# K-means Clustering

## Namen en datum

Stefan Hulspas en Daan Joling   
18-maart-2014

## Doel

Het doel van het experiment is om te weten te komen hoe een K-Means clustering werkt. Om hier achter te komen hebben we een test set gemaakt. De test set bestaat uit een aantal afbeeldingen. Om deze test set laten we ons filter los. Met de resultaten gaan we kijken wat voor effect het filter nou heeft op de afbeeldingen.

## Hypothese

Door K-means Clustering worden alle pixels ondergebracht aan het x-aantal clusters dat van te voren word vast gesteld. Wij verwachten dat dit dus ook gaat gebeuren met onze test set. De clusters worden elke keer random gekozen, omdat de core pixel elke keer anders is. Na een x-aantal iterations zijn alle pixels goed ondergebracht bij het juiste cluster. Onze onderzoeksvraag is: “wat is de beste K-Means clustering configuratie”.

## Werkwijze

Het programma wat is geschreven gaat de afbeelding bewerken.

## Resultaten

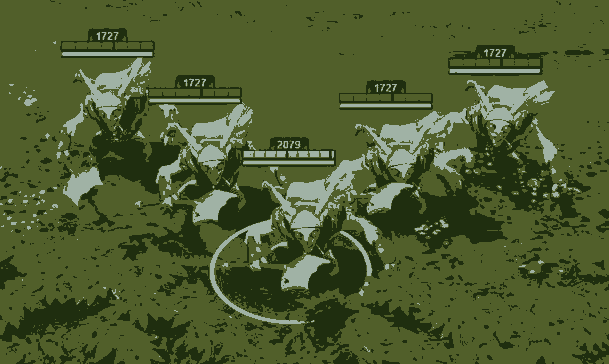


Figure 1 afbeelding met 3 clusters

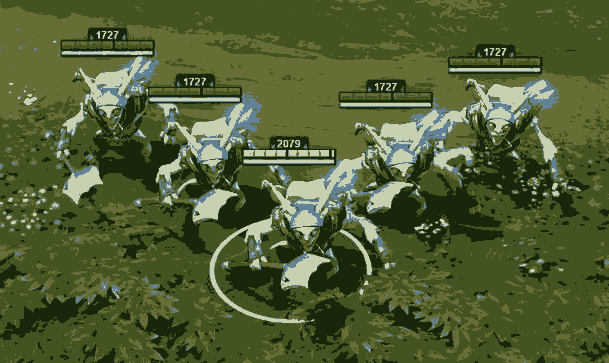


Figure 2 afbeelding met 5 clusters

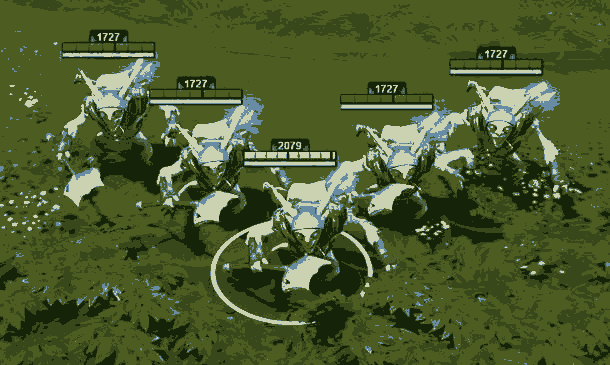


Figure afbeelding met 7 clusters



Figure afbeelding met 20 clusters

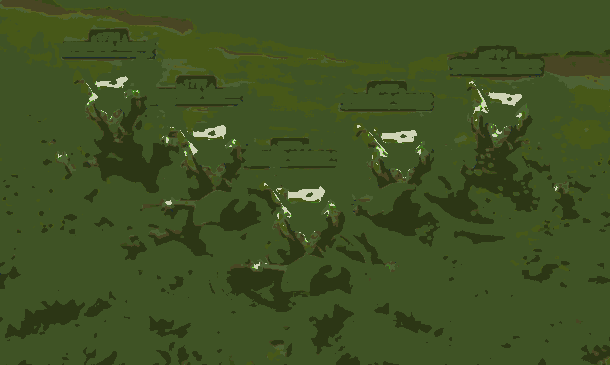


Figure afbeelding met 20 clusters

## Conclusie

Uit het experiment kunnen we afleiden dat de beste configuratie afhangt van hoeveel tijd je hebt. Hoe meer clusters je hebt, desde meer tijd kost het om een goed plaatje te creeren. Bij onze resultaten kunnen we zien dat 7 clusters net op het randje van 25 iteraties zit, hierdoor zijn alle clusters nog niet goed uitelkaar geschoven en kan je grote vlakken met de zelfde kleur vinden. Bij 20 clusters is dit helemaal goed te zien. Het plaatje 4 heeft een paar pogingen gekost om te produceren omdat de clusters random gekozen worden. Dit plaatje verschilt sterk van plaatje 5, waar de clusters dichter bij elkaar gekozen zijn.  
Tot slot zijn wij van mening dat de beste K-Means configuratie afhangt van wat je precies wilt bereiken; een mooi plaatje of onderscheid tussen bepaalde kleuren.