Introduction to Computer Vision
Taught by

- prof. Luc Van Gool
- Prof. Ender Konukoglu
- Guest starring by prof Orçun Göksel

The course comes with a course text that covers most – but not all! – material. Slide decks for all lectures will be made available on eDoz or similar.
We got questions about which course to take

*Computer Vision (D-INFK)*, or

*Image Analysis and Computer vision (this course)*

**IN ANY CASE, DO NOT TAKE BOTH!**

If you took the introductory course on CV at D-INFK, then best take *Computer Vision*

If you did not take that course, then best take *Image Analysis and Computer Vision*
... it is crucial ...
Vision is important

- half our brain is devoted to it
- developed many times during evolution
- it is non-contact
- it can be implemented with high resolution
- works with ambient E-M waves
- yields colour, texture, depth, motion, shape
The central take-home message:

For people vision is their most crucial sense, for good reason
... it is intriguing ...
The perception of intensity
The perception of color

The red squares have equal color…
The perception of length
The perception of length

The horizontal lines are equally long…
The perception of lines being straight
The perception of parallelism
The perception of curvatures

Illusions: interference of differently oriented patterns via adaptation
The perception of motion

The `barber pole’ rotates about the vertical, it does not translate vertically…
It’s not that more context solves it all…

there is literally more than meets the eye, i.e. a lot of massively parallel processing
The perception of intensity
Computer Vision

INTRO

perception applications light
Parallelism again…
Fill-in: averaging of perceived contrast at edges over regions possibly obtained via extrapolation of the edges… in any case such illusion seems to help people to detect patterns in the world.
The role of context

Human vision: Biederman, Bar & Ullman, Palmer, ...

Computer Vision

INTRO

perception applications light
All encircled patterns are identical:

The role of context
The role of context
The role of context
The role of context
The role of context
The role of context

human vision is much more than a bottom-up process of subsequent signal processing steps.
The central take-home message:

**Effective vision needs more than sheer filtering and measuring**
... it is hot ...
The explosion of photography

![Graph showing the number of photos taken over time, with a significant increase in the late 20th century. The graph indicates a sharp rise in 'All photos' and a more gradual rise in 'Analog photos'.]
The explosion of photography

Easier than ever to take a photo
The cost is extremely low (cheap memory)
Most people carry a camera most of the time
The development of computer vision apps

Most early applications were found in production environments, as these allow for controlled conditions and have little uncertainty.

Some areas do not allow for much control: medical IP, remote sensing, surveillance, etc.

Currently CV is conquering the less controllable areas by storm.
Ex App: autonomous vehicles
Ex App: autonomous vehicles

car detection:
Ex App: autonomous vehicles

putting vision modalities together:
Ex: autonomous mobile platform
## Computer Vision

### Ex App: Image Retrieval, Captioning, ...

<table>
<thead>
<tr>
<th>Description</th>
<th>Image Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A person riding a motorcycle on a dirt road.</td>
<td><img src="image1" alt="Person on motorcycle" /></td>
</tr>
<tr>
<td>Two dogs play in the grass.</td>
<td><img src="image2" alt="Dogs playing" /></td>
</tr>
<tr>
<td>A skateboarder does a trick on a ramp.</td>
<td><img src="image3" alt="Skateboarder" /></td>
</tr>
<tr>
<td>A dog is jumping to catch a frisbee.</td>
<td><img src="image4" alt="Dog catching frisbee" /></td>
</tr>
<tr>
<td>A group of young people playing a game of frisbee.</td>
<td><img src="image5" alt="Frisbee game" /></td>
</tr>
<tr>
<td>Two hockey players are fighting over the puck.</td>
<td><img src="image6" alt="Hockey game" /></td>
</tr>
<tr>
<td>A little girl in a pink hat is blowing bubbles.</td>
<td><img src="image7" alt="Blowing bubbles" /></td>
</tr>
<tr>
<td>A refrigerator filled with lots of food and drinks.</td>
<td><img src="image8" alt="Refrigerator" /></td>
</tr>
<tr>
<td>A herd of elephants walking across a dry grass field.</td>
<td><img src="image9" alt="Elephants" /></td>
</tr>
<tr>
<td>A close up of a cat laying on a couch.</td>
<td><img src="image10" alt="Cat" /></td>
</tr>
<tr>
<td>A red motorcycle parked on the side of the road.</td>
<td><img src="image11" alt="Motorcycle" /></td>
</tr>
<tr>
<td>A yellow school bus parked in a parking lot.</td>
<td><img src="image12" alt="School bus" /></td>
</tr>
</tbody>
</table>
Ex App: visual surveillance
Computer Vision

