

Tomasz Kryjak*, Jacek Chodak**

Wykorzystanie efektu jasnej i ciemnej źrenicy w wizyjnym śledzeniu ruchu gałek ocznych***

1. Wprowadzenie

Aplikacje śledzące ruch gałek ocznych wykorzystywane są w wielu dziedzinach i mogą zostać podzielone na dwie kategorie: diagnostyczne i interaktywne. W pierwszym przypadku urządzenie śledzące ruch gałki ocznej (*eye-tracker*) dostarcza różnorodnych informacji związanych z ich ruchem u osoby badanej. W drugim przypadku *eye-tracker* stanowi element interfejsu człowiek-komputer (HCI – *Human Computer Interface*).

W podejściu diagnostycznym wykorzystuje się praktycznie tylko trajektorie ruchu gałek ocznych zarejestrowane jako odpowiedź na konkretne bodźce: prezentowane obrazy, filmy, bądź specjalnie zaprojektowane badanie, np. wodzenie za punktem lub czytanie zniekształconego tekstu. Analiza danych nie musi odbywać się w czasie rzeczywistym i zazwyczaj przebiega po zarejestrowaniu wyniku eksperymentu. Dziedziny, w których wykorzystuje się *eye-tracking*, to neurologia i psychologia (analiza sposobu przetwarzania informacji wizyjnej w korze wzrokowej, analiza zjawiska uwagi [3], analiza czytania, analiza percepcji sceny wizyjnej, analiza zagadnienia wyszukiwania obiektów na scenie oraz analiza pewnych szczególnych typów obrazów np. twarzy, sylwetek itp.). Innym przykładem zastosowania *eye-trackingu* jest inżynieria (lotnictwo – analiza zachowań pilotów w symulatorze, analiza prowadzenia samochodu, projektowanie wyświetlaczy, desek rozdzielczych, kokpitów), marketing i reklama (analiza reklam, stron internetowych, okładek czasopism, gazet, opakowań, ułożenia towaru na półkach sklepowych).

Podejście interaktywne to przede wszystkim tworzenie interfejsów do komunikacji człowieka z komputerem: od prostych, w których kierunek patrzenia „zastępuje” ruch kursora myszy, po bardziej złożone umożliwiające np. pisanie tekstu bądź sterowanie progra-

* Katedra Automatyki, Akademia Górniczo-Hutnicza w Krakowie, kryjak@agh.edu.pl

** Wydział Fizyki i Informatyki Stosowanej, Akademia Górniczo-Hutnicza w Krakowie, chodak@novell.ftj.agh.edu.pl

*** Praca była finansowana ze środków AGH (umowa AGH nr 11.11.120.612 – pierwszy autor), a w przypadku drugiego autora ze środków na naukę w latach 2009–2012 jako projekt badawczy nr N N518 426736

mami/urządzeniami. Interesujące jest też zastosowanie *eye-trakera* do dynamicznego renderowania scen: z dokładnością zależną od punktu skupienia uwagi użytkownika. Interfejsy oparte o śledzenie ruchu oka to także często, jedna z niewielu, szansa dla osób z dużym upośledzeniem ruchowym na komunikację z otoczeniem. Szersze omówienie zastosowań *eye-trackerów* można znaleźć w [4].

Metody śledzenia ruchu gałek ocznych można podzielić na [2, 4]:

- elektrookulograficzne (EOG),
- magnetyczne,
- fotoelektryczne,
- wizyjne.

Metoda EOG polega na pomiarze różnicy potencjałów, jaka występuje między rogówką a siatkówką. Podczas gdy oko się porusza, dipol z nim związany, również zmienia swoją pozycję, co umożliwia rejestrowanie sygnałów oko ruchowych. Główną wadą metody jest czułość rogówkowo-siatkówkowego potencjału na zmianę oświetlenia. Akwizycja sygnałów odbywa się za pomocą elektrod umieszczonych na skórze. Sygnał mierzony jest w stosunku do pozycji głowy, co wymusza użycie drugiego urządzenia do śledzenia ruchów głowy lub jej ustabilizowanie. Jako zaletę należy wymienić możliwość stosowania tej metody u dzieci oraz podczas snu.

Metoda magnetyczna polega na wykorzystaniu soczewki kontaktowej z wtopioną pętlą z cienkiego drutu i umieszczeniu jej w strumieniu zmiennego pola magnetycznego. Analiza otrzymanego sygnału umożliwia odtworzenie ruchu gałki ocznej. Obie powyższe metody charakteryzują się dość dużą inwazyjnością (konieczność naklejania elektrod lub umieszczania w oku soczewki kontaktowej) oraz pozwalają na pomiar położenia oka tylko względem głowy.

Metoda fotoelektryczna polega na pomiarze efektów odbicia światła (widzialnego lub podczerwieni) od powierzchni oka. Przykładem takich systemów są: OBER2 [15] oraz JAZZ Novo [5]. Metoda ta umożliwia rejestrację położenia ze znaczną częstotliwością (nawet 2000 Hz), jednak jest dość podatna na zewnętrzne zakłócenia i wymaga dedykowanego sprzętu.

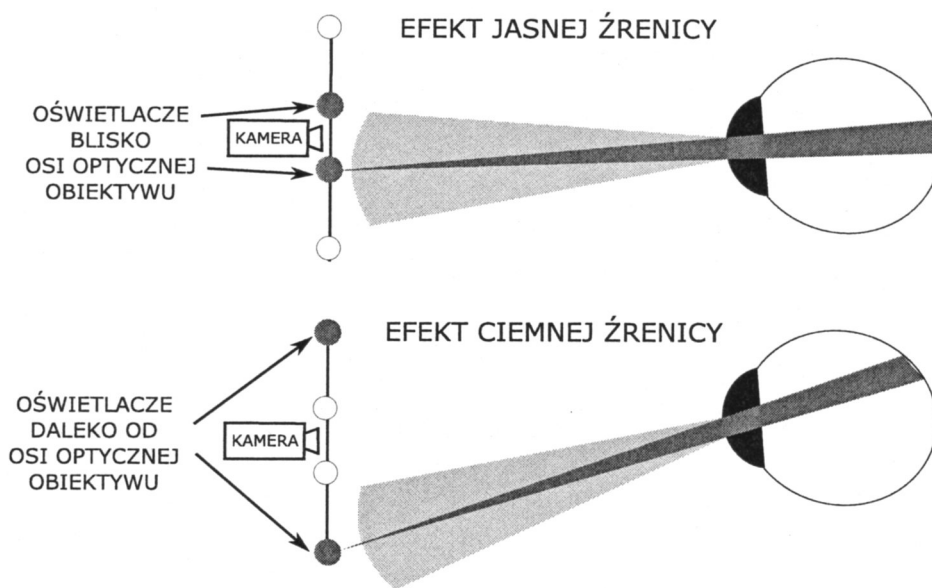
Najpowszechniej stosowaną metodą rejestracji ruchu gałki ocznej jest metoda wizyjna. Polega ona na obserwacji obrazu twarzy i oka za pomocą kamery i wyznaczaniu położenia źrenicy za pomocą algorytmów przetwarzania obrazów cyfrowych. Metoda ta często jest wykorzystywana w systemach dostępnych komercyjnie [1, 18, 21].

