



(19)  
Bundesrepublik Deutschland  
Deutsches Patent- und Markenamt

(10) **DE 602 08 567 T2** 2006.08.17

(12)

## Übersetzung der europäischen Patentschrift

(97) **EP 1 433 310 B1**

(51) Int Cl.<sup>8</sup>: **H04N 5/232** (2006.01)

(21) Deutsches Aktenzeichen: **602 08 567.5**

(86) PCT-Aktenzeichen: **PCT/GB02/04458**

(96) Europäisches Aktenzeichen: **02 762 608.4**

(87) PCT-Veröffentlichungs-Nr.: **WO 2003/032629**

(86) PCT-Anmeldetag: **01.10.2002**

(87) Veröffentlichungstag

der PCT-Anmeldung: **17.04.2003**

(97) Erstveröffentlichung durch das EPA: **30.06.2004**

(97) Veröffentlichungstag

der Patenterteilung beim EPA: **04.01.2006**

(47) Veröffentlichungstag im Patentblatt: **17.08.2006**

(30) Unionspriorität:

**0123921      05.10.2001      GB**

(74) Vertreter:

**Schoppe, Zimmermann, Stöckeler & Zinkler, 82049 Pullach**

(73) Patentinhaber:

**Hewlett-Packard Development Co., L.P., Houston, Tex., US**

(84) Benannte Vertragsstaaten:

**DE, FR, GB**

(72) Erfinder:

**GROSVENOR, David Arthur, Bristol BS36 2AD, GB**

(54) Bezeichnung: **AUTOMATISCHE FOTOGRAFIE**

Anmerkung: Innerhalb von neun Monaten nach der Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents kann jedermann beim Europäischen Patentamt gegen das erteilte europäische Patent Einspruch einlegen. Der Einspruch ist schriftlich einzureichen und zu begründen. Er gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist (Art. 99 (1) Europäisches Patentübereinkommen).

Die Übersetzung ist gemäß Artikel II § 3 Abs. 1 IntPatÜG 1991 vom Patentinhaber eingereicht worden. Sie wurde vom Deutschen Patent- und Markenamt inhaltlich nicht geprüft.

## Beschreibung

**[0001]** Man interessiert sich schon immer für die Automatisierung der Photographie, in den letzten Jahrzehnten schenkte man ihr jedoch immer mehr Aufmerksamkeit. Automatische Belichtungssteuerung und Autofokus sind bei den meisten Kameras nun standardmäßige Merkmale, und viele Kameras weisen andere elektronisch gesteuerte oder elektrische Merkmale auf. Video- und Standbildkameras mit elektronischen Bildsensoren in Form eines oder mehrerer lichtempfindlicher Arrays stehen zunehmend der Öffentlichkeit zur Verfügung und sind populär, da sie Einrichtungen bieten, die bisher nicht verfügbar waren, einschließlich der Möglichkeit, Bilder billig aufzuzeichnen, nur diejenigen Bilder, die als akzeptabel erachtet werden, auszuwählen, zu bearbeiten und zu drucken, und dies unter Verwendung lediglich eines PC und von Zusatzgeräten zu Hause durchzuführen und die Bilder unter Verwendung von E-Mail zu senden und zu empfangen.

**[0002]** Jedoch ist bei Kameras, die „von der Stange“ gekauft werden, allgemein immer noch eine Bedienperson erforderlich, um bezüglich der Zusammensetzung der tatsächlichen Aufnahme, die gemacht wird, Entscheidungen zu treffen, einschließlich Entscheidungen bezüglich der Rahmengenbung (Richtung und Zoom) und der Zeitgebung der Aufnahme. Obwohl die Einführung elektronischer Kameras die Fähigkeit bietet, eine Anzahl von Aufnahmen nacheinander oder sogar einen Videoclip aufzuzeichnen, aus denen bzw. dem anschließend ein geeignetes Bild ausgewählt werden kann, und obwohl man anschließend in einem gewissen Umfang eine Bearbeitung durchführen kann, muss die Bedienperson die Kamera trotzdem noch in die allgemein richtige Richtung richten und ungefähr um den richtigen Zeitpunkt herum die Bedientaste drücken.

**[0003]** Da sowohl elektronische Standbild- als auch Videokameras Sensoren aufweisen, die in der Lage sind, eine Abfolge von Bildsignalen (eine „Bildsequenz“) zu liefern, werden sie hierin grob als „elektronische Kameras“ bezeichnet, und ihr Ausgangssignal wird als „Bildsignal“ bezeichnet, ungeachtet dessen, ob sich ein Standbild oder eine Bildsequenz ergibt, und ungeachtet dessen, ob das Bildsignal an der Kamera oder andernorts aufgezeichnet wird. Bezugnahmen auf eine Aufnahme oder ein Bild sollten so aufgefasst werden, dass sie sowohl Standbilder als auch Bildsequenzen abdecken. Uneingeschränkte Bezugnahmen auf eine „Kamera“ sollten so aufgefasst werden, dass sie sich entweder auf eine elektronische Kamera oder auf eine Kamera, die ein lichtempfindliches Aufzeichnungsmedium verwendet, z.B. eine Filmkamera für Standbildbelichtungen oder eine Filmkamera (eine „photographische“ Kamera), beziehen.

**[0004]** Kamerasteuerungen wie z.B. Autofokus und Belichtung beziehen sich mehr auf die technische Seite des Aufnahmenmachens. Diese haben wenig oder nichts mit der Aufnahmezusammensetzung, einschließlich einer Rahmengenbung und Zeitgebung, zu tun, es besteht derzeit jedoch immer mehr Interesse an diesem Aspekt und an der Bereitstellung von intelligenten Standbild- und Videokameras, die eine Einrichtung zum Aufnahmenmachen oder zur Aufnahmenauswahl aufweisen, bei der eine bewusste Eingabe durch eine Bedienperson entweder weiter reduziert ist oder ganz wegfällt.

**[0005]** Eine mögliche Lösung besteht darin, das gesamte Bildsignal von einer kontinuierlich betriebenen Kamera oder die gesamte Sequenz von einer Filmkamera, d.h. eine Bildsequenz, zu speichern und daraus zu einer späteren Zeit geeignete Einzelbilder oder reduzierte Bildsequenzen manuell auszuwählen. Obwohl dies die Last einer Bedienperson einer Kamera verringert oder das Erfordernis einer Kamera-Bedienperson ganz und gar vermeidet, erfordert dieser Lösungsansatz nicht nur einen großen Umfang an Bildsignal- oder Filmspeicher, sondern er ist auch sehr zeitaufwändig und somit teuer, wenn eine menschliche Bedienperson später während des Auswahlprozesses die gespeicherte Bildsequenz durchsehen muss. Somit ist es äußerst wünschenswert, entweder ein gewisses Maß an Bildauswahl vor der Speicherung zu haben oder, wenn die gesamte Bildsequenz gespeichert wird, mit derselben ein weiteres Signal aufzuzeichnen, das angibt, wo wahrscheinlich die interessantesten Abschnitte der Sequenzen zu finden sind.

**[0006]** Bekannte Typen von automatischen Kameravorrichtungen, bei denen ein gewisser Grad an Aufnahmezusammensetzung bewerkstelligt wird, umfassen ein breites Spektrum. An einem Ende liegen Geräteformen, bei denen die Bedingungen zum Aufnahmenmachen streng vorbestimmt und äußerst stark eingeschränkt sind, z.B. Kabinen zum Schießen von Passphotos, bei denen die Kamera mit einer feststehenden Zeitgebung betrieben wird, nachdem ein in einer im Wesentlichen festgelegten Position sitzender Kunde Geld in die Maschine eingeworfen hat.

**[0007]** Mehr Freiheiten ermöglichen Vorrichtungen, bei denen das Aufnahmenmachen ansprechend auf eine vorbestimmte Bedingung oder einen vorbestimmten Satz von Bedingungen automatisch ausgelöst wird, beispielsweise ansprechend auf das Ergebnis einer algorithmischen Berechnung, die auf das Ausgangssignal eines oder mehrerer Sensoren hin arbeitet. Ein Problem bei Vorrichtungen dieser Art liegt jedoch darin, den Satz von Bedingungen, die notwendig sind, damit es möglich wird, Aufnahmen zu machen, zu definieren.

