

# EINE AM KOPF MONTIERTE VORRICHTUNG FÜR DREIDIMENSIONALEN DISPLAY\*

Ivan E. Sutherland

\* Die Forschungsarbeit, über die im folgenden berichtet wird, fand an der Harvard University statt und wurde zum Teil von der Advanced Research Projects Agency (ARPA) des Department of Defense (KontraktSD265), zum Teil vom Office of Naval Research (ONR 1866 [16]), und teilweise durch ein stehendes Übereinkommen zwischen Bell Telephone Laboratories und dem Harvard Computation Laboratory unterstützt. Auch die Vorarbeiten im Lincoln Laboratory von MIT wurden von ARPA unterstützt.

## Einführung

Die im folgenden beschriebene Einrichtung für dreidimensionale Wiedergabe fußt auf dem Wunsch, dem Benutzer ein perspektivisches Bild darzubieten, das sich verändert, wenn sich dieser bewegt. Das auf der Retina dargestellte Abbild wirklicher Gegenstände ist schließlich auch nur zweidimensional. Folglich können wir, wenn es uns gelingt, geeignete zweidimensionale Bilder auf der Retina des Beobachters zu erzeugen, in ihm die Illusion erwecken, er sähe einen dreidimensionalen Gegenstand. Obwohl eine stereooptische Darbietung für die Erzeugung der Illusion der Räumlichkeit wichtig ist, ist sie doch weniger wichtig als jene Änderungen des gesehenen Bildes, die sich aus einer Kopfbewegung des Betrachters ergeben. Das durch den dreidimensionalen Display dargebotene Bild muß sich daher in genau der gleichen Weise verändern, in der sich das Abbild eines wirklichen Gegenstandes durch eine analoge Kopfbewegung des Betrachters verändern würde. Fachpsychologen wissen längst, daß bewegte perspektivische Bilder auch ohne Stereodarbietung überraschend "räumlich" wirken. Der in der vorliegenden Arbeit beschriebene dreidimensionale Display beruht in seiner Wirkung zum Großteil auf eben diesem "kinetischen Tiefeneffekt".<sup>1</sup>

Beim vorliegenden Projekt sind wir nicht darum bemüht, Bewegungen des Augapfels zu messen. Da solche Messungen äußerst schwierig durchzuführen sind, liegen die Dinge so, daß das dargebotene perspektivische Bild nicht geändert zu werden braucht, wenn der Betrachter seine Augen bewegt, um sich auf den von ihm jeweils gewählten Bildteil zu konzentrieren. Die dargebotene Perspektive ist nur dann zu verändern, wenn der Betrachter seinen Kopf bewegt. Wir messen also tatsächlich nur die Position und Richtung des am Kopf des Benutzers befestigten optischen Systems.

Da das optische System die tatsächliche Bildschirm-Position — und damit den Blickpunkt des Anwenders — bestimmt, entscheiden Lage und Richtung des optischen Systems auch darüber, welche perspektivische Sicht angemessen ist.

Unser Anliegen beim vorliegenden Projekt war es, den Anwender mit wahrnehmbarer dreidimensionaler Information förmlich zu "umgeben". Da wir dabei eine homogene Koordinatendarstellung verwenden,<sup>2,3</sup> können wir dergestalt Gegenstände abbilden, die dem Betrachter nahe erscheinen, aber auch solche, die ihm unendlich weit entfernt erscheinen. Wir können Gegenstände neben oder hinter dem Beschauer darstellen, die für diesen erst sichtbar werden, wenn er seinen Kopf dreht. Der Benutzer ist in der Lage, seinen Kopf etwa einen Meter von der Hauptachse in jeder Richtung zu bewegen, um nahegelegene Gegenstände besser sehen zu können. Er kann sich völlig umwenden und auch seinen Kopf um etwa 30 bis 40 Grad nach unten oder nach oben neigen. Die dargestellten Gegenstände scheinen rund um den Betrachter im Raum zu schweben.

Der Wunsch, den Anwender mit Information zu umgeben, hat uns dazu gezwungen, das "Fensterproblem" zu lösen. Die von uns konstruierte Clipping-Teiler-Hardware eliminiert jene Teile von Linien, die hinter dem Betrachter oder außerhalb seines Gesichtsfeldes liegen. Diese Einrichtung vollzieht auch die für eine wirklich perspektivische Sicht erforderliche Teilung. Der Clipping-Teiler ist imstande, die für die "Beschneidung" irgendeiner Linie erforderlichen Kalkulationen innerhalb von 10 Mikrosekunden durchzuführen. Das entspricht etwa der Geschwindigkeit, mit der ein Hochleistungs-Display Linien auf dem Bildschirm einer Kathodenstrahlröhre zeichnen kann. Der Clipping-Teiler ist im Detail in einem anderen Beitrag<sup>4</sup> in dieser Ausgabe beschrieben. Da der Clipping-Teiler dynamische perspektivische Darstellung dreidimensionaler Zeichnungen und beliebige Vergrößerung zweidimensionaler Zeichnungen gestattet, erscheint er uns als das bis dato wesentlichste Ergebnis der einschlägigen Forschung.

