

Intuïtieve mens-machineinterface met live actieherkenning

Master of Science in de industriële wetenschappen: informatica

Bert De Saffel

prof. dr. ir. Peter Veelaert & prof. dr. ir. Wilfried Philips
ing. Sanne Roegiers & ing. Dimitri van Cauwelaert

04 april 2019

Inhoudsopgave

- ① Context
- ② Probleemstellingen
- ③ Methodologie

Inhoudsopgave

1 Context

2 Probleemstellingen

3 Methodologie

- Machine Learning
- Dataset
- Preprocessing
- Classificatie
- Evaluatie

Context



Context

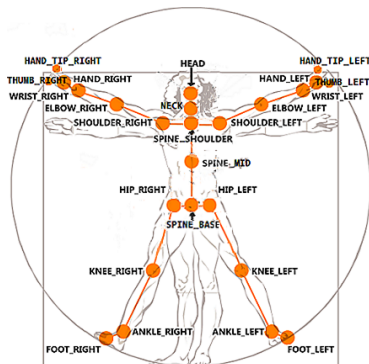
- Hoe besturen?
 - Remote control
 - Autonoom
 - Actieherkenning

Context

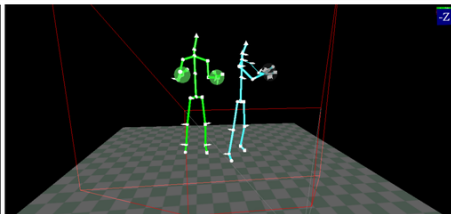
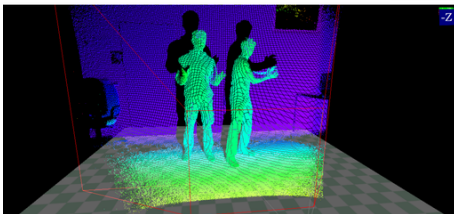
- De verplaatsing van een “robot” uitvoeren met enkel actieherkenning

Context

- De verplaatsing van een “robot” uitvoeren met enkel actieherkenning
- Met de kinect sensor
 - Kan skeletbeelden genereren vanuit RGB-D data



Kinect Sensor



Inhoudsopgave

1 Context

2 Probleemstellingen

3 Methodologie

- Machine Learning
- Dataset
- Preprocessing
- Classificatie
- Evaluatie

Probleemstellingen

- 1 Verschillen in lichaamsbouw mogelijk (klein vs groot)
- 2 Verschillen in camerahoek
- 3 Verschillen in locatie

Probleemstellingen

- ① Verschillen in lichaamsbouw mogelijk (klein vs groot)
- ② Verschillen in camerahoek
- ③ Verschillen in locatie
- ④ Real-time actieherkenning
 - De actie herkennen op het moment dat deze uitgevoerd wordt
 - \approx binnen 5 frames

Onderzoek

- 1 De features moeten rotatie-, lichaams- en plaatsinvariant zijn
- 2 Actie moet vroeg genoeg herkend worden om live te kunnen classificeren

Inhoudsopgave

1 Context

2 Probleemstellingen

3 Methodologie

- Machine Learning
- Dataset
- Preprocessing
- Classificatie
- Evaluatie

Inhoudsopgave

- 1 Context
- 2 Probleemstellingen
- 3 Methodologie
 - Machine Learning
 - Dataset
 - Preprocessing
 - Classificatie
 - Evaluatie

Machine Learning - Classificatieprobleem

- Een verzameling van klassen (labels, uitvoerwaarden, ...)
- **Gegeven een observatie, tot welke klasse behoort ze?**
- Bij actieherkenning:
 - Klassen = acties
 - Observaties = frames

Machine Learning - Features

- Een observatie wordt getransformeerd naar *features*
 - Pixel: RGB-waarden
 - Persoon: leeftijd, geslacht, haarkleur, lengte, ...
- Features op basis van skeletbeelden
 - Elk skelet *joint* wordt gekenmerkt door zijn (x, y, z) coördinaten en (a, b, c, d) quaternionen.
 - Quaternion:

$$\mathbf{q} = a + bi + cj + dk$$

- Het skelet bestaat uit 25 *joints*
- 175-dimensionale *feature vector*

$$\mathbf{f} = (x_1 \quad y_1 \quad \dots \quad y_{25} \quad z_{25} \quad a_1 \quad b_1 \quad \dots \quad c_{25} \quad d_{25})$$

