

Slovenská technická univerzita v Bratislave
Fakulta informatiky a informačných technológií

FIIT-5208-72138

Bc. Tomáš Juhaniak

Vyhodnocovanie kognitívnej záťaže ako súčasť používateľských štúdií

Diplomová práca

Študijný program: Informačné systémy

Študijný odbor: Informačné systémy

Miesto vypracovania: Ústav informatiky, informačných systémov a softvérového inžinierstva, FIIT
STU v Bratislave

Vedúci práce: prof. Ing. Mária Bieliková, PhD

máj 2018

Zadanie diplomovej práce

Meno študenta: **Bc. Tomáš Juhaniak**

Študijný program: Informačné systémy

Študijný odbor: Informačné systémy

Názov práce: **Vyhodnocovanie kognitívnej záťaže ako súčasť používateľských štúdií**

Samostatnou výskumnou a vývojovou činnosťou v rámci predmetov Diplomový projekt I, II, III vypracujte diplomovú prácu na tému, vyjadrenú vyššie uvedeným názvom tak, aby ste dosiahli tieto ciele:

Všeobecný cieľ:

Vypracovaním diplomovej práce preukážte, ako ste si osvojili metódy a postupy riešenia relatívne rozsiahlych projektov, schopnosť samostatne a tvorivo riešiť zložité úlohy aj výskumného charakteru v súlade so súčasnými metódami a postupmi študovaného odboru využívanými v príslušnej oblasti a schopnosť samostatne, tvorivo a kriticky pristupovať k analýze možných riešení a k tvorbe modelov.

Špecifický cieľ:

Vytvorte riešenie zodpovedajúce návrhu textu zadania, ktorý je prílohou tohto zadania. Návrh bližšie opisuje tému vyjadrenú názvom. Tento opis je záväzný, má však rámcový charakter, aby vznikol dostatočný priestor pre Vašu tvorivosť.

Riadte sa pokynmi Vášho vedúceho.

Pokiaľ v priebehu riešenia, opierajúc sa o hlbšie poznanie súčasného stavu v príslušnej oblasti, alebo o priebežné výsledky Vášho riešenia, alebo o iné závažné skutočnosti, dospejete spoločne s Vaším vedúcim k presvedčeniu, že niečo v texte zadania a/alebo v názve by sa malo zmeniť, navrhnete zmenu. Zmena je spravidla možná len pri dosiahnutí kontrolného bodu.

Miesto vypracovania: Ústav informatiky, informačných systémov a softvérového inžinierstva, FIIT STU Bratislava

Vedúci práce: **prof. Ing. Mária Bieliková, PhD.**

Termíny odovzdania:

Podľa harmonogramu štúdia platného pre semester, v ktorom máte príslušný predmet (Diplomový projekt I, II, III) absolvovať podľa Vášho študijného plánu

Predmety odovzdania:

V každom predmete dokument podľa pokynov na www.fiit.stuba.sk v časti: home > Informácie o > štúdiu > organizácia štúdia > diplomový projekt.

V Bratislave dňa 13. 2. 2017

**SLOVENSKÁ TECHNICKÁ UNIVERZITA
V BRATISLAVE**

Fakulta informatiky a informačných technológií
Ilkovičova 2, 842 16 Bratislava 4

1

prof. Ing. Pavol Návrat, PhD.
riaditeľ Ústavu informatiky, informačných systémov
a softvérového inžinierstva

Návrh zadania diplomovej práce

Finálna verzia do diplomovej práce¹

Študent:

Meno, priezvisko, tituly: Tomáš Juhaniak, Bc.
Študijný program: Informačné systémy
Kontakt: juhaniak.tomas@gmail.com

Výskumník:

Meno, priezvisko, tituly: Mária Bieliková, prof. Ing. PhD.

Projekt:

Názov: Vyhodnocovanie kognitívnej záťaže ako súčasť používateľských štúdií
Názov v angličtine: Cognitive load evaluation as a part of user studies
Miesto vypracovania: Ústav informatiky, informačných systémov a softvérového inžinierstva, FIIT STU, Bratislava
Oblasť problematiky: Inteligentné spracovanie informácií

Text návrhu zadania²

Kognitívna záťaž je z pohľadu vyhodnocovania používateľského zážitku jedna z najdôležitejších metrík. Vo fáze identifikácie problémov prameniacych zo zložitosti riešených úloh alebo zle rozvrhnutého rozhrania aplikácií jej úroveň priamo odrzkadľuje úsilie, ktoré používateľ musí vynaložiť pri riešení konkrétneho problému. Táto vlastnosť ju zvýhodňuje oproti iným štandardne používaným metrikám používaným pri vyhodnocovaní používateľského zážitku. Navzdory jej vhodnosti je len veľmi zriedkavo skúmanou metrikou, keďže vyhodnotiť úroveň kognitívnej záťaže v reálnom prostredí webu je súčasnými metódami ťažko dosiahnuteľné. Problém tvorí najmä vplyv viacerých faktorov na dilatáciu zreničky, ktoré sa ťažko dajú oddeliť.

Analyzujte metódy vyhodnocovania kognitívnej záťaže, zamerajte sa pri tom na tie metódy, ktoré využívajú sledovanie dilatácie zreničky, čo predstavuje najmenej intruzívnu metódu využiteľnú pri používateľských štúdiách. Zoberte do úvahy vplyv ďalších faktorov na dilatáciu zreničky, ako napr. svietivosť stimulu či zmeny v prostredí. Analyzujte vlastnosti existujúcich metód, ich presnosť, využiteľnosť a intruzívnosť. Navrhňte vlastnú metódu vyhodnocovania kognitívnej záťaže v reálnom prostredí používateľských štúdií na webe. Porovnajte vlastnosti existujúcich metód s touto metódou s ohľadom na ich použitie v reálnom prostredí. Overte vlastnosti navrhutej metódy pomocou experimentu s úlohami navrhnutými tak, aby vhodne odrzkadľovali úrovne kognitívnej záťaže. Vlastnosti navrhnutého riešenia demonštrujte na identifikácii problémov v skutočnej webovej aplikácii.

¹ Vytlačiť obojstranne na jeden list papiera

² 150-200 slov (1200-1700 znakov), ktoré opisujú výskumný problém v kontexte súčasného stavu vrátane motivácie a smerov riešenia

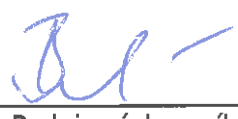
Literatúra³

- Siyuan Chen and Julien Epps. "Automatic classification of eye activity for cognitive load measurement with emotion interference". In: Computer methods and programs in biomedicine, Volume 110, Issue 2, Elsevier, 2013, pp. 111-124.
- WANG, Weihong, et al. Indexing cognitive workload based on pupillary response under luminance and emotional changes. In: Proceedings of the 2013 international conference on Intelligent user interfaces. ACM, 2013. p. 247-256.

Vyššie je uvedený návrh diplomového projektu, ktorý vypracoval(a) Bc. Tomáš Juhaniak, konzultoval(a) a osvojil(a) si ho prof. Ing. Mária Bieliková, PhD. a súhlasí, že bude takýto projekt viesť v prípade, že bude pridelený tomuto študentovi.

V Bratislave dňa 27.1.2017

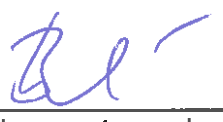

Podpis študenta


Podpis výskumníka

Vyjadrenie garanta predmetov Diplomový projekt I, II, III

Návrh zadania schválený: áno / nie⁴

Dňa: 13.2.2017


Podpis garanta predmetov

³ 2 vedecké zdroje, každý v samostatnej rubrike a s údajmi zodpovedajúcimi bibliografickým odkazom podľa normy STN ISO 690, ktoré sa viažu k téme zadania a preukazujú výskumnú povahu problému a jeho aktuálnosť (uvedte všetky potrebné údaje na identifikáciu zdroja, pričom uprednostnite vedecké príspevky v časopisoch a medzinárodných konferenciách)

⁴ Nehodiace sa prečiarknite

Čestné prehlásenie

Čestne vyhlasujem, že som túto diplomovú prácu vypracoval samostatne, iba s použitím uvedenej literatúry a podľa usmernení vedúceho diplomovej práce.

Bratislava, máj 2018

.....

Bc. Tomáš Juhaniak

Podakovanie

Moje podakovanie patrí prof. Ing. Márii Bielikovej, PhD. za odborné vedenie, pripomienky, rady a ochotu promptne odpovedať na množstvo otázok počas všetkých semestrov pri vypracovávaní tejto diplomovej práce. Ďakujem Ing. Róbertovi Mórovi, PhD., Ing. Petrovi Demčákovi a Bc. Dánielovi Pappovi za spoluprácu pri vykonávaných experimentoch. Za nezištnú pomoc pri nastavovaní experimentov ďakujem aj Patrikovi Brandýsovi. Za konštruktívnu diskusiu ďakujem členom výskumnej skupiny PeWe, s ktorými som neraz preberal možnosti ďalšieho posunu v mojom výskume.

Bc. Tomáš Juhaniak

Anotácia

Slovenská technická univerzita v Bratislave

FAKULTA INFORMATIKY A INFORMAČNÝCH TECHNOLOGIÍ

Študijný program: Informačné systémy

Autor: Bc. Tomáš Juhaniak

Diplomová práca: Vyhodnocovanie kognitívnej záťaže ako súčasť používateľských štúdií

Vedúci diplomovej práce: prof. Ing. Mária Bieliková, PhD

máj 2018

V našej práci riešime možnosti vyhodnocovania kognitívnej záťaže v používateľských štúdiách pomocou dilatácie zrenice. Na základe preukázaných vlastností zrenice v kognitívnej psychológii sme navrhli metódu založenú na absolútnej dilatácii. Tento prístup bol doposiaľ považovaný za veľmi ťažko realizovateľný pre komplexné úlohy v reálnych aplikáciách.

Myšlienkou metódy je odfiltrovanie vplyvu vnímanej svietivosti na základe zobrazeného rozhrania a bodu fixácie pohľadu účastníka používateľskej štúdie. Po odfiltrovaní dilatácie spôsobenej vnímanou svietivosťou považujeme zostatkovú dilatáciu za vplyv kognitívnej záťaže.

Výstupom práce je návrh modelu reakcií zrenice, jeho kalibrácia s dvomi alternatívami výberu tréningových dát a metóda, ktorá dokáže vyhodnotiť prítomnosť kognitívnej záťaže bez ohľadu na svietivosť rozhrania. Preukázali sme funkčnosť navrhnutého riešenia a navrhli postup pre výber vhodnej kalibrácie pre vyhodnocovanie kognitívnej záťaže v používateľských štúdiách v závislosti od technických možností a cieľov výskumníka.

Navrhnuté riešenie sme poskytli výskumníkom v podobe interaktívnej webovej aplikácie, ktorá je schopná spracovať okulometrické dáta viacerých formátov a poskytnúť vyhodnotenie kognitívnej záťaže jednotlivých účastníkov experimentu, či celkovú kvantitatívnu analýzu.

Annotation

Slovak University of Technology in Bratislava

FACULTY OF INFORMATICS AND INFORMATION TECHNOLOGIES

Degree course: Information Systems

Author: Bc. Tomáš Juhaniak

Master Thesis: Cognitive load evaluation as a part of user studies

Supervisor: Prof. Mária Bieliková

may 2018

In our work we solve the possibilities of evaluating the cognitive load in user studies based on pupil dilatation. With reference to demonstrated properties of pupil in the field of cognitive psychology, we proposed and evaluated a method based on absolute pupil dilatation, which works for real world web applications. This task has been so far considered very difficult to realize for the complex tasks in real-world applications.

The idea behind the method is to filter out the effect of perceived luminescence of screen using information about displayed user interface and the fixation point of the user's gaze. After filtering the dilatation caused by perceived luminescence, we consider residual dilatation as a cognitive load.

We have proposed a model of pupil reactions, its calibration with two alternatives of train data selecting, and a method that can evaluate the presence of cognitive load regardless of the luminescence of the interface. We demonstrated the functionality of the proposed solution and we defined a guide to selecting a suitable calibration for cognitive load evaluation in user studies, depending on the technical capabilities and objectives of the researcher.

The proposed solution has been provided to researchers in the form of an interactive web application that is able to process multiple-format oculometric data and it provide an assessment of cognitive load for individual experiment participants or total quantitative analysis.

Obsah

1	Úvod	1
1.1	Cieľ práce	2
1.2	Štruktúra práce	3
2	Testovanie použiteľnosti rozhraní aplikácií	5
2.1	Vyhodnocovanie navrhnutých rozhraní	5
2.2	Testovanie použiteľnosti	5
2.2.1	Subjektívne hodnotenie náročnosti úlohy	5
2.2.2	Sledovanie času vykonávania úlohy	5
2.2.3	Miera správnosti vykonaných úloh	6
2.2.4	Frekvencia žmurkania očí	6
2.2.5	Sledovanie dilatácie zrenice	6
2.2.6	Porovnanie metód pre vyhodnocovanie kognitívnej záťaže	6
2.3	Štatistické prístupy v testovaní použiteľnosti aplikácií	6
3	Zdroje dilatácie zrenice	9
3.1	Kognitívna záťaž a dilatácia zrenice	9
3.2	Iné zdroje dilatácie zrenice	10
3.2.1	Vnímaná svietivosť	10
3.2.2	Emócie	11
3.2.3	Hippus	11
3.3	Diskusia	11
4	Prístupy k odhadu kognitívnej záťaže z dilatácie zrenice	13
4.1	Priebeh dilatácie zrenice pri kognitívnej záťaži	13
4.2	Kvantifikácia kognitívnej záťaže	14
4.3	Referenčná šírka zrenice	14
4.4	Nerovnomerná svietivosť stimulu a okolia	15
4.4.1	Homogénne rozhrania	15
4.4.2	Heterogénne rozhrania	16
4.4.3	Svietivosť okolia monitora	16
4.5	Existujúce metódy spracovania dilatácie zrenice pri práci s počítačom	17
4.5.1	Metóda výpočtu rozdielu maximálnej a minimálnej šírky zrenice	17
4.5.2	Metóda štandardnej odchýlky šírky zrenice	17
4.5.3	Metóda rýchlosti a akcelerácie dilatácie	17
4.5.4	Metódy indexácie kognitívnej aktivity	18
4.6	Metóda modelovaním reakcií zrenice	18
4.6.1	Definícia modelu reakcií zrenice	18
4.6.2	Kalibrácia modelu	20
4.6.3	Normalizácia dilatácie s natrénovaným modelom	21

4.6.4	Výpočet svietivosti pre použitie v modeli	22
4.7	Diskusia	23
5	Metóda odhadovania kognitívnej záťaže pre heterogénne rozhrania založená na modeli reakcií zrenice	25
5.1	Kontext riešeného problému	25
5.2	Kroky riešenia problému	25
5.3	Zdokonalená explicitná kalibrácia	26
5.4	Univerzálna matica projekcie sietnice	27
5.4.1	Bázová matica projekcie sietnice	28
5.4.2	Výrez finálnej matice z bázovej matice projekcie sietnice	29
5.5	Vnímaná svietivosť okolia	30
5.5.1	Kalibrácia svietivosti okolia	30
5.5.2	Celková vnímaná svietivosť	32
5.5.3	Absolútny vplyv okolia na vnímanú svietivosť	33
5.6	Implicitná kalibrácia modelu reakcií zrenice	34
5.6.1	Postupnosť krokov pri implicitnej kalibrácii	34
5.6.2	Kalibrácia bez známej svietivosti okolia	35
5.6.3	Kalibrácia so známou svietivosťou okolia	37
5.6.4	Ohraničenia implicitnej kalibrácie krivky referenčných hodnôt	38
5.7	Návrh kvantitatívneho štatistického spracovania	39
5.7.1	Atomická jednotka kognitívnej záťaže	39
5.7.2	Dvojfázová normalizácia kognitívnej záťaže	40
5.7.3	Kvantitatívne vyhodnotenie používateľských štúdií	41
5.8	Diskusia	42
5.8.1	Prínosy	42
5.8.2	Ohraničenia	43
6	Návrh efektívneho vyhodnocovania kognitívnej záťaže v experimentoch	45
6.1	Predspracovanie okulometrických dát	45
6.2	Priebeh spracovania experimentov	46
6.2.1	Výber hodnôt pre použitie v modeli reakcií zrenice	46
6.2.2	Výpočet vnímanej svietivosti	46
6.3	Diskusia	48
7	Overenie metódy odhadovania kognitívnej záťaže pre heterogénne rozhrania založenej na modeli reakcií zrenice	49
7.1	Zoznam štatisticky odhaliteľných vlastností a stanovených hypotéz	49
7.2	Zoznam použitých datasetov	50
7.3	Zistenie ideálneho parametra sigma matice projekcie sietnice	51
7.4	Zistenie vnímanej svietivosti jednotlivých farieb	53
7.5	Overenie poklesu vnímanej svietivosti v okrajoch monitora	54
7.6	Overenie predpovedateľnosti šírky zrenice pre rozhrania webových sídel	55

7.7	Overenie detekcie kognitívnej záťaže	56
7.7.1	Kognitívne náročné a nenáročné úlohy	56
7.7.2	Overenie detekcie kognitívnej záťaže pri homogénnych stimuloch	57
7.7.3	Overenie detekcie kognitívnej záťaže pri práci s heterogénnymi stimulmi re- prezentujúcimi skutočné rozhrania webových stránok	59
7.7.4	Porovnanie presnosti explicitnej a implicitnej kalibrácie	61
7.8	Diskusia	62
8	Záver	65
8.1	Naplnenie stanovených cieľov	65
8.2	Prínos navrhutej metódy ako celku	68
8.3	Budúca práca	69
	Literatúra	71
	Príloha A - Obsah priloženého média	75
	Príloha B - Detailný opis použitých datasetov	77
	Dataset reakcií zrenice na rôzne svetelné podmienky - D1	77
	Dataset matematických príkladov - D2	78
	Dataset matematických príkladov a čítania textu - D3	80
	Dataset tvorivých úloh s inštrukciami - D4	82
	Dataset skúmania vlastností sietnice oka - D5	83
	Dataset komplexného overenia navrhutej metódy - D6	84
	Príloha C - Etapy riešenia projektu	87
	Prvý semester riešenia	87
	Druhý semester riešenia	88
	Tretí semester riešenia	89
	Príloha D - Navrhnutý softvérový produkt	91
	Vstupné dáta	91
	Technické detaily	92
	Inštalčná príručka	95
	Používateľská príručka	96
	Príloha E - Príspevok na konferenciu IIT.SRC 2018	107

1 Úvod

Elektronické obchody sa predbiehajú v jednoduchosti procesu výberu a nákupu produktov, predajcovia reklamnej plochy sa snažia udržať mieru návratov používateľov a tvorcovia výpočtových aplikácií zjednodušujú, automatizujú a zefektívňujú prácu zamestnancov v širokej palete profesií. Rastúci počet konverzií v elektronických obchodoch, uspokojený návštevník webu či väčšie množstvo vykonanej práce za nižšie náklady sú motivátory v komerčnej sfére, vďaka ktorej sa aj oblasť vyhodnocovania používateľského zážitku a použiteľnosti aplikácií teší veľkému záujmu. Komplikované formuláre, neintuitívne umiestnené prvky rozhrania či nesprávna dekompozícia úloh v navrhnutých postupnostiach krokov rovnako ako ďalšie problémy, s ktorými sa používatelia aplikácií denne stretávajú je lacnejšie identifikovať už vo fáze návrhu aplikácie ako po jej nasadení a strate zákazníkov.

Identifikovať väčšinu najvýznamnejších chýb dokážeme vďaka používateľským štúdiám aj na jednoduchých prototypoch s niekoľko málo používateľmi v počiatočnej fáze práce na vývoji nového softvéru. Rovnako môžeme vykonať rozsiahly experiment na existujúcej implementácii rozhrania a priniesť návrhy na vylepšujúce zmeny.

V závislosti od želaných výstupov sa delia aj prístupy k testovaniu. Technicky najjednoduchšie sú kvalitatívne experimenty, kde používateľ aplikácie priamo či retrospektívne konzultuje jeho pocity a požiadavky s administrátorom experimentu. S rastúcim počtom používateľov, rastie počet identifikovaných nedostatkov aj časová náročnosť spracovania experimentu. Často nás zaujíma úspešnosť používateľa pri prechode konkrétného prípadu použitia. Táto úspešnosť sa u rôznych ľudí môže líšiť a v tejto situácii sa skôr prikloníme ku kvantitatívnemu vyhodnocovaniu aplikácie.

Kognitívna záťaž je z pohľadu vyhodnocovania používateľského zážitku jedna z najdôležitejších metrík. V [Haapalainen et al., 2010] je definovaná ako netriviálna agregácia záťažových faktorov počas vykonávania úlohy. Jej miera reprezentuje vzťah medzi požiadavkami potrebnými na vykonanie úlohy a schopnosťami používateľa tieto požiadavky realizovať.

V literatúre sa často stretáme so zovšeobecnením pojmu kognitívnej záťaže, ktorú môžeme počítať z rôznych metrík, napríklad z miery úspešnosti splnenia alebo času vykonávania úloh. V tomto texte sa snažíme pozerať na kognitívnu záťaž ako na samostatnú metriku psychologického pôvodu. V konečnom dôsledku, v používateľskom testovaní nehľadáme kognitívnu záťaž ale problémy v použiteľnosti. Problémy vyvolávajú kognitívnu záťaž a jej identifikácia nám ich pomáha odhaliť. Ak teda v experimentoch zistíme, že niektorá úloha sa používateľom nepodarila splniť, je trochu nepresné a zbytočné hovoriť o identifikovanej kognitívnej záťaži.

Klasické metriky, ako je nájdnosť elementov, či miera úspešnosti splnenia zadanej úlohy, rovnako dobre identifikujú existujúce problémy v rozhraní, no kognitívna záťaž dokáže konkrétne úlohy ohodnotiť s vyššou informatívnou hodnotou. Napríklad v prípade dvoch úloh so sto percentnou mierou úspešnosti splnenia zadania, môže kognitívna záťaž určiť, ktorá z úloh bola triviálna a pri ktorej boli na používateľa kladené väčšie, aj keď stále zvládnuteľné nároky. Tento rozdiel by sme mohli určiť porovnaním času, za ktorý boli úlohy vykonané, stále však nevieme rozlíšiť, či tento časový rozdiel bol spôsobený dlhším rozmyšľaním alebo množstvom jednoduchých úkonov, ktoré

musel používateľ vykonať. Kognitívna záťaž v každom prípade reprezentuje skutočné zaťaženie používateľa a dokáže identifikovať problémy aj v rýchlo a správne zodpovedaných úlohách.

Skutočnú kognitívnu záťaž vieme merať viacerými senzormi. V práci [Ikehara a Crosby, 2005] nájdeme analýzu rôznych spôsobov, pričom nie všetky sú na používateľské testovanie rovnako vhodné. Medzi vlastnosťami, ktoré nám umožňujú identifikovať záťaž je pohyb očí, šírka zrenice, vodivosť kože, rozdiely v teplote, zmeny v pulze, množstvo kyslíku v krvi alebo tlak pôsobiaci na myš počítača. Práve prvé dve vlastnosti sú v používateľskom testovaní relatívne často využívané. Podľa citovanej publikácie tieto dve metriky dokážu obsiahnuť najviac relevantných aspektov pre používateľské testovanie ako je náročnosť, pozornosť a stres v prípade pohybu očí a únava, náročnosť, emócie či záujem v prípade šírky zrenice.

V našej práci sa úzko zameriavame na dilatáciu zrenice. Tento spôsob, rovnako ako sledovanie pohybu očí, môže byť neintruzívny a to až do takej miery, že v najjednoduchších prípadoch používateľ ani nemusí tušiť, že je jeho záťaž vyhodnocovaná. Táto výhoda pramení z možnosti bezkontaktného snímania šírky zrenice pomocou špeciálnej kamery a taktiež vďaka metódam nevyžadujúcim striktné laboratórne prostredie. Používateľské štúdie sa tak môžu stať masovejšími, a teda presnejšími a lacnejšími. Podobné závery nájdeme aj v závere štúdie [Laeng, Sirois a Gredebäck, 2012].

Ako už bolo spomenuté, dilatácia zrenice môže identifikovať mnoho kognitívnych vlastností. Okrem týchto mentálnych faktorov na ňu výrazne vplýva aj vnímaná svietivosť. Tieto dve vlastnosti robia z dilatácie zrenice zaujímavú no v istej miere nepredvídateľnú veličinu. Aby sme jej správanie dokázali čo najlepšie modelovať, bolo nutné okrem hlbokej analýzy problematiky dilatácie zrenice v odborných publikáciách vykonať aj sériu experimentov.

Experimenty ktoré sme v rámci našej práce vykonali nám umožnili lepšie pochopiť ako sa správa zrenica pri pohľade na rôzne oblasti monitora, ako veľmi na ňu vplýva zmena svietivosti alebo aký vplyv má kontrastné rozloženie elementov v rozhraní a mnohé ďalšie otázky. Vykonali sme dve samostatné série experimentov, participovali sme na experimentoch ďalších výskumov na pôde fakulty a vyčistili sme datasety získané v predchádzajúcej práci. V absolútnych číslach sme tak vyhodnocovali dáta z približne troch stoviek jedinečných pozorovaní. Pri plánovaní a vykonávaní experimentov sme úzko spolupracovali s členmi výskumného centra používateľského zážitku UXI ¹.

1.1 Cieľ práce

Napriek sľubným vlastnostiam sa testovanie kognitívnej záťaže ako súčasť používateľských štúdií využíva len okrajovo. Metódy na jej vyhodnocovanie sú pre tieto štúdie nedostatočné, zložité alebo často nepoužiteľné. Medzi najväčšie problémy radíme neprirodzené kalibračné činnosti alebo použiteľnosť len pre jednoduché rozhrania, ktoré sú príliš vzdialené skutočným aplikáciám. Naším zámerom bolo zjednodušiť a zároveň zefektívniť vykonávanie štúdií novou metódou vyhodnocovania kognitívnej záťaže a umožniť použitie kognitívnej záťaže ako bežnej súčasti štúdií používateľského zážitku. Z tohto dôvodu sme si stanovili nasledujúce ciele.

¹<https://www.pewe.sk/uxi/>

1. Zdokonalit model reakcií zrenice
 - (a) Identifikovať všetky relevantné faktory zmien vo vnímanej svietivosti a preukázať relevantnosť týchto faktorov na experimentálnych dátach
 - (b) Navrhnuť efektívne spôsoby kalibrácie všetkých relevantných faktorov s prihliadnutím na potreby experimentátorov a použiteľnosť týchto procedúr v reálnych štúdiách
2. Navrhnuť vhodný spôsob kvantifikácie kognitívnej záťaže pomocou modelu reakcií zrenice
3. Overiť navrhnuté časti komplexným experimentom s dôrazom na čo najväčšiu podobnosť scenáru experimentu skutočnému používateľskému testovaniu
 - (a) Preukázať schopnosť navrhnutej metódy identifikovať kognitívnu záťaž a klasifikovať úlohy do viacerých tried náročnosti
 - (b) Porovnať rozdiel v predikovanej kognitívnej záťaži pomocou metódy sledovania surovej šírky zrenice a našej metódy
4. Poskytnúť softvérovú platformu umožňujúcu využitie metódy pre identifikáciu kognitívnej záťaže v používateľských štúdiách

1.2 Štruktúra práce

Testovanie použiteľnosti nie je nová disciplína a okrem rôznych techník jej vykonávania sa kladie dôraz aj na korektné štatistické spracovanie objavených vlastností. Rýchlemu náhľadu do používateľského testovania venujeme priestor v kapitole 2. Dilatácia zrenice je jav spôsobovaný viacerými zdrojmi. Jej vzťah s kognitívnou záťažou ako aj vplyvy iných zdrojov analyzujeme v kapitole 3. Dilatácia zrenice ako sprostredkovateľ informácií o kognitívnej záťaži je v používateľských štúdiách často ponímaná veľmi intuitívne a použité procedúry nezriedka vychádzajú priamo z jednoduchých metód používaných v psychológii. Vyhodnocovanie kognitívnej záťaže pomocou dilatácie zrenice od pôvodných metód, cez aktuálne zámery v cudzích prácach až po metódu navrhnutú v našej predchádzajúcej práci opisujeme v kapitole 4.

V kapitole 5 prechádzame k podrobnejšiemu návrhu dôležitých súčastí predmetnej metódy. Zároveň prezentujeme formálnu definíciu, navrhnuté vylepšenia ako aj vhodný spôsob prezentácie výsledkov kognitívnej záťaže vrátane štatistického kvantitatívneho vyhodnotenia. Kapitola obsahuje aj nový alternatívny spôsob kalibrácie modelu reakcií zrenice spolu s vysvetlením jeho výhod a nevýhod oproti štandardnej kalibrácii. Navrhnuté metódy z predchádzajúcich kapitol sme realizovali vo webovej aplikácii, pričom sme identifikovali potrebu niekoľkých optimalizácií, ktoré prezentujeme v kapitole 6. Stanovené hypotézy, analyzované datasety a vyhodnotenie metódy opisujeme v kapitole 7. Zhodnotenie našej práce s identifikovanými prínosmi obsahuje posledná kapitola 8.

Obsah priloženého média sa nachádza v prílohe A. Rozsiahlejší opis datasetov použitých pre vyhodnotenie hypotéz je v prílohe B. Plán prác rozdelený na všetky tri semestre riešenia diplomovej práce obsahuje príloha C. Dokumentáciu k navrhnutému softvérovému produktu, vrátane technických detailov, inštaláčnej a používateľskej príručky uvádzame v prílohe D. Prvé výsledky tejto

práce sme prezentovali na študentskej vedeckej konferencii IIT.SRC, v ktorej sme získali ocenenie za najlepší príspevok v kategórii Inžinierske štúdium. Plné znenie prezentovaného článku sa nachádza v prílohe E.

2 Testovanie použiteľnosti rozhraní aplikácií

2.1 Vyhodnocovanie navrhnutých rozhraní

Pri vyhodnocovaní navrhnutých rozhraní aplikácií sa riadime viacerými prístupmi. Medzi najzákladnejšie činnosti pri vývoji patrí heuristická evaluácia, testovanie použiteľnosti, overenie dodržania štandardov či vykonanie tzv. kognitívnej prechádzky.

Podľa porovnania týchto techník v práci [Jeffries et al., 1991] je zrejmé, že podcenenie ktorejkoľvek z nich môže viesť k neskorému alebo neúspešnému odhaleniu problémov s použiteľnosťou. Výhodou inak ťažkopádnej kognitívnej prechádzky je možnosť jej vykonania priamo softvérovými vývojármi, ktorí môžu rovnako vykonať validáciu systému a jeho rozhrania pomocou všeobecných odporúčaní. Explicitným vykonaním týchto dvoch činností môžeme dostatočne skoro predísť hlavným a často sa opakujúcim problémom. Heuristická evaluácia a testovanie použiteľnosti dokážu identifikovať najviac problémov s vysokou prioritou, pričom výstupom heuristickej evaluácie je navyše odhalenie možných problémov s nízkou prioritou.

2.2 Testovanie použiteľnosti

Testovaniu použiteľnosti ako finančne najnáročnejšej technike [Jeffries et al., 1991] je venované veľké množstvo prác a v snahe získať najrelevantnejšie výsledky máme k dispozícii niekoľko spôsobov. Jedným z najdôležitejších aspektov je identifikácia kognitívnej záťaže [S. Chen, Epps a F. Chen, 2011]. Aj k identifikácii kognitívnej záťaže existuje viacero prístupov, z ktorých každý má svoje špecifiká a rôznu úspešnosť pre odhaľovanie problémov v používateľskom rozhraní. V práci [S. Chen, Epps a F. Chen, 2011] sú analyzované štyri rôzne techniky identifikácie kognitívnej záťaže, ktoré sú porovnávané s jednoduchou ale intruzívnou technikou subjektívneho hodnotenia.

2.2.1 Subjektívne hodnotenie náročnosti úlohy

Subjektívne hodnotenie sa vykonáva explicitným ohodnotením náročnosti každej úlohy každým účastníkom hodnotou z Likertovej škály [Matell a Jacoby, 1971]. Táto technika vyžaduje explicitnú aktivitu človeka, ktorú by za normálnych okolností pri práci s aplikáciou nevykonával. Taktiež je možné vykonať ohodnotenie retrospektívne, kde môže byť skreslenie spôsobené zabúdaním.

2.2.2 Sledovanie času vykonávania úlohy

Pri sledovaní času vykonávania úlohy sa predpokladá, že výskyt problému spomalí prácu účastníka, čo sa prejaví dlhším nameraným časom. Pri tejto technike je potrebné definovať, čo sa považuje za začiatok a koniec úlohy. Pri porovnávaní dvoch odlišných úloh je potrebné rozlišovať, či je dlhší čas vykonávania jednej z úloh spôsobený rozmyšľaním nad problémom alebo väčším množstvom jednoduchých činností.

2.2.3 Miera správnosti vykonaných úloh

Identifikácia problémov pomocou počtu úspešných vykonaní úlohy bola v citovanej práci počítaná pomocou skóre, kde 1 predstavuje úspešne a 0 neúspešne dokončenú úlohu. Táto metrika môže mať nízku rozlišnosť v prípade, že vykonateľnosť porovnávaných úloh nie je prítomnými problémami úplne obmedzená a tie majú skôr vplyv len na zvýšenie kognitívnej záťaže a zhoršeného dojmu z práce s aplikáciou.

2.2.4 Frekvencia žmurkania očí

Aproximácia kognitívnej záťaže pomocou počtu žmurknutí resp. frekvencie žmurkania je založená na okulometrických dátach. Táto technika nebola podľa [S. Chen, Epps a F. Chen, 2011] príliš rozšírená, no bol preukázaný vzťah medzi žmurkaním a zmenami v kognitívnej záťaži.

2.2.5 Sledovanie dilatácie zrenice

Zo záznamov šírky zrenice boli v práci [S. Chen, Epps a F. Chen, 2011] vypočítané priemerné hodnoty pre jednotlivé intervaly úloh. Pri tejto technike je zvlášť potrebné kontrolovať ďalšie vlastnosti zrenice ako napr. reakcie na zmeny vnímanej svetivosti. Práve kontrolovaniu svetelných podmienok sa budeme venovať v ďalších kapitolách.

2.2.6 Porovnanie metód pre vyhodnocovanie kognitívnej záťaže

Vyššie uvedené metódy boli v práci [S. Chen, Epps a F. Chen, 2011] podrobené testu schopnosti rozlišovať rôzne úrovne kognitívnej záťaže na dve až päť tried náročnosti testovaním dvadsiatich participantov. Experiment obsahoval matematické úlohy s piatimi triedami náročnosti. Výsledky spätnej klasifikácie úloh do tried náročnosti na základe získaných dát sú v tabuľke 1.

Tabuľka 1: Správnosť klasifikácie úloh do tried náročnosti. Každý riadok obsahuje počet tried na klasifikáciu.

	Subjektívne hodnotenie	Sledovanie času	Miera správnosti	Okulometrické dáta
5 tried	66%	39%	29%	39%
4 triedy	75%	50%	35%	54%
3 triedy	85%	69%	51%	69%
2 triedy	90%	82%	68%	84%

2.3 Štatistické prístupy v testovaní použiteľnosti aplikácií

Množstvo problémov v použiteľnosti aplikácii môže byť identifikovaných pomocou kvalitatívnych štúdií s použitím intuitívnych postupov pri vyhodnocovaní vykonaného experimentu. Pokiaľ však chceme vyjadriť použiteľnosť návrhu v číslach, či štatisticky porovnať dve verzie systému, musíme

vykonať kvantitatívny experiment so správnym štatistickým vyhodnotením. Veľmi užitočný zdroj na výber správnej štatistiky pre vyhodnocovanie použiteľnosti aplikácií je kniha praktickej štatistiky pre výskum používateľského zážitku, pričom v nasledujúcom texte vychádzame predovšetkým z kapitoly 3 [Sauro a Lewis, 2012b] a kapitoly 5 [Sauro a Lewis, 2012a].

V našej práci sa primárne sústreďíme na vyjadrenie relatívnej náročnosti jednotlivých úloh, ktorými simulujeme prácu človeka v reálnych webových aplikáciách. Druhou nemenej dôležitou štatistickou úlohou je porovnanie dvoch úloh, ktorými simulujeme dve rôzne verzie aplikácie a snažíme sa určiť, či je rozdiel medzi dvoma úlohami signifikantný.

Na štatistické vyjadrenie náročnosti úlohy je vhodné použiť t-intervaly istoty okolo priemernej hodnoty, resp. okolo mediánu v závislosti od tvaru distribúcie hodnôt v populácii.

Pre štatistickú signifikanciu rozdielu v použiteľnosti dvoch konkrétnych návrhov je odporúčaný párový t-test v prípade, ak každý participant vykonal obidve úlohy. Ak každý participant vykonal iba jednu z úloh, je potrebné vykonať nepárový t-test. V záujme lepšieho pochopenia výsledkov kvantitatívneho vyhodnotenia autori zároveň odporúčajú vypočítať aj interval istoty okolo rozdielu v zistenej náročnosti úlohy.

Pokiaľ by výsledky kognitívnej záťaže v populácii neboli normálne distribuované, je potrebné použiť neparametrickú verziu t-testu napríklad Wilcoxon test [Wilcoxon, 1945] pre párové experimenty alebo Mann-Whitney U test [McKnight a Najab, 2010] pre nepárové experimenty.

3 Zdroje dilatácie zrenice

Dilatácia zrenice sa ako predmet výskumu objavuje v publikáciách už 57 rokov [Hess a Polt, 1960], pričom medzi hlavné zdroje dilatácie radíme zmeny svietivosti, kognitívnu záťaž či prítomnosť emócií [Laeng, Sirois a Gredebäck, 2012]. Absolútne hodnoty odmeranej dilatácie bez ďalších informácií nemajú veľkú výpovednú hodnotu, nakoľko rozsah dilatácie alebo priemer zúžených či maximálne rozšírených zreníc sa u každého človeka líši. Tieto tvrdenia podporuje štúdia [Hollingsworth, Bowyer a Flynn, 2009] na vzorke viac ako tisícich snímok zreníc pri svetlých a tmavých podmienkach.

3.1 Kognitívna záťaž a dilatácia zrenice

Súčasná veda disponuje množstvom vedeckých prác dokazujúcich schopnosť merať kognitívnu záťaž pomocou dilatácie zrenice [S. Chen, Epps a F. Chen, 2011; S. Chen a Epps, 2013; Laeng, Sirois a Gredebäck, 2012; Pomplun a Sunkara, 2003; Wang et al., 2013]. Akceptované a efektívne metódy vyhovujú iba psycho-fyziologickým experimentom, pričom ich využiteľnosť vo vyhodnocovaní používateľského zážitku v reálnom prostredí webových aplikácií je výrazne obmedzená. Dôvodom je dynamické rozhranie s nestálou štruktúrou a svietivosťou jednotlivých obrazoviek. Existujúce metódy neošetrujú takto komplexné scenáre, takže je pre ich správne použitie potrebné rozložiť scenár experimentu na jednotlivé kroky. Nevýhodou rozloženia scenára je okrem jeho technicky náročnej prípravy aj prerušovanie prirodzenej kontinuálnej práce používateľa v používateľskej štúdii.

