

**UNIVERSITÀ DI PISA**

---



**FACOLTÀ DI INGEGNERIA**  
**Corso di Laurea Specialistica in Ingegneria Biomedica**

**Riassunto:**

**STUDIO E IMPLEMENTAZIONE DI METODI PER LA  
CLASSIFICAZIONE AUTOMATICA DI MOVIMENTI UMANI  
BASATA SU DATI ACCELEROMETRICI**

**CANDIDATO:**

**ANDREA MANNINI**

**RELATORI:**

**PROF. ANGELO MARIA SABATINI**

**PROF. LUIGI LANDINI**

---

**ANNO ACCADEMICO 2008/2009**

## Introduzione

Questo lavoro si pone come obiettivo lo studio di algoritmi per la classificazione automatica di posture e movimenti eseguiti da un soggetto, mediante elaborazione dei segnali provenienti da accelerometri posti in corrispondenza di determinati punti anatomici. Un sistema di classificazione automatica del movimento è di grande interesse sia in applicazioni di *pervasive computing* che richiedano la conoscenza del contesto per facilitare l'interazione uomo-macchina, sia in biomedicina, per la realizzazione di sistemi *wearable* per la valutazione a lungo-termini di parametri fisiologici e biomeccanici. In questo lavoro ci proponiamo di studiare algoritmi di classificazione *one-shot*, in cui l'esito della classificazione a un certo istante non dipende dalla storia delle classificazioni precedenti, e algoritmi di classificazione *sequenziale* (o *dinamica*) basati sugli *Hidden Markov Model* (HMM), per sfruttare la conoscenza delle statistiche di un *task* risultante dal concatenamento di singole primitive di movimento. Ad oggi sono stati realizzati sistemi di classificazione automatica del movimento che sfruttano approcci molto diversi tra loro in termini di tipologia di sensorizzazione sfruttata, *features*<sup>1</sup> calcolate dai dati e algoritmi di classificazione impiegati. Per quanto riguarda la sensorizzazione, esistono sistemi basati sul *tracking* ultrasonico [1] o sull' utilizzo di piattaforme di forza e di *marker* ottici [2]. Tuttavia la soluzione elettiva per la sua semplicità e i bassi costi, resta quella che impiega accelerometri [3, 4, 5, 6], talvolta unitamente ad altre tipologie di sensori quali giroscopi [7, 8, 9] o sensori di vibrazione [9]. Per quanto riguarda le *features* calcolate e gli algoritmi di classificazione impiegati, si può osservare che in effetti sono state percorse molte strade diverse e non è attualmente possibile affermare che esista un approccio prevalente in prestazioni rispetto agli altri. Allo stato attuale gli HMM sono utilizzati per molte applicazioni di riconoscimento automatico quali ad esempio i sistemi di riconoscimento vocale [10] e il riconoscimento dei gesti degli arti superiori, finalizzato ad applicazioni quali il riconoscimento del linguaggio dei segni [11], o il controllo attraverso il riconoscimento gestuale di sistemi informatici e robotici [12] tuttavia, non sono presenti in letteratura applicazione al riconoscimento di posture e movimenti di tutto il corpo così come sono considerati in questo lavoro di tesi.

## Metodi

Il sistema di classificazione e riconoscimento di movimenti e posture implementato in questo lavoro di tesi è basato su un dataset che è stato reso disponibile dal prof. Stephen Intille [4] attivo sulle tematiche di classificazione automatica *one-shot* mediante analisi accelerometrica. Nel dataset Bao-Intille sono raccolti dati provenienti da 5 accelerometri biassiali fissati su 20 soggetti come in Fig. 1b. Il dataset contiene l' acquisizione supervisionata di 20 movimenti, riportati in Tab. 1 tra i quali in grassetto sono evidenziati i 7 movimenti/posture più significativi, utilizzati per questo lavoro di tesi. Da tale dataset sono calcolate le *features* per la classificazione. In particolare sono state estratte da finestre temporali di 512 campioni (acquisiti a  $\sim 76$  Hz) la componente in continua, l' entropia dei coefficienti nel dominio frequenziale (STFT, finestra rettangolare), l' energia e i coefficienti di correlazione tra i 10 assi accelerometrici. In Fig. 1 è riportato un grafico riassuntivo delle *features* estratte per una specifica acquisizione. Il complesso set di *features* ricavato descrive ogni finestra di dati su uno spazio di dimensione  $d$  pari a 85, per limitare gli effetti sfavorevoli dell' elevata dimensionalità della descrizione dei dati sono applicati algoritmi di selezione delle *features*. A valle della selezione sono state valutate molte tipologie di classificazione *one-shot*, tra le strategie più comuni del riconoscimento *pattern*<sup>2</sup> statistico [15]:

