# Meetrapport week 3-4

## Namen en datum

Lars Veenendaal 1633223 TIV-2A

07-04-14

## Doel

Het doel is om K-Means algorithme te implementeren en zo op te bouwen dat hij zijn optimale middenpunten kan berekenen. Om zo de ideale clusters te generen. Hier voor zal ik moeten experimenteren in hoe ik dit zou moeten opbouwen en ervoor moeten zorgen dat hij niet teveel werk verricht.

## Hypothese

K-Means maakt clusters aan en deze clusters groeien van elke pixel in een afbeelding. Deze pixels worden toegewezen aan clusters waar ze het dichtbijst zijn. Als alles verdeelt word word gekeken naar nieuwe cluster punten en begint het feest opnieuw.

## Werkwijze

Na vele testen in hoe ik de clusters kon laten groeien dynamisch en herschalen. Heb ik besloten om een linkedlist aan te maken en deze te gebruiken voor de clusters en de data die daarin mee gaat.

Verder heb ik er lang op gezeten over hoe ik hem het best ‘overnieuw’ kon laten beginnen. Het komt soms voor dat clusters prachtisch opelkaar liggen en dat moet bijgewerkt worden wat nogal een uitdaging was.

## Resultaten

K-Means 3 clusters 20 Iteraties K-Means 5 clusters 20 iteraties



K-Means 7 Clusters 20 iteraties

 K-Means 10 clusters 1 iteratie

 K-Means 25 clusters 2 iteratie

Verder is duidelijk zichtbaar het verschil wat een paar iteraties maakt. De kleuren worden meer helder en komen los van elkaar na mate van iteraties vorderen.

Maar met elke iteraties neemt de tijd toe. En bij uitstek het meest tijdrovende is alle clusters omzetten naar pixels met hun nieuwe kleuren.

## Verwerking

Voor de verwerking heb ik veel getest op verschillende manieren om de iteraties sneller te krijgen van standaard averaging tot met een kleuren histogram kijken waar de meeste kleuren zitten en basis daarvan je begin posities instellen van de clusters. Maar na veel tweak werk en weer terug brengen naar zijn roots. Is er nu een werkend thans trage k-means algoritme die als zeg ik het zelf mooie plaatjes oplevert.  
  
Verder heb ik procenten bij bepaalde processen gezet om een indicatie te hebben dat het systeem nog wat aan het doen is. Bij een laag aantal clusters kunnen de updates in procenten heel lang op zich laten wachten na in een situatie met veel clusters dan duurd het opsplitsen iets langer maar zie je sneller effect.

## Conclusie

Hoe meer iteraties hoe helderder te splitsing tussen kleuren.   
Hoe meer clusters hoe meer kleuren je krijgt.,  
  
Het is het optisch zeker waard meer clusters te gebruiken maar dan zou ik wel koffie gaan halen. Of je favoriete websites lezen want je bent zo een half uurtje kwijt.

## Evaluatie

Het effect kan problemen hebben als je op voorhand een afbeelding hebt met een aantal kleuren. Als in minder kleuren dan clusters. Dan gaan de clusters groeperen en dan heb je een cluster die eigenlijk verdwijnt.