

Poznań University of Technology
Faculty of Computing and Telecommunications
Institute of Multimedia Telecommunications

Doctoral Dissertation

Modeling of Codecs for 3-Dimensional Video

YASIR ABDULRAHEEM AHMED AL-OBAIDI

Supervisor: prof. dr hab. inż. Marek Domański

Auxiliary supervisor: dr. inż. Tomasz Grajek

Poznań, May 2023

Abstract

The dissertation deals with encoder modeling for multiview plus depth video coding with applications to bitrate control. The dissertation concerns three generations of video encoders (AVC, HEVC, and VVC) thus searching for a general approach for encoder modeling valid for all three generations of video compression encoders. The study comprises both simulcast coding (like basic HEVC) and multiview coding (like MV-HEVC) thus providing also some useful results for monoscopic video encoders. A new approach to bitrate control and bitrate allocation for stereoscopic video plus depth is presented in the dissertation. The proposed approach to bitrate control depends on two models: the bitrate allocation model and the encoder model. This approach is optimized to produce an output bitrate of the video encoder equal to the required bitrate.

The dissertation presents an original unified approach applicable to encoder modeling in all the abovementioned scenarios. The approach consists of the application of the universal encoder model that demonstrates reasonable accuracy of approximation of bitrate as a function of the quantization step for transform coefficients of prediction residuals. Moreover, the problem of the bitrate allocation between a pair of stereoscopic views and the corresponding depth maps is considered in the context of maximization of virtual video quality for the given total bitrate for views and depths. As the results, in the dissertation, the entire original procedure is provided for bitrate estimation for a given quantization step for transform coefficients of prediction residuals. With the use of experimental data, it is also demonstrated the proposed approach is useful for bitrate control for stereoscopic video plus depth.

In practice, the values of the quantization steps are defined by a quantization scale parameter or quantization parameter QP for video. Similarly, a quantization parameter QD for depth is used. In order to derive the bitrate allocation model for stereoscopic video plus depth, the optimum QP - QD pairs are calculated. These optimum QP - QD pairs are the pairs of the quantization parameters for video (QP) and depth (QD) that achieve the best quality of the virtual view for a given bitrate for multiview video plus depth (MVD) sequences. Then, the bitrate allocation model is derived depending on optimum QP - QD pairs. The proposed models are used to estimate the quantization parameter for depth based on the quantization parameter for video components in multiview video and depth maps compression. The proposed models are compared to the models presented in the previous studies: the straightforward approach ($QP = QD$) and the approach presented by the ISO/IEC MPEG group. The efficiency of the proposed method for bitrate allocation between videos and depth maps is also presented in the dissertation.

In the study of the bitrate allocation issue, some sequences present unexpected and surprising behavior in some bitrate range, as an increase in the quality of the virtual views produced from decreasing the bit allocation for the depth component under the video bitrate constancy condition. Therefore the influence of depth map quality on virtual view quality is studied in the dissertation and the respective explanations of the phenomenon are given.

In the dissertation, the encoder model for stereoscopic video plus depth is derived. The proposed model is used to estimate the bitrate or frame size of stereoscopic video plus depth depending on the quantization step size for the video (R-Q model). The accuracy of the encoder

model is carefully studied and demonstrated by the results of extensive experiments with the respective test video sequences.

For the 2D video, this dissertation also presents a new method to compute rate control for HEVC and VVC based on AVC data. This method aims to calculate the proposed models' parameters for HEVC and VVC codecs based on the model parameters for the AVC codec to reduce the required time for bitrate estimation. The effectiveness of the proposed method of rate control for HEVC and VVC is studied in the dissertation.

In order to verify the accuracy of the proposed models to control the bitrate, the relative approximation errors are computed between the experimental data and approximate data calculated by the proposed models. The experiments are performed for a set of well-known and widely accepted test sequences approved by ISO/IEC MPEG experts for the evaluation of new compression techniques. The results prove that the accuracy of the models is sufficient for bitrate control tasks.

Streszczenie

Rozprawa dotyczy modelowania kodeków dla zastosowań w sterowaniu prędkością bitową w kodowaniu wizji wielowidokowej wraz z mapami głębi. Rozprawa traktuje o trzech generacjach koderów wizyjnych (AVC, HEVC i VVC) i poszukuje się w niej ogólnego podejścia do modelowania koderów przydatnego dla wymienionych trzech generacji koderów wizyjnych. Studium obejmuje kodowanie wielowidokowe (jak za pomocą koderów MV-HEVC), a także kodowanie niezależne widoków i głębi (jak za pomocą podstawowych koderów HEVC), i w ten sposób prezentuje wyniki użyteczne także dla kodowania wizji monoskopowej. Rozprawa przedstawia nowe podejście do sterowania prędkością bitową i alokacji bitów pomiędzy widokami i mapami głębi. Zaproponowane podejście wykorzystuje dwa modele: model alokacji bitów i model kodera. Takie podejście jest zaplanowane tak, by można było uzyskiwać założoną prędkość bitową na wyjściu kodera.

