

Večmodalna interakcija v navidezni okoljih

Jože Guna, Jure Žilavec, Matevž Pogačnik, Sašo Tomažič, Andrej Kos in Janez Bešter
Laboratorij za telekomunikacije, Fakulteta za elektrotehniko,
Univerza v Ljubljani, Tržaška cesta 25, 1000 Ljubljana, Slovenija
E-pošta: joze.guna@fe.uni-lj.si

Multimodal interaction in virtual environments

In the following paper we address the issue of multimodal user interfaces for virtual environments. We present the topics of multimodal interaction and virtual environments in general and apply voice, motion and gesture based multimodal interface to a Second Life virtual environment. We present the key system elements and the architecture of the pilot setup. Preliminary user evaluation results are very positive and show great potential for future research.

Keywords: multimodal interaction, gesture, HCI, virtual world

1 Uvod

Komuniciranje in medsebojna interakcija sta eni zmed osnovnih gonilnih sil razvoja človeške družbe. »Človek ne more ne-komunicirati.« se glasi znani rek raziskovalca in teoretika komunikacijskega procesa, Paula Watzlawika [1].

Ljudje ne komuniciramo samo med seboj, t.j. preko interakcij tipa človek-človek (angl. Human-Human Interaction, HHI), temveč tudi z neživimi napravami, kar je predmet raziskovalnega področja interakcij tipa človek-stroj (angl. Human-Computer Interaction, HCI). Raziskovalno področje HCI je močno interdisciplinarno, kjer se med seboj prepletajo telekomunikacijske, računalniške, biološke in psihološke znanosti. Glavni cilji raziskovanja HCI so ljudem čimbolj približati uporabniške vmesnike in interakcije tipa človek-stroj ter jih narediti intuitivne, naravne, varne ter predvsem enostavne za uporabo.

Pri tem je pomembno, da komunikacijo tipa človek-stroj čimbolj približamo medsebojni človeški komunikaciji z uporabo različnih komunikacijskih poti oz. modalnosti. S kombinacijo različnih modalnosti interakcije, kot je npr. uporaba klasične tipkovnice in računalniške miške v kombinaciji z govorom in gestami, postane interakcija med človekom in napravami veliko bolj intuitivna in enostavna. Osnovno področje raziskovanja HCI tako preide v večmodalno [2, 3] HCI (angl. MultiModal HCI, MMHCI).

Kombinacija klasičnih vhodnih vmesnikov ter govornih ukazov in izražanja z motoričnimi kretnjami oz. gestami je še posebno primerna za interakcijo v navidezni svetovih. V tem prispevku se zato

osredotočamo predvsem na inovativne večmodalne načine interakcij v navidezni okoljih.

Pričujoči prispevek je v nadaljevanju organiziran v naslednjem vrstnem redu. Drugo in tretje poglavje predstavlja problematiko večmodalnih uporabniških vmesnikov in navidezni svetov ter motivacijo za izvedbo alternativnih načinov interakcije. Četrto poglavje opisuje uporabljeno metodologijo ter peto pilotsko postavitev sistema. Prispevek zaključimo z diskusijo in sklepnimi mislimi v zadnjem poglavju.

2 Večmodalni uporabniški vmesniki

Večmodalni uporabniški vmesnik z uporabnikom komunicira preko kombinacije različnih komunikacijskih poti oz. modalnosti. Komunikacija je obojesmerna in jo lahko definiramo v smislu:

- komunikacija stroj->človek (izhodni vmesnik) – človeško zaznavanje preko čutil (vid, sluh, otip, vonj in okus), kjer človek sprejema informacije od stroja oz. računalnika in
- komunikacija človek->stroj (vhodni vmesnik) – stroj sprejema informacije od človeka preko različnih senzorjev (npr. kamera - računalniški vid, mikrofoni - razpoznavanje govora, haptični vmesniki – otip, različni biometrični senzorji).

Tipičen vhodno-izhodni grafični uporabniški vmesnik temelji na konceptu WIMP (angl. Windows, Icons, Menus, Pointing device), ki določa osnovni način implementacije z uporabo namizja, oken, ikon in miške za navigacijo. Klasični vizualni vmesniki postanejo bistveno bolj uporabni in učinkoviti z uporabo kombinacije tudi drugih modalnosti interakcije. Koncept računalniške miške in tipkovnice je mogoče dobro dopolniti z govornimi ukazi ter uporabo specializiranih daljinskih upravljalnikov in igralnih palic. Ti pripomočki poleg klasičnega vmesnika z gumbi omogočajo tudi druge modalnosti, kot sta npr. taktilni odziv in upravljanje z gestami [4]. Dobro načrtovan večmodalni uporabniški vmesnik uporabnikom omogoča večjo svobodo pri izbiri načina komunikacije s strojem ter s tem omogoča bistveno bolj naravno in enostavno uporabo sistema [5, 6].

