

# Virtualna stvarnost: hardver u službi zabave

Vladimir Cerić / Ivana Milošević / Mirko Milošević / Nemanja Janković

Audio i video tehnologije

Visoka škola elektrotehnike i računarstva strukovnih studija

Beograd

[ceric@viser.edu.rs](mailto:ceric@viser.edu.rs) / [ivanap@viser.edu.rs](mailto:ivanap@viser.edu.rs) / [mirkom@viser.edu.rs](mailto:mirkom@viser.edu.rs) / [nemanjaj@viser.edu.rs](mailto:nemanjaj@viser.edu.rs)

*Sažetak*—U radu je predstavljena analiza hardvera za virtualnu stvarnost koji se koristi u službi zabave, kroz snimani, a ne računarski generisani sadržaj. Akcenat je dat na primeni monoskopskih i omnidirekcionih kamera u procesu produkcije videa u 360°, odnosno prednosti i mane ovih rešenja kroz sagledavanje njihovih tehničkih karakteristika i nedostataka. Ovakvi video snimci omogućavaju niži stepen virtualne stvarnosti u odnosu na grafički renderovana 3D okruženja. Snimci u 360° omogućuju delimičnu interakciju gledaoca (učesnika) u okvirima već snimljenog video sadržaja pružajući pomeranje tačke gledanja gledaoca, po  $x$  i  $y$  osama, dok je  $z$  osa ograničena na osećaj trodimenzionalnosti sličan onom u klasičnom 3D videu. S druge strane data je analiza hardvera koji se koristi za gledanje (učestvovanje) u virtualnoj stvarnosti (hedsetovi) sa uvidom u specifična rešenja koja za posledicu imaju različita korisnička iskustva.

*Ključne reči*—virtualna stvarnost; video u 360°; monoskopske kamere; omnidirekzione kamere; hedsetovi.

## I. UVOD

Pojam virtualna stvarnost se najpre odnosi na računarski generisana 3D okruženja koja u realnom vremenu omogućavaju korisnicima interakciju sa simuliranim okruženjem. Stepennost stvarnosti ovakvih okruženja razlikuje se u zavisnosti od hardvera i softvera, od jednostavne 3D grafike do detaljnih prizora koja nalikuju filmu. Ipak u poređenju s filmom, virtualnu stvarnost razlikuje stepennost interakcije koju ima, s jedne strane gledalac, u filmu, a s druge strane igrač i učesnik u svetu virtualne stvarnosti. Ubedljivost okruženja virtualne stvarnosti najviše se ogleda u vremenskom i prostornom realizmu [1]. Vremenski realizam se odnosi na događaje koji se odigravaju u realnom vremenu, bez kašnjenja ili elipsi. Prostorni realizam se odnosi na prostor i objekte u prostoru s kojima učesnik, može da interaguje. Kada su ova dva uslova zadovoljena, učesnik utone u virtualno iskustvo, jer stimulansi virtuelne interakcije preuzimaju deo percepcije i kognicije. Virtualni svetovi nasuprot virtuelnoj stvarnosti imaju različit stepennost učešća ljudi u interakciji. Učesnici virtuelnih svetova koriste reprezentaciju putem avatara [2].

U ovom radu biće reči o video snimcima koji se stvaraju za tzv. hedsetove (headsets), izgled hedseta sa pratećim senzorima i kontrolorima dat je na Sl. 1.



Slika 1. Hedset, senzori i kontroleri za virtualnu stvarnost

Ovakvi video snimci, popularno se nazivaju video u 360° (eng. *360 video*) i omogućavaju niži stepennost virtualne stvarnosti u odnosu na grafički renderovana 3D okruženja. Ova virtualna stvarnost omogućava delimičnu interakciju gledaoca (učesnika) u okvirima već snimljenog video sadržaja. Za razliku od grafičkih 3D okruženja, video u 360° omogućava kretanje tačke gledanja gledaoca, po  $x$  i  $y$  osi, dok je  $z$  osa ograničena na osećaj trodimenzionalnosti sličan onom u klasičnom 3D videu.

## II. HARDVER ZA SNIMANJE

Za snimanje videa u 360° koriste se serijski uvezane monoskopske kamere ili omnidirekzione kamere [3], [4]. Monoskopske serijski uvezane i kalibrirane kamere predstavljaju najčešći oblik snimanja videa u 360°. Upravo ovakav način snimanja koristi Guglov (*Google*) servis *Strityju* (Street View). Monoskopsko snimanje obuhvata ravne rendere u 360° kadrova koji se mogu gledati na bilo kom ravnom ekranu ili hedsetu.

