Einführung in die Video-Encodierung

Digitale AV Technik, MIB 5

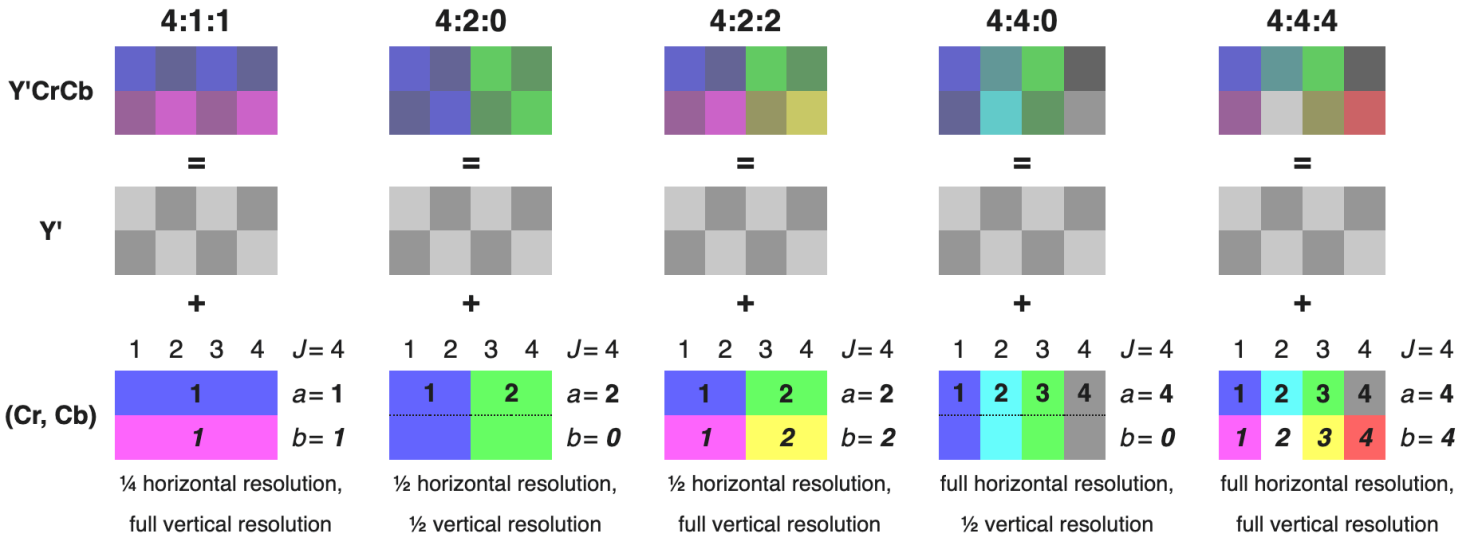
Bildquellen: Christian Feldmann und andere

Konvertierung in den YUV Farbraum

- Helligkeit (Luma) und Farbdifferenzen (Chroma) statt RGB speichern
- Da das Auge vor allem auf die Helligkeit reagiert, kann man bei den Farbdifferenzen reduzieren: Chroma Subsampling

Chroma subsampling im J : a : b Schema

- **J**: Anzahl horizontaler Pixel als Referenz (eigentlich immer == 4)
- **a**: Anzahl der Chrominanzwerte in der ersten Zeile
- **b**: Anzahl der Änderungen der Chrominanzwerte zwischen erster und zweiter Zeile



I, P und B Frames

I-Frames (Intra-coded Frames):

- Vollständige Bilder, unabhängig von anderen Frames.
- Bilden die Grundlage für die Kompression.

P-Frames (Predicted Frames):

- Enthalten Unterschiede zum vorherigen Frame.
- Erfordern weniger Speicherplatz als I-Frames.

B-Frames (Bidirectional Predicted Frames):

- Nutzen von vorherigen und zukünftigen Frames.
- Maximieren die Kompressionseffizienz.

