

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

SCUOLA DI INGEGNERIA E ARCHITETTURA

Corso di Laurea in Ingegneria Informatica

Determinazione della confidenza di mappe depth tramite Deep Learning

Relatore:

Chiar.mo Prof.

Mattoccia Stefano

Presentata da:

Fusco Alessandro

Sessione di ottobre

2016/2017

Indice

1	Introduzione	2
2	Sistemi di visione stereo	4
2.1	Cenni di stereo visione	4
2.2	Sistemi attivi e passivi	5
2.3	Algoritmi di matching stereo	5
2.4	Misure di confidenza per algoritmi stereo	5

Capitolo 1

Introduzione

Nel vasto panorama del digitale, uno degli argomenti più caldi degli anni recenti è l'intelligenza artificiale (AI), che sta ricevendo sempre più attenzione sia dal mondo accademico che da quello industriale.

In particolare il Machine Learning (apprendimento automatico), ovvero quella branca dell'intelligenza artificiale che si occupa di fornire ai calcolatori l'abilità di apprendere senza essere stati esplicitamente programmati, ha subito recentemente un boom esponenziale.

Seppur le basi matematiche del Machine Learning siano state definite e studiate lungo il corso del ventesimo secolo, solo di recente [**krizhevsky2012imagenet**] le si sono potute applicare efficientemente a scopi pratici, grazie al notevole miglioramento delle tecnologie di calcolo parallelo necessarie.

Tra le tante applicazioni di Machine Learning, emerge tra le altre quella relativa alla computer vision. Nel momento in cui si vede necessaria un'interpretazione ad alto livello di quelli che sono i vari bit di informazione di un'immagine o di un video, quando è necessario definire, rilevare o ricostruire proprietà visuali d'interesse inesprimibili proceduralmente, lì i mondi della visione artificiale e dell'apprendimento automatico si vanno a fondere.

La visione stereo è quella branca della computer vision che si occupa della ricostruzione di scene tridimensionali tramite l'acquisizione di immagini bidimensionali, tentando di ricostruire l'informazione perduta nel processo di acquisizione.

L'esigenza di accuratezza nell'impiego di algoritmi stereo (detti di stereo matching) ha portato necessità della creazione di misure di confidenza, ovvero indici che riescano a valutare a monte la correttezza di un processo di ricostruzione 3D arbitrario.

I metodi di misura di confidenza attuali sono molteplici; alcune strategie possono utilizzare nozioni geometriche del sistema stereoscopico; metodi più complessi, quali alberi di classificazione, possono combinare le varie strategie per valutare quale sia la più efficace in un determinato contesto. Tecniche più recenti basate su Deep Learning, anche se computazionalmente più costose, hanno dimostrato però di portare a risultati di qualità notevolmente superiore, in quanto riescono a riconoscere pattern ad alto livello dei possibili problemi che possono indurre un algoritmo stereo a riportare risultati errati, quale la presenza di superfici riflettenti o occlusioni.

Questa tesi vuole essere uno studio su come applicare tecniche di deep learning al fine della determinazione di una misura di confidenza "full-resolution", ovvero calcolata analizzando l'intera immagine stereo.

Capitolo 2

Sistemi di visione stereo

2.1 Cenni di stereo visione

Due degli scopi fondamentali della percezione spazio-visiva umana, sono la determinazione della posizione degli oggetti nell'ambiente, e la distinzione di essi.

La stereopsi, o visione stereoscopica, è ciò che ci permette di percepire la realtà tridimensionalmente, e quindi di assegnare una profondità ai punti dello spazio. Essa risulta dalla fusione delle due immagini leggermente differenti proiettate all'interno dei nostri occhi; un oggetto viene percepito come vicino se, nel confronto tra i segnali generati dalle due retine, il nostro cervello lo vede scostato. Al contrario, minore è lo scostamento, maggiore è la distanza percepita.

La replica di questo interessante meccanismo vede numerosissime applicazioni a livello ingegneristico, dalla guida autonoma [goldberg2002stereo] [broggi2001self] alla realtà aumentata [kanbara2000stereo]; perciò la visione stereo è uno dei principali argomenti di studio della visione artificiale.

2.2 Sistemi attivi e passivi

Le metodologie utilizzate per raccogliere dati riguardo alla configurazione geometrica di una scena tridimensionale si suddividono in due categorie. Quelle di tipo *attivo* usufruiscono di dispositivi che interagiscono fisicamente con l'ambiente circostante, ad esempio tramite laser o ultrasuoni, acquisendo ed elaborando i dati

2.3 Algoritmi di matching stereo

2.4 Misure di confidenza per algoritmi stereo