- classificatori basati sulla costruzione delle densità di probabilità a posteriori delle osservazioni, condizionate rispetto alle categorie (es. classificatori bayesiani, classificatore di Parzen);
- classificatori basati sulla disposizione geometrica dei dati nello spazio delle *features* (*k-Nearest Neighbour*, *Support Vector Machines*) e reti neurali;
- classificatori basati su regole decisionali booleane (alberi decisionali).

Movimenti descritti			
<b>1 camminare</b>	6 mangiare o bere	11 fare stretching	<b>16 posizione sdraiata</b>
2 camminare con oggetti	7 guardare la TV	12 sollevare pesi (allenamento)	17 spazzolarsi i denti
<b>3 posizione seduta</b>	8 leggere	13 strofinare	<b>18 salire le scale</b>
4 lavorare ad un pc	<b>9 correre</b>	14 passare l'aspirapolvere	19 salire su una scala mobile
<b>5 posizione eretta</b>	<b>10 andare in bicicletta</b>	15 piegare la biancheria	20 prendere l' ascensore

Tabella 1. Movimenti descritti nel dataset Bao-Intille, [4]

<sup>1</sup> per *features* si intendono quelle variabili ottenute a partire dai dati, che contengono l' informazione necessaria alla classificazione e pertanto costituiscono le grandezze in ingresso al sistema di classificazione.

<sup>2</sup> un *pattern*, secondo la definizione di Watanabe [14], è "un' entità vagamente definita a cui è possibile dare un nome", mentre il riconoscimento dei *pattern* è definito da Jain come [15], "lo studio di come le macchine osservano l' ambiente, imparano a distinguere i *pattern* di interesse dal loro contesto e traggono ragionevoli decisioni circa le categorie di appartenenza dei *pattern* stessi".

