FACULTEIT INGENIEURSWETENSCHAPPEN
EN ARCHITECTUUR

Computervisie

Bert De Saffel

Master in de Industri Wetenschappen: Informatica Academiejaar 20182019

Gecompileerd op 14 maart 2019

UNIVERSITEIT
GENT

Inhoudsopgave

Hoofdstuk 1

Inleiding

- Verschil tussen computervisie en computergrafiek:
  - Computervisie is de analyse van een beeld of video om zo de structuur en semantiek te achterhalen.
  - Computergrafiek bouwt aan de hand van structuurregels beelden op.
- Verschil tussen detectie en herkenning:
  - Detectie heeft als doel om te achterhalen of een object aanwezig is.
  - Herkenning heeft als doel om dit object te identificeren.

Hoofdstuk 2

Beeldvorming

2.1 Punten

- Elk 2D punt (x, y) kan gerepresenteerd worden via de **homogene coördinaten** $(\lambda x, \lambda y, \lambda)$ voor $\lambda \neq 0$. Homogene coördinaten worden ook **projectieve coördinaten** genoemd, en bevinden zich in het projectieve vlak $\mathcal{P}^2$.
- Het cartesisch punt $(1, 2)$ kan gepresenteerd worden in homogene coördinaten als $(1, 2, 1)$ of zelfs $(2, 4, 2)$.
- Het origineel cartesisch punt kan bekomen worden te delen door λ:

$$(\lambda x, \lambda y, \lambda) = (\frac{\lambda x}{\lambda}, \frac{\lambda y}{\lambda}, \frac{\lambda}{\lambda}) = (x, y)$$

- Een punt kan door oneindig veel homogene coördinaten gerepresenteerd worden.
- Dit kan eenvoudig uitgebreid worden in drie dimensies:

$$(x, y, z, \lambda)$$

2.2 Lijnen

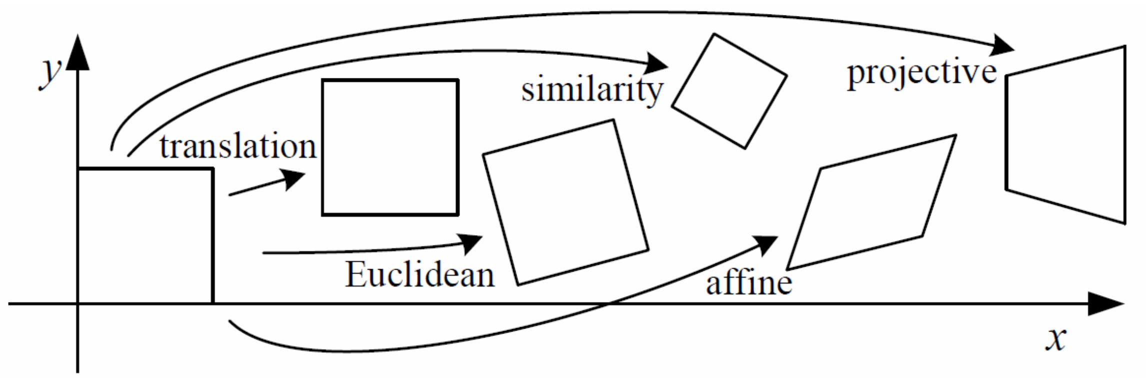
- Een lineaire vergelijking in $\mathcal{R}^2$ kent een aantal problemen:
  - De vergelijking $y = ax + b$ kan geen verticale lijnen voorstellen.
  - De vergelijking $x = ay + b$ kan geen horizontale lijnen voorstellen.
  - De vergelijking $ax + by = 1$ kan geen lijnen door de oorsprong voorstellen.
  - ...
- Een lineaire vergelijking in $\mathcal{P}^2$ heeft de volgende vorm:

$$ax + by + cz = 0$$

  Deze vergelijking kan alle mogelijke lijnstukken voorstellen.
- Deze veralgemening lost ook een aantal geometrische problemen op. In $\mathcal{R}^2$ kunnen twee lijnen ofwel elkaar snijden, ofwel nooit snijden. In $\mathcal{P}^2$ snijden twee lijnstukken altijd, maar kan eventueel in oneindig zijn.

2.3 Twee-dimensionale transformaties

Figuur 2.1 toont alle transformaties mogelijk in de twee-dimensionale ruimte. Elke transformatie



Figuur 2.1: Eenvoudige twee-dimensionale transformaties.

maakt gebruik van één of andere matrix.

2.3.1 Similariteit Transformatie

Deze transformatie combineert eigenlijk de rotatie, translatie en schaling met:

- een schalingsfactor s,
- de rotatiehoek θ,
- en deranslatie t_x en t_y.

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} s\cos\theta & -s\sin\theta & t_x \\ s\sin\theta & s\cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

2.3.2 Affiene transformatie

Deze transformatie laat toe om een figuur scheef te maken terwijl parallellisme bewaart blijft.

