41	3.03	2.44	8.45	**3.57**
		32	1.55	1.24	1.35	1.63	1.16	3.61	**1.75**
		37	0.92	0.71	0.80	0.92	0.63	1.80	**0.96**
	HEVC SCC simulcast	22	5.53	4.62	4.86	6.17	9.46	24.53	**9.20**
		27	2.73	2.36	2.41	3.03	2.41	8.40	**3.56**
		32	1.54	1.25	1.35	1.64	1.15	3.58	**1.75**
		37	0.92	0.72	0.81	0.93	0.63	1.79	**0.97**
	HEVC SCC side-by-side	22	5.03	3.39	4.26	5.74	9.44	23.76	**8.60**
		27	2.23	1.65	1.89	2.57	2.29	7.40	**3.01**
		32	1.15	0.85	0.97	1.26	1.01	2.70	**1.32**
		37	0.64	0.48	0.54	0.66	0.53	1.19	**0.67**
	Multiview HEVC	22	3.43	2.42	2.82	4.29	6.06	14.69	**5.62**
		27	1.56	1.23	1.27	1.96	1.60	4.84	**2.07**
		32	0.82	0.65	0.66	0.97	0.73	1.82	**0.94**
		37	0.46	0.37	0.38	0.52	0.39	0.82	**0.49**

Table A.15.Encoding time [s] values for the experiment evaluating SCC in compression of multiview video (Section 5.7.3).

	Encoder	QP	Balloons	BBB_Butterfly	Kendo	Newspaper	Poznan_Hall2	Poznan_Street	average
All Intra	**HEVC Main simulcast**	22	3193	3263	2632	3984	7635	9546	**5042**
		27	2994	2928	2478	3013	6234	8189	**4306**
		32	2744	2965	2425	2824	6061	7669	**4115**
		37	2385	2613	2313	2625	5847	6940	**3787**
	HEVC Main side-by-side	22	2760	3228	2597	3147	6785	8814	**4555**
		27	2852	2793	2779	2787	6122	7542	**4146**
		32	2914	2659	2472	2575	5874	6715	**3868**
		37	2442	2831	2147	2317	5833	5971	**3590**
	HEVC SCC simulcast	22	4984	4966	4363	7470	11820	19263	**8811**
		27	4307	4350	4073	6196	8681	15930	**7256**
		32	3830	3867	3838	5069	7644	12338	**6097**
		37	3349	3603	3279	4390	7283	10532	**5406**
	HEVC SCC side-by-side	22	6165	3849	5777	7646	15177	26032	**10774**
		27	4931	3259	4802	6232	9269	19778	**8045**
		32	3875	3056	3829	5387	7349	13063	**6093**
		37	3021	3593	3227	3544	6093	8387	**4644**
	Multiview HEVC	22	8750	5647	8285	10773	21965	25064	**13414**
		27	7131	4708	6821	8495	15332	18754	**10207**
		32	5736	4135	5758	6746	11893	13644	**7985**
		37	4548	3906	4982	6254	10401	10180	**6712**
Random Access	**HEVC Main simulcast**	22	9358	9083	8856	7550	21438	21861	**13024**
		27	9020	8476	7793	6774	18021	17973	**11343**
		32	7906	8263	7343	6376	17305	16804	**10666**
		37	7458	7998	6937	6180	16713	16205	**10249**
	HEVC Main side-by-side	22	9001	8796	8545	7536	22950	23336	**13361**
		27	8080	8432	7932	6909	19447	18755	**11592**
		32	7195	7939	7384	6366	17861	17944	**10782**
		37	7002	7713	6834	6140	17722	17188	**10433**
	HEVC SCC simulcast	22	11833	9955	11570	8170	25699	27950	**15863**
		27	8502	8248	8623	6340	15306	16229	**10541**
		32	7093	7178	6902	5051	11931	13032	**8531**
		37	5138	6148	5738	4680	10835	11029	**7261**
	HEVC SCC side-by-side	22	11473	10284	11680	8804	25265	27346	**15809**
		27	8720	8608	9448	6840	14388	16532	**10756**
		32	6866	7224	6815	5096	12529	12862	**8565**
		37	5143	6154	5603	4720	10352	11286	**7210**
	Multiview HEVC	22	8910	9848	9157	8135	21681	21003	**13122**
		27	8402	9348	8426	7679	19217	18473	**11924**
		32	7555	9120	7929	7191	18667	17780	**11374**
		37	7577	8786	7513	6882	17648	17039	**10907**

Table A.16.PSNR [dB] values for the experiment evaluating ASCC in compression of multiview video (Section 6.4.2).