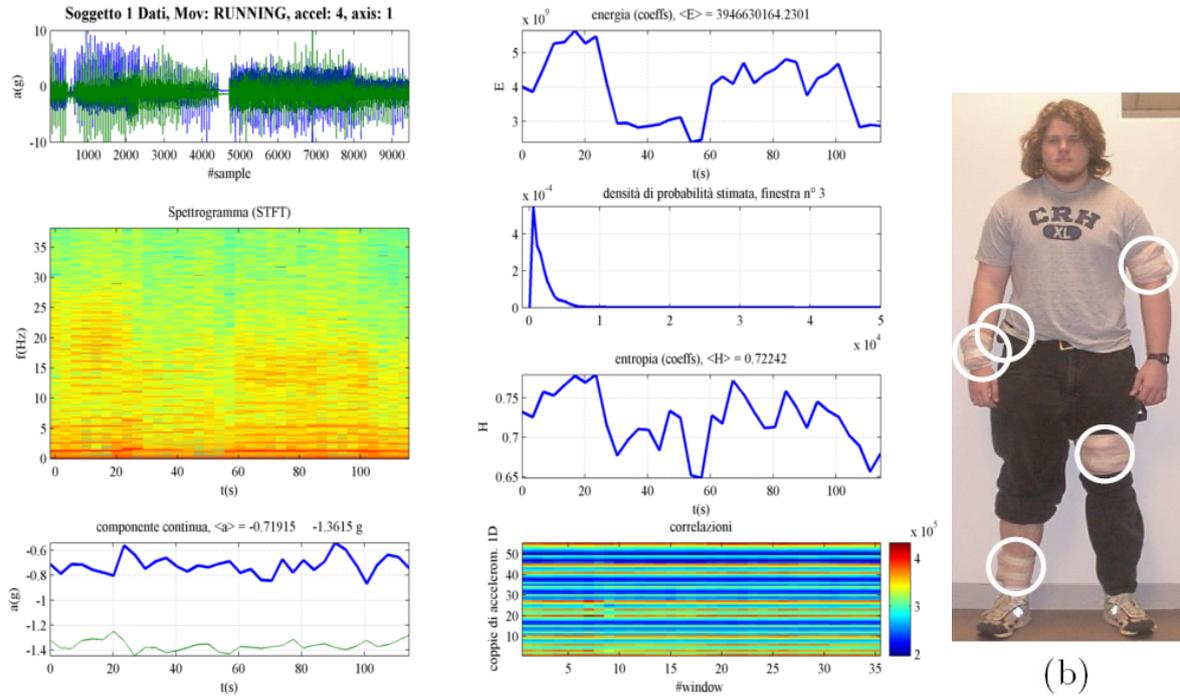


Figura 1. Rappresentazione grafica dell' andamento temporale delle *features* calcolate. Soggetto 1, movimento "corsa", accelerometro 4, asse 1. ; (b) Posizionamento dei sensori su un soggetto. Adattata da [13].

La classificazione *one-shot* è eseguita sfruttando l' approccio *leave-one-window-out* ovvero l' addestramento è eseguito su tutti i soggetti e tutti i dati, ad eccezione dei dati di ogni soggetto relativi alla finestra che si sta classificando. Questo permette di avere un *training set* molto ampio. Al termine del test la classificazione di tutte le finestre e di tutti i soggetti è aggregata per ottenere una descrizione sintetica delle proprietà di ogni algoritmo di classificazione.

Nella seconda fase dello studio si è passati alla classificazione sequenziale. In questo approccio diviene rilevante ottenere informazioni probabilistiche sulle transizioni nel tempo tra una classe di *pattern* e le altre. Inoltre, l' addestramento si è limitato ai dati di ogni singolo soggetto separatamente ed i risultati sono mediati su tutti i soggetti. Per la classificazione dinamica ci si è basati su gli HMM. In particolare è stato sfruttato un HMM a 7 stati, corrispondenti ai 7 movimenti (Fig. 2). Il modello considerato è caratterizzato da densità di probabilità di emissione continue, descritte attraverso delle *gaussian mixture*. Il dataset Bao-Intille non permetterebbe l' addestramento del modello, essendo stato acquisito con il fine di eseguire classificazioni *one-shot* dei dati. Sono state quindi create delle sequenze di dati surrogate ( $S$  sequenze di  $N$  elementi), estraendo finestre relative ai movimenti utilizzati dal dataset, secondo una matrice delle probabilità delle transizioni (TPM) tra ogni movimento e tutti gli altri fissata a priori. Le prime  $P$  sequenze costituiscono il dataset di addestramento e sono costruite utilizzando le prime  $K$  finestre, mentre le sequenze del *test set* sono ottenute utilizzando le rimanenti  $18-K$  finestre di dati. Il modello markoviano è determinato stimando i parametri:

- $\pi$ , probabilità di ogni stato di essere il primo della sequenza;
- TPM, matrice delle probabilità di transizione tra ogni stato e il successivo;
- $\mu$  e  $\Sigma$ , media e matrice di covarianza di ogni elemento delle *mixture* di emissione.

Per ovviare al problema dell' identificazione dei parametri con un dataset comunque limitato, l' addestramento è stato separato in due livelli. Nel primo si individuano le condizioni iniziali per la stima di secondo livello. Nell' addestramento di primo livello, per quanto riguarda l' identificazione di TPM e  $\pi$ , si sfrutta un *Observable Markov*

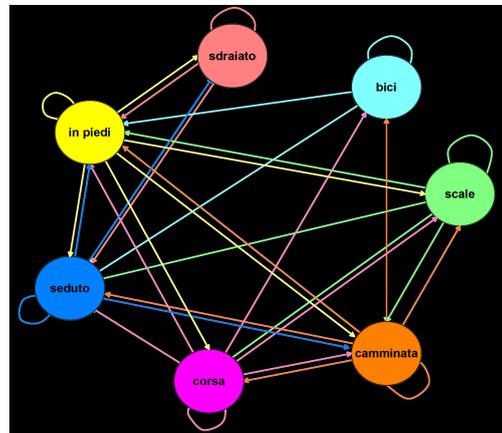


Figura 2. Schema dell' OMM progettato per la creazione delle sequenze di movimenti.

