

(12) **Österreichische Patentanmeldung**

(21) Anmeldenummer: A 50072/2023
(22) Anmeldetag: 07.02.2023
(43) Veröffentlicht am: 15.08.2024

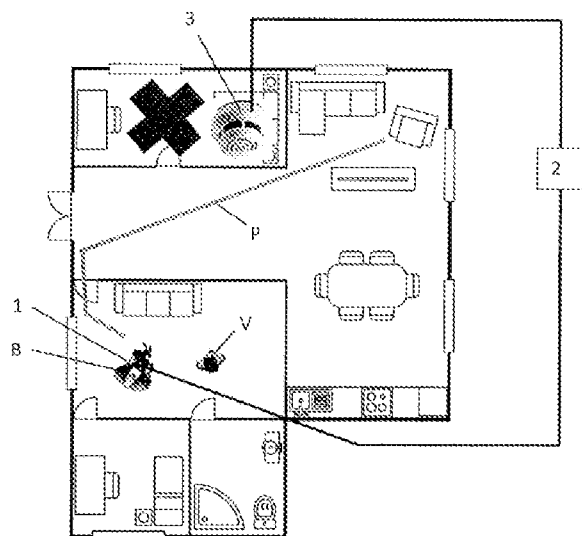
(51) Int. Cl.: **G01C 21/20** (2006.01)
G02B 27/00 (2006.01)
G06F 3/01 (2006.01)
G06F 3/04815 (2022.01)
G06V 20/20 (2022.01)
G06T 17/00 (2006.01)

(71) Patentanmelder:
Peterseil Thomas
4052 Ansfelden (AT)

(74) Vertreter:
Kliment & Henhapel Patentanwaelte OG
1010 Wien (AT)

(54) **VERFAHREN UND VORRICHTUNG ZUM EINBLENDEN EINES VIRTUELLEN OBJEKTS**

(57) Verfahren zum Einblenden eines computergenerierten, virtuellen Objektes (V) in das von einem Benutzer (B) wahrgenommene Sichtfeld einer realen Umgebung, wobei das Einblenden des virtuellen Objektes (V) mittels einer zentralen Datenverarbeitungseinheit (2) in Abhängigkeit von der momentanen Position des Benutzers (B) sowie von der Anwesenheit von Sicht- oder Bewegungshindernissen erfolgt, wobei die Sicht- oder Bewegungshindernisse von der zentralen Datenverarbeitungseinheit (2) anhand eines digitalen Abbilds von Oberflächen der realen Umgebung ermittelt werden. Es wird vorgeschlagen, dass 3D-Bilder angefertigt werden, aus denen von der zentralen Datenverarbeitungseinheit (2) in vorgegebenen Zeitabständen wiederholt ein aktualisiertes digitales Abbild von Oberflächen der realen Umgebung im Sichtfeld ermittelt wird, und mittels Sensoren (3) sichtfeldbezogene Zustandsdaten gewonnen werden, die in der zentralen Datenverarbeitungseinheit (2) verarbeitet werden, wobei mit Hilfe des virtuellen Objektes (V) eine für den Benutzer (B) interaktive und dialogbasierte Wiedergabe der sensorgestützten, sichtfeldbezogenen Zustandsdaten im Sichtfeld des Benutzers (B) erfolgt.



Zusammenfassung:

Verfahren und Vorrichtung Verfahren zum Einblenden eines computergenerierten, virtuellen Objektes (V) in das von einem Benutzer (B) wahrgenommene Sichtfeld einer realen Umgebung, wobei das Einblenden des virtuellen Objekts (V) mittels einer zentralen Datenverarbeitungseinheit (2) in Abhängigkeit von der momentanen Position des Benutzers (B) sowie von der Anwesenheit von Sicht- oder Bewegungshindernissen erfolgt, wobei die Sicht- oder Bewegungshindernisse von der zentralen Datenverarbeitungseinheit (2) anhand eines digitalen Abbilds von Oberflächen der realen Umgebung ermittelt werden. Es wird vorgeschlagen, dass 3D-Bilder angefertigt werden, aus denen von der zentralen Datenverarbeitungseinheit (2) in vorgegebenen Zeitabständen wiederholt ein aktualisiertes digitales Abbild von Oberflächen der realen Umgebung im Sichtfeld ermittelt wird, und mittels Sensoren (3) sichtfeldbezogene Zustandsdaten gewonnen werden, die in der zentralen Datenverarbeitungseinheit (2) verarbeitet werden, wobei mithilfe des virtuellen Objekts (V) eine für den Benutzer (B) interaktive und dialogbasierte Wiedergabe der sensorgestützten, sichtfeldbezogenen Zustandsdaten im Sichtfeld des Benutzers (B) erfolgt.

(Fig. 6)

Die Erfindung betrifft ein Verfahren zum Einblenden eines computergenerierten, virtuellen Objektes in das von einem Benutzer wahrgenommene Sichtfeld einer realen Umgebung, insbesondere in das Sichtfeld einer vom Benutzer getragenen, transparenten Wiedergabeeinheit, wobei das Einblenden des virtuellen Objekts mittels einer zentralen Datenverarbeitungseinheit in Abhängigkeit von der momentanen Position und der Blickrichtung des Benutzers erfolgt, und das Einblenden des virtuellen Objekts zusätzlich in Abhängigkeit von der Anwesenheit von Sicht- oder Bewegungshindernissen erfolgt, wobei die Sicht- oder Bewegungshindernisse von der zentralen Datenverarbeitungseinheit anhand eines digitalen Abbilds von Oberflächen der realen Umgebung ermittelt werden, gemäß dem Oberbegriff von Anspruch 1, sowie eine Vorrichtung zur Durchführung des erfindungsgemäßen Verfahrens nach Anspruch 7.

Verfahren dieser Art werden in so genannten „Mixed-Reality (MR)“-Systemen oder „Augmented Reality (AR)“-Systemen verwendet. Hierbei erfolgt das Einblenden eines computergenerierten, virtuellen Objektes in das vom Benutzer wahrgenommene Blickfeld einer realen Umgebung in Abhängigkeit von der momentanen Position und Blickrichtung des Benutzers, für deren Bestimmung in herkömmlicher Weise unterschiedliche Sensorik wie Beschleunigungssensoren (Accelerometer oder G-Sensorik), mitunter auch in Kombination mit Magnetometern, Kameras und/oder Gyroskopen, sowie GPS-Systeme verwendet werden. Die momentane Position und Blickrichtung des Benutzers entscheiden zunächst darüber, ob überhaupt ein virtuelles Objekt eingeblendet wird. Das Einblenden des virtuellen Objekts in seiner Positionierung und Größe im Blickfeld des Benutzers wird in weiterer Folge maßgeblich in Abhängigkeit von der momentanen Position des beobachtenden Benutzers, insbesondere in Abhängigkeit vom Abstand zum beobachtenden Benutzer vorgenommen, um die realistisch wirkende Illusion eines virtuellen Abstandes zu erzeugen. Hierfür werden insbesondere Wiedergabeeinheiten in Form von Datenbrillen verwendet, also Vorrichtungen, die wie eine Brille getragen werden und in der Lage sind virtuelle Objekte in der realen Welt anzuzeigen und diese Objekte in Relation zu realen

Objekten im Raum zu positionieren. Solche Datenbrillen werden auch als „Augmented-Reality“-Brillen oder „Mixed-Reality“-Brillen bezeichnet. Ein Benutzer, der eine solche Brille aufgesetzt hat, sieht die Umgebung wie durch eine normale Sehbrille, allerdings können ihm (holographisch virtuelle) Objekte in das Sichtfeld eingeblendet werden. Alternativ ist auch die Verwendung von Linsen denkbar, die als Wiedergabeeinheit das Einblenden von virtuellen Objekten in das Blickfeld des Benutzers erlauben.

Ein möglicher Anwendungsbereich besteht etwa in der Ausbildung und im Training von Einsatzkräften der Polizei, der Feuerwehr oder des Militärs. Die Einsatzkräfte bewegen sich dabei bei unterschiedlichen Lichtverhältnissen und mitunter auch unter akustischer Beschallung durch ein Einsatzgebiet und müssen auf unterschiedliche Szenarien richtig reagieren, die durch das Einblenden virtueller Objekte verwirklicht werden. So können etwa im Rahmen einer Ausbildung Einsatzkräfte der Polizei darin geschult werden ein Gebäude nach Tätern zu durchsuchen, wobei die Täter von einem Ausbildungsleiter als virtuelle Objekte hinter Wänden, Türen, Möbelstücken und dergleichen positioniert werden und je nach Position und Blickrichtung des Auszubildenden sichtbar sind oder verdeckt. Im Rahmen einer Ausbildung von Einsatzkräften der Feuerwehr kann das Training etwa darin bestehen einen virtuellen Brandherd in einem Gebäude zu bekämpfen und dabei Gefahren durch Rauchgase, Rauchgasexplosionen oder einstürzende Gebäudeteile zu erkennen und entsprechendes Verhalten zu trainieren. Im Rahmen einer Ausbildung von Einsatzkräften des Militärs kann das Vorgehen gegen feindliche Truppen trainiert werden, die als virtuelle Objekte in Form vorrückender Truppenteile wie Panzer und dergleichen oder auch feindlicher Soldaten eingeblendet werden.

In der US 2002/0196202 A1 und der US 2003/0210228 A1 wird beispielsweise ein Verfahren beschrieben, bei dem Einsatzkräften, die sich in einer Einsatzsituation in einem Gebäudeinneren bewegen, anhand eines dem Einsatzleiter als digitales Abbild vorliegenden Gebäudeplans Informationen in eine Datenbrille eingeblendet werden. Solche Informationen

können etwa Warnmeldungen über mögliche Gefahren sein, oder auch Navigationshilfen in Form von Pfeilen und dergleichen. Ein Einblenden von virtuellen Objekten unter Berücksichtigung einer möglichen Verdeckung durch bauliche Gegebenheiten oder Möbelstücke ist bei einem Verfahren gemäß der US 2002/0196202 A1 und der US 2003/0210228 A1 nicht möglich. In der US 2002/0191004 A1 wird ein weiteres Verfahren zum Einblenden virtueller Objekte zur Ausbildung von Einsatzkräften beschrieben.

In der US 2018/0330184 A1 wird ein Verfahren zur Erstellung eines digitalen Abbilds von Oberflächen eines Gebäudeinneren beschrieben, wobei das Gebäudeinnere mit einem Sensor zur Abstandsmessung vermessen wird. Das Ergebnis der Abstandsmessungen ist eine dreidimensionale Datenpunktwolke, aus der durch Anwendung mathematischer Verfahren der Verlauf von Böden, Decken und Wänden der Innenräume rekonstruiert wird. Ein Einblenden virtueller Objekte ist in der US 2018/0330184 A1 nicht vorgesehen.

In der US 2022/108535 A1 werden Verfahren zum Einblenden eines virtuellen Objekts in Form eines Avatars in das Sichtfeld einer „Augmented-Reality“-Brille beschrieben. In der US 2022/130094 A1 werden Verfahren zur Animation eines Avatars beschrieben, insbesondere zur Steuerung von Bewegungen des Avatars auf Basis eines Datennetzes aus Polygonen ("Navigation mesh", "navmesh"). Die WO 2022/004422 A1 beschäftigt sich mit dem Einblenden von virtuellen Objekten in das Sichtfeld einer „Augmented-Reality“-Brille, wobei externe Daten einbezogen werden, um das Einblenden des virtuellen Objekts realistischer zu gestalten. In der US 2021/038975 A1 wird ein System beschrieben, bei dem Zuschauer mit entsprechenden portablen Endgeräten an AR-Spielen anderer Teilnehmer teilhaben können. In der KR 20210069806 A wird ein Verfahren zur Analyse von Kamerabildern hinsichtlich Raum- und Objekterkennung auf Basis neuronaler Netzwerke vorgeschlagen, wobei die so gewonnenen Raum- und Objektdaten zur verbesserten Einblendung eines virtuellen Objekts verwendet werden.

