

This paper should be cited as: Grobelny, J., Jach, K., Kuliński, M., & Michalski, R. (2006). Śledzenie wzroku w badaniach jakości użytkowej oprogramowania . Historia i mierniki. Unpublished paper presented at Interfejs użytkownika - Kansei w praktyce Conference, Warszawa 2006. Warsaw.

Śledzenie wzroku w badaniach jakości użytkowej oprogramowania. Historia i mierniki

**JERZY GROBELNY, KATARZYNA JACH,
MARCIN KULIŃSKI, RAFAŁ MICHALSKI**

**Laboratorium Ergonomii
Instytut Organizacji i Zarządzania
Politechnika Wroclawska
ul. Smoluchowskiego 25
50-372 Wrocław
tel. +48 71 348 50 50
<http://ergonomia.ioz.pwr.wroc.pl>**

1. Wstęp

Próby śledzenia ruchu gałek ocznych w procesie widzenia podejmowano już w XIX wieku. Przegląd metod używanych do rejestracji ruchu gałek ocznych, począwszy od pierwszych metod inwazyjnych, opierających się na mocowaniu specjalnych nasadek do znieczulonej gałki ocznej, aż po współcześnie stosowane techniki analizy na podstawie zarejestrowanego obrazu wideo, uświadamia postęp, jaki dokonał się w tej dziedzinie. Coraz doskonalsze systemy rejestracji zachowania wzrokowego w połączeniu z rozwijającą się metodyką badań okulograficznych stanowiąc mogą podstawę oceny jakości użytkowej oprogramowania. Analizowane w badaniu okulograficznym psychofizjologiczne parametry pracy wzroku omówiono w oparciu o normę ISO 9241.

2. Przegląd metod okulograficznych

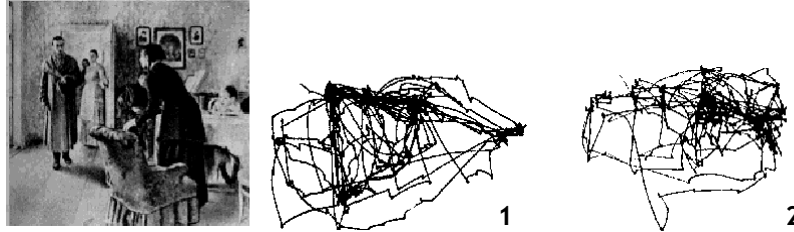
Pierwsze praktyczne próby rejestracji okulograficznych datują się na drugą połowę XIX wieku, kiedy to ruchy gałek ocznych znalazły się w polu zainteresowań ówczesnych psychologów. Użyte przez Javala (1879) lusterka umożliwiały obserwację ruchów oka podczas procesu czytania. Jako pierwszy zwrócił on uwagę na fakt, iż mają one charakter gwałtownych, szarpanych przemieszczeń, po których następuje faza bezruchu gałki ocznej. Przemieszczenia, trwające od 30 do 120 milisekund i charakteryzujące się prędkościami kątowymi rzędu 500 stopni na sekundę (Richardson i Spivey 2004) – co klasyfikuje je jako najszybsze ruchy wykonywane przez ciało człowieka – nazywane są sakkadami, natomiast faza bezruchu, trwająca średnio 200–300 milisekund, podczas której następuje percepcja informacji przez siatkówkę, nosi nazwę fiksacji. Rejestrację ilości sakkad Javal przeprowadzał z użyciem mikrofonu mocowanego do zamkniętej powieki jednego oka, podczas gdy osoba badana odczytywała tekst korzystając z drugiego oka – w ten sposób możliwe było wykrycie każdego przejścia soczewki oka (które posiada w tym miejscu odczuwalne zgrubienie) w okolicach umocowanego mikrofonu. Bardziej

miarodajna metoda opracowana została przez Delabarre'a w roku 1898: przy pomocy gipsu mocował on do znieczulonej kokainą gałki ocznej odpowiednio ukształtowaną nasadkę z otworem, przez który możliwe było czytanie. Nasadka połączona była cienkim drutem z dźwignią, która z kolei przenosiła poziome ruchy oka na obracający się bęben kimografu, umożliwiając ich rejestrację.

