

(12)

Patentschrift

(21) Anmeldenummer: A 858/2008
(22) Anmeldetag: 28.05.2008
(45) Veröffentlicht am: 15.07.2012

(51) Int. Cl. : **H04N 7/18** (2006.01)
G08B 13/196 (2006.01)
G06K 9/34 (2006.01)

(56) Entgegenhaltungen:
WO 2005050972 A2
WO 2006046021 A1
US 2005078853 A1

(73) Patentinhaber:
KIWISecurity SOFTWARE GMBH
1040 WIEN (AT)

(72) Erfinder:
MATUSEK FLORIAN DIPL.ING.
WIEN (AT)
SUTOR STEPHAN DIPL.ING.
PERCHTOLDS DORF (AT)
RÖHR ROBERT DIPL.ING. (FH)
WIEN (AT)
KRAUS KLEMENS DIPL.ING.
WIEN (AT)
HÖRBINGER OLIVER
BADEN (AT)

(54) VERFAHREN ZUR VIDEOANALYSE

(57) Bei einem Verfahren zur Videoanalyse, wobei ein erstes digitales Videosignal, mit aufeinander folgenden digitalen Einzelbildern, in Echtzeit analysiert wird, wobei in jedem Einzelbild wenigstens ein Hintergrundbereich ermittelt wird, wobei gegenüber dem Hintergrundbereich wenigstens ein erster ortswechselnder Vordergrundbereich erkannt wird, und wobei nachfolgend eine erste Meldung erzeugt wird, wobei wenigstens ein erstes ortswechselndes Objekt des ersten ortswechselnden Vordergrundbereichs, mittels eines Trackers erkannt wird, und dessen Bewegungen mit dem Tracker verfolgt werden, wird zur Überwachung bzw. Steuerung einer großen Anzahl an Videokameras mit einer geringen Anzahl Bedienern vorgeschlagen, dass aus einer Gruppe vorgebbarer Tracker eine Mehrzahl Tracker zur Verfolgung der Bewegung des ersten ortswechselnden Objekts ausgewählt wird, dass nachfolgend die Bewegung des wenigstens einen ersten ortswechselnden Objekts mit jedem der ausgewählten Tracker verfolgt und ein Trackingergebnis ermittelt wird, dass zu jedem Tracker eine Fehlerwahrscheinlichkeit ermittelt wird, und dass nachfolgend die Trackingergebnisse der Tracker mit geringen Fehlerwahrscheinlichkeiten zu einem Trackingendergebnis zusammengefasst werden.

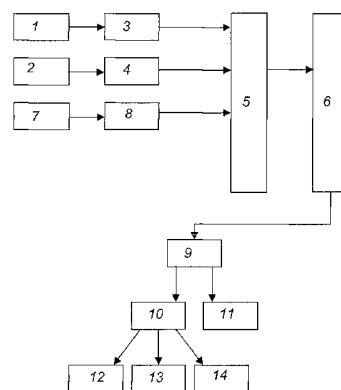


Fig. 1

Beschreibung

[0001] Die Erfindung betrifft ein Verfahren zur Videoanalyse gemäß dem Oberbegriff des Patentanspruches 1.

[0002] Zahlreiche Bereiche des öffentlichen Lebens werden mit Hilfe von Videokameras überwacht, um Verbrecher abzuschrecken bzw. Verbrechen aufzuklären. Die zunehmende Anzahl an Videokameras bringt jedoch das Problem mit sich, dass die derart aufgenommene Bilderflut nicht mehr ausgewertet werden kann. Eine ständige visuelle Überwachung sämtlicher Kameras ist in Anbetracht der Tatsache, dass in einem modernen Großobjekt, wie etwa einem internationalen Flughafen mehrere tausend Kameras installiert sind, rein personell undurchführbar. Ein Mehr an Kameras bringt jedoch nur einen Sicherheitsgewinn, wenn die aufgenommenen Videodaten auch gesichtet bzw. bewertet werden. Vor allem in sicherheitsrelevanten Bereichen, ist eine ständige Sichtung und Bewertung der Videodaten notwendig um einen Schutz von Menschen und Anlagen zu gewährleisten.

[0003] Die WO 2006/046021 A1 beschreibt ein System zur Videodatenverarbeitung, wobei eine Unterscheidung zwischen Vordergrund und Hintergrund vorgesehen ist, sowie eine Methode zum Erkennen von Schatten.

[0004] Aus der US 2005/0078853 A1 ist ein Verfahren bekannt, um einen sich ändernden Vordergrund von einem Hintergrund zu unterscheiden, wobei der Hintergrund ständig neu definiert wird.

[0005] Die WO 2005/050972 A2 offenbart ein videobasiertes Überwachungssystem sowie ein entsprechendes Verfahren, bei welchem zwischen einem Hintergrund und einem Vordergrund unterschieden wird. Die WO 2005/050972 A2 beschreibt weiters das Erkennen und Verfolgen eines ortswechselnden Motivs.

[0006] Aufgabe der Erfindung ist es daher ein Verfahren der eingangs genannten Art anzugeben, mit welchem die genannten Nachteile vermieden werden können, mit welchem eine große Anzahl an Videokameras mit einer geringen Anzahl Bedienern überwacht bzw. gesteuert werden kann, und welches unempfindlich gegen sich ändernde Hintergründe oder Beleuchtungen ist.

[0007] Erfindungsgemäß wird dies durch die Merkmale des Patentanspruches 1 erreicht.

[0008] Dadurch können Videodaten automatisch analysiert und bewertet werden, ohne dass die ständige visuelle Kontrolle durch einen wenigstens humanoiden Bediener notwendig ist. Dadurch kann das Augenmerk eines Bedienorgans auf bestimmte Vordergrundbereiche, welche etwa Objekte oder Gegenstände darstellen, gelenkt werden. Dadurch können auch komplexe Videoanlagen mit einer großen Anzahl an Kameras durch eine geringe Anzahl an Bedienern gesteuert und überwacht werden. Durch die ständige Ermittlung des Hintergrundes werden sich ändernde Hintergründe oder Beleuchtungsänderungen erkannt, und führen nicht zu fehlerhaften Meldungen eines Vordergrundbereichs oder eines Objekts.

[0009] Die Erfindung betrifft weiters einen Datenträger, auf welchem ein Programm zur Ausführung des in den Ansprüchen 1 bis 10 beanspruchten Verfahrens gespeichert ist.

[0010] Die Unteransprüche, welche ebenso wie die unabhängigen Patentansprüche gleichzeitig einen Teil der Beschreibung bilden, betreffen weitere vorteilhafte Ausgestaltungen der Erfindung.

[0011] Die Erfindung wird unter Bezugnahme auf die beigeschlossenen Zeichnungen, in welchen lediglich bevorzugte Ausführungsformen beispielhaft dargestellt sind, näher beschrieben. Dabei zeigt:

[0012] Fig. 1 ein Blockschaltbild einer erfindungsgemäßen Videokameraanordnung;

[0013] Fig. 2 ein Ablaufdiagramm einer bevorzugten Ausbildung des erfindungsgemäßen Verfahrens;

- [0014] Fig. 3 ein Ablaufdiagramm eines Tracker-Managers;
[0015] Fig. 4 ein Ablaufdiagramm eines bevorzugten Verfahrens zur codierten Speicherung einzelner Videoframes;
[0016] Fig. 5 eine Grundrissansicht einer Videokameraanordnung mit drei Videokameras;
[0017] Fig. 6 eine Grundrissansicht einer Videokameraanordnung mit zehn Videokameras;
[0018] Fig. 7 ein Modell der Grundrissansicht gemäß Fig. 6
[0019] Fig. 8 eine bevorzugte Bildschirmansicht eines Modells; und
[0020] Fig. 9 eine bevorzugte Darstellung eines dreidimensionalen Modells einer erfindungsgemäßen Videokameraanordnung.

[0021] Die gegenständliche Erfindung betrifft ein Verfahren zur Videoanalyse, insbesondere zur Videoüberwachung, wobei ein erstes digitales Videosignal, umfassend wenigstens mittelbar aufeinander folgende digitale Einzelbilder, in Echtzeit analysiert wird, wobei in jedem Einzelbild wenigstens ein Hintergrundbereich ermittelt wird, wobei gegenüber dem Hintergrundbereich wenigstens ein erster ortswechselnder Vordergrundbereich erkannt wird, und dass nachfolgend eine erste Meldung erzeugt, und vorzugsweise an eine Kontrollstelle gesendet wird.

[0022] Dadurch können Videodaten automatisch analysiert und bewertet werden, ohne dass die ständige visuelle Kontrolle durch einen humanoiden Bediener notwendig ist. Dadurch kann das Augenmerk eines Bedienorgans auf bestimmte Objekte gelenkt werden. Dadurch können auch komplexe Videoanlagen mit einer großen Anzahl an Videokameras 1, 2, 16 durch eine geringe Anzahl an Bedienern gesteuert und überwacht werden. Durch die ständige Ermittlung des Hintergrundes werden sich ändernde Hintergründe oder Beleuchtungsänderungen erkannt, und führen nicht zu fehlerhaften Meldungen eines Vordergrundbereichs und/oder eines Objekts.

[0023] Ein erfindungsgemäßes Verfahren bzw. eine erfindungsgemäße Videokameraanordnung bzw. Sicherheitsanlage ist zur Verwendung mit digitalen Videokameras 1, 2, 16 vorgesehen, wobei jede Art einer Videokamera 1, 2, 16 vorgesehen sein kann, welche ein digitales Videosignal ausgibt. Bei einem digitalen Videosignal handelt es sich bevorzugt um einen im Wesentlichen binären Datenstrom, welcher eine Reihe aufeinander folgender digitaler Einzelbilder, welche auch als Frame bezeichnet sein können, und gegebenenfalls ein Tonsignal aufweisen, umfasst.

[0024] Erfindungsgemäß ist das Erkennen eines Hintergrundes bzw. vorzugsweise eines ortswechselnden, daher bewegten, hinzukommenden oder verschwindenden, Vordergrundbereichs vorgesehen. Da auch das Zuordnen eines erkannten Vordergrundbereichs in vorgebbare Kategorien vorgesehen sein kann, daher ob es sich bei dem Vordergrundbereich um einen Menschen oder etwa ein Gepäckstück handelt, bzw. in weiterer Folge eine Zusammenführung mehrerer erfindungsgemäß bearbeiteter Videosignale zu einem Modell eines überwachten Bereichs vorgesehen sein kann, ist bevorzugt eine Kalibrierung der durch Videokameras erzeugten Videosignale vorgesehen. Durch eine Kalibrierung der Videosignale ist weiters die Zuordnung eines Punktes auf einem Videobild zu einem Punkt auf einem Plan des überwachten Bereichs möglich.

[0025] Im Folgenden wird die bevorzugte Kalibrierung eines Videosignals beispielhaft beschrieben.

[0026] Linien, die in der realen Welt parallel sind, schneiden sich in fotografischen Abbildungen in Fluchtpunkten. Vorzugsweise sind Videokameras zur Anwendung gemäß der gegenständlichen Erfindung derart ausgerichtet, dass Menschen in der Abbildung im Wesentlichen aufrecht erscheinen, daher der „Up-Vector“ der Videokamera 1, 2, 16 liegt mit dem Blickrichtungsvektor derselben im Wesentlichen in einer Vertikalebene. Somit schneiden sich alle vertikalen Linien (Säulen, Türstöcke, ...) der Abbildung in einem Fluchtpunkt unterhalb der Abbildung, da Videokameras 1, 2, 16, durch die Anordnung an einer Gebäudedecke oder einem Masten meist eine leicht abwärts geneigte Blickrichtung aufweisen. Dieser Fluchtpunkt wird vanishing_south genannt.

[0027] Des Weiteren scheinen in der vom Menschen gestalteten Umgebung oftmals, vor allem innerhalb von Gebäuden, rechteckige (orthogonale) Strukturen auf. Räume sind nahezu immer rechtwinkelig angelegt. Sofern Böden oder Wände eine strukturierte Oberfläche aufweisen, ist diese ebenfalls oftmals parallel bzw. rechtwinkelig zu Wänden eines Raumes ausgerichtet. Auch diese in der Realität parallelen Linien schneiden sich in der Abbildung in Fluchtpunkten, welche vanishing_east und vanishing_west genannt werden. Bei Kenntnis dieser drei Fluchtpunkte, etwa durch Berechnung, ist für eine Kalibrierung des Videosignals weiters die Kenntnis von drei tatsächlichen Längen in Richtung der Fluchtpunkte in dem durch die Videokamera 1, 2, 16 aufgenommenen Aufnahmebereich notwendig. Die weitere Kalibrierung des Videosignals erfolgt vorzugsweise unter Anwendung des sog. Doppelverhältnisses.

