

Achtergrond

TUE plaatst tv-kijker tussen camera's in

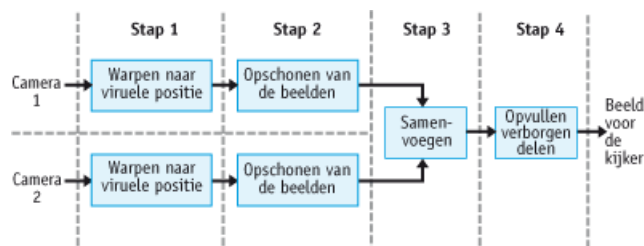
18 april 2010

3D-televisie komt binnenkort goed van de grond, verwachten analisten. Ondertussen wordt druk nagedacht over de volgende stappen. De Nederlandse en Franse partners in het Iglance-project denken dat vrije standpuntkeuze door de kijker daar een van zal zijn. TUE-onderzoeker Sveta Zinger legt uit hoe beelden van verschillende camerastandpunten geïnterpoleerd kunnen worden en hoe ontbrekende delen zijn in te vullen.

De 3D-rage lijkt voor de argeloze consument uit de lucht te komen vallen, maar wie de ontwikkelingen de afgelopen jaren heeft gevolgd, weet dat die al geruime tijd aan het rijpen is. Naast de voortschrijdende displaytechnologie zijn ook de gestage groei in rekenkracht van de tv's en settopboxen en de toename in opslagruimte belangrijke factoren die 3D-weergave mogelijk maken.

En natuurlijk moesten er methodes, technologieën en algoritmes worden ontwikkeld om met content in drie dimensies om te gaan. In oktober 2008 startte het Medea+-programma Iglance, bedoeld om de kenispelen van Eindhoven en Grenoble samen aan 3D-televisie te laten werken. Vanuit Nederland doen Philips Healthcare, Prodrive, Silicon Hive, Task24, de TUE en Verum mee. De Franse partners zijn 4D Solutions, Logica Frankrijk, STMicroelectronics en de onderzoeksinstituten Inria en Tima.

'Het doel is om te kijken hoe we de volgende stap in 3D kunnen zetten', vertelt Peter de With, die het project bij de TUE leidt. 'Dat is onze taak als universiteit; de eerste stap maakt de industrie zelf wel.' Die volgende stap is wat de initiatiefnemers *free viewpoint*-3D noemen, ofwel dat de kijker zelf kan bepalen vanuit welke hoek hij de situatie ziet. Dat gaat verder dan een keuze uit een aanbod van bestaande camerabeelden: de kijker kan zichzelf ook tussen twee camera's in plaatsen.



Bij het bepalen van een virtueel standpunt worden twee bestaande standpunten geïnterpoleerd naar de nieuwe positie, en vervolgens samengevoegd.

Eind dit jaar moet er een prototype bordje liggen dat de benodigde beeldbewerking realtime kan uitvoeren voor een tv-ontvanger. Op de TUE staat een demo die het voorlopige resultaat laat zien. Een 3D-scherm - een prototype van de brillose technologie die Philips een jaar geleden in de ijskast zette - toont een breakdancer die in drie dimensies zijn kunsten uitvoert. Plots verschuift de camera in een vloeiende beweging naar links. Helemaal perfect ziet het beeld er niet uit. Rond een van de schoenveters bijvoorbeeld zweven rare witte vlekken in de lucht. 'Dat komt omdat de diepte-informatie niet goed genoeg is', legt Sveta Zinger uit, die als postdoc verantwoordelijk is voor een groot deel van het free viewpoint-werk.

