



Internal Report 2012-2013-01

Januari 2013

Universiteit Leiden

Opleiding Bachelor Informatica

Image classification
on Flickr
using TOP-SURF Visual Word Kit

Nuno L. D. S. Rico
Studentnummer S9538070

Begeleider: Dr. D.P. Huijsmans

BACHELOR THESIS

Leiden Institute of Advanced Computer Science (LIACS)
Leiden University
Niels Bohrweg 1
2333 CA Leiden
The Netherlands

1 Samenvatting

Afgelopen decennia is registratie, verwerking en presentatie van visueel gerelateerde informatie door computersystemen enorm gegroeid. De toename van het gebruik van foto-, beeld- en video materiaal op het internet, en de groeiende vraag naar classificatie en de registratie van biometrische gegevens, hebben een belangrijke bijdrage geleverd aan de ontwikkeling van de beeldverwerking.

Het adagio '*A picture is worth a thousand words*' herineert ons aan de hoeveelheid informatie welke aanwezig is in een beeld, zowel kwantitatief als kwalitatief. Dit verslag geeft de resultaten en analyse weer van het door mij uitgevoerde onderzoek naar mogelijk vervanging van visual words uit Top-surf [1] voor het handmatige labellen van beelden en in het bijzonder Flickr beelden.

Het onderzoek is uitgevoerd op de MIR Flickr [6] image database van Flickr [5] internet site.

For such an advanced civilization as ours to be without images that are adequate to it is as serious a defect as being without memory.

Werner Herzog, Duitse filmregisseur

Sleutelwoorden: Information Storage and Retrieval, Pattern Recognition, Benchmarking

2 Inleiding

Het gebruik van beeldmateriaal is niet meer weg te denken uit ons dagelijks leven. Foto en video kom je overal tegen, zowel op de voorpagina van tijdschriften als in computer programmas, in elektronische spelletjes en op internetpaginas, enz. De toepassing van visuele en multimedia applicaties zijn legio.

De geleidelijke toename van computer processor snelheid en de afname van kosten voor de aanschaf van intern geheugen bieden nieuwe mogelijkheden voor verbetering van bestaande en de toepassing van nieuwe innovatieve beeldverwerking technieken voor *information retrieval*.

Steeds frequenter worden databases ingezet voor de opslag en het terugvinden van een breed scala aan beeldinformatie, zoals medische gegevens, satellietbeelden, fotos van personen en verkeersbeelden om maar een paar voorbeelden te geven. Een van de grootste uitdagingen op het gebied van beeldverwerking is de weergave en het terugvinden van beelden. De meeste gebruikte methode voor het terugvinden van gelijke beelden is door de image distance measure te berekenen (dit is de afstand tussen de beelden) op basis van gelijkenis van kleur, textuur, vorm of object soort.

Gedigitaliseerde fotos worden op verschillende methodes representeerd en opgeslagen in digitale computersystemen. Een daarvan is Top-surf [1]. De methode ondersteunt een breed scala aan content based image retrieval en is ontworpen om beeld vergelijkingen snel te kunnen uitvoeren. Top-surf is een veel gebruikt image descriptor algoritme dat features of interest points in een beeld tot visual words combineert. Het is flexibel van opzet en biedt de mogelijkheid de descriptor lengte te variëren,

Top-surf wordt opgeleverd als C++ Class library (bibliotheek) formaat met een API Toolkit [9] en werd eerst geïntroduceerd en geïmplementeerd door Bart Thomée in 2010. Sindsdien wordt het breed toegepast voor het terugvinden van gerelateerde beelden.

Deze scriptie geeft de resultaten weer van het onderzoek naar het gebruik van visual words door Top-surf als vervanging van het manueel labelen van Flickr beelden en of een bepaalde combinatie van Flickr tags een zinnige samenhang vertoont met de instellingen van Top-surf visual words woordenboek (dictionary).

In dit bachelor project wordt verder onderzocht:

- Of door slimme parameter instellingen de Top-surf visual word combinaties een opdeling opleveren die op een zelfde manier samenhang vertoont als de Flickr tag(s).
- Wat de invloed is van het in visual words gebruikte aantal clusters en een mogelijke optimum daarin.
- Hoe goed presteert het optimale systeem van Top-surf in termen van nauwkeurigheid (recall en precision) waarbij Flickr tags van de test data als ground truth (grondwaarheid) zullen dienen.
- Zijn er bepaalde type tags die het duidelijk beter of slechter doen dan anderen (b.v. lucht/water/plant in vergelijking met mens/portret/nacht).

De beoordeling van beelden door menselijke gebruikers is subjectief. De relevantie van beelden is geen exacte wetenschap, het is ook niet met zekerheid aan te geven of combinaties van beeld-label juist of niet juist zijn. Hoewel het subjectiviteitsaspect een rol speelt bij de beoordeling van beelden gaat dit projectonderzoek uit van de juistheid van de aan het beeld toegekende labels.

3 MIR Flickr tags

De MIR Flickr beelden collecties uit de sociale fotografie website Flickr.com citeflickr hebben bestanden van 25.000 (25K) en 1.000.000 (1M) beelden en zijn samengesteld uit beelden van Flickr onder de Creative Commons license [7]. MIR Flickr image retrieval beeld verzamelingen worden beschikbaar gesteld via internet.

MIR Flickr 25K is een relatief kleine dataset van een gevarieerde collectie beelden met nauwkeurig beschreven grondwaarheid en wordt met gestandaardiseerde tests geleverd. Het maakt deze verzameling geschikt voor het onderzoek naar de bruikbaarheid van Top-surf voor het zoeken naar beelden en de relatie en samenhang met de door de maker toegekende tags.

Voor de opzet van het onderzoek maak ik gebruik van een beeld bestand van Flickr bestaande uit 25.000 beelden, welke zijn gedownload van The MIR Flickr Retrieval Evaluation website LIACS [6]. Het geselecteerde bestand is niet groot, maar geeft wel een goede representatie van het type beelden dat gebruikt is in onderzoeken naar beeld matching. Het Flickr platform biedt gebruikers de mogelijkheid hun foto's te zoeken op basis van afbeelding tags. Deze tags worden als grondwaarheid beschouwd en geïnterpreteerd als de waarheid. Het gemiddelde aantal tags per beeld is 8,94 [8].

De foto collectie is heel gevarieerd en bestaat uit beelden van vele interessante plaatsen op de wereld. Sommige tags refereren naar abstracte concepten zoals genieten, liefde, reizen of bijvoeglijke naamwoorden (bijv. oud, nieuw, liefelijk). Andere categorien verwijzen naar Flickr gerelateerde terminologien zoals verkennen, mate van belang, geweldig of mislukte foto, natuurschoon, type en merk cameras (Nikon, Canon) [8], enz. De aan de beelden toegekende tags worden als grondwaarheden opgeslagen en zoals de gebruikers de fotos hebben genoemd.

Over het algemeen hebben Flickr beelden meer dan 1 tag, wat de interpretatie van beelden bemoeilijkt. De tag benaming van beelden, gegeven door de maker, is vaak een persoonlijke interpretatie van de foto. Voor de uitvoering van het onderzoek is gekozen om beelden te kiezen met meer dan één tag. Dubbelzinnige beelden met maar één tag kunnen onevenredige invloed hebben op de wijze van clustering en uiteindelijk op de berekende precision en recall waardes.

Aanvankelijk zijn voor het onderzoek beelden geselecteerd met precies drie tags. Combinaties van beelden met meer dan drie tags zouden de interpretatie van de resultaten bemoeilijken. Door het hoge aantal tag combinaties is niet meer te achterhalen wat de

invloed is van een of meerdere tags op de rangschikking van beelden door het algoritme. De MIR Flickr verzameling van 25K beelden kent 24 grondwaarheid (tags) in totaal en heeft 149 groep combinaties van beelden met precies drie tags. Van alle groepen met drie tags hebben sommige groepen maar 1 beeld en maar 2 groepen meer dan 100 beelden.

4 Top-surf Visual words

Top-surf visual words Tool Kit [9] is een open source *image descriptor* woordenboek dat punten van interesse (*features*) van beelden met visual words combineert voor gebruik in content based image retrieval applicaties. Top-surf, voor het eerst gepresenteerd en uitgevoerd door B. Thomée[1] is een high performance compacte descriptor ontworpen voor beeld retrieval toepassingen. Het kunnen variëren van de descriptor lengte en de snelheid waarmee beelden kunnen worden vergeleken, maakt van Top-surf een flexibele kandidaat voor beeldmatching.

Top-surf is gebaseerd op SURF, een robuuste beelddetector en -descriptor, en voor het eerst gepresenteerd door Herbert Bay et al. [18]. SURF maakt gebruik van HAAR wavelet responses in de omgeving van punten van interesse als descriptors. Met behulp van Hessian matrix berekeningen worden de punten van interesse opgespoord en gemeten. Elk interesse punt wordt geassocieerd met een descriptor van 64 elementen. Het aantal punten van interesse is variabel. Voor de vaststelling van de matches tussen twee beelden maakt SURF gebruik van de nearest neighbors (*naaste buren*) ratio matching methode. Elk punt van interesse in de eerste afbeelding vergeleken met alle punten van interesse van het tweede beeld door het berekenen van de Manhattan afstand tussen de descriptors (ook wel *city-block distance* genoemd). Een match wordt gevonden als de afstand tussen een punt in het eerste beeld en een punt in het tweede beeld kleiner is dan 0,65 (drempel waarde) maal de afstand en wanneer elk ander punt in de tweede beeld is overwogen.

Top-surf wordt aangestuurd met behulp van een uitgebreid aantal parameters voor de opbouw van de *visual words* woordenboeken. Top-surf wordt opgeleverd met de OpenSURF [10] image descriptor bibliotheek onder de GNU GPL license, OpenCV [11] bibliotheek onder BSD licentie [3], FLANN [12] approximate nearest neighbor matching using KD-Trees onder BSD licentie en CXIMAGE [13] image representatie onder ZLIB licentie [4].

5 Opzet van het onderzoek

Het gebruik van de Top-surf bibliotheek routines voor de berekening van de afstand tussen beelden bij een grote beeldcollectie is niet erg efficiënt, elk beeld van een geselecteerde groep-tag combinatie wordt met alle andere beelden van de Flickr collectie vergeleken. Dat is een tijdrovend en kostbaar proces in termen van computer verwerkingstijd. Bovendien moet de matching van beelden voor diverse groep-tag combinaties woordenboeken worden herhaald. Het ophalen en de matching van twee beelden met behulp van de bibliotheken van Top-surf neemt 0,46 seconden in beslag. Het gehele proces van beeld voor beeld

vergelijking duurt in totaal 11.500 seconden (3,19 uur) op een collectie van 25.000 beelden.

Uitgaande van een gemiddelde van 69,3 beelden voor een groep-tag bestand, zou de matching van een groep-tag combinatie 221,4 uur in beslag nemen. Uiteindelijk zullen 13 bestanden met groep-tag combinaties op 5 woordenboeken worden geëvalueerd en dat brengt de totale analysetijd op 1.107 uren. Dat is zeer inefficiënt om niet te zeggen ondoenlijk op een desktop computer.

Gelet op de lange doorlooptijd welke nodig is voor de opbouw van de bibliotheken en analyses, is in een vroeg stadium van het onderzoek besloten om het proces voor de verwerking en matching van beelden anders in te richten. Door alle relevante beelddescriptorgegevens te bewaren in een database, wordt de matching en de toegang tot de beelden efficiënter en wint het onderzoek aan snelheid. Voor de analyse van de gegevens en voordat de Manhattan beeldafstand berekening wordt uitgevoerd, worden eerst de database gegevens in het geheugen van de computer geladen.

De grondslag voor de opzet van het onderzoek is de meting van mogelijke overeenstemmingen van Top-surf visual words met de tags in Flickr. Het onderzoek wordt in de volgende stappen uitgevoerd:

- Initiële opbouw van woordenboeken met een gevarieerd aantal clusters.
- Initiële opbouw van databases, één voor elk woordenboek.
- Selectie van groep beelden met precies drie tags en één representatief aantal beelden.
- Afstand berekenen van alle beelden in de groep/tag combinaties naar overige beelden van de MIR-Flickr collectie. Beelden rangschikken op berekende afstand van klein naar groot, kortere afstanden betekent een grote graad van beeld gelijkenis.
- Analyse van de resultaten en selectie van parameterwaardes die de performance van het algoritme positief beïnvloeden. In deze fase van het onderzoek worden de verkregen resultaten geëvalueerd, geanalyseerd en bepaald welke parameters gewijzigd worden teneinde een betere performance te krijgen, waarna stappen 1 tot 5 worden herhaald.

5.1 Inrichting woordenboek

Initiëel worden een aantal Top-surf woordenboeken opgebouwd en ingericht met verschillende cluster waardes. In eerste instantie werd onderzocht wat de invloed is van de parameter grootte op de nauwkeurigheid bij het vinden van beelden van dezelfde groep-tags. Andere parameters worden in dit stadium op de standaard waardes gezet. Standaard waardes zijn waardes beschreven in de API module.

De Top-surf descriptor baseert zich op de bag of words techniek van Philbin et al. [19] voor representatie van gegroepeerde verzamelingen van punten van interesse in clusters. Voor het bepalen van de cluster worden eerst de 100 naaste burens met dichtstbijzijnde

punten van interesse bepaald.