Um wirklich realistische Darstellungen fester, dreidimensionaler Gegenstände zu erhalten, ist erforderlich, das "Problem der verdeckten Linien" zu lösen. Obwohl es relativ leicht ist, die Perspektivenpositionen aller Teile eines komplexen Gegenstandes zu berechnen, ist es doch schwierig zu kalkulieren, welche Teile eines Gegenstandes durch einen anderen verdeckt werden. Von den hierfür derzeit verfügbaren Software-Lösungen<sup>2,5-10</sup> scheinen nur zwei, die Ansätze von MAGI 9 und Warnock<sup>10</sup>, als schließliche Echtzeit-Lösungen für einigermaßen komplexe Situationen in Frage zu kommen. Die von den übrigen entsprechenden Methoden erforderliche Zeit dürfte im Quadrat zum Komplexitätsgrad der Situation anwachsen. Die einzige derzeitige Echtzeit-Lösung des "Problems der verborgenen Linien" bietet ein sehr kostspieliger Spezialcomputer der NASA in Houston<sup>11</sup> und selbst dieser vermag nur relativ einfache Gegenstände darzustellen. Wir sind daher zu dem Schluß gekommen, daß die Darstellung "undurchsichtiger Gegenstände, bei denen die verborgenen Linien eliminiert wurden, jenseits unserer gegenwärtigen Möglichkeiten liegt. Die mit unseren Mitteln dargestellten dreidimensionalen Gegenstände sind sozusagen durchsichtige "Drahtrahmen"-Strichzeichnungen.

## Betrieb und Funktion des Displaysystems

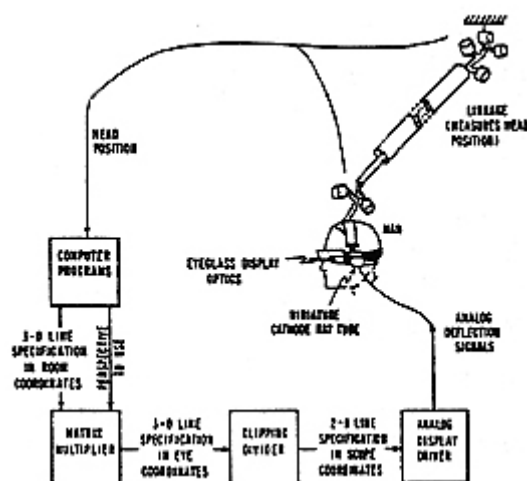


Abb. 1: Die Komponenten eines dreidimensionalen Display-Systems

Um dem Benützer bei Bewegung seines Kopfes sich wandelnde perspektivische Bilder anzubieten, haben wir, wie aus dem Diagramm in Abbildung 1 ersichtlich eine große Vielfalt von Einrichtungen eingesetzt. Am Kopf des Anwenders befindet sich eine Spezialbrille, in der sich zwei Miniatur-Kathodenstrahlröhren befinden. Ein schneller, zweidimensionaler Analog-Strichgenerator versorgt über Transistor-Ablenkverstärker die beiden Miniatur-

Kathodenstrahlröhren mit den erforderlichen Ablenssignalen. Entweder der eine oder der andere Kopf-Positions-Sensor (der eine mechanisch, der andere mit Ultraschall arbeitend) wird eingesetzt, um Ort, Lage und Haltung des Kopfes des Anwenders zu messen.

Bewegt der Betrachter nun seinen Kopf, so verändert sich auch seine Blickrichtung, d.h. sie vollzieht, bezogen auf das Raumkoordinatensystem, eine Drehung. Um nun die Umwandlung von den Raumkoordinaten zu einem auf dem Blickwinkel des Benützers basierenden Koordinatensystem zu erzielen, sind sowohl eine "Übersetzung" als auch eine Drehung erforderlich. Ein Computer nutzt die gemessene Information über Position und Neigung des Kopfes zur Berechnung der Drehungselemente und einer jeden einzelnen Blickrichtung entsprechenden "Übersetzungsmatrix". Anstatt bei Kopfbewegungen des Benützers Informationen<sup>3</sup> im Speicher des Computers zu verändern, formen wir dynamisch Information über Raumkoordinaten in Augenkoordinaten um, während sie dargeboten wird. Eine neue Dreh- und Übersetzungsmatrix wird einmal bei Beginn jeder Bildwiederholung in den digitalen Matrix-Multiplikator geladen. Als Teil des Darbietungsvorgangs werden die Endpunkte von Linien im Raumkoordinatensystem dem Speicher entnommen und vom Matrix-Multiplikator einzeln in das Augenkoordinatensystem "übersetzt". Diese "übersetzten" und gedrehten Endpunkte werdendem digitalen Clipping-Teiler über ein zwischengeschaltetes Pufferglied zugeführt. Der Clipping-Teiler eliminiert nun jegliche außerhalb des Gesichtsfeldes des Benützers liegende Information und berechnet für das verbleibende Datenmaterial das entsprechende perspektivische Bild. Der letztliche Output des Clipping-Teilers besteht aus Endpunkten zweidimensionaler Linien, die in den Rahmenkoordinaten festgelegt sind. Die Spezifikationen der zweidimensionalen Linien werden einer gepufferten Display-Schnittstelle zugeführt, die den analogen Strichzeichnungs-Display treibt.

Wir haben sowohl den Spezial-Digital-Matrix-Multiplikator als auch den Clipping-Teiler gebaut, da kein derzeit verfügbarer Universalcomputer rasch genug ist, ein flackerfreies, dynamisches Schirmbild zu liefern.

