

(12) **Österreichische Patentanmeldung**

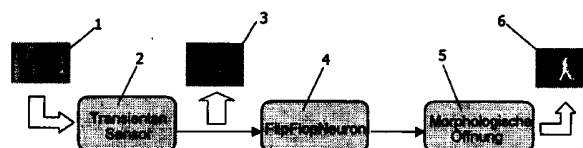
(21) Anmeldenummer: **A 1426/2005** (51) Int. Cl.<sup>8</sup>: **G01D 3/032** (2006.01),  
(22) Anmeldetag: **31.08.2005** **H04N 3/15** (2006.01)  
(43) Veröffentlicht am: **15.02.2007**

(73) Patentanmelder:

ARC SEIBERSDORF RESEARCH GMBH  
A-1010 WIEN (AT)

(54) **VERFAHREN ZUR DETEKTION BEWEGTER OBJEKTE MIT DYNAMISCHER  
HINTERGRUNDUNTERDRÜCKUNG, SOWIE AKTIVES PIXEL FÜR OPTISCHE SENSOREN**

(57) Ein Verfahren zur Detektion bewegter Objekte mit dynamischer Hintergrundunterdrückung umfasst die Erzeugung eines Ausgangssignals vorgegebener Polarität bei Detektion einer Intensitätsänderung an einem aktiven Pixel und die Erzeugung eines Ausgangssignals entgegengesetzter Polarität bei Detektion einer entgegengesetzten Intensitätsänderung am selben Pixel. Um bei wesentlicher Datenmengenreduktion durch Hintergrundadaption auch bewegte Objekte, und vor allem auch Objekte, die im Blickfeld der Kamera vorübergehend zum Stillstand kommen, zu erkennen, wird durch das Ausgangssignal bei Detektion einer Intensitätsänderung an einem aktiven Pixel die kontinuierliche Aussendung von zusätzlichen Ausgangssignalen gleicher Polarität eingeleitet. Dazu weist ein aktives Pixel, umfassend einen Transienten-Sensor und einen analogen Schaltkreis zur Signalvorverarbeitung und zum asynchronen Senden über vorzugsweise das Adress-Event-Protokoll, eine bistabile Schaltung (4) auf, welche durch ein Ausgangssignal (A) mit vorgegebbarer Polarität des Transienten-Sensors (1) aktiviert wird und dann kontinuierlich Ausgangssignale (C) sendet



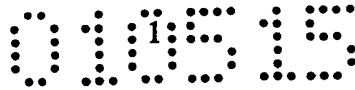
**Zusammenfassung:**

Ein Verfahren zur Detektion bewegter Objekte mit dynamischer Hintergrundunterdrückung umfasst die Erzeugung eines Ausgangssignals vorgegebener Polarität bei Detektion einer Intensitätsänderung an einem aktiven Pixel und die Erzeugung eines Ausgangssignals entgegengesetzter Polarität bei Detektion einer entgegengesetzten Intensitätsänderung am selben Pixel.

Um bei wesentlicher Datenmengenreduktion durch Hintergrundadaption auch bewegte Objekte, und vor allem auch Objekte, die im Blickfeld der Kamera vorübergehend zum Stillstand kommen, zu erkennen, wird durch das Ausgangssignal bei Detektion einer Intensitätsänderung an einem aktiven Pixel die kontinuierliche Aussendung von zusätzlichen Ausgangssignalen gleicher Polarität eingeleitet.

Dazu weist ein aktives Pixel, umfassend einen Transienten-Sensor und einen analogen Schaltkreis zur Signalvorverarbeitung und zum asynchronen Senden über vorzugsweise das Adress-Event-Protokoll, eine bistabile Schaltung (4) auf, welche durch ein Ausgangssignal (A) mit vorgegebbarer Polarität des Transienten-Sensors (1) aktiviert wird und dann kontinuierlich Ausgangssignale (C) sendet.

