

Mozgó személyek követése és 4D vizualizációja Lidar-alapú járáselemzéssel

Nagy Balázs¹, Benedek Csaba¹ és Jankó Zsolt²

¹ Elosztott Események Elemzése Kutatólaboratórium, Magyar Tudományos Akadémia, Számítástechnikai és Automatizálási Kutatóintézet (MTA SZTAKI), 1111, Budapest, Kende utca 13–17, email: {vezetéknév.keresztnév}@sztaki.mta.hu

² Geometriai Modellezés és Számítógépes Látás Kutatólaboratórium, MTA SZTAKI
email: janko.zsolt@sztaki.mta.hu

Absztrakt. A következő tanulmányban Lidar szenzor méréseire alapú járáselemző módszert mutatunk be. A kifejlesztett algoritmus egy 4D vizualizációs keretrendszer része, ahol kültéri környezetben játszódó többszereplős valós jeleneteket elemezhetünk és jeleníthetünk meg interaktívan. A követett személyek trajektóriáinak meghatározása után a rendszer úgynevezett szabad nézőpontú videót hoz létre, amely megjeleníti a geometriailag rekonstruált és textúrázott jelenetet. A megjelenítés során biztosítjuk, hogy az előzőleg egy 4D rekonstrukcióban felvett mozgó avatárjaink kövessék a valóságban megfigyelt szereplők mért útpályáit. A rendszer két különböző moduljában támaszkodunk az emberi járás elemzésével kapható jellemzőkre is. Az egyik feladat a követés során kinyert egyedi járáslenyomatok, mint biometrikus jellemzők felhasználása a járókelők újra-azonosítására, amennyiben időlegesen elhagyták a megfigyelt területet, majd újra feltűntek a szintéren. A másik feladat, hogy a vizuális megjelenítést valósághűbbé téve szinkronizáljuk a megfigyelt szereplők és a mozgatott avatárak lépéseit. Eljárásunkat valós kültéri környezetekben rögzített jeleneteken teszteljük.

1. Bevezetés

Számos alkalmazási területen központi feladat olyan 3D dinamikus jelenetek elemzése, ahol egyszerre több személy mozog egy megfigyelt helyszínen. A probléma többek között megjelenik az intelligens videofelügyelet [1], a videokommunikáció és a kiterjesztett valóság területein is. Egy vizuális méréseken alapuló komplex jelenetfeldolgozó rendszer számos funkciót valósít meg, kezdve a mozgó emberek észlelésétől, a lokalizációtól és a követésen át, egészen a magasabb szintű események felismeréséig, például abnormális események detekciójáig, és a jelenetek hatékony megjelenítéséig.

A videofelügyeleti alkalmazások egyik központi problémája az időlegesen eltűnt szereplők újra-felismerhetősége a követés során. A személyek trajektóriáinak megszakadását okozhatja többek között az alakzatok kölcsönös kitakarása, illetve gyakran előfordul, hogy a célszemély elhagyja a megfigyelt területet, majd később a szintér egy másik pontján újból megjelenik. Az emberek újra-azonosításához [2] biometrikai jellemzőket kell kinyernünk, ezek azonban úgynevezett *gyenge leírók* is lehetnek: segítségükkel elegendő az azonosítást a helyszínen lévő emberek tipikusan kis csoportján

belül elvégezni, nem cél keresés nagy adatbázisokban. Ugyanakkor a biometrikus jellemzőket a személyek természetes viselkedése közben kell lemérni, a felismerést közel valós időben megvalósítva.

A járás, mint biometrikus jellemző felhasználhatóságát széleskörűen elemezték az elmúlt évtizedekben [3–5]. Már az 1960-as évek óta pszichológiai kísérletek igazolják, hogy a legtöbb embert hatékonyan lehet azonosítani a járása alapján [6]. A videó alapú járásfelismerő modulok egyszerűen integrálhatók a megfigyelő rendszerekbe, hiszen a méréshez nincs szükség speciális mérőeszközökre, és a járókelők aktív együttműködését sem igénylik az azonosítási folyamatban. Bár korábban számos járás alapú személyazonosítási módszert javasoltak a szakirodalomban (részletek a 1.1. fejezetben következnek), a legtöbb megoldást csak erősen kontrollált tesztkörnyezetben vizsgálták, ahol az alanyok sétáit előbb egyenként rögzítették, majd a felvett járásmintákat utólag párosították. Valós videofelügyelet során fő kihívás, hogy a járásjellemzőket természetes környezetben kell kinyerni, a helyszínen egyidejűleg több gyalogos is jelen lehet, akik időlegesen takarásba kerülhetnek egymással vagy más tereptárgyakkal. Ahhoz, hogy megőrizzük a rendszer valós idejű feldolgozó képességét, a személyek azonosítását a megfigyelt jelenet közben el kell végezni, miközben tetszőleges gyakorisággal megjelenhetnek új illetve korábban is látott emberek.

A 4D (dinamikus 3D) jeleneteknél további kihívást jelent a mérési eredmények hatékony megjelenítése. Hagyományos videofelügyeleti rendszerekben az operátorok különböző kameraképek párhuzamos figyelésével nagy területet tekintethetnek át, és a takarási problémákat részben megoldja a különböző nézetek együttes vizsgálata. Ugyanakkor a megoldás hatékonysága a kamerák számának növekedésével drasztikusan csökken, mert a strukturálatlan képfolyamok együttes követése megterhelő feladat. A valóságot megjelenítő 4D videoszekvenciák kiterjesztett vizuális élményt nyújthatnak a megfigyelőnek, aki a rekonstruált 4D jeleneteket tetszőleges nézőpontból megtekintheti és elemezheti, igény esetén pedig virtuálisan módosíthatja is. Bár léteznek sztereó-alapú megoldások 4D kültéri jelenetek rekonstrukciójára [7], a feldolgozási folyamat itt nem teljesen automatikus, és létrehozásuk a mai hardvereszközökkel is extrém számítási időt és kapacitást igényel.

Az integrált 4D (i4D) megfigyelő és vizualizációs rendszer [8] dinamikus 3D jelenetek részben valós idejű rekonstrukciójára alkalmas. Az i4D rendszer a feldolgozás során két különböző típusú méréshalmazt használ fel. Az első forrás 4D pontfelhő szekvencia, melyet egy forgó többnyalábos (Rotating Multi-beam, RMB) Lidar szenzor rögzít. A második rendszerkomponens egy beltéri 4D rekonstrukciós stúdió, [9] ahol mozgó emberekről készülnek élethű 4D modellek. A megfigyelő rendszer képes egyszerre több gyalogost detektálni és követni a célterületen, és a felvett jeleneteket 4D-ben megjeleníteni. A rendszer kimeneteként a geometriailag rekonstruált és textúrázott környezetben 4D avatar modellek mozognak, melyek valós időben követik a valóságban megfigyelt járókelők útpályáit.