Hardver monoskopskih kamera omogućava tokom reprodukcije pomeranje tačke gledanja učesnika po prostoru (po  $x$  i  $y$  osi, odnosno tiltovanje i panovanje tačke gledanja), ali bez realne percepcije dubine. U postprodukciji (koja može biti hardverska ili softverska), oponašanje stereoskopskog vida, postiže se kombinovanjem zasebnih video strimova s malim pomeranjem slike, uporedne tehnike su date na Sl. 2.



slika pored slike (SBS)



slika ispod slike (TB)



anaglif

Slika 2. Poređenje tehnika za 3D reprodukciju

Tehnika za 3D reprodukciju slika pored slike (eng. *side-by-side* ili *SBS*), koristi punu rezoluciju slike za oba oka istovremeno, tako što polovina horizontalno skalirane slike biva distribuirana levom, a polovina slike desnom oku. Tako se na primer, *FullHD* video sadržaj rezolucije (1920x1080) zapravo distribuira svakom oku u rezoluciji (960x1080). Kada ovakav signal treba da se reprodukuje na 3D displeju, signal se deli na slike namenjene svakom oku i zatim skalira na punu rezoluciju. Nakon ovoga će svaka zasebna slika biti prikazana u alternaciji sinhrono s aktivnim 3D naočarima.

Druga tehnika za 3D, slika ispod slike, je slična prethodnoj, izuzev što se cela slika za svako oko skalira po vertikalnoj osi kako bi zauzela gornju ili donju polovinu slike.

Najstarija tehnika za 3D koja se danas ređe koristi je anaglif. U ovoj tehnici leva slika se boji u plavo, a desna slika se boji u crveno. Oba videa se zatim postavljaju jedan iznad drugog, što uz nošenje naočara s plavim i crvenim filterom stvara 3D efekat.

#### A. Monoskopski serijski uvezane kamere – GoPro Odyssey

*GoPro Odyssey* se sastoji od 16 sinhronizovanih, serijski povezanih, *GoPro Hero4 Black* kamera. Kamere su postavljene na nosač, koji ima 16 ležišta za svaku pojedinačnu kameru raspoređenih u krug kako bi se obezbedilo: precizno pozicioniranje, adekvatno vidno polje, relativno preklapanje slika s kamera. Jedan od problema ovakve postavke jesu mrtvi uglovi pri snimanju u zenitu i nadiru. Ipak, zahvaljujući tome što nadir nije vidljiv, moguće je sakriti produkcijsku opremu

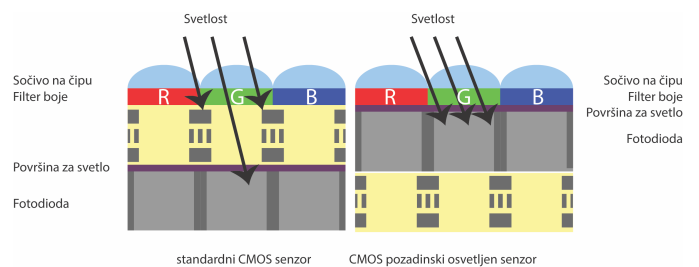
(stativ, napajanje i sl.) i snimatelja izvan vidnog polja. Izgled kamere je dat na Sl. 3.



Slika 3. Snimanje sa *GoPro Odyssey* kamerom

Kamere su povezane UTP kablom da bi se uspostavila sinhronizacija. Neophodno je da svih 16 kamera budu u sistemu i pravilno podešene jer nije moguće snimati sa manjim brojem kamera. *Odyssey* koristi eksterno baterijsko napajanje od 14,4 V, s kojim se povezuje putem četvoropinskog XLR priključka i omogućava autonomiju od 7 sati snimanja ili reprodukcije.