Ex App: Augm. Reality, eg sports
Ex App: motion capture for movies/games
Ex App: computer-assisted surgery
Mobile mapping
The central take-home message:

It is feasible now to let most things see and interpret their environment.
... it needs light ...
And then there was Light…

- no vision without light…
- … because it is influenced by objects
Kickoff: the light, surface, lens & cam
Kickoff: the light, surface, lens & cam
topics

- the nature of light
- interactions with matter
An option on optics

1. Geometrical optics
2. Physical optics, or
3. Quantum-mechanical optics

⇒ wave character
Light as electromagnetic waves
Light as electromagnetic waves

Self-sustaining exchange of electric and magnetic fields

1. wavelength
2. direction
3. amplitude $E$
4. phase
5. direction of polarisation
The spectrum

Normal ambient light is a mixture of wavelengths, polarisation directions, and phases.
## The visible range

<table>
<thead>
<tr>
<th>Wavelength (in nm)</th>
<th>Colour</th>
</tr>
</thead>
<tbody>
<tr>
<td>380 - 450</td>
<td>violet</td>
</tr>
<tr>
<td>450 - 490</td>
<td>blue</td>
</tr>
<tr>
<td>490 - 560</td>
<td>green</td>
</tr>
<tr>
<td>560 - 590</td>
<td>yellow</td>
</tr>
<tr>
<td>590 - 630</td>
<td>orange</td>
</tr>
<tr>
<td>630 - 760</td>
<td>red</td>
</tr>
</tbody>
</table>

**NOTE**: Cameras may have different spectral sensitivities (i.e. also different from human vision)
NOTE: animals may have different spectral sensitivities (i.e. different from human vision), and may also have a Different number of cone types, like 4 in most birds.
Also cams for non-visible `light’, e.g. infrared

Overheating of transformer coils, with far IR

Near infra-red (NIR) space image

NRG -> RGB for visualization (notice the strong reflection in the NIR for vegetation)
Computer Vision

**Interactions with matter**

four types:

<table>
<thead>
<tr>
<th>phenomenon</th>
<th>example</th>
</tr>
</thead>
<tbody>
<tr>
<td>absorption</td>
<td>blue water</td>
</tr>
<tr>
<td>scattering</td>
<td>blue sky, red sunset</td>
</tr>
<tr>
<td>reflection</td>
<td>coloured ink</td>
</tr>
<tr>
<td>refraction</td>
<td>dispersion by a prism</td>
</tr>
</tbody>
</table>

+ diffraction
Interactions with matter

four types:

<table>
<thead>
<tr>
<th>phenomenon</th>
<th>example</th>
</tr>
</thead>
<tbody>
<tr>
<td>absorption</td>
<td>blue water</td>
</tr>
<tr>
<td><strong>scattering</strong></td>
<td>blue sky, red sunset</td>
</tr>
<tr>
<td>reflection</td>
<td>coloured ink</td>
</tr>
<tr>
<td>refraction</td>
<td>dispersion by a prism</td>
</tr>
</tbody>
</table>

+ diffraction
Scattering

3 types depending on relative sizes of particles and wavelengths:

1. small particles: *Rayleigh* (strongly wavelength dependent)
2. comparable sizes: *Mie* (weakly wavelength dependent)
3. Large particles: *non-selective* (wavelength independent)
Wavelength dependence