	Encoder	QP	Balloons	Shark	Kendo	Newspaper	Poznan_Hall2	Poznan_Street	Dancer	GT_Fly	average
All Intra	**HEVC**	25	44.15	43.46	44.88	41.96	42.80	41.21	40.32	41.86	**42.58**
		30	41.90	40.26	42.87	39.17	41.58	38.40	36.83	39.16	**40.02**
		35	39.15	37.18	40.41	36.42	40.05	35.96	33.76	36.50	**37.43**
		40	36.09	34.36	37.64	33.71	38.19	33.58	31.22	33.94	**34.84**
	SCC	25	43.90	42.89	44.74	41.72	42.75	41.13	40.22	41.88	**42.40**
		30	41.56	39.70	42.65	38.87	41.48	38.21	36.74	39.12	**39.79**
		35	38.78	36.66	40.14	36.10	39.94	35.71	33.67	36.40	**37.18**
		40	35.67	33.94	37.32	33.33	38.05	33.26	31.12	33.83	**34.57**
	ASCC	25	42.70	42.37	43.62	40.12	42.05	39.23	39.14	41.18	**41.30**
		30	40.42	39.43	41.56	37.69	40.98	37.22	36.14	38.64	**39.01**
		35	37.74	36.54	39.14	35.20	39.42	35.16	33.34	36.13	**36.58**
		40	34.84	33.90	36.47	32.64	37.50	32.99	30.91	33.69	**34.12**
	Multiview HEVC	25	43.39	42.69	44.25	40.95	42.40	40.20	39.69	41.60	**41.90**
		30	41.15	39.68	42.20	38.36	41.28	37.71	36.47	38.94	**39.47**
		35	38.43	36.73	39.76	35.75	39.77	35.49	33.58	36.32	**36.98**
		40	35.44	34.05	37.07	33.13	37.91	33.23	31.09	33.82	**34.47**
Random Access	**HEVC**	25	43.08	41.67	43.66	41.44	42.10	39.83	38.77	40.12	**41.33**
		30	40.96	38.65	41.47	39.01	41.16	37.85	35.88	37.78	**39.10**
		35	38.33	35.79	38.95	36.43	39.79	35.82	33.34	35.66	**36.76**
		40	35.49	33.25	36.27	33.84	38.07	33.73	31.05	33.53	**34.40**
	SCC	25	43.04	41.41	43.53	41.32	42.10	39.83	38.84	40.17	**41.28**
		30	40.89	38.36	41.30	38.90	41.17	37.85	35.95	37.75	**39.02**
		35	38.24	35.51	38.76	36.30	39.79	35.75	33.35	35.55	**36.66**
		40	35.37	33.07	36.10	33.67	38.05	33.58	31.01	33.41	**34.28**
	ASCC	25	42.31	41.01	42.78	40.44	41.85	39.20	38.24	39.76	**40.70**
		30	40.02	38.11	40.50	38.06	40.88	37.26	35.54	37.43	**38.48**
		35	37.40	35.37	38.04	35.58	39.45	35.30	33.09	35.30	**36.19**
		40	34.61	32.92	35.44	33.07	37.64	33.26	30.83	33.16	**33.87**
	Multiview HEVC	25	42.35	41.03	42.83	40.49	41.87	39.22	38.25	39.79	**40.73**
		30	40.07	38.11	40.54	38.11	40.88	37.28	35.53	37.44	**38.50**
		35	37.45	35.37	38.09	35.62	39.45	35.31	33.09	35.31	**36.21**
		40	34.66	32.93	35.48	33.10	37.66	33.28	30.83	33.16	**33.89**

Table A.17. Bitrate [kbps] values for the experiment evaluating ASCC in compression of multiview video (Section 6.4.2).

