

Corso di Laurea in Ingegneria e Scienze Informatiche

Sviluppo di un sistema di visione artificiale per la rilevazione e localizzazione di marker ArUco in un contesto di aggregate computing

Tesi di laurea in:
PROGRAMMAZIONE AD OGGETTI

Relatore
Prof. Mirko Viroli

Candidato
Bacchini Lorenzo

Correlatore
Dott. Gianluca Aguzzi

Sommario

Il paradigma dell'aggregate computing rappresenta un nuovo approccio alla programmazione distribuita, focalizzando l'attenzione sulla definizione di comportamenti globali che vengono poi interpretati dai singoli dispositivi. L'elaborazione finale non avviene dunque su un sistema esterno, ma sono i dispositivi stessi a comunicare tra loro per prendere decisioni basate sul raggiungimento di un fine comune.

In questo contesto si inserisce un progetto sviluppato dai ricercatori dell'università di Bologna, volto a gestire uno sciame di robot o droni secondo algoritmi distribuiti e comportamenti auto-organizzativi. Generalmente in sistemi come quello sopra citato la localizzazione dei robot avviene grazie a tecnologie di posizionamento satellitari, nel caso in questione però è necessaria l'implementazione di una soluzione in grado di operare anche in ambito indoor, mantenendo costi contenuti e un'infrastruttura quanto più semplice possibile.

Al giorno d'oggi la visione artificiale trova applicazioni in una grande vastità di settori, dalla ricerca all'ambito industriale, con lo scopo di replicare la vista umana. L'uso di tali tecniche è ormai supportato da numerose librerie software ed è largamente diffuso, pertanto si è scelto di analizzare tale approccio al problema posto, sviluppando in questa tesi un sistema di posizionamento basato sul riconoscimento di marker ArUco, ottenuto grazie all'utilizzo della libreria di visione artificiale OpenCV.

I risultati conseguiti in questa dissertazione dimostrano la precisione dei sistemi di localizzazione basati su processi di visione artificiale, analizzando però i limiti di questa tecnologia in termini di condizioni ambientali e contesti di utilizzo.

Alla mia famiglia

Indice

Sommario	iii
1 Introduzione	1
2 Background	5
2.1 Visione artificiale	5
2.1.1 Cos'è la visione artificiale?	5
2.1.2 Come funziona la visione artificiale?	6
2.1.3 Applicazioni e finalità	6
2.1.4 Principali criticità	7
2.1.5 Cenni storici	7
2.2 Marker fiduciari	8
2.2.1 Applicazioni	8
2.2.2 ArUco markers	9
2.3 OpenCV	11
2.3.1 JavaCV	11
2.4 Aggregate Computing	12
2.4.1 Sistema distribuito	12
2.4.2 Aggregate computing	12
2.5 Calibrazione della camera	14
2.5.1 Cos'è la calibrazione?	14
2.5.2 Modello di calibrazione	15
2.5.3 Distorsione della lente	16
3 Analisi	19
3.1 Requisiti funzionali	20
3.2 Requisiti non funzionali	20
3.3 Requisiti tecnologici	21
3.4 Analisi del dominio	21

INDICE

4 Design	23
4.1 ArUco markers	23
4.2 Videocamera	24
4.3 Architettura	24
4.4 Design dettagliato	25
4.4.1 App	26
4.4.2 CameraCalibrator	27
4.4.3 CameraPose	27
4.4.4 ResolutionEnum	29
4.4.5 InputParameters	30
4.4.6 Risoluzione video	31
4.4.7 UI	31
4.4.8 Calcolo della posa	31
5 Implementazione	33
5.1 Creazione del progetto	33
5.2 JavaCV	33
5.3 Avvio dell'app	34
5.4 Calibrazione	35
5.4.1 Rilevazione scacchiera	35
5.4.2 Rifinitura angoli	36
5.4.3 Calibrazione	36
5.5 Posa	36
5.5.1 Cattura del frame	37
5.5.2 Rilevazione marker	38
5.5.3 Calcolo Posa	40
6 Valutazione	41
6.1 Testing	41
6.2 Esempio operativo	42
6.3 Verifica dei requisiti	44
7 Conclusioni	47
7.1 Sviluppi futuri	49

Elenco delle figure

2.1	Marker fiduciari [10]	9
2.2	ArUco Marker 4x4 id: 0	10
2.3	Logo OpenCV	11
2.4	Sistema centralizzato Vs distribuito [18]	13
2.5	Rappresentazione camera con obiettivo stenopeico	15
2.6	Processo di conversione dal mondo reale 3D all'immagine 2D	16
2.7	Distorsione radiale	16
2.8	Distorsione tangenziale	17
3.1	Diagramma a stati semplificato del dominio	22
4.1	Architettura del sistema	25
4.2	Architettura completa	26
4.3	Classe App	26
4.4	Classe CameraCalibrator	27
4.5	Classe CameraPose	29
4.6	Classe ResolutionEnum	30
4.7	Classe InputParameters	30
4.8	Associazione ResolutionEnum	31
4.9	Associazione InputParameters	31
4.10	Diagramma di sequenza del sistema	32
5.1	Posa della camera	37
6.1	Markers posizionati a 30cm sull'asse x	43
6.2	Markers posizionati a 15cm sull'asse y	43
6.3	Markers ruotati rispettivamente di 45 e 90 gradi	44
7.1	Robot identificati da marker ArUco	48
7.2	Sistema completo	49

ELENCO DELLE FIGURE

List of Listings

5.1	Implementazione dipendenza JavaCV	33
5.2	Caricamento classi OpenCV	34
5.3	Configurazione parametri app	34
5.4	Avvio calibrazione e posa	34
5.5	Rilevazione angoli interni della scacchiera	35
5.6	Rifinitura angoli	36
5.7	Calibrazione camera	36
5.8	ResolutionEnum.java	38
5.9	Regolazione tempo di esposizione	38
5.10	Riduzione risoluzione del frame	39
5.11	Rilevazione markers	39
5.12	Traduzione delle posizioni degli angoli	40
5.13	Calcolo Posa	40
6.1	Calcolo errore di riproiezione in fase di calibrazione	42
6.2	Calcolo errore di riproiezione in fase di posa	42

LIST OF LISTINGS

Capitolo 1

Introduzione

Il mondo del lavoro e la vita di tutti i giorni stanno notevolmente cambiando grazie all'avvento di nuove tecnologie, che in maniera più o meno evidente stanno invadendo la nostra quotidianità, ne sono la prova la moltitudine di dispositivi con la quale interagiamo costantemente come smartphone, smartwatch, smartbands, smartglasses ecc. Anche in ambito industriale stiamo assistendo alla formazione di quella che viene comunemente chiamata “Industria 4.0”, la quale promette di portare gli attuali sistemi informatici e di automazione ad un livello successivo basato sulla collaborazione e comunicazione degli stessi. Un tema emergente nell'ambito di sistemi distribuiti collaborativi è rappresentato dal paradigma dell'aggregate computing.

Il paradigma dell'aggregate computing [17] si differenzia dalla programmazione distribuita in quanto maggiore enfasi è posta sulla definizione di un obiettivo globale che il sistema deve raggiungere, piuttosto che sulle regole e sui singoli compiti che ogni nodo deve portare a termine, elevando così il livello di astrazione.

A partire da questo nuovo paradigma i ricercatori dell'università di Bologna hanno sviluppato un software capace di controllare una serie di robot/droni. Sulla base di concetti derivati dalla robotica degli sciami, esso consente ai dispositivi di auto-organizzarsi e conoscere lo stato generale del sistema, scambiando informazioni con i propri vicini. Un aspetto fondamentale del progetto sopracitato è dato dalla possibilità di determinare in ogni istante la posizione dei robot, operazione triviale in ambito outdoor, dove per localizzare i dispositivi è sufficiente un segna-

le di posizionamento satellitare per ottenere dei risultati precisi, dovendo però nel nostro caso utilizzare tale sistema all'interno di edifici, dove un segnale GPS non sarebbe in grado di fornire una precisione sufficiente rende indispensabile esplorare approcci alternativi.

La visione artificiale rappresenta una soluzione economica e poco complessa a tecnologie ben più articolate e soprattutto costose in materia di sistemi di posizionamento indoor. L'obiettivo di questa tesi è quello di approfondire l'utilizzo di tecniche di visione artificiale con l'aiuto di librerie software come OpenCV, con il fine ultimo di ottenere un applicativo in grado di rilevare e localizzare i dispositivi in una determinata scena, così da fornire ai ricercatori dell'università di Bologna un sistema finito da poter integrare all'interno del loro progetto per renderlo utilizzabile anche in contesti indoor.

Di seguito viene fornito il link al sistema sviluppato
<https://doi.org/10.5281/zenodo.14039135>

Struttura della tesi Nel primo capitolo (Background) vengono affrontati i temi utili a comprendere il contenuto di questa dissertazione, andando a definire i concetti principali di visione artificiale e aggregate computing ma anche le singole tecnologie utilizzate in merito a marker fiduciari, calibrazione della camera e OpenCV.

Seguono poi il secondo capitolo (Analisi) dove vengono presentati i requisiti che il sistema finale dovrà soddisfare e il terzo capitolo (Design), dove vengono analizzate le scelte intraprese e l'architettura del software, prima nella sua versione generale poi nel dettaglio.

Il quarto capitolo cioè quello dedicato all'Implementazione riporta tutti i passaggi effettuati per raggiungere l'obiettivo prefissato da questa tesi, analizzando le principali scelte implementative e mostrando al lettore il software degno di nota prodotto.

Il quinto capitolo (Valutazione) mostra una serie di test effettuati sul software durante e al termine del suo sviluppo, che consentono di analizzare i risultati ottenuti, andando anche a valutare se i requisiti posti in fase di analisi siano stati correttamente soddisfatti.

Nel sesto ed ultimo capitolo (Conclusioni) vengono riportate alcune considerazioni sul lavoro svolto e su come esso possa essere ulteriormente sviluppato in futuro, ponendo particolare enfasi sui risultati ottenuti e mostrando un esempio operativo del sistema finale.

Capitolo 2

Background

2.1 Visione artificiale

2.1.1 Cos'è la visione artificiale?

Quando parliamo di visione artificiale o computer vision stiamo considerando un insieme di processi e tecniche che hanno come scopo finale quello di trasformare degli input (solitamente foto o video¹) in una serie di informazioni utili al calcolatore, che possono poi essere utilizzate per prendere decisioni in maniera autonoma, analizzare una situazione o addirittura creare una rappresentazione del mondo reale 3D che ci circonda [5] [12].

Quanto sopra descritto non è troppo diverso da ciò che i nostri occhi fanno tutti i giorni, ed infatti, la visione artificiale nasce proprio per permettere al calcolatore di “vedere” esattamente come un essere umano, in modo da poter interagire con l’ambiente circostante.

¹gli input potrebbero essere anche generati da scanner, sensori LiDaR, radar ecc.

2.1.2 Come funziona la visione artificiale?

Il processo di visione artificiale può essere suddiviso in tre fasi principali:

1. acquisizione di un'immagine;
2. interpretazione e analisi dell'immagine;
3. richiesta di informazioni sull'immagine analizzata.

