

Benutzeroberfläche



Features

- mehrere Filter zur Optimierung der Segmentierung
- Objektfiler zur Entfernung kleinerer Objekte im Vordergrund
- Automatische Distanzeinstellung mit Portrait- und Gruppenmodus
- Gestensteuerung zur Auswahl des Hintergrundbilds
- beliebiges Bild als Hintergrund verwendbar
- Screenshot-Funktion
- WXGA Auflösung (KinBox HD)



**Internationaler Studiengang Medieninformatik
Semesterprojekt Bachelor WS 2012/13
Hochschule für Technik und Wirtschaft Berlin**

Team:
Sarah Schölzel
Leo Liebig
Ramin Pradhan

Betreuender Professor:
Prof. Dr. Klaus Jung

<http://kinbox.f4.htw-berlin.de>



KinBox

Die Software KinBox erkennt mithilfe der Tiefenbildkamera „Kinect“ Personen im Vordergrund eines Bildes und kann diese in Echtzeit freistellen. Der Hintergrund des Ursprungsbildes lässt sich somit durch eine beliebige Grafik ersetzen. Dieses Verfahren könnte eine Alternative zum weit verbreiteten Chroma Key-Verfahren darstellen und Anwendung im TV oder der Videotelefonie finden.

KinBox ist ein Praxisprojekt des Bachelorstudiengangs „Internationaler Studiengang Medieninformatik“ an der HTW Berlin und wurde von Sarah Schölzel, Leo Liebig und Ramin Pradhan umgesetzt. Das Projekt wurde am 10.10.2012 begonnen und am 08.02.2013 im Rahmen der IMI-Showtime präsentiert und abgeschlossen. Betreuender Professor war Prof. Dr. Klaus Jung.

Kinect

Die Kinect liefert kontinuierlich RGB- und Tiefenbilder sowie Skelett-Informationen der Personen vor der Kamera. Das RGB-Bild hat standardmäßig eine Auflösung von 640x480 px bei einer Framerate von 30 Frames pro Sekunde. Alternativ kann ein hochauflöstes Bild mit 1280x960 px verwendet werden, wobei die Framerate dann aber nur 12 fps beträgt.

Zur Erzeugung des Tiefenbilds projiziert die Kinect ein vordefiniertes Punktemuster aus Infrarotstrahlen in den Raum. Dieses Muster wird vom CMOS-Sensor der Kinect wieder aufgenommen und mit dem ausgesendeten Muster verglichen. Anhand der Verzerrungen des aufgenommenen Musters werden anschließend die Tiefen für jeden Pixel berechnet. Die Distanzangabe für jeden Pixel erfolgt in Millimeter und der Messbereich erstreckt sich von 400 bis 8000 mm. Die Kinect liefert 30 Tiefenbilder pro Sekunde, die eine maximale Auflösung von 640x480 px haben. Außerdem können Personen durch ihr Skelett als sogenanntes Skeleton erkannt werden. Dies enthält 3D-Koordinaten für die Position der einzelnen Körperteile. Das Tiefenbild ist die Basis für die auf das RGB-Bild angewendete Alphamaske.

