

Preprint of the paper

Šušić, M., Spasojević, S., Karan, B. "3D data acquisition and practical calibration of Microsoft Kinect camera" (in Serbian: "Akvizicija 3D podataka i praktična kalibracija Microsoft Kinect kamere.") In Proceedings of the 56th ETRAN Conference, Zlatibor, June 11-14, 2012 (Zbornik radova 56. konferencije ETRAN, Zlatibor, 11-14. juna 2012,) RO2.4.1-RO2.4.4, 2012.



This work is licensed under a [Creative Commons - Attribution-Noncommercial-No Derivative Works 3.0 Serbia](https://creativecommons.org/licenses/by-nc-nd/3.0/rs/)

AKVIZICIJA 3D PODATAKA I PRAKTIČNA KALIBRACIJA MICROSOFT KINECT KAMERE

Marko Šušić, *Institut Mihajlo Pupin, Beograd, marko.susic@pupin.rs*

Sofija Spasojević, *Institut Mihajlo Pupin, Beograd, sofija.spasojevic@pupin.rs*

Branko Karan, *Institut tehničkih nauka SANU, Beograd, branko.karan@itn.sanu.ac.rs*

Sadržaj – U radu su prikazani uvodni rezultati analize efikasnosti upotrebe Kinect kamere u realizaciji robotskih zadataka koji se odnose na manipulaciju. Razmotreni su i upoređeni softverski paketi Kinect for Windows i OpenNI sa stanovišta 3D akvizicije podataka. Implementirana je praktična metoda za kalibraciju Kinect kamere. Metoda uključuje identifikaciju unutrašnjih parametara RGB i dubinske kamere, kao i spoljašnjih parametara u vidu relativne transformacije između kamera.

1. UVOD

Kinect predstavlja uređaj nove generacije razvijen od strane Microsoft-a koji u sebi objedinjuje tehnike generisanja i obrade audio i vizuelnih signala. Prvobitno je namenjen komercijalnoj upotrebi u kombinaciji sa konzolom Xbox 360 ili Windows operativnim sistemom [1]. Kinect poseduje mogućnost detekcije ljudskih pokreta i kretanja približno na rastojanju od 0.6m do 4m, a replika tih pokreta može da se prikazuje na ekranu sa kojim je povezan. Kinect uređaj se sastoji od RGB kamere promenljive rezolucije, 3D dubinskog senzora (infracrveni projektor u kombinaciji sa CMOS monohromatskim senzorom – kamera osetljiva na IR zrake koje šalje projektor tj. Dubinska kamera) i širokopojasnog mikrofona, a poseduje i odgovarajući korisnički interfejs. Spoljašnji izgled Kinect uređaja je prikazan na slici 1:



Slika 1. Spoljašnji izgled Kinect uređaja

S obzirom na širok spektar sirovih i obrađenih signala koje Kinect može da pošalje kao izlazne informacije preko USB veze, programeri i istraživači su došli na ideju da bi Kinect mogao da se iskoristi i u svrhe naučnog istraživanja i razvoja. Sa tim ciljem Microsoft je obezbedio softverski razvojni alat *Kinect for Windows SDK*, koji u sebi sadrži razne funkcije i biblioteke na osnovu kojih se vrši komunikacija sa Kinect-om i generisanje izlaznih signala. Međutim, ovaj softverski alat pored brojnih mogućnosti, poseduje i neka ograničenja koja su diktirana politikom kompanije Microsoft. Zbog toga se javila potreba za kreiranjem novih softverskih paketa koji bi omogućili direktniji pristup i uspešniju komunikaciju sa Kinect-om [2].

Jedan od najkompletnijih i najčešće korišćenih softverskih alata koji je razvijen za svrhe komunikacije sa Kinect-om, ali i mnogim drugim hardverskim uređajima je OpenNI. U toku rada će pomenuti softverski paket biti detaljnije analiziran, a biće izvršena i uporedna analiza OpenNI-ja sa Kinect for Windows SDK.