W artykule opisano system do śledzenia ruchu gałki ocznej, który wykorzystuje efekt jasnej i ciemnej źrenicy oraz ponadto rejestruje położenie oka z częstotliwością 50 Hz za pomocą standardowej kamery przemysłowej pracującej w systemie PAL. System stanowi kontynuację prac przedstawionych w artykule [8]. W rozdziale 2 omówiono efekt jasnej i ciemnej źrenicy oraz budowę specjalnego oświetlacza. Rozdział 3 przedstawia sposób wykorzystania półobrazów do uzyskiwania częstotliwości 50 Hz. W rozdziale 4 opisano

integrację oświetlacza z kamerą. W rozdziale 5 omówiono zagadnienie segmentacji źrenicy, a następnie w rozdziale 6 zaproponowano metodę śledzenia ruchów głowy. Rozdział 7 omawia rezultaty przeprowadzonych testów systemu i opis możliwych kierunków rozwoju.

2. Efekt jasnej i ciemnej źrenicy

Opisywany *eye-tracker* wykorzystuje efekt jasnej i ciemnej źrenicy, czyli fakt, że w zależności od sposobu oświetlenia, źrenica na obrazie z kamery jest czarna albo biała. Zjawisko ma swoją genezę w budowie i właściwościach fizycznych gałki ocznej. Jeżeli światło z oświetlacza, odbite od siatkówki oka, wróci do obiektywu kamery, to zarejestrowana źrenica będzie bardzo jasna. Sytuacja taka zachodzi wtedy, gdy oświetlacz umieszczony jest bardzo blisko osi optycznej obiektywu kamery. Efekt ten jest również znany w fotografii jako tzw. efekt czerwonych oczu. Jeżeli oświetlacz jest oddalony od osi optycznej obiektywu, to źrenica na obrazie będzie bardzo ciemna. Schematycznie zjawisko przedstawiono na rysunku 1. Omawiany efekt można też wykorzystać jako element wizyjnego interfejsu człowiek-komputer. Prototyp takiego urządzenia zaprezentowano w [16].

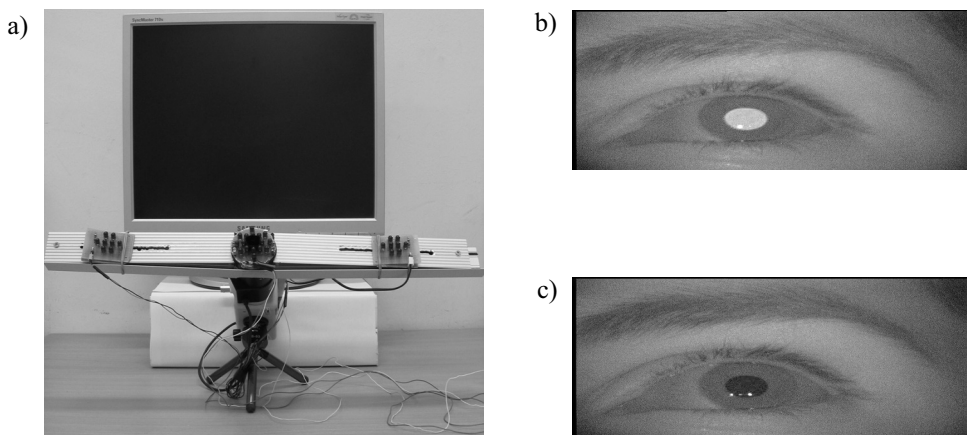


Rys. 1. Efekt jasnej i ciemnej źrenicy

W omawianym *eye-trackerze*, oświetlacz zrealizowano w oparciu o diody emitujące światło o długości fali $\lambda = 980 \text{ nm}$ (w tzw. dalszej podczerwieni). Światło podczerwone wykorzystywane jest zarówno w urządzeniach do śledzenia ruchu gałki ocznej (także komercyjnych) [23, 13, 21], jak i np. oświetlenia obrazu wykorzystywanego w biometrycznej

identyfikacji na podstawie tęczówki oka [10]. Spowodowane jest to następującymi czynnikami: światło podczerwone nie jest widoczne dla oka ludzkiego, zatem nie przeszkadza w prowadzeniu badania oraz dodatkowo, przy wykorzystaniu filtra odcinającego światło widzialne nałożony na obiektyw kamery, pozwala wyeliminować praktycznie wszystkie niepożądane refleksy np. od monitora, oświetlenia w pokoju itp. [8].

W celu wykorzystania efektu jasnej i ciemnej źrenicy skonstruowano oświetlacz składający się z trzech modułów – jednego centralnego w postaci pierścienia okalającego obiektyw (16 diod) oraz dwóch bocznych (8 diod każdy). Pierwsza zaproponowana wersja oświetlacza centralnego miała średnicę 38 mm i zakładana była na obiektyw. W trakcie prowadzonych wstępnych testów okazało się, że efekt jasnej źrenicy udaje się uzyskać, ale tylko wtedy gdy odległość od obiektywu do oka wynosiła ok. 2 m. Dla przewidywanego docelowego zastosowania urządzenia – analizy obrazów lub filmów wyświetlanych na monitorze komputera – okazała się ona za duża. Zdecydowano się zatem na wykonanie drugiego oświetlacza centralnego o średnicy 14 mm i zamontowaniu go na nakładce na obiektyw kamery. Przy tej konfiguracji efekt jasnej źrenicy udało się uzyskać przy znacznie mniejszej odległości oka od obiektywu (ok. 50 cm), co okazało się wystarczające dla przewidywanego zastosowania. Badanie wpływu parametrów oświetlacza i jego geometrii na uzyskiwanie efektu jasnej źrenicy szczegółowo zostało opisane w pracy [14]. Zdjęcie skonstruowanego oświetlacza i stanowiska badawczego zaprezentowano na rysunku 2a, a przykładowe obrazy jasnej i ciemnej źrenicy na rysunku 2b i 2c.



