

# Panoramabilder durch Projektion auf Mannigfaltigkeiten

Seminar 3D - Rekonstruktion  
WS 2002/2003

Sarah Mang

Vortrag : 25.11.2002

## **Zusammenfassung**

Da der Blickwinkel einer Kamera immer viel kleiner ist als der des Menschen, benötigt man für eine optimale Darstellung grosser Szenen Panoramabilder. Da es bei der Erweiterung des Blickfeldes mit optischen Hilfsmitteln, wie zum Beispiel speziellen Linsen, zu starken Störungen am Bildrand kommt, benutzt man Mosaik, um aus mehreren kleinen Bildern eine Panorama-Ansicht zu generieren. In der Vergangenheit benutzte man zum Erstellen von Panoramabildern nur Teilbilder, die durch Rotation der Kamera um ihre eigene Achse oder durch Translation der Kamera parallel zu einer Ebene entstanden waren. Diese beiden Methoden stossen allerdings an ihre Grenzen, wenn es um allgemeinere Transformationen der Kameraposition geht. Dies beschränkt deren Einsatzmöglichkeiten und Nutzbarkeit. Um allgemeinere Transformationen in der Erzeugung von Mosaik-Panoramas zuzulassen, wählt man eine Methode, die nicht mehr das ganze Bild im Mosaik verwendet, sondern nur noch Streifen, und dadurch der Kamerabewegung besser folgen kann.

## Inhaltsverzeichnis

<b>1</b>	<b>Motivation</b>	<b>2</b>
<b>2</b>	<b>Einführung</b>	<b>2</b>
2.1	Scanning-Broom Kamera . . . . .	3
2.2	Multiple Center of Projection (COP) . . . . .	4
<b>3</b>	<b>Projektion auf Mannigfaltigkeiten</b>	<b>5</b>
3.1	Alignment . . . . .	6
3.2	Zusammenfügen der Streifen . . . . .	7
3.3	Beispiele . . . . .	8
<b>4</b>	<b>Projektion auf allgemeine Mannigfaltigkeiten</b>	<b>9</b>
4.1	Allgemeine Mannigfaltigkeiten . . . . .	9
4.2	Projektion gerader Linien . . . . .	9
4.3	Alignment der Bildstreifen . . . . .	10
4.4	Beispiel . . . . .	11
<b>5</b>	<b>Röhrenprojektion</b>	<b>12</b>
5.1	Auswahl der Bildstreifen . . . . .	12
5.2	Kombination der Streifen . . . . .	13
5.3	Beispiel . . . . .	16
<b>6</b>	<b>Parallaxe</b>	<b>17</b>
6.1	Beispiel . . . . .	18
<b>7</b>	<b>Vergleich der Algorithmen</b>	<b>19</b>

## Literatur

[1] Shmuel Peleg und Joshua Herman *Panoramic Mosaics by Manifold Projection*

[2] Assaf Zomet *Mosaicing from a Translating Camera, with Efficient Application of Super-Resolution*

[3] Benny Rousso, Shmuel Peleg, Ilan Finci und Alex Rav-Acha *Universal Mosaicing Using Pipe Projection*

[4] <http://www.cs.huji.ac.il/labs/vision/research/synt.html>

## 1 Motivation

Da das Blickfeld der Kamera immer viel begrenzter ist als das des menschlichen Auges, versucht man durch Hilfsmittel das Bild der Kamera zu vergrössern.

Wenn man dazu optische Hilfsmittel, wie etwa eine Fischaugenlinse, verwendet, kommt es an den Rändern der Bilder zu erheblichen Verzerrungen.

Eine andere Möglichkeit Panorama-Ansichten zu generieren ist das Zusammenfügen mehrerer einzelner Bilder, die mit einer normalen Kamera aufgenommen wurden, zu einer Gesamtansicht, dem Panorama-Mosaik.

Die Einsatzgebiete für Panoramabilder sind vielfältig; hier einige Beispiele:

1. **Die Darstellung von Satellitenaufnahmen oder den Aufnahmen von Überwachungskameras**
2. **Die Videokompression:**  
Der Hintergrund einer Szene wird als Mosaik gespeichert. Nur einmal für die ganze Szene und nicht für jeden Frame. Dazu werden für die einzelnen Frames die Informationen über sich bewegende Objekte gespeichert. MPEG-4 arbeitet mit einem solchen Verfahren.
3. **Die Videoindexierung:**  
Durch das Speichern einer Videoszene in einem Panoramabild kann mit Algorithmen aus der Mustererkennung ein automatischer Index erstellt werden.
4. **Die Unterhaltung:**  
Darstellung von Panorama-Ansichten auf Postern, etc.

## 2 Einführung

Bei der Erstellung eines Mosaiks trifft man auf 3 Probleme:

1. **Die Ausrichtung der Bilder (Alignment):**  
Beim Alignment kommt es darauf an, die Transformation des optischen Flusses von einem Bild zum nächsten zu finden, um Streifen aus den Bildern richtig ins Panorama einzusetzen.
2. **Das Zusammenfügen der Bilder (Cut'n'Paste):**  
Die Bilder, aus denen das Mosaik generiert werden soll, überlappen sich im Allgemeinen. Deshalb muss ein geschickter Weg gefunden werden, um die Ausschnitte für das Panorama-Mosaik zu wählen. Hierbei

kommt es auf die bestmögliche Auflösung und die im Alignment-Schritt berechnete Ausrichtung zur Nachbarregion an.

