



Technical University of Cluj - Napoca
Computer Science Department

Interactiune om-calculator

Curs 3

Aplicatii ale segmentarii bazate pe culoare: detectia mainii si recunoasterea gesturilor



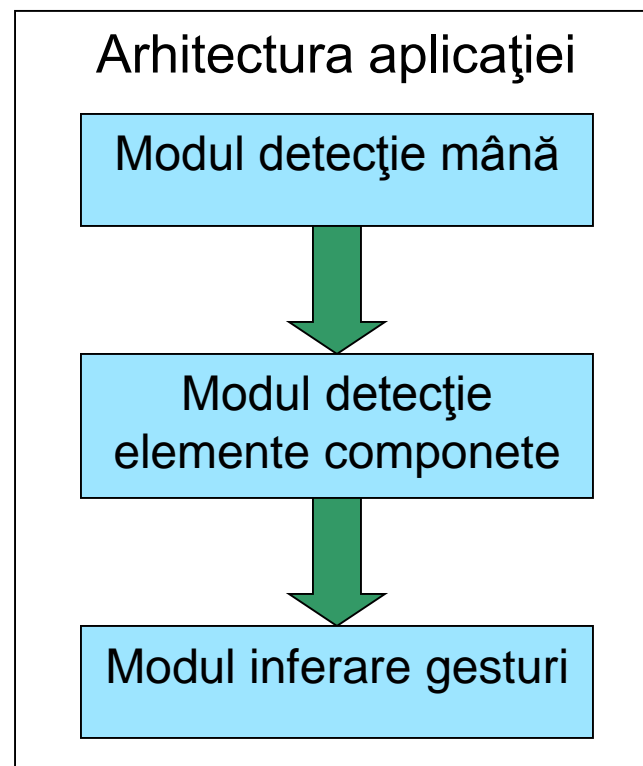
Detecția prezenței mâinii în imagini statice și identificare a degetelor [1]

Posibilități de detecție:

- direcția mâinii, (direcția indicată de mână)
- numărul de degete distincte extinse etc).

Sistem de achiziție:

- cameră web comercială (Low-Cost) care achiziționează imagini RGB la rezoluția de 320x240 pixeli.





Detecția prezenței mâinii în imagini statice și identificare a degetelor

1. Se aplica o procedura de antrenare pentru determinarea distribuției componentelor H și S pt. mana \Rightarrow culori pt. “seed-points”
2. Se aplică o segmentare bazată pe regiuni: algoritm region growing pe componentele H și S doar pt. seed-pints.
3. Pe imaginea rezultată se aplica un algoritm de etichetare, se calculează aria obiectelor etichetate și se elimină obiectele cu aria mică.

Modul detecție mână

RGB \Rightarrow HSI

Filtrare H și S

Segmentare bazată pe regiuni

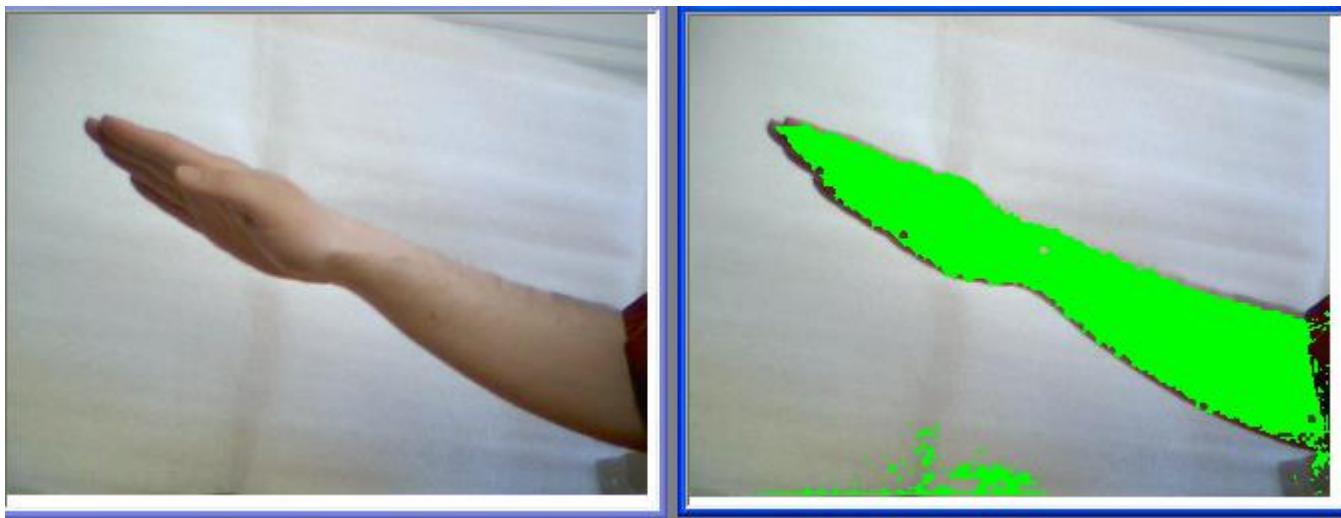
Etichetare obiecte

Eliminare obiecte cu arie mică

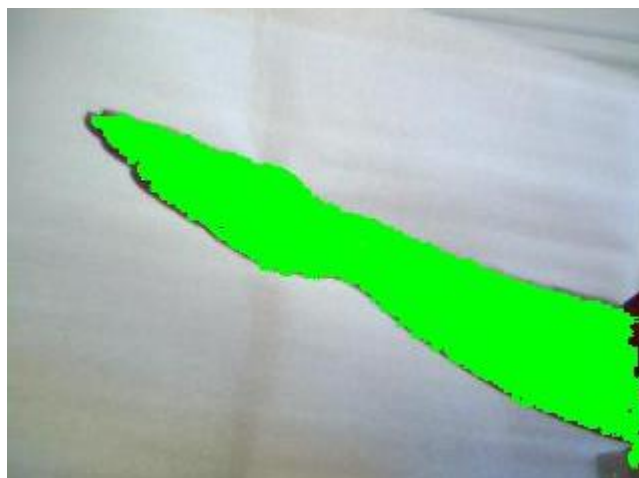


Detecția prezenței mâinii în imagini statice și identificare a degetelor

Rezultate



Pas 2



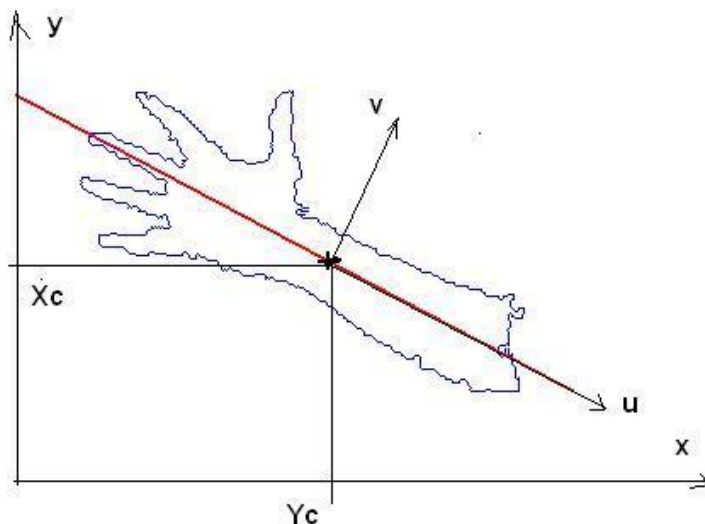
Pas 3



Detecția prezenței mâinii în imagini statice și identificare a degetelor

Detecție CM și axa de alungire

Translația și rotirea sistemelor de coordonate: $(x, y) \Rightarrow (u, v)$



$$R = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}; \text{ unde } \theta \text{ reprezintă unghiul de la axa de alungire.}$$

$$T = \begin{bmatrix} -c_i \\ -r_i \end{bmatrix}; \text{ unde } c_i \text{ și } r_i \text{ sunt coordonatele centrului de masă.}$$

$$\begin{bmatrix} u \\ v \end{bmatrix} = R \left(T + \begin{bmatrix} x \\ y \end{bmatrix} \right); \text{ unde } u \text{ și } v \text{ reprezintă noile valori ale coordonatelor.}$$

Modul detecție elemente componente

Detecție proprietăți
geometrice

Detecție contur

Translația sistemului
de coordonate

Decompoziție
palmă și braț

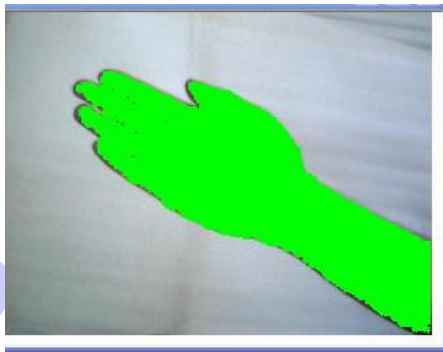
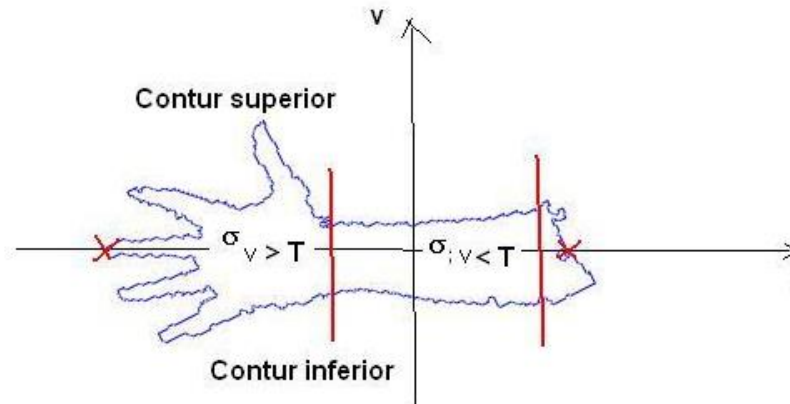
Detecție degete



Detecția prezenței mâinii în imagini statice și identificare a degetelor

Separare palmă / antebrăț

- Determinarea antebratului (delimitat de segmente de contur cvasiliniare)
- Găsirea limitei de separare dintre palmă și antebrăț
- Separare palmă

