

Interfejsy multimodalne



Prof. Andrzej Czyżewski
Politechnika Gdańska, Wydział Systemów Multimedialnych

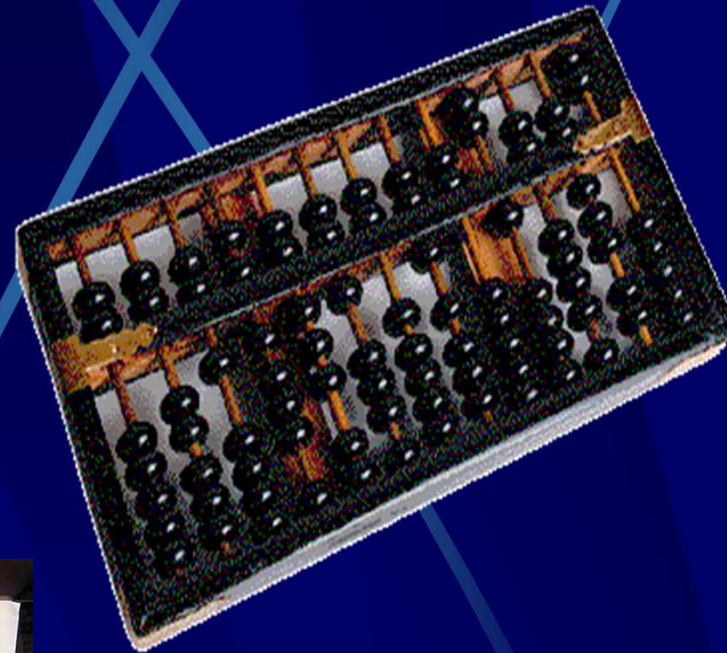
Zawartość prezentacji

- Rozwój interfejsów komputerowych
- Interfejsy unimodalne
- Interfejsy multimodalne (zilustrowane głównie za pomocą technologii opracowanych w KSMM)
- Przyszłość interakcji człowiek-maszyna

**Współdziałanie człowieka z
komputerem
od przeszłość do
teraźniejszości...**

ROZWÓJ INTERFEJSÓW

- W starożytnych Chinach pierwszym znanym interfejsem były skrzela liczydła.
- W latach 60-tych pojawiły się klawiatury maszyn do perforacji kart oraz teletypy



ROZWÓJ INTERFEJSÓW

- Kiedy w latach 70. pojawiły się pierwsze terminale, pojawiła się nagła potrzeba pisania, ponieważ terminale akceptowały tylko taką formę danych wejściowych.
- Pierwszy graficzny system operacyjny został opracowany w latach 80-tych. Interfejs ten wprowadził nas do myszy - zasadniczo prostego urządzenia wskazującego.



ROZWÓJ INTERFEJSÓW

- Kolejnym etapem były obecnie bardzo popularne interfejsy graficzne.
- Szybka ewolucja mocy obliczeniowej w latach 90-tych pozwoliła na stworzenie sprawnego systemu rozpoznawania mowy i tekstu.



Wciąż jest to naturalna ludzka skłonność do mówienia,
gestykulowania i czasami używania pisma odręcznego,
gdy potrzebna jest komunikacja.

W dzisiejszych czasach taka forma komunikacji jest
najbardziej pożądana, a interfejsy wykorzystujące te znane
są jako

INTERFEJSY MULTIMODALNE

POJĘCIA ZWIĄZANE Z INTERFEJSAMI MULTIMODALNYMI

- Interakcja człowiek-maszyna (MMI) - (podczas II wojny światowej)
- Interfejs człowiek-komputer (HCI) - (lata 70.)
- Komunikacja człowiek-maszyna (HMC)
- Percepcyjny interfejs użytkownika (PUI)
- Naturalne systemy interaktywne (NIS)

● **MODALNOŚĆ** - jest to sposób przekazywania i otrzymywania informacji

Istnieje wiele rodzajów modalności:

- język naturalny
- gesty i ruchy rąk
- mowa ciała
- wyrazy twarzy
- styl pisma odręcznego
- ...



MODALNOŚĆ

Człowiek	Komputer
Ruchy rąk lub ramion	Klawiatura, mysz lub trackball, dżojstik, skryptor, tablet graficzny, DataGlove, ekran dotykowy, tablet PC, kamera
Wyrażenia twarzy	Kamera
Głos	mikrofon
Język ciała	Urządzenie do rejestrowania pozycji lub aparat fotograficzny
Zobacz	Kamera, eye-tracker
Skóra lub ciało	Czujniki EMG (elektromiograficzne czujniki sygnału generowanego przez mięśnie)
Fale mózgowe	Czujniki EEG (sygnał elektroencefalograficzny)



Człowiek	Komputer
Oko	ekran, projektor (wyświetlacz)
ręka	Sprzężenie zwrotne siły, generator tekstury (np. alfabet Braile'a)
ucho	Głośnik
ciało	Skafander

- Interfejsy można podzielić na systemy **jednomodalne** - wykorzystujące tylko jeden sposób, np. rozpoznawanie mowy lub tekstu.
- lub systemy **multimodalne** - wykorzystujące kilka trybów jako sygnał wejściowy, np. rozpoznawanie mowy z jednoczesnym uchwyceniem gestów.

PRZYKŁADY ZASTOSOWAŃ SYSTEMÓW MULTIMODALNYCH

Edukacja - specjalne potrzeby (dotychczas rozpowszechnione w większości)

- Dla dzieci z zaburzeniami uwagi (np. **zespołem ADHD**) interfejsy multimodalne dają dużą szansę **na** poprawę ich umiejętności uczenia się - **stymulowanie różnych zmysłów pomaga skupić uwagę.**
- Szkolenie w zakresie koncentracji - stosowanie biofeedbacku
- Gry edukacyjne z interakcją multimodalną
- Inne (zostaną przedstawione w dalszej części)



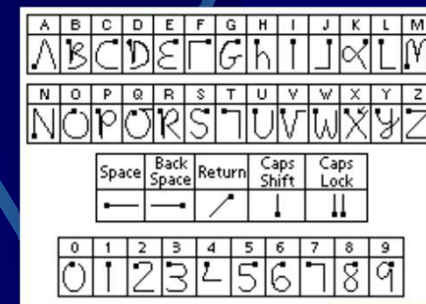
PRZEGLĄD INTERFEJSÓW
JEDNOMODALNYCH

ROZPOZNAWANIE TEKSTU

DWA PODEJŚCIA DO ROZPOZNAWANIA PISMA ODREČZNEGO

1. Zdefiniowany format zapisu (Unistroke, Graffiti)

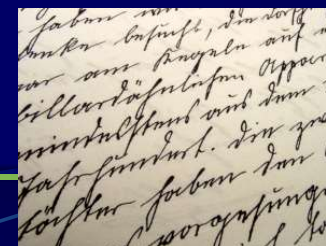
Większość symboli została napisana bez zrywania śladu na tablecie.



Były używane w komputerach **Palm 3Com, Windows CE.**

2. Komputer uczył się pisma odręcznego użytkownika:

Apple Newton



FORENSIC STUDIES FOR CRIMINAL JUSTICE

Developments in Handwriting and Signature Identification in the Digital Age

Heidi H. Harralson

Series Editor: Larry S. Miller





ROZPOZNAWANIE MOWY

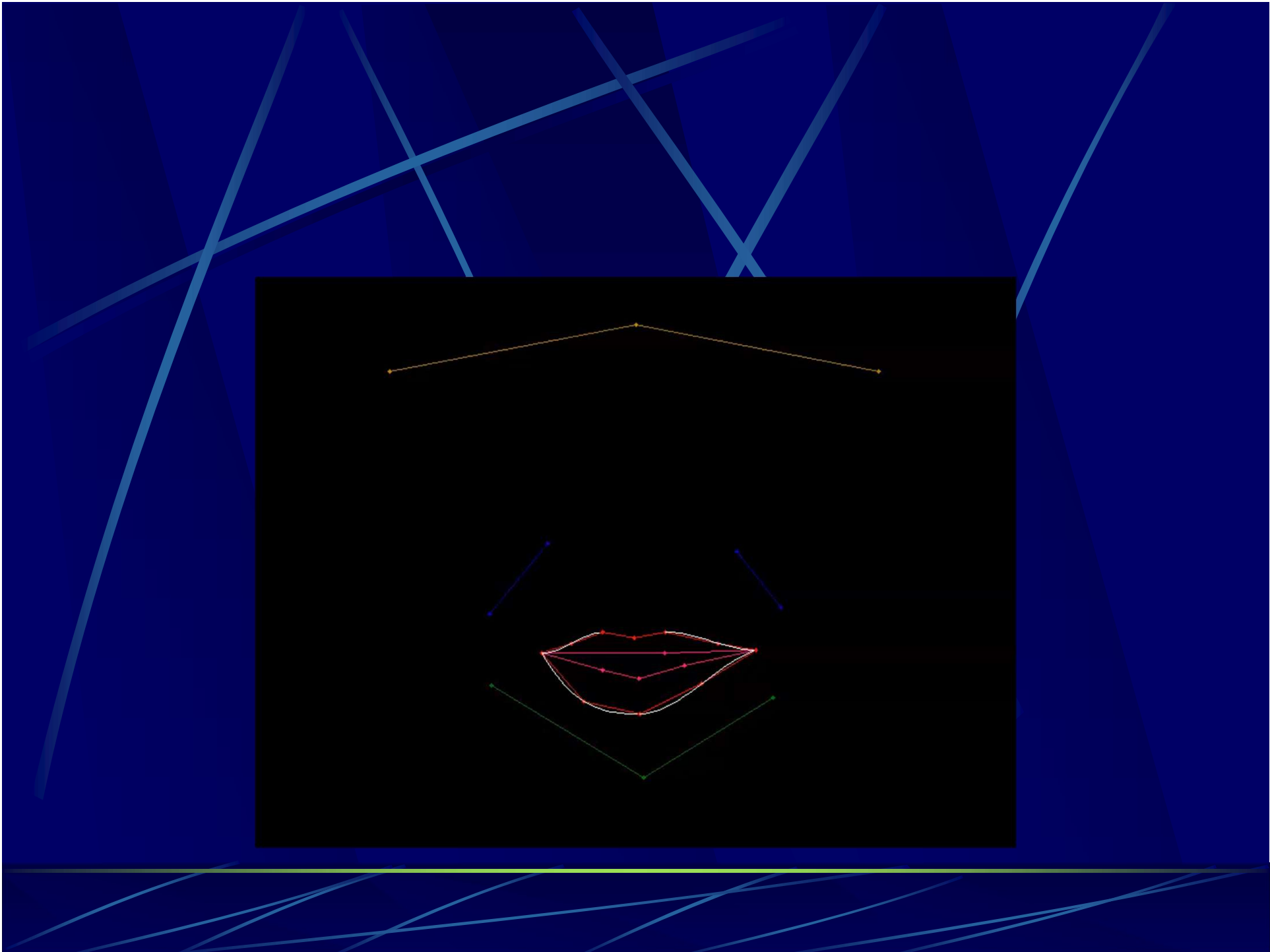
Rozpoznawanie mowy

Dwa podejścia:

- Przed właściwą pracą użytkownik uczy komputer wszystkich niezbędnych słów.
- Baza słów jest zdefiniowana, użytkownik uczy komputer charakterystyki swojej wypowiedzi.

Problemy:

- Mowa ciągła (trudne do odróżnienia pojedyncze słowa)
- Należy zapewnić niezależność mówcy
- Ograniczenia wynikające z wielkości słownika (komputer musi znać wszystkie niezbędne słowa)
- Hałas tła komplikuje rozpoznawanie

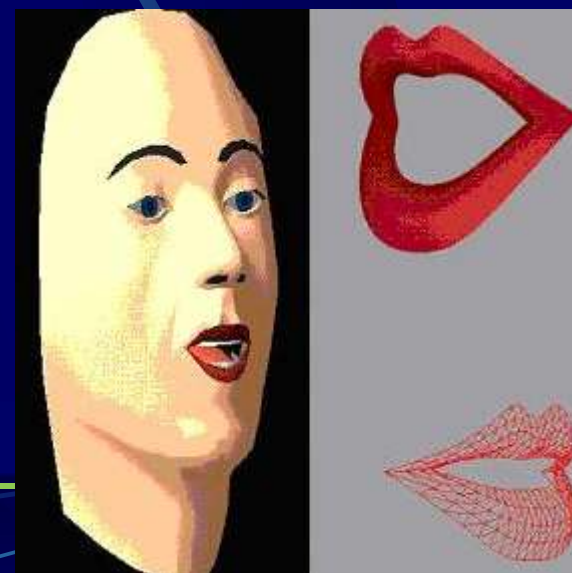


AVATAR

Człowiek < - Komputer (sztuczny rozmówca)

Generowanie obrazu mimiki twarzy, z którego użytkownik może uzyskać dodatkowe informacje (dodatkowy kanał):

