



(19)  
Bundesrepublik Deutschland  
Deutsches Patent- und Markenamt

(10) **DE 60 2004 008 794 T2** 2008.06.12

(12) **Übersetzung der europäischen Patentschrift**

(97) **EP 1 671 276 B1**

(51) Int Cl.<sup>8</sup>: **G06T 15/20** (2006.01)

(21) Deutsches Aktenzeichen: **60 2004 008 794.7**

(86) PCT-Aktenzeichen: **PCT/IB2004/051821**

(96) Europäisches Aktenzeichen: **04 770 055.4**

(87) PCT-Veröffentlichungs-Nr.: **WO 2005/031652**

(86) PCT-Anmeldetag: **22.09.2004**

(87) Veröffentlichungstag  
der PCT-Anmeldung: **07.04.2005**

(97) Erstveröffentlichung durch das EPA: **21.06.2006**

(97) Veröffentlichungstag  
der Patenterteilung beim EPA: **05.09.2007**

(47) Veröffentlichungstag im Patentblatt: **12.06.2008**

(30) Unionspriorität:  
**03103608 30.09.2003 EP**

(84) Benannte Vertragsstaaten:  
**AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB,  
GR, HU, IE, IT, LI, LU, MC, NL, PL, PT, RO, SE, SI,  
SK, TR**

(73) Patentinhaber:  
**Koninklijke Philips Electronics N.V., Eindhoven,  
NL**

(72) Erfinder:  
**VAN GEEST, Bartolomeus W., NL-5656 AA  
Eindhoven, NL; OP DE BEECK, Marc J., NL-5656  
AA Eindhoven, NL**

(74) Vertreter:  
**Volmer, G., Dipl.-Ing., Pat.-Anw., 52066 Aachen**

(54) Bezeichnung: **BILDWIEDERGABE MIT INTERAKTIVER BEWEGUNGSPARALLAXE**

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**

Rot, Grün, Blau und Tiefe (RGBD) ausgetauscht.

**[0001]** Die vorliegende Erfindung bezieht sich auf eine Wiedergabeanordnung zum Schaffen interaktiver Bewegungsparallaxe durch Wiedergabe eines Ausgangsbildes auf Basis von 3D visueller Information.

**[0002]** Die vorliegende Erfindung bezieht sich auf ein Verfahren zum Schaffen interaktiver Bewegungsparallaxe durch Wiedergabe eines Ausgangsbildes auf Basis von 3D visueller Information.

**[0003]** Die vorliegende Erfindung bezieht sich auf ein Computerprogrammprodukt, das von einer Computeranordnung geladen werden soll, mit Instruktionen zum Rendern eines Ausgangsbildes auf Basis von 3D visueller Information, wobei die Computeranordnung Verarbeitungsmittel und einen Speicher aufweist.

**[0004]** In dem Bereich der 3D Visualisierung sind eine Anzahl Tiefenaufrufe bekannt, die an dem 3D Empfinden einen Beitrag liefern. Zwei davon sind Stereoskopie und interaktive Bewegungsparallaxe. Bei Stereoskopie werden den Augen des Zuschauers Bilder angeboten, die einen etwas anderen perspektivischen Gesichtspunkt der visualisierten Szene haben. Bei interaktiver Bewegungsparallaxe sind die visualisierten perspektivischen Gesichtspunkte in Bezug auf die Kopflege des Zuschauers adaptiv.

**[0005]** Untenstehend werden zwei Beispiele der Präsentation dieser Tiefenaufrufe an einen Zuschauer kurz beschrieben. In dem ersten Beispiel wird die dreidimensionale (3D) visuelle Information mit Hilfe eines geometrischen 3D-Modells dargestellt. Der Anwendungsbereich umfasst synthetisierten Inhalt, d.h. Computergraphiken, beispielsweise Spielapplikationen und CAD. In dem vorliegenden Fall werden die zu visualisierenden Szenen durch ein geometrisches 3D-Modell beschrieben, beispielsweise VRML ("Virtual Reality Modeling Language") Information über die Kopflege des Zuschauers, gemessen mit einem sog. "Head-tracker" (Kopffolger), wird verwendet um den Gesichtspunkt als Parameter in der Stereobildsynthese (das Rendern) einzustellen. Das rechte und das linke Bild werden beispielsweise an einem CRT-basierten Monitor "gezeitmultiplext" und ein elektrovisueller Schalter in Kombination mit einer passiven Brille, ermöglicht, auf Basis von Polarisierung, die 3D Visualisierung. Dieser Typ von Visualisierung ist nur illustrativ, es können Alternativen, einschließlich Auto-Stereoskopie, angewandt werden.

**[0006]** Das zweite Beispiel gilt die 3D Visualisierung eines Bildes auf Basis von Inhalt. Die 3D visuelle Information wird mit Hilfe von Bildern und entsprechenden Tiefenkarten präsentiert. Die Daten in diesem Format werden beispielsweise gespeichert und als

**[0007]** Das bedeutet, dass jedem Pixel ein Tiefenwert zugeordnet wird, der den Abstand des entsprechenden Szenenpunktes von der Kamera angibt. Der Tiefenteil in dieser Darstellung kann verschiedenartig erhalten werden. Beispielsweise unmittelbar aufgezeichnet zusammen mit den Bilddaten, und zwar unter Verwendung einer Tiefenkamera oder erhalten aus stereographischen Aufzeichnungen unter Anwendung von Disparitätsschätzung. Die adaptive Synthese von Bildern mit neuen Gesichtspunkten aus diesem Eingangsmaterial wird bewerkstelligt durch die Anwendung sog. Bildwölbungstechniken, beispielsweise wie beschrieben in: "View interpolation for image synthesis", von Shenchang Eric Chen und Lance Williams, in "Computer Graphics Annual Conference Series", "Proceedings of SIGGRAPH 93", Seiten 279–288. Dieses Verwölben stammt im Grunde von der Neuabtastung der Pixel des ursprünglichen Eingangsbild, und zwar in einem Ausmaß, das zu den Tiefenwerten umgekehrt proportional ist, und der nachfolgenden Neuabtastung der erhaltenen Daten her. Wenn dieses Verfahren angewandt wird, tritt ein Problem auf, da die Bilder durch den Wölbungsprozess verzerrt werden. Das Ausmaß an Verzerrung ist abhängig von dem angewandten Offset des Gesichtspunktes, aber auch von dem Bildinhalt: wenn die Tiefendarstellung, d.h. die Tiefenabbildung relativ viele Unterbrechungen aufweist, wird es oft passieren, dass in bestimmten Gebieten des neuen Bildes, Objekte neu erscheinen sollen: Entdeckung. Diese Information ist nicht verfügbar, da das Objekt in dem ursprünglichen Bild versteckt war. Dies lässt Löcher in dem künstlichen Bild, die irgendwie aufgefüllt werden sollen, aber immerhin wird die Bildqualität beeinträchtigt. Das Ausmaß, in dem diese Degradation des Bildes von dem Zuschauer wahrgenommen wird, ist wieder von dem Inhalt abhängig: wenn der Hintergrund um einen Gegenstand eine homogene Art hat, wird das Auffüllen der Löcher mit anderer Hintergrundinformation weniger störend wirken. Wenn auf interaktive Bewegungsparallaxe angewandt, können die Verzerrungen für relativ große Kopfbewegungen, wenn der Zuschauer den Stuhl verlagert, stärker sein.

