

IDENTIFIKIMI I DORËS NËPËRMJET TEKNOLOGJISË MICROSOFT KINECT

*GANI E., KIKA A.

Universiteti i Tiranës, Fakulteti i Shkencave të Natyrës, Departamenti i Informatikës

e-mail: eriglen.gani@fshn.edu.al

Përmbledhje

Gjuha e shenjave është shumë e rëndësishme në përfshirjen e personave që nuk dëgjojnë në shoqëri. Komunikimi ndërmjet tyre dhe personave që dëgjojnë realizohet me ndihmën e interpretuesve të cilat rezultojnë në metoda jo efektive. Përdorimi i teknologjisë për lokalizimin dhe përkthimin e shenjave në mënyrë natyrale do e lehtësonte procesin. Në komunikimin me anë të gjuhës së shenjave, lëvizja e duarve luan një rol shumë të rëndësishëm. Qëllimi i këtij punimi është identifikimi i dorës nëpërmjet teknologjisë Microsoft Kinect. Lokalizimi i dorës në mënyrë natyrale do realizohet nëpërmjet dy mënyrave. Së pari nëpërmjet llogaritjes së statistikave mbi të dhënat e thellësisë së imazheve. Kjo teknikë ndihmon në llogaritjen e vlerave limit, që lejojnë të përcaktohet nëse një piksel është zhurmë apo pjesë e dorës së përdoruesit. Së dyti nëpërmjet lokalizimit fillimisht të skeletit të përdoruesit dhe më pas identifikimit të dorës, shtresat e së cilës ndodhen më afër pajisjes Kinect.

Abstract

Sign Language is very important for the inclusion of the hearing impaired persons in the society. Communication between them and normal people is done by using interpreters which results in non effective methods. Technology would facilitate the process of identification and translation of signs. Hand trajectory plays an important role in the sign language communication process. The purpose of this paper consists of hand identification using Microsoft Kinect technology. Hand identification in a natural way will be done in two ways. In the first technique, statistics can be used while processing depth image data. They allow to determine threshold values that help decide if a pixel is a noise or part of user's hands. In the second technique, firstly the user's skeleton is identified. Then user's hand is identified, the layers of which are situated closer to the Kinect device.

Fjalëkyçe: Ndërfaqe natyrale, microsoft kinect, gjuhë e shenjave, identifikim i dorës.

Hyrje

Personat të cilët kanë probleme me dëgjimin përdorin gjuhën e shenjave si gjuhë primare në komunikim. Për integrimin e personave që nuk dëgjojnë në shoqëri nevojitet një komunikimi ndërmjet tyre dhe personave normalë. Në ditët e sotme ky lloj komunikimi realizohet me ndihmën e interpretuesëve të cilët janë të pakët në numër dhe procesi rezulton jo efektiv (Zico, 2014). Komunikimi me anë të gjuhës së shenjave mundësohet ndërmjet lëvizjes së trupit, lëvizjes së kokës dhe trajektorës së përshkruar nga duart (Sign Language, 2013). Lëvizja e duarve përbën pjesën më të rëndësishme në një

komunikim midis personave që nuk dëgjojnë. Lokalizimi, analizimi dhe përkthimi i shenjave të realizuara me anë të duarve në mënyrë natyrale përbën një nga sfidat në ditët e sotme. Për identifikim e duarve në mënyrë natyrale mund të përdoren teknologji të ndryshme ku secila prej tyre ka përparësi dhe të meta. Përdorimi i kamerave web mundëson kapjen e duarve duke u përqëndruar mbi ngjyrën dhe formën (Cerezo, 2012). Për shkak të diversitetit të ngjyrave dhe formave të duarve të personave nevojitet një trajnim i algoritmit për çdo përdorues. Përdorimi i dorezave speciale të pajisuar me sensorë, lehtëson kapjen e shenjave (Kyatanavar & Futane, 2012), por qëndron si një metodë e kushtueshme e cila mund të përdoret vetëm në projekte të mëdha dhe jo në jetën e përditëshme. Fakti që komunikimi me anë të gjuhës së shenjave nuk realizohet vetëm me anë të lëvizjes së duarve por edhe të kokës dhe trupit bën që përdorimi i dorezave të jetë një metodë jo gjithë-përfshirëse. Teknologjia Microsoft Kinect nëpërmjet përdorimit të kamerave të thellësisë 3D lehtëson kapjen e shenjave dhe nëpërmjet përdorimit të paketave software-ike lejon analizimin e tyre. Teknologjia Microsoft Kinect, për shkak të kostos të ulët, mund të përdoret edhe në komunikimin e përditshëm (Wenjun, 2012).

