

Vind foto's met behulp van foto's

Steeds vaker zoekt een onderzoeker gericht naar één foto. Om dit beeld te vinden moet gezocht worden met de juiste trefwoorden, die alleen doeltreffend zijn wanneer de onderzoeker de structuur en achtergrond van het archief kent. Fotoarchieven zouden daarom actiever moeten nadenken hoe de collectie op andere manieren doorzocht zou kunnen worden, bijvoorbeeld m.b.v. beeldherkenningstechnieken.

Fotodatabanken bedienen verschillende gebruikers met verschillende zoekvragen. Een ware Beatrix-fan zal een avond lang door een databank browsen en klikken op elke foto die zijn aandacht trekt. Een onderzoeker zal gericht zoeken, bijvoorbeeld op datum wanneer hij een artikel schrijft over de manier waarop Beatrix in een bepaald jaar werd geportretteerd en de redacteur van een krant wil wellicht snel die ene foto vinden van de rookbom achter de gouden koets op de huwelijksdag van Beatrix en Claus in 1966. Vooralsnog is er weinig bekend over de wensen van gebruikers van fotoarchieven, maar uit recent onderzoek naar het zoekgedrag van gebruikers bij audiovisuele archieven blijkt dat mensen vaak zeer gerichte zoekvragen hebben.¹ Ze zoeken een bepaald televisiefragment en willen dat zo snel mogelijk vinden. Maar om dit beeldmateriaal te vinden, moet je begrijpen hoe een archief georganiseerd is. Een trefend voorbeeld is de ervaring van Cara Finnegan in het fotoarchief van de Farm Security Administration (FSA), de Amerikaanse overheidsorganisatie die tijdens de grote depressie in de jaren dertig allerlei hervormingsprojecten lanceerde als onderdeel van de 'New Deal'.² Om steun onder de bevolking te krijgen voor de ingrijpende maatregelen werd fotografen onder meer gevraagd om de gevolgen van de 'Dust Bowl' (een opeenvolging van stofstormen op de vlakten in het midden van de Verenigde Staten tussen 1930 en 1941) te fotograferen, resulterend in bekende foto's van onder andere Walker Evans en Dorothea Lange's - voor Amerikanen - iconische foto van de *Migrant mother*. Finnegan deed onderzoek naar de publicatie van de FSA-foto's in tijdschriften uit de jaren dertig en wilde in het archief in Washington DC de foto's traceren om te bekijken of de foto's waren bijgesneden. Ze had een hele duidelijke zoekvraag en wist precies naar welke foto ze op zoek

was. Echter, het kostte haar veel moeite om een foto te vinden van een man met vuile kleren, die poseert op de veranda van zijn huisje. Ze besloot de collectie te doorzoeken met behulp van de trefwoorden die aan de foto's zijn toegekend. Na dagen zoeken had ze de foto nog steeds niet gevonden, totdat ze zich de vraag stelde: 'What is this a picture of?' Pas toen realiseerde ze dat zij al die tijd dacht dat de man het hoofdonderwerp op de foto is. Maar wanneer ze een stapje terugneemt en nadenkt over het doel van de FSA-campagne beseft ze dat het ging om het vastleggen van onder andere de miserabele woonomstandigheden. Zodoende vindt ze haar foto niet in de map 'Persons', maar uiteindelijk in de categorie 'Homes and Living Conditions'.

Achtergronden

Deze anekdote illustreert hoe onderzoekers te werk gaan en hoe fotoarchieven zijn opgebouwd. Bij het beschrijven van foto's zijn we immers afhankelijk van woorden. En wanneer we woorden gebruiken, interpreteren we. Natuurlijk is een aantal feitelijke gegevens eenvoudig op te schrijven, zoals plaats en jaartal. In termen van Roland Barthes *denoteer* je dan de foto.³ Maar het wordt ingewikkelder wanneer de foto's *geconnoteerd* moeten worden, waarbij de associatie



72.021. Corn. Leenheer. Pieter Jelles Troelstra tijdens Rode Dinsdag, Den Haag, 17 september 1912

van de foto wordt genoteerd. Eigenlijk lijkt het heel eenvoudig: waarover gaat de foto? Wat is het centrale onderwerp? Maar kijk eens naar het voorbeeld van een bekende Nederlandse foto van Pieter Jelles Troelstra, waarop hij een van zijn gepassioneerde toespraken houdt, gemaakt op 17 september 1912 tijdens een massale demonstratie voor algemeen kiesrecht in Den Haag.