	Encoder	QP	Balloons	Shark	Kendo	Newspaper	Poznan_Hall2	Poznan_Street	Dancer	GT_Fly	average
All Intra	HEVC	25	16662	54222	10780	26113	16978	65074	104764	55531	**43765**
		30	10420	30623	6629	14820	8635	32526	54049	29010	**23339**
		35	6418	16033	4050	8381	4616	16678	24675	14321	**11896**
		40	4011	8000	2538	4890	2588	8998	10379	6843	**6031**
	SCC	25	12594	27959	8588	21205	15310	56125	60370	29744	**28987**
		30	7106	14131	4834	10705	6746	23378	28367	14537	**13726**
		35	4052	6770	2742	5545	3350	10362	11989	6660	**6434**
		40	2389	3197	1627	3012	1841	4956	5128	2996	**3143**
	ASCC	25	8154	20780	5376	12799	8168	28069	40730	21317	**18174**
		30	4763	11316	3077	6772	4024	12914	20424	10767	**9257**
		35	2801	5712	1803	3726	2097	6389	9357	5245	**4641**
		40	1688	2756	1090	2129	1153	3387	4037	2509	**2343**
	Multiview HEVC	25	9508	22273	6318	15440	10193	37110	45658	23906	**21301**
		30	5468	12007	3539	7945	4698	15283	22292	11831	**10383**
		35	3214	6062	2062	4285	2460	7288	10159	5655	**5148**
		40	1940	2943	1254	2441	1359	3789	4407	2706	**2605**
Random Access	HEVC	25	2209	10391	2076	2426	2373	7265	13833	10425	**6375**
		30	1219	5269	1125	1287	1004	2837	5961	4450	**2894**
		35	702	2645	647	708	517	1349	2707	2078	**1419**
		40	422	1300	391	410	287	700	1251	999	**720**
	SCC	25	1963	6482	1897	2128	2278	6901	10078	7258	**4873**
		30	1003	2978	958	1026	893	2371	4037	3008	**2034**
		35	548	1404	520	525	436	1009	1721	1356	**940**
		40	317	689	302	288	234	481	789	660	**470**
	ASCC	25	1362	4582	1314	1418	1376	3957	6677	4994	**3210**
		30	706	2244	667	699	579	1408	2796	2105	**1400**
		35	395	1103	372	375	301	642	1281	977	**681**
		40	226	532	215	211	166	323	609	471	**344**
	Multiview HEVC	25	1368	4592	1319	1421	1381	3946	6640	5003	**3209**
		30	710	2249	670	702	582	1409	2789	2112	**1403**
		35	399	1108	375	378	304	645	1284	983	**684**
		40	229	536	218	214	170	326	611	474	**347**

Table A.18. Encoding time [s] values for the experiment evaluating ASCC in compression of multiview video (Section 6.4.2).

	Encoder	QP	Balloons	Shark	Kendo	Newspaper	Poznan_Hall2	Poznan_Street	Dancer	GT_Fly	average
All Intra	HEVC	25	4526	12083	4277	5002	10593	13259	14047	12277	**9508**
		30	4210	11170	4132	4443	10240	11344	12310	11008	**8607**
		35	3786	10088	3642	4120	9742	10218	10853	10111	**7820**
		40	3774	9498	3641	3802	9260	10129	10066	9543	**7464**
	SCC	25	18491	43969	14727	25417	23865	62212	81485	39415	**38698**
		30	14445	29217	11820	17835	16588	39359	61752	29648	**27583**
		35	11155	21613	10030	13135	14483	24314	37925	20479	**19142**
		40	8641	15916	8607	9665	12135	16721	30617	14133	**14554**
	ASCC	25	15523	17924	14233	17437	19081	29521	22295	15339	**18919**
		30	11694	14551	11362	12886	13870	18192	16161	11831	**13818**
		35	8718	11789	8386	8922	10125	13160	13534	9924	**10570**
		40	6476	10893	6483	6897	8216	10592	11312	8363	**8654**
	Multiview HEVC	25	14961	17890	14282	16945	20237	33479	24286	16093	**19772**
		30	11410	14761	11089	13092	14371	21639	17684	12012	**14507**
		35	9195	12371	8875	9664	10420	14924	13963	9857	**11159**
		40	7036	11101	7040	7704	7646	11804	11464	8978	**9097**
Random Access	HEVC	25	9502	27869	9436	7089	17351	18395	24685	28859	**17898**
		30	7722	22826	7933	6288	14765	14204	19938	21248	**14366**
		35	6444	18589	6810	5594	13411	13102	16897	17833	**12335**
		40	5790	16565	6257	5222	12760	12400	15251	15062	**11163**
	SCC	25	29092	72791	26395	23422	35499	46779	75575	68570	**47265**
		30	21933	56102	18991	18467	26441	36431	59128	53902	**36424**
		35	17785	38555	14269	15574	23112	29613	44991	42820	**28340**
		40	15411	30665	11853	13547	19728	28009	36396	35678	**23911**
	ASCC	25	13905	34794	14319	11131	17047	22235	29536	28310	**21410**
		30	11303	28407	11761	9345	14328	17502	23217	22486	**17294**
		35	9450	24326	10024	8147	13398	16398	19250	19805	**15100**
		40	8233	21956	8913	7680	12205	15145	18036	17313	**13685**
	Multiview HEVC	25	13757	33077	14276	11059	17162	22183	28630	29970	**21264**
		30	11409	29961	11407	9247	14646	17978	23100	23444	**17649**
		35	9948	25034	10083	8237	13060	16107	20614	19779	**15358**
		40	8261	21383	8745	7710	12392	14022	16866	16474	**13232**

Table A.19. PSNR [dB] values for the experiment evaluating ASCC in compression of screen content video (Section 6.4.3).