Rozprawa prezentuje ogólne podejście do modelowania koderów w wymienionych wyżej scenariuszach zastosowań. To podejście wykorzystuje uniwersalny model kodera, który zapewnia wystarczającą dokładność szacowania prędkości bitowej w zależności od kroku kwantowania współczynników transformaty błędów predykcji. Ponadto problem alokacji bitów pomiędzy widokami pary stereoskopowej oraz odpowiadającymi im mapami głębi jest rozpatrywany w kontekście maksymalizacji jakości syntetycznej wizji wirtualnej dla danej łącznej prędkości bitowej widoków i głębi. W ten sposób rozprawa przedstawia całą procedurę estymacji prędkości bitowej dla danego kroku kwantowania współczynników transformaty błędów predykcji. Z wykorzystaniem danych eksperymentalnych pokazuje się, że zaproponowane podejście jest użyteczne dla sterowania prędkością bitową dla wizji stereoskopowej wraz z głębią.

W praktyce wartości kroków kwantowania są definiowane poprzez parametr skali kwantyzatorów albo parametr kwantyzacji QP dla wizji. Analogicznie wykorzystuje się parametr kwantyzacji QD dla głębi. Aby uzyskać model alokacji bitów dla stereoskopowych sekwencji wizyjnych z mapami głębi w pracy wyznacza się optymalne pary $QP-QD$. Optymalna para $QP-QD$ to taka para wartości parametrów kwantyzacji dla sekwencji wizyjnych (QP) i map głębi (QD), dla której uzyskuje się najlepszą jakość widoku wirtualnego dla danej prędkości bitowej dla sekwencji wielowidokowych z mapami głębi (MVD). Następnie, na podstawie optymalnych par $QP-QD$, wyznacza się model podziału bitów pomiędzy widoki i mapy głębi. Zaproponowane w pracy modele służą do szacowania wartości parametru kwantyzacji dla map głębi na podstawie parametru kwantyzacji dla sekwencji wizyjnych. Zaproponowane modele są porównane z modelami zaprezentowanymi w literaturze: modelem prostym ($QP=QD$) oraz modelem zaprezentowanym przez grupę ekspertów MPEG afiliowaną przy ISO/IEC. W pracy oceniono także efektywność zaproponowanych modeli.

Podczas badań nad problemem alokacji bitów zaobserwowano, że niektóre sekwencje testowe wykazują nieoczekiwane i zaskakujące zachowanie w pewnym zakresie prędkości bitowych, jak np. poprawa jakości widoków wirtualnych uzyskana w wyniku zmniejszenia liczby bitów dla map głębi w trybie stałej prędkości bitowej. W związku z tym, w pracy

zbadano wpływ jakości map głębi na jakość widoków wirtualnych i przedstawiono analizę zjawiska.

W pracy zaproponowano także model koderów dla stereoskopowych sekwencji wizyjnych z mapami głębi. Model ten wykorzystywany jest do oszacowania prędkości bitowej lub rozmiaru ramki w zależności od kroku kwantyzacji dla sekwencji wizyjnej (model $R-Q$). Dokładność modelu koderów została zbadana w trakcie obszernych eksperymentów z sekwencjami testowymi.

Dla dwuwymiarowych sekwencji wizyjnych, niniejsza rozprawa przedstawia również nową metodę sterowania prędkością bitową dla koderów HEVC i VVC w oparciu o dane dla koderów AVC. Metoda ta ma na celu obliczenie parametrów proponowanych modeli dla kodeków HEVC i VVC w oparciu o parametry modelu dla kodeka AVC, co pozwala na znacząco redukcję czasu estymacji parametrów. Skuteczność proponowanej metody sterowania przepływnością dla HEVC i VVC została wykazana w rozprawie.

Aby zweryfikować dokładność proponowanych modeli do sterowania prędkością bitową, obliczono względny błąd aproksymacji pomiędzy danymi eksperymentalnymi a danymi oszacowanymi przez proponowane modele. Eksperymenty przeprowadzono dla zestawu dobrze znanych i powszechnie akceptowanych sekwencji wizyjnych zatwierdzonych przez ekspertów ISO/IEC MPEG do oceny nowych technik kompresji. Wyniki eksperymentów pokazują, że dokładność zaproponowanych modeli jest wystarczająca dla zadań związanych ze sterowaniem koderami wizyjnymi.

List of Publications

- M. Domański, **Y. Al-Obaidi**, T. Grajek, "Universal modeling of monoscopic and multiview video codecs with applications to encoder control," IEEE Int. Conf. Image Proc. (ICIP), Anchorage, Alaska, USA, Sep. 2021. (70 pt)
- **Y. Al-Obaidi**, T. Grajek, "Estimation of the optimum depth quantization parameter for a given bitrate of multiview video plus depth in 3D-HEVC coding," Int. Conf. Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG), Pilsen, Czech Republic, May 2020. (70 pt)
- **Y. Al-Obaidi**, T. Grajek, M. Domański, "Quantization of depth in simulcast and multiview coding of stereoscopic video plus depth using HEVC, VVC and MV-HEVC," Picture Coding Symposium (PCS), Ningbo, China, Nov. 2019. (20 pt)
- **Y. Al-Obaidi**, T. Grajek, "Influence of depth map fidelity on virtual view quality," Int. Conf. Signals and Electronic Systems (ICSES), Kraków, Poland, Sep. 2018. (20 pt)
- **Y. Al-Obaidi**, T. Grajek, O. Stankiewicz, M. Domański, "Bitrate allocation for multiview video plus depth simulcast coding," Int. Conf. Systems, Signals, and Image Proc. (IWSSIP), Maribor, Slovenia, Jun. 2018. (20 pt)