

Computer Vision Challenge - SoSe 2023

Klaus Diepold, Luca Sacchetto, Sven Gronauer
cv.ldv@xcit.tum.de

07.06.2023

Zusammenfassung

Als Teamleiter in der IT-Abteilung eines schwedischen Einrichtungskonzerns werden Sie beauftragt, eine Applikation zu entwickeln, die den Kunden die Möglichkeit gibt, die Inneneinrichtung einer 3D-Szene der eigenen Wohnung zu gestalten. Der Kunde soll dabei einfach selbst aufgenommene Bilder des Raumes in die Applikation hochladen und erhält ein 3D-Modell des Innenraums (siehe Abbildung 1).

1 Datensatz

In Abbildung 2 und Abbildung 3 sehen Sie beispielweise die Aufnahmen einer Halle und des Eingangsbereichs eines Gebäudes aus dem Datensatz ETH3D [1]. Für die Bearbeitung der Challenge stellen wir Ihnen verschiedene Datensätze aus ETH3D auf Moodle zur Verfügung. Für die Begutachtung der von Ihnen abgegebenen Software werden wir auch Bilder von weiteren, unbekannten Innenräumen verwenden. Selbstverständlich ist es Ihnen erlaubt, auch mit Bildern aus anderen Quellen, oder gar selbst aufgenommene Bilder, zu arbeiten. Bedenken Sie, dass der Datensatz ETH3D teilweise auch die Position von Merkmalspunkten, deren 3D-Rekonstruktion sowie Angaben zur Rotation und Translation der Kameras beinhaltet. Diese sind nur für das Debuggen konzipiert und werden in den Bewertungs- und Test-Datensätze nicht vorhanden sein. Ein Beispieldatensatz (siehe Moodle) für Testdaten enthält lediglich die Bilder, die intrinsischen Kameraparameter und die Position der Kamera in Weltkoordinaten.

2 Spezifikation

2.1 Ziel

Ziel der Computer Vision Challenge ist es eine Matlab Applikation zu entwickeln. Ihre Applikation muss in der Lage sein, ein abstraktes 3D-Modell des Innenraums zu generieren. Darüber hinaus soll die Applikation dem Nutzer die Möglichkeit geben sich Informationen über die Maße und Abstände der Objekte anzeigen zu lassen.

2.2 Applikation

Ihre Abgabe muss eine Hauptfunktion namens `main.m` beinhalten, die ausgeführt werden kann, um Ihr Programm zu starten. Alle weiteren Interaktionen mit der von Ihnen entwickelten Applikation müssen über eine **graphische Benutzeroberfläche** (GUI) geschehen. Die GUI muss über folgende Funktionen verfügen:

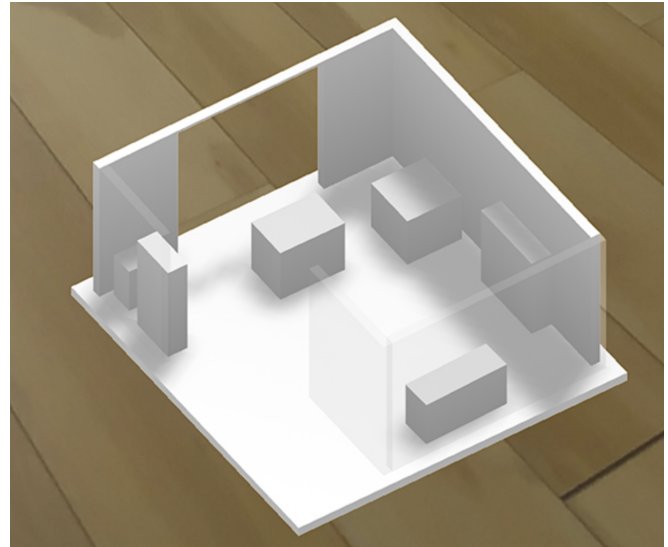


Abbildung 1: Abstraktes 3D-Modell eines Innenraums

- **Auswahl der Bilder:** Ihr Programm sollte dem Anwender die Möglichkeit geben, eine *Szene* oder ein beliebiges anderes Bild (z.B. von der Festplatte) auszuwählen.
- **Einstellung der Ansicht:** Ihr Programm muss es dem Anwender erlauben, die Ansicht des 3D-Modells zu ändern.
- **Anzeige von Maßen und Größen:** Ihr Programm muss es dem Anwender erlauben, Maße und Größen im 3D-Modell zu vermessen.

