Kodierung von Mehrfachansichten

Medieninformatik IL

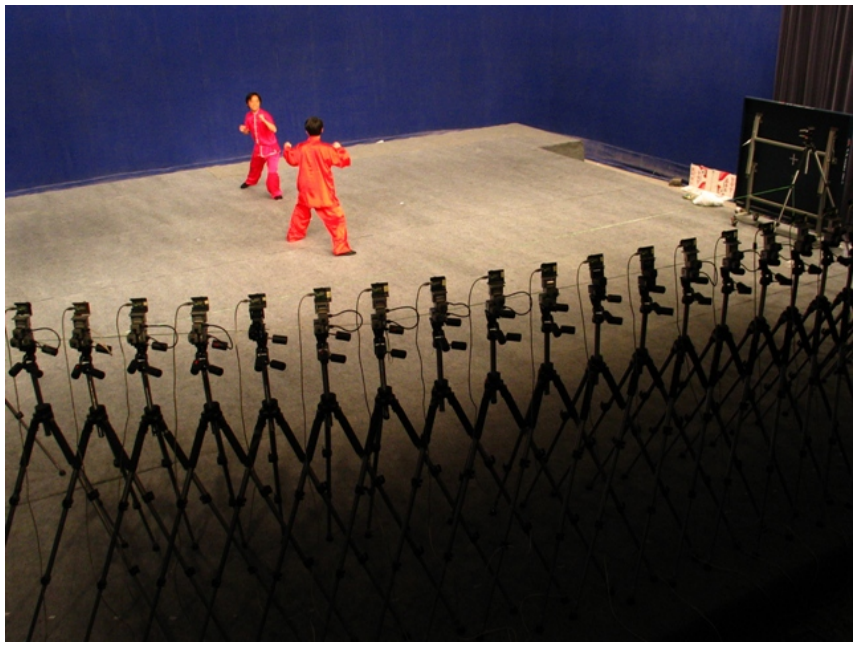
Andreas Unterweger

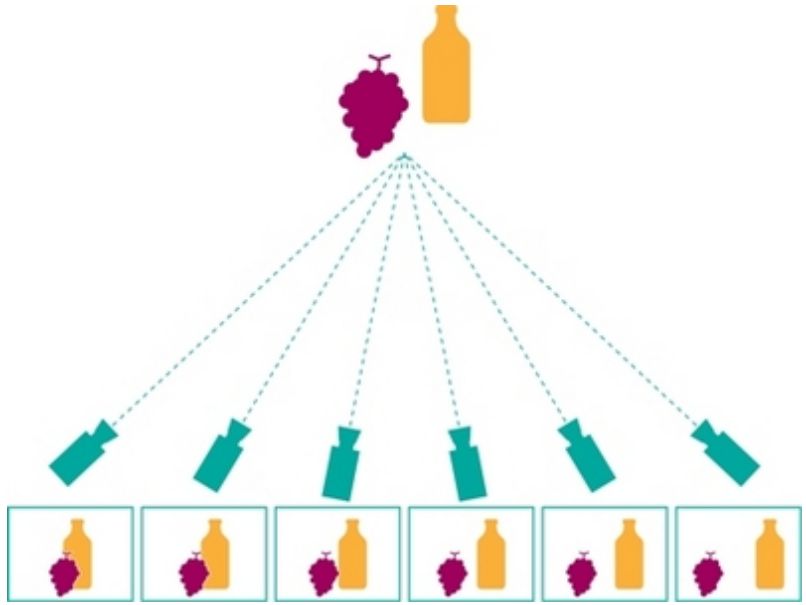
Vertiefung Medieninformatik
Studiengang ITS
FH Salzburg

Wintersemester 2016/17

Überblick Mehrfachansichten

- Mehrfachansichten (engl. *multi-view*): Aufnahmen einer Szene aus mehreren verschiedenen Kamerapositionen
- Sonderfall: Zwei Ansichten (stereo)
- Beispielanwendung: Free-viewpoint TV



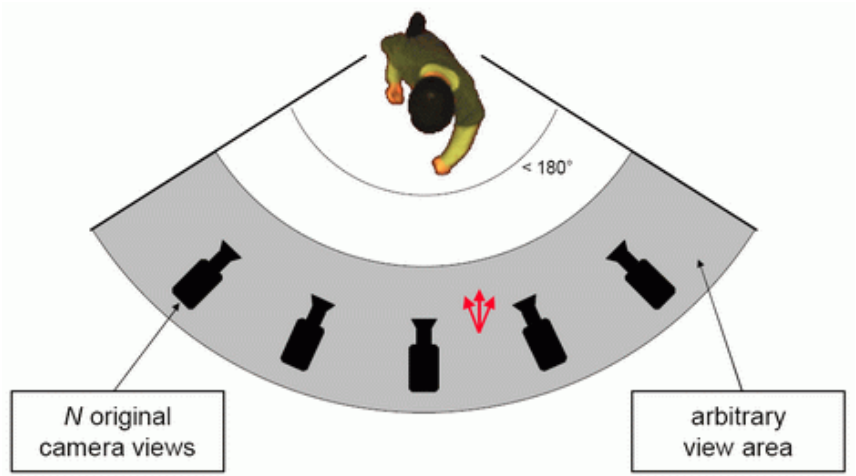


Quellen: Microsoft: Real-time Interactive Multi-view Video System. `http://research.microsoft.com/en-us/projects/imv/` (24.8.2014), 2008; Honekai Media: Feature: Video System Lets You Interactively Change Viewpoint, Generate Dynamic Scenes from any 3D Position. `http://www.techeblog.com/index.php/tech-gadget/feature-video-system-lets-you-interactively-change-viewpoint-generate-dynamic-scenes-from-any-3d-position` (24.8.2014), 2008.