Meranie kognitívnej záťaže vykonávame kontinuálnym meraním šírky zrenice oka používateľa, u ktorej po nástupe kognitívnej záťaže dochádza k rozšíreniu. Vo všeobecnosti teda môžeme povedať, že ak dôjde k nárastu kognitívnej záťaže, dôjde v určitej miere aj ku rozšíreniu zrenice ako fyziologickej reakcii tela [Mathôt a Van der Stigchel, 2015]. Naopak, samotná dilatácia zrenice neindikuje zmenu intenzity kognitívnej záťaže, práve z dôvodu existencie ďalších zdrojov dilatácie. Týmito zdrojmi sú najmä: reakcie na zmeny svietivosti [Ellis, 1981], vplyv emócií [Partala a Surakka, 2003] alebo hippus [Bouma a Baghuis, 1971].

V práci [Beatty, 1982] nájdeme sumarizáciu výstupov mnohých vedeckých prác zameraných na dilatáciu zrenice spôsobenú kognitívnou záťažou. Autor prezentuje štyri hlavné kategórie úloh vyvolávajúcich kognitívnu záťaž. V obrázku (obr. 1) vidíme, že dilatácia zrenice bola najvýraznejšia pri logických úlohách, ďalej úlohách na pamäť a prácu s jazykom a najmenej pri úlohách na pozornosť. Toto poradie nedeklaruje, že niektorý z typov kognitívnych úloh vo všeobecnosti vyvoláva väčšiu kognitívnu záťaž či väčšiu dilatáciu zrenice. Dôležité je, že dilatácia zrenice sa objavuje pri kvalitatívne rôznych typoch úloh. Zaujímavá je najmä veľkosť dilatácie zrenice, ktorá môže dosiahnuť až 0,5 mm. Vo výstupe je zobrazená dôležitá vlastnosť - korelácia maximálnej dilatácie zrenice s objektívnou náročnosťou konkrétnej úlohy daného typu.



Obr. 1: Maximálna dilatácia zrenice spôsobená kognitívnou záťažou vyvolanou riešením úloh pre kvalitatívne odlišné triedy kognitívnych úloh. Zdroj: [Beatty, 1982]

3.2 Iné zdroje dilatácie zrenice

3.2.1 Vnímaná svietivosť

V našej práci sa zameriavame na odfiltrovanie najzávažnejšieho konkurenčného zdroja dilatácie - svietivosti stimulu. Reakcia zrenice na zmenu vnímanej svietivosti je známy jav, ktorý spôsobuje jej zúženie pri intenzívnejších svetelných podmienkach a naopak. Práca [Ellis, 1981] prezentuje model reakcií zrenice ako funkciu svetelných podmienok vyjadrenú vzťahom (1),

$$y = -0.12 + 0.66x - 0.04x^2 \quad (1)$$

kde y predstavuje šírku zrenice a x intenzitu svietivosti. Odozva zrenice na zmenu svietivosti bola kratšia ako 500ms a táto hodnota klesala s narastajúcou intenzitou svietivosti. Roztiahnutie zrenice bolo rýchlejšie ako stiahnutie, pričom rýchlosť opäť závisí od svietivosti stimulu. Najrýchlejšie bolo stiahnutie zrenice pri vysokej intenzite svietivosti s hodnotami okolo 8 mm za sekundu. Najpomalšie bolo roztiahnutie zrenice pri nízkej intenzite svietivosti s hodnotami okolo 0,5 mm za sekundu.

Okrem stimulu, ktorý je zobrazený na monitore počítača, vplývajú na vnímanú svietivosť svetelné podmienky okolia. V záujme dosiahnutia kvalitných výsledkov experimentov je preto dôležité udržiavať konštantné svetelné podmienky prostredia, čím minimalizujeme problematiku dynamickej svietivosti iba na stimul.

3.2.2 Emócie

Emócie vplývajú na šírku zrenice v závislosti od ich intenzity, pričom pozitívne aj negatívne emócie spôsobujú rozšírenie zrenice oproti stavu s neutrálnymi emóciami [Partala a Surakka, 2003]. V štúdiách ako [S. Chen a Epps, 2013] vidíme, že odlíšiť vplyv kognitívnej záťaže a emócií je možné predovšetkým iba pri umelých úlohách. Prezentovaná metóda nie je univerzálna, keďže využíva účelovo vybrané vlastnosti jednotlivých stimulov. Prítomnosť emócií však nebude pri väčšine kontrolovaných experimentov použiteľnosti bežných webových aplikácií výrazná, keďže mnohé aplikácie obsahujú väčšinou emočne neutrálne informácie.

3.2.3 Hippus

Hippus je jav oscilácie šírky zrenice, ktorý vzniká pri nečinnosti bez významnejšieho kognitívneho zaťaženia [Bouma a Baghuis, 1971]. Pokiaľ počas experimentu dôjde k úplnej nečinnosti particpanta, je možné, že hippus spôsobí falošnú identifikáciu kognitívnej záťaže. Úplná a kontinuálna nečinnosť pri aktivitách akou je vykonávanie úloh počas štúdie používateľského zážitku nenastane, preto okrem oddychových intervalov medzi jednotlivými úlohami nemusíme tento vplyv dilatácie očakávať.

3.3 Diskusia

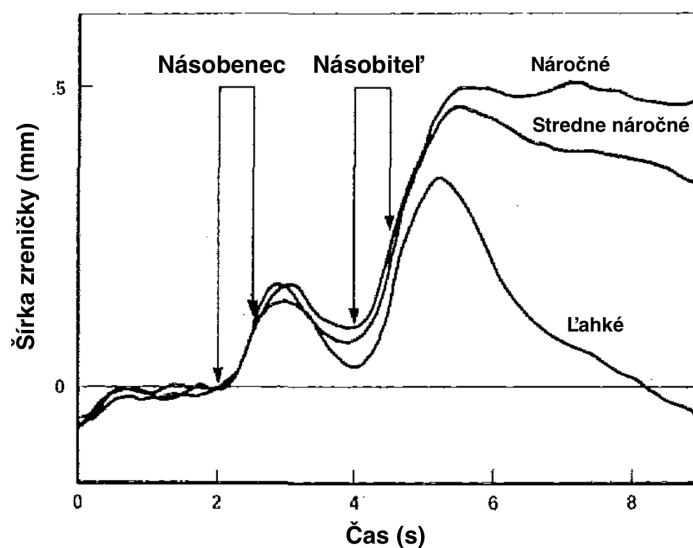
V prípade úspešného odfiltrovania nezanedbateľného príspevku zmeny svietivosti na dilatáciu zrenice, priradujeme zvyšok dilatácie iným fyziologickým faktorom, čiže kognitívnej záťaži, emóciám a nepravdepodobnému javu hippus. Exaktné určenie zdroja dilatácie bez ďalších sledujúcich zariadení môže byť náročné. Je preto dôležité prezentovať výstup metód zakladajúcich sa na dilatácii zrenice, ako všeobecný vplyv mozgovej aktivity a nie samotnej kognitívnej záťaže. Ďalej v našej práci sa budeme o identifikovanej mozgovej aktivite zjednodušene vyjadrovať ako o kognitívnej záťaži, ktorá je v kontexte prezentovaných aplikácií jednotlivých metód jedinou predpokladanou a záujmovou metrikou.

4 Prístupy k odhadu kognitívnej záťaže z dilatácie zrenice

Nutným predpokladom vyhodnotenia kognitívnej záťaže pomocou dilatácie zrenice je možnosť porovnať skutočnú šírku zrenice s jej alternatívou v bezzáťažovom stave za dodržania rovnakých svetelných podmienok. Určenie šírky zrenice v bezzáťažovom stave a transformácia dilatácie zrenice na hodnoty kognitívnej záťaže sú komplikované úlohy. V tejto kapitole analyzujeme základné predpoklady vyhodnocovania kognitívnej záťaže z dilatácie zrenice. V podkapitole (4.5) sa venujeme niekoľkým metódam, ktoré sledujú absolútnu šírku zrenice, pričom identifikujeme ich hlavné nevýhody. Analýza metódy navrhutej v našej predchádzajúcej práci sa nachádza v samostatnej podkapitole (4.6).

4.1 Priebeh dilatácie zrenice pri kognitívnej záťaži

Základným intuitívnym prístupom identifikácie kognitívnej záťaže je výber hodnoty referenčnej šírky zrenice z prvých okamihov zobrazeného rozhrania novej úlohy a sledovanie nasledujúcej dilatácie zrenice. Experimenty tohto prístupu boli sumárne prezentované ešte v roku 1982 v [Beatty, 1982]. Testované osoby dostali za úlohu vynásobiť dve čísla, pričom sa ukázalo, že po úplnom zadaní úlohy (definovaní násobiteľa), nasledovala dilatácia zrenice. Zároveň sa ukázalo, že táto dilatácia je interpretovateľná ako funkcia náročnosti problému. Táto situácia je zobrazená na obrázku (obr. 2).



Obr. 2: Priebeh absolútnej šírky zrenice pri násobení čísel s rôznou náročnosťou. Zdroj: [Beatty, 1982]

Podobný experiment nájdeme aj v [Klingner, Kumar a Hanrahan, 2008]. Tento experiment pracoval so súčasným, neintruzívnym, resp. bezkontaktným okulometrom a preukázal jeho schopnosť vyhodnocovať kognitívnu záťaž. Na rozdiel od náhlavného okulometra, jeho bezkontaktná alternatíva umožňuje merať kognitívnu záťaž používateľa aplikácie prirodzenou formou.

4.2 Kvantifikácia kognitívnej záťaže

Keďže intenzita dilatácie zrenice odzrkadľuje úroveň kognitívnej záťaže [Beatty, 1982], potrebujeme nájsť vzťah, ktorým dokážeme odhalené kognitívne zaťaženie vhodne kvantifikovať. V práci [Iqbal et al., 2005] autor navrhol vzťah (2) na výpočet kognitívnej záťaže normalizáciou rozdielu absolútnych hodnôt šírky zrenice.

$$\text{kognitivna zataz} = \frac{\text{sirka zrenicky} - \text{bazova sirka}}{\text{bazova sirka}} \quad (2)$$

Bázovou šírkou rozumieme šírku zrenice v pokoji za štandardných svetelných podmienok. Samotná práca sa v úvodnej kapitole odkazuje na zdroje, ktoré dokazujú spojitost medzi kognitívnou záťažou a dilatáciou zrenice práve pri kontrolovaných svetelných podmienkach. Otázkou je ako si vysvetliť pojem kontrolované svetelné podmienky, ako ich dosiahnuť a ako sa s tým vysporiadali iné práce. Napríklad v súčasnej práci [Toker, Lallé a Conati, 2017], ktorá detekuje kognitívne zmeny aj podľa tejto jednoduchšej metódy, nenájdeime nijaké riešenie problematiky kontrolovania svetelných podmienok zobrazeného rozhrania. Autorsky blízka práca [Lallé, Conati a Carenini, 2016] opisuje podmienky experimentu, v ktorom ošetrojú neželané reakcie na svietivosť pomocou nemenej intenzity svetla v miestnosti, premenlivé svetlo monitora počítača však ostáva naďalej nevyriešeným problémom. Klasické psychologické štúdie sa tomuto problému vyhýbajú použitím stimulov s konštantnými svetelnými podmienkami, čím sú zaužívané metódy vylúčené z možnosti ich použitia v testovaní skutočných webových aplikácií.

Výsledky práce [Hollingsworth, Bowyer a Flynn, 2009] poukazujú na veľké rozdiely absolútnych hodnôt dilatácie zreníc jednotlivcov. Každý človek má svoj dilatačný rozsah, s rôznymi minimálnymi aj maximálnymi hodnotami absolútnych širok zreníc. Pri viacerých sledovaných osobách narazíme na problém, ako určiť, ktorá z osôb zažila silnejšiu kognitívnu záťaž.

Použitím vzťahu (2) nedokážeme dostatočne presne normalizovať nameranú dilatáciu. Ak odmeriame rovnakú dilatáciu u dvoch rôznych ľudí, kde prvý človek bude mať bázovú šírku zrenice napríklad 2 mm a druhý 3 mm, bude väčšia kognitívna záťaž pripísaná prvému človeku s užšou bázovou zrenicou. U tohto človeka došlo skutočne k väčšej percentuálnej zmene šírky zrenice, vzťah však úplne ignoruje veľkosť dilatačného rozsahu. Na základe zistení v [Hollingsworth, Bowyer a Flynn, 2009] môžeme konštatovať, že dilatačný rozsah nemusí korelovať s bázovou šírkou zrenice, a tak je normalizácia vo vzťahu (2) naozaj len veľmi orientačná.

4.3 Referenčná šírka zrenice

V prípade, že sa v dôsledku zmien v rozhraní aplikácie zmení aj vnímaná svietivosť, bázová šírka zrenice nebude reprezentovať skutočnú šírku zrenice. Keďže sa šírka zrenice prispôbuje zmenám vo vnímanej svietivosti, nestačí nám jediná bázová hodnota. Aby sme mohli identifikovať vplyv kognitívnej záťaže na šírku zrenice, najskôr musíme odfiltrovať vplyv zmenenej svietivosti. Z tohto dôvodu potrebujeme poznať *referenčnú šírku zrenice*, ktorá je bázovou hodnotou pre každú úroveň

intenzity vnímanej svietivosti. V našej práci rozlišujeme pojmy bázová a referenčná šírka zrenice, pričom pri referenčnej šírke zrenice je už započítaný aj vplyv dynamicky zmenenej svietivosti.

Problém dynamickej zmeny svietivosti ilustrujeme na nasledujúcej modelovej situácii. Používateľovi zobrazíme rozhranie s úlohou a vyzveme ho, aby vyriešil zobrazenú úlohu. V prvom okamihu zobrazenia úlohy ešte nie je zrenica ovplyvnená kognitívnou záťažou a môže mať napríklad 2.6 mm a túto hodnotu použijeme ako referenčnú šírku zrenice. Náročnosť vykonávanej úlohy a následný nástup kognitívnej záťaže zapríčiní dilatáciu zrenice až na hodnotu napríklad 3.4 mm. Takto široká zrenica je oproti referenčnému stavu širšia o 0.8 mm. Tento postup zopakujeme pri ďalšej úlohe a zistíme, že referenčná šírka zrenice bola 3.1 mm a zrenica ovplyvnená kognitívnou záťažou bola široká 3.5 mm, čo znamená nárast o 0.4 mm. V našom hypotetickom príklade to znamená, že navzdory väčšej šírke zrenice pri druhej úlohe, určíme ako kognitívne náročnejšiu prvú úlohu, pretože mala väčší rozdiel v dilatácii.

Identifikovať referenčný interval, kedy má používateľ zobrazenú testovaciu úlohu a zároveň sa na nej ešte neprejavila kognitívna záťaž je v dynamicky sa meniacich rozhraniach náročné a v mnohých prípadoch aj neúspešné. Z tohto dôvodu sa vo výskumnej praxi používa metóda vloženia referenčného kalibračného stimulu do scenára pred každú úlohu. Takýto stimul sa líši od skutočného rozhrania úlohy iba v substituovaných znakoch, aby nebolo možné riešiť úlohu už počas zobrazenia kalibračného stimulu [Wang et al., 2013].

Najväčším problémom metódy vloženia stimulu je jej nepoužiteľnosť v reálnych používateľských štúdiách, kde vloženie kalibračného stimulu medzi každý z možných stavov rozhrania nie je z praktického hľadiska možné. V automatizovanom testovaní dynamických aplikácií pomocou metódy vloženia stimulu je preto nutné tento krok ignorovať a nájsť iný spôsob určenia referenčných hodnôt.

4.4 Nerovnomerná svietivosť stimulu a okolia

4.4.1 Homogénne rozhrania

Ideálne rozhranie úloh pre identifikáciu kognitívnej záťaže je jednofarebná obrazovka s krátkym - uprostred obrazovky umiestneným textom úlohy (pozri obr. 3). V tomto prípade je svietivosť obrazovky takmer nemenná a zrenica je počas celého experimentu vystavená konštantným svetelným podmienkam, takže akákoľvek dilatácia zrenice identifikuje iné fyziologické procesy, ako je napr. kognitívna záťaž. Na získanie referenčnej zrenice stačí odmerať šírku zrenice pred zobrazením prvej úlohy. Kognitívna záťaž koreluje s absolútnou šírkou zrenice v reálnom čase.

Ak sa homogénna farba pozadia a jej svietivosť mení pri jednotlivých úlohách, menia sa aj svetelné podmienky pôsobiace na zrenicu oka. Aby sme predišli identifikácii falošnej kognitívnej záťaže pri úlohách s tmavším pozadím, referenčnú šírku zrenice potrebujeme pre každú úroveň svietivosti. Kognitívna záťaž koreluje s upravenou šírkou zrenice podľa referenčnej šírky okrem prvých okamihov zmeny svietivosti rozhrania.

Výskum kognitívnej záťaže s využitím homogénnych obrazoviek je veľmi jednoduchý a používaný vo väčšine psychologických štúdií. V komerčných aplikáciách a veľkých internetových systémoch sa



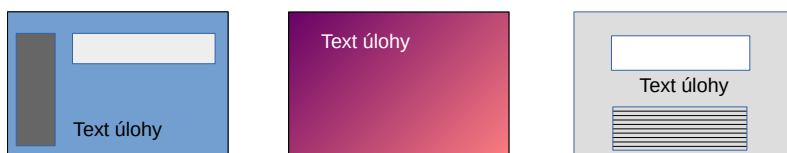
Obr. 3: Ukážka homogénnych rozhraní. Rozhranie má jednotnú podkladovú farbu a minimum textu potrebného na definíciu úlohy, ktorý je umiestnený v strede.

homogénne statické rozhrania prakticky nevyskytujú, takže pri používateľských štúdiách nemôžeme aplikovať existujúce konvenčné prístupy používané vo výskume.

4.4.2 Heterogénne rozhrania

Každé rozhranie obsahujúce navigačné panely, obrázky, tlačidlá a iné komponenty predstavuje heterogénne rozhranie (pozri obr. 4). Výpočet globálnej svietivosti stimulu nepostačuje, keďže vnímaná svietivosť stimulu sa mení s fixáciou pohľadu. Predpokladanú šírku zrenice potrebujeme pre každú oblasť zobrazeného rozhrania. Kognitívna záťaž koreluje so šírkou zrenice po odčítaní referenčnej šírky platnej pre aktuálne sledovanú oblasť.

Rozhranie webovej aplikácie je často dynamické. Jednotlivé komponenty menia svoje pozície či veľkosti, takže vnímané svetelné podmienky závisia okrem fixácie pohľadu aj od času. Referenčnú šírku zrenice potrebujeme pre aktuálne hodnoty vnímanej svietivosti sledovanej scény rozhrania.



Obr. 4: Ukážka heterogénnych rozhraní. Rozhranie má často farebne komplikované pozadie a elementy s rôznou svietivosťou na rôznych miestach.

4.4.3 Svietivosť okolia monitora

Bez ohľadu na to, či zobrazujeme homogénne alebo heterogénne rozhrania, vnímaná svietivosť závisí aj od svietivosti okolia monitora. V jednoduchých experimentoch sú informácie o úlohách umiestnené uprostred obrazovky a testované osoby sú ovplyvnené odlišnou svietivosťou okolia monitora minimálne. Naopak, v reálnych webových aplikáciách je obsah rozmiestnený ľubovoľne takže sa vplyvu odlišnej svietivosti prostredia nevyhneme. Svietivosť prostredia je v laboratórnych podmienkach konštantná, avšak môže byť odlišná od svietivosti zobrazeného rozhrania.

4.5 Existujúce metódy spracovania dilatácie zrenice pri práci s počítačom

V práci [Martinez-Gómez a Aizawa, 2014] je definovaných niekoľko možných prístupov k analýze kognitívnej záťaže pomocou okulometrických dát. Medzi vybranými metódami nájdeme tri, resp. štyri metódy, ktoré sú založené na šírke zrenice. Na analýze frekvencii zmien v šírke zrenice sa zakladá aj alternatívna skupina metód, ktorú opisujeme na záver tejto podkapitoly.

4.5.1 Metóda výpočtu rozdielu maximálnej a minimálnej šírky zrenice

Veľkosťou rozdielu najväčšej a najmenej šírky zrenice z množiny hodnôt zaznamenaných počas trvania jedinej úlohy definujeme jej pravdepodobnú kognitívnu náročnosť. Táto jednoduchá metóda skutočne rieši problém rôznej svietivosti jednotlivých úloh, pretože explicitne určuje referenčnú šírku zrenice, ktorá je relevantná pre danú úlohu. Ak predpokladáme, že vnímaná svietivosť sa počas trvania úlohy zásadne nemenila, nameraný rozdiel skutočne odzrkadľuje mentálnu aktivitu. V skutočnosti však metóda neurčuje kognitívne zaťaženie úlohy ale iba rozsah zmien v kognitívnej záťaži počas vykonávania úlohy. Teoreticky by mohla pre najťažšiu z úloh deklarovať nulovú kognitívnu záťaž, ak by sa participant počas trvania úlohy ani raz nedostal do stavu nízkeho kognitívneho zaťaženia.

Rovnako je dôležité vybrať vhodnú hodnotu, ktorú označíme ako najširšiu, pretože výsledky môžu byť veľmi ľahko znehodnotené náhodnou chybou či krátko trvajúcim dilatačným vzruchom na extrémne vysokú hodnotu. Nakoniec treba poznamenať, že táto metóda dokáže určiť len jedinu hodnotu určenej kognitívnej záťaže pre každú z úloh a nehovorí nič o priebehu záťaže počas úlohy.

4.5.2 Metóda štandardnej odchýlky šírky zrenice

Kognitívna záťaž je identifikovateľná aj pomocou štandardnej odchýlky hodnôt šírky zrenice počas trvania jednej úlohy. Vo svojej podstate je táto metóda vylepšením metódy výpočtu rozdielu maximálnej a minimálnej šírky zrenice, pretože je jej produkt rovnako závislý od veľkosti zaznamenaného dilatačného rozsahu. Navyše, táto metóda môže byť omnoho stabilnejšia pri výskyte chybných a extrémnych hodnôt. Zvyšné problémy však naďalej ostávajú - metóda očakáva nemenné podmienky vnímanej svietivosti a navyše je ešte viac náchylná na nedostatočný pokles kognitívnej záťaže pre stanovenie správnej referenčnej hodnoty. Z tohto dôvodu môže predikovať veľmi nízke kognitívne zaťaženie pri úlohách so stálou úrovňou kognitívnej náročnosti.

4.5.3 Metóda rýchlosti a akcelerácie dilatácie

Posledným pohľadom na možnosti analýzy kognitívnej záťaže v [Martinez-Gómez a Aizawa, 2014] je určenie kognitívnej záťaže pomocou zistenej rýchlosti, resp. akcelerácie dilatácie zrenice. Rýchlosť dilatácie zrenice môže byť vypočítaná ako prvá a akcelerácia ako druhá derivácia záznamu hodnôt širok zreníc počas trvania úlohy. Pozitívom tejto metódy je možnosť získať viac ako jednu hodnotu kognitívnej záťaže pre celú úlohu, avšak ani táto metóda nerieši vplyv zmenenej vnímanej

svietivosti. Zároveň, bez definície ďalších krokov, ako získané hodnoty transformovať na kognitívnu záťaž nie je táto metóda príliš použiteľná. Kognitívna záťaž totiž môže byť prítomná aj pri nulovej rýchlosti a teda aj akcelerácii dilatácie.

4.5.4 Metódy indexácie kognitívnej aktivity

Metóda sledovania rýchlosti a akcelerácie dilatácie je v zdokonalenom prevedení prezentovaná v práci [Marshall, 2002], kde autor kvantifikuje kognitívnu záťaž ako počet náhlych zmien v šírke zrenice. Odhad kognitívnej záťaže pomocou frekvenčnej analýzy nájdeme aj v práci [Gavas, Chatterjee a Sinha, 2017]. Tieto metódy sa už nevenujú pozorovaniu absolútnej šírky zrenice ako takej, ale hľadajú určité vzory v krivke dilatácie ako takej. Aj keď prvá spomínaná práca prezentuje nezávislosť od svetelných podmienok, jej nevýhodou môže byť výstupná metrika s obmedzenou granularitou výsledných hodnôt (identifikovaný počet relevantných zmien za sekundu).

4.6 Metóda modelovaním reakcií zrenice

V testovaní používateľského zážitku webových aplikácií pracujeme s komplikovaným premenlivým heterogénnym rozhraním. Už sme uviedli, že pri tomto type rozhrania potrebujeme na identifikáciu kognitívnej záťaže poznať referenčnú šírku zrenice závislú od momentálne zobrazenej a pohľadom sledovanej oblasti. V tejto situácii musíme dokázať referenčnú šírku zrenice predpovedať podľa lokácie fixácie. Takáto metóda identifikácie záťaže bola ešte v roku 2002 podľa [Marshall, 2002] len veľmi ťažko realizovateľná.

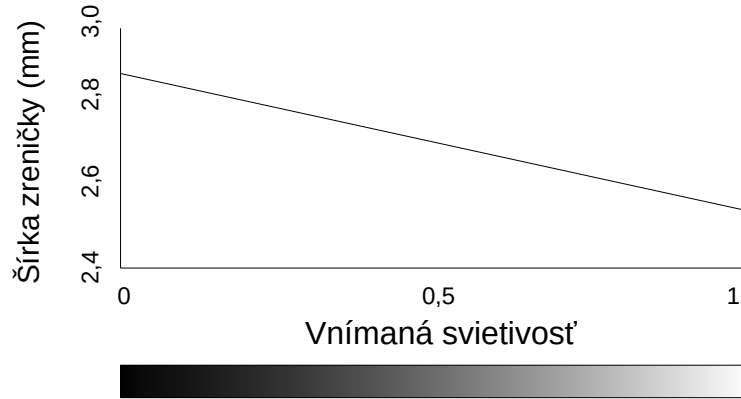
V rámci bakalárskej práce v roku 2016 [Juhaniak, 2016] sa nám podarilo vyvinúť model reakcií zrenice, ktorý nám umožňuje predpovedať referenčnú šírku zrenice na základe zobrazeného rozhrania a fixácie pohľadu v tomto rozhraní. Tento model a jeho kalibráciu sme prezentovali v rámci posterovej sekcie konferencie UMAP [Juhaniak et al., 2016]. Nezávisle môžeme nájsť podobné zámery aj vo výskumných zámeroch ďalších súčasných vedeckých štúdií ako napr. v [Wong, 2016]. Vo zvyšku tejto kapitoly opisujeme najdôležitejšie zistenia a návrhy z našej predchádzajúcej práce.

4.6.1 Definícia modelu reakcií zrenice

Navrhnutý model je zložený z krivky referenčných hodnôt širok zreníc pre jednotlivé úrovne svietivosti a matice projekcie sietnice, ktorá reprezentuje vplyv jednotlivých pixelov na vnímanú svietivosť v závislosti od vzdialenosti k bodu fixácie.

Krivka referenčných hodnôt (pozri obr. 5) je lineárna klesajúca funkcia na intervale $<0, 1>$ podľa vzťahu (3), kde rozsah dilatácie predstavuje rozdiel šírky zrenice nakalibrovaný pri čiernej a bielej obrazovke, svietivosť je hodnota vnímanej svietivosti na intervale $<0, 1>$ a bazová šírka je šírka zrenice nakalibrovaná pri sledovaní bielej obrazovky.

$$sirka\ zrenicky = rozsah\ dilatacie * (1 - svietivost) + bazova\ sirka \quad (3)$$



Obr. 5: Ilustrácia natrénovanej krivky referenčných hodnôt.

Matica projekcie sietnice je reprezentácia ohodnotení príspevkov jednotlivých pixelov k celkovej vnímanej svietivosti. Rozmery matice sú identické s rozlíšením sledovaného stimulu. Pre každý pixel stimulu existuje práve jedna hodnota vplyvu na celkovú vnímanú svietivosť a každá hodnota v matici prislúcha práve jednému pixelu stimulu. Hodnoty matice sú definované dvojrozmerným normálnym rozdelením so stredom v bode fixácie podľa rozloženia svetlocitlivých buniek na sietnici ľudského oka [Purves et al., 2004].

Pixelom v blízkosti bodu fixácie matica predpisuje najvyšší vplyv na celkovú vnímanú svietivosť a s narastajúcou vzdialenosťou od bodu fixácie je vplyv pixelov nižší. Prvky matice v tvare dvojrozmerného normálneho rozdelenia definuje vzťah (4),

$$m_{atica} = \frac{1}{2 * \pi * \sigma^2} * e^{-\frac{x^2 + y^2}{2 * \sigma^2}} \quad (4)$$

kde x a y je horizontálna a vertikálna vzdialenosť pixelu od bodu fixácie a σ je hodnota rozptylu normálneho rozdelenia. Pre maticu platí podmienka, že súčet všetkých hodnôt je rovný jednej, keď je bod fixácie uprostred stimulu (pozri obr. 6).



Obr. 6: Ilustrácia matice projekcie sietnice počas fixácie na stred stimulu.

Takto definovaný model neopisuje situácie ako vypočítať vnímanú svietivosť v momentoch s bodom fixácie mimo stred stimulu. Vyhodenie tohto problému bolo jedným z výsledkov tejto diplomovej práce.

Maticu projekcie sietnice, teda dvojrozmerné normálne rozdelenie, je možné v jazyku Python vytvoriť pomocou knižníc Numpy (np) a Scipy Stats (st). Nasledujúci zdrojový kód (zdrojový kód 1) vytvorí štvorcovú maticu so stranou dlhou 100 px a celkovým súčtom hodnôt rovným jednej. Parameter sigma, ktorého hodnotu získame vo fáze kalibrácie, určuje tvar normálneho rozdelenia výslednej matice. Parameter sigma ovplyvňuje najmä uhlopriečka monitora a jeho vzdialenosť od okraju stola.

Zdrojový kód 1: Ukážka výpočtu centralizovanej matice projekcie sietnice

```
1 matrix_len=100
2 linear_interval = (2 * sigma + 1.0) / matrix_len
3 linspace = np.linspace(-sigma - linear_interval / 2.0, sigma + linear_interval /
4     ↪ 2.0, matrix_len + 1)
5 matrix_1d = np.diff(st.norm.cdf(linspace))
6 matrix_raw = np.sqrt(np.outer(matrix_1d, matrix_1d))
7 matrix = matrix_raw/matrix_raw.sum()
```

4.6.2 Kalibrácia modelu

Krivku referenčných hodnôt kalibrujeme pre každého participanta v rámci používateľského testovania. Kalibráciu krivky referenčných hodnôt tvorí zobrazenie niekoľkých homogénnych obrazoviek, do prostriedku ktorých sa participant musí pozeráť po dobu ich zobrazenia. Interval prepínania obrazoviek trvá 7 sekúnd, čo je dostatočne dlhá doba na prispôsobenie sa zrenice zmeneným svetelným podmienkam [Ellis, 1981].

Poradie zobrazení obrazoviek pri kalibrácii krivky referenčných hodnôt:

- Zobrazenie pokynov pred kalibráciou
- Biela obrazovka (#FFFFFF)
- Šedá obrazovka (#AAAAAA)
- Čierna obrazovka (#000000)
- Modrá obrazovka (#0000FF)
- Šedá obrazovka (#AAAAAA)
- Biela obrazovka (#FFFFFF)

Dodatočné opakovania procedúry môžu zlepšiť kvalitu kalibrácie, avšak veľký počet iterácií môže spôsobiť nervozitu, navýšiť kognitívnu záťaž a znehodnotiť celú kalibráciu.

Priemerné hodnoty šírky zrenice počas zobrazenia jednotlivých kalibračných obrazoviek slúžia ako parametre lineárnej regresie, výsledok ktorej definuje krivku referenčných hodnôt. Napríklad, ak by sme nakalibrovali šírku zrenice pre biele obrazovky na hodnoty blízke napr. 2 mm, pre šedé 2,3 mm a čierne 3 mm, natrénovali by sme model o veľkosti rozsahu dilatácie 1 mm. Týmto modelom

by sme mohli predpovedať šírku zrenice pre teoreticky vypočítanú svietivosť na úrovni 50% podľa vzťahu (3) ako $1 * (1 - 0,5) + 2 = 2,5$ mm.

Parameter sigma, ktorý je potrebný pre vytvorenie matice projekcie sietnice, kalibrujeme pomocou sady stimulov s bielymi centralizovanými kruhmi na čiernom pozadí. Úlohou participanta je opäť sledovať stred obrazovky. Interval prepínania obrazoviek trvá 7 sekúnd.

Poradie zobrazení obrazoviek pri kalibrácii parametra sigma

- Zobrazenie pokynov pred kalibráciou
- Kruh (priemer 1016 px)
- Kruh (priemer 656 px)
- Kruh (priemer 376 px)
- Kruh (priemer 176 px)
- Kruh (priemer 56 px)
- Kruh (priemer 16 px)

Ak poznáme krivku referenčných hodnôt, môžeme odhadnúť svietivosť jednotlivých stimulov pre jednotlivé hodnoty parametra sigma a vykonať numerickú optimalizáciu. Pokiaľ je parameter sigma príliš malý, vplyv vzdialenejších pixelov je podhodnotený a vypočítaná svietivosť je vyššia ako skutočná. Pokiaľ je parameter sigma príliš veľký, dôjde k nadhodnoteniu vzdialených čiernych pixelov a vypočítaná svietivosť je nižšia ako skutočná. V bakalárskej práci sme natrénovali jediný parameter pre celú skupinu participantov. Dnes tento krok hodnotíme ako chybný, pretože to odporuje unikátnosti modelu každého participanta.

4.6.3 Normalizácia dilatácie s natrénovaným modelom

V predchádzajúcej práci sme navrhli riešenie problému nedokonalkej normalizácie referenčnou šírkou zrenice. S odvolaním sa na poznatky z [Pomplun a Sunkara, 2003], kde sa zistilo, že intenzita kognitívnej dilatácie nie je významne ovplyvnená svetelnými podmienkami a na základe takmer lineárneho vzťahu svietivosti k dilatácii zrenice v [Ellis, 1981] sme v predchádzajúcej práci navrhli novú, vylepšenú normalizáciu definovanú vzťahom (5).

$$kognitivna\ zataz = \frac{sirka\ zrenicky - referencna\ sirka}{stimulovana\ sirka - bazova\ sirka} \quad (5)$$

Parameter vzorca - *stimulovaná šírka*, je šírka zrenice pri riešení štandardizovanej úlohy. Štandardizovaná úloha zobrazená na bielom stimule umelo vyvolá čo najvyššiu kognitívnu záťaž testovanej osoby a táto kognitívna záťaž sa stane jednotkou miery kognitívnej záťaže. Jednoduchšie, ale nepresnejšie môže byť do stimulovanej zrenice dosadená šírka zrenice pri sledovaní čiernej obrazovky, pričom dosiahneme maximálne možné rozšírenie zrenice bez kognitívneho zaťaženia. V tomto prípade je jednotkou miery kognitívnej záťaže veľkosť rozsahu dilatácie.

4.6.4 Výpočet svietivosti pre použitie v modeli

Do výpočtu vnímanej svietivosti vstupujú dva parametre: zobrazený stimul a lokácia fixácie v obrazovke.

Postupnosť krokov pri výpočte vnímanej svietivosti:

1. Zisti rozlíšenie zobrazeného stimulu
2. Vytvor maticu projekcie sietnice podľa lokácie fixácie
3. Vypočítaj príspevok každého pixelu v zobrazenom stimule na vnímanú svietivosť
 - (a) Vyber pixel zo zobrazenej obrazovky
 - (b) Vypočítaj svietivosť farby pixelu
 - (c) Vynásob hodnotu svietivosti pixelu prislúchajúcou hodnotou z matice projekcie sietnice
4. Vráť hodnotu vnímanej svietivosti ako súčet príspevkov získaných v predchádzajúcom kroku

Na výpočet svietivosti farieb sa nepodarilo úspešne použiť žiadny zo štandardných existujúcich algoritmov. Prevaha modrej zložky RGB spektra v zobrazenej farbe spôsobovala rozdielne name-
rané hodnoty vnímanej svietivosti u rôznych participantov. Z tohto dôvodu sme v predchádzajúcej práci navrhli algoritmus (zdrojový kód 2), ktorý vyžaduje samostatnú kalibráciu zrenice na modrú farbu. V navrhnutom algoritme preto používame parameter modelu reakcií zrenice na diskrimináciu dominancie modrej zložky podľa vzťahu (6).

$$B_D = \frac{\text{sirka zrenice pri modrej farbe} - \text{bazova sirka}}{\text{rozsah dilatácie}} \quad (6)$$

Zlepšenie odhadu svietivosti farby pomocou navrhnutého algoritmu sme overili experimentálne [Juhaniak, 2016] a prezentovali v práci [Juhaniak et al., 2016].

Zdrojový kód 2: Výpočet svietivosti farby

```
1 #IN 1. CASE
2 if g>=b:
3     return r if r > g else g
4
5 #IN 2. CASE
6 if r > b:
7     return (r-b) + (b-g)*(1-B_D/2) + g
8
9 #IN 3. CASE
10 if r==b:
11     return (b-g)*(1-B_D/2) + g
12
13 #IN 4. CASE
14 if r>g:
15     return (b-r)*(1-B_D) + (r-g)*(1-B_D/2) + g
16
17 #IN 5. CASE
18 return (b-g)*(1-B_D) + g
```

V experimentoch predchádzajúcej práce sme u niekoľkých osôb zistili, že v najhoršom možnom prípade by chyba spôsobená použitím algoritmu MAX_RGB vypočítala vyššiu vnímanú svietivosť s chybou 20% ($B_D = 0.2$).

V zjednodušenej implementácii je možné použiť maximum zo všetkých zložiek rgb spektra konkrétnej farby ako svietivosť konkrétnej farby (algoritmus MAX_RGB). Použitím tohto zjednodušenia sa však zvyšuje šanca falošne identifikovanej kognitívnej záťaže pri rozhraniach s modrými odtieňmi.

4.7 Diskusia

Dostatok vedeckých prác zameraných na identifikáciu kognitívnej záťaže pomocou dilatácie zrenice nás ubezpečuje, že tento zaujímavý prístup má stabilné a overené základy. Vyhodnotenie kognitívnej záťaže pomocou absolútnej dilatácie zrenice bolo doposiaľ v štúdiách používateľského zážitku neuskutočniteľné. Existencia viacerých aktuálnych vedeckých prác nám však naznačuje, že táto oblasť môže v najbližšej dobe výrazne posilniť úroveň poznatkov a preklenúť do bežnej praxe výskumu používateľského zážitku.

Naša predchádzajúca práca tento cieľ značne priblížila. Napriek tomu, aj navrhnutý model reakcií zrenice vykazuje niekoľko problémov, z ktorých najväčším je nedokonalá presnosť v predpovedaní šírky zrenice. Ukázalo sa, že dolná hranica chyby predpovedanej šírky zrenice môže dosiahnuť až 5% celkového dilatačného rozsahu oka. Jedným zo zámerov tejto diplomovej práce je zistenie, či je táto chyba vo väčšej miere spôsobená nedostatkami v návrhu alebo už narazila na hranice možností vyhodnocovania kognitívnej záťaže pomocou dilatácie zrenice, ktorá sa nemusí správať dokonalo podľa vyslovene algoritmických pravidiel.

