

Metoda psychodiagnostyczna oceny interfejsu graficznego użytkownika: zróżnicowanie przepływu informacji w układzie człowiek-komputer

Walery Susłow, Krzysztof Kadowski, Aleksy Patryn, Michał Statkiewicz

Politechnika Koszalińska
Wydział Elektroniki i Informatyki
Ul. Śniadeckich 2, 75-543 Koszalin, swalover@tu.koszalin.pl

Streszczenie

Prezentowane są wyniki badań doświadczalnych wybranych grup użytkowników komputerów w ramach zestawienia skomputeryzowanych testów psychodiagnostycznych. Celem badań jest analiza przepływu informacji w układzie człowiek-komputer w trakcie wykonania testów., W pierwszej kolejności autorów interesowały różnice osobowe wykazywane w parametrach wynikowych oraz w tempie przebiegu testu. Uzyskane dane interpretowane są pod kątem doskonalenia metody weryfikacji użyteczności nowych rozwiązań projektowych w dziedzinie programowych interfejsów graficznych użytkownika.

Wstęp

Autorzy wykorzystują zestaw testów psychodiagnostycznych do obiektywizacji oceny użyteczności rozwiązań projektowych (technicznych) na etapie tworzenia nowych elementów oraz układów graficznego interfejsu użytkownika GUI [1]. Stosowana metodyka oceny użyteczności [2-4] przewiduje porównanie statystycznych wyników pomiarów psychodiagnostycznych w grupie reprezentującej potencjalnych użytkowników. Uczestnicy badań wykonują wielokrotne czynności związane z odbiorem informacji, rozpoznawania obiektów lub podejmowania decyzji w obecności lub bez nowych rozwiązań projektowych (preferowany jest styl nie informowania użytkownika na temat ewentualnych różnic w porównywanych rozwiązaniach). Uwzględnienie występujących w grupie różnic osobowych na etapie wnioskowania co do użyteczności badanych rozwiązań technicznych GUI ma, zdaniem autorów, zwiększyć trafność podejmowanych decyzji projektowych.

Opis skomputeryzowanych pomiarów psychodiagnostycznych

Poniżej przedstawione są opisy dwóch wybranych rodzajów testów psychodiagnostycznych, które były wykorzystane do prowadzenia badań. Ogólnie można powiedzieć, że każdy z nich składał się z szeregu okresów, zawierających jednakową sekwencję działań wykonywanych przez użytkownika komputera, jak: percepcja sygnału pochodzącego od komputera, podjęcie decyzji i reakcja manualna, rejestrowana przez komputer. Testy istotnie różnią się sygnałem pobudzającym (wizualny lub dźwiękowy), wymaganym sposobem reakcji - manipulacji z myszą lub klawiaturą oraz prowadzeniem pomiarów w regularnych lub nieregularnych odstępach czasu.

Algorytm badania każdego użytkownika zawierał następujące czynności:

1. Instruktaż dla użytkownika
2. Pobranie podstawowych danych osobowych
3. Przeprowadzenie właściwego testu

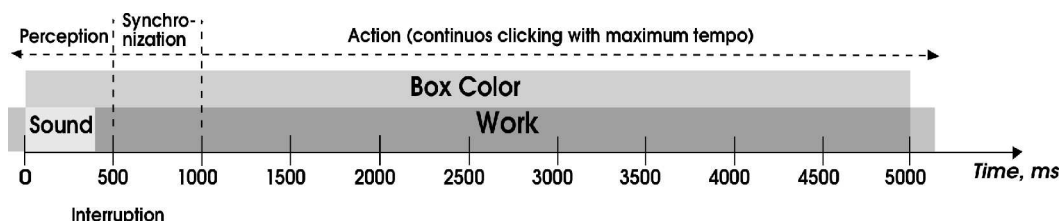
Indywidualny instruktaż przed testem oraz nadzór nad każdym użytkownikiem podczas testu były konieczne dla zapewnienia wiarygodnych wyników badań.