Beispiel Free-viewpoint TV

- Freie Wahl der Kameraposition bei der Wiedergabe
- Zusätzliche Ansichten durch virtuelle Kameras (interpoliert)

→ Virtuelle Kamerafahrten möglich

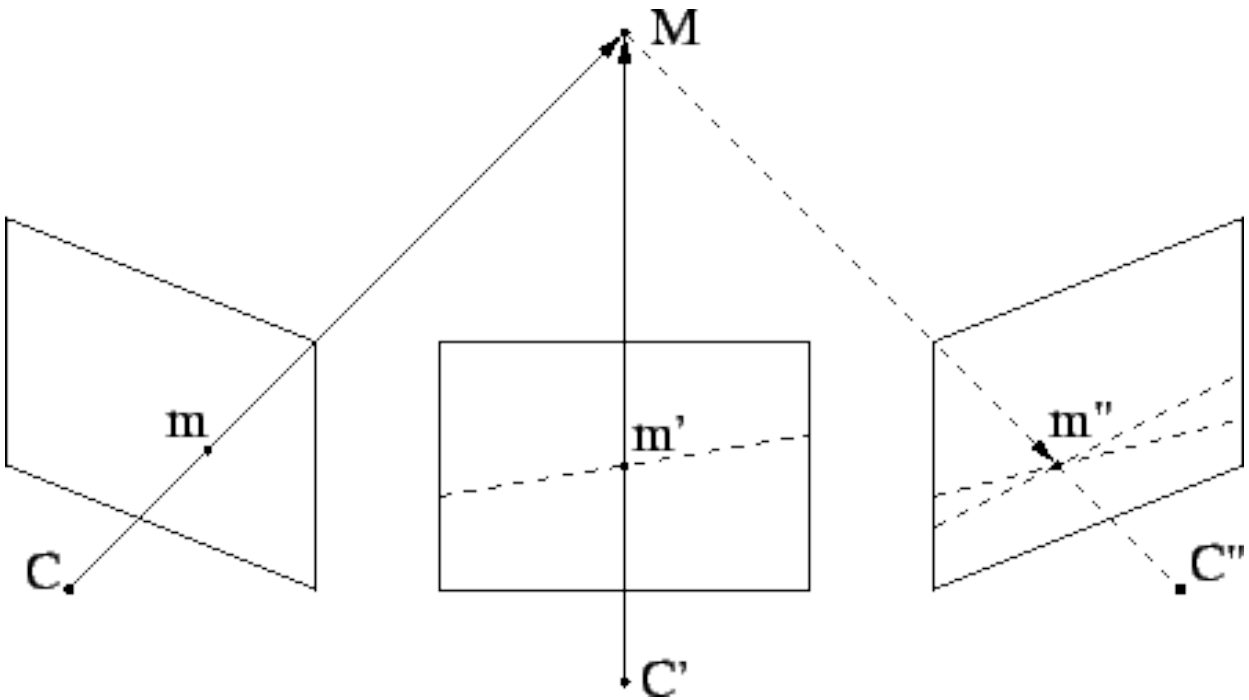


Quellen: Moshe, Y.: Multi-view Video Compression for H.264. http://www-sipl.technion.ac.il/Info/News&Events_1_e.php?id=292 (24.8.2014), 2006; NLT Technologies: NLT Technologies Introduces High-Resolution Multi-View 2D/3D Display with HxDP Technology for Enhanced Stereoscopic Viewing Experience. http://www.nlt-technologies.co.jp/en/release/release_120514.html (24.8.2014), 2012.

Vor- und Nachteile von mehr als zwei Ansichten I

- Einfachere Bestimmung von Punktkorrespondenzen (mit Einschränkungen); Beispiel mit drei Ansichten:
  - Punkt m liegt auf Epipolarlinie in (mittlerer und) rechter Ansicht
  - Punkt m' liegt auf Epipolarlinie in rechter Ansicht

