

UNIVERSITÀ POLITECNICA DELLE MARCHE
FACOLTÀ DI INGEGNERIA

CORSO DI LAUREA IN INGEGNERIA INFORMATICA E
DELL'AUTOMAZIONE



**Implementazione di un algoritmo di
identificazione della persona
utilizzando frame di profondità**

—

***Implementation of a depth-based human
identification algorithm***

RELATORE:
Prof. Ennio Gambi

CORRELATORE:
Prof.ssa Susanna Spinsante
Ing. Samuele Gasparrini (?)

TESI DI LAUREA DI:
Ilario Pierbattista

ANNO ACCADEMICO 2014/2015

Indice

1	Introduzione	3
1.1	Human Sensing	3
1.1.1	Human Sensing	3
1.1.2	Stato dell'arte	3
1.2	Panoramica Generale	4
1.2.1	Introduzione al lavoro di Zhu Wong	4
1.2.2	Configurazione dell'Hardware	4
1.2.3	<i>Head and Shoulders Profile</i>	4
1.2.4	Flusso di Lavoro	5
2	Haar-Like Features	6
2.1	Definizione	6
2.1.1	Richiamo: cosa è una feature (caratteristica)	6
2.1.2	Wavelet di Haar	6
2.1.3	Formula di calcolo standard	7
2.1.4	Cosa mette in evidenza la feature di Haar	7
2.1.5	Formula di calcolo invariante ai resize	7
2.1.6	Vantaggi	7
2.2	Immagine Integrale	7
2.2.1	Definizione rigorosa dell'immagine integrale	7
2.2.2	Complessità computazionale generale	8
2.3	Decision Stump	8
2.3.1	Problema: utilizzare le feature	8
2.3.2	Definizione di albero decisionale	8
2.3.3	Definizione di decision stump	8
3	L'Algoritmo di Allenamento: Adaboost	9
3.1	Apprendimento Supervisionato <i>Ensamble</i>	9
3.1.1	Apprendimento Supervisionato	9
3.1.2	<i>Adaptive Boosting</i>	10
3.2	Dataset di Allenamento	10
3.2.1	Categorie di Classificatori	10
3.2.2	Preparazione dei Dataset	10
3.2.3	Preprocessing	10
3.3	<i>Strong Learner</i>	11
3.3.1	Procedura di estrazione del classificatore forte	11

3.4	<i>Weak Learner</i>	11
3.4.1	Procedura di estrazione del classificatore debole	11
3.4.2	Valutazione della complessità computazionale	11
4	Validazione e Regolazione dei Classificatori	12
4.1	Criteri di Valutazione	12
4.2	Dataset di Validazione	12
4.2.1	Criteri di creazione delle registrazioni	12
4.2.2	Altre caratteristiche	12
4.3	Massimizzazione all' <i>Accuracy</i>	12
4.3.1	Parametri liberi del classificatore	12
4.3.2	Algoritmo di ricerca della soglia e del NWL ottimi	12
4.4	Analisi dei Risultati	12
5	Rilevamento	13
5.1	Tecnica di Rilevamento	13
5.1.1	Detection Window	13
5.1.2	Rilevazione su frame	13
5.2	Selezione della Finestra Migliore	13
5.2.1	Introduzione al problema	13
5.2.2	Algoritmo di selezione	13
5.3	Confronto con l'Algoritmo G-C	13
6	Conclusioni	14
	Appendici	15
A	Software Sviluppato	16
A.1	Componenti	16
A.1.1	Creatore dei Dataset	16
A.1.2	Allenamento	16
A.1.3	Tuning, Testing, Rilevamento	16
A.2	Tecnologie utilizzate	16
A.2.1	C++ e Matlab	16
A.2.2	Git e Github [Opzionale]	16
A.3	Proposte di miglioramento	16
B	Cenni del funzionamento del sensore Kinect	17

Capitolo 1

Introduzione

1.1 Human Sensing

1.1.1 Human Sensing

5 Definizione

L'insieme di tecniche di riconoscimento della presenza di una persona nello spazio prendono il nome di tecniche di *human sensing*. Sensori di vario tipo vengono utilizzati nelle tecniche di riconoscimento. Una volta acquisite le informazioni dai sensori, vanno elaborate da un apposito algoritmo per rilevare la presenza e la posizione della persona
10 nell'ambiente.

Contesti Applicativi

La possibilità di rilevare la presenza di persone, rende le applicazioni di *human sensing* perfette per le applicazioni di sorveglianza. Sistemi di *people counting* sono utili per la conduzione di indagini di mercato. Dispositivi in grado di rilevare la presenza di corpi
15 umani in contesti di crisi sono utilizzati nelle attività di *search rescue*. Le applicazioni di *human tracking* sono utili anche negli ambienti assistivi automatizzati al fine di monitorare le attività dell'assistito.

