

# Az emberi mozgás követése és felismerése

## Human motion tracking and recognition

VAJDA Tamás

Sapientia Erdélyi Magyar Tudományegyetem  
Műszaki és Humántudományok Kar Marosvásárhely, ROMÁNIA  
vajdat@ms.sapientia.ro

### Abstract

*From the entertainment industry to healthcare, there is a need to track and recognize human movement. As a result, a lot of scientist are researching and developing methods to track and recognize human motion. In this article, we present the most relevant methods used to track human movement and discuss their advantages and disadvantages.*

### Kivonat

*A szórakoztató ipartól egészen az egészségügyig szükség van arra, hogy az emberi mozgást kövessük és felismerjük. Ennek eredményeképpen rengeteg kutató foglalkozik olyan módszerek kifejlesztésével, amelyeknek a célja az emberi mozgás követése és felismerése. Ebben a cikkben bemutatjuk azokat a főbb módszereket, amelyeket emberi mozgás követésére használunk, valamint kitérünk ezeknek a hátrányaira és előnyeire.*

**Kulcsszavak:** mozgás követése, inerciális szenzor, képfeldolgozás

### 1. Bevezető

A fejlesztéseknek egyik legfontosabb célkitűzése, hogy az emberi munkát a lehető legjobban leegyszerűsítse. Mivel legtöbb tevékenységben az ember kulcsszerepet játszik magától adódik, hogy szükség van egy olyan módszerre, amely követi és felismeri az emberi tevékenységeket, viselkedésformákat. Az emberi viselkedés felismerésére szükség van a szórakoztató iparban, sportban, egészségügyi szolgáltatásokban, megfigyelő rendszerekben [1, 2].

Minden egyes alkalmazási területnek megvan a maga sajátossága emiatt különböző módszereket fejlesztettek ki a kutatók, hogy a lehető legjobb eredményeket tudják elérni az emberi mozgás felismerésében.

Ahhoz, hogy az emberi mozgást viselkedést fel tudjuk ismerni szükségünk van először is adatra, amelyek leírják az emberi mozgást térben és időben. Tulajdonképpen ez az adat begyűjtés meghatározza, hogy mennyire lehet pontos a viselkedés vagy a mozgás felismerése. Az adatok gyűjtésének a módszerét meghatározza az adott alkalmazás típusa, az adott helyzet megszorításai. Az adatok gyűjtésére videokamerát, valamilyen viselhető szenzorokat vagy ezek együttesét használjuk. Az elmúlt évtizedek domináns emberi mozgás követő rendszerei vizuális követően rendszerek voltak [1, 2]. A vizuális rendszerek tulajdonképpen a megfigyelő rendszerek továbbfejlesztése volt, ahol az emberi munkát egy algoritmussal helyettesítettük. Ezekben az esetekben adott volt kamerarendszer és ezt bővítettük ki egy szoftveres rendszerrel. A vizuális rendszerek legfontosabb előnye, hogy nem szükséges az egyén közvetlen hozzájárulása, egyidejűleg sok egyént lehet követni emiatt rendszerek olcsók voltak. Jelenleg mivel a viselhető szenzorok árai nagyon lecsökkentek egyre elterjedtebbek az olyan módszerek ahol az emberek mozgását viselhető szenzorok segítségével követjük. Ezek a rendszerek feltételezik a megfigyelt egyén aktív hozzájárulását, a követéshez azáltal, hogy hajlandó viselni ezeket az eszközöket.

## 2. Vizuális követés

A vizuális rendszerek nyers adatait a kamerák szolgáltatják. Ezeket az adatokat videó szekvenciáknak nevezzük, amelyek tartalmazzák a kamerák által látott objektumok térbeli illetve időbeli mozgását. A legnagyobb kihívás ezekben a rendszerekben, hogy az adatokból pontosan nyerjük ki az információt, jelen esetben tudjuk meghatározni hogy hol található az ember és az emberi testrészek az adott képkockán. Hogy ezt hogyan valósítjuk meg azt a módszer határozza meg. Leggyakoribb módszerek először megpróbálják leszűkíteni az osztályozandó adatot ezt nevezzük detektálásnak. A második lépés, hogy ezeket a megszürt adatokat osztályozzuk és meghatározzuk, hogy embert vagy emberi testrészt tartalmaznak. Ezt megismételjük egymás után az összes képkockára így kapjuk meg a mozgás szekvenciákat, amelyet a későbbiekben egy újabb osztályozó fogja feldolgozni és meghatározni a hozzá tartozó mozgást vagy viselkedést.

