Affine und perspektivische Kamerabewegungen

Samir Chaturvedi

SEMINAR
Algorithmen zur Erzeugung von Panoramabildern

eingereicht am
Lehrstuhl Praktische Informatik IV

Prof. Dr. Wolfgang Effelsberg

Mathematik und Informatik

Universität Mannheim

im Mai 2008

Inhaltsverzeichnis

Kapitel 1

Einleitung

Wie setzt man Bilder zusammen, so dass man daraus ein Panoramabild erzeugen kann. Dazu muss die Perspektive der vorhandenen Bilder verändert werden und, falls beim Photographieren kein Stativ verwendet wurde, müssen die Bilder rotiert bzw. nach oben oder unten verschoben werden. Dies alles erreicht man durch den Einsatz von affinen bzw. projektiven Modellen. Im Laufe dieser Arbeit werden wir diese Modelle näher kennen lernen und erfahren, welches Modell uns welche Dienste zur Verfügung stellt.

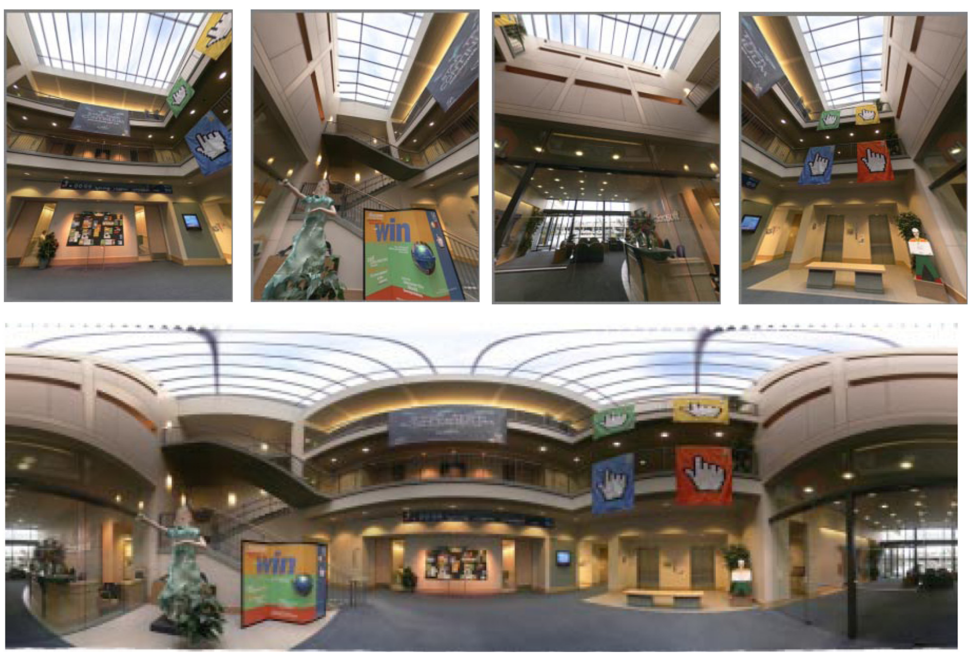


Abbildung 1.1: Panorama Erzeugung hat Einsatz eines projektiven Modells [6]

Kapitel 2

Projektive Geometrie

2.1 Euklidischer Raum

Elementargeometrie wird normalerweise im euklischen Raum beschrieben, der eine direkte Beschreibung des Raumes ist, den wir mit unserer menschlichen Intuition empfinden. Hier werden die Punkte im n-dimensionalen euklidischen Raum $\mathbb{E}^n$ einfach als Vektoren der Länge n dargestellt. Jedoch hat diese einfache Definition des euklidischen Raumes auch einige Nachteile. Erstens gibt es keine Auffassung über die Punkte in unendlich, so dass dies als ein Spezialfall behandelt werden muss. Dieses Problem entsteht z.B., wenn es den Durchschnitt der parallelen Linie berechnet, der nicht im euklischen Raum definiert ist. Der zweite Nachteil kommt zustande, wenn wir eine geometrische Transformation verwenden um eine Gegenstandsbewegung zu beschreiben. Sogar für grundlegende Arten der Bewegung werden verschiedene Formeln benötigt. Verschiebung wird unter Verwendung einer Vektoraddition beschrieben, Rotation als eine Matrixmultiplikation, und eine perspektivische Projektion der Punkte auf eine Ebene benötigt eine Division.

2.2 Projektiver Raum

Ein projektiver Raum besteht aus Geraden und Punkten sowie zwei Relationen. Die erste Relation gibt an, welche zwei Geraden einen eindeutigen Schnittpunkt haben und die zweite Relation gibt an, welche Punkte eine Verbinungsgerade besitzen. Dabei ist zu beachten, dass die Geraden alle Ursprungsgeraden sind. Das Konzept der projektiven Ebene liefert eine alternative einfache Formel, um alle Arten von steifen Bewegungen und perspektivischen Projektionen in einer einheitlichen Formel zu beschreiben. Außerdem sind die Punkte in unendlich ein wesentlicher Bestandteil des Projektiven Raumes und müssen nicht speziell betrachtet werden.