Svaka *Hero4 Black* kamera sadrži CMOS senzor od 12 megapiksela, veličine 6,17 mm x 4,55 mm, s fiksiranim otvorom blende  $f/2,8$ . Ugrađen je pozadinski osvetljen (BSI) CMOS senzor [3], šematski prikazan na Sl. 3, koji drugačijim rasporedom elemenata omogućava prihvatanje veće količine svetla, zbog čega su pojačane performanse u uslovima slabe vidljivosti. Kod ovakvih senzora, elementi su raspoređeni tako da se poluprovodničke komponente i veze između njih nalaze iza fotokatodnog sloja, tako da ne ometaju put svetlosti. Ovaj senzor prvi put je implementirao Sony, 2009. godine, u ovoj veličini, dok je tzv. *fullframe* senzor prvi put korišćen 2015. godine. Ovaj senzor *Hero4 Black* kamere omogućava snimanje u maksimalnoj rezoluciji od 3840x2160 piksela (4K) u 30 slika/sekundi.



Slika 4. Poređenje standardnog i BSI CMOS senzora

Ukratko, *GoPro Odyssey* koristi *Google assembler Jump*, koji 16 različitih video snimaka sa *Hero4 Black* kamera spaja u finalnu sliku koja ima softverski određenu dubinu, tj. u stereoskopski virtualni video u 360°.

Assembler analizira sliku tražeći karakteristične tačke u sadržaju zahvaljujući kojima po spoljnoj liniji vidnog polja

svakog sočiva spaja različite video snimke. U procesu dodavanja dubine (3D efekta), assembler zbog različitih pozicija objektiva u prostoru i karakteristika svake pojedinačne kamere, radi grubo centriranje i pozicioniranje slika, korekciju boja (balans bele) i kompenzaciju različitih ekspozicija delova scene koja se snima. 3D centriranje radi putem uočavanja razlike dubine objekata u sceni (između susednih objektiva), što omogućava stvaranje interpoliranih tačaka gledanja, što daje 3D efekat sceni. Svaki zasebni video strim poseduje specifične metapodatke koji su namenjeni postprodukciji, odnosno spajanju u jedan video.

Snimanje videa uz pomoć *GoPro Odyssey*, daje najbolje rezultate kada se objekat nalazi na udaljenosti većoj od jednog metra. Takođe, pri spajanju videa moguće su greške: kada su tanki objekti koji se snimaju u brzom pokretu ili je pozadina koja se snima sačinjena od šara ili istih predmeta (objekata) koji se uzastopno ponavljaju.

Svaka *GoPro Hero4 Black* kamera je unapred podešena na rezoluciju 2704x2028 (2,7K) u odnosu strana 4:3, pri snimanju u 30 ili 25 slika u sekundi.

*Google* assembler *Jump* nakon spajanja video strimova u jednu datoteku omogućava rezolucije od od 2Kx2K do 8Kx8K. Izlazni format je MPEG-4 s protokom od 150 do 600 Mbit/s.

### B. Omnidirekzione kamere – Ricoh Theta S

*Ricoh Theta S* je omnidirekciona kamera koja se u sebi sadrži dva identična sočiva, postavljena na suprotnim stranama kućišta. Da bi se eliminisali mrtvi uglovi pri snimanju, sočiva su izbačena izvan konstrukcije kućišta kamere, prikazano na Sl. 4. Ovakva konstrukcija kamera omogućava snimanje u zenitu i nadiru jer su pojedinačni uglovi snimanja svakog objektiva 190° [3]. Na ovaj način, uglovi snimanja svakog objektiva obezbeđuju po 10° preklapanja na spoljnim ivicama polusfere tako da se ostvari kvalitetno spajanje i video od punih 360°. Za razliku od *GoPro Odyssey* kamera, *Ricoh Theta S* u sebi sadrži softver za spajanje dva video strima u jedan snimak.



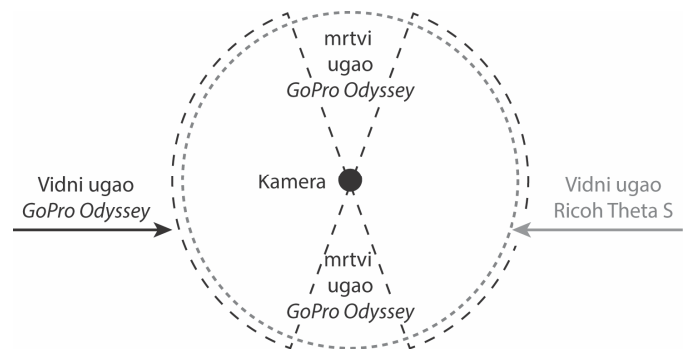
Slika 5. Spoljašnost sočiva *Ricoh Theta S* omnidirekzione kamere

Svaki objektiv *Ricoh Theta S* kamere sadrži standardni CMOS senzor od 12 megapiksela, veličine 6,17 mm x 4,55 mm, s fiksiranim otvorom blende f/2,0. Objekat koji se snima

može biti najmanje udaljen od objektiva 10 cm. Zbog same konstrukcije omnidirekcionih kamera neizbežna je pojava izobličenja na objektima koji su izuzetno blizu (npr. ruka snimatelja) [4][5].

