View Morphing

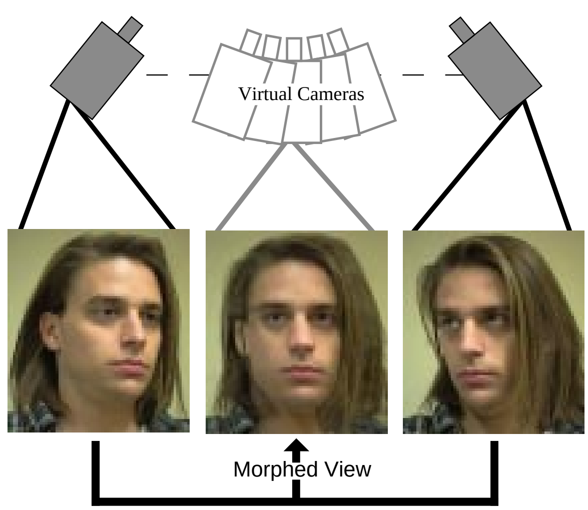
Steven M. Seitz Charles R. Dyer

Department of Computer Sciences
University of Wisconsin–Madison
publiziert in SIGGRAPH ’96, pp. 21 – 30

Abhandlung von Christian Limpach
Eidgenössische Technische Hochschule Zürich
27. Mai 1997

Zusammenfassung

Mittels Image Morphing ist es möglich überzeugende Übergänge zwischen 2-dimensionalen Objekten zu erzeugen. Verändern der Objektposition oder des Beobachtungspunktes führen jedoch zu unnatürlichen Verzerrungen, welche nur schwierig zu vermeiden sind. Das Paper von Seitz und Dyer stellt eine Methode vor, welche es mittels einfachen geometrischen Projektionen ermöglicht Übergänge zu erzeugen, bei welchen sich die Kamerasicht oder die Objektpositionierungen ändern. Diese Technik wird View Morphing genannt. Dabei werden die Bilder zuerst transformiert. Dann werden die Zwischenbilder für den Übergang erzeugt, welche dann in einem letzten Schritt wieder zurücktransformiert werden. Diese Methode kann auf jegliches Bildmaterial angewendet werden, da keine 3D Information über den Bildinhalt benötigt wird. Weil es möglich ist sowohl Änderungen in der Kamerasicht als auch in den Objektpositionierungen durchzuführen können mittels View Morphing viele interessante 3-dimensionale Effekte anhand einfacher Bildtransformationen berechnet werden.



1 Einführung

Bei Morphing werden weiche Übergänge zwischen Bildern berechnet. Bisher wurde dafür hauptsächlich eine 2D Interpolation durchgeführt, welche die Formen und Farben des ersten Bildes an die des letzten Bildes anpasst. Mit dieser Technik ist es möglich atemberaubende Spezialeffekte zu erzeugen, welche auf den Betrachter einen realistischen Eindruck machen. Es ist erstaunlich, dass Übergänge, die nur mittels 2D Berechnungen erzeugt wurden, eine natürliche 3D Transformation vermitteln können und sehr nützlich, da so aufwendige 3D Berechnungen vermieden werden können.

Jedoch garantieren aktuelle Methoden nicht, dass die erzeugte Transformation natürlich aussieht, es bleibt dem Benutzer überlassen zu entscheiden, ob ein Morph ausreichend realistisch aussieht und gegebenenfalls die Parameter anzupassen, um das beste Resultat zu erreichen. Vorallem Veränderungen der Kamerasicht oder der Objektpositionierung liefern meistens schlechte Resultate, insbesondere im Fall von eigentlich einfachen 3D Transformationen wie Translationen oder Rotationen.

Seitz und Dyer beseitigen dieses Problem durch eine Erweiterung, View Morphing. Dabei werden für den Übergang neue Ansichten der Objekte berechnet; dies liefert Resultate, die vergleichbar sind mit dem Bewegen der Kamera oder der Objekte selbst. So ist es möglich sämtliche 3D projektive Abbildungen zu simulieren. Weil als Zwischenschritt die gleichen Techniken wie beim Image Morphing benutzt werden, kann man auch bei View Morphing zwischen voneinander unterschiedlichen Bildern interpolieren. Also kann man gleichzeitig die Form, die Farbe und den Aufbau verändern.