[0028] Wie bereits dargelegt kann der Einsatz einer jeden Art einer Videokamera 1, 2, 16 vorgesehen sein, wobei besonders bevorzugt der Einsatz feststehender Videokameras 1, 2, 16 bzw. sog. PTZ-Kameras 7 vorgesehen ist, wobei PTZ für Pan-Tilt-Zoom steht. Derartige PTZ-Kameras 7 sind motorisch schwenkbar und weisen einen vorgebbaren Zoombereich auf. Die Einstellbarkeit der Schwenk- und Kippmechanismen sind in der Regel derart genau, dass vordefinierte Punkte mit einer derartigen PTZ-Kamera 7 sehr genau angesteuert werden können. Derartige PTZ-Kameras 7 werden derart kalibriert, dass als Erstes die Position der PTZ-Kamera 7 im dreidimensionalen Raum gegenüber einem orthogonalen Achsenkreuz ermittelt wird. Dabei ist vorgesehen, dass die PTZ-Kamera 7 nacheinander auf einen Ursprungspunkt und auf drei Endpunkte im Raum ausgerichtet wird, welche zusammen ein orthogonales Achsenkreuz mit bekannten Längen zwischen dem Ursprung und den Endpunkten bilden, wobei jeweils die drei Relativwinkel zwischen der Ausrichtung auf den Ursprungspunkt und einen der drei Endpunkte gemessen werden, und aus diesen Relativwinkeln die Position der PTZ-Kamera 7 im dreidimensionalen Raum gegenüber dem orthogonalen Achsenkreuz ermittelt wird.

[0029] Diese Ermittlung der Position der PTZ-Kamera 7 erfolgt zufolge geometrischer bzw. trigonometrischer Berechnungen. Hiezu sei lediglich erwähnt, dass sich die Position der PTZ-Kamera 7 am Schnittpunkt dreier Tori befindet, welche durch Rotation jeweils eines Kreissegments jeweils um eine der drei orthogonal aufeinander angeordneten Achsen des orthogonalen Achsenkreuzes gebildet werden. Da dies verfahrenstechnisch sehr aufwendig ist, ist bevorzugt vorgesehen, dass - anstelle dem Schneiden dreier Tori - durch das Schneiden von Parallelkreisen auf einer Hilfskugeloberfläche, deren Mittelpunkt der Ursprungspunkt ist, die Position der PTZ-Kamera 7 im dreidimensionalen Raum gegenüber dem orthogonalen Achsenkreuz ermittelt wird. Dies ist möglich, da die Rotationsachsen der drei Tori auch gleichzeitig die Achsen des orthogonalen Achsenkreuzes sind. Hiedurch kann schnell und ohne großen Aufwand die Position der PTZ-Kamera 7 gegenüber dem orthogonalen Achsenkreuz ermittelt werden.

[0030] Bei dem erfindungsgemäßen Verfahren wird in jedem Einzelbild wenigstens ein Hintergrundbereich sowie ein erster ortswechselnder Vordergrundbereich ermittelt, wobei der erste ortswechselnde Vordergrundbereich sämtliche Bildbereich umfasst, welche nicht dem Hintergrundbereich zugeordnet werden. Daher, welche sich innerhalb der durch das erste Videosignal abgedeckten Sequenz derart bewegen, dass diese aufgrund der weiters erläuterten besonders bevorzugten Ausbildung des erfindungsgemäßen Verfahrens nicht dem Hintergrundbereich zugeordnet werden können. Dabei ist vorgesehen, dass die Werte, die jeder einzelne Pixel eines Kamerabilds über einen definierbaren Zeitraum in der Vergangenheit angenommen hat analysiert werden.

[0031] Bei einem erfindungsgemäßen Verfahren, bei welchem Einzelbilder mit einer vorgebbaren Anzahl an Pixel verarbeitet werden, wobei jeder Pixel wenigstens einen Wert, insbesondere einen Graustufenwert und/oder einen Farbwert, aufweist, ist daher bevorzugt vorgesehen, dass der Wert eines ersten Pixels eines ersten Einzelbildes mit einer vorgebbaren Anzahl Vergleichswerten verglichen wird, dass dem ersten Pixel der Wert eines ersten Vergleichswertes, aus der vorgebbaren Anzahl Vergleichswerten, zugeordnet wird, wenn der Wert des ersten Pixels innerhalb eines vorgebbaren Toleranzbereichs um den ersten Vergleichswert liegt, dass eine Auftretenshäufigkeit des ersten Vergleichswertes um einen vorgebbaren ersten Betrag erhöht wird, und dass der Vergleichswert mit der höchsten Auftretenshäufigkeit dem Hinter-

grundbereich zugeordnet wird. Dadurch wird ein dynamischer Hintergrund ermittelt, welcher stabil gegenüber zeitlich langsam fortschreitenden Änderungen bzw. gegenüber statistisch häufig auftretenden Veränderungen ist. Ein derart generierter bzw. ermittelter Hintergrund ist etwa stabil gegenüber tageszeitlich veränderten Lichtbedingungen. Weiters werden etwa sich leicht verändernde Bereiche, wie etwa ein im Wind bewegter Baum als Hintergrund erkannt, und führen nicht zu einer Alarmierung der Kontrollstelle.

[0032] Für jeden Pixel des Einzelbildes wird die Möglichkeit geschaffen, eine vorgebbare Anzahl Vergleichswerte und deren Auftretenshäufigkeit zu speichern bzw. aufzubewahren. In der Folge wird das beispielhaft für die Anzahl fünf beschrieben.

[0033] Wenn ein neuer Frame bzw. Einzelbild des Videosignals zur Verarbeitung durch das erfindungsgemäße Verfahren bereitgestellt wird, wird der Wert eines jeden Pixels mit den jeweils vorzugsweise fünf Vergleichswerten verglichen. Wenn der Wert innerhalb eines vorgebbaren Toleranzbereichs um einen der vorzugsweise fünf Vergleichswerte liegt, wird dieser Wert diesem gespeicherten Vergleichswert zugeordnet. Sobald ein passender gespeicherter Vergleichswert gefunden wurde, wird das Verfahren unterbrochen, und die weiteren Vergleichswerte werden nicht mehr überprüft.

[0034] Wenn der aktuelle Wert eines Pixels einem gespeicherten Vergleichswert zugeordnet werden kann, wird die Auftretenshäufigkeit dieses Vergleichswerts um einen vorgebbaren ersten Betrag, etwa um eins, erhöht. Somit erhalten diejenigen gespeicherten Vergleichswerte nach vielen Frames die größere Auftretenshäufigkeit, die häufiger die beste Approximation des aktuellen Werts eines Pixels waren. Unter der Annahme, dass in einer Videosequenz bzw. einem Videosignal am häufigsten jene Werte eines Pixels an einer spezifischen Stelle auftauchen, die zum Hintergrundbereich des Einzelbildes (und nicht zu bewegten bzw. ortsändernden Vordergrundbereichen und/oder Objekten) gehören, besitzt nun jener gespeicherte Wert für diesen Pixel die größte Auftretenshäufigkeit, der der Hintergrundfarbe bzw. -graustufe am ehesten entspricht.

[0035] Bevorzugt ist vorgesehen, dass, wenn der Wert des ersten Pixels außerhalb vorgebbarer Toleranzbereiche um die Vergleichswerte liegt, dem Vergleichswert mit der geringsten Auftretenshäufigkeit der Wert des ersten Pixels zugeordnet wird. Dadurch kann eine ständige Anpassung der Vergleichswerte an die sich ändernden Inhalte des Videosignals erreicht werden.

[0036] Gemäß einer besonders bevorzugten Ausbildung der Erfindung ist vorgesehen, dass ein Maximalwert als obere Grenze der Auftretenshäufigkeit vorgegeben wird. Sollte die Auftretenshäufigkeit eines gespeicherten Vergleichswerts diesen Maximalwert übersteigen, wird die Auftretenshäufigkeit auf den Maximalwert gesetzt. Diese Maßnahme definiert damit auch eine Lernrate; da ein neu in dem Videosignal aufscheinendes Objekt eben so viele Frames bzw. Einzelbilder benötigt, um die Auftretenshäufigkeit des ihm zugehörigen gespeicherten Vergleichswerts auf den Maximalwert zu bringen und damit diesen gespeicherten Wert als Hintergrund zu etablieren.

[0037] Weiters ist bevorzugt vorgesehen, dass die Auftretenshäufigkeit jedes Vergleichswertes nach einer vorgebbaren Anzahl nachfolgender Einzelbilder um einen vorgebbaren zweiten Betrag reduziert wird. Dadurch ist ein ständiges „Nachlernen“ durch das erfindungsgemäße Verfahren gegeben. Eine vorgebbare Anzahl der nachfolgenden Einzelbilder von drei würde z.B. bedeuten, dass bei jedem dritten ankommenden Frame die Auftretenshäufigkeit aller gespeicherter Vergleichswerte um den zweiten Betrag, etwa um eins, reduziert werden. Dadurch wird die Auftretenshäufigkeit eines Vergleichswerts, der auf dem Maximalwert war und daher als Hintergrund bzw. Hintergrundbereich deklariert wurde, stetig auf die Auftretenshäufigkeit null reduziert, wenn er nicht mehr „erneuert“ wird.

[0038] In Weiterbildung des erfindungsgemäßen Verfahrens, bei welchem ein zweites Einzelbild in einem vorgebbaren Zeitbereich nachfolgend dem ersten Einzelbild aufgenommen wird, ist bevorzugt zur Bestimmung eines ortswechselnden Vordergrundbereichs vorgesehen, dass

der Hintergrundbereich des ersten Einzelbildes von einem Hintergrundbereich des zweiten Einzelbildes abgezogen wird, und die derart gebildete Differenz wenigstens dem ersten ortswechselnden Vordergrundbereich zugeordnet wird.

[0039] Ein einem Aufnahmebereich zugefügtes Objekt, welches anfangs als Teil des Vordergrundbereiches erkannt werden wird, wird mit Fortlaufen des Videosignals nach einer gewissen Zeitdauer als Hintergrund bzw. als Teil des Hintergrundbereiches erkannt werden. Zum Erkennen, ob ein solches Objekt einem Aufnahmebereich zugefügt oder entfernt worden ist, ist bevorzugt vorgesehen, dass das erfindungsgemäße Verfahren mehrmals, etwa zweimal, parallel durchgeführt wird, wobei jeweils unterschiedliche erste Beträge vorgesehen sind. Daher bei den beiden parallel ablaufenden Verfahren werden Veränderungen unterschiedlich schnell als Hintergrundbereich erkannt bzw. definiert. Daher ist ein schnell „lernendes“ und ein langsam „lernendes“ Verfahren vorgesehen. Dadurch ist auch eine Zuordnung möglich, ob ein Objekt dem Aufnahmebereich zugefügt oder entfernt worden ist. Sollte ein Objekt entfernt worden sein, ist es bei dem langsam „lernenden“ Verfahren noch dem Vordergrundbereich zugeordnet, während dieser Bereich im schnell „lernenden“ Verfahren bereits dem Hintergrund zugeordnet ist. Sollte ein Objekt hinzugefügt worden sein, ist dieses bei dem langsam „lernenden“ Verfahren noch nicht dem Vordergrundbereich zugeordnet, sondern der betreffende Bereich noch als Hintergrundbereich definiert, während das Objekt im schnell „lernenden“ Verfahren bereits dem Vordergrundbereich zugeordnet ist.

[0040] Besonders bevorzugt ist hiebei vorgesehen, dass sämtliche Pixel bewertet werden, welche innerhalb einer geometrischen Struktur, etwa einem Rechteck, um einen erkannten Vordergrundbereich angeordnet sind. Dadurch kann ein direkter Vergleich mit dem umgebenden Hintergrundbereich durchgeführt werden, wodurch die Fehlerwahrscheinlichkeit des Verfahrens gesenkt werden kann.

