

Stereo-Algorithmen

Christian Winkler

9. Februar 2005

Inhaltsverzeichnis

1 Stereo-Algorithmen	2
2 Arbeitsweise klassischer Stereo-Algorithmen	2
2.1 Korrespondenzbasierte Verfahren	2
2.2 Intensitätsbasierte Verfahren	3
2.3 Globale Verfahren	4
2.4 Probleme der klassischen Verfahren	4
3 Stereo-Sehen durch Coherence Detection	5
3.1 Arbeitsweise von Coherence Detection	5
3.2 Vorteile von Coherence Detection	6
4 Quellen	8

1 Stereo-Algorithmen

Stereo-Algorithmen werden benötigt, um aus zwei Stereobildern Tiefeninformationen zu gewinnen. Diese Tiefeninformationen entsprechen der Distanz zwischen zwei korrespondierenden Pixeln in den Stereobildern, der Disparität. Für das gesamte Bild wird eine Disparitätenkarte erstellt, die für jedes Pixel die Disparität speichert. In weiteren Verfahren werden die Disparitätenkarten zu diskreten Tiefeninformationen umgerechnet.

Es existieren viele verschiedene Stereo-Algorithmen. In dieser Arbeit wird die Arbeitsweise von drei klassischen Verfahren vorgestellt sowie ein neuer Algorithmus, der mehrere klassische Methoden vereint.

2 Arbeitsweise klassischer Stereo-Algorithmen

Ein Stereo-Algorithmus arbeitet meist auf einem kleinen Bildausschnitt des gesamten Bildes; dies macht die Bestimmung der Disparität genauer und schneller. Die ersten hier vorgestellten Algorithmen arbeiten auf diese Weise, die globalen Verfahren benutzen jedoch meist das gesamte Bild als Grundlage.

2.1 Korrespondenzbasierte Verfahren

Korrespondenzbasierte Verfahren zur Bestimmung der Disparität beziehen sich auf die Übereinstimmung verschiedener Merkmale der Stereobilder. Dieser Algorithmus kann in vier Grundschritte zerlegt werden. Die später vorgestellten Verfahren arbeiten auf ähnliche Weise, ohne dass dies in dieser Arbeit explizit dargestellt wird.

Zu Beginn werden die verschiedenen Merkmale eines Bildausschnittes extrahiert. Es handelt sich um Merkmale wie zum Beispiel Kantenverläufe, Linienenden oder andere geometrische Figuren (Kreise, Quadrate). Des Weiteren ist eine genaue Lokalisation dieser Merkmale wichtig, da daraus die Disparität berechnet wird.

Wenn alle Merkmale des Bildes erkannt und lokalisiert wurden, muss die Zuordnung der Merkmale zwischen den Bildern geschehen. Das heißt, dass möglichst jedem Merkmal des einen Bildes ein Merkmal des anderen Bildes zugeordnet werden soll. Um diese Aufgabe zu vereinfachen, gibt es vier Strategien.

1. Hat man ein Merkmal in der linken, oberen Ecke des linken Bildes gefunden, sucht man das korrespondierende Merkmal nicht in der rechten Ecke des rechten Bildes, sondern ebenfalls links. Da die Verschiebung zwischen den Bildausschnitten meist nur wenige Pixel beträgt, kann man auf diese Weise den Suchraum sehr gut einschränken.
2. Eine weitere Erleichterung wird durch die Zuordnung anhand von Merkmalseigenschaften geschaffen, das heißt, dass man beispielsweise Linienendpunkte nur Linienendpunkten zuordnet.

3. Außerdem sollte man die Ordnung und Eindeutigkeit der Merkmale beachten, was bedeutet, dass ein Merkmal, welches sich rechts von einem bereits zugeordneten Merkmal befindet, seine Korrespondenz auch rechts des bereits zugeordneten Merkmals hat. Zudem wird jedem Merkmal genau ein korrespondierendes Merkmal zugeordnet, wobei eventuelle Verdeckungen beachtet werden müssen.
4. Eine weitere Hilfe bei der Zuordnung der Merkmale ist das Vorgehen mit einer Grob-zu-Fein-Strategie, bei der die Merkmale erst mit einer groben Auflösung zugeordnet werden, welche im Verlauf der Berechnung feiner wird.