Rys. 2. Zdjęcie oświetlacza i stanowiska badawczego (a).
Przykładowy obraz jasnej (b) i ciemnej (c) źrenicy

Podstawową zaletą efektu jasnej i ciemnej źrenicy jest ułatwiona segmentacja obszaru źrenicy dzięki wykorzystaniu obrazu różnicowego. Konsekwencją ułatwionej segmentacji jest możliwość redukcji natężenia oświetlenia podczerwonego i większa niezawodność całego systemu.

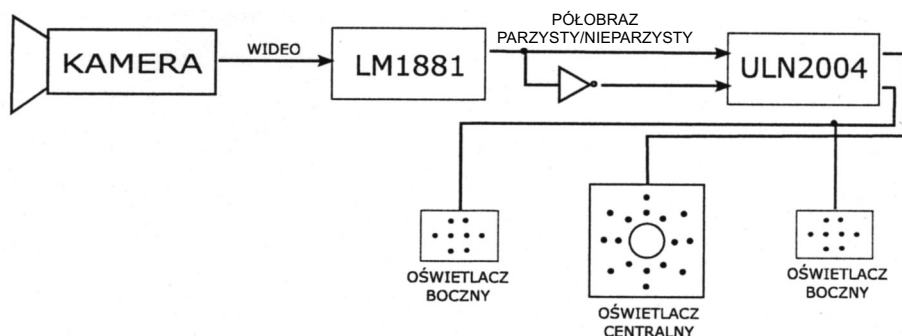
3. Rejestracja położenia oka z częstotliwością 50 Hz

Podstawowym mankamentem wizyjnego śledzenia ruchu gałki ocznej jest zwykle niska częstotliwość próbkowania systemu. Dla typowej, analogowej, kamery przemysłowej pracującej w systemie PAL można uzyskać najwyżej 25 klatek na sekundę. Dla popularnych i tanich kamer USB częstotliwość ta jest zwykle jeszcze niższa. Wprawdzie istnieją kamery, które są w stanie zarejestrować nawet 1000 klatek na sekundę [12, 17], ale są dość drogie, mają nietypowe interfejsy cyfrowe (np. Camera Link) i wymagają specjalnych urządzeń do akwizycji (*frame-grabberów*).

W opisywanym systemie zdecydowano się wykorzystać fakt, że kamera analogowa nie transmituje pełnej ramki obrazu, a dwa kolejne półobrazy (naprzemiennie linie parzyste i nieparzyste). Analizując półobrazy można, kosztem rozdzielczości pionowej, dwukrotnie zwiększyć częstotliwość próbkowania i uniknąć problemu usuwania przeplotu (*deinterlacing*). W systemie zastosowano monochromatyczną, przemysłową kamerę firmy YOKO (z czujnikiem CCD Sony Super HAD o rozdzielczości 600 linii) z obiektywem o ogniskowej 5–100 mm i przysłonie $F = 1,6$. Konieczność zastosowania czujnika monochromatycznego wynika z faktu, że czujniki kolorowe wyposażone są w filtr odcinający poczerwień.

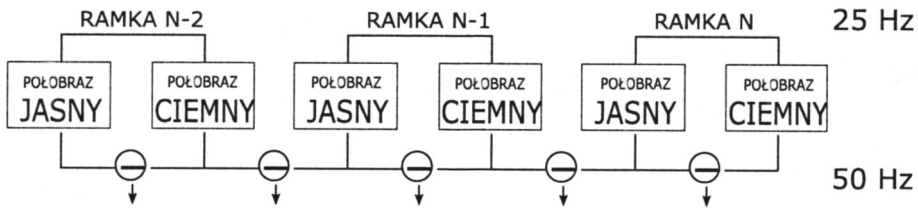
4. Integracja systemu

Opisane w rozdziałach 2 i 3 elementy, tj. oświetlacz pozwalający uzyskać efekt jasnej i ciemnej źrenicy oraz akwizycję z częstotliwością 50 Hz, można zintegrować [13]. Zdecydowano się tak sterować oświetlaczem, aby kolejne półobrazy zawierały naprzemienne efekt jasnej i ciemnej źrenicy. Do sterowania oświetlaczami zestawiono prosty układ elektroniczny (schemat na rys. 3). Jego podstawowym elementem był moduł LM1881 – separator sygnału synchronizacji z analogowego sygnału wideo. Na jednym z wyjść, moduł generował binarny sygnał, określający czy jest transmitowany półobraz parzysty, czy nieparzysty. Ponadto zastosowano jeszcze inwerter (GK72707N) i moduł Darlingтона (ULN2004AN) do załączania diod.



Rys. 3. Schemat sterownika oświetlaczami

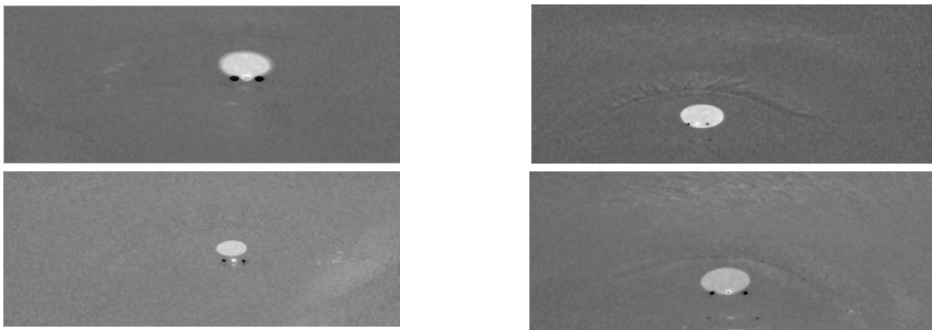
Przeprowadzone testy pokazały, że układ elektroniczny spełnia swoje zadanie i na kolejnych półobrazach widoczna jest naprzemiennie jasna i ciemna żrenica (przykład na rys. 2b i 2c). Ideę, w jaki sposób na podstawie półobrazów uzyskać częstotliwość próbkowania 50 Hz, zaprezentowano na rysunku 4.