**[0008]** Unter bestimmten Umständen sind die vor-

geschriebenen Bedingungen physisch vorbestimmt. Sicherheitskameras, die ihren Betrieb ansprechend auf eine Alarmeingabe beginnen, und Verkehrsgeschwindigkeitskameras, die darauf reagieren, dass ein Fahrzeug zu schnell fährt, fallen in diese Kategorie. Es ist auch bekannt, Aufnahmen von Besuchern bei einer Fahrt in einem Themenpark zu machen, indem eine geeignet positionierte feststehende Kamera ausgelöst wird, die betätigt wird, wenn der Wagen, in dem der Besucher fährt, einen Schalter auslöst, beispielsweise indem er einen Infrarotstrahl durchschneidet. Derartige Systeme sind zum Beispiel in der internationalen Patentanmeldung Nr. WO 98/10358 (Goldberg), der europäischen Patentanmeldung Nr. 0 953 935 (Eastman Kodak), der US-Patentanmeldung Nr. 5,655,053 (Renie) und der US-Patentanmeldung Nr. 5,576,838 (Renie) offenbart.

**[0009]** Andere Systeme lassen deutlich mehr Freiraum, durch Erkennung oder Erfassung eines Merkmals, das dem abzubildenden Merkmal zugeordnet ist. Somit offenbart die US-Patentschrift Nr. 5,844,599 (Hildin) ein Videokamerasystem zur Verwendung bei einer Gruppe von Menschen um einen Konferenztisch, wobei jede Person ein Mikrofon trägt und die Kamera angeordnet ist, um zu reagieren, indem sie direkt auf die Person, die jeweils gerade spricht, gerichtet wird und die Bewegungen dieser Person nachverfolgt. Dies ist ein System, das mehr zum Betrachten gedacht ist als zur Bildaufzeichnung. Die europäische Patentanmeldung Nr. 0 660 131 (Osen) beschreibt ein System, bei dem globale Ortungssysteme verwendet werden, um ein Merkmal, z.B. ein Luftfahrzeug, nachzuverfolgen, so dass Kameras automatisch darauf gerichtet werden können.

**[0010]** Diesbezüglich verwendet man zunehmend Markierungen zum Identifizieren von Merkmalen, die abgebildet werden/oder zum Auslösen der Kamera. Dies ist in der internationalen Patentanmeldung Nr. WO 98/10358 (Goldberg), die oben erwähnt wurde, und in der US-Patentschrift Nr. 5,694,514 (Evans) der Fall, um zu ermöglichen, dass Aufnahmen unterschiedlicher Merkmale ohne weiteres identifiziert werden können. Die internationale Patentanmeldung Nr. WO 98/10358 (Goldberg) beschreibt eine Anzahl von Kamerasystemen, die die Verwendung einer Markierung umfassen, um eine Kamera zu betätigen und um Daten zu liefern, um das Bild der die Markierung tragenden Person zuzuordnen; bei einer Variante wird Gesichtserkennung verwendet, um das Merkmal zu identifizieren, wobei ausgehend von einem Bild und anderen relevanten Daten, die anfänglich in das System eingespeist werden, gearbeitet wird.

**[0011]** Im Vergleich zu einem persönlichen Kamerabenutzer, der eine in der Hand zu haltende Kamera zu jederzeit und immer dann verwenden kann, wenn er Aufnahmen von etwas machen möchte, das er als

interessant betrachtet, bleiben jedoch im gesamten oben erwähnten Stand der Technik beträchtliche Einschränkungen bezüglich der Art und Weise des Aufnahmenmachens bestehen. Die meisten automatischen Kameravorrichtungen verwenden entweder eine feststehende Kamera und/oder weisen enge Einschränkungen bezüglich der Umstände auf, unter denen die Kamera ausgelöst wird, z.B. ansprechend auf eine Markierung oder eine an Bedingungen geknüpfte oder eine Alarmeingabe, die einer eingeschränkten Bedingung oder einem eingeschränkten Satz von Bedingungen zugeordnet ist.

**[0012]** Bei kommerziellen Einrichtungen wie z.B. Themenparks sind die Kameras im Wesentlichen feststehend und positioniert, um unter bestimmten Umständen ein Bild aufzunehmen, z.B. wenn erwartet wird, dass sich zumindest manche der gespeicherten Bilder als ausreichend zusammengesetzt und ansprechend erweisen werden, so dass sie für einen potentiellen Käufer interessant sind.

**[0013]** Es wäre wünschenswert, in der Lage zu sein, eine Kamera zu liefern, die tragbar ist und die verschiedene Bilder ungeachtet der Position der Person automatisch aufnehmen kann, wie es für eine manuell bediente Kamera der Fall wäre.

**[0014]** Somit könnte man die Überlegung anführen, dass ein weiteres wünschenswertes Merkmal einer automatischen Kameravorrichtung die Bereitstellung eines Signals (eines „Ausgeprägtheits“-Signals) ist, das angibt, wenn das betrachtete Bild vielleicht nicht ein Bild eines vorbestimmten Typs ist, jedoch trotzdem für eine Person, z.B. den Kamerabenutzer, interessant sein kann. Dieses Signal könnte verwendet werden, um einen selektiven Kamerabetrieb oder eine selektive Bildspeicherung in Echtzeit auszulösen, oder es könnte in Verbindung mit einem kontinuierlich gespeicherten Bildsignal in einem späteren Bildauswahlprozess verwendet werden; wenn das Signal mehr als zwei Werte aufweist, d.h. die relative Wahrscheinlichkeit angibt, dass sich das Bild auf ein interessantes Merkmal bezieht, so könnte eine Kombination dieser Techniken eingesetzt werden, nämlich eine Speicherung aller Signale, die möglicherweise von Interesse sind, während andere Signale zurückgewiesen werden, und eine spätere Auswahl aus den gespeicherten Signalen, wobei anfänglich diejenigen Teile des Signals bevorzugt werden, die als interessanter angegeben werden. Ein derartiges Merkmal wird bezüglich der vorliegenden Erfindung betrachtet.

**[0015]** Es ist bekannt, dass Töne, die von einer Person stammen, ein Hinweis auf die Reaktion der Person auf die Umgebung sein können, in der sie sich befindet. Derartige Reaktionen können auf ein bestimmtes Ereignis oder auf das allgemeine Ambiente der Umgebung erfolgen. Die Reaktion könnte bei-

spielsweise eine Reaktion auf eine betrachtete oder sichtbare Szene oder ein betrachtetes oder sichtbares Ereignis sein, auf eine Interaktion mit einer anderen Person, auf Töne, ob sie plötzlich (z.B. ein Knall) oder anhaltend (z.B. Musik) sind, und auf andere Reize wie z.B. diejenigen, die erfahren werden, wenn man an einem Ereignis teilnimmt oder wenn man sich amüsiert, z.B. bei einer Fahrt auf einem Rummelplatz. Wie auch immer die Reaktion ausfällt, wenn sie auftritt, kann die durch die Person betrachtete Szene entweder die Ursache der Reaktion darstellen oder, zu einem späteren Zeitpunkt, zumindest an die Ursache der Reaktion erinnern. Demgemäß werden bei Ausführungsbeispielen der vorliegenden Erfindung die von einer Person ausgehenden Töne bei der Erzeugung eines Ausgeprägtheitssignals verwendet. Die Erfindung lässt sich besonders bei Kameras verwenden, die tragbar sind oder von einer Person mit sich geführt werden, ist jedoch nicht hierauf beschränkt und kann bei Kameras verwendet werden, die an leblosen Objekten, die sich bewegen können, oder an feststehenden Positionen angebracht sind oder die eine festgelegte Bahn verfolgen.

**[0016]** Bei einem ersten Aspekt liefert die vorliegende Erfindung eine automatische Bilderzeugungsvorrichtung, die eine Kamera, eine Schallerfassungseinrichtung, die in der Lage ist, Klänge von einer vorbestimmten Person vorzugsweise zu empfangen oder zu identifizieren und ansprechend darauf ein Audiosignal zu erzeugen, eine Audioeinrichtung zum Bereitstellen eines Ausgeprägtheitssignals ansprechend auf das Audiosignal sowie eine Steuereinrichtung, die auf das Ausgeprägtheitssignal anspricht, um die Kamera zu steuern, umfasst, wobei die Audioeinrichtung dazu angeordnet ist, das Ausgeprägtheitssignal bereitzustellen, indem sie auf zumindest eines der folgenden anspricht: (a) die Prosodie der Sprache, (b) nicht-sprachliche stimmliche Klänge und (c) nicht-stimmliche Klänge.

**[0017]** Bei einem zweiten Aspekt liefert die Erfindung eine automatische Bilderzeugungsvorrichtung gemäß Anspruch 4.

**[0018]** Die Erfindung erstreckt sich auf die Verfahren zur Bilderzeugung unter Verwendung einer Vorrichtung gemäß dem ersten und dem zweiten Aspekt.