Wenn alle Merkmale zugeordnet sind und die Disparitäten bestimmt wurden, kann man anhand der Kamerageometrie die Tiefeninformation berechnen.

Als letzten Schritt muss man die Oberflächenstruktur der Objekte des Bildes durch Interpolation berechnen, da man durch die vorangegangenen Berechnungen nicht für alle Punkte Informationen erhält.

2.2 Intensitätsbasierte Verfahren

Es existieren verschiedene Verfahren zur Bestimmung der Disparität anhand von Intensitätswerten. Der hier vorgestellte Algorithmus bezieht sich auf die Korrelation, also die Übereinstimmung zwischen zwei Grauwertverläufen. Ähnlich einer Filterfunktion wird für jedes Pixel ein Grauwertverlauf über einem Bildausschnitt bestimmt. Dieser wird mit dem Grauwertverlauf des entsprechenden Pixels des anderen Bildes verglichen und so verschoben, dass die größte Ähnlichkeit besteht. Diese wird durch die quadratische Abweichung der Grauwertverläufe berechnet. So erhält man für jedes Pixel die Verschiebung zwischen den beiden Bildern; daraus erstellt man die Disparitätenkarte.

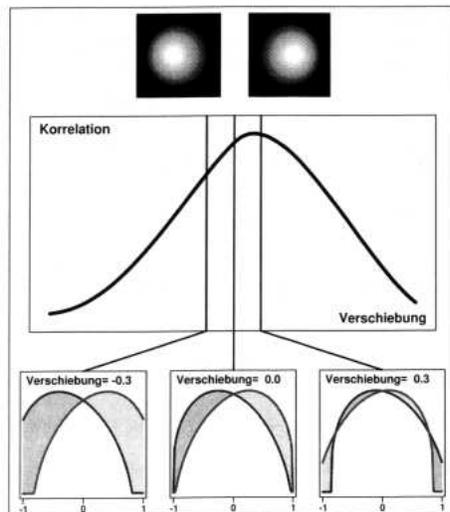


Abbildung 1: intensitätsbasiertes Verfahren

2.3 Globale Verfahren

Globale Verfahren eignen sich für Anwendungen, die keine hochauflösenden Tiefenkarten benötigen. Dies sind beispielsweise einfache Hinderniserkennungen, bei denen nur wichtig ist, ob ein Objekt in kritischer Entfernung steht oder nicht.

Für solche Anwendungen nutzt man die oben beschriebenen klassischen Verfahren, lässt sie aber auf großen Bildausschnitten oder dem gesamten Bild arbeiten.

Ein anderer Algorithmus, der auch zu den globalen Verfahren zählt, ist die Inverse Perspektive.

Bei dieser Methode berechnet man ein künstliches Bild, indem man das Bild der linken Kamera auf die Grundebene projiziert und dieses Bild dann aus Sicht der rechten Kamera betrachtet. Das künstlich erzeugte Bild wird vom realen rechten Bild subtrahiert und man erhält ein Bild, welches nur Objekte enthält, die ober- oder unterhalb der Grundebene liegen. Dadurch lassen sich Objekte, die ein Hindernis darstellen, erkennen.

