

Ein embedded Echtzeit-Videosystem für die Zählung von Fahrzeugen im Innen- und Außenbereich für den 24-Stunden-Betrieb

Ming-Yee Chiu, Setrix Inc.

Allgemeines

Beim neuen Setrix S-37000 handelt es sich um ein embedded Videosystem, mit integriertem CMOS Bildsensor in WebCam Qualität und einem RISC Prozessor um im Innen- und Außenbereich Fahrzeuge zuverlässig zu zählen. Die Herausforderung dieser Anwendung, speziell im Außenbereich war es, einen visuellen Algorithmus zu entwickeln, der sowohl bei Tag und Nacht als auch während des Hell/Dunkel-Übergangszeiten (z. B. Tagesanbruch und Dämmerung) funktioniert.

Das visuelle System ist außerdem in der Lage sich an verschiedenste Lichtverhältnisse anzupassen (von sonnigen Sommern bis zu Schnee im Winter). Das System ist sowohl im Innen- als auch in Außenbereich im Einsatz. Im Außenbereich wurde z. B. ein Projekt, bestehend aus 15 Video-Sensoren auf einem Firmengelände von ca. 1 km², realisiert. Das System ist im täglichen Einsatz und versorgt ein Parkleitsystem mit aktuellen Werten.

In einem der Projekte für den Innenbereich sorgen 7 Sensoren für die Zonen- und Gassenzählung in einer Parkgarage und bieten so mehr Flexibilität und Service.

1. Einleitung

Diese visuelle Anwendung wurde mit dem Ziel entwickelt, durch neue und innovative Sensorik, Betreibern von Parkgaragen oder Parkplätzen bessere Effizienz beim Betrieb und mehr Service für die Kunden zu ermöglichen. Autofahrer werden so über freie Parkplätze in Parkhäusern oder Bereichen/Etagen einer Parkgarage, sowie auf Parkplätzen informiert und geleitet.

Setrix bietet dazu einen visuellen und netzwerkfähigen Sensor zum Zählen von Fahrzeugen auf Straßen, Ein- und Ausfahrten von Parkplätzen und Parkhäusern, oder auch in Bereichen von Parkanlagen (Zonen und Gassen).

Traditionell wird eine Technik mit Induktionsschleifen, Ultraschallsensoren oder Lichtschranken für das Zählen von Fahrzeugen verwendet. Diese erprobten Techniken haben aber ihre Einschränkungen. Visuelle Sensorik hat den Vorteil einer einfachen Installation, verschleißfreiem Betrieb und hoher Lebensdauer. Aufgrund der modernen Systemarchitektur bietet die neue Sensorik mehr Flexibilität für Programmierung, Diagnose, Kommunikation und mehr Intelligenz beim Zählen von

Fahrzeugen und Fahrzeugklassen. Das visuelle embedded System von Setrix besteht aus einem integrierten CMOS Sensor in WebCam Qualität auf einer offenen Computer Architektur mit Standard Hardwarekomponenten, um ein preiswertes und effizientes System anbieten zu können. Damit werden neue Möglichkeiten für den Einsatz von Parkleitsystemen geschaffen, wie z. B. auf Firmengeländen, an Veranstaltungsorten oder aber auch für den preiswerten Einstieg in Zählsysteme. Es ist eine kostengünstige Alternative zur Einzelplatzüberwachung.

Die Software ist optimiert um 30 Bilder/Sek. zu verarbeiten. Mit dieser Verarbeitungsgeschwindigkeit können Fahrzeuge bis zu 90 km/h detektiert werden.

2. Funktionsprinzip

Das Prinzip der Fahrzeugzählung ist ähnlich wie bei der Lichtschranke. Wie in Abbildung 1 dargestellt, wird das Videosystem an einer erhöhten Position auf einer Seite der Fahrbahn montiert. Es ist auf einen speziellen Marker, der sich auf der Gegenseite nahe der Fahrbahn befindet, gerichtet