Unsere Geräte können 3000 Display-Zeilen mit einer Bildfrequenz von 30 pro Sekunde liefern, so daß pro Zeile nur wenig über 10 Mikrosekunden erforderlich sind. Vektorenfolgen in Kettenform, bei denen der Beginn eines Vektors mit dem Ende des Vorhergehenden identisch ist, lassen sich dabei etwas effizienter verarbeiten als Einzellinien. Nehmen wir aber für jede Linie zwei Endpunkte an, so muß der Matrix-Multiplikator die Koordinatentransformation in etwa 5 Mikrosekunden pro Endpunkt schaffen. Jede einzelne Matrix-Multiplikation erfordert 16 akkumulierende Multiplikationen. Das bedeutet also einen Durchsatz von 3,000.000 Multiplikationen pro Sekunde. Der vom Multiplikator getrennte asynchrone Clipping-Teiler arbeitet mit etwa der gleichen Geschwindigkeit und verarbeitet zwei Endpunkte in etwas über 10 Mikrosekunden. Anders als die bestimmte Zeit, die eine Matrix-Multiplikation erfordert, hängt beim Clipping-Teiler die Verarbeitungszeit von den zu verarbeitenden Daten ab.

Das Zeiterfordernis des Analog-Liniengenerators wird durch die Länge der zu erzeugenden Linie bestimmt, wobei die kürzeste etwa 3, die längste dagegen etwa 36 Mikrosekunden erfordert. Der Durchschnitt liegt bei etwa 10 Mikrosekunden.

Der Matrix-Multiplikator, Clipping-Teiler und der Linien-Generator sind miteinander in einer Art "Pipeline" verbunden. Die Daten durchfließen das System in präzise sequenz geregelter Form. Jede der drei Einheiten ist unabhängig zeitgesteuert und besorgt ihre eigene Input- und Output-Synchronisation. Jede Einheit prüft eine "Input Flag" (Eingangskennung), die den Zugang von Daten ankündigt. Diese Daten werden zurückgehalten, bis die Einheit zu ihrer

Übernahme bereit ist. Bei Übernahme jedes einzelnen Informationsteiles liest die Einheit auch gleichzeitig eine "Anweisung", aus der sie erkennt, was mit dem betreffenden Datum zu tun ist. Hat die Einheit ein Datum angenommen, so wird die betreffende "Input Flag" gelöscht. Sobald die Anweisung befolgt ist, erfolgt die Ausgabe in Output-Linien und eine "Output Flag" zeigt an, daß die Daten nun zur Weitergabe bereit sind. In einigen Fällen beginnt die Einheit mit der Verarbeitung der nächsten Information, ehe noch der letzte Output weitergeleitet wurde.

In diesem Falle hält sie in der Berechnung inne, falls eine weitere Verarbeitung den letzten noch vorhandenen Output zerstören würde. Der ordnungsgemäße Durchsatz durch das System ist völlig gesichert, da die Output Flag jeder Einheit gleichzeitig als Input Flag für die nachfolgende Einheit dient. Der durchschnittliche Datendurchsatz des gesamten Systems entspricht etwa dem durchschnittlichen Datendurchsatz der langsamsten Einheit. Der durchschnittliche Standard-Durchsatz liegt bei etwa 10 Mikrosekunden pro Linie.

Der Computer dieses Systems wird nur dafür verwendet, die vom Sensor für die Kopfposition kommende Information einmal pro Bild zu verarbeiten und die dreidimensionale Zeichnung zu speichern und zu manipulieren. Kein derzeit verfügbarer Universalcomputer wäre rasch genug, mit den für eine dynamische perspektivische Darstellung erforderlichen Perspektivenberechnungen fertig zu werden. Ein Displaykanal-Prozessor dient dazu, dem Speicher jene Zeichendaten zu entnehmen, die für die Neuberechnung und "Auffrischung" des Schirmbildes erforderlich sind. Der Kanal-Prozessor läßt sich in verschiedener Weise konfigurieren, so daß es auch möglich wird, den Matrix-Multiplikator und den Clipping-Teiler unabhängig voneinander einzusetzen. So etwa kann der Matrix-Multiplikator in einem direkten Speicher-zu-Speicher-Modus eingesetzt werden, was die arithmetische Verarbeitungskapazität des angeschlossenen Computers wesentlich steigert. Zum Zwecke der zweidimensionalen Präsentation ist es auch möglich, den Matrix-Multiplikator zu übergehen und so den Clipping-Teiler und den Display direkt mit den entsprechenden Input-Signalen zu versorgen. Diese Optionen waren für eine unabhängige Fehlersuche und Störungsbereinigung an den einzelnen Einheiten unerlässlich.

### **Darbietung von Bildern für den Anwender**

Das optische System des Kopfgeräts vergrößert die Bilder auf den Bildschirmen jeder der beiden winzigen Kathodenstrahlröhren, so daß ein etwa achtzehn Zoll vor jedem der beiden Augen des Anwender wahrgenommenes Bild dargeboten wird. Jedes dieser beiden wahrgenommenen Bilder hat annähernd die Größe einer herkömmlichen Bildschirmröhre. Innerhalb der von den Miniatur-Bildschirmröhren dargebotenen synthetischen Information hat der Anwender ein Gesichtsfeld von 40 Grad. Halbverspiegelte Flächen der Prismen, durch die der Benutzer blickt, ermöglichen ihm, gleichzeitig sowohl die Schirmbilder als auch tatsächlich im Raum befindliche Gegenstände zu sehen. So kann man das dargestellte Material entweder sozusagen frei im Raum schweben lassen, oder aber es tatsächlich vorhandenen Landkarten, Arbeitsflächen von Schreibtischen, Wänden oder der Tastatur einer Schreibmaschine überlagern.