Az i4D rendszer eredeti verziójának [8] két fő korlátja volt. Egyrészt a kizárólag *rövidtávú* objektumkövetést tett lehetővé, ezért amikor megszakadt egy követett alakzat útpályája, az újbóli megjelenésekor mindig új személyként azonosította a járókelőt. Az újrafelismerés feladatát részben kezelte Benedek kiterjesztése [10], amely a Lidar szenzor pontfelhőiből nyert ki két gyenge biometrikai azonosítót: a becsült magassá-

got, és az intenzitáshisztogramot. Az előző két leíró azonban könnyen összekeverheti a személyeket, ha azok magassága és öltözete hasonló. Az RMB Lidar szenzor függőleges felbontása meglehetősen alacsony (0.4 fok a mérés során használt Velodyne HDL 64-E szenzornál [10]), ami miatt a biztonságos elválasztáshoz legalább 6-8 cm magasságkülönbségnek kell lennie a két megkülönböztetni kívánt személy között. Egy másik probléma, hogy a használt szenzor intenzitás csatornája kalibrálatlan, vagyis a mért intenzitásértékek nem feltétlenül feleltethetők meg adott típusú ruhaanyagoknak, ráadásul az intenzitás pontos értéke függ a szenzortól való távolságtól és a lézersugár lokális beesési szögétől. A fenitek miatt célszerűtül dinamikusan biometrikus jellemzők, esetünkben a járás, felhasználását is a felismerési folyamatban.

Az i4D rendszer [8] vizualizációs modulja is adott lehetőséget a járásanalízis alapú továbbfejlesztésre. Az avatárok itt mindig az útpálya érintője által meghatározott irányba fordulva követik a valós személyek trajektóriáit, viszont az animált lábak mozgása nincs szinkronban a valóságban megfigyelt járás ütemével. Az animáció során valójában a 4D stúdióban felvett két lépést ismételtük folyamatosan, figyelmen kívül hagyva a járás tényleges frekvenciáját és fázisát. Mivel az aktuális járásfázist ki tudjuk nyerni a Lidar szenzor által mért adatokból, a fázis információt felhasználhatjuk a mérés szinkronizált járásanimáció létrehozásához.

1.1. Járáselemzéssel kapcsolatos munkák a szakirodalomban

A következő fejezetben áttekintjük a szakirodalomba megtalálható kép alapú járásfelismerő módszereket, fókuszálva az általunk bemutatott Lidar-alapú technikákra.

Számos metódus monokuláris optikai kamerák felvételeit próbálja felhasználni detekciós problémák megoldására. Azonban csak nagyon egyedi környezetben feltételezhetjük, hogy a megfigyelt objektumok mindig oldalnézetből látszanak [11, 12], ezért a fő kutatási feladat egy olyan nézet-invariáns reprezentáció kidolgozása, amely mindig lehetővé teszi a járás jellemzők könnyű kinyerését. A Kusakunniran és Tsai módszerében [13] bemutatott nézet-transzformációs modell egy többrétegű megközelítésen alapszik, míg a [14, 15] által ismertetett módszerek egy energiaképpel reprezentálják a járást. Kale [16] a Dynamic Time Warping (DWT) jelfeldolgozó algoritmust felhasználva egy tér-időbeli sziluett összehasonlítást hajtott végre, [17] egyszerű sziluett átlagolásból kinyert jellemzőket használ fel. Számos módszer egy kanonikus alakzat-representációval írja le az objektumokat [18–20]. A fent említett módszerek nyilvános, több-nézőpontú járás adatbázisokon [21] próbálják maximalizálni a detekciós eredményeket. Ilyen referencia adathalmazok a CASIA járás adatbázis [22], az USF adatbázis [23] és a CMU Motion of Body (MoBo) adatbázis [24]. A fent említett adatbázisok sok gyalogosról tartalmaznak különböző nézőpontokból rögzített mozgás szekvenciákat, de a felvételek erősen kontrolláltak, jó megvilágítású beltéri vagy kültéri környezetben készültek. A megfigyelt szereplők többnyire fix pályákat követnek [23], vagy futópadon mozognak [24]. Ezek a megkötések nagyban leegyszerűsítik a felismerést, szemben a valós környezetekkel, ahol a követett objektumok tetszőlegesen mozoghatnak. A HID-UMD [25] adatbázis sokkal általánosabb, kültéri környezetben rögzített járás szekvenciákat tartalmaz, melyek változó nézőponttal, kameratávolsággal és háttér paraméterekkel rendelkeznek. Azonban, a fent említett adatbázisokhoz hasonlóan, a videó szekvenciák egyszerre csak egy gyalogost tartalmaznak, ami lényegesen könnyebbé teszi a jó minőségű sziluett

kinyerést. Az eddig említettekkel ellentétben, valós dinamikus jelenetekben egyidejűleg több járókelőre kell számítani, akik gyakran takarásban vannak egymással, ezért a detekciós folyamat egyik kulcskérdése a megfelelő minőségű sziluettek kinyerése. Léteznek olyan rendszerek [25], melyek a takarás problémáját, több különböző nézőpont fúziójával oldják meg, azonban ehhez telepíteni és kalibrálni kell egy összetett kamera-rendszert, ami sok alkalmazási helyzetben nem lehetséges.

Másik lehetőség, hogy mélységinformációt is kinyerünk a helyszínről, sztereo kamerák vagy Time-of-Flight (ToF) szenzorok felhasználásával. Az olcsó és széles körben elterjedt Kinect szenzort több járás elemzéssel foglalkozó munka során felhasználták [26–29], és kapcsolódó referenciaadatbázist is publikáltak már [30]. Azonban a Kinect szenzorok a kis látómezőjük (0,8m-4,0m látótávolság) és a direkt napfényben való alacsony teljesítményük miatt kevésbé alkalmasak kültéri alkalmazásoknál. A [31, 32] tanulmányokban egy 2D lézer szkennert használtak a járás karakterisztikájának mérésére, azonban a szenzor csak 2D-ban rögzít adatokat, így nem alkalmas sem objektumok klasszifikálására, sem nagy területek megfigyelésére.

A Velodyne HDL-64E Lidar szenzor 64 lézerforrással rendelkezik és maximum 15Hz-es sebességgel képes 360 fokos látószögű 3D-s pontfelhő szekvenciákat rögzíteni, mely pontfelhők átlagosan 65 ezer pontot tartalmaznak. A lézercamera maximális látótávolsága 120 méter, ami nagy kültéri környezet megfigyelésére is alkalmassá teszi a technológiát. Továbbá a szenzor gyorsan telepíthető, újrainstallálás és kalibrálás nélkül mozgatható új helyszínre. A pontfelhő térbeli felbontása viszont relatíve alacsony és inhomogén. A szenzortól távolodva jelentősen csökken a pontok száma, és az egyes szkennelési körgyűrűk távolsága is jelentősen növekszik. Méréseink alapján, 10-20 méteres rádiuszon belül egy gyalogos pontfelhőjének a mérete átlagosan 180 és 500 pont között változik, ami két nagyságrenddel kisebb, mint a Kinect szenzor esetében (10-20 ezer pont/gyalogos), míg a sztereo kamerák mérései is lényegesen több pontot tartalmaznak [33].

