

Indice

Introduzione.....	6
Capitolo 1: Realtà virtuale e computer vision.....	10
Introduzione.....	10
1.1 Realtà virtuale immersiva.....	11
1.2 Dispositivi per l'interazione.....	12
1.3 Gradi di immersione.....	14
1.4 Realtà mista.....	15
1.5 Realtà aumentata.....	16
1.6 Principali funzionalità delle applicazioni di A.R.....	18
1.7 Dispositivi di visualizzazione per i sistemi di A.R.....	18
Computer Vision.....	20
Definizione.....	20
2.1 Intelligenza artificiale.....	20
2.2 Origini delle IA.....	20
2.3 Alan Turing.....	21
2.4 Applicazioni pratiche delle IA.....	22
2.5 Visione: Dalla biologia all'elettronica.....	23
2.6 Il processo di digitalizzazione.....	24
2.7 Elaborazione delle immagini.....	30
3 La libreria OpenCV.....	32
Capitolo 2: Riconoscimento facciale.....	34
Introduzione.....	34
1 Cenni storici.....	34
2 Applicazioni commerciali.....	35
3 Algoritmi per il riconoscimento dei volti.....	38
3.1 Estrattori di features.....	38
3.2 Estrazione dei volti.....	38
4 Algoritmi di riconoscimento facciale.....	43
4.1 Eigenfaces.....	43
4.2 Fisherfaces.....	46

4.3	Confronto tra Eigenfaces e Fisherfaces.....	49
4.4	Local Binary Patterns.....	49
4.5	LBP per il Face Recognition.....	52
Capitolo 3: Emotional Recognition.....		54
Introduzione.....		54
1	Analisi delle espressioni.....	55
1.1	Rabbia.....	55
1.2	Disgusto.....	56
1.3	Felicità.....	57
1.4	Tristezza.....	57
1.5	Sorpresa.....	58
1.6	Disprezzo.....	59
1.7	Paura.....	60
2	Metodi per il riconoscimento.....	60
2.1	Tracciamento dei punti sul volto.....	60
2.2	Approccio statico: Reti Bayesiane e TAN.....	64
2.3	Approccio statico: Reti neurali.....	65
2.4	Approccio dinamico: Modelli di Markov Nascosti.....	65
3	Il metodo FACS.....	68
4	Il Cohn-Kanade AU-Coded Facial Expressions Database.....	71
4.1	Specifiche del database.....	71
5	Progettazione del Sistema.....	73
Capitolo 4: Implementazione.....		74
1.	Introduzione.....	74
2.	Immagini di training.....	75
3.	Training tramite labels.....	76
4.	Fase di riconoscimento.....	77
5.	Person-independent test.....	79
5.1	Person-independent test con un soggetto di training.....	79
5.2	Person-independent test con 2 soggetti di training.....	81
5.3	Person-independent test con 3 soggetti di training.....	83
5.4	Person-independent test: Conclusioni.....	85

6.	Person-dependent recognition.....	85
6.1	Acquisizione delle immagini.....	86
6.2	Test con un soggetto.....	87
6.3	Test con 2 soggetti.....	89
6.4	Person-dependent recognition: Conclusioni.....	90
7.	Altre applicazioni.....	91
	Conclusione.....	94
	Riferimenti.....	96

Introduzione

Questo lavoro di tesi ha come scopo uno studio sulla Computer Vision e sulle possibilità che offre. In particolare lo scopo finale del lavoro è quello di progettare un sistema che, in tempo reale e tramite la webcam installata sul pc, sia in grado di riconoscere le espressioni facciali dell'utente.

Le espressioni facciali sono da sempre il più potente mezzo di comunicazione. Da qui la famosa frase “uno sguardo vale più di mille parole”. Per certi versi si tratta di un linguaggio universale come la matematica.

La scelta dell'argomento è dovuta principalmente alla mia natura empatica nei confronti di altri individui, oltre alla mia passione per il disegno, in cui una buona rappresentazione dell'espressione del soggetto è probabilmente il punto centrale di un disegno valido.