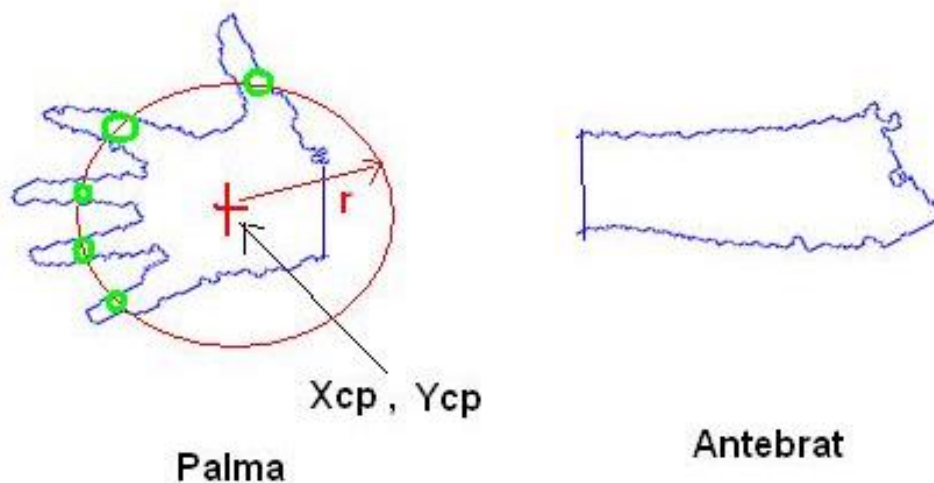




Detecția prezenței mâinii în imagini statice și identificare a degetelor

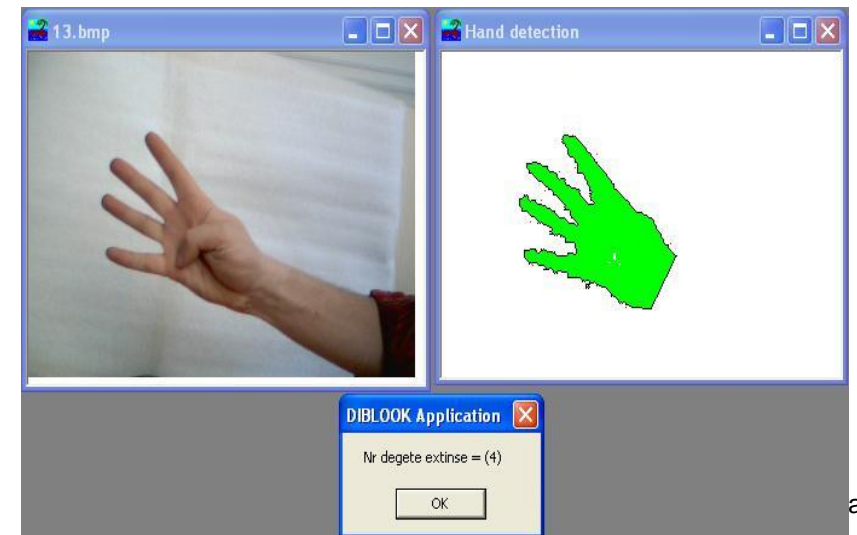
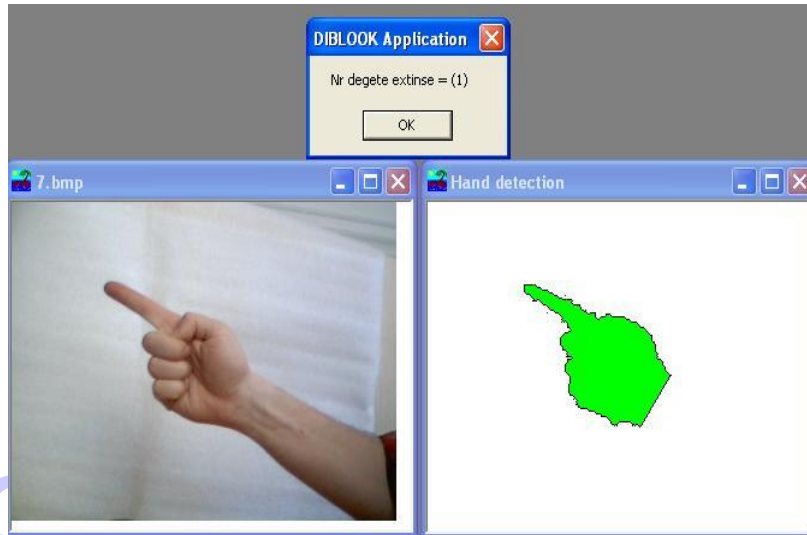
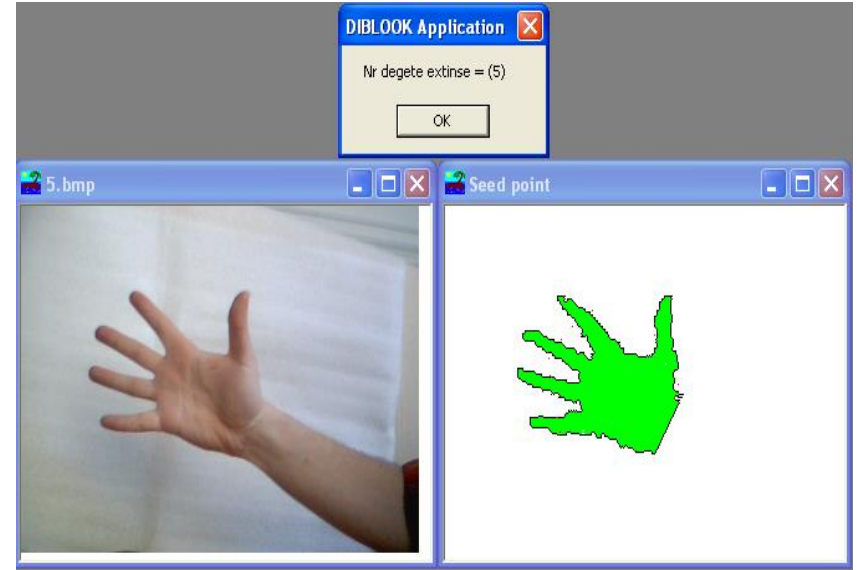
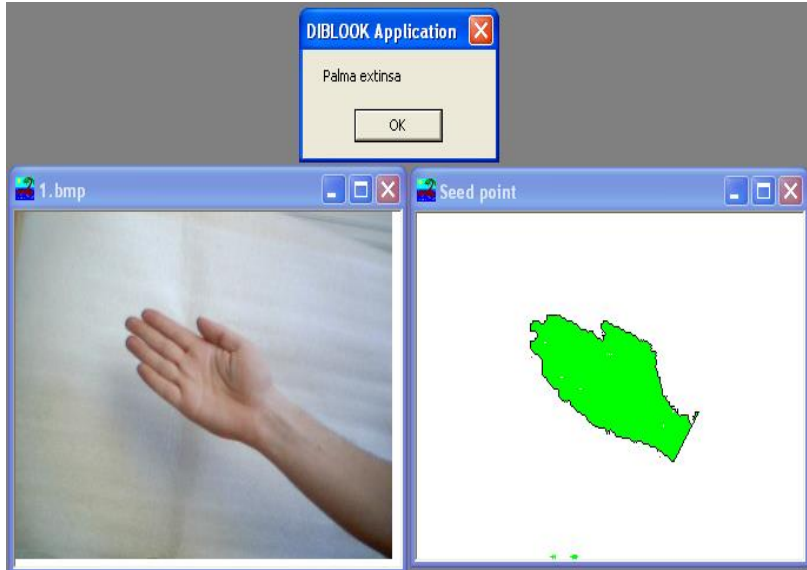
Detecție degete (extinse)

- Recalculare centru de masă și axa de alungire.
- Algoritm de detecție a degetelor extinse
 - se traseaza cercuri cu raze de diferite valori
 - se calculeaza numarul maxim de intersectii cu conturul
 - se infereaza numarul de degetee extinse





Detecția prezenței mâinii în imagini statice și identificare a degetelor



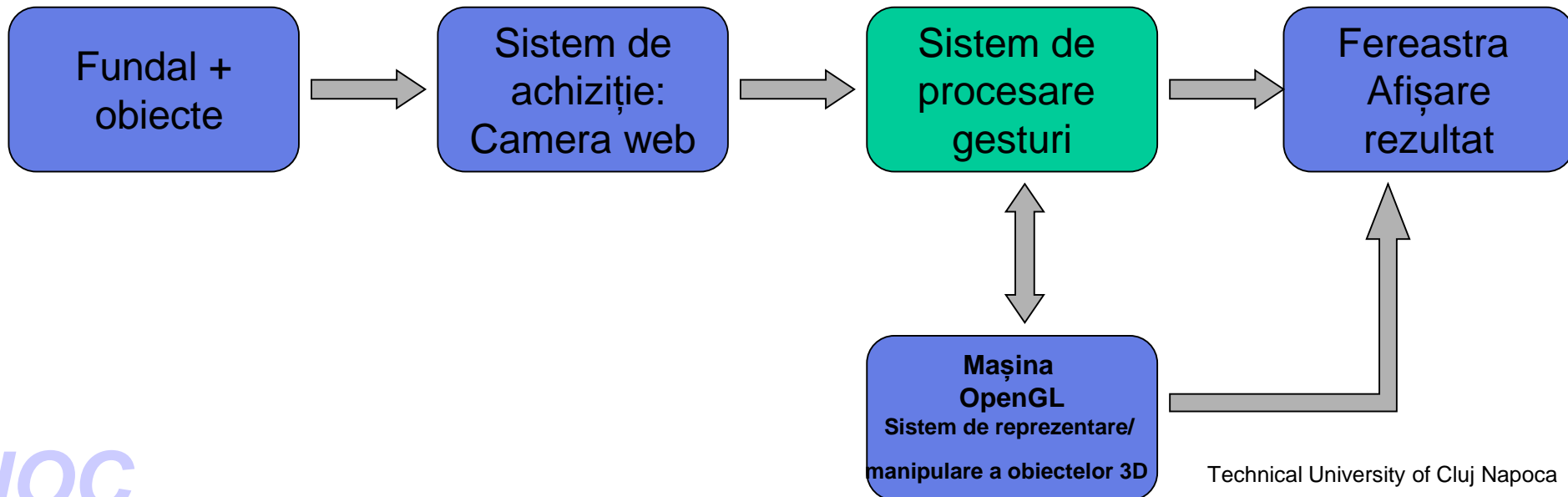


Detectarea si interpretarea gesturilor mâinii din achiziția video in timp real [2]

Specificatii

- Functionarea in timp real (imagini achizitionate de la o camera web)
- Segmentarea mainii si a degetelor
- Interpretarea unui numar limitat de gesturi