2.3 Homogene Koordinaten

Bei einer Transformation werden meistens mehrere Operationen hintereinander durchgeführt. Daraus ergibt sich meisten nicht nur Matrizienmultiplikation sondern auch Matrizenaddition. Man verwendet homogene Koordinaten um aus aufeinander folgenden Matrizenmultiplikationen- und Additionen eine einizge Matrix zu ermitteln. Das bei einer Transformation nicht nur ein Punkt, sondern mehrere Punkte transformiert werden, kann man dadurch vermeiden, dass die Transformationsmatrix mehrmals berechnet werden muss. In n-dimensionalen Euklidischenraum $\mathbb{E}^n$ wird jeder Punkt als Vektor der Länge n beschrieben, in der jeder Komponent eine Position entlang der Koordinatenachse ist. Demgegenüber werden die Punkte im n-dimensionalen projektiven Raum $\mathbb{P}^n$ durch $n+1$-dimensionalen Vektoren dargestellt. Der Aufbau ist so, dass jeder Punkt $(x_1, ..., x_n)^\top$ in $\mathbb{E}^n$ einem eindimensionalen Unterraum $(wx_1, ..., wx_n, w)^\top$ in $\mathbb{P}^n$ mit dem freien Skalierungsparameter $w \neq 0$ entspricht. Der euklidische Raum $\mathbb{E}^n$ kann in $\mathbb{P}^n$ auf eine einfache Art eingebunden werden, indem man eine kanonische injektion $\mathbb{E}^n \ni (x_1, ..., x_n)^\top \mapsto (x_1, ..., x_n, 1)^\top \in \mathbb{P}^n$ verwendet. In der Rückwertsrichtung können euklidische Koordinaten durch diese Funktion zugewonnen werden

$$\mathbb{P}^n \ni (x_1, ..., x_n, w)^\top \mapsto (\frac{x_1}{w}, ..., \frac{x_n}{w})^\top \in \mathbb{E}^n \tag{2.1}$$

Diese projektive Darstellung wird wie als homogene Koordinaten des Punktes bezeichnet. Eine direkte Folgerung aus der Definition ist die wichtigste Eigenschaft, damit die homogene Koordinaten bei der Skalierung unverändert bleiben. So stellen $(x_1, ..., x_n, w)^\top$ und $(\lambda x_1, ..., \lambda x_n, \lambda w)^\top$ den gleichen Punkt für alle $\lambda \neq 0$ dar. D.h., jeder euklidische Punkt p wird durch eine Äquivalenzklasse $E(p) \subset \mathbb{P}^n$ im projektiven Raum dargestellt. Alle Vektoren in der Äquivalenzklasse können durch eine konstante Skalierung der Koordinaten erhalten werden. Projektive Punkte mit $w = 0$ stellen *ideale* Punkte in unendlich mit keine Übereinstimmung in Euklidische Raum dar. Für diese Punkte in unendlich zeigen die ersten n Komponenten der Vektor, die Richtung an, in der der Punkt liegt. Der Nullvektor $(0, ..., 0)^\top$ ist nicht in $\mathbb{P}^n$ eingeschlossen. Den Einsatz von homogenen Koordinaten werden wir im Abschnitt der affinen Modelle sehen. Hier kann die Relation zwischen Punkten im euklidischen Raum $\mathbb{E}^2$ und entsprechenden Punkten im projektiven Raum $\mathbb{P}^2$ leicht sichtbar gemacht werden. Wir zeichnen $\mathbb{P}^2 = \{(x, y, w)^\top\}$ als dreidimensionalen Raum mit den Parametern x, y, w. In diesem Raum kann die Äquivalenzklasse, die den Punkt p im $\mathbb{P}^2$ enthält, als Linie, die durch den Ursprung und den Punkt p geht, dargestellt werden. Der Durchschnitt dieser Linie mit der Ebene $w = 1$ definiert die Koordinaten in der euklidischen Ebene. Da eine Skalierung des homogenen Koordinaten Punktes p durch eine Nicht-Null Konstante nur den Punkt entlang der Linie verschiebt, bleibt seine Position auf der euklidischen Ebene gleich. Dieses erklärt die Skalierungsinvarianz der homogenen Koordinaten. Indem man die spezielle Darstellung mit der letzten Koordinate $w = 1$ wählt, wird klar, dass die euklidische Ebene selbst in dem projektiven Raum als eine Ebene an $w = 1$ eingebettet werden kann. Ideale Punkte p mit $w = 0$ entsprechen den Linien, die zur euklidischen Ebene parallel sind verlaufen, und sich idealerweise im unendlich schneiden. Da diese Linien zur euklidischen Ebene $w = 1$ parallel sind verlaufen, sind die idealen Punkte nicht Teil des euklidischen Raumes. Dies stimmt mit der Definition des euklidischen Raumes überein, da es keine Informationen über die Punkte in unendliche enthält. Um während der Berechnung eine Maßeinheit zu haben $\|p\| = 1$, versucht man homogene Koordinaten zu normieren. Bzgl. der oben genannten Visualisierungsmethode hat diese den Effekt, dass alle Punkte in den homogenen Koordinaten um den Ursprung von $\mathbb{P}^2$ auf dem einheitlichen Bereich liegen. Bei Koordinaten mit erheblich unterschiedliche Größen können Probleme auftreten. Wichtig ist zu wissen, dass die Normierungsverfahren an unterschiedlichen Zeitpunkten in der Berechnung angewendet werden können.

