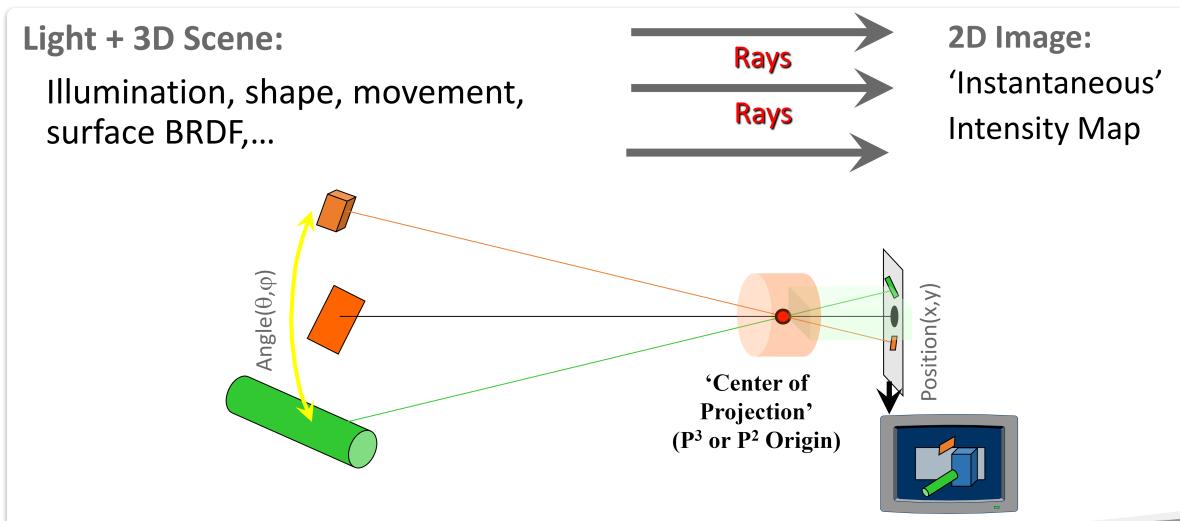


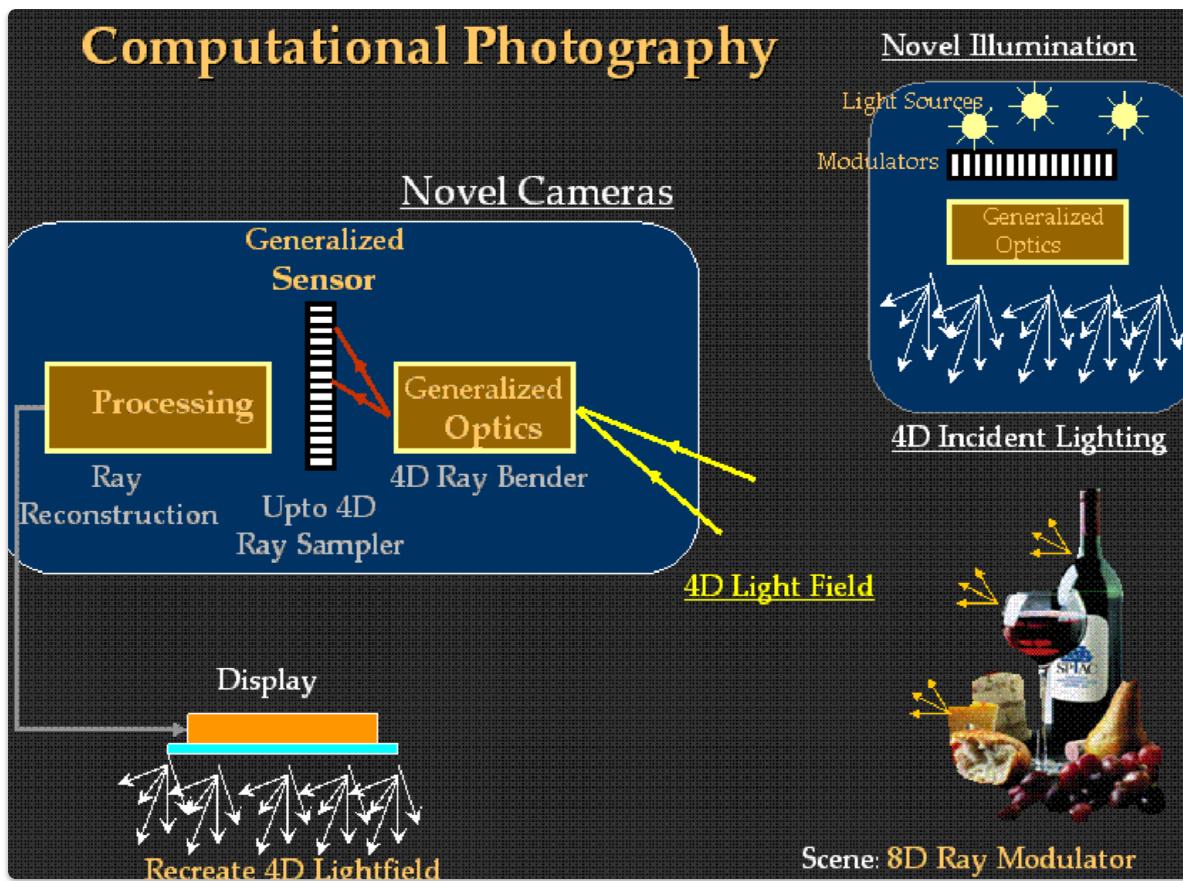
12. Computational Photography

[EVC_Skriptum_CV](#), p.60

Computational Photography ist ein Forschungsgebiet, das die **Verbreitung von Digitalkameras** nutzt, um die alltägliche Fotografie zu verbessern.