In der österreichischen Patentanmeldung AT 523.953 des Anmelders sowie der korrespondierenden europäischen

Patentanmeldung EP 21735597.3 des Anmelders wurde ein Verfahren zum Einblenden eines computergenerierten, virtuellen Objektes in das von einem Benutzer wahrgenommene Sichtfeld eines realen Gebäudeinneren vorgeschlagen, bei dem das virtuelle Objekt gemäß der momentanen Sichtachsen des Benutzers unter Berücksichtigung verdeckender Elemente wie Wände, Türe, Raumteiler oder auch Möbel korrekt eingeblendet werden kann. Hierbei erfolgt das Einblenden des virtuellen Objekts zusätzlich in Abhängigkeit von der Anwesenheit eines Sicht- oder Bewegungshindernisses für den Benutzer, wobei das Sicht- oder Bewegungshindernis anhand eines aus planaren Flächenelementen gebildeten, digitalen Abbilds von Oberflächen des Gebäudeinneren ermittelt wird. Eine Möglichkeit besteht darin, dass ein solches digitales Abbild bereits in Form eines digitalen Gebäudeplans existiert, wobei der digitale Gebäudeplan neben Wänden, Raumteiler oder Türen auch bereits die Positionierung von Möbelstücken umfasst. Eine weitere Möglichkeit zur Erstellung eines digitalen Abbilds des Gebäudeinneren besteht darin, dass es durch eine Vermessung des Gebäudeinneren mit einer zur Messung von Tiefeninformation fähigen Kamera ermittelt wird. In diesem Fall kann die zur Messung von Tiefeninformation fähige Kamera an der Wiedergabeeinheit montiert sein, mit der das Gebäudeinnere durchschritten wird. Eine zur Messung von Tiefeninformation fähige Kamera wird auch als 3D-Kamera bezeichnet und kann etwa als ToF ("Time-of-Flight")-Kamera ausgeführt sein. Bei ToF-Kameras handelt es sich um Kameras, die nicht nur ein 2D-Bild aufnehmen, sondern für jedes Aufnahmepixel auch eine Tiefeninformation messen. Unter Tiefeninformation werden Informationen über die Distanzen zwischen den einzelnen Objekten einer Szene und der ToF-Kamera verstanden. Mithilfe bekannter mathematischer Verfahren kann aus den so gewonnenen „Scans“ an Tiefeninformationen ein digitales Abbild gewonnen werden, das die Oberflächen des Gebäudeinneren mithilfe von Flächenelementen approximiert. Diese Flächenelemente können etwa als planare Flächenelemente in Form von Dreiecken, Vierecken oder Sechsecken vorliegen, die die zu approximierende Oberfläche lückenlos nachbilden. Die gemeinsamen Eckpunkte dieser Flächenelemente werden dabei auch

als Knotenpunkte („vertices“) bezeichnet, die dem Betrachter zugewandte Seite der Flächenelemente als „faces“ und die Gesamtheit der Knotenpunkte auch als „mesh“. Eine weitere Möglichkeit bekannter mathematischer Verfahren besteht darin den Flächenelementen auch eine Orientierung zuzuschreiben, indem eine Flächennormale auf das betreffende Flächenelement errechnet wird. Das oben genannte Verfahren der AT 523.953 und der EP 21735597.3 bedient sich dieser Möglichkeiten, indem zunächst ein aus planaren Flächenelementen gebildetes digitales Abbild von Oberflächen des Gebäudeinneren ermittelt wird. In weiterer Folge werden jedoch nur jene Flächenelemente herangezogen, die zwischen einer Bodenebene des Gebäudeinneren und einer angenommenen Augenhöheebene liegen, die eine Maximalhöhe für die nachfolgenden Verfahrensschritte definiert. Unter einer Augenhöheebene wird dabei in der Regel eine horizontale Ebene verstanden, die sich in Augenhöhe befindet. Die genaue Höhenlage ist dabei nicht wesentlich, sie wird beispielsweise mit einer Höhe von 1,80m angenommen. Das digitale Abbild des Gebäudeinneren wird somit gewissermaßen in Augenhöheebene „abgeschnitten“. Des Weiteren werden jene Flächenelemente ermittelt, deren Flächennormale horizontal oder vertikal verlaufen, also Flächenelemente, deren Orientierung vertikal oder horizontal ist. Zudem wird eine horizontale Schichtung zwischen der Bodenebene des Gebäudeinneren und der angenommenen Augenhöheebene mit einer Mehrzahl von Schichtlagen einer vorgegebenen Schichtdicke vorgenommen.

Die Kategorisierung der Flächenelemente und die Vornahme einer Schichtung werden in weiterer Folge zur Ermittlung von Sicht- oder Bewegungshindernissen verwendet, indem zum Einen vertikale Flächenelemente, die Teil des digitalen Abbilds sind, als Sicht- oder Bewegungshindernisse interpretiert werden. Diese Sicht- oder Bewegungshindernisse werden in der Regel Wände, Raumteiler oder hohe Möbelstücke wie Schränke und dergleichen repräsentieren. Zum Anderen werden horizontale Flächenelemente oder eine Gesamtheit aneinandergrenzender horizontaler Flächenelemente in derselben Schichtlage als Sicht- oder Bewegungshindernisse ermittelt, falls in deren Umfangsbereich horizontale Nebenflächenelemente auffindbar

sind, die um mehr als eine Schichtlage unterhalb des horizontalen Flächenelements oder der Gesamtheit aneinandergrenzender horizontaler Flächenelemente liegen, oder in anderen Worten, wenn in deren Umfangsbereich keine horizontalen Nebenflächenelemente auffindbar sind, die maximal eine Schichtlage unterhalb oder oberhalb des horizontalen Flächenelements oder der Gesamtheit aneinandergrenzender horizontaler Flächenelemente liegen. Diese Sicht- oder Bewegungshindernisse werden in der Regel niedrige Möbelstücke wie Tische und dergleichen repräsentieren.

Mithilfe der so ermittelten Flächenelemente des digitalen Abbilds als Sicht- oder Bewegungshindernisse kann das Einblenden eines virtuellen Objekts in Abhängigkeit von der Anwesenheit des Sicht- oder Bewegungshindernisses auf einfache Weise erfolgen, indem im Rahmen eines Berechnungsalgorithmus Verbindungslinien zwischen einem der momentanen Position und Blickrichtung entsprechenden Augpunkt des Benutzers und Bildpunkten des virtuellen Objekts errechnet und innerhalb des Abstandes zwischen dem Augpunkt und den Bildpunkten des virtuellen Objekts das Vorliegen von Schnittpunkten der Verbindungslinien mit den als planare Flächenelemente vorliegenden Sicht- oder Bewegungshindernissen ermittelt werden, wobei bei Vorliegen eines Schnittpunkts für einen Bildpunkt des virtuellen Objekts der betreffende Bildpunkt im Sichtfeld des Benutzers nicht oder nicht sichtbar eingeblendet wird, und bei Fehlen eines Schnittpunkts für einen Bildpunkt des virtuellen Objekts der betreffende Bildpunkt sichtbar eingeblendet wird. Der Augpunkt ist ein Begriff aus der Zentralperspektivität und bezeichnet in herkömmlicher Weise einen Punkt im Raum, der der Position eines Benutzers entspricht und dem die „Sehstrahlen“ entspringen. Zwischen diesem Augpunkt und Bildpunkten des darzustellenden, virtuellen Objektes werden zunächst Verbindungslinien errechnet. In weiterer Folge wird ermittelt, ob innerhalb des Abstandes zwischen dem Augpunkt und den Bildpunkten des virtuellen Objekts Schnittpunkte der Verbindungslinien mit einem als planares Flächenelement vorliegenden Sicht- oder Bewegungshindernis vorliegen. Es wird somit geprüft, ob die Verbindungslinie eines Bildpunktes mit dem Augpunkt ein als

Sicht- oder Bewegungshindernis bewertetes Flächenelement schneidet. Im zutreffenden Fall wird der betreffende Bildpunkt des virtuellen Objekts im Sichtfeld des Benutzers nicht oder nicht sichtbar eingeblendet. Eine nicht sichtbare Einblendung kann etwa über eine transparente Einblendung erfolgen, bei der der betreffende Bildpunkt für den beobachtenden Benutzer ebenfalls nicht sichtbar wird. Falls kein solcher Schnittpunkt vorliegt, wird der betreffende Bildpunkt sichtbar eingeblendet. Der Berechnungsvorgang beginnend mit der Errechnung der Verbindungslinie und der Klärung, ob ein Schnittpunkt vorliegt, benötigt pro Bildpunkt nur wenige Millisekunden. Der Berechnungsvorgang kann somit rasch genug abgeschlossen werden, um Echtzeit-Anwendungen und eine oftmalige Wiederholung im Zuge der Bewegung des beobachtenden Benutzers durch das Gebäudeinnere zu ermöglichen.

Im Rahmen eines solchen Verfahrens wurde vom Anmelder aufgrund der als planare Flächenelemente des digitalen Abbilds vorliegenden Sicht- oder Bewegungshindernisse eine Anwendung zur Schulung von Einsatzkräften vorgeschlagen, indem virtuelle Objekte wie beispielsweise Darstellungen bewaffneter Täter, die sich hinter Sichthindernissen verbergen, dem Benutzer in das Sichtfeld der Wiedergabeeinheit in Abhängigkeit von der Anwesenheit eines Sicht- oder Bewegungshindernisses eingeblendet werden. Die Wiedergabe virtueller Objekte kann durch das Einblenden von Navigationshilfen begleitet sein, die auf Basis des vorliegenden digitalen Abbilds den Einsatzkräften etwa den schnellsten Weg zu einem Zielpunkt unter Berücksichtigung von Bewegungshindernissen anzeigen, oder bei schlechten Sichtverhältnissen eine sichere Bewegung im Gebäudeinneren ermöglichen.

Das sichere Bewegen in einem Einsatzgebiet erfordert in der Praxis aber nicht nur das Auffinden eines bestimmten Zieles, sondern auch die Verarbeitung einer Fülle an situationsabhängigen Informationen, die in weiterer Folge auch als sichtfeldbezogene Zustandsdaten bezeichnet werden. Dabei handelt es sich etwa im Falle von Einsatzkräften der Feuerwehr beispielsweise um physikalische Zustandsdaten wie Temperatur oder Sichtweite am Einsatzort, um chemische Zustandsdaten wie

Schadstoffkonzentrationen oder Brandgeruch, um bauliche Zustandsdaten etwa zu Eigenschaften von Baumaterialien oder zu statischen Gegebenheiten am Einsatzort, oder bei Einsätzen im Freien beispielsweise auch um meteorologische Zustandsdaten. Die hier als Zustandsdaten bezeichneten Informationen werden in herkömmlicher Weise zumeist durch bloße Beobachtung oder auch Erfahrung der Einsatzkräfte gewonnen. Mitunter werden für spezifische Messungen auch Messgeräte von den Einsatzkräften mitgeführt. Für einen sicheren und effizienten Einsatz wäre es freilich wünschenswert über ein Maximum an solchen Zustandsdaten zu verfügen. Andererseits dürfen Einsatzkräfte durch eine Fülle an zu verarbeitenden Informationen auch nicht überfordert werden. Ein Überangebot an ungefilterter Information kann sogar hinderlich für den Einsatz sein.

Es besteht daher das Ziel der Erfindung darin unter Nutzung von „Augmented-Reality“-Systemen Verfahren und Vorrichtungen zu schaffen, mit denen beispielsweise im Rahmen eines sicherheitskritischen Einsatzes ein Maximum an situationsbezogenen Daten gewonnen und einem Benutzer so benutzerfreundlich wie möglich bereitgestellt werden kann.