Oczywiste wady pierwszych metod okulograficznych, takie jak duży stopień inwazyjności, utrudnianie ruchów oraz fizyczne obciążenie gałki ocznej doprowadziły wkrótce do opracowania udoskonalonych rozwiązań. Dodge i Cline (1901) jako pierwsi wykorzystali fakt odbijania się światła od powierzchni rogówki i stworzyli nieinwazyjną metodę rejestracji na materiale światłoczułym. Oświetlane pionowo oko odbijało promienie w różnych kierunkach, w zależności od jego chwilowego położenia (zmiany kąta odbicia wynikają z kształtu gałki ocznej, która nie jest idealną sferą oraz z faktu, iż nie obraca się ona idealnie wokół swego geometrycznego środka). Światło odbite rejestrowane było poprzez poziomą szczelinę na jednostajnie przesuwanej się płycie światłoczułej, dostarczając w ten sposób informacji o wpływającym podczas badania czasie oraz poziomej amplitudzie ruchów oka. Udoskonalona przez Gillilanda (1921) metoda umożliwiła pierwsze w świecie symultaniczne rejestracje ruchów w poziomie jednego oraz ruchów w pionie drugiego oka. W latach 30 opracowano metodę rozszczepiania odbitego od powierzchni rogówki strumienia światła, co umożliwiło dwuwymiarową rejestrację ruchów pojedynczego oka (Buswell 1935).

Równoległe z powyższymi metodami rozwijano technologię badań elektrookulograficznych. Elektrookulografia bazuje na fakcie występowania różnicy potencjałów pomiędzy tylną i przednią częścią oka, wynikającej wprost z elektrycznej aktywności siatkówki. Istnienie tego zjawiska odkryto u zwierząt w połowie XIX wieku, natomiast pierwsze próby jego wykorzystania w rejestrowaniu ruchów gałki ocznej przeprowadzono w latach 20 wieku XX (Marg 1951). We wszystkich odmianach tego podejścia wykorzystywane są elektrody rozmieszczane na twarzy w bezpośrednim sąsiedztwie gałki ocznej, przy czym w najprostszym przypadku są to dwie elektrody: jedna mocowana po stronie nosowej, a druga po stronie skroniowej, umożliwiające rejestrację zmian potencjału pod wpływem poziomych ruchów oka. W celu badania ruchów w obu płaszczyznach konieczne jest zastosowanie dwóch par elektrod – dodatkowe dwie umieszczone są odpowiednio powyżej i poniżej gałki ocznej.

Lata 60 XX wieku to powrót technik inwazyjnych. Opracowano wtedy m. in. soczewkę kontaktową wyposażoną w cewki, wykorzystując do pomiaru pozycji oka zjawisko indukcji elektromagnetycznej. Wyprowadzone z soczewki przewody pozwalały na rejestrację zmian indukowanego w cewkach prądu pod wpływem ruchów oka pozostającego w generowanym wokół głowy badanego polu elektromagnetycznym. Z kolei Yarus (1965) zastosował w swoich badaniach mocowaną podciśnieniowo soczewkę kontaktową wyposażoną w niewielkie lustro, które odbijało rzucane w stronę oka światło (rys. 1).



Rys. 1. Aktywność oczu zarejestrowana przez Yarbusa (1965) podczas obserwacji obrazu „Nieoczekiwany gość” autorstwa I. J. Repina. Ścieżki uzyskane odpowiednio podczas zwykłego przeglądania (1) i próby oceny statusu majątkowego rodziny (2)