[0041] Fig. 2 zeigt ein Ablaufdiagramm einer besonders bevorzugten Ausbildung des vorstehend beschriebenen erfindungsgemäßen Verfahrens.

[0042] In weiterer Folge sollen innerhalb eines erkannten Vordergrundbereichs einzelne, insbesondere körperlich, zusammengehörende Objekte, etwa Menschen, Fahrzeuge oder bewegte Gegenstände, wie etwa Koffer oder Gepäckwagen, erkannt werden, und die erkannten Objekte durch eine Videosequenz bzw. wie an weiterer Stelle beschrieben, auch über mehrere Videosignale bzw. Videosequenzen unterschiedlicher Videokameras 1, 2, 16 verfolgt werden. Das Erkennen eines Objekts und/oder das Verfolgen bereits bestimmter bzw. erkannter Objekte in einem Videosignal wird als Tracking bezeichnet. Hiefür sind Trackingverfahren bzw. sog. Tracker vorgesehen. Es ist bevorzugt vorgesehen, dass innerhalb des Vordergrundbereichs ortswechselnde Objekte mittels eines Mean-Shift-Trackers erkannt werden, und/oder deren Bewegungen mittels eines Mean-Shift-Trackers oder eines Optical-Flow-Trackers, verfolgt werden, wobei jeweils auch andere Tracker bzw. Trackingverfahren vorgesehen sein bzw. angewendet werden können.

[0043] Um die Vorteile verschiedener Tracker bzw. Trackingverfahren zu vereinen, ist ein sog. Tracker-Manager vorgesehen, welcher die Ergebnisse parallel laufender Trackingverfahren zusammenfasst und ein Trackingergebnis ermittelt. Dabei ist insbesondere vorgesehen, dass aus einer Gruppe vorgegebbarer Tracker eine Anzahl Tracker zur Verfolgung der Bewegung des ersten ortswechselnden Objekts ausgewählt wird, dass nachfolgend die Bewegung des wenigstens einen ersten ortswechselnden Objekts mit jedem der ausgewählten Tracker verfolgt und ein Trackingergebnis ermittelt wird, dass zu jedem Tracker eine Fehlerwahrscheinlichkeit ermittelt wird, und dass nachfolgend die Trackingergebnisse der Tracker mit geringen Fehlerwahrscheinlichkeiten zu einem Trackingergebnis zusammengefasst werden.

[0044] Fig. 3 zeigt ein Ablaufdiagramm eines derartigen Tracker-Managers. Als erster Schritt werden die zu verwendeten Tracker ausgewählt, wobei die bisherige, daher bei einem bereits laufenden Verfahren, die bis zu diesem Zeitpunkt erreichte, Leistungsfähigkeit und/oder Fehlerwahrscheinlichkeit zur Auswahl herangezogen werden. Zusätzlich werden Tracker vorgezogen, die bereits eine definierte Anzahl von vorherigen Frames abgearbeitet haben, um die

Zuverlässigkeit des Trackers zu erhöhen.

[0045] Im nächsten Schritt werden die Abarbeitungsschritte der einzelnen Tracker gestartet, wobei diese zeitlich parallel oder hintereinander ausgeführt werden können. Das Ergebnis des Arbeitsschritts sind n Sets von Objekten, die nun zu einem stimmigen einzelnen Endergebnis zusammengeführt werden müssen; dies erfolgt in den nachfolgenden Schritten.

[0046] Je nach Tracker können unterschiedliche Eingaben die Effizienz und Zuverlässigkeit der Trackingverfahren beeinflussen, was zu fehlerhaften Ergebnissen führen kann; ein Beispiel ist eine zu hohe Personendichte bei einem Einsatz des sog. Mean-Shift-Trackers. Um dem entgegen zu wirken werden vorzugsweise für jeden Tracker Parameter definiert durch die es möglich ist eine Fehlerwahrscheinlichkeit zu ermitteln, die sich auf den gesamten Aufnahmebereich oder auch auf lediglich einen Bildabschnitt beziehen. Bei einer zu hohen, vorermittelten Fehlerwahrscheinlichkeit werden die so gekennzeichneten Bereiche für diesen Tracker nicht für das Endergebnis verwendet. Sollte die Fehlerwahrscheinlichkeit über einen bestimmten Zeitraum gleichermaßen hoch sein, wird die gesamte Abarbeitung dieses Trackers für einen vorgebbaren Zeitabschnitt eingestellt.

[0047] Nach diesem Schritt werden die einzelnen Objekte der unterschiedlichen Tracker untersucht um eine Wahrscheinlichkeit zu ermitteln, dass diese dasselbe physikalische Objekt darstellen. Wird eine vorgebbare Wahrscheinlichkeit ermittelt, dass es sich um dasselbe physikalische Objekt handelt, werden diese Objekte zu einem einzigen Objekt zusammengefasst. Die so zusammengefassten Objekte werden als Eingabemenge für einen Arbeitsfluss verwendet, der für die temporäre Zuordnung der Objekte zuständig ist, d.h. welche gefundenen Objekte im vorherigen Frame sind vergleichbar mit den gefundenen Objekten im derzeitigen Frame.

[0048] Wie bereits dargelegt kann der Einsatz eines jeden Trackers bzw. eines jeden Trackingverfahrens vorgesehen sein. Nachfolgend werden zwei lediglich besonders bevorzugte Tracker, der Mean-Shift-Tracker und der sog. Optical-Flow-Tracker in jeweils besonders bevorzugten Ausbildungen näher beschrieben.

[0049] Die Grundzüge des Mean-Shift-Trackers sind dem Fachmann bekannt. Der Mean-Shift-Tracker wird gemäß gegenständlicher bevorzugten Ausbildung auf sog. Differenzbilder angewandt. Dabei handelt es sich um Einzelbilder, welche keine Hintergrundbereiche aufweisen, bzw. welche lediglich die erkannten Vordergrundbereiche aufweisen bzw. sofern vorgesehen auch tatsächlich darstellen. Derartige Differenzbilder weisen einen hohen Kontrast auf, was für die Anwendung des Mean-Shift-Trackers vorteilhaft ist, und zu geringen Fehlern des Mean-Shift-Trackers führt.

[0050] Auf die in einem Differenzbild aufscheinenden Vordergrundbereiche wird der Meanshift Clustering Algorithmus angewendet um die Zentren (bzw. Schwerpunkte) der einzelnen Vordergrundbereiche zu lokalisieren und derart körperlich zusammengehörige Objekte zu bestimmen. Dabei werden für Kernel bzw. Blöcke mit einer vorgebbaren Anzahl Pixel, welche jeweils wenigstens bereichsweise einen Vordergrundbereich umfassen, jeweils Vektoren bzw. Gradienten ermittelt, welche in Richtung der größten Helligkeit weisen. Derart können bei einem Differenzbild einzelne Objekte erkannt und verfolgt werden. Da ein herkömmlicher Mean-Shift-Tracker sehr arbeitsaufwendig und somit langsam ist, ist bevorzugt vorgesehen, diesen unter Anwendung eines sog. Integralbilds anzuwenden. Dadurch kann die ansonsten zeitaufwendige Berechnung der einzelnen Gradienten wesentlich beschleunigt werden, wodurch die Möglichkeit geschaffen wird, das erfindungsgemäße Verfahren mit vertretbarem Aufwand in Echtzeit durchführen zu können.

[0051] Bekannte Mean-Shift-Tracker weisen den Nachteil auf, dass alle Pixel einen gleichen Stellenwert aufweisen. Für die Ermittlung der Gradienten ist es jedoch vorteilhaft, wenn Pixel in der Mitte eines Kernels stärker gewichtet werden, als solche an einem Randbereich. Es ist daher bekannt, die Pixel innerhalb eines Kernels zu Gewichten. Die Gewichtung mit herkömmlichen Gewichtungsfunktionen ist jedoch mit dem Nachteil verbunden, dass dann die - für die Verarbeitung in Echtzeit vorteilhafte - Anwendung des Integralbildes unmöglich ist. In Weiterbil-

derung der Erfindung ist daher bevorzugt vorgesehen, dass jeder Kernel bzw. Block mit wenigstens einer vorgebbaren Exponentialfunktion gewichtet wird. Exponentialfunktionen sind Funktionen, welche sich selbst in jeweils unterschiedlichen Wertebereichen ähnlich sind, sodass weiterhin die Ermittlung der Gradienten zufolge eines Integralbildes möglich ist. Durch die Gewichtung mit einer Exponentialfunktion können die Vorteile der Anwendung eines Integralbildes mit den Vorteilen der Gewichtung der einzelnen Pixel innerhalb eines Kernels kombiniert werden.

[0052] Wie vorstehend dargelegt, werden durch den Mean-Shift-Tracker jeweils Vektoren bzw. Gradienten ermittelt, welche in Richtung der größten Helligkeit weisen. Die Vektoren bzw. Gradienten weisen dabei einen Endpunkt auf, wobei bei einem idealen Verhalten des Mean-Shift-Trackers die Endpunkte sämtlicher einem Objekt zuordenbaren Vektoren in einem Punkt, dem Zentrum bzw. Schwerpunkt, konvergieren sollten, wodurch dann auch die Zuordnung einzelner Teilbereiche eines Vordergrundbereichs zu einem bestimmten Objekt möglich ist. Unter realen Bedingungen, bei tatsächlich implementierten Mean-Shift-Tracker, kommt es jedoch sehr selten zu einer derartigen Bündelung der Endpunkte der Vektoren in einem Punkt. Vielmehr sind diese Endpunkte der Vektoren in einem Bereich um das Zentrum bzw. den Schwerpunkt angeordnet. Dadurch kann es aber zu fehlerhaften Zuordnungen einzelner Blöcke zu einem Objekt kommen. Es hat sich gezeigt, dass die Genauigkeit des erfindungsgemäßen Verfahrens deutlich gesteigert werden kann, wenn einem Endpunkt des zu jedem Block ermittelten Vektors eine vorgebbare zweite Umgebung zugeordnet wird, und diese zweiten Umgebungen ebenfalls gemäß dem Mean-Shift-Verfahren bzw. -Tracker verarbeitet werden, wobei die vorstehend beschriebenen bevorzugten Ausführungen des Mean-Shift-Verfahrens vorgesehen sein können. Der Mean-Shift-Tracker wird dabei auf ein erstes Zwischenbild angewendet, welches lediglich aus den - gegebenenfalls bereichsweise überlappenden - zweiten Umgebungen aufgebaut ist. Bei der bevorzugten Implementierung des erfindungsgemäßen Verfahrens in Form eines Computerprogrammprodukts kann bevorzugt vorgesehen sein, dass dieses erste Zwischenbild nicht tatsächlich visuell dargestellt bzw. ausgegeben wird, sondern lediglich mathematisch beschrieben wird, und/oder imaginär in Form vorgebbarer Speicherbelegungen in einem Computer existiert. Dadurch kann erreicht werden, dass die derart ermittelten neuen Endpunkte der Vektoren wesentlich besser konvergieren, und eine deutlich bessere Zuordnung der einzelnen ursprünglichen Blöcke zu einem ortswechselnden Objekt erreicht werden. Dadurch wird die Leistungsfähigkeit des Mean-Shift-Trackers deutlich verbessert.

[0053] Der Optical-Flow-Tracker ist ein Verfahren, welches nicht wie der Mean-Shift-Tracker versucht einzelne Objekte isoliert zu verfolgen sondern einen sog. 'optischen Fluss' analysiert. Somit ist dieses Verfahren auch vor allem in Situationen, bei welchen aufgrund der Informationsdichte im Bild das Erkennen einzelner Objekte/Personen nicht mehr möglich ist, immer noch möglich Bewegungsrichtungen zu erkennen. Weiters ist es in manchen Situationen nicht von Interesse jedes Objekt einzeln verfolgen zu können. Optical-Flow-Tracker sind besonders vorteilhaft bei der Bewertung und Verfolgung von Menschenströmen.

