Professor Luce MORIN
INSA Rennes, IETR,
luce.morin@insa-rennes.fr

23 August 2023

Reviewer's opinion
on Ph.D. dissertation authored by
Yasir Abdulraheem Ahmed AL-OBAIDI
entitled:
Modeling of Codecs for 3-Dimensional Video

1. Problem and its impact

Immersive or interactive visualization of real content videos requires generating new views of the observed scene, different from the initially acquired views. In order to generate these "virtual" or "synthesized" views, information on the geometry of the acquired scene is needed. One of the most practical solutions for including geometry information is to build and transmit a multi-view plus depth (MVD) representation of the acquired data, and then use these views and depth maps at the decoder side to synthetize virtual views using depth image based rendering approaches. Both views and depth maps may be transmitted using state-of-the-art codecs designed for classical 2D video coding in a simulcast approach, or a dedicated MVD codec.
In his thesis, Yasir AL-OBAIDI has addressed the problem of rate control and rate allocation in MVD coding. Indeed, as depth maps and views are of different nature, bitrate allocation between views and depth data is not straightforward. Moreover, it should target to maximize the quality of the synthesized views. In practise, rate allocation and rate control are achieved by fixing the quantization parameter for both views and depth. The models linking quantization parameter and bitrate that were designed for classical 2D images are not valid for depth maps. Thus, finding new models specific to MVD data is necessary in order to achieve rate control for MVD transmission. The problems addressed in this dissertation thus have direct practical impact in applications using MVD coding, for instance in the case of immersive videos, interactive free viewpoint television (FTV) or multi-view displays.

2. Contribution

The two main contributions presented in the dissertation are, first, the design and evaluation of several methods for bitrate allocation between video and depth, and second, a rate control method for stereoscopic video plus depth coding. Two secondary contributions are also made : a study of the influence of depth map quality on the quality of the synthesized views, and rate control for HEVC based on AVC data.

In the following, these contributions are more precisely described and commented, following the order of the dissertation chapters.

Chapter 1 presents the context, motivation and general assumptions of the study in a clear and direct way. Chapter 2 is a literature survey while Chapter 3 presents the experimental set-up (see below sections 3 and 4 for more comments on these chapters).

Chapter 4 presents the first scientific contribution, consisting of a rate-control model for HEVC and VVC codecs based on the AVC rate control model. First, an experimental study shows the similarity of the R-Q model curves for a fixed video sequence, whatever the codec. The author does not provide an interpretation for such behaviour. However, this observation enables to use the AVC model parameters in order to estimate the R-Q model for HEVC and VVC. Based on these observations, Yasir AL-OBAIDI proposes a linear relationship between parameter *a* for AVC and parameter *a* for HEVC/VVC respectively.

Extensive experimental results are provided on the accuracy of the proposed model, with details on all sequences of the data set, which are provided in the appendix. Results show that the proposed method enables a significant gain of computing time, considering the number of encodings necessary to estimate the rate control R-Q model and the relative factor of 5 and 10 respectively between AVC encoding time and HEVC/VVC encoding time. Such a contribution has thus straightforward practical interest.

Chapter 5 presents one of the key contributions of the dissertation. In this chapter, Yasir AL-OBAIDI proposes several models to estimate bitrate allocation between views and depth maps, with the objective of optimizing the quality of synthesised views. The used assumption that the same *QP* value is applied on all depth images is plausible. By varying bitrate allocation between views and depth, Yasir AL-OBAIDI demonstrates that allocation choice does have an impact on the quality of the virtual views synthesized from decoded views and depth maps : this first experiment shows that it is necessary to build a model allowing to choose the adequate bitrate allocation between views and depth.

To build such a model, optimal view and depth quantization parameters pairs *(QP,QD)* are estimated from the rate-distortion Pareto curve of the synthesized views. Yasir AL-OBAIDI proposes to estimate the value of depth quantization parameter *QD* as a function of the value of view quantization parameter *QP*. He does not provide the intuition for choosing such a relationship, probably starting from the basic identity function *QD=QP* used in most existing MVD compression schemes.

The candidate experimentally validated the proposed model on verification data different from the training data, for codec-specific models and a global model for all codecs. Comparison with the reference choice *QD=QP*, and optimal choice shows that the proposed models improve rate-distortion performance with respect to the reference, PSNR wise, and also IV-PSNR wise. The proposed specific and global models also overcome alternative state-of-the-art depth and view allocation approaches. The used methodology is very clearly explained. The dissertation provides very complete results, with tables and figures allowing to quantize the gains obtained with respect to the reference and state-of-the-art methods.