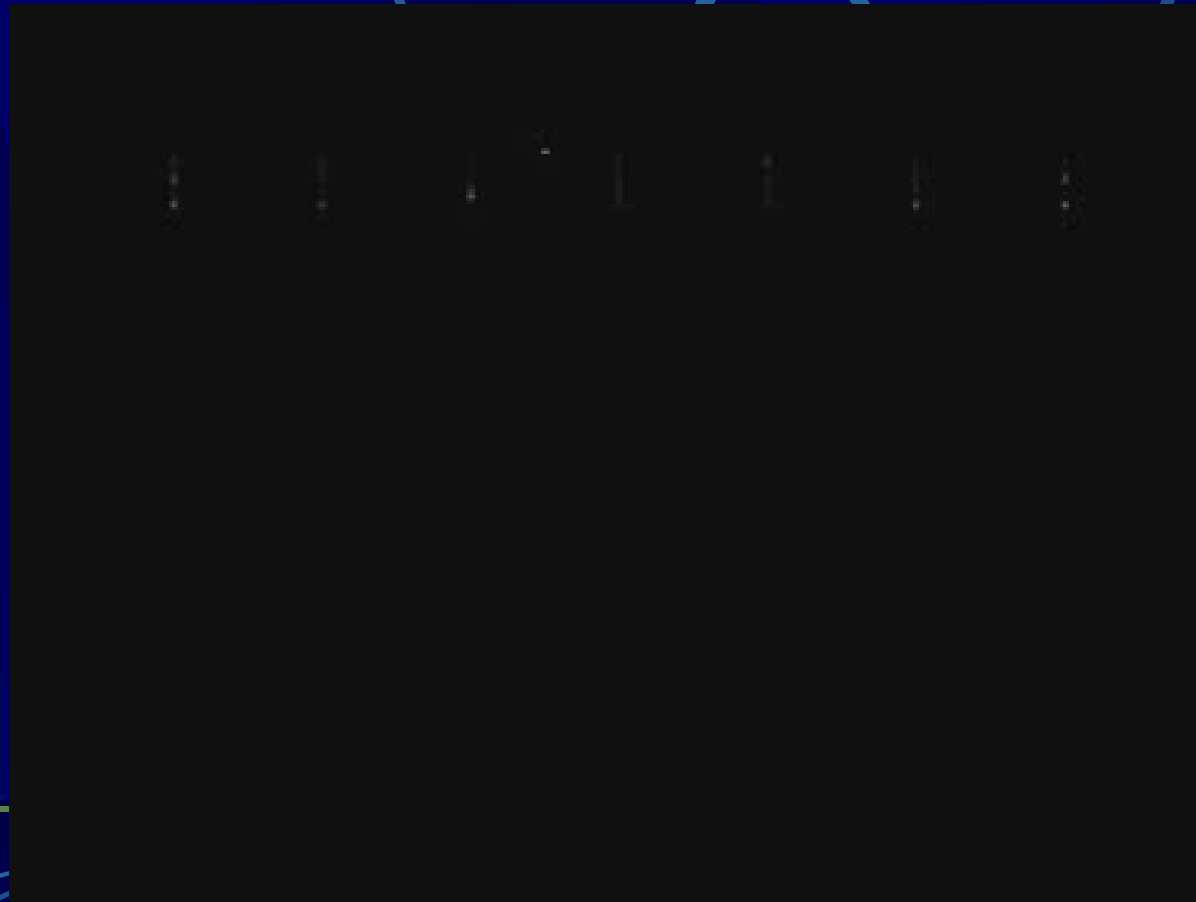
- W 1994 roku interfejs systemu informacyjnego produktów SONY posiada model twarzy, który zbudowany jest z 500 wielokątów, co pozwala na wprowadzenie **26 różnych wyrazów twarzy**.
- Szybki rozwój animacji 3D pozwala teraz na wygenerowanie naturalnie wyglądającego mimika twarzy wraz z realistycznymi szczegółami (np. mrugnięciami oczu, poceniem się, włosami itp.)
- Wirtualni prezenterzy (również w Polsce)



model twarzy zbudowany z wielokątów...



Animacja 3D pozwala na
wygenerowanie naturalnie
wyglądającej twarzy...



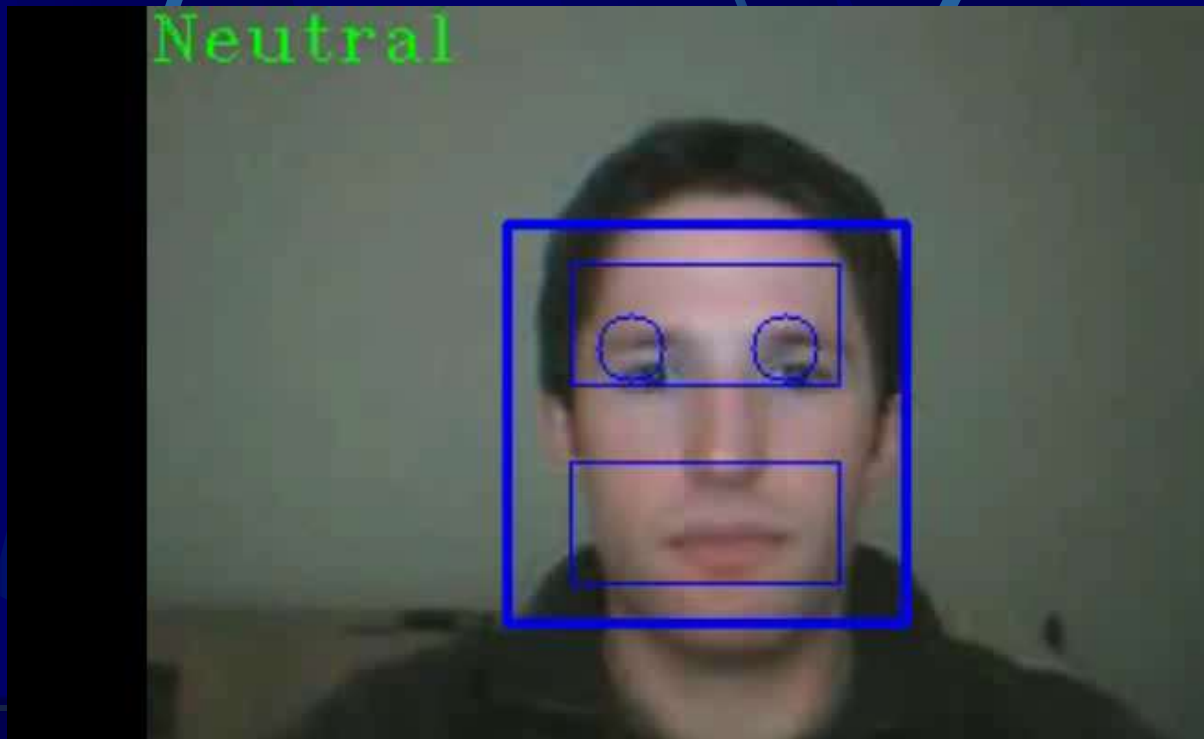
Człowiek (wyrażający nastroje) Komputer

Rozpoznawanie mimiki twarzy to złożony problem, który łączy przetwarzanie i rozpoznawanie obrazu z fizjologią i psychologią.

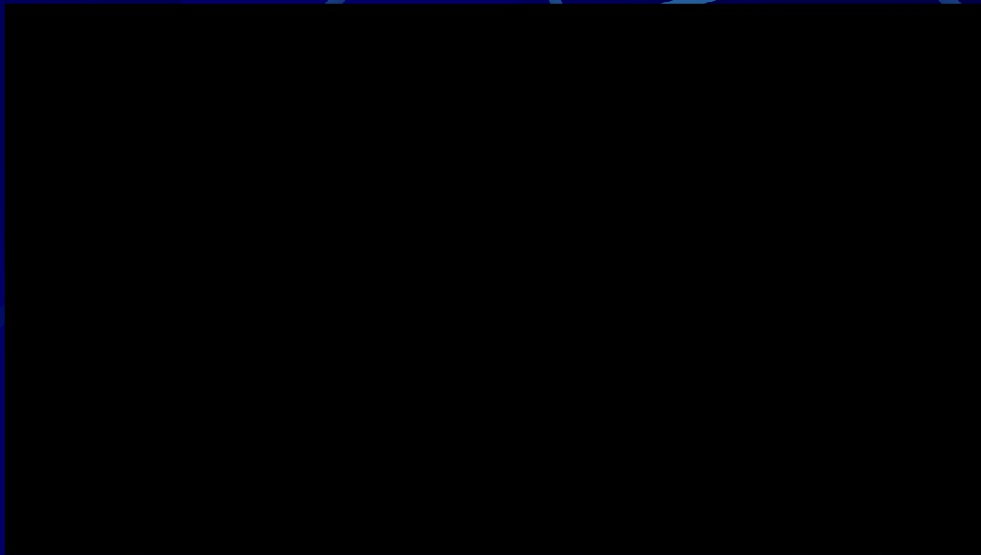
- Istnieją laboratoryjne wersje systemów analizy wyrazu twarzy



Neutral



Rozpoznawanie emocji



NEUTRAL
FEMALE, 22

NEUTRAL
FEMALE, 22

NEUTRAL
FEMALE, 22

NEUTRAL
FEMALE, 22

NEUTRAL
FEMALE, 22

- Anger (17%)
- Happiness (15%)
- Disgust (21%)
- Surprise (26%)
- Sadness (16%)
- Fear (5%)

Visage Technologies

ROZPOZNAWANIE GESTÓW:

- Gesty, ruchy rąk lub język ciała opisują stan człowieka i jego uczucia. Dlatego rozpoznawanie gestów pozwala na zrozumienie czyichś intencji.
- Dobrym przykładem użyteczności takiej umiejętności jest **komunikacja osób niesłyszących z językiem migowym.**



- Naturalnie rozpoznawanie wszystkich możliwych gestów jest bardzo trudnym zadaniem, zwłaszcza, że niektóre gesty mają różne znaczenia dla różnych osób.



- Zasadniczo istnieją dwie kategorie rozpoznawania gestów dłoni:
 - w oparciu o obraz ruchu dłoni lub
 - bez rejestracji obrazu.
- Pierwsze badania nad rozpoznawaniem gestów zostały przeprowadzone bez rejestracji obrazu. Na przykład tzw. rękawiczki do danych (**DataGlove**) z czujnikami, potrafiły odróżnić około 10 pozycji poruszającej się dłoni.



Scena z "Minority Report"
Tom Cruise używa rękawic DataGlove.

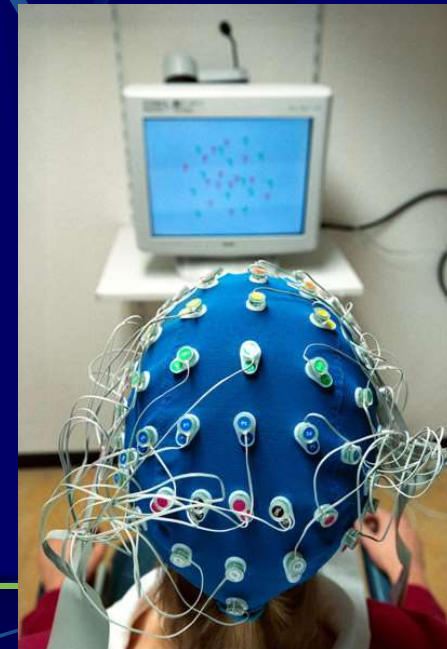
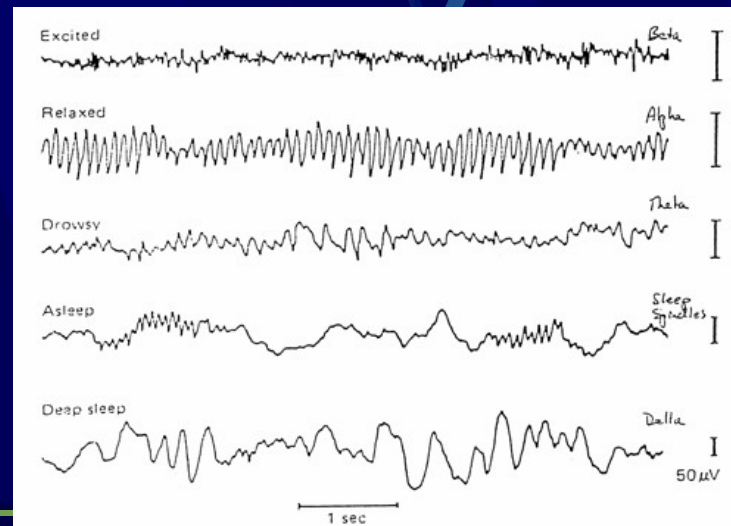
- Obecnie większość badań koncentruje się na rozpoznawaniu opartym na przetwarzaniu obrazu
- Wiele wysiłku wkłada się w eliminację konieczności stosowania wszystkich połączeń kablowych, czujników, rękawic czy innych dodatkowych narzędzi.

ZASTOSOWANIA OPARTE NA BIOFEEDBACKU

- Dzięki zastosowaniu multimodalnych interfejsów, które obsługują różne rodzaje **sprzężenia zwrotnego opartego na sygnałach biometrycznych**, człowiek może świadomie kontrolować swój stan.
- Sygnały są rejestrowane za pomocą różnych czujników (np. przyspieszeniomierzy, żyroskopów, czujników EMG, czujników galwanicznej reakcji skóry, elektrod EEG, elektrod EKG).

SYSTEMY NEUROKONTROLI

W aplikacji neurofeedback zazwyczaj nagrany sygnał jest podzielony na oddzielne **rytmy mózgu** i można podjąć inną decyzję, na podstawie której rytm dominuje.