Machine Learning - Classificatie

- De *feature vector* kan als input dienen voor een classifier
- Welke classifier is de beste? → Evalueren a.d.h.v. een *confusion matrix*:

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

- Precision = $\frac{TP}{TP+FP}$
- Recall = $\frac{TP}{TP+FN}$
- F1 score = $2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}}$

Inhoudsopgave

- 1 Context
- 2 Probleemstellingen
- 3 Methodologie
 - Machine Learning
 - **Dataset**
 - Preprocessing
 - Classificatie
 - Evaluatie

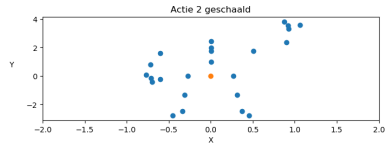
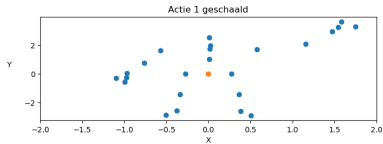
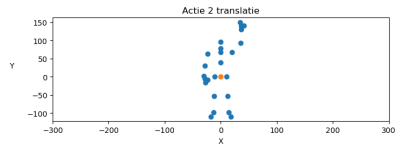
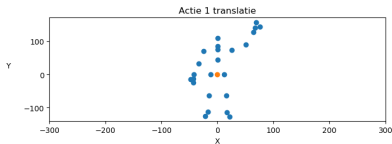
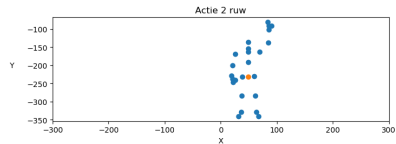
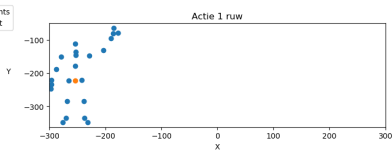
Dataset

- 9 acties, 4 personen
- Opgenomen met de Kinect
- *Leave-One-Out Cross Validation*
 - Trainingset: $n - 1$ personen
 - Testset: 1 persoon
 - Evalueer classifieer met deze twee sets
 - Herhaal n keer zodat elke persoon één keer in de testset zit
 - De score van de classifieer is het gemiddelde van de scores

Inhoudsopgave

- 1 Context
- 2 Probleemstellingen
- 3 Methodologie
 - Machine Learning
 - Dataset
 - Preprocessing
 - Classificatie
 - Evaluatie

Preprocessing



Preprocessing

1. Plaats-invariantie → Translatie

- *Spine base joint* als oorsprong:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$$

met x_0, y_0, z_0 de drie-dimensionale coördinaten van de *Spine base joint*

Preprocessing

1. Plaats-invariantie → Translatie

- *Spine base joint* als oorsprong:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$$

met x_0, y_0, z_0 de drie-dimensionale coördinaten van de *Spine base joint*

2. Schaal-invariantie → Vectorsnormalisatie

- Elk component van elke positievector delen door lengte van de *neck joint* positievector:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \frac{x}{\|n\|} \\ \frac{y}{\|n\|} \\ \frac{z}{\|n\|} \end{pmatrix}$$

met

$$\|n\| = \sqrt{(neck_x)^2 + (neck_y)^2 + (neck_z)^2}$$

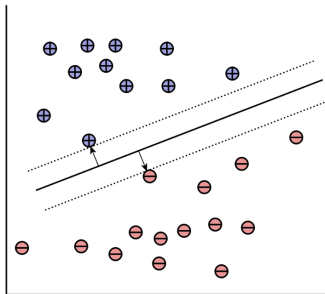
3. Rotatie-invariantie → Lokaal skeletcoördinatensysteem (X' , Y' , Z')
- X' -as = de as door de *Right Hip joint (RH)* en *Left Hip joint (LH)*
 - Y' -as = de as door de *Spine Base joint (SB)* en *Spine Mid joint (SM)*
 - Z' -as = orthogonaal met X' en Y'

Inhoudsopgave

- 1 Context
- 2 Probleemstellingen
- 3 Methodologie
 - Machine Learning
 - Dataset
 - Preprocessing
 - **Classificatie**
 - Evaluatie

Classificatie

- Enerzijds vergelijken verschillende classifiers
 - *K-Nearest Neighbors*
 - Vergelijk de k dichtste feature vectoren, de gelabelde klasse is diegene die het meest voorkomt
 - *Support Vector Machine*
 - Zoek een hypervlak dat de positieve observaties van de negatieve onderscheidt.



