



Quelle: Dmitri Melekhin (Stock)

# Beyond Pixel Count and Video Resolution

Ein bildinhaltsunabhängiges Messverfahren zur Video-Auflösungsbestimmung

Bildqualität und Video-Auflösung sind wichtige, aber auch aufwendig zu bestimmende Parameter in einem Ende-zu-Ende Video-Workflow. Der Beitrag wird einen Überblick über Einflussgrößen auf diese Qualitätsmerkmale geben und betrachtet wichtige Prozessschritte von der Bildgebung bis zum Anzeigegerät. Dabei steht im Vordergrund, wie man die Sicht auf die Bildqualität in den Bereichen Bild-Akquise, Postproduktion und Distribution vereinheitlichen kann, um eine bessere Customer Experience bieten zu können. Weiterhin wird ein neues bildinhaltsunabhängiges Messverfahren zur Video-Auflösungsbestimmung vorgestellt, das ohne Mess-Target auskommt.

## Bildqualität

Wenn man von Videoauflösung spricht, werden oft die Zeilenzahlen 576p, 720p, 1080p und 2160p als Kennzeichnung für eine bestimmte Videoauflösung genutzt. Es gibt aber einen großen Unterschied zwischen der Pixel-Anzahl, die sich aus dem Bildseitenverhältnis und der Anzahl der Zeilen ergibt und der tatsächlichen

► Image quality and video resolution are important but also complex parameters to determine in an end-to-end video workflow. This article will give an overview of the factors influencing these quality characteristics and will look at important process steps from imaging to the display device. The focus will be on how to unify the view of image quality, in the areas of image acquisition, post-production and distribution, in order to provide a better customer experience. Furthermore, a new image content-independent measurement method for video resolution determination is presented, which does not require a measurement target.

Videoauflösung. Die Videoauflösung gibt die höchste Frequenz an, die mit einem Bildsignal übertragen werden kann. Auch heute noch werden Linienpaare pro Bildbreite oder -höhe für die Auflösungsbestimmung genutzt. Hierbei wird ein Testsignal abgetastet und dann die MTF (Modulation Transfer Function) bestimmt.

Nutzt man digital erzeugte Testsignale, um die Performance eines Filmscanners oder einer Kamera zu bestimmen, ist der Aufwand erheblich. Es muss ein geeignetes Test-Target vorhanden sein und im besten Fall eine automatische Auswertung der Testergebnisse vorgenommen werden. Bei den meisten Filmscannern werden Flächensensoren genutzt. Bei digitalen Kameras hat sich das Bayer-Pattern in unterschiedlichen Varianten durchgesetzt, bei dem pro Pixel jeweils nur eine Farbinformation abgetastet wird. Die fehlenden zwei Farben werden dann in der Kamera oder in der Postproduktion aus den Umgebungspixeln interpoliert.

Diese Interpolation von Farbwerten bringt eine Ungenauigkeit mit sich, die wesentlich in der Größe durch die bei der Aufnahme entstandenen Rauschteile und dem

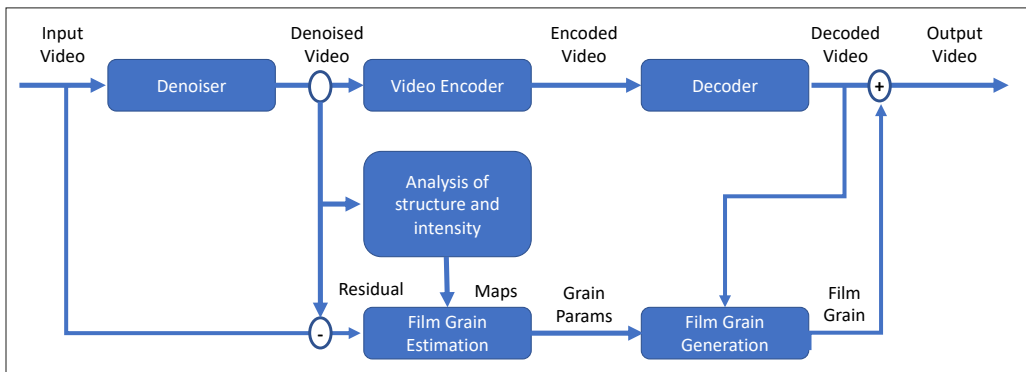


Abbildung 1:  
AV1 Film Grain  
Synthesis Workflow

verwendeten De-Bayering-Algorithmus bestimmt wird. Flächensensoren haben dagegen einen Bildqualitätsvorteil, da hier keine Interpolation notwendig ist.