A cikkben áttekintjük a Velodyne Lidar szenzor alkalmazási lehetőségeit a vizuális járás elemzéssel kapcsolatban. A tanulmányban bemutatott rendszer valós kültéri környezeteket felügyel, egyidejűleg több járókelővel a szintéren. A cél a gyalogosok újraazonosítása és a jelenetek realisztikus 4D-s megjelenítése. A Velodyne Lidar kamerát korábban használták már gyalogos detekcióra és követésre [10, 34], azonban ismereteink szerint ez az első tanulmány, ami az emberi járás elemzéséhez használja fel a szenzort. A szenzor alacsony térbeli felbontása és a különböző takarási jelenségek következtében kialakuló hiányos alakzatformák miatt úgy döntöttünk, hogy modell nélküli megközelítést alkalmazunk. Ellentétben ezzel, [27, 35, 36] munkák egy strukturált test modellt illesztnek a detektált objektumra. Például a [35] egy él-különbség térképen alapuló Particle Swarm optimalizációs módszert használ a modell illesztéséhez, de ehhez nagyon jó minőségű sziluettre van szükség. Esetünkben viszont azt tapasztaltuk, hogy olyan általános jellemzőkre kell támaszkodnunk, mint például a sziluett szélességének a dinamikája. A munkafolyamat során zaj-toleráns, robosztus jellemzők kinyerésére és a járás jellemzők más biometrikus leírókkal történő hatékony integrálására törekedtünk. A cikkben áttekintjük a Velodyne Lidar szenzor alkalmazási lehetőségeit a vizuális járás elemzéssel kapcsolatban. A tanulmányban bemutatott rendszer valós kültéri környezetet felügyel, egyidejűleg több járókelővel a szintéren. A cél a gyalogosok újraazo-

nosítása és a jelenetek realizisztikus 4D-s megjelenítése. A Velodyne Lidar kamerát korábban használták már gyalogos detekcióra és követésre [10, 34], azonban ismereteink szerint ez az első tanulmány, ami az emberi járás elemzéséhez használja fel a szenzort. A szenzor alacsony térbeli felbontása és a különböző takarási jelenségek következtében kialakuló hiányos alakzatformák miatt úgy döntöttünk, hogy modell nélküli megközelítést alkalmazunk. Ellentétben ezzel, [27, 35, 36] munkák egy strukturált test modellt illesztnek a detektált objektumra. Például a [35] egy él-különbség térképen alapuló Particle Swarm optimalizációs módszert használ a modell illesztéséhez, de ehhez nagyon jó minőségű sziluettre van szükség. Esetünkben viszont azt tapasztaltuk, hogy olyan általános jellemzőkre kell támaszkodnunk, mint például a sziluett szélességének a dinamikája. A munkafolyamat során zaj-toleráns, robosztus jellemzők kinyerésére és a járás jellemzők más biometrikus leírókkal történő hatékony integrálására törekedtünk.

2. A javasolt módszer bemutatása

A cikkünkben szereplő járáselemzést az i4D rendszerbe [8] is beintegráltuk. Az i4D rendszer célja dinamikus kültéri környezetek automatikus feltérképezése és 4D megjelenítése, ahol egyszerre akár több járkáló ember is szerepelhet. Ebben a fejezetben áttekintést adunk az i4D keretrendszeréről. A 4D mérés során egy RMB Lidar szenzor felvételt készít a környezetéről dinamikus pontfelhő-szekvencia formájában, egy fix pozícióból. A mérési adatok feldolgozásával a környező tér statikus részéről a rendszer egy 3D modellt készít, valamint detektálja és követi a mozgó embereket. Minden egyes gyalogost egy ritka, mozgó ponthalmazzal reprezentál és a hozzá tartozó útvonallal, trajektóriával. Ezután a gyalogoshoz tartozó ritka ponthalmazt egy, a 4D stúdióban [9] elkészített avatárral helyettesíti. Végül a felvett jelenet integrált 4D modelljét tetszőleges nézőpontból megjeleníthetjük.

2.1. Előtér–háttér elkülönítés

Az RMB Lidar szenzor 360 fokos látószögű, szabálytalan pontfelhő adatszekvenciákat rögzít. A dinamikus előtér és a statikus háttér elkülönítése a 3D adatokon egy valószínűségi módszerrel [37] történik. A valós idejű futás biztosítása érdekében a szabálytalan pontfelhőt egy hengerfelületre vetítjük, miáltal a mélységkép tartománya egy szabályos rács lesz, és így a szegmentációt egy kétdimenziós képtéren hajthatjuk végre.

Az egyes pixelekhez tartozó mélységértékeket Gauss eloszlások keverékeként („Mixture of Gaussians”) modellezzük, és a standard módszerhez [38] hasonlóan frissítjük az egyes paramétereket. A háttérrel a legnagyobb súllyal szereplő Gauss-féle komponensek modellezik, és a kiugró értékek (outlierek) detektálása lehetővé teszi a mozgó régiók kinyerését. Azonban a fenti sémát alkalmazva előfordulhatnak tévedések, a diszkrétizált látószög kvantálási hibája és a háttér villogása, azaz a növényzet mozgása miatt. Ezek a tévedések jelentősen csökkenthetőek egy dinamikus MRF modellel [37], ami a háttér és előtér osztályokat térbeli és időbeli jellemzőkkel írja le. Mivel az MRF modell a mélységkép terében van definiálva, ezért a 2D képszegmentációt egy 3D pontosztályozási lépésnek is követnie kell, hogy a 3D–2D leképezésből származó bizonytalanságot lokális térbeli szűréssel feloldjuk. Térbeli előtérmodellt használva eltávolítjuk a

nem releváns háttérmozgások nagy részét, amit többnyire a fák mozgó lombkoronái okoznak.

2.2. Járókelő-detekció és többalanyos követés

Ebben a részben a járókelő-detekcióért és -követésért felelős modult mutatjuk be. A modul bemenete az RMB Lidar pontfelhő szekvencia, ahol minden egyes pontot megjelölünk előtérként vagy háttérként. A kimenet az előtér egyes régióinak klasztereiből áll, úgy, hogy az ugyanahhoz az emberhez tartozó pontok ugyanazt a címkét kapják a teljes szekvencia során. Ezen felül minden egyes járókelőhöz generálunk egy 2D trajektóriát.

Első lépésként a pontfelhő azon régióit, melyeket előtérként detektáltunk, különálló ponthalmazokba csoportosítjuk, amik maguk a mozgó személyek lesznek. A talaj síkjára egy szabályos rácsot illesztünk, majd az előtér régiót levetítjük erre a rácsra. Morfológiai szűrők felhasználásával a képtérben meghatározzuk az egyes emberekhez tartozó, térben összefüggő alakzatokat. Ezek után a rendszer kinyeri a megfelelő méretű összefüggő komponenseket, azokat, amik méretben egy alsó és felső küszöbérték közé esnek. Minden kinyert alakzat középpontjára úgy tekint a rendszer, mint az emberek lábainak a pozíciójára a talaj síkján. Megjegyezzük, hogy az összefüggő gyalogos alakzatok egybeolvadhatnak, míg a részben takarásban lévő személyeket eltüntetheti, vagy több részre bonthatja az algoritmus. Ahelyett, hogy heurisztikus módszereket alkalmaznánk az ilyen jellegű hibák kiszűrésére külön-külön minden egyes képkockára, mi egy robusztus „multi-tracking” modult fejlesztettünk ki, amely hatékonyan kezeli a fent említett problémákat a teljes szekvenciát figyelembe véve.