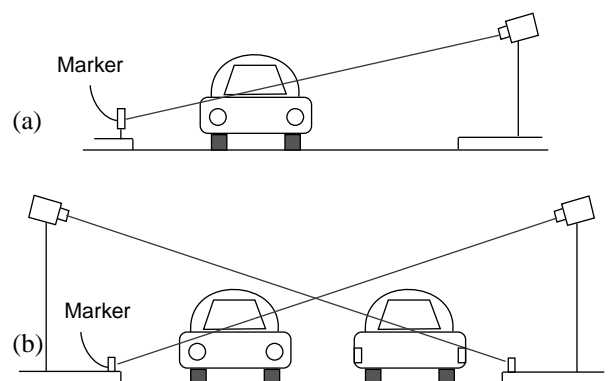


Figure 1 Geometrie für die Fahrzeugzählung

Die Linie zwischen Kamera und Marker bestimmt eine virtuelle "Sichtlinie". Wenn kein Objekt den Zählpunkt passiert, wird der Marker vom Videosystem detektiert. Sobald ein Objekt, Auto, LKW, Fußgänger oder ein Fahrrad diese Sichtlinie blockiert, bemerkt das System dieses unmittelbar. Mit zwei aufeinander folgenden Bildern werden dann der Typ des blockierenden Objektes

und die Bewegungsrichtung errechnet. Für erlaubte Fahrzeugklassen (z. B. Autos, LKWs,) wird dann die Zahl der Fahrzeuge in einer Parkanlage erhöht oder verringert, abhängig von der Bewegungsrichtung (ausfahrend – einfahrend). Fußgänger oder falls gewünscht Fahrräder und Motorräder können von der Zählung ausgeschlossen werden. Der visuelle Algorithmus ermöglicht Autos und LKWs von Motorrädern, Fahrrädern und Fußgängern zu unterscheiden.

Nach der Objektklassifizierung prüft das System ob der Marker wieder sichtbar ist. Falls nicht, wiederholt das System die Bewegungsverarbeitung und die Objektunterscheidung bis der Marker wieder sichtbar wird. Ein "Marker-Block-Ereignis" kann höchstens eine Objektzählung generieren.

Abb.1(a) Hier wird nur ein Videosystem für die Fahrzeugzählung in beiden Richtungen benötigt.

Abb. 1(b) Zwei Videosysteme werden für eine breite Fahrbahn benötigt. Falls ein großes Fahrzeug (LKW) die virtuelle "Sichtlinie" von beiden Videosystemen blockiert, wird durch eine Koordination beider Videosysteme in einem zusätzlichen Verarbeitungsschritt eine Doppelzählung ausgeschlossen.

3. Der 24-Stunden-Marker und seine Erkennung

Wie in Abb. 2 gezeigt benutzen wir ein spezielles von Setrix entwickeltes Schwarz/Weiß Muster als Marker. Das Muster besteht aus 8 strahlenförmigen Segmenten ausgehend von einem Zentrum. Dieses Muster ist ein spezieller Barcode, welcher gleichzeitig zur Identifizierung und Lokalisierung von Objekten benutzt werden kann [2]. Durch die Benutzung der Hough Transformation, kann das Muster auch in schwierigen Umgebungen zuverlässig detektiert werden. Wir nennen dieses Schwarz/Weiß Muster "Sunny-Pattern".

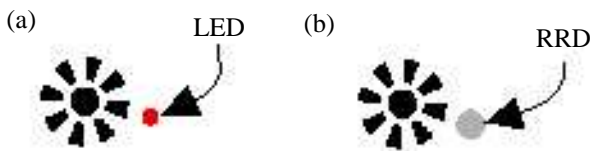


Abb. 2 Der 24-Stunden Marker

Die Bildauflösung für die Marker Detektion ist 320x240 Pixel. Während der anfänglichen Kalibrierungsphase wird das Sunny-Pattern zunächst durch die Verarbeitung des Bildinhaltes mit der Hough Transform Methode erkannt und lokalisiert. Fünf Parameter: die x-, y-Position, der Haupt- und Neben-Radius und die Ellipsen-Winkel des Musters werden ermittelt und

gespeichert. Diese Operation wird nur einmal ausgeführt. Für mehr Information zum Algorithmus siehe [2].

Abb. 3 zeigt die Marker Setups bei Installationen in der Praxis. Die Markerposition ist nah dem horizontalen Zentrum des Bildes. Das Bild zeigt auch die existierende Umgebung und unterschiedlichen Lichtverhältnisse in denen das Videosystem arbeitet.



Abb. 3 Die Platzierung der Marker in der Praxis

Für die Detektion des Markers im Betrieb benutzen wir einen vereinfachten Algorithmus, welcher nur die lokale Umgebung um das Sunny-Pattern auswertet.

Zuerst wird die Intensität des Profils entlang einer elliptischen Pfades abgetastet, der die schwarzen und weißen Segmente des Musters schneidet. Dazu werden die in der Kalibrierungsphase extrahierten fünf Parameter herangezogen. Danach werden die Kantenpositionen der 1-D Intensität-Profile extrahiert und geprüft, ob es 16 Übergänge mit ungefähr gleichem Abstand gibt. Wenn ja, wurde das Sunny Pattern erkannt.

