
MASTERARBEIT

Herr B.Eng.
Felix Oberhardt

**Subjektive Beurteilung
höherer Bildwiederhol-
frequenzen im Fernsehen
mit Auswirkungen gängiger
Codieralgorithmen**

Mittweida, 2015

MASTERARBEIT

Subjektive Beurteilung höherer Bildwiederhol- frequenzen im Fernsehen mit Auswirkungen gängiger Codieralgorithmen

Autor:

**Herr B.Eng.
Felix Oberhardt**

Studiengang:

Industrial Management

Seminargruppe:

ZM13w2-M

Erstprüfer:

Herr Prof. Dipl.-Ing. Hans-Joachim Götz

Zweitprüfer:

Frau Dipl.-Ing. (FH) Dagmar Driesnack

Einreichung:

München, 30.10.2015

Verteidigung/Bewertung:

München, 2015

MASTER THESIS

Subjective Evaluation of High Frame Rate Television with Impacts of Common Coding Algorithms

author:

**Mr. B.Eng.
Felix Oberhardt**

course of studies:

Industrial Management

seminar group:

ZM13w2-M

first examiner:

Prof. Dipl.-Ing. Hans-Joachim Götz

second examiner:

Dipl.-Ing. (FH) Dagmar Driesnack

submission:

Munich, 2015-10-30

defence/ evaluation:

Munich, 2015

Bibliografische Beschreibung:

Oberhardt, Felix:

Subjektive Beurteilung höherer Bildwiederholfräquenzen im Fernsehen mit Auswirkungen gängiger Codieralgorithmen. – 2015. – XIII, 59, V S.

Mittweida, Hochschule Mittweida, Institut für Technologie- und Wissenstransfer, Masterarbeit, 2015

Referat:

Die Arbeit beschäftigt sich mit höheren Bildwiederholfräquenzen als einem von mehreren Systemfaktoren innerhalb des UHD-TV-Standards und ihrem Mehrwert gegenüber dem aktuell geltenden und weitläufig praktizierten HDTV-Standard. Es werden die Vor- und Nachteile verschiedener Methoden zur Erzeugung höherer Bildwiederholfräquenzen evaluiert und insbesondere auf die Art und Weise der Quellencodierung als Grundlage für eine Verbreitung zum Fernsehzuschauer eingegangen.

Vorwort

Die vorliegende Arbeit entstand am *Institut für Rundfunktechnik (IRT)* und alle Inhalte beruhen demzufolge auf Recherchen und Tests, die am *IRT* durchgeführt wurden. Das *IRT* ist die Forschungs- und Entwicklungseinrichtung der öffentlich-rechtlichen Rundfunkanstalten in Deutschland, Österreich und der Schweiz. Es ist maßgeblich an der Entwicklung und Beurteilung neuer Technologien im Rundfunk beteiligt und veröffentlicht regelmäßig aktualisierte technische Richtlinien für die öffentlich-rechtlichen Rundfunkanstalten. Mit dieser Arbeit soll die Wahl einer geeigneten Bildwechselfrequenz als Teil des aktuell gültigen UHD-Standards vereinfacht und auf mögliche Probleme hingewiesen werden. Dazu dienen die zum Zeitpunkt des Verfassens der Arbeit am IRT vorhandenen technischen Hilfsmittel, mit denen die Arbeitsschritte innerhalb der Produktions- und Ausstrahlungsketten in den deutschen Rundfunkanstalten modellhaft nachgestellt werden können. Ergebnisse und Handlungsempfehlungen werden beispielhaft auf einen möglichen zukünftigen UHDTV-Regelbetrieb abgebildet.

Inhaltsverzeichnis

Inhaltsverzeichnis	I
Abbildungsverzeichnis	IV
Tabellenverzeichnis	VI
Abkürzungsverzeichnis	VII
Formelverzeichnis	IX
1 Einleitung	1
1.1 <i>Fragestellungen</i>	1
1.2 <i>Vorgehensweise</i>	2
2 Produktion von HFR	3
2.1 <i>Definition von Bewegungsartefakten</i>	3
2.1.1 <i>Flimmern</i>	3
2.1.2 <i>Stroboskopeffekt</i>	4
2.1.3 <i>Bewegungsunschärfe</i>	5
2.2 <i>Aufnahme und Wiedergabe</i>	5
2.2.1 <i>Verfügbarkeit von Kameras und Displays</i>	5
2.2.2 <i>UHDTV ohne Zeilensprungverfahren</i>	6
2.2.3 <i>Veränderbare Parameter</i>	7
2.3 <i>Codierung</i>	10
2.3.1 <i>Codec</i>	10
2.3.2 <i>Relevante Codingparameter</i>	11
3 Gegenwärtiger Forschungsstand	13
3.1 <i>Nachrichtentechnische Gegebenheiten</i>	13
3.2 <i>Physiologische Gegebenheiten</i>	13
3.3 <i>Subjektive Wahrnehmung</i>	15
3.4 <i>Testaufbau zur Geschwindigkeit bewegter Objekte</i>	18
4 Testumgebung	20
4.1 <i>Testsequenzen</i>	20

4.2	<i>Aufbereitung und Wiedergabe</i>	24
4.2.1	Downsampling.....	24
4.2.2	Interpolation	26
4.2.3	Videoserver	27
4.2.4	Displays.....	28
4.3	<i>Bewertungsmethoden</i>	29
4.3.1	DSCQS	30
4.3.2	Expert Viewing	31
5	Tests zur zeitlichen Auflösung	33
5.1	<i>Auswirkungen unterschiedlicher Shuttereinstellungen</i>	35
5.1.1	Durchschnittliche Bewertung über alle Sequenzen.....	36
5.1.2	50 Hz bei 25 %, 50 % und 100 % Shutter.....	37
5.1.3	100 Hz bei 25 %, 50 % und 100 % Shutter.....	38
5.1.4	200 Hz bei 50 % und 100 % Shutter	39
5.2	<i>Auswirkungen einer Zwischenbildberechnung</i>	39
5.2.1	Durchschnittliche Bewertung über alle Sequenzen.....	40
5.2.2	Interpolation von 50 Hz zu 100 Hz in Abhängigkeit der Shuttereinstellung	41
5.2.3	Interpolation von 50 Hz zu 200 Hz	42
5.2.4	Interpolation von 100 Hz zu 200 Hz in Abhängigkeit der Shuttereinstellung ..	43
6	Tests zur Codierung	44
6.1	<i>Vorbetrachtung und Testvoraussetzungen</i>	44
6.1.1	Praxisbezug	44
6.1.2	Werkzeuge für Encoding und Decoding.....	44
6.2	<i>Testdurchführung</i>	45
6.2.1	Encoding	45
6.2.2	PSNR-Berechnung.....	47
6.2.3	Expert Viewings zur subjektiven Beurteilung	48
7	Ergebnisse	54
7.1	<i>Tests zur Bildwiederholfrequenz</i>	54
7.2	<i>Tests zur Codierung</i>	55
7.3	<i>Weitere Ergebnisse aus EBU-Tests</i>	56
7.4	<i>Empfehlungen für Produktionen in HFR</i>	57
7.4.1	Sport- und Showformate	57
7.4.2	Dokumentationen	57
7.4.3	Szenische Produktionen	58

7.5	<i>Ansätze zur weiterführenden Forschung</i>	58
7.5.1	Auswahl des Testmaterials	58
7.5.2	Prozessoptimierung bei der Aufbereitung des Testmaterials.....	58
7.5.3	Temporal Scalability Coding	59
7.5.4	Wiederholung der Tests mit UHD-Displays.....	59
7.5.5	Implementierung zusätzlicher Bildoptimierungsoptionen	59
	Literaturverzeichnis	XI
	Anhang	
	Selbstständigkeitserklärung	

Abbildungsverzeichnis

Abbildung 1: Unbewegter Siemensstern	8
Abbildung 2: Rotierender Siemensstern bei verschiedenen Shutterwerten (1/50, 1/200, 1/1000 Sekunde) und Aufnahme in 1080p/25	8
Abbildung 3: Rotierender Siemensstern bei verschiedenen Rotationsgeschwindigkeiten (0,15, 0,375, 0,94 U/s) und Aufnahme in 1080p/25	9
Abbildung 4: Rotierender Siemensstern bei verschiedenen Rotationsgeschwindigkeiten (0,15, 0,375, 0,94 U/s) und Aufnahme in 720p/120	9
Abbildung 5: Zeitlich skalierbare Codierung mit HEVC	12
Abbildung 6: Ergebnisse der subjektiven Tests der BTF-HFR	16
Abbildung 7: Testergebnisse der Ericsson HFR-Kontributionstests mit dem HEVC Range Extension Profile Main 4:2:2 10	17
Abbildung 8: Screenshots der verwendeten Testsequenzen. 1 Hürdenlauf, 2 Stufenbarren, 3 Dreisprung, 4 Reitsport, 5 5000m-Lauf	23
Abbildung 9: SI und TI der ausgewählten Testsequenzen	24
Abbildung 10: Downsampling durch Image Averaging	25
Abbildung 11: Downsampling durch Frame Dropping	25
Abbildung 12: Messreihe zur Leuchtdichte der HFR-Displays in Abhängigkeit der LED-Backlight-Einstellung	29
Abbildung 13: Bewertungsskala für DSCQS	31
Abbildung 14: Räumliche Gegebenheiten beim subjektiven Test am IRT	34
Abbildung 15: Downsampling anhand der verwendeten Testsequenzen	35
Abbildung 16: Durchschnitt aller fünf Sequenzen ohne Interpolation	37

Abbildung 17: Durchschnitt je Sequenz ohne Interpolation.....	38
Abbildung 18: Interpolation anhand der verwendeten Testsequenzen	40
Abbildung 19: Durchschnitt aller fünf Sequenzen mit Interpolation.....	41
Abbildung 20: Durchschnitt je Sequenz mit Interpolation.....	42
Abbildung 21: Interpolationsartefakte.....	43
Abbildung 22: Festgelegte Codierdatenraten für alle Sequenzen.....	46
Abbildung 23: Durchschnitt aller fünf Sequenzen ohne Interpolation.....	47
Abbildung 24: Durchschnitt aller fünf Sequenzen mit Interpolation.....	48
Abbildung 25: Grafische Darstellung des Vergleichs (a).....	49
Abbildung 26: Grafische Darstellung des Vergleichs (b).....	50
Abbildung 27: Grafische Darstellung des Vergleichs (c).....	51
Abbildung 28: Grafische Darstellung des Vergleichs (d).....	51
Abbildung 29: Grafische Darstellung des Vergleichs (e).....	53

Tabellenverzeichnis

Tabelle 1: Kameramodelle für HFR-Produktion	6
Tabelle 2: Objektgeschwindigkeiten im Test	19
Tabelle 3: Technische Daten der verwendeten Testsequenzen	22
Tabelle 4: Protokoll des Vergleichs (a).....	49
Tabelle 5: Protokoll des Vergleichs (b).....	50
Tabelle 6: Protokoll des Vergleichs (c).....	51
Tabelle 7: Protokoll des Vergleichs (d).....	52
Tabelle 8: Protokoll des Vergleichs (e).....	53

Abkürzungsverzeichnis

4EVER	For Enhanced Video Experience
AVI	Audio Video Interleave
BBC	British Broadcasting Corporation
BTF	Broadcast Technology Future Group
cd	Candela
CBR	Constant Bit Rate
CFF	Critical flicker frequency
CRT	Cathode Ray Tube
CTB	Coding Tree Block
DPX	Digital Picture Exchange
DSCQS	Double-Stimulus Continuous Quality-Scale
DVB	Digital Video Broadcasting
EBU	European Broadcasting Union
EOTF	Electro-Optical Transfer Function
GOP	Group of Pictures
HDR	High Dynamic Range
HDTV	High Definition Television
HEVC	High Efficiency Video Coding
HFR	High Frame Rates
HM	HEVC Test Model
HMI	Hydrargyrum Medium-Arc Iodide
IEC	International Electrotechnical Commission
IRT	Institut für Rundfunktechnik
ISO	International Organization for Standardization
ITU	International Telecommunication Union
LCD	Liquid Crystal Display
LED	Light Emitting Diode
MCFI	Motion-Compensated Frame Interpolation
MPEG	Moving Picture Experts Group
MSE	Mean Square Error
NHK	Nippon Hōsō Kyōkai
PAL	Phase Alternating Line

PSNR	Peak Signal-to-Noise Ratio
QP	Quantization Point
SAMVIQ	Subjective Assessment of Multimedia Video Quality
SDI	Serial Digital Interface
SDR	Standard Dynamic Range
SI	Spatial Perceptual Information
SMPTE	Society of Motion Picture and Television Engineers
SRG SSR	Schweizerische Radio- und Fernsehgesellschaft
StatMux	Statistischer Multiplex
TI	Temporal Perceptual Information
U	Umdrehungen
UHDTV	Ultra High Definition Television
WCG	Wide Color Gamut

Formelverzeichnis

[1] Grenzgeschwindigkeit
$$v_{grenz} = 2 \cdot \text{Bildbreite} \cdot \frac{f_{Bild}}{N \cdot A}$$

A Belichtungszeit als Faktor der maximal möglichen Belichtungsdauer

f_{Bild} Bildwechselfrequenz in Bilder pro Sekunde

N Anzahl der horizontalen Pixel

v_{grenz} entspricht maximaler Schwenkgeschwindigkeit in [Bildbreite/Sekunde]

[2] Nyquist-Shannonsches Abtasttheorem
$$f_s > 2 \cdot f_{max}$$

f_s Abtastfrequenz

f_{max} maximal vorkommende Frequenz des Nutzsignals

[3] Spitzen-Signal-Rausch-Verhältnis
$$PSNR = 10 \cdot \lg \left(\frac{I_{max}^2}{MSE} \right) \text{ [dB]}$$

I_{max} max. Signalauslenkung; abhängig von Quantisierung und Signalbereich

MSE Mean Square Error; quadratischer Mittelwert des Fehlers aus codiertem Bild und Referenz

[4] Kreisbogen
$$b = \pi \cdot r \cdot \frac{\alpha}{180^\circ}$$

r Radius; entspricht dem Betrachtungsabstand in [Pixel]

[5] Objektgeschwindigkeit
$$v = \frac{b \cdot 180^\circ}{\pi \cdot r \cdot t} \text{ [°/s]}$$

b Kreisbogen; entspricht der zurückgelegten Bildschirmbreite in [Pixel]

t Zeit in [s]

1 Einleitung

Diverse internationale Standardisierungsgremien fokussieren sich auf die Verbesserung des Fernseherlebnisses. Aus diesem Grund werden viele innovative Technologien getestet und mehr als 20 Jahre nach der Entwicklung des gegenwärtigen Fernsehstandards wurde mittlerweile ein neuer Standard eingeführt, der unter der Bezeichnung Ultra High Definition Television (UHDTV) bekannt ist. Die Recommendation ITU-R BT.2020 der *International Telecommunication Union* beinhaltet eine Reihe von Verbesserungen gegenüber dem gegenwärtigen hochauflösenden Fernsehsystem (HDTV). In erster Linie werden höhere örtliche Bildauflösungen definiert, sodass nicht mehr 720 oder 1080 Zeilen pro Bild genutzt werden, wie der momentane Standard nach Rec. ITU-R BT.709 vorschreibt, sondern 2160 oder 4320 Zeilen, gleichbedeutend mit vier oder sechzehn Mal so vielen Bildpunkten wie bei HDTV. Neben der erhöhten Pixelanzahl werden noch einige weitere Parameter neu definiert. Neue Koordinaten für die Farbwerte von Rot, Grün und Blau führen zu einem erweiterten Farbraum (wide color gamut – WCG) und somit können sehr viel mehr Farben repräsentiert werden. Die Quantisierung des Fernsehsignals von derzeit 8 Bit in der Ausstrahlung soll künftig mit 10 oder 12 Bit erfolgen, und da somit mehr Quantisierungsstufen verfügbar sind, kann der Dynamikumfang des Signals durch Verwendung alternativer Übertragungskennlinien sowohl in den Kameras als auch in den Displays vergrößert werden (high dynamic range – HDR). Ein weiterer Parameter in ITU-R BT.2020 ist die Bildwiederholfrequenz, die sich seit den frühen sechziger Jahren, als Walter Bruch das sogenannte Phase Alternating Line (PAL) System mit seiner Übertragung von 50 Halbbildern pro Sekunde entwickelte, nicht verändert hat. Der neue Fernsehstandard spezifiziert Bildwiederholfrequenzen von bis zu 120 Hz bzw. 100 Hz in den Ländern, die das PAL-System nutzten. Die Wirkung dieser sowie noch weitaus höherer Bildraten auf die subjektive Wahrnehmung der Zuschauer soll in dieser Arbeit analysiert werden.

1.1 Fragestellungen

Seit der Einführung des UHDTV Standards wurde eine Vielzahl an Tests durchgeführt, die den Mehrwert der einzelnen Parameter ermitteln sollen. Hauptsächlich bezogen sich diese auf die höhere örtliche Auflösung, aber auch zu höheren Bildwiederholfrequenzen (High Frame Rates – HFR) wurden bereits einige Experimente realisiert. Die größte Schwierigkeit dabei besteht in der Verfügbarkeit geeigneter Displays zu Testzwecken, weshalb Forschungsanstalten oft gezwungen waren und sind, PC-basierte Abspielsoftware und Displays, die hauptsächlich aus dem Computerspielbereich stammen, zu verwenden. Deshalb ist die Bewertung des tatsächlichen Mehrwerts von „echtem“ HFR ohne jegliche Interpolation von Zwischenbildern eine Kernfrage dieser Arbeit. Anschließend sollen die Auswirkungen verschiedener Interpolationsarten evaluiert werden. Weiterhin soll herausgefunden werden, ob HFR nur funktioniert, wenn es durch die komplette Produktionskette hindurch

gleichermaßen verwendet wird, oder ob ein Kompromiss aus Bildrate, Shutter und Interpolation möglich ist. Dabei wird auch die Abschaffung des Zeilensprungverfahrens im UHDTV-Standard diskutiert. Die Durchführung von Tests zur Codierung ist insofern von Bedeutung, da es bisher relativ wenige Aussagen zur benötigten Übertragungsbandbreite von Fernsehen in HFR gibt.

1.2 Vorgehensweise

Nachdem die Motivation des Autors, zentrale Fragestellungen und die Vorgehensweise erläutert wurden, folgt ein Überblick zu allgemeinen Fragen der Produktion von HFR-Inhalten, der einige grundlegende Definitionen von Fehlern beinhaltet, die bei der Betrachtung der zeitlichen Auflösung von Bewegtbildern auftreten können. Das Kapitel behandelt außerdem die Themen Aufzeichnung und Wiedergabe von höheren Bildwiederholfrequenzen und schließt mit Erläuterungen zur Codierung solcher HFR-Inhalte ab.

Das dritte Kapitel zeigt auf, welche Forschungsergebnisse bereits von diversen Unternehmen, Einrichtungen und Arbeitsgruppen erzielt wurden, und der Autor schlussfolgert daraus die Art und Notwendigkeit weiterführender Experimente.

Der darauffolgende Teil der Arbeit beinhaltet eine detaillierte Beschreibung der Testumgebung und dabei u.a. Informationen zur Produktion der genutzten Testsequenzen und zum verwendeten Ausspielsystem. Außerdem wird auf die verwendeten Bewertungsmethoden eingegangen, da hierbei nach ITU-Standards gearbeitet wird, um die Reproduzierbarkeit zu gewährleisten.

Kapitel 5 beginnt mit Erläuterungen zur Durchführung der o. g. Tests zur zeitlichen Auflösung. An den entsprechenden Stellen werden die durch die zuvor beschriebenen Bewertungsmethoden erzielten Testergebnisse interpretiert und diskutiert.

Das nachfolgende Kapitel enthält Erläuterungen zu den Tests zur Codierung von HFR und soll mit Aussagen zum voraussichtlich benötigten, zusätzlichen Datenratenbedarf bei der Übertragung von HFR-Inhalten abschließen. Parallel zum Verfassen dieser Arbeit werden von der Arbeitsgruppe Broadcast Technology Future (BTF) der *European Broadcasting Union (EBU)* ebenfalls Tests zur Codierung von HFR durchgeführt. Unter ähnlichen Voraussetzungen sollen die Tests in dieser Arbeit erfolgen, sodass die Ergebnisse damit überprüft und verifiziert werden können. In welchem Maße die Testdurchführung an die der *EBU* angelehnt ist, wird an entsprechender Stelle genauer erläutert.

Im abschließenden Kapitel werden wesentliche Ergebnisse der beiden Testdurchläufe zusammengefasst und auf deren Grundlage praxisnahe Empfehlungen zu HFR-Produktionen gegeben. Die Arbeit endet mit einer Reihe von Vorschlägen für weiterführende Untersuchungen, die auf den in dieser Arbeit erzielten Ergebnissen aufbauen.

2 Produktion von HFR

In diesem Kapitel sollen einige grundlegende Informationen zur Produktion von zeitlich hochaufgelösten Bewegtbildinhalten und den Bewegungsartefakten, die dabei auftreten können, gegeben werden. Weiterhin wird ein Überblick zu den möglichen Codierstandards gegeben.

2.1 Definition von Bewegungsartefakten

Der folgende Abschnitt beschreibt die wesentlichsten, subjektiv wahrnehmbaren Bewegungsartefakte und wie diese sich gegenseitig beeinflussen. Verschiedene Begriffe werden oft gleichbedeutend verwendet, weshalb es erforderlich erscheint, eine eindeutige Erklärung der Begriffe zu geben, die in dieser Arbeit wiederkehrend vorkommen. Die Benennung orientiert sich an der Publikation von Katy Noland, die insbesondere „flicker, strobing [und] motion blur“¹, also Flimmern, Stroboskopeffekt sowie Bewegungsunschärfe beschreibt.

2.1.1 Flimmern

Flimmern beschreibt im Allgemeinen das subjektive Empfinden einer sich periodisch ändernden Bildhelligkeit² und steht in Relation zum prozentualen Anteil des Bildschirms am menschlichen Gesichtsfeld. Auch die Helligkeit und das Tastverhältnis des Flimmerns sind wichtige Faktoren, die die Intensität der Wahrnehmung beeinflussen. Eine Verringerung der Bildschirmgröße oder eine Vergrößerung des Betrachtungsabstands bewirken eine Abschwächung des Flimmerns.³ Nichtsdestotrotz ist Flimmern heutzutage nicht so problematisch wie zu Zeiten von Röhrenmonitoren (Cathode Ray Tubes – CRT), da moderne Flachbildmonitore aus physikalischen Gründen weniger anfällig für solche Störeffekte sind. Bei Bildwiederholfräquenzen oberhalb von 50 Hz sollte im Normalfall kein Flimmern auftreten. Teilweise kann es zu Flimmern kommen, wenn sehr helle Bilder auf Plasmadisplays wiedergegeben werden, da diese ähnliche Eigenschaften aufweisen wie Röhrenmonitore.

¹ Noland, Katy C. (2014): The Application of Sampling Theory to Television Frame Rate Requirements. Research & Development White Paper WHP 282. British Broadcasting Corporation. London, S. 2. Online verfügbar unter <http://downloads.bbc.co.uk/rd/pubs/whp/whp-pdf-files/WHP282.pdf>, zuletzt geprüft am 12.02.2015

² vgl. Mücher, Michael (2013): Broadcast Fachwörterbuch. 7000 Begriffe und Abkürzungen ; Fernsehen, Videotechnik, Audioteknik, Multimedia. 17., überarb. u. erw. Aufl. ; Stand 2. Januar 2013. Hamburg: BET Michael Mücher. S. 245

³ vgl. Noland 2014, S. 2

Weitaus problematischer wird es, wenn sich während der Produktion die Bildfrequenz insofern von der Frequenz umliegender Kunstlichtquellen unterscheidet, dass eine Frequenz kein Vielfaches der anderen Frequenz darstellt. Zusätzlich wirken sich auch bestimmte Shuttereinstellungen negativ aus.

Die *British Broadcasting Corporation (BBC)* hat dazu Tests durchgeführt, u.a. um die Anfrage der *EBU* an die *ITU* zur Miteinbeziehung der 100 Hz Bildwechselfrequenz in Rec. BT.2020 zu bekräftigen. Dabei wurden verschiedene Lichtquellen bei 50 Hz Netzfrequenz mit Bildwechselfrequenzen bis zu 120 Hz aufgezeichnet und anschließend mit 50 Bildern pro Sekunde betrachtet. Die Testergebnisse zeigen, dass Lampen mit Glühdraht immer dann flimmern, wenn mit 120 Hz gearbeitet wird und die Leistung verhältnismäßig gering bleibt. Bei 12 kW Glühlampen war kein Flimmern mehr erkennbar.⁴

Ebenfalls getestet wurde eine 1,2 kW Halogen-Metaldampf Lampe (Hydrargyrum Medium-Arc-Length Iodide – HMI) in den drei Konfigurationen mit ausgangsseitig 50 Hz, 75 Hz und 1000 Hz. Wurde die Lampe mit 50 Hz betrieben, trat sowohl bei einer Kameraaufnahme von 100 Hz als auch bei 120 Hz ein teilweise sehr starkes Flimmern auf. Im 75 Hz Modus wurde das Flimmern etwas abgeschwächt und im Highspeedmodus mit 1000 Hz verschwand es komplett für beide Bildfrequenzen bei der Aufnahme.⁵

Im Test mit Gasentladungslampen konnte beobachtet werden, dass bei einem Betrieb mit magnetischem Vorschaltgerät und Aufnahme mit beiden Bildwiederholraten ein Flimmern auftritt, dieses aber bei 100 Hz im Gegensatz zur 120 Hz Aufzeichnung schwächer wahrgenommen wird. Bei Nutzung eines elektronischen Vorschaltgerätes, das als Transformator für die Netzspannung fungiert, war die 100 Hz Aufnahme komplett flimmerfrei, während bei 120 Hz nach wie vor ein leichtes Flimmern zu sehen war.⁶

Die Lampe mit lichtemittierenden Dioden (Light Emitting Diode – LED) verursachte im Test der *BBC* bei keiner Aufnahme Frequenz ein Flimmern, was womöglich daran liegt, dass diese Lampen eine Vorschalt elektronik besitzen, die entweder Gleichstrom oder eine enorm hohe Ausgangsfrequenz erzeugen, was das Flimmern weitestgehend beseitigt.⁷

2.1.2 Stroboskopeffekt

Ein störender Stroboskopeffekt tritt insbesondere dann auf, wenn sich schnelle Objekte innerhalb einer Szenerie bewegen und die Frequenz der Bildaufzeichnung für eine flüssige

⁴ vgl. Quedsted, Andy; Salmon, Richard (2013): BBC interim report of the impact of camera frame rates and 50 Hz lighting. Hg. v. BBC Research & Development. London, S. 4 ff.

⁵ vgl. ebd. S. 5 f.

⁶ vgl. ebd. S. 4 f.

⁷ vgl. ebd. S. 6 f.

Darstellung der Bewegung nicht ausreicht, was ein Ruckeln oder auch schattenartige Geisterbilder um ein Objekt herum bewirkt. Eine Kamera mit einer sehr kurzen Belichtungszeit, auch Verschlusszeit oder Shutter genannt, im Vergleich zur absoluten Bilddauer fängt die Bewegung eines Objekts ein, welches auf dem Kamerasensor scharf abgebildet wird. Dieser Vorgang wiederholt sich periodisch während des gesamten Bewegungsablaufs und führt zu einer relativ langen Zeitspanne zwischen zwei aufeinander folgenden aufgezeichneten Bewegungsphasen. Das Ergebnis bei der Wiedergabe der Sequenz auf einem Display ist eine Bewegung, bei der das Objekt viele Male nacheinander kurz erscheint und seine Bewegung nicht flüssig, sondern ruckelnd wirkt.

2.1.3 Bewegungsunschärfe

Bewegte Objekte können nur aufgezeichnet werden, indem der Kamerasensor für eine endliche Zeit innerhalb der Dauer eines Bildes belichtet wird. Je länger die Belichtungszeit desto größer ist der Effekt, dass die Objekte in der resultierenden Aufnahmesequenz verwischt erscheinen. Das bedeutet, eine geringere Bewegungsunschärfe, z. B. durch Verkürzung des Shutters, verstärkt möglicherweise den Stroboskopeffekt. Die einzig zufriedenstellende Möglichkeit, beide Effekte gleichermaßen zu reduzieren, besteht in der Erhöhung der Bildwiederholfrequenz, sodass mehr Bewegungsphasen pro Zeitabschnitt aufgezeichnet werden und diese Phasen relativ kurz bleiben. Kapitel 2.2.2 beschreibt einige Tests zur Verdeutlichung der Wechselwirkung dieser Störeffekte.

Eine Unterscheidung muss bei der Betrachtung zwischen Bewegungsunschärfe im Fernsehen einerseits und im Kino andererseits gemacht werden. Da bei Kinoproduktionen grundsätzlich mit einer Bildwechselfrequenz von 24 Vollbildern, seltener auch 48 Vollbildern, gearbeitet wird, stellt sich unweigerlich eine Bewegungsunschärfe bei schnellen Schwenks ein. Maskiert wird dieser Effekt durch eine sehr kleine Schärfenebene, auf die bei Kinofilmen besonders viel Wert gelegt wird, um einen typischen Look zu erzielen. Schnelle Objekte innerhalb einer Szene werden dabei bewusst in die Unschärfe gelegt und der störende Aspekt der Bewegungsunschärfe kommt weniger zur Geltung als dies im Fernsehen der Fall ist.

2.2 Aufnahme und Wiedergabe

Dieses Kapitel beinhaltet eine Übersicht der derzeitig verfügbaren Kameraausrüstung und Displays für die Produktion und Wiedergabe von HFR-Inhalten. Zudem werden die möglichen veränderbaren Parameter bei einer HFR-Aufzeichnung kurz beschrieben.

