



(10) **DE 10 2020 111 955 A1** 2021.11.04

(12) **Offenlegungsschrift**

(21) Aktenzeichen: **10 2020 111 955.0**

(22) Anmeldetag: **04.05.2020**

(43) Offenlegungstag: **04.11.2021**

(51) Int Cl.: **G06K 9/62 (2006.01)**

(71) Anmelder:
**Airbus Defence and Space GmbH, 85521
Ottobrunn, DE**

(74) Vertreter:
**LKGLOBAL | Lorenz & Kopf PartG mbB
Patentanwälte, 80333 München, DE**

(72) Erfinder:
Hiebl, Manfred, 82024 Taufkirchen, DE

(56) Ermittelter Stand der Technik:

DE	699 11 780	T2
US	10 311 334	B1

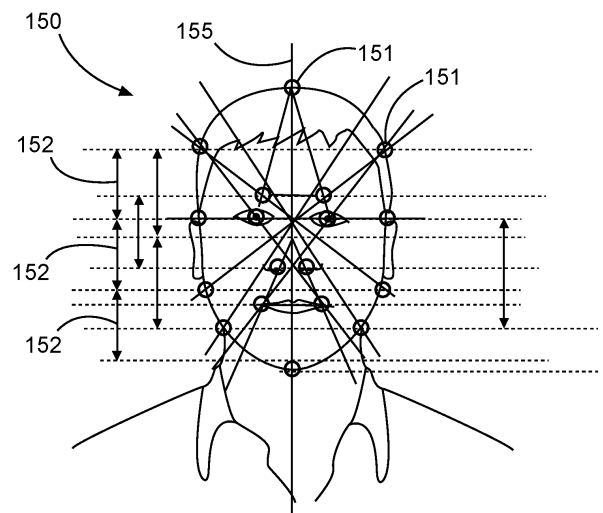
**Polygon mesh. In: Wikipedia, Die freie
Enzyklopädie. Bearbeitungsstand: 29.03.20220,
11:21 UTC. URL: [https://en.wikipedia.org/w/index.
php?title=Polygon_mesh&oldid=947958119](https://en.wikipedia.org/w/index.php?title=Polygon_mesh&oldid=947958119)
[abgerufen am 09.03.2021].**

Prüfungsantrag gemäß § 44 PatG ist gestellt.

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen.

(54) Bezeichnung: **Gesichtserkennung mit neuronalen Netzen**

(57) Zusammenfassung: Eine Objektidentifizierungseinheit enthält ein künstliches neuronales Netzwerk und ist ausgestaltet, menschliche Gesichter zu identifizieren. Hierzu wird ein Gesicht in eine Anzahl von Dreiecken unterteilt. Der relative Anteil der Fläche jedes Dreiecks an der Summe der Flächen aller Dreiecke wird ermittelt, um einen Drehwinkel des Gesichts zu ermitteln. Sodann wird der relative Anteil der Fläche jedes Dreiecks an der Summe der Flächen aller Dreiecke auf eine drehinvariante Größe des Gesichts normiert. Die normierte Fläche der Dreiecke wird dem künstlichen neuronalen Netzwerk zugeführt, um eine Person zu identifizieren.



Beschreibung

Technisches Gebiet

[0001] Die vorliegende Beschreibung betrifft eine Objektidentifizierungseinheit zum Identifizieren von Objekten und ein entsprechendes Verfahren zum Identifizieren von Objekten mit einem künstlichen neuronalen Netzwerk. Die Objektidentifizierungseinheit verwendet ein künstliches neuronales Netzwerk, um Objekte mit einer hohen Zuverlässigkeit auch dann zu erkennen, wenn die Objekte relativ zu einer Bilderfassungseinheit gedreht sind.

Technischer Hintergrund

[0002] Künstliche neuronale Netze können verwendet werden, um Objekte zu klassifizieren und/oder zu identifizieren. Für die Identifizierung einer Person nimmt deren Gesicht eine Kernrolle ein. Das Identifizieren einer Person beinhaltet insbesondere das Zuordnen eines Gesichts zu einer bestimmten Person.

[0003] Damit ein künstliches neuronales Netz eine Person identifizieren kann, muss das künstliche neuronale Netz mit einem Abbild dieser Person trainiert worden sein. Dies passiert in der sogenannten Trainingsphase. Sodann ist das künstliche neuronale Netz in der Lage, ein zu identifizierendes Abbild einer Person mit allen trainierten Personen zu vergleichen und diejenige Person zu präsentieren, welcher das zu identifizierende Abbild am Nächsten kommt.

[0004] An das Abbild einer Person, welches für den Identifizierungsvorgang verwendet wird, können gewisse Vorgaben bzw. Anforderungen gestellt werden. Ein Beispiel hierfür sind die sogenannten biometrischen Bilder, auf denen das Gesicht einer Person einen bestimmten Bereich ausfüllen muss (Anforderung an die Größe des Gesichts auf dem Abbild) und auch aus einem vorgegebenen Blickwinkelbereich erfasst werden muss (Anforderung an die Perspektive).

Darstellung der Erfindung

[0005] Es kann als Aufgabe betrachtet werden, die Zuverlässigkeit der Identifizierung von Personen mit einem künstlichen neuronalen Netzwerk zu verbessern.

[0006] Diese Aufgabe wird gelöst durch den Gegenstand der unabhängigen Ansprüche. Weitere Ausführungsformen ergeben sich aus den abhängigen Ansprüchen sowie aus der folgenden Beschreibung.

[0007] Gemäß einem Aspekt ist eine Objektidentifizierungseinheit zum Identifizieren von Objekten angegeben. Die Objektidentifizierungseinheit weist ein künstliches neuronales Netzwerk, KNN, auf. Das KNN enthält eine Eingabeschicht und eine Ausgabeschicht, wobei das KNN ausgestaltet ist, Objekte basierend auf Eingaben an die Eingabeschicht zu identifizieren, wobei die Eingabeschicht eine Mehrzahl von Eingabeneuronen aufweist, und wobei die Eingabeschicht ausgeführt ist, einen Eingabevektor zu erhalten, wobei der Eingabevektor Objektidentifizierungsinformationen zu dem zu identifizierenden Objekt enthält. Das KNN ist ausgeführt, basierend auf den Objektidentifizierungsinformationen ein Objekt zu identifizieren, wobei die Objektidentifizierungsinformationen einer zweidimensionalen Abbildung des zu identifizierenden Objekts entsprechen. Die Objektidentifizierungseinheit ist konfiguriert, die zweidimensionale Abbildung in eine Mehrzahl von Dreiecken, die durch charakteristische Punkte des zu identifizierenden Objekts an ihren Ecken begrenzt sind, zu unterteilen. Die Objektidentifizierungseinheit ist konfiguriert, aus Symmetrieeigenschaften des zu identifizierenden Objekts und relativen Flächenanteilen der einzelnen Dreiecke an der Gesamtsumme der Fläche der Dreiecke einen Drehwinkel des zu identifizierenden Objekts in der zweidimensionalen Abbildung zu ermitteln und die Flächen der einzelnen Dreiecke mit einem Korrekturfaktor zu multiplizieren, um korrigierte Flächen der Dreiecke zu erhalten, welche einer Symmetriebedingung entsprechen bzw. die Symmetriebedingung erfüllen. Die Objektidentifizierungseinheit ist konfiguriert, die korrigierten Flächen der Dreiecke auf eine drehinvariante Größe des zu identifizierenden Objekts zu normieren und die normierten Flächen der Dreiecke dem KNN als Eingabevektor zuzuführen.

[0008] Die hier beschriebene Objektidentifizierungseinheit kann für das Identifizieren verschiedener Objekte verwendet werden, insbesondere solcher Objekte, die symmetrisch mit Bezug zu einer Symmetrieachse sind. Bei diesen Objekten kann es sich um Gesichter von Menschen, Abbildungen von Tieren und Pflanzen oder von Maschinen handeln. Im Folgenden wird die Objektidentifizierungseinheit in Verbindung mit dem Erkennen von Gesichtern bzw. Identifizieren von Personen beschrieben. Es ist allerdings zu verstehen, dass dies lediglich ein

Anwendungsbeispiel ist und die Allgemeingültigkeit der Verwendung der Objektidentifizierungseinheit für das Identifizieren beliebiger Objekte durch die Bezugnahme auf ein menschliches Gesicht nicht eingeschränkt wird.

[0009] Eine Herausforderung bei der Identifizierung von Personen ist es, wenn das Gesicht einer Person auf einer fotografischen Abbildung von einer Bilderfassungseinheit, wie beispielsweise einer Kamera für bewegte oder unbewegte Bilder, gedreht ist, d.h. dass das Gesicht nicht frontal erfasst wurde. Eine solche Drehung des Gesichts erschwert die automatisierte maschinelle Erkennung des Gesichts, insbesondere wenn die Vergleichsabbildungen der Personen frontal oder in einem Blickwinkel, welcher von dem Blickwinkel der zu identifizierenden Abbildung abweicht, erfasst wurden.

[0010] Die hier beschriebene Objektidentifizierungseinheit löst diese Herausforderung, indem Abschnitte eines zu identifizierenden Abbilds eines Gesichts zunächst mit einem Korrekturfaktor versehen werden, um die durch die Drehung verursachte Abweichung der Dimensionen des Gesichts auszugleichen. Nachfolgend werden die Abschnitte des zu identifizierenden Abbilds des Gesichts normiert, um Größenunterschiede in den zu vergleichenden Abbildungen zu kompensieren. Insbesondere wird der relative Anteil der Fläche jedes Dreiecks an der Summe der Flächen aller Dreiecke ermittelt, um einen Drehwinkel des Gesichts zu ermitteln. Sodann wird der relative Anteil der Fläche jedes Dreiecks an der Summe der Flächen aller Dreiecke auf eine drehinvariante Größe des Gesichts normiert. Die normierte Fläche der Dreiecke wird dem künstlichen neuronalen Netzwerk zugeführt, um eine Person zu identifizieren.

[0011] Insbesondere ist vorgesehen, dass ein Abbild eines Gesichts in mehrere Dreiecke unterteilt wird, wobei jedes Dreieck von drei charakteristischen Punkten in dem Abbild des Gesichts definiert wird. Mindestens einige dieser Dreiecke werden so in dem Abbild des Gesichts platziert, dass jeweils zwei Dreiecke achsensymmetrisch bezogen auf eine vertikale oder horizontale Achse des Gesichts liegen. Dieser Anordnung der Dreiecke liegt der Gedanke zu Grunde, dass ein Gesicht bezogen auf eine vertikale Mittelachse symmetrisch ist. Demnach wird unterstellt, dass Dreiecke in den beiden Gesichtshälften, welche von denselben oder einander entsprechenden charakteristischen Punkten definiert werden, gleich groß sind. Dies wird vorliegend als Symmetrieeigenschaft oder Symmetriebedingung bezeichnet. Ist diese Symmetriebedingung nicht erfüllt, kann davon ausgegangen werden, dass das Gesicht auf der Abbildung gedreht ist. Auf Grund der bekannten Symmetrieeigenschaft eines menschlichen Gesichts kann diese Drehung allerdings ausgeglichen werden, indem die Flächen bzw. Größen von symmetrischen Dreiecken in Relation zueinander gesetzt werden.