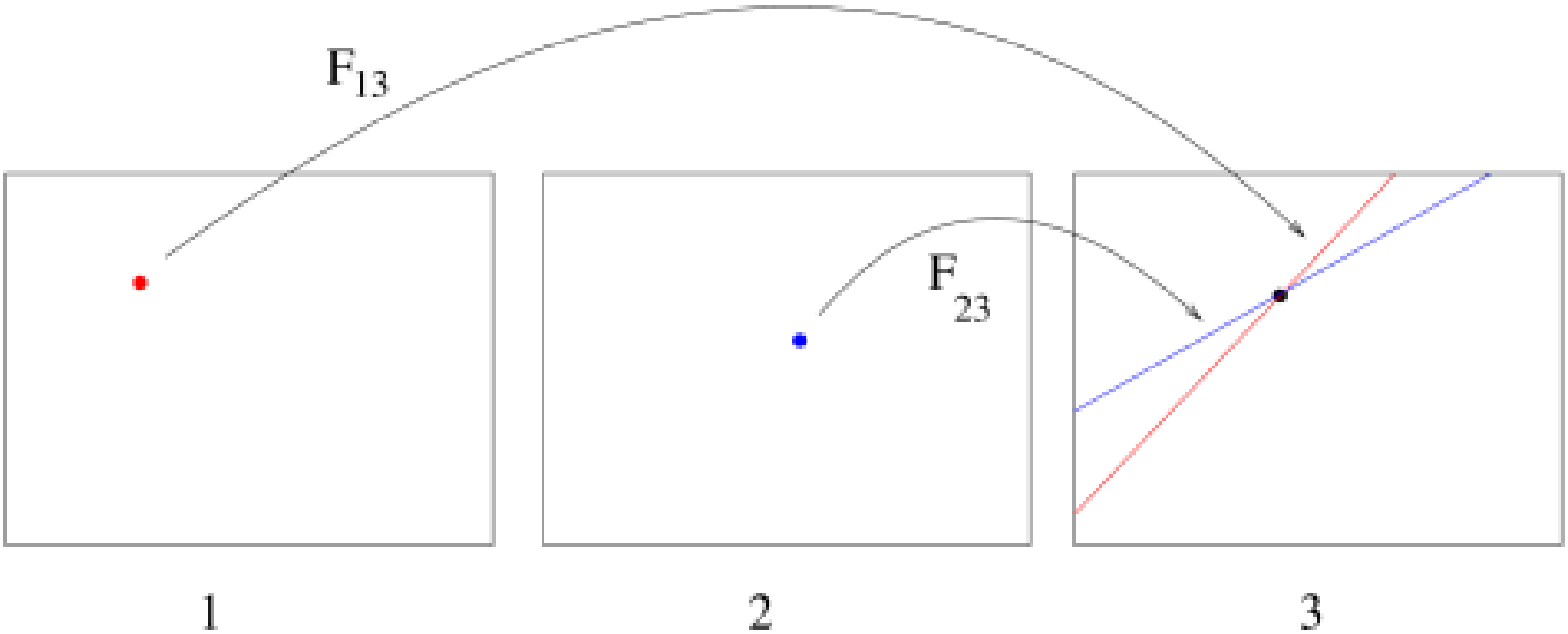
  $\rightarrow$ Schnittpunkt der beiden Epipolarlinien bestimmt Lage von m''
  - Voraussetzung: Epipolarlinien sind nicht parallel zueinander



Quelle: Pollefeys, M.: Three view geometry. `http://www.cs.unc.edu/~marc/tutorial/node45.html` (24.8.2014), 2002.

Vor- und Nachteile von mehr als zwei Ansichten II

- Validierung von Korrespondenzen und anderen geometrischen Zusammenhängen über zusätzliche essenzielle und Fundamentalmatrizen möglich (über Stereobildpaare)
- Aber: Höherer Rechenaufwand und höhere Hardwarekosten



Quelle: Fusiello, A.: Multiple View Geometry. `http://www.3dflow.net/elementsCV/S5.xhtml` (24.8.2014), 2011.

Überblick zur Kodierung von Mehrfachansichten

- Kategorisierung
  - Kodierkomplexität
  - Dekodierkomplexität
  - Ansichtenrekonstruktionskomplexität
  - 2-D-Kodierungskompatibilität
  - Effektiv verfügbare Auflösung
- Arten der Mehrfachansichtenkodierung (Auswahl):
  - Unabhängige Ansichtenkodierung
  - Ansichtenkodierung mit Abhängigkeiten
  - **Ansichtenkodierung mit Tiefenbild (engl. depth map)**
- Ausgewählte Kodiertechniken:
  - JPEG Stereoscopic (zwei Bilder nebeneinander)
  - JPEG Multi-Picture Format (mehrere aneinandergehängte Bilder)
  - **Multi-view Coding (MVC)**
  - **Frame Packing (H.264-Erweiterung)**

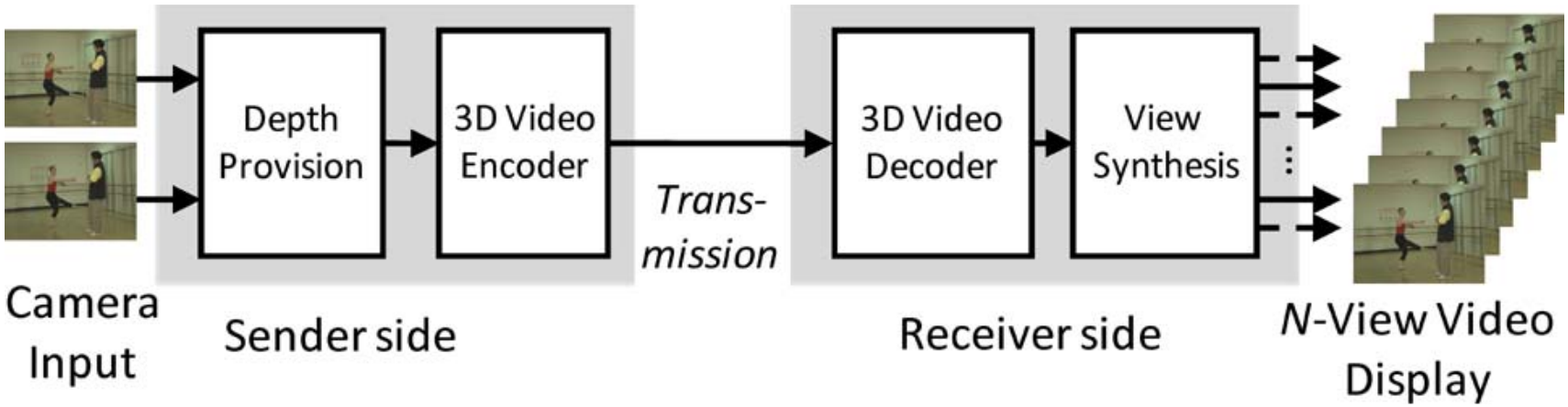
Mehrfachansichtenkodierung mit Tiefenbildern I

- Tiefenbild: Pendant zu Disparitätsbild: Helligkeit gibt z-Koordinate an
- Einschränkungen: Diskrete Werte zwischen z_{min} und z_{max}
- Praktisch: Rektifizierte Bilder mit signalisierten Kameraparametern

Quelle: Müller, K., Merkle, P. und Wiegand, T.: 3-D Video Representation Using Depth Maps. In Proceedings of the IEEE, vol. 9, no. 4, pp. 643-656, 2011.

