**Note – Progetto geometric and 3d computer vision**

**Librarie richieste:**

* Numpy: array, operazioni matematiche
* Cv2: operazioni computer vision
* Glob: import di liste di file
* Math: Distanza tra due punti

**File del progetto (eseguirli nel seguente ordine):**

* Save\_frames: salva i frames di tutti i video (compreso quello della calibrazione)
* Calibration\_camera: calibrazione della camera
* Remove\_background: colora di nero l’oggetto e lascia lo sfondo invariato
* Elaborate\_video: elabora tutti i frames e crea il modello 3d

**Note dell’implementazione:**

* Per ogni frame, nella fase di rimozione dello sfondo, ho eliminato anche il bicchiere ma per fare questo ho dovuto includere i metodi erode e dilate che hanno rovinato la precisione del modello 3d ma per lo scopo di questo progetto va bene.
* Per ogni video salvo solo 40 frames che rappresentano l’intero giro dell’oggetto perché altrimenti l’esecuzione è troppo lunga. Con solo questi frames, il tempo totale di esecuzione (salvare i frames, rimuovere lo sfondo, elaborare il video, salvare immagini per visual debugging, creazione del modello 3d) è di circa 20 minuti

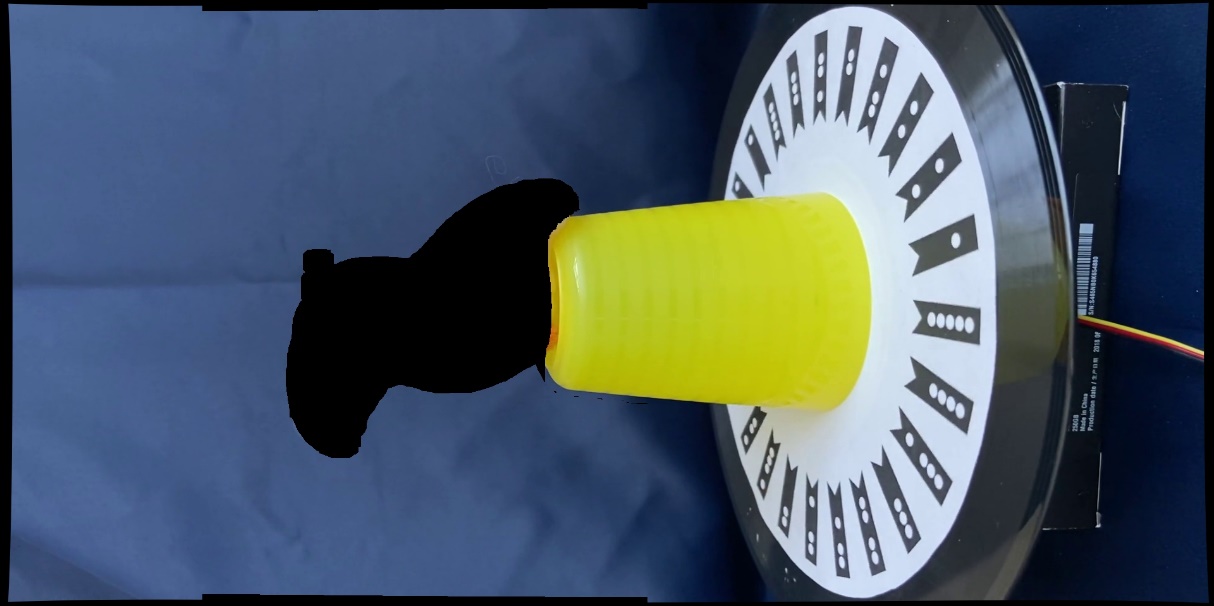
**Visual debug informations:**

* Binary threshold per identificare i poligoni e cerchi del marker

Immagine che contiene testo, ingranaggio

Descrizione generata automaticamente

* L’intero oggetto in nero lasciando invariato lo sfondo. Per fare questo ho dovuto ritagliare una sezione dell’immagine e poi identificare quale è il bicchiere per lasciare solo l’oggetto tramite una “sottrazione d’immagini”. Questo lo si può vedere nelle immagini seguenti. Altrimenti potevo farlo come quello che ci ha mostrato in aula lasciando un pezzo del bicchiere sotto all’oggetto e non servivano i metodi erode e dilate e di conseguenza c’era un modellino 3d impeccabile ma ho preferito farlo in questo modo come “altra sfida del progetto” a discapito dell’accuratezza.