[0012] Die Flächen von zwei Dreiecken, welche symmetrisch mit Bezug auf eine Mittelachse des Gesichts angeordnet sind, können verwendet werden, um einen Quotienten zu bilden. Dieser Quotient sagt aus, um wie viel die Flächen dieser Dreiecke voneinander abweichen. Bei einem an sich symmetrischen Gegenstand, wie einem menschlichen Gesicht, müssten die Flächen gleich groß sein und der Quotient wäre 1. Ist das Gesicht jedoch gedreht, nimmt der Quotient einen anderen Wert an. Der anzuwendende Korrekturfaktor ist also so bemessen, dass die Fläche eines in der Abbildung verfälscht dargestellten Dreiecks auf seinen tatsächlichen Wert bzw. seinen tatsächlichen Anteil an der Gesamtfläche des Gesichts (hier geht es um den relativen Anteil eines Dreiecks an der Summe der Flächen aller Dreiecke, in welche ein Gesicht unterteilt wurde) korrigiert wird. Durch diesen Ansatz wird eine Drehung des Gesichts kompensiert, was eine automatisierte maschinelle Erkennung mittels eines künstlichen neuronalen Netzwerks erleichtert.

[0013] In einem weiteren Schritt wird die dergestalt korrigierte Fläche der Dreiecke auf eine drehinvariante Größe des Gesichts normiert. Als drehinvariante Größe kann beispielsweise ein horizontaler oder vertikaler Abstand zwischen zwei Punkten in dem Gesicht verwendet werden, welcher sich bei einer Drehung des Gesichts nicht verändert. Beispielsweise verändert sich der vertikale Abstand zwischen der Höhe der Augen und dem Zentrum der Oberlippe nicht, wenn der Kopf um die Hochachse gedreht wird. Auch andere vertikale Abstände in dem Gesicht ändern sich nicht, wenn der Kopf lediglich um die Hochachse gedreht wird. Somit können solche Größen, welche unabhängig von einer Drehung sind, herangezogen werden, um die Flächen der Dreiecke zu normieren.

[0014] Es ist also vorgesehen, eine zweistufige Korrektur eines Abbilds eines Gesichts auszuführen, um die Identifizierung einer Person mit einem künstlichen neuronalen Netzwerk zuverlässiger zu gestalten: zunächst wird das Gesicht in Dreiecke unterteilt, welche von charakteristischen Punkten begrenzt sind, und zwar so, dass mindestens einige der Dreiecke achsensymmetrisch bezogen auf eine Achse des Gesichts sind. Die zueinander achsensymmetrischen Dreiecke sind bei einem Abbild des Gesichts ohne Drehung gleich groß. Gilt diese Symmetriebedingung in einem Abbild nicht, ergibt sich, dass das Gesicht gedreht ist. Aus dem Größenverhältnis zueinander achsensymmetrischer Dreiecke kann ein Korrekturfaktor ermittelt werden, mithilfe dessen der relative Anteil der Fläche eines Dreiecks an der Summe der Flächen aller Dreiecke ermittelt werden

kann, und zwar für den Fall, dass das Gesicht nicht gedreht ist. In einem zweiten Schritt werden die korrigierten Flächen der Dreiecke auf eine drehinvariante Größe in dem Gesicht normiert. Durch diesen Ansatz wird ein korrigiertes und normiertes Abbild eines Gesichts erstellt, so dass für das Gesicht die Drehung kompensiert wird und es in eine gemeinsame Bildebene mit den trainierten Abbildern von Gesichtern gebracht wird.

[0015] Grundsätzlich ist das KNN aus einer Eingabeschicht, einer oder mehreren Zwischenschichten, und einer Ausgabeschicht aufgebaut. Es ist denkbar, dass das KNN ein zweischichtiges KNN ohne verdeckte Schicht ist. Jede Schicht enthält eine Mehrzahl von Neuronen, wobei jedes Neuron einer Schicht mit allen Neuronen der nächsten Schicht verbunden ist. Ein KNN kann mit einer Vielzahl von Eingabedaten trainiert werden und so konfiguriert sein, dass es einmal gelernte Objekte wiedererkennt und einen entsprechenden Ausgabewert liefert. Vorliegend wird das KNN mit Abbildungen von Gesichtern trainiert, welche frontal abgebildet sind und bei denen die Gesichtsgröße festgelegt ist. Auch diese Gesichter sind in eine Vielzahl von Dreiecken unterteilt, so dass die Flächen der Dreiecke, der relative Anteil der Fläche jedes Dreiecks an der Gesamtsumme der Flächen der Dreiecke, sowie die Normierung der Flächen der Dreiecke auf eine drehinvariante Größe des Gesichts durch das KNN als Identifizierungsparameter verwendet werden.

[0016] Das künstliche neuronale Netzwerk kann auf einem Rechner oder einem Rechnerverbund ausgeführt werden, wobei der Rechner oder der Rechnerverbund (dann mittels paralleler Abarbeitung der Anweisungen) das KNN in verschiedenen Konfigurationen ausführen kann. Beispielsweise kann das KNN auf einem Computer, einem programmierbaren Logikgatter (FPGA) oder einem Prozessor ausgeführt werden. Wenn das KNN in verschiedenen Konfigurationen ausgeführt wird, ändert dies nicht zwingend etwas an der Hardware, auf der das KNN ausgeführt wird. Vielmehr ändern sich beispielsweise die Konfiguration der einzelnen Neuronen und/oder die Gewichtung der Informationsübertragung zwischen Neuronen verschiedener Schichten.

[0017] Das KNN ist beispielsweise ausgestaltet, eine Mehrzahl verschiedener Konfigurationen einzunehmen, wobei jede Konfiguration des KNN einer trainierten Person entspricht und wobei das KNN einen bestimmten Eingabevektor mehreren Konfigurationen aus der Mehrzahl verschiedener Konfigurationen zuführt. Insbesondere ist die Information der einzelnen Konfigurationen in den Gewichten der einzelnen Neuronen enthalten und stellt ein Minimum auf einer mehrdimensionalen Hyperfläche dar. Dieses Minimum drückt die minimale Abweichung zwischen einem zu identifizierenden Gesicht und einem oder mehreren damit verglichenen Gesichtern dar. Bei sehr ähnlichen Gesichtern, wie es z.B. bei Zwillingen der Fall sein kann, kann es jedoch zu Verwechslungen kommen (auch als „false positive“ benannt). So kann die Ähnlichkeit des zu identifizierenden Abbilds mit verschiedenen dem KNN bekannten Gesichtern ermittelt werden. Das Gesicht mit der höchsten Ähnlichkeit (oder der geringsten Unähnlichkeit) wird dann von dem KNN als Ergebnis des Identifizierungsvorgangs ausgegeben.

[0018] Jede Konfiguration des KNN entspricht einer Person und ist darauf ausgelegt, dass das KNN diese Person wiedererkennt, und zwar unter verschiedenen Perspektiven auf das Gesicht der Person und bei variierenden Abständen der Person von der erfassenden Bilderfassungseinheit.

[0019] In einer Ausführungsform ist die drehinvariante Größe ein vertikaler oder horizontaler Abstand zwischen zwei charakteristischen Punkten des zu identifizierenden Objekts.

[0020] Je nach Drehung des Gesichts in dem Abbild sind entweder horizontale Abstände oder vertikale Abstände zwischen unveränderlichen Punkten von der Drehung nicht beeinflusst. Vorliegend können sowohl horizontale als auch vertikale Abstände als drehinvariante Größe verwendet werden.

[0021] In einer weiteren Ausführungsform ist die Objektidentifizierungseinheit konfiguriert, in einer Trainingsphase mit einer Mehrzahl von Abbildungen von Objekten trainiert zu werden, wobei die Abbildungen die Objekte ohne Drehwinkel zeigen.

[0022] Durch den oben beschriebenen Ansatz ist es möglich, das Abbild eines gedrehten Gesichts unter Verwendung eines künstlichen neuronalen Netzwerks mit solchen Bildern von Gesichtern zu vergleichen, welche frontal und ohne Drehung des Gesichts aufgenommen wurden.

[0023] In einer weiteren Ausführungsform ist die Symmetriebedingung eine Achsensymmetrie bezogen auf eine Achse des zu identifizierenden Objekts.

[0024] Der hier beschriebene Ansatz kann bevorzugt für solche Objekte verwendet werden, welche mindestens eine Achsensymmetrie haben. Das menschliche Gesicht beispielsweise ist in der Regel symmetrisch be-

zogen auf eine vertikale Mittelachse. Diese Symmetrieeigenschaft erlaubt es, Drehungen des Gesichts um eine Hochachse zu erkennen, weil die Flächen von Dreiecken auf der linken Gesichtshälfte von den Flächen von entsprechenden Dreiecken auf der rechten Gesichtshälfte (oder umgekehrt) abweichen, obwohl sie aufgrund der Symmetrieeigenschaft gleich sein müssten.

[0025] Für Objekte, welche achsensymmetrisch bezogen auf eine horizontale Achse sind, gilt die gleiche Aussage für die obere und die untere Hälfte des Abbilds des Objekts. Hierbei kann eine Drehung des Objekts um die Querachse kompensiert werden. Als Beispiel für diesen Anwendungsfall kann ein Flugkörper oder Lenkflugkörper genannt werden, der von der Seite betrachtet wird und bei dem das Höhenruder nicht angestellt ist.

[0026] In einer weiteren Ausführungsform ist die Objektidentifizierungseinheit konfiguriert, die zweidimensionale Abbildung so in eine Mehrzahl von Dreiecken zu unterteilen, dass mindestens einige der Dreiecke symmetrisch zu einer Achse des zu identifizierenden Objekts sind.

[0027] Bevorzugt wird eine vollständig symmetrische Aufteilung vorgenommen. Dies mag hilfreich und vorteilhaft sein. Es ist jedoch auch möglich, ohne eine vollständige symmetrische Unterteilung des Objekts eine Identifikation auszuführen, indem auf die symmetrischen Dreiecke zugegriffen wird.

[0028] In einer weiteren Ausführungsform ist die Objektidentifizierungseinheit ausgeführt, den Korrekturfaktor zu ermitteln, indem für jeweils ein Dreieckspaar bestehend aus zwei einander entsprechenden Dreiecken ein Größenverhältnis ermittelt wird.