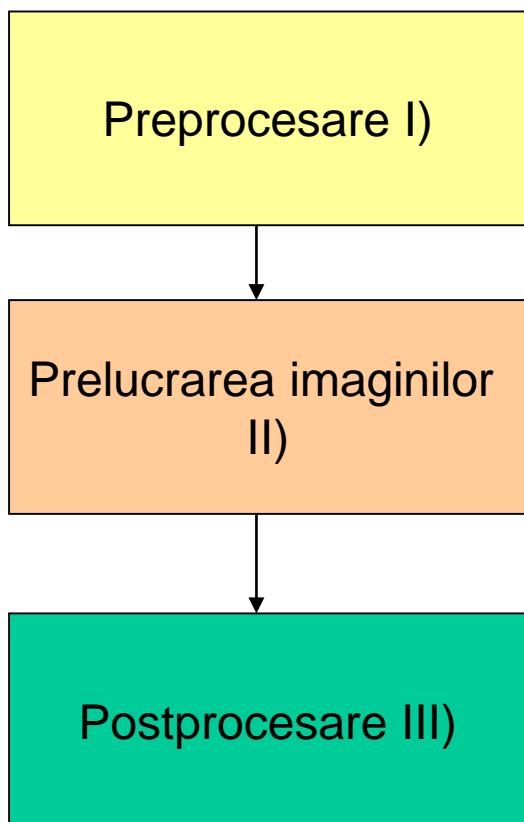
Arhitectura aplicației



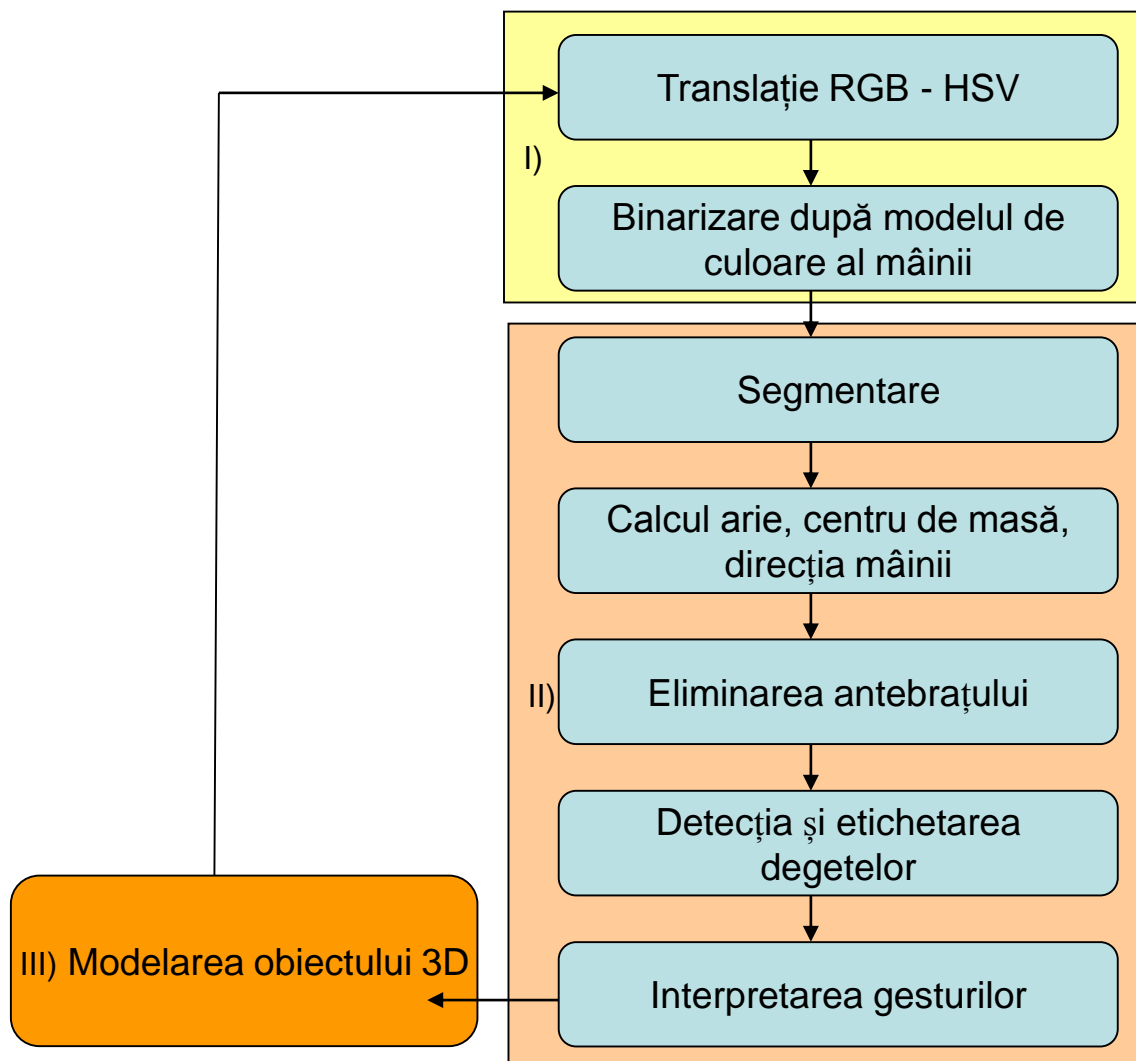


Detectarea și interpretarea gesturilor mâinii din achiziția video în timp real

Schema conceptuală



Schema conceptuală detaliată

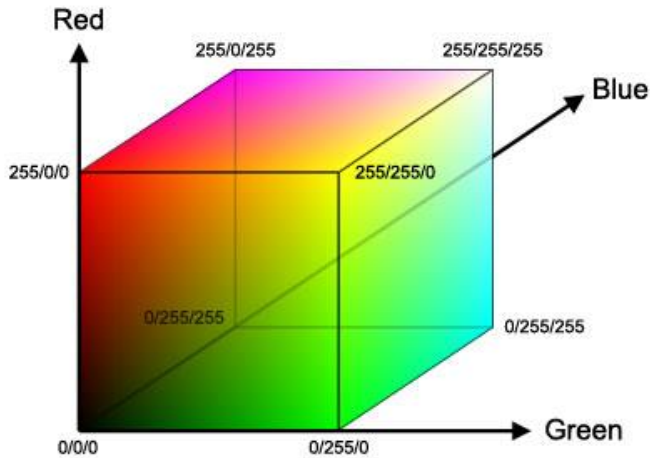




Detectarea și interpretarea gesturilor mâinii din achiziția video în timp real

Preprocesare

Translația din spațiul RGB în spațiul HSV (Hue, Saturation, Value)



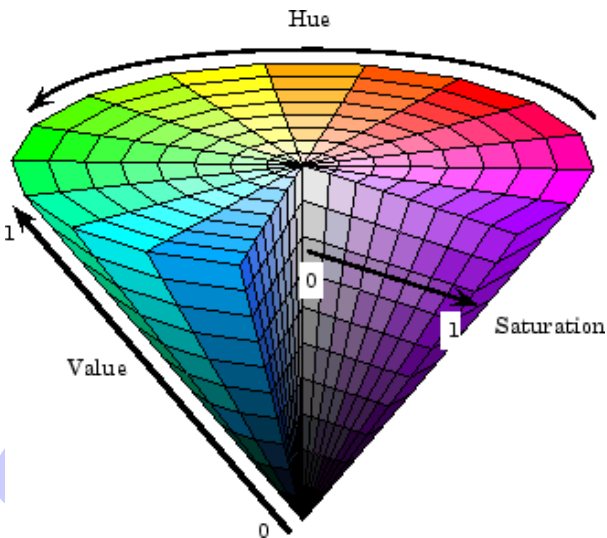
Conversia între modelele de culoare se face utilizând formulele (vezi C2):

$$V = \frac{1}{3}(R + G + B)$$

$$S = 1 - \frac{3}{(R + G + B)}[\min(R, G, B)]$$

$$H = \begin{cases} \theta & \text{daca : } G \geq B \\ 2\pi - \theta & \text{daca : } G \leq B \end{cases}$$

$$\theta = \arccos\left(\frac{\frac{1}{2}[(R - G) + (R - B)]}{\sqrt{[(R - G)(R - G) + (R - B)(G - B)]}}\right)$$

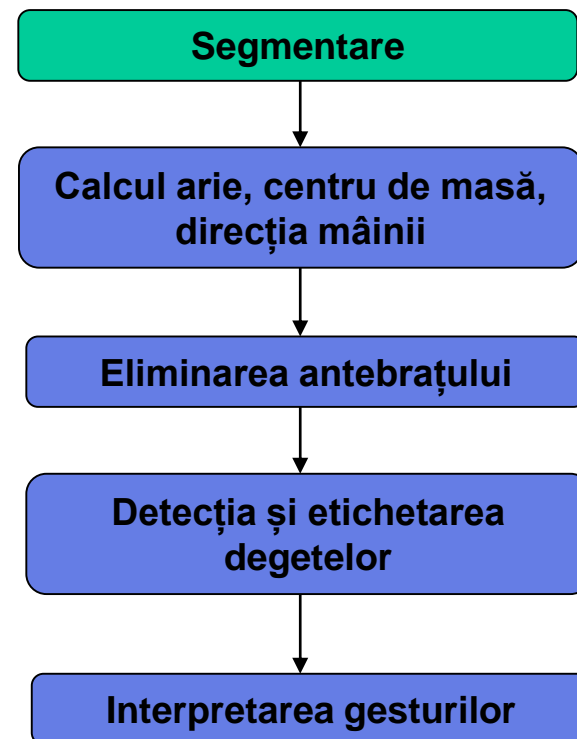
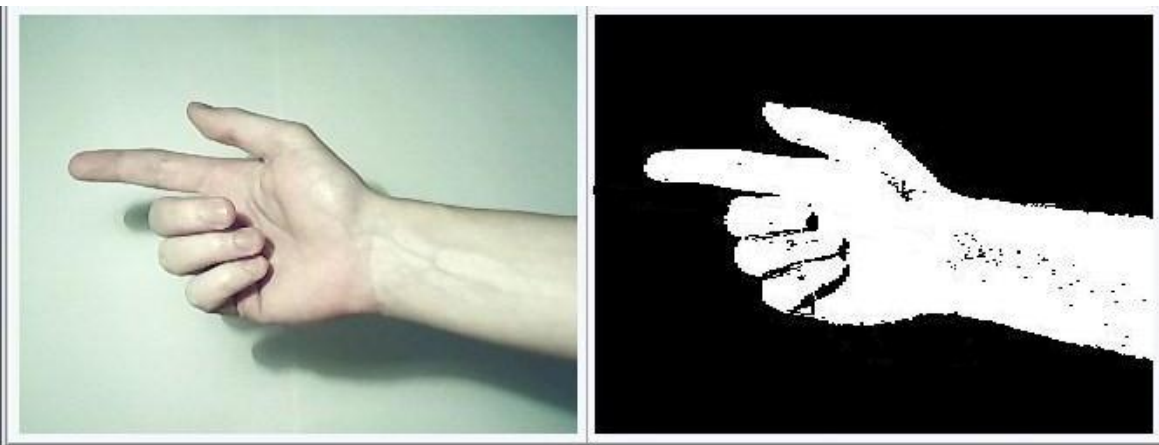




Detectarea și interpretarea gesturilor mâinii din achiziția video în timp real

Segmentare

Postprocesari pt. eliminarea zgomotelor (operatii morfologice + clustering în spațial imaginii)





Detectarea și interpretarea gesturilor mâinii din achiziția video în timp real

Calculul ariei, centrului de masă și a direcției mâinii

-Aria se calculează ca suma pixelilor obiectului .