### 3. Das Glätten der Übergänge:

Die Bilder können nach dem Zusammenfügen Farbkanten an den Übergängen aufweisen, selbst wenn die Objekte perfekt passen. Das liegt an veränderten Lichtverhältnissen oder internen Kameraparametern. Diese Farbkanten werden geglättet, um den Eindruck eines Gesamtbildes nicht zu stören.

Klassische Methoden zum Generieren von Mosaiken sind die Rotation der Kamera um ihre eigene Achse oder die Kamera parallel zu einer Ebene zu verschieben. Die Kamerabewegung ist bei beiden Methoden beschränkt. Dazu kommt bei der Rotation noch der erhöhte Informationsbedarf. Ohne Informationen über interne Kameraparameter (Brennweite) kann das Mosaik nicht erstellt werden. Durch diese Bedingungen sind die Einsatzmöglichkeiten für diese Algorithmen sehr beschränkt. Im allgemeinen Fall sind Kameraparameter nicht bekannt und die Kamerabewegung ist beliebig. Moderne Mosaik-Algorithmen kommen auch mit diesen Gegebenheiten zurecht.

Bei klassischen Algorithmen werden die Bilder alle an einem Referenzbild ausgerichtet. Diese Vorgehensweise führt im Allgemeinen zu einer Verringerung der Auflösung. Ausserdem setzten sich Fehler im Alignment über das ganze Panoramabild hinweg fort, was zu einem massiv gestörten Bild führen kann. Moderne Algorithmen richten Bilder deshalb paarweise aneinander aus, um Alignmentfehler zu vermeiden. Tritt jetzt noch ein Fehler im Alignment auf, betrifft dieser nur 2 Streifen im Mosaik und nicht das komplette Panorama.

## 2.1 Scanning-Broom Kamera

Ein eindimensionales Array (Scanning-Broom oder Push-Broom Kamera) tastet die Bilder ab und extrahiert die Informationen für das Mosaik (Abb. ??). Es findet die Grenzen der Bildstreifen und die Streifen mit der optimalen Auflösung. Um diese Informationen überhaupt zu detektieren, muss der Scanning Broom immer quer (im  $90^\circ$  Winkel) zum optischen Fluss stehen (Abb. ??). Steht er parallel zum optischen Fluss, kann er nichts aufnehmen. Der Scanning-Broom kann über dem Bild beliebig rotiert und verschoben werden, um die geeigneten Streifen zu finden und richtig zum optischen Fluss zu stehen.

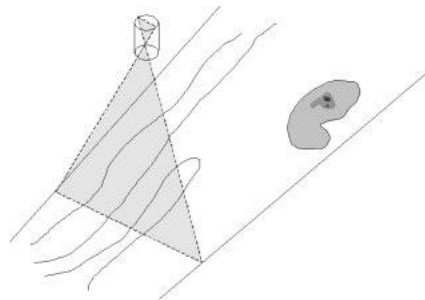


Abbildung 1: Bilderfassung einer Scanning-Broom Kamera

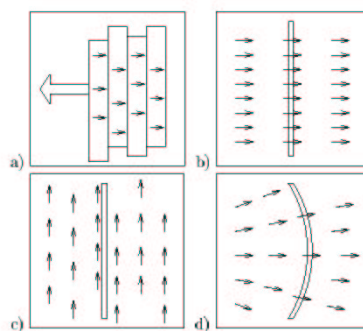


Abbildung 2: Die verschiedenen Formen eines Scanning-Brooms. In (c) wird nichts selektiert, da der Scanning-Broom nicht quer zum optischen Fluss steht.

## 2.2 Multiple Center of Projection (COP)

In den klassischen Verfahren werden alle Bilder an einem einzigen Referenzbild ausgerichtet und auf dessen Projektionsebene abgebildet (single COP). Dadurch kommt es bei Bildern, die weiter vom Referenzbild entfernt liegen, zu Schrumpfungen (Abb. ??). Dieses Schrumpfen kann nicht wieder rückgängig gemacht werden und hat so einen Rückgang der Auflösung zur Folge. Dadurch, daß alle Bilder nur an einem Referenzbild ausgerichtet werden, fallen Fehler im Alignment auch viel stärker ins Gewicht und es kann zu erheblichen Störungen des Panoramas kommen.

Bei den im Folgenden besprochenen Methoden wird multiple COP verwendet um die Bildqualität zu optimieren und Fehlern im Alignment vorzubeugen. Beim multiple COP werden die Bilder auf mehrer Projektionsebenen abgebildet. Dadurch wird Schrumpfen verhindert und die Auflösung im Panorama sinkt nicht unter die Auflösung in der original Bildsequenz.

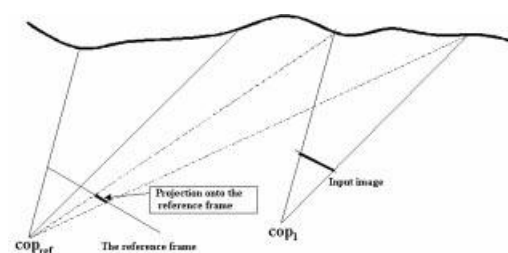


Abbildung 3: Multiple Center of Projection

### 3 Projektion auf Mannigfaltigkeiten

Die Projektion auf Mannigfaltigkeiten lässt beliebige Rotationen und Translationen der Kamera zu. Für das Alignment müssen in den Bildern nicht Überlappungen passend ausgerichtet werden, sondern es reicht, wenn die Grenzen der ausgewählten Streifen übereinstimmen.

