

Computer Vision Challenge - SoSe 2024

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

12. Juni 2024



Abbildung 1: *Saint Jerome in His Study*, H. Steenwick

Zusammenfassung

Im Rahmen der Computer Vision Vorlesung haben wir kennen gelernt wie eine 3D-Rekonstruktion mit Stereobildern durchgeführt werden kann. Jedoch können auch aus einem Einzelbild verschiedene Ansichten generiert werden. In der diesjährigen Challenge schauen wir einen dieser Ansätze an: [Tour into the Picture \[1\]](#).

1 Datensatz

In Abbildung 1 sehen Sie das Ölgemälde *Saint Jerome in His Study* von H. Steenwick (1580-1649) als exemplarisches Beispiel für eine Aufnahme aus der *Ein-Punkt-Perspektive* eines Raumes. Für die Bearbeitung der Challenge stellen wir Ihnen einen Datensatz von unterschiedlichen Bildern derartiger *Szenen* zur Verfügung. Die Komplexität der Szene steigt inkrementell in den Abbildungen 3a-3f.

2 Spezifikation

2.1 Ziel

Ziel der Computer Vision Challenge ist es eine Matlab Applikation zu entwickeln. Ihre Applikation muss in der Lage sein, verschiedene Perspektiven eines Raumes auf Grund eines Einzelbildes zu erstellen und zu visualisieren.

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:

- **Auswahl der Bilder:** Ihr Programm sollte dem Anwender die Möglichkeit geben, eine *Szene* oder ein beliebiges anderes Bild (z.B. von der Festplatte) auszuwählen.
- **Selektion:** Ihr Programm sollte dem Nutzer die Möglichkeit geben, verschiedenartige Punkte zu definieren, z.B. einen Fluchtpunkt und vier Punkte zum Einrahmen einer Ebene für eine Homographie. Eventuell könnte es von Vorteil sein Objekte als Vordergrund zu maskieren. Siehe auch: *spidery mesh interface* und die Details aus [1].
- **Einstellung der Perspektive:** Ihr Programm muss es dem Anwender erlauben, die Ansicht im Raum zu ändern.

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 zureckkommen, welche die folgenden Eigenschaften aufweisen:

- Unterschiedliche Bildgrößen
- Bilder mit uneindeutiger Ebene (z.B. Szene der Metrostation *Komsomolskaya* in Moskau, Abb. 3e)
- Bilder mit uneindeutigem Fluchtpunkt (z.B. Aufnahme der *Sagrada Familia*, Abb. 3f)
- Unbekannte Bilder (welche nicht im Datensatz enthalten sind, sondern im Testset)

3 Poster Session

3.1 Poster

Erstellen Sie für die Poster Session am **10.07.2024** ein Poster im PDF-Format. Wegen der großen Teilnehmerzahl wird die Poster Session über Zoom stattfinden. Das Poster stellt eine Art Abschlusspräsentation Ihres Gruppenprojektes dar und muss die folgenden Punkte enthalten:

- Verwendete Methoden und Ansätze, welche in der finalen Software implementiert sind (inkl. Referenzen)
- Pipeline der Datenverarbeitung/Algorithmen



Abbildung 2: Visualisierung des *Spidery-Mesh*

- Resultate und Ergebnisse anhand von mindestens 2 Szenen:
 1. eine Szene aus dem Datensatz (3a-3f).
 2. eine selbstausgesuchte Szene, welche **nicht** aus dem Datensatz (3a-3f) entspringt.
- 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 Be-gutachtung 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 benötigt 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

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
- Vorhandensein einer Code-Dokumentation (z.B. in der Form von Kommentaren)
- Vorhandensein und Bedienbarkeit einer GUI
- Generierung von verschiedenen Ansichten aus dem ausgewähltem Einzelbild

Ihr Code wird nach der Abgabe auf seine Komplettheit und Ausführbarkeit auf den bereits bekannten als auch auf unbekannten Szenen überprüft. Weitere Bewertungsaspekte sind die Erfüllung aller im Abschnitt 2 genannten Punkte.

4.2 Qualitätsmerkmale

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

- Qualität der Visualisierung (Visueller Eindruck)
- Komplexität der Szene mit denen Ihr Programm gut umgehen kann (z.B. uneindeutige Flächen bzw. Fluchtpunkt)
- Sehr guter Umgang mit unterschiedlichen *Szenen*
- Laufzeit des Programms
- 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 2024a 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 **09.07.2024 um 23:59 Uhr**. Die Abgabe erfolgt via Moodle. 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] Youichi Horry, Ken-Ichi Anjyo, and Kiyoshi Arai. Tour into the picture: using a spidery mesh interface to make animation from a single image. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 225–232, 1997.



(a) Komplexität: Einfach – *Einfacher Raum*



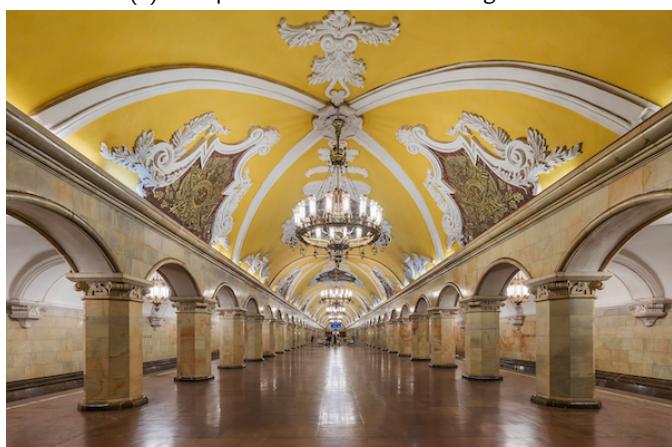
(b) Komplexität: Einfach – *Ölgemälde*



(c) Komplexität: Mittel – *Einkaufsgalerie*



(d) Komplexität: Mittel – *Uhrenturm*



(e) Komplexität: Schwierig – *Metrostation*



(f) Komplexität: Schwierig – *Kathedrale*

Abbildung 3: Datensatz mit steigender Komplexität der Szene.