Test Teppinga (aplikacja TepTest), w wersji opartej o klasyczną metodę bezkomputerową, polegał na rejestracji w określonym czasie ilości kliknięć¹ przyciskiem myszy, kiedy kursor wskazuje na

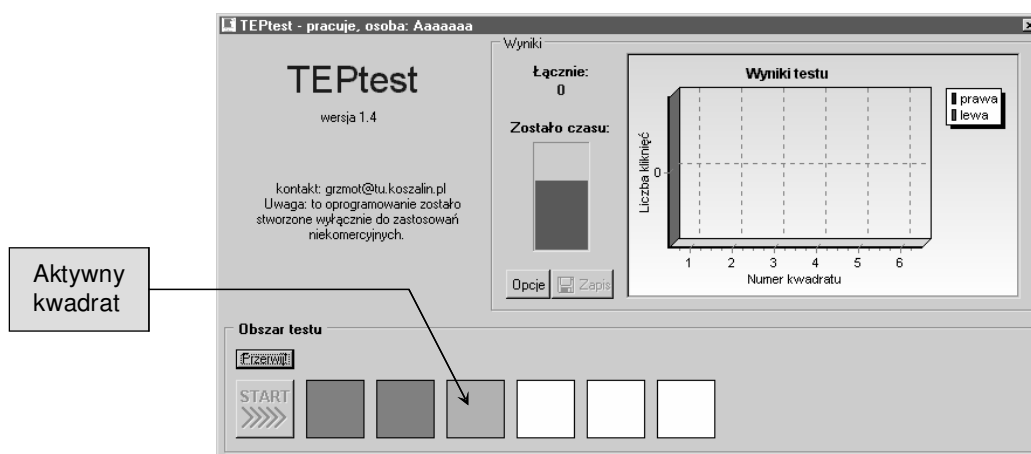
¹ Kliknięcie (ang. click), dźwiękonaśladowczy wyraz określający operację ręczną, polegającą na jednokrotnym naciśnięciu klawisza myszy (Zdzisław Płoski, "Słownik Encyklopedyczny - Informatyka").

aktywny obszar w oknie programu testującego, przy tym instrukcja użytkownika przed testowaniem nastawiała go na maksymalnie możliwe tempo klikania. Pomiar wykonywano dla obu rąk. Stosowanych było 6 obszarów aktywnych w kształcie kwadratu o wymiarach około 2 stopni kątowych na monitorze 15-calowym, przy rozdzielczości 800x600. Na każdy z kwadratów przydzielony był czas pracy 5s, z czego wynika łączny czas trwania aktywnej pracy użytkownika około 60s.

Przebieg testu zobrazowany jest schematycznie na rys. 1, a widok głównego okna aplikacji testującej pokazany jest na rys. 2. Zmiana aktywnego obszaru sygnalizowana była podświetleniem kolorem błękitnym kolejnego dotychczas nieaktywnego obszaru oraz dodatkowo dźwiękiem o częstotliwości 650Hz i długości trwania 350ms. W trakcie pracy z użytkownikiem program rejestrował również czas każdego kliknięcia mierzony od początku testu z rozdzielczością 1ms.



Rys. 1 Przebieg okresu testowania użytkownika komputera za pomocą aplikacji TepTest.