**[0019]** Wie später im Einzelnen beschrieben wird, ist eine bevorzugte Schallerfassungseinrichtung ein Schallleitmikrophon, das dafür ausgelegt ist, am Kopf einer Trägerperson (der „vorbestimmten Person“) angebracht zu werden, um vorzugsweise Klänge zu erfassen, die von dem Kopf und/oder der Kehle der Trägerperson emittiert werden, und ansprechend darauf ein Audiosignal zu erzeugen. Jedoch kennt man andere Techniken zum vorzugsweisen Empfangen oder Identifizieren von Klängen von einer vorbestimmten Person, und eine Nutzung derselben statt einer Nut-

zung eines Schallleitmikrophons fällt in den Schutzbereich der vorliegenden Erfindung. Diese Techniken umfassen die Verwendung einer Spracherkennungssoftware, die auf ein Mikrophon einwirkt, das in der Lage ist, beispielsweise Klänge von mehr als einer Person zu empfangen – eine derartige Software kann dahingehend trainiert werden, Sprache, insbesondere von einem spezifischen Benutzer, zu erkennen. Trotzdem kann die Verwendung eines Schallleitmikrophons oft eine besonders einfache und effektive Art und Weise liefern, die beabsichtigte Funktion zu erfüllen.

**[0020]** Angebrachte Schallleitmikrophone, die allen drei Aspekten der Erfindung gemein sind, sind hinreichend bekannt und umfassen normalerweise einen Schwingungswandler zum Erfassen von Schallschwingungen, die durch den Körper der Trägerperson, bei am Kopf oder am Ohr angebrachten Mikrophonen durch den bzw. die Schädelknochen, geleitet werden. Man unterscheidet sie von Mikrophonen, die Schall aufnehmen, der durch die Luft übertragen wird.

**[0021]** Übliche Beispiele von am Ohr angebrachten Mikrophonen sind in der internationalen Patentanmeldung Nr. WO 86/07157 (ACR Electronics); der US-Patentschrift Nr. 4,588,867 (Konomi) und in der US-Patentschrift Nr. 5,692,059 (Kruger) beschrieben. Andere am Schädel angebrachte Mikrophone sind oben auf dem Kopf befestigt. All diese umfassen einen Schwingungswandler zum Aufnehmen von Klängen, die direkt durch den Kopf übertragen werden, normalerweise mittels Leitung durch Knochen. Eine weitere typische Form eines am Kopf angebrachten Mikrophons ist in physischem Kontakt mit der Kehle angebracht. Klänge, die durch die Luft übertragen werden, können durch diese Mikrophone ebenfalls erfasst werden. Schaltungen zum Betreiben derartiger Mikrophone sind beispielsweise in der internationalen Patentanmeldung Nr. WO 00/21194 (Resound Corporation) und in der europäischen Patentanmeldung Seriennummer 0 984 660 (Nippon Telephone and Telegraph Corporation) offenbart.

**[0022]** Man kennt Kameras, die durch ein Schallsignal aktiviert werden. Beispiele sind in den US-Patentschriften Nrn. 5,014,079 und 5,107,288, die beide an die Konica Corporation erteilt wurden; und in der US-Patentschrift Nr. 5,162,832 (Konishoroku Photo Ind.) offenbart, im Allgemeinen sind sie jedoch dahingehend angeordnet, lediglich auf ein bestimmtes Klangmerkmal, in diesen bestimmten Fällen z.B. die Intensitätshüllkurve (oder den Schalldruck), anzusprechen, und das Mikrophon befindet sich an der Kamera und ist kein Schallleitmikrophon.

**[0023]** Aus der US-Patentschrift Nr. 5,844,599 (Lucent Technologies Inc.) kennt man ferner ein automatisches Kamerasystem zur Verwendung bei Konfe-

renzen, wo jeder Teilnehmer mit einem Mikrophon ausgestattet wird, das als Markierung fungiert, wenn diese Person spricht, wobei das System so ausgelegt ist, dass die Kamera automatisch auf den Sprechenden bzw. den am lautesten Sprechenden gerichtet wird. Obwohl dieses System persönliche Mikrophone beinhaltet, sind sie nicht speziell Schallleitmikrophone, und der einzige Schallparameter, der verwendet wird, ist die Schallintensität.

**[0024]** Keiner dieser Aspekte des Standes der Technik bezieht sich auf die Verwendung eines Audiosignals von einem am Kopf oder an der Kehle angebrachten Schallleitmikrophon zur Steuerung des Betriebs einer Kamera oder Steuerung eines aus derselben produzierten Bildsignals.

**[0025]** Ferner gibt es Systeme zur Klassifizierung von Klängen, und die oben erwähnte Audioeinrichtung können eine Audioanalyseeinrichtung zum Analysieren des Signalverlaufs des Audiosignals umfassen, um es einer eines Satzes von vorbestimmten Kategorien zuzuordnen.

**[0026]** Die US-Patentschrift Nr. 5,712,953 (Langs) offenbart ein System zum Klassifizieren eines Audiosignals als Musik oder Nicht-Musik und bezieht sich auf einen anderen Aspekt dieses Standes der Technik. Für die vorliegende Erfindung vielleicht relevanter ist, dass die US-Patentschrift Nr. 5,918,223 (Muscle Fish) ein System für eine detailliertere Klassifizierung eines Audiosignals durch einen Vergleich mit gegebenen Schallsignalen offenbart. Das System sei bei Multimediadatenbankanwendungen und Internet-Suchmaschinen nützlich beim Finden von Audiodatendateien mit einem Klang, der ähnlich den gegebenen Klängen ist. Als Beispiel ist die Identifizierung von Applaus und Beifallsrufen auf einem Soundtrack angegeben. Durch den Benutzer können Klangsparten wie z.B. „Vogelklänge“ oder „Rockmusik“ erzeugt werden.

**[0027]** Bei einer grundlegenden Form der Vorrichtung der Erfindung kann die Audioanalyseeinrichtung dazu angepasst sein, lediglich eine bestimmte Klangform, z.B. Gelächter, die von der vorbestimmten Person ausgeht, zu erkennen, und in einem derartigen Fall besteht der elementarste Satz von Kategorien aus „Gelächter“ und „Sonstiges“.

**[0028]** ES ist jedoch möglich, eine Anzahl von Klängen, die von der vorbestimmten Person ausgehen, zu kategorisieren, wenn auch manche mit mehr Schwierigkeiten als andere zu kategorisieren sind, um somit die Bandbreite identifizierbarer Kategorien und/oder von Ereignissen, auf die sich die Kategorien beziehen, zu vergrößern. Klänge, die von dem Kopf einer vorbestimmten Person ausgehen, können in Nicht-Sprache und Sprache aufgeteilt werden, die getrennt betrachtet werden müssen.

**[0029]** Übliche Nicht-Sprache-Klänge umfassen stimmliche Klänge wie z.B. Gelächter, Weinen, Summen, Husten und Niesen und nicht-stimmliche Klänge wie z.B. Schniefen, Kratzen, Essen, Zähneknirschen und Aufstoßen.

**[0030]** Sprache selbst kann in Wörter und die Art und Weise, wie sie gesprochen werden (Prosodie) aufgeteilt werden. Worterkennungstechniken sind ausreichend fortgeschritten, so dass viele Wörter genau erkannt werden können. Die Prosodie umfasst Merkmale wie das Stimmintensitätsmuster (angehobene oder abgesenkte Stimme) oder ein Muster von Variationen, wie es z.B. dem Lesen einer Abfolge von Sätzen oder eines Gedichts, dem Stellen einer Frage oder dem Treffen einer Aussage zugeordnet ist; Zögern; die Art und Weise, wie Worte gesprochen werden, d.h. die unterschiedlichen Betonungen, die Silben zukommen; die Frequenzen in der Stimme; die Sprechgeschwindigkeit; die Dauer von Teilen der Sprache. Sie besteht aus einem phonologischen Aspekt, einem Identifizieren diskreter abstrakter Spracheinheiten, und einem phonetischen Aspekt, der sich auf kontinuierlich veränderliche akustische Informationen, die derartigen Einheiten zugeordnet sind, bezieht, und können die Stimmung der vorbestimmten Person, z.B. Aufregung, Ärger oder Entspannung, widerspiegeln.

**[0031]** Manche Klänge, z.B. Singen, können als sowohl in Sprache- als auch in Nicht-Sprache-Zuordnungen fallend betrachtet werden.

