

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
22. Mai 2003 (22.05.2003)

PCT

(10) Internationale Veröffentlichungsnummer
WO 03/042802 A2

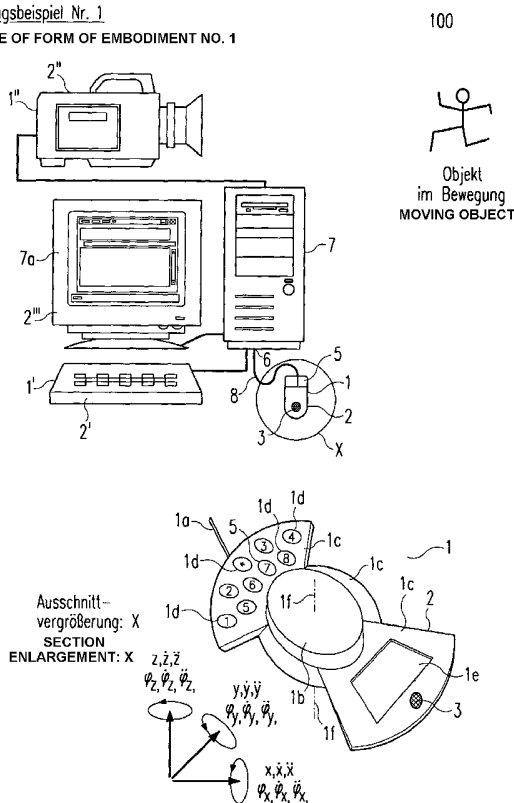
- (51) Internationale Patentklassifikation⁷: G06F 3/00
- (21) Internationales Aktenzeichen: PCT/EP02/12918
- (22) Internationales Anmeldedatum:
18. November 2002 (18.11.2002)
- (25) Einreichungssprache: Deutsch
- (26) Veröffentlichungssprache: Deutsch
- (30) Angaben zur Priorität:
101 56 457.0 16. November 2001 (16.11.2001) DE
- (71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): 3DCONNEXION GMBH [DE/DE]; An der Hartmühle 8, 82229 Seefeld (DE).
- (72) Erfinder; und
- (75) Erfinder/Anmelder (nur für US): GOMBERT, Bernd [DE/DE]; 3Dconnexion GmbH, An der Hartmühle 8, 82229 Seefeld (DE).
- (74) Anwalt: RUPP, Christian; Mitscherlich & Partner, Postfach 33 06 09, 80066 München (DE).
- (81) Bestimmungsstaaten (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

[Fortsetzung auf der nächsten Seite]

(54) Title: INPUT DEVICE, WEBCAM AND SCREEN HAVING A VOICE INPUT FUNCTION

(54) Bezeichnung: EINGABEGERÄT, WEBCAM UND BILDSCHIRM MIT SPRACHEINGABEFUNKTION

Ausführungsbeispiel Nr. 1
EXAMPLE OF FORM OF EMBODIMENT NO. 1



(57) Abstract: The invention relates to a manual input device (1, 1') for producing control signals for real and/or virtual objects. According to the invention, a microphone (3) is integrated into the housing (2, 2') of the manual input device, for recording and converting acoustic voice control signals of a user, thus relieving the user of the handling of the input device (1, 1'). The interface (8) between said input device (1, 1') and a computer (7) can also be used for the transmission of acoustic voice control signals or the evaluation of the same, reducing the number of occupied connections (6) of the computer (7). Furthermore, a processing unit (4), such as an ASIC, can be provided in the housing of the manual input device (1, 1'), said processing unit subjecting the output signals of the microphone (3) to an algorithm for automatic speaker identification, and voice recognition, analysis and/or interpretation. In this case, not only acoustically or electrically converted signals are transmitted from the manual input device (1, 1') to the calculation unit (7), but rather instructions which can be directly processed by a computer (7). The inventive manual input device (1, 1') can be advantageously operated with driver software which automatically activates a voice control function (recording and conversion of acoustic voice control signals), as soon as the manual input device (1, 1') is connected to a connection (6) provided therefore of the calculation unit (7). The voice control activation function can be directly integrated into the driver software which is used to convert the control signals for the cursor or object control system. According to the invention, the converted and processed voice control signals, together with the remaining control instructions, can be transmitted to the calculation unit (7) via a common wire-bound or wireless interface (8).

[Fortsetzung auf der nächsten Seite]



WO 03/042802 A2



(84) **Bestimmungsstaaten** (*regional*): ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

— *ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts*

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

(57) **Zusammenfassung:** Bei der vorliegenden Erfindung handelt es sich um ein manuelles Eingabegerät (1, 1') zur Erzeugung von Ansteuerungssignalen für reale und/oder virtuelle Objekte. Erfindungsgemäss ist in das Gehäuse (2, 2') des manuellen Eingabegerätes ein Mikrofon (3) zur Aufnahme und Umsetzung von akustischen Sprachsteuersignalen eines Benutzers integriert. Auf diese Weise wird der Benutzer von der Handhabung eines Eingabegerätes (1, 1') entlastet. Die Schnittstelle (8) des Eingabegerätes (1, 1') zu einem Computer (7) kann dabei auch für die Übertragung von akustischen Sprachsteuersignalen bzw. deren Auswertung verwendet werden. Auf diese Weise wird die Anzahl der belegten Anschlüsse (6) des Computers (7) verringert. In dem Gehäuse des manuellen Eingabegerätes (1, 1') kann weiterhin eine Verarbeitungseinheit (4), wie beispielsweise ein ASIC, vorgesehen sein, der die Ausgangssignale des Mikrofons (3) einem Algorithmus zur automatischen Sprecheridentifikation, Spracherkennung, -analyse und/oder -interpretation unterzieht. In diesem Fall werden also von dem manuellen Eingabegerät (1, 1') nicht nur akustisch bzw. elektrisch umgesetzte Signale zu der Recheneinheit (7) hin übertragen, sondern vielmehr direkt von einem Computer (7) verarbeitbare Befehle vermittelt. Das manuelle Eingabegerät (1, 1') kann vorteilhaft zusammen mit einer Treibersoftware betrieben werden, welche eine Sprachsteuerungsfunktion (Aufnahme und Umsetzung von akustischen Sprachsteuersignalen) automatisch aktiviert, sobald das manuelle Eingabegerät (1, 1') an einem dazu vorgesehenen Anschluss (6) der Recheneinheit (7) angeschlossen wird. Die Sprachsteuerungs-Aktivierungsfunktion kann dabei direkt in die Treibersoftware integriert werden, welche zur Umsetzung der Ansteuerungssignale für die Cursor- oder Objektsteuerung vorgesehen ist. Erfindungsgemäss können die umgesetzten und verarbeiteten Sprachsteuersignale zusammen mit den übrigen Ansteuerbefehlen über eine gemeinsame drahtgebundene oder drahtlose Schnittstelle (8) zu der Recheneinheit (7) übertragen werden.

Eingabegerät, Webcam und Bildschirm mit Spracheingabefunktion

Die vorliegende Erfindung bezieht sich auf ein manuelles Eingabegerät zur Erzeugung von Ansteuersignalen für reale und/oder virtuelle Objekte, einen Bildschirm sowie auf einen Internet-Kameraaufsatz für eine Recheneinrichtung.