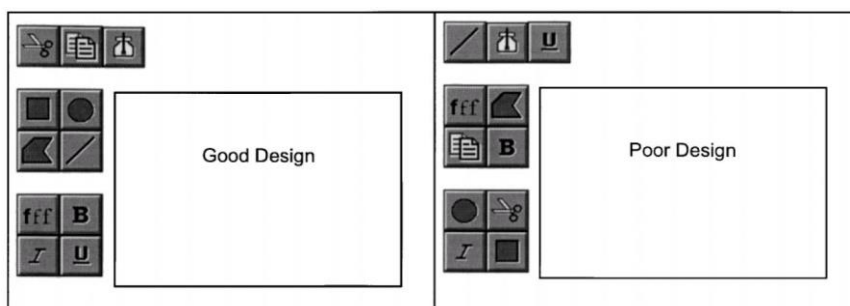
Podstawową wadą wszystkich metod opracowanych do końca lat 60. była niemożność analizowania ruchów gałki ocznej na bieżąco, podczas prowadzenia badania. Rejestrowane dane umożliwiały odtworzenie chwilowej pozycji oka czy kierunku obserwacji dopiero po przeprowadzeniu ich żmudnej obróbki. Przełomem było zastosowanie technik telewizyjnych oraz komputerów, przez co możliwa stała się obróbka sygnałów oraz prezentacja wyników w czasie rzeczywistym, w trakcie prowadzenia badania. Za protoplastę współcześnie używanych urządzeń do rejestracji ruchów gałki ocznej uważa się „Oculometer”, stworzony w latach 70. na potrzeby sił powietrznych USA (Merchant i in. 1974). Jego twórcy opracowali szereg komputerowych algorytmów, służących do rozpoznawania źrenicy na obrazie wideo, wyznaczania jej geometrycznego środka oraz obliczania odległości między nim a tzw. odbiciem rogówkowym, co pozwalało wyznaczyć kierunek patrzenia osoby badanej. Metoda ta, w udoskonalonej formie, jest wykorzystywana do dziś w systemach do eyetrackingu opartych na komputerach PC. W użyciu pozostają również, głównie jednak na polu diagnostyki oraz badań medycznych, urządzenia oparte na zjawisku indukcji elektromagnetycznej oraz techniki elektrookulograficzne.

3. Przegląd wskaźników charakteryzujących aktywność wzrokową

W normie ISO 9241 użyteczność definiuje się w trzech wymiarach: skuteczności, czyli dokładności i kompletności osiągnięcia celów przez użytkowników, efektywności lub wydajności, czyli relacji skuteczności i zużytych zasobów oraz zadowolenia lub satysfakcji użytkownika, wynikającej z używania produktu. Badanie okulograficzne może służyć do analizy stopnia spełnienia każdego z tych wymiarów. Nasuwającym się miernikiem skuteczności jest binarny wskaźnik realizacji zadania wzrokowego przez użytkownika. Mierniki efektywności koncentrują się na pomiarze realizacji zadania w wymiarze czasowym i/lub przestrzennym. Do pomiaru satysfakcji wykorzystać można różnice wielkości źrenic podczas badania.

Nowoczesne systemy okulograficzne współpracujące z dedykowanym oprogramowaniem pozwalają rejestrować i analizować, a także prezentować w odpowiedniej formie rezultaty śledzenia zachowania wzroku. Idea analizy okulograficznej bazuje na założeniu sformułowanym przez Justa i Carpentera (1976) stwierdzającym, że podczas wykonywania zadań wzrokowych lokalizacja fiksacji wskazuje na obszar zainteresowania (ang. area of interest, w skrócie AOI) obserwatora. System monitorowania ruchu gałek ocznych rejestruje z odpowiednią częstotliwością punkty zatrzymania wzroku (gaze-

points). Krokiem pierwszym analizy jest rozróżnienie sekwencji fiksacji (dłuższego spoglądania w stałym kierunku) i sakkad – szybkiego zmieniania kierunku patrzenia. Następnie w analizie wykorzystuje się różne wskaźniki charakteryzujące zarejestrowane położenia wzroku w czasie realizowanych zadań. Podstawowe wskaźniki zaproponowali w tej dziedzinie Goldberg i Kotval (1999), badając wrażliwość przebiegu procesu poszukiwania informacji na bardziej lub mniej logiczne ukształtowanie interfejsu graficznego. Przedmiotem badań był interfejs pokazany na rys. 2.