[0054] Wie eingangs dargelegt weisen bekannte Tracker den Nachteil auf, dass mit diesen ein Auffinden von Bewegungsmustern in einem ersten digitalen Videosignal in Echtzeit, daher im Wesentlichen nur sehr geringfügig gegenüber dem Auftreten eines Ereignisses zeitverzögert, nicht möglich ist. Die Erfindung betrifft daher weiters ein Verfahren zum Auffinden von Bewegungsmustern in einem ersten digitalen Videosignal, wobei das erste digitale Videosignal, umfassend wenigstens mittelbar aufeinander folgende digitale Einzelbilder, in Echtzeit analysiert wird, wobei in einem ersten Einzelbild wenigstens ein erster ortswechselnder Vordergrundbereich mit einer vorgebbaren Anzahl erster Markierungen vorgebbare belegt wird, wobei nachfolgend die Relativbewegung der ersten Markierungen zwischen dem ersten Einzelbild und einem diesem nachfolgenden zweiten Einzelbild ermittelt werden, wobei nachfolgend jeder ersten Markierung jeweils eine vorgebbare erste Umgebung zugeordnet wird, wobei das erste und/oder wenigstens ein zweiter ortswechselnder Vordergrundbereich mit einer vorgebbaren Anzahl zweiter Markierungen vorgebbare belegt wird, wobei die zweiten Markierungen, welche innerhalb von Schnittbereichen einer vorgebbaren Anzahl sich überschneidender erster Umge-

bungen angeordnet sind, entfernt werden, wobei nachfolgend die Relativbewegung der ersten und zweiten Markierungen zwischen dem zweiten Einzelbild und einem diesem nachfolgenden dritten Einzelbild ermittelt werden, und wobei die Relativbewegungen der ersten und/oder zweiten Markierungen als erstes Bewegungsmuster ausgegeben werden. Dadurch ist ein Auffinden von Bewegungsmustern in einem ersten digitalen Videosignal beim Auftreten größerer bewegter Objektmengen in Echtzeit möglich. Dadurch ist eine EDV-mäßige Implementierung des erfindungsgemäßen Verfahrens mit vertretbarem Hardwareaufwand möglich. Dadurch werden vorhandene Hardwareressourcen besonders effektiv genutzt.

[0055] Als Bewegungsmuster wird dabei vorzugsweise die Richtung, Geschwindigkeit und gegebenenfalls Beschleunigung eines ortswechselnden Objekts verstanden.

[0056] Bei dem erfindungsgemäßen Verfahren werden auf Vordergrundbereichen, welche in dem ersten Einzelbild als ortswechselnd erkannt sind, Markierungen aufgebracht. Bevorzugt ist vorgesehen, dass Vordergrundbereiche zufolge des vorstehend beschriebenen Verfahrens erkannt werden. Es kann jedoch auch vorgesehen sein, sofern die erste Videokamera auf einen Ursprungsbereich gerichtet ist, welcher mit hoher Wahrscheinlichkeit ein Ursprung sämtlicher in einem Videosignal aufscheinender bewegter Objekte bzw. Vordergrundbereiche ist, etwa eine Personenzutrittsschleuse, dass auf eine separat implementierte Erkennung ortswechselnder Vordergrundbereiche verzichtet wird, und sämtliche Bereiche in diesem Ursprungsbereich mit Markierungen belegt werden.

[0057] Auf dem ersten Einzelbild werden erste Markierungen in vorgebbarer Verteilung angeordnet, wobei lediglich die ersten Markierungen weiterverarbeitet werden, welche auf einem ortswechselnden Vordergrundbereich angeordnet werden. Erfindungsgemäß kann vorgesehen sein, dass die Markierungen in jeder vorgebbaren Form verteilt auf dem wenigstens einen ersten Vordergrundbereich angeordnet werden. Bevorzugt ist vorgesehen, dass die ersten und/oder zweiten Markierungen in Form eines, vorzugsweise gleichmäßigen oder vorgebbbar gewichtet verteilten, Rasters angeordnet werden, wobei besonders bevorzugt vorgesehen ist, Bildausschnitte, in welchen mit dem Auftreten bewegter bzw. ortsändernder Objekte bzw. Vordergrundbereiche gerechnet wird, mit einem dichteren Raster an Markierungen zu belegen.

[0058] Erfindungsgemäß ist vorgesehen, dass die Relativbewegung der ersten Markierungen zwischen dem ersten Einzelbild und einem diesem nachfolgenden zweiten Einzelbild ermittelt werden, wobei besonders bevorzugt vorgesehen ist, dass die Relativbewegungen der ersten und/oder zweiten Markierungen mittels eines Verfahrens zur Schätzung, insbesondere zur Bestimmung, eines optischen Flusses in dem ersten Videosignal, vorzugsweise ein Optical-Flow-Tracker, insbesondere zufolge des Horn-Schunk-Verfahrens oder des Kanade-Lucas-Tomasi-Verfahrens, ermittelt wird.

[0059] Den derart weiterbewegten ersten Markierungen wird weiters eine erste Umgebung zugeordnet, welche als, vorzugsweise gewichtetes, Rechteck, insbesondere als Quadrat, oder als, vorzugsweise gewichtete, Kegelschnittfläche, insbesondere als Kreis oder Ellipse, ausgebildet ist. In einem nächsten Verfahrensschritt wird das derart bearbeitete zweite Einzelbild wiederum mit zweiten Markierungen in der beschriebenen Art belegt, wobei jedoch neue zweite Markierungen nur auf Vordergrundbereichen angeordnet werden, wenn an der betreffenden Stelle an welcher eine zweite Markierung vorgesehen ist, kein Schnittbereich einer vorgebbaren Anzahl sich überlappender bzw. überschneidender erster Umgebungen ist. Daher wenn eine zweite Markierung an einem ortswechselnden Vordergrundbereich angeordnet werden soll, diese Stelle aber bereits durch den Schnittbereich von, beispielsweise drei, ersten Umgebungen abgedeckt ist, so wird diese zweite Markierung nicht gesetzt. Anschließend wird das erfindungsgemäße Verfahren von Beginn weg sinngemäß für die weiters nachfolgenden Einzelbilder in Form einer Schleife wiederholt.

[0060] Bevorzugt ist vorgesehen, dass zu jeder Markierung und jeder berechneten Relativbewegung eine Fehlerstatistik ermittelt wird, und Markierungen, deren Fehlerstatistik eine vorgebbare Grenze übersteigt, entfernt werden. Weiters wird vorzugsweise die Größe der Relativbewegung betrachtet, und Markierungen, deren Relativbewegung unter ein vorgebbares Maß fällt,

entfernt. Dadurch kann erreicht werden, dass sich nicht bewegende Markierungen, etwa solche welche fehlerhaft gesetzt wurden, ausscheiden und der Arbeitsaufwand bei Durchführung des Verfahrens gering gehalten werden kann. Zusätzlich wird vorzugsweise das Alter jeder Markierungen über einen vorgebbaren Zeitraum bzw. eine vorgebbare Anzahl an aufeinander folgenden Einzelbildern aufgezeichnet, und Markierungen, deren Alter eine vorgebbare Grenze, beispielsweise 50 aufeinander folgende Einzelbilder, übersteigt, entfernt. Auch dadurch kann eine ständige Regeneration der verarbeiteten Markierungen erreicht werden, und der Arbeitsaufwand bei Durchführung des Verfahrens gering gehalten werden.

[0061] Zur Auswertung der ermittelten Bewegungsmuster ist bevorzugt vorgesehen, dass wenigstens das zweite Einzelbild in vorgebar anordenbare Segmente vorgebbarer Form und Größe unterteilt wird, dass wenigstens die ersten Bewegungsmuster innerhalb eines jeden Segments bewertet werden, und dass für jedes Segment ein Segmentbewegungsmuster ermittelt wird. Daher wird wenigstens das zweite Einzelbild, vorzugsweise auch sämtliche weiteren Einzelbilder, vorgebar segmentiert, wobei die Segmente in Form und Größe vorgebar sind. Vorzugsweise werden die Segmente an den zu erwartenden Bildinhalt angepasst. Daher Bereiche, in welchen nicht mit Vordergrundbereichen bzw. Objekten zu rechnen ist, welche für einen Überwacher von Interesse sein dürften, werden größere Segmente angeordnet, welche in Bereichen, welche möglichst genau analysiert werden sollen, mehrere kleine Bereiche angeordnet werden. Innerhalb eines jeden derartigen Segments wird vorzugsweise ein Richtungshistogramm erstellt wird, in welchem jede ermittelte Bewegungsrichtung, vorzugsweise bewertet bzw. gewichtet, eingetragen wird. Dadurch kann verhindert werden, dass mehrere große ermittelte Bewegungsrichtungen ein Segmentbewegungsmuster ergeben würden, welches die tatsächlichen Bewegungsrichtungen nicht mehr wiedergibt. Dadurch ist eine schnelle und doch differenzierte Bewertung der Bewegungsabläufe innerhalb des ersten Videosignals möglich.

[0062] Aus dem ersten Bewegungsmuster und/oder aus den Segmentbewegungsmustern wird vorzugsweise wenigstens ein erstes Hauptbewegungsmuster ermittelt. Hierzu ist kann ebenfalls vorgesehen sein, dass ein Richtungshistogramm erstellt wird, in welchem jedes Segmentbewegungsmuster, vorzugsweise bewertet bzw. gewichtet, eingetragen wird. Dadurch kann weiters ein globales Hauptbewegungsmuster ermittelt werden. Weiters ist vorzugsweise vorgesehen, gezielt nach bestimmten Bewegungsrichtungen bzw. Objekten mit bestimmten Bewegungsrichtungen innerhalb des Videosignals zu suchen, etwa nach Objekten, welche sich entgegen einem Hauptbewegungsmuster bewegen, daher etwa sehr schnell sind (laufen) und/oder sich etwa auf einer Rolltreppe entgegen der Fahrtrichtung fortbewegen. Derartige abweichende Bewegungsmuster lassen den Verdacht auf eine sicherheitsrelevante Situation zu.

[0063] Besonders bevorzugt ist vorgesehen, dass die beschriebenen Verfahren als Computerprogrammprodukt umgesetzt ist, welches direkt in den internen Speicher eines Computers geladen werden kann, oder auf einem Datenträger abgelegt ist, und Softwarecodeabschnitte umfasst, mit denen die Schritte eines der erfindungsgemäßen Verfahren ausgeführt werden, wenn das Computerprogrammprodukt auf einem Computer ausgeführt wird. Ein Computer ist hierbei vorzugsweise jegliche Vorrichtung, welche einen geeigneten Mikroprozessor, MikroController und/oder digitalen Signalprozessor (DSP), sowie die weiters notwendigen Speicher und/oder einen Datenbus umfasst. Dabei kann etwa auch vorgesehen sein, dass ein Mikrocontroller und/oder ein digitaler Signalprozessor innerhalb eines sog. PLD (Programmable Logic Device) oder FPGA (Field Programmable Gate Array) ausgebildet ist. Neben herkömmlichen Personal Computern bzw. Großrechnern sind dabei weiters vor allem Ausbildung eines Computers als sog. Embedded System vorgesehen.

[0064] Die Erfindung betrifft weiters eine Videokameraanordnung mit wenigstens einer ersten und einer zweiten Videokamera 1, 2, wobei die erste und die zweite Videokamera 1, 2 wenigstens bereichsweise den selben Bereich aufnehmen, wobei die erste Videokamera 1 an eine erste Datenverarbeitungseinheit 3 und die zweite Videokamera 2 an eine zweite Datenverarbeitungseinheit 4, insbesondere umfassend wenigstens einen digitalen Signalprozessor, angeschlossen sind, zum Erkennen eines Hintergrundes und/oder eines ortswechselnden Vordergrundbereichs und/oder eines Objekts. Eine derartige Videokameraanordnung ist besonders

gut für die Umsetzung eines erfindungsgemäßen Verfahrens geeignet, da jeweils einer Videokamera 1, 2 eine Datenverarbeitungseinheit 3, 4 zugeordnet ist. Die mit der Umsetzung des erfindungsgemäßen Verfahrens verbundene hohe Rechenleistung kann somit durch Bereitstellung separater Datenverarbeitungseinheit 3, 4, 8 bereitgestellt werden. Dabei kann auch die Kalibrierung der einzelnen Videokameras 1, 2, 7, 16 durch die Datenverarbeitungseinheiten 1, 2, 8 erfolgen.