V predchádzajúcej práci sme vykonávali experimenty len na stimuloch s centralizovaným obsahom. Toto zjednodušenie bolo nutné, pretože univerzálna metóda výpočtu vnímanej svietivosti nebola dokončená, keďže chýbala definícia vytvorenia matice projekcie sietnice v prípadoch, keď je lokácia fixácie mimo stredu stimulu. Pôvodný návrh mal iba zjednodušenú podobu vhodnú len pre teoretické a základné experimentálne prístupy.

Nevýhodou pôvodného návrhu metódy výpočtu kognitívnej záťaže pomocou modelu reakcií zrenice je striktna definovaná kalibrácia s pevne definovanými stimulmi. V skutočných experimentoch je pritom veľmi dôležité zamerať pozornosť účastníkov na skúmané oblasti v experimente a minimalizovať potrebné kalibračné procedúry.

5 Metóda odhadovania kognitívnej záťaže pre heterogénne rozhrania založená na modeli reakcií zrenice

5.1 Kontext riešeného problému

Používateľská štúdia sa skladá z úloh, ktoré musí participant vykonať na testovanom systéme. Úlohy sú obvykle vopred definované a cieľom používateľskej štúdie je identifikovať problémy, ktoré by mohli pri testovanej verzii systému nastať. Jedna úloha v používateľskom testovaní môže predstavovať vykonanie jedného prípadu použitia, resp. ich kombinácie v nadväzujúcej postupnosti.

Typickým príkladom je testovanie novej verzie systému internetového bankovníctva, kde dostanú participantí radu úloh v určenom poradí, napr.:

- Prihlásenie do systému
- Zobrazenie zostatku na bežnom účte
- Zadanie trvalého príkazu
- Obnova PIN platobnej karty
- Vyhľadanie detailov o úverových možnostiach
- Odhlásenie zo systému

Každá z týchto úloh má jasne definované ohraničenie začiatku a úspešného splnenia. Cieľom používateľského testovania je v tomto prípade zistiť náročnosť vykonania jednotlivých úloh, ktorá môže byť identifikovaná priebehom kognitívnej záťaže detekovanej počas trvania úlohy. Pri riešení zadania sa participant môže dostať na ľubovoľnú podstránku, čo môže ovplyvniť svetelné podmienky vplývajúce na dilatáciu zrenice.

Pri testovaní funkčnosti novej metódy je potrebné použiť umelé úlohy so zámerne zavedenou náročnosťou do niektorých z nich, pričom úlohou systému na vyhodnocovanie kognitívnej záťaže je správne identifikovať úlohy, pri ktorých mali participantí najvyššiu kognitívnu záťaž. Týmto úlohami môže byť riešenie matematických príkladov, ktoré majú rovnako ako skutočné úlohy, pevne definované podmienky začiatku (prvé zobrazenie príkladu) a konca úlohy (zadanie odpovede) a zároveň sa vyskytujú na popredí heterogénnych a komplexných rozhraní.

5.2 Kroky riešenia problému

Spracovanie informácií z experimentu sa v návrhu metódy delí do niekoľkých logických celkov. Táto kapitola prezentuje vylepšenia jednotlivých súčastí pôvodného návrhu našej metódy, ktorá je opísaná v kapitole 4.6. V podkapitole 5.3 sa venujeme zdokonalenému trénovaniu krivky referenčných hodnôt, ktoré je dôležité pre správne predpovedanie šírky zrenice podľa vnímanej svietivosti. Vnímaná svietivosť sa podľa pôvodného návrhu metódy vypočítava pomocou matice projekcie sietnice, ktorej generický návrh prezentujeme v podkapitole 5.4. S vnímanou svietivosťou súvisí vnímaná

svietivosť prostredia, ktorá sa mení s každou zmenou fixácie. Návrh odhadu vnímanej svietivosti prostredia riešime v podkapitole 5.5. Explicitná kalibrácia môže v niektorých situáciách diskvalifikovať použitie navrhnutej metódy v skutočných používateľských štúdiách. Z tohto dôvodu sme navrhli alternatívnu verziu implicitnej kalibrácie, ktorú prezentujeme v podkapitole 5.6. Na konci procesu vyhodnocovania experimentu získame kontinuálne krivky predpokladanej a skutočne nameranej šírky zrenice pre jednotlivé úlohy. Pre účely efektívneho vyhodnotenia experimentu sme v kapitole 5.7 navrhli postup agregácie týchto surových hodnôt do zrozumiteľných ohodnotení jednotlivých úloh s korektným štatistickým vyhodnotením.

5.3 Zdokonalená explicitná kalibrácia

Najdôležitejšia súčasť modelu reakcií zrenice je krivka referenčných hodnôt. V pôvodnom návrhu sme túto krivku reprezentovali pomocou lineárnej funkcie vnímanej svietivosti. Po analýze dát v experimentoch sme dospeli k záveru, že pôvodné zjednodušenie modelovania krivky referenčných hodnôt pomocou lineárnej funkcie zavádza chybu do predpovedania šírky zrenice. Táto chyba sa prejavovala predpovedaním širšej zrenice pre určité úrovne vnímanej svietivosti, čo malo za následok odhadovanie záporných hodnôt kognitívnej záťaže.

Tento problém navrhujeme riešiť kvadratickým modelovaním krivky referenčných hodnôt (pozri obr. 7). Touto úpravou by zdanlivo nemalo dôjsť k zhoršeniu stavu keďže lineárna krivka sa dá vyjadriť aj pomocou kvadratickej definície. V skutočnosti sa však vystavujeme väčšiemu riziku pretrénovania modelu. Keďže v explicitnej kalibrácii zobrazujeme z časových dôvodov len malý počet kalibračných stimulov, jedinou zle natrénovanou hodnotou môžeme výrazne zmeniť tvar výslednej krivky. Tento problém sa v prípade lineárnej regresie nemusel prejaviť z dôvodu nižšej prispôbitelnosti lineárnej krivky prípadným odchýlkam. Z tohto dôvodu sme navrhli rozšírenie procedúry tréningu o autoregulačné kroky.



Obr. 7: Ilustrácia natrénovanej kvadratickej krivky referenčných hodnôt.

Autoregulačné kroky sú zmeny tréningových dát takým spôsobom, ktorý identifikuje neštandardné hodnoty a upraví ich do teoreticky akceptovateľného tvaru. Teoreticky neakceptovateľný tvar je prípad, ak sa v kalibračných dátach objaví širšia zrenica nameraná pri bielej obrazovke ako pri

šedej či tmavej obrazovke a naopak. Ďalší teoretický predpoklad hovorí o tom, že ak pri bielej obrazovke nameriame dve rôzne šírky zrenice, vyššiu prioritu v trénoch má užšia zrenica, pretože jej širšia alternatíva musela byť ovplyvnená kognitívnou záťažou alebo sa nestihla dostatočne zúžiť po dočasnom sledovaní iného tmavého objektu ako kalibračného stimulu. Analogicky, pri trénoch minimálnej vnímanej svietivosti preferujeme širšie hodnoty zrenice.

Kroky optimalizovanej kalibrácie:

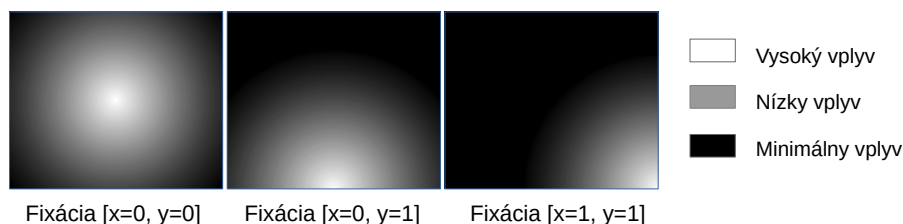
1. Natrénuj iniciálnu kvadratickú krivku referenčných hodnôt
2. Zmeň prioritu trénoch hodnotám podľa ich relevancie voči iniciálnej kvadratickej krivke
3. Natrénuj druhú verziu kvadratickej krivky referenčných hodnôt
4. Skontroluj, či natrénovaná krivka neobsahuje lokálny extrém
5. Ak natrénovaná krivka obsahuje lokálny extrém
 - (a) Ak obsahuje lokálne maximum - zisti rozdiel medzi lokálnym maximom a referenčnou šírkou zrenice pre minimálnu svietivosť a oprav tieto hodnoty o dve tretiny zisteného rozdielu, inak zisti rozdiel medzi referenčnou šírkou zrenice pre maximálnu svietivosť a lokálnym minimom a oprav tieto hodnoty o dve tretiny zisteného rozdielu
 - (b) Vyber tri jedinečné hodnoty do kalibrácie
 - i. Opravenú hodnotu lokálneho extrému
 - ii. Opravenú hodnotu opravovaného okraja krivky
 - iii. Pôvodnú hodnotu neopravovaného okraja predpovedaním tejto hodnoty pomocou druhej verzie krivky referenčných hodnôt
 - (c) Natrénuj tretiu verziu kvadratickej krivky referenčných hodnôt pomocou troch vybraných hodnôt
 - (d) Skontroluj, či natrénovaná krivka neobsahuje lokálny extrém
 - (e) Ak natrénovaná krivka obsahuje lokálny extrém
 - i. Natrénuj lineárnu krivku referenčných hodnôt pomocou troch vybraných hodnôt
 - ii. Spriemeruj tieto hodnoty s ich alternatívami v lineárnej krivke referenčných hodnôt a použi ich v trénoch štvrtej verzie krivky referenčných hodnôt
 - iii. Skontroluj, či natrénovaná krivka neobsahuje lokálny extrém
 - iv. Ak natrénovaná krivka obsahuje lokálny extrém, vráť lineárnu krivku referenčných hodnôt
6. Vráť poslednú kvadratickú verziu krivky referenčných hodnôt

5.4 Univerzálna matica projekcie sietnice

V predchádzajúcej práci sme navrhli maticu projekcie sietnice. Základnou úlohou tejto matice je určiť vplyv jednotlivých pixelov na vnímanú svietivosť v závislosti od ich vzdialenosti k bodu fixácie. Keďže sme si v predchádzajúcej práci vystačili s centralizovanými stimulmi, stačilo vytvoriť jej najjednoduchšiu - centralizovanú alternatívu. Vytvorenie centralizovanej matice ako aj kalibráciu parametra sigma sme vysvetlili v kapitole 4.6.1 a 4.6.2.

V predchádzajúcej práci sme definovali podmienku, že pokiaľ je matica orientovaná na stred stimulu, súčet všetkých hodnôt v matici musí byť rovný jednej. Táto podmienka zabezpečuje, že biely stimul

bude mať hodnotu vnímanej svietivosti rovnú jednej. Pokiaľ je bod fixácie totožný so stredom stimulu, súčet vzdialeností jednotlivých pixelov k bodu fixácie je minimálny a práve táto situácia reprezentuje najväčšiu možnú vnímanú svietivosť. S narastajúcou vzdialenosťou bodu fixácie od stredu stimulu narastá aj súčet vzdialeností jednotlivých pixelov k bodu fixácie, čoho dôsledkom je klesajúca celková vnímaná svietivosť pixelov stimulu (pozri obr. 8). V praxi to znamená, že vnímaná svietivosť bieleho stimulu je rovná jednej iba v jedinom bode uprostred obrazovky a minimálna v rohoch obrazovky.



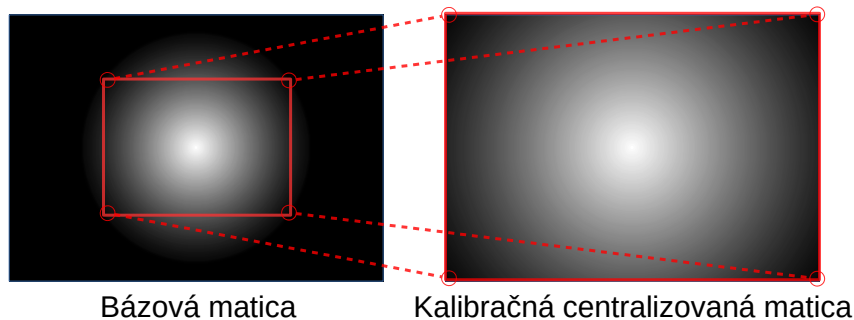
Obr. 8: Ilustrácia matice projekcie sietnice pre charakteristické body fixácie. V prípade fixácie v rohu obrazovky je vplyv pixelov v opačnom rohu obrazovky výrazne nižší ako v situácii s fixáciou uprostred obrazovky.

Vygenerovať necentralizovanú maticu projekcie sietnice nie je triviálna úloha. Necentralizovaná matica musí mať rovnaký tvar normálneho rozdelenia ako centralizovaná matica, pre ktorú dokážeme natrénovať optimálny parameter sigma. Stred tohto rozdelenia musí kopírovať bod fixácie. Navrhli sme vylepšenie algoritmu z kapitol 4.6.1 a 4.6.2, ktoré umožňuje vygenerovať maticu na základe dodatočného vstupného parametra - bodu fixácie.

Po zistení parametra sigma, ktorý určuje tvar normálneho rozdelenia matice projekcie sietnice vytvoríme takú maticu, aby sme z nej dokázali vyrezať finálnu maticu pre ľubovoľný bod fixácie. Z tohto dôvodu musí byť matica dvojnásobnej výšky a dvojnásobnej šírky stimulu. Túto maticu nazývame *bázová matica projekcie sietnice*.

5.4.1 Bázová matica projekcie sietnice

Bázová matica reprezentuje rozloženie svetlocitlivých buniek sietnice na ploche väčšej ako je veľkosť obrazovky. Oproti základnej centralizovanej matici je štvornásobne väčšia - 2 krát vyššia a 2 krát širšia (pozri obr. 9). Táto vlastnosť nám umožňuje vybrať v matici všetky také podmnožiny, ktoré reprezentujú vplyv pixelov na vnímanú svietivosť, pokiaľ je bod fixácie v rozmedzí stimulu. Nutnou podmienkou správnosti bázevej matice je súčet všetkých pixelov rovný jednej na takej podmnožine bázevej matice, že táto podmnožina je finálnou centralizovanou maticou projekcie sietnice. Finálna centralizovaná matica je využitá pri kalibrácii modelu podľa návrhu v kapitole 4.6.2. Súčet všetkých hodnôt v celej bázevej matici je väčší ako jedna, pretože matica je väčšia ako rozlíšenie stimulu, ktorého maximálna svietivosť je jedna.



Obr. 9: Vzťah medzi bázovou maticou a kalibračnou (centralizovanou) maticou. Bázová matica je rozšírením základnej centralizovanej matice projekcie sietnice.

V nasledujúcej ukážke zdrojového kódu (zdrojový kód 3) prezentujeme algoritmus vybudovania bázovej matice projekcie sietnice úpravou algoritmu z predchádzajúcej práce, ktorý opisujeme v kapitole 4.6.1. Modifikovaný algoritmus vybuduje maticu so stranami dvojnásobne väčšími ako má stimul. Matica sa normalizuje súčtom centralizovanej podmnožiny.

Zdrojový kód 3: Generovanie bázovej matice vnímania svetlice oka

```

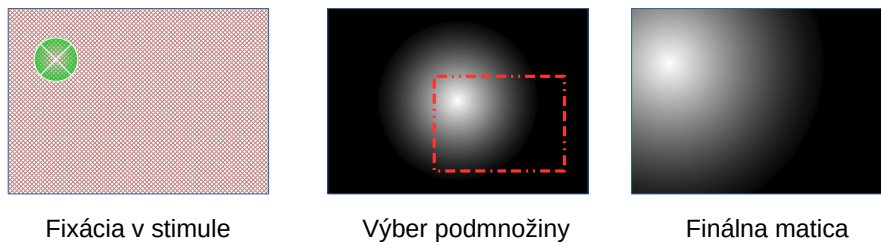
1 matrix_len=final_matrix_len * 2
2 sigma=3.2 * 2
3 linear_interval = (2 * sigma + 1.0) / matrix_len
4 linspace = np.linspace(-sigma - linear_interval / 2.0, sigma + linear_interval /
   ↳ 2.0, matrix_len + 1)
5 matrix_1d = np.diff(st.norm.cdf(linspace))
6 matrix_raw = np.sqrt(np.outer(matrix_1d, matrix_1d))
7 base_matrix = matrix_raw/matrix_raw[0.5*final_matrix_len::1.5*final_matrix_len
   ↳ ][0.5*final_matrix_len::1.5*final_matrix_len].sum()

```

5.4.2 Výrez finálnej matice z bázovej matice projekcie sietnice

Finálnu maticu projekcie sietnice pre konkrétny bod rozhrania generujeme vždy tak, aby sa stred normálneho rozdelenia váh matice nachádzal v bode fixácie. Keďže má bázová matica dvojnásobnú výšku a šírku stimulu, je zaručené, že ľubovoľný výrez s veľkosťou stimulu bude mať aspoň jeden bod dotýkajúci sa prvého kvadrantu v bázovej matici. Týmto bodom je ľavý horný roh výrezu a nazývame ho identifikátor matice, pretože jeho poloha v bázovej matici jednoznačne identifikuje finálnu maticu projekcie sietnice. Na určenie pozície tohto bodu používame vzťah (7), kde sú parametrami rozlíšenie stimulu a bod fixácie v stimule podľa dimenzii d . Ilustrácia tohto procesu je na (obr. 10). Generovanie matice vykonáva kód (zdrojový kód 4).

$$identifikator\ matice_d = rozlisenie_d - fixacia_d \quad (7)$$



Obr. 10: Ilustrácia vygenerovania finálnej matice projekcie sietnice pre necentralizovanú fixáciu.

Zdrojový kód 4: Generovanie finálnej matice projekcie sietnice oka so stredom normálneho rozdelenia v bode fixácie

```
1 matrix = base_matrix[final_matrix_len-fixation_x::2*final_matrix_len-fixation_x][
    ↪ final_matrix_len-fixation_y::2*final_matrix_len-fixation_y]
```

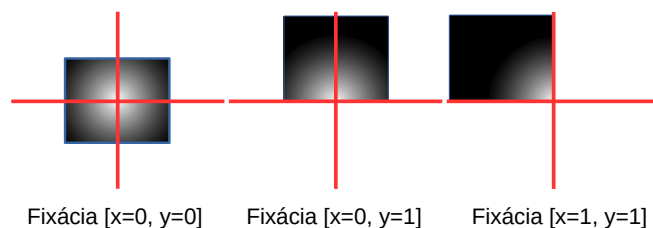
5.5 Vnímaná svietivosť okolia

Vnímaná svietivosť primárne závisí od svietivosti stimulu, no nezanedbateľný vplyv má aj svietivosť okolia monitora. S klesajúcou vzdialenosťou bodu fixácie od okraja monitora sa znižuje vplyv svietivosti stimulu podľa matice definovanej v podkapitole 5.4, no inverzne k tomuto poklesu narastá vplyv svietivosti okolia. Pokiaľ sa participant pozerá do jedného z rohov obrazovky, prostredie zaberá až tri kvadranty zorného poľa (obr. 11), čím sa stáva vplyv okolia na vnímanú svietivosť nezanedbateľnou skutočnosťou, ktorú je potrebné reflektovať vo výpočte vnímanej svietivosti.

Zatiaľ čo relatívnu svietivosť monitora je možné vypočítať na základe zobrazeného stimulu, svietivosť okolia monitora je určená vlastnosťami prostredia a na meranie jej hodnôt by bol pri experimentoch nutný ďalší hardvér, čo je pre reálne štúdie v skupinách ťažko realizovateľná požiadavka. Z tohto dôvodu musí byť svietivosť okolia monitora konštantná. Pokiaľ splníme túto podmienku, je možné svietivosť okolia natréňovať a započítať ju do výpočtu vnímanej svietivosti. Dodržanie rovnakej svietivosti môže čiastočne narušiť laboratórium s väčším množstvom počítačov využívaných počas experimentu. V tomto prípade sú zmeny vo vnímanej svietivosti minimálne, no pokiaľ je to možné, treba sa týmto situáciám vyhýbať. Dobrou praxou je vypnúť alebo stabilizovať obraz dataprojektora, ktorý má v uzavretej miestnosti po klasickom osvetlení najvyšší vplyv na svietivosť v prostredí.

5.5.1 Kalibrácia svietivosti okolia

Kalibráciu svietivosti okolia vykonávame sledovaním bielej obrazovky s fixáciou blízko ku okraju monitora. Vybrali sme 8 reprezentatívnych bodov, pričom 4 sú v rohoch monitora, kedy plocha monitora patrí len do jedného kvadrantu zorného poľa a 4 sú uprostred hrán medzi jednotlivými rohmi, kedy plocha monitora patrí do dvoch kvadrantov (pozri obr. 11).



Obr. 11: Ilustrácia poklesu vnímanej svietivosti monitora so zvyšujúcou sa vzdialenosťou fixácie od stredu stimulu. Pri centralizovanej fixácii zasahuje monitor do všetkých kvadrantov zorného poľa, pri fixácii uprostred jedného z okrajov sa monitor nachádza v dvoch a v prípade fixácie v rohu zasahuje monitor iba do jedného kvadrantu zorného poľa.

Postup kalibrácie svietivosti okolia:

1. Zobrazenie pokynov pred kalibráciou
2. Zobrazenie bielych stimulov s fixačným bodom
 - (a) Horná hrana uprostred
 - (b) Pravý horný roh
 - (c) Pravá hrana uprostred
 - (d) Pravý dolný roh
 - (e) Dolná hrana uprostred
 - (f) Ľavý dolný roh
 - (g) Ľavá hrana uprostred
 - (h) Ľavý horný roh
3. Výpočet vnímanej svietivosti okolia pre každú nameranú hodnotu
 - (a) Výpočet vnímanej svietivosti z monitora v konkrétnom bode fixácie ako súčet hodnôt v príslušnej matici projekcie sietnice
 - (b) Reverzné určenie celkovej vnímanej svietivosti pomocou krivky referenčných hodnôt podľa nameranej šírky zrenice
 - (c) Výpočet vnímanej svietivosti okolia rozdielom zistenej vnímanej svietivosti a vnímanej svietivosti z monitora
 - (d) Výpočet objektívnej svietivosti okolia normalizáciou vnímanej svietivosti okolia doplnkom vnímanej svietivosti monitora
4. Výber mediánu z vypočítaných hodnôt ako reprezentatívnu hodnotu svietivosti okolia

Výpočet vnímanej svietivosti okolia sa skladá z troch dôležitých krokov, počnúc určením vnímanej svietivosti z monitora. Pokiaľ máme správne natrénovaný parameter sigma, dokážeme určiť vplyv svietivosti monitora pomocou matice projekcie sietnice. Keďže sme použili biely stimul, každý pixel prispieva k vnímanej svietivosti s maximálnym využitím potenciálu definovanom v matici. V takomto prípade suma všetkých hodnôt v matici prislúchajúcej vybranému bodu fixácie pri okraji obrazovky predstavuje vnímanú svietivosť monitora.

Druhý krok je reverzné určenie celkovej vnímanej svietivosti podľa nameranej šírky zrenice. Tento krok vyžaduje predchádzajúce vykonanie kalibrácie krivky referenčných hodnôt opísanej v kapi-

tole 4.6.2. Pokiaľ by bola svietivosť okolia nulová, šírka zrenice by zodpovedala úrovni vnímanej svietivosti z monitora. Ak je nameraná šírka zrenice užšia ako zodpovedajúca hodnota svietivosti v krivke referenčných hodnôt, je to dôsledok zvýšenej vnímanej svietivosti zásluhou svietivosti okolia. Celkovú vnímanú svietivosť pri okraji monitora určíme tak, že zistíme ktorá úroveň vnímanej svietivosti predpokladá skutočne nameranú šírku zrenice.

Tretím krokom je výpočet vnímanej svietivosti okolia pomocou vzťahu (8).

$$vnimana\ svietivosť = vnimana\ svietivosť\ monitora + vnimana\ svietivosť\ okolia \quad (8)$$

Celkovú vnímanú svietivosť sme určili na základe odmeranej šírky zrenice a vnímanú svietivosť monitora sme vypočítali podľa bodu fixácie na bielom stimule, takže vnímaná svietivosť okolia je rozdiel celkovej vnímanej svietivosti a vnímanej svietivosti monitora.

Na záver je potrebné normalizovať vnímanú svietivosť okolia doplnkom ku vnímanej svietivosti monitora, čím získame všeobecný vplyv okolia monitora. Normalizácia je potrebná z dôvodu, že vplyv svietivosti okolia sme kalibrovali pre konkrétne body fixácie umiestnené v okrajoch monitora, takže zistená svietivosť je len relatívna k týmto bodom. Napríklad, ak by sme kalibrovali vplyv svietivosti okolia uprostred hrany okraju monitora a vplyv vnímanej svietivosti monitora by mal hodnotu 0,5, zistený vplyv okolia monitora by bol iba polovičný z hodnoty skutočnej svietivosti okolia monitora. V našom ilustračnom prípade by sme preto vydělili zistený vplyv okolia hodnotou 0,5 a získali skutočný vplyv okolia monitora na vnímanú svietivosť podľa vzťahu (9).

$$svietivosť\ okolia = vnimana\ svietivosť\ okolia / (1 - potencialna\ svietivosť\ monitora) \quad (9)$$

5.5.2 Celková vnímaná svietivosť

V podkapitole 5.4 sme definovali výpočet vnímanej svietivosti monitora pre ľubovoľný bod fixácie a v časti 5.5.1 sme vysvetlili postup ako kalibrovat svietivosť okolia monitora. Celkovú vnímanú svietivosť vypočítame podľa vzťahu (8). Prvý parameter, vnímanú svietivosť monitora, vypočítame podľa už známeho postupu využitia matice projekcie sietnice, pričom druhý parameter definujúci vnímanú svietivosť okolia, je rovnako závislý od bodu fixácie. Vnímanú svietivosť okolia získame pomocou vzťahu (10), kde svietivosť okolia predstavuje normalizovanú hodnotu získanú počas kalibrácie svietivosti okolia a potenciálna svietivosť monitora je suma všetkých hodnôt v prislúchajúcej matici vplyvu váh na vnímanú svietivosť.

$$vnimana\ svietivosť\ okolia = svietivosť\ okolia * (1 - potencialna\ svietivosť\ monitora) \quad (10)$$

Navrhnutý postup je možné použiť na výpočet vnímanej svietivosti pre ľubovoľný bod fixácie. Pokiaľ je bod fixácie uprostred stimulu, potenciálna svietivosť monitora je rovná jednej, takže vnímaná

svietivosť okolia je rovná nule. Vzťah ostáva konzistentný s tvrdením, že relatívna svietivosť okolia pri fixácii uprostred stimulu je nulová a zároveň platí, že s klesajúcou potenciálnou svietivosťou monitora narastá vnímaná svietivosť okolia.

5.5.3 Absolútny vplyv okolia na vnímanú svietivosť

Absolútny vplyv okolia monitora v skutočnosti nie je nikdy rovný nule, preto sa v kontexte modelu reakcií zrenice vždy vyjadrujeme len o svietivosti okolia relatívnej ku kalibrácii. Hodnota vnímanej svietivosti okolia uprostred monitora, nemôže byť nikdy menšia v iných bodoch fixácie, takže nepatrná a konštantná časť vplyvu okolia je reflektovaná kalibráciou.

Na obrázku (obr. 11) sme ilustrovali hraničné prípady, kedy má okolie väčší vplyv na vnímanú svietivosť ako samotný monitor. Pokiaľ sa človek pozerá do úplného rohu monitora, vplyv svietivosti okolia na vnímanú svietivosť je v absolútnych číslach viac ako 75%, pretože okolie zaberá tri celé kvadranty zorného poľa a časť kvadrantu, v ktorom je monitor umiestnený. Napriek tomu, ak sčítame všetky hodnoty v prislúchajúcej matici projekcie sietnice, zistená hodnota je väčšia ako 0,25 a dosadením tejto hodnoty do vzťahu (10) vypočítame vnímanú svietivosť okolia menšiu ako 0,75.

Zmienená vlastnosť je dôsledkom deklarácie modelu reakcií zrenice, ktorá hovorí o tom, že vnímaná svietivosť uprostred monitora je rovná jednej, pokiaľ monitor zobrazuje biely stimul. Časť absolútnej vnímanej svietivosti okolia, ktorú sme takto zaradili do svietivosti monitora sa prejavuje ako kladný rozdiel medzi kalkulovanou a absolútnou vnímanou svietivosťou monitora v rohu obrazovky.

V predchádzajúcej práci sme pre podmienky v laboratóriu našej fakulty natrénovali parameter sigma s hodnotou 3,2. Vygenerovaná matica projekcie sietnice pre túto hodnotu definuje vnímanú svietivosť monitora pri fixácii v jeho rohu na úrovni približne 0,27. Doplnok k tejto hodnote je 0,73 takže konštantná vnímaná svietivosť okolia je pre túto konkrétnu hodnotu parametra sigma viac ako 2%. Kladná dolná hranica konštantnej vnímanej svietivosti okolia je nutným dôsledkom správneho použitia modelu reakcií zrenice. Inými slovami, súčet hodnôt v matici definovanej pre fixáciu v rohu obrazovky nesmie byť za žiadnych okolností menší než 0,25 pre ľubovoľný natrénovaný parameter sigma.

Aj keď je vplyv okolia na vnímanú svietivosť čiastočne závislý od parametra sigma a teda je pre každého účastníka rozdielny, nikdy ho nemôžeme zanedbať, pretože by sme predpovedali falošnú kognitívnu záťaž v rohoch obrazovky na úrovni približne 70 a viac percent rozsahu dilatácie (po prvej fáze normalizácie). Teoreticky by sme svietivosť okolia mohli ignorovať pri úplne nule svietivosti okolia, avšak toto zjednodušenie by predstavovalo zbytočne veľkú komplikáciu pri vykonávaní experimentov. Navyše zavedenie ďalšieho kontrastného prvku (tmavého prostredia), by viedlo k zbytočným chybám.

5.6 Implicitná kalibrácia modelu reakcií zrenice

Model reakcií zrenice umožňuje neintruzívne vyhodnocovanie kognitívnej záťaže v používateľských štúdiách a nekladie nijaké podmienky na úpravu stimulov, resp. testovaných webových aplikácií, vďaka čomu je možné použiť metódu v pestrej škále používateľských testovaní. Ak by metóda nevyžadovala explicitnú kalibráciu, bolo by možné zaradiť metriku kognitívnej záťaže medzi ostatné automaticky získavané metriky používateľských štúdií. Okrem toho, odstránením explicitnej kalibrácie reakcií zrenice je umožnená zjednodušená príprava a vykonávanie experimentov využívajúcich okulometer a tiež zabránenie vzniku nežiadúceho efektu neprirodzeného správania používateľov v experimentálnom prostredí.

Model reakcií zrenice predpisuje opakované zobrazenie kalibračných stimulov s rôznou úrovňou svietivosti. Počas explicitnej kalibrácie predpokladáme minimálnu prítomnosť kognitívneho zaťaženia účastníka, a preto nameraná šírka zrenice pre jednotlivé hodnoty svietivosti predstavuje v prevažnej miere prispôbenie zrenice oka svetelným podmienkam. Kvalita tejto kalibrácie závisí od počtu vstupných hodnôt pri trénovaní referenčnej krivky pomocou kvadratickej regresie. Z praktických a časových dôvodov je počet natrénovaných hodnôt malý.

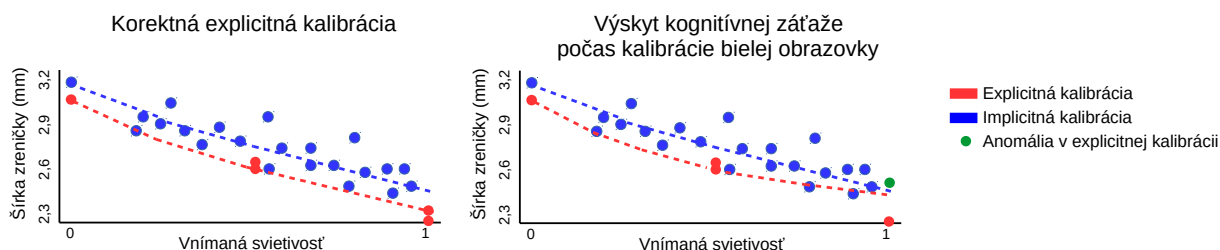
Hlavnou myšlienkou návrhu implicitne kalibrovaného modelu reakcií zrenice je zmena vstupných dát - použitie hodnôt z celého trvania experimentu. Nahradením stimulov síce porušíme základný predpoklad, že referenčná krivka zrenice má reprezentovať šírky zrenice pre jednotlivé úrovne svietivosti bez prítomnosti významnej kognitívnej záťaže, ako protihodnotu získame neporovnateľne väčšie množstvo vstupných dát do fázy kalibrácie. Explicitná kalibrácia by mala zabráť len veľmi málo času, zatiaľ čo samotný experiment môže trvať aj viac ako 20 minút. Pozorovaním zrenice počas tak dlhej doby úplne eliminujeme náhodné anomálie, ktoré sa môžu vyskytnúť aj počas explicitnej kalibrácie a významne tak ovplyvniť natrénovanú krivku referenčných hodnôt a následné vyhodnotenie (pozri obr. 12). Ďalšou výhodou je fakt, že pri kalibrácii kladieme dôraz na trénovanie tých hodnôt vnímanej svietivosti, ktoré sa v experimente nachádzajú najčastejšie. Väčšina webových aplikácií totiž dodržiava rovnakú schému rozloženia prvkov a farieb naprieč celým rozhraním.

V ideálnom prípade takto natrénovaná krivka referenčných hodnôt zrenice reprezentuje predpokladané šírky zrenice pre konkrétnu svietivosť a za vplyvu priemerného kognitívneho zaťaženia.

5.6.1 Postupnosť krokov pri implicitnej kalibrácii

Nahradením explicitnej kalibrácie za implicitnú sa myslí odlišný postup trénovania modelu reakcií zrenice pre dva zo štyroch základných parametrov modelu, a to krivka referenčných hodnôt a svietivosť okolia. Parameter sigma týmto spôsobom nie je možné kalibrovať a je potrebné použiť optimalizovanú hodnotu pre konkrétne laboratóriá získanú samostatnou kalibráciou. Pre diskriminátor modrej farby navrhujeme použiť hodnotu z experimentov s explicitnou kalibráciou alebo ak taká nie je dostupná, je možné použiť hodnotu 0,15, ktorá je najvhodnejšou aproximáciou získanou v predošlej práci na základe dát od 63 účastníkov.

Keďže implicitná kalibrácia nezávisí od vopred definovaných stimulov a vynútených fixácií, nie je



Obr. 12: Ilustrácia kvadratickej regresie pri kalibrácii kriviek referenčných hodnôt pomocou explicitnej a implicitnej kalibrácie. Krivky referenčných hodnôt sú v ideálnom prípade identické s konštantným posunom o hodnotu dilatácie zrenice pri priemernej kognitívnej záťaži (obrázok vľavo). Ak dôjde k náhlemu výskytu kognitívnej záťaže počas explicitnej kalibrácie (zelený bod vpravo), krivka referenčných hodnôt sa pre nedostatok ďalších validných bodov nemusí správne natrénovať.

možné využiť matematické predpoklady použité pri klasickom tréningu modelu reakcií zrenice. V explicitnej kalibrácii trénujeme krivku referenčných hodnôt s využitím predpokladov vďaka použitiu špeciálnych stimulov, bez nutnosti poznať svietivosť okolia. Svietivosť okolia následne trénujeme až v druhom kroku s využitím už natrénovanej krivky referenčných hodnôt.

Pretože sa v implicitnej kalibrácii nenachádzajú centralizované fixácie na biele stimuly, navrhli sme dva odlišné spôsoby ako natrénovať model reakcií zrenice. Prvý spôsob aproximuje explicitnú kalibráciu, pričom akceptuje možné riziko väčšej chybovosti, druhý spôsob navrhuje natrénovať konštantnú svietivosť okolia vopred na malej vzorke účastníkov pomocou explicitnej kalibrácie.

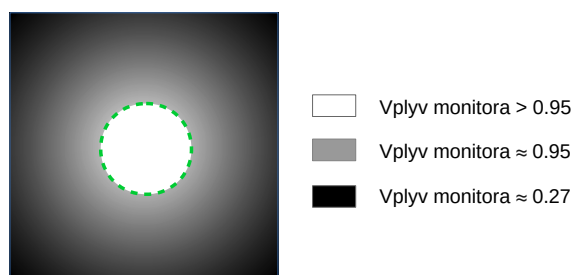
5.6.2 Kalibrácia bez známej svietivosti okolia

Prvou úlohou pri aproximácii explicitnej kalibrácie je získanie krivky referenčných hodnôt zrenice. Krivku referenčných hodnôt zrenice trénujeme pomocou kvadratickej regresie, kde vstupujú dvojice hodnôt, vnímaná svietivosť a prislúchajúca šírka zrenice. Pri aproximácii explicitnej kalibrácie nepoznáme skutočnú vnímanú svietivosť, pretože ak kalibrujeme na základe hodnôt z necentralizovaných fixácií, musíme uvažovať aj vplyv svietivosti okolia.

Pokiaľ vypočítame vnímanú svietivosť bez ohľadu na bod fixácie a svietivosť okolia iba odhadneme, môže to mať vážne následky na kvalitu kalibrácie. V kapitole 5.5 sme ukázali, že pri sledovaní niektorých oblastí monitora môže byť vplyv okolia na celkovú vnímanú svietivosť väčší ako vplyv monitora. Zlý odhad svietivosti spôsobí nesprávne výpočty celkovej vnímanej svietivosti, ktoré vnášajú do sústavy matematických vzťahov nekonzistenciu. Z tohto dôvodu musíme z kalibrácie vylúčiť tie hodnoty, ktoré boli významne ovplyvnené svietivosťou okolia.

Čím je pomer vnímanej svietivosti z monitora k vnímanej svietivosti okolia vyšší, tým menšia je chyba v určení celkovej vnímanej svietivosti. Ak si stanovíme hornú hranicu chyby, ktorú budeme v prípade nevhodne odhadnutej svietivosti okolia akceptovať, môžeme presne definovať podmnožinu fixačných bodov vhodných na kalibráciu krivky referenčných hodnôt. Táto množina zahŕňa všetky body fixácie v prostrednej časti monitora, pre ktoré platí, že suma hodnôt v prislúchajúcej matici

projekcie sietnice je väčšia ako doplnok akceptovateľnej chyby. Napríklad, ak si stanovíme maximálnu chybu na úrovni 5%, akceptujeme tie body fixácie, pre ktoré je maximálna možná svietivosť monitora väčšia alebo rovná 0,95 (pozri obr. 13).



Obr. 13: Podmnožina fixačných bodov vhodných na kalibráciu pri stanovenej maximálnej chybe na úrovni 5% vnímanej svietivosti. Vďaka pomerne veľkému sklonu normálneho rozdelenia hodnôt v matici je bezpečná kalibračná plocha dostatočne veľká s priemerom až 34% dĺžky hrany monitora, takže vzniká vysoká pravdepodobnosť, že kalibrácia krivky referenčných hodnôt bude kvalitná.

