



Aan
VCA Stakeholders

Van
Drs. J.H.C. van Rest

Onderwerp
Terminologie en Taxonomie van Video Content Analyse

BU 1 Waarnemingssystemen
Oude Waalsdorperweg 63
Postbus 96864
2509 JG Den Haag

www.tno.nl

T +31 70 374 00 00
F +31 70 328 09 61
info-DenV@tno.nl

Datum
12 juli 2010

Onze referentie
<vnr-ext>

E-mail
jeroen.vanrest@tno.nl

Doorkiesnummer
+31 15 269 29 73

Inhoudsopgave

Inhoudsopgave	1
Leeswijzer	1
Inleiding en stand van zaken	1
VCA Systeem.....	2
Functionaliteit van VCA	2
Fysieke implementatie van VCA	4
Kwaliteit	4
VCA Waardeketen	5
Internationaal.....	5

Leeswijzer

Deze white paper is tot stand gekomen naar aanleiding van een discussie met leveranciers en eindgebruikers over de objectieve kwaliteitsmeting van Video Content Analyse in het domein van maatschappelijke veiligheid.

Inleiding en stand van zaken

Automatisch waarnemen aan menselijk gedrag wordt steeds vaker toegepast, zo ook in het veiligheidsdomein. Een van de technieken om dit mee te doen is Video Content Analyse (VCA), ook wel Video Analytics genaamd.

In het veiligheidsdomein in Engeland wordt een groei van de inzet van video content analyse (VCA) gerapporteerd door de BSIA¹. De technologie is ook sterk in ontwikkeling bij bedrijven en kennisinstellingen, getuige het grote aantal onderzoeksprogramma's. *CANTATE*, *VICOMO* en *Efficiënt Zoeken in Digitale Beelden* zijn slechts enkele voorbeelden waar Nederlandse partijen bij betrokken zijn.

Ook in Nederland zijn recent enkele forse cameraprojecten gestart of uitgevoerd, bijvoorbeeld in het openbaar vervoer², de luchthavensector en bij de beveiliging van nationale infrastructuur.

¹ BSIA reports an increase in the use of Video Content Analysis, 15 maart 2010,
<http://www.bsia.co.uk/aboutbsia/cctv/O5E926740891>

² Toespraak pNCTb tijdens de uitreiking van de subsidiebeschikkingen van het CTOV-programma,
http://www.nctb.nl/Actueel/Toespraken/NCTb/toespraak_090526.aspx?cp=91&cs=25499



Datum
12 juli 2010

Onze referentie
<vnr-ext>

Blad
2/5

De ontwikkelingen worden gedreven door sterke technologische en bedrijfsmatige trends. De wet van Moore³ zorgt voor steeds meer processorkracht, resolutie en bandbreedte die er samen voor zorgen dat er steeds meer resources voor VCA beschikbaar komen. De voortdurende zoektocht naar hogere efficiëntie en effectiviteit zorgt er voor dat de vraag naar geoperationaliseerde intelligentie voortdurend groeit. Het is te verwachten dat beide trends onverlet door zullen zetten.

In het veiligheidsdomein speelt de mens in de rol als beveiligger een hele belangrijke rol. Dit wordt geïllustreerd door de grote hoeveelheden menskracht in de private beveiliging, maar ook om de roep naar Meer Blauw op Straat. Beeldinformatie zal mede door haar toegankelijkheid voor de mens een geliefde databron in dit domein blijven.

VCA Systeem

Een VCA systeem is een systeem dat voor een deel van zijn functie gebruik maakt van de analyse van video signalen. We onderscheiden daarbij functionaliteit, de fysieke implementatie en de kwaliteitsaspecten.

Functionaliteit van VCA

VCA is het automatisch interpreteren van video signalen⁴. Het kan als zodanig gezien worden als de automatisering van de biologisch visuele cortex. Taxonomisch gezien is het een soort beeldverwerking, en daarmee wordt het gezien als een vorm van kunstmatige intelligentie.

Er zijn vele functionaliteiten die te maken zijn met VCA. Voorbeelden zijn bewegingsdetectie, gezichtsdetectie en –herkenning, en tracking. In enge zin bouwt VCA een interne representatie van een geobserveerde wereld op vanuit ruwe sensordata. In ruimere zin probeert ze op basis van deze representatie ook tot een begrip te komen van de situatie. Dit wordt ook wel het overbruggen van de semantic gap⁵ genoemd. Voor de interne representatie zijn voor sommige functionaliteiten databeschrijvingen bekend, zoals voor de functionaliteit *tracking* de PETS Benchmark Data⁶ een databeschrijving voorstelt. Op basis van deze interne representatie en begrip van de situatie – zoals bijvoorbeeld detectie afwijkend gedrag- wordt complexere functionaliteit gebouwd. Voorbeelden hiervan zijn perimeterbeveiliging en toegangscontrole.

