

Horváth Kristóf Attila

Intelligens színpadvilágítás

TDK dolgozat

Konzulensek:

Dr. Szabó Sándor

Lendvai Károly

Milánkovich Ákos

2012

Tartalomjegyzék

1. Bevezető.....	3
1.1. Motiváció.....	3
1.2. Rendszerkomponensek.....	3
2. Kinect.....	4
2.1. A Kinect rövid áttekintése.....	4
2.2. Kinect for Windows.....	5
2.3. A készülék felépítése.....	5
2.3.1. Érzékelők.....	6
2.4. Az érzékelők működési elvei.....	8
2.4.1. Nyalábformálás.....	8
2.4.2. Mélységérzékelés.....	10
2.5. Kinect for Windows fejlesztői környezet.....	13
2.5.1. Rendszerkomponensek.....	13
2.5.2. Adatfolyamok.....	14
2.5.3. Csontvázkövetés.....	14
2.5.4. Arcok követése.....	18
2.6. Kinect tesztprogram.....	19
3. A Kinect pontossága.....	21
3.1. A mérés leírása.....	21
3.2. Mérés egyhelyben.....	23
3.3. Mélységi mérések.....	24
3.4. Nyers mélységi adatfolyam.....	28
3.5. Oldalirányú mérések.....	30
3.6. Összegzés, a Kinect színpadi alkalmazásának korlátjai.....	32
4. Színpadvilágítás.....	33
4.1. A színpadvilágítás rövid története.....	33
4.2. Színpadi vezérléstechnika.....	35
4.3. Lámpavezérlés.....	38
4.4. Lámpavezérlő tesztprogram.....	42
5. Követés.....	43
5.1. Koordinátarendszerek.....	44
5.2. Transzformációs modell.....	45
5.3. Követő program.....	46
6. Kalibráció.....	46
6.1. A kalibráció elve.....	46
6.2. Kalibráció a gyakorlatban.....	47
7. Összegzés.....	48
8. Továbbfejlesztési lehetőségek.....	49
9. Forrásjegyzék.....	50
10. Ábrajegyzék.....	52
11. Táblázatok jegyzéke.....	53

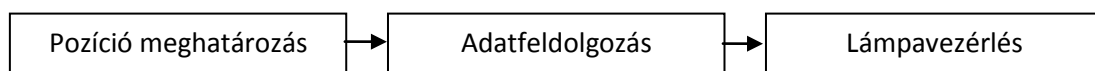
1. Bevezető

1.1. Motiváció

A színpadi világítástechnika gyakorta alkalmaz fejgépeket, mely berendezések célja, hogy a fontosabb színészeket egy fénykörrel megvilágítva kiemelje környezetükből. Manapság ezeket a lámpákat manuálisan, technikusok pozicionálják, és követik velük a színészek mozgását. A világítástechnikában azonban már évek óta alkalmaznak robotlámpákat, melyek automatikusan, a vezérlőpult utasításainak megfelelően képesek adott irányba fordulni, színt, effektet váltani, stb. Összevetve a fejgépeket és a robotlámpákat, jogosan merül fel az ötlet, hogy a ma már oly fejlett robotlámpák kiválthatnák a fejgépeket egy olyan rendszer segítségével, mely képes a színpadon mozgó színészek követésére, és a lámpák irányítására.

1.2. Rendszerkomponensek

A fejlesztendő rendszer három részre bontható. Először detektálni kell a követendő objektum pozícióját. Ezt követően meg kell határozni a robotlámpák vezérléséhez szükséges vezérlőinformációkat. Végül ezen adatokat ki kell küldeni a lámpaparknak.



1. ábra Rendszerkomponensek

A projekt megvalósítása során a színészek pozíciójának meghatározására számtalan lehetőség kínálkozik: Wi-fi, Bluetooth, UWB vagy ultrahang alapú háromszögelés, inerciális helymeghatározó eszközök, képfeldolgozás – csak néhányat említve. A projekt első rendszerkomponenseként a Kinectet választottam, mely kamerák segítségével, képfeldolgozás útján határozza meg az objektumok pozícióját. Azért esett erre az eszközre a választás, mert – bár maga az eszköz meglehetősen bonyolult – felhasználása, programozása viszonylag könnyű a biztosított fejlesztőeszközök segítségével, közvetlenül támogatja emberalakok követését, és nagy pontosság érhető el vele.

A lámpapark vezérlése a DMX512 szabványon keresztül történik, mely mind a mai napig a legelterjedtebb lámpavezérlő szabvány, számos későbbi szabvány is erre épül, melyből több kompatibilis is ezzel – köszönhetően a szabvány rendkívül egyszerű felépítésének.

2. Kinect

2.1. A Kinect rövid áttekintése

A Microsoft 2010-ben jelentette meg Kinect elnevezésű eszközét, melyet eredetileg az Xbox 360-as játékkonzoljukhoz fejlesztettek ki. Elnevezése a „kinetic” és „connect” szavak összetételéből fakad. A cél egy olyan periféria kialakítása volt, melynek segítségével a játékos kontrollerek kézbevétele nélkül irányíthatja a játékokat. [1]



2. ábra Kinect [2]

Korábban a Nintendo Wii-vel együttműködő Remote elnevezésű controller volt forradalmi újítás a játékiparban, ugyanis ezzel a controllerrel nyomógombok, joystick-ok helyett magával az eszköz mozgásával lehetett a konzolnak utasításokat adni. Bizonyos játékoknál rendkívül hatékonyan kihasználható ez az újfajta irányítási metódus, azonban a controller jelenléte zavaró, a rendszer csak a controller mozgását képes érzékelni, a játékos többi testrészének mozgásait nem, ráadásul a játék hevében, a kellenénél hevesebben meglendített controller olykor a felhasználó kezéből kicsúszva a televízió képernyőjében landolt, hatalmas kárt okozva ezzel. [3] [4]



3. ábra Wii Remote [5]

A Microsoft Kinect elnevezésű berendezése a fenti problémáktól mentes, ugyanis a játékosok kezében egyáltalán nincsenek vezérlők, a készülék az előtte állók egész testét érzékeli, a testrészek megfelelő mozgásaival lehet a programokat irányítani. A kézben tartott kontrollereket elhagyva lehetőség nyílik a többi testrészt is bevonni a mozgásba, így természetesen a technológiát kihasználó legtöbb program sport tematikájú, a velük történő játék az egész testet megmozgatja – rafting, tenisz, ping-pong, gátfutás, gerelyhajítás, bokszt, röplabda, foci is megtalálható a játékkonzolra írt programok között, csak a legfontosabbakat említve. [6]

2.2. Kinect for Windows

A Kinectben rejlő lehetőségek azonban nem merülnek ki a játékkonzolok világában, ezt felismerve a Microsoft 2012 februárjában piacra dobta Kinect for Windows elnevezésű termékét, melyet a korábbi verzióval ellentétben már számítógépes használathoz készítettek. A Kinectre történő fejlesztést ösztönözvén ingyenes fejlesztői környezetet biztosít a vállalat, ami remélhetőleg rövid idő múlva számtalan program megszületését fogja eredményezni, ugyanis jelen pillanatban még nem érhetők el Windowson futtatható, a Kinect képességeit használó szoftverek, egyelőre a Kinect for Windows-t is csak fejlesztőknek, vállalatoknak kínálja a cég, a végfelhasználóknak történő értékesítés még várat magára. [7]

A Kinect for Windows, levetkőzvé a játékkonzolok világának behatároltságát, már alkalmas lehet bizonyos speciális műszaki problémák megoldására, ilyen például a színészek követése is jelen projektben.

Mivel eredetileg az Xbox 360 játékkonzolhoz fejlesztették ki a Kinectet, és az abba beépített számítási kapacitás erősen korlátozott, ott elengedhetetlen volt a kiegészítő hardveres feldolgozás. A modern asztali számítógépek valószínűleg elboldogulnának a nyers adatfolyamok feldolgozásával is, de még itt is jelentős teljesítménynövekedés érhető el a készülék beépített hardveres megoldásainak köszönhetően.

2.3. A készülék felépítése

A Kinect tulajdonképpen több érzékelő egybeépítve, kiegészítvén a témérdek adat kiértékelését gyorsító hardveres feldolgozóegységgel. Éppen ezért talán meglepő a berendezést egyszerűen szenzornak titulálni, azonban a hivatalos források is többször hivatkoznak így az eszközre.

Az eszköz egy fejrészből, egy talpazatból, és az abból kivezető kábelből áll, melyen keresztül egy USB porthoz kapcsolódhat. Bár a csatlakozó szabványos USB kábel, az SDK dokumentációja felhívja a figyelmet arra, hogy a kommunikáció eltér a szabványban rögzítettektől. Az eszköz áramfelvétele is

jelentős, így a Kinectből kilépő kábel a számítógéppel, és egy tápon keresztül az elektromos hálózattal is kapcsolatban áll. A talpazat és a fej bezárt szöge egy motorral változtatható, így a fej elevációs szöge tetszőlegesen beállítható $\pm 27^\circ$ között, aminek aktuális értéke a fejben található eszközökkel mérhető is. Ennek az a jelentősége, hogy az érzékelő 60 centiméter és 2 méter közötti magasságban tetszőlegesen elhelyezhető, így ennek megfelelően szoftverek fejlesztése során gondoskodni kell arról, hogy a szenzor magasságától függetlenül, az elevációs szög állításának segítségével a kamerák látótere mindig lefedje az előtte állókat. Tulajdonképpen minden fontosabb elem, az érzékelők és a feldolgozó egység is a fejben található. [8]

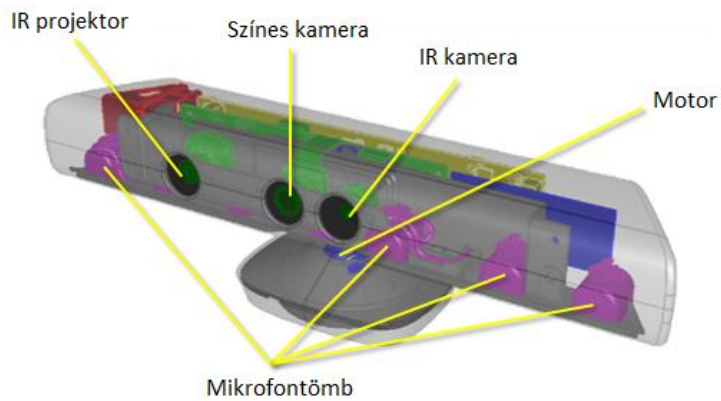


4. ábra A Kinect belseje [9]

A szenzoroktól származó adatfolyamokat feldolgozó hardveres egység lelke az izraeli PrimeSense által kifejlesztett PS1080-A2 kódú chip, melyhez egy 64MB-os, Hynix gyártmányú DDR2 memóriablokk kapcsolódik. Az eszköz hűtésére egy ventilátor is helyet kapott a többi alkatrész mellett – az Xbox 360-as játékkonzolok esetében túlmelegedési problémák léptek fel, azonban a Kinect esetében a hűtés remekül működik. [9]

2.3.1. Érzékelők

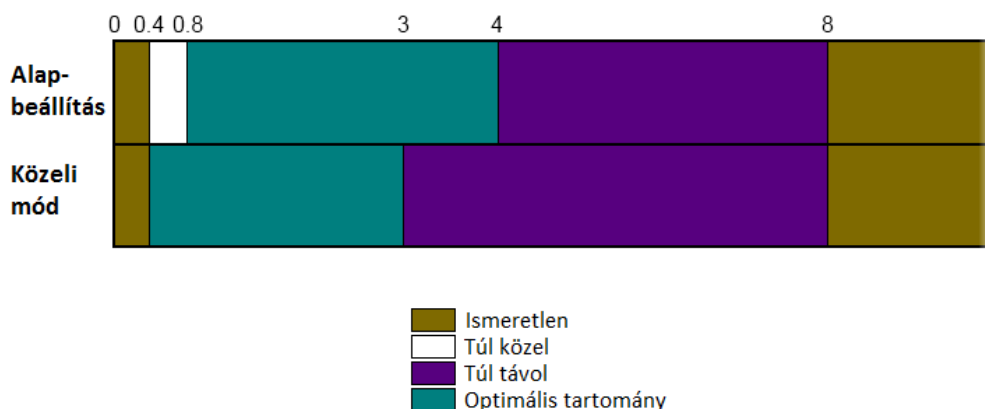
A különböző érzékelők képi- és hanginformációkat is rögzítenek. A képi jelfeldolgozás alapja az előlapon található két kamera. Az egyik egy teljesen közönséges színes VGA kamera (640*480 pixeles felbontással). A belátott tartomány vízszintesen 43° , míg függőlegesen 57° , a képfrissítés gyakorisága 30 kép másodpercenként. [10]



5. ábra Érzékelők és a motor elhelyezkedése [10]

A másik is egy hasonló VGA kamera, azonban ez az infravörös tartományban érzékel, és a három színinformáció rögzítése helyett monokróm üzemmódban működik. Ehhez tartozik a készülék előlapjának átellenes oldalán található IR projektor, ennek segítségével a készülék a környezetére apró infravörös pontokat vetít egy pszeudorandom minta alapján. Az infravörös kamerát sokszor csak mélységérzékelő kamerának hívják a különböző leírásokban, és valóban, olyan bitfolyam érkezik a berendezés felől, melyben minden pixelhez egy-egy távolságérték van rendelve.

A mélységérzékelés a belátott szögtartomány korlátjain kívül további megszorításokat tartalmaz. A 40 cm-nél közelebbi és 8 m-nél távolabbi objektumok teljesen értelmezhetetlenek, és optimális feldolgozás csak 80 cm és 4 m között lehetséges. A Kinect for Windows egyik újítása, hogy bevezették a közeli módot, ugyanis rengeteg visszajelzés érkezett arról, hogy kis lakásban rendkívül körülményes a megfelelő távolság kialakítása. Ebbe az üzemmódba átkapcsolva az optimális tartomány 40 cm-től 3 m-ig terjed.



6. ábra Működési tartományok [10]

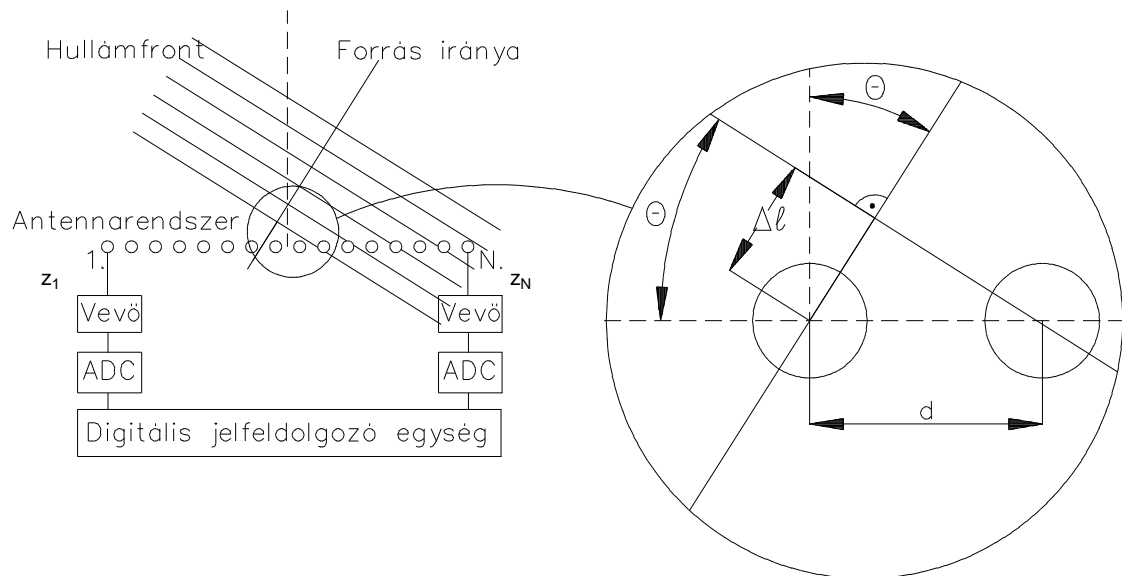
A hanginformációkat négy mikrofon dolgozza fel. Ezek mindegyike 24 bites A/D átalakítóhoz kapcsolódik. A mikrofon sor konstrukciójából fakadóan képes nyálábformálásra, ami egyrészt a

visszhang- és zajnyomáshoz, másrészt pedig a forrás irányának felderítésére alkalmas, ami több játékos nyomon követése esetén elengedhetetlen. Nyalábformálással 11 előre rögzített iránykarakterisztika közül lehet választani $\pm 50^\circ$ között 10° -os lépésekben, valamint lehetőség van az optimális karakterisztika automatikus felderítésére. A Microsoft Speech API használatával lehetőség nyílik beszédfelismerésre is. [10]

