



Politechnika Gdańska
Wydział Elektroniki, Telekomunikacji i Informatyki
Katedra Systemów Multimedialnych

ul. Narutowicza 11/12, 80-952 Gdańsk; tel. (58) 347-13-01; fax (58) 347-11-14;
e-mail: ksm@sound.eti.pg.gda.pl <http://www.multimed.org/>

RAPORT WEWNĘTRZNY
z realizacji pracy badawczo-wdrożeniowej

tytuł raportu:

**Raport przejściowy pracy nad
SYSTEMEM BADANIA FIKSACJI WZORKU
(część dotycząca algorytmu)**

nr raportu MM001/08

praca zrealizowana w ramach projektu: Multimodal

W okresie od 21-02-2008 do 06-05-2008

Kierownik projektu: prof. dr hab. inż. Andrzej Czyżewski

Opiekun projektu: mgr inż. Bartosz Kunka

Autor: **Mariusz Kurkowski**
mariuszk@sound.eti.pg.gda.pl

1. Spis treści

2. Wstęp
3. Cel realizowanej pracy
4. Opis merytoryczny pracy
5. Stan aktualny algorytmu
6. Podsumowanie
7. Bibliografia
8. Załączniki

2. Wstęp

Praca powstaje w grupie 4 osobowej w której skład wchodzi:

Mgr inż. Bartosz Kunka, mgr inż. Maciej Kulesza, Rafał Rybacki i Mariusz Kurkowski. Dwóch ostatnich realizuje jednocześnie na podstawie ów pracy dyplom magistra.

3. Cel realizowanej pracy

Postawionym celem jest zrealizowanie w pełni funkcjonalnego algorytmu śledzenia punktu fiksacji wzorku, działającego na hardware budowanym równolegle przez członka zespołu.

Z założenia system w pierwszej fazie ma funkcjonować w środowisku Matlab. Równolegle odpowiednio część softwarowa i hardware mają być do siebie dopasowywane. W kolejnych fazach pracy nastąpi dopracowywanie algorytmu, by w ostatniej fazie zaimplementować i zoptymalizować algorytm w środowisku C++. Ma on być oddzielną aplikacją dostępną dla przeciętnego użytkownika, dostosowaną bezpośrednio do tworzonego sprzętu (kamera, 5LED...).

Streszczenie części dyplomowej:

W pracy dokonano przeglądu dyplomowej dokonano przeglądu w zakresie inwazyjnych i nieinwazyjnych metod śledzenia punktu fiksacji wzroku. Dodatkowo przedstawiono przykładowy system do prowadzenia ćwiczeń zręcznościowych przy pomocy komputera. W części praktycznej zaimplementowano wybrane algorytmy wykorzystujące metody przetwarzania obrazu oraz opracowano prototypowe stanowisko komputerowe, służące nieinwazyjnemu śledzeniu punktu fiksacji wzroku na monitorze komputerowym przy pomocy kamery internetowej działającej w zakresie podczerwieni.

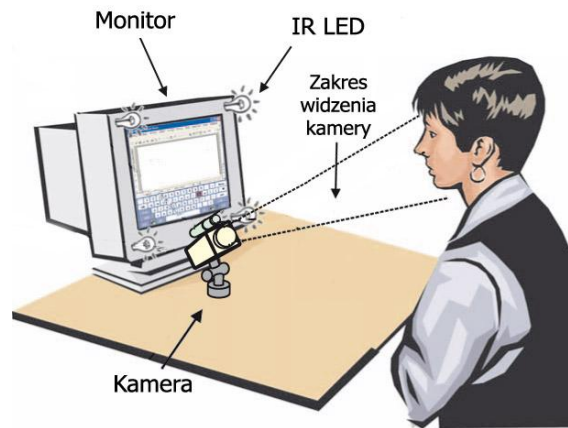
4. Opis merytoryczny pracy

Typowe systemy wykrywają położenie gałki ocznej i jej orientację, co przekłada się na niebezpieczeństwo niedokładnych pomiarów, a co za tym idzie dużą niewiarygodność miejsca punktu fiksacji wzroku. Dodatkowo początkowe systemy, oparte w szczególności o źródło IR, wymagały stałej, a co więcej, znanej odległości głowy od kamery. W proponowanym przez nas rozwiązaniu nie potrzebujemy informacji na temat pozycji oka, system jest w dużym stopniu odporny na ruchy głowy. Jest to niewątpliwie wielką jego zaletą, zwłaszcza przy zastosowaniu systemu do terapii ludzi młodych. Co ważne, system ten nie wymaga utrzymania niezmienniej odległości głowy od monitora - źródeł IR.