La lettura delle espressioni è senza dubbio frutto dell'abitudine. Un bambino non è in grado di riconoscere le espressioni facciali di un altro individuo. Solo con l'abitudine, dopo aver visto migliaia di volte quella data espressione sul volto di molti individui, si è in grado di riconoscerla. L'approccio utilizzato per questo lavoro di tesi si basa sullo stesso concetto, ovvero “istruire” il calcolatore a riconoscere una data espressione partendo da numerosi volti che la ritraggono.

Il primo capitolo tratta la Realtà Virtuale e la Computer Vision in maniera generale. Nella prima parte vi è uno sguardo ai concetti di Realtà Virtuale Immersiva, con i vari gradi di immersione, Realtà Aumentata e Realtà Mista e ai dispositivi utilizzati per la loro realizzazione, tra cui dispositivi di motion tracking e motion capture, e i dispositivi di visualizzazione, tra cui dispositivi Head Mounted Display(HMD) Video See Through e Optical See Through. Nella seconda viene trattata la Visione Artificiale, partendo dalla branca di cui fa parte, l'Intelligenza Artificiale, trattata dal punto di vista storico con la figura di Alan Turing. Vengono poi prese in considerazione le connessioni tra la visione biologica e quella artificiale artificiale, la digitalizzazione e l'elaborazione delle immagini, tra cui l'estrazione dei contorni. Viene poi introdotta la libreria utilizzata per la realizzazione dell'architettura software, ovvero OpenCV nella sua versione Java (JavaCV).

Nel secondo capitolo viene trattato il riconoscimento facciale dal punto di vista storico, con le principali applicazioni in commercio, tra cui TrueFace e FaceIt, sviluppate da Miros e Visionics. Si passa poi per il rilevamento dei volti attraverso estrazione delle

features con l'algoritmo di Viola-Jones e le cascate di classificatori, con il relativo codice Java tramite libreria OpenCV per dare un primo sguardo agli oggetti che entrano in gioco nella realizzazione di applicazioni basate sulla computer vision. Vengono infine studiati dal punto di vista matematico e pratico i principali algoritmi di riconoscimento facciale contenuti nella libreria OpenCV, ovvero Eigenfaces, Fisherfaces e Local Binary Patterns. Il terzo capitolo tratta le espressioni facciali, con uno sguardo ai principali movimenti muscolari delle espressioni di base, tra cui rabbia, felicità, tristezza, paura, disprezzo, disgusto e sorpresa, per poi passare ai principali e più utilizzati algoritmi per il riconoscimento delle espressioni, ovvero il tracciamento di punti sul volto, Reti Bayesiane, TAN, Reti Neurali, Modelli di Markov Nascosti e il metodo FACS, con le sue principali unità d'azione. Inoltre viene sviscerato il dataset utilizzato per le operazioni di training del nostro sistema, il Cohn-Kanade Facial Expressions Database, con le varie specifiche interne. Infine viene dato un primo abbozzo di quella che sarà la progettazione del nostro sistema di riconoscimento delle espressioni facciali, che utilizza alcuni degli strumenti studiati in precedenza, quali estrattori di features tramite descrittori xml, riconoscimento facciale tramite Eigenfaces, Fisherfaces e Local Binary Patterns e reti bayesiane.

Il quarto capitolo riguarda l'implementazione del sistema. Viene inizialmente spiegato l'approccio utilizzato per estrarre il volto tramite cascata di classificatori e Region of Interest. Viene poi introdotto il concetto di label e lo pseudocodice per il training. Si passa poi alla fase di riconoscimento e al relativo pseudocodice. Seguono poi una serie test indipendenti dalla persona (utilizzando quindi soggetti di training diversi da quelli dei test) per valutare il comportamento del sistema in presenza di un soggetto sconosciuto. I test sono effettuati in presenza di uno, due o tre soggetti di training, per valutare le differenze nel riconoscimento in tutti e tre i casi. Si passa quindi al riconoscimento dipendente dalla persona, ovvero il riconoscimento in cui il soggetto dei test è anche un soggetto di training. Quest'ultimi sono effettuati su due soggetti che si alternano di fronte alla webcam.