2.4. Az érzékelők működési elvei

2.4.1. Nyalábformálás

A Kinect esetében a mikrofonosor akusztikus nyalábformálásra képes. Nagyfrekvenciás rendszerek esetében szintén szükséges lehet a nyalábformálás – körsugárzó antennák helyett valamilyen irányítottággal rendelkező antennák sokhelyütt előfordulnak. Erre megoldást jelenthet egy megfelelő geometriával rendelkező apertúra antenna, például egy forgásparaboloid antenna, azonban antennarendszer használatával is kialakítható irányítottág. A különböző antennarendszerek leírása gazdag irodalommal rendelkezik, mikrofonosorok analóg módon tárgyalhatók, azonban ekkor elektromágneses hullámok helyett hanghullámok képezik a vizsgálat tárgyát. [11]



7. ábra θ szögből érkező síkhullám [11]

Induljunk ki egy lineáris ekvidisztáns antennarendszerből, mely N elemű. (Tehát N darab antenna egy vonalban, egymástól egyenlő távolságra helyezkedik el.) Érkezzék θ szögből egy síkhullám. Így az egyes antennákon adott időpillanatban a vett jel csupán fázisában különbözik. A fáziskülönbség az úthosszkülönbségre vezethető vissza, mely a

$$\Delta l = d \cdot \sin \theta$$

képletből számítható. Ebből adódik az elemek közötti fáziskülönbség:

$$\Delta\Phi = \Delta l \cdot \frac{2\Pi}{\lambda} c = d \cdot \frac{2\Pi}{\lambda} \cdot \sin\theta$$

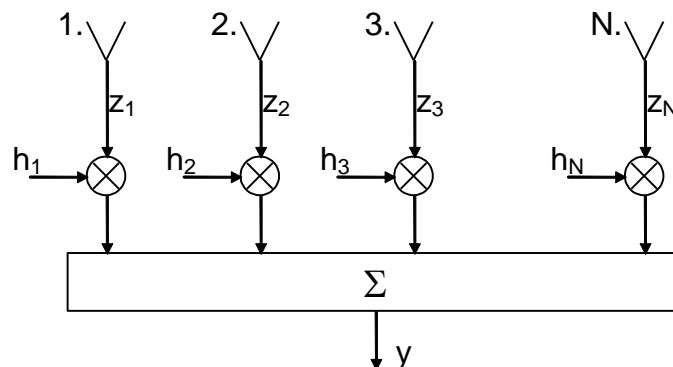
Így a k. elem fázisa:

$$\Phi_k = k \cdot \Delta\Phi = k \cdot 2\Pi \cdot \frac{d}{\lambda} \cdot \sin\theta$$

A vett jel vektora az eddigiek alapján a következő:

$$\mathbf{s}^T(\theta) = [1 \quad e^{-\Delta\Phi} \quad e^{-2\Delta\Phi} \quad e^{-3\Delta\Phi} \quad \dots]$$

A cél egy olyan vevőberendezés létrehozása, mely a θ szögből érkező síkhullámokra a legérzékenyebb. Ez a vett jel fázismenetének kompenzálásával oldható meg. Erre az alábbi elrendezés lehet alkalmas:



8. ábra Antennarendszer [11]

Az antennákról érkező jelet egy-egy konstans állandóval szorozzuk, majd egy összegzőn keresztül kapjuk a kimenetet. Azonos szorzók esetén a merőlegesen beérkező síkhullám okozza a kimeneten a legnagyobb jelet, hiszen ekkor fázismódosítás nélkül, fázisban adódnak össze az egyes antennára beérkező jelek. θ szög felé irányítottság esetén a szorzóknak úgy kell viselkedniük, hogy az ebből a szögből beeső síkhullámok a szorzókkal kompenzálva az összegző bemenetére már fázisban érkezzenek. Ebből a kompenzáló súlyvektorra a következő adódik:

$$h(\theta) = \mathbf{s}^*(\theta)$$

Mivel $\mathbf{s}^T(\theta)$ a fentiek alapján

$$s^T(\theta) = [1 \quad e^{-\Delta\Phi} \quad e^{-2\Delta\Phi} \quad e^{-3\Delta\Phi} \quad \dots]$$

így a kompenzáló súlyvektor:

$$h^T(\theta) = [1 \quad e^{+\Delta\Phi} \quad e^{+2\Delta\Phi} \quad e^{+3\Delta\Phi} \quad \dots]$$

A Kinect esetében lényeges különbséget az antennák mikrofonokra cserélése nem jelent. Az viszont a fenti számítási lépésekhez képest újdonság, hogy az eszköz nem ekvidisztáns mikrofonosort alkalmaz. Ebben az esetben a probléma megoldása természetesen némileg bonyolódik, azonban létrehozható olyan konstrukció, mely egy ekvidisztáns mikrofonosor tulajdonságaival bír, azonban „foghíjas”, és ebből fakadóan kevesebb elemet tartalmaz. A nagyfrekvenciás irodalom ezt a konstrukciót minimum redundanciájú antennának nevezi. [12]

2.4.2. Mélységérzékelés

Mint sok más technikai megoldás esetében, a Kinect mérnökei is biológiai ismeretek alapján oldották meg az eszköz mélységérzékelését. Ebből kifolyólag érdemes röviden kitérni az emberi mélységérzékelésre.

Optikai mélységészlelés

Az optikai mélységészlelés csoportjába a két szem konvergenciája és a szemlencse akkomodációja tartozik. Bár egyik sem tudatosul bennünk, agyunkba eljutnak ezen információk, és távolságérzetet okoznak. A konvergencia esetén minél párhuzamosabb a két szem tengelye, annál messzebbre nézünk, közelebbi pontokra történő fixálás esetén a két szem egyre inkább az orr irányába fordul. Az akkomodáció a szem fénytörő rendszerének egyetlen változtatható elemének, a szemlencse domborúságának alkalmazkodását jelenti a fókuszban tartani kívánt távolsághoz. [13]

Képi mélységészlelés

A látás során a szem optikai rendszerén túl agyunk magából a képi információból is következtet a különböző objektumok távolságára.

- Atmoszférikus perspektíva: a távolban lévő tárgyakat elmosódottabbnak látjuk
- Takarás: az eltakart tárgy az eltakaró tárgy mögött található
- Relatív nagyság: a közelebbi tárgyak nagyobbak
- Relatív függőleges pozíció: a horizonthoz közelebbi tárgyak messzebb helyezkednek el
- Lineáris perspektíva: távolodó párhuzamos vonalakat összetartónak érzékelünk

Rengeteg optikai csalódás a fenti jelenségeken alapszik, többek között a Csodák Palotájában is felépített Ames-szoba, mely a perspektíva eszközével azt a hatást kelti, mintha az oldalirányba mozgó személy magassága folyamatosan változna. [14]



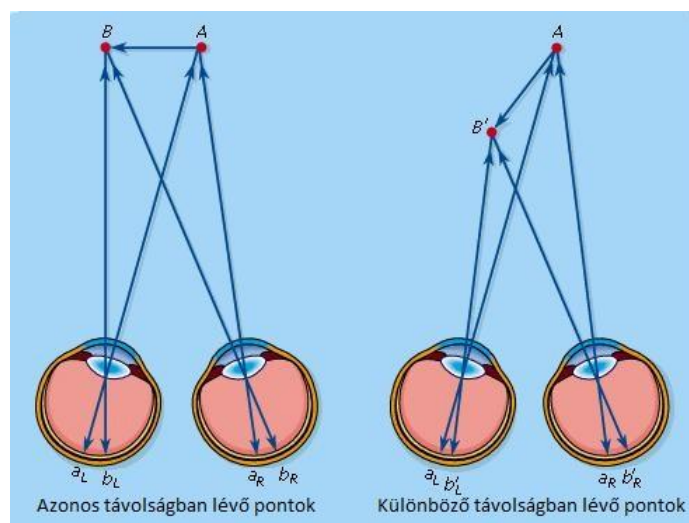
9. ábra A Cirkus Palotájában található Ames-szoba [14]

Mozgásparallaxis

Fejünket mozgatva is mélységinformációhoz juthatunk. Legjobb szemléltető példa erre az, hogyha kinézünk egy mozgó autóból, a fixált ponttól közelebb elhelyezkedő tárgyak az autó haladási irányával ellentétesen, míg a távolabb elhelyezkedő tárgyak azzal megegyező irányba mozognak. [13]

Binokuláris látás

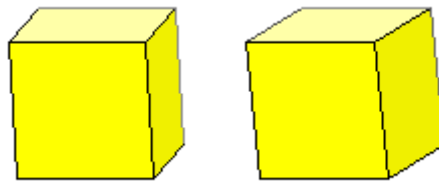
A legfontosabb mindközül azonban a két szemmel történő mélységérzékelés, térérzékelésünk leginkább erre támaszkodik. A jelenség alapja a retinális diszparitás. A szemlélőtől azonos távolságban lévő tárgyak képe a bal és jobb szem retináján is egymástól azonos távolságban váltanak ki ingerületet. Amennyiben az egyik tárgy közelebb van, úgy ez a szimmetria sérül, az eltérés nagyságából számítja ki agyunk a tárgyak egymáshoz képesti viszonyát. Amennyiben a fixált pont távolságától a másik tárgy távolsága egy bizonyos mértéknél jobban eltér, a tárgy két retinán létrejött képei már nem is fuzionálhatók, két képet látunk. [15]



10. ábra Retinális diszparitás [13]

A jelenség könnyen demonstrálható: helyezzük egyik ujjunkat a következő sorra, majd a sor olvasása közben (ez biztosítja a fixáció megtartását) közelítsünk ujjunkkal az arcunk felé! Szinte azonnal kettőt fogunk látni ujjunkból, és a közelítés mértékétől függően a két kép távolsága egyre növekszik. Bár zavarónak tűnhet a jelenség, a mindennapok során ritkán vehető észre, ugyanis az éppen feldolgozandó részletet mindig fixálja látórendszerünk, a retinális diszparitás szerepe pedig a térérzékelésben mutatkozik meg.

Amennyiben a két szemünk számára eltérő képet tudunk mutatni, lehetővé válik a fenti jelenséget kihasználva háromdimenziós érzet kiváltása. Az alábbi ábrából is egy térhatású kép alakul ki, amennyiben sikerül a kocka két nézetét külön-külön szemmel néznünk. Ehhez mindössze annyit szükséges tenni, hogy a távolba kell meredni, azaz a kép mögött, egy képzeletbeli távoli pontra kell fixálni, egészen addig, míg a két nézet fedésbe kerül – ekkor valójában egyik szemünk az egyik, míg a másik szemünk a másik képre fog tekinteni, agyunk pedig a látottak alapján összeállítja a térhatású kockát.



11. ábra A binokuláris látás segítségével térhatású kocka

A manapság egyre népszerűbb 3D filmvetítési technika is a binokuláris látást használja ki. Kétséggel ez a legmarkánsabb térérzékelési rendszerünk, azonban nem szabad megfeledkeznünk arról, hogy a fent leírt egyéb módszerek segítségével is mélységinformációhoz jutunk. A moziban állandóan a vászonra fókuszálunk, a fixálás csak a rendező által fókuszban tartott képrészletnél hozhat kielégítő eredményt, nem működik a mozgásparallaxis, stb. Ezek mind ellentétes információkat küldenek az agyunknak, és bár a binokuláris látás elnyomja ezeket az ingereket, így nem fog megszűnni a ránk erőltetett térérzékelés, ennek ellenére hányingert, fejfájást okozhat. [16]

A Kinect mélységérzékelése

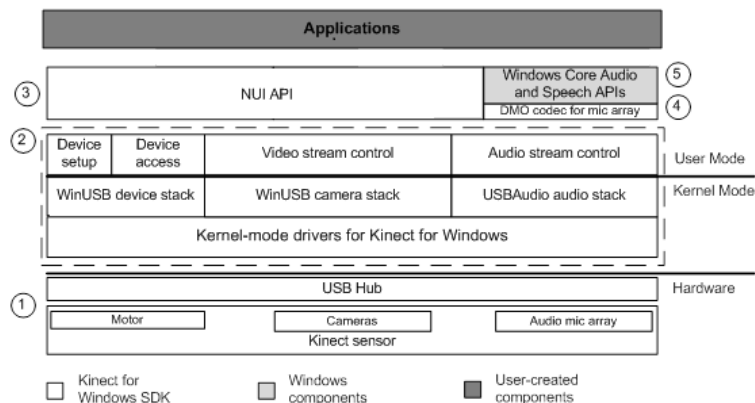
A távolságérzékelés tulajdonképpen a Kinect esetében is a binokuláris látáson alapszik. A jelenség két kamerával is szimulálható, a két képen előforduló azonos részletek pozíciójának különbségéből mélységinformáció nyerhető. A Kinect mélységérzékelésében azonban még egy csavar található: csupán egy kamera képe áll rendelkezésre, legalábbis ami a mélységérzékelést illeti. Azonban a sztereo feldolgozáshoz szükséges másik képet éppen az infravörös kamera vetíti a környezetre, így ez is rendelkezésre áll a mélységinformáció számításához. [17]

2.5. Kinect for Windows fejlesztői környezet

2.5.1. Rendszerkomponensek

A Microsoft által biztosított ingyenes fejlesztőkörnyezet számtalan módon segíti a programozók munkáját a Kinectet használó alkalmazások fejlesztése közben. Letöltése és telepítése két részletben történik. A Kinect for Windows SDK tartalmazza a driver-eket, valamint az eszközzel kapcsolatot tartó függvénykönyvtárakat, interfészeket. A Kinect for Windows Developer Toolkit további hasznos eszközöket nyújt a fejlesztéshez: példa forráskódokat, a Kinect Studio elnevezésű monitorozó eszközt, valamint a Face Tracking SDK-t, mellyel elérhetővé válik a későbbiek során bemutatásra kerülő arckövető funkció.

A szenzor felől érkező adatfolyamokat a különböző programok az úgynevezett Natural User Interface (NUI) segítségével érhetik el. Az SDK teljes architektúráját az alábbi ábra szemlélteti:



12. ábra SDK architektúra [10]

1. Hardver komponensek: Ebbe a csoportba tartoznak az érzékelők tárgyalásánál bemutatott kamerák és a mikrofonsor, az elevációs szöveget beállító motor, valamint a számítógéppel kapcsolatot teremtő USB csatoló.
2. Driver-ek: Lehetővé teszik a hang és videó adatfolyamok kezelését, biztosítják a mikrofonsor elérését a Windowsban található szabványos audio API-k segítségével, valamint lehetővé teszik, hogy egy alkalmazás több Kinecthez is kapcsolódhasson egyidejűleg. (Azonban egy Kinect egyszerre csak egy alkalmazással állhat kapcsolatban.)
3. Natural User Interface: Függvénykönyvtár az adatfolyamok kezeléséhez, valamint a csontvázkövetéshez.
4. DirectX Media Object: Ez a komponens valósítja meg a mikrofonsor nyálábformálását, és határozza meg a hangforrás irányát.
5. Windows 7 függvénykönyvtárak: A Windows 7 SDK, illetve a Microsoft Speech SDK segítségével szintén elérhetők az eszköz bizonyos funkciói.

A függvénykönyvtárak C++, C# és Visual Basic nyelvekre implementáltak. [10]

2.5.2. Adatfolyamok

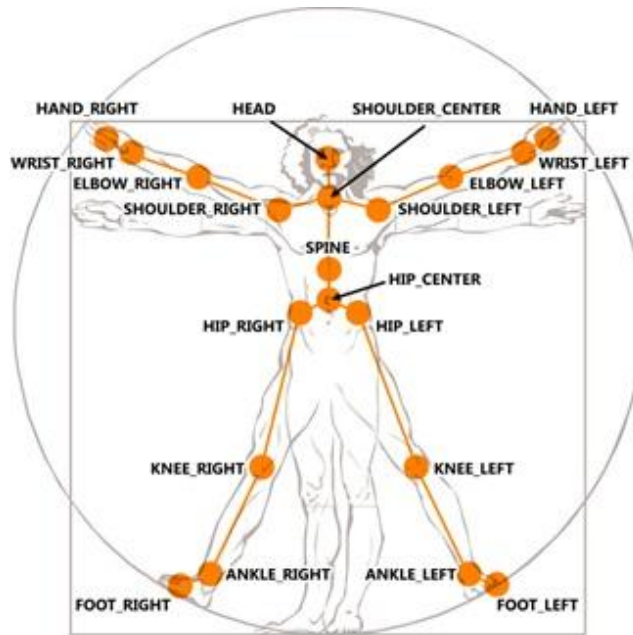
A Natural User Interface segítségével elérhető a Kinect által szolgáltatott különböző adatfolyamok. A NUI API inicializálásakor meg kell határozni a használni kívánt adatfolyamokat, a későbbiekben másik adatfolyamot elérni csak az eszköz újrainicializálása után lehet. Az érzékelőktől hang-, színes kép- és mélységi adatfolyam kérhető le. A függvénykönyvtár rengeteg eszközt nyújt az adatfolyamokkal történő manipulálásra is.