Dieses Ziel wird durch die Merkmale von Anspruch 1 und Anspruch 7 erreicht. Anspruch 1 bezieht sich auf ein Verfahren zum Einblenden eines computergenerierten, virtuellen Objektes in das von einem Benutzer wahrgenommene Sichtfeld einer realen Umgebung, insbesondere in das Sichtfeld einer vom Benutzer getragenen, transparenten Wiedergabeeinheit, wobei das Einblenden des virtuellen Objekts mittels einer zentralen Datenverarbeitungseinheit in Abhängigkeit von der momentanen Position und der Blickrichtung des Benutzers erfolgt, und das Einblenden des virtuellen Objekts zusätzlich in Abhängigkeit von der Anwesenheit von Sicht- oder Bewegungshindernissen erfolgt, wobei die Sicht- oder Bewegungshindernisse von der zentralen Datenverarbeitungseinheit anhand eines digitalen Abbilds von Oberflächen der realen Umgebung ermittelt werden. Erfindungsgemäß wird dabei vorgeschlagen, dass mittels einer zur Messung von Tiefeninformation fähigen Kamera von dem der momentanen Position und Blickrichtung entsprechenden Sichtfeld des Benutzers 3D-Bilder angefertigt werden, aus denen von der

zentralen Datenverarbeitungseinheit mithilfe bekannter mathematischer Verfahren in vorgegebenen Zeitabständen wiederholt ein aktualisiertes digitales Abbild von Oberflächen der realen Umgebung und somit der Sicht- oder Bewegungshindernisse im Sichtfeld ermittelt wird, und mittels Sensoren sichtfeldbezogene Zustandsdaten gewonnen werden, die in der zentralen Datenverarbeitungseinheit verarbeitet werden, wobei mithilfe des virtuellen Objekts eine für den Benutzer interaktive und dialogbasierte Wiedergabe der sensorgestützten, sichtfeldbezogenen Zustandsdaten im Sichtfeld des Benutzers erfolgt. Im Rahmen der Erfindung werden somit zunächst sensorgestützte, sichtfeldbezogene Zustandsdaten gewonnen, also Zustandsdaten, die für das jeweilige Sichtfeld relevant sind und mittels Sensoren gewonnen werden. Diese Sensoren können beispielsweise für einen sicherheitskritischen Einsatz von Einsatzkräften eigens mitgeführt oder installiert werden, oder es wird auf bereits vorhandene Sensoren zur Überwachung von Örtlichkeiten wie Brandmelder, Bewegungssensoren, Sicherheitskameras und dergleichen zurückgegriffen. Bei diesen Sensoren kann es sich auch um luftgestützte Sensoren handeln, bei denen etwa mithilfe von Drohnen sichtfeldrelevante Daten gesammelt werden, beispielsweise im Rahmen einer Luftaufklärung in militärischen Anwendungen, oder um bodengestützte Sensoren wie beispielsweise Überwachungskameras einer vorhandenen Sicherheitsarchitektur. Diese Zustandsdaten werden in weiterer Folge im Sichtfeld des Benutzers mithilfe des virtuellen Objekts in Form einer für den Benutzer interaktiven und dialogbasierten Wiedergabe bereitgestellt. Die erfindungsgemäße Wiedergabe der Zustandsdaten ist einerseits *interaktiv* ausgeführt, also als Wiedergabe, die eine wechselseitige Bezugnahme auf Ein- und Ausgaben von Daten und Befehlen von Benutzer und zentraler Datenverarbeitungseinheit ermöglicht, und andererseits *dialogbasiert*, also durch sprachlichen Ausdruck in schriftlicher oder mündlicher Form. Diese dialogbasierte, interaktive Wiedergabe wird erfindungsgemäß mithilfe des virtuellen Objekts ermöglicht, das dem Benutzer in sein Sichtfeld eingeblendet wird und das ihm somit den Zugriff auf die von den externen Sensoren

gewonnenen und von der zentralen Datenverarbeitungseinheit verarbeiteten und beispielsweise als Datenbanken aufbereiteten Zustandsdaten ermöglicht. Durch den dialogbasierten Zugriff kann der Benutzer beispielsweise mithilfe einer Sprachsteuerung gewünschte sichtfeldbezogene Zustandsdaten abrufen, also etwa durch die mündlich gestellte Frage „*Wie hoch ist die momentane Kohlenmonoxid-Konzentration?*“ oder einfach „*Kohlenmonoxid?*“. Dieser Sprachbefehl wird von einem Mikrofon aufgenommen, das der Benutzer trägt, und über das der mündlich geäußerte Eingabebefehl der zentralen Datenverarbeitungseinheit übermittelt wird. Die zentrale Datenverarbeitungseinheit ruft Messdaten eines entsprechenden Messgeräts ab und gibt sie über das virtuelle Objekt im Sichtfeld des Benutzers wieder. Die Wiedergabe kann etwa optisch im Sichtfeld des Benutzers über eine dialogbasierte schriftliche Wiedergabe von Zustandsdaten erfolgen, wie sie von Chatbots und dergleichen bekannt ist, oder akustisch über ein vom Benutzer mitgetragenes Wiedergabegerät wie Kopfhörer und dergleichen, wobei die akustische Wiedergabe über eine vom Benutzer dem virtuellen Objekt zugeordnete sprachliche Ausgabe erfolgt. Im erstgenannten Fall stellt das virtuelle Objekt selbst eine schriftliche oder sonstige graphische Darstellung der wiedergegebenen Zustandsdaten dar, und im zweitgenannten Fall erfolgt die Wiedergabe der Zustandsdaten in einer vom virtuellen Objekt unterstützten Form, indem der Benutzer die Wiedergabe dem virtuellen Objekt zuordnet und der Benutzer das virtuelle Objekt somit als Ausgabeschnittstelle wahrnimmt. Diese Zuordnung kann etwa durch geeignete graphische Aufbereitung des virtuellen Objekts erfolgen. Eine Möglichkeit besteht etwa darin die dialogbasierte, interaktive Wiedergabe mithilfe eines virtuellen Objekts mit einem antropomorphen Aussehen zu unterstützen. Ein solches computergeneriertes Objekt mit einem antropomorphen Aussehen, das zu einer dialogbasierten Interaktion mit dem Benutzer fähig ist, wird mitunter auch als Avatar bezeichnet. In der Regel werden dabei Avatare nicht statisch dargestellt, sondern als Bewegtbild animiert, wofür in der Regel so genannte „3D-Echtzeit-Engines“ verwendet werden. Auf diese Weise lassen sich Interaktionen auch optisch darstellen, wodurch für den Benutzer der Eindruck

einer Interaktion verstärkt wird. Der Eindruck der Interaktion kann dabei auf vielfältige Weise erreicht werden, beispielsweise durch eine dem virtuellen Objekt zugeordnete Sprachfärbung, oder durch Augenkontakt eines antropomorph ausgeführten virtuellen Objekts mit dem Benutzer und entsprechende Mundbewegungen. Das virtuelle Objekt wird somit vom Benutzer als Benutzerschnittstelle wahrgenommen, auch wenn der physische Datenaustausch zwischen dem Benutzer und der zentralen Datenverarbeitungseinheit mithilfe eines Mikrofons und eines Lautsprechers, oder auch mithilfe von Gestenerkennung und dergleichen erfolgen kann. Die Benutzerschnittstelle kann in diesem Sinn auch als virtuelle Benutzerschnittstelle bezeichnet werden.

Die Erfindung sieht des Weiteren vor die Gewinnung und Wiedergabe der sensorgestützten Zustandsdaten mit iterativ präzisierten räumlichen Daten des Sichtfeldes zu kombinieren. Diese Maßnahme dient einerseits dazu ein aktualisiertes digitales Abbild von Oberflächen der realen Umgebung und somit der Sicht- oder Bewegungshindernisse im Sichtfeld des Benutzers zu erhalten, weil sich die physische Umgebung vor allem unter realen Einsatzbedingungen von Einsatzkräften ständig ändern kann. Die automatisiert wiederholte Erhebung räumlicher Daten und das damit verbundene Erfassen von Gefahrenquellen und von Bewegungshindernissen sind für eine genaue Lagebilderstellung entscheidend. Diese Daten werden an die zentrale Datenverarbeitungseinheit übermittelt, die diese räumlichen Daten des Sichtfeldes mit den sichtfeldbezogenen Zustandsdaten kombiniert und auswertet. Vorzugsweise erfolgt die Bewertung derselben durch die zentrale Datenverarbeitungseinheit automatisiert, da im Echteinsatz keine Zeit besteht diese Daten manuell in ein Lagebild einzupflegen, und der Einsatz autonomer Systeme und automatischer Detektion die Risiken für Infrastrukturen und Einsatzteams verringert. Die Erfindung ermöglicht es somit durch die Kombination von multimodalen Daten von unterschiedlichen Sensorsystemen mit den in Echtzeit erhobenen räumlichen Daten eine automatische Lagebilderstellung vorzunehmen, wie noch näher ausgeführt werden wird. Wie die praktische Erfahrung zeigt, ist es

mitunter erst durch eine Kombination der vor Ort in Echtzeit gewonnenen Daten wie persönliche Wahrnehmung oder 3D-Bildern der Szenerie mit Daten externer Sensoren oder vorhandenen digitalen Plänen möglich, ein korrektes Lagebild inklusive aller erkannten Gefahren und Hindernissen zu erstellen. Diese Daten dienen als Basis für Entscheidungen der zentralen Datenverarbeitungseinheit, um Einsatzkräfte vor Gefahren zu schützen, Bewegungsrouten zu berechnen oder Anweisungen geben zu können.

Andererseits dient die Maßnahme der iterativ präzisierten räumlichen Daten des Sichtfeldes aber auch dazu die Wiedergabe des virtuellen Objekts zu optimieren, vor allem wenn es sich dabei um ein virtuelles Objekt mit antropomorpher Gestalt handelt, das im Sichtfeld des Benutzers bewegt animiert wird und von Sichthindernissen stets unterschiedlich verdeckt wird. Dabei ist darauf hinzuweisen, dass erfindungsgemäß die wiederholte Messung von Tiefeninformation und die Anfertigung von 3D-Bildern zur Gewinnung eines aktualisierten digitalen Abbildes von Oberflächen der realen Umgebung von dem der momentanen Position und Blickrichtung entsprechenden Sichtfeld des Benutzers erfolgt. Das Sichtfeld des Benutzers ist deshalb entscheidend, weil der Benutzer naturgemäß seine Aufmerksamkeit und damit sein Sichtfeld auf die für ihn momentan relevante Szenerie richtet und eine genaue Analyse des Sichtfeldes daher besonders wichtig ist. Zudem erfolgt auch das Einblenden des virtuellen Objekts für die Wiedergabe von Zustandsdaten im Sichtfeld des Benutzers, sodass die genaue Kenntnis der Sicht- und Bewegungshindernisse für eine optimale Einblendung des virtuellen Objekts förderlich ist, wie noch näher ausgeführt werden wird. Unter dem Sichtfeld wird dabei in herkömmlicher Weise die räumliche Ausdehnung des binokularen Gesichtsfelds eines Menschen verstanden, also der Summe der beiden monokularen Gesichtsfelder des jeweils rechten und linken Auges allein. Bei einem Erwachsenen beträgt die horizontale Ausdehnung des binokularen Gesichtsfelds etwa 200° (100° nach links und 100° nach rechts), und die vertikale Ausdehnung etwa 60° – 70° nach oben und 70° – 80° nach unten. Dieses Sichtfeld stellt den Gesamtbereich dar, in dem ohne Zuhilfenahme von Augen- und Kopfbewegungen visuelle

Wahrnehmung möglich ist. Die vom Benutzer erfindungsgemäß getragene Wiedergabeeinheit wird dieses Sichtfeld vorzugsweise kaum einschränken. Innerhalb eines zentralen Bereichs des Sichtfeldes ist ein klares Erkennen mit maximaler Wahrnehmungsqualität hinsichtlich Sehschärfe, Mustererkennung und Farbsehen möglich. Dieser zentrale Bereich des Sichtfeldes entspricht einem horizontalen und vertikalen Seh- oder Blickwinkel von jeweils etwa 40-55°. Der horizontale und vertikale Seh- oder Blickwinkel ist dabei der Winkel zwischen zwei gegenüberliegenden Randpunkten eines Gegenstands und dem Augpunkt in einer horizontalen und einer vertikalen Ebene. Erfindungsgemäß wird wiederholt ein präzisiertes digitales Abbild von Oberflächen des Sichtfeldes erhoben. Dabei muss es sich nicht um das gesamte Sichtfeld handeln. So kann etwa der Bildwinkel der erfindungsgemäß verwendeten 3D-Kameras im Zuge des bildgebenden Verfahrens zur Abbildung eines Ausschnitts einer Szenerie in derselben Größenordnung wie der zentrale Bereich des Sichtfeldes mit maximaler Wahrnehmungsqualität liegen, also im Bereich von 40-55°.

