

*Interaktivni in uporabniku prilagodljivi
informacijski vmesniki*

Robert Ravnik

DOKTORSKA DISERTACIJA

PREDANA

FAKULTETI ZA RAČUNALNIŠTVO IN INFORMATIKO

KOT DEL IZPOLNJEVANJA POGOJEV ZA PRIDOBITEV NAZIVA

DOKTOR ZNANOSTI

S PODROČJA

RAČUNALNIŠTVA IN INFORMATIKE



Ljubljana, 2013

IZJAVA O AVTORSTVU

Izjavljam, da sem avtor doktorske disertacije z naslovom Interaktivni in uporabniku prilagodljivi informacijski vmesniki, ki sem jo izdelal samostojno pod vodstvom mentorja in da ta ne vsebuje materiala, ki bi ga kdorkoli predhodno že objavil ali oddal v obravnavo za pridobitev naziva na univerzi ali na drugem visokošolskem zavodu, razen v primerih, kjer so navedeni viri. Soglašam z javno objavo elektronske oblike doktorske disertacije, ki je identična s tiskano obliko doktorske disertacije.

— Robert Ravnik —

november 2013

ODDAJO SO ODOBRILI

dr. Franc Solina

redni profesor za računalništvo in informatiko

MENTOR IN ČLAN OCENJEVALNE KOMISIJE

dr. Franc Jager

redni profesor za računalništvo in informatiko

PREDSEDNIK OCENJEVALNE KOMISIJE

dr. Franjo Pernuš

redni profesor za elektrotehniko

ZUNANJI ČLAN OCENJEVALNE KOMISIJE

Univerza v Ljubljani, Fakulteta za elektrotehniko

dr. Vesna Žabkar

redna profesorica za ekonomijo

ZUNANJA ČLANICA OCENJEVALNE KOMISIJE

Univerza v Ljubljani, Ekonomska fakulteta

PREDHODNA OBJAVA

Izjavljam, da so bili rezultati obravnavane raziskave predhodno objavljeni/sprejeti za objavo v recenzirani reviji ali javno predstavljeni v naslednjih primerih:

- [1] R. Ravnik, F. Solina. Audience measurement of digital signage: Quantitative study in real-world environment using computer vision. *Interacting with Computers*, 25(3):218–228, 2013. doi: [10.1093/iwci/iwso23](https://doi.org/10.1093/iwci/iwso23)
- [2] R. Ravnik, F. Solina. Interactive and audience adaptive digital signage using real-time computer vision. *International Journal of Advanced Robotic Systems*, 10(107):1–7, 2013. doi: [10.5772/55516](https://doi.org/10.5772/55516)
- [3] R. Ravnik, B. Batagelj, B. Kverh and F. Solina. Dynamic anamorphosis as a special, computer generated user interface. *Interacting with Computers*, 2013. doi: [10.1093/iwci/iwt027](https://doi.org/10.1093/iwci/iwt027)

Potrjujem, da sem pridobil pisna dovoljenja vseh lastnikov avtorskih pravic, ki mi dovoljujejo vključitev zgoraj navedenega materiala v pričujočo disertacijo. Potrjujem, da zgoraj navedeni material opisuje rezultate raziskav, izvedenih v času mojega podiplomskega študija na Univerzi v Ljubljani.

POSVETILO

☐ ☐ ▲ ☐ ☐ ☐ ☐ ▲
☐ □ ☐ ☐ < ☐ ☐ ☐
☐ ▲ ☐ ▲ □ ☐ ☐ ▲
☐ L ▲ ☐ ☐ ☐ ▲ ☐
< ☐ ☐ ☐ ▲ ☐ ▲ V
☐ < □ ☐ ☐ ☐ ☐ ▲
☐ L ☐ ☐ ☐ ☐ ▲ ☐
☐ ▲ > □ ☐ ☐ ☐ ▲

POVZETEK

V doktorskem delu razvijemo interaktivni in uporabniku prilagodljiv informacijski vmesnik na osnovi povezave metod računalniškega vida in strojnega učenja. Pri tem uporabimo metodo za segmentacijo slike na osnovi Gaussovih mešanice, algoritem za detekcijo obrazov Viola & Jones, metodo za poravnavo obrazov z modeli aktivnega videza ter algoritem POSIT za oceno orientacije obrazov. Nabor metod se izkaže za ključnega pri razmerju med natančnostjo in časovno zahtevnostjo, saj je vmesnik zasnovan za delovanje v realnem času na običajni strojni opremi. Razvit vmesnik uporabimo za študijo treh izbranih problemov: razvoj dinamične anamorfoze, izvedbo kvantitativne študije gledanosti v realnem okolju in modeliranje procesa nakupnega odločanja. V prvem sklopu raziskav z implementacijo vmesnika za digitalno anamorfozo pokažemo novo interaktivno metodo, ki se z anamorfčno transformacijo slike v realnem času prilagaja položaju uporabnika v prostoru. Eksperimentalno z vmesnikom potrdimo, da je koncept uporaben za izboljšanje očesnega stika pri videokonferencah. V drugem sklopu študija gledanosti kvantitativno ovrednoti pozornost uporabnikov in pokaže statistično značilne razlike vezane na demografske in predvajalne parametre. Posebej zanimivo se pokaže konverzija med stopnjami gledanosti. V tretjem sklopu raziskav pa z razvitim vmesnikom modeliramo nakupni proces. Gre za interdisciplinarno raziskavo, kjer z uporabo strojnega učenja in razvitega vmesnika analiziramo in modeliramo nakupno odločitev in vloge v nakupnem procesu. Splošneje razvit vmesnik predstavlja doprinos na področju interakcija človek-računalnik (HCI) in kaže zanimive možnosti, tako znanstvene kot aplikativne rabe, kot so odprt problem uporabniške slepote, razvoj novih interaktivnih metod za podajanje vsebin in kvantitativna analiza obnašanja uporabnikov.

Ključne besede: informacijski vmesniki, metode interaktivnosti, analiza gledanosti

ABSTRACT

This doctoral thesis develops an interactive and user-adaptive information interface based on computer vision and machine learning methods. The selected methodology is chosen to optimise the balance between accuracy and time complexity, subjected to design-aim to perform in real-time and using conventional hardware. More particularly, we use the Mixture of Gaussians method for background segmentation, Viola & Jones method for face detection algorithm, Active Appearance Models for face alignment and POSIT algorithm for head pose estimation. The developed interface is used as the key research tool to explore three currently open problems in the field of human-computer interaction: dynamic anamorphosis, quantitative audience measurement study of digital signage in real-world environment, and modeling of the purchase decision process. In the first study, we develop a new interactive method -the dynamic anamorphosis- that in real time, adapts the projected image to the observer's position by using anamorphic transformation. As an interesting application, we show that dynamic anamorphosis notably improves eye-contact in videoconferencing. In the second study, we use the developed interface to perform a quantitative audience measurement field study, which evaluates the user attention. The study shows significant attentional differences related to demographic and broadcasting features; especially interesting are conversion rates of the attention phases. In the third study, the interface is applied to model the purchase decision process, which is an interdisciplinary study, where data collected with the developed interface and subjected to machine learning are combined to model and analyze the decision and roles in a purchasing process. Finally, more generally, the developed system presents a contribution to the field of human-computer interaction and shows further possibilities for scientific use and applications, such as open problem of display blindness, development of new interactive methods for broadcasting of relevant content, and quantitative analysis of user behavior.

Key words: information interfaces, interaction design, audience measurement

ZAHVALA

Zahvaljujem se mentorju prof. dr. Francu Solini za vodenje in usmerjanje v času doktorskega študija. Zahvaljujem se tudi sodelavcem iz Laboratorija za računalniški vid FRI UL, posebno Bojanu Kverhu za njegovo delo in nasvete pri anamorfični transformaciji. Za pripombe in komentarje, ki so pomagali izboljšati disertacijo, se zahvaljujem prof. dr. Francu Jagru, prof. dr. Franju Pernušu in prof. dr. Vesni Žabkar. Slednji še dodatna hvala za vzpodbudo in usmerjanje pri raziskavah modeliranja nakupnega procesa.

Posebna zahvala za vsestransko podporo gre moji družini. Miha, hvala za vse nasvete in pomoč. Urša, hvala za neskončno podporo in razumevanje v času nastajanja tega dela.

KAZALO

<i>Povzetek</i>	<i>i</i>
<i>Abstract</i>	<i>iii</i>
<i>1 Uvod</i>	<i>1</i>
1.1 Uvod	2
1.2 Motivacija	3
1.3 Prispevki k znanosti	3
1.4 Pregled doktorske disertacije	4
<i>2 Osnovni koncepti in ozadje</i>	<i>5</i>
2.1 Informacijski vmesniki	6
2.2 Pregled področja	9
2.3 Računalniški vid	14
2.3.1 Segmentacija ozadja	14
2.3.2 Detekcija, poravnava in ocenjevanje orientacije obraznih slik .	15
2.3.3 Analiza demografskih značilnic	16
<i>3 Metodologija - Adaptivni informacijski vmesnik</i>	<i>19</i>
3.1 Arhitektura adaptivnega informacijskega vmesnika	20
3.2 Prostorsko-časovne značilnice uporabnika	22
3.2.1 Segmentacija ozadja in sledenje aktivnih regij	24
3.2.2 Detekcija obraza	26
3.2.3 Poravnava in ocena 3D orientacije obraza	28
3.2.4 Ocena 3D položaja uporabnika	31

3.3	Modeliranje demografskih značilnic	34
3.4	Kontekstno-odvisna izbira predvajane vsebine	38
3.5	Pravni vidiki in zasebnost	39
4	<i>Dinamična anamorfoza</i>	41
4.1	Anamorfoza	42
4.2	Dinamična anamorfoza	44
4.3	Aplikacija - Izboljšanje očesnega stika pri videokonferencah	48
4.4	Aplikacija - Interaktivna instalacija	52
5	<i>Merjenje in kvantitativna analiza gledanosti</i>	55
5.1	Eksperiment v realnem okolju	56
5.2	Meritve in verifikacija rezultatov	58
5.3	Rezultati analize gledanosti	59
5.3.1	Čas prisotnosti	60
5.3.2	Čas kontaktne možnosti	62
5.3.3	Čas pozornosti	64
5.4	Stopnje gledanosti in konverzija pozornosti med uporabniki	66
5.5	Informacijska analiza gledanosti	70
5.5.1	Graf interakcij med značilnicami	71
5.5.2	Analiza interakcij med značilnicami	72
6	<i>Modeliranje procesa nakupnega odločanja</i>	75
6.1	Proces nakupnega odločanja	76
6.2	Modeliranje procesa nakupnega odločanja	77
6.3	Modeliranje nakupnega procesa nakupa s strojnim učenjem	86
7	<i>Zaključek</i>	91
	<i>Literatura</i>	95

Uvod

1.1 Uvod

Razvoj informacijske tehnologije in strojne opreme danes stalno ustvarjata nove načine podajanja in sprejemanja informacij, in osrednji del tega procesa so informacijski vmesniki. Po ACM klasifikaciji za področje računalništva informacijski vmesniki zajemajo področja multimedijskih informacijskih sistemov, metode interakcije in uporabniške vmesnike [2]. Pod izrazom informacijski vmesnik tako razumemo sistem, ki uporabniku preko informacijske naprave, običajno zaslona, omogoča dostop do vira informacij. V doktorski disertaciji se bomo posebej osredotočili na digitalne zaslone, ki se nahajajo na javnih mestih, in omogočajo prikaz multimedijskih vsebin, ki se na predvavalno mesto običajno prenašajo preko omrežja [52, 59, 77, 91, 111, 162].

Uporaba digitalnih zaslonov v pretočnih okoljih omogoča dinamičen in interaktiven način podajanja informacij uporabniku [17, 126, 188]. Vendar potencial prikazovanja relevantnih vsebin v realnem okolju pogosto ni izpolnjen, kar pripelje do t.i. efekta uporabniške slepote (angl. *display blindness*), ko se na zaslonu prikazujejo generične in za uporabnike nezanimive vsebine [131]. Predlagani so bili številni izboljšani načini komunikacije med uporabnikom in digitalnimi zasloni, ki izkoriščajo različne modalnosti in dodatne naprave [43, 57, 72, 155]. Poleg dobro sprejetih načinov interakcije [122, 183] pa pogosto uporabniki niti ne prepoznajo možnosti interakcije, ki jim jo zaslone ponujajo [130]. Pojav se imenuje interakcijska slepota (angl. *interaction blindness*) in kaže na problem, kako učinkovito vzbuditi zanimanje uporabnikov [136].

Informacijski vmesnik naj bi bil primarno namenjen podajanju informacij, vendar je ta proces podajanja ključno odvisen od značilnic uporabnika, torej ali je informacija zanj primerna ali ne. Da bi naredili takšne sisteme čim bolj uporabne in učinkovite, se uporablja k uporabniku usmerjeno načrtovanje (angl. *user-centered design*), kjer se prikazovane vsebine prilagajajo trenutnemu uporabniku in kontekstu predvavalnega mesta [11, 29, 119, 160, 161, 184]. Integracija prilagojenega prikazovanja vsebin z mehanizmom vseprisotnega računalništva vodi do razvoja t.i. ambientalne inteligence (angl. *ambient intelligence*) [1, 128].

1.2 Motivacija

Raziskovalno področje informacijskih vmesnikov za podajanje vsebin v urbanih okoljih je mlado in se danes hitro razvija [17, 43, 57, 72, 97]. Predlagani so številni novi koncepti informacijskih vmesnikov, ki bi omogočali bolj učinkovito podajanje vsebine. Velik problem obstoječih znanstvenih študij in obstoječih konceptov načrtovanja informacijskih vmesnikov je, da vsi uporabljajo striktno kvalitativne podatke za modeliranje vedenjske analize uporabnikov [81, 131], čeprav je to ključno za določitev učinkovitosti in uspešnosti vmesnika. Tako se tipično uspešnost sistema meri preko podatkov, ki so bili pridobljeni s predhodnimi ali naknadnimi intervjuji uporabnikov in so kot taki tipično močno subjektivni in zato neprimerni.

V doktorski disertaciji bomo naslovili vprašanje, kako zasnovati adaptivni informacijski vmesnik, ki bo strojno z uporabo kamere in programsko z uporabo metod računalniškega vida spremljal dejansko dogajanje na predvajanem mestu in hkrati temu primerno podajal informacijske vsebine, kar je nov in originalen pristop k razvoju informacijskih vmesnikov, ki inherentno že v samem konceptu zaobide danes vroče probleme npr. uporabniške in interakcijske slepote. Tovrstni pristop omogoči tako interaktivno in kontekstno-odvisno prikazovanje vsebin, kot tudi zajem kvantitativnih podatkov obnašanja uporabnikov pred zaslonom. Na osnovi zajetih podatkov statistično analiziramo značilnice gledanosti, kvantitativno ovrednotimo efekt uporabniške slepote in poiščemo statistično značilne vedenjske vzorce. Modeliranje obnašanja uporabnikov povežemo z metodami strojnega učenja ter tako omogočimo raziskave interdisciplinarnih problemov, kot je ekonomsko pomemben proces nakupnega odločanja.

1.3 Prispevki k znanosti

Poglavitni prispevki doktorske disertacije k znanosti so:

- RAZVOJ INFORMACIJSKEGA VMESNIKA ZA INTERAKTIVNO IN ADAPTIVNO PRIKAZOVANJE VSEBIN V REALNEM ČASU; opišemo metodologijo in implementiramo sistem, ki z uporabo metod računalniškega vida omogoča prilagojeno prikazovanje vsebin glede na časovne, prostorske in demografske značilnice uporabnika. Omogoča tudi kvantitativno statistično vrednotenje gledanosti prikazane vsebine.

- **DINAMIČNA ANAMORFOZA**; uvedemo novo metodo prostorske interakcije z uporabnikom, ki razširja koncept perspektivne anamorfične transformacije z realnočasnim spremljanjem uporabnika in prilagajanjem vsebine. Pokažemo tudi možno uporabo v videokonferencah.
- **KVANTITATIVNA ANALIZA GLEDANOSTI ZASLONOV V REALNEM OKOLJU**; z uporabo razvitega vmesnika izvedemo prvo tovrstno študijo, ki ovrednoti dejansko obnašanje ljudi in pozornost, ki jo ljudje namenijo tovrstnim zaslonom na javnih površinah. Ponudimo tudi možno razlago opaženih vedenjskih vzorcev na osnovi informacijske analize.
- **MODELIRANJE PROCESA NAKUPNEGA ODLOČANJA**; na osnovi podatkov zajetih pri analizi gledanosti v realnem okolju vzpostavimo model strojnega učenja, ki opiše ekonomsko pomemben nakupni proces in vloge nakupnega odločanja.

1.4 Pregled doktorske disertacije

Doktorska disertacija je organizirana naslednje. V poglavju 2 podamo pregled literature, opredelimo širše raziskovalno področje in izpostavimo danes odprte probleme na področju interakcija človek-računalnik. Poglavje 3 predstavi metodologijo razvoja predlaganega informacijskega vmesnika, kjer za modeliranje časovnih, prostorskih in demografskih značilnic uporabimo povezane metode računalniškega vida in strojnega učenja. V poglavju 4 opišemo metodo prostorske interakcije na osnovi dinamične anamorfične transformacije in njeno možno aplikacijo za izboljšavo videokonferenčne uporabniške izkušnje. Poglavje 5 predstavi rezultate kvantitativne študije gledanosti v realnem okolju. Statistično značilne vedenjske vzorce, ki se pokažejo iz eksperimentalno zajetih podatkov, umestimo v obstoječe raziskave na področju, in na osnovi analize interakcij med značilnicami predlagamo njihovo razlago. Poglavje 6 predstavi interdisciplinarni problem modeliranja procesa nakupnega odločanja in predstavi predlagano rešitev na osnovi zajetih podatkov gledanosti in strojnega učenja. Poglavje 7 poda ključne doktorske disertacije in predlaga izhodišča za nadaljnje raziskave.

Osnovni koncepti in ozadje

V tem poglavju najprej opredelimo področje področje informacijskih vmesnikov in njegovo vpetost v področje interakcije človek-računalnik. Sledi pregled *state-of-the-art* raziskav na področju adaptivnih in interaktivnih informacijski vmesnikov. Zadnji razdelek poglavja pa umesti metode računalniškega vida, povezane z obravnavano problematiko informacijskih vmesnikov.

2.1 Informacijski vmesniki

Raziskovalno področje informacijskih vmesnikov sodi v širše področje raziskav interakcije človek-računalnik (angl. *human-computer interaction*). Po definiciji, ki jo navaja ACM kurikulum, je interakcija človek-računalnik veda, ki se ukvarja z načrtovanjem, evaluacijo in implementacijo interaktivnih računalniških sistemov namenjenim za človekovo uporabo ter s študijami pojavov, ki spremljajo te sisteme [75]. Gre za vedo, ki predstavlja presečišče med računalniškimi vedami, kot so umetna inteligenca, vseprisotno računalništvo in računalniški vid, z drugimi znanstvenimi disciplinami, med njimi psihologijo, kognitivnimi znanostmi in ergonomijo [169, 178]. Shematski prikaz znanstvenih disciplin, ki skupaj interdisciplinarno tvorijo področje interakcija človek-računalnik (HCI) je podan na Sliki 2.1.

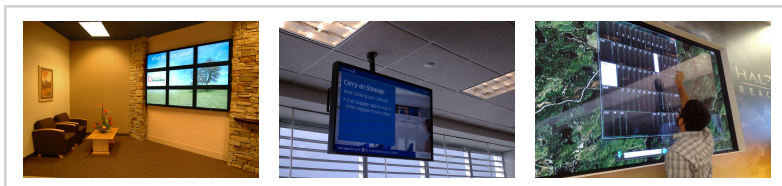


Slika 2.1

Umestitev interakcije človek-računalnik (HCI) med povezana področja. Gre za interdisciplinarno področje, ki povezuje računalniške znanosti z kognitivni znanostmi, psihologijo, ergonomijo, grafičnim načrtovanjem in drugimi vedami.

Informacijski vmesnik je sistem, ki uporabniku preko informacijske naprave omogoča dostop do vira informacij. Med informacijske vmesnike tako neposredno sodijo multimedijski informacijski sistemi ter načrtovanje uporabniških vmesnikov in načinov manipulacije [2]. V nadaljevanju bomo definicijo omejili na interaktivne multi-

medijske sisteme, ki se nahajajo na javnih in pretočnih površinah in podajajo informacijo preko digitalnih zaslonov [52, 91, 95, 111, 162]. Danes jih srečamo na letališčih, avtobusnih in železniških postajah, bolnišnicah, nakupovalnih središčih, športnih dvoranh, poslovnih prostorih in drugih javnih površinah. Rast števila tovrstnih sistemov se danes stalno povečuje [11] in naj bi po napovedih podjetja Intel leta 2015 zajemala 22 milijonov aktivnih digitalnih zaslonov za prikazovanje vsebin [192]. Slika 2.2 prikazuje nekatere tipične primere uporabe informacijskih vmesnikov na javnih površinah.



Slika 2.2

Informacijski vmesniki za prikazovanje vsebine so vedno bolj pogosti v urbanem okolju. Običajno jih srečamo na javnih in pretočnih površinah.

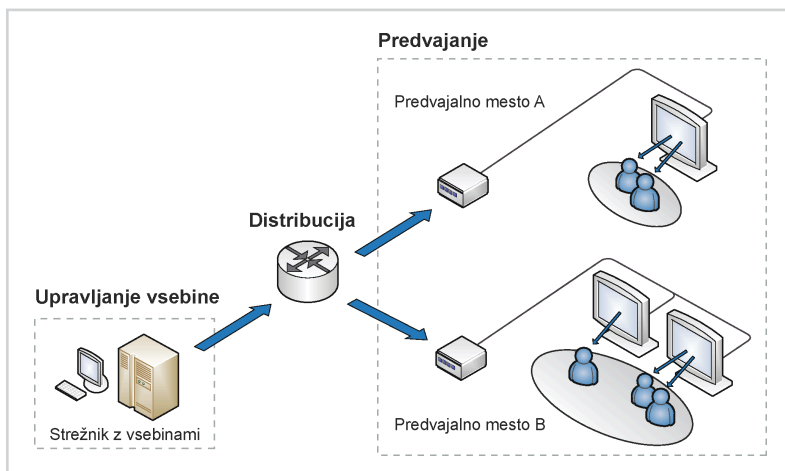
Zasloni za predvajanje multimedijskih vsebin so običajno del bolj razvejanega sistema. Slika 2.3 prikazuje tipično arhitekturo celotnega informacijskega vmesnika, ki zajema tri osnovne sklope [111, 162]:

- **UPRAVLJANJE VSEBINE**; se običajno izvaja centralizirano in poteka na strežniškem delu sistema. Zajema pripravo in ustrezno obdelavo multimedijskih vsebin, katere sistem pred distribucijo prevede v ustrezne formate. Upravljalec na osnovi želenih ciljev sestavi predvajalni seznam, vezan na specifično predvajalno mesto.
- **DISTRIBUCIJA** vsebin in predvajalnega seznama poteka od osrednjega strežnika z vsebinami do predvajalnega mesta. Prenos se lahko izvaja preko lokalnega ali prostranega računalniškega omrežja (angl. *wide area network*). Prenesejo se le tiste vsebine, katere so na aktualnem predvajalnem seznamu, a še niso v lokalni podatkovni shrambi predvajalnega mesta.
- **PREDVAJANJE** vsebine na samem predvajalnem mestu izvaja predvajalnik. Nabor vsebine se izbere glede na aktualni predvajalni seznam. Vsebine se prikazujejo izmenično, vsaka z lastno frekvenco, ki jo določa število ponovitev znotraj izbranega časovnega intervala.

Na sliki 2.3 ključno opazimo, da tok informacij poteka enosmerno, in sicer od centra za upravljanje vsebine do predvajalnika na predvajalnem mestu.

Slika 2.3

Arhitektura tipičnega informacijskega vmesnika za predvajanje vsebin. Vsebine se hranijo na centralnem strežniku, s katerega se preko omrežja skupaj s predvajalnim seznamom distribuirajo na predvajalno mesto. Modre puščice označujejo tok informacij, ki pomembno poteka enosmerno od centralnega strežnika do predvajalnega mesta.



Predstavljena arhitektura informacijskega vmesnika omogoča doseganje ciljev na podlagi upravljanja predvajalnih značilnic (angl. *broadcasting features*) *BF*, ki omogočajo prilagajanje predvajanja vsebine glede na [111]:

- nabor izbranih vsebin (vezan na aktualni predvajalni seznam),
- zaporedje in frekvenco izbranih vsebin,
- časovni raspored prikazovanja (vezan na uro, dan v tednu, mesec oziroma določen datum),
- zgodovino predvajanja (minimalni oziroma maksimalni čas od zadnje pojavitve vsebine),
- organizacijo zaslonov v skupine (vezano na njihovo geografsko lokacijo in položaj).

Predvajalne značilnice določi upravljalec glede na pričakovan pretok občinstva, ki gre v nekem časovnem intervalu (npr. dnevu) mimo predvajalnega mesta. Pričakovan pretok je običajno posplošena ocena števila ljudi, vezana na širše področje okoli predvajalnega mesta (npr. dnevno število obiskovalcev celotnega nakupovalnega središča).

2.2 Pregled področja

Začetki raziskav področja informacijskih vmesnikov segajo v leto 2004, ko Churchill in sodelavci predstavijo sistem digitalnih panelov postavljenih na različnih mestih znotraj delovne organizacije. Sistem so poimenovali *Plasma Posters*. Namenjen je bil odprtemu in neformalnemu deljenju informacij in multimedijskih vsebin med zaposlenimi znotraj delovnih skupin, kot tudi med večjimi organizacijskimi enotami. Uporabniki so sistem sprejeli kot dodaten komunikacijski kanal poleg spleta in elektronske pošte [30].

Müller in sodelavci predstavijo podoben sistem, postavljen v univerzitetno okolje. Z več vrstami namenskih zaslonov naslovijo zahteve in pričakovanja različnih uporabnikov. Izpostavijo tudi problem aktualnosti in ažurnosti vsebine. Uporabnik bo namenil zaslonu več časa v kolikor pričakuje, da mu bo le-ta posredoval nove informacije. Kot ključni parameter, da uporabniki ohranijo pozornost in prepoznajo uporabnost sistema, izpostavijo čas posodabljanja informacij [129].

Da informacijski vmesniki niso zgolj domena urbanega okolja pokažejo Taylor in sodelavci [182]. S sistemom digitalnih zaslonov, ki se nahajajo v ruralnem okolju pokažejo, da lahko tovrstni sistemi služijo pri spodbujanju lokalne identitete in povezovanju lokalne skupnosti [183]. Cheverst s sodelavci obravnava vpliv družbenega in organizacijskega konteksta na postavljen sistem zaslonov [29]. Primerjava dveh sistemov, postavljenih na različnih lokacijah in okolju, pokaže različne vedenjske vzorce uporabnikov.

Davies s sodelavci predlaga možnost združevanja omrežij digitalnih zaslonov v odprte mreže zaslonov [43]. Z uveljavitvijo odprte arhitekture bi informacijski vmesniki za predvajanje uporabljali različne, decentralizirane vire vsebine (ne zgolj svojega) ter uporabnikom omogočali podporne storitve [120]. Predlagana je bila tudi že virtualizacijska aplikacijska platforma, ki bi omogočila neodvisen razvoj aplikacij in storitev za velika omrežja digitalnih zaslonov [108].

José in sodelavci se ukvarjajo z udeležbo, lastništvom in nadzorom vsebine v okviru omrežij javnih zaslonov, na katere lahko uporabniki na daljavo dodajajo vsebino [86, 87]. Predstavijo dve različni paradigmi objavljanja vsebin in kako le-te vplivata na vsebino, ki so jo ljudje pripravljene objaviti in dojemanje predvajane informacije. Izpostavijo tudi razlike med upravljalcem in uporabniki sistema pri objavljanju vsebin [77].

Pozornost, ki jo uporabniki namenijo digitalnim zaslonom, je eden ključnih parametrov informacijskih vmesnikov. Primerjalna študija, ki so jo izvedli Huang in sodelavci

[81] pokaže, da gre za kompleksen proces, ki je odvisen od več dejavnikov, med njimi: prostorska umestitev zaslona, velikost zaslona, format predvajane vsebine ter njena sporočilna vrednost. Uporabniki se zavedajo prednosti in možnosti, ki jih za podajanje informacij ponujajo digitalni zasloni in si želijo relevantnih vsebin [81, 131]. Pogosto neizpolnjena pričakovanja ob predvajanju generičnih vsebin pa vodijo do ignoriranja digitalnih zaslonov, kar opisuje pojav uporabniške slepote [121, 131].

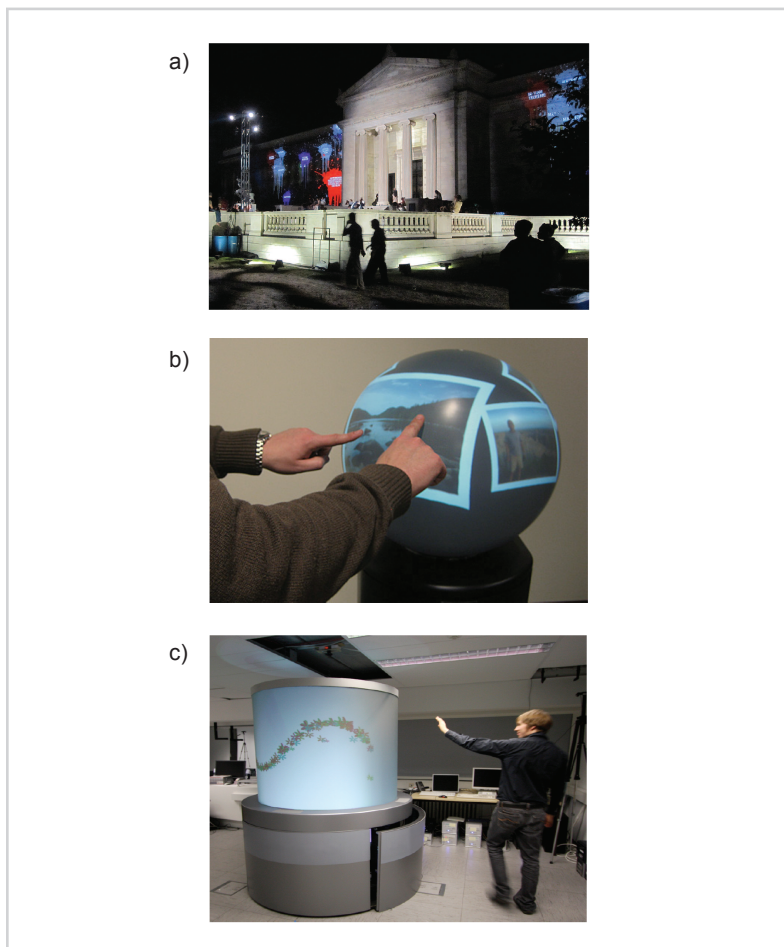
Številne raziskave na področju informacijskih vmesnikov se ukvarjajo z načrtovanjem in razvojem interaktivnega podajanja vsebine na javnih površinah [86, 95]. Tako so se predlagali različni vmesniki za interakcijo z uporabniki, na primer uporabniški vmesniki, ki temeljijo na prepoznavi gibov z rokami [28, 73]. Ren in sodelavci predstavijo uporabo gibov z rokami kot možnost manipulacije 3D vizualizacije [155]. Izvedena je bila uporabniška študija v kateri Hespanhol in sodelavci pokažejo, kako nabor in predvidena kompleksnost gibov z rokami vplivata na uporabniško izkušnjo [74]. Načrtovanje interakcije mora biti primerno zadani nalogi informacijskega vmesnika.

Za interakcijo lahko uporabimo tudi druge modalnosti, kot npr. smer pogleda [205], položaj telesa in dotik [126]. Ostkamp in sodelavci uvedejo koncept *vizualnih podarkov*, s katerimi dosežejo večjo učinkovitost podajanja informacij [139]. Predlagani so bili tudi načini interakcije preko mobilnih naprav [17, 127]. Fisher in sodelavci so koncept informacijskega vmesnika uporabili za razvoj interaktivnega multimedijskega pročelja stavbe [57] s katerim lahko uporabniki manipulirajo preko mobilnega telefona (glej Sliko 2.4a).

Primer informacijskega vmesnika, ki združuje uporabo večdotične interakcije in 3D stereoskopske vizualizacije predstavijo Hatchet in sodelavci [72]. Uporabnik preko dotika manipulira z vsebino na 2D površini, ki se na stereoskopskem zaslonu vizualizira v plavajočih 3D objektih. Za delovanje mora uporabnik nositi stereo očala, ki omogočajo globinsko uporabniško izkušnjo obogatene resničnosti [71].

Raziskave zajemajo tudi informacijske vmesnike, ki za podajanje in manipulacijo informacij uporabljajo neravne zaslone [81]. Beyer in sodelavci razvijejo sistem, ki temelji na uporabi valjastega zaslona [14]. Uporabniška študija pokaže, da valjasti zasloni pokažejo specifično obnašanje uporabnikov [15], ki imajo možnost interakcije s premikanjem po prostoru okoli njih (glej Sliko 2.4c). Lin in sodelavci predstavijo interaktiven cevast zaslon [107] na osnovi uporabe projektorja in konveksnega ogledala. Benko in sodelavci predstavijo sistem *Sphere*, sferični zaslon, ki omogoča večdotično interakcijo [13] in obravnava značilne lastnosti (uporabnik ne vidi celotnega zaslona,

vmesnik nima glavnega uporabniškega položaja in orientacije, gladek prehod med vertikalnimi in horizontalnimi površinami), ki ga razlikujejo od ostalih vmesnikov (glej Sliko 2.4b).