Mehrfachansichtenkodierung mit Tiefenbildern II

- Kodierung von Mehrfachansichten mit je einem Tiefenbild
- Dekodierung auf Empfängerseite mit zwei prinzipiellen Möglichkeiten
  - Ausgabe einer der aufgenommenen Ansichten (über Disparität)
  - Erzeugung zusätzlicher (virtueller) Ansichten aus den bestehenden (Ansichtensynthese, engl. *view synthesis*)
- Typischerweise nur wenige aufgenommene Ansichten $\rightarrow$ Genauigkeit der Tiefenbilder hat großen Einfluss auf Wiedergabe-/Synthesequalität



Quelle: Müller, K., Merkle, P. und Wiegand, T.: 3-D Video Representation Using Depth Maps. In Proceedings of the IEEE, vol. 9, no. 4, pp. 643-656, 2011.

Möglichkeiten zur Kodierung von Tiefenbildern (Auswahl)

- Verlustlose Kompression
  - Vorteil: Bestmögliche Wiedergabe-/Synthesequalität
  - Nachteil: Sehr speicherintensiv
- Lineare Quantisierung mit Entropiekodierung
  - Vorteil: Einfach zu implementieren
  - Nachteil: Artefakte im Tiefenbild
- Kodierung nach H.264 (o.ä.)
  - Vorteile: Bestehende Encoder wiederverwendbar; speicherschonend
  - Nachteil: Kompressionsartefakte im Tiefenbild
- Spezielle Kodierung (z.B. Platelets)
  - Vorteil: Weniger und vor allem weniger sichtbare Artefakte
  - Nachteil: Zusätzliche Implementierung notwendig

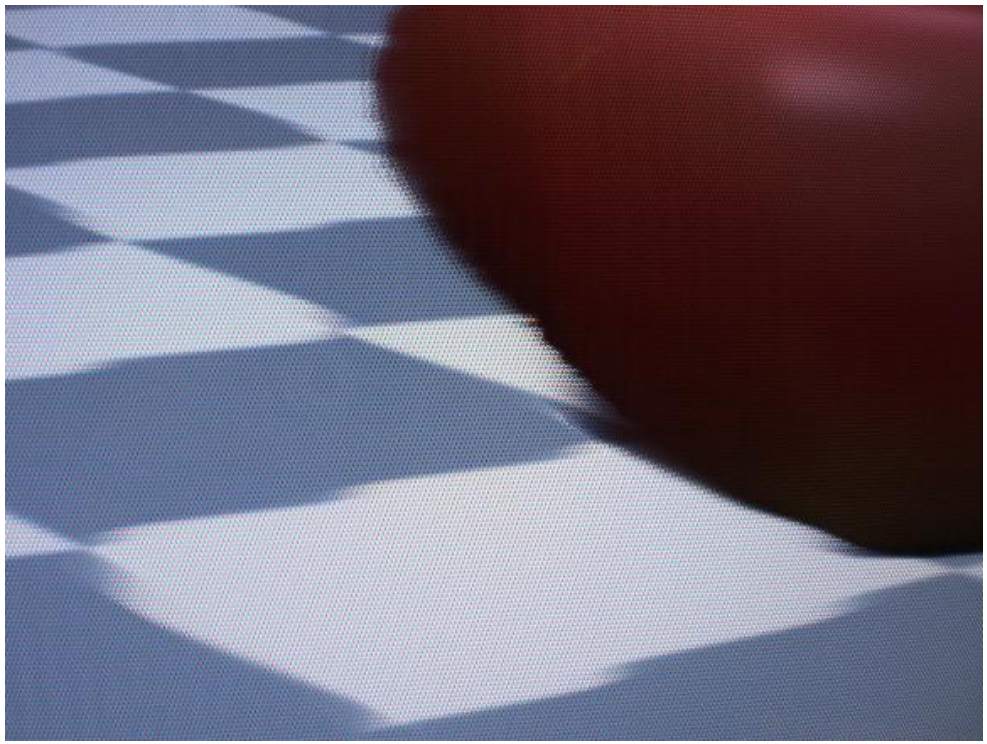
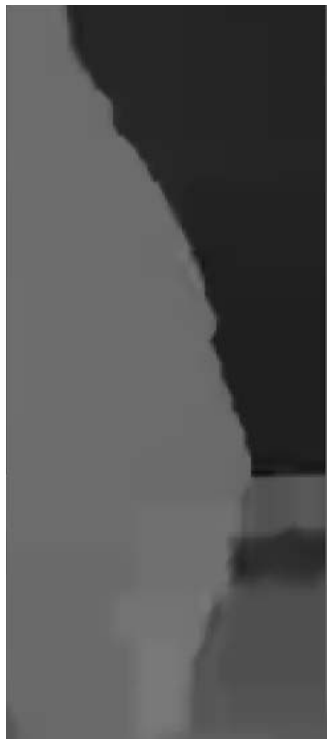
Artefakte durch Tiefenbildkompression (Auswahl) I

- Puppentheatereffekt (Reduktion der dargestellten Tiefen $\rightarrow$ wenige klar voneinander unterscheidbare Tiefenebenen wie in einem Puppentheater) durch lineare Quantisierung des Tiefenbildes

Quelle: Boev, A., Hollosi, D. und Gotchev, A.: Classification of stereoscopic artifacts. `http://sp.cs.tut.fi/mobile3dtv/results/tech/D5.1_Mobile3DTV_v1.0.pdf` (24.8.2014), 2008.