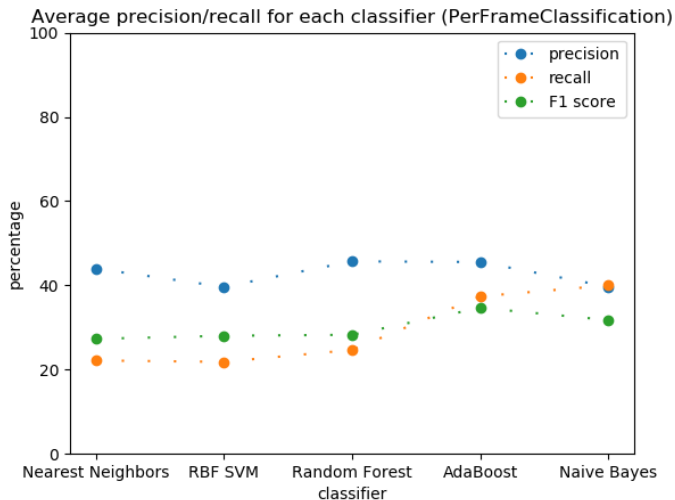
Classificatie

- Anderzijds verschillende classificatiestrategieën toepassen
 - Frame per frame classificeren zonder temporaal aspect
 - Buffer bijhouden van 30 opeenvolgende frames, met *majority voting* de actie bepalen
 - Buffer bijhouden van 30 opeenvolgende frames, met een gewogen vote de actie bepalen
 - Gewichten op actie plaatsen
 - Stopactie heeft voorrang

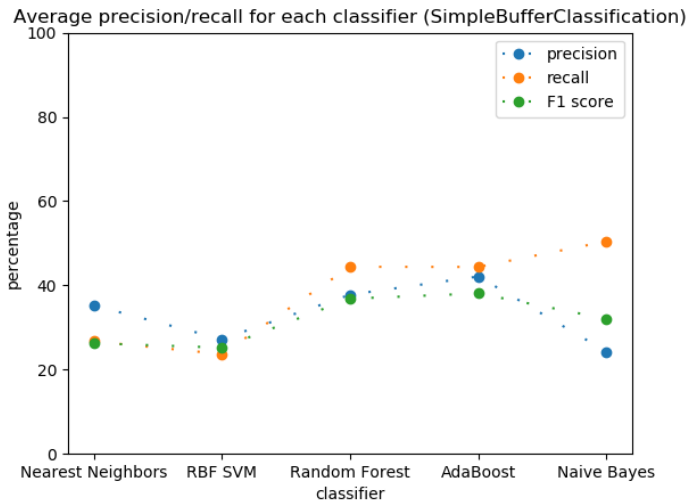
Inhoudsopgave

- 1 Context
- 2 Probleemstellingen
- 3 Methodologie**
 - Machine Learning
 - Dataset
 - Preprocessing
 - Classificatie
 - Evaluatie**

Evaluatie

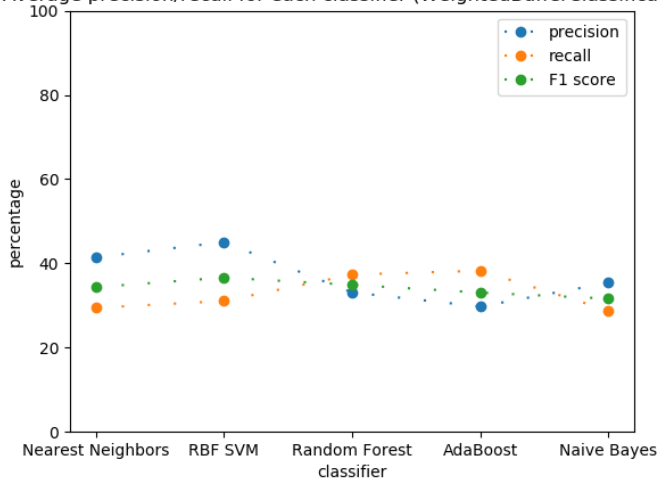


Evaluatie



Evaluatie

Average precision/recall for each classifier (WeightedBufferClassification)



Vragen, opmerkingen, ...?