Top-surf ontleent zijn naam aan de keuze om de top 100 beelden (TOP) te nemen voor bepaling van het cluster center. Elke cluster wordt bijgewerkt door de gemiddelde afstand te berekenen van de huidige locatie naar die van zijn naaste burens. Om de stabiliteit van de cluster center te waarborgen wordt dit proces een aantal keer (iteraties) herhaald (standaard 1.000). Opsporen van de burens in een hoge dimensionale ruimte neemt veel tijd in beslag. Het proces wordt versneld door gebruik te maken van de naaste burens benaderingstechniek en met behulp van KD-Trees [20].

De TD-IDF (Term Frequency - Inverse Document Frequency) wordt door Top-surf gebruikt als wegingsfactor om aan te geven hoe belangrijk een kenmerk (interesse punt) is in een cluster. TD-IDF vermindert het gewicht van het kenmerk dat vaak voorkomt in de beeld-database en verhoogt het gewicht van het kenmerk dat zelden voorkomt. De TD-IDF waarde is de verhouding van het aantal keren dat een kenmerk voorkomt in de cluster ten opzichte van de frequentie van het kenmerk in de verzameling van beelden. Dit verrekent het feit dat sommige kenmerken over het algemeen vaker voorkomen dan anderen.

Het woordenboek bestaat uit vier bestanden waarin cluster aantal, de TD-IDF waardes van de clusters, de KD-tree verzameling van visual words en de visual words zelf zijn opgeslagen.

Voor de opbouw van het visual words woordenboek maakt Top-surf gebruik van parameters die via de API module worden aangestuurd. Maximaal aantal clusters, aantal naaste burens per cluster, aantal iteraties en aantal punten van interesse per beeld zijn de belangrijkste parameters voor de opbouw van het woordenboek.

Top-surf heeft een aantal interne parameters die niet beïnvloedbaar zijn van buitenaf zoals het maximale aantal punten van interesse ingesteld op 33.55 miljoen en het aantal beelden voor het bepalen van de inverse document feature ingesteld op 2.500. Top-surf kijkt alleen naar de eerste top 100 "visual words" van een beeld voor de berekening van de descriptor. Hieraan ontleend Top-surf ook zijn naam.

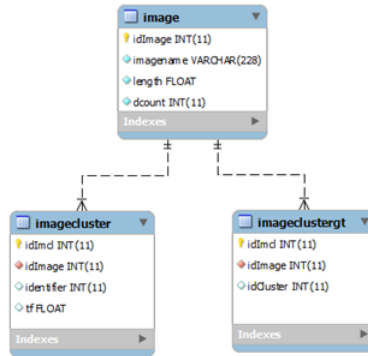
5.2 Inrichting database

Voor het verloop van het onderzoek is het belangrijk dat de visual words snel kunnen worden gevonden. Relevante beeldinformatie en de visual words worden in een database opgeslagen die de toegang tot de beeldinformatie bevordert. Cluster en beeld descriptor informatie van het woordenboek worden bewaard in een MySQL [15] relationele database. Dat vereenvoudigt de toegang tot beeld descriptors en de snelheid van beeld matching gedurende het onderzoek.

Voor de opbouw van de database is de C++ applicatie Imagecla3.exe ontworpen. Met behulp van deze applicatie wordt de database met descriptor en visual word gegevens gevuld.

De Top-surf DLL bibliotheek biedt een aantal functies aan voor de berekening van de afstand tussen de beelden. Het aanroepproces van de API functies is traag en niet erg

efficiënt. Elk beeld in de selectie dient te worden vergeleken met alle andere beelden in de verzameling. Door de folder referentie, evenals de descriptor gegevens in de database, te bewaren is het niet meer nodig de bibliotheek routines aan te roepen. Omdat de bibliotheken niet alleen voor de opbouw van de descriptor zorgden, maar ook voor de berekening van de beeldafstand tussen de visual words zijn de routines voor de afstands berekening in de analyse applicatie nagebouwd. Deze aanpak zorgde voor aanzienlijke versnelling van het matchingsproces.



Figuur 1: Database model en tabel indeling

Het relationeel model van de database bestaat uit drie hoofd tabellen *image*, *imagecluster* en *imageclustertgt*. De tabellen bevatten beeldgegevens zoals de referentie naar de folder locatie op schijf, descriptor lengte, aantal interesse punten en de cluster- en grondwaarheden van de beelden. De aanroep van *DLL* functies van de Top-surf bibliotheek vindt plaats tijdens de opbouw van de database, eenmalig. Door middel van een dll functie worden tijdens de opbouw fase van de database de beeld-descriptors en visual words opgehaald voor alle beelden van de collectie. De referentie naar de visual word identifiers en de term/frequency (tf) waarden, nodig voor de latere berekening van beeldafstand, worden in de tabel *imagecluster* in main memory opgeslagen.

Vanaf het begin van het onderzoek was duidelijk dat er minimaal vijf databases dienden te worden gebouwd. Het was noodzakelijk snel te kunnen schakelen van de ene naar de andere database voor analyse van de beelden. Om dit proces zo efficiënt mogelijk te laten verlopen is er voor gekozen om een basis database te bouwen, eenmalig en zonder cluster en grondwaarheid gegevens en met alleen de grondwaarheden (*imageclustertgt*). De enige gegevens die veranderden bij de opbouw van de woordenboeken waren de referenties naar de visual words in de tabel *Imagecluster*. Voor het vullen van de basis database is de Perl applicatie *Imageclal.pl* ontworpen. De parameters *\$datapath* en *\$labelpath* sturen de fysieke locatie aan van waar de beelden en label (tag) grondwaarheid referenties zich bevinden. Een eenmaal opgebouwde database voor een woordenboek-cluster combinatie kan in MySQL eenvoudig en snel worden geladen.

5.3 Selecteren van groep en tags voor analyse

Bij de opzet van het onderzoek is eerst een selectie gemaakt van de combinaties van groepen beelden met precies drie tags en meer dan 20 beelden. Bij de selectie van groep-tag combinaties is speciale aandacht gegeven aan de selectie van zo uitgebreid mogelijke variaties van tags (labels). Er is ervoor gekozen om de groep-tag combinaties te selecteren met meer dan 20 beelden. Groepen met weinig beelden zouden een onevenredige invloed hebben op de verkregen onderzoeksresultaten.

Geselecteerde groepen met precies 3 tags voor het onderzoek:

group	cluster1	cluster 2	cluster 3	number images	Cluster 1 name	Cluster 2 name	Cluster 3 name
1	15	19	23	43	plant_life	sky	tree
2	15	20	23	54	plant_life	structures	tree
3	13	19	20	49	night	sky	structures
4	12	14	16	39	male	people	portrait
5	12	14	20	85	male	people	structures
6	10	12	14	133	indoor	male	people
7	7	10	14	164	female	indoor	people
8	7	14	16	94	female	people	portrait
9	7	14	20	62	female	people	structures
10	5	19	20	71	clouds	sky	structures
11	1	6	10	67	animals	dog	indoor
12	1	6	20	21	animals	dog	structures
13	1	15	23	20	animals	plant_life	tree
Total number of images				902			

Figuur 2: Groepen in onderzoek

Het aantal grondwaarheid tags per beeld verschilt enorm. Voor de extractie van alle beelden met precies het gewenste aantal tags (grondwaarheden) is de Perl applicatie *groupclusters.pl* ontworpen. De selectie aantal tags per beeld wordt aangestuurd door de parameter *nr_clusters* te veranderen.

5.4 Prestatie meting

Voor patroonherkenning worden *precision en recall* gebruikt voor de meting van de prestatie van de gevonden beelden (*image retrieval*). Precision geeft het aantal opgehaalde exemplaren aan die relevant zijn, recall geeft de fractie aan van relevante gevallen die zijn opgehaald. Zowel precisie en recall zijn gebaseerd op de mate van relevantie. De met Top-surf teruggehaalde beelden worden in vier klassen verdeeld:

- *true positives*, relevante beelden geclassificeerd als (*hits*) en die tot de *class of interest* behoren.
- *false positives*, beelden onjuist geclassificeerde als *hits* en die niet tot de *class of interest* behoren.
- *false negatives*, relevante beelden niet geclassificeerd als *hits* maar die wel tot de *class of interest* behoren.

- *true negatives*, niet relevante beelden en die terecht niet tot de *class of interest* zijn geclassificeerd.

		ground truth	
		class of interest	class not of interest
label given	predicted positive	True Positive	False Positive
by algorithm	predicted negative	False Negative	True negative

Figuur 3: Classificatie van beelden (relevantie)

Voor elke groep-tag combinatie wordt met behulp van Precision en Recall de afstand tussen de beelden uitgerekend. De verkregen resultaten geven de bruikbaarheid en performance weer van Top-surf voor het vinden van handmatig gelabelde beelden in Flickr.

Precision geeft de proportie van de gevonden beelden aan die positieve voorspellingen zijn, ofwel voldoende gelijkenis tonen (true positives) met het referentiebeeld. Precision meet het vermogen van het systeem om relevante beelden te vinden en presenteren.

$$Precision = \frac{| \text{true positives} |}{| \text{bekeken top aantal beelden} |}$$

Recall geeft de proportie relevante beelden aan die zijn gevonden (true positives) ten opzichte van de target beelden ofwel positieve voorspellingen (true positives en false negatives). Recall meet het vermogen van een systeem om alle relevante informatie te presenteren en geeft aan hoe compleet het algoritme is.

$$Recall = \frac{| \text{true positives} |}{| \text{true positives} | + | \text{false negatives} |}$$

$$Recall = \frac{| \text{true positives} |}{| \text{Groepsgrootte} |}$$

Voor elk beeld van de groep/tag en alle beelden van het Flickr bestand wordt de afstand naar alle andere beelden in de database berekend. De beelden worden op beeld afstand van kleiner naar groter gesorteerd en gerangschikt. **Omdat het aantal onderzochte top beelden gelijk is aan het totaal aantal beelden in de gezochte groep, komen de verkregen waardes voor Precision en Recall met elkaar overeen. In een ideaal systeem waar alleen naar de groeps grootte wordt gekeken, hebben precision en recall hetzelfde waarde (ideaal systeem als toets criterium), zie figuur [17].** De performance wordt getoetst door de precisie afstanden van het Top-surf algoritme met de random search te vergelijken.

$$Random\ precision = \frac{|\text{true positives}|}{|\text{aantal beelden in de Flickr collective}|}$$

Het *gemiddelde precision* van een groep tag wordt berekend met behulp van de MAP (*mean average precision*) methode gebruikt.

$$MAP = \frac{\sum_{i=1}^n precision(i)}{N}, \text{ n is het totale aantal beelden in de groep}$$

Voor de berekening van de afstand tussen de beelden van MIR Flickr wordt de *Manhattan distance* gebruik, ook wel *city-blok* afstand genoemd. Hiermee wordt de afstand berekend tussen twee punten $x = x_1, \dots, x_n$ en $y = y_1, \dots, y_n$ als de som van de absolute waarde van het verschil tussen de coördinaten.

$$d_{L_1}(x, y) = \sum_{i=1}^n |x_i - y_i|$$

5.5 Verloop van de evaluatie procedure en uitwerking

Het onderzoek is in fases opgezet en bestaat uit twee te onderscheiden fases. In de eerste fase van het onderzoek worden de Top-surf parameters op standaard waarden gezet waarbij alleen het aantal clusters verandert. Elk woordenboek wordt bewaard voor latere analyse. In de eerste fase zijn vijf woordenboeken opgebouwd met het aantal clusters op respectievelijk 25, 100, 500, 1.000 en 5.000. Overige parameters zijn met de volgende standaard waarden geïnitieerd: *nearest neighbors* = 100, *iterations* = 100, *aantal interesse punten* = 25.

Voor de tweede fase worden eerst de woordenboeken met de beste prestatiewaarden geselecteerd voor de overige parameters te gaan variëren. Het uiteindelijke doel is een woordenboek genereren met grondwaarheden die zoveel mogelijk overeenkomen met gerelateerde Flickr tags. De prestatie wordt bepaald door de Manhattan afstand te meten tussen het te onderzoeken beeld naar alle andere beelden van de Flickr verzameling. Beelden worden gesorteerd en gerangschikt op oplopende volgorde van de Manhattan afstand (gelijkzijnde beelden vooraan). Voor de verschillende groep/tags worden gemiddelde precision en recall uitgerekend en in twee grafieken geplot, *prestatie per cluster* en *prestatie per groep-tag*.

Deze systematisch aanpak geeft een goed beeld van de invloed van de parameters op de opbouw van de woordenboeken en de effectiviteit bij het zoeken naar groep-tag combinaties van beelden in Flickr. De bouw van de woordenboeken en bijbehorende databases neemt veel tijd in beslag. Bovendien is de volgorde voor de opbouw van de parameters willekeurig. In dit project ben ik initieel uitgegaan van cluster variatie maar dat had ook

een willekeurig andere parameter kunnen zijn.