(Fig. 1)

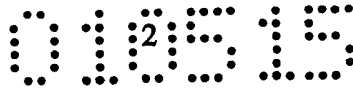


Die Erfindung betrifft ein Verfahren zur Detektion bewegter Objekte mit dynamischer Hintergrundunterdrückung, umfassend die Erzeugung eines Ausgangssignals vorgegebener Polarität bei Detektion einer Intensitätsänderung an einem aktiven Pixel und die Erzeugung eines Ausgangssignals entgegengesetzter Polarität bei Detektion einer entgegengesetzten Intensitätsänderung am selben Pixel, ein aktives Pixel für optische Sensoren, umfassend einen Transienten-Sensor und einen analogen Schaltkreis zur Signalvorverarbeitung und zum asynchronen Senden über vorzugsweise das Adress-Event-Protokoll, sowie einen Matrix-Sensor mit dynamischer Hintergrundunterdrückung, umfassend eine Anzahl von in mehreren Zeilen angeordneten aktiven Pixel und eine Auswerte-Elektronik.

Eines der Hauptprobleme im Echtzeit-Überwachungsbereich (real-time surveillance) ist eine ausreichend schnelle und robuste Extraktion relevanter Objekte. Da es sich um einen frühen Schritt in der Datenverarbeitungskette handelt, hat das kritischen Einfluss auf die Leistungsfähigkeit des Gesamtsystems. Üblicherweise erfolgt eine Subtraktion des statischen Hintergrunds (background subtraction => foreground segmentation), wie beispielsweise in "Foreground segmentation using adaptive mixture models in color and depth", M. Harville, G. Gordon and J. Woodfill, Proceedings of the IEEE Workshop on Detection and Recognition of Events in Video 2001, pp. 3-11, beschrieben ist. Trotz dessen Wichtigkeit gilt dieses Problem in der klassischen Bildverarbeitung noch immer als schlecht gelöst.

Die meisten gebräuchlichen Algorithmen basieren entweder auf der Analyse der Video-Bilder unter Berücksichtigung verschiedener Merkmale wie Farbe, Bewegung, Textur, usw., wofür ein Beispiel in "A robust spatio-temporal video object segmentation scheme with prestored background information," J. Pan, C. Lin, C. Gu, and M. Sun, Proc. IEEE Int. Symp. Circuits and Systems, pp. 803-806, May 2002, beschrieben ist, oder verwenden verschiedene Pixel-Statistikmodelle, wie etwa "Background subtraction techniques", A. McIvor, Proceedings of Image & Vision Computing New Zealand 2000(IVCNZ'00). Beide Methoden sind sehr rechenaufwendig was eine Anwendung solcher Algorithmen im Echtzeit-Überwachungsbereich sehr einschränkt.

Bekannt sind andererseits neuromorphe/bioinspirierte Sensoren wie Transienten-Sensor oder Silicon-Retina als potentielle Grundlage für weitere Entwicklungen, die intrinsische Eigenschaften wie Adaption und Kontrastsensitivität (Invarianz gegenüber Lichtverhältnissen) bieten. Ein Transienten-Bildsensor, wie beispielsweise in den folgenden Publikationen von T. Delbrück and C. A. Mead: "Analog VLSI phototransduction by continuous-time, adaptive, logarithmic photoreceptor circuits, in Vision Chips: Implementing vision algorithms with analog VLSI circuits", C. Koch and H. Li editors, IEEE Computer Society Press, pp. 139-161, 1995 ; T. Delbrück and C. A. Mead: "Time-derivative adaptive silicon photoreceptor array", Infrared Sensors: Detectors, Electronics, and Signal Processing,



1541: 92-99, Jay Jayadev, T.S. (Eds.), 1991 und J. Kramer: "An integrated optical transient sensor", IEEE Transactions on Circuits and Systems II, Analog and Digital Signal Processing, 49:(9) 612-628, Sep, 2002, beschrieben, reagiert nur auf die Änderung des Intensitätswerts eines Pixels, und sendet diese Information asynchron über das so genannten Adress-Event-Protokoll, erläutert etwa in "VLSI Analogs of Neuronal Visual Processing: A Synthesis of Form and Function", Misha Mahowald, Ph.D. Thesis, Computation and Neural Systems, California Institute of Technology, 1992. Dadurch werden nur bewegte Kanten im Bildfeld sichtbar und somit erfolgt automatisch eine dynamische Unterdrückung/Ausblendung des unbewegten Hintergrunds.