KURZBESCHREIBUNG DES STANDES DER TECHNIK

Manuelle Eingabegeräte zur Eingabe von Informationen, wie etwa Tastaturen, Mäuse, Trackballs und Joysticks sind heute weit verbreitet. Sie werden verwendet, um Positionsmarken (engl.: „Cursors“), Mauszeiger etc. zu steuern, um z.B. durch eine virtuelle Szene navigieren zu können oder virtuelle Objekte, die auf einem Bildschirm grafisch dargestellt werden, zu bewegen. Während Tastaturen und Mäuse eine feste Ablagefläche benötigen, um in effizienter Weise bedient werden zu können, ist es bei Verwendung eines Touchscreens bzw. eines Touch-Panels möglich, direkt auf Objekte, die auf dem Bildschirm abgebildet sind, mit dem Finger zu zeigen, ohne weitere platzraubende Zusatzgeräte auf dem Schreibtisch zu benötigen.

Nach dem Stand der Technik sind heute verschiedene Lösungen für das Problem der Echtzeit-Bewegungssteuerung virtueller Objekte verfügbar, wobei jede dieser Lösungen für einen speziellen Anwendungszweck optimiert ist. Um einige der wichtigsten dieser Lösungen, welche für das Verständnis der vorliegenden Erfindung von Relevanz sind, ansatzweise erklären zu können, ist es notwendig, kurz auf ihre wichtigsten technischen Merkmale einzugehen.

Eine Möglichkeit zur Echtzeit-Bewegungssteuerung virtueller Objekte hat sich in letzter Zeit durch die Entwicklung von Eingabegeräten für Computer ergeben, die die gleichzeitige Eingabe von Ansteuersignalen mehrerer, voneinander unabhängiger Freiheitsgrade ermöglichen. Die dadurch geschaffenen Möglichkeiten übersteigen bei weitem diejenigen,

die beispielsweise bei der Verwendung einer Maus bestehen, die lediglich zweidimensional (z.B. auf der Ablagefläche eines Schreibtischs) gesteuert werden kann. Zwar ist es auch bekannt, eine Maus beispielsweise mit zusätzlichen Schaltern zu versehen, indessen haben diese Schalter den Nachteil, dass sie nicht die Eingabe von Analogdaten ermöglichen, sondern vielmehr auf Binärdaten (Ein/Aus) beschränkt sind.

Aus dem Stand der Technik sind auch verschiedene Eingabegeräte bekannt, die analoge Ansteuersignale mit verschiedenen, voneinander unabhängigen Freiheitsgraden erzeugen können, wobei jedes dieser Analogsignale somit als Parameterwert bei einer Steuerung virtueller Objekte verwendet werden kann. Derartige manuell steuerbare Eingabesysteme, die eine Navigation in drei Dimensionen erlauben, werden heute in einer Reihe der unterschiedlichsten technischen Anwendungsfelder erfolgreich eingesetzt.

Beispielsweise ist aus der Patentschrift US-A-5,757,360 ein eiförmiges Eingabegerät für Computer bekannt, das durch eine Hand des Benutzers frei im Raum bewegt werden kann, seine momentanen Positionen, Bewegungsrichtungen, Geschwindigkeiten und Beschleunigungen ermittelt und diese kinematischen Daten drahtlos zu einem Computer überträgt. Dabei wird ein analoger Bewegungsablauf in Form eines Bewegungsmusters identifiziert, woraus Bewegungsbefehle abgeleitet und in eine animierte Grafikdarstellung umgesetzt werden. Die Bewegungsmuster werden dann mit Hilfe eines Mustererkennungsalgorithmus automatisch erkannt. Überdies werden Steuerbefehle erzeugt. Der Nachteil dieses Verfahrens besteht darin, dass es nicht frei spezifizierbar ist, da Bewegungsabläufe des Benutzers, die durch das Eingabegerät analog erfasst werden, korrespondierenden Bewegungsabläufen von gespeicherten Bewegungssequenzen einer animierten Grafikdarstellung zugeordnet werden und nur als solche dargestellt werden können.

Eingabegeräte, die über manuell zu betätigende Kraft-/Momenten-Sensoren verfügen, sind beispielsweise aus den Patentschriften DE 36 11 336 C2, DE 37 64 287 sowie EP 0 979 990 A2 bekannt.

Aus der zuletzt genannten europäischen Patentschrift EP 0 979 990 A2 ist bekannt, einen derartigen Kraft-/Momenten-Sensor zum Steuern von Bedienelementen eines realen oder

virtuellen Misch- bzw. Steuerpults zu verwenden, beispielsweise um neuartige Farb-, Licht- und/oder Tonkompositionen zu kreieren und zu gestalten. Hierbei kann in vorteilhafter Weise die intuitive räumliche Steuerung in drei translatorischen sowie drei rotatorischen Freiheitsgraden auf ein stufenloses räumliches Mischen oder Steuern einer großen Anzahl von optischen und/oder akustischen Parametern übertragen werden. Zur Steuerung wird auf die Bedienoberfläche des Eingabegeräts ein Druck ausgeübt und dadurch ein Impuls erzeugt, der mit Hilfe des Kraft-/Momenten-Sensors erfasst und in ein aus einem Kraft- und einem Momentenvektor bestehendes Vektorpaar umgesetzt wird. Werden dabei bestimmte charakteristische Impulsvorgaben erfüllt, kann beispielsweise eine objektspezifische Steueroperation und/oder eine technische Funktion durch Schalten in einen Aktivierungszustand ausgelöst bzw. durch Schalten in einen Deaktivierungszustand wieder beendet werden.

Aus dieser Druckschrift ist weiterhin bekannt, den besagten Kraft-/Momenten-Sensor als Bedienelement eines 3D-Eingabegeräts seitlich an einem Touchscreen anzubringen, so dass die Längsachse des Bedienelements parallel zu der Anzeige- und Bedienfläche des Touchscreens liegt. Damit ist jedoch der Nachteil verbunden, dass die Blickrichtung auf den Touchscreen nicht mit der Längsachse des Bedienelements zusammenfällt. Das hat zur Folge, dass die Hand-Augen-Koordination des Benutzers erschwert wird, da die Richtungen von Steuerbewegungen des Bedienelements und angezeigten Objektbewegungen auf dem Bildschirm bei ungünstiger Achsenbelegung des Bedienelements nicht übereinstimmen.

Bei der Verwendung von dreidimensionalen Eingabegeräten, wie beispielsweise Kraft-/Momentensensoren, im Bereich von CAD-Anwendungen (engl.: „Computer-Aided Design“) ist die Arbeitsweise weit verbreitet, dass der Benutzer in der einen Hand das dreidimensionale Eingabegerät und mit der anderen Hand eine normale Computermaus (zweidimensionale Computermaus) bedient. Gleichzeitig ist es für die Ausführung von CAD-Tätigkeiten typisch, dass an bestimmten Stellen des bearbeiteten Objekts numerische Eingaben (Maßangaben etc.) getätigt werden müssen.

Bei der oben angesprochenen zweihändigen Arbeitsweise im CAD-Bereich tritt dann also das Problem auf, dass der Benutzer die Hand von einem der genannten Eingabegeräte lösen

muss, um beispielsweise mittels einer Tastatur die Eingabe numerischer Information oder von (Shortcut-)Befehlen auszuführen. Danach muss die Hand von der Tastatur wiederum zu dem Eingabegerät zurückgeführt werden. Es ist offensichtlich, dass dieser Wechsel der Handposition Störungen des Arbeitsablaufs mit sich bringt, insbesondere wenn man be-
5 berücksichtigt, dass der Blick des Benutzers nach Möglichkeit nicht von der Bildschirmoberfläche abgelenkt werden soll.

