

Analisi del movimento dell'uomo e applicazioni alla riabilitazione e alla medicina

A. Barla, M. Casadio, N. Noceti, M. Moro, F. Odone, L. Rosasco

DIBRIS Università degli Studi di Genova

{cognome.nome}@unige.it

Abstract

L'obiettivo della ricerca presentata è lo sviluppo di metodi di computer vision e machine learning per l'analisi senza marker del movimento dell'uomo in contesti di riabilitazione. Queste tecniche di intelligenza artificiale possono portare significativi cambiamenti nella pratica clinica e facilitare la progettazione e l'implementazione di sistemi di analisi del movimento ecologici, non invasivi, con una ridotta polarizzazione e un potenziale miglioramento della velocità di risposta rispetto alle tecniche tradizionali.

1 Introduzione

Misurare e comprendere il movimento umano è cruciale in diversi campi, dalle neuroscienze, alla riabilitazione e alla biomeccanica sportiva. Negli ultimi decenni questo campo di ricerca ha compiuto notevoli progressi [Colyer *et al.*, 2018].

In clinica e nell'ambito del controllo motorio, la tecnologia che fornisce il riferimento per misure quantitative e accurate è basata su sistemi marker, ma evidenzia limiti oggettivi: i marker sono invasivi, il loro numero e la loro posizione sono determinati a priori, richiedono complesse sessioni di acquisizione. Inoltre richiedono dispositivi costosi con diverse telecamere a infrarossi. Una conseguenza rilevante è il potenziale disagio che procurano e il conseguente disturbo alla naturalezza dei movimenti. Infine, queste tecniche sono computazionalmente onerose nel tempo e nello spazio. Di conseguenza non vengono comunemente adottate in tempo reale all'interno di sessioni di riabilitazione, né sono applicabili ad ambienti domestici. La videografia è un metodo semplice e ampiamente utilizzato dagli operatori del settore della riabilitazione per osservare e registrare i movimenti umani in ambienti naturali, soprattutto utilizzata per valutazioni portate avanti manualmente dall'operatore. Viceversa, non sono ad oggi disponibili sistemi per l'estrazione automatica o semi-automatica di misure quantitative affidabili da sequenze video ai fini del controllo motorio; le tecniche disponibili richiedono conoscenze tecniche a priori e sono poco adatte alla pratica clinica in quanto il loro livello di affidabilità in questo ambito è ancora da verificare [Mathis *et al.*, 2018].

Al tempo stesso, lo stato dell'arte dell'intelligenza artificiale apre diverse possibilità, verso lo sviluppo di sistemi au-

tomatici in grado di fornire misure quantitative significative, precise e affidabili del movimento umano, anche in situazioni non controllate. In primo luogo, i recenti progressi nell'ambito della stima della posa senza marker basati su computer vision e deep learning ci permettono di estrarre informazioni sul movimento da video. Inoltre, grazie al machine learning possiamo analizzare i pattern di movimento osservati fornendo supporto all'interpretazione dei dati e alle decisioni cliniche.

In questo contributo presentiamo gli obiettivi generali della nostra ricerca in questo ambito e quelli più specifici all'interno progetto MOREorLESS¹, la cui ambizione è quella di progettare e realizzare metodi markerless basati su video e tecniche basate su analisi dati per modellare l'evoluzione del movimento nel tempo e in grado di fornire una valutazione quantitativa del comportamento umano in contesti naturali (ad esempio ambienti domestici). Questa ricerca faciliterà lo sviluppo di sistemi che offrono vari vantaggi: minori costi delle apparecchiature, un bias ridotto, elaborazione più breve e tempi di post-elaborazione, con un potenziale di analisi in tempo reale; interpretabilità migliorata; portabilità in ambienti domestici o domestici; la possibilità di estrarre misure quantitative e significative da grandi quantità di registrazioni video già disponibili. Immaginiamo un futuro in cui una persona con disabilità motoria, o una persona che pratica uno sport possa ricevere informazioni accurate, quantitative e significative sulle sue prestazioni, insieme a una classificazione del suo comportamento, che evidenzia i punti di forza e di debolezza, da semplici sequenze video.

2 Obiettivi di ricerca

Il nostro punto di partenza è un problema classico dell'analisi del movimento, la *gait analysis* o analisi della camminata, alla quale sono associati modelli standard e protocolli di valutazione basata su marker; inoltre in questo caso i parametri cinematici sono ben definiti e noti, sia per controlli sani che per persone afflitte da deficit motori dovuti a problematiche ortopediche o neurologiche. Inoltre su questo dominio c'è un'ampia disponibilità di dati benchmark e dati che possono essere messi a nostra disposizione da partner clinici. In questo ambito identifichiamo tre problemi principali:

¹SEED 2018, Fondi di Ricerca di Ateneo 2018 Università degli Studi di Genova

- **O1. Studio e sviluppo di metodi per la stima del movimento senza marker.** Ci poniamo l'obiettivo di identificare elementi anatomici rilevanti in immagini e ricondurli ad un sistema di riferimento tridimensionale, andando ad analizzare la posizione e movimento degli elementi e parametri cinematici quali angoli su diversi piani di riferimento.
- **O2. Studio e sviluppo di metodi di machine learning a supporto delle decisioni cliniche.** Apprenderemo modelli del movimento di controlli sani, con lo scopo di comprendere le differenze rispetto a persone affette da deficit motori. Questo studio sarà condotto su dati attualmente disponibili provenienti dalla pratica clinica tradizionale e da metodi marker e la sua potenzialità di generalizzazione verrà estesa alle misure ottenute dai metodi marker-less sviluppati in O.1.
- **O3. Validazione dei metodi su diversi scenari applicativi nella pratica clinica.** Trasferiremo i metodi studiati nell'ambito dell'analisi della camminata ai movimenti delle braccia o delle dita, in riferimento ad appropriati deficit motori. Un'ulteriore direzione di interesse, riguarda l'analisi efficiente [Rudi *et al.*, 2017] di grandi quantità di video già disponibili nei domini clinici e sportivi e molto spesso sotto utilizzati.