Kinect je posebno veliku primenu i značajne rezultate ostvario u oblasti robotike. Algoritmi prepoznavanja pokreta i gestova koji su implementirani u Kinect-u, otvorili su brojne mogućnosti njegove primene kod upravljanja kretanjem robota pomoću gestovnih komandi. Takođe, razvijen je i princip upravljanja robotom pomoću govornih komandi. Korišćenjem RGB kamere u kombinaciji sa dubinskim senzorima, dobijaju se informacije o 3D okruženju robota na osnovu kojih robot može da obilazi prepreke u prostoru, dohvata i prenosi objekte ili obavlja kretanje sa nekim drugim ciljem u svojoj okolini. Za neke robotske zadatke, kao što je npr. prepoznavanje geometrijskih oblika u 3D prostoru, dohvatanje objekata i precizna manipulacija robotske ruke, posebno je od interesa poznavanje koordinata karakterističnih 3D tačaka u prostoru. Da bi se ove tačke egzaktno odredile, neophodno je poznavanje određenih parametara kamere, koji se dobijaju procesom kalibracije kamere. Upravo zbog neophodnosti realizacije ovog procesa, predmet ovog rada će između ostalog biti i kalibracija Kinect kamere.

2. OPENNI I POREĐENJE SA SOFTVERSKIM PAKETOM KINECT FOR WINDOWS SDK

OpenNI je softverski paket nastao sa ciljem poboljšanja prirodnog korisničkog interfejsa koji se koristi za interakciju sa hardverskim uređajima, pristup izlaznim signalima uređaja i kreiranje aplikacija koje koriste ove uređaje. Osnovna svrha i ideja nastanka paketa OpenNI je formiranje standardnog izvornog koda koji je javno dostupan za gore pomenute namene. Korisniku su na ovaj način dostupne mnogobrojne biblioteke, rutine, funkcije, razne strukture podataka, klase i promenljive za rad sa Kinect-om i sličnim uređajima. Shodno svojim potrebama, korisnik može da menja, koriguje i unapređuje kod u cilju rešavanja postavljenih zadataka.

Osnovna prednost OpenNI-ja u odnosu na Kinect for Windows SDK je mogućnost realizacije OpenNI aplikacija za rad na više operativnih sistema – Windows, Linux i Mac OSX, dok je Microsoft-ov SDK prilagođen samo Windows operativnom sistemu. Još jedna prednost OpenNI-ja koja se u nekim slučajevima pokazala kao značajna je mogućnost pristupa sirovim signalima sa infracrvenog senzora [3]. U ovom radu kalibracija Kinect kamere je realizovana pomoću metoda koji koristi dva tipa slika – slike u boji generisane pomoću RGB kamere i monohromatske slike dobijene pomoću kamere koja je osetljiva na IR zrake projektora (dubinska kamera). Jasno je da se ovaj zadatak ne bi mogao realizovati korišćenjem Kinect for Windows SDK.

Međutim, pri rešavanju zadataka koji koriste audio signale sa Kinect-a ili se odnose na probleme prepoznavanja govora, više se preporučuje i zastupljenija je upotreba paketa Kinect for Windows SDK [4]. Razlog je kompletniji pristup audio signalima i efikasnija obrada govora, kao i mogućnost eliminacije šuma. Ipak, najveći deo dostupnih aplikacija poseduju i jedan i drugi softverski paket, a performanse tih aplikacija su u većini slučajeva iste ili postoje neke minimalne razlike. To su npr. aplikacije koje se odnose na detekciju skeletona, pristup RGB i dubinskoj mapi i mnoge druge. Tačnost u praćenju kretanja skeletona je identična, a razlike postoje kod mogućeg izbora rezolucije generisanih slika sa RGB ili dubinske kamere.

Na kraju se može zaključiti da softverski paket treba izabrati na osnovu konkretnog zadatka koji se želi realizovati pomoću Kinect-a. Ipak, za većinu zadataka OpenNI i Kinect for Windows SDK su jednako efikasni, a na korisniku je da odluči koji softverski paket će koristiti.

3. PROCES KOMUNIKACIJE I GENERISANJA IZLAZNIH SIGNALA SA KINECT-A POMOĆU OPENNI SOFTVERSKOG PAKETA

Kao što je ranije pomenuto, OpenNI se između ostalog koristi za komunikaciju sa Kinect uređajem i generisanje izlaznih signala. Što se tiče izlaznih signala, moguće je pristupiti signalu sa RGB kamere i signalu sa dubinske kamere. Mogu se generisati slika u boji, IR mapa i dubinska mapa. Dubinska mapa je nastala procesiranjem i odgovarajućom obradom IR mape, i na osnovu nje se može zaključiti da li su objekti u vidnom polju senzora bliže ili dalje Kinect senzoru (svetliji objekti na dubinskoj slici su bliže senzoru, a tamniji dalje). Ukoliko su neki objekti suviše blizu (do 60 cm) ili suviše daleko (preko 4 m) od senzora, oni ne pripadaju vidnom polju senzora i na slici su pikseli koji reprezentuju te objekte crni. Ukoliko ne postoji informacija o udaljenosti nekih tačaka u prostoru, odgovarajući pikseli koji predstavljaju te tačke na slici su takođe crni. Primeri generisanih mapa i slika sa Kinect-a su dati na slikama 2-4:



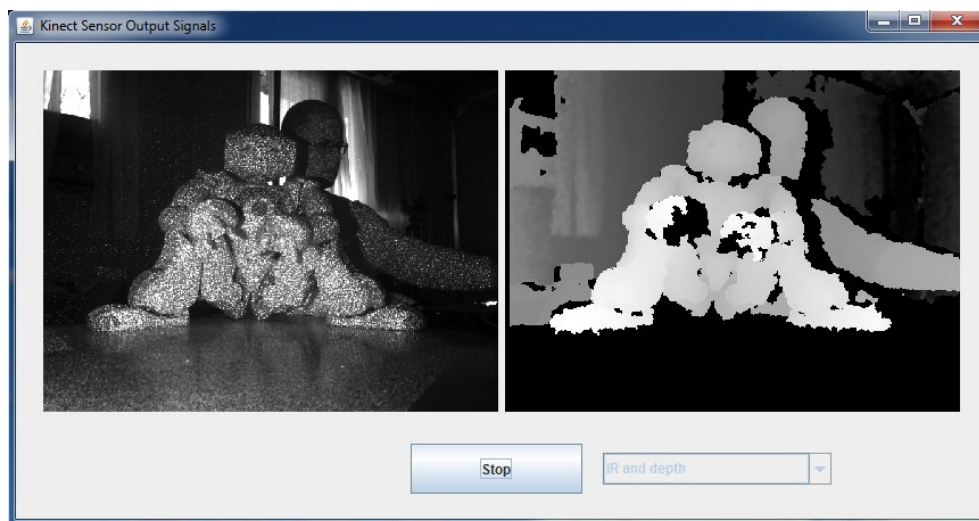
Slika 2. RGB slika sa Kinect kamere



Slika 3. IR slika sa Kinect senzora



Slika 4. Dubinska slika dobijena procesiranjem IR mape



Slika 5. Grafički interfejs za prikaz izlaznih signala sa Kinect-a

Gore pomenuti tokovi podataka sa Kinect-a (RGB, IR i dubinski) generišu maksimalno 30 slika u sekundi. Međutim, to je slučaj kada se dopusti tok samo jednog tipa signala ili odgovarajućeg para dva tipa signala. Sasvim solidno funkcionišu RGB signal i dubinski u paru, kao i IR i dubinski signal istovremeno. Problem se javlja kod pokušaja istovremenog generisanja RGB, IR i dubinskog signala, kao i kod generisanja RGB i IR signala u istim trenucima vremena. U poslednja dva pomenuta slučaja, broj slika u sekundi naglo opada i ovakav način generisanja izlaznih signala sa Kinect-a nije zastupljen u praksi. Grafički interfejs za prikaz i generisanje izlaznih signala sa Kinect-a, prikazan je na slici 5.

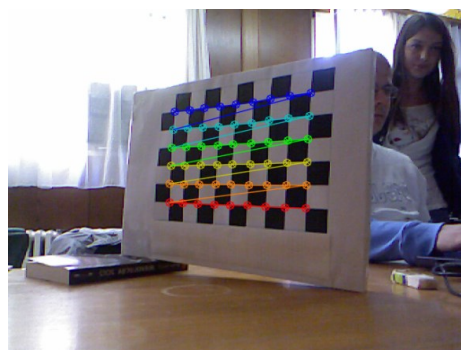
4. KALIBRACIJA KINECT KAMERE

Proces kalibracije kamere u opštem slučaju podrazumeva izračunavanje parametara kamere koji su od interesa za realizaciju algoritama koji koriste signale generisane sa kamere, bilo da su to pojedinačne slike ili sekvence slika. Svi parametri kamere se dele na spoljašnje i unutrašnje parametre. Unutrašnji parametri su žižna daljina, pozicija optičkog centra kamere i parametri distorzije. Spoljašnji parametri se odnose na parametre koji su potrebni za transformaciju koordinatnog sistema kamere u globalni koordinatni sistem i predstavljaju ih matrice rotacije i translacije. U slučaju Kinect-a, s obzirom da je potrebno kalibrisati RGB i dubinsku kameru, rezultati će se dobiti u formi unutrašnjih parametara za svaku od njih ponaosob, dok će se spoljašnji parametri u vidu matrica rotacije i translacije odnositi na relativni položaj kamera kada se globalni sistem veže za jednu od kamera.