Rys. 4. Uzyskiwanie częstotliwości 50 Hz z różnic półobrazów

5. Śledzenie żrenicy i odbłasku

Wykorzystanie efektu jasnej i ciemnej żrenicy znacząco ułatwia segmentację obszaru żrenicy. Poprzez odjęcie półobrazu z ciemną żrenicą od półobrazu z jasną żrenicą uzyskuje się wyróżnienie obszaru żrenicy i odbłasków od oświetlaczy (tzw. *glintów*, obrazów *Purkinjego*). Przykłady zaprezentowano na rysunku 5.



Rys. 5. Obraz różnicowy z widoczną żrenicą i odbłaskami

Wyraźny obraz żrenicy ma szereg konsekwencji dla projektowania algorytmów segmentacji. Przede wszystkim, niezależnie od wykorzystywanej metody, można stwierdzić, że wyniki wykrywania i śledzenia żrenicy będą dokładniejsze i mniej zakłócone. Dodatkowo zastosowanie mało złożonych obliczeniowo metod powinno dać dobre rezultaty. Najprostszy algorytm segmentacji żrenicy to zastosowanie binaryzacji połączonej z przetwarzaniem rezultatu za pomocą metod morfologicznych oraz indeksacją. Stosowane są też bardziej złożone rozwiązania, takie jak algorytm rozchodzących się promieni (*starburst*), detekcja krawędzi i dopasowanie elipsy [9].

W pierwszym etapie prac sprawdzono wpływ wstępnej filtracji obrazów na jakość otrzymywanego obrazu różnicowego i dalszą segmentację. Zastosowanie oświetlacza pracującego w dalszej podczerwieni i filtra odcinającego światło widzialne (obecnie w postaci prześwietlonej, kolorowej, kliszy fotograficznej) spowodowało, prócz minimalizacji wpływu zewnętrznych warunków oświetleniowych na przebieg badania, znaczne zaszumienie obrazów. Sprawdzono działanie szeregu znanych z literatury [20, 6] algorytmów filtracji: z maską uśredniającą, z maską Gaussa oraz medianową. Analiza wyników pokazała, że zadawalające rezultaty daje filtracja medianowa o rozmiarze okna 5×5 pikseli. Na dalszym etapie prac okazało się, że wstępna filtracja (całych półobrazów) jest jednym z bardziej czasochłonnych elementów algorytmu i zdecydowano się ją pominąć. Spowodowało to bardzo nieznaczne pogorszenie jakości segmentacji, a w sposób istotny przyspieszyło działanie całego algorytmu.

W opisywanej wersji systemu, zadaniem algorytmu jest śledzenie położenia środka źrenicy. Punkt ten można wyznaczyć jako środek ciężkości obszaru źrenicy lub jako środek elipsy, w której zawiera się obszar źrenicy. W przyszłości planowane jest dodanie śledzenia odbłasków od oświetlacza centralnego i bocznych – informacje o ich położeniu często wykorzystywane są w metodach kalibracji *eye-trackerów* [7]. Dla każdego automatycznego algorytmu śledzącego bardzo ważne są dwie sytuacje: początkowe znalezienie śledzonych obiektów (inicjalizacja, kalibracja – w idealnej sytuacji powinna być w pełni automatyczna) oraz obsługa zdarzenia polegającego na „zgubieniu” śledzonego obiektu – tj. wykryciu takiej sytuacji, a następnie podjęciu próby ponownego odnalezienia szukanego obiektu. W przypadku śledzenia ruchu gałki ocznej wspomniane zdarzenie występuje dość często: osoba badana co jakiś czas mruga, co powoduje zasłonięcie oka i „zgubienie” zarówno źrenicy, jak i odbłasków. Po drugie, przy niekorzystnym ułożeniu osoby badanej względem oświetlacza i jednoczesnym patrzeniu w górę, wszystkie trzy odbłaski mogą zaniknąć. Ponadto jeżeli dopuści się pewne minimalne ruchy głowy, to przy użytej kamerze, o rozdzielczości 720×576 pikseli nietrudno o zniknięcie oka z pola obserwacji. Ponieważ opisanych powyżej sytuacji raczej nie da się uniknąć, algorytm powinien zasygnalizować ich wystąpienie i odrzucić błędne fragmenty trajektorii ruchu oka.

Podczas prac przyjęto założenie, że pierwsza ramka analizowanej sekwencji jest prawidłowa: oko jest widoczne, odbłaski znajdują się w okolicy źrenicy oraz obraz jest względnie ostry. Dla takiej ramki obliczany jest obraz różnicowy i przeprowadzane jest poszukiwanie odbłasku od oświetlacza centralnego. W pierwszym kroku wyszukiwane jest maksimum poziomu jasności na obrazie i na jego podstawie określany jest próg binaryzacji – jako 0,8 wartości maksymalnej. Takiemu postępowaniu towarzyszy założenie, że analizowany odbłask jest najjaśniejszym punktem na obrazie różnicowym, co jest prawdziwe dla znakomitej większości analizowanych obrazów, a w szczególności dla obrazów uznawanych jako referencyjne. Zdarzają się jednak obrazy, na których wstępują odbłaski, np. od oprawek okularów itp. Opracowanie metod analizy dla tych sytuacji przewiduje się jako dalszy etap prac na systemem. W kolejnych krokach wykonywana jest binaryzacja

(z wyliczonym progiem), indeksacja oraz liczenie środków ciężkości obiektów oraz ich powierzchni. Ostatecznie za poszukiwany odbłask przyjmuje się obiekt, który ma powierzchnię mniejszą niż zadany próg (150 pikseli) i jednocześnie jest najjaśniejszy

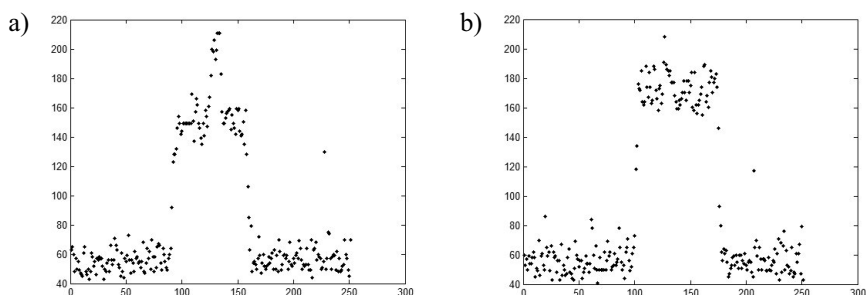
Wyznaczone położenie odbłasku od oświetlacza centralnego stanowi pierwsze położenie środka obszaru zainteresowania (ROI – *Region of Interest*) ze źrenicą. Zdecydowano się na realizację wszystkich operacji na ograniczonym fragmencie obrazu aby przyspieszyć ich wykonywanie się oraz zwiększyć niezawodność algorytmu. Po wstępnych badaniach rozmiar ROI ustalono na 250×250 pikseli. W kolejnych iteracjach algorytmu za współrzędne środka ROI uznawano uzyskany w poprzedniej iteracji środek ciężkości źrenicy.