Encoder	QP	All Intra		Random Access	
		SCC	ASCC	SCC	ASCC
ChineseEditing	25	47.10	47.17	48.23	48.39
	30	42.95	43.08	44.08	44.24
	35	38.59	38.75	39.72	39.94
	40	33.98	34.22	35.17	35.42
MissionControlClip3	25	47.00	47.01	47.45	47.47
	30	43.19	43.20	43.89	43.91
	35	39.22	39.23	40.14	40.17
	40	35.12	35.16	36.22	36.24
sc_console	25	54.98	55.11	54.44	54.54
	30	49.89	49.96	49.07	49.13
	35	44.60	44.72	43.54	43.67
	40	38.69	39.10	38.05	38.28
sc_desktop	25	51.44	51.73	52.71	52.98
	30	47.06	47.20	48.30	48.51
	35	42.09	42.26	43.50	43.78
	40	37.00	37.17	38.23	38.46
sc_flyingGraphics	25	49.01	48.83	44.56	44.51
	30	44.22	44.10	39.48	39.45
	35	39.50	39.49	35.07	35.07
	40	34.87	34.92	31.35	31.38
sc_map	25	47.72	47.71	48.64	48.68
	30	43.01	43.02	43.95	44.02
	35	38.58	38.63	39.39	39.47
	40	34.78	34.81	35.49	35.55
sc_programming	25	48.47	48.51	47.98	48.00
	30	44.41	44.47	43.66	43.70
	35	40.26	40.32	39.51	39.55
	40	35.93	36.03	35.74	35.81
sc_robot	25	42.52	42.52	40.96	40.96
	30	38.62	38.62	37.70	37.70
	35	35.62	35.62	35.14	35.14
	40	33.14	33.15	32.97	32.97
sc_web_browsing	25	50.95	51.22	52.03	52.27
	30	46.52	46.62	47.59	47.86
	35	41.89	41.95	43.24	43.45
	40	37.03	37.17	38.38	38.56
Basketball_Screen	25	46.41	46.40	46.64	46.66
	30	42.84	42.86	43.33	43.36
	35	39.18	39.20	39.88	39.91
	40	35.42	35.44	36.28	36.35
ChinaSpeed	25	45.21	45.22	42.74	42.74
	30	41.37	41.37	38.96	38.95
	35	37.75	37.76	35.41	35.41
	40	34.33	34.33	32.32	32.33
MissionControlClip2	25	46.22	46.20	46.36	46.36
	30	42.85	42.84	43.06	43.05
	35	39.44	39.42	39.71	39.71
	40	35.97	35.97	36.52	36.56
SlideShow	25	50.97	50.96	50.31	50.33
	30	47.39	47.38	46.63	46.63
	35	43.85	43.83	43.08	43.09
	40	39.97	40.02	39.55	39.58

Table A.20.Bitrate [kbps] values for the experiment evaluating ASCC in compression of screen content video (Section 6.4.3).

Encoder	QP	All Intra		Random Access	
		SCC	ASCC	SCC	ASCC
ChineseEditing	25	112008	117355	7117	7487
	30	85120	88979	5419	5660
	35	66106	69394	4113	4310
	40	49926	52969	3065	3242
MissionControlClip3	25	56212	57692	3838	3914
	30	39270	40307	2467	2521
	35	27227	28119	1628	1674
	40	18667	19555	1096	1138
sc_console	25	32477	41177	5314	5978
	30	26265	30077	4105	4467
	35	21109	24081	3151	3407
	40	16401	18883	2441	2645
sc_desktop	25	36047	42440	2865	3265
	30	28573	32240	2273	2480
	35	23252	26495	1864	2040
	40	19181	22374	1551	1717
sc_flyingGraphics	25	58146	63313	26891	28452
	30	43029	46115	15769	16537
	35	31799	34625	9375	9855
	40	23246	26034	5842	6192
sc_map	25	36522	37298	2779	2819
	30	24950	25477	1846	1878
	35	15860	16362	1161	1191
	40	10052	10391	720	742
sc_programming	25	25668	26790	4067	4135
	30	18499	19326	2364	2409
	35	13690	14435	1356	1393
	40	10096	10764	836	873
sc_robot	25	27780	27783	4067	4070
	30	13956	13973	1621	1621
	35	6753	6785	677	678
	40	3366	3398	318	320
sc_web_browsing	25	6394	7200	369	415
	30	4596	5207	268	296
	35	3448	4035	201	228
	40	2638	3287	153	182
Basketball_Screen	25	100573	103824	7705	7870
	30	68156	70545	4693	4816
	35	47387	49382	3036	3134
	40	33322	35274	2040	2128
ChinaSpeed	25	19399	19474	4900	4909
	30	12645	12708	2496	2502
	35	8171	8231	1275	1279
	40	5252	5310	671	676
MissionControlClip2	25	84664	84940	7972	7992
	30	56706	56890	4747	4755
	35	36662	36742	2816	2829
	40	23182	23424	1721	1741
SlideShow	25	4138	4116	631	629
	30	2656	2604	372	369
	35	1744	1727	229	227
	40	1135	1143	149	149