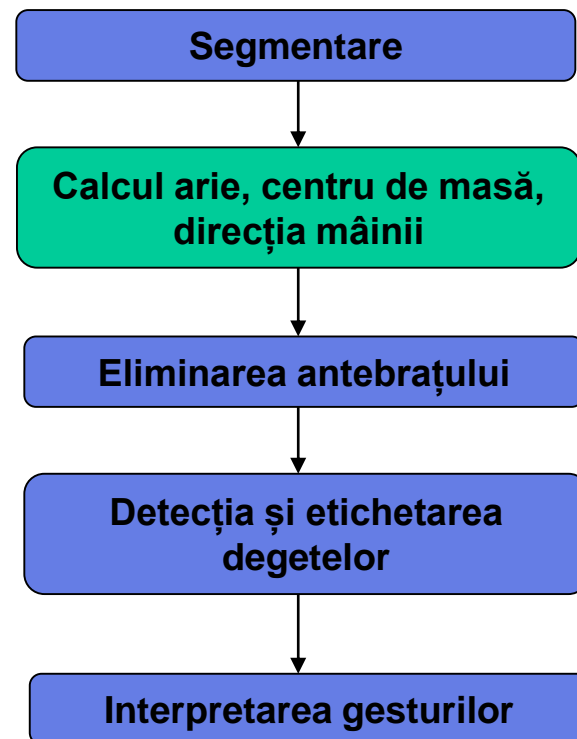
-Pentru a calcula coordonatele **centrului de masă** se calculează mai întâi momentul de-a lungul axelor Ox, și Oy .

$$\bar{i} \cdot \sum_{i=1}^n \sum_{j=1}^m b(i, j) = \sum_{i=1}^n \sum_{j=1}^m ib(i, j) \quad \bar{j} \cdot \sum_{i=1}^n \sum_{j=1}^m b(i, j) = \sum_{i=1}^n \sum_{j=1}^m jb(i, j)$$

$$\bar{i} = \frac{\sum_{i=1}^n \sum_{j=1}^m ib(i, j)}{A} \quad , \quad \bar{j} = \frac{\sum_{i=1}^n \sum_{j=1}^m jb(i, j)}{A}$$

-**Axa de alungire** o reprezintă direcția cu cel mai mic moment de ordin 2 (axa de inerție minimă)

$$\tan(2\phi_i) = 2 \frac{\sum_{r=0}^{N-1} \sum_{c=0}^{N-1} (r - \bar{r}_i)(c - \bar{c}_i) I_i(r, c)}{\sum_{r=0}^{N-1} \sum_{c=0}^{N-1} (c - \bar{c}_i)^2 I_i(r, c) - \sum_{r=0}^{N-1} \sum_{c=0}^{N-1} (r - \bar{r}_i)^2 I_i(r, c)}$$





Detectarea și interpretarea gesturilor mâinii din achiziția video în timp real

Decompoziția mâinii în elemente (mână, degete, antebraț)

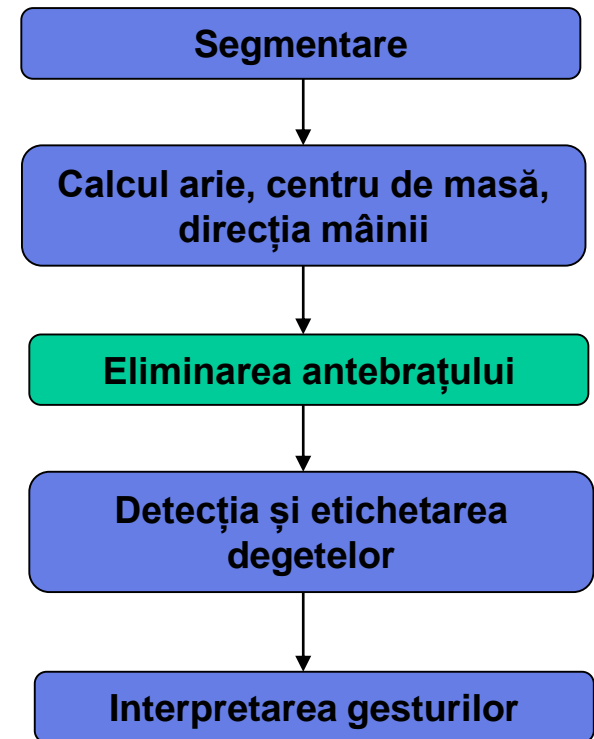
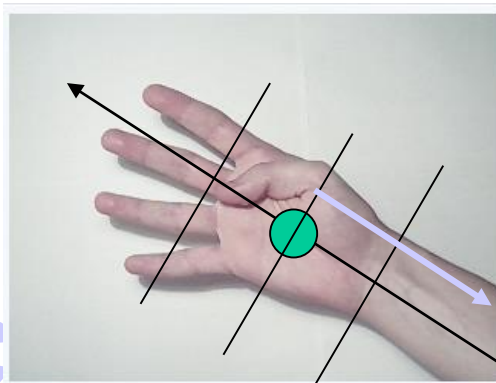
- Translatarea originii imaginii în centrul de greutate al obiectului + rotație

$$R = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \quad T = \begin{bmatrix} -ci \\ -ri \end{bmatrix}$$

$$\begin{bmatrix} u \\ v \end{bmatrix} = R \left(T + \begin{bmatrix} x \\ y \end{bmatrix} \right)$$

Eliminarea antebrațului

- Se realizează aplicând un algoritm de determinare a monotoniei proiecției verticale

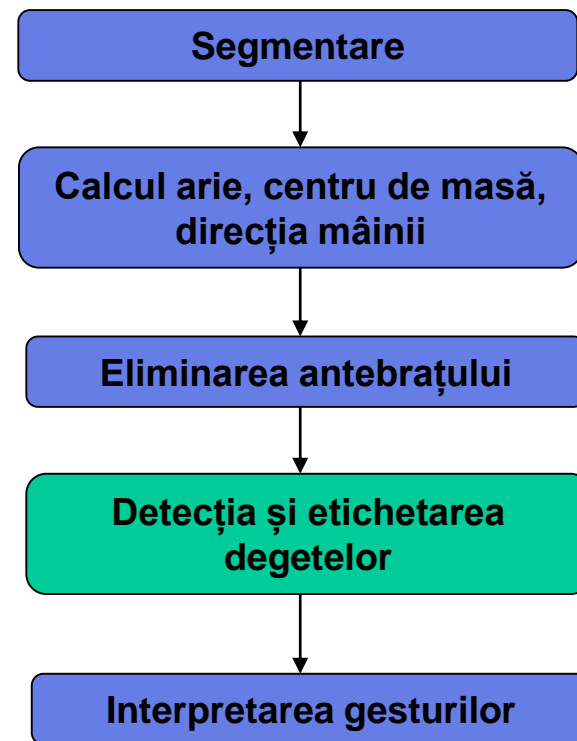




Detectarea si interpretarea gesturilor mâinii din achiziția video in timp real

Detecția si etichetarea degetelor

- Se determină centrul de masă al suprafeței rămase (palma + degete)
- Se trasează un cerc de rază = $\text{arie_palmă}/190$ (valoare determinată experimental)
- Se elimină punctele din interiorul cercului, rămânând doar degetele
- Se aplică algoritmul de etichetare bazat pe clase de echivalența

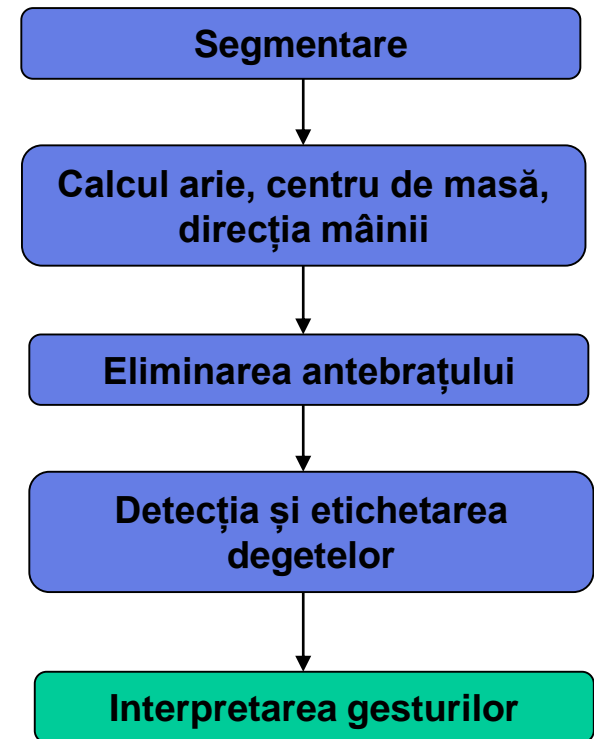




Detectarea si interpretarea gesturilor mâinii din achiziția video in timp real

Interpretarea gesturilor

- 1 deget detectat => starea nr.1 (S1)
 - transformare de perspectivă, adică rotația obiectului din scenă
- 2 degete detectate => starea nr.2 (S2)
 - scalarea obiectului din scenă
- 3 degete detectate => starea nr.3 (S3)
 - obiectul din scena se va schimba
- 4 degete detectate => starea idle (Si)
 - mediator de tranziții între restul stărilor
- 5 degete detectate => starea nr.5 (S5)
 - transformata pe culoare a obiectului din scenă





Detectarea si interpretarea gesturilor mâinii din achiziția video in timp real

Spatiu 3D de manipulare

Webcam Window -- no processing

Starea detectata S1

Image Processing Result Window

Panou de control

Selectie mijloc achizitie

- USB Webcam 1
- USB Webcam 2
- AVI Video (Xvid)

Cale video: inputXvid

Rata Detectie Gesturi

RDG: 5

Panou control Hue

Hue1_min	168
Hue1_max	180
Hue2_min	0
Hue2_max	50

Quit

Using Philips web... 1. Franz Ferdinand... Liceanta IfanView test_cons (Runnin... 6 test_cons 21:00



Detectarea si interpretarea gesturilor mâinii din achiziția video in timp real

The interface consists of several windows:

- Spatiu 3D de manipulare:** A 3D environment showing a large, multi-colored cube (rainbow gradient) on a black background.
- Panou de control:** A control panel with the following settings:
 - Selectie mijloc achizitie:
 - USB Webcam 1
 - USB Webcam 2
 - AVI Video (Xvid)
 - Cale video: InputXvid
 - Rata Detectie Gesturi: RDG: 5
 - Panou control Hue:
 - Hue1_min: 168
 - Hue1_max: 180
 - Hue2_min: 0
 - Hue2_max: 50
 - Quit button
- Webcam Window -- no processing:** Shows the original video frame of a hand gesture. Text below reads "Starea detectata S2".
- Image Processing Result Window:** Shows the detected gesture area as a dark shape on a black background.
- hsv-img:** Shows the processed binary image of the hand gesture.

The Windows taskbar at the bottom shows the following open applications: Using Philips web..., 1. Franz Ferdinand..., Liceanta, *C:\Users\Cristian..., test_cons (Runnin..., and 6 test_cons. The system clock shows 21:01.