Nella fase di rilevazione come sopra citato è possibile utilizzare diversi tipi di strumenti come fotocamere o videocamere, ma è nella fase centrale che il processo può differire maggiormente. Infatti, l'interpretazione dell'immagine viene effettuata secondo algoritmi che possono essere anche molto diversi in base al loro scopo, negli ultimi anni inoltre si stanno facendo largo tecnologie innovative come l'intelligenza artificiale, il machine learning e il deep learning per poter intraprendere decisioni e svolgere compiti in modo autonomo senza il bisogno dell'intervento umano.

2.1.3 Applicazioni e finalità

Di seguito sono elencate alcune applicazioni della visione artificiale:

- **classificazione di immagini:** la classificazione di immagini consiste nell'associare un'etichetta o una classe all'immagine in input. Un esempio pratico è un sistema in grado di classificare i vari animali di uno zoo;
- **rilevazione di oggetti:** la rilevazione di oggetti è il processo di individuazione e localizzazione di oggetti all'interno delle immagini. Ad esempio un sistema di rilevazione può essere addestrato a individuare e seguire tutte le persone nel campo visivo di una telecamera;
- **suddivisione di immagini in sezioni da analizzare:** il processo di segmentazione mira a suddividere un'immagine in gruppi di pixel così da poter assegnare un'etichetta o classe ad ogni gruppo. Grazie a questa tecnica, pixel con la stessa etichetta condivideranno caratteristiche simili. Un'applicazione tipica è riscontrabile in ambito medico nella rilevazione di masse tumorali nelle immagini diagnostiche;

2.1. VISIONE ARTIFICIALE

- **riconoscimento facciale:** l'obiettivo del riconoscimento facciale è quello di individuare un volto in un'immagine. vi sono numerosi esempi di utilizzo di tale tecnologia, soprattutto per quanto riguarda sistemi di sicurezza e sorveglianza o per l'accesso a contenuti privati;
- **ricostruzione di ambienti 3D:** grazie all'utilizzo di più camere e/o all'acquisizione di una serie di immagini è possibile effettuare stime di strutture 3D. Queste tecniche possono essere utilizzate da robot in grado di muoversi per rilevare eventuali ostacoli;
- **guida autonoma:** le tecniche di visione artificiale sono largamente utilizzate in veicoli a guida autonoma per classificare gli oggetti circostanti e rilevare altre auto o pedoni.

La lista dei possibili utilizzi è ovviamente molto vasta ma quelli riportati sopra sono tra i più gettonati sia in ambito professionale che di ricerca.

2.1.4 Principali criticità

Tutte le operazioni che caratterizzano un sistema di visione artificiale possono essere largamente influenzate da una serie di condizioni interne o esterne con il risultato che il nostro sistema potrebbe non operare come previsto.

Un esempio di condizioni esterne che possono influenzare il comportamento del nostro sistema sono sicuramente l'illuminazione, la prospettiva ed eventuali occlusioni dell'immagine in input, che possono portare ad una maggiore difficoltà di rilevazione e riconoscimento, per quanto riguarda invece i parametri interni possiamo considerare la risoluzione della camera che stiamo utilizzando, l'algoritmo di elaborazione e la complessità (in termini di numero di pixel da elaborare) dell'immagine ottenuta, come parametri che possono variare anche di molto la velocità e la precisione del nostro sistema.

2.1.5 Cenni storici

I primi articoli prodotti riguardanti la visione artificiale risalgono agli anni '60 dove però l'idea di poter acquisire immagini ed elaborarle, facendone comprendere il

contenuto all’elaboratore era ancora troppo precoce per l’hardware a disposizione. Solo intorno agli anni ‘80 si sono iniziati a vedere i primi sviluppi significativi grazie all’introduzione della trasformata di Hough² e dei primi algoritmi di riconoscimento ottico dei caratteri optical character recognition (OCR). Dagli anni ‘90 sino ai primi anni ‘00 l’attenzione si è spostata sullo sviluppo di algoritmi di machine learning, questo ha permesso nel 2001 di sviluppare il primo algoritmo di riconoscimento facciale [16].

Ad oggi la visione artificiale adotta tecniche e processi completamente differenti rispetto a quelli visti nei suoi primi anni di sviluppo, facendo largo uso di reti convoluzionali e dell’intelligenza artificiale (ormai ampiamente utilizzabile data la sempre crescente potenza di calcolo dei dispositivi) che le permettono non solo di essere più veloce ma anche di garantire una precisione dei risultati molto maggiore grazie anche al vasto numero di dati a disposizione.

2.2 Marker fiduciari

I marker fiduciari sono degli oggetti che, posti all’interno del campo visivo di una fotocamera possono essere utilizzati come punti di riferimento.

Gli scopi principali di questi marker comprendono la calibrazione della camera, la localizzazione, il tracking e la rilevazione di oggetti.

Alcuni possibili tipi di marker fiduciari sono riportati nella figura 2.1.

2.2.1 Applicazioni

- **fisica:** per ottenere posizioni e riferimenti di oggetti;
- **realità aumentata:** utilizzo dei marker come “ancore” così da sapere dove posizionare gli elementi virtuali nel mondo reale;
- **circuiti stampati:** identificano dei pattern così che i macchinari possano operare sui circuiti in maniera autonoma selezionando i componenti univocamente.

²https://en.wikipedia.org/wiki/Hough_transform

2.2. MARKER FIDUCIARI

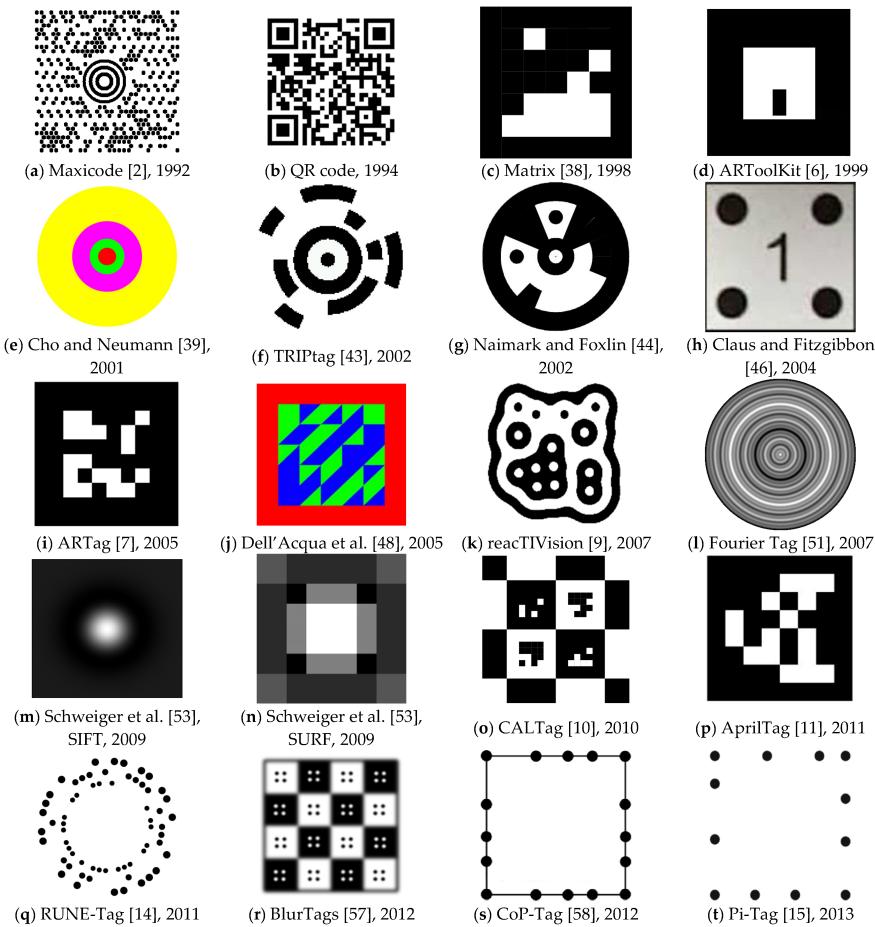


Figura 2.1: Marker fiduciari [10]

Vedi video che riporta l'utilizzo di marker fiduciari per un'applicazione di realtà aumentata³.

2.2.2 ArUco markers

I marker Augmented Reality University of Cordova (ArUco) sono una tipologia di marker fiduciari binari molto utilizzata in ambiente di visione artificiale.

La caratteristica principale di questo tipo di marker risiede nel fatto di essere estremamente versatile, dove, con il termine versatile si vuole sottolineare il fatto che questa tipologia di marker non necessita di particolari condizioni di luce ne di

³<https://youtu.be/nsu9tNIJ6F0?si=MB4zTZP5yiC09aW7>

grandi capacità di elaborazione per essere processata, in quanto, la complessità di ogni marker può variare in base alle esigenze (vedi figura 2.2).

Un altro aspetto fondamentale dei marker ArUco è dato dal fatto che, attraverso sistemi di posa della camera, utilizzata per ottenere le posizioni dei marker, è possibile calcolare anche la loro rotazione.

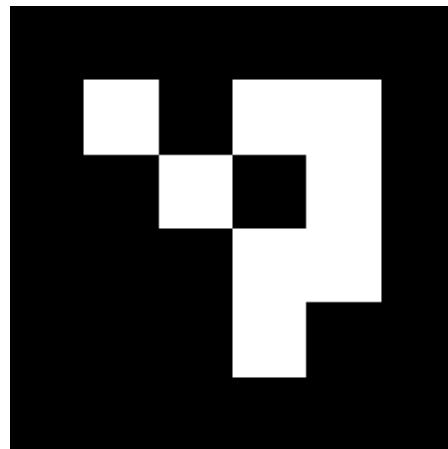


Figura 2.2: ArUco Marker 4x4 id: 0

Come funzionano?

All'interno del marker abbiamo dei quadrati bianchi e neri che, in base alla loro disposizione identificano univocamente il marker stesso, il numero di questi quadrati e la loro disposizione può variare in base al “dizionario” selezionato.

Dizionario

Ogni ArUco marker appartiene ad un dizionario il quale ne specifica la dimensione e la disposizione interna dei quadrati bianchi e neri. Modificando il dizionario possiamo ottenere marker più grandi e complessi che consentono rilevazioni più precise, riducendo la possibilità di confondere tra loro due marker simili, d'altro canto utilizzando marker più piccoli e semplici, nonostante si vada incontro ad una riduzione della precisione, si guadagna in velocità di elaborazione [13].

DICT_4X4_100⁴

⁴dizionario con marker 4x4 e con una quantità massima di marker istanziabili pari a 100.

2.3 OpenCV

Open Source Computer Vision Library (OpenCV) [14] è una libreria realizzata inizialmente da Intel pensata per fornire funzioni di visione artificiale che facilitassero il compito degli sviluppatori, garantendo loro un'infrastruttura stabile sulla quale poter lavorare 2.3.

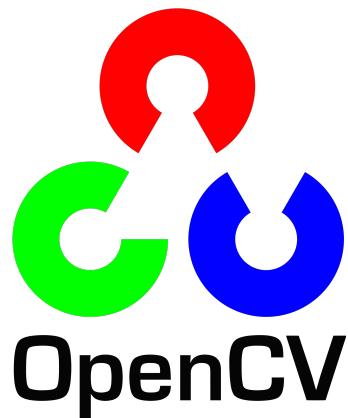


Figura 2.3: Logo OpenCV

OpenCV è una libreria multi-piattaforma rilasciata su licenza open-source, questo ha portato durante gli anni ad una crescita continua grazie anche all'aiuto della community, che ha permesso alla libreria in questione di diventare una tra le più importanti nel suo ambito [4] [15].