Eine denkbare Lösung wäre die Eingabe der numerischen Daten mittels Sprachsteuersignalen (engl.: „Voice Control“). Hierzu muss der Benutzer gemäß dem Stand der Technik
10 auf seinem Kopf einen Aufsatz tragen, der ein Mikrofon trägt. Dieser Aufsatz muss von dem Benutzer getragen werden (engl.: „Headset“). Auch diese Lösung weist ergonomische Probleme auf, da einerseits nunmehr der Benutzer insgesamt vier Eingabegeräte, nämlich den Kraft-/Momentensensor, die normale Computermaus, die Tastatur und das
15 Mikrofon handhaben muss. Darüber hinaus kann auch das Problem auftreten, dass die Anzahl der Anschlüsse beispielsweise auf der Rückseite des Computers, an dem die genannten vier Geräte angeschlossen werden sollen, nicht ausreichend sind, oder aber Doppelbelegungen vorkommen, falls wenigstens zwei der genannten Geräte den selben Anschlussstyp (z.B. USB, RS 232 C) benötigen.

20 AUFGABE DER VORLIEGENDEN ERFINDUNG

Es ist daher Aufgabe der vorliegenden Erfindung, die obigen Probleme durch entsprechende konstruktive Änderungen seitens eines manuellen Eingabegerätes zu lösen.

25 Diese Aufgabe wird erfindungsgemäß durch die Merkmale der unabhängigen Ansprüche gelöst. Die abhängigen Ansprüche bilden den zentralen Gedanken der Erfindung in besonders vorteilhafter Weise weiter.

ZUSAMMENFASSENDER DARSTELLUNG DER VORLIEGENDEN ERFINDUNG

30

Erfindungsgemäß ist ein manuelles Eingabegerät zur Erzeugung von Ansteuerungssignalen für reale und/oder virtuelle Objekte vorgesehen. Das manuelle Eingabegerät kann somit nicht nur zur Ansteuerung von Computern, sondern auch zur Ansteuerung von Robo-

tern oder anderen realen Objekten verwendet werden. Erfindungsgemäß ist in das Gehäuse des manuellen Eingabegerätes ein Mikrofon zur Aufnahme und Umsetzung von akustischen Sprachsteuersignalen eines Benutzers in elektrische Signale integriert. Durch diese vorteilhafte Ausführungsform werden mittels einer einzigen Maßnahme zwei Vorteile erzielt: Einerseits wird der Benutzer durch die Integration des Mikrofons in das Gehäuse des manuellen Eingabegerätes von der Handhabung eines Eingabegerätes entlastet. Andererseits kann die Schnittstelle des Eingabegerätes auch für die Übertragung von akustischen Sprachsteuersignalen bzw. deren Auswertung verwendet werden. Somit wird die Anzahl der belegten Schnittstellen des Computers verringert. Erfindungsgemäß wird somit also durch eine einzige einfache Maßnahme die Vermeidung zweier Nachteile aus dem Stand der Technik erreicht.

In dem Gehäuse des manuellen Eingabegerätes kann weiterhin eine Verarbeitungseinheit, z.B. ein anwendungsspezifischer integrierter Halbleiterbaustein (ASIC), vorgesehen sein, der die Ausgangssignale von dem Mikrofon einem Algorithmus zur automatischen Sprecheridentifikation, Spracherkennung, -analyse und/oder -interpretation unterzieht. Wird dabei ein von der Stimme des jeweiligen Sprechers unabhängiger Spracherkennungsalgorithmus realisiert, kann die Implementierung der automatischen Sprecheridentifikation entfallen. Von dem manuellen Eingabegerät werden also nicht nur akustisch bzw. elektrisch umgesetzte Signale zu der Recheneinheit hin übertragen, sondern vielmehr direkt von einem Computer verarbeitbare Befehle vermittelt. Unter „Sprachsteuerung“ ist im Übrigen auch die Eingabe von (beispielsweise numerischen) Werten oder Texten zu verstehen. Sprachsteuerung ist also auch im Sinne von „Spracheingabe“ (Eingabe von Informationen und Steuerbefehlen) zu verstehen.

Das manuelle Eingabegerät kann beispielsweise eine herkömmliche Computermaus, ein Kraft-/Momentensensor oder eine Tastatur sein.

Die Sprachsteuerungsfunktion, d.h. die Aktivierung der Aufnahme und Umsetzung von akustischen Sprachsteuersignalen durch die Kombination aus Eingabegerät und Mikrofon kann dabei durch Betätigung des manuellen Eingabegerätes (z.B. mittels einer speziellen Taste) aktivierbar sein.

Das manuelle Eingabegerät kann zusammen mit einer Treibersoftware betrieben werden, welche die Sprachsteuerungsfunktion (Aufnahme und Umsetzung von akustischen Sprachsteuersignalen) automatisch aktiviert, sobald das manuelle Eingabegerät an einem dafür vorgesehenen Anschluss der Recheneinheit angeschlossen wird. Dies stellt eine besonders vorteilhafte Lösung dar, weil zusammen mit dem Anschließen des Eingabegerätes nicht nur die Eingabe von üblichen Ansteuerbefehlen (Cursor- bzw. Mauszeigersteuerung etc.) für die Recheneinheit ermöglicht wird, sondern gleichzeitig auch die Spracherkennungsfunktion bereitgestellt wird.

- 10 Die Sprachsteuerungs-Aktivierungsfunktion kann dabei direkt in die Treibersoftware integriert werden, die zur Umsetzung der übrigen Ansteuersignale (Cursor- oder Objektsteuerung) vorgesehen ist. Dies hat den Vorteil, dass der Benutzer nicht zwei verschiedene Softwarepakete für das manuelle Eingabegerät und die Sprachsteuerung installieren muss, sondern vielmehr durch die Installation einer einzigen Treibersoftware gleichzeitig die Ansteuersignalfunktion wie auch die Sprachsteuerungsfunktion installiert werden kann.

Erfindungsgemäß können die umgesetzten, interpretierten und verarbeiteten Sprachsteuersignale zusammen mit den übrigen Ansteuerbefehlen (beispielsweise für die Cursorsteuerung, etc.) über eine gemeinsame drahtgebundene oder drahtlose Schnittstelle zu einer Recheneinheit übertragen werden. Somit wird erfindungsgemäß nicht eine zusätzliche Schnittstelle für die Übertragung der aufgenommenen und in vom Computer ausführbare Befehle umgesetzten akustischen Sprachsteuersignale benötigt, wodurch, wie bereits oben ausgeführt, die Anzahl der Schnittstellen reduziert wird. Ein solches Eingabegerät ist insbesondere im Bereich CAD, Desktop Publishing oder dergleichen vorteilhaft.

Gemäß einem noch weiteren Aspekt der vorliegenden Erfindung ist ein Internet-Kameraaufsatz für eine Recheneinrichtung vorgesehen. Derartige Kameraaufsätze werden oft auch als Webcams bezeichnet. Erfindungsgemäß ist in das Gehäuse des Internet-Kameraaufsatzes ein Mikrofon integriert. Alternativ kann auch ein Mikrofon fest mit dem Gehäuse der Webcam verbunden sein bzw. an diesem angebracht sein.