[0029] Somit wird ein virtuelles Abbild bzw. werden Dreiecksgrößen der Abbildung ohne Drehwinkel erstellt.

[0030] Gemäß einem weiteren Aspekt ist ein Verfahren zum Identifizieren von Objekten mit einem künstlichen neuronalen Netzwerk, KNN, angegeben. Das Verfahren weist die folgenden Schritte auf: Unterteilen einer zweidimensionalen Abbildung eines zu identifizierenden Objekts in eine Mehrzahl von Dreiecken, die durch charakteristische Punkte des zu identifizierenden Objekts an ihren Ecken begrenzt sind; Ermitteln eines Drehwinkels des zu identifizierenden Objekts in der zweidimensionalen Abbildung basierend auf Symmetrieeigenschaften des zu identifizierenden Objekts und relativen Flächenanteilen der einzelnen Dreiecke an einer Gesamtsumme der Fläche der Dreiecke; Ermitteln von korrigierten Flächen der Dreiecke, indem die Flächen der Dreiecke mit einem Korrekturfaktor multipliziert werden, wobei die korrigierten Flächen der Dreiecke einer Symmetriebedingung entsprechen; Normieren der korrigierten Flächen der Dreiecke auf eine drehinvariante Größe des zu identifizierenden Objekts; und Zuführen der normierten korrigierten Flächen an das KNN, um basierend darauf einen Abgleich von Objekten vorzunehmen und diese zu identifizieren.

[0031] Für das Verfahren gelten die gleichen Ausführungen, wie sie mit Bezug zu der Objektidentifizierungseinheit weiter oben oder unten dargelegt sind. Um unnötige Wiederholung zu vermeiden, werden die funktionalen Merkmale der Objektidentifizierungseinheit an dieser Stelle nicht wiedergegeben. Jedenfalls können die funktionalen Merkmale der Objektidentifizierungseinheit als Verfahrensschritte implementiert werden und umgekehrt. Beispielhaft werden hier einige der funktionalen Merkmale als Verfahrensschritte angegeben, was jedoch nicht als Einschränkung dahingehend verstanden werden darf, dass andere funktionale Merkmale nicht als Verfahrensschritte implementiert werden können.

[0032] In einer Ausführungsform des Verfahrens ist die drehinvariante Größe ein vertikaler oder horizontaler Abstand zwischen zwei charakteristischen Punkten des zu identifizierenden Objekts.

[0033] Ein horizontaler Abstand bei einer Drehung des Objektes/Gesichts um die Hochachse ist allerdings erst nach einer Korrektur des Drehwinkels um die Hochachse drehinvariant. Allgemein ausgedrückt ist ein vertikaler Abstand drehinvariant gegenüber Drehungen um die Hochachse und ein horizontaler Abstand drehinvariant gegenüber Drehungen um die Querachse.

[0034] In einer weiteren Ausführungsform ist die Symmetriebedingung eine Achsensymmetrie bezogen auf eine Achse des zu identifizierenden Objekts.

Figurenliste

[0035] Nachfolgend wird anhand der beigefügten Zeichnungen näher auf Ausführungsbeispiele eingegangen. Die Darstellungen sind schematisch und nicht maßstabsgetreu. Gleiche Bezugszeichen beziehen sich auf gleiche oder ähnliche Elemente. Es zeigen:

Fig. 1 eine schematische Darstellung eines künstlichen neuronalen Netzwerks für eine Objektidentifizierungseinheit zum Identifizieren von Objekten.

Fig. 2 eine schematische Darstellung einer Objektidentifizierungseinheit.

Fig. 3 eine schematische Darstellung eines menschlichen Gesichts mit charakteristischen Punkten und dadurch definierte Dreiecke.

Fig. 4 eine Unterteilung des Gesichts aus **Fig. 3** in eine Vielzahl von Dreiecken.

Fig. 5 eine schematische Darstellung eines künstlichen neuronalen Netzwerks;

Fig. 6 eine schematische Darstellung der Drehung eines Abbildes eines menschlichen Gesichts;

Fig. 7 eine schematische Darstellung eines Verfahrens zum Identifizieren von Objekten mit einem künstlichen neuronalen Netzwerk.

Detaillierte Beschreibung von Ausführungsbeispielen

[0036] **Fig. 1** zeigt eine schematische Darstellung eines künstlichen neuronalen Netzwerks, KNN, 100. Das KNN 100 weist eine Eingabeschicht 110 mit mehreren Eingabeneuronen 111 bis 114 auf. Beispielsweise enthält die Eingabeschicht 110 eine Anzahl von Eingabeneuronen, an welche Identifizierungsparameter eines Gesichts übertragen werden. Als Identifizierungsparameter werden die korrigierten und normierten Dreiecksflächen verwendet, wie weiter oben beschrieben. Das KNN kann ohne Zwischenschicht oder mit einer oder mehreren Zwischenschichten ausgestattet sein. Vorliegend sind zwei Zwischenschichten 120A und 120B gezeigt. Weiterhin enthält das KNN 100 eine Ausgabeschicht 130 mit Ausgabeneuronen 131 bis 134.

[0037] Es ist ein Kennzeichen des KNN, dass jedes Neuron einer Schicht mit allen Neuronen der nachfolgenden (in der Darstellung der **Fig. 1** weiter rechts befindlichen) Schichten verbunden ist. In einer Trainingsphase werden der Eingabeschicht des KNN Objektidentifizierungsdaten zugeführt und das KNN wird darauf trainiert, um die zugeführten Objektidentifizierungsdaten sämtlicher Objekte zu erkennen und den jeweiligen Personen zuzuweisen. Die Objektidentifizierungsdaten sind in den Gewichten des KNN enthalten und äußern sich als Minima auf einer mehrdimensionalen Hyperfläche, wobei jedem der trainierten Referenzbilder genau ein (Haupt-)Minimum entspricht. Bei der Identifizierung wird eines dieser Minima durch Backpropagation angesteuert, und zwar das Gesicht der gesuchten bzw. zu identifizierenden Person. Basierend auf diesem Training werden die Neuronen so vernetzt, dass sie ein einmal gelerntes Objekt auch in anderen Objektidentifizierungsdaten erkennen und an der Ausgabeschicht 130 einen entsprechenden Ausgabewert liefern.

[0038] **Fig. 2** zeigt eine Objektidentifizierungseinheit 1 mit einer Datenquelle 50, einem künstlichen neuronalen Netzwerk 100, wie beispielsweise in **Fig. 1** gezeigt, einer Ausgabeeinheit 60, einem Datenspeicher 70 und einer Steuereinheit 80.

[0039] Die Datenquelle 50 liefert Bilder von einem oder mehreren Gesichtern an das KNN 100, wo die Bilder einem Identifizierungsvorgang unterlaufen. Die Datenquelle 50 kann beispielsweise eine Kamera sein oder ein Datenspeicher, in welchem Bilder hinterlegt sind und an das KNN 100 ausgegeben werden können. Die Datenquelle 50 kann mit dem KNN 100 direkt oder über ein Datenübertragungsnetz verbunden sein. Die Übertragung der Bilder von der Datenquelle 50 an das KNN 100 kann von der Steuereinheit 80 initiiert oder gesteuert werden.

[0040] Das KNN 100 kann auf einen Datenspeicher 70 zugreifen. In dem Datenspeicher 70 können beispielsweise die Parameter von verschiedenen Konfigurationen des KNN 100 abgelegt sein. Mithilfe dieser Konfigurationen können die Gewichtungen für die einzelnen Neuronen geladen werden, um ein Bild einer Person mit dem Bild aller trainierten Personen abzugleichen. Das Ergebnis des Identifizierungsvorgangs wird von dem KNN 100 an die Ausgabeeinheit 60 übermittelt.

[0041] **Fig. 3** zeigt beispielhaft die Festlegung von charakteristischen Punkten in einem menschlichen Gesicht. Die charakteristischen Punkte sind jeweils mit einem Kreis 151 gekennzeichnet. Einige dieser Kreise 151 sind über Linien miteinander verbunden, so dass diese Liniendreiecke definieren, von denen manche für das

Identifizieren einer Person herangezogen werden können. Weiterhin sind in **Fig. 3** einige vertikale Abstände **152** gezeigt. Wird das Gesicht in **Fig. 3** um die vertikale Achse **155** gedreht, so ändern sich die vertikalen Abstände **152** nicht.

[0042] **Fig. 4** zeigt das Gesicht aus **Fig. 3** mit hervorgehobenen Dreiecken und nummerierten Eckpunkten dieser Dreiecke. Die Auswahl und Festlegung der Dreiecke in **Fig. 4** ist lediglich als Beispiel zu verstehen. Dreiecke können auf andere Weise und in einer anderen Anzahl in das Gesicht bzw. die Abbildung des Gesichts gelegt werden.

[0043] In **Fig. 4** ist zu erkennen, dass einige der Dreiecke in den beiden Gesichtshälften bezogen auf die vertikale Mittelachse **155** des Gesichts symmetrisch zueinander sind. So ist beispielsweise das durch die Punkte 2, 3, 4 definierte Dreieck eine Achsenspiegelung des durch die Punkte 16, 17, 18 definierten Dreiecks (diese beiden an der Mittelachse gespiegelten Dreiecke werden als einander entsprechende Dreiecke bezeichnet). Ähnliche Symmetrieeigenschaften finden sich auch bei anderen Dreiecken. Einander entsprechende Dreiecke sind durch gleiche Form und Fläche gekennzeichnet (d.h. diese Dreiecke werden durch eine Drehung um 180° aufeinander abgebildet) und finden sich in vielen oder gar den meisten menschlichen Gesichtern.

[0044] Im Falle eines um die Hochachse gedrehten Gesichts sind die Flächen einander entsprechender Dreiecke in einer zweidimensionalen Abbildung des gedrehten Gesichts nicht gleich. Wenn das Gesicht in **Fig. 4** beispielsweise leicht nach links gedreht wird, wird die Fläche des Dreiecks 16, 17, 18 kleiner als die Fläche des Dreiecks 2, 3, 4. Das Verhältnis dieser beiden Flächen kann verwendet werden, um den Drehwinkel des Gesichts zu bestimmen. Um die Zuverlässigkeit bzw. Genauigkeit der Bestimmung des Drehwinkels des Gesichts zu erhöhen, können mehrere oder alle einander entsprechenden Dreiecke der beiden Gesichtshälften herangezogen und in ihrer Größe miteinander verglichen werden. Hierbei wird bevorzugt nicht die absolute Fläche der Dreiecke verwendet, sondern der Anteil der Fläche eines Dreiecks an der Summe der Fläche aller Dreiecke, welche in das menschliche Gesicht gelegt sind.

[0045] Bezugnehmend auf **Fig. 3** und **Fig. 4** wird an dieser Stelle der Ansatz für die Identifizierung von Personen beispielhaft beschrieben.

