Entwicklungen bei sehenden Rechnern

Autor(en): **Novotný, Radomir**

Objekttyp: **Article**

Zeitschrift: **Bulletin.ch : Fachzeitschrift und Verbandsinformationen von Electrosuisse, VSE = revue spécialisée et informations des associations Electrosuisse, AES**

Band (Jahr): **105 (2014)**

Heft 10

PDF erstellt am: **31.05.2024**

Persistenter Link: https://doi.org/10.5169/seals-856306

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.
Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.
Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

Ein Dienst der *ETH-Bibliothek*
ETH Zürich, Rämistrasse 101, 8092 Zürich, Schweiz, www.library.ethz.ch

http://www.e-periodica.ch

TECHNOLOGIE **PANORAMA**

Entwicklungen bei sehenden Rechnern

European Conference on Computer Vision ECCV in Zürich

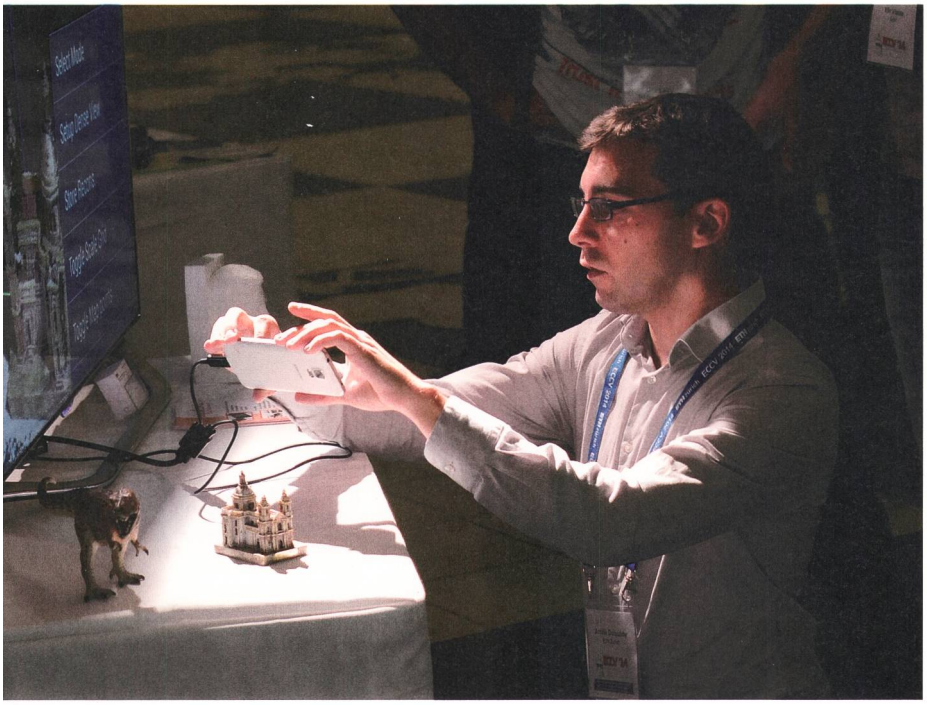
Seit 1990 findet alle zwei Jahre die europäische Konferenz für Computer Vision statt. Zürich war der Gastgeber der 13. Austragung dieser Top-Konferenz. Vom 6. bis 12. September trafen sich Forscher, um aktuelle Erkenntnisse zu diskutieren. Im Gespräch blickt Prof. Marc Pollefeys, ETH Zürich, auf bisherige Errungenschaften zurück und weist auf kommende Trends hin.

Radomír Novotný

Seit Jahren befasst sich ETH-Informatikprofessor Marc Pollefeys mit der Bildverarbeitung. Als Mitorganisator der ECCV hat er Einblick in eine äusserst vielfältige Welt der Computer Vision, denn insgesamt gingen bei den Veranstaltern 1444 Papers ein, von denen 363 angenommen und in Poster Sessions und Vorträgen interessierten Forschern erläutert wurden.

Bulletin SEV/VSE: Was sind die Themen der ECCV 2014?