- **Ziele:**
 - Neue Aufnahme- und Bearbeitungsmethoden entwickeln, die normalerweise fotografisch nicht möglich wären.
 - Beispiele: Aufnahme bei extremem Kontrastumfang, verbesserte verwackelte Bilder, Komposition mehrerer Personen oder Objekte zu einem neuen Gesamtbild, Entfernen störender Personen oder Ähnliches.
 - Übergeordnetes Ziel ist es, neue Funktionalitäten und Anwendungsgebiete für die Fotografie zu erschließen.
- **Grundlage:** Mittels algorithmischer ("Computational") Ansätze werden die physikalischen Grenzen traditioneller Projektionskameras überwunden.
 - Ermöglicht Bilderzeugung, die mit herkömmlichen Kameras nicht realisierbar wäre (z.B. große Tiefenschärfe, hohe Auflösung, geringer Verlust der dritten Dimension).
 - Durch Kombination von Software, digitalen Sensoren, moderner Optik sowie intelligenter Beleuchtung verschiebt Computational Photography die Grenzen der traditionellen Fotografie.

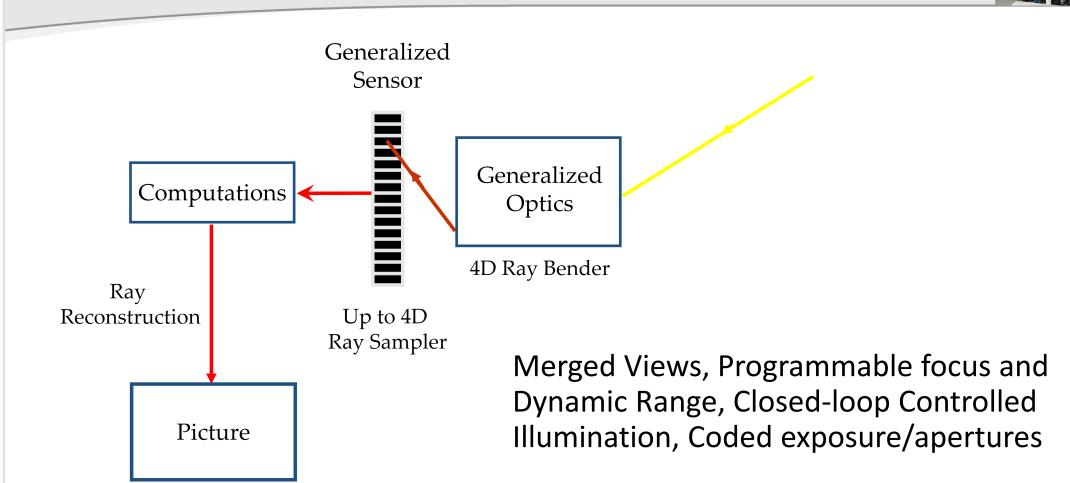




Moderne Kameras vs. Traditionelle Kameras

- **Moderne Kameras (Computational Photography):**
 - Komplexe Produkte, die eine über hundertjährige technische Evolution durchlaufen haben.
 - Grundlegender optischer Aufbau aus Linse(n), Blende und Sensor ist praktisch unverändert.
 - Aktuelle Systeme sind in der Lage, Leistungen zu erbringen, die das menschliche Auge übertreffen.

Computational Photography: Optics, Sensors and Computations



- **Traditionelle Kameras:**

- Nachteile, besonders bei der Bildaufnahme von traditionellen Kameras:
Führen dem Verlust einer Dimension der Szenengeometrie zu.

