

Friedrich-Alexander-Universität Erlangen-Nürnberg

**Lehrstuhl für Multimediakommunikation und
Signalverarbeitung**

Prof. Dr.-Ing. André Kaup

Forschungspraktikum

**Codierung medizinischer Volumen durch
H.265/HEVC mittels geeigneter
Kostenfunktionen in der
Raten-Verzerrungs-Optimierung**

Tayseer Al Sabbagh

Oktober 2018

Betreuerin: Karina Jaskolka

Erklärung

Ich versichere, dass ich die vorliegende Arbeit ohne fremde Hilfe und ohne Benutzung anderer als der angegebenen Quellen angefertigt habe, und dass die Arbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen hat und von dieser als Teil einer Prüfungsleistung angenommen wurde. Alle Ausführungen, die wörtlich oder sinngemäß übernommen wurden, sind als solche gekennzeichnet.

Ort, Datum

Unterschrift

Inhaltsverzeichnis

Abkürzungsverzeichnis	II
Formelzeichen	III
1 Einleitung	1
2 Theoretischer Teil	2
2.1 HEVC Coder im Überblick	2
2.2 Rate-Distortion Optimization (RDO)	3
2.3 Die Verzerrungsfunktionen	4
2.4 Bjøntegaard Delta-Bitrate und Quantizierungsparameter QP	7
3 Experimentelle Arbeit und Ergebnisse	8
4 Fazit	16
Abbildungsverzeichnis	16
Tabellenverzeichnis	17
Literaturverzeichnis	19

Abkürzungsverzeichnis

CTU	Coding Tree Units
CU	Coding Units
DST	Discrete Sine Transform
HEVC	High Efficiency Video Coding
NCC	Normalized Cross Correlation
PU	Prediction Units
RDO	Rate-Distortion Optimization
SAD	Sum of Absolute Difference
SSD	Sum of Squared Differences
ZNCC	Zero Mean Normalized Cross Correlation
ZSAD	Zero Mean Summ of Absolut Difference
ZSSD	Zero Mean Sum of Squared Difference

Formelzeichen

λ	Lagrange multiplier
D	Distortion
R	Bitrate
QP	Quantization Parameter
I	Licht Intensität des Pixels

Kapitel 1

Einleitung

Heutzutage werden Video Dateien komprimiert und durch Internet bzw. Datenübertragungsmittel weitergeleitet. Ein großes Problem im medizinischen Bereich ist derzeit die Übertragung von medizinischen Volumen komprimiert und möglichst fehlerfrei mit den ganzen detaillierten Informationen, die der Diagnose dienen können. In dieser Arbeit wird die Methode der Berechnung der Kosten des HEVC Coders mittels anderer Verzerrungsfunktionen durchgeführt, um das Raten-Verzerrungs-Verhältnis des Code durch die Änderung in der automatischen Auswahl von Partitionsmethoden verbessern zu können.

Kapitel 2

Theoretischer Teil

2.1 HEVC Coder im Überblick

Dieser Algorithmus wird in 4G Video, IPTV benutzt und hat die typische Bitrate (500 kbit/s -10 Mbit/s). Wie in Abbildung 2.1 gezeigt wurde, enthält der HEVC Algorithmus unterschiedliche Schritte[1], [2] :

- Transformation: Diskrete Sinustransformation (DST).
- Skalierung und Quantisierung.
- Sample Adaptive Offset Loop Filter (SAO).
- Inter Prädiktion mittels Bewegungsschätzung und Bewegungskompensation.
- Intra Prädiktion: Hier wird jedes Frame zu gleichgroßen Einheiten in Coding Tree Units (CTUs) geteilt. Dann wird jedes CTU durch Quadtree Methode zu Coding Units (CUs) geteilt. Danach wird jedes CUs zu Prediction Units (PUs) geteilt. Jedes PU ist von 4x4 bis zu 32x32 definiert (siehe Abbildung 2.2).
- Entropy Codierung: CABAC.

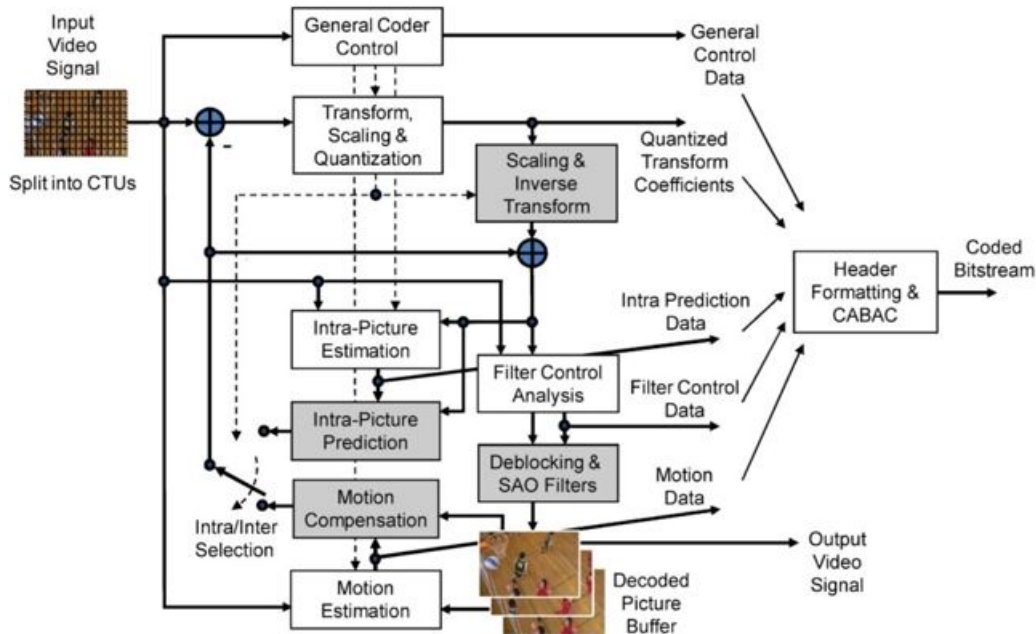


Abbildung 2.1: Ablauf des HEVC Coders [2]

2.2 Rate-Distortion Optimization (RDO)

HEVC benutzt zwei Arten von Prädiktion: Intra- und Inter-Prädiktion. In demselben Frame wird das Intra-Prädiktion benutzt. Um die (PUs) zu codieren, gibt es 35 unterschiedliche Prädiktion Mods. Die Größen der (PU) variieren im Bereich (4x4 bis 32x32). Um die beste Kombination aus Prädiktionsmodus passender PU-Größe finden zu können, wird die Rate-Distortion Optimization (RDO) auf jedes PU angewendet [3].