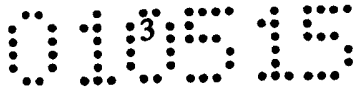
Allerdings bringt das auch wesentliche Nachteile mit sich, da die bewegten Objekte durch ihre Umrisse/Kanten dargestellt werden, d.h. Objekte aus dem Nahblickfeld des Sensors zerfallen in viele nicht zusammenhängende Konturen, was das Clustering/Gruppieren erheblich erschwert. Überdies ist der Transienten-Sensor nicht in der Lage Objekte zu erkennen die plötzlich zum Stillstand kommen. Besonders in dem Zusammenhang mit Tracking/Verfolgung entpuppt sich das als ein wichtiges Problem (verschwinden von Personen, Fahrzeugen, usw...).

Die Aufgabe der vorliegenden Erfindung war daher ein Verfahren sowie eine Vorrichtung anzugeben, die bei wesentlicher Datenmengenreduktion durch Hintergrundadaptation auch die Fähigkeit aufweist, bewegte Objekte, und vor allem auch Objekte die im Blickfeld der Kamera vorübergehend zum Stillstand kommen, als solche zu erkennen.

Zur Lösung dieser Aufgabe ist das eingangs beschriebene Verfahren dadurch gekennzeichnet, daß durch das Ausgangssignal bei Detektion einer Intensitätsänderung an einem aktiven Pixel die kontinuierliche Aussendung von zusätzlichen Ausgangssignalen gleicher Polarität eingeleitet wird. Die Interpretation der Bilddaten und eine genaue Untersuchung der detektierten Bildobjekte erfolgt anschließend in der nachgeschalteten Erkennungsstufe und sind nicht Gegenstand dieser Erfindung. Die Aussendung von Ausgangssignalen wird somit von der Vorderkante des Objektes gestartet und solange fortgesetzt, bis die nachlaufende Hinterkante des Objektes auftaucht. Auf diese Weise wird der Innenraum zwischen Objektkonturen in dem Transient-Sensor-Bildfeld befüllt und das Objekt flächenhaft dargestellt.

Vorzugsweise werden die zusätzlichen Ausgangssignale mit fester, vorzugsweise einstellbarer Taktrate gesendet.

Wenn gemäß einer vorteilhaften Ausführungsform der Erfindung die Aussendung der zusätzlichen Ausgangssignale durch das Ausgangssignal entgegengesetzter Polarität am gleichen Pixel wieder beendet wird, werden somit Adress-Events, vorzugsweise in einer



festen (einstellbaren) Taktrate, geliefert, bis ein gegenpoliges Adress-Event am Eingang auftaucht und die Schaltung zurücksetzt. Das erfindungsgemäße Pixel wird derart polaritätssensitiv betrieben und reagiert auf die Polaritätsänderung.

Gemäß einer weiteren vorteilhaften Ausführungsform der Erfindung ist vorgesehen, daß die Aussendung der zusätzlichen Ausgangssignale nach Ablauf einer einstellbaren Zeitspanne automatisch beendet wird. Damit ist ein zusätzlicher Interner Reset-Mechanismus gegeben, der das Zurücksetzen übernimmt, wenn über eine bestimmte Zeit (hängt vor allem von der charakteristischen Zeitskala der Bewegungen in der Szene) keine Aktivität (Events) am Eingang und Ausgang vorhanden ist.

Vorzugsweise wird auf die Gesamtheit aller Pixel die morphologische Öffnung angewendet und werden verbleibende einzelne nichtzusammenhängende Pixel weggefiltert. Dadurch werden die großen Strukturen der Bildmenge von innen abgerundet und kleine Strukturen beseitigt.

Ein aktives Pixel ist zur Lösung der eingangs gestellten Aufgabe erfindungsgemäß dadurch gekennzeichnet, daß der Schaltkreis eine bistabile Schaltung enthält, welche durch ein Ausgangssignal mit vorgebar polarität des Transienten-Sensors aktiviert wird und dann kontinuierlich Ausgangssignale sendet. Das erfindungsgemäße Pixel wird somit von der Vorderkante des Objektes aktiviert und bleibt so lange aktiv, bis die nachlaufende Hinterkante des Objektes auftaucht, um derart den Innenraum zwischen Objektkonturen zu befüllen und das Objekt flächenhaft darzustellen.