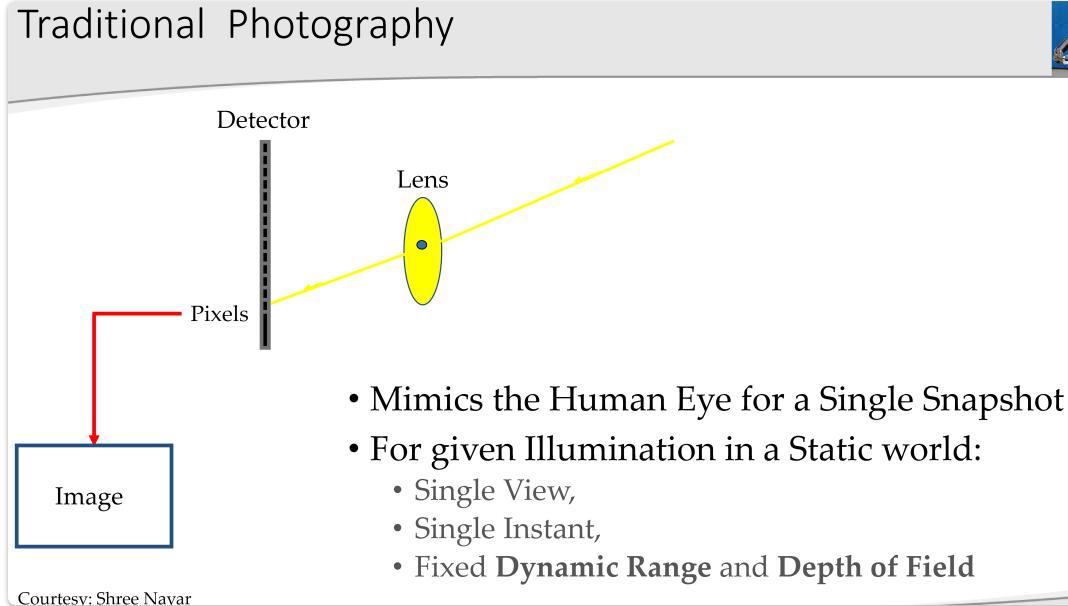
Der Sensor summiert einfallende Lichtstrahlen aus allen Einfallswinkeln zu einer Intensität.

Informationen über die Reflexionseigenschaften von Szenenoberflächen (charakterisiert durch die Bidirectional Radiance Distribution Function, kurz BRDF) gehen verloren.

Lassen sich aus der Bestrahlungsstärke auf der Bildebene nicht die strahlungsdichten Objekte bestimmen.

* Die **Objektreffektivität** und der **unbekannten Bestrahlungsstärke** des Objekts werden zusammengesetzt dargestellt, was die Identifizierung und Klassifizierung von Objekten erschwert ("Helligkeits-Dilemma").

Traditional Photography



Zusätzliche optische Komponenten und deren Vorteile

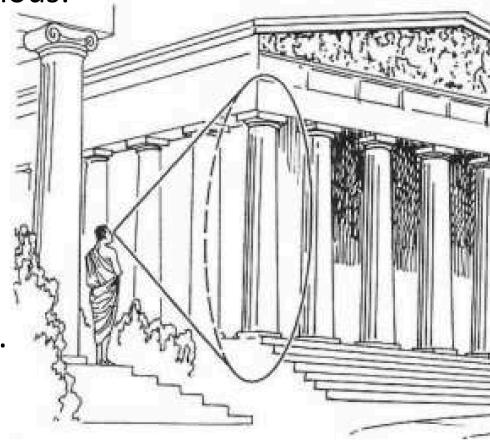
- **Abbildung in Kombination mit einer Blende zur Reduktion der Tiefunschärfe:**
 - Führt zu einem Verlust brauchbarer Informationen für weite Tiefenbereiche der Szene, da nur ein kleiner Bereich scharf abgebildet wird.
 - Verfügbare Sensoren integrieren über ein breites Spektrum der Wellenlängen einfallendes Licht.
 - Messung des spektralen Dimensions von Oberflächen (z.B. Farben) an unterschiedlichen Orten.
 - Zeit (Belichtungszeit) wird ebenfalls in die Dimension des Lichts integriert (zeitliche Auflösung geht verloren).
 - Räumliche Auflösung und die Quantisierung zu einer begrenzten Helligkeitsauflösung sind ebenfalls reduziert.
 - Dies führt zu einer **Reduktion des Signalumfangs (Dynamic Range)**.

- Core ideas are ancient, simple, and seem obvious:

- **Lighting:** ray sources
- **Optics:** ray bending / folding devices
- **Sensor:** measure light
- **Processing:** assess it
- **Display:** reproduce it

