

Virtuelle Gerüste

Heimo Müller, Behnam Tabataba

(Dr. Heimo Müller¹¹, Behnam Tabataba; Institut für Informationssysteme, Joanneum Research, Steyrergasse 17, A-8010 Graz)

Abstract

Virtual Reality benötigt nicht immer die teuersten Computer und aufwendigsten Betrachtungsgeräte. Mit Hilfe eines Stereobildbetrachters in Pocketform kann man simulierte Architektur in einfachster Weise transportieren und einer breiten Öffentlichkeit zugänglich machen.

Das in Joanneum Research vorhandene Know-how aus High-Tech Anwendungsgebieten, wie Raumfahrt und wissenschaftliche Visualisierung, wurde in einem Architekturprojekt des Europäischen Kulturmonats angewandt. Die Aufgabe lautete, virtuelle, mit Literaturzitate versehene Gerüste in Graz zu installieren. Die Stereobildtechnik fand hierbei äußerst erfolgreich in einem neuen Anwendungsgebiet ihren Einsatz.

Die Projektergebnisse wurden neben den Stereobildern noch in einem Kinokurzfilm, in einer Ausstellung des Stadtmuseums Graz und auf 200 Plakatwänden gezeigt. Weiters wurden die Architektursimulationen und die darauf erhaltenen Reaktionen in Buchform veröffentlicht und an europäische Kulturpolitiker verteilt.

1. EINLEITUNG

„Im Zentrum Fedoras, der Metropole aus grauem Stein, steht ein metallener Palast mit einer Glaskugel in jedem Zimmer. In jeder Kugel erblickt man beim Hineinsehen eine blaue Stadt, das Modell für ein anderes Fedora“ schreibt Italo Calvino in seinem Roman „Die unsichtbaren Städte“. Fedora war auch der Titel eines Architektur- und Literaturprojektes, welches sich mit dem Stadtraum als Kulturraum beschäftigte, durchgeführt im Europäischen Kulturmonat der Landeshauptstadt Graz.

Auf überdimensionalen Gerüsten wurden Literaturzitate mit Bezug zum Standort präsentiert. Das Gerüst als Werkzeug des Bauens und Veränderns wird als Transportmedium - vom privaten Raum eines Buches in den öffentlichen Stadtraum - für Literatur genutzt. Die Gerüste wurden jedoch nicht in den globalen Nord-Süd und Ost-West Achsen in eine reale Architektur umgesetzt, sondern in Form einer „Utopie“ so der künstlerische Leiter Architekt Ernst Giselsbrecht, über ein multimediales Netzwerk, bestehend aus Film, Plakaten, Kinospot, Video und Faxkommunikation, präsentiert.

Mit Hilfe von Visualisierungsprogrammen wurden die Gerüste in Stereobilder eingerechnet und auf die verschiedensten Ausgabemedien konvertiert. Folgende Arbeitsschritte waren dazu notwendig: Aufnahme der Stereobilder, photogrammetrische Auswertung (Bestimmung der Kameraparameter), Automatisches Erstellen von Masken, Modellieren der Gerüste, Rendering der Bilder, Montage der computer-generierten Bilder mit den Originalbildern und schließlich die Ausgabe auf verschiedenste Medien. Hervorzuheben ist dabei die Herstellung von Stereobildern und die Produktion eines 35mm Kinokurzfilmes. Eine Pocketversion der Stereobetrachter ermöglicht es, den Stadtraum mitzunehmen bzw. an interessierte Personen zu versenden.

2. BESTIMMUNG DER KAMERAPARAMETER

2.1. Perspektivische Transformation

Bei der Bestimmung der Kameraparameter gehen wir von einer perfekten perspektivischen Transformation aus. Diese Annahme ist bei professionellen und halbprofessionellen Aufnahmegegeräten vertretbar. Bei

¹¹ Studium der technischen Mathematik an der TU-Graz, Dissertationsstudium an der TU-Wien betreut von Prof. Werner Purgathofer und Prof. Walter Kropatsch.