Detectarea si interpretarea gesturilor mâinii din achiziția video in timp real

The interface consists of several windows:

- Spatiu 3D de manipulare**: A 3D environment showing a large, multi-colored plane (rainbow gradient) on a black background.
- Webcam Window -- no processing**: A video feed showing a hand gesture. Below the video, the text "Starea detectata S3" is displayed.
- Image Processing Result Window**: A window showing the processed video feed, where the hand is highlighted in white against a black background.
- Panou de control**: A control panel with the following settings:
 - Selectie mijloc achizitie:
 - USB Webcam 1
 - USB Webcam 2
 - AVI Video (Xvid)
 - Cale video:
 - Rata Detectie Gesturi:
 - RDG:
 - Panou control Hue:
 - Hue1_min:
 - Hue1_max:
 - Hue2_min:
 - Hue2_max:
 - Quit button

The Windows taskbar at the bottom shows the following open applications: Using Philips web..., 1. Franz Ferdinand..., Liceanta, Clipboard03 - Irf..., test_cons (Runnin..., and 6 test_cons. The system clock shows 21:02.



Detectarea si interpretarea gesturilor mâinii din achiziția video in timp real

Spatiu 3D de manipulare

Webcam Window -- no processing

Starea detectata S4

Image Processing Result Window

hsv-img

Panou de control

Selectie mijloc achizitie

- USB Webcam 1
- USB Webcam 2
- AVI Video (Xvid)

Cale video: inputXvid

Rata Detectie Gesturi

RDG: 5

Panou control Hue

Hue1_min	168
Hue1_max	180
Hue2_min	0
Hue2_max	50

Quit

Using Philips web... 1. Eve 6 - Inside O... Liceanta *C:\Users\Cristian... test_cons (Runnin... 6 test_cons EN 21:03



Detectarea si interpretarea gesturilor mâinii din achiziția video in timp real

The screenshot displays a software interface for real-time hand gesture detection. The main window, titled "Spatiu 3D de manipulare", shows a large red 3D hand model against a black background. A control panel titled "Panou de control" is overlaid on the main window, containing the following settings:

- Selectie mijloc achizitie:
 - USB Webcam 1
 - USB Webcam 2
 - AVI Video (Xvid)
- Cale video: inputXvid
- Rata Detectie Gesturi:
 - RDG: 5
- Panou control Hue:
 - Hue1_min: 168
 - Hue1_max: 180
 - Hue2_min: 0
 - Hue2_max: 50
- Quit button

Three smaller windows are visible on the right side:

- "Webcam Window -- no processing": Shows the original video frame with a hand. Text below reads "Starea detectata S5".
- "Image Processing Result Window": Shows the detected hand area as a white shape on a black background.
- "hsv-img": Shows the processed image with the hand area highlighted in white.

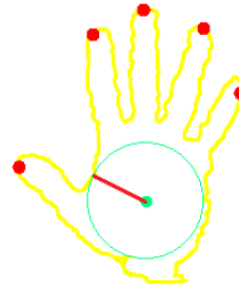
The Windows taskbar at the bottom shows the following open applications: "Using Philips web...", "1. [Buffer: 0%] Eve...", "Licenta", "*C:\Users\Cristian...", "test_cons (Runnin...", and "6 test_cons". The system clock shows 21:03.



Metoda alternative de detectie a degetelor

1. Clasificarea punctelor ce aparțin mâinii (segmentare) \Rightarrow img. binara
2. Detectia conturului mainii (ex: met. chain code)
3. Detectia punctelor din interiorul conturului (Obiect \ Contur)
4. Detectia centrului palmei
5. Găsirea vârfurilor degetelor

4. Centrul palmei: reprezintă centrul celui mai mare cerc care poate fi desenat în interiorul palmei.

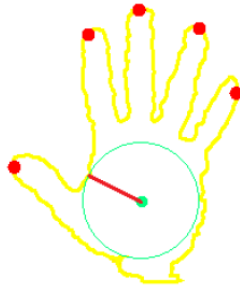


Pentru a calcula centrul palmei, se calculează distanțele minime dintre punctele interioare și punctele conturului mâinii. Punctul care corespunde maximului acestor distanțe reprezintă centrul palmei.



Metoda alternative de detectie a degetelor

4. Centrul palmei: reprezintă centrul celui mai mare cerc care poate fi desenat în interiorul palmei.



Optimizarea algoritmului:

- când se calculează distanța minimă a unui punct interior față de fiecare punct de pe contur, dacă minimul curent este mai mic ca maximul curent, putem fi siguri că acest punct nu este centrul, și nu mai este nevoie să calculăm distanța pentru restul punctelor din contur.
- știind că punctele interioare și cele de contur sunt foarte apropiate unele de altele, putem crește performanța prin executarea algoritmului pe 1 din N puncte consecutive. Eroarea care apare este neglijabilă, iar eficiența este crescută de $1/N^2$.



Metoda alternative de detectie a degetelor

4. Detectia centrului palmei – metoda alternativa:

Robust Hand Gesture Recognition Based on FingerEarth Mover's Distance with a Commodity Depth Camera, Zhou Ren Junsong Yuan, Zhengyou Zhang

http://web.cs.ucla.edu/~zhou.ren/Ren_Yuan_Zhang_MM11short.pdf

- Se calculeaza transformata distanta (DT) a pixelilor din interiorul conturului

http://users.utcluj.ro/~igiosan/Resources/PRS/L4/lab_04e.pdf

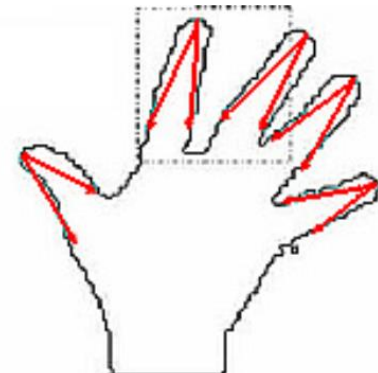
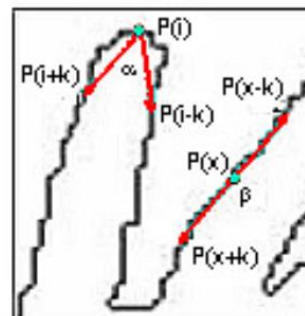
- Pixelul cu cea mai mare valoare a DT va fi centrul palmei





Metoda alternative de detectie a degetelor

5. Gasirea varfului degetelor: algorimul k-curvature descris în [3]



- Pt. fiecare punct ce aparține conturului $P(i)$, consideram 2 puncta ajutatoare: $P(i-k)$ și $P(i+k)$. Cu aceste 3 puncte vom genera 2 vectori și vom calcula unghiul minim format. Vectorii sunt formați de $P(i-k) - P(i)$ și $P(i+k) - P(i)$. Dacă unghiul este mai mic decât o valoare α , atunci punctul $P(i)$ reprezintă vârful unui deget. Valori potrivite pt. $k \sim$ rezoluție imagine și $\alpha = 40$.
- Pentru a evita găsirea unui vârf fals reprezentat de o vale, vom compara distanța dintre centrul palmei și punctul găsit, cu distanța dintre centrul palmei și cele două puncte ajutatoare. Dacă prima distanță este mai mică, avem un vârf fals.



Detectia si interpretarea gesturilor bazata pe makere de culoare [4]

Specificarea cerintelor si solutia aleasa

Cerințele temei:

- Detectia gesturilor mainii. Se va folosi ca mijloc de achizitie o camera web. Se vor folosi markere de diferite culori pe varful degetelor pentru a facilita detectia. Markerele vor fi izolate de restul obiectelor din imagine.

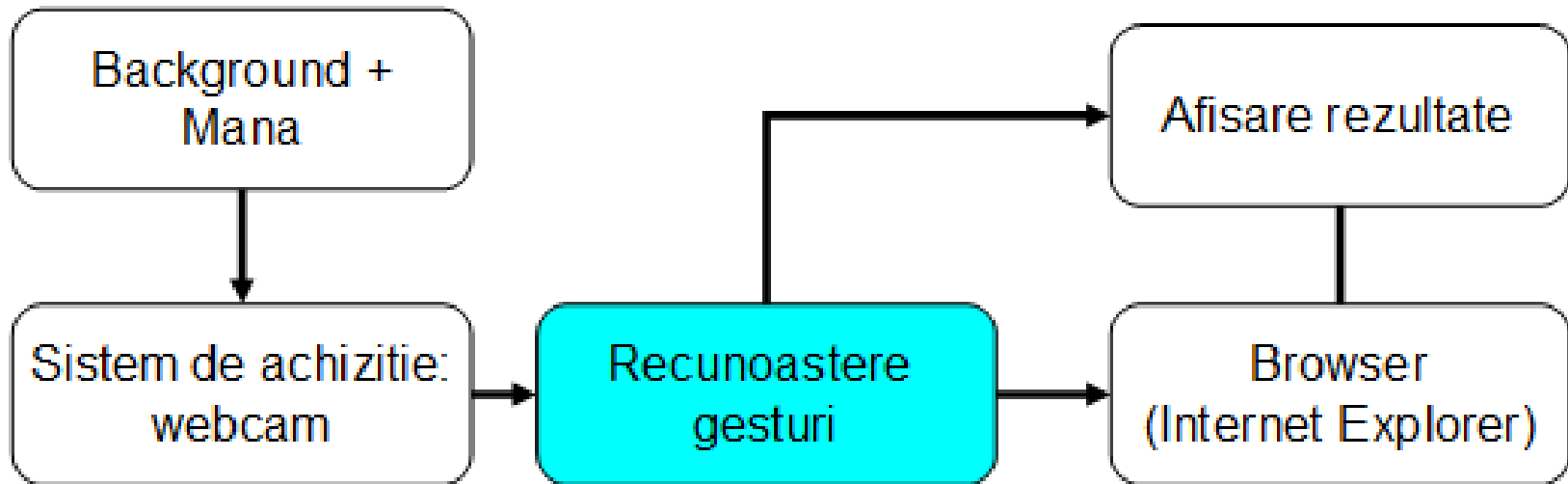
Solutia aleasa:

- Utilizarea unei camera web cu ajutorul careia se vor achizitiona imagini RGB de dimensiunea 640x480
- Pentru implementarea sistemului s-a utilizat mediul de dezvoltare Microsoft Visual Studio 2008 si bibliotecile OpenCV (doar pt. achizitia de imagini in timp real)