Table A.21. Encoding time [s] values for the experiment evaluating ASCC in compression of screen content video (Section 6.4.3).

Encoder	QP	All Intra		Random Access	
		SCC	ASCC	SCC	ASCC
ChineseEditing	25	55785	62980	9287	10712
	30	51116	58577	9123	9955
	35	48530	55544	8096	9208
	40	43745	47907	7807	8635
MissionControlClip3	25	37169	45036	14322	15525
	30	33742	40537	11940	13117
	35	28259	34715	10160	10788
	40	25898	29968	8582	9769
sc_console	25	26271	36459	27086	31521
	30	25583	31907	25757	29603
	35	23542	31301	23342	26069
	40	21821	27187	20704	24273
sc_desktop	25	33875	39601	11076	11867
	30	32097	40465	10747	11782
	35	30909	38506	9373	11177
	40	26667	33770	9038	10449
sc_flyingGraphics	25	21177	25763	46738	67792
	30	19925	23876	38783	54058
	35	17959	21347	33632	40124
	40	16093	19133	27840	33675
sc_map	25	19102	22423	5474	6199
	30	17085	20318	4951	5452
	35	15177	17539	4296	4766
	40	13099	14938	3610	3833
sc_programming	25	15632	18492	13528	15575
	30	13907	15487	11680	13006
	35	13029	15965	9285	10583
	40	11842	14340	7948	8993
sc_robot	25	12599	14157	13857	17015
	30	10374	11623	10385	10865
	35	8065	9152	6482	8432
	40	6067	7027	5573	6982
sc_web_browsing	25	5893	8148	968	1083
	30	5633	7313	902	1070
	35	5035	5959	846	999
	40	4701	5563	684	858
Basketball_Screen	25	35544	43427	12642	13614
	30	30693	38875	10104	11270
	35	27711	33262	7921	9096
	40	25387	30288	7146	7702
ChinaSpeed	25	15617	17423	33372	45562
	30	13882	15464	26384	33689
	35	11058	13228	19784	24109
	40	10130	11621	15439	16471
MissionControlClip2	25	33742	39817	14712	17158
	30	27923	34056	12746	14187
	35	24777	29968	11019	11644
	40	22373	25580	9259	9857
SlideShow	25	7995	10361	8765	9769
	30	7525	8630	6109	8600
	35	6974	8686	5502	7469
	40	5968	7396	4411	6472

Table A.22. PSNR [dB] values for the experiment evaluating ASCC in compression of immersive video, part 1 of 2 (Section 6.4.4).