Slika 2.4

Sorodne raziskave na področju informacijskih vmesnikov. a) Multimedijско pročelje s katerim uporabniki manipulirajo preko mobilnega telefona in pošiljanjem kratkih besedilnih sporočil. b) Večdotični sferični zaslon osnovi infrardeče kamere, ki omogoča več-uporabniško interakcijo. c) Valjast zaslon omogoča uporabnikom prostorsko interakcijo. Slike so povzete po [13, 14, 57].

Sorodne raziskave predstavi Koppel s sodelavci, ki preučuje vpliv kombinacije povezanih zaslonov na uporabniško izkušnjo. Skupino šestih zaslonov postavijo v različne

konfiguracije: frontalno, konkavno in heksagonalno postavitev. Pojavijo se značilni vedenjski vzorci, ki pokažejo, kako uporabnik opazi postavitev zaslonov, razvije motivacijo in nadaljnjo interakcijo [90].

Müller in sodelavci predstavijo ugotovitve eksperimentalne študije, kako mimoidoči opazijo in se odzovejo na interaktivno vsebino javnih zaslonov v izložbah [130]. Za komunikacijo z uporabniki zasnujejo interaktiven sistem, ki za povratno informacijo uporablja vizualizacijo na osnovi slučajnega gibanja mimoidočih. Za predstavitev uporabnika uporabijo: zrcalno sliko uporabnika, obris njegove podobe in predstavitev brez slike. Študija pokaže, da je predstavitev z zrcalno sliko k interakciji privabila 90% več uporabnikov kot konvencionalna uporaba napisanih navodil.

Pomanjkanje splošno sprejetega modela za vrednotenje digitalnih sistemov za prikazovanje vsebine in različnih ciljev, ki jim te sistemi sledijo (npr. optimizacija časa pozornosti in interakcije uporabnikov, izbira čim boljše metode interakcije, idr.) sta privedla do kategorizacije vrednotenja informacijskih vmesnikov [7]. Alt in sodelavci opišejo 7 kategorij, katere v celoti opredelijo delovanje tovrstnih sistemov: uporabniška izkušnja, obnašanje uporabnikov, sprejemljivost za uporabnika, storilnost, učinkovitost sistema, vprašanje zasebnosti in vpliv na družbo [7].

Michelis izpostavi problem primerjave različnih interaktivnih sistemov. V ta namen predlaga ogrodje za generalizacijo interakcije uporabnikov [122], ki ga imenuje *Audience Funnel*. Ogrodje definira 6 stopenj uporabniške interakcije in je namenjeno sistematični primerjavi različnih uporabniških študij interaktivnih digitalnih zaslonov.

Sorodno pojavu uporabniške slepote, primerjalne študije interaktivnih zaslonov pokažejo na nov pojav t.i. interakcijske slepote. Ta se navezuje na dejstvo, da uporabniki pogosto ne opazijo interaktivnih možnosti, ki jih zaslon ponuja [136]. Ojala in sodelavci pokažejo, da uporaba nekonvencionalnega predmeta (angl. *curiosity object*), ki v ljudeh vzbudi radovednost, poveča stopnjo interaktivnosti, kot tudi povzroči spremembo gibanja ljudi v okolici zaslona [136].

Sklop raziskav se nanaša tudi na kontekstualno-odvisne sisteme, ki omogočajo adaptivno predvajanje vsebine. Kriterij za prilagajanje je vezan na specifične lastnosti predvajalnega mesta, lahko pa tudi samega uporabnika [82]. Predlagan je bil sistem *Mobi-DiC*, ki se spremlja gledanost uporabnikov preko kuponov [128]. Na podlagi odziva se prilagodi strategija predvajanja glede na okolje predvajalnega mesta. Satoh predlaga platformo za kontekstno-odvisno predvajanje vsebin [160], ki opredeli razmerje med ljudmi, virtualnimi vsebinami in fizičnim prostorom predvajanja [161]. Vsebine in-

formacijskih vmesnikov se tako lahko prilagajajo lokalnemu okolju ali pa naslavljajo globalno publiko [6].

Sorodni koncepti adaptivnega predvajanja vsebine se razvijajo tudi pri televizijskem predvajanju. Mei in sodelavci predlagajo sistem za kontekstualno prekrivanje televizijskih vsebin [119], ki samodejno poišče nabor regij videa in določi časovni okvir, v katerem jih lahko prekrije z oglaševalsko vsebino. Na osnovi analize strukture video vsebine, detekcije obrazov in besedila ter vizualnih poudarkov izbere regijo, ki minimizira invazivnost dodane vsebine. Teixeira in sodelavci predstavijo koncept kontekstualne interakcije preko televizije, ki modelira uporabnika na osnovi množice akcij daljinskega upravljalnika [184].

Raziskave informacijskih vmesnikov vključujejo tudi področje trženja za namene poslovne inteligence [24]. Študija, ki so jo med obiskovalci nakupovalnih središč izvedli Newman in sodelavci pokaže, da digitalni sistemi za prikazovanje vsebin v trgovinah ustvarjajo t.i. ugodno *nakupovalno klimo* (angl. mall atmospherics), ki posredno vpliva na nakupne odločitve [48, 133]. V nakupovalnih središčih se pojavlja vedno več sistemov, ki s pomočjo tehnologije omogočajo pomoč pri nakupni odločitvi [141, 142]. Uporabnikom preko tehnoloških rešitev ponujajo dodatne informacije o izdelkih na samem prodajnem mestu, kar izboljša nakupovalno izkušnjo [47, 143] in vpliva na nakupovalno vedenje [5]. Nakupovalci se najbolj odzovejo na vsebino, ki se navezuje na njihovo trenutno aktivnost in potrebo. Relevantno sporočilo ob ustreznem času vpliva na povečanje prodaje [25]. Spremljanje dogajanja v nakupovalnih središčih z namenom kategorizacije obiskovalcev in personaliziranega predvajanja vsebin pa mora upoštevati pričakovana etična merila [49].

Pregled področja zaključimo z raziskavami, ki se nanašajo na ustrezno upravljanje z zasebnostjo adaptivnih sistemov za prikazovanje vsebine [21, 101, 172]. Nadzor nad zasebnostjo je ključen za zagotavljanje zaupanja med uporabniki in novim medijem, ki omogoča vedno več personalizirane, kontekstno-odvisne vsebine in interaktivnosti. Geiger predstavi model za upravljanje zasebnosti informacijskih vmesnikov, kjer izpostavi, da mora sistem zagotavljanja zasebnosti preseči možnost posredne identifikacije uporabnika [63].

2.3 Računalniški vid

Računalniški vid je področje, ki vključuje metode za zajem, procesiranje, analizo in razumevanje slik [58, 163, 181]. V splošnem govorimo o visoko dimenzijskih podatkih realnega sveta, iz katerih želimo razbrati numerične oziroma simbolne informacije za izvajanje nadaljnjih odločitev [176]. Področje temelji na razvoju metod strojnega zaznavanja, ki modelirajo sposobnosti človeškega vida [163]. Razumevanje slike tipično predstavlja ekstrahiranje simbolnih informacij iz slikovnih podatkov z uporabo modelov, ki temeljijo na načelih geometrije, fizike, statistike in teorije učenja [58].

Računalniški vid se uporablja pri reševanju različnih problemov, kot so: avtomatizacija in integracija nadzornih procesov, detekcija dogodkov, modeliranje objektov in okolja, razvoj uporabniških vmesnikov za interaktivne sisteme, navigacija avtonomnih vozil in avtomatski pregled v proizvodnji [181]. Področje računalniškega vida je široko, zato se bomo v nadaljevanju omejili na predstavitve in pregled tistih problemov in metod računalniškega vida, ki so relevantni za interaktivne in adaptivne informacijske vmesnike [54]. Ključno v vmesniku bo realnočasno in adaptivno modeliranje uporabnika, kar bo osrednje vodilo pri izboru metod za: (i) segmentacijo ozadja, (ii) detekcijo, poravnavo in ocenjevanje orientacije obraznih slik ter (iii) analizo demografskih značilnic, kot sta spol in starost.

2.3.1 Segmentacija ozadja

Segmentacija ozadja zaporedja slik se uporablja za razlikovanje med regijami slike, na katerih se nahajajo objekti in na katerih je ozadje. Regije objektov in ozadja so med seboj disjunktno in skupaj tvorijo pokrite slike. Najenostavnejši način za modeliranje ozadja je, da zajamemo sliko ozadja, ki ne zajema nobenega objekta. Vendar to ni mogoče v vseh situacijah, kjer lahko pride do prekrivanja objektov, spremembe svetlobnih pogojev oziroma tudi sprememb samega ozadja [19].

Wren in sodelavci predstavijo metodo, ki modelira vsako točko ozadja neodvisno od ostalih. Model temelji na tem, da v vsaki točki išče idealno prileganje Gaussove funkcije gostote verjetnosti za zadnjih n vrednosti [193]. Pri uporabi časovnega povprečja lahko uporabimo tudi druge pristope [146]. Za modeliranje ozadja lahko uporabimo mediano zadnjih n vrednosti, ki daje stabilne rezultate tudi pri pomanjšani velikosti osnovne slike [38]. Stauffer in Grimson predlagata metodo, ki vsako točko modelira kot mešanico Gaussovih funkcij (angl. *Mixture of Gaussians*) in omogoča stalno apro-

ksimacijo za posodobljanja modela [174]. Zivkovic predlaga dinamično prilagajanje števila komponent posamezni točki slike, kar izboljša natančnost in zmanjša računsko zahtevnost [207, 208]. Predlagani so bili tudi neparametrični modeli, ki temeljijo na aproksimaciji gostote porazdelitve verjetnosti z jedri (angl. *kernel density approximation*) [53] in pristopi z uporabo teorije grafov [168]. Comanicu in Meer predlagata metodo za segmentacijo slik, ki temelji na *mean-shift* algoritmu [35]. Predlagane so bile tudi metode, ki izkoristijo prostorsko lokalnost variacije slik in za osnovni element ozadja namesto točke uporabijo blok velikosti $n \times n$ točk [165]. Oliver in sodelavci predlagajo modeliranje ozadja na osnovi razcepa z lastnimi vrednostmi [137]. Maddalena in Petrosino opišeta metodo za segmentacijo ozadja z uporabo umetnih nevronskih mrež [113]. Bolj podroben pregled metod je podan v [19, 23, 146].

2.3.2 Detekcija, poravnava in ocenjevanje orientacije obraznih slik

Iz slike ospredja (angl. *foreground*) želimo identificirati množico regij, ki predstavljajo ljudi oziroma njihove obraze. Detekcija mora upoštevati spremembe položaja v prostoru, orientacijo in svetlobne pogoje. Gre za zahteven problem, saj so obrazi ne-togi (angl. *non-rigid*) in imajo na slikah visoko stopnjo variabilnosti glede na velikost, obliko, barvo in teksturo [195].

Predlagani so bili številni pristopi za detekcijo obrazov [79, 93, 203] na osnovi obraznih značilnic [200], barve kože [37, 118] in teksture [40]. Liu predstavi metodo za frontalno detekcijo obrazov, ki temelji na Bayesovih diskriminantnih funkcijah in modelira klasifikator obraza z multivariantno normalno porazdelitvijo [109]. Feraud in sodelavci predstavijo metodo, ki za detekcijo uporablja model umetnih nevronskih mrež [56, 157]. Viola in Jones predlagata metodo za detekcijo obrazov, ki temelji na uvedbi Haarovih značilnic, integrala slike in meta učnega algoritma AdaBoost [186]. Predlagane so bile metode, ki temeljijo na statističnih značilnicah: analiza spektralnih histogramov [189], histogram orientiranih gradientov [41] in lokalni binarni vzorci [204]. Predlagane so bile metode, ki omogočajo detekcijo obraza z različnih pogledov [104, 194].

Detektiran obraz želimo spremljati in podrobno analizirati. Nad regijo slike na kateri je bil detektiran obraz se zato izvede poravnava obraza in ocena orientacije obraza. Predlagane so bile številne metode in pristopi, ki na različen način obravnavajo problem sledenja, poravnave in ocene orientacije obraza [132].

Metode predloge videza (angl. *appearance template*) iščejo ujemanje med novo sliko

obrazu z množico pripravljenih primerov izmed katerih vsak pripada anotirani diskretni orientaciji [16, 167]. Jones in Viola predlagata metodo zbirke detektorjev (angl. *detector array*), kjer je v množici detektorjev vsak naučen specifične diskretne orientacije obraza [85]. Nelinearni regresijski modeli iščejo preslikavo med sliko oziroma značilnicami obraza in merami orientacije obraza [104, 138, 175]. Ocenjevanje orientacije obraza lahko zvezno modeliramo z uporabo linearnih in nelinearnih podprostorskih metod [156, 167]. Krüger in Sommer predstavita metodo za učinkovito ocenjevanje orientacije glave z uporabo mrež Gaborjevih valovnih elementov [94]. Predlagane so bile številne metode, ki modelirajo obrazno strukturo kot ne-tog, prilagodljiv objekt [36]. Orientacija obraza se oceni na nivoju značilnic oziroma ob parametrov modela [158]. Med metode prilagodljivih modelov spadajo modeli aktivnega videza [36, 159] in modeli aktivne oblike [179, 180]. Wang in sodelavci opišejo geometrično metodo, ki določi orientacijo obraza na osnovi relativne konfiguracije obraznih značilnic kot so oči, usta in konica nosu [187]. Skupina sledilnih metod oceni orientacijo obraza na osnovi spremembe položaja glave v danem zaporedju slik [100, 124].

2.3.3 Analiza demografskih značilnic

V informacijskem vmesniku bodo ključne tudi metode, s katerimi lahko ocenimo demografske parametre osebe, kot sta na primer spol in starost. Predstavljene so bile metode računalniškega vida, ki za modeliranje spola uporabljajo različne modalnosti [134]. Bourdev in sodelavci za učenje klasifikatorja spola uporabijo množico krpic (angl. *patch*) predstavljeno s HOG značilnicami, barvnim histogramom in značilnicami kože [18]. Med modalnostmi za klasifikacijo spola na osnovi vizualne informacije najdemo tudi človeško hojo (angl. *gait*) [80, 201].

Za nas so posebej zanimive metode, ki delujejo na osnovi registrirane obrazne slike [115, 149], saj se ta pristop najbolj umešča v naš informacijski vmesnik. Moghadam in Yang razvijeta sistem za detekcijo obrazov in primerjata različne klasifikatorje spola [123]. Yang in sodelavci predstavijo podrobno analizo, kako različni pristopi k normalizaciji vplivajo na klasifikacijsko natančnost pri določanju spola [196]. Baluja in Rowley predstavita učinkovito metodo za klasifikacijo spola, ki temelji na algoritmu AdaBoost [10] in ga primerjata z metodo podpornih vektorjev ob spreminjanju velikosti, translacije in rotacije vhodnih podatkov. Shan razvije metodo za klasifikacijo spola obraznih slik iz realnega okolja, ki temelji na lokalnih binarnih vzorcih [166]. Predlagani so bili tudi klasifikatorji spola na osnovi obrazne slike različnih pogledov na osnovi

lokalnih binarnih vzorcev združenimi z značilnicami intenzitete in oblike [4] oziroma kontrastno informacijo [199]. Mäkinen in Raisamo izvedeta sistematično primerjavo različnih metod za poravnavo obraznih slik in klasifikacijo spola [115] na podatkovni zbirki FERET [144, 145].

Ocenjevanje starosti je z metodami strojnega učenja, podobno kot pri spolu, možno na osnovi različnih modalnosti. Predlagani so bili številni pristopi, med njimi tudi na osnovi hoje [110] in glasu [27]. Večina metod pa za oceno uporabi obrazne slike [60]. Ocenjevanje starosti lahko obravnavamo kot večrazredni klasifikacijski problem, kjer starost ocenimo znotraj določenih starostnih razredov, ali pa kot regresijski problem, kjer starost modeliramo z zvezno vrednostjo.

Metoda, ki jo predlagata Kwon in Lobo [99], sodi med začetke strojne klasifikacije starosti na osnovi obraznih slik in temelji na povezavi matematičnega modela z antropometrično razvojno teorijo obraza [55], ki opisuje rast glave od otroštva do odraslih let. Ramanathan in Chellappa opišeta 8 razmerij med merami razdalje s katerimi modelirata staranje obraza [150]. Poleg antropometričnih modelov so bili razviti tudi modeli ocenjevanja starosti na osnovi aktivnih modelov videza [36]. Lanitis in sodelavci razširijo koncept aktivnih modelov videza in z uvedbo funkcije staranja opišejo povezavo med starostjo posameznika in parametričnim opisom obraza [102]. Geng in sodelavci predlagajo obrazne vzorce staranja, ki zajemajo zaporedje obraznih slik iste osebe, urejenih v časovnem redu [65]. Gre za metodo EM (angl. *expectation-maximization algorithm*) iterativnega strojnega učenja, ki modelira starost z uporabo vzorcev starostnih podprostorov. Guo in sodelavci razvijejo model na osnovi mnogoterostnega učenja (angl. *manifold learning*), ki upošteva ortogonalnost, lokalnost in ohranjanje perspektive [69]. Dehshibi in Bastanfard pokažeta vpliv obraznih krivulj in gub na klasifikacijsko natančnost pri oceni starosti [44]. Predlagani so bili tudi modeli, ki najprej izvedejo klasifikacijo spola in šele nato oceno starosti, saj se vzorci staranja med moškimi in ženskami razlikujejo [62].

Regresijsko napovedovanje starosti je bilo predstavljeno z uporabo modelov linearne regresije [102], večkratne linearne regresije [61] in regresijskih podpornih vektorjev [69, 70]. Ni in sodelavci predlagajo robustno več-instančno regresijo, ki naslovi problem osamelih primerov. Pri izgradnji univerzalnega modela za ocenjevanje starosti obrazov uporabijo obširno zbirko obrazov zajetih s spleta [135].



*Metodologija - Adaptivni
informacijski vmesnik*

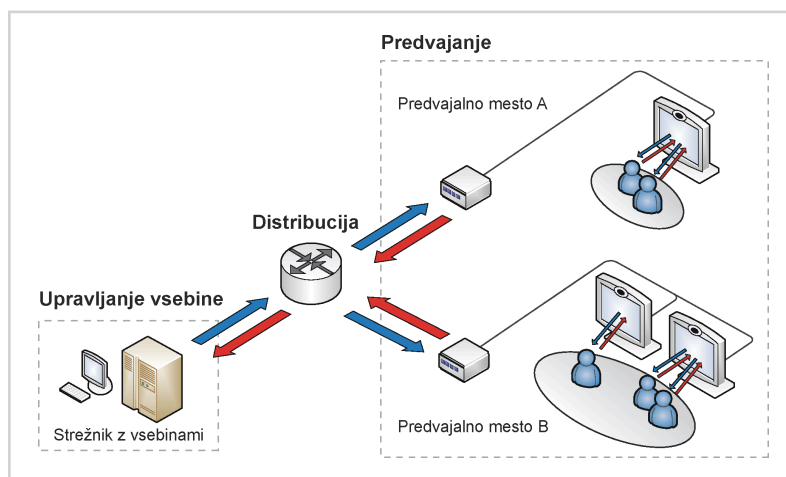
V poglavju opišemo uporabljeno metodologijo in razvoj predlaganega informacijskega vmesnika. Razdelek 3.1 razširi koncept sistema za *slapo* predvajanje vsebine na adaptivni informacijski vmesnik, ki pri izbiri predvajane vsebine poleg predvajalnih značilnic upošteva tudi uporabniške značilnice. Nato predstavimo nabor metod računalniškega vida in strojnega učenja, s katerimi modeliramo časovne, prostorske in demografske značilnice uporabnikov. V razdelku 3.4 opredelimo kontekstualno-odvisno izbiro vsebin, ki upošteva predvajalne in uporabniške značilnice. Razdelek 3.5 zaključí poglavje z umestitvijo predlagane metodologije na področje upravljanja in varovanja zasebnosti.

3.1 Arhitektura adaptivnega informacijskega vmesnika

Pri razvoju adaptivnega informacijskega vmesnika sledimo petim ključnim značilnostmi, ki jih morajo izpolnjevati prihajajoče inteligentne tehnologije [1]: vgrajenost in integracija sistema v okolje, kontekstna-ozaveščenost oziroma prepoznavanje trenutne situacije, personalizacija vezana na uporabnikove potrebe, adaptivnost in vnaprejšnja pripravljenost. Koncept razširimo na interaktivni in adaptiven multimedijiški sistem, ki omogoča kontekstualno prikazovanje vsebine na osnovi analize trenutnega stanja pred zaslonom. Jedro sistema -adaptivnega informacijskega vmesnika- je povezava metod računalniškega vida in strojnega učenja, ki se v realnem času odzovejo na dogajanje, ki ga spremljamo preko kamere vgrajene v okvir zaslona. Arhitekturo predlaganega adaptivnega sistema prikazuje Slika 3.1. Shema je podobna arhitekturi informacijskega vmesnika, ki je predstavljena na Sliki 2.3, s ključno razliko, da je tukaj pretok informacij krožen in v obe smeri: od upravljalnega središča do predvajalnega mesta in preko odziva uporabnikov nazaj.

Opisana metodologija interaktivnega in adaptivnega sistema za prikazovanje vsebine sledi principom načrtovanja uporabniških vmesnikov. Schneiderman definira 8 visokonivojskih konceptov [169], ki upoštevajo psihofizične lastnosti ljudi. Med njimi izpostavimo konsistentnost uporabniškega vmesnika in zagotavljanja relevantne povratne informacije. Razvit adaptivni sistem omogoča nadzor uporabnika preko interakcije in uporablja konsistenten vmesnik ter standardizirano platformo za podajanje informacij [116]. Uporaba adaptivnega sistema je prilagodljiva, preprosta za uporabnika in omogoča intuitivno uporabniški izkušnjo [178]. Glede na specifiko vmesnika pa skozi optimizacijo učinkovitosti sistema opredelimo nov koncept, ki zajema maksimizacijo pozornosti uporabnika.

Poleg predvajalnih značilnic, ki smo jih predstavili v razdelku 2.1, predlagana me-



Slika 3.1

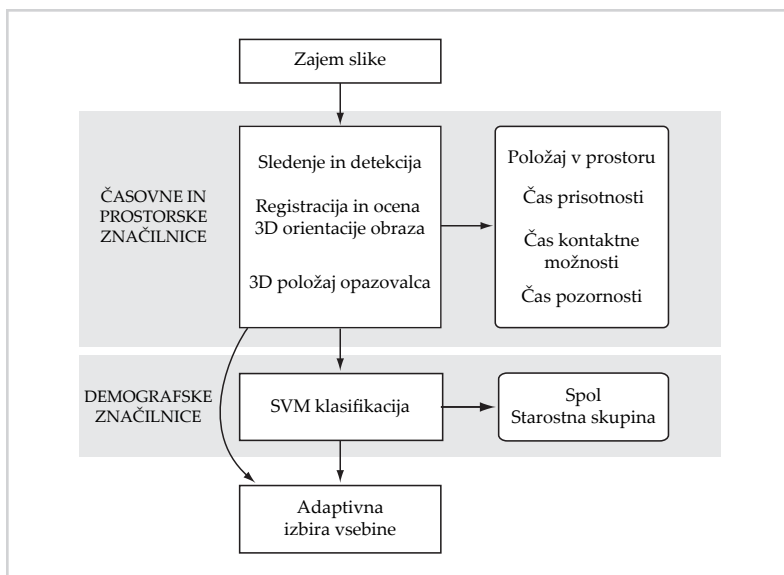
Arhitektura interaktivnega informacijskega vmesnika za predvajanje prilagojenih vsebin. Zasloni so opremljeni s kamero, ki z metodami računalniškega vida in strojnega učenja v realnem času analizira dogajanje na predvajalnem mestu. Tok informacij je dvosmeren in se prilagaja kontekstu, ki ga definirajo trenutni uporabniki. Shema razširja koncept informacijskega vmesnika predstavljenega na Sliki 2.3.

todologija uvaja še uporabniške značilnice (angl. *audience features*). Uporabniške značilnice *AF* so namenjene modeliranju lastnosti uporabnika ter posledično kontekstno odvisni izbiri predvajalnih vsebin. Definiramo tri skupine uporabniških značilnic [153]:

- **PROSTORSKE ZNAČILNICE** opisujejo trenutni položaj uporabnika v prostoru pred zaslonom in omogočajo prostorsko interaktivno podajanje vsebine.
- **ČASOVNE ZNAČILNICE** so primarno namenjene analizi gledanosti in primerjavi učinkovitosti različnih sistemov za prikazovanje vsebin. Opisujejo čas prisotnosti uporabnika na predvajalnem mestu, čas kontaktne možnosti z zaslonom ter čas pozornosti, ki ga je le-ta namenil vsebini zaslona.
- **DEMOGRAFSKE ZNAČILNICE** omogočajo bolj ozko segmentiranje uporabnikov in opisa prikazanih vsebin. Zajemajo spol uporabnika in uvrstitev v starostno skupino.

V nadaljevanju predlagamo nabor metod računalniškega vida in strojnega učenja, s katerimi modeliramo uporabniške značilnice. Zahtevamo, da morajo metode delovati v realnem času na običajni strojni opremi, zato izbira metod upošteva čim bolj ugodno razmerje med robustnostjo in časovno zahtevnostjo. Predpostavimo, tudi da je za zajem

slike na voljo ena kamera, kar v več pogledih oteži določanje prostorskih značilnic. Slika 3.2 podaja shematski prikaz v vmesniku uporabljene metodologije.



Slika 3.2

Struktura adaptivnega informacijskega vmesnika. Za modeliranje časovnih in prostorskih značilnic uporabimo nabor metod računalniškega vida. Demografske značilnice uporabnikov določimo s klasifikatorji strojnega učenja. Postopek izbire vsebine upošteva tako predvajalne kot tudi uporabniške značilnice.

3.2 Prostorsko-časovne značilnice uporabnika

Prostorske značilnice opisujejo položaj uporabnika v prostoru predvajalnega mesta. Za slon z vgrajeno kamero postavimo v središče koordinatnega sistema. Uporabljene časovne značilnice so: čas prisotnosti, čas kontaktne možnosti in čas pozornosti.

Čas prisotnosti uporabnika (angl. *dwell time*) opredelimo kot čas, ki ga uporabnik preživi na področju informacijskega vmesnika. Upošteevamo, da lahko isti uporabnik področje začasno zapusti in se nanj nato tudi vrne. Čas prisotnosti uporabnika $T_{d,i}$ zato definiramo kot vsoto vseh časovnih intervalov prisotnosti uporabnika:

$$T_{d,i} = \sum_{j=1}^{N_{d,i}} \tau_{d,j}, \quad (3.1)$$

kjer i označuje uporabnika, $N_{d,i}$ število intervalov prisotnosti uporabnika i in $\tau_{d,j}$ tra-

janje posameznega intervala. Definiramo tudi povprečni čas prisotnosti \bar{T}_d vseh uporabnikov kot:

$$\bar{T}_d = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^{N_{d,i}} \tau_{d,j} = \frac{1}{N} \sum_{i=1}^N T_{d,i} \quad (3.2)$$

in povprečno število intervalov prisotnosti uporabnikov \bar{N}_d kot:

$$\bar{N}_d = \frac{1}{N} \sum_{i=1}^N N_{d,i}, \quad (3.3)$$

kjer N označuje število vseh prisotnih uporabnikov v opazovanem časovnem obdobju.

Čas kontaktne možnosti uporabnika (angl. *opportunity to see, opportunity to contact*) definiramo kot celotni čas, v katerem je imel uporabnik možnost opaziti predvajano vsebino na zaslonu. Čas kontaktne možnosti uporabnika $T_{o,i}$ izračunamo kot:

$$T_{o,i} = \sum_{j=1}^{N_{o,i}} \tau_{o,j}, \quad (3.4)$$

kjer i označuje uporabnika ($1 \leq i \leq N$), $N_{o,i}$ število intervalov kontaktne možnosti uporabnika i in $\tau_{o,j}$ trajanje posameznega intervala. Po definiciji so $\tau_{o,j}$ podintervali $\tau_{d,j}$, iz česar sledi da je $T_{o,i} \leq T_{d,i}$. Definiramo še povprečni čas kontaktne možnosti \bar{T}_o vseh uporabnikov kot:

$$\bar{T}_o = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^{N_{o,i}} \tau_{o,j} = \frac{1}{N} \sum_{i=1}^N T_{o,i} \quad (3.5)$$

in povprečno število intervalov kontaktne možnosti \bar{N}_o kot:

$$\bar{N}_o = \frac{1}{N} \sum_{i=1}^N N_{o,i} \quad . \quad (3.6)$$

Čas pozornosti uporabnika (angl. *attention time*) pa definiramo kot čas, ki ga je uporabnik porabil za aktivno spremljanje vsebine na zaslonu. Čas pozornosti uporabnika $T_{a,i}$ definiramo kot vsoto vseh časovnih intervalov pozornosti uporabnika:

$$T_{a,i} = \sum_{j=1}^{N_{a,i}} \tau_{a,j}, \quad (3.7)$$

kjer i označuje uporabnika ($1 \leq i \leq N$), $N_{a,i}$ število intervalov pozornosti uporabnika i in $\tau_{a,j}$ trajanje posameznega intervala. Definiramo tudi povprečni čas pozornosti \bar{T}_a vseh uporabnikov kot:

$$\bar{T}_a = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^{N_{a,i}} \tau_{a,j} = \frac{1}{N} \sum_{i=1}^N T_{a,i} \quad (3.8)$$

in povprečno število intervalov pozornosti vseh uporabnikov \bar{N}_a kot:

$$\bar{N}_a = \frac{1}{N} \sum_{i=1}^N N_{a,i} \quad . \quad (3.9)$$

V nadaljevanju predstavimo nabor metod računalniškega vida, ki jih uporabimo za določitev vseh pravkar opisanih časovnih značilnic.

3.2.1 Segmentacija ozadja in sledenje aktivnih regij

Čas prisotnosti uporabnika na predvajanem mestu modeliramo z njegovo prisotnostjo na zajeti sliki, kjer se uporabnik ali več uporabnikov hkrati tipično giba in premika. V fazi predprocesiranja segmentacija ozadja zmanjša prostor slike na množico aktivnih regij, katere uporabimo pri nadaljnji analizi. Pregledovanje zgolj nekaterih regij in ne celotne slike pospeši delovanje vmesnika.

Za segmentacijo ozadja uporabimo adaptivni model mešanice Gaussovih porazdelitev (angl. *Mixtures of Gaussians, MOG*) [207, 208]. Model se v primerjavi z drugimi sorodnimi metodami in pristopi za segmentacijo ozadja slike pokaže za eno bolj robustnih in učinkovitih metod [23].

Metoda modelira vsako slikovno točko z njeno intenziteto v RGB barvnem prostoru [19]. Verjetnost, da se v izbrani točki pojavi trenutna vrednost, se v več dimenzijskem prostoru izračuna kot:

$$P(X_t) = \sum_{i=1}^K \omega_{i,t} \eta(X_t, \mu_{i,t}, \Sigma_{i,t}), \quad (3.10)$$

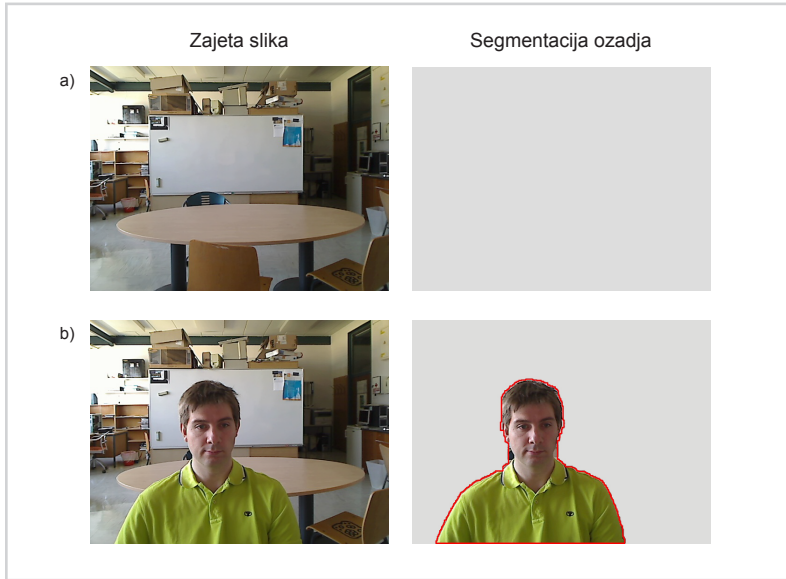
kjer K predstavlja število porazdelitev, $\omega_{i,t}$ utež i -te Gaussove porazdelitve v času t s povprečno vrednostjo $\mu_{i,t}$ in standardnim odklonom $\Sigma_{i,t}$. η predstavlja funkcijo gostote Gaussove porazdelitve:

$$\eta(X_t, \mu, \Sigma) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}(X_t - \mu) \Sigma^{-1} (X_t - \mu)} \quad . \quad (3.11)$$

Vsako slikovno točko torej opiše K Gaussovih porazdelitev. Ob času $t + 1$ se za vsako slikovno točko izračuna ujemanje z njenimi Gaussovimi porazdelitvami na osnovi Mahalanobis-ove razdalje [114]. V primeru ujemanja točko razglasimo za ozadje in prilagodimo parametre porazdelitve, kateri se najbolj prilega. Če razdalja ne preseže pragu, točko označimo kot del ospredja. Porazdelitev, ki ima najmanjše ujemanje z trenutno vrednostjo pa zamenjamo z novo [19]. Število komponent, ki opisujejo zgodovino vrednosti vsake slikovne točke, se prilagaja vsaki točki posebej. Prilagajanje števila komponent vpliva na manjši čas procesiranja, izboljša pa tudi rezultate segmentacije [207].