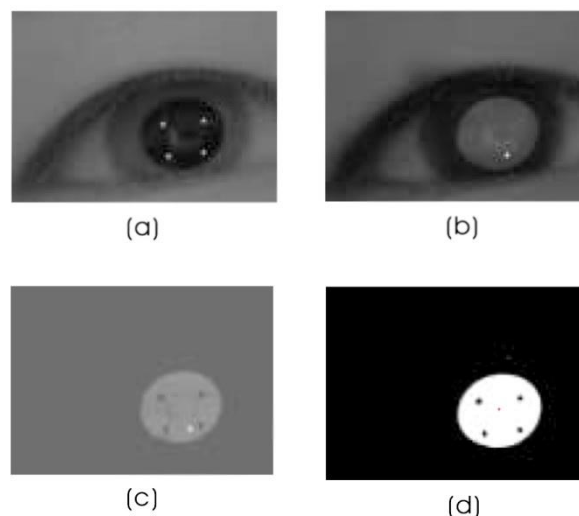
System zbudowany z 5 LED IR oraz 2 kamer [111]

Cztery, z pięciu źródeł IR, umieszczone są dokładnie na narożnikach ekranu. Dzięki nim uzyskujemy odbicia na powierzchni oka - glinty. Piąta natomiast znajduje się jak najbliżej osi kamery, służy do uzyskania tzw. jasnego oka (z ang. bright-eye effect). Źródła światła są na przemian włączane i wyłączane, tworząc 5 glintów na powierzchni oka. Na podstawie odbić oraz środka źrenicy zostaje wyznaczony punkt fiksacji wzroku [111].



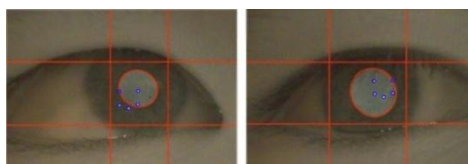
Budowa stanowiska [222]

Używana kamera ma za zadanie rejestrować obraz uzyskany na powierzchni oka, rejestruje glinty. I tak na odbieranym obrazie widoczne są odbicia z 4 LED umieszczonych na narożnikach monitora, oraz tzw. jasna źrenica wraz z widocznym glintem, pochodzącym od źródła oświetlenia znajdującego się na osi kamery. Dokonując prostej operacji matematycznej, dostajemy obraz, który następnie po wstępnym progowaniu, przedstawiany jest jako obraz binarny.



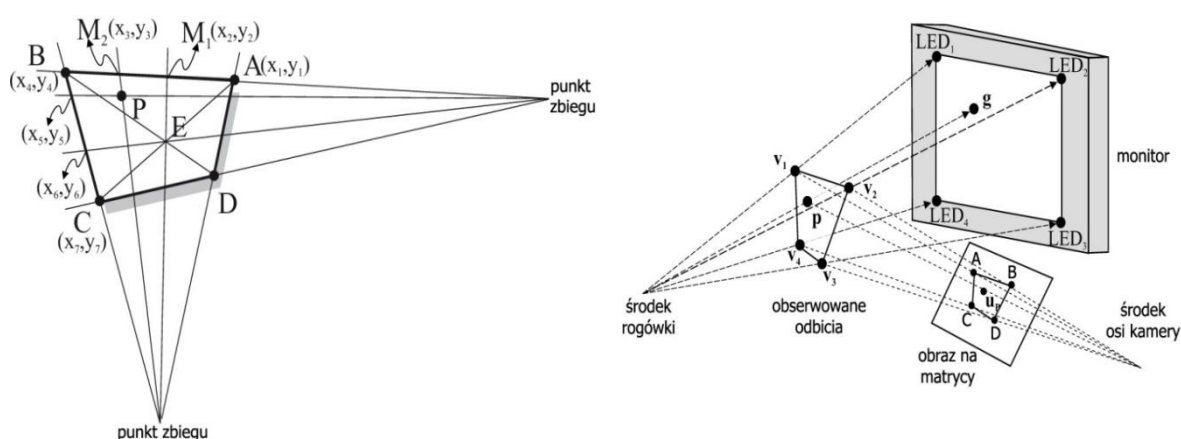
Obraz rejestrowany przez kamerę; a-ciemna źrenica, b-jasna źrenica, c-obraz różnicowy, d – obraz różnicowy binarny, z punktem oznaczającym środek źrenicy [222].