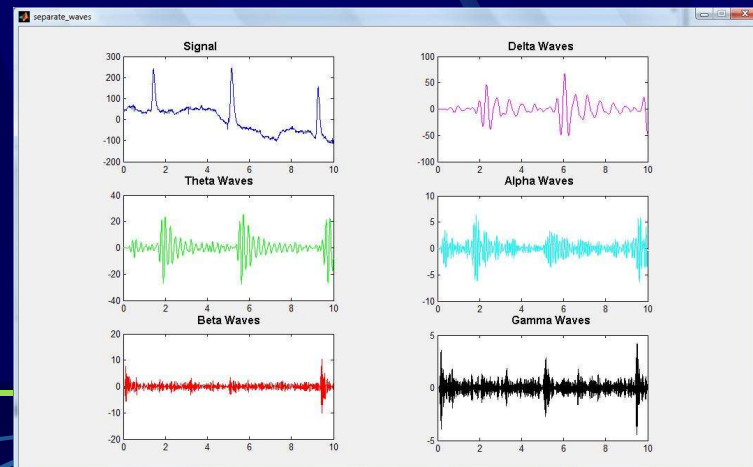
INTERFEJS OPARTY NA BIOFEEDBACKU DO SYNCHRONIZACJI PÓŁKUL

- System synchronizacji półkul wykorzystuje **hełm do pomiaru aktywności mózgu**

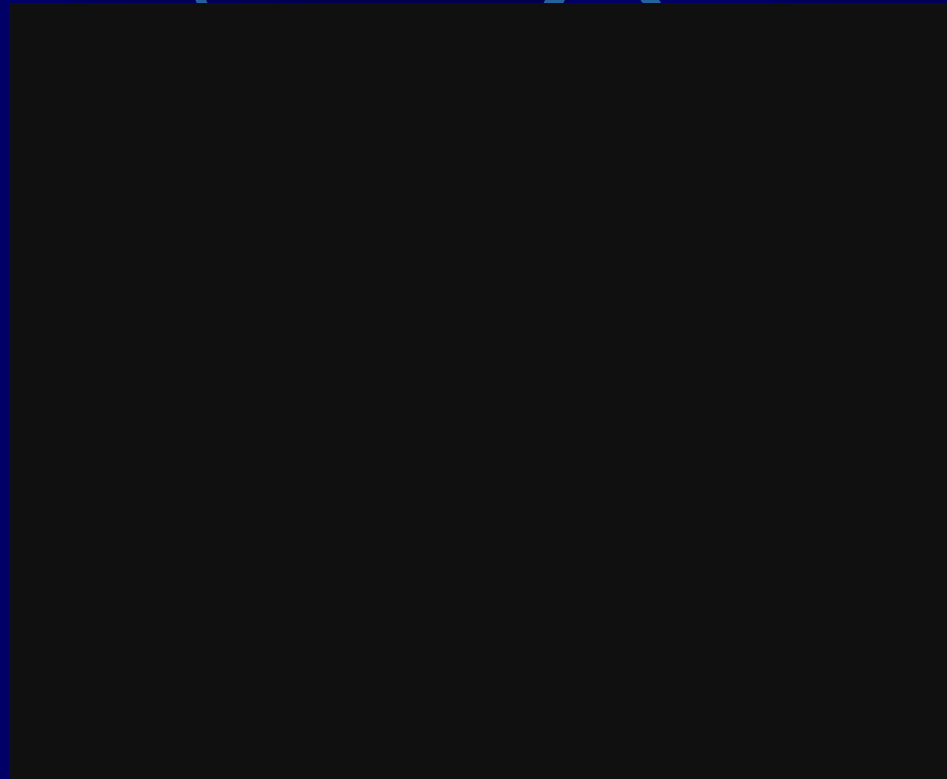


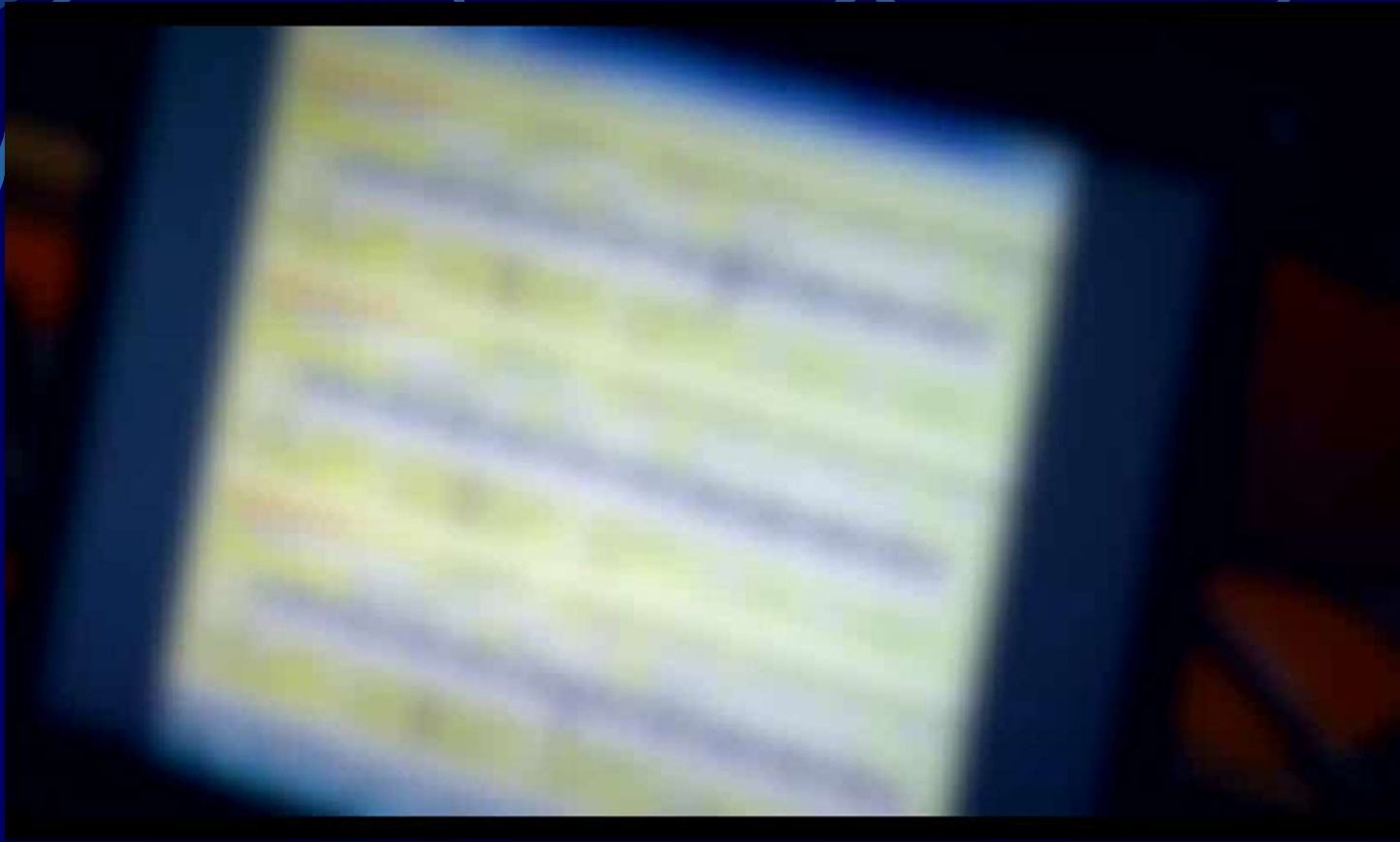
INTERFEJS OPARTY NA BIOFEEDBACKU DO SYNCHRONIZACJI PÓŁKUL

- Artefakty są usuwane z sygnału poprzez zastosowanie filtracji adaptacyjnej.
- Sygnał jest podzielony na oddzielne rytmy mózgowie (np. delta, alfa, beta)
- Interfejs zapewnia wizualną informację zwrotną o stanie synchronizacji



INTERFEJS OPARTY NA
BIOFEEDBACKU DO SYNCHRONIZACJI
PÓŁKUL MÓZGOWYCH

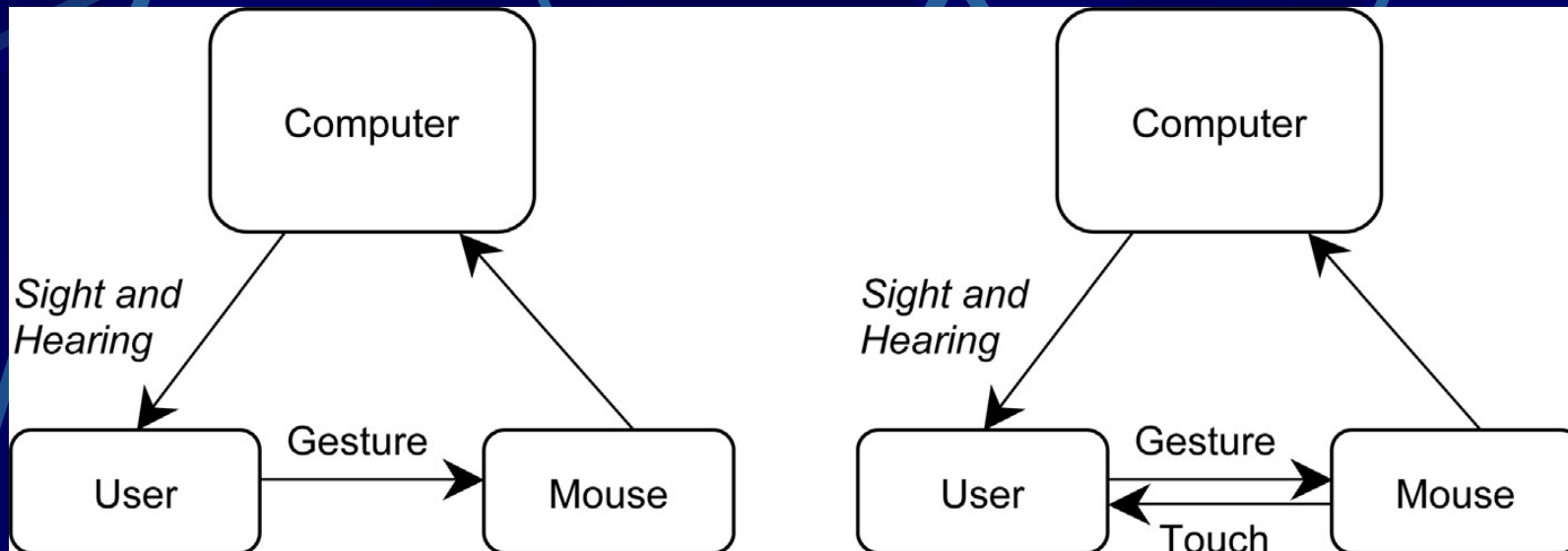




INTERFEJSY HAPTYCZNE

- **Haptyczne interfejsy** umożliwiają komunikację człowiek-maszyna poprzez **dotyk** i najczęściej w odpowiedzi na ruchy użytkownika.
- Haptyczne interfejsy generują mechaniczne sygnały, które **stymulują ludzkie kanały kinestetyczne i dotykowe**. Powszechnie ich użycie związane jest z rozpoznawaniem gestów.

INTERFEJSY HAPTYCZNE



Cechą wyróżniającą interfejsy haptyczne jest jednoczesna wymiana informacji pomiędzy użytkownikiem a maszyną.



PRZEGLĄD INTERFEJSÓW
MULTIMODALNYCH

Komputery w sali lekcyjnej

Może pociągać za sobą następujące niekorzystne skutki:

- izolować uczniów
- odwrócić ich uwagę od nauczyciela
- zerwać więzi emocjonalne między uczniami
- zapobiegać socjalizacji podczas lekcji
- niekorzystnie zmieniać nawyki pracy zespołowej
- pogorszenie ostrości wzroku
- negatywny wpływ na postawę ciała
- patologia nadgarstków



Skomputeryzowana sala lekcyjna?



Zdalny
Kontroler
dla monitora

Zaawansowany wyświetla
Monitor



Teza: rozwój technologii może prowadzić nas do bardziej naturalnego sposobu korzystania z komputerów w ogóle, również w klasach.