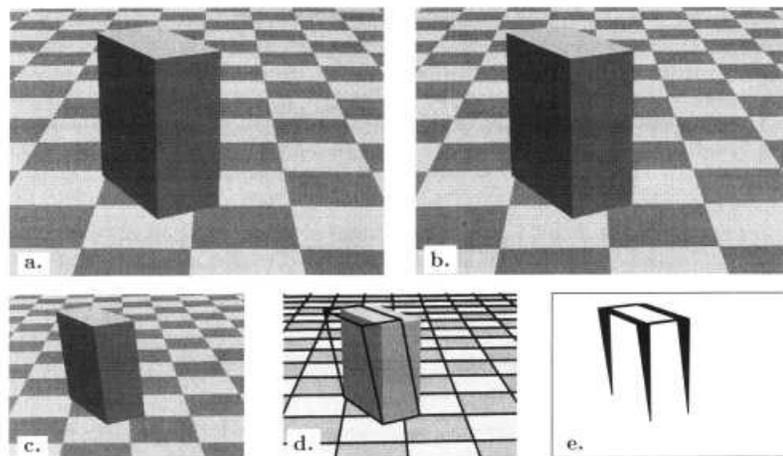


Abbildung 2: Inverse Perspektive

2.4 Probleme der klassischen Verfahren

Alle hier vorgestellten Algorithmen beinhalten für sich verschiedene Probleme. Oft arbeitet ein System für eine bestimmte Anwendung sehr gut, für eine andere Anwendung eignet sich ein anderer Algorithmus jedoch wesentlich besser. Man kann nie sagen, welches der beste Algorithmus für alle Anwendungen ist, da die Wahl des Verfahrens stark von den zu bearbeitenden Daten abhängt.

Bei den merkmalsbasierten Verfahren kann es zum Beispiel oft dazu kommen, dass man keine Übereinstimmungen findet, außerdem bieten sie nur approximierete Ergebnisse. Alle Methoden sind sehr rechenintensiv und können durch Aliasing-Effekte verfälscht werden. Unter Aliasing-Effekten versteht man in der Stereo-Vision die fälschliche Zuordnung von Bildausschnitten. Abbildung 3 zeigt, dass die Zuordnung von Grauwertverläufen bei kontinuierlichen Werten

kein Problem darstellt (Bild links). Bei diskreten Werten wird die korrekte Zuordnung jedoch erschwert. Ist die Verschiebung zwischen den Bildern nicht allzu groß (Bild mitte), kann noch richtig zugeordnet werden. Sobald der Schritt zwischen den Bildern jedoch zu groß wird (Bild rechts), ist eine richtige Zuordnung nicht mehr möglich.

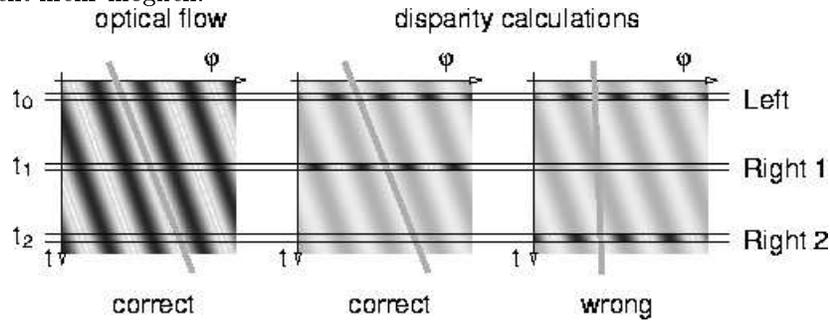


Abbildung 3: Aliasing-Effekt

3 Stereo-Sehen durch Coherence Detection

Beim Stereo-Sehen durch Coherence Detection handelt es sich um einen neu entwickelten Algorithmus, der eine schnelle Verarbeitung von Stereobildern möglich macht. Entwickelt wurde das Verfahren von Rolf Henkel an der Universität Bremen. Die Firma "3d Image Processing", für die Rolf Henkel zur Zeit arbeitet, bietet eine Komplettlösung für dieses Verfahren unter dem Namen "COBAM" (Coherence Based Modelling) an.

3.1 Arbeitsweise von Coherence Detection

Das kohärenzbasierte Verfahren kann man sich als ein Netzwerk von Disparitätsberechnungsfunktionen (*disparity estimator*) vorstellen. Jeder dieser disparity estimators entspricht der Anwendung eines klassischen Verfahrens. Die estimators sind in sogenannten *disparity stacks* geordnet, alle estimators einer Ebene sind identisch, die einzelnen Ebenen dürfen jedoch voneinander abweichen.