Rys. 2. Interfejsy badane przez Goldberga i Kotwała (1999)

Tak zwany „dobry projekt” (po lewej stronie rys. 2) wykonano grupując logicznie poszczególne ikony reprezentujące narzędzia edycji. Projekt „zły” miał te same narzędzia rozmieszczone losowo. Proponowane charakterystyki – mierniki procesu ruchu gałek ocznych – obejmowały dwa wymiary tego procesu: czasowy i przestrzenny. Analizom poddano zarówno trzy rodzaje otrzymanych danych: punkty zatrzymania wzroku (gaze-points) rejestrowane z częstotliwością 60 Hz, fiksacje (zbiory takich punktów leżących w promieniu 40 pikseli) oraz sakkady. Uzupełniona klasyfikacja mierników zaproponowana przez cytowanych autorów została pokazana na rys. 3. Autorzy stwierdzili istotność różnic wszystkich wymienionych poniżej wskaźników dla projektów interfejsów pokazanych na rys. 2.

	Czasowe		Przestrzenne
Punkty zatrzymania wzroku (częstotliwość próbkowania 60 Hz)	Czas przebywania w AOI		Długość ścieżki
	Czas realizacji ścieżki (czas obsługi)		Powłoka wypukła ścieżki skanowania
	Czas pierwszego punktu zatrzymania wzroku w AOI		Gęstość przestrzenna
	Gęstość przejść		
	Udział fiksacji w AOI w ogólnej liczbie fiksacji		
Fiksacje	Liczba fiksacji		
		Czas trwania fiksacji	Liczba fiksacji w AOI
Sakkady	Iloraz fiksacji do sakkad		Liczba sakkad Czas trwania sakkad

Rys. 3. Klasyfikacja rodzajów mierników ścieżek poszukiwania wzrokowego według Goldberga i Kotwała (1999) – uzupełniona.

Ilość fiksacji

Ilość fiksacji jest miernikiem negatywnie skorelowanym z efektywnością poszukiwania wzrokowego. Dobra organizacja przekazu informacyjnego powinna pozwolić zmniejszyć ilość fiksacji, w trakcie których odczytuje się fragmenty interfejsu o zakresie obszaru odpowiadającym kątom dobrego widzenia (1–5°). Według Jacoba i Karna (2003) ilość fiksacji jest wskaźnikiem wykorzystywanym najczęściej w badaniach okulograficznych. Graf i Kruger (1989) rozróżniają dodatkowo fiksacje przypadkowe (krótsze niż 240 ms) i celowe (dłuższe niż 320 ms).

Czas trwania fiksacji

Analiza całkowitego lub średniego czasu trwania fiksacji opiera się na założeniu, że im dłuższy czas fiksacji, tym większa trudność zadania wzrokowego (Fitts 1954, Yarbus 1965, Goldberg i Kotval 1998). Z badań Renshawa i in. (2004) wynika, że wydłużenie czasu fiksacji świadczy o problemach z przyswojeniem informacji.

Ilość sakkad

Ilość sakkad jest ściśle związana z organizacją przestrzenną informacji. Częste przeskoiki będą wymuszane niejasną organizacją albo niewłaściwą organizacją otoczenia funkcjonalnie powiązanych elementów.