Der erfindungsgemäß genannte, vorgegebene Zeitabstand der wiederholten Gewinnung eines aktualisierten digitalen Abbildes kann variieren und der Dynamik der Situation sowie der vorhandenen Rechenleistung angepasst sein. Zum Zeitpunkt der Anmeldung sind beispielsweise Zeitabstände von etwa drei Sekunden machbar, das heißt, dass alle drei Sekunden ein aktualisiertes digitales Abbild des Sichtfeldes des Benutzers erfolgt. Der Vorgang der Aktualisierung des digitalen Abbildes kann beschleunigt und verbessert werden, indem mithilfe aktueller 3D-Bilder („scan“) ein Differenzbild zu einem bereits vorliegenden digitalen Abbild erstellt wird. Falls dabei in Teilbereichen signifikante Abweichungen ermittelt werden, wird das digitale Abbild lediglich in diesen Teilbereichen anhand der aktuellen Daten aktualisiert.

Bei den Zustandsdaten kann es sich wie erwähnt um physikalische Zustandsdaten wie z.B. Temperatur oder Sichtweite, oder um chemische Zustandsdaten wie z.B. Schadstoffkonzentrationen handeln. Zustandsdaten dieser Art sind beispielsweise für Einsatzkräfte der Feuerwehr

entscheidend und werden wie erwähnt entweder von bereits installierten Sensoren erhoben, oder von Sensoren, die von den Einsatzkräften mitgeführt werden. Ein weiteres Beispiel sind sicherheitsrelevante Zustandsdaten beispielsweise für Einsatzkräfte der Polizei. Zustandsdaten dieser Art werden etwa von Videokameras einer bereits vorhandenen Sicherheitsarchitektur erhoben, mit denen zunächst Personen detektiert werden und in weiterer Folge durch Abgleich mit einem digitalen Abbild der realen Umgebung ein mögliches Schussfeld ermittelt wird. Sicherheitsrelevante Zustandsdaten könnten somit beispielsweise die Information *„Sichtfeld quert das mögliche Schussfeld einer unbekannten Person“* beinhalten. Des Weiteren könnte im Rahmen des erfindungsgemäßen Verfahrens auf bereits vorhandene Brand- oder Rauchdetektoren als externe Sensoren zurückgegriffen werden, die einen möglichen Brandherd detektieren. Sicherheitsrelevante Zustandsdaten könnten in diesem Fall die Information *„Sichtfeld befindet sich im Bereich ansteigender Rauchgaskonzentration“* beinhalten. Ein mögliches Anwendungsbeispiel luftgestützter Sensoren wie Drohnen und dergleichen wäre beispielsweise im militärischen Bereich im Rahmen von Aufklärungsaktivitäten eines Gefechtsbereiches möglich, wobei sicherheitsrelevante Zustandsdaten beispielsweise die Information *„Sichtfeld befindet sich im Schussfeld feindlicher Artillerie“* sein können.

Des Weiteren könnte es sich bei den Zustandsdaten auch um physiologische Zustandsdaten wie z.B. Puls, Blutdruck, Sauerstoffsättigung und dergleichen handeln. Zustandsdaten dieser Art sind für Einsatzkräfte der Rettung entscheidend und werden an einem Patienten von medizinischen Sensoren erhoben. Bei den sichtfeldbezogenen Zustandsdaten handelt es sich somit um physiologische Zustandsdaten eines Patienten im Sichtfeld des Benutzers.

Wie bereits ausgeführt wurde, kann die Wiedergabe von Zustandsdaten akustisch über ein vom Benutzer mitgetragenes Gerät wie Kopfhörer, Ohrlautsprecher und dergleichen erfolgen, wobei die akustische Wiedergabe über eine vom Benutzer dem virtuellen Objekt zugeordnete, sprachliche Ausgabe erfolgt.

Diese Zuordnung kann etwa durch geeignete graphische Aufbereitung des virtuellen Objekts erfolgen, beispielsweise in Form eines virtuellen Objekts mit einem antropomorphen Aussehen. Zudem kann vorzugsweise vorgesehen sein, dass der benutzerseitige Befehl zur Wiedergabe von Zustandsdaten oder des virtuellen Objekts mithilfe einer sprachgesteuerten Eingabeschnittstelle zur zentralen Datenverarbeitungseinheit erfolgt. Durch den dialogbasierten Zugriff kann der Benutzer beispielsweise mithilfe der Sprachsteuerung gewünschte sichtfeldbezogene Zustandsdaten abrufen, also etwa durch die mündlich gestellte Frage „*Wie hoch ist die momentane Kohlenmonoxid-Konzentration?*“ oder einfach „*Kohlenmonoxid?*“. Dieser Sprachbefehl wird von der als Mikrofon ausgeführten Eingabeschnittstelle aufgenommen, das der Benutzer trägt und über das der mündlich geäußerte Eingabebefehl der zentralen Datenverarbeitungseinheit übermittelt wird. Die sprachgesteuerte Eingabeschnittstelle kann auch zur Eingabe von Befehlen zur Wiedergabe des virtuellen Objekts verwendet werden. Ein als Avatar ausgeführtes virtuelles Objekt kann etwa über den sprachgesteuerten Befehl „*Komm mit!*“ dazu aufgefordert werden dem Benutzer zu folgen.

Des Weiteren wird vorgeschlagen, dass ein automatisierter Abruf von Zustandsdaten in denselben vorgegebenen Zeitabständen oder in anderen vorgegebenen Zeitabständen wie jenen der wiederholten Aktualisierungen des digitalen Abbildes erfolgt. Somit kann eine automatisierte Wiedergabe von Zustandsdaten erfolgen, beispielsweise über den optisch oder akustisch wiedergegebenen Hinweis „*Achtung, Schadstoffkonzentration erhöht sich!*“ oder „*Sauerstoffsättigung des Patienten sinkt!*“. Der Abruf kann beispielsweise im Takt der Iterationen der Aktualisierungen des digitalen Abbildes erfolgen.

Vorzugsweise erfolgt die Verarbeitung der mittels Sensoren gewonnenen sichtfeldbezogenen Zustandsdaten in der zentralen Datenverarbeitungseinheit mithilfe eines künstlichen neuronalen Netzwerks. Ein solches künstlich neuronales Netzwerk wird auch als „Künstliche Intelligenz“ oder „KI“ bezeichnet. Das beispielsweise als Avatar ausgeführte

virtuelle Objekt greift somit auf künstliche Intelligenz zurück, um durch so genanntes „Machine Learning“ und an sich bekannter Klassifikationsverfahren Gefahrenquellen zu identifizieren, sowie durch autonome Agentensysteme eine Lagebewertung automatisiert durchzuführen. Alle zur Verfügung stehenden Informationen, die mithilfe der externen Sensoren oder durch wiederholte Erhebung aktualisierter räumlicher Daten gewonnen werden, können von der zentralen Datenverarbeitungseinheit verwendet werden, um eine aktuelle Gefahren einschätzung treffen und Einsatzkräfte über den Avatar als Benutzerschnittstelle entsprechend informieren zu können. Des Weiteren kann von der zentralen Datenverarbeitungseinheit eine Indoor- oder Outdoornavigation unter Berücksichtigung aller erkannten erfassten Gefahren vorgenommen werden. Ebenso kann der Benutzer über den Avatar durch den permanenten Abgleich der Sensorik mit den erst bei der Echtansicht gewonnenen Daten informiert werden, wo sich im Raum Gegenstände, Möbel oder dergleichen befinden.

Wie bereits erwähnt wurde, kann das virtuelle Objekt mit einem antropomorphen Aussehen versehen werden, das zu einer dialogbasierten Interaktion mit dem Benutzer fähig ist. In der Regel werden dabei Avatare nicht statisch dargestellt, sondern als Bewegtbild animiert, wofür in der Regel so genannte „3D-Echtzeit-Engines“ verwendet werden. Auf diese Weise lassen sich Interaktionen auch optisch darstellen, wodurch für den Benutzer der Eindruck einer Interaktion verstärkt wird. Um eine korrekte Bewegung des Avatars im Sichtfeld des Benutzers sicherzustellen, bei der Sicht- und Bewegungshindernisse umgangen werden, wird vorgeschlagen, dass im digitalen Abbild horizontale Freibereiche ermittelt werden, die in vertikaler Projektion frei von Sicht- oder Bewegungshindernissen sind, und eine im Sichtfeld des Benutzers zu vollziehende Bewegung des virtuellen Objekts von einem Anfangsort zu einem Zielort von der zentralen Datenverarbeitungseinheit in Form eines kürzesten und innerhalb der Freibereiche liegenden Bewegungspfades errechnet wird. Die Ermittlung der horizontalen Freibereiche kann wie bereits erwähnt wurde anhand der horizontalen Flächenelemente des digitalen Abbilds erfolgen, wobei eine Bodenfläche als Gesamtheit

aneinandergrenzender horizontaler Flächenelemente einer untersten Schichtlage ermittelt wird, die in vertikaler Projektion frei von Sicht- oder Bewegungshindernisse ist. Dieser Freibereich liegt im digitalen Abbild als Gesamtheit entsprechender Knotenpunkte („mesh“) vor. Diese Knotenpunkte („mesh“) können als so genanntes „Navigation mesh“ den oben erwähnten „3D-Echtzeit-Engines“ bereitgestellt werden, die es in bekannter Weise zur Bewegungssteuerung von als Bewegtbild animierten virtuellen Objekten („Avatare“) verwenden. Das „Navigation mesh“ kann dabei gemäß der erfindungsgemäß vorgesehenen Aktualisierungen des digitalen Abbilds in vorgegebenen Zeitabständen erneuert werden, um auf zwischenzeitlich verschobene Objekte wie Stühle und dergleichen reagieren zu können. Auf diese Weise kann aufgrund des iterativ präzisierten digitalen Abbildes eine lebensecht scheinende Bewegung des virtuellen Objekts unter Vermeidung von Bewegungshindernissen und eine korrekte Verdeckungsrechnung für das Einblenden des Avatars in Abhängigkeit von Sichthindernissen vorgenommen werden.

Des Weiteren kann vorgesehen sein, dass im digitalen Abbild horizontale Freibereiche wie oben beschrieben ermittelt werden, die in vertikaler Projektion frei von Sicht- oder Bewegungshindernissen sind, und eine im Sichtfeld des Benutzers zu vollziehende Bewegung des virtuellen Objekts von einem Anfangsort zu einem Zielort von der zentralen Datenverarbeitungseinheit in Form eines innerhalb der Freibereiche liegenden Bewegungspfades unter Einbeziehung der sensorgestützten Zustandsdaten als Randbedingung errechnet wird. Das als Avatar ausgeführte virtuelle Objekt bewegt sich somit vom Anfangsort zum Zielort unter Einbeziehung der Zustandsdaten, also beispielsweise unter Vermeidung eines Brandherdes. Mit anderen Worten erfolgt die Ermittlung eines optimalen Bewegungspfades aus der Vielzahl möglicher Bewegungspfade innerhalb der Freibereiche durch Bewertung eines Bewegungspfades anhand vorgegebener Kriterien, beispielsweise eines möglichen Sicherheitsrisikos entlang eines Bewegungspfades, das sich durch das Queren eines möglichen Schussfeldes einer nicht identifizierten Person einstellen könnte. Der optimale Bewegungspfad ist jener, bei

dem die vorgegebenen Kriterien bestmöglich erfüllt werden, also beispielsweise ein kleinstmögliches Sicherheitsrisiko einstellt. Dieser optimale Bewegungspfad kann dem Benutzer als Navigationshilfe bereitgestellt werden. Eine Möglichkeit besteht etwa darin, dass sich das als Avatar ausgeführte virtuelle Objekt entlang des optimalen Bewegungspfades bewegt. Da sich die Bewegung des Avatars im Sichtfeld der Person vollzieht, kann die Person dem Avatar einfach folgen. In einem solchen Anwendungsfall werden die vorgegebenen Zeitabstände, mit denen die zentrale Datenverarbeitungseinheit aus den gemessenen 3D-Bildern wiederholt ein aktualisiertes digitales Abbild von Oberflächen der realen Umgebung und somit der Sicht- oder Bewegungshindernisse im Sichtfeld ermittelt, entsprechend kürzer gewählt werden, beispielsweise unter 3 Sekunden, etwa im Sekundentakt.