Die Ausrichtung an einem Referenzbild, wie sie bei den klassischen Methoden verwendet wird, wird hier durch das paarweise Ausrichten zweier Bildsegmente abgelöst. Durch diese Technik wird das Alignment lokal und Fehler setzen sich nicht über das ganze Mosaik fort, sondern betreffen nur die Grenze, an der der Fehler aufgetreten ist.

Die Projektionsfläche ist allgemeiner gewählt als eine Ebene. Dadurch kann der Kamera mehr Bewegungsfreiheit zugestanden werden (Abb. ??).

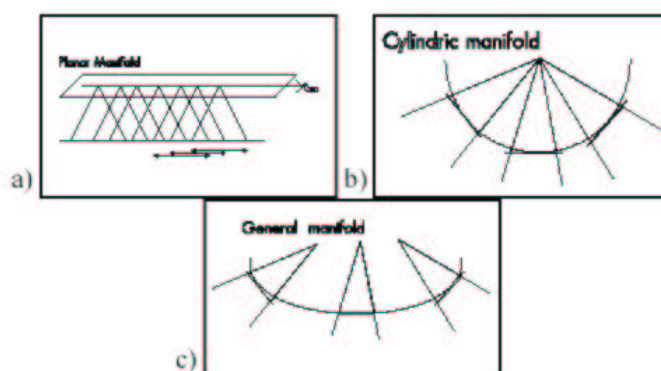


Abbildung 4: Die Projektion auf 3 Beispielmannigfaltigkeiten: (a) entspricht der klassischen Translation, (b) entspricht der klassischen Rotation und (c) der Projektion auf Mannigfaltigkeiten.

### 3.1 Alignment

Um die Bildausschnitte für das Mosaik zu finden werden die Bilder mit dem Scanning-Broom abgetastet. Beim Abtasten kann der Scanning-Broom beliebig rotiert und verschoben werden um immer quer zum optischen Fluss zu stehen.

Die selektierten Arrays ergeben zusammengesetzt immer ein brauchbares Mosaik. Das Zusammenfügen dieser Streifen stellt im Allgemeinen allerdings ein Problem dar. Um die eindimensionalen Arrays aneinander zu fügen, benötigt man genaue Informationen über die Bewegungen der Kamera. Dieses Verfahren kann zum Beispiel bei Satelliten, die ihre Aufnahmen über einen Aussenspiegel aufnehmen, dessen Bewegung genau bekannt ist, verwendet werden. Im Allgemeinen steht die Information über die Kamerabewegung allerdings nicht zur Verfügung. Und ohne diese Zusatzinformationen ist es unmöglich die Streifen korrekt zusammen zusetzen, da in den Streifen selbst zu wenig Information enthalten ist.

Im zweidimensionalen Fall, bei dem man die ganzen Bilder berücksichtigt,

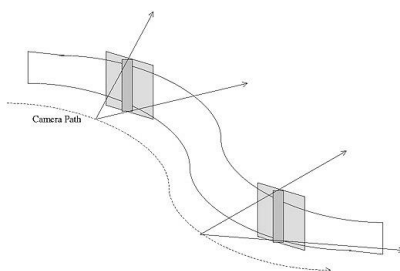


Abbildung 5: Ausrichten der Center-Strips unter Berücksichtigung der ganzen Bilder. So wird die Kamerabewegung berechnet wenn diese Information nicht vorliegt um die eindimensionalen Arrays zusammen zufügen

kann man die fehlenden Informationen berechnen. Deshalb führt man das Alignment mit sich überlappenden Bildern durch (Abb. ??). Mit diese Überlappungen kann man die Kamerabewegung berechnen. Man wählt Bildstreifen, die zwischen zwei Grenzen (eindimensionale Array-Streifen) liegen, um sie im Mosaik zu verwenden. Um die bestmögliche Bildqualität im Panorama zu erreichen, wählt man unter sich überlappenden Streifen immer den, der am nächsten zur Bildmitte des original Bildes liegt (center strip). In der Mitte jenes Bildes treten die wenigsten, durch interne Kameraparameter verursachten, Störungen auf. Ausserdem wählt man den Streifen, der die höchste Auflösung an der Stelle in der Szene hat.



### 3.2 Zusammenfügen der Streifen



Abbildung 6: Links ist das Alignment perfekt, die Pixelwerte werden gemittelt. Rechts ist das Alignment nicht so gut, nur Originalpixelwerte aus den Streifen werden benutzt

Das Alignment der Bilder ist perfekt, wenn breitere Streifen, es müssen nicht alle Bilder komplett sein, gut angepasst sind. In diesem Fall wird man um die Farbkanten an den Streifengrenzen zu entfernen, die Pixelwerte an sich überlappenden Stellen mitteln. Dadurch kommt es zu einem glatten Übergang an den Grenzen, Die Nähte zwischen den Bildstreifen sind nicht mehr so auffällig, da die Pixelwerte global (von allen sich überlappenden Regionen) und nicht lokal (nur von einem einzigen Bildstreifen) abhängig sind.