2.2.1 Verfügbarkeit von Kameras und Displays

Bereits seit einigen Jahren werden Kameras, die eine Aufzeichnung von mehr als 50 Bildern pro Sekunde ermöglichen, besonders für Sportproduktionen vielfach genutzt. Im Gegensatz

zu echten HFR-Produktionen werden die aufgezeichneten Sequenzen dabei allerdings weiterhin mit 50 Bildern pro Sekunde und somit in Zeitlupe wiedergegeben. Großveranstaltungen aus jüngster Zeit, wie die Olympischen Winterspiele in Sotschi oder die Fußball-Weltmeisterschaft in Brasilien haben gezeigt, dass immer mehr sogenannte High Speed Kameras für die Live-Produktion eingesetzt werden. Diese Kameras erlauben das Aufnehmen von 1000 Bildern pro Sekunde und mehr, wodurch besonders beeindruckende Zeitlufenaufnahmen entstehen.

Andere Anwendungsbereiche für High Speed Kameras sind z. B. die Automobilindustrie, wo Kameras in den Forschungsabteilungen zur Visualisierung technischer Mängel verwendet werden, oder die Optoelektronik, wo hochfrequente Laser zu Testzwecken mit Kameras aufgezeichnet werden. Der Nachteil dieser Kameratypen besteht darin, dass deren mögliche Bildwiederholfrequenz oftmals von der örtlichen Auflösung der resultierenden Sequenz abhängig ist. Daher sind sie nicht praktikabel für die Broadcastindustrie mit ihren festgelegten Auflösungen in HD- oder UHD-Standards.

Dennoch gibt es viele Hersteller, die HFR-Kameras für die verschiedenen üblichen Fernsehauflösungen anbieten. Beispielmuster kommen vorrangig aus dem Bereich der digitalen Kinematografie, doch auch Fernsehkameras für den Live-Betrieb beherrschen zum Teil schon die höheren Bildwechselfrequenzen. Tab. 1 gibt einen Überblick einiger kürzlich veröffentlichter Modelle.

Hersteller	Modell	Max. örtliche Auflösung	Max. Bildrate
AJA	Cion	3840 x 2160	120 Hz
ARRI	Alexa	1920 x 1080	120 Hz
	Amira	1920 x 1080	200 Hz
Grass Valley	LDX XtremeSpeed	1920 x 1080	300 Hz
Panasonic	VariCam 35	3840 x 2160	120 Hz
Sony	F55	1920 x 1080	180 Hz
	F65	3840 x 2160	240 Hz
VisionResearch	Phantom Flex4K	3840 x 2160	1000 Hz

Tabelle 1: Kameramodelle für HFR-Produktion

Demnach ist das Problem beim Umgang mit HFR-Inhalten weniger die Produktion sondern vielmehr die Übertragung und Wiedergabe, da für die enorm hohen Datenraten und die vielfach erhöhte Bewegungsauflösung keine konventionellen Schnittstellen und Displays genutzt werden können.

2.2.2 UHDTV ohne Zeilensprungverfahren

Wie eingangs erwähnt (vgl. 1.1), beinhaltet ITU-R BT.2020 im Gegensatz zum Vorgängerstandard BT.709 keinerlei Hinweise auf die Produktion und Wiedergabe mit Halbbildern. In den letzten Jahren ist bereits einer der ursprünglichen Beweggründe, das Zeilensprungver-

fahren zu verwenden, nach und nach aus der Rundfunklandschaft verschwunden – es werden keine Wiedergabegeräte mit Kathodenstrahlröhre mehr gebaut, die aufgrund ihrer Flimmeranfälligkeit den Ausschlag für die Entwicklung des Zeilensprungverfahrens vor fast einhundert Jahren gaben. Heutige Displays arbeiten durchweg progressiv, das heißt bei Zuführung eines Signals mit Halbbildern muss vor der Wiedergabe ein Deinterlacing erfolgen, was eine zusätzliche potenzielle Fehlerquelle bedeutet. Der HEVC-Codierstandard, der im weiteren Verlauf dieses Kapitels noch näher beschrieben wird und der eine Möglichkeit zur effizienten Codierung von UHD-Material bietet, beinhaltet zudem keine speziellen Werkzeuge für die Codierung von Interlaced-Material mehr.

Dennoch spricht nach wie vor der um 100 % erhöhte Bandbreitenbedarf bei der Übertragung von 50 Vollbildern anstelle von 50 Halbbildern pro Sekunde gegen ein sofortiges Ende der Halbbildübertragung. Immerhin arbeiten einige Rundfunkanstalten produktionsseitig erst seit wenigen Jahren im HD-Standard 1080i/25. Ein sukzessiver Übergang zu einem rein progressiven Fernsehsystem in der Produktion und der Ausstrahlung ist aus den o. g. Gründen dennoch empfehlenswert.

2.2.3 Veränderbare Parameter

Neben der Bildwiederholrate gibt es mit der Belichtungszeit eine weitere kameraseitige Größe, mit der das Verhalten bei der Aufnahme bewegter Objekte beeinflusst werden kann. Die Belichtungszeit lässt sich über eine mechanische Umlaufblende, insbesondere bei Filmkameras, oder mithilfe eines elektronischen Shutters in der Kamera definieren. Bei Filmkameras findet sich häufig die Bezeichnung in Grad, die sich auf den nicht durch die Umlaufblende verdeckten Sektor bezieht. Eine Shuttereinstellung von 180° bei einer Bildrate von 50 Hz ist demzufolge gleichbedeutend mit einer Belichtungszeit von $1/100$ s. Ebenfalls geläufig ist eine relative Angabe der Belichtungsdauer zur Bilddauer. Für das eben genannte Beispiel bedeutet dies einen Shutter von 50 %.

Eng verbunden mit diesen Einstellungen ist auch die Lichtmenge, die auf den Sensor trifft. Jede Verkürzung der absoluten Belichtungszeit, sei es durch eine höhere Bildwechselfrequenz oder eine Verringerung des Shutters, führt demzufolge auch zu einer geringeren Lichtmenge. Neben der Gefahr eines zunehmenden Stroboskopeffekts bei Verkürzung der relativen Belichtungszeit besteht also zusätzlich das Risiko eines verstärkten Bildrauschens bis hin zur Unterbelichtung des Bildes.

Um die beiden genannten Kameraparameter und deren Einfluss auf die Aufzeichnung bewegter Objekte besser verstehen zu können, werden im Folgenden Tests beschrieben, die die Wechselwirkung von Bewegungsunschärfe, Shutter und schließlich Bildwiederholfrequenz belegen. Die getroffenen Aussagen beziehen sich dabei auf Einzelframes von Testsequenzen, anhand derer eine störende Bewegungsunschärfe ausgemacht werden kann. Stroboskopeffekte und Flimmern werden jedoch vorerst vernachlässigt, da sich diese Effekte erst bei der tatsächlichen Betrachtung der Sequenz am Display zeigen und nicht in einzeln extrahierten Bildern.

Abb. 2 zeigt den Einfluss der Shuttereinstellung auf die Bewegungsunschärfe bei konstanter Objektgeschwindigkeit im Bild. Dabei wurde ein mit konstanter Geschwindigkeit rotierender Siemensstern mit einer unveränderlichen Bildfrequenz von 25 Bildern pro Sekunde bei verschiedenen Shutterwerten, also Belichtungszeiten, aufgenommen. Zum Vergleich dazu ist in Abb. 1 der unbewegte Stern zu sehen.



Abbildung 1: Unbewegter Siemensstern

Zur Aufzeichnung wurde das Kameramodell EOS 5D Mark III von *Canon* genutzt. Das linke Bild in Abb. 2 stammt aus einer Sequenz mit $1/50$ s Belichtung, wobei die einzelnen Segmente des Siemenssterns offensichtlich verwischt erscheinen. Das Bild in der Mitte weist eine geringere Bewegungsunschärfe auf, bewirkt durch eine kürzere Belichtungszeit von $1/200$ s und im rechten Bild, das aus einer Sequenz mit $1/1000$ s Belichtung stammt, ist keine wahrnehmbare Bewegungsunschärfe ersichtlich.



Abbildung 2: Rotierender Siemensstern bei verschiedenen Shutterwerten ($1/50$, $1/200$, $1/1000$ Sekunde) und Aufnahme in 1080p/25

Im Gegensatz dazu wurde getestet, welche wahrnehmbare Bewegungsunschärfe bei verschiedenen Geschwindigkeiten des rotierenden Siemenssterns und gleichbleibender Belichtungszeit auftreten. Aus Gründen der Reproduzierbarkeit der jeweiligen Geschwindigkeiten wurde die Skala des verwendeten Funktionsgenerators, der die Eingangsfrequenz des antreibenden Schrittmotors erzeugt, auf die Einheit Umdrehungen pro Sekunde (U/s) umgerechnet. Abb. 3 zeigt aus einer 1080p/25-Sequenz exportierte Einzelframes mit $1/50$ s Belichtung, wobei die Rotationsgeschwindigkeit schrittweise gesteigert wurde.



Abbildung 3: Rotierender Siemensstern bei verschiedenen Rotationsgeschwindigkeiten (0,15, 0,375, 0,94 U/s) und Aufnahme in 1080p/25

Bei geringer Drehzahl von 0,15 U/s ist nahezu kein Auflösungsverlust erkennbar. Die Abbildung in der Mitte weist bereits einen wahrnehmbaren Auflösungsverlust auf, welcher bei 0,94 U/s, der höchstmöglichen Drehzahl im Testaufbau, deutlich zunimmt. Um einen Vergleich mit höheren Bildwechselfrequenzen ziehen zu können, wurde der Test mit Sequenzen in 720p/120 wiederholt. Letztere wurden mit einem *Apple iPad Air 2* (Modell: MGKL2FD/A) erstellt. Wie aus Abb. 4 ersichtlich ist, tritt eine deutlich wahrnehmbare Bewegungsunschärfe erst bei sehr viel höherer Drehzahl auf und zwar bei 0,94 U/s für das aufgezeigte Beispiel.



Abbildung 4: Rotierender Siemensstern bei verschiedenen Rotationsgeschwindigkeiten (0,15, 0,375, 0,94 U/s) und Aufnahme in 720p/120

Das beweist bereits, dass nur eine erhöhte Bildwechselfrequenz sowohl scharfe als auch flüssige Bewegungsabläufe garantieren kann.⁸

⁸ vgl. Noland 2014, S. 3

2.3 Codierung

Wie bei allen Fernsehinhalten so soll auch bei Produktionen in HFR das Ziel sein, einen Mehrwert beim Zuschauer zu Hause zu schaffen. Dazu müssen die Inhalte über die gängigen Distributionskanäle dorthin übertragen werden, was eine Quellencodierung aufgrund der begrenzten Bandbreite unumgänglich macht. Da HFR momentan in erster Linie als Systemfaktor für UHDTV standardisiert ist, muss davon ausgegangen werden, dass erhöhte Bildwechselfrequenzen gemeinsam mit erhöhten spatialen Auflösungen genutzt werden. Daher sollten auch Codierstandards zum Einsatz kommen, die effizient genug erscheinen, die enormen Datenmengen eines unkomprimierten UHD-Signals auf ein Minimum zu reduzieren und dabei den Mehrwert von HFR beizubehalten, um die Bildqualität zum Konsumenten zu bringen.

2.3.1 Codec

Der aktuell am häufigsten mit UHDTV in Verbindung gebrachte Codierstandard, der auch schon für diverse Testausstrahlungen genutzt wurde, ist High Efficiency Video Coding (HEVC). Die Bezeichnung der *ITU* für den Codec lautet H.265, die der *ISO/IEC* lautet 23008-2, bei MPEG selbst heißt es MPEG-H Part 2. Im Vergleich zum Vorgänger H.264 soll eine bis zu 50 % bessere Codiereffizienz bei subjektiver Betrachtung erreicht werden, möglich durch neue Kompressionswerkzeuge und die Weiterentwicklung bekannter Techniken. Demnach sind beispielsweise variable Größen von Coding-Tree-Blocks (CTB) zwischen 8x8 und 64x64 Bildpunkten möglich, während der Vorgängerstandard H.264 lediglich 16x16 Pixel große Makroblöcke unterstützte. Zudem enthält der neue Codec eine verbesserte Intra-Prädiktion, also Bewegungsschätzung innerhalb eines Bildes, da 35 statt 10 Prädiktionsmodi verwendet werden können und deshalb in feinere Winkel aufgelöst wird.⁹ Mit HEVC können Videos mit einer Auflösung von bis zu 8192 x 4320 Pixeln codiert werden, womit auch UHD-2 abgedeckt wäre. Bildwechselfrequenzen sind bis zu 300 Hz für HD und UHD-1 und 120 Hz für UHD-2 standardisiert.¹⁰ Die *ITU* empfiehlt HEVC laut Rec. BT.2073 für die Übertragung von stark komprimierten Inhalten in HDTV und UHDTV.¹¹

⁹ vgl. Bross, Benjamin; Marpe, Detlev; Schwarz, Heiko (2013): "High-Efficiency Video Coding" - (HEVC)-Standard. Vorstellung des neuen Kompressionsstandards. In: *FKT 67* (1-2), S. 56–63

¹⁰ vgl. International Telecommunication Union (2014): High efficiency video coding. Series H: Audiovisual and Multimedia Systems. Infrastructure of audiovisual services - Coding of moving video. Recommendation ITU-T H.265. Hg. v. Telecommunication Standardization Sector of ITU. International Telecommunication Union. Genf, S. 240

¹¹ vgl. International Telecommunication Union (2015): Use of the high efficiency video coding (HEVC) standard for UHDTV and HDTV broadcasting. Recommendation ITU-R BT.2073-1. Hg. v. Radiocommunication Sector of ITU. International Telecommunication Union. Genf

2.3.2 Relevante Codingparameter

Bei der Codierung von Videodaten gibt es eine Reihe von Parametern, die Einfluss auf die resultierende Qualität haben. An dieser Stelle soll ein kurzer Überblick gegeben werden, inwieweit sich diese auf HFR-Inhalte auswirken können.

GOP-Struktur

Die Group of Pictures (GOP) beschreibt die Abhängigkeiten zwischen benachbarten Einzelbildern. I-Frames werden nur innerhalb eines Bildes codiert, während P-Frames nur auf zeitlich vorhergehende Bilder und B-Frames auf zeitlich vorhergehende oder nachfolgende Bilder verweisen. Eine Open GOP bedeutet in diesem Zusammenhang, dass B-Frames sich auch auf Bilder benachbarter GOPs beziehen können. Die Reihenfolge der Darstellung der Bilder unterscheidet sich folglich von der Reihenfolge ihrer Codierung. Sequenzen mit vielen B-Frames sind schwieriger zu codieren und führen zu einem größeren Delay, benötigen aber gleichzeitig auch weniger Bandbreite, da es weniger I-Frames zu codieren gibt. Der Bandbreitenbedarf lässt sich noch weiter minimieren, indem die GOP dynamisch konfiguriert wird und ihre Länge somit je nach Komplexität der zu codierenden Bildinhalte variiert. Encoder verfügen über eine Funktion zur Angabe der GOP-Länge und -Struktur, so ist für TV-Distributionsencoder eine Länge von etwa einer halben Sekunde üblich. Der HEVC-Standard verwendet für die GOP und deren Elemente eine andere Syntax, die grundlegenden Funktionen sind aber mit o. g. vergleichbar. Für die Codierung von HFR-Sequenzen ließe sich schlussfolgern, dass aufgrund der größeren Zahl der Bilder längere GOPs ohne ein spürbar größeres Delay möglich sind. Konfiguriert man beim Encoding einer Sequenz mit 50 Hz etwa eine GOP-Länge von 32, würde sich für 200 Hz und einer GOP-Länge von 128 dasselbe Verhalten bei der Codierung einstellen.

Zeitliche Skalierbarkeit

Die *ITU*-Empfehlung BT.2073 beinhaltet neben den empfohlenen Datenraten auch ein präferiertes Codierschema für die Übertragung von HFR (vgl. Abb. 5). Das Ziel ist eine Abwärtskompatibilität zukünftiger HFR-Ausstrahlungen zu konventionellen TV-Decodern, die durch eine unabhängige Codierung der zusätzlichen Bilder in eine separate Schicht im Bitstrom möglich wird. Diese Funktion wird seitens des HEVC-Standards unterstützt. Encodiert wird der HFR-Strom in ein sogenanntes Subset (oder Enhancement Layer), während der Sub-Bitstream (oder Base Layer) mit der konventionellen Bildrate daraus abgeleitet wird. Zum Decodieren benötigte Referenzbilder werden dabei ausschließlich in den Sub-Bitstream geschrieben, damit konventionelle Decoder den Bitstrom in jedem Fall decodieren können. Die beiden Layer weisen folglich sowohl unterschiedliche GOP-Strukturen, als auch verschiedene Shutter auf (vgl. 4.2.1), weshalb eventuell auftretende Stroboskopeffekte durch Interpolation im Display kompensiert werden müssen.

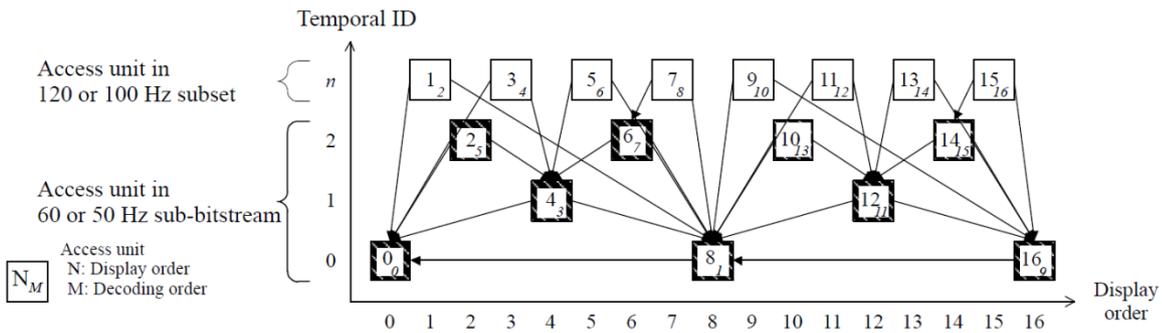


Abbildung 5: Zeitlich skalierbare Codierung mit HEVC¹²

Datenrate

Diese Größe gibt an, wie viele Daten pro Sekunde bei der Codierung zur Verfügung gestellt werden. Komplexe Bewegungsabläufe und detailreiche Bilder benötigen eine höhere Codierdatenrate als beispielsweise Standbildaufnahmen. Verschiedene Codecs arbeiten verschieden effizient, so benötigt HEVC beispielsweise bis zu 50 % weniger Datenrate als der Vorgängerstandard H.264 bei gleicher subjektiver Bildqualität bzw. liefert bei gleicher Datenrate eine subjektiv doppelt so hohe Bildqualität. Bei der Codierung von Rundfunkinhalten für die Ausstrahlung wird logischerweise versucht, eine hohe Qualität mit der geringstmöglichen Datenrate zu erzielen, da Übertragungskapazitäten immer begrenzt sind.

Vereinzelt wird davon ausgegangen, dass bei einer Codierung von HFR gegenüber konventionellen Bildraten nur wenig mehr bzw. sogar weniger Datenrate benötigt wird. Grund für diese Annahme ist die geringere Differenz zwischen den einzelnen Quellbildern bei HFR und deren geringerer Auflösungsverlust aufgrund der kürzeren Integrationszeit in der Kamera. Eine erste Bestätigung dieser Tendenz ist die Tatsache, dass zur Übertragung von 1080p/50 weit weniger als die doppelte Datenrate einer Übertragung in 1080i/25 benötigt wird.¹³

In ITU-R Rec. BT.2073 werden für verschiedene örtliche Auflösungen und Bildraten Empfehlungen zur Datenrate gegeben. Die dort aufgeführten Werte sind Maximalwerte, die bei der Übertragung kritischer Sequenzen noch eine ausreichend hohe Qualität liefern. Für 1080p/50 werden dort 10–15 Mbit/s und für 2160p/50 eine Spanne von 30–40 Mbit/s empfohlen. Zusätzlich werden bereits Angaben für Sequenzen in UHD und HFR gemacht, für die bis 120 Hz laut Empfehlung etwa 35–50 Mbit/s Codierdatenrate veranschlagt werden sollten.¹⁴

¹² International Telecommunication Union 2015, S. 4

¹³ vgl. Roberts, Alan (2009): Circles of confusion. Television, how it works, why it works this way, how to get the best from it. Genf: European Broadcasting Union

¹⁴ vgl. International Telecommunication Union 2015, S. 2

3 Gegenwärtiger Forschungsstand

Neben den Aktivitäten des *IRT* waren und sind weitere Institutionen und Arbeitsgruppen damit beschäftigt, den Mehrwert höherer Bildwechselfrequenzen zu erforschen. In diesem Kapitel sollen daher bisherige Forschungsergebnisse erläutert werden, die die Vorgehensweise des Autors in dieser Arbeit beeinflussen.

3.1 Nachrichtentechnische Gegebenheiten

Am *IRT* wurden die Probleme eingehend untersucht, die aufgrund der allgemein üblichen realen Szenengestaltung auftreten. Demnach kommt es in der Praxis regelmäßig bereits bei der Aufzeichnung zu einem Auflösungsverlust, da Bewegungsartefakte bei schnellen Kameraschwenks oftmals vernachlässigt werden. Die Grenzgeschwindigkeit für einen horizontalen Schwenk über eine Szenerie ergibt sich aus der Bildbreite, die über die örtliche Auflösung klar definiert ist, sowie der Aufnahmefrequenz und der verwendeten Shuttereinstellung (vgl. Formel 1)¹⁵.

$$v_{\text{grenz}} = 2 \cdot \text{Bildbreite} \cdot \frac{f_{\text{Bild}}}{N \cdot A}$$

Formel 1: Grenzgeschwindigkeit

Wird demzufolge bei einer konventionellen HD-Produktion über nahezu die komplette Bilddauer belichtet, was einer Shuttereinstellung von 100 % oder 360° entspricht (vgl. 2.2.2), muss ein Schwenk über die komplette Szene fast 20 Sekunden dauern, damit die volle HD-Auflösung erhalten bleibt. Diese Herangehensweise berücksichtigt rein nachrichtentechnische Gesichtspunkte und ist daher objektiv. Außerdem liegt der Fokus hierbei allein auf der Bewegungsunschärfe. Weitere störende Effekte werden nicht speziell betrachtet.

3.2 Physiologische Gegebenheiten

Katy Noland verfolgt dagegen in ihrer Arbeit verschiedene Ansätze zur Ermittlung der benötigten Bildwechselfrequenz. Neben dem Abtasttheorem nach Nyquist und Shannon

¹⁵ vgl. Driesnack, Dagmar; Gierlinger, Friedrich; Schäfer, Rainer (2014): Höhere Bildwechselfrequenz für Ultra-HD? In: *FKT* 68 (12), S. 641–644

(vgl. Formel 2)^{16,17} werden physiologisch-optische Modelle herangezogen, die unterschiedliche Aussagen zur benötigten Frequenz für Bewegungen erlauben. Zum einen für Objektbewegungen, die mit den Augen verfolgt werden und zum anderen solche, bei denen die Augen statisch bleiben.

$$f_s > 2 \cdot f_{max}$$

Formel 2: Nyquist-Shannonsches Abtasttheorem

Soll das Abtasttheorem nicht verletzt werden, dürfen sich Objekte wie erwartet nur mit sehr geringer Geschwindigkeit bewegen. Zur verlustfreien Abbildung von Objekten, die sich mit 3,4°/s bewegen, werden bereits etwa 200 Hz benötigt. Die Werte decken sich hierbei mit denen von Driesnack et al. Wird das Modell zur Kontrastempfindlichkeitsfunktion mit in die Berechnung einbezogen, ergeben sich für dieselbe Objektgeschwindigkeit Bildwechselfrequenzen von etwa 140 Hz bei den örtlichen Auflösungen von HD, UHD-1 und UHD-2 für deren jeweils optimalen Betrachtungsabstand. Für eine Verfolgung des Objekts mit den Augen wurden ca. 700 Hz errechnet, die ebenfalls für alle genannten örtlichen Auflösungen und Betrachtungsabstände gelten.¹⁸ Diese Aussagen stützen sich zwar auf diverse Modelle, die wiederum auf Untersuchungen der menschlichen Wahrnehmung basieren, sind jedoch aus Mangel an Wiedergabegeräten nicht durch subjektive Tests belegt.

Während der kontinuierlichen Verfolgung von Objekten erreicht das Auge Geschwindigkeiten von maximal 30–80°/s¹⁹, wobei die Augengeschwindigkeit nur etwa 82 % der Geschwindigkeit des verfolgten Objekts beträgt. Wird der Unterschied zu groß, wird er durch Sakkaden, also spontane willkürliche Augenbewegungen, aufgeholt.²⁰ Bei einer maximalen Objektgeschwindigkeit von mehr als $(80/0,82) = 97,56°/s$ ist der Mensch folglich nicht mehr in der Lage ein sich bewegendes Objekt kontinuierlich im Bereich des schärfsten Sehens im Auge, der Fovea centralis, abzubilden.²¹ Diese Maximalgeschwindigkeit bildet damit den Grenzwert, bis zu dem es theoretisch möglich ist, ein bewegtes Objekt scharf zu sehen.

¹⁶ vgl. Nyquist, Harry (1928): Certain Topics in Telegraph Transmission Theory. In: *Transactions of the American Institute of Electrical Engineers* 47 (2), S. 617–644

¹⁷ vgl. Shannon, Claude Elwood (1948): A Mathematical Theory of Communication. In: *The Bell System Technical Journal* 27 (3, 4), S. 379-423, 623-656

¹⁸ vgl. Noland 2014, S. 15

¹⁹ Koesling, Hendrik (2011/2012): Augenbewegungen und visuelle Aufmerksamkeit. Visuelle Aufmerksamkeit und Blickbewegungen. Universität Bielefeld, AG Neuroinformatik. Bielefeld, 2011/2012. Online verfügbar unter <http://www.techfak.uni-bielefeld.de/~ihkoesli/vab2011/07-vab2011-hk-augenbewegungen.pdf>, zuletzt geprüft am 11.05.2015

²⁰ vgl. Daly, Scott (2001): Engineering Observations from Spatiovelocity and Spatiotemporal Visual Models. In: *Vision Models and Applications to Image and Video Processing*, S. 179–200

²¹ vgl. Noland 2014, S. 13

Der zweite Grenzwert, der deutlich darunter liegt, beschreibt die Maximalgeschwindigkeit ohne nachrichtentechnisch bedingten Auflösungsverlust, entsprechend den Aussagen von Noland und Driesnack et al. – dieser beträgt ca. 3,4°/s. Innerhalb dieser Spanne liegt der zu ermittelnde Wert für einen gerade wahrnehmbaren Auflösungsverlust durch Bewegungsunschärfe, der im Kapitel 3.4 ermittelt werden soll.

Auch für die Wahrnehmung von Flimmern existieren physiologische Modelle. Die japanische öffentlich-rechtliche Rundfunkanstalt *Nippon Hōsō Kyōkai (NHK)* unternahm Versuche, um die sogenannte Flimmerfusionsfrequenz (Critical Flicker Frequency – CFF) des menschlichen Sehsinnes zu ermitteln. Sie beschreibt die Frequenz, bei der aufeinander folgende Lichtblitze als kontinuierliches Leuchten wahrgenommen werden und hat damit direkten Einfluss auf die notwendige, flimmerfreie Bildwechselfrequenz im Fernsehen. Die Untersuchungen ergaben, dass die CFF mit steigender Leuchtdichte des Bildschirms und dessen Anteil am Gesichtsfeld zunimmt (vgl. 2.1.1) und eine Bildfrequenz von mindestens 80 Hz notwendig ist, damit der Großteil der Menschen das Flimmern nicht wahrnimmt.²² Dieser Wert liegt über den für HDTV standardisierten 50 Hz, weshalb die Einführung von HFR, sei es mit einer gleichzeitigen Erhöhung der örtlichen Auflösung oder ohne, in jedem Fall eine Minimierung oder Aufhebung des Flimmerns bewirkt.

3.3 Subjektive Wahrnehmung

Bisher wurden von mehreren Institutionen subjektive Tests durchgeführt, um den Mehrwert von HFR gegenüber konventionellen Bildwechselfrequenzen zu ermitteln.

Das *IRT* beteiligt sich innerhalb der *EBU* an Projekten der Arbeitsgruppe BTF, zu denen auch subjektive Tests zu HFR gehören, die Mitte des Jahres 2013 am *IRT* durchgeführt wurden. Dabei wurden fünf ausgewählte Testsequenzen einer Reihe von Experten und Nicht-Experten vorgestellt, wobei jeweils die Bildfrequenz und die Shuttereinstellung variiert werden konnte. Im Mittel über alle Sequenzen verbesserte sich die bewertete Qualität sowohl für 120 Hz gegenüber 60 Hz als auch für 240 Hz gegenüber 120 Hz um jeweils eine Stufe, was auf einer fünfstufigen, adjektivisch bestimmten Bewertungsskala gemäß Rec. ITU-R BT.500²³ eine recht deutliche Qualitätssteigerung bedeutet. Auch eine Halbierung der Shutter-Apertur wurde im Durchschnitt besser bewertet als eine Belichtung über die gesamte Bilddauer. Bei separierter Betrachtung der einzelnen Sequenzen gab es hierbei Unterschiede (vgl. Abb. 6). Die von *NHK* zur Verfügung gestellten und im Test verwen-

²² vgl. Sugawara, Masayuki (2011): Psychophysical requirements for higher frame rates in future television. DCS 2011 - Advances in Image and Sound, 3D, 4K and beyond. NHK. NAB Show, SMPTE. Las Vegas, 10.04.2011

²³ vgl. International Telecommunication Union (2012): Methodology for the subjective assessment of the quality of television pictures. Recommendation ITU-R BT.500-13. Hg. v. Radiocommunication Sector of ITU. International Telecommunication Union. Genf, S. 20

deten Displays arbeiteten durchweg in 240 Hz, weshalb bei niedrigeren Bildraten Quellbilder wiederholt werden mussten. Es besteht die Vermutung, dass die Ergebnisse dadurch etwas beeinträchtigt wurden.