HS e Computer Vision

Le applicazioni di human sensing che utilizzano sensori di acquisizione *visiva* risolvono
20 problemi di computer vision. Lo scopo della *computer vision* è quello di riprodurre la vista umana. L'obiettivo di tale riproduzione non si limita alla semplice acquisizione di una rappresentazione bidimensionale di una regione di spazio, ma mira all'interpretazione del relativo contenuto.

1.1.2 Stato dell'arte

25 Pedestrian Detection and Counting

Papageorgiou et Al [6] hanno sviluppato un sistema di riconoscimento e conteggio di pedoni a partire da immagini RGB.

Face Recognition

30 Viola e Jones [9] hanno proposto un framework per il riconoscimento dei volti all'interno di immagini RGB. Al momento è il sistema più solido nel suo contesto.

Kinect: a serious game

35 Gli ambiti d'utilizzo del dispositivo Kinect, nota periferica legata a sistemi videoludici, si stanno espandendo costantemente. La quantità e la qualità dei sensori di cui è equipaggiato, il costo relativamente contenuto e l'evoluzione di framework e toolkit di sviluppo, lo rendono un dispositivo particolarmente versatile e adatto allo studio di problemi di computer vision.

1.2 Panoramica Generale

1.2.1 Introduzione al lavoro di Zhu Wong

Elenco delle tecnologie coinvolte

40 Kinect per l'acquisizione

Adaboost per l'allenamento

Affermare la notevole affinità con Viola Jones

1.2.2 Configurazione dell'Hardware

Sensore utilizzato

45 Descrizione sommaria del Kinect

Breve esposizione delle caratteristiche tecniche del sensore

Presentazione delle immagini di profondità

Configurazione Top-Down

Esposizione della configurazione hardware

50 Vantaggi del top-down rispetto al rilevamento frontale

1.2.3 *Head and Shoulders Profile*

L'attività di riconoscimento è un'attività di classificazione

Cosa bisogna riconoscere?

Il riconoscimento è un'attività affine alla discriminazione

55 Definizione di classi di oggetti: classificazione

La classificazione si basa sulla misurazione di alcune caratteristiche

Discriminare in base a cosa?

Le caratteristiche di un oggetto sono la base su cui differenziare un oggetto da un altro

60 Caratteristiche del profilo HASP in linguaggio naturale

Come si può descrivere un'immagine di profondità?

Le tre caratteristiche

1.2.4 Flusso di Lavoro

Definizione dei moduli funzionali

65 Allenamento

Rilevamento

Allenamento

Input: Dataset

Adaboost

70 Output: Classificatore

Rilevamento

Input: Frame

Algoritmo di rilevamento

Output: finestre di rilevamento

Capitolo 2

Haar-Like Features

2.1 Definizione

2.1.1 Richiamo: cosa è una feature (caratteristica)

Cosa sono le caratteristiche di un oggetto

80 Le caratteristiche dipendono da cosa si vuole evidenziare

Le caratteristiche dipendono da cosa si ha a disposizione

2.1.2 Wavelet di Haar

Le feature di Haar derivano dalle wavelet di Haar

Definizione informale delle wavelet di Haar

85 Chi le ha sviluppate

Cosa sono (base ortonormale spazio funzionale)

Rappresentazione dei segnali (fourier duale)

Wavelet di Haar e DWT

Utilizzi commerciali DWT (JPEG2000)

90 Utilizzo delle dwt per il pattern recognition

2.1.3 Formula di calcolo standard

Rappresentazione visuale

Formula generale

Altri tipi di feature (OpenCv)

95 Tipi di feature utilizzate

2.1.4 Cosa mette in evidenza la feature di Haar

Immagini normali (Viola Jones)

Immagini di profondità (Zhu Wong)