Vorteilhafterweise ist vorgesehen, daß die bistabile Schaltung in aktivem Zustand Ausgangssignale mit fester, vorzugsweise einstellbarer, Taktrate sendet.

Gemäß einer vorteilhaften Ausführungsform der Erfindung ist vorgesehen, daß die bistabile Schaltung durch ein Ausgangssignal mit entgegengesetzter Polarität des Transienten-Sensors wieder deaktivierbar ist. Das erfindungsgemäße Pixel ist also polaritätssensitiv und reagiert auf die Polaritätsänderung.

Vorzugsweise ist vorgesehen, daß die bistabile Schaltung einen Reset-Kreis enthält, der die Aussendung der zusätzlichen Ausgangssignale nach einer einstellbaren Zeitspanne automatisch beendet. Damit ist ein zusätzlicher interner Reset-Mechanismus gegeben, der das Zurücksetzen übernimmt, wenn über eine bestimmte, hauptsächlich von der charakteristischen Zeitskala der Bewegungen in der Szene abhängige, Zeit keine Aktivität am Eingang und Ausgang vorhanden ist.

Ein Matrix-Sensor (mit Zeilen- und Spaltenweiser Anordnung der Pixel) ist zur Lösung der erfindungsgemäßen Aufgabe gekennzeichnet durch aktive Pixel, die gemäß einem der vorhergehenden Absätze aufgebaut sind.

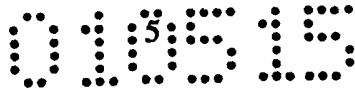
In der nachfolgenden Beschreibung soll die Erfindung anhand eines bevorzugten Beispiels und unter Bezugnahme auf die beigefügten Zeichnungen näher erläutert werden.

Dabei zeigt die Fig. 1 das Konzept der Bildverarbeitung in Form eines Ablaufdiagramms, Fig. 2 ist eine Darstellung der prinzipiellen Funktionsweise der Erfindung, Fig. 3 zeigt die Wirkungsweise des erfindungsgemäßen Verfahrens und der Vorrichtung für bewegte und unbewegte Objekte am Beispiel der Verarbeitung von Bildinformation, und Fig. 4 ist eine Darstellung der Funktion des erfindungsgemäßen aktiven Pixels in Form einer Finite-State-Machine.

Die Erfindung stellt das Konzept eines so genannten „hintergrundadaptiven“ Pixels mit Adress-Event-Interface dar, welches bereits eine entsprechende Nachverarbeitung der von einem Transienten-Sensor kommenden Address-Events beinhaltet. Das Konzept setzt sich aus drei in der Fig. 1 dargestellten Verarbeitungsstufen dar. Die grundlegende Idee dabei ist die Annahme dass jedes bewegte Objekt im Bildfeld 1 des Transienten-Sensors 2 eine Kontur erzeugt die sich aus zwei Kanten zusammensetzt (eine Vorder- und einer Hinterkante) mit entgegengesetzter Polarität (eine positive Intensitätswertänderung am Pixel wird von einer negativen gefolgt und umgekehrt). Der Transienten-Sensor reagiert auf die Änderungen des Intensitätswertes, d.h. es werden nur die Kanten des Objektes im Bildfeld sichtbar, so daß sich eine Darstellung 3 ergibt. Das Signal des Transienten-Sensors wird in der zweiten Verarbeitungsstufe, die ebenfalls noch im aktiven Pixel selbst erfolgt, durch ein sogenanntes Flip-Flop-Neuron (FFN) 4 verarbeitet.