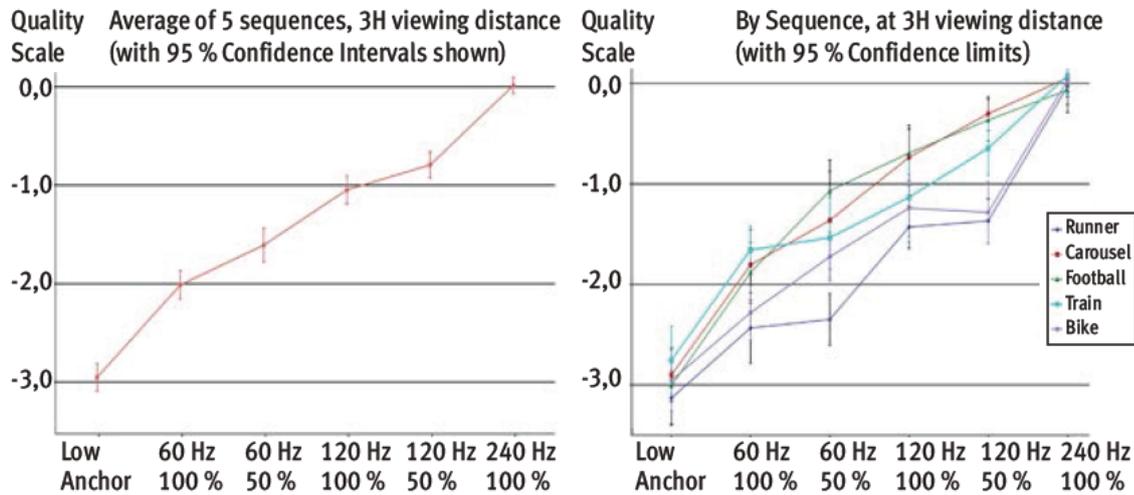


Abbildung 6: Ergebnisse der subjektiven Tests der BTF-HFR²⁴

Auch das französische Telekommunikationsunternehmen *Orange* führte im Rahmen des Forschungsprojekts For Enhanced Video Experience (4EVER) solche Tests durch. Hierbei wurde verstärkt auch darauf geachtet, inwieweit sich eine Bewegungsinterpolation innerhalb des Displays auswirkt. Anders als bei den Tests der BTF-HFR verwendete man ein 60 Hz- und ein 120 Hz-Display sowie mit Subjective Assessment of Multimedia Video Quality (SAMVIQ) ein anderes Bewertungsverfahren. Die Wiedergabe von nativem 120 Hz-Material auf dem 120 Hz-Display wurde als exzellent bewertet, während 60 Hz-Sequenzen auf dem 60 Hz-Display etwa 20 Bewertungspunkte weniger erhielten und damit als gut angesehen wurden. Erstaunlicherweise wurde die Wiedergabe von 60 Hz-Material mit 50 % Shutter und Bewegungsinterpolation auf dem HFR-Display sehr ähnlich bewertet wie die bereits erwähnte exzellente 120 Hz-Wiedergabe.²⁵ Das deutet darauf hin, dass womöglich in Zukunft je nach Produktion unterschieden werden kann, ob in HFR produziert wird oder die zusätzlichen Bilder interpoliert werden können.

Nachdem Mitte 2014 zwanzig sogenannte Range Extension Profiles²⁶ zur professionellen Anwendung von HEVC zum Standard hinzugefügt wurden, führte das schwedische Unternehmen *Ericsson* Untersuchungen durch, um die benötigten Datenraten bei der Kontribution von UHD-Material mit HFR zu ermitteln. Zur Codierung der Testsequenzen wurde das

²⁴ Driesnack et al. 2014, S. 643

²⁵ vgl. Clare-Charrier, Maryline (2015): HDR, HFR and BT.2020. Exploring content and visual comfort. EBU Production Technology Seminar 2015. Orange Labs. European Broadcasting Union, Technology & Innovation Department. Genf, 29.01.2015

²⁶ vgl. International Telecommunication Union 2014, S. 228 ff.

Main 4:2:2 10 Profile verwendet, wobei zusätzlich zwei verschiedene GOP-Strukturen eingesetzt wurden, um unterschiedliche Delay-Modi zu repräsentieren. Die Testergebnisse wurden rein durch Berechnung des durchschnittlichen Spitzen-Signal-Rausch-Verhältnisses (Peak Signal-to-Noise Ratio – PSNR) ermittelt und spiegeln damit keinen subjektiven Eindruck wider, sondern das Verhältnis von Codierdatenrate und dem mittleren Fehler über alle codierten Einzelbilder gegenüber der Quelle (vgl. Formel 3).

$$PSNR = 10 \cdot \lg \left(\frac{I_{max}^2}{MSE} \right) \text{ [dB]}$$

Formel 3: Spitzen-Signal-Rausch-Verhältnis

Die beiden verwendeten Testsequenzen „ReadySetGo“ und „Beauty“ werden durch eine Forschungsgruppe der Tampere Universität in Finnland frei zur Verfügung gestellt.²⁷ Ersterer beinhaltet eine Startszene aus einem Pferderennen, die in Zeitlupe wiedergegeben wird. „Beauty“ zeigt das Gesicht einer blonden Frau in einer schwach beleuchteten Umgebung, ebenfalls mit verlangsamter Wiedergabe. Die Ergebnisse sind in Abb. 7 dargestellt.

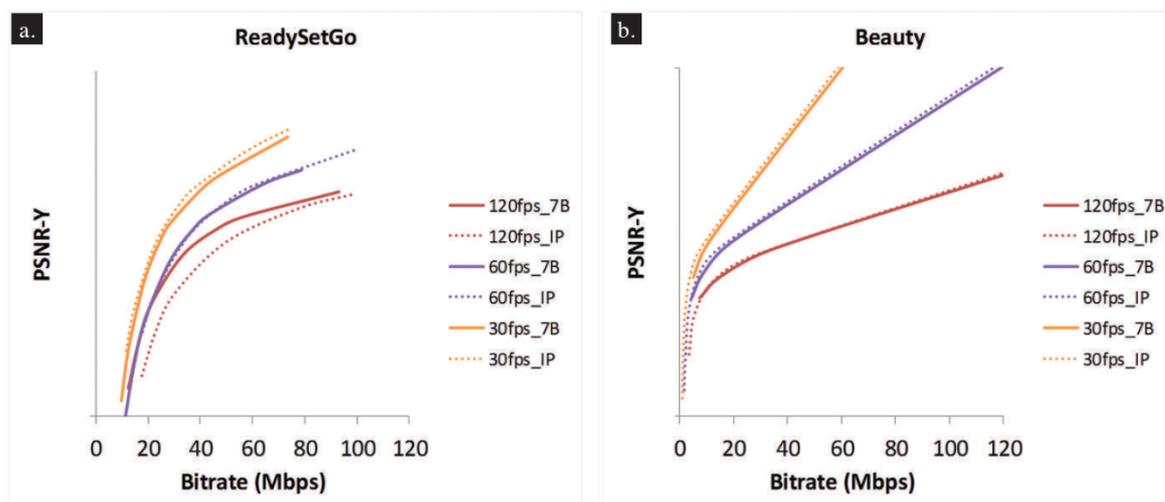


Abbildung 7: Testergebnisse der Ericsson HFR-Kontributionstests mit dem HEVC Range Extension Profile Main 4:2:2 10²⁸

Die Sequenz „ReadySetGo“ benötigte für dieselbe objektive Qualität beim Übergang von 60 Hz zu 120 Hz ca. 72 % mehr Datenrate bei einer GOP mit stetig wechselnden I- und P-Frames und etwa 42 % mehr Datenrate bei einer GOP-Struktur mit sieben B-Frames zwischen den Referenzbildern. Für die zweite Testsequenz musste die Datenrate verdoppelt werden, um bei 120 Hz gegenüber 60 Hz denselben PSNR-Wert zu erzielen. Erklärt wird

²⁷ vgl. Ultra Video Group (Hg.): Test Sequences. Tampere University of Technology. Online verfügbar unter <http://ultravideo.cs.tut.fi/#testsequences>, zuletzt geprüft am 20.05.2015

²⁸ Goldman, Matthew; Litwic, Lukasz; Baumann, Olie (2015): Ultra-HD Content Acquisition and Exchange. Using HEVC Range Extensions. In: *SMPTE Motion Imaging Journal* 124 (3), S. 28–36

dies durch den generell geringeren Signal-Rausch-Abstand in dieser Sequenz, der dem Encoder das Erkennen von zeitlichen Redundanzen erschwert. Zudem beinhaltet die Sequenz kaum Bewegung, weshalb wenig Datenrate zur Errechnung von Bilddifferenzen aufgebracht werden muss und die Unterschiede in den getesteten GOP-Strukturen keine Rolle spielen.²⁹

3.4 Testaufbau zur Geschwindigkeit bewegter Objekte

Anhand des bereits in Kapitel 2.2.3 beschriebenen Testaufbaus wird versucht, eine Aussage über die subjektiv wahrnehmbaren Störeffekte bewegter Objekte innerhalb einer Szene zu treffen und den Geschwindigkeitswert zu ermitteln, bis zu dem subjektiv kein Auflösungsverlust wahrnehmbar ist. Da der Siemensstern in diesem Versuchsaufbau vertikal bildfüllend eingerichtet wird, entspricht sein Durchmesser der Bildhöhe. Die Drehzahl wird variiert durch Veränderung der Eingangsfrequenz des Schrittmotors, der die Rotation des Sterns bewirkt. Die Eingangsfrequenz wurde im Test schrittweise erhöht und immer der jeweilige Winkel φ gemessen, der vom äußersten Punkt des Sterns in einer Sekunde überstrichen wird. Anhand der Untersuchung der jeweiligen Drehzahl bei den verschiedenen Eingangsfrequenzen können Geschwindigkeitsangaben gemacht werden, die repräsentativer erscheinen. Wird der Unterschied in der Wahrnehmung von Kreisbewegungen im Gegensatz zu linearen Bewegungen vernachlässigt, ist eine Drehung des Siemenssterns um 204° gleichbedeutend mit einer zurückgelegten Strecke des äußersten Punktes des Sterns von der Länge einer Bildbreite. Aus dem Winkel φ geteilt durch die ermittelten 204° lässt sich also recht einfach die Geschwindigkeit des Punktes in Bildbreite pro Sekunde ermitteln. Diese Einheit verwenden Driesnack et al. bekanntlich ebenfalls zur Angabe der Grenzgeschwindigkeit. Folglich ergibt sich die Objektgeschwindigkeit in Pixel pro Sekunde aus dem Produkt der Objektgeschwindigkeit in Bildbreite pro Sekunde und 1920, also der Anzahl der horizontalen Pixel.

Mit diesen Voraussetzungen und der Annahme eines optimalen Betrachtungsabstandes ohne Auflösungsverlust von der dreifachen Bildschirmhöhe (3 H) kann die Formel

$$b = \pi \cdot r \cdot \frac{\alpha}{180^\circ}$$

Formel 4: Kreisbogen

schließlich in die Formel

$$v = \frac{b \cdot 180^\circ}{\pi \cdot r \cdot t} \text{ [°/s]}$$

Formel 5: Objektgeschwindigkeit

²⁹ vgl. Goldman et al. 2015, S. 34f.

übersetzt und die Objektgeschwindigkeit in Grad pro Sekunde berechnet werden. Letztere Einheit hat besondere Relevanz für die Betrachtung von menschlichen Augenbewegungen. Die mit diesem Testaufbau mögliche Minimalgeschwindigkeit liegt bei $9,15^\circ/\text{s}$ (vgl. Tab. 2).

f [kHz]	φ [$^\circ/\text{s}$]	Drehzahl [U/s]	v [Bildbreite/s]	für 1080p		für 720p	
				v [Pixel/s]	v [$^\circ/\text{s}$]*	v [Pixel/s]	v [$^\circ/\text{s}$]*
1,5	55	0,15	0,27	518	9,17	346	9,17
2	70	0,19	0,34	659	11,65	439	11,65
2,5	85	0,24	0,42	800	14,15	533	14,15
3	100	0,28	0,49	941	16,64	627	16,64
3,5	120	0,33	0,59	1129	19,97	753	19,97
4	135	0,38	0,66	1271	22,47	847	22,47
4,5	155	0,43	0,76	1459	25,80	973	25,80
5	175	0,49	0,86	1647	29,13	1098	29,13
5,5	190	0,53	0,93	1788	31,62	1192	31,62
6	210	0,58	1,03	1976	34,95	1318	34,95
6,5	225	0,63	1,10	2118	37,45	1412	37,45
7	240	0,67	1,18	2259	39,94	1506	39,94
7,5	260	0,72	1,27	2447	43,27	1631	43,27
8	275	0,76	1,35	2588	45,77	1725	45,77
8,5	285	0,79	1,40	2682	47,43	1788	47,43
9	305	0,85	1,50	2871	50,76	1914	50,76
9,5	320	0,89	1,57	3012	53,26	2008	53,26
10	340	0,94	1,67	3200	56,59	2133	56,59

* bei optimalem Betrachtungsabstand für HD

Tabelle 2: Objektgeschwindigkeiten im Test

Für die Betrachtung der Testsequenzen stehen am *IRT* zwei identische und kalibrierte Plasmascreens bereit, für die jegliche Bildoptimierungsmodi deaktiviert wurden. Die mit 25 Vollbildern aufgenommene Testsequenz weist bereits bei der geringstmöglichen Objektgeschwindigkeit von $9,15^\circ/\text{s}$ sehr deutliche Stroboskopeffekte auf.

Im Vergleich dazu wurde die Sequenz in 720p/120 betrachtet, indem die 3D-Funktion des genutzten Plasmascreens zu Hilfe genommen und zwei 720p/60-Videostreams im Frame Packing Verfahren dargestellt wurden. Die Grenzggeschwindigkeit, bei der eine störende Unschärfe wahrgenommen werden konnte, beträgt etwa $20\text{--}30^\circ/\text{s}$. Höhere Geschwindigkeiten führten zu deutlich sichtbaren Doppelkonturen der einzelnen Segmente des Siemenssterns. Projiziert man das Ergebnis auf die reale Szenengestaltung würde dies bedeuten, dass bei einer Bildwechselfrequenz von 120 Hz ein Schwenk über die komplette Bildbreite in ein bis zwei Sekunden ohne deutlich wahrnehmbaren Auflösungsverlust möglich ist. Da der beschriebene Testaufbau einen solchen Schwenk allenfalls abstrahiert darstellen kann, müssen die Ergebnisse anhand gängiger Testsequenzen überprüft werden.

4 Testumgebung

Dieses Kapitel beschreibt die am *IRT* vorhandene Testumgebung und geht dabei auf die verwendeten Testsequenzen, die technischen Einrichtungen zur Wiedergabe sowie die Methoden zur subjektiven Beurteilung ein. Wie eingangs erwähnt, finden vor allem die Tests zur Codierung in Anlehnung an parallel stattfindende Untersuchungen durch die Arbeitsgruppe BTF der *EBU* statt.

4.1 Testsequenzen

Um neue Technologien in einem Umfeld zu testen, das die momentane Produktionslandschaft bestmöglich abbildet, besteht immer die Nachfrage nach Testsequenzen, die nicht unter Studiobedingungen aufgenommen wurden und auf besonders kritische Situationen abzielen, die eine neue Technologie sofort an ihre Grenzen bringen. Bedarf besteht vor allem an solchen Inhalten, die das Bild des täglichen Fernsehprogramms prägen. Allen voran stehen Sportprogramme, die aufgrund ihrer schnellen, dynamischen Bewegungsabläufe besonders stark von HFR profitieren können. Problematisch ist dabei die Rechtfertigung, da die Produktionsfirmen und Rundfunkanstalten diese Inhalte nicht ohne weiteres zur Verfügung stellen.

Die *EBU* hat deshalb gemeinsam mit der *Schweizer Radio- und Fernsehgesellschaft (SRG SSR)* während der European Athletics Championships im August 2014 in Zürich einige Testsequenzen mit verschiedenen Kameras aufgenommen, die insbesondere zur Qualitätsbewertung der Parameter für UHD-1 Phase 2, also HFR, HDR und WCG, dienen sollen. Mit einer *ARRI Amira* wurden HD-Sequenzen mit 100 und 200 Hz Bildwechselfrequenz aufgezeichnet, eine *Phantom Flex4K* von *VisionResearch* wurde zur Aufzeichnung von Sequenzen in 4K und 100, 120, 300 sowie 1000 Hz verwendet. Des Weiteren standen mit der *LDX XtremeSpeed* und einem Prototypen der *LDX WorldCam* zwei Kameras von *Grass Valley* zur Verfügung, die auf den Live-Betrieb ausgelegt sind. Während die *LDX XtremeSpeed* 150 Bilder pro Sekunde in HD aufzeichnen konnte, lag das Hauptaugenmerk bei dem Prototypen auf HDR in 1080p/50, was durch Implementierung einer alternativen Transferfunktion nach dem Standard der *Society of Motion Picture and Television Engineers (SMPTE) ST 2084* bewirkt wurde und für diese Arbeit weniger von Bedeutung ist. Das Rohmaterial wurde bei *ARRI* in München einer Farbkorrektur unterzogen, geschnitten und das 4K-Material der *Phantom Flex4K* der UHD-Bildbreite von 3840 Pixeln angepasst. Die Testsequenzen werden über die Webseite der *EBU Technology and Innovation* in mehreren Sets zur Verfügung gestellt. Das Public Test Set beinhaltet eine Auswahl von acht Clips mit je einer Sequenz aus jeder der möglichen Kombinationen aus örtlicher Auflösung, Bildwechselfrequenz und Dynamikumfang. *EBU*-Mitglieder erhalten Zugriff auf das komplette Test Set mit 86 Sequenzen. Diese liegen für die HDR-Sequenzen im Format OpenEXR vor,

einem HDR-Grafikformat mit unkomprimierten Bilddaten als 16-Bit-Gleitkommazahl pro Farbkanal. Die Gammakennlinie wurde für diese Sequenzen linearisiert, damit eine einwandfreie Wiedergabe durch selbstständige Auswahl der passenden elektro-optischen Transferfunktion (EOTF) ermöglicht wird. Die Aufnahmen in Standard Dynamic Range (SDR) stehen ebenfalls unkomprimiert im Digital Picture Exchange (DPX) Format, mit 10 Bit Quantisierung und 4:2:2 Farbunterabtastung zur Verfügung.

Um eine gewisse Vielfalt an Testmaterial sicherstellen zu können, werden weitere ausgewählte Sequenzen verwendet, die von verschiedenen Partnern innerhalb der *EBU* stammen. Von der *BBC* werden UHD-Sequenzen zur Verfügung gestellt, die von den Commonwealth Games 2014 in Glasgow stammen. Diese wurden dort mithilfe einer FT-One Kamera des Herstellers *For-A* in 400 Hz aufgezeichnet.³⁰

4EVER liefert Sequenzen mit diversen Sportinhalten wie Ballett, Basketball, Hip Hop, Reitsport und Rugby. Aufgezeichnet wurden diese im HD-Format unter Verwendung einer *Sony F55* mit Bildwechselfrequenzen von 100 und 200 Hz.

Aus dieser großen Menge an Sequenzen wurden von der BTF-HFR wiederum neun Sequenzen in die engere Wahl genommen und letztlich fünf Sequenzen selektiert, die für Codingtests verwendet werden. Für diese Arbeit wurden aus den neun Testsequenzen aus der Vorauswahl ebenfalls fünf Sequenzen für subjektive Tests ausgewählt. Die Auswahl unterscheidet sich dabei etwas von der *EBU*-Gruppe. Tab. 3 gibt Auskunft über die Herkunft und die technischen Parameter der finalen Auswahl. Deren aufnahmeseitige Bildrate liegt bei 200 und 400 Hz, woraus durch Downsamplingprozesse niedrigere Vergleichsbildraten erzeugt werden können (vgl. 4.2.1).

Da am Ende der Codingtests Aussagen zu künftig möglichen Ausstrahlungsparametern getroffen werden sollen, wird einheitlich eine 4:2:0 Farbunterabtastung vorgenommen. Aus Gründen der Verfügbarkeit an Displays, einigte man sich auf die Verwendung der örtlichen Auflösung für HDTV (vgl. 4.2.3). Für Sequenzen, die in UHD-Auflösung vorliegen, kommt ein Skalierungsalgorithmus zur Anwendung, der die örtliche Auflösung auf das HD-Format mit 1080 Zeilen bringt. Bei der Auswahl der Sequenzen wurde darauf geachtet, dass eine Länge von zwölf Sekunden je Sequenz bei Echtzeitwiedergabe nicht deutlich über- oder unterschritten wird. Diese Länge entspricht ungefähr der empfohlenen Länge für subjektive Bewertungsmethoden, die zehn Sekunden beträgt (vgl. 4.3.1 und 4.3.2). Gemäß diesen Empfehlungen wird auch von jeder Sequenz eine Version in 1080i/25 erzeugt, die dann als sogenannter Low Anchor dient, also deutlich geringere Qualität aufweist, als die restlichen Vergleichssequenzen und damit eine im Vergleichstest klar identifizierbare untere Grenze

³⁰ vgl. Pindoria, Manish (2014): Higher frame rate television for future broadcasts at the Commonwealth Games - "But I already have a 100Hz TV... ?". Hg. v. BBC Research & Development. Online verfügbar unter <http://www.bbc.co.uk/rd/blog/2014-08-higher-frame-rate-television-for-future-broadcasts-at-the-commonwealth-games-but-i-already-have-a-100hz-tv>, zuletzt aktualisiert am 12.08.2015, zuletzt geprüft am 30.09.2015

bildet. Der Low Anchor wurde mithilfe einer Software zur Erzeugung von Interlaced-Sequenzen aus progressivem Material erstellt.

Sequenz	örtliche Auflösung	Bildrate	Länge	Herkunft
Hürdenlauf	3480x2160	400 Hz	14 s	BBC
Stufenbarren	3480x2160	400 Hz	13 s	BBC
Dreisprung	1920x1080	200 Hz	10 s	EBU
Reitsport	1920x1080	200 Hz	12 s	4EVER
5000m-Lauf	3480x2160	200 Hz	12 s	BBC

Tabelle 3: Technische Daten der verwendeten Testsequenzen

Um auch einen bildlichen Gesamteindruck der verwendeten Testsequenzen zu vermitteln, sind in Abb. 8 Screenshots der fünf Sequenzen zu sehen.

Die Sequenz „Hürdenlauf“ wurde von einer höher positionierten Kamera mit Teleobjektiv aufgezeichnet. Zu Beginn der Sequenz ist im sehr detailreichen Hintergrund eine Tribüne mit Publikum sichtbar. Im weiteren Verlauf erfolgt ein langsamer vertikaler Schwenk nach unten, während die Sportler über die aufgestellten Hürden springen.

Die nächste Sequenz zeigt eine Kür am Stufenbarren, wobei die Kamera den Bewegungen der Turnerin nur leicht folgt und ansonsten recht statisch bleibt. Das Augenmerk liegt auf der gleichmäßigen Kreisbewegung sowie der schnellen Drehbewegung beim Abgang. Der Hintergrund bleibt die ganze Zeit über sehr dunkel und zieht die Aufmerksamkeit daher nicht so sehr auf sich, wie beispielsweise in der Sequenz „Hürdenlauf“.

In der Sequenz „Dreisprung“ sind zwei Sportlerinnen zu sehen, die beim Anlauf und letztlich beim Sprung in die Sandgrube aufsichtig und mittels eines horizontalen Kameraschwenks von rechts nach links verfolgt werden. Während des Anlaufs passieren sie unter anderem einen Tisch, an dem einige Wertungsrichter sitzen. Der Hintergrund besteht aus der Rasenfläche des Stadions.

Die vierte verwendete Sequenz wurde während einer Reitsportveranstaltung aufgezeichnet. Die Kamera befindet sich in Normalsicht und schwenkt horizontal von rechts nach links mit, während die Reiterin mit ihrem Pferd einige Hindernisse passiert. Dabei variiert die Schwenkgeschwindigkeit. Da die Szene in einem Reitstadion stattfindet, sind im Hintergrund während der gesamten Sequenz gut gefüllte Zuschauerränge zu sehen.

Zuschauerränge dominieren auch das Bild der Sequenz „5000m-Lauf“. Der Kameraschwenk folgt den Athleten hierbei jedoch mit relativ gleichmäßiger Geschwindigkeit von rechts nach links und die Perspektive ist aufsichtig. Ein weiteres Merkmal im Hintergrund sind die prägnanten Werbeschriftzüge unterhalb der Tribünen.

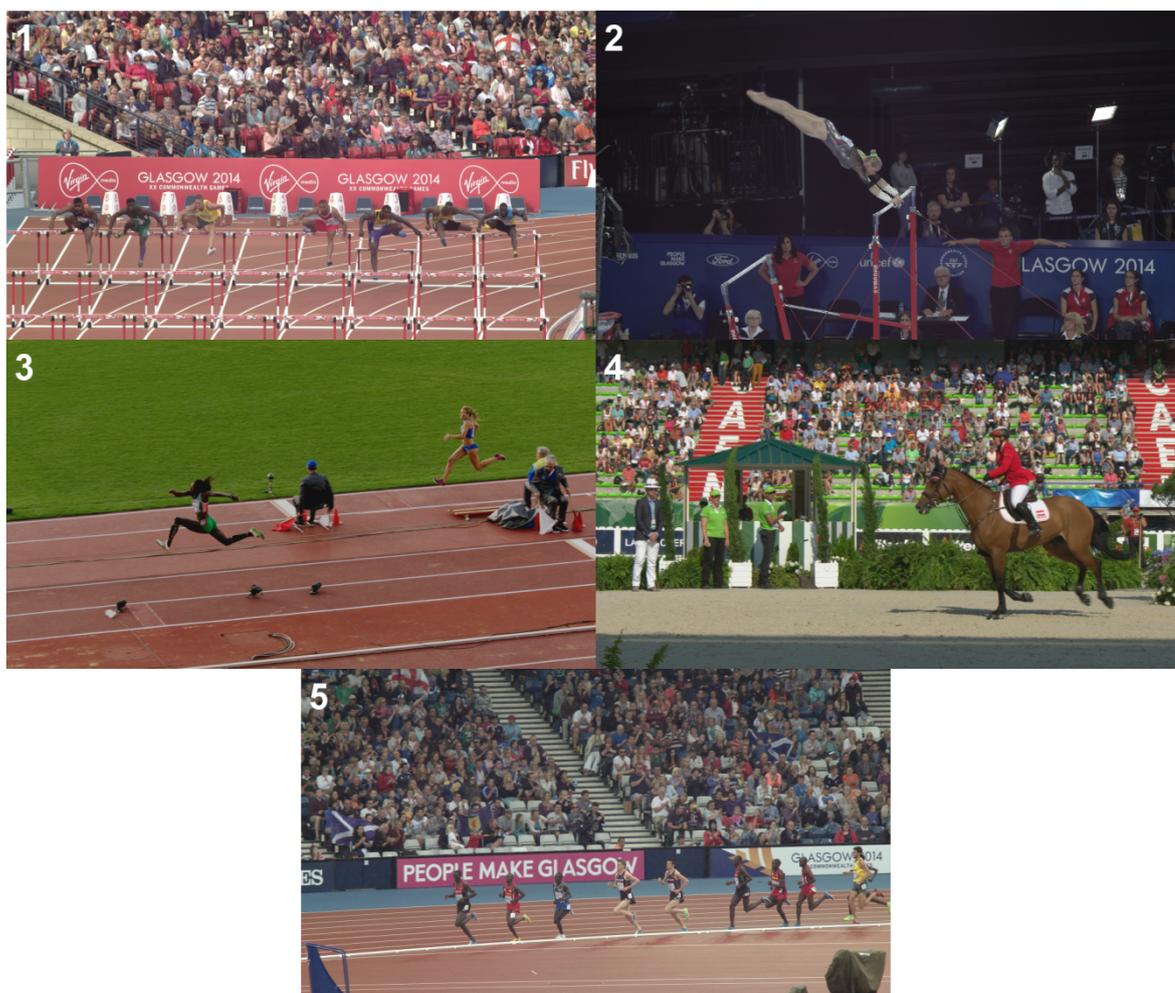


Abbildung 8: Screenshots der verwendeten Testsequenzen. 1 Hürdenlauf, 2 Stufenbarren, 3 Dreisprung, 4 Reitsport, 5 5000m-Lauf

Um bei der Auswahl der Testsequenzen neben der rein subjektiven Betrachtung auch über technisch vergleichbare objektive Auswahlkriterien zu verfügen, wurden für alle Sequenzen die Spatial Perceptual Information (SI) und die Temporal Perceptual Information (TI) ermittelt. Diese beiden Indizes geben in Form eines Zahlenwerts eine objektive Auskunft über die örtlichen Details innerhalb eines Bildes sowie die zeitlichen Differenzen zwischen mehreren Bildern. Die Berechnung der SI basiert auf dem Sobel-Operator, einem Hochpassfilter zur Kantendetektion. Die TI wird über Unterschiede der Luminanzwerte von Pixeln an gleicher Stelle innerhalb eines Bildes aber über den Zeitraum mehrerer Bilder ermittelt. Das Hinzuziehen dieser objektiven Parameter bei der Auswahl von Testmaterial wird für die in Kapitel 4.3.2 genannte Testmethode empfohlen und in Rec. ITU-T P.910 näher spezifiziert.³¹ Die SI und TI der für die Codingtests ausgewählten Testsequenzen sind in Abb. 9 dargestellt.

³¹ vgl. International Telecommunication Union (2008): Subjective video quality assessment methods for multimedia applications. Recommendation ITU-T P.919. Hg. v. Telecommunication Standardization Sector of ITU. International Telecommunication Union. Genf, S. 4f.