Artefakte durch Tiefenbildkompression (Auswahl) II

- Räumliche Blockbildung, Unschärfe und Überschwingen (durch verlustbehaftete Kodierung nach H.264 u.ä.)
- Hochfrequentes Quantisierungsrauschen (engl. *mosquito noise*)

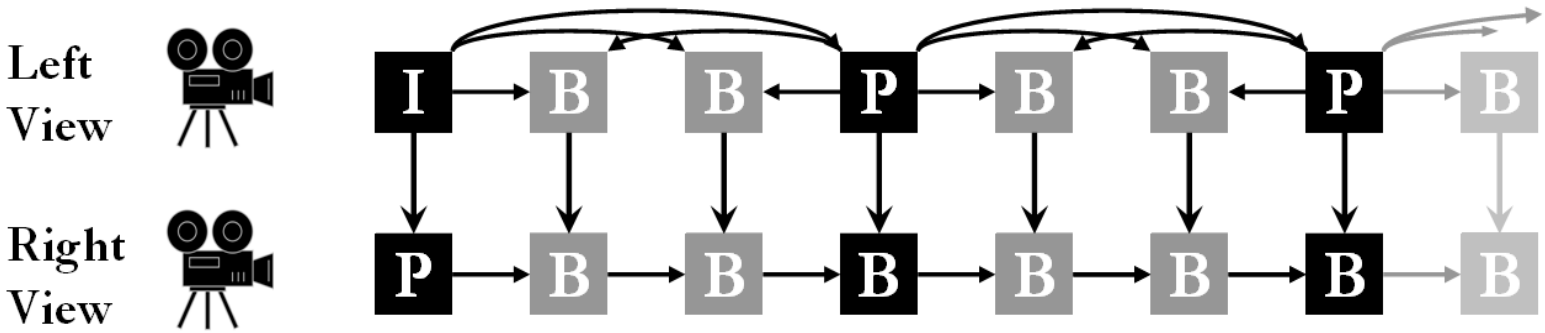
Quellen: Boev, A., Hollosi, D. und Gotchev, A.: Classification of stereoscopic artifacts. `http://sp.cs.tut.fi/mobile3dtv/results/tech/D5.1_Mobile3DTV_v1.0.pdf` (24.8.2014), 2008; Müller, K., Merkle, P. und Wiegand, T.: 3-D Video Representation Using Depth Maps. In Proceedings of the IEEE, vol. 9, no. 4, pp. 643-656, 2011.

Einschub: Kurzwiederholung H.264

- Videokompression nach H.264
  - Blockweise Transformation und Quantisierung
  - Motion Estimation und Motion Compensation
  - Intra- vs. Interprädiktion
  - I- vs. P- vs. B-Bilder
  - Coding Order vs. Display Order
- Bitstromformat nach H.264 Annex B
  - Kleinste Einheit: NAL Unit (NALU)
  - NALU-Header mit Typenidentifikation
  - Wichtige NALU-Typen: SPS, PPS, (IDR-)Slices
- Ergänzungen:
  - Decoded Picture Buffer (DPB) speichert zuvor kodierte Referenzbilder (vereinfacht) nach FIFO-Prinzip
  - NALU-Typ SEI (Supplemental Enhancement Information): Optionale Daten, die vom Decoder ignoriert werden können