2.1.5 Formula di calcolo invariante ai resize

100 Anticipazione del problema del ridimensionamento

Formula: Normalizzazione sull'area

2.1.6 Vantaggi

Differenze di intensità vs Valutazione dei singoli pixel

Differenze di intensità vs Estrazione dei contorni

105 Estrema efficienza computazionale

2.2 Immagine Integrale

2.2.1 Definizione rigorosa dell'immagine integrale

Problema: efficienza nel calcolo di somme di pixel

Complessità computazionale del calcolo *ignorante*

110 Soluzione: rendere queste somme subito disponibili

Definizione immagine integrale

Formula di calcolo della somma dei pixel in un'area

Enunciazione della formula

Dimostrazione della formula

115 Complessità computazionale del calcolo della feature

2.2.2 Complessità computazionale generale

Complessità del calcolo dell'immagine integrale

Convenienza del calcolo dell'immagine integrale

2.3 Decision Stump

120 2.3.1 Problema: utilizzare le feature

È necessario un meccanismo primitivo per utilizzare le feature

Bisogna discriminare in base al valore

2.3.2 Definizione di albero decisionale

2.3.3 Definizione di decision stump

125 Radice: Test, funzione booleana

Foglie: risultati possibili

Formule di calcolo binaria

Formula di calcolo unica: polarità

Capitolo 3

130 L'Algoritmo di Allenamento: Adaboost

3.1 Apprendimento Supervisionato *Ensamble*

3.1.1 Apprendimento Supervisionato

Definizione

135 Obiettivo

Spazio delle Ipotesi

Esempi di Supervised learning

Algoritmi di Sup.Learning

Maggiori campi applicativi

140 Concetti di base

Overfitting

Ensamble Learning

3.1.2 *Adaptive Boosting*

Algoritmi di Boosting

145 Aptive: adattabilità

Strong learner e Weak learner

3.2 Dataset di Allenamento

3.2.1 Categorie di Classificatori

Variabilità della forma HASP

150 Variazione dell'orientazione

Variazione derivata dalla distorsione prospettica

Definizione delle categorie di classificatori

Categorie: Verticale e Orizzontale

Categorie alternative: Obliquo, a zone

155 3.2.2 Preparazione dei Dataset

Acquisizioni

Soggetti, percorsi

Acquisizione delle registrazioni

Ritaglio dei samples

160 Trainset Creator

3.2.3 Preprocessing

Resize

Nearest Neighbour

Altri algoritmi di resize

165 Conversione delle distanze

3.3 *Strong Learner*

3.3.1 Procedura di estrazione del classificatore forte

3.4 *Weak Learner*

3.4.1 Procedura di estrazione del classificatore debole

170 3.4.2 Valutazione della complessità computazionale

Capitolo 4

Validazione e Regolazione dei Classificatori

4.1 Criteri di Valutazione

175 4.2 Dataset di Validazione

4.2.1 Criteri di creazione delle registrazioni

4.2.2 Altre caratteristiche

4.3 Massimizzazione all'*Accuracy*

4.3.1 Parametri liberi del classificatore

180 Numero di weak learner

Soglia del classificatore

4.3.2 Algoritmo di ricerca della soglia e del NWL ottimi

4.4 Analisi dei Risultati

Capitolo 5

185 Rilevamento

5.1 Tecnica di Rilevamento

5.1.1 Detection Window

5.1.2 Rilevazione su frame

Resize detection window

190 Slide detection window

5.2 Selezione della Finestra Migliore

5.2.1 Introduzione al problema

5.2.2 Algoritmo di selezione

5.3 Confronto con l'Algoritmo G-C

Capitolo 6

Conclusioni

Appendici

Appendice A

Software Sviluppato

- 200 **A.1 Componenti**
 - A.1.1 Creatore dei Dataset**
 - A.1.2 Allenamento**
 - A.1.3 Tuning, Testing, Rilevamento**
- A.2 Tecnologie utilizzate**
 - 205 **A.2.1 C++ e Matlab**
 - A.2.2 Git e Github [Opzionale]**
- A.3 Proposte di miglioramento**

Appendice B

Cenni del funzionamento del sensore Kinect

Bibliografia

- [1] Thomas H Cormen. Introduction to algorithms. 2009.
- [2] Ronald A Fisher. The use of multiple measurements in taxonomic problems. *Annals of eugenics*, 7(2):179–188, 1936.
- 215 [3] Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of computer and system sciences*, 55(1):119–139, 1997.
- [4] Alfred Haar. Zur theorie der orthogonalen funktionensysteme. *Mathematische Annalen*, 69(3):331–371, 1910.
- 220 [5] Michael Oren, Constantine Papageorgiou, Pawan Sinha, Edgar Osuna, and Tomaso Poggio. Pedestrian detection using wavelet templates. In *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*, pages 193–199. IEEE, 1997.
- 225 [6] Constantine P Papageorgiou, Michael Oren, and Tomaso Poggio. A general framework for object detection. In *Computer vision, 1998. sixth international conference on*, pages 555–562. IEEE, 1998.
- [7] ITUT Rec. T. 800— iso/iec 15444-1,“. *Information technology—JPEG*, 2000.
- [8] Stuart Russell and Peter Norvig. Artificial intelligence: a modern approach. 1995.
- 230 [9] Paul Viola and Michael J Jones. Robust real-time face detection. *International journal of computer vision*, 57(2):137–154, 2004.
- [10] Lei Zhu and Kin-Hong Wong. Human tracking and counting using the kinect range sensor based on adaboost and kalman filter. *Advances in Visual Computing*, pages 582–591, 2013.