Die Erfindung bezieht sich gleichermaßen auf in Bildschirmgehäuse integrierte Mikrofone zur Spracheingabe.

KURZBESCHREIBUNG DER ZEICHNUNGEN

5

Weitere Eigenschaften, Merkmale, Vorteile und Zweckmäßigkeiten der zugrunde liegenden Erfindung ergeben sich aus den untergeordneten abhängigen Ansprüchen sowie aus der folgenden, detaillierten Beschreibung von vier verschiedenen Ausführungsbeispielen zur Erzeugung von Steuersignalen für die Echtzeit-Bewegungssteuerung von realen bzw. 10 virtuellen Objekten, welche in den folgenden Zeichnungen abgebildet sind. Dabei zeigt

Fig. 1 ein erstes Ausführungsbeispiel, bei dem in das Gehäuse einer Computermouse ein Mikrofon zur Aufnahme und Umsetzung von akustischen Sprachsteuersignalen in elektrische Signale, ein ASIC zur automatischen Sprecheridentifikation, Spracherkennung, -analyse und -interpretation sowie eine Taste zur manuellen Aktivierung der Sprachsteuerungsfunktion integriert sind,

Fig. 2 ein zweites Ausführungsbeispiel, bei dem in das Gehäuse einer Computertastatur ein Mikrofon zur Aufnahme und Umsetzung von akustischen Sprachsteuersignalen in elektrische Signale, ein ASIC zur automatischen Sprecheridentifikation, Spracherkennung, -analyse und -interpretation und eine Taste zur manuellen Aktivierung der Sprachsteuerungsfunktion integriert sind,

Fig. 3 ein drittes Ausführungsbeispiel, bei dem in das Gehäuse eines Internet-Kameraaufsatzes (engl.: „Webcam“) ein Mikrofon zur Aufnahme und Umsetzung von akustischen Sprachsteuersignalen in elektrische Signale, ein ASIC zur automatischen Sprecheridentifikation, Spracherkennung, -analyse und -interpretation und eine Taste zur manuellen Aktivierung der Sprachsteuerungsfunktion integriert sind, und

Fig. 4 ein viertes Ausführungsbeispiel, bei dem in das Gehäuse eines Computerbildschirms ein Mikrofon zur Aufnahme und Umsetzung von akustischen Sprachsteuersignalen in elektrische Signale, ein ASIC zur automatischen

Sprecheridentifikation, Spracherkennung, -analyse und -interpretation und eine Taste zur manuellen Aktivierung der Sprachsteuerungsfunktion integriert sind.

DETAILLIERTE BESCHREIBUNG DER ERFINDUNG

Im Folgenden werden die Funktionen der in vier verschiedenen Ausführungsbeispielen
5 der vorliegenden Erfindung enthaltenen Baugruppen, wie in den Figuren 1 bis 4 abgebildet, näher beschrieben. Die Bedeutung der mit Bezugszeichen versehenen Symbole in den Figuren 1 bis 4 kann der beigefügten Bezugszeichenliste entnommen werden.

In Fig. 1 bis 4 ist ein Computer 7 mit einem Bildschirm 7a, einer Maus bzw. einem Kraft/
10 /Momentensensor 1 sowie einer Tastatur 1' als manuellen Eingabegeräten zur Eingabe von Steuerbefehlen und einem Internet-Kameraaufsatz 1'' (engl.: „Webcam“) abgebildet. Dieser Computer 7 stellt nur ein Beispiel für eine Umgebung dar, in der reale und/oder virtuelle Objekte angesteuert werden können. Zum Beispiel kann auf dem Computer 7 ein CAD-Programm installiert sein.

15 Erfindungsgemäß sind die manuellen Eingabegeräte 1, 1' und/oder die Webcam 1'' dahingehend erweitert, das sie, wie in Fig. 1 bis 4 abgebildet, jeweils ein integriertes Mikrofon 3 aufweisen. Mittels dieses Mikrofons 3 können Sprachsteuersignale des Benutzers aufgenommen, in elektrische Signale umgesetzt und mittels einer Verarbeitungseinheit 4
20 (Mikroprozessor), z.B. einem anwendungsspezifischen integrierten Halbleiterbaustein (ASIC), in geeignete Befehle zur Ansteuerung des Computers 7 umgesetzt werden. Der ASIC 4 unterzieht die Ausgangssignale des Mikrofons 3 einem Algorithmus zur automatischen Sprecheridentifikation (sofern nicht ein von der Stimme des jeweiligen Sprechers unabhängiger Spracherkennungsalgorithmus implementiert ist), Spracherkennung und -
25 analyse und interpretiert sie als Steuersignale für den angeschlossenen Rechner.

Vorzugsweise sind die Steuersignale derart, dass sie mit der Konfiguration/Einstellung oder dem Betrieb des Eingabegeräts oder der Webcam zusammenhängen. Es können also per Sprache Einstellungen (Belegung der Freiheitsgrade des Eingabegeräts etc.) vorge-
30 nommen werden oder Funktionen ausgelöst werden. In diesem Fall hat die Treiber-

Software (üblicherweise auf dem Rechner) für das Eingabegerät bzw. die Webcam eine Funktion, die die Auswertung der sprachgesteuerten Steuersignale ermöglicht.

Der Prozessor 4 für die Sprachsignalauswertung ist dabei vorzugsweise in dem Gehäuse
5 des Eingabegeräts bzw. der Webcam angeordnet, üblicherweise als eigener Baustein separat zu den Prozessoren für die Bildsignalauswertung bzw. die Auswertung der Signale von dem Eingabegerät. Vorzugsweise kann die Sprachsignalauswertung als Firmware hinterlegt sein.

10 Die umgesetzten und ausgewerteten Befehle werden dann über dieselbe Schnittstelle 8 zu dem Computer 7 hin übertragen, welche auch für die Übertragung der übrigen Ansteuerbefehle über die Maus 1 bzw. die Tastatur 1' zur Steuerung des Cursors bzw. des Mauszeigers vorgesehen ist.

15 Somit wird vorteilhafterweise für die Eingabe der Sprachsteuersignale und die Eingabe der übrigen Ansteuerbefehle nur ein einziger (bspw. USB-)Anschluss 6 des Computers 7 belegt. Die Umsetzung der Sprachsteuersignale in vom Computer 7 ausführbare Steuerbefehle kann natürlich auch im Computer 7 selbst erfolgen. In diesem Fall werden (digitalisierte) Sprachsignale über die Schnittstelle 8 übertragen.

20 Auf dem manuellen Eingabegerät 1, 1' kann die Sprachsteuerungsfunktion mittels einer speziellen Taste – Taste 5 im Falle der Maus 1 bzw. Taste 5' im Falle der Tastatur 1' – aktiviert werden.