Mitarbeiter am Institut für digitale Bildverarbeitung und Computer Graphik (DIBAG) sowie am Institut für Informationssysteme (IIS) der Joanneum Research Graz.

Dokumenteditor von ISO/IEC JTC1 SC24 (Computer Graphic and Image Processing) WG7, Forschungsschwerpunkte: Datenaustausch, Computer Graphik und Digitale Medien.

anderen handelsüblichen Geräten, vor allem bei Weitwinkelaufnahmen, treten Verzerrungen auf, die unter Zuhilfenahme von zusätzlichen Paßpunkten extra berechnet werden müssen.

Die perspektivische Transformation (bei Vernachlässigung der Verzerrungen) wird durch folgende 9 Parameter beschrieben (siehe Abbildung 1):

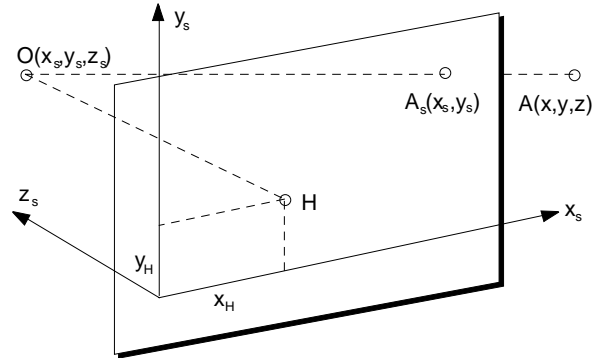


Abbildung 1: Perspektivische Transformation

- Parameter der äußeren Orientierung
 - Kameraposition (x_0, y_0, z_0)
 - Orientierung des Bildkoordinatensystems (φ, θ, ψ)
- Parameter der inneren Orientierung
 - Position der Kamera im Bildkoordinatensystem, d.h., Lage des Hauptpunktes und die Brennweite (x_H, y_H, c)

Die Abbildung von Weltkoordinaten (x, y, z) in die Bildkoordinaten (x_s, y_s) wird somit durch folgende Gleichungen definiert:

$$x_s = x_H - c \frac{(x_0 - x)m_{11} + (y_0 - y)m_{12} + (z_0 - z)m_{13}}{(x_0 - x)m_{31} + (y_0 - y)m_{32} + (z_0 - z)m_{33}}$$

$$y_s = y_H - c \frac{(x_0 - x)m_{21} + (y_0 - y)m_{22} + (z_0 - z)m_{23}}{(x_0 - x)m_{31} + (y_0 - y)m_{32} + (z_0 - z)m_{33}}$$

mit

$$[m_{ij}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \psi & \sin \psi \\ 0 & -\sin \psi & \cos \psi \end{bmatrix} \cdot \begin{bmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{bmatrix} \cdot \begin{bmatrix} \cos \varphi & \sin \varphi & 0 \\ -\sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Für die Bestimmung der obengenannten 9 Kameraparameter werden somit zumindest 5 räumlich gut verteilte Paßpunkte benötigt (2 Gleichungen pro Paßpunkt). Die Lösung des nichtlinearen, überbestimmten Gleichungssystems erfolgt etwa durch das Newton-Raphson-Verfahren mit gleichzeitigem Fehlerausgleich nach Gauß (L_2 -Norm). Es ist jedoch zu beachten, daß für dieses iterative Verfahren ein sicherer Startvektor benötigt wird, wofür man geschätzte Werte für die Kameraparameter einsetzen kann (ungefähre Position des Aufnahmegeräts, Brennweite aus technischen Unterlagen, Hauptpunkt in der Bildmitte, φ und θ gleich Null bei horizontalen Aufnahmen und ψ geschätzt).