Die Bilddaten werden pixelweise verarbeitet, wobei sie diagonal durch das Netzwerk laufen. Hinzu kommt, dass beide Bilder um eine bestimmte Anzahl von Pixeln (*preshift*) verschoben werden, und diese parallel durch das Netzwerk laufen. Dadurch erreicht man eine Vergrößerung des Disparitätsbereichs und wirkt Fehlberechnungen durch Auftreten des oben beschriebenen Aliasing-Effektes entgegen bzw. nutzt diesen Effekt aus.

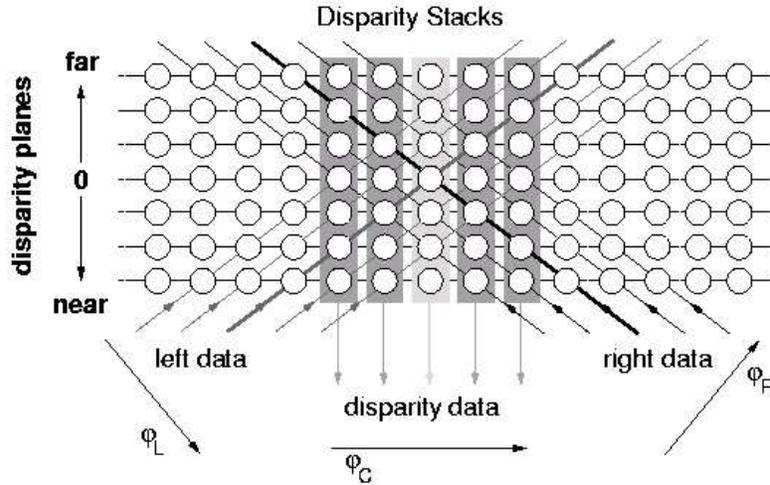


Abbildung 4: Netzwerkstruktur von Coherence Based Stereo Vision

Bei jedem Verarbeitungsschritt berechnet jeder disparity estimator ein Ergebnis für das entsprechende Pixel. Durch die Struktur ist der Arbeitsbereich $D_i = [d_i^{min}, d_i^{max}]$ jedes estimators festgelegt. Das bedeutet, dass es sich um eine korrekt berechnete Disparität handelt, wenn der Wert innerhalb des Intervalls liegt. Der Stack wird nun in zwei Klassen C und \bar{C} eingeteilt, wobei C alle estimators enthält, deren Disparitätswert innerhalb des Intervalls liegt ($d \in D_i$). Für estimators i, j der Klasse C gilt also: $d_i \approx d \approx d_j$. Dann wird der Mittelwert der Klasse C gebildet und dieser wird als Ergebnis des Stacks ausgegeben. Für die weitere Verarbeitung wird ein Wert aus allen Stackergebnissen bestimmt. Auf das genaue Verfahren wird in dem Artikel, der dieser Ausarbeitung zugrunde liegt, jedoch nicht weiter eingegangen. Dieser Wert ist die Disparität an dem bestimmten Pixel und wird in die Disparitätenkarte eingetragen.

Um eine größere Genauigkeit zu erzielen, werden mehrere Durchläufe mit verschiedenen disparity estimators durchgeführt.

3.2 Vorteile von Coherence Detection

Der größte Vorteil, den dieser Algorithmus bietet, ist seine Geschwindigkeit. Angelehnt an das menschliche Sehen kann dieser Algorithmus als neuronales Netz implementiert werden, wobei die Disparitätsberechnungsfunktionen als einfache Filteroperationen realisiert werden können. Durch den Aufbau der Stacks kann das System leicht 'nach oben' erweitert werden, um weitere estimators mit in die Berechnung einzubeziehen.

Da die disparity estimators im Einzelnen den klassischen Verfahren entsprechen, durch den Aufbau jedoch über mehrere Verfahren gemittelt wird, ist das

System sehr flexibel in der Anwendung und robust gegenüber einzelner Fehlrechnungen.