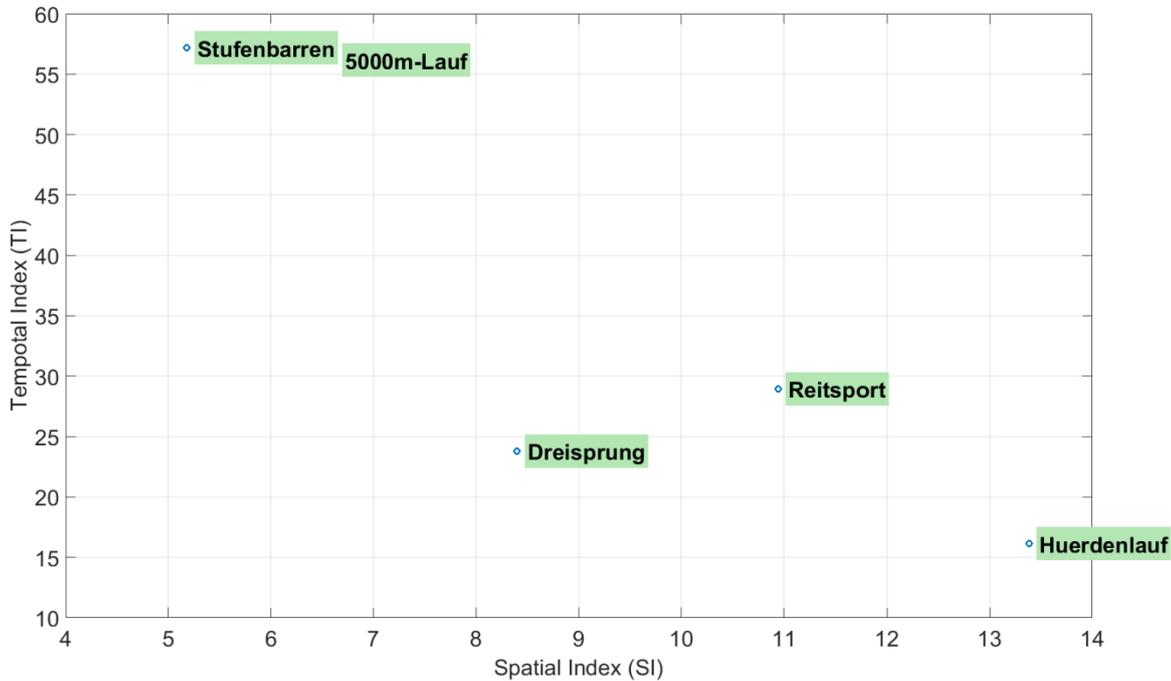


Abbildung 9: SI und TI der ausgewählten Testsequenzen

4.2 Aufbereitung und Wiedergabe

Um die Sequenzen in HFR optimal mit niedrigeren Bildwechselfrequenzen und Shuttereinstellungen vergleichen zu können, müssen sie einem Downsampling-Prozess unterzogen werden. Dieser wird in diesem Kapitel dargestellt. Umgekehrt müssen die Auswirkungen einer Berechnung von Zwischenbildern untersucht werden, einer Technik, die heutige Displays schon einsetzen, um Bewegungsartefakte auf der Wiedergabeseite zu reduzieren. Wie diese Berechnung simuliert wird, beschreibt dieses Kapitel. Außerdem wird erklärt, wie durch Modifikation des Playout-Servers die Wiedergabe mit Bildwechselfrequenzen bis 200 Hz ermöglicht wird.

4.2.1 Downsampling

Seitens der *BBC* wird für das Downsampling von HFR-Sequenzen eine bestimmte Vorgehensweise empfohlen. Diese beinhaltet zuerst eine standardkonforme Farbraumkonvertierung ausgehend vom Y'CbCr-Signal mit einer 4:2:2 Unterabtastung hin zu R'G'B' mit voller Farbauflösung in 4:4:4. Die fehlenden Samples werden mithilfe eines digitalen Filters interpoliert. Anschließend findet eine Linearisierung statt, indem die Umkehrung der Gammafunktion des Eingangssignals angewendet wird. Das Resultat ist folglich ein lineares R'G'B'-Signal mit voller Farbauflösung.

Erst danach findet das eigentliche Downsampling der dynamischen Auflösung statt. Dabei gibt es zwei mögliche Anpassungen, welche in Abb. 10 und Abb. 11 anhand eines Downsampling von 200 Hz auf 50 Hz dargestellt sind. Soll lediglich die Bildwechselfrequenz

verändert und der Shutter beibehalten werden, ist es notwendig, aus den binären Werten mehrerer aufeinanderfolgender Einzelbilder einen Durchschnittswert zu ermitteln. Dieser Vorgang wird auch als Image Averaging bezeichnet. Der Shutter des resultierenden Bildes verhält sich in diesem Fall wie die Summe der Shutterwerte der Quellbilder.

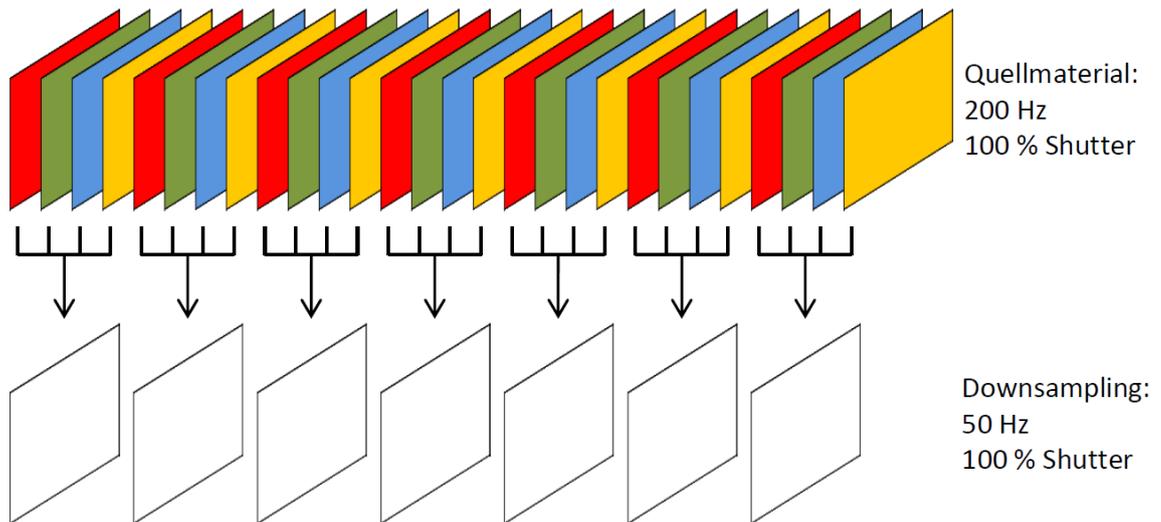


Abbildung 10: Downsampling durch Image Averaging

Werden einzelne Bilder der Quellsequenz entfernt, bleiben die Shutter gleich, was bei weniger Bildern pro Sekunde in einem kürzeren Shutter bezogen auf die komplette Sequenz resultiert. Bei diesem sogenannten Frame Dropping wird im Gegensatz zum Image Averaging nahezu keine Rechenleistung benötigt, da lediglich Einzelbilder entfernt werden.

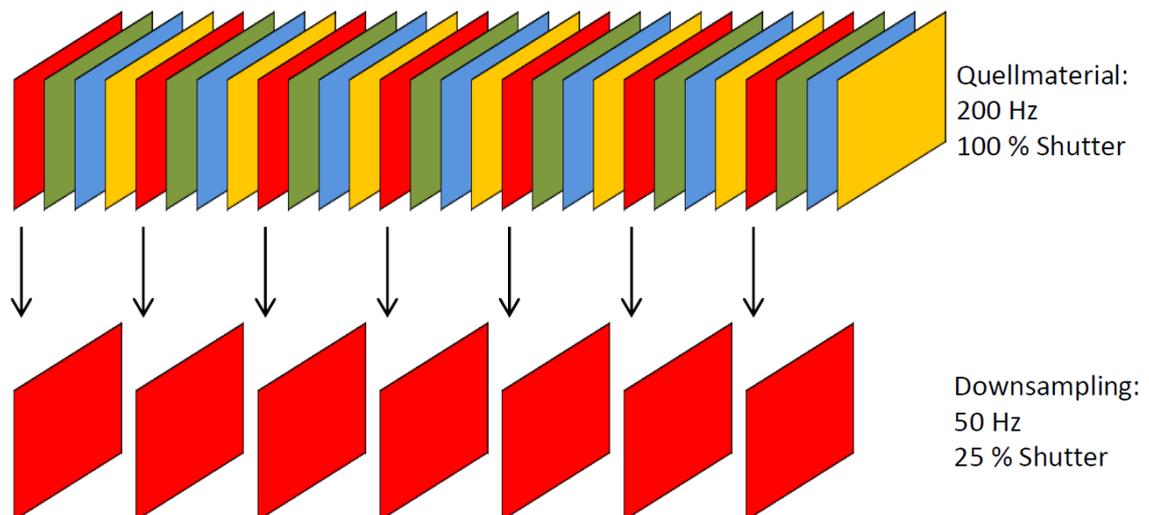


Abbildung 11: Downsampling durch Frame Dropping

Am *IRT* steht zur Aufbereitung der Testsequenzen eine Software der *BBC* zur Verfügung, welche die o. g. Prozesse automatisiert ausführt, nachdem Bildwechselfrequenz und Shutter am Eingang und die gewünschten Parameter am Ausgang definiert worden sind. In der ersten Version der Software, die der Autor für die Tests nutzt, werden als Eingangsformate

entweder v210 (Y'CbCr, 10 Bit, 4:2:2) oder uyvy (Y'CbCr, 8 Bit, 4:2:2) akzeptiert, wobei alle Bilddaten in eine einzelne Datei gepackt und sämtliche Headerinformationen entfernt sein müssen, da diese sonst fälschlicherweise vom Programm als Bilddaten interpretiert werden würden. Am Ausgang wird immer ein UYVY-File erzeugt. Beispielhaft für eine Konvertierung von Material in 200 Hz zu 50 Hz wird nachfolgend der dafür nötige Kommandozeilen-ausdruck mit den wesentlichsten Parametern aufgeführt.

```
hfrDownsampleCombined -f 2538 -x 1920 -y 1080 -o 50@50 -t 50 [Eingabedatei].uyvy 200 200 [Ausgabedatei] uyvy
```

Der Parameter `-f` beschreibt dabei die Anzahl der Bilder der Eingabedatei. Die örtliche Auflösung wird mit `-x` und `-y` definiert, bevor mit `-o` die gewünschte Bildwechselfrequenz und der Shutter der Ausgabedatei festgelegt werden. Die Frequenz, mit der das Ausgabe-File wiedergegeben werden soll, lässt sich mit `-t` definieren. Es folgt der Pfad der Eingabedatei sowie deren Bildfrequenz und Shutter bevor Ausgabepfad und `-format` festgelegt werden.

4.2.2 Interpolation

Eine Berechnung von Zwischenbildern, auch Bewegungsinterpolation genannt, erfolgt bereits seit geraumer Zeit in den heute existierenden Consumerdisplays. Die Qualität der Interpolation schwankt dabei sehr stark und unterscheidet sich oftmals deutlich von der hardwarebasierten Frameratekonvertierung, wie sie im Live-Broadcastbereich zur Anwendung kommt. Da die verwendeten Displays keine interne Interpolation, also die automatische Berechnung von Zwischenbildern, unterstützen, muss hierbei auf spezielle Software zurückgegriffen werden, die diese Berechnung beherrscht. Die Tests sollen möglichst nah an eine zukünftige Wohnzimmersituation herankommen, weshalb bewusst auf die Verwendung eines hardwarebasierten Konverters verzichtet und die Softwarelösung Alchemist On Demand von *Snell* (seit September 2015: *Snell Advanced Media*) gewählt wurde. Das Qualitätsniveau der Interpolation sollte ungefähr dem einer in zukünftigen Displays implementierten Interpolation entsprechen.

Die Software unterstützt unter anderem die Konvertierung von unkomprimiertem Videomaterial in 1080p/25 nach 1080p/50, wobei die Methode der Zwischenbildberechnung wählbar ist. Die einfachste Möglichkeit besteht in der Verdopplung aller Bilder, was der Darstellung von Bewegungen nicht zu Gute kommt und unter Umständen sogar Stroboskopeffekte bewirkt. Bessere Ergebnisse bringt eine lineare Interpolation, wobei neue Pixelwerte aus den vorhandenen Werten der horizontal und vertikal benachbarten Pixel ermittelt werden. Im Test wurde dagegen eine bewegungskompensierende Interpolation (Motion-Compensated Frame Interpolation – MCFI) verwendet. Diese nutzt die Phasenverschiebung der Frequenz im Bezug auf die Zeit als Bewegungsindikator. Nach einer Fouriertransformation erfolgt ein Phasenabgleich der Schwingungen aufeinanderfolgender Bilder, dies ist gleichzusetzen mit einer Analyse der Bewegung. Die Phasenverschiebung wird durch Subtraktion der Spektren beider Bilder bestimmt, bevor wiederum durch eine inverse Fouriertransformation eine

Korrelationsoberfläche entsteht. Anhand von Veränderungen dieser Oberfläche können Bewegungen erkannt und in der Zwischenbildberechnung kompensiert werden. Die Interpolation durch Phasenkorrelation, wie diese Methode auch genannt wird, gilt als die effektivste im Vergleich zu anderen Techniken.³²

Dieses Feature kann für die Zwecke dieser Arbeit genutzt werden, da es der Software egal ist, in welcher Bildfrequenz das Quellmaterial vorliegt, solange in den Metadaten 25 Bilder pro Sekunde erkannt werden. Die Anpassung dieses sich im Header des Videocontainer befindlichen Metadatum kann mit dem am *IRT* verwendeten Videosever vorgenommen werden.

Im Vorfeld der eigentlichen Tests wurde bereits ein Vergleich zwischen einem aktuell auf dem Markt befindlichen Consumerdisplay mit Flüssigkristallanzeige (Liquid Crystal Display – LCD), das eine zuschaltbare Funktion zur Verbesserung der Bewegungsdarstellung ausweist, sowie dem HFR-Display gezogen. Es stellte sich heraus, dass das Consumergerät durch interne Zwischenbildberechnung und verschiedenartige Modulation der LED-Hintergrundbeleuchtung die Darstellungsdauer von Einzelbildern verkürzt und damit eine subjektiv bessere Bewegungsauflösung vermittelt. Gleichmäßige Schwenks etwa weisen dadurch eine wahrnehmbar geringere Bewegungsunschärfe auf.

4.2.3 Videosever

Das standardmäßige Playout-System für Videoinhalte am *IRT* ist die Mastering-Plattform Clipster von *Rohde & Schwarz DVS*. Die interne Verarbeitung erfolgt in einzelnen Files für jedes Bild einer Sequenz. Auf zwei Videosevern werden alle für die Vergleichstests relevanten Testsequenzen vorbereitet und wiedergabebereit abgelegt. Für die Weitergabe an die o. g. *BBC*-Software, die das Downsampling übernimmt, müssen die Einzelbildsequenzen zuerst in eine Datei gepackt werden. Dies funktioniert mittels der *IRT*-eigenen Software *DpxConvert*, die aus Clipster exportierte DPX-, oder YUV-Sequenzen in einen Audio Video Interleave (AVI) Container schreibt. Das Entfernen der Headerdaten erfolgt dann mittels der Open-Source-Software *FFmpeg*. Die dazu benötigte Kommandozeile lautet

```
ffmpeg -i [Eingabedatei].avi -f rawvideo -pix_fmt uyvy422 [Ausgabedatei].uyvy
```

Umgekehrt ist mit *DpxConvert* auch das Erzeugen von DPX-Sequenzen aus einzelnen gepackten YUV-Files möglich, was für die Wiedergabe von Files aus der *BBC*-Software über Clipster notwendig ist.

³² vgl. Sauter, Martin (2008): Einfluss von Bildformat- und Bildwechselfrequenzkonvertierung auf die Qualität von HDTV-Produktionen. Diplomarbeit. Fachhochschule Wiesbaden, Wiesbaden. Informationstechnologie und Elektrotechnik. S. 23 ff.

Eine unkomprimierte Übertragung von HD-Material in 200 Hz benötigt eine Datenrate von ca. 12 Gbit/s. Die Ausgabe erfolgt über vier Serial Digital Interface (SDI) Schnittstellen mit jeweils 3 Gbit/s Datendurchsatz. Da Clipster standardmäßig zwar eine UHD-Ausgabe in Form von vier HD-Quadranten, jedoch keine native HFR-Ausgabe unterstützt, müssen vorher diverse Anpassungen der Sequenzen vorgenommen werden. Eine am *IRT* programmierte Software ermöglicht das Mapping von vier aufeinanderfolgenden Einzelbildern einer HFR-Sequenz in die vier Quadranten eines UHD-Signals, sodass die 200 Hz nativ von Clipster ausgespielt werden können.

4.2.4 Displays

Für die geplante Durchführung der subjektiven Tests wurden bereits 2014 mehrere Displays in Auftrag gegeben, welche eine Wiedergabe von bis zu 240 Bildern pro Sekunde erlauben sollen. Das Berliner Unternehmen *MikroM* kooperierte für dieses Projekt mit dem TV-Gerätehersteller *Loewe* und modifizierte ein entsprechendes Display mit zusätzlichen Baugruppen. Als örtliche Auflösung wurde vorerst HD geordert, um den Implementierungsaufwand und den damit verbundenen Realisierungszeitraum möglichst gering zu halten. Mitte Juni 2015 wurde ein erster Prototyp ans *IRT* ausgeliefert, Ende August folgte dann ein serienreifes zweites Gerät.

Die Displays werden über vier 3G-SDI-Eingänge bespielt und mittels Kommandozeilentools auf dem integrierten Linux-System konfiguriert und gesteuert. In seinem Arbeitsbereich von etwa 192 bis 240 Hz stellt das Display die zugeführten Bilder nativ dar.³³ Wird eine geringere Bildfrequenz zugespielt, erfolgt intern eine automatische Bildwiederholung. Bei einer Wiedergabe von Quellen in z. B. 100 Hz wird jedes Bild doppelt, bei 50 Hz vier Mal und bei 25 Hz sogar acht Mal wiedergegeben. Damit bei subjektiven Vergleichstests verschiedener Bildfrequenzen nicht permanent eine Neukonfiguration des Displays nötig ist, werden die Bilder der Sequenzen unter 200 Hz im Vorfeld mithilfe der *BBC*-Software gedoppelt bzw. vervierfacht.

Die beiden Displays unterscheiden sich in ihrem Design, da das Panel des Prototyps zusätzlich durch eine Glasfront bedeckt wird. Diese Front beeinflusst die maximale Helligkeit, weshalb eine Messreihe (vgl. Abb. 12) durchgeführt und die Hintergrundbeleuchtung des zweiten Displays in den entsprechenden Tests mittels o. g. Kommandozeilentools angepasst wird.

³³ vgl. MikroM GmbH (Hg.) (2015): Kolibri HFR Display. User Manual. Berlin. S. 13

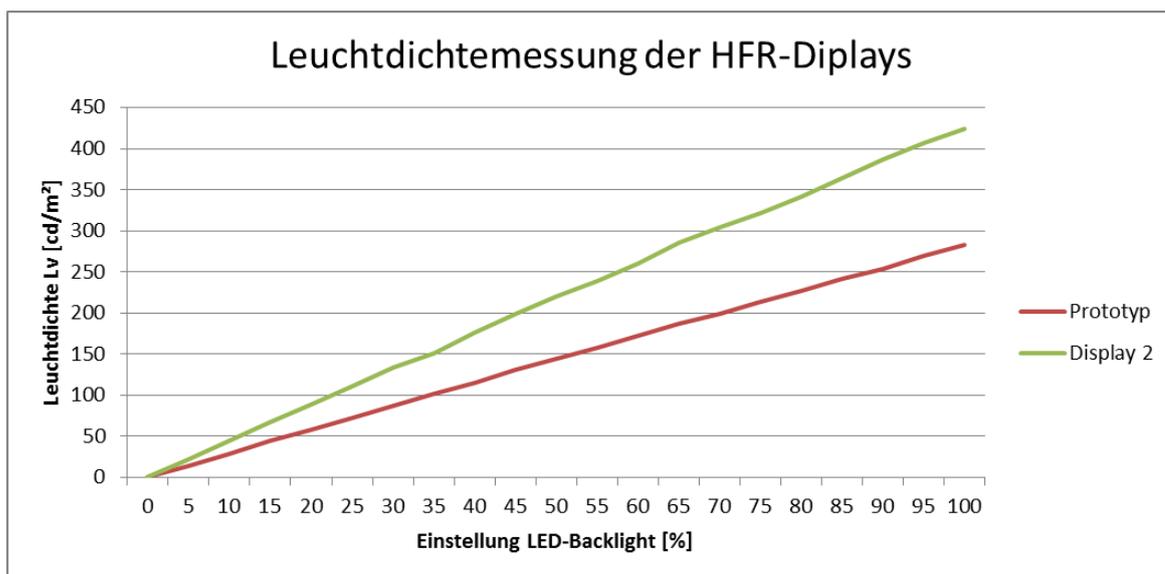


Abbildung 12: Messreihe zur Leuchtdichte der HFR-Displays in Abhängigkeit der LED-Backlight-Einstellung

Als Räumlichkeit für die Tests ist das TV-Studio des *IRT* prädestiniert, das gemäß den Empfehlungen von ITU-R BT.500 ausgestattet und eingemessen wurde. Die Testumgebung genügt den dort genannten Anforderungen an eine Laborumgebung

4.3 Bewertungsmethoden

Die Methoden zur subjektiven Bewertung von Wiedergabesystemen sind in bestimmten Standards festgelegt, um eine Vergleichbarkeit der Ergebnisse sicherzustellen. Die *EBU* bewertet in ihren Codingtests für gewöhnlich nach den Methoden Double-Stimulus Continuous Quality-Scale (DSCQS) oder Subjective Assessment of Multimedia Video Quality³⁴ (SAMVIQ), für die jeweils mindestens 15 Teilnehmer empfohlen werden. Für die subjektiven Tests in dieser Arbeit werden durch die DSCQS-Methode bestimmte Kombinationen aus Bildwiederholrate, Shutter und Interpolation ermittelt, die in den nachfolgenden zweiten Test, der auch die Codierung beinhaltet, einbezogen werden. Letzterer findet in Form von Expert Viewings statt, die einer Bewertung durch eine kleinere Gruppe bedürfen. Die Methodik der Expert Viewings wird ebenso wie die o. g. subjektiven Vergleichsmethoden durch die *ITU* standardisiert. Nachfolgend werden die Rahmenbedingungen, Durchführungsprozeduren und Auswertungsschritte der hier verwendeten Methoden näher beschrieben.

³⁴ vgl. International Telecommunication Union (2007): Methodology for the subjective assessment of video quality in multimedia applications. Recommendation ITU-R BT.1788. Hg. V. Radiocommunication Sector of ITU. International Telecommunication Union. Genf.

4.3.1 DSCQS

Diese Methode wird in Rec. ITU-R BT.500 definiert und basiert auf einem Schema, nach dem jedes Testsequenzenpaar, bestehend aus der Referenz und der Vergleichssequenz den Teilnehmern zweimal in Folge vorgespielt wird. Die Testteilnehmer sollen vor den Vergleichstests im Hinblick auf ihre Sehschärfe und ihr Farbsehvermögen mithilfe gängiger Testcharts geprüft werden. Außerdem sollen ihr Alter, Geschlecht und das für den Test relevante und gegebenenfalls vorhandene Expertenwissen dokumentiert werden.

Für die eigentlichen Testsessions sind zwei Varianten möglich. Die erste Variante berücksichtigt eine Testperson pro Session, während bei der zweiten Variante mehrere Personen zugleich an einer Session teilnehmen. Letztere Variante wird in dieser Arbeit bevorzugt betrachtet. Eine Testsession soll maximal bis zu 30 Minuten dauern. In diesem Zeitraum werden die Teilnehmer vorerst mit einigen Sequenzen auf den Test vorbereitet und lernen den generellen Ablauf sowie die Testmethode kennen.

Zu Beginn der ersten Session werden etwa fünf Dummy-Sequenzen eingefügt, deren Wertungen nicht in das Gesamtergebnis einfließen dürfen. Sie dienen ausschließlich dazu, den Testpersonen einen konstanten Bewertungsmaßstab anzutrainieren. Finden mehrere Sessions statt, genügen zu Beginn der folgenden Durchläufe etwa drei Dummy-Sequenzen. Im Anschluss werden die für das Testergebnis relevanten Sequenzen präsentiert, wobei die Reihenfolge pseudo-zufällig erfolgt und sich von Session zu Session ändern kann. Im weiteren Verlauf erscheinen die Dummy-Sequenzen noch einmal, wobei deren Bewertung diesmal einbezogen wird. Des Weiteren sollen mögliche Faktoren, wie etwa Gewöhnung oder Müdigkeit der Testteilnehmer, ausgeschlossen werden. Wie oben erwähnt wird jedes Paar aus Referenz und Vergleichssequenz den Teilnehmern zwei Mal vorgestellt, damit zuerst ein Eindruck gewonnen und anschließend die Bewertung vorgenommen werden kann. Eine Länge von zehn Sekunden je Sequenz gilt als Minimum. Zwischen den einzelnen Sequenzen wird ein Graubild eingefügt, das durch einen Videopegel von etwa 200 mV erzeugt wird. Die Testpersonen bewerten die Qualität der ihnen gezeigten Testsequenzen auf einer stufenlosen vertikalen Skala, die zur Orientierung in fünf gleichgroße Segmente eingeteilt ist, welche die in Abb. 13 dargestellten, allgemeinen Qualitätsstufen repräsentieren.

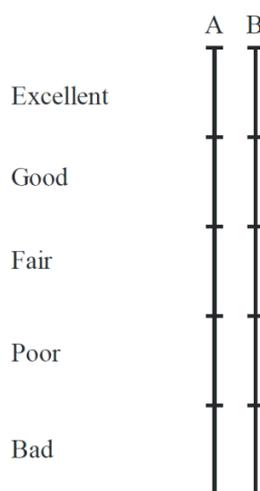


Abbildung 13: Bewertungsskala für DSCQS³⁵

Für die Auswertung gilt, dass bei der Qualitätsbewertung immer nur das aktuell wiedergegebene Sequenzenpaar in Relation gesetzt wird. In Rec. ITU-R BT.500 werden auch die Methoden der anschließenden Datenanalyse und -auswertung genau spezifiziert. Die Eintragungen der Testteilnehmer in die stufenlose Skala werden dabei in Messwerte zwischen null und einhundert übertragen und die Differenz gebildet. Neben den Mittelwerten aller Teilnehmer für jede getestete Sequenz wird auch ein sogenanntes Confidence Interval von 95 % jedes Messwerts im Ergebnis angegeben.³⁶

4.3.2 Expert Viewing

Die Rec. ITU-R BT.1663 definiert die Voraussetzungen und Durchführungsbestimmungen von sogenannten Expert Viewings zur Qualitätsbeurteilung digital wiedergegebener, großflächiger Bewegtbildinhalte. Letztere werden im Standard als solche Inhalte beschrieben, die in Kinos oder Veranstaltungsräumen präsentiert werden. Die Methode an sich ist aber ohne weiteres auch auf Videoinhalte anwendbar, die auf kleineren Wiedergabesystemen von Fernsehzuschauern angesehen werden. Expert Viewings werden insbesondere empfohlen, wenn der absehbare Zeit- und Kostenaufwand zu groß wird, um gemäß DSCQS oder anderer Methoden zu bewerten.³⁷

Die zu testenden Sequenzen werden von fünf bis sechs Personen, die Erfahrung mit subjektiver Beurteilung solcher Testsysteme aufweisen aber nicht direkt an der Entwicklung

³⁵ International Telecommunication Union 2012, S. 15

³⁶ vgl. International Telecommunication Union 2012, S. 13 ff.

³⁷ vgl. International Telecommunication Union (2003): Expert viewing methods to assess the quality of systems for the digital display of large screen digital imagery in theatres. Recommendation ITU-R BT.1663. Hg. v. Radiocommunication Sector of ITU. International Telecommunication Union. Genf, S. 2 ff.

ihrer Technologie beteiligt sind, bewertet. Dabei werden die Referenz und die entsprechende Vergleichssequenz paarweise wiedergegeben und in der Phase 1 des Tests zwei Mal wiederholt. Die Reihenfolge der Sequenzen wird zufällig bestimmt und eine Diskussion unter den Experten ist ausdrücklich erlaubt. Das Gesehene wird anhand eines festgelegten Wortschatzes bewertet, wobei die Deutlichkeit des Unterschieds zwischen den beiden Vergleichsbildern herausgestellt wird. Am *IRT* haben sich dafür die Begrifflichkeiten aus der fünfstufigen „*EBU*-Skala“ bewährt. Ins Deutsche übersetzt bedeuten diese so viel wie „nicht wahrnehmbar“, „gerade wahrnehmbar“, „wahrnehmbar“, „leicht störend wahrnehmbar“ und „deutlich störend wahrnehmbar“.³⁸ In Phase 2 des Tests werden die Sequenzen nach Inhalt sortiert und das Ranking verfeinert. Hierbei dürfen einzelne oder alle Sequenzen beliebig oft wiederholt werden, bis alle Experten mit der Bewertung konform gehen.

³⁸ vgl. International Telecommunication Union 2012, S. 11

5 Tests zur zeitlichen Auflösung

In diesem Kapitel sollen zuerst die verschiedenen Möglichkeiten der Erzeugung und Wiedergabe von HFR-Material betrachtet und bewertet werden. Ziel ist es herauszufinden, wie hoch die Bildwechselfrequenz für eine zukünftige UHD Phase 2 sein sollte, damit ein deutlich sichtbarer Mehrwert erkennbar ist. Dabei soll auch geklärt werden, ob es notwendig ist, diese Bildfrequenz nativ von der Produktion bis zur Wiedergabe aufrecht zu erhalten oder ob es gegebenenfalls ausreicht, produktionsseitig Abstriche zu machen und erst bei der Wiedergabe die benötigten zusätzlichen Bilder künstlich zu erzeugen. Auch die Rolle der bei der Aufnahme eingestellten Belichtungszeit in Form des Shutter wird genauer untersucht.

Eckdaten zum subjektiven Test am IRT

Im September 2015 werden die subjektiven Tests mit unkomprimierten HFR-Sequenzen an insgesamt fünf Tagen am *IRT* durchgeführt. Dabei werden die fünf in Kapitel 4.1 genannten Sequenzen in 18 verschiedenen Varianten, die in den nachfolgenden Kapiteln 5.1 und 5.2 näher erläutert werden, den Probanden vorgestellt. Die etwa 90 Sequenzen werden dabei auf drei halbstündige Sessions zu jeweils 30 Sequenzen aufgeteilt, wobei aus organisatorischen Gründen womöglich nicht alle Probanden an allen drei Sessions teilnehmen können. Daher müssen mehr Teilnehmer eingeplant werden, deren Bewertungen dann zusammengefasst werden können. Ziel ist es, etwa 20 Bewertungen für alle drei Sessions aufzunehmen, da gemäß dem *ITU*-Standard zu DSCQS mindestens 15 valide Testpersonen benötigt werden und besonders deutlich abweichende Ergebnisse nicht in das Gesamtergebnis einfließen dürfen.