Das Flip-Flop-Neuron 4 besteht aus einer bistabilen Schaltung, die von einem am Eingang eintreffendem Adress-Event zum laufenden Aussenden von Adress-Events angeregt wird. Diese Einheit 4 liefert Adress-Events, in einer festen (einstellbaren) Taktrate, bis ein gegenpoliges Adress-Event am Eingang auftaucht und die Schaltung zurücksetzt, d.h. die FFN-Einheit ist Polaritätssensitiv und reagiert auf die Polaritätsänderung. Dadurch wird der Innenraum zwischen den Objektkonturen befüllt. In der dritten Verarbeitungsstufe 5 schließlich, die entweder ebenfalls am aktiven Pixel vorgesehen oder aber in einer nachgeordneten digitalen Schaltung realisiert sein kann, wird das Ausgangssignal durch Anwendung der morphologischen Öffnung ausgebessert, so daß sich schließlich die Darstellung des Objektes wie in Bildfeld 6 ergibt.

Durch die Erweiterung des aktiven Pixels um ein sogenanntes Flip-Flop-Neuron, also einer Einheit mit Polaritätsgedächtnis-Effekt, die von der Vorderkante des Objektes aktiviert wird, und so lange aktiv bleibt bis die nachlaufende Hinterkante des Objektes auftaucht, ergeben sich die nachfolgend unter Bezugnahme auf die Fig. 2 erläuterten Effekte und Vorteile. Der Transienten-Sensor 2 reagiert wie im obersten Diagramm (a) dargestellt auf die Änderungen des Intensitätswerts am Pixel und generiert einen Address-Events-Datenstrom.



Eine positive Flanke erzeugt – wie im mittleren Diagramm (b) erläutert ist – ON-events (positive Ausgangssignale) A und die negative Flanke erzeugt OFF-Events (negative Ausgangssignale) B. Dieses Address-Event-Signal A, B wird von der FFN-Einheit 4 verarbeitet. Die FFN-Einheit 4 wird durch das erste eintreffende Event A aktiviert und wird erst durch das nachkommende gegenpolige Event B wieder zurückgesetzt, was im untersten Diagramm (c) der Fig. 2 dargestellt ist. Zwischen diesen Events A und B sendet die FFN-Einheit 4 Adress-Events C aus.

Zusätzlich kann durch einen internen Reset-Mechanismus gewährleistet sein, dass die FFN-Einheit 4 zurückgesetzt wird, wenn über eine bestimmte Zeit (hängt vor allem von der charakteristischen Zeitskala der Bewegungen in der Szene) keine Aktivität (Events) am Eingang und Ausgang vorhanden ist. Auf diese Weise wird der Innenraum zwischen Objektkonturen in dem Transient-Sensor-Bildfeld befüllt und man bekommt „spikende“ Flächen, wie sie auf der linken Seite in Fig. 3 dargestellt sind. Hier ist eine Momentaufnahme eines sich bewegenden Objekts dargestellt, wobei aus der Videoszene (obere Bilder) mittels Simulation eines Transienten-Sensors 1 AE-Daten berechnet wurden (schwarz-weiße Punkte in den mittleren Bildern) und anschließend das Bildfeld am Ausgang des Hintergrundadaptiven Sensor (weiße Punkte in den unteren Bildern).

Der zweite entscheidende Vorteil des erfindungsgemäßen Verfahrens ist, dass dadurch auch stehendbleibende Objekte im Bildfeld sichtbar bleiben, wie in Fig. 3 auf der rechten Seite zu erkennen ist. Hier ist eine Momentaufnahme zu dem Zeitpunkt dargestellt, wenn das Objekt stehen bleibt.

In der dritten und letzten Verarbeitungsstufe, werden verbleibende aktive Pixel (d.h. Pixel die zu keinem Objekt gehören, entstanden durch Rauschen oder fehlende gegenpolige Events, durch die Anwendung der morphologischen Öffnung mit entsprechender Maske (Filter-Kerne = kernel) weggefiltert.

Da das Flip-Flop-Neuron 4 einen Gedächtnis-Effekt aufweist, lässt sich die Funktionsweise auch in einem Zustandsdiagramm (siehe Fig. 4) darstellen. Die Kantenbeschriftung (IN/OUT) bezieht sich jeweils auf eine Änderung des Einganges (IN), die einen Zustandübergang auslöst und den zusätzlich am Ausgang ausgegebenen Wert (OUT). Der interne Reset-Mechanismus wird durch die Einführung der Zähl-Variablen (cnt) berücksichtigt.