Gratis

In principe zijn free viewpoint-interpolatie en 3D-weergave twee verschillende dingen, maar zowel vanuit de methode als de toepassing ligt combinatie voor de hand. 'Je kunt het idee ook in 2D uitvoeren, maar wij zijn erg enthousiast over de 3D-kant', zegt De With. Waar de initiatiefnemers naartoe willen, is dat de kijker vanaf de bank straks bepaalt vanuit welke hoek hij een sportwedstrijd ziet. Of dat de regisseur van een film na de opnames nog virtueel met de camera's kan schuiven. Maar denk bijvoorbeeld ook aan trainingen in 3D-simulatoren of medische toepassingen.

Het virtuele standpunt wordt gereconstrueerd door twee videobeelden aan beide kanten van de kijker te interpoleren - extrapolatie is niet mogelijk. Daarvoor is het nodig om te weten hoe ver elke pixel van de camera af ligt. Deze informatie kan bijvoorbeeld tijdens de opnames verzameld zijn met sensoren of achteraf via beeldanalyse worden geschat. Bij 3D-animaties zit die informatie er 'gratis' bij.

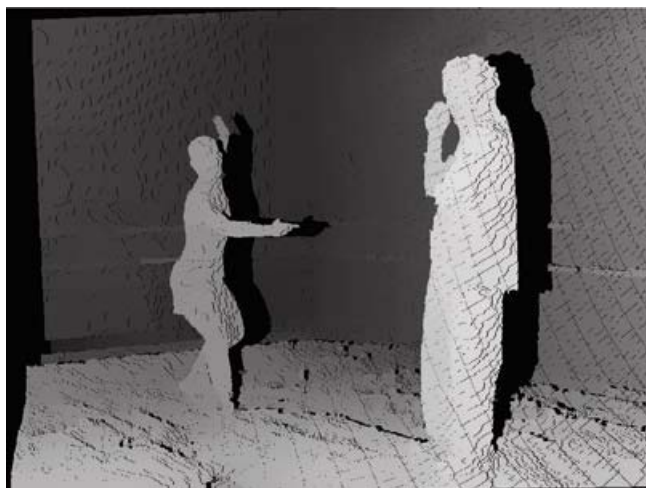
Daarnaast moet van alle camera's bekend zijn waar ze staan ten opzichte van elkaar en welke richting ze op kijken. Ook is er nog wat informatie nodig over de beeldparameters, zoals de brandpuntafstand. Op die manier kan het systeem voor elk virtueel standpunt bepalen welke camera's er het dichtst in de buurt staan en vervolgens een synthetisch beeld opstellen.

Gloed

Het opstellen van het nieuwe beeld gaat in twee stappen. Eerst worden beide beelden *gewarpt*: via een matrixtransformatie wordt becijferd hoe het eruit zou zien vanuit het virtuele standpunt. Vervolgens worden deze twee gewarpte beelden samengesmolten.

De interpolatie kent twee belangrijke bronnen van fouten: *warp cracks* en *occlusies*. De eerste hebben te maken met afronden. Van elke pixel worden tijdens de vervorming naar het virtuele standpunt de nieuwe coördinaten berekend als drijvende-kommagetal. Pixelposities moeten echter gehele getallen zijn. Tijdens het afronden verschuiven de pixels dus een beetje en kan het gebeuren dat er op sommige punten geen beeldpunten zijn. Daardoor ontstaan als het ware scheurtjes in het beeld.

Een vrij eenvoudige mediaanfiltering bleek voldoende om die scheurtjes weg te poetsen. Mediaanfilters zijn populaire algoritmes in beeldverwerking om ruis te verminderen. Van elke pixel worden de waarde en die van zijn burens op een rijtje gezet, waarna de mediaan wordt gekozen als nieuwe waarde. Een voordeel is dat dat relatief weinig rekenkracht vergt; er wordt alleen met gehele getallen gewerkt en er hoeft niet aan gerekend te worden. 'Bovendien laat het de randen intact', licht Zinger de keus verder toe. 'Er zijn heel veel andere methodes die *smoothing* hebben, zoals Gauss-filters. Dat betekent dat je randen verliest en daarmee kwaliteit van het beeld.'