De parameter aantal iterations werd niet gewijzigd. Top-surf heeft geen optimale verdeling kunnen vinden onder de 1.000 iterations voor de verdeling van de parameter *nearest neighbors*. Er is verder geen reden om aan te nemen dat een hogere waarde een betere clustering van *nearest neighbors* zou geven. Naast de Manhattan afstand kan Top-surf ook de cosinus afstand gebruiken voor beeldvergelijking. In dit project wordt de cosinusafstand niet gebruikt. De Manhattan afstand is de kortste afstand tussen kenmerken in het beeld.

5.6 Aanpak onderzoek

Bij de opzet van het onderzoek naar de bruikbaarheid van Top-surf bij voor het selecteren van Flickr beelden is zorgvuldig gekeken naar de gevolgen van de gekozen aanpak op de te verwachten resultaten. De keuze om eerst te variëren met het cluster aantal is logisch gelet op het feit dat Top-surf visual word aan een cluster toekent en die wordt opgebouwd uit het gemiddelde van alle "visual words" van de beelden in de cluster.

De voordelen en nadelen van deze aanpak zijn als volgt te noemen:

De voordelen en nadelen van het gekozen aanpak	
Voordelen	gebruikt van precision en recall als standaard methoden voor de performance meting
	gefaseerde parameter variatie
	alle beelden in de onderzochte database van Flickr zijn gelabeld (=grondwaarheden)
	alleen de standaard parameters van TOP-SURF wijzigen (geen inhoudelijke wijzigingen in het algoritme)
Nadelen	opbouw van meerdere woordenboeken en databases in elke fase van het onderzoek
	relatief kleine beeld collectie van Flickr (+/- 25.000)
	label van beelden in Flickr (= grondwaarheid) op basis van menselijk oordeel (subjectief)
	een aantal groep-tag verzamelingen gebruik in het onderzoek hebben een relatief klein aantal beelden

Figuur 4: aanpak onderzoek

5.7 Rol van parameters

Zoals eerder is aangegeven spelen de parameters in Top-surf een zeer belangrijke rol. Hoewel Top-surf aangestuurd wordt door een aantal standaard parameters die te beïnvloeden zijn van buiten af, kent Top-surf interne parameters die ook een rol spelen bij de inrichting van de visual words en woordenboeken. De interne parameters zijn voor de uitvoer van het onderzoek niet gewijzigd.

Parameters in TOP-SURF		initiele waarden	maximale waarden
extern	aantal clusters in de woordenboek	instelbaar (standaard 100)	< aantal features
	aantal nearest neighbors voor het bepalen van cluster centrum	instelbaar	<= aantal beelden in de collectie
	maximaal aantal iterations nodig voor het bepalen en optimaliseren van een cluster centrum	instelbaar (standaard 1000)	onbeperkt
	aantal interesse punten ook wel "features" genoemd	instelbaar (standaard 25)	>= aantal clusters < maximaal aantal features
	afstandsberekening methodiek Manhattan distance of cosinus distance	Manhattan	n.v.t.
intern	aantal beelden relevant voor selectie van clusters (TOP), daar is ook de naam van TOP-SURF aan ontleent	100	n.v.t.
	maximaal aantal beelden gebruikt voor de bouw van de tf-idf tabel	2.500	<= aantal beelden in de collectie
	maximaal aantal interesse punten (features) die TOP-SURF aan kan	33.550.000	33.550.000
	threshold waarde ratie (d1/d2) voor het bepalen van een match tussen twee beelden	0,65	< 1
	element descriptor lengte in bytes	64	onbeperkt

Figuur 5: Overzicht van parameters in Top-surf

6 Experimenten

Voor de uitvoering van het onderzoek is **geen scheiding gemaakt tussen de training dataset en de test dataset**. De volledige beeld-collectie van Flickr is gebruikt voor de opbouw van de woordenboeken alsook voor de analyse van de prestaties. **Dat betekent dat de woordenboeken niet voor een specifieke type groep-tag zijn geoptimaliseerd.**

6.1 Opbouw woordenboeken

Initieel zijn vijf woordenboeken opgebouwd met parameters ingesteld met de standaard waarden van Top-surf, figuur 6. Voor de opbouw van de woordenboeken is de C++ applicatie *imagecla2.exe* ontworpen. De *imagecla2.exe* applicatie roept de Top-surf bibliotheek aan met de ingestelde parameters waarden en genereert een "visual words" woordenboek.

Opbouw dictionaries								
	number of dicts	nearest clusters	number of iterations	number of points	number of images	average time image / sec	elapsed time/hours	elapsed time/sec
1	25	100	1000	25	24582	0,55	3,7	13418
2	100	100	1000	25	24582	0,55	3,8	13552
3	500	100	1000	25	24582	0,59	4,0	14391
4	1000	100	1000	25	24582	0,61	4,2	14967
5	5000	100	1000	25	24581	0,66	4,5	16164
totaal aantal uren							20,1	

Figuur 6: Opgebouwde woordenboeken in fase 1

De blueprint van de "visual word" van een cluster wordt opgebouwd door het gemiddelde te berekenen van alle visual words van naaste burens. Gezien het grote aantal mogelijke parameter combinaties, werd eerst een selectie gemaakt van de te variëren parameters. De woordenboeken van de eerste fase verschillen in het aantal clusters die zij bezitten,

overige parameters zoals het aantal naaste buren voor de clustering, het aantal iteraties voor het bepalen van cluster centers en het aantal punten van interesse zijn vastgezet.

6.2 Opbouw databases

Met de basis database als referentie zijn initieel vijf databases aangemaakt, één voor elk woordenboek-database en cluster combinatie, figuur 7. De woordenboek-database combinatie wordt maar een keer opgebouwd. Eenmaal opgebouwd wordt de woordenboek-database combinatie herhaald gebruikt voor de analyse van de verschillende groep-tag combinaties.

Opbouw databases									
number of dicts	number of clusters	nearest neighbors	number of iterations	number of points	number of images	number of visual-words	average time image / sec	elapsed time/hours	elapsed time/sec
1	25	100	1000	25	24582	323436	0,50	3,4	12367
2	100	100	1000	25	24582	932984	0,50	3,4	12392
3	500	100	1000	25	24582	1756972	1,22	8,3	29937
4	1000	100	1000	25	24582	2008044	1,40	9,6	34428
5	5000	100	1000	25	24581	2151863	1,54	10,5	37871
totaal aantal uren								35,3	

Figuur 7: Opgebouwde databases in fase 1

6.3 Tussentijdse resultaten

Na de opbouw van de woordenboeken en databases met de parameters ingesteld op de standaard waarden en gevarieerd aantal cluster grootte, zijn voor alle 13 geselecteerde groepen/tags, analyses uitgevoerd. Het doel van de analyse is, de identificatie van welke cluster combinaties de beste prestatie resultaten leveren met Top-surf voor de match met Flickr tag beeldcombinaties.

Voor de uitvoer van de beeld analyse is de C++ applicatie *imageclares.exe* ontworpen. Met behulp van deze applicatie wordt een analyse gedaan naar de performance voor alle beelden van een groep. De afstand tussen de beelden wordt berekend, geanalyseerd, gesorteerd en gerangschikt op aflopende afstand waarden, beelden die gelijkenis vertonen met het geanalyseerde beeld vooraan. Op groep/tag niveau worden Precision en Recall gemiddelden berekend.

Imageclares.exe maakt drie bestanden aan met de berekende waarden van random, precision en recall, als volgt:

- Het bestand *im[beeld nummer].txt* bevat de random, precision en recall waarden van het geanalyseerde beeld van een groep-tag combinatie ten opzichte van alle andere beelden in de beeldcollectie.
- Het bestand *pr-file-imageres[cluster x,y,z].txt* bevat het aantal relevante gevonden beelden en de gemeten random, gemiddelde precision en recall waarden voor alle geanalyseerde beelden van de cluster/groep x,y,z.

- Het bestand *prt-imageres-cluster[nr].txt* bevat de totaal random en gemiddelde precision en recall waarden van alle cluster groepen.

D:/data/mirflickr/images1/im1085.jpg						
id	imagename	absolute distance	Relevant	Random	Precision	Recall
0	D:/data/mirflickr/images1/im1085.jpg	0,00000	0			
1	D:/data/mirflickr/images3/im10010.jpg	1,58352				
2	D:/data/mirflickr/images5/im23523.jpg	1,58638				
3	D:/data/mirflickr/images5/im24943.jpg	1,64846				
4	D:/data/mirflickr/images1/im4621.jpg	1,65386				
5	D:/data/mirflickr/images5/im24811.jpg	1,67260				
6	D:/data/mirflickr/images1/im2776.jpg	1,68214				
7	D:/data/mirflickr/images2/im9762.jpg	1,69188				
8	D:/data/mirflickr/images2/im6163.jpg	1,70090				
9	D:/data/mirflickr/images1/im4759.jpg	1,70351				
10	D:/data/mirflickr/images3/im12942.jpg	1,71084				
11	D:/data/mirflickr/images3/im14297.jpg	1,71428				
12	D:/data/mirflickr/images3/im11792.jpg	1,71738				
13	D:/data/mirflickr/images5/im21371.jpg	1,72270				
14	D:/data/mirflickr/images1/im491.jpg	1,72774				
15	D:/data/mirflickr/images4/im18067.jpg	1,73018				
16	D:/data/mirflickr/images1/im4528.jpg	1,73404				
17	D:/data/mirflickr/images4/im16065.jpg	1,73855				
18	D:/data/mirflickr/images5/im23678.jpg	1,74201				
19	D:/data/mirflickr/images1/im2029.jpg	1,74251	1			
totaal			1	0,00082	0,05263	0,05263

Figuur 8: Prestatie waarden voor beeld im1085.jpg

Neem de groep-tag *pr-file-imageres_1-15-23.txt* aan als voorbeeld, figuur 9:

In totaal bevat de deze groep-tag precies 19 beelden. Na berekening van de absolute afstand van het beeld *im1085.jpg* (figuur 8) naar alle andere beelden in de woordenboek/-database werden de beelden gerangschikt op absolute afstand in oplopende volgorde. Het Top-surf algoritme heeft twee beelden gevonden die na sortering tot de eerste 19 beelden behoren: *im1085.jpg* en *im2029.jpg*. 19 is het aantal beelden in de groep-tag. Het geanalyseerde beeld *im1085.jpg*, eerste beeld in het output bestand, is het controle beeld en wordt niet meegenomen voor de berekening van de prestatie (precision en recall). De berekende waarden voor precision en recall $1/19 = 0,05263$ zijn hoger dan random met waarde 0,00082, figuur 8. De prestatie waarde van de algoritme is 64,1 keer beter dan random, Top-surf presteert beter in dit geval.

In de verdere analyse wordt voor alle beelden in een groep-tag combinatie de precision en recall waarden berekend ten opzicht van alle beelden in de Flickr beeldcollectie.

Voorbeeld van precision en recall voor de beelden in cluster groep-tag combinatie 1, 15, 23 en met woordenboek van 1.000 clusters, 100 nearest neighbors, 1.000 iterations en 25 interesse punten, figuur 9.

Bij de analyse van de gemiddelde beeldafstand in deze cluster groep zijn drie positieve gevallen (true positives) gevonden. Top-surf vindt meer dan één match (met uitzondering van eigen beeld) in de geselecteerde groep-tag voor de beelden *im94.jpg*, *im1085.jpg*, *im22742.jpg*. Bovendien is in alle drie gevallen maar één beeld gevonden van de 19 met gelijkwaardige kenmerken.