De foto is op dit moment beschikbaar bij in ieder geval drie Nederlandse archieven. Alle drie geven ze de foto verschillende trefwoorden. Het IISG zegt dat de foto over 'algemeen kiesrecht' gaat, het Tresoar zegt dat het een 'actiefoto' is en bij Spaarnestad Photo gaat de foto over 'politieke partijen' en 'socialdemocratie in Nederland'.⁴ Dit soort verschillende beschrijvingen is te verklaren door rekenschap te geven van de achtergrond van het archief. Aangezien deze foto is ingevoerd in een database, zal iedereen hem eenvoudig kunnen vinden, doordat alle drie de archieven natuurlijk ook hebben opgeschreven dat Troelstra op de foto staat. Maar voor Cara Finnegan was het helaas niet zo eenvoudig om de man met vuile kleren te vinden, poserend op de veranda van zijn huisje. Al was het alleen maar omdat ze niet wist hoe de man heette en - belangrijker nog - ze niet wist met welke woorden ze moest zoeken. Ze betoogt dan ook dat gebruikers van archieven de achtergrond van het archief moeten kennen om de 'retorica' van de beschrijvingen van de foto's te kunnen doorgronden. Niet voor niets heeft ze haar ervaringen beschreven in een artikel dat gepubliceerd is in het wetenschappelijke tijdschrift *Rhetoric & Public Affairs*. Het heeft ook iets paradoxaals: beelden moeten worden beschreven met woorden en daarbij wordt praktisch altijd geïnterpreteerd. Zowel onderzoekers als archivariissen hebben daardoor constant te maken met de thesaurus waarmee een foto is beschreven. Schrijver en criticus Allan Sekula beschrijft het werk van de Fransman Alphonse Bertillon en de Brit Francis Galton, die beiden een systeem ontwikkelden om sociaal afwijkend gedrag op foto's te kunnen zien. Sekula betoogt echter dat het onvermijdelijk is om verschillende systemen te ontwikkelen om foto's te beschrijven: '... it is absurd to imagine a dictionary of photographs, unless one is willing to disregard the specificity of individual images in favor of some model of typicality (...)'⁵

Thesauri

Zodoende zien we alleen al in Nederland allerlei verschillende thesauri om beelden te omschrijven. Spaarnestad Photo gebruikt de eigen ontwikkelde CODEC (Concrete Decimale Classificatie) waarmee foto's worden ingedeeld op basis van een combinatie van getallen. Een alternatief is *Iconclass*, waarmee beeldelementen worden beschreven. Dit systeem is in de jaren vijftig van de vorige eeuw ontwikkeld door de Leidse hoogleraar Henri van de Waal en wordt tegenwoordig onderhouden door het Rijksbureau voor Kunsthistorische Documentatie (RKD).⁶ Daarnaast wordt ook de *Gemeenschappelijke Thesaurus Audiovisuele Archieven* (GTAA) gebruikt.⁷ Deze is ontwikkeld door het Nederlands Instituut voor Beeld en Geluid en EYE (voormalig Filmmuseum) en wordt sinds kort ook gebruikt door het Nationaal Archief.⁸ Welke thesaurus er ook wordt gebruikt, interpretatie met behulp van tekst blijft onvermijdelijk. Sekula droomt dan ook over: 'a machine, or rather a clerical apparatus, a filing system, which allows the operator/researcher/editor to retrieve the individual instance from the huge quantity of images contained within the archive.'⁹