Die Bildschirme der an dem optischen System angebrachten Miniatur-Röhren liefern ein Bild, das etwa einen halben Zoll (ca. 13 mm) im Quadrat mißt.

Da diese Röhren eine Nenn-Punktauflösung von etwa sechs Zehntel Mil aufweisen, so entspricht die Auflösung des vom Anwender wahrgenommenen Bildes dem üblicher normal großer Bildschirme etwa gleichwertig. Jede der beiden Kathodenstrahlröhren ist von einem

sorgfältig schutzgeerdeten Metallmantel umschlossen, um den Anwender vor Schlüssen im Hochspannungssystem zu schützen. Geerdete Abschirmung der gesamten Hochspannungsverkabelung sorgt für zusätzliche Sicherheit.

Es hat sich gezeigt, daß die Miniatur-Kathodenstrahlröhren leicht zu steuern sind. Bei diesen Röhren ist sowohl die Ablenkung als auch die Bündelung elektrostatisch. Da ihre Ablenkplatten Signalspannungen in der Größenordnung von nur 300 V erfordern, sind die transistorbestückten Ablenkverstärker in Schaltung und Ausführung relativ unkompliziert.

Komplementär symmetrische Emitter-Followers dienen als Treiber für vier kleine Koaxialkabel, die den Verstärker mit jeder der beiden Kathodenstrahlröhren verbinden. Ablenkungs- und Verstärkungssignale für die beiden Miniatur-Kathodenstrahlröhren werden von einem handelsüblichen Analog-Linienzeichen-Display. Dieser ist in der Lage, lange Linien in 36 Mikrosekunden (Nennzeiten) und kurze Linien sogar in nur 3 Mikrosekunden (Nennzeiten) zu zeichnen.

Der Analog-Linienzeichen-Generator akzeptiert Bildinformation im Koordinatensystem der Miniatur-Bildschirmröhre. Er erhält zweidimensionale Rahmenkoordinaten für die Endpunkte jedes darzustellenden Liniensegments. Er verbindet diese Endpunkte auf dem zweidimensionalen Bildschirm mit kontinuierlichen geraden Linien. So ermöglichen der Linienzeichen-Generator, die transistorbestückten Ablenkverstärker, die Miniaturbildschirmröhre und das am Kopf befestigte optische System in ihrem Zusammenwirken die Darbietung beliebiger zweidimensionaler Zeichnungen für den Anwender.

### **Sensor für die Bestimmung von Kopfposition und Haltung**

Aufgabe des Sensors ist es, Kopfposition und Blickrichtung des Anwenders zu messen und an den Computer zu melden. Der Kopfhaltungs-Sensor soll dem Anwender ein vernünftiges Maß an Bewegungsfreiheit gewähren. Im weiteren Verlauf unserer Arbeit möchten wir selbstverständlich den Anwender in die Lage versetzen, im Raum frei herumzugehen. Jetzt aber erlaubt unsere Anfangsausrüstung lediglich ein Arbeitsvolumen für Kopfbewegung und Lageänderung von etwa sechs Fuß (nicht ganz 2 m) Durchmesser und etwa drei Fuß (91,5 cm) Höhe. Innerhalb dieses Volumens kann sich der Anwender völlig frei bewegen, kann sich ganz um seine eigene Achse drehen, und kann den Kopf (und damit den Blickwinkel) um bis zu 40 Grad nach oben oder nach unten neigen. Jenseits dieser Grenzen kann Kopfposition und Neigung vom Sensor nicht mehr erfaßt und gemessen werden. Wir nehmen aber an, daß es schließlich möglich sein wird, das Bewegungsfeld des Anwenders einfach dadurch zu vergrößern, daß man den Oberteil des Sensors an der Decke durch eine Art von einem Servomotor oder Schrittmotor angetriebene Laufkatze entsprechend der veränderten Position des Kopfes bewegen läßt. Da die Position des Kopfes mit Bezug auf den Sensor bekannt ist, dürfte es relativ leicht sein, den Sensor so stets lotrecht über dem Kopf zu halten.

Es ist wichtig, daß die Messungen der Kopfposition mit relativ hoher Auflösung gemacht werden. Letztlich streben wir eine Genauigkeit von einem Hundertstel Zoll und einem Zehntausendstel der Drehbewegung an. Ein noch höheres Auflösungsvermögen ist nicht sinnvoll, da die digital/analog-Konversion des Display-Systems selbst schon ein "Korn" etwa dieser Größenordnung aufweist. Schwieriger schon ist es, die Genauigkeitserfordernisse für den Kopfpositions-Sensor festzulegen. Da die Miniatur-Bildschirmröhren und das am Kopf des Benützers befestigte optische System zusammen eine "Nadelkissenverzerrung" von etwa 3 Prozent aufweisen, ist es möglich, daß die dem Benützer dargebotene visuelle Information

bis zu drei Zehntel Zoll von ihrer Soll-Position abweicht. Demgemäß müßte unser Kopfpositions-Sensor über eine Genauigkeit in der Größenordnung von einem Zehntel Zoll verfügen, obwohl durchaus brauchbare Ergebnisse schon mit weniger genauer Kopfpositions-Information zu erzielen sind.