Das vorgestellte Konzept kann grundsätzlich als Vorstufe dienen für alle höhere Bildverarbeitungsoperationen wie Clustering, Objektklassifizierung oder Verfolgung (tracking). In erste Linie wegen der hohen Zeitauflösung des Sensors, gegeben durch Datenmengenreduktion und asynchrone Informationsübertragung, eignet sich dieses Konzept besonders für zeitkritische Anwendungen. Weitere mögliche Einsatzgebiete sind

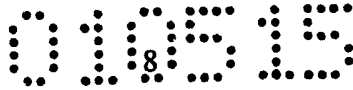
010515

Verkehrsüberwachung (Erkennung von Gefahrsituation entstanden durch plötzlich stehen bleibende Fahrzeuge) oder in der Sicherheitsüberwachung (zB. Erkennung von stehengelassenen Gegenständen).

Patentansprüche:

**Patentansprüche:**

1. Verfahren zur Detektion bewegter Objekte mit dynamischer Hintergrundunterdrückung, umfassend die Erzeugung eines Ausgangssignals vorgegebener Polarität bei Detektion einer Intensitätsänderung an einem aktiven Pixel und die Erzeugung eines Ausgangssignals entgegengesetzter Polarität bei Detektion einer entgegengesetzten Intensitätsänderung am selben Pixel, dadurch gekennzeichnet, daß durch das Ausgangssignal bei Detektion einer Intensitätsänderung an einem aktiven Pixel die kontinuierliche Aussendung von zusätzlichen Ausgangssignalen gleicher Polarität eingeleitet wird.
2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die zusätzlichen Ausgangssignale mit fester, vorzugsweise einstellbarer Taktrate gesendet werden.
3. Verfahren nach Anspruch 1 oder 2, dadurch gekennzeichnet, daß die Aussendung der zusätzlichen Ausgangssignale durch das Ausgangssignal entgegengesetzter Polarität am gleichen Pixel wieder beendet wird.
4. Verfahren nach einem der Ansprüche 1, 2 oder 3, dadurch gekennzeichnet, daß die Aussendung der zusätzlichen Ausgangssignale nach Ablauf einer einstellbaren Zeitspanne automatisch beendet wird.
5. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, daß auf die Gesamtheit aller Pixel die morphologische Öffnung angewendet wird und verbleibende einzelne nichtzusammenhängende Pixel weggefiltert werden.
6. Aktives Pixel für optische Sensoren, umfassend einen Transienten-Sensor und einen analogen Schaltkreis zur Signalvorverarbeitung und zum asynchronen Senden über vorzugsweise das Adress-Event-Protokoll, dadurch gekennzeichnet, daß der Schaltkreis eine bistabile Schaltung (4) enthält, welche durch ein Ausgangssignal (A) mit vorgegebbarer Polarität des Transienten-Sensors (1) aktiviert wird und dann kontinuierlich Ausgangssignale (C) sendet.



7. Aktives Pixel nach Anspruch 6, dadurch gekennzeichnet, daß die bistabile Schaltung (4) in aktivem Zustand Ausgangssignale (C) mit fester, vorzugsweise einstellbarer, Taktrate sendet.
8. Aktives Pixel nach Anspruch 6 oder 7, dadurch gekennzeichnet, daß die bistabile Schaltung (4) durch ein Ausgangssignal (B) mit entgegengesetzter Polarität des Transienten-Sensors (1) wieder deaktivierbar ist.
9. Aktives Pixel nach Anspruch 6, 7 oder 8, dadurch gekennzeichnet, daß die bistabile Schaltung (4) eine Reset-Kreis enthält, der die Aussendung der zusätzlichen Ausgangssignale (C) nach einer einstellbaren Zeitspanne automatisch beendet (Fig. 4).
10. Matrix-Sensor mit dynamischer Hintergrundunterdrückung, umfassend eine Anzahl von in mehreren Zellen angeordneten aktiven Pixel und eine Auswerte-Elektronik, dadurch gekennzeichnet, daß die aktiven Pixel gemäß einem der Ansprüche 6 bis 9 aufgebaut sind.