Multi-view Coding (MVC) I

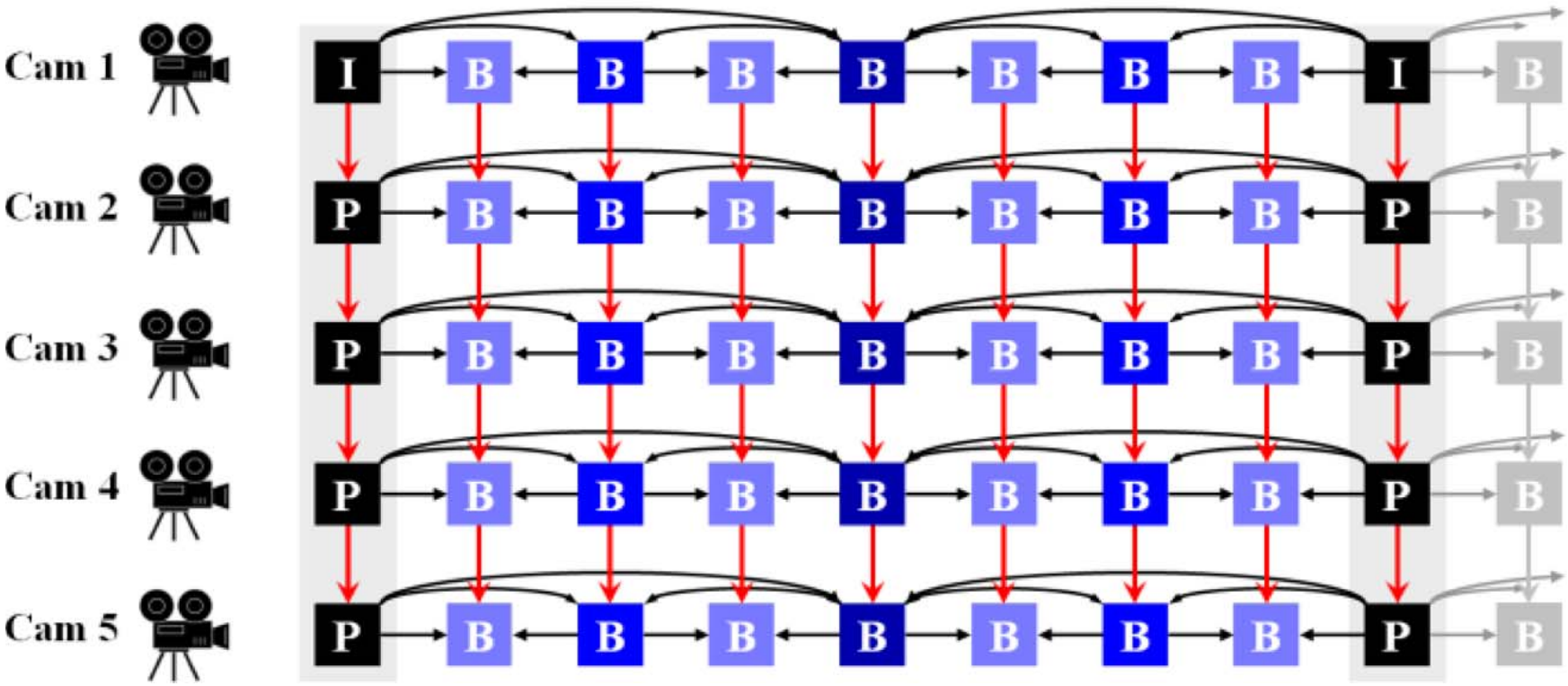
- Prinzipieller Aufbau (spezifiziert in H.264 Annex H):
  - (Unabhängige) H.264-kompatible Grundansicht (engl. *base view*) → H.264-Decoder kann die Grundansicht dekodieren (ignoriert den Rest)
  - Daten weiterer Ansichten durch neue NALU-Typen
  - Metadaten zu neuen Ansichten durch zusätzliche SPS (mit neuem Typ)
  - Redundanzen zwischen Ansichten werden durch Inter-Ansichtenprädiktion (engl. *inter-view prediction*) ausgenutzt
  - Andere Ansichten werden als Referenzen temporär in DPB eingefügt
  - Eigene Profile (z.B. Stereo High für Blu-ray: nur zwei Ansichten)



Quelle: Vetro, A., Wiegand, T. und Sullivan, G. J.: Overview of the Stereo and Multiview Video Coding Extensions of the H.264/MPEG-4 AVC Standard. In Proceedings of the IEEE, vol. 99, no. 4, pp. 626-642, 2011.

Multi-view Coding (MVC) II

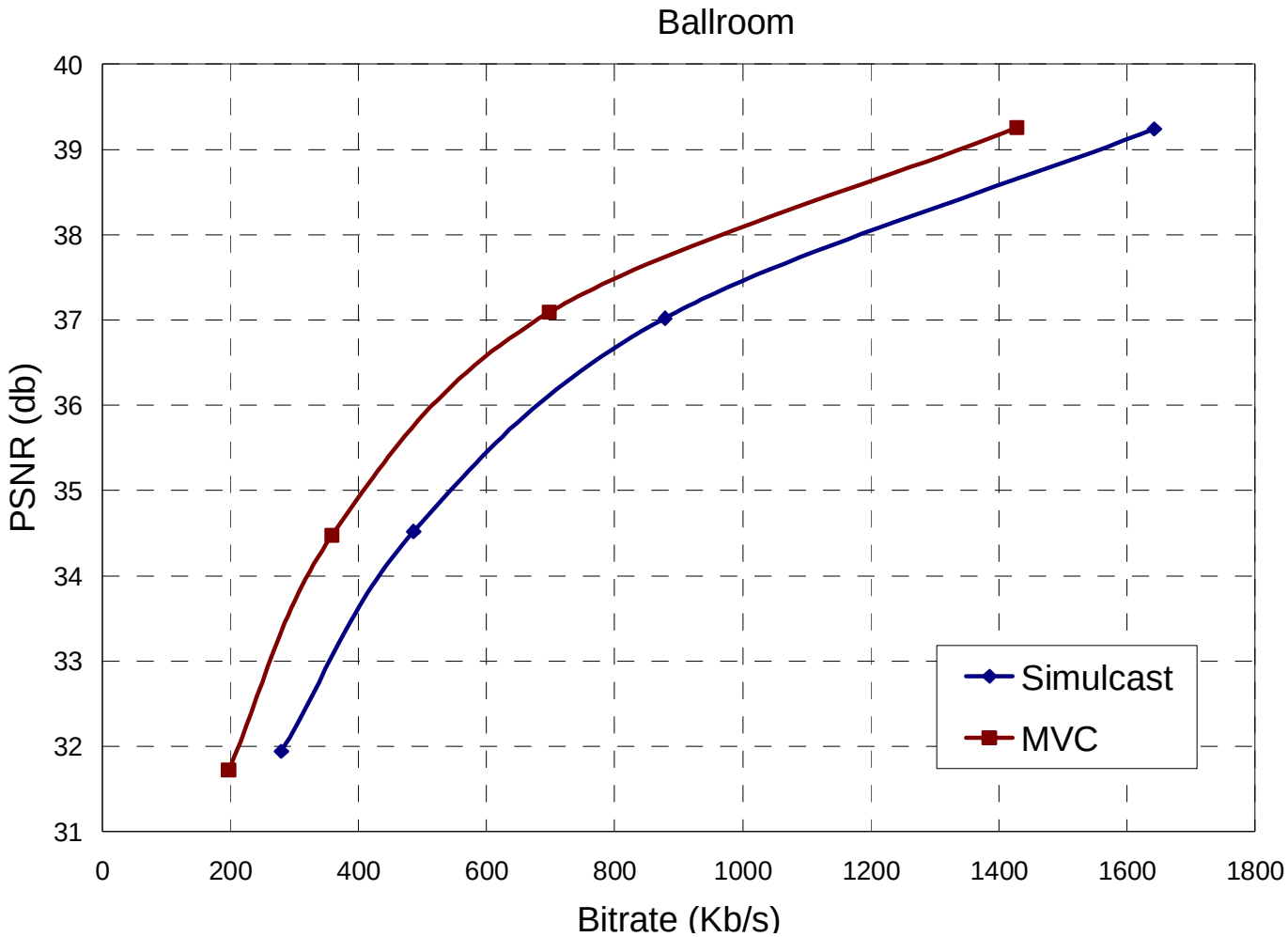
- Zur Reduktion der Komplexität: Inter-Ansichtenprädiktion (rot) nur zwischen Bildern, die den gleichen Darstellungszeitpunkt haben
- Reguläre Prädiktion nach wie vor möglich $\rightarrow$ RDO