Ist das Alignment nicht so gut, können durch das Mitteln der Pixelwerte erhebliche Störungen auftreten. Es könnten Pixel die zu einem anderen Objekt gehören, in die Mittelwertberechnung einbezogen werden, die dann den Farbwert dieses Objektpixels verfälschen könnten. Deshalb wird in diesem Fall kein Mittelwert gebildet, sondern man nimmt den besten Streifen (aus dem Alignment-Schritt) und setzt diesen unverändert in das Mosaik ein. Die entstehenden Farbkanten können dann durch Farbwertinterpolation geglättet werden.

### 3.3 Beispiele



Abbildung 7: Panoramamosaik durch Projektion auf Mannigfaltigkeiten



Abbildung 8: Die ausgefransten Ränder entstehen durch die Bewegung der Kamera

Die Bilder entstanden durch das Mosaikverfahren mit Projektion auf Mannigfaltigkeiten. Die Streifen werden aneinandergereiht ohne auf die Mosaikränder zu achten. Deshalb folgt das Mosaik genau der Kamerabewegung. Das sieht man an den unregelmässigen schwarzen Rändern am oberen und unteren Bildrand.

## 4 Projektion auf allgemeine Mannigfaltigkeiten

Die Geometrie einer Szene entspricht im Allgemeinen nicht einer Ebene. Deshalb erfolgt bei dieser Methode die Projektion des Mosaiks auf eine, der Geometrie der Szene angepasste, Mannigfaltigkeit. So würde das Mosaik einer Tunneldurchfahrt zum Beispiel auf eine Mannigfaltigkeit, die der Form des Tunnels entspricht, projiziert. Durch das geschickte Wählen der Projektionsfläche kommt es zu visuell angenehmen Bildern mit natürlichem Aussehen. Eine Besonderheit der Projektion auf allgemeine Mannigfaltigkeiten ist das Anpassen des Mosaiks, so dass die Ober- und Unterkante des Bildes relativ gerade verlaufen. Drifts werden ausgeglichen. Das trägt auch sehr zur Verbesserung des visuellen Eindrucks bei.

Um den Algorithmus richtig anwenden zu können müssen 3 Annahmen erfüllt sein:

1. **In der Kamerabewegung gibt es keine abrupten Bewegungswechsel.**

Zum Beispiel darf es nicht vorkommen, daß erst gezoomt wird und dann eine Verschiebung entlang einer Ebene nach links folgt. Das würde einer Richtungsänderung von  $90^\circ$  entsprechen.

2. **Die internen Kameraparameter bleiben konstant.**

Das heißt, man fängt nicht mitten in einer anderen Bewegung an zu zoomen oder wechselt die Brennweite zwischen den Aufnahmen.

3. **Die Drifts der Kamerabewegung können berechnet und eliminiert werden.**

Die Rotation, die für den Drift verantwortlich ist, wird berechnet und rückgängig gemacht. Nur so kommt es zu visuell angenehmen Bildern.

### 4.1 Allgemeine Mannigfaltigkeiten

Mannigfaltigkeiten müssen nicht notwendigerweise Ebenen sein. Sie können sich jeder beliebigen Kurve (forming curve) anpassen (Abb. ??). Bei der Projektion auf allgemeine Mannigfaltigkeiten wird diese Tatsache zur Verbesserung des visuellen Eindrucks ausgenutzt. Die Mannigfaltigkeit, auf die am Ende das Mosaik projiziert wird, entspricht der Geometrie der Szene. Dadurch sieht das entstehende Panorama natürlich aus.

### 4.2 Projektion gerader Linien

Bei der perspektivischen Projektion, die einer Kameraaufnahme zugrunde liegt, kommt es vor, daß Linien, die in der realen Welt parallel verlaufen

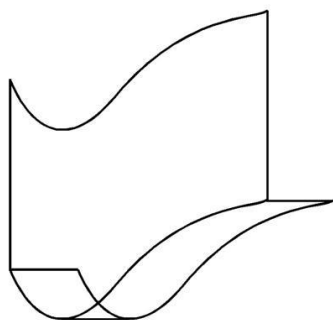


Abbildung 9: allgemeine Mannigfaltigkeit mit einer Forming-Curve in der Form eines  $L$

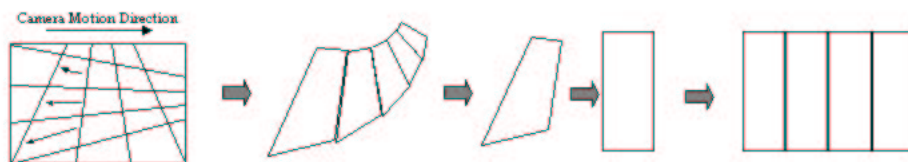


Abbildung 10: Um die Bildstreifen gut zusammenfügen zu können, muß der optische Fluß parallel sein. Das erreicht man indem man die Verzerrungen der perspektivischen Projektion, die bei der Kamera-Aufnahme entstehen, rückgängig macht. Im original parallele Linien müssen wieder gerade gebogen werden.

nach der Projektion nicht mehr parallel sind. Das heißt, der optische Fluß ist nicht parallel. Diese optische Verzerrung wird rückgängig gemacht und das Mosaik wird so zusammengesetzt, daß ein paralleler optischer Fluß entsteht (Abb. ??).