³ *Cramming more components onto integrated circuits*, G.E. Moore, Electronics Magazine 19 April 1965

⁴ http://en.wikipedia.org/wiki/Video_Content_Analysis

⁵ http://en.wikipedia.org/wiki/Semantic_gap

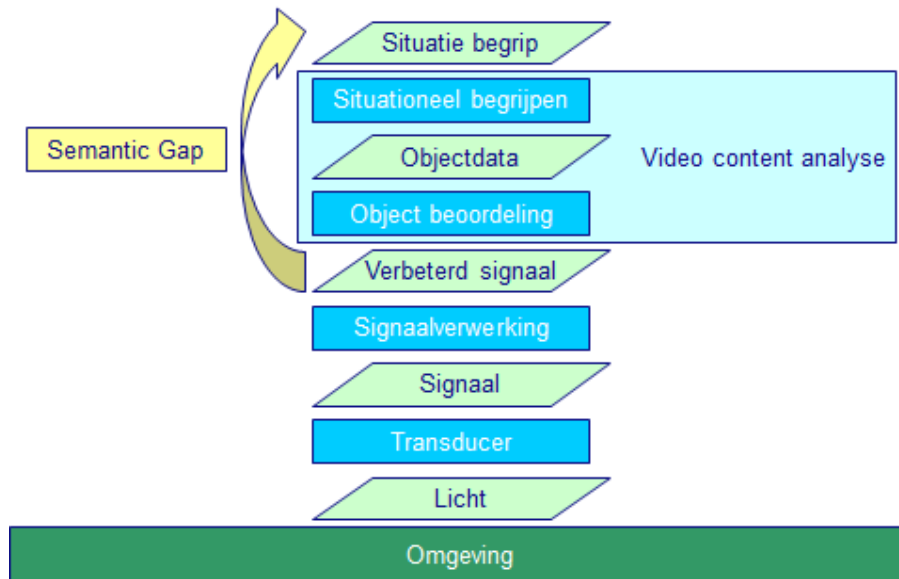
⁶ Performance Evaluation of Tracking and Surveillance, <http://www.cvg.cs.rdg.ac.uk/slides/pets.html>



Datum
12 juli 2010

Onze referentie
<vnr-ext>

Blad
3/5



6

Den Haag, vrijdag 25 juni 2010



Figuur 1 - Functionele decompositie van video content analyse

De kwaliteit van VCA hangt onder andere af van de kwaliteit van hetingangssignaal. Dit kan verbeterd worden door technieken als superresolutie, stabilisatie en contrastverbetering toe te passen. Op zich voegen deze geen informatie toe aan het signaal, maar ze worden typisch wel sterk geassocieerd met VCA.

Figuur 1 illustreert hoe de deelfunctionaliteiten samen hangen. Ze is geïnspireerd op IEEE papers over datafusie⁷. Merk op dat dit niets zegt over waar de functionaliteit zich fysiek bevindt: in een camera, of juist centraal. Het zegt dus ook niets over de onderliggende communicatie structuur. Tenslotte laat het ook de kardinaliteit in het midden. Een rijk voorbeeld van een ingevulde functionele keten kan gevonden worden voor identificatie op basis van gezichtsherkenning:

Transducer: Een persoon loopt langs 2 camera's waarvan de onderlinge posities bekend zijn. Het licht wordt opgevangen door de camera's en gestitched tot 1 beeld met hogere resolutie en vanuit een beter viewpoint dan met 1 camera mogelijk was geweest.

Signaalverwerking: Het is een hele donkere dag, dus het blijkt nodig om ruisreductie en contrastverbetering toe te passen om robuust te zijn tegen de dynamische lichtomstandigheden.

Object beoordeling: In het beeld wordt gezocht naar betekenisvolle objecten, in dit geval gezichten, en daarna onderdelen van gezichten. De locatie van de neus, de ogen en de mondhoeken worden bepaald.

Situatie begrijpen: Het actuele gezicht wordt op basis van de gevonden kenmerken vergeleken met een database van gezichten. Er komt een ranking van meest

⁷ Designing Networked Adaptive Interactive Hybrid Systems, Kester, L.J.H.M., IEEE MFI 2008



Datum
12 juli 2010

Onze referentie
<vnr-ext>

Blad
4/5

gelijke gezichten uit. Aan deze gezichten zijn identiteiten gekoppeld en de identiteit van de eerste persoon op die lijst wordt als resultaat gegeven.