Warunkiem koniecznym poprawności działania metody wyznaczania środka źrenicy w oparciu o współrzędne środka ciężkości jest dokładna segmentacja obszaru źrenicy. Wstępna analiza obrazów zamieszczonych na rysunku 5 może sugerować, że wystarczający okaże się prosty algorytm binaryzacji ze stałym progiem. Przeprowadzone testy na sekwencjach zarejestrowanych w różnych warunkach oświetleniowych dla 5 osób pokazały, że globalna jasność obrazu zmienia się w zależności od położenia źrenicy i zwykle progowanie nie sprawdza się.

Na podstawie eksperymentów przeprowadzonych na sekwencjach testowych zaproponowano metodę automatycznego wyznaczania progu binaryzacji, która uwzględni poziom jasności źrenicy z poprzedniej ramki obrazu oraz aktualne parametry obrazu w ROI. Podstawą działania metody jest analiza dwóch elementów: maksymalnych wartości jasności w wierszach i globalnej maksymalnej jasności w ROI. W przypadku przebiegu maksymalnej jasności w wierszach mogą wystąpić dwie sytuacje: odbłask od środkowego oświetlacza ma wyraźnie większą jasność niż źrenica oraz jasność odbłasku i źrenicy jest podobna. Obie sytuacje zaprezentowano odpowiednio na rysunkach 6a i 6b.

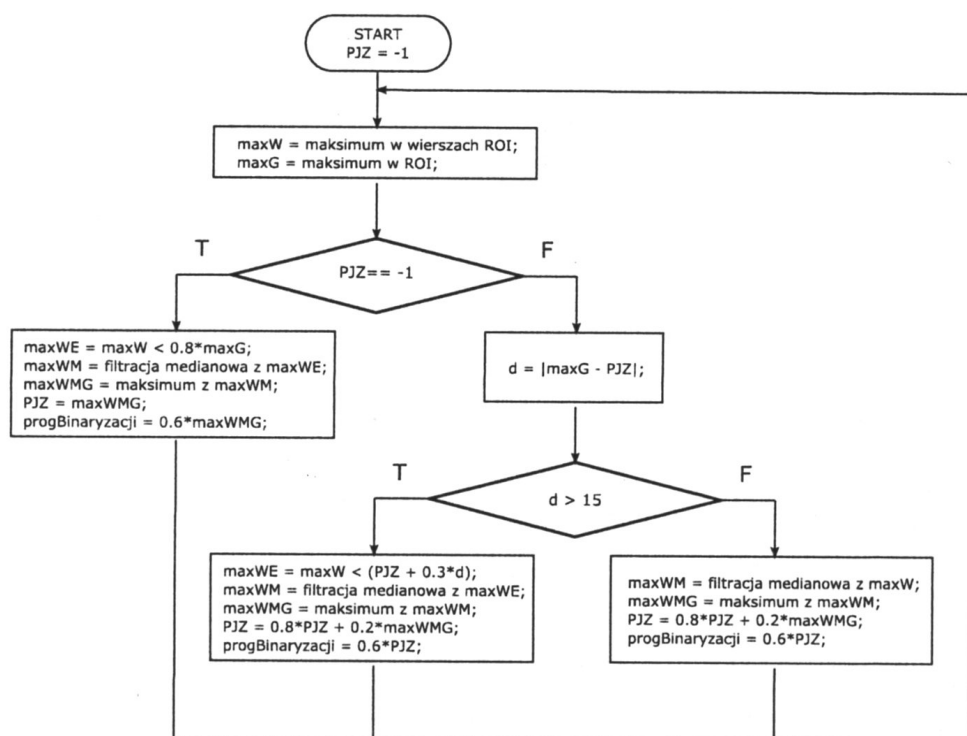
Przy inicjalizacji algorytmu, kiedy poziom jasności źrenicy (**PJZ**) nie jest znany, próg wyznaczany jest następująco: w pierwszym kroku eliminowane są piksele związane z odbłaskiem (za takie uznawane są wszystkie o wartości większe niż 0,8 maksimum przebiegu (**maxG**)), następnie przebieg poddawany jest filtracji medianowej i ostatecznie próg binaryzacji wyznaczany jest jako 0,6 maksymalnej wartości zmodyfikowanego przebiegu. Dodatkowo za pierwszy poziom jasności źrenicy przyjmuje się maksymalną wartość z przebiegu.

Dla kolejnych ramek możliwe są dwie sytuacje: przebieg ma postać podobną do zaprezentowanej na rysunku 6a lub 6b. Wykrycie konkretnego przypadku następuje na podstawie analizy odległości pomiędzy poziomem jasności źrenicy z poprzedniej iteracji a maksymalną wartością przebiegu (**d**). Jeżeli odległość jest większa od zadanego progu (eksperymentalnie dobrana wartość 15) to zakłada się, że na przebiegu występuje odbłask i należy go wyeliminować (za należące do odbłasku uznaje się piksele o jasności większej niż **PJZ** + $0,3 \times \mathbf{d}$). Następnie wykonywana jest filtracja medianowa i wyznaczane jest maksimum z przebiegu (**M**). Na podstawie wartości **M** wyliczany jest nowy poziom jasności źrenicy: **PJZ** = $0,8 \times \mathbf{PJZ} + 0,2 \times \mathbf{M}$, a ostatecznie próg binaryzacji ustalany jako $0,6 \times \mathbf{PJZ}$. Wszystkie współczynniki zostały dobrane empirycznie.