Kalibráciu krivky referenčných hodnôt vykonáme nad všetkými hodnotami vypočítanej svietivosti pre tie momenty, kedy bola fixácia zrenice vnútri vyhradenej plochy. Výpočet vnímanej svietivosti vykonávame klasickým postupom výpočtu vplyvu svietivosti monitora a svietivosť monitora môžeme vďaka jeho malému vplyvu odhadnúť empiricky, napr. 0,8.

V druhom kroku vykonávame kalibráciu skutočnej svietivosti okolia. Postup tejto kalibrácie je rovnaký ako pri explicitnej kalibrácii okolia, kde si vystačíme so vzťahom (9). Jediným rozdielom sú vstupné kalibračné hodnoty, ktoré nepochádzajú z pevne definovaných bodov. Odporúčame použiť také hodnoty, kedy fixácia pohľadu bola v blízkosti okrajov monitora. Tieto body môžeme definovať plochou, pre ktorú platí, že vplyv svietivosti monitora pri fixácii na bod z tejto plochy je menší ako 0,6. Vo všeobecnosti neodporúčame túto plochu zväčšovať, pretože by sme vo vzťahu (9) dostali príliš malý deliteľ, ktorý znásobuje efekt prípadných anomálií.

Implicitná kalibrácia modelu reakcií zrenice bez známej svietivosti okolia:

1. Kalibrácia krivky referenčných hodnôt
 - (a) Zvoľ predpokladanú hodnotu svietivosti okolia
 - (b) Definuj maximálnu euklidovskú vzdialenosť od stredu monitora, pre ktorú je fixácia kalibračne validná
 - (c) Iteruj cez všetky trojice (zobrazený stimul, fixácia, šírka zrenice), ktorých euklidovská vzdialenosť fixácie od stredu stimulu je menšia ako maximálna hranica
 - i. Vypočítaj vnímanú svietivosť stimulu
 - ii. Vypočítaj celkovú vnímanú svietivosť
 - iii. Zaraď vypočítanú svietivosť a príslušnú šírku zrenice do množiny pre kvadratickú regresiu
 - (d) Zisti parametre krivky referenčných hodnôt pomocou kvadratickej regresie nad pripravenou množinou hodnôt
2. Kalibrácia skutočnej svietivosti okolia
 - (a) Definuj minimálnu euklidovskú vzdialenosť od stredu monitora, pre ktorú je fixácia kalibračne validná
 - (b) Iteruj cez všetky trojice (zobrazený stimul, fixácia, šírka zrenice) v priebehu experimentu, ktorých euklidovská vzdialenosť fixácie od stredu stimulu je väčšia ako minimálna hranica
 - i. Vypočítaj vnímanú svietivosť stimulu
 - ii. Urči celkovú vnímanú svietivosť podľa šírky zrenice a krivky referenčných hodnôt
 - iii. Vypočítaj normalizovanú svietivosť okolia a zaraď ju do množiny zistených svietivostí okolia
 - (c) Vyber medián z usporiadanej množiny zistených svietivostí okolia ako natrénovanú svietivosť okolia

5.6.3 Kalibrácia so známou svietivosťou okolia

Svietivosť okolia je v štandardných podmienkach konštantná, pričom v laboratórnych podmienkach je pomerne jednoduché dosiahnuť konštantnú svietivosť okolia aj naprieč experimentami s rôznym časovým odstupom. Zatiaľ čo krivka referenčných hodnôt je rôzna pre rôznych účastníkov experimentu a môže byť odlišná aj medzi rôznymi experimentami, svietivosť okolia závisí len od zdrojov svetla a odrazových vlastností materiálov v miestnosti.

Spojením explicitnej a implicitnej kalibrácie môžeme vyťažiť z výhod oboch prístupov. Implicitná kalibrácia so známou svietivosťou okolia navrhuje nakalibrovať vnímanú svietivosť okolia na malej vzorke účastníkov pomocou presnejšej explicitnej kalibrácie. Ak sa explicitná kalibrácia urobí správne a zistená hodnota svietivosti okolia bude štatisticky signifikantná, v ďalších experimentoch s rovnakými podmienkami prostredia už tréning neprinesie zmenu hodnôt. S vopred známou svietivosťou okolia v nasledujúcich experimentoch je predmetom kalibrácie len krivka referenčných hodnôt.

Kalibrácia krivky referenčných hodnôt so známou svietivosťou okolia plne využíva potenciál implicitnej kalibrácie, pretože do vstupnej množiny pre kvadratickú regresiu je možné použiť všetky

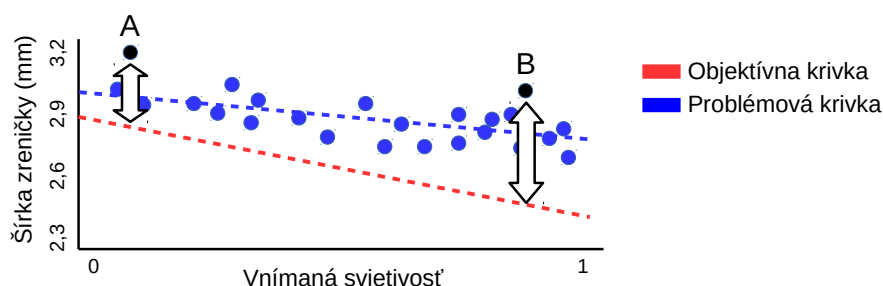
hodnoty z celého experimentu. Okrem nutnosti zistiť svietivosť okolia monitora samostatným sedením, je tento prístup presnejší a technicky jednoduchší ako plne automatická alternatíva.

Implicitná kalibrácia modelu reakcií zrenice so známou svietivosťou okolia:

1. Zisti štatisticky signifikantnú svietivosť okolia v samostatnej explicitnej kalibrácii
2. Iteruj cez všetky trojice (zobrazený stimul, fixácia, šírka zrenice), ktorých fixácia je na ploche monitora
 - (a) Vypočítaj vnímanú svietivosť stimulu
 - (b) Vypočítaj celkovú vnímanú svietivosť
 - (c) Zaraď vypočítanú svietivosť a príslušnú šírku zrenice do množiny pre kvadratickú regresiu
3. Zisti parametre krivky referenčných hodnôt pomocou kvadratickej regresie nad pripravenou množinou hodnôt

5.6.4 Ohraničenia implicitnej kalibrácie krivky referenčných hodnôt

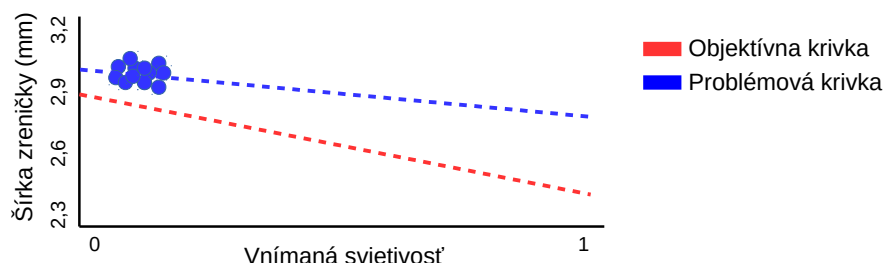
Implicitná kalibrácia nemusí správne fungovať za každých podmienok. Úrovně svietivosti kalibrujeme s možnou prítomnosťou kognitívnej záťaže, pričom predpokladáme, že rozloženie kognitívne náročných úloh v experimente nijako nesúvisí so svietivosťou rozhrania, na ktorom sú zobrazené. Pokiaľ by kognitívna náročnosť úlohy narastala/klesala s úrovňou vnímanej svietivosti, kalibrácia by podcenila/preceniла vplyv svietivosti. Ilustrácia na (obr. 14) zobrazuje implicitnú kalibráciu, kde skutočná kognitívna náročnosť úloh koreluje so stúpajúcou vnímanou svietivosťou, takže kognitívne náročnejšia úloha na bielom pozadí je považovaná za rovnako náročnú ako je tomu v prípade úlohy na tmavom pozadí.



Obr. 14: Ilustrácia nekorelujúceho posunu implicitne kalibrovanej krivky referenčných hodnôt voči objektívnej alternatíve v prípade, že kognitívna náročnosť úloh narastala so svietivosťou stimulov na ktorých bola zobrazená. Úlohy (A) aj (B) majú rovnakú veľkosť posunu v dilatácii oproti natrénovanej krivke referenčných hodnôt, avšak ich objektívna náročnosť ilustrovaná vertikálnou šípkou nie je rovnaká.

Implicitná kalibrácia môže zlyhať aj v takom prípade, že celé rozhranie obsahuje väčšinu úloh s jednou úrovňou svietivosti a existuje len jedna alebo veľmi malé množstvo úloh s inou úrovňou svietivosti. V tomto prípade sa môže krivka referenčných hodnôt pretrénovať pre úroveň svietivosti s dominantným zastúpením úloh a nereflektovať skutočný sklon krivky referenčných hodnôt zrenice, ako je to znázornené na obrázku (15). Ak aj tento problém nastane, krivka referenčných hodnôt

bude predpovedať nesprávne hodnoty pre úrovně svietivosti, ktoré sa v experimente prakticky nevyskytujú, pričom predpoveď pre dominantnú vzorku svietivosti ostane presná.



Obr. 15: Ilustrácia situácie, keď krivka referenčných hodnôt nekoreluje s objektívnou krivkou v prípade, ak nie sú v experimente prítomné stimuly významne odlišnej svietivosti.

5.7 Návrh kvantitatívneho štatistického spracovania

Navrhnutá metóda v každom časovom okamihu odhaduje úroveň vnímanej svietivosti sieťnicou oka a predpovedá očakávanú šírku zrenice. Následne pre každý časový okamih určí kognitívne zaťaženie ako rozdiel skutočnej a predpovedanej šírky zrenice a tieto namerané hodnoty agreguje do reprezentatívnej hodnoty pre každú úlohu. Keďže sa snažíme podporiť aj kvantitatívnu analýzu používateľských testovaní, agregované hodnoty jednotlivých úloh od viacerých používateľov agregujeme štatistickými metódami, čím získame objektívne ohodnotenie náročnosti testovaných úloh.

5.7.1 Atomická jednotka kognitívnej záťaže

Pôvodná metóda výpočtu záťaže modelovaním reakcií zrenice určuje len algoritmus získania aktuálnej prítomnej kognitívnej záťaže pre každú časovú jednotku vyprodukovanú okulometrom. V našej práci primárne využívame 60 Hz okulometre Tobii X2-60. Takýto okulometer je schopný zaznamenávať aktuálnu šírku zrenice šesťdesiat krát za sekundu. Keďže aj tie najkratšie štandardné úlohy, pre ktoré je možné vyhodnocovať spôsobenú kognitívnu záťaž majú minimálne niekoľko sekúnd, v každej úlohe získame stovky či tisíce záznamov vyhodnotení kognitívnej záťaže. Tieto záznamy tvoria užitočnú krivku priebehu kognitívnej záťaže, no z pohľadu automatického vyhodnotenia experimentu je taktiež vhodné vypočítať jedinou reprezentatívnu hodnotu záťaže pre každú úlohu.

Aby sme získali jedinou hodnotu vyhodnotenej kognitívnej záťaže pre každú z úloh, uvažovali sme nad rôznymi možnosťami najvhodnejšej transformácie množiny na jedinou hodnotu. Výber maxima môže kognitívnu záťaž skresliť, keďže je náchylný na šum v meraní šírky zrenice. Jediný nesprávne identifikovaný záznam zrenice, resp. krátkodobé s úlohou nesúvisiace zvýšenie dilatácie by spôsobilo znehodnotenie výsledku. Naopak výber mediánu môže mať za následok podhodnotenie kognitívnej záťaže v prípadoch, keď sa silná dilatácia spôsobená záťažou prejaví až v druhej polovici trvania úlohy. Aritmetický priemer môže byť v tomto prípade najvhodnejším riešením. Vďaka vysokej

frekvencii získaných záznamov ho ojedinele vysoké hodnoty príliš neskreslia a zároveň vychádza z každej hodnoty v množine. Výber aritmetického priemeru tiež nájdeme v ďalších aj technicky prísnejších okulometrických štúdiách ako napr. [Hayes a Petrov, 2016].

Použitie aritmetického priemeru nemusí byť vhodné ak je úloha príliš dlhá a nevieme ju rozdeliť na menšie podúlohy. Po konzultácii tohto problému s psychológom zaoberajúcim sa vyhodnocovaním kognitívnej záťaže sme dospeli k záveru, že namiesto generickej metódy určenia jedinej hodnoty pre rozsiahlu úlohu je vhodnejšie poskytnúť experimentátorovi krivku priebehu kognitívnej záťaže. Túto krivku je potom možné analyzovať manuálne alebo pomocou špecifického skriptu podľa požiadaviek kognitívneho psychológa.

5.7.2 Dvojfázová normalizácia kognitívnej záťaže

V pôvodnom návrhu hovoríme o normalizácii hodnoty rozdielu medzi skutočnou šírkou zrenice a jej referenčnou hodnotou pomocou dilatačného rozsahu pri zmenách vnímanej svietivosti alebo pomocou dilatačného rozsahu pri zmenách kognitívnej záťaže. Prvý spôsob môže byť problémom, pretože dilatačný rozsah pri zmenách vnímanej svietivosti môže byť u rôznych participantov rôzny. Naopak, normalizácia pomocou dilatačného rozsahu podľa kognitívnej záťaže spôsobenej pri maximálnom možnom úsilí sa zdá teoreticky ideálna, no experimenty ukázali, že dosiahnuť maximálne úsilie na špeciálnej kalibračnej úlohe môže byť veľmi problematické.

Na vyriešenie tohto problému sme navrhli dvojfázovú normalizáciu kognitívnej záťaže. Prvá fáza normalizácie kognitívnej záťaže je vykonaná vydelením rozdielu skutočnej a referenčnej hodnoty pomocou dilatačného rozsahu pri zmenách svietivosti. Krivky kognitívnej záťaže po prvej fáze normalizácie môžu byť použité najmä na porovnanie dvoch nezávislých experimentov jedného participanta.

Druhá fáza je normalizácia hodnôt pomocou obmedzeného dilatačného rozsahu pri zmenách kognitívnej záťaže v konkrétnom experimente. Tento obmedzený rozsah vychádza z usporiadanej množiny všetkých hodnôt zistenej kognitívnej záťaže po prvej fáze normalizácie. Rozsah je definovaný ako rozdiel 0,95 a 0,05 kvantilov.

Dvojfázovú normalizáciu vyjadrujú vzťahy (11) a (12),

$$A_t = \frac{Z_t - R_t}{Z_{cierna} - Z_{biela}} \quad (11)$$

$$B_t = \frac{A_t}{A_{k0,95} - A_{k0,05}} \quad (12)$$

kde Z je množina skutočných a R je množina referenčných hodnôt širok zrenice a index t je čas záznamu. Z_t je nameraná šírka zrenice v čase t , R_t je referenčná šírka zrenice v čase t . Z_{cierna} je skutočná šírka zrenice pri zobrazení čiernej obrazovky a Z_{biela} je skutočná šírka zrenice pri zobrazení bielej obrazovky. A je množina atomických hodnôt relatívnej kognitívnej záťaže po prvej fáze normalizácie a B je množina atomických hodnôt relatívnej kognitívnej záťaže po druhej fáze

normalizácie. A_t predstavuje hodnotu množiny A v čase t , $A_{k0,95}$ je 0,95 kvantil množiny A a $A_{k0,05}$ je 0,05 kvantil množiny A . B_t je hodnota množiny B v čase t .

Použitím druhej fázy normalizácie je zaručené, že 90% hodnôt bude ohodnotených číslom z intervalu $<0, 1>$ a zároveň ohodnotenie jednotlivých úloh bude dostatočne rozlíšiteľné. Pokiaľ by sme použili úplný rozsah, rozlíšnosť väčšiny úloh by bola omnoho menšia z dôvodu možného výskytu malého počtu extrémne vysokých alebo extrémne nízkych hodnôt zistenej kognitívnej záťaže. Ďalšia výhoda druhej fázy normalizácie je garancia ohodnotenia najťažších úloh (rozumej úloh okolo 0,95 kvantilu) pre jednotlivých participantov hodnotami blízkymi číslu 1, naopak najľahšie úlohy (0,05 kvantil) budú ohodnotené číslami blízkymi 0. Táto vlastnosť odstraňuje problém vysokej štandardnej odchýlky medzi participantami s rôznym dilatčným rozsahom. Treba však poznamenať, že druhá fáza normalizácie je vhodná len pre úlohy, v ktorých hľadáme rozdiely v náročnosti úloh jediného experimentu, pretože určuje kognitívnu záťaž relatívnu k úlohám ktoré sa v experimente vyskytli.

5.7.3 Kvantitatívne vyhodnotenie používateľských štúdií

Po vyhodnotení a normalizovaní atomických hodnôt kognitívnej záťaže jednotlivých participantov pre každú vykonanú úlohu, sa opäť dostávame do situácie, kedy máme pre jednu úlohu viac hodnôt kognitívnej záťaže. Výber maxima zamietame z dôvodu vysokej pravdepodobnosti výskytu extrémne silnej kognitívnej záťaže v malej časti datasetu. Táto vlastnosť je často prítomná aj pri nevhodne zvolenej referenčnej úlohe, napríklad, ak sa príliš podhodnotila kognitívna záťaž referenčnej úlohy aspoň pri jednom účastníkovi.

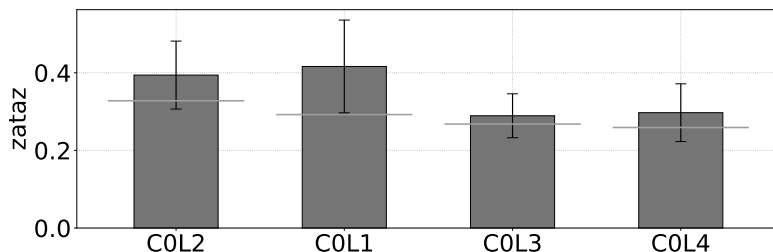
Po druhej fáze normalizácie a pri dodržaní základných podmienok na štatistické testovanie ² predpokladáme, že zistená kognitívna záťaž má naprieč datasetom používateľov normálne rozdelenie ³. Na základe odporúčania z [Sauro a Lewis, 2012b] je v tomto prípade možné použiť vhodnú alternatívu t-testu. Na základe toho sme navrhli ako metriku kvantitatívneho vyhodnotenia používať 95% t-interval istoty. Pohľad na rozdiely v intervaloch istôt medzi jednotlivými úlohami ponúka rýchly spôsob identifikácie problémových úloh. Pre štatisticky exaktnejšie porovnanie citovaná publikácia odporúča párový, resp. nepárový t-test, podľa vlastností datasetu. V prípade, že tvar výslednej kognitívnej záťaže nezodpovedá normálnemu rozdeleniu, navrhujeme použiť Wilcoxon test pre párové a Mann-Whitney U test pre nepárové datasety. Analýzu štatistických prístupov sme opísali v kapitole 2.3.

Ako vhodnú metódu získania finálnej a reprezentatívnej hodnoty kognitívnej náročnosti úlohy sme navrhli použitie mediánu. Medián reprezentovaný horizontálnou čiarou v grafe (obr. 16) v podstate ignoruje extrémne hodnoty na okrajoch normálneho rozdelenia, ktoré môže byť v niektorých

²Výber participantov do experimentu by mal reprezentovať populáciu, ktorá bude systém používať. Treba klásť dôraz na nezávislosť úloh v experimente, prípadne postupnosť úloh definovať náhodne pre každého participanta. Odporúča sa väčší počet participantov (aspoň 20).

³Ak toto nie je splnené, použijeme alternatívne metriky neparametrických verzií t-testov. Finálny výber vhodnej metriky závisí od testu normality. Poskytnúť generické rozhodovanie na základe testovania normality rôznych scenárov siahá nad rámec tejto práce. Táto práca poskytuje všetky alternatívne metriky, z ktorých výber správnej je závislý od podmienok experimentu.

prípadoch rozťahnuté určitým smerom. Vhodné je aj určenie priemernej hodnoty s vypočítaným intervalom istoty. Pre ilustráciu rozdielu kognitívnej záťaže medzi dvomi skupinami úloh pomocou intervalov istoty okolo priemeru nie je podľa [Sauro a Lewis, 2012b] potrebné vykonať normalizáciu do štandardného normálneho rozdelenia.



Obr. 16: Ilustrácia zistenej kognitívnej záťaže pre štyri hypotetické skupiny úloh. Priemer reprezentuje výška stĺpca so znázorneným 95%-intervalom istoty. Horizontálna čiara reprezentuje hodnotu mediánu. Skupina C0L1 má vyšší priemer ako skupina C0L2, avšak C0L1 má aj väčší interval istoty a menšiu hodnotu mediánu. Tento stav je dôsledkom prítomnosti niekoľkých extrémne veľkých hodnôt v určenej kognitívnej záťaži niekoľkých participantov v skupine úloh C0L1. Medián nie je ovplyvnený extrémnymi hodnotami niekoľkých participantov a môže poskytnúť realistickejší pohľad na určenie objektívnej náročnosti úlohy.

5.8 Diskusia

V tejto práci sme navrhli robustnú metódu na vyhodnocovanie kognitívnej záťaže schopnú pracovať s heterogénnymi rozhraniami, ktorá vychádza z našej predchádzajúcej práce prezentovanej v [Juhaniak, 2016].

5.8.1 Prínosy

Pôvodnú metódu sme zdokonalili použitím kvadratického tvaru krivky referenčných hodnôt a návrhom autoregulačných krokov pre jej presnejšiu kalibráciu. Navrhli sme generický výpočet matice projekcie sietnice pre ľubovoľný bod fixácie, vďaka čomu je možné odhadovať vnímanú svietivosť pre ľubovoľné oblasti záujmu v rozhraní aplikácie. Zároveň sme navrhli problém výpočtu vnímanej a kalibrácie celkovej svietivosti okolia.

Aby bolo možné našu metódu použiť aj bez narušenia scenáru experimentu explicitnými kalibračnými úlohami, navrhli sme spôsob implicitnej kalibrácie s dvomi variantami. Prvý návrh umožňuje vykonať kalibráciu na ľubovoľnom datasete bez predošlej znalosti svietivosti okolia. Tento spôsob odporúčame použiť, pokiaľ vyhodnocujeme dáta z experimentov, ktoré neobsahujú požadované kalibračné stimuly. Druhý, presnejší prístup vyžaduje znalosť svietivosti okolia monitora vopred. Tento spôsob odporúčame použiť, pokiaľ dokážeme zabezpečiť trvale nemenné podmienky svietivosti okolia počas všetkých vykonávaných štúdií. Ak to situácia dovoľuje, odporúčame používať explicitnú kalibráciu s viacerými iteráciami kalibračnej procedúry.

Taktiež sme navrhli spôsob agregácie získaných výsledkov pre automatizované a štatisticky správne kvantitatívne vyhodnotenie experimentu. Agregácia výsledkov je možná len pre označovaný dataset. Metóda umožňuje aj manuálny náhľad na surové dáta pre vyhodnotenie expertnou analýzou. Surovými dátami rozumieme návrh modifikácie záznamu šírky zrenice na záznamy kognitívnej záťaže v dvoch fázach normalizácie.

Navrhnutá metóda je vnútorne konzistentná a experimentálne overiteľná po častiach aj v celku. Kroky navrhutej metódy považujeme za dostatočne samostatné, takže je možné navrhnutú metódu ďalej rozširovať alebo upravovať jej jednotlivé súčasti.

5.8.2 Ohraničenia

Úspešnosť metódy je závislá od správnej kalibrácie modelu reakcií zrenice. Ak nie je explicitná kalibrácia úspešná alebo možná, je potrebné použiť alternatívny návrh implicitnej kalibrácie, ktorá môže skresliť výsledky dvoma spôsobmi.

Prvý potenciálny problém je korelácia úrovne kognitívnej záťaže so zmenou svietivosti rozhrania. Tento problém sa môže vyskytnúť ak sa špecifický typ prípadu použitia (napr. vyplnenie formulára) vyskytuje na rozhraniach odlišnej svietivosti ako zvyšok prípadov použitia systému.

Druhý problém, ktorým je kalibrácia len na obmedzenom intervale možnej svietivosti sa môže vyskytnúť v systémoch s približne rovnakou úrovňou svietivosti všetkých podstránok. V tomto prípade sa správne identifikuje kognitívna záťaž u väčšiny prípadov použitia, no je možné, že úlohy s výnimočnou úrovňou svietivosti budú ohodnotené nesprávne.

Ďalším problémom môžu byť zmeny svietivosti okolia. V prípade významného zníženia svietivosti okolia po fáze kalibrácie môže dôjsť k falošnej identifikácii kognitívnej záťaže.

6 Návrh efektívneho vyhodnocovania kognitívnej záťaže v experimentoch

Nástroj na vyhodnocovanie kognitívnej záťaže pomocou modelu reakcií zrenice pracuje s veľkými objemami dát. Aby bolo možné tieto dáta spracovávať časovo a pamäťovo efektívne, každá implementácia nástroja musí obsahovať radu optimalizácií, ktoré umožnia efektívne spracovanie veľkého množstva okulometrických a grafických dát v súlade s definovaným použitím modelu reakcií zrenice. V tejto kapitole uvádzame najdôležitejšie procesné a optimalizačné vlastnosti použité v našej implementácii metódy založenej na modeli reakcií zrenice ako všeobecný základ každej potenciálnej implementácie. V kapitole sa viackrát odkazujeme na konkrétny použitý modul pre spracovanie videa, ktorý používame na iteráciu jednotlivých záznamov obrazovky, efektívne zmeny rozlíšenia a porovnávanie vzájomnej podobnosti po sebe nasledujúcich záznamov. Tento modul je možné nahradiť ľubovoľnou efektívnou alternatívou poskytujúcou obdobnú funkcionálnu.

6.1 Predspracovanie okulometrických dát

Okulometre zvyčajne neprodukujú kontinuálne validné hodnoty šírky zrenice. Najväčším problémom nekvality dát sú chýbajúce alebo abnormálne hodnoty šírky zreníc v jednotlivých záznamoch. Abnormálne hodnoty sú ojedinelé výkyvy v hodnote šírky zrenice v jednom alebo malom počte po sebe idúcich záznamov, pričom ich hodnota presahuje fyziologické možnosti oka. Tieto abnormálne hodnoty sa často vyskytujú v intervaloch s množstvom chýbajúcich záznamov šírky zrenice. Ich ďalšou vlastnosťou môže byť krátka sekvencia hodnôt s kontinuálnym poklesom alebo nárastom hodnôt šírky zrenice. Validné, neprerušované intervaly hodnôt obsahujú drobné poklesy v nameranej šírke zrenice aj pri trende nárastu šírky zrenice a naopak.

Chýbajúce hodnoty nahrádzame štandardnou interpoláciou medzi dvomi validnými bodmi. Ako už bolo spomenuté, úplný výpadok hodnôt často predchádza a nasleduje postupnosť nevalidných záznamov. Interpolácia medzi najbližšími hodnotami ohraničujúcimi okno chýbajúcich dát by vo väčšine prípadov spôsobila ešte viac nepresností. Prvý validný bod preto identifikujeme ako bod, za ktorým nasledujú ešte dva mikro nárasty a dva mikro poklesy šírky zrenice v neprerušovanej postupnosti záznamov. Identifikácia druhého validného bodu je analogická, pričom v tomto prípade platia rovnaké podmienky pre body nachádzajúce sa pred vybraným bodom.

Chýbajúce hodnoty bodov fixácie interpolujeme pomocou poslednej a prvej zaznamenatej hodnoty bez ďalších špeciálnych ohraničení.

6.2 Priebeh spracovania experimentov

6.2.1 Výber hodnôt pre použitie v modeli reakcií zrenice

Pre použitie modelu reakcií zrenice musíme každému záznamu vo vstupných okulometrických dátach definovať trojicu hodnôt:

1. Reprezentačná šírka zrenice
2. Reprezentačný bod fixácie
3. Zobrazený stimul

Reprezentačná šírka zrenice a reprezentačný bod fixácie sú priemerné hodnoty záznamov pre ľavé aj pravé oko. V prípade, že jedna z hodnôt bola vypočítaná len pomocou interpolácie, používame váhovaný priemer preferujúci to oko, ktoré malo validné dáta. Váhu definuje dĺžka interpolovaného intervalu druhého oka. V kontexte modelu reakcií zrenice neuvažujeme prítomnosť sakád, pretože dôležitou vlastnosťou je určenie oblasti, ktorá má rozhodujúci vplyv na celkovú vnímanú svietivosť a nie identifikácia bodov, kde sa pohľad používateľa zastavil.

V prípade spracovávania datasetu založenom na stimuloch projektu je výber zobrazeného stimulu jednoduchý, pretože tento dataset musí obsahovať identifikátor zobrazeného stimulu pre každý záznam. V prípade datasetu založenom na videozázname vykonávame paralelnú iteráciu cez jednotlivé záznamy obrazoviek pomocou modulu OpenCV. Iterovanie cez záznamy okulometra a videozáznamu vykonávame pomocou cyklickej operácie výberu aktuálneho stimulu podľa nasledujúceho pseudokódu (5):

Zdrojový kód 5: Pseudokód synchronizácie okulometrického záznamu a videozáznamu

```
1 while not exists(next_screen) or offset(next_screen) < offset(record):  
2     if exists(next_screen):  
3         actual_screen = next_screen  
4     else:  
5         actual_screen = get_next_screen()  
6  
7     next_screen = get_next_screen()
```

Týmto spôsobom dokážeme synchronizovať iteráciu videozáznamu obrazovky a okulometrických dát bez ohľadu na rozdielnu frekvenciu záznamov jednotlivých vstupných prúdov či prípadných chýb v nedodržaní vlastného frekvenčného taktu. V skutočnosti majú okulometrické záznamy vyššiu frekvenciu ako prislúchajúce videozáznamy obrazovky, takže jeden záznam obrazovky obvyčajne pripadne viacerým okulometrickým záznamom.

6.2.2 Výpočet vnímanej svietivosti

Vnímaná svietivosť obrazovky sa podľa návrhu v metóde vykonáva na základe hodnôt každého z pixelov, čím sa stáva spracovanie experimentu výpočtovo náročné najmä pri datasetoch založených

na videozázname obrazovky. V prípade neoptimalizovanej implementácie musíme pre každý zo záznamov obrazovky vykonať radu netriviálnych úloh.

1. Vygenerovanie matice projekcie sietnice
2. Výpočet vnímanej svietivosti každého pixelu (až 2073600 iterácií v prípade FullHD rozlíšenia)
 - (a) Výpočet vnímanej svietivosti farby
 - (b) Vynásobenie váhou z matice projekcie sietnice
3. Výpočet vnímanej svietivosti okolia
4. Súčet príspevkov všetkých pixelov a vnímanej svietivosti okolia

Problém neoptimalizovaného prístupu tkvie v množstve vnorených iterácií, uprostred ktorých sa vykonávajú viaceré matematické operácie. Štandardný kvantitatívny experiment s 20 účastníkmi v trvaní 30 minút s frekvenciou 20 záznamov obrazovky a s rozlíšením dvoch miliónov pixelov na každý záznam predstavuje takmer 1,5 bilióna vnútorných iterácií. Takáto vysoká zložitosť by našu implementáciu celkom úplne diskvalifikovala od reálneho používania, takže sme navrhli niekoľko optimalizácií.

Zníženie rozlíšenia videozáznamu je prvou nutnou optimalizáciou. Už v jednoduchej implementácii v rámci predchádzajúcej práce sme dospeli k číslu 100x100 pixelov. Zníženie rozlíšenia efektívne vykonáva priamo modul OpenCV. Takto dramatické zníženie rozlíšenia stále poskytuje až 10000 vstupných hodnôt, z ktorých každá môže nadobúdať 255^3 hodnôt. Chyba v predpovedaní dilatácie na základe vnímanej svietivosti o niekoľko rádov prevyšuje rozlíšnosť vnímanej svietivosti takto upraveného videa. Z tohto dôvodu je rozdiel medzi predpovedanou svietivosťou zmenšeného a reálneho videozáznamu prakticky nulový avšak náš algoritmus sme zrýchlili až 200 krát v prípade videozáznamov s vysokým rozlíšením.

Neodporúčame ďalšie extrémne znižovanie rozlíšenia videa napriek dostatočnej rezerve v rozlíšiteľnosti vnímanej svietivosti. Veľkú úlohu vo výpočte svietivosti zohráva práve fixačný bod, ktorý má pri rozlíšení 100x100 pixelov a 30 cm širokej obrazovke plochu 9 mm štvorcových a pri každom znižovaní rozlíšenia rastie kvadraticky.

Videozáznam obrazovky môže obsahovať množstvo za sebou idúcich rovnakých snímok, kde jediným meniacim sa prvkom je iba kurzor myši. V takýchto prípadoch je zbytočné počítat vnímanú svietivosť pre tú istú obrazovku opakovane. Navrhujeme preskakovať záznamy obrazovky v prípadoch, že sa zmenia o menej ako 2.5% a zároveň nedôjde k výraznej zmene fixácie pohľadu používateľa.

Výpočet podobnosti dvoch záznamov obrazovky je výpočtovo náročná, no optimalizovaná funkcia v module OpenCV, s ktorou sa nám pri vysokej miere preskakovania záznamov obrazovky podarilo zrýchliť priebeh spracovania exportu na jednu desatinu času procesora. Výpočet podobnosti dvoch záznamov obrazoviek používa metódu normovaného rozdielu štvorcov podľa vzťahu (13),

$$zmena = \frac{\sum_{x,y} (V(x,y) - Z(x,y))^2}{\sqrt{\sum_{x,y} V(x,y)^2 * \sum_{x,y} Z(x,y)^2}} \quad (13)$$

kde V predstavuje vzorový, teda posledný ohodnotený záznam obrazovky a Z predstavuje prešetrovaný záznam obrazovky, x a y predstavujú súradnice pixelu v xovej a ypsilónovej osi.

Generovanie matice projekcie sietnice a výpočet vnímanej svietivosti okolia sú tiež netriviálne operácie, ktoré musia operovať s rovnakým rozlíšením ako optimalizované rozlíšenie stimulu. Pamäťovo nenáročnou optimalizáciou je vypočítanie matice vnímanej svietivosti okolia pre každý z možných bodov fixácie. Vďaka tejto matici je v priebehu experimentu možné zistiť vnímanú svietivosť okolia iba pomocou indexov z aktuálneho bodu fixácie. Kompletná matica s rozlíšením 100x100 pixelov zaberie len 80kb pamäte v prípade 64 bitovej presnosti dátového typu double. Pamäťovo náročnejšie, no stále efektívne je predpočítanie všetkých možných matíc projekcie sietnice. V tomto prípade je pamäťová náročnosť až 800Mb, pretože pre každý z možných bodov fixácie uchováame celú maticu hodnôt.

Kompletné iniciálne generovanie matíc má zmysel len pre experimenty s dĺžkou viac ako 3 minúty, ak predpokladáme že frekvencia okulometrického záznamu je 60 záznamov za sekundu a nechceme robiť výpočty matíc počas hlavného spracovávania datasetu, hlavne pri živom spracovávaní experimentu z prúdu dát. V akejkoľvek analýze po vykonaní experimentu je vhodné vypočítavať a ukladať vygenerované matice len po prvom dopyte konkrétneho bodu fixácie, čím predchádzame zbytočnému generovaniu nevyužitých matíc.

6.3 Diskusia

Schopnosť aplikovať teoretické vlastnosti navrhutej metódy do reálneho nástroja schopného spracovať často nedokonalé datasety je dôležitou súčasťou použiteľnosti konceptu skúmania kognitívnej záťaže pomocou absolútnej dilatácie šírky zrenice. V tejto kapitole sme prezentovali základné možnosti riešenia nevalidných vstupných dát ako aj problémov s efektivitou spracovávania objemných datasetov. Uvedenými optimalizačnými návrhmi sme dokázali zefektívniť spracovanie experimentov aj na výpočtových strojoch s parametrami súčasných osobných počítačov v čase kratšom ako samotné trvanie experimentu.

7 Overenie metódy odhadovania kognitívnej záťaže pre heterogénne rozhrania založenej na modeli reakcií zrenice

Vylepšený návrh metódy sme rozčlenili do samostatných celkov, pre ktoré sme vykonali experimentálne overenie. Získať rozsiahly experimentálny dataset nie je jednoduché, pretože pre vykonanie experimentu je potrebný špeciálny hardvér a softvér prístupný na fakulte. Z tohto dôvodu sme v našej práci okrem vykonávania nových experimentov zhromaždili aj datasety z predchádzajúcej práce, kde sme sa zvlášť venovali oprave rôznych technických chýb, vďaka čomu sme získali ďalšie použiteľné dáta.

Pre úplnosť a vďaka možnosti použiť viac validných dát sme vykonali vyhodnotenie nielen pridaných súčastí návrhu, ale aj pôvodných zistení z predchádzajúcej práce. Vlastnosti navrhutej metódy sme rozdelili do niekoľkých častí, ktoré je možné samostatne natrénovať, resp. overiť štatistickými metódami.

7.1 Zoznam štatisticky odhaliteľných vlastností a stanovených hypotéz

1. Zistenie ideálneho parametra sigma matice projekcie sietnice
 - (a) Hypotéza: Maticou projekcie sietnice dokážeme aproximovať rozloženie svetlocitlivých buniek v oku
 - (b) Overenie: Funkcia chyby vnímanej svietivosti podľa zvoleného parametra sigma matice projekcie sietnice má jediné lokálne minimum a toto minimum patrí sigme s kladnou hodnotou.
2. Overenie zníženej vnímanej svietivosti pre farby s dominanciou modrej zložky
 - (a) Hypotéza: Vnímaná svietivosť modrej farby je nižšia ako vnímaná svietivosť iných základných farieb, ktoré spĺňajú podmienku $\text{MAX}(R,G,B) = 255$
 - (b) Overenie: Nameraná šírka zrenice pri zobrazení modrej farby zodpovedá menšej hodnote svietivosti podľa krivky referenčných hodnôt v porovnaní s ostatnými sledovanými farbami.
3. Overenie poklesu vnímanej svietivosti v okrajoch monitora
 - (a) Hypotéza: Vnímaná svietivosť monitora sa mení s narastajúcou vzdialenosťou fixácie od prostriedku monitora
 - (b) Overenie: Pri zobrazení bielej obrazovky má priemer zrenice najväčšie hodnoty pri sledovaní rohu obrazovky, menšie hodnoty pri sledovaní prostriedku hrán a najmenšie pri sledovaní prostriedku obrazovky.

4. Overenie predpovedateľnosti šírky zrenice pre rozhrania webových sídel
 - (a) Hypotéza: Využitím modelu reakcií zrenice dokážeme predpovedať šírku zrenice pri zobrazení komplexných stimulov
 - (b) Overenie: Predpovedaná šírka zrenice koreluje so skutočne nameranou šírkou zrenice pri zobrazení rôznych webových stránok, čo sa prejaví rovnomerne predpovedanou kognitívnu záťažou pre úlohy s rovnakou náročnosťou ale rôznou svietivosťou stimulu.
5. Overenie detekcie kognitívnej záťaže pri práci s homogénnymi stimulmi
 - (a) Hypotéza: Využitím navrhutej metódy dokážeme detekovať prítomnosť kognitívnej záťaže pri zobrazení homogénnych stimulov
 - (b) Overenie: Identifikovaná kognitívna záťaž koreluje s objektívnou náročnosťou matematických úloh zobrazených na bielom pozadí s rovnakými svetelnými podmienkami.
6. Overenie detekcie kognitívnej záťaže pri práci s heterogénnymi stimulmi reprezentujúcimi skutočné rozhrania webových stránok
 - (a) Hypotéza: Využitím navrhutej metódy dokážeme detekovať prítomnosť kognitívnej záťaže pri zobrazení heterogénnych stimulov
 - (b) Overenie: Identifikovaná kognitívna záťaž koreluje s objektívnou náročnosťou matematických úloh zobrazených na popredí skutočných webových stránok s rozostretým obsahom a rôznymi svetelnými podmienkami.