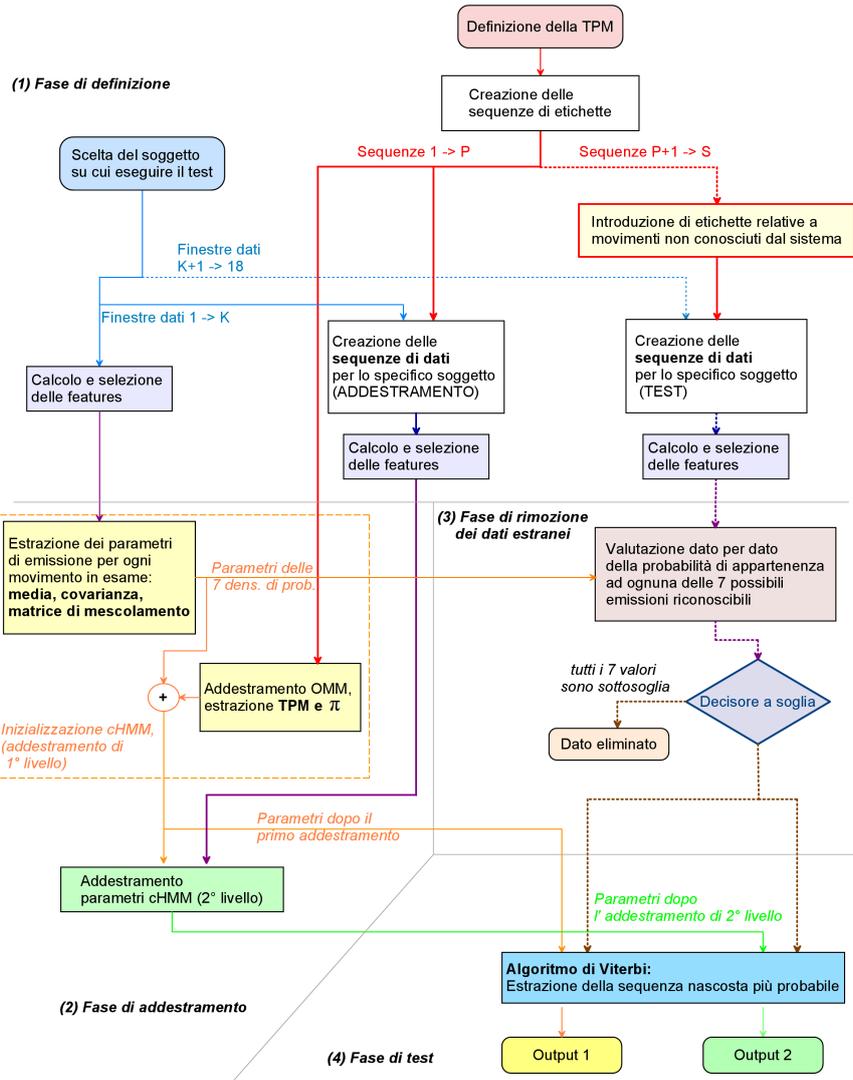


Figura 3. Schema concettuale dell' algoritmo di classificazione automatica sequenziale del movimento.

