



BUDAPESTI MŰSZAKI ÉS GAZDASÁGTUDOMÁNYI EGYETEM
VILLAMOSMÉRNÖKI ÉS INFORMATIKAI KAR
IRÁNYÍTÁSTECHNIKA ÉS INFORMATIKA TANSZÉK

Tekintetkövető rendszer fejlesztése képfeldolgozási alapokon

DIPLOMATERV

Készítette
Kovács Balázs

Konzulens
Kertész Zsolt

2012. május 10.

Tartalomjegyzék

Tartalomjegyzék	I
Hallgatói nyilatkozat	III
Kivonat	IV
Abstract	V
Bevezető	1
1. Felhasználási lehetőségek	2
1.1. Kutatási felhasználás	2
1.1.1. Kognitív pszichológia	2
1.1.2. Érzelemdetektálás	2
1.1.3. Orvosi felhasználás	2
1.2. Gyakorlati felhasználás	3
1.2.1. Webergonómia	3
1.2.2. Vezetésbiztonság	3
2. A tekintetkövetésről	4
2.1. Tekintetkövetési módszerek	4
2.1.1. Az ideális tekintetkövető	4
2.1.2. Optikai elvű követés	6
2.1.3. Elektromos potenciál-alapú követés	8
2.1.4. Követés speciális kontaktlencse használatával	9
2.1.5. A tekintetkövetési módszerek összehasonlítása	9
3. Szemmozgások	10
3.1. Osztályozás	10
3.1.1. Szakkádok	10
3.1.2. Lassú követések	11
3.1.3. Fixációk	12
3.1.4. Nystagmus	12
3.2. Összegzés	12
4. Hardverek	13
4.1. Optikai megfontolások	13
5. Megvalósítás	15

5.1. Az OpenCV könyvtár	15
5.2. Pszichológiai bemutató	18
5.3. Webergonómiai bemutató	20
Irodalomjegyzék	21

HALLGATÓI NYILATKOZAT

Alulírott *Kovács Balázs*, szigorló hallgató kijelentem, hogy ezt a diplomatervet meg nem engedett segítség nélkül, saját magam készítettem, csak a megadott forrásokat (szakirodalom, eszközök stb.) használtam fel. Minden olyan részt, melyet szó szerint, vagy azonos értelemben, de átfogalmazva más forrásból átvettem, egyértelműen, a forrás megadásával megjelöltem.

Hozzájárulok, hogy a jelen munkám alapadatait (szerző(k), cím, angol és magyar nyelvű tartalmi kivonat, készítés éve, konzulens(ek) neve) a BME VIK nyilvánosan hozzáférhető elektronikus formában, a munka teljes szövegét pedig az egyetem belső hálóján keresztül (vagy autentikált felhasználók számára) közzétegye. Kijelentem, hogy a benyújtott munka és annak elektronikus verziója megegyezik. Dékáni engedéllyel titkosított diplomatervek esetén a dolgozat szövege csak 3 év eltelte után válik hozzáférhetővé.

Budapest, 2012. május 10.

Kovács Balázs
hallgató

Kivonat

Jelen dokumentum egy a Budapesti Műszaki és Gazdaságtudományi Egyetem Irányítástechnika és Informatika Tanszékén készített diplomaterv *„Tekintetkövető rendszer fejlesztése képfeldolgozási alapokon”* témában.

Abstract

The current paper is a thesis from Department of Control Engineering and Information Technology at Budapest University of Technology and Economics in the topic of „*Development of a Gaze Tracking System Based on Image Processing*”.

Bevezető

Hová merült el szép szemed világa?

Mi az, mit kétes távolban keres?

...

Tekintetkövető rendszer fejlesztése összetett feladat. A probléma mind magas, mind alacsony absztrakciós szintről megközelítve megfelelő szakmai felkészültséget követel meg: a rendszer megtervezéséhez, összeállításához és implementációjához olyan önálló mérnöki munka szükséges, ami a ideálissá tette számomra a feladatot diplomatervem témájaként. Nem mellesleg az elkészült rendszer remélhetőleg sok hasznos kutatási vagy ipari felhasználásban bizonyulhat hasznosnak.

Dolgozatom első fejezeteiben főként a téma szakirodalmát, valamint a szóba jöhető elméleti témákat tekintem át. A dolgozat második felében az előzetes megfontolások alapján elkészített rendszer felépítését dokumentálom, majd a mérési eredmények analízisével értékelem az elkészült munkát.

1. fejezet

Felhasználási lehetőségek

A tekintet megfelelő minőségű és robusztus követésének számos gyakorlati felhasználása lehetséges. Elég csak a *perceptuális* (észlelési), vagy *kognitív* (megértési) kutatási területekre gondolni, ahol például az olvasás, vagy az alvás folyamatának vizsgálatánál bizonyulhat hasznosnak.

1.1. Kutatási felhasználás

1.1.1. Kognitív pszichológia

Olvasási folyamatok terén a szakkádok gyors és pontos követése válhat lehetővé, amely a kognitív folyamatok elemzésében komoly segítséget nyújthat. Alvásvizsgálat terén leginkább – nevéből adódóan – a REM (Rapid Eye Movement) fázis kötődik a szemmozgás követéséhez, ebben az alkalmazásban azonban értelemszerűen optikai elvű követés nem jöhet szóba.

1.1.2. Érzelemdetektálás

A pupillaátmérő nem csak a fénymennyiség-változás hatására módosulhat. Érzelmi, izgalmi állapotok is előidézöik a változást, mint például félelem, idegesség vagy öröm [?]. Ez a jelenség szintén potenciális felhasználási lehetőségeket rejt magában. Mivel a pupillareflex akaratlagosan nem koordinálható, állandó fénymennyiség mellett a pupillaméret változásának figyelésével detektálhatóvá válhatnak a fent említett érzelmi állapotok. Ehhez a változás mértékének és sebességének pontos mérése szükséges, ami azonban kellően nagy sebességű kamerával és elfogadható számítási teljesítményt nyújtó hardverrel kielégítő minőségben megtehető lehet. Az alkalmazás ráadásul nem feltétlenül igényli a szem közvetlen közlről (például fejre erősített kamerával) történő felvételét. Megfelelően nagy felbontású forrás esetén az arc-, majd szemrégio automatikus szegmentálása után a felismert zónát felhasználva, akár távolról is történhet a pupillareflex vizsgálata.

1.1.3. Orvosi felhasználás

Egyes betegségek is okozhatják a pupilla rendellenes méretét vagy viselkedését. Például a „miosis”, azaz a szem összehúzódása nemcsak a fent említett okokra vezethető

vissza. Rendellenes összehúzódás alakulhat ki bizonyos patológiai állapotok, gyógyszerek, vagy mérgek hatására, sőt a mikrohullámú sugárzásnak kitett szervezet is produkálja ezt a tünetet. A „mydrasis” (a pupilla tágulása) során ugyancsak nem megszokott viselkedés alakulhat ki bizonyos gyógyszerek vagy kábítószeres használatkor, de akár komoly fizikai trauma hatására is a normálisnál jelentősebb mértékű vagy időtartamú lehet a pupilla tágulata. A két szem eltérő méretű pupillája (az „anisocoria”) olyan betegségek meglétét jelezheti, mint a Horner-, vagy az Adieszindróma [?]. Orvosi szempontból is van tehát mit vizsgálni: a pupilla követésével egyes betegségek, állapotok felismerése, vagy alakulásuk megfigyelése laikus és orvos számára is automatizálható, megkönnyíthető lehet.