### 4.3 Alignment der Bildstreifen

Wie auch im vorigen Kapitel werden die Bildstreifen paarweise aneinander ausgerichtet (Abb. ??). Dabei bleibt eine Grenze  $B_i$  des Streifens immer unverändert. Durch diese Eigenschaft des Verfahrens ist gewährleistet, daß Fehler im Alignment sich nicht fortpflanzen. Nur die linke Grenze (oder Vorderseite) des Streifens aus dem Bild  $I_k$  wird an die rechte Grenze (oder auch Rückseite)  $B_{k-1}$  des Streifens  $R_{k-1}$  angepasst.

Die Breite der Streifen  $R_i$  hängt von der Geschwindigkeit der Kamerabewegung ab. Je höher die Geschwindigkeit, desto breiter der Streifen. Die Ge-

schwindigkeit für  $R_{k-1}$  lässt sich durch den Vergleich der Position der Kante  $B_{k-1}$  in den Bildern  $I_{k-1}$  und  $I_k$  berechnen.

Die obere und die untere Kante jedes Streifens wird durch eine vertikale Spalte im Bild gelegt, um Drifts des Panoramas zu vermeiden.

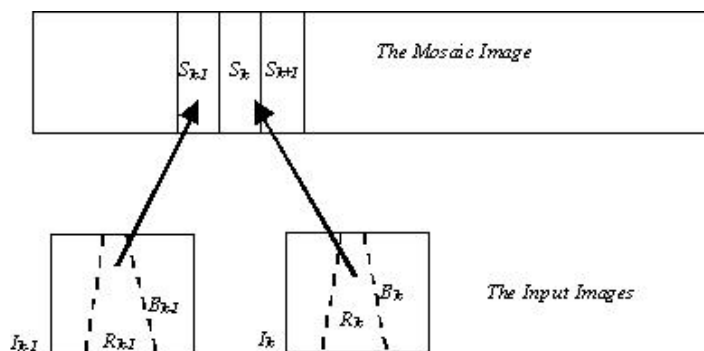


Abbildung 11: An jedem Streifen wird im Alignmentschritt nur eine Kante angepasst. So setzen sich Fehler im Alignment nicht über das ganze Panorama fort.

#### 4.4 Beispiel

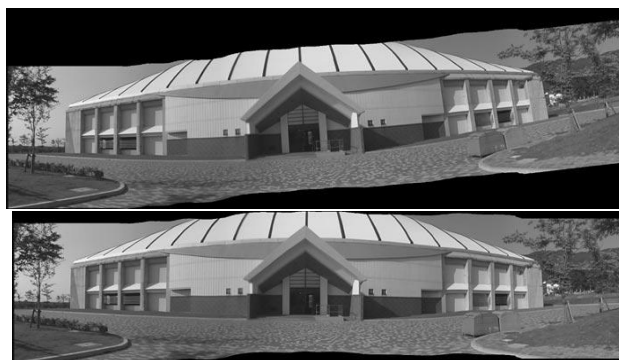


Abbildung 12: Das obere Bild entstand durch Projektion auf Mannigfaltigkeiten, das unter durch die Projektion auf allgemeine Mannigfaltigkeiten. Man sieht deutlich, daß der Drift im unteren Bild nicht mehr auftritt.

## 5 Röhrenprojektion

Einer der Nachteile bei der Projektion auf Mannigfaltigkeiten ist, daß Zoom und Vorwärtsbewegungen nicht dargestellt werden können. Das liegt daran, daß radialer, optischer Fluss nicht mit einem geraden Scanning-Broom detektiert werden kann.

In der Röhrenprojektion kann sich der Scanning-Broom (Kap. ??) beliebig verformen und so auch radialen, optischen Fluss erfassen.

Die Projektion erfolgt auf eine Mannigfaltigkeit in der Form einer Röhre, deren Achse mit der Blickrichtung der Kamera übereinstimmt. Durch die Form der Mannigfaltigkeit ist eine gute Verarbeitung radial detektierter Streifen gegeben.

### 5.1 Auswahl der Bildstreifen

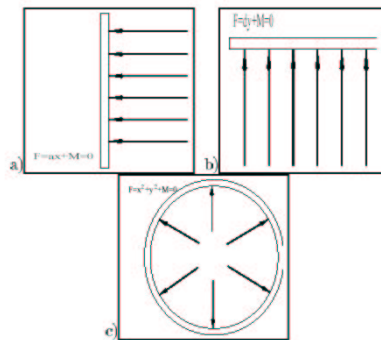


Abbildung 13: Die Form der Streifen hängt nur von der Bewegung im Bild, dem optischen Fluss ab.