Less haze in the infrared (long wavelengths -> little scatter)
Looking through clouds by radar (even longer wavelengths)
NOTE: without scatter we would wander mainly in the dark
Atmospheric showcase

**Rayleigh:**
Tyndall effect (blue sky)
Red, setting sun

**Non-selective:**
Grey clouds

**Mie:**
Coloured cloud from volcanic eruption
Interactions with matter

four types:

<table>
<thead>
<tr>
<th>phenomenon</th>
<th>example</th>
</tr>
</thead>
<tbody>
<tr>
<td>absorption</td>
<td>blue water</td>
</tr>
<tr>
<td>scattering</td>
<td>blue sky, red sunset</td>
</tr>
<tr>
<td>reflection</td>
<td>coloured ink</td>
</tr>
<tr>
<td>refraction</td>
<td>dispersion by a prism</td>
</tr>
</tbody>
</table>

+ diffraction
Mirror reflection
Mirror reflection

Angle of reflection = angle of incidence
Mirror reflection: dielectric

Polarizer at **Brewster angle**

Full reflection at grazing angles
Mirror reflection: conductor

- strong reflectors (under all angles)
- more or less preserve polarization
Roughness of surfaces leads to ‘diffuse’ reflection

(a) Mirror or ‘specular’ reflection, (b) diffuse reflection
... and to mixed reflection for most real surfaces

three types of reflection:

- diffuse
- specular
- mixed

Note: Lambertian example of diffuse reflection
Spectral reflectance
e.g. vegetation

![Graph showing spectral reflectance with curves for upper and lower leaf sides.](image)
Ideally: spectral BRDF at all points known

BRDF = bidirectional reflectance distribution function
## Interactions with matter

four types:

<table>
<thead>
<tr>
<th>phenomenon</th>
<th>example</th>
</tr>
</thead>
<tbody>
<tr>
<td>absorption</td>
<td>blue water</td>
</tr>
<tr>
<td>scattering</td>
<td>blue sky, red sunset</td>
</tr>
<tr>
<td>reflection</td>
<td>coloured ink</td>
</tr>
<tr>
<td>refraction</td>
<td>dispersion by a prism</td>
</tr>
</tbody>
</table>

+ diffraction
Refraction
Refraction

$n_1 \sin \Delta_i = n_2 \sin \Delta_t$

Snell’s law
Dispersion

Refraction is more complicated than mirror reflection: the path orientation of light rays is changed depending on material AND wavelength !!!
Interactions with matter

four types:

<table>
<thead>
<tr>
<th>phenomenon</th>
<th>example</th>
</tr>
</thead>
<tbody>
<tr>
<td>absorption</td>
<td>blue water</td>
</tr>
<tr>
<td>scattering</td>
<td>blue sky, red sunset</td>
</tr>
<tr>
<td>reflection</td>
<td>coloured ink</td>
</tr>
<tr>
<td>refraction</td>
<td>dispersion by a prism</td>
</tr>
</tbody>
</table>

+ diffraction
Absorption

Dissipation of wavelengths specific for the medium

Based on resonance frequencies of molecules -> peaks
Holes in sky light spectrum observed by Fraunhofer
The solar spectrum

Peaks around 500nm, hence human sensitivity for that part of the spectrum

Spectral composition of light above atmosphere

Spectral composition of light below atmosphere

… some wavelengths get strongly weakened due to absorption
Acquisition of Images
Acquisition of images

We focus on:

1. illumination
2. cameras
Acquisition of images

We focus on:

1. **illumination**
2. **cameras**
Acquisition of images

We focus on:

1. illumination
2. cameras
illumination
Illumination

Well-designed illumination often is key in visual inspection.