- **Ancient Greeks:**

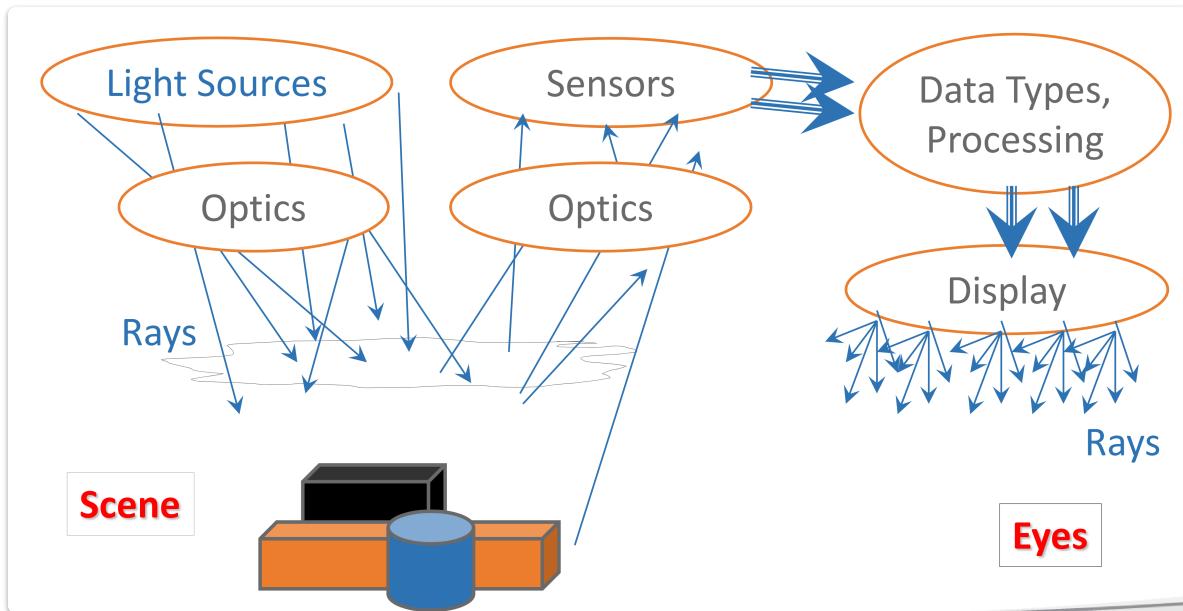
'Eye rays' wipe the world to feel its contents...



12.2 Lichtfeld

[EVC_Skriptum_CV](#), p.60, p.61

Das Lichtfeld ist eine Funktion, die die **Lichtmenge** beschreibt, die an jedem Punkt des dreidimensionalen Raums aus allen Richtungen einfällt.



- In der Computational Photography betrachtet man das Lichtfeld, das auf einen Punkt fällt, d.h., **welche Strahlen an diesen Punkt reflektiert werden**.
- **Objekte** werden als **Volumen** betrachtet.
- Einfallendes Beleuchtungsfeld wird in ein Beleuchtungsfeld umgewandelt, welches vom Objekt reflektiert wird.

4D-Lichtfeld-Parametrisierung

[EVC_Skriptum_CV](#), p.61

Einfallende Beleuchtung kann mittels eines **4D-Lichtfelds** parametrisiert werden.



- Man stellt sich um das Objekt eine Kugel vor, die das Objekt einschließt.
- **Position auf der Kugel:** (u_i, v_i)
- **Richtung des eintreffenden Lichts:** (θ_i, ϕ_i)
- **Einfallendes Lichtfeld:** $R_i(u_i, v_i, \theta_i, \phi_i)$

Das ausgehende Licht kann ebenfalls mittels eines **4D-Lichtfelds** parametrisiert werden:

- **Was das Licht die umgebende Oberfläche verlässt:** $R_r(u_r, v_r, \theta_r, \phi_r)$

Hier gabs dann noch ein Recap zur [PLenopischen Funktion: EVC-CV12-Computational Photography_S2025_Slides, p.11](#)

Reflexionsfeld (8D-Funktion)

Die **Reflexionseigenschaften eines Objekts** können mittels einer **achtdimensionalen Funktion** - dem **Reflexionsfeld** - charakterisiert werden.

- Das Reflexionsfeld codiert für jeden einfallenden Lichtstrahl die 4D Reflexion des Lichtstrahls.
- Es beinhaltet die nötigen Informationen, um das Objekt unter verschiedenen Beleuchtungsgegebenheiten und von verschiedenen Blickwinkeln aus zu rendern.

4D-Lichtfeld-Kameras (Plenoptische Kameras)

4D-Lichtfeld-Kameras sind nicht nur in der Lage, Ort und Intensität der Projektion der von einer Szene ausgehenden Lichtstrahlen zu detektieren, sondern auch den **Einfallswinkel** auf der Kamera-Ebene zu quantifizieren.



- Die so gewonnene **4D-Abtastung des Lichtfelds** einer Szene ermöglicht einen vielfältigen Informationsgewinn.
- **Aufbau:** 4D-Lichtfeld-Kameras verwenden ein **Mikrolinsen-Array**, das direkt auf dem Sensor aufgebracht wird.
- **Funktionsweise:** Die eintreffenden Lichtstrahlen werden durch die einzelnen Pixel im Sensorbereich unter jeder Mikrolinse nach ihrem Einfallswinkel quantifiziert, um ein 4D-Lichtfeld einer Szene zu rekonstruieren.
- **Nachteil:** Die direkt erreichbare **Ortsauflösung** entspricht der **Anzahl der Mikrolinsen**. Deren Durchmesser (in Pixeln) bestimmt wiederum die **Winkelauflösung**.
 - Je größer die Winkelauflösung, desto geringer ist die Ortsauflösung und umgekehrt.