Die Form des Bildstreifens hängt von der Bewegung im Bild, dem optischen Fluss, ab. Diese Bewegung berechnet sich wie folgt:

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} x_n - x_{n-1} \\ y_n - y_{n-1} \end{pmatrix} = \begin{pmatrix} a + bx_n + cy_n \\ d + ex_n + fy_n \end{pmatrix} \quad (1)$$

Hierbei sind  $P_n = (x_{n-1}, y_{n-1})$  und  $P_{n-1} = (x_n, y_n)$  korrespondierende Punkte in den Bildern  $I_{n-1}$  und  $I_n$ . Der optische Fluss wird durch den Vektor  $(u, v)$  als Funktion der Pixelposition  $(x_n, y_n)$  dargestellt. Gesucht ist nun  $\mathcal{F}(x, y) = 0$  (Gl. ??), eine Linie, die im rechten Winkel zum optischen Fluss steht. Die Normale dieser Linie hat die selbe Richtung wie  $(u, v)$ . Wenn  $c = e$

gilt und wir integrieren, erhalten wir die Gleichung des Scanning-Brooms (Abb. ??):

$$0 = \mathcal{F}(x, y) = ax + dy + \frac{b}{2}x^2 + \frac{c+e}{2}xy + \frac{f}{2}y^2 + M \quad (2)$$

$M$  ist hier das Entscheidungsmerkmal, das die Streifen bei mehreren Möglichkeiten selektiert.  $M$  sollte auf die maximale Anzahl der Pixel auf dieser Linie im Bild gesetzt werden. So wählt man den Streifen mit der besten Auflösung. Bringt auch das keine eindeutige Entscheidung, wählt man den Streifen, der näher an der Bildmitte liegt, da dadurch Störungen, die durch die Kameralinse verursacht wurden, minimiert werden.

Die ausgewählte Linie ist bei der Röhrenprojektion der Schnitt zwischen Röhre und Bildebene.  $M$  ist äquivalent zur Wahl des Radius  $R$  der Röhre.

Bei Zoom oder Vorwärtsbewegung geht der optische Fluss radial vom Fokus nach aussen. Diese Bewegung kann von einem normalen Scanning-Broom nicht detektiert werden, da es keine Gerade gibt, die quer zu Punkten auf einer Kurve steht. Um dies Informationen doch zu erkennen, ist es in der Röhrenprojektion möglich, den Scanning-Broom zu krümmen, so daß er senkrecht auf dem optischen Fluss steht. (Abb. ??)

Ein Streifen, der später im Mosaik verarbeitet wird, wird immer von zwei detektierten Kanten,  $\mathcal{F}_n(x_n, y_n) = 0$  und  $\mathcal{F}'_{n+1}(x_n, y_n) = 0$  begrenzt. Dabei entspricht die Kante  $\mathcal{F}'_{n+1}(x_n, y_n) = 0$  in Bild  $I_n$  der Kante  $\mathcal{F}_{n+1}(x_{n+1}, y_{n+1}) = 0$  aus dem Bild  $I_{n+1}$  und die Kante  $\mathcal{F}_n(x_n, y_n) = 0$  aus  $I_n$  korrespondiert mit  $\mathcal{F}'_{n-1}(x_{n-1}, y_{n-1}) = 0$  aus  $I_{n-1}$ . Durch diese Wahl der Streifen wird sichergestellt, daß keine Informationen doppelt verarbeitet werden oder gar verloren gehen.

Die Breite der Streifen hängt allein von der Geschwindigkeit der Bewegung der Kamera von einem Bild zum nächsten ab. Je schneller die Bewegung, desto breiter der Streifen.

## 5.2 Kombination der Streifen

Der radiale, optische Fluss muss nun noch so transformiert werden, daß er parallel wird. Nur so kann man die Streifen richtig zusammenfügen, da es sonst durch die perspektivische Projektion in den Bildern zu Störungen im zusammengesetzten Mosaik käme. Das erreicht man, indem man das zweidimensionale Bild auf einen dreidimensionalen Zylinder, die Röhre, projiziert. Die Achse der Röhre  $\hat{s}$  entspricht der Blickrichtung der Kamera. Sie geht durch das optische Zentrum der Kamera  $O = (0, 0, 0)$  und durch den Zoomfokus  $S = (s_x, s_y, f_c)$  wobei  $f_c = S/|S|$  gilt,  $f_c$  also die Brennweite der Kamera ist. Jeder Punkt  $P = (x, y, f_c)$  im Bild wird auf einen zugehörigen Punkt  $Q$

auf der Röhre projiziert. Die Entfernung von  $Q$  zur Achse  $\hat{s}$  beträgt genau den Radius der Röhre  $R$ .

Nach der Projektion auf die Röhre sind alle Pixel  $Q$  parallel zur Achse  $\hat{s}$  ausgerichtet. Somit hat man einen parallelen, optischen Fluss.

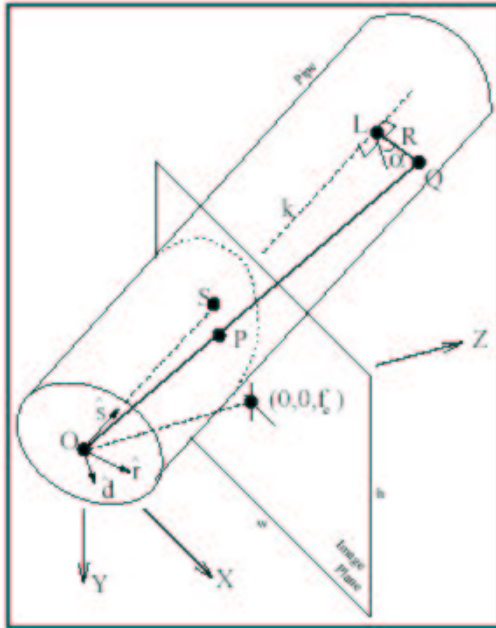


Abbildung 14: Die Röhrenprojektion

Um die Streifen richtig zusammenzufügen, müssen sie nur auf der Röhre an die richtige Stelle gerückt werden. Diese Projektion und das Zusammenfügen durch Verschiebung sind verlustfrei. Wodurch eine optimale Bildauflösung gewährleistet wird.