Wird das Bildsignal in einem unkomprimierten RAW-Format aufgenommen und werden diese RAW-Daten vor dem De-Bayering korrekt denoised, so erhält man nicht nur ein im Rauschverhalten optimales Bildsignal, sondern auch eine erheblich höhere Bildauflösung, die sich aus dem dann optimalen Interpolationsvorgang des De-Bayering ergibt. Die Ungenauigkeiten, die das Rauschen bei der Interpolation erzeugt, können so auf ein Minimum reduziert werden. Die Nutzung von verlustbehafteten RAW-Formaten ist aus patentrechtlichen Gründen sehr gering, wenn auch diese Patentvergabe hin und wieder in Frage gestellt wird.

Die höchste Bildqualität und Auflösung ergibt unkomprimiertes RAW z. B. als DNG-Sequenz in einer Kombination mit einem hochwertigen RAW-Video-Denoising außerhalb der Kamera (z. B. 9x7 Kamera Pawel Achtel<sup>1)</sup> und IRIS.RAW<sup>2)</sup>). Wird das (RAW) Kamerasignal verlustbehaftet komprimiert, so werden die Rauschanteile moduliert und es entsteht ein Shift der Frequenzen in andere Frequenzbereiche. Natürlich werden Denoising-Verfahren auch in Kameras verwendet. Diese kommen aber unter anderem aus Performance-Gründen der Kameraplattformen nicht an die Bild-Qualität der außerhalb der Kameras angewandten Verfahren heran. Ein bezüglich der Bildauflösung sehr schlechtes Verfahren ist es, schon in der Kamera Grain-Anteile fest dem Bildsignal zu überlagern, um dem DP die Möglichkeit zu geben, einen Filmlook fest vorzugeben. Hierbei können gegebenenfalls entstandene Denoising-Artefakte mit Hilfe der künstlich erzeugten Grain-Anteile überdeckt werden. Infolge dessen wird aber die Bildauflösung und der Dynamikumfang nachhaltig negativ beeinflusst.

## Die Auswirkungen von Rauschen auf die Video-Codierung

Videoräuschanteile und Grain bei der Distribution führen zu einer verminderten Videoqualität, verminderten Videoauflösung und zu erhöhten Bandbreitenanforderungen. Videocodecs erkennen und kodieren überflüssige Informationen. Je höher die Komprimierung ist, desto mehr der weniger wichtigen Informationen werden entfernt. Rauschen im Videoinhalt reduziert die zeitliche Redundanz und zwingt den Encoder dazu, mehr Originalinhalte zu entfernen. Die Fähigkeit des Encoders,

Bewegungen in einer Videosequenz zu erkennen und die Bewegungsvektoren korrekt zu definieren, kann durch Rauschen stark beeinflusst werden. Die Genauigkeit der Bewegungsvektoren ist für einen qualitativ hochwertigen Videocodierungsprozess unerlässlich. Wenn Rauschen vorhanden ist, ist der Encoder gezwungen, das Rauschen als Bewegung zu interpretieren, was zu fehlerhaften Bewegungsvektoren führt und der Grund für die erhöhten Bandbreitenanforderungen bei verrauschten Videosignalen ist.

Wenn der Intra/Inter-Entscheidungsmechanismus erkennt, dass die Bewegungsschätzung aufgrund des gegebenen Rauschniveaus zu schwach ist, wird nur eine Intra-Vorhersage genutzt. Ein Verlust der Inter-Vorhersage führt zu einem Qualitätsverlust an den entsprechenden Stellen im Video.

Eine Anwendung eines Highfidelity Video Noise Managements vor dem Codierungsprozess führt zu einer höheren Videoauflösung, einer besseren Bildqualität und optimierten Video-Bitraten. Auch bei Filmen mit hohen Grain-Anteilen geht der Filmlook nicht verloren, wenn das Denoising-Verfahren dazu in der Lage ist. Dieses ist aber nicht für alle am Markt erhältlichen Video Noise Management-Lösungen der Fall.