Czasami bywa tak, że w zakwizowanym obrazie glinty wychodzą poza obszar źrenicy. Jednak dzięki modyfikacją algorytmu opisywanego w artykule z *IEEE Non-intrusive Eye Gaze Estimation using a Projective Invariant under Head Movement*, proponowany przez nas system jest nieczuły na tego typu sytuacje.



Możliwe nieprawidłowości [102]

Gdy osoba obserwuje monitor, punkt fiksacji jest zawsze w wielokącie tworzonym przez uzyskane z odbić punkty referencyjne –A, B, C, D, zgodnie z rys 3.8. Metoda ta upraszcza znacznie obliczenia związane z wyliczaniem geometrycznych zależności pomiędzy okiem, kamerą i monitorem. Korzysta z zależności położenia odbić promieni LED na powierzchni tęczówki i środka źrenicy w wielokącie, a ich odzwierciedleniami na ekranie.



Schemat projekcji systemu 5LED [222]

LED_{1,2,3,4} – źródła IR LED, umieszczone na narożnikach ekranu,

A,B,C,D – kolejne gliny, zarejestrowane przez kamera na podstawie obserwowanych na powierzchni oka ludzkiego punktów referencyjnych V₁,V₂,V₃,V₄,

P – środek źrenicy,

E – punkt przecięcia prostych łączących przeciwległe gliny w wielokącie,

M_{1,2,3,4} – punkty tworzone przez proste przechodzące przez odpowiednie punkty zbiegu oraz wcześniej wyznaczone punkty E i P,

Punkty zbiegu – punkty przecięcia prostych łączących przyległe gliny,

g – punkt fiksacji wzroku (x_g,y_g),

x_g,y_g –współrzędne punktu fiksacji wzroku.

W celu wyliczenia punktu fiksacji wzroku na monitorze komputera, należy postępować zgodnie z niżej opisanymi równaniami. By obliczyć współrzędną odciętą należy rozwiązać równanie 1.1 i kolejne, kończąc na równaniu 1.3.

Na początku należy wyliczyć stosunek punktów M_1, M_2, A, B :

$$CRx = \frac{(x_1y_2 - x_2y_1)(x_3y_4 - x_4y_3)}{(x_1y_3 - x_3y_1)(x_2y_4 - x_4y_2)} \quad \text{Równanie 1.1}$$

W podobny sposób wyliczamy, stosunek wymiarów ekranu:

$$CRx = \frac{\left(\frac{w-x_g}{2}\right)x_g}{\left(\frac{w-x_g}{2}\right)\frac{w}{2}} = \frac{x_g}{w-x_g} \quad \text{Równanie 1.2}$$

Ponieważ oba stosunki są równe, otrzymujemy:

$$x_g = \frac{w \cdot CRx}{1 + CRx} \quad \text{Równanie 1.3}$$

co stanowi współrzędną odciętą punktu fiksacji wzroku na monitorze.

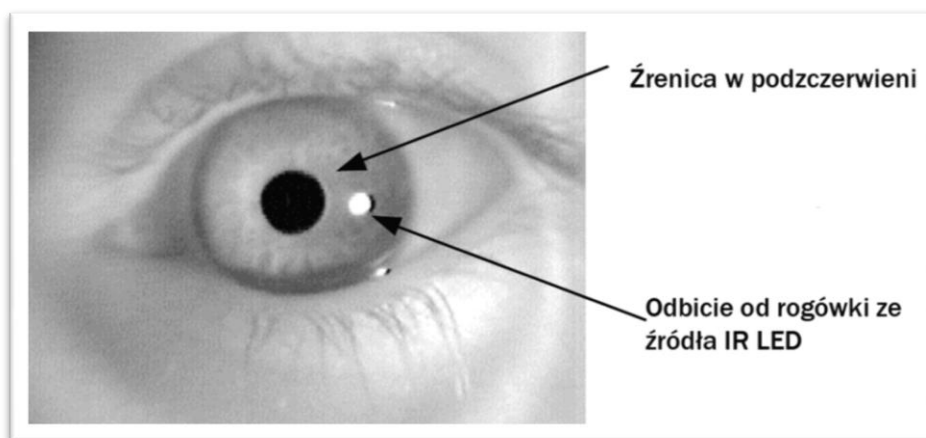
Analogicznie wyliczamy kolejną współrzędną:

$$y_g = \frac{h \cdot CRy}{1 + CRy} \quad \text{Równanie 1.4}$$

Na podstawie uzyskanych pomiarów, skorygowanych wcześniej o dane uzyskane w procesie kalibracji, dostajemy współrzędne punktu fiksacji wzroku na monitorze komputera.