Az adatfolyamok egyes adatkeretei szelektív hívással vagy eseménymodell segítségével érhetőek el. Előbbi esetben a forráskódban elhelyezett megfelelő függvény a keret rendelkezésre állásáig felfüggeszti a program végrehajtását, míg utóbbi esetben eseménykezelő függvények társíthatók az adatfolyamokhoz, melyek a keretek elkészültekor automatikusan hívódnak meg. [10]

2.5.3. Csontvázkövetés

A Kinect talán legnagyobb sikerű funkciója, hogy képes az előtte állókat automatikusan detektálni, és különböző testrészeit követni. A rendszer egy-egy játékos 20 testrészének pozícióját határozza meg minden képkocka alapján, ezek az alábbiak:

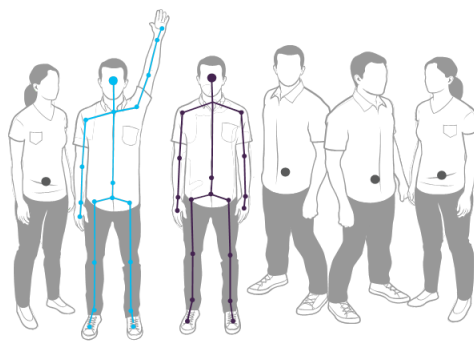
- Fej
- Jobb kézfej
- Bal kézfej
- Jobb csukló
- Bal csukló
- Jobb könyök
- Bal könyök
- Jobb váll
- Bal váll
- Vállvonal közepe (nyak)
- Gerinc
- Csípő közepe (köldöktájék)
- Csípő jobboldala
- Csípő baloldala
- Jobb térd
- Bal térd
- Jobb boka
- Bal boka
- Jobb lábfej
- Bal lábfej



13. ábra A csontvázkövetés által támogatott testrészek [10]

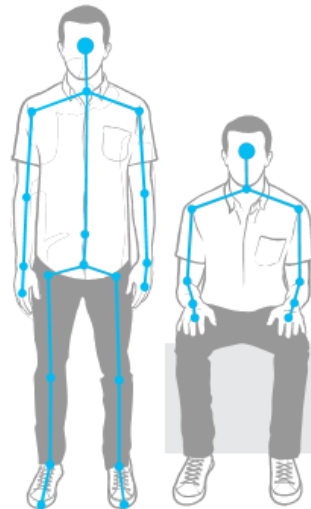
Az ábra a megfelelő függvényekben közvetlenül alkalmazható elnevezéseket tartalmazza. Fontos megjegyezni, hogy az ujjak pozíciójáról semmiféle információt nem szolgáltat a szenzor.

A Kinect számítási kapacitása is véges, éppen ezért a követhető emberalakok száma limitált, nem csak a kamerák által belátott terület geometriai tulajdonságai szabnak határt ennek. Egyidejűleg 6 csontváz követhető, azonban ebből aktívan csupán kettő. Passzív követés esetén a csontváz egészéről kapunk csak információt, az egész csontvázat egyetlen pont pozíciója írja le csupán.



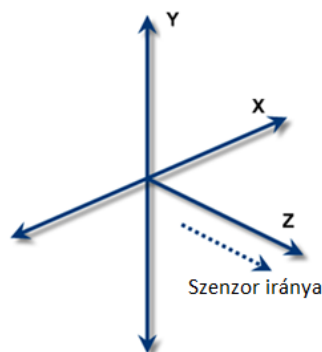
14. ábra Aktív és passzív követés [10]

Aktív követés esetén érhető el a fentebb felsorolt 20 testrész egyedi meghatározása. A csontvázkövetés funkció csak álló emberalakokat képes teljesen felismerni, fekvő személyek követése egyáltalán nem támogatott, ülő testhelyzetben pedig a felsőtestről, karokról kapunk információt 10 pont erejéig. További megkötés, hogy a készülék fele kell nézni, teljesen oldalvást, vagy háttal állva a működés bizonytalan.



15. ábra Álló és ülő mód [10]

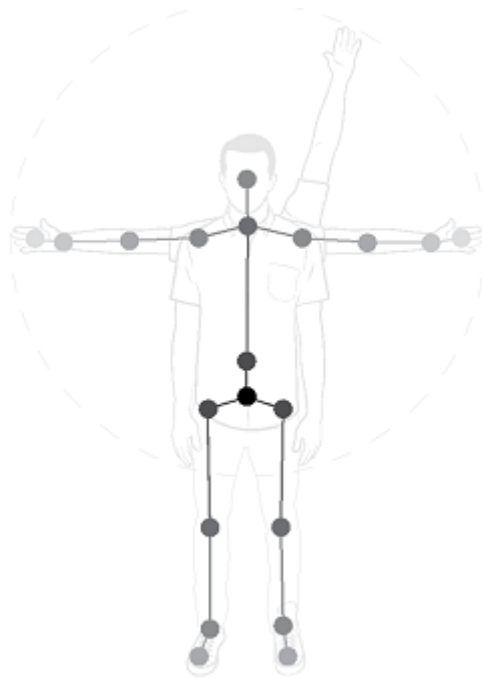
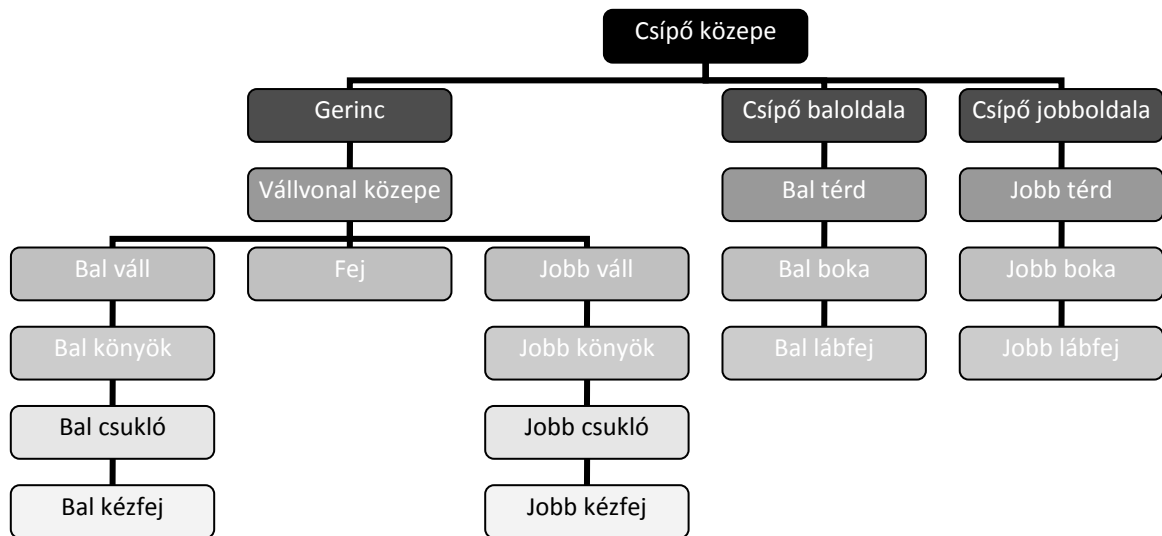
A testrészek követése során a pozíciók egy a készülékhez rögzített koordinátarendszerbe képződnek le. A koordinátarendszer Descartes-féle, az x , y , z tengely ebben a sorrendben jobbsodrású rendszert alkot. A z tengely a szenzor érzékelési irányába néz, az x tengely a két kamerát összekötő egyenesen helyezkedik el, és a készülékkel szembe fordulva jobb oldalt található a pozitív tartománya. Nem szabad megfeledkeznünk arról, hogy egy motor segítségével a berendezés fejének elevációja változtatható, ekkor az x tengely körül elfordul az y és a z tengely. A csontvázak és testrészek pozícióinak értékei várhatóan az x tengely mentén $-2,2$ m-től $2,2$ m-ig, az y tengely mentén $-1,6$ m-től $1,6$ m-ig, míg a z tengely mentén 0 -tól 4 m-ig terjedhetnek az érzékelő korlátjaiból fakadóan.



16. ábra A Kinect koordinátarendszere

A testrészek pozícióján túl lehetőség van a csontok orientációjának lekérdezésére is. A reprezentált csontok nem valódi, biológiailag értelmezhető csontok, hanem a megfelelő testrészeket összekötő szakaszok, azonban az ezek segítségével felépíthető csontváz tökéletesen alkalmas egy virtuális test felépítéséhez, a játékok során alkalmazott avatárok létrehozásához nem szükséges ennél részletesebb, biológiailag pontosabb leírás.

A csontok reprezentációjához a korábban bemutatott 20 testrész egy fa gráfba rendezett, melynek gyökere a csípő középrésze, a csontok pedig a gráf élei. Ülő testhelyzetben történő követésnél a gráf gyökere a vállvonal középrésze.

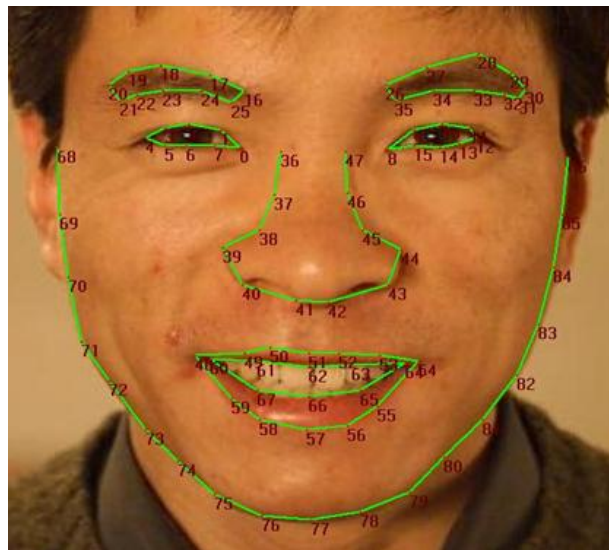


17. ábra A testrészek hierarchikus szervezése [10]

A csontok orientációját a gráf megfelelő éléhez tartozó két pont közül a gyermekelem tárolja. A szülőelem orientációjához képesti eltérés kódolt. A játékos abszolút orientációját a gyökérem (a csípő közepe, ülő testhelyzet esetén a váll közepe) tartalmazza, a kamerával tökéletesen szembefordult állapothoz képesti eltérés kódolt. [10]

2.5.4. Arcok követése

A Face Tracking SDK segítségével lehetőség nyílik nemcsak az érzékelő előtt állók csontvázát, hanem az arcukat is nyomon követni. A függvények itt is a csontvázkövetésnél már bemutatott koordináta-rendszert alkalmazzák. A különböző metódusok a Kinect színes kép és mélységi adatfolyamával dolgoznak, így ezek minősége jelentősen meghatározza a végeredményt. A jó eredmény érdekében az alakoknak jól megvilágítottaknak, a képnek élesnek kell lennie. Természetesen annál pontosabb lesz az arc egyes részleteinek detektálása, minél nagyobb területet foglal el a képből.

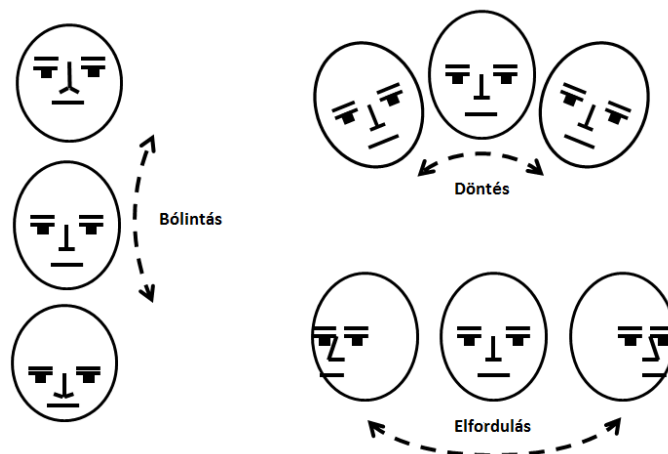


18. ábra Arckövetés által meghatározott pontok [10]

Az arc feldolgozása során számos információ kérdezhető le. 87 pont kerül detektálásra, ezek egy tömb elemeiként egyenként elérhetők. Ezen pontok csak síkbeli információval rendelkeznek, a tömb elemei a szenzortól megkapott RGB kép egy-egy pixelére mutatnak. (A képen 13 pont nincs jelölve.) A fej bólintása, döntése és elfordulása szintén lekérdezhető, az eredmény fokokban értelmezendő.

	Pozitív irány	Működési tartomány	Optimális tartomány
Bólintás	Felfelé	20°	10°
Döntés	Bal váll felé	90°	45°
Elfordulás	Baloldalra	45°	30°

1. táblázat A fej különböző mozgásai



19. ábra A fej különböző mozgásai [10]

A Candide3 modell alapján a száj nyitottsága, az ajkak összeszorításának mértéke, a szemöldökök helyzete, és még néhány dinamikus paraméter szintén egyből a programozó rendelkezésére áll, éppúgy, mint a szem méreteire, távolságára, az orr pozíciójára, stb. vonatkozó statikus paraméterek. Ezen statikus paraméterek alkalmasak arra, hogy a kamera elé később visszatérő személyt arcvonásai alapján felismerhesse a rendszer.

Ezen kívül – szintén a Candide3 modell alapján – egy háromdimenziós modell is megalkotható az arcról, bár ez a modell nem jön létre automatikusan az arckövető függvények hívásakor, hanem a korábban említett dinamikus és statikus Candide3 paraméterekből számítható, az API tartalmazza az ehhez szükséges függvényeket is. [18]

2.6. Kinect tesztprogram

A projekt megvalósítása során a Kinect csontvázkövetés funkciója a legkézenfekvőbb megoldás. A Kinect kapcsolódó funkcióinak feltérképezésére egy tesztprogramot írtam, mely egy csontvázat, és annak egy adott testrészét monitorozza, az eredményeket vagy szöveges módban, vagy a testrész pozícióját grafikus úton jeleníti meg.

A Kinect for Windows SDK használatához a Visual Studio program projektjének beállításain néhány változtatást kell alkalmazni. Egyrészt a megfelelő header fájlok eléréséhez az „Include Directories” mezőhöz hozzá kell adni a $\$(KINECTSDK10_DIR)\inc$ bejegyzést. Másrészt az SDK programkönyvtárát is láthatóvá kell tenni a linker számára, így az „Additional Dependencies” mezőhöz hozzá kell adni a $\$(KINECTSDK10_DIR)\lib\x86\Kinect10.lib$ fájlt.

A Natural User Interface használatához ki kell adni az `#include <NuiApi.h>` fordítói direktívát. A `Windows.h` header fájljal hasonlóképpen kell eljárni (akkor is, ha csak interfészen belüli függvények szerepelnek), egyébként a fordítás sikertelen lesz.

A program először a kinect.config fájlból konfigurációs adatokat olvas ki. Minden indulás alkalmával bekérni ezen adatokat hosszadalmas procedúra volna, forráskódban történő rögzítésük pedig a gyakori módosítások miatt a fordítási idő miatt körülményes, ezért hoztam létre a konfigurációs fájlt. Mivel az adatok beolvasása a Kinect programozása szempontjából nem releváns, így ennek lépéseit nem részletezem.

A szenzor használatához először inicializálni kell azt. Több szenzor is használható egy programmal, azonban egy készüléket csak egy program használhat. Éppen ezért az inicializálás erőforrásként le is foglalja az eszközt, így más program ezután már nem képes végrehajtani rajta az inicializációt, míg fel nem szabadítja az előző program. Amennyiben csak egy Kinectet használunk a rendszerben, lehetőség lenne INuiSensor objektum nélkül megoldani a feladatot, azonban számomra átláthatóbb kódot eredményezett az objektumhoz rendelés, és az IntelliSense használatát is megkönnyítette. Éppen ezért az inicializálás előtt a NuiCreateSensorByIndex(0,&szenzor) függvény segítségével a „szenzor” változóhoz rendeltem a 0. azonosítójú eszközt. Ezt követhette a szenzor->NuiInitialize (NUI_INITIALIZE_FLAG_USES_SKELETON) függvény, mely az inicializációt végrehajtotta. Az argumentumként megadott flag biztosítja, hogy a csontvázkövetést a későbbiekben használhassuk. Mindkét függvény hibakóddal történő visszatérése megszakítja a program működését. A program végén az erőforrást a NuiShutdown függvény szabadítja fel.