**[0008]** Es ist nun u. a. eine Aufgabe der vorliegenden Erfindung, eine Wiedergabeanordnung der eingangs beschriebenen Art zu schaffen, die vorgesehen ist zum Rendern eines Vorgabebildes entsprechend eines vorbestimmten Bildes der 3D visuellen Information, wenn ein Zuschauer während einer bestimmten Zeitspanne kaum bewegt.

**[0009]** Diese Aufgabe der vorliegenden Erfindung wird dadurch erfüllt, dass die Wiedergabeanordnung Folgendes umfasst:

- erste Empfangsmittel zum Empfangen eines ersten Signals, das die 3D visuelle Information

darstellt,

- zweite Empfangsmittel zum Empfangen eines zweiten Signals, das Positionsinformation eines Zuschauers des Ausgangsbildes darstellt, als eine Funktion der Zeit, wobei die Positionsinformation gegenüber der Wiedergabeanordnung relativ ist,
- Filtermittel zur Hochpassfilterung des zweiten Signals, was zu einem dritten Signal führt;
- Aufbereitungsmittel zum Aufbereiten des Ausgangsbildes auf Basis des ersten Signals und der jüngsten Werte des dritten Signals; und
- Wiedergabemittel zum Wiedergeben des Ausgangsbildes.

**[0010]** Ein wichtiger Aspekt der vorliegenden Erfindung ist die Filterung des zweiten Signals, das Lageninformation des Zuschauers des Bildes darstellt. Durch Filterung des zweiten Signals gibt es keine lineare Beziehung zwischen der wirklichen Lageninformation und dem Ausgang der Renderingmittel, es gibt aber eine Beziehung zwischen der Änderung der wirklichen Lageninformation je Zeiteinheit und dem Ausgang der Renderingmittel. Das bedeutet, dass wenn die Änderung wirklicher Lageninformation während einer bestimmten Zeitperiode Null ist, d.h. wenn die Geschwindigkeit des Zuschauers Null ist, ist der Ausgang der Filtermittel gleich Null. Dadurch werden die Renderingmittel das Vorgabebild entsprechend einer Vorgabelageninformation, und zwar einem vorbestimmten Bild der 3D visuellen Information, rendern. Andererseits, wenn die Änderung der wirklichen Lageninformation während einer bestimmten Zeitperiode relativ groß ist, d.h. wenn die Geschwindigkeit und/oder die Beschleunigung des Zuschauers relativ groß ist, ist der Ausgang der Filtermittel relativ hoch, was zu einer Folge von gerenderten Ausgangsbildern führt, entsprechend relativ großen Winkeln in Bezug auf das Vorgabebild. Der Vorteil der Wiedergabeanordnung nach der vorliegenden Erfindung ist, dass sie vorgesehen ist um auf schnelle Kopfbewegungen des Zuschauers, entsprechend Bewegungen zum Observieren interaktiver Bewegungsparallaxe, zu reagieren, während sie vorgesehen ist um ein bevorzugtes Vorgabebild wiederzugeben, wenn eine kürzlich erfolgte Bewegung nicht als solche gemeint war sondern beispielsweise dadurch verursacht wird, dass nur eine andere Haltung angenommen wird oder dass nur der Stuhl, in dem der Zuschauer sitzt, verlagert wird. In dem letzteren Fall wird die Wiedergabeanordnung ggf. in einen Zustand gehen, in dem das genannte Vorgabebild wiedergegeben wird, wenn nach der kürzlich erfolgten Bewegung der Zuschauer eine Zeitlang sich kaum bewegt hat.

**[0011]** Die 3D visuelle Information könnte verschiedenartig dargestellt werden: als ein 3D-Modell in VRML, als einen Satz von Volumenelementen, als einen Satz von Oberflächenbeschreibungen oder als ein Bild plus Tiefenabbildung.

**[0012]** In einer Ausführungsform der Wiedergabeanordnung nach der vorliegenden Erfindung umfasst die 3D visuelle Information ein Eingangsbild und eine entsprechende Tiefenabbildung und das Eingangsbild und das Ausgangsbild entsprechen einander im Wesentlichen um einen vorbestimmten Wert des dritten Signals, während für einen weiteren Wert des dritten Signals das Ausgangsbild ein anderes Gesicht auf die Szene darstellt als ein erstes Gesicht auf die Szene entsprechend dem Eingangsbild. Mit anderen Worten: die Wiedergabeanordnung nach der vorliegenden Erfindung gibt ein Ausgangsbild mit minimaler Verzerrung wieder. Folglich ist die Bildqualität optimal, wenn der Zuschauer sich einige Zeit nicht verlagert hat. Es können zwischen dem Eingangsbild und dem Ausgangsbild geringfügige Unterschiede sein, d.h. die Bilder stimmen miteinander im Wesentlichen überein und sind einander nicht unbedingt gleich. Diese Unterschiede könnten beispielsweise durch geringfügige Wölbungsvorgänge, durch Quantisierung oder durch andere Bildverarbeitungsvorgänge, durchgeführt zum Berechnen des Ausgangsbildes auf Basis des Eingangsbildes, verursacht werden.

**[0013]** Eine Ausführungsform der Wiedergabeanordnung nach der vorliegenden Erfindung umfasst weiterhin Beschneidemittel zum Beschneiden des dritten Signals zwischen einer unteren Grenze und einer oberen Grenze. Das dritte Signal, das von dem Kopffolger herrührt, wird derart gefiltert, dass relativ große Standpunktverschiebungen vermieden werden. Dies vermeidet die assoziierte Verzerrung auf Kosten der Standpunktsanpassungsfähigkeit für relativ große Kopfbewegungen.

**[0014]** Eine Ausführungsform der Wiedergabeanordnung nach der vorliegenden Erfindung umfasst weiterhin Inhaltsanalysiermittel, vorgesehen zum Ermitteln eines Maßes eines Satzes von Maßen mit einem ersten Maß entsprechend der Anzahl Unterbrechungen in der Tiefenabbildung, ein zweites Maß entsprechend der Homogenität des Hintergrundes des Eingangsbildes und ein drittes Maß entsprechend der Anzahl Löcher in dem Ausgangsbild, wobei die Inhaltsanalysiermittel (**128**) zum Anpassen der Grenzfrequenz der Filtermittel (**122**) und/oder zum Ändern der unteren und/oder oberen Grenze der Beschneidemittel (**124**) vorgesehen ist. Die angewandte Steuerung ist vorzugsweise wie folgt:

- die Inhaltsanalysiermittel sind vorgesehen zur Steigerung der unteren Grenze und/oder zur Verringerung der oberen Grenze, wenn das erste Maß relativ hoch oder das zweite Maß relativ niedrig ist oder das dritte Maß relativ hoch ist; und
- die Inhaltsanalysiermittel sind vorgesehen um die Grenzfrequenz der Filtermittel zu verringern, wenn das erste Maß relativ hoch oder das zweite Maß relativ niedrig ist oder das dritte Maß relativ hoch ist.

**[0015]** Auf alternative Weise wird das Steuersignal offline ermittelt und in die 3D visuelle Information als Metadaten eingebettet.

**[0016]** Vorzugsweise ist die Wiedergabeordnung eine Mehrfachwiedergabeordnung, vorgesehen zum Rendern eines weiteren Ausgangssignals und zum Wiedergeben des Ausgangsbildes in einer ersten Richtung zum Wiedergeben des weiteren Ausgangsbildes in einer zweiten Richtung. Mit anderen Worten, es ist vorteilhaft, die vorliegende Erfindung in einer 3D Wiedergabeordnung, die auch als stereoskopische Wiedergabeordnung bezeichnet wird, anzuwenden.