Źródło oświetlenia

W systemie wykorzystane jest zjawisko odbicia światła od rogówki, które jest widoczne jako jasny punkt (ang. *glint*) – punkt referencyjny. Na podstawie ich obserwacji i pewnych zależności matematycznych, określa się punkt fiksacji wzroku



Oko w podczerwieni [333]

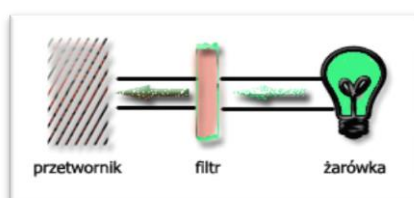
Wybór źródła oświetlenia powierzchni oka wydawała się być stosunkowo oczywisty. Większość istniejących rozwiązań opiera się o źródło bliskiej podczerwieni. Najlepiej gdy będzie to długość 850nm, ponieważ krzem, z którego wykonuje się obiektywy kamer ma maksimum czułości dla tej długości fali.

Zaletami używania promieni IR są min:

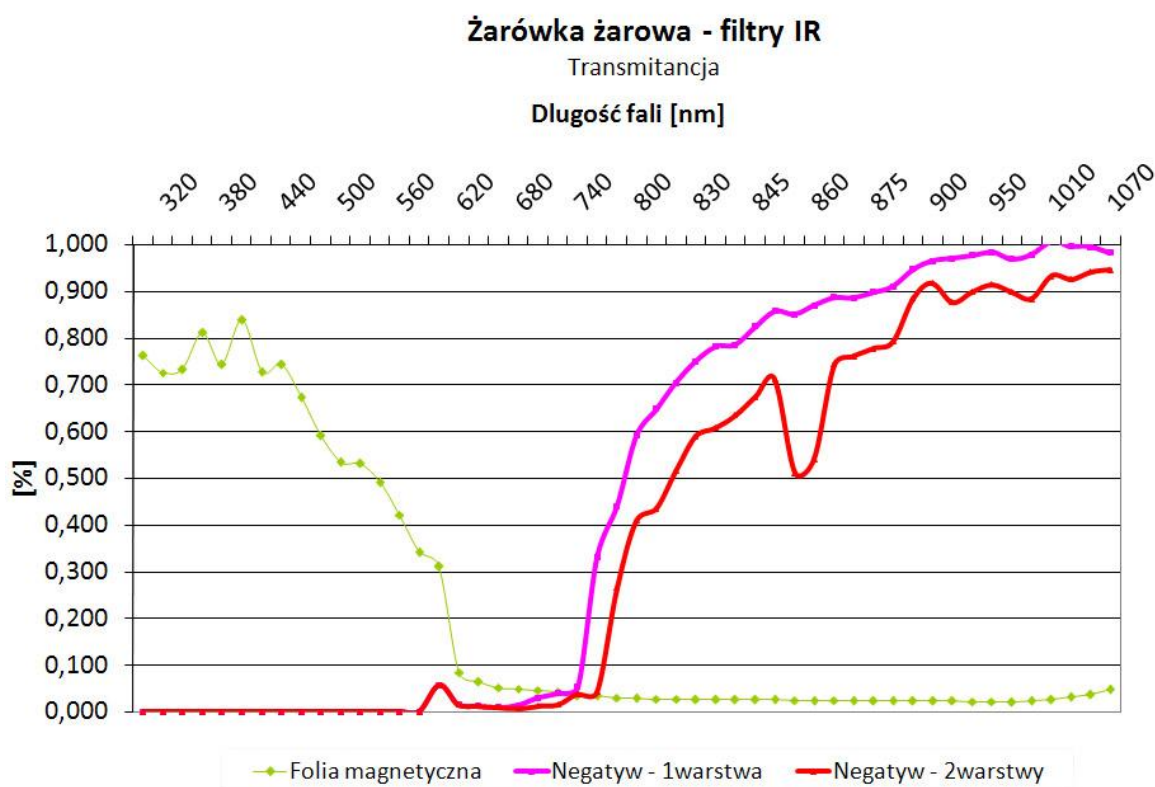
- ✦ względna niewrażliwość oka ludzkiego na pasmo bliskiej podczerwieni,
- ✦ lepszy kontrast poszczególnych składowych klatki, a co za tym idzie łatwiejsze wykrywanie brzegów poszczególnych części oka oraz glintów.
- ✦ dostępność źródeł LED,
- ✦ czułość matrycy kamery pozwalająca rejestrować dany zakres fali,
- ✦ relatywnie proste stworzenie urządzenia rejestrującego odbicia – kamery.
- ✦

Musimy uważać, by natężenie promieniowania nie było zbyt wysokie, aby nie uszkodzić rogówki oka osoby użytkującej system. Więcej na ten temat możemy przeczytać w polskich normach dotyczących pracy z laserami [555].