Weil keine Kenntnisse über die 3D Form der Objekte benötigt wird, sind virtuelle Manipulationen von unbekannten Objekten, welche nur als Zeichnungen oder Photografien gegeben sind möglich.

Es gibt zwar noch andere Methoden um eine Veränderung der Kamerasicht zu erreichen. Beim View Morphing jedoch wird versucht natürliche Übergänge zu erzeugen. So kann man erstens die Kamerabewegung als Linie zwischen den beiden optischen Zentren der Bilder wählen und so das Verfahren vereinfachen. Zweitens ist das Verfahren ausreichend flexibel, so dass Übergänge zwischen zwei völlig unterschiedlichen Bildern berechnet werden können. Dies war bei bisherigen Verfahren meistens nicht der Fall. Ausserdem werden beim View Morphing bekannte Image Morphing Techniken benutzt, welche übernommen, bzw. erweitert werden können.

2 Image Morphing

Image Morphing nennt man die Klasse von Techniken, welche benutzt werden, um Übergänge zwischen Bildern zu berechnen. Es gibt mehrere Methoden, die alle auf der Interpolation von Pixelpositionen und Farben zwischen 2 Bildern basieren. Es scheint, dass es kein universelles Kriterium gibt, um die Qualität eines Morphs zu bewerten. Jedoch kann man sich die Frage stellen, ob die 3D Form der Objekte beim erzeugten Übergang erhalten bleibt. Dies ist meistens eben nicht der Fall, in den Zwischenbildern erkennt man daher oft Objekte, welche eine andere Form haben als die ursprünglichen Objekte.

Ein Morph besteht aus zwei Bildern $\mathcal{I}_0$ und $\mathcal{I}_1$ sowie zwei Übereinstimmungsabbildungen $C_0 : \mathcal{I}_0 \Rightarrow \mathcal{I}_1$ und $C_1 : \mathcal{I}_1 \Rightarrow \mathcal{I}_0$, welche eine komplette Übereinstimmung zwischen Punkten in den beiden Bildern beschreiben. Es sind zwei Abbildungen nötig, weil Punkte zusammen fallen können. In der Praxis legt der Benutzer Bereiche fest, die sich entsprechen. Die restlichen Übereinstimmung werden dann interpoliert. Die Interpolationsfunktion berechnet sich wie folgt:

$$W_0(\mathbf{p}_0, s) = (1-s)\mathbf{p}_0 + sC_0(\mathbf{p}_0) \tag{1}$$
$$W_1(\mathbf{p}_1, s) = (1-s)C_1(\mathbf{p}_1) + s\mathbf{p}_1 \tag{2}$$

W_0 und W_1 berechnen die Verschiebung der Punkte aus $\mathcal{I}_0$ und $\mathcal{I}_1$ für die Zwischenstufe $s \in [0,1]$. Mittels der Verschiebfunktionen erhält man zwei Zwischenbilder. Um das interpolierte Bild $\mathcal{I}_s$ zu berechnen, wird für jeden Punkt in $\mathcal{I}_s$ die Farbe als Mittel der Farbe der entsprechenden Punkte in den Zwischenbildern berechnet.

Die starken 3D Verzerrungen erklären sich wie folgt: Zwei verschiedene Ansichten eines planaren Objektes sind mittels einer 2D projektiven Abbildung ineinander überführbar. Aus der projektiven Abbildung lassen sich die benotigten Übereinstimmungsabbildungen berechnen. Die allgemeine Form einer solchen 2D projektiven Abbildung ist:

$$H(x,y) = \left(\frac{ax+by+c}{gx+hy+i}, \frac{dx+ey+f}{gx+hy+i}\right)$$

Da die Summe zweier solcher Ausdrücke im allgemeinen einen Bruch mit quadratischen Ausdrücken ergibt ist diese keine projektive Abbildung. Daher sind die interpolierten Bilder keine projektiven Abbildung des Objektes. Visuell äussert sich dies meistens als Verzerrung.