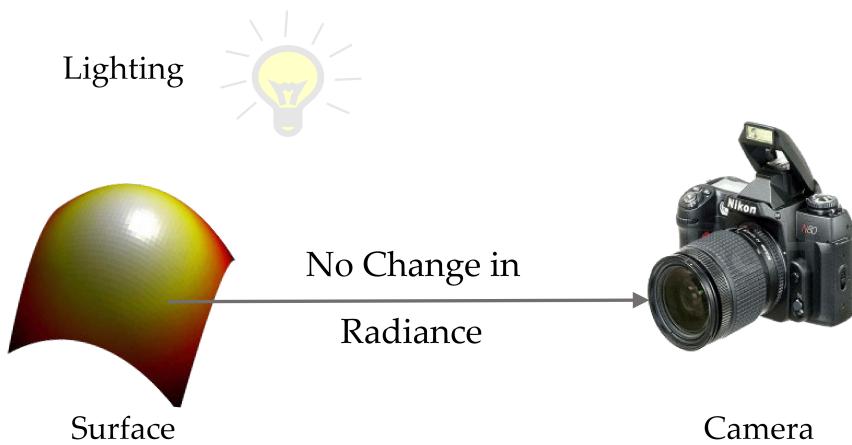
Ray

- Let's not worry about time and color:



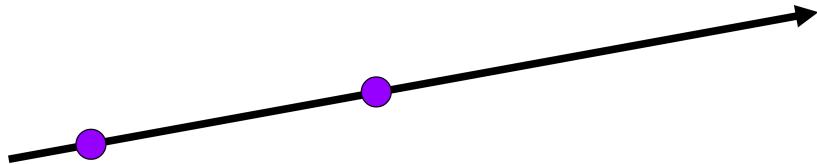
- 5D
 - 3D position
 - 2D direction

How can we use this?



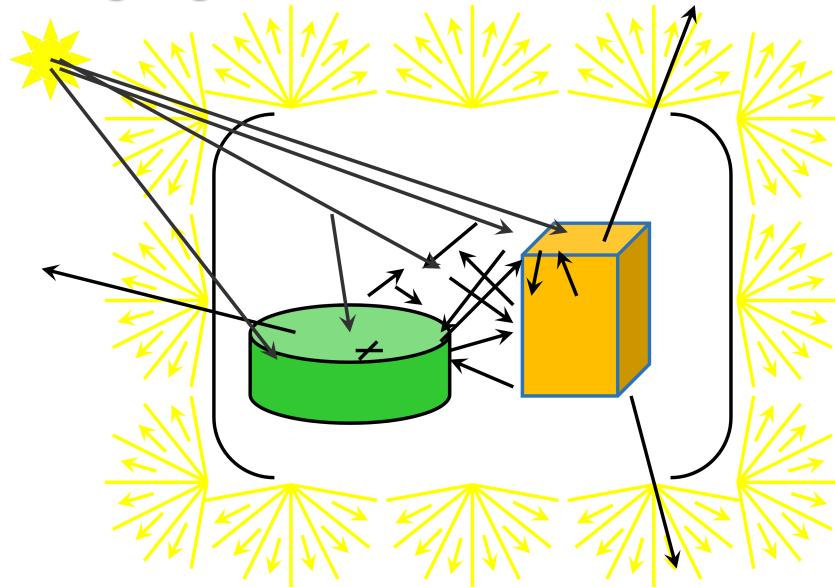
Ray Reuse

- Infinite line
 - Assume light is constant (vacuum)



- 4D
 - 2D direction
 - 2D position
 - Non-dispersive medium

- Measure all the **outgoing** light rays.



(u_i, v_i) indicate the position on the surface where the light enters,

(θ_i, ϕ_i) indicate the direction in which it enters.

(u_r, v_r) indicate the position on the surface where the light leaves,

(θ_r, ϕ_r) indicate the direction in which it leaves.



$$(u_i, v_i, \theta_i, \phi_i ; u_r, v_r, \theta_r, \phi_r)$$

8D reflectance field

Since it is linear, we can represent as a matrix

$$R(u_i, v_i, \theta_i, \phi_i ; u_r, v_r, \theta_r, \phi_r)$$

360 x 180 x 180 x 180 x 360 x 180 x 180 x 180

= 4.4e18 measurements

x 6 bytes/pixel (in RGB 16-bit)

= 26 exabytes (billion GB)

= 1.3 million 20TB hard drives



(für nur ein Bild)

High Dynamic Range (HDR)

[EVC_Skriptum_CV, p.61](#)

High Dynamic Range (HDR) beschreibt eine Fülle an Techniken, die im Vergleich zu herkömmlichen Techniken für digitale Bildgebung oder fotografischen Methoden einen **größeren Bereich zwischen den hellsten und dunkelsten Regionen eines Bildes** darzustellen ermöglichen.

- **Ziel von HDR-Bildern:** Den Helligkeitsbereich in realen Szenen genauer zu repräsentieren.
- **Möglicher Helligkeitsbereich:** Erstreckt sich von direktem Sonnenlicht bis hin zu schwachem Sternenlicht.
- **Erzeugung:** Erzielt durch die Aufnahme von mehreren Bildern mit unterschiedlicher Belichtung der gleichen Szene.
- **HDR kompensiert Detailverlust**, der bei Nicht-HDR-Kameras mit einem einzigen Belichtungslevel auftritt (dort erhält man ein Bild mit Detailverlust in hellen oder dunklen Bildbereichen).
- **Ergebnis:** Ein Bild, dessen dunkle als auch helle Bildbereiche im Detail darstellbar sind.