In Ermangelung plausibler Schätzwerte für den Startvektor kann man sich der Methode der „Direkten Linearen Transformation“ (DLT) bedienen. Es handelt sich dabei um eine lineare Transformation $\mathbf{R}^3 \mapsto \mathbf{R}^2$, beschrieben durch die Gleichungen

$$x_s = \frac{b_{11}x + b_{12}y + b_{13}z + b_{14}}{b_{31}x + b_{32}y + b_{33}z + 1}$$

$$y_s = \frac{b_{21}x + b_{22}y + b_{23}z + b_{24}}{b_{31}x + b_{32}y + b_{33}z + 1}$$

Die 11 unbekannt DLT-Parameter b_{ij} werden aus mindestens 6 (räumlich gut verteilten) Paßpunkten bestimmt. Unter der Annahme, daß es sich bei dieser linearen Abbildung um eine perspektivische Transformation handelt, können nun die inneren und äußeren Orientierungsparameter wie folgt berechnet werden:

- Hilfswerte:

$$\lambda^2 = b_{31}^2 + b_{32}^2 + b_{33}^2$$

$$\lambda = +\sqrt{\lambda^2}$$

- Parameter der inneren Orientierung

$$x_H = \frac{b_{11}b_{31} + b_{12}b_{32} + b_{13}b_{33}}{\lambda^2}$$

$$y_H = \frac{b_{21}b_{31} + b_{22}b_{32} + b_{23}b_{33}}{\lambda^2}$$

$$c_1 = \sqrt{\left| \frac{b_{11}^2 + b_{12}^2 + b_{13}^2}{\lambda^2} - x_H^2 \right|}$$

$$c_2 = \sqrt{\left| \frac{b_{21}^2 + b_{22}^2 + b_{23}^2}{\lambda^2} - y_H^2 \right|}$$

$$c = \frac{c_1 + c_2}{2}$$

- Parameter der äußeren Orientierung
 - Kameraposition

$$\begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} = - \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}^{-1} \begin{bmatrix} b_{14} \\ b_{24} \\ 1 \end{bmatrix}$$

- Rotationsmatrix

$$m_{3k} = -\lambda b_{3k}$$

$$m_{1k} = (\lambda b_{1k} + x_H m_{3k}) / c$$

$$m_{2k} = (\lambda b_{2k} + y_H m_{3k}) / c$$

2.2. Disparitätsfeld und epipolare Geometrie

Die Grundlage für die Bestimmung des Tiefenmodells ist die Parallaxe der korrespondierenden Punkte in den digitalisierten Stereo-Bildern. Sie äußert sich in diesem Fall als Verschiebung der Pixelkoordinaten in den rechnerisch zunächst nicht korrelierten Links- und Rechtsaufnahmen. Um den Suchbereich für den Stereo-Matcher einzuschränken, ist es nun notwendig, die Stereo-Bilder in einer epipolaren Geometrie zu korrelieren, so daß während eine Zeile im linken Bild verarbeitet wird, die korrespondierende Zeile im rechten Bild bekannt ist. Man erhält die epipolare Geometrie, in dem man mittels einer affinen Abbildung das rechte Bild in die Geometrie des linken Bildes transformiert. Die Grundgleichung für diese affine Transformation besteht aus einem rotatorischen (\mathbf{A}) und einem translatorischen Anteil (\mathbf{T}):

$$\begin{bmatrix} x_R \\ y_R \end{bmatrix} = \mathbf{A} \cdot \begin{bmatrix} x_L \\ y_L \end{bmatrix} + \mathbf{T}$$

mit

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \text{ und } \mathbf{T} = \begin{bmatrix} a_{01} \\ a_{02} \end{bmatrix}$$

Für die Bestimmung der 6 unbekannt Parameter benötigt man mindestens 3 Paßpunkte. Als Paßpunkte wählt man vorzugsweise mehrere in genügender Entfernung liegende Punkte (verschwindende Parallaxen; für die bei uns verwendeten Bildformate von 11 x 6 cm und eine Digitalisierungsauflösung von 1200 dpi war dies eine Entfernung von etwa 100 Metern), und löst das dabei entstehende überbestimmte Gleichungssystem mit einem Fehlerausgleich nach Gauß.