Quelle: Müller, K., Merkle, P. und Wiegand, T.: 3-D Video Representation Using Depth Maps. In Proceedings of the IEEE, vol. 9, no. 4, pp. 643-656, 2011.

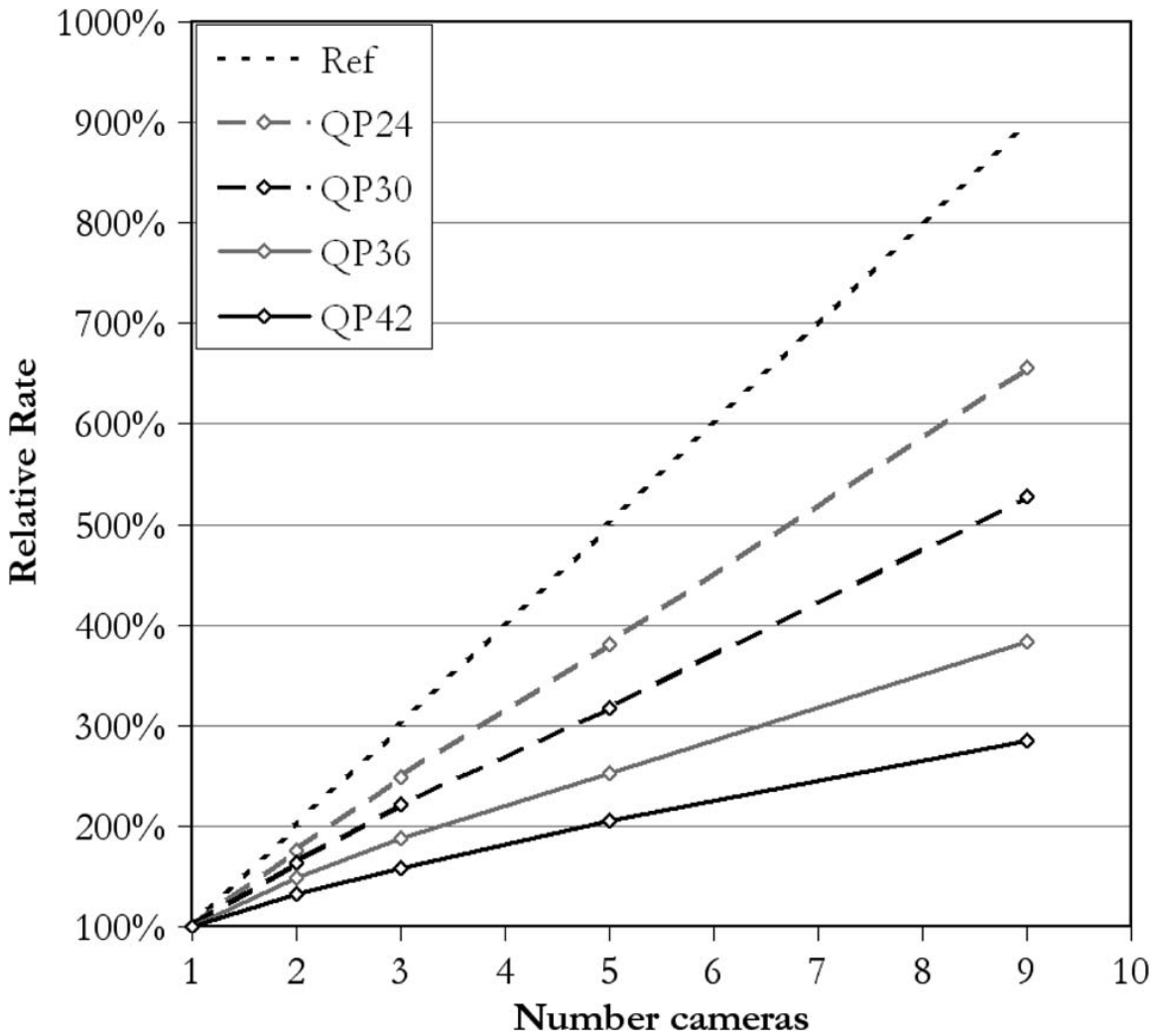
Multi-view Coding (MVC) III

Vergleich zu Simulcast (unabhängige Ansichtenkodierung nach H.264):



Quelle: Vetro, A., Wiegand, T. und Sullivan, G. J.: Overview of the Stereo and Multiview Video Coding Extensions of the H.264/MPEG-4 AVC Standard. In Proceedings of the IEEE, vol. 99, no. 4, pp. 626-642, 2011.