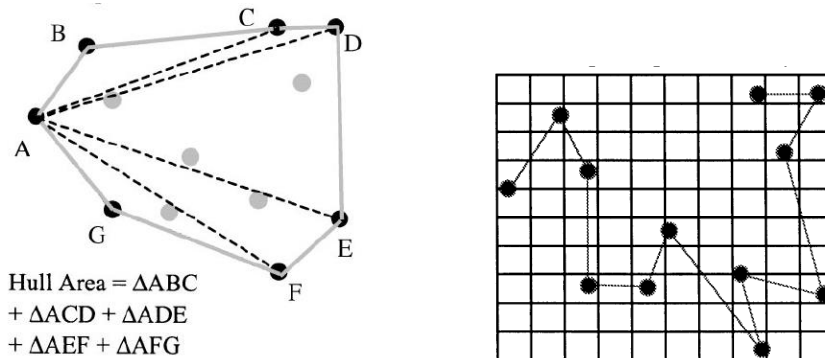
Długość i czas realizacji ścieżki

Badania wrażliwości charakterystyk ścieżki poszukiwań wzrokowych na graficzny format wykresów zostały pokazane w pracy Renshawa i in. (2004). Rejestrowane dane dla 20 osób pozwoliły wyznaczyć szereg wskaźników dla ścieżek skanowania wzrokowego, z których najbardziej istotna okazała się długość ścieżki skanowania wzrokowego. Jest dość oczywistym stwierdzenie, że im krótsza jest rzeczywista ścieżka skanowania wzrokowego, tym lepiej realizowano zadanie wyszukiwania. Ścieżka idealna to sekwencja prostych linii pomiędzy elementami koniecznymi do odnalezienia. Interfejs może być więc oceniany wskaźnikiem „produktywności wyszukiwania wzrokowego”, określonym jako relacja rzeczywistej długości ścieżki do długości po liniach prostych łączących tylko konieczne obiekty. Wg badań Goldberga i Kotwała wskaźnik długości jest bardziej wrażliwy na jakość projektu, niż tradycyjnie używany miernik czasu realizacji ścieżki (czasu obsługi).

Inne autorskie wskaźniki, zaproponowane przez Goldberga i Kotwała do pomiaru wymiaru przestrzennego ruchu gałek ocznych, to powłoka wypukła ścieżki skanowania, gęstość przestrzenna oraz gęstość przejść.

Powłoka wypukła ścieżki skanowania

Powłoka wypukła ścieżki skanowania (convex hull area) definiowana jest jako wielkość powierzchni, jaką zajmuje powłoka wypukła zawierająca wszystkie zarejestrowane punkty ścieżki poszukiwania (rys. 4a).



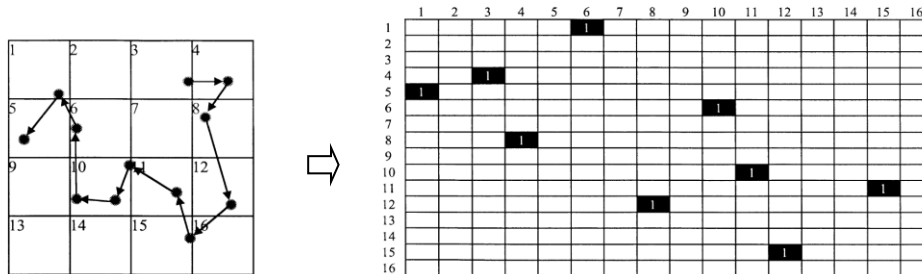
Rys. 4. Powłoka wypukła ścieżki skanowania (a) oraz gęstość przestrzenna ścieżki skanowania (b) (Goldberg, Kotval 1999)

Wskaźnik gęstości przestrzennej

Wskaźnik gęstości przestrzennej (spatial density) określa na ile równomiernie rozkładają się kierunki spojrzeń badanych. Wskaźnik oblicza się jako stosunek ilości komórek regularnej siatki (przypisanej do ekranu) zajętych przez ścieżkę do liczby wszystkich komórek siatki (rys. 4b). Gęstość przestrzenna ścieżki poszukiwań pokazanej na rys. 4 wynosi 12%. Im bardziej skoncentrowane są ścieżki poszukiwań, tym łatwiejszy – bo przestrzennie ukierunkowany – jest interfejs.

Gęstość przejść

Macierz przejść albo też macierz powiązań jest znanym pojęciem ergonomicznym. Pokazuje powiązania pomiędzy każdą parą obiektów. Powiązania mogą mieć różnorodny charakter. W przypadku poszukiwania wzrokowego macierz przejść jest definiowana dla tak zwanych obszarów zainteresowania, które w systemach śledzenia pracy wzroku zwykle definiuje się za pomocą prostokątów. Wskaźnik gęstości przejść wyraża stosunek pełnych komórek macierzy do wszystkich jej komórek. Na rys 5. pokazano sposób zdefiniowania i opisu macierzy przejść dla regularnego wzorca obszarów zainteresowań. Wartość wskaźnika w tym przypadku wynosi 14% (9 wypełnionych komórek na wszystkie 64 komórki). Im wyższy wskaźnik, czyli im więcej nawrotów – przejść w jedną i drugą stronę, tym większe trudności ze zidentyfikowaniem elementów czy też określeniem kolejności procesów.