Predprocesiranje slike ozadja zaključimo z zaporedjem morfoloških operacij. Nad sliko segmentiranega ospredja najprej izvedemo operacijo razjedanja (angl. *erosion*), ki odstrani osamele točke. Sledi dvojna ponovitev operacije razširjanja (angl. *dilation*), ki zapolni vrzeli v regijah ospredja. Postopek zaključimo z operacijo razjedanja, ki odpravi prekomerno povečanje regij ospredja na račun druge operacije razširjanja. Celoten postopek predprocesiranja poteka na pomanjšani sliki (skalirni faktor = 0.6). Tako dosežemo približno dvakrat nižji čas izvajanja predprocesiranja in ohranimo zadovoljivo natančnost. Primer predprocesiranja zajete slike je predstavljen na Sliki 3.3. Segmentirana slika služi kot vhod metode za sledenje uporabnikov.

Za sledenje uporabnika uporabimo metodo hitrega ujemanja s predlogo (angl. *Fast Template Matching*) [22]. Predlogo definira zaključena aktivna regija uporabnika, ki jo dobimo po segmentaciji slike (glej Sliko 3.3b). Posodabljanje predloge se časovno prilagaja gibanju uporabnika, za katerega hranimo seznam lastnosti najbolj značilnih predlog. Predlogo modeliramo z ekviliziranim barvnim histogramom [20]. Histogram trenutne predloge dodamo v seznam, ko razlika med njim in vsemi preteklimi histogrami uporabnika preseže prag, določen z Bhattacharyya-evo razdaljo [88].



Slika 3.3

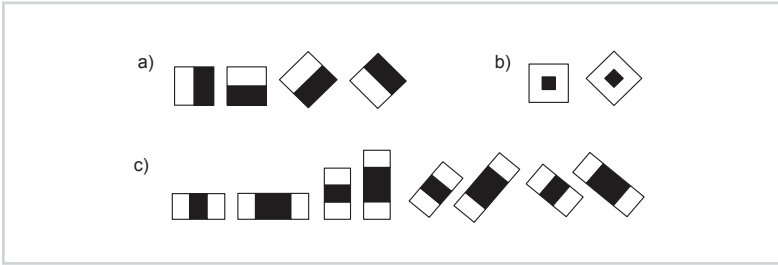
Predprocesiranje vhodne slike. a) Slika praznega predvajalnega mesta (levo) in rezultat segmentacije ozadja (desno). b) Predvajalno mesto na katerega je prišel uporabnik (levo). Regija osrednja, ki jo dobimo po segmentaciji ozadja z adaptivnim modelom mešanice Gaussovih porazdelitev (desno).

3.2.2 Detekcija obraza

Čas kontaktne možnosti modeliramo z metodo za detekcijo obraza v frontalnem in profilnem položaju. Za detekcijo obrazov uporabimo statistični pristop, ki sta ga predlagala Viola in Jones [186] ter kasneje analizirala in razširila Lienhart in Maydt [106]. Metodo izberemo zato, ker ponuja dobro razmerje med hitrostjo in zanesljivostjo [115]. Metoda deluje v realnem času in detektira obraze neodvisno od njihovega položaja in velikosti. Natančnost metode je ocenjena na 98% [105], kar zadošča za potrebe našega sistema.

Metoda uvede preproste značilnice pravokotnih oblik imenovane Haar značilnice, ki tvorijo osnovo modela za opis obraznega klasifikatorja. Haar značilnica je binarna funkcija, katere prisotnost ugotovimo s primerjavo povprečne sivinske vrednosti sosednjih regij. Prag primerjave se določi v fazi učenja klasifikatorja. Primer razširjenega nabora Haar značilnic je prikazan na Sliki 3.4.

Prisotnost Haar značilnic na regiji poljubne velikosti in položaja lahko učinkovito preverimo z uporabo integralne slike II . Točka integralne slike II na koordinatah (x,y)



Slika 3.4

Razširjen nabor Haar značilnic, ki opisujejo različne vzorce. a) Robne značilnice. b) Značilnice obkroženega središča (angl. *center-surround features*). c) Linijske značilnice.

predstavlja vsoto vrednosti vseh točk, ki jih omejuje pravokotnik med $(0, 0)$ in (x, y) . Definiramo jo kot:

$$\Pi(x, y) = \sum_{i=1}^x \sum_{j=1}^y I(i, j) \quad (3.12)$$

kjer $I(i, j)$ predstavlja sivinsko vrednost v točki (i, j) osnovne slike I . Integralno sliko lahko izračunamo z enim samim prehodom slike na osnovi rekurzivne definicije:

$$S(x, y) = S(x, y - 1) + I(x, y), \quad (3.13)$$

$$\Pi(x, y) = \Pi(x - 1, y) + S(x, y), \quad (3.14)$$

kjer $S(x, y)$ predstavlja kumulativno vsoto vrstice. Robna pogoja rekurzije sta:

$$S(x, -1) = 0,$$

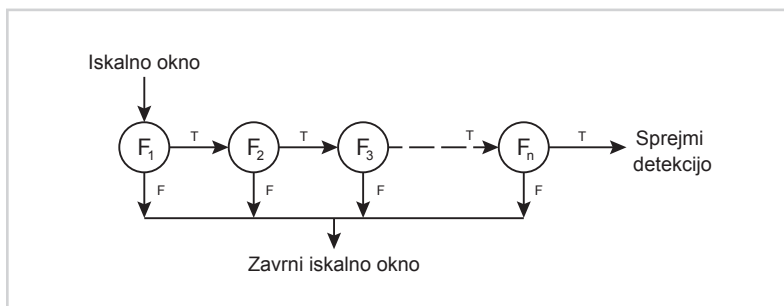
$$\Pi(-1, y) = 0.$$

Z uvedbo integralne slike lahko izračunamo vsoto sivinskih točk poljubne pravokotne regije s 4 naslavljajči integralne slike in 3 računskimi operacijami. To predstavlja izreden prihranek pri računanju in preverjanju prisotnosti značilnic različnih velikosti in oblik [103].

V fazi učenja metoda na osnovi Haarovih značilnic in algoritma AdaBoost zgradi množico šibkih klasifikatorjev (angl. *weak classifier*), ki jih nato združi v kaskado močnih klasifikatorjev F_k . Šibek klasifikator f_l predstavlja binarni pragovni klasifikator, katerega uspešnost presega vsaj klasifikacijsko natančnost naključnega klasifikatorja.

Slika 3.5

Kaskada klasifikatorjev. Preverjanje se ustavi, če iskalno okno zavrne vsaj eden od močnih klasifikatorjev F_k . Iskalno okno, ki je uspešno prešlo vse šibke klasifikatorje metoda označi za obraz.



Močni klasifikator dobimo kot linearno kombinacijo šibkih klasifikatorjev:

$$F_k = \text{sign}(w_1 f_1 + w_2 f_2 + \dots + w_m f_m), \quad (3.15)$$

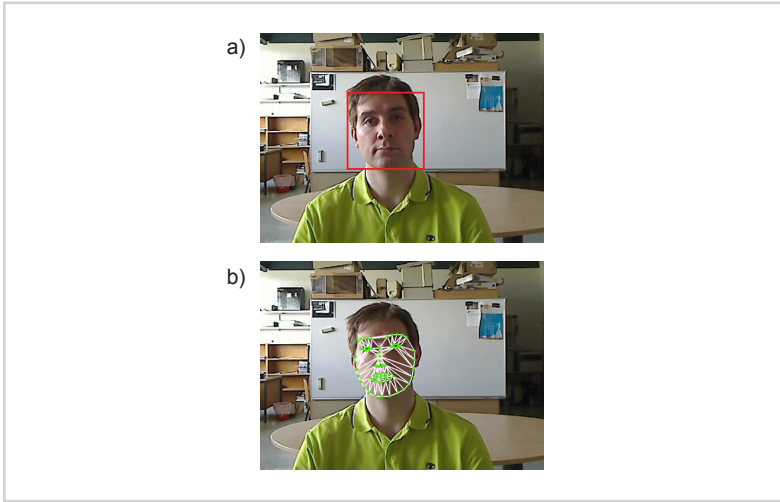
kjer sign predstavlja funkcijo predznaka, F_k močni klasifikator, f_l šibek klasifikator in w_l njegovo utež. Najbolj selektivni klasifikatorji so uvrščeni na začetek kaskade, saj s tem zmanjšamo čas detekcije. Pri detekciji vsak izmed F_k klasifikatorjev analizira iskalno okno. Detektor zavrne iskalno okno, če ga zavrne vsaj eden od klasifikatorjev. Okno sprejme le v primeru, ko uspešno preide celotno verigo klasifikatorjev. Shema kaskade prikazuje Slika 3.5.

Pri detekciji uporabimo kaskadi namenjeni frontalni detekciji in detekciji profila obraza [20]. Metodo za detekcijo obrazov uporabimo na segmentirani sliki ozadja. Kaskada klasifikatorjev hitro zavrne uniformne dele ozadja, kar izboljša časovno učinkovitost izvajanja metode. Detektor vrne položaj obrazov, katerih velikost ni manjša od 20x20 slikovnih točk. Primer detektiranega obraza je prikazan na Sliki 3.6a.

3.2.3 Poravnava in ocena 3D orientacije obraza

Čas pozornosti uporabnika določimo na osnovi njegovega položaja in orientacije glave. V nadaljevanju opišemo postopek poravnave obraza in ocene orientacije glave.

Cilj poravnave obraza je poiskati preslikavo med 2D modelom oblike oziroma videza od detektirano sliko obraza. Poravnava nam služi kot vmesni korak pri oceni orientacije obraza in modeliranju demografskih značilnic. Za poravnavo obraza izberemo modele aktivnega videza (angl. *active appearance models*), ki predstavljajo generativno in parametrično metodo za modeliranje vizualnih pojavov [36, 117]. Metoda poveže



Slika 3.6

Detekcija in poravnava obraza. a) Rezultat detekcije obraza z metodo Viola in Jones. b) Poravnava obraza z uporabo modelov aktivnega videza. 66 značilnih točk obraza je označenih z zeleno barvo.

variacijo oblike in videza v enoten statističen model. Izhaja iz modelov aktivne oblike [179, 180], vendar za primerljivo uspešnost zahteva manjši nabor parametrov.

Metoda definira geometrijo obraza kot nabor vnaprej določenih značilnih točk (angl. *landmarks*), ki opisujejo fizično konsistentne lokacije obraza (npr. konica nosu, rob očesa, rob veke). Število značilnih točk, na katerih zgradimo model je odvisno od specifične aplikacije. Natančnost in robustnost poravnave narašča skupaj s številom točk, saj njihove ločene ocene krepijo druga drugo skozi model relativne prostorske odvisnosti. Z naraščajočim številom točk pa linearno narašča tudi časovna zahtevnost metode [9].

Modeli aktivnega videza omogočajo linearno variacijo oblike. Obliko s tako definiramo kot vsoto med osnovno obliko s_0 in linearno kombinacijo vektorjev oblike s_i :

$$s = s_0 + \sum_{i=1}^n p_i s_i, \quad (3.16)$$

kjer p_i predstavljajo parametre oblike, n pa število vektorjev oblike [117]. Metoda na podoben način modelira tudi videz obraza.

Videz $A(\mathbf{x})$, zajema vse slikovne točke, ki ležijo znotraj osnovne mreže \mathbf{s}_0 in je definiran kot:

$$A(\mathbf{x}) = A_0(\mathbf{x}) + \sum_{i=1}^m \lambda_i A_i(\mathbf{x}), \quad (3.17)$$

kjer λ_i predstavljajo parametre videza, $A_0(\mathbf{x})$ osnovni videz obraza, m število slik v fazi učenja, $A_i(\mathbf{x})$ pa slike videza iz faze učenja. Instanca modela aktivnega videza je definirana z naborom parametrov oblike in videza, kateri služijo za krivljenje (angl. *warping*) videza A iz osnovne mreže \mathbf{s}_0 v obliko \mathbf{s} . Operacijo krivljenja označimo z $\mathbf{W}(\mathbf{x}; \mathbf{p})$, instanco modela pa z M .

Prilagajanje modela (angl. *fitting*) je optimizacijski problem, pri katerem minimiziramo razliko med vhodno sliko obraza in instanco modela aktivnega videza [158, 159]. Za dano vhodno sliko $I(\mathbf{x})$ želimo poiskati optimalne parametre oblike \mathbf{p} in videza λ . To pomeni, da stremimo k čim večji podobnosti med sliko $I(\mathbf{x})$ in instanco modela, torej $M(\mathbf{W}(\mathbf{x}; \mathbf{p})) = A(\mathbf{x})$. Optimizacijski kriterij za prilagajanje modela tako predstavlja minimizacijo vsote:

$$\sum_{\mathbf{x} \in \mathbf{s}_0} \left[A_0(\mathbf{x}) + \sum_{i=1}^m \lambda_i A_i(\mathbf{x}) - I(\mathbf{W}(\mathbf{x}; \mathbf{p})) \right]^2, \quad (3.18)$$

ki teče po vseh točkah \mathbf{x} v osnovni mreži \mathbf{s}_0 . V splošnem je optimizacija parametrov videza λ linearna, optimizacija parametrov oblike \mathbf{p} pa nelinearna [117]. Primer prilagajanja modela aktivnega videza na regijo, kjer je bil detektiran obraz prikazuje Slika 3.7. Model predstavljajo zelene točke, povezane z belimi črtami.

Po uspešni poravnavi obraza in postavitvi značilnih točk z uporabo POSIT algoritma ocenimo 3D orientacijo obraza [45]. Ime algoritma je kratica, ki označuje ocenjevanje orientacije na osnovi ortografije z iteracijami (angl. *Pose from Orthography and Scaling with Iterations*). Algoritem oceni orientacijo objekta na osnovi posamezne slike. Metoda hitro konvergira, ni odvisna od začetne ocene orientacije objekta in omogoča izvajanje v realnem času.

Metoda predpostavlja, da smo na objektu detektirali in locirali vsaj 4 ne-koplanarne točke ter da poznamo njihovo relativno lego na objektu. Najprej z uporabo skalirane ortografske projekcije in reševanjem linearnega sistema aproksimira perspektivno projekcijo in poišče rotacijsko matriko ter translacijski vektor. Sledi iterativno izboljševanje ocenjene orientacije, ki skalirano ortografsko projekcijo še bolj prilagodi značilnim

točkam [66].

Rezultat ocene orientacije obraza prikazuje Slika 3.7. Na osnovi izračunanje rotacijske matrike vizualiziramo osi naklona (angl. *pitch*), odklona (angl. *yaw*) in nagiba (angl. *roll*) obraza z rdečo, rumeno in modro črto.



Slika 3.7

Ocena orientacije obraza. a) Referenčna slika (levo) in ocena orientacije obraza (desno). Uporabnik ima obraz obrnjen proti zaslonu. Osi naklona, odklona in nagiba so označene z rdečo, rumeno in modro črto. b) Referenčna slika (levo) in ocena orientacije obraza (desno), ko ima uporabnik glavo obrnjeno poševno navzgor.

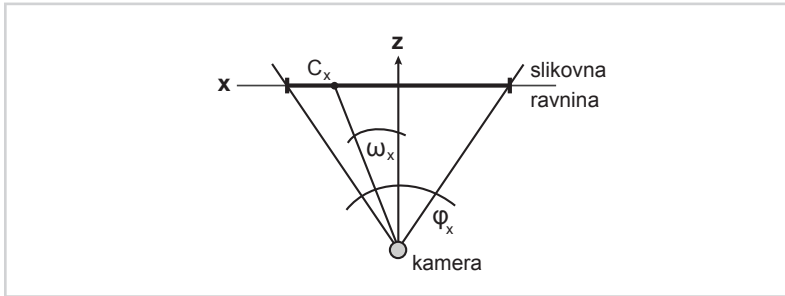
3.2.4 Ocena 3D položaja uporabnika

Položaj uporabnika v prostoru omogoča številne načine prostorske interakcije. Opredelimo ga z kotom med kamero in uporabnikom ter razdaljo med uporabnikom in kamero. Predstavimo učinkovito in hitro metodo, ki z uporabo ene kamere oceni oba parametra in določi položaj uporabnika [153].

Najprej izračunamo kot med uporabnikom in kamero. Poravnava obraza določi položaj 66 značilnih obraznih točk. Vsako oko opisuje 6 značilnih točk, ki tvorijo konveksni mnogokotnik okrog očesne votline. Središčno točko očesa izračunamo kot centroid mnogokotnika. Središčni točki levega in desnega očesa označimo z E_1 in E_2 .

Slika 3.8

Ocena 3D položaja uporabnika. Prikaz horizontalne ravnine (x, z), kjer ω_x označuje kot med uporabnikom in optično osjo kamere, φ_x horizontalni zorni kot kamere in C_x projekcijo središčne točke obraza na slikovni ravnini.



Na osnovi središčnih točk očesa definiramo središčno točko obraza C , katere koordinate na sliki izračunamo kot razpolovišče med točkama E_1 in E_2 :

$$C_x = \frac{E_{1,x} + E_{2,x}}{2} \quad (3.19)$$

$$C_y = \frac{E_{1,y} + E_{2,y}}{2} \quad (3.20)$$

Uporabimo še horizontalni zorni kot kamere φ_x in vertikalni kot kamere φ_y . Vrednosti φ_x in φ_y sta vezani na tehnične lastnosti kamere in se ju pridobi na podlagi tehničnih listin. Horizontalni kot ω_x predstavlja kot med uporabnikom in optično osjo kamere v ravnini (x, z). Vertikalni kot ω_y pa definira kot med uporabnikom in kamero v ravnini (y, z). Izračunamo ju lahko kot:

$$\omega_x = \arctan \left(2 \tan \left(\frac{\varphi_x}{2} \right) \cdot \left(\frac{C_x}{I_w} - 0.5 \right) \right) \quad (3.21)$$

$$\omega_w = \arctan \left(2 \tan \left(\frac{\varphi_y}{2} \right) \cdot \left(\frac{C_y}{I_h} - 0.5 \right) \right), \quad (3.22)$$

kjer I_w predstavlja širino slike, I_h pa višino slike v točkovnih elementih. Geometrijska shema horizontalne ravnine je predstavljena na Sliki 3.8. Odebeljena črta označuje slikovno ravnino, na kateri se nahaja središčna točka obraza C_x . Optično os kamere označuje z .

Dogson predstavi rezultate študije v kateri ugotavlja, da ima večina odraslih ljudi skoraj enako medzenično razdaljo [50]. Ta je pri več kot 90% odraslih med 57mm in 69mm in v povprečju ~ 63 mm. Razdalja torej pri večini odraslih ljudi ne odstopa

od povprečja za več kot 10%. V nadaljevanju v implementaciji našega vmesnika predpostavimo konstantno medočesno razdaljo $IPD = 63\text{mm}$, katero pri analizi slike definiramo kot evklidsko razdaljo med središčnima točkama E_1 in E_2 :

$$IPD = \|E_1 - E_2\| \quad (3.23)$$

Medočesna razdalja je obratno sorazmerna razdalji med obrazom in kamero. Definiramo aproksimacijsko funkcijo razdalje \hat{F}_{dist} kot:

$$\hat{F}_{dist} = A/x + B, \quad (3.24)$$

kjer x predstavlja ocenjeno medočesno razdaljo v slikovnih točkah, A in B pa sta konstanti vezani na tehnične lastnosti kamere (resolucija, karakteristike leče). Pri vseh opisanih eksperimentih je bila uporabljena kamera Logitech WebCam Pro 9000, katere horizontalni zorni kot znaša 63.1° , vertikalni zorni kot pa 49.4° . Kamera zajema sliko pri resoluciji 1600×1200 točkovnih točk. Z uporabo standardnih numeričnih paketov in primerjalnimi testi določimo parametra $A = 21500$ in $B = 7.78$.

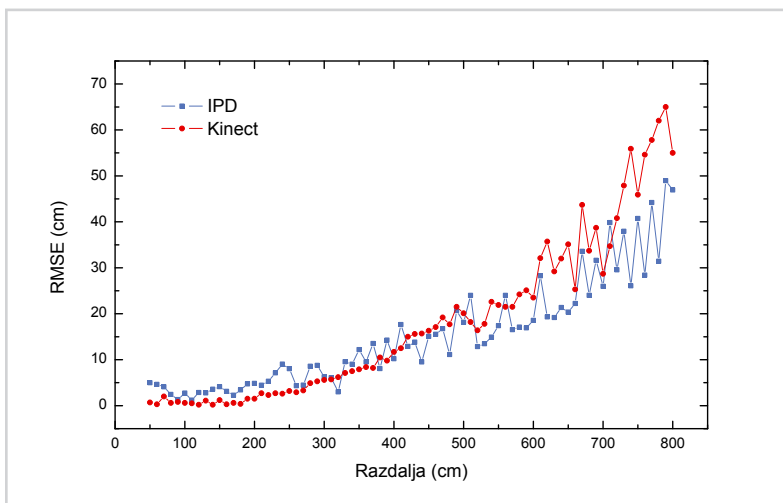
Natančnost predlagane metode ovrednotimo z eksperimentalno primerjavo. Za referenčno oceno vzamemo meritve razdalje pridobljene z globinskim RGB-D senzorjem Kinect, ki je uveljavljena platforma za interaktivne uporabniške vmesnike [130]. Osnovni programski vmesnik za dostop do podatkov globinskega senzorja vrača podatke do razdalje $\sim 4\text{m}$, kar je nominalna razdalja, ki jo Kinect uradno podpira. Senzor sicer vrača podatke do razdalje $\sim 9\text{m}$, zato za zajem podatkov s Kinecta uporabimo programski vmesnik OpenNI [8].

Testirali smo pet različnih uporabnikov pri hoji po ravni liniji, z obrazom obrnjenim proti kameri našega razvitega sistema in proti kameri Kinecta. Kamera je bila postavljena nad Kinect, oba pa sta bila poravnana vzdolž smeri hoje. Na tleh so bile v ravni liniji nameščene oznake, ki so označevale 10cm intervale od razdalje 0.5m do 8m. Ko je uporabnik dosegel oznako sta oba sistema zabeležila ocenjeno razdaljo. Slika 3.9 prikazuje koren srednje kvadratne napake (RMSE) pri zajetih meritvah različnih razdalj [153].

Primerjava natančnosti pokaže, da povprečna napaka (koren srednje kvadratne napake) med razdaljo od 0.5m do 8m znaša 18.7cm za Kinect in 19.3cm za predlagano metodo. Kinect vrača stabilne rezultate do razdalje 4m, pri večji razdaljah pa napaka narašča eksponentno. Obnašanje senzorja Kinect se sklada z ugotovitvami v [8]. Iz-

Slika 3.9

Primerjava merjenja razdalje z uporabo senzorja Kinect (rdeča) in metode IPD (modra). Koren srednje kvadratne napake (RMSE), se povečuje skupaj z razdaljo. Napaka obeh sistemov na danem intervalu merjenja ostaja v primerljivih okvirjih.

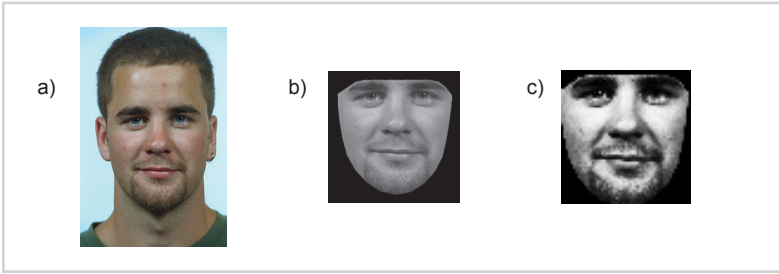


merjena napaka obeh sistemov narašča z razdaljo, vendar ostaja v primerljivih okvirjih. Relativna napaka predlagane metode za ocenjevanje razdalje znaša 4.3%. Metoda daje zadovoljivo natančnost in je tako časovno kot strojno nezahtevna, saj ne predvideva posebne opreme (stereo vid) in teče v realnem času.

3.3 Modeliranje demografskih značilnic

Na osnovi obrazne slike modeliramo osrednji demografski značilnici, to sta spol in starostna skupina uporabnika. Detektirano sliko obraza najprej poravnamo in jo s pomočjo krivljenja pretvorimo v poravnano frontalno obliko (glej Sliko 3.10b). Sledi postopek normalizacije, v katerem izvedemo izravnavo histograma sivinskih vrednosti in skaliranje na velikost 50x50 slikovnih točk. To velikost normalizirane slike izberemo na podlagi obstoječih študij optimalnih primerov [115, 123]. Primer normaliziranega obraza je prikazan na Sliki 3.10c.

Normaliziran obraz projeciramo v nižje dimenzijski prostor z uporabo metode glavnih komponent. Metoda glavnih komponent (PCA) je podprostorska metoda, ki z minimizacijo rekonstrukcijske napake in maksimizacijo variance projekcije določi ortogonalno bazo lastnih vektorjev, s katerimi opišemo nabor vhodnih slik [103, 181].



Slika 3.10

Stopnje normalizacije obraza. a) Vhodna slika. b) Poravnava obraza v frontalno obliko. c) Normalizirana oblika obraza. Osnovna slika je del podatkovne zbirke slik FERET [144, 145].

V fazi učenja iz množice M vhodnih slik izračunamo povprečno sliko \bar{x} :

$$\bar{x} = \frac{1}{M} \sum_{i=1}^M \mathbf{x}_i, \tag{3.25}$$

kjer \mathbf{x}_i predstavlja sivinske vrednosti slike, zapisane v vektorski obliki. Nato definiramo kovariančno matriko \mathbf{Q} kot:

$$\mathbf{Q} = \mathbf{A}\mathbf{A}^T, \tag{3.26}$$

kjer je $\mathbf{A} = [\mathbf{x}_1 - \bar{x}, \mathbf{x}_2 - \bar{x}, \dots, \mathbf{x}_M - \bar{x}]$. Z razcepom na singularne vrednosti izračunamo lastne vektorje in lastne vrednosti matrike \mathbf{Q} . Lastni vektorji matrike \mathbf{Q} so netrivialne rešitve ($\nu \neq 0$) karakteristične enačbe:

$$(\mathbf{Q} - \lambda \mathbf{I})\nu = 0, \tag{3.27}$$

kjer λ označuje lastno vrednost, ν lastni vektor, \mathbf{I} pa enotsko matriko. Lastne pare uredimo padajoče glede na velikost lastne vrednosti, tako da velja $\lambda_1 \geq \lambda_2 \geq \dots \lambda_M$. Efektivno po velikosti večja lastna vrednost pomeni, da pripadajoči lastni vektor opisuje več informacije. Za normalizirano sliko \mathbf{x}_j lahko izračunamo njeno projekcijo g_{ji} na lastne vektorje sistema:

$$g_{ji} = \nu_i^T \cdot (\mathbf{x}_j - \bar{x}), \tag{3.28}$$

kjer ν_i označuje enega od M lastnih vektorjev. Izračunane projekcije uredimo v vektor projekcij $\mathbf{g}_j = [g_{j1}, g_{j2}, \dots, g_{jM}]$. Sliko \mathbf{x}_j v vektorski obliki sedaj lahko rekonstruiramo kot:

$$\mathbf{x}_j = \bar{x} + \sum_{i=1}^M g_{ji} \nu_i, \tag{3.29}$$

kjer \bar{x} predstavlja povprečno sliko sistema, g_{ji} i -to komponento projekcije \mathbf{g}_j in ν_i je i -ti lastni vektorje kovariančne matrike \mathbf{Q} . Pri slikah večjih dimenzij se izračun lastnih vrednosti in lastnih vektorjev izvede preko kovariančne matrike $\hat{\mathbf{Q}} = \mathbf{A}^T \mathbf{A}$, kar zmanjša prostorsko in časovno zahtevnost učenja.

Vektor projekcije \mathbf{g}_j nam služi kot abstrahiran zapis normalizirane slike. Z uporabo metode podpornih vektorjev smo sliko pretvorili v nižje dimenzijsko obliko, ki je primerna za strojno učenje klasifikatorjev demografskih značilnic uporabnikov.

Primer povprečnega vektorja \bar{x} , ki ga pretvorimo nazaj v slikovno obliko je prikazan na Sliki 3.11a, najbolj značilni lastni vektorji sistema, urejeni glede na lastne vrednosti, pa na Sliki 3.11b.

Za učenje in evaluacijo metode uporabimo podatkovno zbirko FERET [144, 145]. Zbirka zajema 2.413 obraznih slik 856 različnih oseb. Zbirka zajema slike ljudi z različnimi modalnostmi in vsakega pri dveh različnih starostih. Poleg slikovnega gradiva vključuje tudi anotacijske podatke o starosti in spolu osebe na fotografiji. Za evaluacijo sistema smo uporabili 845 normaliziranih frontalnih obraznih slik, ki jih z uporabo metode glavnih komponent opišemo z vektorjem projekcij.

Za modeliranje demografskih lastnosti primerjamo klasifikatorje spola in klasifikatorje starostne skupine. Podatke uredimo v 7 starostnih skupin, ki zajemajo starost: 1-14 let, 15-24 let, 25-34 let, 35-44 let, 45-54 let, 55-64 let in več kot 64 let. Z uporabo 10-kratnega prečnega preverjanja primerjamo različne algoritme strojnega učenja, in sicer naivni Bayesov klasifikator (NB), metodo K-najbližjih sosedov (kNN), metodo podpornih vektorjev (MPV) in metodo naključnih gozdov (NG). Pri evaluaciji



Slika 3.11

Metoda glavnih komponent. a) Povprečni vektor sistema \bar{x} , pretvorjen v slikovno obliko. b) Lastni vektorji sistema z največjimi lastnimi vrednostmi.

uporabimo ogrodje za podatkovno rudarjenje Orange [46]. Parametri metod strojnega učenja obdržijo privzete vrednosti. Rezultati prečnega preverjanja za klasifikatorje spola so predstavljeni v Tabeli 3.1, za klasifikatorje starostne skupine pa v Tabeli 3.2. Senzitivnost (angl. *sensitivity*) označuje delež pravilno klasificiranih pozitivnih primerov, specifičnost (angl. *specificity*) pa delež pravilno klasificiranih negativnih primerov.

Tabela 3.1

Primerjava različnih algoritmov strojnega učenja za klasifikacijo spola na osnovi normaliziranih obranih slik. Primerjava z 10-kratnim prečnim preverjanjem pokaže, da največjo klasifikacijsko točnost ($CA_{MPV} = 0.952$) doseže metoda podpornih vektorjev.

Metoda	Klasifikacijska točnost	Senzitivnost	Specifičnost
NB	0.938	0.951	0.923
kNN	0.875	0.915	0.824
MPV	0.952	0.963	0.939
NG	0.931	0.932	0.931

V množici 845 primerov je bilo 376 žensk in 469 moških fotografij. Ciljni razred binarnega klasifikatorja spola je večinski razred, tj. moški spol. Med klasifikatorji spola se najbolje izkaže metoda podpornih vektorjev, saj doseže klasifikacijsko točnost $CA_{MPV} = 0.952$. Visoka senzitivnost in specifičnost kažeta na visok odstotek pravilno klasificiranih pozitivnih in negativnih primerov. Metodi naivnega Bayesa ($CA_{NB} = 0.938$) in naključnih gozdov ($CA_{NG} = 0.931$) dosežeta primerljivo natančnost, medtem ko je metoda K-najbližjih sosedov bistveno manj uspešna ($CA_{kNN} = 0.875$).

Starostna distribucija primerov, ki smo jih uporabili za evaluacijo, je bila razdeljena na: 14 oseb v razredu 1-14 let, 390 oseb v razredu 15-24 let, 227 oseb v razredu 25-34 let, 132 oseb v razredu 35-44 let, 61 oseb v razredu 45-54 let, 19 oseb v razredu 55-64 let in 2 osebi v razredu nad 64 let. Gre za neenakomerno porazdelitev, ki je posledica demografskih značilnosti sodelujočih pri nastajanju podatkovne zbirke FERET.

Tudi v primeru klasifikatorja starostne skupine se za najboljšo izkaže metoda podpornih vektorjev, ki doseže klasifikacijsko točnost $CA_{MPV} = 0.917$. Sledijo metoda naključnih gozdov z $CA_{NG} = 0.909$, metoda naivnega Bayesa ($CA_{NB} = 0.859$) in metoda K-najbližjih sosedov ($CA_{kNN} = 0.823$). Senzitivnost in specifičnost sta vezana na večinski razred, tj. starostna skupina od 15-24 let.

Tabela 3.2

Primerjava različnih algoritmov strojnega učenja za klasifikacijo starostnih skupin na osnovi normaliziranih obranih slik. Primerjava z 10-kratnim prečnim preverjanjem pokaže, da največjo klasifikacijsko točnost ($CA_{MPV} = 0.917$) doseže metoda podpornih vektorjev.