3 View Morphing

Beim View Morphing werden die Probleme des Image Morphing vermieden, d.h. die Objekte bewegen sich ohne ihre Form zu verändern zwischen ihren Positionen in den beiden Bildern.

Um einen Morph zu berechnen braucht man zwei Bilder $\mathcal{I}_0$ und $\mathcal{I}_1$, die dazugehörenden Projektionsmatrizen $\mathbf{\Pi}_0$ und $\mathbf{\Pi}_1$ und Korrespondenzen zwischen Pixeln in den beiden Bildern. Es wird also keine Information über die 3D Form der dargestellten Objekte benötigt. Die Projektionsmatrizen können z.B. entweder aus der Kameraposition oder mittels der 3D Position von wenigen Bildpunkten berechnet werden.

Die Korrespondenzen zwischen Pixeln in den beiden Bildern werden aus Informationen die der Anwendern liefert mittels Interpolation berechnet. Es hat sich gezeigt,

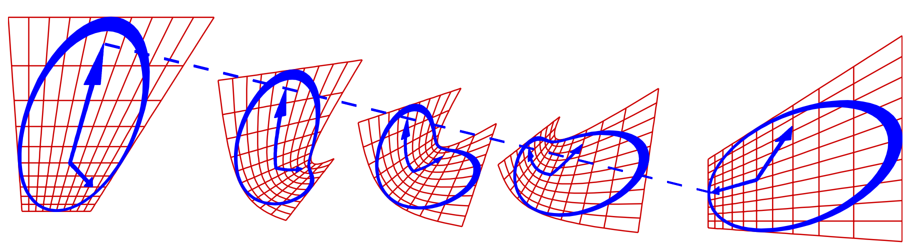


Abbildung 1: Ein Form-verzerrender Morph: Die Uhr wird bei der Interpolation verzerrt. Die gestrichelte Linie zeigt den geraden Weg, welchen der Zeiger während der Transformation nimmt.

dass minimal ungenaue Korrespondenzen noch ausreichen, um visuell korrekt erscheinende Übergänge zu berechnen. Bei grösseren Ungenauigkeiten in den Korrespondenzen treten sichtbar Fehler auf: Bildteile erscheinen z.B. aus dem Nichts (*ghosting*) oder der Übergang ist nicht mehr Form behaltend. Weitere Fehler sind möglich, wenn Teile nur in einem der beiden Bilder sichtbar sind, dies führt auch zu ghosting oder zu Löchern.

Im weiteren Verlauf wird die Projektionsmatrix anhand der Kameraposition gewählt: $\mathbf{\Pi} = [\mathbf{H}| - \mathbf{HC}]$. Der Vektor $\mathbf{C}$ entspricht den Kamerakoordinaten, die 3×3 Matrix $\mathbf{H}$ gibt die Position und Orientierung des Blickfeldes der Kamera an.

3.1 Parallele Ansichten

Abbildung 2 zeigt die Ausgangssituation für einen Morph zwischen zwei Bildern, die parallel sind. Dies entspricht einerseits dem Bewegen des Objektes, andererseits dem Bewegen der Kamera, jeweils parallel zum Blickfeld der Kamera. Da beide Bewegungen das gleiche Resultat liefern, reicht es aus nur die Bewegung der Kamera zu behandeln.