A gyalogoskövető modul a rövid (STA) illetve a hosszú (LTA) távú követési modellt kombinálja. Az STA minden aktuálisan járókelőnek detektált objektumot megpróbál összeegyeztetni a követő modul által tárolt jelenlegi objektum trajektóriákkal, egyszerűen a levetített 2D középpontok pozíciójának az összehasonlításával. Az STA folyamatnak folytatnia kell tudnia az adott trajektóriát, még akkor is, ha a takarások miatt néhány képkockán keresztül a detektor nem találja meg a megfelelő objektumot. Ezekben az esetekben, a követés időbeli folytonossága miatt, a hiányzó detekciókat a trajektóriából becsült pozíciókkal tölti fel az algoritmus. Másrészt, az LTA modul felelős olyan jellemzők kinyeréséért, amiket azoknak az objektumoknak az újraazonosításánál lehet felhasználni, melyeket az STA modul elvesztett a több képkockán átívelő folyamatos takarások, vagy a megfigyelt terület elhagyása miatt. Ezért az elvesztett objektumokat egy archív objektum listába helyezzük, melyet periodikusan vizsgál az LTA folyamat. Az LTA feladata az is, hogy felismerje, ha egy teljesen új személy jelenik meg a színtéren, akit ezelőtt még nem regisztrált a követő modul.

Rövid távú hozzárendelés (STA) modul: Az STA modul megkapja a mért talaj pozíciókat és minden képkockánál három alpműveletet hajt végre: adat hozzárendelése, Kálmán szűrő javítása és Kálmán szűrővel új pozíció becslése. A hozzárendelés során az aktuálisan mért pozíciókat a modul összeegyezteti a regisztrált objektum trajektóriákkal, a Magyar módszer alapján [39]. Ezután a becsült objektum pozíciókat kor-

rigálja a tényleges pozíciókkal és végül megbecsüli a későbbi pozíciókat, amiket visszacsatol a hozzárendelési eljárásba. Az algoritmus képes kezelni a hamis pozitív eredményeket, valamint egy szekvencián belüli követés kezdetét és végét. A várt takarások és zajhatások következtében a *pár nélkül maradt* trajektóriák nem kerülnek azonnal törlésre, hanem egy inaktív jelölést kapnak, amely állapotban maximum TSIL időnyi időkeretet tölthetnek. Az inaktív trajektóriák is részt vesznek az STA folyamatban, azért, hogy a trajektóriák Kálmán szűrője frissítve legyen az aktuális pozíciók utolsó becsült értékével. A *pár nélkül maradt* mérési eredmények új trajektóriák lehetséges kezdő értékei, ezért a rendszer új objektum követést kezd el velük, melyet tovább vizsgál a következő iterációk során.

Hosszú távú hozzárendelés (LTA) modul: A [10] cikkben a szerzők két statikus pontfelhőleíró használják a személy újraazonosításhoz. Először is megfigyelték, hogy mivel az emberek ruházata különböző anyagú lehet, ezért az RMB Lidar szenzorral mért intenzitás adatok különböző statisztikai karakterisztikákat mutatnak különböző embereknél. Minden egyes követett objektumról legalább 100 képkocka felhasználásával intenzitás hisztogramot készít a rendszer, és jellemzőként használja az összehasonlításoknál. A kísérleti mérések megmutatták, hogy a normalizált intenzitás hisztogramok Bhattacharyya távolságát felhasználva hatékonyan megállapítható, hogy két objektum egy személyhez tartozik-e vagy sem.

A második jellemző a személy becsült magassága, amit minden egyes képkockára kiszámol az algoritmus a mért legalacsonyabb és legmagasabb pont különbségként, majd magassághisztogramot állít elő belőlük. A személy adott magasságát az aktuális magassághisztogram legmagasabb értékéből becsüli az algoritmus.

Mivel mind a két jellemző időbeli statisztikából származik, ezért egy újonnan megjelenő objektum először egy inicializálási állapotba kerül, ahol a hosszú távú hozzárendelés hisztogramjait meg lehet konstruálni. Egy adott számú időkeret után azok, amelyek teljesítették az LTA folyamatot, átkerülnek az azonosított objektumok közé. Csak akkor fogadunk el egy LTA általi objektum-újraazonosítást, ha mind az intenzitás- és mind a magassághisztogram nagy hasonlóságot mutat. Ha az LTA folyamat az új objektumot egyik archív objektumhoz sem tudta megfelelően társítani, akkor egy egyedi új azonosítót kap, hiszen ezelőtt még nem szerepelt a színtéren.

Követési folyamat Az előzőekben ismertetett STA és LTA modulok alapján a követés folyamata egy véges állapotú géppel írható le [10]. Egy aktuális követett objektum adott állapota az STA szerint lehet *Aktív* vagy *Inaktív*, míg az LTA szerint *Azonosított* vagy *Inicializálási* fázisban lévő. Ezzel a két bináris paraméterrel négy állapot különíthető el. Az *aktív* és *inaktív* állapotok közötti átmenetet az STA modul vezérli, attól függően, hogy az aktuális mérést sikerült-e már létező trajektóriához párosítani. Azon *Azonosított* objektumok, amik több, mint TSIL képkockán át *inaktívak*, az archív listába kerülnek: az *Archív* objektumok nem vesznek részt az STA folyamatban, de később újra lehet őket azonosítani az LTA folyamat során. Azok az objektumok, melyek TSIL időt eltöltenek az *Inicializálási–Inaktív* szakaszban, törlésre kerülnek és nem vesznek részt többet a követési folyamatban. Ezek a törölt trajektóriák általában valamilyen mérési zajnak

felelnek meg vagy túl rövidek ahhoz, hogy megbízható jellemzőket szolgáltatassanak az LTA folyamat számára.

Az LTA újraazonosítási folyamat azoknál az objektumoknál fejeződik be, amelyek legalább TinitL képkockányi időt töltenek az *Inicializálási–Aktív* állapotban. Ha van sikeres párosítás valamelyik archív objektummal, akkor az új és a megfeleltetett objektum trajektóriája összefésülésre kerül és a hiányzó trajektória részeket kiegészítjük interpolációval. Ezután az LTA által párosított *archív* objektum átkerül az *Azonosított–Aktív* állapotba és az új objektum törlésre kerül a duplikálódások megelőzése végett. Másrészt, ha az LTA folyamat nem talál társítást, akkor az új objektum átkerül az *Azonosított–Aktív* állapotba, megtartva saját azonosítóját.

2.3. Mozgó avatár generálás és 4D jelenet vizualizáció

Relatív kis objektumok, mint a járókelők, nem rekonstruálhatók kellő részletességgel a Lidar szenzor adataiból, mivel a pontfelhő túl ritka és csak 2,5D információk van a környezetről. Ezért az i4D rendszer egy 4D rekonstrukció stúdiót használ megfelelő részletességű, textúrázott dinamikus modellek létrehozására. Egy ilyen stúdió hardver- és szoftverrendszerét a [9] cikkben láthatjuk. A mozgó avatárokat tetszőleges 3D háttérmodellbe helyezhetjük (pontfelhő, háló vagy textúrázott modell), mely vagy manuálisan egy CAD rendszerben, vagy a Lidar mérései alapján automatikus környezet-térképezéssel készül [8].