Detectia si interpretarea gesturilor bazata pe makere de culoare

Arhitectura conceptuala a aplicatiei



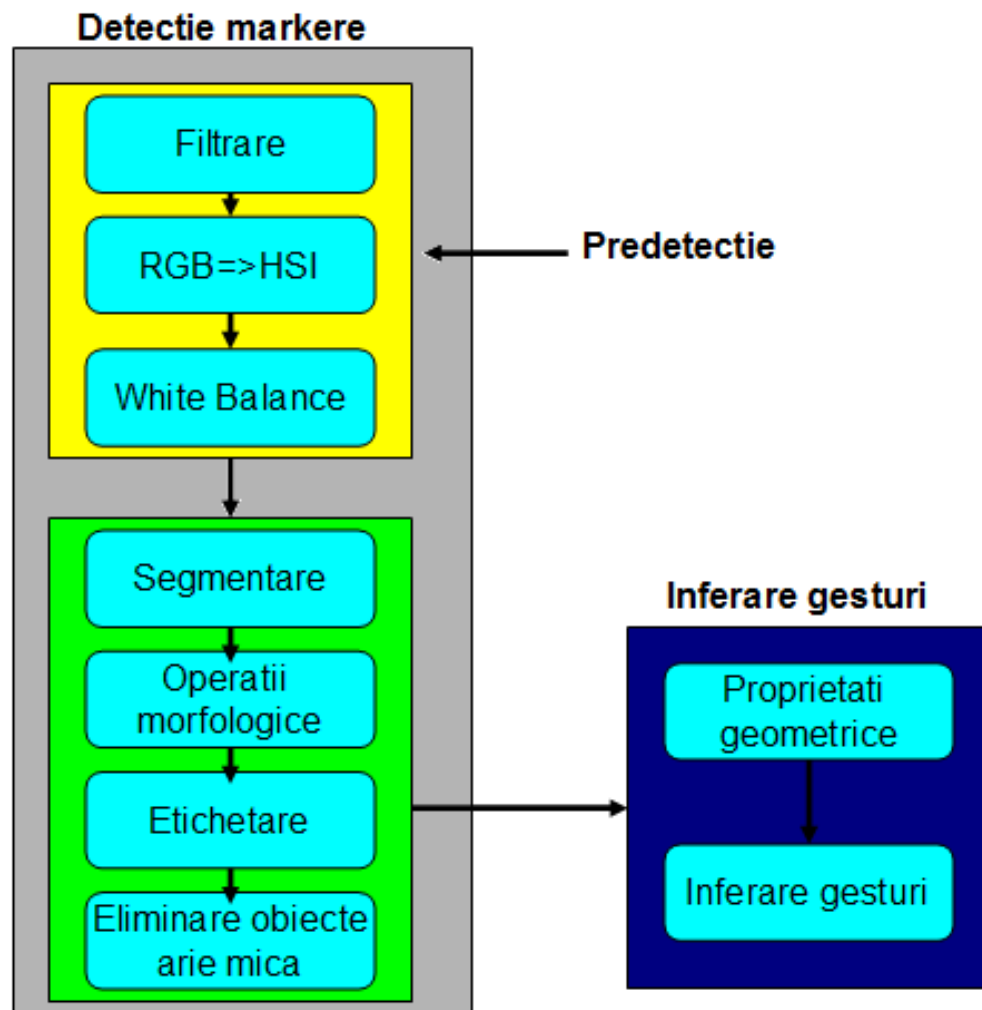


Detectia si interpretarea gesturilor bazata pe makere de culoare

Modulul de recunoastere si interpretare a gesturilor

In cadrul acestui modul pentru fiecare imagine (frame) primita de la webcam se aplica o serie operatii:

- predetectie
- detectie
- inferare gesturi





Detectia si interpretarea gesturilor bazata pe makere de culoare

Predectia

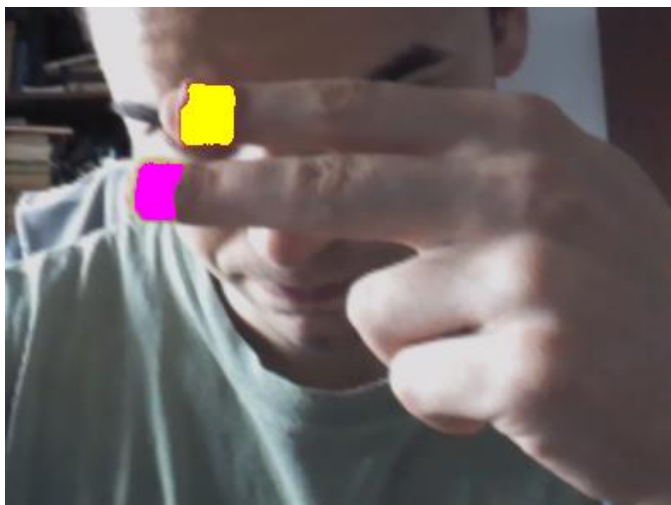
- **White balance**: operatie necesara pentru a reda corect culorile (<http://www.digitalcameraworld.com/2014/01/31/white-balance-explained-how-your-camera-corrects-the-colour-of-different-kinds-of-lighting/>, https://en.wikipedia.org/wiki/Color_balance)
- **Filtrarea imaginii**: se realizeaza prin convolutia imaginii cu un filtru gaussian – pt a elimina zgomotul (sursa de informatie nedorita).
- **Transformarea RGB => HSI**: spatiul de culoarea HIS (hue, saturatie, intensitate) este invariant la variatiile de iluminare ale scenei; se construiesc mai usor relatiile intre pixeli.



Detectia si interpretarea gesturilor bazata pe makere de culoare

Detectia propriu zisa

- **Segmentarea imaginii:** se aplica un algoritm de segmentare prin care se marcheaza toti pixelii din imagine care au valoarea *hue* intr-un interval (obtinut prin antrenare).
- **Operatii morfologice:** prin operatia morfologica de *deschidere* se vor elimina pixelii singulari, iar prin cea de *inchidere* se vor umple golurile din obiecte.
- **Etichetarea obiectelor:** se realizeaza pintr-un algoritm de tichetare cu clase de echivalenta, in urma etichetarii obiectelor fiecare obiect va putea fi tratat individual;
- **Eliminarea obiecte de arie mica:** aria unui obiecte este egala cu numarul de pixeli ce il formeaza; obiectele de arie mica sunt presupuse zgomote



Rezultatul in urma operatiilor de mai sus:
-markerele de pe degete sunt izlotate
-sunt reprezentate prin galben si violet



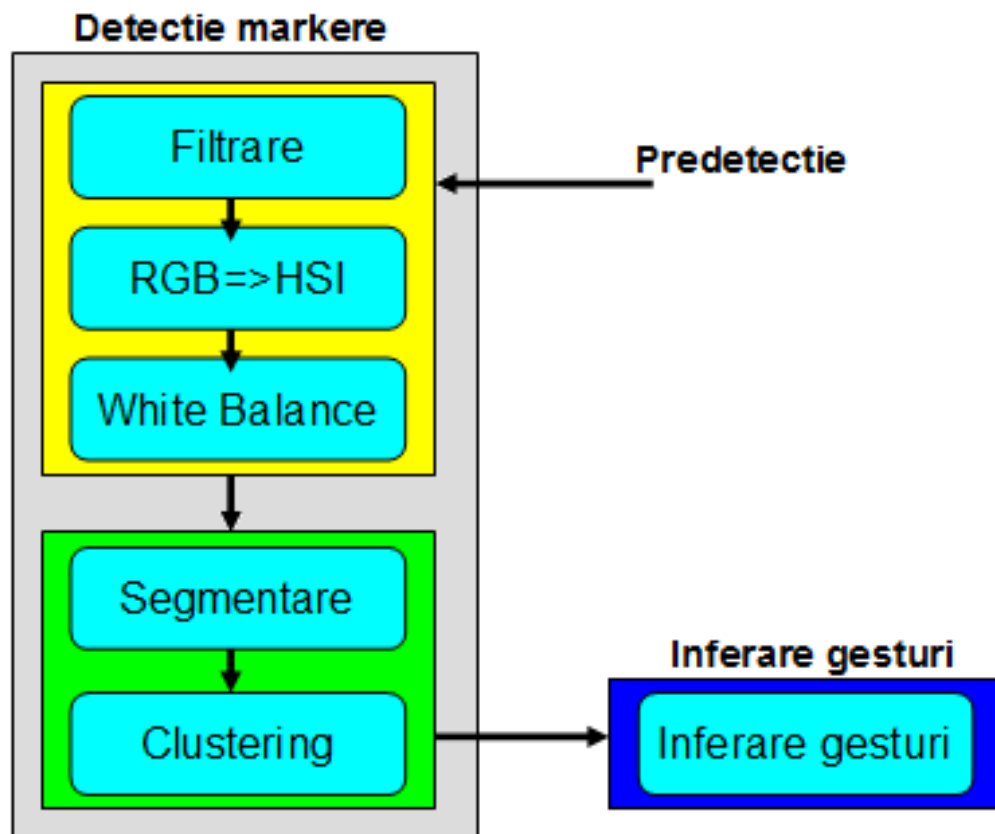
Detectia si interpretarea gesturilor bazata pe makere de culoare

Solutia imbunatatita de segmentare

Daca lumina este focalizata pe markere , zone de pixeli de pe markere se satureaza foarte mult \Rightarrow obiecte "sparte". In urma etichetarii \Rightarrow un numar mai mare de obiecte decat cel real (fiecare marker era reprezentat de 2 sau mai multe obiecte in loc de unul cum este normal).

<i>Obiect rosu</i>						<i>Obiect rosu spart</i>					
0	0	0	0	0	0	0	0	0	0	0	0
0	1	1	1	1	0	0	1	1	0	1	0
0	1	1	1	1	0	0	1	0	0	1	0
0	1	1	1	1	0	0	0	0	1	1	0
0	0	0	0	0	0	0	0	0	0	0	0

In aceasta situatie este necesara abordarea unei metode de grupare Spatiala (clustering) a zonelor cu densitate mare de pixeli

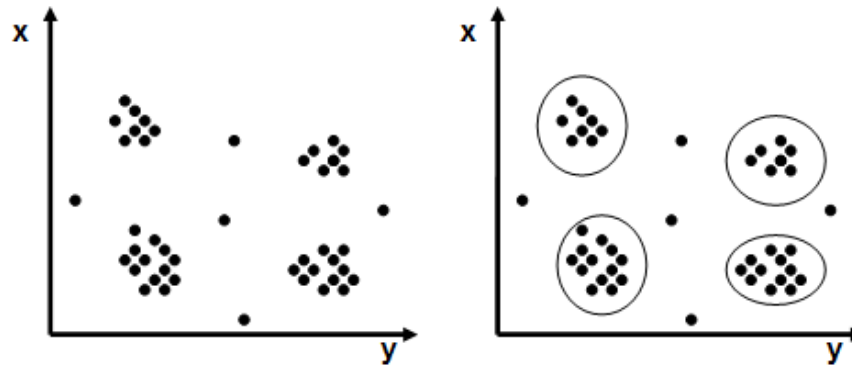




Detectia si interpretarea gesturilor bazata pe makere de culoare

Imbunatatirea segmentarii prin clustering

- Pentru a gasi zone in imagine cu densitate mare de pixeli de nuante apropiate am folosit/implementat algoritmul nesupervizat MBSAS (modified basic sequential algorithmic scheme) [5], care presupune determinarea clusterelor intr-o prima faza dupa care fiecare punct neassignat inca este adaugat celui mai apropiat cluster.



