

[Faculty of Science
Kennispunt Bètawetenschappen]

Opsporing misbruik beeldmateriaal

Prestaties van gezichtsherkenningssoftware

Yannick van Welzen

KPB/11-12



Universiteit Utrecht

Kennispunten slaan een brug tussen universiteit en maatschappij. Zij behandelen onderzoeksvragen van bedrijven, overheden en maatschappelijke organisaties.

Colofon

Rapportnummer	KPB/11-12
ISBN	978-90-79589-14-2
Verschenen	Juni 2011
Druk	Eerste
Titel	Opsporing misbruik beeldmateriaal Prestaties van gezichtsherkenningsoftware
Auteur	Y. van Welzen
Begeleider	prof. dr. R.C. Veltkamp, Multimedia en Geometrie, Dept. Informatica, Universiteit Utrecht
Projectcoördinator	drs. J. van Winden, Kennispunt Bètawetenschappen, Universiteit Utrecht
Opdrachtgever	Stichting AMBI, het Burgercomité tegen Onrecht en Open Ogen – bureau voor training en voorlichting
Projectnummer	KPB/1027
Uitgever	Kennispunt Bètawetenschappen, Universiteit Utrecht Leuvenlaan 4, 3584 CE Utrecht. tel. 030-253 7363 www.uu.nl/beta/kennispunt
Coverfoto	Shannon Kringen, Flickr
Disclaimer	Het is niet toegestaan (gedeelten van) deze uitgave te vermenigvuldigen door middel van druk, fotokopie, microfilm, of op welke andere wijze dan ook. Overname van gedeelten van de tekst, mits met bronvermelding, is wel toegestaan. Toezending van een bewijsexemplaar wordt zeer op prijs gesteld.

Inhoudsopgave

Dankwoord	5
Samenvatting	7
1 Inleiding	9
1.1 Motivatie	9
1.2 Doelstelling	9
1.3 Opdrachtgever	9
1.4 Onderzoeksvraag	10
1.5 Afbakening	10
1.6 Opbouw van het verslag	10
2 State-of-the-Art	11
3 Aanpak	13
3.1 Foto's	13
3.2 Experiment 1	15
3.3 Experiment 2	16
3.4 Prestatie-indicatoren	17
4 Resultaten	19
4.1 Experiment 1	19
4.2 Experiment 2	20
5 Conclusie	23
6 Evaluatie/discussie	25
Literatuurlijst	27

Dankwoord

Dank aan prof. dr. Veltkamp voor de begeleiding en dr. Spacek van de University of Essex voor het beschikbaar maken van zijn database.

Yannick van Welzen

Utrecht, mei 2011

Samenvatting

Compromitterende foto's van slachtoffers van *happy-slapping*, zedendelicten of loverboys verschijnen steeds vaker op het internet. Als de mogelijkheid bestaat om deze beelden eenvoudig op te sporen en te laten verwijderen zou dit veel leed besparen. Om dit te bereiken is het nodig dat slachtoffers goed automatisch herkend kunnen worden op dit beeldmateriaal. Dit onderzoek is een eerste verkenning van de mogelijkheden om dit voor elkaar te krijgen met behulp van gezichtsherkenningsoftware. De prestaties van drie gratis gezichtsherkenningapplicaties zijn vergeleken en er zijn experimenten uitgevoerd met een zelf gebouwde applicatie.

De prestaties van Pitt Patt Facesort, Fotobounce en Google Picasa zijn vergeleken door zoekopdrachten te geven naar personen uit verschillende fotodatabases (samen 515 foto's). PittPatt Facesort heeft het hoogste gezichtsherkenningpercentage (88,4%) en koppelde 114 keer de verkeerde naam aan een foto. Fotobounce had het laagste herkenningpercentage (51,4%) en stelde in totaal 2108 keer een verkeerde persoon voor. Google Picasa herkende 76,9% van de gezichten zonder één keer een fout voorstel te doen. Voor alle programma's geldt dat een hoger aantal referentiefoto's een positieve invloed heeft op het herkenningpercentage. Ook scoorden alle drie de applicaties beter voor foto's die genomen zijn in een gecontroleerde omgeving dan foto's genomen in een ongecontroleerde omgeving. Daarnaast werden frontale foto's beter herkend dan zijwaartse foto's en scoorden over- of onderbelichte foto's ondergemiddeld, evenals foto's van personen met een hoofddekseel of bril.

In de zelf gebouwde applicatie is onderzocht of de software beter presteert als een referentiefoto van het slachtoffer aan de trainingsset wordt toegevoegd. Dit was inderdaad het geval. Het gemiddelde herkenningpercentage nam toe met 9%. De gemiddelde positie van de laatste voorgestelde foto waarop de persoon juist herkend was, nam echter ook toe, van positie 219 naar 226.

Het onderzoek laat zien dat de onderzochte software zeker in staat is gezichten van personen te herkennen en terug kan vinden in een database. De applicaties scoren echter niet goed in ongecontroleerde omgevingen; iets wat wel belangrijk is voor de toepassing bij compromitterende foto's van slachtoffers. Omdat een toename in het aantal referentiefoto's een positief effect heeft op het herkenningpercentage, wordt aangeraden om meerdere foto's van het slachtoffer te nemen en te gebruiken voor de zoekopdracht. Als het mogelijk is zouden deze foto's ook aan de trainingsset toegevoegd moeten worden. Verder onderzoek moet uitwijzen of het toevoegen van zijwaartse foto's de prestaties voor de herkenning van zijwaartse foto's verbetert.

Voor de uiteindelijke toepassing is het ook belangrijk om onderzoek te doen naar het doorzoeken van het internet in plaats van een (kleine) fotodatabase met de gezichtsherkenningsoftware.

Hoofdstuk 1

Inleiding

1.1 Motivatie

De laatste jaren wordt steeds vaker beeldmateriaal van mensen die slachtoffer zijn geworden van zedendelicten of *happy-slapping* op het internet gezet. Van mensenhandelaren en loverboys is bekend dat zij compromitterend beeldmateriaal gebruiken als chantagemiddel om hun slachtoffers te dwingen zich te prostitueren of andere handelingen te verrichten. Het komt ook voor dat er wordt ingebroken op een webcam van een slachtoffer waarmee ongemerkt beeldmateriaal wordt verkregen. Dit materiaal wordt vervolgens op het internet gezet of het slachtoffer wordt gechanteerd met de dreiging het beeldmateriaal op internet te zetten.