The light was good, but the hot wax was a problem...
Illumination techniques

Simplify the image processing by controlling the environment

An overview of illumination techniques:

1. back-lighting
2. directional-lighting
3. diffuse-lighting
4. polarized-lighting
5. coloured-lighting
6. structured-lighting
7. stroboscopic lighting
Back-lighting

lamps placed behind a transmitting diffuser plate, light source behind the object

generates high-contrast silhouette images, easy to handle with *binary vision*

often used in inspection
Example backlighting
Directional and diffuse lighting

**Directional-lighting**

- generate sharp shadows
- generation of specular reflection (e.g. crack detection)
- shadows and shading yield information about shape

**Diffuse-lighting**

- illuminates uniformly from all directions
- prevents sharp shadows and large intensity variations over glossy surfaces:
  - all directions contribute extra diffuse reflection, but contributions to the specular peak arise from directions close to the mirror one only
Crack detection
Example directional lighting
Example diffuse lighting
Polarized lighting

2 uses:

1. to improve contrast between Lambertian and specular reflections

2. to improve contrasts between dielectrics and metals
Polarised lighting

polarizer/\text{analyzer} configurations

\text{law of Malus}:

\[ I(\theta) = I(0) \cos^2 \theta \]
Polarized lighting

2 uses:

1. to improve contrast between Lambertian and specular reflections

2. to improve contrasts between dielectrics and metals
Polarized lighting

specular reflection keeps polarisation: diffuse reflection depolarises

suppression of specular reflection:

polarizer/analyzer crossed prevents the large dynamic range caused by glare
Example pol. lighting (pol./an.crossed)
Polarized lighting

2 uses:

1. to improve contrast between Lambertian and specular reflections

2. to improve contrasts between dielectrics and metals
Reflection : dielectric

Polarizer at *Brewster angle*
strong reflectors
more or less preserve polarization
Polarised lighting

distinction between specular reflection from dielectrics and metals;
works under the Brewster angle for the dielectric
dielectric has no parallel comp.; metal does
suppression of specular reflection from dielectrics:

polarizer/analyzer aligned
distinguished metals and dielectrics
Example pol. lighting (pol./an. aligned)
Coloured lighting

highlight regions of a similar colour

with band-pass filter: only light from projected pattern (e.g. monochromatic light from a laser)

differentiation between specular and diffuse reflection

comparing colours $\Rightarrow$ same spectral composition of sources!

spectral sensitivity function of the sensors!
Example coloured lighting
Structured and stroboscopic lighting

spatially or temporally modulated light pattern

**Structured lighting**

e.g. : 3D shape : objects distort the projected pattern  
(more on this later)

**Stroboscopic lighting**

high intensity light flash

to eliminate motion blur
Stroboscopic lighting
App: vegetable inspection (colored light + polarization)
cameras
Optics for image formation

the pinhole model:
Optics for image formation

the pinhole model:

hence the name: CAMERAr
obscura
Optics for image formation

the pinhole model:

\[ \frac{X_i}{X_o} = \frac{Y_i}{Y_o} = \frac{f}{-Z_o} = -m \]

\((m = \text{linear magnification})\)
Camera obscura + lens
The thin-lens equation

lens to capture enough light:

\[ \frac{1}{Z_O} - \frac{1}{Z_i} = \frac{1}{f} \]

assuming
- spherical lens surfaces
- incoming light ± parallel to axis
- thickness << radii
- same refractive index on both sides
The depth-of-field

Only reasonable sharpness in Z-interval

\[ \Delta Z_0^- = Z_0 - Z_0^- = \frac{Z_0(Z_0 - f)}{Z_0 + f \frac{d}{b - f}} \]

decreases with \( d \), increases with \( Z_0 \)

strike a balance between incoming light (\( d \)) and
large depth-of-field (usable depth range)
The depth-of-field

\[ \Delta Z^-_0 = Z_0 - Z^-_0 = \frac{Z_0(Z_0 - f)}{Z_0 + f \frac{d}{b} - f} \]

Similar expression for \( Z^+_O - Z_O \)
The depth-of-field

\[ \Delta Z_0^- = Z_0 - Z_0^- = \frac{Z_0 (Z_0 - f)}{Z_0 + f \frac{d}{b - f}} \]