De berekende precision en recall waarde van $0,00082 / 0,00789 = 9,62$, voor de cluster groep-tag combinatie 1-15-23 (figuur 9) is door de verrichtingen van Top-surf 9,62 keer

pr-file-mageres_1_15_23.txt						
nr	image	tag	relevant	random	precision	recall
1	D:/data/mirflickr/images1/im94.jpg	19	2	0.00082	0.05263	0.05263
2	D:/data/mirflickr/images5/im22235.jpg	19	1	0.00082	0.00000	0.00000
3	D:/data/mirflickr/images5/im24542.jpg	19	1	0.00082	0.00000	0.00000
4	D:/data/mirflickr/images1/im287.jpg	19	1	0.00082	0.00000	0.00000
5	D:/data/mirflickr/images1/im2029.jpg	19	1	0.00082	0.00000	0.00000
6	D:/data/mirflickr/images4/im19198.jpg	19	1	0.00082	0.00000	0.00000
7	D:/data/mirflickr/images1/im1085.jpg	19	2	0.00082	0.05263	0.05263
8	D:/data/mirflickr/images3/im13679.jpg	19	1	0.00082	0.00000	0.00000
9	D:/data/mirflickr/images4/im16984.jpg	19	1	0.00082	0.00000	0.00000
10	D:/data/mirflickr/images5/im22742.jpg	19	2	0.00082	0.05263	0.05263
11	D:/data/mirflickr/images2/im8410.jpg	19	1	0.00082	0.00000	0.00000
12	D:/data/mirflickr/images2/im7649.jpg	19	1	0.00082	0.00000	0.00000
13	D:/data/mirflickr/images4/im15775.jpg	19	1	0.00082	0.00000	0.00000
14	D:/data/mirflickr/images5/im24996.jpg	19	1	0.00082	0.00000	0.00000
15	D:/data/mirflickr/images1/im157.jpg	19	1	0.00082	0.00000	0.00000
16	D:/data/mirflickr/images2/im7473.jpg	19	1	0.00082	0.00000	0.00000
17	D:/data/mirflickr/images4/im19711.jpg	19	1	0.00082	0.00000	0.00000
18	D:/data/mirflickr/images2/im8043.jpg	19	1	0.00082	0.00000	0.00000
19	D:/data/mirflickr/images5/im20052.jpg	19	1	0.00082	0.00000	0.00000
20	D:/data/mirflickr/images5/im20526.jpg	19	1	0.00082	0.00000	0.00000
Totaal					0.00789	0.00789

Figuur 9: Performance waarden voor groep 1, 15 en 23

beter dan een random search.

id	clusters	nearest neighbors	iterations	number points	File-cluster	Random	Precision	Recall	% TOP-SURF beter dan Random	number images in cluster groep
86	5000	100	1000	25	imageres_15_19_23.txt	0,00175	0,02159	0,02159	12,34	43
40	1000	100	1000	25	imageres_1_15_23.txt	0,00082	0,00789	0,00789	9,62	20
7	25	100	1000	25	imageres_13_19_20.txt	0,00200	0,01063	0,01063	5,32	49
47	1000	100	1000	25	imageres_15_19_23.txt	0,00175	0,00831	0,00831	4,75	43
8	25	100	1000	25	imageres_15_19_23.txt	0,00175	0,00831	0,00831	4,75	43
87	5000	100	1000	25	imageres_15_20_23.txt	0,00220	0,00908	0,00908	4,13	54
33	500	100	1000	25	imageres_13_19_20.txt	0,00200	0,00765	0,00765	3,83	49
20	100	100	1000	25	imageres_13_19_20.txt	0,00200	0,00723	0,00723	3,62	49
34	500	100	1000	25	imageres_15_19_23.txt	0,00175	0,00609	0,00609	3,48	43
27	500	100	1000	25	imageres_1_15_23.txt	0,00082	0,00263	0,00263	3,21	20
14	100	100	1000	25	imageres_1_15_23.txt	0,00082	0,00263	0,00263	3,21	20
45	1000	100	1000	25	imageres_12_14_20.txt	0,00346	0,01106	0,01106	3,20	85
48	1000	100	1000	25	imageres_15_20_23.txt	0,00220	0,00699	0,00699	3,18	54
84	5000	100	1000	25	imageres_12_14_20.txt	0,00346	0,01092	0,01092	3,16	85

Figuur 10: Prestaties woordenboeken met de beste resultaten in fase 1

id	clusters	nearest neighbors	iterations	number points	File-cluster	Random	Precision	Recall	% TOP-SURF beter dan Random	number images in cluster groep
1	1000	100	1000	25	imageres_1_6_10.txt	0,00273	0,00452	0,00452	1,66	67
2	1000	100	1000	25	imageres_1_6_20.txt	0,00086	0,00000	0,00000	-1,00	21
3	1000	100	1000	25	imageres_1_15_23.txt	0,00082	0,00789	0,00789	9,62	20
4	1000	100	1000	25	imageres_5_19_20.txt	0,00289	0,00121	0,00121	-0,58	71
5	1000	100	1000	25	imageres_7_10_14.txt	0,00668	0,00382	0,00382	-0,43	164
6	1000	100	1000	25	imageres_7_14_16.txt	0,00383	0,00046	0,00046	-0,88	94
7	1000	100	1000	25	imageres_7_14_20.txt	0,00253	0,00264	0,00264	1,04	62
8	1000	100	1000	25	imageres_10_12_14.txt	0,00542	0,00319	0,00319	-0,41	133
9	1000	100	1000	25	imageres_12_14_16.txt	0,00159	0,00067	0,00067	-0,58	39
10	1000	100	1000	25	imageres_12_14_20.txt	0,00346	0,01106	0,01106	3,20	85
11	1000	100	1000	25	imageres_13_19_20.txt	0,00200	0,00595	0,00595	2,98	49
12	1000	100	1000	25	imageres_15_19_23.txt	0,00175	0,00831	0,00831	4,75	43
13	1000	100	1000	25	imageres_15_20_23.txt	0,00220	0,00699	0,00699	3,18	54

Figuur 11: Performance waarden voor woordenboek met 1.000 clusters

6.4 Resultaten fase 1

De analyse naar de beste prestaties van Top-surf geeft geen duidelijke voorrang aan voor een of ander woordenboek. Bij de eerste top drie prestaties komen wij woordenboeken tegen met 5.000, 1.000 en zelfs 25 clusters. Het beste match resultaat wordt bereikt met een woordenboek van 5.000 clusters op de groep-tag combinatie *plant_life*, *sky* en *tree* (tags 15, 19 en 23), voor deze groep-tag combinatie is Top-surf 12,34 beter dan random search. Ook voor groep-tag combinaties van *animals*, *plant_life* en *tree* (tags 1, 15 en 23) met een woordenboek van 1.000 clusters is Top-surf 9,62 beter dan random search.

Top-surf presteert beter voor bepaalde groep-tags combinaties dan anderen. De groep-tag *plant_life*, *sky* en *tree* komt vier keer voor bij de eerste 10 beste prestaties, zie figuur 10, terwijl de groep-tag *animals*, *dogs* en *structures* (tags 1, 6 en 20) slecht scoort met Top-surf, zie figuur 12.

De prestatie resultaten voor een specifiek woordenboek zijn niet eenduidig. Neem een woordenboek met 1.000 clusters als voorbeeld dan presteert Top-surf 4,75 keer beter dan random search als de beste prestatie voor de groep-tag *plant_life*, *sky* en *tree* en het slechtste presteert voor de groep-tag *animals*, *dogs* en *structures* waarin geen enkele match is gevonden, zie figuur 11

Het is interessant om op te merken dat meerdere combinaties met de tags *plant_life* (15) en *tree* (23) vooraan zitten in de prestatie meting. De genoemde tags presteren goed in combinatie met de tags *sky* (19), *animals* (1) en *structures* (20) ongeacht het aantal clusters in het woordenboek, zie figuur 10.

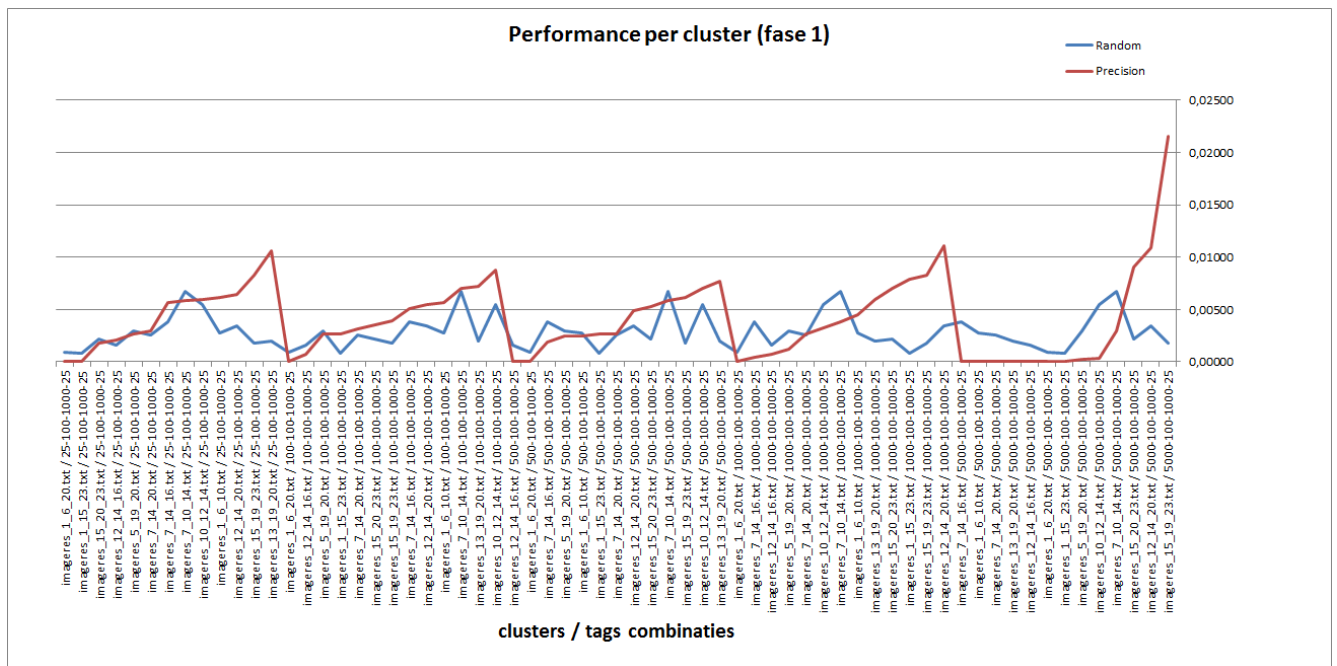
Woordenboeken met groep-tag combinaties van *animals*, *dogs* en *structures* (tags 1, 6 en 20) leveren een negatieve score met Top-surf (100% slechter dan random search), figuur 12. Verder zien wij dat een woordenboek met 5.000 clusters zowel de beste als de slechtste prestaties levert. Wij kunnen hier voorzichtig concluderen dat de omvang van de cluster geen grote impact heeft op de prestaties van Top-surf en dat Top-surf beter presteert voor

een bepaald soort goep tags dan anderen.

Clusters-Neighbors-Iterations-Points	File-precision-per-cluster	Random	Precision	Recall	number of times Topsurf beter dan random	CL1	CL2	CL3	CL1- Name	CL2- Name	CL3-Name
5000-100-1000-25	imageres_1_15_23.txt	0,00082	0,00000	0,00082	-1,00	1	15	23	animals	plant_life	tree
5000-100-1000-25	imageres_1_6_20.txt	0,00086	0,00000	0,00086	-1,00	1	6	20	animals	dog	structures
5000-100-1000-25	imageres_1_6_10.txt	0,00273	0,00000	0,00273	-1,00	1	6	10	animals	dog	indoor
1000-100-1000-25	imageres_1_6_20.txt	0,00086	0,00000	0,00086	-1,00	1	6	20	animals	dog	structures
500-100-1000-25	imageres_12_14_16.txt	0,00159	0,00000	0,00159	-1,00	12	14	16	male	people	portrait
500-100-1000-25	imageres_1_6_20.txt	0,00086	0,00000	0,00086	-1,00	1	6	20	animals	dog	structures
100-100-1000-25	imageres_1_6_20.txt	0,00086	0,00000	0,00086	-1,00	1	6	20	animals	dog	structures
25-100-1000-25	imageres_1_15_23.txt	0,00082	0,00000	0,00082	-1,00	1	15	23	animals	plant_life	tree
25-100-1000-25	imageres_1_6_20.txt	0,00086	0,00000	0,00086	-1,00	1	6	20	animals	dog	structures

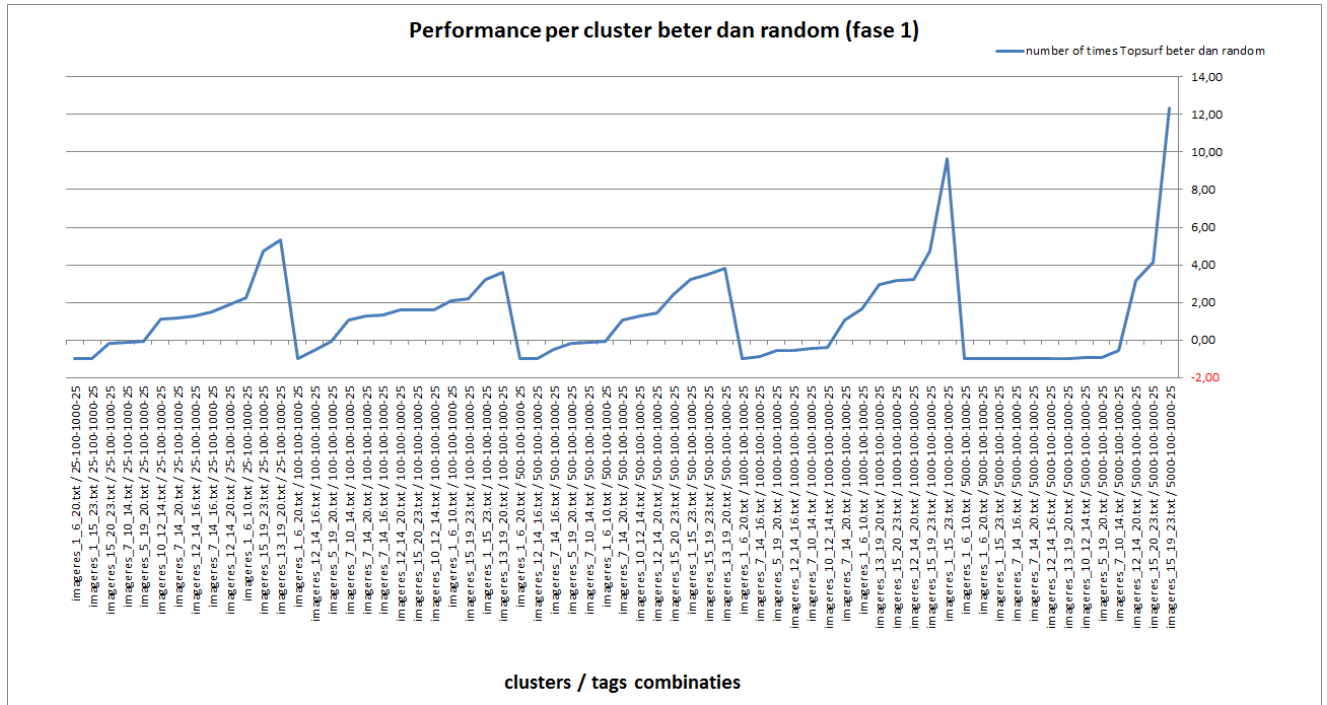
Figuur 12: Prestatie cluster/tags slechter dan random

In de grafieken "Prestatie per cluster", figuur 13 en "Prestatie per cluster beter dan random", laat figuur 14 zien dat de woordenboeken gelijkwaardige resultaten tonen voor de geselecteerde groep-tag combinaties. Er zijn twee uitschieters voor de woordenboeken met 1.000 en 5.000 clusters waar de prestaties van Top-surf aanzienlijk beter zijn dan random search maar nog steeds te laag voor het zoeken en matchen van relevante beelden in de Flickr beelden database.