Für die Wahrnehmung und Messung der Kopfposition haben wir zwei verschiedene Methoden erprobt. Bei der ersten der beiden hängt ein mechanischer Arm von der Decke. Dieser Arm kann von einem vertikalen Angelpunkt in der Decke aus jede Drehbewegung ausführen. Der Arm selbst verfügt über zwei Kardangelenke (je eines am oberen und am unteren Ende) und einen teleskopartig gleitenden Mittelteil, um die sechs für die Messung von Drehung und Übertragung erforderlichen Bewegungen ausführen zu können. Die Lage jedes der Gelenke wird gemessen und dem Computer über einen digitalen Achsenpositions-Kodierer zugeführt.

Der mechanische Kopfpositions-Sensor ist relativ schwer und unbequem in der Anwendung. Die von ihm erlangte Information dagegen läßt sich leicht in eine für die Erzeugung der Perspektiventransformation geeignete Form konvertieren. Wir haben das Gerät gebaut, um über eine verlässliche Methode zur Messung der Kopfposition zu verfügen. Daneben haben wir aber auch den nicht gepulsten Ultraschall-Kopfpositions-Sensor entwickelt und konstruiert.

$$X_s = \frac{x'}{r'} \cotan \frac{\alpha}{2}$$

$$Y_s = \frac{y'}{r'} \cotan \frac{\alpha}{2}$$

Abb. 2: Das Logik-System des Ultraschall-Kopfpositions-Sensors

Drei Sender, die Ultraschall in den Frequenzen 37, 38,6 und 40,2 kHz emittieren, sind an dem am Kopfe des Benützers angebrachten optischen System montiert. Vier Ultraschall-Empfänger sind in quadratischer Anordnung an der Decke des Raumes angebracht. Jeder dieser Empfänger ist, wie in Abbildung 2 gezeigt, mit einem Verstärker und drei Filtern verbunden, so daß Phasenänderungen im ausgestrahlten Ultraschall auf zwölf verschiedenen Wegen gemessen werden können. Die für jeden Ultraschall-Weg gemessene Phasenverschiebung wird vom Computer als eine diskrete 5bit-Zahl gelesen. Der Computer zählt wesentliche Phasenveränderung und verfolgt damit Bewegungen, die das Ausmaß einer Wellenlänge überschreiten.

Anders als beim Lincoln Wand<sup>12</sup>, bei dem es sich um ein gepulstes Ultraschallsystem handelt, arbeitet unser Ultraschall-Kopfpositions-Sensor mit kontinuierlicher Schallemission. Wir haben diese Form gepulstem Ultraschall vorgezogen, weil es auf dem Markt relativ preiswerte Engpaß-Transducer gibt, aber auch um Interferenzen durch gepulste Geräusche (etwa von Schreibmaschinen) zu vermeiden, die beim Lincoln Wand Schwierigkeiten bereitet hatten.

Die Verwendung von kontinuierlichem Ultraschall führt andererseits aber auch zu einer Zweideutigkeit der Meßergebnisse. Obwohl der Ultraschall-Kopfpositions-Sensor zwölf Messungen durchführt, aus denen sich Informationen über die Kopfposition und Haltung

ergeben, besteht doch in jeder dieser Messungen eine Zweideutigkeit hinsichtlich der Wellenlänge. Wohl werden die Messungen innerhalb einer Welle mit hoher Genauigkeit durchgeführt, doch bleibt offen, welche Welle gemessen wurde. Da die Schallwellenlänge in Luft bei 40 kHz etwa ein Drittel Zoll beträgt, ist jede der zwölf Messungen in Ein-Drittel-Zoll-Intervallen zweideutig. Da der Computer auf vollständige Phasenveränderungen anspricht, erscheint die Zweideutigkeit der Meßergebnisse in Form eines konstanten Fehlers in der gemessenen Entfernung. Man kann diesen Fehler sozusagen als Ausgangsfehler ("initialization error") des Systems betrachten. Dieser stellt den Unterschied zwischen der Schätzung des Computers hinsichtlich der Ausgangs-Wegstrecke und der tatsächlichen Länge des Ausgangs-Weges dar.

Unserer Ansicht nach ist das Problem der "Ausgangsfehler" durch Nutzung der Durchführung von zwölf Messungen inhärenten geometrischen Redundanz zu bereinigen. Wir haben viel Mühe aufgewendet, um für den Ultraschall-Koppositions-Sensor Programme zu schreiben. Diese Programme umfassen auch mehrere Methoden zur Eliminierung der Zweideutigkeit der Meßergebnisse. Obwohl wir bereits einige ermutigende Ergebnisse erzielt haben, ist es noch nicht möglich, einen umfassenden Bericht über den Ultraschall-Koppositions-Sensor zu veröffentlichen.