Die Ausgangsposition der Kamera ist $\mathbf{C}_0$ und dies ist zugleich der Ursprung des Koordinatensystems. Das Koordinatensystem ist desweiteren so orientiert, dass die Koordinaten von $\mathbf{C}_1$ gleich $(C_X, C_Y, 0)$ sind. Die Brennweite der Kamera verändert sich von f_0 zu f_1. Die Projektionsmatrizen sehen dann so aus:

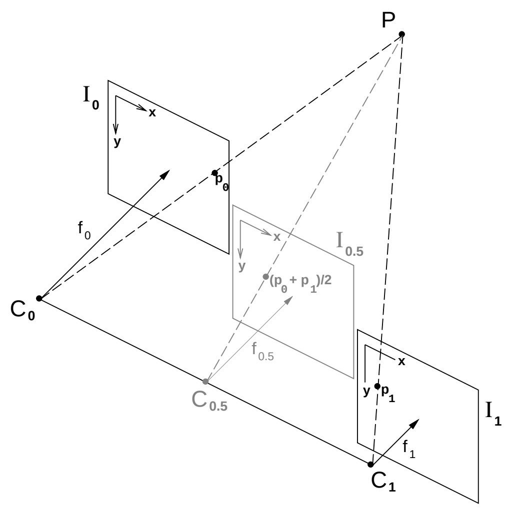


Abbildung 2: Morph zwischen parallelen Ansichten.

$$\mathbf{\Pi}_0 = \begin{bmatrix} f_0 & 0 & 0 & 0 \\ 0 & f_0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \qquad \mathbf{\Pi}_1 = \begin{bmatrix} f_1 & 0 & 0 & -f_1 C_X \\ 0 & f_1 & 0 & -f_1 C_Y \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Wenn $\mathbf{p}_0 \in \mathcal{I}_0$ und $\mathbf{p}_1 \in \mathcal{I}_1$ die jeweiligen Projektionen des Punktes $\mathbf{P} = [XYZ1]^T$ sind, dann berechnet sich die Interpolation von $\mathbf{p}_0$ und $\mathbf{p}_1$ wie folgt:

$$\begin{aligned} (1-s)\mathbf{p}_0 + s\mathbf{p}_1 &= (1-s)\frac{1}{Z}\mathbf{\Pi}_0\mathbf{P} + s\frac{1}{Z}\mathbf{\Pi}_1\mathbf{P} \\ &= \frac{1}{Z}\mathbf{\Pi}_s\mathbf{P} \end{aligned} \tag{3}$$

mit

$$\mathbf{\Pi}_s = (1-s)\mathbf{\Pi}_0 + s\mathbf{\Pi}_1 \tag{4}$$

Es wurde also jetzt eine neue Ansicht berechnet, mit Projektionsmatrix $\mathbf{\Pi}_s$. $\mathbf{\Pi}_s$ ist eine lineare Interpolation der Projektionsmatrizen $\mathbf{\Pi}_0$ und $\mathbf{\Pi}_1$ und stellt eine Kamera mit Position $\mathbf{C}_s$ und Brennweite f_s dar:

$$\mathbf{C}_s = (s\,C_X, s\,C_Y, 0) \tag{5}$$

$$f_s = (1-s)f_0 + sf_1 \tag{6}$$

Der Morph besteht nun aus einer Kamerabewegung entlang der Geraden $\overline{\mathbf{C}_0\mathbf{C}_1}$ und kontinuierlichem Vergrössern oder Verkleinern.

3.2 Nicht parallele Ansichten

Um auch zwischen Bildern mit nicht parallelen Ansichten morphen zu können, muss man die Bildern reprojezieren.

3.2.1 Bildreprojektion

Zwei Bilder $\mathcal{I}$ und $\hat{\mathcal{I}}$ mit identischem optischen Zentrum können ineinander gewandelt werden. Aus den Projektionsmatrizen $\mathbf{\Pi} = [\mathbf{H}| - \mathbf{HC}]$ und $\hat{\mathbf{\Pi}} = [\hat{\mathbf{H}}| - \hat{\mathbf{H}}\mathbf{C}]$ ergibt sich folgende Beziehung für die beiden Projektionen $\tilde{\mathbf{p}} \in \mathcal{I}$ und $\tilde{\hat{\mathbf{p}}} \in \hat{\mathcal{I}}$ des Punktes $\mathbf{P}$:

$$\hat{\mathbf{H}}\mathbf{H}^{-1}\tilde{\mathbf{p}} = \hat{\mathbf{H}}\mathbf{H}^{-1}\mathbf{\Pi}\mathbf{P} \tag{7}$$

$$= \hat{\mathbf{\Pi}}\mathbf{P} \tag{8}$$

$$= \tilde{\hat{\mathbf{p}}} \tag{9}$$

Die 3×3 Matrix $\hat{\mathbf{H}}\mathbf{H}^{-1}$ ist eine Projektion, welche die Bildebene $\mathcal{I}$ auf die Bildebene $\hat{\mathcal{I}}$ projiziert.

