



AKADEMIA GÓRNICZO-HUTNICZA
IM. STANISŁAWA STASZICA W KRAKOWIE

System śledzenia oczu, twarzy i ruchów użytkownika komputera za pośrednictwem kamery internetowej i pozycjonujący oczy cyberagenta internetowego na oczach i akcjach użytkownika

Mirosław Słysz

Promotor: dr Adrian Horzyk

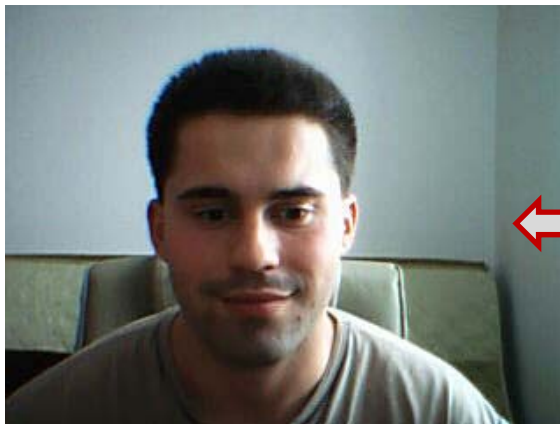


Cel pracy

- **System śledzenia twarzy użytkownika komputera za pomocą kamery internetowej**
 - szybki i efektywny algorytm lokalizacji twarzy
 - lokalizacja partii twarzy (oczu i ust)
 - rozpoznawanie mimiki
- **Animacja cyberagenta internetowego**
 - pozycjonowanie oczu na użytkowniku komputera
 - odpowiedź na uśmiech użytkownika komputera, uśmiechem cyberagenta
 - odpowiedź na mrugnięcie użytkownika komputera, mrugnięciem cyberagenta
- **Program zrealizowany jako aplikacja internetowa**

Założenia dla algorytmu lokalizacji twarzy

- Wydajny (konieczność lokalizacji twarzy w czasie rzeczywistym na obrazie video)
- Efektywny dla obrazów z kamery internetowej niskiej jakości, wykonanych w otoczeniu o nierównym oświetleniu



Obraz o nierównym oświetleniu, wykonany w pomieszczeniu, w którym po jednej ze stron komputera znajduje się okno

Algorytm lokalizacji twarzy

Opracowany algorytm lokalizacji twarzy oparty został o cechy, które można wskazać na obrazie w skali szarości oraz na obrazie binarnym

Przykład:

Obszar zawierający oczy, na obrazie w skali szarości posiada niższą wartość (jest ciemniejszy) od obszaru położonego pod oczami

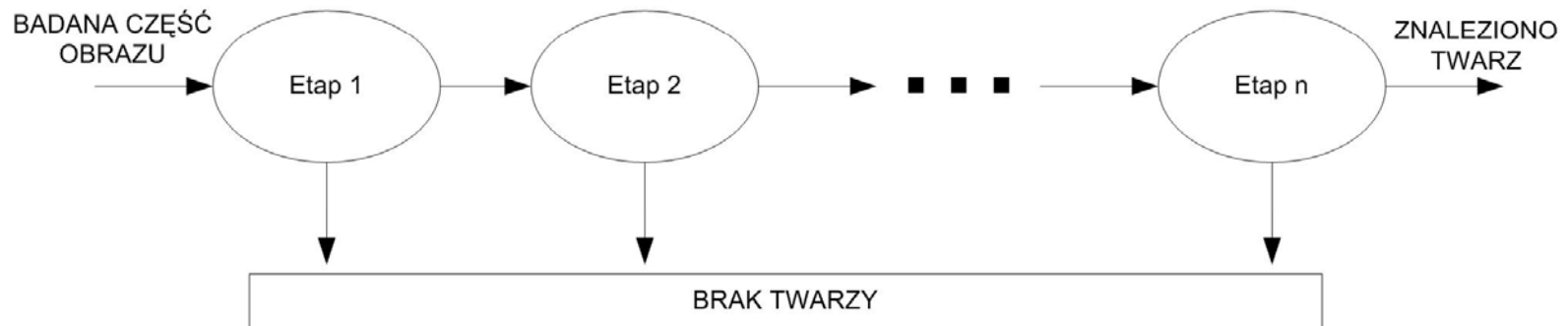


Algorytm lokalizacji twarzy c.d.

- Twarz poszukiwana jest na całej powierzchni badanego obrazu 320x240 pikseli – sprawdzając części (klatki) o rozmiarach od 50x50 do 180x180 pikseli z krokiem 5 pikseli
- Algorytm lokalizacji twarzy składa się z 25 etapów (klasyfikatorów), w których sprawdzane są empirycznie dobrane cechy

Algorytm lokalizacji twarzy c.d.

- Kaskadowa forma algorytmu zapewnia wysoką wydajność. Jeżeli w badanej części obrazu nie zostanie odnaleziona cecha jednego z etapów, to zostaje uznane, że dana część nie zawiera twarzy



Algorytm lokalizacji twarzy c.d.

Cechy poszukiwane są na obrazie w skali szarości oraz obrazie binarnym

Wartości piksela (x,y) obrazu w skali szarości wylicza się ze wzoru

$$\text{gray}(x,y) = 0.3 * R(x,y) + 0.59 * G(x,y) + 0.11 * B(x,y)$$

gdzie R, B, G – wartości składowe obrazu w formacie RGB

Wartość piksela (x,y) obrazu binarnego wylicza się na podstawie wartości danego piksela na obrazie skali szarości, porównując tę wartość z pewnym progiem (treshold)

$$\text{binary}(x,y) = \begin{cases} 1 & \text{gdy } \text{gray}(x,y) \leq \text{threshold}(x,y) \\ 0 & \text{w innym przypadku} \end{cases}$$

Algorytm lokalizacji twarzy c.d.