Az inicializálás után a program ellenőrzi a fej elevációs szögét a NuiCameraElevationGetAngle függvény segítségével. Amennyiben eltérés mutatkozik a kívánt értékkel, a NuiCameraElevationSetAngle függvény segítségével az új értéket állítja be az eszközön.

Ezután ciklikusan elemzi a Kinect által szolgáltatott adatokat. Először a NuiSkeletonGetNextFrame(0,&csontvazallapot) lekéri a csontvázak aktuális állapotát. Itt az első paraméter egy késleltetés, csak ennyi idő után fogunk állapotleírást kapni. Természetesen a paraméter 0 értéke esetén is meg kell várnunk az adatok rendelkezésére állásának idejét. A NUI_SKELETON_FRAME típusú csontvazallapot nevű változóban tárolódnak az adatok. Az objektum SkeletonData tömbjében az egyes csontvázak NUI_SKELETON_DATA típusú objektumban találhatóak. Ezen objektum eTrackingState tagváltozója NUI_SKELETON_TRACKED értékű, amennyiben az adott csontváz aktívan követett, egy ciklusban a program kiválasztja a tömb első ilyen objektumát, és a későbbiekben ezzel dolgozik.

A csontváz koordinátáit a Position.x, Position.y, és Position.z tartalmazzák, a választott testrészt pedig a SkeletonPositions[követendoTestreszID].x, .y és .z változók. Szöveges módban ezen adatok a konzol képernyőjére kiírásra kerülnek. Grafikus módban a testrész x és y koordinátája alapján 24 sor magas és 80 oszlop széles területen az értékek megfelelő átskálázásával egy X-et rajzol ki a program,

míg a 25. sorban > jelek mennyisége jelzi a z koordináta értékét. Egy adatgyűjtő mód is része a programnak, ekkor a mért értékek egy .m kiterjesztésű fájlba kerülnek, a Matlab program által értelmezhető formában, a későbbi kiértékelés megkönnyítése érdekében. A mért értékeken kívül egy referenciaérték is bekerül a fájlba, ezt minden egyes mérési pont előtt a program bekéri a felhasználótól, ez biztosítja azt is, hogy az egyes mérési pontok csak a felhasználó beavatkozására kerüljenek rögzítésre.

3. A Kinect pontossága

3.1. A mérés leírása

Mivel a Kinectet eredetileg szórakoztatóelektronikai eszköznek tervezték, jogosan merül fel a kérdés, pontossága megfelelő-e komolyabb alkalmazások számára. A Kinect legfőbb erénye az emberalakok automatikus detektálása, ráadásul a projekt megvalósítása során is ezt a funkciót használom ki, éppen ezért pontosságának ellenőrzése is emberi alakok felismerése közben a legcélravezetőbb. Könnyen belátható, hogy élő emberrel elvégezni a mérést egyrészt körülményes volna, másrészt lehetetlen volna biztosítani a referenciatávolságok pontos betartását az állandó mozgolódás miatt. Éppen ezért szükség van egy próbababára, melyet a Kinect emberalakként ismer fel, azonban a mérés szempontjából hátrányos emberi tulajdonságokat nélkülözi. Egy ruhaboltbeli baba vélhetően tökéletesen megfelelt volna erre a célra, azonban ez nem állt rendelkezésre, építeni kellett egy saját mérőbabut.



20. ábra Mérőbábu

A szerkezet vázát egy ruhatartó adta, erre volt ráterítve egy kabát, mely a törzset, ujjai pedig egyszersmind a baba karjait reprezentálta. A ruhatartó kerete a kabát alatt lábként funkcionált, bár egy valóságos lábhoz képest valamelyest vékonyabb a keret vastagsága. A lábfej kiterjedését a ruhatartóra felragasztott két papucs jelenítette meg.



21. ábra A mérőbábu feje

Kardinális kérdés volt a fej pontos kialakítása. A fej szemből nézeti kontúráját egy körülvágott kartonlap alkotja, később ez a lap került rögzítésre a ruhatartó felső részén. A lapon színes ceruzával megrajzoltak az ajkak, a szemek, a szempillák, valamint a szemöldökök. Arcunk egyik legmarkánsabb eleme orrunk, hiszen mélységi kiterjedése ennek a legnagyobb. Mivel a Kinect mélységi információt is mér, az orrot rajzolás helyett egy száraz kenyérvég alkotta, mely egy papírból készült támasztószervezet segítségével vette fel a kívánt pozíciót. A fej tetejére a csőszerelésből jól ismert kóc került hajként.



22. ábra Mérési elrendezés

A mérési elrendezésben a számítógéphez kapcsolt Kinect előtti térben helyezkedett el a próbababa, a méréshez szükséges referenciatávolságokat egy mérőszalag segítségével lehetett beállítani. A mérőszalag a képen a mélységi mérésekhez van beállítva, az x koordináta mérésénél természetesen merőleges irányban volt kifeszítve. A Kinect y koordinátájának pontossága nem képezte vizsgálat tárgyát, ugyanis túlságosan bonyolult feladat lett volna a baba függőleges irányba történő mozgatása – ráadásul úgy, hogy a referenciaértékek pontosan beállíthatók legyenek. Szerencsére a próbababát emberi alakként detektálta a műszer, azonban néha a felismerés sikertelen volt, ami megnehezítette a mérést. Az érzékelési tartományok határán a mérés megkezdése szinte lehetetlen volt, a babát egy alkalmasabb pozícióban fel kellett ismertetni a rendszerrel (olykor a karok mozgatása is szükséges volt ehhez), majd csak ezt követően lehetett a tartomány határához kitolva megkezdni az adatgyűjtést. További problémát jelentett, hogy a mérés során az eszköz mindvégig a próbababa helyzetét detektálja, az esetlegesen pozicionálás céljából feltűnő egyéb emberalakok ne zavarják össze a mérést. Ennek a kritériumnak a mérések jelentős részénél csak úgy lehetett eleget tenni, hogy a pozicionálást végző személy mindvégig a próbababa mögött, annak takarásában tartózkodott.

3.2. Mérés egyhelyben

A méréssorozat kezdéseként a bábút egy-egy rögzített pozícióban hagyva mértem a pozícióját. A mérési eredmények a Kinect koordinátarendszere szerint értelmezendők, azaz az x koordináta az oldalirányú-, az y koordináta a függőleges elmozdulást, míg a z koordináta a mélységi információt hordozza. Mindkét mérési pont a készülék előtt, a kamerák tengelyével egyvonalban helyezkedett el, az egyik 2 m, míg a másik 3,5 m távolságban. A csípő közepét jelző pont (azaz a fő referenciapont) adatait lekérve, a mérési adatokat kiértékelve a következő eredmények adódtak eredményül az egyes koordináták szórására:

	2 m-es távolságban	3,5 m-es távolságban
x koordináta	0,45 mm	2,0 mm
y koordináta	0,61 mm	2,4 mm
z koordináta	0,23 mm	0,53 mm

2. táblázat Szórások (csípő közepe)

Amennyiben a csípő közepe helyett a fej pozícióját jelző pont helyzetét vizsgáltuk, abban az esetben a szórások az alábbiak:

	2 m-es távolságban	3,5 m-es távolságban
x koordináta	1,2 mm	7,2 mm
y koordináta	1,5 mm	3,6 mm
z koordináta	1,8 mm	10,6 mm

3. táblázat Szórások (fej)

A szórások milliméteres, vagy annál kisebb nagyságrendbe estek. Egyedül a fej mérése esetén, 3,5 m-es távolságban a z koordináta mért eredményeinek szórása éri el a centiméteres nagyságrendet, azonban egy szórakoztatóelektronikai eszköztől lévén szó ez az eredmény is kiválónak mondható.

Érdeemes megfigyelni, hogy a 3,5 m-es távolságban elvégzett mérések minden esetben rosszabb eredményt adtak a 2 m-nél elhelyezett mérőponthoz képest. Ennek az lehet a magyarázata, hogy a kamerákkal történő észlelés miatt a közelebbi objektumok felbontása nagyobb, hiszen a kamera egyes pixeleire eső fénynyalábok a tér egy adott szögtartományából érkeznek a kamerába, attól távolodva pedig a pixelhez tartozó szögtartomány alatt látott terület nagysága folyamatosan nő.

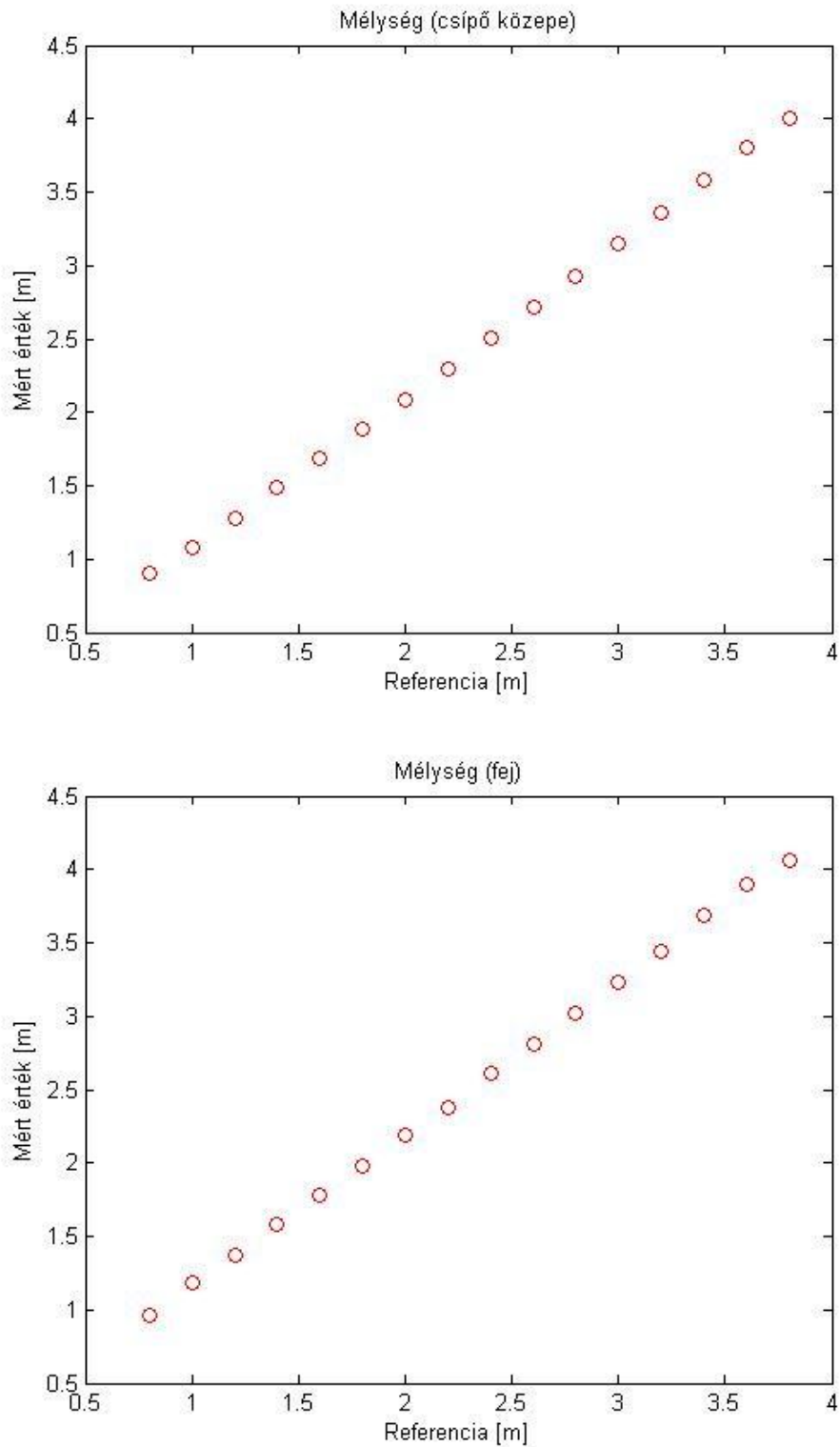
Az egyhelyben történt mérések – bár bizonyos szempontból ezek is tanulságosak – tulajdonképpen a további mérések pontossági követelményeinek megállapítására szolgáltak. Ugyanis ebben az esetben a mért értékek egymástól való eltérései egyedül a Kinectnek tudhatók be, a mért környezet a mért adatok felvétele között nem változott. Éppen ezért az ezt követő mérések során legalább ekkora pontosságot kell elérni, hogy a bábu különböző mérési pontokra történő pozicionálásának hibája ne zavarja meg a mérést. Ezt csak részben sikerült teljesíteni – a másfél méteres próbababát minden igyekezetem ellenére lehetetlen lett volna 10^{-4} m-es pontossággal pozicionálni, azonban a milliméteres pontossági kritériumot sikerült teljesíteni. Ezt figyelembe kell venni a további mérések értékelésénél.

3.3. Mélységi mérések

A Kinect képes az előtte lévő objektumok készüléktől mérhető távolságát rögzíteni, ez a funkció a csontvázkövetés esetében is megvalósul, így méréssorozatokat ennek vizsgálatával folytattam. A mérés során a mérőbaba az eszköz előtti vonalban, a Kinect dokumentációjában rögzített 80 cm és 4 m közötti optimális érzékelési tartományban mozgott, és 20 cm-es közönként mértem le a baba pozícióját a Kinect segítségével. Bár a dokumentáció 4 méteres távolságig határozza meg az optimális

tartományt, ebben a mérési pontban már nem tudtam rögzíteni eredményt, minden esetben 3,8 m volt az utolsó, még felvett érték.

A mérést a csípő közepe és fej referenciaponttal is elvégeztem, az alábbi eredmények adódtak:



23. ábra Mélységi adatok

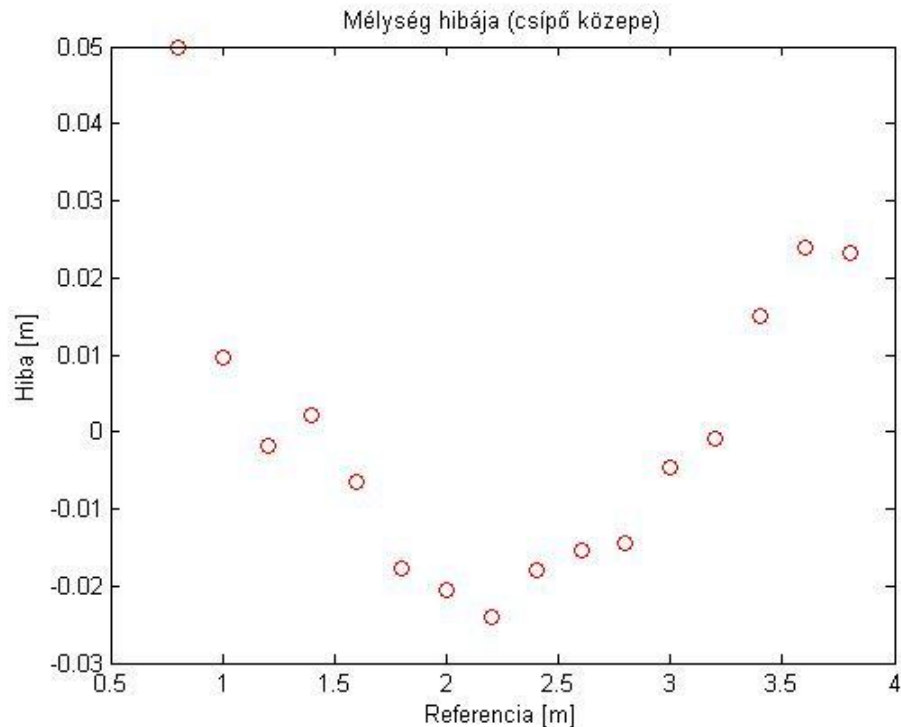
Rájuk nézve, mindkét grafikon a várakozásoknak megfelelő, lineáris jelleget mutatja. A legkisebb négyzetek módszerével illesztett egyenesek az alábbi paraméterekkel rendelkeznek:

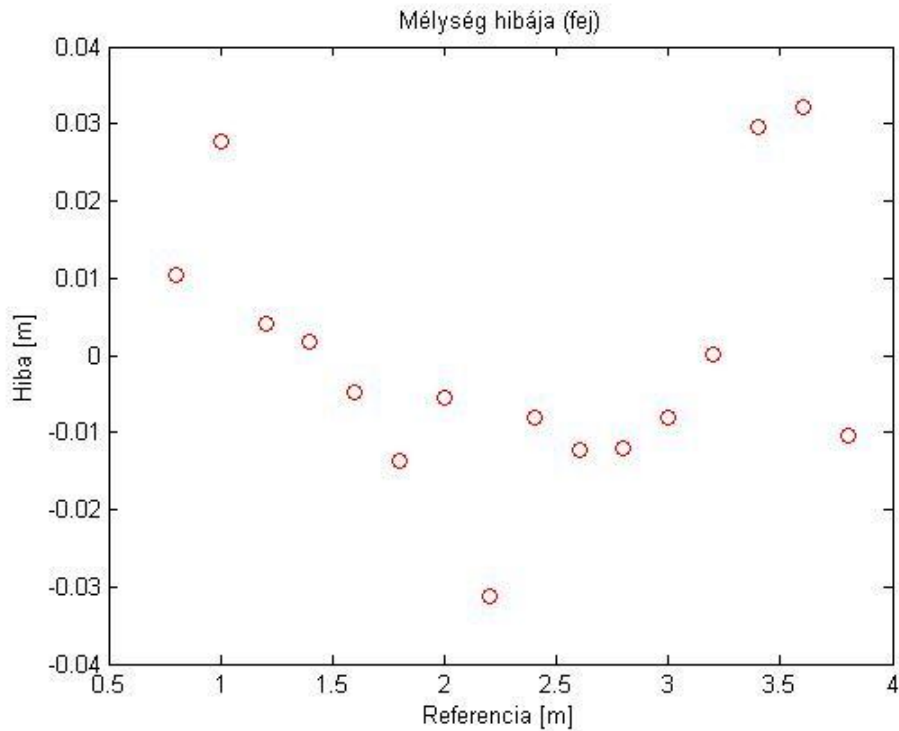
	Csípő közepe	Fej
Ofszet	0,0255 m	0,1197 m
Merekség	1,0419	1,0395

4. táblázat Mélységmérés illesztett egyenseinek paraméterei

Érdekes, hogy a fej mérése esetén az ofszet majdnem 10 centiméterrel nagyobb, azonban ez vélhetően annak köszönhető, hogy a papírból készített fejnek mélységi kiterjedése (az orrot leszámítva) tulajdonképpen nincs, a ruhakeretre történő rögzítésnek köszönhetően pedig a kabát elülső vonalához, így a detektált csípő közepe ponthoz képest ennyivel hátrébb helyezkedett el. A merekség mindkét mért pont esetében közel 1, ami megnyugvásra adhat okot, hiszen ez azt jelenti, hogy a Kinect az általa mért paramétereket megfelelően skálázza méter egységekbe.