25 Im übrigen ist es auch möglich, die Funktion der Maus 1 mittels Sprachbefehlen aktiv zu schalten („aufzuwecken“), was ein Beispiels dafür ist, wie Sprachbefehle für die Funktion eines Eingabegeräts spezifisch verwendet werden können

Im Übrigen ist es auch vorteilhaft, wenn die Treibersoftware für das Eingabegerät die
30 Sprachsteuerungsfunktion automatisch aktiviert, sobald das Eingabegerät 1, 1' bzw. die Webcam 1'' an den entsprechenden Anschluss 6 (z.B. einen USB-Anschluss 6 bzw. eine serielle RS 232-Schnittstelle etc.) des Computers 7 angeschlossen wird. Insbesondere kann in die Treibersoftware bereits die Sprachsteuerungsfunktion bzw. die Sprach-

steuerungs-Aktivierungsfunktion integriert sein, was den Installationsaufwand für den Benutzer verringert. Mit dem Anschluss des manuellen Eingabegerätes 1, 1' steht dann gleichzeitig auch die Sprachsteuerung bereit.

- 5 Anstelle der in Fig. 1 bis 4 dargestellten, drahtgebundenen Schnittstelle 8, welche zur gemeinsamen Übertragung der in elektrische Signale umgesetzten, erkannten, sprachlich analysierten und interpretierten Sprachsteuersignale zusammen mit den übrigen Ansteuerbefehlen dient, kann selbstverständlich auch eine aus dem Stand der Technik bekannte drahtlose Schnittstelle verwendet werden. Für diese drahtlose Schnittstelle eignet sich u.a.
- 10 eine Infrarot-Schnittstelle oder eine Schnittstelle nach dem Bluetooth-Standard.

3DConnexion GmbH
P 26028 DE

5 Bezugszeichenliste

Nr.	Technische Funktions- bzw. Systemkomponente
1	Kraft-/Momentensensor bzw. Maus des Computers 7 als manuelles Eingabegerät für Steuerbefehle
1'	Tastatur des Computers 7 als manuelles Eingabegerät für Steuerbefehle und Informationen
1''	Internet-Kameraaufsatz (engl.: „Webcam“) des Computers 7
1a	Kabelanschluss zum Computer 7
1b	Bedienteil des manuellen Eingabegeräts 1
1c	Basisplatte des manuellen Eingabegeräts 1
1d	Funktionstasten der Basisplatte 1c
1e	Anzeigevorrichtung mit Touchscreen, integriert in die Basisplatte 1c oder das Bedienteil 1b des manuellen Eingabegeräts 1
1f	(gedachte) vertikale Längsachse des Bedienteils 1b
2	Gehäuse des Kraft-/Momentensensors bzw. der Maus 1
2'	Gehäuse der Tastatur 1'
2''	Gehäuse des Internet-Kameraaufsatzes 1''
2'''	Gehäuse des Bildschirms 7a
3	Mikrofon, integriert in das Gehäuse 2, 2', 2'' bzw. 2''' des Kraft-/Momentensensors bzw. Maus 1, der Tastatur 1', des Internet-Kameraaufsatzes 1'' und/oder des Bildschirms 7a
4	Verarbeitungseinheit (Mikroprozessor), realisiert als anwendungsspezifischer integrierter Halbleiterbaustein (ASIC), welche die vom Mikrofon 3 erfassten Sprachsteuersignale eines Benutzers einem Algorithmus zur automatischen Sprecheridentifikation, Spracherkennung, -analyse und/oder -interpretation unterzieht
5	Taste der Computermaus 1 zur manuellen Aktivierung der Sprachsteuerungsfunktion
5'	Taste der Tastatur 1' zur manuellen Aktivierung der Sprachsteuerungsfunktion
5''	Taste des Internet-Kameraaufsatzes 1'' zur manuellen Aktivierung der Sprachsteuerungsfunktion
5'''	Taste des Computer-Bildschirms 7a zur manuellen Aktivierung der Sprachsteuerungsfunktion

Nr.	Technische Funktions- bzw. Systemkomponente
6	USB-Anschluss bzw. serielle RS 232-Schnittstelle des Computers 7 zur Eingabe von Sprach- und/oder sonstigen Ansteuerbefehlen
7	Recheneinheit (Computer) zur Echtzeit-Bewegungssteuerung von realen und/oder mit Hilfe eines Bildschirms 7a grafisch visualisierten virtuellen Objekten
7a	Bildschirm des Computers 7
8	gemeinsame (drahtgebundene oder drahtlose) Schnittstelle zur Übertragung von verarbeiteten und erkannten Sprachsteuersignalen und/oder übrigen Ansteuerbefehlen über das manuelle Eingabegerät 1, 1' bzw. 1'' zum Computer 7
100	erstes Ausführungsbeispiel eines Systems zur Erzeugung von Steuersignalen für die Echtzeit-Bewegungssteuerung von realen bzw. virtuellen Objekten gemäß der vorliegenden Erfindung, bei dem in das Gehäuse einer Computermaus ein Mikrofon zur Aufnahme und Umsetzung von akustischen Sprachsteuersignalen in elektrische Signale, ein ASIC zur automatischen Sprecheridentifikation, Spracherkennung, -analyse und -interpretation sowie eine Taste zur manuellen Aktivierung der Sprachsteuerungsfunktion integriert sind
200	zweites Ausführungsbeispiel eines Systems zur Erzeugung von Steuersignalen für die Echtzeit-Bewegungssteuerung von realen bzw. virtuellen Objekten gemäß der vorliegenden Erfindung, bei dem in das Gehäuse einer Computertastatur ein Mikrofon zur Aufnahme und Umsetzung von akustischen Sprachsteuersignalen in elektrische Signale, ein ASIC zur automatischen Sprecheridentifikation, Spracherkennung, -analyse und -interpretation und eine Taste zur manuellen Aktivierung der Sprachsteuerungsfunktion integriert sind
300	drittes Ausführungsbeispiel eines Systems zur Erzeugung von Steuersignalen für die Echtzeit-Bewegungssteuerung von realen bzw. virtuellen Objekten gemäß der vorliegenden Erfindung, bei dem in das Gehäuse eines Internet-Kameraaufsatzes (engl.: „Webcam“) ein Mikrofon zur Aufnahme und Umsetzung von akustischen Sprachsteuersignalen in elektrische Signale, ein ASIC zur automatischen Sprecheridentifikation, Spracherkennung, -analyse und -interpretation und eine Taste zur manuellen Aktivierung der Sprachsteuerungsfunktion integriert sind
400	viertes Ausführungsbeispiel eines Systems zur Erzeugung von Steuersignalen für die Echtzeit-Bewegungssteuerung von realen bzw. virtuellen Objekten gemäß der vorliegenden Erfindung, bei dem in das Gehäuse eines Computerbildschirms ein Mikrofon zur Aufnahme und Umsetzung von akustischen Sprachsteuersignalen in elektrische Signale, ein ASIC zur automatischen Sprecheridentifikation, Spracherkennung, -analyse und -interpretation und eine Taste zur manuellen Aktivierung der Sprachsteuerungsfunktion integriert sind

Nr.	Technische Funktions- bzw. Systemkomponente
X	<p data-bbox="411 277 1447 360">dreidimensionale Ansicht des manuellen Eingabegeräts 1 und der Basisplatte 1c mit einem 3D-Koordinatensystem, in dem die sechs Freiheitsgrade</p> $x, y, z \text{ [m]} \text{ und } \varphi_x, \varphi_y, \varphi_z \text{ [rad]}$ <p data-bbox="411 517 1447 600">des manuellen Eingabegeräts 1 sowie deren erste und zweite zeitliche Ableitungen</p> $\dot{x}, \dot{y}, \dot{z} \text{ [m}\cdot\text{s}^{-1}] \text{ und } \dot{\varphi}_x, \dot{\varphi}_y, \dot{\varphi}_z \text{ [rad}\cdot\text{s}^{-1}] \text{ sowie}$ $\ddot{x}, \ddot{y}, \ddot{z} \text{ [m}\cdot\text{s}^{-2}] \text{ und } \ddot{\varphi}_x, \ddot{\varphi}_y, \ddot{\varphi}_z \text{ [rad}\cdot\text{s}^{-2}]$ <p data-bbox="411 801 644 837">eingezeichnet sind</p>

5

Ansprüche

1. Manuelles Eingabegerät zur Erzeugung von Steuersignalen für die Echtzeit-Bewegungssteuerung von realen und/oder mit Hilfe einer Anzeigevorrichtung (7a) grafisch visualisierten virtuellen Objekten,
10 dadurch gekennzeichnet,
dass in das Gehäuse (2, 2') des manuellen Eingabegeräts (1, 1') ein Mikrofon (3) zur Aufnahme und Umsetzung von akustischen Sprachsteuersignalen in elektrische Signale integriert ist.