Festlegung von Frame-Typen

Group of Pictures (GOP)

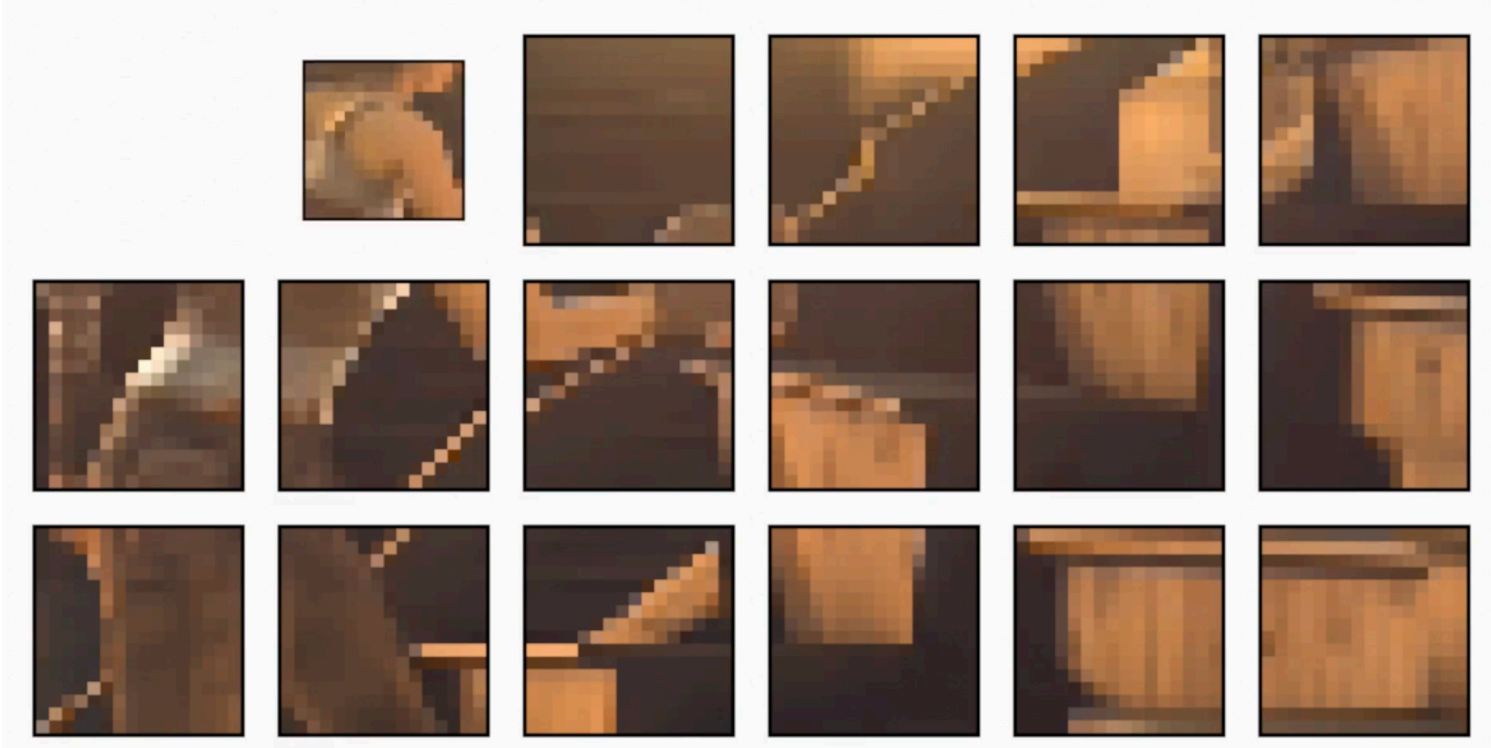
- Struktur von Frames in einer Sequenz:
  - Beispiel: `IBBPBBPBB`
  - Startet immer mit einem I-Frame.

Erkennen von Szenenwechseln

- Automatische Analyse des Videos:
  - Plötzliche Änderungen in Helligkeit oder Bewegung → Neuer I-Frame.