2.3.1 JavaCV

JavaCV è un wrapper per la libreria OpenCV, in grado di fornire, oltre a tutte le funzionalità di OpenCV, anche molti metodi per processare immagini e video, questo fa sì che con una singola dipendenza si sia in grado di utilizzare tutte le funzionalità in maniera agevolata per gli sviluppatori Java attraverso il wrapper, senza dover interagire con il codice nativo di OpenCV [3].

2.4 Aggregate Computing

Prima di poter affrontare il concetto di aggregate computing è necessario introdurre brevemente i sistemi distribuiti, costituendo questi ultimi il fondamento concettuale su cui l'approccio dell'aggregate computing si basa.

2.4.1 Sistema distribuito

Un sistema distribuito rappresenta una rete di processi o “nodi di elaborazione” che svolgono calcoli individualmente ma che appaiono dall'esterno come un unico sistema.

Questo tipo di approccio è caratterizzato da una grande eterogeneità sia a livello hardware che software, la quale porta il sistema ad essere aperto a tecnologie diverse, che comunicano tra loro attraverso lo scambio di messaggi per far apparire il sistema distribuito come un'entità unica.

L'aspetto fondamentale di questi sistemi riguarda il fatto che ad ogni singolo nodo⁵ viene affidato un compito da espletare, il nodo in questione può comunicare con gli altri al fine di portare a termine il suo incarico e può essere a sua volta interrogato dagli altri per fornire i risultati della sua elaborazione. Tutto questo può essere visto dall'esterno come un unico sistema che però al suo interno è composto da una serie di dispositivi con compiti differenti che contribuiscono al raggiungimento di uno scopo comune 2.4.

2.4.2 Aggregate computing

Il termine Aggregate computing rappresenta un approccio emergente nell'ambito della gestione di sistemi distribuiti complessi [17] che sposta l'attenzione dal calcolo effettuato da un singolo nodo al far emergere un comportamento globale del sistema sulla base dei calcoli effettuati sui singoli nodi.

A primo avviso, la differenza tra i due sembra marginale, ma, se scendiamo più nel dettaglio, possiamo notare alcune differenze: un sistema distribuito focalizza la propria attenzione sulla suddivisione dei compiti ai nodi, che possono comunicare tra loro per ottenere un fine comune. Un sistema di aggregate computing pone

⁵ dispositivo visto come elemento di una rete.

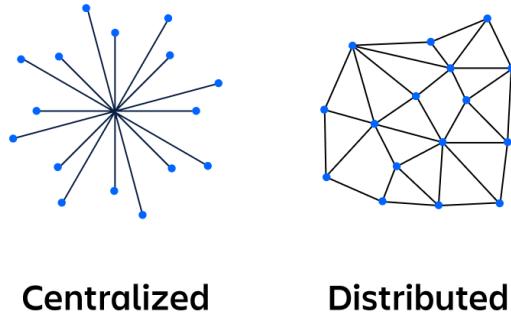


Figura 2.4: Sistema centralizzato Vs distribuito [18]

maggior enfasi sullo scambio di informazioni tra nodi vicini, i quali, comunicando tra loro, fanno emergere una serie di comportamenti complessi che descrivono il sistema nella sua totalità. Nel contesto di aggregate computing un nodo non è in grado di comunicare con tutti i nodi, ma solamente con quelli a lui vicini, attraverso i quali può progredire nel suo compito e derivare informazioni sulla rete globale. Questo suggerisce quindi, che il comportamento generale del sistema visto dall'esterno emerge da interazioni prettamente locali al suo interno.

Vantaggi

I vantaggi offerti dall'approccio dell'aggregate computing sono molteplici e spesso dipendono proprio dalla modalità operativa del sistema, in questo caso quindi dalla possibilità di derivare comportamenti emergenti sulla base di interazioni locali:

- **scalabilità:** l'approccio dell'aggregate computing è caratterizzato da interazioni locali, non vi è un sistema centrale che coordina tutti i nodi. Nel caso quindi si necessiti di un numero maggiore di nodi, sarà sufficiente inserirli nella rete: saranno poi i nodi stessi a comunicare con i propri vicini per contribuire al sistema di aggregate computing;
- **tolleranza ai guasti:** in un sistema di aggregate computing il calcolo è distribuito sui vari nodi della rete, di conseguenza, la rimozione di alcuni di essi non compromette il comportamento generale del sistema;

- **auto-organizzazione:** grazie alle interazioni locali, un sistema di aggregate computing è in grado di riorganizzarsi in base al numero di nodi a disposizione, mantenendo comunque il comportamento globale.

Applicazioni

Le principali applicazioni di sistemi di aggregate computing possono essere ricercate oggi in ambiti quali:

- **reti di sensori:** un insieme di sensori che vengono utilizzati per monitorare una certa area geografica della quale vogliamo estrarre informazioni in maniera aggregata quali temperatura pressione atmosferica o precipitazioni [2];
- **robotica collettiva:** per robotica collettiva si intende tutta quella serie di applicazioni che consentono a robot e droni di auto-organizzarsi e coordinarsi secondo sciami o schemi predefiniti [1];
- **smart cities:** con il termine smart cities possiamo riassumere le operazioni volte a gestire l'infrastruttura cittadina come l'illuminazione, il controllo di semafori e del traffico ed il monitoraggio della folla in luoghi pubblici [6];
- **sistemi IoT:** con sistemi IoT si intendono una serie di dispositivi eterogenei, connessi tra loro, che interagiscono per fornire servizi a livello globale [7].

2.5 Calibrazione della camera

Di seguito sono riportate alcune informazioni di base per comprendere il concetto di calibrazione della camera

2.5.1 Cos'è la calibrazione?

La calibrazione è un processo volto a ottenere i parametri della lente e del sensore della videocamera/fotocamera, tali parametri potranno poi essere impiegati per correggere la distorsione della lente, calcolare distanze, ricostruire l'immagine 3D di partenza ecc [11].

2.5.2 Modello di calibrazione

Esistono diversi modelli di calibrazione in base al tipo di camera che si intende utilizzare, nel nostro caso possiamo ipotizzare l'utilizzo di uno dei modelli più semplici e cioè quello di una fotocamera con obiettivo stenopeico 2.5, nella quale però teniamo conto anche della distorsione della lente [9] [8].

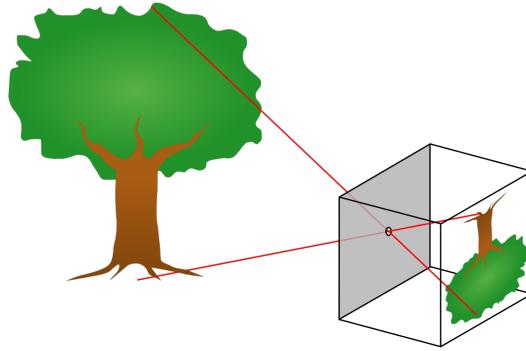


Figura 2.5: Rappresentazione camera con obiettivo stenopeico

Il parametro più importante è la matrice della camera “camera matrix” che mappa le coordinate del mondo 3D nell’immagine 2D catturata, questa matrice è ottenuta a partire dai parametri intrinseci e estrinseci:

- **parametri estrinseci:** rappresentano la posizione e rotazione della camera nel mondo reale;
- **parametri intrinseci:** consentono di trasformare le coordinate della camera 3D nelle coordinate dell’immagine 2D.

L’obiettivo della calibrazione è proprio quello di determinare questi due parametri così da poter calcolare la matrice della camera, va però precisato che solo nel caso in cui si debbano utilizzare più camere è necessario includere i parametri estrinseci nella matrice, questo perché le diverse camere devono conoscere le rispettive posizioni, in caso contrario (in presenza di una sola camera) la matrice sarà definita dai soli parametri intrinseci in quanto sarà presente un solo sistema di riferimento (**come nel caso in esame**) 2.6.

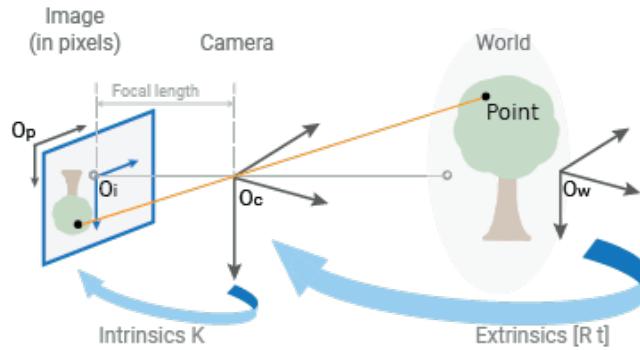


Figura 2.6: Processo di conversione dal mondo reale 3D all’immagine 2D

2.5.3 Distorsione della lente

Un altro aspetto fondamentale da tenere in considerazione, riguarda la distorsione introdotta dalla lente della camera, nel modello utilizzato infatti si trascura l’uso della lente ma in situazioni reali l’obiettivo gioca un ruolo cruciale, poiché introduce due tipi di distorsione:

- **distorsione radiale:** condizione per la quale i raggi di luce vengono curvati in maniera non uniforme, deformando l’immagine finale a partire dal centro dell’obiettivo 2.7;

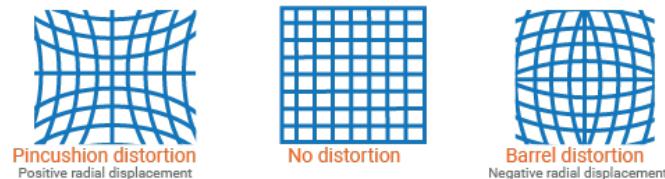


Figura 2.7: Distorsione radiale

- **distorsione tangenziale:** distorsione che ha luogo nel caso in cui la lente e il piano che si vuole fotografare non sono paralleli 2.8.

2.5. CALIBRAZIONE DELLA CAMERA

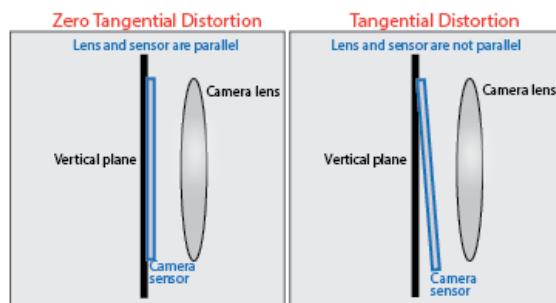


Figura 2.8: Distorsione tangenziale

Capitolo 3

Analisi

Il progetto si pone l’obiettivo di creare un sistema software sviluppato in java o altro linguaggio che sia in grado di essere eseguito sulla Java Virtual Machine (JVM), con il compito di localizzare e tracciare gli spostamenti di alcuni dispositivi¹ nel campo visivo di una videocamera.

Per poter analizzare correttamente il sistema in questione è necessario però fare un passo indietro, per introdurre il contesto nel quale dovrà inserirsi:

l’argomento di questa tesi rappresenta solo un componente di un progetto più complesso, sviluppato dai ricercatori dell’università di Bologna, con l’obiettivo di gestire una serie di dispositivi (robot, droni ecc.), connessi tra loro, secondo il paradigma dell’aggregate computing. Questo approccio permette ai dispositivi della rete di auto-organizzarsi presentando le caratteristiche tipiche della robotica degli sciami.

Il vincolo riguardante il linguaggio da utilizzare, viene dal fatto che il software attualmente usato dai ricercatori è eseguito dalla JVM e di conseguenza per avere un’interazione più semplice tra i due componenti è importante che anche la parte di localizzazione sia scritta con un linguaggio compatibile.