3 Navidezna okolja

Navidezno okolje oz. svet (angl. Virtual World) je oblika socialne omrežne skupnosti, ki je največkrat

realizirano kot računalniška simulacija. Uporabniki v navideznem svetu nastopajo kot avatarji, ki predstavljajo njihovo podobo in osebnost (i.e. alter ego) v simuliranem navideznem okolju. Poleg živih uporabnikov so v navideznih svetovih pogosti tudi računalniško simulirani neživi igralci in elementi okolja. Glavni poudarek v navideznih svetovih je na interakciji in komunikaciji med živimi uporabniki samimi ter z okoljem.

Po načinu tehnične izvedbe navidezni svetovi uporabnikom omogočajo različno stopnjo realizma in vključenosti v okolje (angl. Immersion). Prvi navidezni svetovi so bili realizirani v obliki MUD (angl. Multi-User Dungeon) [7] večigralskih iger v zelo enostavnem tekstovnem načinu. Naslednji nivo realizma je omogočala izvedba v dvodimenzionalni grafični obliki z dodanih zvokom in efekti. Sodobni navidezni svetovi spadajo danes med računalniško najbolj izpopolnjene in zahtevne aplikacije. Tipično so izvedeni v vrhunski tridimenzionalni grafični podobi, s prostorskim zvokom in efekti. Dodatno stopnjo realizma omogoča delni 3D način prikaza s stereovizijo ter uporaba naprednih vhodno-izhodnih naprav, kot so prilagojeni igralni pripomočki z možnostjo zaznavanja gibanja, pripomočki s haptično-taktilnim odzivom ter vmesniki za upravljanje z mislimi in zaznavanje čustev uporabnika [8].

Najbolj znani navidezni svetovi spadajo v skupino masovnih večigralskih omrežnih iger MMORPG (angl. Massively Multiplayer Online Role Playing Game). Gre za tehnično izjemno dodelane aplikacije, ki pa poleg elementov zabave, posegajo tudi na druga področja človeškega delovanja, kot so socialno, pravno in ekonomsko. Najbolj znani predstavniki so WoW (angl. World of Warcraft) z več kot 10 milijoni aktivnih igralcev, LOTRO (ang. Lord of the Rings Online) ter Entropia. Slednji omogoča tudi povezavo med navidezno in realno ekonomijo preko vgrajenega borznega in poslovnega modela.

Poleg izrazito zabavno usmerjenih navideznih svetov drugo skupino predstavljajo navidezni svetovi s poudarjenimi socialnimi in izobraževalnimi aspekti. Med najbolj znanimi je svet Second Life (Slika 1) podjetja Linden Labs [9]. Podobno kot pri drugih navideznih svetovih, so uporabniki v Second Life predstavljeni z njihovimi avatarji. Razlike se kažejo v prilagodljivosti in stopnji personalizacije, ki je tu bistveno bolj poudarjena. Second Life nima vnaprej strogo začrtane osrednje zgodbe. Namesto tega predstavlja skupno platformo za medsebojno komunikacijo, predstavitev, izobraževanje in zabavo. Predstavlja možno alternativo svetovnemu spletu v 3D obliki. Posamezniki in organizacije imajo namesto predstavitev v obliki spletnih strani kar svoje virtualne tridimenzionalne posesti. Posebnost predstavlja tudi virtualna valuta Linden dolar (L\$), ki je osnova vgrajenega ekonomskega sistema. Možna je obojesmerna izmenjava monetarnih sredstev – iz realnega sveta v navideznega in obratno. Second Life

ima velik potencial [10], ni pa še polno zaživel. Trenutno je v Second Life približno milijon (aktivnih) uporabnikov.

Posebno zanimiv način uporabe je e-izobraževanje v navideznem svetu, ker fizične omejitve iz realnega sveta ne veljajo v virtualnem. Medsebojno lahko komunicira in si izmenjuje ideje in vsebine skoraj poljubno število uporabnikov v samo z domišljijo omejenem okolju. Primer tovrstne uporabe predstavlja Open University v Second Life svetu [11]. Za uporabnike pa sta pomembni tako vsebinska kot tudi sama tehnična izvedba. Pri tem sta ključnega pomena uporabniški vmesnik in način interakcije z drugimi uporabniki ter okoljem.

Večmodalni vmesniki ter načini interakcije so še posebni primerni za implementacijo uporabniških vmesnikov v navideznih okoljih, kar je motivacija za opisani prispevek.