Továbbra is orvosi területen maradva a szemmozgás követése és regisztrálása *ontoneurológiai* vizsgálatokban is szerepet kaphat. Az ilyen vizsgálat célja az egyensúlyszerv működésének megfigyelése és értékelése. Az összetett vizsgálat egyes fázisaiban a szemmozgások követése fontos információt hordoz az alany állapotáról, ugyanis a szemmozgató és az egyensúlyi információkat szállító idegpályák szoros kapcsolatban állnak egymással.

1.2. Gyakorlati felhasználás

1.2.1. Webergonómia

Ugyancsak felkapott kutatási terület manapság a *webergonómia* területe. Kellően pontos követéssel vizsgálható lehet, hogy a dizájnereknek mennyire sikerült a felhasználók igényeinek megfelelő felületet alkotniuk: könnyen eligazodnak-e rajta, esetleg idejük nagy részét a rossz tervezés következtében kaotikus bolyongással töltik.

1.2.2. Vezetésbiztonság

Vezetésbiztonsági alkalmazásban is elképzelhető lehet a pupillakövetés alkalmazása. A követés során mintegy járulékos információként mérhetjük a pislogások gyakoriságát és hosszát, ezzel felismerhetővé válhat a gépjárművezetés közben lankadó figyelem, és jelezhető, ha fennáll az elalvás veszélye. Az eljárás így hasznosnak bizonyulhat már meglévő elalvásdetektálási módszerek [?] kiegészítéseként, tovább javítva azok megbízhatóságát.

2. fejezet

A tekintetkövetésről

A tekintetkövető rendszer fejlesztésének megkezdése előtt számos elméleti és gyakorlati szempont veendő figyelembe, hogy a rendszer által nyújtott kritériumok a lehető legjobban közelítsék az elvártat. Nem kerülhető meg az általános technikák, vagy már meglévő megoldások áttekintése a szakirodalomból, aminek segítségével képet alkothatunk a kutatási téma jelenlegi állásáról, és a gyakorlati szempontok figyelembe vételével dönthetjük el, hogy mely módszer mellett tesszük le végül a voksunkat.

A fejezet 2.1. szakaszában mindenekelőtt a jelenleg elérhető tekintetkövetési módszereket veszem górcső alá, majd hasonlítom össze őket, hogy kellően megalapozott döntést hozhassak a később felhasználni kívánt technikával kapcsolatban.

2.1. Tekintetkövetési módszerek

Bár már lassan másfél évtizedes, a témában mégis hiánypótló munka Arne John Glenstroupnak és Theo Engell-Nielsennek, a dán Københavns Universitet (Kopenhágai Egyetem) hallgatóinak diplomamunkája [1]. Dolgozatuk jelen szempontból fontos második fejezetét a modern tekintetkövetési technikák bemutatásának és összehasonlításának szentelik, meglátásaik pedig mind a mai napig helytállóak.

Ebben a szakaszban főleg az ő munkájuk alapján szeretném összefoglalni a tekintet követésére felhasználható technikákat, néhol a lényeg kiemelésével, ahol pedig szükséges, a hivatkozott cikk keletkezése óta az idő múlásával érvénytelenné vált adatok, paraméterek aktualizálásával.

2.1.1. Az ideális tekintetkövető

Manapság számos módszer kínálkozik a tekintet követésére. De mégis milyen követelményeket kell teljesíteni az „ideális” tekintetkövetőnek? Scott és Findlay 1993-as munkájukban [2] Hallett eredményeit [3] figyelembe véve meghatározták az ideális tekintetkövető eszköz (vagy rendszer) paramétereit és tulajdonságait.

Ezek szerint az ideális tekintetkövető eszköznek a következő 12 pont által támasztott követelményeket kell teljesítenie:

- a. az arc és a fejrégió könnyen hozzáférhető maradjon

- b. ne legyen fizikailag kapcsolatban a vizsgált személlyel
- c. ha szükséges, képes legyen stabilizálni a kapott eredményeket
- d. a rendszer *pontossága* néhány százalékos (1–2 szögperces) eltérést engedjen meg
- e. támogasson legalább egy szögperces *felbontást* másodpercenként, hogy a szempozíció legkisebb változása is követhető legyen; a felbontásnak csak az érzékelő eszköz zaja szabjon határt
- f. támogasson kellően széles *dinamika-tartományt* a szem mozgásainak leképezéséhez
- g. a rendszer időbeli dinamikája legyen megfelelő (jó erősítés, kis fázistolás)
- h. nyújtson *valósídejű* válaszidőt
- i. legyen *invariáns* mindhárom forgási és eltolási szabadságfokra
- j. egyszerűen kiterjeszthető legyen mindkét szem feldolgozására (*binokuláris* vizsgálat)
- k. *kompatibilis* meglévő legyen fej- és testfelvételek használatával
- l. tesztalanyok széles skáláján (pl. nem, kor, rassz szerint, vagy szemüvegesek és szemüveg nélküliek körében) használható legyen

A fent felsorolt követelmények valóban az ideális esetet testesítik meg. Sokszor egy követelmény enyhítésével, vagy figyelmen kívül hagyásával más követelmények kielégítése jelentősen egyszerűsödik. Ha például a *b.* pont szerinti fizikai kontaktust mégis megengedjük, rögzíthetjük a követésre használt kamerát a vizsgálni kívánt alany fejéhez (természetesen kellően diszkrét és kényelmes módon). Ebben az esetben a *i.* pont szerinti invariancia biztosítása máris egyszerűbb feladat, mint külső nézőpontból geometriai transzformációk segítségével végezni ugyanezt.

Több más, látszólag egymásnak ellentmondó követelményt is észrevehetünk az ideális tekintetkövető eszköz 12 pontja között. Komoly megoldandó mérnöki feladatot jelent az ütközések feloldásával az ideális eszközt különböző, valós életben használható megoldásokká alakítani.

A gyakorlatban használható tekintetkövetési megoldások egy lehetséges csoportosítása az alábbi elveken működő iránymeghatározásokat tartalmazza:

1. a szem felületének (vagy a felület speciális megvilágításának) alapján, optikai elven
2. a szem körüli bőrfelület elektromos potenciáljának mérésével
3. speciális kontaktlencse használatával

Már csekély megfontolással látszik, hogy minden csoportnak vannak előnyei és hátrányai. Például az 1. csoportba tartozó megoldások igénylik a legkevésbé – többnyire egyáltalán nem – a fizikai kontaktust a vizsgált személlyel. Azonban elképzelhető, hogy ennek az az ára, hogy az ebbe a csoportba tartozó megoldások nem veszik fel a versenyt a másik két alapelv szerint fejlesztett rendszerekkel pl. a pontosság, vagy valamely más kritérium tekintetében.