*Ricoh Theta S*, omogućava snimanje do 65 minuta u FullHD kvalitetu sa 16Mbit/s.

Snimljeni video je punih 360°, bez stereoskopskog 3D efekta. Međutim, *YouTube* servis podržava format ove kamere, i prilagođava video stereoskopskoj reprodukciji. Omogućavajući pomeranje tačke gledanja po  $x$  i  $y$ , dok osećaj trodimenzionalnost gotovo da ne postoji, odnosno vezan je isključivo za snimljeni sadržaj.



Slika 6. Dvodimenzionalni prikaz uglova snimanja kamera *GoPro Odyssey* i *Ricoh Theta S*

Na Sl. 4. dat je uporedni dvodimenzionalni prikaz uglova snimanja *GoPro Odyssey* i *Ricoh Theta S* kamera, uz naznačene mrtve uglove kod monoskopski serijski uvezanih kamera.

### III. HARDVER ZA REPRODUKCIJU

Osnovni hardver predstavljaju tzv. hedsetovi odnosno naočare koje omogućavaju reprodukciju sadržaja za virtualnu realnost. U pogledu konstrukcije hedsetovi se dele na one koji imaju ugrađene displeje i senzore i hedsetove koji služe kao konstrukcija u koju se postavljaju nezavisni uređaji (najčešće smartfon). Ova hardverska razlika ogleda se najpre u kvalitetu slike, jer displeji smartfona mogu imati ukupno veću, ali u stereoskopskom gledanju znatno manju rezoluciju od hedsetova sa ugrađenim zasebnim displejima za svako oko. Najveću popularnost na tržištu, naročito uzimajući u obzir procenat korisnika smartfona imaju jeftini hedsetovi koji u sebi nemaju ugrađene displeje i senzore [2].

#### A. Hedsetovi kao konstrukcija u koju se postavljaju nezavisni uređaji

Osnovni hardver predstavljaju tzv. hedsetovi odnosno naočare koje omogućavaju reprodukciju sadržaja za virtualnu realnost. Postoje različiti hedsetovi kojima najčešće odgovaraju smartfoni određenog proizvođača što je vezano za dimenzije aparata, ali i softver (uz podržani hardver) koji je razvijen za tu vrstu hedsetova. Najjeftinije rešenje razvio je *Google* (proizvod *Cardboard* od kartona, prikazan na Sl. 7), hedset koji se jednostavno sklapa po perforiranim ivicama i odgovara različitim tipovima smartfona. Danas postoje brojne ergonomski i tehnički unapređene varijante (npr. *DodoCase*).



Slika 7. Google Cardboard sa postavljenim telefonom

U oba slučaja ovi headsetovi se u potpunosti oslanjaju na senzore koji su ugrađeni u smartfon izuzev sočiva koja omogućavaju stereoskopski vid. Sočiva svih headsetova su bikonveksna, a u zavisnosti od dimenzija i fokalne dužine (ali i dimenzija smartfona) menjaju se svi parametri slike.

#### B. Headsetovi sa ugrađenim displejima i sensorima

Headsetovi sa ugrađenim displejima i sensorima omogućavaju uživanje u punoj virtualnoj realnosti (renderovana 3D okruženja) koja uključuje osećaj utonulosti (eng. *immersion*) u virtualni svet, odnosno punu interakciju u virtualnom okruženju. Najčešće ovakvi headsetovi dolaze sa ugrađenim žiroskopom i akcelerometrom, uz senzore praćenja i kontrolere [2].

Na tržištu postoje različiti headsetovi koji se razlikuju po rezolucijama ekrana, po osvežavanju slike, vidnom polju i pratećim sensorima i kontrolerima. Najkompletnije rešenje za virtualnu realnost na tržištu je *HTC Vive* koji poseduje ukupnu rezoluciju od 2160x1200 piksela, odnosno 1080x1200 po displeju, osvežavanje slike od 90 slika po sekundi i vidno polje od preko 110°. Za *SteamVR* kontrolere, potrebno je do 3x3 m prostora u kome se interaguje. Postoje rešenja sa znatno većom rezolucijom, poput *StarVR*, koji nudi ukupnu rezoluciju 5k i 210° vidnog polja, uz manje osvežavanje slike i bez kontrolera.