### 2.1. Az emberek észlelése

Általában ez egy opcionális lépés, aminek a célja az, hogy csökkentjük annak az adatmennyiséget, amelyet az osztályozó fel kell, dolgozzon. A leggyakrabban abból indulnak ki a kutatók, hogy az adatgyűjtő egység egy rögzített kamera. Ebben az esetben a háttér egy konstans adat lesz, amelyik nem, vagy csak kicsit változik az egymás utáni képkockákban. A másik feltevés, amelyet használni szoktak a detektáláshoz, hogy két egymás utáni képkockában az objektumok pozíciói csak keveset változnak [6].

Leggyakrabban használt módszerek [3, 8, 10]

- képkockák közötti eltérés módszere
- optikai folyam
- háttér kivonás

### 2.2. Az objektumok osztályozása

Az objektum osztályozása során eldöntjük a bemenetről, hogy embert vagy emberi testrészt-e vagy nem. Rendszerektől függően a bemenet az lehet a detekció eredménye vagy minden egyes képrészlet képből [6].

Az osztályozó algoritmusok bemeneti tulajdonságok alapján lehetnek [6, 7]:

- Forma alapú
- Szín alapú
- Mozgás alapú
- Textura alapú

A bemeneti módszer feldolgozása alapján [4, 6]:

- Statisztikai
- Mesterséges intelligencia alapú módszerek

### 2.3 Objektum követése

Az objektum követésnek a célja, hogy egymás követő képkockákban szegmentálni tudjuk a cél objektumot a kép többi részétől. a követés az egy bonyolult feladat mivel fennáll az információ veszteség vagy hiány dinamikus forma zaj takarás.

Az objektum követési módszerek lehetnek:

*-Pont alapú*

A pont alapú módszerek esetében a követést egy-egy tulajdonság követésre vezeti vissza[3,4, 5].

A leggyakrabban alkalmazott módszerek:

- Kálmán szűrő
- Rézecskeszűrő
- Multi-hypotézis szűrő

*-Mag (Kernel) alapú*

Mag alapú módszerek esetében a pont jellegű tulajdonságok helyett forma alapú tulajdonságokat követünk amelyek egyszerű kernelekkel leírhatók. például ellipszis szögletes tulajdonságok [9].

Leggyakrabban alkalmazott módszerek:

- egyszerű mint a találatás
- Mean shift módszer
- Support Vector taltatás
- réteg alapú találatás

*-Körvonal alapú*

Sok esetben az objektumok nem írhatók le egyszerű formákkal ilyen esetekben szokták használni a körvonal alapú követést [6].

Leggyakrabban alkalmazott módszerek

- körvonal találatás
- forma találatás

### 3. Viselhető szenzor alapú követés

Annak köszönhetően az elmúlt években jelentő fejlődés történt szenzorok területén, amelyek kisebbek és olcsóbbak lettek, a kutatók figyelmé a hordozható szenzorok felé irányulnak az emberi mozgás követés területén. A szenzorok segítségével nagyon sok a vizuális követés esetén megoldhatatlan problémát ki lehet küszöbölni, ilyen probléma, például amikor egy végtag takarásban van. A hordozható szenzorokat általában szenzor hálózatba használják és így adnak lehetőséget arra, hogy részletesen követni tudjuk test mozgását. A módszer legnagyobb hátránya, hogy a követendő egyénnek viselnie kell ezeket a szenzorokat [12, 13, 14]

Leggyakrabban használt hordozható szenzorok az inerciális szenzorok. Egy inerciális szenzor csak egyetlen emberi testrész mozgását tudja követni ezért ahhoz, hogy megkapjuk egy teljes emberi test mozgását szükségünk van legalább 9-10 szenzorra van szükségünk, de alkalmazástól függően sok mindent meg lehet oldani kevesebb szenzor segítségével is. A felhasznált szenzorok száma attól függ, hogy mennyire részletesen akarom követni a különböző testrészek mozgását például egy járásnak az elemzéséhez nincs szükségünk a kar mozgásának követésére így 4 szenzorral kevesebb szükséges. Általában az inerciális szenzorok mellett magnetométert [11] is használunk, ami segít abban, hogy megbecsüljük az inerciális szenzor orientációját és elhelyezzük a valós világ koordináta rendszerében.

Az adatok elemzésében és a helyességük ellenőrzésében az úgynevezett exoskeleton segít. Az exoskeleton az emberi test kinetikus modellje, amely tartalmazhatja különböző szenzorok elhelyezkedését valamint megszorításokat, amely meghatározza a lehetséges elmozdulásokat. Exoskeletonot nemcsak a hordozható szenzorok esetében használunk, hanem nagyon sok vizuális követési rendszer is alkalmazza.