L’aspetto che ha portato alla stesura di questa tesi, si deve ricercare nel fatto che questo tipo di sistema deve poter funzionare anche in applicazioni indoor, nelle quali, l’utilizzo del GPS per tenere traccia degli spostamenti dei dispositivi

¹oggetti di ogni genere e tipo dei quali si vuole conoscere posizione e rotazione chiamati più volte dispositivi in questa dissertazione.

3.1. REQUISITI FUNZIONALI

è ostacolato da numerosi fattori quali pareti e strutture che possono ridurne il segnale.

3.1 Requisiti funzionali

RF1 il sistema deve poter rilevare i dispositivi;

RF2 il sistema deve poter calcolare distanza, posizione e angolo di rotazione dei dispositivi rispetto ad un punto fissato nello spazio;

RF3 il sistema dovrà poter essere interrogato per ottenere i risultati computati riguardo al posizionamento dei dispositivi;

RF4 ogni dispositivo deve poter essere identificato univocamente.

3.2 Requisiti non funzionali

RNF1 il sistema deve poter rilevare i dispositivi ad una distanza di due/tre metri;

RNF2 il sistema deve essere sufficientemente reattivo da rilevare e tracciare movimenti repentini da parte dei dispositivi, con velocità nell'ordine di 0,5 m/s;

RNF3 il sistema deve rimanere stabile ed affidabile anche in presenza di condizioni di scarsa illuminazione;

RNF4 il sistema deve essere efficiente al punto che il tempo impiegato per stimare la posizione dei dispositivi non influenzi il tempo di esecuzione del software in cui è utilizzato;

RNF5 il sistema deve essere sufficientemente preciso da garantire un corretto funzionamento generale, senza incorrere in rilevazioni fuorvianti che porterebbero al crash dello stesso.

3.3 Requisiti tecnologici

RT1 il sistema deve essere scritto in java o un linguaggio che sia in grado di essere eseguito dalla JVM;

RT2 la videocamera utilizzata per la rilevazione dei dispositivi deve avere una risoluzione minima FullHD².

3.4 Analisi del dominio

Come già sopracitato, l'oggetto di questa tesi andrà a costituire parte di un sistema più ampio, il quale, una volta assemblato, sarà in grado di:

1. **rilevare e tracciare gli spostamenti dei dispositivi:** in questa fase il progetto sviluppato nella presente dissertazione si occuperà di ottenere un flusso di immagini della scena in cui i dispositivi possono muoversi. Grazie a tali immagini sarà in grado di rilevare i nodi e salvare le informazioni sulla loro posizione e rotazione per le successive fasi;
2. **inviare ai vari nodi le informazioni acquisite durante la fase di rilevazione:** da questo punto il controllo passa al sistema creato dai ricercatori dell'università di Bologna ed in particolare i dati raccolti sul posizionamento vengono inviati ai vari nodi senza essere elaborati in maniera centralizzata;
3. **far comunicare tra loro i nodi vicini:** in questa fase i nodi scambieranno informazioni con i propri vicini secondo le regole descritte dal programma aggregato;
4. **far organizzare in maniera autonoma i dispositivi:** grazie alle informazioni raccolte e alle interazioni locali dei singoli dispositivi, questi ultimi saranno in grado di organizzarsi in maniera autonoma secondo algoritmi di aggregate computing.

Vedi diagramma esemplificativo 3.1

²risoluzione di 1920 x 1080 pixel.

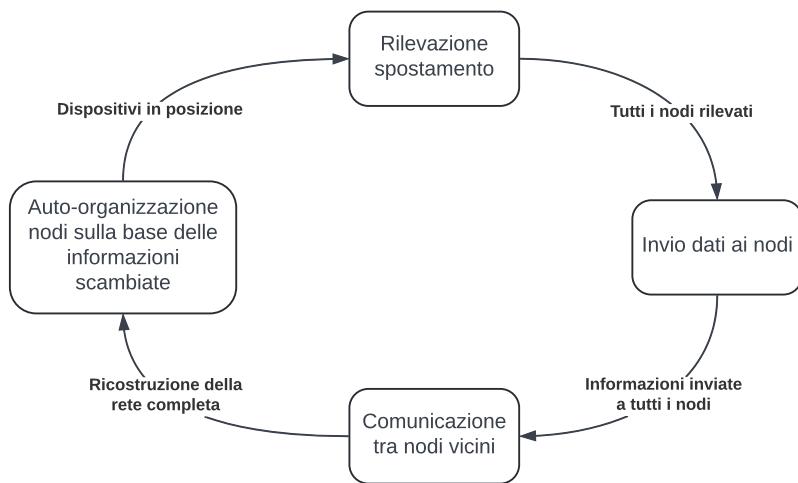


Figura 3.1: Diagramma a stati semplificato del dominio

Capitolo 4

Design

L'ambito di questa tesi è focalizzato sullo sviluppo della componente di visione artificiale utile a rilevare e seguire i dispositivi nel campo visivo della videocamera.

Nelle seguenti sezioni vengono illustrate le scelte, operate in osservanza dei requisiti enunciati nelle sezioni 3.1 e 3.2 che costituiranno poi l'implementazione del software oggetto della tesi.

Prima però di poter approfondire l'architettura del sistema è necessario stabilire come rilevare e identificare quelli che finora abbiamo genericamente chiamato “dispositivi”.

4.1 ArUco markers

Dopo un'attenta ricerca volta a determinare quale fosse il metodo più efficace per poter identificare un qualsiasi oggetto nel campo visivo di una videocamera mi sono imbattuto nei marker ArUco 2.2.2, una tipologia di marker fiduciari 2.2 particolarmente utilizzati in contesti di visione artificiale.

La motivazione che mi ha spinto ad utilizzare i marker ArUco deriva dai numerosi vantaggi che questa tecnologia offre. I marker sono estremamente semplici e scalari, è infatti possibile scegliere fra numerose versioni che si differenziano in base al dizionario utilizzato 2.2.2, inoltre i marker ArUco garantiscono una buona affidabilità anche in condizioni di scarsa illuminazione, scenario da non sottovallutare in una applicazione di visione artificiale. Un altro vantaggio è sicuramente

dato dalla loro velocità di rilevazione che permette quindi l'uso di un gran numero di marker sulla scena senza rallentamenti, infine ogni marker restituisce un ID univoco, utile ad identificare i diversi dispositivi come richiesto da requisito RF4.

Per poter utilizzare i marker ArUco è stata utilizzata la libreria offerta da OpenCV 2.3.

4.2 Videocamera

Per poter rispettare i requisiti imposti in fase di analisi, riguardanti la videocamera da utilizzare per ottenere il flusso video, la scelta è ricaduta su una webcam FullHD capace di registrare a 30/60fps. Tale scelta permette di contenere i costi, garantire una risoluzione sufficiente al riconoscimento degli oggetti e avere un campo visivo ampio.

4.3 Architettura

Premessa: *L'architettura riguarderà il solo sistema di visione artificiale, realizzato da me come software stand-alone che è poi stato rielaborato dai ricercatori per poterlo integrare con il loro progetto, non verrà trattata l'architettura del sistema finale se non attraverso alcuni cenni utili a comprendere le scelte intraprese.*

I diagrammi che seguono descrivono l'architettura del sistema, esclusi quelli delle classi tutti gli altri diagrammi riportano solo i metodi principali e non sono quindi esaustivi.

Finora sono state presentate soluzioni che riguardano la videocamera e l'utilizzo di marker ArUco per rilevare e identificare gli oggetti nella scena¹, dal punto di vista del software invece per ottenere risultati conformi ai requisiti e di conseguenza un sistema capace di operare anche in condizioni di scarsa luminosità, con un buon livello di precisione e a distanze considerevoli, si è reso necessario l'impiego di sistemi di calibrazione della videocamera oltre che ovviamente a funzionalità per determinare la posizione degli oggetti.

¹campo visivo della videocamera.

Il software da me realizzato si compone di tre parti principali descritte da tre classi:

- App;
- CameraCalibrator;
- CameraPose.

`App` si occupa di orchestrare le diverse fasi di avvio dell'applicazione, consentendo in una fase iniziale di calibrare la camera attualmente in uso attraverso la classe `CameraCalibrator` e in un secondo momento di calcolare la posa della camera rispetto ai marker sulla scena con la classe `CameraPose` 4.1.

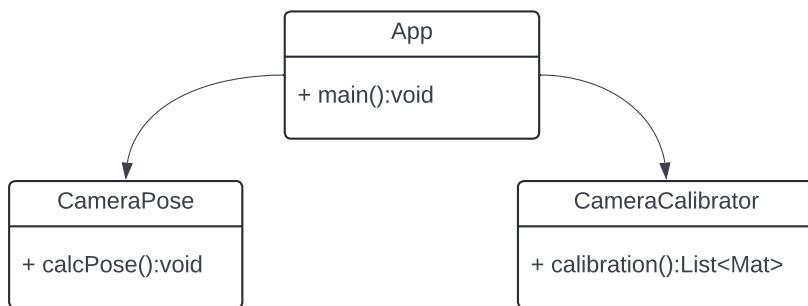


Figura 4.1: Architettura del sistema

4.4 Design dettagliato

Entrando più nel dettaglio dell'architettura, il sistema si compone anche di altre due classi denominate: `ResolutionEnum` e `InputParameters` andando quindi a definire il sistema nella sua totalità come in figura 4.2, dove le due classi sono rispettivamente una enumerazione che racchiude tutte² le risoluzioni che la camera può utilizzare ed una Graphical User Interface (GUI) che permette all'utente di configurare il software in base alle proprie preferenze in fase di avvio.

²sono riportate solo le risoluzioni più comuni.

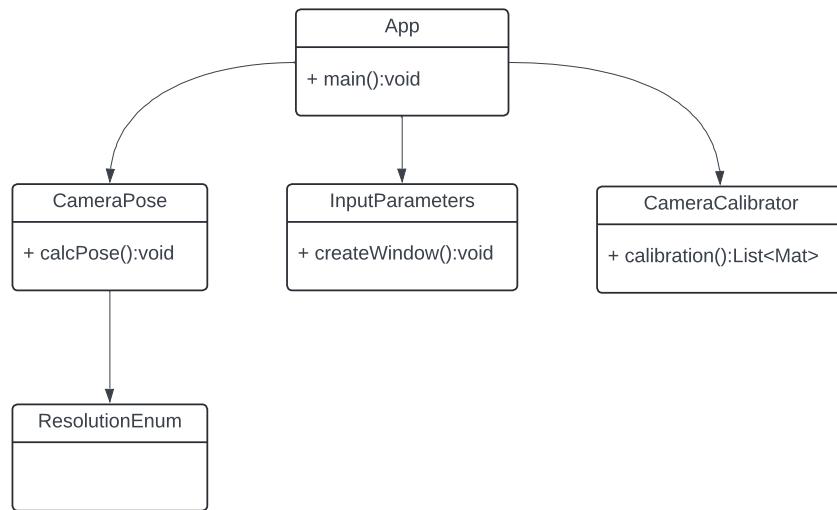


Figura 4.2: Architettura completa

4.4.1 App

Questa classe rappresenta il punto di ingresso dell'applicazione e si occupa di istanziare tutte le classi necessarie per il corretto funzionamento del sistema all'interno del main 4.3.

Il metodo `convertDictionary(String dictionaryType)` è un metodo di utility necessario a convertire una stringa che rappresenta un dizionario di marker, in un intero valido e significativo per il sistema.