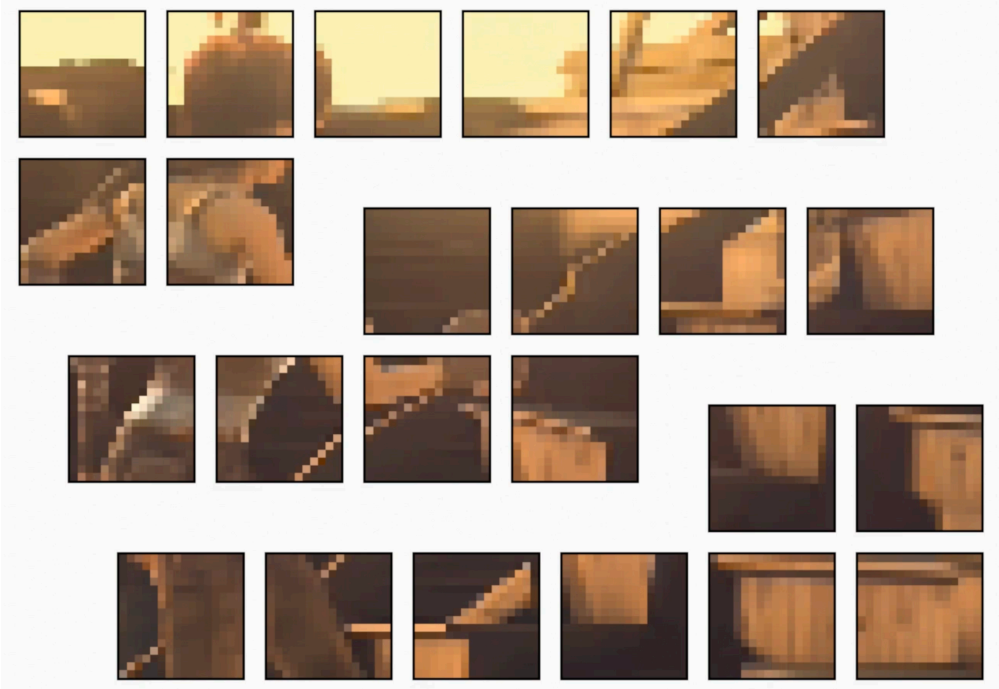
Blockbildung

- Das Bild wird in einzelne Blöcke unterteilt, die nacheinander abgearbeitet werden.



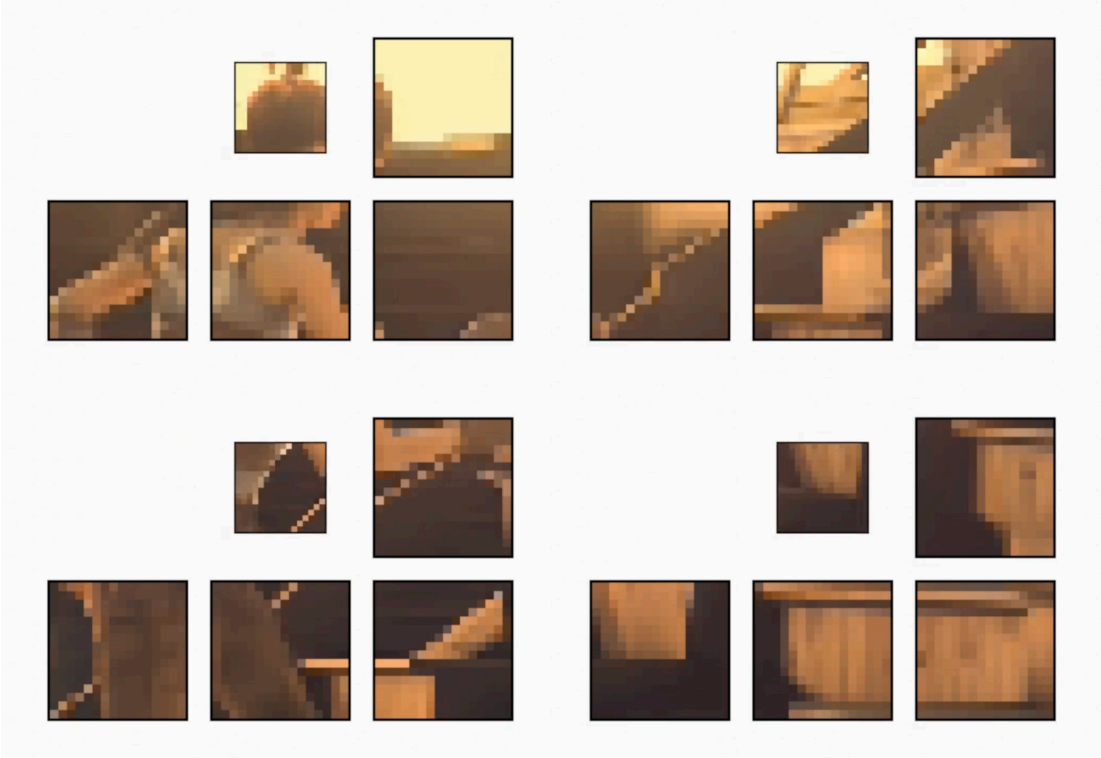
Parallelverarbeitung 01

- **Slices:**
  - Teilbilder innerhalb eines Frames, die unabhängig verarbeitet werden.