Anzeige von HDR-Bildern auf Geräten mit kleinem dynamischen Bereich

Um HDR-Bilder auf Geräten mit einem kleinen dynamischen Bereich anzuzeigen, werden **Techniken zur Abbildung von Farbtönen (engl. Tone Mapping)** benötigt.

- **Ziel von Tone Mapping:** Den gesamten Kontrast reduzieren.
- **Grund:** Ausdrücke, CRT- oder LCD-Monitore und Projektoren haben einen begrenzten dynamischen Bereich, der für die Darstellung der vollen Bandbreite der Lichtintensitäten ungeeignet ist.
- **Funktionsweise:** Tone Mapping behandelt die **Kontrastreduktion** der Lichtintensitäten der Szene zu den anzeigbaren Wertebereichen.
- **Wichtige Aspekte dabei:** Bilddetails und die Farbwirkung sollen möglichst erhalten bleiben, um den originalen Szeneninhalt wahrnehmen zu können.

Beispiel zu ToneMapping:

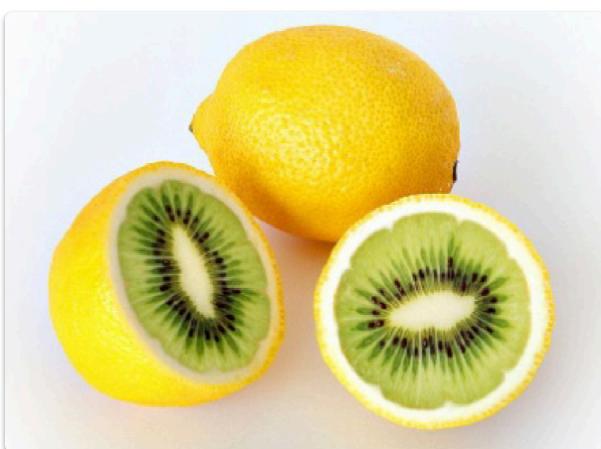


Fotomontage/Komposition

EVC_Skriptum_CV, p.61

Fotomontage (auch bekannt als "Komposition" oder "image compositing") bezeichnet das Schneiden und Zusammenfügen von verschiedenen Fotos zu einem zusammengesetzten Bild.

- **Realisierung:** Heutzutage mittels moderner Bildbearbeitungssoftware.
- **Ziel:** Visuelle Elemente von unterschiedlichen Quellen werden in einem einzigen Bild kombiniert, um die **Illusion zu schaffen, dass alle diese Elemente Teil derselben Szene sind.**





- Bei Filmaufnahmen ist die Komposition unter verschiedenen Bezeichnungen bekannt:
 - "chroma key"
 - "blue screen"
 - "green screen" und andere Bezeichnungen.

Image Inpainting

[EVC_Skriptum_CV](#), p.62

Inpainting ist ein Bildbearbeitungsprozess, der bei der **Rekonstruktion von schlecht erhaltenen oder nicht vorhandenen Bild- oder Videoteilen** Anwendung findet.

- Analogie zur Malerei:** Im Bereich der Malerei wäre dies die Arbeit eines ausgebildeten Bildrestaurators.
- In der digitalen Welt (auch bekannt als Bild- oder Videointerpolation):** Bezieht sich auf die Anwendung hochentwickelter Algorithmen, um **verlorene oder fehlerhafte Teile von Bilddaten zu ersetzen**.



- Anwendungsbeispiele:**
 - Entfernung eines Objekts aus einer Szene**, um ein beschädigtes Bild zu retuschieren.
 - Für Fotografie und Film:**
 - Beseitigung von Zerfallsmerkmalen (z.B. Risse in Fotografien, Kratzer und Staubflecken im Film).
 - Hinzufügen oder Entfernen von Bildteilen (z.B. Aufnahmedatum).
- Generelles Ziel:** Ein verändertes Bild zu produzieren, in dem die Inpaint-Region mit dem restlichen Bild so verschmolzen wird, dass ein gewöhnlicher Betrachter **nicht bemerkt**,

dass das Bild bearbeitet wurde.

Warping

EVC_Skriptum_CV, p.62

Warping ist der Prozess der **Bildmanipulation**, bei dem jede im Bild vorkommende Form **räumlich signifikant verändert** wird.

