

# Live actieherkenning met de Kinect sensor in Python

**Bert De Saffel**

Student number: 01614222

Supervisors: Prof. dr. ir. Peter Veelaert, Prof. dr. ir. Wilfried Philips

Counsellors: ing. Sanne Roegiers, ing. Dimitri Van Cauwelaert

Master's dissertation submitted in order to obtain the academic degree of  
Master of Science in Information Engineering Technology

Academic year 2018-2019

# Hoofdstuk 1

## Gerelateerde werken

- Bron [1]
  - voorstel van een methode om op een accurate manier de 3D posities van de joints te bepalen, vanuit slechts één dieptebeeld, zonder temporale informatie
  - Het bepalen van lichaamsdelen is invariant van pose, lichaamsbouw, kleren, etc...
  - Kan runnen aan 200 fps
  - Wordt effectief gebruikt in de Kinect software (onderzoeksteam is van Microsoft)
  - Een dieptebeeld wordt gesegmenteerd in verschillende lichaamsdelen, aangegeven door een kleur, op basis van een kansfunctie; Elke pixel van het lichaam wordt apart behandeld en gekleurd. Een verzameling van dezelfde kleuren wordt een joint
  - Aangezien tijdsaspect weg is, is er enkel interesse in de statische poses van een frame. Verschillen van pose in twee opeenvolgende frames is minuscule zodat die genegeerd worden
- Bron [2]
  - ✓ Bevat bruikbare datasets van skelet-, diepte- en kleurenbeelden
  - Ook hier praten ze over de vaak voorkomende uitdagingen: Intra-en interklasse variaties, de omgeving en de grootte van de verzameling van acties die er eigenlijk bestaan.
  - Hier tonen ze ook weer het nut van de kinect sensor aan, en gebruiken de kinect
  - Ze geven een nieuw algoritme om menselijke actieherkenning uit te voeren vanuit een dieptebeeld, een view-invariante representatie van poses en het systeem werkt real-time.
  - Histogram gebaseerde representatie van 3D poses (HOJ3D genoemd) = partitie van 3D ruimte in  $n$  "bins", gebruik maken van een bolcoördinatensysteem. Selectie van 12 joints die een compacte representatie van het skelet weergeven. (hand en pols, voet en enkel worden gecombineerd).
  - Het centrum van deze 3D ruimte is de heup joint. Er is ook een vector  $\alpha$ , parallel met de grond, door de heup (van links naar rechts), en een vector  $\theta$  loodrecht op de grond en door het centrum Deze 3D ruimte (figuur 3b) wordt opgesplitst in  $n$  partities.  
Voor  $\theta$ : [0, 15], [15, 45], [45, 75], [75, 105], [105, 135], [165, 180]. (7 bins)  
Voor  $\alpha$ : 30 graden voor elke bin, dus 12 bins.  
in totaal  $7 * 12 = 84$  bins  
Via deze bolcoördinaten kan elke 3D joint gelokaliseerd worden in een unieke bin
  - De 3 joints die gebruikt worden om het bolcoördinatenstelsel te oriënteren staan uiteraard vast. De overige 9 joints worden onderverdeeld in één van de 84 bins.

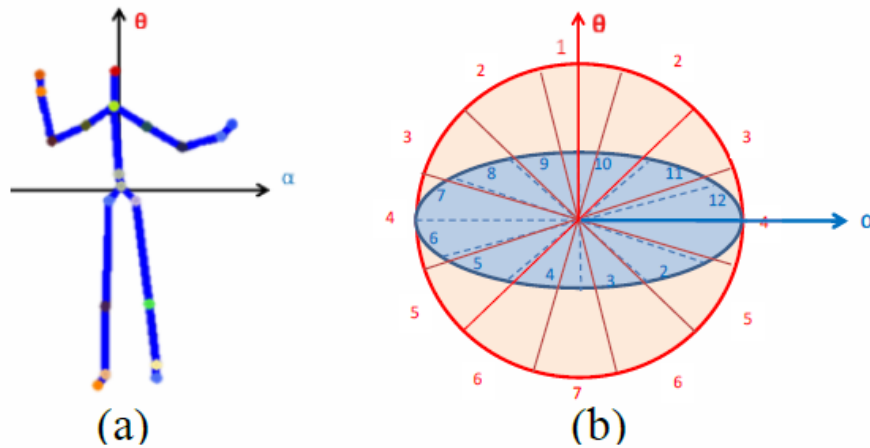


Figure 3: (a) Reference coordinates of HOJ3D. (b) Modified spherical coordinate system for joint location binning.

- Om de representatie robust te maken, wordt één enkele joint over verschillende, naburige bins verdeeld (8 burens), op basis van gewichtsfunctie:

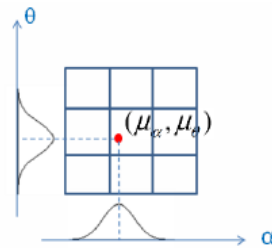


Figure 4: Voting using a Gaussian weight function.

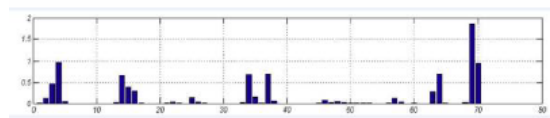


Figure 5: Example of the HOJ3D of a posture.

- Linear discriminant analysis (LDA) wordt toegepast om dominante features eruit deze histogram te halen.
- Ze beweren sneller te zijn dan bron [3]
- Bron [3]
  - Actieherkenning met behulp van reeksen van dieptebeelden
  - Gaan ervan uit dat efficiënte tracking van skeletbeelden nog niet mogelijk is. (is gepubliceerd zelfde jaar dat Kinect beschikbaar was, 2010)
  - Hun oplossing is dus niet gebaseerd op het tracken van de skeletbeelden

- Bron [4]

– s

- Bron [5]

–

# Bibliografie

- [1] J. Shotton, A. Fitzgibbon, A. Blake, A. Kipman, M. Finocchio, B. Moore, and T. Sharp, “Real-time human pose recognition in parts from a single depth image.” IEEE, June 2011, best Paper Award. [Online]. Available: <https://www.microsoft.com/en-us/research/publication/real-time-human-pose-recognition-in-parts-from-a-single-depth-image/>
- [2] L. Xia, C. Chen, and J. Aggarwal, “View invariant human action recognition using histograms of 3d joints,” in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on*. IEEE, 2012.
- [3] W. Li, Z. Zhang, and Z. Liu, “Action recognition based on a bag of 3d points.” IEEE, June 2010, pp. 9–14. [Online]. Available: <https://www.microsoft.com/en-us/research/publication/action-recognition-based-bag-3d-points/>
- [4] G. M. Sean Ryan Fanello, Illaria Gori, “Keep it simple and sparse: Real-time action recognition,” 2013.
- [5] J. Han, L. Shao, D. Xu, and J. Shotton, “Enhanced computer vision with microsoft kinect sensor: A review,” *IEEE Transactions on Cybernetics*, vol. 43, no. 5, pp. 1318–1334, Oct 2013.