Bij het herpositioneren van pixels in de nieuwe afbeelding ontstaan warp cracks. Deze zijn weg te poetsen met een vrij eenvoudig mediaanfilter.

Occlusies zijn fundamenteel van aard. Een camera kan nooit weten hoe het er achter een object uitziet. Als hij virtueel van positie verschuift, komen er dus delen in beeld waar geen informatie over is.

Het samenvoegen van beide geïnterpoleerde camerabeelden lost dat gedeeltelijk op; de camera's zien allebei immers een ander stukje achtergrond. Meestal blijven er echter nog stukjes over die geen van beide camera's zien. Om die in te vullen, hebben de onderzoekers een trucje bedacht. De missende informatie hoort doorgaans bij de achtergrond - de voorgrond staat immers in vol zicht van de camera. Vanuit elke pixel zonder waarde wordt in de rondte gezocht naar bestaande beeldpunten, en uiteindelijk wordt de pixel met de grootste diepte gekozen als surrogaat.

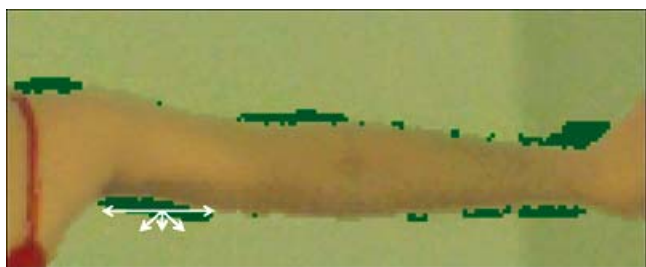
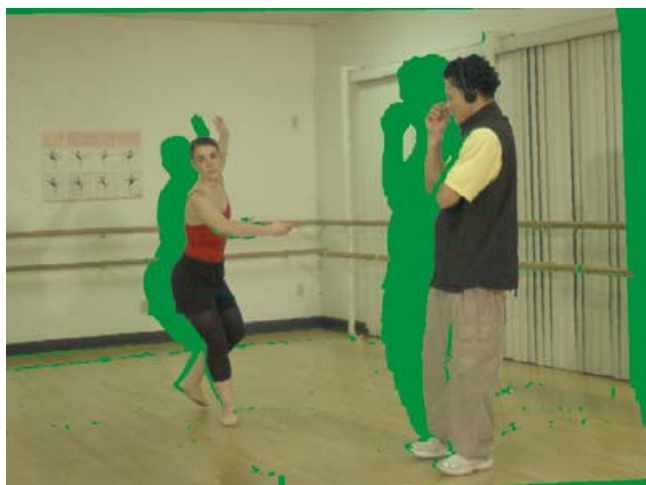
Er zijn meer storingsbronnen. Onbetrouwbare diepte-informatie resulteert in verstoringsen. Dat speelt vooral een rol als deze gegevens achteraf uit de beelden geëxtraheerd worden via beeldanalyse. Vooral bij de randen van objecten ontstaan er problemen. 'Randen zijn in beelden nooit echt abrupt maar altijd een beetje fuzzy', vertelt Zinger. Dat maakt het lastig om de uitsnede te bepalen. Bij een te ruime keuze komt een deel van de achtergrond mee met het object en kan een 'gloed' in de achtergrondkleur eromheen ontstaan. Bij een te nauwe uitsnede blijft er juist een deel van de rand op de achtergrond achter, een soort spookbeeld. De oplossing ligt in het gebruik van zowel een nauwe als een wijde uitsnede. Bij het wegpoetsen van het object wordt dit zo breed mogelijk uitgesneden, maar van het object zelf wordt alleen de nauwe uitsnede getoond.

Nét realtime

Om met voldoende kwaliteit te interpoleren, moeten de camera's niet al te ver uit elkaar staan. Voor een beetje fatsoenlijke keuze zijn dus flink wat camera's nodig, bij voorkeur in een halve cirkel om de scène heen. Dat is minder problematisch dan het lijkt. Bij grote evenementen staan nu al tientallen camera's en naarmate studio's zich bewuster worden van 3D, raken dergelijke opstellingen steeds meer in zwang.