Metoda	Klasifikacijska točnost	Senzitivnost	Specifičnost
NB	0.859	0.892	0.920
kNN	0.823	0.890	0.872
MPV	0.917	0.954	0.973
NG	0.909	0,941	0.966

3.4 Kontekstno-odvisna izbira predvajane vsebine

Informacijski vmesniki omogočajo predvajanje vsebin na osnovi predvajalnih značilnic BF . Običajno upoštevajo lokacijo zaslona in vnaprej določen predvajalni seznam vsebin znotraj izbranih časovnih intervalov. Z adaptivnim informacijskim vmesnikom vpeljemo še uporabniške značilnice AF , ki so vezane na trenutno stanje na predvajalnem mestu in opisujejo časovne (čas prisotnosti, čas kontaktne možnosti, čas pozornosti), prostorske (razdalja, položaj) in demografske (spol, starostna skupina) lastnosti prisotnih uporabnikov. In prav zaradi vpeljave uporabniških značilnic je potrebno razširiti koncept predvajanja in izbire predvajane vsebine. Vsakemu predvajalnemu elementu določimo opisnik vsebine (angl. *content descriptor*), ki predstavlja množico predvajalnih in uporabniških značilnic. Opisnik vsebine CD definiramo v skladu z [3] kot:

$$CD = \{\{\alpha_i AF_i\} \cap \{\beta_j BF_j\} : i, j \geq 0\}, \quad (3.30)$$

kjer α_i in β_j označujejo uteži uporabniških in predvajalnih značilnic. Definiramo tudi opisnik stanja (angl. *state descriptor*), ki hrani značilnice vseh trenutno prisotnih uporabnikov pred zaslonom. Opisnik stanja SD opredelimo kot:

$$SD = \{AF_k : k \geq 0\}, \quad (3.31)$$

kjer k teče po vseh relevantnih uporabniških značilnicah prisotnih uporabnikov. Izbor naslednje predvajane vsebine, ki upošteva trenutni kontekst, definiramo kot minimalno razdaljo med opisnikom trenutnega stanja SD in opisnikom posamezne vsebi-

ne CD_j :

$$j_{next} = \arg \min_{j \leq n} dist(SD, CD_j), \quad (3.32)$$

kjer $dist$ predstavlja funkcijo razdalje, n pa število predvajalnih vsebin, ki jih imamo na voljo. Za funkcijo razdalje lahko uporabimo različne mere razdalje oziroma metode strojnega učenja.

3.5 Pravni vidiki in zasebnost

Ustrezno in transparentno ravnanje z zasebnimi podatki je ključnega pomena za zagotavljanje zaupanja med uporabniki in adaptivnimi sistemi za prikazovanje vsebine. Model upravljanja zasebnosti informacijskih vmesnikov mora preseči možnost posredne identifikacije uporabnika [63].

Če želimo, da sistem uporabniku prikaže relevantno vsebino, se mora prilagoditi uporabniku. Naša predstavljena arhitektura informacijskega vmesnika, ki za prilagajanje vsebine in interakcijo uporablja računalniški vid, upošteva načelo vgrajene zasebnosti (angl. *privacy-by-design*) in zagotavlja varno ter ustrezno ravnanje s podatki, ki bi omogočali osebno identifikacijo. Koncept vgrajene zasebnosti [21, 101, 172] določa, da je skrb za ustrezno ravnanje z zasebnimi podatki vključena v celoten proces načrtovanja in končno arhitekturo informacijskega sistema. Pri tem pa ob upoštevanju varovanja zasebnosti, ne zmanjšamo funkcionalnosti in učinkovitosti delovanja sistema.

Osebne podatke v našem sistemu predstavlja video posnetek dogajanja pred zaslonom. Procesiranje slike se izvaja v realnem času na samem predvajalnem mestu, zato ni potrebe, da bi se slikovni posnetki hranili oziroma pošiljali v obdelavo preko omrežja. Predvajalna enota zavrže vse slikovno gradivo takoj po njegovi obdelavi. Zabeležijo se le podatki gledanosti zaslona, ki se preko varne povezave shranjujejo na centralnem strežniku. Predlagana arhitektura tako omogoča zajem relevantnih podatkov za izvedbo posplošene statistične analize gledanosti zaslona, ne da bi tem hranili osebne podatke oziroma poizkušali identificirati posamezne uporabnike.



Dinamična anamofaza

Predlagamo novo prostorsko interaktivno metodo -t.i. dinamično anamorfozo-, ki se z uporabo dinamične anamorfične transformacije prilagaja premikanju uporabnika po prostoru [151]. V poglavju najprej predstavimo (statično) anamorfozo, nato pa koncept anamorfoze razširimo v dinamično anamorfozo, formaliziramo anamorfično transformacijo in predstavimo rezultate. Metoda se pokaže kot nova možnost za izboljšanje očesnega stika pri videokonferencah.

4.1 Anamorfoza

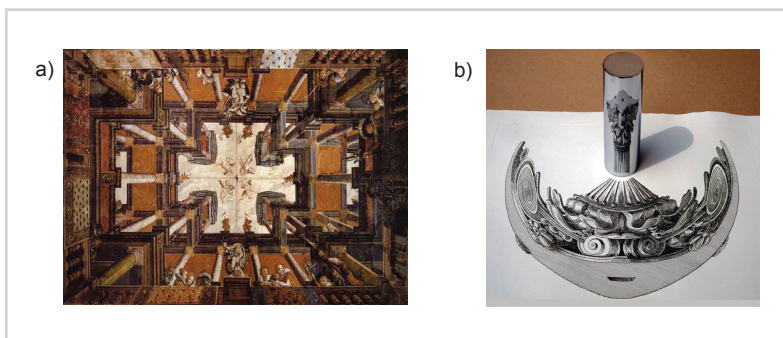
Ena od značilnosti človeškega zaznavanja je, da naši možgani stalno urejajo vizualne signale na pravilen, urejen in simetričen način, kot to utemeljujejo Gestalt načela zaznavne organizacije [89]. Tako na primer vidimo pravilne pravokotne oblike in podobe, kljub temu da so le-te večino časa popačene zaradi perspektivne projekcije in se stalno spreminjajo ob gibanju po prostoru. Opisani pojav imenujemo stalnost oblike (angl. *shape constancy*) [147].

Anamorfoza je izkrivljena projekcija, ki od gledalca zahteva, da za prikaz nedeformirane slike zavzame specifičen položaj v prostoru oziroma za-to uporabi posebno napravo. Če je prikaz nedeformirane slike odvisen od gledišča (angl. *viewpoint*) uporabnika govorimo o perspektivni anamorfozi (angl. *perspectival anamorphosis*). V kolikor pa za prikaz nepopačene slike potrebujemo specifičen predmet (npr. zrcala različnih oblik), ki odbija svetlobo, govorimo o odbojni oziroma zrcalni anamorfozi (angl. *catoptric anamorphosis*) [164]. Slika 4.1 prikazuje primer perspektivne in odbojne anamorfoze.

Perspektivna anamorfoza ali anamorfična projekcija sega v umetnost 15. stoletja, ki

Slika 4.1

Perspektivna in zrcalna anamorfoza. a) Primer perspektivne anamorfoze je tudi Celjski strop Stare Grofije v Celju. Poslikava ravne površine stropa z določenega gledišča tvori iluzionistični efekt, da se strop navidezno razširja proti nebu. b) Zrcalna anamorfoza. Madžarski umetnik István Orosz je za prikaz nedeformirane slike uporabil zrcalo cilindrične oblike.



temelji na uporabi linearne perspektive [32, 33]. Klasična linearna perspektiva temelji na Evklidski paradigmi in predpostavlja, da svetloba potuje v ravnih linijah, ki se odbijajo od objekta. Odbiti žarki sekajo projekcijsko ravnino in na njej pustijo natančno podobo osnovnega objekta. Slike običajno gledamo frontalno oziroma pod omejenim razponom zornih kotov. Anamorfične slike pa se gledalcu običajno pokažejo v nedeformirani obliki, ko se jih gleda pod izrazito ostrim zornim kotom. Če bi uporabnik anamorfično sliko v takšnem primeru gledal frontalno, bi bila tako popačena, da motiva ne bi prepoznal.

Verjetno najbolj znan primer anamorfoze v umetnostni zgodovini predstavlja slika *Ambasadorja*, nemškega slikarja Hansa Holbeina iz leta 1533 [185], predstavljena na Sliki 4.2. Primer anamorfične projekcije predstavljajo tudi deformirane slike oziroma oznake, ki se uporabniku prikažejo v nepopačeni obliki, ker so projicirane na nagnjeno površino. Na primer prometni znaki, ki so naslikani na cestno površino, so pogosto tudi v resnici bolj podolgovate oblike. Tako se z gledišča udeleženca v prometu, ki se znaku približuje, pokažejo v bolj pravilnem in prepoznavnem razmerju. Primer talne označbe, ki je prilagojena pričakovanemu zornemu kotu gledalca, je prikazan na Sliki 4.3. Anamorfoza pa se uporablja tudi za namene oglaševanja. Virtualno oglaš-



Slika 4.2

Hans Holbein, *The Ambassadors*, 207x209cm, National Gallery, London. Diagonalna lisa na dnu slike se pokaže kot človeška lobanja, če sliko gledamo od desno zgoraj [185].

vanje omogoča vstavljanje oglasnih sporočil, kot so npr. logotipi blagovnih znamk in promocija produktov, v prenos ali posnetek televizijskega programa. Predvsem pri prenosu športnih dogodkov, ki so obdani s številnimi oglasnimi panoji, se njihova vsebina lahko zamenja in prekrije z novo [190]. Prekrivajočo sliko je potrebno deformirati, da se natančno prilaga panoju, istočasno pa upoštevati gledišče kamere. S tem se ustvari efekt, da je predmet na sliki dejansko prisoten kot del scene. Vsebina sporočil pa se tako lahko prilagaja potrebam različnih oglaševalskih segmentov in trgov po svetu.



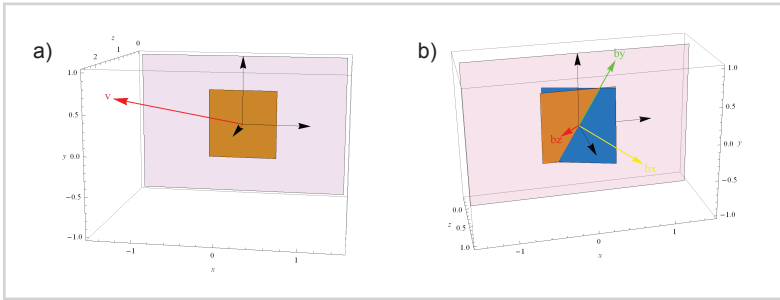
Slika 4.3

a) Frontalni pogled talne oznake iz katerega je razvidno, da je oblika kolesa podolgovata. b) Pogled na isto talno oznacbo z pričakovanega gledišča. Oznaka kolesa se prikaže v bolj pravilnem razmerju.

4.2 Dinamična anamorfoza

Za prikaz nedeformirane statične anamorfične slike mora uporabnik zavzeti specifično gledišče in sliko gledati z vnaprej določene smeri. Vendar pa za prikaz anamorfično deformirane slike lahko uporabimo video projektor oziroma večji zaslon, kar da popolnoma novo modalnost, namreč da anamorfično sliko prilagajamo položaju uporabnika. Z gibanjem uporabnika po prostoru se stalno dinamično izračunava anamorfična transformacija in ustrezno prilagodi deformacija slike, tako da na novo deformirana slika za opazovalca učinkovito ostaja enaka. Opisan koncept, ki razširja perspektivno anamorfozo s prilagajanjem deformiranja slike položaju uporabnika, imenujemo dinamična anamorfoza.

Za doseganje stalnosti slike mora sistem v realnem času spremljati položaj opazovalca in izračunati anamorfično transformacije osnovne slike prilagojeno trenutnemu gledišču uporabnika. Dinamična anamorfoza je prostorsko-ozaveščen (angl. *position-aware*) vmesnik, ki združuje lokalizacijo uporabnika in krivljenje slike, ki se prilagaja



Slika 4.4

Dinamična anamorfoza. a) Gledišče opazovalca \mathbf{v} je označeno z rdečo, osnovna nedeformirana slika pa z rjavo barvo. b) Modra ploskev prikazuje transformirano osnovno sliko na navidezni ravnini, ki jo napenjata vektorja \mathbf{b}_y in \mathbf{b}_x . Vektor \mathbf{b}_z je pravokoten na navidezno slikovno ravnino in kaže v smeri gledišča uporabnika.

spreminjajočemu se položaju uporabnika. Ocenjevanje 3D položaja uporabnika smo opisali v razdelku 3.2.4, zato v nadaljevanju opišemo še postopek za anamorfično deformacijo slike.

Za dinamično prilagajanje perspektivne anamorfoze formaliziramo anamorfično transformacijo. Predpostavimo, da se slika I prikazuje na slikovni ravnini, ki leži v ravnini $z = 0$. Izhodišče koordinatnega sistema postavimo v središče slikovne ravnine. Slika I , ki zajema dimenzije $I_w \times I_h$ slikovnih točk, ima pred transformacijo središče poravnano z izhodiščem koordinatnega sistema in se razteza po definiciji normirano od -0.5 do 0.5 v x in y smeri.

Iščemo transformacijo, ki bo transformirala slikovno ravnino ($z = 0$) na tak način, da bo vektor, ki povezuje medočesno točko uporabnika in središče projicirane slike na navidezni slikovni ravnini pravokoten na transformirano ravnino (glej Sliko 4.4). Anamorfična transformacija zajema dva koraka. Osnovno sliko najprej transformiramo na navidezno slikovno ravnino, katere normala je vzporedna vektorju medočesne točke uporabnika. Navidezno sliko nato projiciramo na realno slikovno ravnino in tako za uporabniku ustvarimo efekt, kot da bi bila prikazana na navidezni slikovni ravnini (glej Sliko 4.5).

Postopek anamorfične transformacije in projekcije formaliziramo v sledeče zaporedje korakov:

1. Vektor gledišča uporabnika $\mathbf{v} = (v_x, v_y, v_z)$ normiramo in označimo z \mathbf{b}_z .
2. Izračunamo $\mathbf{b}_z \times (0, 0, 1)$, kjer \times predstavlja vektorski produkt. Dobljeni vektor normiramo in ga označimo z \mathbf{b}_y . Vektor \mathbf{b}_y definira os, okoli katere je potrebno zavrteti sliko.

3. Definiramo vektor \mathbf{b}_x kot $\mathbf{b}_y \times \mathbf{b}_z$.
4. Sedaj lahko konstruiramo matriko \mathbf{R} , ki bo poravnala vektor \mathbf{b}_y z vektorjem $(0, 1, 0)$ kot:

$$\mathbf{R} = \begin{bmatrix} \mathbf{b}_x \\ \mathbf{b}_y \\ \mathbf{b}_z \end{bmatrix} \quad (4.1)$$

Vektorje \mathbf{b}_x , \mathbf{b}_y in \mathbf{b}_z si v tem primeru predstavljamo kot podmatrike dimenzije 1×3 .

5. V naslednjem koraku konstruiramo matriko \mathbf{P} , ki bo zavrtila koordinatni sistem okrog osi \mathbf{b}_y za enak kot, kot je med vektorjema $(0, 0, 1)$ in \mathbf{b}_z :

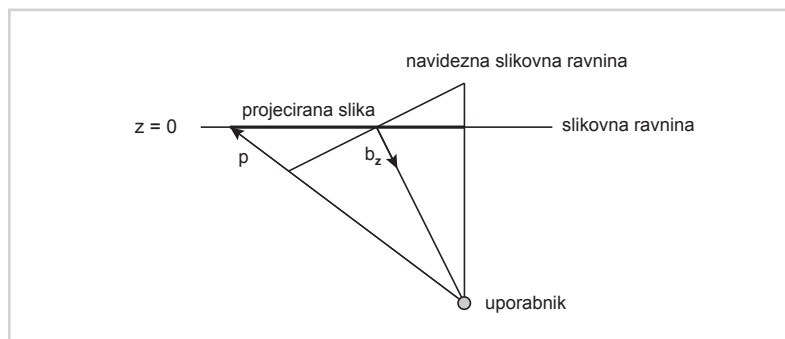
$$\mathbf{P} = \begin{bmatrix} c & 0 & -s \\ 0 & 1 & 0 \\ s & 0 & c \end{bmatrix} \quad (4.2)$$

kjer je $c = \mathbf{b}_z \cdot (0, 0, 1)$ in $s = (0, 0, 1) \cdot \mathbf{b}_x$.

6. Potrebujemo še matriko, ki bo vektor $(1, 0, 0)$ poravnala nazaj z \mathbf{b}_x . Omenjeno operacijo izvede inverz matrike \mathbf{R} . Matrika \mathbf{R} je rotacijska matrika, zato njen inverz izračunamo kar kot njeno transponirano vrednost; velja $\mathbf{R}^{-1} = \mathbf{R}^T$.
7. Definiramo anamorfično matriko \mathbf{T} , ki z množenjem omenjenih treh matrik transformira osnovno sliko na navidezno ravnino: $\mathbf{T} = \mathbf{R}^T \cdot \mathbf{P} \cdot \mathbf{R}$.
8. Naj $\mathbf{i} = (i_x, i_y, i_z)$ predstavlja koordinate slikovne točke slike I . Položaj vsake slikovne točke na navidezni slikovni ravnini lahko sedaj izračunamo z množenjem njenega vektorja \mathbf{i} z matriko \mathbf{T} . Na osnovi načela podobnih trikotnikov za projekcijo točke z navidezne slikovne ravnine nazaj na realno slikovno ravnino uporabimo vektor, ki poteka od gledišča uporabnika do $\mathbf{T} * \mathbf{i}$ in se nadaljuje do realne slikovne ravnine. Projekcijo točke \mathbf{p} izračunamo kot:

$$\mathbf{p} = (\mathbf{T} * \mathbf{i} - \mathbf{v}) \frac{v_z}{v_z - t_z}, \quad (4.3)$$

kjer t_z označuje z komponento vektorja $\mathbf{T} * \mathbf{i}$ (glej Sliko 4.5). Položaj projekcijskih točk v koordinatnem sistemu realnega sveta lahko izračunamo kot $\mathbf{v} + \mathbf{p}$.



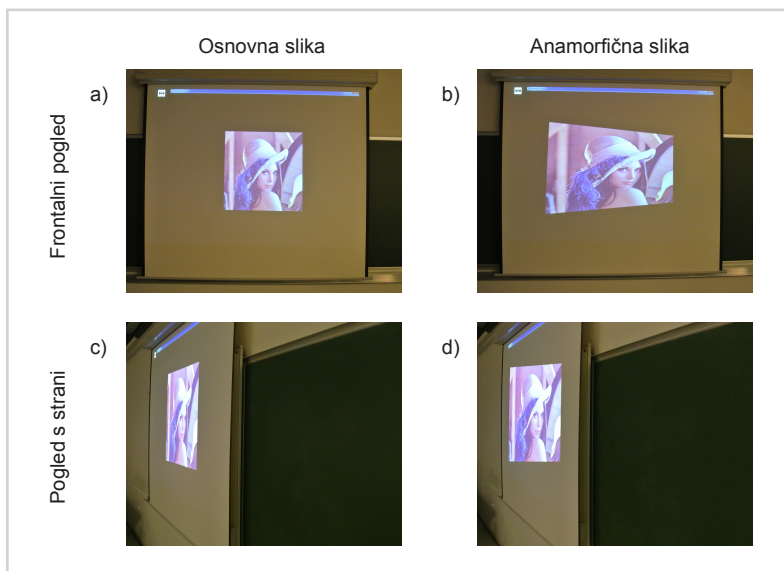
Slika 4.5

Schema formalizacije anamorfične transformacije. Prikazan je shematski pregled v ravnini $x-z$.

Implementacijo anamorfične transformacije smo izvedli z uporabo knjižnice OpenCV [9, 20]. Gre za odprtokodno knjižnico, ki ponuja širok nabor metod in algoritmov računalniškega vida in omogoča realnočasno procesiranje video posnetkov. Za testiranje smo uporabili računalnik z procesorjem Intel Core2 Quad Q9000 (2.00 GHz), 4GB delovnega pomnilnika in grafično kartico NVidia Quadro FX 2700M. Na dani konfiguraciji lahko z anamorfično transformacijo prilagajamo video HD ločljivosti v realnem času.

Rezultate metode smo določili tudi eksperimentalno. Izbrali smo dve gledišči: frontalno gledišče in gledišče postavljeno relativno v desno glede na središče zaslona. Slika 4.6a prikazuje osnovno sliko, ki jo vidi uporabnik iz frontalnega gledišča. Osnovna, netransformirana slika z stranskega gledišča pa je predstavljena na Sliki 4.6c. Osnovno sliko transformiramo in prilagodimo glede na gledišče na desni. Slika 4.6b prikazuje frontalni pogled na transformirano sliko. Slika je podaljšana v horizontalni smeri, saj je prilagojena na gledišče postavljeno v desno. Slika 4.6d pa prikazuje pogled na transformirano sliko z gledišča na desni.

Razvita prostorska metoda omogoča nadaljnje aplikacije. Med njimi izpostavimo možnosti interaktivnih prostorov, npr. inteligentne nadzorne sobe, kjer se vizualizacija trenutnega stanja sistema in meritev stalno prilagaja položaju uporabnika. Digitalna anamorfoza omogoča tudi aplikacijo v interaktivnih umetniških postavitvah, kjer projicirana slika oziroma video posnetek spremlja uporabnika v enaki obliki ne glede na njegov položaj v prostoru. Kot prostorska metoda omogoča inovativne pristope v oglaševanju. Možna uporaba pa zajema tudi prilagajanje videokonferenčne slike z namenom izboljšanja očesnega stika, kar bolj podrobno predstavimo v naslednjem razdelku.

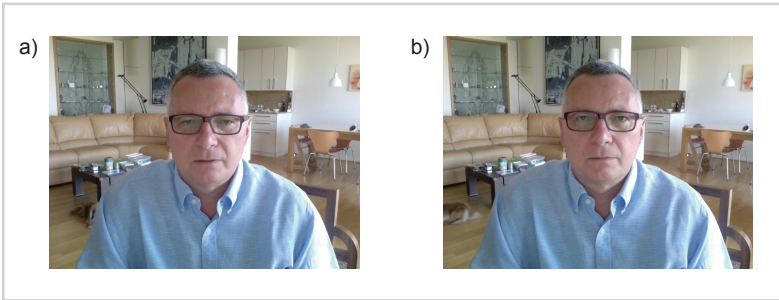


Slika 4.6

Rezultati dinamične anamorfoze. a) Osnovna slika, kot jo vidi opazovalec iz frontalnega gledišča. b) Transformirana slika kot jo vidi gledalec iz frontalnega gledišča. c) Osnovna slika kot jo vidi opazovalec iz strani. d) Transformirana slika kot jo vidi opazovalec iz strani.

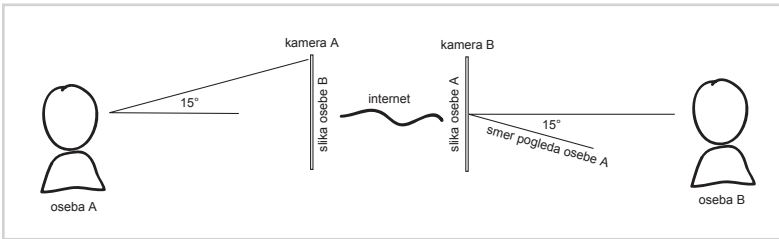
4.3 Aplikacija - Izboljšanje očesnega stika pri videokonferencah

Videokonference omogočajo zvočno in slikovno komunikacijo med ljudmi. Z uporabo telekomunikacijskega omrežja lahko sodelujoči komunicirajo med seboj sicer krajevno oddaljenih lokacij. Odprto vprašanje, ki se nanaša na uporabniško izkušnjo ob uporabi videokonference, je izguba očesnega stika med sodelujočimi [64, 68, 98]. Problem je izrazit pri komunikaciji med dvema posameznikoma, saj se pomanjkanje očesnega stika inherentno povezuje z prikritimi nameni in posledično zmanjšuje stopnjo zaupanja med sogovorniki [12]. Posameznikove oči so med videokonferenco običajno usmerjene v sredino računalniškega zaslona, kjer je prikazan obraz sogovornika, kamera pa je običajno postavljena nad zaslonom. Študije so pokazale, da ljudje manj zaupajo komunikaciji preko video povezave, kjer je kamera postavljena nad zaslonom, kot zgolj zvočni komunikaciji ali komunikaciji preko elektronske pošte [12]. Stokes pokaže, da je izguba očesnega kontakta opazna, če vertikalni kot med kamero in očmi ter smerjo pogleda v zaslon presega 5° [177]. Pri uporabi povprečno velikega računalniškega zaslona na normalni razdalji gledanja je ta kot običajno med 15° in 20° , kar ima



Slika 4.7

Problem manjkajočega očesnega stika pri videokonferencah. Obe sliki uporabnika sta bile zajeti s kamero nameščeno nad 27 palčni zaslon pri normalni delovni razdalji. a) Uporabnik gleda v središče zaslona. b) Uporabnik gleda naravnost v kamero.



Slika 4.8

Oseba A gleda osebo B na računalniškem zaslonu pred seboj (levo). Kamera A je postavljena nad računalniškim zaslonom, zato oseba B pri spremljanju slike osebe A ne ujame sogovornikove smeri pogleda (desno).

za posledico izgubo očesnega stika med sogovornikoma [198]. Opisana problematika je predstavljena na Sliki 4.7 in Sliki 4.8.

Podoben problem nastopi, če se obraz našega sogovornika prikazuje v majhnem oknu na zaslonu večjih dimenzij in je kot med sogovornikovim obrazom v oknu in položajem kamere večji od 5° . Uporabniki mobilnih naprav običajno nimajo teh težav, saj je kot med središčem zaslona in uporabnikovim obrazom običajno manjši od 5° .

Problem manjkajočega očesnega stika je že dolgo znan. Na začetku so bile predlagane številne rešitve, ki so temeljile na specialni strojni opreми kot npr: uporaba polovičnega ogledala (angl. *half mirror*), delilcev žarka (angl. *beam splitters*) in kamere, integrirane v središče zaslona. Tovrstne rešitve naslavlajo problem tako, da želijo poravnati os kamere s središčno osjo računalniškega zaslona. Zahtevajo specializirane sisteme, ki pa so dragi in večinoma na splošno nedostopni.

S povečevanjem računske zmogljivosti osebnih računalnikov so bili predlagani programski pristopi, ki zajemajo manipulacijo slike sogovornika [64]. Za generiranje navideznega pogleda s sredine zaslona so bile predlagane metode, ki uporabljajo konfiguracijo od dveh [140] pa vse do osem [51] kalibriranih kamer. Opisani sistemi delujejo

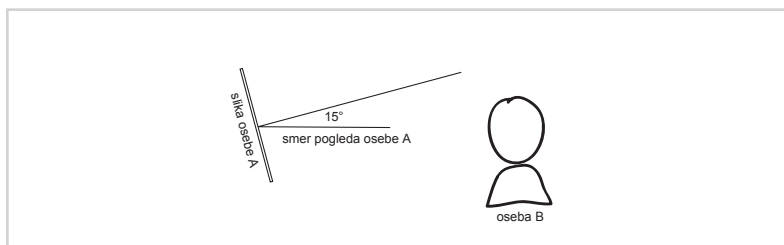
na osnovi stereo ujemanja in preobrazbe slike sogovornika (angl. *image morphing*). Predlagan je bil tudi sistem z eno kamero, ki z afino transformacijo prilagodi določene regije obraza in oči [197].

Uporabniki danes uporabljajo videokonferenčne programe na namiznih in prenosnih računalnikih, zato se tovrstne rešitve izkažejo za prekompleksne. Microsoftovi raziskovalci so prvi poročali o neformalnih opažanjih, da z rotacijo slike lahko vplivamo na dojetje smeri pogleda sogovornika, tudi ko le-ta ne gleda naravnost v zaslon. Tehnično poročilo sklenejo z ugotovitvijo, da so za določitev povezave med rotacijo in zaznavanjem smeri pogleda potrebne nadaljnje raziskave [206]. Njihova opažanja smo povezali s konceptom dinamične anamorfne transformacije in razvili metodo, ki uporabnikoma videokonference omogoča izboljšano dojetje očesnega stika [151, 170].

Z uporabo anamorfne transformacije sliko sogovornika rotiramo pod zmernim kotom okoli osi x . Slika se tako deloma deformira. Stopnja deformacije zavisi od rotacijskega kota, ki pa je odvisen od konfiguracije videokonference (postavitev kamere, razdalja med uporabnikom in zaslonom). Shemo transformacije prikazuje Slika 4.9. Deformirano sliko našega sogovornika dojemamo podobno kot osnovno, nedeformirano sliko, saj naša vizualna percepcija sama poskrbi za korekcijo perspektivne deformacije. Zanimalo pa nas je, ali takšna transformacija vpliva tudi na dojetje smeri pogleda in posledično očesnega stika.

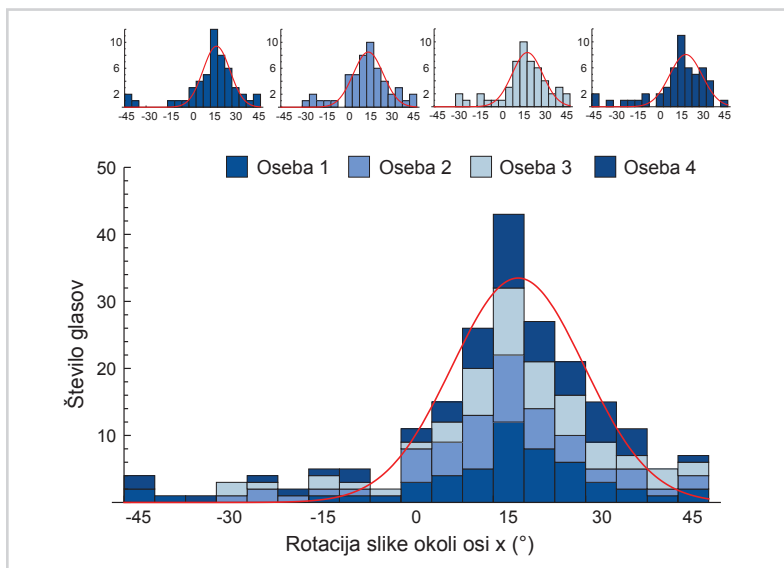
Slika 4.9

Z transformacijo slike osebe A se spreminja tudi dojetje smeri pogleda osebe B. Glede na eksperimentalne rezultate se uporabniška izkušnja medočesnega stika izboljša.



Preizkus hipoteze smo izvedli eksperimentalno. Zajeli smo slike 4 različnih oseb, ki so sedele pred 27 palčnim zaslonom in gledale v sredino računalniškega zaslona (primer na Sliki 4.7a). Kamera, ki je zajela slike, je bila postavljena nad zaslonom. Razvili smo spletno aplikacijo, ki uporabniku omogoča, da sam interaktivno izbere kot anamorfne transformacije, ki mu najbolj odgovarja. Ob izbiri kota se deformirana slika sproti posodablja, kar omogoča lažjo primerjavo in iskanje optimalne vrednosti kota. Izbira

kota je bila omejena na interval $[-45^\circ, 45^\circ]$, z začetnim stanjem 0° . Vzorec sodelujočih zajema 54 oseb, med njimi so bili večinoma dodiplomski in podiplomski študenti. Sodelujoče smo prosili naj z manipulacijo slike izberejo tisto, pri kateri zaznajo najboljši očesni stik z osebo na sliki. Čas za izbiro vrednosti ni bil omejen. Dva sodelujoča sta podala oceno, da nobena transformacija slike ne izboljša očesnega stika. Rezultati preostalih 52 udeležencev pa so predstavljeni na Sliki 4.10.



Slika 4.10

Rezultati eksperimenta v katerem so sodelujoči iskali vrednost transformacijskega kota ob katerem so zaznali najboljši medočesni stik z osebo na sliki. 20% vseh glasov sodelujočih zajema vrh histograma pri 15° . Skoraj polovica vseh glasov leži znotraj intervala med 10° in 20° .

Skupni podatki zajemajo 208 izbranih kotov transformacije, ki jih je 52 udeležencev podalo za 4 različne primere. Zgornji (manjši) histogrami na Sliki 4.10 prikazujejo histogram izbranih vrednosti za vsak primer posebej, osrednji histogram pa zajema kumulativne podatke za vse 4 primere skupaj. Rdeča črta označuje interpolacijo podatkov na osnovi Gaussove funkcije. Povprečna vrednost kota v eksperimentu znaša $13,2^\circ$, mediana $15,3^\circ$, standardni odklon $17,8^\circ$, povprečje standardnih odklonov posameznega sodelujočega pa $11,6^\circ$.

Vrh histograma pokaže, da je 20% (43) glasov sodelujočih pri vrednosti kota 15° določilo optimalno deformacijo slike. Skoraj polovica glasov, tj. 42% (88), zajema vrednost med 10° in 20° in več kot 63% (132) kot med 5° in 25° . Histogram pokaže

izrazit odklon zabeleženih vrednosti v smer pozitivne vrednosti kota transformacije. Zvonasta oblika in jasen vrh histograma pa nakazujeta, da je večina sodelujočih zaznala optimum medočesnega kontakta pri kotu transformacije $\sim 15^\circ$. Zanimivo, da ta vrednost ustreza kotu med smerjo pogleda in kamero nad zaslonom, če uporabnik sedi pred običajno velikim računalniškim zaslonom pri oddaljenosti ~ 60 cm in gleda v središče zaslona. Primer videokonferenčne slike uporabnika na Sliki 4.7a, katero smo glede na rezultate zbranih glasov optimalno korigirali z anamorfično transformacijo, torej zavrteli za 15° v smeri x , je prikazan na Sliki 4.11.