Die Erfindung bezieht sich des Weiteren auf eine Vorrichtung umfassend eine tragbare, transparente Wiedergabeeinheit zum Einblenden eines computergenerierten, virtuellen Objektes in das von einem Benutzer wahrgenommene Sichtfeld einer realen Umgebung, insbesondere in das Sichtfeld der vom Benutzer getragenen, transparenten Wiedergabeeinheit, sowie eine mit der Wiedergabeeinheit verbundene zentrale Datenverarbeitungseinheit zum Generieren und Einblenden des virtuellen Objektes in das vom Benutzer wahrgenommene Sichtfeld einer realen Umgebung. Erfindungsgemäß wird hierbei vorgeschlagen, dass zumindest eine zur Messung von Tiefeninformation fähige Kamera zur Anfertigung von 3D-Bildern von dem der momentanen Position und Blickrichtung entsprechenden Sichtfeld des Benutzers vorgesehen ist, und die zentrale Datenverarbeitungseinheit ausgelegt ist aus den 3D-Bildern mithilfe bekannter mathematischer Verfahren in vorgegebenen Zeitabständen wiederholt ein aktualisiertes digitales Abbild von Oberflächen der realen Umgebung im Sichtfeld zu ermitteln, und Sensoren zur Ermittlung sichtfeldbezogener Zustandsdaten vorgesehen sind, die mit der zentralen Datenverarbeitungseinheit verbunden sind, wobei mithilfe des virtuellen Objekts eine für den Benutzer interaktive und dialogbasierte Benutzerschnittstelle zur Wiedergabe der sensorgestützten, sichtfeldbezogenen

Zustandsdaten im Sichtfeld des Benutzers verwirklicht ist. Die obigen Ausführungen zum erfindungsgemäßen Verfahren treffen sinngemäß auch auf die erfindungsgemäße Vorrichtung zu.

Vorzugsweise wird vorgeschlagen, dass für den benutzerseitigen Befehl zur Wiedergabe von Zustandsdaten oder des virtuellen Objekts eine sprachgesteuerte Eingabeschnittstelle zur zentralen Datenverarbeitungseinheit vorgesehen ist.

Des Weiteren kann vorgesehen sein, dass ein künstliches neuronales Netzwerk (KI) zur Verarbeitung der mittels Sensoren gewonnenen sichtfeldbezogenen Zustandsdaten in der zentralen Datenverarbeitungseinheit vorgesehen ist.

Die Erfindung wird in weiterer Folge anhand von Ausführungsbeispielen mithilfe der beiliegenden Figuren näher erläutert. Es zeigen hierbei die

Fig. 1 eine schematische Übersicht über den grundsätzlichen Ablauf der Generierung eines digitalen Abbilds im Rahmen des erfindungsgemäßen Verfahrens,

Fig. 2 eine schematische Ansicht eines Raumplans eines Gebäudeinneren, das mithilfe einer Datenbrille zur Generierung eines digitalen Abbilds und von Referenzbildern vermessen wird,

Fig. 3 ein Beispiel eines real ermittelten digitalen Abbilds eines Gebäudeinneren,

Fig. 4 eine schematische Übersicht über den grundsätzlichen Ablauf der Ermittlung von Sicht- oder Bewegungshindernissen im Rahmen des erfindungsgemäßen Verfahrens,

Fig. 5 eine schematische Darstellung zur Erläuterung der horizontalen Schichtung zur Ermittlung der Sicht- oder Bewegungshindernisse, und die

Fig. 6 eine schematische Ansicht eines Raumplans eines Gebäudeinneren zur Erläuterung des erfindungsgemäßen Verfahrens und der erfindungsgemäßen Vorrichtung.

Zunächst wird anhand der Fig. 1 bis 5 erläutert, wie im Rahmen der Erfindung ein digitales Abbild gewonnen und aktualisiert

wird und daraus Sicht- und Bewegungshindernisse für das Einblenden eines virtuellen Objekts V ermittelt werden, und danach wird anhand der Fig. 6 mithilfe eines Ausführungsbeispiels erläutert, wie die erfindungsgemäße Wiedergabe von Zustandsdaten erfolgt. Die Fig. 1 zeigt zunächst eine schematische Übersicht über den grundsätzlichen Ablauf der Generierung eines digitalen Abbilds eines Gebäudeinneren im Rahmen des erfindungsgemäßen Verfahrens. Zunächst ist ein digitales Abbild der dem Rauminnen zugewandten Oberflächen des Gebäudeinneren zu erstellen. Gemäß Fig. 1 wird hierfür eine Wiedergabeeinheit 1 in Form einer Datenbrille verwendet, die mit einer Kamera zur Aufnahme zweidimensionaler Bilder, mit einer 3D-Kamera zur Aufnahme von mit Tiefeninformation versehenen Bildern sowie mit einer Sensorik zur Ermittlung von Position und Blickrichtung versehen ist und in der Fig. 1 oben rechts angedeutet ist. Mit einer solchen Wiedergabeeinheit 1 wird das Gebäudeinnere durchschritten, das beispielsweise einen Grundriss gemäß der Fig. 2 aufweist. Die Fig. 2 zeigt schematisch einige Positionen eines Benutzers B mit einer solchen Wiedergabeeinheit 1, die vom Benutzer B im Zuge des Durchschreitens der Räume nacheinander eingenommen werden. Das Ergebnis des Durchschreitens ist eine Vielzahl an mit Tiefeninformation versehenen Bildern der Raumansichten, die sich dem Benutzer während des Durchschreitens bieten, und die an eine zentrale Datenverarbeitungseinheit 2 gesendet werden. Bei der zentralen Datenverarbeitungseinheit 2 handelt es sich in der Regel um ein externes Gerät wie beispielsweise ein stationärer oder tragbarer Computer oder Tablet, sie könnte aber auch in der Wiedergabeeinheit 1 integriert sein. Mithilfe bekannter mathematischer Verfahren kann in der zentralen Datenverarbeitungseinheit 2 aus den so gewonnenen „Scans“ an Tiefeninformationen ein digitales Abbild gewonnen werden, das die dem Benutzer jeweils zugewandten Oberflächen des Gebäudeinneren mithilfe von Flächenelementen F approximiert. Diese Flächenelemente F sind in der Fig. 2 als planare Flächenelemente F in Form von Dreiecken angedeutet, die die zu approximierenden Oberflächen lückenlos nachbilden. Eine weitere Möglichkeit bekannter mathematischer Verfahren besteht

darin den Flächenelementen F auch eine Orientierung zuzuschreiben, indem eine Flächennormale auf das betreffende Flächenelement F errechnet wird. Auf diese Weise können Flächenelemente F als vertikal orientierte Flächenelemente F_v oder als horizontal orientierte Flächenelemente F_h identifiziert werden, aus denen in weiterer Folge Sicht- oder Bewegungshindernisse rekonstruiert werden können, wie später anhand der Fig. 4 erläutert werden soll. Das Ergebnis des „Scans“ ist ein dreidimensionaler „Mesh“ des Gebäudeinneren. Sobald eine gewisse Differenz der Höhenkoordinate festgestellt wird, erfolgt die Generierung eines neuen 3D-Modells, da dies als Wechsel des Stockwerkes interpretiert wird. Durch Visualisierung der bereits gescannten Flächen lässt man den Benutzer B wissen, was bereits erfasst wurde und wo noch Lücken im Scan sind. Nachdem der Benutzer B zufrieden ist, werden die 3D-Modelle für jedes Stockwerk am Server (Tablet-Applikation) oder lokal gespeichert. Darüber hinaus können zusätzlich von jedem Raum ein oder mehrere Referenzbilder beispielsweise als RGB-Bilder aufgenommen werden, die später als Wiedererkennungspunkte dienen sollen („optische Marker“) und in einer Datenbank der zentralen Datenverarbeitungseinheit 2 hinterlegt werden. Diesen Referenzbildern werden dabei die jeweiligen, bekannten Aufnahmepositionen zugeordnet, wie sie in der Fig. 2 durch die Rauten angedeutet sind. Anhand dieses Vorgangs kann auch festgestellt werden, welche Räume miteinander verbunden sind. Somit können direkt mögliche Raumwechsel eingeschränkt werden, was bei der Bestimmung des aktuellen Raumes im Positionsbestimmungsverfahren hilfreich sein kann.

Nachdem das digitale Abbild des Gebäudeinneren, die Referenzbilder sowie die Sicht- und Bewegungshindernisse ermittelt wurden, ist die zentrale Datenverarbeitungseinheit 2 für einen Einsatz zum Einblenden virtueller Objekte bereit, die in einer Datenbank der zentralen Datenverarbeitungseinheit 2 gespeichert sind. Bei einem Einsatz in einem bislang noch nicht durchschrittenen Gebäudeinneren kann das digitale Abbild anhand eines bekannten Gebäudeplans erfolgen, der mithilfe von 3D-Bildern der vom Benutzer mitgeführten 3D-Kamera ergänzt wird. Bei einem Einsatz im offenen Gelände kann das digitale

Abbild anhand eines bekannten Umgebungsplans erfolgen, der mithilfe von 3D-Bildern der vom Benutzer mitgeführten 3D-Kamera ergänzt wird. Mittels eigener Software lassen sich auf dem digitalen Abbild virtuelle Objekte V platzieren.

Im Zuge eines Trainings oder Einsatzes bewegt sich ein Benutzer B mit einer Wiedergabeeinheit 1 etwa in Form einer Datenbrille, die mit einer Kamera zur Aufnahme zweidimensionaler Bilder, mit einer 3D-Kamera zur Aufnahme von mit Tiefeninformation versehenen Bildern sowie mit Sensorik zur Ermittlung von Position und Blickrichtung des Benutzers B versehen ist, durch das Gebäudeinnere. Diese Wiedergabeeinheit 1 ist in der Fig. 1 oben links angedeutet. Bei der Durchführung des Trainings oder Einsatzes werden die Szenario-Informationen wieder von der zentralen Datenverarbeitungseinheit 2 an die Applikation der Wiedergabeeinheit 1 übertragen. Damit das Szenario korrekt abläuft, muss das aktuelle Koordinatensystem der Wiedergabeeinheit 1 mit jenem synchronisiert werden, das für die Erstellung des digitalen Abbilds verwendet wurde. Hierfür können die zuvor aufgenommenen Referenzbilder verwendet werden. Liefert die Kamera der Wiedergabeeinheit 1 für eine Ansicht ein Bild, das zuvor aufgenommen wurde, kann die Abweichung der beiden Koordinatensysteme errechnet werden. Die Referenzbilder können auch dazu verwendet werden, das Szenario in eingeschränkter Weise fortzuführen, sollte sich die Wiedergabeeinheit 1 nicht mehr orientieren können, da man durch deren Wiedererkennung auf die Position und Orientierung des Benutzers B rückschließen kann. Alternativ wäre es auch denkbar an mehreren Orten sogenannte „Spatial Anchors“ zu erstellen. Durch Vergleich der Anchor-Transformationen lässt sich ebenfalls eine Differenz der Startpunkte und der Orientierung errechnen.

Die Position der Wiedergabeeinheit 1 kann an der zentralen Datenverarbeitungseinheit 2 im digitalen Abbild des Raumplanes dargestellt werden.

Mithilfe der als planare Flächenelemente F des digitalen Abbilds vorliegenden Sicht- oder Bewegungshindernisse kann nun das Einblenden eines virtuellen Objekts V in Abhängigkeit von

der Anwesenheit eines Sicht- oder Bewegungshindernisses auf einfache Weise erfolgen, indem im Rahmen eines Berechnungsalgorithmus Verbindungslinien zwischen einem der momentanen Position und Blickrichtung des Benutzers B entsprechenden Augpunkt des Benutzers B und Bildpunkten des virtuellen Objekts V errechnet und innerhalb des Abstandes zwischen dem Augpunkt und den Bildpunkten des virtuellen Objekts V das Vorliegen von Schnittpunkten der Verbindungslinien mit den als planare Flächenelementen F vorliegenden Sicht- oder Bewegungshindernissen ermittelt werden, wobei bei Vorliegen eines Schnittpunkts für einen Bildpunkt des virtuellen Objekts V der betreffende Bildpunkt im Sichtfeld des Benutzers B nicht oder nicht sichtbar eingeblendet wird, und bei Fehlen eines Schnittpunkts für einen Bildpunkt des virtuellen Objekts V der betreffende Bildpunkt sichtbar eingeblendet wird. Das virtuelle Objekt V wird somit in Abhängigkeit von der Anwesenheit eines Sicht- oder Bewegungshindernisses entweder zur Gänze, teilweise oder überhaupt nicht sichtbar eingeblendet.