In der euklidischen Geometrie werden Linien als $ax+by+c = 0$ definiert, wobei a, b, c die Parameter der Linie sind. Diese Parameter bleiben unverändert. Auch bei der Multiplikation mit einem konstanten Wert wie $w \neq 0$, der $w(ax) + w(by) + wc = 0$ die gleiche Linie beschreibt. Wenn wir die Linienparameter mit Hilfe des Vektors $l = (a, b, c)^\top$ bezeichnen und die Punkte zur Spezifizierung von homogenen Koordinaten $p = (wx, wy, w)^\top$ verwenden, können wir die gleiche Linie als $l^\top.p = 0$ beschreiben. Diese Konstruktion stellt jede Linie in $\mathbb{E}^2$ als eine Ebene in $\mathbb{P}^2$ dar, da sie den Ursprung durchläuft und einen Normalvektor hat. Diese Ebene schneidet die euklidische Ebene $w = 1$ in der gewünschten Linie. Da die Norm der Ebene normalerweise irrelevant ist, bleiben die Linienparameter auch zur Skalierung mit einem Nicht-Null Wert unverändert. Ähnlich zu den Punkten an unendlich, ist eine spezielle Linie durch $l_\infty = (0, 0, 1)^\top$ in unendlich definiert. Solange $l_\infty^\top.(x, y, w)^\top = 0$ genau dann ist, wenn $w = 0$ ist, ist die Linie l_∞ eine Reihe von Punkten in unendlich. Wichtig ist zu wissen, dass die Linie l_∞ ,wie die Punkte in unendlich, keine Übereinstimmung mit euklidischer Geometrie hat. In unserer Visualisierung, in der wir den Vektor von Linienparametern betrachten, definiert l_∞ eine Ebene, die parallel zur euklidischen Ebene ist, und sich idealerweise in unendlich schneidet. Linien l_1 und l_2 beschreiben zwei Linien in $\mathbb{P}^2$. Im Raum (x, y, w) können zwei Ebenen durch den Ursprung mit Normalvektoren l_1 und l_2 dargestellt werden. Der Schnittpunkt p von zwei Linien l_1,l_2 muss auf beiden Ebenen im Raum (x, y, w) liegen. Das bedeutet, dass $p^\top.l_1 = 0$ und $p^\top.l_2 = 0$ gelten muss. Mit anderen Worten, p ist orthogonal zu beide Linien l_1 und l_2. Wir können dies wie folgt berechnen: $p = l_1 \times l_2$. Wenn er homogene Koordinaten verwendet, ist der Durchschnitt von zwei Linien wohl definiert und ergibt einen idealen Punkt in unendlich. In diesem Fall hat der Schnittpunkt $w = 0$. Die anderen zwei Koordinaten zeigen die Richtung an, den der Schnittpunkt in unendlich Punkte p_1 und p_2 beschreiben zwei Punkte in $\mathbb{P}^2$. Daraus folgt ein ähnliches Ergebnis wie oben. Um projektive Parameter von der Linie zwischen p_1 und p_2 zu finden, müssen wir eine Ebene finden, in dem die Punkte p_1 und p_2 sowie der Ursprung enthalten sind. Wir spezifizieren diese Ebene durch seine normale Ebene l, d.h., dass $l^\top.p_1 = 0$ als auch $l^\top.p_2 = 0$ gelten müssen. Daraus können wir l durch die Verwendung von Kreuzprodukten ermitteln $l = p_1 \times p_2$. Die Linie zwischen zwei Punkten in unendlich ergibt die Linie l_∞.

2.4 Geometrische Transformation in 2-D

Die Projektive Transformation wird häufig verwendet, um Bewegungen eines Gegenstandes als geometrische Transformation in Bilder zu beschreiben. Wir zeigen zuerst einige grundlegende Eigenschaften von projektiven Transformationen und stellen die affine Bewegung als eine wichtige Unterklasse der projektiven Transformation vor. Die Klasse der affinen Bewegung erlaubt uns, ein intuitives Verständnis der physikalischen Bedeutung der Tansformationsparameter zu entwickeln. Schließlich sehen wir, wie das projektive Bewegungsmodell, das für die Bewegungsanalyse in einer inhomogenen Formalisierung verwendet wird, direkt von der homogenen Definition abgeleitet werden kann.

2.4.1 Projektive Transformation

Eine projektive Transformation wird als lineare Transformation zwischen homogenen Koordinaten definiert. Wir betrachten nur nicht degenerierte Fälle in denen die Transformation umkehrbar ist. Da die Transformation linear und umkehrbar ist, kann sie als eine nichtsinguläre Matrix $H = \{h_{jk}\}$ geschrieben werden. Für einen projektiven Raum $\mathbb{P}^n$ hat eine Matrix die Größe $(n+1) \times (n+1)$. Im Falle von $\mathbb{P}^2$ erhalten wir folgendes:

$$\begin{pmatrix} x' \\ y' \\ w' \end{pmatrix} = \begin{bmatrix} h_{00} & h_{01} & h_{02} \\ h_{10} & h_{11} & h_{12} \\ h_{20} & h_{21} & h_{22} \end{bmatrix} \begin{pmatrix} x \\ y \\ w \end{pmatrix} \tag{2.2}$$

Die Einträge der Matrix $\{h_{jk}\}$ sind die Transformationsparameter. Bei der Verwendung von homogenen Koordinaten wird die Transformationsmatrix nur bis zur einem Skalierungsfaktor definiert. Folglich; in $\mathbb{P}^2$ hat die Transformation nur 8 Parameter obwohl die Transformationsmatrix 9 Elemente hat. Eine Kollineation in der Ebene wird als eine Transformation dieser Abbildungslinien auf Linien definiert. Es kann gezeigt werden, dass jede Kollineation als projektive Umwandlung geschrieben werden kann, und umgekehrt. Für eine nichtsingulär gegebene Transformationsmatrix $H = \{h_{kl}\}$ können wir leicht zeigen, dass die Linien immer auf Linien abgebildet werden. Um dies zu prüfen, betrachten wir $\{p_i\}$ eine Reihe von Punkten, die alle auf der Linie l liegen, so dass $l^\top.p_i = 0$ gilt. Da wir annehmen, dass H umkehrbar ist, entspricht dieses $l^\top(H^{-1}H)p_i = 0$. Beim umklammern können wir diese Gleichung ermitteln $(l^\top H^{-1})(Hp_i) = 0$. Aber dies bedeutet, dass die umgewandelten Punkte Hp_i alle auf der Linie $l^\top H^{-1}$ liegen. So erhalten projektive Transformationen die Linien. Die Umkehrrichtung gibt an, dass jede Kollineation als projektive Transformation geschrieben wird. Für eine Transformation $p' = Hp$ zwischen Punkten, können wir eine entsprechende Transformation finden, die die Linienparameter des Vektors l zu Linienparametern l' der transformierten Linie abbildet. Es kann folgendes ermittelt werden:

$$l'^\top = l^\top H^{-1} \quad \textit{oder entsprechend} \quad l' = H^{-\top} l \tag{2.3}$$

Entsprechend können wir sagen, dass für die Punkttransformation H, die entsprechenden Linientransformationen $H^{-\top}$ ist.