Wanneer personen op foto's herkend kunnen worden, kan er een lijst gemaakt worden met foto's waarop de betreffende persoon staat, om vervolgens deze foto's van het internet af te laten halen. Dit zorgt dat er minder mogelijkheden zijn om slachtoffers te chanteren.

Dit rapport beschrijft een eerste verkenning naar mogelijkheden om beeldmateriaal van slachtoffers automatisch te herkennen. Daartoe is een onderzoek opgezet dat uit een aantal verschillende experimenten bestaat.

1.2 Doelstelling

Het primaire doel van dit onderzoek is inzicht te krijgen in de mogelijkheden om mensen die op foto's staan te herkennen. Er zal gebruik worden gemaakt van gezichtsherkenning. Wanneer er succesvol een afbeelding van een slachtoffer is gedetecteerd, kan/moet de beheerder van de site worden aangesproken om het betreffende materiaal op verzoek van het slachtoffer, de hulpverlener of de politie van de site af te halen.

1.3 Opdrachtgever

De formele opdrachtgever is het Kennispunt Bètawetenschappen, wat onderdeel is van de Universiteit Utrecht. Het Kennispunt helpt bedrijven, overheden en maatschappelijke organisaties bij het ontwikkelen van nieuwe kennis. Daarnaast biedt het studenten en onderzoekers de ervaring om hun kennis op een directe, waardevolle manier in te zetten voor de maatschappij. Het Kennispunt heeft van Stichting AMBI, het Burgercomité tegen Onrecht en van het bureau voor voorlichting en training Open Ogen een verzoek tot onderzoek ontvangen en dit neergelegd bij de afdeling Informatica van de Universiteit Utrecht. Vervolgens is er een masterstudent Game & Media Technology gevonden die dit onderzoek wilde uitvoeren.

1.4 Onderzoeksvraag

De hoofdvraag luidt:

“In hoeverre is het mogelijke om personen op foto’s te herkennen, en - in het geval van mislukte herkenning - waarom wordt de persoon niet altijd herkend?”

Om hier inzicht in te krijgen wordt gekeken naar bestaande softwaresystemen, en is er een kleine testapplicatie gebouwd met behulp van een softwarebibliotheek.

1.5 Afbakening

Bij de startbijeenkomst is naar voren gekomen dat het belangrijk is om gezichten te herkennen. De nadruk ligt minder op het detecteren van dezelfde beelden; het betreft de herkenning van personen en in mindere mate de herkenning van kopieën van bestaande foto’s. In de praktijk komt het geregeld voor dat het beeldmateriaal bewerkt wordt met behulp van grafische bewerkingsprogramma’s. Hierdoor is het moeilijk om kopieën te detecteren. Het is echter zeer onwaarschijnlijk dat deze grafische bewerkingsoperaties zullen worden toegepast op de gezichten van de slachtoffers, aangezien dat zal resulteren in een verminderde mogelijkheid om te chanteren; het slachtoffer is dan minder herkenbaar.

Daarnaast kan met behulp van gezichtsherkenning ook onbekend beeldmateriaal van slachtoffers worden opgespoord.

Er zal gebruik gemaakt worden van bestaande software. Er wordt een selectie gemaakt van applicaties die onder de loep worden genomen, dit in verband met beperkingen, waaronder de aanschafprijs.

Stichting AMBI, het Burgercomité tegen Onrecht en Open Ogen zijn ook geïnteresseerd in persoonsherkenning op films. Gezien de gelimiteerde tijd die dit onderzoek in beslag kan nemen, is dit niet behandeld in het huidige onderzoek.

1.6 Opbouw van het verslag

Het vervolg van het verslag ziet er als volgt uit. In hoofdstuk 2 worden een aantal verschillende technieken op het gebied van gezichtsherkenning besproken. Hoofdstuk 3 gaat verder in op de onderzoeks aanpak. Het beschrijft de methoden van de experimenten en gaat in op de gebruikte fotodatabases en prestatie-indicatoren. In hoofdstuk 4 worden de resultaten van de experimenten gepresenteerd. Hoofdstuk 5 en 6 zijn respectievelijk de conclusie en evaluatie/discussie.

Hoofdstuk 2

State-of-the-art

Het herkennen van personen met behulp van gezichtsherkenning is een bekend probleem waar al decennia lang onderzoek naar wordt gedaan. Huidige methodes voor het herkennen van gezichten zijn robuust, maar niet foutloos. Af en toe komt het voor dat een persoon verkeerd geïdentificeerd wordt. Ook nu wordt nog actief onderzoek gedaan naar steeds betere en efficiëntere methodes. Dit hoofdstuk beschrijft enkele recente onderzoeken op het gebied van gezichtsherkenning.

Chai et al. (2007) hebben een techniek ontwikkeld waardoor gezichten sneller herkend kunnen worden. Een van de bottlenecks van gezichtsherkenning is de verschillende invalshoeken in de foto's. Chai en collega's hebben een methode ontwikkeld die vanuit een niet-frontale invalshoek een frontaal beeld genereert. Hierdoor is het eenvoudiger om gezichtsspecifieke informatie te berekenen, met als gevolg dat personen beter herkend worden. De foto's zijn nu genormaliseerd; alle gezichten die herkend moeten worden hebben dezelfde invalshoek. Er moet echter wel vermeld worden dat de conversie van een niet-frontaal beeld naar een frontaal beeld gedeeltelijk berust op kansrekening. Dit houdt in dat het frontale beeld dat gegenereerd wordt uit twee verschillende niet-frontale beelden niet altijd precies hetzelfde is.

Ling et al. (2010) hebben onderzoek gedaan naar de verificatie van personen. Hierbij wordt vastgesteld of wel of niet dezelfde persoon staat op twee foto's die een aantal jaar na elkaar zijn gemaakt. Ling et al. gebruikten een *Gradient Orientation Pyramid* om de foto's te representeren en in combinatie met een *Support Vector Machine* worden personen vaker herkend dan wanneer gebruik gemaakt wordt van oudere methoden, ondanks de leeftijdsverschillen in de foto's. Uit hetzelfde onderzoek is gebleken dat het lastiger is om mensen te herkennen wanneer zij een bril dragen dan wanneer zij van kapsel zijn veranderd.