Fysieke implementatie van VCA

De kwaliteit van het videosignaal is op zijn beurt weer afhankelijk van de sensor en de omstandigheden rond de sensor: fixatie, belichting, resolutie, camerahoek en focusgebied. De functionaliteit van het tellen van mensen is bijvoorbeeld makkelijker te maken met een cameraopstelling die boven op de mensen kijkt, dan met een camera die van opzij kijkt, omdat dan occlusie door mensen onderling uitgesloten kan worden.

Een ander element van de fysieke implementatie heeft betrekking op de locatie van de processing. Deze kan bijvoorbeeld dicht bij de sensor -dus decentraal oftewel gedistribueerd- geïmplementeerd zijn, maar deze kan ook juist centraal gebeuren. Beide hebben voor- en nadelen in verband met benodigde bandbreedte, databeveiliging, onderhoudbaarheid, robuustheid en te realiseren functionaliteit.

Kwaliteit

De primaire kwaliteitsfactor van VCA is de correctheid van de interne representatie ten opzichte van de realiteit, niet alleen voor een object, maar voor de hele situatie. Afgeleiden hiervan zijn bijvoorbeeld de false matches / false non-matches bij identificatie, en de hoeveelheid trackfragmentatie bij tracking. Initiatieven als TRECVID⁸, de Face Recognition Vendor Test⁹ en i-Lids¹⁰ hebben voor specifieke functionaliteiten op specifieke datasets een annotatie-proces om deze correctheid te bepalen. Om dit te doen hebben ze dus datastandaarden moeten afspreken om tot een vergelijk te kunnen komen.

Andere prestatiefactoren liggen op het gebied van snelheid en robuustheid in termen van de kwaliteit van het ingangssignaal.

Een indirecte manier om naar de kwaliteit van VCA te kijken is middels het track-record. Hierbij zien we grote verschillen tussen verschillende vormen van VCA. Bijvoorbeeld, nummerplaatherkenning en gezichtsherkenning hebben een beter track-record dan tracking. Maar binnen één functionaliteit is ook weer onderscheid te maken. Bijvoorbeeld het track-record van gezichtsherkenning in een scenario van toegangscontrole versus het track-record van gezichtsherkenning bij inbeslagnames van videomateriaal in zedenzaken. Deze zijn bijvoorbeeld verschillend omdat de fysieke omstandigheden rond de sensoren typisch anders zijn, of omdat de bijvoorbeeld de mensen in een bepaalde use case wel of juist niet meewerken met het systeem.

Met betrekking tot datasets zouden er dus 2 verschillende datasets voor deze 2 verschillende use cases moeten komen: één met mensen die meewerken en van

⁸ TREC Video Retrieval Evaluation, <http://trecvid.nist.gov/>

⁹ Face Recognition Vendor Test, <http://www.frvt.org/>

¹⁰ Image Library for Intelligent Detection Systems, <http://www.ilids.co.uk/>



Datum
12 juli 2010

Onze referentie
<vnr-ext>

Blad
5/5

dichtbij in de camera kijken, en een ander van mensen die op grotere afstanden poseren, of niet zelf met de camera bezig zijn, of zelfs met een bewegende camera.

VCA Waardeketen

De waardeketen van VCA is te typeren door 3 rollen te onderscheiden: de producent van VCA, de system integrator die VCA als onderdeel inzet en de eindgebruiker. Dit is een versimpeling van de werkelijkheid, maar voor het doel van deze paper voldoende.

Sommige organisaties verenigen meerdere rollen in zich. Er zijn echter geen organisaties die alle drie de rollen in zich verenigen. Er zijn dus altijd meerdere partijen nodig om tot een complete oplossing te komen.

De politie verenigt bijvoorbeeld in zich de rollen eindgebruiker en system integrator. Voor heimelijke observatie bijvoorbeeld zijn ze zelf system integrator, voor de surveillance toepassing zullen ze echter wel eerder vertrouwen op een externe system integrator. Er zijn ook system integrators die ook producent zijn van VCA. Hier is Bosch een voorbeeld van.

Internationaal

De markt voor VCA is internationaal. Producenten komen uit binnen- en buitenland, en leveren in binnen- en buitenland. Inkopers kijken internationaal. De bestaande buitenlandse testdatasets zoals i-Lids en TRECVID zijn in Nederland bekend, maar zouden wellicht vaker gebruikt kunnen worden om de transparantie in de markt te vergroten.

Het maakt bij bovenstaande categorisering van producten en productfamilies niet uit waar de betreffende testdataset vandaan komt of wie hem beheert, zolang de onafhankelijkheid van het testproces maar gewaarborgd blijft.