15

2. Manuelles Eingabegerät nach Anspruch 1,
dadurch gekennzeichnet,
dass in das Gehäuse (2, 2') weiterhin eine Verarbeitungseinheit (4) integriert ist, die Ausgangssignale des Mikrofons (3) einem Algorithmus zur automatischen Sprecheridentifikation, Spracherkennung und -analyse unterzieht und als Steuersignale für die Echtzeit-
20 Bewegung mindestens eines der grafisch visualisierten Objekte interpretiert.

3. Manuelles Eingabegerät nach Anspruch 2,
dadurch gekennzeichnet,
25 dass die Verarbeitungseinheit (4) als anwendungsspezifischer integrierter Halbleiterbaustein (ASIC) ausgeführt ist.

4. Manuelles Eingabegerät nach einem der vorhergehenden Ansprüche,
dadurch gekennzeichnet,
30 dass es sich dabei um eine Computermaus bzw. einen Kraft-/Momentensensor (1) oder um eine Tastatur (1') handelt.

5. Manuelles Eingabegerät nach einem der vorhergehenden Ansprüche,

dadurch gekennzeichnet,

dass die Sprachsteuerungsfunktion durch Betätigung (5, 5') des manuellen Eingabegeräts (1, 1') aktivierbar ist.

- 5 6. Manuelles Eingabegerät nach einem der vorhergehenden Ansprüche, gekennzeichnet durch eine Treibersoftware, welche die Sprachsteuerungsfunktion automatisch aktiviert, sobald das Eingabegerät (1, 1') an einem Anschluss (6) einer Recheneinheit (7) angeschlossen wird.

10

7. Manuelles Eingabegerät nach Anspruch 6, dadurch gekennzeichnet, dass die Funktion zur automatischen Aktivierung der Sprachsteuerungsfunktion in die Treibersoftware integriert ist, die zur Umsetzung der übrigen Ansteuersignale mittels Computermaus (1) und Tastatur (1') vorgesehen ist.

15

8. Manuelles Eingabegerät nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass die in elektrische Signale umgesetzten, erkannten, sprachlich analysierten und interpretierten Sprachsteuersignale zusammen mit den übrigen Ansteuerbefehlen über eine gemeinsame drahtgebundene oder drahtlose Schnittstelle (8) zu einer Recheneinheit (7) übertragen werden.

20

9. Internet-Kameraaufsatz für eine Recheneinrichtung (7) zur Echtzeit-Bewegungssteuerung von realen und/oder mit Hilfe einer Anzeigevorrichtung (7a) grafisch visualisierten virtuellen Objekten, dadurch gekennzeichnet, dass in das Gehäuse (2'') des Internet-Kameraaufsatzes (1'') zur Aufnahme und Umsetzung von akustischen Sprachsteuersignalen in elektrische Signale ein Mikrofon (3) integriert bzw. ein Mikrofon (3) fest an dem Gehäuse (2'') angebracht ist.

25

30

10. Internet-Kameraaufsatz nach Anspruch 9, dadurch gekennzeichnet,

dass in sein Gehäuse (2'') weiterhin eine Verarbeitungseinheit (4) integriert ist, die Ausgangssignale des Mikrofons (3) einem Algorithmus zur automatischen Sprecheridentifikation, Spracherkennung und -analyse unterzieht und als Steuersignale für die Echtzeit-Bewegung mindestens eines der grafisch visualisierten Objekte interpretiert.

5

11. Internet-Kameraaufsatz nach Anspruch 10, dadurch gekennzeichnet, dass die Verarbeitungseinheit (4) als anwendungsspezifischer integrierter Halbleiterbaustein (ASIC) ausgeführt ist.

10

12. Internet-Kameraaufsatz nach einem der Ansprüche 9 bis 11, dadurch gekennzeichnet, dass die Sprachsteuerungsfunktion durch Betätigung einer im Gehäuse (2'') des Internet-Kameraaufsatzes (1'') integrierten Taste (5'') aktivierbar ist.

15

13. Internet-Kameraaufsatz nach einem der Ansprüche 9 bis 12, gekennzeichnet durch eine Treibersoftware, welche die Sprachsteuerungsfunktion automatisch aktiviert, sobald der Internet-Kameraaufsatz (1'') an einem Anschluss (6) der Recheneinheit (7) angeschlossen wird.

20

14. Internet-Kameraaufsatz nach einem der Ansprüche 9 bis 13, dadurch gekennzeichnet, dass die Funktion zur automatischen Aktivierung der Sprachsteuerungsfunktion in die Treibersoftware integriert ist, die zur Umsetzung der übrigen Ansteuersignale vorgesehen ist.

25

15. Internet-Kameraaufsatz nach einem der Ansprüche 9 bis 14, dadurch gekennzeichnet, dass die in elektrische Signale umgesetzten, erkannten, sprachlich analysierten und interpretierten Sprachsteuersignale zusammen mit den übrigen Ansteuerbefehlen über eine gemeinsame drahtgebundene oder drahtlose Schnittstelle (8) zu der Recheneinheit (7) übertragen werden.

30

16. Anzeigevorrichtung für eine Recheneinrichtung (7) zur grafischen Visualisierung der Echtzeit-Bewegungssteuerung von realen und/oder virtuellen Objekten, dadurch gekennzeichnet,
- 5 dass in das Gehäuse (2'') der Anzeigevorrichtung (7a) ein Mikrofon (3) integriert bzw. ein Mikrofon (3) fest an ihrem Gehäuse (2'') angebracht ist.
17. Anzeigevorrichtung nach Anspruch 16, dadurch gekennzeichnet,
- 10 dass in ihr Gehäuse (2'') weiterhin eine Verarbeitungseinheit (4) integriert ist, die Ausgangssignale des Mikrofons (3) einem Algorithmus zur automatischen Sprecheridentifikation, Spracherkennung und -analyse unterzieht und als Steuersignale für die Echtzeit-Bewegung mindestens eines der grafisch visualisierten Objekte interpretiert.
- 15 18. Anzeigevorrichtung nach Anspruch 17, dadurch gekennzeichnet, dass die Verarbeitungseinheit (4) als anwendungsspezifischer integrierter Halbleiterbaustein (ASIC) ausgeführt ist.
- 20 19. Anzeigevorrichtung nach einem der Ansprüche 16 bis 18, dadurch gekennzeichnet, dass die Sprachsteuerungsfunktion durch Betätigung einer im Gehäuse (2'') der Anzeigevorrichtung (7a) integrierten Taste (5'') aktivierbar ist.
- 25 20. Anzeigevorrichtung nach einem der Ansprüche 16 bis 19, gekennzeichnet durch eine Treibersoftware, welche die Sprachsteuerungsfunktion automatisch aktiviert, sobald die Anzeigevorrichtung (7a) an einem Anschluss (6) einer Recheneinrichtung (7) angeschlossen wird.
- 30 21. Anzeigevorrichtung nach Anspruch 20, dadurch gekennzeichnet,

dass die Funktion zur automatischen Aktivierung der Sprachsteuerungsfunktion in die Treibersoftware integriert ist, die zur Umsetzung der übrigen Ansteuersignale mittels Computermaus (1) und Tastatur (1') vorgesehen ist.

