

Human detection using depth information by Kinect

Door: Boris

Methodes om personen te herkennen in afbeeldingen en video's [Referenties: 1,2,6,13]. Veel problemen geweest met het scannen van RGB afbeeldingen die door de achtergrond niet accuraat waren of veel rekenkracht kostten. Veel research is de afgelopen tijd gedaan naar het herkennen van posen en menselijke lichaamsdelen. Hierbij zijn verschillende methodes gebruikt: stereo camera's [3,4,15], time-of-flight diepte camera's [5,7,11,16]. In hun eigen algoritme combineren ze 2D edge detector en een 3D shape detector om hoek informatie en de gerelateerde diepte verandering informatie samen te voegen en op die manier hoofden te detecteren.

In dit paper is onderzoek gedaan naar een menselijke detectiemethode voor de Kinect van de Xbox360. Hiervoor zijn meerdere algoritmes gebruikt die met behulp van de diepte beelden van de Kinect personen proberen te herkennen. Het herkenningsproces is opgedeeld in een aantal stappen, deze stappen zijn als volgt:

1. Invoer diepte afbeelding.
2. Data opschonen door de missende waarden van de diepte afbeelding in te vullen met nearest neighbours en vervolgens een 4x4 median filter over de bitmap heen halen.
3. Een 2d afbeelding maken van de randen die te vinden zijn in de diepte afbeelding met behulp van Canny edge detector.
4. Berekenen van de parameters van het hoofd (grote, diepte, breedte, positie, middelpunt).
5. Genereren van een 3D van het hoofd & fit.
6. De persoon scheiden van de achtergrond.
7. De persoon en zijn contouren detecteren.
8. Tracking van de persoon

Dit algoritme heeft moeite wanneer de hoofden van personen onderbroken worden door andere voorwerpen of personen, en wanneer personen zich niet volledig in beeld bevinden.

In de toekomst willen de onderzoekers werken aan een combinatie van dit algoritme met een ander algoritme dat andere lichaamsdelen kan herkennen. Op die manier hopen ze een persoon ook te kunnen herkennen als een deel van het lichaam niet in beeld is.

Slimme technieken:

nearest neighbours algorithm om missende waardes uit de diepte array op te vullen. Mediaan filter met 4x4 window om de diepte array 'smooth' te maken.

gebruik randen informatie om regio's met personen te detecteren, hierbij streven ze naar een zo laag mogelijke false negative rate. Canny edge detector wordt gebruikt om randen te mappen vanaf een bepaalde grote, dit geeft beter weer waar personen zich bevinden in de afbeelding. Van de afbeelding die hieruit overblijft wordt een distance map gemaakt. Met behulp van een binary head template kunnen ze echte hoofden uit de foto detecteren. Door deze template te draaien worden personen in allerlei houdingen gedetecteerd.

Daarbij wordt het centrum van het hoofd berekent en op basis van die gegevens en de diepte konden ze bepalen of de breedte en hoogte van het hoofd accuraat werden gemeten.

Het is bijna onmogelijk om uit een diepte afbeelding te meten waar de voeten van een persoon de grond raken omdat de diepte van beiden ongeveer gelijk zijn. Dit hebben ze gecorrigeerd door ervan uit te gaan dat voeten altijd rechtopstaande objecten zijn.

Nadat de contouren van de lichaamsdelen die andere objecten aanraken wat duidelijker zijn gemaakt, wordt met behulp van een region growing algorithm bepaald wat het menselijk lichaam is op de foto. Hierin is het eerder bepaalde centrum van het hoofd, het beginpunt.

Tenslotte tracking, in een diepte afbeelding kan je alleen gebruik maken van de diepte die objecten. Er wordt van uitgegaan dat de objecten op een soepele manier van positie en snelheid wisselen. Met de huidige coördinaten en snelheid van een persoon en die van de vorige frame kan worden getraceerd hoe objecten zich bewegen.