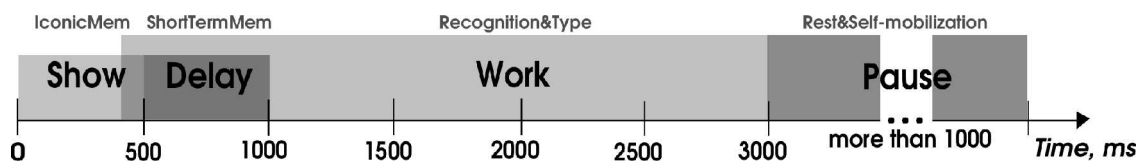


Rys. 2 Widok głównego okna aplikacji TepTest podczas wykonania pomiaru

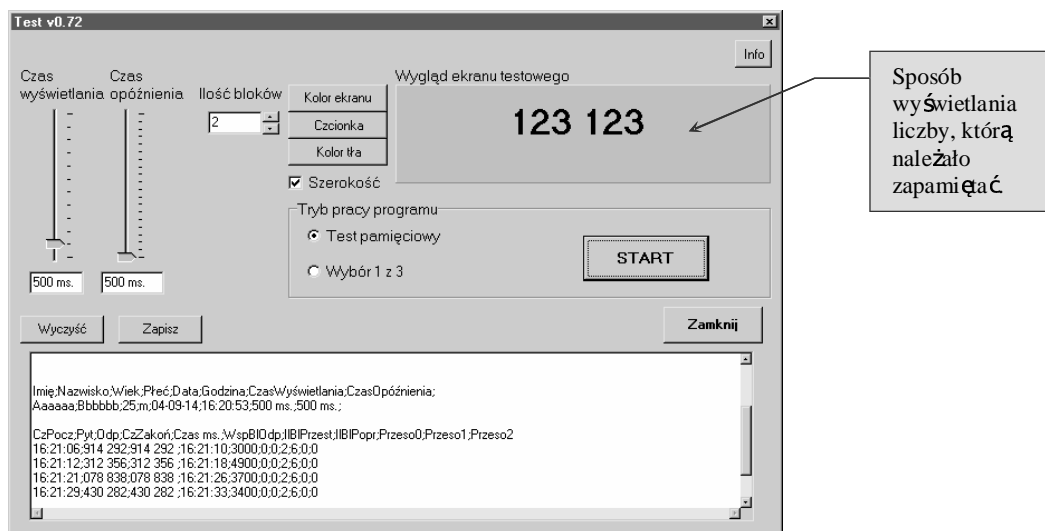
Podstawowa metoda analizy wyników testu Teppinga polegała na ocenie kształtu krzywych reprezentujących wykres zmiany tempa pracy użytkownika w kolejnych kwadratach. Kształt tak uzyskanych krzywych ma być według modelu Ilina skorelowany z „siłą” układu nerwowego badanego człowieka.

Test pamięciowy (aplikacja VerbTypeTest) przewidywał nieco odmienne zadanie dla badanego użytkownika. Celem testu było rozpoznawanie i zapamiętywanie wyświetlanych przez krótki czas liczby, a następnie po krótkiej przerwie wprowadzanie jej do komputera z klawiatury. Rozpoczęcie każdej sekwencji działań sygnalizowane było tylko pojawieniem się nowej liczby na pustym dotąd ekranie, co wymagało od użytkownika skupienia się na zmianach zachodzących na monitorze. Wykorzystywany był tryb pełnoekranowy z możliwością zmiany kolorów tła i czcionki. Wyświetlana liczba generowana była przy użyciu generatora liczb pseudolosowych, za każdym razem wyświetlane były 2 grupy po 3 cyfry (patrz rys. 4). Całkowity czas trwania testu ustalony był na 10 minut, w związku z czym w zależności od tempa pracy użytkownika wyświetlane było od 200 do 300 liczb. Graficzna ilustracja przebiegu jednego okresu testowania pokazana została na rys. 3, a widok głównego okna aplikacji testującej przedstawia rys. 4.

Podczas badań program rejestrował czas rozpoczęcia pokazu kolejnej liczby i czas zatwierdzenia przez użytkownika odpowiedzi przyciskiem „Enter”, mierzony od początku testu z dokładnością 100ms. Rejestrowana była zarówno sekwencja liczb wyświetlonych oraz wprowadzonych przez użytkownika.



Rys. 3 Przebieg okresu testowania użytkownika komputera za pomocą aplikacji VerbTypeTest.