Parallelverarbeitung 02

- **Tiles:**
  - Rechteckige Regionen im Bild, um Parallelisierung zu verbessern.

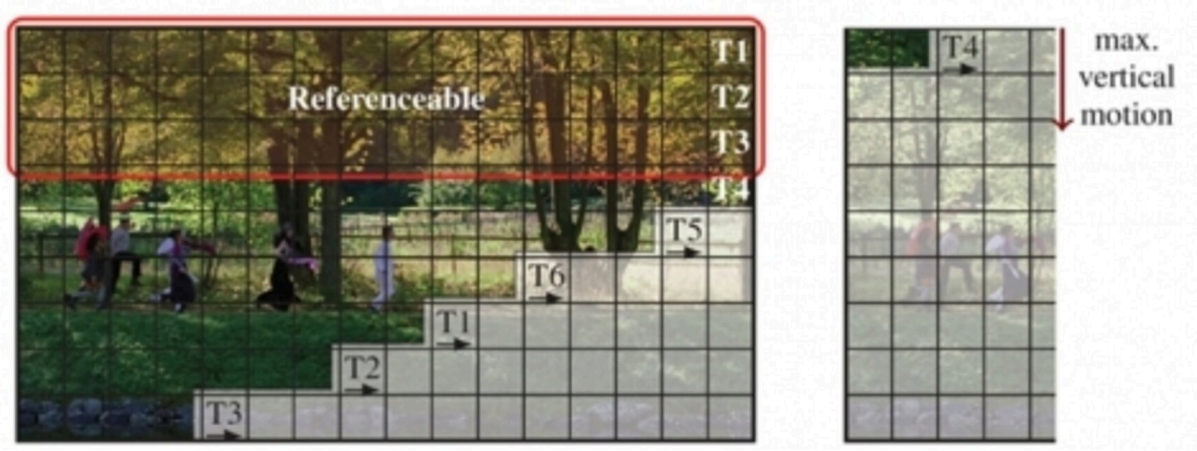


Parallelverarbeitung 03

- **Hierarchische Blöcke**
  - ermöglichen noch mehr Flexibilität

Parallelverarbeitung 04

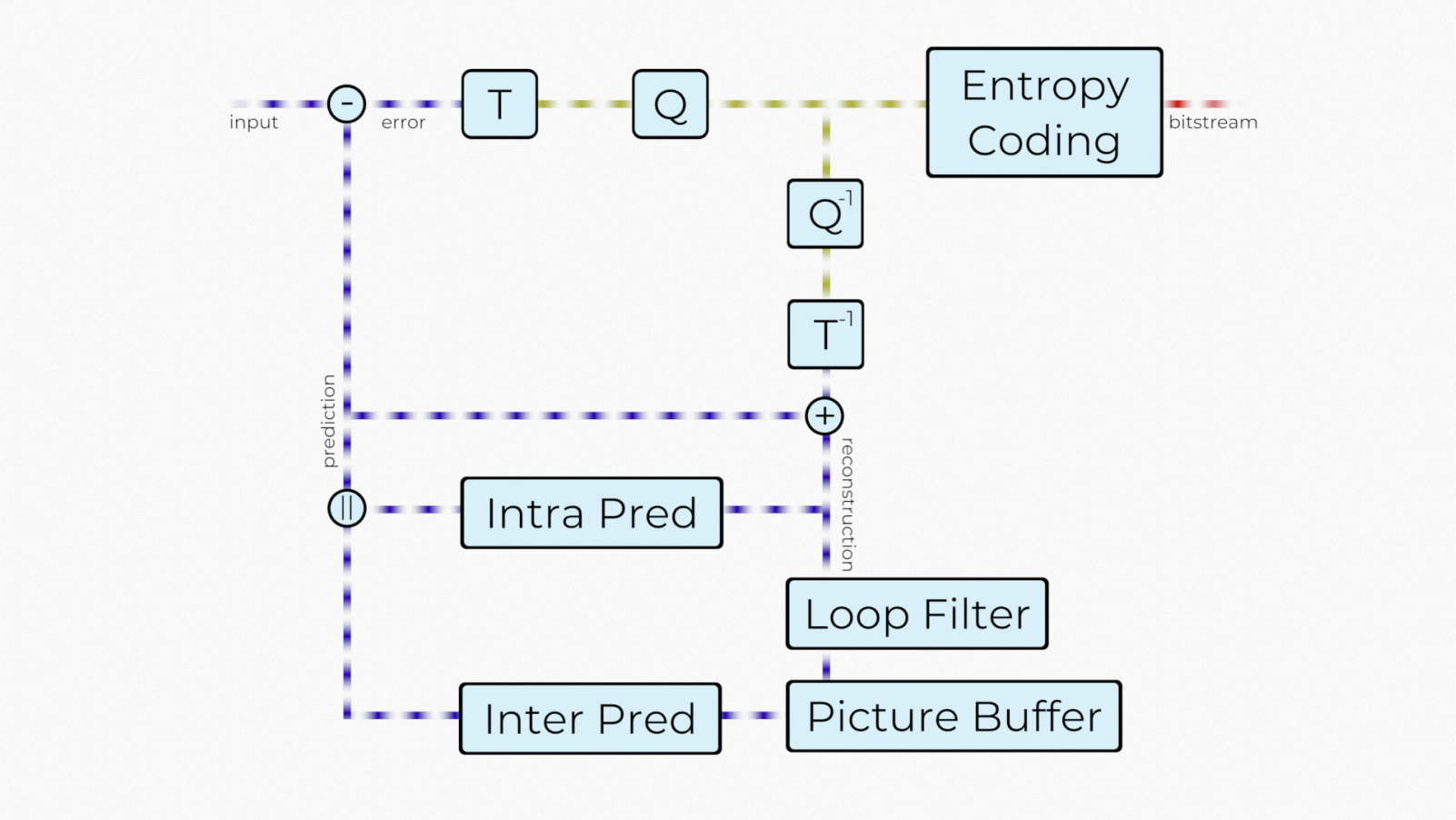
- **Wavefront Parallel Processing**:
  - Verarbeitung startet an einer Ecke des Frames und breitet sich wellenartig aus.
  - Es muss entweder der linke oder obere Nachbar bekannt sein, dann kann der nächste Block verarbeitet werden.