**[0017]** Es ist eine weitere Aufgabe der vorliegenden Erfindung, ein Verfahren der eingangs beschriebenen Art zu schaffen, ein Vorgabebild entsprechend einem vorbestimmten Bild der 3D visuellen Information zu rendern, wenn ein Zuschauer während einer bestimmten Zeitspanne kaum bewegt.

**[0018]** Diese Aufgabe wird nach der vorliegenden Erfindung dadurch erfüllt, dass das Verfahren Folgendes umfasst:

- das Empfangen eines ersten Signals, das die 3D visuelle Information darstellt,
- das Empfangen eines zweiten Signals, das Positionsinformation eines Zuschauers des Ausgangsbildes darstellt, als eine Funktion der Zeit, wobei die Positionsinformation gegenüber der Wiedergabeordnung relativ ist,
- das Hochpassfiltern des zweiten Signals, was zu einem dritten Signal führt;
- das Aufbereiten des Ausgangsbildes, und zwar auf Basis des ersten Signals und neuer Werte des dritten Signals; und
- das Wiedergeben des Ausgangsbildes.

**[0019]** Es ist eine weitere Aufgabe der vorliegenden Erfindung, ein Computerprogrammprodukt der eingangs beschriebenen Art zu schaffen, um ein Vorgabebild entsprechend einem vorbestimmten Bild der 3D visuellen Information zu rendern, wenn ein Zuschauer sich während einer bestimmten Zeitspanne kaum bewegt.

**[0020]** Diese Aufgabe der vorliegenden Erfindung wird dadurch erfüllt, dass das Computerprogrammprodukt, nachdem es geladen worden ist, die genannten Verarbeitungsmittel mit der Fähigkeit versieht, Folgendes durchzuführen:

- das Empfangen eines ersten Signals, das die 3D visuelle Information darstellt,
- das Empfangen eines zweiten Signals, das Positionsinformation eines Zuschauers des Ausgangsbildes darstellt, als eine Funktion der Zeit, wobei die Positionsinformation gegenüber der Wiedergabeordnung relativ ist,
- das Hochpassfiltern des zweiten Signals, was

zu einem dritten Signal führt;

- das Aufbereiten des Ausgangsbildes, und zwar auf Basis des ersten Signals und neuer Werte des dritten Signals.

**[0021]** Modifikationen der Wiedergabeordnung und Variationen davon können Modifikationen und Variationen davon des Verfahrens und des Computerprogrammproduktes entsprechen, wie diese beschrieben worden sind.

**[0022]** Ausführungsbeispiele der vorliegenden Erfindung sind in der Zeichnung dargestellt und werden im Folgenden näher beschrieben. Es zeigen:

**[0023]** [Fig. 1](#) eine schematische Darstellung einer Ausführungsform der Wiedergabeordnung nach der vorliegenden Erfindung,

**[0024]** [Fig. 2](#) drei verschiedene Ausgangsbilder, die mit Hilfe der Wiedergabeordnung nach [Fig. 1](#) erzeugt worden sind,

**[0025]** [Fig. 3](#) eine schematische Darstellung einer Ausführungsform einer stereoskopischen Wiedergabeordnung nach der vorliegenden Erfindung,

**[0026]** [Fig. 4](#) eine schematische Darstellung der Übertragungskennlinie der Beschneideeinheit; und

**[0027]** [Fig. 5](#) ein von dem Kopffolger geliefertes Kopffolgersignal und das aus diesem Kopffolgersignal hergeleitete hochpassgefilterte Signal.

**[0028]** In den Figuren sind gleiche Elemente mit denselben Bezugszeichen angegeben.

**[0029]** [Fig. 1](#) zeigt schematisch eine Ausführungsform der Wiedergabeordnung **100** nach der vorliegenden Erfindung. Die Wiedergabeordnung **100** ist vorgesehen zum Wiedergeben eines Ausgangsbildes auf Basis von 3D visueller Information und gelieferter Lageninformation. Typischerweise ist die Wiedergabeordnung **100** mit einem Kopffolger **108** verbunden, der vorgesehen ist zum Ermitteln der Lage **102** eines Beobachters **104**, d.h. eines Zuschauers, gegenüber der Wiedergabeordnung **100**. Auf alternative Weise umfasst die Wiedergabeordnung **100** einen derartigen Kopffolger **108**. Die Lage **102** des Beobachters **104** kann durch ein Ultraschallfolgesystem angastet werden oder der Beobachter **194** kann einen Magneten tragen um seine Position **102** gegenüber dem Magnetenfolgesystem anzugeben. Bei einer weiteren Ausführungsform können eine oder mehrere Kameras das Beobachtungsgebiet abtasten um die Position des Beobachters zu ermitteln, beispielsweise durch Liefern von Bilddaten zu einem System, das die Augen des Beobachters erkennt. In noch einer anderen Ausführungsform trägt der Beobachter **104** einen Reflektor,

der elektromagnetische Energie, wie IR-Energie, reflektiert. Eine abtastende IR-Quelle und ein IR-Detektor oder eine IR-Weitwinkelquelle und ein abtastender IR-Detektor ermitteln die Position des Reflektors, der vorzugsweise zwischen den Augen des Beobachters **104** getragen wird.

**[0030]** Die Wiedergabeordnung **100** umfasst Folgendes:

- eine erste Eingangseinheit **101** zum Empfangen eines ersten Signals 3DV, das die 3D visuelle Information darstellt;
- eine zweite Eingangseinheit **116** zum Empfangen eines zweiten Signals P, das die Positionsinformation des Beobachters als Funktion der Zeit darstellt;
- eine Hochpassfiltereinheit **122** zur Hochpassfilterung des zweiten Signals P, was zu einem dritten Signal PF führt;
- eine Renderingeinheit **118** zum Rendern des Ausgangsbildes auf Basis des ersten Signals 3DV und des dritten Signals PF; und
- eine Wiedergabeordnung **112** zur Wiedergabe des Ausgangsbildes.

**[0031]** Die Wiedergabeordnung **100** umfasst ggf. eine Beschneideeinheit **124** zum Beschneiden des zweiten Signals PF und eines dritten Signals zwischen einer unteren Grenze und einer oberen Grenze, was zu einem vierten Signal PFC führt.

**[0032]** Die Wiedergabeordnung **100** umfasst ggf. eine Signaltransformationseinheit **126** zum transformieren des vierten Signals PFC in ein fünftes Signal PP mit Werten, die zum Rendern geeignet sind. Die Transformation könnte eine Skalierung oder eine Abbildung zwischen Koordinatensystemen aufweisen, beispielsweise zum Transformieren von Weltkoordinaten des Beobachters in Bildkoordinaten der 3D visuellen Information oder von Kartesischen Koordinaten in Polenkoordinaten.

**[0033]** Die Wirkungsweise der Wiedergabeordnung **100** wird nachstehend im Zusammenhang mit [Fig. 1](#) und [Fig. 2](#) näher beschrieben. [Fig. 2](#) zeigt drei verschiedene Ausgangsbilder **200–204**, die mit Hilfe der Wiedergabeordnung **100** nach der vorliegenden Erfindung erzeugt worden sind. Es wird nun vorausgesetzt, dass das erste Signal 3DV ein Eingangsbild und eine entsprechende Tiefenabbildung aufweist. Es wird auch vorausgesetzt, dass der Beobachter **104** sich an einer bestimmten Stelle **102** zu einem bestimmten Zeitpunkt vor der Wiedergabeordnung **112** befindet. Diese bestimmte Stelle **102** entspricht dem räumlichen Ursprung des Koordinatensystems des Kopffolgers **108**. Die Wiedergabeordnung **100** gibt ein erstes Bild **200** der Ausgangsbilder wieder. Das erste Bild **200** der Ausgangsbilder stellt einen Teil einer Person dar, beispielsweise einen Kopf **208**, Schultern **210** und einen rechten Arm

**212**. Es sieht aus, als könne der Beobachter **104** die Person recht in die Augen **206** schauen. Das erste Bild **200** der Ausgangsbilder ist im Wesentlichen gleich dem Eingangsbild, das der Wiedergabeordnung **100** zugeführt wird.

