Skalierbare Videokodierung am Beispiel SVC Medieninformatik IL

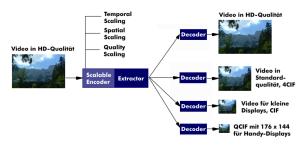
Andreas Unterweger

Vertiefung Medieninformatik Studiengang ITS FH Salzburg

Wintersemester 2021/22

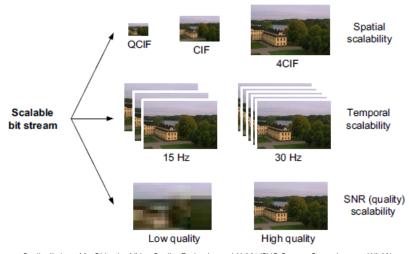
Überblick zu skalierbarer Videokodierung I

- Kodierung von Videos mit der Möglichkeit zur Dekodierung von festgelegten Teilen (so genannten Layern)
- Jeder Layer fixiert drei Skalierbarkeitsdimensionen:
 - Zeitlich (engl. temporal)
 - Örtlich (engl. spatial)
 - Qualitativ (engl. quality)



Quelle: Unbekannt: SVC (scalable video coding). http://www.itwissen.info/definition/lexikon/SVC-scalable-video-coding.html (30.8.2014), 2014.

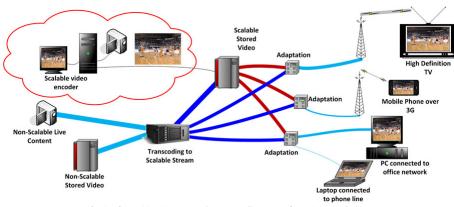
Überblick zu skalierbarer Videokodierung II



Quelle: Ibekwe, M.: Objective Video Quality Evaluation and H.264/SVC Content Streaming over WLANs. http://access.feld.cvut.cz/view.php?cisloclanku=2013010001 (30.8.2014), 2013.

Vor- und Nachteile von skalierbarer Videokodierung

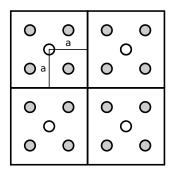
- Vorteil: Bandbreitenersparnis durch empfängerseitige Adaptierung statt senderseitiger Mehrfachkodierung (bei Redundanzausnutzung)
- Nachteil: Empfänger- oder netzwerkseitige Unterstützung notwendig

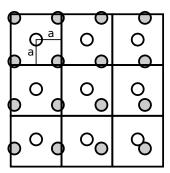


Quelle: Queen Mary University of London: Efficient and Scalable Video Coding. http://www.eecs.qmul.ac.uk/research/impact/aceSVC (30.8.2014), 2014.

Überblick zu Bildskalierung

- Ziel: Abtastrate (Pixelabstand 2a) eines Bildes nachträglich ändern
- Fallunterscheidung anhand des Skalierungsfaktors $s=rac{1}{rac{\partial_{neu}}{\partial_{alt}}}=rac{a_{alt}}{a_{neu}}$:
 - s > 1: Hinaufskalierung (engl. *upsamling*)
 - ullet s < 1: Herunterskalierung (engl. downsampling)
 - $s \in \mathbb{Q} \cap]0;1[$: Einfacher Fall: Unterabtastung (Aliasingfilter notwendig)
 - $s \in \mathbb{R}^+ \setminus \{0\}$: Allgemeiner Fall (schwieriger)



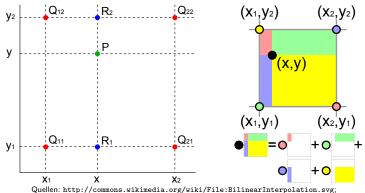


Skalierungsmethoden

- Allgemeiner Ansatz zur Herunterskalierung (meist kombiniert):
 - Interpolation
 - Tiefpassfilterung (um Aliasingartefakte zu vermeiden)
 - Unterabtastung
- Allgemeiner Ansatz zur Hinaufskalierung (meist kombiniert):
 - Ergänzung der Originalpixel um Nullfolgen
 - Interpolation
 - Tiefpassfilterung (zur Glättung)
- Übliche 2-D-Interpolationsverfahren (Auswahl):
 - Nächster Nachbar (engl. nearest neighbor)
 - Bilinear
 - Bikubisch
 - Spline (stückweise Polynome mit Übergangsbedingungen)
 - Lanczos
- Kernidee bei Interpolation: Werte für "Zwischenpixel" bestimmen

Bilineare Interpolation

- 1. Gerade $\overline{Q_{11}Q_{21}}$ konstruieren und R_1 linear interpolieren
- 2. Gerade $\overline{Q_{12}Q_{22}}$ konstruieren und R_2 linear interpolieren
- 3. Gerade $\overline{R_1R_2}$ konstruieren und P linear interpolieren
- ightarrow Pixel werden über jeweils mit P eingeschlossene Flächen gewichtet