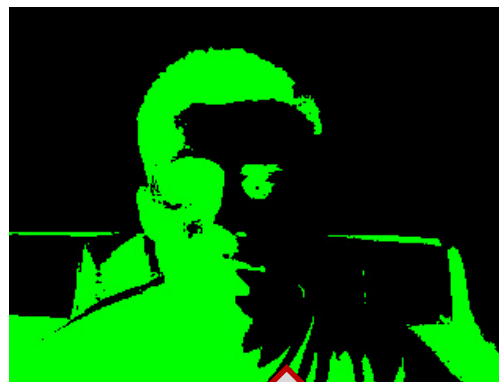
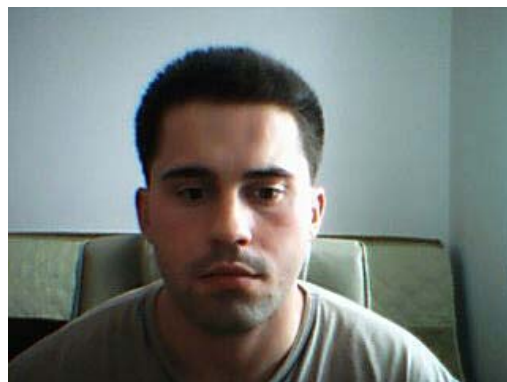
Po binaryzacji standardową metodą obrazu wykonanego w otoczeniu o równym oświetleniu wyraźnie zaznaczają się rysy twarzy. Można to wykorzystać do ustalenia wielu cech



Obraz po binaryzacji
standardową metodą

Algorytm lokalizacji twarzy c.d.

Po binaryzacji obrazu wykonanego w otoczeniu o nierównym oświetleniu rysy twarzy nie zostają wyłonione



Obraz o nierównym oświetleniu po binaryzacji standardową metodą

Algorytm lokalizacji twarzy c.d.

Pomimo nierównego oświetlenia pewne informacje zostają zachowane – wartość pikseli obszaru oczu na obrazie w skali szarości posiadają najniższą wartość w swym lokalnym otoczeniu.

Wykorzystując tę własności, opracowano algorytm binaryzacji ze zmiennym progiem, wyliczanym dla każdego piksela obrazu. Pozwala to wyłonić cechy twarzy również dla zdjęć wykonanych w pomieszczeniu o nierównym oświetleniu



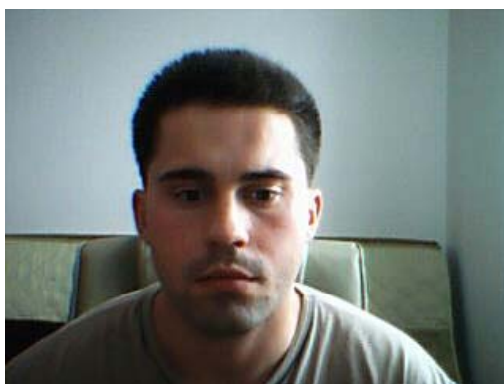
Algorytm lokalizacji twarzy c.d.

Wartość pikseli obrazu binarnego wyliczana jest ze wzoru:

$$\text{binary}(x,y) = \begin{cases} 1 & \text{gdy } \text{gray}(x,y) \leq 0.95 * \text{srednia}(x,y) \\ 0 & \text{w innym przypadku} \end{cases}$$

$\text{srednia}(x,y)$ – średnia wartość powierzchni obrazu w skali szarości, o rozmiarze 7x7 pikseli, o środku w punkcie (x,y)

Algorytm lokalizacji twarzy c.d.



↑

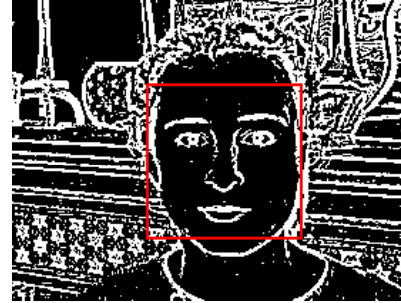
Obraz o nierównym oświetleniu
po binaryzacji metodą ze
zmiennym progiem

Etapy algorytmu lokalizacji twarzy

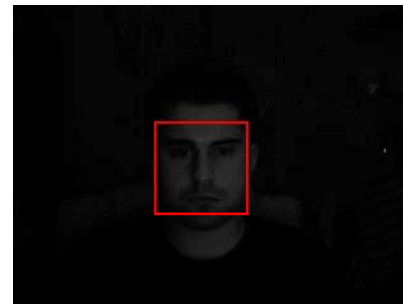
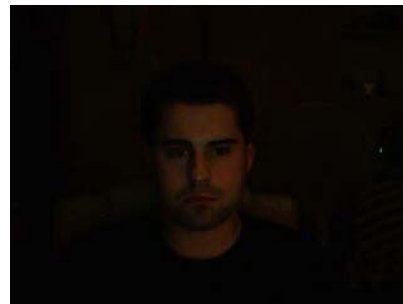
- Na poszczególnych etapach algorytmu lokalizacji twarzy wyliczane są klasyfikatory - wartości obszarów o określonych rozmiarach i położeniu
- Suma klasyfikatorów wyznacza „wartość” klatki – na jej podstawie wybierana jest najlepsza klatka w przypadku, gdy dla jednej twarzy znaleziono kilka klatek
- Klasyfikatory dzielą się na:
 - klasyfikatory mocne – ich wartość musi spełniać pewną zależność, w przeciwnym wypadku klatka zostaje odrzucona
 - klasyfikatory słabe – ich wartość służy do sumarycznej oceny badanej klatki

Etapy algorytmu lokalizacji twarzy c.d.