A munkafolyamat utolsó lépése az egyes rendszerkomponensek integrálása és az integrált modell megjelenítése. A sétáló gyalogos modellek egy rekonstruált környezetbe kerülnek, mégpedig úgy, hogy a modellek lábainak középpontja követi a Lidar pontfelhőszekvenciákból kinyert trajektóriákat. Még a Kálmán szűrő alkalmazásával is a kinyert 2D nyers objektum pályák kissé zajosnak bizonyultak, ezért mi egy 80 százalékos tömörítést alkalmaztunk a görbékre a Fourier jellemzők terében [40]. Ennek eredménye az 1. ábrán látható. Az animáció alatt azzal a feltételezéssel élünk, hogy a gyalogos előrefele halad a trajektóriája mentén. A személyek felülnézeti irányát a 2D simított trajektória változásából számoljuk.

3. Lidar alapú járás analízis

Ebben a fejezetben a pontfelhő alapú járásanalízisre fókuszálunk, az i4D keretrendszert további két új jellemzővel kiegészítve:

- A személykövetést megvalósító rendszerkomponensben (2.2. fejezet) a *hosszútávú hozzárendelés modul* (LTA) kiegészítjük járás-alapú biometrikus jellemzőkkel
- A 4D jelenet vizualizációs modulban (2.3. fejezet) szinkronizáljuk a megfigyelt gyalogosok pontfelhő szekvenciákból lement lépéseit a megjelenített sétáló stúdió objektumok animált lábmozgásával.

A fenti célok eléréséhez a meglévő rendszer két elemét használtuk fel:

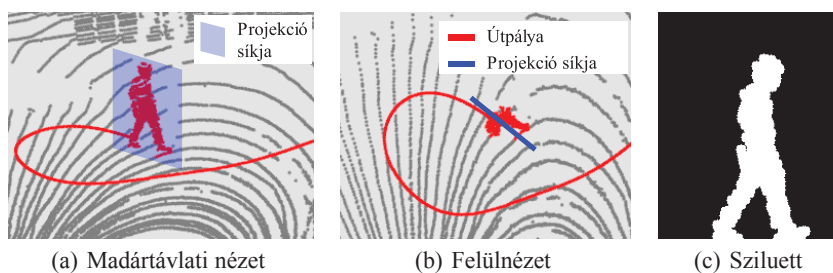
- A személyek adott időkerethez tartozó pontfelhőit, melyek az előtér-detekció (2.1. fejezet) és a gyalogosszeparálás (2.2. fejezet) lépéseket követően elérhetők.

- A követés során kinyert majd simított objektumtrajektóriákat, melyeket a gyalogos aktuális felülnézeti orientációjának becsléséhez használunk.

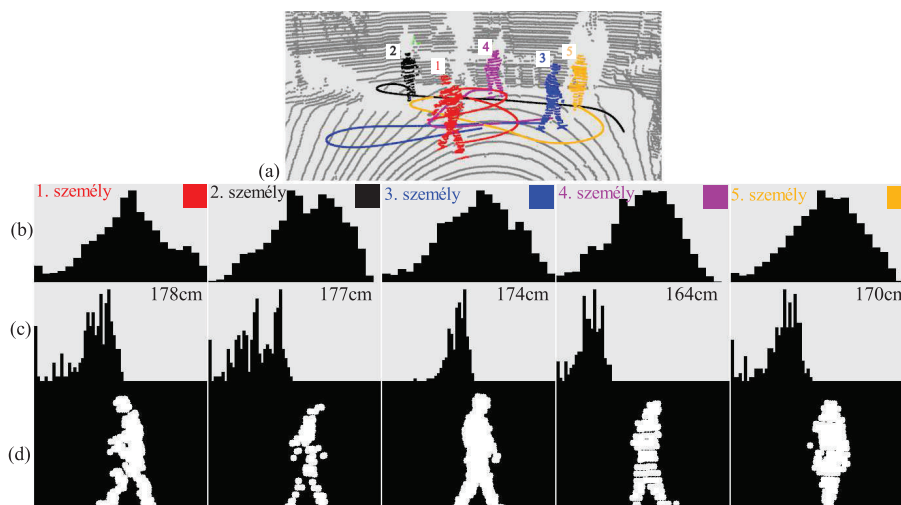
A járásvizsgálat során 2D sziluett alapú megközelítést követtünk. Ahogy az 1.1. fejezetben részleteztük, számos korábbi módszer alapszik a mért vagy interpolált 3D sziluettek oldalnézeti projektálásán. Rendszerünkben az RMB Lidar pontfelhő szekvenciákat felhasználva az oldalnézeti sziluett automatikusan becsülhető (1 ábra). Itt azzal a feltételezéssel élünk, hogy az emberek a szintéren mindig előre nézve haladnak, követve a trajektóriájukat. Ezért egy adott időkereten belül valamennyi személy pontfelhőjét levetítjük arra a síkra, amely metszi a személy aktuális talppontját, merőleges a talaj síkjára, és felülnézetből tekintve párhuzamos a (Fourier leírókkal) simított trajektória lokális érintő vektorával (1(a),(b) ábra). Az így projektált pontfelhő több elkülönülő régiót tartalmazhat a képsíkon, melyeket morfológiai műveletekkel alakítunk összefüggő 2D sziluetté (1(c) ábra).

A sziluettkinyerés során számos gyakorlati nehézséggel kell szembenéznünk, melyeket a 2 ábra (a) és (d) sorai szemléltetnek. Egyrészt a szenzortól távol helyezkedő emberek pontthalmaza ritka, ezért a sziluettjeikben szakadások lehetnek (2. és 3. személyek a 2(d) ábrán). Másrészt, azoknál az embereknél, akik épp a szenzor felé vagy annak háttal fordulva sétálnak, a 2,5D mérés egy elől- vagy hátulnézetet eredményez, ahol a lábak részben kitakarhatják egymást (5. személy). Továbbá néhány sziluett-rész takarásban lehet más emberek vagy tereptárgyak által. Bár munkánk során kipróbáltunk több szofisztikált sziluett-alapú jellemzőkinyerést a hagyományos (optikai) videofeldolgozás szakirodalmából, kísérleteink azt mutatták, hogy a fent említett mellékhatások miatt a rendelkezésünkre álló Lidar-alapú sziluetteken nem tudunk megfelelő pontosságú azonosítást végrehajtani. Ezért úgy döntöttünk, hogy egyszerűen származtatható robotus járásjellemzőkre támaszkodunk, melyeket együttesen használunk fel korábbi fejezetben bevezetett az intenzitás hisztogrammal (2(b) ábra) és magasság becsléssel (2(c) ábra) [10], az LTA azonosítási folyamat során.