Verschiedene abgetastete Intensitätsprofile mit unterschiedlichen Radien werden für eine robuste, zuverlässige Detektion ausprobiert. Außerdem wird die zentrale Position des Sunny Pattern von +1 bis -1 Pixeln in der x- und y-Richtung variiert, um Positionsänderung der Kamera bei z. B. Temperaturschwankungen zu handhaben.

Die Kantendetektion für die 1-D Intensitäts-Profile muss empfindlich genug sein, um auch die Kanten mit geringem Kontrast zu erkennen. Diese Situation tritt z. B. ein, wenn ein Sunny Pattern an einem sonnigen Tag im Schatten eines Baumes ist. Die Detektion benötigt ebenfalls eine Anpassung an wechselnden Schattenwurf von Bäumen auf den Marker.

In der Nacht bieten wir zwei Lösungen für den Marker. Eine Lösung ist eine batteriebetriebene LED (Abb. 2a). Die andere (Abb. 2b) ist ein Rückstrahler (Retro Reflective Disc – RRD), der den Strahl einer Reihe von Infrarot LED's in den Videosensor reflektiert. Die LED oder RRD sind an einer festen Position des Sunny Pattern befestigt. Die LED oder RRD sind während der Nacht kleine, leuchtende Punkte. Um die LED/RRD zu detektieren, suchen wir die Pixel mit maximaler Intensität nahe der eingebauten LED/RRD. Wenn diese Spitzenintensität (Peak) einen bestimmten Wert überschreitet, bewegen wir uns vom Ort mit der höchsten Intensität nach außen entlang von acht Winkelrichtungen um herauszufinden, wo die Intensität um einen festen Prozentsatz zum Peak abfällt. Wenn diese Punkte innerhalb einer kurzen Entfernung vom Peak sind, dann handelt es sich um ein LED/RRD. Dieser einfache Algorithmus ist schnell und effektiv. Gelegentlich jedoch kann ein Lichtreflex von vorbeifahrenden Fahrzeugen in der LED/RRD Suchregion auftreten und als LED/RRD identifiziert werden. Dies generiert fehlerhafte Marker-Erkennungen und teilt ein Marker-Blockierungsereignis in zwei oder mehr Ereignisse auf. Zur Abhilfe dieses Problems prüfen wir die Dauer der Sichtbarkeit des Markers. Ist diese unter einem Grenzwert, dann vereinen wir diese Ereignisgruppe wieder in ein Ereignis.

4. Die Verarbeitung der Bewegungsrichtung der Objektunterscheidung

So bald der Marker blockiert ist startet der Bewegungsrichtungs-Prozess, der eine reduzierte Auflösung von 160x120 Pixeln (erweitert zu 160x128 Pixeln) verwendet.

Abb. 4(a) zeigt zwei aufeinander folgende Bilder eines vorbeifahrenden Autos.

Der Bewegungsrichtungs-Prozess basiert auf einem bekannten Bewegungs-Algorithmus [3] der von den meisten Video-Kompressions-Standards benutzt wird. Das Bild wird in 10x8 Makroblocks (MBs) mit je 16x16 Pixeln aufgeteilt. Der Bewegungsvektor wird basierend auf SAD (Sum of Absolute Difference) berechnet. Vor Berechnung der Bewegung führen wir einen Screening Prozess durch, der die MBs für Bilder mit wenig Veränderungen im Bildinhalt oder kleinen vertikalen Strukturen ignoriert – angedeutet durch einen Punkt in Abb. 4(b), weil die Bewegung hauptsächlich horizontal erfolgt.

Abb. 4(b) zeigt die berechneten X-Motion Komponenten der 10x6 MBs. Die Zahlen geben den Pixel-Shift zwischen zwei Bildern an und sind proportional zur Geschwindigkeit.

Da die Bewegung des Fahrzeuges eine Starrkörperbewegung ist, können wir durch die Selektion der MBs mit ungefähr der gleichen horizontalen Bewegung eine Konsistenzprüfung durchführen. Diese MBs werden "Konsistenz-Makroblocks" genannt.

Abb. 4(c) zeigt das Ergebnis. Wie sie aus diesem Beispiel entnehmen können, hat die Verwischung im Bild einen geringen Effekt auf das Ergebnis.