Sposób wyświetlania liczby, którą należało zapamiętać

Rys. 4 Widok głównego okna aplikacji VerbTypeTest przed rozpoczęciem testu pamięciowego.

Zastosowane aplikacje testujące to programy komputerowe optymalizowane do pracy na platformie MS Windows. Aplikacje uruchamiane były na komputerach klasy PC pracujących pod kontrolą systemów Windows 98SE lub Windows 2000, w typowej konfiguracji sprzętowej. Stanowiska wykorzystywane w trakcie badań są typowymi stanowiskami do pracy przy komputerze, znajdującymi się w pomieszczeniach dydaktycznych uczelni. Badania odbywały się na wybranych grupach: uczniów szkoły podstawowej, studentów studiów dziennych oraz zaocznych kierunków technicznych, pracowników naukowych uczelni.

Podejście informacyjne do analizy danych

Formalna analiza przepływu informacji w układzie człowiek-komputer została przeprowadzona przy założeniu, iż układ ten w trakcie wykonywania testów pracuje w trybie okresowym ustabilizowanym. Szacowany stały przepływ informacji od komputera do człowieka za jeden okres w teście Teppingu uznano za 1 bit/okres, bo chodzi o sprzężenie typu „dotyk-klik” lub „dźwięk-klik” (nie wiadomo, które z nich jest dominujące w organizacji rytmu pracy podczas intensywnego klikania przyciskiem myszy). Test ten przewiduje pracę układu raczej zbliżoną do pracy automatu mechanicznego – pracuje głównie układ mięśniowy, przerwania stanu ustalonego występują podczas zmiany aktywnego kwadratu.

W VerbTypeTest wartość stałego przepływu informacji za okres oszacowano na poziomie 20,6 bitów/okres. W tym 19,6 bitów przekazywano jest poprzez liczbę sześciocyfrową (3,3 bity na symbol przy alfabecie 10-cio cyfrowym), a 1 bit sygnalizujący o rozpoczęciu kolejnego okresu pracy poprzez pokazanie się pasku do drukowania na ekranie. W tym teście od użytkownika wymaga się zaawansowanej pracy pamięci krótkotrwałej wraz z siecią neuronową rozpoznającą obiekty wizualne

oraz z układami odpowiedzialnymi z przekazaniem symboli rozpoznanych poprzez czynności manualne do komputera za pośrednictwem klawiatury.

Na rys. 5 przedstawione są wyniki badań przeprowadzonych dla tej samej osoby w obu testach. Badane tutaj parametry charakteryzują się pewną stałością, można więc powiedzieć, że są charakterystyczne dla wybranej osoby. Przepływ informacji Inf_i od komputera do jednostki za i -okres liczony był według następującego wzoru:

$$Inf_i = I_0 + I_m \cdot \frac{t_i - t_\infty}{t_0 - t_\infty}$$

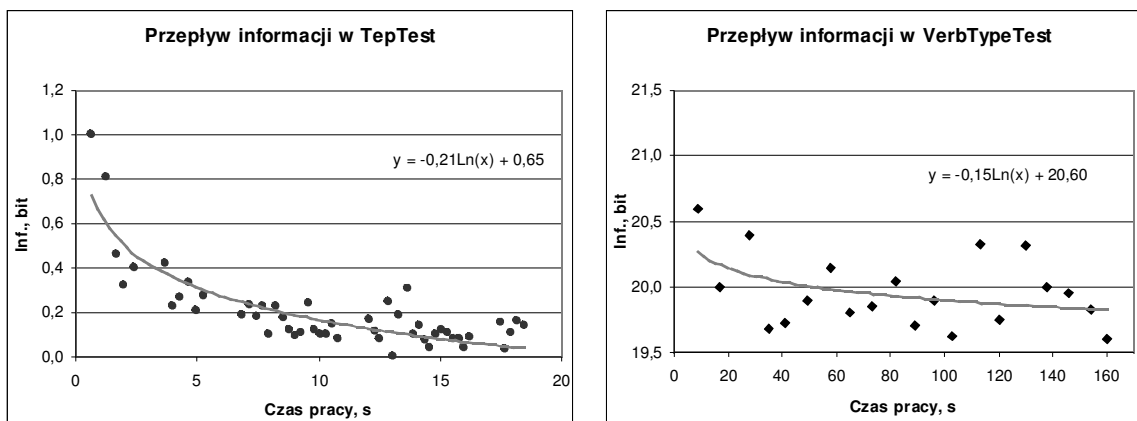
gdzie: I_0 - informacja przesyłana ze stałym tempo w trakcie testowania,

