

# Élpont osztályozáson alapuló robusztus tekintetkövetés

HELFENBEIN TAMÁS

*Ipari Kommunikációs Technológiai Intézet, Bay Zoltán Alkalmazott Kutatási Közalapítvány  
helfenbein@ikti.hu*

Lektorált

*Kulcsszavak: pupilladetektálás, pupillakövetés, tekintetkövetés, képfeldolgozás*

**A billentyűzet és az egér után a modern számítógépek perifériájává vált maga az ember is. Az emberi mozdulatokat felismerő perifériák mellett a tekintetkövető rendszerek is egyre nagyobb szerephez jutnak az ember-gép kapcsolat megvalósításában. A mobilis, képfeldolgozáson alapuló tekintetkövető eszközök hatékonyságát a változó fényviszonyok nagymértékben csökkentik. A hagyományos technológiák ilyen körülmények között nem használhatók, ezért olyan módszert dolgoztunk ki, mely képes az emberi szem pupilla-paramétereinek robusztus meghatározására változó fényviszonyok között is. A módszer vizsgálatokor bizonyítást nyert, hogy az kiválóan alkalmas mobilis tekintetkövető rendszerekben történő használathoz.**

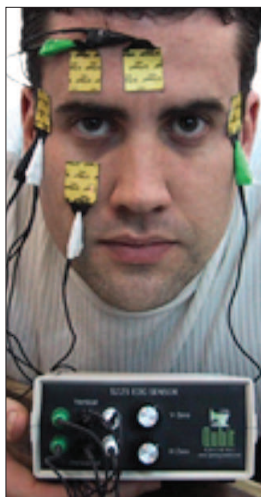
## 1. Bevezetés

A tekintet megfigyelése az emberiség igen korai időszaka vezethető vissza. Az ember különleges tulajdonsága, hogy képes észlelni egy másik ember tekintetének irányát, illetve saját tekintetét áthelyezni az adott objektumra. Ehhez kapcsolódó ősi kézzel az ujjal való mutató, ami arra szolgál, hogy közöljük embertársunkkal, hova összpontosítsa figyelmét. A tekintet abszolút iránya a fej orientációjából, illetve az ehhez képest relatívan mozgó szem irányultságából tevődik össze. A szem mozgásának vizsgálatára az 1900-as évekig közvetlen emberi megfigyeléssel volt lehetőség. A technika fejlődésével egyre összetettebb tekintetkövető eszközök láttak napvilágot, de az igazi áttörést a digitális számítógépek felhasználása jelentette. Lehetőség nyílt tekintetkövető rendszerek használatára az orvoslásban, hadiiparban, biztonságtechnikában és a mindennapi életben is.

## 2. Tekintetkövető rendszerek

A tekintetkövető rendszereket különböző rendszertípusokba sorolhatjuk. Az első rendszertípus egy szemre rögzített paszszív eszköz (kontaktlencse, tükrök, mágneses érzékelő) helyzetét méri. Ennek előnye, hogy pontos és gyakori mérést tesz lehetővé, de az eszköz viselete kellemetlen lehet [1].

A következő rendszertípus a szemkörnyékre helyezett elektródákra épít (1. ábra).



1. ábra  
Elektrookulogram mérése

Az elektródákból származó elektromos jelek változása alapján képes a szemmozgás érzékelésére. Az ilyen rendszerek csukott szem esetén és teljesen sötétben is kifogástalanul működnek [2].



2. ábra  
Mobil és rögzített optikai tekintetkövető eszközök

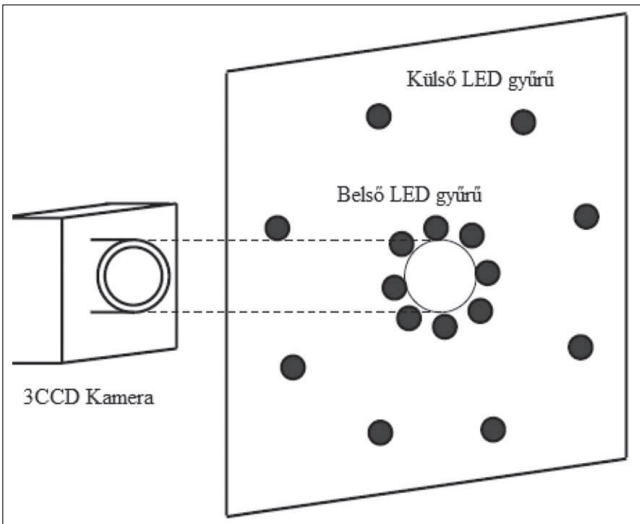
A harmadik típus a kontaktusmentes optikai méréseken alapuló rendszereket foglalja magában. Ebben az esetben a szemet tipikusan infravörös fényforrással megvilágítják és a szemről visszaverődő fényt optikai érzékelőkkel (kamerával) detektálják [3]. Kialakítást tekintve megkülönböztethetünk rögzített, valamint mobilis eszközöket (2. ábra). A rögzített rendszerek képesek a tekintet abszolút irányát megadni, míg a mobilis eszközök az alany fejpozíciójához képest relatívan adják meg a szem irányultságát. Speciális esetben a mobilis eszközök is képesek abszolút irány mérésére akkor, ha valamilyen ismert helyzetű pontokhoz kalibrálva végzik a detektálást.