Ennek megfelelően szükség van a lehetőségek számba vételére, értékelésére, majd összehasonlítására, hogy a legmegfelelőbb technikát tudjuk kiválasztani, ha tekintetkövető eszköz (alkalmazás) fejlesztésébe fogunk. A következő szakaszokban ezért röviden ismertetem a fenti kategóriákba eső megoldások alapjait, majd összehasonlítom jellemző tulajdonságaikat, valamint a használatukkal elérhető fontos pontossági- és sebességértékeket.

2.1.2. Optikai elvű követés

Az optikai elvű követés során – a nevéből adódóan – optikai úton próbáljuk meghatározni a tekintet irányát. Ez történhet speciális megvilágítás nélkül, pl. az írisz (pontosabban az írisz és az ínhártya közti határvonal, a limbus), vagy a pupilla követésével.

A másik lehetséges megoldásra a szem speciális, többnyire infravörös fénnel történő megvilágítása, majd a szemgolyó felületén megjelenő tükröződések, visszaverődések vizsgálata kínálkozik. Az infravörös megvilágítás előnye a látható fénnel szemben egyrészt az, hogy nem zavarja a vizsgált személyt, nem vonja el a figyelmét, másrészt pedig infraszűrők használatával a látható fény tartományába eső változásokkal a megfelelő körülmények között (lehetőleg kevés napfény a magas infratartalma miatt) invariánssá tehető.

Limbuskövetés

Az írisz és az ínhártya közötti határvonal a többnyire nagy intenzitásváltozás miatt könnyen detektálható lehet. Ugyanakkor meg kell jegyeznünk, hogy az esetek jelentős részében a limbus számottevő része lehet a szemhéjak takarásában. Ennek megfelelően a technika csak vízszintes helyzet és mozgás követésére alkalmas kielégítő pontossággal [2].

Klasszikus esetben ez a követési technika a limbus *relatív* fejhez képest helyzetén alapul, ezért a vizsgálat közben a fejmozgások teljes hanyagolását, esetleg a tekintetkövető eszköz fejhez rögzítettségét igényli.

Pupillakövetés

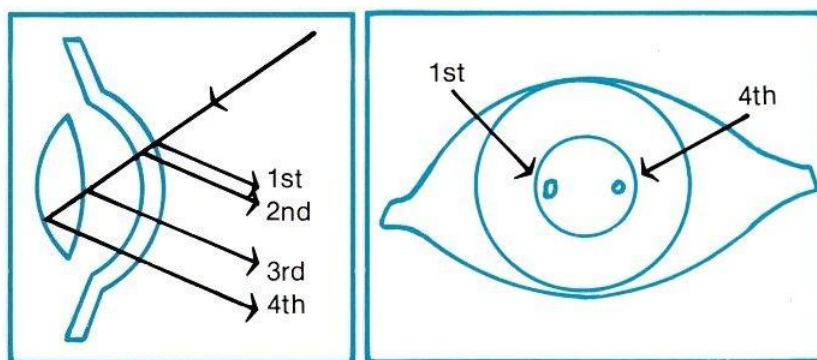
A pupillakövetési technika hasonló az előbb említett limbuskövetéshez, de némi többet előnyt is magában hordoz. Az egyik ilyen előny, hogy a pupilla közel sincs olyan nagy részben a szemhéjak takarásában, mint a limbus, ennek következtében függőleges irányú követés is megvalósítható lehet. A másik előny, hogy az írisz és a pupilla közti határvonal jóval élesebb, mint az írisz-ínhártya határa, ezért jobb felbontással, pontosabban tudjuk meghatározni a nézeti irányt.

Az előző technikához képest azonban meg kell említeni a felmerülő nehézségeket is. A kontraszt ugyan lehet, hogy nagyobb az írisz-pupilla határon, azonban

az intenzitáskülönbség nem olyan jelentős, mint a limbus környezetében. Az sem elhanyagolható szempont, hogy a pupilla átmérője is szükségszerűen kisebb, mint az íriszé, vagyis a forrásnak relatívan nagyobb felbontásúnak kell lennie, hogy azonos pixelnyi méretű pupillát és íriszt detektálhassunk; ellenkező esetben a kisebb abszolút méret a pontosság rovására mehet.

Visszaverődés-alapú követés

A szemet (infravörös) fénnel megvilágítva több tükröződés is megfigyelhető lesz a szemlencse és a szaruhártya határán: ezek az úgynevezett *Purkinje-képek* (lásd 2.1. ábra).



2.1. ábra. A *Purkinje-képek* elhelyezkedése. Forrás: <http://www.fourward.com/dconcept.htm>

Ezen visszavert képek intenzitása sorrendben egyre csökken, az első azonban (közkeletű nevén az úgynevezett „csillanás”, angolul *glint*) még viszonylag egyszerűen detektálható. Infravörös megvilágításban ugyancsak egyszerű a megfelelő kamerával a pupilláról visszavert fény detektálása – a pupilla a környezeténél jóval nagyobb mértékben veri vissza az infravörös fényt, az infraképen egy „fényes”, kontrasztos objektumot alkotva.

A fenti két objektum egymáshoz viszonyított *relatív* helyzetéből következtethetünk a tekintet irányára, ugyanis az első Purkinje-kép világos pontja, illetve a pupilla kontrasztos ellipszise egymással összefüggésben mozdulnak el a fej vagy a szem mozgásának következtében.

Az előző bekezdésben foglaltakból látszik, hogy a technika nem igényeli a fej mozgatlanságát, vagy a képfelvevő eszköz fejhez rögzítését. Ezen előnye azonban egy megkötést is magával hordoz: egyszerű algoritmusok segítségével kb. $\pm 12\text{--}15$ foknyi szabadsága van a felhasználónak a fejmozgásokra [2], nagyobb mértékű mozgások esetén ugyanis komplexebb matematikai számítások szükségesek a követett objektumok mozgásának modellezéséhez.

A visszaverődés-alapú megoldásokkal rokonítható még az a módszer, amikor szintén az első Purkinje-kép erős visszaverődését detektálva a képen, kivágunk egy viszonylag kicsi (párszor tíz pixel nagyságrendű) részt, a csillanással a középpontban. Az így kapott képről egy neurális hálózat dönti el, hogy milyen nézeti irányhoz tartozik [4].

A neurális hálózatot természetesen be kell tanítani a használhatóság érdekében. A betanítási procedúra első lépéseként tanítóképeket kell generálni, általános esetben pár percet vesz igénybe, ami alatt a követnie kell egy, a képernyőn megjelenő jelzést. A tanítóképek birtokában ezután a hálózat betanítható, ez a jelenlegi technikai szint mellett kb. tíz perces nagyságrendű időt vesz igénybe. Azonban azonos körülmények és tesztszemély esetén a tanítást nem kell máskor újra elvégezni.

A neurális hálózat előnye a betanítás után, használat közben mutatkozik meg leginkább: felépítéséből adódóan a háló nagyon rövid idő alatt „döntést tud hozni”, így a valósidejű feldolgozási sebesség kritériuma mindenképpen kielégített lesz. A döntés eredménye ráadásul akár közvetlenül felhasználható: mindössze úgy kell megterveznünk a rendszert, hogy a neurális háló kimenet közvetlenül egy kétdimenziós koordinátát szolgáltatson a felhasznált képernyő koordináta-rendszerében.

