

## Bildunterschriften zur Kodierung von Mehrfachansichten

- **Folie 2:** Durch die Verwendung mehrerer halbkreisförmig angeordneter Kameras können zusätzliche Blickwinkel abgebildet werden.
- **Folie 3:** Neben den Aufnahmen physischer Kameras können in einem beschränkten Bereich (links grau gekennzeichnet) zusätzliche virtuelle Kameras (rotes Koordinatenkreuz links) platziert werden. Durch Interpolation der vorhandenen Daten werden zusätzliche Ansichten und dadurch virtuelle Kamerafahrten (rechts) möglich.
- **Folie 4:** Der Punkt  $M$  wird von der linken Kamera  $C$  auf  $m$  und von der mittleren Kamera  $C'$  auf  $m'$  projiziert. Die beiden Epipolarlinien von  $m$  und  $m'$  in der rechten Kamera  $C''$  können geschnitten werden und ergeben dadurch theoretisch den projizierten Punkt  $m''$  ohne zusätzliches Stereo-Matching.
- **Folie 5:** Die Fundamentalmatrix  $F_{13}$ , die die Epipolargleichung zwischen Kamera 1 und Kamera 3 beschreibt, kann mit der Fundamentalmatrix  $F_{23}$ , die die Epipolargleichung zwischen Kamera 2 und Kamera 3 beschreibt, kombiniert werden, um aus einem bekannten Punktepaar (rot und blau) aus zwei Ansichten den korrespondierenden Punkt (schwarz) in der dritten Ansicht zu berechnen. Dieser entspricht dem Schnittpunkt der sich aus den Epipolargleichungen ergebenden Epipolarlinien für den roten und den blauen Punkt.
- **Folie 7:** Die linke und die rechte Ansicht (oben) haben voneinander unabhängige Tiefenbilder (unten). Die Helligkeit jedes Pixels eines Tiefenbildes gibt z-Koordinate bzw. die Entfernung ebendieses Pixels an.
- **Folie 8:** Durch die Tiefenermittlung auf Senderseite und eine anschließende Kodierung können Mehrfachansichten effizient übertragen und beim Empfänger dekodiert werden. Je nach Wiedergabetechnologie können zusätzliche Ansichten über Ansichtensynthese aus den empfangenen Daten interpoliert werden.
- **Folie 10:** Wird ein Tiefenbild (links) linear quantisiert, entsteht der so genannte Puppentheater effekt, bei dem die Entfernungen benachbarter Objekte nicht fließend ineinander übergehen, sondern abrupt zwischen diskreten Werten springen (rechts) – ähnlich der beschränkten Anzahl von Bühnenebenen bei einem Puppenspiel.
- **Folie 11:** Kompressionsartefakte im Tiefenbild äußern sich bei der Darstellung (links) an den Rändern von Objekten, z.B. in Form von so genanntem Depth Blocking (Stufen im Schachbrettmuster) und Depth Ringing (Übergang zwischen Ball und Schachbrettmuster) inkl. Depth Blurring. Quantisierungsrauschen, das im Tiefenbild selbst (Mitte) kaum wahrnehmbar ist, erzeugt an den Rändern zwischen Objekten mit großem

Entfernungsunterschied deplatzierte Pixel (rechts), z.B. schwarze Vordergrundpixel im beigen Hintergrund.

- **Folie 13:** Während die linke Ansicht (oben) regulär und für sich kodiert wird, hängt die rechte Ansicht (unten) von der linken ab (durch vertikale Pfeile angedeutet). Eine Prädiktion ist daher nur von der unabhängigen (oben) zur abhängigen Ansicht (unten) möglich – nicht umgekehrt.
- **Folie 14:** MVC erlaubt komplexe Prädiktionsstrukturen mit mehreren Ansichten, sofern die Ansicht von Kamera 1 unabhängig vom Rest kodiert wird, alle anderen Ansichten dieselbe Kodierreihenfolge verwenden und Inter-Ansichtenprädiktion auf Frames mit demselben Darstellungszeitpunkt beschränkt bleibt (angedeutet durch rote Pfeile). Die Kodierreihenfolge der Ansicht von Kamera 1 gibt daher auch die Kodierreihenfolge der anderen Ansichten vor.
- **Folie 15:** Im Vergleich zu Simulcast, der unabhängigen Kodierung aller Ansichten, ist MVC etwas effizienter, da es Teile der Redundanzen zwischen den Ansichten durch Prädiktion verringert.
- **Folie 16:** Mit der Anzahl der Ansichten steigt der Speicherbedarf selbst mit MVC praktisch linear – auch unterschiedliche QPs (verschieden strichlierte Linien) ändern daran nichts. Im Vergleich zu Simulcast (*Ref*) ist die relative Platzersparnis immer konstant.
- **Folie 18:** Die linke (dunkelgrau) und rechte Ansicht (hellgrau) können per Frame Packing in verschiedenen Anordnungen gemeinsam übertragen werden: a) vertikal mit halbiertem horizontaler Bildauflösung, b) horizontal mit halbiertem vertikaler Bildauflösung, c) als blockweises Schachbrettmuster mit Reduktion der Bildauflösung um den Faktor zwei, d) blockweise vertikal mit halbiertem horizontaler Bildauflösung, e) blockweise horizontal mit halbiertem vertikaler Bildauflösung, f) abwechselnd mit jeweils halbiertem Bildwiederholrate.
- **Folie 19:** Horizontales Frame Packing (links) verringert die vertikale Bildauflösung und belässt die horizontale. Vertikales Frame Packing (rechts) hingegen verringert die horizontale Bildauflösung und belässt die vertikale.