[0065] In Weiterführung der Erfindung ist bevorzugt vorgesehen, dass die erste und die zweite Datenverarbeitungseinheit 3, 4 an eine erste Zentraldatenverarbeitungseinheit 5 angeschlossen sind, zur Bildung eines Modells 18 der von der ersten und zweiten Videokamera 1, 2 aufgenommenen Bereiche, wobei weiters bevorzugt vorgesehen sein kann, dass die erste Zentraldatenverarbeitungseinheit 5 an einen Datenspeicher 6 angeschlossen ist. Dadurch können die Daten, welche durch jede der Videokameras 1, 2, 7, 16 ermittelt werden in einem einfachen und übersichtlichen Modell 18 zusammengestellt werden.

[0066] Fig. 1 zeigt ein Blockschaltbild einer besonders bevorzugten entsprechenden Anordnung, wobei noch weitere Baugruppen vorgesehen und dargestellt sind. Eine erste und zweite Videokamera 1, 2 sind jeweils an eine erste bzw. zweite Datenverarbeitungseinheit 3, 4 angeschlossen. Weiters ist eine PTZ-Kamera 7 vorgesehen, welche an eine dritte Datenverarbeitungsanlage 8 angeschlossen ist, welche neben der Umsetzung des erfindungsgemäßen Verfahrens weiters auch für die Kalibrierung und Steuerung der PTZ-Kamera 7 vorgesehen sein kann. Zusätzlich zu den dargestellten drei Videokameras 1, 2, 7 kann jede vorgebbare Anzahl weiterer Videokameras 16 vorgesehen sein, wobei durchaus die Anordnung vieler tausend Videokameras 1, 2, 7, 16 vorgesehen sein kann. Weiters kann vorgesehen sein, neben den beschriebenen Videokameras 1, 2, 7, 16 auch Videokameras 1, 2, 7, 16 vorzusehen, welche bereits eine integrierte Datenverarbeitungseinheit 3, 4, 8 umfassen. Weiters können sog. mobile Überwachungseinheiten vorgesehen sein, welche jeweils wenigstens eine Videokamera 1, 2, 7, 16 sowie eine Datenverarbeitungseinheit 3, 4, 8 und einen Datenspeicher 6 umfassen, und nur ausgelesen werden, wenn dies durch ein Ereignis, etwa einen Terroranschlag bedingt ist. Zusätzlich kann jede Kombination der vorstehend beschriebenen Anordnungen unterschiedlicher Videokameras 1, 2, 7, 16 vorgesehen sein, etwa die Kombination aus statischen Videokameras 1, 2 und PTZ-Kameras 7.

[0067] Die Datenverarbeitungseinheiten 3, 4, 8 der einzelnen Videokameras 1, 2, 7, 16 sind an eine erste Zentraldatenverarbeitungseinheit 5 angeschlossen.

[0068] In diesem Zusammenhang kann vorgesehen sein, dass wenigstens zwei digitale Videosignale, welche jeweils einem wenigstens bereichsweise überlappenden Aufnahmebereich zugeordnet sind, zu einem Modell 18 des gesamt durch die Videosignale beschriebenen Aufnahmebereichs 15 zusammengeführt werden, und dass Modell 18 ausgegeben, insbesondere angezeigt, wird, wobei besonders bevorzugt vorgesehen sein kann, dass das Modell 18 als dreidimensionales Modell 19 erstellt wird. Zum Zusammenführen der, durch die einzelnen Videokameras 1, 2, 7, 16 ermittelten Videosignale, bzw. insbesondere der Informationen, ob Bereiche eines Bildes Hintergrundbereich oder Vordergrundbereich sind, sowie die Verfolgung detektierter Objekte quer durch eine von mehreren Videokameras 1, 2, 7, 16 abgedeckten Aufnahmebereichs ist eine Kalibrierung der einzelnen Videokameras 1, 2, 7, 16 vorteilhaft. Weiters ist ein globales Koordinatensystem notwendig, um die einzelnen Videosignale bzw. die aus diesen extrahierten bzw. gewonnenen Daten in einem einzigen Modell 18 zu verbinden. Dadurch kann einem Benutzer einer derartigen Videokameraanordnung die Möglichkeit gegeben werden sich schnell und intuitiv in einem videoüberwachten Bereich zu recht zu finden, wobei der Benutzer immer die Daten erhält, welche mit hoher Wahrscheinlichkeit die für den Benutzer interessantesten Daten sind.

[0069] In der ersten Zentraldatenverarbeitungseinheit 5 wird aus den von den Daten der einzelnen Datenverarbeitungseinheiten 3, 4, 8 ein Modell 18 des gesamten, durch die Videokameras 1, 2, 7, 16 überwachten Bereichs gebildet. Fig. 5 zeigt eine Anordnung mit drei Videokameras 1, 2, 16 in einer schematischen Grundrissdarstellung, wobei neben den Videokameras 1, 2, 16

auch deren Aufnahmebereiche 15 sowie markante Gebäudekanten 20 dargestellt sind. Fig. 6 zeigt eine Anordnung mit zehn Videokameras 1,2, 16, wobei durch die Doppelpfeile 17 jeweils überlappende Aufnahmebereiche verdeutlicht werden. Aus Gründen der Übersichtlichkeit wurde auf eine vollständige Bezeichnung sämtlicher Doppelpfeile verzichtet. Den neben der ersten und zweiten Videokamera 1,2 weiters in den Fig. 5 und 6 dargestellten Videokameras 16 wurde dabei einheitlich das Bezugszeichen 16 zugewiesen. Fig. 7 zeigt ein Modell 18 der Anordnung gemäß Fig. 6, wobei an die Stelle der einzelnen Videokameras 1, 2, 16 gemäß der Grundrissansicht aus Fig. 6 jeweils die durch die betreffenden Videokameras 1,2, 16 aufgenommenen Bilder gesetzt werden.

[0070] Das durch die Zentraldatenverarbeitungseinheit 5 erstellte Modell 18 wird in einem Datenspeicher 6 abgelegt, wobei aus Sicherheitsgründen die codierte Speicherung der entsprechenden Daten vorgesehen sein kann. In Fig. 4 ist ein Ablauf eines bevorzugten Verfahrens zur codierten Speicherung der einzelnen Videoframes dargestellt.

[0071] Die Zentraldatenverarbeitungseinheit 5 und/oder der Datenspeicher 6 sind mit einem Anwendungsserver 9 verbunden, welcher die ermittelten Daten für die Bearbeitung durch einen Benutzer aufbereitet, und mit einer Benutzerschnittstelle 10 bzw. einer mobilen Schnittstelle 11 in Verbindung steht. Durch die mobile Schnittstelle 11 ist ein Zugriff auf das System etwa durch einen sich bewegendenden Wächter möglich. Dadurch ist die videotechnische Verfolgung eines Zielobjekts unmittelbar durch eine Zugriffsmannschaft, etwa ein Polizeisondereinsatzkommando, möglich.

[0072] Die Benutzerschnittstelle 10 liefert die Daten an eine Einheit zur sog. Ereignisdarstellung 12. Eine solche Darstellung ist beispielhaft in Fig. 8 dargestellt. Auf einem symbolisch dargestellten Bildschirm werden etwa die Videokameras 1, 2, 16 gemäß der Darstellung in Fig. 7 angezeigt, wobei durch die Ereignisdarstellung 12 Videosignale, welche durch das erfindungsgemäße Verfahren erkannte Ereignisse umfassen, zentral vergrößert dargestellt werden. Darüber hinaus ermöglicht dies die gezielte und intuitive Verfolgung eines Zielobjekts durch einen Bereich, welcher durch mehrere Videokameras 1, 2, 16 überwacht wird. Die in einer Bildschirmansicht gemäß Fig. 8 dargestellten Bilder einzelner Videokameras sind an der Stelle eines Grundrisses des überwachten Bereichs dargestellt, an welchem die Videokamera 1,2, 16 angeordnet ist, oder den diese aufnimmt.

[0073] Dabei ist bei einer Ausführung der Erfindung, bei welcher ein Grundrissmodell und ein dreidimensionales Modell 19 erstellt wird, vorgesehen, dass sowohl das Grundrissmodell als auch das dreidimensionale Modell 19 im Wesentlichen gleichzeitig angezeigt werden, und dass durch Auswahl eines der wenigstens zwei digitalen Videosignale im Grundrissmodell oder im dreidimensionalen Modell 19 das Videosignal im Grundrissmodell und im dreidimensionalen Modell 19 ausgewählt wird. Dabei kann etwa vorgesehen sein, dass durch Auswahl einer Videokamera 1, 2, 16 im dreidimensionalen Modell 19, etwa mittels einer Maus oder durch Berühren des entsprechenden Bereichs eines Touchscreens, auch das Bild derselben Videokamera 1, 2, 16 im Grundrissmodell ausgewählt und auf der zentralen Ansicht vergrößert dargestellt wird. Dadurch ist ein besonders schnelles und intuitives Arbeiten möglich. Dadurch können mit einer geringen Anzahl an Benutzern eine große Anzahl Videokameras 1, 2, 16 überwacht werden. Dadurch ist weiters nur eine geringe Anlernzeit für die Benutzer notwendig.

[0074] An die Benutzerschnittstelle 10 ist weiters eine Einheit 13 zur Darstellung einer dreidimensionalen Ansicht bzw. eines dreidimensionalen Modells 19 des überwachten Bereichs angeschlossen. Fig. 9 zeigt beispielhaft eine derartige Darstellung. Hierdurch wird einem Benutzer die Möglichkeit gegeben, sich schnell und intuitiv in einem überwachten Raum zu orientieren. Dies ist vor allem vorteilhaft bzw. wichtig bei Videokameraanordnungen bzw. Sicherheitsanlagen, bei welchen eine sehr große Anzahl an Videokameras 1, 2, 7, 16 durch eine begrenzte Anzahl an Personen überwacht werden muss, sowie beim Einsatz von Personal, welches mit den räumlichen Gegebenheiten nur unzureichend vertraut ist, wie dies etwa beim Einsatz von Polizeieinsatzkräften im Zuge der Verbrechensbekämpfung gegeben ist.

[0075] Zusätzlich besteht die Möglichkeit, dass an die Benutzerschnittstelle 10 eine Einheit 14

zur statistischen und/oder forensischen Analyse der ermittelten Daten angeschlossen ist. Hierdurch besteht eine weitere Möglichkeit, auf außergewöhnliche bzw. auffällige Ereignisse bzw. Verhaltensweisen aufmerksam zu werden.

[0076] In Weiterbildung des vorstehend beschriebenen erfindungsgemäßen Verfahrens kann in diesem Zusammenhang vorgesehen sein, dass wenigstens zwei digitale Videosignale, welche jeweils einem wenigstens bereichsweise überlappenden Aufnahmebereich 15 zugeordnet sind, zu einem Modell 18 der gesamten durch die Videosignale beschriebenen Aufnahmebereiche 15 zusammengeführt werden, und dass Modell 18 ausgegeben, insbesondere angezeigt, wird, wobei besonders bevorzugt vorgesehen sein kann, dass das Modell 18 als dreidimensionales Modell 19 erstellt wird. Zum Zusammenführen der durch die einzelnen Videokameras 1, 2, 7, 16 ermittelten Videosignale, bzw. insbesondere der Informationen, ob Bereiche eines Bildes Hintergrundbereich und/oder Vordergrundbereich sind, sowie die Verfolgung detektierter Objekte quer durch die von mehreren Videokameras 1, 2, 7, 16 abgedeckten Aufnahmebereiche 15 ist eine Kalibrierung der einzelnen Videokameras 1, 2, 7, 16 vorteilhaft. Weiters ist ein globales Koordinatensystem notwendig, um die einzelnen Videosignale bzw. die aus diesen extrahierten bzw. gewonnenen Daten in einem einzigen Modell 18 zu verbinden. Dadurch kann einem Benutzer einer derartigen Videokameraanordnung die Möglichkeit gegeben werden sich schnell und intuitiv in einem videoüberwachten Bereich zu recht zu finden, wobei der Benutzer immer die Daten erhält, welche mit hoher Wahrscheinlichkeit die für den Benutzer interessantesten Daten sind.

[0077] Die Erfindung betrifft weiters eine Sicherheitsanlage, welche eine erfindungsgemäße Videokameraanordnung aufweist und/oder ein erfindungsgemäßes Verfahren anwendet, wobei insbesondere vorgesehen ist, dass eine derartige Sicherheitsanlage weiters wenigstens einen Sensor ausgewählt aus der Gruppe: Lichtschranke, Bewegungsmelder, Trittschallsensor, Vibrationssensor, Kontaktsensor, Infrarotsensor, umfasst. Bevorzugt kann dabei vorgesehen sein, dass der wenigstens eine Sensor an die Zentraldatenverarbeitungseinheit angeschlossen ist, wodurch die durch die Sensoren gewonnenen Daten ebenfalls bei der Bildung des Modells berücksichtigt werden können.