**[0032]** Bei der Erfindung kann die durch die Audioanalyseeinrichtung durchgeführte Kategorisierung zusätzlich eine Funktion der Gesamtsignal(Hüllkurven)amplitude sein, obwohl dies in jeglichem Fall, in dem Sprache beteiligt ist, als Teil der Prosodie betrachtet werden kann.

**[0033]** Somit kann die Audioanalyseeinrichtung zur Erkennung von stimmlichen Klängen und zum Erkennen dessen, ob die Klänge einer Sprache oder einer Nicht-Sprache zugeordnet sind, angeordnet sein. Wenn Sprache im Spiel ist, kann die Audioanalyseeinrichtung für eine Worterkennung und/oder für die Erkennung von Merkmalen, die der Prosodie zugeordnet sind, angeordnet sein. Jedoch kann die Audioanalyseeinrichtung alternativ oder zusätzlich dazu zur die Erkennung von nicht-sprachlichen stimmlichen Klängen oder zur Erkennung von nicht-stimmlichen Klängen angeordnet sein.

**[0034]** Zusätzlich kann die Audioanalyseeinrichtung dahingehend angeordnet sein, die Gesamtsignalamplitude oder -hüllkurve zu berücksichtigen, wenn sie dem Audiosignal eine Kategorie zuweist. Beispielsweise kann sie dazu angeordnet sein, zwischen lautem Gelächter und leisem Gelächter zu unterscheiden.

**[0035]** Die Audioanalyse kann mit einer vorbestimmten Serie von Referenzsignalverläufen ausgestattet sein, bezüglich derer die Kategorisierung durchzuführen ist. Alternativ oder zusätzlich dazu und vorzugsweise kann sie so angeordnet sein, dass sie trainiert werden kann, neue Klänge zu erkennen, so dass sie an die spezifischen Eigenheiten der vorbestimmten Person angepasst werden kann.

**[0036]** Eine Worterkennung kann in Bezug auf eine freiwillige und unfreiwillige Kamera- oder Bildsignalsteuerung nützlich sein. Worte und Ausdrücke wie z.B. „Hallo“, „Auf Wiedersehen“, „Wow“ und „Mach schon“ können bezüglich einer unfreiwilligen Einleitung eines Bildaufnahmeverganges nützlich sein, wenn man es mit Leuten zu tun hat oder auf ein Ereignis wartet, wohingegen ein Befehl wie z.B. „Nimm das (auf)“ („Take that“) als absichtlicher Befehl, eine Aufnahme zu machen, erkannt werden mag. Worte wie z.B. „Ja“, wenn sie mit einem Anflug von Triumph ausgesprochen werden, oder „Nein“, wenn sie im Zweifel ausgesprochen werden, sind Beispiele einer mit Prosodie kombinierten Worterkennung. Lachen und Summen können als zwei Beispiele von nicht-sprachlichen stimmlichen Klängen herangezogen werden, die das Aufnahmenmachen oder Auswählen von Aufnahmen einleiten könnten. Lautes Gelächter kann eine bedingungslose derartige Einleitung bewirken, wohingegen sanfteres Lachen eventuell mit einem weiteren an eine Bedingung geknüpften Signal kombiniert werden muss, um zu bestimmen, ob dies bewirkt wird oder nicht.

**[0037]** Ein Ausprägtheitssignal kann gemäß der auf diese Weise durchgeführten Kategorisierung entwickelt werden. In seiner einfachsten und möglicherweise praktischsten Form ist das Ausprägtheitssignal binär, so dass ansprechend auf den von dem Mikrophon stammenden Klang jederzeit ein „Ja“- oder „Nein“-Signal ausgegeben wird. Dies kann alleine oder in einer logischen Kombination mit anderen Steuersignalen zur Verwendung als ein Steuersignal eingesetzt werden. Letzteres kann beispielsweise zum Steuern der Kamera oder des Bildsignals verwendet werden, beispielsweise indem man der Kamera sagt, sie soll eine Aufnahme machen oder eine Bilderzeugungssequenz beginnen, oder indem man ein Gatter für eine selektive Übertragung, Anzeige oder Aufzeichnung des Bildsignals betreibt.

**[0038]** Alternativ dazu kann das Ausprägtheitssignal eine Bandbreite von diskreten oder kontinuierlichen Werten aufweisen. Dies kann für jemanden nützlich sein, der eine vollständige Bildsequenz durchsieht, entweder zum Identifizieren bestimmter Klänge, die einem bestimmten Wert des Ausprägtheitssignals zugeordnet sind, oder zum Suchen nach den interessantesten Klängen, wie sie durch ein vorbestimmtes Ordnen von Werten, die unterschiedlichen Klängen zugewiesen sind, definiert sind. Auch

in einer solchen Form könnte es beispielsweise gemäß einem vorbestimmten Algorithmus auf numerische Weise mit anderen Signalen kombiniert werden, um ein Steuersignal des Typs zu erzeugen, der in dem vorstehenden Paragraph erwähnt wurde, möglicherweise nachdem das resultierende Signal einer Schwellwertbewertung unterzogen wurde.

**[0039]** Um ein numerisches Ausprägtheitssignal abzuleiten, können verschiedene erkannte Klänge (und dies könnte umfassen, dass ähnliche Klänge, die unterschiedliche Intensitäten aufweisen, als unterschiedliche erkannte Klänge aufgefasst werden) unterschiedlichen Werten zugewiesen werden. Dies kann ausreichend sein, oder, wenn die Intensität nicht berücksichtigt wurde, kann der zugewiesene Wert gemäß der Intensität des Klanges kontinuierlich modifiziert werden.

**[0040]** Bei einer bevorzugten Form der Erfindung trägt die vorbestimmte Person, vorzugsweise ein Träger eines Schalleitmikrophons, die Kamera an sich, und vorzugsweise wird die Kamera so getragen, dass sie von dem Träger aus nach vorne zeigt. Obwohl die Kamera an der Brust angebracht sein könnte, z.B. als Anstecker, ist besonders bevorzugt, dass die Kamera mit einer Befestigungsvorrichtung versehen ist, um sie oder zumindest die Bilderzeugungsoptik derselben am Kopf der Person anzuordnen, beispielsweise im Rahmen einer Brille oder in einer Kopfbedeckung wie z.B. einem Hut. Auf diese Weise folgt die Ausrichtung der Kamera präziser der Blickrichtung des Trägers. Wenn es angebracht ist, kann das Gesichtsfeld der Kamera beträchtlich größer gestaltet sein als das der Trägerperson, möglicherweise um die gesamte Bandbreite der Augenbewegung abzudecken. Dies ist besonders vorteilhaft, wenn die Anzahl verfügbarer Pixel, z.B. eines Sensorarrays bei einer Halbleiterkamera, entsprechend erhöht ist, um die Auflösung für eine Ansicht aufrechtzuerhalten, die an das normale menschliche Gesichtsfeld angenähert ist.

**[0041]** Zum Zweck einer echten unabhängigen Tragbarkeit bezüglich des ersten Aspekts der Erfindung ist die vorbestimmte Person ferner mit einem Aufzeichnungsgerät zum Aufzeichnen sowohl der Bild- als auch der Audiosignale ausgestattet, z.B. zur späteren Entwicklung des Ausprägtheitssignals zum Steuern eines wiedergegebenen Bildsignals oder zum Angeben von Teilen eines wiedergegebenen Bildsignals, die für den Betrachter von Interesse sein könnten.

**[0042]** Zum Zweck einer echten unabhängigen Tragbarkeit bezüglich des zweiten und des dritten Aspekts der Erfindung kann die vorbestimmte Person auch mit der Rudioeinrichtung zum Bereitstellen des Ausprägtheitssignals ausgestattet sein. Dieses kann dann neben dem vollständigen Bildsignal aufgezeichnet werden, z.B. zum Zweck einer Verwen-

derung, die in dem vorstehenden Abschnitt erwähnt wurde, oder zur Steuerung des Bildsignals, zum Beispiel indem es zu einer Anzeige oder einem Aufzeichnungsgerät durchgeschaltet wird, oder zur Kamerasteuerung oder -betätigung.

