
Übung zur Vorlesung

Rechnersehen 2

SS 2015

Übungsblatt 2: Kamerakalibrierung (und SIFT-Merkmale)

Ausgabe: 06.05.2015

Abgabe: 20.05.2015

Aufgabe 1 Wiederholung RS1: Merkmale detektieren

(2 Punkte)

Auf der Seite <http://www.vlfeat.org/~vedaldi/code/siftpp.html> steht eine Implementierung der SIFT-Merkmale zur Verfügung. Diese ist im unten angegebenen Pfad auch lokal installiert. Ein mitgeliefertes Beispielprogramm kann eigenständig .ppm-Bilder einlesen und die gefundenen Merkmale in Textform ausgeben. Alternativ kann das Paket auch als MATLAB-Bibliothek VLFeat in eigenen Programmen verwendet werden.

Verwenden Sie diese Implementierung, um in einem Bild Schlüsselpunkte zu detektieren und entsprechende Deskriptoren zu extrahieren. Stellen Sie die gefundenen Punkte in einem Bild als Pfeile dar. Die Längen der Pfeile sollen dabei die Skalierungen (Scale) der Merkmale angeben und die Pfeilrichtungen den Orientierungen entsprechen.

Aufgabe 2 Wiederholung RS1: Korrespondenzen finden

(3 Punkte)

Gegeben seien nun zwei Bilder und darin detektierte SIFT-Merkmale. Ermitteln Sie korrespondierende Punkte zwischen diesen Bildern. Hierfür wird einerseits ein geeignetes Abstandsmaß für die Deskriptoren benötigt, wie z.B. der euklidische Abstand. Zum anderen wird eine Zuordnungsstrategie benötigt, wie z.B. Nächster-Nachbar-Zuordnung. Wählen Sie entsprechende Verfahren für Ihre Implementierung und begründen Sie die Wahl.

Stellen Sie die gefundenen Korrespondenzen geeignet dar.

Aufgabe 3 Extrinsische Parameter

(10 Punkte)

Bei der Kamerakalibrierung soll eine Abbildung von Weltkoordinaten in Bildkoordinaten bestimmt werden. Hierfür werden meist eine Menge von Weltkoordinaten mit zugehörigen Bildkoordinaten bestimmt, beispielsweise Punkte in einem Kalibrieremuster. Zwei Kalibrierverfahren basierend auf solchen Korrespondenzen wurden von Tsai [2] angegeben.

Im Gegensatz zu dem in der Vorlesung vorgestellten Fall mit 3-D Punkten in allgemeiner Lage, liegen in dieser Übungsaufgabe alle 3-D Punkte in einer Ebene (koplanarer Fall). Als zusätzliche Einschränkung wird ein bekanntes Ansichtenverhältnis $\alpha = 1$, $f_x = f_y = f$, bekannter (affiner) Scherungsparameter $s = 0$ und ein bekannter Bildhauptpunkt in der Bildmitte angenommen.

Zunächst werden die extrinsischen Parameter der Rotation und Translation bestimmt, indem die Transformationsvorschrift von Welt- in Kamerakoordinaten aufgestellt wird:

$$x_c = r_{11}x_w + r_{12}y_w + r_{13}z_w + t_x \quad (1)$$

$$y_c = r_{21}x_w + r_{22}y_w + r_{23}z_w + t_y \quad (2)$$

$$z_c = r_{31}x_w + r_{32}y_w + r_{33}z_w + t_z \quad (3)$$

Die Weltkoordinaten sind in den Gleichungen mit (x_w, y_w, z_w) und die Kamerakoordinaten mit (x_c, y_c, z_c) bezeichnet. Im koplanaren Fall lässt sich eine Vereinfachung vornehmen und man setzt $z_w = 0$. Die Kamerakoordinaten werden anschließend in die folgenden Abbildungsgleichungen eingesetzt:

$$x_s = (x_i - o_x) = f \frac{x_c}{z_c} \quad y_s = (y_i - o_y) = f \frac{y_c}{z_c} \quad (4)$$

Dabei ist der Punkt (o_x, o_y) der Bildhauptpunkt, was im koplanaren Fall mit dem Mittelpunkt des Bildes gleichgesetzt wird. Nach Auflösen und Gleichsetzen des Nenners z_c ergibt sich daraus die Gleichung (vergleiche Vorlesung, Kapitel 4, Seite 28):

$$x_s f (r_{21}x_w + r_{22}y_w + t_y) - y_s f (r_{11}x_w + r_{12}y_w + t_x) = 0 \quad (5)$$

Für N Punkte folgt ein homogenes lineares Gleichungssystem, welches sich mit Standardverfahren (SVD) bis auf eine unbekannte Skalierung lösen lässt. Der erste Teil der gesuchten Parameter ergibt sich somit aus:

$$\begin{pmatrix} x_s^{(1)} x_w^{(1)} & x_s^{(1)} y_w^{(1)} & x_s^{(1)} & -y_s^{(1)} x_w^{(1)} & -y_s^{(1)} y_w^{(1)} & -y_s^{(1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_s^{(n)} x_w^{(n)} & x_s^{(n)} y_w^{(n)} & x_s^{(n)} & -y_s^{(n)} x_w^{(n)} & -y_s^{(n)} y_w^{(n)} & -y_s^{(n)} \end{pmatrix} \begin{pmatrix} r_{21} \\ r_{22} \\ t_y \\ r_{11} \\ r_{12} \\ t_x \end{pmatrix} = 0 \quad (6)$$