* Ho ritagliato la giusta porzione di foto per avere uno sfondo di colore blu omogeneo e mediante il range sono riuscito ad eliminarlo interamente

Immagine che contiene silhouette

Descrizione generata automaticamente

* Ho identificato il bicchiere giallo del bicchiere (il bicchiere non ha una bella forma perché ho ridimensionato la foto per farla stare nel documento)

Immagine che contiene testo

Descrizione generata automaticamente

* Poi ho identificato i poligoni del marker e identificato alcuni poligoni specifici tramite il loro codice binario

Immagine che contiene testo

Descrizione generata automaticamente

* Poi quando ho trovato la corrispondenza tra posizione 2d e 3d di ciascun punto, ho proiettato tutti i poligoni (anche quelli nascosti dal bicchiere) per vedere se era corretta la correlazione

Immagine che contiene testo

Descrizione generata automaticamente

* Poi ho disegnato anche i relativi codici binari in ciascun poligono proiettato

Immagine che contiene testo

Descrizione generata automaticamente

* Poi ho disegnato il rettangolo attorno all’oggetto per capire quali punti dovevo proiettare (il rettangolo tanto più grande perché non volevo cambiare la sua dimensione tra i 4 video e quindi dovevo farci stare dentro qualunque tipo di oggetto)

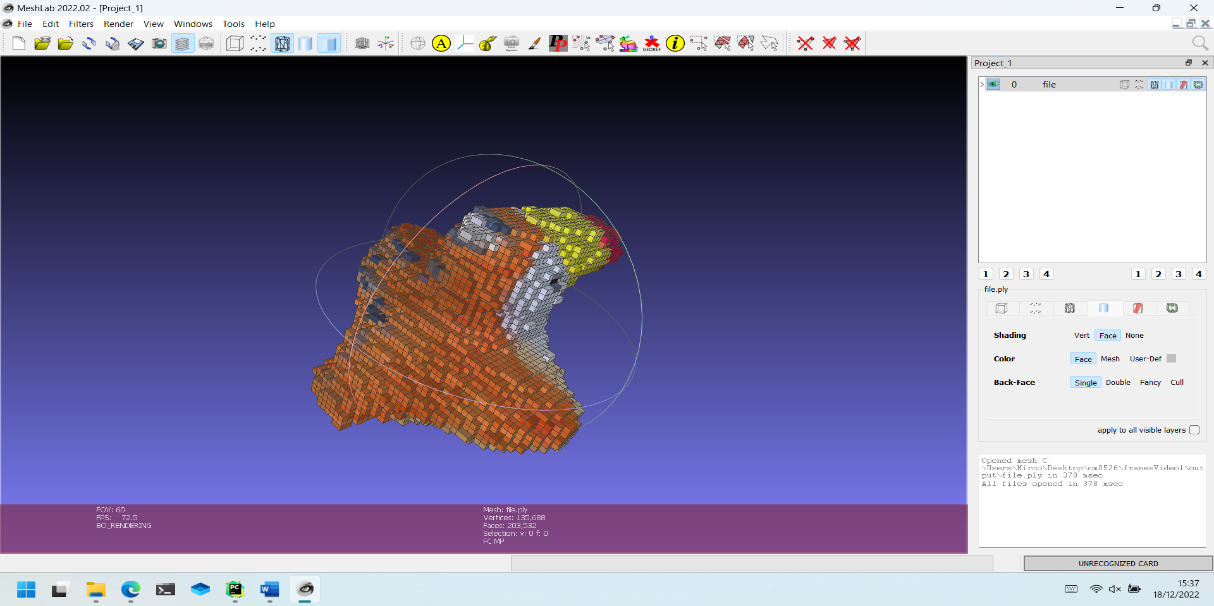
Immagine che contiene testo

Descrizione generata automaticamente

* Infine, ho creato il modellino 3d e allego alcune foto.