Park & Jain (2010) hebben onderzoek gedaan naar het gebruik van biometrische kenmerken zoals sproeten en littekens. De biometrische kenmerken kunnen helpen gezichten te herkennen die gedeeltelijk bedekt zijn, of om bijvoorbeeld eeneiige tweelingen van elkaar te onderscheiden. Park & Jain hebben een methode ontwikkeld die gebruik maakt van een *Active Appearance Model*, *Laplacian-of-Gaussian blob* detectie en morfologische operatoren. Uit hun onderzoek is gebleken dat de correcte herkenningratio sterk omhoog gaat bij een combinatie van de ontwikkelde methode en conventionele gezichtsherkenning algoritmes. Ze claimen dan ook dat een dergelijke combinatie de prestatie van een state-of-the-art gezichtsherkenner kan verbeteren.

Een ander onderzoek is gedaan naar gezichtsherkenning op lage resolutie foto's. Door het beperkt aantal pixels in lage resolutie foto's is het niet altijd mogelijk om de benodigde informatie op te halen. Hennings-Yeomans et al. (2008) hebben een methode ontwikkeld die de gezichtsherkenning sterk verbetert wanneer meerdere foto's van lage resolutie gebruikt worden. Conventionele methodes maken gebruik van superresolutie foto's, waarbij het aantal

pixels groter wordt door middel van interpolatie, maar vatbaar is voor fouten. Of zij *downgraden* de trainingfoto's naar een lage resolutie, wat een negatieve invloed heeft op het herkenningspercentage. De methode die Hennings-Yeomans et al. hebben ontwikkeld maakt gebruik van superresolutie modellen en gezichtskenmerken die geëxtraheerd zijn uit de lage-resolutie foto's.

In 2006 zijn de resultaten van de *Face Recognition Grand Challenge* gepubliceerd. Het primaire doel van deze Grand Challenge was het promoten en verbeteren van gezichtsherkenning (www.frtv.org^a). Er werd onderzoek gedaan naar gezichtsherkenning voor hoge resolutieplaatjes, 3D-modellen en irisscans. Eerder, in 2002, was de *Face Recognition Vendor Test* uitgevoerd. Het doel van de Grand Challenge was om de resultaten van deze Vendor Test te verbeteren. In 2006 werd een tweede *Face Recognition Vendor Test* in het leven geroepen, waarin experimenten werden uitgevoerd om te kijken of dit doel was gehaald (www.frtv.org^b). Uit de resultaten bleek dat dit doel is bereikt voor foto's die gemaakt zijn in een gecontroleerde omgeving en 3D-modellen. De herkenning van hoge resolutie foto's, en iris scans was vergelijkbaar met de resultaten uit 2002. In een experiment waarin de prestaties van gezichtsherkenning algoritmes werden vergeleken met de prestaties van mensen, deden de algoritmes het beter dan de mensen.

Bovenstaande onderzoeken illustreren dat het onderzoeksgebied continu in beweging is en dat er vooruitgang geboekt wordt. De prestatie van gezichtsherkenning software wordt dankzij dergelijke onderzoeken steeds beter. Gezichtsherkenning anno 2011 is al redelijk goed, maar het is te verwachten dat gezichtsherkenning in de toekomst nog accurater wordt, waardoor personen nog vaker worden herkend.

Hoofdstuk 3

Aanpak

Om een duidelijk beeld te krijgen van de mogelijkheden, worden verschillende applicaties onderworpen aan experimenten. Er worden twee experimenten uitgevoerd. Het eerste experiment wordt uitgevoerd op bestaande applicaties. Het tweede experiment wordt uitgevoerd op een applicatie die zelf ontwikkeld is met behulp van een *Computer Vision Library*. De applicaties die onder de loep zijn genomen zijn:

Experiment 1

Google Picasa

Fotobounce

Pitt Patt FaceSort

Experiment 2

Eigen applicatie

3.1 Foto's

Er zal gebruik gemaakt worden van bestaande gezichtsdata bases, en van een zelf verzamelde collectie. Deze data bases bevatten gezichten van personen onder verschillende omstandigheden. De data bases die gebruikt worden voor het huidige onderzoek zijn de Yale Face Database en een database die samengesteld is door Dr. L. Spacek. Er zal ook gebruik gemaakt worden van individueel van internet verzamelde foto's van bekende personen. Er is voor bekende personen gekozen omdat van hen veel foto's vanuit verschillende invalshoeken te vinden zijn. Hierdoor kan goed vastgesteld worden in welke situaties de gezichten daadwerkelijk herkend worden. De foto's verschillen van elkaar, zodat vastgesteld kan worden in welke situaties de personen herkend worden. Het betreft de volgende situaties:

- Foto's van verschillende afmetingen.
- Foto's met verschillend contrast
- Foto's vanuit een ander gezichtspunt
- Foto's met andere belichting (donker/licht/verschillende richtingen)
- Foto's waarop de persoon een hoofddekse l of bril draagt
- Foto's met verschillende gezichtsuitdrukkingen

Verder zullen de foto's ingedeeld worden in de volgende categorieën:

1. Hoofddekse l / bril
- 2) Gezichtshaar

- 3) Geen hoofddekseel / bril of gezichtshaar
- 4) Onderbelichting
- 5) Normale belichting
- 6) Overbelichting
- 7) Lachen
- 8) Geen lachen

Het is mogelijk dat één foto onder meerdere categorieën valt. Bij elke foto zal een ja/nee antwoord worden gegeven; wordt de persoon wel of niet herkend? Door de situaties één voor één af te gaan, wordt een antwoordenmatrix verkregen. Deze matrix zal gebruikt worden om een antwoord te geven op de onderzoeksvraag.

3.1.1 Databases

Het totaal aantal gebruikte foto's komt neer op 515, waarvan 91 foto's van vrouwen en 424 foto's van mannen zijn. Dit aantal is verspreid over drie fotodatabases.