Ważnym aspektem przy wyborze źródła oświetlenia była czułość kamery na dane pasmo promieniowania. Zostały przeprowadzone badania z użyciem prostych filtrów: negatywu kliszy fotograficznej oraz foli magnetycznej z dyskietki komputerowej 5 i 1/4 cala. Schematycznie przebieg badania przedstawiono na rys.3.10. Źródłem promieniowania była żarówka żarowa. Strumień światła przechodzący przez specjalnie przygotowany kanał, tak aby do przetwornika nie docierały inne fale, trafia na badany filtr. Następuje odcięcie niepotrzebnych składowych widma promieniowania, po czym analiza uzyskanych wartości. Otrzymane wyniki przedstawione zostały na wyk. 1.



Badanie filtrów



Pomiar charakterystyk filtrów w paśmie widzialnym i bliskiej podczerwieni

Z uzyskanych pomiarów wynika, że nie ma potrzeby używania dedykowanych filtrów, gdyż nie potrzebujemy odcięcia całego pasma widzialnego. Jedna warstwa negatywu kliszy fotograficznej w pełni wystarcza do zastosowań w systemie.

Kalibracja systemu

Położenie punktu odbicia nie ulega zmianie przy ruchach oka, jednak przy ruchach głowy następują pewne odchyłki. Stąd by zniwelować efekty ruchu głowy stosuje się określone wartości korekcyjne, wykorzystywane w samym algorytmie w celu poprawy dokładności systemu. Parametry przekształcenia definiowane są w procesie kalibracji.

Odpowiednie wartości kalibracyjne potrzebne są również ze względu na krzywizny płaszczyzny oka. Powierzchnia oka nie jest płaska, stąd w procesie kalibracji systemu należy bezwarunkowo pamiętać o uwzględnieniu tego faktu.

Kalibracja systemu jest również potrzebna choćby z samego faktu wprowadzenia szerokości i wysokości ekranu monitora. Jest to najczęściej operacja automatyczna, pobierana na podstawie rozdzielczości monitora.

Budowa stanowiska

Stanowisko powinno być sporządzone wg pewnych określonych zasad. Niespełnienie zaleceń może znacznie wpłynąć na dokładność systemu.

System zbudowany jest z 4 diod umieszczonych na monitorze, w wersji roboczej przyklejanych na przylepcu opasającym krawędzie monitora. Kamera ustawiona powinna być pod monitorem w niedalekiej jego odległości, najlepiej w połowie jego szerokości...



Stanowisko robocze, widoczna 'jasna źrenica'

Założenia sprzętowe

Kamera

Podstawowym założeniem jest wykorzystanie powszechnie dostępnej kamery internetowej, o rozdzielczości co najmniej 720lini, nie posiadającej w sobie filtra odcinającego pasmo podczerwone. Ewentualnym rozwiązaniem jest, usunięcie takiego filtra. Do kamery montowany będzie nowy filtr przepuszczający pasmo bliskiej podczerwieni, zwłaszcza w okolicach promieniowania użytych diod $\sim 850\text{nm}$.

Diody

Używane diody powinny promieniować z odpowiednią mocą, tak aby nie stanowić zagrożenia dla osoby badanej, zgodnie z polskimi normami dla laserów klasy 1 lub 1M, jednocześnie na tyle mocno by odpowiednio oświetlić oko. Maksymalna dopuszczalna ekspozycja żywej tkanki, szczególnie czułego oka, nie może zostać przekroczona. Dodatkowo kąt promieniowania wybranych diod musi zapewniać dobre oświetlenie oka, nawet w sytuacji odchylenia głowy w bok.

Stanowisko komputerowe

Zapewnione musi zostać stanowisko wyposażone w monitor LCD, USB 2.0 oraz odpowiedniej klasy jednostkę obliczeniową typu PC z systemem operacyjnym Windows.