Slika 4.11

Sliko 4.7a, ki prikazuje uporabnika videokonference, smo z anamorfično transformacijo zavrteli za 15° okoli osi x . Glede na eksperimentalno zajete rezultate, takšna transformacija optimalno izboljša medočesni stik.

4.4 Aplikacija - Interaktivna instalacija

Koncept dinamične anamorfoze smo razvili tudi v kontekstu interaktivne instalacije. Osnovo za projiciranje slike predstavlja človeški obraz, z pogledom usmerjenim naravnost v kamero. Uporabimo kratek izsek iz filma *1984* (glej Sliko 4.12), posnetega po istoimenskem romanu pisatelja George-a Orwella, objavljenega v letu 1949. Interaktivna instalacija predvideva video projekcijo, kjer se pogled Velikega brata neprestano prilagaja položaju uporabnika v prostoru ter s tem ustvarja občutek stalnega in vseprisotnega nadzora [34].

Na projekcijskem mestu se naenkrat nahaja zgolj en uporabnik, kateremu se prilagaja transformacija posnetka. Klasična anamorfoza predvideva, da uporabnik zavzame specifično gledišče. Opisana projekcija pa uporabi dinamično anamorfično transformacijo in izzove intuitivni občutek za prostorsko orientacijo, saj uporabnik kljub premikanju

po prostoru, ves čas vidi enako sliko, prilagojeno trenutnemu gledišču.



Slika 4.12

Anamorflična transformacija slike Velikega brata iz prepoznavnega filma 1984, posnetega po istoimenski knjigi pisatelja George-a Orwella, kjer se pogled Velikega brata neprestano prilagaja položaju uporabnika v prostoru.



*Merjenje in kvantitativna
analiza gledanosti*

Med motivacijo v razdelku 1.2 smo predstavili predloge številnih avtorjev po izvedbi kvantitativne študije [81, 131, 133], ki bi analizirala dejansko obnašanje uporabnikov pred digitalnimi zasloni in ne zgolj njihovega dojemanja le-tega. V nadaljevanju opišemo eksperiment, ki smo ga izvedli z razvitim adaptivnim informacijskim vmesnikom v realnem okolju, kjer -kot eni svetovno prvih- izmerimo in kvantitativno analiziramo gledanost vmesnika. Eksperimentalne podatke statistično ovrednotimo in na podlagi analize interakcij med atributi predstavimo prvo kvantitativno študijo gledanosti zaslonov [152].

5.1 Eksperiment v realnem okolju

Razviti sistem smo uporabili za izvedbo eksperimenta v realnem okolju. Za prikazovanje in obdelavo slike smo uporabili 24 palčni računalniški zaslon Sony Vaio VP-CL135FX/B z že vgrajeno strojno opremo za procesiranje (4-jedrni procesor Intel Q8400S, 6 GB pomnilnika). Zajem slike se je vršil preko kamere Logitech WebCam Pro 9000.

Eksperiment je potekal v butični trgovini v središču Ljubljane. Večino tlorisa je zajemal osrednjih nakupovalni del ($\sim 35\text{m}^2$), ki se je nahajal med vhodom in blagajniškim pultom. Za blagajniškim pultom je bil dodaten manjši prostor za pomerjanje oblačil. Za izvedbo eksperimenta smo namenoma izbrali manjšo trgovino, saj je lahko ena kamera pokrila celotno nakupovalno površino. Omenimo še, da glavni asortiment trgovine zajema modna in športna oblačila višjega cenovna razreda, kar lahko posredno vpliva na zajete demografske in vedenjske meritve.

Cilj eksperimenta je bila izvedba kvantitativne študije gledanosti digitalnega zaslona. Pri postavitvi zaslona in pripravi predvajanih vsebin smo optimizirali kriterije, ki po [81, 131] maksimizirajo čas pozornosti obiskovalca. Za optimalni položaj zaslona smo izbrali polico poleg blagajniškega pulta. Zaslon je bil obrnjen proti vhodu in se nahajal na višini oči, kar je omogočalo, da je bil viden s celotnega nakupovalnega dela trgovine. Nekaj tipičnih posnetkov eksperimenta prikazuje Slika 5.1.

Pomemben kriterij, ki vpliva na pozornost gledalcev, je prikazovana vsebina. Zanimalo nas je, kako različni tipi vsebine vplivajo na vedenje ljudi, zato smo pripravili dva različna scenarija prikazovanja. Prvi scenarij je zajemal prikazovanje t.i. *statične vsebine*. Pripravili smo računalniško predstavitev, ki je zajemala 20 prosojnic. Prosojnice so večinoma prikazovale slike asortimenta trgovine ter nekatere prepoznavne osebnosti, ki so bila vključena v promocijo te znamke oblačil. Časovni interval prikazovanja med

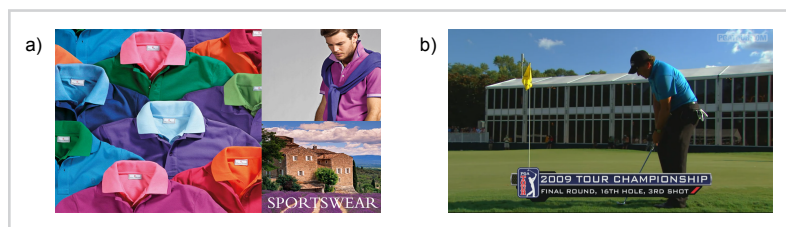


Slika 5.1

Eksperiment z razvitim adaptivnim informacijskim vmesnikom v realnem okolju. (a) Tipičen posnetek trgovine, v kateri je potekal eksperiment. (b) Uporabnik, ki gleda blago trgovine in ima možnost opaziti vsebino na zaslonu. (c) Uporabnik, ki je usmeril svojo pozornost na predvajano vsebino. (d) Sledenje aktivne regije s uporabo odstranjevanja ozadja.

prosojnicami je bil 10 sekund. V drugem scenariju smo želeli spremljati odziv gledalcev na *dinamično vsebino*. Ta scenarij je zajemal tri video posnetke v skupni dolžini ~ 30 minut. Posnetki so prikazovali izseke iz različnih športnih aktivnosti, povezanih z ponudbo trgovine. Pri pripravi vsebin za oba scenarija smo po najboljših močeh upoštevali tudi kriterij živahnih barv (angl. *colorful criterion*), čustvenega naboja (angl. *emotional criterion*) in estetike (angl. *aesthetic look criterion*). Primer vsebine za oba scenarija predstavlja Slika 5.2.

Širši pogled na ustrezno izvajanje video nadzora javnih površin z uporabo predstavljene metodologije smo predstavili v razdelku 3.5. Skupaj z upravljalcem trgovine smo zagotovili, da je celoten eksperiment potekal skladno z Zakonom o varstvu osebnih podatkov in ustrezal ostalim veljavnim zakonsko predpisanim normam, ki urejajo področje varovanja zasebnosti v Republiki Sloveniji [202]. Obiskovalci trgovine so bili o izvajanju ustrezno obveščeni že pred vstopom v trgovino.



Slika 5.2

Scenarija prikazovanja vsebine uporabljena pri eksperimentu v realnem okolju. (a) Statični tip vsebine, sestavljen iz 20 prosojnic. (b) Dinamični tip vsebine, zajema video posnetke s športno vsebino, ki se nanašajo na asortiment trgovine.

5.2 Meritve in verifikacija rezultatov

Eksperiment smo izvajali 23 dni ter analizirali 214 ur video vsebine. Z uporabo sistema smo zajeli obnašanje in odziv gledanosti 1306 oseb. Pred nadaljnjo analizo sta vse avtomatsko pridobljene časovne in demografske podatke verificirala dva človeška ocenjevalca, kar smo v nadaljevanju lahko uporabili za določitev natančnosti delovanja vmesnika.

Pri anotaciji demografskih podatkov (spol in starost) je v nekaterih primerih prišlo do odstopanj med ocenjevalcema. V tem primeru sta si ocenjevalca izmenjala argumente in konsenzno določila končno vrednost značilnice. Za prikaz usklajenosti med dvema človeškima ocenjevalcema -že pred konsenzno argumentacijo- smo uporabili mero Cohen-ova kappa κ , ki predstavlja koeficient skladnosti med ocenjevalcema vrednosti kategorične spremenljivke [26]. Koeficient definiramo kot:

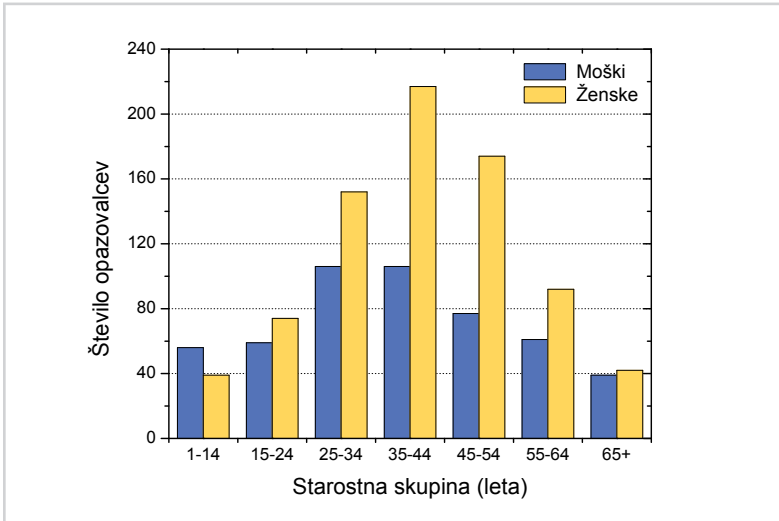
$$\kappa = \frac{Pr(a) - Pr(e)}{1 - Pr(e)}, \quad (5.1)$$

kjer $Pr(a)$ predstavlja relativno strinjanje med ocenjevalcema, $Pr(e)$ pa verjetnost naključnega strinjanja, izračunana na podlagi ocen vsakega ocenjevalca posebej. Mera κ zavzame vrednost 1 v kolikor sta ocenjevalca popolnoma usklajena in 0, v kolikor skladnost med ocenjevalcema ne presega verjetnosti naključnega strinjanja. Pri verifikaciji podatkov zajetih v eksperimentu določimo $\kappa_{spol} = 1.0$ pri klasifikaciji spola in $\kappa_{starost} = 0.91$ pri klasifikaciji v starostne skupine.

V fazi predobdelave smo iz podatkov izločili tiste primere, ki so se nanašali na zaposlene v trgovini, kot tudi podatke 12 osamelcev (angl. *outliers*), katerih izmerjene časovne vrednosti so za več kot 30-krat presegle povprečje. Vse analize v nadaljevanju poglavja se tako nanašajo na vzorec, ki zajema $N = 1294$ oseb.

Natančnost naučenih klasifikatorjev smo primerjali z anotiranimi podatki ocenjevalcev. Primerjava pokaže 86.6% klasifikacijsko natančnost pri klasifikaciji spola in 77.1% klasifikacijsko natančnost pri klasifikaciji gledalcev v starostne skupine. Procesiranje zajete slike z uporabo dveh jeder procesorja Intel Q8400 (2.66 GHz) je potekalo na ~21 FPS, kar omogoča spremljanje dogajanja pred zaslonom in prilagajanje vsebine v realnem času.

Demografska analiza pokaže strukturo obiskovalcev trgovine, v kateri je potekal eksperiment. Od skupno 1294 oseb jih je bilo 790 (61%) ženskega spola, preostalih 504 (39%) pa so bili moški. Starost opazovalcev obravnavamo kot intervalno spremenljivko



Slika 5.3

Distribucija opazovalcev v eksperimentu razdeljena po demografskih značilnicah (spol in starostna skupina). Primerjava po spolu pokaže, da je trgovino v času eksperimenta obiskalo več žensk (61%) kot moških (39%). Zastopanost opazovalcev glede na njihovo starostno skupino pokaže izrazit vrh v starostnem razredu 35–44 let (25%).

in jo razdelimo v 7 intervalnih razredov. Zastopanost znotraj posameznega intervala je bila: 95 (7%) oseb v skupini 1–14 let, 133 (10%) oseb v skupini 15–24 let, 258 (20%) oseb v skupini 25–34 let, 323 (25%) oseb v skupini 35–44 let, 251 (19%) oseb v skupini 45–54 let, 153 (12%) oseb v skupini 55–64 let in 81 (6%) oseb v skupini 65 in več let. Celotno strukturo, razdeljeno po spolu in starosti prikazuje Slika 5.3.

5.3 Rezultati analize gledanosti

Vsako od zajetih časovnih značilnic predstavimo z opisno statistiko, ki zajema povprečno vrednost μ_x , mediano M_x in standardno deviacijo σ_x . Izvedemo tudi statistično analizo za preverjanje signifikantnosti vpliva demografskih značilnic in tipa prikazovane vsebine. Za preverjanje oblike porazdelitve izračunamo koeficient asimetrije unimodalno porazdeljenih spremenljivk γ_1 kot:

$$\gamma_1 = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \mu_x)^3}{\left(\frac{1}{N} \sum_{i=1}^N (x_i - \mu_x)^2\right)^{3/2}}, \quad (5.2)$$

kjer x_i predstavlja i -to vrednost v vzorcu. Koeficient asimetrije (angl. *skewness*) lahko zavzame vrednosti intervala $(-\infty, +\infty)$. Za normalno porazdelitev velja asimetrija

$\gamma_1 = 0$, oddaljevanje od 0 pa kaže na naraščajočo asimetričnost porazdelitve. Če je vrednost γ_1 negativna govorimo o pozitivni (levi) asimetričnosti, ko pa je $\gamma_1 > 0$ pa o negativni (desni) asimetričnosti porazdelitve.

Poleg asimetrije preverimo tudi koeficient sploščenosti porazdelitve γ_2 . Sploščenost (angl. *excess kurtosis*) za vzorec N vrednosti izračunamo kot:

$$\gamma_2 = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \mu_x)^4}{\left(\frac{1}{N} \sum_{i=1}^N (x_i - \mu_x)^2\right)^2} - 3 \quad (5.3)$$

in lahko zavzame vrednosti $[-2, \infty)$. Sploščenost primerjamo s sploščenostjo normalne porazdelitve, za katero velja da je $\gamma_2 = 0$.

Pri statistični analizi zajetih podatkov smo uporabili Kruskal-Wallisov H test. Gre za neparametrični statistični test, ki ugotavlja statistično značilne razlike med dvema ali več skupinami [96, 173]. Metoda ne predpostavlja normalne porazdelitve vzorcev populacije. Postopek izračuna poteka tako, da podatke celotnega vzorca uredimo v enotno ranžirno vrsto, nato pa izračunamo vrednost statistike H kot:

$$H = (N - 1) \frac{\sum_{i=1}^g n_i (\bar{r}_i - \bar{r})^2}{\sum_{i=1}^g \sum_{j=1}^{n_i} (r_{ij} - \bar{r})^2}, \quad (5.4)$$

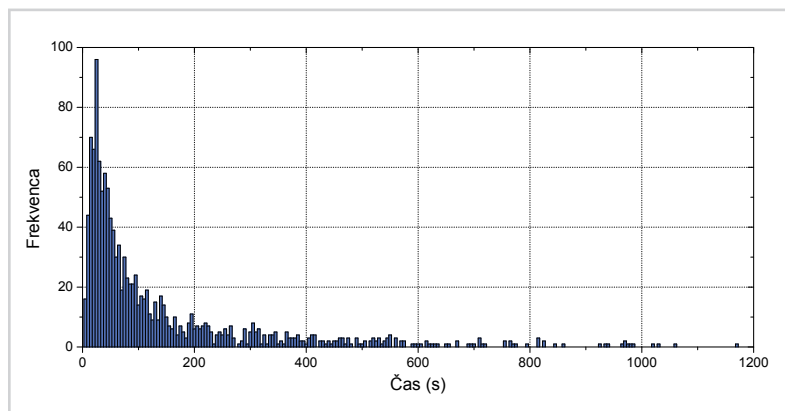
kjer g predstavlja število skupin v vzorcu, n_i število enot v skupini i , r_{ij} rang enote j v skupini i znotraj celotnega vzorca, $\bar{r}_i = \frac{\sum_{j=1}^{n_i} r_{ij}}{n_i}$ predstavlja povprečni rang skupine i in $\bar{r} = \frac{1}{2}(N + 1)$, ki predstavlja povprečno vrednost ranga vseh enot. Nato p-vrednost aproksimiramo kot $Pr(\chi_{g-1}^2 \geq H)$. Metodo smo izbrali, ker omogoča preizkus na več kot dveh skupinah (pri primerjavi starostnih skupin jih imamo 7) in ker ne predpostavlja normalne porazdelitve vzorca, kar smo opazili v zajetih podatkih.

5.3.1 Čas prisotnosti

Čas prisotnosti označuje skupni čas, ki ga uporabnik porabi na področju informacijskega vmesnika. V primeru eksperimenta čas prisotnosti uporabnika $T_{d,i}$ predstavlja skupni čas, ki ga je oseba preživela v vidnem polju kamere informacijskega vmesnika na osrednjem nakupovalnem prostoru trgovine. Nekateri obiskovalci so v okviru obiska trgovine osrednji prostor začasno zapustili ter se nato vrnili.

Iz eksperimentalno zajetih podatkov izračunamo povprečni čas prisotnosti obiskovalca, ki znaša $\bar{T}_d = 144s$, mediano $M_d = 64.8s$ in standardni odklon $\sigma_d = 198s$.

Negativna asimetrija ($\gamma_1 = 2.83$) in koničasta oblika ($\gamma_2 = 10.2$) porazdelitve nakazujeta izrazito odstopanje zajetih podatkov od normalne porazdelitve. Podrobno porazdelitev časa prisotnosti prikazuje histogram z velikostjo koša $b_w = 5s$ na Sliki 5.4.



Slika 5.4

Porazdelitev časa prisotnosti ($b_w = 5s$). Povprečni čas prisotnosti 1294 oseb znaša 144s in mediano 64.8s. V dveh minutah od vstopa je trgovino zapustilo 873 (68%) obiskovalcev. Oseb, ki so v trgovini ostale 5 minut ali več je bilo 189 (15%) in 53 (4%) takšnih, ki so v trgovini preživeli 10 minut in več.

Opisna statistika časa prisotnosti in rezultati statistične analize Kruskal-Wallis (K-W) testa so predstavljeni v Tabeli 5.1. Rezultati dvosmernega K-W testa ($\alpha = 0.05$) so predstavljeni z vrednostjo testa H, stopnjami svobode ν in pripadajočo p-vrednostjo.

Primerjava obiskovalcev glede na spol pokaže, da imajo v povprečju moški večji čas prisotnosti (156s) kot ženske (137s). Razliko smo potrdili s Kruskal-Wallis (K-W) testom, ki pokaže statistično značilno razliko povprečnih rangov obeh skupin ($H(1)=4.25$, $p=0.039$). Primerjava starostnih skupin pokaže, da ima skupina opazovalcev starih od 15 do 24 let izrazito manjši čas prisotnosti od povprečja (101s v primerjavi s povprečjem 144s). Statistično značilno razliko potrди tudi K-W test ($H(6)=20.4$, $p=0.002$). Glede na rezultate K-W testa ($H(1)=1.48$, $p=0.223$), scenarij predvajane vsebina pričakovano nima statistično značilnega vpliva na čas prisotnosti opazovalcev.

Statistično značilno razliko v času prisotnosti med moškimi in ženskami bi lahko iskali v številu kratkih obiskov trgovine. Analiza pokaže, da je v času trajanja eksperimenta 51% vseh žensk in le 44% vseh moških imelo čas prisotnosti krajši od 60s. Razliko v starostnih skupinah bi lahko potrdilo dejstvo, da je trgovina primarno namenjena populaciji od 25 do 55 let, kar je tudi razvidno iz Slike 5.3.

Tabela 5.1

Kvantitativni rezultati merjenja in statistična analiza časa prisotnosti. Povprečna vrednost, mediana in standardni odklon so podani v sekundah. Opisna in statistična analiza pokaže statistično značilno razliko, če vzorec razdelimo glede na spol in starostno skupino.

Spr.	Vrednost	n_i	μ_x	M_x	σ_x	H	ν	p
spol	moški	504	156	73.4	204	4.25	1	0.039
	ženski	790	137	60.3	193			
starostna skupina	1-14	95	148	60.1	186	20.4	6	0.002
	15-24	133	101	43.7	146			
	25-34	258	154	64.4	222			
	35-44	323	138	68.7	191			
	45-54	251	163	86.8	206			
	55-64	153	157	72.1	213			
	65+	81	124	67.6	158			
vsebina	statična	141	58.1	193	635.1	1.48	1	0.223
	dinamična	629	148	72.3	202			
skupaj		1294	144	64.8	198			

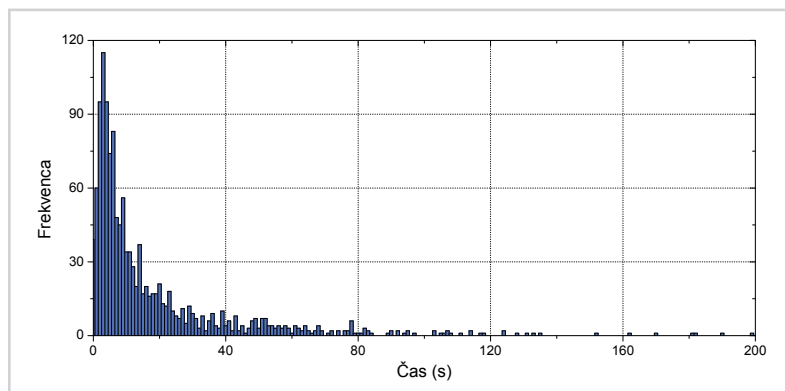
5.3.2 Čas kontaktne možnosti

Čas kontaktne možnosti zajema skupni čas, ko je uporabnik imel možnost opaziti vsebino na zaslonu. V primeru eksperimenta, $T_{o,i}$ predstavlja skupni čas, ki ga je oseba preživela v osrednjem nakupovalnem prostoru trgovine, z obrazom vsaj deloma obrnjena proti informacijskemu vmesniku in tako imela možnost opaziti zaslon.

Povprečni čas kontaktne možnosti eksperimentalno zajetih podatkov znaša $\mu_x = 17.6s$, z mediano $M_x = 8.38s$ in standardnim odklonom $\sigma_x = 24.8s$. Porazdelitev je negativno simetrična ($\gamma_1 = 3.12$) in koničaste oblike ($\gamma_2 = 13.2$). Odstopanje od normalne porazdelitve je še bolj izrazito kot pri porazdelitvi časa prisotnosti. Histogram porazdelitve časa kontaktne možnosti ($b_w = 1s$) prikazuje Slika 5.5.

Analiza časa kontaktne možnosti pokaže, da se je zaslon v povprečju 4.9 krat pojavil v vidnem polju obiskovalca ($\overline{N}_o = 4.9$). Glede na povprečen čas kontaktne možnosti ($\overline{T}_o = 17.6s$) in povprečen čas prisotnosti ($\overline{T}_d = 144s$) pa lahko sklepamo, da je imel obiskovalec možnost kontakta z zaslonom 12% celotnega časa, ko je bil prisoten v prostoru pred zaslonom. Primerjava obiskovalcev glede na spol pokaže, da imajo v

povprečju moški večji čas kontaktne možnosti (21s) kot ženske (16s). Razliko smo potrdili s K-W testom, ki potrjuje statistično značilno razliko povprečnih rangov obeh skupin ($H(1)=32.4, p < 0.0001$) in je predstavljen v Tabeli 5.2.



Slika 5.5

Porazdelitev časa kontaktne možnosti ($b_w = 1s$). V povprečju so imeli obiskovalci 17.6s možnosti kontakta z zaslonom (mediana znaša 8.38s). 960 obiskovalcev (74%) je imelo čas kontaktne možnosti manjši od 20s, 1070 (83%) manjši od 30s in 1213 (94%) manjši od 60s.

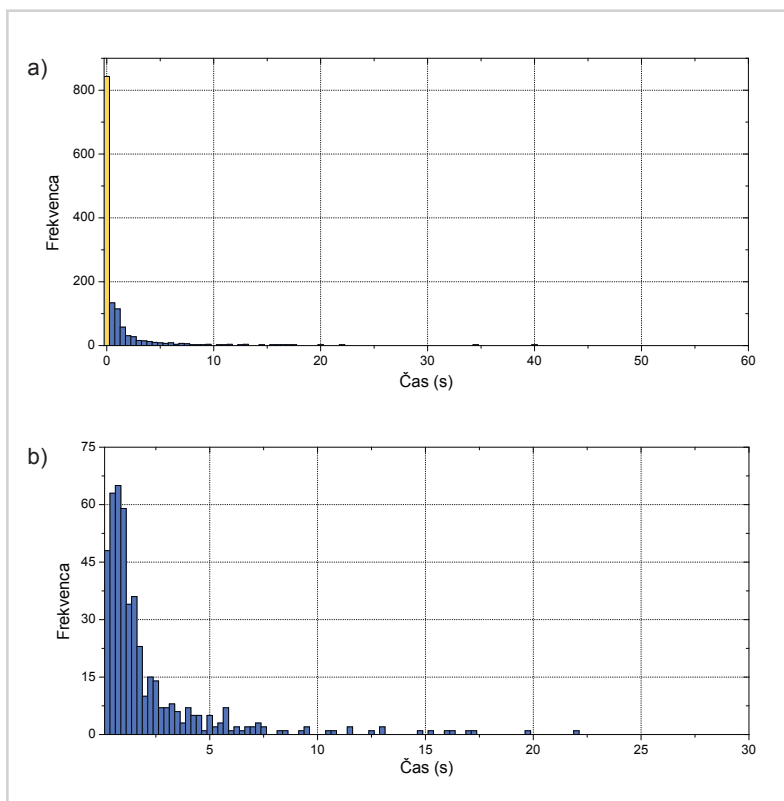
Tabela 5.2

Kvantitativni rezultati merjenja in statistična analiza časa kontaktne možnosti. Povprečna vrednost, mediana in standardni odklon so podani v sekundah. Opisna in statistična analiza pokaže statistično značilno razliko, če vzorec razdelimo glede na spol.

Spr.	Vrednost	n_i	μ_x	M_x	σ_x	H	ν	p
spol	moški	504	20.9	10.4	27.7	32.4	1	<0.0001
	ženski	790	15.6	6.88	22.6			
starostna skupina	1-14	95	17.9	8.15	26.7	6.77	6	0.343
	15-24	133	14.5	6.15	19.2			
	25-34	258	18.9	8.87	27.1			
	35-44	323	16.2	7.95	24.1			
	45-54	251	18.7	8.20	26.1			
	55-64	153	20.4	10.4	26.3			
65+	81	15.7	8.55	18.3				
vsebina	statična	665	16.7	8.1	22.7	0.85	1	0.357
	dinamična	629	18.6	8.8	26.7			
skupaj		1294	17.6	8.38	24.8			

5.3.3 Čas pozornosti

Čas pozornosti $T_{a,i}$ predstavlja skupni čas, ko je oseba i svojo pozornost namenila predvajani vsebini na zaslonu. Iz eksperimentalno zajetih podatkov izračunamo povprečni čas pozornosti obiskovalca, ki znaša $\bar{T}_a = 0.72s$, mediano $M_a = 0.0s$ in standardni odklon $\sigma_a = 1.91s$. Negativna asimetrija ($\gamma_1 = 5.46$) in koničasta oblika ($\gamma_2 = 39.8$) porazdelitve nakazujeta izrazito odstopanje zajetih podatkov od normalne porazdelitve. Podrobno porazdelitev časa pozornosti prikazujeta histograma na Sliki 5.6. Opisna statistika porazdelitve časa pozornosti in rezultati dvosmernega K-W testa ($\alpha = 0.05$) so predstavljeni v Tabeli 5.3.



Slika 5.6

Porazdelitev časa pozornosti. a) Histogram časa pozornosti ($b_w = 1.0s$), ki zajema podatke vseh 1294 obiskovalcev. Oranžno obarvani stolpec predstavlja 835 (65%) obiskovalcev, ki zaslon niso pogledali niti enkrat ($T_{a,i} = 0.0s$). b) Histogram časa pozornosti ($b_w = 0.25s$), ki zajema podatke 459 (35%) obiskovalcev, ki so zaslon pogledali vsaj enkrat ($N_{a,i} > 0.0s$). Med njimi jih je 161 (12%) vsebino na zaslonu pogledalo vsaj dvakrat in 80 (6%) trikrat ali več. Iz histogramov je razvidno, da obiskovalci zaslonu niso posvetili veliko pozornosti, kar potrjuje in kvantitativno ovrednoti efekt uporabniške slepote.

Tabela 5.3

Kvantitativni rezultati merjenja in statistična analiza časa pozornosti. Povprečna vrednost, mediana in standardni odklon so podani v sekundah. Opisna in statistična analiza pokaže statistično značilno razliko, če vzorec razdelimo glede na spol, starostno skupino in scenarij predvajanje vsebine.

Spr.	Vrednost	n_i	μ_x	M_x	σ_x	H	ν	p
spol	moški	504	1.19	0.0	2.61	71.9	1	<0.0001
	ženski	790	0.42	0.0	1.19			
starostna skupina	1-14	95	2.39	0.55	4.54	37.6	6	<0.0001
	15-24	133	0.70	0.0	1.41			
	25-34	258	0.60	0.0	1.35			
	35-44	323	0.42	0.0	1.19			
	45-54	251	0.67	0.0	1.69			
	55-64	153	0.68	0.0	1.73			
	65+	81	0.66	0.0	1.29			
vsebina	statična	665	0.60	0.0	1.49	5.71	1	0.017
	dinamična	629	0.86	0.0	2.27			
skupaj		1294	0.72	0.0	1.91			

Analiza pokaže, da je 35% vseh obiskovalcev trgovine pogledalo zaslon vsaj enkrat ($N_{a,i} > 1$). 12% jih je zaslon pogledalo dvakrat ali več, 6% pa trikrat ali več. Povprečni čas pozornosti znaša $\bar{T}_a = 0.7s$.

Primerjava obiskovalcev glede na spol pokaže, da so moški v povprečju zaslonu namenili več pozornosti kot ženske: 48% vseh moških in le 27% vseh žensk je pogledalo v zaslon vsaj enkrat. Razlika je evidentna tudi pri povprečnem času pozornosti, saj le-ta pri moških znaša 1.2s, pri ženskah pa 0.4s. Statistično značilno razliko je potrdil tudi K-W preizkus ($H(1)=71.9, p < 0.0001$).

Primerjava po starostnih skupinah pokaže, da ima starost obiskovalca velik vpliv na pričakovan čas pozornosti. K-W test pokaže statistično značilno razliko v distribucijah skupin ($H(6)=37.6, p < 0.0001$). Tudi primerjava povprečnega časa pozornosti v Tabeli 5.3 pokaže izrazito odstopanje skupine 1-14 let od ostalih skupin, kar smo statistično potrdili z uporabo dvostranskega Steel-Dwass-Critchlow-Flinger *post hoc* preizkusa [78]. Predvidevamo, da je razlog za izrazito odstopanje najmlajše starostne skupine v ponudbi trgovine. Asortiment trgovine je bil namenjen odraslim obiskoval-

cem, zaradi česar so mlajši obiskovalci več pozornosti namenili prikazovani vsebini na zaslonu.

Predvajanje statične in dinamične vsebine pričakovano ni prineslo značilnih odstopanj pri času prisotnosti in času kontaktne možnosti. V primeru časa pozornosti pa opazimo, da predvajanje dinamične vsebine bolj pritegne pozornost obiskovalcev kot predvajanje statične vsebine. Povprečni čas pozornosti se je ob predvajanju dinamične vsebine povečal za 43% glede na statično vsebino. Rezultati se skladajo z ugotovitvami kvalitativnih študij informacijskih vmesnikov [48, 81, 131], kot tudi psiholoških raziskav usmerjanja pozornosti (angl. *attention capture*) [76, 154]. Statistično značilno razliko distribucij obeh predvajanih scenarijev je potrdil tudi K-W test ($H(1)=5.71$, $p=0.017$).

5.4 Stopnje gledanosti in konverzija pozornosti med uporabniki

Študijo gledanosti informacijskega vmesnika želimo razširiti z analizo razmerij med zajetimi časovnimi značilnicami (čas prisotnosti, čas kontaktne možnosti in čas pozornosti). Gre za prvo tovrstno referenčno oceno razmerij na osnovi kvantitativno zajetih eksperimentalnih podatkov, ki se lahko tudi neposredno uporabi pri pripravi novih vsebin in načrtovanju podajanja vsebine poljubnih vmesnikov oziroma zaslonov.