**[0043]** Jedoch kann jegliches der Audio-, Ausgeprägtheits- und Bildsignale zu einer entfernten Stelle übertragen werden, wenn es angebracht ist. Eine Gelegenheit, bei der eine Übertragung des Audio- oder Ausgeprägtheitssignals angebracht ist, ist, wenn die vorbestimmte Person die Kamera nicht an sich trägt. Beispielsweise kann die Vorrichtung ein Teil einer Installation sein, bei der die Kamera feststehend ist, um eine vorbestimmte Szene zu betrachten, und die Person dann als Photograph agiert und den Betrieb der Kamera willkürlich (Wortbefehle) oder unwillkürlich (nicht-sprachliche Klänge) steuert oder das Durchschalten des Bildsignals steuert, oder zumindest die Ausgeprägtheit eines aufgezeichneten Signals angibt. Dies könnte bei einer besonderen Gelegenheit wie z.B. einer Geburtstagsfeier von Nutzen sein. Es ist zu beachten, dass die Kamera nicht feststehend sein muss und beispielsweise eine vorbestimmte Spur verfolgen könnte, wie bei einer Fahrt auf einem Rummelplatz, wobei die vorbestimmte Person verfolgt wird. Andere geeignete Anordnungen und Verwendungszwecke der Erfindung werden dem Leser zweifellos einleuchten.

**[0044]** Obwohl ein notwendiger Teil bevorzugter Ausführungsbeispiele der Erfindung die Verwendung eines Schallleitmikrophons ist, sollte man beachten, dass ein oder mehrere weitere Mikrophone bereitgestellt werden können und dass zumindest ein solches weiteres Mikrophon ein „herkömmliches“ Mikrophon zum Empfangen von Schall über die Luft sein kann.

**[0045]** Das Signal von dem „herkömmlichen“ Mikrophon kann aufgezeichnet werden, z.B. mit dem Bild- bzw. Schallleitmikrophonsignal, um es zu einem späteren Zeitpunkt wiederzugeben. Diesbezüglich sollte man bedenken, dass es normaler wäre, das Signal von einem „herkömmlichen“ Mikrophon zu Hörzwecken wiederzugeben, wobei das Signal von dem Leitmikrophon zu diesem Zweck zu einem eher unechten Klang führt.

**[0046]** Alternativ oder zusätzlich dazu kann es mit dem Signal von dem Schallleitmikrophon kombiniert werden. Beispielsweise kann es möglich sein, einen beträchtlichen Teil des Schalls, der durch eine Person erzeugt wird, die ein Schallleitmikrophon an sich trägt, von dem durch das „herkömmliche“ Mikrophon erzeugten Signal zu subtrahieren, wobei dabei Klänge übrig bleiben würden, die als um den Träger des Schallleitmikrophons herum vorhandene Klänge betrachtet werden könnten. Dies würde auch ermöglichen, abzuleiten, ob die Person, die das Leitmikrophon trägt, sich in einer Unterhaltung mit jemand an-

derem befand, z.B. durch Analysieren oder Vergleichen des Musters des Umgebungsschalls und dessen, der durch das Leitmikrophon erfasst wurde. Dies könnte ein Merkmal sein, das das Ausgeprägtheitssignal bestimmt, d.h. Letzteres könnte mittels einer Analyse der Ausgangssignale einer Kombination von Mikrophenen abgeleitet werden, und eine derartige Kombination kann ein „herkömmliches“ Mikrophon umfassen.

**[0047]** Die Kamera kann in der Lage sein, Standbilder und/oder Bildsequenzen festzuhalten. Sie kann ein „herkömmliches“ Mikrophon umfassen, wie es in dem vorstehenden Abschnitt erwähnt wurde.

**[0048]** Dem Leser wird einleuchten, dass die Audioanalyse in den meisten Fällen auf effizientere und effektivere Weise durch die Verwendung von Fourier-Techniken, die an sich hinreichend bekannt sind, durchgeführt werden kann.

**[0049]** Weitere Merkmale und Vorteile der Erfindung ergeben sich bei Betrachtung der beigefügten Patentansprüche, auf die der Leser verwiesen wird, und bei einer Lektüre der folgenden ausführlicheren Beschreibung eines bevorzugten Ausführungsbeispiels der Erfindung, die unter Bezugnahme auf die beiliegenden Zeichnungen gegeben wird, bei denen die [Fig. 1](#) bis [Fig. 4](#) ein erstes bis viertes Ausführungsbeispiel der Erfindung in schematischer Form zeigen und [Fig. 5](#) eine Audioanalyse in Form eines Phantomdiagramms beispielhaft darstellt.

**[0050]** Dort, wo es angebracht ist, bezeichnen gleiche Bezugszeichen ähnliche Posten in jeder der Zeichnungen.

**[0051]** Bei [Fig. 1](#) erzeugt eine am Kopf angebrachte elektronische Kamera **1** ein Bildsignal **2** zur Speicherung in einem Mehrkanal-Aufzeichnungsgerät **3**. Ein am Ohr angebrachtes Schallleitmikrophon **4** erzeugt ein Audiosignal **5** zum synchronen Aufzeichnen mit dem Bildsignal an dem Aufzeichnungsgerät **3**. Anschließend wird das abgespielte Audiosignal **9** während des Abspielens an eine Schallkategorisierungsschaltung **10** geleitet, die je nachdem, ob die Form des Schallsignalverlaufs annähernd einer einer Anzahl von gespeicherten Signalverlaufsformen entspricht, ein binäres (Ausgeprägtheits-)Signal **11** erzeugt. Das Signal **11** wird verwendet, um ein Gatter **7** zur selektiven Übertragung **8** des abgespielten Bildsignals **6** zu betreiben. Bei einer Variante wird das Signal **11** dazu verwendet, ein Signal an einen menschlichen Betrachter des abgespielten und angezeigten Bildsignals **6** zu betreiben, um anzugeben, dass besondere Aufmerksamkeit darauf gelenkt werden sollte.

**[0052]** Bei [Fig. 2](#) wird das Signal **5** von einem am Schädel angebrachten Schallleitmikrophon **4** in der

Schaltung **10** kategorisiert, und eine Veränderung des binären (Ausgeprägtheits-)Ausgangssignals **11** in einer Richtung wird verwendet, um eine photographische oder elektronische Kamera **1** zu betätigen, um sie zu veranlassen, eine einzelne Aufnahme zu machen oder eine Bildsequenz (ein Video oder Filmclip) zu beginnen. In dem letzteren Fall geht der Clip weiter, bis das Signal **11** den Zustand in die andere Richtung wechselt. Auch wird in dem Fall einer elektronischen Kamera **1** das Bildsignal **2** bei **12** aufgezeichnet oder angezeigt.

**[0053]** Bei [Fig. 3](#) wird das Ausgangssignal **5** eines an der Kehle angebrachten Schalleitmikrophons **4** in der Schaltung **10** kategorisiert, um ein Ausgeprägtheitssignal **13** zur Steuerung eines Gatters **7** zu erzeugen, das das Bildsignal **2** selektiv an ein Aufzeichnungsgerät oder eine Anzeige **12** überträgt. Das Signal **13** kann ein binäres Signal sein und kann direkt zu dem Gatter **7** gekoppelt sein, gemäß der Darstellung ist es jedoch mit einem oder mehreren anderen Signalen **16** in einer Schaltung **14** kombiniert, die ein binäres Signal **15** zum Steuern des Gatters **7** erzeugt.

**[0054]** In einer ersten Form dieser Schaltung sind das Signal **13** und das bzw. die Signal(e) **16** binäre Signale und sind in der Schaltung **14** logisch kombiniert. Beispielsweise kann das Signal **13** mit einem Signal **16** UND-verknüpft sein, das angibt, dass die Aufnahmenzusammensetzung (die durch eine räumliche Segmentierung des Bildsignals **2** erhalten wird) akzeptabel ist, und das Ergebnis kann mit einem Signal **16** ODER-verknüpft sein, das die Kamera anweist, in jedem Fall eine Aufnahme zu machen. Das letztgenannte Signal könnte in der Tat ein Signal von der Schaltung **10** auf eine Erfassung eines Befehls wie z.B. „Nimm das (auf)“ („Take that“) hin sein, dies ist jedoch nicht unbedingt der Fall.

**[0055]** Bei einer zweiten Form dieser Schaltung weist die Schaltung **10** dem Audiosignal **5** gemäß der Klangart, die erfasst wurde, Werte (das Ausgeprägtheitssignal) zu, und diese Werte werden in einer Algorithmusschaltung **14** gemäß einem vorbestimmten Algorithmus mit numerischen Signalen **16** von anderen Quellen kombiniert, wobei das Ausgangssignal einer Schwellwertbildung unterzogen wird, um das binäre Signal **15** zu liefern. In diesem Fall wird einem aufhebenden Befehl wie z.B. „Nimm das (auf)“ („Take that“) ein extremer Wert zugewiesen, um den Wert des Signals **15** ungeachtet der Werte der Signale **16** zu ermitteln.