Minimalne wymagania sprzętowe: Intel Core duo / Athlon x2, 1Gb Ram, 100mb HDD

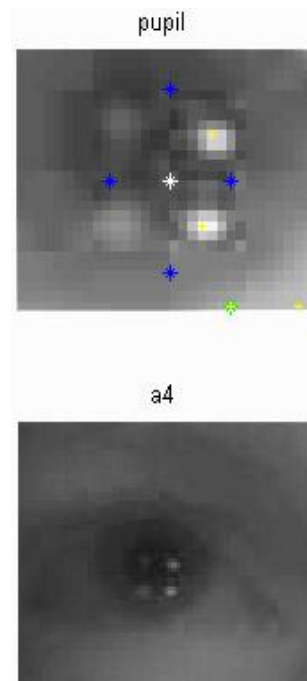
5. Stan aktualny algorytmu

Aktualnie algorytm gdy zostanie mu podana klatka z jasną i ciemną źrenicą niezawierającą dodatkowych artefaktów radzi sobie bardzo dobrze z wykrywaniem glintów i dość dobrze ze środkiem źrenicy. Natomiast w przypadku dodatkowych artefaktów lub zbyt dużego rozjaśnienia twarzy wynik działania algorytmu jest często niezadowalający. Cały czas trwają prace nad wprowadzaniem poprawek jednocześnie ze strony hardware jak i software.

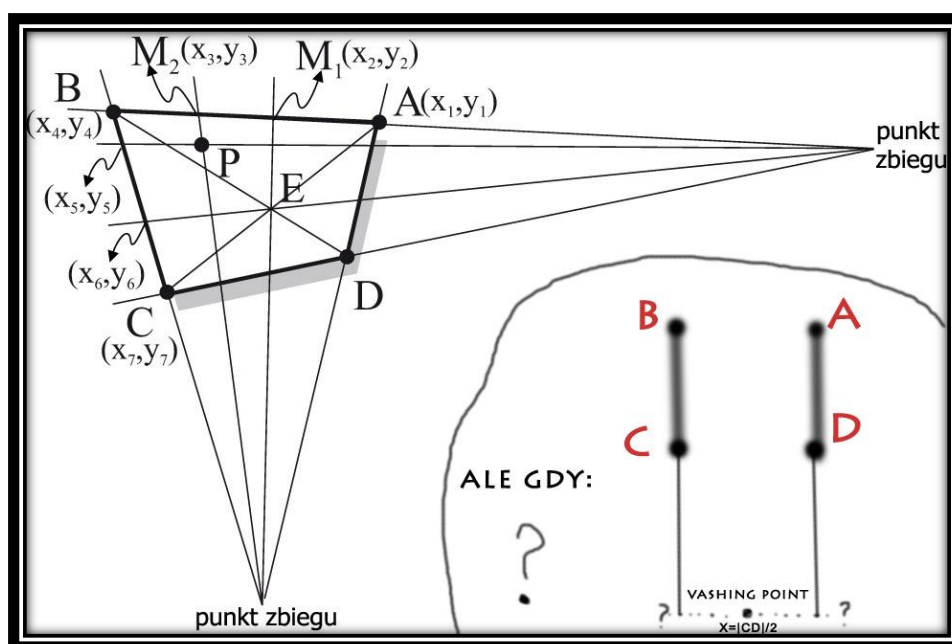
Algorytm pokazuje w przybliżeniu punkt fiksacji wzorku, jednak nie zawsze w pełni poprawny. Jest cały czas w fazie testowania i dopracowywania oraz dostosowywania do akwizowanego z kamery IR obrazu, często niestety z dużymi artefaktami. Dodatkowo trwają prace nad stosowaniem kamer różnej rozdzielczości i soczewek powiększających rejon zainteresowania co wpływa niekorzystnie na sam algorytm. Za każdym razem trzeba nieco zmieniać pewne wartości progowe stosowane przy wykrywaniu obszaru oka i podobnych. Więcej wywodów na ten temat znajduje się dalej.

Poniżej znajdują się przeprowadzone wstępnie wyniki oparte o algorytm 4glintowy. Przypomnę że najważniejszym do wykrycia jest środek źrenicy, glinty są o wiele mniej ważne. Wykrycie dokładne środka źrenicy spędza sen z powiek, jak z resztą widać... Nawet najmniejsze przesunięcie środka wpływa niekorzystnie na wyniki. Tutaj też kłania się proces kalibracji w którym to powinno się dokładnie przeanalizować położenie 4 glintów na obrazie i na podstawie ich położenie korygować wyniki. 'Opisane' w *Non-intrusive Eye Gaze Estimation without Knowledge of Eye Pose.pdf*. Stąd też wyniki są nieco przekłamane, ale da się zauważyć powtarzalność błędu, I miejmy nadzieję da to się skorygować (przesunąć wynik o pewną wartość) po procesie kalibracji.