I_m - współczynnik skalujący,

t_0, t_∞ - odpowiednio czas jednego cyklu na początku i na końcu „idealnego” testu.

Ponieważ test Teppingu nie wymaga zaangażowania dużych zasobów intelektualnych i poznawczych. Na początku pracy z nieznaną aplikacją konieczność synchronizacji pracy wymaga od użytkownika dokładnie 1 bit/cykl, wyuczona czynność potrzebuje już synchronizacji raz na około 5-10 cykli.

Test pamięciowy bardziej obciąża pamięć krótkotrwałą oraz układ rozpoznawczy, który stosuje się do identyfikacji pokazywanych liczb. Zmniejszenie czasu wykonania pojedynczego cyklu wskazuje na zachodzące po stronie użytkownika zjawisko uczenia się czynności powtarzającej się. Jednak, natężenie strumienia informacyjnego odbieranego przez użytkownika w tym teście jest bardziej stabilne, bo zależy w pierwszej kolejności od konieczności rozpoznawania 6-cyfrowej liczby losowej. To oznacza, że składowa czasu reakcji na rozpoznawanie obiektów nie może się znacząco zmniejszyć. Główną przyczyną niestabilności pracy użytkownika podczas wykonania testu pamięciowego jest, naszym zdaniem subiektywna nierównomierność informacyjna strumienia liczb losowych, generowanych w aplikacji testującej. Na przykład, niektóre osoby po wykonaniu testu mówiły, że dla nich liczby zawierające powtarzające się lub symetryczne kombinacje cyfr były łatwiejsze do zapamiętania. Analiza zarejestrowanych wyników potwierdza taką tendencję - liczby te powodowały zauważalnie krótsze czasy reakcji, co można interpretować jako mniejszą zawartość informatyczną dla układu rozpoznawczego jednostki.

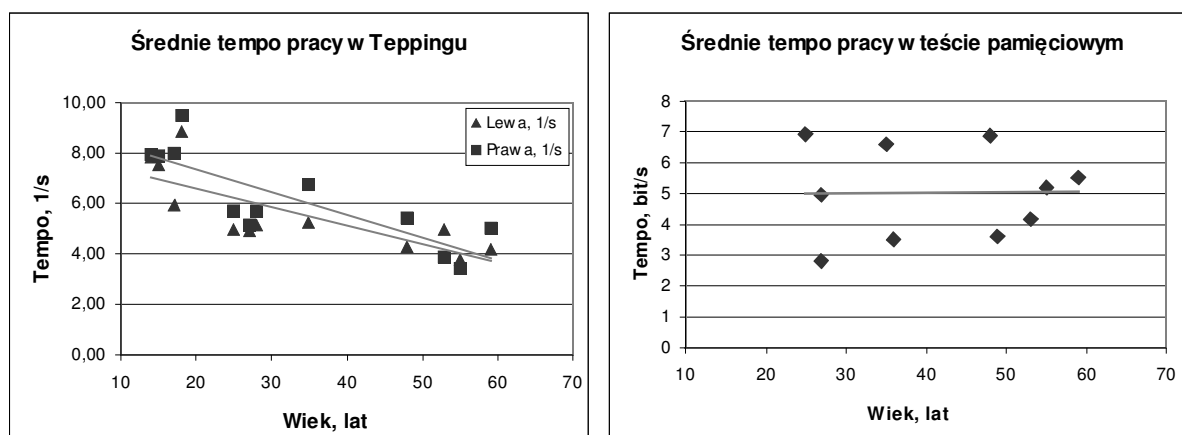


Rys.5. Analiza trendu w zmianach przepływu informacji pomiędzy komputerem a użytkownikiem podczas wykonywania różnych zadań