Intra-Prädiktion hat die Bedeutung, dass PU von den benachbarten PUs durch das Fehlersignal rekonstruiert wird, um die statische Abhängigkeit zwischen PUs zu verringern. Bei jedem Prädiktionsverfahren von jedem PU werden die Schritte in Abbildung 2.3 nachverfolgt. Am Ende des Prozesses ergibt sich das predicted PU. Die Auswertung des Verfahrens wird durch die Bitrate (R) und die Verzerrung (Distortion) (D) mit der

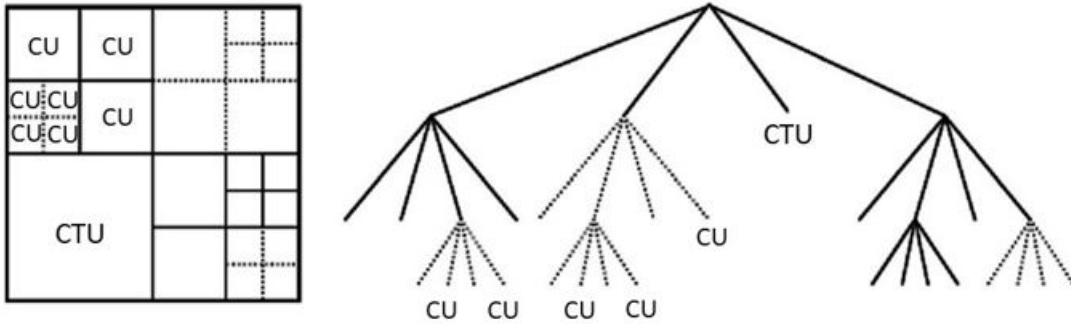


Abbildung 2.2: Frame durch Quadtree Method geteilt [1]

folgenden Gleichung berechnet:

$$J = D + \lambda R. \quad (2.1)$$

Lambda ist die Lagrange multiplier und ist vom Quantizierungsparameter (QP) und der Konstante C abhängig:

$$\lambda = C \cdot 2^{\frac{(QP-12)}{3}} \quad (2.2)$$

2.3 Die Verzerrungsfunktionen

Die Verzerrung bezieht sich auf den Fehler zwischen dem prädizierten PU und dem originalen. Die Verzerrung wird im aktuellen Zustand durch „the sum of squared differences (SSD)“ berechnet. Hier wird die Differenz zwischen den Pixeln quadriert.

$$SSD = \sum |I_{Original}(x, y) - I_{Predicted}(x, y)|^2 \quad (2.3)$$

Das prädizierten PU ist von dem prädiktionsmodus abhängig und es wird aus dem benachbarten decodierten PU berechnet [3]. Diese SSD Funktion ergibt den Fehler von den ganzen Pixeln im Block. Jedoch gibt es noch viele weitere Kostenfunktionen, die in Betracht gezogen werden sollen. Das könnte eine große Änderung in den Kosten des HEVC in der Codierung von medizinischen Dateien verursachen.[3] Deshalb werden in dieser Arbeit neue Kostenfunktionen benutzt, um die Verzerrung berechnen zu können.

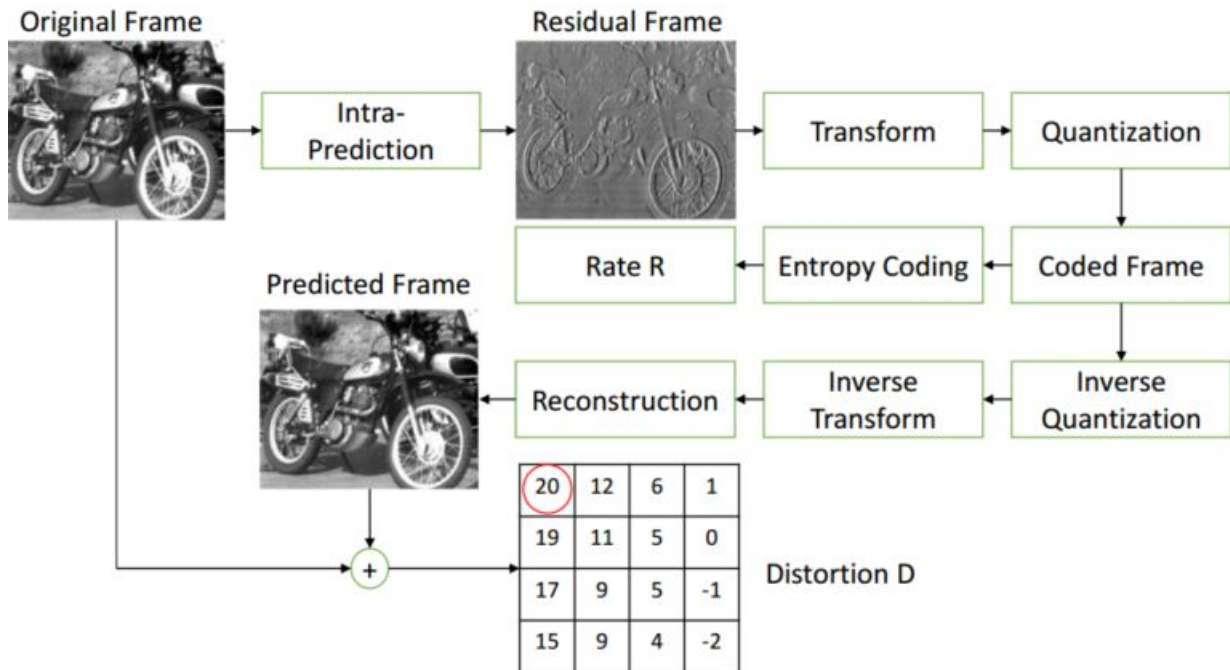


Abbildung 2.3: Diagramm der RDO in HEVC für Intra Prädiktion [3]