7.2 Zoznam použitých datasetov

V tejto časti sumárne uvádzame vlastnosti získaných datasetov (tab. 2), ktoré sme použili vo fáze návrhu alebo overenia navrhutej metódy. Datasety, resp. ich jednotlivé intervaly striktne členíme na tréningové a overovacie dáta, aby sme zamedzili možnosti závislostí a následného nesprávneho vyhodnotenia. Dva z datasetov sme vytvorili výhradne pre túto prácu. Dva datasety pochádzajú z našej predchádzajúcej práce. Datasety z našej predchádzajúcej práce sme očistili od rôznych artefaktov spôsobených chybou, čím sme zvýšili počet validných vzoriek v týchto experimentoch. Zvyšné experimenty sme získali zaradením kalibračných alebo iných stimulov do experimentov iných autorov.

Všetky experimenty boli vykonané v laboratóriu pre skupinové testovanie vo výskumnom centre používateľského zážitku a interakcie na FIIT STU (UXI). Pre vykonanie experimentov sa použila prvá verzia systému pre vykonávanie okulometrických štúdií UXR alebo softvér Tobii Studio. Podrobný opis všetkých experimentov je uvedený v prílohe B.

Tabuľka 2: Zoznam použitých datasetov. Datasety (D) boli vykonané vo viacerých sedeniach s celkovým počtom unikátnych participantov (K). Autor označuje zodpovednú osobu za organizáciu účastníkov prípadne zber dát. UXI predstavuje participáciu viacerých členov výskumnej skupiny UXI. Autor UK predstavuje študentov a vedeckých pracovníkov Univerzity Komenského, ktorí vykonávali experimenty psychologického charakteru v ktorých sa nachádzali aj stimuly vhodné pre náš výskum.

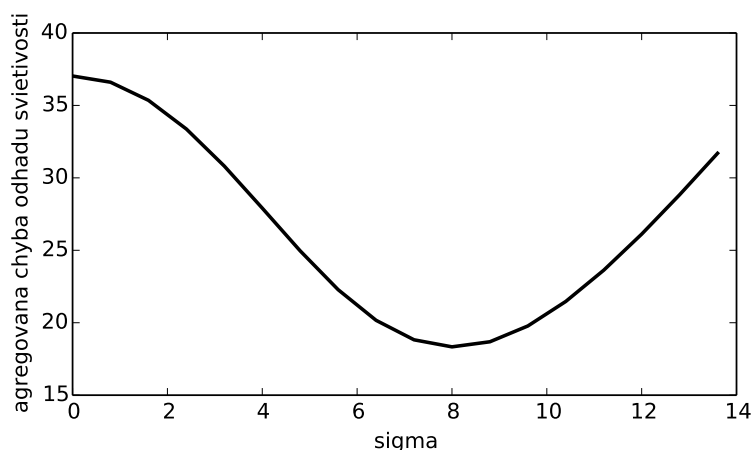
D	K	Téma	Obdobie	Autor
D1	63	Návrh modelu	BP (2016)	Tomáš Juhaniak + UXI
D2	19	Matematické úlohy	BP (2016)	Tomáš Juhaniak + UXI
D3	36	Čítanie, pamäť, matematické úlohy	DP (2017)	Tomáš Juhaniak + UXI
D4	39	Svietivosť okolia	DP (2017)	UK + Róbert Móro
D5	9	Tvar sietnice	DP (2017)	Dániel Papp + UXI
D6	21	Komplexné overenie	DP (2018)	Tomáš Juhaniak + UXI

7.3 Zistenie ideálneho parametra sigma matice projekcie sietnice

Hypotéza: Maticou projekcie sietnice dokážeme aproximovať rozloženie svetlocitlivých buniek v oku.

Overenie: Funkcia chyby vnímanej svietivosti podľa zvoleného parametra sigma matice projekcie sietnice má jediné lokálne minimum a toto minimum patrí sigme s kladnou hodnotou.

Skúmaná hypotéza vyjadruje predpoklad, že príspevok pixelov ku celkovej vnímanej svietivosti dokážeme aproximovať pomocou matice ohodnotení s normálnou 2D distribúciou váh s maximom v bode fixácie. Hypotéza môže byť overená, ak dokážeme numericky optimalizovať parameter sigma (pozri obr. 17), ktorý definuje normálne rozdelenie matice projekcie sietnice na takých dátach, kde sumárna svietivosť monitora je konštantná ale vnímaná svietivosť monitora závisí od bodu fixácie. K tejto formulácii overenia sme dospeli na základe výsledkov, ktoré naznačujú nezanedbateľný vplyv potenciálneho výskytu kognitívnej záťaže počas sledovania kalibračných stimulov.



Obr. 17: Ilustrácia numerickej optimalizácie parametra sigma pre jedného participanta.

Na dátach z datasetu (D1), ktorý obsahuje stimuly s kalibračnými kruhmi pre tréning parametra sigma a štandardné stimuly pre tréning krivky referenčných hodnôt, sme natrénovali model

reakcií zrenice pre 62 participantov. Zistili sme, že až u 16 participantov sme natrénovali optimálnu sigma s hodnotou 0, ktorá definuje takú maticu projekcie siete, kde všetky pixely obrazovky majú rovnaký vplyv na vnímanú svietivosť. Tento stav mohol nastať práve preto, že participant mal počas kalibrácie rozšírenú zrenicu vplyvom prítomnej kognitívnej záťaže a numerická optimalizácia preferovala čierne pixely, ktoré boli za bielym kalibračným kruhom.

Pri ďalších 6 participantoch došlo k natrénovaniu najvyššej povolenej hodnoty sigma, ktorú sme stanovili na hodnotu 13,6. V tomto prípade mohlo dôjsť k nesprávnej kalibrácii krivky referenčných hodnôt vplyvom kognitívnej záťaže. Ak je kognitívna záťaž prítomná pri trénovaní krivky referenčných hodnôt, numerická optimalizácia preferuje biele stimuly v bode fixácie, aby dosiahla neobjektívne vyššiu vnímanú svietivosť a kompenzovala menšie hodnoty nameranej šírky zrenice. Pre zvyšných 41 participantov sme natrénovali validné lokálne minimum v rozmedzí stanoveného rozsahu, pričom až 36 z nich malo optimálnu sigma v intervale 4,0 až 9,6.

Obdobný test sme vykonali aj nad dátami z datasetu (D6), kde sa nachádzali kalibračné body pre tréovanie vnímanej svietivosti okolia. Numerická optimalizácia v týchto bodoch môže fungovať lepšie, ak sú svetelné podmienky prostredia konštantné a rovnomerné, pretože takouto kalibráciou sme schopný dosiahnuť dve resp. tri úrovne vnímanej svietivosti bez zmeny celkovej vnímanej svietivosti obrazovky. Z 21 participantov mal iba jeden sigma s hodnotou 0 a traja participant mali sigma na hornej hranici povoleného rozsahu. 16 participantov malo optimálnu sigma v intervale 2,4 až 6,4. 1 participant mal optimálnu sigma s hodnotou 9,6.

Optimalizáciu parametra sigma sme sledovali aj na datasete (D5) s ôsmimi participantami, kde sa nachádzali čierne štvorce umiestnené striedavo nad, pod a vedľa centralizovaného bodu fixácie. V tomto experimente boli optimálne parametre sigma blízke hornej hranici povoleného rozsahu s hodnotou 13,6. Napriek tomu iba jeden participant dosiahol túto hranicu. Piaty participant mali ideálnu sigma v intervale 11,2 až 12,8 a jednému participantovi bola natrénovaná hodnota 8,0.

Sumárne výsledky prezentujeme v tabuľke (tab. 3).

Tabuľka 3: Histogramy ideálneho parametra sigma v troch experimentoch. Povolene hodnoty pre numerickú optimalizáciu boli násobky 0,8 v intervale $<0; 13,6>$. Tabuľka obsahuje len tie hodnoty, ktoré sa natrénovali aspoň v jednom experimente.

	0,0	2,4	3,2	4,0	4,8	5,6	6,4	7,2	8,0	8,8	9,6	10,4	11,2	12,0	12,8	13,6
D1	16	0	1	5	4	6	3	6	4	4	4	1	1	0	2	6
D6	1	3	1	2	7	2	1	0	0	0	1	0	0	0	0	3
D5	0	0	0	0	0	0	0	0	1	0	0	0	1	2	3	1

Zhodnotenie: Na veľkej vzorke participantov a pri použití rôznych kalibračných stimulov sme preskúmali možnosti kalibrácie parametra sigma. Zo získaných dát je zrejmé, že pri kalibrácii parametra sigma je dôležité vyvarovať sa nožnej chybe pod vplyvom kognitívnej záťaže. Platnosť stanovenej hypotézy je možná ale pre jej overenie bude nutné vykonať novú sériu experimentov. Navrhujeme použiť stimuly podobné stimulom v datasete (D5), kde bude čierny štvorec meniť svoju pozíciu tak, aby sa vyskytoval v rôznych vzdialenostiach od bodu fixácie. Ďalej navrhujeme, aby sa pre všetky stimuly s čiernym štvorcom na bielom pozadí vyskytli v experimente rovnaké stimuly s invertovanými farbami. Takéto stimuly by mohli odstrániť problém s pretrénovaním parametra

sigma na zvýšenú/zníženú úroveň kognitívnej záťaže oproti záťaži prítomnej pri kalibrácii krivky referenčných hodnôt.

7.4 Zistenie vnímanej svietivosti jednotlivých farieb

Hypotéza: Vnímaná svietivosť modrej farby je nižšia ako vnímaná svietivosť iných základných farieb, ktoré spĺňajú podmienku $\text{MAX}(\text{R}, \text{G}, \text{B}) = 255$.

Overenie: Nameraná šírka zrenice pri zobrazení modrej farby zodpovedá menšej hodnote svietivosti podľa krivky referenčných hodnôt v porovnaní s ostatnými sledovanými farbami.

Po opravení datasetu (D1) sa počet validných záznamov rozšíril o 10 dosiaľ nevyužitých meraní. Opätovne sme analyzovali vplyv základných RGB farieb na vnímanú svietivosť. Dáta jedného participanta boli z vyhodnocovania vylúčené, pretože pri jednom zo sledovaných stimulov jeho fixácia na stimul netrvala ani jednu sekundu. Vo vyhodnotení sme analyzovali relatívne šírky zreníc pre 62 účastníkov nasledujúcim postupom:

1. Pre každého účastníka zisti vnímanú svietivosť jednotlivých farieb vrátane duplicitných zobrazení tej istej farby
 - (a) Vyber záznamy pre obidve biele, šedé a jednu čiernu obrazovku a natrénuj krivku referenčných hodnôt
 - (b) Odmeraj šírku zrenice pre každý homogénny stimul
 - (c) Zisti svietivosť definovanú v krivke referenčných hodnôt, ktorá je ekvivalentná skutočne nameranej šírke zrenice.
2. Zo získaných hodnôt vypočítaj medián, aritmetický priemer a interval istoty aritmetického priemeru.

Kompletné výsledky prezentujeme v tabuľke (tab. 4).

Zo získaných výsledkov vyvodzujeme niekoľko vlastností reakcií zrenice. Štatistickú signifikanciu výsledkov sme potvrdili vykonaním neparametrického štatistického testu (Wilcoxon).

1. Modrá farba má signifikantne ($p < 0,001$) nižšiu vnímanú svietivosť ako iné základné farby, ktoré spĺňajú podmienku: $\text{MAX}(\text{R}, \text{G}, \text{B}) = 255$.
2. Purpurová farba má signifikantne ($p < 0,001$) nižšiu vnímanú svietivosť ako iné základné farby, ktoré spĺňajú podmienku: $\text{MAX}(\text{R}, \text{G}, \text{B}) = 255$ okrem modrej farby a zároveň má signifikantne ($p < 0,001$) vyššiu vnímanú svietivosť ako modrá farba.
3. Šírka zrenice sa prispôsobuje svietivosti vyžarovanej monitorom (biela, šedá a čierna obrazovka). Rozdiely vo vnímanej svietivosti všetkých troch farieb sú signifikantné ($p < 0,001$).

Zhodnotenie: Výsledky na čiastočne rozšírenom datasete nezmenili trend znižovania vnímanej svietivosti pri dominancii modrého kanála a submisii zeleného kanála, ktorý sme prezentovali už v predchádzajúcej práci ako diskriminátor modrej farby. Po vyladení algoritmu na vyhodnocovanie

Tabuľka 4: Štatistické výsledky vnímanej svietivosti farieb na vzorke 62 účastníkov. Biele a čierne stimuly nedosahujú svoje teoretické hodnoty, pretože tréning krivky referenčných hodnôt preferuje menšie šírky zrenice pri kalibrácii bielej obrazovky a väčšie šírky zrenice pri kalibrácii čiernej obrazovky. Toto správanie je očakávané, pretože priemerná šírka zrenice je ovplyvnená kognitívnou záťažou viac ako minimálna šírka zrenice. Ďalšie skreslenie môže byť spôsobené odlišným časovým intervalom ohraničujúcim hodnoty vstupujúce do kalibrácie a intervalom zobrazenia stimulu.

Farba	RGB kód	medián	priemer	95% t-interval istoty
Biela	FFFFFF	0,916	0,913	0,899 - 0,928
Červená	FF0000	0,950	0,956	0,908 - 1,004
Zelená	00FF00	0,886	0,879	0,841 - 0,917
Modrá	0000FF	0,696	0,647	0,598 - 0,696
Žltá	FFFF00	0,927	0,935	0,891 - 0,980
Purpurová	FF00FF	0,724	0,745	0,700 - 0,791
Azúrová	00FFFF	0,882	0,899	0,863 - 0,934
Šedá	AAAAAA	0,648	0,654	0,639 - 0,670
Čierna	000000	-0,007	-0,010	-0,027 - 0,006

vnímanej svietivosti a zaradení väčšieho množstva participantov do kvantitatívneho vyhodnotenia navrhujeme novú odporúčanú hodnotu základného diskriminátora modrej farby s hodnotou 0,304. Táto hodnota je rozdielom maximálnej možnej vnímanej svietivosti a mediánu vnímanej svietivosti modrej farby na vzorke 62 participantov.

7.5 Overenie poklesu vnímanej svietivosti v okrajoch monitora

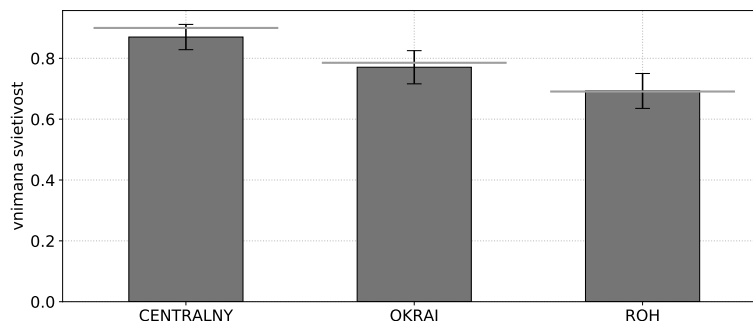
Hypotéza: Vnímaná svietivosť monitora sa mení s narastajúcou vzdialenosťou fixácie od prostriedku monitora.

Overenie: Nameraná šírka zrenice pri zobrazení bielej obrazovky a sledovaní rohu obrazovky zodpovedá menšej hodnote svietivosti podľa krivky referenčných hodnôt v porovnaní so sledovaním ostatných fixačných bodov. Analogický vzťah musí platiť aj pri sledovaní prostriedku hrany monitora vzhľadom k fixačnému bodu v prostriedku obrazovky.

Dataset (D3) obsahuje 36, dataset (D4) 39 a dataset (D6) ďalších 21 okulometrických záznamov, v ktorých sa účastníci pozerali do kalibračných bodov v okrajoch monitora. V datase (D3) bol z vyhodnotenia vylúčený jeden záznam z dôvodu technickej chyby nahrávacieho systému. Spolu sme tak získali dataset obsahujúci okulometrické dáta pre 95 pozorovaní účastníkov.

Vypočítali sme medián, priemer a 95% interval istoty priemernej vnímanej svietivosti pre všetky body fixácie v rohoch obrazovky, na okrajoch a v strede obrazovky. K priemerným hodnotám pre jednotlivé fixačné body sme vypočítali p-hodnoty nulových hypotéz, ktoré hovoria, že vnímaná svietivosť pri sledovaní týchto fixačných bodov je rovnaká. P-hodnoty sme vypočítali pomocou neparametrického párového štatistického testu (Wilcoxon). Keďže tvar rozdelenia hodnôt v jednotlivých skupinách pozorovaní je veľmi blízky normálnemu rozdeleniu, vypočítali sme aj p-hodnoty párovým t-testom. Oba testy významne preukázali ($p < 0,001$), že vnímaná svietivosť sa znižuje

s narastajúcou vzdialenosťou od prostriedku stimulu. Výsledky ilustruje graf (obr. 18).



Obr. 18: Porovnanie vnímanej svietivosti pri pozorovaní stredu, okraja a rohu obrazovky. Priemer vnímanej svietivosti pre bod fixácie je vyjadrený stĺpcom s vyznačeným 95%-intervalom istoty a hodnotou mediánu (horizontálna čiara).

Zhodnotenie: Vnímaná svietivosť obrazovky sa mení s narastajúcou vzdialenosťou fixácie od prostriedku stimulu, čo sme preukázali štatisticky signifikantným výsledkom. Tento výsledok je dosiahnuteľný iba v prípadoch, keď svietivosť okolia dosahuje nižšie hodnoty ako svietivosť monitora, čo je štandardná vlastnosť laboratórií na skúmanie používateľského zážitku.

Výsledok nás oprávňuje vyjadriť tvrdenie, že kognitívna záťaž nemôže byť interpretovaná iba pomocou skutočnej šírky zrenice, napriek zachovaniu rovnakej svietivosti stimulu, ak používateľ mení fixáciu pohľadu medzi stredom a okrajmi obrazovky.

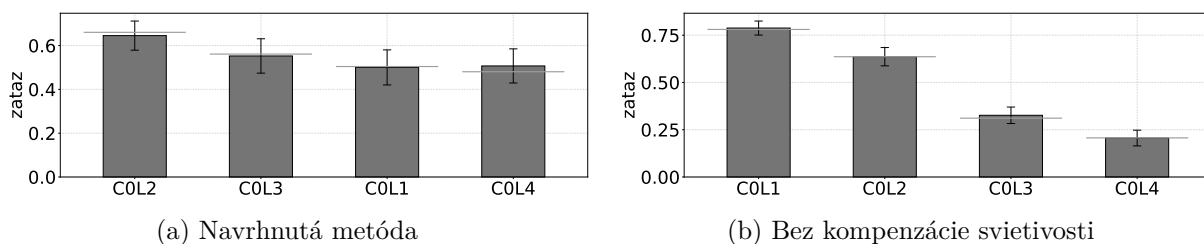
7.6 Overenie predpovedateľnosti šírky zrenice pre rozhrania webových sídel

Hypotéza: Využitím modelu reakcií zrenice dokážeme predpovedať šírku zrenice pri zobrazení komplexných stimulov.

Overenie: Predpovedaná šírka zrenice koreluje so skutočne nameranou šírkou zrenice pri zobrazení rôznych webových stránok, čo sa prejaví rovnomerne predpovedanou kognitívnou záťažou pre úlohy s rovnakou náročnosťou ale rôznou svietivosťou stimulu.

V datasete (D6) sa nachádza 72 matematických príkladov, z čoho 24 príkladov bolo jednoduchých (vyžadovalo sa sčítanie štyroch binárnych cifier). Pri týchto úlohách predpokladáme relatívne nízke kognitívne zaťaženie, takže dominantný vplyv na dilatáciu zrenice má vnímaná svietivosť webových stránok na pozadí riešených úloh. 24 jednoduchých úloh sme rozdelili na štyri skupiny podľa úrovne svietivosti webovej stránky alebo stimulu na pozadí. Každá zo sledovaných skupín má rovnakú náročnosť, takže v prípade dokonalej predpovede šírky zrenice pre každý zo zobrazených stimulov by mala byť odhadnutá približne rovnaká úroveň kognitívneho zaťaženia. Kognitívnu záťaž sme odhadli pomocou navrhutej metódy a tiež pomocou neupravenej šírky zrenice (pozri obr. 19).

Zhodnotenie: V porovnaní grafov v (obr. 19) vidieť významný rozdiel medzi úspešnosťou metódy a základným stavom, kedy sme svietivosť obrazovky vôbec neodhadovali. Ukázalo sa, že ak ignorujeme svietivosť stimulu, najtmavším obrazovkám je nesprávne odhadnutá najvyššia kognitívna



Obr. 19: Porovnanie odhadnutej kognitívnej záťaže metódou sledovania absolútnej dilatácie zrenice s využitím a bez využitia modelu reakcií zrenice pre úlohy s najmenšou obtiažnosťou. Úlohy sú rozdelené do skupín podľa svietivosti webovej stránky na ktorej sú zobrazené. Priemer odhadnutej záťaže pre skupinu úloh je vyjadrený stĺpcom s vyznačeným 95%-intervalom istoty a hodnotou mediánu (horizontálna čiara). C0 označuje skupinu úloh s binárnymi ciframi. $L < N >$ označuje úroveň svietivosti webovej stránky, kde $N=1$ je najtmavší web a $N=4$ je biele pozadie.

záťaž. Naopak, pomocou navrhnutej metódy sa podarilo predpovedať nárast priemeru zrenice pri tmavých (L1) weboch, čo spôsobilo, že úlohám na bielej obrazovke (L4) sa podarilo správne odhadnúť rovnakú kognitívnu záťaž ako úlohám na tmavých weboch. Pri weboch so strednou (L3) a vysokou (L2) úrovňou svietivosti model predpovedal čiastočne užšiu zrenicu, čo spôsobilo malý nárast v predpovedanej kognitívnej záťaži.

Experiment vyžadoval správne predpovedanie šírky zrenice na najväčšom možnom rozpätí vnímanej svietivosti, pričom takto významné rozdiely nie sú bežne prítomné v štandardných webových aplikáciách, ktoré dodržiujú uniformný vzhľad naprieč celou aplikáciou. Napriek významným rozdielom vo vnímanej svietivosti sa podarilo predpovedať šírku zrenice pomerne stabilne a metóda splnila naše očakávania. Predmetom ďalších vylepšení je otázka, prečo model reakcií zrenice významne preceňuje svietivosť stredne tmavých webových stránok.

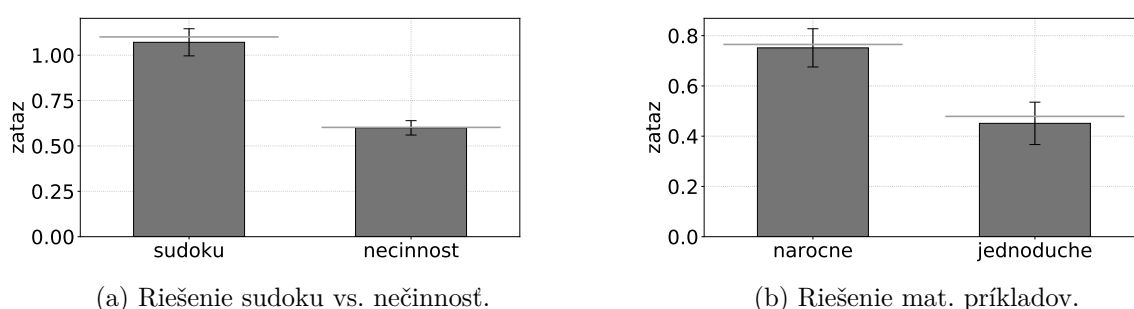
7.7 Overenie detekcie kognitívnej záťaže

7.7.1 Kognitívne náročné a nenáročné úlohy

Najjednoduchšie porovnanie detekovanej kognitívnej záťaže sme vypracovali, podobne ako v predchádzajúcej práci, nad dátami z riešenia jednoduchého sudoku v datasete (D1). V predchádzajúcej práci sme vypočítali iba kvartily nad množinami hodnôt záťaže pre jednotlivých účastníkov, v súčasnej práci sme vykonali porovnanie 95% t-intervalu istoty na bohatšej verzii datasetu. V grafe (obr. 20a) je signifikantný rozdiel medzi detekovanou záťažou úlohy so sudoku a kognitívnu záťažou odhadnutou pri sledovaní rôznych webových stránok, kedy participanti nemali zadanú nijakú úlohu. Môžeme konštatovať, že medián detekovanej záťaže kognitívne náročnej úlohy sa vylepšením metódy významne nezmenil na rozdiel od poklesu chýb pri stimuloch bez predpokladanej záťaže. Do vyhodnotenia bolo z datasetu (D1) zaradených 48 participantov, ktorí videli všetky webové stránky a riešili jednoduché sudoku.

V datasete (D2) 19 participantov počítalo matematické príklady so štyrmi úrovňami obtiažnosti na pozadí webov s rôznou svietivosťou. Z dôvodu technickej chyby nahrávacieho zariadenia bol z vyhodnotenia vylúčený jeden participant.

Štyri úrovne úloh sme rozdelili do dvoch skupín: jednoduché a náročné. Skupinu jednoduchých úloh tvoria iba príklady na súčet jednociferných čísel. Súčty dvojciferných a trojciferných čísel predstavuje skupina náročných príkladov. Pre tieto dve skupiny sme vytvorili štatistické vyhodnotenie kognitívnej záťaže pomocou mediánu ako aj priemernú záťaž a 95% intervaly istoty ilustrovanú grafom (obr. 20b). Ukázalo sa, že zistená kognitívna záťaž pre skupinu náročných príkladov je signifikantne väčšia ako pre skupinu jednoduchých úloh.



Obr. 20: Grafy odhadnutej kognitívnej záťaže pri úlohách s rozdielnymi kognitívnymi nárokmi. Priemer odhadnutej záťaže pre skupinu úloh je vyjadrený stĺpcom s vyznačeným 95%-intervalom istoty a hodnotou mediánu (horizontálna čiara).

Tento experiment simuluje A/B testovanie webových aplikácií. Skupina náročných úloh resp. sudoku simuluje skupinu úloh v zle navrhutej verzii A a skupina jednoduchých úloh resp. nečinnosť simuluje skupinu úloh v lepšie navrhutej verzii B. Signifikantný rozdiel medzi verziami uisťuje, že ak bolo nastavenie experimentu v súlade so základnými predpokladmi správnosti A/B testovania, verzia B vyvoláva u používateľov menšiu kognitívnu záťaž.

7.7.2 Overenie detekcie kognitívnej záťaže pri homogénnych stimuloch

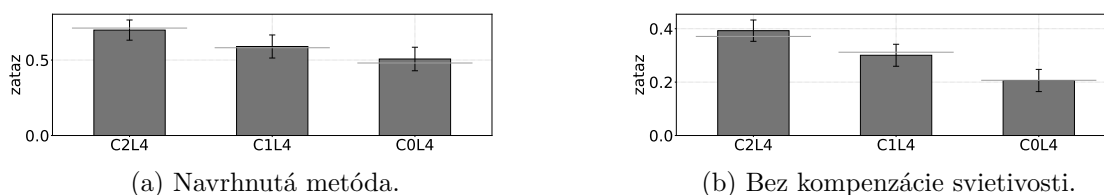
Hypotéza: Využitím navrhutej metódy dokážeme detekovať prítomnosť kognitívnej záťaže pri zobrazení homogénnych stimulov.

Overenie: Identifikovaná kognitívna záťaž koreluje s objektívnou náročnosťou matematických úloh zobrazených na bielom pozadí s rovnakými svetelnými podmienkami.

Cieľom tohto overenia je potvrdiť predpoklady získané v analýze dilatácie zrenice, podľa ktorých šírka zrenice koreluje s intenzitou kognitívnej záťaže a potvrdiť schopnosť okulometra tieto zmeny detekovať s dostatočnou presnosťou.

V datasete (D6) sa nachádza 72 úloh s tromi úrovňami náročnosti. Úlohy boli zobrazené v náhodnom poradí. 18 úloh bolo zobrazených iba na bielom homogénnom pozadí. Skupinu najľahších úloh tvorí 6 príkladov na sčítanie troch binárnych čísel. Teoreticky náročnejšiu skupinu tvorí 6 príkladov

na sčítanie jednociferných čísel, pričom pri každom príklade došlo k prechodu cez desiatku. Najnáročnejšiu skupinu tvorí 6 príkladov na sčítanie dvojčiferných čísel, pričom aj pri týchto príkladoch došlo ku prechodu cez desiatku. Do vyhodnotenia sme zaradili všetkých 21 participantov a odhadli kognitívnu záťaž pomocou navrhnutej metódy a výsledok porovnali s odhadom bez uvažovania svietivosti stimulu (pozri obr. 21).



Obr. 21: Grafy odhadnutej kognitívnej záťaže pri úlohách s rozdielnymi kognitívnymi nárokmi, ktoré boli zobrazené na bielom homogénnom pozadí. Priemer odhadnutej záťaže pre skupinu úloh je vyjadrený stĺpcom s vyznačeným 95%-intervalom istoty a hodnotou mediánu (horizontálna čiara). C0 sú úlohy s binárnymi ciframi, C1 sú jednociferné príklady a C2 sú dvojčiferné príklady. L4 sú úlohy zobrazené na stimuloch s najvyššou možnou svietivosťou.

Výsledky podľa očakávaní určili správne poradie náročnosti jednotlivých úloh obidvoma metódami, pretože uvažovať svietivosť nebolo pri tomto type experimentu nutné. V grafoch (obr. 21) je jediným rozdielom odlišná hodnota relatívnej kognitívnej záťaže pre jednotlivé skupiny úloh. Tento rozdiel je spôsobený druhou fázou normalizácie, z čoho je možné dedukovať, že pri metóde bez kompenzácie svietivosti došlo k natrénovaniu väčšieho kognitívneho dilatáčného rozsahu, pretože táto metóda odhaduje vyššiu kognitívnu záťaž pre ťažké úlohy zobrazené na čiernych stimuloch ako nami navrhnutá metóda.

Zaujímavý je pohľad na štatistickú signifikanciu rozdielu odhadnutej náročnosti medzi pozorovanými skupinami úloh. Distribúcia odhadov kognitívnej záťaže v jednotlivých skupinách je v tvare normálneho rozdelenia a vykonané párové t-testy potvrdili štatistickú signifikanciu aj medzi najpodobnejšou skupinou jednociferných a binárnych súčtov ($p=0,0032$). Významnú signifikanciu výsledkov potvrdzuje aj neparametrický test (Wilcoxon: $p=0,0071$).

Zhodnotenie: Týmto experimentom sa jednoznačne preukázala vysoká citlivosť dilatácie zrenice, ktorá je schopná detekovať aj tie najjemnejšie rozdiely v kognitívnej záťaži na vzorke 21 participantov, čo je prijateľné množstvo participantov pre testovanie používateľského zážitku a použiteľnosti webových aplikácií. Pre dosiahnutie takto vysokej presnosti aj pri reálnych weboch je potrebné poznať takmer dokonalý model reakcií zrenice na zmeny v svietivosti. V kapitole 7.6 sme demonštrovali presnosť súčasného modelu reakcií zrenice, ktorý nám ešte neumožňuje tak citlivé porovnávanie, no ukazuje že tento cieľ nemusí byť nedosiahnuteľný.

7.7.3 Overenie detekcie kognitívnej záťaže pri práci s heterogénnymi stimulmi reprezentujúcimi skutočné rozhrania webových stránok

Hypotéza: Využitím navrhutej metódy dokážeme detekovať prítomnosť kognitívnej záťaže pri zobrazení heterogénnych stimulov.

Overenie: Identifikovaná kognitívna záťaž koreluje s objektívnou náročnosťou matematických úloh zobrazených na popredí skutočných webových stránok s rozostretým obsahom a rôznymi svetelnými podmienkami.

Cieľom tohto overenia je potvrdiť schopnosť navrhutej metódy vysporiadať sa s rôznymi svetelnými podmienkami, ktoré vytvárajú skutočné rozhrania webových aplikácií a správne detekovať kognitívnu záťaž počas práce s týmito rozhraniami.

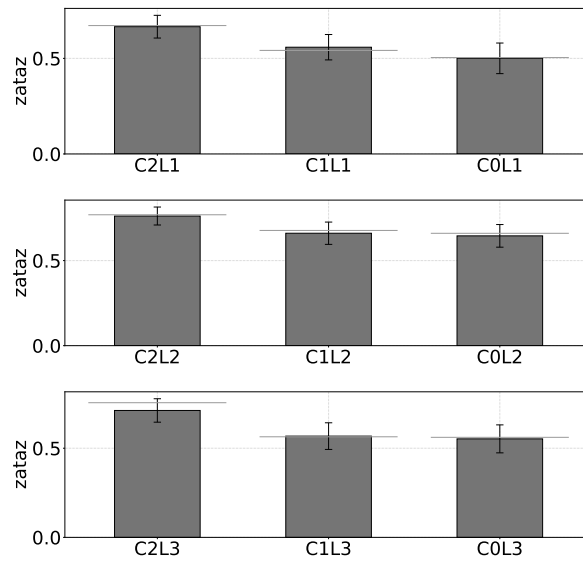
V datasete (D6) sa nachádza 72 úloh s tromi úrovňami náročnosti zobrazených v náhodnom poradí. 54 úloh bolo zobrazených na popredí skutočných komerčných webových aplikácií, ktorých obsah bol upravený tak, aby tento obsah nebol čitateľný. 18 úloh bolo zobrazených na skupine tmavých rozhraní (skupina L1), 18 úloh bolo zobrazených na skupine polotmavých rozhraní (skupina L2) a 18 úloh bolo zobrazených na skupine svetlých rozhraní (skupina L3). V každej skupine boli tri rôzne rozhrania. Polovica úloh bola centralizovaná a pri druhej polovici sa používatelia museli pozeráť do rôznych častí monitora. V každej skupine úloh boli príklady troch úrovní náročnosti. Kombinácie všetkých vlastností boli rovnomerne distribuované a úlohy sa používateľom zobrazovali v náhodnom poradí.

Prvá séria vyhodnotení má za cieľ ukázať, že navrhnutá metóda po spracovaní komplexných heterogénnych stimulov s ľubovoľnými bodmi fixácie dokáže identifikovať rozdiel v kognitívnej záťaži. Tento typ experimentu sa najviac podobá skutočnému používateľskému vyhodnocovaniu, pretože každá skupina rozhraní má približne rovnakú sumárnu svietivosť ale rozdielny obsah, čím simulujeme jednotlivé podstránky systému. Výsledky odhadnutej kognitívnej záťaže pre jednotlivé skupiny úloh sú v (obr. 22).

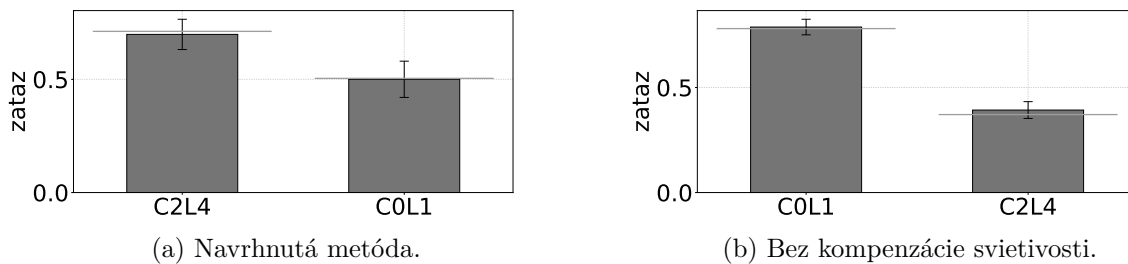
Odhady kognitívnej záťaže pre jednotlivé skupiny dopadli podľa očakávaní, úlohy boli správne zoradené podľa kognitívnej náročnosti a skupine najťažších príkladov bola odhadnutá signifikantne ($p < 0,005$) väčšia kognitívna záťaž. Podotýkame, že podobné výsledky by sme dosiahli aj bez využitia modelu reakcií zrenice, no vystavili by sme sa riziku nesprávneho odhadu kognitívnej záťaže pri ojedinelých výskytoch tmavšieho rozhrania.

Asi najdôležitejším prínosom metódy je objektívny odhad kognitívnej záťaže aj pri úlohách s odlišnou úrovňou svietivosti. Metódu sme podrobili testu správneho odlíšenia náročnosti úloh s rôznou svietivosťou a výsledky ilustrujeme na príklade najextrémnejších situácií, kde metódy ignorujúce svietivosť stimulu úplne zlyhávajú (pozri obr. 23). Týmito situáciami sú porovnania dvoch úloh, kde je jednoduchá úloha zobrazená na tmavom pozadí a náročná úloha je zobrazená na svetlom pozadí. Dominancia reakcií zrenice na svetelné podmienky spôsobí, že pri ťažšej úlohe je šírka zrenice užšia ako pri úlohe jednoduchšej.

Zhodnotenie: Získaný výsledok dokazuje schopnosť našej metódy vysporiadať sa s odlišnými svetelnými podmienkami zobrazeného rozhrania a štatisticky signifikantne ($p < 0,001$) určiť nároč-



Obr. 22: Grafy odhadnutej kognitívnej záťaže pre každú skupinu svietivosti rozhraní (L1, L2 a L3). V každej skupine sú vyčlenené tri úrovne náročnosti úloh: C0 sú úlohy s binárnymi ciframi, C1 sú jednociferné príklady a C2 sú dvojciferné príklady.



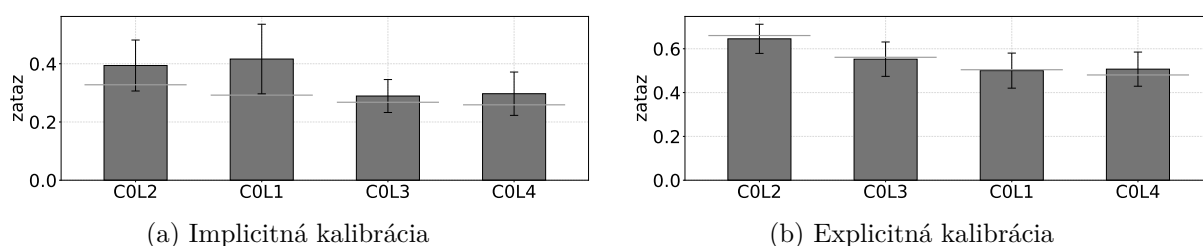
Obr. 23: Grafy odhadnutej kognitívnej záťaže pri úlohách s najväčším vplyvom rušivej svietivosti. Navrhnutá metóda dokázala odfiltrovať vplyv svietivosti a správne určiť kognitívnu záťaž pri jednotlivých úlohách. Pri metóde bez kompenzácie vplyvu svietivosti došlo k nesprávnemu odhadu vysokej kognitívnej záťaže pre ľahké úlohy a podceneniu kognitívnej záťaže ťažkých úloh. Priemer odhadnutej záťaže pre skupinu úloh je vyjadrený stĺpcom s vyznačeným 95%-intervalom istoty a hodnotou mediánu (horizontálna čiara). C0 sú úlohy s binárnymi ciframi a C2 sú dvojciferné príklady. L4 sú úlohy zobrazené na stimuloch s najvyššou možnou svietivosťou a L1 sú tmavé weby.

nejšiu úlohu. Zároveň sme preukázali, akú významnú chybu môže spôsobiť ignorovanie svietivosti rozhrania. Tento problém sa môže v praxi prejavovať v rôznych situáciách, napr. ak niekto preukáže signifikantne väčšiu šírku zrenice pri rozhraní s tmavou nápovedou ako pri rozhraní bez nápovedy. Potenciálny výskumník sa tak môže mylne domnievať, že nápoveda rušila používateľa natoľko, že ju odstráni, čím môže paradoxne dôjsť k zhoršeniu situácie.