Përdorimi i kamerave 3D të teknologjisë Microsoft Kinect lejon identifikimit e dorës si pjesë e trupit më afër pajisjes Kinect, pas identifikimit të skeletit të përdoruesit. Në këtë mënyrë shmangen problemet e humbjes së imazhit “dorë” si pasojë e ndërhyrjeve të sendeve midis pajisjes Kinect dhe dorës së përdoruesit. Microsoft Kinect i bashkëngjijt çdo pikseli të imazhit të gjeneruar, informacion mbi thellësinë duke lejuar në këtë mënyrë përdorimin e statistikave mbi të dhënat e thellësisë së imazheve, për llogaritjen e vlerave limit që ndihmojnë të përcaktohet nëse një piksel është zhurmë apo pjesë e dorës së përdoruesit (James & Jarrett, 2012).

Mjedisi i punës

Zhvillimet e fundit në teknologjitë e kamerave të thellësisë 3D si Microsoft Kinect kanë krijuar mundësi të reja në avancimin e programeve multimedialë. Microsoft Kinect ka ndryshuar mënyrën se si përdoruesit tentojnë të luajnë. Përdorimi i sensorëve mundëson kapjen e lëvizjeve të trupit, kokës, duarve dhe përkthimin e tyre në komanda të interpretueshme për kompjuterin. Kinect mundëson një komunikim natyral ndërmjet përdoruesit dhe kompjuterit (Wenjun, 2012). Impakti i Kinect-it nuk ka qëndruar vetëm në industrinë e lojrave, por ka kaluar kufijt e tyre duke gjetur përdorim edhe në fushën akademike. Kostoja e ulët ka lejuar integrimin e teknologjisë në fusha si robotika, shkencat kompjuterike dhe ka mundësuar mënyra të reja në komunikimin kompjuter-njeri. (Wenjun, 2012). Në shtresën më të ulët ndodhet hardware-i cili konsiston në sensorin e ngjyrave RGB, përgjegjës për gjenerimin e imazheve me ngjyra në një rezolucion 1280x960, sensorin e emetimit infrared dhe sensorin e thellësisë përgjegjës për llogaritjen e distancës së objekteve nga pajisja Kinect, motorin e zhvendosjes dhe sensorin audio. (Windows kinect components, 2015).

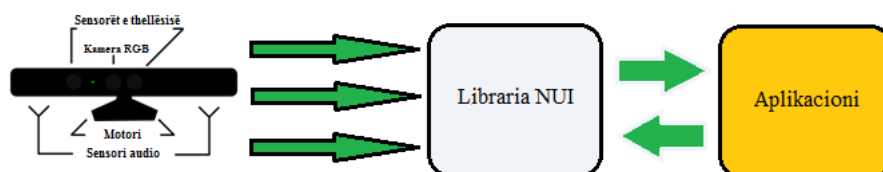


Figura 1. Modulet kryesore të sistemit Microsoft Kinect

Për të forcuar edhe më shumë efektin e Kinect është e nevojshme një shtresë e ndërmjetme e cila lehtëson kapjen dhe përpunimin e të dhënave që mundësohen nga sensorët hardware-ik. Ekzistojnë disa produkte me kod të hapur që mund të përdoren në fushën e bashkëveprimit njeri-kompjuter të cilat përfshijnë OpenNI, OpenKinect dhe Microsoft SDK (Zico, 2014). Aplikacionet e përdorura në këtë punim janë dizenuar të funksionojë mbi framework-un Microsoft SDK 1.8.0.0.