A módszer egy másik előnye, hogy nem igényel közeli, nagy felbontású képet a szemrégióról, egy átlagos felbontású kamerával, pl. kartávolságból is elegendő nagyságú lesz a szemrégió. Ennek egyik folyományaként – hogy nagyobb látószögű képek használhatók – adódik, hogy viszonylag nagy mozgási szabadsága lehetséges a vizsgált személy fejének tekintetében, anélkül, hogy a kamerát újra kéne pozicionálnunk.

Ennek a szabadságnak azonban ára van: ha a kalibrálási fázis során több fejpozícióból is rögzítettünk tanító képeket, akkor a betanítás után a neurális hálózat jó eséllyel fogja felismerni különböző fejpozíciókban is a tekintet irányát; a mozgási szabadság azonban a pontosság csökkenésével jár. A pontosság egyébként is érzékeny pontja az eljárásnak, ezt pedig éppen úgy növelhetjük, ha tanítás és előhívás közben *nem* engedjük meg, hogy a fejpozíció változzon.

Purkinje-képek követése

Egy további lehetséges optikai elven működő tekintetkövetési technikát mutatott be Müller *et al* 1993-ban [5], „*Kettős Purkinje-kép*” (Dual-Purkinje Image) módszer néven. Az eljárás lényege, hogy a már említett első és negyedik Purkinje-kép egymáshoz viszonyított helyzetéből számítja a tekintet irányát. Megfelelő matematikai modell esetén a módszer rendkívül pontos, azonban a negyedik Purkinje-kép alacsony intenzitása miatt hatványozottan érzékeny a megvilágítási problémákra.

2.1.3. Elektromos potenciál-alapú követés

Az eddigiektől egy merőben eltérő megközelítése a problémának az *elektro-okulográfia* (EOG). Az eljárásból nyerhető elektro-okulogram rögzítése pl. megismerési és kognitív folyamatok, a vizuális információfeldolgozás, vagy az alvásvizsgálat terén szokásos.

A módszer a szemgolyó elülső és hátulsó pólusa közötti potenciálkülönbség mérésén alapul, amellyel mind függőleges, mind vízszintes irányban követhető a szem mozgása, de csak körülbelül 1–2 fok pontossággal. A használata azonban korántsem mondható egyszerűnek: az potenciálváltozást érzékelő elektródák (lásd 2.2.) miatt szükséges fizikai kontaktuson kívül a rendszer kalibrációja rendkívül hosszadalmas, és hozzáértést igénylő feladat.



2.2. ábra. Az EOG eljárás elektródái. Forrás: <http://www.metrovision.fr/mv-po-notice-us.html>

2.1.4. Követés speciális kontaktlencse használatával

A szem helyzetének, ebből közvetve a tekintet irányának számításához speciális kontaktlencsét is használhatunk. Az egyik lehetséges megoldás, hogy a lencse anyagában olyan véseteket alakítanak ki (tipikusan valamilyen könnyen felismerhető mintázatot), amelyek a fénytörés felhasználásával megkönnyítik a szem helyzetének meghatározását.

Ha azonban egy megfelelően kis méretű indukciós tekercset is sikerült a kontaktlencse anyagába ágyazni, akkor a fej körül generált nagyfrekvenciás elektromágneses mezőkkel az tekercs helyzete közvetlenül is könnyen meghatározható. Az eljárás körülményessége azonban kétséggé teszi a technika laboratóriumokon kívüli felhasználását.

2.1.5. A tekintetkövetési módszerek összehasonlítása

A 2.1. táblázat tartalmazza az ebben a szakaszban felsorolt tekintetkövetési módszerek összehasonlítását. Az összehasonlítás szempontjai a módszer által igényelt kontaktust, az általa nyújtott pontosságot, illetve felbontást (hogy a bemeneti képen mekkora minimális elmozdulás jelent változást a kimeneti pozícióban) jelentik.

2.1. táblázat. A felsorolt tekintetkövetési módszerek összehasonlítása.

	kontaktus	pontosság	felbontás
limbuskövetés	pl. áll-tartó	1–7°	0,1°
pupillakövetés	nincs	0,003°	0,005°
visszaverődés alapú	nincs	0,5–2°	jó
neurális hálózat	nincs	1,5°	–
kettős Purkinje-képek	nincs	0,017°	0,25°
elektro-okulográfia	elektródák	±1,5–2°	jó
kontaktlencse	kontaktlencse	0,08°	0,017°

3. fejezet

Szemmozgások

A szív érez, a szem kutat

...

A szemmozgások többsége – így a későbbi vizsgálatok szempontjából fontosak is – a retinán elhelyezkedő sárgafolt (*fovea centralis*, lásd 3.1. ábra) helyzetének megváltoztatására szolgál. A látóideg kivezetése mellett elhelyezkedő sárgafolton csak tömötten egymás mellé rendeződött csapok találhatók. A csapok jó fényviszonyok mellett nagyon magas felbontásban képesek színek gazdag skáláját feldolgozni, ezért a legélesebb látás a sárgafoltra vetített kép esetén lehetséges. Érthető tehát, hogy a sárgafoltot olyan pozícióba célszerű mozgatni, hogy a figyelem tárgyának képe ide vetüljön. Az ehhez szükséges pozicionáló típusú szemmozgások közé tartoznak a *szakkádok*, a *lassú követések*, és – bár meglepőnek tűnik – a *fixációk* is, ezekkel a fejezet további szakaszaiban részletesebben foglalkozok.

A *fovea* pozicionálását elősegítő mozgások mellett természetesen más is szerepet kap a kívánt kép előállításában. A teljesség igénye nélkül, ilyenek például a szem divergenciáját illetve konvergenciáját beállító mozgások (mélységérzékelés), valamint azok a reflexszerű mozgások, amelyek az egyensúlyszervvel összhangban beállítják a szemeket a fej térbeli orientációjának megfelelően.

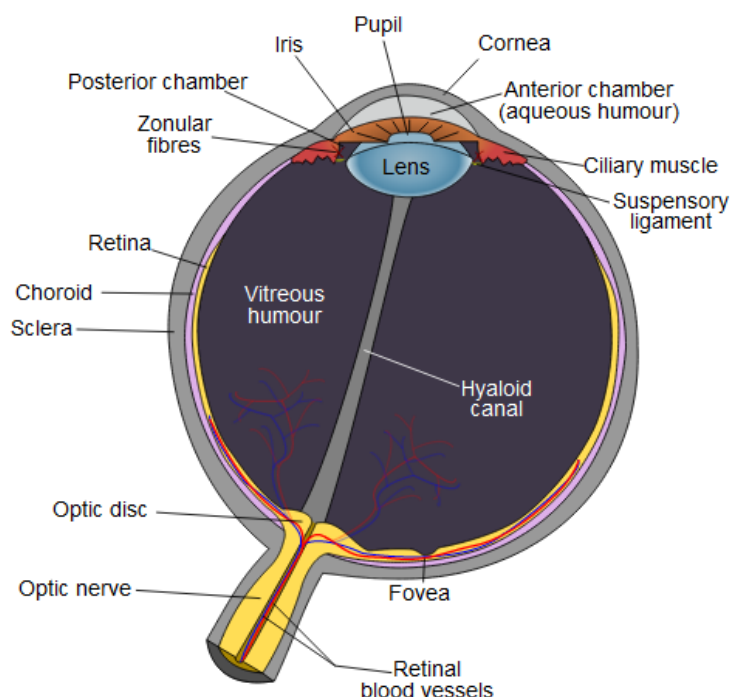
A fejezet 3.1. szakaszában sorra veszem, és bemutatom a pozicionáló típusú szemmozgások alapjait, a mérnöki megközelítésből jelentéktelen részleteket mellőzve. Végül a 3.2. szakaszban összefoglalom a megszerzett ismereteket, valamint hogy a megismert mozgások milyen követelményeket, illetve korlátozásokat támasztanak a realizálandó rendszer tervezése során.