**[0056]** Bei [Fig. 4](#) wird das Signal **5** von dem am Ohr angebrachten Schalleitmikrophon **4** über eine drahtlose Kopplung **18**, **19**, z.B. über eine Hochfrequenz, Infrarot oder eine Induktion, an ein Mehrkanal-Aufzeichnungsgerät **3** übertragen, das auch das Bildsignalausgangssignal **2** einer feststehenden Kamera **17**

und (ein) Steuersignal(e) **16** von zumindest einer weiteren Quelle empfängt. Anschließend werden die Signale als Bildsignal **6** und Steuersignale **20** abgespielt, wobei die Letzteren durch eine Schaltung **14** verarbeitet werden und dazu verwendet werden, das Bildsignalgatter **7** wie bei [Fig. 3](#) zu steuern.

**[0057]** Bei einer Variante der [Fig. 4](#) ist die Kamera an einem Waggon eines Rummelplatzgefährts befestigt und beobachtet den Träger des Mikrophons während der Fahrt. Eines der Signale **16** ist ein Freigabesignal, das für einen Zeitraum vorgesehen ist, der den Zeitraum, während dessen die Fahrt im Gange ist, und einen kurzen Zeitraum davor und danach umfasst.

**[0058]** [Fig. 5](#) zeigt in schematischer Form eine Audioanalyseanordnung zum Bereitstellen eines Ausgeprägtheitssignals **27**. Das Audiosignal **5** wird durch eine Signalverlaufsform-Analyseschaltung **21** und eine Hüllkurvenamplitude-Analyseschaltung **25** empfangen.

**[0059]** Die Schaltung **21** ermittelt die Art des Signalverlaufs und erzeugt einen Satz von vier Signalen **29** bis **32** je nachdem, ob der Typ als stimmliche Sprache, stimmliche Nicht-Sprache, als nicht-stimmlich bzw. unbestimmt ermittelt wird. Jedes der Signale **29** bis **31** wird verwendet, um eine jeweilige Schaltung **22** bis **24** zu befähigen, den Grad der Übereinstimmung der Signalverlaufsform mit jedem eines jeweiligen Satzes von gespeicherten Signalverläufen zu bestimmen. Die im Betrieb befindliche der Schaltungen **22** bis **24** geht dann dazu über, ein Markierungssignal **33** bis **35** zu erzeugen, das den gespeicherten Signalverlauf, dem der ankommende Signalverlauf am ehesten entspricht, eindeutig zu identifizieren, vorausgesetzt, dass eine ausreichende Entsprechung vorliegt, und liefert andernfalls ein Markierungssignal, das keine Entsprechung bedeutet, was dem Signal **32** von der Schaltung **21** gleichkommt.

**[0060]** Die Signale **33** bis **35** werden mit einem digitalisierten Volumen/Amplitude-Signal **36** von der Schaltung **25** kombiniert und verwendet, um eine zweidimensionale Nachschlagtabelle zu adressieren. Für ein binäres Ausgangssignal **27** aus der Tabelle ist der gesamte Tabelleninhalt entweder Null oder Eins, für ein numerisches Ausgeprägtheitsausgangssignal kann die Tabelle jedoch eine Bandbreite von Zahlen enthalten.

**[0061]** Wenn bei Nichtvorliegen jeglicher Angabe des Gesamtsignalverlaufs die Bestimmung der Hüllkurvenamplitude ein Problem ist, werden die Schaltungen **21** und **25** gekoppelt, um die Bestimmung einer Hüllkurvenamplitude zu ermöglichen, nachdem die Signalverlaufsform ermittelt wurde, wie durch die gestrichelte Linie **28** gezeigt ist. Alternativ oder zusätzlich dazu können die Ausgangssignale **33** bis **35**

zu der Schaltung **25** gekoppelt werden.

**[0062]** Zusätzlich dazu können dort, wo die Prosodie wichtig ist, das Ausgangssignal **33** der Schaltung **22** zusammen mit dem Signal **5** mit einer Prosodiebestimmungsschaltung gekoppelt sein, um ein alternatives numerisches Ausgangssignal zum Eingeben in die Tabelle **26** zu liefern, wobei das numerische Ausgangssignal in diesem Fall das Wort und die zugeordnete Prosodie angibt.

### Patentansprüche

1. Automatische Bilderzeugungsvorrichtung, die eine Kamera, eine Schallerfassungseinrichtung, die in der Lage ist, Klänge von einer vorbestimmten Person vorzugsweise zu empfangen oder zu identifizieren und ansprechend darauf ein Audiosignal zu erzeugen, eine Audioeinrichtung zum Bereitstellen eines Ausgeprägtheitssignals ansprechend auf das Audiosignal sowie eine Steuereinrichtung, die auf das Ausgeprägtheitssignal anspricht, um die Kamera zu steuern, umfasst, wobei die Audioeinrichtung dazu angeordnet ist, das Ausgeprägtheitssignal bereitzustellen, indem sie auf zumindest eines der folgenden anspricht: (a) die Prosodie der Sprache, (b) nicht-sprachliche stimmliche Klänge und (c) nicht-stimmliche Klänge.

2. Vorrichtung gemäß Anspruch 1, bei der die Kamera zur Bildaufzeichnung auf ein lichtempfindliches Medium angeordnet ist.

3. Vorrichtung gemäß Anspruch 1, bei der die Kamera eine elektronische Kamera ist, die ein Bildsignal liefert.

4. Automatische Bilderzeugungsvorrichtung, die eine elektronische Kamera, die ein Bildsignal liefert, eine Schallerfassungseinrichtung, die in der Lage ist, Klänge von einer vorbestimmten Person vorzugsweise zu empfangen oder zu identifizieren und ansprechend darauf ein Audiosignal zu erzeugen, eine Audioeinrichtung zum Bereitstellen eines Ausgeprägtheitssignals ansprechend auf das Audiosignal sowie eine Steuereinrichtung, die auf das Ausgeprägtheitssignal anspricht, um das Bildsignal zu steuern, umfasst, wobei die Audioeinrichtung dazu angeordnet ist, das Ausgeprägtheitssignal bereitzustellen, indem sie auf zumindest eines der folgenden anspricht: (a) die Prosodie der Sprache, (b) nicht-sprachliche stimmliche Klänge und (c) nicht-stimmliche Klänge.

5. Vorrichtung gemäß Anspruch 4, die ferner eine Aufzeichnungseinrichtung zum Aufzeichnen des Audiosignals zusammen mit dem Bildsignal sowie eine Abspielereinrichtung zum Abspielen des Audio- und des Bildsignals umfasst, wobei die Audioeinrichtung angeordnet ist, um ansprechend auf das abgespielte Audiosignal das Ausgeprägtheitssignal zu liefern,

und wobei die Steuereinrichtung auf das Ausgeprägtheitssignal anspricht, um das abgespielte Bildsignal zu steuern.

6. Vorrichtung gemäß Anspruch 4 oder 5, bei der die Steuereinrichtung ein Gatter zum selektiven Übertragen des Bildsignals umfasst.

7. Vorrichtung gemäß Anspruch 6, bei der die Steuereinrichtung eine Kombinationseinrichtung zum Kombinieren des Ausgeprägtheitssignals mit anderen Steuersignalen, um ein Gattersteuerausgangssignal zum Steuern des Gatters zu erzeugen, umfasst.

8. Vorrichtung gemäß Anspruch 6, bei der das Gatter angeordnet ist, um direkt durch ein binäres Ausgeprägtheitssignal gesteuert zu werden.

9. Vorrichtung gemäß einem der Ansprüche 3 bis 8, die eine Aufzeichnungseinrichtung zum Aufzeichnen des gesteuerten Bildsignals umfasst.

10. Vorrichtung gemäß einem der vorhergehenden Ansprüche, bei der die Audioanalyseeinrichtung ferner angeordnet ist, um ein Ausgeprägtheitssignal mittels Worterkennung zu liefern.

11. Vorrichtung gemäß einem der vorhergehenden Ansprüche, bei der die Audioeinrichtung eine Audioanalyseeinrichtung zum Analysieren des Signalverlaufs des Audiosignals umfasst, um es als in eine eines Satzes von vorbestimmten Kategorien fallend zu kategorisieren.

12. Vorrichtung gemäß Anspruch 11, bei der die Audioanalyse eine Serie von Referenzsignalverläufen enthält, bezüglich derer die Kategorisierung durchzuführen ist.

13. Vorrichtung gemäß einem der vorhergehenden Ansprüche, bei der die Audioanalyse so angeordnet ist, dass sie trainiert werden kann, neue Klänge zu erkennen.