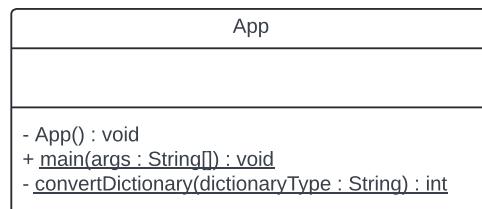


Figura 4.3: Classe App

4.4.2 CameraCalibrator

La classe `CameraCalibrator` ricopre un ruolo molto importante, il suo obiettivo è quello di calibrare la camera 2.5 affinché le immagini ottenute consentano di stabilire correttamente la posizione della stessa rispetto agli oggetti nella scena 4.4.

Il metodo `calibration()` si occupa di calcolare e restituire i parametri della camera, utilizzando le immagini di calibrazione fornite.

Il metodo `getImageFiles(String directoryPath)` viene utilizzato all'interno di `calibration()` per ottenere tutti i percorsi assoluti delle immagini di calibrazione.

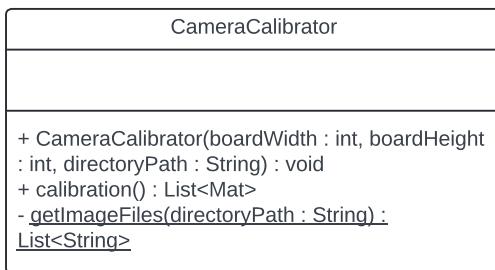


Figura 4.4: Classe CameraCalibrator

4.4.3 CameraPose

La classe `CameraPose` rappresenta il fulcro del progetto, grazie ai metodi sviluppati è possibile ottenere la posa della camera rispetto agli oggetti identificati dai marker ArUco. Le informazioni ottenute riguardano distanza, posizione, rotazione e identità dei marker, inoltre è anche possibile visualizzarne graficamente lo spostamento e l'orientamento 4.5.

Di seguito viene presentato l'elenco di operazioni che la classe deve eseguire per calcolare la posa e successivamente vengono analizzati i singoli metodi:

1. cattura del frame 5.5.1;
2. rilevazione marker 5.5.2;

3. inizio posa;

- (a) calcolo posa 5.5.3;
- (b) trascrizione posizione e rotazione di ogni marker sul frame;
- (c) calcolo errore di riproiezione.

4. restituzione frame aumentato³.

- **getArucoDetector()**: metodo che crea, modifica e restituisce il componente in grado di rilevare i marker.
- **getCamera()**: metodo che personalizza e restituisce l'istanza di una camera connessa all'elaboratore.
- **getCanvas(String title)**: metodo che crea e restituisce una finestra visualizzabile dall'utente.
- **calcSinglePose(VideoCapture cap)**: metodo in grado di calcolare e restituire la posa della camera per un singolo frame. Il frame viene ottenuto dal parametro cap passato al metodo per ridurre il tempo di apertura di una nuova camera.
- **calcSinglePose()**: metodo in grado di calcolare e restituire la posa della camera per un singolo frame. In questo caso ad ogni chiamata del metodo viene creata una nuova un'istanza della camera.
- **calcPose()**: metodo che calcola e visualizza all'interno di una finestra la posa della camera per un flusso video proveniente dalla stessa.
- **prepareMatrixForPoseEstimation(List<Mat>corners, int i, MatOfPoint2f cornerPoints)**: metodo che converte la matrice degli angoli dei marker in un tipo utilizzabile per calcolare la posa della camera.

³un frame aumentato è un frame al quale sono stati sovrapposti elementi attraverso operazioni di visione artificiale.

- **drawAxis(Mat image, Mat rvec, Mat tvec, float length)**: metodo che permette di aggiungere informazioni sui marker (posizione e rotazione) nella finestra che visualizza la posa della camera.
- **calculateReprojectionError(MatOfPoint3f objPoints, MatOfPoint2f imgPoints, Mat rvec, Mat tvec)**: metodo utilizzato per calcolare l'errore di riproiezione durante la fase di posa della camera
- **rescalePoints(List<Mat>corners)**: metodo utilizzato per correggere la posizione degli angoli dei marker nel caso in cui il rilevazione sia avvenuta su una immagine di risoluzione ridotta (rispetto alla risoluzione delle immagini di calibrazione).

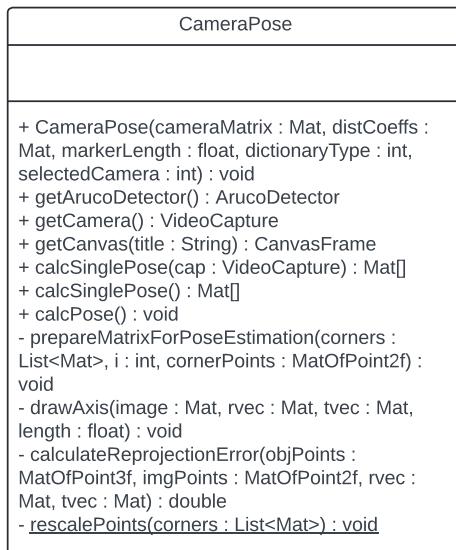


Figura 4.5: Classe CameraPose

4.4.4 ResolutionEnum

Enumerazione in grado di fornire i formati più utilizzati per flussi video 4.6.

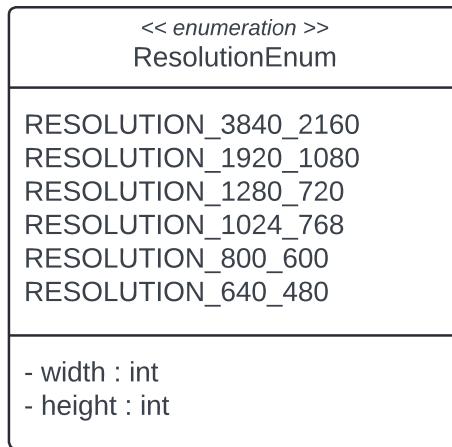


Figura 4.6: Classe ResolutionEnum

4.4.5 InputParameters

La classe `InputParameters` rappresenta una GUI con la quale l'utente finale può interagire, al fine di configurare il programma in base alle proprie specifiche e preferenze 4.7.

Il metodo `createWindow()` crea la finestra attraverso la quale l'utente può personalizzare il sistema al suo avvio.

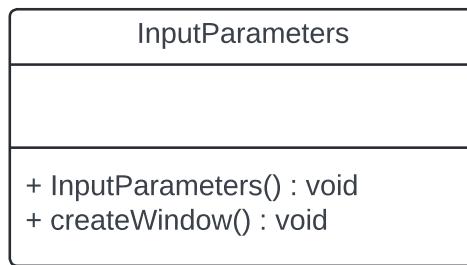


Figura 4.7: Classe InputParameters

4.4.6 Risoluzione video

Nell'immagine 4.8 viene mostrata l'associazione tra la classe `CameraPose` e la classe `ResolutionEnum`, grazie alla quale la prima riesce ad ottenere l'elenco delle possibili risoluzioni video utilizzate dalla camera.



Figura 4.8: Associazione ResolutionEnum

4.4.7 UI

Nell'associazione in figura 4.9 App fa uso della classe `InputParameters` per poter istanziare correttamente i parametri del sistema sulla base delle preferenze dell'utente.



Figura 4.9: Associazione InputParameters

4.4.8 Calcolo della posa

Di seguito viene mostrato un diagramma di sequenza esaustivo, che rappresenta l'intera serie di operazioni svolte dal sistema per calcolare la posa della camera 4.10.

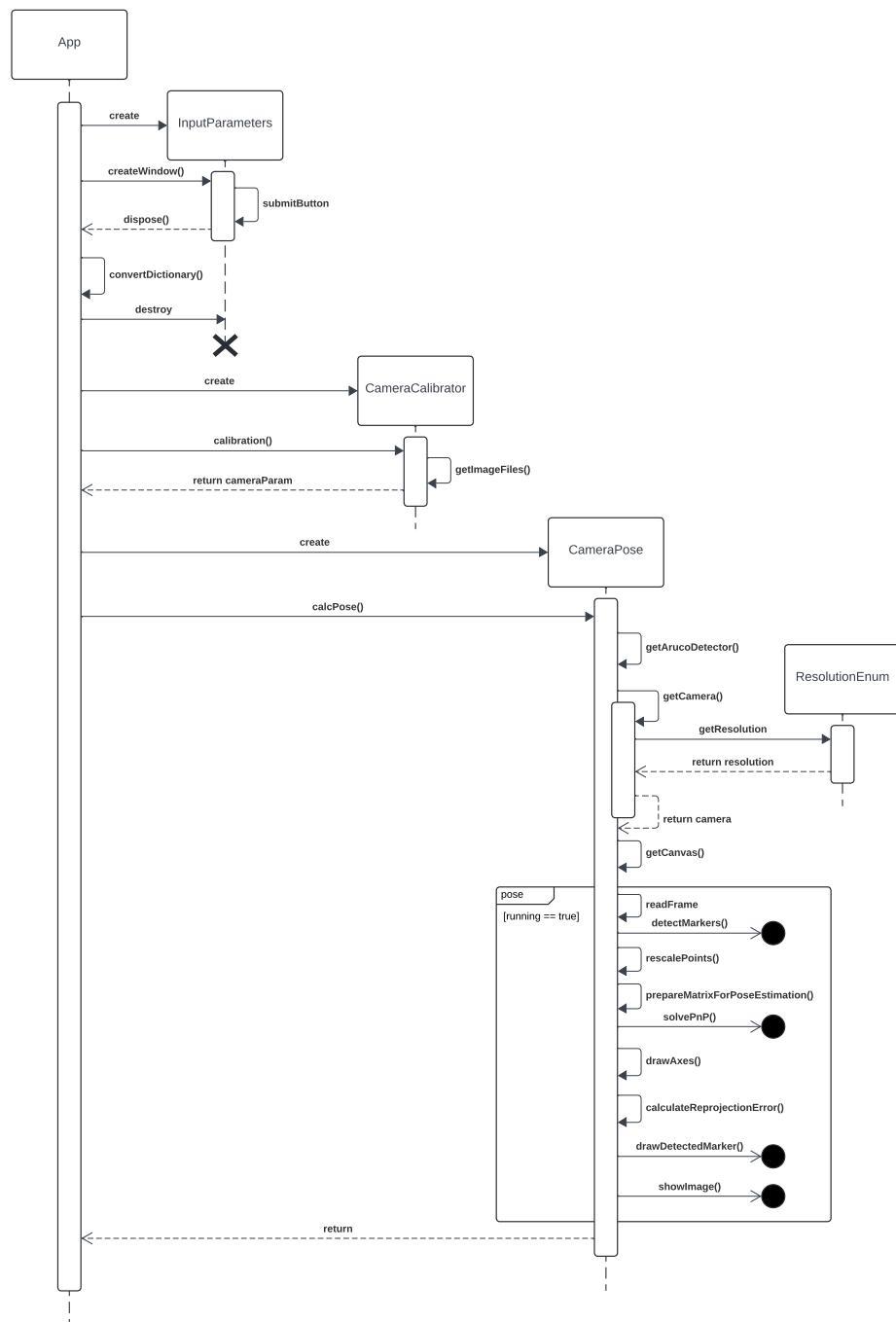


Figura 4.10: Diagramma di sequenza del sistema

Capitolo 5

Implementazione

In questo capitolo verranno presentate alcune scelte implementative effettuate durante lo sviluppo software, con particolare attenzione alle fasi di calibrazione e posa della camera.