Een dergelijke machine lijkt op de portals die op dit moment worden ontwikkeld, vooral op Europese schaal, zoals *EUScreen* voor televisie en de *European Film Gateway* voor films, die het materiaal tevens beschikbaar maken voor *Europeana*, de portal om digitaal materiaal van Europese bibliotheken, archieven, musea en audiovisuele collectie te verkennen.¹⁰ De initiatieven voor een dergelijke samenwerking op het gebied van de fotografie komen in Europees verband langzaam van de grond en zullen, net als de reeds bestaande initiatieven, vooral gebruik maken van de tekst en metadata die bij de foto geleverd wordt. Dit voldoet dan nog lang niet aan de droom van Sekula, maar wellicht bieden de ontwikkelingen in het veld van beeldherkenning een eerste aanzet om wel te denken over de realisering van Sekula's 'machine'.

Beeldherkenning

Reeds sinds het einde van de jaren zestig worden de mogelijkheden van beeldherkenning verkend.¹¹ In 2000 schreef Annemieke de Jong in het *Tijdschrift voor Mediageschiedenis* over de mogelijkheden van beeldherkenning om collecties te beschrijven.¹² Bij het grote publiek zijn de mogelijkheden van beeldherkenning pas recent bekend geworden door de introductie van de techniek in consumentensoftware, zoals

iPhoto en Google's Picasa, dat gezichten kan herkennen in een persoonlijke fotocollectie, of websites als *Tineye* of *Google Images*, waar je een foto kunt uploaden en automatisch identieke of soortgelijke foto's op het internet gevonden worden.¹³

Ook binnen de informatica wordt gekeken naar de mogelijkheden van beeldherkenning voor erfgoedinstellingen en geesteswetenschappers. Het project *BeeldCanon* is een voorbeeld van een dergelijke studie, waarbij informatici onderzoeken hoe bekende Nederlandse en Vlaamse gebouwen op beeldmateriaal automatisch herkend kunnen worden.¹⁴ Onlangs is het Europese project *AXES - Access to Audiovisual Archives* gestart, waarbij bestaande technieken als spraak- en ook beeldherkenning worden gecombineerd om audiovisuele archieven op een vernieuwende manier te ontsluiten.¹⁵

De Amerikaanse hoogleraar Lev Manovich gebruikt soortgelijke technieken om trends te ontdekken in grote databestanden en richt zich daarbij vooral op culturele producten in wat hij 'cultural analytics' noemt.^{16,17} Bijzonder hoogleraar Geschiedenis van de Radio en Televisie Huub Wijffjes riep in zijn oratie op om dergelijke technieken ook eindelijk eens toe te passen binnen de geesteswetenschappen.¹⁸ Een oproep die ik van harte ondersteun. Want wat zou het fijn zijn geweest voor Cara Finnegan wanneer zij de door haar gezochte foto op een site kon inladen, waarna alle FSA-foto's doorzocht werden en ze de foto in kwestie meteen zou vinden. Het zou de machine zijn waar Sekula over droomt en die op dit moment nog geen realiteit is, maar wellicht wel snel gerealiseerd kan worden.

Kansen voor fotoarchieven

Een dergelijke machine zou niet alleen de gebruikers van fotoarchieven geweldig helpen om de door hen gezochte foto's te vinden, maar kan ook de archieven helpen. Stel je voor dat een recent verworven collectie door de 'beeldherkenner' gehaald kan worden, waarna de software automatisch de mensen en gebouwen op de foto herkent en suggesties geeft voor hun namen. Het zou het werk van annoteerders een stuk makkelijker maken. Is dit toekomstmuziek? Natuurlijk. Maar de eerste initiatieven op dit gebied vinden reeds plaats, getuige de eerder genoemde projecten *BeeldCanon* en *AXES*. Hierbij wordt concreet geëxperimenteerd om te bekijken hoe dergelijke technieken ingezet kunnen worden teneinde collecties