1. Sum of Absolute Difference (SAD):

Hier wird der Unterschied zwischen Pixel von den originalen und prädizierten PUs berechnet und addiert. Diese Methode ist wenig Zeitaufwendig, einfach zu implementieren und sie ist ideal für das „Real Time System“. Aber diese Methode wird von den Ausreißenden Werte beeinflusst.

$$SAD = \sum |I_{Original}(x, y) - I_{Predicted}(x, y)| \quad (2.4)$$

Bei SAD und SSD ergeben zwei PUs mit denselben Pixel-Werten die Verzerrung Null. Dieses Problem wurde durch ZSSD und ZSAD gelöst.

2. Zero Mean Sum of Absolut Difference (ZSAD):

Hier wird SAD berechnet, aber der Mittelwert des PU wird von jedem Pixel subtrahiert. Das gibt dieser Methode die Fähigkeit, die lineare Intensitätsschwan-

kungen zu verhindern.

$$ZSAD = \sum |I_{Original}(x, y) - M_{Original} - I_{Predicted}(x, y) + M_{Predicted}| \quad (2.5)$$

3. Zero Mean Sum of Squared Difference (ZSSD):

Hier wird SSD berechnet, aber der Mittelwert des PU wird von jedem Pixel subtrahiert und mitquadrirt. Das wird zu der Rechenkomplexität eingefügt.

$$ZSSD = \sum (I_{Original}(x, y) - M_{Original} - I_{Predicted}(x, y) - M_{Predicted})^2 \quad (2.6)$$

4. Normalized Cross Correlation (NCC):

Hier wird die Summe von der Multiplikation zwischen dem Pixel in dem originalen PU und dem Pixel in dem predictierten PU durch die Standardabweichung zwischen den originalen und predictierten PU geteilt. Diese Methode besitzt eine höhere Rechenkomplexität. Aber sie ist stabil gegen lineare Variationen in unterschiedlichen Beleuchtungsvoraussetzungen.

$$NCC = \frac{\sum \{I_{Original}(x, y) \cdot I_{Predicted}(x, y)\}}{\sqrt{\sum I_{Original}^2(x, y) \cdot \sum I_{Predicted}^2(x, y)}} \quad (2.7)$$

5. Zero Mean Normalized Cross Correlation (ZNCC):

Hier wird das NCC berechnet mit dem Unterschied, dass der Mittelwert von der Intensität jedes Pixels subtrahiert wird. Diese Methode ist stabil gegen die Intensitätsverzerrungen, aber besitzt höhere Rechenkomplexität als NCC.

$$ZNCC = \frac{\sum \{[I_{Original}(x, y) - M_{Original}] \cdot [I_{Predicted}(x, y) - M_{Predicted}]\}}{\sqrt{\sum [I_{Original}(x, y) - M_{Original}]^2 \cdot \sum [I_{Predicted}(x, y) - M_{Predicted}]^2}} \quad (2.8)$$

NCC und ZNCC ergeben normalerweise gute Ergebnisse, wenn die Größe des PU klein ist[4].

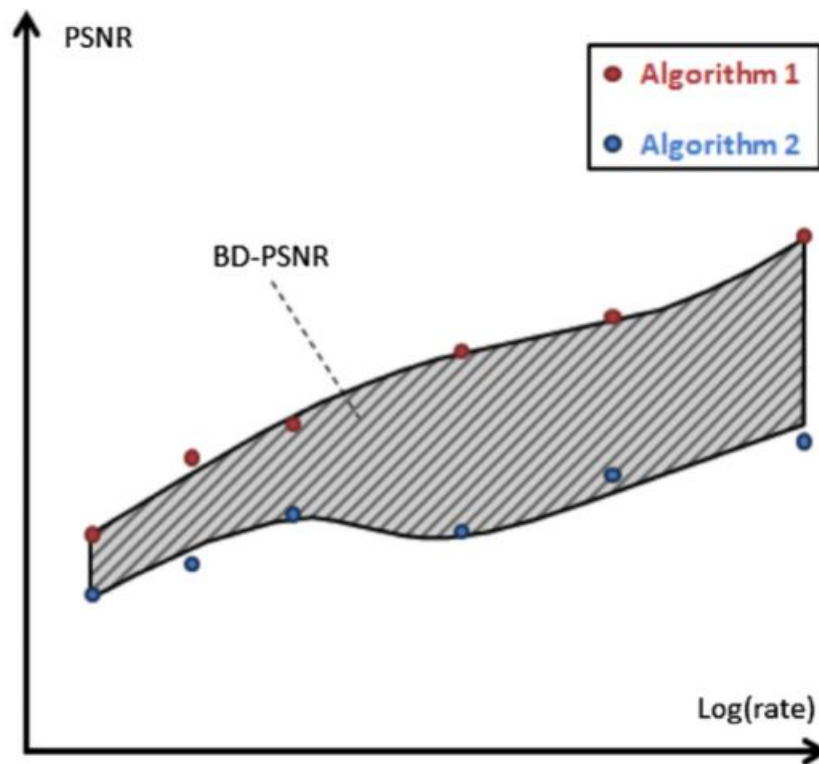


Abbildung 2.4: Beispiel von Bjøntegaard delta

2.4 Bjøntegaard Delta-Bitrate und Quantizierungsparameter QP

Die Bjøntegaard Methode wird oft benutzt, um die Effizienz der Decodierung zwischen unterschiedlichen Algorithmen herausfinden zu können. Das Bjøntegaard Delta berechnet die Differenz zwischen zwei Raten-Verzerrungs-Kurven. Dabei wird die Fläche zwischen den Kurven durch das Integrationsintervall geteilt [5], siehe Abbildung 2.4.