[0046] Zunächst wird das Gesicht in eine endliche Anzahl elementarer Grunddreiecke zerlegt, deren Eckpunkte zueinander in einem für jeden Menschen eigentümlichen Abstandsverhältnis stehen. Der Idee einer Gesichtserkennung liegt der Gedanke zugrunde, dass es mit Ausnahme von eineiigen Zwillingen keine zwei gleichen Gesichter gibt. Vergleichen wir zwei Gesichter miteinander, so stellen wir fest, dass Menschen unterschiedlich lange Nasen, unterschiedliche Augenabstände oder unterschiedliche Mundbreiten haben. Diese Merkmale werden biometrische Daten genannt. Ein Beispiel dafür liefert **Fig. 3** und **Fig. 4**.

[0047] So lassen sich etwa anhand markanter Punkte wie Augen- und Nasenmitte, Kinnschuppe etc. oder charakteristischer Abmessungen wie Ohrenabstand, Nasen- oder Oberlippenlänge Linien durchs Gesicht ziehen, die ein biometrisches Profil eines Menschen ergeben und somit ein unveränderliches Körpermerkmal darstellen. Wenn man eine charakteristische Abmessung, etwa die Distanz der Augenlinie (horizontale Linie, welche die Augen miteinander verbindet) zur Nasenlinie (horizontale Linie, welche die Nasenlöcher miteinander verbindet) als Normierungsgröße verwendet, werden sich sämtliche Menschen in mindestens einem, aber höchstwahrscheinlich in mehreren, möglicherweise sogar vielen Merkmalen von der Referenzperson unterscheiden.

[0048] Wählen wir nun eine endliche Anzahl charakteristischer Punkte im Gesicht (Kreise **151** in **Fig. 3**), so lassen sich aus jeweils drei Punkten Dreiecksflächen bilden, die eine eindeutig definierte Flächennormale besitzen. Die Darstellung eines Gesichts ist nichts anderes als die Projektion einer räumlich gekrümmten Oberfläche in die Bildebene, wobei zwischen der dreidimensionalen Flächennormalen A und der Flächennormalen der projizierten Fläche A_n der Winkel θ liegt. Jedes Flächenelement auf einem Foto ist also um den Kosinus des Winkels, den der Normalenvektor auf der dreidimensionalen Gesichtsoberfläche mit dem Lot senkrecht zur Bildebene einnimmt, kleiner als die tatsächliche Oberfläche des dreidimensionalen Gesichts.

[0049] Da ein menschliches Gesicht mehrfach und unterschiedlich gekrümmt ist, wird es zunächst in eine ausreichende Zahl von Flächenelementen zerlegt. Als Zerlegung der Gesichtsfläche wird ohne Beschränkung der Allgemeinheit die in **Fig. 4** dargestellte Anordnung aus 25 Dreiecken, die nahtlos aneinander anschließen, gewählt.

[0050] Dabei ist jeder der insgesamt 27 Eckpunkte wenigstens einmal Ausgangs- oder Endpunkt einer der 25 Dreiecksflächen $A_{i,i+1,i+2}$, die in aufsteigender Folge wie folgt angeordnet werden können:

$$A_{123}, A_{234}, A_{345}, \dots, A_{j,j+1,j+2}, \dots, A_{m,m+1,m+2}.$$

[0051] Um also eine Fläche in m Dreiecke zerlegen zu können, werden $m+2$ Eckpunkte benötigt. In dem vorliegenden Fall ist $m = 25$. Hinsichtlich der Zerlegung ist auch noch bevorzugt, diese so zu wählen, dass eine fortlaufende Reihe von Punkten entsteht, bei denen möglichst keiner ausgelassen wird oder einer nicht Eckpunkt eines Grunddreiecks ist. Doppelzählungen sind ausdrücklich erlaubt, bisweilen sogar unvermeidbar. **Fig. 4** zeigt eine Zerlegung mit 18 Eckpunkten und 9 weiteren Punkten, die doppelt vergeben wurden und links von dem Gesicht angeführt sind.

[0052] Sämtliche Teildreiecke der Bildebene addieren sich nun nach folgender Formel zu einer Gesamtfläche:

$$A = \sum_{j=1}^m A_{j,j+1,j+2},$$

[0053] Die Gesamtfläche wird als Bezugsgröße gebraucht, weil die Neuronen zu Vergleichszwecken als Werte kleiner oder gleich 1 behandelt werden. Für die normierten Dreiecksflächen gilt dann der Ausdruck

$$a_{j,j+1,j+2} \equiv \frac{A_{j,j+1,j+2}}{A}.$$

[0054] Zur Berechnung einer Dreiecksfläche im dreidimensionalen Raum benötigen wir die Flächennormale, die sich als Kreuzprodukt der beiden Abstandsvektoren $\Delta \mathbf{r}_{ij}$ und $\Delta \mathbf{r}_{ik}$ von einem gemeinsamen Eckpunkt i zu jedem der zwei anderen Eckpunkte j und k im Dreieck darstellen lässt, d.h.

$$\mathbf{A}_{ijk} = \frac{1}{2} \Delta \mathbf{r}_{ij} \times \Delta \mathbf{r}_{ik} = \frac{1}{2} \begin{pmatrix} \Delta x_{ij} \\ \Delta y_{ij} \\ \Delta z_{ij} \end{pmatrix} \times \begin{pmatrix} \Delta x_{ik} \\ \Delta y_{ik} \\ \Delta z_{ik} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \Delta y_{ij} \Delta z_{ik} - \Delta z_{ij} \Delta y_{ik} \\ \Delta z_{ij} \Delta x_{ik} - \Delta x_{ij} \Delta z_{ik} \\ \Delta x_{ij} \Delta y_{ik} - \Delta y_{ij} \Delta x_{ik} \end{pmatrix}$$

[0055] Für ebene Vektoren in der x-y-Ebene verbleibt von diesem Vektor nur die z-Komponente

$$A_{ijk} = \frac{1}{2} |\Delta \mathbf{r}_{ij} \times \Delta \mathbf{r}_{ik}| = \frac{1}{2} \left| \begin{pmatrix} \Delta x_{ij} \\ \Delta y_{ij} \\ 0 \end{pmatrix} \times \begin{pmatrix} \Delta x_{ik} \\ \Delta y_{ik} \\ 0 \end{pmatrix} \right| = \frac{1}{2} |\Delta x_{ij} \Delta y_{ik} - \Delta y_{ij} \Delta x_{ik}|,$$

wobei Betragsstriche gesetzt sind, da Flächen auch negativ sein können. Im Einzelnen gelten für die jeweiligen Abstandsdifferenzen die Relationen

$$\Delta x_{ij} = x_j - x_i, \quad \Delta x_{ik} = x_k - x_i,$$

$$\Delta y_{ij} = y_j - y_i, \quad \Delta y_{ik} = y_k - y_i,$$

$$\Delta z_{ij} = z_j - z_i, \quad \Delta z_{ik} = z_k - z_i.$$

[0056] Um einer Vielzahl möglicher Gesichter Rechnung zu tragen, führen wir außerdem für jedes einzelne Gesicht ein Trainingsmuster mit dem Index p ein, was schematisch in **Fig. 5** gezeigt ist.

[0057] Um Laufindizes einzusparen, definieren wir die Ausgangsneuronen α_{pj} , sowie die Sollausgangswerte $\hat{\alpha}_{pj}$ zunächst durch

$$a_{pj} \equiv a_{p,j,j+1,j+2} = \frac{A_{p,j,j+1,j+2}}{A_p} = \frac{|\Delta x_{p,j,j+1} \Delta y_{p,j,j+2} - \Delta y_{p,j,j+1} \Delta x_{p,j,j+2}|}{\sum_{j=1}^m A_{p,j,j+1,j+2}},$$

$$\hat{a}_{pj} \equiv \hat{a}_{p,j,j+1,j+2} = \frac{\hat{A}_{p,j,j+1,j+2}}{\hat{A}_p} = \frac{|\Delta \hat{x}_{p,j,j+1} \Delta \hat{y}_{p,j,j+2} - \Delta \hat{y}_{p,j,j+1} \Delta \hat{x}_{p,j,j+2}|}{\sum_{j=1}^m \hat{A}_{p,j,j+1,j+2}}$$

für $j = 1, 2, \dots, m$. In dem zweischichtigen Netz aus **Fig. 5** gilt also α_{pi} der Input des Neurons i in der Eingabeschicht, wobei $1 \leq i \leq m$, a_{pj} der Output des Neurons j in der Ausgabeschicht, wobei $1 \leq j \leq m$, net_{pj} der Netto-Input des Neurons j in der Ausgabeschicht aus Schicht 1, wobei $1 \leq j \leq m$, w_{ij} das Verbindungsgewicht von Neuron i nach j an der Schnittstelle.

[0058] Nachdem das Netz mit einer Vielzahl von Gesichtern angelernt wurde (**Fig. 4** gibt ein Beispiel für ein in das Netz eingespeistes bzw. angelerntes Gesicht), kann es auf andere konkrete Fälle angewandt werden.

$$a_{pj} = \frac{A_{pj}^\perp}{A_p^\perp} = \begin{cases} \frac{A_{pj} \cos(\theta_{pj} - \theta_p)}{\sum_{j=1}^m A_{pj}^\perp} & \text{für } j \in \{1, 2, 3, 4, 5, 6, 7, 8\}, \\ \frac{A_{pj} \cos \theta_{pj}}{\sum_{j=1}^m A_{pj}^\perp} & \text{für } j \in \{9, 19, 20, 21, 22, 23, 24, 25\}, \\ \frac{A_{pj} \cos(\theta_{pj} - \theta_p)}{\sum_{j=1}^m A_{pj}^\perp} & \text{für } j \in \{10, 11, 12, 13, 14, 15, 16, 17\}, \end{cases}$$

wobei $A_{p,18}^\perp = 0$. Zur expliziten Unterscheidung von den dreidimensionalen Flächenelementen haben wir die zweidimensionalen Projektionen mit A_{pj}^\perp anstelle von A_{pj} , bezeichnet. Ferner gilt

$$\hat{a}_{pj} = \frac{\hat{A}_{pj}^\perp}{\hat{A}_p^\perp} = \frac{\hat{A}_{pj} \cos \theta_{pj}}{\sum_{j=1}^m \hat{A}_{pj}^\perp}$$

mit $\hat{A}_{p,18}^\perp = 0$ und \hat{A}_{pj}^\perp anstelle von A_{pj} .

[0059] **Fig. 6** zeigt am Beispiel zweier Flächenelemente 1 und 17 anschaulich, wie sich die Auswirkungen einer Kopfdrehung um die Hochachse kompensieren lassen.