Mithilfe der als planare Flächenelemente F des digitalen Abbilds vorliegenden Sicht- oder Bewegungshindernisse kann aber auch eine Anwendung als Navigationshilfe erfolgen, indem es sich bei den in Abhängigkeit von der Anwesenheit eines Sicht- oder Bewegungshindernisses eingeblendeten virtuellen Objekten V um Navigationshilfen handelt, die dem Benutzer B in das Sichtfeld der Wiedergabeeinheit 1 eingeblendet werden. Diese Navigationshilfen können den Einsatzkräften etwa den schnellsten Weg zu einem Zielpunkt unter Berücksichtigung von Bewegungshindernissen anzeigen, oder bei schlechten Sichtverhältnissen eine sichere Bewegung im Gebäudeinneren ermöglichen, wie anhand der Fig. 6 noch genauer beschrieben wird.

In der Fig. 1 ist ferner ein Pfeil zu sehen, der zurück zur Generierung eines digitalen Abbilds führt. Dieser Pfeil soll andeuten, dass mittels der an der Wiedergabeeinheit 1 befestigten und zur Messung von Tiefeninformation fähigen Kamera der momentanen Position und Blickrichtung des Benutzers B entsprechende 3D-Bilder angefertigt werden, aus denen ein

aktualisiertes digitales Abbild ermittelt werden kann. Da sich die momentane Ansicht eines Raumbereiches im Vergleich zu einem vorab erstellten digitalen Abbild auch ändern kann, indem etwa ein Möbelstück verschoben oder eine Tür geschlossen wurde, wird das digitale Abbild anhand der vom Benutzer B momentan wahrgenommenen, realen Umgebung aktualisiert. Diese Aktualisierung wird in vorgegebenen Zeitabständen wiederholt durchgeführt, wobei der vorgegebene Zeitabstand der wiederholten Gewinnung eines aktualisierten digitalen Abbildes variieren kann und der Dynamik der Situation sowie der vorhandenen Rechenleistung angepasst sein wird. Zum Zeitpunkt der Anmeldung sind beispielsweise Zeitabstände von etwa drei Sekunden machbar, das heißt, dass alle drei Sekunden ein aktualisiertes digitales Abbild des Sichtfeldes des Benutzers B erfolgt. Der Vorgang der Aktualisierung des digitalen Abbildes kann beschleunigt und verbessert werden, indem mithilfe aktueller 3D-Bilder („scan“) ein Differenzbild zu einem bereits vorliegenden digitalen Abbild erstellt wird. Falls dabei in Teilbereichen signifikante Abweichungen ermittelt werden, wird das digitale Abbild lediglich in diesen Teilbereichen anhand der aktuellen Daten aktualisiert.

In weiterer Folge wird anhand der Fig. 4 und 5 die Ermittlung von Sicht- und Bewegungshindernissen aus dem digitalen Abbild erläutert. Wie bereits erwähnt wurde, kann mithilfe bekannter mathematischer Verfahren aus den zuvor beschriebenen Scans des Gebäudeinneren ein digitales Abbild als so genannter „Mesh“ gewonnen werden, das die Oberflächen des Gebäudeinneren mithilfe von planaren Flächenelementen F („Faces“) approximiert. Mithilfe einer Applikation, die auf der zentralen Datenverarbeitungseinheit 2 läuft, werden die Meshes, die von der Wiedergabeeinheit 1 erhalten wurden, bearbeitet und daraus ein digitales Abbild erstellt. Die Vertices und Faces werden anhand ihrer Höhe in Stockwerke eingeteilt. Danach werden bei den gescannten Meshes pro Stockwerk alle Vertices und Faces oberhalb einer Augenhöhenebene A , die der durchschnittlichen Augenhöhe entspricht, entfernt, sowie um fehlerhafte Faces bereinigt. Die genaue Höhenlage der Augenhöhenebene A ist dabei nicht wesentlich, sie wird beispielsweise mit einer Höhe von 1,80m

angenommen. Das digitale Abbild des Gebäudeinneren wird somit gewissermaßen in Augenhöheebene A „abgeschnitten“. Was sich in Augenhöhe oder knapp darunter befindet wird als Teil eines Sicht- oder Bewegungshindernisses, also Wände, Einrichtung oder sonstiges, gewertet. Was sich in Bodennähe befindet, wird als solcher eingestuft und damit als passierbar bewertet. Hierbei wird wie folgt vorgegangen. Zunächst werden jene Flächenelemente F ermittelt, deren Flächennormale horizontal oder vertikal verlaufen, also Flächenelemente F , deren Orientierung vertikal oder horizontal ist. In weiterer Folge wird eine horizontale Schichtung zwischen der Bodenebene B des Gebäudeinneren und der angenommenen Augenhöheebene A mit einer Mehrzahl von Schichtlagen S_i ($i=1, \dots, N$) einer vorgegebenen Schichtdicke vorgenommen. Im gezeigten Beispiel der Fig. 5 werden etwa sechs Schichtlagen S_i ($N=6$) mit einer Schichtdicke von jeweils 30cm verwendet.

Zum Einen werden nun vertikale Flächenelemente Fv_j ($j=1, \dots, M$) als Sicht- oder Bewegungshindernisse ermittelt. Diese Sicht- oder Bewegungshindernisse werden in der Regel Wände, Raumteiler oder hohe Möbelstücke wie Schränke und dergleichen repräsentieren. Zum Anderen werden horizontale Flächenelemente Fh_k ($k=1, \dots, P$) oder eine Gesamtheit aneinandergrenzender horizontaler Flächenelemente Fh_k in derselben Schichtlage als Sicht- oder Bewegungshindernisse ermittelt, falls in deren Umfangsbereich horizontale Nebenflächenelemente Fn auffindbar sind, die um mehr als eine Schichtlage unterhalb des horizontalen Flächenelements Fh_k oder der Gesamtheit aneinandergrenzender horizontaler Flächenelemente Fh_k liegen. Diese Sicht- oder Bewegungshindernisse werden in der Regel niedrige Möbelstücke wie Tische und dergleichen repräsentieren, im Beispiel der Fig. 5 ist etwa ein Tisch angedeutet. Die so als Sicht- oder Bewegungshindernis identifizierte Gesamtheit aneinandergrenzender horizontaler Flächenelemente Fh_k können etwa einen Höhenabstand von 90-120 cm von der Bodenfläche N aufweisen, wie im Beispiel der Fig. 5. Diese Schichtung ermöglicht es zudem, dass horizontale Flächenelemente Fh_l ($l=1, \dots, Q$), die sich in vertikal aufeinanderfolgenden Schichtlagen S_i befinden, als Stufen identifiziert werden können, wie ebenfalls in der Fig. 5

angedeutet ist. In diesem Sonderfall werden horizontale Flächenelemente Fh_1 als Sichthindernisse, nicht aber als Bewegungshindernisse interpretiert.

Horizontale Faces werden somit als Sitz- oder Tischflächen, Böden und ähnliches kategorisiert. Vertikale Faces werden hingegen als Wände und somit als nicht passierbare Sicht- oder Bewegungshindernisse kategorisiert. Durch Projektionen der Vertices und der Faces auf eine 2D-Fläche lässt sich ein Plan der gescannten Umgebung erstellen. Ein Beispiel eines solchen Plans ist in Figur 3 zu sehen. Anhand der Höhe und Ausrichtung der Normalvektoren der Faces kann die Visualisierung der Faces verfeinert werden, um Höhenunterschiede zwischen den Sicht- oder Bewegungshindernissen hervorheben.

Die Gesamtheit aneinandergrenzender horizontaler Flächenelemente Fh_1 einer untersten Schichtlage kann als Bodenfläche N bewertet werden. Jener Anteil der Bodenfläche N , der in vertikaler Projektion frei von Sicht- oder Bewegungshindernisse ist, wird in weiterer Folge als Freibereich bezeichnet. Dieser Freibereich liegt im digitalen Abbild als Gesamtheit entsprechender Knotenpunkte („mesh“) vor, die in weiterer Folge als so genanntes „Navigation mesh“ zur Bewegungssteuerung eines virtuellen Objekts („Avatar“) verwendet werden kann. Hierfür sind so genannte „3D-Echtzeit-Engines“ bekannt, die anhand eines solchen „Navigation mesh“ die Bewegungssteuerung von als Bewegtbild animierten virtuellen Objekten („Avatare“) vornehmen, wie in weiterer Folge anhand der Fig. 6 erläutert wird.

Die Fig. 6 zeigt den Grundriss eines Gebäudeinneren analog zur Fig. 2. Die Fig. 6 zeigt des Weiteren die momentane Position eines Benutzers B mit einer Wiedergabeeinheit 1, mit der ihm ein virtuelles Objekt V eingeblendet wird. Der Grundriss liegt als digitales Abbild beispielsweise aufgrund eines bekannten Gebäudeplans vor, oder wurde durch vorheriges Durchschreiten anhand der mitgeführten 3D-Kameras erstellt. Das digitale Abbild wird in beiden Fällen wie erwähnt anhand der mitgeführten 3D-Kameras laufend aktualisiert und durch die Detektion der Anwesenheit und Positionierung von Gegenständen wie Möbel und dergleichen präzisiert.

Des Weiteren wird auf externe Sensoren 3 zugegriffen. In der Fig. 6 ist beispielsweise ein Bewegungssensor in einem Raum im linken oberen Bereich des Gebäudeplans eingezeichnet, der als externer Sensor 3 dient. Dieser Sensor 3 detektiert die Anwesenheit sich bewogender Objekte oder Personen und liefert ein Signal bei Anwesenheit sich bewogender Objekte oder Personen. Der externe Sensor 3 ist mit der zentralen Datenverarbeitungseinheit 2 verbunden, in der das vom Sensor 3 gelieferte Signal verarbeitet wird. Vorzugsweise erfolgt die Verarbeitung der mittels der Sensoren 3 gewonnenen sichtfeldbezogenen Zustandsdaten in der zentralen Datenverarbeitungseinheit 2 mithilfe eines künstlichen neuronalen Netzwerks. Ein solches künstlich neuronales Netzwerk wird auch als „Künstliche Intelligenz“ oder „KI“ bezeichnet.

Die zentrale Datenverarbeitungseinheit 2 steuert ferner das Einblenden des computergenerierten, virtuellen Objektes V in das vom Benutzer B wahrgenommene Sichtfeld gemäß der momentanen Sichtachsen des Benutzers B unter Berücksichtigung der Anwesenheit von Sicht- oder Bewegungshindernissen wie oben beschrieben. Eine Möglichkeit besteht etwa darin das computergenerierte virtuelle Objekt V mit einem antropomorphen Aussehen auszustatten, das als Bewegtbild so animiert wird, dass auf die Eingabe von Sprachbefehlen des Benutzers B reagiert wird. Ein solches virtuelles Objekt V wird im Folgenden auch als Avatar bezeichnet. Auf diese Weise wird für den Benutzer der Eindruck einer Interaktion erzeugt. Um eine korrekte Bewegung des virtuellen Objekts V im Sichtfeld des Benutzers B sicherzustellen, wird auf die oben beschriebene Bodenfläche N des digitalen Abbilds des Gebäudeplans der Fig. 6 als Gesamtheit aneinandergrenzender horizontaler Flächenelemente Fh_1 einer untersten Schichtlage zurückgegriffen, die in vertikaler Projektion frei von Sicht- oder Bewegungshindernisse ist. Dieser Freibereich liegt im digitalen Abbild als „Navigation mesh“ vor, mit dem mithilfe eigener Software, etwa die oben erwähnten „3D-Echtzeit-Engines“, eine Bewegungssteuerung des als Bewegtbild animierten virtuellen Objektes V vorgenommen wird. Das „Navigation mesh“ kann dabei in vorgegebenen Zeitabständen

erneuert werden, um auf zwischenzeitlich verschobene Objekte wie Stühle und dergleichen reagieren zu können. Auf diese Weise kann aufgrund des iterativ präzisierten digitalen Abbildes und der Animation durch die zentrale Datenverarbeitungseinheit 2 eine lebensecht scheinende Bewegung des virtuellen Objekts V unter Vermeidung von Bewegungshindernissen und eine korrekte Verdeckungsrechnung für das Einblenden des virtuellen Objekts V in Abhängigkeit von Sichthindernissen vorgenommen werden.