3.2.2 Der drei-Schritte Algorithmus

Mittels der Reprojektion kann jetzt bei nicht parallelen Ansichten die gleiche Methode wie bei parallelen Ansichten benutzt werden. Hier wird jetzt die X-Achse des Koordinatensystems parallel zur Geraden $\overline{\mathbf{C}_0\mathbf{C}_1}$ gewählt. Die beiden anderen Achsen sollten so gewählt werden, dass durch die Bild Reprojektion möglichst geringe Verzerrungen entstehen. Eine einfache Wahl, welche gut funktioniert, ist die Y-Achse parallel zum Kreuzprodukt der Normalen der beiden Bildebenen zu wählen.

Um das Bild $\mathcal{I}_s$ zu berechnen sind folgende Schritte notwendig:

1. Mit den Projektionen $\mathbf{H}_0^{-1}$ auf $\mathcal{I}_0$ und $\mathbf{H}_1^{-1}$ auf $\mathcal{I}_1$ werden die Bilder $\hat{\mathcal{I}}_0$ und $\hat{\mathcal{I}}_1$ erzeugt.

2. $\hat{\mathcal{I}}_s$ werden durch lineare Interpolation die Punkte und deren Farben aus den Bildern $\hat{\mathcal{I}}_0$ und $\hat{\mathcal{I}}_1$ berechnet.

3. Mit der Projektion $\mathbf{H}_s$ auf $\hat{\mathcal{I}}_s$ wird das Bild $\mathcal{I}_s$ erzeugt.

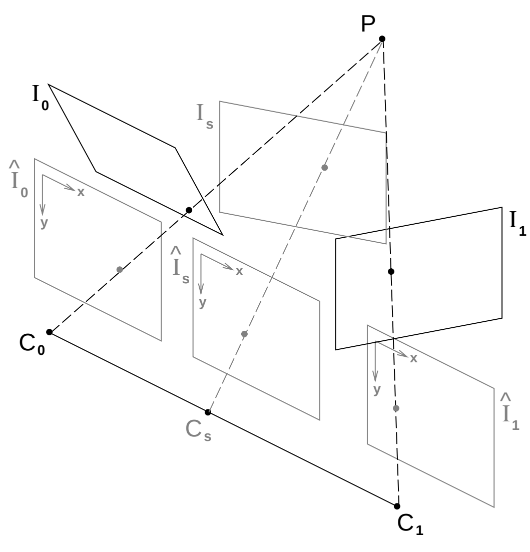


Abbildung 3: Morph in 3 Schritten.

Durch den 1. Schritt werden die Bildebenen aligniert. Da die Projektionsmatrizen der Bilder des ersten Schritts $\hat{\mathbf{\Pi}}_0 = [\mathbf{I}| - \mathbf{C}_0]$ und $\hat{\mathbf{\Pi}}_1 = [\mathbf{I}| - \mathbf{C}_1]$ mit der 3×3 Einheitsmatrix $\mathbf{I}$ eine spezielle Form haben, sind in $\mathcal{I}_0$ und $\mathcal{I}_1$ sich entsprechende Punkte jeweils in der gleichen Scanlinie. Dies erlaubt es die Interpolation eine Scanlinie nach der anderen zu berechnen und so entsprechend optimierte Algorithmen einzusetzen.

Da in den 3 Schritte die Bilder mehrmals umgerechnet werden, können Qualitätsverluste auftreten. Solche können vermieden werden, indem man die 3 Schritte zu einem zusammenfasst.