Módszerünkben három különböző járás jellemzőt nyertünk ki és használtunk fel: domináns lépés hossz (L , cm-ben mérve), lépés frekvencia (F , egységnyi idő alatti lépés ciklusok száma), és a nyitott lábak által bezárt maximális mért szög (A). Meg-



1. ábra: Sziluett projekció: (a) követett személy pontfelhője és sziluettprojekció síkja madártávlatból (b) a projekciós sík felülnézetből a simított trajektória érintőjét követve (c) projektált sziluett maszk a képsíkon



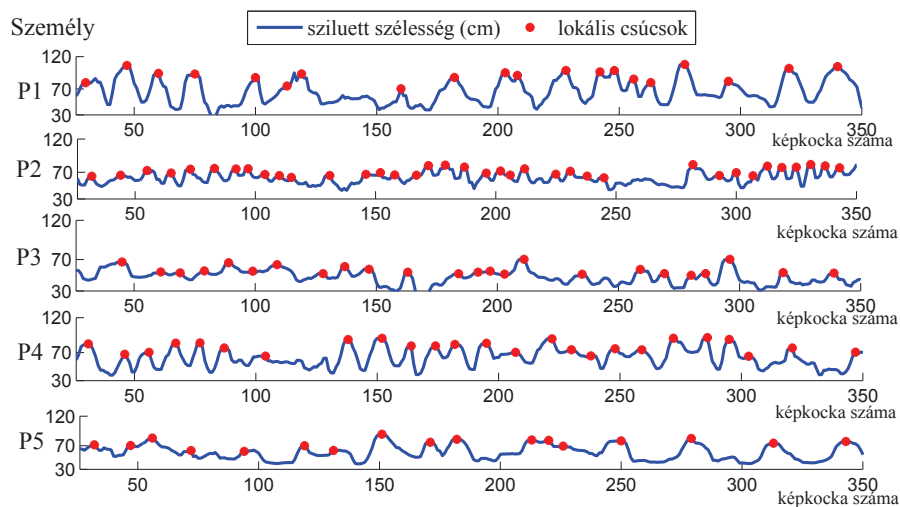
2. ábra: Újraazonosításhoz használt jellemzők (Nyár teszt szekvencia). Sor (a): Pontfelhő időkeret, szeparált emberekkel és trajektóriákkal Sor (b): térbeli intenzitás histogram Sor (c): időbeli magasság histogram Sor (d): projektált sziluett, mint járáslenyomat

közelítésünk modellmentes, tehát a felsorolt paramétereket közvetlenül a sziluettből, csontváz illesztés nélkül nyerjük ki. A lépéshossz és a frekvencia jellemzőket az alsó sziluett harmad 2D befoglaló téglalapjának időbeli követésével közelítjük. A zajcsökkentéshez medián szűrőt használunk, majd kinyerjük a befoglaló téglalapok szélesség sorozatának lokális maximumait, és minden személynél egy időbeli histogramot készítünk a kinyert értékekből. A histogram csúcsainak lokalizációjával becsüljük a domináns lépéshosszakat. A frekvencia kinyerésénél hasonlóan a lépéshossz megállapításához, egy időbeli histogramon meghatározzuk a domináns időperiódusokat a lokális maximumok között eltelt időből (időkeretek száma). A lábak által bezárt szög kiszámításához meghatározzuk a sziluett középpontját és azt a két pontot, ahol a lábak érintkeznek a talajjal. Ezután kiszámoljuk azt a szöveget, amit a középpontból az egyik illetve a másik lábpontba húzott egyenes határoz meg. Mindhárom paramétert a $[-1, 1]$ intervallumra normalizáljuk, majd a következő képlet alapján számolunk egy egyesített járásjellemzőt: $f = 1,5L + 1,5F + A$. Az f járás jellemzőt az LTA részben használjuk fel, az emberek újraazonosítási folyamatában az intenzitás histogram és a magasság jellemzőkkel együttvéve.

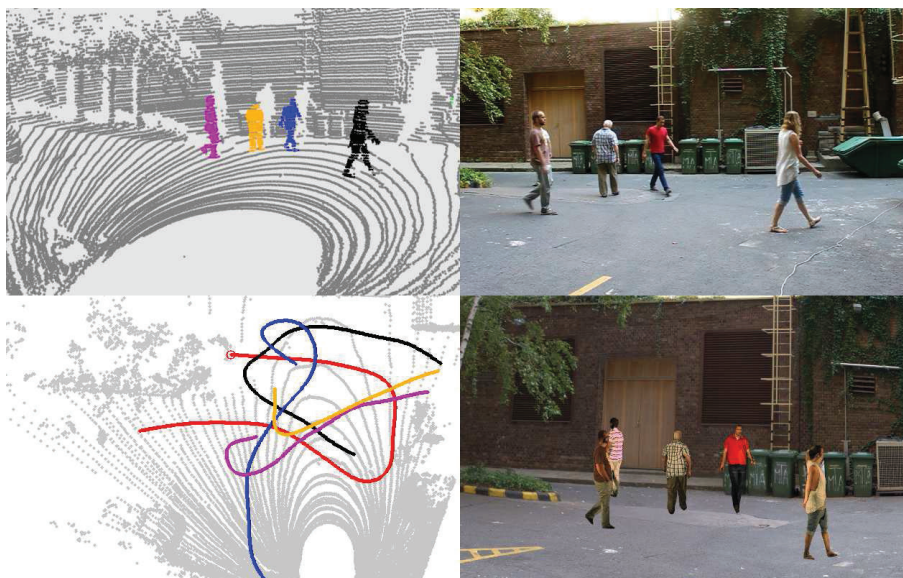
A járás periódusának kinyerése szükséges a megfigyelt és az animált lábak mozgásának a szinkronizálásához is a vizualizációs modulban. Ezt szintén a sziluett szélességi változásaiból nyerjük ki, ahol a histogramban két egymás utáni csúcs hely közötti idősor egy fél lépésciklust fed le.

4. Teszteredmények

A bemutatott módszert három, kültéren rögzített szekvencián teszteltük, melyek a Velodyne HDL-64E RMB Lidar szenzorral készültek. A rögzítés során a Lidar szenzor 15



3. ábra: Öt ember (P1-P5) sziluett szélességi szekvenciája, a Járás minta adathalmazon. P1: nagy lépések kis frekvenciával. P2: kocogás. P3: totyogás (séta kis lépésekkel). P4: normál séta. P5: szaggatott járás (aszimmetrikus lépések).



4. ábra: A rekonstrukciós munkafolyamat bemutatása. Bal fent: bemeneti pontfelhő szekvenciája, az előtér meghatározásával és a multi-objektum követés alkalmazásával. Bal lent: a követett alanyok felülnézeti trajektóriája. Jobb fent: referencia videó a megfigyelt területről. Jobb lent: a rekonstruált helyszín, a 4D stúdióban készített mozgó avatárokkal, ahol az avatárok lábainak a mozgása szinkronizálva van a Lidar szenzor által megfigyelt valós lépésperiódusokkal.

képkocka/másodperc felvételi sebességgel forgott mind a három szekvencia esetében. Mindhárom felvétel esetén egyidejűleg több gyalogos sétált a megfigyelt területen, majd egy kis ideig elhagyták a pásztázott részt és később különböző sorrendben újra megjelentek a szintér egy tetszőleges pontján. A kísérlet fő célja az volt, hogy a felvételek első és második szakaszában megfigyelt járásjellemzők alapján újra tudjuk azonosítani az embereket.

A *Járás minta* szekvencián (5 ember, 3:13 perc, 2995 képkocka) a tesztalanyokat arra kértük, hogy különböző karakterisztikájú járással mozogjanak, mint például normál járás, kocogás, szaggatott járás, és totyogás. Ez a tesztszekvencia egy egyszerűbb példa, amit a járásjellemzők tesztelésére használtunk, hiszen itt szignifikáns a különbség az egyes járások között. Ezzel szemben a *Nyár* (5 ember, 2:43 perc, 2434 képkocka) és a *Tél* (6 ember, 3:01 perc, 2710 képkocka) tesztalanyokon a megfigyelt alanyok természetesen sétáltak, így ezeken az adatokon lehetőségünk nyílt a járásjellemzők alapján történő szeparálási képességet ellenőrizni valószerű környezetben.