Dodatkowo jak już pewnie każdy z członków zespołu zdażył się zorientować w uzyskanych zdjęciach często są artefakty. Próbuje z nimi walczyć ‘softwarowo’ aczkolwiek niektóre wprowadzają znaczne błędy i są praktycznie nie do obejścia w tego typu algorytmie. Przypomnę że zgodnie z początkowymi wersjami algorytmu, ROI – oko, wykrywamy na podstawie najjaśniejszego punktu obrazu i później pierwszy z glintów na podobnej zasadzie. Niestety często są dodatkowe odbicia które powodują znaczne utrudnienia w wyznaczeniu obszaru z okiem. Tu dodam że można stosować projekcję poziomą i pionową w celu znalezienia oka, jednak jest to wolne, jak i również z podobnymi błędami w dalszej kolejności. Kolejnym utrudnieniem jest to że często dolna część policzka – zaraz pod okiem jest bardziej oświetlona niż niektóre glinty. Co jest dużym problemem w przypadku niepoprawnego ROI (przy tego typu artefaktach jest to dość częste, tak że obszar zainteresowania zawiera też pewną część wybiegającą nieco za same oko – dolną powiekę). Widać to na zdjęciu obok.



Jak można się też domysleć, w dokumentacji z IEEE na której się opieramy nie napisali zbyt dużo szczegółów, stąd też trzeba posługiwać się sporą fantazją w budowie algorytmu, co nie jest takie proste w pewnych przypadkach. Stąd też zakładałam pewne błędy pomiarowe zwłaszcza przy wykrywaniu tzw. punktów *vashing points* (punkty zbieżności). Poniższy rysunek może coś wyjaśnić. Tworzone są pewne punkty wirtualne na podstawie położenia glintów. Gdy glinty są o podobnych współrzędnych tak że układają się w pewnych charakterystyczny sposób, można tylko zgadywać jak obliczono współrzędne vashing points... Jedna z nich jest oczywista (środek), ale druga...



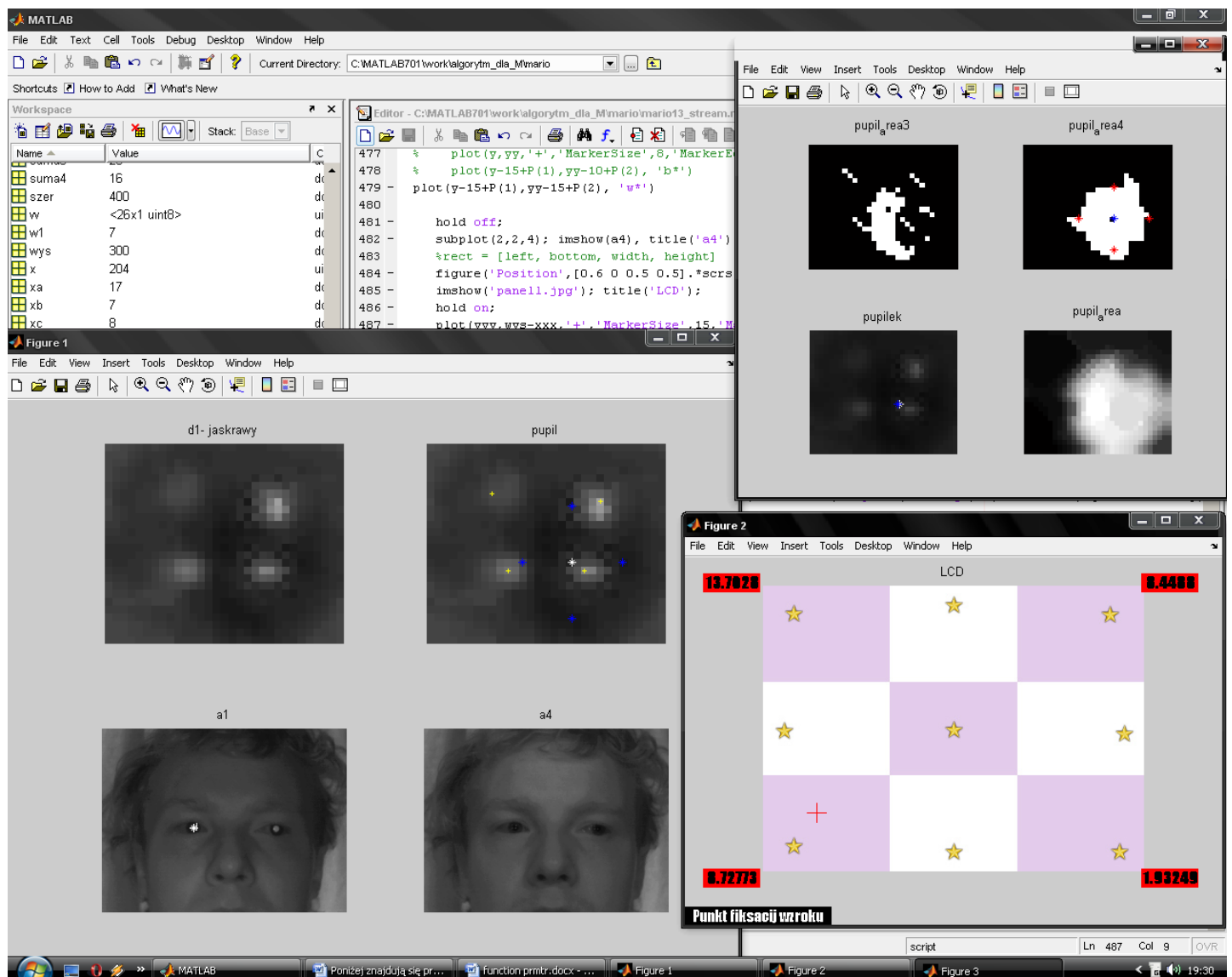
Poniżej znajdują się otrzymane wyniki na podstawie zdjęć 800x600 x 12