**[0034]** Danach bewegt der Beobachter **104** schnell in eine durch einen ersten Pfeil **105** angegebenen Richtung. Der Kopffolger **108** detektiert die Bewegung und liefert das zweite Signal P auf entsprechende Weise. Das zweite Signal wird mit Hilfe der Hochpassfiltereinheit **122** hochpassgefiltert. Das zweite Signal wird mit Hilfe der Hochpassfiltereinheit **122** hochpassgefiltert. Das Ausgangssignal der Hochpassfiltereinheit **122** wird ggf. zugeschnitten und transformiert und ggf. der Renderingeinheit **118** zugeführt. Folglich startet die Renderingeinheit **118** die Berechnung einer Reihe von Ausgangsbildern auf Basis des Eingangsbildes, der Tiefenabbildung und der gefilterten Positionsinformation. Jedes der Ausgangsbilder basiert auf einem anderen Wert des verarbeiteten Signals entsprechend der Positionsinformation. Die Ausgangsbilder werden vorzugsweise berechnet, wie in: "View interpolation for image synthesis", von Shenchang Eric Chen und Lance Williams, in "Computer Graphics Annual Conference Series", "Proceedings of SIGGRAPH 93", Seiten 279–288 beschrieben. Ein zweites Bild **204** der Reihe von Ausgangsbildern, wie an der Wiedergabeordnung **112** wiedergegeben, ist in [Fig. 2](#) dargestellt. Dieses zweite Bild **204** der Ausgangsbilder stellt den Teil der Person dar, d.h. den Kopf **208**, die Schultern **210** und den rechten Arm **212**. Nun scheint es, als könne der Beobachter **104** nicht die Person recht in die Augen **206** schauen, sondern es sieht aus, als habe die Person den Kopf etwa nach links gedreht.

**[0035]** Wenn der Beobachter **104** danach relativ schnell in der entgegengesetzten Richtung bewegt, d.h. in der Richtung, angegeben durch den zweiten Pfeil **103**, wird ein gleicher Prozess durchgeführt. Die Folge ist, dass dem Beobachter **104** ein drittes Bild **204** der Ausgangsbilder gezeigt wird. Dieses dritte Bild **204** der Ausgangsbilder stellt auch den Teil der Person dar, d.h. den Kopf **208**, die Schultern **210** und den rechten Arm **212**. Auch hier sieht es aus, als könne der Beobachter **104** nicht die Person recht in die Augen **206** sehen. Nun aber sieht es aus, als habe die Person den Kopf **208** etwas nach rechts gedreht.

**[0036]** Die Beschneideeinheit **214** wird das dritte hochpassgefilterte Signal PF beschneiden, wenn es vorbestimmte Schwellen übersteigt. Folglich wird dem Beobachter **104** dasselbe dritte Bild **204** der Ausgangsbilder für die beiden Positionen **107** und **109** entsprechend den Abständen  $d_1$  und  $d_2$  in Bezug auf den Ursprung **102** präsentiert.

**[0037]** Wie oben beschrieben wird wegen der Bewegungen dem Beobachter **104** verschiedene Aus-

gangsbilder **200–204** entsprechend verschiedenen Gesichter auf eine Szene präsentiert. In diesem Beispiel umfasst die Szene eine sprechende Person. Dieses Bildpräsentationsphänomen wird als interaktive Bewegungsparallaxe bezeichnet.

**[0038]** Es wird vorausgesetzt, dass der Beobachter sich an einer zweiten Stelle **107** befindet und sich eine Zeitlang, beispielsweise 1–5 Sekunden, nicht berührt hat. Dadurch ist der Wert des hochpassgefilterten dritten Signals PF gleich Null. Die Renderingeinheit **118** wird das Vorgabeausgangsbild erzeugen, d.h. das erste Bild **200** der Ausgangsbilder.

**[0039]** Wenn der Beobachter startet, sich aus der zweiten Lage **107** in einer durch den zweiten Pfeil **103** angegebenen Richtung zu verlagern, wird dem Beobachter das dritte Bild **204** der Ausgangsbilder präsentiert. Wenn der Beobachter startet, sich von der zweiten Lage **107** aus in der durch den ersten Pfeil **105** angegebenen entgegengesetzten Richtung verlagert, wird dem Beobachter das zweite Bild **202** der Ausgangsbilder präsentiert.

**[0040]** Die erste Eingangseinheit **101**, die zweite Eingangseinheit **116**, die Hochpassfiltereinheit **122**, die Beschneideeinheit **124**, die Renderingeinheit **118** und die Signaltransformationseinheit **126** können unter Verwendung eines einzigen Prozessors implementiert werden. Normalerweise werden diese Funktionen unter Ansteuerung eines Softwareprogrammproduktes durchgeführt. Während der Durchführung wird normalerweise das Softwareprogrammprodukt in einen Speicher, wie einen RAM, geladen und von hieraus durchgeführt. Das Programm kann aus einem Hintergrundspeicher, wie einem ROM, einer Festplatte oder einem magnetischen und/oder optischen Speicher geladen werden, oder es kann über ein Netzwerk, wie Internet, geladen werden. Gegebenenfalls schafft eine applikationsspezifische integrierte Schaltung die beschriebene Funktionalität.

**[0041]** [Fig. 3](#) zeigt schematisch eine Ausführungsform einer stereoskopischen Wiedergabeordnung **300** nach der vorliegenden Erfindung. Die Wirkungsweise dieser Ausführungsform **200** ist der Wirkungsweise der Ausführungsform **100**, wie diese im Zusammenhang mit den [Fig. 1](#) und [Fig. 2](#) beschrieben worden ist, im Wesentlichen gleich. Einige Unterschiede werden nachstehend näher beschrieben.

**[0042]** Die stereoskopische Wiedergabeordnung **300** umfasst eine Renderingeinheit **118** zum Rendern eines Linksauge-Ausgangsbildes und eine weitere Renderingeinheit **120** zum Rendern eines Rechtsauge-Ausgangsbildes, wobei das Linksauge-Ausgangsbild und das Rechtsauge-Ausgangsbild ein Stereopaar bilden. Die beiden Ausgangsbilder des Stereopaars werden berechnet, wie im Zusammenhang mit [Fig. 1](#) beschrieben worden ist, sei es, dass

für die Renderingeinheit **118** und die weitere Renderingeinheit **120** verschiedene Lageninformationssignale PPL und PPR geliefert werden. Der Unterschied zwischen diesen zwei Signalen PPL und PPR bezieht sich auf den Abstand (oder den vorausgesetzten Abstand) zwischen den Augen des Beobachters **104**. Das Linksauge-Ausgangsbild und das Rechtsauge-Ausgangsbild werden mit Hilfe der Multiplexereinheit **114** gezeitmultiplext und an der CRT-basierten Wiedergabeordnung **112** wiedergegeben. Der elektrooptische Schalter **110** ermöglicht in Kombination mit einer passiven Brille **106** auf Basis von Polarisation die stereoskopische Visualisierung. Dieser Visualisierungstyp ist nur illustrativ, es können Alternativen, einschließlich Auto-Stereoskopie, angewandt werden.