Vizsgáljuk meg a mért értékek illesztett egyenestől való eltérését:

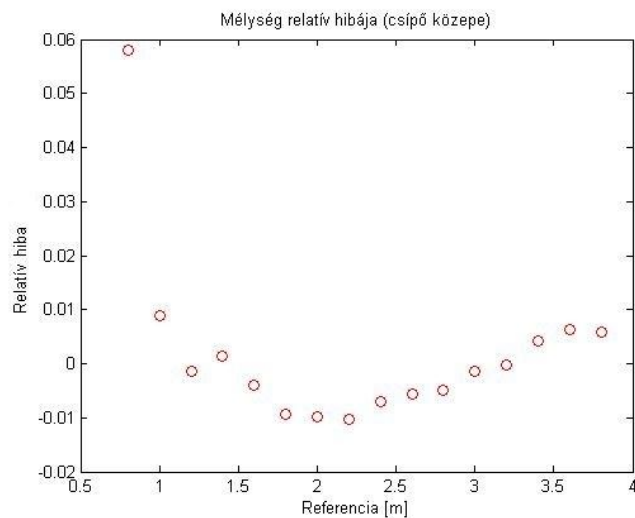


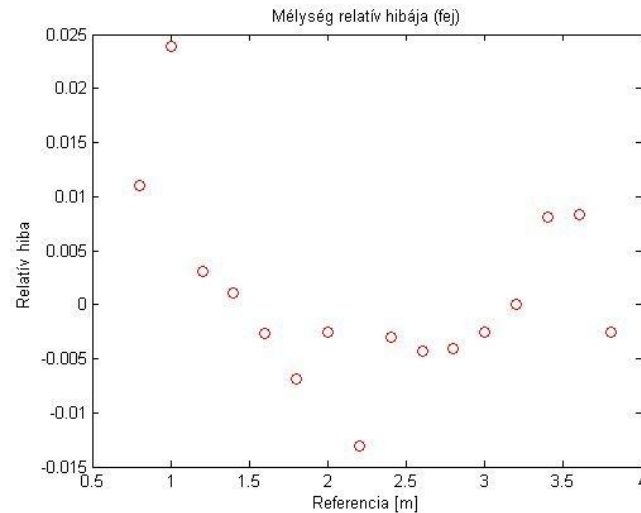


24. ábra Mélységmérés hibái

A csípő közepe esetén a legpontatlanabb a 0,8 m-es referenciapontban felvett érték, azonban ez is mindösszesen 5 centiméterrel tér el az illesztett egyenestől. Nem szabad megfélekednünk arról sem, hogy ez a mérési pont az optimális tartomány széle, így némileg pontatlanabb eredmények itt megengedhetők.

Bár a Kinect legtöbb alkalmazása esetében az abszolút hiba fontosabb paraméter, hiszen a készüléktől mérhető távolságunktól függetlenül várunk el egy bizonyos pontosságot, vizsgáljuk meg a mélységérzékelés relatív hibáját is:





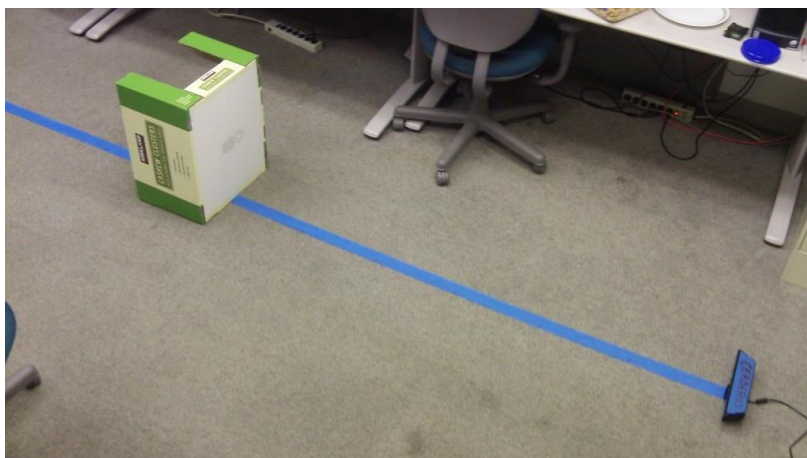
25. ábra Mélységmérés relatív hibái

Megfigyelhető, hogy a korábban már említett 80 centiméternél felvett értéket, vagy a fej mérésénél az 1 m távolságban mért értéket leszámítva minden eredmény 1%-os pontosságon belül van, vagy éppen csak túllépi azt.

Továbbá az is érdekes, hogy mindkét esetben a mérési tartomány szélein a mért értékek az illesztett egyenes megfelelő pontjaihoz képest nagyobbak, míg a mérési tartomány közepén, 2 méter környékén kisebbek.

3.4. Nyers mélységi adatfolyam

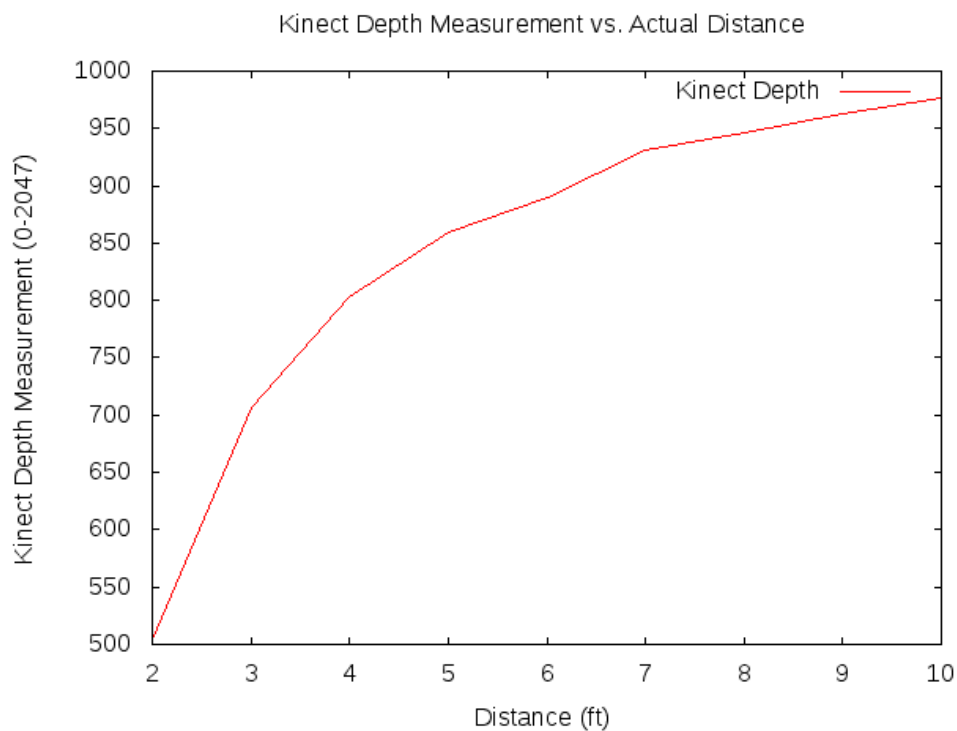
Nathan Crock emberi alakok vizsgálata helyett a Kinect nyers mélységi adatfolyamát vette górcső alá. A mérési elrendezés nem sokban különbözött az általunk felállítottól, a legmarkánsabb különbség az, hogy a mérőbábút egy kartondoboz váltja fel, melynek sima hátlapja kiválóan alkalmas a mélységi mérések lefolytatásához. [19]



26. ábra Nathan Crock mérési elrendezése [19]

A nyers adatfolyam 11 bites értékeket tartalmaz, amivel elvileg 2048 db különböző távolságérték lenne reprezentálható. Crock 2 és 10 láb között lábanként (körülbelül 30 cm) vett fel mérési pontokat. [19]

A lejjebb közölt ábrán jól látható, hogy a nyers adatfolyam egyáltalán nem lineáris, inkább logaritmikus jelleget mutat. Ez illeszkedik ahhoz a korábban már bemutatott jelenséghez, hogy a távolabbi objektumok felbontása kisebb, ezért pozíciójuk csak pontatlanabban mérhető ebben a tartományban, így ideálisabbnak tűnik a logaritmikus skála, amely a közelebb lévő és pontosabban detektált objektumokhoz több, míg a távolabbi és ezért kevésbé pontosan detektált objektumokhoz kevesebb különböző távolságértéket rendel.



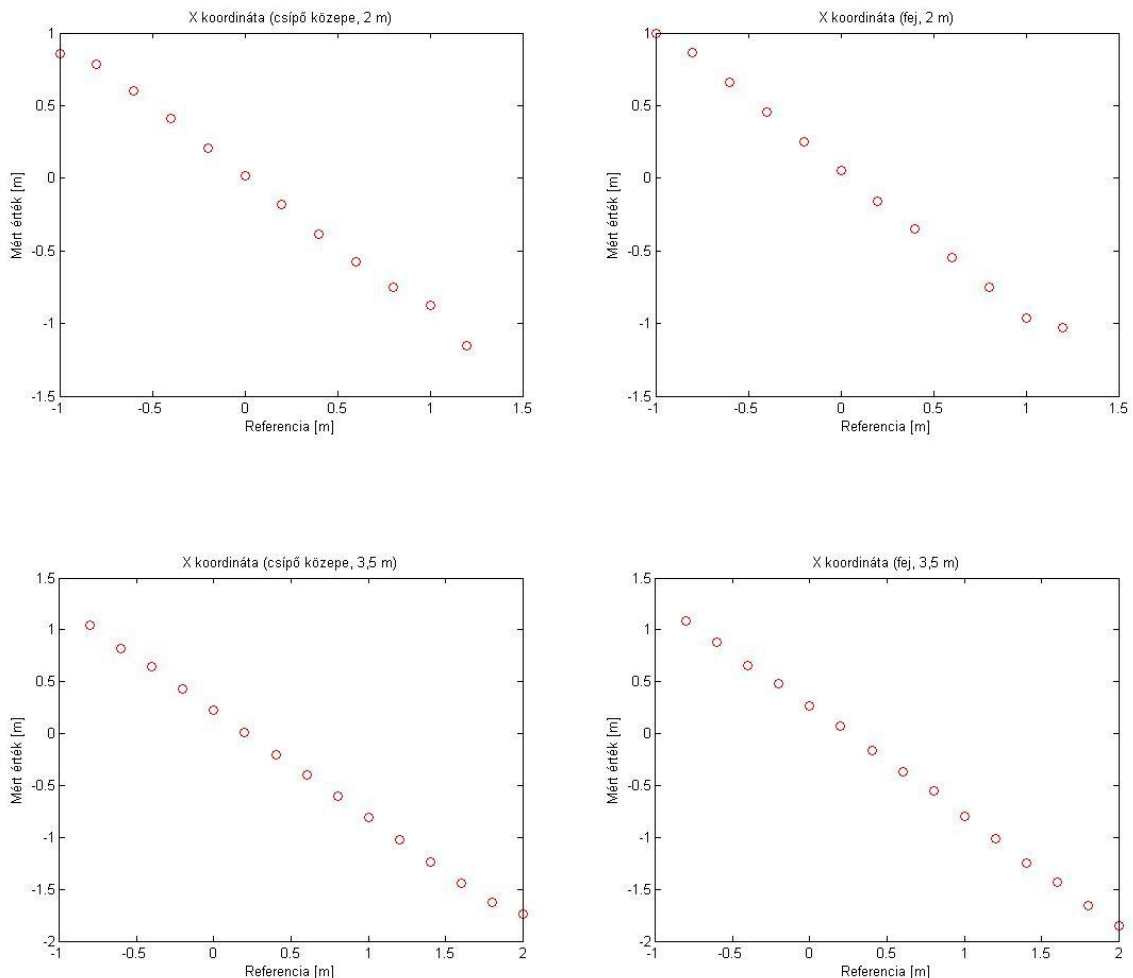
27. ábra Nathan Crock mérési eredménye [19]

Meglepő, hogy 2 lábhoz (ami már az optimálisan érzékelhető tartomány határánál közelebb, de még az értelmezhető tartományon belül van) már az 504-es értéket rendeli a szenzor. 10 lábhoz pedig a 976-os érték társul – ha feltételezzük, hogy a logaritmikus skála lineárisra válik, még akkor sem érjük el az értelmezhető tartomány felső határában, azaz 8 méteres távolságban a 11 bitből számítható 2047-es utolsó távolságértéket.

3.5. Oldalirányú mérések

A mélység érzékelésének pontossága után vegyük szemügyre az oldalirányú elmozdulások detektálásának kérdését. Két méréssorozatot folytattam le: először a kamerától 2 méter, majd 3 és fél méter távolságban helyeztem el a kamera irányára merőlegesen a mérőszalagot, és ennek mentén mozgott a bábu. Mindkét esetben a csípő közepének, majd a fejnek a pozícióját is megmértem.

A közelebbi és a távolabbi mérésnél is a kamera tengelyétől egy méterre jobbról (a kamera felé fordulva) kezdtem a mérést, majd 20 centiméterenként haladtam a másik irányba, egészen addig, ameddig a csontvázkövetés eredményt adott. Ez a 2 méterre lévő mérések esetében 1,2 méterre balra a kamera tengelyétől, míg a 3,5 méterre lévő mérések esetében 2 méternél jelentette az utolsó mérési pontot, így a közelebbi mérésnél 12, míg a távolabbinál 16 érték került rögzítésre.