### Die Perspektiven-Transformation

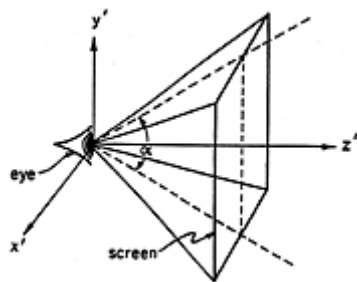
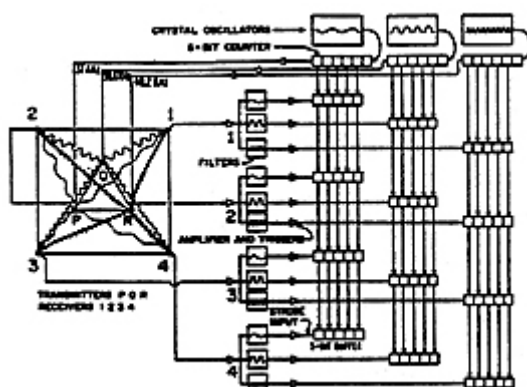


Abb. 3: Das auf der Augenposition des Betrachters basierende  $x', y', z'$  Koordinatensystem

Ein perspektivisches Abbild dreidimensionaler Information zu erzeugen, ist relativ leicht. Nehmen wir einmal an, die Information werde in einem auf dem Auge des Betrachters basierenden Koordinatensystem dargeboten, wie dies aus Abbildung 3 ersichtlich ist.

Betrachtet man die zweidimensionalen Rahmenkoordinaten als von  $-1$  bis  $+1$  reichend, so zeigen einfache geometrische Überlegungen, daß die Position, an der ein bestimmter Punkt auf dem Bildschirm darzustellen ist, zur Position dieses Punktes im dreidimensionalen Raum in folgender Beziehung steht:



Wünscht man eine orthogonale Projektion, so läßt sich dies dadurch erreichen, daß man den Wert für  $z'$  konstant hält. Da die perspektivische (oder orthogonale) Projektion einer Geraden im dreidimensionalen Raum wieder eine Gerade ist, ist die Division durch die  $z'$ -Koordinate nur für die Endpunkte der geraden Linie erforderlich.

Der zweidimensionale Analog-Liniengenerator kann den Mittelteil einer dreidimensionalen Linie einfach durch Aufzeichnen einer zweidimensionalen Geraden füllen. Der digitale Perspektiven-Generator berechnet nur die Werte der Endpunkt-Koordinaten einer Geraden.

Im Computer wird die für den dreidimensionalen Display erforderliche Information in einem festen dreidimensionalen Koordinatensystem gespeichert. Da dieses Koordinatensystem auf dem den Benutzer umgebenden Raum beruht, nennen wir es das "Raum"-Koordinatensystem. Die Zeichendaten im Raum-Koordinatensystem sind als homogene Koordinaten dargestellt. Das bedeutet, daß jeder dreidimensionale Punkt oder jedes Ende einer dreidimensionalen Geraden in Form von vier separaten Zahlen gespeichert wird. Die ersten drei dieser vier Zahlen entsprechen den üblichen  $x$ -,  $y$ - und  $z$ -Koordinaten des dreidimensionalen Raumes. Die vierte, meist mit  $w$  bezeichnete Koordinate, ist ein Maßstabsfaktor, dem zu entnehmen ist, ein wie großer Wert für  $x$ ,  $y$  oder  $z$  eine Streckeneinheit darstellt. Weit entferntes Material kann so leicht dadurch dargestellt werden, daß man den Maßstabsfaktor  $w$  klein macht. Unendlich weit entfernte Punkte werden durch Gleichsetzung des Wertes  $w$  mit Null dargestellt. In diesem Fall bedeuten die ersten drei Raumkoordinaten nur die Richtung auf diesen Punkt hin.

Nahegelegene Punkte stellt man üblicherweise dadurch dar, daß man dem Maßstabsfaktor  $w$  seinen größtmöglichen Wert gibt. In diesem Fall sind dann die übrigen drei Koordinaten nur die vertrauten Fixpunkt-Darstellungen von  $x$ ,  $y$  und  $z$ .

### **Der Matrix-Multiplikator**

Wir haben den Matrix-Multiplikator für die dynamische Konversion von Information vom fixen Raum-Koordinatensystem zum bewegten "Augen"-Koordinatensystem gebaut. Der Matrix-Multiplikator speichert eine Vier-Mal-Vier-Matrix von 18-bit-Fixpunkt-Zahlen. Da die Zeichendaten in Form homogener Koordinaten dargestellt werden, dient die einfache Vier-Mal-Vier-Matrix-Multiplikation sowohl der "Übersetzung" (Umsetzung) als auch der Drehung (Rotation).<sup>2</sup> Der Matrix-Multiplikator akzeptiert die vier 18-bit-Zahlen, die einen Endpunkt darstellen, wobei er diese als Vier-Komponenten-Vektor behandelt, der nun mit der Vier-Mal-Vier-Matrix multipliziert wird.

Das Ergebnis ist ein Vier-Komponenten-Vektor, dessen Einzelkomponenten auf 20 bits reduziert wurden. Der Matrix-Multiplikator gibt nun dieses 80-bit-Ergebnis innerhalb von annähernd 5 Mikrosekunden an den Clipping-Teiler weiter. Somit finden im Matrix-Multiplikator rund drei Millionen Skalar-Multiplikationen pro Sekunde statt.

Der Matrix-Multiplikator verfügt für jede Kolonne über ein eigenes Multiplikator-Modul. Jedes Modul besteht aus einem Akkumulator, einem Teilproduktverzeichnis, Speicherkapazität für vier Matrix-Elemente in der betreffenden Kolonne und die Multiplikations-Logik. Die Eingaben einer Reihe der Matrix stellen gleichzeitig vier getrennte Multiplikanden dar. Eine Einzelkomponente des empfangenen Vektors dient als gemeinsamer Multiplikator. Solchermaßen werden die vier Multiplikationen für eine einzelne Reihe gleichzeitig durchgeführt. Um die Verarbeitungsgeschwindigkeit noch weiter zu erhöhen,



werden die bits eines Multiplikators in Vierergruppen anstatt jeweils einzeln geprüft, um Additions-Arrays mit Mehrfach-Eingang zu steuern.