7.7.4 Porovnanie presnosti explicitnej a implicitnej kalibrácie

Všetky doterajšie vyhodnotenia sme robili na základe explicitnej kalibrácie. Pre určité prípady, kedy nie je možné použiť explicitnú kalibráciu sme navrhli možnosť implicitnej kalibrácie modelu reakcií zrenice v dvoch alternatívach. V našej implementácii sme vytvorili hybridnú verziu, ktorá kombinuje kalibráciu krivky referenčných hodnôt na obmedzenej ploche uprostred stimulu a používa natrénovanú svietivosť okolia získanú v explicitnej kalibrácii.

Vykonalí sme rovnaké overenie schopnosti predpovedať šírku zrenice pre rôzne úrovne svietivosti webových stránok ako v podkapitole 7.6. Výsledky prezentujeme v konfrontácii s explicitnou kalibráciou v grafoch na (obr. 24).



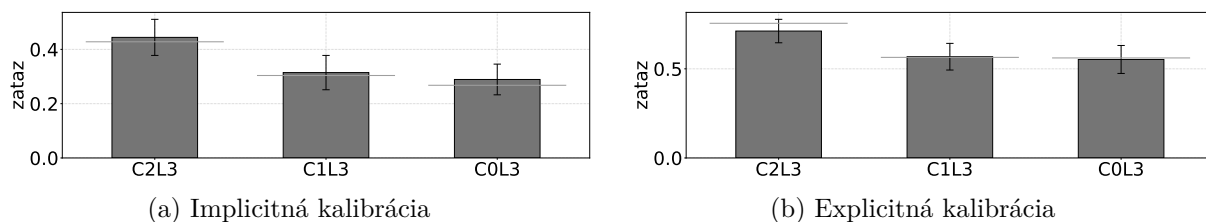
Obr. 24: Porovnanie odhadnutej kognitívnej záťaže s implicitnou a explicitnou kalibráciou modelu reakcií zrenice pre úlohy s najmenšou obtiažnosťou. Úlohy sú rozdelené do skupín podľa svietivosti webovej stránky na ktorej sú zobrazené. Priemer odhadnutej záťaže pre skupinu úloh je vyjadrený stĺpcom s vyznačeným 95%-intervalom istoty a hodnotou mediánu (horizontálna čiara). C0 označuje skupinu úloh s binárnymi ciframi. L<N> označuje úroveň svietivosti webovej stránky, kde N=1 je najtmavší web a N=4 je biele pozadie.

Z porovnania oboch typov kalibrácie modelu, ktorý predpovedal šírky zrenice a následne kognitívnu záťaž pre jednotlivé skupiny svietivosti vyvodzujeme, že implicitná kalibrácia je do určitej miery schopná natrénovať model reakcií zrenice, tento však nie je tak presný ako jeho explicitne kalibrovaná alternatíva. Najväčší problém vidíme vo väčších intervaloch istoty predpovedanej záťaže. Väčšie intervaly istoty znamenajú väčší rozptyl predpovedaných hodnôt a teda väčšiu chybovosť v jednotlivých prípadoch.

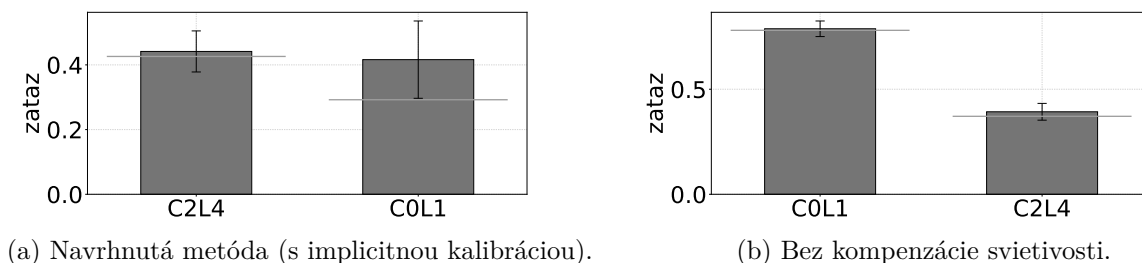
Vykonalí sme tiež odhady kognitívnej záťaže pre všetky tri úrovne náročnosti pre skupinu svetlých webových stránok (L3). Výsledky získané pomocou implicitnej kalibrácie porovnávame s výsledkami tej istej skupiny úloh ohodnotenej explicitne kalibrovaným modelom (pozri obr. 25).

Vykonalí sme aj porovnanie v najextrémnejšej situácii s vysokou svietivosťou pri náročných a nízkou svietivosťou pri jednoduchých úlohách podobne ako na konci podkapitoly 7.7.3. V tomto prípade porovnávame výsledky našej metódy s implicitnou kalibráciou a metódy bez kompenzácie svietivosti (obr. 26), keďže chceme ukázať, aký efekt má výber našej metódy pokiaľ nie je možnosť vykonať experiment s explicitnou kalibráciou a rozhodujeme sa, či použiť model reakcií zrenice.

Zhodnotenie: Metóda na základe porovnania mediánov odhadnutej kognitívnej záťaže správne určila náročnejšiu úlohu aj pri úlohách s najväčším kontrastom v záťaži a svietivosti, no priemerné hodnoty sú veľmi podobné a rozdiel nie je štatisticky signifikantný. Z celkového hľadiska je implicitne



Obr. 25: Porovnanie odhadnutej kognitívnej záťaže s implicitnou a explicitnou kalibráciou modelu reakcií zrenice pre úlohy zobrazené na skupine svetlých webových stránok. Úlohy sú rozdelené do skupín podľa náročnosti. Priemer odhadnutej záťaže pre skupinu úloh je vyjadrený stĺpcom s vyznačeným 95%-intervalom istoty a hodnotou mediánu (horizontálna čiara). C0 sú úlohy s binárnymi ciframi, C1 sú jednociferné príklady a C2 sú dvojciferné príklady.



Obr. 26: Grafy odhadnutej kognitívnej záťaže pri úlohách s najväčším vplyvom rušivej svetivosti. Implicitná kalibrácia dokázala natrénovať model reakcií zrenice do takej miery, že bol schopný čiastočne kompenzovať vplyv svetivosti. Priemer odhadnutej záťaže pre skupinu úloh je vyjadrený stĺpcom s vyznačeným 95%-intervalom istoty a hodnotou mediánu (horizontálna čiara). C0 sú úlohy s binárnymi ciframi a C2 sú dvojciferné príklady. L4 sú úlohy zobrazené na stimuloch s najvyššou možnou svetivosťou a L1 sú tmavé weby.

kalibrovaný model reakcií zrenice lepšou alternatívou ako úplné ignorovanie vnímanej svetivosti. Jeho použitie preto naďalej odporúčame v prípadoch, keď nie je možné nastaviť vyhodnocovanie tak, aby sme model natrénovali explicitne. V porovnaní s explicitne kalibrovaným modelom, implicitne kalibrovaný model podáva horšie výsledky.

7.8 Diskusia

Na dátach zo šiestich datasetov, v ktorých sa vykonalo 187 jedinečných okulometrických pozorovaní práce človeka s počítačom sme vykonali niekoľko experimentov, ktoré potvrdzujú funkčnosť nami navrhnutej metódy na vyhodnocovanie kognitívnej záťaže. V našej práci bolo nutné priebežne skúmať vlastnosti dilatácie zrenice a navrhovať riešenia, ktoré by mohli pomôcť k dosiahnutiu hlavného stanoveného cieľa - návrhu čo najpresnejšej metódy na odhadovanie kognitívnej záťaže vrátane spoľahlivého modelu reakcií zrenice. Z tohto dôvodu bolo potrebné vykonať viaceré experimenty s rôznym zameraním. Posledný experiment slúžil na komplexné overenie navrhnutého riešenia a väčšiny stanovených hypotéz.

Medzi najväčšie poznatky získané vyhodnotením experimentov radíme experimentálny odhad vnímanej svietivosti okolia za štandardných svetelných podmienok pre konkrétne podmienky fakultného laboratória pre výskum používateľského zážitku. Podarilo sa nám dokázať, že vnímaná svietivosť v blízkosti okraja monitora je menšia, čo sa prejavilo rozšírenou zrenicou. Ďalším zaujímavým výsledkom je preukázanie vysokej citlivosti reakcií zrenice na zmenu v kognitívnej náročnosti. Aj keď, zistený výsledok iba potvrdil desaťročia overenú schopnosť zrenice reagovať na kognitívne zmeny, naše vyhodnotenie stanovilo hranicu, pokiaľ je možné vylepšovanie navrhnutej metódy iba pomocou zdokonaľovania modelu reakcií zrenice.

Najdôležitejším overením bolo porovnanie odhadnutej kognitívnej záťaže pre úlohy zobrazené na stimuloch so svietivosťou pôsobiaceou proti smeru očakávanej zmeny šírky zrenice. Úspešne sme preukázali, že navrhnutá metóda sa dokáže vysporiadať s rôznymi svetelnými podmienkami a identifikovať relatívnu kognitívnu záťaž, ktorá sa vyskytuje pri úlohách v rámci jedného experimentu. Úspešnosť metódy sme porovnali s alternatívou úplného ignorovania svietivosti stimulu a vykonali sme porovnanie úspešnosti metódy s použitím explicitne aj implicitne kalibrovaného modelu. Najlepšie výsledky dosiahla metóda pri použití explicitnej kalibrácie.

8 Záver

V našej práci, ktorá nadviazala na výstup predchádzajúcej práce sme objasnili a obohatili niektoré časti modelu reakcií zrenice, jeho kalibrácie, či jednotlivých častí pri vyhodnocovaní kognitívnej záťaže. Za najväčší prínos pokladáme návrh vytvorenia matice projekcie sietnice, ktorú dokážeme generovať pre ľubovoľný bod fixácie. Ďalším prínosom je navrhnutý spôsob automatického štatistického vyhodnocovania, ktorý sme úspešne implementovali do novovytvorenej webovej aplikácie, ktorá je schopná zbierať a vyhodnocovať široké spektrum výskumných experimentov minimálne v rámci fakulty. Analyzovali sme možnosti implicitnej kalibrácie modelu reakcií zrenice, ktorou sa môže zjednodušiť použitie navrhnutej metódy zisťovania kognitívnej záťaže v používateľských štúdiách.

Opravili sme datasety z predchádzajúcej práce a tiež sme vykonali alebo asistovali pri niekoľkých experimentoch. Pre návrh a overenie súčastí prezentovanej v tejto práci sme tak získali kvalitné datasety okulometrických záznamov 187 účastníkov. Výsledky navrhnutej metódy sú uspokojivé, pričom vidíme potenciál pre ďalšie zdokonalenie.

8.1 Naplnenie stanovených cieľov

Na začiatku práce sme si stanovili niekoľko cieľov, ktoré požadovali vypracovanie konkrétnych príspevkov k poznaniu alebo analýzu vhodných praktík pre oblasť používateľského testovania. Naším hlavným zámerom bolo pretaviť ideu modelu reakcií zrenice z predchádzajúcej práce do uceleného návrhu metódy v bežnej výskumnej praxi.

Identifikácia relevantných faktorov vplyvujúcich na vnímanú svietivosť. Pôvodný návrh modelu reakcií zrenice sme rozšírili o ďalší dôležitý parameter - vnímanú svietivosť okolia. Rozvinuli sme teóriu o celkovej vnímanej svietivosti s ohľadom na necentralizované body fixácie, kedy je potrebné uvažovať svietivosť okolia. Experimentálne sme overili, že svietivosť okolia ovplyvňuje vnímanú svietivosť a jej vplyv narastá so vzdialenosťou fixácie od prostriedku obrazovky.

Skúmali sme aj problematiku reprezentácie matice projekcie sietnice ako druhého parametra modelu. Zistili sme, že parameter sigma, ktorý definuje hodnoty matice a jeho určenie má výrazný vplyv na správnosť odhadnutej svietivosti, je potrebné trénovať komplexnejším spôsobom. Dôležitým poznatkom je aj zistenie, že vnímaná svietivosť monitora nemá lineárnu koreláciu so šírkou zrenice prispôbenú na konkrétnu hodnotu svietivosti. V našej práci sme preto pristúpili ku kvadratickej interpretácii krivky referenčných hodnôt.

Návrh efektívnej kalibrácie faktorov pôsobiacich na vnímanú svietivosť. Vypracovali sme efektívny spôsob kalibrácie svietivosti okolia pomocou genericky generovanej matice projekcie sietnice a sústavy jednoduchých rovníc, ktoré deterministicky určujú pomer vplyvu svietivosti okolia k celkovej vnímanej svietivosti. Vďaka týmto vzťahom a s opretím sa o základnú hypotézu modelu reakcií zrenice je možné nielen určiť vnímanú svietivosť v konkrétnom bode fixácie ako agregát oboch zdrojov vnímanej svietivosti ale zároveň aj kalibrovať vnímanú svietivosť okolia.

Navrhli sme viackrokovú optimalizáciu kalibrácie krivky referenčných hodnôt. Po zmene kalibračnej metódy z lineárnej regresie na kvadratickú sa vystavuje krivka referenčných hodnôt väčšiemu riziku chybnéj kalibrácie v dôsledku náhodného výskytu kognitívnej záťaže. Kalibračnú metódu sme obohatili o optimalizačné kroky, ktoré majú identifikovať a pozmeniť teoreticky neprípustné hodnoty kalibrovanej šírky zrenice s potenciálnym ovplyvnením kognitívnou záťažou.

Po vyhodnotení experimentov sme dospeli k záveru, že navrhovaný spôsob kalibrácie parametra sigma matice projekcie sietnice môže byť veľmi náchylný na náhodný výskyt kognitívnej záťaže. Práve kalibráciu tohto parametra, spolu s kalibráciou krivky referenčných hodnôt, považujeme za najdôležitejšie súčasti navrhutej metódy, ktoré po zdokonalení môžu priniesť významné zlepšenie metódy.

Návrh kvantitatívneho vyhodnotenia experimentov. Navrhli sme podrobný postup transformácie získaných výsledkov od jednotlivých hodnôt šírky zrenice až po sumárne vyhodnotenie jednotlivých úloh v používateľskom testovaní. Navrhli sme dvojfázovú normalizáciu výsledkov, ktorá odstraňuje problémy s veľkou variabilitou schopností a biologických vlastností testovaných participantov. Po analýze dát získaných našou metódou a s inšpiráciou v najzaužívanejších a odporúčaných metódach vyhodnocovania používateľského zážitku sme navrhli postup pre korektné štatistické vyhodnotenie experimentov spolu s vhodnými štatistickými testami.

Preukázanie schopnosti navrhutej metódy identifikovať a klasifikovať kognitívnu záťaž do viacerých tried náročnosti. Navrhnutá metóda dokáže rozlíšiť kognitívnu záťaž medzi úlohami s rôznou náročnosťou. Usporiadáním testovaných úloh od najproblémovnejších po najjednoduchšie umožňujeme používateľovi vizuálne porovnať, ktorá z testovaných úloh, či prípadov použitia si vyžaduje najväčšiu pozornosť pri zlepšovaní testovaných aplikácií. Vďaka dvojfázovej normalizácii v rámci jedného experimentu je zaručená klasifikácia úloh na (relatívne) ťažké, či ľahké. Granularitu tried si môže určiť výskumník pomocou hraníc v ohodnotení úloh relatívnou kognitívnou záťažou s ohľadom na rozsah intervalov istoty. Schopnosť klasifikovať náročnosť úloh sme signifikantne dokázali pre skupinu dvojciferných a jednociferných matematických príkladov. Zároveň sme preukázali možnosti metódy v dosiahnutí omnoho vyššej granularity pomocou štatisticky významného rozdielu jednociferných a binárnych matematických príkladov zobrazených na homogénnom pozadí.

Porovnanie rozdielu v predikovanej kognitívnej záťaže pomocou metódy sledovania surovej šírky zrenice a našej metódy. Preukázali sme významný rozdiel v predikovaní kognitívnej záťaže ťažkej úlohy zobrazenej na svetlom stimule a jednoduchej úlohy zobrazenej na tmavej webovej stránke. Naša metóda dokázala spoľahlivo predpovedať dominantný nárast zrenice pri tmavom stimule, pričom menší nárast šírky zrenice pri svetlom stimule správne identifikovala ako nadpriemerný a ťažkej úlohe predpovedala najvyšší vplyv záťaže. Výsledok sme porovnali s metódou bez kompenzácie vnímanej svietivosti, ktorá v tomto teste zlyhala a predpovedala záťaž u jednoduchej úlohy. Vykonali sme aj validáciu, či metóda neklasifikuje úlohy na základe svietivosti stimulu. Naša metóda preukázala pomerne stabilný odhad kognitívnej záťaže pre rôzne úrovne vnímanej svietivosti, pričom metóda bez kompenzácie svietivosti chybné reflektovala každú skupinu svietivosti ako samostatnú úroveň kognitívnej záťaže. Kognitívna záťaž úloh rovnakej náročnosti, ale rôznej svietivosti bola odhadnutá pre všetky skupiny v intervale $<0,500; 0,645>$, čo je akceptovateľný rozptyl

pri priemernej veľkosti intervalu istoty na úrovni 0,152. Pre porovnanie, v prípade ignorovania vnímanej svietivosti bol rozptyl v rovnakom vyhodnotení v intervale $<0,206; 0,788>$ pri priemernej veľkosti intervalu istoty len 0,083.

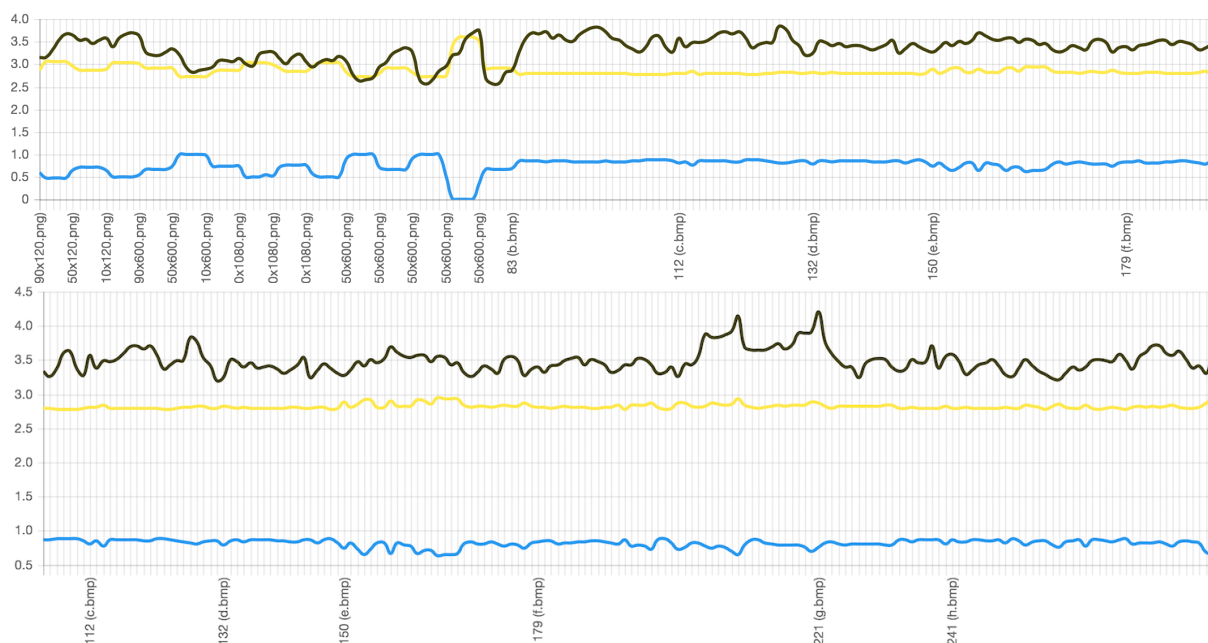
Realizácia softvérovej platformy umožňujúcej použitie metódy pre identifikáciu kognitívnej záťaže v používateľských štúdiách vo forme webovej aplikácie. Navrhnuté súčasti metódy sme paralelne implementovali vo webovej aplikácii, pričom sme identifikovali niekoľko technických optimalizácií pre praktickú uskutočniteľnosť vyhodnocovania záťaže pomocou našej metódy. Zámer sa nám podarilo naplniť (dokumentácia k vybudovanému systému je k dispozícii v prílohe D). V krátkosti demonštrujeme vlastnosti navrhnutého riešenia na identifikácii problémov v skutočnej webovej aplikácii.

Participant dostal za úlohu vyhľadať mobilné telefónne číslo a otváracie hodiny v stredu pre každú prevádzku na portáli pre rozvoj lokálnych podnikov. Experiment sme vykonali pomocou softvéru Tobii Studio a export získaných dát obsahoval sekvenciu šírky zrenice a sekvenciu fixácií participanta počas vykonávania úloh spolu s obrazovkami testovanej aplikácie a označeným časom prechodu na novú stránku. Tento export sme nahrali do systému, definovali kalibračné obrazovky a spustili vyhodnotenie našou metódou s využitím explicitnej kalibrácie.

Cieľom tohto experimentu bolo naznačiť potenciál navrhnutého riešenia pre štúdie použiteľnosti a identifikácii problémov v reálnych webových aplikáciách na základe identifikácie kognitívnej záťaže. Vykonali sme kvalitatívnu analýzu na jednom participantovi. V detaile exportu sledovaného experimentu si môžeme pozrieť priebeh šírky zrenice v priebehu času spolu s vnímanou svietivosťou a odhadnutou šírkou zrenice (pozri obr. 27).

V prvej časti experimentu sme zobrazili participantovi stimuly explicitnej kalibrácie (približne polovica prvej časti grafu). Po počiatočnom rozptýlení participanta, čo sa prejavilo širšou zrenicou pre prvé tri stimuly, sledujeme veľmi presnú predpoveď šírky zrenice, takže môžeme predpokladať správne natrénovanie modelu. Okamžite po zobrazení prvého webu v 83. sekunde je vidieť značný nárast šírky zrenice oproti predpokladanej hodnote, čo znamená, že kognitívna záťaž participanta skutočne vzrástla.

Jedna z podstránok neobsahovala žiadnu poznámku o otváracích hodinách. Táto stránka začína približne v polovici druhej časti grafu v 179. sekunde. Podľa dilatácie zrenice môžeme zhodnotiť, že niekoľko sekúnd participant nevykazoval výnimočne zvýšenú kognitívnu záťaž a pracoval na zadanej úlohe. Približne v polovici úlohy participant identifikoval problém, keďže sa mu podarilo nájsť telefónne číslo, ale nepodarilo sa mu nájsť otváracie hodiny. Tento problém sa prejavil okamžitým nárastom šírky zrenice do najextrémnejších hodnôt v celom experimente, no predpokladaná šírka zrenice si zachovala približne rovnaké hodnoty. Na základe toho môžeme usudzovať, že rozšírenie zrenice spôsobila kognitívna záťaž a nie sledovanie tmavého objektu. V kalibračnej fáze sa vyskytoval rovnaký prudký nárast šírky zrenice, ktorý bol však správne predpokladaný, pretože sa vyskytol pri sledovaní čiernej obrazovky. Zaujímavý je aj okamžitý pokles záťaže na priemerné hodnoty v okamihu, keď sa participant rozhodol pokračovať na ďalšiu úlohu v 221. sekunde.



Obr. 27: Záznam obrazovky z webovej aplikácie na vyhodnocovanie kognitívnej záťaže. Záznam je pre lepšiu čitateľnosť rozdelený na dve časti a zobrazuje skutočnú šírku zrenice (čierna krivka), odhadnutú vnímanú svietivosť (modrá krivka) a predpovedanú šírku zrenice (žltá krivka). Hodnoty na ypsilónovej osy predstavujú šírku zrenice v milimetroch, resp. hodnotu vnímanej svietivosti. Na xovej osy sú poznámky o zmene zobrazenej stránky alebo kalibračného stimulu spolu s časom od začiatku experimentu v sekundách.

8.2 Prínos navrhutej metódy ako celku

Metóda vyhodnocovania kognitívnej záťaže pomocou absolútnej dilatácie zrenice, ktorej vývoj sme začali už v predchádzajúcej práci, má vysoký potenciál použitia v reálnych používateľských štúdiách. Dilatácia zrenice je iba jedným z mnohých parametrov implicitnej používateľskej spätnej väzby, ktorá môže byť použitá samostatne alebo v kombinácii s ďalšími indikátormi kognitívnej záťaže. Našu metódu môžeme chápať nielen ako spôsob na jednorázové vyhodnotenie kognitívnej záťaže, ale môže byť použitá aj vo fáze predspracovania dát v iných metódach.

Metóda deklaruje očistenie sekvencie zaznamenananej šírky zrenice od vplyvu svietivosti pri sledovaní heterogénnych stimulov. Jadro metódy vyžaduje iba dva vstupy: šírku zrenice a zobrazený stimul, pričom metóda môže byť použitá v prúdovom spracovaní dát. Toto umožňuje používať navrhnutú metódu s drobnými úpravami aj na vyhodnocovanie kognitívnej záťaže človeka pracujúceho s počítačom aj v reálnom čase. Vyhodnocovanie kognitívnej záťaže v reálnom čase nám ponúka množstvo ďalších možných využití metódy aj mimo oblasť používateľského testovania.

8.3 Budúca práca

V ďalšej práci sa otvára priestor jednak na zdokonalenie modelu reakcií zrenice a jeho kalibrácie, ako aj pre rozšírenie aplikovateľnosti použitia metódy.

V kalibrácii modelu má metóda čiastočné rezervy pri správnom odhade šírky zrenice pre niektoré úrovne svietivosti. Asi najmenšiu úspešnosť má metóda pri kalibrácii parametru sigma pre vybudovanie matice projekcie sietnice. Po vyriešení týchto dvoch problémov predpokladáme pomerne veľký nárast presnosti, čo umožní používať metódu pri kvantitatívnom vyhodnotení aj s menším počtom participantov.

Ďalšou oblasťou môže byť návrh možností úpravy metódy pre jej použitie vo vyhodnocovaní kognitívnej záťaže v reálnom čase.

Literatúra

- [1] Chen, Siyuan a Julien Epps (2013). “Automatic classification of eye activity for cognitive load measurement with emotion interference”. In: *Computer methods and programs in biomedicine* Vol. 110. No. 2, s. 111–124.
- [2] Chen, Siyuan, Julien Epps a Fang Chen (2011). “A comparison of four methods for cognitive load measurement”. In: *Proceedings of the 23rd Australian Computer-Human Interaction Conference*. ACM, s. 76–79.
- [3] Beatty, Jackson (1982). “Task-evoked pupillary responses, processing load, and the structure of processing resources.” In: *Psychological bulletin* Vol. 91. No. 2, s. 276.
- [4] Bouma, H a LCJ Baghuis (1971). “Hippus of the pupil: Periods of slow oscillations of unknown origin”. In: *Vision Research* Vol. 11. No. 11, s. 1345–1351.
- [5] Ellis, CJ (1981). “The pupillary light reflex in normal subjects”. In: *British Journal of Ophthalmology* Vol. 65. No. 11, s. 754–759.
- [6] Gavas, Rahul, Debatri Chatterjee a Aniruddha Sinha (2017). “Estimation of cognitive load based on the pupil size dilation”. In: *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, s. 1499–1504.
- [7] Haapalainen, Eija, SeungJun Kim, Jodi F Forlizzi a Anind K Dey (2010). “Psycho-physiological measures for assessing cognitive load”. In: *Proceedings of the 12th ACM international conference on Ubiquitous computing*. ACM, s. 301–310.
- [8] Hayes, Taylor R a Alexander A Petrov (2016). “Mapping and correcting the influence of gaze position on pupil size measurements”. In: *Behavior Research Methods* Vol. 48. No. 2, s. 510–527.
- [9] Hess, Eckhard H a James M Polt (1960). “Pupil size as related to interest value of visual stimuli”. In: *Science* Vol. 132. No. 3423, s. 349–350.
- [10] Hollingsworth, Karen, Kevin W Bowyer a Patrick J Flynn (2009). “Pupil dilation degrades iris biometric performance”. In: *Computer vision and image understanding* Vol. 113. No. 1, s. 150–157.
- [11] Ikehara, Curtis S a Martha E Crosby (2005). “Assessing cognitive load with physiological sensors”. In: *Proceedings of the 38th Annual Hawaii International Conference on System Sciences. HICSS’05*. IEEE, 295a–295a.
- [12] Iqbal, Shamsi T, Piotr D Adamczyk, Xianjun Sam Zheng a Brian P Bailey (2005). “Towards an index of opportunity: understanding changes in mental workload during task execution”. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, s. 311–320.
- [13] Jeffries, Robin, James R Miller, Cathleen Wharton a Kathy Uyeda (1991). “User interface evaluation in the real world: a comparison of four techniques”. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, s. 119–124.
- [14] Juhaniak, Tomas (2016). “Kognitívne zaťaženie človeka pri používaní aplikácií”. Bakalárska práca. Ilkovičova 2, 842 16 Bratislava 4: Fakulta informatiky a informačných technológií STU v Bratislave.

- [15] Juhaniak, Tomas, Patrik Hlavac, Robert Moro, Jakub Simko a Maria Bielikova (2016). “Pupillary Response: Removing Screen Luminosity Effects for Clearer Implicit Feedback”. In: *UMAP 2016: Posters, Demos, Late-breaking Results and Workshop Proc. of the 24rd Conf. on User Modeling, Adaptation, and Personalization*, s. 2.
- [16] Klingner, Jeff, Rakshit Kumar a Pat Hanrahan (2008). “Measuring the task-evoked pupillary response with a remote eye tracker”. In: *Proceedings of the 2008 symposium on Eye tracking research & applications*. ACM, s. 69–72.
- [17] Laeng, Bruno, Sylvain Sirois a Gustaf Gredebäck (2012). “Pupillometry a window to the preconscious?” In: *Perspectives on psychological science* Vol. 7. No. 1, s. 18–27.
- [18] Lallé, Sébastien, Cristina Conati a Giuseppe Carenini (2016). “Prediction of individual learning curves across information visualizations”. In: *User Modeling and User-Adapted Interaction* Vol. 26. No. 4, s. 307–345.
- [19] Marshall, Sandra P (2002). “The index of cognitive activity: Measuring cognitive workload”. In: *Proceedings of the 2002 IEEE 7th conference on Human factors and power plants*. IEEE, s. 7–5.
- [20] Martinez-Gómez, Pascual a Akiko Aizawa (2014). “Recognition of understanding level and language skill using measurements of reading behavior”. In: *Proceedings of the 19th international conference on Intelligent User Interfaces*. ACM, s. 95–104.
- [21] Matell, Michael S a Jacob Jacoby (1971). “Is there an optimal number of alternatives for Likert scale items? Study I: Reliability and validity”. In: *Educational and psychological measurement* Vol. 31. No. 3, s. 657–674.
- [22] Mathôt, Sebastiaan a Stefan Van der Stigchel (2015). “New light on the mind’s eye: The pupillary light response as active vision”. In: *Current directions in psychological science* Vol. 24. No. 5, s. 374–378.
- [23] McKnight, Patrick E a Julius Najab (2010). “Mann-Whitney U Test”. In: *Corsini Encyclopedia of Psychology*.
- [24] Partala, Timo a Veikko Surakka (2003). “Pupil size variation as an indication of affective processing”. In: *International journal of human-computer studies* Vol. 59. No. 1, s. 185–198.
- [25] Pomplun, Marc a Sindhura Sunkara (2003). “Pupil dilation as an indicator of cognitive workload in human-computer interaction”. In: *Proceedings of the International Conference on HCI*. Citeseer, s. 542–546.
- [26] Purves, D., G.J. Augustine, D. Fitzpatrick, W.C. Hall, A.-S. LaMantia, J.O. McNamara a S.M. Williams (2004). “Anatomical distribution of rods and cones”. In: *Neuroscience, 3rd edition*. Sinauer Associates, s. 244–245.
- [27] Sauro, Jeff a James R Lewis (2012a). “How Precise Are Our Estimates? Confidence Intervals”. In: *Quantifying the User Experience: Practical Statistics for User Research*. Elsevier, s. 19–40.
- [28] — (2012b). “Is There a Statistical Difference between Designs?” In: *Quantifying the User Experience: Practical Statistics for User Research*. Elsevier, s. 63–103.
- [29] Toker, Dereck, Sébastien Lallé a Cristina Conati (2017). “Pupillometry and Head Distance to the Screen to Predict Skill Acquisition During Information Visualization Tasks”. In: *Proceedings of the 22nd International Conference on Intelligent User Interfaces*. ACM, s. 221–231.

- [30] Wang, Weihong, Zhidong Li, Yang Wang a Fang Chen (2013). “Indexing cognitive workload based on pupillary response under luminance and emotional changes”. In: *Proceedings of the 2013 International Conference on Intelligent user interfaces*. ACM, s. 247–256.
- [31] Wilcoxon, Frank (1945). “Individual comparisons by ranking methods”. In: *Biometrics bulletin* Vol. 1. No. 6, s. 80–83.
- [32] Wong, Hoe Kin (2016). “Developing instantaneous eye-activity based task analysis”. In: *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications*. ACM, s. 323–324.

Príloha A - Obsah priloženého média

1. `webova_aplikacia`
adresár s implementáciou webovej aplikácie na automatické spracovanie experimentov a vyhodnocovanie kognitívnej záťaže
2. `stimuly_experimentu`
adresár stimulov navrhnutých pre explicitnú kalibráciu modelu
3. `dataset`
dataset komplexného autorského experimentu
4. `diplomova_praca.pdf`
text diplomovej práce vo formáte `pdf`
5. `obsah_nosica.pdf`
obsah priloženého média

Príloha B - Detailný opis použitých datasetov

Dataset reakcií zrenice na rôzne svetelné podmienky - D1

Autor: Tomáš Juhaniak

Spoluautori: Členovia výskumnej skupiny UXI

Obdobie vykonania: Bakalárska práca

Rok vykonania: 2016

Téma: Zistenie vlastností reakcií zrenice pri zobrazení rôznych stimulov

Počet participantov: 63

Dataset bol v súčasnej práci opravený. Predmetom opravy boli záznamy desiatich účastníkov, ktoré boli v predchádzajúcej práci vyradené zo štatistického spracovania.

Dataset bol vytvorený pomocou infraštruktúry UXR. 48 účastníkov začalo riešiť, prípadne vyriešilo zjednodušené sudoku po skončení bloku základných obrazoviek a ostatné stimuly boli zobrazené všetkým účastníkom. Poradie obrazoviek bolo rovnaké pre všetkých účastníkov. Sumárne vlastnosti sledovaného bloku v experimente opisuje tabuľka (tab. 5). Tento blok bol vždy prvým alebo jediným v celom sedení bez predchádzajúceho vykonávania experimentálnych úloh participanta.

Tabuľka 5: Základné parametre datasetu D1. Dataset obsahuje stimuly, pri ktorých sa používateľ pozerá do prostriedku monitora. Okrem úlohy riešenia sudoku je predpokladané kognitívne zaťaženie používateľov minimálne.

Stimul / úloha	počet účastníkov	počet zobrazení	poradie stimulu
Centralizovaný kruh 1016px	63	63	1
Centralizovaný kruh 656px	63	63	2
Centralizovaný kruh 376px	63	63	3
Centralizovaný kruh 176px	63	63	4
Centralizovaný kruh 56px	63	63	5
Centralizovaný kruh 16px	63	63	6
Webové sídlo (10x)	63	63	7 - 16
Homogénny stimul #FFFFFF	63	$63 \cdot 4 = 252$	17, 21, 23, 31
Homogénny stimul #AAAAAA	63	$63 \cdot 2 = 126$	18, 20
Homogénny stimul #000000	63	$63 \cdot 2 = 126$	19, 22
Homogénny stimul #FF0000	63	$63 \cdot 2 = 126$	24, 33
Homogénny stimul #00FF00	63	$63 \cdot 2 = 126$	25, 30
Homogénny stimul #0000FF	63	$63 \cdot 2 = 126$	26, 35
Homogénny stimul #00FFFF	63	$63 \cdot 2 = 126$	27, 36
Homogénny stimul #FFFF00	63	$63 \cdot 2 = 126$	28, 32
Homogénny stimul #FF00FF	63	$63 \cdot 2 = 126$	29, 34
Zjednodušené sudoku	48	48	37

Dataset matematických príkladov - D2

Autor: Tomáš Juhaniak

Spoluautori: Členovia výskumnej skupiny UXI

Obdobie vykonania: Bakalárska práca

Rok vykonania: 2016

Téma: Overenie schopnosti navrhnuť metódy rozlíšiť rôzne úrovne kognitívnej záťaže počas zobrazenia stimulov so skutočnými webovými stránkami

Počet participantov: 19

Dataset obsahuje základnú explicitnú kalibráciu, bez kalibračných stimulov pre zistenie vnímanej svietivosti okolia, po ktorej nasledujú inštrukcie a blok matematických príkladov so štyrmi úrovňami obtiažnosti. Matematický blok obsahuje 15 príkladov, prvé tri boli vybrané ako tréningové. Zvyšných 12 príkladov obsahuje po tri príklady pre každú zo štyroch skupín náročnosti. Každá z úloh bola zobrazená na popredí jedného z desiatich stimulov zobrazujúcich webové sídla alebo na jednoduchých stimuloch s tmavošedým alebo svetlošedým pozadím. Poradie úloh bolo rovnaké pre všetkých účastníkov a náročnosť úloh nebola usporiadaná.

Náročnosť úloh bola špecifikovaná podľa počtu čísel v každom zo štyroch čísel, ktoré musel participant sčítavať a pamätať si priebežný výsledok. Čísla v prvej úrovni obsahovali len jednu cifru s binárnymi hodnotami, druhá úroveň používala jednu cifru s dekadickými hodnotami, tretia úroveň dve a najťažšia tretia úroveň obsahovala až trojciferné čísla. Niekoľko participantov považovalo tretiu úroveň za tak náročnú, že sa ju ani nepokúšali riešiť, takže pri nej nemuseli mať významnú kognitívnu záťaž.

Dataset bol vytvorený pomocou infraštruktúry UXR. Okrem vyššie spomenutých úloh obsahuje ešte 4 ďalšie príklady na veľmi svetlých stimuloch, ktoré sa pre rôzne nastavenia monitorov rôznych účastníkov nezobrazili vždy správne a musia byť z vyhodnocovania vyradené.