A last contribution presented in Chapter 5 is a model of order 3, defined by 4 parameters, to estimate view bitrate over total bitrate as a function of *QP*, based on optimal *(QP, QD)* pairs. It is not very clear whether this model is linked to the previous models or if it is an alternative approach.

For some of the test sequences, it was observed in Chapter 5 that an increase in rate would produce a decrease in view synthesis quality. Chapter 6 aims to study and explain this unusual behavior. This study is a very interesting and original contribution, as this type of behavior is typical in MVD compression and is problematic in a rate-distortion optimization context. In this study, Yasir AL-OBAIDI chooses to add high-frequency noise to depth maps. He demonstrates that if the depth map

contains high-frequency noise, a compressed, and thus low-pass filtered depth map will produce a synthesized view of higher quality. It is not said if such a behavior is also observed for other quality metrics (IV-PSNR, visual quality…).
One can wonder about the relative impact of the different algorithms implied in the process: depth estimation algorithm, compression algorithms and artifacts, view synthesis algorithms, type of added noise. In particular, it would have been interesting to test different types of noise such as low-pass filtering or geometric noise (image motion), to comfort the conclusions of the study.
Another point is the type of test video sequences, either natural or computed-generated, which has an impact on the precision of the uncompressed depth map. This is briefly mentioned in Chapter 6's conclusion, but the relation between sequence type and rate-distortion behavior could have been further studied.
Overall, the study in Chapter 6 is very interesting. It is not mentioned if its conclusions may be used for the targeted problem of rate allocation and rate control.

In Chapter 7, Yasir AL-OBAIDI presents a method to perform rate control for MVD videos at GOP and frame levels, by using the R-Q model proposed by Grajek on MVD sequences. The initial 3-parameter model, and reduced versions with 2 or 1 content dependent parameters are derived and evaluated. The experimental results show that this model provides an accurate enough estimation of the R-Q curve for each of the tested sequences.

In chapter 8, Yasir AL-OBAIDI proposes to combine the previously proposed models to build a rate control method in the CBR (constant bitrate) context, for MVD videos compression with the different considered coding schemes: HEVC simulcast, VVC simulcast, MV-HEVC et 3D-HEVC. The description of the method is not as clear as in other chapters (see Section 3 for comments) and experiments are not as extensive as in previous chapters. However, this chapter provides a main contribution of the dissertation, as it demonstrates the validity of the previously proposed models for the practical issue of rate control in MVD coding, which was the main objective of the thesis. A few comments can be made. The experiments show that the proposed method enables to predict the target QP, with more precision as the initial QP value is closer to the target QP, which might be a limitation in practice. The precision is evaluated on the QP value and not on the target bitrate, which is the final objective, and for which a precision in percent would seem more natural and easy to appreciate. Only encoder models with 1 parameter were used in these experiments. It would have been interesting to compare the results obtained with the 2- and 3-parameter models. No quantitative comparison is provided with respect to alternative rate control state-of-the-art method, which would have been interesting to quantitatively evaluate the gains provided by the proposed method.
Overall, the contribution of Chapter 8 validates the models proposed in the dissertation as a practical method to implement rate control for MVD video coding.

Chapter 9 provides a summary of the dissertation, first reminding the main contributions. Well-chosen future work tracks are also proposed.

3. Correctness

The dissertation is very clearly structured, well written and very easy to read. Working assumptions are clearly stated. Yasir AL-OBAIDI has employed a clear and sound methodology for theoretical and experimental parts for each of his contributions. All necessary details are provided in a synthetic way, on both test data set, used algorithms, and experimental results. Result analysis is comprehensive and well organized. Chapter 3 providing an overview on the test sequences and the experimental methodology is very helpful in making clear the methodology used in the dissertation. The appendix providing curves and tables for all test sequences enables to have a complete reporting of the experimental results.

A few minor comments are given bellow, rather for suggestions of improvement, but they do not affect the correctness of the dissertation.

In Chapter 3, it would be useful to indicate if test sequences are natural or computer generated contents, together with information on the characteristics of depth contents (depth range and complexity, quality of original depth map). A typical view and depth map for each test sequence (in the appendix if needed) would also be an interesting information for the reader.