Rys. 6. Przebieg maksymalnej jasności w wierszach; z wyraźnym odbłaskiem (a), bez widocznego odbłasku (b)

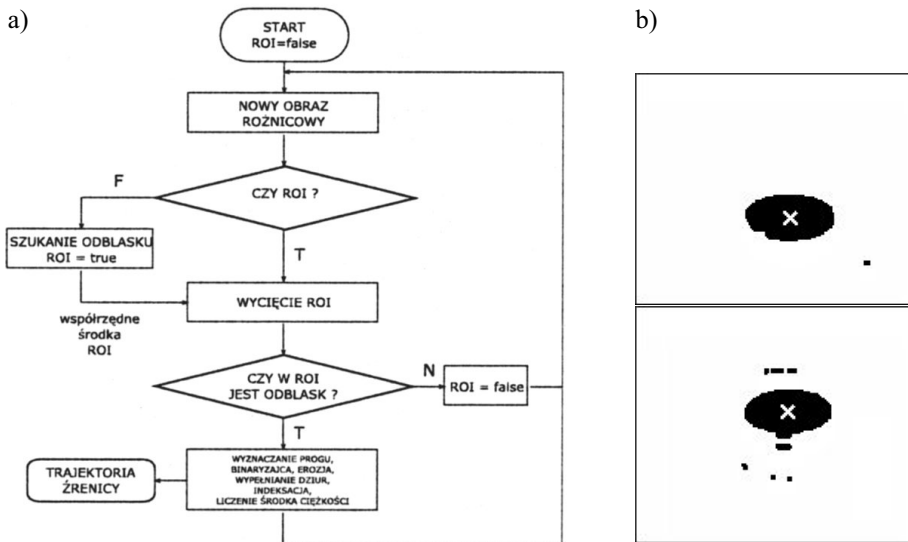
W przypadku odległości d mniejszej od progu, zakłada się, że odbłask nie jest wyróżniony i z opisanego wyżej postępowania wyklucza się krok eliminacji odbłasku. Algorytm w postaci schematu blokowego został przedstawiony na rysunku 7. Testy, przeprowadzone na zarejestrowanych sekwencjach wideo, wykazały, że zaproponowana metoda wyznaczania progu binaryzacji działa poprawnie i jest odporna na zmiany jasności obrazu.



Rys. 7. Wyznaczanie progu binaryzacji

Obraz ROI po binaryzacji podlegał dalszemu przetwarzaniu: morfologicznej dylatacji z maską 3×3 w celu usunięcia pojedynczych pikseli oraz morfologicznemu wypełnianiu dziur w celu eliminacji miejsc w których widoczne były odbłaski od oświetlaczy bocznych (na obrazie różnicowym miały kolor czarny). Następnie dokonywano indeksacji i dla wyznaczonych obiektów obliczano środek ciężkości i powierzchnię. Za źrenicę uznawany był obiekt o największej powierzchni, a za jej położenie środek ciężkości tego obiektu.

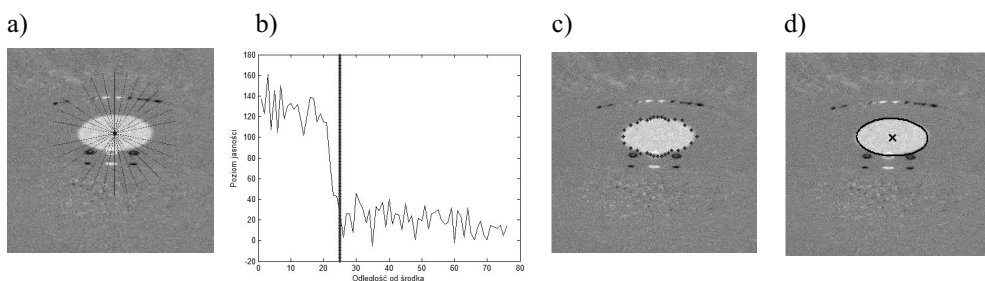
Na podstawie analizy zarejestrowanych sekwencji wideo stwierdzono, że w przypadku gdy oko nie jest widoczne, głównie z powodu mrugnięcia, znikają również odbłaski od oświetlaczy. Zdecydowano się zatem uznawać za poprawne takie obrazy ROI, w których maksymalny poziom jasności przekracza pewien ustalony próg. Ta bardzo prosta metoda okazała się względnie skuteczna, jednakże aby uczynić system bardziej niezawodnym, należałoby w przyszłości rozbudować algorytm klasyfikujący czy na obrazie znajduje się otwarte oko. W przypadku gdy stwierdzono brak odbłasku w ROI, algorytm przechodził do poszukiwania odbłasku na całym obrazie (podobnie jak w przypadku pierwszej ramki), a punkty na trajektorii środka ciężkości źrenicy i odbłasku oznaczane były jako niepoprawne. Algorytm, w postaci schematu blokowego, przedstawiony został na rysunku 8a, a przykładowe obrazy po segmentacji źrenicy na rysunku 8b.



Rys. 8. Przetwarzanie kolejnych obrazów różnicowych (a), przykładowe wyniki (b)

W literaturze spotyka się wyznaczanie środka źrenicy na podstawie środka elipsy, która opisuje obszar źrenicy [9] lub tęczówki [22]. Zdecydowano się zatem zaimplementować wspomnianą metodę i wstępnie porównać jej działanie z metodą opartą na wyliczaniu środka ciężkości. W pierwszym etapie konieczne było znalezienie punktów, które leżą na granicy źrenicy i tęczówki. Wykorzystano algorytm rozchodzących się promieni (*starburst*).

Następnie na podstawie przebiegu jasności wzdłuż promieni automatycznie wyznaczano położenie największego spadku jasności. Otrzymany w ten sposób zbiór punktów leżących na krawędzi źrenicy stanowił wejście dla metody dopasowywania źrenicy. Zdecydowano się wykorzystać metodę zaproponowaną przez [19]. Środek uzyskanej elipsy uznawany był za położenie środka źrenicy. Kolejne etapy działania algorytmu zaprezentowano na rysunku 9.



Rys. 9. Algorytm dopasowania elipsy: analizowane promienie (a), przykładowy przebieg jasności wzdłuż promienia ze znaną krawędzią źrenicy (b), wykryte punkty na krawędzi źrenicy (c), dopasowana elipsa i znaleziony środek źrenicy (d)

Obie porównywane metody generują podobne trajektorie środka źrenicy. Zauważono jednak, że w przypadku metody opartej o dopasowanie elipsy częściej następuje nieznaczna zmiana położenia środka – spowodowana nie ruchem źrenicy, a zmianą oświetlenia. Dodatkowo metoda ta jest bardziej złożona obliczeniowo. W dalszym etapie prac planowane jest ulepszenie metody opartej o wyznaczanie elipsy, tak aby generowane za jej pomocą trajektorie były dokładniejsze niż te uzyskiwane w oparciu o środki ciężkości.