[0060] Das linke Halbbild weist für den Winkel $\theta = 0$ eine vollkommene Symmetrie auf, wogegen sich die Kopfdrehung im rechten Halbbild in unterschiedlich großen Flächendreiecken rechts und links der vertikalen Symmetrieachse des Gesichts auswirkt. Bei einer Linksdrehung des Kopfes müssen der Projektionswinkel θ_j der Flächennormalen und der Drehwinkel θ voneinander subtrahiert werden, da die Dreiecksflächen dem Betrachter zugewandt sind und rechts von der Symmetrieachse liegen. Sind sie hingegen links von der Symmetrieachse, müssen sie addiert werden. Die Kopfdrehung der Soll-Bilder kann daher durch Mittelwertbildung von zusammengehörigen Flächenelementen rechts und links der Gesichtsmitte rechnerisch eliminiert werden. Flächen wiederum, die kein Pendant haben, bleiben ungemittelt, wobei dies nur dann erfolgt, wenn ein Gesicht nicht vollständig in zueinander symmetrische Dreiecke unterteilt wurde. In **Fig. 6** ist die Geometrie der

Kopfbewegung aus der Draufsicht von oben dargestellt, also mit dem Gesicht nach unten. Das Verfahren wird nachfolgend beschrieben.

[0061] Dazu wird im Folgenden der Index p weggelassen, weil nur ein einziges Bild betrachtet wird. Für $j \in \{1, 2, 3, 4, 5, 6, 7, 8\}$ gilt im Falle der biometrischen Draufsicht ($\theta = 0$) für die Referenzflächen

$$\hat{a}_j = \hat{a}_{18-j} = \frac{\hat{A}_j^\perp + \hat{A}_{18-j}^\perp}{2\hat{A}^\perp} = \frac{\hat{A}_j \cos \theta_j + \hat{A}_{18-j} \cos \theta_{18-j}}{2\hat{A}^\perp} = \frac{\hat{A}_j \cos \theta_j}{\hat{A}^\perp}$$

bzw.

$$\hat{a}_j = \frac{\hat{A}_j^\perp}{\hat{A}^\perp} = \frac{\hat{A}_j \cos \theta_j}{\hat{A}^\perp} \text{ für } j \in \{9, 18, 19, 20, 21, 22, 23, 24, 25\}.$$

[0062] Weil davon ausgegangen werden kann, dass, wie in **Fig. 6.** dargestellt, $\theta_j = \theta_{18-j}$ und $A_j = A_{18-j}$, können die symmetrischen Flächen für $j \in \{1, 2, 3, 4, 5, 6, 7, 8\}$ im allgemeinen Fall einer leichten Schrägsicht gemittelt werden, d.h.

$$a_j = a_{18-j} = \frac{A_j^\perp + A_{18-j}^\perp}{2A^\perp} = \frac{A_j \cos(\theta_j - \theta) + A_{18-j} \cos(\theta_{18-j} + \theta)}{2A^\perp} = \frac{A_j \cos \theta_j \cos \theta}{A^\perp}.$$

[0063] Für die asymmetrischen Flächen gilt stattdessen

$$a_j = \frac{A_j^\perp}{A^\perp} = \frac{A_j \cos \theta_j \cos \theta}{A^\perp} \text{ für } j \in \{9, 18, 19, 20, 21, 22, 23, 24, 25\}.$$

[0064] Wenn beide Gleichungen durcheinander dividiert werden, werden Relationen erhalten, aus denen der Drehwinkel θ bestimmt werden kann,

$$\cos \theta = \frac{A_j^\perp + A_{18-j}^\perp}{\hat{A}_j^\perp + \hat{A}_{18-j}^\perp} \frac{\hat{A}_j}{A_j} \text{ für } j \in \{1, 2, 3, 4, 5, 6, 7, 8\},$$

$$\cos \theta = \frac{A_j^\perp}{\hat{A}_j^\perp} \frac{\hat{A}_j}{A_j} \text{ für } j \in \{9, 18, 19, 20, 21, 22, 23, 24, 25\}.$$

[0065] Für normierte Gesichtsflächen ist $A_j = A_j$ usw. Damit können die Winkel direkt mittels der Relationen

$$\cos \theta = \frac{a_j A^\perp}{\hat{a}_j \hat{A}^\perp} = \frac{A_j^\perp + A_{18-j}^\perp}{\hat{A}_j^\perp + \hat{A}_{18-j}^\perp} \text{ bzw. } \cos \theta = \frac{A_j^\perp}{\hat{A}_j^\perp}$$

berechnet werden. Zur Normierung der Bilder wird der Abstand verwendet

$$h_0 = \left| \frac{\hat{\mathbf{r}}_{15} + \hat{\mathbf{r}}_5}{2} - \frac{\hat{\mathbf{r}}_{13} + \hat{\mathbf{r}}_7}{2} \right| = \frac{1}{2} \sqrt{(\hat{x}_{15} + \hat{x}_5 - \hat{x}_{13} - \hat{x}_7)^2 + (\hat{y}_{15} + \hat{y}_5 - \hat{y}_{13} - \hat{y}_7)^2}$$

zwischen Augen- und Nasenmitte aus dem Referenzbild. Im tatsächlich vorgelegten Sollbild lautet dieser Abstand beispielsweise

$$h = \left| \frac{r_{15} + r_5}{2} - \frac{r_{13} + r_7}{2} \right| = \frac{1}{2} \sqrt{(x_{15} + x_5 - x_{13} - x_7)^2 + (y_{15} + y_5 - y_{13} - y_7)^2}.$$

[0066] Um nun ein tatsächlich auszuwertendes Bild zu normieren, müssen alle Positionsdaten mit dem Faktor $\kappa = h_0/h$ multipliziert werden, d.h. $\kappa < 1$, wenn $h > h_0$, und $\kappa \geq 1$, wenn $h < h_0$. So ergibt sich beispielsweise für eine Vergleichsperson ein κ von 0,7173 und damit ein korrigierter Fehler von $E = 0,0012898$, während das κ von einem anderen Bild gleich 1,0139 ist und somit einem Fehler von $E = 0,0018017$ entspricht. Das Bild mit dem höheren Fehler weicht also von der Vergleichsperson um 39,7 % ab, was eine deutliche Unterscheidung ermöglicht.

[0067] Nachfolgend wird noch der Trainingsalgorithmus beschrieben.

[0068] Ausgangspunkt für die Herleitung eines Backpropagation-Algorithmus ist die Minimierungsvorschrift des Lernfehlers

$$E_p = \frac{1}{2} \sum_{j=1}^m (\hat{a}_{pj} - a_{pj})^2$$

[0069] in einem Muster p . Der nun folgende Ansatz basiert auf der Überlegung, dass die Berechnung des Gradienten für die Gesamtfehlerfunktion E durch Summation der Gradienten über alle Fehlerfunktionen E_p der einzelnen Trainingsmuster erfolgen kann, da der Gesamtfehler E nichts anderes ist als die zu minimierende Summe

$$E = \sum_{p=1}^q E_p = \frac{1}{2} \sum_{p=1}^q \sum_{j=1}^m (\hat{a}_{pj} - a_{pj})^2 = \min$$

aller q Lernfehler. Da der Lernfehler E abnehmend ist, d.h. $\partial E / \partial w_i < 0$, muss für $i = 1, 2, \dots, n$ gelten:

$$\Delta w_i = -\gamma \frac{\partial E}{\partial w_i} = -\gamma \sum_{p=1}^q \frac{\partial E_p}{\partial w_i},$$

wobei der Proportionalitätsfaktor γ die Lernrate angibt. Die Modifikationsregel für alle n Verbindungsgewichte kann demnach geschrieben werden als

$$\Delta \mathbf{w} = -\gamma \nabla E(\mathbf{w}) = -\gamma \left(\frac{\partial E}{\partial w_1}, \frac{\partial E}{\partial w_2}, \dots, \frac{\partial E}{\partial w_n} \right) = -\gamma \left(\sum_{p=1}^q \frac{\partial E_p}{\partial w_1}, \sum_{p=1}^q \frac{\partial E_p}{\partial w_2}, \dots, \sum_{p=1}^q \frac{\partial E_p}{\partial w_n} \right).$$

[0070] Die Änderung des Gewichtsvektors ist folglich proportional zum negativen Gradienten der Fehlerfunktion. Der Netto-Input der Ausgangsneuronen errechnet sich gemäß der linearen Standard-Propagierungsfunktion zu

$$net_{pj} = \sum_{i=1}^m w_{ij} a_{pi},$$

wobei w_{ij} das Gewicht zwischen dem sendenden Neuron i und dem empfangenden Neuron j ist. Die Besonderheit von mehrlagigen Perzeptren liegt in der Verwendung einer nichtlinearen Sigmoidfunktion

$$a_{pj} = f(net_{pj}) = f\left(\sum_{i=1}^m w_{ij} a_{pi}\right) = \frac{1}{1 + e^{-net_{pj}}}$$

zur Aktivierung der Neuronen. Der Fehler im Trainingsmuster p über alle Ausgabeneuronen muss also nach Einsetzen der Aktivierungsfunktion minimal werden, d.h.

$$E_P = \frac{1}{2} \sum_{j=1}^m (\hat{a}_{pj} - a_{pj})^2 = \frac{1}{2} \sum_{j=1}^m \left(\hat{a}_{pj} - f \left(\sum_{i=1}^m w_{ij} a_{pi} \right) \right)^2 \rightarrow \min.$$

[0071] Verbal lässt sich das Lernproblem etwa dadurch beschreiben, dass die Verbindungsgewichte w_{ij} so bestimmt werden sollen, dass die quadratische Gesamtfehlerdifferenz zwischen Soll- und Ist-Ausgabe (wobei letztere eine nichtlineare Funktion des gewichteten Netto-Inputs darstellt) für alle Ausgangsneuronen minimal wird. Mithin ist also der euklidische Abstand zwischen Zielvektor und Ausgangsvektor zu minimieren. Die notwendige Bedingung zur Lösung der Extremwertaufgabe hat zur Konsequenz, dass der Gradient der Fehlerfunktion null werden muss, d.h.

$$\nabla E(W) = \left(\frac{\partial E}{\partial w_{11}}, \frac{\partial E}{\partial w_{12}}, \dots, \frac{\partial E}{\partial w_{mm}} \right) = 0.$$

[0072] Hierbei können die Koeffizienten für sämtliche modifizierbaren Verbindungsgewichte übersichtlicher als Konnektionsmatrix dargestellt werden:

$$W = \begin{pmatrix} w_{11} & \dots & w_{1j} & \dots & w_{1m} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ w_{i1} & \dots & w_{ij} & \dots & w_{im} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ w_{m1} & \dots & w_{mj} & \dots & w_{mm} \end{pmatrix}.$$

[0073] Dem Gradientenabstiegsverfahren liegt nun die Idee zugrunde, die Veränderung der einzelnen Verbindungsgewichte proportional zum negativen Gradienten der Fehlerfunktion vorzunehmen:

$$\begin{aligned} \frac{\partial E_p}{\partial w_{ij}} &= \frac{1}{2} \sum_{l=1}^m \frac{\partial E_p}{\partial (\hat{a}_{pl} - a_{pl})} \frac{\partial}{\partial \text{net}_{pl}} (\hat{a}_{pl} - f(\text{net}_{pl})) \frac{\partial \text{net}_{pl}}{\partial w_{ij}} \\ &= \sum_{l=1}^m \frac{\partial E_p}{\partial a_{pl}} \frac{\partial f(\text{net}_{pl})}{\partial \text{net}_{pl}} \frac{\partial \text{net}_{pl}}{\partial w_{ij}}, \end{aligned}$$

wobei die Beziehung

$$\frac{\partial E_p}{\partial (\hat{a}_{pl} - a_{pl})} = - \frac{\partial E_p}{\partial a_{pl}} = \hat{a}_{pl} - a_{pl}$$

verwendet wurde. Wäre die Aktivierungsfunktion linear, so ließe sich jedes mehrschichtige Backpropagation-Netz auf ein zweistufiges Netz ohne verdeckte Schicht reduzieren, da eine hintereinandergeschaltete lineare Propagierungs- und Aktivierungsfunktion zu einer einzigen linearen Funktion zusammengefasst werden können.