Sequence	QP	SCC	ASCC					
			QPel IBC accuracy for views. FPel for depth	**QPel IBC accuracy for depth. FPel for views**	**QPel IBC accuracy**	**QPel + tile-based IBC for first views atlas**	**QPel + tile-based IBC for all views atlases**	**QPel + tile-based IBC for all atlases**
ClassroomVideo	25	34.40	34.40	34.40	34.40	34.38	34.38	34.38
	27	34.36	34.36	34.36	34.36	34.35	34.35	34.35
	30	33.41	33.41	33.41	33.41	33.39	33.38	33.38
	33	31.70	31.70	31.70	31.70	31.64	31.63	31.63
TechnicolorMuseum	21	32.06	32.06	32.06	32.06	32.06	32.06	32.05
	27	31.49	31.49	31.49	31.49	31.49	31.49	31.49
	33	29.73	29.73	29.73	29.73	29.72	29.72	29.71
	37	27.22	27.21	27.22	27.21	27.22	27.22	27.20
Fan	30	28.70	28.70	28.70	28.70	28.70	28.70	28.69
	38	28.50	28.50	28.50	28.50	28.51	28.51	28.51
	45	27.96	27.96	27.95	27.96	27.96	27.95	27.94
	48	26.70	26.70	26.69	26.69	26.68	26.67	26.65
OrangeKitchen	14	33.28	33.28	33.28	33.28	33.29	33.28	33.28
	21	32.87	32.87	32.87	32.87	32.87	32.87	32.86
	27	31.73	31.73	31.72	31.72	31.73	31.73	31.71
	33	29.69	29.69	29.68	29.69	29.67	29.65	29.63
TechnicolorPainter	22	37.75	37.75	37.75	37.75	37.76	37.76	37.76
	28	36.94	36.94	36.94	36.94	36.93	36.93	36.93
	35	35.07	35.07	35.08	35.08	35.05	35.05	35.04
	44	31.98	31.97	31.98	31.97	31.94	31.93	31.93
IntelFrog	30	30.86	30.86	30.86	30.86	30.84	30.84	30.84
	36	30.28	30.28	30.27	30.27	30.27	30.27	30.28
	43	29.07	29.07	29.07	29.06	29.07	29.07	29.05
	47	26.57	26.57	26.57	26.57	26.55	26.54	26.53
Poznan_Carpark	22	36.34	36.35	36.34	36.35	36.36	36.36	36.35
	26	35.48	35.48	35.48	35.48	35.51	35.51	35.50
	32	33.51	33.52	33.51	33.51	33.54	33.54	33.54
	39	30.42	30.40	30.41	30.40	30.40	30.40	30.39
Chess	11	36.50	36.50	36.49	36.50	36.49	36.49	36.49
	18	35.92	35.92	35.92	35.92	35.91	35.90	35.90
	25	34.49	34.50	34.50	34.51	34.50	34.49	34.47
	31	32.29	32.30	32.28	32.28	32.26	32.24	32.23
Group	24	30.53	30.53	30.53	30.53	30.53	30.53	30.51
	30	29.81	29.81	29.80	29.80	29.80	29.80	29.76
	35	28.09	28.09	28.09	28.09	28.08	28.08	28.03
	40	26.25	26.24	26.24	26.23	26.22	26.22	26.18
Poznan_Fencing2	22	35.14	35.14	35.14	35.14	35.13	35.12	35.12
	25	34.72	34.72	34.72	34.72	34.72	34.72	34.72
	32	33.67	33.68	33.67	33.68	33.67	33.67	33.68
	41	31.48	31.47	31.49	31.48	31.47	31.47	31.46
Poznan_Hall2	15	41.02	41.01	41.01	41.01	41.01	41.01	41.02
	23	40.46	40.46	40.45	40.45	40.46	40.46	40.44
	31	39.01	38.99	39.00	38.98	39.00	38.99	38.98
	40	36.17	36.16	36.17	36.16	36.16	36.15	36.15
Poznan_Street	20	36.55	36.55	36.55	36.55	36.54	36.54	36.55
	24	35.68	35.67	35.67	35.67	35.69	35.69	35.69
	29	33.90	33.92	33.90	33.92	33.91	33.91	33.91
	34	31.18	31.20	31.17	31.19	31.16	31.16	31.17

Table A.23.PSNR [dB] values for the experiment evaluating ASCC in compression of immersive video, part 2 of 2 (Section 6.4.4).

Sequence	QP	SCC	ASCC					
			QPel IBC accuracy for views. FPel for depth	QPel IBC accuracy for depth. FPel for views	QPel IBC accuracy	QPel + tile-based IBC for first views atlas	QPel + tile-based IBC for all views atlases	QPel + tile-based IBC for all atlases
ChessPieces	4	36.47	36.47	36.47	36.47	36.47	36.47	36.46
	11	36.13	36.13	36.11	36.11	36.11	36.11	36.09
	18	34.99	34.99	35.04	35.04	35.04	35.04	35.01
	26	33.09	33.08	33.08	33.07	33.08	33.05	33.03
TechnicolorHijack	16	39.66	39.66	39.66	39.66	39.66	39.66	39.66
	22	39.04	39.05	39.04	39.04	39.04	39.03	39.01
	29	37.43	37.44	37.41	37.42	37.40	37.39	37.38
	38	34.77	34.77	34.77	34.77	34.74	34.72	34.70
Mirror	25	30.22	30.22	30.22	30.22	30.22	30.22	30.22
	30	30.01	30.01	30.01	30.01	30.01	30.01	30.01
	35	29.39	29.39	29.39	29.39	29.40	29.40	29.40
	40	27.98	27.98	27.98	27.97	27.99	27.98	27.98

Table A.24.Bitrate [kbps] values for the experiment evaluating ASCC in compression of immersive video, part 1 of 2 (Section 6.4.4).