Encoding Loop

1. **Blockaufteilung:**
    - Das Bild wird in kleinere Blöcke (z. B. 16x16 Pixel) unterteilt.
2. **Vorhersage:**
    - **Intra-Prediction:** Innerhalb desselben Frames.
    - **Inter-Prediction:** Zwischen verschiedenen Frames.
3. **Berechnung der Differenz:**
    - Differenz zwischen Originalblock und Vorhersage wird quantisiert und komprimiert.

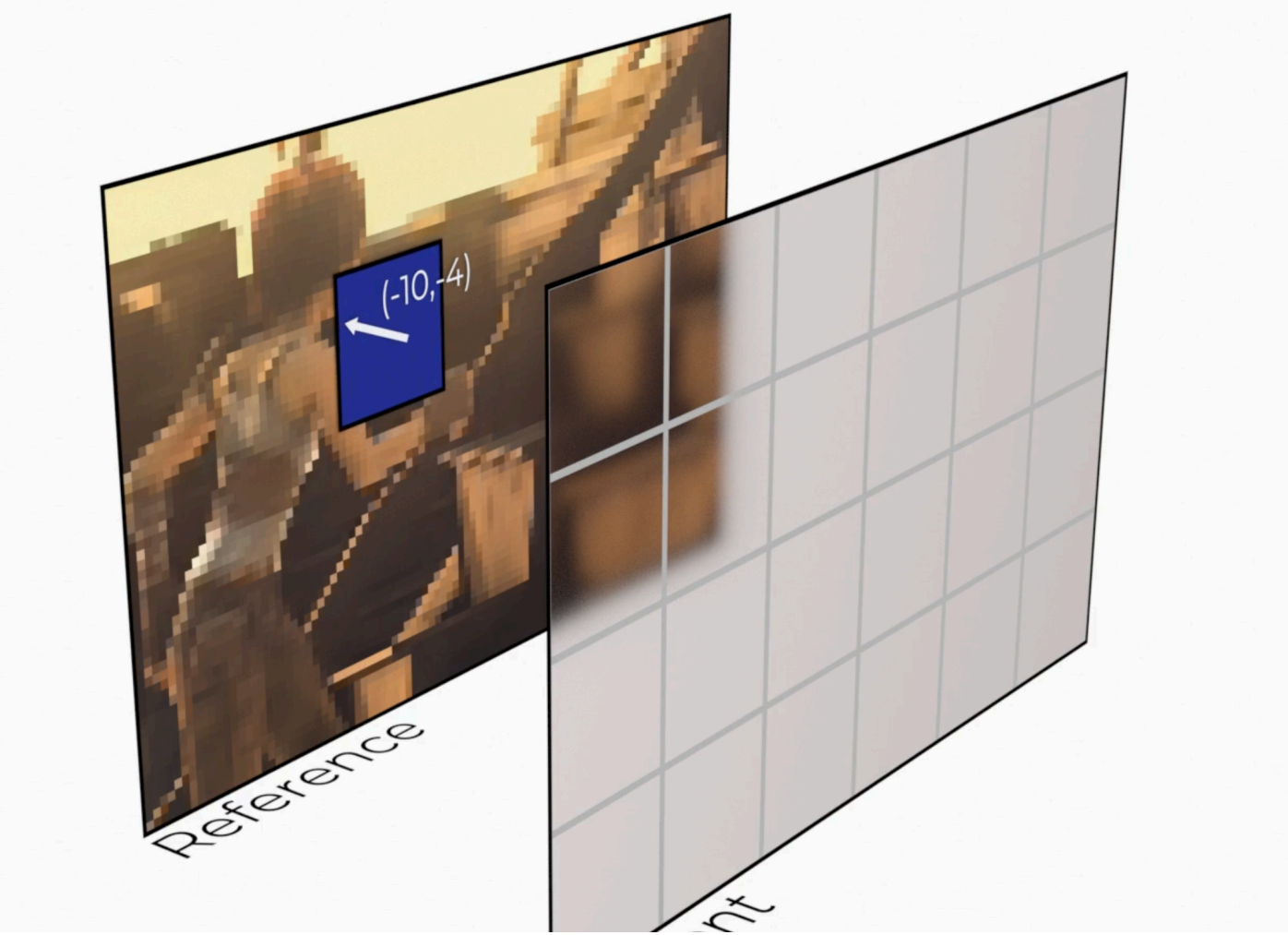
Encoding Loop Übersicht



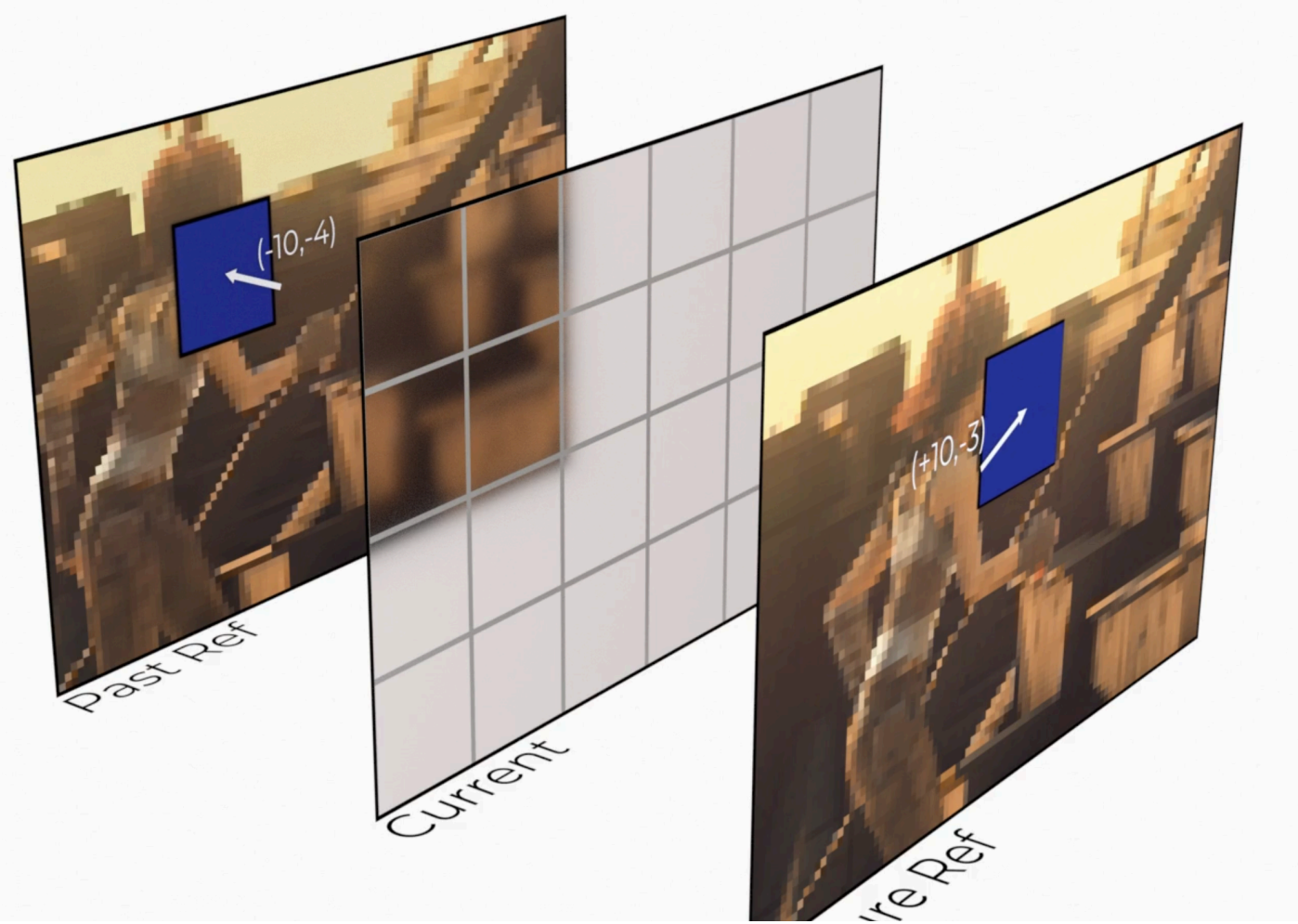
Intra-Prediction

- **Direction estimation**: Man versucht Bewegungsvektoren einzelner Blöcke zu finden
- Dann wird auch hier nur die Differenz gespeichert, die sehr klein sein sollte

Inter-Prediction mit P Frame



Inter-Prediction mit B Frame



Inter-Prediction

Sum of Absolute Transformed Differences (SATD):

- Ein Maß für die Ähnlichkeit von Blöcken.
- Berechnet die Differenz zwischen Original und Vorhersage nach Transformation.
- Wichtig für die Auswahl der besten Kompressionsstrategie.

Decoding Loop