*Model* (OMM) analogo a quello di Fig. 2 che costituisce lo strato nascosto dell' HMM globale. Utilizzando le sequenze di etichette dei movimenti relative alle  $P$  sequenze di addestramento, si stimano i parametri dell' OMM attraverso un algoritmo di *Expectation Maximization*. Per quanto riguarda i parametri che descrivono le emissioni del modello, si addestrano separatamente i parametri di ogni singolo stato, utilizzando le prime  $K$  finestre relative ad ognuno dei 7 movimenti. L' addestramento di secondo livello consiste nell' applicazione dell' algoritmo di Baum-Welch alle sequenze del *training set* e fornisce una nuova stima di tutti i parametri. Utilizzando il modello addestrato è stato possibile ricavare la sequenza degli stati nascosti (e quindi dei movimenti) più probabile, mediante l' applicazione dell' algoritmo di Viterbi. I risultati sono confrontati con quelli della classificazione *one-shot* effettuata con un classificatore GMM (*gaussian mixture model*) con parametro  $M$  (numero di componenti della *mixture*) fissato allo stesso valore rispetto al numero di gaussiane di emissione del modello.

Lo studio effettuato si basa su dati surrogate, tuttavia l' applicazione in un contesto reale non può prescindere dal fatto che, oltre ai dati utili per la classificazione, si otterrebbero anche dati non classificabili, dovuti a movimenti diversi dai 7 considerati (*dati spuri*). Si è quindi introdotto nel sistema un algoritmo di rimozione dei dati spuri che valutando la probabilità di ottenere ogni singolo dato come emissione di uno dei 7 stati, elimina i dati non classificabili attraverso un decisore a soglia. Tale soglia è determinata valutando la sensibilità e la specificità della rimozione, su curve ROC. In Fig. 3 è rappresentato uno schema concettuale dell' algoritmo di classificazione automatica sequenziale descritto. In ultima analisi, per valutare la robustezza del sistema di classificazione, si è valutata la risposta ad alterazioni nella stima dei parametri, in particolare della TPM. Per quantificare l' alterazione introdotta su ogni riga della TPM rispetto alla stima derivante dall' addestramento di primo livello è stata utilizzata la *distanza di Kullback-Liebler (KL)*. Si sono quindi studiate le percentuali di classificazione medie ottenute per diversi gradi di alterazione della TPM, utilizzando un livello di addestramento o entrambi.

## Risultati e Discussione

Per quanto riguarda la rappresentazione dei dati, attraverso algoritmi di selezione sequenziale (*sequential forward selection* SFS, *sequential forward floating selection* SFFS) sono ricavate descrizioni dei dati su spazi di dimensione  $d$  pari a 56 e 17 (Fig. 4). Le prestazioni dei diversi classificatori *one-shot* sono state valutate in base alle matrici di confusione ottenute ed in termini di sensibilità e specificità della classificazione. In Tab. 2 sono riportati i risultati ottenuti con i classificatori kNN, Parzen e SVM che hanno permesso i risultati migliori.

<i>Features</i>	$P_{parzen}$	$P_{SVM}$	$P_{k-NN}$
56	99.33%	94.09%	<b>99.72%</b>
17	99.60%	99.44%	<b>99.80%</b>

Tabella 2: Risultati della classificazione *one-shot* per i due set di *features*.

La classificazione dinamica si è basata su dataset di addestramento più limitati, dovendo considerare una classificazione soggetto per soggetto. Tale scelta causa inevitabilmente un abbassamento delle prestazioni della classificazione dovuta al fatto che il *training-set* è più limitato. In Tab. 3 è riportato il confronto tra classificatore dinamico e GMM *one-shot*, basato sulla stessa configurazione di emissione. Il confronto  $P_{HMM}$  Vs.  $P_{GMM}$  mostra l'influenza delle informazioni disponibili sulla dinamica del movimento a livello di TPM. Un elemento critico nel progetto di un classificatore è il valore del rapporto  $n/d$  ( $n$ : numerosità del *training-set*) [15]: le prestazioni dei classificatori studiati in questo lavoro sono sorprendentemente buone, in relazione al valore modesto del rapporto  $n/d$ , che ad esempio per  $K = 7$  è pari a 49/17. Il sistema di rimozione dei dati spuri riesce nella rimozione con sensibilità del 96.4% e specificità del 93.6% (Fig. 5), e riporta la percentuale di riconoscimento a valori prossimi al 95%, partendo da una degradazione netta di questa.