Immagine che contiene testo, screenshot, elettronico, schermo

Descrizione generata automaticamenteImmagine che contiene testo, screenshot, elettronico, schermo

Descrizione generata automaticamenteImmagine che contiene testo, screenshot, elettronico, schermo

Descrizione generata automaticamenteImmagine che contiene testo, screenshot, elettronico, schermo

Descrizione generata automaticamenteImmagine che contiene testo, screenshot, monitor, elettronico

Descrizione generata automaticamente

**Problemi del progetto:**

* Ho creato un rettangolo e non un cubo per la projection dei punti in quanto ho cercato di minimizzare al massimo il numero di punti per velocizzare l’algoritmo ed inoltre per non cambiare la dimensione del rettangolo in relazione al video in input ho dovuto creare un rettangolo dove ci stanno tutti gli oggetti.
* Ho cercato di eliminare interamente anche il bicchiere dal modello 3d, che però questo ha implicato l’uso di dilate ed erode in quanto alcune parti degli oggetti avevano la stessa colorazione del bicchiere e di conseguenza usando questi metodi ho ridotto enormemente la qualità dell’oggetto (questo lo si può vedere dalle prossime foto) che però per questo progetto va bene lo stesso in quanto bisogna dimostrare di saper applicare le diverse tecniche di computer vision.
* Inoltre, negli oggetti 1 e 3 ci sono delle etichette blu ed avendo la stessa colorazione dello sfondo, bisogna rimuoverle e questo lo ho fatto con il metodo erode ovviamente a discapito di precisione ed accuratezza ma avendo almeno un modello 3d senza buchi.
* Stessa cosa del punto precedente anche per l’oggetto numero 4 in quanto alcuni pixel hanno il codice rgb contenuto nel range dello sfondo e di conseguenza anche in questo caso ho dovuto usare il metodo erode peggiorando l’accuratezza dell’oggetto.
* Inoltre, il mio algoritmo di Pose estimation usa un metodo basato sul counting di cerchi pieni all’interno di diverse zone del poligono (per esempio se ho il poligono con cinque cerchi bianchi, il suo codice sarà 212) (le spiego meglio durante l’orale come funziona perché scriverlo diventa troppo lungo e difficile da capire). Questo algoritmo ha un problema quando in alcuni (pochissimi) frames vengono identificati i poligoni, mediante il loro codice binario, in una sola zona del marker e di conseguenza c’è un errore massimo di projection di 3/4 pixel in quel singolo frame (lo si può vedere nella prossima immagine). Per risolverlo servirebbe una prima pose estimation per identificare i primi poligoni e poi riapplicare un’altra pose estimation per cercare i poligoni vicini a quelli identificati. Però per lo scopo di questo progetto, il numero molto limitato di frame “rovinati” e visto il risultato finale che ottengo va bene lo stesso il mio algoritmo.

Immagine che contiene testo

Descrizione generata automaticamente

* Inoltre, dalla prossima immagine si può osservare che l’oggetto numero 2 ha una colorazione gialla in molte parti (ho usato il range per poter identificare la parte del bicchiere) e di conseguenza per poter tenere le parti dell’oggetto ho dovuto usare molte iterazioni di erode e dilate che hanno rovinato il modello 3d (lo si può vedere nelle prossime immagini). Di conseguenza potevo lasciare in minima parte il bicchiere per non dover rovinare l’immagine ma ho preferito farlo in questo modo per aggiungere un ulteriore step al progetto.

Immagine che contiene testo

Descrizione generata automaticamente

Immagine che contiene testo, screenshot, elettronico, schermo

Descrizione generata automaticamenteImmagine che contiene testo, screenshot, elettronico, schermo

Descrizione generata automaticamente

Immagine che contiene testo, screenshot, elettronico, schermo

Descrizione generata automaticamente

Immagine che contiene testo, screenshot, elettronico, schermo

Descrizione generata automaticamente