Die Lösung des Gleichungssystems (6) führt zu den vier bis auf Skalierung bekannten Einträgen \tilde{r}_{11} , \tilde{r}_{12} , \tilde{r}_{21} und \tilde{r}_{22} . Dank der Orthonormalität der Rotationsmatrix R lassen sich daraus auch die übrigen erschließen. Auch die Skalierung k der Lösung v des homogenen Gleichungssystems (6) wird durch diese Bedingung festgelegt. Eine Herleitung findet sich z.B. in [1]:

$$b = -(\tilde{r}_{11}^2 + \tilde{r}_{12}^2 + \tilde{r}_{21}^2 + \tilde{r}_{22}^2) \quad c = (\tilde{r}_{11}\tilde{r}_{22} - \tilde{r}_{21}\tilde{r}_{12})^2$$

$$k^2 = \frac{-b + \sqrt{b^2 - 4c}}{2} \quad (7)$$

Genaugenommen kann auch das Vorzeichen von k bestimmt werden. Dies entscheidet über ein Rechts-/Links-Koordinatensystem, ist in dieser Aufgabe aber nicht relevant. Unabhängig vom Vorzeichen lassen sich die Elemente der Rotationsmatrix R (erneut bis auf Vorzeichen) berechnen aus:

$$\tilde{r}_{11}^2 + \tilde{r}_{12}^2 + \tilde{r}_{13}^2 = k^2 \quad \tilde{r}_{21}^2 + \tilde{r}_{22}^2 + \tilde{r}_{23}^2 = k^2 \quad (8)$$

$$\text{sgn}(r_{23}) = -\text{sgn}\left(\frac{r_{11}r_{21} + r_{12}r_{22}}{r_{13}}\right)$$

Der dritte Vektor r_3 folgt aus dem Kreuzprodukt der beiden anderen Vektoren. Nicht vergessen werden sollte eine Normierung der Vektoren r_i auf die Länge 1 (Skalierung k). Zwei der Translationsparameter können nach dieser korrekten Skalierung von v auf $\frac{v}{k}$ ebenfalls aus Gleichung (6) abgelesen werden.

In den unten angegebenen Pfaden finden sich Bilder mit zugehörigen Punktkorrespondenzen. Es finden sich jeweils ganzzahlige Einträge für die Felder eines Schachbrettes und Pixelkoordinaten, die sub-pixel genau die entsprechenden Bildpunkte angeben. Schreiben Sie ein Programm, das diese gegebenen Paare von Bild- und Weltkoordinaten einliest und die oben genannten Kameraparameter bestimmt. Entsprechende Punktepaare finden Sie in den unten angegebenen Dateien, welche sich beispielsweise einfach mit dem Befehl `textscan` einlesen lassen.

Aufgabe 4 Weitere Parameter

(5 Punkte)

Schließlich fehlen noch f und t_z . Um auch diese zu bestimmen, setzen wir Gleichung (2) in (4) ein und erhalten (vergleiche Vorlesung, Kapitel 4, Seite 34):

$$y_s(r_{31}x_w + r_{32}y_w + t_z) - f(r_{21}x_w + r_{22}y_w + t_y) = 0 \quad (9)$$

und daraus mit n bekannten Punkten das inhomogene lineare Gleichungssystem:

$$\begin{pmatrix} y_s^{(0)} & -(r_{21}x_w^{(0)} + r_{22}y_w^{(0)} + t_y) \\ \vdots & \vdots \\ y_s^{(n)} & -(r_{21}x_w^{(n)} + r_{22}y_w^{(n)} + t_y) \end{pmatrix} \begin{pmatrix} t_z \\ f \end{pmatrix} = \begin{pmatrix} -y_s^{(0)}r_{31}x_w^{(0)} - y_s^{(0)}r_{32}y_w^{(0)} \\ \vdots \\ -y_s^{(n)}r_{31}x_w^{(n)} - y_s^{(n)}r_{32}y_w^{(n)} \end{pmatrix} \quad (10)$$

Schreiben Sie ein Programm, dem die Ergebnisse aus der obigen Aufgabe zur Verfügung stehen und das nun noch die Parameter t_z und f berechnet. In der Originalarbeit von Tsai wird weiterhin erwähnt, falls sich ein negativer Wert für f ergibt, sind die Vorzeichen in der Rotationsmatrix R falsch gewählt. Die obige Wahl des Vorzeichens von r_{13} bzw. r_{23} muss in diesem Fall geändert werden, womit sich gleichzeitig die Vorzeichen von r_{31} und r_{32} umkehren. Berücksichtigen Sie in Ihrem Programm auch diesen Fall.

In den unten angegebenen Verzeichnisse finden Sie auch Vergleichswerte für die Kameraparameter. Dabei ist zu beachten, dass kleine Abweichungen durchaus normal sind und bei den Vergleichswerten eine Normierung der Bildkoordinaten auf den Bereich -1 bis 1 vorgenommen wurde.

Viel Spaß und Erfolg!

Literatur

- [1] Berthold K.P. Horn. Tsai's camera calibration method revisited, 2000.
http://people.csail.mit.edu/bkph/articles/Tsai_Revisited.pdf.
- [2] Roger Y. Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Journal of Robotics and Automation*, 3(4):323–344, August 1987.

Pfade:

SIFT++:	http://www.robots.ox.ac.uk/~vedaldi/code/siftpp.html
	/home/sickert/uebung-rechnersehen2/siftpp
Eingabebilder:	/home/sickert/uebung-rechnersehen2/02-sift/input
Eingabebilder Kalibrierung	/home/sickert/uebung-rechnersehen2/03-calib/input0/ /home/sickert/uebung-rechnersehen2/03-calib/input1/