

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
- Feature transformatie
- Classificatie
- Evaluatie

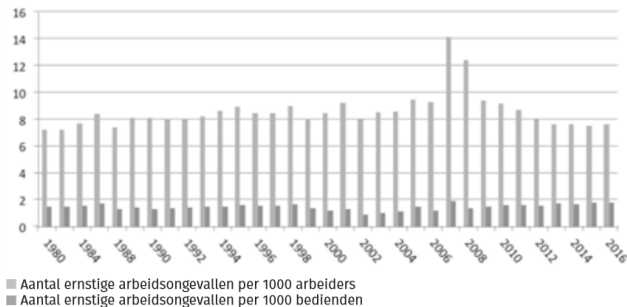
Context

- Oorzaken van ernstige arbeidsongevallen in 2015
 - 1 Verlies van controle over een machine of voertuig
 - 2 Uitglijden of struikelen
 - 3 Het tillen of neerzetten van lasten
 - 4 Vrijkomen van giftige producten

Context

- Oorzaken van ernstige arbeidsongevallen in 2015
 - ① Verlies van controle over een machine of voertuig
 - ② Uitglijden of struikelen
 - ③ Het tillen of neerzetten van lasten
 - ④ Vrijkomen van giftige producten
- Gevolgen
 - Langdurige ongeschiktheid
 - Permanente letsels
 - Sterfgeval

Context



Figuur: Frequentiegraad ernstige arbeidsongevallen in de privésector.

Context

- Mogelijke oplossing
 - Het inzetten van robotica in gevaarlijke omgevingen

Context

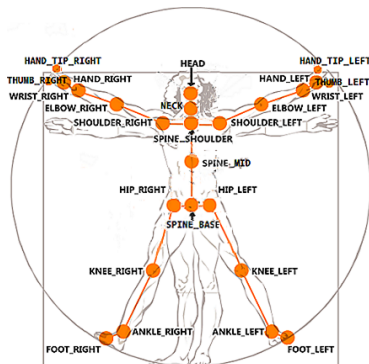
- Mogelijke oplossing
 - Het inzetten van robotica in gevaarlijke omgevingen
 - Hoe besturen?
 - Remote control
 - Autonoom
 - Actieherkenning

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



Inhoudsopgave

1 Context

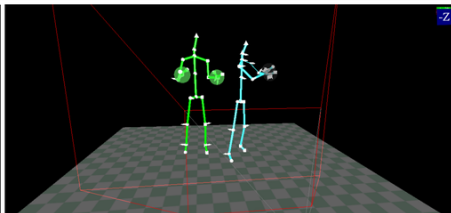
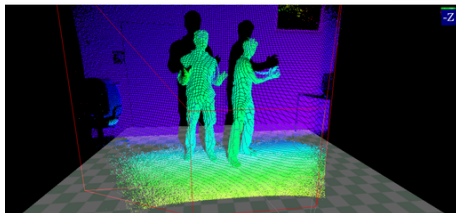
2 Probleemstellingen

3 Methodologie

- Machine Learning
- Dataset
- Preprocessing
- Feature transformatie
- Classificatie
- Evaluatie

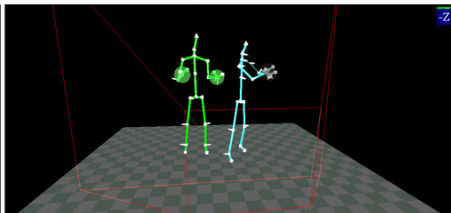
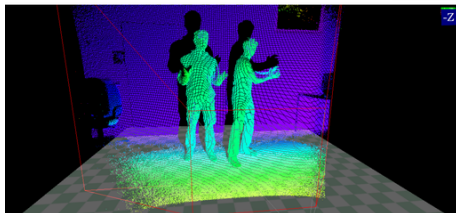
Probleemstellingen

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



Probleemstellingen

- ① Verschillen in lichaamsbouw mogelijk (klein vs groot)
- ② Verschillen in camerahoek
- ③ Real-time actieherkenning
 - De actie herkennen op het moment dat deze uitgevoerd wordt



Onderzoek

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

Inhoudsopgave

1 Context

2 Probleemstellingen

3 Methodologie

- Machine Learning
- Dataset
- Preprocessing
- Feature transformatie
- Classificatie
- Evaluatie

Inhoudsopgave

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

Machine Learning - Classificatieprobleem

- Een verzameling van klassen (labels, uitvoerwaarden, ...)
- **Gegeven een observatie, tot welke klasse behoort deze observatie?**
- 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:

$$\mathbf{q} = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$$

- 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 eender welke 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
 - Feature transformatie
 - Classificatie
 - Evaluatie