Sumárne vlastnosti experimentu opisuje tabuľka (tab. 6).

Tabuľka 6: Základné parametre datasetu D2. Dataset obsahuje stimuly, pri ktorých sa používateľ pozerá do prostriedku monitora. Kognitívne zaťaženie je predpokladané najmä pri matematických úlohách, pričom jeho úroveň by mala korelovať s obtiažnosťou úlohy.

Stimul / úloha	počet účastníkov	počet zobrazení	poradie stimulu
Homogénny stimul #FFFFFF	19	$19 \cdot 2 = 38$	1, 6
Homogénny stimul #AAAAAA	19	$19 \cdot 2 = 38$	2, 5
Homogénny stimul #000000	19	19	3
Homogénny stimul #0000FF	19	19	4
Inštrukcie	19	19	7
Trénovacie príklady	19	$19 \cdot 3 = 57$	8 - 10
Úloha - obtiažnosť 0	19	$19 \cdot 3 = 57$	18, 22, 24
Úloha - obtiažnosť 1	19	$19 \cdot 3 = 57$	12, 15, 23
Úloha - obtiažnosť 2	19	$19 \cdot 3 = 57$	13, 17, 19
Úloha - obtiažnosť 3	19	$19 \cdot 3 = 57$	11, 21, 25
Úloha - biele pozadie	19	$19 \cdot 4 = 76$	14, 16, 20, 26

Dataset matematických príkladov a čítania textu - D3

Autor: Tomáš Juhaniak

Spoluautor: Lukáš Rešutík

Obdobie vykonania: Diplomová práca

Rok vykonania: 2017

Téma: Porovnanie detekcie kognitívnej záťaže pri rôznych typoch úloh

Počet participantov: 36

Dataset obsahuje kompletnú explicitnú kalibráciu modelu reakcií zrenice vrátane ôsmych stimulov pre natrénovanie vnímanej svietivosti okolia monitora.

Po kalibrácii nasleduje prvý blok úloh, v ktorých si účastník musel prečítať text vo vyhradenom časovom intervale, po ktorom sa jedno slovo v texte zakrylo. Účastník musel označiť chýbajúce slovo výberom jednej z troch možností alebo neodpovedať, ak si odpovedou nebol istý. Po každej odpovedi sa na obrazovke zobrazilo jedno slovo, ktoré si účastník musel zapamätať a po piatich textových úlohách vybral z 10 možností tie slová ktoré si zapamätal. Textové úlohy sa skladali z jednej trénovacej päťice textov a štyroch náhodne usporiadaných päťíc, kde každá päťica mala rôzne dĺžky textov na prečítanie a zapamätanie.

Posledným blokom boli matematické príklady s rôznou náročnosťou, ktoré boli prerušované zobrazením slov na zapamätanie, podobne ako v predchádzajúcom bloku textových úloh. Blok matematických príkladov začínal sekvenciou trénovacích príkladov, za ktorými nasledovalo päť sekvencií príkladov, pričom každá sekvencia mala odlišnú náročnosť. Poradie úrovni náročnosti sekvencií bolo náhodné.

Dataset bol vytvorený pomocou infraštruktúry UXR a má niekoľko chýbajúcich dát pri väčšine účastníkov. Chýbajúce dáta vznikli chybou nahrávacieho softvéru, pričom dáta prestali byť ukladané od momentu prvého výpadku. Výpadok nastal u každého účastníka v rôznom čase, takže početnosť záznamov v jednotlivých blokoch klesá podľa poradia bloku v scenári experimentu. Kalibračné dáta sú kompletné. Prvý výskyt chyby sa objavil ku koncu druhého bloku u 4 účastníkov. Najmenej dát obsahujú sekvencie matematických príkladov. Druhou chybou sú náhodne chýbajúce dáta označujúce intervaly zobrazenia možností pre označenie zakrytého slova.

Sumárne vlastnosti experimentu opisuje tabuľka (tab. 7).

Tabuľka 7: Základné parametre datasetu D3. Dataset obsahuje stimuly, pri ktorých sa používateľ pozeral do rôznych častí monitora, najmä pri explicitnej kalibrácii a textových úlohách. Dataset priniesol prvý náhľad na vnímanú svietivosť v okrajoch monitora a možnosť porovnať kognitívnu záťaž matematických a textových úloh na popredí skutočných webových stránok.

Stimul / úloha	počet účastníkov	počet zobrazení	poradie stimulu
Homogénny stimul #FFFFFF	36	$2 \cdot 36 = 72$	1, 6
Homogénny stimul #AAAAAA	36	$2 \cdot 36 = 72$	2, 5
Homogénny stimul #000000	36	36	3
Homogénny stimul #0000FF	36	36	4
Biely stimul fixácia v lavo	36	36	7
Biely stimul fixácia v lavo hore	36	36	8
Biely stimul fixácia v hore	36	36	9
Biely stimul fixácia v pravo hore	36	36	10
Biely stimul fixácia v pravo	36	36	11
Biely stimul fixácia v pravo dole	36	36	12
Biely stimul fixácia v dole	36	36	13
Biely stimul fixácia v lavo dole	36	36	14
Text čítanie	36	856	-
Výber chýbajúceho slova	36	448	-
Matematický príklad	36	732	-

Dataset tvorivých úloh s inštrukciami - D4

Autor: Študenti a výskumný pracovníci UK

Spoluautor: Róbert Móra

Obdobie vykonania: Diplomová práca

Rok vykonania: 2017

Téma: Vnímaná svietivosť okolia

Počet participantov: 39

Dataset bol vytvorený v rámci psychologického experimentu zameraného na tvorivosť a obsahuje explicitnú kalibráciu a dáta pre 15 úloh rozdelených do troch blokov a tri inštrukčné obrazovky na začiatku každého bloku. Obsah inštrukcií zaberá veľkú časť obrazovky. Stimuly úloh obsahujú niekoľko kruhovo zoradených predmetov s jedným centrálnym predmetom. Úlohou účastníkov bolo vymyslieť čo najväčší počet možných použití centrálného predmetu v kombinácii s niektorým zo sekundárnych predmetov. Explicitná kalibrácia bola vykonaná po skončení všetkých troch blokov úloh. Poradie blokov bolo náhodné, stimuly vnútri blokov boli zoradené pre všetkých používateľov rovnako.

Dataset bol vytvorený pomocou softvéru Tobii Studio. 38 účastníkov absolvovalo všetky tri bloky úloh, dáta jedného z účastníkov chýbajú pri 9 úlohách a jednom zobrazení inštrukcií.

Sumárne vlastnosti experimentu opisuje tabuľka (tab. 8).

Tabuľka 8: Základné parametre datasetu D4. Dataset obsahuje stimuly, pri ktorých sa používateľ pozerá do rôznych častí monitora. Rozdiely v objektívnej kognitívnej záťaži jednotlivých úloh nie sú očakávané. Dataset poskytuje významný počet účastníkov, ktorí absolvovali explicitnú kalibráciu.

Stimul / úloha	počet účastníkov	počet zobrazení	poradie stimulu
Homogénny stimul #FFFFFF	39	$2 \cdot 39 = 78$	19, 24
Homogénny stimul #AAAAAA	39	$2 \cdot 39 = 78$	20, 23
Homogénny stimul #000000	39	39	21
Homogénny stimul #0000FF	39	39	22
Biely stimul fixácia vľavo	39	39	25
Biely stimul fixácia vľavo hore	39	39	26
Biely stimul fixácia hore	39	39	27
Biely stimul fixácia vpravo hore	39	39	28
Biely stimul fixácia vpravo	39	39	29
Biely stimul fixácia vpravo dole	39	39	30
Biely stimul fixácia dole	39	39	31
Biely stimul fixácia vľavo dole	39	39	32
Tvorivá úloha	38 - 39	$39 \cdot 15 - 9 = 576$	2-6, 8-12, 14-18
Inštrukcie na začiatku bloku	39	$39 \cdot 3 - 1 = 116$	1, 7, 13

Dataset skúmania vlastností sietnice oka - D5

Autor: Dániel Papp

Spoluautori: Členovia výskumnej skupiny UXI

Obdobie vykonania: Diplomová práca

Rok vykonania: 2017

Téma: Skúmanie reakcií zrenice pri zobrazení rovnakého kontrastného prvku v rôznych polohách vzhľadom k bodu fixácie

Počet participantov: 9

Dataset obsahuje kalibračné stimuly pre tréning krivky referenčných hodnôt a špeciálne stimuly pre preskúmanie vlastností sietnice oka. Tieto stimuly obsahujú čierny štvorec na bielom pozadí. Pri každom stimule bol vyznačený fixačný bod v prostriedku obrazovky. Čierny štvorec menil svoju pozíciu okolo tohto fixačného bodu, pričom sa zobrazil nad, pod, vpravo a vľavo od fixačného bodu. Pozície čierneho štvorca sa menili v smere a následne v protismere hodinových ručičiek v dvoch iteráciách. Spolu sa tak každá pozícia čierneho štvorca vyskytla v sekvencii práve štyrikrát. Čierny štvorec bol vždy v rovnakej vzdialenosti od fixačného bodu. Dataset bol vytvorený pomocou softvéru Tobii Studio.

Sumárne vlastnosti experimentu opisuje tabuľka (tab. 9).

Tabuľka 9: Základné parametre datasetu D5. Dataset obsahuje stimuly, pri ktorých sa používateľ pozerá do prostriedku monitora. Dataset obsahuje pozorovanie dilatácie zrenice deviatich účastníkov pri zobrazení špeciálnych stimulov na odhalenie vlastností sietnice oka. Sekvencia zobrazení štvorcov sa vykonala štyrikrát, pričom nepárne iterácie sa zobrazovali v poradí definovanom v tabuľke a párne iterácie sa zobrazovali v opačnom poradí. Stimuly sa zobrazili v usporiadanom poradí ako: s1, s2, s3, s4, s1, s4, s3, s2, s1, s2, s3, s4, s1, s4, s3, s2, kde s1 predstavuje štvorec hore, s2 predstavuje štvorec vpravo, s3 predstavuje štvorec dole a s4 predstavuje štvorec vľavo.

Stimul / úloha	počet účastníkov	počet zobrazení
Inštrukcie	9	9
Čierny štvorec (hore)	9	$9 * 4 = 36$
Čierny štvorec (vpravo)	9	$9 * 4 = 36$
Čierny štvorec (dole)	9	$9 * 4 = 36$
Čierny štvorec (vľavo)	9	$9 * 4 = 36$
Homogénny stimul #FFFFFF	9	$9 * 2 = 18$
Homogénny stimul #AAAAAA	9	$9 * 2 = 18$
Homogénny stimul #0000FF	9	9
Homogénny stimul #000000	9	9

Dataset komplexného overenia navrhnutej metódy - D6

Autor: Tomáš Juhaniak

Spoluautori: Členovia výskumnej skupiny UX1

Obdobie vykonania: Diplomová práca

Rok vykonania: 2018

Téma: Komplexné overenie schopnosti navrhnutej metódy detekovať kognitívnu záťaž pri úlohách rôznej náročnosti a stimuloch s rôznou úrovňou svietivosti

Počet participantov: 21

Dataset obsahuje explicitnú kalibráciu a 72 matematických príkladov v troch úrovniach náročnosti a štyroch úrovniach vnímanej svietivosti. Stimul prvej úrovne svietivosti bola biela homogénna obrazovka. Stimuly v druhej úrovni svietivosti boli tri svetlé webové stránky. Tretiu úroveň tvorili tri webové stránky so stredne tmavým dizajnom a štvrtú úroveň tvorili tri tmavé webové stránky. Na popredí týchto stimulov boli zobrazené matematické príklady v dvoch pozičných verziách. V prvej verzii bolo zadanie príkladu aj výber správnej odpovede uprostred obrazovky. V druhej verzii bolo zadanie príkladu rozmiestnené po okrajoch obrazovky a výber správnej odpovede bol uprostred obrazovky. Všetci účastníci mali zobrazené rovnaké úlohy, ktoré boli náhodne usporiadané pre každého používateľa. Na všetkých stimuloch boli zobrazené príklady každej úrovne v oboch verziách rozmiestnenia. Na bielej obrazovke sa zobrazil až trojnásobok príkladov v záujme zachovania počtu príkladov pre každú úroveň svietivosti. Matematickým príkladom predchádzali tri tréningové príklady.

Webové stránky na pozadí úloh boli skutočné rozhrania z bankovej, mediálnej a voľnočasovej komerčnej oblasti. Webové stránky boli upravené náhodným preusporiadaním malých fragmentov tak, aby nebolo možné čítať ich obsah, čo by mohlo spôsobiť rozptýlenie participanta a neželanú kognitívnu záťaž. Preusporiadanie fragmentov nenarušilo celkovú kompozíciu stránky, pretože jednotlivé fragmenty ostali v tesnej blízkosti svojej pôvodnej polohy. Dataset bol vytvorený pomocou softvéru Tobii Studio.

Sumárne vlastnosti experimentu opisuje tabuľka (tab. 10).

Tabuľka 10: Základné parametre datasetu D6. Dataset obsahuje stimuly, pri ktorých sa participant pozerali do rôznych častí monitora. Homogénny stimul s bielou farbou, ktorý obyčajne trénujeme pomocou dvoch zobrazení sa prvý krát vyskytol uprostred a druhý krát na konci kalibrácie svietivosti okolia.

Stimul / úloha	počet účastníkov	počet zobrazení
Biely stimul fixácia vľavo hore	21	21
Biely stimul fixácia hore	21	21
Biely stimul fixácia vpravo hore	21	21
Biely stimul fixácia vľavo	21	21
Biely stimul fixácia uprostred	21	$21 * 2 = 42$
Biely stimul fixácia vľavo dole	21	21
Biely stimul fixácia dole	21	21
Biely stimul fixácia vpravo dole	21	21
Homogénny stimul #AAAAAA	21	$21 * 2 = 42$
Homogénny stimul #0000FF	21	21
Homogénny stimul #000000	21	21
Matematická úloha	21	$21 * 72 = 1512$

Príloha C - Etapy riešenia projektu

Postup prác nášho výskumu je rozdelený do troch etáp podľa príslušných semestrov, ktorých plán uvádzame v nasledujúcich tabuľkách. Plán je prezentovaný vo verzii pôvodného zámeru a ku každej časti plánu pripájame zhodnotenie do akej miery sa ho podarilo naplniť.

Prvý semester riešenia

Tabuľka 11: Plán prác v prvom semestri riešenia. Vytvorené: február 2017.

Týždeň semestra	Úloha	Rozšírené zadanie
1 - 4	Hlbšia analýza relevantných súčastí	Vykonať hlbšiu analýzu tých oblastí súčasného stavu poznania, ktoré sú relevantné k existujúcemu návrhu modelu reakcií zreničky a iné prístupy k vyhodnocovaniu kognitívnej záťaže pomocou absolútnej dilatácie zreničky
2 - 5	Analýza modelu reakcií zreničky	Nájsť hlavné nedostatky predchádzajúceho návrhu, zistiť nejasnosti a navrhnúť postup opráv
4 - 7	Implementácia jadra webovej aplikácie	Navrhnuť a implementovať webovú aplikáciu schopnú automatizovať doterajšie schopnosti metódy použitia modelu reakcií zreničky
6 - 8	Príprava experimentu	Pripraviť experiment s kognitívnymi úlohami zameraný na riešenie zistených nedostatkov
6 - 11	Písanie diplomovej práce	Vytvoriť prvú verziu textu diplomovej práce na základe preštudovaných vlastností a navrhnutého postupu vylepšení
8 - 11	Analýza štatistického vyhodnotenia	Analyzovať možnosti štatistického vyhodnotenia experimentov rovnakého typu
9 - 11	Vykonanie experimentov	Vykonať pripravené experimenty s čo najväčším možným počtom účastníkov a zozbierať získané dáta
11 - 12	Finalizácia a zhodnotenie	Finalizovať prvú verziu textu diplomovej práce a skontrolovať získaný dataset

V prvej etape sme nadviazali na výstup predchádzajúcej práce. Vykonali sme hlbšiu analýzu oblastí v relevantných častiach nášho návrhu. Začali sme vyvíjať webovú aplikáciu na vyhodnocovanie experimentov, vykonali sme prvý experiment a analyzovali možnosti štatistického vyhodnotenia experimentov zisťujúcich kognitívnu záťaž pomocou našej metódy.

Stanovený plán sa nám podarilo naplniť s drobnými odchýlkami v čase vykonávania niektorých činností. Implementáciu webovej aplikácie sme začali už v druhom semestri a priebežne sme ju vy-

lepšovali až do začiatku finalizácie dokumentu v jedenástom týždni semestra. Príprava experimentu trvala až do jeho vykonania v desiatom týždni.

Druhý semester riešenia

Tabuľka 12: Plán prác v druhom semestri riešenia. Vytvorené: apríl 2017.

Týždeň semestra	Úloha	Rozšírené zadanie
1 - 2	Vytvoriť experimentálne stimuly	Vytvoriť nové prepracované experimentálne stimuly
2 - 5	Implementácia štatistického vyhodnotenia	Implementovať automatizované štatistické vyhodnocovanie experimentov podľa predchádzajúceho návrhu
2 - 6	Formálna definícia modelu do textu DP	Formálne prepracovať navrhnutý model, kalibráciu a metódu použitia vrátane ďalšieho návrhu vylepšení priamo do textu diplomovej práce
3 - 8	Návrh alternatívnej kalibrácie	Navrhnuť odstránenie problémov s existujúcou kalibráciou, novým spôsobom kalibrácie
3 - 12	Iniciácia experimentov	Iniciovať čo najväčšie možné zaradenie nových experimentálnych stimulov do experimentov vykonávaných na fakulte
6 - 10	Implementácia alternatívnej kalibrácie	Implementovať navrhnuté riešenia súčasne navrhovanej alternatívnej kalibrácie
7 - 8	Písanie textu o implicitnej kalibrácii	Formálne opísať návrh alternatívnej kalibrácie do textu DP
9	Stanovenie hypotéz	Stanoviť hypotézy pre navrhnuté vylepšenia a celkový koncept vyhodnocovania kognitívnej záťaže metódou absolútnej dilatácie zreničky
9 - 11	Vyhodnocovanie experimentov	Vyhodnotiť získané datasety a overenie časti stanovených hypotéz
11	Návrh dodatočných vylepšení	Navrhnuť riešenia prípadných nedostatkov zistených v overení
12	Finalizácia dokumentu	Finalizovať druhú verziu textu diplomovej práce spolu s čiastočným vyhodnotením experimentov a zostavenie plánu pre poslednú záverečnú etapu

V druhej etape sme pripravili prepracované stimuly pre experimenty. Získali sme dva nové datasety a iniciovali ďalšie experimenty s neskoršou dobou ukončenia. Navrhli sme nový spôsob kalibrácie. Skompletizovali sme formálnu definíciu modelu reakcií zrenice a všetkých jej súčastí a vylepšení. Upravili sme webovú aplikáciu tak, aby bola schopná pracovať s niekoľkými novými návrhmi vylepšenia. Definovali sme výskumné hypotézy. Analyzovali a opravili sme získané datasety. Vykonali sme vyhodnotenie štyroch datasetov v kontexte niektorých zo stanovených hypotéz.

Stanovený plán sme splnili s miernym posunutím začiatku plánovaných experimentov až do šiesteho týždňa. Štatistické vyhodnotenie sme implementovali v priebehu druhého týždňa v časovom predstihu.

Tretí semester riešenia

Tabuľka 13: Plán prác v treťom semestri riešenia. Vytvorené: december 2017.

Týždeň semestra	Úloha	Rozšírené zadanie
1 - 2	Opravenie krivky referenčných hodnôt	Zmeniť tréning krivky referenčných hodnôt z lineárnej na kvadratickú
1 - 9	Iniciácia experimentov	Iniciovať čo najväčšie možné zaradenie experimentálnych stimulov do experimentov vykonávaných na fakulte
2 - 3	Analýza dostupných datasetov	Zozbierať všetky externe vykonávané experimenty obsahujúce vhodné dáta pre naše hypotézy a analýza chýbajúcich vlastností v datasetoch
3 - 4	Príprava záverečných experimentov	Pripraviť potrebné experimenty pre spoľahlivé overenie stanovených hypotéz
3 - 6	Implementácia chýbajúcich vlastností	Implementovať všetky chýbajúce časti formálneho návrhu a pripraviť aplikáciu na efektívne a finálne vyhodnotenie hypotéz
5 - 7	Analýza iných možností detekcie záťaže	Analyzovať iné možnosti detekcie kognitívnej záťaže s porovnateľnou mierou intruzívnosti
5 - 9	Vykonávanie vlastných experimentov	Vykonať potrebné experimenty pre doplnenie chýbajúcich datasetov
7 - 10	Záverečné vyhodnotenia experimentov	Vykonať záverečné overenie všetkých hypotéz a zostaviť vyhodnotenie
8 - 10	Porovnanie metódy	Vytvoriť záverečné porovnanie našej metódy s inými možnosťami vyhodnocovania kognitívnej záťaže ako súčasti používateľských štúdií
11 - 12	Finalizácia dokumentu	Finalizovať a zdokonaľiť všetky výstupy našej práce

V poslednej etape sme vykonali samostatný experiment pre overenie hypotéz, ktoré sa nám nepodarilo overiť pomocou existujúcich datasetov. Prvou úlohou bolo opravenie krivky referenčných hodnôt v modeli reakcií zrenice. Implementovali sme všetky chýbajúce súčasti návrhu, aby bolo možné overiť stanovené hypotézy. Vykonali sme záverečné overenie stanovených hypotéz a overili sme úspešnosť nášho riešenia. Naše riešenie sme porovnali s alternatívou bez uvažovania vnímanej svietivosti stimulov. Dokončili sme navrhnutú webovú aplikáciu a predstavili ju ostatným výskumníkom na fakulte.

Plán sme dokázali naplniť s miernym časovým natiahnutím niektorých úloh, pri ktorých sa vyskytli ďalšie zistené nedostatky. Opravenie krivky referenčných hodnôt prinieslo problém citlivejšieho tréningu, ktoré sa nám podarilo vylepšiť v ôsmom týždni semestra. Iniciácia zaradenia stimulov do experimentov na fakulte priniesla v tejto etape už iba jeden kvalitný neautorský dataset. Implementácia chýbajúcich vlastností trvala spolu s vylepšením tréningu krivky referenčných hodnôt do ôsmeho semestra. Drobné opravovanie chýb pokračovalo až do desiateho týždňa.

Príloha D - Navrhnutý softvérový produkt

Pre maximálnu jednoduchosť použitia našej metódy na fakulte sme vyvinuli webovú aplikáciu pre kompletný manažment nahrávania, spracovávania a vyhodnotenia experimentov. V tejto prílohe predstavujeme softvérový produkt určený výskumníkom používateľského zážitku.

Aplikácia je naprogramovaná v jazyku Python 2.7 s použitím rámca Django a umožňuje používateľovi vytvárať a spravovať projekty s prislúchajúcimi datasetmi experimentov. Na vykonávanie náročných úloh spracovania je použitý Celery worker, vďaka ktorému je spracovanie viaczáramových experimentov paralelizovateľné. Pridelovanie a udržiavanie fronty úloh má na starosti Rabbit MQ. Perzistenciu dát zabezpečuje relačná databáza Postgres. Efektívne spracovanie videa zabezpečuje modul OpenCV. Na generovanie štatistických grafov používame knižnicu Matplotlib a na samotnú štatistickú analýzu používame knižnicu ScipyStats. Pre jednoduché úlohy v rozhraní používame prvú verziu Angular-u.

Cieľom navrhnutého riešenia je poskytnúť výskumníkom parametrizovateľný vyhodnocovač kognitívnej záťaže v experimentoch pomocou modelu reakcií zrenice bez nutnosti vlastnej implementácie. Veľký dôraz kladíme aj na informatívnu spätnú väzbu systému v prípade nesprávneho dátového vstupu, s cieľom jasne identifikovať problémy v datasetoch, s ktorými sme sa počas výskumnej práce stretli.

Vstupné dáta

Po vytvorení projektu, ktorý má združovať spolu súvisiace dátové exporty, je možné do systému nahrávať dáta exportované pomocou softvéru Tobii Studio alebo infraštruktúrou UXR. Aplikácia podporuje hromadné nahrávanie exportovaných súborov. Podľa zostavenia scenára experimentu rozlíšujeme dve skupiny vstupných údajov.

Dataset založený na zázname obrazovky. Z priebehu práce používateľa v používateľskej štúdii sa okrem okulometrických dát môžu uchovávať aj iné záznamy. Záznam obrazovky je pre účel vyhodnotenia vnímanej svietivosti ideálnym vstupom, pretože dokonale uchováva svetelné podmienky v každom časovom okamihu, vrátane najmenších detailov, ako sú animácie a iné dynamické správanie rozhrania.

Vstupnými dátami pre vyhodnotenie kognitívnej záťaže jediného účastníka sú v tomto prípade záznam obrazovky a prislúchajúci záznam okulometra obsahujúci šírky zreničiek, body fixácie a časovú pečiatku každej položky záznamu. V prípade potreby je možné samostatným súborom označiť intervaly úloh, pre ktoré sa má vypočítať priemerná kognitívna záťaž. Aplikácia umožňuje nahrávanie experimentov ako zip archív jednotlivých súborov pre každého účastníka zvlášť.

Odporúčame nahrávať záznam obrazovky v nízkom rozlíšení a s nízkou frekvenciou záznamov obrazovky za sekundu. 5 obrázkov za sekundu je dostatočne veľká frekvencia na úspešné predpovedanie vnímanej svietivosti.

Dataset založený na stimuloch projektu. Dataset založený na zázname obrazovky má nevýhodu vo veľkom objeme uchovávaných dát. Pokiaľ vykonávame experiment na ručne pripravených stimuloch alebo v aplikácii, ktorá je schopná zasielať informácie o vykonávaných akciách do programu spracúvajúceho okulometrický záznam, môžeme namiesto videozáznamu použiť sadu stimulov predstavujúcich rozhranie pre jednotlivé úlohy. Rovnakú sadu stimulov, resp. jej podmnožinu, vidia v jednom experimente viacerí participanti, takže nie je potrebné uchovávať grafické dáta pre jednotlivých participantov zvlášť.

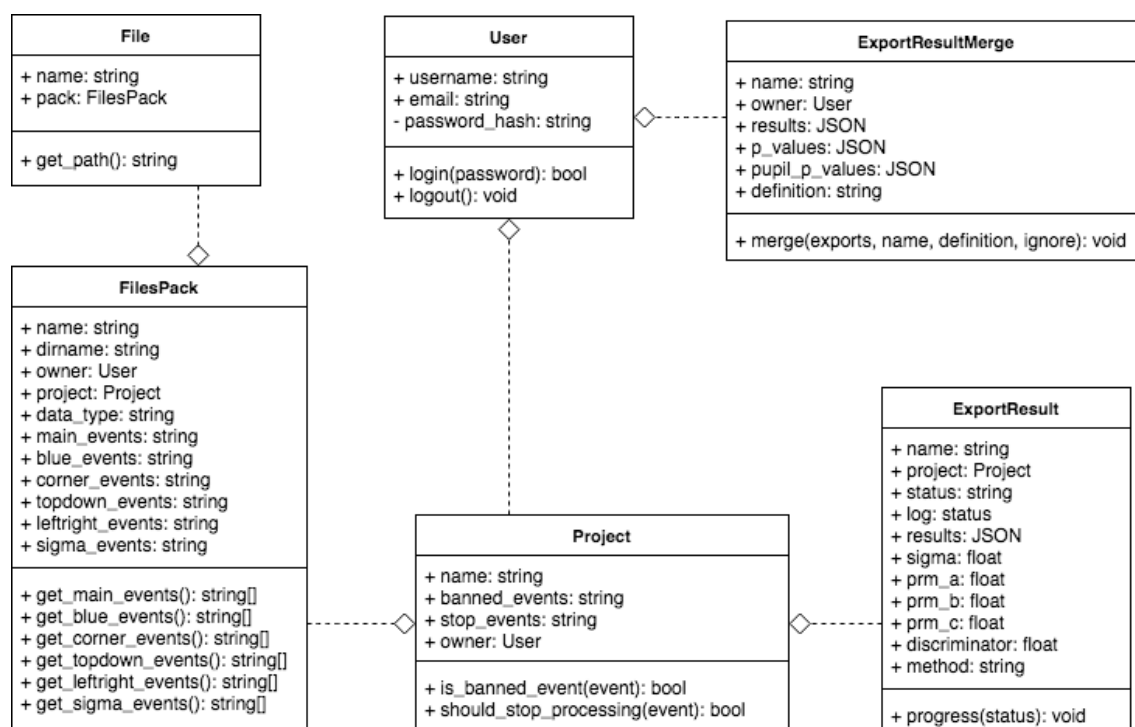
Vstupnými dátami pre vyhodnotenie kognitívnej záťaže jediného participanta je v tomto prípade iba záznam okulometra obsahujúci šírky zreničiek, body fixácie a aktuálne zobrazený identifikátor stimulu a vyžiadaná podmnožina stimulov, ktorá je do systému nahraná iba jedenkrát pre celý projekt. Aplikácia umožňuje nahrávanie experimentov ako zip archív sady experimentov od viacerých používateľov a zip archív sady stimulov pre prislúchajúci projekt.

Technické detaily

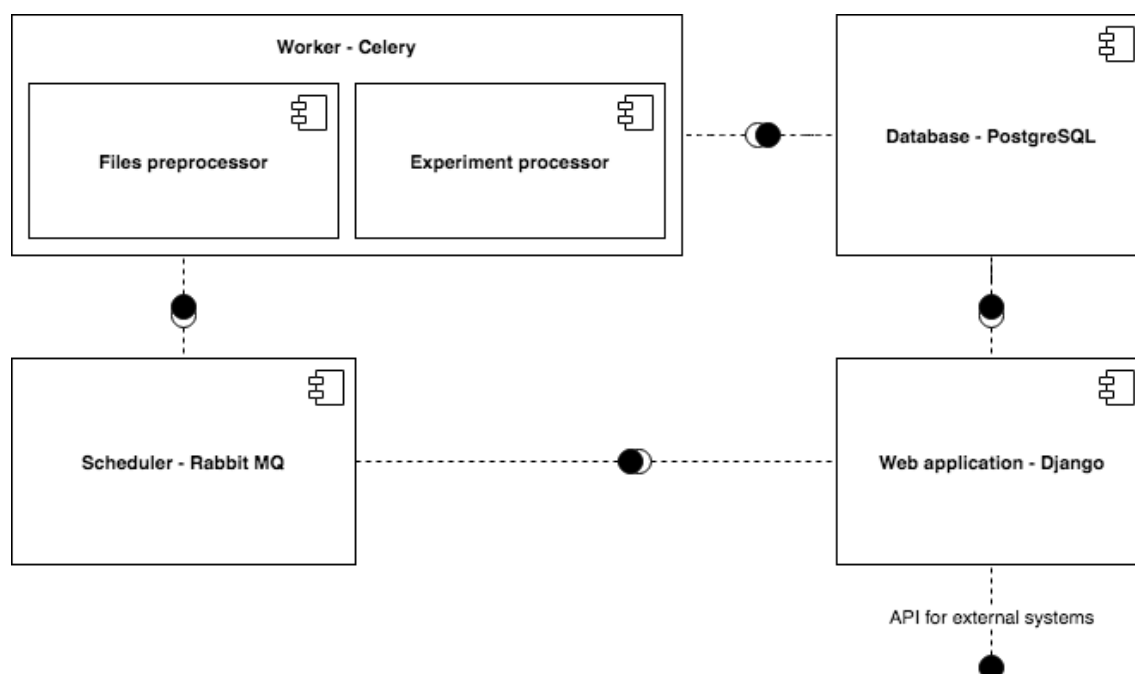
Navrhnutý softvér operuje nad niekoľkými entitami, ktoré udržujú všetky nastavenia projektov a výsledky vyprodukované systémom. Najvýznamnejšie triedy systému opisuje nasledujúci diagram (obr. 28). Ústrednou triedou je používateľ, ku ktorému sa viaže celý obsah v systéme. Používateľ môže vytvárať ľubovoľné množstvo projektov. Každý projekt sa skladá z viacerých záznamov experimentov (FilesPack) a niekoľkých exportov (ExportResult). Keďže jeden záznam experimentu sa môže skladať z viacerých súborov, systém reprezentuje záznam experimentu pomocou zhukovacej triedy, ktorá zabezpečuje prístup k súborovému systému a organizáciu množstva súborov do logických štruktúr. Export, ako výsledok spracovania jediného záznamu experimentu, môže existovať aj po zmazaní záznamu experimentu, z ktorého bol vytvorený. Z tohto dôvodu sa exporty vedú ako samostatné entity patriace priamo pod projekt. Kvantitatívne vyhodnotenie reprezentuje trieda ExportResultMerge. Táto trieda nespadá pod konkrétny projekt, pretože môže byť vytvorená z exportov viacerých projektov.

Celý systém tvoria štyri hlavné komponenty (obr. 29). Základným komponentom je jadro webovej aplikácie, ktorá má na starosti interakciu s používateľom, spravovanie obsahu a kontrolu stavu. Ďalším komponentom je vykonávač exportov. Vykonávačov môže byť viacero a v prípade, že vykonávač nemá aktívnu úlohu, môže prijať nový príkaz na vykonanie exportu. Tento komponent združuje dva podkomponenty, prvý má na starosti načítanie a predspracovanie okulometrických súborov do pamäte, nad ktorými druhý komponent môže vykonať analýzu a samotné spracovanie. Pridelovanie úloh na spracovanie má na starosti plánovač, ktorý prijíma požiadavky od webovej aplikácie a podľa dostupnosti vykonávačov rozdeľuje úlohy. Vykonávač zapisuje všetky svoje výsledky do databázy, z ktorej webová aplikácia čerpá informácie o stave spracovania a následne pripravené výsledky.

Zdrojový kód je organizovaný podľa štandardov rámca Django. Aplikácia sa skladá z troch adresárov. Adresár *core* obsahuje zdrojový kód logiky spracovania experimentov. Adresár *management* obsahuje kód zabezpečujúci biznis logiku. Adresár *uxjour* obsahuje základné nastavenia aplikácie ako celku a nastavenia vzájomnej komunikácie komponentov. Najzaujímavejšie časti kódu sa nachádzajú v adresári *core*.



Obr. 28: Diagram tried navrhnutého systému.



Obr. 29: Diagram komponentov navrhnutého systému.

V nasledujúcej ukážke zdrojového kódu (zdrojový kód 6) prezentujeme logiku predpovedania šírky zrenice na základe svietivosti a natrénovaného modelu.

Zdrojový kód 6: Predikovanie šírky zrenice na základe vnímanej svietivosti a času. Predikcia reflektuje aj také detaily, ako je maximálna rýchlosť dilatácie zrenice pri náhlej zmene svietivosti.

```

1  if milliseconds is None:
2      return self.a*(lum**2) + self.b*lum + self.c
3
4  predicted = self.a*(lum**2) + self.b*lum + self.c
5
6  if self.last_diameter_time[sigma]:
7      if predicted < self.last_diameter[sigma] - settings.CONSTRICTION_MM_PER_MSEC*(
8          ↪ milliseconds - self.last_diameter_time[sigma]):
9          self.last_diameter[sigma] = self.last_diameter[sigma] - settings.
10             ↪ CONSTRICTION_MM_PER_MSEC*(milliseconds - self.last_diameter_time[
11             ↪ sigma])
12     elif predicted > self.last_diameter[sigma] + settings.DILATION_MM_PER_MSEC*(
13         ↪ milliseconds - self.last_diameter_time[sigma]):
14         self.last_diameter[sigma] = self.last_diameter[sigma] + settings.
15             ↪ DILATION_MM_PER_MSEC*(milliseconds - self.last_diameter_time[sigma])
16     else:
17         self.last_diameter[sigma] = predicted
18 else:
19     self.last_diameter[sigma] = predicted
20
21 self.last_diameter_time[sigma] = milliseconds
22 return self.last_diameter[sigma]

```

Najkritickejší výpočet, ktorý zaberá približne 75% času spracovávania experimentu je vypočítanie vnímanej svietivosti na základe zobrazeného stimulu a bodu fixácie. Nasledujúca ukážka zdrojového kódu (zdrojový kód 7) zobrazuje optimalizovaný zápis krátkej metódy na výpočet vnímanej svietivosti pre jednu časovú jednotku.

Zdrojový kód 7: Predpovedanie vnímanej svietivosti je jedna z najdôležitejších súčastí navrhutej metódy. Tento výpočet je extrémne zaťažujúci, takže jeho vykonávanie z väčšej časti zabezpečujú štandardné funkcie jazyka Python.

```

1  def compute_total_luminosity(self, image, x, y, sigma, frame_id):
2      # function returns sum of perceived luminosity of each pixel
3      key = str(x) + "_" + str(y) + "_" + str(sigma)
4      if not key in self.kernels_cache:
5          self.kernels_cache[key] = self.kernel_factory(x, y, sigma)
6
7      return (
8          (map(lambda row: map(self.color, row), image)*self.kernels_cache[key]).sum
9          ↪ ()/255.0
10         + self.outside_lum(self.kernels_cache[key], sigma)
11     )

```

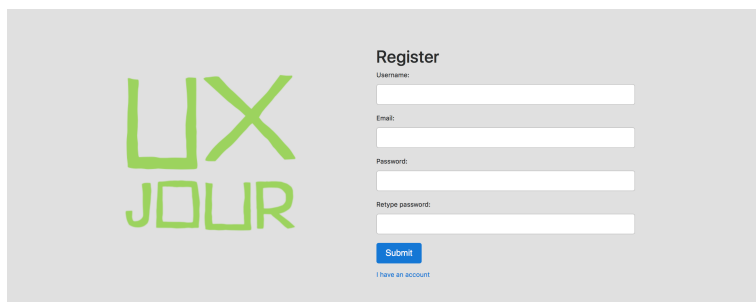
Inštalačná príručka

Inštalačná príručka predpokladá predchádzajúcu inštaláciu jazyka **Python 2.7** a manažéra Python balíkov **pip**. Inštaláciu systému odporúčame vykonať do vyhradeného **pip prostredia**. Návod je primárne určený pre os **Linux**.