Ist noch in der Post eine dem Celluloid-Film ähnliche Grain-Struktur zu erkennen, so bleiben vom Korn beim End-Gerät nur noch ein tieffrequentes Störsignal, Banding und Blocking-Artefakte übrig. Erst bei ungewöhnlich hohen Bitraten in der Distribution würde sich dieser Effekt ohne Video Noise Management leicht verbessern. Unabhängig von der Encoder-Technologie ist es also sinnvoll, für alle Distributionskanäle eine Übertragung der Rausch- und Grainanteile über ein Metadatenmodell vorzusehen. Eine entsprechende Spezifikation ist in Vorbereitung. AV1 ist die erste Video-Encoder-Technologie, bei der „Film Grain Synthesis“ verpflichtend im Decoder vorhanden ist. Bei der Referenzimplementierung (libaom) aber auch anderen AV1-Encodern ist bereits ein Video Denoising als Pre-Processing vor dem Encoder mit implementiert. Allerdings ist die Qualität dieses Denoising für professionelle Anwendung oft nicht ausreichend. Dabei wird vor dem Video-Encoding das Filmmaterial einem Denoising unterzogen und erst nach der Dekodierung wird das Grain wieder über Parameter hinzugefügt.

1) Pawel Achtel, 9x7 Camera: <http://24x7.com.au>

2) Wavelet Beam, RAW Video Denoising: <https://waveletbeam.com/index.php/iris>



**DIRK HILDEBRANDT**

CTO Wavelet Beam

➔ [www.waveletbeam.com](http://www.waveletbeam.com)

Mit „Film Grain Synthesis“ (Bild 1)<sup>3)</sup> ist es möglich, dass ein Film-Look für Broadcast und OTT erzeugt werden kann, der weit über das hinausgeht, was man heute an den Bildschirmen erleben kann. Auch hier gilt, dass das verwendete Denoising hochqualitative Ergebnisse liefern sollte und die Gefahr besteht, dass hinzugefügte Rausch- und Grainanteile nur über eine geringe Bildauflösung oder Artefakte des Denoisings hinwegtäuschen sollen.

Der große Vorteil, wenn man Bildinhalt und Rauschanteile voneinander trennt, ist es, dass eine bis zu 30-prozentige Bitrateneinsparung<sup>4)</sup> bei der Distribution erreicht werden kann. Kostenvorteile bei der Distribution, aber auch eine bessere Streaming-Fähigkeit können so erzielt werden.

Die Nutzung von „Film Grain Synthesis“ schon beim Kamerasignal oder beim Intermediate-Format würde zu einer wesentlich höheren Bildqualität bei der Distribution führen. Hier gilt es, die schon vorhandenen Spezifikationen für die unterschiedlichen Workflowschritte zu adaptieren und ein übergreifendes Verständnis für diese Problematik zu etablieren.

### Videoauflösung

Die Videoauflösung ist nicht an eine Zeilenzahl gebunden, sondern sie kann von Bild zu Bild oder Szene zu Szene variieren. Dabei erreicht sie in den wenigsten Fällen die durch die Pixelanzahl und dem Bildseitenverhältnis sich ergebende maximale mögliche Auflösung. Die jeweils genutzte Videokamera und die Qualität der Objektive bilden dabei die Grundlage dessen, welche maximale Videoauflösung erreicht werden kann. Die Videoauflösung wird weiterhin durch ein fehlendes oder unzureichendes Video Noise Management und die gewählten Lossy Video Encoding-Formate für die Akquisition und Postproduktion bestimmt. Jeder zusätzliche lossy Re-Encoding-Schritt kann dabei die Videoauflösung nachhaltig beeinflussen. Vorsicht ist auch bei dem technischen Begriff „Visually Lossless Compression“ geboten. Auch wenn man ohne Messtechnik in der Bildakquisition keine wesentliche Bildverschlechterung wahrnehmen kann, so werden doch die Rauschanteile in andere Frequenzbereiche transformiert. Die Auswirkungen auf die Bildqualität ergeben sich dann in der Postproduktion oder Distribution der Videosignale. Die Rauschanteile können dann nur noch bedingt zur Noise-Analyse des Videosignals genutzt werden und mindern die Qualität der Ergebnisse des Video Noise-Managements. Findet in der Postproduktion kein Video Noise-Management statt, kann der verlorengegangene Headroom der Rauschanteile nicht für eine zusätzliche Kontrast- und HDR-Optimierung genutzt werden.