Napjainkban a tekintetkövető rendszerek területén az egyik legelterjedtebb megoldás a kamerával történő mérés.

### 3. Pupilladetektálás kamerával

Képfeldolgozási szempontból a pupilla detektálásának inputját három alapvető típusra bonthatjuk: sötét pupillakép, világos pupillakép és hibrid [4].

Sötét pupillakép esetén a megvilágító fényforrás az optikai tengelytől egy adott távolságra található. Ekkor a pupilla a kialakult képen sötét színű. Világos pupillakép esetén az optikai tengelyhez közeli, azzal párhuzamos irányultságú fényforrásokat használunk. Ilyenkor a retináról visszaverődő fény miatt a pupilla világos lesz.



3. ábra Hibrid elrendezés

A hibrid megoldásokban mindkét módszert alkalmazzák (3. ábra). Megkülönböztethetünk idő- és hullámhosszosztásos technikákat [4]. Időosztásos esetben a két különböző fényforrást az egymás utáni képekre szinkronizáltan, felváltva használják, így világos és sötét pupillakép egyaránt keletkezik. Ez meglehetősen előnyös, hiszen a világos és sötét képek differenciáján (jó eséllyel) csak a pupilla látszik. Hátrány, hogy a pupilla a két kép készítése között elmozdulhat. E hibátényező mértéke függ a mintavételezés frekvenciájától is.

A hullámhosszosztásos megoldás esetében ezt a tényezőt kívánják kiküszöbölni úgy, hogy a két képet optikai szűrőkkel, egy időpillanatban veszik fel. Ekkor az optikai tengelyhez közeli és távoli fényforrásoknak különböző hullámhosszakot választanak, majd a visszavert fényt 3CCD kamerával érzékelik. Ennek előnye, hogy a két különböző hullámhosszú visszavert fény (az-

az a sötét és világos pupillakép) külön csatornán jelenik meg azonos időpillanatban.

A továbbiakban csak a sötét pupillaképet használó módszereket részletezzük. Számos megoldásban feltételezik, hogy a pupilla fényessége nagymértékben eltér a pupillakép egyéb részeinek fényességétől. Ez lehetőséget ad olyan eljárások használatára, melyek hisztogramküszöbözést alkalmaznak a pupilla, mint objektum detektálására. Sok esetben ezek a megoldások nem használhatók, mivel nem állapítható meg megfelelő küszöbszint. Itt lépnek színre azon eljárások, melyek csak a pupilla lokális tulajdonságait veszik figyelembe, mégpedig azt, hogy a pupilla fényessége eltér a környezetének fényességétől. Ez ad lehetőséget az éldetektáláson alapuló technikáknak. Ide sorolhatjuk a Hough-transzformáció [5] valamely típusát vagy annak továbbfejlesztését, illetve az aktív kontúrokat [6] használó módszereket is.