Rys. 5. Macierz przejść dla 16 regularnych obszarów zainteresowania. (Goldberg i Kotval 1999)

Oprócz wskaźników opisanych i zbadanych przez Goldberga i Kotwała w analizie danych z badań eyetrackingowych wykorzystuje się mierniki opisujące w wymiarze przestrzennym i czasowym aktywność wzrokową osób badanych w wybranych obszarach szczególnie interesujących badacza (np. obiekty graficzne lub tekstowe, paski narzędzi, formy reklamowe). Obiekty te – tzw. obszary zainteresowania – opisywane są różnie w różnych systemach okulograficznych (np. AOI, Lookzones). Zgodnie z przytoczoną powyżej hipotezą Justa i Carpentera (1976) lokalizacja fiksacji wskazuje na obszar zainteresowania obserwatora. Stąd też badane wskaźniki to np. procentowy udział fiksacji w obszarze (obszarach) zainteresowań w ogólnej liczbie fiksacji czy też liczba punktów fiksacyjnych w obszarze (obszarach) zainteresowań. Innym, często wyróżnianym, miernikiem jest sumaryczny czas spędzony w obszarze (obszarach) zainteresowań. Zakłada się tutaj, że czas przeznaczony przez użytkownika na analizę wzrokową obszaru zainteresowania jest zależny od przypisanej temu obszarowi ważności (Fitts i in. 1950). Ciekawym, choć rzadko używanym miernikiem jest wskaźnik czasu pierwszego punktu zatrzymania wzroku w wybranym obszarze zainteresowania (Ellis i in. 1998).

Wymienione wskaźniki można analizować oddzielnie dla każdego obszaru zainteresowania lub łącznie. Analiza rozkładów czasowych i przestrzennych dla poszczególnych obszarów zainteresowania może wspomagać ocenę użyteczności poszczególnych elementów interfejsu komputerowego.

Średnica źrenic

Innym aspektem badań zachowania wzroku w trakcie interakcji jest rejestracja rozmiaru średnicy źrenic. Eyetracker, monitorując położenie źrenic, rozpoznaje także ich szerokość. Parametr ten jest niezależny od świadomości człowieka, a jego zmiany są wywoływane nie tylko różnicami natężenia oświetlenia czy też ogólnym stanem fizycznym organizmu (które to parametry pozostają względnie stałe podczas badania), ale także procesami poznawczymi i emocjonalnymi. Według badań Partala i Surakka (2003) średnica źrenic może być parametrem mierzącym subiektywne nastawienie do oprogramowania, jednak parametr ten nie informuje o kierunku tego nastawienia – efekty bodźców pozytywnych i negatywnych są podobne. Siegle i in. (2003) podkreślają jednak, że poza składową emocjonalną istotny wpływ na szerokość źrenic mają procesy poznawcze, a szczególnie obciążenie pamięci krótkotrwałej. Zatem zmiany średnicy źrenic mogą pełnić rolę wskaźnika fizjologicznego oceniającego aktywność mózgu wywołaną procesami poznawczymi i emocjonalnymi.

4. Podsumowanie

Metody okulograficzne pozwalają uzupełnić badania użyteczności o wiedzę obiektywną, uzyskaną w oparciu o analizę rzeczywistych zachowań osób badanych a nie tylko ich subiektywnych opinii. Metody te, dzięki zmniejszającej się inwazyjności, stają się szerzej dostępne, a wciąż rozwijająca się metodyka badań pozwala na ich komplementarność.

Omówione wskaźniki charakteryzujące aktywność wzrokową pozwalają kompleksowo przeanalizować wyniki badania okulograficznego i zastosować to badanie do oceny jakości użytkowej oprogramowania.