6. Rejestracja ruchów głowy

Rejestracja ruchów głowy jest drugim, obok niezawodnej segmentacji źrenicy, ważnym elementem systemu śledzącego ruch gałki ocznej, wykorzystywanym podczas kalibracji. Z przyczyn technicznych dość trudno jest wykonać system, który jednocześnie zapewniłby wysoką dokładność i dopuszczał znaczną swobodę ruchu głowy osoby badanej. W płaszczyźnie równoległej do ekranu problemem jest rozdzielczość kamery – jeżeli dopuści się znaczny zakres ruchu, to rozmiar źrenicy (mierzony w pikselach) będzie bardzo mały, ponieważ pole widzenia kamery musi być znaczne (zakłada się nieruchome zamocowanie kamery). Spowoduje to spadek dokładności wyznaczania środka źrenicy. Dopuszczenie ruchu w osi prostopadłej do ekranu wymaga mechanizmu pozwalającego na wyznaczenie odległości oka osoby badanej od obiektywu (np. wykorzystując stereowizję [11]) oraz obiektywu o odpowiednich parametrach: zmienna, regulowana ogniskowa lub duża głębia ostrości.

Z uwagi na opisane powyżej problemy w prezentowanej wersji systemu zdecydowano się wykorzystać statyw, który unieruchamia głowę osoby badanej. Wpływ na przyjęcie takiego rozwiązania miały przede wszystkim dość niska rozdzielczość wykorzystywanej kamery (720×576 pikseli) oraz parametry obiektywu, które powodowały, że potencjalna swoboda ruchu osoby badanej byłaby bardzo nieznaczna i dość często występowałoby znikanie oka z kadru kamery. Podczas prowadzonych eksperymentów stwierdzono, że pomimo wykorzystania statywu, osoba badana wykonywała nieznaczne ruchy głową. Przy niewielkiej rozdzielczości systemu (obszar w jakim poruszał się środek ciężkości źrenicy miał rozmiar ok. 90×30 pikseli), ruchy miały niepomijalny wpływ na otrzymywaną trajektorię. Zdecydowano się zatem zaimplementować mechanizm rejestracji tych nieznacznych ruchów, wykorzystując proste śledzenie wzorca, aby w przyszłości wykonać kompensację ruchów głowy.

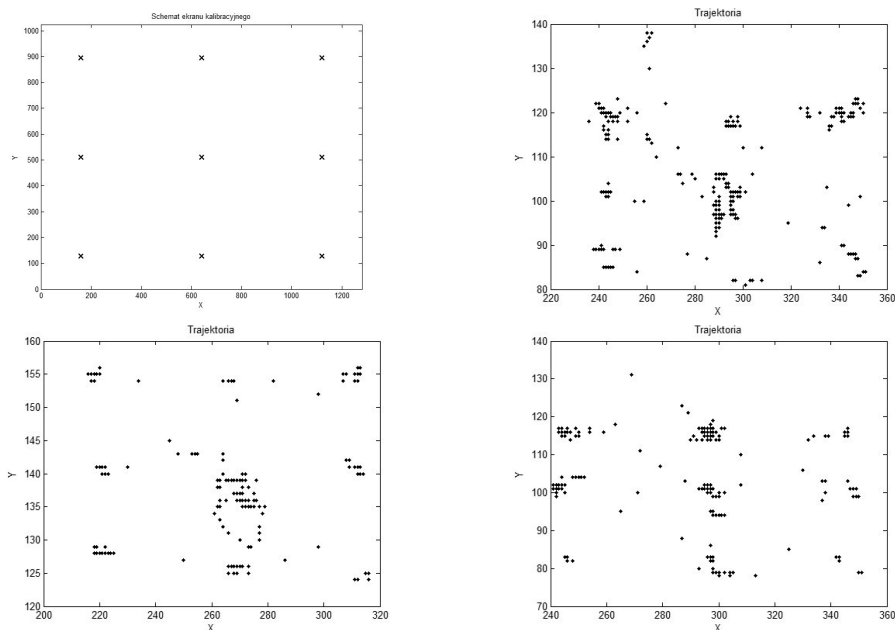
Jako wzorec wybrano symbol w kształcie krzyżyka. Marker ten przyklejany był osobie badanej na zewnętrznej części nosa. Ważnym aspektem jest taki wybór umiejscowienia markera, aby nie poruszał się on podczas mrugania, a jednocześnie znajdował się bardzo blisko oka. Pierwsze położenie markera (na ramce referencyjnej) wyznaczane było na podstawie analizy jasności i kształtu. Na kolejnych ramkach marker był śledzony z wykorzystaniem algorytmu SAD (*Sum of Absolute Difference* – suma różnic wartości bezwzględnych). W otoczeniu poprzedniego położenia wzorca wybierane były okna i obliczane były wartości SAD, współrzędne dla których wartość SAD była najmniejsza stanowiły nowe położenie markera. Testy wykazały, że rozwiązanie zapewnia wystarczającą dokładność śledzenia markera i pozwala na kompensację niewielkich ruchów głowy osoby badanej.

7. Testy, wnioski i dalsze kierunki prac

Dla opisanego systemu przeprowadzono testy będące wstępem do zaimplementowania skutecznej metody kalibracji. Na ekranie monitora wyświetlano kolejno 9 punktów, zadaniem osoby badanej było kierowanie wzroku na aktualnie widoczny punkt. Na rysunku 10 przedstawiono schemat ekranu kalibracyjnego i przykładowe trajektorie uzyskane dla różnych osób. Opisany *eye-tracker* spełnia dwa podstawowe założenia: wykorzystuje efekt jasnej i ciemnej źrenicy oraz zapewnia częstotliwość próbkowania 50 Hz wykorzystując półobrazy. Podczas prac napotkano i rozwiązano szereg problemów zarówno na etapie konstrukcji oświetlaczy i ich sterownika, jak i przy opracowywaniu algorytmu segmentacji źrenicy.