3.2.3 Änderungen in der Sichtbarkeit

Wenn Teile nur in einem der beiden Bilder sichtbar sind, kann dies zu Löchern oder zu Bildpunkten führen, für welche es mehrere richtige Farben gäbe.

Falls Tiefeninformation vorhanden ist, kann das zweite Problem behoben werden, indem die näherliegende Farbe gewählt wird, falls diese Information fehlt, dann kann die Farbe des Punktes gewählt werden, welcher im Originalbild am nächsten zum Problempunkt liegt.

Löcher können nicht so einfach eliminiert werden. Es wäre möglich eine Hintergrundfarbe zu definieren und diese dann zu benutzen, oder zwischen anliegenden Punkten zu interpolieren, oder auf Information aus zusätzlichen Bildern zurückzugreifen, falls diese vorhanden ist.

3.2.4 Erzeugung des Morphs

Man muss eine Sequenz von Projektionsmatrizen $\mathbf{\Pi}_s = [\mathbf{H}_s| - \mathbf{H}_s\mathbf{C}_s]$ zwischen $\mathbf{\Pi}_0$ und $\mathbf{\Pi}_1$ finden. Da die Kamera sich entlang einer Geraden bewegt, bleibt nur $\mathbf{H}_s$ zu wählen.

Eine Möglichkeit, welche ein natürliches Resultat liefert, ist zwischen den Bildebenen mittels einer Achsenrotation zu interpolieren. Die Rotationsachse und der Rotationswinkel sind dann:

$$\begin{aligned} \mathbf{D} &= \mathbf{N}_0 \times \mathbf{N}_1 & \qquad \mathbf{N}_0\text{und}\mathbf{N}_1\text{sind die} \\ \phi &= \cos^{-1}(\mathbf{N}_0.\mathbf{N}_1) & \qquad \text{Normalenvektoren von}\mathcal{I}_0\text{und}\mathcal{I}_1. \end{aligned}$$

4 Projizierende Transformation

View Morphing kann auch benutzt werden um einen Übergang zwischen zwei Ansichten desselben Objektes, die durch eine 3D projektive Abbildung voneinander abhängen, zu berechnen. Der Vorteil gegenüber Image Morphing hierbei ist, dass Bildattribute, wie z.B. gerade Linien, nicht verzerrt werden.

Eine 3D projektive Abbildung wird als 4×4 Matrix $\mathbf{T}$ mit homogenen Koordinaten dargestellt. Ein Punkt $\mathbf{P}$ wird hierbei in den Punkt $\hat{\mathbf{Q}}$ umgerechnet, $\hat{\mathbf{Q}} = \mathbf{TP}$, woraus man mittels Division durch die 4. Koordinate $\mathbf{Q}$ erhält. Nun kann man diese 3D projektive Abbildung mit der Transformation der Kamera verschmelzen, d.h. für den Punkt $\mathbf{P}$ gilt dann: $\hat{\mathbf{p}} = \mathbf{\Pi}(\mathbf{TP})$ oder mit einer neuen Projektionsmatrix $\hat{\mathbf{\Pi}} = \mathbf{\Pi T}$, $\hat{\mathbf{p}} = \hat{\mathbf{\Pi}}\mathbf{P}$.

Wegen dieser Umformung können jetzt beliebige Transformationen, welche durch eine 3D projektive Abbildung beschrieben werden, mittels des 3-Schritte Algorithmus berechnet werden.

Ein Problem bleibt noch, nämlich um die Projektionsmatrix zu bestimmen, müsste für die beiden Ausgangsbilder die Kamerapostion bekannt sein. Auch hierfür gibt es eine Lösung. Das Resultat von Schritt 1 sollten nämlich zueinander parallele Ansichten der Ausgangsbilder sein, in welchen sich entsprechende Punkte in der gleichen Scanlinie befinden. D.h. die Projektionsmatrizen für Schritt müssen so gewählt werden, dass diese Bedingung erfüllt ist. Wenn mindestens 8 Punktzusammengehörigkeiten bekannt sind, ist dies schon möglich.