### 4. Az IKTIVision rendszer

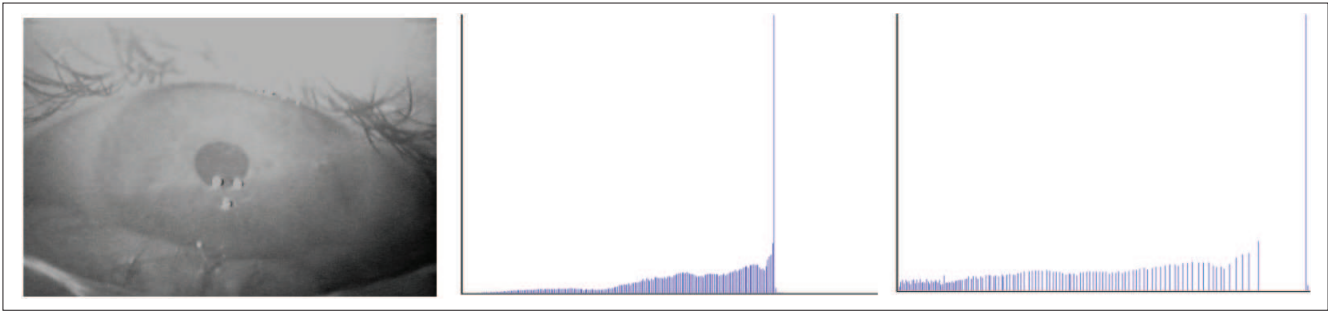
A FormaDisplay Kft. 2008-tól az Új Magyarország Fejlesztési Terv részeként új vizsgálati módszer kialakítását végezte (KMOP-1.1.1-08/1-2008-0038). E projekt keretein belül a Bay Zoltán Alkalmazott Kutatási Közalapítványban kifejlesztettünk egy IKTIVision nevű tekintetkövető rendszert, mely egy viselhető szemüvegen található két kamerából, egy képdigitalizáló eszközből, illetve egy 1.6 GHz-es, egymagos processzorral és 1GB RAM-al rendelkező netbook-ból áll.

A szemkamera (4. ábra) mellett LED megvilágítás található, mely egy féligáteresztő tükrön keresztül a viselő személy szemét világítja meg. A kamera ennek fényét érzékeli a tükrön keresztül. A tárgykamera előre néz, így a kinyert kamerakép a viselő személy által látott képet reprezentálja.

A mobilis tekintetkövető rendszert a mozgó felhasználó tekintetének és érzelmi állapotának megfigyelésére terveztük úgy, hogy a felhasználót a viselt rendszer a legkevésbé befolyásolja. Az alkalmazási terület megköveteli az alacsony fogyasztású és feldolgozási sebességgel rendelkező eszközök alkalmazását, így a felhasznált algoritmusokat is e követelményekhez igazítottuk. Követelmény volt a valós idejű feldolgozás úgy, hogy párhuzamosításra (többmagos processzorok, grafikus kártya) a fogyasztási követelmények miatt nincs lehetőség.



4. ábra Az IKTIVision rendszer felépítése



5. ábra Egy pupillakép, hisztogramja és kiegyenlített hisztogramja

## 5. Pupilladetektálás és követés

Egy mobilis tekintetkóvető rendszer ki van téve a környezeti fényviszonyok változásából adódó zajhatásoknak, melynek következtében a pupillakép kontrasztja annyira lecsökkenhet, hogy a különféle globális mértékeket használó, küszöbözésen alapuló technikák rendre elbuknak a pupilla paramétereinek meghatározásában (5. ábra).

A legtöbb módszer laborkörülmények között jó megoldást ad, de a fényviszonyok változásakor már nem képes a pupilla detektálására. Az IKTIVision rendszerben olyan algoritmusokat dolgoztuk ki, mely a fenti problémákra megoldást kínálva képes a pupilla paramétereinek és a tekintet irányának robusztus meghatározására.

A képfeldolgozási algoritmus két fő lépése (i) a kalibrációs és (ii) a mérési fázis (6. ábra). A kalibrációs lépéseket a mérés megkezdése előtt, kontrollált fényviszonyok mellett kell elvégezni. A kalibrációs fázis fő feladata, hogy egy közelítő transzformációt adjon a két kamerakép között. Ennek menete, hogy a tárgyképen ismert koordinátájú pontokhoz rögzítjük a szemképen a megfelelő pupillavektorokat ( $\underline{V}_P$ ). A pupilla vektor a LED-ek által kijelölt kamera optikai tengelyéből ( $\underline{V}_{LC}$ ) a pupilla középpontjába ( $\underline{V}_{PC}$ ) mutató vektor (7. ábra).