Bildqualitätsmessverfahren wie PSNR, VMAF oder SSIM werden in der Distribution genutzt, um das Intermediate-Format mit dem Distributionsformat zu vergleichen. Bei den für die Distribution vorgesehenen Bitraten wird das Bildsignal schon so hoch komprimiert, dass Messverfahren wie PSNR nur eingeschränkt brauchbare Ergebnisse liefern. VMAF wurde von Netflix initial dafür entwickelt, um für jeden Content die optimale Bitrate zu finden (Per Titel/ Per Shot Encoding). Hohe Bewegungsanteile, Grain- und Rauschanteile im Video führen zu hohen Bandbreitenanforde-

rungen und digital erzeugte Inhalte ohne Rauschen hingegen zu niedrigen. VMAF bietet auch bei hohen Kompressionsraten noch Qualitätsabstufungen, dennoch kann es aber nur eingeschränkt für Encoder-Vergleiche genutzt werden. Die VMAF-Messergebnisse sind über Manipulationen des Bildsignals im Encoding-Prozess beeinflussbar<sup>5)</sup>. Sicherlich wird an einer robusteren VMAF-Version gearbeitet, aber eine Manipulation der Ergebnisse kann zurzeit nur bedingt ausgeschlossen werden. Weiterhin wird daran gearbeitet, eine Lösung dafür zu finden, wie VMAF bei der Nutzung von „Film Grain Synthesis“ eingesetzt werden kann.

Bei Qualitätsbeurteilungen wird das Intermediate-Format als Referenz genutzt. Die von der Zeilenzahl unabhängige Bildauflösung des Intermediate-Formates geht dabei aber nicht mit in das Messergebnis ein. Bei der Beurteilung der VMAF-Werte ist also nicht bekannt, für welche Videoauflösung die Messergebnisse gelten. Messergebnisse des im Folgenden vorgestellten Messverfahrens zur „bildinhaltsunabhängigen Messung von Video-Auflösungen“ zeigen, dass die Bildauflösung zwischen 60 Prozent und 100 Prozent der durch die Zeilenzahl maximal möglichen Auflösung schwankt.

Weiterhin ist bei einer VMAF-Auswertung nicht bekannt, welche Videoauflösung das Video-Testmaterial besitzt. Ein wesentlich in der Videoauflösung reduziertes Referenzmaterial führt dazu, dass die getesteten Encoder am Eingang nur eine reduzierte Frequenzbandbreite verarbeiten müssen. Hieraus ergeben sich tendenziell zu gute VMAF-Ergebnisse, da der Encoder mit einer reduzierten Bandbreite am Eingang bessere Encoding-Ergebnisse liefert. Ist dann eine Postproduktionslinie in der Lage, ein Intermediate-Format mit höheren Video-Auflösungen zu erzeugen, so könnten die Qualitätsgewinne mit schlechteren VMAF-Werten bestraft werden, da die genutzten Encoder mit der nun erhöhten Frequenzbandbreite am Eingang schlechtere Ergebnisse liefern könnten.

### Bildinhaltsunabhängiges Messverfahren zur Video-Auflösungsbestimmung

Das Einbringen von Mess-Targets in das Bildsignal ist aufwendig und spiegelt auch nicht die im Alltag zu erreichenden Videoauflösungen wider. Ein neues bildinhaltsunabhängiges Messverfahren zur Video-Auflösungsbestimmung, das in die Bildsignalkette eingebunden werden kann, wurde von der Firma Wavelet Beam entwickelt und soll im Folgenden vorgestellt werden.

Das Messverfahren kann sowohl auf RAW-Videodaten als auch auf bereits de-bayertes Videomaterial (Bild 2: Pos 2) angewendet werden. In einem weiteren Schritt wird das Videosignal einer Transformation (Bild 2: Pos 3) in den Frequenz- oder Zeitfrequenzbereich unterzogen. Da die Videoauflösung in den Farbkanälen unterschiedlich sein kann, sollte die Transformation pro Farbkanal vorgenommen werden. Dann werden die Koeffizienten analysiert (Bild 2: Pos 4) und nach Frequenzanteilen sortiert. In einem fünften Schritt (Bild 2: Pos 5) wird dann mit Hilfe von z. B. statistischen Methoden die obere Grenzfrequenz ermittelt. Die horizontale und vertikale Pixelanzahl definiert dabei die maximale Videoauflösung. Eine zusätzliche Normierung (Bild 2: Pos 6) der Messergebnisse ermöglicht den Einsatz des Messverfahrens an unterschiedlichen Stellen im Workflow.