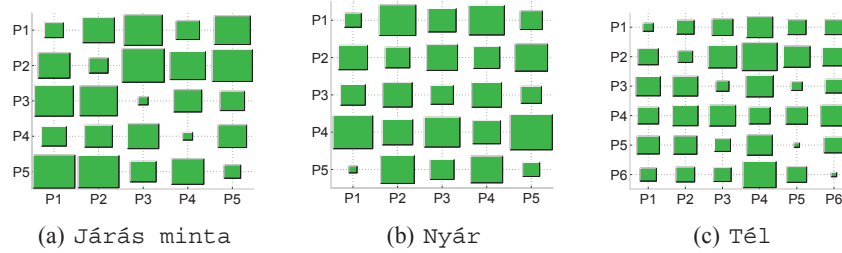
A 3 ábrán a *Járás minta* szekvencián kinyert sziluettisélességek időbeli változását figyelhetjük meg, ahol a kinyert lokális maximumokat piros pontok jelölik. Megfigyelhetjük, hogy a járás görbék eléggé zajosak, köszönhetően a valószerű kísérleti környezet különböző tényezőinek, úgy mint az alacsony pontsűrűség, takarások és a 2,5D adat korlátai. Azonban a járásjellemzők néhány karakterisztikája így is szépen látszik: az első személy (P1) lassan sétált nagy lépésekkel, míg a második személy (P2) járásgörbéjében sok közeli, kicsi amplitúdójú csúccsal találkozunk, ami kocogásra utal. P3 kis lépésekkel sétált, de kisebb frekvenciával, mint P2. P4 járása szabályosnak tűnik, míg P5 járásmintájában szimetriát fedezhetünk fel, ami például utalhat valamilyen sérülésre.

Ezt követően egy kvantitatív kiértékelést hajtottunk végre az emberek járásanalíziséből számolt f érték segítségével, amit a 3. fejezetben mutattunk be. Minden egyes szekvencián kiszámoltuk az f értéket, mind a szekvenciák első fázisában, mind az alanyok újramejelenése után. Az összes elő- és utóazonosítási fázisból származó járásmintát páronként összehasonlítottuk az f értékek távolsága alapján. A kinyert távolságértékeket az 5 ábra szemlélteti, ahol az optimális egyezést az átlókban várjuk. Ez tisztán kirajzolódik a *Járás minta* ábráján (5(a)(a) ábra). Habár a megfeleltetési mátrix a *Nyár* és a *Tél* adathalmazokon nem ennyire egyértelmű, de a Magyar módszer [10] felhasználásával ezek is tökéletes eredményt adnak.

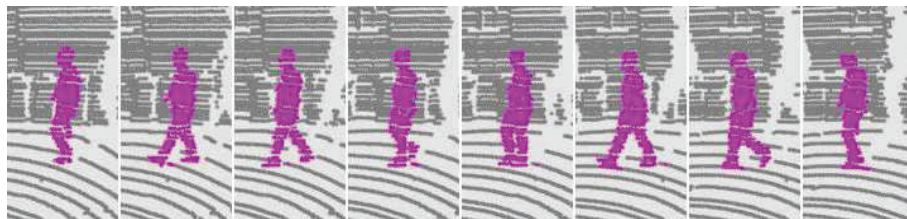
A bemutatott járásjellemzőket egyszerű módon integrálhatjuk az i4D rendszerbe, az LTA követési folyamat támogatása gyanánt. Kísérleteink kimutatták, hogy a *Járás minta* szekvencián, ha az újraazonosításnál csak az intenzitás és magasság jellemzőkre hagyatkozunk, ahogyan ez a [10] cikkben is szerepel, akkor két hasonló magasságú és hasonló ruházatú embert összekever az algoritmus. Ez a hiba a járásjellemzők felhasználásával kiküszöbölhető.

A rendszer vizualizációs moduljába implementáltuk a mért értékek és az avatárok lépésének a szinkronizációját. Egy példa látható a 6 ábrán, ahol a Lidar pontfelhő képkockáit a hozzá tartozó referencia képfelvétel és a 4D stúdióban animált avatárokkal készített jelenet követi. A 4 ábra a teljes felismerés és megjelenítés folyamatát mutatja be. Az Olvasó a következő linken találhat demonstrációs videót:

<http://vimeo.com/user32136096/videos>.



5. ábra: Az f járás jellemző kvantitatív kiértékelése a Járás minta, Nyár és Tél szekvenciákon. Minél kisebb egy téglalap, annál jobb a két alany párosítása. Az optimális párosítást a főátló reprezentálja.



(a) Pontfelhő szekvencia (felismeréshez)



(b) Videó szekvencia (nem használt, csak vizuális referencia)



(c) Szintetizált 4D séta, járás periódusa a Lidar megfigyeléséhez van szinkronizálva

6. ábra: Egymás utáni képkockák a (a) Lidar és a (b) videó szekvenciákból, és a 4D szintetizált jelenet, ahol a lábak szinkronizálva vannak az eredeti méréssel.

5. Konklúzió

Cikkünkben új algoritmust és kísérleteket mutattunk személyek járásuk alapján történő korlátozott azonosítására RMB Lidar pontfelhősorozatokot felhasználva. Rendszerünk valós kültéri környezetben egyidejűleg képes több gyalogost detektálni és követni. A kinyert járásjellemzőket a személyek újraazonosításához és a jelenet 4D megjelenítésének életszerűbbé tételéhez használtuk fel. Kvantitatív kiértékelést végeztünk három különböző mérési szekvencián, és demonstráltuk a bemutatott megközelítés előnyeit valamint későbbi felhasználási lehetőségeit 4D megfigyelő és vizualizációs rendszerekben.

Köszönetnyilvánítás

A munkát részben támogatta az OTKA #101598 posztdoktori projekt.

Irodalom

1. Roth, P., Settgast, V., Widhalm, P., Lancelle, M., Birchbauer, J., Brandle, N., Havemann, S., Bischof, H.: Next-generation 3D visualization for visual surveillance. In: IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS). (2011) 343–348
2. Baltieri, D., Vezzani, R., Cucchiara, R., Utasi, Á., Benedek, C., Szirányi, T.: Multi-view people surveillance using 3D information. In: Proc. International Workshop on Visual Surveillance at ICCV, Barcelona, Spain (2011) 1817–1824
3. Boyd, J., Little, J.: Biometric gait recognition. In Tistarelli, M., Bigun, J., Grosso, E., eds.: Advanced Studies in Biometrics. Volume 3161 of Lecture Notes in Computer Science. Springer Berlin Heidelberg (2005) 19–42
4. Zhang, Z., Hu, M., Wang, Y.: A survey of advances in biometric gait recognition. In Sun, Z., Lai, J., Chen, X., Tan, T., eds.: Biometric Recognition. Volume 7098 of Lecture Notes in Computer Science. Springer Berlin Heidelberg (2011) 150–158
5. Wang, J., She, M., Nahavandi, S., Kouzani, A.: A review of vision-based gait recognition methods for human identification. In: International Conference on Digital Image Computing: Techniques and Applications (DICTA), Sydney, Australia (2010) 320–327
6. Murray, M.P.: Gait as a total pattern of movement. *American Journal of Physical Medicine* **46** (1967) 290–333
7. Kim, H., Guillemaut, J.Y., Takai, T., Sarim, M., Hilton, A.: Outdoor dynamic 3-d scene reconstruction. *IEEE Trans. on Circuits and Systems for Video Technology* **22** (2012) 1611–1622
8. Benedek, C., Jankó, Z., Horváth, C., Molnár, D., Chetverikov, D., Szirányi, T.: An integrated 4D vision and visualisation system. In: International Conference on Computer Vision Systems (ICVS). Volume 7963 of Lecture Notes in Computer Science. Springer, St. Petersburg, Russia (2013) 21–30
9. Hapák, J., Jankó, Z., Chetverikov, D.: Real-time 4D reconstruction of human motion. In: Proc. 7th International Conference on Articulated Motion and Deformable Objects (AMDO 2012). Volume 7378 of Springer LNCS. (2012) 250–259
10. Benedek, C.: 3D people surveillance on range data sequences of a rotating Lidar. *Pattern Recognition Letters* **50** (2014) 149–158 Special Issue on Depth Image Analysis.