**[0043]** Die stereoskopische Wiedergabeordnung **300** umfasst weiterhin einen Bildinhaltsanalysator **128**, der zur Steuerung der Beschneideeinheit **124** und der Hochpassfiltereinheit **122** vorgesehen ist. Das Verhalten der Wiedergabeordnung **300** ist derart, dass eine geeignete Bildqualität der Ausgangsbilder angestrebt wird. Das bedeutet, dass im Falle der Ausgangsbilder mit erwarteter niedriger Qualität die Beschneideeinheit **124** den linearen Teil **406** der Transformationscharakteristik **400** verengt. Das Verengen des linearen Teils entspricht der Abnahme des maximalen Ausgangswertes  $C_{max}$  **402** und/oder Zunahme des minimalen Ausgangswertes  $C_{min}$  **404**. Die Erwartung kann auf der Anzahl bei der Wölbung des Eingangsbildes in die Ausgangsbilder gezählter Löcher oder den Analysen des Hintergrundes der Eingangsbilder basieren. Analysen des Hintergrundes umfassen Strukturanalysen, beispielsweise mit Hilfe von Hochpassfilterung des Eingangsbildes, mit einem etwaigen nachfolgenden Schwellenbestimmungsvorgang. Das Vorhandensein relativ vieler HF-Anteile ist ein Hinweis auf einen detaillierten Hintergrund.

**[0044]** Vorzugsweise wird während der Wölbung Information über den Hintergrund verwendet. Ein bekanntes Verfahren zum Reduzieren der Verzerrungsprobleme ist das Bild samt Tiefe mit Information über versteckten Gebieten zu ergänzen. Derartige Information ist für Bild samt Tiefeninformation verfügbar, erhalten aus Stereoaufzeichnungen. Weiterhin machen immer mehr Filme Gebrauch von "Chromakey"-Verfahren. Dies ist ein Verfahren, wobei die Darsteller vor einem blauen oder grünen im Filmstudio vorgesehenen Hintergrund spielen. In einem späteren, dem Editierstadium wird der ursprüngliche blaue oder grüne Hintergrund durch den beabsichtigten Hintergrund ersetzt (ausgetastet), der auf allen Arten von Filmmaterial beruhen kann, beispielsweise Aufnahmen, oder sogar vom Computer erzeugtem Material. Für derartige Fälle ist der komplette Hintergrund, einschließlich der Teile, die von Darstellern abgedeckt sind, verfügbar und kann in Kombina-

tion mit dem Bild samt Tiefeninformation ausgetauscht werden. Der Videocodierungsstandard MPEG-4 unterstützt derartige Ergänzungen durch Verwendung sog. Verbesserungsschichten.

**[0045]** Wie bereits erwähnt, ist das Verhalten der Wiedergabeordnung **300** derart, dass eine geeignete Bildqualität der Ausgangsbilder angestrebt wird. Das bedeutet, dass die Hochpassfiltereinheit **122** schneller reagiert um zu dem Vorgabeausgangsbild **200** zurückzukehren im Falle von Ausgangsbildern mit erwarteter niedrigerer Qualität. Die Erwartung kann auf der Anzahl Löcher basiert sein, die während der Verwölbung des Eingangsbildes in die Ausgangsbilder oder der Analysen des Hintergrundes der Eingangsbilder gezählt wurden. Die Schätzung der Anzahl Unterbrechungen in der Tiefenabbildung ist eine andere Art und Weise der Quantifizierung der erwarteten Bildqualität.

**[0046]** Obschon die 3D visuelle Information als Bild samt Tiefenabbildung in den Ausführungsformen der Wiedergabeordnung **100** und **200**, wie im Zusammenhang mit [Fig. 1](#) und [Fig. 3](#) beschrieben worden ist, geliefert wird, dürfte es einleuchten, dass alternative Ausführungsformen imstande sind, die 3D visuelle Information zu empfangen, die auf eine andere Art und Weise dargestellt wird, beispielsweise als ein 3D-Modell in VRML, als einen Volumensatz von Voxeln oder als einen Satz von Oberflächenbeschreibungen. In dem Fall werden von den Renderingeinheiten **118** und **120** andere Renderingtypen durchgeführt.

**[0047]** Gegebenenfalls wird die Filterkennlinie der Hochpassfiltereinheit **122** auf Basis der Beschnideeinheit **124** gesteuert. Die Grenzfrequenz wird abhängig von der Tatsache, ob der Eingang PF der Beschnideeinheit **124** beschnitten ist oder nicht, angepasst. Außerdem wird bevorzugt, dass die Hochpassfiltereinheit **122** ein sog. asymmetrisches Verhalten aufweist, beispielsweise schnelle Reaktionen auf Bewegungen aber langsame Reaktionen auf stationäre Zustände.

**[0048]** Die Wiedergabeordnung könnte ein Teil eines Videokonferenzsystems, einer Konsumentenordnung, wie eines Fernsehers oder einer Spieleinrichtung sein.

**[0049]** [Fig. 5](#) zeigt ein (Eingangs) Kopffolgersignal P, das von einem Kopffolger **108** geliefert wird, und das (Ausgangs) hochpassgefilterte Signal PF, hergeleitet von diesem Kopffolgersignal P. Das angewandte Filter ist ein Hochpassfilter mit einer Grenzfrequenz von 0,05 Hz. Aus [Fig. 5](#) ist deutlich ersichtlich, dass das hochpassgefilterte Signal PF mit dem Kopffolgersignal P während der Zeit = 0 bis zur Zeit = 5 Sekunden durchaus übereinstimmt. Nach der Zeit = 6 Sekunden verwandelt sich das hochpassgefilterte

Signal PF langsam in den Vorgabewert, der zu der betreffenden Position **102** entsprechend dem räumlichen Ursprung des Koordinatensystems des Kopffolgers **108** gehört. Mit anderen Worten der NF-Teil in dem Kopffolgersignal P, der einem räumlichen Offset von etwa 0,2–0,25 m entspricht, wird unterdrückt.

**[0050]** Es sei bemerkt, dass die oben genannten Ausführungsformen die vorliegende Erfindung illustrieren statt begrenzen und dass der Fachmann imstande sein wird, im Rahmen der beiliegenden Patentansprüche alternative Ausführungsformen zu entwerfen. In den Patentansprüchen sollen eingeklammerte Bezugszeichen nicht als den Anspruch begrenzend betrachtet werden. Das Wort "enthalten" schließt das Vorhandensein von Elementen oder Verfahrensschritten, die nicht in dem Anspruch erwähnt werden, nicht aus. Das Wort "ein" vor einem Element schließt das Vorhandensein einer Anzahl derartiger Elemente nicht aus. Die vorliegende Erfindung kann mit Hilfe von Hardware mit verschiedenen einzelnen Elementen und mit Hilfe eines auf geeignete Art und Weise programmierten Computers implementiert werden. In den Einrichtungsansprüchen, in denen verschiedene Mittel nummeriert sind, können verschiedene dieser Mittel von ein und demselben Hardware-Item verkörpert werden. Die Verwendung des Wortes erste, zweite und dritte usw. gibt nicht eine gewisse Reihenfolge an. Diese Wörter sollen als Name betrachtet werden.

