

## IMAGE STITCHING

### Übungsblatt 3, Computer Vision

Aufbauend auf der Python-Funktion zur projektiven Entzerrung soll nun ein Panorama aus einer Serie von Bildern eines (möglichst ebenen) Objekts erstellt werden. Dazu gehen wir wie folgt vor:

1. Wie bei der projektiven Entzerrung vermessen wir in jedem Bild mindestens 4 Punkte, für die wir geschätzte Weltkoordinaten festlegen. Sinnvollerweise sollten einige dieser Punkte im Überlappungsbereich beider Bilder liegen.
2. Mit Hilfe der projektiven Entzerrung werden die Bilder in das Weltkoordinatensystem transformiert.
3. Die transformierten Bilder werden miteinander in einem Bild verschmolzen.

Idealerweise sollte jeder Pixel in den überlappenden Bereichen die gleiche Intensität haben, aber dies ist in der Realität nicht der Fall. Gründe dafür sind unterschiedliche Belichtungszeiten oder Blendenwahl, Vignettierung (Abnahme der Intensität zu den Bildrändern hin), Schätzungsfehler bei der projektiven Entzerrung, radiale Verzerrung der Kameraoptik usw.

Eine einfache Strategie ist die gewichtete Verschmelzung der Bilder. Dabei wird jeder Pixel mit seinem Abstand zu den Mittellinien gewichtet, so daß Pixel in der Bildmitte Gewicht 1 und Pixel an den Rändern Gewicht 0 haben. Für ein Bild mit Breite  $M$  und Höhe  $N$  und ein Pixel an der Position  $i, j$  berechnet sich das Gewicht als

$$w_{ij} = (1 - \frac{2}{M}|i - \frac{M}{2}|) \cdot (1 - \frac{2}{N}|j - \frac{N}{2}|).$$

In überlappenden Bildbereichen wird entweder nur der Pixel mit dem höheren Gewicht übernommen, oder eine gewichtete Summe aus beiden Pixeln berechnet. Im ersten Fall besteht die Gefahr, daß sichtbare Diskontinuitäten entstehen, im zweiten Fall können aufgrund von Registrierungsfehlern Details verloren gehen. Die beste Vorgehensweise ist daher das sogenannte *multi-band blending*: Mit Hilfe eines Tiefpasses wird das Bild in einen hoch- und einen tieffrequenten Anteil zerlegt. Der tieffrequente Anteil, also die groben Strukturen, werden durch gewichtete Mittelung vereinigt, der hochfrequente Details durch Beibehaltung des Pixels mit dem höchsten Gewicht.

#### Aufgaben:

- a. Erweitern Sie Ihre Python-Funktion zur projektiven Entzerrung so, daß sie mehr als 4 Passpunkte verarbeiten kann.
- b. Nehmen Sie mit einer Kamera Ihrer Wahl eine Serie von mindestens 4 Aufnahmen eines ebenen Objektes (z.B. Fassaden, Häuserfront) von unterschiedlichen Standpunkten her auf. Legen Sie eine Liste von Passpunkten mit geschätzten Weltkoordinaten an, wobei pro Bild mindestens 4 Passpunkte (besser mehr) vorhanden sein sollten. Idealerweise sollten jeweils mindestens 2 Passpunkte im Überlappungsbereich liegen.

c. Schreiben Sie eine Funktion, die ein projektiv entzerrtes Bildpaar miteinander in ein neues, größeres Bild verschmelzen kann. Der Benutzer soll dabei wählen können, ob jeweils nur die Pixel mit dem größeren Gewicht beibehalten werden, oder ob ein gewichtetes Mittel aus allen Pixeln berechnet wird.

d. Schreiben Sie eine weitere Funktion, die ein projektiv entzerrtes Bildpaar in einen Hoch- und einen Tiefpaßanteil zerlegt und dann über *multi-band blending* vereinigt. Der Tiefpaß (z.B. ein Gauß-Tiefpaß oder ein gleitender Mittelwert) kann hierzu mit der Matlab-Funktion *filter2* realisiert werden. Den Hochpaßanteil erhält man einfach durch Subtraktion des Tiefpaßanteiles vom Eingangsbild.

*Nützliche Python-Funktionen:* Die Pseudoinverse einer Matrix wird mithilfe von *numpy.linalg.pinv* berechnet.