Yale Face Database

In totaal bevat de database 165 zwart-wit foto's, 11 foto's van 15 personen. De personen hebben verschillende gezichtsuitdrukkingen per foto en de foto's verschillen van configuratie: frontale belichting, belichting van links, belichting van rechts, met/zonder bril, blij, droevig, normaal, vermoeid, verrast en een knipoog. De droevige gezichtsuitdrukking is voor persoon 4 mislukt en vervangen door de normale gezichtsuitdrukking. De personen op de foto's zijn allemaal volwassenen van verschillende etnische afkomsten en de leeftijden verschillen van jonge volwassenen tot ouderen. De foto's zijn 320 x 243 pixels. De verschillende personen worden in dit onderzoek subject01 tot subject15 genoemd.

Dr. L. Spacek database

In totaal bevat de database 7900 foto's, 20 foto's van 395 personen. Er zijn foto's van mensen van verschillende afkomst en de database bevat foto's van zowel mannen als vrouwen. De meerderheid van de afbeeldingen zijn gemaakt van eerstejaars studenten van tussen de 18 tot 20 jaar maar er zijn ook foto's van oudere personen. De database bevat tevens foto's van mensen met een bril en/of baard. De foto's zijn in kleur en het formaat van de foto's is 24 bit JPEG. De belichting is kunstmatig en is voor elke foto hetzelfde. De foto's zijn 180 x 200 pixels.

Er zijn 10 willekeurige personen uit deze database gekozen, waarvan alle 20 foto's per persoon zijn gebruikt. Hierdoor komt het aantal foto's dat uit deze database gebruikt zal worden neer op 200 foto's. De personen die voor dit onderzoek gebruikt zijn heten Boylee, Cywan, Dfhodd, Howar, Jross, Kouri, Ldgodd, Mizli, Mkosto en Pjrand.

Individueel verzamelde foto's

Er zijn 30 foto's per bekend persoon verzameld, 20 frontaal, 5 van de linker- en 5 van de rechterkant. De personen van wie foto's zijn verzameld zijn: Bill Clinton, Barack Obama, Mahmoud Ahmedinejad, Daniel Radcliffe als Harry Potter en Martin Jol. De persoonskeuze is gebaseerd op een variatie van etnische achtergrond, haargroei en hoofddekseels. De foto's zijn genomen vanuit verschillende invalshoeken, hebben een variërende belichting en zijn van verschillende afmetingen. De foto's zijn gedownload van Google afbeeldingen.

3.2 Experiment 1

In experiment 1 worden de prestaties van drie bestaande gratis applicaties met elkaar vergeleken: Google Picasa, Fotobounce en PittPatt Facesort. Daarnaast wordt in dit experiment onderzocht of er met behulp van meerdere referentiefoto's meer foto's worden gevonden waarop de juiste persoon te zien is.

3.2.1 Applicaties voor gezichtsherkenning

De herkenning van personen geschiedt door gebruik te maken van gezichtsherkenning. Er wordt gebruik gemaakt van drie bestaande applicaties. Het betreft Google Picasa, Fotobounce en PittPatt Facesort. Deze applicaties zijn ontworpen om gemakkelijk foto's te groeperen in albums. De gebruiker kan door zijn eigen mappenstructuur bladeren om vervolgens een foto te selecteren en te benoemen. Dit handmatige proces wordt 'taggen' genoemd. De foto die in een album staat, zal in het verdere verloop van dit rapport een referentiefoto genoemd worden. Voor de eerste referentiefoto in een album is het van belang dat de persoon op de foto recht in de camera kijkt. Hierdoor kan de software preciezer te werk gaan en is het makkelijker om de kenmerken te extraheren en foto's te vinden waarop dezelfde persoon staat.

Zodra de foto benoemd is, neemt deze plaats in het bijbehorende album. Op dit moment zullen de drie applicaties de gezichten op de foto detecteren om vervolgens kenmerken van de gezichten te extraheren. Wanneer dit gedaan is, gebruiken zij deze kenmerken om in de overige foto's die in de mappenstructuur te vinden zijn op te zoek gaan naar gezichten met dezelfde kenmerken. Wanneer de bovengenoemde applicaties gezichten hebben gevonden die voldoen aan deze kenmerken, kan de gebruiker zelf kiezen of de gevonden foto's deel uit moeten maken van het album.

3.2.2 Fases

Dit experiment kent drie herkenningsfases. In de eerste fase wordt alleen de referentiefoto gebruikt als zoekopdracht. De verschillende applicaties zullen nadat de zoekopdracht is uitgevoerd een aantal foto's tonen. Vervolgens is het aan de gebruiker om de voorgestelde foto's door te nemen en wel/niet te accepteren. Er wordt bijgehouden welke foto's voorgesteld worden en deze informatie zal aan het einde van het experiment geanalyseerd worden. De lijst met voorgestelde foto's wordt doorlopen en de foto's die in het album horen (de foto's met dezelfde persoon als de persoon op de referentiefoto) worden één voor één aan het album toegevoegd. Wanneer alle voorgestelde foto's zijn toegevoegd aan het album, begint de tweede fase. In de tweede fase van het experiment worden ook de foto's die geaccepteerd zijn in de eerste fase gebruikt als zoekopdracht. In deze fase worden dus meerdere referentiefoto's gebruikt. Wederom wordt bijgehouden welke foto's door de drie applicaties voorgesteld worden en worden de correcte voorgestelde foto's toegevoegd aan het album. Dit proces wordt nog eens herhaald voor de derde fase. Wanneer de derde fase compleet doorlopen is voor een gegeven persoon, wordt het album van deze persoon verwijderd. Hierdoor zal de applicatie de ene persoon niet meer herkennen wanneer een ander persoon herkend zou moeten worden. Dit creëert een eerlijk herkenningsproces met een gelijk aantal opties voor alle personen.