Multi-view Coding (MVC) IV



Quelle: Müller, K., Merkle, P. und Wiegand, T.: 3-D Video Representation Using Depth Maps. In Proceedings of the IEEE, vol. 9, no. 4, pp. 643-656, 2011.

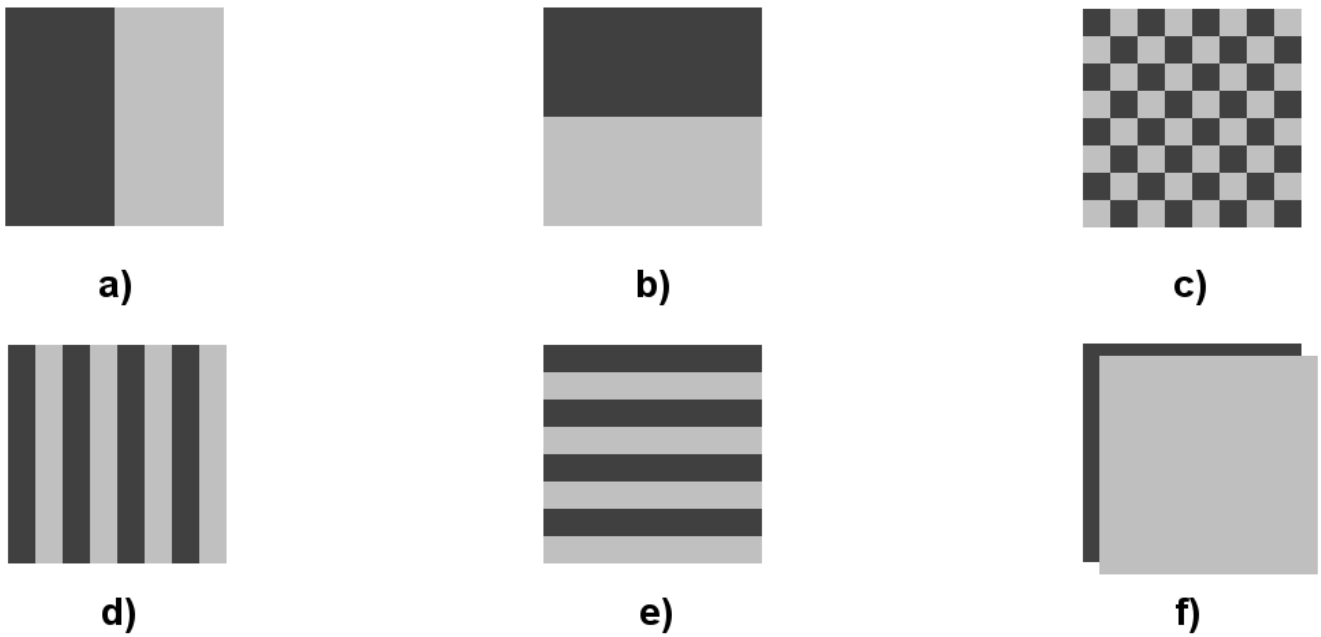
Überblick Frame Packing

- Multiplexing von mehreren Bildern in einem Trägerbild
- Reguläre Kodierung des (gemultiplexten) Bildes möglich
- Fokus: Erweiterung von H.264 (Frame Packing Arrangement SEI)

- Funktionsweise:
  - Encoder: Multiplexing; Signalisierung der Multiplexingart über SEI
  - Decoder: Extraktion der Multiplexingart; Demultiplexing
- Vorteile:
  - Auf Bitstromebene abwärtskompatibel
  - Verwendung bestehender Coding Tools $\rightarrow$ minimaler Implementierungsmehraufwand
- Nachteile:
  - Ohne Decoderunterstützung nicht abwärtskompatibel darstellbar
  - Reduktion der Bildwiederholrate oder Auflösung pro Bild
  - Übersprechen (Kompressionsartefakte) durch Prädiktion

Mögliche Frame Packing Arrangements

- a)-e): Halbierung der horizontalen bzw. vertikalen Bildauflösung pro Ansicht zur Einhaltung der Levelgrenzen $\rightarrow$ Rekonstruktion bei Wiedergabe (mit Unterabtastungsartefakten)
- f): Halbierung der Bildwiederholrate pro Ansicht zur Einhaltung der Levelgrenzen (bei Reserven nicht unbedingt notwendig)



Beispiele für Frame Packing

- Beispiele für horizontale (a)) und vertikale (b)) Frame Packings:

Adaptiert von: Sinofsky, S.: Building a rich and extensible media platform. `http://blogs.msdn.com/b/b8/archive/2012/06/08/building-a-rich-and-extensible-media-platform.aspx` (24.8.2014), 2012.

- Anmerkung: Einige Frame Packings werden direkt von HDMI 1.4a (und seinen Nachfolgern) unterstützt $\rightarrow$ Direkte Wiedergabe möglich

Danke für die Aufmerksamkeit!

Fragen?