1. **Dekomprimierung:**
    - Rückgängig machen der Quantisierung.
2. **Inverses Transformieren:**
    - Rekonstruktion des Bildsignals aus den Differenzen.
3. **Vorhersage-Rekonstruktion:**
    - Kombination aus vorhersagten und dekodierten Daten ergibt das Bild.

Unterschiede zwischen H.261 (1988) und H.264 (2004)

H.261	H.264
Blöcke: 8x8	Flexiblere Blockgrößen
Kein B-Frame-Support	Unterstützung für B-Frames
Einfaches Bewegungsmodell	Fortgeschrittene Bewegungskompensation
Niedrigere Kompression	Höhere Kompression durch CABAC

Unterscheidung Format, Container, Codec

Formell muss man zwischen *video coding format* und *codec* unterscheiden. Das Format enthält nur eine Spezifikation, der Codec benötigt noch eine Implementierung. Des Weiteren gibt es *video container formats* also Formate, die nur den Container beschreiben und damit wie die Daten gespeichert werden.

Auf Wikipedia findet man ausführliche Listen der gängigen Formate und Codecs im Vergleich:

- Video container formats
- Audio coding formats
- Video codecs

FourCC

Oft findet man noch einen FourCC, also einen vierstelligen Schlüssel für den verwendeten Codec in der Videodatei.

Die Idee kam 1985 von Apple und wurde dann von Electronic Arts auf dem Amiga eingesetzt und später wurden neue Video Codecs von Microsoft zur Kennzeichnung registriert. Heute werden an verschiedenen Stellen Listen geführt, aber diese Kennzeichnung ist eigentlich nicht mehr notwendig.

Multiplexing

In der Videotechnik wird Software oder Hardware, die das Zusammenführen von Audio und Videosignalen in einen Container umsetzt, auch gern als *multiplexer* oder kurz *muxer* (bzw. *demuxer* für das Extrahieren) bezeichnet.

Generell bedeutet Multiplexing aber das Zusammenführen von Signalen. Dies ist zum Beispiel notwendig um mehrere Signale über ein Kabel zu übertragen.

Wenn der Kanal genug Bandbreite hat, kann man time division multiplexing (TDM) machen, also die Signale in kurzen Zeitabständen hintereinander versenden.