W czerwonych narożnikach znajdują się obliczone odległości od wyznaczonego punktu fiksacji do odpowiednich glintów (przy odczycie zamieniać horyzontalnie).

W obrazie 'pupil' na niebiesko pokazane punkty skrajne elipsy opisanej na wykrytej źrenicy (obszar pupil_area4 jest to obszar wyznaczony na podstawie binaryzacji i zamknięcia obszaru). Natomiast na biało wykryty dwiema metodami środek źrenicy. Na żółto zaznaczono wykryte glinty.

Po nazwie obrazu można odczytać gdzie osoba patrzy. I tak nazwa 12_800_c_11.jpg oznacza że patrzy w lewą stronę na wysokości środka monitora: c-center, l-left...

(obraz ma lepszą jakość po powiększeniu).



próbka 12_800_d_11.jpg

6. Podsumowanie

W ramach projektu został napisany wstępnie algorytm służący wykrywaniu punktu fiksacji wzroku na monitorze komputera. W chwili obecnej cały czas trwają prace w celach dopracowania algorytmu, udoskonalenia wykrywania środka źrenicy, zmniejszenia czułości algorytmu na powstałe przy akwizycji obrazu zakłócenia – artefakty, próbach analizy drugiego oka gdy analizowane okaże się obciążone zbyt dużą liczbą artefaktów, zmniejszenia naświetlenia okolic oka, udoskonaleniu obrazu różnicowego i inne. Dodatkowo istnieje zależność działania algorytmu od równoległe budowanej części hardware systemu.

7. Bibliografia

- [111] D.H. Yoo, M.J. Chung, D.B. Ju, I.H. Choi. Non-intrusive Eye Gaze Estimation using a Projective Invariant under Head Movement
- [222] D. H. Yoo and M. J. Chung. Non-intrusive Eye Gaze Estimation without Knowledge of Eye Pose
- [333] J. S. Babcock and J.B. Pelz. Building a lightweight eye tracking headgear
- [444] J.Kim, K. Park, G. Khang. A Method for Size Estimation of Amorphous Pupil in 3-Dimensional Geometry
- [555] POLSKA NORMA. PN-EN 60825-1
- [666] R. M. Mimica and C.H. Morimoto. A Computer Vision Framework for Eye Gaze Tracking
- [777] Z. Zhu and Q. Ji. Eye gaze tracking under natural head movements
- [888] J. Wang, R. Venkateswarlu, E. Sung. Eye Gaze Estimation from a Single Image of One Eye
- [999] S. Rakshit, D. M. Monro. Iris image selection and localization based on analysis of specular reflection

8. Załączniki

W plikach wymienionych w poniższej tabeli zawarto pliki źródłowe oraz próbki na których zostają prowadzone testy.

Tab. 8.1 Lista załączników

Nazwa i adres pliku	Zawartość
funkcje_fiksacja.rar http://www.bestq.ovh.org/funkcje_fiksacja.rar	M. Files MATLAB
probki_9.zip http://www.bestq.ovh.org/probki_9.zip	Próbki