Das Bjøntegaard Delta Bitrate Rechenmethode ergibt negative Werte, die sich darauf beziehen, wie viel sich die Bitrate reduziert hat. Positive Werte beziehen sich auf die Zunahme von der Bitrate für denselben PSNR.

Das Encoding in HEVC wird durch QP kontrolliert. QP ist der Index, um die „Scaling Matrix“ abzuleiten. QP bestimmt den passenden Quantisierungsschritt. Je größer der QP ist, desto größer ist der Quantisierungsschritt.

Kapitel 3

Experimentelle Arbeit und Ergebnisse

Hier wurde die HM test software version 16.9 modifiziert, als eine Simulation der Arbeit des HEVC Coders. Die Testdatei sind 3D medizinische Volume mit 8 bpp (CT Aufnahme eines Schädels). Die unterschiedliche Rechenmethoden der Verzerrung wurden erfolgreich in dem File „TComRdCost.cpp“ implementiert. Unterschiedliche Größen der PUs im Bereich zwischen (4x4 Pixel) bis zu (64x64 Pixel) wurden in Betracht gezogen. Um die Rate und die Qualität zu kontrollieren, wird die folgende QPs im Konfiguration File „intra_main_rext.cfg“ ausgewählt (12, 17, 22, 27).

Hier wurde das medizinische Volumen durch HM 16.9 unter unterschiedlichen RDO Rechenmethoden codiert. Die Ergebnisse ist die resultierende Bitrate und die gesamte PSNR zwischen dem eingegebenen Volumen und dem resultierenden Volumen. Abbildung 3.1 zeigt die Ergebnisse

Um die Ergebnisse zwischen den unterschiedlichen Verzerrungsfunktionen zu evaluieren, wird das erklärte Bjøntegaard-Delta berechnet. Die Bitrate- und die PSNR-Ergebnisse wurden für jedes QP in einem Array angelegt. Dann werden diese Arrays der Funktion „Bjøntegaard“ in Matlab mit dem Mode „Rate“ eingegeben. Die Funktion „Bjøntegaard“ evaluiert die Verzerrungsfunktionen durch den Vergleich mit „SSD“,

	QP	SSE	SAD	NCC	HADS
Bitrate	12	2726,4788180	2655,9940890	2664,0157640	2802,7605910
	17	1785,2877000	1675,7448280	1683,0177340	1723,7064040
	22	998,1724140	859,1290640	861,3802960	873,6581280
	27	572,3812810	520,5507390	522,8965520	523,2423650
PSNR	12	52,5178100	52,0996940	51,5066650	52,6419990
	17	47,9493000	47,1867090	46,7583740	47,5663840
	22	43,8279980	42,7465350	42,5530410	42,9499740
	27	40,9623330	40,0170280	39,8596670	40,1700620
	QP	ZSSE	ZSAD	ZNCC	
Bitrate	12	2748,9546800	2743,9399010	2586,9990150	
	17	1816,1113300	1747,8029560	1650,1921180	
	22	1026,9694580	903,4344830	853,3970440	
	27	591,0088670	532,6916260	520,7684730	
PSNR	12	52,4189680	52,0644640	51,3845870	
	17	47,8940130	47,4021000	46,6022200	
	22	43,7793290	42,9717410	42,5443520	
	27	40,8875890	40,1527020	39,8791280	

Abbildung 3.1: Zahlenergebnisse (Bitrate und PSNR) resultierend aus dem Test

die im aktuellen Zustand benutzt wird. Die Bjøntegaard-Rate wird in der Tabelle 3.1 gezeigt. Je größer die Rate ist, desto schlechter ist das PSNR für dieselbe Bitrate. In den Abbildungen 3.2 bis 3.7 sind die Raten-Verzerrungs-Kurven abgebildet.

Verzerrungsfunktionen	SSE	SAD	NCC	HADS	ZSSE	ZSAD	ZNCC
Bjøntegaard Rate [%]	reference	3,3050	8,7571	1,8628	2,8933	4,9931	8,2880

Tabelle 3.1: Bjøntegaard-Rate

In den Abbildungen und Tabellen kann man sehen, dass keine der implementierten Kostenfunktionen eine Verbesserung in der Raten-Verzerrung-Kurve bringt. Am schlechtesten hat NCC mit 8,75% abgeschlossen.