5.1 Creazione del progetto

Il software è stato sviluppato all'interno di un progetto Java utilizzando Gradle con uno script di build in Kotlin.

5.2 JavaCV

Durante l'implementazione è stato usato il wrapper di JavaCV 2.3.1 al posto di OpenCV nativo.

La dipendenza necessaria per utilizzare JavaCV è stata inserita nel file build.gradle.kts listing 5.1

Listing 5.1: Implementazione dipendenza JavaCV

```
1 implementation ("org/bytedeco/javacv-platform:1.5.10")
```

5.3 Avvio dell'app

Come prima operazione all'avvio dell'applicazione vengono caricate le classi di OpenCV in maniera dinamica listing 5.2, dopodiché `App.java` si occuperà di istanziare e lanciare la classe `InputParameter.java` utile a configurare i parametri dell'app come mostrato dal codice listing 5.3, infine sarà il turno di calibrazione e posa della camera, che, come è possibile notare da listing 5.4 vengono chiamate sequenzialmente al fine di passare i parametri della camera 2.5 alla fase di posa.

Listing 5.2: Caricamento classi OpenCV

```
1 static {
2     Loader.load(opencv_java.class);
3 }
```

Listing 5.3: Configurazione parametri app

```
1 final InputParameters ip = new InputParameters();
2 ip.createWindow();
3
4 // Wait for the user to set the parameters
5 final int squaresX = ip.getSquaresX();
6 final int squaresY = ip.getSquaresY();
7 final float markerLength = ip.getMarkerLength();
8 final String directoryPath = ip.getDirectoryPath();
9 final int selectedCamera = ip.getCameraIndex();
10 final int dictionaryType = convertDictionary(ip.getDictionaryType());
11 // Close the input parameters window
12 ip.close();
```

Listing 5.4: Avvio calibrazione e posa

```
1 //Camera calibration
2 final CameraCalibrator cc = new CameraCalibrator(squaresX, squaresY, directoryPath
    );
3 final List<Mat> cameraParam = cc.calibration();
4
5 //Camera pose estimation
6 final CameraPose cp = new CameraPose(cameraParam.get(0), cameraParam.get(1),
    markerLength, dictionaryType, selectedCamera);
7 cp.calcPose();
```

5.4 Calibrazione

La fase di calibrazione della camera rappresenta un aspetto fondamentale del sistema finale in quanto consente di ottenere risultati precisi e affidabili nelle fasi successive.

Per effettuare la calibrazione OpenCV mette a disposizione varie strategie, è possibile infatti utilizzare una scacchiera classica, oppure una scacchiera ArUco o una semplice tabella di marker ArUco, nel sistema sviluppato si è optato per la prima soluzione, una scacchiera classica infatti offre un'alta precisione in situazioni ideali e semplicità di implementazione, fattori determinanti visto l'utilizzo di immagini pre acquisite ottenute in condizioni ottimali.

Il ruolo principale di questa classe quindi, è quello di restituire alla chiamata del metodo `Calibration()` la matrice della camera e i coefficienti di distorsione della lente, per fare ciò sono stati utilizzati i metodi di JavaCV come mostrato nelle seguenti sezioni.

5.4.1 Rilevazione scacchiera

Nel codice listing 5.5 si è fatto uso del metodo `findChessboardCorners()` per ottenere le posizioni degli angoli interni della scacchiera, sulla base di un'immagine della stessa convertita in bianco e nero per ridurre la complessità computazionale.

Listing 5.5: Rilevazione angoli interni della scacchiera

```
1 final Mat image = Imgcodecs.imread(filePath);
2 final Mat grayImage = new Mat();
3 //Converting to a gray scale image to reduce computational complexity
4 Imgproc.cvtColor(image, grayImage, Imgproc.COLOR_BGR2GRAY);
5
6 final MatOfPoint2f imageCorners = new MatOfPoint2f();
7 //Scanning the image to extract the corners
8 final boolean found = Calib3d.findChessboardCorners(grayImage, boardSize,
    imageCorners);
```

5.4.2 Rifinitura angoli

Si è ritenuto necessario rifinire le posizioni degli angoli della scacchiera per aumentarne ulteriormente la precisione, analizzando l'immagine a livello di sub-pixel¹ con il metodo `cornerSubPix()` listing 5.6.

Listing 5.6: Rifinitura angoli

```
1 //finishing the corners with the cornerSubPix method
2 if (found) {
3     Imgproc.cornerSubPix(grayImage, imageCorners, new Size(WIN_X_SIZE, WIN_Y_SIZE),
4                           new Size(ZERO_X_ZONE, ZERO_Y_ZONE),
5                           new TermCriteria(TermCriteria.EPS + TermCriteria.COUNT, MAX_ITERATION,
6                                             ACCURACY));
7
8     imagePoints.add(imageCorners);
9     objectPoints.add(objectPoint);
}
```

5.4.3 Calibrazione

La calibrazione vera e propria è svolta dal metodo `calibrateCamera()`, che oltre a popolare la matrice della camera e dei coefficienti di distorsione, restituisce l'errore quadratico medio di riproiezione dei punti del mondo reale nell'immagine analizzata listing 5.7.

Listing 5.7: Calibrazione camera

```
1 //calibration of the camera
2 final double rms = Calib3d.calibrateCamera(objectPoints, imagePoints, boardSize,
3                                              cameraMatrix, distCoeffs, rvecs, tvecs);
```

5.5 Posa

La classe `CameraPose.java` si occupa di calcolare la posa della camera rispetto ai marker ArUco e, successivamente, di visualizzare in un ambiente di realtà aumentata la posizione e la rotazione dei marker 5.1.

¹tecnica che permette di analizzare lo spazio vuoto tra i pixel, vedi: <https://www.pomeas.com/download/2021/5/4/what-subpixel/> per maggiori informazioni.

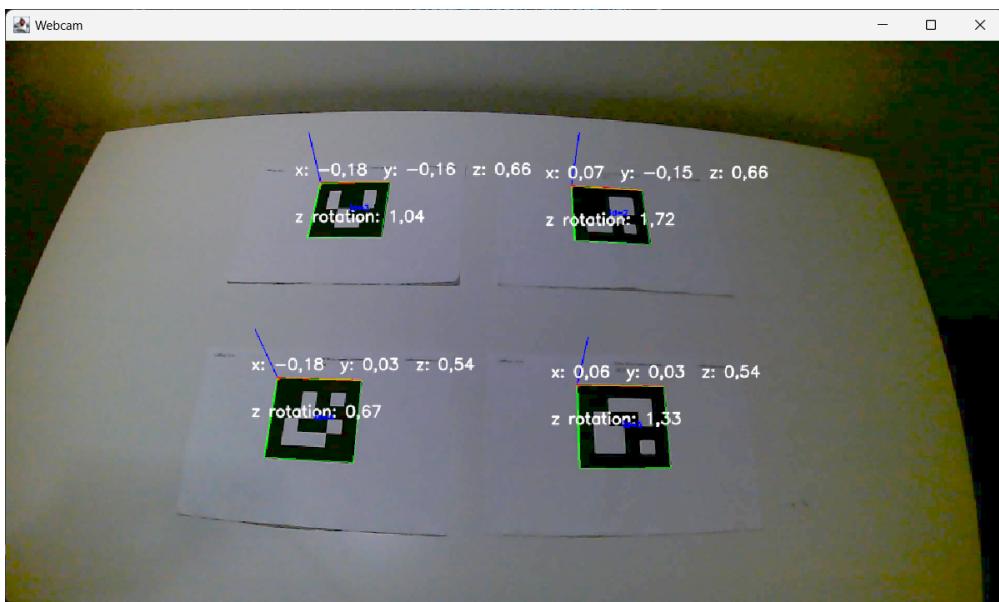


Figura 5.1: Posa della camera

Di seguito viene descritta l'implementazione delle principali azioni svolte dalla classe `CameraPose.java` come descritto in 4.4.3.

5.5.1 Cattura del frame

Durante la fase di cattura del frame sono state operate alcune scelte per garantire il corretto funzionamento del sistema.

Risoluzione

La classe utilizzata per la cattura del frame apriva la camera con una risoluzione standard, la quale però non combaciava con quella delle immagini utilizzate per la calibrazione, questa discrepanza faceva sì che la successiva fase di posa lavorasse su immagini con risoluzioni differenti da quelle attese producendo errori nella stima.

Per risolvere tale problema ho creato l'enumerazione `ResolutionEnum.java` riportata in maniera semplificata² dal seguente codice listing 5.8 in modo da poter

²L'enumerazione presentata al lettore contiene solo alcune risoluzioni tra le quali il sistema può scegliere, la classe utilizzata dal software è più complessa e contiene attributi e metodi per fornire l'altezza e la larghezza di ogni risoluzione.

sempre impostare la risoluzione massima possibile della webcam (questo perché è logico pensare di avere a disposizione immagini per la calibrazione alla maggior risoluzione possibile).

Listing 5.8: ResolutionEnum.java

```
1 public enum ResolutionEnum {
2     RESOLUTION_3840_2160(3840, 2160),
3     RESOLUTION_1920_1080(1920, 1080),
4     RESOLUTION_1280_720(1280, 720),
5     RESOLUTION_1024_768(1024, 768),
6     RESOLUTION_800_600(800, 600),
7     RESOLUTION_640_480(640, 480);
8 }
```

Esposizione

La reattività del sistema in termini di capacità di tracciamento degli oggetti è risultata essere scarsa nei primi test qualitativi effettuati, per risolvere tale problema si è operato sul tempo di esposizione della camera, che ha permesso di incrementare notevolmente la reattività a discapito di una riduzione del range operativo³, rendendo il sistema in grado di rilevare i marker solo in condizioni di luce più elevate.

listing 5.9

È possibile modificare il valore dell'esposizione della camera per ottenere un sistema più reattivo decrementandolo, o per poter operare in condizioni di scarsa luminosità incrementandolo.

Listing 5.9: Regolazione tempo di esposizione

```
1 //Costants for the camera
2 private static final int CAMERA_EXPOSURE = -6;
3
4 //Setting the camera exposure to reduce the motion blur
5 capture.set(Videoio.CAP_PROP_EXPOSURE, CAMERA_EXPOSURE);
```

5.5.2 Rilevazione marker

Durante il calcolo della posa è emerso che gran parte del tempo veniva impiegato nella fase di rilevazione del marker, rallentando l'intero processo. Per risolvere

³intervallo nel quale uno strumento può operare correttamente.

tale problema si è deciso di modificare la dimensione del frame ottenuto prima di rilevare i marker, dividendo altezza e larghezza per un fattore arbitrario listing 5.10, in modo da passare al metodo `detectMarkers()` un’immagine più semplice da elaborare e di conseguenza velocizzare la rilevazione listing 5.11.

Adottando questa soluzione bisogna però fare attenzione a due aspetti:

1. **risoluzione troppo bassa:** dividendo il frame ottenuto per un fattore arbitrario si rischia di ottenere immagini di bassa risoluzione, questo può portare ad una perdita significativa della precisione in casi in cui i marker sono piccoli o peggio all’impossibilità di rilevarli;
2. **posizioni degli angoli dei marker falsate:** avendo in questo caso effettuato la rilevazione su un’immagine ristretta, le posizioni degli angoli dei marker saranno relative a tale immagine, per poter tornare ad utilizzare il frame di partenza nelle successive fasi è sufficiente chiamare il metodo `rescalePoints()` listing 5.12, in grado di tradurre le posizioni degli angoli dei marker ottenute sull’immagine ristretta, nelle coordinate dell’immagine originale.