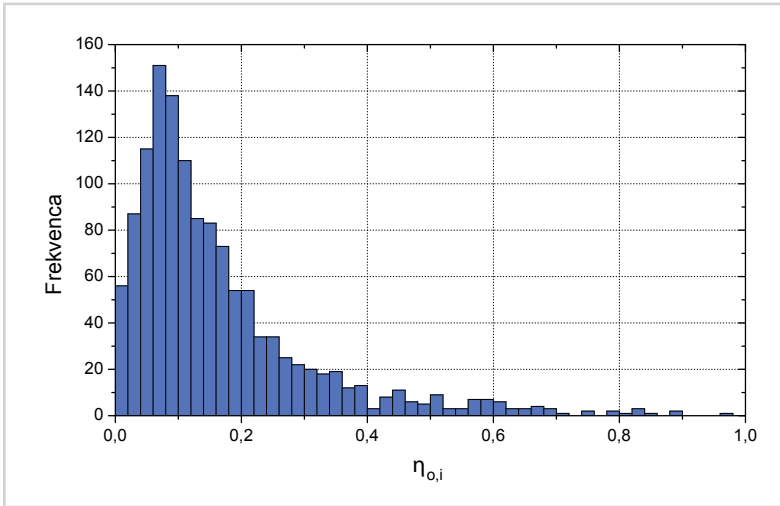
Definiramo razmerje med časom, ki ga je uporabnik preživel na področju informacijskega vmesnika in časom kontaktne možnosti $\eta_{o,i}$ kot:

$$\eta_{o,i} = \frac{\sum_{j=1}^{N_{o,i}} \tau_{o,i}}{\sum_{k=1}^{N_{d,i}} \tau_{d,k}} = \frac{T_{o,i}}{T_{d,i}}, \quad (5.5)$$

kjer $T_{o,i}$ predstavlja skupen čas kontaktne možnosti, $T_{d,i}$ pa skupen čas prisotnosti uporabnika i . Definiramo tudi povprečno razmerje med časom kontaktne možnosti in prisotnosti $\bar{\eta}_o$:

$$\bar{\eta}_o = \frac{1}{N} \sum_{i=1}^N \eta_{o,i}, \quad (5.6)$$

kjer N predstavlja število uporabnikov. Histogram porazdelitve $\eta_{o,i}$ ($b_w = 0.02$), ki zajema podatke vseh eksperimentalno zajetih podatkov, je prikazan na Sliki 5.7. Izračunamo povprečno vrednost razmerja med časom kontaktne možnosti in časom prisotnosti $\bar{\eta}_o = 0.16$, mediano $M_{\eta_{o,i}} = 0.12$ in standardni odklon $\sigma_{\eta_{o,i}} = 0.14$. Vrh histograma predstavlja koš na razponu $0.06 \leq \eta_{o,i} < 0.08$ in zajema 11.7% (151)



Slika 5.7

Porazdelitev razmerja med časom kontaktne možnosti in časom prisotnosti ($b_w = 0.02$). Tipičen obiskovalec je imel možnost opaziti vsebino zaslona $\sim 15\%$ časa od celotnega časa, ki ga preživel na področju pred zaslonom.

oseb. Razpon med $0.04 \leq \eta_{o,i} < 0.10$ zajema 31.2% (404) oseb. Za več kot polovico udeležencev eksperimenta, to je 52.7% (682), pa velja $0.04 \leq \eta_{o,i} < 0.16$. Povprečni obiskovalec trgovine je imel možnost opaziti predvajano vsebino $\sim 15\%$ vsega časa, ki ga je preživel na področju informacijskega vmesnika.

Opredelimo še razmerje med časom pozornosti in časom kontaktne možnosti $\eta_{a,i}$ uporabnika i kot:

$$\eta_{a,i} = \frac{\sum_{j=1}^{N_{a,i}} \tau_{a,i}}{\sum_{k=1}^{N_{o,i}} \tau_{o,k}} = \frac{T_{a,i}}{T_{o,i}} \quad (5.7)$$

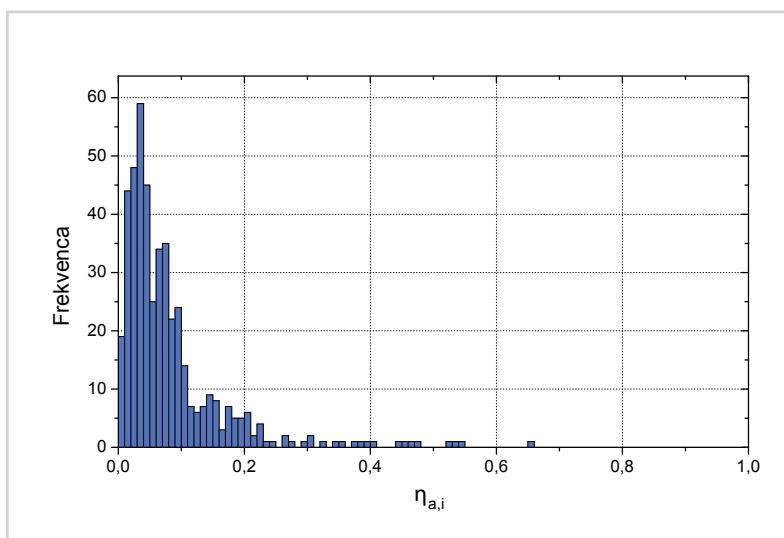
in povprečno razmerje med časom pozornosti in časom kontaktne možnosti $\bar{\eta}_a$ vseh uporabnikov kot:

$$\bar{\eta}_a = \frac{1}{N} \sum_{i=1}^N \eta_{a,i} \quad (5.8)$$

Pri analizi $\eta_{a,i}$ smo uporabili zgolj podatke 461 gledalcev, ki so v zaslon pogledali vsaj enkrat ($N_{a,i} > 0$). Histogram porazdelitve $\eta_{a,i}$ ($b_w = 0.01$) je prikazan na Sliki 5.8. Povprečna vrednost razmerja med časom pozornosti in časom kontaktne možnosti je $\bar{\eta}_a = 0.08$, mediana $M_{\eta_{a,i}} = 0.05$ in standardni odklon $\sigma_{\eta_{a,i}} = 0.08$.

Slika 5.8

Porazdelitev razmerja med časom pozornosti in časom kontaktne možnosti ($b_w = 0.01$). $\eta_{a,i}$ doseže vrh na intervalu $0.03 \leq \eta_{a,i} < 0.04$, ki predstavlja 12.8% gledalcev, ki so zaslon pogledali vsaj enkrat.



Histogram doseže vrh pri razponu koša $0.03 \leq \eta_{a,i} < 0.04$, ki zajema 12.8% (59) gledalcev. S širitvijo intervala na razpon $0.02 \leq \eta_{a,i} < 0.05$ zajamemo 32.9% (152) gledalcev, 47.9% (221) pa jih je imelo $\eta_{a,i}$ v razponu $0.01 \leq \eta_{a,i} < 0.01$.

Za nadaljnjo analizo učinkovitosti razvitega sistema uporabimo konverzijsko analizo stopenj gledanosti. Ustrezna primerjava gledanosti različnih informacijskih vmesnikov (različne lokacije in postavitve, namen vmesnika, predvajane vsebine, konfiguracija strojne opreme, idr.) zahteva enoten in sistematičen pristop. Predstavimo konverzijsko analizo stopenj gledanosti, ki sledi smernicam analize interaktivnosti predlaganim v [122]. Glede na opažanja v eksperimentu definiramo 5 stopenj gledanosti, ki opisujejo zaporedna stanja pozornosti gledalca:

- **PRISOTNOST**; definira stopnjo gledanosti osebe, ki je prisotna na predvajalnem mestu. Področje informacijskega vmesnika je odvisno od specifične situacije postavitve zaslona in običajno zajema prostor okoli zaslona, kjer oseba še lahko vidi predvajano vsebino. V primeru eksperimenta je predvajalni prostor pokrival celotno trgovino in tako kot *prisotne* štejemo vse osebe, ki so vanjo vstopile v času eksperimenta.

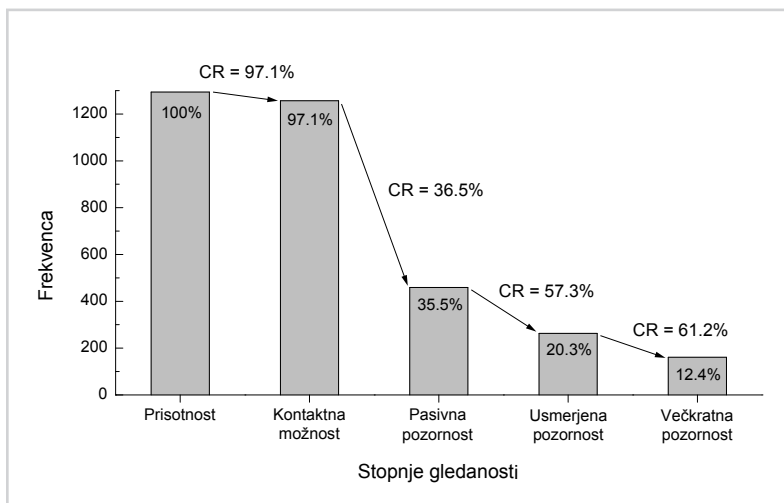
- **KONTAKTNA MOŽNOST**; označuje stopnjo gledanosti, ko je oseba imela možnost videti vsebino na zaslonu. V primeru eksperimenta to stopnjo gledanosti definiramo tedaj, ko imela oseba število kontaktnih možnosti $N_{o,i} \geq 1$.
- **PASIVNA POZORNOST**; označuje stopnjo gledanosti, ko oseba s pogledom preleti zaslon in za trenutek usmeri svojo pozornost na prikazovano vsebino. Pasivna pozornost je običajno nenamerna in zajema le bežen pogled. V primeru eksperimenta pasivno pozornost definiramo, ko je število pozornosti uporabnika $N_{a,i} \geq 1$.
- **USMERJENA POZORNOST** oziroma aktivna pozornost; označuje stopnjo, ko oseba usmeri svojo pozornost neposredno na predvajano vsebino in jo nato zavedoma spremlja. Ustrezen časovni interval pozornosti nakazuje, da gre za namerno spremljanje vsebine in ne zgolj bežen pogled. V eksperimentu smo stopnjo usmerjene pozornosti zabeležili osebam, ki so imele čas pozornosti $T_{a,i} \geq 1.0s$.
- **VEČKRATNA POZORNOST**; predstavlja stopnjo gledanosti, ko oseba večkrat usmeri svojo pozornost neposredno na predvajano vsebino zaslona. V eksperimentu smo jo določili ljudem, ki so imeli število intervalov pozornosti $N_{a,i} \geq 2$.

V času trajanja eksperimenta je bilo v trgovini prisotnih 1294 oseb. Možnost stika s predvajano vsebino je imelo 1257 (97.1%) oseb. Izmed teh jih 459 (35.5%) umestimo v stopnjo pasivne pozornosti, saj so vsaj bežno pogledali v zaslon. Stopnja usmerjene pozornosti vključuje 263 (20.3%) oseb, 161 (12.4%) oseb pa je svojo pozornost na zaslon usmerilo večkrat. Slika 5.9 prikazuje frekvence gledalcev, ki so prešli prag posamezne stopnje gledanosti. Poleg relativne frekvence prikazuje še stopnjo konverzije (angl. *conversion rate*) med posameznimi stopnjami gledanosti.

Razvidno je, da upada gledalcev med stopnjo prisotnosti in stopnjo kontaktne možnosti skoraj ni, saj stopnja konverzije znaša 97.1%. Razlog za to je najverjetneje v velikosti tlorisa trgovine, ki je skoraj v celoti predstavljal tudi predvajalno mesto in pa ustrezni postavitvi zaslona na vidno mesto znotraj trgovine. Podatki pa pokažejo izrazito mejo med stopnjo kontaktne možnosti in pasivno pozornostjo, saj stopnja konverzije znaša le 36.5%. Približno dve tretjini gledalcev nista zaslonu namenili niti bežnega pogleda, kar se sklada s primerljivo 33% stopnjo konverzije predstavljeno v [122]. Od 459 uporabnikov, ki so zaslonu namenili posredno pozornost, jih je 57.3% obdržalo

Slika 5.9

Konverzija med stopnjami gledanosti. Konverzija med stopnjo prisotnosti in kontaktno možnostjo znaša 97.1%, kar nakazuje na primerno postavitev in vidnost zaslona. Opazen padec konverzije (CR = 36.5%) med stopnjo kontaktno možnostjo in stopnjo pasivne pozornosti razložimo s pojavom uporabniške slepote. Stopnja konverzije med nadaljnji stopnjami gledanosti ohranja višjo in bolj stabilno, približno 60% vrednost.



pozornost na zaslonu vsaj 1s. Podobno opazimo tudi relativno veliko stopnjo konverzije med usmerjeno in večkratno pozornostjo, ki znaša 61.2%.

Padec med stopnjama kontaktno možnostjo in pasivne pozornosti pripisujemo pojavu uporabniške slepote [131], ki kvantitativno ovrednoti mejo apriornega nezanimanja ljudi za digitalne zaslone. Med nadaljnji stopnjami gledanosti pa se ponovno pokaže višja in relativno stabilna stopnja konverzije gledanosti (~60%).

5.5 Informacijska analiza gledanosti

Poleg predstavljene statistične obdelave zajetih časovnih in demografskih značilnic nas je zanimalo, ali lahko dobljene vedenjske vzorce tudi razložimo. Primarno nas je zanimalo, katere značilnice in v kakšni medsebojni povezavi vplivajo na to, da je oseba vsaj enkrat usmerila svojo pozornost na prikazovano vsebino na zaslonu. V ta namen smo izvedli analizo grafa interakcij med značilnicami (angl. *interaction graph analysis*), ki pokaže izrazite interakcije med posameznimi značilnicami [83].

V nadaljevanju najprej opišemo osnovne količine informacijske teorije [67] in predstavimo definiramo graf interakcij med značilnicami [84]. Sledi informacijska analiza gledanosti, v kateri s pomočjo zgrajenega grafa interakcij pokažemo interakcijske povezave med zajetimi podatki.

5.5.1 Graf interakcij med značilnicami

Graf interakcij med značilnicami [84] temelji na Shannonovi entropiji, ki v teoriji informacij predstavlja mero nedoločenosti [67]. Shannonova entropija H diskretne naključne spremenljivke X , ki lahko zavzame vrednosti $\{x_1, x_2, \dots, x_K\}$, je definirana kot:

$$H(X) = \sum_{x \in X} p(x) \log_2 \frac{1}{p(x)} \quad (5.9)$$

kjer $p(x)$ predstavlja verjetnost pojavitve vrednosti x . Definiramo tudi medsebojno informacijo ali informacijski prispevek (angl. *information gain*), ki nam pove koliko informacije ena spremenljivka vsebuje o drugi. Informacijski prispevek naključnih diskretnih spremenljivk X in Y definiramo kot:

$$I(X; Y) = \sum_{x \in X, y \in Y} p(x, y) \log_2 \frac{p(x, y)}{p(x)p(y)} \quad (5.10)$$

kjer $p(x, y)$ označuje verjetnost, da se vrednosti x in y pojavita hkrati [67]. V kolikor med X in Y obstaja korelacija, informacijski prispevek zavzame pozitivno vrednost, če pa sta spremenljivki medsebojno neodvisni in je $p(x, y) = p(x)p(y)$, pa vrednost nič. Relativno medsebojno informacijo (angl. *relative mutual information*), izraženo v odstotkih, izračunamo kot:

$$I_{rel}(X; Y) = \frac{I(X; Y)}{H(Y)} \quad (5.11)$$

Definiramo tudi pogojno medsebojno informacijo (angl. *conditional mutual information*) med tremi naključnimi diskretnimi spremenljivkami X , Y in Z kot:

$$I(X; Y|Z) = \sum_{x \in X, y \in Y, z \in Z} p(x, y, z) \log_2 \frac{p(x, y|z)}{p(x|z)p(y|z)} \quad (5.12)$$

kjer $p(x|z)$ predstavlja pogojno verjetnost. Pogojna medsebojna informacija ima negativno vrednost in pokaže, kako Z vpliva na interakcijo med X in Y . Na osnovi pogojne medsebojne informacije in informacijskega prispevka definiramo informacijo interakcije (angl. *interaction information*) med diskretnimi naključnimi spremenljivkami X , Y in Z kot:

$$I(X; Y; Z) = I(X; Y|Z) - I(X; Y), \quad (5.13)$$

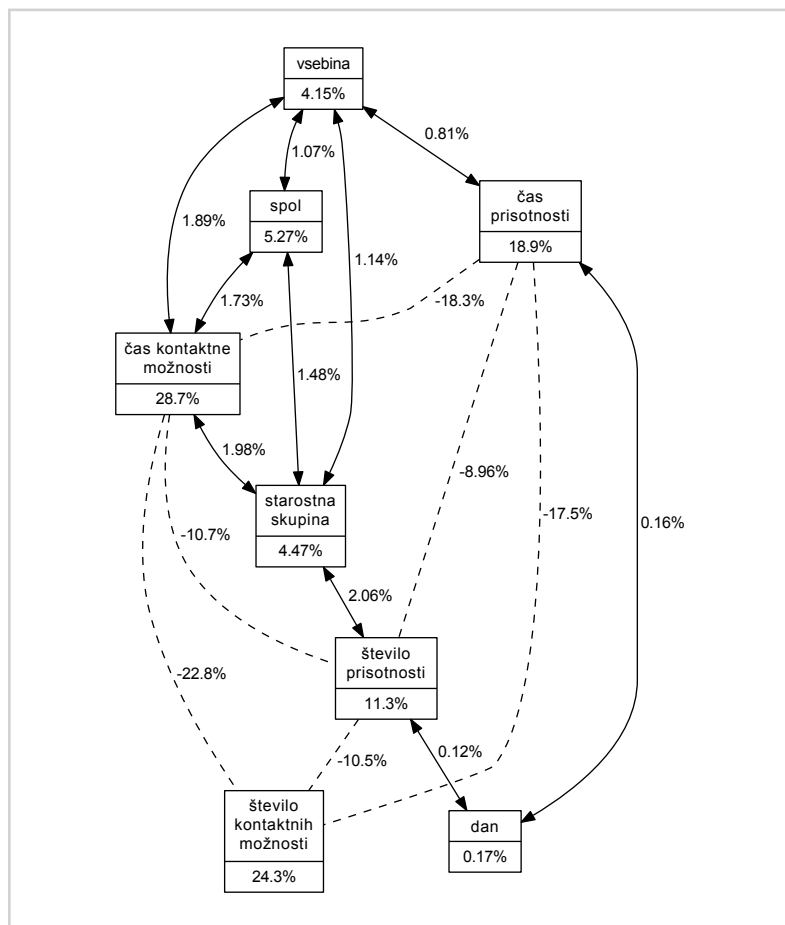
ki jo lahko razumemo kot mero informacije, skupno hkrati vsem trem spremenljivkam [84]. Interakcijska informacija lahko zavzame nenegativno ali negativno vrednost. Vse predstavljene mere so simetrične.

Graf interakcij je vizualizacija interakcij med značilnicami, katerega definira množica vozlišč ter povezav med njimi. Vsako vozlišče grafa ustreza posamezni značilnici, kateri poleg imena pripišemo tudi njeno relativno skupno informacijo z razredno spremenljivko. Povezava med značilnicami X in Y pri razredni spremenljivki Z pa predstavlja relativno vrednost $I(X; Y; Z)$ glede na entropijo razredne spremenljivke $H(Z)$. Negativna povezava, označena z neusmerjeno črtkano črto, med vozliščema pomeni, da si dve značilnici deloma delita enako informacijo razreda. Pozitivna povezava, označena z dvosmerno polno črto, pa označuje relativno količino nove informacije, ki jo dobimo ob poznavanju povezanih značilnic [83, 84]. Primer takega grafa je prikazan na Sliki 5.10.

5.5.2 Analiza interakcij med značilnicami

Iz podatkov gledanosti, ki so bili zajeti pri eksperimentu izračunamo graf interakcij med značilnicami. Pri izračuni smo uporabili programsko ogrodje Orange, namenjeno strojnemu učenju [39, 46]. Zanima nas, katere značilnice in v kakšni meri vplivajo na to ali neka oseba usmeri svojo pozornost na vsebino zaslona. Graf interakcij predpostavlja diskretne značilnice. Tako značilnici čas pozornosti in število intervalov pozornosti nadomestimo z novo, binarno razredno spremenljivko *pogledal-zaslon*, katero definiramo kot logično izjavo $N_{a,i} > 0$. Ostale zvezne značilnice (čas prisotnosti, čas kontaktne možnosti) pa pretvorimo v diskretne z uporabo MDL diskretizacije [31]. Graf interakcij med značilnicami je predstavljen na sliki 5.10. Zaradi boljše preglednosti so med vozlišči prikazane zgolj najbolj značilne povezave.

Najbolj informativna značilnica je *čas kontaktne možnosti*, ki sama izloči 28.7% negotovosti ali bo neka oseba pogledala zaslon ali ne. Druga najbolj informativna značilnica je *število intervalov kontaktnih možnosti*, ki označuje kolikokrat je imela oseba možnost opaziti zaslon, in sama izloči 2.4.3% negotovosti razreda. Negativna interakcija med tema dvema značilnicama (črtkana črta v grafu) označuje, da *število kontaktnih možnosti* zmanjša skupno entropijo razreda za le $24.3 - 22.8 = 1.5\%$ ob že predhodno znanem *času kontaktne možnosti*. Značilnica *spol* samostojno zmanjša entropijo razreda za 5.27%. Če pa upoštevamo pozitivno interakcijo med *časom kontaktne možnosti* in *spolom* (dvosmerna polna črta), vidimo da skupaj izločita $28.7 + 5.27 + 1.73 = 35.7\%$



Slika 5.10

Graf interakcij med značilnicami za binarni razredno spremenljivko *pogledal-zaslon*. Za najbolj informativne značilnice se izkažejo *čas kontaktne možnosti* (28.7%), *število kontaktnih možnosti* (24.3%) in *čas prisotnosti* (18.9%). Negativne interakcije (črtkana črta) med njimi nakazujejo, da si delijo večino skupne informacije razreda.

entropije razreda.

Najbolj informativne značilnice v grafu interakcij so: *čas kontaktne možnosti* (28.7%), *število intervalov kontaktne možnosti* (24.3%), *čas prisotnosti* (18.9%) in *število intervalov prisotnosti* (11.3%). Dobljene rezultate razložimo s sklepom, da več časa ko oseba preživi v trgovini oziroma ima zaslon v svojem vidnem polju, večja je verjetnost, da bo nanj usmerila pozornost. Med omenjenimi značilnicami opazimo negativne povezave,

ki nakazujejo, da je med njimi večino skupne informacije.

Značilnice *spol* (5.27%), *starostna skupina* (4.47%) in *tip vsebine* (4.15%) so tudi zanimive. Iz pozitivnih interakcij med njimi sklepamo, da so med seboj povezane. Tako se nedoločenost ali bo oseba pogledala na zaslon ali ne, ob znanem *spolu* in *tipu vsebine* zmanjša za $5.27 + 4.15 + 1.07 = 10.5\%$, ob znanem *tipu vsebine* in *starostni skupini* pa za 9.76%.

V grafu opazimo še značilnico *dan*, ki označuje dan v tednu, ko je oseba obiskala trgovino. Značilnica je bila dodana grafu interakcij za preverjanje veljavnosti interakcijske analize. Predvidevali smo, da dan v tednu ni koreliran s tem, ali je oseba pogledala zaslon ali ne. Interakcijska analiza potrди domnevo, saj ob znanem *dnevu* izločimo le 0.17% entropije razreda.

*Modeliranje procesa
nakupnega odločanja*

6

Razviti sistem za adaptivno prikazovanje vsebine in podatke gledanosti uporabimo za interdisciplinarno raziskavo modeliranja nakupnega procesa. V nadaljevanju najprej opredelimo nakupni proces in tipične vloge, ki se v njem pojavijo. Sledi analiza procesa nakupnega odločanja iz eksperimentalno zbranih podatkov in predstavitev rezultatov modeliranja nakupnega procesa s strojnim učenjem.

6.1 *Proces nakupnega odločanja*

Proces nakupnega odločanja opisuje zaporedje postopkov, ki ga opravi porabnik pri izbiri in odločitvi za nakup določenega izdelka oziroma storitve [148]. Razumemo ga lahko tudi kot proces odločanja in reševanja problema, kjer porabnik z premišljenim in zavestnim ravnanjem zadovolji svoje potrebe [42]. Rezultat procesa nakupnega odločanja predstavlja odločitev porabnika ali se bo odločil za nakup izdelka oziroma storitve.

Proces nakupnega odločanja ima po Kotler in Keller-ju pet stopenj [92]. Prva stopnja je prepoznavanje problema oziroma potrebe, ki jo lahko sprožijo notranji ali zunanji dejavniki. V drugi stopnji sledi iskanje informacij, ki lahko zajema zgolj višji nivo pozornosti (angl. *heightened attention*) pri stiku z tematiko problema oziroma se stopnjuje v aktivno iskanje (angl. *active information search*). Tretja stopnja predstavlja vrednotenje alternativ, ki običajno vključuje primerjavo različnih možnosti in lastnosti na osnovi modelov pričakovane vrednosti in prepričanja. V četrti stopnji nakupne odločitve se opredeli ponudnika, kraj, čas, vrednost, način in količina izbranega izdelka oziroma storitve. Zadnja, peta stopnja procesa nakupnega odločanja pa opisuje ponakupno vedenje in odzive [92].

Ločimo različne vrste nakupnega odločanja: odločanje iz navade, zoženi proces odločanja in razširjeni proces odločanja. Vrste se med seboj razlikujejo po vrednosti in pogostosti nakupa, zajemajo pa tudi različno intenzivne stopnje vpetosti in časa, vloženega v nakupno odločitev [171].

Na nakupni proces in vedenje vplivajo dejavniki nakupne odločitve, ki jih lahko razdelimo v skupine kulturnih, družbenih in osebnih dejavnikov. Kulturni dejavniki zajemajo kulturni okvir, pripadnost določenemu družbenemu razredu in subkulturi. Družbene dejavnike definirajo položaj in vloga posameznika, družina in referenčne skupine, ki imajo posreden in neposreden vpliv na nakupno vedenje. Osebnih dejavnikov pa opredelijo življenjski slog, poklic, premoženjsko stanje, osebnost in samopodoba [92, 171].

V nakupnem procesu je lahko udeležena ena, lahko pa tudi več oseb. Definiramo nakupno enoto, ki zajema sodelovanje množice oseb pri nakupni odločitvi. Vsaka izmed oseb lahko zavzame različno nakupno vlogo [92]:

- **POBUDNIK** je oseba, ki prva prične z iskanjem rešitve problema in predlaga nakup izdelka ali storitve. Vlogo pobudnika lahko zavzame uporabnik, lahko pa tudi drugi udeleženci nakupne enote.
- **VPLIVNEŽ** je oseba, katere mnenje ali stališče vpliva na nakupno odločitev. Vpliv običajno poteka s podajanjem informacij o lastnostih izdelka in vrednotenjem alternativ.
- **UPORABNIK** je oseba, kateri je izdelek ali storitev namenjena. Običajno sodeluje pri definiciji zahtevanih lastnosti izdelka ali storitve.
- **ODLOČEVALEC** narekuje končno odločitev pri izbiri med različnimi izdelki. Odločitev sprejme na osnovi zahtevanih lastnosti izdelka oziroma storitve.
- **NAKUPOVALEC** ima formalno avtoriteto, da neposredno opravi nakup in plača za blago oziroma storitev. Določi tudi pogoje nakupa, kot je na primer način plačila.
- **SPREMLJEVALEC** oziroma *pasivni vplivnež* je oseba, ki je del nakupne enote, a ni aktivno udeležena v nakupni proces.

Nabor vseh vlog nakupnega procesa ni enolično določen in se v literaturi razlikuje. Pobudnik, vplivnež, uporabnik, odločevalec in nakupovalec sodijo med osnovne vloge nakupnega procesa in se pojavijo v večini opredelitev [92, 171]. Vlogo spremljevalca vpeljemo sami na osnovi analize nakupnega procesa opisane v nadaljevanju. Gre za osebe, ki imajo pri nakupu pasivno vlogo in bi jih ob upoštevanju zgolj najbolj ustaljenih vlog morali izključiti iz analize.

6.2 Modeliranje procesa nakupnega odločanja

Podatke iz študije gledanosti opisane v poglavju 5 uporabimo za analizo procesa nakupnega odločanja. V ta namen anotiramo podatke gledanosti za vsakega udeleženca eksperimenta z novimi značilnicami, ki opisujejo nakupno enoto, velikost nakupne

enote, opravljen nakup in vloge v procesu nakupnega odločanja. Pri anotaciji sta sodelovala dva ocenjevalca, ki sta v primeru nesoglasja, po medsebojni argumentaciji sprejela skupno oceno.

Značilnica nakupne enote oziroma skupine predstavlja enolično določeno številko nakupne enote, kateri pripada določena skupina oseb. Velikost nakupne enote definiramo s številom oseb v nakupni enoti. Opravljen nakup pomeni ali se je nakupna enota odločila za nakup ali ne. V kolikor je skupina opravila nakup, za vsakega udeleženca ocenimo še vsako izmed šestih definiranih vlog v procesu nakupnega odločanja.

V času eksperimenta je bilo opravljenih 140 nakupov. Izmed $N = 1294$ oseb je bilo v nakupni proces vpetih $N_{buy} = 246$ oseb. Porazdelitev odločitve za nakup udeležencev glede na demografske značilnice, velikost skupine (nakupne enote) je prikazana v Tabeli 6.1. Z n_{all} označimo število oseb, katerih značilnica zadostuje izbrani vrednosti. Označo n_{buy} priredimo številu oseb, ki izpolnjujejo vrednost značilnice in so bile udeležene v proces nakupnega odločanja, ki se je končal z nakupom. Z p_n označimo delež/verjetnost pojavitve posamezne vrednosti spremenljivke glede na celotno število udeležencev $p_n = n_{all}/N$. Definiramo tudi delež/verjetnost pojavitve vrednosti značilnice znotraj oseb, ki so bile udeležene v nakupni proces kot: $p_{buy} = n_{buy}/N_{buy}$. Z p_{all} pa označimo delež/verjetnost oseb z dano značilnico vključena v nakupni proces glede na samo zastopanost te značilnice, tj. $p_{all} = n_{buy}/n_{all}$.

Primerjava udeležbe v nakupnem procesu pokaže skoraj popolno uravnoteženost nakupno konverzijo med spoloma. Delež žensk v celotnem eksperimentu zajema 61% ($p_n = 0.61$), delež tistih vključenih v nakupni proces pa 63% ($p_{buy} = 0.63$). Vseh moških je bilo 39% ($p_n = 0.39$), vključenih v nakupni proces pa 37% ($p_{buy} = 0.37$). Razmerje med vsemi sodelujočimi in tistimi udeleženi v nakup se glede na spol skoraj ohranja. Od 504 sodelujočih moških jih je bilo 18% ($p_{all} = 0.18$) udeleženi v nakupnem procesu. Pri ženskah je bila udeležba skoraj primerljiva, saj znaša 20% ($p_{all} = 0.20$).

Pregled po starostni skupini tudi pokaže določeno mero uravnoteženosti med razredi, saj se vrednost deleža oseb vključenih v nakupni proces glede na velikost razreda p_{all} pri večini giblje okoli 20%. Izjema je skupina sodelujočih starih med 15 in 24 let, kjer je delež približno dvakrat manjši in znaša 9% ($p_{all} = 0.09$).

Analiza velikosti nakupne skupine pokaže, da je v času trajanja eksperimenta predvajalno mesto obiskalo 438 posameznikov, 309 nakupnih enot z dvema osebama, 55 nakupnih enot s tremi osebami, 17 nakupnih enot s štirimi osebami in 1 z petimi

Tabela 6.1

Porazdelitev procesa nakupnega odločanja glede na demografske značilnice in velikost nakupne enote. Primerjava po spolu pokaže, da je delež moških in žensk, ki so bili udeleženi v nakupni proces približno enak (~20%). Delež sodelujočih starih med 15 in 24 let je približno dvakrat manjši od ostalih starostnih skupin. Z naraščanjem velikosti skupine pa postopoma raste tudi delež skupin, ki so odločile za nakup.

Spr.	Vrednost	n_{all}	n_{buy}	p_n	p_{buy}	p_{all}
spol	moški	504	92	0.39	0.37	0.18
	ženski	790	154	0.61	0.63	0.20
starostna skupina	1-14	95	20	0.07	0.08	0.21
	15-24	133	12	0.10	0.05	0.09
	25-34	258	54	0.20	0.22	0.21
	35-44	323	60	0.25	0.24	0.19
	45-54	251	53	0.19	0.22	0.21
	55-64	153	30	0.12	0.12	0.20
	65+	81	17	0.06	0.07	0.21
velikost skupine	1	438	57	0.34	0.23	0.13
	2	618	124	0.48	0.50	0.20
	3	165	57	0.13	0.23	0.35
	4	68	8	0.05	0.03	0.12
	5	5	0	0.004	0.0	0.0
skupaj		1294	246			

osebami. Zanimiv je trend, ki pokaže da se z naraščanjem velikosti nakupne skupine poveča tudi možnost za nakup. Od 438 posameznikov se jih je 13% ($p_{buy} = 0.13$) odločilo za nakup. Od 618 oseb, ki predstavljajo vse nakupne enote z dvema osebama, se jih je 20% odločilo za nakup ($p_{buy} = 0.20$). Delež doseže vrh pri skupinah, ki so jih sestavljale tri osebe, in znaša 35% ($p_{buy} = 0.35$). Število oseb, ki so bile vključene v večje nakupne enote (tj. s 4 in 5 osebami), predstavlja ~5% (73) od vseh sodelujočih, kar pa ne predstavlja dovolj velikega vzorca za statistično značilne sklepe. Tudi razmerje med $p_n = 0.13$ in $p_{buy} = 0.23$ pri nakupnih enotah s tremi osebami pokaže na veliko stopnjo konverzije.