Sequence	QP	SCC	ASCC					
			QPel IBC accuracy for views, FPel for depth	QPel IBC accuracy for depth, FPel for views	QPel IBC accuracy	QPel + tile-based IBC for first views atlas	QPel + tile-based IBC for all views atlases	QPel + tile-based IBC for all atlases
ClassroomVideo	25	185613	185624	185601	185612	185832	186142	186937
	27	19563	19504	19530	19471	19478	19553	19886
	30	4477	4498	4462	4483	4510	4550	4765
	33	2194	2200	2194	2199	2202	2224	2355
TechnicolorMuseum	21	43378	43225	43294	43141	43581	43725	44251
	27	19583	19617	19556	19591	19745	19815	20008
	33	7063	7110	7056	7103	7142	7196	7310
	37	2386	2388	2384	2386	2403	2446	2516
Fan	30	157168	156570	157267	156669	157566	159508	162667
	38	64315	64346	64354	64385	64799	65554	68060
	45	30687	30825	30740	30878	30996	31226	32707
	48	17312	17346	17389	17423	17456	17530	18214
OrangeKitchen	14	21130	20152	21127	20149	21687	22062	22746
	21	9702	9529	9668	9495	10033	10174	10424
	27	4496	4550	4482	4536	4588	4671	4833
	33	2398	2413	2387	2401	2396	2450	2562
TechnicolorPainter	22	75286	74699	73764	73178	72291	72314	75150
	28	25897	25562	25338	25002	24710	24785	25742
	35	11696	11687	11558	11550	11366	11420	11886
	44	6390	6393	6378	6381	6300	6335	6564
IntelFrog	30	311619	311276	308184	307841	308431	310012	315711
	36	76344	76446	74892	74994	74812	75170	76916
	43	25700	25999	25431	25730	25515	25627	26429
	47	12098	12124	12053	12079	11872	11922	12312

Table A.25. Bitrate [kbps] values for the experiment evaluating ASCC in compression of immersive video, part 2 of 2 (Section 6.4.4).

Sequence	QP	SCC	ASCC					
			QPel IBC accuracy for views, FPel for depth	QPel IBC accuracy for depth, FPel for views	QPel IBC accuracy	QPel + tile-based IBC for first views atlas	QPel + tile-based IBC for all views atlases	QPel + tile-based IBC for all atlases
Poznan_Carpark	22	77999	77612	76352	75965	75864	75980	78312
	26	21371	21359	20758	20745	20196	20233	20774
	32	8294	8352	8176	8234	7863	7896	8144
	39	4081	4091	4048	4059	3904	3931	4038
Chess	11	20060	19305	20028	19273	19700	20204	20780
	18	8715	8653	8688	8625	8669	8788	9075
	25	3941	3947	3938	3944	3956	4005	4193
	31	2141	2143	2147	2149	2154	2173	2296
Group	24	68574	68515	68532	68473	68396	68630	69906
	30	28205	28223	28181	28199	28156	28287	28712
	35	9972	9994	9963	9986	9988	10083	10354
	40	4425	4425	4434	4434	4441	4485	4655
Poznan_Fencing2	22	83128	82837	81939	81648	81670	81918	83994
	25	20678	20679	20044	20045	20054	20182	21131
	32	9067	9091	8868	8893	8902	8967	9492
	41	4886	4893	4833	4839	4852	4893	5164
Poznan_Hall2	15	48358	47633	45107	44382	44812	45021	49872
	23	22427	22361	20174	20108	20080	20151	22640
	31	9861	9871	9144	9154	9100	9141	10167
	40	5279	5278	5087	5086	5066	5095	5421
Poznan_Street	20	67442	67137	65780	65476	64919	65000	67207
	24	16351	16285	15435	15369	14681	14717	15427
	29	5715	5874	5437	5595	5091	5119	5364
	34	2614	2688	2531	2606	2360	2382	2429
ChessPieces	4	12769	12311	12733	12275	12529	12885	13428
	11	6322	6285	6299	6262	6303	6421	6690
	18	3393	3394	3389	3391	3409	3476	3643
	26	2073	2074	2070	2071	2080	2113	2236
TechnicolorHijack	16	27186	26762	27179	26754	27016	27492	28642
	22	12634	12611	12618	12595	12650	12798	13455
	29	6975	6984	6961	6971	6984	7050	7485
	38	4509	4512	4496	4499	4510	4531	4832
Mirror	25	62769	61034	62662	60927	62669	64939	65598
	30	23476	22988	23423	22935	23929	24901	25216
	35	9305	9465	9285	9445	9762	9892	10138
	40	4113	4139	4110	4137	4191	4208	4367

Table A.26. Encoding time [s] values for the experiment evaluating ASCC in compression of immersive video, part 1 of 2 (Section 6.4.4).