[0074] Die zur Erfüllung der Aufgabenstellung benötigte partielle Ableitung der obigen Aktivierungsfunktion lautet

$$\frac{\partial f(\text{net}_{pl})}{\partial \text{net}_{pl}} = \frac{e^{-\text{net}_{pl}}}{(1 + e^{-\text{net}_{pl}})^2} = f^2(\text{net}_{pl}) \left(\frac{1}{f(\text{net}_{pl})} - 1 \right) = f(\text{net}_{pl}) (1 - f(\text{net}_{pl})).$$

Wird nun die Eingabe an das KNN

$$net_{pl} = \sum_{k=1}^m w_{kl} a_{pk}$$

nach w_{ij} abgeleitet, wird schließlich erhalten

$$\begin{aligned} \frac{\partial net_{pl}}{\partial w_{ij}} &= \sum_{k=1}^m \frac{\partial}{\partial w_{ij}} a_{kl} a_{pk} = \sum_{k=1}^m \frac{\partial w_{kl}}{\partial w_{ij}} \frac{\partial}{\partial w_{kl}} w_{kl} a_{pk} \\ &= \sum_{k=1}^m \frac{\partial w_{kl}}{\partial w_{ij}} \frac{\partial}{\partial w_{kl}} \sum_{i=1}^m w_{il} a_{pi} = \sum_{k=1}^m \frac{\partial w_{kl}}{\partial w_{ij}} \frac{\partial net_{pl}}{\partial w_{kl}}. \end{aligned}$$

[0075] Wenn nun alle bislang erhaltenen Ausdrücke in den Ausdruck für den Gradienten der obigen Fehlerfunktion eingesetzt werden, lautet dieser am Ende

$$\frac{\partial E_p}{\partial w_{ij}} = - \sum_{l=1}^m (\hat{a}_{pl} - a_{pl}) f(net_{pl}) (1 - f(net_{pl})) \sum_{k=1}^m \frac{\partial w_{kl}}{\partial w_{ij}} \frac{\partial net_{pl}}{\partial w_{kl}}.$$

Zusammengefasst ergibt sich mit

$$\frac{\partial net_{pl}}{\partial w_{kl}} = a_{pk} \quad \text{und} \quad \frac{\partial w_{kl}}{\partial w_{ij}} = 1,$$

und weil $i = k$ und $j = l$ ist, der Ausdruck

$$\frac{\partial E_p}{\partial w_{ij}} = - (\hat{a}_{pj} - a_{pj}) a_{pj} (1 - a_{pj}) a_{pi}.$$

Als Gewichtsadaption wird schließlich erhalten

$$\Delta w_{ij} = - \gamma \sum_{p=1}^q \frac{\partial E_p}{\partial w_{ij}} = \gamma \sum_{p=1}^q (\hat{a}_{pj} - a_{pj}) a_{pj} (1 - a_{pj}) a_{pi}.$$

[0076] Die bisherigen allgemeinen Überlegungen lassen sich nun auf eine konkrete Startsituation übertragen, was eine schrittweise Anpassung der Gewichte ermöglicht. Dabei können die Anfangswerte w_{ij}^0 beliebig gewählt werden. Mit dem entsprechenden Netz-Input

$$net_{pj}^0 = \sum_{i=1}^m w_{ij}^0 a_{pi}$$

lässt sich direkt der nullte Output

$$a_{pj}^0 = f\left(\sum_{i=1}^m w_{ij}^0 a_{pi}\right) = \frac{1}{1 + e^{-\sum_{i=1}^m w_{ij}^0 a_{pi}}}$$

angeben, und die nullte Gewichtsadaption lautet

$$\Delta w_{ij}^0 = \gamma \sum_{p=1}^q (\hat{a}_{pj} - a_{pj}^0) a_{pj}^0 (1 - a_{pj}^0) a_{pi}.$$

[0077] Damit ergeben sich die Verbindungsgewichte im nächsten Lernschritt aus

$$w_{ij}^1 = w_{ij}^0 + \Delta w_{ij}^0,$$

und mit

$$a_{pj}^1 = f\left(\sum_{i=1}^m w_{ij}^1 a_{pi}\right) = \frac{1}{1 + e^{-\sum_{i=1}^m w_{ij}^1 a_{pi}}}$$

folgt als erstes Differenzgewicht

$$\Delta w_{ij}^1 = \gamma \sum_{p=1}^q (\hat{a}_{pj} - a_{pj}^1) a_{pj}^1 (1 - a_{pj}^1) a_{pi}.$$

[0078] Die zweite Näherung lautet damit

$$w_{ij}^2 = w_{ij}^1 + \Delta w_{ij}^1,$$

und mit

$$a_{pj}^2 = f\left(\sum_{i=1}^m w_{ij}^2 a_{pi}\right) = \frac{1}{1 + e^{-\sum_{i=1}^m w_{ij}^2 a_{pi}}}$$

folgt das zweite Differenzgewicht

$$\Delta w_{ij}^2 = \gamma \sum_{p=1}^q (\hat{a}_{pj} - a_{pj}^2) a_{pj}^2 (1 - a_{pj}^2) a_{pi}.$$

[0079] Analog ergibt sich das dritte Gewicht

$$w_{ij}^3 = w_{ij}^2 + \Delta w_{ij}^2.$$

[0080] Allgemein gilt, wenn w_{ij}^n bekannt ist, eine iterative Abfolge folgender Größen:

$$net_{pj}^n = \sum_{i=1}^m w_{ij}^n a_{pi},$$

$$a_{pj}^n = f\left(net_{pj}^n\right) = \frac{1}{1 + e^{-net_{pj}^n}},$$

$$\Delta w_{ij}^n = \gamma \sum_{p=1}^q (\hat{a}_{pj} - a_{pj}^n) a_{pj}^n (1 - a_{pj}^n) a_{pi},$$

$$w_{ij}^{n+1} = w_{ij}^n + \Delta w_{ij}^n.$$

[0081] Die Anwendung neuronaler Netzwerke gliedert sich in zwei Phasen, und zwar in eine Lern- und eine Anwendungsphase. In der Lernphase werden dem Netz eine Menge Trainingsgesichter präsentiert, aufgrund derer die Verbindungsgewichte der Neuronen so justiert werden, dass das Netz dem gewünschten Verhalten genügt und die Differenz zwischen dem Ausgabewert und dem gewünschten Sollwert minimal wird. In der Anwendungsphase werden erneut Inputvektoren angeboten, wobei die Ausgabevektoren unter Abruf des erlernten Wissens vom Netz selbständig berechnet werden.

[0082] Ergänzend ist darauf hinzuweisen, dass „umfassend“ oder „aufweisend“ keine anderen Elemente oder Schritte ausschließt und „eine“ oder „ein“ keine Vielzahl ausschließt. Ferner sei darauf hingewiesen, dass Merkmale oder Schritte, die mit Verweis auf eines der obigen Ausführungsbeispiele beschrieben worden sind, auch in Kombination mit anderen Merkmalen oder Schritten anderer oben beschriebener Ausführungsbeispiele verwendet werden können. Bezugszeichen in den Ansprüchen sind nicht als Einschränkung anzusehen.

Bezugszeichenliste

1	Objektidentifizierungseinheit
50	Datenquelle
60	Ausgabeeinheit
70	Datenspeicher
80	Steuereinheit
100	künstliches neuronales Netzwerk
110	Eingabeschicht
111-114	Eingabeneuronen
120	Zwischenschicht
130	Ausgabeschicht
131-136	Ausgabeneuronen
150	zweidimensionale Abbildung
151	charakteristische Punkte
152	vertikale Abstände
155	vertikale Symmetrieachse
200	Verfahren
210-250	Verfahrensschritte

Patentansprüche

1. Objektidentifizierungseinheit (1) zum Identifizieren von Objekten, aufweisend:
ein künstliches neuronales Netzwerk, KNN, (100) mit einer Eingabeschicht (110) und einer Ausgabeschicht (130), wobei das KNN (100) ausgestaltet ist, Objekte basierend auf Eingaben an die Eingabeschicht zu identifizieren;
wobei die Eingabeschicht (110) eine Mehrzahl von Eingabeneuronen (111 - 114) aufweist;
wobei die Eingabeschicht ausgeführt ist, einen Eingabevektor zu erhalten, wobei der Eingabevektor Objektidentifizierungsinformationen zu dem zu identifizierenden Objekt enthält;
wobei das KNN (100) ausgeführt ist, basierend auf den Objektidentifizierungsinformationen ein Objekt zu identifizieren;
wobei die Objektidentifizierungsinformationen einer zweidimensionalen Abbildung (150) des zu identifizierenden Objekts entsprechen;
wobei die Objektidentifizierungseinheit (1) konfiguriert ist, die zweidimensionale Abbildung (150) in eine Mehrzahl von Dreiecken, die durch charakteristische Punkte (151) des zu identifizierenden Objekts an ihren Ecken begrenzt sind, zu unterteilen;
wobei die Objektidentifizierungseinheit (1) konfiguriert ist, aus Symmetrieeigenschaften des zu identifizierenden Objekts und relativen Flächenanteilen der einzelnen Dreiecke an der Gesamtsumme der Fläche der Dreiecke einen Drehwinkel des zu identifizierenden Objekts in der zweidimensionalen Abbildung (150) zu ermitteln

und die Flächen der einzelnen Dreiecke mit einem Korrekturfaktor zu multiplizieren, um korrigierte Flächen der Dreiecke zu erhalten, welche einer Symmetriebedingung entsprechen;
 wobei die Objektidentifizierungseinheit (1) konfiguriert ist, die korrigierten Flächen der Dreiecke auf eine drehinvariante Größe des zu identifizierenden Objekts zu normieren;
 wobei die Objektidentifizierungseinheit (1) konfiguriert ist, die normierten Flächen der Dreiecke dem KNN (100) als Eingabevektor zuzuführen.