- Analiza wartości badanej klatki na obrazie binarnym

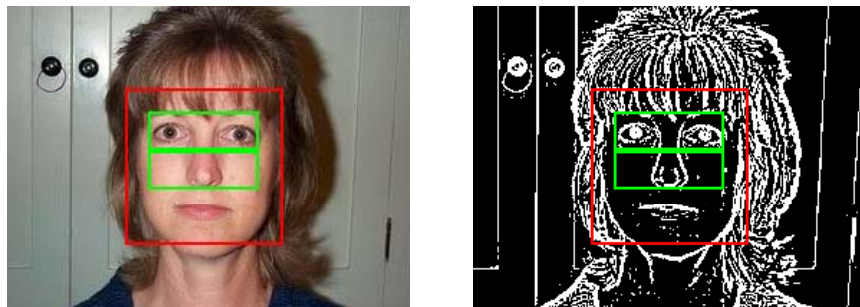


- Analiza wartości badanej klatki na obrazie binarnym

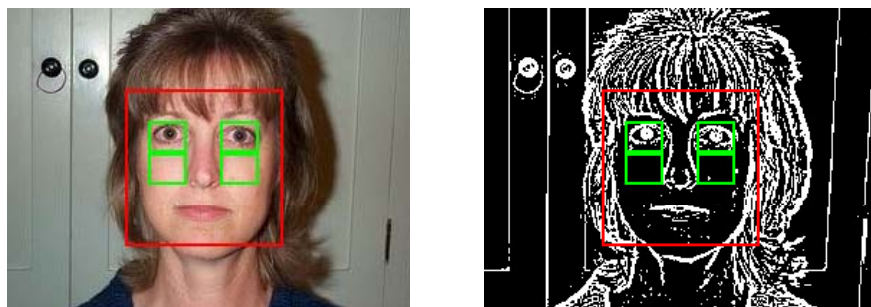


Etapy algorytmu lokalizacji twarzy c.d.

- Analiza dużego obszaru oczu oraz pod oczami na obrazie binarnym

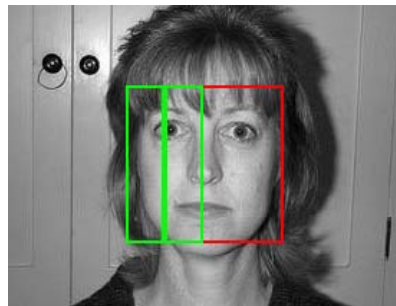


- Analiza małych obszarów oczu oraz pod oczami na obrazie binarnym



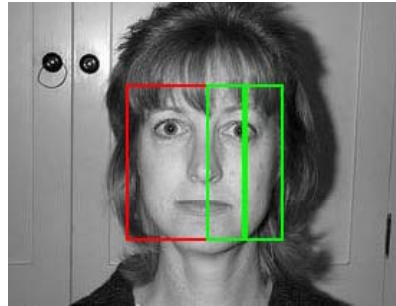
Etapy algorytmu lokalizacji twarzy c.d.

- Analiza małych obszarów oczu oraz pod oczami na obrazie w skali szarości
- Analiza lewych ćwiartek klatki na obrazie w skali szarości

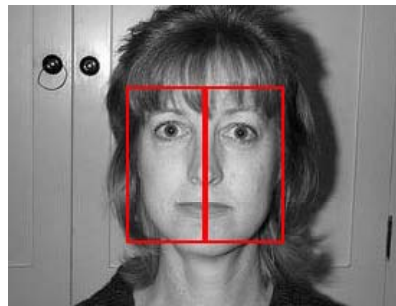


Etapy algorytmu lokalizacji twarzy c.d.

- Analiza prawych ćwiartek klatki na obrazie w skali szarości

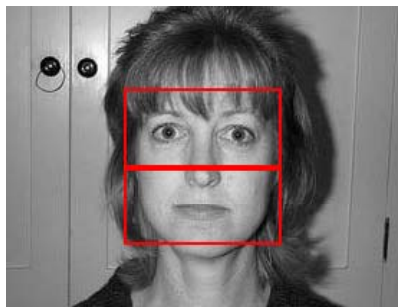


- Analiza pionowych połówek klatki na obrazie w skali szarości

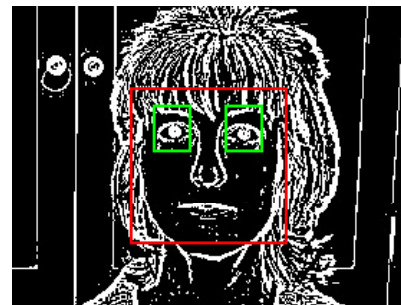
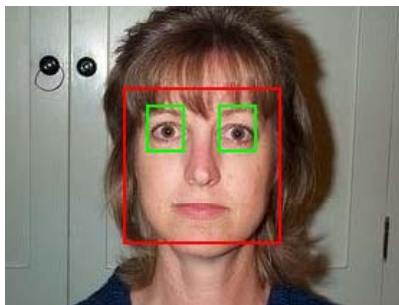


Etapy algorytmu lokalizacji twarzy c.d.

- Analiza poziomych połówek klatki na obrazie w skali szarości

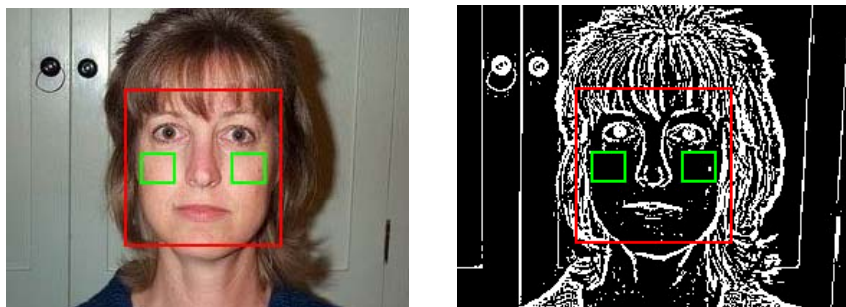


- Analiza wartości obszarów oczu i brwi na obrazie binarnym

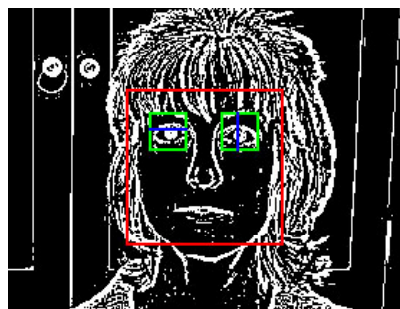


Etapy algorytmu lokalizacji twarzy c.d.