Olyan algoritmusokat (komponenseket) dolgoztunk ki, melyek képesek egy adott pupillaképen a kalibrációs fázisban meghatározni a kamera optikai tengelyét (LD: LED Detektor), valamint a pupilla középpontját és átmérőjét (PDK: Pupilla Detektor és Követő). Ha adott egy tárgyképbeli képpont, akkor a Kalibrátor komponens a fenti

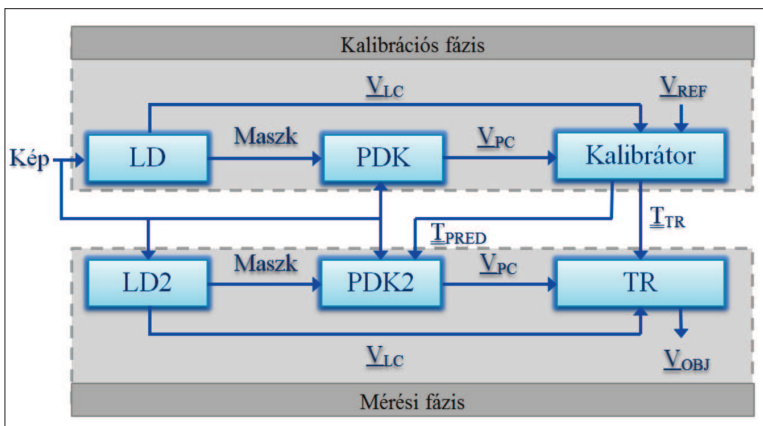
adatok felhasználásával képes eltárolni a paramétereiket, majd legalább három kalibrációs pont ( $\underline{V}_{REF}$ ) megadása után megkeresni a két kamerakép közti transzformációt.

Az LD küszöbözésen alapuló módszerrel keresi meg a lehetséges LED fényforrások képét. A bináris képet szegmentáljuk és kiválasztjuk a LED-eknek megfelelő objektumokat úgy, hogy a priori ismereteinket használjuk fel. A LED fényforrások területére, torzultságára és geometriai elhelyezkedésére vonatkozó paraméterek alapján választjuk ki a legmegfelelőbb jelölteket (8. ábra).

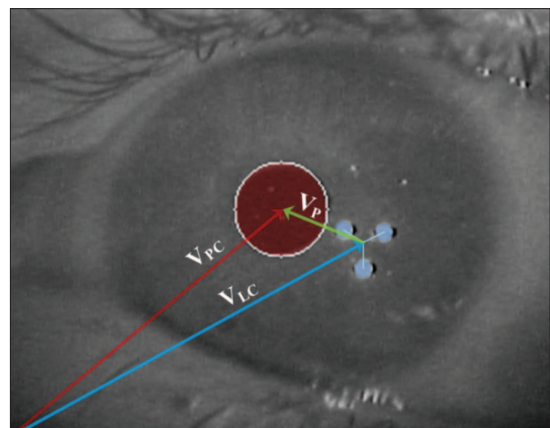
A PDK első körben a pupilla detektálását végzi el. A kép zajszűrése után Canny-élkeresést [7] végzünk a képen (9. ábra bal felső képei). Az élképről eltávolítjuk a LED okozta éleket az LD által kiadott maszk segítségével. A kialakult élkép nem csak a pupillát tartalmazza, ezért szükség van az élpontok szűrésére (9. ábra jobb felső képei). Az eredeti és az élkép felhasználásával soronként végigmegyünk a képen és kiszűrjük azon élpont párokat, melyek a pupilla kezdetét és végét jelenthetik (azaz megfelelő távolságban lévő, nagy abszolút meredekségű átmenetek).

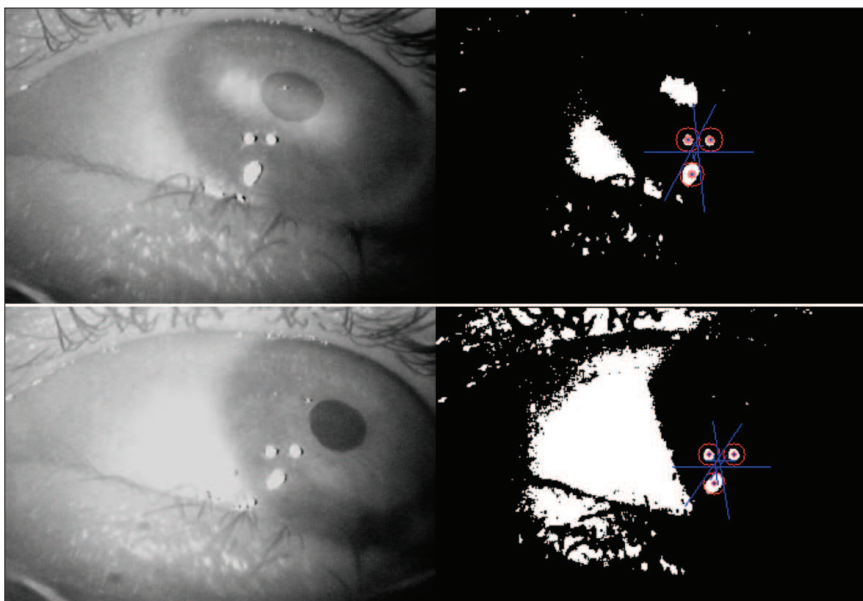
A szűrés után az élpontok klaszterezése következik. Az osztályozásban azt vesszük figyelembe, hogy a pupilla kontúrpointjai egy kisebb szakadásokkal tördelt, viszonylag egybefüggő kontúrnak felelnek meg és a gradiensének irányváltozása közelítőleg konstans. A kialakult pontokat kiegészítjük a kapcsolódó, megfelelő gradiens-változással rendelkező élpontokkal (9. ábra bal alsó képei). Az így kialakult élpont halmazra ellipszist illesztünk (9. ábra jobb alsó képei).