Ein wichtiges Alleinstellungsmerkmal ist dabei die Bildinhaltsunabhängigkeit des Messverfahrens. Das bedeutet, dass auch eine sehr kleine Anzahl von hochfrequenten Bildanteilen bereits für die Bestimmung der Bildauflösung

3) Andrey Norkin and Neil Birkbeck, Film Grain Synthesis for AV1 Video Codec: [https://norkin.org/pdf/DCC\\_2018\\_AV1\\_film\\_grain.pdf](https://norkin.org/pdf/DCC_2018_AV1_film_grain.pdf).

4) Andrey Norkin & Liwei Guo from Netflix, DEMUXED 2019: <https://www.youtube.com/watch?v=JU-sE14U9xM&t=299s>

5) Zhi Li, Netflix, On VMAF's property in the presence of image enhancement operations

ausreicht. Ein Target im Bild ist nicht notwendig und ermöglicht dadurch eine kontinuierliche Überwachung des Bildsignals in der Produktion. Bei der Bildakquisition kann mit diesem neuen Verfahren für jede Kameraposition die erreichte Bildschärfe beurteilt werden oder eine automatische QA von externen Videomaterial unterstützen.

In der Distribution hingegen kann dieses Verfahren zur Auswahl der Testclips genutzt werden. So besteht nicht mehr die Gefahr, dass bei Bildqualitätsbeurteilungsverfahren wie VMAF, SSIM oder PSNR die Messergebnisse eine zu geringe Bildauflösung widerspiegeln. Dieses wiederum hat natürlich auch Auswirkungen auf die Beurteilungen der VMAF- oder PSNR-Werte, denn diese können jetzt auf die tatsächliche Videoauflösung bezogen werden. Eine VMAF-Auswertung könnte in Zukunft also so aussehen: Der Encoder A hat gegenüber dem Encoder B bei einer Bitrate von 13200 Kbit/s eine um 1,5 VMAF-Punkte bessere Bildqualität bei einer mittleren Videoauflösung von 3,24K horizontal (1,63K vertikal).

Das ist wichtig, da in der Praxis teilweise Empfehlungen ausgesprochen werden, die besagen, dass man das Intermediate-Videomaterial gerne noch weiter in der Bitrate reduzieren kann, da die VMAF-Messergebnisse bei der Distribution nur eine geringe Änderung zu dem geringer komprimierten Videomaterial liefern würde. Eine solche Aussage ist äußerst kritisch, da sich die VMAF-Differenzwerte auf die wirkliche und gemessene Bildauflösung beziehen sollten und nicht nur auf die Anzahl der Pixel im Bild.

Wird ein Intermediate-Format einer zu hohen Kompression unterzogen, geht auch Auflösung verloren und eine weitere Reduzierung der Bitrate des Intermediate Formats ergibt nur eine geringe VMAF-Differenz. Dieses aber tatsächlich nur bei einer dann wesentlich reduzierten Bildauflösung. Leider fehlt bei fast allen VMAF-Auswertungen die Angabe des Messfehlers und damit die Möglichkeit, auch die VMAF-Ergebnisse korrekt auszuwerten.

## Endgeräte

Endgeräte wie z. B. UHD-Fernseher und hochauflösende Second Screen Devices sind in Deutschland schon weit verbreitet. Die mögliche Bildqualität wird aber durch die IP-Infrastruktur in vielen Wohnung und Häusern sowie durch weitere Bildverarbeitungsschritte nach dem Dekodieren des Bildsignals weiter gemindert. Bei Fernsehgeräten sind die von den Herstellern gewählten Default-Werte für Farbe, Schärfe und Kontrast oft viel zu hoch gewählt. Weiterhin ist das Denoising und De-Blocking als Default voreingestellt und hat oft nur eine sehr geringe Qualität<sup>7)</sup>. Grundsätzlich muss man aber auch festhalten, dass ein Denoising am Ende der Distribution nach dem Decoder die schlechteste aller Optionen ist. Werden verrauschte Bildsignale für die Distribution encodiert, so benötigen sie eine wesentlich höhere Bitrate, um einen vergleichbaren Qualitätseindruck im Gegensatz zu rauscharmen Signalen zu erreichen. Die von Netflix angegebene, durch ein Video Noise-Management mögliche Reduzierung der benötigten Bitrate von bis zu 30 Prozent konnte bei Wavelet Beam in der Praxis bestätigt werden.