- Analiza obszarów policzków na obrazie binarnym

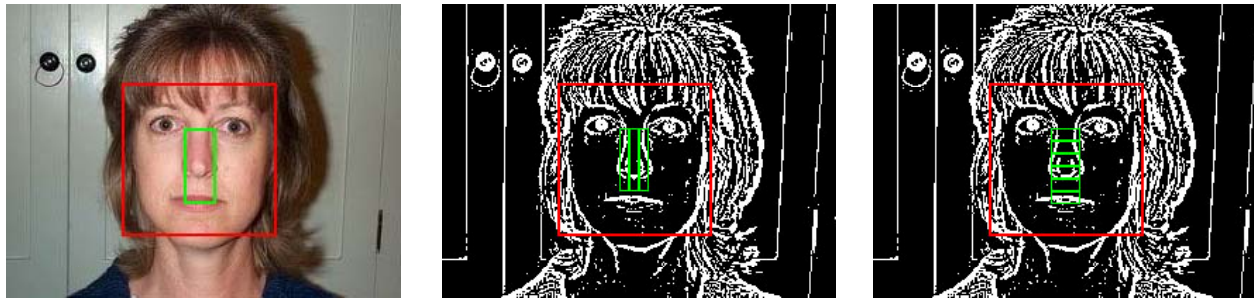


- Analiza regionów oczu na obrazie binarnym

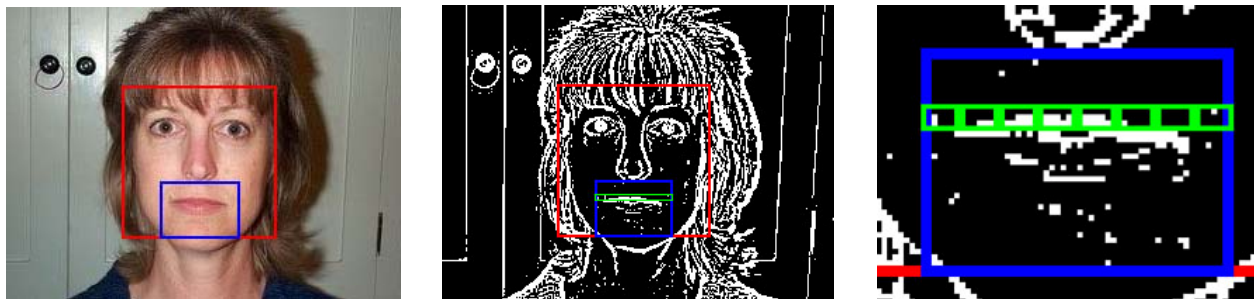


Etapy algorytmu lokalizacji twarzy c.d.

- Analiza regionu nosa na obrazie binarnym

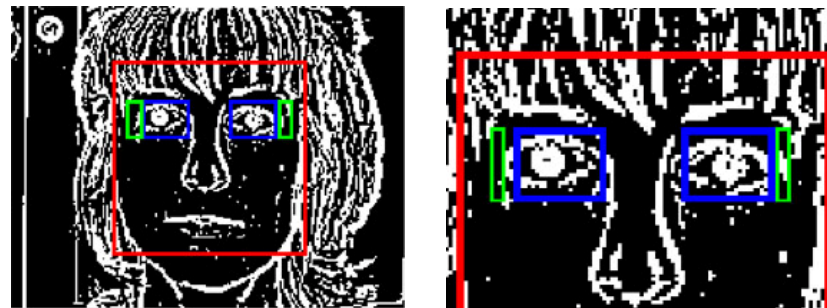


- Analiza obszaru ust na obrazie binarnym



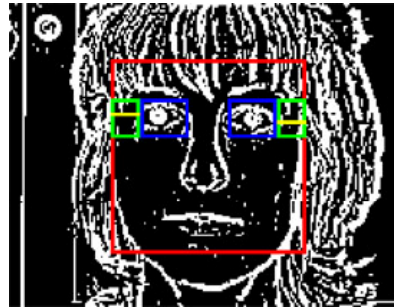
Etapy algorytmu lokalizacji twarzy c.d.

- Analiza różnicy wartości obszaru ust i policzków na obrazie binarnym
- Analiza zewnętrznych części obszarów oczu - w poziomie na obrazie binarnym

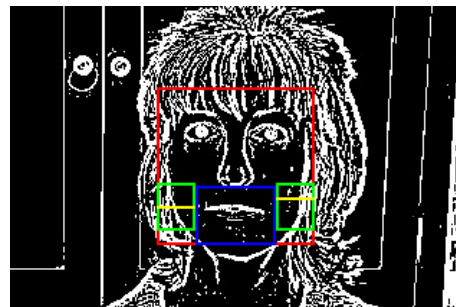


Etapy algorytmu lokalizacji twarzy c.d.

- Analiza zewnętrznych części obszarów oczu w pionie na obrazie binarnym

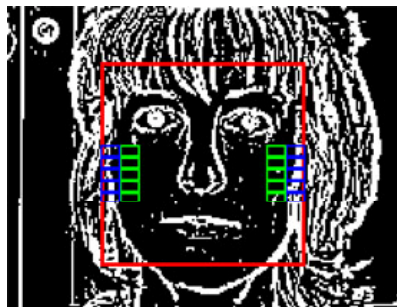


- Analiza zewnętrznych części obszaru ust w pionie na obrazie binarnym

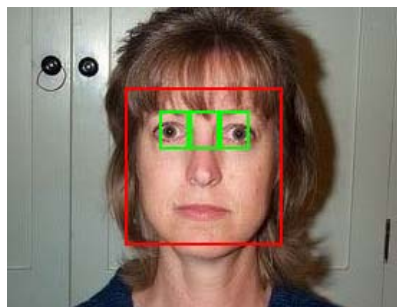


Etapy algorytmu lokalizacji twarzy c.d.