Marc Pollefeys: Heute geht es hauptsächlich darum, Informationen aus Bildern zu extrahieren. Früher war man bezüglich der Art der Informationen ziemlich eingeschränkt; nun hat man die Möglichkeit, geometrische Informationen, Formen, Bewegungen, Objektarten und Szenen zu erkennen. Als Beispiel wurde ein Automobil-Projekt von Daimler vorgestellt, bei dem nicht nur Fussgänger erkannt werden, sondern auch versucht wird, ihre Absichten zu deuten, indem man beispielsweise ihre Blickrichtung berücksichtigt. Bei vielen Projekten wurde die Detailtiefe deutlich gesteigert, indem verschiedene Methoden kombiniert wurden. Zudem sind heutige Lösungen auch fähig, in «unkontrollierten» Umgebungen, d.h. in Umgebungen, bei denen beispielsweise die Position der Objekte frei ist, zu funktionieren. Wir sind aber noch weit von der menschlichen visuellen Wahrnehmung entfernt. Sie dient uns heute als Inspiration. Man weiss, was möglich wäre. Der Mensch kann mit seinem Gehirn, das nur 20 W verbraucht, sehr anspruchsvolle visuelle Aufgaben lösen. Der Heilige Gral wäre sozusagen eine Lösung, die bei gleicher Leistungsaufnahme wie das Gehirn eine vergleichbare visuelle Leistungsfähigkeit aufweisen würde. Das ist das Ziel.



No

Eine ETH-Lösung ermöglicht das Generieren von 3D-Modellen mittels Smart Phone.

Gibt es Fortschritte bezüglich praktischen Anwendungen?

Es gibt dank leistungsfähigen Rechnern neue Anwendungsmöglichkeiten beispielsweise im erwähnten Automobilbereich und bei Smart Phones.

Gibt es Technologien, die hier an der ECCV in Zürich nicht präsent sind?

Ja, das Imaging im medizinischen Bereich. Dies ist eigentlich eine sehr spezifische Fragestellung, denn da geht es um ein Abbilden vom Inneren eines Lebewesens, das andere Lösungen erfordert. An unserer Konferenz befassen wir uns hauptsächlich mit dem Erfassen des «sichtbaren» Äusseren. Gewisse Lösungen sind natürlich verwandt und können in beiden Bereichen eingesetzt werden, aber die spezifischen medizinischen Methoden werden an entsprechenden Konferenzen behandelt.

Was sind die bisherigen grossen Meilensteine der Computervision?

In den 1990er-Jahren hat man an den Grundlagen gearbeitet: Wie extrahiert man geometrische Information aus Bildern auf automatische Weise? Man hat sich auch mit der Beschreibung von interessanten Punkten befasst und wie man sie in Bildern mit verschiedenen Perspektiven erkennen und verbinden kann. Im darauffolgenden Jahrzehnt wurde diese Methode der interessanten Punkte erfolgreich erweitert: Durch die Kombination solcher Verfahren konnten nun komplexere Probleme gelöst werden. Die Anwendbarkeit in Alltagssituationen wurde markant verbessert. Man hat sich zudem damit befasst, wie man Objekte und Objektklassen erkennen kann. Mit statistischen Verfahren konnte man Objekte auf flexiblere Weise beschreiben. Vieles davon wurde durch eine deutlich höhere Rechenleistung und Speicherkapazität er-

TECHNOLOGIE **PANORAMA**

Minh Tran, ETH Zürich

Mitorganisator Marc Pollefeys weist auf die Entwicklungen bei der Computer Vision hin.

möglicht. Im letzten Jahrzehnt kamen noch Machine-Learning-Techniken hinzu. Generische Beschreibungen wurden realisierbar und haben einen grossen Einfluss ausgeübt.

Wo liegen die aktuellen Herausforderungen?

Im Entwickeln von Algorithmen, die allgemein einsetzbar sind. Oft sind Computer-Vision-Algorithmen spezifisch auf eine bestimmte Situation zugeschnitten und müssen für breitere Anwendungsgebiete mit anderen Algorithmen kombiniert werden. Das Lösen generischer Aufgaben ist immer noch sehr schwierig. Die benötigte Datenmenge nimmt exponentiell zu, wenn man alle denkbaren Szenarien berücksichtigt. Die Suche nach der optimalen Repräsentation für vielseitig einsetzbare Lösungen ist die grösste Herausforderung.

Welche Anwendungen werden schon bald im Alltag anzutreffen sein, die heute erst im Labor existieren?

Ganz bestimmt Lösungen für die Automobil-Industrie, die helfen sollen, Unfälle zu vermeiden. Die Systeme werden immer leistungsfähiger bei der Erkennung von gefährlichen Situationen. Heute können die Systeme nur in einer beschränkten Anzahl Situationen eingreifen, denn man will absolut sicher sein, dass die Entscheidung stimmt. Ein abruptes Bremsen des Systems in Situationen, in denen es nicht nötig ist, wäre verheerend. Es ist immer ein Abwägen zwischen falschen Alarmen und wirklichen Gefahren. Da darf kein Fehler passieren, auch wenn unterschiedliche Szenarien berücksichtigt werden sollen. Künftig kommt noch das automatische Fahren hinzu, beispielsweise auf Autobahnen, wobei der Fahrer in Zweifelsfällen übernehmen muss. Weitere Anwendungen werden bei Mobiltelefonen stattfinden, die beispielsweise Zusatzinformationen liefern können – zur geografischen Orientierung (Weghinweise als Pfeile auf Bildschirmausschnitt) oder bei der Gerätewartung, indem die Tätigkeitsschritte dem konkreten Bild des Geräts überblendet werden. Man sieht dann direkt, welcher Knopf betätigt werden muss oder welche Schrauben gelöst werden sollen. Statt ein Youtube-Wartungsvideo anzuschauen, erhält man die Instruktionen am konkreten Objekt. Die physische Welt wird viel stärker mit der virtuellen verbunden sein.

