

Recognizing Cars

Louka Dlagnekov, Serge Belongie

André Justus

Seminar Visuelle Überwachung

Review

Kernaussage

In dem Paper wird ein System zur Identifizierung von Autos vorgestellt; hierzu werden sowohl das Nummernschild, als auch die Automarke und das Modell erkannt.

Relevanz in Bezug auf das Seminar

Es wird ein Verfahren vorgestellt, Autos zu entdecken und zu identifizieren. Damit wird eine weitere Möglichkeit geboten, Personen zu überwachen. Diskussionswürdig sind hierbei Nutzen und Risiken von mehreren kombinierten Überwachungssystemen, bei denen die Identifizierung von Autos eine Rolle spielen kann.

Der kürzlich vorgestellte AdaBoost Algorithmus wird bei der Entdeckung des Nummernschildes in einem Videostream verwendet. Somit dient der Beitrag nochmals zur Veranschaulichung des Algorithmus an einem Beispiel.

Das Paper steht außerdem im Zusammenhang mit dem Vortrag "Automatic License Plate Recognition", der am 24. Mai zu hören war; und gibt zusammen mit diesem, Einblicke in den aktuellen Forschungsstand, Möglichkeiten und Aussichten eines Teilbereichs der Visuellen Überwachung, der zumindest für Autofahrer in Zukunft sicher an Bedeutung gewinnen wird.

Zusammenfassung

Im Paper werden Verfahren dargestellt, die dazu dienen, Autos zu identifizieren. Dazu werden einerseits das Nummernschild, andererseits aber auch die Marke und das Modell des Autos verwendet. Zunächst besteht das Problem darin, überhaupt ein Nummernschild, bzw. dessen Position im Bild zu entdecken. Wurde nun ein Nummernschild entdeckt, wird versucht, es zu erkennen; d.h. es wird versucht, die Zeichen korrekt zu entziffern. Getrennt hierzu befasst sich das Paper mit der Erkennung von Marke und Modell. Hierzu wird, basierend auf der Position eines erkannten Nummernschildes, ein Bildausschnitt um das Nummernschild herum genommen, und mit zuvor erfassten Daten verglichen.

Das System ist hierbei an einen bestimmten Aufbau gebunden. Als Szenario diente ein Stoppschild, vor dem Autos gehalten haben. Die Kamera war so angebracht, dass man das Heck der Autos gut im Bild hatte. Allerdings wird keine teure oder spezielle Hardware benötigt. Verwendet wurde eine günstige Kamera der Auflösung 640x480, ohne weiteres spezielles Zubehör.

Die Nummernschilderkennung basiert auf dem AdaBoost Algorithmus, welcher anhand von 2400 Rechteckfeatures trainiert wird. Zur Erhöhung der Performance wurden Integralbilder und kaskadierte Klassifikatoren verwendet. Die Erkennungsrate beträgt 96,67% bei einer False Positive Rate von 0,002 %. Der Algorithmus kann 10 Bilder der Auflösung 640x480 pro Sekunde verarbeiten. Zu beachten ist, dass die False Positives im Zuge der Erkennung im nächsten Schritt noch einmal minimiert werden.

Bei der Nummernschilderkennung wurden Sequenzen von erkannten Nummernschildern extrahiert. Da die False Positives meist unbeständig waren, sind an dieser Stelle noch einige heraus gefallen, da Sequenzen, die keinen flüssigen Bewegungsablauf aufwiesen, verworfen worden sind. Diese Sequenzen wurden nach Nummernschildern abgesucht. Wird eins entdeckt, wird entweder ein neuer Tracker erstellt, oder aber das Nummernschild einem existierenden Tracker zugewiesen, wenn es im Frame davor an einer ähnlichen Position im Bild zu finden war. Ähnlich bedeutet, dass die Position des Nummernschildes im aktuellen Frame, im Vergleich zu dem vorherigen Frame

- einen Abstand kleiner eines Parameters ‚a‘, oder
- einen Abstand kleiner eines Parameters ‚b‘, wobei die Translation der bisherigen Bewegungsrichtung des Autos folgt.

Wobei $b > a$ ist.

Hat ein Tracker am Schluss weniger als 5 Frames, wurde er verworfen. Dies hat wiederum verbleibende False Positives eliminiert.