4.1 Aussehen des Morphs

Um das Aussehen des Morphs zu bestimmen, muss noch die Projektionsmatrix $\mathbf{H}_s$ bestimmt werden. Anstatt die Matrix direkt anzugeben, können auch in $\mathcal{I}_0$ und $\mathcal{I}_1$ je 4 Punkte gewählt werden, welche sich entsprechen, z.B. Punkte, die eine Fläche definieren, oder die Endpunkte von 2 Linien. Danach muss der Benutzer noch die Position dieser 4 Punkte im Resultatbild $\mathcal{I}_{0.5}$ angeben. Die Position dieser Punkte in den restlichen Resultatbildern kann dann interpoliert werden. Abbildung 4 veranschaulicht diese Methode.

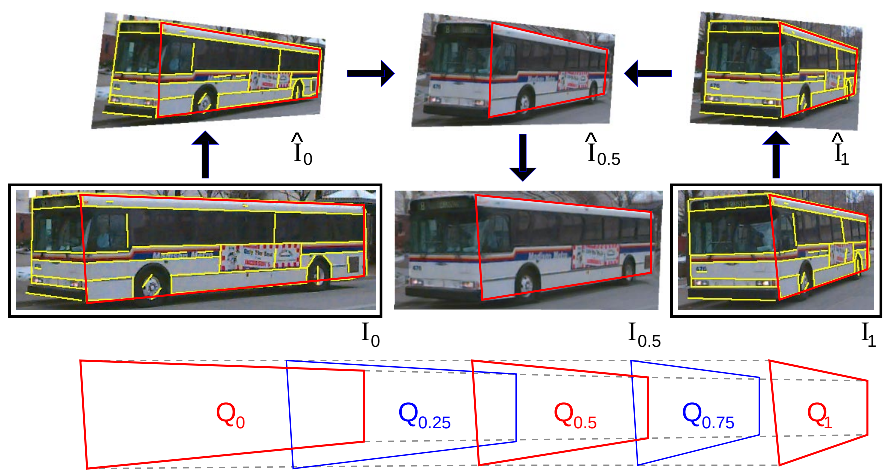


Abbildung 4: Aussehen des Morphs: Der Benutzer bestimmt die Position der Seitenwand des Busses im den Bildern $\mathcal{I}_0$, $\mathcal{I}_1$ und $\mathcal{I}_{0.5}$.

4.2 View Morphing ohne Schritt 1

Die Berechnung der Projektionsmatrix für den 1. Schritt des 3-Schritte Algorithmus ist bei Bildern, welche sich völlig unterscheiden nicht mehr ohne Fehler automatisch durchführbar. Der 3. Schritt sollte jedoch weiterhin durchgeführt werden, weil dadurch unnatürliche Verzerrungen eliminiert werden können.

Bei Bildern, welche schon fast identisch ausgerichtet sind, kann der 1. Schritt ebenfalls weggelassen werden.

5 Schlussfolgerung

Mittels View Morphing ist es möglich, realistische Bildübergänge zu berechnen, ohne Zugriff auf die 3D-Daten der Objekte zu haben. Es können sowohl Kamerabewegungen simuliert, als auch Bildübergänge zwischen z.B. Photos, Zeichnungen oder berechneten Bildern von identischen und nicht identischen Objekten berechnet werden.

Es gibt zwei mögliche Methoden um die Bildübergänge zu steuern, einerseits können die Kamerapositionen verwendet werden, andererseits können vom Benutzer im Bild definierte Punkte über eine ebenfalls vom Benutzer angegebene Bahn bewegt werden.

Die besten Resultate erzielt man, wenn jeweils in beiden Bildern ähnliche Bildteile sichtbar (bzw. nicht sichtbar) sind. Oft kann hier jedoch der Fehler reduziert werden, indem die fehlenden Bildteile anderen Bildteilen zugeordnet werden. Ansonsten könnten diese Fehler auch mittels zusätzlichen Bilddaten beseitigt werden.