- Analiza bocznych krawędzi ramki na obrazie binarnym

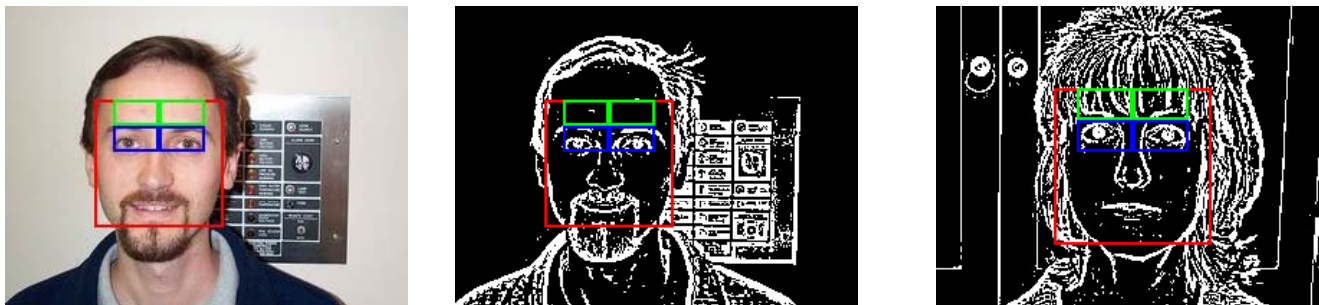


- Analiza obszaru oczu i pomiędzy oczami na obrazie binarnym

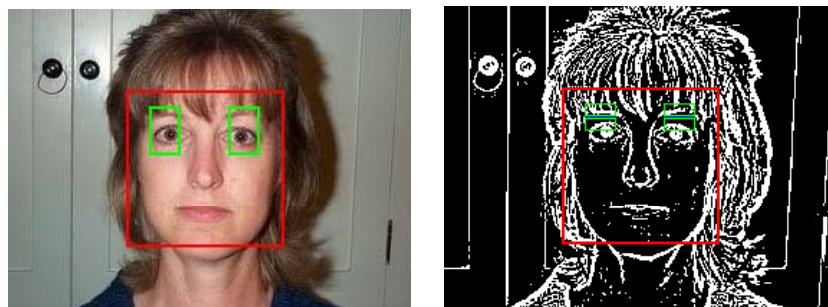


Etapy algorytmu lokalizacji twarzy c.d.

- Analiza obszaru oczu oraz czoła na obrazie binarnym



- Analiza obszaru oczu i brwi na obrazie binarnym





Lokalizacja twarzy w czasie rzeczywistym

Dla zwiększenia wydajności algorytmu wprowadzono wykrywanie ruchu

- W celu lokalizacji twarzy przeszukiwane są jedynie obszary, w których wystąpił ruch
- Jeżeli w otoczeniu ramki ze zlokalizowaną twarzą nie wystąpił ruch, twarz zostaje zachowana w poprzednim położeniu

Lokalizacja twarzy w czasie rzeczywistym c.d.

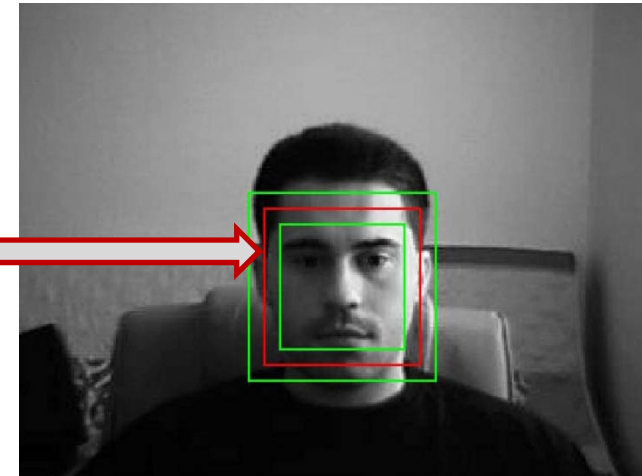
1) Wyznaczenie binarnego obrazu ruchu



Lokalizacja twarzy w czasie rzeczywistym c.d.

2) Wyliczenie wartości otoczenia ramki ze zlokalizowaną twarzą na binarnym obrazie ruchu

- Kolor czerwony – wykryta ramka z twarzą
- Wartość obszaru pomiędzy zielonymi ramkami na binarnym obrazie ruchu określa czy twarz uległa przemieszczeniu



3) Jeżeli wystąpił ruch – ponownie zlokalizuj twarz

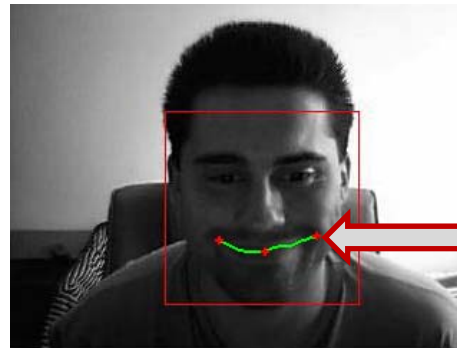
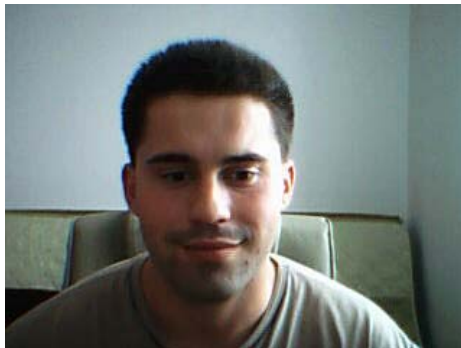
Lokalizacja ust

- Na jednym z etapów algorytmu lokalizacji twarzy wyznaczane jest przybliżone położenie ust.
- W pobliżu wyznaczonej pozycji poszukiwany jest zespół trzech prostokątów, dla których suma stosunków prostokąta środkowego do górnego oraz środkowego do dolnego posiada najmniejszą wartość.
- Środkowy prostokąt z trzech odnalezionych wyznacza pozycję przestrzeni pomiędzy ustami.