6. ábra Az IKTIVision rendszer feldolgozó modulja



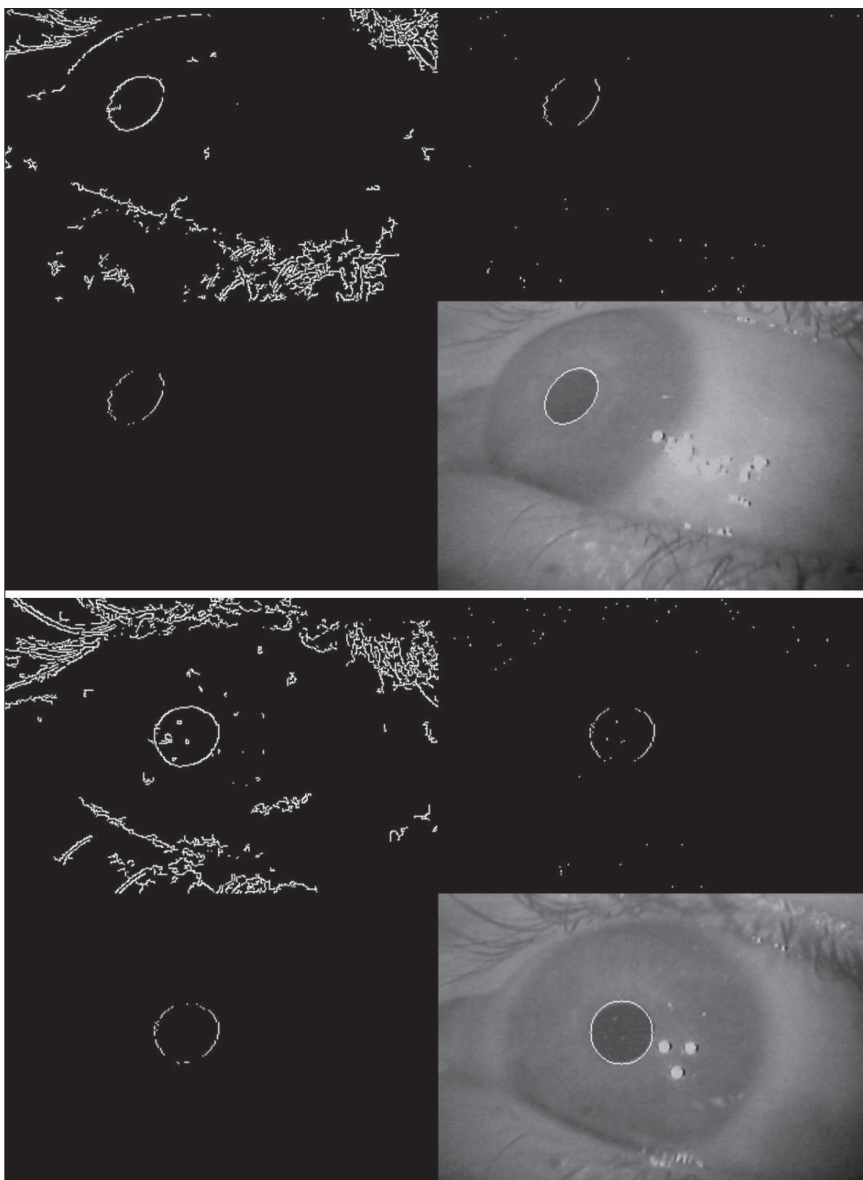
7. ábra A pupilla vektor





8. ábra LD működése

9. ábra Élpontok szűrése és osztályozása



Ha a megelőző időpillanatokban ismertek a pupilla paraméterei, akkor a következő időpillanatra becslést a-dunk Kalman-szűrő használatával [8]. Ezek után már nem a teljes pupillaképen, hanem csupán a pupilla feltételezett helyén végezzük a detektálást.