Image filtering: change **range** of image

$$g(x) = T(f(x))$$

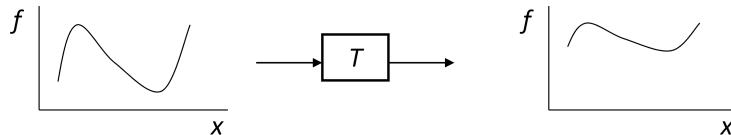


Image warping: change **domain** of image

$$g(x) = f(T(x))$$

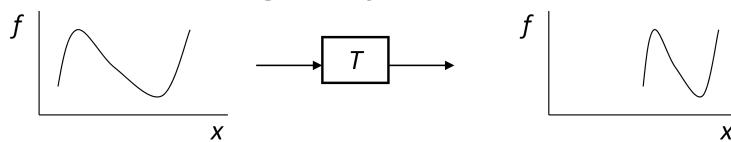


Image warping: change **domain** of image



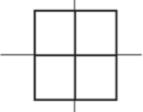
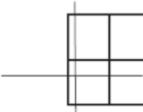
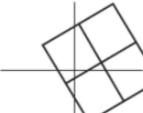
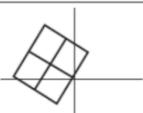
$$g(x) = f(T(x))$$



- **Anwendung:**

- **Verzerrung von Bildern korrigieren.**
- **Kreative Zwecke** (z.B. für Morphing).

- **Funktionsweise:** Beim reinen Warping wird der **Punkt auf einen anderen Punkt abgebildet, ohne die Farbe des Punktes zu ändern**.
- **Grundlage:** Warping kann auf jeder **Funktion** basieren, die (von einem Teil) einer Ebene auf eine Ebene abbildet.
- **Injektive Funktion:** Bei einer injektiven Funktion kann das Original wieder rekonstruiert werden.
- **Bijektive Funktion:** Wenn die Funktion bijektiv ist, kann das Bild invers transformiert werden.

Identity	$W(x) = x$	
Translation	$W(x; t) = x - t$	
Rigid	$W(x; R, t) = Rx - t$	
Similarity	$W(x; R, a, t) = aRx - t$	
Affine	$W(x; A, t) = Ax - t$	

where $\alpha \in \mathcal{R}$, $A \in \mathcal{R}^{2 \times 2}$, $t \in \mathcal{R}^2$, $R \in \mathcal{R}^{2 \times 2}$, $|R| = 1$

- **Mathematische Beschreibung:**

- Gegeben sei ein gesampeltes Bild $I(x)$.
- **Warping** ist die **räumliche (geometrische) Deformation** des Quellbildes $I_s(x)$.
- Das Ergebnis $I_d(x)$ wird **Zielbild** genannt.
- Das Warping wird mittels der **Warpingfunktion** $W(x; p)$ mit den Parametern p bestimmt.
- Für jedes Pixel x in I_d gibt uns die Warpingfunktion an, woher dieses Pixel im Quellbild I_s stammt: $I_d(x) = I_s(W(x; p))$.

- **Zwei Formen von Transformationen:**

- **Affine Transformationen**
- **Projektive Transformationen** (auch perspektivische Transformation oder Homographie genannt).

Affine Transformation

Die affine Transformation ist eine Transformation, die **gerade Linien und Distanzverhältnisse zwischen Punkten, die auf einer Linie liegen, erhält**.

- **Beispiel:** Der Mittelpunkt eines Liniensegments bleibt auch nach der Transformation dessen Mittelpunkt.
- **Wird nicht zwingend erhalten:** Längen und Winkel.
- **Typische affine Transformationen:** Translation (Verschiebung), Rotation (Drehung), Scherung und ähnliche Transformationen.
- **Kombinationen** von affinen Transformationen sind ebenfalls affin.

$$\begin{bmatrix} x' \\ y' \\ w \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ w \end{bmatrix}$$

Projektive Transformation

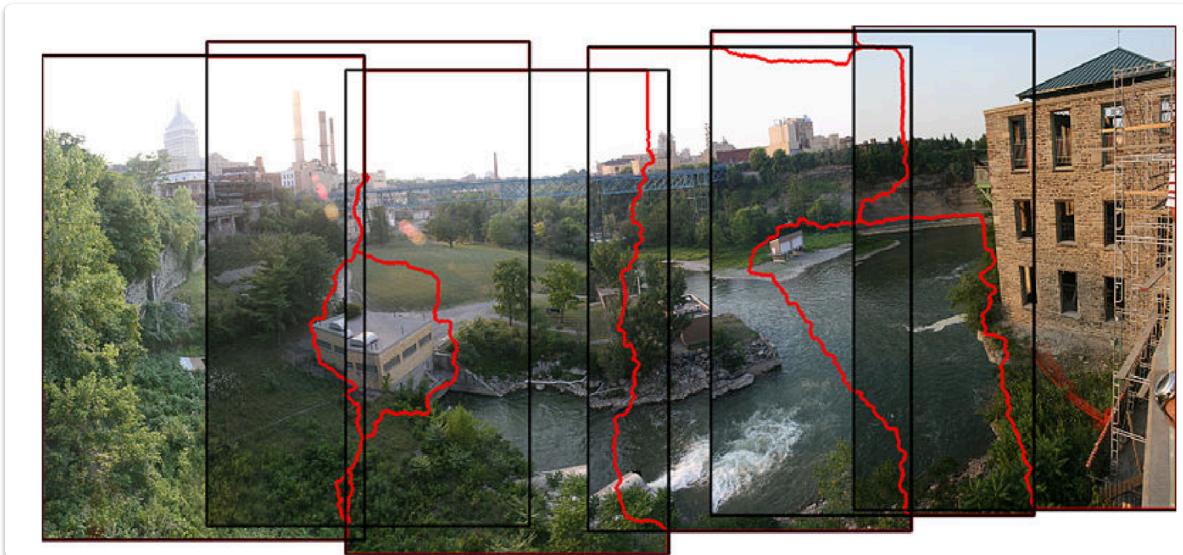
Die projektive Transformation einer Ebene ist eine Transformation, die in der **projektiven Geometrie** verwendet wird.