Vor der ersten Teilnahme einer jeden Testperson wird diese mittels einer Snellen Testtafel auf ihr Sehvermögen geprüft. Anschließend finden sich die Probanden im TV-Studio ein, wobei maximal sieben Personen an einer Session teilnehmen können. Die Anordnung der Sitzplätze erfolgt in zwei Reihen. Die erste Reihe, bestehend aus vier Plätzen, befindet sich im Abstand der dreifachen Bildschirmhöhe vor dem Display, was dem optimalen Betrachtungsabstand für HDTV entspricht. Die zweite Reihe aus drei Sitzplätzen befindet sich dahinter und weist einen ungefähren Betrachtungsabstand von 2,6 m zum Display auf. Diese Entfernung repräsentiert den durchschnittlichen Abstand zum Fernsehgerät zu Hause und geht auf empirische Erhebungen zurück, die 2015 von der *BBC* durchgeführt wurden.³⁹

³⁹ vgl. Noland, Katy C.; Truong, Louise H. (2015): A Survey of UK Television Viewing Conditions. Research & Development White Paper WHP 287. British Broadcasting Corporation. London, S. 38. Online verfügbar unter <http://downloads.bbc.co.uk/rd/pubs/whp/whp-pdf-files/WHP287.pdf>, zuletzt geprüft am 11.09.2015

Der Vergleich mit einer *BBC*-Studie von 2004 bringt die Erkenntnis, dass sich der Betrachtungsabstand in den letzten zehn Jahren trotz größer werdender Displays um 7 cm verringert hat.⁴⁰ Um eine uneingeschränkte Sicht zu gewährleisten, sitzen die Probanden in der zweiten Reihe erhöht. Das Display befindet sich vor einer grauen Wand und der Raum ist bis auf eine diffuse Beleuchtung mit ca. 6500 K abgedunkelt. Die räumlichen Gegebenheiten sind in der Abb. 14 dargestellt.

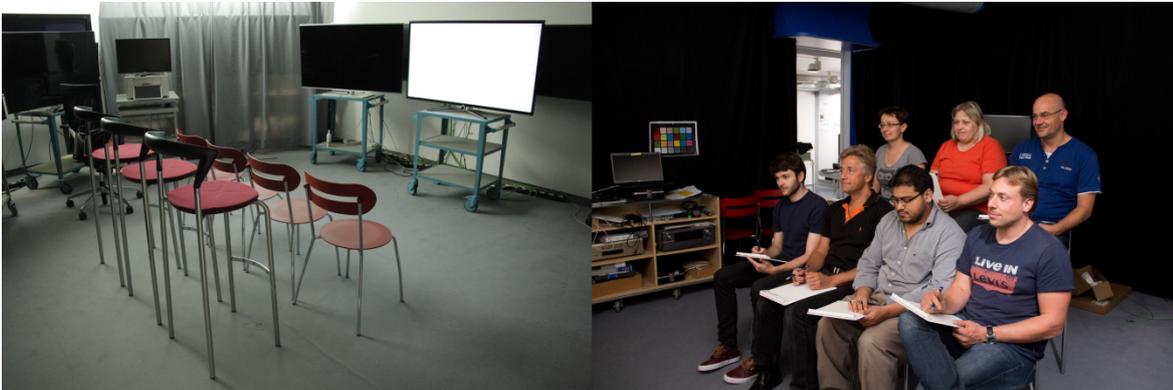


Abbildung 14: Räumliche Gegebenheiten beim subjektiven Test am IRT

Damit eventuelle Abweichungen in der späteren Auswertung nachvollzogen werden können, wird neben dem Namen, dem Datum, der Uhrzeit und der Testgruppe auch die Sitzposition auf dem Bewertungsbogen (vgl. Anhang 1) dokumentiert. Der generelle Ablauf und die Bewertungsmethodik werden den Teilnehmern unmittelbar vor der Session detailliert erläutert und auftretende Fragen beantwortet. Anhand der Sequenzen „Speerwurf“ und „Netball“, die nicht zu den finalen fünf Testsequenzen gehören, wird die Bewertungsmethode vor Beginn der eigentlichen Session trainiert. Dabei wird „Speerwurf“ in Form eines Vergleichs von 50 Hz und 100 % Shutter zur Referenz in 200 Hz und 100 % Shutter vorgestellt, um den Probanden die sichtbaren Unterschiede in der Bewegungsdarstellung aufzuzeigen, die bei dieser Sequenz in einer deutlich wahrnehmbaren Hintergrundunschärfe während eines horizontalen Schwenks ersichtlich sind. Im Gegensatz dazu zeigt sich bei der Sequenz „Netball“ in Form eines Vergleichs zwischen 100 Hz und 50 % Shutter, interpoliert zu 200 Hz, und der Referenz kaum ein Unterschied. Dies soll den Testteilnehmern klar machen, dass auch scheinbar identische Sequenzenpaare im Test vorkommen können.

Die ersten fünf Sequenzen in Session 1 dienen als Dummy-Sequenzen, weshalb aus dieser Session effektiv 25 Bewertungen einfließen. Die nachfolgenden Sessions 2 und 3 beginnen mit jeweils zwei Dummy-Sequenzen, folglich können hierbei 28 Bewertungen entnommen werden.

⁴⁰ vgl. Tanton, Nick E. (2004): Results of a survey on television viewing distance. Research & Development White Paper WHP 090. British Broadcasting Corporation. London, S. 4. Online verfügbar unter <http://downloads.bbc.co.uk/rd/pubs/whp/whp-pdf-files/WHP090.pdf>, zuletzt geprüft am 11.09.2015

5.1 Auswirkungen unterschiedlicher Shuttereinstellungen

In diesem Test wird untersucht, wie sich eine Veränderung der Shuttereinstellung in der Kamera auf die subjektive Wahrnehmung von Bewegungen in der Szene auswirkt. Die verschiedenen Shutter wurden wie in Kapitel 4.2.1 beschrieben erzeugt, da die Sequenz andernfalls mit einem Kamerarig, aufgebaut aus zwei identischen Kameras, hätte aufgezeichnet werden müssen.

Ähnliche Tests wurden bereits zu einem früheren Zeitpunkt mit anderen Testsequenzen und Bildwiederholraten von 60 Hz bis 240 Hz durchgeführt. Entsprechende Ergebnisse wurden im Kapitel 3.3 dieser Arbeit erläutert, weshalb ein erneuter Vergleichstest der Shuttereinstellungen nicht so sehr im Fokus steht und in erster Linie zur Information dienen soll.

Aus den vorhandenen Quellsequenzen, die in bis zu 400 Hz vorliegen, konnten in der Vorbereitung des Tests insgesamt acht verschiedene Kombinationen aus Bildwechselfrequenz und Shutter erzeugt werden. Die Abb. 15 zeigt die bereits in Kapitel 4.2.1 beschriebenen Downsamplingmaßnahmen anhand der verwendeten Testsequenzen. Die Sequenzen „Hürdenlauf“ und „Stufenbarren“ wurden in 400 Hz aufgenommen, die restlichen drei in 200 Hz. Deshalb ist die Erzeugung von 200 Hz mit 50 % Shutter sowie 100 Hz mit 25 % Shutter nur bei diesen zwei Sequenzen möglich.

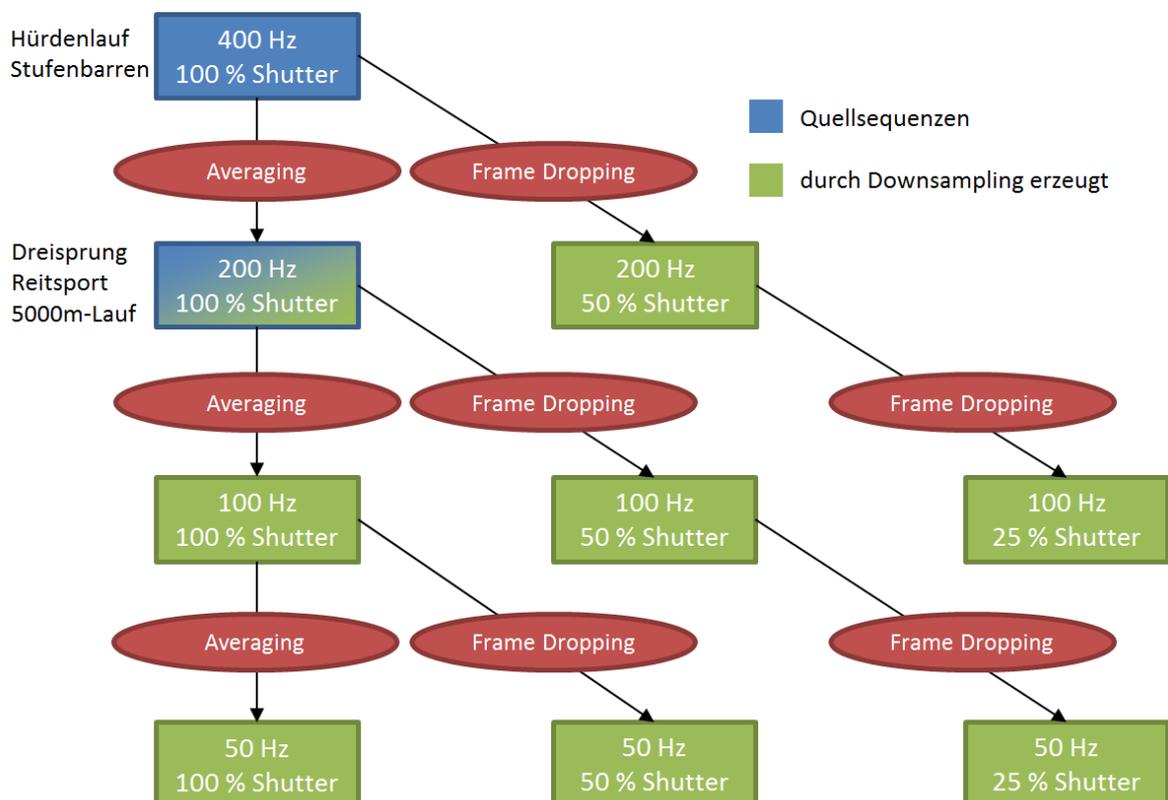


Abbildung 15: Downsampling anhand der verwendeten Testsequenzen

Einzig die beiden Sequenzen in 400 Hz können nicht in den Testsessions untersucht werden. Somit werden alle in Grün eingefärbten Varianten getestet. Die Sequenzen in 100 Hz

und 50 Hz werden, wie bereits im Kapitel 4.2.4 angedeutet, mithilfe des *BBC* Downsampling-Tools vorbereitet, damit deren Bilder doppelt bzw. vierfach wiedergegeben werden und somit alle Clips bei der Wiedergabe mit 200 Bildern pro Sekunde vorliegen. Dies vereinfacht die Wiedergabe mittels Clipster enorm.

5.1.1 Durchschnittliche Bewertung über alle Sequenzen

Abb. 16 zeigt die Auswertung des subjektiven Tests in Form der Mittelwerte der Bewertungen aller Probanden über alle Sequenzen für die jeweilige Kombination aus Bildwechselfrequenz und Shutter. Wie erwartet geht eine Verdopplung der Bilder pro Sekunde mit einer Verbesserung des subjektiven Eindrucks einher. Beim Übergang von 50 Hz zu 100 Hz verbessert sich die wahrgenommene Qualität um ca. eine Stufe auf der Skala (vgl. 4.3.1). Die Erhöhung der Bildrate von 100 Hz zu 200 Hz bringt noch einmal eine Verbesserung um etwa eine halbe Qualitätsstufe. Dies könnte darauf hindeuten, dass eine Bildwiederholrate von 200 Hz ausreichend ist, um eine Bewegtbildwiedergabe flüssig und ohne Bewegungsunschärfen darzustellen – ab einem bestimmten Punkt wäre demzufolge eine Erhöhung der Bilder pro Sekunde bei den verwendeten Testsequenzen nicht mehr wahrnehmbar.

Eine Verlängerung des Shutters führt bei 50 Hz ebenfalls zu einer leichten, statistisch messbaren Verbesserung während sich eine Veränderung dieses Parameters bei 100 Hz nur minimal bis gar nicht auswirkt. Dies bekräftigt die oben getroffene Aussage zur abnehmenden Empfindlichkeit der Wahrnehmung von Veränderungen bei größer werdenden Bildraten. Ebenso verhält es sich, wenn der Shutter bei 200 Hz variiert wird. Wird nur über die halbe Bilddauer belichtet, was einem Shutter von 50 % entspricht, ist die subjektive Qualität nahezu dieselbe wie bei einem Shutter von 100 % bei 200 Hz, was in diesem Test die Referenz für alle Vergleiche bildete. Es sei an dieser Stelle noch einmal darauf hingewiesen, dass nur für zwei der fünf getesteten Sequenzen eine Variante in 200 Hz mit 50 % Shutter erstellt werden konnte. Anhand dieses Diagramms lassen sich noch keine Aussagen treffen, die auf ein breites Spektrum an Bewegtbildinhalten übertragen werden können, da hierfür mindestens eine detaillierte Betrachtung der Bewertungen je Sequenz erforderlich ist.

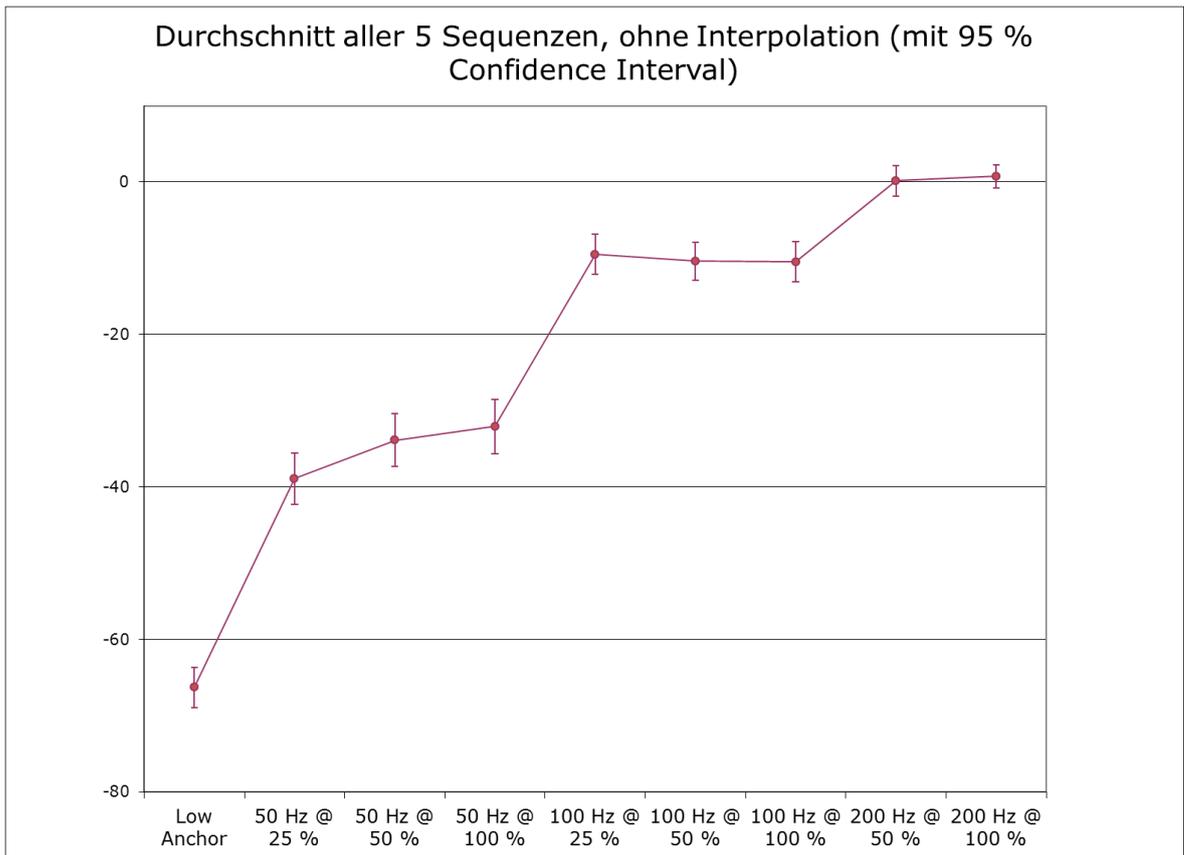


Abbildung 16: Durchschnitt aller fünf Sequenzen ohne Interpolation

5.1.2 50 Hz bei 25 %, 50 % und 100 % Shutter

Wurden die Sequenzen im Test mit lediglich 50 Hz Bildwechselfrequenz wiedergegeben, fiel deren Bewertung immer schlechter aus, als bei höheren Frequenzen. Die Differenz variierte dabei zum Teil sehr stark in Abhängigkeit des Inhalts der Sequenzen, wie Abb. 17 zeigt. „5000m-Lauf“ und „Reitsport“ ähneln sich hinsichtlich ihrer Schwenkrichtung sowie des Hintergrunds, wo bei beiden Sequenzen Tribünen mit vielen Zuschauern zu sehen sind. Dennoch fällt auf, dass vor allem bei 50 % Shutter die Bewertungen sehr stark auseinander gehen während beide Sequenzen bei 25 % und 100 % Shutter als nahezu gleichwertig gesehen wurden. Der Grund dafür liegt allem Anschein nach in der ungleichmäßigen Schwenkgeschwindigkeit der Sequenz „Reitsport“. Im Vergleich zu 25 % Shutter nimmt die Bewegungsunschärfe bei 50 % Shutter deutlich zu und variiert im Laufe der Sequenz mit der Geschwindigkeit des Schwenks. Diese störende Unschärfe kommt hierbei deutlicher zur Geltung als beispielsweise in der Sequenz „5000m-Lauf“, wo die ganze Zeit über mit konstanter Geschwindigkeit geschwenkt und die stärkere Unschärfe somit verdeckt wird. Umgekehrt wird ein stärkerer Stroboskopeffekt in Sequenzen mit gleichmäßigen Schwenks offenbar als störender bewertet. Ähnliches lässt sich auch in den Bewertungen der Sequenz „Dreisprung“ erkennen. Auch diese beinhaltet einen gleichmäßig ausgeführten Kameraschwenk. Zudem weist die Rasenfläche im Hintergrund zwar feine Details auf, wird aber farblich mehr als homogene Fläche akzeptiert, weshalb auch hier eine Hintergrundunschärfe in der Bewegung weniger auffällt als in der Sequenz „Reitsport“.

Sequenzen mit relativ statischen Einstellungen bzw. langsam bewegten Kameranischen, wie etwa die Sequenzen „Stufenbarren“ und „Hürdenlauf“, hinterließen bei den Testteilnehmern den subjektiv besten Eindruck, wenn die Shuttereinstellung mit 100 % gewählt wurde. Die Bewegungen der fokussierten Objekte erscheinen dabei flüssig und aufgrund geringer Kamerabewegung fällt auch die Unschärfe im Hintergrund geringer aus. Insbesondere bei der Sequenz „Stufenbarren“ stehen die Personen im Hintergrund zusätzlich im Halbdunkel und lenken deshalb noch weniger die Aufmerksamkeit der Testpersonen auf sich. Interessanterweise besitzt „Stufenbarren“ den höchsten TI-Wert aller fünf Sequenzen im Test (vgl. Abb. 9) und beinhaltet folglich unter objektiven Gesichtspunkten die meiste Bewegung.

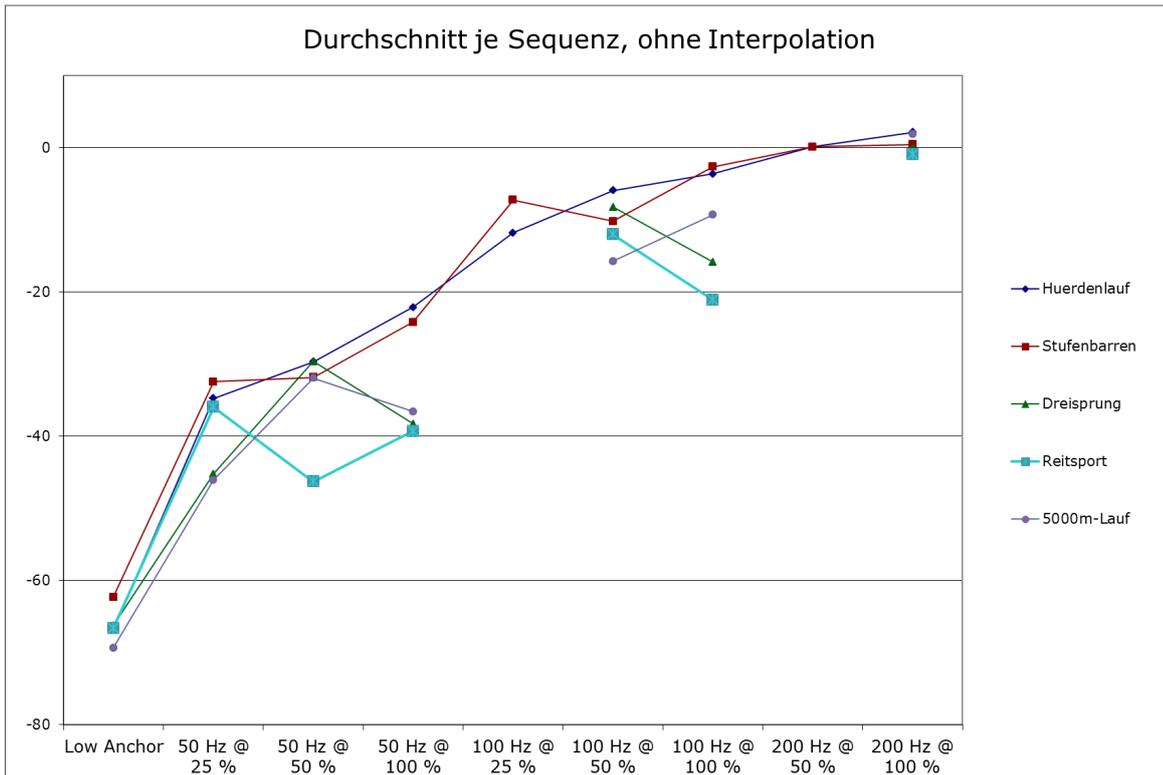


Abbildung 17: Durchschnitt je Sequenz ohne Interpolation

5.1.3 100 Hz bei 25 %, 50 % und 100 % Shutter

Die Abb. 17 macht weiterhin deutlich, dass die Sequenzen „Hürdenlauf“ und „Stufenbarren“ auch bei einer Verdopplung der Bildwechselfrequenz auf 100 Hz die besten Bewertungen erhielten, wenn der Shutter mit 100 % sehr lang gewählt wurde. Wie schon im obigen Abschnitt erläutert, wird dies damit begründet, dass diese beiden Sequenzen als subjektiv eher bewegungsarm gesehen werden, da die Kamera kaum oder nur sehr langsam bewegt wird. Die Sequenzen „Dreisprung“ und „Reitsport“ verzeichnen erwartungsgemäß Qualitätseinbußen bei einer Verdopplung des Shutters von 50 % auf 100 %, da die Bewegungsunschärfe zunimmt. Unerwartet ist allerdings die etwas bessere Bewertung der 100 % Shutter im Vergleich zu 50 % Shutter bei der Sequenz „5000m-Lauf“, obwohl diese viele Ähnlichkeiten zur Sequenz „Reitsport“ aufweist. Ein Expert Viewing zur Evaluierung dieses Qualitätsunterschieds brachte die Erkenntnis, dass der Gesamteindruck bei 50 % Shutter

wegen der geringeren Bewegungsunschärfe ein besserer ist als bei 100 %. Dass dieser Vergleich von den Probanden im Mittel umgekehrt bewertet wurde, liegt möglicherweise an besonders großen Abweichungen in der Bewertung einzelner Testpersonen vom Durchschnitt.

5.1.4 200 Hz bei 50 % und 100 % Shutter

An früherer Stelle im Kapitel 5.1 deutete sich keine wahrnehmbare Qualitätsverbesserung einer Wiedergabe in 200 Hz an, wenn mit halbiertem Shutter gearbeitet wurde. Diese Vermutung kann nach Betrachtung der detaillierten Ergebnisse je Sequenz bestätigt werden. Die zwei relevanten Sequenzen weisen nahezu die identischen Bewertungen bei 50 % Shutter auf wie bei der Referenz mit 100 % Shutter. Aufgrund der bereits getroffenen Aussagen bezüglich der geringen Kamerabewegung in diesen Sequenzen sowie der geringen Unterschiede verschiedener Shutter bei besonders hohen Bildwechselfrequenzen ist dieses Ergebnis nicht überraschend. Unglücklicherweise fehlen Ergebnisse für 200 Hz bei 50 % Shutter für die Sequenzen „5000m-Lauf“ und „Reitsport“, die im bisherigen Testverlauf und insbesondere bei Shuttervergleichen innerhalb einer gegebenen Bildwechselfrequenz teilweise recht deutliche Unterschiede aufwiesen.

5.2 Auswirkungen einer Zwischenbildberechnung

Auch dieser Test wurde in ähnlicher Form schon einmal im Rahmen des 4EVER Projekts durchgeführt, wenngleich die Sequenzen dabei mit maximal 120 Hz untersucht wurden. Wesentliche Ergebnisse sind im Kapitel 3.3 zusammengefasst. Es kam bereits eine Interpolation von Zwischenbildern zur Anwendung, die sich äußerst positiv auf das verwendete 60 Hz Material auswirkte, wenn letzteres mit 50 % Shutter produziert wurde. Demzufolge soll ein ähnlicher Vergleich zeigen, inwieweit eine Abhängigkeit vom Inhalt einer Sequenz besteht. Weiterhin wird untersucht, ob es sinnvoll ist, zwei Interpolationsschritte nacheinander durchzuführen und somit beispielsweise 50 Hz Archivmaterial in der Ausstrahlung der Sendeanstalt oder im Display des Zuschauers auf 200 Hz Wiedergabefrequenz zu bringen.

Die Abb. 18 verdeutlicht das Vorgehen bei der Erzeugung der interpolierten Testsequenzen. Ausgehend von den Sequenzen in 50 Hz und 100 Hz mit den verschiedenen Shuttereinstellungen werden entweder ein oder zwei Interpolationsschritte ausgeführt. Alle blauen Rechtecke symbolisieren eine Variante einer Sequenz, die in den subjektiven Test aufgenommen wird. Einzig für „Hürdenlauf“ und „Stufenbarren“ existiert nach dem Downsampling, dargestellt in Abb. 15, jeweils eine Variante in 100 Hz mit 25 % Shutter. Deshalb kann auch nur für diese beiden Sequenzen eine zu 200 Hz interpolierte Variante erstellt werden.

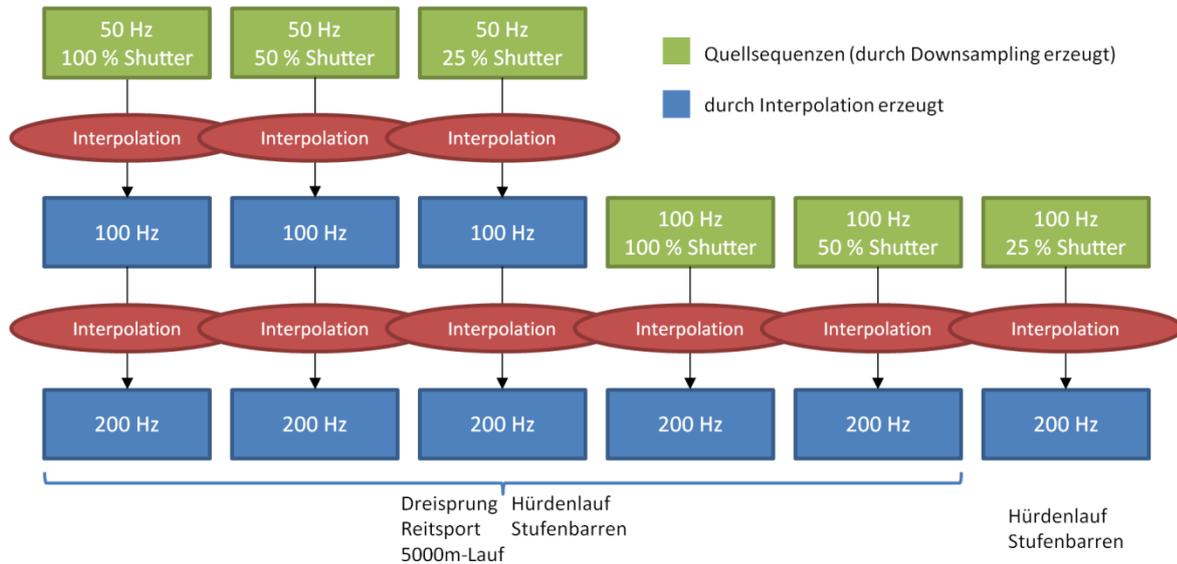


Abbildung 18: Interpolation anhand der verwendeten Testsequenzen

5.2.1 Durchschnittliche Bewertung über alle Sequenzen

In der ganzheitlichen Auswertung, dargestellt in Abb. 19, zeigt sich bereits, dass die Sequenzen, die durch bewegungskompensierende Bildinterpolation von 50 Hz auf 100 Hz gebracht wurden, fast um eine ganze Qualitätsstufe besser bewertet wurden als die nativen 50 Hz-Sequenzen, vorausgesetzt es wurde nicht über die komplette Bilddauer belichtet. Dennoch liegen die nativen 100 Hz-Sequenzen noch um etwa eine halbe Qualitätsstufe über den interpolierten. Die Qualitätsdifferenz zu den erneut interpolierten und damit wiedergabeseitig auf 200 Hz hochgerechneten Sequenzen fällt hingegen äußerst ernüchternd aus. Eine mögliche Fehlerquelle könnte der Workflow sein, nach dem bei dieser zweifachen Interpolation vorgegangen wurde – die finalen Sequenzen konnten nicht in einem Schritt aus den 50 Hz-Quellen erzeugt werden, sondern es wurden in einem Zwischenschritt zunächst Sequenzen in 100 Hz erstellt und anschließend die Interpolation wiederholt. Möglicherweise könnten durch eine unmittelbare Berechnung von drei Zwischenbildern bessere Ergebnisse erzielt werden.