A kalibráció lehetőséget biztosít arra, hogy a mérési fázis folyamán a pupilla paramétereinek egy részét (a pupilla középpontját) megbecsüljük. Mindezt úgy tehetjük meg, hogy a kalibrációs pontok megadásakor a *Kalibrátor* eltárolja az optikai tengely helyzetét, illetve a pupilla középpontjának koordinátáit (ld. 7. ábra), majd legalább három pont megadása után egy közelítő transzformációt számít ( $\underline{T}_{\text{PRED}}$ ).

Ez a transzformáció leírja, hogy ha az optikai tengely képe a

$$\underline{V}_{\text{LC}} = (x_{\text{LC}}, y_{\text{LC}})^T$$

pupillakép pontban található, akkor a pupilla középpontja megközelítőleg

$$\underline{V}_{\text{PC}} = (x_{\text{PC}/S}, y_{\text{PC}/S})$$

pontban található, ahol

$$(x_{\text{PC}}, y_{\text{PC}}, s)^T = \underline{T}_{\text{PRED}} (x_{\text{LC}}, y_{\text{LC}}, 1)^T.$$

A mérési fázisban szintén meg kell határozni az optikai tengely helyzetét (*LD2*), valamint a pupilla paramétereit (*PDK2*). Itt azonban rendelkezésre áll a pupilla középpontjának becsült pozíciója, így a keresési terület szűkíthető. A pupilla detektálása az előző módszerhez hasonlóan működik, de csak a szűkített területet vizsgálja. Sikeres detektálás után a következő keresési terület itt is becsülhető Kalman-szűrő segítségével. Először a két keresési metszeten végezzük el a detektálást. Ha ez nem vezet eredményre, akkor a keresési területek unióján végezzük el ugyanezt. Ez a kitétel növeli a feldolgozási sebességet, ugyanakkor biztosítja a robusztusságot is a mérési fázisban.

A tekintet iránya a *PDK2* által meghatározott  $\underline{T}_{\text{TR}}$  transzformációs mátrix alkalmazásával kapható meg:

$$\underline{V}_{\text{OBJ}} = (x_{\text{OBJ}/S}, y_{\text{OBJ}/S})^T,$$

ahol  $(x_{\text{OBJ}}, y_{\text{OBJ}}, s)^T = \underline{T}_{\text{TR}} (x_P, y_P, 1)^T.$

## 6. Eredmények

A rendszer működését többféle környezetben vizsgáltuk. A rendszer valós időben végezte a pupillakövetést úgy, hogy az eredményeket és a mozgóképet fájlba rögzítette.

Környezet	Sikeresség	Effektív sikeresség
Kontrollált fényviszonyok	88,2%	98,3%
Változó fényviszonyok	69,8%	81,6%

1. táblázat Mérési eredmények

2. táblázat Kalibrációs idők

Rendszer	Kalibráció	T <sub>min</sub>	T <sub>avg</sub>
ASL Mobil Eye	Manuális	92 s	211 s
IKTIVision	Manuális	40 s	124 s
IKTIVision	Automata	15 s	89 s

Az 1. táblázat foglalja össze a mérések eredményeit. A sikeresség oszlopban a sikeres mérések számának aránya látható az összes méréshez képest, százalékban. Az effektív sikeresség oszlopban csak azon mérések szerepelnek, ahol a pupilla a képen látható volt és a mérési tartományon belül helyezkedett el.

A képdigitalizáló eszköz 640x480-as felbontásban biztosítja a kameraképeket, melyet 29,97 képkocka/másodperc frekvenciával képes a számítógép felé továbbítani. Mérés közben a rendszer a pupilla paramétereit és a tárgykamera képét a merevlemezre menti PICVIDEO MJPEG kódolásban.

A szoftver kétféle kalibrációs módszert is támogat: (i) manuális és (ii) automatikus.

A manuális kalibráció során a kezelő kijelöli a tárgykamera képén azt a pontot, ahova az eszközt viselő felhasználó éppen néz. Legalább három (nem egy egyenesre eső) pont sikeres felvétele után a rendszer a kalibrációs paramétereket elmenti. Az automatikus kalibráció közben a felhasználó egy kalibrációs mintát (10. ábra) néz, melyet a rendszer automatikusan felismer. A kezelőnek ebben az esetben csak megerősítenie kell, hogy a felhasználó a megfelelő kalibrációs pontra néz a táblán.