- **Wird nicht erhalten:** Größen oder Winkel.
- **Anwendung:** Jedes Bildpaar, das dieselbe ebene Fläche im Raum zeigt, steht mittels einer projektiven Transformation zueinander in Beziehung.
- **Praktische Anwendungen:**
Bildrektifizierung (Entzerrung von Bildern).
Bildregistrierung (Ausrichtung von Bildern zueinander).
* Berechnung der Kamerabewegung zwischen zwei Bildern.

$$\begin{bmatrix} x' \\ y' \\ w \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} x \\ y \\ w \end{bmatrix}$$

Bildmosaik (Image Mosaicing / Stitching)

EVC_Skriptum_CV, p.63



Der Prozess zur Erstellung eines **Bildmosaiks** (auch oft als **Stitching** bezeichnet) kombiniert mehrere Bilder mit überlappenden Bereichen zu einem **hochauflösenden Panoramabild**.

- **Voraussetzung für korrekte Überlagerung (ohne Parallaxenfehler):** Die Kamera muss um den **Fokuspunkt** rotiert werden.
- **Der Prozess kann in drei Bestandteile aufgeteilt werden:**
 1. **Bildregistrierung**
 2. **Kalibrierung**

3. Blending

1. Bildregistrierung

Bei der Bildregistrierung werden **korrespondierende lokale Merkmale** zwischen den Bildern bestimmt, um die Bilder korrekt "übereinander legen" zu können.

2. Kalibrierung

Die Kalibrierung versucht, die **Differenzen zwischen einem idealen Linsenmodell und dem realen Linsensystem der Kamera(s)** zu minimieren. Das bedeutet, sie korrigiert:

- Optische Defekte (wie **Verzerrungen**).
- Unterschiedliche **Belichtungszeiten** der Bilder.
- **Vignettierung** (Randabdunkelung).
- **Chromatische Aberrationen** (Farbsäume).

3. Blending ("Verschmelzen")

Das Blending korrigiert die im Kalibrierungsschritt festgemachten Abweichungen und **bildet die einzelnen Aufnahmen auf ein gemeinsames Ausgabebild ab**.

- **Farben werden zwischen den Bildern angepasst**, um die Beleuchtungsunterschiede auszugleichen.
- Die Bilder müssen so miteinander verschmolzen werden, dass **keine Übergänge (engl. seams)** von einem Einzelbild zum nächsten sichtbar sind.

Image Morphing

[EVC_Skriptum_CV](#), p.63

Image Morphing ist ein spezieller Bildeffekt, der ein Bild **nahtlos in ein anderes verwandelt**.

- **Anwendung:** Oft in surrealen Sequenzen oder im Fantasyfilm-Genre verwendet, um z.B. eine Person in eine andere zu verwandeln.
- **Traditionelle Methode:** Durch Aus- und Einblendtechniken (engl. *fading*) im Film erzielt.
- **Seit den frühen 1990ern:** Immer häufiger Computertechniken verwendet, die überzeugendere Ergebnisse lieferten.



- **Heutige Methode (Computertechniken):** Beim Überblenden werden die Bilder gleichzeitig anhand von **markierten korrespondierenden Punkten verzerrt**.
- **Beispiel (Gesichts-Morphing):**
 - Umwandlung eines Gesichts in ein anderes.
 - **Keypoints** im ersten Gesicht (z.B. die Kontur der Nase oder Position der Augen) werden markiert. Diese Keypoints müssen auch im zweiten Bild existieren.
 - Anschließend wird das erste Gesicht **verzerrt**, um die Form des zweiten Gesichts zu erhalten.
 - **Gleichzeitig** wird das erste Gesicht **ausgeblendet** und das zweite Gesicht **eingebettet**.
- **Höher entwickelte Überblendungstechniken (später):** Verschiedene Teile eines Bildes werden graduell auf das andere überblendet, anstatt das gesamte Bild auf einmal zu wandeln.

Idea #1: Cross-Dissolve



- Interpolate whole images:

$$\text{Image}_{\text{halfway}} = (1-t) * \text{Image}_1 + t * \text{Image}_2$$
- This is called **cross-dissolve** in the film industry
- But what if the images are not aligned?

- Align first, then cross-dissolve
 - Alignment using global warp – picture still valid



Testähnliches Beispiel

- Benennen Sie die dargestellten Verfahren der Computational Photography:

Eingabebild(er)	Ergebnis	Verfahren
		<u>Image Inpainting</u>
		<u>Image Morphing</u>
		<u>Panoramic Images</u>
		<u>HDR – High Dynamic Range Imaging</u>