3.3 Experiment 2

Dit experiment heeft betrekking op de eigen applicatie. De software maakt gebruik van een bestaande *softwarelibrary*. Deze *library* gebruikt een *Principal Component Analysis*-techniek voor de herkenning van gezichten. Voordat de applicatie in staat is om gezichten te herkennen is het vereist om een aantal foto's op te geven als trainingset. Er wordt een trainingset bestaande uit vijftig foto's van willekeurige mensen gebruikt. Hieruit destilleert het programma een 'gemiddeld gezicht'. Vervolgens wordt een referentiefoto aangeboden van een persoon die in de database herkend moet gaan worden, bijvoorbeeld van een slachtoffer. De software berekent dan hoeveel deze referentiefoto afwijkt van dit 'gemiddelde gezicht' en van de foto's uit de trainingset. Gezichten uit een database, die ongeveer een gelijke afwijking hebben als de referentiefoto, lijken veel op het gezicht van het slachtoffer en zijn dus mogelijk van dezelfde persoon.

Meer wiskundig geformuleerd projecteert de applicatie de foto's op hun bijbehorende Eigenspace. Dit proces genereert een gemiddelde foto en een aantal Eigenspace plaatjes: G, E_1, \dots, E_k .

Vervolgens wordt door de gebruiker een referentiefoto gekozen van een persoon. Deze foto maakt geen deel uit van de trainingset. Deze foto wordt ook geprojecteerd op de Eigenspace:

$$R_i = G + w_{i,1}E_1 + \dots + w_{i,k-1}E_{k-1}$$

Waarin R_i de foto is uit de referentieset, G de gemiddelde foto is van de gehele trainingset, $w_{i,1}E_1$ het product is van een scalar $w_{i,1}$ behorende bij de eerste Eigenspace projectie, etc, en k het aantal foto's in de trainingset is. Aangezien $G, E_1 \dots E_{k-1}$ constant zijn, kunnen de foto's ook worden gerepresenteerd als een lijst met gewichten:

$$R_i = [w_{i,1} \dots w_{i,k-1}]$$

Deze lijst wordt gebruikt als een query op een opgegeven dataset. Deze dataset bestaat uit een aantal gezichtsdata bases die beschreven worden in de volgende paragraaf, maar in de praktijk zou de opgegeven dataset het internet kunnen zijn. Ook de individuele foto's uit de dataset worden geprojecteerd op de Eigenspace en gerepresenteerd als een lijst met gewichten. Deze lijst wordt vergeleken met de lijst behorende bij de referentiefoto.

Het vergelijken komt neer op het berekenen van een afstand tussen beide lijsten. Deze afstand is gedefinieerd als Euclidische afstand tussen $[w_{i,1} \dots w_{i,k-1}]$ en $[w_{l,1} \dots w_{l,k-1}]$:

$$\sqrt{\sum_{j=1}^{k-1} (w_{i,j} - w_{l,j})^2}$$

Waar $w_{i,j}$ een gewicht is behorende bij een foto uit de dataset.

Hoe kleiner de afstand, hoe meer de huidige foto lijkt op de referentiefoto. Wanneer de gehele dataset is doorlopen, wordt de dataset olopend gesorteerd op afstand en komen de foto's die het meest overeenkomen met de referentiefoto vooraan in de lijst te staan.

Een interessante vraag is of de gezichtsherkenning verbetert wanneer de referentiefoto wel deel uitmaakt van de trainingset. Dit zal ook onderzocht worden in dit experiment. Dit experiment bestaat dus uit twee delen. In deel 1 bevat de trainingset niet de referentiefoto, in deel 2 bevat de trainingset wel de referentiefoto. Afhankelijk van de resultaten kan geconcludeerd worden of het toevoegen van de referentiefoto aan de referentieset invloed heeft op de effectiviteit van de gezichtsherkenning.

3.4 Prestatie-indicatoren

Als prestatie-indicator voor het eerste experiment wordt het herkenningspercentage berekend. Dit is gedefinieerd als:

$$h = \frac{g}{t} * 100\%,$$

waar h staat voor het herkenningspercentage, g staat voor het aantal gevonden relevante beelden en t staat voor het totaal aantal relevante beelden. Voor het eerste experiment wordt gekeken naar de herkenningspercentages van de verschillende fases, categorieën en databases.

Als prestatie-indicator voor het tweede experiment wordt de *recall* berekend. *Recall* is gedefinieerd als:

$$\text{recall} = \frac{v}{c} * 100\%,$$

waar v staat voor het aantal correct herkende personen en c staat voor het totaal aantal foto's van de persoon in de dataset.

Naast *recall* wordt ook geteld hoe vaak een programma een verkeerde naam koppelt aan een foto. Deze foute suggesties zullen in het vervolg van het verslag misclassificatie worden genoemd.

Voor experiment 2 worden tevens de *first tier recall* ($s = c$) en de *second tier recall* ($s = 2c$) berekend. Hierbij is s de *scope* en c representeert het totaal aantal foto's van een persoon in de dataset. De *scope* geeft aan hoeveel foto's uit de resultatenlijst van de zoekopdracht meegenomen worden om het percentage correct herkende personen te bepalen. Als in de *first tier* de positie van een foto van een persoon buiten de *scope* valt, zal deze foto niet meetellen voor het aantal correct herkende personen. Wanneer de *scope* bijvoorbeeld dertig is, zullen alleen de foto's op de eerste dertig posities worden meegenomen in de berekeningen. De overige foto's zullen worden genegeerd bij het bepalen van de waarde van v (zie formule).

Als laatste prestatie-indicator zal gebruik gemaakt worden van de *last ranked image*. Dit is de positie van laatste foto van de persoon op de referentiefoto. Deze drie metrieken zijn niet van toepassing op het eerste experiment, aangezien de informatie die benodigd is voor het berekenen ervan niet beschikbaar is.

Hoofdstuk 4

Resultaten

4.1 Experiment 1

Tabel 1 toont de totale herkenningpercentages van de drie applicaties en de herkenningpercentages per applicatie.

Tabel 1: Gezichtsherkenningpercentages van Picasa, Fotobounce en PittPatt Facesort per fotocategorie.