**Ex 1:** microscopes -> small DoF

**Ex 2:** special effects -> flood miniature scene with light
Deviations from the lens model

3 assumptions:

1. all rays from a point are focused onto 1 image point

2. all image points in a single plane

3. magnification is constant

deviations from this ideal are *aberrations*
Aberrations

2 types:

1. geometrical

2. chromatic

*geometrical*: small for paraxial rays

*chromatic*: refractive index function of wavelength (Snell’s law !!)
Geometrical aberrations

- spherical aberration
- astigmatism
- radial distortion
- coma

The most important type
Spherical aberration

rays parallel to the axis do not converge

outer portions of the lens yield smaller focal lengths
Radial Distortion

magnification different for different angles of inclination

barrel  none  pincushion
Radial Distortion

magnification different for different angles of inclination

barrel  none  pincushion

The result is pixels moving along lines through the center of the distortion – typically close to the image center – over a distance $d$, depending on the pixels’ distance $r$ to the center

$$d = (1 + \kappa_1 r^2 + \kappa_2 r^4 + \ldots)$$
Radial Distortion

magnification different for different angles of inclination

This aberration type can be corrected by software if the parameters \(( \kappa_1, \kappa_2, \ldots )\) are known.
Radial Distortion

magnification different for different angles of inclination

Some methods do this by looking how straight lines curve instead of being straight
Chromatic aberration

Rays of different wavelengths focused in different planes cannot be removed completely but **achromatization** can be achieved at some well chosen wavelength pair, by combining lenses made of different glasses.

Sometimes **achromatization** is achieved for more than 2 wavelengths.
the figure shows wavelengths that materials let pass

additional considerations:
humidity and temperature resistance, weight, price,...
Cameras

we consider 2 types:

1. CCD
2. CMOS
Cameras

CCD = Charge-coupled device
CMOS = Complementary Metal Oxide Semiconductor
CCD

separate photo sensor at regular positions
no scanning

charge-coupled devices (CCDs)

area CCDs and linear CCDs

2 area architectures:

interline transfer and frame transfer

- photosensitive
- storage
The CCD (inter-line) camera
CMOS

Same sensor elements as CCD
Each photo sensor has its own amplifier
  More noise (reduced by subtracting ‘black’ image)
  Lower sensitivity (lower fill rate)
Uses standard CMOS technology
  Allows to put other components on chip
  ‘Smart’ pixels
Resolution trend in mobile phones

*Volume and revenue opportunity for high resolution sensors*

Source: TSR, CCD/CMOS Area Image Sensor Market Analysis, dated June 2011
CCD vs. CMOS

- Niche applications
- Specific technology
- High production cost
- High power consumption
- Higher fill rate
- Blooming
- Sequential readout

- Consumer cameras
- Standard IC technology
- Cheap
- Low power
- Less sensitive
- Per pixel amplification
- Random pixel access
- Smart pixels
- On chip integration with other components

2006 was year of sales cross-over
CCD vs. CMOS

- Niche applications
- Specific technology
- High production cost
- High power consumption
- Higher fill rate
- Blooming
- Sequential readout

- Consumer cameras
- Standard IC technology
- Cheap
- Low power
- Less sensitive
- Per pixel amplification
- Random pixel access
- Smart pixels
- On chip integration with other components

In 2015 Sony said to stop CCD chip production.
Colour cameras

We consider 3 concepts:

1. Prism (with 3 sensors)
2. Filter mosaic
3. Filter wheel
Prism colour camera

Separate light in 3 beams using dichroic prism
Requires 3 sensors & precise alignment
Good color separation
Prism colour camera
Filter mosaic

Coat filter directly on sensor

Demosaicng (obtain full colour & full resolution image)
Color filters lower the effective resolution, hence **microlenses** often added to gain more light on the small pixels.
Filter wheel