Figuur 13: Prestaties per cluster (fase1 1)

Zoals eerder is aangegeven zien wij dat bepaalde groep-tag combinaties terugkeren in de beste prestaties per cluster waarbij wij de indruk krijgen dat de cluster "grootte" weinig



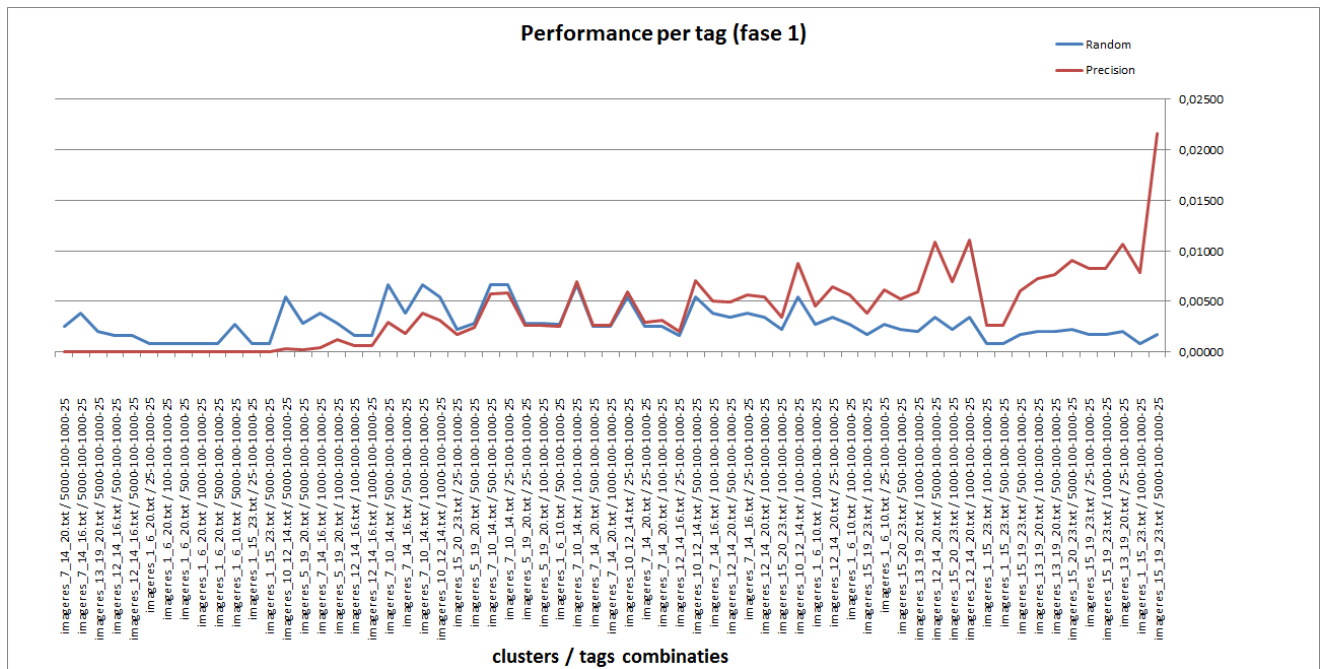
Figuur 14: Prestatie per cluster beter dan random (fase 1)

tot geen invloed heeft op de prestaties, zie figuur 16. Top-surf presteert het best in combinaties waar *plant_life en tree* (tags 15 en 23) in voorkomen, zie figuur 15.

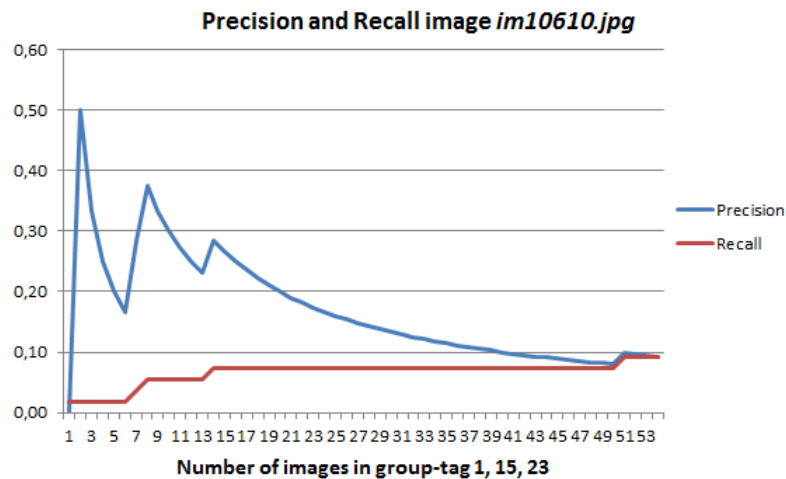
clusters	File-cluster	% TOP-SURF beter dan Random	Rank	File-cluster	% TOP-SURF beter dan Random	Rank	File-cluster	% TOP-SURF beter dan Random	Rank
25	imageres_13_19_20.txt	5,32	3	imageres_15_19_23.txt	4,75	4	imageres_1_6_10.txt	2,24	12
100	imageres_13_19_20.txt	3,62	7	imageres_1_15_23.txt	3,21	9	imageres_15_19_23.txt	2,22	13
500	imageres_13_19_20.txt	3,83	6	imageres_15_19_23.txt	3,48	8	imageres_1_15_23.txt	3,21	9
1000	imageres_1_15_23.txt	9,62	2	imageres_15_19_23.txt	4,75	4	imageres_12_14_20.txt	3,20	10
5000	imageres_15_19_23.txt	12,34	1	imageres_15_20_23.txt	4,13	5	imageres_12_14_20.txt	3,16	11

Figuur 15: Best prestaties per cluster in fase 1

De resultaten van de verschillende groep-tags t.o.v. het ideale systeem (Precision = Recall) laten geen optimaal systeem zien. In het voorbeeld van de berekende gemiddelde precision en recall voor image *im10610.jpg* is geen optimaal kruispunt gevonden. Dat wil niet zeggen dat er geen kruispunt is voor andere groep-tag combinaties maar dit voorbeeld toont aan dat een optimaal kruispunt niet altijd te vinden is, figuur 17.



Figuur 16: prestatie per groep-tag (fase1)



Figuur 17: voorbeeld: Precision = Recall voor image im10610.jpg

6.5 Opbouw woordenboeken en database - fase 2

In de tweede fase van het project zijn een aantal nieuwe woordenboeken en de bijbehorende database aangemaakt. Omdat er geen duidelijke aanwijzing was om één specifieke parameter te variëren ten opzichte van alle anderen, heb ik er voor gekozen om het aantal naaste burens en interesse punten te variëren. Op deze wijze zijn vijf nieuwe woordenboeken en databases aangemaakt.

Er zijn drie woordenboeken aangemaakt met de parameter vaste aantal clusters op 1.000. Voor de drie woordenboeken is gevarieerd met de parameters naaste burens op 200, interesse punten op 50, voor het laatste woordenboek zijn de parameters naaste burens en interesse punten op respectievelijk 50 en 75 gezet. Voor het vierde woordenboek is de parameter aantal interesse punten op 50 gezet voor een woordenboek met 5.000 clusters.

Opbouw dictionaries									
	number of	nearest		number of	number of	aantal	average time	elapsed	elapsed
dicts	clusters	neighbors	iterations	points	images	visualwords	image / sec	time/hours	time/sec
1	1000	50	1000	75	24582		0,60	4,1	14749
2	1000	100	1000	50	24582		0,55	3,8	13520
3	1000	200	1000	25	24582		0,75	5,1	18437
4	5000	100	1000	50	24582		0,85	5,8	20895
totaal aantal uren								18,8	

Figuur 18: Opgebouwde woordenboeken in fase 2

Opbouw databases									
	number of	nearest		number of	number of	number of	average time	elapsed	elapsed
dicts	clusters	neighbors	iterations	points	images	visual-words	image / sec	time/hours	time/sec
1	1000	50	1000	75	24582	1846255	0,71	4,8	17453
2	1000	100	1000	50	24582	1910875	0,66	4,5	16224
3	1000	200	1000	25	24582	2068995	1,50	10,2	36873
4	5000	100	1000	50	24582	2146528	1,80	12,3	44248
totaal aantal uren								31,9	

Figuur 19: Opgebouwde databases in fase 2

6.6 Resultaten van fase 2

Na uitvoer van de analyse van de beeldcollectie op de nieuwe aangemaakte woordenboeken valt het volgende op:

- Beide woordenboeken met de parameter combinatie (cluster, naaste-burens, iteraties, interesse punten) 5000-100-1000-25 en 5000-100-1000-50 geven hoge prestatie scores respectievelijk 12,34 en 11,39 op de tag combinatie *plant life*, *sky* en *tree* (tags 15, 19 en 23), zie figuur 20.

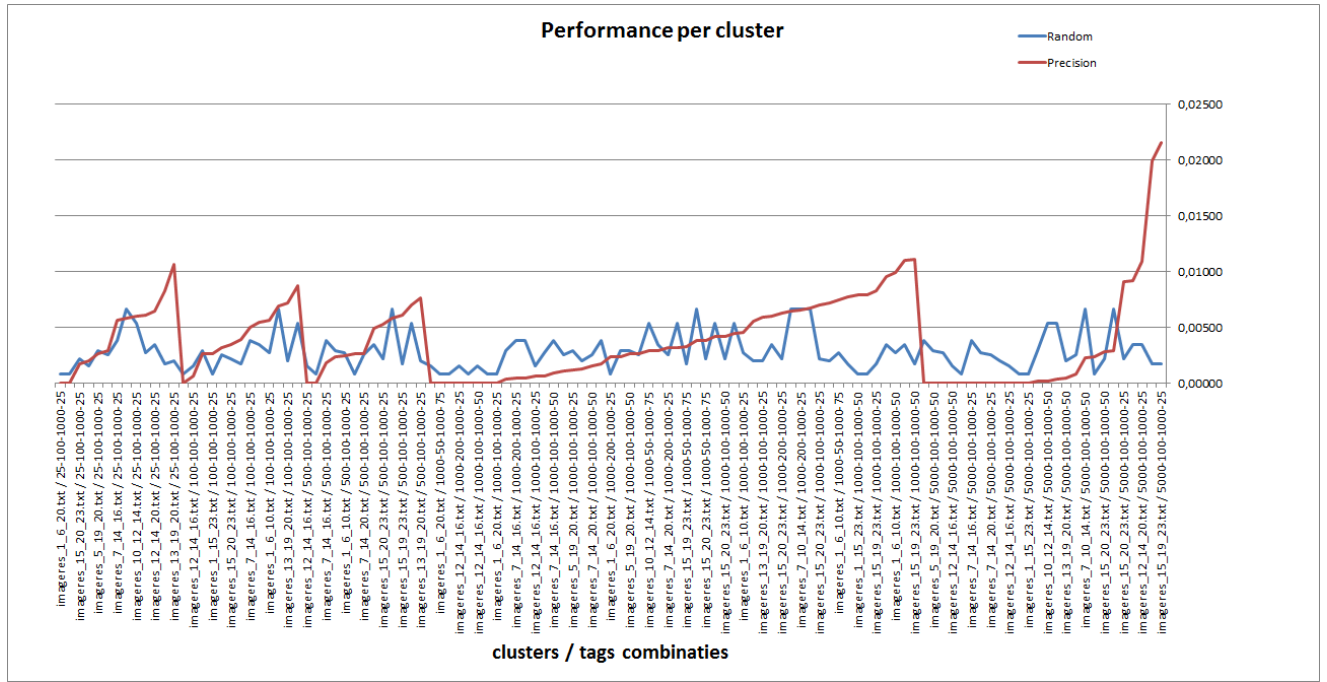
- Ook de woordenboeken met de parameter combinaties 1000-100-1000-25 en 1000-100-1000-50 geven hetzelfde prestatie score op de tag combinatie *animals, plant_life en tree* (tags 1, 15 en 23), zie figuur 20.