In Chapter 5, schemes from Figure 5.2 should all have exactly the same scale in order to facilitate visual comparison.

In Chapter 6, Figure 6.4, magenta line does not show measure points, contrary to other curves.

A noise with amplitude 1 to 3 is added to the depth map, however the range of depth values is not provided. Moreover, it is not indicated if this noise is added on the depth value or on the depth map value which is usually an inverse function of depth, as illustrated by provided depth images where larger (brighter) values correspond to closer points.

In Chapter 9, the bitrate control algorithm is not clear for some steps:

- Step 1 : it is not clear whether $QD_{initial}$ is computed from $QP_{initial}$ using one of the models presented in Chapter 5, and if so, which model is used (specific or general).
- Step 3 : it is not mentioned how values *b* and *c* are obtained, and in the case they are fixed, which values are chosen.
- Step 4 : it is confusing as two steps are mixed up in the description : the computations of $Q_{calculated}$ and the computation of $Bitrate_{goal}$. Furthermore $Bitrate_{goal}$ is expected to be an input of the algorithm, but I understand that the algorithm input is QP_{goal}, from which QD_{goal} and $Bitrate_{goal}$ are deduced, using a Chapter 5 model (which one?) and by running a MVD encoding, respectively. This is done before step 1, while in step 4, $Q_{calculated}$ is computed from 8.2 as a function of $Bitrate_{goal}$, with *b* and *c* still fixed to same values as Step 3. So, if this is right, it would be much clearer to separate Step 4 into 2 steps, one being before Step 1.

In section 8.3, it is claimed that the goal is to reach the target bitrate $Bitrate_{goal}$, but indeed, what is done is to scan a set of QP_{goal} values and to try to reach these QP_{goal} values through the proposed algorithm.

4. Knowledge of the candidate

The dissertation shows that Yasir AL-OBAIDI has general knowledge in the field of Information and Communication Technology, more specifically in the domain of video compression and parametric model fitting and estimation.

More precisely, Chapter 2 is a literature survey that provides a well-written and synthetic overview of the main issues in multi-view video plus depth coding. The main characteristics of classical hybrid coding methods and the data representations suitable for immersive video are briefly presented. A focus is made in this chapter on rate control, bit allocation, and multi-view video plus depth compression, which are the core topics addressed in the dissertation. The key issues are well introduced and the literature referenced in this chapter is complete, including seminal papers, as well as recent publications. Chapter 3 presents the methodology used for the experiments. The goal of the

experiments are clearly settled, the data (test sequences) and used metrics are presented in details and the choices are justified. This chapter also shows that the candidate is able to build a well-designed experimental setup and produce sound and clearly defined results.
In Chapters 5, 7 and 8, Yasir AL-OBAIDI also shows that he masters model estimation methods for data fitting, as well as the evaluation of such parametric models.

5. Conclusion

As a conclusion, Yasir AL-OBAIDI has tackled a scientific issue with practical meanings in the domain of rate allocation and rate control for multi-view plus depth video coding. He has proposed several original and meaningful scientific contributions, validated by extensive and well-designed experiments. These contributions have led to publications in international conferences. I thus consider that Yasir AL-OBAIDI's knowledge and contribution presented in this dissertation, his thesis work is certainly worthy to be defended in order to be granted the PhD title in Information and Communication Technology.

Taking into account what I have presented above and the requirements imposed by Article 13 of *the Act of 14 March 2003 of the Polish Parliament on the Academic Degrees and the Academic Title* (with amendments)[1], my evaluation of the dissertation according to the three basic criteria is the following:

A. Does the dissertation present an original solution to a scientific problem? (the selected option is marked with **X**)

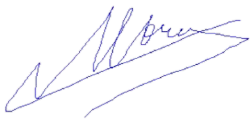
X				
Definitely YES	*Rather yes*	*Hard to say*	*Rather no*	*Definitely NO*

B. After reading the dissertation, would you agree that the candidate has general theoretical knowledge and understanding of the discipline of **Information and Communication Technology**, and particularly the area of Video compression ?

X				
Definitely YES	*Rather yes*	*Hard to say*	*Rather no*	*Definitely NO*

C. Does the dissertation support the claim that the candidate is able to conduct scientific work?

X				
Definitely YES	*Rather yes*	*Hard to say*	*Rather no*	*Definitely NO*



Signature

[1] http://www.nauka.gov.pl/g2/oryginal/2013_05/b26ba540a5785d48bee41aec63403b2c.pdf