2.4.2 Affine Modelle

Wir behandeln jetzt 2-dimensionale affine Transformationen. Dies sind Abbildungen, die ein Objekt in ein anderes Objekt überführen. Die wichtigsten Transformationen sind:

- Translation
- Rotation
- Skalierung
- Scherung

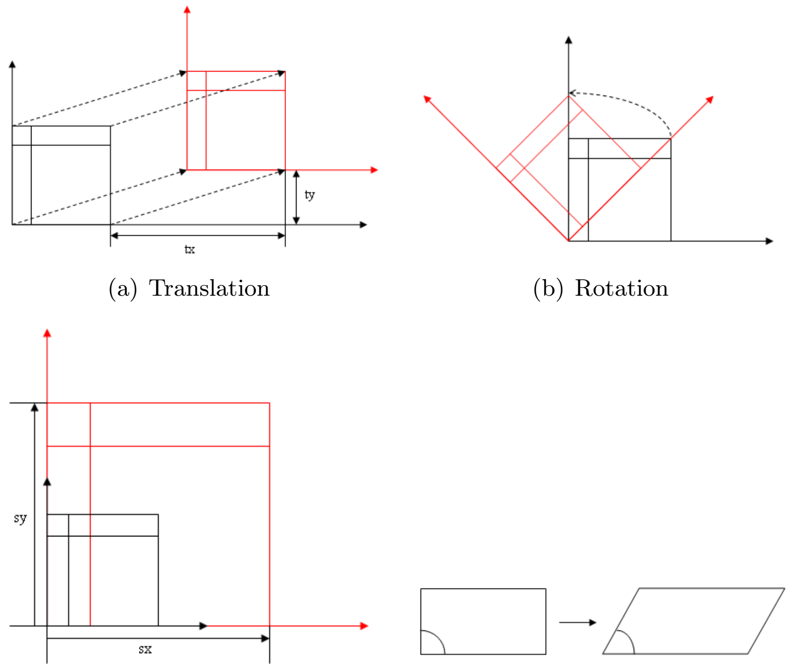
Wir legen aber vorher fest, was man unter einer Transformation versteht. Eine Transformation T ist eine Abbildung $v \mapsto T(v) = v'$. Eine Transformation T wird linear genannt, falls für Vektor v und w gilt:

$$T(v + w) = T(v) + T(w)$$
$$T(\mu v) = \mu T(v)$$

Von der Elementargeometrie wissen wir, dass die affine Bewegung in $\mathbb{E}^2$ geschrieben werden als

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix}, \tag{2.4}$$

so dass er eine 2×2 Transformationsmatrix $A = \{a_{kl}\}$ und einen Verschiebungsvektor $t = (t_x, t_y)^{\mathrm{T}}$ enthält. Die Matrix A kann immer in einer Reihe von elementaren Operationen, Rotationen, Skalierungen und Reflexionen faktorisiert werden.



(a) Translation (b) Rotation

(c) Skalierung (d) Scherung

Abbildung 2.1: Transformationen[1]

Die Transformationsmatrizen für jede dieser grundlegenden Operation werden unten definiert. Es ist leicht zu sehen, dass all diese elementaren Operationen, parallele Linien auf parallelen Linien abbilden. Infolgedessen bietet diese Eigenschaft die Möglichkeit, dass die affine Transformation in eine Reihe von elementaren Operationen zerlegt werden kann. Unter Verwendung von homogenen Koordinaten, können wir den Verschiebungsvektor in die Matrix integrieren, um daraus folgende Formel zu ermitteln [2]

$$\begin{pmatrix} x' \\ y' \\ w' \end{pmatrix} = \begin{bmatrix} a_{00} & a_{01} & t_x \\ a_{10} & a_{11} & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ w \end{pmatrix} \tag{2.5}$$

Translation

$$x' = x + t_x$$
$$y' = y + t_y$$

[1]Quelle:http://de.wikipedia.org/wiki/Koordinatentransformation

Man nennt $\begin{pmatrix} t_x \\ t_y \end{pmatrix}$ den Translationsvektor [1] [5].

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{2.6}$$

Rotation

Das Rotationszentrum sei $(0, 0)$. Vor der Rotation gilt:

$$x = r \cdot \cos(\gamma)$$

$$y = r \cdot \sin(\gamma)$$

Nach der Rotation gilt [1] [5]:

$$x' = r \cdot \cos(\gamma + \mu) = x \cdot \cos(\mu) - y \cdot \sin(\mu)$$

$$y' = r \cdot \sin(\gamma + \mu) = y \cdot \cos(\mu) + x \cdot \sin(\mu)$$

Somit lautet die Transformationsmatrix wie folgt:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} \cos(\mu) & -\sin(\mu) & 0 \\ \sin(\mu) & \cos(\mu) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}. \tag{2.7}$$