Rotate multiple filters in front of lens
Allows more than 3 colour bands

Only suitable for static scenes
# Computer Vision

## Prism vs. mosaic vs. wheel

<table>
<thead>
<tr>
<th></th>
<th>Prism</th>
<th>Mosaic</th>
<th>Wheel</th>
</tr>
</thead>
<tbody>
<tr>
<td>approach</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td># sensors</td>
<td>High</td>
<td>Average</td>
<td>Good</td>
</tr>
<tr>
<td>Resolution</td>
<td>High</td>
<td>Low</td>
<td>Average</td>
</tr>
<tr>
<td>Cost</td>
<td>High</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Framerate</td>
<td>High</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Artefacts</td>
<td>Low</td>
<td>Aliasing</td>
<td>Motion</td>
</tr>
<tr>
<td>Bands</td>
<td>3</td>
<td>3</td>
<td>3 or more</td>
</tr>
</tbody>
</table>

- High-end cameras
- Low-end cameras
- Scientific applications
Geometric camera model

perspective projection

(Man Drawing a Lute, woodcut, 1525, Albrecht Dürer)
Models for camera projection

the pinhole model revisited:

center of the lens = center of projection

notice the virtual image plane

this is called *perspective* projection
Models for camera projection

We had the virtual plane also in the original reference sketch:
Perspective projection

- origin lies at the center of projection
- the $Z_c$ axis coincides with the optical axis
- $X_c$-axis $||$ to image rows, $Y_c$-axis $||$ to columns
Perspective projection

\[ u = f \frac{X}{Z} \quad \text{and} \quad \ v = f \frac{Y}{Z} \]
Pseudo-orthographic projection

\[ u = f \frac{X}{Z} \quad \quad v = f \frac{Y}{Z} \]

If \( Z \) is constant \( \Rightarrow x = kX \) and \( y = kY \),
where \( k = \frac{f}{Z} \)

i.e. *orthographic* projection + a scaling

Good approximation if \( f/Z \pm \) constant, i.e. if objects are small compared to their distance from the camera
Pictoral comparison

Pseudo-orthographic

Perspective
Projection matrices

the perspective projection model is incomplete:
what if:

1. 3D coordinates are specified in a
   world coordinate frame

2. Image coordinates are expressed as
   row and column numbers

We will not consider additional refinements, such as radial distortions,...
Projection matrices

\[ u = f \frac{\langle r_1, P - C \rangle}{\langle r_3, P - C \rangle} \]
\[ v = f \frac{\langle r_2, P - C \rangle}{\langle r_3, P - C \rangle} \]

\[ u = f \frac{r_{11}(X - C_1) + r_{12}(Y - C_2) + r_{13}(Z - C_3)}{r_{31}(X - C_1) + r_{32}(Y - C_2) + r_{33}(Z - C_3)} \]
\[ v = f \frac{r_{21}(X - C_1) + r_{22}(Y - C_2) + r_{23}(Z - C_3)}{r_{31}(X - C_1) + r_{32}(Y - C_2) + r_{33}(Z - C_3)} \]
Projection matrices

Image coordinates are to be expressed as *pixel coordinates*

\[
\begin{align*}
    x &= k_x u + s v + x_0 \\
    y &= k_y v + y_0
\end{align*}
\]

with:

→ \((x_0, y_0)\) the pixel coordinates of the principal point

→ \(k_x\) the number of pixels per unit length horizontally

→ \(k_y\) the number of pixels per unit length vertically

→ \(s\) indicates the skew; typically \(s = 0\)
Projection matrices

Image coordinates are to be expressed as *pixel coordinates*

\[
\begin{align*}
    x &= k_x u + s v + x_0 \\
    y &= k_y v + y_0
\end{align*}
\]

with:

**NB1:** often only integer pixel coordinates matter
Projection matrices

Image coordinates are to be expressed as *pixel coordinates*

\[
\begin{align*}
x &= k_x u + s v + x_0 \\
y &= k_y v + y_0
\end{align*}
\]

with:

**NB2**: \(k_y/k_x\) is called the *aspect ratio*
Projection matrices

Image coordinates are to be expressed as *pixel coordinates*

\[
\begin{align*}
x &= k_x u + s v + x_0 \\
y &= k_y v + y_0
\end{align*}
\]

with:

**NB3**: $k_x, k_y, s, x_0$ and $y_0$ are called *internal camera parameters*
Projection matrices

Image coordinates are to be expressed as **pixel coordinates**

\[
\begin{align*}
x &= k_x u + s v + x_0 \\
y &= k_y v + y_0
\end{align*}
\]

with:

**NB4**: when they are known, the camera is **internally calibrated**
Projection matrices

Image coordinates are to be expressed as **pixel coordinates**

\[
x = k_x u + s v + x_0 \\
y = k_y v + y_0
\]

with:

**NB1:** often only integer pixel coordinates matter

**NB2:** \(k \) \(x/y\) is called the aspect ratio

**NB3:** \(k_x, k_y, s, x_0, y_0\) are called internal camera parameters

**NB4:** when they are known, the camera is internally calibrated

**NB5:** vector \(C\) and matrix \(R \in SO(3)\) are the external camera parameters
Projection matrices

Image coordinates are to be expressed as *pixel coordinates*

\[
\begin{aligned}
x &= k_x u + s v + x_0 \\
y &= k_y v + y_0
\end{aligned}
\]

with:

**NB6**: when these are known, the camera is *externally calibrated*
Projection matrices

Image coordinates are to be expressed as \textit{pixel coordinates}

\[
\begin{align*}
x &= k_x u + s v + x_0 \\
y &= k_y v + y_0
\end{align*}
\]

with:

\textbf{NB7 :} \textit{fully calibrated} means internally and externally calibrated
Homogeneous coordinates

Often used to linearize non-linear relations

2D
\[
\begin{pmatrix}
x \\ y \\ z
\end{pmatrix}
\rightarrow
\begin{pmatrix}
x/z \\ y/z
\end{pmatrix}
\]

3D
\[
\begin{pmatrix}
X \\ Y \\ Z \\ W
\end{pmatrix}
\rightarrow
\begin{pmatrix}
X/W \\ Y/W \\ Z/W
\end{pmatrix}
\]

Homogeneous coordinates are only defined up to a factor.
Projection matrices

\[
\begin{align*}
    u &= f \frac{r_{11}(X - C_1) + r_{12}(Y - C_2) + r_{13}(Z - C_3)}{r_{31}(X - C_1) + r_{32}(Y - C_2) + r_{33}(Z - C_3)} \\
    v &= f \frac{r_{21}(X - C_1) + r_{22}(Y - C_2) + r_{23}(Z - C_3)}{r_{31}(X - C_1) + r_{32}(Y - C_2) + r_{33}(Z - C_3)}
\end{align*}
\]

Exploiting homogeneous coordinates:

\[
\begin{pmatrix}
    u \\
    v \\
    1
\end{pmatrix}
= \begin{pmatrix}
    f & r_{11} & f & r_{12} & f & r_{13} \\
    f & r_{21} & f & r_{22} & f & r_{23} \\
    r_{31} & r_{32} & r_{33}
\end{pmatrix}
\begin{pmatrix}
    X - C_1 \\
    Y - C_2 \\
    Z - C_3
\end{pmatrix}
\]
Projection matrices

\[
\begin{cases}
x = k_x u + s v + x_0 \\
y = k_y v + y_0
\end{cases}
\]

Exploiting homogeneous coordinates:

\[
\begin{pmatrix}
x \\ y \\ 1
\end{pmatrix} =
\begin{pmatrix}
k_x & s & x_0 \\ 0 & k_y & y_0 \\ 0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
u \\ v \\ 1
\end{pmatrix}
\]
Projection matrices