Identifikimi i dorës nëpërmjet përdorimit të statistikave të thellësisë

Paketa software-ike SDK e ofruar nga Microsoft lejon analizimin e të dhënave të thellësisë. Çdo pikseli, pjesë e imazheve të gjeneruar nga pajisja Kinect, i bashkëngjitet informacion mbi nivelin e thellësisë sipas formatit të mëposhtëm.

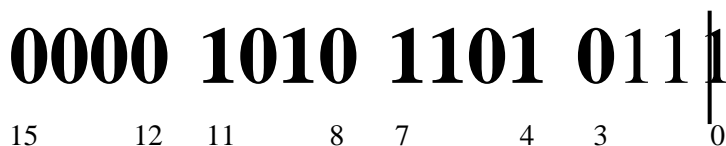


Figura 2. Bitet e informacionit mbi thellësinë

Trembëdhjetë bitet e para mbartin informacion mbi thellësinë në një interval të shtrirë nga [0-4095] (James & Jarrett, 2012). Për këtë punim kemi realizuar disa eksperimente të cilat paraqesin shpërndarjen statistikore të të dhënave, ndërmjet tabelave dhe grafikëve dy-dimensionalë. Fokusi i eksperimenteve është shpërndarja e të dhënave të thellësisë. Mbi çdo frame të gjeneruar nga pajisja Kinect llogarisim në kohë reale numrin e pikseleve për çdo shtresë. Të dhënat janë paraqitur në formë tabelare dhe grafike si më poshtë:

Intervali i shtresave të thellësisë	Denduria e pikave
[0-992]	0
[993-1167]	4006
[1168-1246]	0
[1247-4094]	70592

Figura 3. Shpërndarja e dendurisë së pikave sipas shtresave të thellësisë

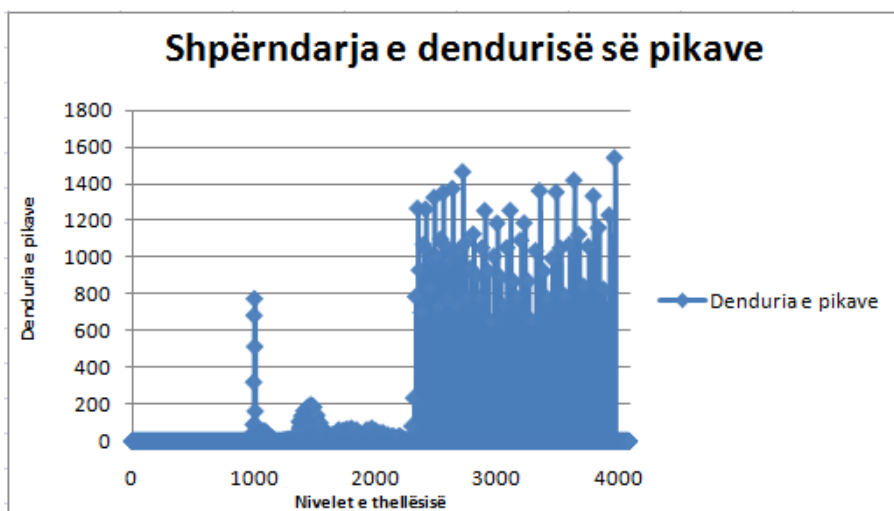


Figura 4. Shpërndarja e dendurisë së pikave sipas shtresave të thellësisë

Nga interpretimi i vlerave është e dukshme që shtresat e ndodhura në intervalin [993-1167] përfaqësojnë objektin e parë të ndodhur përpara pajisjes Kinect, ku në rastin e eksperimentit është dora e përdoruesit. Shtresat e tjera përfaqësojnë pikselet e ndodhur pas dorës së përdoruesit. Lokalizimi i shtresës së parë nga ku fillon denduria e pikave, dhe shtimi i një konstante prej 100 shtresash mundëson kapjen e dorës së përdoruesit. Për këtë qëllim është realizuar një aplikacion, i bazuar në teknologjinë .Net, i cili në kohë reale mundëson shfaqjen e dorës së përdoruesit.