Die Position von  $Q$  auf der Röhre wird durch  $k$  und  $\alpha$  bestimmt. Dabei ist  $k$  die zugehörige Position auf der Achse  $\hat{s}$  und  $\alpha$  ist der Winkel des Punktes zu  $\hat{d}$ .  $\hat{d}$ ,  $\hat{r}$  und  $\hat{s}$  spannen dabei das Koordinatensystem der Röhre auf.

Die Position von  $Q$  auf der Röhre berechnet sich wie folgt:

$$Q = (Q_x, Q_y, Q_z) = k\hat{s} + R \cos(\alpha)\hat{d} + R \sin(\alpha)\hat{r} \quad (3)$$

Der zugehörige Bildpunkt genügt der Gleichung:

$$P = (x, y, f_c) = \left( \frac{f_c Q_x}{Q_z}, \frac{f_c Q_y}{Q_z}, f_c \right) \quad (4)$$



Pixel, deren Abstand von der Achse  $\hat{s}$  im Bild kleiner als  $R$  war, werden auf der Röhre vergrössert abgebildet. Bei der Rückprojektion auf eine Ebene wird diese Skalierung allerdings verlustfrei rückgängig gemacht. Ist der Abstand zu  $\hat{s}$  aber grösser als  $R$ , schrumpfen die Objekte bei der Projektion auf die Röhre. Das Schrumpfen ist nicht mehr umkehrbar und führt so zu einer Verminderung der Bildauflösung und somit der Bildqualität. Um diesen Effekt zu umgehen, wählt man  $R$  unter Berücksichtigung der Höhe  $h$  und der Breite  $w$  des Originalbildes:

$$R = \sqrt{f_c^2 + \left(\frac{w}{2}\right)^2 + \left(\frac{h}{2}\right)^2} \quad (5)$$

Um die bestmögliche Auflösung im Mosaik zu erhalten, wählt man die Streifen aus der Menge der sich überlappenden Regionen aus, die lokal die beste Auflösung haben. Dafür nutzt man  $M$  (Gl. ??). Die beste Auflösung ist immer an der Schnittstelle von Röhre und Bild ( $Q_z = f$ ). Je grösser der Abstand zur Schnittstelle, desto schlechter die Auflösung. Geht man noch dazu nach aussen, steigt die Warscheinlichkeit von Verzerrungen durch interne Kameraparameter.

Um die Fortsetzung von Alignmentfehlern zu verhindern, wird auch hier beim Zusammenfügen nur eine Kante der Streifen verändert und die andere,  $\mathcal{F}_n$ , bleibt original, senkrecht zum optischen Fluss.

Bei dem eigentlichen Cut'n'Paste-Vorgang müssen die Streifen gerade sein und nicht gebogen wie sie im Alignmentschritt waren. Deshalb werden sie zum Zusammensetzen geradegebogen. Das führt zu erheblichen Verzerrungen im Panorama. Deshalb wird das Mosaik vor der visuellen Ausgabe wieder gekrümmt, wie es im Alignmentschritt war. Durch das Zurückbiegen des Mosaiks kommt es zu gekrümmten Rändern (rechts und links) des Panoramas.

### 5.3 Beispiel



Abbildung 15: Oben sind 2 der Ausgangsbilder um das Panorama unten zu schaffen. Man beachte, daß im Panorama der obere und der untere Rand Himmel darstellt. Da trifft sich das Bild in der Röhre.

Wenn man die 2 oberen Bilder vergleicht bemerkt man, daß das untere der beiden ein Ausschnitt des oberen ist (Die Bäume unten rechts sind im oberen Bildhintergrund). Das sind 2 Aufnahmen, die in eine Röhre, deren Achse parallel zur Strasse verläuft, projiziert werden. Das untere Bild ist das Ergebnis der Röhrenprojektion. Der obere und der untere Bildrand stellt Himmel da. Hier trifft sich die die Abbildung in der Röhre. Durch die Projektion auf die Röhre wird der optische Fluss parallel. Das heißt die perspektivische Projektion, wird herausgerechnet. Dadurch

kann man den Blickpunkt ändern wie es hier der Fall ist. Nun blickt man nicht mehr die Strasse hinunter, sondern seitlich auf die Strasse.

Die gebogenen Ränder am linken Bildrand des Panoramas kommen durch das Biegen der Bildstreifen im Cut'n'Paste-Schritt. Beim zurück biegen entstehen diese Kanten.

## 6 Parallaxe

Durch grosse Tiefenunterschiede in einer Szene kommt es zu Parallaxe-Effekten. Parallaxe nennt man die relative, optische Verschiebung von Objekten gegeneinander durch die Bewegung des Betrachters.

Es gibt keine Transformation, die den optischen Fluss der ganzen Szene richtig darstellen könnte. Man kann entweder die Objekte im Hintergrund richtig zusammenfügen und staucht somit die Objekte im Vordergrund, oder man passt den Vordergrund richtig zusammen und hat bei den Objekten im Hintergrund Verdopplungseffekte (Objektteile kommen in verschiedenen Bildern an verschiedenen, relativen Positionen vor). Diese Phänomene treten vor allem bei Streifen auf, die breiter als ein Pixel sind.