Het aantal interesse punten heeft een geringe invloed op de prestaties van Top-surf. De aangemaakte woordenboeken met 75 interesse punten en 200 naaste burens komen tijdens de analyse niet voor in de hogere prestatie ranking. Blijkbaar is de impact van deze twee parameters op de prestatie van Top-surf niet groot. Er is weinig verschil in de prestaties van een woordenboek met 5.000 clusters en 25 *interesse punten* (score 12,34) en een woordenboek met 5.000 clusters en 50 *interesse punten* (score 11,39), bovendien er is helemaal geen verschil in een woordenboek met 1.000 clusters en 25 *interesse punten* (score 9,62) met een woordenboek met 1.000 clusters en 50 *interesse punten* (score 9,62), figuur 20.

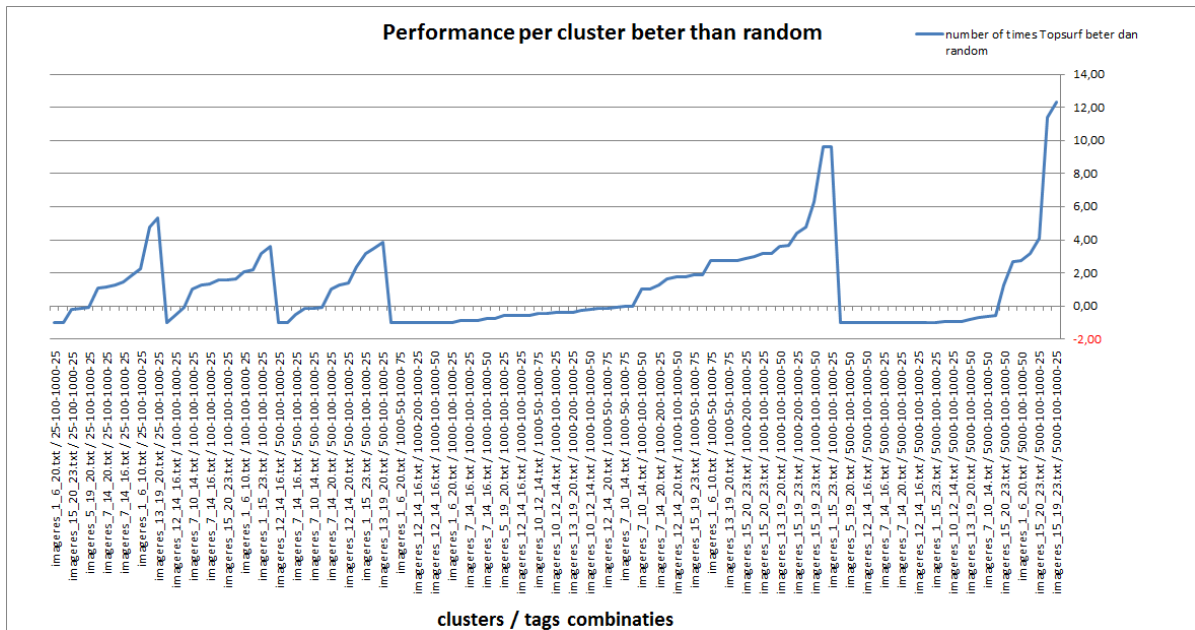
Clusters-Neighbors-Iterations-Points	File-precision-per-cluster	Random	Precision	Recall	number of times Topsurf beter dan random	CL1	CL2	CL3	CL1-Name	CL2-Name	CL3-Name
5000-100-1000-25	imageres_15_19_23.txt	0,00175	0,02159	0,02159	12,34	15	19	23	plant_life	sky	tree
5000-100-1000-50	imageres_15_19_23.txt	0,00175	0,01993	0,01993	11,39	15	19	23	plant_life	sky	tree
1000-100-1000-25	imageres_1_15_23.txt	0,00082	0,00789	0,00789	9,62	1	15	23	animals	plant_life	tree
1000-100-1000-50	imageres_1_15_23.txt	0,00082	0,00789	0,00789	9,62	1	15	23	animals	plant_life	tree
1000-100-1000-50	imageres_15_19_23.txt	0,00175	0,01107	0,01107	6,33	15	19	23	plant_life	sky	tree
25-100-1000-25	imageres_13_19_20.txt	0,00200	0,01063	0,01063	5,32	13	19	20	night	sky	structures

Figuur 20: Prestatie top 5 cluster-tags beter dan random (fase 2)

In de grafiek "prestatie per cluster over fase 1 en 2", figuren 21 en 22, wordt duidelijk zichtbaar dat de woordenboeken van 1.000 en 5.000 clusters met gewijzigde parameters geen wezelijke verbetering geven van de Top-surf prestaties, hoewel een kleine verbetering aanwezig is in de prestatie van het woordenboek met 1.000 clusters en 50 interesse punten. De verschillen in prestatie tussen random search en Top-surf zijn minimaal voor de verschillende woordenboeken. De grootte van de clusters hebben geringe invloed bij het vinden van gelijke beelden (matching). Echter voor bepaalde groep-tag combinaties levert Top-surf betere resultaten op dan random search.



Figuur 21: Prestatie per cluster (fase 1 en 2)



Figuur 22: prestatie per cluster beter dan random (fase 1 en 2)

6.7 Eerste opzet werkend via database queries

Bij de initiële opzet van het onderzoek werden de descriptors een voor een uit de database gelezen. Dit proces neemt veel tijd in beslag, er wordt een query op de database uitgevoerd elke keer dat een beeld nodig is voor de berekening van de afstand tussen de beelden. Analyse van een groep beelden zoals bijvoorbeeld *imageres_10_12_14.txt* met

133 beelden duurt enkele uren. Als wij bedenken dat voor elk beeld in de verzameling van de geselecteerde groepen de afstand wordt berekend naar alle beelden in de Flickr beeld collectie en dat voor elk woordenboek, dan kunnen we alleen concluderen dat deze methodiek tijdrovend en inefficiënt is.

Het aantal queries op de database wordt geschat op 22,5 miljoen, uitgaande van een totaal van 902 te matchen beelden in de geselecteerde bestanden en een database met 25.000 beelden zou de analyse van een fase in totaal 150 uur duren.

6.8 Versnelde opzet middels main memory queries

Om het analyse proces te versnellen is er voor gekozen om de descriptor en de visual words informatie van alle beelden in de database te laden in het geheugen van de computer. De descriptor en visual words identifiers van alle beelden in de database zijn initieel in een array op de heap geladen. Door middel van een verwijzing naar het array worden de beelddescriptors opgehaald die nodig zijn voor de berekening van de afstand tussen de beelden. De verwerking en analyse van 902 beelden van 13 groep-tags en 9 woordenboeken duurt 1,44 uur. Dat is 515 keer sneller dan door het gebruik van database-queries.

Deze aanpak levert enorme tijds winst op, de analyse van alle 13 groep-tags op alle woordenboeken neemt ongeveer 45 minuten in beslag op een ACER Aspire PC met een AMD Athlon 64x2 dual core 32 bits processor van 2,1 Ghz en 4GB interne geheugen. De keerzijde van deze aanpak is dat het aantal beelden dat in het geheugen kan worden geladen begrensd wordt door het beschikbare geheugen.

```
// images information stored in memory
struct _imagedesc {
    int idImage;                // idImage
    char imagename[128];        // imagename
    int identifier[MAX_CLUSTER_QTY]; // topsurf identifiers (clusters) array
    float tf[MAX_CLUSTER_QTY];  // term frequency
    int descriptor_count;        // descriptor count
    float descriptor_length;     // for cosinus comparisation
    int clustergt_count;         // number of gt clusters
    int array_clusters[MAX_CLUSTER_GT_QTY]; // ground truth clusters
    float distance_absolute;     // calculated absolute distance
    float distance_cosinus;     // calculated cosinus distance
};
```

De beeldafstand berekening routines zijn embedded in de Top-surf bibliotheek en maken gebruik van de interne technische Top-surf omgeving. Omdat de beeld informatie uit de database wordt gehaald en geen gebruik maakt van de Top-surf API, zijn de routines *CompareDescriptorsAbsolute()* en *CompareDescriptorsCosine()* herschreven voor het gebruik van de array en berekening van de beeldafstand.

6.9 Multiple tag vergelijkingen

Voor het onderzoek zijn groepen geselecteerd met precies 3 tags. De analyse van resultaten voor beelden met meer dan 3 tags wordt lastig om te achterhalen wat de invloed

is van de gecombineerde tags op de prestatie van Top-surf voor de specifieke groep beelden.

CL1	CL1-name	CL2	CL1-name	CL3	CL1-name
15	plant_life	19	sky	23	tree
15	plant_life	20	structures	23	tree
12	male	14	people	20	structures
7	female	10	indoor	14	people
5	clouds	19	sky	20	structures
10	indoor	12	male	14	people
13	night	19	sky	20	structures
12	male	14	people	16	portrait
7	female	14	people	20	structures
7	female	14	people	16	portrait
1	animals	15	plant_life	23	tree
1	animals	6	dog	20	structures
1	animals	6	dog	10	indoor

Figuur 23: Groep-tag cmbinaties geselecteerd voor het onderzoek

6.10 Single tag vergelijkingen

Tot nu toe is de evaluatie van Top-surf prestaties uitgevoerd zonder specifieke training op een of meerdere groep-tag combinaties middels een deel van de test set.

Om de invloed van specifieke training binnen Top-surf te onderzoeken, heb ik twee woordenboeken aangemaakt. Het eerste woordenboek is aangemaakt zonder het gebruik van training set en de tweede met het gebruik van training set.

Voor de analyse van de resultaten heb ik voor de groep-tag (*animals*) gekozen met 197 beelden met precies 1 tag. De groep-tag is in twee gelijke helften verdeeld, de eerste helft (99 beelden) is als training set gebruikt voor de opbouw van het woordenboek, de tweede helft (98 beelden), is als test set meegenomen in de collectie van beelden voor de analyse van de resultaten.

De analyse is uitgevoerd op de twee genoemde woordenboeken met identieke parameters, alleen het aantal naaste burens verschilt, zie figuur 24.

woordenboek	clusters	nn	iterations	points	random	precision	recall	prestatie
standaard	25	100	1000	25	0,00403	0,00866	0,00866	2,15
getraind	25	50	1000	25	0,00405	0,00495	0,00495	1,22

Figuur 24: prestatie standarad en getraind woordenboek op groep-tag animals

Het getrainde woordenboek levert met een score van 1,22 geen beter resultaat dan de niet getrainde met een score van 2,15. De verschil in prestaties is relatief klein maar het geeft wel aan dat een getraind Top-surf woordenboek niet beter presteert dan een niet getraind.

7 Conclusies

Top-surf is een zeer uitgebreide *high performance* toolkit voor visual words, ontworpen voor inzet in een breed scala toepassingen. Deze flexibiliteit en brede inzet van Top-surf heeft een prijs. Top-surf presteert beter voor bepaalde groep-tag combinaties dan bij anderen ongeacht welk woordenboek wordt toegepast. De omvang van de cluster grootte en andere parameters van Top-surf hebben een geringe impact op de prestaties.

De prestaties van Top-surf zijn wisselvallig. Voor sommige groep-tags levert Top-surf een betere prestatie dan random search, maar niet voor alle groep-tags. De samenstelling van het woordenboek is hierin ook niet bepalend. Wij hebben gezien dat de impact van bijvoorbeeld het aantal *interesse punten* (features) of het aantal *naaste burens* vrijwel geen invloed heeft op de prestaties. Groep-tag combinaties met *plant_life en tree* scoren beter dan random search, terwijl groep-tag combinaties met *animals, dogs en structures* scoren duidelijk slechter presteren dan random search.

Ook voor combinaties waarin Top-surf het best scoort, is het aantal gevonden beelden van Flickr op basis van de grondwaarheden hooguit 2%, zie figuur 30.

In het voorbeeld van een getraind woordenboek met precies 1 groep-tag is de prestatie van Top-surf niet beter dan random, zie figuur 24. De verkregen scores zijn te laag om Top-surf te kunnen gebruiken voor het zoeken van Flickr tag(s) op basis van de grondwaarheden.

7.1 Aanbevelingen en suggesties voor vervolg onderzoek

Het uitgevoerde onderzoek naar de bruikbaarheid van Top-surf voor het vinden van gelijkwaardige beelden is kleinschallig en beperkt. Een dieper en uitgebreider onderzoek naar Top-surf zal ongetwijfeld een beter inzicht geven in de mogelijkheden en onmogelijkheden van de Top-surf toolkit.

Hierbij een overzicht van de aanbevelingen voor een vervolg onderzoek:

- Onderzoek uitvoeren op een uitgebreide beeld collectie van bijvoorbeeld 1.000.000 beelden gelabeld met grondwaarheden. Het onderzoek naar de bruikbaarheid van Top-surf voor het vinden van gelijkwaardige beelden is uitgevoerd op een relatief kleine beeldcollectie van 25.000 beelden.
- Uitgebreid gebruik van de externe parameter opties in de verschillende fases van het onderzoek. Alle parameter mogelijkheden onderzoeken neemt veel tijd in beslag, bovendien zijn de beschikbare middelen (computer power) beperkt.
- Onderzoeken van de impact van de interne parameters op de prestaties. Bijvoorbeeld door het inperken of uitbreiden van de lengte van de visual words of door het aantal beelden voor de tf-idf waarden te wijzigen.
- Onderzoeken van het gebruik van de cosinus afstand voor de beeldafstand berekening. Wat is dan de invloed op de prestaties.
- Onderzoeken doen naar andere beeld collecties. Zijn de prestaties beter of slechter en wat kunnen wij daaruit concluderen.