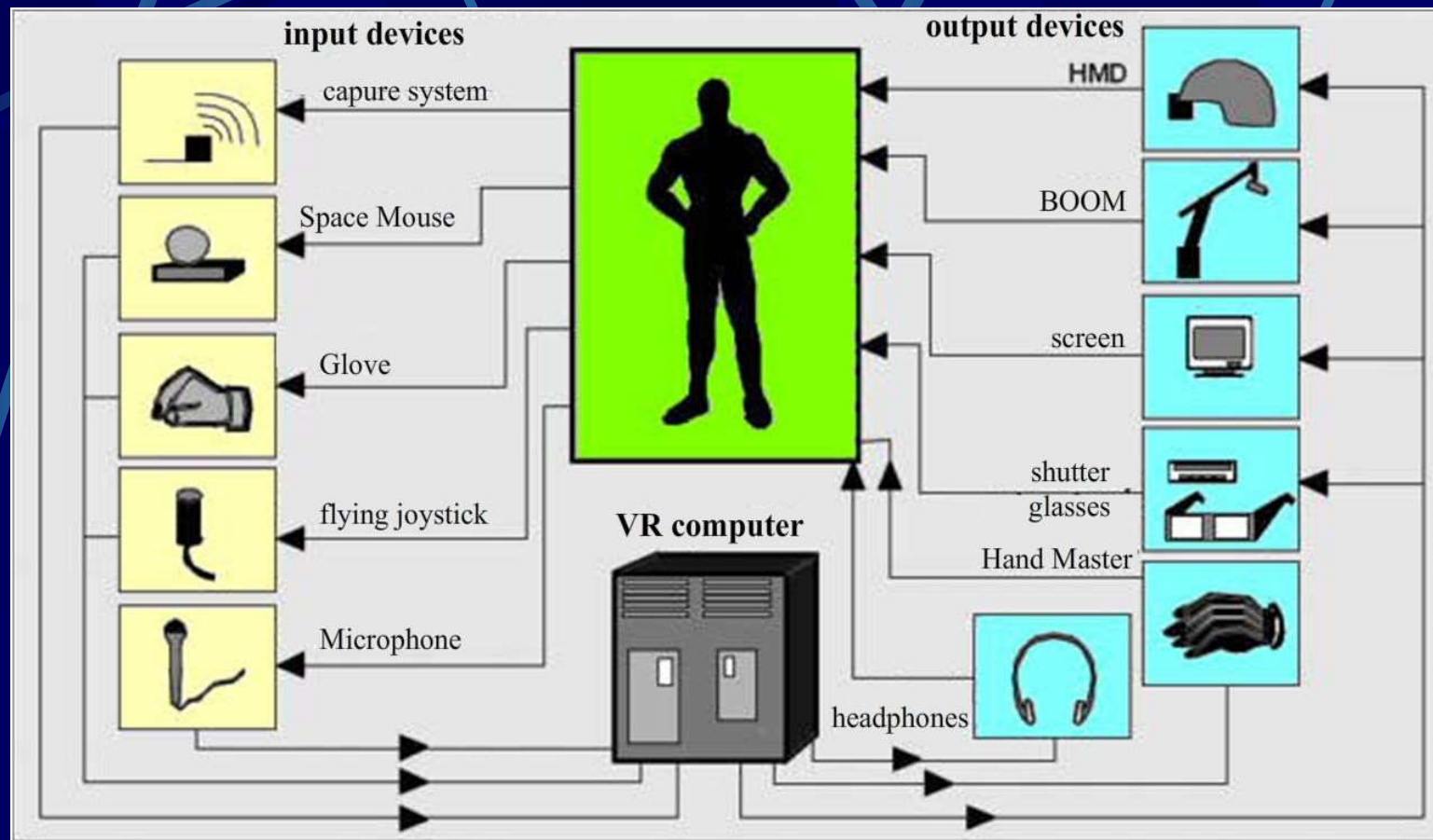
TRÓJWYMIAROWY
INTERFEJS MULTIMODALNY

To jest wirtualna rzeczywistość

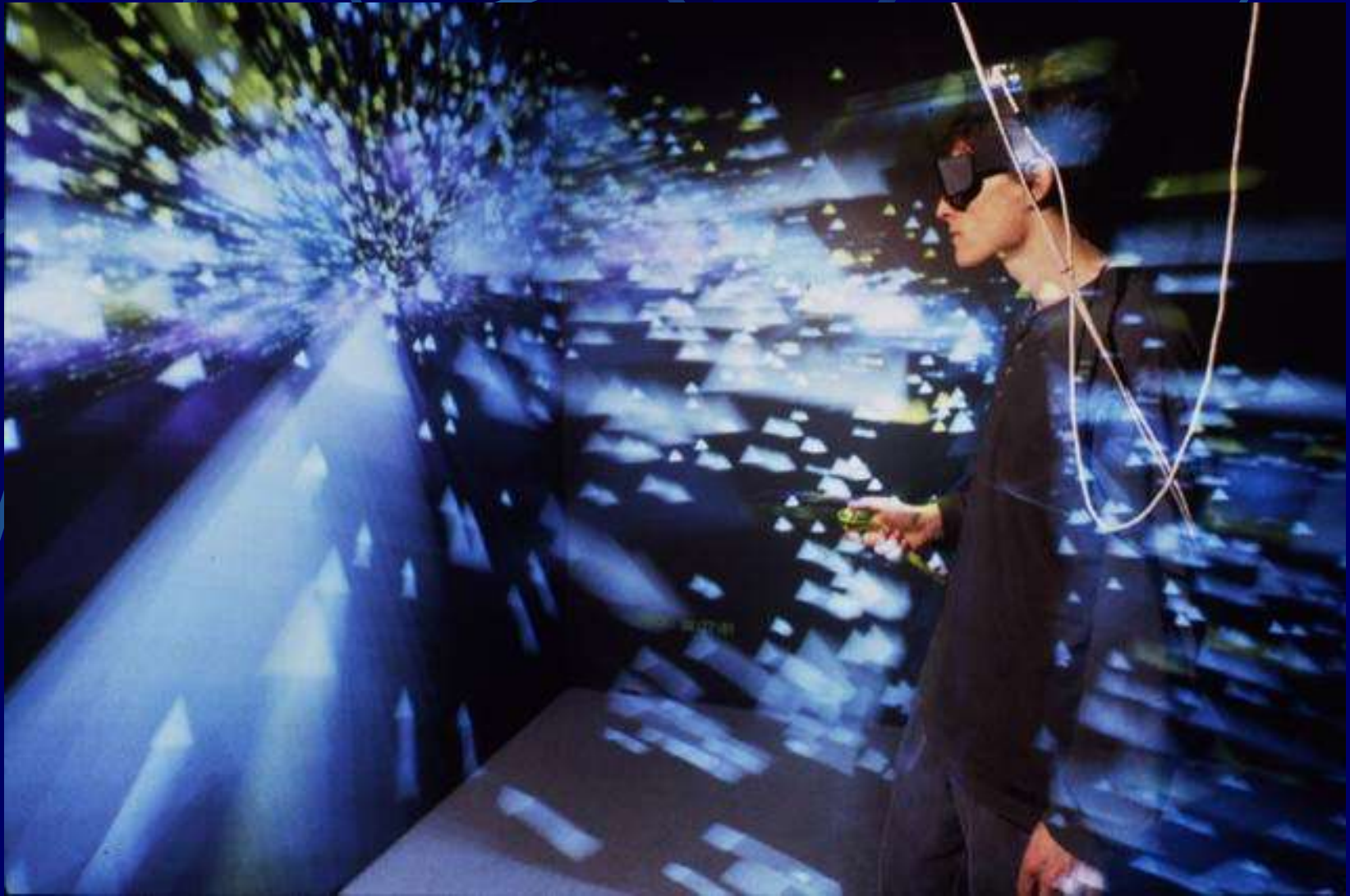
- **Wirtualna rzeczywistość** - technika, która tworzy iluzję rzeczywistości poprzez odpowiednie pobudzanie ludzkich zmysłów za pomocą różnych urządzeń podłączonych do komputera.
- Obecnie możliwe jest stworzenie realistycznego efektu trójwymiarowego obrazu.
- Sygnał stymulacyjny może być odpowiednio modyfikowany dzięki **odbieraniu naturalnych reakcji człowieka**. Obecnie mogą być odbierane proste reakcje, takie jak ruchy rąk lub głowy.

- **Czujniki ruchu** - analizują położenie i orientację obiektów w przestrzeni i wysyłają zebrane informacje do komputera.
- **Rękawice cyfrowe** - działają podobnie jak czujniki ruchu, ale ograniczają się tylko do ruchów rąk, czasami wbudowany jest moduł oporowy.
- **Manipulatory** - trójwymiarowa mysz, joystick, itp.

SCHEMAT ŚRODOWISKA RZECZYWISTOŚCI WIRTUALNEJ



Systemy immersyjne

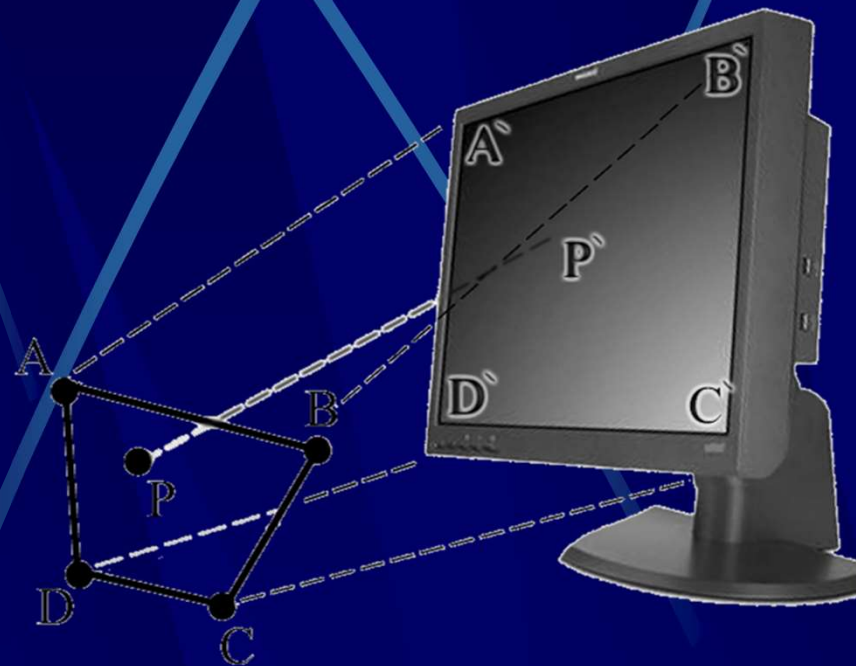
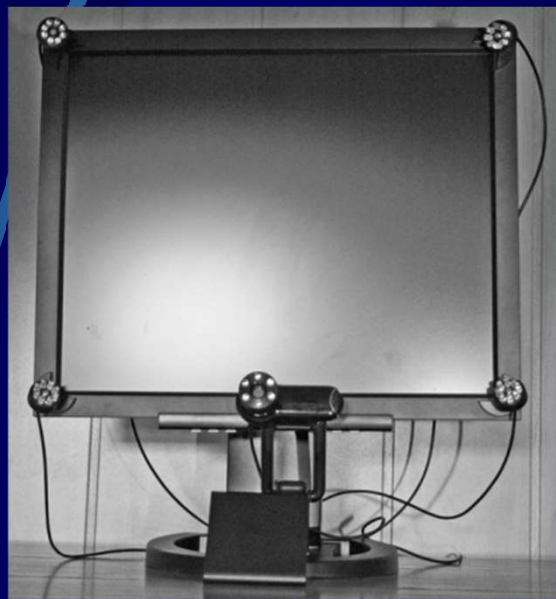