Listing 5.10: Riduzione risoluzione del frame

```

1 //Resize the frame to speed up the marker detection
2 final Mat reducedFrame = new Mat();
3 Imgproc.resize(frame, reducedFrame, new Size((double) frame.width() / SCALE,
4                                         (double) frame.height() / SCALE));

```

Listing 5.11: Rilevazione markers

```

1 final Mat gray = new Mat();
2 Imgproc.cvtColor(reducedFrame, gray, Imgproc.COLOR_BGR2GRAY);
3
4 //Corners and ids of the detected markers
5 final List<Mat> corners = new ArrayList<>();
6 final Mat ids = new Mat();
7
8 //Detecting the markers
9 arucoDetector.detectMarkers(gray, corners, ids);
10 if (!corners.isEmpty()) {
11     rescalePoints(corners);
12 }

```

Listing 5.12: Traduzione delle posizioni degli angoli

```
1 private static void rescalePoints(final List<Mat> corners) {
2     for (final Mat corner : corners) {
3         for (int i = 0; i < CORNER_NUMBER; i++) {
4             final double[] data = corner.get(0, i);
5             data[0] *= SCALE;
6             data[1] *= SCALE;
7             corner.put(0, i, data);
8         }
9     }
10 }
```

5.5.3 Calcolo Posa

Questa fase rappresenta il fine ultimo del sistema e cioè stabilire una relazione tra la posizione della camera e la posizione dei marker, grazie ad OpenCV questa operazione può essere svolta con il metodo `solvePnP()` che restituisce i vettori di rotazione e traslazione in grado di tradurre le coordinate 3D nelle coordinate 2D della camera listing 5.13.

Come è possibile vedere da 4.5 la classe `CameraPose.java` mette a disposizione altri due metodi per la stima della posa: `calcSinglePose(VideoCapture cap)` e `calcSinglePose()` in grado ottenere la posa per un solo frame e ritornare i parametri che descrivono rotazione, traslazione e id di ogni marker rilevato. I due metodi differiscono solamente per un parametro che rappresenta l'interfaccia usata per ottenere il frame, il metodo che ne è sprovvisto dovrà ad ogni chiamata istanziare un `VideoCapture`, rallentando le successive fasi di calcolo della posa in attesa del frame.

Listing 5.13: Calcolo Posa

```
1 Calib3d.solvePnP(objPoints, cornerPoints, cameraMatrix,
2     new MatOfDouble(distCoeffs), rvec, tvec,
3     false, Calib3d.SOLVEPNP_ITERATIVE
4 );
```

Capitolo 6

Valutazione

Durante e al termine dello sviluppo del sistema, sono state impiegate diverse tecniche di valutazione del software prodotto, qui di seguito viene illustrata la fase di testing, un esempio operativo del sistema finale e la verifica dei requisiti.

6.1 Testing

Una delle fasi di testing più importanti riguarda il calcolo dell'errore di riproiezione¹, tale parametro è stato calcolato in due fasi: durante la calibrazione per verificare che i parametri della camera fossero stati ottenuti correttamente e durante la fase di posa controllando che le immagini dei marker acquisite rispecchiassero quelle attese.

In fase di calibrazione il calcolo avviene subito prima di ritornare i valori della camera e quindi dopo aver analizzato tutte le immagini listing 6.1 durante la stima della posa invece il calcolo dell'errore di riproiezione avviene singolarmente per ogni marker e solo successivamente viene calcolato l'errore quadratico medio sul valore totale listing 6.2.

Con la webcam FullHD utilizzata e 40 immagini di calibrazione il sistema riporta un errore di riproiezione di 0.44 pixel, anche se questo valore può ovviamente variare considerevolmente in base alla camera e al set di immagini utilizzato.

¹L'errore di riproiezione indica quanto l'immagine ottenuta si discosti da quella attesa, il risultato è dato in pixel.

Durante la posa della camera l'errore ottenuto si aggira tra 0.5 e 2.0 pixel, valori accettabili per l'obiettivo del sistema.

Listing 6.1: Calcolo errore di riproiezione in fase di calibrazione

```

1 //Reprojection error calculation
2 double totalError = 0;
3 double totalPoints = 0;
4 for (int i = 0; i < objectPoints.size(); i++) {
5   final MatOfPoint2f projectedPoints = new MatOfPoint2f();
6   Calib3d.projectPoints(new MatOfPoint3f(objectPoints.get(i)), rvecs.get(i), tvecs
7     .get(i), cameraMatrix, new MatOfDouble(distCoeffs), projectedPoints);
8   final MatOfPoint2f imgPoints = new MatOfPoint2f(imagePoints.get(i));
9   final double error = Core.norm(imgPoints, projectedPoints, Core.NORM_L2);
10  totalError += error * error;
11  totalPoints += objectPoints.get(i).total();
12 }
13 final double meanError = Math.sqrt(totalError / totalPoints);
14 LOGGER.info("Mean Reprojection Error: " + meanError);

```

Listing 6.2: Calcolo errore di riproiezione in fase di posa

```

1 private double calculateReprojectionError(final MatOfPoint3f objPoints,
2   final MatOfPoint2f imgPoints, final Mat rvec, final Mat tvec) {
3   final MatOfPoint2f projectedPoints = new MatOfPoint2f();
4   Calib3d.projectPoints(objPoints, rvec, tvec, cameraMatrix, new MatOfDouble(
5     distCoeffs), projectedPoints);
6
6 // Calculating reprojection error
7   final double error = Core.norm(imgPoints, projectedPoints, Core.NORM_L2);
8   projectedPoints.release();
9   return error;
10 }

```

6.2 Esempio operativo

Per comprendere a pieno le misurazioni effettuate di seguito è necessario tenere presente che il sistema traccia i marker a partire dall'angolo in alto a sinistra, di conseguenza le relative posizioni dei marker sono calcolate su quell'angolo.

Di seguito è mostrato un esempio operativo del sistema, sono state effettuate tre catture alle immagini ottenute dal software, la prima vede due marker posizionati

6.2. ESEMPIO OPERATIVO

uno affianco all'altro ad una distanza di 30cm 6.1, la seconda vede due marker posizionati uno sotto l'altro ad una distanza di 15cm 6.2, l'ultima invece vede due marker ruotati, il primo di 45° mentre il secondo di 90° 6.3, tutte le misure delle distanze tra i marker sono comprovate da una scala graduata con risoluzione al cm^2 .

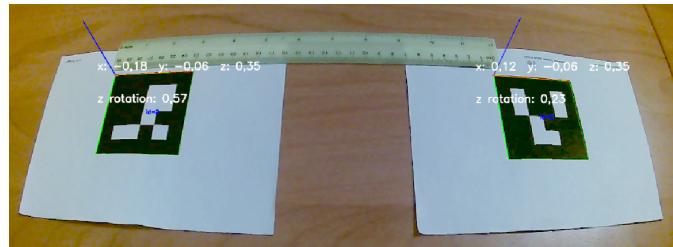


Figura 6.1: Markers posizionati a 30cm sull'asse x

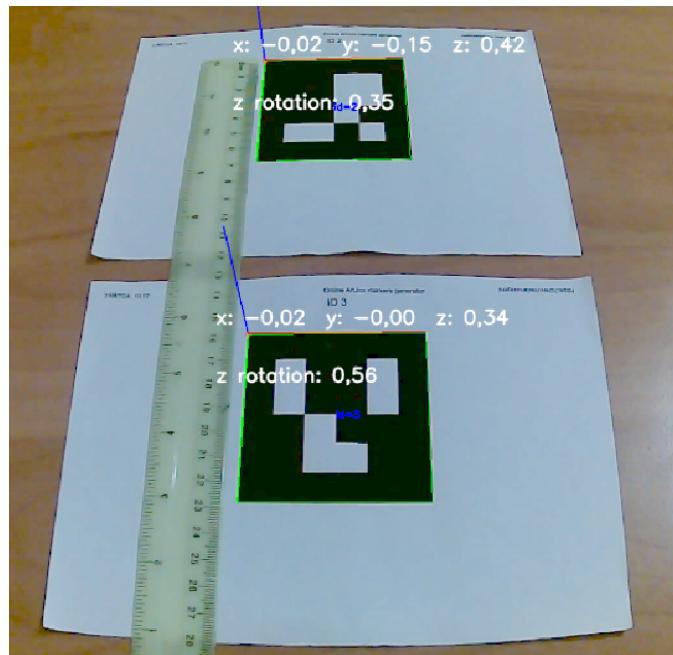


Figura 6.2: Markers posizionati a 15cm sull'asse y

²Purtroppo a causa della qualità della camera utilizzata risulta difficile comprendere le posizioni dei marker rispetto alla scala graduata, tenere quindi conto che lo strumento utilizzato ha un range operativo che va da 0 a 30cm.

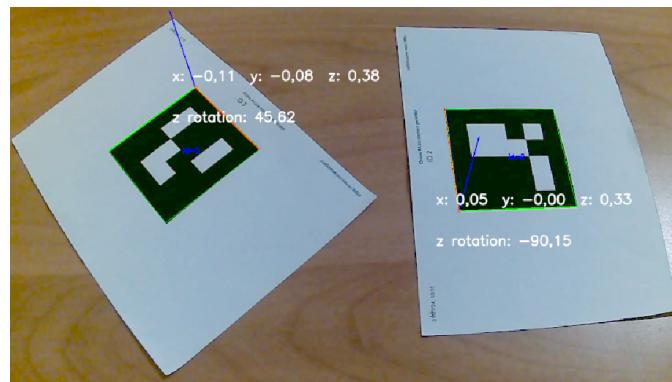


Figura 6.3: Markers ruotati rispettivamente di 45 e 90 gradi

6.3 Verifica dei requisiti

In ultimo è stato verificato che tutti i requisiti posti in fase di analisi fossero stati soddisfatti.

RF1 il sistema rileva correttamente i marker ArUco nel campo visivo della camera;

RF2 il sistema è in grado di calcolare la posizione e la rotazione dei marker rispetto alla camera;

RF3 il sistema fornisce i metodi `calcSinglePose(VideoCapture cap)` e `calcSinglePose()` per poter ottenere i risultati computati (posizione, rotazione, identificativo) sotto forma di matrici;

RF4 grazie all'utilizzo di marker ArUco ogni dispositivo viene identificato tramite un id univoco.

RNF1 il sistema è in grado di rilevare i marker ad una distanza di 2/3 metri, gli unici vincoli riscontrati sono le dimensioni dei marker che devono essere di almeno 6/7cm (misura del lato del marker) e l'illuminazione della scena che deve essere sufficiente;

6.3. VERIFICA DEI REQUISITI

RNF2 il sistema è in grado di tracciare lo spostamento dei marker in maniera molto efficace a patto che la l'esposizione della camera sia regolata correttamente e che le condizioni di illuminazione lo permettano;³

RNF3 il sistema resta affidabile anche in condizioni di scarsa luminosità purché il tempo di esposizione venga incrementato rendendo di conseguenza il sistema meno reattivo;

RNF4 il sistema è rapido ed efficiente, è stato infatti misurato che con un numero di circa 30 marker sulla scena il tempo di cattura dell'immagine, sommato al tempo di calcolo della posizione di ogni marker non supera mai i 50ms (test eseguiti su un laptop di fascia media);

RNF5 il sistema è preciso come dimostrato dall'esempio operativo 6.2 e dal testing del sistema 6.1.