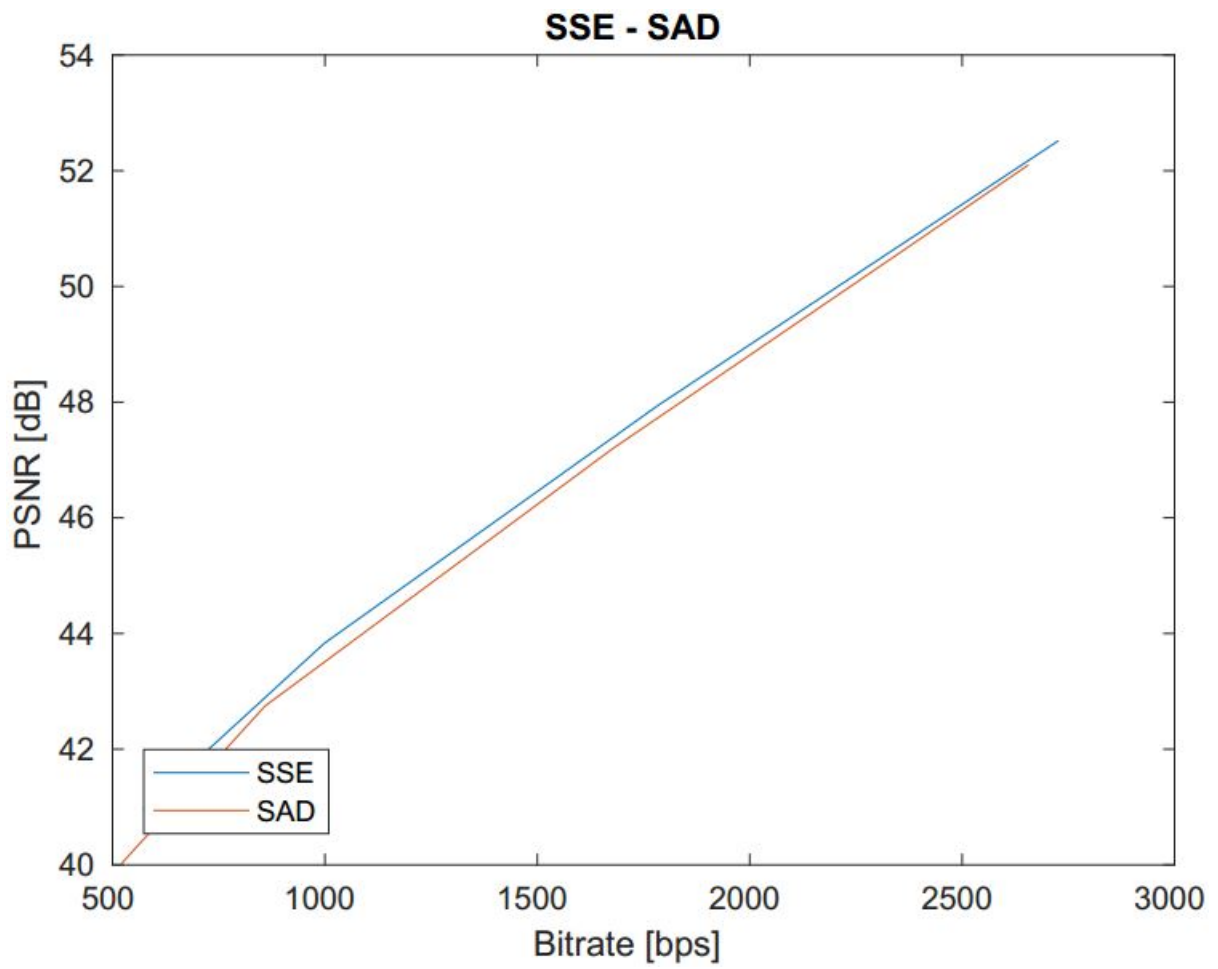


Abbildung 3.2: SSE-SAD

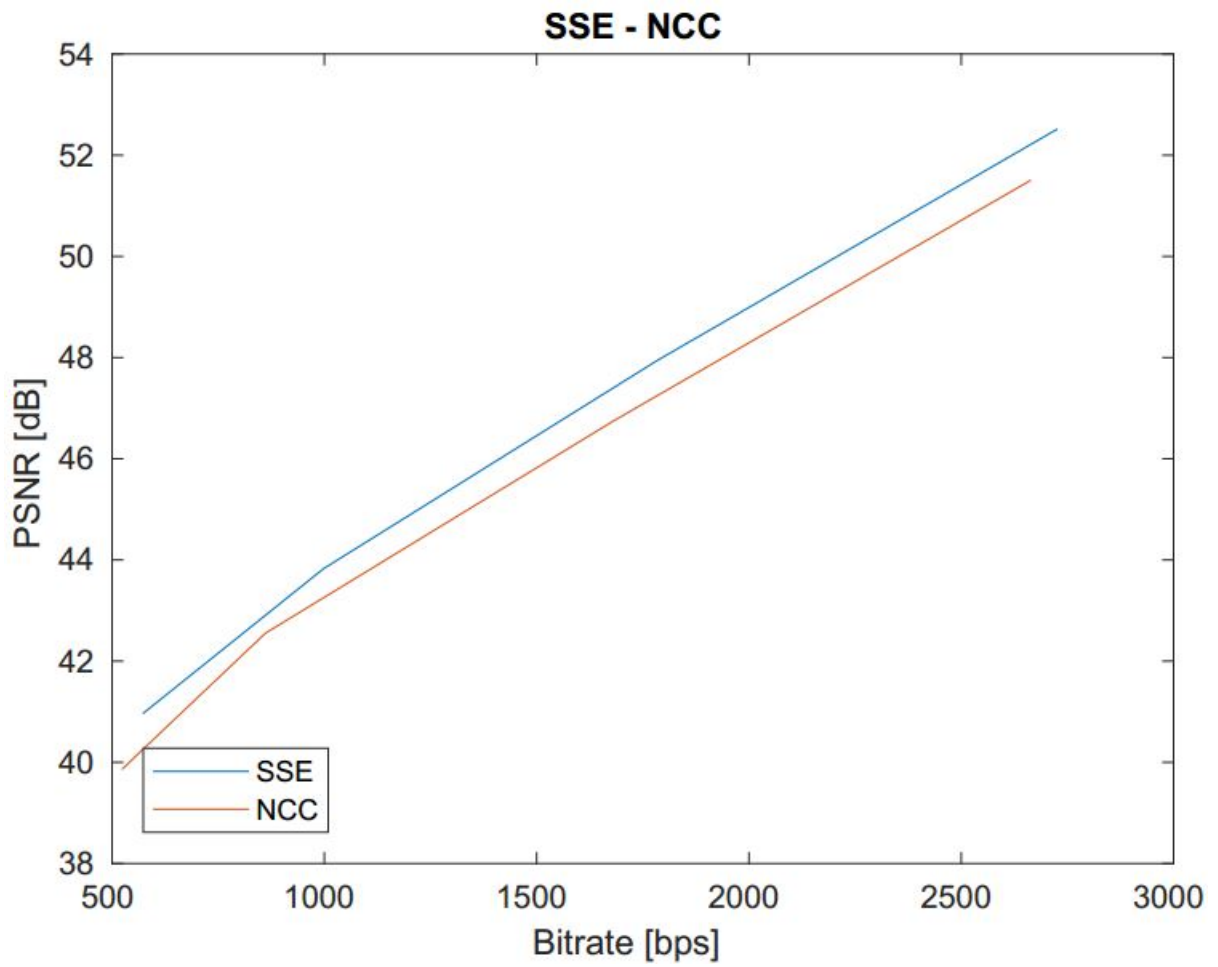


Abbildung 3.3: SSE-NCC

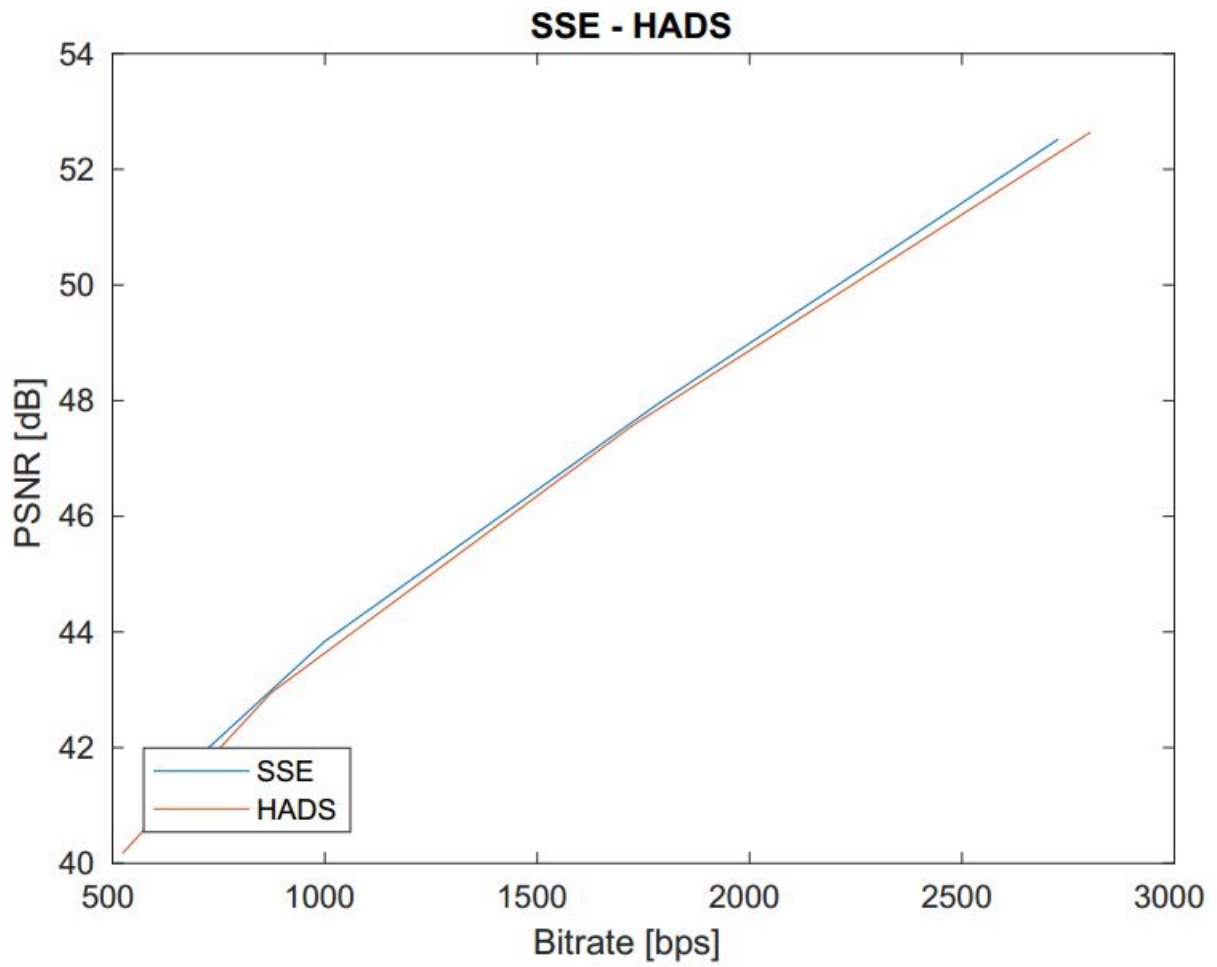


Abbildung 3.4: SSE-HADS

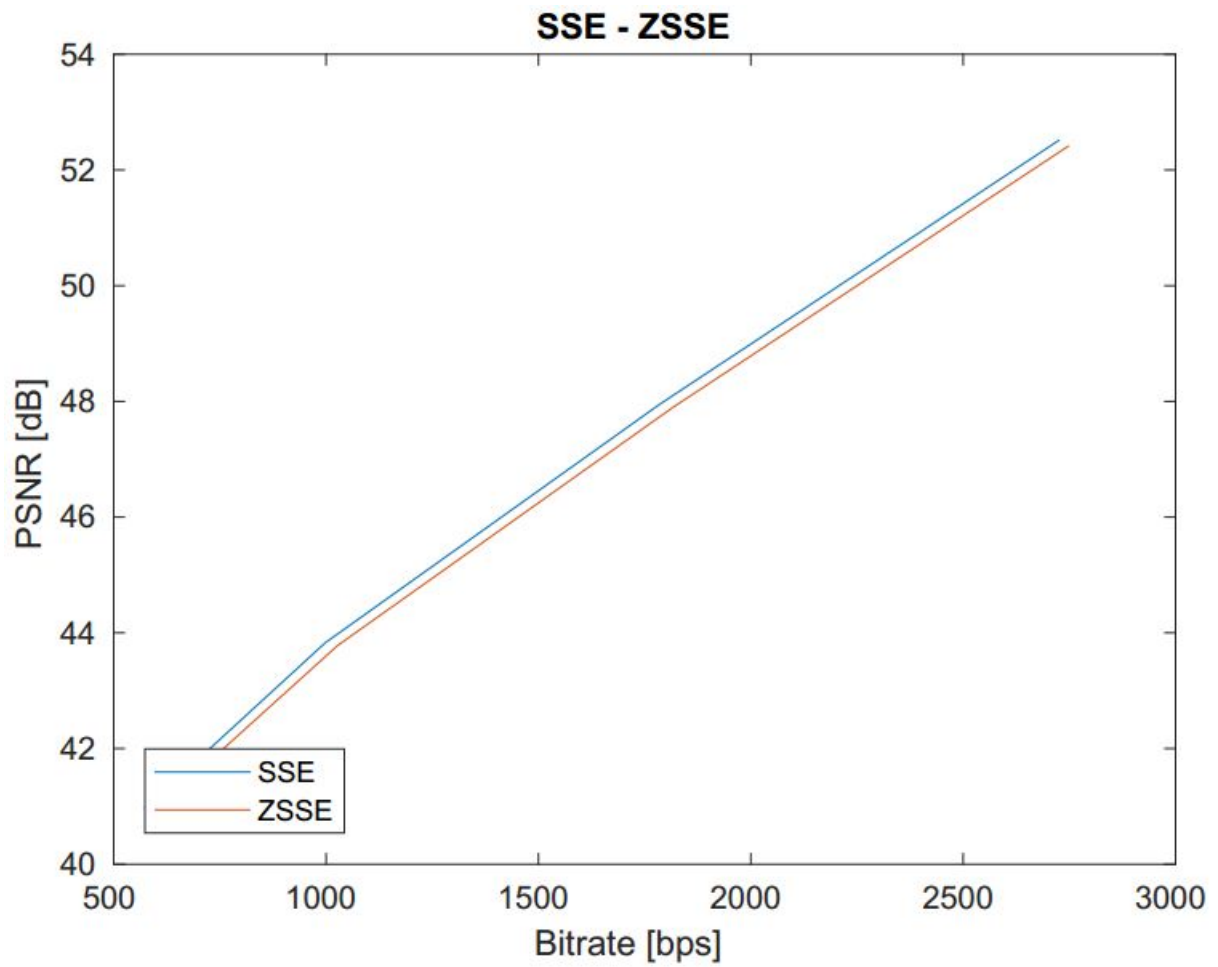


Abbildung 3.5: SSE-ZSSE

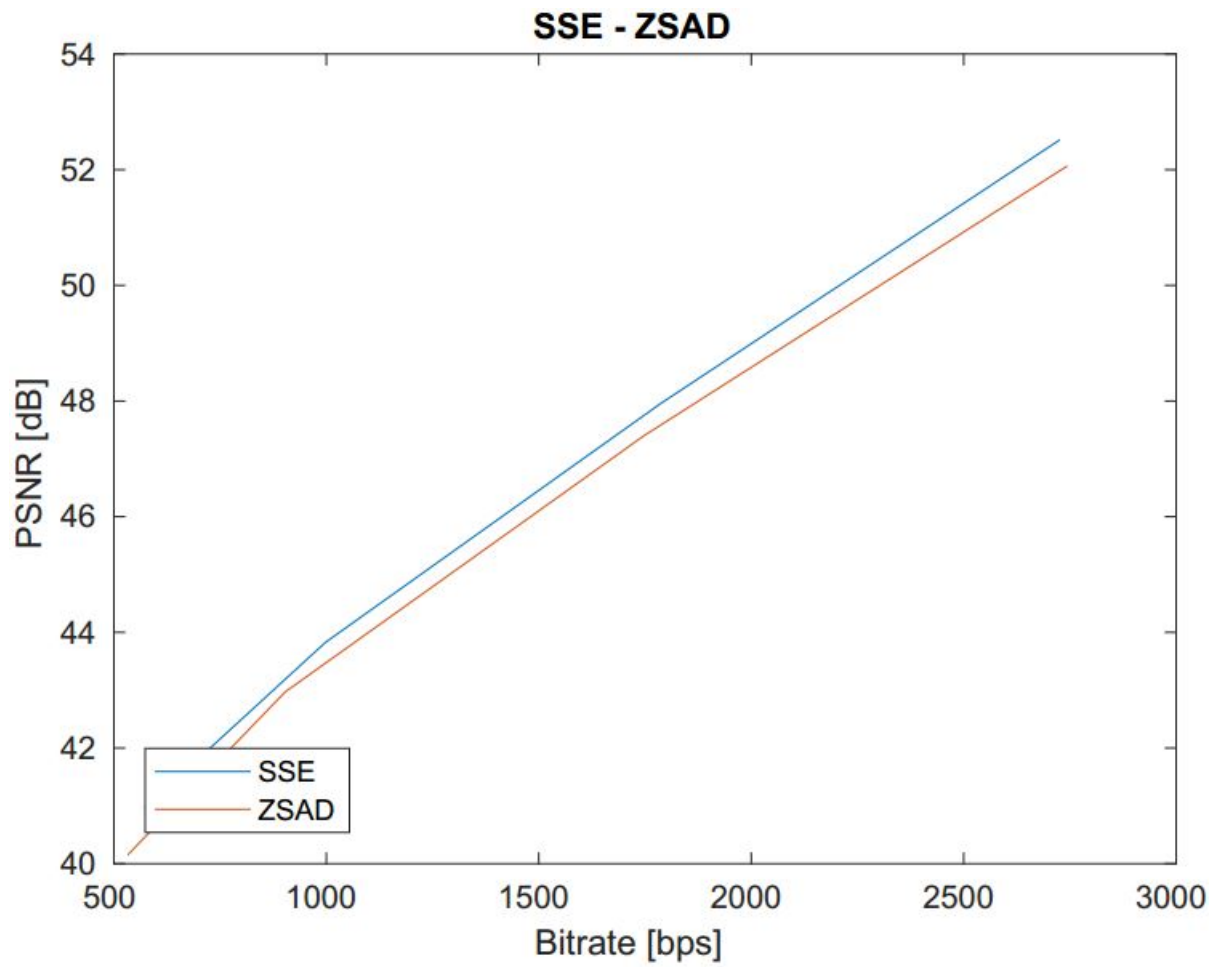


Abbildung 3.6: SSE-ZSAD

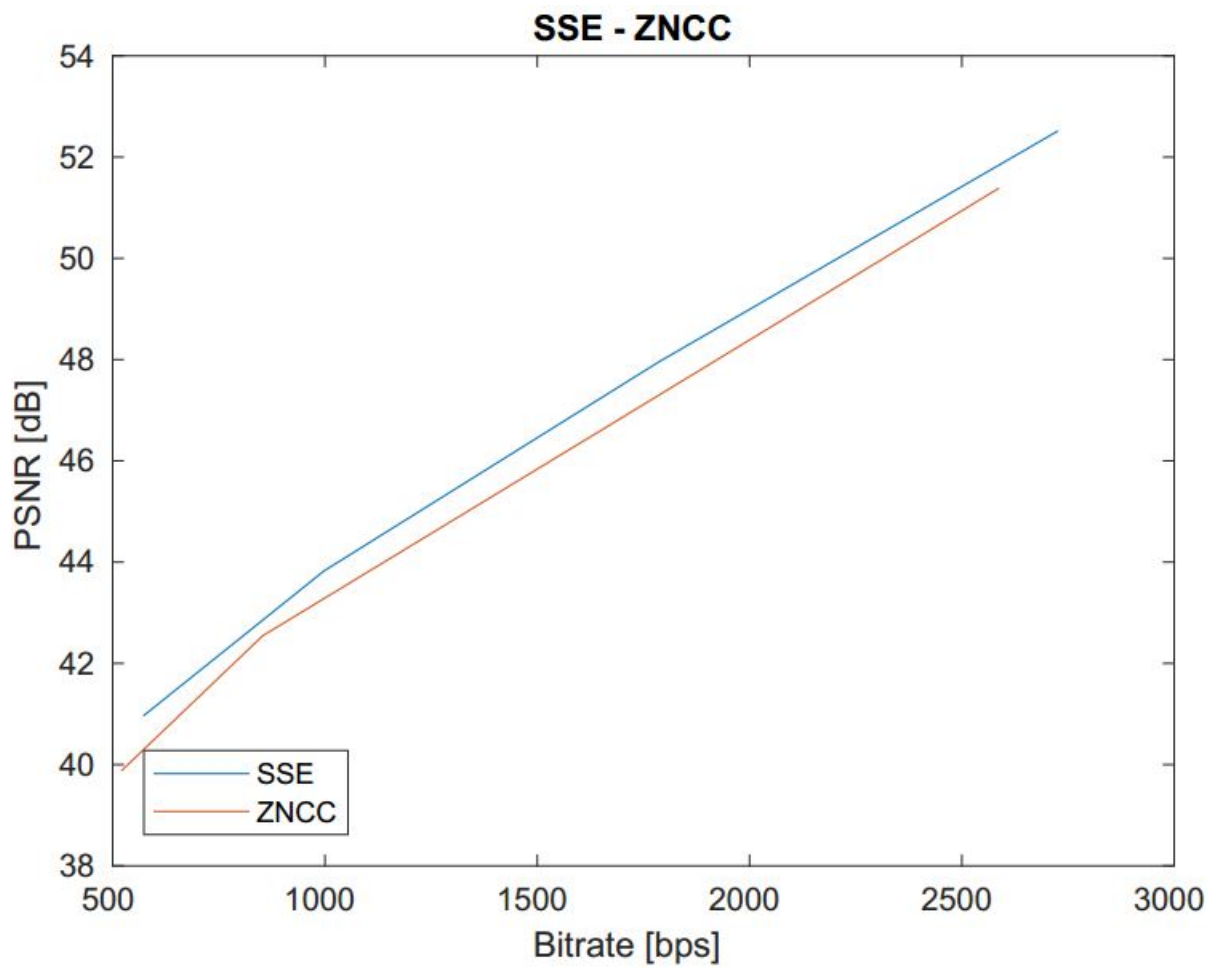


Abbildung 3.7: SSE-ZNCC

Kapitel 4

Fazit

In dieser Arbeit wurden die Rechenmethoden von den Kosten der Codierung in HEVC durch die Einfügung neuer Verzerrungsfunktionen modifiziert. Die Verzerrungsfunktionen