#### 3.1 Az emberi test kinematikai modellje

Minden inerciális szenzor alapú követő rendszer feltételezi, hogy létezik egy merev kinetikai test modell. Legegyszerűbb modelleket arra használjuk, hogy rögzítsük a szenzorok egymáshoz való viszonyulását. Azokban az esetekben amikor több testrészt követünk akkor általában kinetikus lánc modelleket használunk. Általában a modell minden egyes elemét külön követjük szenzorral. Kinetikus lánc jobban kiaknázható azáltal, hogy a kapcsolódási pontokhoz hozzárendelhetünk megszorításokat. Ezek a megszorítások segítenek abban, hogy a követett mozgás megfeleljen a valóságban annak ellenére, hogy nagyon sok esetben a szenzorok által begyűjtött adott zajos vagy adott esetben hibás. A kinematikai modellt legtöbbször Cartesian térben ábrázolják. A kinematikai modell elemei közötti elfordulásokat Euler szögek segítségével írjuk le. Ez egy intuitív megközelítés mivel a szögek fizikai jelentéssel bírnak, viszont probléma szabadságfok elvesztése (Gimbal lock) vagyis szingularitások léphetnek fel ebben a modellben. Egy másik megközelítés amikor az Euler szögek helyett quaternionokat használunk. A quaternionok felhasználásával megalkotott modell nem szenved a szingularitási problémáktól viszont nem jeleníthető meg egy az egyben a fizikai modell. [16, 17, 18]

Kinematikai megszorítások egyes esetekben a szenzor fúziós algoritmusokban vannak beépítve de ugyanúgy találkozunk olyan megoldásokkal amikor fúzió után alkalmazzuk a megszorításokat.

Szenzor fúziós algoritmusok eredménye a legtöbb esetben random változó és ezekre megszorításokat alkalmazni vagy limitálni ezek értékeit nagyon kényes feladat.[19, 20, 21, 22]

A szabad szegmensek modellje egy olyan model amely rendelkezik kétféle típusú megszorítással: vannak kemény megszorítások ilyen például a különböző végtagok egymáshoz való kapcsolódása és vannak gyenge megszorítások mint például a könyökre vonatkozó megszorítások, mivel a könyök nem egy tökéletes csuklózület.

#### 4 Következtetések

Az emberi mozgás követése és felismerése egy olyan terület, amelyet folyamatosan intenzíven kutattak az elmúlt három évtizedben és kutatnak ma is annak köszönhetően, hogy nagyon nagy az alkalmazási lehetősége. Annak ellenére, hogy számos kutató dolgozik a feladaton még számos olyan terület van valami fejlesztésre és kutatásra szorul.

Mind a két megközelítés, amelynek a célja az emberi mozgás követése és felismerésére megvan a jogosultsága, alkalmazási területe. A vizuális követés legnagyobb alkalmazási területe az a megfigyelő rendszerek. Ennek előnye, hogy nem szükséges a megfigyelt egyéneknek szenzort viselniük viszont az egyik legnagyobb problémája, ha valami eltakarja megfigyelt személyt. A takarás kiküszöbölésének egyik megoldása, az hogy több kamerát szerelnek fel, viszont a kamerák számával növekszik a feldolgozandó adatmennyiség. A vizuális követés másik hátránya, hogy a feldolgozandó adatmennyisége sokkal nagyobb, mint a szenzor alapú követés esetében. A hordozható szenzor alapú mozgás követés legnagyobb előnye az, hogy pontosabban tudja követni az emberi test mozgását sokkal kisebb a feldolgozandó adatmennyiség és nem szenved a takarás problémájától. A megközelítés legnagyobb hátránya az, hogy a megfigyelt egyének viselniük kell a szenzort, aminek a felhelyezése nagyon sok esetben elég körülményes. A másik hátránya az, hogy szükség van vezeték nélküli hálózatokra amelyek vételezik és továbbítják a szenzor háló által küldött adatokat.