Im vorliegenden Beispiel der Fig. 6 könnte es sich etwa um ein Szenario für Einsatzkräfte der Polizei handeln. Der Benutzer B ist Teil der Einsatzkräfte und befindet sich in einem Gebäude, in dem ein bewaffneter Täter vermutet wird. Der Auftrag lautet das Gebäude zu durchsuchen, unbeteiligte Personen zu evakuieren und den Täter zu stellen. Der Benutzer B ist hierfür mit der Wiedergabeeinheit 1 ausgerüstet, die mit der zentralen Datenverarbeitungseinheit 2 verbunden ist, sodass auch seine Position der zentralen Datenverarbeitungseinheit 2 jederzeit bekannt ist. Die zentrale Datenverarbeitungseinheit 2 ist des Weiteren mit externen Sensoren 3 verbunden, beispielsweise mit dem in der Fig. 6 dargestellten Bewegungssensor.

Der Sensor 3 liefert nun ein Signal für die Anwesenheit einer Person im linken oberen Raum des Gebäudeplans an die zentrale Datenverarbeitungseinheit 2. Die zentrale Datenverarbeitungseinheit 2 detektiert die Möglichkeit, dass es sich um den bewaffneten Täter handeln könnte, sowie die Nähe des Benutzers B, dessen Sichtfeld bei Verlassen seines in der Fig. 6 gezeigten, momentanen Standortes in das Schussfeld des möglichen Täters gelangen könnte. Die zentrale Datenverarbeitungseinheit 2 übermittelt somit eine Warnung an die Wiedergabeeinheit 1 als sicherheitsrelevante Zustandsdaten „Sichtfeld könnte das mögliche Schussfeld einer unbekannten Person queren“ und steuert die Wiedergabe dieser Zustandsdaten mit Unterstützung des virtuellen Objekts V. Das als Avatar ausgeführte virtuelle Objekt V könnte sich etwa unter gleichzeitiger akustischer Wiedergabe der Warnung über einen

vom Benutzer B getragenen Ohrlautsprecher zum Benutzer B wenden.

In weiterer Folge kann die mit künstlicher Intelligenz versehene zentrale Datenverarbeitungseinheit 2 auch bei der weiteren Handlungsplanung unterstützen. So kann anhand der gestellten Aufgabe unbeteiligte Personen zu evakuieren und den Täter zu stellen eine bessere Positionierung des Benutzers B überlegt werden. Die zentrale Datenverarbeitungseinheit 2 errechnet hierfür mögliche Bewegungspfade P innerhalb der Freibereiche unter Einbeziehung der Zustandsdaten, also unter Verringerung eines möglichen Sicherheitsrisikos entlang eines Bewegungspfad P, das sich durch das Queren eines möglichen Schussfeldes der nicht identifizierten Person einstellen könnte. Andererseits sollte der Bewegungspfad P auch die Identifizierung einer unbeteiligten Person erlauben. Der optimale Bewegungspfad P ist jener, bei dem sich also beispielsweise ein kleinstmögliches Sicherheitsrisiko einstellt, der aber dennoch eine Identifizierung der Person im linken oberen Raum gestattet. Dieser optimale Bewegungspfad P ist in der Fig. 6 eingezeichnet und kann dem Benutzer B als möglicher Bewegungspfad P vorgeschlagen werden. Die Wiedergabe kann wiederum über das als Avatar ausgeführte virtuelle Objekt V erfolgen, indem sich der Avatar beispielsweise über entsprechende Handbewegungen bemerkbar macht und dem Benutzer B entweder über eine akustische Wiedergabe über einen Ohrlautsprecher oder lautlos über Handbewegungen ein mögliches Verlassen der Position und die geplante Bewegungsrichtung vorschlägt. Der Benutzer B kann seine Zustimmung durch den mündlich artikulierten Eingabebefehl „OK“ signalisieren, der von einem vom Benutzer B mitgeführten Mikrofon der zentralen Datenverarbeitungseinheit 2 übermittelt wird. Die zentrale Datenverarbeitungseinheit 2 startet daraufhin die Animation des virtuellen Objekts V in Form einer Bewegung entlang des errechneten optimalen Bewegungspfad P. Die Navigation des Benutzers B entlang des errechneten optimalen Bewegungspfad P kann somit einfach dadurch erfolgen, indem der Benutzer B dem Avatar folgt.

Mithilfe der Erfindung wird somit ein Verfahren bereitgestellt, bei dem unter Nutzung von „Augmented-Reality“-Systemen beispielsweise im Rahmen eines sicherheitskritischen Einsatzes ein Maximum an situationsbezogenen Daten gewonnen und einem Benutzer B benutzerfreundlich bereitgestellt werden kann. Anwendungen der Erfindung sind dabei auch im privaten Bereich denkbar, indem von einem personalisierten Avatar sichtfeldbezogene Zustandsdaten von externen Sensoren für den Benutzer B erhoben werden und in Form einer interaktiven und dialogbasierten Wiedergabe erfindungsgemäß bereitgestellt werden.

Patentansprüche:

1. Verfahren zum Einblenden eines computergenerierten, virtuellen Objektes (V) in das von einem Benutzer (B) wahrgenommene Sichtfeld einer realen Umgebung, insbesondere in das Sichtfeld einer vom Benutzer (B) getragenen, transparenten Wiedergabeeinheit (1), wobei das Einblenden des virtuellen Objekts (V) mittels einer zentralen Datenverarbeitungseinheit (2) in Abhängigkeit von der momentanen Position und der Blickrichtung des Benutzers (B) erfolgt, und das Einblenden des virtuellen Objekts (V) zusätzlich in Abhängigkeit von der Anwesenheit von Sicht- oder Bewegungshindernissen erfolgt, wobei die Sicht- oder Bewegungshindernisse von der zentralen Datenverarbeitungseinheit (2) anhand eines digitalen Abbilds von Oberflächen der realen Umgebung ermittelt werden, **dadurch gekennzeichnet**, dass mittels einer zur Messung von Tiefeninformation fähigen Kamera von dem der momentanen Position und Blickrichtung entsprechenden Sichtfeld des Benutzers (B) 3D-Bilder angefertigt werden, aus denen von der zentralen Datenverarbeitungseinheit (2) mithilfe bekannter mathematischer Verfahren in vorgegebenen Zeitabständen wiederholt ein aktualisiertes digitales Abbild von Oberflächen der realen Umgebung und somit der Sicht- oder Bewegungshindernisse im Sichtfeld ermittelt wird, und mittels Sensoren (3) sichtfeldbezogene Zustandsdaten gewonnen werden, die in der zentralen Datenverarbeitungseinheit (2) verarbeitet werden, wobei mithilfe des virtuellen Objekts (V) eine für den Benutzer (B) interaktive und dialogbasierte Wiedergabe der sensorgestützten, sichtfeldbezogenen Zustandsdaten im Sichtfeld des Benutzers (B) erfolgt.
2. Verfahren nach Anspruch 1, **dadurch gekennzeichnet**, dass der benutzerseitige Befehl zur Wiedergabe von Zustandsdaten oder des virtuellen Objekts (V) mithilfe einer sprachgesteuerten Eingabeschnittstelle zur zentralen Datenverarbeitungseinheit (2) erfolgt.

3. Verfahren nach Anspruch 1 oder 2, **dadurch gekennzeichnet**, dass ein automatisierter Abruf von Zustandsdaten in denselben vorgegebenen Zeitabständen oder in anderen vorgegebenen Zeitabständen wie jenen der wiederholten Aktualisierungen des digitalen Abbildes erfolgt.
4. Verfahren nach einem der Ansprüche 1 bis 3, **dadurch gekennzeichnet**, dass die Verarbeitung der mittels Sensoren (3) gewonnenen sichtfeldbezogenen Zustandsdaten in der zentralen Datenverarbeitungseinheit (2) mithilfe eines künstlichen neuronalen Netzwerks („KI“) erfolgt.
5. Verfahren nach einem der Ansprüche 1 bis 4, **dadurch gekennzeichnet**, dass im digitalen Abbild horizontale Freibereiche ermittelt werden, die in vertikaler Projektion frei von Sicht- oder Bewegungshindernissen sind, und eine im Sichtfeld des Benutzers (B) zu vollziehende Bewegung des virtuellen Objekts (V) von einem Anfangsort zu einem Zielort von der zentralen Datenverarbeitungseinheit (2) in Form eines kürzesten und innerhalb der Freibereiche liegenden Bewegungspfades (P) errechnet wird.
6. Verfahren nach einem der Ansprüche 1 bis 5, **dadurch gekennzeichnet**, dass im digitalen Abbild horizontale Freibereiche ermittelt werden, die in vertikaler Projektion frei von Sicht- oder Bewegungshindernissen sind, und eine im Sichtfeld des Benutzers (B) zu vollziehende Bewegung des virtuellen Objekts (V) von einem Anfangsort zu einem Zielort von der zentralen Datenverarbeitungseinheit (2) in Form eines innerhalb der Freibereiche liegenden Bewegungspfades (P) unter Einbeziehung der sensorgestützten Zustandsdaten als Randbedingung errechnet wird.
7. Vorrichtung umfassend eine tragbare, transparente Wiedergabeeinheit (1) zum Einblenden eines computergenerierten, virtuellen Objektes (V) in das von einem Benutzer (B) wahrgenommene Sichtfeld einer realen Umgebung, insbesondere in das Sichtfeld der vom Benutzer (B) getragenen, transparenten Wiedergabeeinheit (1),

sowie eine mit der Wiedergabeeinheit (1) verbundene zentrale Datenverarbeitungseinheit (2) zum Generieren und Einblenden des virtuellen Objektes (V) in das vom Benutzer (B) wahrgenommene Sichtfeld einer realen Umgebung, **dadurch gekennzeichnet**, dass zumindest eine zur Messung von Tiefeninformation fähige Kamera zur Anfertigung von 3D-Bildern von dem der momentanen Position und Blickrichtung entsprechenden Sichtfeld des Benutzers (B) vorgesehen ist, und die zentrale Datenverarbeitungseinheit (2) ausgelegt ist aus den 3D-Bildern mithilfe bekannter mathematischer Verfahren in vorgegebenen Zeitabständen wiederholt ein aktualisiertes digitales Abbild von Oberflächen der realen Umgebung im Sichtfeld zu ermitteln, und Sensoren (3) zur Ermittlung sichtfeldbezogener Zustandsdaten vorgesehen sind, die mit der zentralen Datenverarbeitungseinheit (2) verbunden sind, wobei mithilfe des virtuellen Objekts (V) eine für den Benutzer (B) interaktive und dialogbasierte Benutzerschnittstelle zur Wiedergabe der sensorgestützten, sichtfeldbezogenen Zustandsdaten im Sichtfeld des Benutzers (B) verwirklicht ist.

8. Vorrichtung nach Anspruch 7, **dadurch gekennzeichnet**, dass für den benutzerseitigen Befehl zur Wiedergabe von Zustandsdaten oder des virtuellen Objekts (V) eine sprachgesteuerte Eingabeschnittstelle zur zentralen Datenverarbeitungseinheit (2) vorgesehen ist.
9. Vorrichtung nach Anspruch 7 oder 8, **dadurch gekennzeichnet**, dass ein künstliches neuronales Netzwerk („KI“) zur Verarbeitung der mittels Sensoren (3) gewonnenen sichtfeldbezogenen Zustandsdaten in der zentralen Datenverarbeitungseinheit (2) vorgesehen ist.