Text in der Zeichnung

#### [Fig. 4](#)

Ausgang PFC  
Eingang PF

#### [Fig. 5](#)

Filter für das gefilterte Kopffolgersignal  
Kopfposition x  
Eingang  
Ausgang

### Patentansprüche

1. Wiedergabeordnung (**100, 300**) zum Schaffen einer interaktiven Bewegungsparallaxe durch Wiedergabe eines Ausgangsbildes auf Basis von visueller dreidimensionaler Information wobei die Wiedergabeordnung (**100, 300**) Folgendes umfasst:  
– erste Empfangsmittel (**101**) zum Empfangen eines ersten Signals (3DV), das die 3D visuelle Information darstellt,  
– zweite Empfangsmittel (**116**) zum Empfangen eines zweiten Signals (P), das Positionsinformation eines Zuschauers des Ausgangsbildes darstellt, als eine Funktion der Zeit, wobei die Positionsinformation gegenüber der Wiedergabeordnung (**100, 300**) rela-

tiv ist,

- Aufbereitungsmittel (118) zum Aufbereiten des Ausgangsbildes;
- Wiedergabemittel (112) zum Wiedergeben des Ausgangsbildes,

**dadurch gekennzeichnet**, dass

- die Anordnung weiterhin Filtermittel (122) zur Hochpassfilterung des zweiten Signals (P) aufweist, was zu einem dritten Signal (PF) führt, das die Änderung der wirklichen Positionsinformation je Zeiteinheit darstellt; und
- dass die Aufbereitungsmittel (118) das Ausgangsbild auf Basis des ersten Signals (3DV) und junger Werte des dritten Signals (PF) derart aufbereiten, dass wenn die Änderung der wirklichen Positionsinformation Null ist, die Aufbereitungsmittel das Ausgangsbild entsprechend einer vorgegebenen Positionsinformation aufbereiten, und wenn die Änderung der wirklichen Positionsinformation relativ groß ist, das Ausgangsbild entsprechend relativ großen Winkeln in Bezug auf das vorgegebene Bild aufbereitet wird.

2. Wiedergabeordnung (100, 300) nach Anspruch 1, wobei die visuelle 3D Information ein Eingangsbild und eine entsprechende Tiefenabbildung aufweist.

3. Wiedergabeordnung (100, 300) nach Anspruch 2, wobei für einen vorbestimmten Wert des dritten Signals (PF) das Eingangsbild und das Ausgangsbild einander im Wesentlichen entsprechen.

4. Wiedergabeordnung (100, 300) nach Anspruch 2, die weiterhin Beschneidemittel (124) aufweist zum Beschneiden des dritten Signals zwischen einer unteren Grenze (404) und einer oberen Grenze (402).

5. Wiedergabeordnung (300) nach Anspruch 2 oder 4, die weiterhin Inhaltsanalysiermittel (128) aufweist, vorgesehen zum Ermitteln eines Maßes eines Satzes von Maßen mit einem ersten Maß entsprechend der Anzahl Unregelmäßigkeiten in der Tiefenabbildung, einem zweiten Maß entsprechend der Homogenität des Eingangsbildes und einem dritten Maß entsprechend der Anzahl Löcher in dem Ausgangsbild, wobei die genannten Inhaltsanalysiermittel (128) dazu vorgesehen sind, die Grenzfrequenz der Filtermittel (122) anzupassen und/oder die untere und/oder obere Grenze der Beschneidemittel (124) zu ändern.

6. Wiedergabeordnung (300) nach Anspruch 5, wobei die Inhaltsanalysiermittel (128) dazu vorgesehen sind, die untere Grenze zu erhöhen und/oder die obere Grenze zu verringern, wenn das erste Maß relativ hoch oder das zweite Maß relativ niedrig ist oder das dritte Maß relativ hoch ist.

7. Wiedergabeordnung (100, 300) nach Anspruch 5, wobei die Inhaltsanalysiermittel (128) dazu vorgesehen sind, die Grenzfrequenz der Filtermittel (124) zu erhöhen, wenn das erste Maß relativ hoch ist oder das zweite Maß relativ niedrig ist oder das dritte Maß relativ hoch ist.

8. Wiedergabeordnung (300) nach Anspruch 1, wobei die Wiedergabeordnung (300) eine Mehrfachbildwiedergabeordnung ist, vorgesehen zum Aufbereiten eines weiteren Ausgangsbildes und zum Wiedergeben des Ausgangsbildes in einer ersten Richtung und zum Wiedergeben des weiteren Ausgangsbildes in einer zweiten Richtung.

9. Verfahren zum Schaffen einer interaktiven Bewegungsparallaxe durch Wiedergabe eines Ausgangsbildes auf Basis von visueller 3D Information wobei das Verfahren Folgendes umfasst:

- das Empfangen eines ersten Signals, das die 3D visuelle Information darstellt,
- das Empfangen eines zweiten Signals, das Positionsinformation eines Zuschauers des Ausgangsbildes darstellt, als eine Funktion der Zeit, wobei die Positionsinformation gegenüber der Wiedergabeordnung (100, 300) relativ ist,
- das Aufbereiten des Ausgangsbildes; und
- das Wiedergeben des Ausgangsbildes, dadurch gekennzeichnet, dass
- das Verfahren weiterhin Hochpassfilterung des zweiten Signals aufweist, was zu einem dritten Signal führt, das die Änderung der wirklichen Positionsinformation je Zeiteinheit darstellt; und dass
- das Ausgangsbild auf Basis des ersten Signals und junger Werte des dritten Signals derart aufbereitet wird, dass wenn die Änderung der wirklichen Positionsinformation Null ist, das Ausgangsbild entsprechend einer vorgegebenen Positionsinformation aufbereitet wird, und wenn die Änderung der wirklichen Positionsinformation relativ groß ist, das Ausgangsbild entsprechend relativ großen Winkeln in Bezug auf das vorgegebene Bild aufbereitet wird.

10. Computerprogrammprodukt, das von einer Computeranordnung geladen werden soll zum Schaffen einer interaktiven Bewegungsparallaxe, mit Instruktionen zum Aufbereiten eines Ausgangsbildes auf Basis von visueller dreidimensionaler Information wobei die Computeranordnung Verarbeitungsmittel und einen Speicher umfasst, wobei das Computerprogrammprodukt nach dem Laden die genannten Verarbeitungsmittel mit der Fähigkeit versieht, Folgendes durchzuführen:

- das Empfangen eines ersten Signals, das die 3D visuelle Information darstellt,
- das Empfangen eines zweiten Signals, das Positionsinformation eines Zuschauers des Ausgangsbildes darstellt, als eine Funktion der Zeit, wobei die Positionsinformation gegenüber der Wiedergabeordnung (100, 300) relativ ist,



- die Hochpassfilterung des zweiten Signals, was zu einem dritten Signal führt, das die Änderung der wirklichen Positionsinformation je Zeiteinheit darstellt; und
- das Aufbereiten des Ausgangsbildes auf Basis des ersten Signals und junger Werte des dritten Signals derart, dass wenn die Änderung der wirklichen Positionsinformation Null ist, das Ausgangsbild entsprechend einer vorgegebenen Positionsinformation aufbereitet wird, und wenn die Änderung der wirklichen Positionsinformation relativ groß ist, das Ausgangsbild entsprechend relativ großen Winkeln in Bezug auf das vorgegebene Bild aufbereitet wird.