28. ábra X koordináta mérési eredményei

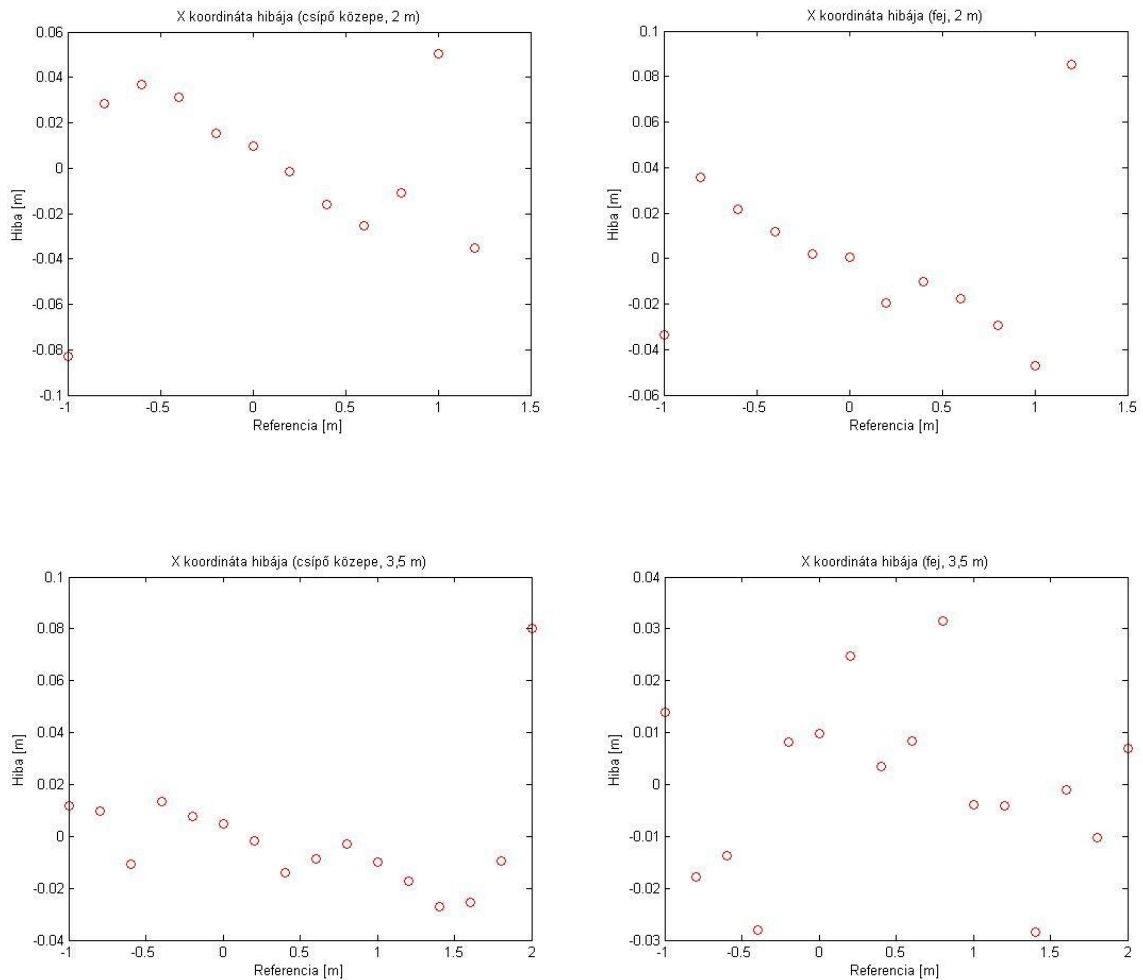
A későbbiekben az eredményekre illesztett egyenesek negatív meredeksége annak köszönhető, hogy a mérés közben alkalmazott referenciaértékek pozitív iránya pont fordítva lett meghatározva a Kinect koordinátarendszeréhez képest. A meghatározott meredekségek:

	Csípő közepe	Fej
2 méteres távolságban	-0,9347	-0,9722
3,5 méteres távolságban	-1,0206	-1,0566

5. táblázat X koordináta mérése során illesztett egyenesek meredeksége

2 méteres távolságban a valóságosnál némileg kisebb oldalirányú távolságokat becsül a Kinect, ezzel szemben a 3,5 méternél elvégzett mérés esetében a távolságokat felülbecsülte.

Az illesztett egyenestől való eltéréseket az alábbi grafikonok szemléltetik:



29. ábra X koordináta hibái

Jól megfigyelhető, hogy a kamera által belátott tartomány határához érve a mérés bizonytalanná válik. Ennek az az oka, hogy a csak részletében látott emberalakokat is feldolgozza a rendszer, azonban a nem látott részletek kiegészítése nem mindig megfelelő, így részben fedett emberalakok esetén az egyes testrészek pozícióinak detektálása teljesen kiszámíthatatlan.

3.6. Összegzés, a Kinect színpadi alkalmazásának korlátjai

Bár a projekt megvalósítását jelentősen megkönnyítette a Kinect alkalmazása, komoly hátrányokkal is számolnunk kell. A berendezés egy szórakoztatóelektronikai eszköz, ipari alkalmazhatóságával kapcsolatban kérdések merülnek fel, robusztussága, megbízhatósága valószínűleg nem elégítenék ki egy valódi alkalmazás követelményeit.

A Kinect pozíciódetektáló rendszere meglehetősen hatékonyan működik. A különböző irányokba történő mérések esetén a jellemző hibaértékek 1-2 cm körül mozogtak. Ez rendkívül precíz helymeghatározást tesz lehetővé, ráadásul közvetlenül lekérdezhető a színész fejének pozíciója, mely a projekt szempontjából kulcsfontosságú. Azonban az eszköz által érzékelt terület meglehetősen kicsi, mindössze egy közepes méretű szoba lefedésére alkalmas – az érzékelési tartomány elfér egy 4 m oldalélű kockában, ami még az olyan kisméretű színpadok méreteit sem fedi le, mint például a Centrál Színház Kisszínpada, mely körülbelül 8*5 méteres. Ez a korlát a maximálisan együtt kezelhető 4 db érzékelő csatlakoztatása esetén sem enyhül jelentősen, egy nagyobb terület lefedésének kérdése nem megoldott, rendkívül bonyolult probléma volna akár egy ilyen kis színpad megfelelő belátása is, nem beszélve arról, hogy a szenzorok megfelelő elhelyezése nem is biztos, hogy lehetséges volna. Ráadásul a több eszköz koordináta-rendszereinek illesztése nem támogatott, arról szintén gondoskodni kellene. [20]

További komoly probléma, hogy az alakzat-felismerés csak álló és – korlátozott mértékben – ülő emberalakokra, és a kamera felé nézve működik, színházi előadások alkalmával viszont gyakorta vesznek fel egyéb testhelyzetet is a szereplők. A szereplők azonosítása is nehézkes – ugyan a Face Tracking SDK segítségével lehetséges arcfelismerés, az érzékelési tartomány szélén megjelenő színészek azonnali azonosítása lehetetlen, ez csak a kamerához közelebb érve válna lehetővé, viszont a színre lépő színészeket késlekedés nélkül a megfelelő világításképpel kellene megvilágítani.

Összességében elmondható, hogy a rendszer egészének működése remekül demonstrálható a Kinect segítségével, ipari felhasználásra azonban nem ajánlott.

4. Színpadvilágítás

4.1. A színpadvilágítás rövid története

Már a római színház esti előadásainál is alkalmaztak fáklyákat, olajlámpákat, sőt, attól sem riadtak vissza, hogy egy drámai jelenetben a látvány kedvéért olykor-olykor a díszletet is felgyútsák. Azonban a római színház esetében még nem beszélhetünk a klasszikus értelemben vett színpadi világítástechnikáról, hiszen a fénysugarak irányításának, szabályozásának lehetősége ekkor még nem állt fent. Erre a 16. században került először sor – a fényt gyertyákkal állították elő, színezésüket különböző folyadékszűrőkkel oldották meg. A 17. századtól csillárokkal világították meg a színházakat. A nézőtérben a világítás folyamatos volt, a színpadot azonban el lehetett sötétíteni, ehhez a zsinórpádlásról leengedhető különböző szerkezeteket konstruáltak. A 19. században a gázlámpák hódítottak. Bár felforrósították a levegőt, és a gáz kellemetlen szaga is érezhető volt a levegőben, mégis előszeretettel alkalmazták, ugyanis a gáz áramlási sebességének változtatásával könnyen lehetett a fényerőt változtatni. A lámpákhoz futó gázcsövek egy helyen futottak össze, ahol szelepekkel lehetett állítani a különböző lámpákhoz eljutó gáz mennyiségét – tulajdonképpen ez volt az első „fényvezérlőpult”. Az elektromossággal begyűjthető gázlámpák újabb áttörést hoztak, hiszen mindeközéig a fényforrások bekapcsolása komoly nehézségekbe ütközött. [21]

A 19. sz. második felétől ívlámpát használnak különleges megvilágításokhoz, így például a párizsi operában a fejszereket üzemeltették ennek segítségével. Az izzószálalás elektromos színpadvilágítás elterjedése sokkal biztonságosabbá teszi a színpadot – nem véletlenül alakult ki a színpadot a nézőtérrel elválasztó tűzzáró vasfüggöny, korábban sok tüzesetet okozott a veszélyes színpadvilágítás. Kizárólag elektromos világítást először 1881-ben a Savoy Színházban alkalmaznak, bár a színház átadásakor még gázlámpás világítás működött, ugyanis elővigyázatosságból ez a rendszer is beépítésre került. Az év végétől kezdve azonban nem használták tovább a gázlámpákat.

A fényforrások és az optika tovább fejlődött. Gondot okozott a nehezen hozzáférhető helyen lévő lámpák színének módosítása, így automatikus színváltókat kezdtek használni, majd megjelentek az intelligens fényvetők, melyek a megvilágítás számos paraméterét képesek automatikusan, távoli vezérlés hatására beállítani. Ehhez különböző sötétítő egységeket, színtárcsákat, gobókat (ábrát tartalmazó maszkokat), íriszblendéket és nagyító lencsákat tartalmaznak, melyeket léptető motorok állítanak a megfelelő pozícióba.



30. ábra Fényjáték intelligens fényvetőkkel

Az intelligens fényvetőknek két főbb típusa létezik, melyek a fénycsóva iránybeállításának módjában térnek el. Először a mozgótükrös fényvetők jelentek meg. Itt a lámpatestből kilépő fény egy tükörről verődik vissza a színpadra, ennek a tükörnek a forgatásával lehet a fénynyalábot irányítani. A konstrukció előnye, hogy a kis tükröt könnyű mozgatni, így gyors beállási idő érhető el, azonban hátránya a tükör tökéletlen reflexiója miatti fényintenzitás-csökkenés, valamint a bevilágítható térszögtartomány korlátossága. A lámpák általában 150° -ot képesek átfogni vízszintesen, 110° -ot függőlegesen.

Különösképpen ez utóbbi hátrány kiküszöbölésére jöttek létre a mozgófejes fényvetők. Ebben az esetben a lámpafej egészének orientációja változik meg a talphoz képest, így függőlegesen 250° a szokásos szögtartomány, míg vízszintesen 450° , így teljesen körbe tudnak fordulni. Vannak robotlámpák, melyeknél úgy alakították ki a mechanikát, hogy vízszintesen a fenti szögtartomány sem legyen korlát, és a tengelyük körül akárhányszor körbe tudnak fordulni. További előny a mozgótükrös fényvetőkhöz képest, hogy a tükör hiánya miatt, annak reflexió tényezőjével sem kell számolni. Legnagyobb hátránya a mozgófejes kivitelnek, hogy az egész lámpafej mozgatása miatt jóval nagyobb motorokat kell alkalmazni, melyek egyrészt zajosabbak, másrészt még így is csak mintegy tized akkora sebesség érhető el a tükör mozgatásához képest. Éppen ezért mindkét konstrukció megtalálható a piacon, bár a mozgófejes változatok egyre csendesebbek és gyorsabbak, így kezdenek előtérbe kerülni. [21]



31. ábra Mozgótükrös- és mozgófejes fényvető

4.2. Színpadi vezérléstechnika

A vezérlés igénye a kezdetekkor a fényerőszabályzásban merült ki, a ma használatos rendszerek is magukon hordozzák az ebből eredő korlátokat. A lámpák teljesítményszabályzása úgynevezett dimmek segítségével lehetséges. Ezek olyan eszközök, melyek a vezérlőjelek függvényében az egyes lámpákhoz a megfelelő villamos teljesítményt képesek juttatni. Leggyakrabban triakkal fázishasítást alkalmaznak, illetve újabban az IGBT segítségével megvalósított váltóáramú szaggatók is egyre elterjedtebbek.

A dimmek vezérlése analóg illetve digitális úton is történhet, bár az analóg megoldás szinte teljesen eltűnt. Ennek oka, hogy analóg vezérlés esetén az egyes lámpacsoportokhoz tartozó információt egy feszültség szint közvetíti, így minden egyes lámpacsoport vezérléséhez külön ér szükséges, ami megnehezíti a kábelezést, ráadásul a környezeti zavarok is gondot okozhatnak. Digitális vezérlés esetén, egyetlen érpáron átvihető az összes eszköz üzemeltetéséhez szükséges információ. [21]

Mindmáig a DMX512 a legelterjedtebb szabvány a színpadi világítástechnika körében, és többnyire az új fejlesztések is ehhez a szabványhoz nyúlnak vissza. Az USITT (The United States Institute of Theatre Technology) 1986-ban fejlesztette ki a protokollt. Nevét a „Digital Multiplex” kifejezés alapján kapta, ezzel is utalva arra a korábbi megállapításra, hogy több lámpacsoport jele együtt, digitálisan kerül átvitelre. Az 512 arra utal, hogy ennyi csatornát, azaz alapesetben ennyi lámpacsoportot képes irányítani. A robotlámpák megjelenésekor meg kívánták tartani a már bevált protokollt, viszont az intelligens fényvetők esetében a fényerősségen kívül rengeteg egyéb paraméter állítható, melyekhez mind-mind külön csatornát kell rendelni. Ráadásul egy csatornán csupán 8 bit információ kerül átvitelre, ami például a korszerű robotlámpák pozícióállításának pontosságigényét nem elégíti ki, ilyen esetekben egy funkcióhoz is felhasználnak két csatornát. A fentiek figyelembevételével egy-egy ilyen lámpa akár 30 csatornát is lefoglalhat. Ebben az esetben az 512

maximális csatorna korlátot jelenthet, ekkor a vezérlőpultból több, független DMX vonalat szoktak indítani. [22]

A DMX512 az EIA/TIA-485 szabványon (sokszor RS-485-ként emlegetik) alapszik. Ez egy aszinkron, szimmetrikus kommunikációs protokoll. Egy master egység lehet, az abból kiinduló vonalra kapcsolódnak a slave eszközök, a vonal végét a reflexiók elkerülése végett le kell zárni – a leggyakoribb megoldás az, hogy a két szimmetrikus jelvezeték közé 120 ohmos ellenállást kötnek. A DMX512 az EIA/TIA-485-nél szigorúbb előírásokat állít fel a vonal hosszával kapcsolatban, csupán 1000 lábat (~300 m) engedélyez, a vevők között 300 láb távolság lehet. Összesen 32 eszköz kapcsolódhat a vonalra. [22] [23]

Az előírt kábel az RS-485 szerinti, ezen kívül a DMX512 specifikálja a csatlakozók típusát. Az ötlábú XLR csatlakozóra esett a választás, melyből csak három láb használt, a maradék kettő fenntartott későbbi felhasználás céljára – például visszafele irányuló kommunikáció valósítható meg itt (üzemidő, hőmérsékleti adatok lekérése, stb.). Mivel két láb a csatlakozón nem használt, akadnak megoldások, melyek a háromlábú XLR csatlakozót alkalmazzák, ezek azonban nem kompatibilisek a szabványos eszközökkel. [24]

Lábszám	Jel
1	Föld
2	Adat -
3	Adat +
4	Fenntartott (Adat -)
5	Fenntartott (Adat +)

6. táblázat XLR csatlakozó lábkiosztása [24]

A szabvány által előírt jelzési sebesség 250 kbaud, azaz a szimbólumidő 4 μ s. Az adatátvitel során alkalmazott protokoll meglehetősen hasonlít az RS-485 szabvány által előírthoz, azonban a DMX esetében egy multiframe struktúra is megtalálható az egyes csatornák adatainak átviteléhez. A keretszervezést egy speciális kezdőbájt indítja, majd sorban következnek az egyes csatornák (legfeljebb 512). Nincs szükség csatornánkénti címzésre, ráadásul nem kötelező elküldeni az összes csatorna konfigurációját, bármelyik után befejezhető az átvitel, ezáltal jelentősen növekszik az

adatebesség. Az egyes bájtok átvitele egy startbittel (0) kezdődik, ezután nyolc adatbit következik, melyek közül az LSB kerül először küldésre, legutoljára az MSB. Paritásbit nincs, a keretet két darab stopbit (1) zárja. A paritásbit hiánya a felhasználást tekintve nem okoz különösebb problémát, hiszen az esetek döntő többségében annyira gyorsan küldjük újra a csatornainformációkat, hogy egy esetleges hibás érték hatását észre sem lehet venni. Azonban az átvitel megbízhatatlansága miatt a szabvány kifejezetten tiltja alkalmazását olyan esetekben, ahol a hibás átvitel sérüléseket okozhat (például színpadgépészet, pirotechnika). [22]