3.1. Osztályozás

3.1.1. Szakkádok

A szakkád mindkét szem gyors, egyidejű, azonos irányú mozgását jelenti. Célja a fixáció áthelyezése az egyik tárgyról a másikra. A mozgás során észlelt jellegzetes „ugráló” mintáról kapta a nevét: a francia *saccader* szó jelentése „rángatás”, „ugrás”.

A szakkádok lehetnek akaratlagosak és reflexszerűek is. Az egyensúlyszervből érkező jelre például reflexszerűen aktiválódhatnak, míg a tekintetünk, figyelmünk áthelyezése nyilvánvalóan akaratlagos cselekedet. Időtartamban a szakkádok körülbelül



3.1. ábra. Az emberi szem felépítése. Forrás: <http://en.wikipedia.org/wiki/Eye>

10 és 100 ms közé tehetők. Ez kellően rövid ahhoz, hogy az agy a gyakorlatban ne észlelje, hogy a mozgás alatt nem történik információfelvétel – vagyis a szakkádikus mozgás közben pár pillanatra effektíve vakok vagyunk. [6]

A szakkádok végrehajtásáról megoszlanak a vélemények. Korábban a szakkádokat ballisztikusnak tartották, vagyis amint a következő fixációs pont helye kiszámításra kerül (nagyjából 200 ms idő alatt), a mozgás már nem megszakítható vagy megváltoztatható. [7] Ezt a feltevést támasztotta alá a tény, hogy a végrehajtás 10-100 ms-os időtartama alatt vizuális visszacsatolásra nincs elegendő idő. Léteznek feltevések azonban, amelyek szerint nincs szükség vizuális visszacsatolásra a végrehajtás közbeni megváltoztatáshoz, így a mozgás ballisztikussága (a nagy sebességek miatt) csak látszólagos. [8]

3.1.2. Lassú követések

A lassú követés (*smooth pursuit*) mozgó tárgyak vizuális követésére szolgál. Az ilyen típusú mozgás könnyen modellezhető egy negatív visszacsatolású szabályozóval. [7]

Egy bizonyos sebességig a szem képes tisztán lassú követést használni a fixáció tárgyán tartására, azonban a $30^\circ/\text{s}$ sebességnél gyorsabban mozgó tárgyak esetén a megfelelő követéshez általában szakkádok beiktatása szükséges. Itt jegyzendő meg, hogy a lassú követés vízszintes és függőleges irányban nem szimmetrikus: a legtöbb ember a vízszintes mozgásokat jobban, míg a függőleges mozgásokat kevésbé tudja lekövetni (ahol „jó” követésen azt értjük, hogy nem szükséges szakkádok beiktatása). A szakkádokkal ellentétben a lassú követést egyértelműen megváltoztathatja az érzékelt vizuális visszacsatolás, a mozgás nem ballisztikus.

Érdekesség, hogy a legtöbben vizuális stimuláció (tényleges mozgó tárgy) nélkül

nem tudnak lassú követéses szemmozgást előidézni, csak rövid szakkádok gyors egymásutánját. Szintén megemlítendő, hogy a lassú követés bár hasonlóan tűnik ahhoz, amikor a fej mozgását „kompenzálva” egy álló tárgy képét fixáljuk a látómezőn, a két mozgás alapvetően különbözik: már abban is, hogy míg az egyik akaratlagos, a másik reflexszerűen hajtódik végre.

3.1.3. Fixációk

A fixációs mozgások a retina stabilizálására szolgálnak, aminek köszönhetően az álló objektumokat tudjuk figyelmünk középpontjában tartani. Az előző szakasz alapján azt gondolhatnánk, hogy a fixációt ugyanazok az idegi pályák aktiválhatják, mint amelyek a lassú követésért felelősek, mindössze a „mozgás” sebessége zérus.

A fixáció azonban több, mint nulla sebességű mozgás: a fixációk közben *tremor*, *drift* és *mikroszakkád* jellegű mozgások váltják folyamatosan egymást. Az állandó mozgást a fényérzékelő sejtek felépítése indokolja. Ha a kép pár másodpercnél tovább marad teljesen változatlanul a retinán, a sejtek telítésbe kerülnek, és a tekintet elhomályosul.

3.1.4. Nystagmus

A nystagmus szakkádok és lassú követések egymásutánja, nem akaratlagos szemmozgás. Előidéződhet optokinetikus úton, illetve patológiásan is (pl. kábítószer-használat hatására). Patológiás nystagmus jelenléte orvosi szempontból lehet jóindulatú, de akár mélyebb neurológiai problémákra is utalhat.

3.2. Összegzés

A fejezet első szakaszában röviden bemutattam a *fovea* pozicionálására szolgáló szemmozgásokat, amelyek alapszintű megismerése elengedhetetlen a tekintetkövető rendszer fejlesztése során. A rendszernek – ahogy ez a mozgások tulajdonságaiból kitűnik – mind a gyorsaság, mind a pontosság tekintetében komoly követelményeket kell teljesítenie.

Az elérendő *gyorsasághoz* a szakkádok, mint a leggyorsabb sebességű mozgások időtartama (10-100 ms) adhat támpontot. Egy átlagos kamerakép 60 Hz-es frissítése legjobb esetben 16 ms időközönkénti vizsgálatot tesz lehetővé. Ez az eredmény egy tipikus drágább, digitális webkamera 30 képkocka/másodperces képfrissítésével máris 33 ms-ra csökken, nem is beszélve a 30 FPS-nél gyengébb teljesítményű kamerákról. Tisztában kell lennünk tehát azzal, hogy átlagos képrögzítő hardver használatával előfordulhat olyan szituáció, hogy fizikailag képtelenek vagyunk a nagyon gyors szakkádok pontos követésére.

A *pontosság* tekintetében azt kell figyelembe vennünk, hogy (pl. olvasásnál) a két egymást követő fixációs pont közti eltérés akár szögperces nagyságrendű is lehet. Ezek korrekt elkülönítéséhez nyilvánvalóan minél nagyobb felbontás szükséges. Minél kisebb felbontásúak a felvételek (azaz a szem, illetve a pupilla megjelenítésére relatíve kevesebb képpont kerül), annál nagyobb elmozdulás jut egyetlen képpontra, ami a feldolgozás során a legkisebb elkülöníthető egység.