	Picasa (%)	Fotobounce (%)	PittPatt Facesort (%)	Gemiddelde over de applicaties (%)
1. Hoofddeksel/bril	47,9	17,0	68,8	44,8
2. Gezichtshaar	86,4	41,9	93,0	73,6
3. Geen hoofddeksel/bril of gezichtshaar	78,4	60,0	90,5	76,6
4. Onderbelichting	66,3	44,9	84,2	65,6
5. Normale belichting	83,8	57,1	91,4	75,0
6. Overbelichting	36,4	6,3	65,6	36,1
7. Lachen	78,4	41,2	93,1	71,2
8. Geen lachen	76,8	54,3	88,0	73,1
Categorieën gemiddeld	76,9	51,4	88,4	72,2
Vrouwen*	100	80,2	100	93,4
Mannen*	94,5	64,6	98,8	86,0

*Bij 'vrouwen' en 'mannen' zijn de individueel verzamelde foto's buiten beschouwing gelaten.

Uit deze gegevens blijkt dat categorie 1 (hoofddeksel/bril) en categorie 6 (overbelichting) de laagste herkenningpercentages hebben. De foto's uit categorie 4 (onderbelichting) scoren ook ondergemiddeld. Tevens is het herkenningpercentage voor mannen en vrouwen gegeven. Bij de berekening van deze percentages is de database met individueel verzamelde foto's buiten beschouwing gelaten omdat deze database enkel mannen bevat. Dit verklaart het hogere percentage voor de gemiddelde waarden in de tabel.

Fotobounce is het programma dat het minst goed gezichten herkent; 51,4% van de gezichten werd herkend en in totaal stelde het programma 2108 keer een verkeerd persoon voor. Picasa herkende 76,9% van de gezichten zonder één keer een fout voorstel te doen. PittPatt Facesort herkende 88,4% van de gezichten en koppelde in totaal 114 keer de verkeerde naam aan een foto.

Tabel 2 toont de herkenningpercentages per programma per fase. De resultaten komen overeen met de verwachtingen: Er worden meer personen herkend naarmate er meer referentiefoto's worden gebruikt om te zoeken.

Tabel 2: Gezichtsherkenningpercentages per fase voor Picasa, Fotobounce en PittPatt Facesort

	Picasa (%)	Fotobounce (%)	PittPatt Facesort (%)
1 ^e fase	67,4	26,3	85,3
2 ^e fase	75,3	42,7	88,4
3 ^e fase	76,9	51,4	88,4

Interessant om te zien is hoe de verschillende programma's scoren per database. In het huidige experiment zijn er twee gecontroleerde databases (de Yale Face database en de L. Spacek database) en één ongecontroleerde database (de individueel verzamelde foto's database). Tabel 3 toont de resultaten.

Tabel 3: Gezichtsherkenningpercentages per fotodatabase voor Picasa, Fotobounce en PittPatt Facesort.

	Picasa (%)	Fotobounce (%)	PittPatt Facesort (%)
Yale Face	91	53	98
L. Spacek	100	81	100
Individueel verzamelde foto's	34	14	64

Uit deze gegevens blijkt dat de programma's de gezichten uit de ongecontroleerde database slechter herkennen dan de foto's uit de gecontroleerde databases. De ongecontroleerde database bevat foto's vanuit meerdere gezichtspunten, terwijl de foto's uit de gecontroleerde databases allemaal frontaal genomen zijn. Er is daarom ook gekeken hoe de programma's scoren op frontaal genomen foto's en foto's genomen vanaf de zijkant van de individueel verzamelde foto's. Onderstaande tabel geeft het aantal herkende foto's en het herkenningpercentage voor zowel de frontale als zijwaartse foto's voor deze database per applicatie.

Tabel 4: Het herkenningpercentage en aantal herkende frontale en zijwaartse foto's uit de database van individueel verzamelde foto's voor Picasa, Fotobounce en PittPatt Facesort.

	Picasa	Fotobounce	PittPatt Facesort
Frontaal herkend (aantal)	39	19	82
Zijwaarts herkend (aantal)	2	0	14
Frontaal herkend (%)	39 %	19%	82%
Zijwaarts herkend (%)	4%	0%	28%

Uit bovenstaande tabel blijkt dat het voor alle geteste applicaties moeilijker is om foto's waar slechts de zijkant van een gezicht op te zien is te herkennen dan foto's waarop het gehele gezicht te zien is. Dit kan meerdere redenen hebben. In de evaluatie/discussie zal hier op worden ingegaan.

4.2 Experiment 2

In totaal zijn 500 foto's gebruikt. Om gezichtsherkenning toe te passen op deze foto's was het van belang dat de gezichten gedetecteerd werden. De eigen applicatie heeft 530 gezichten gevonden in de gehele dataset. Echter, de gezichtsdetectie werkte niet foutloos; soms werden er ook niet-gezichten als gezichten beschouwd. Deze niet-gezichten zijn genegeerd in de resultaten.

Tabel 5: De 1st en 2nd tier recall en de positie van de *last ranked image* voor foto's van verschillende personen in de eigen applicatie. De referentiefoto van de gezochte persoon werd niet aan de trainingset van de applicatie toegevoegd.

	Recall 1st tier (%)	Recall 2nd tier (%)	Last ranked image
Individueel verzamelde foto's			
1st tier scope = 30; 2nd tier scope = 60			
Bill Clinton	30	43,3	359
Barack Obama	26,7	26,7	520
Ahmedinejad	3,3	6,6	465
Harry Potter	13,3	23,3	464
Martin Jol	23,3	30	441
Yale Face Database			
1st tier scope = 10; 2nd tier scope = 20			
Subject01	10	10	384
Subject02	50	70	107
Subject03	10	10	427
Subject04	70	80	64
Subject05	40	50	151
Subject06	20	60	38
Subject07	10	10	316
Subject08	30	60	314
Subject09	40	40	185
Subject10	10	30	302
Subject11	80	90	59
Subject12	10	10	507
Subject13	30	60	103
Subject14	20	20	62
Subject15	0	0	479
L. Spacek Database			
1st tier scope = 19; 2nd tier scope = 38			
Boylee	94,7	100	21
Cywan	94,7	100	21
Dfhodd	52,6	52,6	223
Howar	100	100	20
Jross	94,7	100	22
Kouri	36,8	42,1	158
Ldgodd	63,2	100	33
Mizli	36,8	78,9	44
Mkosto	26,3	36,8	147
Pjrand	84,2	89,5	121
Gemiddelde over alle databases	40,4	51,1	219

Tabel 5 bevat de gegevens die horen bij deel 1 van het tweede experiment (referentiefoto niet in de trainingset). Het 1st tier gemiddelde was 40,4%, het 2nd tier gemiddelde 51,1%, en de gemiddelde *last ranked image* stond op de 219^e plaats in de resultatenlijst.