Aus 100 Hz-Quellen nach 200 Hz interpolierte Inhalte erreichen beinahe das Qualitätsniveau der Referenz, unter der Voraussetzung eines möglichst kurzen Shutter. Dies könnte schon ein Indiz für eine bandbreiteneffizientere Ausstrahlung von HFR auf mittel- oder langfristige Sicht sein. Um klarer differenzierte Aussagen treffen zu können, bedarf es wiederum einer detaillierten Auswertung der Ergebnisse für die einzelnen Sequenzen.

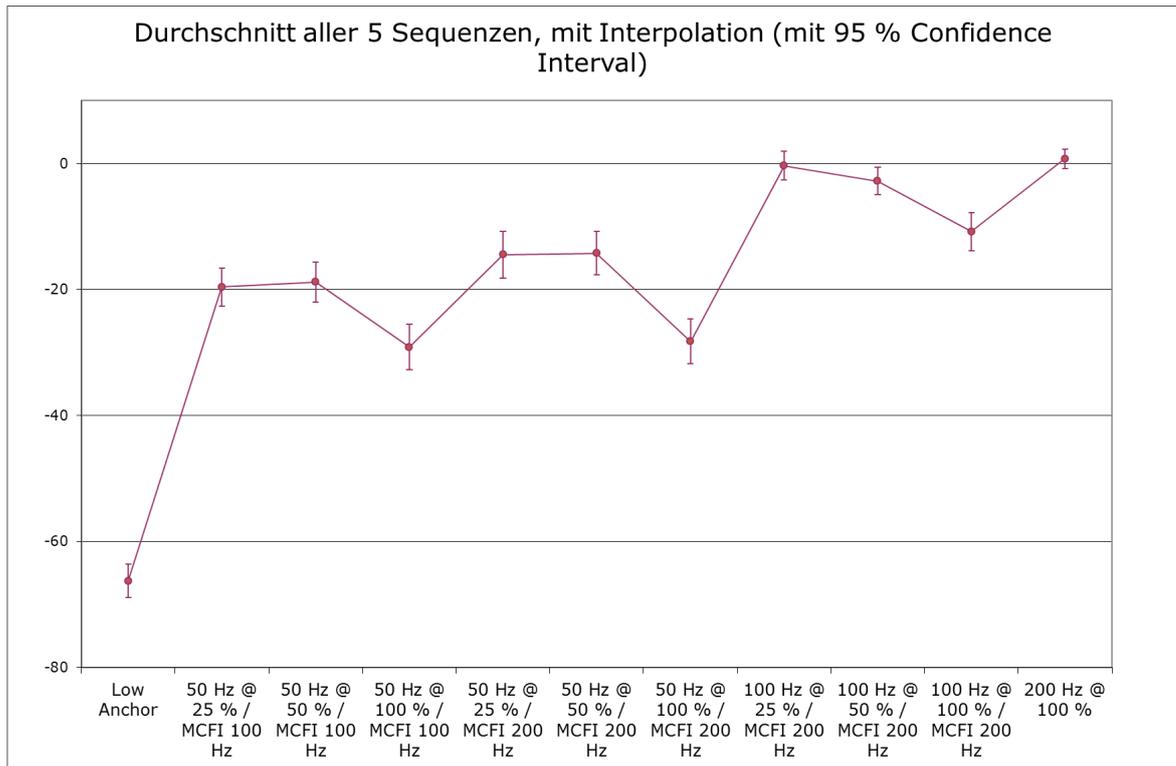


Abbildung 19: Durchschnitt aller fünf Sequenzen mit Interpolation

5.2.2 Interpolation von 50 Hz zu 100 Hz in Abhängigkeit der Shuttereinstellung

Die Frage nach der optimalen Shuttereinstellung des Quellmaterials ist im Hinblick auf die Bewegungsinterpolation von besonderem Interesse. Die Ergebnisse des subjektiven Tests, die im Diagramm in Abb. 20 dargestellt sind, lassen dabei eine Tendenz hin zu kürzeren Shuttern bei bewegungsintensiven Inhalten erkennen. Dies lässt sich dadurch erklären, dass durch einen generell schärferen Gesamteindruck, erzielt durch kürzere Shutter, mehr hochfrequente Anteile im Signal vorhanden sind, die der Interpolationsalgorithmus als Objektkanten und Details besser in Bewegungen übersetzen kann. Hohe Bewegungsunschärfe, wie sie lange Shutter zur Folge haben, führt zu fehlenden Strukturen und Kontrastarmut im Bild, gleichzusetzen mit niedrigen Frequenzen im Signal. Bewegungen können schlechter detektiert und mittels Berechnung von Zwischenbildern kompensiert werden. Zu beobachten ist dieses Verhalten insbesondere in den Sequenzen „Reitsport“ und „5000m-Lauf“.

Die einzige Sequenz, bei der die zu 100 Hz interpolierte Variante aus einer 50 Hz-Quelle mit 100 % Shutter als optimal bewertet wurde, ist „Hürdenlauf“. Die Sequenz beinhaltet bekanntlich als einzige im Test einen vertikalen Kameraschwenk, welcher sehr langsam und gleichmäßig ausgeführt wird. Hierbei sind nur eine minimale Bewegungsunschärfe und wenige bis gar keine Artefakte sichtbar. Eine mögliche Erklärung kann durch die Größe der möglichen Fehlerquellen für die Interpolation innerhalb des Bildes gegeben werden. Anders als beispielsweise bei der Sequenz „Stufenbarren“, wo die Turnerin als nahezu einziges

bewegtes Objekt im Bild sehr präsent ist, stehen bei „Hürdenlauf“ mehrere Läufer im Fokus, deren schnelle Arm- und Beinbewegungen mit abrupten Richtungswechseln die Interpolation an ihre Grenzen bringen können. Die betreffenden Bildbereiche sind aber kleiner als bei „Stufenbarren“ und zusätzlich örtlich voneinander getrennt, sodass die Bereiche, in denen Artefakte auftreten können, weniger in den Mittelpunkt rücken und alles in allem ein guter Gesamteindruck entsteht. Die kreisförmige Bewegung der Turnerin am Stufenbarren kann dagegen sehr gut mit den Augen verfolgt werden, wodurch die üblichen Artefakte einer fehlerhaften Interpolation wie z. B. doppelte oder verwaschene Konturen leichter erkennbar sind.

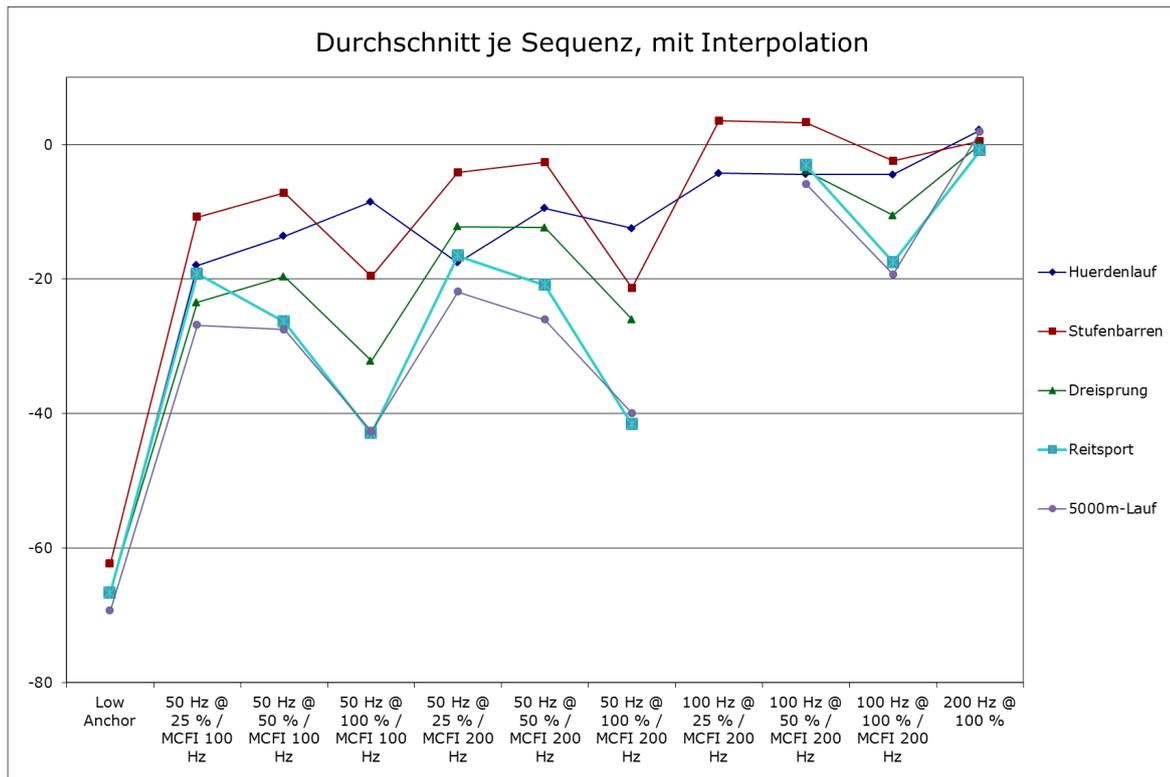


Abbildung 20: Durchschnitt je Sequenz mit Interpolation

5.2.3 Interpolation von 50 Hz zu 200 Hz

Wie bereits zu Beginn des Kapitels 5.2.1 erläutert, ist der Qualitätsunterschied der einfach interpolierten Sequenzen zu den Sequenzen, die durch wiederholte Interpolation erstellt wurden, äußerst marginal. Bei näherer Betrachtung der sequenzspezifischen Auswertung fällt auf, dass die Qualität der von 50 Hz nach 200 Hz interpolierten Sequenz „Stufenbarren“ bei langem Shutter sogar schlechter ausfällt als die der einfachen Interpolation. Ein Expert Viewing bestätigte dieses Verhalten. Die durch fehlerhafte Bewegungsschätzung auftretenden Artefakte, konkret sind dies Unschärfeerscheinungen und Doppelkonturen im Bereich um die Turnerin, werden durch die wiederholte Interpolation noch verstärkt. Besonders auffällig ist dies unmittelbar vor bzw. während des Abgangs vom Sportgerät, weil dann die Geschwindigkeit der Turnerin am höchsten ist.

Die Sequenz „Dreisprung“ verzeichnet im Mittel durch eine erneute Interpolation den größten Qualitätsgewinn, der ungefähr bei einem halben Punkt gegenüber der einfachen Interpolation liegt. Am deutlichsten ist dieser bei einem Shutter von 50 %, weshalb dieser Vergleich in einem Expert Viewing nachträglich betrachtet wurde. Tatsächlich erscheinen die Personen am Tisch zwischen den beiden Anlaufbahnen schärfer, wenn die zweistufige Interpolation vorgenommen wurde. Die auftretenden Artefakte um die Sportlerinnen, durch den einfarbigen Hintergrund gut erkennbar, sind bei beiden Sequenzen allerdings ähnlich stark ausgeprägt.

5.2.4 Interpolation von 100 Hz zu 200 Hz in Abhängigkeit der Shuttereinstellung

Die interpolierten 200 Hz-Sequenzen aus Quellen in 100 Hz erwiesen sich im Test als qualitativ beinahe gleichwertig zu den Referenzsequenzen. Die Sequenzen „Stufenbarren“ und „Hürdenlauf“ konnten recht erfolgreich nach 200 Hz interpoliert werden, wobei die Shuttereinstellung der Quelle eine untergeordnete Rolle spielt. Für 100 % Shutter bei der 100 Hz-Quelle waren laut den Testergebnissen nur leichte Qualitätsverluste gegenüber 50 % Shutter erkennbar. Grund hierfür ist auch die geringe Kamerabewegung in diesen Sequenzen. Anders bei den Sequenzen mit mehr Kamerabewegung „Dreisprung“, „Reitsport“ und „5000m-Lauf“. Die Vermutung liegt nah, dass ein längerer Shutter in der Quellsequenz hier zu deutlich mehr Bewegungsunschärfe nach der Interpolation führt. Dies konnte durch Expert Viewings auch bestätigt werden. Die Interpolationsartefakte sind beispielsweise in der Sequenz „Reitsport“ die gleichen wie vor der Interpolation, jedoch nimmt die Bewegungsunschärfe im Hintergrund sichtbar ab. Das Gegenbeispiel ist die Sequenz „Stufenbarren“, wo im direkten Vergleich der auf 200 Hz hochgerechneten Varianten aus einer Quelle mit 50 % sowie 100 % Shutter bei 100 Hz kein subjektiver Unterschied festzustellen ist. Insgesamt kann bei allen Sequenzen ein verstärktes Auftreten von Interpolationsartefakten bei großen Kontrastwechseln oder einer sich schnell ändernden Bewegungsrichtung von Objekten beobachtet werden, beispielsweise wenn sich in der Sequenz „5000m-Lauf“ die Läufer an einer hellen Werbebande vorbeibewegen. Gut erkennbar ist dieses Verhalten in der Abb. 21.



Abbildung 21: Interpolationsartefakte

6 Tests zur Codierung

Nachdem im vorhergehenden Kapitel die grundlegenden Möglichkeiten der Erzeugung von HFR-Inhalten innerhalb der Produktionskette betrachtet wurden, soll dieses Kapitel die Codierung mit einbeziehen. Es sollen Fälle untersucht werden, die auf die Praxis übertragen werden können. Deshalb sind einige Vorbetrachtungen nötig, um die HFR-Workflows optimal auf eine mögliche zukünftige Produktionslandschaft abbilden zu können.

6.1 Vorbetrachtung und Testvoraussetzungen

6.1.1 Praxisbezug

Die momentan bei den öffentlich-rechtlichen Rundfunkanstalten in Deutschland vorherrschende Bildwechselfrequenz liegt bei 50 Hz. Beinhaltet ein künftiger Produktionsstandard auch HFR, müssen die höheren Bildraten für das bestehende Archivmaterial der Rundfunkanstalten anderweitig erzeugt werden. Dies kann entweder senderseitig vor der Ausstrahlung erfolgen oder durch Interpolationsprozesse in den Displays der Zuschauer. Beide Wege müssen durch Tests untersucht werden, um feststellen zu können, ob es Qualitätsunterschiede gibt und wo diese liegen.

6.1.2 Werkzeuge für Encoding und Decoding

Zum Zeitpunkt des Verfassens der Arbeit waren weder Encoder noch Decoder am Markt verfügbar, die bereits eine Codierung mit zeitlicher Skalierbarkeit, wie in ITU-R BT.2073 empfohlen und in Kapitel 2.3.2 erläutert, offiziell unterstützen. Deshalb werden alle Testsequenzen als einzelne Transportströme codiert. Dabei wird der Offline-Encoder TITAN File von *Ateme* in der Softwareversion 3.7.2.0 eingesetzt, der sich bereits in früheren Tests am *IRT* als qualitativ hochrangiger HEVC-Encoder erwiesen hat. Dem Encoder werden eingangsseitig die erzeugten Sequenzen unkomprimiert mit einer 10 Bit Quantisierung und 4:2:2 Farbunterabtastung bereitgestellt. Als Containerformat dient *Apples* Quicktime. Die Transportströme am Ausgang liegen in 10 Bit und 4:2:0 vor.

Decodiert werden diese mittels der Software zur Elementarstromanalyse von *Tektronix* MTS4EA in YUV-Dateien, die ausschließlich die Videorohdaten beinhalten. Diese müssen anschließend noch mittels FFmpeg wiederum in einen MOV-Container geschrieben werden, damit sie mit Clipster weiterverarbeitet werden können.

In den Tests der *EBU* wird dagegen sowohl für das Encoding als auch für das Decoding die HEVC-Referenzimplementierung vom *Fraunhofer Heinrich Hertz Institut* mit der Bezeichnung HEVC Test Model (HM) verwendet. Diese dient in erster Linie zu Testzwecken der im

Standard definierten Vielfalt an Kompressionswerkzeugen und ist nicht vorrangig auf eine optimale Bildqualität hin optimiert. Da keine Rate Control stattfindet, kann nicht die Datenrate explizit festgelegt werden, sondern lediglich ein Basiswert für diverse Quantisierungsparameter (Quantization Point – QP). Ein PSNR-Vergleich mit der *Ateme*-Software anhand einer Sequenz und zweier Bitraten bzw. QP zeigt leichte Vorteile des HM, die zumindest für die stärker komprimierten Sequenzen auch subjektiv bestätigt werden können. Bei einer höheren Bitrate bzw. geringeren QP erscheinen die beiden Encoder subjektiv identisch.

6.2 Testdurchführung

Im Folgenden wird der Ablauf des Tests erläutert. Zu Beginn erfolgt die Auswahl geeigneter Codierdatenraten und die Festlegung der weiteren Parameter, die in Kapitel 2.3.2 bereits allgemein beschrieben wurden. Anschließend wird die Bildqualität objektiv bewertet, bevor eine Korrektur anhand subjektiver Kriterien vorgenommen wird. Die Sequenzen, die durch zwei Interpolationsschritte erzeugt wurden, sollen für die Codingszenarien nicht mehr betrachtet werden, da diese minimale bis keine Verbesserungen gegenüber den Sequenzen aufweisen, deren Bildfrequenz durch Interpolation verdoppelt wurde (vgl. 5.2.1 und 5.2.2).

6.2.1 Encoding

Anhand der Abb. 9 wurde bereits ersichtlich, dass die ausgewählten Sequenzen teils große Unterschiede in ihren örtlichen Details und der Komplexität von Bewegungen aufweisen. Die Codierung erfolgt daher nicht mit denselben Datenraten über alle Sequenzen, sondern wird für jede Sequenz differenziert betrachtet. Dadurch werden die Ergebnisse realistischer im Hinblick auf eine zukünftige Ausstrahlung innerhalb eines statistischen Multiplex (StatMux). Um dennoch eine gewisse Einheitlichkeit zu wahren, wurden über alle Sequenzen und Bildwiederholraten fünf konstante Datenraten (Constant Bit Rate – CBR) beibehalten, aus denen für jede Sequenz gemäß ihrer Komplexität drei ausgewählt wurden. In Abb. 22 sind diese grafisch dargestellt. Die vorangestellte Bildwiederholrate beschreibt die Bildrate, in der die jeweilige Testsequenz unmittelbar vor dem Encoding vorliegt. Folglich werden beispielsweise die Parameter für eine 100 Hz-Quellsequenz gewählt, wenn diese durch Interpolation aus einer 50 Hz-Sequenz entstanden ist.

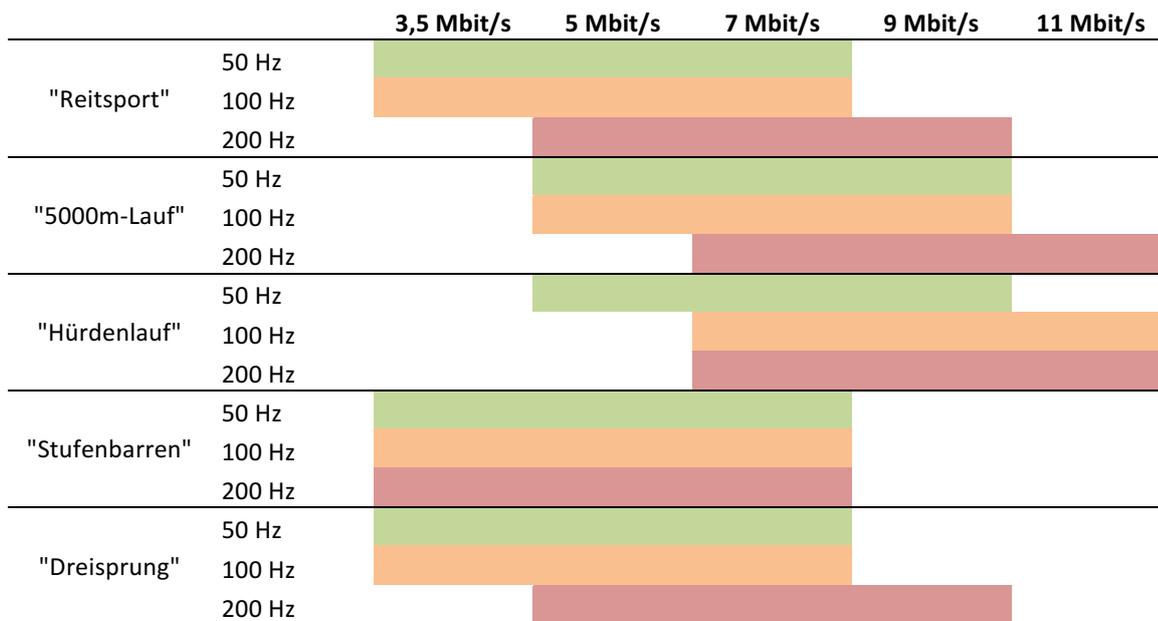


Abbildung 22: Festgelegte Codierdatenraten für alle Sequenzen

Die Auswahl erfolgte, nachdem alle Sequenzen mit 2,5–5 Mbit/s codiert und daraufhin subjektiv anhand ihrer Komplexität bewertet wurden. Bei einer Datenrate von 2,5 Mbit/s traten bei fast allen Sequenzen deutlich sichtbare Codierartefakte in Form von Blockingstrukturen auf, weshalb diese sehr geringe Datenrate für die weiteren Untersuchungen ausgelassen wurde. Vergleicht man die berechneten SI- und TI-Werte aus Abb. 9 mit den oben dargestellten Bitraten, wird erkennbar, dass teilweise trotz objektiv sehr hohem örtlichen Detailreichtum und sehr viel Bewegung innerhalb einer Sequenz eine geringe Datenrate bereits ausreicht, um die Sequenz subjektiv fehlerfrei codieren zu können. Die Sequenz „Stufenbarren“ beispielsweise weist den höchsten gemessenen Wert auf, was die Bewegung betrifft, kann aber in allen Bildwiederholraten mit 3,5–7 Mbit/s hinreichend fehlerfrei codiert werden. Dies liegt im Wesentlichen in der Art der Bewegung. Die Turnerin in dieser Sequenz bewegt sich recht gleichmäßig und gerichtet, schnelle Richtungsänderungen sind selten. Des Weiteren ist der Hintergrund verhältnismäßig homogen und detailarm, sodass jegliche Datenrate effizient für die sich bewegende Turnerin eingesetzt werden kann. Heutige Encodergenerationen sind mit sehr guten Prädiktionsalgorithmen ausgestattet, welche solche Arten von Bewegungen gut voraussagen und kompensieren können. In der Sequenz „Hürdenlauf“ dagegen ändert sich permanent und rasant schnell die Bewegungsrichtung der Arme der Athleten, weshalb diese Sequenz sehr schwierig zu codieren ist und mehr Datenrate benötigt, obwohl der Wert des Temporal Index deutlich unter dem der Sequenz „Stufenbarren“ liegt.

In Kapitel 2.3.2 wurden einige Aussagen zur GOP-Struktur gemacht, die als Grundlage für die Dimensionierung dieses Parameters dienen sollen. Wie dort bereits erwähnt, kann bei einer Codierung von HFR-Material eine längere GOP gewählt werden. Für die Tests in dieser Arbeit werden ausgehend von der Bildfrequenz des Quellmaterials feste GOP-Längen von jeweils ca. einer halben Sekunde im Encoder konfiguriert. In früheren Encodergenerationen des Herstellers *Cisco* wurde bei Verwendung einer dynamischen GOP ein Vielfaches

von 16 als Länge empfohlen, um bestmögliche Ergebnisse zu erzielen. Diese Handhabung hat sich seitdem in mehreren *IRT*-Tests bewährt und wird auch in diesem Test so beibehalten. Deshalb ergeben sich für 50 Hz-Quellen die GOP-Länge 32, für 100 Hz-Quellen eine Länge von 64 und für Quellen in 200 Hz dementsprechend 128 Bilder in einer GOP.

6.2.2 PSNR-Berechnung

Den ersten Anhaltspunkt für die Bewertung liefern die berechneten PSNR-Werte der codierten Sequenzen zur jeweiligen unkomprimierten Quelle, wobei nur die Luminanzkomponenten miteinander verglichen werden. Am *IRT* liegt für diese Zwecke eine eigens programmierte Software vor, die eine PSNR-Berechnung aller Einzelframes nach der Formel 3 vollzieht, wobei anschließend der Mittelwert gebildet werden muss. Sinnvoll ist nur die PSNR-Betrachtung der Sequenzen ohne Interpolation sowie derer, die nach der Interpolation codiert werden (vgl. 6.1.1).

Im Versuch werden alle erzeugten Sequenzen mit den Datenraten aus Abb. 22 codiert. Die entsprechenden Diagramme befinden sich in den Anhängen 2–6. Abb. 23 zeigt in Form eines Diagramms die dabei ermittelten durchschnittlichen PSNR über alle Sequenzen vorerst ohne Interpolation, in Abb. 24 wurde die Interpolation einbezogen. Jede Sequenz wurde bekanntlich nur mit drei verschiedenen Datenraten codiert, die fehlenden Werte wurden mittels potenzieller Regression aus den bereits vorhandenen Werten berechnet. So erhält man eine gute Näherung der Ergebnisse, die eine reale Codierung mit anschließender PSNR-Berechnung ebenfalls bringt.

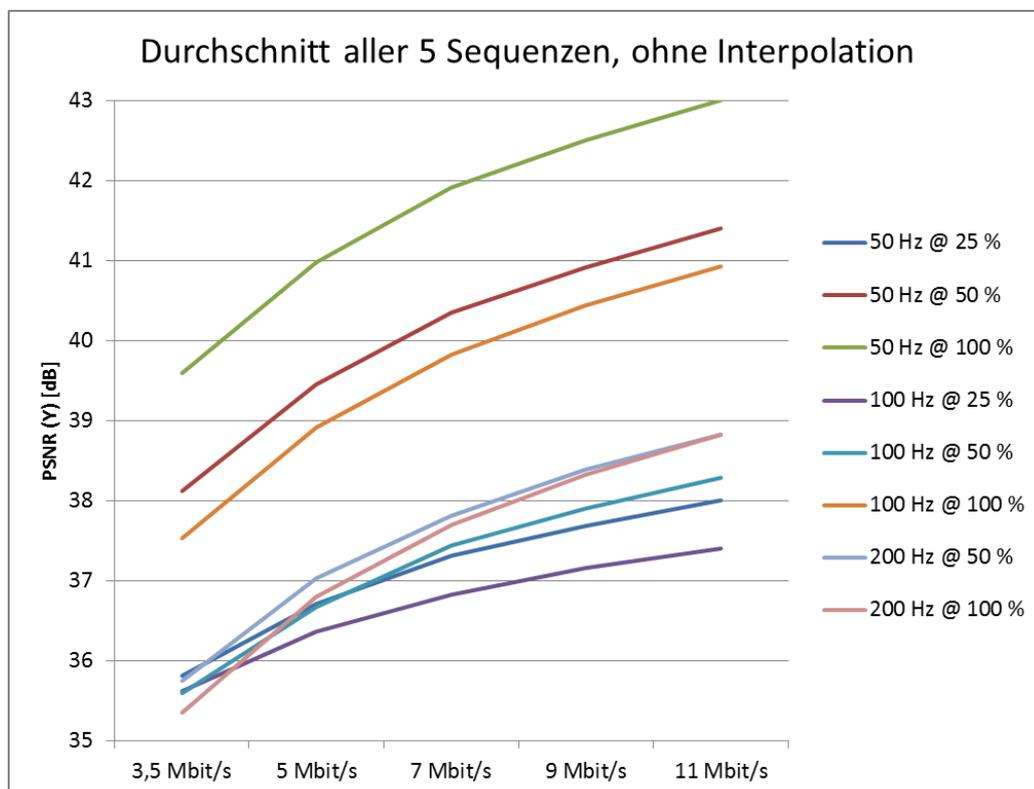


Abbildung 23: Durchschnitt aller fünf Sequenzen ohne Interpolation

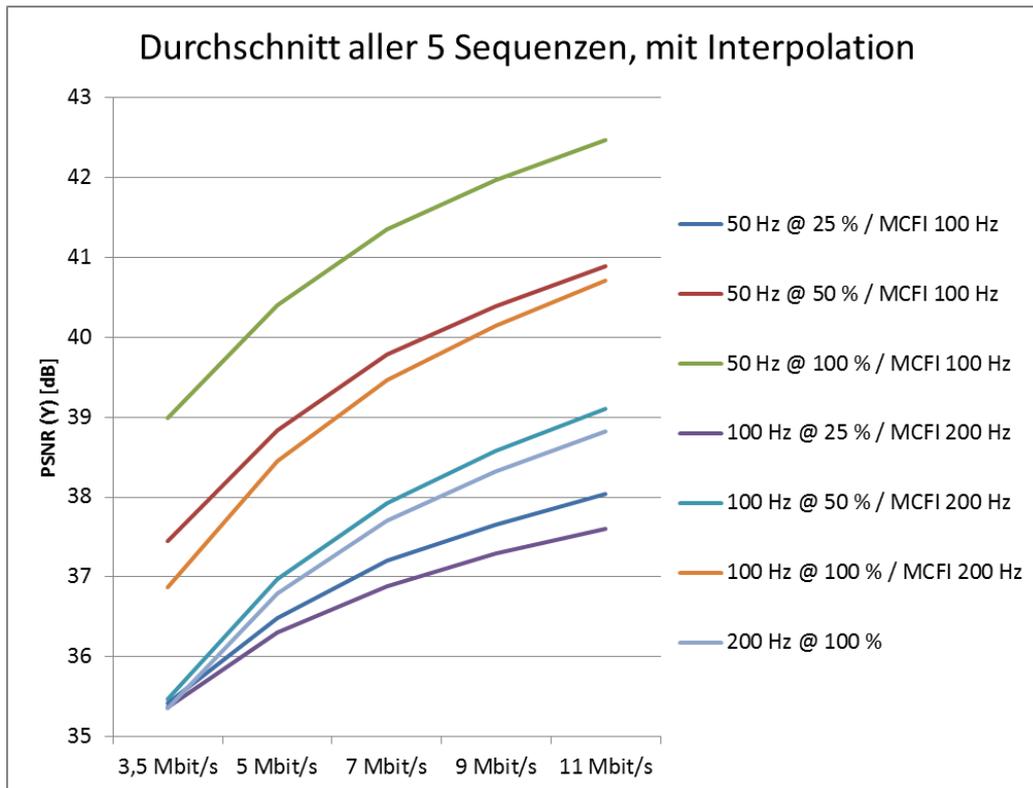


Abbildung 24: Durchschnitt aller fünf Sequenzen mit Interpolation

Die objektiv mitunter sehr großen Unterschiede liegen zum Teil an der Art des Downsamplings. Die Sequenzen wurden ursprünglich mit 200 Hz bzw. 400 Hz und 100 % Shutter produziert. Das bedeutet, zur Erzeugung von 50 Hz bei 100 % Shutter wurden vier bzw. acht Einzelbilder miteinander verrechnet (vgl. 4.2.1), was zu mehr Bewegungsunschärfe vor der Codierung führen musste. Die fehlenden Details fordern den Encoder nur sehr wenig, was zu einem geringeren mittleren Fehler und schließlich zu einem höheren PSNR-Wert führt. Sequenzen, bei denen die niedrigeren Bildraten größtenteils durch Frame Dropping entstanden sind wie beispielsweise 50 Hz oder 100 Hz bei jeweils 25 % Shutter, liegen mit ihren PSNR-Werten deutlich näher an denen der 200 Hz Sequenzen. Daher sollten diese für einen vorläufigen objektiven Vergleich bevorzugt werden. Eine detaillierte Betrachtung ist nicht zuletzt wegen der unterschiedlichen ursprünglichen Bildwechselfrequenz bei der Produktion unabdingbar und wird im nachfolgenden Kapitel in Form einer subjektiven Testauswertung erfolgen.