Važna karakteristika svih headsetova s ugrađenim ekranima je podrška za 6DOF (eng. *6 degrees of freedom*) - odnosno, podršku za transliranje (napred/nazad, gore/dole, levo/desno) i rotiranje unutar virtualnog sveta.

#### IV. ZAKLJUČAK

U ovom radu predstavljena su najpopularnija hardverska rešenja za snimanje i reprodukciju videa u 360° koji omogućavaju uživanje i zabavu u nižem stepenu virtualne stvarnosti u odnosu na grafički renderovana 3D okruženja. U zavisnosti od tehnologija koje se koriste u produkcijskom i reprodukcijom lancu, video sadržaji namenjeni uživanju u virtualnoj stvarnosti, zapravo omogućavaju različite stepene osećaja *utonulosti* u virtualni svet.

Treba napomenuti, da virtualna realnost, ali i videi u 360° kod nekih gledalaca u veoma kratkom vremenskom periodu izazivaju osećaj opšte nelagodnosti, glavobolju, mučninu, zamor, dezorijentaciju i sl. Naučnici se još uvek ne slažu oko razloga nastanka ovih tegoba jer su veća istraživanja još uvek u

toku, a rezultati pojedinačnih studija u manjoj ili većoj meri kontradiktorni. Veruje se da u tehničkom pogledu, ove zdravstvene tegobe stvaraju se zbog razlike u subjektivnom i objektivnom pokretu, vidnom polju i uglovima gledanja. Na pojačavanje zdravstvenih tegoba utiče i vreme provedeno u virtualnom svetu, a najveći problem predstavljaju ekstremni uglovi gledanja i rakursi, naročito ukoliko se radi o virtualnim svetovima koji prikazuju ekstremitete avatara (ili snimatelja/učesnika) u izobličnim proporcijama.

Trenutno, rešenja koja postoje na tržištu nude pod veoma sličnim opisima različita korisnička iskustva. Tako treba obratiti pažnju na mogućnosti osećaja dubine slike (z osa), ugao snimanja, rezoluciju i ugao reprodukcije i ukupni kvalitet u pogledu osvežavanja slike, izobličjenja (uglovi i spajanje različitih video strimova), ali i same ergonomije headsetova s jedne strane, a mogućnosti snimanja i spajanja videa na strani produkcije s druge strane.

#### ZAHVALNICA

Ovaj rad rađen je u okviru istraživanja na projektu III44006 Ministarstva za nauku i tehnološki razvoj Srbije.

#### LITERATURA

- [1] Liang, J., C. Shaw and M. Green, "On temporal-spatial realism in the virtual reality environment," UIST '91 Proceedings of the 4th annual ACM symposium on User interface software and technology, 19-25, 1991
- [2] S. M. LaValle, "Virtual Reality", Cambridge University Press, Cambridge, 2017
- [3] H. Maître, "From Photon to Pixel: The Digital Camera Handbook", ISTE Ltd and John Wiley & Sons, Inc, London and Hoboken, 2015
- [4] L. Puig, J. J. Guerrero, "Omnidirectional Vision Systems: Calibration, Feature Extraction and 3D information", Springer, London, 2013
- [5] D. Stump, "Digital Cinematography," Focal Press. New York and London, 2014

#### ABSTRACT

This paper presents an analysis of virtual reality hardware used for entertainment, through filmed content instead of computer generated content. The accent is given to the use of monoscopic and omnidirectional cameras in 360 video production and pros/cons of these solutions through overview of their technical specs and design flows. This kind of video footage allows for a lower degree of virtual reality opposed to rendered 3D environments. 360 degrees' virtual reality videos enable limited interaction of viewers (participants) in prerecorded video footage, which is reflected in possibility for the viewers to move their point of view along x and y axes, while z axis is limited to the sense of 3D that is known from classic (flat) 3D video recording and projection. On the other hand, the analysis of hardware (headsets) used for viewing (participating) in virtual reality is given with respect to specific solutions that have an impact to end-user experience.

#### VIRTUAL REALITY: HADWARE FOR ENTERTAINEMENT

Vladimir Cerić  
 Ivana Milošević  
 Mirko Milošević  
 Nemanja Janković