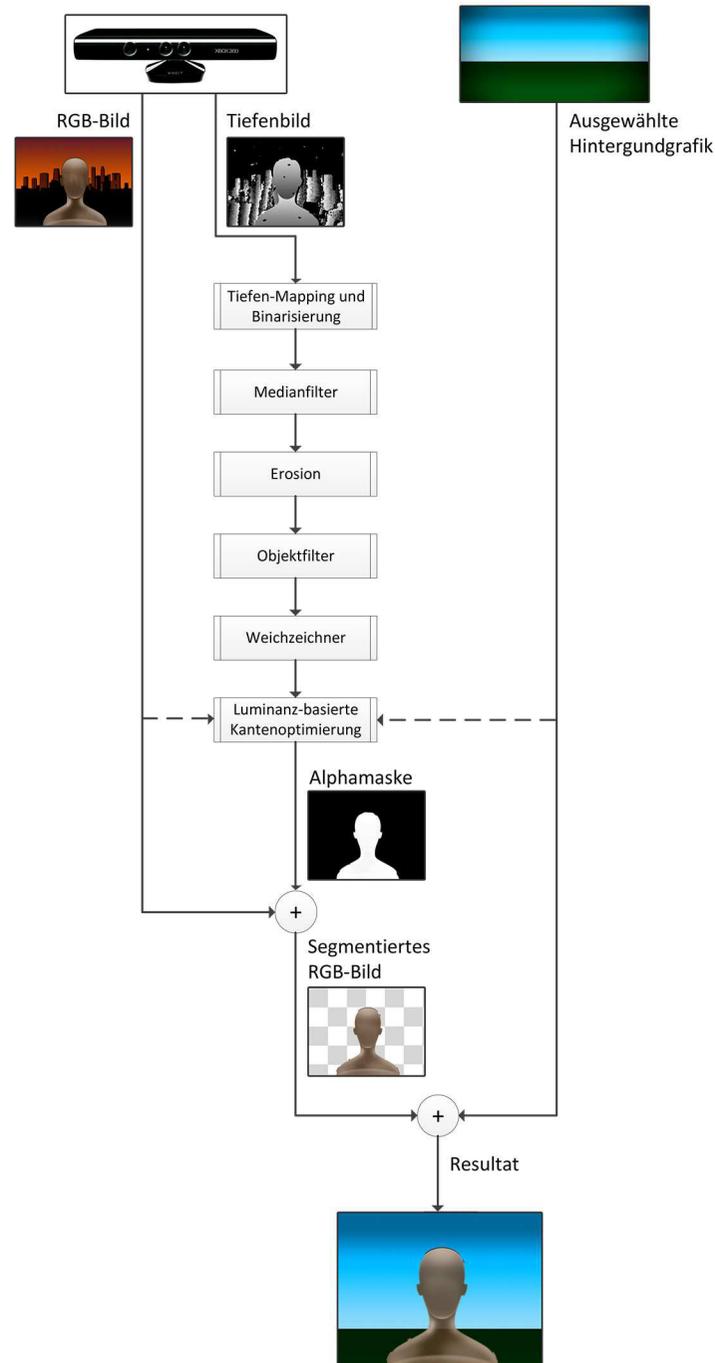
Filter

Die Filter in KinBox helfen eine optisch möglichst ansprechende Alphamaske für die Segmentierung zu erstellen. Weiße Flächen in der Alphamaske sind im Resultat sichtbar und schwarze Flächen werden durch das ausgewählte Hintergrundbild ersetzt.



Das Ausgangsbild nach der Binarisierung.

Ablaufdiagramm



Medianfilter

Der Medianfilter entfernt Rauschen in der Alphamaske, d.h. einzelne weiße Pixel in schwarzen Flächen werden schwarz und einzelne schwarze Pixel in weißen Flächen werden weiß. Dieser Filter vermeidet außerdem, dass der später folgende Flood-Fill-Algorithmus Störungen als Flächen erkennt und sich dessen Laufzeit erhöht.



Rauschen und kleine Störungen wurden gefiltert.

Erosion

Durch die Erosion werden weiße Flächen in der Alphamaske verkleinert. Das ist notwendig, da in einem späteren Verarbeitungsschritt die Alphamaske weichgezeichnet wird. Die Weichzeichnung erweitert weiße Flächen und würde ohne vorherige Erosion nicht erwünschte Pixel des RGB-Ursprungsbilds durchscheinen lassen.



Schrumpfen: Die ursprüngliche weiße Fläche ist hier zur Verdeutlichung rot dargestellt.

Objektfilter

Der Objektfilter basiert auf dem Flood-Fill-Algorithmus und filtert kleine, störende Objekte. Er erkennt zusammenhängende weiße Flächen in der Alphamaske und berechnet deren Größe. Flächen, deren Pixelanzahl kleiner ist als der festgelegte Schwellwert, werden schwarz gefärbt und sind somit im Resultat der Segmentierung nicht mehr sichtbar.



Die Hand am rechten Rand wurde entfernt, da ihre Fläche in Pixeln unter dem definierten Schwellwert lag.

Weichzeichner

Dieser Box-Filter zeichnet die Alphamaske weich und verbessert damit den Übergang von segmentierten Objekten auf den ausgewählten Hintergrund. Der Filter hat einen 5x5 Pixel Kernel, der aus Performancegründen in einen horizontalen und vertikalen Kernel unterteilt ist.



Weiche Kontur nach dem Weichzeichnen.

Luminanz-basierte Kantenoptimierung

Diese Optimierung vermindert störende Kanten, die auftreten können, wenn Pixel im Randbereich der Segmentierung deutlich heller oder dunkler als der Hintergrund sind. Beim Mischen des segmentierten Bilds und des Hintergrundes werden die Luminanzen des Vordergrund- und Hintergrundpixels verglichen. Übersteigt die Differenz einen festgelegten Schwellwert, dann wird der Alphawert des Vordergrundpixels in Abhängigkeit der Differenz reduziert.



Entfernung weißer Pixel des Ursprungsbildes im Randbereich der Segmentierung.