A vizsgálatok során kiderült, hogy a kalibráció gyorsan elvégezhető. Mindkét kalibráció típusához szükséges időt vizsgáltuk a gyakorlati mérések során is, az eredmények összevetésre kerültek a piacon igen elterjedt ASL Mobil Eye mobilis tekintetkövető rendszer kalibrációs időivel. A 2. táblázat mutatja a minimális (T<sub>min</sub>) és az átlagos (T<sub>avg</sub>) kalibrációs időket.

A szoftver-erőforrás felhasználása is tesztelésre került. A 3. táblázatban látható számítógép konfigurációkon vizsgáltuk a szoftvert.

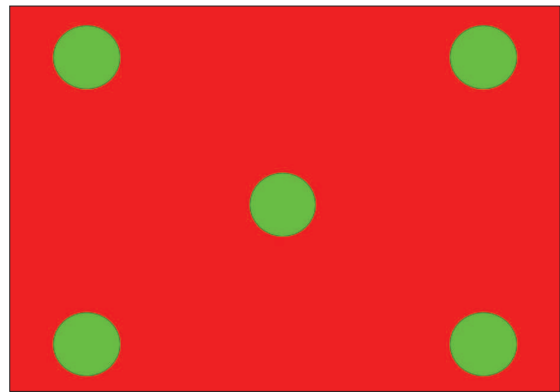
Az erőforrások felhasználása szempontjából három figyelemre méltó esetre bontható a vizsgálat: (i) a pupilla-detektálás, (ii) az automatikus kalibráció és (iii) a felvétel (mérés) esete.

Az automatikus kalibráció alkalmával ugyanis a tárgykamera képe is feldolgozásra kerül, ami növeli az erőforrás felhasználás mértékét. Mérés közben pedig a pupilla paraméterek meghatározása mellett a videófolyam tömörítésére és merevlemezre írására is szükség van, ami szintén extra erőforrás felhasználásával jár.

Az egyes konfigurációk átlagos processzor használatát mutatja százalékosan a 4. táblázat, a felhasznált fizikai memória mennyiségét pedig az 5. táblázatban olvashatjuk. Érdemes megjegyezni, hogy a tesztek során a hagyományos algoritmusok a „B” és „C” konfigurációkon nem voltak használhatók 30 képkocka/másodperc sebességnél.

A pupilla követése miatt azon képkockákra, melyeknél ismert a pupilla előző helyzete és/vagy a pupilla helyzetbecslő helyesen lett kalibrálva, a detektálás a kép csak egy részén történik. Ekkor a processzorhasználat csökkenése volt megfigyelhető.

A mérések során a „B” és a „C” jelű konfigurációk átlagos üzemidejét is megvizsgáltuk. Mindkét konfiguráció-

10. ábra  
Automatikus kalibrációs minta

Konfig.	Típus	CPU	CPU sebesség	CPU magok	Fizikai memória	Operációs rendszer
A	Acer TimelineX	Intel Core i5	2.3 GHz	2	4 GB	Windows 7 (64bit)
B	Gigabyte BookTop	Intel Atom	1.6 GHz	1	1 GB	Windows 7 (32bit)
C	Gigabyte TouchNote	Intel Atom	1.6 GHz	1	1 GB	Windows XP (32bit)

3. táblázat  
A tesztekben használt hordozható számítógép konfigurációk

4. táblázat

A konfigurációk átlagos processzorhasználatát

Konfig.	Pupilla-detektálás	Automatikus kalibráció	Felvétel (mérés)
A	2,9%	3,6%	8,2%
B	42,8%	46,9%	47,3%
C	41,1%	46,2%	46,8%

5. táblázat

A konfigurációk átlagos fizikai memória használatát

Konfig.	Pupilla-detektálás	Automatikus kalibráció	Felvétel (mérés)
A	46 884 KB	48 012 KB	41 780 KB
B	32 904 KB	32 932 KB	32 544 KB
C	33 231 KB	33 272 KB	32 767 KB

ban megközelítőleg 4400 mAh kapacitású akkumulátor található, ezzel a rendszer átlagosan 92 perc folyamatos mérést tett lehetővé az akkumulátorok töltése nélkül.