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} a & b & t_x \\ c & d & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

2.3.3 Projectieve transformatie

$$\begin{pmatrix} kx' \\ ky' \\ k \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

Transformatie	Matrix	Degrees of Freedom	Behoudt
Translatie	$[I\|t]_{3\times4}$	3	Oriëntatie
Rotatie	$[I\|t]_{3\times4}$	6	Lengte
Similariteit	$[sR\|t]_{3\times4}$	6	Hoeken
Affien	$[A]_{3\times4}$	12	Parallellisme
Projectie	$[\widetilde{H}]_{4\times4}$	15	Rechte lijnen

Tabel 2.1: Hiërarchie van drie-dimensionale transformaties.

2.4 Drie-dimensionale transformaties

Tabel 2.1 toont de hiërarchie van drie-dimensionale transformaties.

2.4.1 Drie-dimensionale rotatie

- Elke rotatie in drie dimensies is rond een as $\hat{\mathbf{n}}$
- Een verzameling van rotaties rond verschillende assen kan vervangen worden door één rotatie rond één as.
- Rotaties in drie dimensies is <u>niet commutatief</u>.
- Stel:

$$[\hat{\mathbf{n}}]_\times = \begin{bmatrix} 0 & -\hat{n}_z & \hat{n}_y \\ \hat{n}_z & 0 & -\hat{n}_x \\ -\hat{n}_y & \hat{n}_x & 0 \end{bmatrix}$$

  dan kan elk kruisproduct $\hat{\mathbf{n}} \times \mathbf{v}$ geschreven worden als

$$\hat{\mathbf{n}} \times \mathbf{v} = [\hat{\mathbf{n}}]_\times \mathbf{v}$$

- De rotatiematrix rond een as $\hat{\mathbf{n}}$ en hoek *theta*:

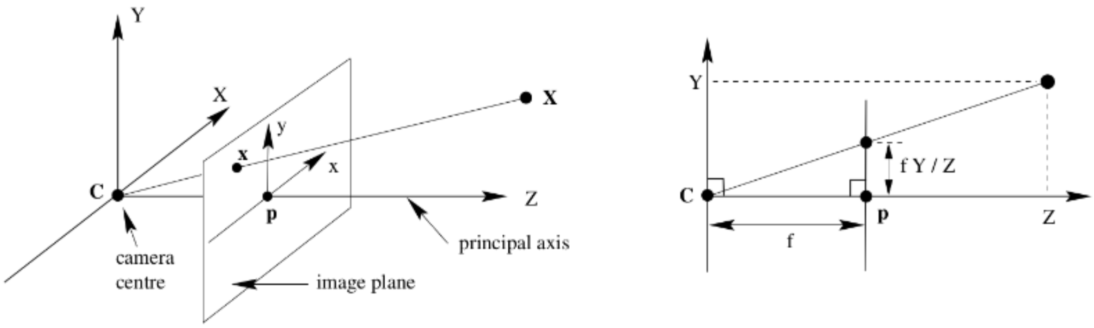
$$R(\hat{\mathbf{n}}, \theta) = \mathbf{I} + \sin\theta[\hat{\mathbf{n}}]_\times + (1 - \cos\theta)[\hat{\mathbf{n}}]_\times^2$$

  Dit staat bekend als de **formule van Rodrigues**.
- Elke beweging van een camera kan geschreven worden als de combinatie van een rotatie met rotatiematrix $\mathbf{R}$ en een translatie met translatiematrix T:

$$\begin{bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix}$$

2.5 Pinhole model

- Er wordt een oogpunt C verondersteld.
- Op een afstand f bevindt er zicht een projectievlak p. Op een lengte Z en hoogte Y bevindt er zich een punt in drie dimensies (analoog voor het xz-vlak, maar dan met lengte Z en hoogte X).
- Dit driedimensionaal punt kan op het projectievlak geprojecteerd worden door $y = f\frac{Y}{Z}$ en $x = f\frac{X}{Z}$.



Figuur 2.2: Het pinhole model, uitgewerkt voor de yz-vlak.

- Als we nu rekening houden met homogene coördinaten:

$$Z\mathbf{x} = Z \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = K_f \mathbf{X}$$

De deling door Z verdwijnt en de projectie is nu een lineaire transformatie.

2.6 Projectieve transformaties

-

Hoofdstuk 3

ROC curves

3.1 Inleiding

- ROC = Receiver Operating Characteristic.
- Een eenvoudige manier om classifiers te evalueren.

3.2 Classifiers

- Het toekennen van een klasse uit een verzameling van voorgedefinieerde klassen aan een object.
- Een **Binaire classifier** kent slechts twee klassen.

3.3 Binaire classifiers

- Stel twee klassen α en β. We zijn nu geïnteresseerd of een object in klasse α zit.
- Een **False Positive** komt voor wanneer het object de klasse α krijgt, terwijl hij β is.
- Een **False Negative** komt voor wanneer het object de klasse β krijgt, terwijl hij α is.
- Binaire **confusion matrix**:

Klasse	Voorspelde klasse	
	positief	negatief
positief (#P)	#TP	#FN
positief (#N)	#FP	#TN

Klasse	Voorspelde klasse	
	positief	negatief
positief (#P)	#TP	#P - #TP
positief (#N)	#FP	#N - #FP

- De **true positive rate (TPR)** is:

$$TPR = \frac{\#TP}{\#P}$$

- De **false positive rate (FPR)** is:

$$FPR = \frac{\#FP}{\#N}$$