4. fejezet

Hardverek

4.1. Optikai megfontolások

Az optikai megfontolások közül az első a tekintetkövetés során felhasznált fénytartomány, ami lehet a látható fény tartománya, valamint az infratartomány. A második fontos szempont pedig a kamera látószöge, ami alapjaiban határozza meg a felhasználótól megkívánt beavatkozás mértékét, valamint a szem felbontását (ezzel közvetve az elérhető maximális pontosságot).

A **látható fényben** történő követés egyértelmű előnye, hogy nem igényel speciális hardvert, a legtöbb, a piacon kapható analóg vagy digitális (web)kamera ebben a tartományban „lát”. Sőt, a digitális eszközökben a CMOS/CCD érzékelő elé szinte minden esetben ún. infratükröt (felülvágó szűrőt, amely kirekeszti a látható fény feletti spektrumot) helyeznek, ezeket az eszközöket tehát sikerrel csak látható fényben használhatjuk. A látható fény használatakor azonban nem feledkezhetünk meg annak hátrányairól sem: a követéshez analizálandó kép meglehetősen érzékeny lesz a megfelelő megvilágításra. A fényviszonyok változásának kiküszöbölése komolyabb előfeldolgozást kíván, és extrém esetekben nem is mindig teljesíthető.



4.1. ábra. Az emberi szem képe infravörös fényben. Forrás: <http://bit.ly/aKwzsZ>

Az **infratartománynak** a látható fénnel szemben vannak előnyei, és hátrányai is. Az előnyök közé tartozik, hogy a pupilla képe a infra megvilágításban nagyon kontrasztos (lásd 4.1. ábra), jóval könnyebben szegmentálható az írisztől (különösen sötétebb szemű alany esetén), mint látható fényben. Az infrás megvalósítás hátránya azonban, hogy a megfelelő működéshez sötétet igényel (igaz, ezzel összhangban a látható megvilágítási változásokra invariáns), de sötétben, a tisztán infra megvilágítás megvalósítása nehéz feladat, különösen a piacon kapható egyszerűbb infra LED-ekkel, vagy LED-es reflektorokkal.

A követéshez használt kamera optikája lehet egyrészt **nagylátószögű**. Ez azzal az előnnyel jár, hogy a vizsgálat alatt álló személy fejmozgása szabadabb lehet, nincs szükség a fejpozíció fixálására (pl. álltámasszal). A szemrégió külön szegmentálható, majd ezen régióon belül a pupillát követve valósulhat meg a tekintetkövetés. A módszer hátránya, hogy a szemrégió nem tölti be az egész képet, vagyis nem használja ki maximálisan a rendelkezésre álló felbontást. Ennek következtében a tekintetkövetés pontossága jelentős mértékben romolhat, esetleg lehetetlenné is válhat.

További lehetőség a **teleobjektívek**, vagyis zoomoptikák használata. Teleobjektívvel a szemrégió „közel hozható” még nagy távolságról is, hogy teljesen kitöltse a feldolgozandó képet, kihasználva annak teljes felbontását. Ez azonban azzal a hátránnyal jár, hogy a fejet fix pozícióba kell kényszeríteni (pl. egy álltámassz segítségével), ugyanis a fej mozgásával a pupillarégió kikerülhet a kamera látószögéből.

Végül lehetőség nyílik egyfajta „hibrid” megoldás használatára is, amelynél

5. fejezet

Megvalósítás

A rendszer szoftverét C++ nyelven fejlesztettem, az OpenCV gépi látás könyvtár, illetve a Qt felhasználói toolkit segítségével. A feladat megoldása során törekedtem az objektum-orientált paradigma szem előtt tartására, valamint hogy az elkészült rendszer a lehető legkönnyebben testre szabható, illetve igény esetén kiegészíthető legyen.

5.1. Az OpenCV könyvtár

Az OpenCV¹ egy nyílt forráskódú gépi látás (computer vision) könyvtár. Elsődleges célja, keretet nyújtani **valósídejű** képfeldolgozási alkalmazások fejlesztésére. A könyvtár szabadon letölthető és felhasználható a BSD licenc² keretein belül.

Az OpenCV projekt hivatalosan 1999-ben indult az Intel kezdeményezésében. A nagyközönségnek a 2000. évi „*IEEE Conference on Computer Vision and Pattern Recognition*” konferencián mutatkozott be, majd öt béta-verziót követően 2006-ban jutott el az 1.0-ás hivatalos kiadásig. A fejlesztése itt úgy tűnt, hogy megáll, de végül a projektet a Willow Garage³ robotikai kutatólabor 2008-ban szárnyai alá vette, és azóta is aktív fejlesztés alatt áll. 2008 októberében az 1.1-es verzióval közel egy időben látott napvilágot az első hivatalos OpenCV-vel foglalkozó könyv „*Learning OpenCV: Computer Vision with the OpenCV Library*” címmel Gary Bradski és Adrian Kaehler fejlesztők tollából [?]. Az egy évvel később, 2009 októberében megjelent 2.0-ás verzióval a projekt nagy fejlődésen esett át. Ebben a verzióban található meg először a C++ és Python interfész (ez a meglévő C mellett már három hivatalosan fejlesztett interfészt jelent), amely az egyszerűbb kezelhetőség, új függvények mellett a meglévő eljárások teljesítmény tekintetében – különösen többmagos rendszereken – jobb implementációját kínálja a felhasználóknak.

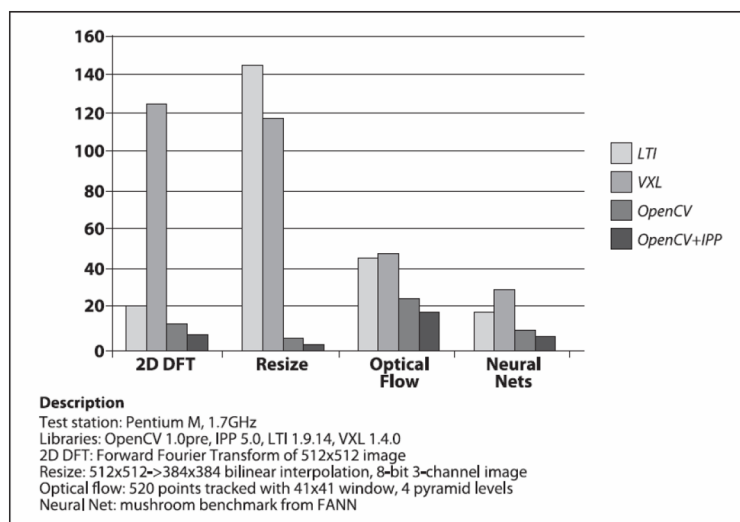
Az OpenCV jelenlegi 2.1-es verziója elérhető FreeBSD, Linux, Mac OS és Windows operációs rendszerek alá. Széles körben, mondhatni világszerte használt, felhasználói táborra több, mint 40 000 főt számlál. Köszönhető ez többek között annak, hogy felhasználási lehetőségei igencsak sokrétűek: több, mint 500 optimalizált algoritmust

¹<http://opencv.willowgarage.com/>

²<http://www.lininfo.org/bsdlicense.html>

³<http://www.willowgarage.com/>

kínál annak érdekében, hogy „ne kelljen újra feltalálnunk a kereket”. Sebesség tekintetében érződik a kipróbált, optimalizált algoritmusok használata: az OpenCV a jelenleg elérhető leggyorsabb alternatíva gépi látás terén (5.1. ábra). Teljesítménye azonban adott esetben még tovább növelhető, mivel ha Intel IPP⁴ (Integrated Performance Primitives) támogatást észlel, az abban található szálakra optimalizált algoritmusok használatát fogja preferálni.