- 5 22. Anzeigevorrichtung nach einem der Ansprüche 16 bis 21,
dadurch gekennzeichnet,
dass die in elektrische Signale umgesetzten, erkannten, sprachlich analysierten und interpretierten Sprachsteuersignale zusammen mit den übrigen Ansteuerbefehlen über eine
gemeinsame drahtgebundene oder drahtlose Schnittstelle (8) zu einer Recheneinheit (7)
10 übertragen werden.
23. Verwendung eines manuellen Eingabegerätes (1, 1') nach einem der Ansprüche 1 bis 8 in Kombination mit einem CAD-Programm.
- 15 24. Verwendung eines manuellen Eingabegerätes (1, 1') nach einem der Ansprüche 1 bis 8 in Kombination mit einem Desktop Publishing-Programm.
25. Verwendung eines Internet-Kameraaufsatzes (1'') nach einem der Ansprüche 9 bis 15 in Kombination mit einem CAD-Programm.
- 20 26. Verwendung eines Internet-Kameraaufsatzes (1'') nach einem der Ansprüche 9 bis 15 in Kombination mit einem Desktop Publishing-Programm.
27. Verwendung einer Anzeigevorrichtung nach einem der Ansprüche 16 bis 22 in Kombination mit einem CAD-Programm.
- 25 28. Verwendung einer Anzeigevorrichtung nach einem der Ansprüche 16 bis 22 in Kombination mit einem Desktop Publishing-Programm.

Ausführungsbeispiel Nr. 1

100

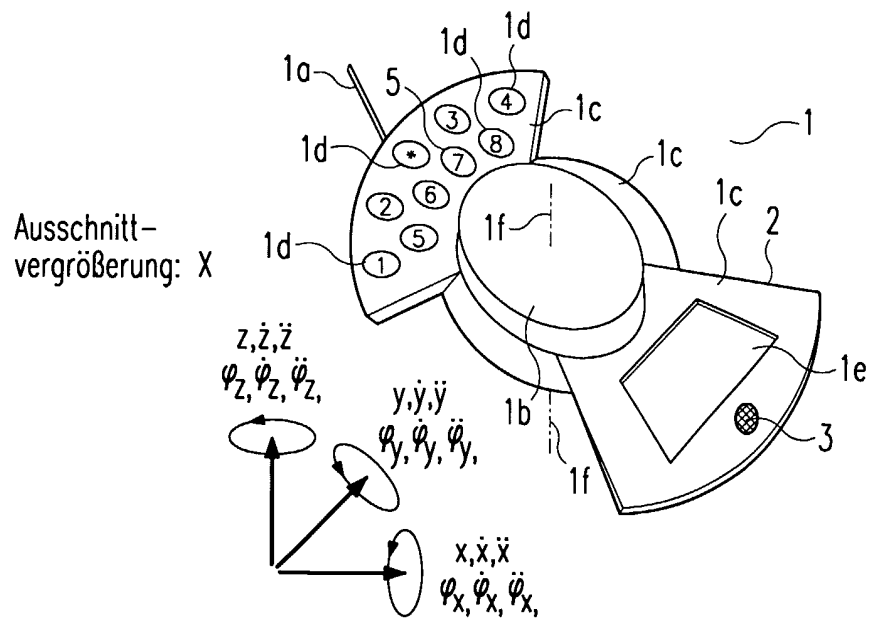
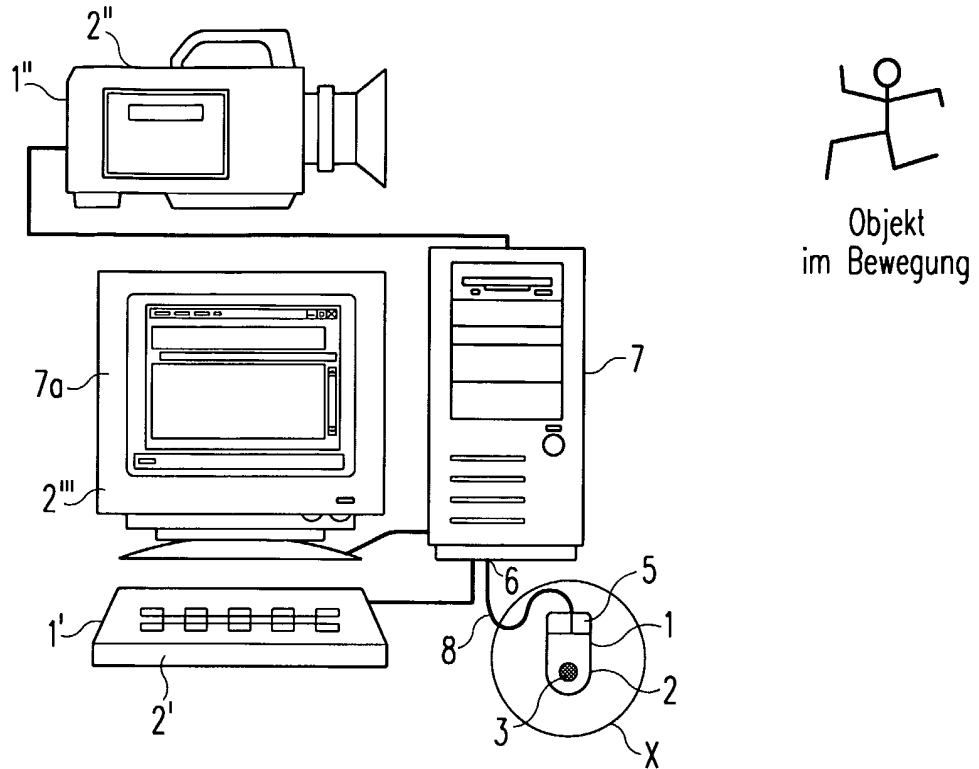