Voor de ene opname zal de interpolatie ook soepeler verlopen dan voor de andere. 'Denk aan film. Daar heb je vaak bij dialogen een stabiele achtergrond, terwijl de voorgrond beweegt. Dat zijn situaties waarbij je over het algemeen zeer goed de diepte kan reconstrueren', vertelt De With. 'Bij sport loopt juist alles en iedereen door elkaar. Dat is veel complexer.'

Of het ook in alle gevallen gewenst is dat een kijker het standpunt kiest, is een ander verhaal. De Iglance-deelnemers zijn nauw betrokken bij de discussies van de MPEG-standaardisatiegroep over het hoe en wat van free viewpoint-bepalingen. Waarschijnlijk kan de contentleverancier straks aangeven in welke mate een kijker het standpunt kan aanpassen.



Een camera kan niet zien wat er achter een object is, dus bij het interpoleren vallen er gaten. Dat wordt opgevangen bij het samenvoegen van de twee geïnterpoleerde camerabeelden, die elkaar in de regel aanvullen. Om de laatste stukjes in te vullen, wordt in alle richtingen gezocht naar het meest op de achtergrond liggende beeldpunt.

Een ander onderdeel van het project is om in kaart te brengen wat er nou aan hardware nodig is. 'Dat zijn adembenemende getallen', zegt De With. 'Bij ST, dat het in zijn chipsets moet gaan verwerken, schrokken ze zich in het begin een hoedje. Maar bij de eerste 100-hertz-tv's was dat niet anders. Bij de introductie moesten er ineens hele rekenintensieve

taken worden uitgevoerd, terwijl dat tegenwoordig gewoon een subsysteempje in de chip is geworden.'

De

Voor implementatie op een FPGA is Silicon Hive verantwoordelijk. Op dit moment kunnen de algoritmes net realtime op het programmeerbare silicium draaien. De onderzoekers verwachten de komende tijd echter nog wel het een en ander te kunnen schaven aan de snelheid.

Ook op het gebied van de kwaliteit zal nog het een en ander verbeterd moeten worden. 'We zijn nog niet klaar', zegt De With. 'Als dat bordje er is, duurt het nog een paar jaar voordat het in tv's zit. En dat moet ook; eerst moet er geld verdiend worden aan normaal 3D.'

[Pieter Edelman](#)

[Terug naar overzicht](#)

TUE is uiteraard niet de enige die aan dit soort technologie werkt. Een stukje zuidelijker sleutelt bijvoorbeeld ook Imec aan virtuele camerastandpunten. In tegenstelling tot Iglance richt het onderzoekscentrum zich op programmamakers. Sportwedstrijden of concerten worden met enkele tientallen camera's rondom opgenomen, waarna de regisseur achteraf een willekeurig standpunt kan kiezen. Anders dan bij Iglance hoeft de virtuele camera dus niet strikt tussen twee camera's in te staan. 'De broadcaster kan bijvoorbeeld na een voetbalmatch de hele wedstrijd vanuit het standpunt van de topscorer genereren', noemt projectleider Johan de Geyter als voorbeeld.

Het systeem gebruikt standaard HD-camera's, veertig in totaal voor een voetbalwedstrijd. Diepte-informatie wordt via beeldverwerking verkregen. Daarmee zijn gelijk ook allerlei leuke statistieken te genereren, zoals de buitenspelijn of de snelheid van een speler. Tussen 12 en 15 april toont Imec de mogelijkheden op de NAB-show voor broadcasters in Las Vegas.

© Bits & Chips | Deze pagina op internet: <http://www.bits-chips.nl/nieuws/achtergrond/bekijk/artikel/tue-plaatst-tv-kijker-tussen-cameras-in.html>