Zróznicowanie wyników testowania dla próbek statystycznych

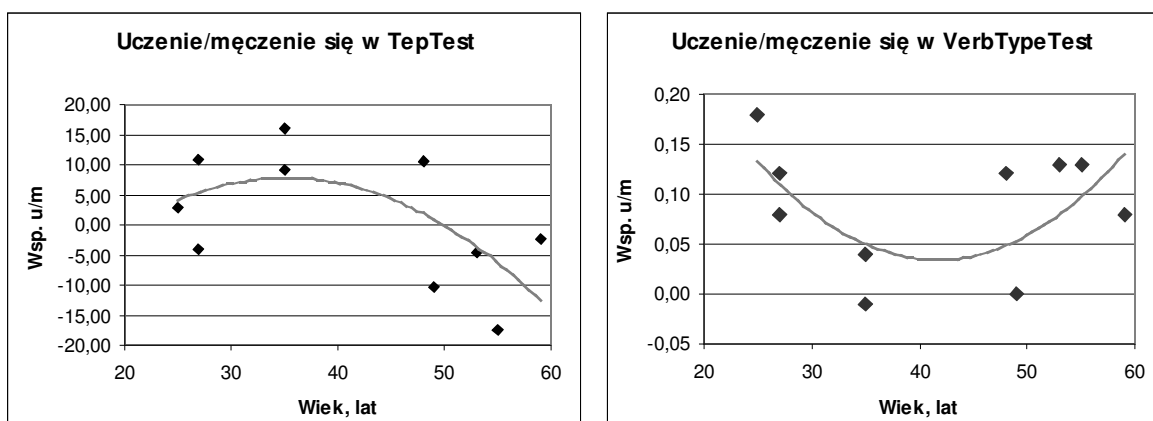
Dane uzyskane w obu testach analizowano, biorąc pod uwagę tak parametry elementarne, np. czas pojedynczego kliknięcia, jak i wyodrębniając ogólne tendencje, np. trend w zmianach parametrów okresowych podczas całego testu. Z obserwacji zachowań uczestników badań oraz z analizy teoretycznej zjawisk zachodzących w działającym układzie człowiek-komputer wynikało, że cechy indywidualne mogą mieć wpływ na przebieg wszystkich etapów testów od percepcji poprzez podejmowanie decyzji do reakcji podejmowanej przez użytkownika.

Wiek użytkowników był rozpatrywany, przypuszczalnie jako jeden z głównych czynników różnicujących reakcje badane w obydwu testach. Podstawą ku temu były znane z psychologii inżynierskiej dane o zmianach czasów elementarnych reakcji na światło i dźwięk, oraz o zmianach wrażliwości kanałów informacyjnych jednostki z wiekiem [5]. Słuszność przypuszczenia potwierdzają parametry statystyczne zebranych próbek, najbardziej charakterystyczne wyniki dają metoda klasyczna analizy testu Teppinga oraz analiza informacyjna. Czynnikiem silnie skorelowanym z wiekiem był w badanych grupach poziom wykształcenia, który prawdopodobnie wywiera znaczący wpływ na strategię uczenia w powyższych testach oraz na wyniki testu Teppinga oceniane w sposób klasyczny. Skumulowane wyniki porównania tempa pracy użytkowników w różnym wieku podane są na rys. 6. Jak można zauważyć, wskaźnik tempa klikania w teście Teppingu reprezentujący manualne zdolności jednostki maleje z wiekiem, przy tym zauważalnych zmian w tempie wykonania testu pamięciowego dla tej samej grupy nie da się wyróżnić. Jest to naturalne, bo w drugim teście nie sprawność manualna, a raczej sprawność pamięciowa lub umysłowa decydują o czasie koniecznym do wykonania zadania.