Slika 1: Navidezni svet Second Life

4 Večmodalna interakcija v svetu SecondLife

Klasičen pristop navigacije v okolju Second Life poteka z uporabo računalniške tipkovnice in miške. Ta kombinacija sicer omogoča polno funkcionalnost, je pa glede na specifični namen (npr. zgolj osnovna komunikacija in raziskovanje navideznega sveta) oz. za določene ciljne uporabniške skupine (npr. mlajši, starejši, ljudje s posebnimi potrebami) pogosto prezahtevna za uporabo. Zato postavimo hipotezo, da je računalniški pristop smiselno dopolniti oz. v celoti zamenjati z bolj naravnim in enostavnejšim vmesnikom. Prav tako postavimo hipotezo, da je smiselna uporaba več različnih modalnosti pri interakciji človek-stroj.

V naši pilotski postavitvi smo kot primarno krmilno napravo uporabili kombinacijo Nintendo Wii [12] daljinskega upravljalnika WiiMote z dodatkom Wii Nunchuk. Kontroler WiiMote je zelo zmogljiva večmodalna vhodno/izhodna naprava v obliki enostavnega daljinskega upravljalnika, prikazanega na sliki 2. Kljub igralnemu izvoru, gre za zelo prilagodljivo napravo, ki se uporablja v številnih izobraževalnih, rehabilitacijskih in znanstvenih aplikacijah [13, 14].



Slika 2: Nintendo WiiMote

WiiMote združuje različne modalnosti in sicer [15]:

- Vhodni vmesniki:
 - triosni pospeškometer z območjem zaznavanja pospeška reda +/- 3g, 8 bitno natančnostjo ter vzorčenjem s frekvenco 100 Hz,
 - infrardeča kamera ločljivosti 1024x768 slikovnih elementov z vzorčno frekvenco 100 Hz,
 - 8 gumbov in digitalni smerni križ.
- Izhodni vmesniki:
 - taktilni povratni odziv preko vgrajenega vibratorja,
 - zvočnik,
 - svetlobni odziv preko štirih led diod.
- Razširitveni vtič za različne dodatke (npr. dodatni pospeškometer, giroskop, igralna palica, merilnik pulza, itd.).



Slika 3: Wii Nunchuk

Wii Nunchuk (Slika 3) je dodatni vmesnik, ki se priključi neposredno na WiiMote. Vgrajeno ima analogno krmilno palico, dva gumba ter dodatni pospeškometer. Primarno je namenjen uporabi v levi roki in sekundarni navigaciji.

Za zajem zvoka smo uporabili klasični računalniški mikrofون.

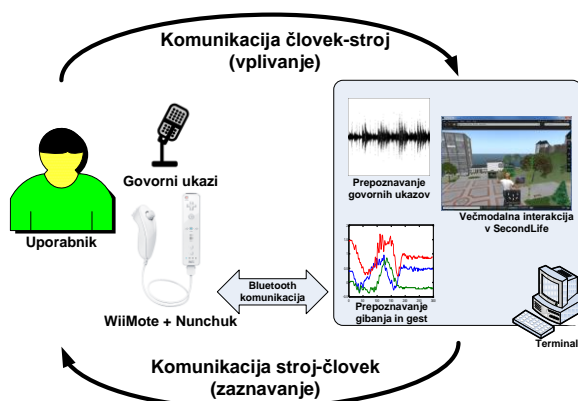
V enoten vmesnik smo združili naslednje modalnosti:

- krmiljenje gibanja uporabniškega avatarja s pomočjo gumbov in digitalnega smernega križa,
- krmiljenje gibanja uporabniškega avatarja s pomočjo analogne krmilne palice Wii Nunchuk,
- izvajanje akcij z uporabo gest,
- izvajanje akcij preko govornih ukazov,
- taktilni povratni odziv preko vibratorja,
- svetlobni povratni odziv preko LED luči,
- zvočni povratni odziv preko vgrajenega zvočnika v WiiMote.

5 Pilotska postavitve

V pilotski postavitvi se je okolje Second Life izvajalo na osebem računalniku z WindowsXP operacijskim sistemom.

WiiMote in Nunchuk z osebnim računalnikom komunicirata preko Bluetooth protokola. Interakcija med krmilnimi napravami in aplikacijo Second Life je bila sprogramirana preko namenske skripte v posebni aplikaciji. Za sintezo in prepoznavo govora je bila uporabljena programska knjižnica podjetja Microsoft. Arhitektura sistema je prikazana na sliki 4.



Slika 4: Arhitektura sistema

Uporabnik svojega avatarja upravlja preko smernega križa na WiiMote krmilniku, kjer puščici levo-desno pomenita obračanje levo in desno ter puščici gor-dol dviganje oz. spuščanje. Za osnovno premikanje (levo, desno, naprej, nazaj) je uporabljena analogna krmilna palica na krmilniku Nunchuk. Določene akcije, npr. letenje, je mogoče izvesti preko gest. Gesta, kjer uporabnik sunkovito premakne WiiMote v smeri navzgor tako aktivira oz. deaktivira letenje avatarja. Prav tako je mogoče upravljanje z govornimi ukazi, kjer npr. ukaz »jump« pomeni skok. Uporabnika med uporabo aplikacije prikazuje slika 5.