Es benötigt vom Anwender einiges an Fachwissen, um einen Fernseher optimal einzustellen. Eine Hilfe stellt dabei der „Filmmaker Mode“<sup>8)</sup> dar, der über nur einen

Menüpunkt oder eine Taste der Fernbedienung einige der Default-Einstellungen deaktivieren kann. Ein Automatikmode, der anhand des Genres oder eines DVB-Deskriptors den „Filmmaker Mode“ pro Sendung automatisch aktivieren oder deaktivieren kann, wäre sehr wünschenswert. Durch einen dynamischen Film Maker Mode könnten so Sportsendungen weiterhin von den höheren Bildwiederholfrequenzen profitieren aber eben auch Spielfilme, die für eine Wiedergabe von 24 fps vorgesehen sind, den so genannten Soap-Opera-Effekt vermeiden. Da Fernseher und Second Screen Devices oft über WLAN angeschlossen sind, besteht die Gefahr, dass der Abstand zwischen Repeater oder Router zu groß ist und nur eine sehr eingeschränkte Bitrate zur Verfügung steht. Eine um einige Prozentpunkte zu niedrigere WLAN-Bitrate kann den Ausschlag geben, ob ein Programm in maximal UHD oder doch nur in 1080p empfangen werden kann. Für Broadcast Services besteht diese Gefahr grundsätzlich nicht, allerdings muss erst noch nachgewiesen werden, welche Bildqualität die Encoder-Strecken tatsächlich liefern.

Das Live Video-Encoding bietet in der Regel zurzeit noch wesentlich schlechte Enkodier-Effizienz gegenüber dem Offline-Encoding für OTT. Zurzeit liegt die Bildqualität für OTT sogar über der der Broadcast-Channels. Eine höhere Encodierdauer beim Offline-Encoding ist dabei ausschlaggebend für die Encodier-Effizienz. Da OTT-Services oft auch auf Second Screen Devices konsumiert werden und daher der Abstand zwischen Bildschirm und Betrachter unter 70 Zentimeter sinkt, kann diese höhere Videoqualität auch vom Betrachter wahrgenommen werden. Je kleiner der Betrachtungsabstand, desto höher ist die mögliche wahrnehmbare Auflösung bei Video. Hierzu kommt noch, dass auf Second Screen Devices nur eine geringe Manipulation des Bildsignals vorgenommen wird. Auch hier werden Schwarz- und Kontrastwerte verändert, aber Echtzeit Video-Denoising-Verfahren werden bei Second-Screen Devices eher nicht zu finden sein und somit muss mit einer weiteren Reduzierung der Videoauflösung bei OTT nicht gerechnet werden.

Der DVB-Broadcast steht bezüglich der Sendeinhalte und der Bildqualität in der Konkurrenz zu den OTT-Angeboten. Ein Großteil der Jugendlichen in Deutschland nutzt fast ausschließlich OTT-Dienste, an denen sich der Broadcast-Bereich messen lassen muss. Nur ein sehr geringer Teil der Broadcast-Inhalte sind tatsächlich Live. Deshalb kann heute schon durch hocheffiziente HPC Video Noise-Management-Lösungen ein Großteil des Programms homogenisiert werden. Eine Reduzierung der Rauschanteile bewirkt dann bei anspruchsvollen Szenen, dass die Bildqualität dennoch nicht abfällt. Das Video Noise-Management sollte natürlich nicht das Bildsignal einfach nur glätten und damit auch den Filmlook zerstören. „Film Grain Synthesis“ wie bei