14. Vorrichtung gemäß einem der Ansprüche 11 bis 13, bei der die Audioanalyseeinrichtung angeordnet ist, um die Gesamtsignalamplitude oder -hüllkurve zu berücksichtigen, wenn sie dem Audiosignal eine Kategorie zuweist.

15. Vorrichtung gemäß einem der Ansprüche 11 bis 14, bei der die Audioanalyseeinrichtung angeordnet ist, um ein Ausgeprägtheitssignal zu liefern, das einen Wert aufweist, der zumindest teilweise durch die zugewiesene Kategorie des Audiosignals bestimmt wird.

16. Vorrichtung gemäß Anspruch 15, bei der die Audioanalyseeinrichtung ein binäres Ausgeprägtheitssignal gemäß der zugewiesenen Kategorie des

Audiosignals liefert.

17. Vorrichtung gemäß Anspruch 15, bei der die Audioanalyseeinrichtung ein Ausgeprägtheitssignal liefert, das einen einer Mehrzahl von Werten gemäß der zugewiesenen Kategorie des Audiosignals aufweist.

18. Vorrichtung gemäß Anspruch 15, bei der die Audioanalyseeinrichtung angeordnet ist, um ein Ausgeprägtheitssignal zu liefern, das in Abhängigkeit von der zugewiesenen Kategorie diskret variabel und in Abhängigkeit von der Gesamtsignalamplitude oder -hüllkurve des Audiosignals kontinuierlich variabel ist.

19. Vorrichtung gemäß einem der vorhergehenden Ansprüche, bei der die Schallerfassungseinrichtung ein Schalleitmikrophon aufweist, das dazu ausgelegt ist, im Bereich des Kopfes oder der Kehle der vorbestimmten Person getragen zu werden, um vorzugsweise Klänge zu erfassen, die von dem Kopf des Trägers emittiert werden, und um ansprechend darauf das Audiosignal zu erzeugen.

20. Vorrichtung gemäß Anspruch 19, die ferner ein herkömmliches Mikrophon zum Empfangen von Klängen über die Luft aufweist.

21. Vorrichtung gemäß Anspruch 20, die eine Einrichtung zum Aufzeichnen des Signals aus dem herkömmlichen Mikrophon umfasst.

22. Vorrichtung gemäß Anspruch 20 oder 21, bei der die Audioanalyseeinrichtung angeordnet ist, um das Signal von dem herkömmlichen Mikrophon zu berücksichtigen, wenn sie das Audiosignal als in eine eines Satzes von vorbestimmten Kategorien fallend kategorisiert.

23. Automatische Bilderzeugungsvorrichtung gemäß einem der vorhergehenden Ansprüche, bei der die Kamera mit einer Befestigungseinrichtung zum Befestigen derselben an einer vorbestimmten Person ausgestattet ist.

24. Automatische Bilderzeugungsvorrichtung gemäß einem der Ansprüche 1 bis 22, bei der die Kamera mit einer Befestigungseinrichtung zum Befestigen zumindest der Bilderzeugungsoptik derselben an dem Kopf einer vorbestimmten Person ausgestattet ist.

25. Vorrichtung gemäß einem der Ansprüche 21 bis 22, bei der die Kamera an einer anderen Kameraposition als an der vorbestimmten Person angebracht ist.

26. Vorrichtung gemäß Anspruch 1 oder Anspruch 21, bei der die Aufzeichnungseinrichtung mit

einer Befestigungseinrichtung zum Befestigen derselben an einer vorbestimmten Person ausgestattet ist.

27. Vorrichtung gemäß Anspruch 1 oder 21, bei der die Aufzeichnungseinrichtung an einer von der Schallerfassungseinrichtung entfernten Aufzeichnungsgerätposition angeordnet ist, und die eine Einrichtung zum drahtlosen Übertragen des Audiosignals von der Schallerfassungseinrichtung an die Aufzeichnungsgerätposition umfasst.

28. Vorrichtung gemäß einem der vorhergehenden Ansprüche, bei der die Audioeinrichtung an einer von der Schallerfassungseinrichtung entfernten Audioeinrichtungsposition angebracht ist, und die eine Einrichtung zum drahtlosen Übertragen des Audiosignals an die Audioeinrichtungsposition umfasst.

29. Vorrichtung gemäß einem der vorhergehenden Ansprüche, bei der die Steuereinrichtung an einer von der Audioeinrichtung entfernten Steuerungsposition angeordnet ist, und die eine Einrichtung zum drahtlosen Übertragen des Ausgeprägtheitssignals von der Audioeinrichtung an die Steuerungsposition umfasst.

30. Ein Verfahren zum Bilderzeugen, das die Schritte eines Betätigens einer Kamera, um ein Bildsignal zu erzeugen, eines vorzugsweise Empfangens oder Identifizierens von Klängen von einer vorbestimmten Person, eines Erzeugens eines Audiosignals ansprechend darauf, und eines Erzeugens eines Ausgeprägtheitssignals ansprechend auf das Audiosignal, und eines Steuerns der Kamera ansprechend auf das Ausgeprägtheitssignal umfasst, wobei das Ausgeprägtheitssignal durch zumindest eines der folgenden bestimmt wird: (a) die Prosodie der Sprache, (b) nicht-sprachliche stimmliche Klänge und (c) nicht-stimmliche Klänge.

31. Ein Verfahren gemäß Anspruch 30, bei dem die Kamera eine elektronische Kamera ist, die ein Bildsignal liefert.

32. Ein Verfahren zum Bilderzeugen, das die Schritte eines Betätigens einer elektronischen Kamera, um ein Bildsignal zu erzeugen, eines vorzugsweise Empfangens oder Identifizierens von Klängen von einer vorbestimmten Person, eines Erzeugens eines Audiosignals ansprechend darauf, und eines Erzeugens eines Ausgeprägtheitssignals ansprechend auf das Audiosignal, und eines Steuerns des Bildsignals ansprechend auf das Ausgeprägtheitssignal umfasst, wobei das Ausgeprägtheitssignal durch zumindest eines der folgenden bestimmt wird: (a) die Prosodie der Sprache, (b) nicht-sprachliche stimmliche Klänge und (c) nicht-stimmliche Klänge.

33. Ein Verfahren gemäß einem der Ansprüche

30 bis 32, bei dem der Schritt des vorzugsweise Empfangens oder Identifizierens unter Verwendung eines Schalleitmikrophons durchgeführt wird, das im Bereich des Kopfes oder der Kehle der vorbestimmten Person angebracht ist, um Klänge, die von dem Kopf der Person emittiert werden, vorzugsweise zu erfassen.

34. Ein Verfahren gemäß einem der Ansprüche 30 bis 33, bei dem der Schritt des Erzeugens des Ausgeprägtheitssignals ein Analysieren des Signalverlaufs des Audiosignals umfasst, um es als in eine eines Satzes von vorbestimmten Kategorien fallend zu kategorisieren.

35. Ein Verfahren gemäß einem der Ansprüche 30 bis 34, das den Schritt des Aufzeichnens und des Abspielens des Audio- und des Bildsignals vor dem Erzeugen des Ausgeprägtheitssignals umfasst.