Um die Parallaxe-Störungen auszugleichen, benötigt man also von den kritischen Stellen ein Pixel breite Streifen. Dann ist die Positionsänderung des Betrachters weniger abrupt und die daraus resultierende Verschiebung nicht so schwerwiegend. Um diese  $N$  Ein-Pixel-Streifen zu bekommen, die den ursprünglichen  $N$ -Pixel-Streifen ersetzen sollen, erzeugt man mittels Bildinterpolation  $N$  Zwischenbilder, aus denen man dann die passenden Ein-Pixel-Streifen herausnimmt.

## 6.1 Beispiel

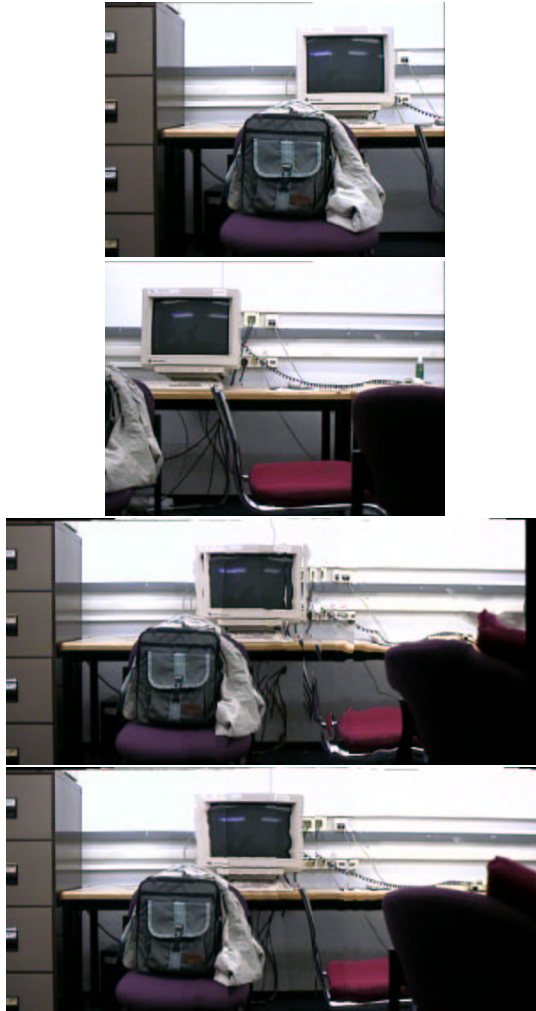


Abbildung 16: Oben sind 2 der Ausgangsbilder um die Panoramas unten zu schaffen. Das erste Panorama weist starke Parallaxe-Effekte auf. Im zweiten Panorama wurden diese, durch Streifen aus Zwischenbildern, ausgeglichen

Aus den oberen beiden kleinen Bildern werden die unteren beiden Mosaik generiert. Man sieht bei dem oberen Mosaik die Verdopplungseffekte am Computerbildschirm und den daneben liegenden Steckdosen besonders deutlich.

Im darunter liegenden, korrigierten Mosaik ist die Störung nicht mehr so schwerwiegend. Hier wurden Streifen aus generierten Zwischenbildern ein-

gesetzt. Dadurch ist die Parallaxebewegung (Verschiebung der Gegenstände gegeneinander) nicht so abrupt. Es kommt zu einem glatteren Übergang.

## 7 Vergleich der Algorithmen

Die Lokalität des Alignments ist ein grosser Vorteil der Projektion auf Mannigfaltigkeiten gegenüber den klassischen Mosaik-Algorithmen. Dadurch, daß die Streifen immer nur paarweise aneinander angepasst werden und an jedem Streifen nur eine Kante verändert wird, können sich Fehler im Alignment nicht über das ganze Mosaik fortsetzen. Sie betreffen immer nur lokal die zwei Streifen, bei deren Alignment der Fehler aufgetreten ist.

Durch die Verwendung des Multiple-COP-Ansatzes wird die Auflösung der Originalbilder im Mosaik erhalten.

Die einfache Projektion auf Mannigfaltigkeiten schlägt allerdings fehl, wenn in der Szene Zoom oder Vorwärtsbewegung vorkommen. Dafür ist sie für Szenen ohne Zoom beziehungsweise Vorwärtsbewegung sehr effizient zu implementieren, so daß sie, ohne Qualitätsverbesserungen, wie Farbabgleiche und ähnliches, in Echtzeit laufen kann.

Der grosse Vorteil der Röhrenprojektion ist die Handhabung von Zoom und Vorwärtsbewegung durch das Biegen des Scanning-Brooms. Ein Nachteil sind jedoch die verbogenen Ränder bei der Ausgabe, die sehr unnatürlich aussehen, da die Geometrie einer Szene im Allgemeinen nicht einer Röhre entspricht.

In der Projektion auf allgemeine Mannigfaltigkeiten wird die Projektionsfläche der Geometrie der Szene angepasst. Dadurch sehen die so entstandenen Panoramabilder sehr natürlich aus.

Auch Drifts in der Kamerabewegung werden ausgeglichen, was dem Bild ein sehr angenehmes Aussehen verleiht.

Mit Hilfe von Bildinterpolation können Zwischenbilder erstellt werden, die bei Parallaxe-Effekten zur Verbesserung der Bildqualität eingesetzt werden.