Des Weiteren wird ein Gütemaß (*verification count*) für die berechnete Disparität mitgeliefert; es müssen lediglich die estimators gezählt werden, die der Klasse C bei der Berechnung zugeordnet werden. Liegt die Anzahl unter einem bestimmten Wert, wird diese Berechnung als ungültig eingestuft. ($\frac{N(C)}{N(C \cup C)}$ mit $N(C)$ als Anzahl der estimators der Klasse)

Ein weiterer Vorteil ist, dass die zyklische Sicht als Nebeneffekt entsteht. Die zyklische Sicht entspricht der Sicht einer Kamera, die zwischen den zwei Kameras des Stereosystems angebracht ist. Mit diesem Algorithmus kann jede beliebige Blickrichtung zwischen den Stereokameras simuliert werden. Jeder disparity stack entspricht einer dieser Blickrichtungen. Zur Berechnung wird der Mittelwert der Pixelwerte (I_i^L, I_i^R) der disparity estimators der Klasse C berechnet:

$$I^C = \langle I_i^L + I_i^R \rangle_{i \in C}$$

Zur Veranschaulichung zeigt Abbildung 5 die verschiedenen Blickrichtungen, die zwischen den Ursprungsbildern berechnet werden können.

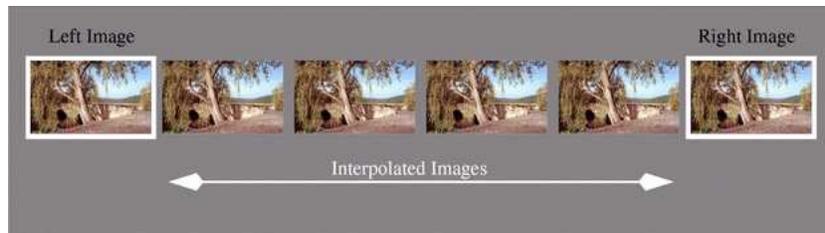


Abbildung 5: die zyklische Sicht

Abbildung 6 zeigt das rechte und das linke Bild eines Stereo-Systems (A, B), sowie einige Ergebnisse, die mit *Coherence Based Stereo Vision* errechnet wurden. In Ergebnisbild E sind sehr viele Störungen zu sehen. Bild F zeigt Bild-daten mit niedrigen verification count, wohingegen in Bild G diese "unsicheren" Ergebnisse ausgeblendet sind. C und D zeigen den Unterschied zwischen einfachem Übereinanderlegen der Bilder (C) und der zyklischen Sicht (D).

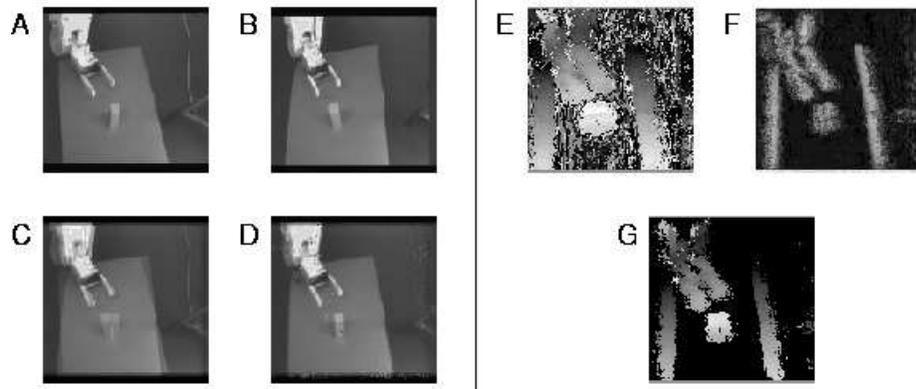


Abbildung 6: Ergebnisse von Coherence Based Stereo Vision

4 Quellen

- -<http://axon.physik.uni-bremen.de/research/stereo/index.html>
Internetpräsenz von Rolf Henkel an der Universität Bremen zum Thema "Stereo Vision", Dezember 2004
- -Mallot, Hanspeter A.: "Sehen und die Verarbeitung visueller Information", Vieweg, 2000
- -<http://www.3d-ip.com>
Internetpräsenz der Firma "3d Image Processing", Dezember 2004