Figura 5. Imazhet e kapura ndërmjet përdorimit të statistikave

Në eksperimentin e dytë kemi shtimin e një objekti midis përdoruesit dhe pajisjes Kinect. Është lehtësisht e dukshme pamundësia për lokalizimin e dorës duke marrë parasysh që shtimi i objekteve mund të jetë një numër i panjohur.

Intervali i shtresave të thellësisë	Denduria e pikave
[0-865]	0
[866-1090]	24121
[1091-1249]	626
[1247-4094]	67288

Figura 6. Shpërndarja e dendurisë së pikave sipas shtresave të thellësisë

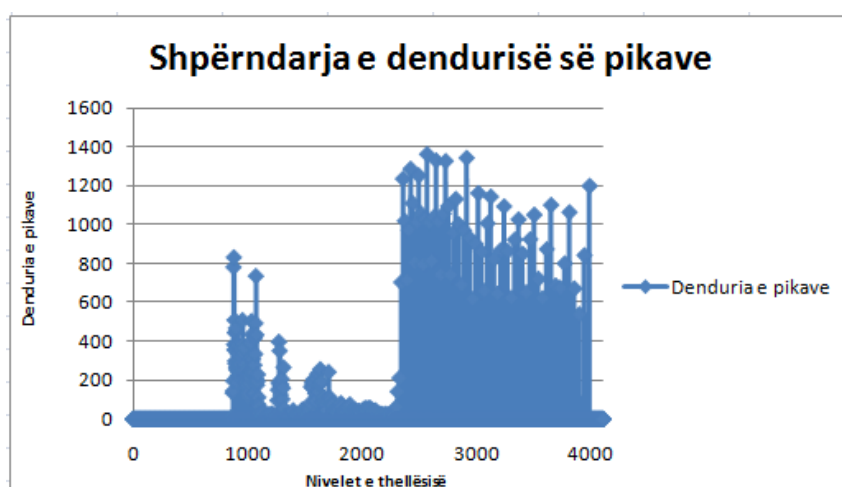


Figura 7. Shpërndarja e dendurisë së pikave sipas shtresave të thellësisë



Figura 8. Shpërndarja e dendurisë së pikave sipas shtresave të thellësisë

Përdorimi i statistikave mbi llogaritjen e limiteve të thellësisë duke u bazuar në dendurinë e pikave merr kuptim në rastet kur përpara pajisjes Kinect, nuk vendosim asnjë objekt tjetër përveç përdoruesit.

Identifikimi i dorës ndërmjet identifikimit të skeletit të përdoruesit

Identifikimi i dorës ndërmjet përdorimit të statistikave të thellësisë nuk jep rezultate të sakta në momentin kur një objekt i tretë depërton midis përdoruesit dhe pajisjes Kinect. Kjo për arsye të vendosjes së një shtrese thellësie përpara dorës së përdoruesit. Për rritjen e saktësisë së rezultateve fokusohemi vetëm në shtresat që i përkasin përdoruesit (James & Jarrett, 2012). Pajisja Kinect, përveç bashkëngjites së informacionit mbi nivelin e thellësisë, çdo pikseli i bashkëngjitet edhe informacion mbi indeksin e përdoruesit.