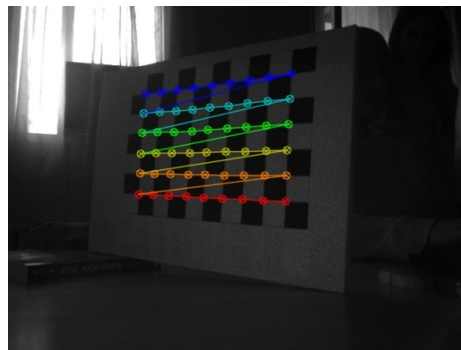
Algoritmi kalibracije koriste kao ulazne podatke slike kalibracionih objekata. To su objekti unapred poznatih dimenzija koji imaju istaknuta obeležja, ivice ili karakteristične tačke. Jedan od najpogodnijih i najčešće korišćenih kalibracionih objekata je šahovska tabla. Potrebno je snimiti šahovsku tablu u više različitih položaja. Za svaki položaj fiksirane šahovske table potrebno je generisati jednu za drugom RGB i IR sliku. IR slika u sebi sadrži refleksije projektovanih tačaka, što može negativno da utiče na procesiranje slike tokom algoritma, konkretno na automatsku ekstrakciju tačaka od interesa. Zbog toga se pri generisanju IR slika prekriva IR projektor, a rezultat su IR slike bez izraženih osvetljenja. One su na prvi pogled prilično tamne i šahovska tabla na njima je jedva vidljiva. Međutim, to za algoritam kalibracije ne predstavlja problem, s obzirom da on poseduje sposobnost ekstrakcije tačaka od interesa i u slučaju takvih slika. Automatska ekstrakcija tačaka od interesa na RGB i dobijenoj IR slici predstavljena je na slikama 6 i 7:

Tabela 1. Unutrašnji parametri RGB kamere

Unutrašnji parametri RGB kamere	
Fx_rgb	5.41598694e+002
Fy_rgb	5.50454102e+002
Cx_rgb	3.16908051e+002
Cy_rgb	2.07033234e+002



Slika 6. Ekstrakcija tačaka od interesa na RGB slici



Slika 7. Ekstrakcija tačaka od interesa na IR slici

Neke metode za kalibraciju Kinect RGB i dubinske kamere koriste dubinske slike umesto IR slika [5]. Međutim, u tom slučaju je moguće detektovati samo četiri karakteristične tačke šahovske table i to četiri krajnja temena table. S obzirom na ovu činjenicu, zbog manjka raspoloživih tačaka potrebno je generisati veliki broj slika i to nije jedini nedostatak ove metode. Već je pomenuto da se u dubinskoj slici mogu javiti šumovi, što može uticati na netačno detektovanje krajnjih temena table, a ova odstupanja se mogu reflektovati na tačnost rezultata kalibracije.

Algoritam opisan u ovom radu pored toga što koristi IR slike, koristi i neke funkcije OpenCV biblioteke [6]. Korišćene funkcije su se pokazale kao veoma efikasne i sveukupno su uticale na veću tačnost rezultata kalibracije.

Nakon ekstrakcije karakterističnih tačaka na svim generisanim parovima RGB-IR slika primenom optimizacionog algoritma uz postavljeni kriterijum minimalne greške dobijaju se parametri kamera prikazani u tabelama 1-6:

Tabela 2. Unutrašnji parametri dubinske kamere

Unutrašnji parametri dubinske kamere	
Fx_d	6.19996643e+002
Fy_d	6.30794006e+002
Cx_d	3.04181335e+002
Cy_d	1.71219696e+002

Tabela 3. Parametri distorzije RGB kamere

Parametri distorzije RGB kamere	
K _{1_rgb}	2.75411397e-001
K _{2_rgb}	-4.86645162e-001
P _{1_rgb}	-4.59333546e-002
P _{2_rgb}	-5.94996754e-003
K _{3_rgb}	0

Tabela 4. Parametri distorzije dubinske kamere

Parametri distorzije dubinske kamere	
K _{1_d}	-1.02668023e-002
K _{2_d}	1.50729537e-001
P _{1_d}	-3.91754471e-002
P _{2_d}	-2.92079337e-003
K _{3_d}	0

Tabela 5. Relativna transformacija između kamera – matrica rotacije

Relativna transformacija između kamera – matrica rotacije		
9.99971151e-001	3.18455906e-003	-6.89391652e-003
-3.42341512e-003	9.99384344e-001	-3.49174216e-002
6.77847583e-003	3.49400155e-002	9.99366403e-001