#### IRODALOM

- [1] Fortino, G.; Giannantonio, R.; Gravina, R.; Kuryloski, P.; Jafari, R. Enabling effective programming and flexible anagement of efficient body sensor network applications. *IEEE Trans. Hum.-Mach. Syst.* 2013, 43, 115–133.
- [2] Chen, M.; Gonzalez, S.; Vasilakos, A.; Cao, H.; Leung, V.C. Body area networks: A survey. *Mob. Netw. Appl.* 2011, 16, 171–193.
- [3] Himani S. Parekh, Darshak G. Thakore, Udesang K. Jaliya “A Survey on Object Detection and Tracking Methods”, *International Journal of Innovative Research in Computer and Communication Engineering*, Vol. 2, Issue 2, February 2014.
- [4] P. Bilinski, F. Bremond, and M. Kaaniche. Multiple object tracking with occlusions using HOG descriptors and multi resolution images. In *The International Conference on Imaging for Crime Detection and Prevention (ICDP)*, London, UK, 2009.
- [5] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *The Proceeding of The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 511–518, April 2003.
- [6] Shraddha Kothiya Kinjal, Mistree Kinjal Mistree A review on real time object tracking in video sequences, *Conference: 2015 International Conference on Electrical, Electronics, Signals, Communication and Optimization (EESCO) 2015*
- [7] Mrinali M. Bhajibhakare, Pradeep K. Deshmukh, To Detect and Track Moving Object for Surveillance System, *International Journal of Innovative Research in Computer and Communication Engineering*, Vol. 1, Issue 4, June 2013.
- [8] Nishu Singla Motion Detection Based on Frame Difference Method *International Journal of Information & Computation Technology*. ISSN 0974-2239 Volume 4, Number 15 (2014)
- [9] Rahul Mishra, Mahesh K. Chouhan, Dr. Dhiiraj Nitnawwre, “Multiple Object Tracking by Kernel Based Centroid Method for Improve Localization”, *International Journal of Advanced Research in Computer Science and Software Engineering*, July-2012, pp 137-140.
- [10] Denis Fortun, Patrick Bouthemy, Charles Kervrann. Optical flow modeling and computation: a survey. *Computer Vision and Image Understanding*, Elsevier, 2015, 134, pp.21
- [11] Zhang Z.Q., Wu J.K. A novel hierarchical information fusion method for three-dimensional upper limb motion estimation. *IEEE Trans. Instrum. Meas.* 2011;60:3709–3719.
- [12] Ahmad N., Ghazilla R.A.R., Khairi N.M., Kasi V. Reviews on various inertial measurement unit (IMU) sensor applications. *Int. J. Signal Proc. Syst.* 2013;1:256–262. doi: 10.12720/ijsp.1.2.256-262.

- [13] Patel S., Park H., Bonato P., Chan L., Rodgers M. A review of wearable sensors and systems with application in rehabilitation. *J. Neuroeng. Rehabil.* 2012;9:21. doi: 10.1186/1743-0003-9-21
- [14] Gravina R., Alinia P., Ghasemzadeh H., Fortino G. Multi-sensor fusion in body sensor networks: State-of-the-art and research challenges. *Inf. Fusion.* 2017;35:68–80. doi: 10.1016/j.inffus.2016.09.005
- [15] Picerno P., Cereatti A., Cappozzo A. Joint kinematics estimate using wearable inertial and magnetic sensing modules. *Gait Posture.* 2008;28:588–595. doi: 10.1016/j.gaitpost.2008.04.003
- [16] To G., Mahfouz M.R. Quaternionic Attitude Estimation for Robotic and Human Motion Tracking Using Sequential Monte Carlo Methods with von Mises-Fisher and Non Uniform Densities Simulations. *IEEE Trans. Biomed. Eng.* 2013;60:3046–3059. doi: 10.1109/TBME.2013.2262636.
- [17] Wang J.J., Wang J., Sinclair D., Watts L. A neural network and Kalman filter hybrid approach for GPS/INS integration; Proceedings of the 12th IAIN World Congress, 2006 International Symposium on GPS/GNSS; Jeju, Korea. 18–20 October 2006; pp. 18–20.
- [18] Pons-Moll G., Baak A., Gall J., Leal-Taixe L., Muller M., Seidel H., Rosenhahn B. Outdoor human motion capture using inverse kinematics and von mises-fisher sampling; Proceedings of the 2011 IEEE International Conference on Computer Vision (ICCV); Barcelona, Spain. 6–13 November 2011; pp. 1243–1250.
- [19] Lunge H.J., Veltink P.H., Baten C.T. Ambulatory measurement of arm orientation. *J. Biomech.* 2007;40:78–85. doi: 10.1016/j.jbiomech.2005.11.011
- [20] Peppoloni L., Filippeschi A., Ruffaldi E., Avizzano C.A. A novel 7 degrees of freedom model for upper limb kinematic reconstruction based on wearable sensors; Proceedings of the 2013 IEEE 11th International Symposium on Intelligent Systems and Informatics (SISY); Subotica, Serbia. 26–28 September 2013; pp. 105–110.
- [21] Cutti A.G., Giovanardi A., Rocchi L., Davalli A., Sacchetti R. Ambulatory measurement of shoulder and elbow kinematics through inertial and magnetic sensors. *Med. Biol. Eng. Comput.* 2008;46:169–178. doi: 10.1007/s11517-007-0296-5
- [22] Zhang Z.Q., Wong W.C., Wu J.K. Ubiquitous human upper-limb motion estimation using wearable sensors. *IEEE Trans. Inf. Technol. Biomed.* 2011;15:513–521. doi: 10.1109/TITB.2011.2159122.
- [23] Mihelj M. Inverse kinematics of human arm based on multisensor data integration. *J. Intell. Robot. Syst.* 2006;47:139–153. doi: 10.1007/s10846-006-9079-8.