Vpliv števila oseb v nakupni skupini na časovne značilnice (čas prisotnosti, čas kontaktne možnosti in čas pozornosti) prikazuje Tabela 6.2. Za vsako časovno značilnico izvedemo tudi dvosmerni Kruskal-Wallisov neparametrični test ($\alpha = 0.05$), s katerim

ugotavljamo statistično značilno razlike med dvema ali več skupinami. Število ljudi v posamezni skupini označuje n_{size} , povprečno vrednost μ_x , mediano M_x , standardni odklon σ_x , vrednost Kruskal-Wallis (K-W) testa H, število stopenj prostosti ν in p prilagojeno p-vrednost.

Tabela 6.2

Porazdelitev časovnih značilnic glede na velikost nakupne enote. Povprečna vrednost, mediana in standardni odklon so podani v sekundah. Primerjava pokaže povezavo med naraščanjem velikosti skupine in vseh časovnih značilnic (časa prisotnosti, časa kontaktno možnosti in časa pozornosti). Statistično značilno razliko rangov za vsako časovno značilnico glede na velikost nakupne enote potrди tudi K-W test.

	Spr.	Vr.	n_{size}	μ_x	M_x	σ_x	H	ν	p
čas prisotnosti	velikost skupine	1	433	117	51.4	181	33.4	4	<0.0001
		2	617	156	71.4	202			
		3	168	185	91.2	219			
		4	71	129	75.8	193			
		5	5	36.6	40.3	11.9			
	skupaj	1294	144	64.8	198				
čas kontak. možn.	velikost skupine	1	433	15.0	6.45	23.4	14.0	4	0.007
		2	617	18.9	9.2	25.1			
		3	168	20.7	9.6	27.4			
		4	71	16.9	8.1	24.4			
		5	5	6.04	4.3	4.09			
	skupaj	1294	17.7	8.38	24.8				
čas pozornosti	velikost skupine	1	433	0.51	0.0	1.69	21.7	4	0.0002
		2	617	0.74	0.0	1.89			
		3	168	1.16	0.0	2.32			
		4	71	0.89	0.0	2.16			
		5	5	0.08	0.0	0.18			
	skupaj	1294	0.72	0.0	1.91				

Velikost skupine vpliva na čas prisotnosti uporabnika. Zanimivo je, da imajo najmanjši čas prisotnosti osebe, ki so nakupovale same ($\mu_x = 117s$) in znaša skoraj ~20% manj od skupnega povprečja. Povprečni čas prisotnosti na osebo nato postopoma narašča z velikostjo skupine. Nakupna enota z dvema osebama je bila na predvajanem mestu v povprečju prisotna 156s, nakupna enota s tremi osebami pa 185s. Pri na-

knupnih enotah s štirimi udeleženci čas ponovno pade. Statistično značilno razliko v porazdelitvah potrди tudi K-W test ($H(4)=33.4, p < 0.0001$).

Podoben trend opazimo tudi pri času kontaktne možnosti z najmanjšo vrednostjo pri enoosebnih nakupnih enotah ($\mu_x = 15.0s$) in največjo pri nakupnih enotah s tremi osebami ($\mu_x = 20.7s$). Signifikantno razliko v porazdelitvah časa kontaktne možnosti glede na velikost skupine pokaže tudi K-W test ($H(4)=14.0, p=0.007$).

Efekt naraščanja časa je še bolj izrazit pri času pozornosti. Čas pozornosti oseb, ki so bile na predvajalnem mestu same, je namreč za ~30% manjši ($mu_x = 0.51s$) od skupnega povprečja. Povprečni čas pozornosti nakupnih enot s tremi člani je od skupnega povprečja večji za kar 60% ($\mu_x = 1.16s$). Statistično značilno razliko potrди tudi K-W test ($H(4)=21.7, p=0.0002$).

Vsem $N_{buy} = 246$ osebam, ki so bile udeležene v nakupni proces določimo tudi vloge v nakupnem procesu. Vsaka oseba nakupne enote je v nakupnem procesu zavzela vsaj eno od šestih vlog. Tabela 6.3 prikazuje porazdelitev nakupnih vlog in deleže/verjetnost njihove pojavitve. Oznaka n označuje število primerkov v dani skupini, p pa delež/verjetnost pojavitve te skupine. Indeksna oznaka *init* se nanaša na vlogo pobudnika, *inf* na vlogo vplivneža, *dcd* na vlogo odločevalca, *prch* označuje nakupovalca, *usr* uporabnika in *pssv* vlogo spremljevalca.

Primerjava porazdelitve vlog glede na spol pokaže značilno razliko pri vlogah vplivneža in spremljevalca. Delež moških, ki so zavzeli vlogo vplivneža znaša 8% ($p_{inf} = 0.08$) od vseh moških udeleženih v nakupni proces, ko je med tem pri ženskah ta odstotek skoraj petkrat večji in znaša 39% ($p_{inf} = 0.39$). Razmerje se obrne pri vlogi spremljevalca (pasivni vplivnež), katero je zavzelo 30% ($p_{pssv} = 0.30$) vseh moških in 8% žensk ($p_{pssv} = 0.08$).

Analiza starostnih skupin pokaže izrazito visok odstotek vlog spremljevalcev za najmlajšo skupino udeležencev v starostnem razredu 1-14 let, saj le-ta znaša 95% ($p_{pssv} = 0.95$). Omenjena starostna skupina tudi ni vplivala na noben opravljen nakup ($p_{inf} = 0.0$). Podoben trend se pokaže tudi pri starostnem razredu od 15-24 let, saj delež sodelujočih, ki so zavzeli vlogo spremljevalca znaša 42% ($p_{pssv} = 0.42$). Razlago iščemo v asortimentu trgovine, kjer je potekal eksperiment, saj je bil primarno namenjen odraslim osebam. To potrđijo tudi visoki deleži vlog pobudnika, odločevalca, nakupovalca in uporabnika pri vseh nadaljnjih starostnih skupinah.

Tabela 6.3

Porazdelitev vlog procesa nakupnega odločanja glede na demografske značilnice (spol in starostna skupina) in velikost nakupne enote za $N_{buy} = 246$ oseb, ki so bile udeležene v nakupni proces. Primerjava porazdelitve vlog glede na spol pokaže značilno razliko pri vlogah vplivneža in spremljevalca. Izrazit je tudi trend naraščanja pojavitve vloge spremljevalca z velikostjo nakupne skupine.

Spr.	Vrednost	n_{init}	n_{inf}	n_{dcd}	n_{prch}	n_{usr}	n_{pssv}	p_{init}	p_{inf}	p_{dcd}	p_{prch}	p_{usr}	p_{pssv}
spol	moški	48	7	53	56	57	28	0.52	0.08	0.58	0.61	0.62	0.30
	ženski	106	60	93	85	76	12	0.69	0.39	0.60	0.55	0.49	0.08
starostna skupina	1-14	1	0	1	1	1	19	0.05	0.0	0.05	0.05	0.05	0.95
	15-24	5	2	5	4	5	5	0.42	0.17	0.42	0.33	0.42	0.42
	25-34	39	18	35	36	33	1	0.72	0.33	0.65	0.67	0.61	0.02
	35-44	41	18	39	41	33	5	0.68	0.30	0.65	0.68	0.55	0.08
	45-54	40	16	36	31	32	4	0.76	0.30	0.68	0.59	0.60	0.08
	55-64	18	9	19	20	18	3	0.60	0.30	0.63	0.67	0.60	0.10
65+	10	4	11	8	11	3	0.59	0.24	0.65	0.47	0.65	0.18	
velikost skupine	1	57	2	56	55	50	0	1.0	0.04	0.99	0.97	0.88	0.0
	2	73	47	68	64	63	14	0.59	0.38	0.55	0.52	0.51	0.11
	3	22	16	20	20	18	23	0.39	0.28	0.35	0.35	0.32	0.40
	4	2	2	2	2	2	3	0.25	0.25	0.25	0.25	0.25	0.38
skupaj		154	67	146	141	133	40	0.63	0.27	0.59	0.57	0.54	0.16

Pregled po velikosti nakupnih enot pokaže trend naraščanja vloge spremljevalca z velikostjo skupine. Verjetnost, da oseba v nakupovalni enoti z dvema osebama zavzame vlogo spremljevalca znaša 11% ($p_{pssv} = 0.11$), v enoti s tremi osebami pa 40% ($p_{pssv} = 0.40$). Podobna vrednost (38%) se pokaže tudi v pri skupinah štirih oseb, vendar zaradi majhnega vzorca, ki zajema le 2 skupini, vrednost upoštevamo zgolj okvirno. Večja zastopanost vloge spremljevalca razloži tudi trend naraščajočih časov pozornosti, opazen v Tabeli 6.2. Verjetnost za pojavitev vlog pobudnika, odločevalca, nakupovalca in uporabnika pa pada približno linearno obratno sorazmerno z velikostjo nakupne enote.

Kumulativni pregled pokaže, da so v vsaki nakupni skupini, ki se je odločila za nakup, običajno prisotne vloge pobudnika, odločevalca in nakupovalca. Vlogi vplivneža in spremljevalca pa se pojavita pri nakupnih enotah z več osebami. To se sklada z opredelitvijo procesa nakupnega odločanja in razloži manjši delež oseb, ki so zavzele vlogo vplivneža (27%) in spremljevalca (16%). Manjša odstopanja pri enoosebnih nakupnih skupinah so posledica izločitve 12 osamelcev iz osnovnega nabora podatkov opisanega

Tabela 6.4

Prisotnost vlog nakupnega procesa glede na čas prisotnosti. Statistična analize ne pokaže značilnih razlik med prisotnostjo vlog. Povprečna vrednost, mediana in standardni odklon so podani v sekundah.

Vloga	Vr.	n_{role}	μ_x	M_x	σ_x	H	ν	p
pobudnik	da	154	414.9	337.2	293.8	0.009	1	0.926
	ne	92	402.3	336.1	275.5			
vplivnež	da	67	439.9	401.0	276.1	1.35	1	0.245
	ne	179	399.1	309.3	290.4			
uporabnik	da	133	403.9	325.2	295.4	0.50	1	0.479
	ne	113	417.7	363.1	277.1			
odločevalec	da	146	416.3	337.2	305.4	0.03	1	0.861
	ne	100	401.3	345.1	257.9			
nakupovalec	da	141	414.1	335.5	309.1	0.15	1	0.704
	ne	105	405.1	363.1	254.6			
spremljevalec	da	40	394.7	316.1	262.2	0.15	1	0.696
	ne	206	413.2	343.4	291.6			
skupaj		246	410.2	337.2	286.6			

v razdelku 5.2.

Zanimala nas je tudi povezava med časovnimi značilnicami in prisotnostjo posamezne vloge v nakupnem procesu. Vpliv prisotnost posamezne vloge glede na čas prisotnosti prikazuje Tabela 6.4. Statistična analiza pokaže, da prisotnost nobene od šestih vlog značilno ne vpliva na čas prisotnosti, kar je razvidno tudi iz primerjave povprečne vrednosti in mediane za različne vloge.

V Tabeli 6.5 predstavimo vpliv prisotnosti posamezne vloge nakupnega procesa na čas kontaktne možnosti. Statistična analiza z K-W testom izpostavi vlogo spremljevalca, ki ima nadpovprečno visok čas prisotnosti ($\mu_x = 52.3s$) glede na skupno povprečje ($\mu_x = 43.4s$). Razlago iščemo v definiciji vloge. Spremljevalec se lahko pojavi zgolj v nakupnih enotah z dvema ali več osebama. Med nakupom sodeluje kot pasivni udeleženec in svoje zanimanje usmerja prosto po predvajalnem mestu in ne v konkretno izbiro blaga. Statistično značilno razliko potrди tudi K-W test ($H(4)=4.28, p=0.039$).

Tabela 6.5

Prisotnost vlog nakupnega procesa glede na čas kontaktne možnosti. Prisotnost vloge spremljevalca se kaže v nadpovprečno visokem času kontaktne možnosti, kar potrди tudi statistična analiza. Povprečna vrednost, mediana in standardni odklon so podani v sekundah.

Vloga	Vr.	n_{role}	μ_x	M_x	σ_x	H	ν	p
pobudnik	da	154	41.3	33.9	32	0.31	1	0.576
	ne	92	46.7	37.3	41.8			
vplivnež	da	67	42.3	33.3	38.9	0.89	1	0.344
	ne	179	43.7	35.6	34.9			
uporabnik	da	133	43.1	33.1	35.6	0.01	1	0.941
	ne	113	43.6	37.5	36.6			
odločevalec	da	146	43.4	33.4	35.1	0.05	1	0.830
	ne	100	43.2	37.3	37.5			
nakupovalec	da	141	43.9	34.5	36.4	0.02	1	0.892
	ne	105	42.5	35.7	35.5			
spremljevalec	da	40	52.3	41.9	41.5	4.28	1	0.039
	ne	206	41.6	33.2	34.6			
skupaj		246	43.3	35.2	36.1			

Zadnjo časovno značilnico predstavlja čas pozornosti. Prisotnost vlog nakupnega procesa glede na čas pozornosti so predstavljene v Tabeli 6.6. Povprečni čas pozornosti nakupnih enot, katerih nakupni proces se je zaključil z nakupom, znaša 1.81s, kar je izrazito večje od skupnega povprečja 0.72s.

Tabela 6.6

Prisotnost vlog nakupnega procesa glede na čas pozornosti. Prisotnost vloge pobudnika, uporabnika odločevalca ali nakupovalca se pokaže v statistično značilnem manjšem času pozornosti. Pri prisotnosti vloge spremljevalca pa čas pozornosti za nekajkrat večji od povprečja. Povprečna vrednost, mediana in standardni odklon so podani v sekundah.

Vloga	Vr.	n_{role}	μ_x	M_x	σ_x	H	ν	p
pobudnik	da	154	1.33	0.65	2.43	12.7	1	0.0003
	ne	92	2.63	1.50	3.51			
vplivnež	da	67	1.31	0.55	1.86	1.66	1	0.197
	ne	179	2.01	0.90	3.24			
uporabnik	da	133	1.39	0.65	2.55	6.49	1	0.011
	ne	113	2.31	1.15	3.29			
odločevalec	da	146	1.37	0.65	2.50	9.06	1	0.003
	ne	100	2.46	1.25	3.41			
nakupovalec	da	141	1.40	0.65	2.59	8.58	1	0.003
	ne	105	2.37	1.20	3.29			
spremljevalec	da	40	4.51	3.38	4.34	40.7	1	<0.0001
	ne	206	1.29	0.60	2.26			
skupaj		246	1.81	0.90	2.94			

Analiza vloge pobudnika pokaže, da prisotnost vloge zmanjša čas pozornosti. Ta znaša 1.33s, kar je skoraj še enkrat manj od časa pozornosti oseb, ki vloge pobudnika niso zavzele (2.63s). Značilno razliko potrди tudi K-W test ($H(1)=12.7$, $p=0.0003$). Pojav si razlagamo z večjo stopnjo osredotočenosti pobudnika za nakup, ki posledično pušča manj časa za stvari, ki niso neposredno vezane na nakupni proces.

Podoben trend se pokaže tudi pri vlogah uporabnika, odločevalca in nakupovalca, kjer prisotnost vloge pomeni manjši čas pozornosti. Statistično značilnost razlike za omenjene vloge potrди tudi K-W test.

Prisotnost vloge spremljevalca pokaže na izrazito visok čas pozornosti, ki znaša kar $\mu_x = 4.51s$ in za več kot šestkrat presega skupno povprečje (0.72s) in skoraj trikrat povprečje oseb, udeleženih v nakup (1.81s). Statistično značilnost z visoko stopnjo zaupanja potrди tudi K-W test ($H(1)=40.7$, $p<0.0001$). Razlog za visoko razliko v času pozornosti, podobno kot pri času kontaktne možnosti, iščemo v pasivni naravi vloge. Oseba, ki pri nakupnem procesu ni sodelovala aktivno je več časa namenila prikazovani vsebini.

6.3 Modeliranje nakupnega procesa nakupa s strojnim učenjem

Na osnovi zbranih podatkov preizkusimo različne metode strojnega učenja za izdelavo klasifikatorja nakupne odločitve. V času trajanja eksperimenta smo zajeli podatke 1294 oseb, od tega jih je bilo 19% (246) vključenih v nakupne enote, ki so se odločile za nakup. Izhodišče za našo analizo predstavlja 81% ($100\% - 19\% = 81\%$) natančnost, ki ustreza apriorni verjetnosti porazdeljenosti razreda. Osnovni nabor podatkov zajema število intervalov in kumulativno vsoto za vsako časovno značilnico (čas prisotnosti, čas kontaktne možnosti in čas pozornosti), spol, starostno skupino, tip predvajane vsebine in velikost nakupne skupine.

Tabela 6.7 prikazuje rezultate 10-kratnega prečnega preverjanja za izbrane metode strojnega učenja: večinski klasifikator (Večina), naivni Bayesov klasifikator (NB), metodo K-najbližjih sosedov (kNN), metodo podpornih vektorjev (MPV) in metodo naključnih gozdov (NG). Ciljni razred predstavlja odločitev za nakup. Večinski klasifikator doseže 81% klasifikacijsko točnost in predstavlja referenčno vrednost za oceno ostalih metod. Za najbolj uspešno se izkaže metoda podpornih vektorjev (MPV), ki doseže klasifikacijsko točnost 88.6% in tako izboljša napoved večinskega klasifikatorja za 7.6%. Metoda naključnih gozdov se izkaže kot druga najboljša, saj doseže 87.3% klasifikacijsko točnost. Ostale metode strojnega učenja presežejo izhodiščno klasifikacijsko točnost za ~5%.

Modeliranje z metodami strojnega učenja izvedemo tudi za vloge v nakupnem procesu. Na podlagi opažanj in možnosti iz osnovnega nabora podatkov gledanosti izpeljemo nove hevristične značilnice. Binarna logična značilnica *v-skupini* označuje ali je bila v nakupni enoti več kot ena oseba. Na osnovi zabeleženih intervalov časa prisotnosti definiramo numerično značilnico *vstopil*, ki označuje vrstni red vstopanja posamezne osebe znotraj nakupne enote v trgovino. Iz nje izpeljemo binarno logično značilnico *vstopil-prvi*, ki označuje ali je določena oseba v nakupni enoti prva vstopila v trgovino.

Tabela 6.7

Primerjava različnih algoritmov strojnega učenja za klasifikacijo procesa nakupnega odločanja. Primerjava z 10-kratnim prečnim preverjanjem pokaže, da največjo klasifikacijsko točnost ($CA_{MPV} = 0.886$) doseže metoda podpornih vektorjev, ki izboljša natančnost večinskega klasifikatorja za 7.6%.

Metoda	Klasifikacijska točnost	Senzitivnost	Specifičnost
Večina	0.810	0.000	1.000
NB	0.867	0.768	0.890
kNN	0.851	0.492	0.935
MPV	0.886	0.594	0.954
NG	0.873	0.394	0.986

Po enakem ključu definiramo še značilnici *odšel* in *odšel-zadnji*. Nabor hevrističnih značilnic zaključujeta še razmerje med časom prisotnosti in časom kontaktne možnosti $\eta_{o,i}$ in razmerje med časom kontaktne možnosti in časom pozornosti $\eta_{a,i}$.

Evaluacijo algoritmov strojnega učenja izvedemo z 10-kratnim prečnim preverjanjem. Vsak algoritem preizkusimo na osnovnem naboru značilnic in na razširjenem naboru značilnic z dodanim naborom hevrističnih značilnic. Klasifikacijska natančnost za posamezni algoritem in dano vlogo v nakupnem procesu je prikazana v Tabeli 6.8. Primer, ko je algoritma deloval na razširjenem naboru značilnic označimo z indeksom *h*.

Ponovno se za najbolj robustno in uspešno metodo izkaže metoda podpornih vektorjev (MPV). V primeru klasifikacije z osnovnimi podatki gledanosti (zgornja polovica Tabele 6.8) doseže MPV najboljšo klasifikacijsko točnost pri napovedovanju vlog pobudnika (69.2%), vplivneža (74.8%), uporabnika (72.8%), nakupovalca (72.4%) in spremljevalca (91.0%). Za vlogo odločevalca pa se s klasifikacijsko točnostjo 63.5% za najbolj uspešno izkaže metoda naključnih gozdov. Klasifikacijska natančnost najboljših metod, ki so delovale na naboru osnovnih značilnic gledanosti, v povprečju za vseh 6 vlog preseže točnost večinskega klasifikatorja za 8.9%.

Spodnji del Tabele 6.8 predstavlja klasifikacijsko natančnost izbranih metod strojnega učenja na podatkih z dodanimi hevrističnimi atributi. Najboljša dosežena klasifikacijska natančnost se izboljša pri vseh napovedovanju vseh šestih vlog. Metoda podpornih vektorjev z uporabo hevrističnih atributov v povprečju izboljša klasifikacijsko natančnost za 14.9% glede na referenčno oceno večinskega klasifikatorja.

Tabela 6.8

Klasifikacijska točnost algoritmov strojnega učenja za napovedovanje vlog nakupnega procesa z uporabo 10-kratnega prečnega preverjanja. Metode, katerim smo za klasifikacijo dodali hevristične atribute so označene z indeksom h . Za najbolj uspešno in robustno se izkaže metoda podpornih vektorjev. Če klasifikacijsko točnost povprečimo čez vseh šest vlog, metoda podpornih vektorjev z dodanimi hevrističnimi značilnicami doseže za 14.9% boljšo klasifikacijsko točnost od večinskega klasifikatorja, ki predstavlja osnovno referenčno oceno.

Metoda	Vloga v procesu nakupnega odločanja					
	Pobudnik	Vplivnež	Uporabnik	Odločevalec	Nakupovalec	Spremljevalec
Večina	0.626	0.728	0.541	0.593	0.573	0.837
NB	0.679	0.735	0.606	0.614	0.614	0.882
kNN	0.659	0.740	0.630	0.603	0.651	0.865
MPV	0.692	0.748	0.728	0.599	0.724	0.910
NG	0.651	0.724	0.611	0.635	0.653	0.861
NB _h	0.728	0.736	0.658	0.719	0.682	0.857
kNN _h	0.747	0.757	0.682	0.695	0.744	0.878
MPV _h	0.793	0.768	0.731	0.764	0.755	0.918
NG _h	0.711	0.724	0.686	0.670	0.614	0.846

Predstavljeni rezultati izhajajo iz podatkov analize gledanosti in so neposredno vezani na konkretno predvajalno mesto, kjer so bili zajeti. Gre za specifične podatke, ki so odvisni od lokacije in časovnega intervala zajema. Rezultate analize procesa nakupnega odločanja in modeliranje z uporabo strojnega učenja je zato potrebno razumeti kot predstavitev trendov vezanih na trgovino, v kateri je potekal eksperiment. Seveda pa razvita metodologija adaptivnega vmesnika omogoča, da se meritve in modeliranje ponovno opravijo na poljubnem predvajalnem mestu.



Zaključek

7

V doktorski disertaciji razvijemo adaptivni informacijski vmesnik in raziščemo procese, kjer prilagajanje vsebine ključno vpliva na učinkovitost in delovanje takega vmesnika, in sicer: uvedemo koncept digitalne anamorfoze, izvedemo kvantitativno analizo gledanosti ter modeliramo proces nakupnega odločanja. Razviti uporabniku prilagodljiv interaktivni sistem deluje na osnovi povezave metod računalniškega vida in strojnega učenja.

Adaptivnost in prilagodljivost sistema dosežemo z vpeljavo uporabniških značilnic. Pri opredelitvi metodologije in izbiri metod za modeliranje uporabniških značilnic se kot ključno pokaže razmerje med natančnostjo in časovno zahtevnostjo, saj je vmesnik zasnovan za realnočasno izvajanje na običajni konfiguraciji strojne opreme. Za segmentacijo slike uporabimo metodo Gaussovih mešanic, za detekcijo obrazov pa algoritem Viola & Jones. Detektiran obraz pretvorimo v poravnano normalizirano obliko z modeli aktivnega videza. Oceno orientacije obraza pridobimo z uporabo algoritma POSIT. Prostorske značilnice uporabnika modeliramo z metodo, ki temelji na medočesni razdalji. Vrednost demografskih značilnic uporabnika -spol in starostno skupino- določimo iz normalizirane in abstrahirane obrazne slike uporabnika s strojnim učenjem, kjer se za stabilno in robustno metodo izkaže metoda podpornih vektorjev.

Razviti vmesnik uporabimo za razvoj nove prostorske interaktivne metode, dinamične anamorfoze, kjer z uporabo anamorfične transformacije adaptivno prilagajamo osnovno sliko položaju uporabnika. Osrednje pri razvoju je formalizacija anamorfične transformacije slike. Predlagamo možno aplikacijo dinamične anamorfoze in sicer kot način za izboljšanje očesnega stika pri videokonferencah. Z dinamično anamorfozo izboljšanje očesnega stika eksperimentalno ovrednotimo in na osnovi študije ocenjevanja 208 slik tudi potrdimo.

Z razvitim vmesnikom izvedemo kvantitativno študijo in analizo gledanosti informacijskega vmesnika v realnem okolju. Študija je vključila podatke gledanosti 1294 oseb in ključno ovrednotila efekt uporabniške slepote. Konverzija stopenj gledanosti pokaže, da je informacijskemu vmesniku pasivno pozornost namenilo zgolj 35.5% oseb, usmerjeno pozornost pa 20.3% oseb. Izmerjeni povprečni čas pozornosti je bil 0.7s. Statistično opredelimo tudi značilne razlike gledanosti glede na demografske in predvajalne značilnice, kjer npr. kvantitativno pokažemo, da moški na tem konkretnem predvajalnem mestu, ki ga predstavlja trgovina z oblačili višjega cenovnega razreda, namenijo informacijskemu vmesniku skoraj trikrat več pozornosti kot ženske. Analiza grafa interakcij splošneje pokaže, da sta čas kontaktne možnosti in čas prisotnosti

uporabnika najbolj informativni značilnici gledanosti.

Na osnovi zajetih podatkov gledanosti modeliramo in izmerimo proces nakupne odločitve. Gre za interdisciplinarno delo računalništva in ekonomije, kjer z razvitim informacijskim vmesnikom in povezanim strojnim učenjem izvedemo analizo in modeliranje nakupne odločitve in vloge posameznika v njej. Klasifikacijska točnost napovedi nakupne odločitve za 7% izboljša apriorno verjetnost razreda. Točnost napovedi prisotnosti vloge nakupnega procesa pa v povprečju za vseh 6 vlog (pobudnik, vplivnež, uporabnik, odločevalec, nakupovalec, spremljevalec) presega apriorno verjetnost za 15%.

Predstavljeni rezultati odpirajo konceptualno nove in drugačne načine za nadgradnjo obstoječih in razvoj novih sistemov za prikazovanje multimedijskih vsebin. Ob ustreznem zagotavljanju zasebnosti lahko služijo kot napredno kvantitativno orodje za analizo gledanosti različnih vrst medijev, modeliranje uporabniških značilnic pa omogoča razvoj novih interaktivnih načinov podajanja informacij. Možna aplikacija sistema sega tudi na področje telemetrije, kjer bi se razviti koncepti na primer lahko uporabili za analizo gledanosti televizijskih kanalov.

Opisano delo odpira tudi serijo novih znanstvenih izzivov. Nadaljnje delo z razvitim vmesnikom lahko poteka v smeri širše analize gledanosti in obnašanja, ki bi lahko pokazala značilne vedenjske vzorce in družbene implikacije. Analiza pozornosti in možnost adaptivnih predvajalnih scenarijev bi lahko privedla do nadaljnjega razvoja na področju prikazovanja vsebin za informacijske vmesnike. Raziskave se lahko osredotočijo tudi na razvoj povratnih mehanizmov, ki so hkrati učinkoviti, a tudi dovolj preprosti in razumljivi za uporabnika. Nov pomemben korak pri določanju maksimalne učinkovitosti informacijskih vmesnikov bi predstavljala tudi analiza vloge položaja zaslona, velikosti zaslona in adaptivne vsebine. Splošneje pa uporaba predlagane metodologije lahko služi tudi kot osrednji element in osnova za primerjalne raziskave različnih dosedaj ločeno raziskovanih konceptov.

Kot sklepne misli, zapišimo mnenja nekaterih vodilnih raziskovalcev s področja ambientalne inteligence, npr. Krumma, ki pravi, da bo vseprisotno računalništvo, predvsem na javnih in pretočnih vsebinah povzročilo naslednji velik tehnološki preskok [95]. Pravtako je eden pionirjev vseprisotnega računalništva, Mark Weiser, že leta 1991 v svojem znanem članku *The Computer of the 21st Century* predlagal tri osnovne tehnološke platforme, ki bodo v prihodnosti podpirale moderno digitalno življenje [191]. Po analogiji z anglosaškimi dolžinskimi merskimi enotami je tehnologijo

razdelil v tri razrede -palec, čevelj, jard- glede na njihovo osnovno uporabo. In res, palec danes predstavljajo mobilni telefoni. Mesto naprav v dimenzijah čevlja so zasedli tablični računalniki in bralniki elektronskih knjig. Zadnja platforma, reda velikosti jarda, pa nastaja danes v obliki pametnih televizorjev in interaktivnih zaslonov oziroma vmesnikov [188], kar je ena od sklepnih indikacij teže predstavljenega doktorskega dela.

LITERATURA

- [1] E. Aarts. Ambient intelligence: a multimedia perspective. *IEEE Multimedia*, 11(1):12–19, 2004.
- [2] ACM. *The ACM Computing Classification System (CCS)*. Association for Computing Machinery, New York, 1998. <http://www.acm.org/class/>.
- [3] G. Adomavicius and A. Tuzhilin. Context-aware recommender systems. In *Recommender systems handbook*, pages 217–253. Springer, 2011.
- [4] L. A. Alexandre. Gender recognition: A multiscale decision fusion approach. *Pattern Recognition Letters*, 31(11):1422–1427, 2010.
- [5] G. Allenby. The effects of in-store displays and feature advertising on consideration sets. *International Journal of Research in Marketing*, 12(1):67–80, May 1995.
- [6] F. Alt, J. Müller, and A. Schmidt. Advertising on public display networks. *IEEE Computer*, 45(5):50–56, 2012.
- [7] F. Alt, S. Schneegaß, A. Schmidt, J. Müller, and N. Memarovic. How to evaluate public displays. In *Proceedings of the 2012 International Symposium on Pervasive Displays, PerDis '12*, pages 17:1–17:6, New York, NY, USA, 2012. ACM.
- [8] M. Andersen, T. Jensen, P. Lisouski, A. Mortensen, M. Hansen, T. Gregersen, and P. Ahrendt. Kinect depth sensor evaluation for computer vision applications. Technical Report ECE-TR-6, Department of Engineering, Aarhus University, 2012.
- [9] D. L. Baggio, S. Emami, D. M. Escrava, K. Ievgen, N. Mahmood, J. Saragih, and R. Shilkrot. *Mastering OpenCV with Practical Computer Vision Projects*. Packt Publishing, Limited, 2012. recommended: advanced OpenCV project support/examples inc. iOS and Android examples.
- [10] S. Baluja and H. A. Rowley. Boosting sex identification performance. *International Journal of Computer Vision*, 71(1):111–119, 2007.
- [11] C. Bauer and S. Spiekermann. Conceptualizing context for pervasive advertising. In Müller et al. [125], pages 159–183.
- [12] E. Bekkering and J. Shim. Trust in videoconferencing. *Communications of the ACM*, 49(7):103–107, 2006.
- [13] H. Benko, A. D. Wilson, and R. Balakrishnan. Sphere: multi-touch interactions on a spherical display. In S. B. Cousins and M. Beaudouin-Lafon, editors, *UIST*, pages 77–86. ACM, 2008.
- [14] G. Beyer, F. Alt, J. Müller, A. Schmidt, K. Isakovic, S. Klose, M. Schiewe, and I. Hauslen. Audience behavior around large interactive cylindrical screens. In D. S. Tan, S. Amershi, B. Begole, W. A. Kellogg, and M. Tungare, editors, *CHI*, pages 1021–1030. ACM, 2011.
- [15] G. Beyer, F. Köttner, M. Schiewe, I. Hauslen, and A. Butz. Squaring the circle: how framing influences user behavior around a seamless cylindrical display. In Mackay et al. [112], pages 1729–1738.
- [16] D. J. Beymer. Face recognition under varying pose. In *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94, 1994 IEEE Computer Society Conference on*, pages 756–761. IEEE, 1994.
- [17] S. Boring and D. Baur. Making public displays interactive everywhere. *Computer Graphics and Applications, IEEE*, 33(2):28–36, 2013.
- [18] L. Bourdev, S. Maji, and J. Malik. Describing people: A poselet-based approach to attribute classification. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 1543–1550. IEEE, 2011.
- [19] T. Bouwmans, F. E. Baf, and B. Vachon. Background modeling using mixture of gaussians for foreground detection - a survey. *Recent Patents on Computer Science*, 1(3):219–237, 2008.
- [20] G. Bradski and A. Kachler. *Learning OpenCV: Computer Vision with the OpenCV Library*. O'Reilly Media, 2008.