An der ECCV begegnet man der Lightfield-Technologie. Was ist das?

Bei Lightfield-Kameras werden Systeme mit mehreren Linsen eingesetzt, um Lichtstrahlen statt eines einzigen Bildes aufzunehmen. Lichtstrahlen, die aus unterschiedlichen Richtungen kommen, werden getrennt aufgenommen. Dadurch erhält man eine Datensammlung, bei der man den Fokus nachträglich wählen und das gewünschte Bild berechnen lassen kann. Es gibt Systeme mit einer grossen Linse und vor der CCD einer Sammlung kleiner Linsen sowie Systeme, die nur kleine Linsen verwenden. Man gewinnt sozusagen statt einer zweidimensionalen Abbildung eine vierdimensionale, aus der gewünschte zweidimensionale Bilder beliebig kombiniert werden können. Es braucht selbstverständlich viel mehr Speicher, man hat aber andererseits viel mehr Möglichkeiten zur nachträglichen Bildbearbeitung. Zudem gewinnt man zusätzliche Tiefeninformationen.

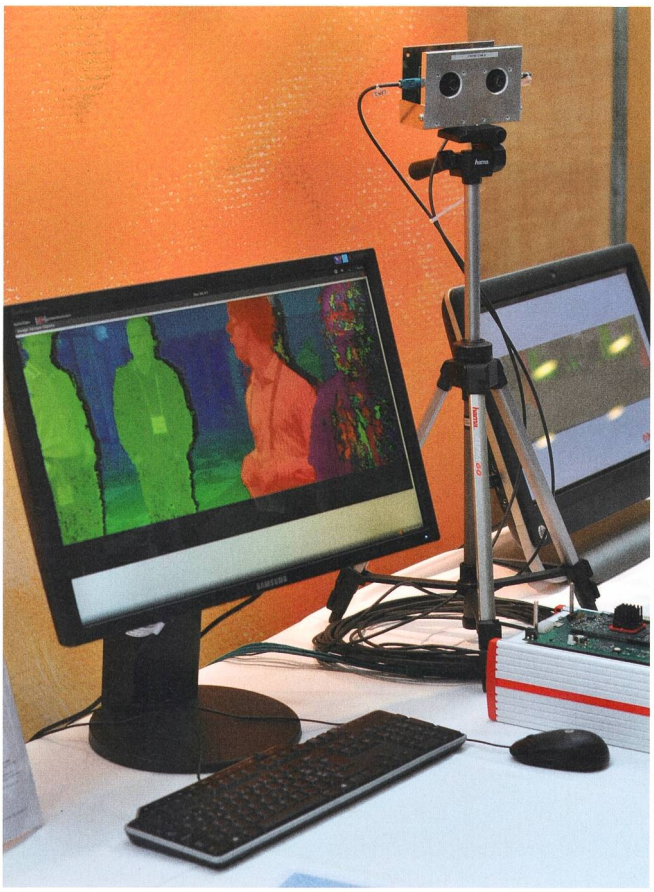
Sehen Sie diese Technologie eher im professionellen Bereich?

Nein, eher im Consumer-Bereich.

Aber das wird den benötigten Speicher vervielfachen. Wird es nicht den Rahmen sprengen?

Nicht unbedingt, denn es könnte sich um einen temporären Speicher handeln. Der Anwender (oder ein Algorithmus) wählt kurz nach dem Fotografieren den gewünschten Fokus aus und speichert das Bild im konventionellen Format – das 4D-Bild kann gelöscht werden. Die Auflösung könnte etwas kleiner sein, aber man gewinnt eine grosse Flexibilität beim Kombinieren der Brennpunkte.

Interview und Übersetzung aus dem Englischen: Radomír Novotný



No

Ein System von Super Computing Systems, das in ähnlicher Form in Fahrzeugen zur Erkennung von Fussgängern eingesetzt werden kann. Die Farbe der Personen symbolisiert ihre Distanz zur Kamera.