6.2.3 Expert Viewings zur subjektiven Beurteilung

Ein reiner PSNR-Vergleich reicht nicht aus, da es sich bekanntlich um ein objektives Maß für die Bildqualität handelt. Erst ein subjektiver Vergleich lässt konkrete Aussagen für Datenratenempfehlungen zu, da bestimmte Fehler innerhalb eines Bildes zwar für die PSNR-Berechnung relevant sind, der menschliche Gesichtssinn jedoch relativ unempfindlich gegenüber dieser Art von Fehlern sein kann.

Von besonderem Interesse für die Auswertung sind die interpolierten Sequenzen, die im Test mit unkomprimiertem Material besonders gut abgeschnitten haben und sich in ihrer subjektiven Qualität nicht sehr stark von nativen 100 Hz bzw. 200 Hz-Sequenzen unterscheiden.

(a) 100 Hz, 50 % Shutter bei senderseitiger Interpolation gegenüber 200 Hz, 100 % Shutter

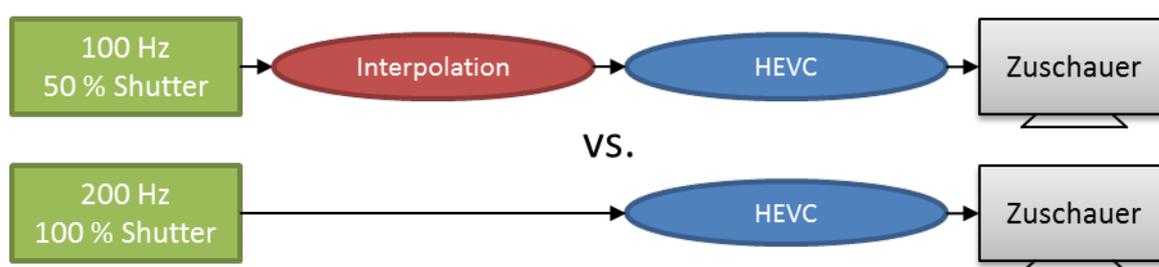


Abbildung 25: Grafische Darstellung des Vergleichs (a)

	100 Hz @ 50 % / MCFI 200 Hz / HEVC	200 Hz @ 100 % / HEVC
"Reitsport"	-gerade wahrnehmbar mehr Codierartefakte bei 5 Mbit/s	-gerade wahrnehmbar weniger Codierartefakte bei 5 Mbit/s
"5000m-Lauf"	-identisch bei 9 Mbit/s gegenüber 7 Mbit/s	
"Hürdenlauf"	-identisch bei 7 Mbit/s gegenüber 7 Mbit/s	
"Dreisprung"	-identisch bei 5 Mbit/s gegenüber 5 Mbit/s	
"Stufenbarren"	-identisch bei 3,5 Mbit/s gegenüber 3,5 Mbit/s	

Tabelle 4: Protokoll des Vergleichs (a)

Dieser Vergleich betrachtet das in Kapitel 6.1.1 beschriebene Szenario einer 200 Hz Ausstrahlung, wobei die Interpolation zu 200 Hz vor der Ausstrahlung also auch vor der Codierung vorgenommen wird. Die Auswertung des ersten Tests in Kapitel 5.2.4 haben bereits gezeigt, dass dabei ein kürzerer Shutter von Vorteil ist. Die anhand der codierten Sequenzen durchgeführten Expert Viewings bieten ein ähnliches Bild – über nahezu alle Sequenzen gibt es keine wahrnehmbaren Unterschiede an Codierartefakten. Die interpolierten Sequenzen von „Hürdenlauf“, „Dreisprung“ und „Stufenbarren“ können jeweils mit derselben Datenrate codiert werden wie die Referenz. Bei der Sequenz „5000m-Lauf“ benötigt die interpolierte Variante etwas mehr Datenrate, sodass sich bei 9 Mbit/s im Vergleich zu 7 Mbit/s der identische subjektive Eindruck ergibt. Vergleicht man die interpolierte Sequenz „Reitsport“ mit der Referenz bei 5 Mbit/s sind Codierartefakte auf beiden Seiten vorhanden, allerdings treten bei der hochgerechneten Sequenz gerade wahrnehmbar mehr dieser Artefakte auf.

Insgesamt ist festzuhalten, dass eine Ausstrahlung von bereits interpolierten Inhalten, wie sie oben beschrieben ist, bis auf wenige Ausnahmen nicht mehr Datenrate benötigt als eine Ausstrahlung von nativen 200 Hz, und die subjektiv wahrgenommenen Codierartefakte für den Konsumenten im Mittel gleich bleiben. Zudem lassen sich alle getesteten Sequenzen

mit 3,5–7 Mbit/s codieren, ohne dass dadurch deutlich wahrnehmbare Codierartefakte auftreten.

(b) 100 Hz, 50 % Shutter bei wiedergabeseitiger Interpolation gegenüber 100 Hz, 50 % Shutter bei senderseitiger Interpolation

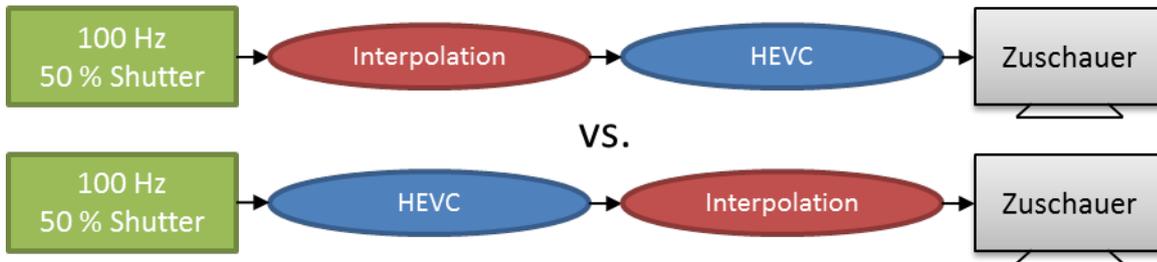


Abbildung 26: Grafische Darstellung des Vergleichs (b)

	100 Hz @ 50 % / MCFI 200 Hz / HEVC	100 Hz @ 50 % / HEVC / MCFI 200 Hz
"Reitsport"	-identisch bei 5 Mbit/s gegenüber 5 Mbit/s	
"5000m-Lauf"	-identisch bei 7 Mbit/s gegenüber 7 Mbit/s	
"Hürdenlauf"	-identisch bei 7 Mbit/s gegenüber 7 Mbit/s und 3 H -bei 7 Mbit/s gerade wahrnehmbar mehr Codier- und Bewegungsartefakte bei < 3 H	
"Dreisprung"	-identisch bei 5 Mbit/s gegenüber 5 Mbit/s	
"Stufenbarren"	-identisch bei 3,5 Mbit/s gegenüber 3,5 Mbit/s	

Tabelle 5: Protokoll des Vergleichs (b)

Diese Gegenüberstellung soll zeigen, ob sich sichtbare Qualitätsunterschiede ergeben, wenn die Interpolation von Programminhalten, die in 100 Hz produziert wurden, nicht vor der Codierung sondern erst im Fernsehgerät erfolgt. Das Expert Viewing beweist aber, dass bei identischer Datenrate und einem Betrachtungsabstand von der dreifachen Bildschirmhöhe kein subjektiver Unterschied auszumachen ist. Einzig die Sequenz „Hürdenlauf“ weist bei einem etwas geringeren Betrachtungsabstand gerade wahrnehmbar mehr Codier- und Bewegungsartefakte in der Variante der wiedergabeseitigen Interpolation auf.

Eine Verschiebung des Interpolationsvorgangs innerhalb der Produktions- und Ausstrahlungskette hat demzufolge keine bis minimale Auswirkungen auf die subjektive Bildqualität nach der Anwendung von Codierung und Interpolation. Die Berechnung von Zwischenbildern einer Quellsequenz, die bereits Codierartefakte enthält, hat somit ähnliche Auswirkungen auf das Endergebnis wie die Anwendung einer Codierung auf evtl. fehlerhaft berechnete Zwischenbilder.

(c) 100 Hz, 100 % Shutter bei wiedergabeseitiger Interpolation gegenüber 100 Hz, 50 % Shutter bei wiedergabeseitiger Interpolation

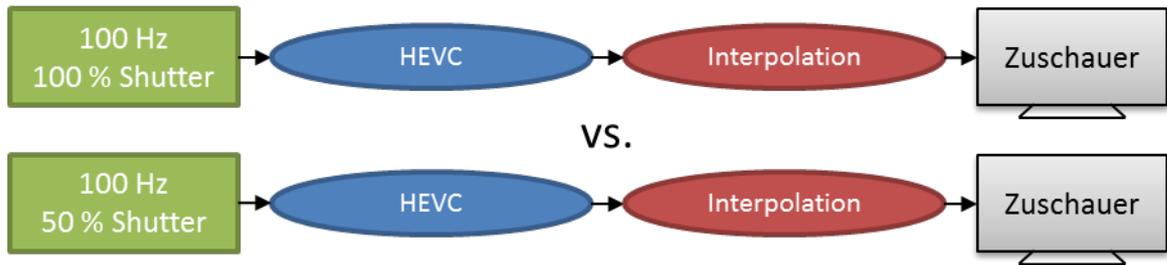


Abbildung 27: Grafische Darstellung des Vergleichs (c)

	100 Hz @ 100 % / HEVC / MCFI 200 Hz	100 Hz @ 50 % / HEVC / MCFI 200 Hz
"Reitsport"		-identisch bei 5 Mbit/s gegenüber 5 Mbit/s
"5000m-Lauf"		-identisch bei 7 Mbit/s gegenüber 7 Mbit/s
		-wahrnehmbar mehr Bewegungsartefakte
"Hürdenlauf"		-identisch bei 7 Mbit/s gegenüber 7 Mbit/s
"Dreisprung"		-identisch bei 5 Mbit/s gegenüber 5 Mbit/s
"Stufenbarren"		-identisch bei 3,5 Mbit/s gegenüber 3,5 Mbit/s

Tabelle 6: Protokoll des Vergleichs (c)

In diesem Vergleich wird lediglich die Shuttereinstellung der gegenübergestellten Sequenzen verändert, um festzustellen, ob ein kürzerer Shutter – gleichzusetzen mit schärferen Bildern – auch tatsächlich mehr Datenrate benötigt. Über alle Sequenzen betrachtet kann dieses erwartete Verhalten nicht bestätigt werden. Größtenteils führen identische Datenraten zu subjektiv identisch wahrgenommenen Codierartefakten, wobei die geringere Schärfe bei 100 % Shutter sowie das generelle Auftreten von Bewegungsartefakten (vgl. 5.2.4) einen reinen Vergleich der Codierung deutlich erschweren.

(d) 50 Hz, 25 % Shutter bei senderseitiger Interpolation gegenüber 100 Hz, 50 % Shutter

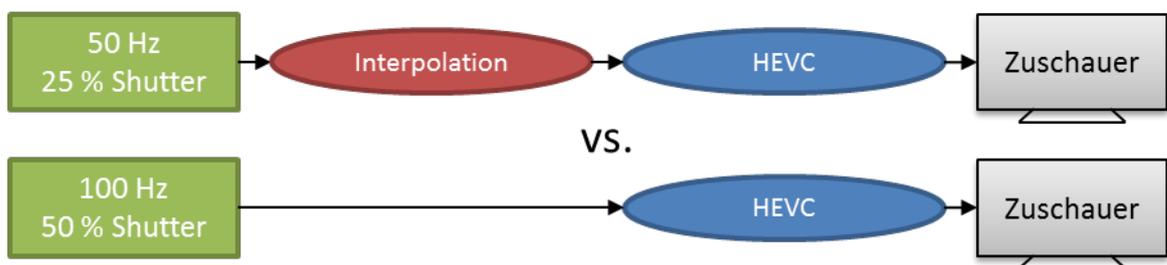


Abbildung 28: Grafische Darstellung des Vergleichs (d)

	50 Hz @ 25 % / MCFI 100 Hz / HEVC	100 Hz @ 50 % / HEVC
"Reitsport"	-gerade wahrnehmbar mehr Codierartefakte bei 5 Mbit/s	-gerade wahrnehmbar weniger Codierartefakte bei 5 Mbit/s
	-identisch bei 7 Mbit/s gegenüber 3,5 Mbit/s	
"5000m-Lauf"	-wahrnehmbar mehr Codierartefakte bei 9 Mbit/s	-wahrnehmbar weniger Codierartefakte bei 5 Mbit/s
"Hürdenlauf"	-identisch bei 11 Mbit/s gegenüber 7 Mbit/s	
"Dreisprung"	-identisch bei 7 Mbit/s gegenüber 3,5 Mbit/s	
"Stufenbarren"	-identisch bei 3,5 Mbit/s gegenüber 3,5 Mbit/s	

Tabelle 7: Protokoll des Vergleichs (d)

Nachdem ein ähnlicher Vergleich bereits für 100 Hz und 200 Hz gezogen wurde, soll hier untersucht werden, ob ein anderes Verhalten für geringere Bildwechselfrequenzen festgestellt werden kann. Um eine optimale Vergleichbarkeit zu gewährleisten, wurde wie bei der Gegenüberstellung (a) auch hier die absolute Belichtungszeit von 1/200 s auf beiden Seiten gleich gewählt. Ein erster Vergleich anhand identischer Datenraten lässt bereits erkennen, dass die mit 50 Hz produzierten Sequenzen subjektiv deutlicher wahrnehmbare Codierartefakte gegenüber den 100 Hz-Sequenzen aufweisen, als dies bei Vergleich (a) mit höheren Bildraten der Fall war. Die Vorteile der 100 Hz bei 50 % Shutter sind über fast alle Sequenzen hinweg gerade wahrnehmbar oder wahrnehmbar. Die interpolierten Sequenzen „Reitsport“, „Dreisprung“ und „5000m-Lauf“ erreichen erst bei der doppelten Datenrate gegenüber den 100 Hz-Sequenzen ein etwa identisches Qualitätsniveau. Der generelle Datenratenbedarf für eine hinreichend artefaktfreie Codierung liegt bei „5000m-Lauf“ noch etwas höher. Die Sequenz „Hürdenlauf“ weist allgemein weniger Bewegungsartefakte in der interpolierten Variante auf und benötigt demzufolge nicht die doppelte Datenrate einer nativen 100 Hz-Sequenz. Für die Sequenz „Stufenbarren“ sind 3,5 Mbit/s für einen subjektiv identischen Bildeindruck ohne Codierartefakte ausreichend. Alle anderen interpolierten Sequenzen benötigen für eine Codierung ohne negativ auffallende Codierartefakte etwa 7–11 Mbit/s. In 100 Hz mit 50 % Shutter hingegen lassen sich diese mit 3,5–7 Mbit/s hinreichend artefaktfrei codieren.

Zusammenfassend lässt sich feststellen, dass ein kürzerer Shutter in Kombination mit einer Berechnung von Zwischenbildern den Encoder deutlich mehr fordert, als dies bei nativem 100 Hz-Material der Fall ist. Ein Produktionsstandard mit 100 aufgenommenen Bildern pro Sekunde wäre folglich einer senderseitigen Hochkonvertierung von 50 Hz-Inhalten vorzuziehen.

**(e) 50 Hz, 50 % Shutter bei senderseitiger Interpolation gegenüber
50 Hz, 50 % Shutter bei wiedergabeseitiger Interpolation**

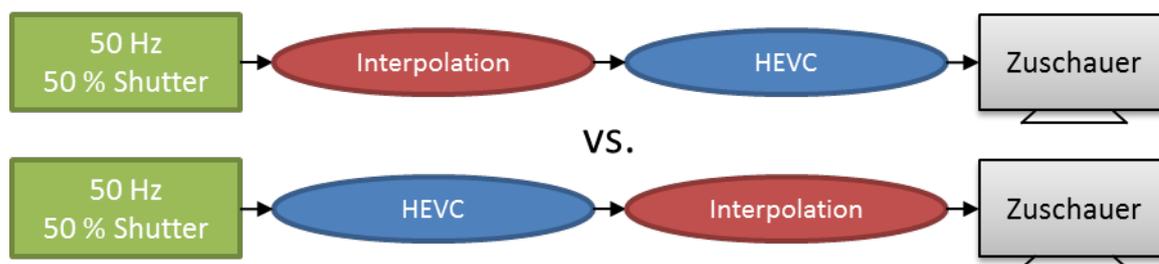


Abbildung 29: Grafische Darstellung des Vergleichs (e)

	50 Hz @ 50 % / MCFI 100 Hz / HEVC	50 Hz @ 50 % / HEVC / MCFI 100 Hz
"Reitsport"	-identisch bei 5 Mbit/s gegenüber 5 Mbit/s	
"5000m-Lauf"	-identisch bei 7 Mbit/s gegenüber 7 Mbit/s	
"Hürdenlauf"	-gerade wahrnehmbar weniger Codierartefakte bei sich stark bewegenden Objekten bei 7 Mbit/s	-gerade wahrnehmbar mehr Codierartefakte bei sich stark bewegenden Objekten bei 7 Mbit/s
"Dreisprung"	-identisch bei 5 Mbit/s gegenüber 5 Mbit/s	
"Stufenbarren"	-identisch bei 3,5 Mbit/s gegenüber 3,5 Mbit/s	

Tabelle 8: Protokoll des Vergleichs (e)

Da der Vergleich zwischen sender- und wiedergabeseitiger Interpolation bereits für Quellsequenzen in 100 Hz durchgeführt wurde, soll dieser Vergleich untersuchen, ob dasselbe Verhalten bei geringeren Bildwechselfrequenzen zu beobachten ist. Allerdings sind auch für eine Interpolation von 50 Hz zu 100 Hz keine deutlichen subjektiven Unterschiede zu erkennen, wenn die Interpolation vor oder nach der Codierung angewendet wird. Eine wiedergabeseitige Interpolation der Sequenz „Hürdenlauf“ führt zu gerade wahrnehmbar mehr Codierartefakten bei sich stark bewegenden Objekten.

7 Ergebnisse

Die wesentlichen Ergebnisse aus den beiden Tests werden in diesem Kapitel zusammengefasst. Weiterhin wird versucht, die erlangten Erkenntnisse auf die Praxis zu übertragen und somit allgemein gültige Handlungsempfehlungen für bestimmte Voraussetzungen im Produktionsalltag zu geben. Auf den Ergebnissen aufbauend werden abschließend einige Ansätze für weiterführende Forschungen vorgeschlagen.

7.1 Tests zur Bildwiederholfrequenz

Mit den vorangegangenen subjektiven Tests konnte der Mehrwert einer erhöhten Bildwiederholfrequenz klar identifiziert werden. Durch eine Erhöhung auf nativ produzierte und wiedergegebene 100 Hz konnte der subjektive Eindruck um eine Qualitätsstufe verbessert werden. Eine nochmalige Erhöhung auf 200 Hz brachte eine Verbesserung um eine weitere halbe Qualitätsstufe (vgl. 5.1.1). Bildinhalte mit einem hohen Bewegungsanteil, insbesondere aufgrund schneller Kameranews, profitierten dabei von einem kürzeren Shutter, während in eher statischen Einstellungen eine längere relative Belichtungszeit zu besseren Ergebnissen führte (vgl. 5.1.2, 5.1.3). Die Schwenkgeschwindigkeit innerhalb einer Szene steht offensichtlich in Wechselwirkung mit dem Grad der störenden Wahrnehmung der verschiedenen Bewegungsartefakte. In gleichmäßigen Schwenks wird eine erhöhte Bewegungsunschärfe weniger störend wahrgenommen als ein gleichzeitig vorhandener Stroboskopeffekt. Variiert die Schwenkgeschwindigkeit des Öfteren während einer Szene, fällt die sich ebenfalls ändernde Bewegungsunschärfe deutlicher ins Gewicht. Zusätzlich wirkt sich auch ein hoher Detailreichtum im Bildhintergrund insofern aus, dass Unschärfe eher als solche erkannt wird (vgl. 5.1.2). Während der Tests wurde ein weiterer, bisher eher unbeachteter Faktor für eine Einführung höherer Bildwechselfrequenzen erkannt – die bessere Lesbarkeit von Schriftzügen, insbesondere auf Werbebanden, könnte HFR für Sponsoren und die gesamte Werbeindustrie attraktiver machen. Letztere könnte in Anbetracht ihres enormen Marktvolumens ein wichtiger Befürworter werden.

Die größten Vorteile einer Zwischenbildberechnung kamen immer dann zur Geltung, wenn die Shuttereinstellung des Quellmaterials gemäß der Art der Bewegung innerhalb der Szene gewählt wurde. Kurze Shutter von 25 % oder 50 % brachten bessere Interpolationsergebnisse für bewegungsintensive Inhalte, ein langer Shutter von 100 % wirkte sich nur dann positiv auf die Interpolation aus, wenn statischere Kameraeinstellungen verwendet wurden und mehrere bewegte Objekte örtlich voneinander getrennt auftraten (vgl. 5.2.2, 5.2.4). Interpolationsartefakte traten zwar bei allen Sequenzen im Test auf, jedoch fielen diese bei den besonders gut bewerteten Sequenzen weniger auf. Vor allem große Kontrastunterschiede oder eine schnelle Änderung der Bewegungsrichtung waren anfällig für diese Art von Artefakten.

7.2 Tests zur Codierung

Die subjektiven Tests zur Codierung von Inhalten in HFR gaben Aufschluss darüber, ob ein erhöhter Bandbreitenbedarf für eine zukünftige Ausstrahlung besteht und wenn ja, wie viel mehr Datenrate letztendlich benötigt wird.

Es konnte durchweg festgestellt werden, dass Sequenzen mit mehr Bildern pro Sekunde dem Encoder nicht mehr Datenrate abverlangen, um einen identischen subjektiven Bildeindruck zu erzielen, sondern eher weniger. Es kommt dabei auf die Art der Erzeugung der höheren Bildwechselfrequenzen an. Während hochgerechnete Sequenzen bereits Interpolationsartefakte besitzen, die eine Codierung erschweren, kann für native HFR-Inhalte die Datenrate gegenüber niedrigeren Bildraten nahezu gleich bleiben. Besonders auffällig war dieses Verhalten im Vergleich (d) des Kapitels 6.2.3, wo 50 Hz-Material mit einem sehr kurzen Shutter von 25 % teilweise die doppelte Datenrate zur Codierung benötigte, um die subjektive Qualität der nativen 100 Hz Sequenzen mit 50 % Shutter zu erreichen. Vergleich (a) fand fast unter den gleichen Bedingungen statt, nur dass dort 100 Hz-Sequenzen hochgerechnet und mit nativen 200 Hz-Sequenzen verglichen wurden. Dabei benötigten erstere im Gegensatz zu Vergleich (d) im Mittel keine höhere Datenrate. Es lässt sich daraus schlussfolgern, dass ein Produktionsstandard mit 100 Vollbildern der beste Kompromiss wäre, da eine Interpolation vor der Ausstrahlung oder im Fernsehgerät zu Hause die Qualität noch einmal deutlich steigern könnte.

Eine Veränderung des Shutters vor der Interpolation und Codierung des Materials hatte kaum Auswirkungen auf die wahrgenommenen Codierartefakte, wie der Vergleich (c) zeigte. Dennoch sollten kürzere Shutter im Hinblick auf die geringeren Bewegungsartefakte bevorzugt werden. Letztere wirkten sich in allen Tests immer negativer auf den Gesamteindruck aus als Codierartefakte.

Die Vergleiche (b) und (e) zeigten, dass eine Verschiebung des Interpolationsprozesses vor bzw. nach die Codierung in allen Bildraten keinen subjektiven Unterschied bewirkt. Die beiden Varianten bringen Vor- und Nachteile mit sich. Würde senderseitig interpoliert werden, wäre die Qualität aufgrund des eingesetzten Algorithmus bei allen Empfängern gleich. Zusätzlich im TV-Gerät aktivierte Bildoptimierungsoptionen könnten die Bildqualität dagegen weniger gut beeinflussen. Bei einer ausschließlichen empfängerseitigen Interpolation kämen gegebenenfalls proprietäre Lösungen verschiedener Hersteller zum Einsatz, wodurch keine einheitliche Bildqualität gewährleistet wäre. Allerdings läge die Entscheidungsmacht beim Konsumenten, wie „scharf“ er sein Bild gern sehen möchte. Weitere Vor- und Nachteile müssten im Zuge einer geplanten HFR-Einführung zum gegebenen Zeitpunkt diskutiert werden.

7.3 Weitere Ergebnisse aus EBU-Tests

Seitens des 4EVER-Projekts, ebenfalls aktiv in der *EBU*-Arbeitsgruppe BTF-HFR, wurden auch subjektive Tests durchgeführt, die den Mehrwert einer HFR-Wiedergabe ermitteln sollten (vgl. 4.1 und 6.1.2). Der Fokus lag bei diesen Untersuchungen auf dem Mehrwert von 100 Hz, sowohl nativ als auch interpoliert, gegenüber 50 Hz. Das angestrebte Ziel ist die Einbeziehung dieser höheren Bildwechselfrequenz in die zukünftigen UHD-Standards des *Digital Video Broadcasting (DVB)* Konsortiums.

Die Tests beinhalteten unkomprimierte und codierte Inhalte, wobei auch eine bewegungskompensierende Zwischenbildberechnung betrachtet wurde. Als Testmethode wurde SAMVIQ gewählt und das verwendete Display war baugleich zum HFR-Display in den *IRT*-Tests. Die Auswahl der Testsequenzen unterschied sich in zwei der fünf Sequenzen – anstelle von „Stufenbarren“ und „Dreisprung“ wurden die Sequenzen „Speerwurf“, die in den *IRT*-Tests nur für die Trainings Sessions genutzt wurde, und „Boxen“ gewählt. Letztere stammt wie „Hürdenlauf“ und „5000m-Lauf“ von den Commonwealth Games 2014 und wurde durch die *BBC* bereitgestellt. Zudem wurden seitens 4EVER weniger Variationen aus Bildwechselfrequenz und Shutter in den Test einbezogen und insgesamt sechs Fälle betrachtet:

- 50 Hz @ 50 %
- 50 Hz @ 100 %
- 50 Hz @ 50 % / MCFI 100 Hz
- 50 Hz @ 100 % / MCFI 100 Hz
- 100 Hz @ 100 %
- 200 Hz @ 100 %

Zum Zeitpunkt des Verfassens dieser Arbeit umfassten die Testergebnisse die Daten von 13 Teilnehmern. Die Tests mit unkomprimiertem Material ergaben einen Mehrwert von ca. 20 Punkten von nativen 100 Hz gegenüber nativen 50 Hz. Das entspricht ziemlich genau der in den Kapiteln 5.1.1 und 7.1 beschriebenen Qualitätsverbesserung um eine Stufe. Die Erhöhung der Bildwechselfrequenz auf 200 Hz brachte einen weiteren, jedoch geringeren Qualitätsvorteil. Auch diese Aussage deckt sich mit den ermittelten Ergebnissen der vorliegenden Arbeit. Dass der Vorteil einer höheren Bildwechselfrequenz und gleichzeitig kürzeren Shuttern in Sequenzen wie „Reitsport“ und „5000m-Lauf“ deutlicher zur Geltung kommt, konnte ebenfalls durch die 4EVER-Tests bestätigt werden. Kürzere Shutter wirkten sich des Weiteren positiv auf eine Interpolation aus. Deren Leistungsfähigkeit war wiederum abhängig von der Art der Bewegung. Im schlechtesten Fall, konkret bei der Interpolation einer 50 Hz-Quelle mit 100 % Shutter in der Sequenz „Hürdenlauf“, erreichte die Qualität der interpolierten Sequenzen nicht die der nativen 50 Hz-Sequenzen. In den Sequenzen „5000m-Lauf“ und „Speerwurf“ hingegen konnte durch Interpolation beinahe die Qualität der nativen

100 Hz-Sequenzen erreicht werden. Insgesamt ließ sich allerdings feststellen, dass durch Interpolation erzeugte 100 Hz im Durchschnitt nicht die Qualität von nativen 100 Hz erreichten.