2. Objektidentifizierungseinheit (1) nach Anspruch 1, wobei das zu identifizierende Objekt ein menschliches Gesicht ist.

3. Objektidentifizierungseinheit (1) nach Anspruch 1 oder 2, wobei die drehinvariante Größe ein vertikaler oder horizontaler Abstand zwischen zwei charakteristischen Punkten des zu identifizierenden Objekts ist.

4. Objektidentifizierungseinheit (1) nach einem der voranstehenden Ansprüche, wobei die Objektidentifizierungseinheit konfiguriert ist, in einer Trainingsphase mit einer Mehrzahl von Abbildungen von Objekten trainiert zu werden, wobei die Abbildungen die Objekte ohne Drehwinkel zeigen.

5. Objektidentifizierungseinheit (1) nach einem der voranstehenden Ansprüche, wobei die Symmetriebedingung eine Achsensymmetrie bezogen auf eine Achse des zu identifizierenden Objekts ist.

6. Objektidentifizierungseinheit (1) nach einem der voranstehenden Ansprüche, wobei die Objektidentifizierungseinheit konfiguriert ist, die zweidimensionale Abbildung so in eine Mehrzahl von Dreiecken zu unterteilen, dass mindestens einige der Dreiecke symmetrisch zu einer Achse des zu identifizierenden Objekts sind.

7. Objektidentifizierungseinheit (1) nach einem der voranstehenden Ansprüche, wobei die Objektidentifizierungseinheit (1) ausgeführt ist, den Korrekturfaktor zu ermitteln, indem für jeweils ein Dreieckspaar bestehend aus zwei einander entsprechenden Dreiecken ein Größenverhältnis ermittelt wird.

8. Verfahren zum Identifizieren von Objekten mit einem künstlichen neuronalen Netzwerk, KNN, wobei das Verfahren die folgenden Schritte aufweist:

Unterteilen (210) einer zweidimensionalen Abbildung eines zu identifizierenden Objekts in eine Mehrzahl von Dreiecken, die durch charakteristische Punkte (151) des zu identifizierenden Objekts an ihren Ecken begrenzt sind;

Ermitteln (220) eines Drehwinkels des zu identifizierenden Objekts in der zweidimensionalen Abbildung basierend auf Symmetrieeigenschaften des zu identifizierenden Objekts und relativen Flächenanteilen der einzelnen Dreiecke an einer Gesamtsumme der Fläche der Dreiecke;

Ermitteln (230) von korrigierten Flächen der Dreiecke, indem die Flächen der Dreiecke mit einem Korrekturfaktor multipliziert werden, wobei die korrigierten Flächen der Dreiecke einer Symmetriebedingung entsprechen;
 Normieren (240) der korrigierten Flächen der Dreiecke auf eine drehinvariante Größe des zu identifizierenden Objekts;

Zuführen (250) der normierten korrigierten Flächen an das KNN, um basierend darauf einen Abgleich von Objekten vorzunehmen und diese zu identifizieren.

9. Verfahren nach Anspruch 8, wobei die drehinvariante Größe ein vertikaler oder horizontaler Abstand zwischen zwei charakteristischen Punkten des zu identifizierenden Objekts ist.

10. Verfahren nach Anspruch 8 oder 9, wobei die Symmetriebedingung eine Achsensymmetrie bezogen auf eine Achse des zu identifizierenden Objekts ist.