Achten Sie bei der Auswahl Ihrer Funktionen auf die Nützlichkeit für eine potentielle Anwendergruppe, die nicht notwendigerweise Experten in Computer Vision sind. Bei der Entwicklung der Applikation dürfen **nur offizielle Mathworks Toolboxes** verwendet werden.

2.3 Bilder

Ihre Applikation muss mit Bildern zurechtkommen, welche die folgenden Eigenschaften aufweisen:

- Unterschiedliche Bildgrößen
- Unbekannte Bilder (welche nicht auf Moodle zur Verfügung stehen)

3 Poster Session

3.1 Poster

Erstellen Sie für die Poster Session am **12.07.2023** ein Poster im PDF-Format. Wegen der großen Teilnehmerzahl wird die



Abbildung 2: Datensatz kicker

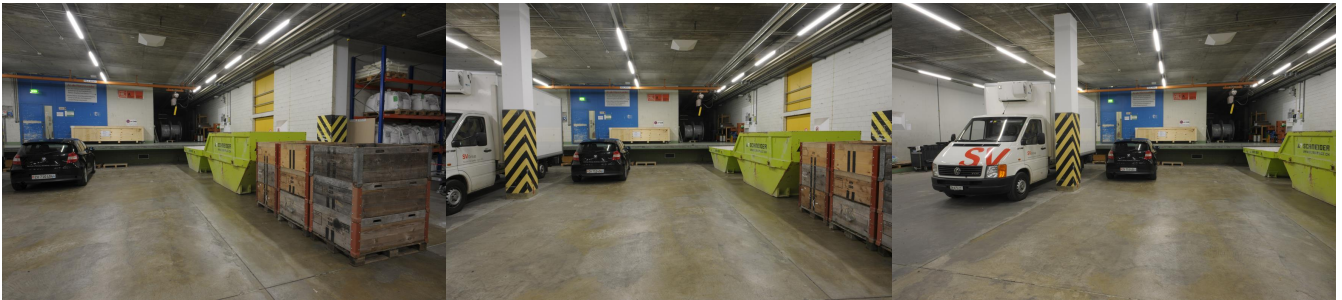


Abbildung 3: Datensatz delivery_area

Poster Session über Zoom stattfinden. Das Poster stellt eine Art Abschlusspräsentation Ihres Gruppenprojektes dar und muss die folgenden Punkte enthalten:

- Eingesetzte Algorithmen (inkl. Referenzen)
- Pipeline der Datenverarbeitung/Algorithmen.
- Resultate und Ergebnisse anhand von mindestens zwei Szenen:
 1. eine Szene aus den auf Moodle verfügbaren Datensätzen.
 2. eine selbstausgesuchte Szene, welche **nicht** auf Moodle verfügbar ist.
- Probleme und Herausforderungen.

Bitte achten Sie auf eine angemessene Detailtiefe, die sich in einer 5-minütigen Vorstellung des Posters sinnvoll vermitteln lässt (nicht zu viele Details). Es gibt für diese Aufgabe keine offizielle Vorlage oder Designbeschränkungen für die Gestaltung des Posters. Bitte wählen Sie selbstständig ein geeignetes Layout und verwenden Sie eine klare, sorgfältige Darstellung. Achten Sie zudem auf die Angabe von Referenzen. Weitere Infos zur Gestaltung des Posters und zur Poster Session erhalten Sie in einer der nächsten Präsenzveranstaltungen. Die erlaubten Sprachen auf dem Poster und in dessen Präsentation sind Deutsch oder Englisch.

Jede Gruppe muss ein Poster erzeugen, aber jedes Mitglied der Gruppe sollte in der Lage sein das Poster zu präsentieren bzw. zu erläutern und Fragen dazu zu beantworten. Während der Poster Session werden alle Poster zeitgleich in dedizierten Breakout-Räumen vorgestellt. Die teilnehmenden Studierenden können sich frei bewegen und sich alle Poster ansehen und diskutieren. Jedes Poster muss allerdings von mindestens einem Gruppen-Mitglied besetzt und präsentiert werden. Alle Team-Mitglieder sind somit aufgerufen eine Schicht als Präsentator zu übernehmen, damit alle Teilnehmer auch alle

Poster besuchen können.