- Un cluster este caracterizat prin :
 - centrul clusterului
 - cardinalul clusterului – numarul de puncte din cluster
- Se vor pastra doar clusterii care au un numar de puncte mai mare ca un *prag*



Detectia si interpretarea gesturilor bazata pe makere de culoare

Inferare gesturi

Pentru a recunoaste si interpreta gesturi trebuie sa gasim pozitia relativa a markerelor in imagine

Proprietatii geometrice: pozitia relativa a unui obiect este data de centrul de masa:

$$\mathbf{r}_i = \frac{1}{M} \sum_j m_j \mathbf{r}_j$$
$$\mathbf{r}_i = \frac{1}{M} \sum_j m_j \mathbf{r}_j$$

Inferare gesturi: pentru a construi gesturi ma voi folosi de:

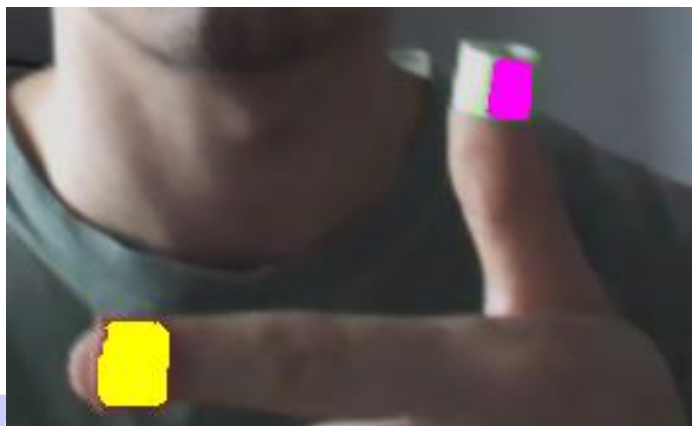
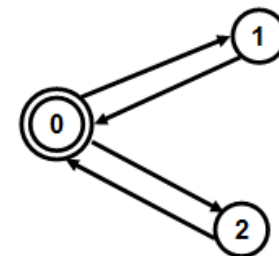
- pozitia markerelor in imagine
- distanta dintre anumite markere
- numarul de markere din imagine
- prezenta sau absenta unor markere de anumita culoare din imagine



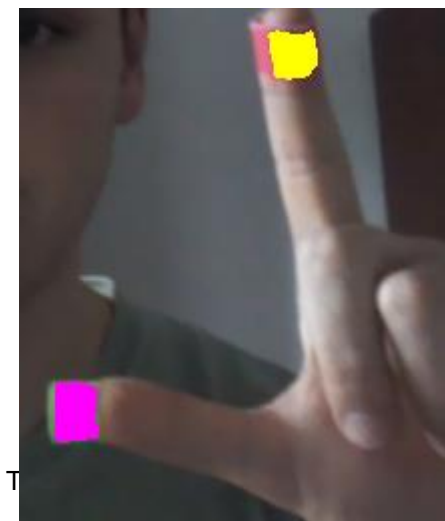
Detectia si interpretarea gesturilor bazata pe makere de culoare

Starile aplicatiei

- Se folosesc 3 culori de markere: rosu, verde si galben
- Aplicatia este construita in jurul a 3 stari principale:
 - **starea 0** – aici se realizeaza **antreanarea** ;din aceasta stare facand anumite gesturi se va trece in celalate 2 stari
 - **starea 1** – in acesta stare se va pune la dispozitia utilizatorului o tastatura virtuala asemanatoare cu tastatura unui telefon mobil cu ajutorul careia se vor scrie diferite caractere; va permite accesarea unei pagini web
 - **starea 2** – se vor recunoaste gesturile: *zoom in*, *zoom out*, si gestul de revenire in starea 0



Daca se introduce un deget cu marker de culoare galbena in dreptunghiul format de celelate doua markere: rosu si verde si daca pozitia celor 2 degete este conform imaginii din dreapta at. se va trece in **starea 1** , iar daca este ca in imaginea din stanga at. se va trece in **starea 2**





Detectia si interpretarea gesturilor bazata pe makere de culoare

Starea 1

- In aceasta stare utilizatorul va putea scrie diferite caractere cu ajutorul unei tastaturi virtuale. Starea 1 este impartita la randul ei in 3 substari.
- Constrangeri:
 - pozitia markerului rosu introduce o relatie de “deget peste” anumit buton
 - prezenta markerului de culoare galbena introduce relatia de “apare” a butonului.
 - numarul de degete din imagini va indica caracterul rezultat sau actiunea rezultata

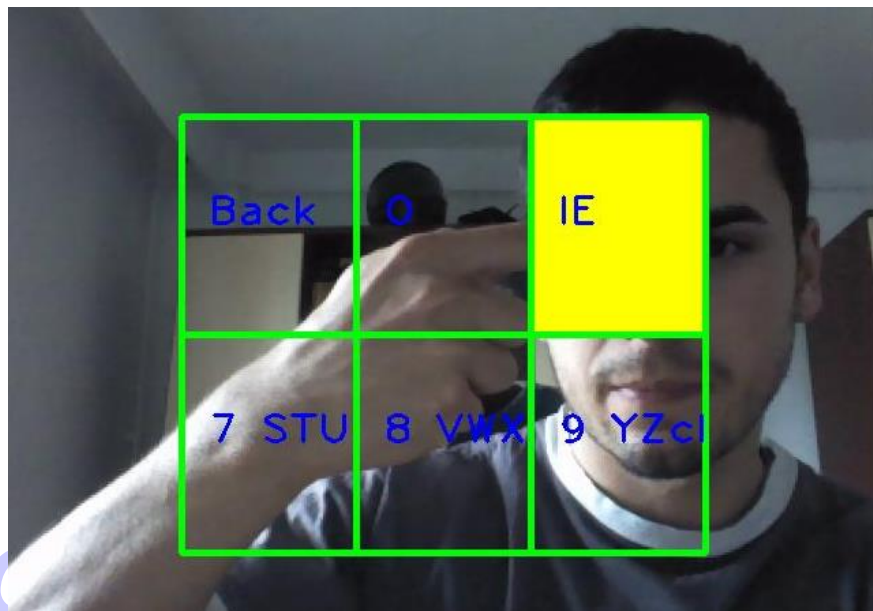
Tasta	Marker rosu	Marker galben	Numar markere	Rezultat
1 ABC	Peste tasta 1	exista	2	1
1 ABC	Peste tasta 1	exista	3	a
1 ABC	Peste tasta 1	exista	5	c
FW	Peste tasta FW	exista	3	Substarea 2 sau 3
Back	Peste Back	exista	3	Substarea 1 sau 2 Sau starea 0



Detectia si interpretarea gesturilor bazata pe makere de culoare

Starea 1

Degetul cu marker de culoare rosie este peste butonul 1, in momentul in care in imagini succesive va fi prezent degetul cu marker galben tasta 1 se va considera apasat, iar numarul de degete din imagini va da caracterul -starea este 1, substarea este 1



Degetul cu marker de culoare rosie este peste butonul **IE** , in momentul in care in imagini succesive va fi prezent degetul cu marker galben iar numarul de degete va fi egal cu 5 , se va accesa o pagina de internet avand URL-ul format din caracterele tastate anterior si care au fost pastrate intr-un vector de caractere

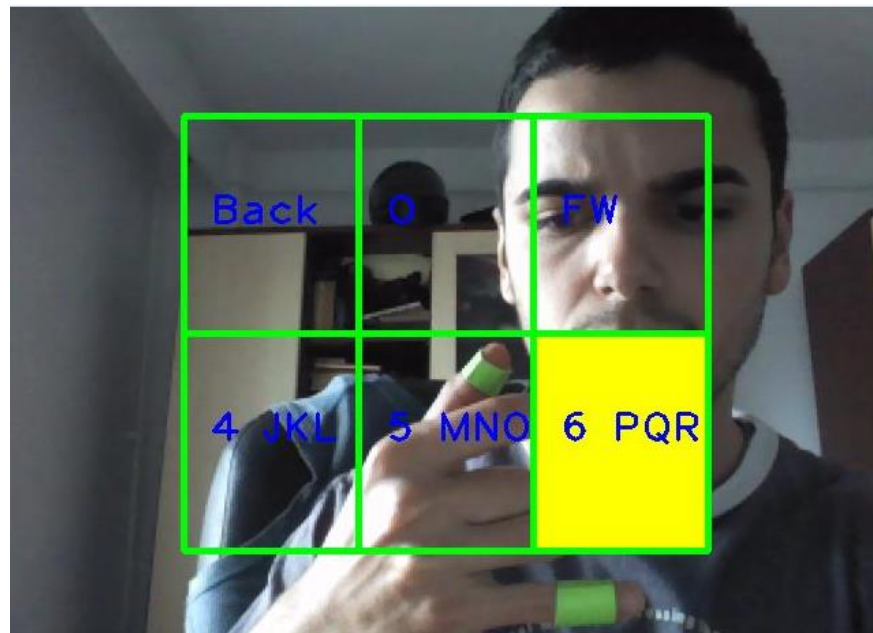
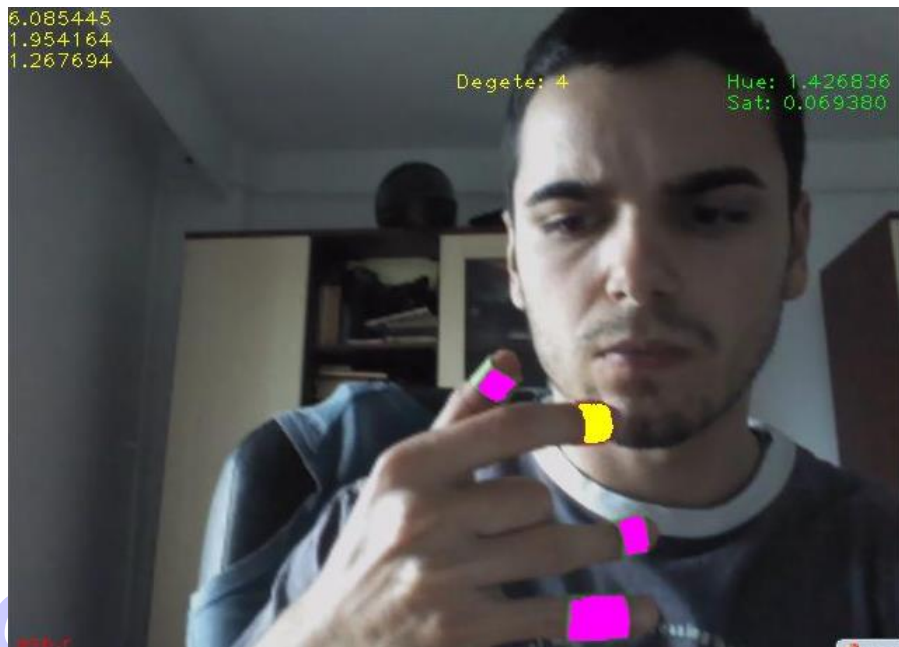


Detectia si interpretarea gesturilor bazata pe makere de culoare

Starea 1

Se folosesc 2 ferestre pentru afisarea rezultatelor:

- in imaginea din stanga sunt indetificate markererele
- se afiseaza valorile *hue* care au fost setate la antrenare- stanga sus
- valorile *hue* si a saturatiei din zona de antrenare sunt afisate in dreapta sus
- numarul de degete: 4 in imaginea de mai jos
- In stanga jos este afisata adresa URL formata prin apasarea a diferitor butoane in imaginea de mai jos este "gsp.ro"
- daca un caracter a fost gresit el se poate sterge- tasta 0...5 degete in imagine





Detectia si interpretarea gesturilor bazata pe makere de culoare

Starea 2

- Sunt neceare 4 degete: 2 degete avand marker rosu, iar 2 degete cu marker verde
- Se va calcula distanta euclidiana intre degetele cu marker de aceeasi culoare:
 - daca distanta intre markererele rosii si markererele verzi va creste in imagini succesive si daca distanta intre markererele verzi este mai mica decat distanta intre degetele rosii atunci gestul va recunoscat ca *zoom in*- se va afisa testul "zoom in" ; Font size-ul pentru textul "zoom in" va creste de fiecare data cand se va face acest gest.