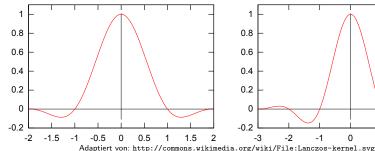
Quener: http://commons.wikimedia.org/wiki/File:Bilinear_interpolation.svg, http://commons.wikimedia.org/wiki/File:Bilinear_interpolation_visualisation.svg

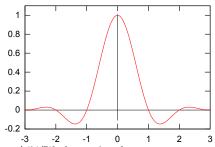
Lanczos-Interpolation I

• Basis: Normalisierte $\frac{\sin(x)}{x}$ -Funktion mit $\frac{\sin(x)}{x}$ -Fensterfunktion:

$$I(x) = \begin{cases} sinc(x) \cdot sinc\left(\frac{x}{b}\right) &, |x| < b \\ 0 &, sonst \end{cases}, b \in \mathbb{N}, sinc(x) = \frac{sin(\pi x)}{\pi x}$$

• Interpolation von $F: f_{interpoliert}(x) \approx \sum_{x'=|x|-b+1}^{\lfloor x \rfloor + b} F(x') \cdot I(x-x')$





Lanczos-Interpolation II

• Erweiterung des 1-D-Falles zu 2-D (kombinierte horizontale und vertikale Interpolation durch Separierbarkeit, ohne Details):

$$f_{interpoliert}(x,y) pprox \sum_{x'=\lfloor x \rfloor-b+1}^{\lfloor x \rfloor+b} \sum_{y'=\lfloor y \rfloor-b+1}^{\lfloor y \rfloor+b} F(x',y') \cdot I(x-x') \cdot I(y-y')$$

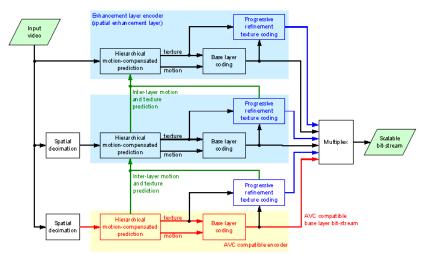
- Vorteile:
 - Gute N\u00e4herung der theoretisch optimalen sinc-Interpolation
 - Erhält Schärfe besser als bilineare und bikubische Filterung
 - Güte durch Anzahl von Nachbarpixeln über Parameter b steuerbar
- Nachteile:
 - ullet Kann in Randfällen negative Werte liefern o Korrektur notwendig
 - Aufwändig zu berechnen
 - ullet Berücksichtigung des Randfalles I(0):=1 (Definition) notwendig



Scalable Video Coding (SVC)

- Standard zur skalierbaren Videokodierung
- Spezifiziert in H.264 Annex G
- Baut auf H.264 auf und ist abwärtskompatibel (ähnlich wie bei MVC über NALU-Erweiterungen gelöst)
- Basislayer (engl. base layer): Layer mit geringster Bildwiederholrate,
 Auflösung und Qualität (voll H.264-konform)
- Verbesserungslayer (engl. enhancement layer): Auf Basislayer aufbauend mit höherer Bildwiederholrate, Auflösung oder Qualität
- → H.264-Decoder kann Basislayer dekodieren
- ightarrow SVC-Decoder kann Basislayer und beliebig viele Verbesserungslayer dekodieren (je nach Benutzervorgabe und Anwendung)

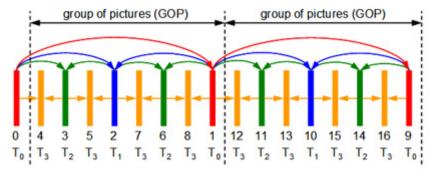
SVC-Architektur



Quelle: Unbekannt: Scalable H.264 Video Coding. http://www-sipl.technion.ac.il/Info/News&Events_1_e.php?id=284 (30.8.2014), 2006.

Zeitliche Skalierbarkeit

- "Gratis" bei entsprechender (hierarchischer) GOP-Struktur
- Frames von zeitlichem Layer T_i , $i \in \{x | 1 \le x \le t_{max}\}$, $t_{max} \in \mathbb{N}$ werden nur aus Frames von zeitlichem Layer T_{i-1} (bidirektional) prädiziert; Frames von T_0 werden von anderen T_0 -Frames prädiziert



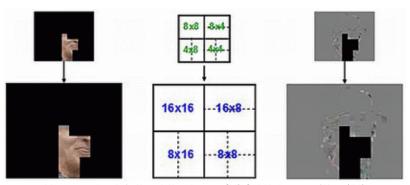
Adaptiert aus: Fraunhofer Heinrich Hertz Institute: SVC: Scalable Extension of H.264/AVC. http://www.hhi.fraunhofer.de/fields-of-competence/image-processing/research-groups/image-communication/video-coding/svc-scalable-extension-of-h264avc.html (30.8.2014), 2014.