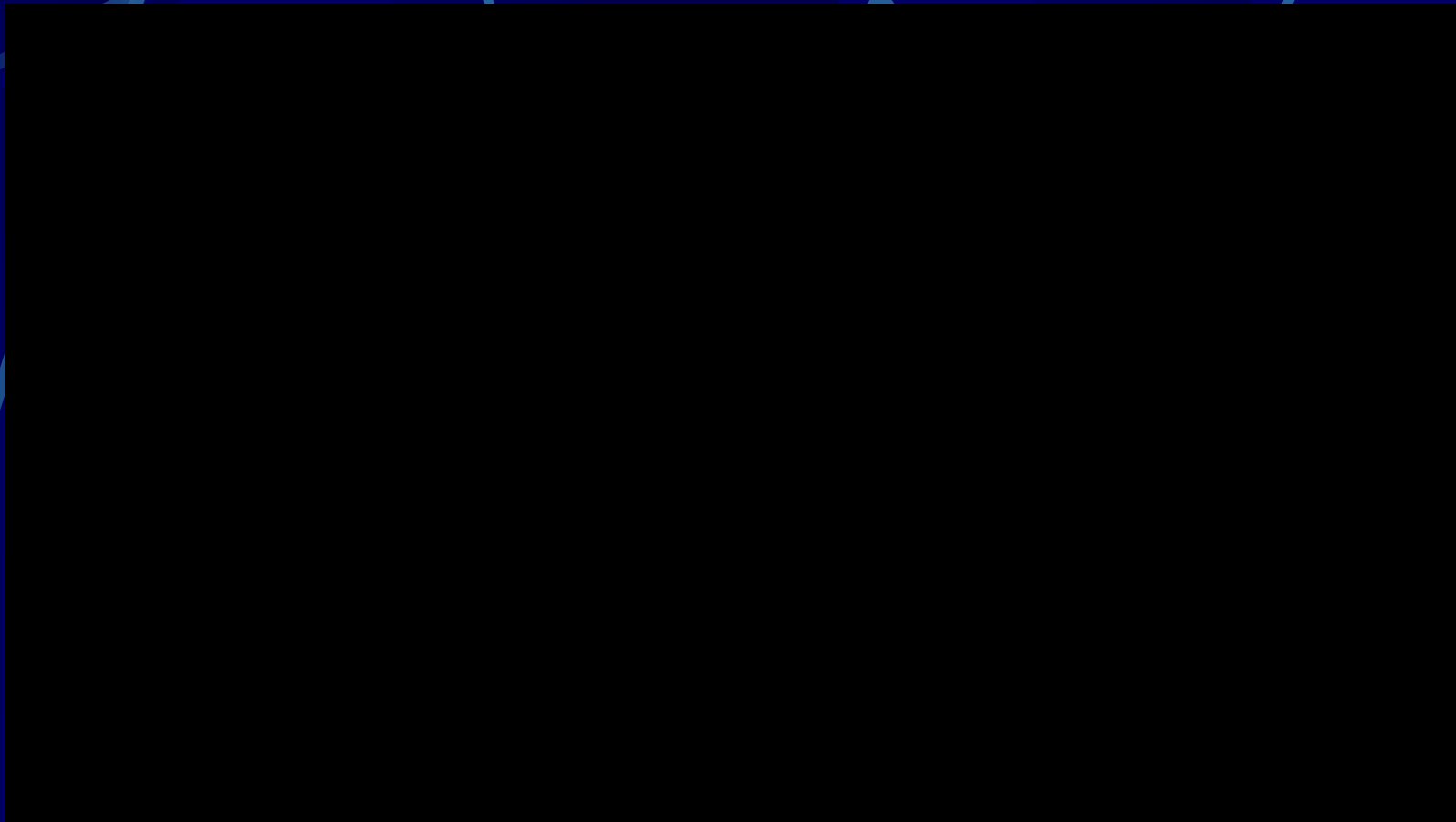
MODALNOŚĆ WZROKOWA – GAZE TRACKER

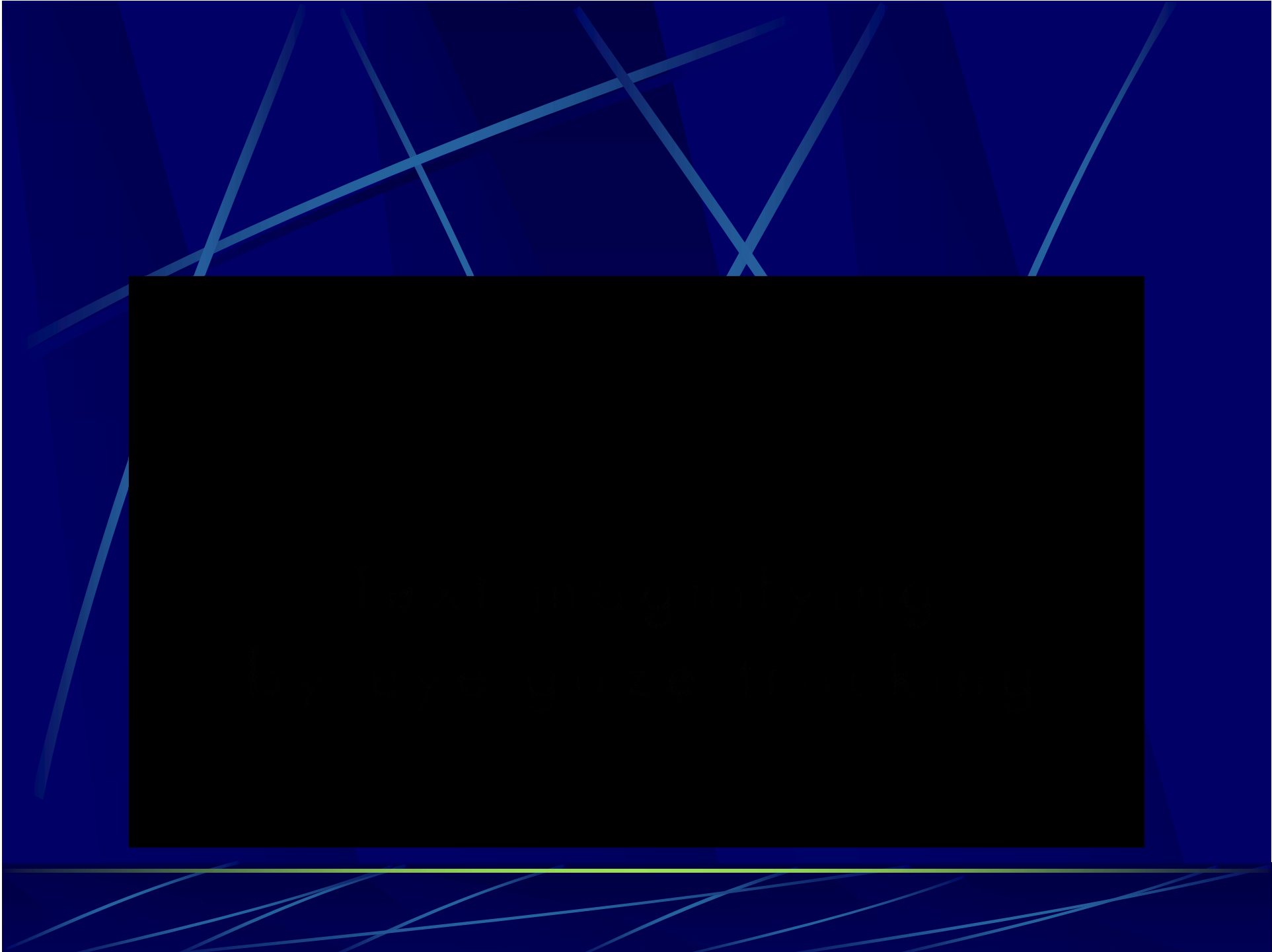
- Cztery moduły diod podczerwieni umieszczone w narożnikach wyświetlacza wytwarzają unikalne refleksy rogówkowe (odblaski).
- Algorytm analizuje każdą klatkę filmową wykonaną przez kamerę obserwującą twarz użytkownika i lokalizującą w niej cztery odbicia.
- Punkt fiksacji jest szacowany dla każdego oka niezależnie, na podstawie informacji o położeniu wszystkich wykrytych punktów: szkieł i ośrodków źrenicowych.

MODALNOŚĆ WZROKOWA – GAZE TRACKER



MODALNOŚĆ WZROKOWA – GAZE TRACKER





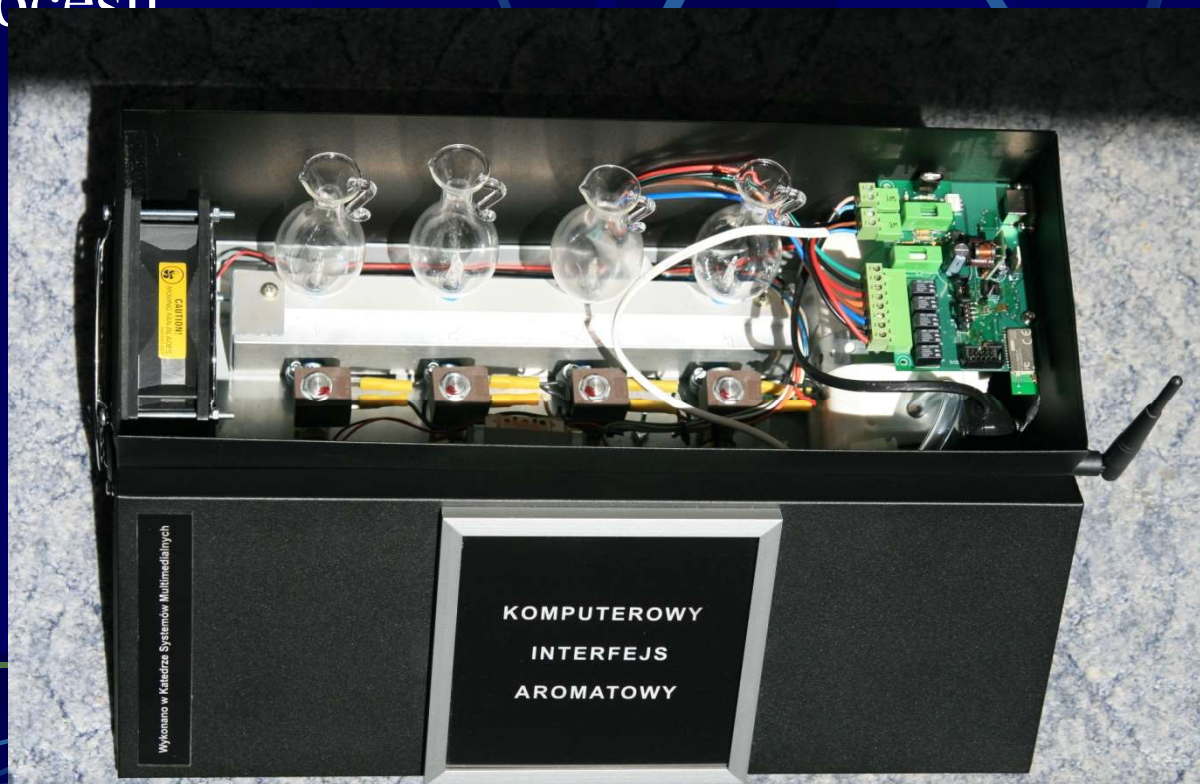


MULTIMODALNY INTERFEJS KOMPUTEROWY EMITUJĄCY ZAPACH

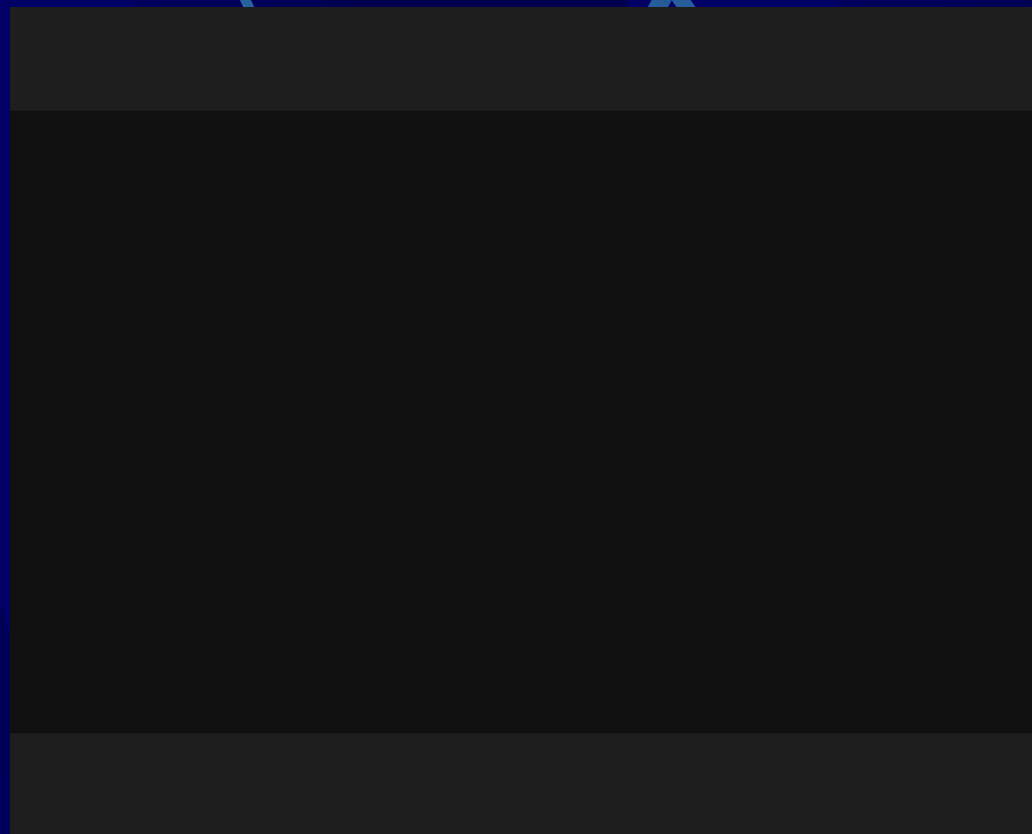
- Interfejs emituje różne zapachy w zależności od preferencji użytkownika, stanu i stanu pomieszczenia (np. lawenda o działaniu uspokajającym, cytrynowy zapach poprawiający koncentrację).
- Interfejs jest przeznaczony dla dzieci z zaburzeniami uwagi, w celu poprawy warunków nauki w klasach.
- Zastosowano innowacyjną metodę dyfuzji zimnego powietrza.

MULTIMODALNY INTERFEJS KOMPUTEROWY EMITUJĄCY ZAPACH

- Emisja zapachu jest kontrolowana przez interfejs użytkownika, co umożliwia pełną regulację procesu