Sistem smo preliminarno preizkusili na petih uporabnikih. Uporabniki so najprej za interakcijo v svetu Second Life uporabljali klasično kombinacijo računalniške miške in tipkovnice. Osnovne naloge so

bile premikanje (hoja, letenje, skakanje) v navideznem svetu ter interakcija z objekti. Nato so uporabniki poskus ponovili z našim predlaganim vmesnikom, ki je obsegal WiiMote in Nunchuk krmilni enoti ter ukazovanje z glasovnimi ukazi. Razliko med obema uporabniškima izkušnjama so uporabniki opisali na tro-nivojski Likertovi skali (slabše, enako, bolje). Vseh pet uporabnikov se je strinjalo, da je predlagani večmodalni pristop interakcije boljši kot klasični.



Slika 5: Uporabnik med uporabo aplikacije

6 Diskusija in sklep

V prispevku smo predstavili alternativne večmodalne načine interakcije v navideznih okoljih. Sistem smo preizkusili na primeru sveta Second Life. Povratni odziv testnih uporabnikov našega večmodalnega sistema je zelo vzpodbuden, saj se jim je naš sistem interakcije zdel bolj intuitiven in enostaven.

Opisani postopek omogoča številne možnosti za izboljšave ter nadaljnje raziskave. Sistem bomo še dodatno nadgradili ter ga hkrati skušali poenostaviti. V raziskavo bomo vključili večje število uporabnikov ter na ta način statistično relevantno skušali ugotoviti primerjavo uporabniške izkušnje med klasičnim in predlaganim alternativnim pristopom.

Zahvala

Znanstveno-raziskovalno delo je bilo sofinancirano s strani Ministrstva za visoko šolstvo, znanost in tehnologijo Republike Slovenije.

Literatura

- [1] P. Watzlawick, J. H. Beavin, D. D. Jackson: *Menschliche Kommunikation*. Stuttgart, Wien: Hans Huber Verlag, 1969
- [2] T. Matthew: *Multimodal Human-Computer Interaction, Real-Time Vision for Human-Computer Interaction*, pg. 269-283, Springer US, 2005
- [3] A. Jaimes, N. Sebe: *Multimodal Human Computer Interaction: A Survey*, IEEE International Workshop on Human Computer Interaction; Beijing 2005
- [4] J. Guna, M. Pogačnik, A. Štern, I. Humar, J. Bešter: *Inovativni vmesniki in načini interakcij*, Zbornik osemnajste mednarodne Elektrotehniške in računalniške

konference ERK 2009, 21. - 23. September 2009, Portorož, Slovenija (ERK_Ref)

- [5] P. Barthelmess, S. Oviatt: *Multimodal Interfaces: Combining Interfaces to Accomplish a Single Task*, HCI Beyond the GUI, 2008, Pages 391-444
- [6] S. Oviatt: *Breaking the Robustness Barrier: Recent Progress on the Design of Robust Multimodal Systems*, *Advances in computers*, vol. 56, 2002, Pages: 305-341
- [7] MUD, Wikipedia, <http://en.wikipedia.org/wiki/MUD>
- [8] OCZ NIA, http://www.ocztechnology.com/products/ocz_peripherals/oczn-neural_impulse_actuator
- [9] LindenLabs Second Life, <http://secondlife.com/>
- [10] A. M. Kaplan, M. Haenlein: *The fairyland of Second Life: Virtual social worlds and how to use them*, *Business Horizons* (2009) 52, 563—572
- [11] LindenLabs Second Life izobraževanje, <http://education.secondlife.com/?lang=en-US>
http://secondlife.grid.net.s3.amazonaws.com/docs/SecondLife_Case_OpenU_EN.pdf
- [12] Nintendo Wii, <http://www.nintendo.com/wii>
- [13] J.A. Vicaria, J.M. Maestre, E.F. Camacho: *Academical and research wiimote applications*, IADIS International Conference Interfaces and Human Computer Interaction 2008
- [14] P.J. Standen, C. Camm, S. Battersby, D.J. Brown, M. Harrison: *An evaluation of the Wii Nunchuk as an alternative assistive device for people with intellectual and physical disabilities using switch controlled software*, *Computers & Education*, 2010
- [15] Lee Johnny C.: *Hacking the Nintendo Wii Remote*, *IEEE Pervasive Computing*, Volume 7, Issue 3, 2008