0000	1010	1101	0111				
15	12	11	8	7	4	3	0

Figura 8. Bitet e informacionit mbi indeksin e përdoruesit

Tre bitet e fundit mbartin informacion mbi indeksin e përdoruesit në një interval të shtrirë nga [0-7]. Në të njëjtën kohë mund të identifikohet më shumë se një përdorues i ndodhur përpara pajisjes Kinect (James & Jarrett, 2012). Në eksperimentin tonë jemi fokusuar në një përdorues të vetëm. Interesohemi për indekse të ndryshëm nga zero. Çdo shtresë me një indeks përdoruesi të barabartë me zero nënkupton një shtresë jashtë trupit fizik të përdoruesit. Për kapjen e skeletit të përdoruesit është e nevojshme aktivizimi i sensorit për kapjen e skeletit së bashku me sensorin e thellësisë.

Më poshtë janë kapur shtresat që i përkasin përdoruesit duke shkëputur në këtë mënyrë çdo shtresë jashtë trupit të tij. Për këtë qëllim është realizuar një aplikacion, i bazuar në teknologjinë .Net. Aplikacioni ka të integruar paketën software-ike Microsoft Kinect SDK v1.8 dhe për të mundësuar kapjen e skeletit të përdoruesit duhet të aktivizohen sensorët e skeletit dhe sensorët e thellësisë. Aktivizimi i sensorëve të skeletit ju jep vlerë tre biteve të fundit që përfaqësojnë indeksin e përdoruesit. Nëpërmjet kodit të aplikacionit, në kohë reale shfaqen vetëm pikselet, për të cilat, bitet e indeksit të përdoruesit përmbajnë një vlerë të ndryshme nga zero.



Figura 9. Identifikimi i skeletit të përdoruesit

Shpërndarja e dendurisë së pikave tregon intervalin e shtresave [1022-1585] që përmban shtresat e përdoruesit dhe çdo shtresë jashtë këtij intervali, nuk përmban asnjë pikë.

Intervali i shtresave të thellësisë	Denduria e pikave
[0-1021]	0
[1022-1585]	28834
[1586-4094]	0

Figura 9. Shpërndarja e dendurisë së pikave sipas shtresave të thellësisë

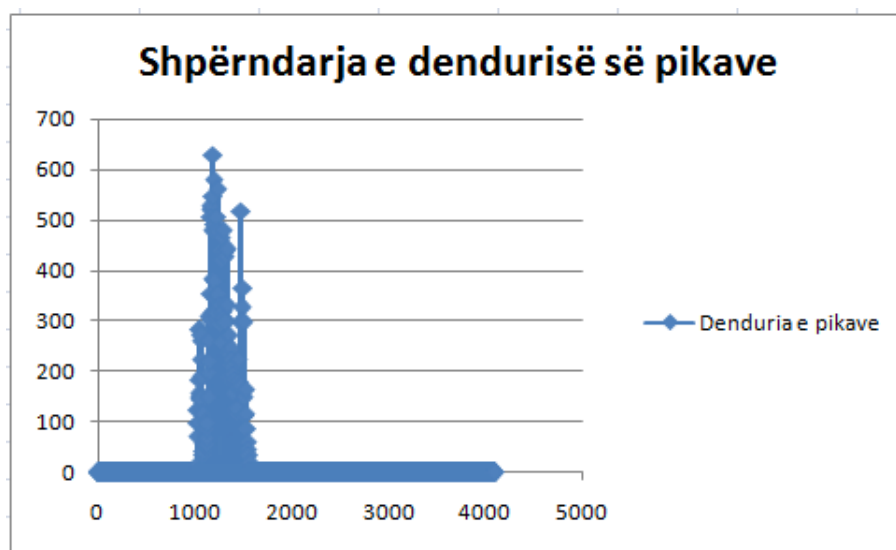


Figura 10. Shpërndarja e dendurisë së pikave sipas shtresave të thellësisë

Duke supozuar që dora e përdoruesit ndodhet në shtresat më të para të thellësisë fokusohemi vetëm në këto shtresa. Figura e mëposhtme jep paraqitjen e dorës së përdoruesit nga ekstraktimi i 100 shtresave të para duke filluar nga shtresa e parë e shfaqjes së skeletit të përdoruesit.