Po pierwsze okazało się, że aby uzyskać efekt jasnej źrenicy z niewielkiej odległości (ok. 50 cm od kamery), oświetlacz musi być umieszczony możliwie blisko osi optycznej obiektywu. Kolejny problem stanowił dobór odpowiedniego filtra odcinającego światło widzialne. Sprawdzono różne rozwiązania i najlepsze efekty uzyskano dla prześwietlonej kolorowej kliszy fotograficznej. Podczas projektowania algorytmu segmentacji z pierwszych analiz wynikało, że prosta binaryzacja ze stałym progiem powinna dać dobre rezultaty – z uwagi na bardzo wyróżniony obszar źrenicy. Dalsze badania wykazały jednak, że

konieczne jest zaimplementowanie bardziej wyrafinowanej metody wyznaczania progu binaryzacji. Praca nad urządzeniem pozwoliła zgromadzić doświadczenie, które w przyszłości pozwoli na znaczne udoskonalenie rozwiązania.



Rys. 10. Schemat ekranu kalibracyjnego i przykładowe trajektorie

Dalsze prace można podzielić na część sprzętową i programową. W części sprzętowej warto rozważyć zastosowanie kamery wysokiej rozdzielczości, pozwalającej na akwizycję obrazu z częstotliwością wyższą niż 25 Hz oraz obiektywu, np. dedykowanego do pracy w podczerwieni. W przyszłości celowe wydaje się też wykorzystanie dodatkowej kamery do detekcji twarzy i kompensacji ruchu głowy bez wykorzystywania markera. W części programowej warto dopracować dopasowanie na podstawie elipsy i algorytm *starburst*. Wyzwaniem jest na pewno zapewnienie niezawodnego działania w różnych warunkach oświetleniowych oraz przy obecności odbłasków. Dodatkowo warto rozważyć wprowadzenie śledzenia odbłasków od oświetlaczy (*glintów*). Pełną funkcjonalność *eye-trakera* zapewni przeniesienie ze środowiska rozwojowego (Matlab) do docelowego np. C/C++.

8. Podsumowanie

W artykule przedstawiono dokładny opis budowy prototypu *eye-trakera* wykorzystującego efekt jasnej i ciemnej źrenicy, o częstotliwości próbkowania 50 Hz. Opisano

konstrukcję oświetlacza, który pozwala na kolejnych półobrazach, uzyskiwać efekt jasnej i ciemnej źrenicy. Zaprezentowano szczegółowy opis zastosowanych algorytmów segmentacji i wyznaczania środka źrenicy, które pozwalają na generowanie trajektorii. Ponadto wykonano próbne eksperymenty z wyznaczaniem środka źrenicy za pomocą algorytmu *starburst* i dopasowywania elipsy oraz zaproponowano prosty mechanizm śledzenia ruchów głowy. *Eye-trakera* przetestowano na 5 osobach, w typowych warunkach oświetleniowych (światło sztuczne i dzienne). Uzyskane wyniki wskazują na duży potencjał zastosowanych rozwiązań i w przyszłości pozwolą na skonstruowanie w pełni funkcjonalnego systemu do wizyjnego śledzenia ruchu gałki ocznej.

Literatura

- [1] ASL – <http://www.asleyetracking.com>.
- [2] Augustyniak P., *Przetwarzanie sygnałów elektrodiagnostycznych*. UWND AGH, Kraków, 2001.
- [3] Augustyniak P., Mikrut Z., *Detection of object salient features based on the observer scanpath analysis*. 3rd European Medical & Biological Engineering Conference : IFMBE European conference on Biomedical engineering, Prague, Czech Republic, 2005.
- [4] Duchowski A., *Eye Tracking Methodology*. Springer-Verlag, London, 2007.
- [5] JAZZ-novo – <http://www.ober-consulting.com/>.
- [6] Gonzalez R., Woods R., *Digital Image Processing*. Prentice Hall, New Jersey, 2008.
- [7] Hammoud R.I., *Passive Eye Monitoring Algorithms, Applications and Experiment*. Springer-Verlag, Berlin, Heidelberg, 2008.
- [8] Kryjak T., *Wizyjne śledzenie ruchu gałek ocznych*. Automatyka (półrocznik AGH), t. 13, z. 3, 2009, 1141–1152.
- [9] Li D., Winfield D., Parkhurst D.J., *Starburst: A hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches*. Proc. of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cvpr'05), Washington, 2005.
- [10] Lorenz M.G., Mengibar-Pozo L., Liu-Jimenez J., Fernandez-Saavedra B., *User-friendly biometric camera for speeding iris recognition system*. 42nd Annual IEEE International Carnahan Conference on Security Technology. ICCST 2008, 2008.
- [11] Matsumoto Y., Zelinsky A., *An algorithm for real-time stereo vision implementation of head pose and gaze direction measurement*. Proc. Fourth IEEE International Conference on Automatic Face and Gesture Recognition, 2000, 499–504.
- [12] Microtron – <http://www.mikrotron.de/>.
- [13] Morimoto C.H., Koons D., Amir A., Flickner M., *Frame-Rate Pupil Detector and Gaze Tracker*. Proc. of the IEEE ICCV'99 frame-rate workshop, 1999.
- [14] Nguyen K., Wagner C., Koons D., Flickner M., *Differences in the infrared bright pupil response of human eyes*. Proc. of the 2002 Symposium on Eye Tracking Research & Applications, New Orleans, 2002, 133–138.
- [15] Ober J., Loska J., *Przeznaczenie i funkcje systemu pomiarowego ruchu oka OBER2*. Materiały z Krajowej Konferencji Techniki Informatyczne w medycynie, 2000.
- [16] Przybyło J., Jabłoński M., Wołoszyn P., *Wizyjny interfejs człowiek-komputer przeznaczony dla użytkowników niepełnosprawnych*. Automatyka (półrocznik AGH), t. 7, z. 3, 2003, 385–398.
- [17] Silicon Imaging – <http://www.siliconimaging.com/>.
- [18] SMI – <http://www.smivision.com/>.

- [19] Taubin G., *Estimation Of Planar Curves, Surfaces And Nonplanar Space Curves Defined By Implicit Equations, With Applications To Edge And Range Image Segmentation*. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 13, no. 11, 1991, 1115–1138.
- [20] Tadeusiewicz R., Korohoda P., *Komputerowa analiza i przetwarzanie obrazów*. Wydawnictwo Fundacji Postępu Telekomunikacji, Kraków, 1997.
- [21] TOBI – <http://www.tobii.com/>.
- [22] Zhu J., Yang J., *Subpixel Eye Gaze Tracking*. Proc. of the Fifth IEEE international Conference on Automatic Face and Gesture Recognition, Washington, 2002.
- [23] Zhu Z., Ji Q., *Eye and gaze tracking for interactive graphic display*. Machine Vision and Applications, vol. 15, Springer-Verlag, 2004, 139–148.