Określanie kształtu ust

- Wyznaczenie środka ust
 - współrzędna x – środek ramki zawierającej twarz
 - współrzędna y – wyznaczona przez pozycję środkowego prostokąta z trzech odnalezionych
- Śledzenie najciemniejszych sąsiadujących pikseli po lewej i prawej stronie od środka ust

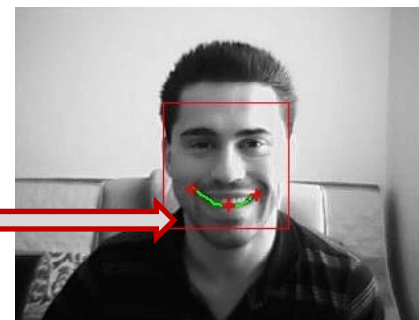


Wykryty kształt ust

Wykrywanie uśmiechu

- Wystąpienie uśmiechu ustalane w zależności od kształtu ust

Krańcowe punkty kształtu ust powyżej środka ust – wykryto uśmiech



- Jeżeli kształt ust odpowiada uśmiechowi i na obrazie zlokalizowany zostanie obszar zawierający zęby to uśmiech zostaje uznany za otwarty

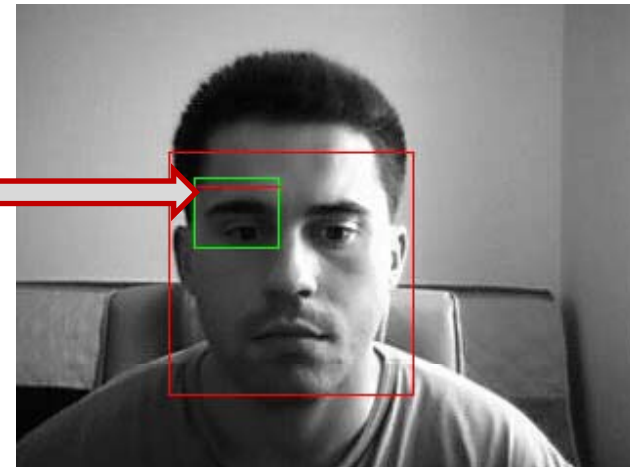
Obszar nad zlokalizowaną pozycją środka ust posiada wysoką wartość na obrazie w skali szarości – na obrazie są widoczne zęby



Lokalizacja oczu

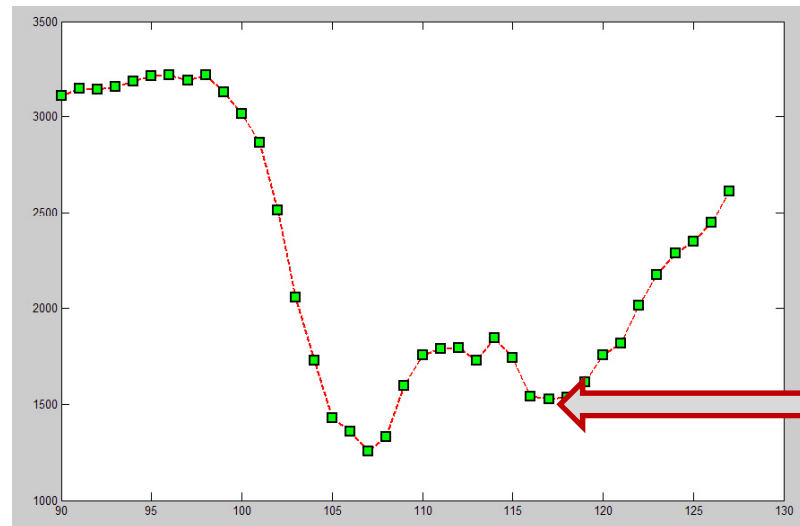
1) Wyliczenie wartości prostokątów o wysokości 1, leżących wzdłuż obszaru zawierającego oczy

- Kolor zielony – przeszukiwany obszar
- Kolor czerwony – przykład prostokąta, dla którego obliczana jest wartość



Lokalizacja oczu c.d.

2) Analiza wyliczonych wartości – drugie minimum lokalne od góry badanego obszaru wyznacza górną krawędź oka



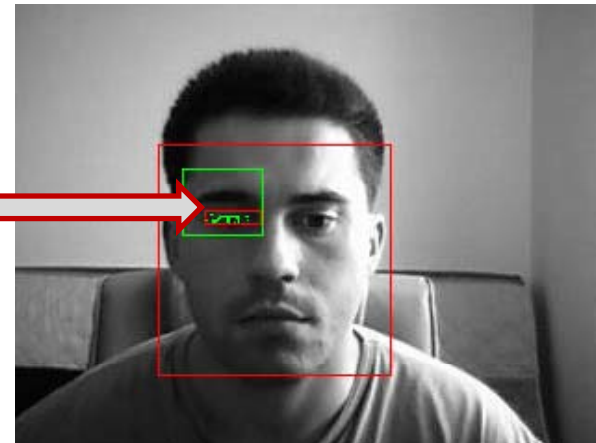
Wartość górnej krawędzi oka

Lokalizacja oczu c.d.