Es folgen 5 Blatt Zeichnungen

Anhängende Zeichnungen

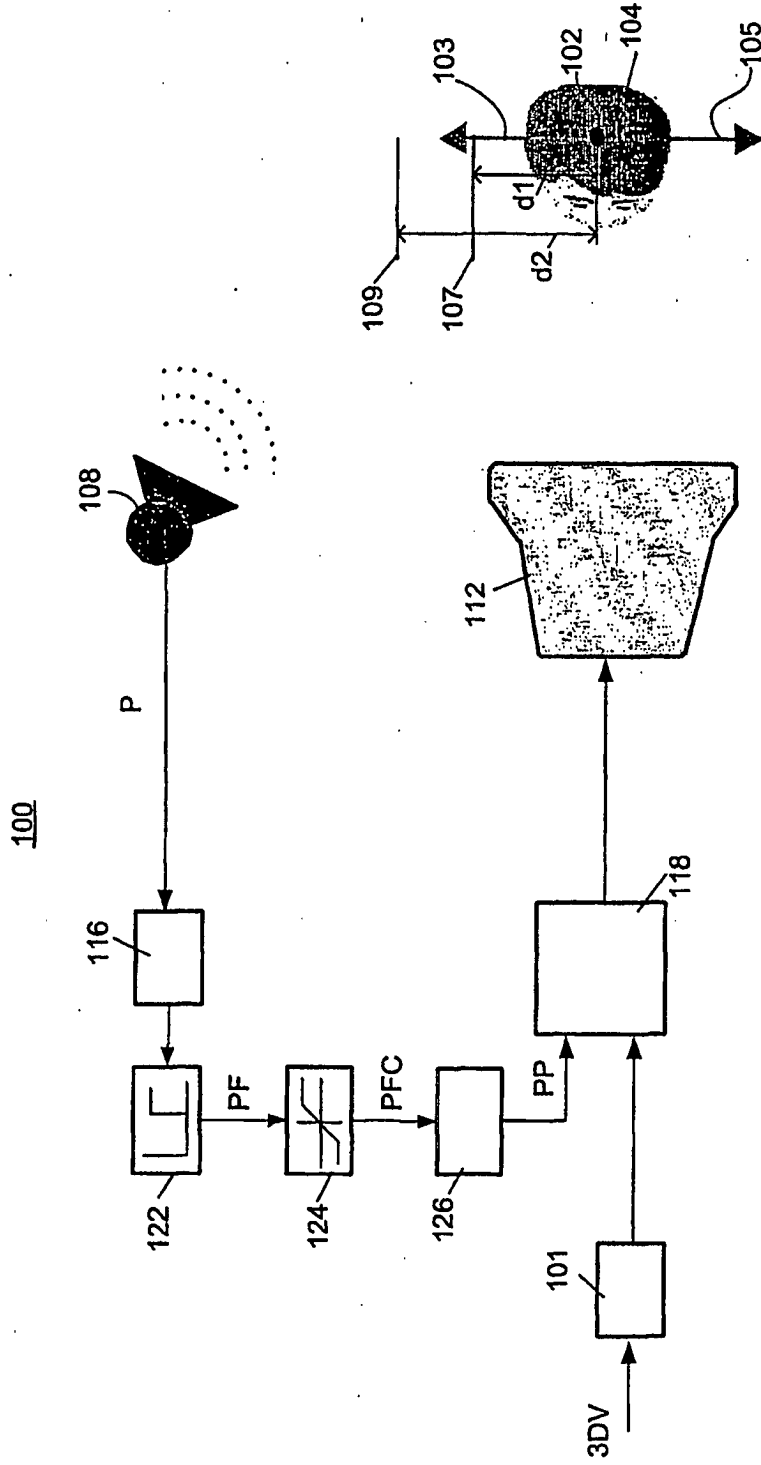


FIG. 1

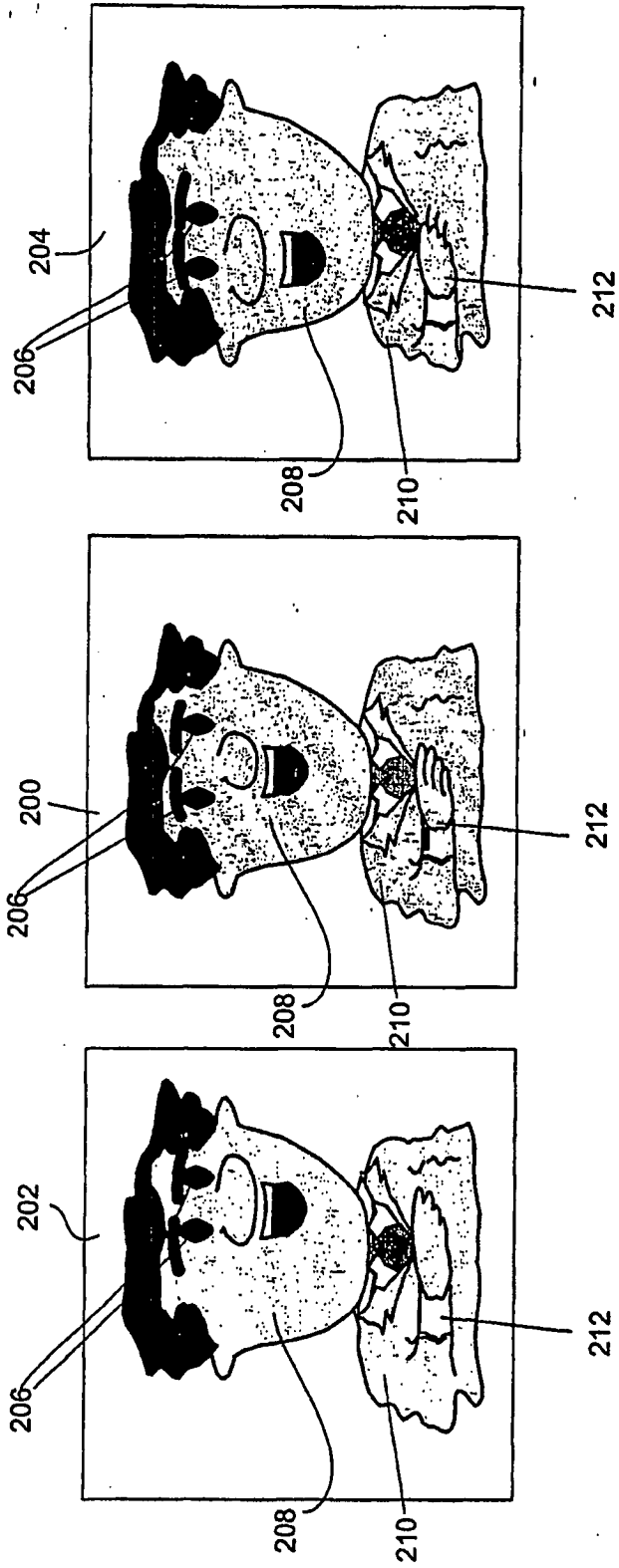


FIG. 2