MULTIMODALNY INTERFEJS KOMPUTEROWY EMITUJĄCY ZAPACH

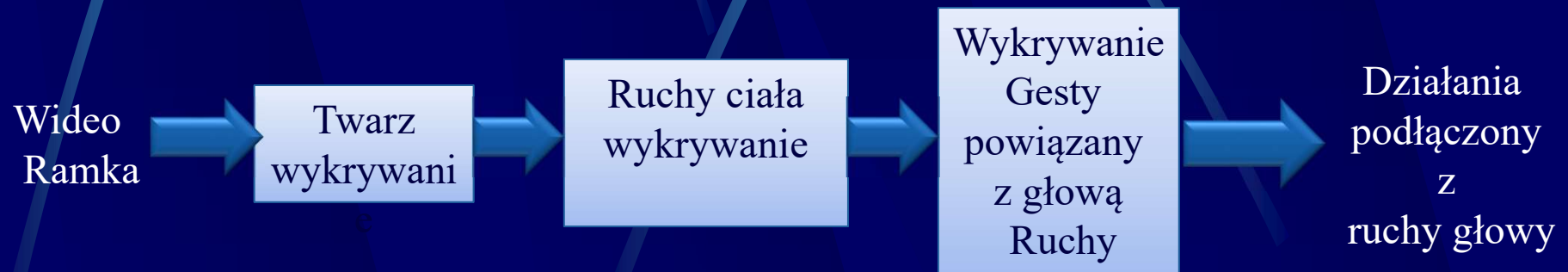


ROZPOZNAWANIE RUCHÓW GŁOWY - PRZEGLĄDARKA MULTIMEDIALNA

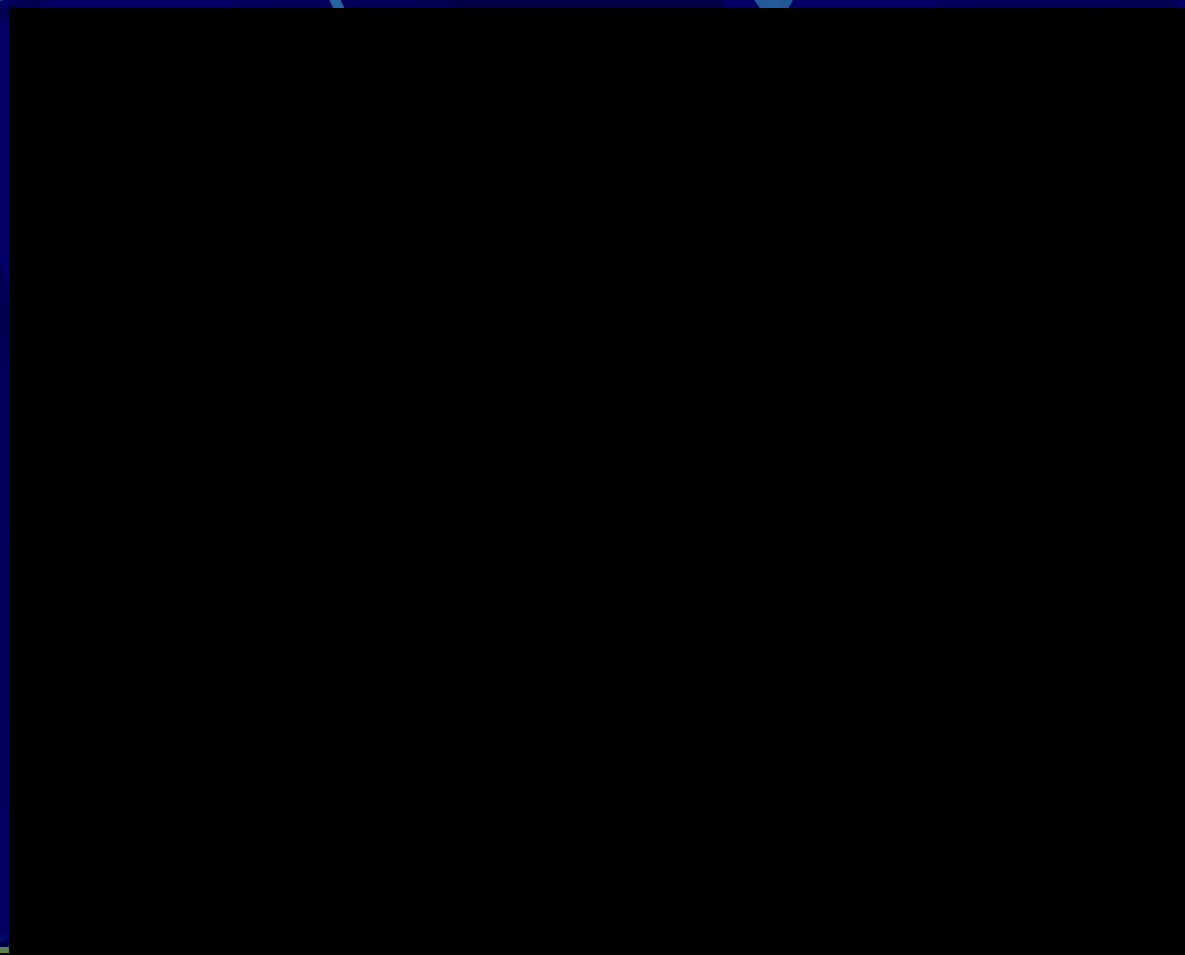
- Interfejs został zaprojektowany do **przeglądania multimediiów**
- Opracowany interfejs pozwala na przeglądanie **statycznych zdjęć, plików wideo i audio**
- Pliki są ułożone w hierarchicznej strukturze
- Użytkownik porusza się po strukturze poruszając głową w lewo, w prawo, w górę i w dół.
- Wykonywana akcja zależy od rodzaju aktualnie przeglądanej pliku (**przewijanie pliku wideo, przejście do następnego obrazu**)

ROZPOZNAWANIE RUCHU GŁOWY - PRZEGLĄDARKA MULTIMEDIALNA

- Ruchy głowy są wykrywane poprzez analizę strumienia wideo z kamery internetowej



ROZPOZNAWANIE RUCHU GŁOWY - PRZEGLĄDARKA MULTIMEDIALNA



USTOMYSZ

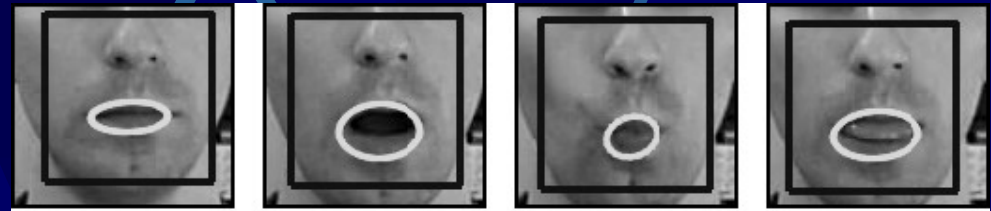
- Może być używany przez osoby z niepełnosprawnymi ruchami rąk
- Pozwala użytkownikowi na pracę na komputerze za pomocą ruchów warg i gestów
- Aplikacja jest uruchomiona na standardowym komputerze PC.
- Obraz z kamery jest przetwarzany
- Wszystkie ruchy ust (głową) są zamieniane na ruchy kursor ekranowy



USTOMYSZ

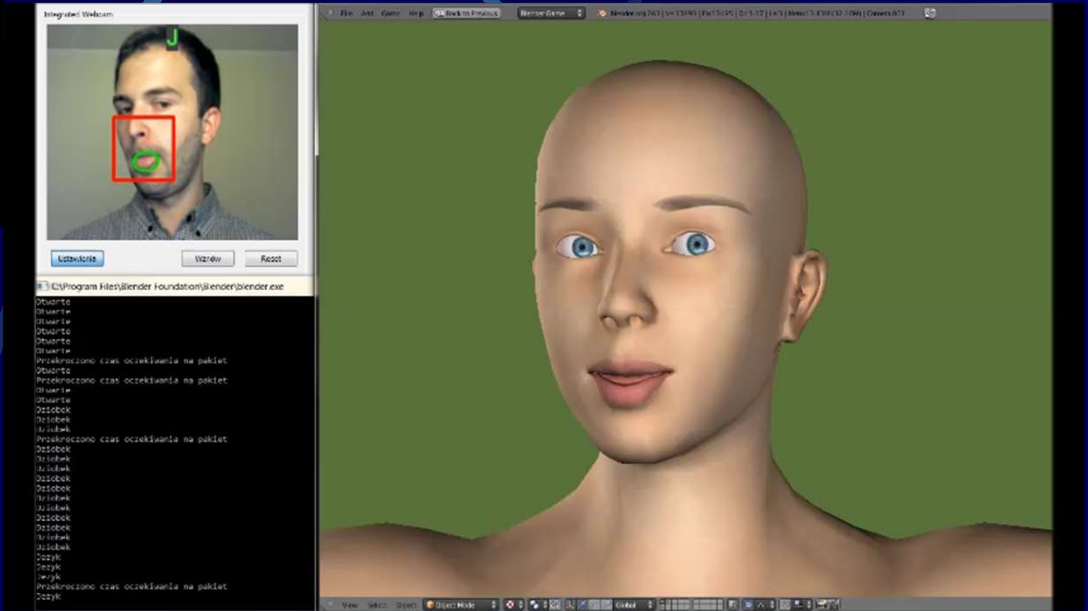
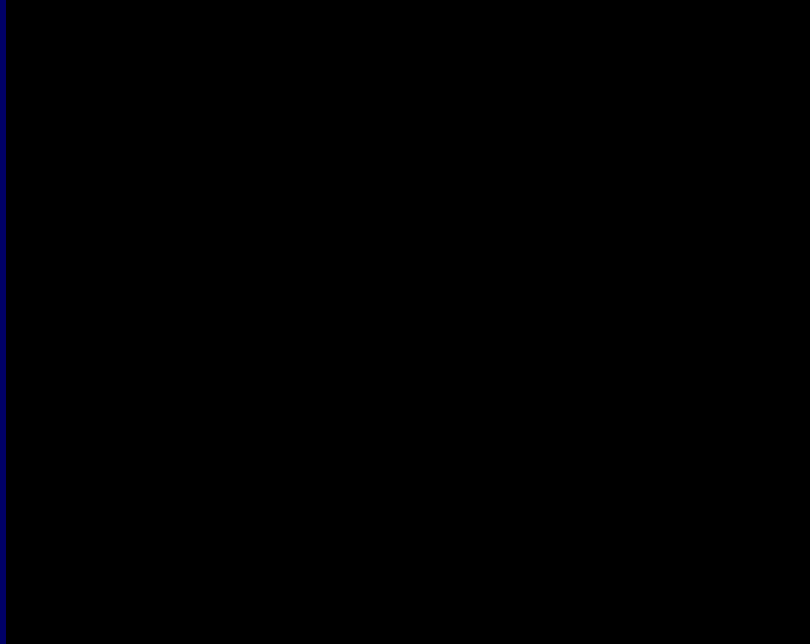
- LipMouse wykrywa gesty ustne:

- Neutralny
- otwarcie ust
- wystający język
- kształtowanie warg



- Każdy gest może być związany z działaniem, które może być dowolnie wybrane przez użytkownika.
- Możliwe działania to **kliknięcie lub dwukrotne kliknięcie różnymi przyciskami myszy oraz poruszanie kółkiem myszy - zarówno w poziomie jak i w pionie.**

STEROWANIE USTAMI



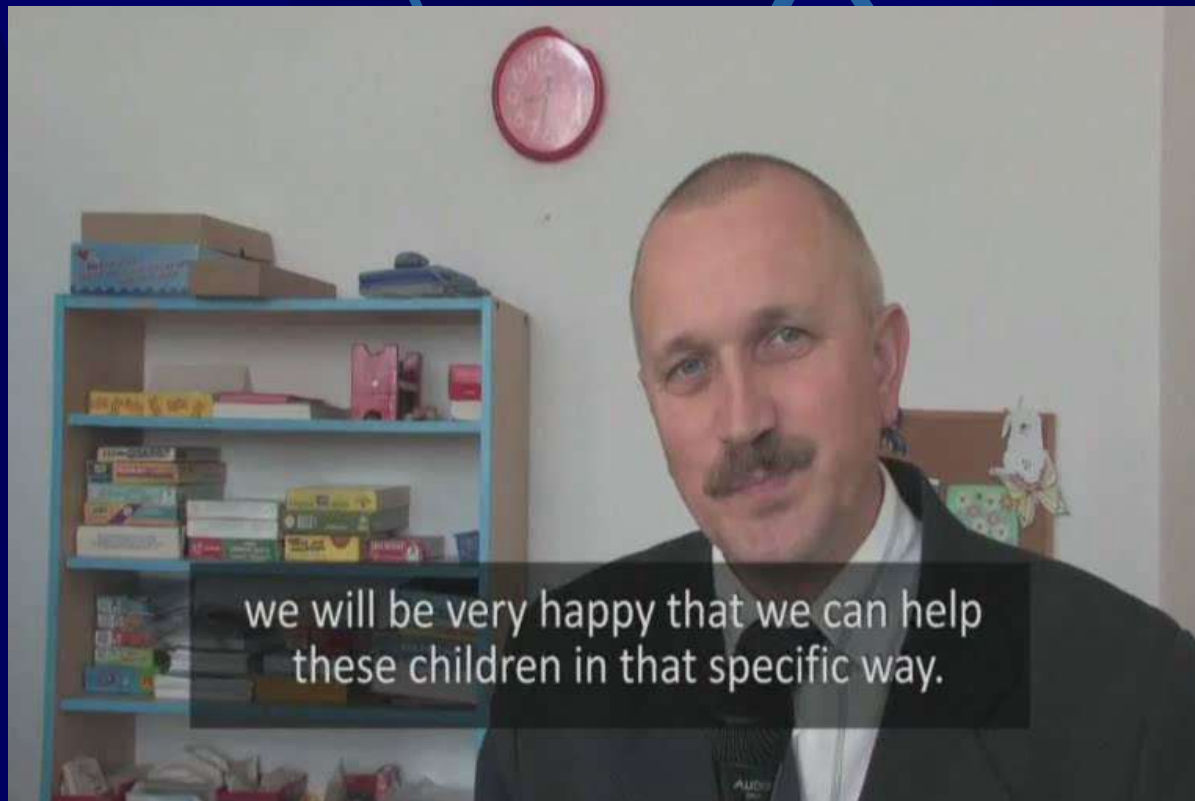


INTELLIGENTNE PIÓRO



- **IntelligentPen** został opracowany głównie dla dzieci z dysleksją.
- Stosowany jest touchpad i specjalny rysik z czujnikami nacisku.
- Starannie opracowane ćwiczenia pomagają dzieciom skoncentrować się na ćwiczeniach
- Dostępne są różne ćwiczenia (np. kolorowanie zdjęć, dopasowywanie numerów)

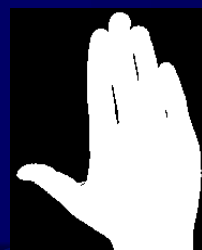
INTELIGENTNE PIÓRO (i LipMouse)



we will be very happy that we can help
these children in that specific way.