5.1. ábra. Az OpenCV teljesítménye az LTI és VXL képfeldolgozási könyvtárakkal összehasonlítva.

Forrás: [?]

A teljes funkcionalitás részletekbe menő bemutatása – mint láthattuk [?] – egy könyvet is megtölt, de a teljesség igénye nélkül tekintsük át, hogy milyen alapvető jellemzői és alkalmazásai vannak a környezetnek:

- alap adatstruktúrák
 - mátrixok, vektorok
- mátrix és vektor manipuláció, lineáris algebra
- dinamikus adatstruktúrák
 - listák, sorok, halmazok
 - gráfok és fák
- kép és videó input/output
 - beolvasás fájlból (kép vagy videó) és kameráról
 - kiírási lehetőség képként vagy videóként
- előfeldolgozás
 - él- és sarokkeresés

⁴<http://software.intel.com/en-us/intel-ipp/>

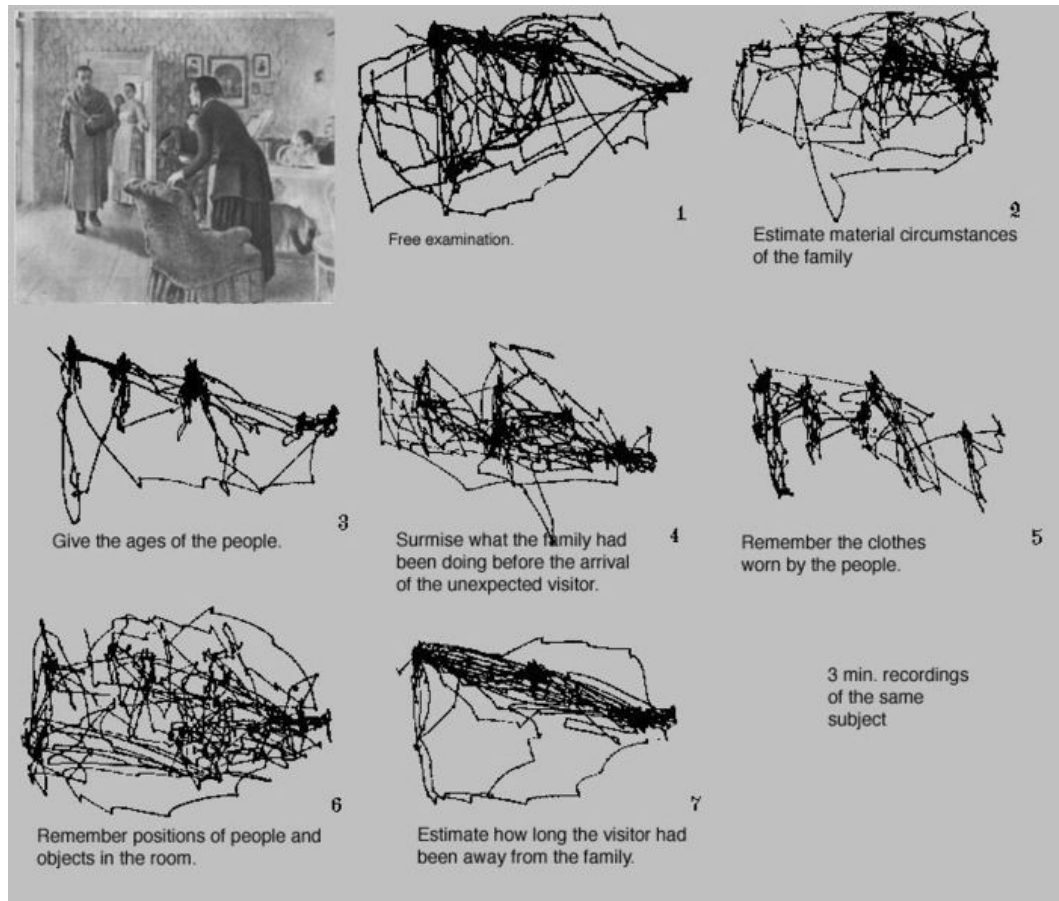
- mintavételezés és interpoláció
- szíkonverzió
- morfológiai operátorok
- struktúraanalízis
 - távolság- és Hough-transzformáció
 - kontúrfeldolgozás
 - sablonillesztés
 - különböző momentumok
 - Delaunay háromszögelés
- kamerakalibráció
 - kalibrációs mintázatok felismerése és követése
 - fundamentális mátrix becslés
 - homográfia becslés
 - sztereó megfeleltetés
- mozgásanalízis
 - optical flow
 - mozgásszegmentálás és -követés
- objektumfelismerés
 - eigen-módszerek
 - rejtett Markov-modell (Hidden Markov Model – HMM)
- GUI és rajzolás
 - kép és videó megjelenítés
 - billentyűzet és egérkezelés
 - egyenes, kör, poligon, szöveg rajzolása

Láthatjuk, hogy a fent felsorolt funkciókkal a gépi látás terén rengeteg egyszerűbb feladatot szinte „egy lépésben”, beépített, optimalizált eljárások segítségével oldhatunk meg. Ha nagyobb szabású projektbe kezdünk, akkor is hasznunkra lehet, hogy részben vagy egészében egy több tízezres felhasználói tábor (melynek jelentős részét aktív kutatók alkotják) visszajelzései alapján fejlesztett környezetre építhetjük munkánkat.

Az OpenCV számos előnye mellett hátrányként említhető meg, hogy segítségével a felhasználói felületet csak nagyon leegyszerűsített módon szabhatjuk testre. Igaz, hogy a könyvtár feladata elsősorban a képfeldolgozás és nem a megjelenítés, így ebből a nézőpontból a beépített kép és videó megjelenítési lehetőség, billentyűzet- és egérkezelés valamint trackbarok (csúsztható kezelőszerv értékek beállítására) létrehozásának lehetősége inkább hozzáadott értékként jelenik meg. Azonban ez nem változtat azon a tényen, hogy ha igény van felhasználóbarát kezelőfelület készítésére, az OpenCV-t integrálnunk kell valamely elterjedt grafikus felhasználói felület toolkittel.

5.2. Pszichológiai bemutató

A rendszer működését úgy demonstráltam, hogy végrehajtottam Alfred L. Yarbus orosz pszichológus 1967-es tanulmánya egy részletét. A kísérletben a kutatók azt bizonyították be, hogy a tesztalanyoknak előzetesen különböző kérdéseket feltéve, a kérdések jelentősen befolyásolták egy kép részleteinek vizsgálatát ahhoz képest, ha csak „szabadon” nézegették azt. A teszteredményeket összefoglaló kép – mint az egyik első a témával foglalkozó eredmény – jól ismert a tekintetkövetéssel foglalkozók körében, ez látható a 5.2. ábrán.