Es folgen 4 Seiten Zeichnungen

Anhängende Zeichnungen

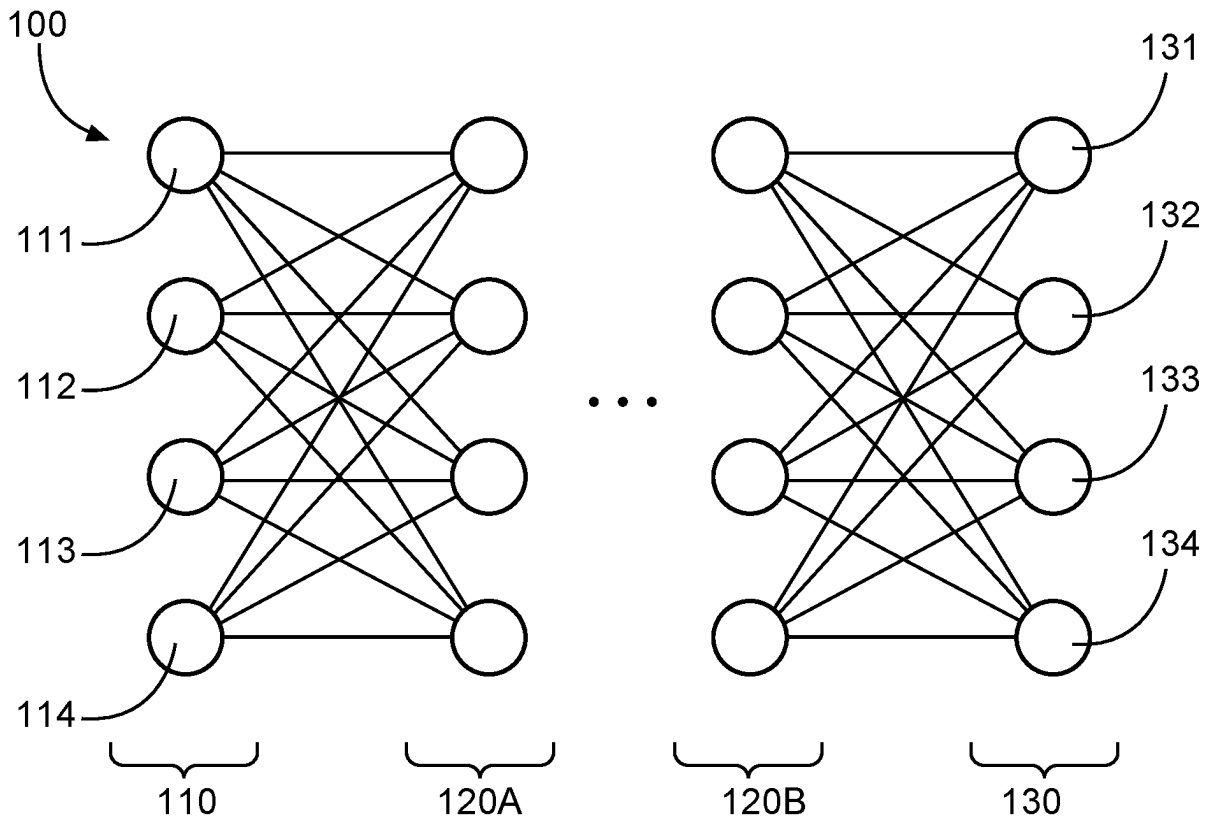


Fig. 1

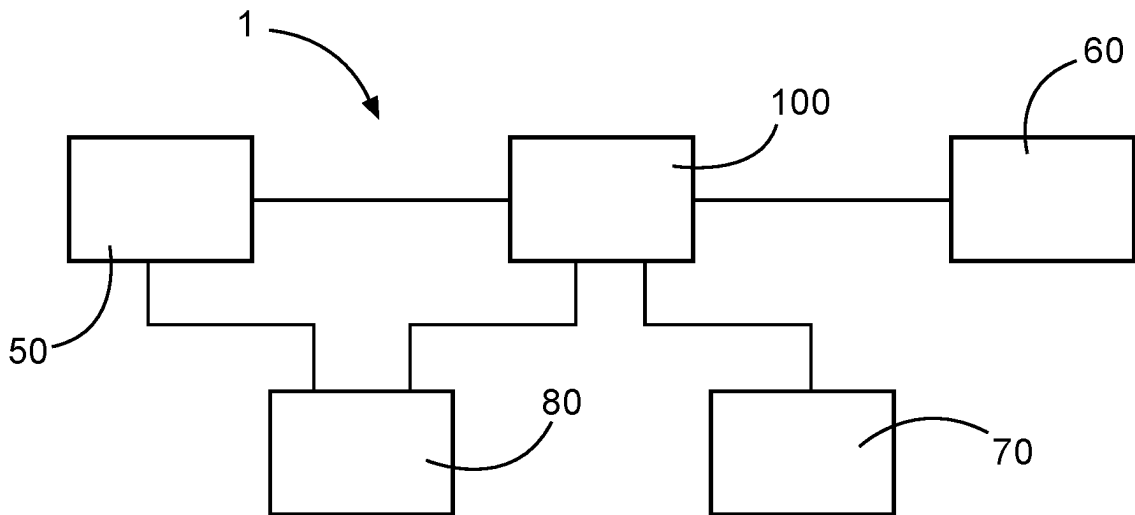


Fig. 2

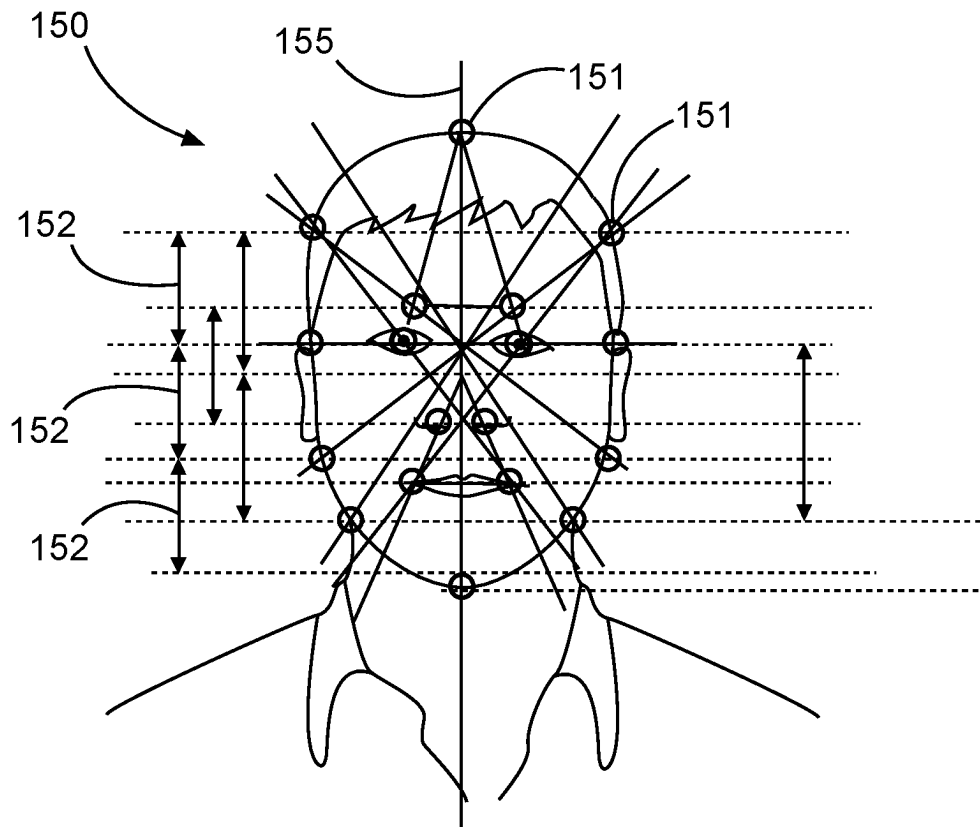


Fig. 3

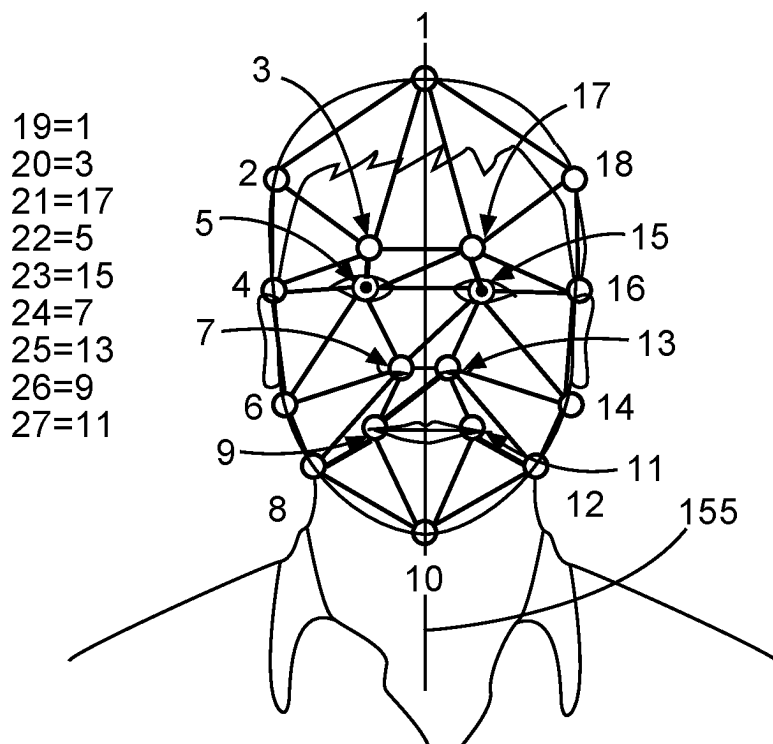


Fig. 4

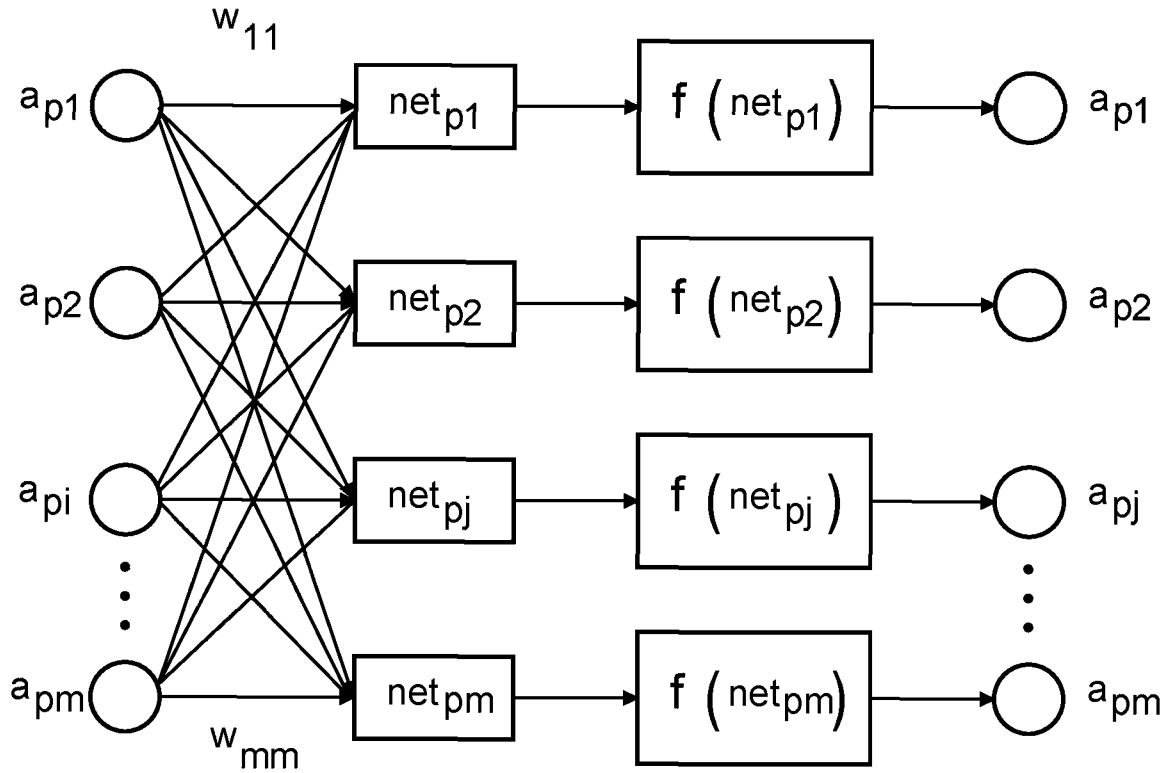


Fig. 5

$$A_1 \cos(\theta_1 - \theta) + A_{17} \cos(\theta_{17} + \theta) = 2A_1 \cos\theta_1 \cos\theta$$

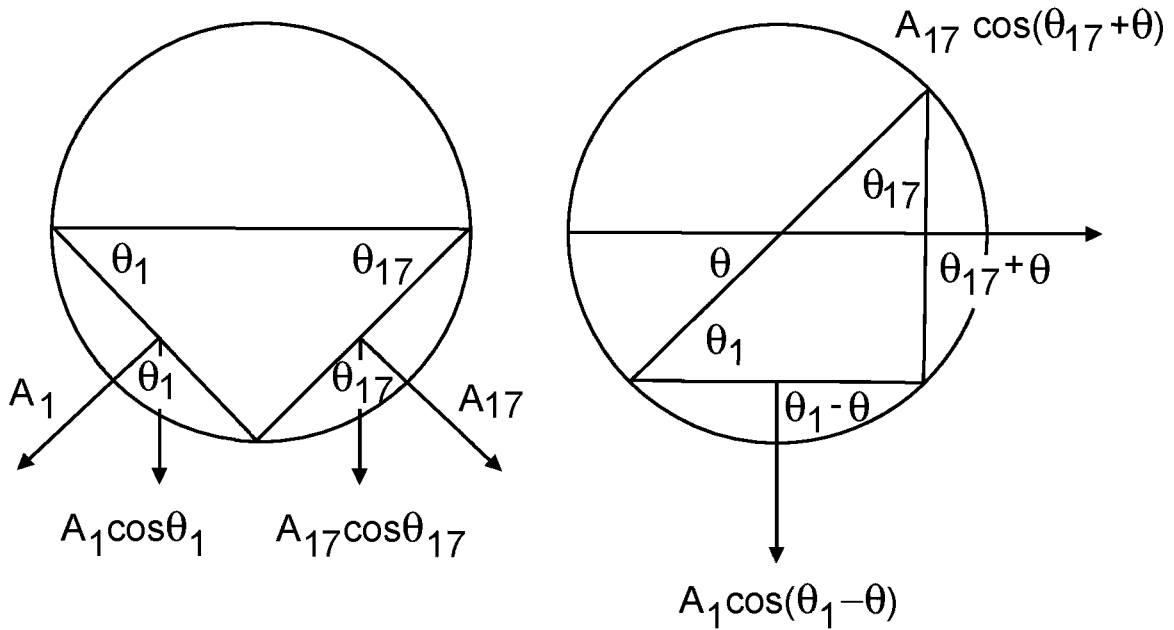


Fig. 6

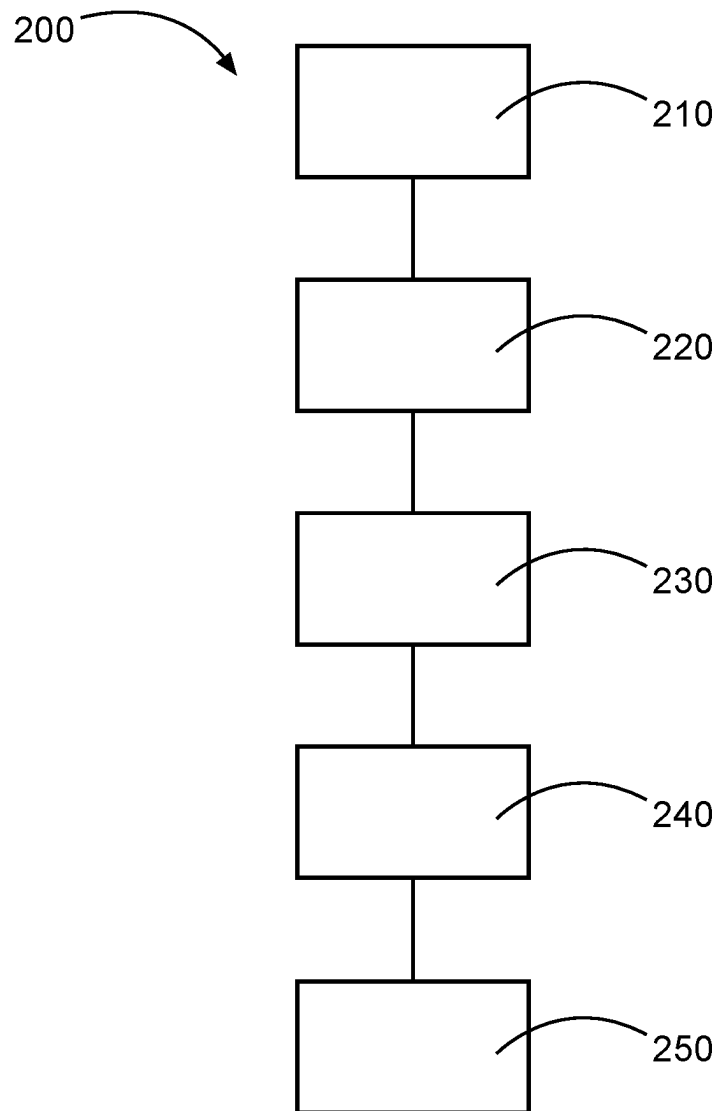


Fig. 7