Onderstaande tabel bevat de gegevens behorende bij deel 2 (referentiefoto wel in trainingset) van het tweede experiment.

Tabel 6: De 1st en 2nd tier recall en de positie van de *last ranked image* voor de verschillende personen in de eigen applicatie. De referentiefoto van de gezochte persoon werd wel aan de trainingset toegevoegd.

	Recall 1st tier (%)	Recall 2nd tier (%)	Last ranked image
Individueel verzamelde foto's			
1st tier scope = 30; 2nd tier scope = 60			
Bill Clinton	10	30	418
Barack Obama	30	36,7	513
Ahmedinejad	10	13,3	454
Harry Potter	10	16,7	514
Martin Jol	23,3	30	481
Yale Face database			
1st tier scope = 10; 2nd tier scope = 20			
Subject01	30	30	148
Subject02	50	60	373
Subject03	10	30	484
Subject04	80	90	25
Subject05	70	80	34
Subject06	50	80	156
Subject07	20	40	129
Subject08	60	60	244
Subject09	50	60	406
Subject10	50	80	105
Subject11	50	80	136
Subject12	30	70	470
Subject13	20	60	177
Subject14	20	30	338
Subject15	20	20	489
L. Spacek database			
1st tier scope = 19; 2nd tier scope = 38			
Boylee	94,7	100	21
Cywan	94,7	100	23
Dfhodd	57,9	73,7	173
Howar	100	100	20
Jross	94,7	100	22
Kouri	47,4	68,4	117
Idgodd	89,5	100	32
Mizli	89,5	100	28
Mkosto	36,8	47,4	147
Pjrand	89,5	94,7	80
Gemiddelde over alle databases	49,6	60	226

Interessant om te zien is dat de gemiddelde 1st tier van deel 2, ongeveer even hoog is als de gemiddelde 2nd tier van deel 1; het verschil bedraagt 1,5%. Daarnaast is het verschil tussen beide 2nd tiers 8,9%. De gemiddelde last ranked image ligt voor het tweede deel van het experiment 7 posities hoger dan de gemiddelde last ranked image van het eerste deel van het experiment. Tevens is duidelijk te zien in de gedetailleerde tabellen dat de foto's uit de individueel verzamelde foto's database minder goed herkend worden (lagere recall) dan de foto's uit de gecontroleerde databases. Dit komt overeen met de bevindingen uit experiment 1.

Hoofdstuk 5

Conclusie

Uit de resultaten blijkt dat de geteste applicaties beter presteren voor gecontroleerde omgevingen dan ongecontroleerde omgevingen. Daarnaast zijn de applicaties niet in staat om accuraat gezichten te herkennen waarvan alleen de zijkant te zien is. Dit heeft consequenties voor de bedoelingen achter dit onderzoek. Wanneer ongecontroleerd beeldmateriaal, het type materiaal dat van de slachtoffers verkregen wordt, op internet wordt gezet, is dit lastiger op te sporen dan gecontroleerde foto's. De kans dat al het compromitterende materiaal wordt gevonden is, wanneer gebruik gemaakt wordt van de software die in het huidige onderzoek gebruikt is, niet optimaal. Er kan dus geen garantie gegeven worden dat al het beeldmateriaal opgespoord zal worden. Aan de andere kant is de kans dat er helemaal niets wordt gevonden klein.

Al het materiaal dat gevonden wordt (ook al is het maar een klein percentage van het totaal), kan van het internet afgehaald worden. Als gevolg heeft de dader minder chantagemateriaal om het slachtoffer te dwingen om ongewenste handelingen te verrichten.

Het aantal foto's dat daadwerkelijk opgespoord wordt, is sterk afhankelijk van de software die gebruikt wordt en van het aantal referentiefoto's. Voor de drie geteste applicaties kan geconcludeerd worden dat het gebruik van een hoger aantal referentiefoto's leidt tot een hoger aantal gevonden relevante foto's. Wanneer een slachtoffer aangifte doet, is het aan te raden om foto's te maken van het slachtoffer en deze te gebruiken als referentiefoto's. Hierdoor zal het aantal gevonden foto's toenemen.

Het tweede experiment heeft bewezen dat er meer foto's worden gevonden wanneer er een foto van de gezochte persoon wordt toegevoegd aan de trainingset. Dit resultaat werd verwacht, aangezien de referentiefoto grotere overeenkomsten heeft met de gezochte persoon dan de andere foto's van willekeurige personen in de trainingset. Hierdoor is het gewicht behorende bij de referentiefoto hoger dan de gewichten van de andere foto's en wordt de persoon beter herkend.

In de praktijk betekent dit dat de foto's die bij een aangifte van het slachtoffer worden gemaakt niet alleen gebruikt moeten worden als referentiefoto's, maar ook moeten worden toegevoegd aan de trainingset van de applicatie die de foto's opspoor. Dit heeft consequenties voor de gebruikte software; de trainingset moet variabel kunnen zijn. Het moet mogelijk zijn om foto's toe te voegen aan of te verwijderen uit de trainingset. Dit was niet mogelijk in Google Picasa, Fotobounce en PittPatt Facesort. Kanttekening bij deze consequentie is dat het alleen van toepassing is op software die gebruikt maakt van Principle Component Analysis. Het is niet duidelijk of Google Picasa, Fotobounce en PittPatt Facesort berusten op deze techniek.

Hoofdstuk 6

Evaluatie/discussie

De onderzochte software is niet in staat om foutloos te werk te gaan. De resultaten voor gecontroleerde en ongecontroleerde omgevingen verschillen. In gecontroleerde omgevingen worden veel gezichten herkend, terwijl in ongecontroleerde omgevingen relatief weinig gezichten worden herkend. Voor de onderzochte software is het van belang dat de foto's goed belicht zijn voor optimale gezichtsherkenning.