Das Disparitätsfeld ist nun definiert als der Verschiebungsvektor $\mathbf{d}_R(\mathbf{p})$, der notwendig ist, um einen vom linken Bild ins rechte transformierten Punkt \mathbf{p}_{LR} in die Position desselben Punktes im rechten Bild \mathbf{p}_R zu versetzen:

$$\mathbf{p}_{LR} = \mathbf{p}_R + \mathbf{d}_R$$

Man beachte, daß wegen der epipolaren Geometrie die y-Komponenten des Vektors \mathbf{d} nur noch verschwindend klein sind. Daher kann das Matching in einem ein-dimensionalen Bereich stattfinden. Die für das Matching verwendeten Algorithmen stammen aus dem Bereich Computervision. Dabei wird die Korrelierung der Bilder in verschiedenen Auflösungsebenen (Pyramidenansatz) vorgenommen. Die dazu benötigte Software wurde am Institut für digitale Bildverarbeitung entwickelt. [PAAR94a,b]

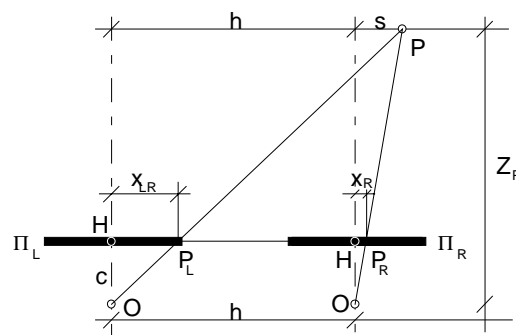


Abbildung 2: Entfernungsberechnung

Mit Kenntnis über die inneren Orientierungsparameter ist es nun möglich, die 3-dimensionalen lokalen Koordinaten der Bildpunkte (vor allem deren Entfernungen) zu bestimmen und diese bei Bedarf mit Hilfe der äußeren Orientierungsparameter in das globale Koordinatensystem zurückzuführen:

$$z = \frac{c \cdot h}{x_{LR} - x_R} = \frac{c \cdot h}{d_R}$$

mit c als Brennweite und h als Achsenabstand der beiden Aufnahmengereäte.

Das für das gesamte Bild berechnete Tiefenmodell wird in einem nächsten Schritt mit der Tiefeninformation aus dem 3D Architekturmodell verglichen. Das Ergebnis ist eine sogenannte Bildmaske, die angibt, welche Bildeinheiten des Hintergrundbildes durch das (virtuell) in die Szene gestellte Gerüst verdeckt werden. Weiters liefern die Konturen der Bildmaske die Stellen, wo die Antialiasing-Algorithmen zur Eliminierung der optischen Unstetigkeiten beim Image-Compositing angewendet werden müssen.

3. MODELERSTELLUNG UND RENDERING

Die einzelnen Gerüste wurden als 3D Architekturmodelle an Hand von Detailplänen konstruiert. Dazu wurde das Animations-, Renderingprogramm Creative Environment (Softimage Inc.) verwendet. Die beiden Kamerapositionen wurden hierbei der photogrammetrischen Auswertung entnommen, und die Koordinatensysteme der Tiefeninformation (Z-Buffer) kalibriert. Nach der Modelerstellung wurden die

Materialeigenschaften definiert und die Lichtverhältnisse und der Schattenwurf ausgehend von den Beleuchtungsverhältnissen der Realbilder definiert. Hierbei wurde auch die Tiefenschärfe und Kameranichtlinearitäten der Computer generierten Bilder an das Stereobildmaterial angepaßt.



Das Rendering der Bilder erfolgte in der höchsten erforderlichen Auflösung. Bildmaterial mit niedriger Auflösung (z.B für Video) wurden durch „downsampling“ (Interpolation) erzeugt. Gegeben durch die sehr hohe Auflösung und die hohen Anforderungen an die Bildqualität waren selbst auf einer High-End Graphik Workstation (Silicon Graphics Crimson VGX, 160 MB) Renderingzeiten von bis zu 6 Stunden für ein Bild notwendig.

Das Zusammensetzen (Image Compsiting) der Hilfer der automatisch erzeugten Bildmasken und eine händische Endkorrektur erfolgte mit dem 2D Program Eddie (Softimage Inc.).