Örtliche Skalierbarkeit I

- ullet Erzeugung verschiedener Layer durch Herunterskalieren des Ausgangsbildes mit Lanczos-3-Filter o niedrigere Auflösungen
- Skalierung typischerweise dyadisch (geviertelte Auflösung)
- Beschränkung: Auflösung muss monoton (über Layer) steigen
- Ohne Details: Ausschneiden und Erweitern von Bildinhalten bei Layerwechsel möglich (engl. extended spatial scalability)
- Inter-Layer-Prädiktion (engl. inter-layer prediction, ILP): Verwendung von Daten "niedrigerer" Layer zur Prädiktion im aktuellen Layer
- Drei ILP-Techniken (teilweise kombinierbar):
 - Inter-Layer-Intraprädiktion (engl. inter-layer intra prediction)
 - Inter-Layer-Modi- und -Bewegungsprädiktion (engl. *inter-layer mode and motion prediction*)
 - Inter-Layer-Differenzprädiktion (engl. inter-layer residual prediction)

Örtliche Skalierbarkeit II

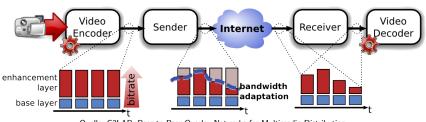
- Modi- und Bewegungsprädiktion: Makroblockpartitionen und MV werden vergrößert übernommen (Differenzen können kodiert werden)
- Differenzprädiktion: MC-Ergebnis wird (bis zum Blockrand) bilinear hinaufskaliert und übernommen (Differenzen können kodiert werden)



Adaptiert aus: Fraunhofer Heinrich Hertz Institute: SVC: Scalable Extension of H.264/AVC. http://www.hhi.fraunhofer.de/fields-of-competence/image-processing/research-groups/image-communication/video-coding/svc-scalable-extension-of-h264avc.html (30.8.2014), 2014.

Qualitative Skalierbarkeit

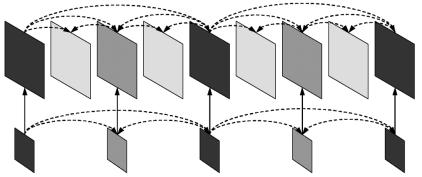
- Unterscheidung: Grob (engl. coarse-grain) und fein (engl. fine-grain)
- Grobe qualitative Skalierbarkeit: Sonderfall von örtlicher Skalierbarkeit mit gleicher Auflösung (keine Skalierung notwendig)
- Feine qualitative Skalierbarkeit: Zusätzliche Koeffizienten(-differenzen) im Bitstrom (ohne Details)



Quelle: C3LAB: Peer to Peer Overlay Networks for Multimedia Distribution. http://c3lab.poliba.it/index.php/OverlayNetworks (30.8.2014), 2007.

Kombinierte Skalierbarkeit I

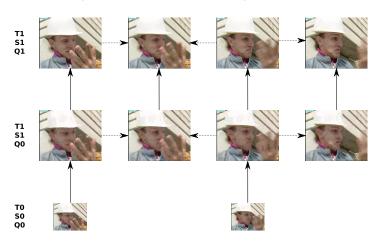
- Kombination von Skalierbarkeitsdimensionen trivial
- Signalisierung von Layern notwendig (über SEI und SPS-Erweiterung)
- Beispiel für Kombination von zeitlicher und örtlicher Skalierbarkeit:



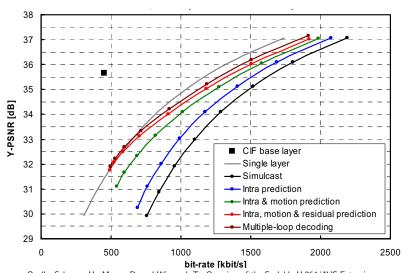
Quelle: Schwarz, H., Marpe, D. und Wiegand, T.: Overview of the Scalable Video Coding Extension of the H.264/AVC Standard. In IEEE Transactions on Circuits and Systems for Video Technology, vol. 17, no. 9, 2007.

Kombinierte Skalierbarkeit II

Beispiel für Kombination von zeitlicher, örtlicher und qualitativer
 Skalierbarkeit (nicht alle Layer abgebildet):

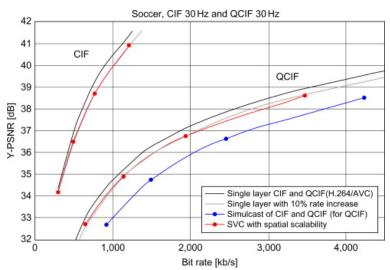


Leistungsfähigkeit I



Quelle: Schwarz, H., Marpe, D. und Wiegand, T.: Overview of the Scalable H.264/AVC Extension. http://iphome.hhi.de/marpe/download/icip06_svc.pdf (30.8.2014), 2006.

Leistungsfähigkeit II



Quelle: J. W. Woods: Digital Video Compression. In Multidimensional Signal, Image, and Video Processing and Coding, 2. Auflage, Kapitel 12, pp. 467-528, 2012.

Danke für die Aufmerksamkeit!

Fragen?