(a)

(b) X-Motion vor der Konsistenzprüfung:

2	.	-19	-17	0
0	-16	-18	-18	-16	-15
0	.	-17	-18	-18	.	-17	-18	-6	.
0	-16	-17	-19	-20	-18	-19	-15	.	-16
24	-16	.	-18	-19	-19	.	.	.	-18
.	-20	-18	-18	-19	-19
.	-15	-20	-20	-17
.	-19	-22	-23	.	-17

(c) X-Motion nach der Konsistenzprüfung:

.	.	-19	-17
.	.	-18	-18
.	.	-17	-18	-18	.	-17	-18	.	.
.	.	-17	-19	-20	-18	-19	.	.	.

```

. . . -18 -19 -19 . . . -18
. . . . . -20 -18 -18 -19 -19
. . . . . . . -20 -20 -17
. . . . . . . . . . .

```

Abb. 4 Bewertung der Bewegung in aufeinander folgenden Bildern

Während der Nacht beruht die Verarbeitung auf der Bewegung von Frontscheinwerfern oder Rücklichtern der Fahrzeuge. Abb. 5 zeigt einige Bilder für die Verarbeitung der Bewegungsrichtung bei Nacht. Wie man aus den Bildern entnehmen kann, werden Teile der Szene manchmal durch das Scheinwerferlicht anderer Fahrzeuge erhellt. Die dünnen horizontalen Linienmuster, die das Bild überlagern, werden vom Bildrauschen des CMOS Sensor verursacht (Verstärkung 7,5 mal höher als normal).



Abb. 5 Nachtszene für Richtungserkennung

Nachts wird eine andere Bewegungsverarbeitung ausgeführt als tagüber. Zuerst wird vom Algorithmus bestimmt, ob die Tagverarbeitung geeigneter ist. Dies geschieht durch das Prüfen, ob der Prozentsatz von hellen Pixeln (Pixel mit Intensität höher als ein Schwellwert) ein Limit überschreitet (z. B. 25%). Wenn ja, wie im mittleren linken Bild von Abb. 5 der Fall, wird die Tagverarbeitung ausgeführt. Ist dies nicht der Fall, wird geprüft, ob die Anzahl von hellen Pixeln für ein Scheinwerferlicht/Rücklicht groß genug ist. Wenn ja, gehen wir in den Scheinwerfer/Rücklicht-Modus und führen die gleichen Bewegungsbewertung und Konsistenzprüfungen aus. Normalerweise ist in dieser Situation die Anzahl von konsistenten MBs gering (siehe linke und rechte Bilder in Abb. 5)

Die Objektunterscheidung basiert auf der Geschwindigkeit und der Größe des blockierenden Objektes und verwendet die 10x8 MBs. Wenn die Geometrie der Installation (Abb. 1) richtig ausgewählt ist, haben die Autos oder LKWs konsistente MBs auf der linken und rechten Teil des Bildes, während bei Fußgänger oder

Fahrrädern die konsistenten MBs nur in der Mitte des Bildes erscheinen (Abb. 6). Deshalb teilen wir die 10x6 MBs in drei Regionen auf: (L)inks, (M)itte und (R)echts (Abb. 7). L- und R-Regionen nehmen 3 Spalten von MBs in Anspruch, während M die mittleren 4 Spalten besetzt. Es gibt 5 mögliche Objekttypen: (1) Fahrzeug_rechts(V2R), (2) Fahrzeug_links(V2L), (3) Fußgänger_rechts(P2R), (4) Fußgänger_links(P2L), und (5) nichts(N).



Abb. 6 Klassifizierung von blockierenden Objekten

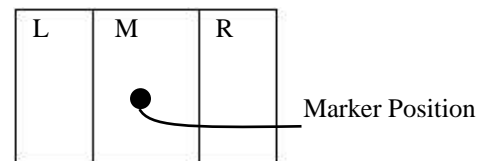


Abb. 7 Einteilung von 10x8 MBs in 3 Regionen

Die Regeln für die Objektunterscheidung sind:

- (1) Wenn die Geschwindigkeit kleiner ist als ein Grenzwert, haben wir Typ N.
- (2) Für Scheinwerfer-Verarbeitung gilt: wenn die Anzahl der konsistenten MBs kleiner 1 ist, haben wir Typ V2R oder V2L abhängig von der Bewegungsrichtung. Andernfalls haben wir Typ N.
- (3) Wenn die horizontale Ausdehnung von konsistenten MBs den Wert A (Autogröße) überschreiten und die Anzahl der konsistenten MBs B überschreiten, haben wir Typ V2R oder V2L.
- (4) Wenn die Anzahl konsistenter MBs in Region R und L den Wert C überschreitet und die Anzahl von konsistenten MBs B überschreitet, haben wir Typ V2R oder V2L.
- (5) Andernfalls ist das Objekt vom Typ P2R oder P2L, abhängig von der Bewegungsrichtung.