Die allgemeine Aussage der Codingtests des 4EVER-Projekts war, dass bei gleichbleibender Datenrate von 3 Mbit/s bessere Ergebnisse bei der Codierung nativer 100 Hz-Sequenzen erzielt werden konnten, als bei der Codierung nativer 50 Hz- oder interpolierter 100 Hz-Sequenzen. Trotz anders gewählter Shutter deckt sich diese Aussage im Wesentlichen mit den Ergebnissen des subjektiven Vergleichs (d) im Kapitel 6.2.3.

7.4 Empfehlungen für Produktionen in HFR

Die Empfehlungen für zukünftige Produktionen HFR orientieren sich nicht nur an einer optimalen Abbildung der Geschehnisse, die den Zuschauer interessieren, sondern auch an den natürlichen Sehgewohnheiten.

7.4.1 Sport- und Showformate

Sportinhalte, wie sie in dieser Arbeit ausführlich getestet wurden, beinhalten für gewöhnlich schnelle Bewegungen und Richtungsänderungen und damit schnelle Kameraschwenks. Durch eine große Schärfentiefe sind viele Hintergrunddetails vorhanden. Außerdem werden Sportstätten seit einigen Jahren mit leistungsstarker Lichttechnik ausgestattet, die auch eine Produktion mit Highspeed-Kameras vereinfacht. Somit sind Sportereignisse prädestiniert für HFR. Aufgrund der Testergebnisse sollte dabei ein Shutter von 50 % gewählt werden und die Bildwechselfrequenz bei mindestens 100 Hz liegen. Die gleichen Empfehlungen gelten für viele Showformate, da auch dort sehr viel künstliches Licht vorhanden ist und von Seiten der Kameraleute teils Unerwartetes mit schnellen Schwenks eingefangen werden muss.

7.4.2 Dokumentationen

Dokumentarische Inhalte umfassen u. a. Features, Doku-Soaps und Reportagen. Dabei spielt die szenische Gestaltung eine größere Rolle als beispielsweise bei Sportinhalten. Der Einsatz von HFR muss hierbei sicherlich je nach Kontext entschieden werden. Prägen eher langsame Schwenks, wenig Bewegung und viele statische Einstellungen das Bild, kann mit einem längeren Shutter von bis zu 100 % gearbeitet werden. Dies ist oftmals sogar notwendig, da bei dokumentarischen Produktionen generell weniger Licht vorhanden ist und teilweise auch nur natürliches Licht zur Verfügung steht. Offene Blenden und damit geringe Schärfentiefe stellen viele Vorzüge von HFR nicht deutlich genug heraus. In diesem Fall sollte mit 50 Vollbildern produziert und bei Bedarf nach 100 Hz interpoliert werden.

7.4.3 Szenische Produktionen

Szenische Produktionen sind geprägt von einer akribisch geplanten Szenengestaltung und einer geringen Schärfentiefe. Es gibt keine überraschenden Momente, die durch unkontrollierte Schwenks kompensiert werden müssen. Verwischte Objekte werden bewusst in die Unschärfe gelegt (vgl. 2.1.3). Zudem ist dieser sogenannte „Filmlook“ bei den Zuschauern etabliert. Höhere Bildwechselfrequenzen im fernsehtechnischen Sinne werden sich demzufolge in diesem Genre wahrscheinlich nicht durchsetzen. Szenische Produktionen sollten mit 50 Bildern pro Sekunde und einem bereits etablierten Shutter von 50 % erstellt werden.

7.5 Ansätze zur weiterführenden Forschung

Aufbauend auf den Ergebnissen dieser Arbeit können weitere Tests zusätzlich Aufschluss geben zu bestimmten Faktoren, die im Zusammenhang mit einer regulären HFR-Ausstrahlung stehen oder diskutiert werden.

7.5.1 Auswahl des Testmaterials

Bei allen zukünftigen Untersuchungen spielt auch die Auswahl des Testmaterials eine wichtige Rolle. Für die Tests in dieser Arbeit wurden fünf Sequenzen genutzt, für die zwar sehr unterschiedliche SI und TI berechnet wurden, die subjektiv aber teilweise Ähnlichkeiten aufwiesen. Drei der fünf Sequenzen beinhalteten einen horizontalen Kameraschwenk. Die Sequenzen „Reitsport“ und „5000m-Lauf“ besaßen zudem einen sehr ähnlichen Hintergrund. Trotzdem konnten die Unterschiede zwischen statischen und dynamischen Einstellungen sowie höheren und geringeren örtlichen Details gut herausgestellt werden.

Aus dem Sportsegment wären weitere Sequenzen, die z. B. Wassersportarten beinhalten, von Interesse. In bewegten Wasserflächen lassen sich Codierartefakte zumeist recht gut erkennen. Wurde in der vorliegenden Arbeit ausschließlich mit Sportinhalten gearbeitet, empfiehlt es sich für spätere Tests auch weiteres fernsehtypisches Material einzubeziehen. Studioproduktionen wie bestimmte Showformate könnten beispielsweise aufgrund der vielfältigen Möglichkeiten des Einsatzes von Licht mehr Aufschluss über die zusätzlich benötigte Lichtmenge bei HFR-Produktionen geben.

7.5.2 Prozessoptimierung bei der Aufbereitung des Testmaterials

Über die praktische Durchführung des Downsamplings wurde in Kapitel 4.2.1 eingehend informiert. Die Quellsequenzen wurden mit Bildraten von 200 Hz oder 400 Hz produziert und daraus die niedrigeren Bildraten erzeugt. Es besteht die Möglichkeit, dass sich der subjektive Eindruck von nativ produzierten im Vergleich zu künstlich erzeugten Bildfrequenzen unterscheidet. Im Rahmen des 4EVER-Projekts wurden bereits diverse Sequenzen mit einem Kamerarig aufgezeichnet, das zwei identische Kameras kombiniert, die auf unterschiedliche Bildraten eingestellt waren. Eine Auswertung liegt derzeit noch nicht vor, kann

aber zum gegebenen Zeitpunkt Aufschluss über die Praktikabilität der hier verwendeten Downsamplingalgorithmen geben.

Ein Punkt der bereits in Kapitel 5.2.1 angedeutet wurde, ist die Vorgehensweise zur Erzeugung von 200 Hz-Sequenzen aus 50 Hz-Quellen durch Interpolation. Im Test wurde diese bekanntlich durch zwei Schritte realisiert, in denen die Bildrate jeweils verdoppelt wurde. Es stellt sich also die Frage, ob bei einer unmittelbaren Berechnung von drei Zwischenbildern evtl. andere Ergebnisse ermittelt werden könnten.

7.5.3 Temporal Scalability Coding

In Bezug auf die HEVC-Codierung mussten alle Sequenzen mangels spezieller Optionen des Encoders als einzelne Transportströme encodiert werden. Kapitel 2.3.2 veranschaulichte dagegen eine präferierte Methode zur Codierung von HFR-Quellen, welche in die Fähigkeit der zeitlichen Skalierbarkeit des HEVC-Standards ausnutzt. Ein Test, der die Auswirkungen dieser Option vor allem auch in Hinblick auf die Interoperabilität von En- und Decoder untersucht, wird für die weiterführende Forschung empfohlen.

7.5.4 Wiederholung der Tests mit UHD-Displays

Die vorangegangenen Untersuchungen wurden alle für eine örtliche Auflösung mit 1080 Zeilen, also dem HD-Format, durchgeführt. Der Standard ITU-R BT.2020 beinhaltet aber die erhöhte Bildrate ausschließlich für die verschiedenen UHD-Formate. Wie im Kapitel 4.2.4 erläutert, sollte der Aufwand für die Entwicklung von geeigneten Displays für diese Tests gering bleiben, weshalb hierfür die HD-Auflösung bevorzugt wurde. Zudem gab es zum Zeitpunkt des Auftrags zum Bau des HFR-Displays noch kein UHD-Panel am Markt, das mit 240 Hz angesprochen werden konnte. Inzwischen gibt es erste Hersteller, die solche Panels anbieten, woraus sich die Möglichkeit ergibt, UHD-HFR-Displays zu bauen, die durchgeführten Tests zu wiederholen und die Ergebnisse abzugleichen.

7.5.5 Implementierung zusätzlicher Bildoptimierungsoptionen

Insgesamt wurde festgestellt, dass selbst die maximale Bildrate im Test noch zu gering war, um Bewegungsunschärfen komplett zu vermeiden. Verdoppelt man die maximale Bildrate im Test, können sicherlich noch bessere Ergebnisse erzielt werden. Da zum jetzigen Zeitpunkt noch keine entsprechenden Geräte verfügbar sind, kann gegebenenfalls auf Techniken abseits einer nativen Darstellung jedes einzelnen Bildes zurückgegriffen werden. Eine Möglichkeit besteht in der Modulation der Hintergrundbeleuchtung. Diese Funktion ist bei vielen aktuellen Consumerdisplays bereits zuschaltbar und verspricht bis zu einer gewissen Grenze eine verbesserte Bewegungsdarstellung (vgl. 4.2.2). Durch eine Implementierung dieser Funktionalität in die für die IRT-Tests genutzten Displays könnte der subjektive Eindruck möglicherweise noch verbessert werden.

Literaturverzeichnis

Monografien und Sammelwerke

MÜCHER, Michael (2013): Broadcast Fachwörterbuch. 7000 Begriffe und Abkürzungen ; Fernsehen, Videotechnik, Audiotechnik, Multimedia. 17., überarb. u. erw. Aufl. ; Stand 2. Januar 2013. Hamburg: BET Michael Mücher.

ROBERTS, Alan (2009): Circles of confusion. Television, how it works, why it works this way, how to get the best from it. Genf: European Broadcasting Union.

Graue Literatur

International Telecommunication Union (2003): Expert viewing methods to assess the quality of systems for the digital display of large screen digital imagery in theatres. Recommendation ITU-R BT.1663. Hg. v. Radiocommunication Sector of ITU. International Telecommunication Union. Genf.

International Telecommunication Union (2007): Methodology for the subjective assessment of video quality in multimedia applications. Recommendation ITU-R BT.1788. Hg. v. Radiocommunication Sector of ITU. International Telecommunication Union. Genf.

International Telecommunication Union (2008): Subjective video quality assessment methods for multimedia applications. Recommendation ITU-T P.919. Hg. v. Telecommunication Standardization Sector of ITU. International Telecommunication Union. Genf.

International Telecommunication Union (2012): Methodology for the subjective assessment of the quality of television pictures. Recommendation ITU-R BT.500-13. Hg. v. Radiocommunication Sector of ITU. International Telecommunication Union. Genf.

International Telecommunication Union (2014): High efficiency video coding. Series H: Audiovisual and Multimedia Systems. Infrastructure of audiovisual services - Coding of moving video. Recommendation ITU-T H.265. Hg. v. Telecommunication Standardization Sector of ITU. International Telecommunication Union. Genf.

International Telecommunication Union (2015): Use of the high efficiency video coding (HEVC) standard for UHD TV and HDTV broadcasting. Recommendation ITU-R BT.2073-1. Hg. v. Radiocommunication Sector of ITU. International Telecommunication Union. Genf.

MikroM GmbH (Hg.) (2015): Kolibri HFR Display. User Manual. Berlin.

NOLAND, Katy C. (2014): The Application of Sampling Theory to Television Frame Rate Requirements. Research & Development White Paper WHP 282. British Broadcasting

Corporation. London. Online verfügbar unter <http://downloads.bbc.co.uk/rd/pubs/whp/whp-pdf-files/WHP282.pdf>, zuletzt geprüft am 12.02.2015.

NOLAND, Katy C.; TRUONG, Louise H. (2015): A Survey of UK Television Viewing Conditions. Research & Development White Paper WHP 287. British Broadcasting Corporation. London. Online verfügbar unter <http://downloads.bbc.co.uk/rd/pubs/whp/whp-pdf-files/WHP287.pdf>, zuletzt geprüft am 11.09.2015.

QUESTED, Andy; SALMON, Richard (2013): BBC interim report of the impact of camera frame rates and 50 Hz lighting. Hg. v. BBC Research & Development. London.

TANTON, Nick E. (2004): Results of a survey on television viewing distance. Research & Development White Paper WHP 090. British Broadcasting Corporation. London. Online verfügbar unter <http://downloads.bbc.co.uk/rd/pubs/whp/whp-pdf-files/WHP090.pdf>, zuletzt geprüft am 11.09.2015.

Hochschulschriften

SAUTER, Martin (2008): Einfluss von Bildformat- und Bildwechselfrequenzkonvertierung auf die Qualität von HDTV-Produktionen. Diplomarbeit. Fachhochschule Wiesbaden, Wiesbaden. Informationstechnologie und Elektrotechnik.

Internetdokumente

PINDORIA, Manish (2014): Higher frame rate television for future broadcasts at the Commonwealth Games - "But I already have a 100Hz TV... ?". Hg. v. BBC Research & Development. Online verfügbar unter <http://www.bbc.co.uk/rd/blog/2014-08-higher-frame-rate-television-for-future-broadcasts-at-the-commonwealth-games-but-i-already-have-a-100hz-tv>, zuletzt aktualisiert am 12.08.2015, zuletzt geprüft am 30.09.2015.

Ultra Video Group (Hg.): Test Sequences. Tampere University of Technology. Online verfügbar unter <http://ultravideo.cs.tut.fi/#testsequences>, zuletzt geprüft am 20.05.2015.

Vorträge und Präsentationen

CLARE-CHARRIER, Maryline (2015): HDR, HFR and BT.2020. Exploring content and visual comfort. EBU Production Technology Seminar 2015. Orange Labs. European Broadcasting Union, Technology & Innovation Department. Genf, 29.01.2015.

KOESLING, Hendrik (2011/2012): Augenbewegungen und visuelle Aufmerksamkeit. Visuelle Aufmerksamkeit und Blickbewegungen. Universität Bielefeld, AG Neuroinformatik. Bielefeld, 2011/2012. Online verfügbar unter <http://www.techfak.uni-bielefeld.de/~ihko-esli/vab2011/07-vab2011-hk-augenbewegungen.pdf>, zuletzt geprüft am 11.05.2015.

SUGAWARA, Masayuki (2011): Psychophysical requirements for higher frame rates in future television. DCS 2011 - Advances in Image and Sound, 3D, 4K and beyond. NHK. NAB Show, SMPTE. Las Vegas, 10.04.2011.

Zeitschriftenaufsätze

BROSS, Benjamin; MARPE, Detlev; SCHWARZ, Heiko (2013): "High-Efficiency Video Coding" - (HEVC)-Standard. Vorstellung des neuen Kompressionsstandards. In: *FKT 67* (1-2), S. 56–63.

DALY, Scott (2001): Engineering Observations from Spatiovelocity and Spatiotemporal Visual Models. In: *Vision Models and Applications to Image and Video Processing*, S. 179–200.

DRIESNACK, Dagmar; GIERLINGER, Friedrich; SCHÄFER, Rainer (2014): Höhere Bildwechselfrequenz für Ultra-HD? In: *FKT 68* (12), S. 641–644.

GOLDMAN, Matthew; LITWIC, Lukasz; BAUMANN, Olie (2015): Ultra-HD Content Acquisition and Exchange. Using HEVC Range Extensions. In: *SMPTE Motion Imaging Journal* 124 (3), S. 28–36.

NYQUIST, Harry (1928): Certain Topics in Telegraph Transmission Theory. In: *Transactions of the American Institute of Electrical Engineers* 47 (2), S. 617–644.

SHANNON, Claude Elwood (1948): A Mathematical Theory of Communication. In: *The Bell System Technical Journal* 27 (3, 4), S. 379-423, 623-656.

Anhang

Bewertungsbogen für subjektive Tests.....A-2

Durchschnittliche PSNR-Werte je Sequenz.....A-3

Subjective Test

Screen

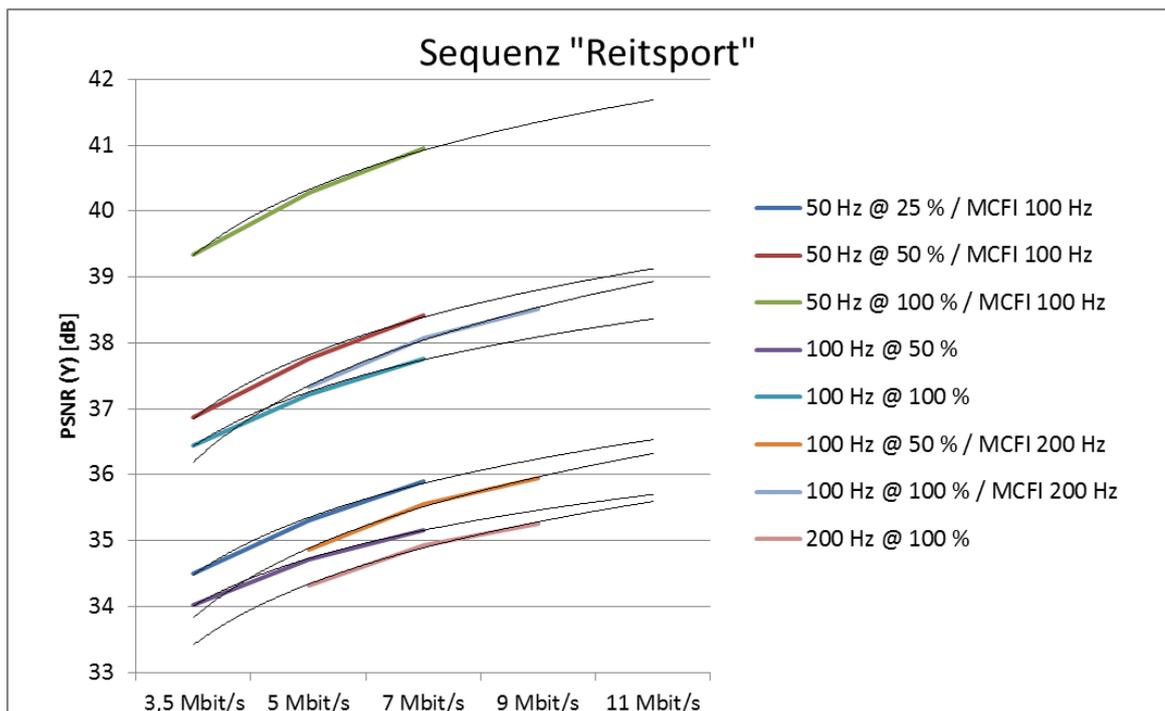
Name: Gruppe Nr.: Session: 1 Sitz (ankreuzen):

Datum: Zeit:

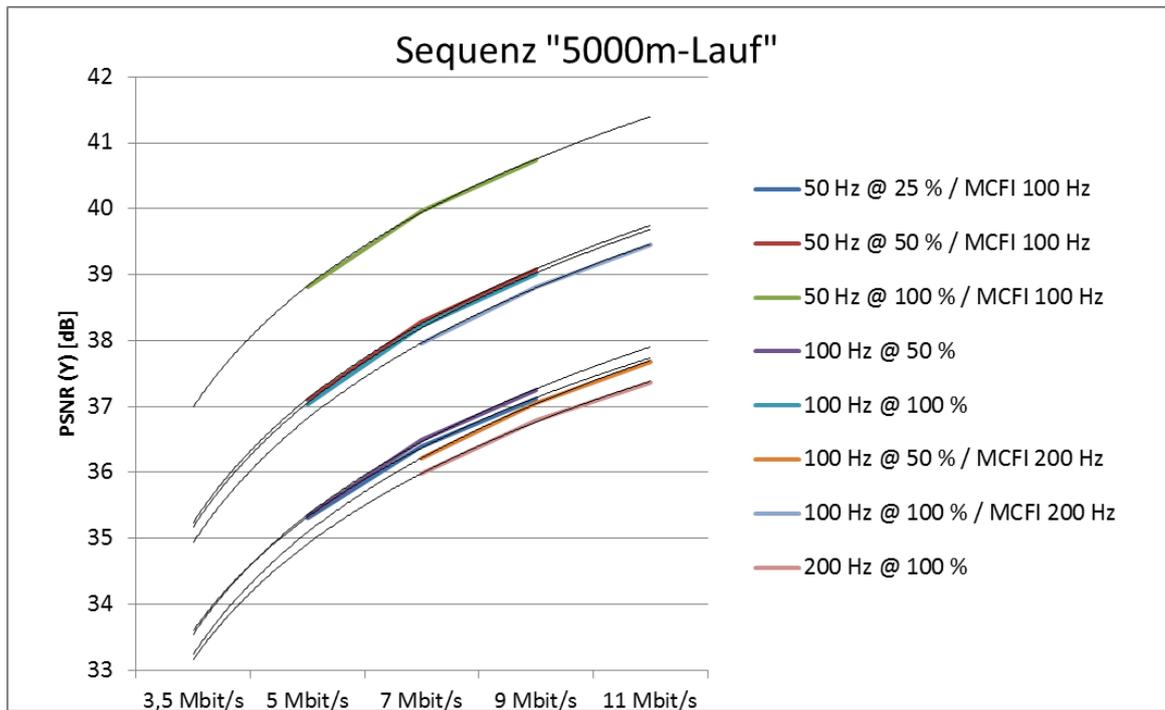
	Beispiel	1 Hürden	2 Barren	3 Dreisprung	4 Reiten	5 5000m	6 Barren	7 Reiten	8 Dreisprung
	A B	A B	A B	A B	A B	A B	A B	A B	A B
excellent									
good									
fair									
poor									
bad									

	9 Hürden	10 Dreisprung	11 5000m	12 Reiten	13 Dreisprung	14 Hürden	15 Barren	16 Dreisprung
	A B	A B	A B	A B	A B	A B	A B	A B
excellent								
good								
fair								
poor								
bad								

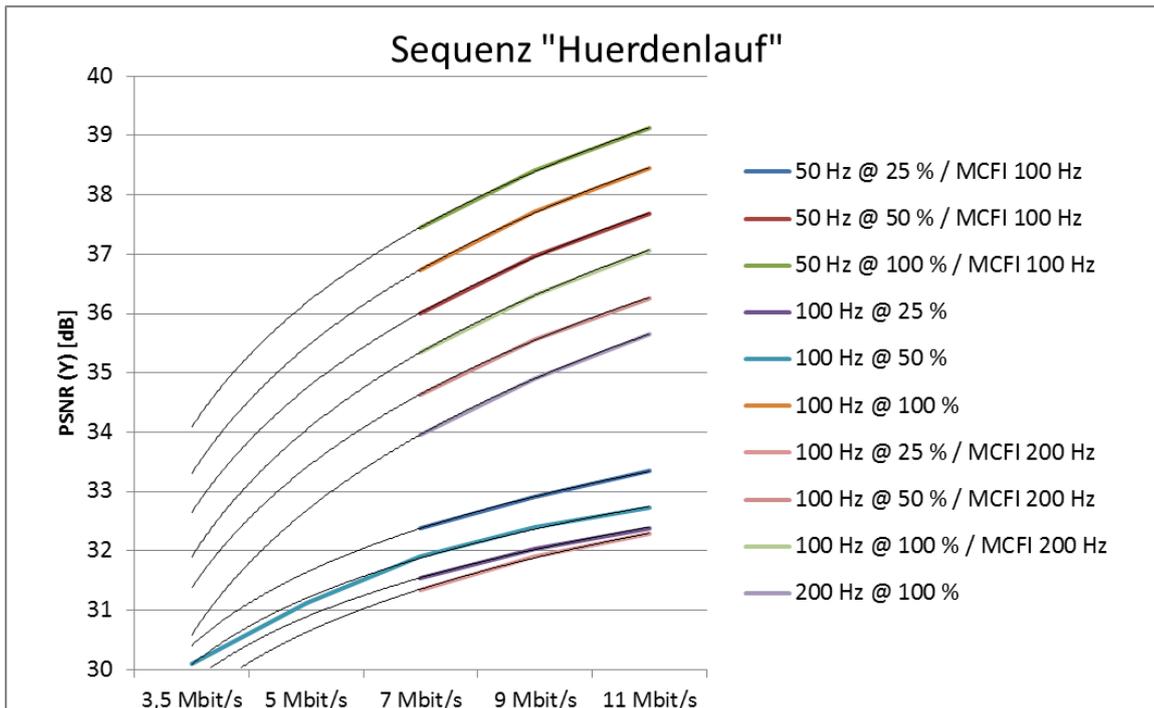
Anhang 1: Bewertungsbogen für subjektive Tests



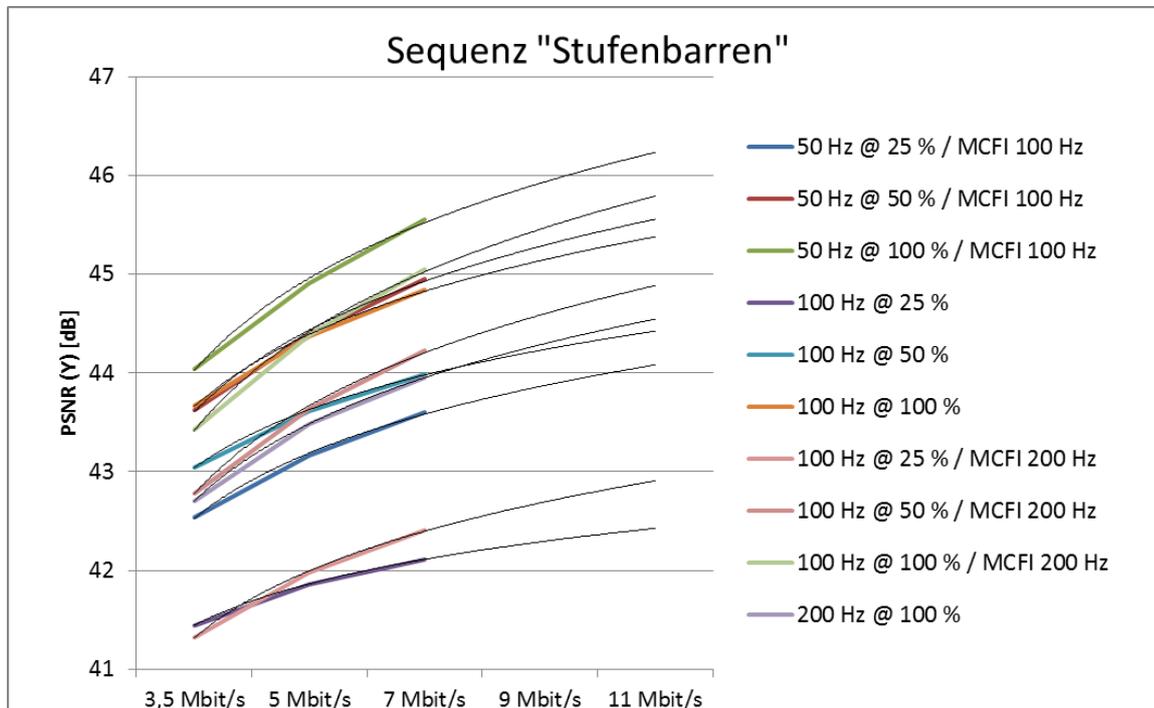
Anhang 2: Durchschnittliche PSNR-Werte der Sequenz "Reitsport"



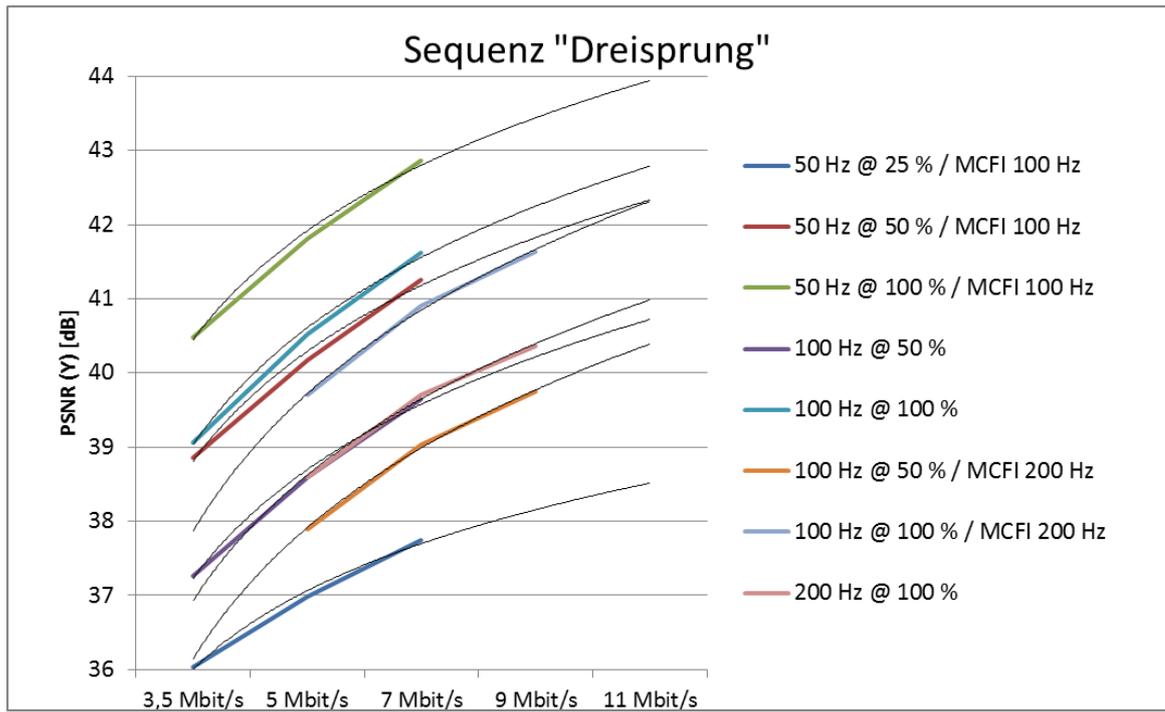
Anhang 3: Durchschnittliche PSNR-Werte der Sequenz "5000m-Lauf"



Anhang 4: Durchschnittliche PSNR-Werte der Sequenz "Hürdenlauf"



Anhang 5: Durchschnittliche PSNR-Werte der Sequenz "Stufenbarren"



Anhang 6: Durchschnittliche PSNR-Werte der Sequenz "Dreisprung"

Selbstständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Stellen, die wörtlich oder sinngemäß aus Quellen entnommen wurden, sind als solche kenntlich gemacht.

Diese Arbeit wurde in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegt.

München, den 30. Oktober 2015

Felix Oberhardt