72.022. Schermafbeelding van de Axes-website

anders te ontsluiten en onderzoekers te ondersteunen in hun zoektocht naar beeldmateriaal. Er is echter één punt van aandacht: Nederlandse fotocollecties worden in dergelijke projecten nog erg weinig gebruikt. Dat is ontzettend jammer vanwege op zijn minst drie redenen. De eerste, en wellicht de belangrijkste, is financieel: op dit moment zijn er verschillende onderzoeksprogramma's die onderzoeksprojecten ondersteunen op het gebied van de e-humanities (onderzoek in de geesteswetenschappen met behulp van digitale technieken). In Nederland loopt bijvoorbeeld al enige tijd het NWO-programma *Continuous Access to Cultural Heritage* (CATCH) waarbij geesteswetenschappers samen met informatici en cultureel-erfgoedinstellingen een vierjarig onderzoeksproject kunnen opzetten om gezamenlijke onderzoeksvragen te onderzoeken.¹⁹ De geesteswetenschapper heeft een relevante onderzoeksvraag, waarna de informatici werken aan een techniek om die vraag te beantwoorden met behulp van de collecties van de erfgoedinstellingen. Een soortgelijk programma is *Clarín*, dat beoogt collecties met elkaar te koppelen dankzij gezamenlijke standaarden om geesteswetenschappelijk onderzoek te faciliteren.²⁰ Ik droom bijvoorbeeld al lange tijd over een portal waar een onderzoeker een zoekterm kan invullen, waarna alle aangesloten fotoarchieven worden doorzocht en de resultaten op één plek getoond worden. Wanneer daar nog eens een beeldherkenner aan gekoppeld kan worden, zijn er helemaal veel onderzoeksmogelijkheden. Dergelijke projecten worden heel behoorlijk financieel ondersteund en hier liggen mooie kansen voor fotoarchieven, omdat tot op heden vooral tekstuele en audiovisuele archieven participeren in dergelijke projecten.

Een tweede reden om aan dit soort projecten mee te doen lijkt dan ook evident: deelnemende instellingen profiteren mee van de technieken die gedurende het project ontwikkeld worden. Technieken die binnen CATCH-projecten zijn ontwikkeld en als geslaagd worden beschouwd, worden verder ontwikkeld in het programma *CATCH-Plus*, zodat ze daadwerkelijk kunnen worden ingezet bij de erfgoedinstellingen.²¹ Tenslotte zorgen dergelijke projecten ook voor een ander gebruik van de collectie, waardoor de bekendheid en wellicht ook het gebruik van de collectie zullen toenemen.

Toekomstdromen

Natuurlijk zijn dit soort projecten niet de heilige graal, net zoals beeldherkennings-technieken nog erg in de kinderschoenen staan. Maar het is wel een feit dat tegenwoordig technieken worden ontwikkeld die we ons tien jaar geleden nog niet eens konden voorstellen. Fotoarchieven zouden zich dan ook moeten aansluiten bij dergelijke projecten om de collecties op andere en vernieuwende manieren te ontsluiten. Dit is natuurlijk makkelijker gezegd dan gedaan. Maar wellicht ligt hier een rol voor het Nederlands Fotogenootschap, zodat de instellingen gezamenlijk kunnen optrekken. Een eerste brainstormsessie van geïnteresseerde instellingen, wetenschappers en informatici zou al een goede stap zijn. Ik meld me hierbij alvast graag aan voor een eerste bijeenkomst.

Martijn Kleppe

Martijn Kleppe werkt aan de Erasmus Universiteit Rotterdam aan een proefschrift over Nederlandse iconofoto's. Sinds januari 2011 werkt hij tevens als senior onderzoeker aan het Europese onderzoeksproject AXES - *Access to Audiovisual Archives*. kleppe@eshcc.eur.nl

Dit artikel is een bewerking van een langer artikel dat in december verschijnt in het Tijdschrift voor Mediageschiedenis.