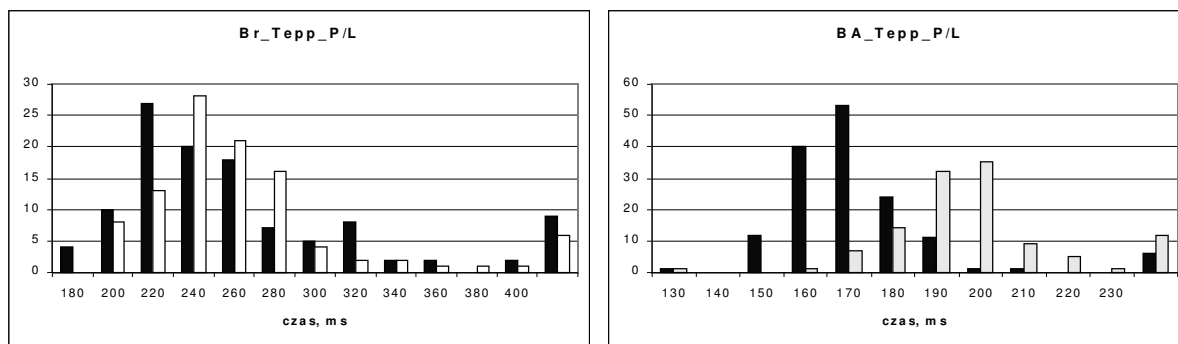
Rys. 6. Średnie tempo pracy w zależności od wieku użytkowników, dla dwóch typów testów

Podobnie reprezentują się wyniki analizy zdolności do uczenia się w trakcie pracy z nową aplikacją. Zdolność ta szacowana była za pomocą współczynnika kształtu krzywej przepływu informacji, która wykazuje najlepsze dopasowanie do danych. Przypuszczalnie parametr ten może być wykorzystany do ujawnienia zjawisk uczenia się lub zmęczenia użytkownika. Jak można zauważyć na rys. 7, generalnie młodzież ma zdecydowanie lepsze osiągnięcia w uczeniu się w stosunku do ludzi w starszym wieku podczas wykonania zadania w Teppingu, o czymś takim trudno jest wnioskować na podstawie wyników testu pamięciowego. Jeśli przeanalizować jednak dane z założenia podziału skali wiekowej na trzy strefy (młodzież, średni wiek i starsi), to można zauważyć tendencję ekstremum akurat w wieku średnim. Wniosek taki opiera się również na ścisłej analizie matematycznej, przy aproksymacji danych za pomocą paraboli wykazywany jest ekstremum wskaźnika uczenia się dla użytkowników w wieku 35-45 lat. Obserwowane zjawisko jest dość interesujące, ale wymaga potwierdzenia.



Rys. 7. Współczynnik uczenia się w trakcie wykonania testów dla użytkowników w różnym wieku.

Zróżnicowanie wyników pomiaru występuje również wewnątrz danych uzyskanych dla jednej osoby, chodzi nam konkretnie o wyniki testu Teppinga dla lewej i prawej ręki (patrz rys. 6). Do porównania czasów reakcji w tej sytuacji najbardziej pasują histogramy, przykłady graficznej reprezentacji danych tego rodzaju dla dwóch wybranych osób podane są na rys. 8. Z obserwacji dokonanych podczas wykonania pomiarów my możemy wnioskować, iż wśród użytkowników łatwo wydzielić osobników ze zdecydowaną asymetrią wyników. Jedna ręka, wiodąca pracuje u nich w zdecydowanie większym tempie. Ale istotnym dla naszych pomiarów okazało się, że ręka ta jest bardziej kontrolowana przez świadomość użytkownika, podczas gdy druga bardziej odzwierciedla obiektywne warunki pracy w środowisku wirtualnym. Można spodziewać się, więc, że ręka niewiodąca dostarczy badaczom danych, mniej „zanieczyszczonych” przez świadome wewnętrzne nastawienia uczestnika testu.