enthalten alle PU-Größen, was dem Coder erlaubt, die passende PU-Größe zu finden. Der Vergleich zwischen den unterschiedlichen Algorithmen wurde mit Hilfe von Bjøntegaard Delta-Bitrate geschafft. Es wurde festgestellt, dass die Verzerrungsfunktionen ZSSE, SAD, ZSAD, HADS, NCC und ZNCC keine Verbesserung in der PSNR für dieselbe Bitrate im Vergleich zu SSE zeigen.

Abbildungsverzeichnis

2.1	Ablauf des HEVC Coders [2]	3
2.2	Frame durch Quadtree Method geteilt [1]	4
2.3	Diagramm der RDO in HEVC für Intra Prädiktion [3]	5
2.4	Beispiel von Bjøntegaard delta	7
3.1	Zahlenergebnisse (Bitrate und PSNR) resultierend aus dem Test	9
3.2	SSE-SAD	10
3.3	SSE-NCC	11
3.4	SSE-HADS	12
3.5	SSE-ZSSE	13
3.6	SSE-ZSAD	14
3.7	SSE-ZNCC	15

Tabellenverzeichnis

3.1	Bjøntegaard-Rate	9
-----	----------------------------	---

Literaturverzeichnis

- [1] A.KAUP: Image and Video Compression. (2018). – Friedrich-Alexander-Universität Erlangen-Nürnberg
- [2] SULLIVAN, G. J. ; OHM, J. ; HAN, W. ; WIEGAND, T.: Overview of the High Efficiency Video Coding (HEVC) Standard. In: *IEEE Transactions on Circuits and Systems for Video Technology* 22 (2012), Dez, Nr. 12, S. 1649–1668
- [3] JASKOLKA, K. ; KAUP, A.: Joint Optimization of Rate, Distortion, and Maximum Absolute Error for Compression of Medical Volumes Using HEVC Intra. In: *Picture Coding Symposium (PCS)*, 2018, S. 126–130
- [4] PATIL, S. ; NADAR, J. ; GADA, J. ; MOTGHARE, S. ; NAIR, S.: Comparison of Various Stereo Vision Cost Aggregation Methods. In: *International Journal of Engineering and Innovative Technology (IJEIT)* 2 (2013), Nr. 8
- [5] BLINDER, D. ; BRUYLANTS, T. ; OTTEVAERE, H. ; MUNTEANU, A. ; SCHELKENS, P.: JPEG 2000-based compression of fringe patterns for digital holographic microscopy. In: *Optical Engineering* 53 (2014)