Die Zeichen eines so erkannten Kennzeichens werden dann anhand des „template matching“ Algorithmus gelesen. Template matching ist ein cross-correlation Ansatz, der alle vorkommenden Zeichen eines Trainings-sets mit allen segmentierten Zeichen des Nummernschildes vergleicht, und diejenigen mit der größten Übereinstimmung auswählt. Pro Zeichen, das im Nummernschild vorkommen kann, wurden 10 Muster für das Trainingset verwendet, so dass man insgesamt 360 Muster im Set hatte (26 Buchstaben, 10 Ziffern). Der Algorithmus benötigt etwa 0,5 Sekunden um ein Nummernschild zu lesen.

Für die Erkennung wurde ein Ähnlichkeitsmaß verwendet. Jede Abweichung eines Zeichens des gelesenen Nummernschildes von einer gesuchten Zeichenfolge bedeutet die Erhöhung des Abstandes um 1, außer das Zeichen war sehr ähnlich dem erkannten, dann gab es keinen Abzug. Eine Ähnlichkeitsgruppe innerhalb der ein Erkennungsfehler zu keiner Erhöhung des Abstandes führt ist z.B. {'O','0','D','Q'}.

Die Ergebnisse des Algorithmus werden auf Seite 5 des Papers gezeigt.

Die Erkennung von Marke und Modell erfolgt mit Hilfe des SIFT Algorithmus.

Durch diesen lassen sich markante Punkte eines Bildes extrahieren. Diese Punkte, Keypoints genannt, sind robust gegenüber Skalierung, Rotation und teilweise robust gegenüber einer Änderung der Lichtverhältnisse. Die Keypoints von einem Trainingsset, bestehend aus 790 Bildern von den Hecks verschiedener Autos, werden vorab berechnet. Von diesen Bildern ist außerdem die Information zu Marke und Modell des Autos verfügbar. Bei einem Vergleich zwischen einem Anfragebild und allen Bildern in der Datenbank, werden die Keypoints des Anfragebildes berechnet und mit den Keypoints aller Bilder in der Datenbank verglichen, so dass das Vergleichsbild mit den meisten Keypointübereinstimmungen die Marke und das Modell des getesteten Autos repräsentiert. Die Erkennungsrate liegt bei 89,5%, kann aber wahrscheinlich durch eine größere Menge von Vergleichsbildern noch deutlich gesteigert werden. Der SIFT Algorithmus performt allerdings noch nicht in Echtzeit. Hierzu wird gesagt, dass noch weitere Forschung nötig sei. Ein guter Ansatz sei es, die Autos nach Typen (z.B. Limousine, SUV, etc.) zu gliedern und den Algorithmus entlang einer Baumstruktur durch die Daten laufen zu lassen. Als weitere Verbesserung wird über die Berücksichtigung von Farbe und Kontur der Autos nachgedacht. Des Weiteren wird auch darauf hingewiesen, dass so ein System nicht nur auf das Stoppschild Szenario begrenzt sein soll.

Praktischer Einsatz

Ein System, wie es im Paper vorgestellt worden ist, wäre, wenn es besser funktionieren würde, gut in der Praxis einsetzbar. Gut geeignet wäre das zum Beispiel für den Polizeieinsatz, u. A. zur Fahndung von, mit dem Auto flüchtigen Personen. Auch automatische Gebührenerfassungssysteme sind denkbar. In den Niederlanden werden Temposünder anhand eines Nummernschilderkennungssystems automatisch zur Kasse gebeten. (vgl. [3]). So ein System sollte mit den Vorgestellten Möglichkeiten der Marken und Modellerkennung erweitert werden, um False Positives zu minimieren.