Skalierung

Der Fixpunkt der Skalierung sei $(0, 0)$. Es gelten folgende Gleichungen:

$$x' = x \cdot s_x$$
$$y' = y \cdot s_y$$

$s_x > 0$ und $s_y > 0$ heißen Skalierungsfaktoren. Das Objekt wird bei $s_x = s_y > 0$ vergrößert und bei $s_x = s_y < 0$ verkleinert. Die Darstellung als Matrix ergibt folgende Gleichung [1] [5]:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}. \tag{2.8}$$

Falls der Fixpunkt nicht bei $(0, 0)$ liegt, wird wie folgt vorgegangen:

1. Translastion Fixpunkt $\rightarrow$ Ursprung
2. Skalierung bzgl. $(0, 0)$
3. Rücktranslation zum ursprünglichen Fixpunkt

Scherung

Der Fixpunkt sei $(0,0)$. Es gelten dann folgende Gleichungen [1] [5]:

$$x' = x + Sch_x \cdot y$$
$$y' = y + Sch_y \cdot x$$

Spezialfälle für Scherung: x-Scherung: $Sch_x = 0 \rightarrow y' = y$ (Scherung in x-Richtung) y-Scherung: $Sch_y = 0 \rightarrow x' = x$ (Scherung in y-Richtung)

Die Matrixdarstellung der Scherung ergibt:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} 1 & Sch_x & 0 \\ Sch_y & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}. \tag{2.9}$$

Wenn man alle oben genannten Operationen aus (2.6), (2.7) und (2.8) zu einer Matrix zusammenfassen würde, erhielt man eine Matrix die wie folgt aussähe:

$$\begin{bmatrix} s_x \cos(\alpha) & s_y \sin(\alpha) & t_x \\ -s_x \sin(\alpha) & s_y \cos(\alpha) & t_y \\ 0 & 0 & 1 \end{bmatrix}. \tag{2.10}$$

Die affine Transformation hat sechs Parameter. Jedoch umfasst die affine Transformation, in dieser allgemeine Form, Bewegungsarten, die keine gültigen Bewegungen der steifen Gegenstände sind. Folglich können wir die Transformation weiter einschränken, indem wir die Scherung ($Sch_x = Sch_y = 0$) verbieten und nur isotrope Skalierung ($s := s_x = s_y$) erlauben. Wenn man dies in (2.10) ändert bleiben als Ergebnis nur vier Parameter übrig, nämlich einer für Rotation, einer für isotrope Skalierung und zwei für den Verschiebungsvektor. In diesem Fall kann die Matrix wie folgt ausgedrückt werden [2]

$$\begin{bmatrix} s \cos(\alpha) & s \sin(\alpha) & t_x \\ -s \sin(\alpha) & s \cos(\alpha) & t_y \\ 0 & 0 & 1 \end{bmatrix}, \tag{2.11}$$

wobei α der Winkel der Rotation ist und s den isotropischen Skalierungsfaktor repräsentiert. Obwohl dies nie in der Praxis zutreffen kann, wird das affine Kameramodell als Approximationswert noch häufig benutzt, wenn der Abstand zwischen Kamera und Szene groß, und die Bewegung klein ist. Wir versuchen die Parameter x' und y' zu ermitteln. Dazu verwenden wir die allgemeine Transformationsformel und erhalten folgendes Ergebnis [2]

$$\begin{pmatrix} x' \\ y' \\ w' \end{pmatrix} = \begin{bmatrix} h_{00} & h_{01} & h_{02} \\ h_{10} & h_{11} & h_{12} \\ h_{20} & h_{21} & h_{22} \end{bmatrix} \begin{pmatrix} x \\ y \\ w \end{pmatrix}. \tag{2.12}$$

Durch ausmultiplizieren der Matrix erhalten wir folgendes [2]:

$$x' = \frac{h_{00}x + h_{01}y + h_{02}}{h_{20}x + h_{21}y + h_{22}}, \qquad y' = \frac{h_{10}x + h_{11}y + h_{12}}{h_{20}x + h_{21}y + h_{22}}. \tag{2.13}$$

Aus oben genannter Matrix sehen wir, dass die letzte Zeile der Matrix immer $(0, 0, 1)$ ist. Daraus können wir die Werte h_{20},h_{21} und h_{22} ermitteln. Wir wissen jetzt $h_{20} = h_{21} = 0$ und $h_{22} = 1$. Wir setzen das in die obere Gleichung ein und erhalten folgende Gleichung:

$$x' = h_{00}x + h_{01}y + h_{02}, \qquad y' = h_{10}x + h_{11}y + h_{12}.$$

Der letzte Schritt ist einfach, nämlich die Transformationsfaktoren aus der Matrix in diese beiden Gleichungen einzusetzen:

$$\begin{aligned} x' &= s\cos(\alpha)x + s\sin(\alpha)y + t_x, \\ y' &= -s\sin(\alpha)x + s\cos(\alpha)y + t_y. \end{aligned} \tag{2.14}$$

Diese beide Formeln werden in affinen Modellen eingesetzt um die neuen Werte für x und y zu berechnen. Wir werden später zeigen, dass die affine Transformation ein Spezialfall des projektiven Transformation sind.