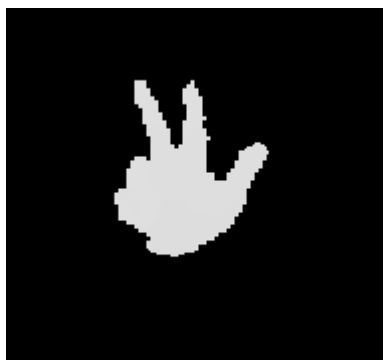


Figura 11. Identifikimi i dorës së përdoruesit ndërmjet identifikimit të skeletit

Përfundime

Pajisja Microsoft Kinect, nëpërmjet sensorëve hardware-ik mundëson kapjen e të dhënave të cilat nëpërmjet pëdorimit të paketave software-ike si Microsoft SDK lejon përpunimin e tyre. Identifikimi i dorës së përdoruesit është i rëndësishëm në komunikimet ndërmjet personave që nuk dëgjojnë pasi në shumicën e rasteve shenjat realizohen me anë të duarve. Kinect nëpërmjet kamerave të thellësisë 3D gjeneron bashkësi frame-sh të cilat

nëpërmjet përpunimit në kohë reale na shërbëjnë për lokalizimin e dorës. Për identifikimin e dorës kemi përdorur dy teknika: Teknika e parë konsiston në përdorimin e statistikave mbi llogaritjen e limiteve të thellësisë. Kjo teknikë është e vlefshme në rastet kur përpara pajisjes Kinect nuk vendosim asnjë objekt tjetër përveç përdoruesit. Teknika e dytë lokalizon fillimisht skeletin e përdoruesit dhe më pas dorën, shtresat e së cilës ndodhen më afër pajisjes Kinect. Teknika e dytë shmang problemet që mund të lindin si pasojë e futjes së objekteve të treta midis përdoruesit dhe pajisjes Kinect. Kufizimi i teknikës së dytë është pozicionimi i përdoruesit në një largësi të mjaftueshme nga Kinect, sa për tu shfaqur i gjithë skeleti.

Puna në të ardhmen

Identifikimi i dorës së përdoruesit të ndodhur përpara pajisjes Kinect lejon aplikimin në kohë reale të teknikave të procesimit të imazhit për të kuptuar numrin e gishtave, e vlefshme në një komunikim ndërmjet personave që nuk dëgjojnë dhe personave që dëgjojnë.

Paketa software-ike SDK lejon kapjen dhe analizimin e të dhënave por nuk mundëson mjete të nevojshme për ruajtjen e imazhit. Gjenerimi i imazheve nga pajisja Kinect është dinamik dhe i vazhdueshëm. Për këtë arsye është i nevojshëm ruajtja e tyre në një hapsirë të dhënash persistente, për analizime të mëvonshme. Zbatimi i teknika të procesimeve të imazheve dhe teknikave të mësimin makinë do mundësonte kuptimin e shenjave të personave që nuk dëgjojnë.

Literatura

Zico Pratama Putra (2014): A Natural User Interface Translation Tool: From Sign Language to Spoken Text and Vice Versa, Master Thesis

Cerezo, F. T. (2012). 3D Hand and Finger Recognition using Kinect. Marrë nga <http://frantracerkinecft.codeplex.com/>, e aksesuar më, dhjetor 2015

Kyatanavar, M. R., & Futane, P. (2012): Comparative Study of Sign Language Recognition Systems. International Journal of Scientific and Research Publications.

Wenjun Zeng (2012): Microsoft Kinect Sensor and Its Effect, Microsoft Research

James Ashley, Jarrett Webb (2012): Beginning Kinect Programming with the Microsoft Kinect SDK: Apress

Kinect for Windows Sensor Components and Specifications:

<https://msdn.microsoft.com/en-us/library/jj131033.aspx>, e aksesuar më, dhjetor 2015

Sign Language, (2013): Marrë nga Wikipedia:

http://en.wikipedia.org/wiki/Sign_language, e aksesuar më, dhjetor 2015