Valutando l'andamento della percentuale di classificazione al variare dell'alterazione introdotta sulla TPM in termini di distanza KL si ottengono andamenti analoghi a quelli riportati in Fig. 6. Limitando l'addestramento al primo livello la degradazione della TPM riduce la percentuale di riconoscimento più velocemente rispetto al caso in cui siano seguiti entrambi i livelli di addestramento. L'addestramento di secondo livello quindi permette di recuperare, seppur parzialmente l'errore introdotto alterando la TPM.

	$P_{GMM}$ ( <i>one-shot</i> )	$P_{HMM}$ (1 liv.)	$P_{HMM}$ (2 liv.)
$M = 1,$ $K = 3$	74.95%	<b>95.01%</b>	80.58%
$M = 1,$ $K = 7$	91.62%	<b>94.94%</b>	80.25%
$M = 1,$ $K = 15$	<b>98.10%</b>	94.42%	80.02%

Tabella 3: Confronto tra percentuale di corretta classificazione *one-shot* GMM e classificazione dinamica HMM, addestramento su singolo soggetto.

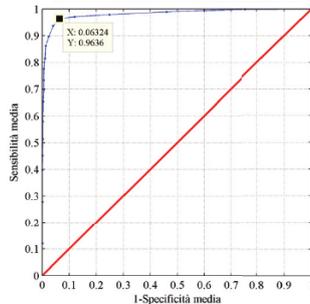


Figure 5. Curva ROC della rimozione di 100 spuri per ognuna delle 20 sequenze di 300 dati: media dei risultati ottenuti con movimenti spuri diversi

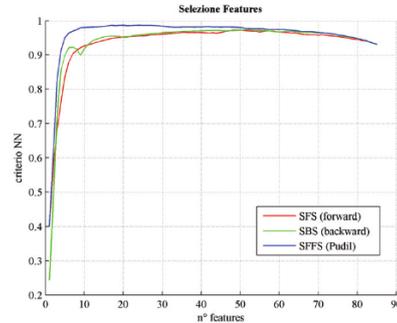


Figure 4. Applicazione degli algoritmi di selezione sequenziale delle *features* sul dataset utilizzando la classificazione  $k$ -NN come criterio.

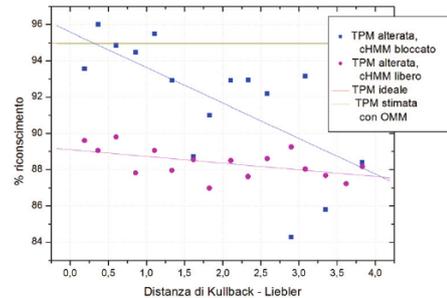


Figure 6. Alterazione della TPM. Esito e regressione lineare dei risultati ottenuti eseguendo (magenta) o meno (blu) l'addestramento di 2° livello. In rosso e verde sono indicati i risultati in caso di TPM ideale ed in caso di stima non alterata.

## Conclusioni e Sviluppi Futuri

Lo studio svolto ha permesso di raggiungere gli obiettivi prefissati. Infatti, in primo luogo è stato possibile ottenere degli algoritmi di classificazione automatica del movimento basati su osservazioni accelerometriche.

Il secondo obiettivo fondamentale ottenuto è stato l' addestramento di un HMM con emissioni continue in grado di effettuare una classificazione di tipo "dinamico", di sequenze di dati accelerometrici anziché di singoli dati. Lo sviluppo inedito di un HMM per la classificazione di posture e movimenti umani completi è stato inoltre esteso realizzando un sistema di rimozione di dati non classificabili (ovvero estranei ai movimenti appresi in fase di addestramento), caratterizzato da elevata sensibilità e specificità nella rimozione di questi.

Poiché tutto lo studio sugli HMM è eseguito su un dataset surrogato, è auspicabile in futuro di poter applicare la trattazione descritta ad un set di dati sperimentali reale. La realizzazione di una campagna sperimentale sufficiente a raccogliere i dati necessari non è difatti proibitiva come si riteneva, in quanto suddividendo il problema dell' identificazione dei parametri del modello in più sotto-problemi si riesce ad addestrare il modello anche con un dataset dimensionalmente molto limitato.