5. Literatura

1. Buswell G.T. (1935). *How people look at pictures: a study of the psychology and perception in art*. Oxford, England: Univ. Chicago Press.
2. Delabarre E.B. (1898). A method of recording eye-movements. *American Journal of Psychology*, 9(4), 572–574.
3. Dodge R., Cline T.S. (1901). The angle velocity of eye movements. *Psychological Review*, 8(2), 145–157.
4. Fitts P.M., (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47, 381–391.
5. Fitts P.M., Jones R.E., Milton J.L. (1950). Eye movements of aircraft pilots during instrument-landing approaches. *Aeronautical Engineering Review*, 9(2), 24–29.
6. Ellis S., Candrea R., Misner J., Craig C.S., Lankford C.P., Hutshinson T.E., (1998). Windows to the soul? What eye movements tell us about software usability. *Proceedings of the Usability Professionals' Association Conference*, 151–178.
7. Gilliland A.R. (1921). Photographic methods for studying reading. *Visual Education*, 2, 21–26.
8. Goldberg J.H., Kotval X.P. (1998). Eye-movement based evaluation of the computer interface. [w:] *Advances in Occupational Ergonomics and Safety*. Kumar S.K. (red.), IOS Press.
9. Goldberg J.H., Kotval X.P. (1999). Computer Interface Evaluation Using Eye Movements: Methods and Constructs. *International Journal of Industrial Ergonomics*, 24(6), 631–645.
10. Graf W., Krueger H. (1989). Ergonomic evaluation of user interfaces by means of eye-movement data. [w:] *Work with Computers: Organizational, Management, Stress and Health Aspects*. Smith M.J., Salvendy G. (red.), Elsevier Science Publishers, Amsterdam.

11. Jacob R.J.K., Karn K.S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises (Section Commentary). [w:] *The mind's eye: cognitive and applied aspects of eye movement research*. Hyona J., Radach R., Deubel H. (red.), Elsevier Science, Amsterdam.
12. Javal E. (1879). *Essai sur la Physiologie de la Lecture*. *Annales D'Oculistique*, 81, 61–73.
13. Marg E. (1951). Development of electro-oculography. Standing potential of the eye in registration of eye movement. *Archives of Ophthalmology*, 45, 169–175.
14. Merchant J., Morrissette R., Porterfield J.L. (1974). Remote measurement of eye direction allowing subject motion over one cubic foot of space. *IEEE Transactions on Biomedical Engineering*, 21(4), 309–317.
15. Partala T., Surakka V. (2003). Pupil size variation as an indication of affective processing. *International Journal of Human Computer Studies*, 59, 1–2, 185–198.
16. PN–EN ISO 9241–1:2001 Wymagania ergonomiczne dotyczące pracy biurowej z zastosowaniem terminali wyposażonych w monitory ekranowe (VDT). Ogólne wprowadzenie.
17. PN–EN ISO 9241–11:2002 (U) Wymagania ergonomiczne dotyczące pracy biurowej z zastosowaniem terminali wyposażonych w monitory ekranowe (VDT). Część 11: Wskazówki dotyczące użyteczności.
18. Renshaw J.A., Finlay J.E., Tyfa D., Ward R.D. (2004). Regressions re-visited: A new definition for the visual display paradigm. CHI 2004, April 24–29, Vienna, Austria.
19. Richardson D.C., Spivey M.J. (2004). Eye-tracking: Characteristics and methods. Research areas and applications. [w:] *Encyclopedia of biomaterials and Biomedical Engineering*. Wnek G., Bowlin G.L. (red.), Marcel Dekker, New York.
20. Siegle G.J., Steinhauer S.R., Carter C.S., Ramel W., Thase M.E. (2003). Do the seconds turn into hours? Relationships between sustained pupil dilation in response to emotional information and self-reported rumination. *Cognitive Therapy and Research*, 27, 365–383.
21. Yarbus A.L. (1965). *Eye Movements and Vision*, Plenum Press, New York