2.4.3 Projektive Modelle

Nachdem wir den Fall von affinen Bewegungen abgedeckt haben, und sie als eine Unterklasse der projektiven Bewegung beschrieben haben, können wir uns den allgemeinen projektiven Transformationen widmen. Wenn wir uns an die Gleichung für affine Bewegungen in homogenen Darstellungen erinnern, sehen wir, dass die letzten Zeilen der Matrix immer $(0, 0, 1)$ sind. Folglich hält $w' = w$ für jede mögliche Transformation, d.h., die Transformation ändert nicht die w-Koordinaten. Bei affinen Modellen wird die Bewegung eines Punktes nur innerhalb einer Ebene durchgeführt, die zur euklidischen Projektionsebene parallel ist. Da keine Änderung in der Höhe eintritt, beobachten wir keinen perspektivischen Projektionseffekt. Wenn wir das Modell so verallgemeinern, dass die letzte Linie willkürliche Werte enthält, haben wir die allgemeine Projektionstransformation. Wie wir zuvor gesehen haben, modelliert diese eine allgemeine Ebene zu Transformationsebene, die die 3D Bewegung einer Ebene, angesehen durch Pinlochkamera, beschreiben darf. Der Unterschied zur affinen Transformation ist, dass parallele Linien im Allgemeinen nach der Transformation nicht parallel bleiben. Stattdessen hat die projektive Transformation die Eigenschaft, dass die parallelen Linien zu Linien abgebildet werden, die sich in einem allgemeinen Fluchtpunkt schneiden. Das schließt natürlich den Spezialfall der parallelen Linien, die parallel bleiben, mit ein, da die parallelen Linien sich in einem idealen Punkt in unendlich schneiden. Da sich die parallelen Linien in einem Fluchtpunkt schneiden, ist es klar, dass die Position des Punktes von der Richtung der Linien in das Bild abhängig sein kann.

Der Horizont

Die Menge der Fluchtpunkte wird Fluchtlinie genannt [2]. In der realen Welt kennen wir diese Linie als den Horizont. Seit die Horizontlinie und sein Verhältnis zur Linie an unendlich eine wichtige Rolle spielt, forscht man in diesem Gebiet, um herauszufinden, wie die Parameter der Horizontlinie von einer Transformationsmatrix ermittelt werden können. Wir bezeichnen die Transformation von Bildkoordinaten auf die Gegenstandsebene mit H. In dieser Formel können wir die Horizontlinie als Punkt p definieren, die durch die Transformation von H zu den Punkten an der Unendlichen abgebildet sind. In Anbetracht dessen, liegen die Punkte (x, y, w) im Unendlich falls $w = 0$, gilt [2]

$$\begin{pmatrix} x' \\ y' \\ w' \end{pmatrix} = \begin{bmatrix} h_{00} & h_{01} & h_{02} \\ h_{10} & h_{11} & h_{12} \\ h_{20} & h_{21} & h_{22} \end{bmatrix} \begin{pmatrix} x \\ y \\ w \end{pmatrix}. \tag{2.15}$$

Von diesem leiten wir die Bedingung ab, die für Punkte auf dem Horizont gilt [2]

$$(h_{20}, h_{21}, h_{22}) \cot p = 0.$$

Da diese Gleichung die gleiche Form wie eine allgemeine Liniengleichung hat, brauchen wir nur die letzte Zeile der Matrix H als auch die Linienparameter für die horizontale Linie l_h zu betrachten. Folglich haben wir das einfache Ergebnis:

$$l_h = (h_{20}, h_{21}, h_{22})^\top.$$

Lassen Sie uns schließlich den speziellen Fall von einer affinen Transformation behandeln. Wir wissen bereits, dass die affine Transformation den Fall Parallelismus enthält. Infolgedessen liegen die Fluchtpunkte immer auf der Linie am unendlich. Diese Beobachtung wird durch die Tatsache gestützt, dass die letzte Zeile der affinen Transformationsmatrix $(0, 0, 1)$ ist, welches l_∞^T entspricht. Folglich bleibt die Position der Linie am unendlich zu irgendwelchen affinen Bewegungen unverändert.

Projektiven Transformation

Die projektive Transformation mit der inhomogenen Formulierung gibt uns folgendes [2]:

$$\begin{pmatrix} x' \\ y' \\ w' \end{pmatrix} = \begin{bmatrix} h_{00} & h_{01} & h_{02} \\ h_{10} & h_{11} & h_{12} \\ h_{20} & h_{21} & h_{22} \end{bmatrix} \begin{pmatrix} x \\ y \\ w \end{pmatrix}. \tag{2.16}$$

Durch Ausmultiplizieren der Matrizen erhalten wir folgendes [4]:

$$\begin{pmatrix} x' \\ y' \\ w' \end{pmatrix} = \begin{bmatrix} h_{00} & h_{01} & h_{02} \\ h_{10} & h_{11} & h_{12} \\ h_{20} & h_{21} & h_{22} \end{bmatrix} \begin{pmatrix} x \\ y \\ w \end{pmatrix} = \begin{pmatrix} h_{00}x + h_{01}y + h_{02} \\ h_{10}x + h_{11}y + h_{12} \\ h_{20}x + h_{21}y + h_{22} \end{pmatrix} \tag{2.17}$$

Wie bereits im Abschnitt homogene Koordinaten erklärt, um die Koordinaten in euklidischen Raum umzuwandeln, teilen wir diese durch dir letzte Zeile der Matrix bzw. w und daraus entstehen folgende Gleichungen [2]:

$$x' = \frac{h_{00}x + h_{01}y + h_{02}}{h_{20}x + h_{21}y + h_{22}}, \qquad y' = \frac{h_{10}x + h_{11}y + h_{12}}{h_{20}x + h_{21}y + h_{22}}. \tag{2.18}$$