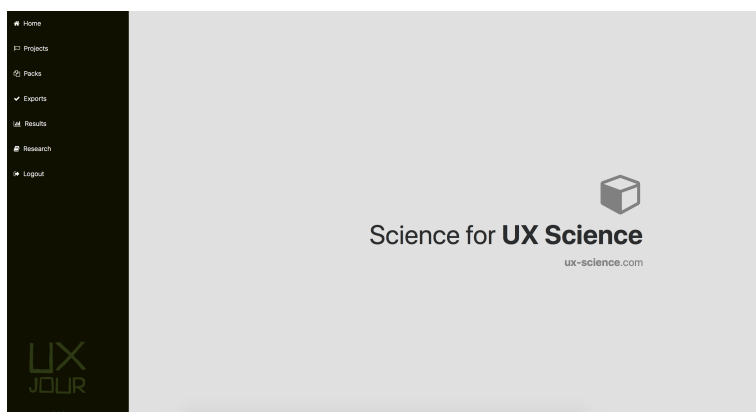
1. Nakopírujte adresár **webova_aplikacia/uxjour** z priloženého média do adresára, v ktorom chcete mať inštaláciu aplikácie
2. Prejdite do hlavnej zložky aplikácie: **cd uxjour/uxjour**
3. Nainštalujte potrebné knižnice: **pip install -r requirements.txt**
4. Nainštalujte OpenCV: **sudo apt-get install python-opencv** (alebo obdobným príkazom v inom operačnom systéme)
5. Nainštalujte plánovač: **sudo apt-get install rabbitmq-server** (alebo obdobným príkazom v inom operačnom systéme)
6. Spustite plánovač: **rabbitmq-server**
7. Spustite vykonávač: **celery -A uxjour worker --loglevel=info**
8. Spustite server: **python manage.py runserver**
9. Otvorte prehliadač na adrese: **http://localhost:8000**

Používateľská príručka

Pomocou navrhnutého systému dokážeme v niekoľkých krokoch analyzovať surové exportované súbory a získať prehľadné kvalitatívne a kvantitatívne vyhodnotenie. Prierez postupnosti prípadov použitia ilustrujeme v nasledujúcich krokoch.



Obr. 30: Dáta každého používateľa systému sú súkromné a pred vstupom do aplikácie je potrebné sa zaregistrovať / prihlásiť.



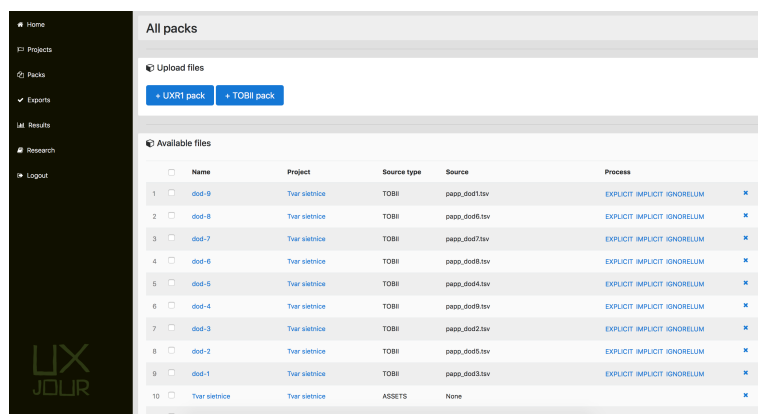
Obr. 31: Po prihlásení sa zobrazí základná obrazovka. V ľavej časti stránky sa nachádza vždy viditeľné menu. Menu obsahuje základné časti systému: domovská stránka, projekty, súbory, exporty, výsledky, relevantné publikácie a odhlásenie.

Project	owner	files	exports
1 Tvar sietečnice	tomas	10	9
2 Kruhy	tomas	40	39
3 Experiment DP1	tomas	36	36
4 Colors	tomas	63	189
5 Math	tomas	19	19
6 Hlavný experiment	tomas	22	63
7 Usage demonstration	tomas	2	1

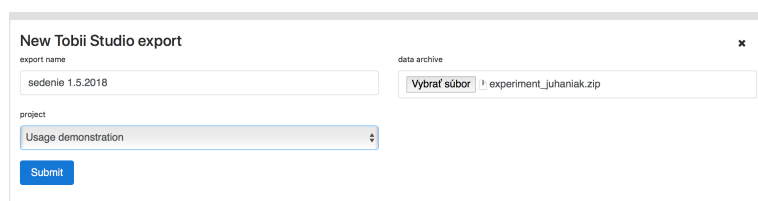
Obr. 32: Najskôr klikneme na položku Projekty, ktoré nás dostane na zoznam všetkých projektov používateľa. Vytváranie nového projektu začína stlačením tlačidla "Pridať nový projekt".

Obr. 33: Pri vytváraní projektu zadáme názov a dôležité poznámky o projekte. Tiež je potrebné vybrať účel projektu (vedecký, edukatívny a komerčný).

Obr. 34: Po vytvorení projektu môžeme pokračovať jeho detailnejším nastavením po kliknutí na jeho meno. V nastavení zadávame riadiace pokyny pre spracovanie exportov v projekte. Môžeme nastaviť, ktoré úlohy sa nemajú vyhodnocovať alebo pri ktorých úlohách zastaviť vyhodnocovanie úplne. Tiež môžeme v súborovom systéme vybrať obrázky jednotlivých úloh, ktoré nahráme ako zip zložku súborov pomenovaných rovnako ako sú pomenované udalosti v zázname okulometra. Tieto súbory nahrávame len v prípade, že sa v projekte nachádzajú súbory bez videozáznamov obrázok.



Obr. 35: Ak máme vytvorený projekt, môžeme do systému nahráť exportované súbory z experimentov, ktoré prislúchajú tomuto projektu. V menu vyberieme položku "Súbory". Na zobrazenej stránke klikneme na tlačidlo pridania nového projektu podľa formátu exportu (UXR alebo Tobii).



Obr. 36: Pri nahrávaní súborov zadáme meno skupiny súborov a v súborovom systéme vyberieme zip súbor so všetkými okulometrickými záznamami. Jednotlivé súbory systém identifikuje a vytvorí pre každý súbor jeden záznam v systéme. Prefix v mene jednotlivých záznamov je meno skupiny súborov. V tejto fáze tiež vyberáme príslušnosť súboru k projektu, ktorý sa po vytvorení nachádza v dostupných možnostiach.

 **Set calibration events**

Basic calibration events

black
gray
white

Blue color events

blue

Sigma calibration events

square1
square2
square3
square4

Corner calibration events

corner1
corner2
corner3
corner4

Obr. 37: Na spodku zoznamu súborov sa nachádza formulár, ktorým môžeme hromadne identifikovať kalibračné stimuly v jednotlivých súboroch. Samostatne označujeme skupinu stimulov pre explicitnú kalibráciu krivky referenčných hodnôt, kalibráciu vnímania modrej farby, kalibráciu parametra sigma matice projekcie sietnice a tiež stimuly s fixačnými bodmi v okrajoch obrazovky.

Top or down calibration events

topfixation
bottomfixation

Left or Right calibration events

leftfixation
rightfixation

Append events

Reset events

Obr. 38: Po zadaní kalibračných pravidiel ich uložíme úplným prepísaním alebo doplnením k predchádzajúcim nastaveniam. Tieto nastavenia sa aplikujú len na označené súbory.

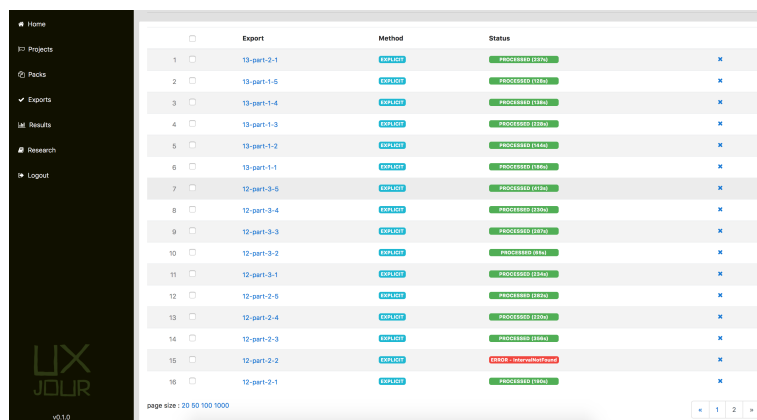
Batch actions

EXPLICIT

IMPLICIT

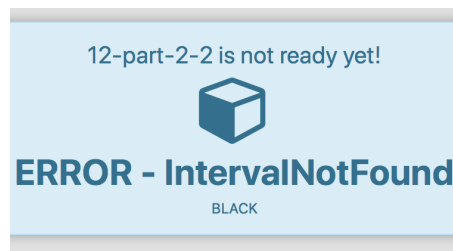
IGNORELUM

Obr. 39: Správnym vyplnením pravidiel sme splnili všetky podmienky pre úspešný export a všetky označené súbory môžeme hromadne spracovať pomocou jednej z metód. Klasické vyhodnotenie spustíme výberom možnosti Explicit.

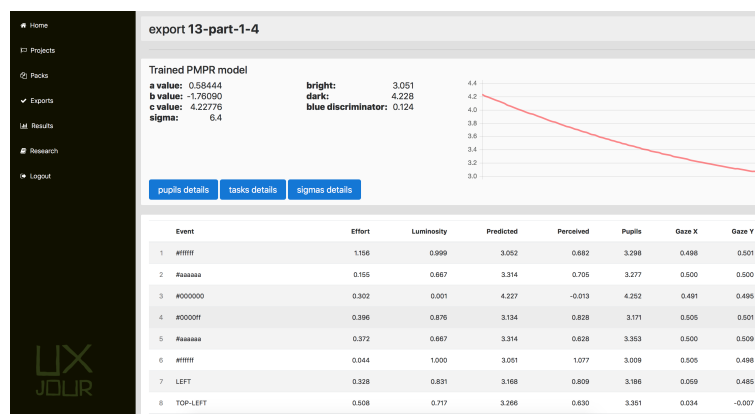


	Export	Method	Status
1	13-part-3-1	EXPLICIT	PROCESSED (OK)
2	13-part-1-6	EXPLICIT	PROCESSED (OK)
3	13-part-1-4	EXPLICIT	PROCESSED (OK)
4	13-part-1-3	EXPLICIT	PROCESSED (OK)
5	13-part-1-2	EXPLICIT	PROCESSED (OK)
6	13-part-1-1	EXPLICIT	PROCESSED (OK)
7	12-part-3-6	EXPLICIT	PROCESSED (OK)
8	12-part-3-4	EXPLICIT	PROCESSED (OK)
9	12-part-3-3	EXPLICIT	PROCESSED (OK)
10	12-part-3-2	EXPLICIT	PROCESSED (OK)
11	12-part-3-1	EXPLICIT	PROCESSED (OK)
12	12-part-2-6	EXPLICIT	PROCESSED (OK)
13	12-part-2-4	EXPLICIT	PROCESSED (OK)
14	12-part-2-3	EXPLICIT	PROCESSED (OK)
15	12-part-2-2	EXPLICIT	ERROR - IntervalNotFound
16	12-part-2-1	EXPLICIT	PROCESSED (OK)

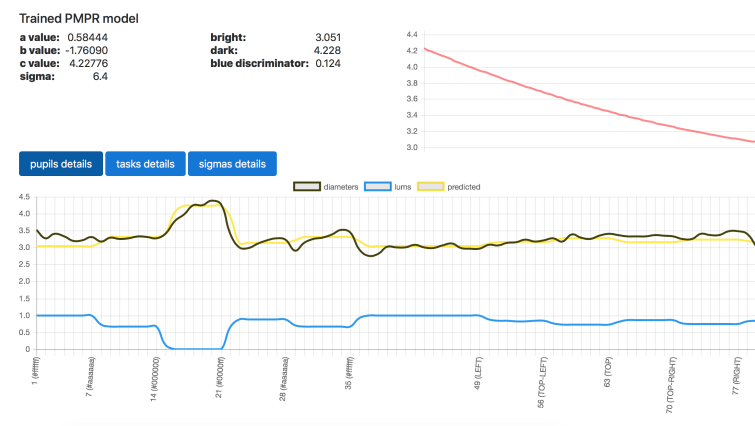
Obr. 40: Prehľad exportovaných súborov nájdeme po výbere položky Exporty v hlavnom menu. Pri jednotlivých exportoch vidíme stav spracovávaní, prípadne výsledok s akým sa spracovávanie skončilo. Ak spracovanie skončilo chybou, status exportu je označený červenou farbou a názvom chyby.



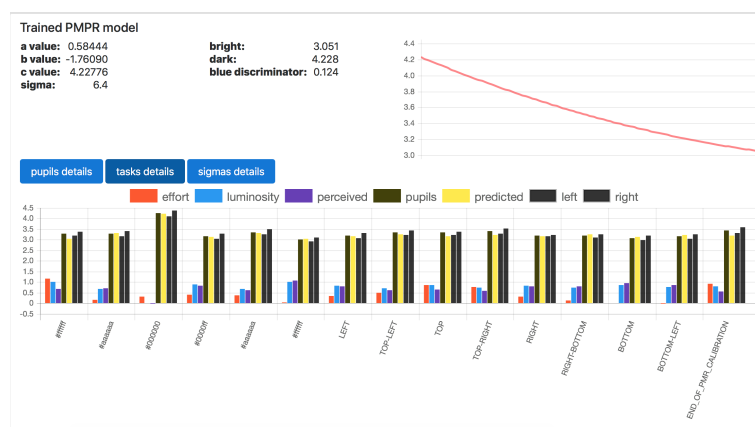
Obr. 41: Pri otvorení neúspešne spracovaného exportu sa nám zobrazí detail chyby. V tomto prípade systém nemohol nájsť čiernu obrazovku v približne definovanom čase podľa súboru okulometrických dát, čím nás upozorňuje na chybu vo vstupných dátach.



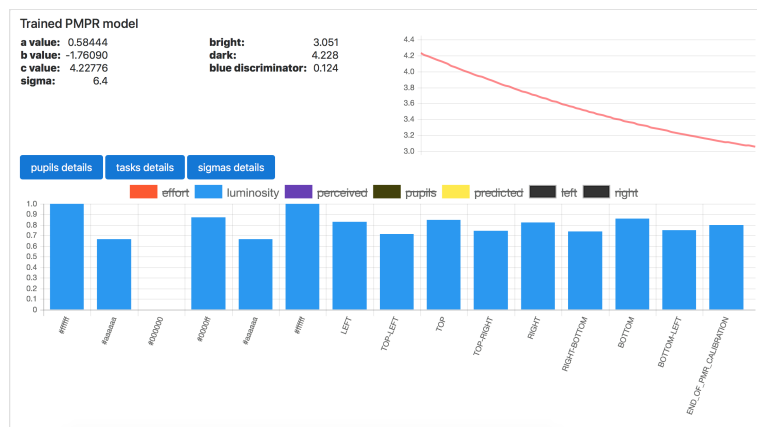
Obr. 42: Ak sa export podaril, v jeho detaile nájdeme všetky dôležité informácie o kalibrácii modelu reakcií zrenice. V pravo hore môžeme vidieť ilustráciu krivky referenčných hodnôt používateľa.



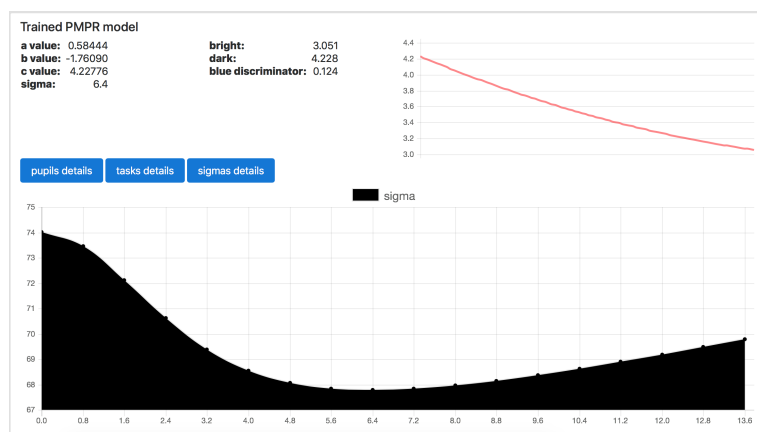
Obr. 43: Systém umožňuje náhľad na skutočnú a predpokladanú šírku zrenice v priebehu času, pomocou ktorej môžeme spraviť kvalitatívnu analýzu experimentu.



Obr. 44: Náhľad na sumárne metriky jednotlivých úloh.



Obr. 45: Náhľad na sumárnu metriku vnímanej svetivosti osobitne.



Obr. 46: Náhľad na detail kalibrácie matice projekcie siete.

	Event	Effort	Luminosity	Predicted	Perceived	Pupils	Gaze X	Gaze Y
1	#fffff	1.156	0.999	3.052	0.682	3.298	0.498	0.501
2	#aaaaa	0.155	0.667	3.314	0.705	3.277	0.500	0.500
3	#00000	0.302	0.001	4.227	-0.013	4.252	0.491	0.495
4	#000ff	0.396	0.876	3.134	0.828	3.171	0.505	0.501
5	#aaaaa	0.372	0.667	3.314	0.628	3.353	0.500	0.509
6	#ffffff	0.044	1.000	3.051	1.077	3.009	0.505	0.498
7	LEFT	0.328	0.831	3.168	0.809	3.188	0.059	0.485
8	TOP-LEFT	0.508	0.717	3.266	0.630	3.351	0.034	-0.007
9	TOP	0.850	0.851	3.152	0.643	3.338	0.507	0.061
10	TOP-RIGHT	0.754	0.747	3.239	0.572	3.412	0.949	0.050
11	RIGHT	0.298	0.823	3.174	0.810	3.185	0.954	0.492
12	RIGHT-BOTTOM	0.117	0.737	3.248	0.792	3.200	0.961	0.961
13	BOTTOM	0.005	0.860	3.146	0.948	3.084	0.503	0.958
14	BOTTOM-LEFT	-0.011	0.752	3.234	0.849	3.154	0.058	0.950
15	END_OF_PMR_CALIBRATION	0.933	0.798	3.203	0.547	3.439	0.200	0.058

Obr. 47: Každý export umožňuje aj detailný prehľad pomocou tabuľky najdôležitejších metrík pre každú úlohu. V tejto tabuľke nájdeme aj najviac sledovanú oblasť obrazovky, ktorú používateľ pri vykonávaní úlohy sledoval.

	Export	Project	Method	Status
1	math-part-7-3	Math	EXPORT	PROCESSED (376x)
2	math-part-7-2	Math	EXPORT	PROCESSED (368x)
3	math-part-7-1	Math	EXPORT	PROCESSED (333x)
4	math-part-6-3	Math	EXPORT	PROCESSED (316x)
5	math-part-6-2	Math	EXPORT	PROCESSED (316x)
6	math-part-6-1	Math	EXPORT	PROCESSED (347x)
7	math-part-5-3	Math	EXPORT	PROCESSED (361x)
8	math-part-5-2	Math	EXPORT	PROCESSED (358x)
9	math-part-5-1	Math	EXPORT	PROCESSED (321x)
10	math-part-4-3	Math	EXPORT	PROCESSED (346x)
11	math-part-4-2	Math	EXPORT	PROCESSED (333x)
12	math-part-4-1	Math	EXPORT	PROCESSED (344x)
13	math-part-3-3	Math	EXPORT	PROCESSED (342x)
14	math-part-3-2	Math	EXPORT	PROCESSED (378x)
15	math-part-3-1	Math	EXPORT	PROCESSED (304x)
16	math-part-2-3	Math	EXPORT	ERROR - too many exports
17	math-part-2-2	Math	EXPORT	PROCESSED (388x)

Obr. 48: V zozname exportov môžeme označiť tie exporty, ktoré chceme vyhodnotiť kvantitatívne. Tieto exporty môžu patriť aj do viacerých projektov.

Merge selected exports

Delete selected exports

Matematicke ulohy BC 4 urovne

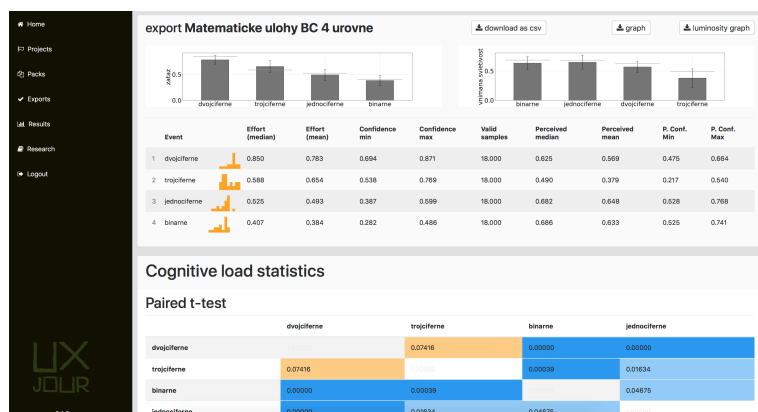
binarne, lvl_0
 jednociferno, lvl_1
 dvojciferno, lvl_2
 trojciferno, lvl_3

☒ Ignore other events

Delete exports

Merge exports

Obr. 49: Na spodku zoznamu exportov sa nachádza formulár pre definovanie kvantitatívneho vyhodnotenia. Tento formulár umožňuje zadať do jedného riadku všetky názvy úloh oddelené čiarkou, ktoré majú byť zaradené do jednej skupiny. Systém podporuje regexový zápis. Skupina úloh bude pomenovaná podľa prvej hodnoty v riadku, aj keď sa v exportoch nenachádza ani jedna takto nazvaná úloha. Týmto spôsobom môžeme zmysluplne pomenovať skupiny úloh. Systém tiež umožňuje vybrať, či sa majú z vyhodnocovania vyradiť úlohy, ktoré sme explicitne nedefinovali. Následne klikneme na tlačidlo "Spojiť exporty". Systém automaticky vyberie skupinu exportov, ktoré obsahujú najväčší počet rovnakých úloh. Táto funkcionality zabezpečuje korektné štatistické vyhodnotenie.



Obr. 50: Po vytvorení kvantitatívneho vyhodnotenia nás systém presmeruje na zoznam vyhodnotení. V detaile vyhodnotenia sa nachádza graf štatistickej analýzy prežitej kognitívnej záťaže pri špecifikovaných úlohách a tiež graf vnímanej svietivosti jednotlivých úloh. Tieto grafy je možné stiahnuť vo formáte pdf. Vyhodnotenie je dostupné aj pomocou prehľadnej tabuľky najdôležitejších metrík. Túto tabuľku je navyše možné stiahnuť vo formáte CSV. Vyhodnotenie tiež obsahuje hĺbkovú analýzu štatistickej signifikancie rozdielu v kognitívnej záťaži a vnímanej svietivosti medzi jednotlivými úlohami. Vypočítané p-hodnoty sú dostupné pre štyri rôzne metódy štatistickej analýzy. Výber správnej metódy je závislý od charakteru dát a experimentu. Používateľovi je tiež umožnený náhľad na rozdelenie hodnôt kognitívnej záťaže v jednotlivých úlohách pomocou histogramu zobrazeného v každom riadku tabuľky.

Cognitive load statistics				
Paired t-test				
	dvojciferna	trojciferna	binarna	jednociferna
dvojciferna		0.07416	0.00000	0.00000
trojciferna	0.07416		0.00039	0.01634
binarna	0.00000	0.00039		0.04675
jednociferna	0.00000	0.01634	0.04675	
Unpaired t-test				
	dvojciferna	trojciferna	binarna	jednociferna
dvojciferna		0.07068	0.00000	0.00009
trojciferna	0.07068		0.00076	0.03749
binarna	0.00000	0.00076		0.12717
jednociferna	0.00009	0.03749	0.12717	
Wilcoxon test				
	dvojciferna	trojciferna	binarna	jednociferna
dvojciferna		0.04286	0.00020	0.00020
trojciferna	0.04286		0.00063	0.01762

Obr. 51: Vyhodnotenie štatistickej analýzy je dostupné v maticovej vizualizácii.

Wilcoxon test				
	dvojčiferné	trojčiferné	binárne	jednociferné
dvojčiferné		0.00185	0.07071	0.00649
trojčiferné	0.00185	1.00000	0.00028	0.00046
binárne	0.07071	0.00028		0.52773
jednociferné	0.00649	0.00046	0.52773	

Mann-Whitney U test				
	dvojčiferné	trojčiferné	binárne	jednociferné
dvojčiferné		0.03444	0.09459	0.06644
trojčiferné	0.03444	1.00000	0.00375	0.00310
binárne	0.09459	0.00375		0.44339
jednociferné	0.06644	0.00310	0.44339	

binárne, lvl_0
jednociferné, lvl_1
dvojčiferné, lvl_2
trojčiferné, lvl_3

Obr. 52: Každé vyhodnotenie obsahuje definíciu skupín, ktorú je možné skopírovať pre vytváranie ďalších vyhodnotení.

Príloha E - Príspevok na konferenciu IIT.SRC 2018

Pripájame plné znenie článku vychádzajúceho z tejto diplomovej práce, ktorý sme publikovali na študentskej vedeckej konferencii IIT.SRC 2018.



Obr. 53: Ocenenie za najlepší príspevok prezentovaný na IIT.SRC 2018 - Informatics and Information Technologies Student Research Conference v kategórii: Inžinierske štúdium.

Automatic Detection of Cognitive Load from Pupil Dilation in Real world Scenarios

Tomáš JUHANIÁK*

*Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
juhaniak.tomas@gmail.com*

Abstract. We propose a method for evaluating the cognitive load in user studies based on pupil dilation. With reference to demonstrated properties of pupil in the field of cognitive psychology, we proposed a method based on absolute pupil dilation considering the context of presented stimuli in real world scenario, e.g. in a UX user study working with an application dynamically changing its user interface. The idea behind our method is to filter out the effect of perceived luminosity using information about displayed user interface and the fixation point of the user's gaze. After filtering the dilation caused by perceived luminosity, we consider residual dilation as a cognitive load.

1 Introduction

User-based studies aimed to detecting problems present in user interfaces can be addressed by various explicit and implicit feedback methods. Basically, problems encountered are reflected in the task completion, prolonging execution time, nervousness, or cognitive load. All these properties can be evaluated with certain accuracy automatically and without explicit feedback. Implicitly obtained feedback does not limit participant's workflow and allows put himself into simulated situation. The cognitive load, as one of the most important metrics in user testing, can be obtained almost without notice of the participant with an eye-tracker and a pupil dilation.

The pupil dilates in case of increased cognitive load and the dilation intensity correlates with the level of cognitive load. The widening of the pupil also occurs in other situations, like increase of the emotional experience or the light reduction perceived on the eye's retina.

In fact, perceived luminosity is constantly changing. The participant constantly changes his view in the

interface with uneven luminosity, which also varies from task to task. This feature prevents the direct determination of cognitive load only by measurements of pupil width, because its value is not the product of a single source.

Our goal is to model the response of the pupil to the light conditions and to predict the width of the pupil at every time of the user study. Before we can determine the value of the perceived luminosity, we need to know what the brightness of the environment is, what interface is displayed on the monitor and what is the point of fixation on the screen.

This paper presents proposed method for cognitive load detection with use of eye-tracker data, the results of experiments on perceived luminosity under different luminosity conditions and cognitive load detection on tasks of varying difficulty and different backgrounds.

2 Related work

In the last hundred years, people have been trying to figure out the formula of the relationship between the lightness and the width of pupil. In fact, there exist a number of factors shaping the reaction of the pupil to the luminosity changes. In [8] the author presents a unified formula for light adapted pupil size. The most important knowledge is the ability to represent the relationship between light conditions and individual's pupil width only by quadratic function.

Evaluating the cognitive load by eye-tracker dilation is nothing new. In works like [9], [2], [4], [7] the authors successfully detected cognitive load in artificial and very simple stimuli. It is proven that the eye-tracker can reliably detect changes in the width of pupil caused by cognitive load, even in several levels proportional to the intensity of the cognitive load.

* Master study programme in field: Information Systems

Supervisor: Professor Mária Bielíková, Institute of Informatics, Information Systems and Software Engineering, Faculty of Informatics and Information Technologies STU in Bratislava

Pupil dilation is present in various types of tasks requiring significant mental activity. In [1] pupil dilation amplitude in different types of tasks is compared. Different types of tasks cause different amplitudes.

The two most important factors influencing the change in pupil width are always present for user application testing. For this reason, it is not easy to determine whether dilation of the pupil has occurred due to a change in luminosity or a change in cognitive load. This problem can be simplified by using homogeneous stimuli background [6] or by using training stimuli before each task [7].

Since the response of the pupil to the change of luminosity and the occurrence of the cognitive load have a different course, it is possible to separate the cognitive-based changes, calculate number of these changes per second and determine cognitive activity as is done in [5].

Limitation of existing approaches is that they do not consider context of real web applications (changing screen luminosity and also the fixation point of the user's gaze). In our first work we proposed first version of pupillary response model for detecting cognitive load which filters out the screen luminosity [3]. Previously proposed method is able to predict pupil diameter only for fixations in the middle of the screen. In this paper we further extend the model and improve its calibration.

3 Pupillary response modelling

We proposed a model, which calculates the luminosity of individual pixels on the screen considering data on the currently displayed stimulus. Subsequently, using the current viewpoint data, we determine the effect of individual pixels on the perceived luminosity as well as the contribution of the luminosity of the environment to the overall perceived luminosity.

After calculating the perceived luminosity, we compare the measured pupil diameter with the expected one predicted by the model. If the predicted value is lower, additional dilation is caused by mental activity. In this way, we divide the reactions of the pupil into changes in luminosity and cognitive load.

3.1 Definition of the model

We define pupillary response model (PRM) as quaternion (its elements are described in further text):

$$PRM = (QRC, RPM, ELum, UCP) \quad (1)$$

Quadratic Reference Curve (QRC) defines pupil reactions for all possible luminosity levels as

$$diameter = ax^2 + bx + c. \quad (2)$$

where a , b and c define trained parameters for quadratic curve and x is perceived luminosity from interval $[0,1]$ (see Fig. 1).

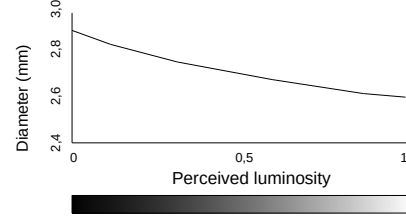


Figure 1. Example of QRC - a personalized curve mapping the luminosity values emitted by the monitor to specific pupil's diameters.

Retinal Projection Matrix (RPM) is a 2D Gauss matrix with mean corresponding to fixation point and with numerically optimized variance parameter σ . Matrix must have the same size as the screen resolution. This matrix represents visual perception of stimulus as it is projected on the retina (see Fig. 2).

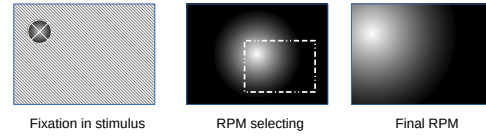


Figure 2. Generating a matrix for concrete fixation point. Matrix showed in the middle is double sized base matrix. Sum of values in centralized subset of the base matrix has to be equal to 1.

Environment Luminosity (ELum) of the monitor's background is dependent on laboratory conditions, so it is included in the calibration characteristics of the model. Environment luminosity is defined as vector of additional luminosity values calibrated at edges and corners of monitor (see Fig. 3).

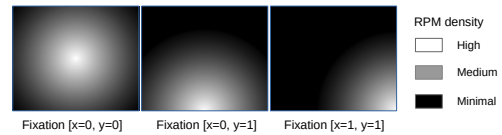


Figure 3. If the fixation is located at the edge of the monitor, the retinal projection is concentrated to a greater extent on the surroundings, so we can not neglect the luminosity of the environment.

Unique Color Perception (UCP) defines perceived luminosity of the RGB color as

$$L_{rgb} = a + b + c + g \quad (3)$$

$$a = \max(0, b - \max(r, g) * L_b) \quad (4)$$

$$b = \max(0, (b - \max(0, b - r) - g) * (\frac{1}{2} + \frac{L_b}{2})) \quad (5)$$

$$c = \max(0, r - \max(g, b)) \quad (6)$$

where r , g and b are components of RGB color and L_b is calibrated value of perceived luminosity for the blue color.

3.2 Explicit calibration procedure

Since we calibrate a reference curve for each individual, we do not need to use a complicated formula dependent on a number of parameters, including externally determined luminosity of the monitor and the environment.

During the calibration the participant looks at the mark in the screen, five seconds for each stimulus. The first second serves to adapt the pupil to the new light conditions and for the remaining four seconds we measure the average diameter of the pupil. There are three types of stimuli in the calibration procedure:

1. Homogeneous stimuli with calibration colors in the following order: white, gray, black, blue, gray and white. Fixation mark is in the middle of monitor.
2. White centered circles with black background with the following diameters: 1016px, 656px, 376px, 176px, 56px and 16px. Fixation mark is in the middle of monitor.
3. White stimulus. Fixation marks are on the edge of monitor. Fixation mark must be distant from the edge by 10% of the width or height of the monitor with following order: top, top right, right, bottom right, bottom, bottom left, left, top left.

Quadratic reference curve QRC is trained in quadratic regression. Train data for regression are records from first sequence, where diameters for white, gray and black colors were measured. White color has luminosity equals 1, black has 0 and gray is depended on the choice of specific RGB color - e.g. #aaaaaa is 0.66. Calibrated diameter for blue color is used for UCP calculation.

Numeric optimization for variance parameter σ of RPM is based on the absolute error of calculated luminosities for screens in the second sequence. For each stimulus in the second sequence pupil diameter is recorded which corresponds to some level of the luminosity in trained QRC. Screen luminosity is calculated as sum of all matrix values which are placed inside white circle area in equally large stimulus.

Third sequence serves for ELum calibration. Since the stimuli in this sequence are white, perceived luminosity should be equal to sum of values in particular retina's projection matrix. Perceived luminosity

calculated from quadratic curve with use of calibrated pupil's diameter will be higher than sum of values in matrix. The difference in the observed and predicted value is the luminosity of the environment at a particular calibration point.

3.3 Implicit calibration

In order to increase effectiveness of the whole procedure and reduce intrusiveness in the user study scenario, we proposed an implicit calibration procedure that does not require any additional calibration stimuli. The main idea is to calibrate the PRM directly on the data from main part of the experiment. In this way, we predict the behavior of the pupil in state of the average cognitive load, but we are still able to identify which tasks were difficult and simple.

4 Evaluation

Until now, we have collaborated on several experiments, based on which we obtained 6 datasets for total cardinality up to 260 participants. Most of the participants were students, men and women aged 19-26.

As we expected, we found, that RPM's parameter σ for different experiment participants can vary. It has been confirmed that the numerical optimization error function has a single local minimum and is convex. Moreover, we found, that is possible to calibrate σ with use of all stimuli in the experiment and there is no need to use special ones.

The same 63 people were shown screens of nine different rgb colors. From the values of white, gray and black we trained a linear reference curve and calculated the relative luminosity of individual colors according to the diameter of the pupils. Results are in Tab. 1. We can notice that the yellow color emitted by the LCD monitor caused a significantly higher perceived luminosity than trained range.

Table 1. RGB colors perception.

RGB	median	mean	95% t-interval
ff0000	0,994	1,010	0,983 - 1,038
00ff00	0,987	0,988	0,960 - 1,015
0000ff	0,838	0,798	0,764 - 0,831
ffff00	1,014	1,026	1,005 - 1,048
ff00ff	0,918	0,901	0,874 - 0,929
00ffff	0,975	0,959	0,930 - 0,987
aaaaaa	0,788	0,777	0,757 - 0,798

Further 75 participants were shown the calibration stimuli from the second calibration sequence for training the ELum vector. We have confirmed that the perceived relative luminosity of the monitor is smaller at the edges of the monitor and the smallest in the corners (see Tab. 2).

Table 2. Perceived luminosity in the edges and corners of the monitor.

fixation	median	mean	95% t-interval
middle	0,909	0,908	0,877 - 0,938
edges	0,860	0,848	0,824 - 0,871
corners	0,822	0,806	0,781 - 0,831
both	0,846	0,827	0,810 - 0,844

19 participants were given the task of counting the sums of four consecutive numbers displayed in the middle of the screen at the forefront of a randomly selected real website. Tasks were categorized into four difficulty groups according to the number of digits in each number and each participant calculated three examples from each level. The sequence of tasks was shuffled and each participant calculated three test tasks before the main sequence of examples.

As expected, the cognitive load for simple tasks was very low, indicating that the diameter for specific luminosity conditions was predicted correctly. The cognitive load recorded for difficult tasks is significantly higher, because predicted pupil diameter for a particular light condition should be about 20 percent of dilation range narrower (see Tab. 3).

Table 3. Predicted cognitive load for simple and difficult tasks with 114 records for each group.

difficulty	median	mean	95% t-interval
simple	0,017	0,028	-0,007 - 0,065
difficult	0,202	0,285	0,220 - 0,350

36 participants were given the task of calculating mathematical examples of different difficulty. Later they read texts of different lengths, in which a word was hidden after a while. Then the participants had to select missing word by click to one of the four provided. As expected, while the reading tasks, significantly lower cognitive load have been detected in comparison with the math tasks. This result matches findings in [1], but our method in addition deals with problematic luminosity interference. Also, the choices of selecting the correct word were a bit easier than mathematical tasks, but this effect is not significantly lower (see Tab. 4).

Table 4. The cognitive load for different task's types.

task	median	mean	95% t-interval
math	0,193	0,232	0,208 - 0,256
selecting	0,159	0,216	0,191 - 0,241
reading	0,044	0,088	0,070 - 0,105

4.1 Conclusion

Until today, we have collected lot of data to uncover the behavior of the various features of the proposed model. We have found that the pupil reactions to the different colors of the LCD monitor has to be calculated according to a special algorithm and no known formulas of color luminosity can be used. Yellow color emitted by the LCD monitor may cause stronger constriction than white color.

We have proven, that the environment luminosity significantly influences pupil reactions and that this effect is most pronounced at the corners of the monitor. We demonstrated the ability to detect cognitive load in stimuli with complex color backgrounds. Also, we evaluated some of the obtained datasets, by which we demonstrate model's capabilities. Our plan for near future is to realize a comprehensive experiment attempting to reveal the prediction accuracy under different conditions.

References

- [1] Beatty, J.: Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological bulletin*, 1982, vol. 91, no. 2, p. 276.
- [2] Chen, S., Epps, J.: Automatic classification of eye activity for cognitive load measurement with emotion interference. *Computer methods and programs in biomedicine*, 2013, vol. 110, no. 2, pp. 111–124.
- [3] Juhaniak, T., Hlavac, P., Moro, R., Simko, J., Bielikova, M.: Pupillary Response: Removing Screen Luminosity Effects for Clearer Implicit Feedback. In: *UMAP 2016 Extended Proc. of the 24rd Conf. on User Modeling, Adaptation, and Personalization*, 2016, p. 2.
- [4] Laeng, B., Sirois, S., Gredebäck, G.: Pupillometry a window to the preconscious? *Perspectives on psychological science*, 2012, vol. 7, no. 1, pp. 18–27.
- [5] Marshall, S.P.: The index of cognitive activity: Measuring cognitive workload. In: *Human factors and power plants, 2002. proceedings of the 2002 IEEE 7th conference on*, IEEE, 2002, pp. 7–5.
- [6] Pomplun, M., Sunkara, S.: Pupil dilation as an indicator of cognitive workload in human-computer interaction. In: *Proceedings of the International Conference on HCI*, Citeseer, 2003, pp. 542–546.
- [7] Wang, W., Li, Z., Wang, Y., Chen, F.: Indexing cognitive workload based on pupillary response under luminance and emotional changes. In: *Proceedings of the 2013 international conference on Intelligent user interfaces*, ACM, 2013, pp. 247–256.
- [8] Watson, A.B., Yellott, J.I.: A unified formula for light-adapted pupil size. *Journal of vision*, 2012, vol. 12, no. 10, pp. 12–12.
- [9] Xu, J., Wang, Y., Chen, F., Choi, E.: Pupillary response based cognitive workload measurement under luminance changes. In: *Human-Computer Interaction-INTERACT 2011*. Springer, 2011, pp. 178–185.