Le applicazioni di sistemi di classificazione del movimento di questo tipo sono molteplici ed in modo particolare sono destinate a divenire parte integrante di sistemi di *intelligenza ambientale e computing ubiquitario* (o *pervasivo*). L' informazione associata al movimento che il soggetto in esame sta compiendo o alla postura può infatti costituire un' importante risorsa per la realizzazione sistemi intelligenti. Ad esempio si può pensare ad applicazioni robotiche "smart" che sfruttino la classificazione automatica del movimento e della postura dell' utilizzatore. Tale informazione permette infatti un incremento sia nelle possibilità di controllo del dispositivo robotico - che diviene in grado di seguire fisicamente l' utilizzatore e comportarsi diversamente in funzione della postura o del movimento di questo - sia nel monitoraggio dell' utilizzatore stesso. Infatti la classificazione ottenuta permette di variare i parametri di sistemi di monitoraggio dei parametri vitali o biomeccanici, quali ad esempio il consumo energetico in funzione dello specifico movimento in atto.

## Riferimenti bibliografici

- [1] M.-C. Su, Y.-Y. Chen, K.-H. Wang, C.-Y. Tew, and H. Huang, "3d arm movement recognition using syntactic pattern recognition," *Artificial Intelligence in Engineering*, vol. 14, pp. 113–118, 2000.
- [2] R. Begg and J. Kamruzzaman, "A machine learning approach for automated recognition of movement patterns using basic, kinetic and kinematic gait data," *Journal of Biomechanics*, vol. 38, pp. 401–408, 2005.
- [3] F. R. Allen, E. Ambikairajah, N. H. Lovell, and B. G. Celler, "Classification of a known sequence of motions and postures from accelerometry data using adapted gaussian mixture models." *Physiol Meas*, vol. 27, pp. 935–951, Oct 2006.
- [4] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in *Pervasive*, 2004, pp. 1–17.
- [5] M. Sekine, T. Tamura, M. Akay, T. Fujimoto, T. Togawa, and Y. Fukui, "Discrimination of walking patterns using wavelet-based fractal analysis," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 10, no. 3, pp. 188–196, Sept. 2002.
- [6] K. Aminian, P. Robert, E. E. Buchser, Rutschmann, D. Hayoz, and . Depairon, "Physical activity monitoring based on accelerometry: validation and comparison with video observation," *Medical and Biological Engineering and Computing*, vol. 37 (1), pp. 304–308, 1999.
- [7] B. Najafi, K. Aminian, A. Paraschiv-Ionescu, F. Loew, C. Bula, and P. Robert, "Ambulatory system for human motion analysis using a kinematic sensor: Monitoring of daily physical activity in the elderly," *Ieee Transactions on Biomedical Engineering*, vol. 50, no. 6, pp. 711–723, 2003.
- [8] S.-W. Lee and K. Mase, "Activity and location recognition using wearable sensors," *Pervasive Computing, IEEE*, vol. 1, no. 3, pp. 24–32, 2002. [Online]. Available: [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=1037719](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1037719)
- [9] N. Noury, "A smart sensor for the remote follow up of activity and fall detection of the elderly," in -, 2002, pp. 314–317.
- [10] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989. [Online]. Available: <http://dx.doi.org/10.1109/5.18626>
- [11] R. Liang and M. Ouhyoung, "A real-time continuous alphabetic sign language to speech conversion VR system," *Computer Graphics Forum*, vol. 14, no. 3, pp. 67–76, 1995. [Online]. Available: [citeseer.ist.psu.edu/liang95realtime.html](http://citeseer.ist.psu.edu/liang95realtime.html)
- [12] J. Yang, Y. Xu, and C. Chen, "Human action learning via hidden markov model," *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, vol. 27, no. 1, pp. 34–44, Jan 1997.
- [13] L. Bao, "Physical activity recognition from acceleration data under semi-naturalistic conditions," Master's thesis, Department of electrical engineering and computer science, MIT, 2003.
- [14] S. Watanabe, *Pattern Recognition: Human and Mechanical*, S. Watanabe, Ed. Wiley, 1985.
- [15] A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: A review," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 1, pp. 4–37, January 2000. [Online]. Available: <http://dx.doi.org/10.1109/34.824819>