- [21] P. Brey. Freedom and privacy in ambient intelligence. *Ethics and Information Technology*, 7(3):157–166, Sept. 2005.
- [22] R. Brunelli. *Template Matching Techniques in Computer Vision: Theory and Practice*. Wiley, 2009.
- [23] S. Brutzer, B. Hoferlin, and G. Heidemann. Evaluation of background subtraction techniques for video surveillance. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 1937–1944, 2011.
- [24] R. R. Burke. The third wave of marketing intelligence. In M. Krafft and M. Mantrala, editors, *Retailing in the 21st Century: Current and Future Trends*, pages 159–171. Springer, 2006.
- [25] R. R. Burke. Behavioral effects of digital signage. *Journal of Advertising Research*, 49(2):180–185, 2009.
- [26] J. Carletta. Assessing agreement on classification tasks: the kappa statistic. *Computational Linguistics*, 22(2):249–254, 1996.
- [27] L. Cerrato, M. Falcone, and A. Paoloni. Subjective age estimation of telephonic voices. *Speech Communication*, 31(2):107–112, 2000.
- [28] Q. Chen, F. Malric, Y. Zhang, M. Abid, A. Cordeiro, E. M. Petriu, and N. D. Georganas. Interacting with digital signage using hand gestures. In M. S. Kamel and A. C. Campilho, editors, *ICIAI*, volume 5627 of *Lecture Notes in Computer Science*, pages 347–358. Springer, 2009.
- [29] K. Cheverst, A. Dix, D. Fitton, M. Rouncefield, and C. Graham. Exploring awareness related messaging through two situated-display-based systems. *Human-Computer Interaction*, 22(1-2):173–220, 2007.
- [30] E. F. Churchill, L. Nelson, L. Denoue, J. Helfman, and P. Murphy. Sharing multimedia content with interactive public displays: a case study. In D. Benyon, P. Moody, D. Gruen, and I. McAra-McWilliam, editors, *Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques*, DIS '04, pages 7–16, New York, NY, USA, 2004. ACM.
- [31] E. J. Clarke and B. A. Barton. Entropy and mdl discretization of continuous variables for bayesian belief networks. *International Journal of Intelligent Systems*, 15(1):61–92, 2000.
- [32] D. L. Collins. Anamorphosis and the eccentric observer: History, technique, and current practice. *Leonardo*, 25(2):179–187, 1992.
- [33] D. L. Collins. Anamorphosis and the eccentric observer: Inverted perspective and construction of the gaze. *Leonardo*, 25(1):72–82, 1992.
- [34] R. T. Collins, A. Lipton, T. Kanade, H. Fujiyoshi, D. Duggins, Y. Tsin, D. Tolliver, N. Enomoto, O. Hasegawa, P. Burt, et al. *A system for video surveillance and monitoring*, volume 2. Carnegie Mellon University, the Robotics Institute Pittsburgh, 2000.
- [35] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(5):603–619, 2002.
- [36] T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(6):681–685, 2001.
- [37] J. L. Crowley and F. Berard. Multi-modal tracking of faces for video communications. In *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*, pages 640–645. IEEE, 1997.
- [38] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati. Detecting moving objects, ghosts, and shadows in video streams. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 25(10):1337–1342, 2003.
- [39] T. Curk, J. Demšar, Q. Xu, G. Leban, U. Petrovič, I. Bratko, G. Shaulsky, and B. Zupan. Microarray data mining with visual programming. *Bioinformatics*, 21:396–398, Feb. 2005.
- [40] Y. Dai and Y. Nakano. Face-texture model based on sgld and its application in face detection in a color scene. *Pattern recognition*, 29(6):1007–1017, 1996.
- [41] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 886–893. IEEE, 2005.
- [42] J. Damjan and S. Možina. *Obmašanje potrošnikov*. Ekonomska fakulteta, 1999.
- [43] N. Davies, M. Langheinrich, R. José, and A. Schmidt. Open display networks: A communications medium for the 21st century. *IEEE Computer*, 45(5):58–64, 2012.
- [44] M. M. Dehshibi and A. Bastanfard. A new algorithm for age recognition from facial images. *Signal Processing*, 90(8):2431–2444, 2010.
- [45] D. DeMenthon and L. S. Davis. Model-based object pose in 25 lines of code. *International Journal of Computer Vision*, 15(1-2):123–141, 1995.
- [46] J. Demšar, B. Zupan, G. Leban, and T. Curk. Orange: From experimental machine learning to interactive data mining. *Knowledge discovery in databases: PKDD 2004*, pages 537–539, 2004.

- [47] C. Dennis, R. Michon, J. J. Brakus, A. Newman, and E. Alamanos. New insights into the impact of digital signage as a retail atmospheric tool. *Journal of Consumer Behaviour*, 11(6):454–466, 2012.
- [48] C. Dennis, A. Newman, R. Michon, J. J. Brakus, and L. T. Wright. The mediating effects of perception and emotion: Digital signage in mall atmospherics. *Journal of Retailing and Consumer Services*, 17(3):205–215, 2010.
- [49] C. A. Dodd, I. Clarke, and M. H. Kirkup. Camera observations of customer behaviour in fashion retailing: Methodological propositions. *International Journal of Retail and Distribution Management*, 26(8):311–317, 1998.
- [50] N. A. Dodgson. Variation and extrema of human interpupillary distance. In *Proc. SPIE 5291*, pages 36–46. International Society for Optics and Photonics, 2004.
- [51] M. Dumont, S. Maesen, S. Rogmans, and P. Bekaert. A prototype for practical eye-gaze corrected video chat on graphics hardware. In P. A. A. Assunção and S. M. M. de Faria, editors, *SIGMAP 2008*, pages 236–243. INSTICC Press, 2008.
- [52] F. Dupin and M. Adolph. *Digital signage: the right information in all the right places*. ITU, 2011.
- [53] A. M. Elgammal, D. Harwood, and L. S. Davis. Non-parametric model for background subtraction. In *Proceedings of the 6th European Conference on Computer Vision-Part II, ECCV '00*, pages 751–767, London, UK, UK, 2000. Springer-Verlag.
- [54] I. Essa. Computers seeing people. *AI Magazine*, 20(1):69–82, Summer 1999.
- [55] L. Farkas. *Anthropometry of the head and face*. Raven Press, 1994.
- [56] R. Feraund, O. J. Bernier, J.-E. Viallet, and M. Collobert. A fast and accurate face detector based on neural networks. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(1):42–53, 2001.
- [57] P. T. Fischer, C. Zöllner, T. Hoffmann, S. Piazza, and E. Hornecker. Beyond information and utility: Transforming public spaces with media facades. *Computer Graphics and Applications, IEEE*, 33(2):38–46, 2013.
- [58] D. A. Forsyth and J. Ponce. *Computer Vision: A Modern Approach*. Prentice Hall Professional Technical Reference, 2002.
- [59] A. Friday, N. Davies, and C. Efstratiou. Reflections on long-term experiments with public displays. *IEEE Computer*, 45(5):34–41, 2012.
- [60] Y. Fu, G. Guo, and T. S. Huang. Age synthesis and estimation via faces: A survey. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(11):1955–1976, 2010.
- [61] Y. Fu and T. S. Huang. Human age estimation with regression on discriminative aging manifold. *Multimedia, IEEE Transactions on*, 10(4):578–584, 2008.
- [62] T. Fujiwara and H. Koshimizu. Age and gender estimation by modeling statistical relationship among faces. In *Quality Control by Artificial Vision*, pages 559–566. International Society for Optics and Photonics, 2003.
- [63] H. L. Geiger. A standard for digital signage privacy. In Müller et al. [125], pages 103–117.
- [64] J. Gemmell, K. Toyama, C. L. Zitnick, T. Kang, and S. Seitz. Gaze awareness for video-conferencing: A software approach. *Multimedia, IEEE*, 7(4):26–35, 2000.
- [65] X. Geng, Z.-H. Zhou, and K. Smith-Miles. Automatic age estimation based on facial aging patterns. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(12):2234–2240, 2007.
- [66] T. Gramegna, L. Venturino, G. Cicirelli, G. Attolico, and A. Distante. Optimization of the posit algorithm for indoor autonomous navigation. *Robotics and Autonomous Systems*, 48(2):145–162, 2004.
- [67] R. M. Gray. *Entropy and information theory*. Springer-Verlag New York, Inc., New York, NY, USA, 1990.
- [68] D. M. Grayson and A. F. Monk. Are you looking at me? eye contact and desktop video conferencing. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 10(3):221–243, 2003.
- [69] G. Guo, Y. Fu, C. R. Dyer, and T. S. Huang. Image-based human age estimation by manifold learning and locally adjusted robust regression. *Image Processing, IEEE Transactions on*, 17(7):1178–1188, 2008.
- [70] G. Guo, Y. Fu, T. S. Huang, and C. R. Dyer. Locally adjusted robust regression for human age estimation. In *Applications of Computer Vision, 2008. WACV 2008. IEEE Workshop on*, pages 1–6. IEEE, 2008.
- [71] M. Hacher, B. Bossavit, A. Cohé, and J.-B. de la Rivière. Toucheo: multitouch and stereo combined in a seamless workspace. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pages 587–592. ACM, 2011.
- [72] M. Hacher, J.-B. de la Rivière, J. Laviolle, A. Cohé, and S. Cursan. Touch-based interfaces for interacting with 3d content in public exhibitions. *Computer Graphics and Applications, IEEE*, 33(2):80–85, 2013.

- [73] J. Hardy, E. Rukzio, and N. Davies. Real world responses to interactive gesture based public displays. In Q. Dai, R. Jain, X. Ji, and M. Kranz, editors, *MUM*, pages 33–39. ACM, 2011.
- [74] L. Hespanhol, M. Tomitsch, K. Grace, A. Collins, and J. Kay. Investigating intuitiveness and effectiveness of gestures for free spatial interaction with large displays. In *Proceedings of the 2012 International Symposium on Pervasive Displays*, PerDis '12, pages 6:1–6:6, New York, NY, USA, 2012. ACM.
- [75] T. T. Hewett, editor. *ACM SIGCHI Curricula for Human-Computer Interaction*. ACM Press, New York, NY, 1992.
- [76] A. P. Hillstrom and S. Yantis. Visual motion and attentional capture. *Perception And Psychophysics*, 55(4):399–411, 1994.
- [77] U. Hinrichs, S. Carpendale, N. Valkanova, K. Kuikkaniemi, G. Jacucci, and A. V. Moere. Interactive public displays. *IEEE Computer Graphics and Applications*, 33(2):25–27, 2013.
- [78] M. Hollander and D. A. Wolfe. *Nonparametric Statistical Methods*. John Wiley & Sons, New York, 1999.
- [79] R.-L. Hsu, M. Abdel-Mottaleb, and A. K. Jain. Face detection in color images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(5):696–706, 2002.
- [80] M. Hu, Y. Wang, Z. Zhang, and D. Zhang. Gait-based gender classification using mixed conditional random field. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 41(5):1429–1439, 2011.
- [81] E. M. Huang, A. Koster, and J. Borchers. Overcoming assumptions and uncovering practices: When does the public really look at public displays? In J. Indulska, D. J. Patterson, T. Rodden, and M. Ott, editors, *Pervasive*, volume 5013 of *Lecture Notes in Computer Science*, pages 228–243. Springer, 2008.
- [82] M.-C. Hwang, N.-H. Kim, C.-S. Park, S.-J. Ko, et al. Person identification system for future digital tv with intelligence. *Consumer Electronics, IEEE Transactions on*, 53(1):218–226, 2007.
- [83] A. Jakulin and I. Bratko. Analyzing attribute dependencies. In N. Lavrac, D. Gamberger, H. Blockeel, and L. Todorovski, editors, *PKDD*, volume 2838 of *Lecture Notes in Computer Science*, pages 229–240. Springer, 2003.
- [84] A. Jakulin and I. Bratko. Testing the significance of attribute interactions. In C. E. Brodley, editor, *ICML*, volume 69 of *ACM International Conference Proceeding Series*, pages 409–416. ACM, 2004.
- [85] M. Jones and P. Viola. Fast multi-view face detection. *Mitsubishi Electric Research Lab TR-20003-96*, 3:14, 2003.
- [86] R. José, H. Pinto, B. Silva, and A. Melro. Pins and posters: Paradigms for content publication on situated displays. *Computer Graphics and Applications, IEEE*, 33(2):64–72, 2013.
- [87] R. José, H. Pinto, B. Silva, A. Melro, and H. Rodrigues. Beyond interaction: tools and practices for situated publication in display networks. In *Proceedings of the 2012 International Symposium on Pervasive Displays*, PerDis '12, pages 8:1–8:6, New York, NY, USA, 2012. ACM.
- [88] T. Kailath. The divergence and bhattacharyya distance measures in signal selection. *Communication Technology, IEEE Transactions on*, 15(1):52–60, 1967.
- [89] K. Koffka. *Principles of Gestalt Psychology*. Harcourt, Brace and World, New York, 1935.
- [90] M. T. Koppel, G. Bailly, J. Müller, and R. Walter. Chained displays: configurations of public displays can be used to influence actor-, audience-, and passer-by behavior. In J. A. Konstan, E. H. Chi, and K. Höök, editors, *CHI*, pages 317–326. ACM, 2012.
- [91] V. Kostakos and T. Ojala. Public displays invade urban spaces. *IEEE Pervasive Computing*, 12(1):8–13, 2013.
- [92] P. Kotler and K. L. Keller. *Marketing Management*. Pearson, 2006.
- [93] J. Kovač, P. Peer, and F. Solina. Human skin colour clustering for face detection. In *EUROCON 2003, Computer as a Tool. The IEEE Region 8*, volume 2, pages 144–148, 2003.
- [94] V. Krüger and G. Sommer. Gabor wavelet networks for efficient head pose estimation. *Image and vision computing*, 20(9):665–672, 2002.
- [95] J. Krumm. Ubiquitous advertising: The killer application for the 21st century. *IEEE Pervasive Computing*, 10(1):66–73, 2011.
- [96] W. Kruskal and W. Wallis. Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association*, 47(260):583–621, 1952.
- [97] K. Kuikkaniemi, G. Jacucci, M. Turpeinen, E. E. Hoggan, and J. Müller. From space to stage: How interactive screens will change urban life. *IEEE Computer*, 44(6):40–47, 2011.
- [98] C. Kuster, T. Popa, J.-C. Bazin, C. Gotsman, and M. Gross. Gaze correction for home video conferencing. *ACM Transactions on Graphics (TOG)*, 31(6):174, 2012.

- [99] Y. H. Kwon and N. da Vitoria Lobo. Age classification from facial images. *Computer Vision and Image Understanding*, 74(1):1–21, 1999.
- [100] M. La Cascia, S. Sclaroff, and V. Athitsos. Fast, reliable head tracking under varying illumination: An approach based on registration of texture-mapped 3d models. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(4):322–336, 2000.
- [101] M. Langheinrich. Privacy by design - principles of privacy-aware ubiquitous systems. In *Ubicomp*, pages 273–291, 2001.
- [102] A. Lanitis, C. J. Taylor, and T. F. Cootes. Toward automatic simulation of aging effects on face images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(4):442–455, 2002.
- [103] S. Li and A. Jain. *Handbook of Face Recognition*. Springer, 2005.
- [104] Y. Li, S. Gong, J. Sherrah, and H. Liddell. Support vector machine based multi-view face detection and recognition. *Image and Vision Computing*, 22(5):413–427, 2004.
- [105] R. Lienhart, A. Kuranov, and V. Pisarevsky. Empirical analysis of detection cascades of boosted classifiers for rapid object detection. In *DAGM-Symposium*, pages 297–304, 2003.
- [106] R. Lienhart and J. Maydt. An extended set of haar-like features for rapid object detection. In *Image Processing. 2002. Proceedings. 2002 International Conference on*, volume 1, pages 1–900. IEEE, 2002.
- [107] J.-Y. Lin, Y.-Y. Chen, J.-C. Ko, H. Kao, W.-H. Chen, T.-H. Tsai, S.-C. Hsu, and Y.-P. Hung. i-m-tube: an interactive multi-resolution tubular display. In W. Gao, Y. Rui, A. Hanjalic, C. Xu, E. G. Steinbach, A. El-Saddik, and M. X. Zhou, editors, *ACM Multimedia*, pages 253–260. ACM, 2009.
- [108] T. Lindén, T. Heikkinen, V. Kostakos, D. Ferreira, and T. Ojala. Towards multi-application public interactive displays. In *Proceedings of the 2012 International Symposium on Pervasive Displays, PerDis '12*, pages 9:1–9:5, New York, NY, USA, 2012. ACM.
- [109] C. Liu. A bayesian discriminating features method for face detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 25(6):725–740, 2003.
- [110] J. Lu and Y.-P. Tan. Gait-based human age estimation. *Information Forensics and Security, IEEE Transactions on*, 5(4):761–770, 2010.
- [111] L. I. Lundström. *Digital Signage Broadcasting: Content Management and Distribution Techniques*. Focal Press Media Technology Professional. Focal Press, 2008.
- [112] W. E. Mackay, S. A. Brewster, and S. Bodker, editors. *2013 ACM SIGCHI Conference on Human Factors in Computing Systems, CHI '13, Paris, France, April 27 - May 2, 2013*. ACM, 2013.
- [113] L. Maddalena and A. Petrosino. A self-organizing approach to background subtraction for visual surveillance applications. *Image Processing, IEEE Transactions on*, 17(7):1168–1177, 2008.
- [114] P. C. Mahalanobis. On the generalized distance in statistics. *Proceedings of the National Institute of Sciences (Calcutta)*, 2:49–55, 1936.
- [115] E. Makinen and R. Raisamo. Evaluation of gender classification methods with automatically detected and aligned faces. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(3):541–547, 2008.
- [116] T. Mandel. *The elements of user interface design*. Wiley computer publishing. Wiley, 1997.
- [117] I. Matthews and S. Baker. Active appearance models revisited. *International Journal of Computer Vision*, 60(2):135–164, 2004.
- [118] S. J. McKenna, Y. Raja, and S. Gong. Tracking colour objects using adaptive mixture models. *Image and vision computing*, 17(3):225–231, 1999.
- [119] T. Mei, J. Guo, X.-S. Hua, and F. Liu. Adon: toward contextual overlay in-video advertising. *Multimedia Syst.*, 16(4-5):335–344, 2010.
- [120] N. Memarovic, M. Langheinrich, and F. Alt. The interacting places framework: conceptualizing public display applications that promote community interaction and place awareness. In *Proceedings of the 2012 International Symposium on Pervasive Displays, PerDis '12*, pages 7:1–7:6, New York, NY, USA, 2012. ACM.
- [121] E. Mencarini, L. Giusti, and M. Zancanaro. An investigation on acceptance and rejection of public displays in a knowledge company. In *Proceedings of the 2012 International Symposium on Pervasive Displays, PerDis '12*, pages 16:1–16:6, New York, NY, USA, 2012. ACM.
- [122] D. Michelis and J. Müller. The audience funnel: Observations of gesture based interaction with multiple large displays in a city center. *International Journal of Human Computer Interaction*, 27(6):562–579, 2011.
- [123] B. Moghaddam and M.-H. Yang. Learning gender with support faces. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(5):707–711, 2002.

- [124] L.-P. Morency, A. Rahimi, N. Checka, and T. Darrell. Fast stereo-based head tracking for interactive environments. In *Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on*, pages 390–395. IEEE, 2002.
- [125] H. J. Müller, F. Alt, and D. Michelis, editors. *Pervasive Advertising*. Human-Computer Interaction Series. Springer, 2011.
- [126] J. Müller, F. Alt, D. Michelis, and A. Schmidt. Requirements and design space for interactive public displays. In A. D. Bimbo, S.-F. Chang, and A. W. M. Smeulders, editors, *ACM Multimedia*, pages 1285–1294. ACM, 2010.
- [127] J. Müller and A. Krüger. How much to bid in digital signage advertising auctions. In *Adjunct proceedings of Pervasive*, volume 2007, 2007.
- [128] J. Müller and A. Krüger. Mobidic: Context adaptive digital signage with coupons. In M. Tscheligi, B. E. R. de Ruyter, P. Markopoulos, R. Wichert, T. Mirlacher, A. Meschtscherjakov, and W. Reitberger, editors, *AMI*, volume 5859 of *Lecture Notes in Computer Science*, pages 24–33. Springer, 2009.
- [129] J. Müller, O. Paczkowski, and A. Krüger. Situated public news and reminder displays. In B. Schiele, A. K. Dey, H. Gellersen, B. E. R. de Ruyter, M. Tscheligi, R. Wichert, E. H. L. Aarts, and A. P. Buchmann, editors, *AMI*, volume 4794 of *Lecture Notes in Computer Science*, pages 248–265. Springer, 2007.
- [130] J. Müller, R. Walter, G. Bailly, M. Nischt, and F. Alt. Looking glass: a field study on noticing interactivity of a shop window. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '12*, pages 297–306. New York, NY, USA, 2012. ACM.
- [131] J. Müller, D. Wilmshann, J. Exeler, M. Buzeck, A. Schmidt, T. Jay, and A. Krüger. Display blindness: The effect of expectations on attention towards digital signage. In H. Tokuda, M. Beigl, A. Friday, A. J. B. Brush, and Y. Tobe, editors, *Pervasive*, volume 5538 of *Lecture Notes in Computer Science*, pages 1–8. Springer, 2009.
- [132] E. Murphy-Chutorian and M. M. Trivedi. Head pose estimation in computer vision: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.*, 31(4):607–626, 2009.
- [133] A. Newman, C. Dennis, L.-T. Wright, and T. Kingh. Shoppers' experiences of digital signage—a cross-national qualitative study. *International Journal of Digital Content Technology and its Applications*, 4(7):50–57, 2010.
- [134] C. B. Ng, Y. H. Tay, and B.-M. Goi. Recognizing human gender in computer vision: a survey. In *PRICAI 2012: Trends in Artificial Intelligence*, pages 335–346. Springer, 2012.
- [135] B. Ni, Z. Song, and S. Yan. Web image mining towards universal age estimator. In *Proceedings of the 17th ACM international conference on Multimedia*, pages 85–94. ACM, 2009.
- [136] T. Ojala, V. Kostakos, H. Kukka, T. Heikkinen, T. Lindén, M. Jurmu, S. Hosio, F. Kruger, and D. Zanni. Multipurpose interactive public displays in the wild: Three years later. *IEEE Computer*, 45(5):42–49, 2012.
- [137] N. M. Oliver, B. Rosario, and A. P. Pentland. A bayesian computer vision system for modeling human interactions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(8):831–843, 2000.
- [138] M. Osadchy, Y. L. Cun, and M. L. Miller. Synergistic face detection and pose estimation with energy-based models. *The Journal of Machine Learning Research*, 8:1197–1215, 2007.
- [139] M. Ostkamp, G. Bauer, and C. Kray. Visual highlighting on public displays. In *Proceedings of the 2012 International Symposium on Pervasive Displays, PerDis '12*, pages 2:1–2:6. New York, NY, USA, 2012. ACM.
- [140] M. Ott, J. P. Lewis, and I. Cox. Teleconferencing eye contact using a virtual camera. In *INTERACT '93 and CHI '93 Conference Companion on Human Factors in Computing Systems*, pages 109–110. ACM, 1993.
- [141] E. Pantano. New technologies and retailing: Trends and directions. *Journal of Retailing and Consumer Services*, 17(3), 2010.
- [142] E. Pantano and V. Corvello. Digital contents management for improving consumers' experience. *Journal of Digital Contents Technology and its Applications*, 4(7), 2010.
- [143] E. Pantano and G. Naccarato. Entertainment in retailing: the influences of advanced technologies. *Journal of Retailing and Consumer Services*, 17(3), 2010.
- [144] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss. The feret evaluation methodology for face-recognition algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1090–1104, 2000.
- [145] P. J. Phillips, H. Wechsler, J. Huang, and P. J. Rauss. The feret database and evaluation procedure for face-recognition algorithms. *Image and vision computing*, 16(5):295–306, 1998.
- [146] M. Piccardi. Background subtraction techniques: a review. In *Systems, man and cybernetics, 2004 IEEE international conference on*, volume 4, pages 3099–3104. IEEE, 2004.

- [147] Z. Pizlo. A theory of shape constancy based on perspective invariants. *Vision Research*, 34(12):1637–1658, 1994.
- [148] V. Potočnik. *Temeljni trženja: s primeri iz prakse*. GV založba, 2002.
- [149] P. Rai and P. Khanna. Gender classification techniques: A review. In *Advances in Computer Science, Engineering & Applications*, pages 51–59. Springer, 2012.
- [150] N. Ramanathan and R. Chellappa. Face verification across age progression. *Image Processing, IEEE Transactions on*, 15(11):3349–3361, 2006.
- [151] R. Ravnik, B. Batagelj, B. Kverh, and F. Solina. Dynamic anamorphosis as a special, computer generated computer generated interface. *Interacting with Computers*, 2013.
- [152] R. Ravnik and F. Solina. Audience measurement of digital signage: Quantitative study in real-world environment using computer vision. *Interacting with Computers*, 25(3):218–228, 2013.
- [153] R. Ravnik and F. Solina. Interactive and audience adaptive digital signage using real-time computer vision. *International Journal of Advanced Robotic Systems*, 10(107):1–7, 2013.
- [154] R. W. Remington, J. C. Johnston, and S. Yantis. Involuntary attentional capture by abrupt onsets. *Perception and Psychophysics*, 51(3):279–290, 1992.
- [155] G. Ren, C. Li, E. O'Neill, and P. Willis. 3d freehand gestural navigation for interactive public displays. *Computer Graphics and Applications, IEEE*, 33(2):47–55, 2013.
- [156] S. T. Roweis and L. K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500):2323–2326, 2000.
- [157] H. A. Rowley, S. Baluja, and T. Kanade. Neural network-based face detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 20(1):23–38, 1998.
- [158] J. Saragih and R. Göcke. Learning aam fitting through simulation. *Pattern Recognition*, 42(11):2628–2636, 2009.
- [159] J. M. Saragih, S. Lucey, and J. F. Cohn. Deformable model fitting by regularized landmark mean-shift. *International Journal of Computer Vision*, 91(2):200–215, 2011.
- [160] I. Satoh. An agent-based framework for context-aware digital signage. In J. C. Augusto, J. M. Corchado, P. Novais, and C. Analide, editors, *ISAmI*, volume 72 of *Advances in Soft Computing*, pages 105–112. Springer, 2010.
- [161] I. Satoh. A framework for context-aware digital signage. In N. Zhong, V. Callaghan, A. A. Ghorbani, and B. Hu, editors, *AMT*, volume 6890 of *Lecture Notes in Computer Science*, pages 251–262. Springer, 2011.
- [162] J. Schaeffler. *Digital Signage: Software, Networks, Advertising, and Displays*. Focal Press, 2008.
- [163] R. Schalkoff. *Digital image processing and computer vision*. John Wiley & Sons, 1989.
- [164] A. Seckel. *Masters of Deception: Escher, Dalí & the Artists of Optical Illusion*. Sterling Publishers, 2004.
- [165] M. Seki, T. Wada, H. Fujiwara, and K. Sumi. Background subtraction based on cooccurrence of image variations. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, volume 2, pages 11–65. IEEE, 2003.
- [166] C. Shan. Learning local binary patterns for gender classification on real-world face images. *Pattern Recognition Letters*, 33(4):431–437, 2012.
- [167] J. Sherrah, S. Gong, and E.-J. Ong. Face distributions in similarity space under varying head pose. *Image and Vision Computing*, 19(12):807–819, 2001.
- [168] J. Shi and J. Malik. Normalized cuts and image segmentation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(8):888–905, 2000.
- [169] B. Shneiderman. *Designing the User Interface*. Addison-Wesley, Reading, 1998.
- [170] F. Solina and R. Ravnik. Fixing missing eye-contact in video conferencing systems. In Z. B. V. Luzar-Stiler, I. Jarec, editor, *Information Technology Interfaces (ITI), Proceedings of the ITI 2011 33rd International Conference on*, pages 233–236, 2011.
- [171] M. Solomon. *Consumer Behaviour: A European Perspective*. Prentice Hall, 2006.
- [172] S. Spiekermann and L. F. Cranor. Engineering privacy. *IEEE Trans. Software Eng.*, 35(1):67–82, 2009.
- [173] J. D. Spurrier. On the null distribution of the kruskal-wallis statistic. *Journal of Nonparametric Statistics*, 15(6):685–691, 2003.
- [174] C. Stauffer and W. E. L. Grimson. Adaptive background mixture models for real-time tracking. In *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on*, volume 2. IEEE, 1999.

- [175] R. Stiefelhagen, J. Yang, and A. Waibel. Modeling focus of attention for meeting indexing based on multiple cues. *Neural Networks, IEEE Transactions on*, 13(4):928–938, 2002.
- [176] G. Stockman and L. G. Shapiro. *Computer Vision*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 1st edition, 2001.
- [177] R. Stokes. Human factors and appearance design considerations of the mod ii picturephone station set. *IEEE Transactions on Communication Technology*, 17(2):318–323, 1969.
- [178] D. Stone, C. Jarrett, M. Woodroffe, and S. Minocha. *User Interface Design and Evaluation*. Morgan Kaufmann, Amsterdam, 2005.
- [179] F. Sukno, J. J. Guerrero, and A. F. Frangi. Projective active shape models for pose-variant image analysis of quasi-planar objects: Application to facial analysis. *Pattern Recognition*, 43(3):835–849, 2010.
- [180] F. Sukno, S. Ordas, C. Butakoff, S. Cruz, and A. F. Frangi. Active shape models with invariant optimal features: Application to facial analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(7):1105–1117, 2007.
- [181] R. Szeliski. *Computer Vision: Algorithms and Applications*. Springer-Verlag New York, Inc., New York, NY, USA, 1st edition, 2010.
- [182] N. Taylor and K. Cheverst. Rural communities and pervasive advertising. In Müller et al. [125], pages 269–286.
- [183] N. Taylor and K. Cheverst. Supporting community awareness with interactive displays. *IEEE Computer*, 45(5):26–32, 2012.
- [184] C. A. C. Teixeira, E. L. Melo, R. G. Cattelan, and M. da Graça Campos Pimentel. Taking advantage of contextualized interactions while users watch tv. *Multimedia Tools Appl.*, 50(3):587–607, 2010.
- [185] D. Topper. On anamorphosis: Setting some things straight. *Leonardo*, 33(2):115–124, 2000.
- [186] P. A. Viola and M. J. Jones. Robust real-time face detection. *International Journal of Computer Vision*, 57(2):137–154, 2004.
- [187] J.-G. Wang and E. Sung. Em enhancement of 3d head pose estimated by point at infinity. *Image and Vision Computing*, 25(12):1864–1874, 2007.
- [188] R. Want and B. N. Schilit. Interactive digital signage. *IEEE Computer*, 45(5):21–24, 2012.
- [189] C. A. Waring and X. Liu. Face detection using spectral histograms and svms. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 35(3):467–476, 2005.
- [190] A. Watve and S. Sural. Soccer video processing for the detection of advertisement billboards. *Pattern Recognition Letters*, 29(7):994–1006, 2008.
- [191] M. Weiser. The computer for the 21st century. *Scientific American*, 265(3):66–75, September 1991.
- [192] J. S. Wilkinson and J. Kolodzy. *Communication Technology Update and Fundamentals*, chapter Digital Signage, pages 133–144. Focal Press, 2012.
- [193] C. R. Wren, A. Azarbayejani, T. Darrell, and A. P. Pentland. Pfinder: Real-time tracking of the human body. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 19(7):780–785, 1997.
- [194] B. Wu, H. Ai, C. Huang, and S. Lao. Fast rotation invariant multi-view face detection based on real adaboost. In *Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on*, pages 79–84. IEEE, 2004.
- [195] M.-H. Yang, D. J. Kriegman, and N. Ahuja. Detecting faces in images: A survey. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(1):34–58, 2002.
- [196] Z. Yang, M. Li, and H. Ai. An experimental study on automatic face gender classification. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, volume 3, pages 1099–1102. IEEE, 2006.
- [197] B. Yip. Face and eye rectification in video conference using affine transform. In *ICIP 2005*, pages 513–516, 2005.
- [198] B. Yip and J. S. Jin. Pose determination and viewpoint determination of human head in video conferencing based on head movement. In Y.-P. P. Chen, editor, *10th International Multimedia Modeling Conference*, pages 130–135. IEEE Computer Society, 2004.
- [199] J. Ylioinas, A. Hadid, and M. Pietikäinen. Combining contrast information and local binary patterns for gender classification. In *Image Analysis*, pages 676–686. Springer, 2011.
- [200] K. C. Yow and R. Cipolla. Feature-based human face detection. *Image and vision computing*, 15(9):713–735, 1997.
- [201] S. Yu, T. Tan, K. Huang, K. Jia, and X. Wu. A study on gait-based gender classification. *Image Processing, IEEE Transactions on*, 18(8):1905–1910, 2009.
- [202] Zakon o varstvu osebnih podatkov. ZVOP-1. *Uradni list Republike Slovenije*, 86, 2004.

- [203] C. Zhang and Z. Zhang. A survey of recent advances in face detection. Technical report, Tech. rep., Microsoft Research, 2010.
- [204] H. Zhang, W. Gao, X. Chen, and D. Zhao. Object detection using spatial histogram features. *Image and Vision Computing*, 24(4):327–341, 2006.
- [205] Y. Zhang, A. Bulling, and H. Gellersen. Sideways: a gaze interface for spontaneous interaction with situated displays. In Mackay et al. [112], pages 851–860.
- [206] C. L. Zitnick, J. Gemmell, and K. Toyama. Manipulation of video eye gaze and head orientation for video teleconferencing. Technical Report MSR-TR-99-46, Microsoft Research, Microsoft Corporation, 1999.
- [207] Z. Zivkovic. Improved adaptive gaussian mixture model for background subtraction. In *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, volume 2, pages 28–31. IEEE, 2004.
- [208] Z. Zivkovic and F. van der Heijden. Efficient adaptive density estimation per image pixel for the task of background subtraction. *Pattern recognition letters*, 27(7):773–780, 2006.