## Die Clipping- oder "Fenster"-Funktion

Die Aufgabe des Clipping-Teilers ist es, dreidimensionale Informationen im Augen-Koordinatensystem aufzunehmen und diese in für die zweidimensionale Darstellung entsprechende Endpunkte umzuwandeln. Sind beide Enden einer Linie sichtbar, hat der Clipping-Teiler lediglich vier Divisionen durchzuführen, nämlich je eine für jede zweidimensionale Koordinate jedes Endes der Linie. Der Clipping-Teiler ist elektronisch hinreichend ausgestattet, um diese vier Divisionen gleichzeitig auszuführen.

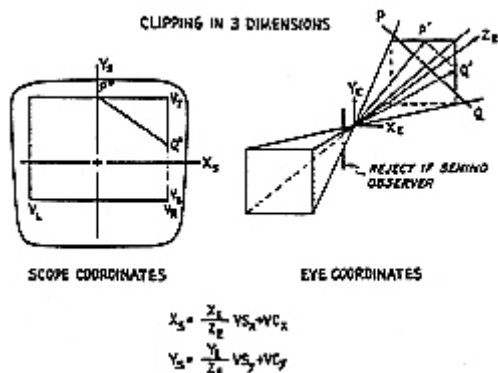


Abb. 4: Clipping und perspektivische, dreidimensionale Projektion

Befinden sich aber die Endpunkte einer Linie nicht innerhalb des Gesichtsfeldes des Betrachters, hat der Clipping-Teiler zu entscheiden, ob sich eventuell irgendein Teil der Linie innerhalb des Gesichtes des Betrachters befindet. Trifft dies zu, so muß der Clipping-Teiler, wie in Abbildung 4 gezeigt, entsprechende Endpunkte für den betreffenden Teil der Linie errechnen. Linien, die sich außerhalb des Gesichtsfeldes des Betrachters oder hinter diesem befinden, müssen eliminiert werden. Funktion und Betrieb des Clipping-Teilers sind in einem eigenen Artikel an anderer Stelle in dieser Publikation dargestellt.

Ebenso wie der Matrix-Multiplikator ist auch der Clipping-Teiler ein unabhängig zeitgesteuertes Digitalgerät, das seine eigene Eingangs- und Ausgangs-Synchronisation besorgt. Er verfügt auch über eigene Eingangs- und Ausgangs-"Flags", die für den geordneten Informationsfluß durch den Clipping-Teiler sorgen. Liegt eine Linie völlig außerhalb des Gesichtsfeldes, so akzeptiert der Clipping-Teiler einen neuen Input, ohne seine "Output Flags" zu zeigen. Dadurch gelangen nur die sichtbaren Teile gänzlich oder teilweise sichtbarer Linien durch den Clipping-Teiler.

## Ergebnisse

Gegen Ende des Jahres 1966 und Anfang 1967 führte ich am Lincoln Laboratory von MIT eine Reihe vorführender Experimente auf dem Gebiete der dreidimensionalen bzw. räumlichen Darstellung durch. Wir arbeiteten damals mit einem relativ primitiven optischen System, durch das sich die Information nur einem Auge des Betrachters darbieten ließ. Der Kopfpositions-Sensor funktionierte gut genug, um die Kopfhaltung einige Minuten hindurch hinlänglich genau zu messen, ehe kumulative Fehler zu stark anwuchsen. Die Koordinatentransformationen und Perspektivenberechnungen wurden von der Software im

TX-2 durchgeführt. Die Clipping-Funktion gab es damals noch nicht: war irgendein Teil einer Linie nicht auf dem Schirm, verschwand die ganze Linie.

Selbst mit diesem relativ primitiven System gelang es, die Illusion der Räumlichkeit zu verwirklichen. Die Anwender nahmen selbstverständlich jene Positionen oder Haltungen ein, die die von ihnen gewünschte Sicht erforderte. So etwa konnte man die "Größe" eines dargestellten Würfels dadurch messen, daß man festhielt, wie weit der Betrachter sich zu bewegen hatte, um seine Sichtlinie mit der linken oder der rechten Seitenfläche des Würfels "fluchten" zu lassen.

Bei diesem vorführenden Experiment stießen wir auf zwei seltsame und bisher unerklärliche Phänomene. Zunächst einmal waren, da die von uns dargebotene Information aus sozusagen durchsichtigen "Drahtrahmen"-Bildern bestand, unterschiedliche bzw. zweideutige Interpretationen des Gesehenen immer noch möglich. So etwa wurde in einem Bild ein kleiner Würfel oberhalb eines größeren plaziert, um den Eindruck eines Kamins auf einem Haus zu vermitteln. Aus Blickpunkten unterhalb der Dachlinie, wo man den "Kamin" sozusagen "von innen" sah, bedurfte es einiger Konzentration, um sich daran zu erinnern, daß der Kamin ja tatsächlich weiter entfernt war als das Gebäude. Die Erfahrung mit wirklichen, physischen Gegenständen ließ uns darauf beharren, daß, sollte der Kamin überhaupt wahrgenommen werden, er sich vor dem Haus befinden müsse.