Noten:

1. Search in audiovisual broadcast archives / Bouke Huurnink. – Amsterdam : [s.n.], 2010. – (proefschrift Universiteit van Amsterdam)
2. What is this a picture of? : some thoughts on images and archives / C.A. Finnegan
In: *Rhetoric & Public Affairs* 9 (2006) 1 ; p. 117
3. The photographic message / Roland Barthes
In: *Image music text* / Roland Barthes. – London : Fontana, 1977 ; p. 17-19
4. IISG: search.iisg.nl/search/search?action=transform&col=marc_images&xsl=marc_images-detail.

- xsl&lang=en&docid=10555691_MARC
Tresoor: dokumentaasje.tresoor.nl/tonen.php?id=11427
Spaarnestad Photo: www.spaarnestadphoto.nl/component/option,com_memorix/task,topview/searchplugin,simple/Itemid,60/lang,nl/mrxpopup,1/CollectionID,1/PhotoID,SFA001010540/RecordID,7995/ResultRecord,7
Alle voor het laatst bezocht op 11 september 2011
5. The body and the archive / A. Sekula
In: *October* (39) 1986, p. 3–64, aldaar 17
 6. Drie eeuwen vaderlandsche geschied-uitbeelding, 1500-1800 : een iconologische studie / H. van de Waal. - 's-Gravenhage : Martinus Nijhoff, 1952.
Meer over ICONCLASS via www.iconclass.nl/about-iconclass/history-of-iconclass. Via de portal Arkyves kan een aantal archieven die ontsloten zijn met Iconclass doorzocht worden. Arkyves is op dit moment te gebruiken bij de EUR, UvA, UU, het Rijksmuseum en de KB
 7. Thesauri gekoppeld / Johan Oomen en Hennie Brugman
In: *Digitale bibliotheek* 2 (2010) 5 ; p. 18-21
 8. Meer informatie over het beschikbaar stellen van de GTAA op www.catchplus.nl/2010/catchplus-maakt-beschikbaarstelling-gtaa-mogelijk
 9. Ibid. noot 5
 10. Meer informatie over de genoemde projecten op www.euscreen.eu, www.europeanfilmgateway.eu en europeana.eu

11. Picture processing by computer / Azriel Rosenfeld
In: *ACM Computer Survey* 1(1969) 3 ; p. 147-176
12. Het beeld, het woord en de algoritmen : mogelijkheden en onmogelijkheden van automatische indexering / Annemieke de Jong
In: *Tijdschrift voor Mediageschiedenis* 3 (2000) 2 (december) ; p. 113-137
13. Meer informatie over de genoemde programma's op picasa.google.com, www.tineye.com en www.google.com/imgph
14. Meer informatie over het project 'Beeldcanon' op staff.science.uva.nl/~gavves/beeldcanon/html/index.html en in: *Landmark image retrieval using visual synonyms* / Efstratios Gavves & Cees G.M. Snoek
In: *Proceedings of ACM Multimedia*, 2010 ; p. 1123-1126
15. Meer informatie over AXES op www.axes-project.eu
16. The language of new media / Lev Manovich. - Cambridge : MIT, 2001
17. Cultural analytics : visualizing cultural patterns in the era of 'more media' / Lev Manovich. – [s.l. : s.n.], 2009
18. Uit de ban van goede en slechte smaak : perspectieven in televisiegeschiedenis / H. Wijffjes. – Amsterdam : Amsterdam University press, 2010 ; p. 18-19
19. www.nwo.nl/catch
20. www.clarin.nl
21. www.catchplus.nl



72.023. Schermafbeelding van gezichtsherkenning met iPhoto

‘Het is zorgelijk dat veel instellingen niet weten wat ze in huis hebben. En dat is toch wél een primaire taak. Je ziet het vooral bij instellingen buiten het circuit van het NFG. Kennis en diepgang schieten daar vaak tekort. Het netwerk dat het NFG biedt is zo goed als uniek in de wereld. Collega's in het buitenland zijn ze er jaloers op’

Mattie Boom, conservator fotografie Rijksmuseum/oud-voorzitter NFG