Tabela 6. Relativna transformacija između kamera – vektor translacije

Relativna transformacija između kamera – vektor translacije		
2.45786929e+000	4.08228412e-002	-1.42872393e+000

U tabeli 1 parametri F_{x_rgb} i F_{y_rgb} predstavljaju žižnu daljinu u pikselima. C_{x_rgb} i C_{y_rgb} se odnose na položaj optičkog centra kamere u pikselima. U tabeli 2 dati parametri imaju isto značenje kao u tabeli 1, samo se odnose na dubinsku kameru. U tabelama 3 i 4 su dati parametri distorzije RGB i dubinske kamere. Parametri K_1 , K_2 i K_3 predstavljaju koeficijente radijalne distorzije, dok se parametri P_1 i P_2 odnose na tangencijalnu distorziju. Tabele 5 i 6 reprezentuju relativni položaj RGB kamere u odnosu na dubinsku kameru, koji je dat matricom rotacije i vektorom translacije (brojke su date u centimetrima).

3D tačke u prostoru se na osnovu poznavanja parametara kamere i dubinskog senzora mogu izračunati na osnovu (1):

$$\begin{aligned} X &= (x - C_{x_d}) * depth(x, y) / F_{x_d} \\ Y &= (y - C_{y_d}) * depth(x, y) / F_{y_d} \\ Z &= depth(x, y) \end{aligned} \quad (1)$$

U prethodnoj relaciji parametri C_{x_d} , C_{y_d} , F_{x_d} i F_{y_d} su dobijeni kalibracijom dubinske kamere, x i y predstavljaju poziciju piksela u dubinskoj mapi, a $depth(x, y)$ je vrednost

dubine za posmatrani piksel, odnosno odgovarajuća vrednost dispariteta.

X , Y i Z su koordinate odgovarajuće 3D tačke u odnosu na globalni koordinatni sistem vezan za Kinect kameru. Parametri distorzije se koriste za svrhe korekcije zakrivljenosti slike.

5. ZAKLJUČAK

Oblast robotike se ubrzano razvija sa pojavom novih tehnologija. Posebno značajan segment robotike predstavlja upravljanje kretanjem robota. Postoji mnogo načina na koje se kretanje može realizovati, ali svaki od njih podrazumeva dobru interakciju robota sa okolinom. Kinect uređaj svojom multifunkcionalnošću pokriva širok spektar mogućih primena u robotici. Pomoću Kinect-a se može realizovati upravljanje kretanjem robota pomoću gestovnih ili glasovnih komandi, obilazak postavljenih prepreka ili dohvatanje objekata od strane robotske ruke. Fundamentalni korak u radu sa Kinect-om predstavlja kalibracija Kinect kamere. Kalibracijom se dobijaju parametri kamere neophodni za realizaciju algoritama robotskih zadataka. Značajnu stavku u radu sa Kinect-om predstavlja i pravilan izbor odgovarajućeg softverskog alata, koji se bira u skladu sa postavljenim zadatkom.

ZAHVALNICA

Ovaj rad je realizovan u okviru projekata III44008 i TR35003 koje finansira Ministarstvo za prosvetu i nauku Republike Srbije.

LITERATURA

- [1] Sean Kean, Jonathan Hall, Phoenix Perry, "Meet the Kinect", *Technology in Action*, 2011.
- [2] Jeff Kramer, Nicolas Burrus, Daniel Herrera C., Florian Ehtler, Matt Parker, "Hacking the Kinect", *Technology in Action*, 2012.
- [3] OpenNI. Internet: <http://openni.org> [25.03.2012.]
- [4] Kinect for windows. Internet: <http://www.microsoft.com/en-us/kinectforwindows/> [26.03.2012]
- [5] Kinect Calibration. Internet: <http://nicolas.burrus.name/index.php> [12.03.2012.]
- [6] Gary Bradski, Adrian Kaehler, "Learning OpenCV", *O'Reilly*, 2008.

Abstract – The paper presents introductory results on analyzing effectiveness of using Kinect camera in realization of robotic manipulation tasks. Two available communication interfaces, Kinect for Windows and OpenNI are considered and compared from the standpoint of 3D data acquisition. A practical method for calibration of Kinect camera is implemented. The method involves identification of internal parameters of Kinect RGB and depth camera and external parameters in the form of relative transformation between the cameras.

3D DATA ACQUISITION AND PRACTICAL CALIBRATION OF MICROSOFT KINECT CAMERA

Marko Šušić
Sofija Spasojević
Branko Karan