Sequence	QP	SCC	ASCC					
			QPel IBC accuracy for views, FPel for depth	QPel IBC accuracy for depth, FPel for views	QPel IBC accuracy	QPel + tile-based IBC for first views atlas	QPel + tile-based IBC for all views atlases	QPel + tile-based IBC for all atlases
ClassroomVideo	25	17290	18284	17314	18308	7418	5577	5715
	27	6556	6982	6577	7002	3699	2586	2732
	30	4228	4596	4248	4617	2916	1972	2130
	33	3580	3868	3599	3887	2599	1752	1913
TechnicolorMuseum	21	5130	5512	5150	5531	3847	2213	2381
	27	4480	4909	4498	4928	3511	1996	2169
	33	4238	4652	4257	4671	3293	1843	2013
	37	4012	4417	4031	4435	3092	1702	1869
Fan	30	13835	14452	13948	14564	10500	5952	5121
	38	10048	10525	10155	10631	8079	4821	3999
	45	7047	7422	7144	7519	6085	4022	3250
	48	5226	5544	5322	5640	4783	3450	2717
OrangeKitchen	14	3751	4094	3768	4110	3365	2094	2301
	21	3385	3717	3402	3733	3122	1939	2141
	27	3102	3415	3118	3431	2948	1802	2010
	33	2860	3182	2876	3198	2756	1676	1884
TechnicolorPainter	22	12775	13852	12914	13991	11378	6579	5018
	28	9040	9967	9179	10106	8843	4804	3489
	35	7517	8394	7666	8544	7685	4030	2869
	44	6602	7417	6746	7560	6885	3502	2452
IntelFrog	30	32319	34095	32463	34239	20970	12611	9754
	36	19712	20945	19812	21046	14791	8257	6069
	43	13491	14544	13628	14681	11491	6271	4540
	47	10217	11186	10365	11334	9417	5150	3636
Poznan_Carpark	22	10298	10805	10351	10857	6365	5303	4604
	26	5447	5743	5503	5799	4136	3381	2872
	32	4041	4316	4103	4378	3368	2758	2349
	39	3374	3626	3427	3679	2929	2405	2064
Chess	11	4244	4533	4269	4558	3849	2352	2499
	18	3735	3993	3759	4016	3472	2115	2260
	25	3397	3660	3415	3677	3214	1944	2095
	31	3173	3411	3196	3433	3002	1796	1948
Group	24	7880	8310	7921	8351	5978	3397	3347
	30	6509	6921	6553	6965	4943	2838	2785
	35	5830	6193	5866	6229	4334	2515	2467
	40	5268	5655	5307	5694	3898	2236	2187
Poznan_Fencing2	22	16191	17174	16293	17276	11844	7378	6318
	25	8498	9058	8583	9143	7301	4744	3857
	32	6363	6829	6444	6910	5782	3827	3067
	41	5159	5551	5241	5634	4840	3257	2567
Poznan_Hall2	15	8772	9080	8809	9117	7455	5891	4238
	23	6760	7116	6726	7082	6030	4734	3439
	31	5371	5709	5386	5724	4960	3839	2856
	40	4473	4769	4536	4832	4221	3223	2455

Table A.27. Encoding time [s] values for the experiment evaluating ASCC in compression of immersive video, part 2 of 2 (Section 6.4.4).

Sequence	QP	SCC	ASCC					
			QPel IBC accuracy for views, FPel for depth	QPel IBC accuracy for depth, FPel for views	QPel IBC accuracy	QPel + tile-based IBC for first views atlas	QPel + tile-based IBC for all views atlases	QPel + tile-based IBC for all atlases
Poznan_Street	20	8548	9373	8563	9389	5664	4940	4300
	24	4820	5156	4836	5172	3723	3145	2676
	29	3658	3933	3679	3954	3028	2531	2193
	34	3065	3350	3097	3382	2641	2172	1912
ChessPieces	4	4426	4784	4447	4805	4056	2319	2468
	11	3917	4282	3938	4303	3683	2110	2256
	18	3624	3967	3648	3992	3448	1960	2108
	26	3423	3767	3442	3786	3244	1822	1973
TechnicolorHijack	16	6225	6899	6291	6965	6704	3353	3129
	22	5326	5998	5395	6068	6037	2985	2744
	29	4689	5357	4752	5419	5507	2696	2458
	38	4238	4860	4291	4913	5082	2461	2237
Mirror	25	8294	8503	8327	8536	7418	3849	3688
	30	5604	5700	5625	5721	5111	3012	2907
	35	4211	4296	4230	4314	3973	2573	2499
	40	3455	3534	3471	3549	3309	2250	2189