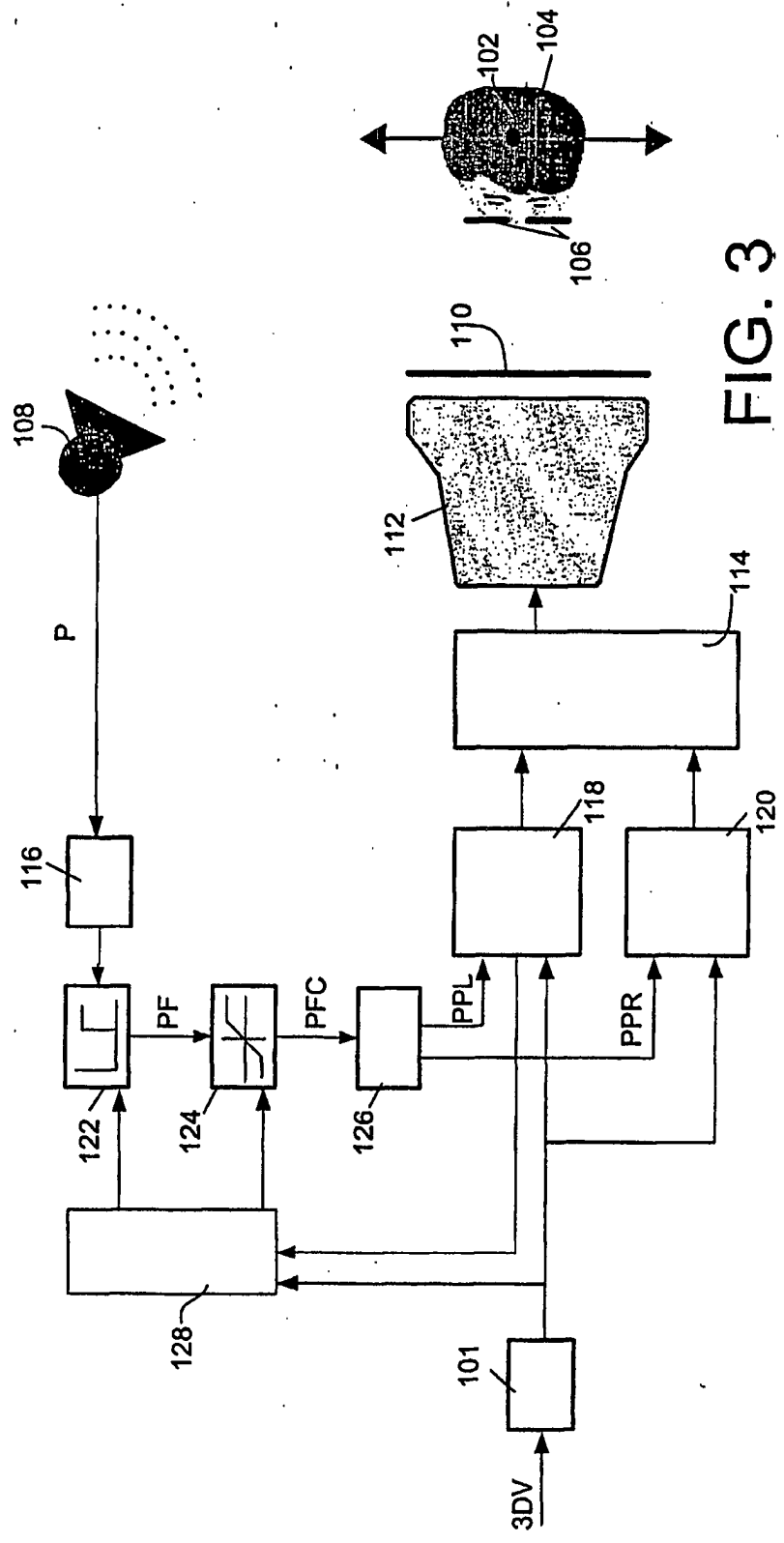


FIG. 3

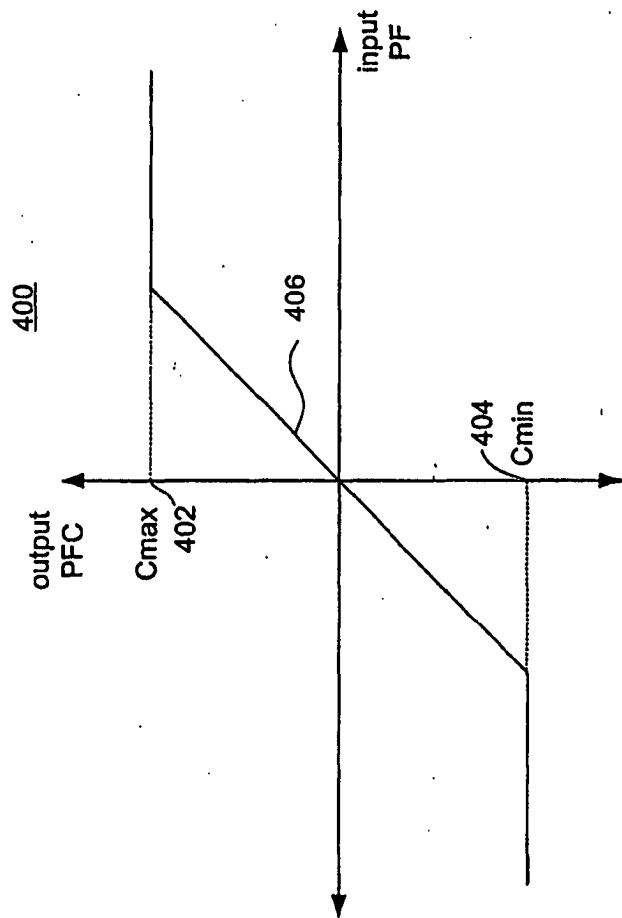


FIG. 4

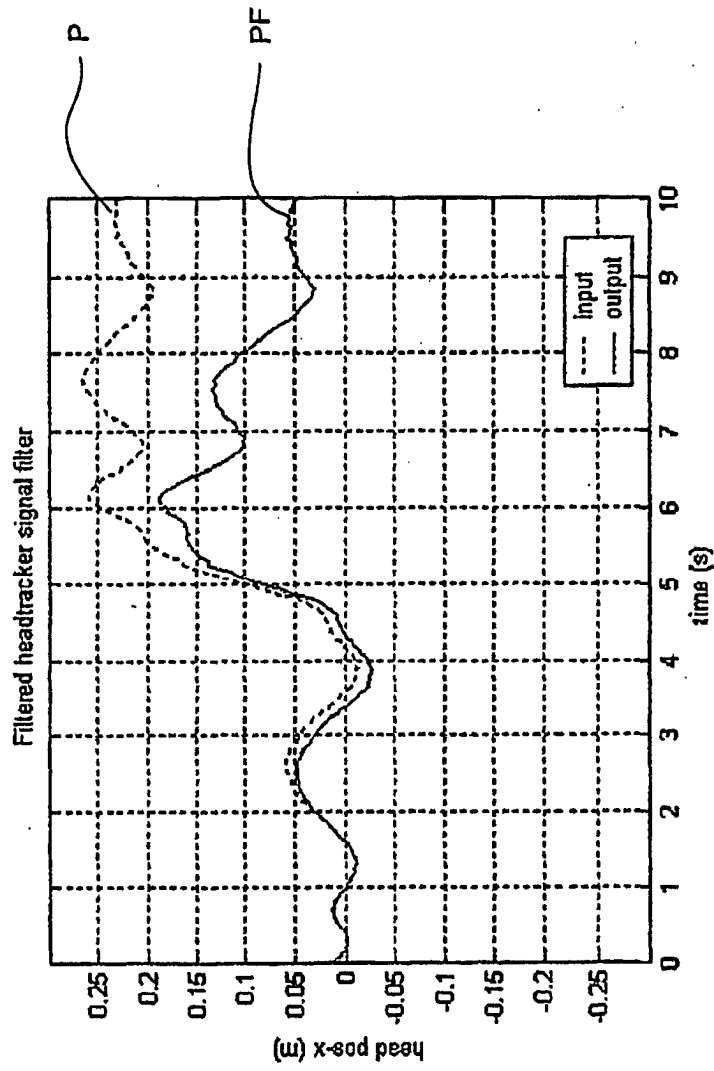


FIG. 5