RT1 il sistema è stato scritto in linguaggio Java;

RT2 è stata utilizzata una webcam FullHD 30fps durante lo sviluppo del sistema.

³l'esposizione e l'illuminazione sono parametri strettamente correlati in quanto decrementando l'esposizione la luce catturata dal sensore potrebbe non essere più sufficiente, di conseguenza l'illuminazione della scena deve essere maggiore per garantire il funzionamento del sistema.

6.3. VERIFICA DEI REQUISITI

Capitolo 7

Conclusioni

In conclusione si può ritenere l'obiettivo di questa tesi soddisfatto, lo scopo era infatti quello di sviluppare un sistema di visione artificiale in grado di rilevare e tracciare dei dispositivi nel campo visivo di una videocamera, sistema da integrare poi ad un progetto sviluppato dai ricercatori dell'università di Bologna, volto a controllare uno sciame di robot in un contesto indoor.

Durante lo sviluppo software sono state affrontate alcune sfide utili a comprendere il funzionamento della camera, delle immagini ottenute da essa e di come queste vengano rappresentate internamente al calcolatore, per fare ciò sono state di grande aiuto le funzionalità offerte da OpenCV, grazie alle quali il progetto ha potuto prendere forma e concretizzarsi.

Particolare attenzione è stata posta a rendere il sistema utilizzabile in molteplici condizioni che possono variare per illuminazione, velocità di rilevazione e numero di dispositivi nella scena.

Come mostrato da 6.3 tutti i requisiti sono stati soddisfatti anche se, per rispettare alcuni di essi è stato necessario scendere a compromessi in termini di precisione/illuminazione, è infatti emerso che a causa delle proprietà fisiche della camera per poter rilevare i dispositivi con un'elevata velocità di movimento è necessario avere un tempo di esposizione ridotto, il quale però riduce inevitabilmente la luce catturata dal sensore, richiedendo un ambiente più luminoso per il corretto funzionamento del sistema.

Qui di seguito sono mostrati i risultati del progetto completo una volta integrato a quello sviluppato dai ricercatori dell'università di Bologna.

I risultati proposti si riferiscono alla manifestazione della Notte dei ricercatori svoltasi il 27 settembre 2024, durante la quale è stato mostrato al pubblico il progetto finale.

Nella prima immagine è possibile vedere i robot utilizzati, ognuno con il proprio marker ArUco, già disposti in formazione attorno ad un robot centrale 7.1, nella seconda immagine invece è riportato il sistema completo dove, a terra sono presenti i robot mentre in alto è possibile notare una camera connessa ad un computer grazie alla quale è possibile ottenere le posizioni dei robot 7.2.



Figura 7.1: Robot identificati da marker ArUco

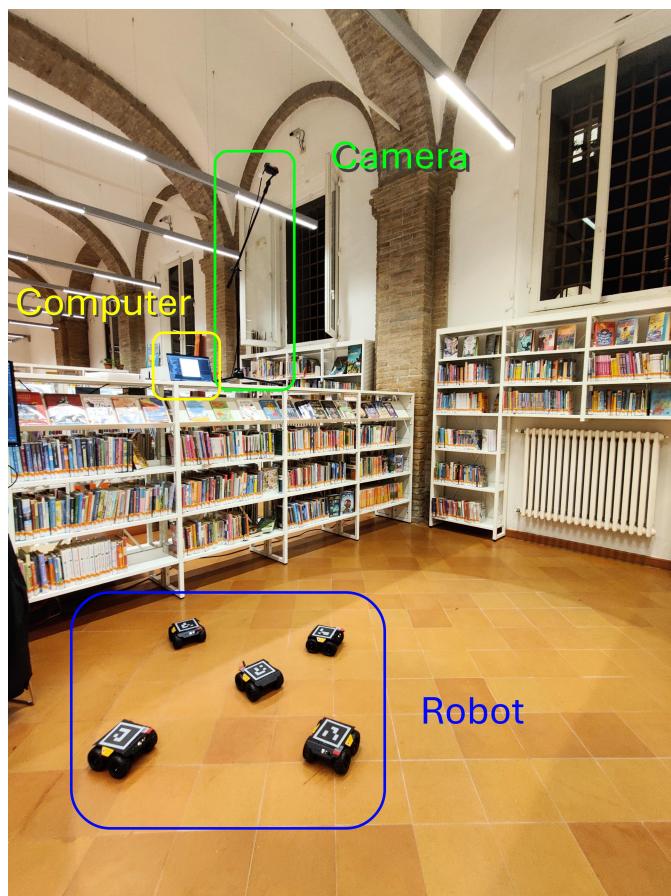


Figura 7.2: Sistema completo

7.1 Sviluppi futuri

Qui di seguito sono proposti alcuni possibili sviluppi futuri:

- miglioramento hardware (calcolatore e fotocamera) per rendere il sistema più rapido e preciso;
- aggiunta di più camere per ampliare il campo visivo e garantire una stima della profondità più accurata;
- utilizzo di sistemi alternativi per calibrare la camera come ChArUco boards, così da permettere calibrazioni precise e allo stesso tempo flessibili in caso di occlusioni.

7.1. SVILUPPI FUTURI

Bibliografia

- [1] Gianluca Aguzzi, Roberto Casadei e Mirko Viroli. “MacroSwarm: A Field-Based Compositional Framework for Swarm Programming”. In: *Coordination Models and Languages*. A cura di Sung-Shik Jongmans e Antónia Lopes. Cham: Springer Nature Switzerland, 2023, pp. 31–51. ISBN: 978-3-031-35361-1.
- [2] Gianluca Aguzzi et al. “Dynamic Decentralization Domains for the Internet of Things”. In: *IEEE Internet Computing* 26.6 (nov. 2022), pp. 16–23. ISSN: 1941-0131. DOI: 10.1109/MIC.2022.3216753.
- [3] Samuel Audet. *javacv*. URL: <https://github.com/bytedeco/javacv?tab=readme-ov-file> (visitato il 06/10/2024).
- [4] D.L. Baggio. *OpenCV 3.0 Computer Vision with Java*. Packt Publishing, 2015. ISBN: 9781783283989. URL: <https://books.google.it/books?id=LFTICgAAQBAJ>.
- [5] G. Bradski e A. Kaehler. *Learning OpenCV: Computer Vision with the OpenCV Library*. O'Reilly Media, 2008, pp. 2–5. ISBN: 9780596554040. URL: <https://books.google.it/books?id=seAgi0fu2EIC>.
- [6] Roberto Casadei et al. “A development approach for collective opportunistic Edge-of-Things services”. In: *Information Sciences* 498 (2019), pp. 154–169. ISSN: 0020-0255. DOI: <https://doi.org/10.1016/j.ins.2019.05.058>. URL: <https://www.sciencedirect.com/science/article/pii/S002002551930461X>.

BIBLIOGRAFIA

- [7] Roberto Casadei et al. “Modelling and simulation of Opportunistic IoT Services with Aggregate Computing”. In: *Future Generation Computer Systems* 91 (2019), pp. 252–262. ISSN: 0167-739X. DOI: <https://doi.org/10.1016/j.future.2018.09.005>. URL: <https://www.sciencedirect.com/science/article/pii/S0167739X18307246>.
- [8] Wikipedia contributors. *Pinhole camera*. 2024. URL: https://en.wikipedia.org/wiki/Pinhole_camera (visitato il 14/10/2024).
- [9] Wikipedia contributors. *Pinhole camera model*. 2024. URL: https://en.wikipedia.org/wiki/Pinhole_camera_model (visitato il 13/10/2024).
- [10] Milan Košták e Antonín Slabý. “Designing a Simple Fiducial Marker for Localization in Spatial Scenes Using Neural Networks”. In: *Sensors* 21.16 (2021), p. 7. ISSN: 1424-8220. DOI: 10.3390/s21165407. URL: <https://www.mdpi.com/1424-8220/21/16/5407>.
- [11] MathWorks. *What Is Camera Calibration?* 2024. URL: <https://it.mathworks.com/help/vision/ug/camera-calibration.html> (visitato il 13/10/2024).
- [12] Microsoft. *Che cos’è la visione artificiale?* URL: <https://azure.microsoft.com/it-it/resources/cloud-computing-dictionary/what-is-computer-vision#classificazione-degli-oggetti> (visitato il 22/09/2024).
- [13] Brandon Minor. *Reverse-Engineering Fiducial Markers For Perception*. URL: <https://www.tangramvision.com/blog/reverse-engineering-fiducial-markers-for-perception> (visitato il 22/09/2024).
- [14] OpenCV.org. *OpenCV*. URL: <https://opencv.org> (visitato il 06/10/2024).
- [15] OpenCV.org. *opencv*. URL: <https://github.com/opencv/opencv.git> (visitato il 11/11/2024).
- [16] The Sama Team. *Computer Vision: History & How it Works*. URL: <https://www.sama.com/blog/computer-vision-history-how-it-works> (visitato il 22/09/2024).

BIBLIOGRAFIA

- [17] Mirko Viroli et al. “From distributed coordination to field calculus and aggregate computing”. In: *Journal of Logical and Algebraic Methods in Programming* 109 (2019), p. 1. ISSN: 2352-2208. DOI: <https://doi.org/10.1016/j.jlamp.2019.100486>. URL: <https://www.sciencedirect.com/science/article/pii/S235222081930032X>.
- [18] Kev Zettler. *What is a distributed system?* URL: <https://www.atlassian.com/microservices/microservices-architecture/distributed-architecture> (visitato il 07/10/2024).

BIBLIOGRAFIA

Ringraziamenti

Al termine di questa tesi mi sembra doveroso ringraziare tutte le persone che mi hanno aiutato a raggiungere questo obiettivo.

In primis ci tengo a ringraziare il Prof. Mirko Viroli, sempre disponibile e professionale, grazie al quale ho avuto la possibilità di prendere parte ad un progetto interessante e innovativo, che mi ha permesso di espandere le mie conoscenze e consolidare le competenze acquisite durante questi tre anni. Un ringraziamento speciale va al Dott. Gianluca Aguzzi che mi ha seguito durante tutta la stesura della tesi, fornendomi sempre un aiuto ove necessario in maniera pronta e gentile.

Non posso non ringraziare tutta la mia famiglia, in particolare i miei genitori e mio fratello, che in questi anni hanno sempre saputo comprendermi e anche quando le mie ansie prendevano il sopravvento erano sempre pronti ad ascoltarmi e rassicurarmi.

Grazie ai miei genitori per aver supportato ogni mia decisione e per avermi reso quello che sono, grazie a mio fratello con cui so di potermi sempre confidare, grazie ai miei nonni per essere fonte di ispirazione e per credere sempre in me.

Un ringraziamento speciale va ai miei amici più stretti, quelli con cui sono cresciuto e quelli che ho avuto la fortuna di conoscere più avanti, con i quali ho affrontato mille avventure e grazie ai quali riesco sempre ad essere me stesso e a dimenticarmi delle mie preoccupazioni.

Un grazie ai miei colleghi dell'università con cui ho condiviso gioie e dolori di questo percorso universitario, dai pranzi assieme sino ai progetti d'esame più difficili. Vi ringrazio di aver reso le mie giornate in università meno impegnative e di avermi sempre aiutato in situazioni di difficoltà.