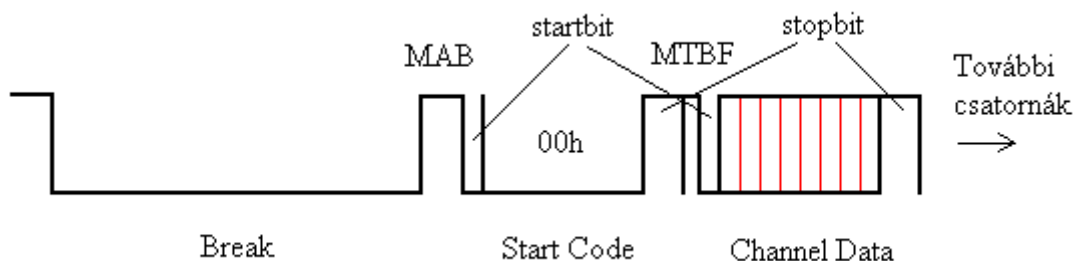
A csatorna inicializálására a Break jelzés szolgál. Ez legalább 88 μ s-ig tartó alacsony jelszintet jelent. Ezt követi a Mark After Break (MAB), mely jelzi a vevőkészülékek számára a soron következő multiframe indulását. Ez legalább 8 μ s magas jelszint. Ezután a fentebb leírt bájtátviteli protokoll szerint küldi a DMX vezérlő a multiframe kezdőbájtját, mely dimmereknek szóló adatcsomag esetén 00h. Ezt a részt (a start- és stopbitekkel együtt) a szabvány Start Code (SC) névvel illeti. Jelen pillanatban minden DMX átvitel (a robotlámpáknak szóló is) a 00h kezdőbájtot alkalmazza, ugyanis nincs más definiálva. A DMX512-ből kialakult új szabványok (pl: RDM) viszont használnak más kezdőbájtokat, ugyanis így kompatibilitás érhető el a régi eszközökkel, hiszen ezen eszközök az új üzeneteket tartalmazó csomagokat az általuk hibásnak értelmezett kezdőbájt miatt eldobják. A Start Code után Channel Data (CD) blokkok következnek, melyben az egyes csatornák bájtjai kerülnek átvitelre. A Start Code után, illetve a Channel Data blokkok között Mark Time Between Frames (MTBF) szerepelhet, mely tulajdonképpen a stopbitek meghosszabbítása, legfeljebb egy másodpercig magas szinten maradhat a vonal. Az utolsó csatorna elküldése után, azaz a csomagok között is szerepel egy hasonló elem, a Mark Time Between Packets (MTBP). Az időzítések az alábbi táblázatban foglalhatók össze. [24] [25]

	Tartalom	Méret (szimbólum)	Időtartam
Break	00..	22 – 250000	88 μ s – 1 s
Mark After Break	11(11..)	2 – 250000	8 μ s – 1 s
Start Code	0 00000000 11	11	44 μ s
Mark Time Between Frames	(11..)	0 – 250000	0 – 1 s
Channel Data	0 xxxxxxxx 11	11	44 μ s

Mark Time Between Packets	(11..)	0 – 250000	0 – 1 s
----------------------------------	--------	------------	---------

7. táblázat DMX512 keretszerkezése

Az alábbi ábra az eddigiek szemléltetésére szolgál:



32. ábra DMX512 jelalak [22]

4.3. Lámpavezérlés

A projekt során egy DTS Vinciscan intelligens fényvetőt bírtam működésre. A lámpa mozgótükrös kivitelű, tartalmaz egy szín- és egy gobótárcsát. Ennek megfelelően a manapság használatos lámpákhoz képest nagyon kevés csatorna felhasználásával programozható, két csatornát foglal le a tükrök irányítása, és egyet-egyét a két tárcsa, így csupán 4 csatorna szükséges a működéséhez. Bár a fényvetőhöz nem állt rendelkezésre gépkönyv, és kora, illetve gyártásának megszűnése miatt az interneten sem lehetett ráakadni, a felderített csatornakiosztás az alábbi:

Csatorna	Funkció
1	Vízszintes mozgás
2	Függőleges mozgás
3	Színváltás
4	Gobóváltás, stroboszkóp effekt

8. táblázat DTS Vinciscan fényvető csatornakiosztása

A lámpán beállítható az első csatorna DMX vonalbeli ofsztetje, így a fenti csatornaszámok relatív értendők. Ugyancsak állítható a vízszintes és függőleges csatorna felcserélése, illetve a pozícionáló csatornákon belüli invertálás – ekkor az irányok megfordulnak. A vízszintes csatorna (invertálás nélkül, a lámpára felülről ránézve) a matematikai pozitív irány szerinti, azaz 0-s vezérlés esetén éri el

a fénycsóva a teljes jobb-, míg 255 esetén a teljes bal kitérést. A függőleges eltérés egyre nagyobb értékei esetén a fénycsóva egyre lejjebb kúszik.

A szín- és gobócsatorna kiosztása az alábbi:

DMX érték	Szín
0-17	Fehér
18-35	Világoskék
36-53	Sötétkék
54-71	Vörös
72-89	Zöld
90-107	Bíbor
108-125	Sárga
126-143	Enciánkék
144-161	Rózsaszín
162-179	Narancs
180-197	Világoszöld
198-223	Sötétkék
224-231	Forgó tárcsa (1)
232-239	Forgó tárcsa (2)
240-247	Forgó tárcsa (3)
248-255	Forgó tárcsa (4)

9. táblázat DTS Vinciscan fényvető színcsatorna kiosztása

DMX érték	Gobó
0-9	Elsötétítés
10-19	Nincs gobó
20-29	Negyedelt négyzet
30-39	Csillag körvonal
40-49	Csillagok
50-59	Csillag
60-69	Hajlított vonalak
70-79	Négy függőleges vonal
80-89	Nyolc pötty
90-99	Vízszintes sáv
100-109	S alak
110-119	Háromszög körvonal
120-129	Körgyűrű
130-139	Virág
140-149	Kör alakban kis körök
150-159	Három kör háromszög alakban
160-169	Három kör vízszintesen
170-179	Kisméretű kör
180-189	Közepes méretű kör

190-199	Elsötétítés
200-209	Nincs gobó
210-217	Stroboszkóp (1)
218-225	Stroboszkóp (2)
226-233	Stroboszkóp (3)
234-241	Stroboszkóp (4)
242-255	Nincs gobó

10. táblázat DTS Vinciscan fényvető gobócsatorna kiosztása

A számítógéppel történő vezérléshez az FTDI USB to RS485 Serial Converter Cable termékét használtam. Ennek oka a DMX512 és RS-485 szabványok korábban már leírt hasonlósága. A kábel egyik fele egy USB csatlakozóban végződik, melyben egy lapka található, amely az USB-n kapott adatok alapján meghajtja az RS-485 vonalat. A kábel másik oldalán csatlakozó nélkül, szabadon hozzáférhető az erek. Mivel a DMX szabvány XLR csatlakozókat használ, a kábelnek erre az oldalára ráforrasztottam egy megfelelő 5 lábú csatlakozót. A kábel ereinek és a csatlakozó lábainak összerendelését az alábbi táblázatban foglalom össze: [26]

Név	Szín	XLR lábszám
GND	Fekete	1
Terminator 1	Barna	-
POWER	Vörös	-
Data +	Narancs	3
Data -	Sárga	2
Terminator 2	Zöld	-

11. táblázat Az FTDI kábel ereinek és az XLR csatlakozó lábainak összerendelése



33. ábra USB-DMX kábel

Készítettem egy csatlakozót a vonal lezárásához is, a szokásos módon, a 2. és 3. láb közé beforrasztottam egy 120 ohmos ellenállást. Azonban ezt a lezárást használva, a lámpával történő kommunikáció sikertelen volt. Egy véletlen folytán rájöttem, hogy a saját konfigurációm akkor működik megfelelően, ha a lezáró ellenállás az 1. és 3. láb között található.



34. ábra Csatlakozók

4.4. Lámpavezérlő tesztprogram

Az FTDI chippel a D2XX driver segítségével lehet kommunikálni, melyet az FTD2XX.DLL implementál. Ennek eléréséhez a linker számára hozzá kell adni az ftd2xx.lib könyvtárat, forrásainkhoz pedig csatolni kell az ftd2xx.h header fájlt. A driver függvények használnak a window.h-ban elérhető függvényeket is, éppen ezért ezt is hivatkozni kell. [27]

A lámpavezérlés tesztelésére megírt programom a DMX vonal inicializálásával indul. Először az FT_Open(0,&ftHandle) függvény meghívásával az ftHandle mutatóhoz rendeli az eszközt. A függvény

visszatérési értéke egy státuszkód, melynek 0 értéke esetén volt sikeres a művelet. Ezt követően az `FT_SetBaudRate(ftHandle,250000)` függvénnyel a szabványban előírt 250 kbaud-ra állítja az átvitel sebességét. Végül a bájtviteli paramétereket állítja be az `FT_SetDataCharacteristics(ftHandle,FT_BITS_8,FT_STOP_BITS_2,FT_PARITY_NONE)` függvény segítségével: 8 adatbit alkot egy keretet, két stopbit zárja a kereteket, nincs paritásbit. A kód végén az `FT_Close(ftHandle)` függvény szabadítja fel az erőforrást.

Két tömböt alkalmaz a program az adatok tárolására. A `TxBuffer` tömb mindig az aktuálisan lámpára kerülő adatsort tartalmazza, míg a `TxAlap` tömb referenciaértékeket tárol. Ennek a csatornák végigpásztázásánál van jelentősége. A tömb elemeinek típusa `char`, hiszen egy bájt hosszú adatokat kell a DMX vonalon kiadni. Mivel a referenciaértékek az egyszerű kezelhetőség érdekében szövegfájlba mentődnek, és a program indulásakor onnan is töltődnek vissza, több esetben is szükséges a karakter és szám típusok közötti oda-vissza kasztolás, mely bizonyos esetekben a típusok szélességének differenciája miatt konverziós problémákat vet fel, melynek leküzdésére bitmaszkolások szolgálnak. Bár a lámpa vezérlése csak négy csatornán történik, és a program funkcionalitása is ebben merül ki, a tömbök tartalmazzák a nem használt csatornákat is, ezeket a program alapesetben 0-val tölti fel, de a `dmxertekek.txt` fájlban felülírhatók, ugyanis induláskor a fájl tartalma alapján töltődnek fel a tömbök.

A tömbök inicializálása után egy egyszerű kezelőfelület generálódik, melyen egy karakter hosszú parancskódokkal lehet navigálni. A 'k' leütésével kilép a program. A 'd' elindít egy bemutatót, mely végigpásztazza csatornákat, így a lámpa összes paraméterére fény derül. Az 'e', 'a', 's', és 'g' gombok hatására az eleváció, azimut, szín és gobó csatornákat lehet beállítani, majd a lámpa állapotát egy DMX üzenet azonnal frissíti, a `dmxertekek.txt` fájl pedig felülíródik a `mentKiir` függvény meghívása által. Előbbi két paraméter értékei relatív állíthatók az aktuális értékhez képest, a szín és gobó beállításoknál ellenben abszolút csatornaértéket kell megadni. Az 'r' parancs az aktuális referenciaértékeket menti fájlba és juttatja a DMX vonalra.

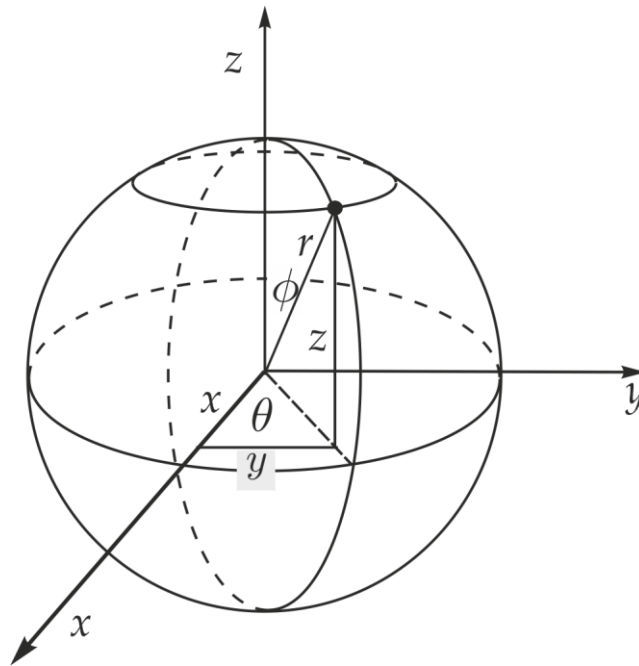
5. Követés

Az eddigi fejezetekben láthattuk a rendszer két szélső komponensét: a Kinect segítségével megtörténik az emberalakok helyzetének detektálása, illetve a DMX vonalon lehetséges a lámpák vezérlése. A követéshez azonban szükség van még a két komponens összekötésére.

5.1. Koordinátarendszerek

A Kinect Descartes-féle koordinátarendszerben ábrázolja az objektumokat. Ennek a koordinátarendszernek az alapja három ortogonális tengely, melyek – általában jobbsodrású rendszert alkotva – rendre x , y és z . Egy pont helyzete három koordinátával írható le: x a pont előjeles távolsága az y és z tengelyek által kifeszített síktól, ugyanígy y és z koordináták is a másik két tengely által kifeszített síktól mérhető távolságokat jelölik.

A lámpák vezérléséhez szükséges információk leírására a gömbi koordinátarendszer a legalkalmasabb. A gömbi koordinátarendszerben egy pontot három paraméter határoz meg: r , θ , ϕ . Az r megadja az origótól mért távolságot, azaz r a gömb sugara. θ a pont x - y síkú vetülete és az origót összekötő egyenes hajlásszöge az x tengelytől, azimut szögnek nevezik. Végül a ϕ a pontot az origóval összekötő egyenes hajlásszöge a z tengelytől, ez az elevációs szög. [28]



35. ábra Gömbi koordinátarendszer [28]

A földrajzi szélességi és hosszúsági fokok is ebben a rendszerben értendők, amennyiben z koordináta az Északi-sark felé mutat – annyi különbséggel, hogy a szélességi fok az elevációs értékből a $90^\circ - \phi$ összefüggésből számítható, így értéke éppen az Egyenlítőnél 0° , a sarkoknál pedig $\pm 90^\circ$. A hosszúsági fokok referenciája a Greenwichen átfutó főkör.

Amennyiben egy gömbi és egy Descartes-féle koordinátarendszer a fenti ábrán látható módon kapcsolódik, az alábbi módon térhetünk át gömbi koordinátából közönséges koordinátákra:

$$\begin{aligned}x &= r \sin \phi \cos \theta \\y &= r \sin \phi \sin \theta \\z &= r \cos \phi\end{aligned}$$

A megvalósítandó feladathoz azonban éppen a Descartes-féle koordinátákból kell gömbi koordinátákra áttérni, ekkor az alábbi összefüggések alkalmazhatók:

$$\begin{aligned}r &= \sqrt{x^2 + y^2 + z^2} \\ \theta &= \arctan\left(\frac{y}{x}\right) \\ \phi &= \arccos\left(\frac{z}{r}\right)\end{aligned}$$

5.2. Transzformációs modell

A lámpavezérlés során, a detektált objektum Kinect által szolgáltatott három Descartes-féle koordinátájából meg kell határozni a lámpa függőleges és vízszintes elfordulásának vezérlőinformációit. A Kinect koordinátarendszerében y koordináta tartozik a függőleges irányhoz – a szokásos jelölésmód használhatósága, és az átláthatóság érdekében átneveztem a koordinátákat, hogy a z tengely jelölje a függőleges irányt. Mivel a fényvető nem a Kinect által használt koordinátarendszer origójában van, első lépésként x , y és z koordinátákhoz rendre Δx , Δy és Δz eltolásokat adva el kell érni, hogy a kapott x' , y' és z' már egy olyan Descartes-féle koordinátarendszerben határozzák meg az objektum pozícióját, melynek origójában a lámpa található. A fentiek alapján $(-\Delta x; -\Delta y; -\Delta z)$ a lámpa pozíciója a Kinecthez rögzített koordinátarendszerben.

Ezt követően történik az előző fejezetben már látott gömbi koordinátákra áttérés. A kapott r érték jelen projektben nem kerül felhasználásra, ugyanis a teszteléshez használt lámpa nem képes a kibocsájtott fénynyaláb kúpszögét állítani. Amennyiben ez lehetséges volna, lehetővé válna a színészt körülvevő fénykör állandó nagyságon tartása, a kúpszög r függvényében történő állítása segítségével. Ha a lámpa ténylegesen azimut és elevációs szögeket dolgozna fel, a gömbi koordinátákra történő áttérés során kapott θ és ϕ közvetlenül képezhetné a lámpavezérlő információt. Azonban a helyzet ennél bonyolultabb: mind a vízszintes, mind a függőleges elmozdulás vezérlésére (a, e) a lámpa egy 8 bites értéket vár, azaz egy 0-tól 255-ig terjedő számot. Mivel a két tengely irányába az elfordulás maximális nagysága nem egyezik meg, ráadásul a nullapontok helyzete is önkényes, mindkét irányba bevezetendő egy skálafaktor és egy ofszet (f, g) , mely a kivezérelt szögtartományt leképzi a $[0; 255]$ tartományra.