Thus, we have:

\[
\begin{pmatrix}
\tau \\
u \\
v \\
1
\end{pmatrix}
= 
\begin{pmatrix}
f \r_{11} \\
f \r_{12} \\
f \r_{13} \\
r_{31} \\
r_{32} \\
r_{33}
\end{pmatrix}
\begin{pmatrix}
X - C_1 \\
Y - C_2 \\
Z - C_3
\end{pmatrix}
\]

\[
\begin{pmatrix}
x \\
y \\
1
\end{pmatrix}
= 
\begin{pmatrix}
\begin{pmatrix}
k_x \\
s \\
x_0
\end{pmatrix} \\
0 \\
0 \\
0 \\
0 \\
1
\end{pmatrix}
\begin{pmatrix}
\tau \\
u \\
v \\
1
\end{pmatrix}
\]
Projection matrices

Concatenating the results:

\[
\tau \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} k_x & s & x_0 \\ 0 & k_y & y_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} f & r_{11} & f & r_{12} & f & r_{13} \\ f & r_{21} & f & r_{22} & f & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \begin{pmatrix} X - C_1 \\ Y - C_2 \\ Z - C_3 \end{pmatrix}
\]

Or, equivalently:

\[
\tau \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} k_x & s & x_0 \\ 0 & k_y & y_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \begin{pmatrix} X - C_1 \\ Y - C_2 \\ Z - C_3 \end{pmatrix}
\]
Projection matrices

Re-combining matrices in the concatenation:

\[
\begin{pmatrix}
  x \\
  y \\
  1
\end{pmatrix} =
\begin{pmatrix}
  k_x & s & x_0 \\
  0 & k_y & y_0 \\
  0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
  f & 0 & 0 \\
  0 & f & 0 \\
  0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
  r_{11} & r_{12} & r_{13} \\
  r_{21} & r_{22} & r_{23} \\
  r_{31} & r_{32} & r_{33}
\end{pmatrix}
\begin{pmatrix}
  X - C_1 \\
  Y - C_2 \\
  Z - C_3
\end{pmatrix}
\]

yields the calibration matrix \( K \):

\[
K =
\begin{pmatrix}
  k_x & s & x_0 \\
  0 & k_y & y_0 \\
  0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
  f & 0 & 0 \\
  0 & f & 0 \\
  0 & 0 & 1
\end{pmatrix} =
\begin{pmatrix}
  f & k_x & f & s & x_0 \\
  0 & f & k_y & y_0 \\
  0 & 0 & 0 & 1
\end{pmatrix}
\]
We define

\[
\begin{pmatrix}
x \\
y \\ 1
\end{pmatrix}
, \quad
\begin{pmatrix}
X \\
Y \\
Z
\end{pmatrix}
, \quad
\begin{pmatrix}
X \\
Y \\
Z \\
1
\end{pmatrix}
\]

yielding

\[
\rho p = KR^t(P - C)
\]
for some non-zero \(\rho \in \mathbb{R}\)

or,

\[
\rho p = K\begin{pmatrix}
R^t \\
-R^tC
\end{pmatrix}\tilde{P}
\]

or,

\[
\rho p = (M \mid t)\tilde{P} \quad \text{with} \quad \text{rank } M = 3
From object radiance to pixel grey levels

After the geometric camera model... a photometric camera model

2 steps:

1. from object radiance to image irradiance
2. from image irradiance to pixel grey level
we look at the irradiance that an object patch will cause in the image

assumptions:
radiance $R$ assumed known and
object at large distance compared to the focal length

Is image irradiance directly related to the radiance of the image patch?
The viewing conditions

\[ I = R \frac{A_I}{f^2} \cos^4 \alpha \]

the cos^4 law
The $\cos^4$ law cont’d

Especially strong effects for wide-angle and fisheye lenses
From irradiance to gray levels

\[ f = g I^\gamma + d \]

Gain

“gamma”

Dark reference
From irradiance to gray levels

\[ f = g I^\gamma + d \]

- set w. size diaphragm close to 1 nowadays
- signal w. cam cap on
- Gain
- “gamma”
- Dark reference