5.2. ábra. Yarbus '67-es kísérletének eredménye

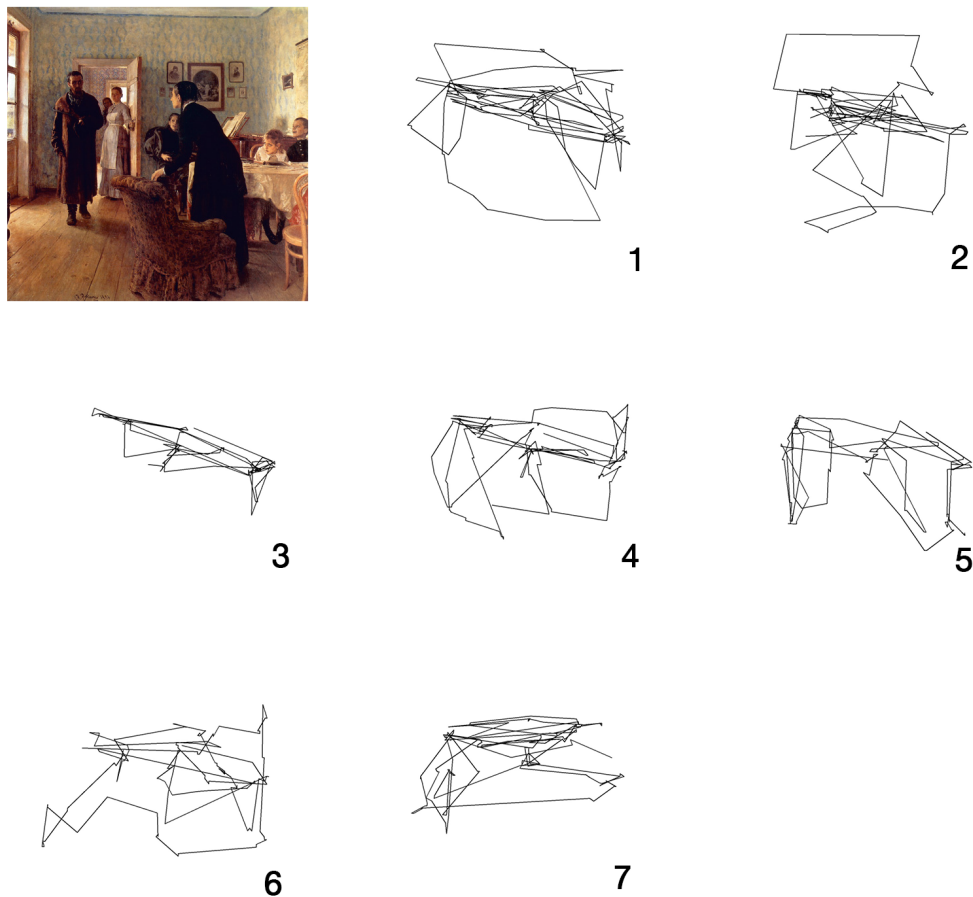
Látható, hogy a kísérlet végrehajtásához szükség volt a tekintet követésére. Ekkor még a kornak megfelelően nem álltak rendelkezésre kifinomult módszerek: az alanyokat egy meglehetősen kényelmetlen acélszerkezethez rögzítve vizsgálták. A bemutató alkalmazásomban arra kerestem a választ, hogy lehetséges-e hasonló méréseket elvégezni az általam fejlesztett rendszerrel.

Az általam tesztelt alanyoknak a következő kérdésekre kellett válaszolniuk a tesztkép (Repin: Váratlan utazó) rövid vizsgálata után. A vizsgálatot minden kérdésben 200 beérkezett érvényes mérési pontig folytattam, ez a felhasznált webkamera, illetve a feldolgozási sebesség mellett nagyjából kérdésenként 15–20 másodpercet vett igénybe.

1. szabad nézelődés

2. „Milyen anyagi körülmények között él a család?”
3. „Adja meg az egyes szereplők életkorát!”
4. „Mit csinálhattak a szereplők, mielőtt az utazó betoppant?”
5. „Milyen ruhát viselnek a kép szereplői?”
6. „Próbáljon megjegyezni minél több strukturális részletet (személyek, tárgyak pozíciója)!”
7. „Mennyi ideig lehetett távol az utazó a családtól?”

A kérdésekre nem volt jó, vagy rossz válasz, a kísérlet minden alanynál csak a feladatnak megfelelő figyelmi területek változását vizsgálja. A vizsgálatom eredménye a ábrán követhető.



5.3. ábra. Eredmények a saját rendszerrel

Látható – bár ez nem volt feltétel – hogy a mérési eredmények ezen alany esetén meglehetősen jól fedik Yarbus tesztalányának eredményeit. Az viszont mindenképpen kijelenthető, hogy szignifikáns különbség van az ábra első (szabad nézelődés), valamint többi része között. Nem céлом a kísérlet pszichológiai eredményeit értékelni, az viszont bizonyos, hogy a rendszer hasonló kísérletek elvégzését támogatja.

5.3. Webergonómiai bemutató

Ahogy az előző fejezetemben már említettem, a tekintet követése a webergonómia területén is fontos kísérletek elvégzésére ad lehetőséget. A rendszer ilyen irányú képességeit demonstrálandó, elkészítettem az pszichológiai kísérlet adatait hő térképen (heat map) megjelenítő funkciót is, ennek eredménye egy teszt képre a ábrán látható.



5.4. ábra. Hő térképes megjelenítés a pszichológiai teszt egy adathalmazára („Adja meg a szereplők életkorát!”)

Minden vizsgálatban célszerű a lehető leginkább informatív megjelenítést választani a rendelkezésre álló adathalmaz alapján. A képen egyre pirosabb színnel jelöltem a kép frekventáltan vizsgált területeit, például egy webergonómiai vizsgálatban az adatok effajta vizualizációja lehet a legcélszerűbb.

Irodalomjegyzék

- [1] Arne John Glenstrup and Theo Engell-Nielsen. Eye Controlled Media: Present and Future State. Master's thesis, University of Copenhagen, 1995.
- [2] D. Scott and J. M. Findley. Visual search, eye movements and display units. Technical report, University of Durham, South Road, Durham DH1 3LE, UK, 1993.
- [3] P. E. Hallett. *Handbook of Perception and Human Performance I*, chapter Eye movements, pp. 10.25-10.28. 1986.
- [4] S. Baluja and D. Pomerleau. Non-intrusive gaze tracking using artificial neural network. Technical report, School of Computer Science, Carnegie Mellon University, Pittsburgh PA, USA, 1993.
- [5] G. d'Ydewalle P. U. Muller, D. Cavegn and R. Groner. *Perception and Cognition*, chapter A comparison of a new limbus tracker, corneal reflection technique, purkinje eye tracking and electro-oculography, pp. 393-401. Elsevier Science Publishers, B.V, 1993.
- [6] Shebilske W. L. and D.F. Fisher. Understanding Extended Discourse Through the Eyes: How and Why. *Eye Movements and Psychological Functions: International Views (pp. 303-314)*., 1983.
- [7] R. H. S. Carpenter. *Movements of the Eyes*. London: Pion, 1977.
- [8] J. D. Cook D. A. Robinson D. S. Zee, L. M. Optican and W. K. Engel. Slow Saccades in Spinocerebellar Degeneration. *Archives of Neurology*, 33, 243-251., 1976.