A fentiek alapján a transzformációs modell az alábbi összefüggésekben foglalható össze:

$$x' = x + \Delta x \quad y' = y + \Delta y \quad z' = z + \Delta z$$

$$r = \sqrt{x'^2 + y'^2 + z'^2} \quad \theta = \arctan\left(\frac{y'}{x'}\right) \quad \Phi = \arccos\left(\frac{z'}{r}\right)$$

$$a = f_1 \cdot \theta + g_1 \quad e = f_2 \cdot \Phi + g_2$$

5.3. Követő program

A követés demonstrálására írt programom a már korábban bemutatott függvényeket, megoldásokat alkalmazza az adatok Kinect felől történő eléréséhez, és a lámpavezérléshez. Inicializálja a Kinectet és a DMX512 vonalmeghajtást, majd ciklikusan gyűjteni kezdi az adatokat a szenzor felől, az ezekből feldolgozott vezérlőjeleket pedig minden ciklusban a kimenetre küldi. Amennyiben érvénytelen adat érkezik a bemenetről, azzal a feltételezéssel élünk, hogy éppen egyáltalán nem tartózkodik követendő objektum a megfigyelt területen, így ebben az esetben a kimenet gobócsatornájának 0 értékre állítása elsötétíti a lámpát. Ellenkező esetben ennek a csatornának 10-es értékre állítása teljes megvilágítást eredményez, az azimut és elevációs csatornák pedig az első megtalált emberalak fejének pozíciója alapján értékelődnek ki, a transzformációs modell szerint. A hét modellparamétert (Δx , Δy , Δz , f_1 , g_1 , f_2 , g_2) fájlból olvassa be a program indulásakor.

A követőrendszer működésének egyetlen komoly problémája a sebessége. A tesztelés során nyilvánvalóvá vált, hogy gyors mozgások esetén a fénycsóva valamelyest lemarad a mozgó testhez képest. Ennek oka valószínűleg a Kinectben keresendő, ugyanis a többi rendszerkomponens (transzformációs modell számításai, DMX vezérlés) sebessége lényegesen nagyobb.

6. Kalibráció

6.1. A kalibráció elve

Jogosan merül fel a kérdés: a szükséges modellparaméterek honnan állnak rendelkezésre? A korai tesztelések során a Kinectet a lámpára helyeztem, így az első három paraméterből kettőt nullának feltételeztem, a harmadikat pedig (a függőleges eltérést) lemértem, mely 0,1 m-nek adódott. Az elfordulások átskálázását szabályozó további négy paramétert empirikus módszerrel, döntően próbálgatással határoztam meg. Ez a módszer számomra is körülményes volt, valós körülmények között kivitelezhetetlen lenne. Biztosítani kell egy eljárást, mellyel a szükséges modellparaméterek egyszerűbben meghatározhatók – egyéb mérőeszközök igénybevétele, illetve próbálgatás nélkül.

A kalibráció a következőképpen történik: a színész egymás után néhány tetszőlegesen kiválasztott kalibrációs helyre áll, eközben a világosító technikus minden egyes esetben a lámpát manuálisan vezérelve a színész fejére állítja a fénykört, és az összes kalibrációs pont adatait rögzíti. A kalibrációs pontok rögzítése során olyan adathalmaz keletkezik, melyben a Descartes-féle koordináta-rendszerben meghatározott koordináták mellett ugyanezen pontokhoz tartozó lámpavezérlő információk is ismertek. A cél ezek alapján meghatározni a transzformációs modellparamétereket.

A transzformációs modell tárgyalásánál felsorolt 8 egyenlet mindegyike mátrixegyenletként fogható fel, hiszen egyszerre több kalibrációs pont transzformációs összefüggéseit vizsgáljuk. Így tulajdonképpen egy $8 \cdot k$ egyenletből álló nemlineáris egyenletrendszer adódik, ahol k a felvett kalibrációs pontok száma. Az egyenletrendszerben $5 \cdot k$ ismert mennyiség (x, y, z, a, e), a meghatározandó 7 modellparaméter ($\Delta x, \Delta y, \Delta z, f_1, g_1, f_2, g_2$) és további $6 \cdot k$ ismeretlen ($x', y', z', r, \theta, \phi$) található. Így az egyenletek és ismeretlenek számára felírható a $8k > 7 + 6k$ egyenlőtlenség, ennek megoldása pedig azt jelzi, hogy 4 illetve több kalibráló pont esetén – ha a kapott egyenletek függetlenek – megoldható az egyenletrendszer.

6.2. Kalibráció a gyakorlatban

A probléma megoldására a Matlab `fsolve` függvényét alkalmaztam. A modellben használt utolsó két egyenletet nullára rendeztem, majd implementáltam a modell-egyenleteket egy függvényben, melynek bemenő paraméterei a modellparaméterek és a kalibrációs pontok adatai, visszatérési értékei pedig a nullára rendezett egyenletek nem nulla oldala által meghatározott érték. A nullával való egyenlőség ugyanis csak akkor állna fenn, ha a függvény paramétereként helyes modellparamétereket kapna. Az `fsolve`-val történő megoldáskeresésnek pontosan ez az alapja: a bemenő modellparamétereket addig iterálja a metódus, míg a visszatérési érték nulla (illetve egy bizonyos tűréshatáron belül) lesz.

A valóságos környezetben történő tesztelés előtt egy Matlabban elkészített modellt használva próbáltam ki az eljárást, amiben – a várakozásoknak megfelelően – 4 vagy több kalibrációs pont esetén adódott a helyes transzformációs modell. A kalibrációs adatokat zajjal terhelve is sikerült a kalibrálás.

A kalibrációs eljárás valóságban történő kipróbálásához szükség volt egy programra, mellyel felvehetők a kalibrációs adatok. A program a Kinect- és a lámpavezérlő tesztprogramban megismert módszerekkel lehetővé teszi a lámpa manuális ráirányítását a kalibrációs pontokra, majd rögzíti a lámpavezérlő információkat, illetve a Kinect által aktuálisan mért értékeket egy Matlab által könnyen

értelmezhető fájlba. A kalibráció pontosságát elősegítendő a lámpa gobótárcsája teljes fénykör vetítése helyett az eljárás során csak egy kis fénynyalábot enged ki.

A matematikai modellel ellentétben valóságos kalibrációs eljárás során 4 pont felvétele nem bizonyult elegendőnek, melynek egyrészt a felvett pontok által meghatározott egyenletek nem kellő függetlensége, másrészt a kalibráció során fellépő hiba (Kinect pontatlansága, kalibráló technikus hibája, lámpavezérlő információk 8 bites ábrázolásából fakadó kis pontossága) lehet az oka. 5-6 pont felvétele esetén azonban az adódó transzformációs paramétereket kipróbálva a követés általában megfelelő volt, mindazonáltal a Matlab fsolve függvénye ebben az esetben is „No solution found” üzenettel tért vissza, ugyanis az iteráció során a függvény visszatérési értéke nem közelítette meg kellően a nullát. Valószínűleg ezt a jelenséget is az eljárás során fellépő különböző hibák okozzák. Mindenesetre az iteráció során kapott eredmény minden esetben ellenőrizendő a követő programmal, hogy a megfelelő működést produkálja-e.

7. Összegzés

A projekt megvalósítása során először a pozíció meghatározó rendszerkomponenst vettem szemügyre. A színészek követésére a Microsoft Kinect érzékelőjét használtam, így áttekintettem ennek képességeit. Majd egy tesztprogramot írtam a Kinect lehetőségeit feltérképezendő – természetesen a program elsősorban a projekthez kapcsolódó csontvázkövetés képességeit vizsgálja. Végül ennek segítségével egy méréssorozatot folytattam le, mely segítségével meggyőződtem az eszköz megfelelő pontosságáról.

A pozíció meghatározó rendszerkomponens taglalása után a rendszer másik oldala, a lámpavezérlés következett. Röviden áttekintettem a színpadi világítástechnikát, illetve a vezérléshez használt DMX512 szabványt. Az FTDI USB to RS485 Serial Converter Cable megfelelő preparálásának segítségével létrehoztam egy csatolót, mellyel lehetőségessé vált számítógép segítségével irányítani a lámpákat, ennek tesztelésére egy újabb programot írtam, mely képes a teszteléshez rendelkezésre álló DTS Vinciscan típusú intelligens fényvető minden funkcióját vezérelni.

Végül a két oldal közötti kapcsolatot kellett megteremteni, azaz a Kinect által detektált Descartes-féle koordináta-rendszerbeli pontokhoz lámpavezérlő információt rendelni. Ehhez először létrehoztam egy transzformációs modellt, majd ezt implementálva megalkottam a követő programot, mely biztosítja a projekt célját, azaz a színészek követését.

A transzformációs modell alkalmazásához azonban 7 modellparaméter szükséges, melyek leírják egyrészt a rendszer geometriáját (a Kinect és a lámpa kapcsolatát), másrészt a lámpa pozícionáló

rendszerének tulajdonságait. A paraméterek meghatározásához megalkottam egy kalibrációs rendszert, mely néhány pontra történő rávilágítás során rögzített pozíció- és lámpavezérlő információ segítségével kiszámolja a transzformációs paramétereket.

8. Továbbfejlesztési lehetőségek

A kalibrációhoz és a rendszer teljes körű kezeléséhez szükséges egy könnyen kezelhető felhasználói felület, nem várható el a felhasználóktól, hogy a példaprogramok során használt konzolos felületről vezéreljék a rendszer működését, ráadásul a kalibráció során a Matlabot is alkalmazzák.

Felkészíthető a rendszer több lámpa vezérlésére is – a DMX512 szabvány segítségével nagyon egyszerűen adhatók a rendszerhez további fényvetők, a programban több egyidejű transzformációs modell kezelését kell implementálni. Nem csak a lámpák, hanem a szereplők számának növelése is elképzelhető, a Kinect által aktívan követett mindkét emberalakot meg lehet világítani.

Kiküszöbölendő, vagy legalábbis csökkentendő a rendszer késleltetése. Elképzelhető, hogy ennek egyetlen módja a Kinect lecserélése, de hosszú távú fejlesztési irányként ez amúgy is ildomos, a 3.6. fejezetben felsorolt hátrányai miatt.

9. Forrásjegyzék

- [1] Stephanie Crawford: How Microsoft Kinect Works
<http://electronics.howstuffworks.com/microsoft-kinect.htm>
- [2] Kinect for Windows
<http://www.microsoft.com/en-us/kinectforwindows/>
- [3] Nintendo Wii
<http://www.nintendo.com/wii>
- [4] Wii Accidents Montage
<http://www.youtube.com/watch?v=HAHlqX6ISCs>
- [5] <http://www.gameshop.hu/img/wii-remote-controller.jpg>
- [6] Kinect megateszt
http://www.gamekapocs.hu/cikk/1495/kinect_megateszt
- [7] Barna József: A Kinecttel újra izgalmassá válik a PC
http://itcafe.hu/hir/microsoft_kinect_for_windows_pc_sdk.html
- [8] Kinect érzékelő kézikönyv
- [9] Darabjaira szedték a Kinect szenzort
http://www.gamekapocs.hu/hir/20889/darabjaira_szedtek_a_kinect_senzort
- [10] Kinect for Windows Software Development Kit Documentation
<http://www.microsoft.com/en-us/kinectforwindows/develop/resources.aspx>
- [11] Seller Rudolf: Antennarendszerek alapjai
<http://hvt.bme.hu/~seller/Radar/Antennarendszer.DOC>
- [12] Seller Rudolf: Antenna topológiák
<http://hvt.bme.hu/~seller/Adaptív/Jegyzet/5fejezet.doc>
- [13] Encyclopedia Britannica – Human Eye – Binocular Vision
<http://www.britannica.com/EBchecked/topic/1688997/human-eye/64969/Binocular-vision?anchor=toc64969>
- [14] Csodák Palotája – Ames-szoba
<http://www.csodakpalotaja.hu/index.php/hu/ames-szoba>
- [15] Alan Johnston: Depth Perception
<http://www.psychol.ucl.ac.uk/alan.johnston/Depth.html>
- [16] Hányingerkeltők és fárasztóak a 3D-s filmek
http://szamitastechnika.network.hu/blog/szamitastechnika_klub_hirei/hanyingerkeltok-es-farasztoak-a-3-d-s-filmek

- [17] How Kinect depth sensor works – stereo triangulation?
<http://mirror2image.wordpress.com/2010/11/30/how-kinect-works-stereo-triangulation/>
- [18] Candide – a parameterized face
<http://www.icg.isy.liu.se/candide/>
- [19] Nathan Crock: Kinect Depth vs. Actual Distance
<http://mathnathan.com/2011/02/03/depthvsdistance/>
- [20] Centrál Színház – Műszaki adatok
<http://www.centralszinhaz.hu/muszaki-adatok>
- [21] Pelyhe János: Világítástechnikai jegyzet
http://www.pelyhe.hu/index.php?option=com_phocadownload&view=category&download=86:08-2-szinpadvilagitas-es-vezerlestechika
- [22] Ujjal Kar: DMX512
<http://www.dmx512-online.com/>
- [23] Guidelines for Proper Wiring of an RS-485 (TIA/EIA-485-A) Network
<http://www.maxim-ic.com/app-notes/index.mvp/id/763>
- [24] Dmx 512 Handbook
<http://www.elationlighting.com/pdffiles/dmx-101-handbook.pdf>
- [25] RDM Protocol
<http://www.rdmprotocol.org>
- [26] USB to RS485 Serial Converter Cable – Datasheet
http://www.ftdichip.com/Support/Documents/DataSheets/Cables/DS_USB_RS485_CABLES.pdf
- [27] Software Application Development – D2XX Programmer's Guide
[http://www.ftdichip.com/Support/Documents/ProgramGuides/D2XX_Programmer's_Guide\(FT_000071\).pdf](http://www.ftdichip.com/Support/Documents/ProgramGuides/D2XX_Programmer's_Guide(FT_000071).pdf)
- [28] Dr. Kovács Emőd: Komputergrafika – Matematikai alapok
Gömbi koordináta-rendszer
<http://aries.ektf.hu/~hz/pdf-tamop/pdf-01/html/ch03.html#id672217>

10. Ábrajegyzék

1. ábra Rendszerkomponensek	3
2. ábra Kinect [2]	4
3. ábra Wii Remote [5]	4
4. ábra A Kinect belseje [9].....	6
5. ábra Érzékelők és a motor elhelyezkedése [10]	7
6. ábra Működési tartományok [10]	7
7. ábra θ szögéből érkező síkhullám [11]	8
8. ábra Antennarendszer [11]	9
9. ábra A Csodák Palotájában található Ames-szoba [14].....	11
10. ábra Retinális diszparitás [13]	11
11. ábra A binokuláris látás segítségével térhatású kocka.....	12
12. ábra SDK architektúra [10]	13
13. ábra A csontvázkövetés által támogatott testrészek [10].....	15
14. ábra Aktív és passzív követés [10]	15
15. ábra Álló és ülő mód [10]	16
16. ábra A Kinect koordinátarendszere	16
17. ábra A testrészek hierarchikus szervezése [10].....	17
18. ábra Arckövetés által meghatározott pontok [10]	18
19. ábra A fej különböző mozgásai [10]	19
20. ábra MÉRŐBÁBU.....	21
21. ábra A mérőbábu feje.....	22
22. ábra MÉRÉSI ELRENDEZÉS	22
23. ábra Mélységi adatok	25
24. ábra Mélységmérés hibái	27
25. ábra Mélységmérés relatív hibái	28
26. ábra Nathan Crock mérési elrendezése [19]	28
27. ábra Nathan Crock mérési eredménye [19]	29
28. ábra X koordináta mérési eredményei.....	30
29. ábra X koordináta hibái	31
30. ábra Fényjáték intelligens fényvetőkkel.....	34
31. ábra Mozgótükrös- és mozgófejes fényvető	35
32. ábra DMX512 jelalak [22]	38
33. ábra USB-DMX kábel	42
34. ábra Csatlakozók.....	42
35. ábra Gömbi koordinátarendszer [28]	44

11. Táblázatok jegyzéke

1. táblázat A fej különböző mozgásai	18
2. táblázat Szórások (csípő közepe).....	23
3. táblázat Szórások (fej)	24
4. táblázat Mélységmérés illesztett egyenseinek paraméterei.....	26
5. táblázat X koordináta mérése során illesztett egyenesek meredeksége	31
6. táblázat XLR csatlakozó lábkiosztása [24]	36
7. táblázat DMX512 keretszerkezése	38
8. táblázat DTS Vinciscan fényvető csatornakiosztása	38
9. táblázat DTS Vinciscan fényvető színcsatorna kiosztása	39
10. táblázat DTS Vinciscan fényvető gobócsatorna kiosztása	41
11. táblázat Az FTDI kábel ereinek és az XLR csatlakozó lábainak összerendelése.....	41