- Dieper onderzoek naar beeldcollecties van een specifieke groep-tag. Uit het onderzoek is gebleken dat Top-surf beter presteert voor bepaalde groep-tag combinaties.
- Top-surf is gebaseerd op de visual words aanpak van Surf en histogram gelateerd. Onderzoek doen naar combinatie van Top-surf met andere beeld herkenning algoritmen zoals *shape recognition*, *neurale netwerken* en *genetisch algoritmen* .
- Onderzoek van de impact van de getrainde en niet getrainde groep-tag combinaties op de prestatie van Top-surf.

8 Dankbetuiging

Ik zou graag mijn dankbaarheid willen uiten aan Dr. Nies Huijsmans voor zijn begeleiding, supervisie en geleverde feedback. Zijn kennis, inzicht en adviezen hebben mij enorm geholpen tijdens het onderzoek.

Referenties

- [1] Bart Thomée, Erwin M. Bakker, Michael S. Lew. Top-surf: A Visual Words Toolkit, 2010.
- [2] Bart Thomée. A Picture is worth a thousand word, Content-based image retrieval techniques, pages 95-106, 2010.
- [3] BSD 3-Clause License. <http://opensource.org/licenses/BSD-3-Clause>
- [4] ZLIB License. http://www.gzip.org/zlib/zlib_license.html
- [5] Flickr images database. www.flickr.com.
- [6] The MIR Flickr Retrieval Evaluation. MirFlickr-25K database. Collection of 25.000 database images including image tags. <http://press.liacs.nl/mirflickr/dlform.php>.
- [7] M. J. Huiskes, M. S. Lew (2008). The MIR Flickr Retrieval Evaluation. ACM International Conference on Multimedia Information Retrieval (MIR'08), Vancouver, Canada (bib) <http://creativecommons.org/> http://press.liacs.nl/mirflickr/#sec_extension
- [8] Mark Huiskes, Bart Thomée, Michael Lew. The MIRFLICKR Retrieval Evaluation. <http://press.liacs.nl/mirflickr/>
- [9] Top-surf library Version 1.4.2 - 11 June 2011. <http://press.liacs.nl/researchdownloads/topsurf>
- [10] Chris Evans. The OpenSURF Computer Vision Library. <http://www.chrisevansdev.com/computer-vision-opensurf.html>
- [11] OpenCV is released under a BSD license. <http://opencv.willowgarage.com/wiki/>
- [12] FLANN - Fast Library for Approximate Nearest Neighbors. <http://people.cs.ubc.ca/~mariusm/index.php/FLANN/FLANN>
- [13] C++ image processing and conversion library. <http://www.xdp.it/cximage.htm>
- [14] Microsoft Visual C++ 2010 Express Compiler. <http://www.microsoft.com/visualstudio/en-us/products/2010-editions/visual-cpp-express>
- [15] MySQL. MySQL database Version 5.5 DPL License, 2012, <http://www.mysql.com>
- [16] Comprehensive Perl Archive Network. <http://www.cpan.org>

- [17] Dionysius P. Huijsmans and Nicu Sebe, Member IEEE How to Complete Performance Graphs in Content-Based Image retrieval: Add Generality and Normalize Scope. <http://doi.ieeecomputersociety.org/10.1109/TPAMI.2005.30>
- [18] Bay, H., Ess, A., Tuytelaars, T., and van Gool, L. 2008. Speed-up robust features (SURF). *Computer Vision and Image Understanding*, 110(3), 346-359.
- [19] Philbin, J., Chum, O., Isard, M., Sivic, J., and Zisserman, A. 2007. Object retrieval with large vocabularies and fast spatial matching. In *Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition*, 1-8.
- [20] Muja, M., and Lowe, D.G. 2009. Fast approximate nearest neighbors with automatic algorithm configuration, in *Proceedings of the 2009 International Conference on Computer Vision Theory and Applications*, 331-340.

Bijlagen

Voorbeeld beelden van MIR Flickr data set gebruikt in het onderzoek.



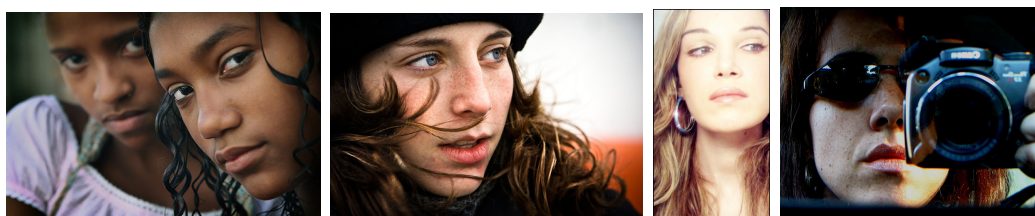
Figuur 25: Beelden groep-tag animals (1)



Figuur 26: Beelden groep-tag animal, plant_life, tree (1, 15 en 23)



Figuur 27: Beelden groep-tag plant_life, sky, tree (15, 19 en 23)



Figuur 28: Beelden groep-tag female, people, portrait (7, 14 en 16)

fase	id	Clusters	Neighbors	Iterations	Points	File-precision-per-cluster	Random	Precision	number images	Precision - Random	number of times Topsurf better dan	CL1	CL2	CL3	CL1-Name	CL2-Name	CL3-Name
1	86	5000	100	1000	25	imageres_15_19_23.txt	0,00175	0,02159	43	0,01984	12,34	15	19	23	plant_life	sky	tree
2	111	5000	100	1000	50	imageres_15_19_23.txt	0,00175	0,01993	43	0,01818	11,39	15	19	23	plant_life	sky	tree
1	40	1000	100	1000	25	imageres_1_15_23.txt	0,00082	0,00789	20	0,00707	9,62	1	15	23	animals	plant_life	tree
2	53	1000	100	1000	50	imageres_1_15_23.txt	0,00082	0,00789	20	0,00707	9,62	1	15	23	animals	plant_life	tree
2	60	1000	100	1000	50	imageres_15_19_23.txt	0,00175	0,01107	43	0,00932	6,33	15	19	23	plant_life	sky	tree
1	7	25	100	1000	25	imageres_13_19_20.txt	0,00200	0,01063	49	0,00863	5,32	13	19	20	night	sky	structures
1	47	1000	100	1000	25	imageres_15_19_23.txt	0,00175	0,00831	43	0,00656	4,75	15	19	23	plant_life	sky	tree
1	8	25	100	1000	25	imageres_15_19_23.txt	0,00175	0,00831	43	0,00656	4,75	15	19	23	plant_life	sky	tree
2	73	1000	200	1000	25	imageres_15_19_23.txt	0,00175	0,00775	43	0,00600	4,43	15	19	23	plant_life	sky	tree
1	87	5000	100	1000	25	imageres_15_20_23.txt	0,00220	0,00908	54	0,00688	4,13	15	20	23	plant_life	structures	tree
1	33	500	100	1000	25	imageres_13_19_20.txt	0,00200	0,00765	49	0,00565	3,83	13	19	20	night	sky	structures
2	54	1000	100	1000	50	imageres_1_6_10.txt	0,00273	0,00995	67	0,00722	3,64	1	6	10	animals	dog	structures
1	20	100	100	1000	25	imageres_13_19_20.txt	0,00200	0,00723	49	0,00523	3,62	13	19	20	night	sky	structures
2	59	1000	100	1000	50	imageres_13_19_20.txt	0,00200	0,00723	49	0,00523	3,62	13	19	20	night	sky	structures
1	34	500	100	1000	25	imageres_15_19_23.txt	0,00175	0,00609	43	0,00434	3,48	15	19	23	plant_life	sky	tree
1	27	500	100	1000	25	imageres_1_15_23.txt	0,00082	0,00263	20	0,00181	3,21	1	15	23	animals	plant_life	tree
1	14	100	100	1000	25	imageres_1_15_23.txt	0,00082	0,00263	20	0,00181	3,21	1	15	23	animals	plant_life	tree
1	45	1000	100	1000	25	imageres_12_14_20.txt	0,00346	0,01106	85	0,00760	3,20	12	14	20	male	people	structures
1	48	1000	100	1000	25	imageres_15_20_23.txt	0,00220	0,00699	54	0,00479	3,18	15	20	23	plant_life	structures	tree
1	84	5000	100	1000	25	imageres_12_14_20.txt	0,00346	0,01092	85	0,00746	3,16	12	14	20	male	people	structures
1	46	1000	100	1000	25	imageres_13_19_20.txt	0,00200	0,00595	49	0,00395	2,98	13	19	20	night	sky	structures
2	74	1000	200	1000	25	imageres_15_20_23.txt	0,00220	0,00629	54	0,00409	2,86	15	20	23	plant_life	structures	tree
2	106	5000	100	1000	50	imageres_1_6_20.txt	0,00086	0,00238	21	0,00152	2,77	1	6	20	animals	dog	structures
2	68	1000	200	1000	25	imageres_1_6_20.txt	0,00086	0,00238	21	0,00152	2,77	1	6	20	animals	dog	structures
2	97	1000	50	1000	75	imageres_13_19_20.txt	0,00200	0,00553	49	0,00353	2,77	13	19	20	night	sky	structures
2	71	1000	200	1000	25	imageres_12_14_20.txt	0,00346	0,00952	85	0,00606	2,75	12	14	20	male	people	structures
2	92	1000	50	1000	75	imageres_1_6_10.txt	0,00273	0,00746	67	0,00473	2,73	1	6	10	animals	dog	indoor
2	109	5000	100	1000	50	imageres_12_14_20.txt	0,00346	0,00924	85	0,00578	2,67	12	14	20	male	people	structures
1	35	500	100	1000	25	imageres_15_20_23.txt	0,00220	0,00524	54	0,00304	2,38	15	20	23	plant_life	structures	tree
1	2	25	100	1000	25	imageres_1_6_10.txt	0,00273	0,00611	67	0,00338	2,24	1	6	10	animals	dog	indoor
1	21	100	100	1000	25	imageres_15_19_23.txt	0,00175	0,00388	43	0,00213	2,22	15	19	23	plant_life	sky	tree
1	15	100	100	1000	25	imageres_1_6_10.txt	0,00273	0,00565	67	0,00292	2,07	1	6	10	animals	dog	indoor
2	61	1000	100	1000	50	imageres_15_20_23.txt	0,00220	0,00419	54	0,00199	1,90	15	20	23	plant_life	structures	tree
2	98	1000	50	1000	75	imageres_15_19_23.txt	0,00175	0,00332	43	0,00157	1,90	15	19	23	plant_life	sky	tree
1	6	25	100	1000	25	imageres_12_14_20.txt	0,00346	0,00644	85	0,00298	1,86	12	14	20	male	people	structures
2	99	1000	50	1000	75	imageres_15_20_23.txt	0,00220	0,00384	54	0,00164	1,75	15	20	23	plant_life	structures	tree
2	58	1000	100	1000	50	imageres_12_14_20.txt	0,00346	0,00602	85	0,00256	1,74	12	14	20	male	people	structures
1	41	1000	100	1000	25	imageres_1_6_10.txt	0,00273	0,00452	67	0,00179	1,66	1	6	10	animals	dog	indoor
1	17	100	100	1000	25	imageres_10_12_14.txt	0,00542	0,00877	133	0,00335	1,62	10	12	14	indoor	male	people
1	22	100	100	1000	25	imageres_15_20_23.txt	0,00220	0,00349	54	0,00129	1,59	15	20	23	plant_life	structures	tree
1	19	100	100	1000	25	imageres_12_14_20.txt	0,00346	0,00546	85	0,00200	1,58	12	14	20	male	people	structures
1	12	25	100	1000	25	imageres_7_14_16.txt	0,00383	0,00561	94	0,00178	1,46	7	14	16	female	people	portrait
1	32	500	100	1000	25	imageres_12_14_20.txt	0,00346	0,00490	85	0,00144	1,42	12	14	20	male	people	structures
1	25	100	100	1000	25	imageres_7_14_16.txt	0,00383	0,00503	94	0,00120	1,31	7	14	16	female	people	portrait
1	30	500	100	1000	25	imageres_10_12_14.txt	0,00542	0,00701	133	0,00159	1,29	10	12	14	indoor	male	people
2	112	5000	100	1000	50	imageres_15_20_23.txt	0,00220	0,00280	54	0,00060	1,27	15	20	23	plant_life	structures	tree
1	5	25	100	1000	25	imageres_12_14_16.txt	0,00159	0,00202	39	0,00043	1,27	12	14	16	male	people	portrait
1	26	100	100	1000	25	imageres_7_14_20.txt	0,00253	0,00317	62	0,00064	1,25	7	14	20	female	people	structures
2	78	1000	200	1000	25	imageres_7_14_20.txt	0,00253	0,00317	62	0,00064	1,25	7	14	20	female	people	structures
1	13	25	100	1000	25	imageres_7_14_20.txt	0,00253	0,00291	62	0,00038	1,15	7	14	20	female	people	structures
1	4	25	100	1000	25	imageres_10_12_14.txt	0,00542	0,00598	133	0,00056	1,10	10	12	14	indoor	male	people
1	52	1000	100	1000	25	imageres_7_14_20.txt	0,00253	0,00264	62	0,00011	1,04	7	14	20	female	people	structures
1	39	500	100	1000	25	imageres_7_14_20.txt	0,00253	0,00264	62	0,00011	1,04	7	14	20	female	people	structures
1	24	100	100	1000	25	imageres_7_10_14.txt	0,00668	0,00696	164	0,00028	1,04	7	10	14	female	indoor	people
1	63	1000	100	1000	50	imageres_7_10_14.txt	0,00668	0,00670	164	0,00002	1,00	7	10	14	female	indoor	people
2	76	1000	200	1000	25	imageres_7_10_14.txt	0,00668	0,00658	164	-0,00010	-0,01	7	10	14	female	indoor	people
2	101	1000	50	1000	75	imageres_7_10_14.txt	0,00668	0,00643	164	-0,00025	-0,04	7	10	14	female	indoor	people
1	28	500	100	1000	25	imageres_1_6_10.txt	0,00273	0,00249	67	-0,00024	-0,09	1	6	10	animals	dog	indoor
1	23	100	100	1000	25	imageres_5_19_20.txt	0,00289	0,00262	71	-0,00027	-0,09	5	19	20	clouds	sky	structures
1	10	25	100	1000	25	imageres_5_19_20.txt	0,00289	0,00262	71	-0,00027	-0,09	5	19	20	clouds	sky	structures
2	62	1000	100	1000	50	imageres_5_19_20.txt	0,00289	0,00262	71	-0,00027	-0,09	5	19	20	clouds	sky	structures
1	11	25	100	1000	25	imageres_7_10_14.txt	0,00668	0,00584	164	-0,00084	-0,13	7	10	14	female	indoor	people
1	37	500	100	1000	25	imageres_7_10_14.txt	0,00668	0,00580	164	-0,00088	-0,13	7	10	14	female	indoor	people
2	96	1000	50	1000	75	imageres_12_14_20.txt	0,00346	0,00294	85	-0,00052	-0,15	12	14	20	male	people	structures
1	36	500	100	1000	25	imageres_5_19_20.txt	0,00289	0,00241	71	-0,00048	-0,17	5	19	20	clouds	sky	structures
2	100	1000	50	1000	75	imageres_5_19_20.txt	0,00289	0,00241	71	-0,00048	-0,17	5	19	20	clouds	sky	structures
2	56	1000	100	1000	50	imageres_10_12_14.txt	0,00542	0,00444	133	-0,00098	-0,18	10	12	14	indoor	male	people
1	9	25	100	1000	25	imageres_15_20_23.txt	0,00220	0,00175	54	-0,00045	-0,20	15	20	23	plant_life	structures	tree
2	69	1000	200	1000	25	imageres_10_12_14.txt	0,00542	0,00416	133	-0,00126	-0,23	10	12	14	indoor	male	people
2	72	1000	200	1000	25	imageres_13_19_20.txt	0,00200	0,00128	49	-0,00072	-0,36	13	19	20	night	sky	structures
2	65	1000	100	1000	50	imageres_7_14_20.txt	0,00253	0,00159	62	-0,00094	-0,37	7	14	20	female	people	structures
1	43	1000	100	1000	25	imageres_10_12_14.txt	0,00542	0,00319	133	-0,00223	-0,41	10	12	14	indoor	male	people
1	50	1000	100	1000	25	imageres_7_10_14.txt	0,00668	0,00382	164	-0,00286	-0,43	7	10	14	female	indoor	people
2	94	1000	50	1000	75	imageres_10_12_14.txt	0,00542	0,00290	133	-0,00252	-0,46	10	12	14	indoor	male	people
1	38	500	100	1000	25	imageres_7_14_16.txt	0,00383	0,00183	94	-0,00200	-0,52	7	14	16	female	people	portrait
2	102	1000	50	1000	75	imageres_7_14_16.txt	0,00383	0,00172	94	-0,00211	-0,55	7	14	16	female	people	portrait
1	88	5000	100	1000	25	imageres_7_10_14.txt	0,00668	0,00296	164	-0,00372	-0,56</						