- 3) Binaryzacja obszaru oka
- 4) Wyznaczenie obszaru o największej ilości pikseli o wartości 1 na obrazie binarnym

Kolor zielony – piksele o wartości 1 na obrazie binarnym

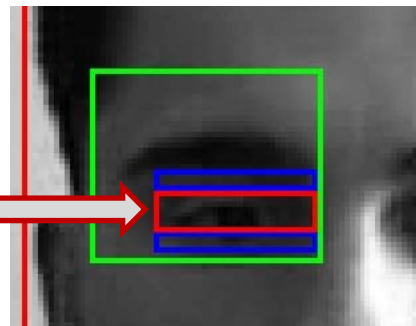
Kolor czerwony – ramka z wykrytym okiem



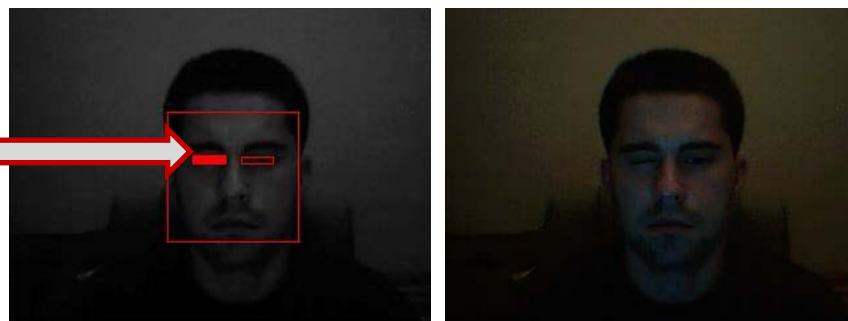
Określanie stanu oczu

Stan oka (zamknięte/otwarte) wyznaczone jest na podstawie stosunku wartości powierzchni wykrytego oka oraz powierzchni nad o pod nią

Powierzchnie, na podstawie których wartości określany jest stan oka



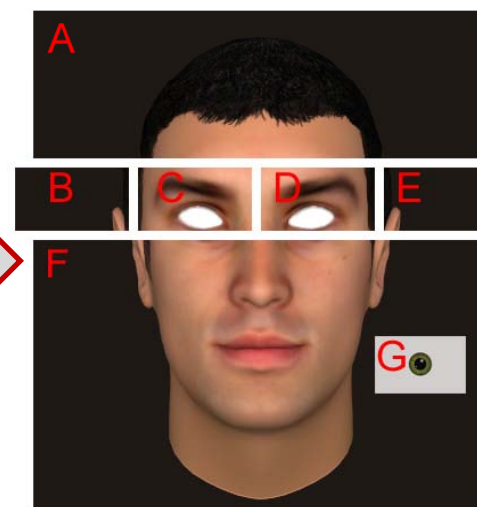
Stan oka ustalony jako zamknięte



Animacja postaci

- Zrealizowana jako seria szybko zmienianych klatek obrazu
- W celu ograniczenia ilości wykorzystywanych obrazów, twarz cyberagenta podzielona została na części

A, B, E – niezmiennie obszary
C i D – animacja oczu
F – animacja uśmiechu
G – animacja ruchu źrenicy oka



Zrealizowane animacje

- Ruch źrenicy oka – przemieszczanie bitmapy ze źrenicą
- Mrugnięcie zwykłe – realizowane w losowo wybranym czasie (symuluje naturalny odruch mrugania)

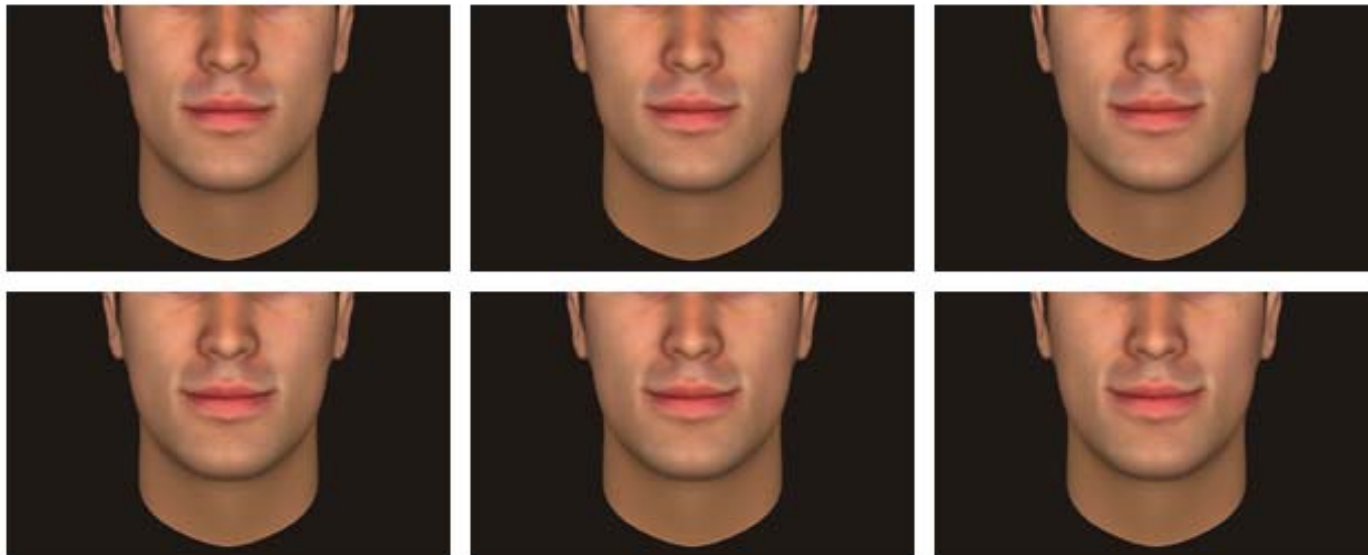


- Mrugnięcie specjalne tzw „oczko”