4. AUSGABEMEDIEN

Zur Präsentation der Ergebnisse wurden Video und 35 mm Film, Stereobilder, Plakate (200 verteilt über das Stadtgebiet von Graz) und eine Ausstellung im Grazer Stadtmuseum, welche mittels Telekommunikation das Projekt Fedora einer größeren, überregionalen Öffentlichkeit präsentierte, herangezogen.

Die Vielzahl der Präsentationsmedien erforderte die Umwandlung der Bilder auf entsprechende Auflösungen (von 720 x 576 bis 5400 x 2700 Bildpunkten) und Farbtiefen (8 Bit pro Bildelement = Monochrom, 16 Bit pro Bildelement = YUV und 24 Bit pro Bildelement = RGB).

5. ZUSAMMENFASSUNG UND AUSBLICK

Ein semiautomatisches Verfahren zur Auswertung von Stereobildern und zur Bestimmung der Tiefeninformation wurde vorgestellt. Die Tiefeninformation eines Stereobildes wurde verwendet um eine Bildmaske für das Zusammensetzen eines Stereobildes mit den computer-generierten Architekturmodellen zu erzeugen. Die Ausgabe der so gewonnenen Stereobilder, wieder im Stereoformat oder als konventionelle dynamische bzw. statische Bilder, wurde beschrieben.

Eine Integration der Ergebnisse der Sterobildauswertung in ein Architekturinformationssystem und das automatische Erzeugen von 3D Modellen (Koordinaten und Textur) aus den Stereobildern ist die logische Erweiterung der vorgestellten Konzepte.

6. DANKSAGUNG

Unser Dank gilt allen Projektbeteiligten, insbesondere DI Ernst Giselbrecht für Idee und Konzeption, Werner Schrenpf (die ORGANISATION) für Projektkoordination und Ausstellungsgestaltung, Johannes Pötscher (Stereobildtechnik Pötscher) für die Aufnahme der Stereobilder, Michael Gruber (Institut für Computergrafik der TU Graz) für seine Hilfe bei der Stereobildauswertung, Gerhard Kohlmaier, Samir Madwar und Thomas Bauer (Digital Canvas) für die Unterstützung in graphischen Belangen, sowie unseren Kollegen Werner Haas, Harald Mayer, Georg Thallinger (Institut für Informationssysteme) und Gerhard Paar (Institut für digitale Bildverarbeitung) von Joanneum Research für Diskussion und Anregungen.

7. LITERATUR

- [ALVE89] Alvertos Nicos, Brzakovic Dragana, Gonzalez Rafael: Camera Geometries for Image Matching in 3D Machine Vision, Transaction on Pattern Analysis and Machine Intelligence, Vol. 11. No. 9 September 1989, pp. 897-914.
- [ALVE86] Alvertos Nicos, Brzakovic Dragana, Gonzalez Rafael: Correspondance in pairs of images aquired by camera diplacement in depth, in SPIE Vol. 726, Intelligent Robots and Computer Vision, Oct. 1986, pp.131-136.
- [BOPP78] Bopp Hanspeter, Krauss Herbert: An Orientation and Calibration for Non-Topographic Applications, in Photogrammetric Engineering and Remote Sensing, Vol. 44, No. 9, September 1978, pp. 1191-1196.
- [ERTL91] Ertl Gerhard, Heimo Müller, Tabatabai Behnam: Move-X, A System for Combining Video Films and Computer Animation, Proceedings of Eurographics '91, pp. 305-314.
- [HWAN80] Hwang J.J.: Computer Stereo vision for three-dimensional object location, Dep. Elec. Eng., Univ. Tennessee, Tech. Report, August 1980.
- [PAAR94a] Paar Gerhard, Sidla Oliver, Classen Hans-Josef, Parkens Steven, Gillions Andrews: Planetary Body High Resolution 3D Modeling, ESA Phase 1 Report, JOANNEUM RESAERCH 1994.
- [PAAR94b] Paar Gerhard, Ulm Michael, Sidla Oliver: Vision Based Navigation for Moon Landing, Final Report ESTEC 142958, JOANNEUM RESAERCH 1994.