A, B, und C sind Grenzwertzahlen. Diese Unterscheidungsregeln nach "Regionen" führen nicht immer zu einem korrekten Resultat. Zum Beispiel kann ein schwarzes Auto bei geringen Lichtverhältnissen eine geringe Anzahl von konsistenten MBs haben (kleiner B) und deshalb als Fußgänger falsch klassifiziert werden.

Die Unterscheidung nach "Regionen" wird während der Markerblockadephase für jedes Bild ausgeführt. Das heißt wir verfeinern die Klassifizierung mit der Kombination der Verarbeitung über mehrere Bilder. Wir benutzen einen Wahlmechanismus um eine zuverlässigen Objekttyp zu erhalten, speziell für die Unterscheidung von Fahrrädern oder Motorrädern bis zu Fußgängern anstelle von Autos oder LKWs.

5. Der Algorithmus während der Tag/ Nacht-Übergänge

Abb. 8 zeigt Bilder während der Tag/Nachtübergänge (Morgen- und Abenddämmerung) welche in mittleren geographischen Breiten ungefähr 0,5–1 Std. dauern. Während dieser Zeit wechseln die Lichtverhältnisse dramatisch.



Abb. 8 Bilder während der Nacht/Tagwechsel Phase

Bei der Marker Detektion ist die LED/RRD Detektions-Methode nur aktiv, wenn die Szene beginnt dunkel zu werden. Dieser Zeitpunkt wird bestimmt durch die Überwachung der Parameter der automatischen Belichtungssteuerung des CMOS Sensors. Falls sie aktiv ist, wird zuerst die LED/RRD Detektion ausprobiert. Falls diese fehlschlägt, wird der Sunny Pattern Detektion ausgeführt. Diese überlappende Detektion arbeitet extrem gut.

Für den Übergang der Bewegungsverarbeitung benutzt das System den Zeitpunkt wenn die Szene dunkler wird. Für den Fall in Abb. 8 wird nur im links dargestellten Ereignis der Nachtalgorithmus ausgeführt.

6. Real-time Performance

Alle Echtzeit Videosysteme müssen sich der Herausforderung der unterschiedlichen Anforderung an die

Rechenleistung stellen. Diese Anwendung ist keine Ausnahme. Während der Markerblockadephase ist die Anforderung an die Rechenleistung groß, da die Berechnung von Bewegungsrichtung, Objektunterscheidung und Marker-Detektion für jedes Bild ausgeführt wird. Wir puffern die Bilder, so dass mehr Zeit für die Ausführung von Marker-Blockade-Ereignissen zur Verfügung steht. Der Algorithmus ist so implementiert das die Ausführung ausgewogen geschieht. Wir sind in der Lage 30 Frame/Sek. zu verarbeiten und auch in der Rush-Hour 30-40 Autos/Minute zu detektieren.

7. Das visuelle Sensorsystem

Das Videosystem (Abb. 9) hat eine Größe von 20x12x10 cm. Die Hardware besteht aus einer Kamera, einer RISC CPU, 16 MB Flash und 16 MB RAM. Der Stromverbrauch beträgt max. 2 Watt. Die Videosensoren können zu einem Parkleitsystem über ein RS-485 Interface verbunden werden. Einige Daten aus der Praxis: die Länge eines seriellen Kabels beträgt bis zu 4,2 km, zwei visuelle Sensoren an einem Messpunkt zählen täglich ca. 2.000 Fahrzeuge.

Die Erfahrungen aus den Installationen zeigen eine gute Zählgenauigkeit. Durch die ausgezeichneten Diagnosemöglichkeiten, können Erkenntnisse aus der Praxis, für unsere Weiterentwicklung der Software und Hardware genutzt werden.



Abb. 9 Das embedded Video-System

8. Referenzen

- [1] W. Wolf, B. Ozer, and T. Lv, "Smart cameras as embedded systems", IEEE Computer, 35(9), 2002, pp. 48-53.
- [2] M.Y. Chiu, T. Spindler, "Hough Transform for Aggregate Shapes and the Design of a High-Density Barcodable Marker", submitted to CVPR 2004.
- [3] J.R. Jain and A.K. Jain, "Displacement measurement and its application in interframe image coding", IEEE Trans. Commu. COM_29, 1981, pp1799-1808.