

Literatuurstudie

Olivier Van den Eede

3 december 2018

1 Indoor navigatie & visie

Op visie gebaseerde navigatie is een onderwerp dat zeer vaak onderzocht wordt.

2 Object detection

Een belangrijk aspect van dit onderzoek is het detecteren van individuele objecten in het beeld van 1 enkele RGB camera. De te detecteren objecten zijn op voorhand vastgelegd, en zijn afhankelijk van de ruimte waarin de robot zich bevindt.

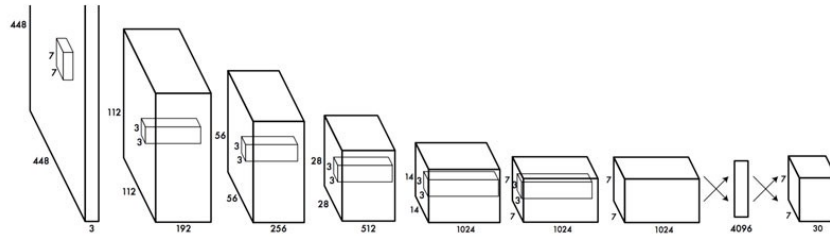
In de logistieke gangen van een ziekenhuis zijn er heel wat objecten te zien die we kunnen detecteren, een kleine selectie van deze objecten zijn.

- Pictogrammen
- Brandblussers
- Deurklinken

Voor deze objecten gaan we kijken naar detectie technieken uit de traditionele beeldverwerking, en naar meer *state of the art* technieken.

2.1 Traditionele object detectie

In openbare gebouwen zijn er heel wat pictogrammen te vinden zoals nooduitgang, hoogspanning en brandblusser. Deze pictogrammen hebben steeds een specifieke vorm, kleur en symbool. De literatuur leert ons weinig over pictogramdetectie, maar pictogrammen kunnen wel vergeleken worden met verkeersborden die bijna dezelfde kenmerken hebben. De aanpak van [1] is om 2 soorten features in een beeld te onderscheiden. In eerste instantie detecteren ze vormen op basis van kleur randen en anderzijds wordt de afbeelding omgezet naar HSI waaruit enkel de hue gebruikt wordt. De hue is de belangrijkste component voor het onderscheiden van kleuren omdat er zo geen rekening wordt gehouden met de hoeveelheid licht en schaduwen. Een recenter onderzoek [5] bouwt voort op deze technieken, maar bereken de Histogram of Oriented Gradients (HOG) features van het



Figuur 1: De lagen van een CNN volgens het YOLO [4] detection system.

beeld. Vervolgens wordt er gebruik gemaakt van een Support Vector Machine (SVM) om te bepalen waar er zich een match bevindt.

Vervolgens kunnen de vorm en kleur features gecombineerd worden om de plaats voor een mogelijke match te vinden. Eens er een mogelijke bounding box gevonden is, kan er geprobeerd worden een template te matchen om het effectieve pictogram te achterhalen. Het grootste probleem bij de techniek van [1] is dat hun gebruikte template matching techniek niet robuust is voor schaal invarianties. Bij [5] maken ze voor de herkenningfase gebruik van SIFT[3] features en kleur informatie. Hierbij worden de SIFT features van de kandidaat matches en de templates vergeleken, en er wordt een gemiddelde genomen van de verschillen tussen hue, saturation en value. Door middel van RANSAC en een threshold wordt er bepaald welke matches gebruikt worden.

2.2 Convolutional neural network

Een convolutional neural network of CNN is een supervised deep learning techniek die gebruikt kan worden om complexere beeldinterpretatie te doen. Een CNN kan bestaan uit meerdere lagen die meestal een combinatie zijn van 'convolutional-layers' en 'fully connected-layers'. Elk van deze lagen bevat een aantal neuronen met elk een eigen set van gewichten. Het doel van een CNN is om de gewichten zodanig bij te stellen zodat data die aan de eerste laag gegeven wordt een verwacht resultaat geeft aan de laatste laag.

3 Object tracking

4 Image segmentation

Het correct segmenteren van de beelden zal een belangrijke rol spelen. Niet in elk beeld zal er een distinctief object aanwezig zijn om te detecteren. Daarom is het belangrijk om de vloer van de muren te kunnen onderscheiden. Een eenvoudige approach zou kunnen zijn om via K-means een verdeling van een beeld te doen en met een soort regressie de regio's te labelen. Volgens [6] werkt de K-means aanpak met een op textuur en kleur gebaseerde aanpak redelijk goed, maar wordt steeds de muur verbonden met het plafond omwille van kleur en textuur gelijkenissen. Hun regressie gebaseerde labeling techniek blijkt echter een slechte oplossing. Verder zoals [2] aangeeft zijn reflecties en overbelichting eigenschappen van indoor omgevingen die het moeilijk kunnen maken om een correcte segmentatie te doen.

Referenties

- [1] Chiung Yao Fang, Sei Wang Chen, and Chiou Shann Fuh. Road-sign detection and tracking. *IEEE Transactions on Vehicular Technology*, 52(5):1329–1341, 2003.
- [2] Yinxiao Li and Stanley T. Birchfield. Image-based segmentation of indoor corridor floors for a mobile robot. *IEEE/RSJ 2010 International Conference on Intelligent Robots and Systems, IROS 2010 - Conference Proceedings*, pages 837–843, 2010.
- [3] D G Lowe. Object recognition from local scale-invariant features. In *Proceedings of the Seventh IEEE International Conference on Computer Vision*, volume 2, pages 1150–1157 vol.2, sep 1999.
- [4] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun 2016.
- [5] S. J. Zabihi, S. M. Zabihi, S. S. Beauchemin, and M. A. Bauer. Detection and recognition of traffic signs inside the attentional visual field of drivers. *IEEE Intelligent Vehicles Symposium, Proceedings*, (Iv):583–588, 2017.
- [6] Zhong-Ju Zhang. Wall, floor, ceiling, object region identification from single image.