11. Li, Y., Yin, Y., Liu, L., Pang, S., Yu, Q.: Semi-supervised gait recognition based on self-training. In: International Conf. Advanced Video and Signal-Based Surveillance (AVSS), Beijing, China (2012) 288–293
12. Makihara, Y., Mannami, H., Tsuji, A., Hossain, M., Sugiura, K., Mori, A., Yagi, Y.: The ou-isir gait database comprising the treadmill dataset. *IPSN Trans. on Computer Vision and Applications* **4** (2012) 53–62
13. Kusakunniran, W., Wu, Q., Zhang, J., Li, H.: Cross-view and multi-view gait recognitions based on view transformation model using multi-layer perceptron. *Pattern Recognition Letters* **33** (2012) 882 – 889
14. Han, J., Bhanu, B.: Individual recognition using gait energy image. *IEEE Trans. Pattern Analysis and Machine Intelligence* **28** (2006) 316–322
15. Kusakunniran, W., Wu, Q., Li, H., Zhang, J.: Multiple views gait recognition using view transformation model based on optimized gait energy image. In: International Conference on Computer Vision Workshops, Kyoto, Japan (2009) 1058–1064
16. Kale, A., Cuntoor, N., Yegnanarayana, B., Rajagopalan, A., Chellappa, R.: Gait analysis for human identification. In Kittler, J., Nixon, M., eds.: *Audio- and Video-Based Biometric Person Authentication*. Volume 2688 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg (2003) 706–714
17. Liu, Z., Sarkar, S.: Simplest representation yet for gait recognition: averaged silhouette. In: International Conference on Pattern Recognition. Volume 4., Cambridge, UK (2004) 211–214 Vol.4
18. Liu, N., Tan, Y.: A novel study and analysis on segmental gait sequence recognition. In: IEEE International Conf. Acoustics, Speech and Signal Processing (ICASSP), Prague, Czech Republic (2011) 1473–1476
19. Kusakunniran, W., Wu, Q., Zhang, J., Ma, Y., Li, H.: A new view-invariant feature for cross-view gait recognition. *IEEE Trans. Information Forensics and Security*, **8** (2013) 1642–1653
20. Hu, H.: Multiview gait recognition based on patch distribution features and uncorrelated multilinear sparse local discriminant canonical correlation analysis. *IEEE Trans. Circuits and Systems for Video Technology* **24** (2014) 617–630
21. Shirke, S., Pawar, S., Shah, K.: Literature review: Model free human gait recognition. In: International Conf. Communication Systems and Network Technologies (CSNT), Bhopal, India (2014) 891–895
22. Zheng, S., Zhang, J., Huang, K., He, R., Tan, T.: Robust view transformation model for gait recognition. In: International Conference on Image Processing (ICIP). (2011)
23. Sarkar, S., Phillips, P., Liu, Z., Vega, I., Grother, P., Bowyer, K.: The humanID gait challenge problem: data sets, performance, and analysis. *IEEE Trans. on Pattern Analysis and Machine Intelligence* **27** (2005) 162–177
24. Gross, R., Shi, J.: The CMU Motion of Body (MoBo) Database. Technical Report CMU-RI-TR-01-18, Robotics Institute, Pittsburgh, PA (2001)
25. Sharma, S., Shukla, A., Tiwari, R., Singh, V.: View variations effect in gait recognition and performance improvement using fusion. In: IEEE International Conf. Recent Advances in Information Technology (RAIT), Dhanbad, India (2012) 892–896
26. Tang, J., Luo, J., Tjahjadi, T., Gao, Y.: 2.5D multi-view gait recognition based on point cloud registration. *Sensors* **14** (2014) 6124–6143
27. Gabel, M., Renshaw, E., Schuster, A., Gilad-Bachrach, R.: Full body gait analysis with kinect. In: Proceedings of EMBC 2012, Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (2012)
28. Whytock, T., Belyaev, A., Robertson, N.: Dynamic distance-based shape features for gait recognition. *Journal of Mathematical Imaging and Vision* (2014) 1–13

29. Hofmann, M., Bachmann, S., Rigoll, G.: 2.5D gait biometrics using the depth gradient histogram energy image. In: Biometrics: Theory, Applications and Systems (BTAS), 2012 IEEE Fifth International Conference on. (2012) 399–403
30. Hofmann, M., Geiger, J., Bachmann, S., Schuller, B., Rigoll, G.: The TUM gait from audio, image and depth (GAID) database: Multimodal recognition of subjects and traits. *J. Vis. Comun. Image Represent.* **25** (2014) 195–206
31. Pallejá, T., Teixidó, M., Tresanchez, M., Palacín, J.: Measuring gait using a ground laser range sensor. *Sensors* **9** (2009) 9133–9146
32. Teixidó, M., Mercé, P., Pallejá, T., Tresanchez, M., Nogués, M., Palacín, J.: Measuring oscillating walking paths with a LIDAR. *Sensors* **11** (2011) 5071–5086
33. Ryu, J., Kamata, S.: Front view gait recognition using spherical space model with human point clouds. In: IEEE International Conf. Image Processing (ICIP), Brussels, Belgium (2011) 3209–3212
34. Spinello, L., Luber, M., Arras, K.: Tracking people in 3D using a bottom-up top-down detector. In: IEEE International Conference on Robotics and Automation (ICRA), Shanghai, China (2011) 1304–1310
35. Kwolek, B., Krzeszowski, T., Michalczuk, A., Josinski, H.: 3D gait recognition using spatio-temporal motion descriptors. In: Intelligent Information and Database Systems. Volume 8398 of Lecture Notes in Computer Science. Springer International Publishing (2014) 595–604
36. Mansour, R.: A robust approach to multiple views gait recognition based on motion contours analysis. In: National Workshop on Information Assurance Research (WIAR). (2012) 1–7
37. Benedek, C., Molnár, D., Szirányi, T.: A dynamic MRF model for foreground detection on range data sequences of rotating multi-beam lidar. In: International Workshop on Depth Image Analysis, LNCS, Tsukuba City, Japan (2012)
38. Stauffer, C., Grimson, W.E.L.: Learning patterns of activity using real-time tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22** (2000) 747–757
39. Kuhn, H.W.: The Hungarian method for the assignment problem. *Naval Research Logistic Quarterly* **2** (1955) 83–97
40. Zhang, D., Lu, G.: A comparative study of fourier descriptors for shape representation and retrieval. In: Asian Conference on Computer Vision (ACCV, Springer (2002) 646–651