Da die Parameter $\{h_{kl}\}$ im Vergleich zu einer Gesamtskalierung unverändert sind, ist die praktische Normierung anzuwenden $h_{22} = 1$. Besonders in der Literatur über Bewegungsabschätzung für Videokodierung, wird normalerweise die inhomogene Formel benutzt. Jedoch soll man nicht vergessen, dass diese Normierung nur möglich ist wenn $h_{22} \neq 0$. Es ist wichtig zu wissen, in welchen Fällen h_{22} dem Wert Null entspricht, um zu entscheiden, wann die inhomogene Formel anwendbar ist. Die Überprüfung von der Linie in unendlich gibt uns mehr Einblick in den Fall. Wir wissen, dass die Horizontlinie im Photo wie folgt lautet: $(h_{20}x + h_{21}y + h_{22})$. Der spezifische Fall $h_{22} = 0$ verursacht, dass die Horizontlinie die Form $h_{20}x + h_{21}y = 0$ hat, die entlang (?) der Linie ist, die den Ursprung durchläuft. Infolgedessen ist die Normalisierung $h_{22} = 1$ genau dann zulässig, wenn die Horizontlinie den Ursprung umfasst. Wenn diese Situation in einer spezifischen Anwendung auftritt, sollte die inhomogene Formel nicht benutzt werden. Im Falle der Beschreibung von Objektbewegung durch eine projektive Transformation zwischen aufeinander folgende Frames, ist die Bewegung so klein, dass die Normierung in der Regel kein Problem sein wird. Andererseits, wenn die Bewegung über einen große Zeitabstand geschätzt wird, kann dies doch ein Problem sein. Ignorieren wir diese Fälle, in denen die Normalisierung $h_{22} = 1$ unzulässig ist, erhalten wir die meist benutzte Formel der projektiven Bewegung [2]

$$x' = \frac{h_{00}x + h_{01}y + h_{02}}{h_{20}x + h_{21}y + 1}, \qquad y' = \frac{h_{10}x + h_{11}y + h_{12}}{h_{20}x + h_{21}y + 1}. \tag{2.19}$$

Oder, wenn verschiedene Symbole für die Parameter Transformation verwenden werden, um Ihre Relation zur affinen Transformation zu verbessern, schreiben wir dies wie folgt [2]

$$x' = \frac{h_{00}x + h_{01}y + t_x}{h_{20}x + h_{21}y + 1}, \qquad y' = \frac{h_{10}x + h_{11}y + t_y}{h_{20}x + h_{21}y + 1}. \tag{2.20}$$

Wir sehen, dass die affine Transformation als spezieller Fall von der projektiven Transformation für $h_{20} = h_{21} = 0$ zustande kommt. Weil der Nenner im affinen Fall verloren geht, ist die projektive Transformation eine nicht lineare Transformation. Dividiert man Zähler und Nenner durch $h_{22} \neq 0$, wird die Transformation um einen Parameter vereinfacht. Wir schreiben die h_{kj} Parameter um $b_{kj} = \frac{h_{kj}}{h_{22}}$. Wir erhalten [2]:

$$x' = \frac{b_{00}x + b_{01}y + b_{02}}{b_{20}x + b_{21}y + 1}, \qquad y' = \frac{b_{10}x + b_{11}y + b_{12}}{b_{20}x + b_{21}y + 1}. \tag{2.21}$$

Wir schreiben diese Gleichungen um [4]:

$$\begin{aligned} x'.(b_{20}x + b_{21}y + 1) &= b_{00}x + b_{01}y + b_{02}, \\ y'.(b_{20}x + b_{21}y + 1) &= b_{10}x + b_{11}y + b_{12}. \end{aligned} \tag{2.22}$$

Ausmultiplizieren [4]

$$\begin{aligned} b_{20}x'x + b_{21}x'y + x' &= b_{00}x + b_{01}y + b_{02}, \\ b_{20}y'x + b_{21}y'y + y' &= b_{10}x + b_{11}y + b_{12}. \end{aligned} \tag{2.23}$$

Trennen der unbekannten von bekannten Größen ergibt [4]:

$$\begin{aligned} x' &= b_{00}x + b_{01}y + b_{02} - b_{20}x'x - b_{21}x'y, \\ y' &= b_{10}x + b_{11}y + b_{12} - b_{20}y'x - b_{21}y'y. \end{aligned} \tag{2.24}$$

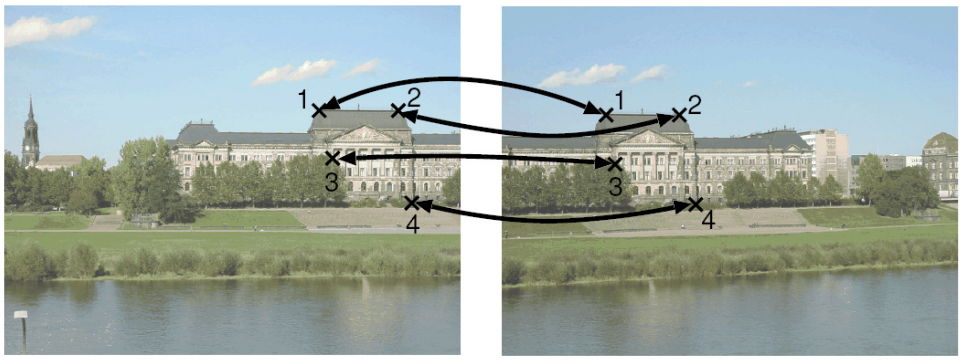
Mit minddestens 4 Punktpaaren (Korrespondierende Punkte) zwischen Quellbild und Zielbild erhält man 8 Gleichungen zur Bestimmung der 8 Paramter $b_{00}, ..., b_{21}$. Wenn $k = 1...n$ dann gilt [4] [3]:

$$\begin{aligned} x'_k &= b_{00}x_k + b_{01}y_k + b_{02} - b_{20}x'_k x_k - b_{21}x'_k y_k, \\ y'_k &= b_{10}x_k + b_{11}y_k + b_{12} - b_{20}y'_k x_k - b_{21}y'_k y_k. \end{aligned} \tag{2.25}$$