Fig. 1

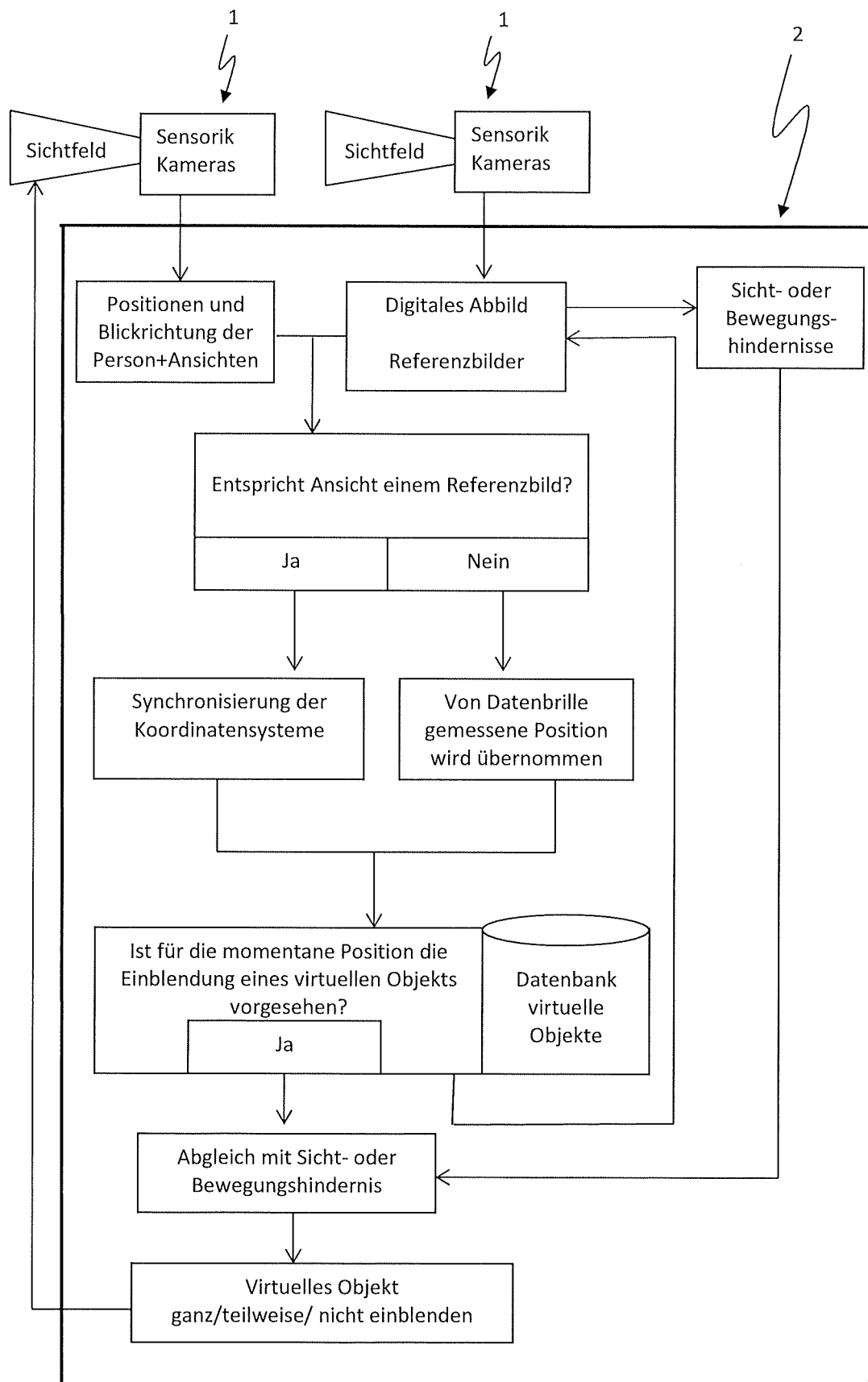


Fig. 2

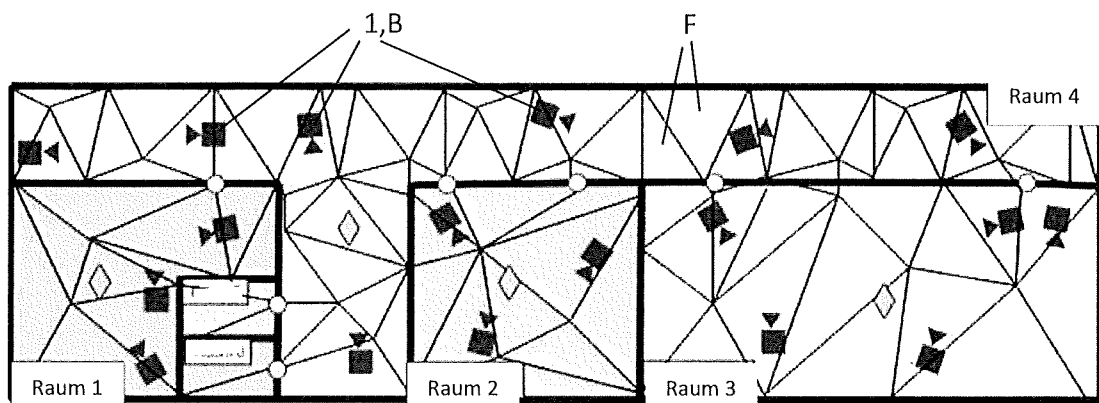


Fig. 3

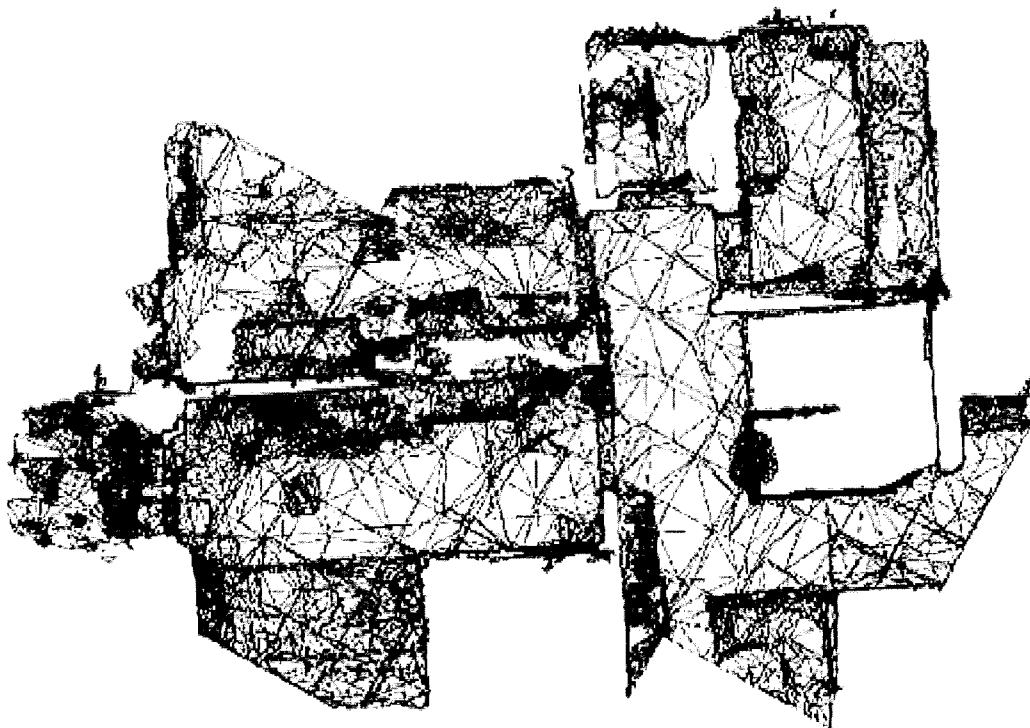


Fig. 4

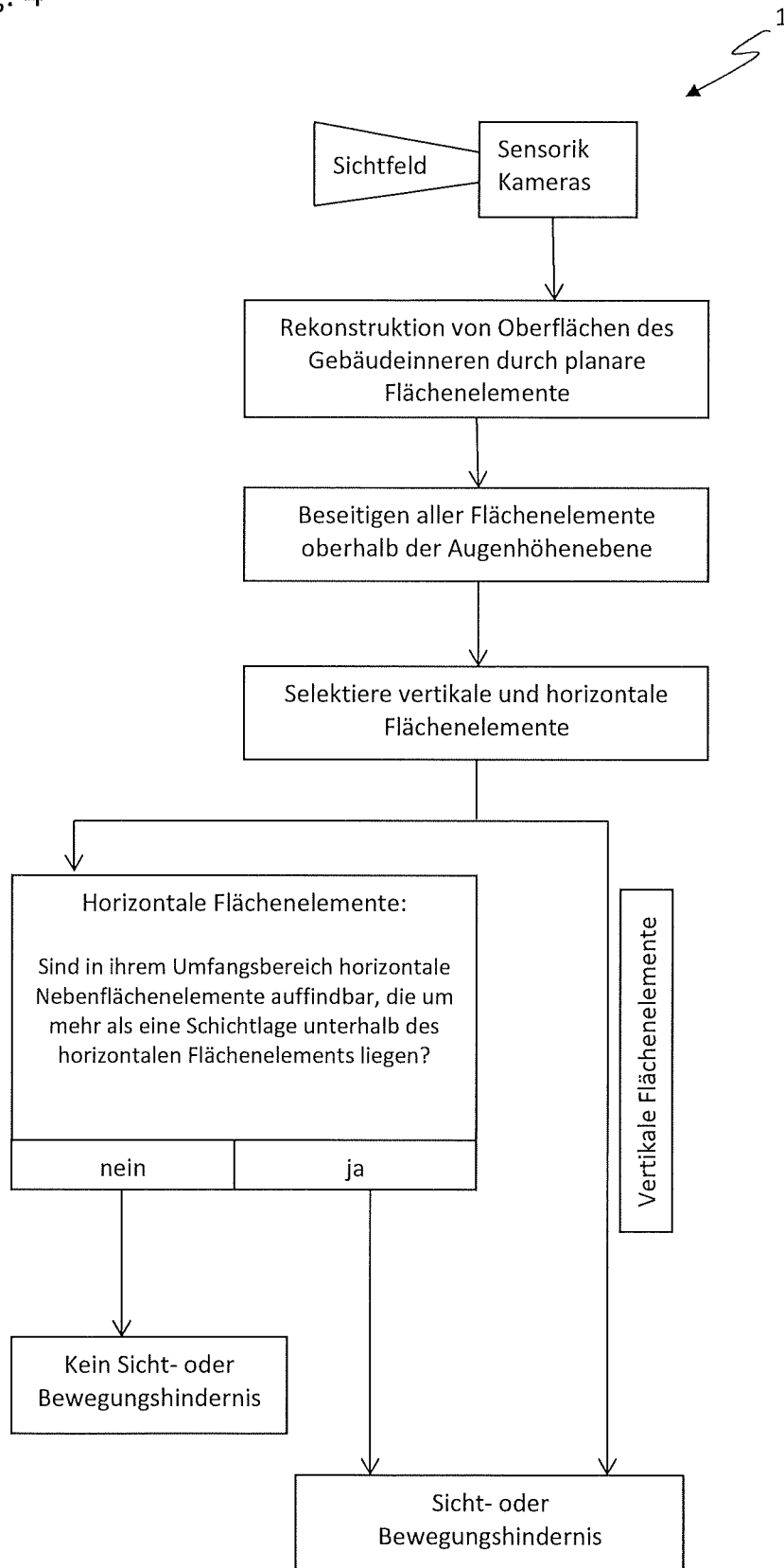


Fig. 5

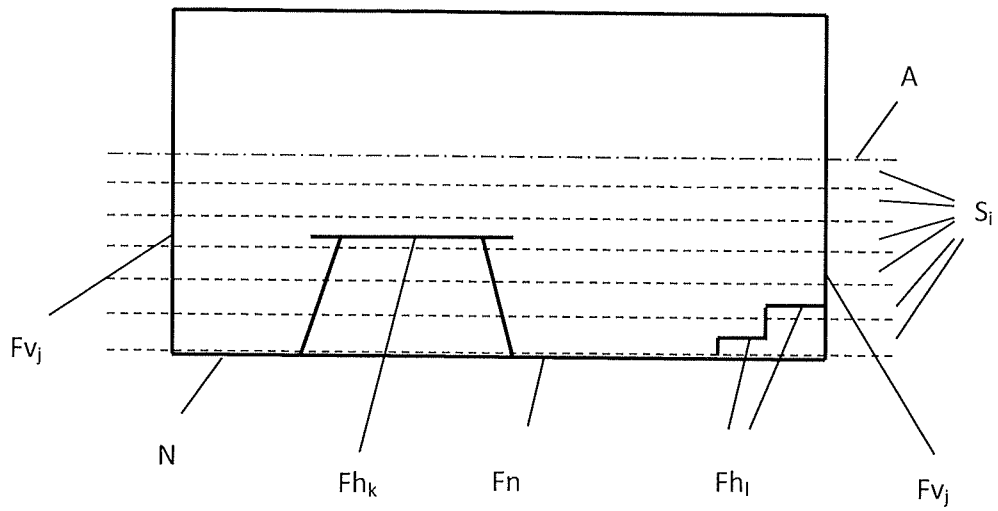


Fig. 6