[0078] Weitere erfindungsgemäße Ausführungsformen weisen lediglich einen Teil der beschriebenen Merkmale auf, wobei jede Merkmalskombination, insbesondere auch von verschiedenen beschriebenen Ausführungsformen, vorgesehen sein kann.

Patentansprüche

1. Verfahren zur Videoanalyse, insbesondere zur Videoüberwachung, wobei ein erstes digitales Videosignal, umfassend wenigstens mittelbar aufeinander folgende digitale Einzelbilder, in Echtzeit analysiert wird, wobei in jedem Einzelbild wenigstens ein Hintergrundbereich ermittelt wird, wobei gegenüber dem Hintergrundbereich wenigstens ein erster ortswechselnder Vordergrundbereich erkannt wird, und wobei nachfolgend eine erste Meldung erzeugt, und vorzugsweise an eine Kontrollstelle gesendet wird, wobei wenigstens ein erstes ortswechselndes Objekt, welches Teil des ersten ortswechselnden Vordergrundbereichs ist, mittels eines Trackers, insbesondere einem Mean-Shift-Tracker oder einem Optical-Flow-Tracker, erkannt wird, und dessen Bewegungen mit dem Tracker verfolgt werden, **dadurch gekennzeichnet**, dass aus einer Gruppe vorgegebener Tracker eine Mehrzahl Tracker zur Verfolgung der Bewegung des ersten ortswechselnden Objekts ausgewählt wird, dass nachfolgend die Bewegung des wenigstens einen ersten ortswechselnden Objekts mit jedem der ausgewählten Tracker verfolgt und ein Trackingergebnis ermittelt wird, dass zu jedem Tracker eine Fehlerwahrscheinlichkeit ermittelt wird, und dass nachfolgend die Trackingergebnisse der Tracker mit geringen Fehlerwahrscheinlichkeiten zu einem Trackingergebnis zusammengefasst werden.
2. Verfahren nach Anspruch 1, wobei die Einzelbilder eine vorgebbare Anzahl Pixel aufweisen, und wobei jeder Pixel wenigstens einen Wert, insbesondere einen Graustufenwert und/oder einen Farbwert, aufweist, **dadurch gekennzeichnet**, dass der Wert eines ersten

Pixels eines ersten Einzelbildes mit einer vorgebbaren Anzahl Vergleichswerten verglichen wird, dass dem ersten Pixel der Wert eines ersten Vergleichswertes, aus der vorgebbaren Anzahl Vergleichswerten, zugeordnet wird, wenn der Wert des ersten Pixels innerhalb eines vorgebbaren Toleranzbereichs um den ersten Vergleichswert liegt, dass eine Auftretenshäufigkeit des ersten Vergleichswerts um einen vorgebbaren ersten Betrag erhöht wird, und dass der Vergleichswert mit der höchsten Auftretenshäufigkeit dem Hintergrundbereich zugeordnet wird.

3. Verfahren nach Anspruch 2, **dadurch gekennzeichnet**, dass, wenn der Wert des ersten Pixels außerhalb vorgegebbarer Toleranzbereiche um die Vergleichswerte liegt, dem Vergleichswert mit der geringsten Auftretenshäufigkeit der Wert des ersten Pixels zugeordnet wird.
4. Verfahren nach Anspruch 3, wobei ein zweites Einzelbild in einem vorgebbaren Zeitbereich nachfolgend dem ersten Einzelbild aufgenommen wird, **dadurch gekennzeichnet**, dass der Hintergrundbereich des ersten Einzelbildes von einem Hintergrundbereich des zweiten Einzelbildes abgezogen wird, und die derart gebildete Differenz wenigstens dem ersten ortswechselnden Vordergrundbereich zugeordnet wird.
5. Verfahren nach Anspruch 2, 3 oder 4, **dadurch gekennzeichnet**, dass die Auftretenshäufigkeit jedes Vergleichswertes nach einer vorgebbaren Anzahl nachfolgender Einzelbilder um einen vorgebbaren zweiten Betrag reduziert wird.
6. Verfahren nach einem der Ansprüche 1 bis 5, wobei ein Mean-Shift-Tracker angewendet wird, bei welchem Blöcke umfassend eine vorgebbare Anzahl Pixel gemäß dem Mean-Shift-Verfahren unter Anwendung eines Integralbilds verarbeitet werden, **dadurch gekennzeichnet**, dass die einzelnen Blöcke mit wenigstens einer vorgebbaren Exponentialfunktion gewichtet werden.
7. Verfahren nach einem der Ansprüche 1 bis 6, wobei ein Mean-Shift-Tracker angewendet wird, bei welchem Blöcke umfassend eine vorgebbare Anzahl Pixel gemäß dem Mean-Shift-Verfahren unter Anwendung eines Integralbilds verarbeitet werden, und wobei mit dem Mean-Shift-Tracker, zur Ermittlung eines Zentrums und/oder Schwerpunkts des wenigstens einen ersten ortswechselnden Objekts, für die einzelnen Blöcke jeweils ein Vektor ermittelt wird, welcher in Richtung eines Bereiches um das Zentrum bzw. den Schwerpunkt weist, **dadurch gekennzeichnet**, dass einem jeden Endpunkt des zu jedem Block ermittelten Vektors eine vorgebbare zweite Umgebung zugeordnet wird, dass die zweiten Umgebungen in einem ersten Zwischenbild angeordnet werden, und auf das erste Zwischenbild der Mean-Shift-Tracker angewendet wird.
8. Verfahren nach einem der Ansprüche 1 bis 7, **dadurch gekennzeichnet**, dass wenigstens zwei digitale Videosignale, welche jeweils einem wenigstens bereichsweise überlappenden Aufnahmebereich (15) zugeordnet sind, zu einem Modell (18) des gesamt durch die Videosignale beschriebenen Aufnahmebereichs (15) zusammengeführt werden, und das Modell (18) ausgegeben, insbesondere angezeigt, wird.
9. Verfahren nach Anspruch 8, **dadurch gekennzeichnet**, dass das Modell (18) als dreidimensionales Modell (19) und/oder als Grundrissmodell erstellt wird.
10. Verfahren nach Anspruch 9, wobei ein Grundrissmodell und ein dreidimensionales Modell (19) erstellt wird, **dadurch gekennzeichnet**, dass sowohl das Grundrissmodell als auch das dreidimensionale Modell (19) im Wesentlichen gleichzeitig angezeigt werden, und dass durch Auswahl eines der wenigstens zwei digitalen Videosignale im Grundrissmodell oder im dreidimensionalen Modell (19) das Videosignal im Grundrissmodell und im dreidimensionalen Modell (19) ausgewählt wird.
11. Datenträger **dadurch gekennzeichnet**, dass auf diesem ein Programm zur Ausführung des in den Ansprüchen 1 bis 10 beanspruchten Verfahrens gespeichert ist.