Een bottleneck in de onderzochte software is de herkenning van zijwaartse foto's. Waar frontale foto's over het algemeen redelijk goed worden herkend, voornamelijk in gecontroleerde omgevingen, wordt slechts een laag percentage van het totale aantal zijwaartse foto's herkend. Een reden hiervoor is dat er minder informatie te verkrijgen is uit een zijwaartse foto dan een frontale foto. Door dit gebrek aan informatie is het lastiger om accuraat vast te stellen wie er op de foto staat, met als gevolg dat de applicaties niet zeker genoeg zijn en daardoor geen suggestie doen over wie er op de foto staat. Een andere reden is dat de gezichtsdetectie minder goed is voor zijwaartse gezichten. Om gezichten op foto's te detecteren worden *classifiers* gebruikt en deze *classifiers* zijn dikwijls ontwikkeld voor frontale foto's en niet voor zijwaartse foto's. In de praktijk kan het dus zo zijn dat er op een zijwaartse foto een gezicht is, maar dat de gezichtsdetectie dit niet herkent als een gezicht met als gevolg dat er geen gezichtsherkenning gedaan kan worden.

Uit experiment 2 is gebleken dat het toevoegen van een foto van de betreffende persoon aan de trainingset een positief effect heeft op de gezichtsherkenning. In het experiment werd de recall hoger wanneer een foto van de betreffende persoon deel uitmaakte van de trainingset. De last ranked image nam echter met 7 posities toe. Dit heeft als gevolg dat er iets langer gezocht moet worden voordat alle plaatjes gevonden zijn.

Belangrijk om te vermelden is dat er in het huidige onderzoek gebruik is gemaakt van gratis software. De mogelijkheid bestaat dat commerciële software andere, wellicht betere, resultaten opgeleverd zou hebben. De bevindingen van dit onderzoek zijn dus alleen van toepassing op de onderzochte software. Het zou te ver gaan om de resultaten door te trekken naar gezichtsherkenningsssoftware in het algemeen.

Een andere kanttekening die bij het huidige onderzoek geplaatst kan worden is dat de verschillende applicaties drempels gebruiken als 'zekerheidsindicator' voordat zij een naam koppelen aan een gezicht. Deze drempels verschillen per applicatie. Omdat de drempel voor Fotobounce niet in te stellen is, staan alle drempels op de standaardwaarde. Voor Picasa is dit 85 (van 0-100) en voor PittPatt Facesort is dit 7 (van 0-10). Aangezien niet nagegaan kan worden hoe de berekening van de drempelwaarde in zijn werk gaat, kan niet vastgesteld worden of een drempelwaarde van 80 in Picasa gelijk staat aan een drempelwaarde van 8 in PittPatt Facesort. Wel is duidelijk dat deze drempel van invloed is op de prestatie. In een vervolgonderzoek zou bijvoorbeeld de optimale

drempelwaarde bepaald kunnen worden en gemeten worden hoe sterk deze drempel de gezichtsherkenning beïnvloed.

Een ander interessant vervolgonderzoek zou zich kunnen richten op de optimalisatie van de referentiefoto's. Hoeveel referentiefoto's zouden moeten worden toegevoegd aan de trainingset voor optimale gezichtsherkenning? Hoe goed werkt de gezichtsherkenning als de trainingset alleen uit referentiefoto's bestaat? Van welke invalshoeken moeten deze referentiefoto's worden genomen? Worden zijwaartse foto's beter herkend wanneer zijwaartse foto's worden toegevoegd aan de trainingset? Hoe goed zal het werken op echt beeldmateriaal van slachtoffers, in plaats van kunstmatig testmateriaal? Dit zijn allemaal vragen die nog beantwoord moeten en zullen bijdragen aan een betere gezichtsherkenning.

Voor de uiteindelijke toepassing om foto's van slachtoffers op het internet op te sporen, zou ook het doorzoeken van het internet mogelijk moeten zijn. Verder zou er meer onderzoek moeten worden gedaan naar algoritmes die de gezichtsherkenning op onder- en overbelichte foto's verbeteren, naar algoritmes die de gezichtsherkenning op foto's waarop mensen staan die een hoofddekseel of bril dragen verbeteren en naar gezichtsdetectie algoritmes voor zijwaarts genomen foto's. Tot slot zou er een soortgelijk onderzoek als dit gedaan kunnen worden naar gezichtsherkenning in video's. Een mogelijkheid is om vanuit videodata een 3D reconstructie van het gezicht te maken, en die te vergelijken met willekeurige standen van het gezicht. Daar wordt op dit moment onderzoek naar gedaan door Universiteit Utrecht en het Nederlands Forensisch Instituut voor forensische toepassingen.

Literatuurlijst

Wetenschappelijke literatuur

X. Chai, S. Shan, X. Chen, W. Gao, "Locally Linear Regression for Pose-Invariant Face Recognition", *IEEE transactions on image processing*, vol.16, no. 7, pp. 1716-1725, 2007

P. Hennings-Yeomans, S. Baker, B.V.K.V. Kumar, "Recognition of Low-Resolution Faces Using Multiple Still Images and Multiple Cameras", *IEEE Second International Conference on Biometrics: Theory, Applications and Systems*, 2008

H. Ling, S. Soatto, N. Ramanathan, D. Jacobs, "Face Variation across Age Progression using Discriminative Methods", *IEEE transactions on information forensics and security*, 2010

U. Park, A. Jain, "Face Matching and Retrieval Using Soft Biometrics", *IEEE transactions on information forensics and security*, 2010

Websites

www.frtv.org^a

Geraadpleegd in januari 2011: <http://www.frvt.org/FRGC/>

www.frtv.org^b

Geraadpleegd in januari 2011: <http://www.frvt.org/FRVT2006/docs/FRVT2006andICE2006LargeScaleReport.pdf>

Gezichtsdata bases

Yale Face database

Geraadpleegd in december 2010: <http://cvc.yale.edu/projects/yalefaces/yalefaces.html>¹

Dr. L. Spacek database

Geraadpleegd in december 2010: <http://cswww.essex.ac.uk/mv/allfaces/index.html>

¹ Bij publicatie van dit rapport was deze database niet meer online beschikbaar. Een vergelijkbare database is Yale Face Database B. Deze is beschikbaar op <http://vision.ucsd.edu/~leekc/ExtYaleDatabase/Yale%20Face%20Database.htm>