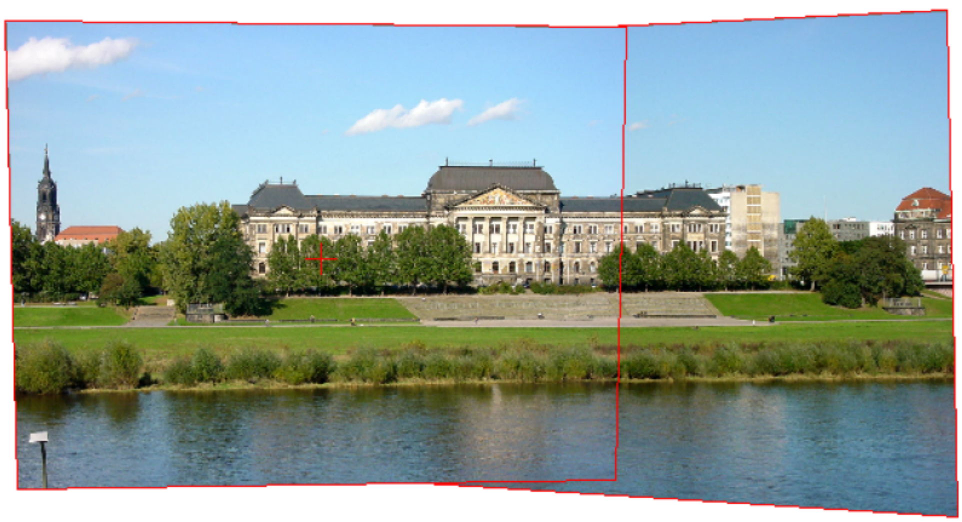
Schreiben wir das Ergebnis. Als eine Matrixmultiplikation gilt [4]:

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x'_1 x_1 & -x'_1 y_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -y'_1 x_1 & -y'_1 y_1 \\ \cdots \\ \cdots \\ x_n & y_n & 1 & 0 & 0 & 0 & -x'_n x_n & -x'_n y_n \\ 0 & 0 & 0 & x_n & y_n & 1 & -y'_n x_n & -y'_n y_n \end{bmatrix} \begin{pmatrix} b_{00} \\ b_{01} \\ b_{02} \\ b_{10} \\ b_{11} \\ b_{12} \\ b_{20} \\ b_{21} \\ b_{22} \end{pmatrix} = \begin{pmatrix} x'_1 \\ y'_1 \\ \cdots \\ \cdots \\ x'_n \\ y'_n \end{pmatrix}. \tag{2.26}$$

Wir lösen dieses lienare Gleichungssystem und erhalten die transformierenden Punkte x' und y'. Die Abbildung (2.2) zeigt ein Beispiel für den Einsatz von projektiven Modellen zur Erstellung von Panoramabildern. Wie bereits erwähnt, sucht man in zwei Bildern 4 korrespondierende Punkte im gemeinsamen Teil beider Bilder (2.2)(a), löst (2.26) für diese Punkte und somit werden die Bilder aneinander angepasst, so dass man daraus ein Panoramabild wie in (2.2)(b) erzeugen kann.



(a) Zwei Bilder mit 4 korrespondierenden Punkten



(b) Beide Bilder können mit Hilfe des projektiven Modells zusammengesetzt werden.

Abbildung 2.2: Anwendung [2]

Kapitel 3

Zusammenfassung

In dieses Arbeit werden zuerst euklidische, affine und projektive Räume vorgestellt. Um verschiedene Arten der Transformationen besser zu verstehen, wurde auf Homogene Koordinaten genauer eingegangen, um aus Matrixaddition und -multiplikation nur eine einzelne Matrixmultiplikation zu erzeugen. Um über affine Transformationen zu sprechen, war es notwendig auf verschiedene Transformationsarten wie Verschiebung, Drehung, Skalierung und Scherung einzugehen. Beim affine Kameramodell werden die bereits bekannten Transformationen zusammengefasst und es enstand daraus eine Matrix die alle Transformationsarten enthielt. Um Perspektivenänderungen zu berücksichtigen, hat affine Kameramodell mit 6 Parametern nicht ausgereicht. Das Projektive bzw. perspektivische Kameramodell hat 8 Paramatern. Um das projektive Kameramodell anzuwenden benötigt man 4 korrespondierende Punkte. Dabei handelt es sich um 4 gemeinsame Punkte in zwei unterschiedlichen Bildern, die man auswählt. Mit Hilfe dieser Punkte kann man die beiden Bilder transformieren, so dass daraus eine Panoramabild ensteht. Wie man diese Punkte auswählt ist nicht Inhalt dieser Arbeit. Es wurde ebenfalls gezeigt, dass das affine Kameramodell eine Unterklasse des projektiven Kameramodells ist.

Literaturverzeichnis

[1] M. Alexa. Perspektivische Transformation und Projektion. Vorlesungsskript, Computergraphik I, Technische Universität Berlin, 2008.

[2] D. Farin. *Automatic Video Segmentation Employing Object/Camera Modeling.* PhD thesis, Technische Universiteit Eindhoven, Einhoven, The Netherlands, 2005.

[3] R. Klauer. Ebene koordinatentransformationen. Vorlseungsskript, Geoinformatik, Fachhochschule München, 2007.

[4] A. Meisel. 3D-Bildverarbeitung. Vorlesungsskript, 3D-Bildverarbeitung, Hochschule für Angewandte Wissenschaften Hamburg, 2008.

[5] W. Purgathofer. Geometrische Transformationen. Vorlesungsskript, Computergraphik 1, Technische Universität Wien, 2007.

[6] R. Szeliski and H.-Y. Shum. Creating full view panoramic image mosaics and environment maps. Technical report, Microsoft Research, 1997.