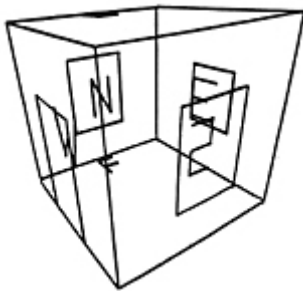


Abb. 5: Vom Computer gezeigte perspektivische Ansicht eines Zyκλο-Hexan-Moleküls

Ein weiteres merkwürdiges Phänomen trat bei der Darstellung der topographischen Verbindungsstruktur von Zyκλο-Hexan (siehe Abbildung 5) auf. Betrachter, denen die merkwürdig wellig-hexagonale Struktur dieses Moleküls nicht vertraut war, deuteten seine Form falsch. Da ihre Sicht des Gegenstandes auf bestimmte Richtungen beschränkt war, konnten sie das Molekül nicht "von oben" sehen. Nun ist es aber die Draufsicht, bei der sich die hexagonale Form am deutlichsten zeigt. Andererseits erkannten mit Molekularstrukturen vertraute Betrachter den "Gegenstand" ohne weiteres als Zyκλο-Hexan.

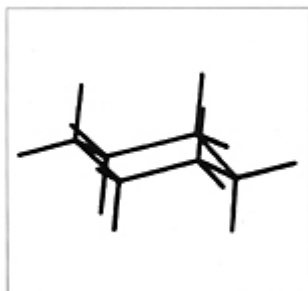


Abb. 6: Vom Computer gezeigte perspektivische Ansicht des von außen gesehenen "Raumes"

In neueren Experimenten mit dem verbesserten optischen System und sehr wesentlich erweiterter Computer-Leistungsfähigkeit wurden zwei Arten von Gegenständen dargestellt.

Bei einem derartigen Test wird ein "Raum" dargestellt, der den Benutzer völlig umgibt. In Abbildung 6 sehen wir sozusagen eine "Außenansicht" dieses Raumes. Der Raum besteht aus vier mit N, S, E und W bezeichneten Wänden, einer mit C bezeichneten Decke, und einem mit F bezeichneten Fußboden. Der Betrachter gewöhnt sich relativ rasch daran, sich innerhalb des vom Computer dargebotenen Raumes zu befinden, und vermag durch Kopfdrehung jeden gewünschten Teil des Raumes zu sehen. In einem anderen Test wurde inmitten des Arbeitsbereiches des Benützers ein kleiner Würfel dargestellt. Der Proband kann ihn von jeder gewünschten Seite betrachten.

Bis dato am überraschendsten für uns war die stark positive Reaktion der Benutzer auf gute Stereo-Darbietung. Das binokulare optische System bietet jedem der beiden Augen voneinander völlig unabhängige Bilder dar. Überdies ist das System mit einer mechanischen Einrichtung versehen, die es erlaubt, das Gerät auf den persönlichen Pupillenabstand verschiedener Benutzer einzustellen. Daneben gestatten Software-Einstellungen in unseren Testprogrammen, den für die Stereo-Berechnungen verwendeten Wert der virtuellen Augenseparation individuell anzupassen. Mit diesen beiden Anpassungsmöglichkeiten ist es relativ einfach, hervorragende Stereo-Darbietungen zu präsentieren. Der Stereovision fähige Betrachter waren sich alle über den Realismus der erzielten Bilder einig.

### **Danksagung**

Als ich die Arbeit am 3D-Kopf-Display begann, hatte ich noch keine Ahnung, wie anstrengend diese Arbeit sein würde. Das Projekt wäre ohne den Geist und die Einstellung der zahlreichen Menschen, die schließlich damit zu tun bekamen, viele Male gescheitert. Der Ultraschall-Sensor zur Ermittlung der Kopfposition und Haltung wurde von Charles Seitz und Stylianos Pezaris im Lincoln Laboratory von MIT konstruiert und entwickelt und steht uns zum dauernden Gebrauch dank der Unterstützung der Lincoln Group 23 zur Verfügung. Seitz, als Angestellter der Harvard University, entwickelte in der Folge den Matrix-Multiplikator. Robert Stroull, ein hochbegabter Harvard-Student im letzten Studienjahr, simulierte, entwickelte im wesentlichen, baute zum Teil, und eliminierte schließlich die Fehler und Schwächen des Clipping-Teilers. Zwei Studenten im höheren Studienabschnitt, Ted Lee und Dan Cohen, waren das ganze Projekt über wesentlich an den Entwicklungsarbeiten beteiligt. Unsere zahlreichen Diskussionen über perspektivische Darstellung, Clipping, die Algorithmen, um normalerweise nicht sichtbare Linien zum Verschwinden zu bringen, und andere Sachthemen bilden eine der stimulierendsten Lehr- und Lernerfahrungen meines bisherigen Berufslebens. Ted Lees Programme für die Stereo-Darstellung gekrümmter bzw. sphärischer Oberflächen bildeten die Grundlage zahlreicher Experimente. Cohens Programme für die Erprobung des gesamten Systems setzen uns überhaupt erst in die Lage, Demonstrationen des Systems zu geben. Danken möchte ich hier auch Quintin Foster, der die Konstruktion, den Bau und die Entstörung der Anlagen überwachte.

Schließlich gebührt unser Dank auch dem sogenannten "project engineer" Steward Ogden, tatsächlich der oberste Administrator, der uns allen die Lasten des Papierkrams vom Leibe hielt, so daß wir mit der eigentlichen Arbeit vorankommen konnten.