Fig. 1

Ausführungsbeispiel Nr. 2

200

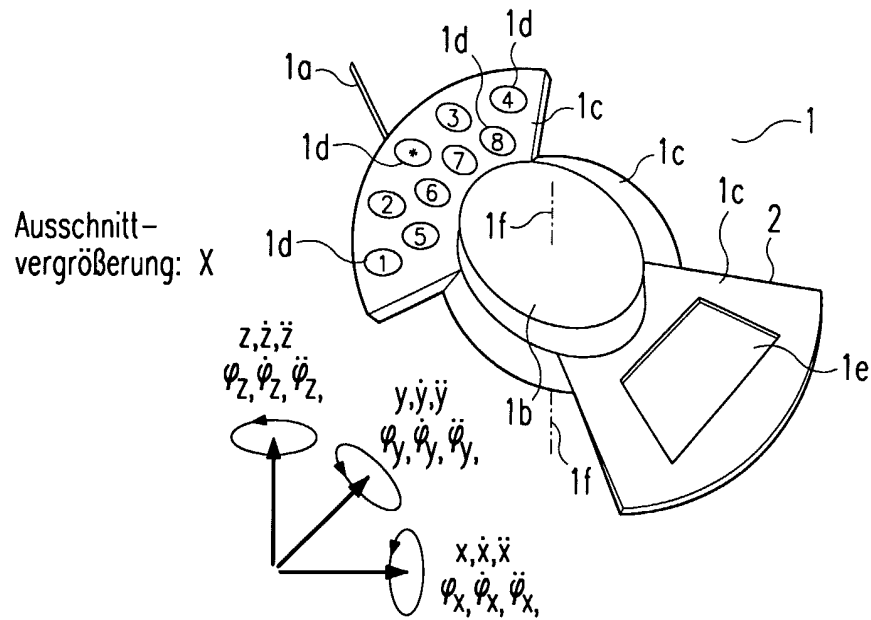
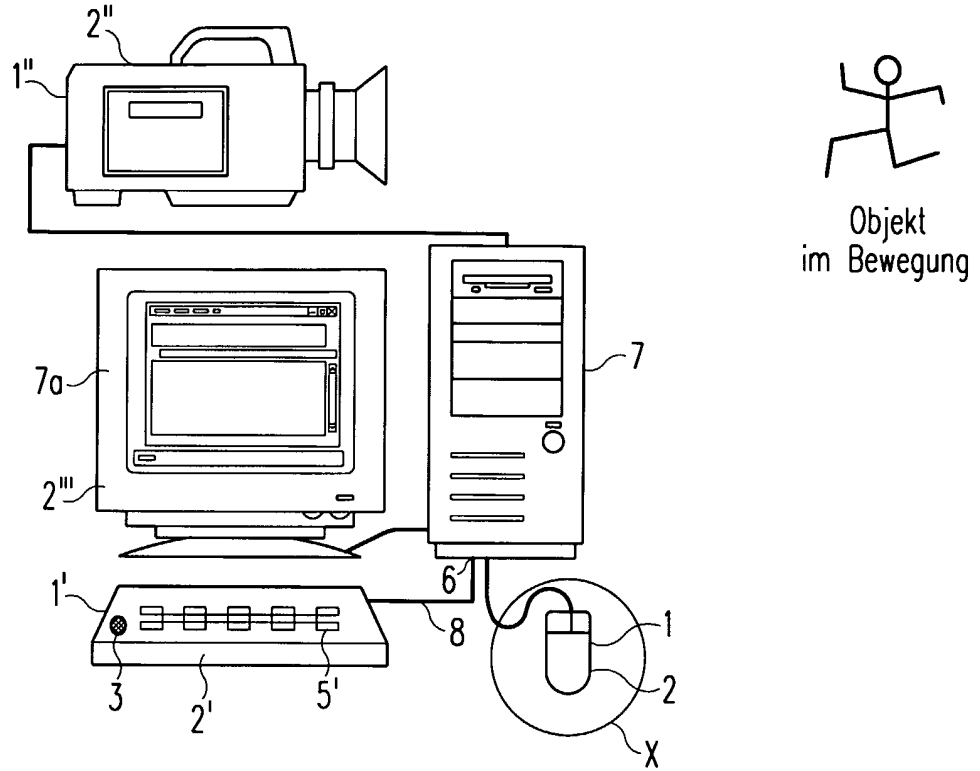


Fig. 2

Ausführungsbeispiel Nr. 3

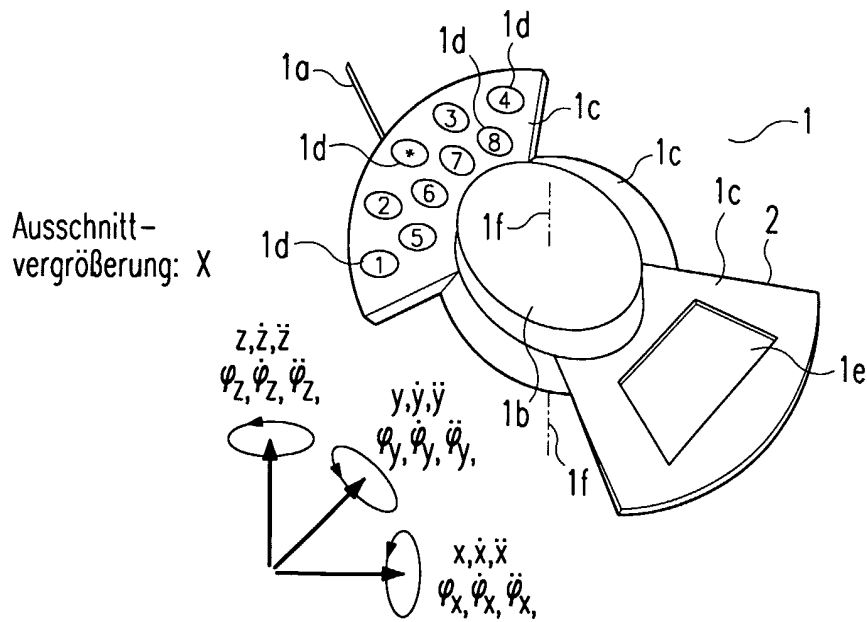
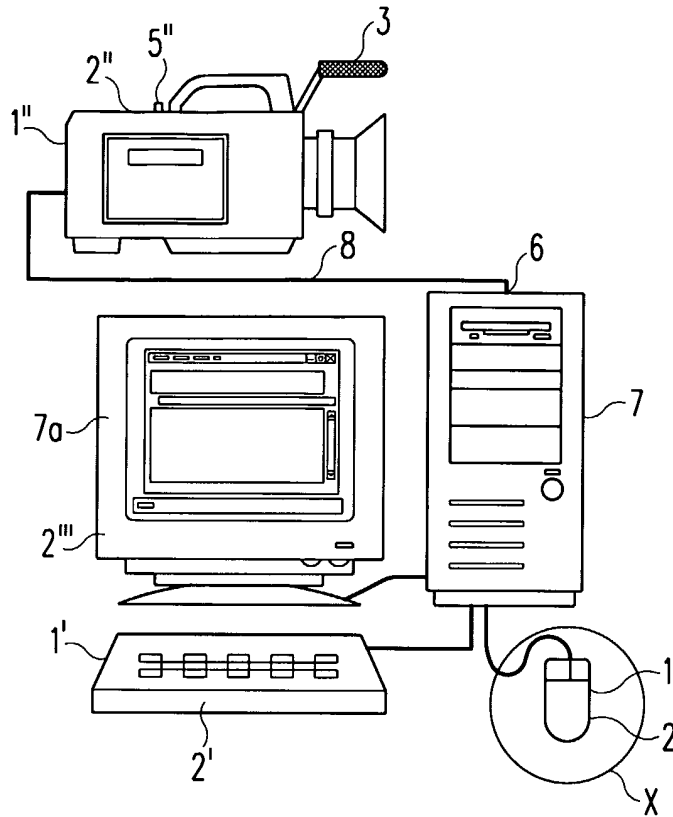


Fig. 3