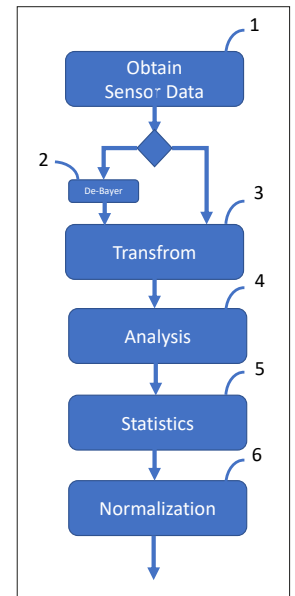


Abbildung 2: Wavelet Beam: Ein Bildinhaltsunabhängiges Messverfahren zur Video-Auflösungsbestimmung

6) Wavelet Beam, Copyright Q2/2023, Image content-independent measurement method for video resolution determination

7) HDTVTest, TV Settings That Should be ILLEGAL: <https://www.youtube.com/watch?v=uGfT746TJu0&t=143s>

8) FILMMAKER MODE: <https://filmmakermode.com>

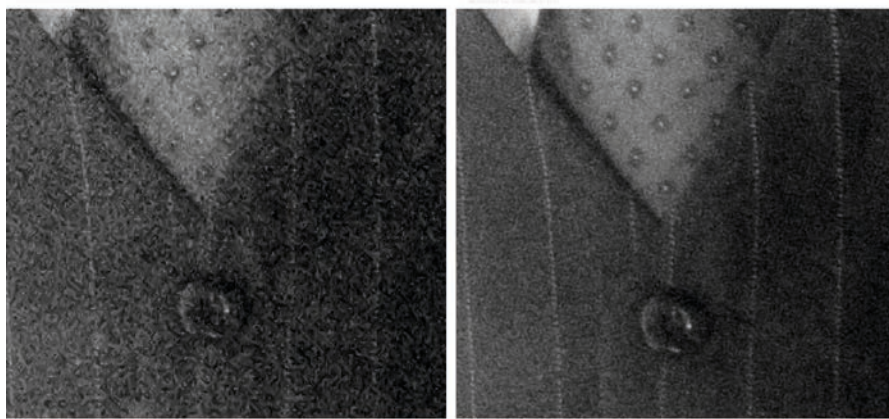


Abbildung 3:  
Video Noise Management  
kombiniert mit KI-Upscaling.

Quelle des Bildmaterials:  
NETFLIX, 4k Referenzfilm „Meridian“ und  
8K AI Up-Scaling

AV1 bietet heute schon die Verfügbarkeit einer End-zu-End Distributions-Kette für alle AV1-fähigen Endgeräte. Auch ohne „Film Grain Synthesis“ sollte ein hochwertiger Bildeindruck gewährleistet werden. Für 24/7 Broadcast-Kanäle ist die Performance aber auch der Energiebedarf sehr wichtig. Ein Load-Balancing zwischen CPU und GPU ist dabei die Schlüsseltechnologie. Nur wenn auf einer GPU mehrere Files gleichzeitig verarbeitet werden können und auch mehrere GPU's pro Server genutzt werden, kann die benötigte Hardware effizient genutzt werden.

### KI-basiertes Video Up-Scaling

Traditionelle Video-Upscaling-Filter (Bicubic oder Lanczos) verwenden Interpolationstechniken, um die Bildgröße zu erhöhen, indem sie vorhandene Pixelwerte extrapolieren und die Lücken zwischen den Pixeln füllen. Diese Methode kann zu einem Verlust von Bildqualität und Details führen, da sie keine Informationen über das tatsächliche Bild enthält. Im Gegensatz dazu nutzt KI-basiertes Video-Upscaling maschinelles Lernen, um ein tiefes Verständnis des zugrundeliegenden Bildes zu entwickeln und die fehlenden Pixel basierend auf diesem Wissen zu generieren. Dies wird erreicht, indem große Datenmengen an hochwertigen Bildern und Videos verwendet werden, um ein neuronales Netzwerk zu trainieren, das in der Lage ist, komplexe Muster zu erkennen und das Bild entsprechend aufzufüllen. Durch diese Methode kann die KI genauere Vorhersagen treffen und eine höhere Bildqualität und Detailschärfe erreichen, als herkömmliche Upscaling-Filter. Darüber hinaus, kann die KI auch spezifische Merkmale eines Videos oder einer Szene erkennen und diese entsprechend verbessern, um ein besseres Gesamtbild zu erzielen. Insgesamt bietet KI-basiertes Video-Upscaling ein höheres Maß an Präzision und Qualität als herkömmliche Methoden.