Dataset

- Onderzoek naar intuïtieve handelingen
- ≈ 30 FPS

Inhoudsopgave

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

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:

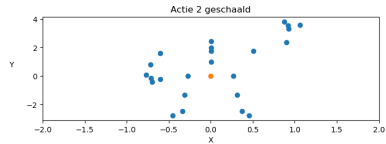
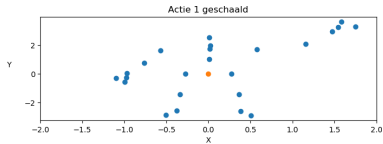
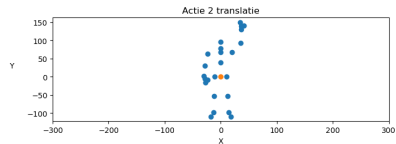
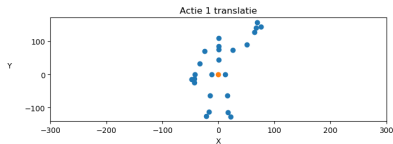
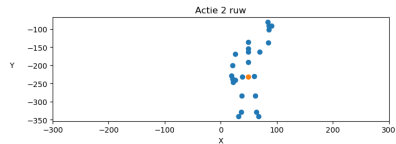
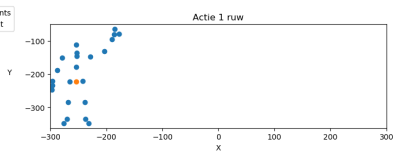
$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \frac{x}{\|n\|} \\ \frac{y}{\|n\|} \\ \frac{z}{\|n\|} \end{pmatrix}$$

met

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

3. Rotatie-invariantie \rightarrow 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'

Preprocessing



Inhoudsopgave

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

Feature transformatie

- Momenteel: *Features* zijn invariant t.o.v. locatie, schaal en camerahoek

Inhoudsopgave

1 Context

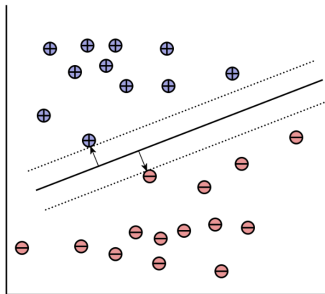
2 Probleemstellingen

3 Methodologie

- Machine Learning
- Dataset
- Preprocessing
- Feature transformatie
- **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 frames, met *majority voting* de actie bepalen
 - Buffer bijhouden van 30 frames, met een gewogen vote de actie bepalen

Inhoudsopgave

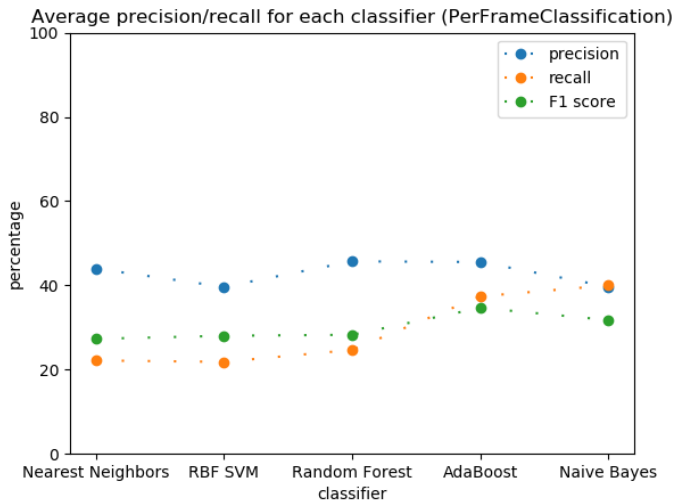
1 Context

2 Probleemstellingen

3 Methodologie

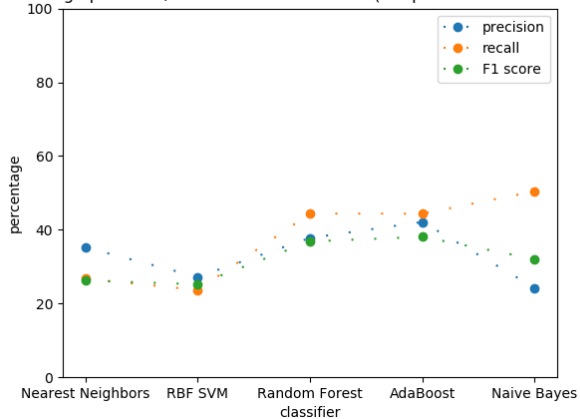
- Machine Learning
- Dataset
- Preprocessing
- Feature transformatie
- Classificatie
- Evaluatie

Evaluatie



Evaluatie

Average precision/recall for each classifier (SimpleBufferClassification)



Vragen, opmerkingen, ...?