fase	id	Clusters	Neighbors	Iterations	Points	File-precision-per-cluster	Random	Precision	number images	Precision - Random	number of times Topsurf beter dan random	CL1	CL2	CL3	CL1-Name	CL2-Name	CL3-Name
1	11	25	100	1000	25	imageres_7_10_14.txt	0,00668	0,00584	164	-0,00084	-0,13	7	10	14	female	indoor	people
1	37	500	100	1000	25	imageres_7_10_14.txt	0,00668	0,00580	164	-0,00088	-0,13	7	10	14	female	indoor	people
2	96	1000	50	1000	75	imageres_12_14_20.txt	0,00346	0,00294	85	-0,00052	-0,15	12	14	20	male	people	structures
1	36	500	100	1000	25	imageres_5_19_20.txt	0,00289	0,00241	71	-0,00048	-0,17	5	19	20	clouds	sky	structures
2	100	1000	50	1000	75	imageres_5_19_20.txt	0,00289	0,00241	71	-0,00048	-0,17	5	19	20	clouds	sky	structures
2	56	1000	100	1000	50	imageres_10_12_14.txt	0,00542	0,00444	133	-0,00098	-0,18	10	12	14	indoor	male	people
1	9	25	100	1000	25	imageres_15_20_23.txt	0,00220	0,00175	54	-0,00045	-0,20	15	20	23	plant_life	structures	tree
2	69	1000	200	1000	25	imageres_10_12_14.txt	0,00542	0,00416	133	-0,00126	-0,23	10	12	14	indoor	male	people
2	72	1000	200	1000	25	imageres_13_19_20.txt	0,00200	0,00128	49	-0,00072	-0,36	13	19	20	night	sky	structures
2	65	1000	100	1000	50	imageres_7_14_20.txt	0,00253	0,00159	62	-0,00094	-0,37	7	14	20	female	people	structures
1	43	1000	100	1000	25	imageres_10_12_14.txt	0,00542	0,00319	133	-0,00223	-0,41	10	12	14	indoor	male	people
1	50	1000	100	1000	25	imageres_7_10_14.txt	0,00668	0,00382	164	-0,00286	-0,43	7	10	14	female	indoor	people
2	94	1000	50	1000	75	imageres_10_12_14.txt	0,00542	0,00290	133	-0,00252	-0,46	10	12	14	indoor	male	people
1	38	500	100	1000	25	imageres_7_14_16.txt	0,00383	0,00183	94	-0,00200	-0,52	7	14	16	female	people	portrait
2	102	1000	50	1000	75	imageres_7_14_16.txt	0,00383	0,00172	94	-0,00211	-0,55	7	14	16	female	people	portrait
1	88	5000	100	1000	25	imageres_7_10_14.txt	0,00668	0,00296	164	-0,00372	-0,56	7	10	14	female	indoor	people
1	44	1000	100	1000	25	imageres_12_14_16.txt	0,00159	0,00067	39	-0,00092	-0,58	12	14	16	male	people	portrait
1	18	100	100	1000	25	imageres_12_14_16.txt	0,00159	0,00067	39	-0,00092	-0,58	12	14	16	male	people	portrait
2	103	1000	50	1000	75	imageres_7_14_20.txt	0,00253	0,00106	62	-0,00147	-0,58	7	14	20	female	people	structures
1	49	1000	100	1000	25	imageres_5_19_20.txt	0,00289	0,00121	71	-0,00168	-0,58	5	19	20	clouds	sky	structures
2	114	5000	100	1000	50	imageres_7_10_14.txt	0,00668	0,00228	164	-0,00440	-0,66	7	10	14	female	indoor	people
2	116	5000	100	1000	50	imageres_7_14_20.txt	0,00253	0,00079	62	-0,00174	-0,69	7	14	20	female	people	structures
2	67	1000	200	1000	25	imageres_1_6_10.txt	0,00273	0,00068	67	-0,00205	-0,75	1	6	10	animals	dog	structures
2	64	1000	100	1000	50	imageres_7_14_16.txt	0,00383	0,00092	94	-0,00291	-0,76	7	14	16	female	people	portrait
2	110	5000	100	1000	50	imageres_13_19_20.txt	0,00200	0,00043	49	-0,00157	-0,79	13	19	20	night	sky	structures
2	75	1000	200	1000	25	imageres_5_19_20.txt	0,00289	0,00040	71	-0,00249	-0,86	5	19	20	clouds	sky	structures
1	51	1000	100	1000	25	imageres_7_14_16.txt	0,00383	0,00046	94	-0,00337	-0,88	7	14	16	female	people	portrait
2	77	1000	200	1000	25	imageres_7_14_16.txt	0,00383	0,00046	94	-0,00337	-0,88	7	14	16	female	people	portrait
1	88	5000	100	1000	25	imageres_5_19_20.txt	0,00289	0,00020	71	-0,00269	-0,93	5	19	20	clouds	sky	structures
1	82	5000	100	1000	25	imageres_10_12_14.txt	0,00542	0,00034	133	-0,00508	-0,94	10	12	14	indoor	male	people
2	107	5000	100	1000	50	imageres_10_12_14.txt	0,00542	0,00023	133	-0,00519	-0,96	10	12	14	indoor	male	people
1	79	5000	100	1000	25	imageres_1_15_23.txt	0,00082	0,00000	20	-0,00082	-1,00	1	15	23	animals	plant_life	tree
1	1	25	100	1000	25	imageres_1_15_23.txt	0,00082	0,00000	20	-0,00082	-1,00	1	15	23	animals	plant_life	tree
2	104	5000	100	1000	50	imageres_1_15_23.txt	0,00082	0,00000	20	-0,00082	-1,00	1	15	23	animals	plant_life	tree
2	66	1000	200	1000	25	imageres_1_15_23.txt	0,00082	0,00000	20	-0,00082	-1,00	1	15	23	animals	plant_life	tree
2	91	1000	50	1000	75	imageres_1_15_23.txt	0,00082	0,00000	20	-0,00082	-1,00	1	15	23	animals	plant_life	tree
1	80	5000	100	1000	25	imageres_1_6_10.txt	0,00273	0,00000	67	-0,00273	-1,00	1	6	10	animals	dog	indoor
2	105	5000	100	1000	50	imageres_1_6_10.txt	0,00273	0,00000	67	-0,00273	-1,00	1	6	10	animals	dog	indoor
1	81	5000	100	1000	25	imageres_1_6_20.txt	0,00086	0,00000	21	-0,00086	-1,00	1	6	20	animals	dog	structures
1	42	1000	100	1000	25	imageres_1_6_20.txt	0,00086	0,00000	21	-0,00086	-1,00	1	6	20	animals	dog	structures
1	29	500	100	1000	25	imageres_1_6_20.txt	0,00086	0,00000	21	-0,00086	-1,00	1	6	20	animals	dog	structures
1	16	100	100	1000	25	imageres_1_6_20.txt	0,00086	0,00000	21	-0,00086	-1,00	1	6	20	animals	dog	structures
1	3	25	100	1000	25	imageres_1_6_20.txt	0,00086	0,00000	21	-0,00086	-1,00	1	6	20	animals	dog	structures
2	55	1000	100	1000	50	imageres_1_6_20.txt	0,00086	0,00000	21	-0,00086	-1,00	1	6	20	animals	dog	structures
2	93	1000	50	1000	75	imageres_1_6_20.txt	0,00086	0,00000	21	-0,00086	-1,00	1	6	20	animals	dog	structures
1	83	5000	100	1000	25	imageres_12_14_16.txt	0,00159	0,00000	39	-0,00159	-1,00	12	14	16	male	people	portrait
1	31	500	100	1000	25	imageres_12_14_16.txt	0,00159	0,00000	39	-0,00159	-1,00	12	14	16	male	people	portrait
2	108	5000	100	1000	50	imageres_12_14_16.txt	0,00159	0,00000	39	-0,00159	-1,00	12	14	16	male	people	portrait
2	57	1000	100	1000	50	imageres_12_14_16.txt	0,00159	0,00000	39	-0,00159	-1,00	12	14	16	male	people	portrait
2	70	1000	200	1000	25	imageres_12_14_16.txt	0,00159	0,00000	39	-0,00159	-1,00	12	14	16	male	people	portrait
2	95	1000	50	1000	75	imageres_12_14_16.txt	0,00159	0,00000	39	-0,00159	-1,00	12	14	16	male	people	portrait
1	85	5000	100	1000	25	imageres_13_19_20.txt	0,00200	0,00000	49	-0,00200	-1,00	13	19	20	night	sky	structures
2	113	5000	100	1000	50	imageres_5_19_20.txt	0,00289	0,00000	71	-0,00289	-1,00	5	19	20	clouds	sky	structures
1	89	5000	100	1000	25	imageres_7_14_16.txt	0,00383	0,00000	94	-0,00383	-1,00	7	14	16	female	people	portrait
2	115	5000	100	1000	50	imageres_7_14_16.txt	0,00383	0,00000	94	-0,00383	-1,00	7	14	16	female	people	portrait
1	90	5000	100	1000	25	imageres_7_14_20.txt	0,00253	0,00000	62	-0,00253	-1,00	7	14	20	female	people	structures

Figuur 30: ranking groep-tag combinaties beter dan random