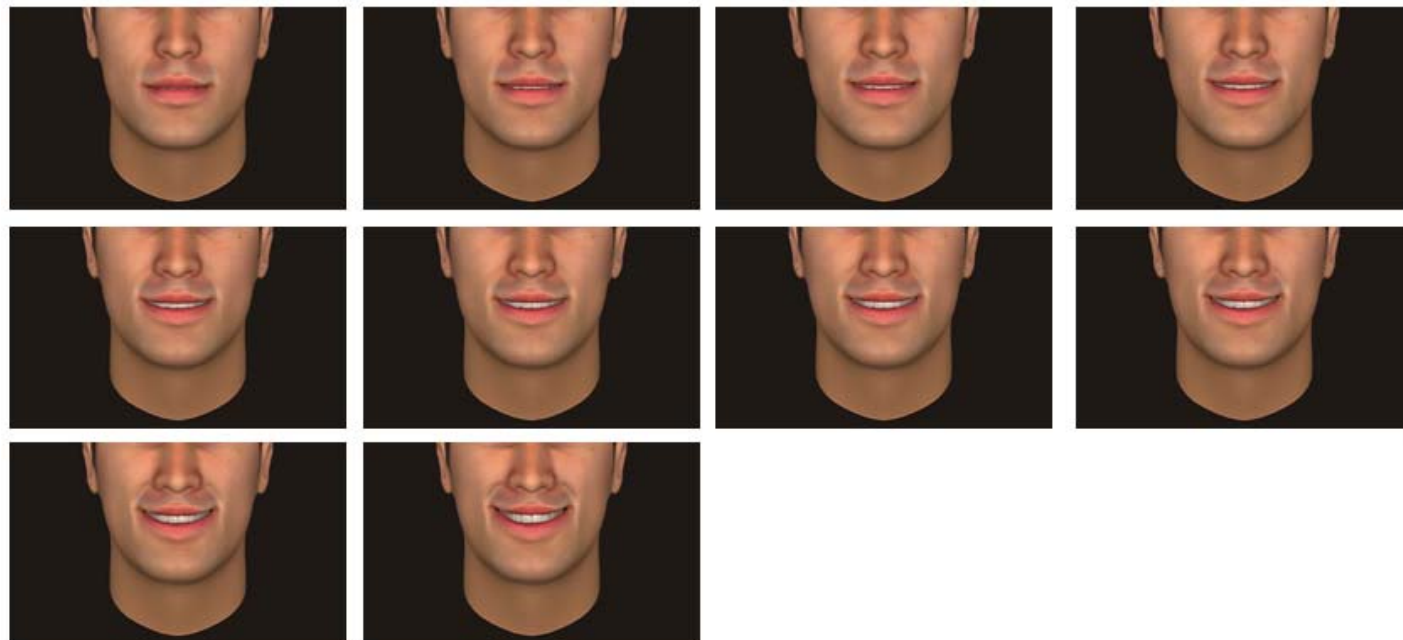
Zrealizowane animacje c.d.

- Uśmiech zamknięty



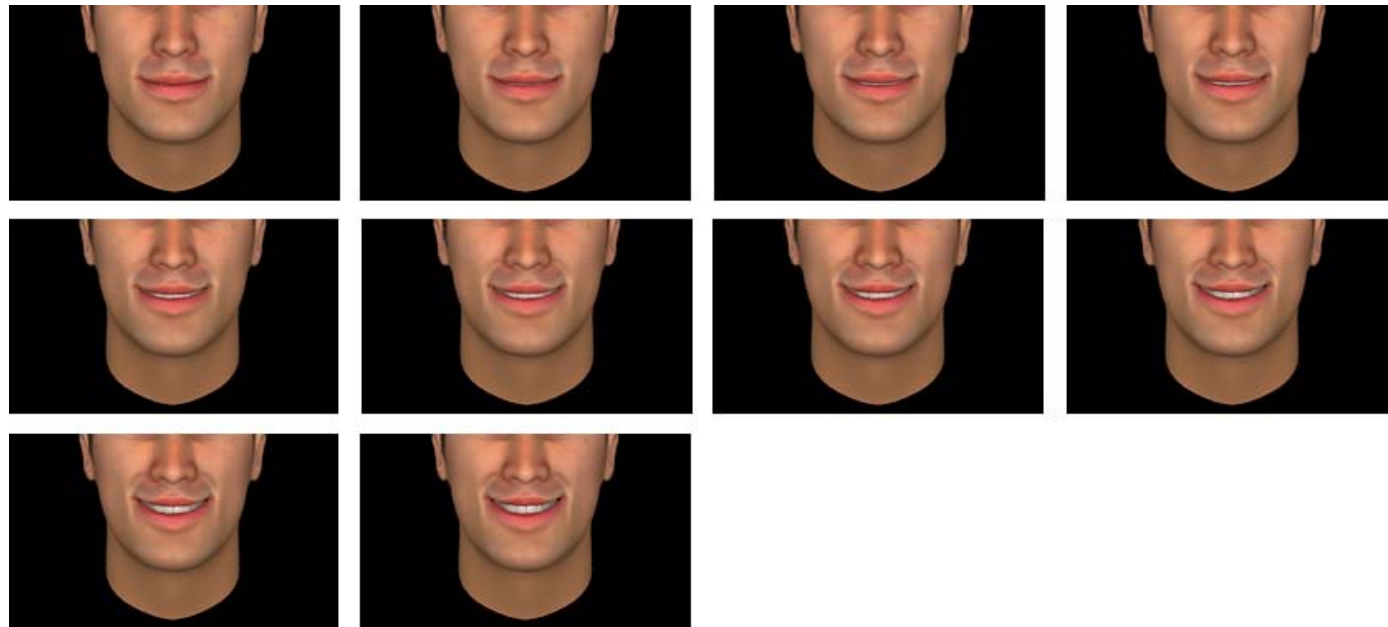
Zrealizowane animacje c.d.

- Uśmiech otwarty



Zrealizowane animacje c.d.

- Przejście pomiędzy uśmiechem otwartym a zamkniętym



Aplikacja

- Zrealizowana w języku Java jako Applet
- Dodatkowo przygotowana strona internetowa



Testy aplikacji

- Skuteczna i szybka lokalizacja twarzy
- Szybka odpowiedź cyberagenta na ruch oraz mimikę użytkownika komputera





Zrealizowane cele

- Skuteczny i wydajny algorytm lokalizacji twarzy, efektywny w złych warunkach oświetlenia
- Skuteczna lokalizacja i określanie stanu oczu oraz ust
- Płynna animacja postaci cyberagenta – zachowane pewne naturalne odruchy ludzkie



Koniec

Dziękuję za uwagę