Diskussion

Im Paper wird eine gute Idee aufgezeigt, Autos sowohl anhand des Nummernschildes, als auch anhand der Informationen über Marke und Modell zu identifizieren. In diesem Zusammenhang ist auch die Gruppierung von ähnlichen Zeichen nützlich, um z.B. für eine Polizeifahndung effektiv eingesetzt werden zu können, wo sowohl Mensch als auch Maschine das Kennzeichen nicht 100%ig richtig erkennen müssen. Auch die Verwendung des vorgestellten Abstandsmaßes ist diesbezüglich eine gute Idee, da man Autos mit ähnlichen Kennzeichen anhand Marke und Modell soweit von einander unterscheiden könnte, dass man nur vergleichsweise wenige False Positives, bei einer hohen Erkennungsrate hätte. Allerdings funktioniert der Nummernschilderkenner meines Erachtens so schlecht, dass man damit kein sinnvoll funktionierendes System aufbauen kann. Kalifornische Kennzeichen haben sieben Zeichen. Der Nummernschilderkenner hat in den durchgeführten Tests eine Wahrscheinlichkeit von mehr als 30%, vier bis sieben der Zeichen falsch zu erkennen. Und dass, ohne Fehler innerhalb von Ähnlichkeitsgruppen zu berücksichtigen. Irritiert hat mich, dass die Fehlerrate des Algorithmus im Paper als durchaus gut beschrieben wird.

Der im Paper "Automatic License Plate Recognition" vorgestellte Nummernschilderkenner weist eine Erkennungsrate von über 90% für eine exakte Übereinstimmung auf. Der hier vorgestellte, im Vergleich dazu, gerade einmal 11,8%. Einzig die langsamere Verarbeitungsgeschwindigkeit von ca. 2,5 Sekunden pro Nummernschild (auf einem Pentium IV, 1,6 Ghz) ist ein Nachteil dieses Algorithmus. Durch fehlende Angaben über die Hardware ist allerdings kein quantitativer Vergleich möglich. Vielleicht ist es möglich, einen Algorithmus

zu finden, der Vorteile aus beiden Algorithmen zieht. Sollte dann die Verarbeitungsgeschwindigkeit das einzige Problem sein, könnte man dem mit spezieller Hochleistungshardware begegnen, um ein funktionierendes und robustes Echtzeitsystem aufzubauen.

Die Marken und Modellerkennung an sich liefert ganz gute Ergebnisse, die sicherlich mit den vorgestellten Verbesserungsvorschlägen sowohl von der Performance, als auch von der Erkennungsrate noch deutlich verbessert werden können. Dennoch bin ich skeptisch, ob das System auf "normaler Hardware" mit einem stark vergrößertem Vergleichsdatensatz in Echtzeit laufen wird, insbesondere wenn es nicht nur darum geht, die Autos von hinten zu erkennen, sondern aus jeder Position; so dass das System auch in beliebigen anderen Szenarien zu verwenden wäre. Dann wäre es außerdem sehr bedenklich, ob annähernd hohe Erkennungsraten, wie in dem vorgestellten statischen Szenario, erreicht werden können.

Fazit

Ich persönlich finde die Idee der Kombination aus einem Nummernschilderkenner mit Fehlertoleranz und einem Automarken- und -modellerkenner gut. Allerdings muss in dem Bereich noch viel geforscht werden, damit ein Erkennungssystem mit einer deutlich geringeren Fehlerrate und einem dynamischeren Einsatzgebiet entsteht, und damit öffentlich eingesetzt werden kann.

Risiken

Falls ein ausgereiftes System, basierend auf einer Weiterentwicklung der vorgestellten Technik bundesweit aufgebaut werden würde, wäre es möglich, Bewegungsprofile einer Person aufzuzeichnen und für Dinge zu ge-, bzw. zu missbrauchen die Außerhalb des (der Öffentlichkeit bekannten) eigentlichen Einsatzgebietes des Systems lägen. Ein einleuchtendes Beispiel wäre die aktuelle BND Affäre, bei der Journalisten durch Wanzen ausspioniert worden sind. Die Ausnutzung von Bewegungsprofilen wäre dabei ein Mittel das sicherlich auch benutzt worden wäre, würde ein solches System existieren.

Referenzen

- [1] Louka Dlagnekov, Serge Belongie. Recognizing Cars. UCSD CSE Tech Report CS2005-0833
- [2] Shyang-Lih Chang, Li-Shien Chen, Yun-Chung Chung, and Sei-Wan Chen. Automatic License Plate Recognition. IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, VOL. 5, NO. 1, MARCH 2004
- [3] Louka Dlagnekov. Video-based Car Surveillance: License Plate, Make, and Model Recognition, U.C. San Diego. Masters Thesis, 2005
- [4] http://vision.ucsd.edu/car_rec.html
- [5] David G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. Computer Science Department University of British Columbia Vancouver, B.C., Canada.