3.2 Live-Demo

Während der Präsentation des Posters muss eine Live-Demo der entwickelten Applikation präsentiert werden. Für die Begutachtung der von Ihnen abgegebenen Software werden wir auch Bilder von weiteren, unbekannten *Szenen* verwenden – sogenannte Testset Bilder, die Sie im Rahmen einer Live-demo während der Poster Session vorführen. Die Testset Daten werden wir Ihnen am Tag vor der Poster Session zur Verfügung stellen.

4 Bewertungskriterien

Dieser Abschnitt geht auf die *vorläufigen* Bewertungskriterien ein. Bedenken Sie, dass die Computer Vision Challenge ein Gruppenprojekt ist, d.h. die Gruppe wird einheitlich benotet und jedes Mitglied ist mitverantwortlich für die gesamte Abgabe.

In der Bewertung der Challenge werden sowohl die Poster Session als auch die abgegebene Software gleichwertig gewichtet.

4.1 Mindestanforderungen an Software

Diese Anforderungen müssen gegeben sein, um die Challenge zu bestehen.

- Ausführbarkeit des Programms auf einem beliebigen System (Windows, Mac oder Unix)
- Vollständigkeit des Programms bei Abgabe
- Codequalität und Kommentare
- Vorhandensein und Bedienbarkeit einer GUI.

- Grobe Rekonstruktion der Szene (z.B. Boxen anstatt feiner Strukturen).
- Vermessung von Maßen und Größen im 3D-Modell.
- Einstellung der Ansicht des 3D-Modells.

Ihr Code wird nach der Abgabe auf seine Komplettheit und Ausführbarkeit auf den bereits bekannten als auch auf unbekannten Szenen überprüft.

4.2 Qualitätsmerkmale der Software

Mit diesen Merkmalen verbessern Sie die Bewertung der Computer Vision Challenge.

- Qualität der Visualisierung (Visueller Eindruck)
- Komplexität der rekonstruierten Szene (Objekte, Texturen, Strukturen, etc.).
- Sehr guter Umgang mit unterschiedlichen *Szenen*
- Laufzeit des Programms (Je schneller, desto besser)
- Benutzerfreundlichkeit und Übersichtlichkeit der GUI

4.3 Poster Session

Die folgenden Punkte beziehen sich auf die Inhalte und der Darstellung ihres Posters:

- Inhalte und Übersichtlichkeit des Posters
- Qualität der Poster Darstellung/Präsentation
- Beantwortung von Fragen
- Qualität der Live-Demo

5 Abgabe

Geben Sie nur funktionierenden Code ab. Testen Sie Ihr Programm vorher auf einem unbekannten Rechner (Windows und Unix) z.B. im Eikon, ob dieses dort auch ausführbar ist und zum gewünschten Ergebnis führt. Zum Testen wird die Matlab Version 2023a verwendet. Wenn Sie mit einer stark abweichenden Matlab Version arbeiten, vergewissern Sie sich, dass keine grundlegenden Funktionen im Vergleich zu dieser Version geändert wurden. Komprimieren Sie Ihre Abgabe in einem *.zip-Archiv und geben Sie diese Datei auf Moodle für Ihre Gruppe ab.

Abgabefrist für das Poster und die Software ist der **11.07.2023 um 23:59 Uhr**. Nachträgliche Abgaben können nicht berücksichtigt werden!

Software:

- **Abgabeformat:** Sie können nur *.zip-Dateien abgeben. Achten Sie auf eine eindeutige Benennung nach dem Schema *GXX.zip* wobei *XX* für Ihre Gruppennummer steht! Sind Sie also in Gruppe 42 geben Sie eine Datei ab, die *G42.zip* heißt. Die Abgabe von Gruppe 2 heißt entsprechend *G02.zip*. Die Abgabe erfolgt über Moodle.
- **Readme:** Erstellen Sie eine *Readme.txt* Datei, die eine Anleitung Ihrer Applikation und eine Liste aller für

die Ausführung der Applikation notwendigen Toolboxen beinhaltet.

- **Uploadgröße:** Die Uploadgröße ist auf 100MB beschränkt.

Poster:

- **Abgabeformat:** Das Poster ist einzureichen als einseitiges PDF auf Moodle.

Literatur

- [1] Thomas Schöps, Johannes L. Schönberger, Silvano Galliani, Torsten Sattler, Konrad Schindler, Marc Pollefeys, and Andreas Geiger. A multi-view stereo benchmark with high-resolution images and multi-camera videos. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.