Es ist wichtig, vor der KI-basierten Hochskalierung ein Video-Denoising durchzuführen, da das Rauschen in einem Video es den Hochskalierungsalgorithmen erschweren kann, qualitativ hochwertige Ergebnisse zu erzielen. Wenn ein Video Rauschen enthält, kann der Hochskalierungsalgorithmus das Rauschen fälschlicherweise als tatsächliche Bilddetails interpretieren.

Durch Entfernen des Rauschens aus dem Video vor der Hochskalierung kann dem Hochskalierungsalgorithmus ein sauberes und hochwertiges Videosignal zur Verfügung gestellt werden, das zu besseren Hochskalierungsergebnissen und einer besseren visuellen Qualität des hochskalierten Videos führt. Bei der Kombination von KI-basierten UP-Scaling und Video Noise-Management kann es zu Superresolution-Effekten kommen, die im Bild 3 eindrucksvoll zu sehen sind.

### Zusammenfassung

Die Video-Auflösung des zur Erzeugung der Distribution-Encodes genutzten Intermediate-Formates ist in der Regel nicht bekannt. Oft ist hier die tatsächliche Video-Auflösung um mehr als 30 Prozent niedriger als die, die sich rechnerisch aus der Pixelanzahl und dem Bildseitenverhältnis ergeben würde.

Das oben vorgestellte Verfahren zu Video-Auflösungsbestimmung kann dazu genutzt werden, um Demomaterial auszuwählen, „Per Titel/Shot Encoding“ zu unterstützen oder auch, um den Schärfepeak einer Kameraeinstellung zu bestimmen.

Weiterhin ist es überall dort hilfreich, wo Betrachtungsabstände und deren Einfluss auf die Wahrnehmung von z. B. 4K/8K HDR-Videoqualität wissenschaftlich untersucht werden sollen.

Es wurde gezeigt, welche unterschiedlichen Faktoren beim Ende-zu-Ende Workflow die Bildqualität beeinflussen kann. Die fehlende Angabe der Videoauflösung und Fehlertoleranz bei der Bildqualitätsbeurteilung macht es vielen Marketingabteilungen leicht, ihre Produkte gegenüber der Konkurrenz durch vermeidliche Leistungsmerkmale abzugrenzen.

Hierzu passt sehr gut ein Zitat von Steve Yedlin:

„The dominant narrative is that as long as we know this one number. We don't need to know anything else about the complex pipeline that defines our master image, because that one number, we believe somehow corresponds to an audience's perceptual experience of sharpness and clarity.“

Ein gutes Broadcast- und VOD-Produkt zeichnet sich durch viele Faktoren aus, aber eine gute Audio- und Videoqualität sollten immer das Ziel sein und als Qualitätsmerkmal dienen. Nicht nur „weil wir es können“, sondern auch weil es wirtschaftlich sehr attraktiv ist. ➤

Linkverzeichnis:

1. Pawel Achtel, 9x7 Camera: <http://24x7.com.au>
2. Wavelet Beam, RAW Video Denoising: <https://waveletbeam.com/index.php/iris>
3. Andrey Norkin and Neil Birkbeck, Film Grain Synthesis for AV1 Video Codec: [https://norkin.org/pdf/DCC\\_2018\\_AV1\\_film\\_grain.pdf](https://norkin.org/pdf/DCC_2018_AV1_film_grain.pdf)
4. Andrey Norkin & Liwei Guo from Netflix, DEMUXED 2019: <https://www.youtube.com/watch?v=JU-sE14U9xM&t=299s>
5. Zhi Li, Netflix, On VMAF's property in the presence of image enhancement operations
6. Wavelet Beam, Copyright Q2/2023, Image content-independent measurement method for video resolution determination: <https://waveletbeam.com/index.php/news/65-image-content-independent-measurement-method-for-video-resolution-determination>
7. HDTVTest, TV Settings That Should be ILLEGAL: <https://www.youtube.com/watch?v=uGfT746TJu0&t=143s>
8. FILMMAKER MODE: <https://filmmakermode.com>