Wien, 31. August 2005

ARC Seibersdorf research GmbH  
in Wien (AT)

vertreten durch:

Patentanwälte  
KLEIN, PINTER  
& LAMINGER OEG  
Prinz Eugen Straße 70

Wien  
Mag. Norbert LAMINGER  
Patentanwalt



FIG. 1

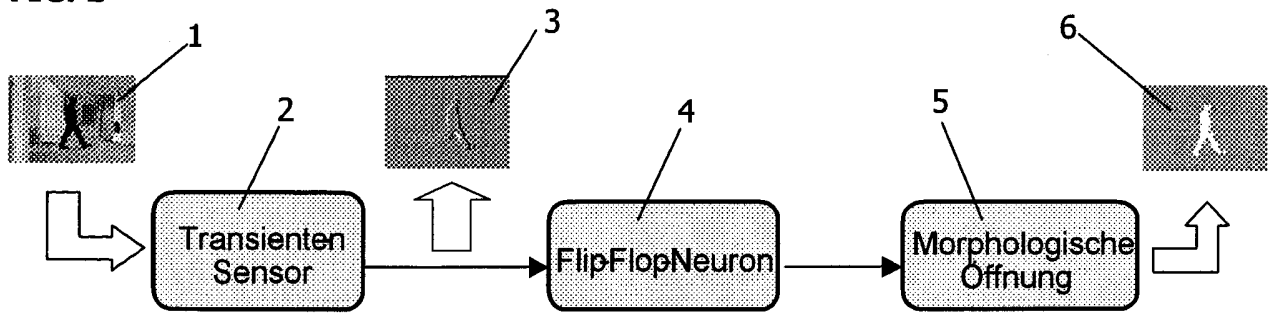


FIG. 2

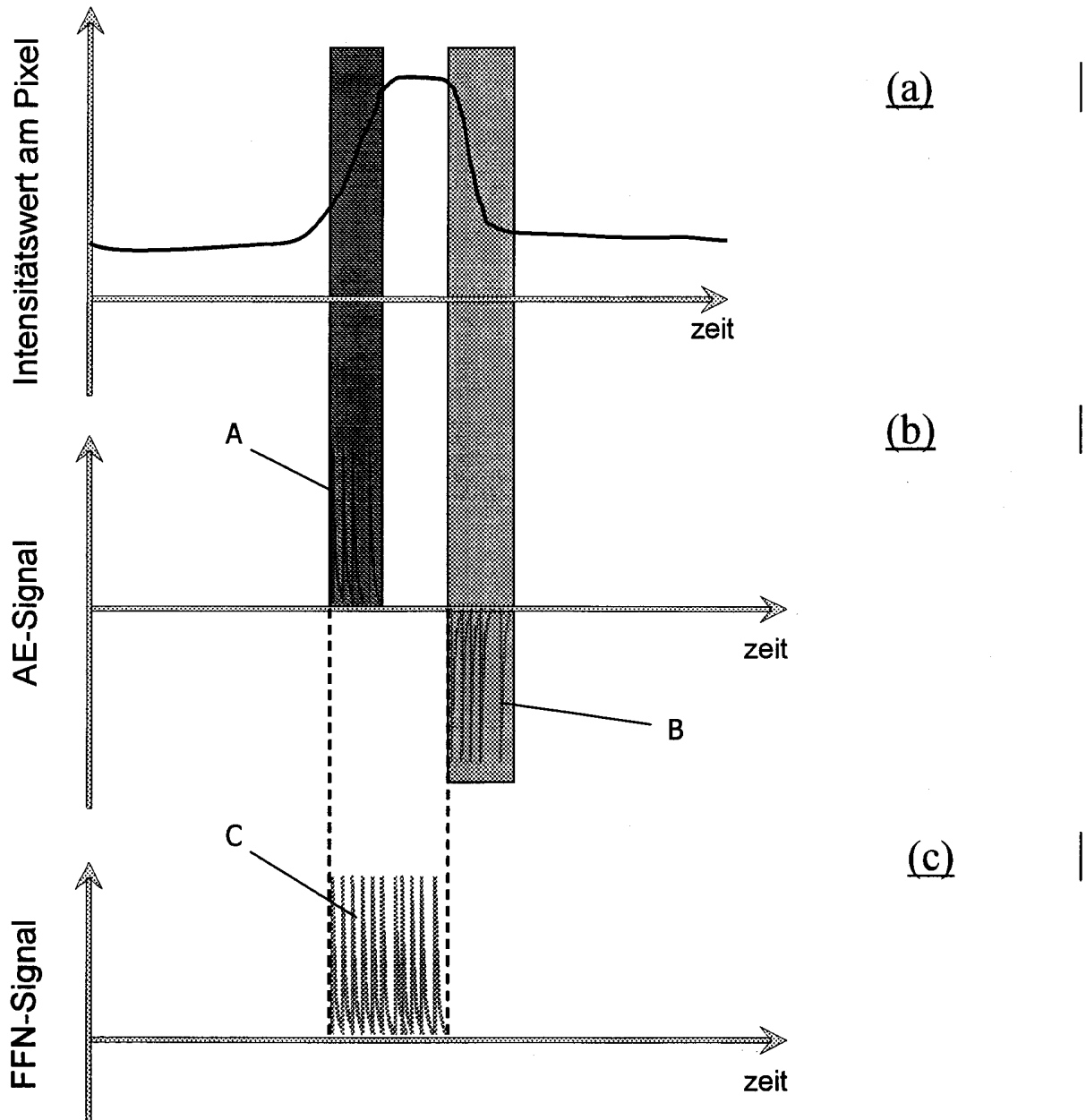


FIG. 3

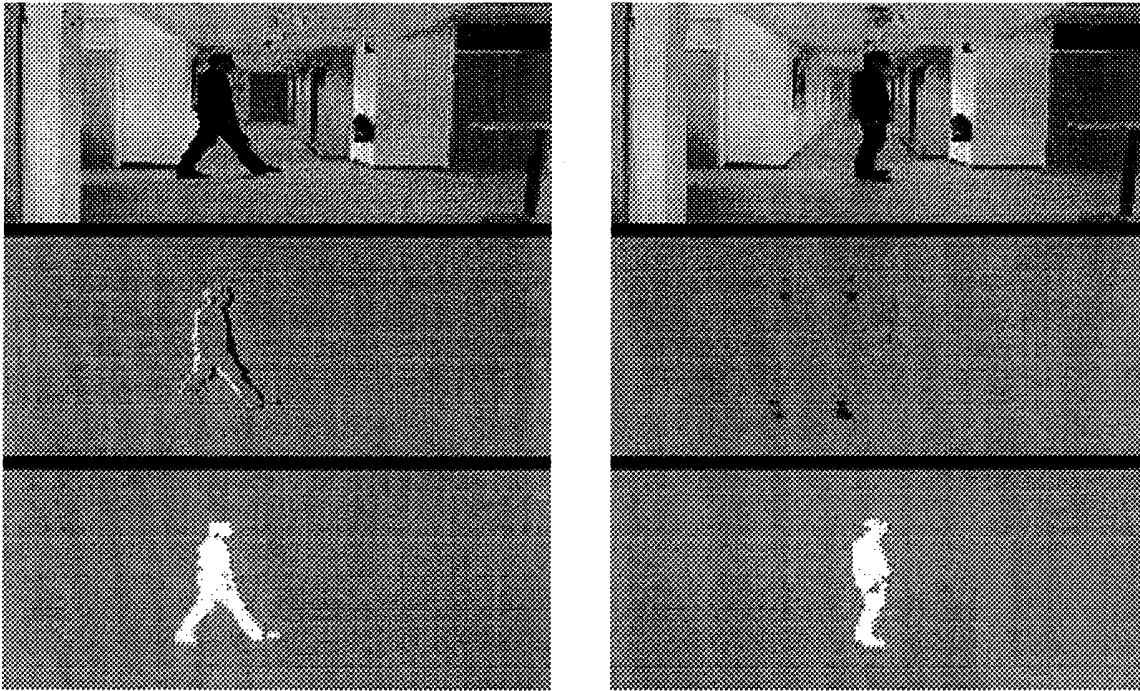


FIG. 4