## 7. Összefoglalás

Olyan mobilis tekintetkövető rendszerek számára dolgoztunk ki a pupilla paraméterek új mérési módszereit, amelyek esetén a hagyományos globális, küszöbözésen alapuló technikák rendre elbuknak. Az élpontok szűrésén és osztályozásán alapuló pupilladetektáló algoritmus lehetőséget biztosít a paraméterek robusztus meghatározására, melyre nagy szükség van a mozgó tekintetkövetésre jellemző változó fényviszonyok miatt. A bemutatott rendszert kutatási projektekből, ipari és kereskedelmi mérésekben egyaránt használatba vették, ahol a mérési módszer bizonyította alkalmazhatóságát a tekintetkövetés és a pupilla paraméter meghatározás feladataira.

A rendszer továbbfejlesztését két lépcsőben képzeljük el. A meglévő rendszerben infravörös sávszűrés használatát kívánjuk megvalósítani és a szoftvert a kialakult új mérési elrendezéshez igazítani. A rendszer robusztussága a szűrés segítségével tovább növelhető.

Majd egy új tekintetkövető eszköz kidolgozását tűzzük ki célul, mely a hibrid technológiát használja fel a pupillaparaméterek meghatározására. Ezzel lehetővé válhat a mobil tekintetkövetés beágyazott eszközökkel történő megvalósítása, mely nagymértékben növeli az eszköz hordozhatóságát. Egy ilyen eszköz a hibrid megoldás alacsony feldolgozási igényével felvértezve egy új kategóriát teremthet a mobil tekintetkövető rendszerek területén.

### A szerzőről



**HELFENBEIN TAMÁS** 2008-ban szerezte meg MSc fokozatát a Budapesti Műszaki és Gazdaságtudományi Egyetemen. 2006-2008-ig a Magyar Tudományos Akadémia Központi Fizikai Kutatóintézetében (KFKI) végzett nagyméretű adathalmazok információs- és adatbázis technológiájára vonatkozó kutatás-fejlesztési feladatokat. 2006-tól a Forrai Magániskola és Kéttannyelvű Középiskola óraadó tanára adatbázis- és információs rendszerek témakörében. 2008-2011-ig a BME doktori iskoláját végezte, emellett mindvégig a Bay Zoltán Alkalmazott Kutatási Közalapítvány (BZAKA) ösztöndíjas kutatója volt. Az egyetemen és az alapítványban a gépi látás és képfeldolgozás témakörre mellett számos vezeték nélküli szenzorrendszerekkel és ambiens intelligenciával kapcsolatos nemzetközi kutatási és fejlesztési projektben is részt vett. 2011-től a BZAKA kutatója, ahol a képfeldolgozás, gépi látás és a vezeték nélküli szenzorrendszerek témakörében végez kutatási tevékenységet.

### Irodalom

- [1] Plotkin, A., Shafir, O., Paperno E., Kaplan, D.M., "Magnetic Eye Tracking: A New Approach Employing a Planar Transmitter", IEEE Transactions On Biomedical Engineering, Vol. 57, No. 5., pp.1209–1215., 2010.
- [2] Lin, M., Li, B., "A Wireless EOG-based Human Computer Interface", 3rd International Conference on Biomedical Engineering and Informatics, pp.1794–1796., 2010.
- [3] Gang, W., "Gaze Tracking Using One Fixed Camera", 7th International Conference on Control, Automation, Robotics And Vision, pp.1409–1414., 2002.
- [4] Morimoto, C.H., Koons, D., Amir, A., Flickner, M., "Pupil detection and tracking using multiple light sources", Image and Vision Computing 18, pp.331–335., 2000.
- [5] Duda, R.O., Hart, P.E., "Use of the Hough Transformation to Detect Lines and Curves in Pictures", Comm. ACM, Vol. 15, pp.11–15., 1972.
- [6] Bastos, C.A. C.M., Ren, T.I., Calvalcanti, G.D.C., "A Combined Pulling & Pushing and Active Contour Method for Pupil Segmentation", IEEE Trans.: Acoustics Speech and Signal Proc., pp.850–853., 2010.
- [7] Canny, J., "A Computational Approach To Edge Detection", IEEE Trans. Pattern Analysis and Machine Intelligence, 8(6), pp.679–698., 1986.
- [8] Kalman, R.E., "A New Approach to Linear Filtering and Prediction Problems", Transactions of the ASME – J. of Basic Engineering, 82 (Series D), pp.33–45., 1960.