Es folgen 2 Blatt Zeichnungen

Anhängende Zeichnungen

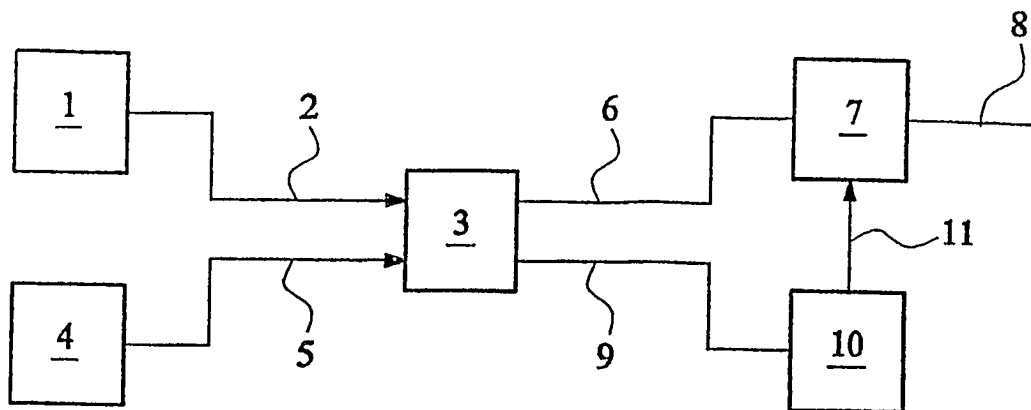


FIG. 1

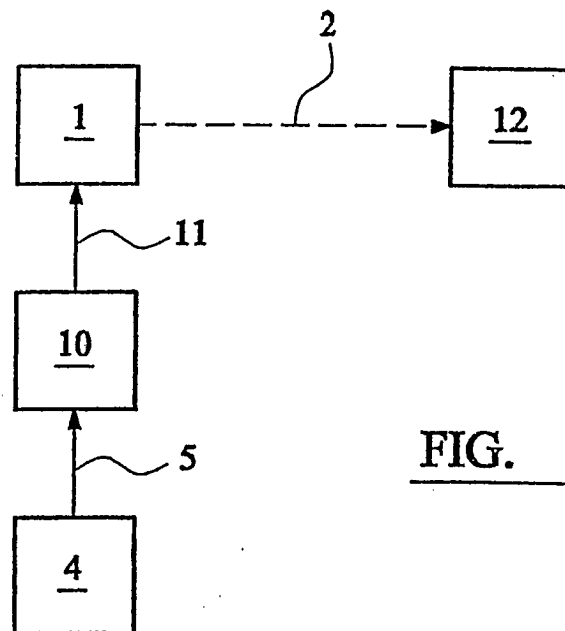


FIG. 2

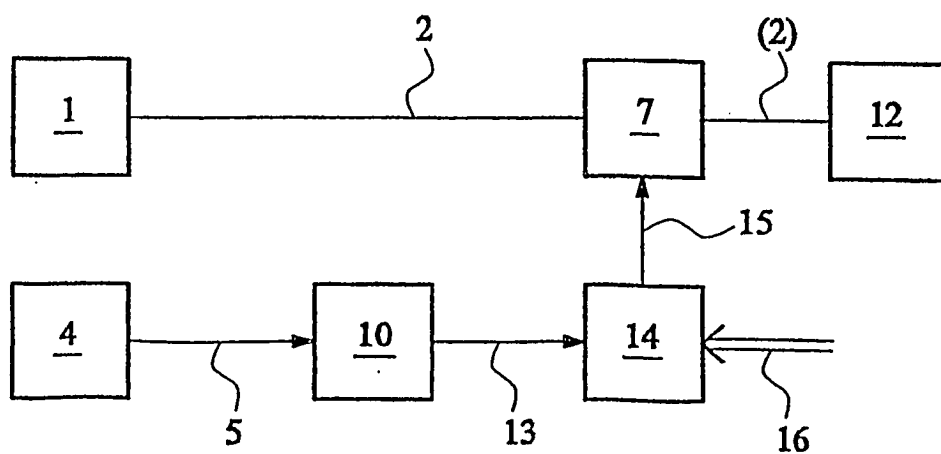


FIG. 3

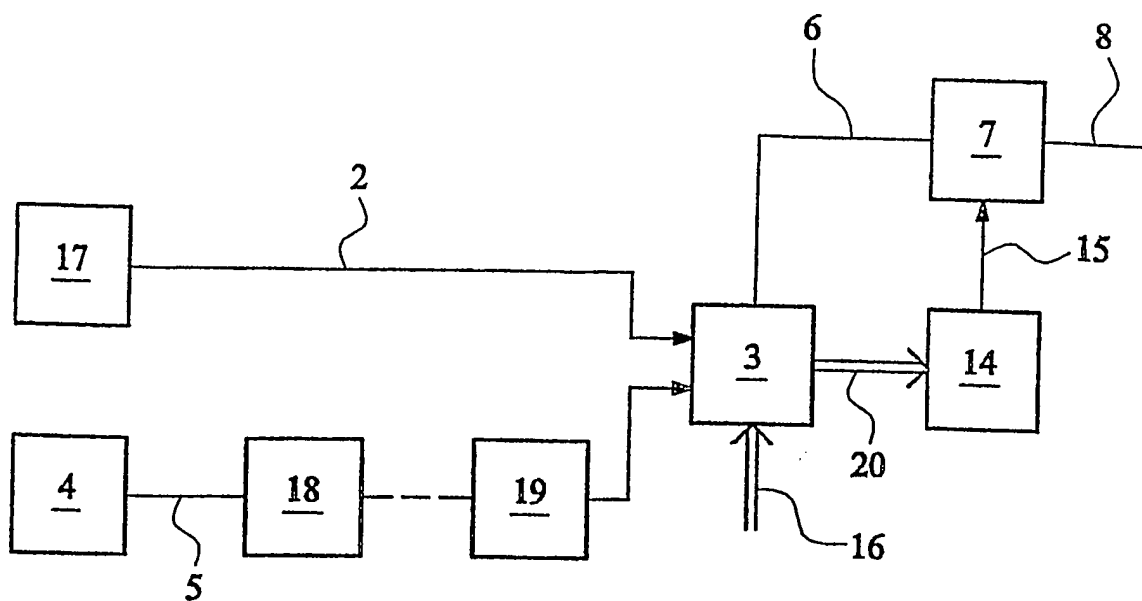


FIG. 4

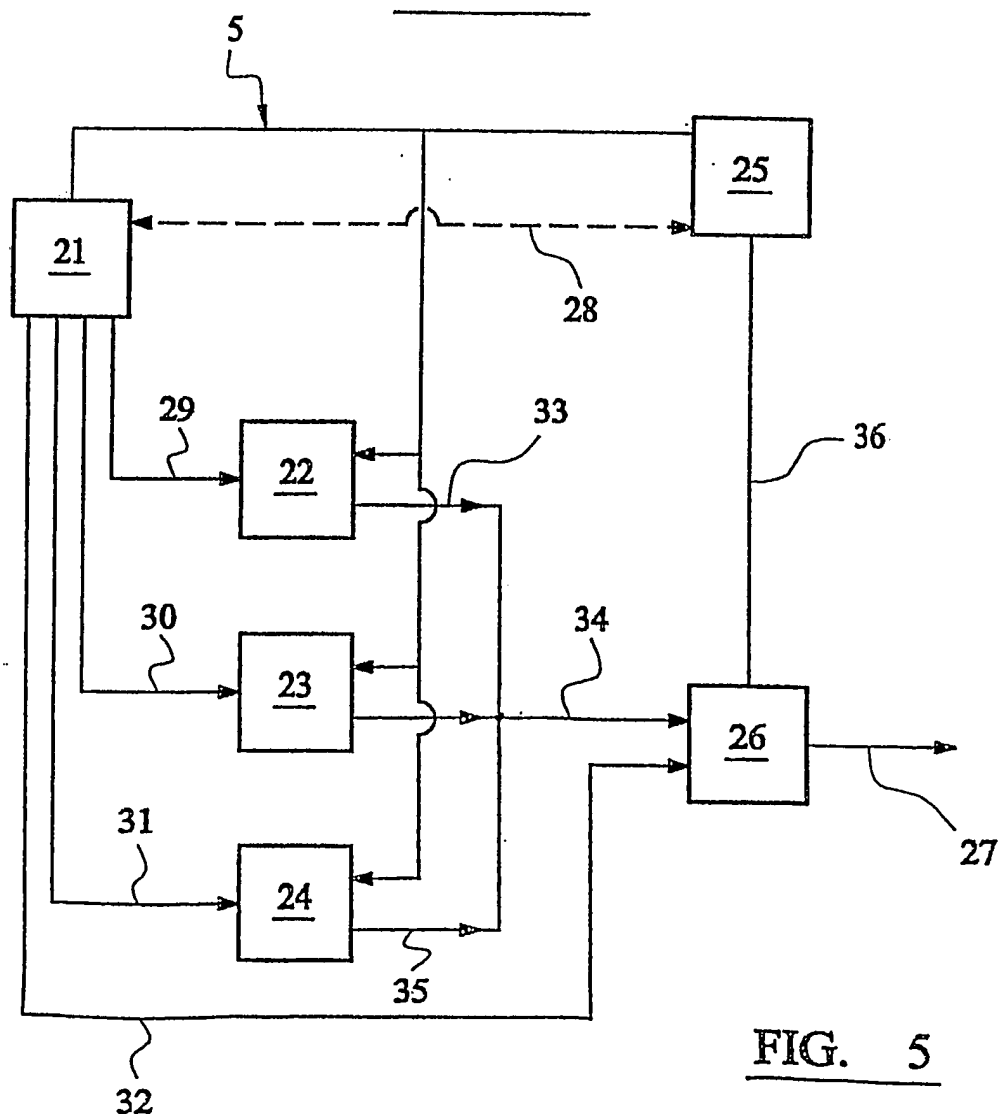


FIG. 5