ROZPOZNAWANIE GESTÓW - WIRTUALNY TOUCHPAD

- Umożliwia sterowanie aplikacjami komputerowymi za pomocą gestów ręcznych uchwyconych za pomocą typowej kamery internetowej



Parkinson Disease Rating

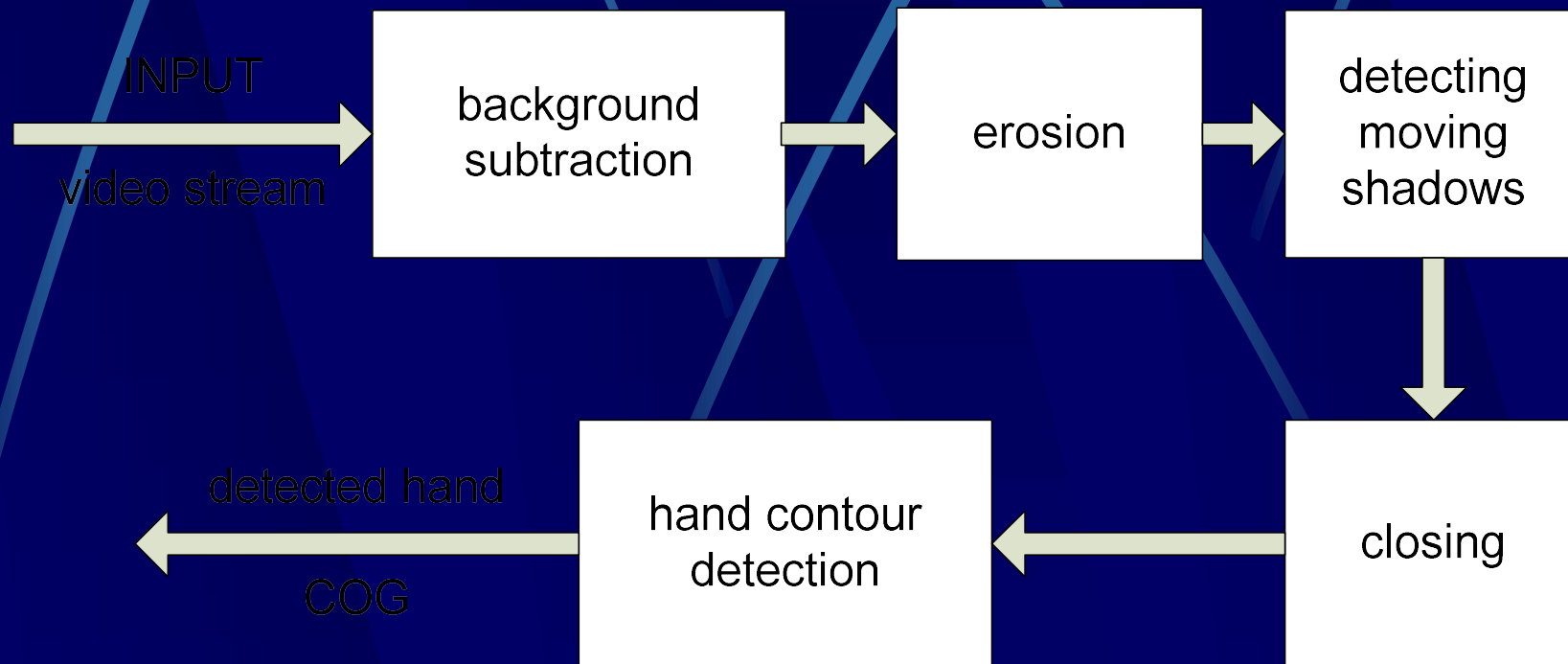
(UPDRS scale); 3D hand model; comparisons...



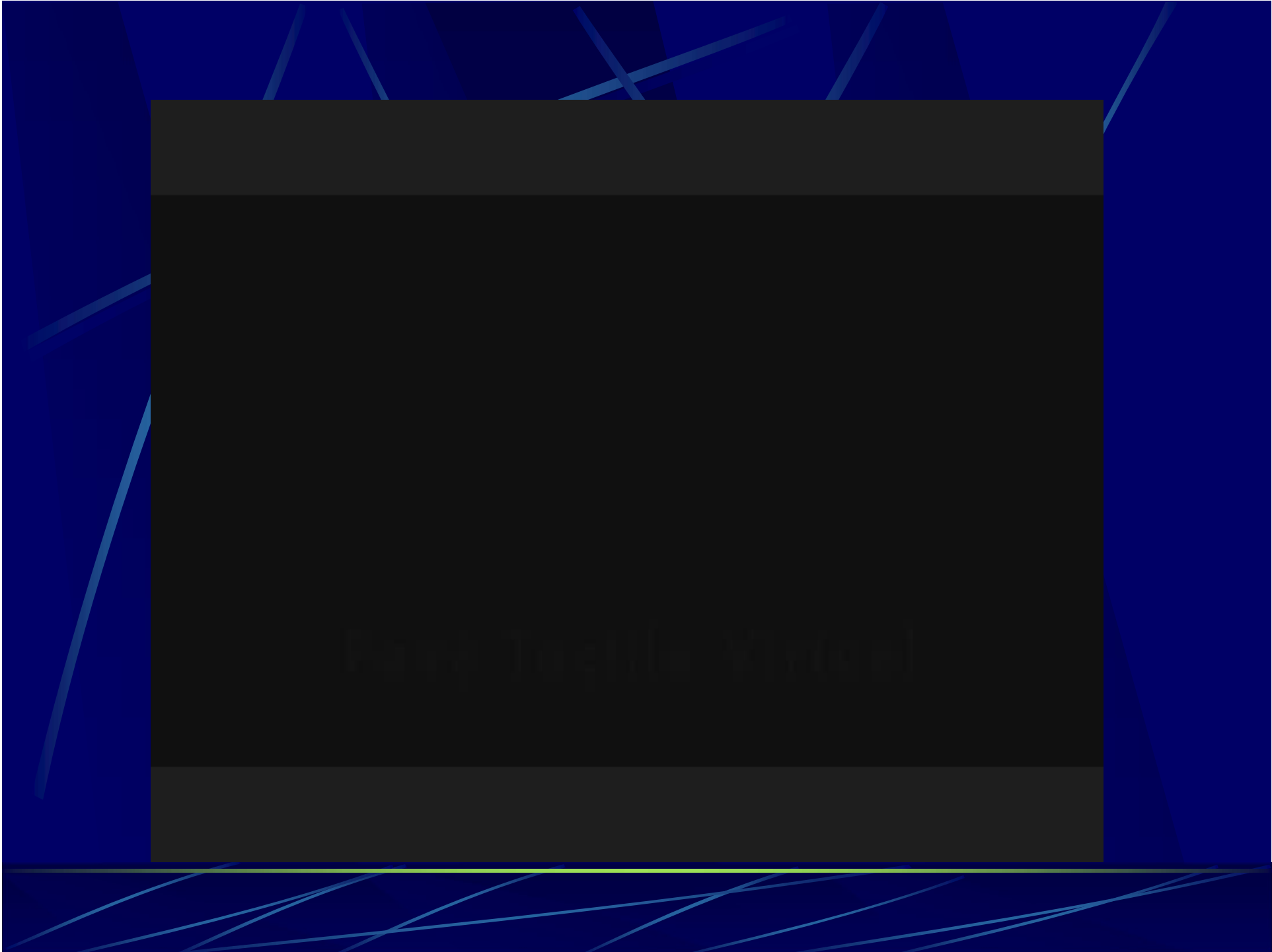
UPDRS = Universal Parkinson Disease Rating Scale



WIRTUALNY TOUCHPAD - WYKRYWANIE RĄK

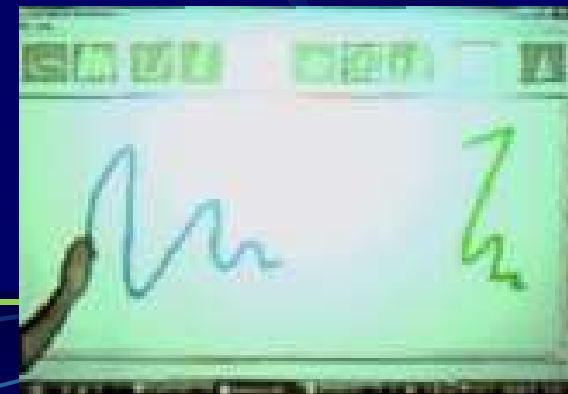




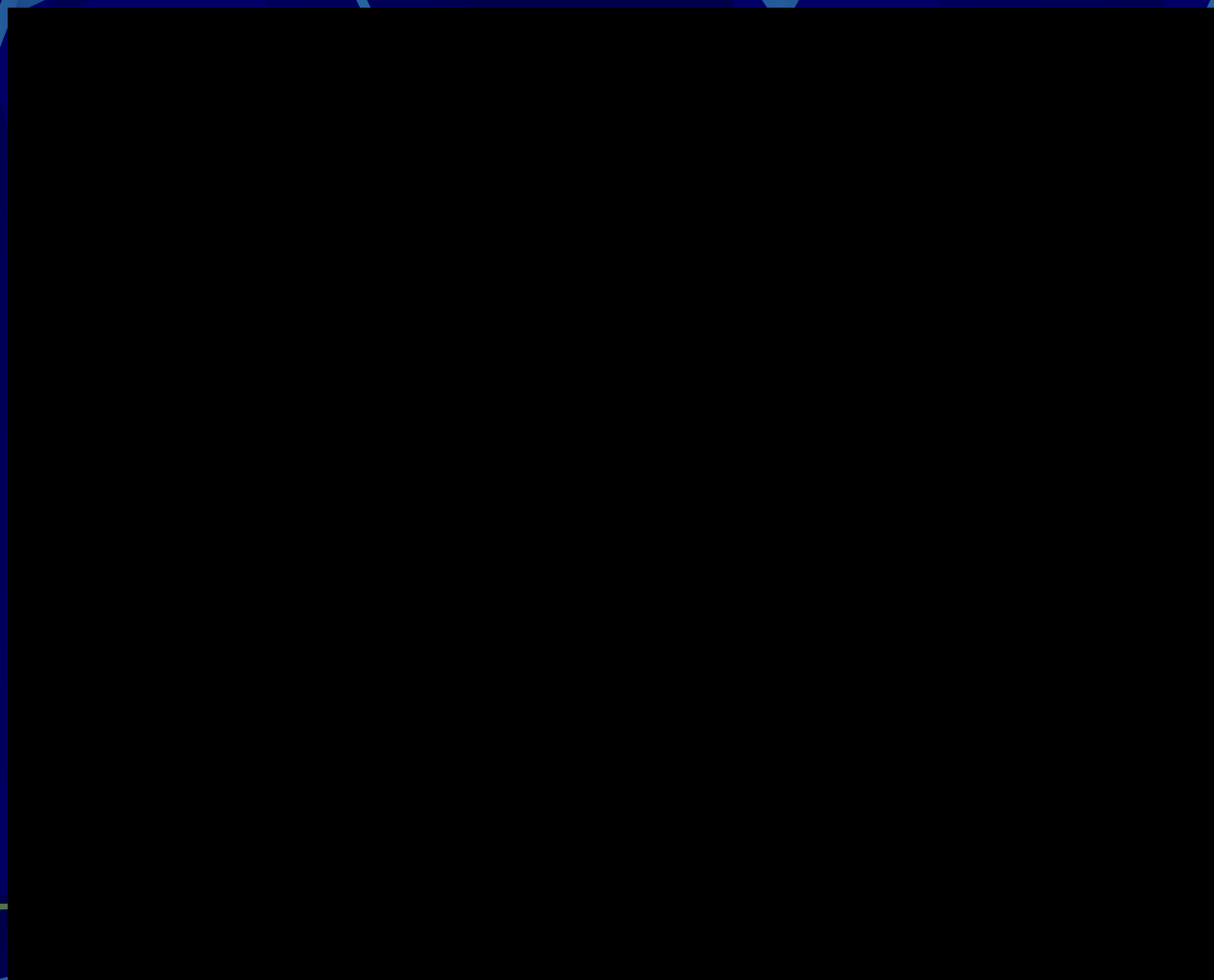


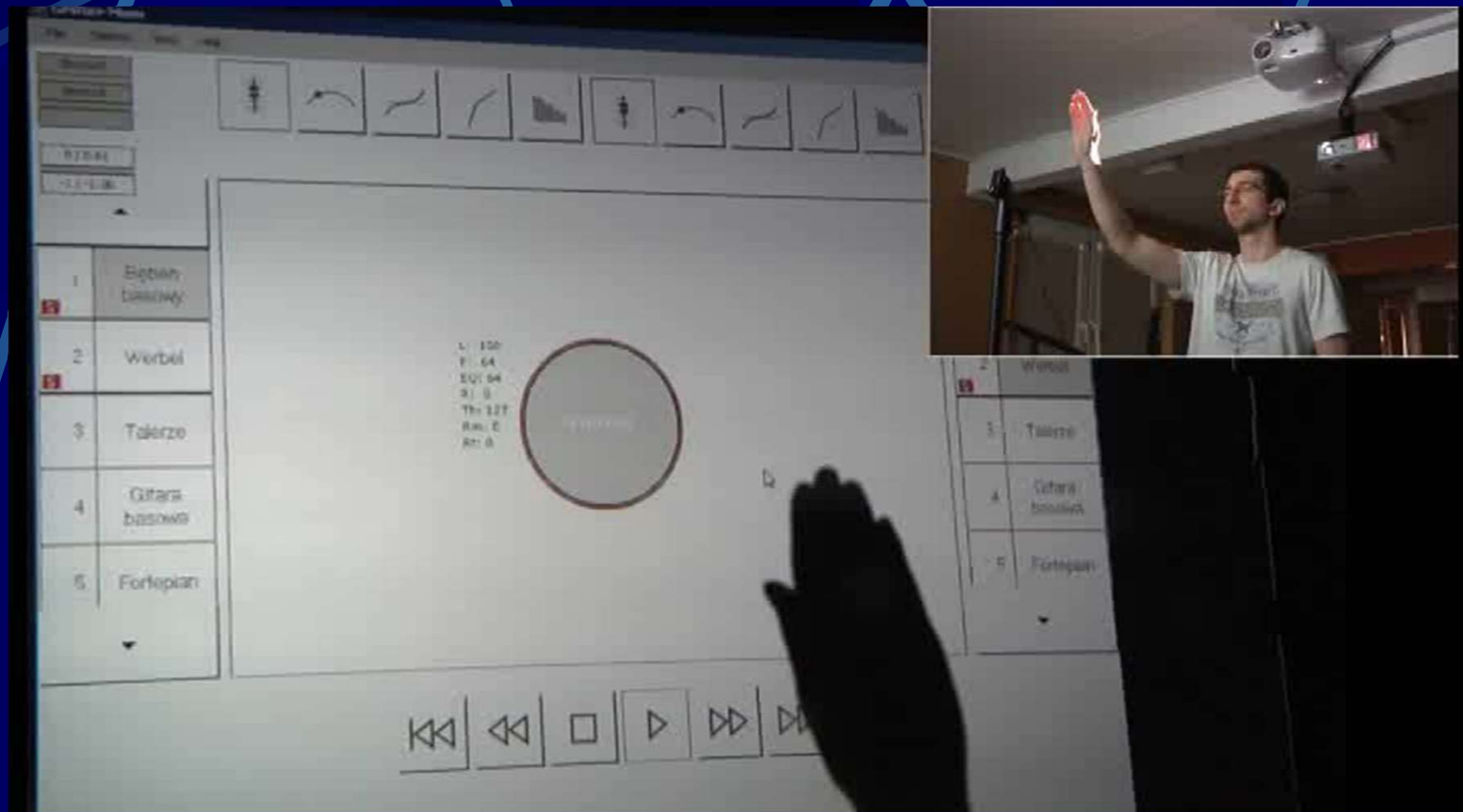
ROZPOZNAWANIE GESTÓW - TABLICA INTERAKTYWNA