Hierzu 8 Blatt Zeichnungen

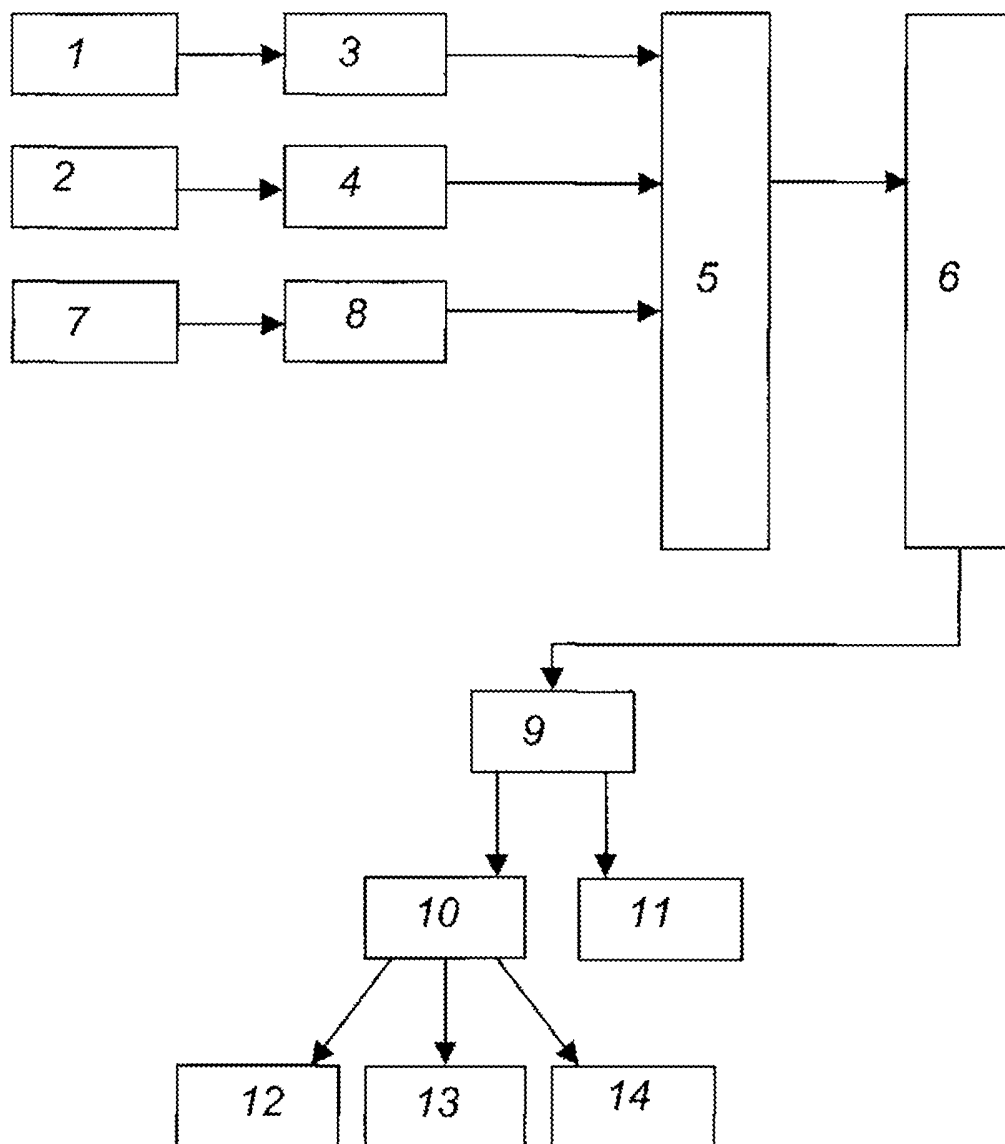


Fig. 1

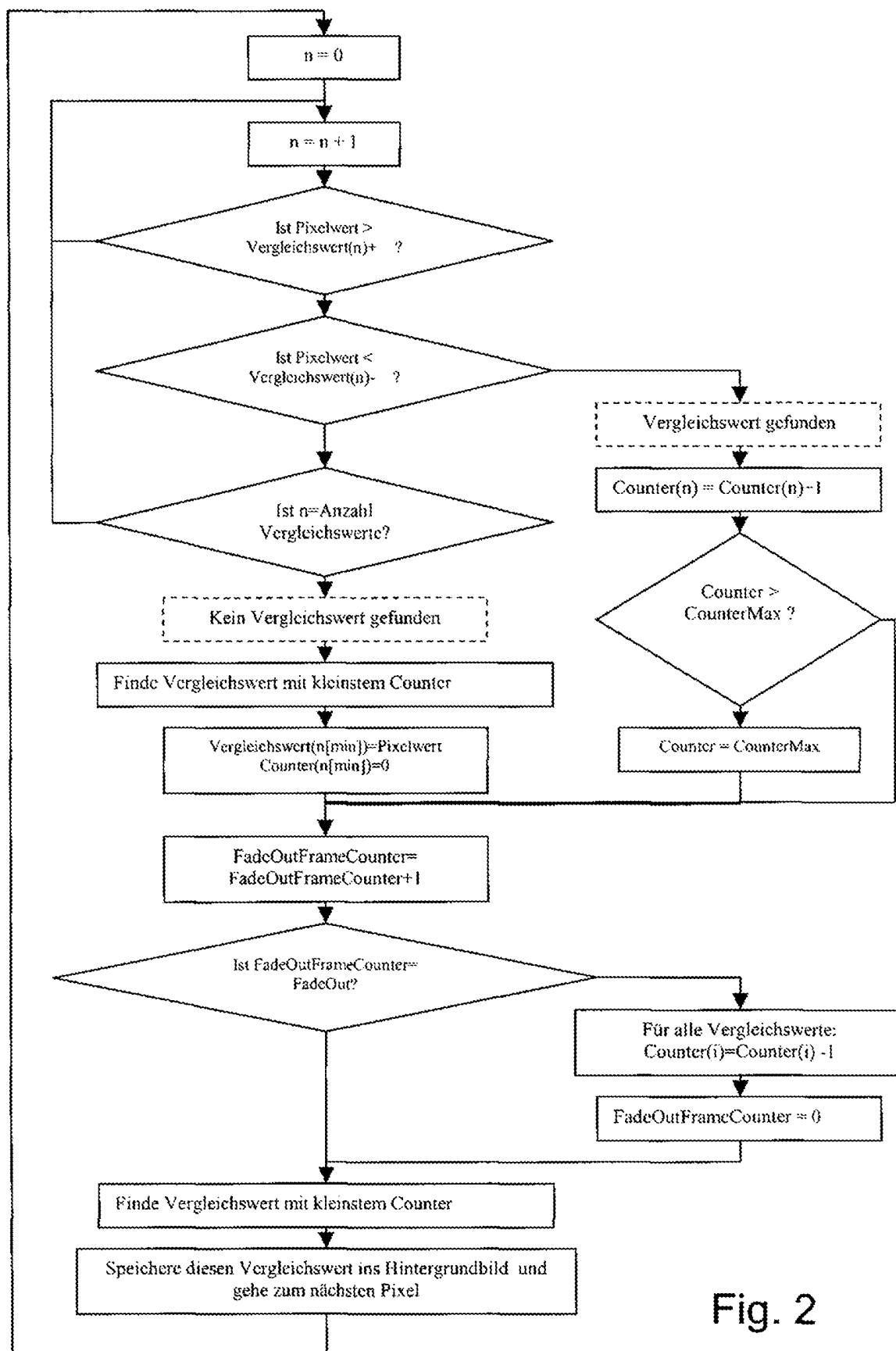


Fig. 2

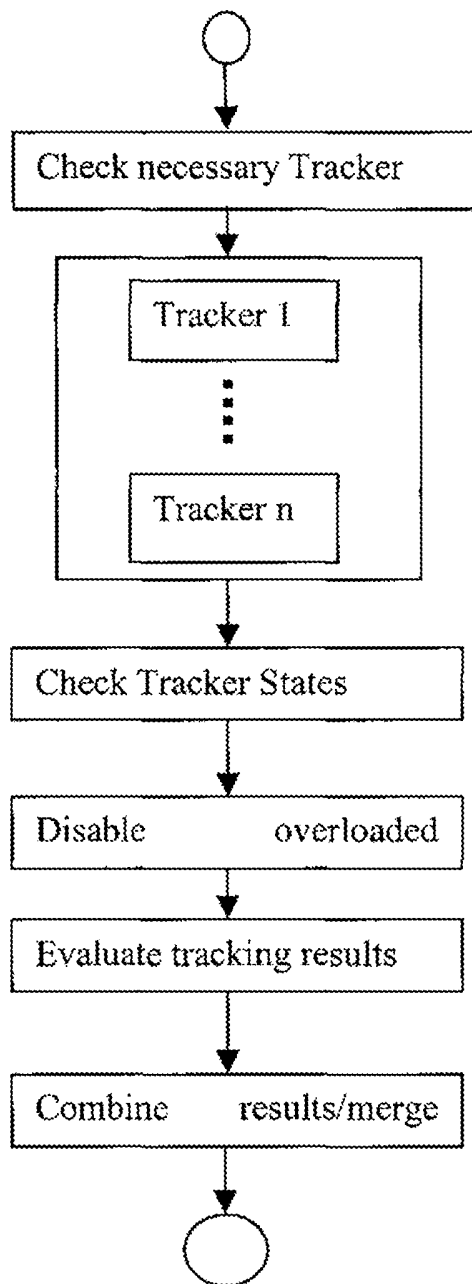


Fig. 3

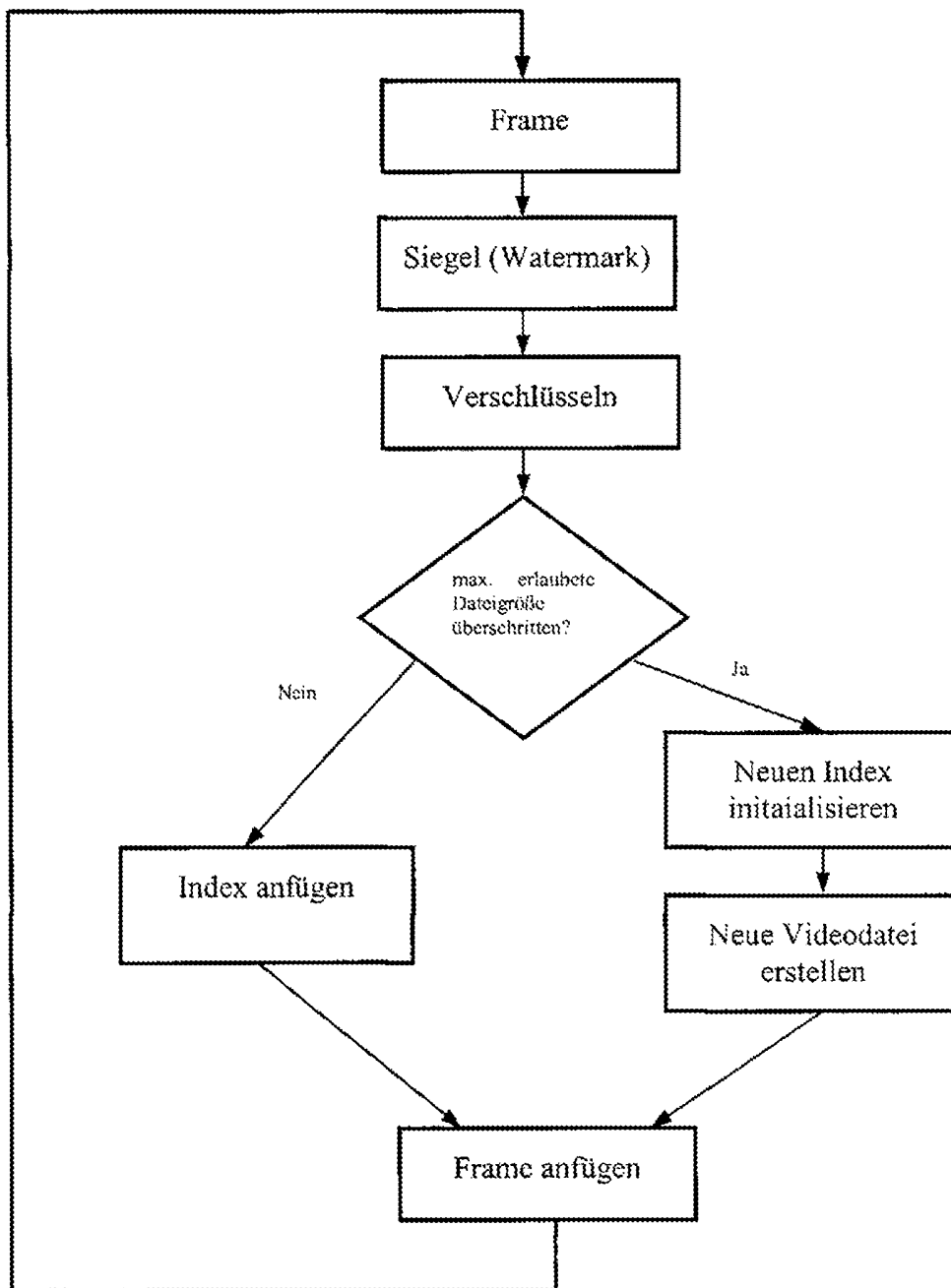


Fig. 4