Rys. 8. Wykrywanie różnic osobowych za pomocą histogramu (ciemne słupki – prawa ręka, jasne – lewa).

Podsumowanie

Stwierdzone różnice w percepcji informacji i reagowaniu na nie wśród badanych użytkowników prowadzą do wniosku o konieczności zróżnicowania projektowanych interfejsów użytkownika w celu zwiększenia komfortu i wydajności pracy dla wszystkich przedstawicieli grupy docelowej. Ze względu na zaobserwowane w badanych grupach zróżnicowanie strategii uczenia się i przyswajania informacji, dopasowaniu do wybranej grupy użytkowników muszą podlegać, naszym zdaniem nie tylko elementy GUI, ale przede wszystkim metody pracy z daną aplikacją. Takie podejście powinno być stosowane zwłaszcza w przypadku aplikacji, w których rytm pracy dla użytkowników ma być narzucany przez aplikację jak np. w testach wiedzy i umiejętności.

Wniosek powyższy stanowi jednak wyłącznie ogólną wskazówkę dotyczącą projektowania GUI, przez co jego użyteczność jest dość ograniczona. W przypadku poszukiwania wydajnych interfejsów użytkownika wygodniejsze może być opieranie się na zestawie ścisłych wzorców opisujących wybrane rozwiązania i ich zastosowanie w konkretnych sytuacjach niż korzystanie ze zbioru

ogólnych reguł postępowania [6]. Z tego względu autorzy proponują utworzenie pewnej liczby wzorców opisujących elementy interfejsu użytkownika oraz algorytmy obsługi aplikacji. Wzorce takie mogą być zoptymalizowane np. pod kątem dopasowania do określonych grup użytkowników, wydzielonych na podstawie wieku lub wykształcenia.

Literatura

1. Spolsky J. User Interface Design for Programmers. - Apress, 2001, s 12-30.
2. Patryn A., Statkiewicz M., Susłow W. – Testy psychodiagnostyczne aparaturowe narzędziem projektanta aplikacji e-learningu. W materiałach III międzynarodowej konferencji naukowo-metodycznej "Distance learning – środowisko edukacyjne XXI wieku", Mińsk, Białoruś, 13-15 listopada 2003 r. – Mińsk, wyd. BGUIR, 2003, s. 481-484.
3. Patryn A., Statkiewicz M., Susłow W. – Perspektywa optymalizacji aplikacji komputerowych dydaktycznych poprzez modelowanie układu „uczeń-komputer”. Kognitywistyka i Media w Edukacji, tom 7, nr 1 (2003), s. 49-64.
4. Statkiewicz M., Susłow W. – Tworzenie inteligentnych interfejsów użytkownika do aplikacji edukacyjnych jednym z aspektów respektowania osobowości ucznia. W monografii pokonferencyjnej I Międzynarodowej Konferencji Naukowej "Edukacja XXI wieku" (10-12 grudnia 2003, Trzcianka), Wydawnictwo Dom Wydawniczy ELIPSA, Warszawa - Poznań 2004, tom 3, s. 96-104.
5. Ananiev B. G. Człowiek jako przedmiot poznania (w jęz. rosyjskim). – St. Petersburg, wyd. Piter, 2002.
6. M. van Welie, G.C. van der Veer, A. Eliëns (2000) Patterns as Tools for User Interface Design In: Vanderdonckt, J., Farenc, C. (eds), International Workshop on Tools for Working with Guidelines 7-8 October 2000, Biarritz, France, Springer Verlag, ISBN 1-85233-355-3, pp. 313-324.