- Zastosowanie zwykłej kamerki i projektora multimedialnego do dynamicznego rozpoznawania gestów
- Nie są potrzebne żadne rękawice ani dodatkowe narzędzia
- 10 gestów zostało zdefiniowanych w systemie
- Gesty wykonywane dwiema rękami kojarzone są z typowymi dla systemu poleceniami.



ROZPOZNAWANIE GESTÓW - TABLICA INTERAKTYWNA

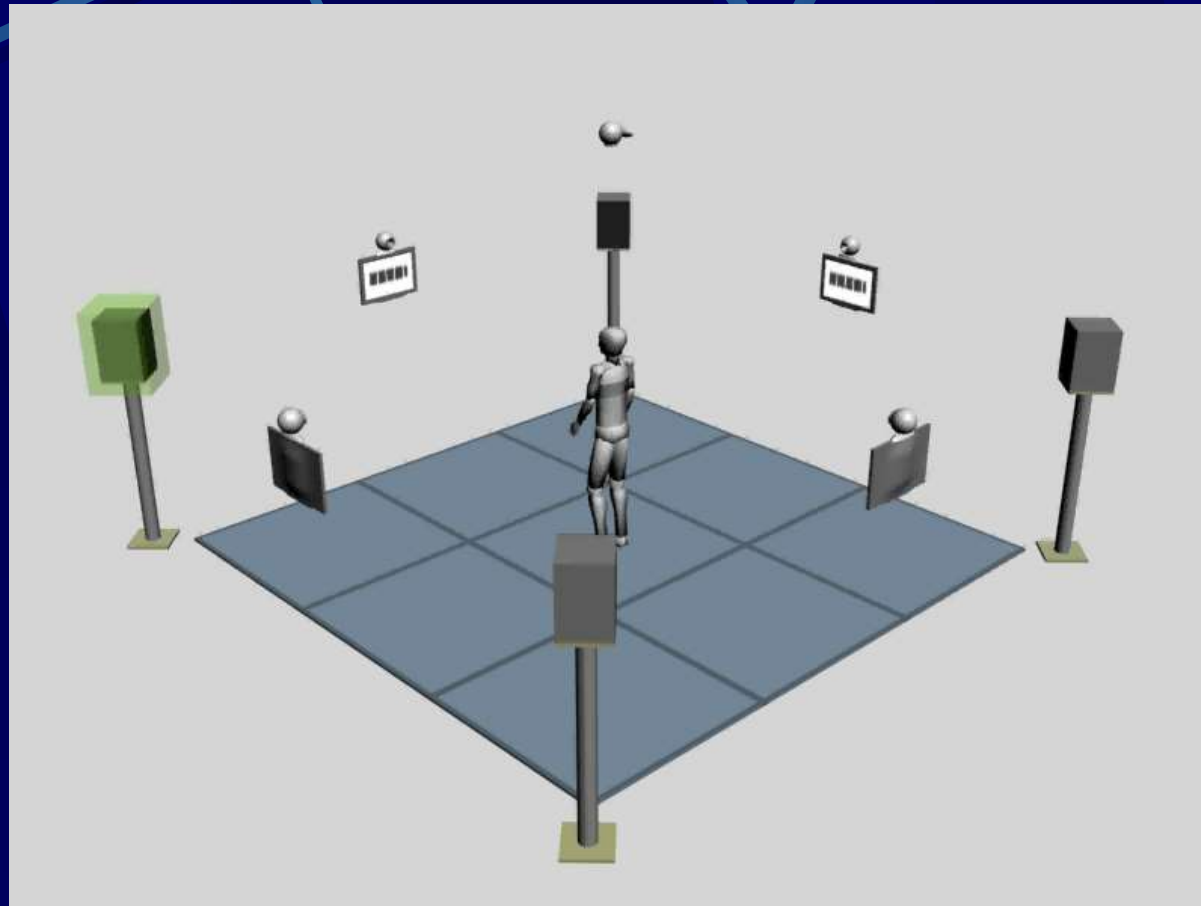




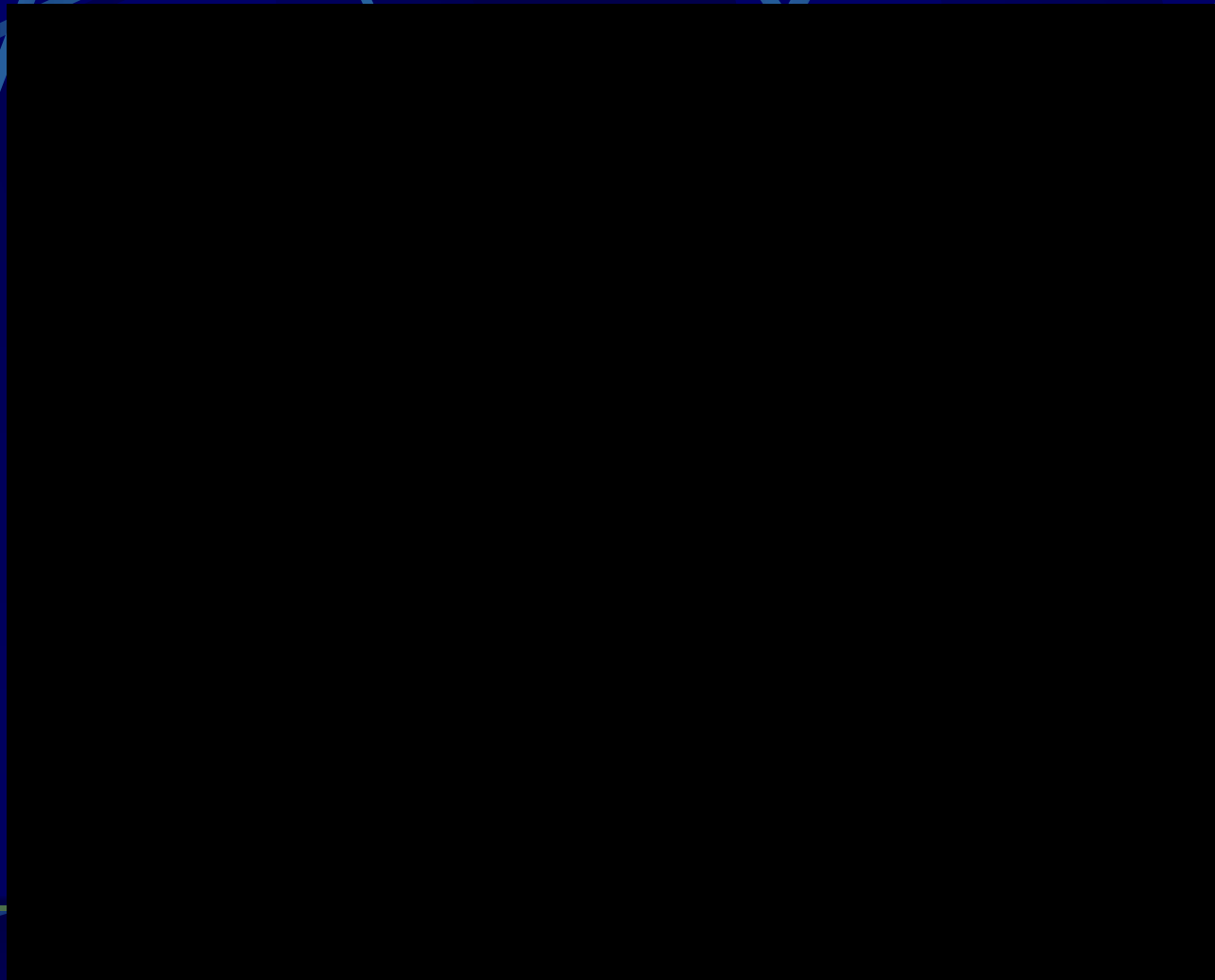
POLISENSORYCZNY SYSTEM STYMULACJI

- Polisensoryczny system stymulacji został opracowany, aby pomóc dzieciom z zaburzeniami koncentracji
- Połączenie sprzężenia zwrotnego **wizualnego i dźwiękowego** ze sterowaniem interfejsem za pomocą różnych ruchów ciała (np. skoków) zachęca do skupienia uwagi
- Zaprojektowano różne ćwiczenia

Wdrożenie multimodalne (i rozszerzenie)
metody Dirka J. Bakker: **Metronom interaktywny**



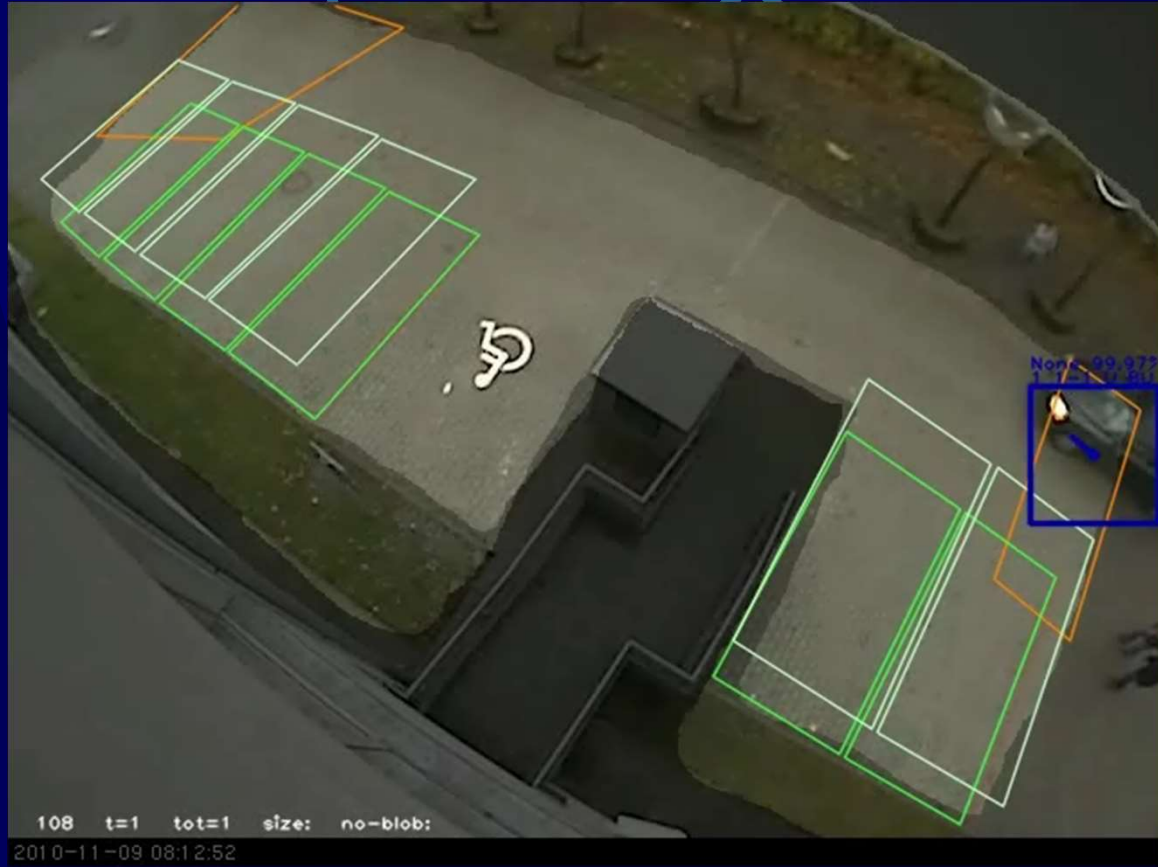
POLISENSORYCZNY SYSTEM STYMULACJI



Interfejs wizyjny



Chcemy połączyć technologię komputerową ze światem materialnym.







Dziękuję za uwagę

- www.multimed.org

- E-mail:

prof. Andrzej Czyżewski ac@pg.edu.pl