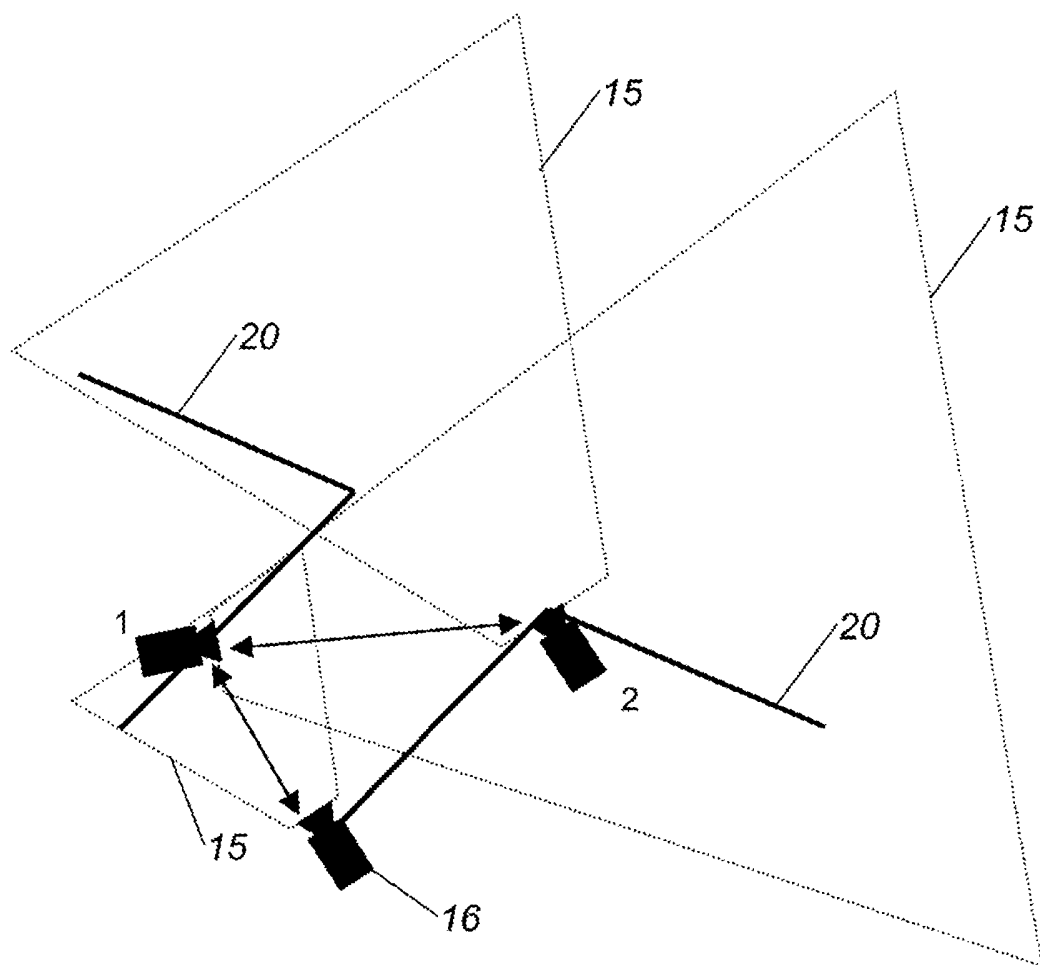


Fig. 5

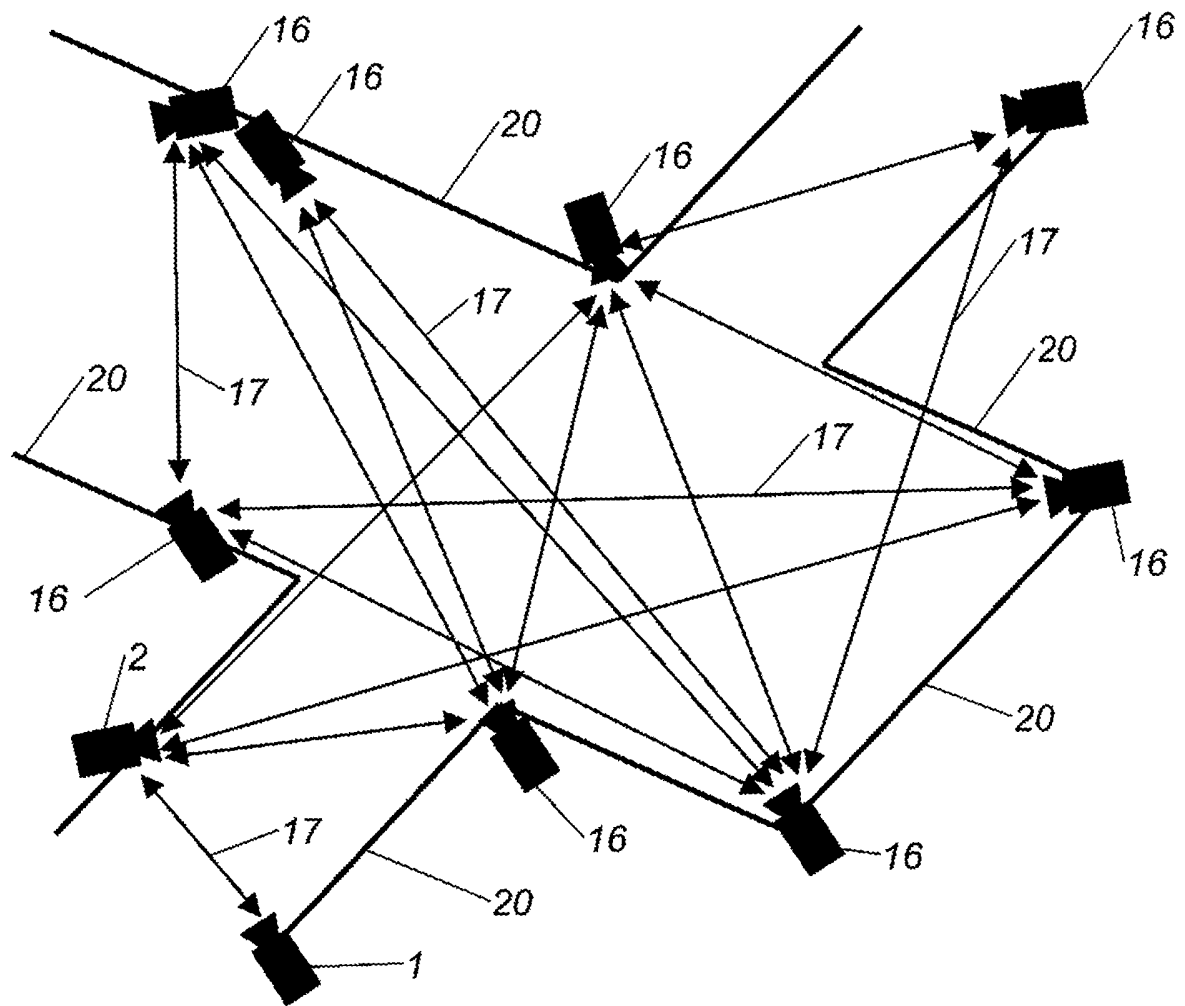


Fig. 6

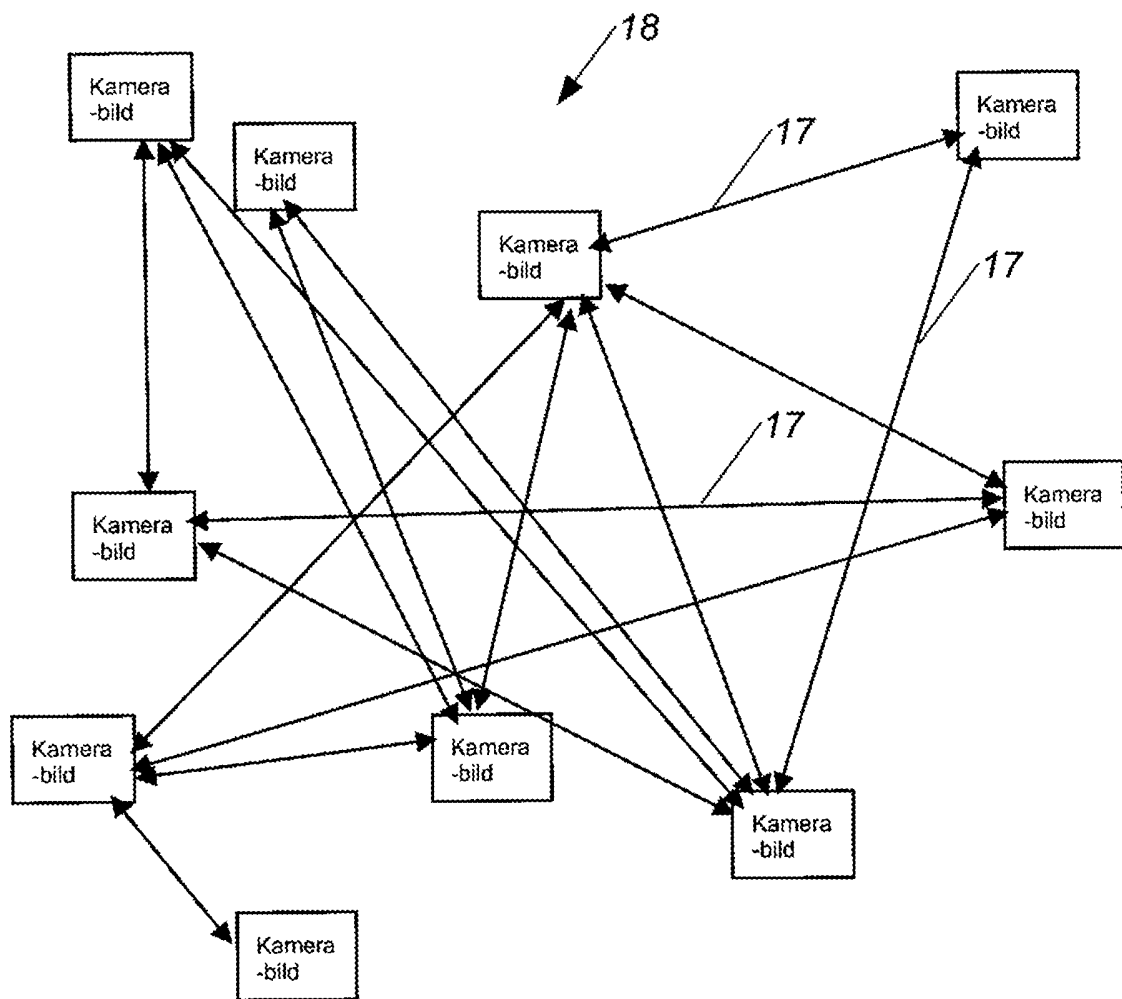


Fig. 7

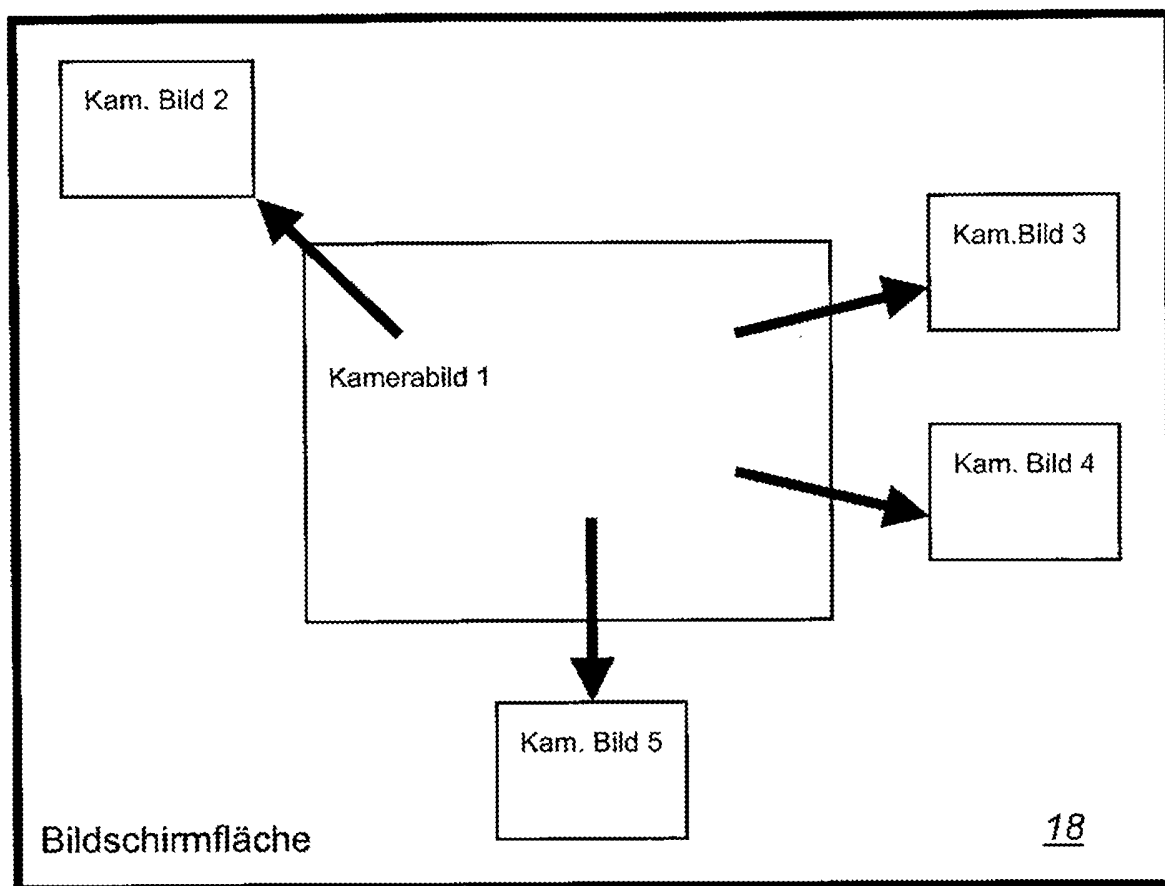


Fig. 8

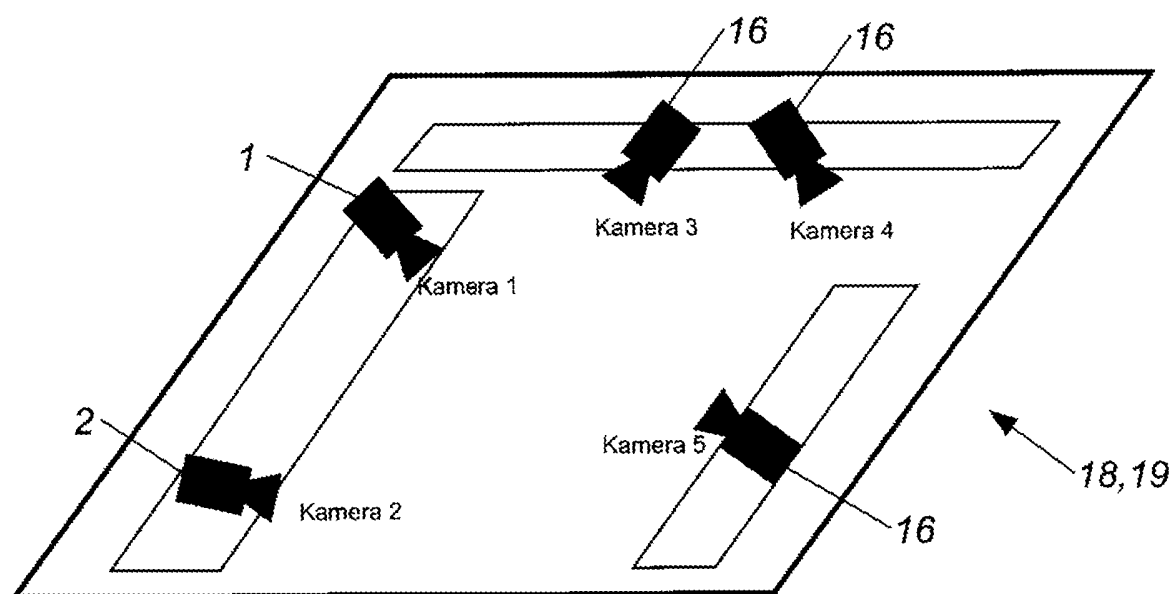


Fig. 9