3 O1: metodi e risultati preliminari

In questa sezione ci concentriamo sull'obiettivo O1, attualmente in fase di studio. La nostra analisi parte da tecniche recentemente proposte per la stima della posa 2D o 3D da sequenze video – si veda ad esempio [Insafutdinov *et al.*, 2016; Cao *et al.*, 2017; Luvizon *et al.*, 2018]. Anche se il loro potenziale applicativo al campo della biomeccanica è ancora da verificare, recenti studi riportano risultati preliminari incoraggianti. In [Mathis *et al.*, 2018] viene presentato un metodo di stima della posa ispirato a DeeperCut [Insafutdinov *et al.*, 2016] con applicazione a studi di controllo motorio e analisi di comportamento in topi. Osservazione cruciale è il fatto che la tecnica proposta, chiamata DeepLabCut, può raggiungere eccellenti prestazioni su problemi di stima della posa attraverso un training set di dimensione molto limitata (meno di 200 immagini) Questo studio è il punto di partenza della nostra ricerca. DeepLabCut utilizza una variante di deep residual network (ResNet-50) che permette l'estrazione di determinate caratteristiche dell'immagine. Queste caratteristiche possono essere controllate in fase di training e corrispondono alle feature rilevanti nell'analisi del movimento. A questo si aggiunge uno strato finale deconvoluzionale che permette di ottenere densità di probabilità spaziali.

Attualmente stiamo valutando il metodo su dati provenienti da dati di pazienti affetti da stroke e su un insieme di controlli. I dati sono costituiti da coppie di sequenze video (una frontale e una laterale) sincronizzate, a partire dalle quali stimare punti di riferimento virtuali (sul piano immagine). Per alcuni video selezioniamo 5 frame e li annotiamo manualmente, identificando i punti anatomici di interesse utili poi per l'estrazione di parametri cinematici associati al cammino (anca, ginocchio, caviglia e punta del piede). La rete viene allenata

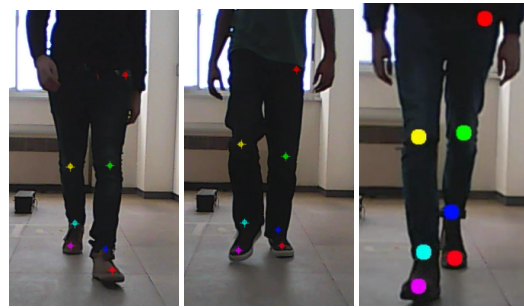


Figura 1: Rilevamento automatico feature anatomiche: a sinistra e al centro, fotogrammi di training (+ i punti annotati manualmente e le stime ottenute); a destra un frame di test

con un centinaio di fotogrammi annotati. I parametri della rete sono stati inizializzati con quelli ottenuti allenando la rete su ImageNet. I risultati preliminari mostrano un'accuratezza nella localizzazione dei punti di interesse molto promettente. A partire da questi deriviamo le posizioni dei punti nel sistema di coordinate 3D e infine deriveremo le posizioni dei giunti e le rotazioni nel sistema di coordinate anatomico. L'analisi sperimentale viene condotta in riferimento ad un gold standard già disponibile: misure derivate da 21 marker riflettenti, posizionati rispetto al protocollo Davis [Davis *et al.*, 1991] e acquisiti da un sistema mo-cap formato da 8 telecamere infrarossi. La figura 1 mostra alcuni risultati preliminari dell'architettura allenata su sequenze video acquisite in laboratorio con video di qualità medio-bassa.

Riferimenti bibliografici

- [Cao *et al.*, 2017] Z. Cao, T Simon, S Wei, e Y Sheikh. Real-time multi-person 2d pose estimation using part affinity fields. In *CVPR*, 2017.
- [Colyer *et al.*, 2018] S Colyer, M Evans, D Cosker, e A Salo. A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system. *Sports Med*, 2018.
- [Davis *et al.*, 1991] R.B. Davis, S Ounpuu, D Tyburski, e R Gage. A gait analysis data collection and reduction technique. *Human Movement Science*, 10, 1991.
- [Insafutdinov *et al.*, 2016] Insafutdinov, L Pishchulin, B Andres, M Andriluka, e B Schiele. Deeppercut: a deeper, stronger, and faster multi-person pose estimation model. In *ECCV*, 2016.
- [Luvizon *et al.*, 2018] D. C. Luvizon, D Picard, e H Tabia. 2d/3d pose estimation and action recognition using multitask deep learning. In *CVPR*, 2018.
- [Mathis *et al.*, 2018] A. Mathis, P Manidamma, K Cury, T ABe, e V Murthy. Deeplabcut: markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, 2018.
- [Rudi *et al.*, 2017] A Rudi, L Carratino, e L Rosasco. Falcon: An optimal large scale kernel method. In *Advances in Neural Information Processing Systems*, 2017.