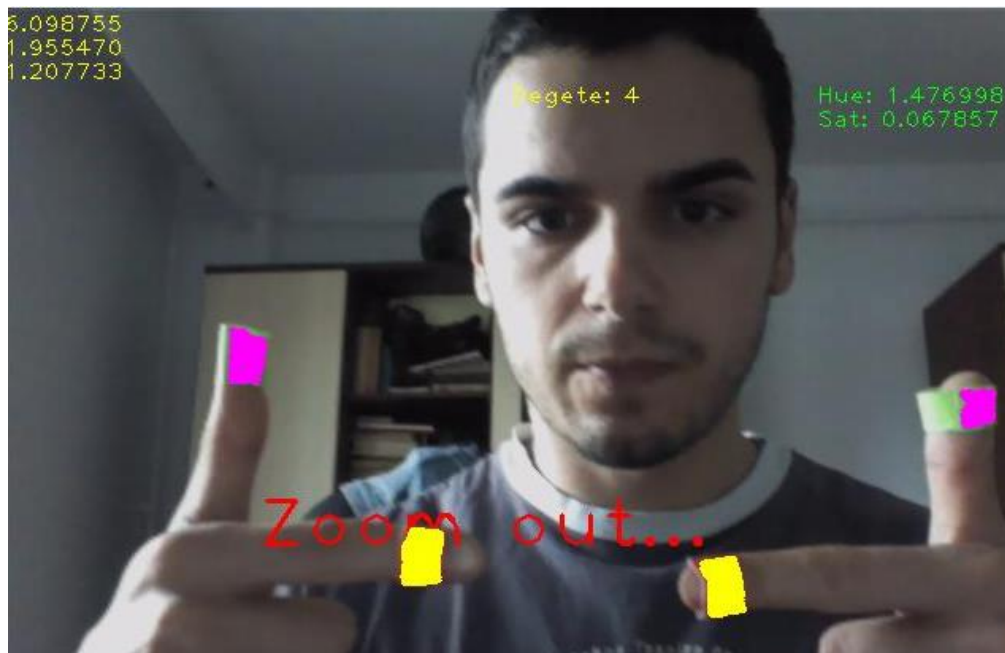




Detectia si interpretarea gesturilor bazata pe makere de culoare

Starea 2

- Sunt neceare 4 degete: 2 degete avand marker rosu, iar 2 degete cu marker verde
- Se va calcula distanta euclidiana intre degetele cu marker de aceeasi culoare:
 - daca distanta intre markerele rosii si markerele verzi va scadea in imagini succesive si daca distanta intre markerele verzi este mai mare decat distanta intre degetele rosii atunci gestul va recunoscat ca *zoom out*- se va afisa testul zoom out ; Font size-ul pentru textul “zoom out” va scadea de fiecare data cand se va face acest gest.





Detectia si interpretarea gesturilor bazata pe makere de culoare

Starea 2

- Pentru revenirea in starea initiala 0 se va face urmatorul gest:
 - toate degetele cu marker vor fi apropiate astfel incat distanta intre markere rosii si distanta intre markerele verzi sa fie mai mica decat un *prag*- se vor apropia cat de mult posibil cele 4 markere.





Detectia mainii si a gesturilor cu senzori de profunzime

Senzorii Intel RealSense

<https://software.intel.com/en-us/realsense/devkit>

F200: <https://software.intel.com/en-us/blogs/2015/01/26/can-your-webcam-do-this>

Intel RealSense SDK: Hand Tracking module [F200]

The SDK hand tracking module provides real-time 3D hand motion tracking, using a single depth sensor. The hand module can track one or two hands, providing precise joint-level locations and positions. The module can also identify “gestures”, which are certain significant hand postures or motions, for example a wave, tap or thumbs-up sign.

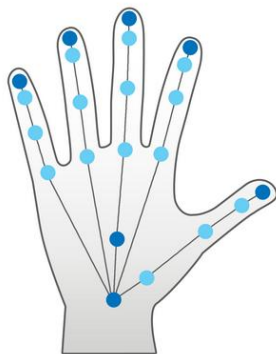


Detectia mainii si a gesturilor cu senzori de profunzime

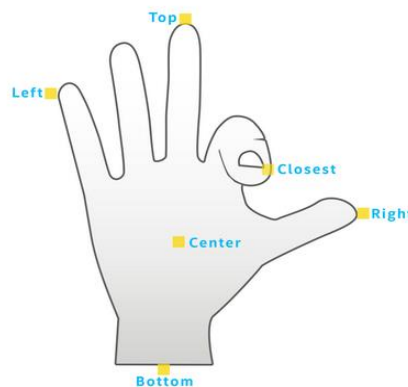
Intel RealSense SDK: Hand Tracking module [F200]

Tracking Modes: The hand module has two main tracking modes, which differ by the information they provide and the computation resources that they require:

- **Full-hand** - returns the full 3D skeleton of the hand, including all 22 joints, fingers information, gestures, and more.
- **Extremities** – returns the general location of the hand, its silhouette, and the extremities of the hand (6 points): the hand's top-most, bottom-most, right-most, left-most, center and closest (to the sensor) points).



Full mode



Extremities mode

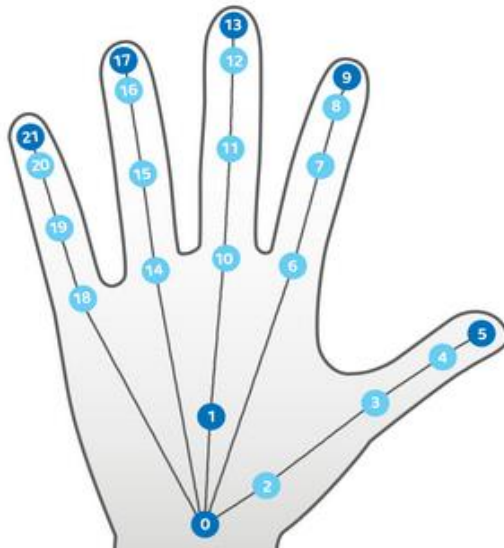


Detectia mainii si a gesturilor cu senzori de profunzime

Intel RealSense SDK: Hand joints

#	Joint Name	Joint Parent	#	Joint Name	Joint Parent
0	JOINT_WRIST	NA	11	JOINT_MIDDLE_JT1	10
1	JOINT_CENTER	0	12	JOINT_MIDDLE_JT2	11
2	JOINT_THUMB_BASE	0	13	JOINT_MIDDLE_TIP	12
3	JOINT_THUMB_JT1	2	14	JOINT_RING_BASE	0
4	JOINT_THUMB_JT2	3	15	JOINT_RING_JT1	14
5	JOINT_THUMB_TIP	4	16	JOINT_RING_JT2	15
6	JOINT_INDEX_BASE	0	17	JOINT_RING_TIP	16
7	JOINT_INDEX_JT1	6	18	JOINT_PINKY_BASE	0
8	JOINT_INDEX_JT2	7	19	JOINT_PINKY_JT1	18
9	JOINT_INDEX_TIP	8	20	JOINT_PINKY_JT2	19
10	JOINT_MIDDLE_BASE	1	21	JOINT_PINKY_TIP	20

Table 13: The Joint Name and Index.



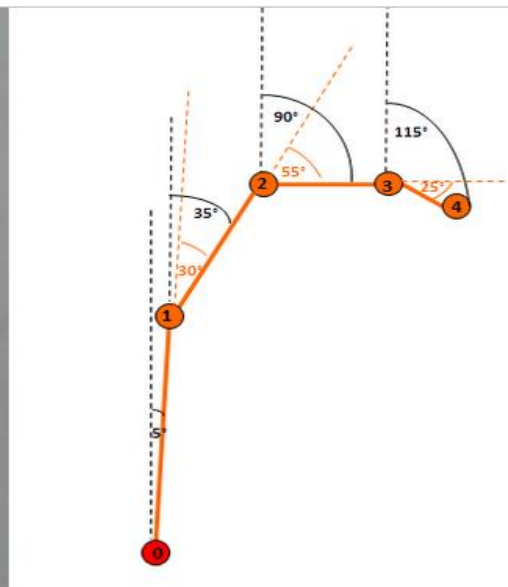
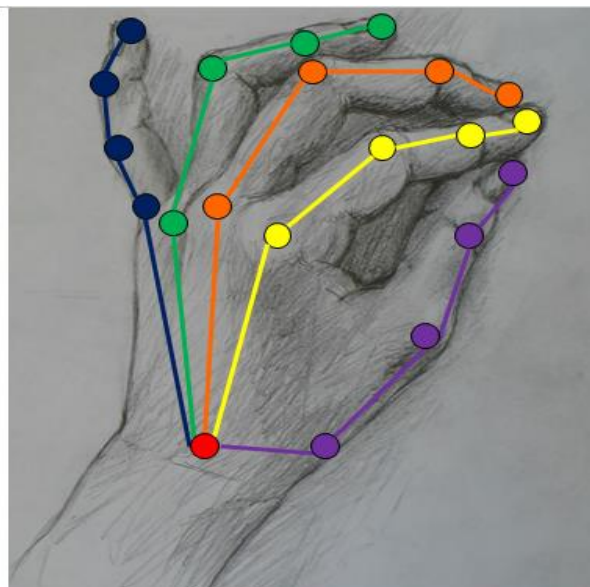
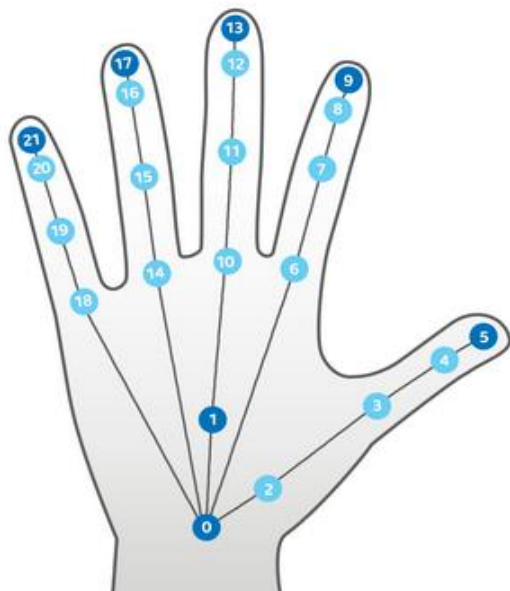


Detectia mainii si a gesturilor cu senzori de profunzime

Intel RealSense SDK: Local and Global Joint Rotation

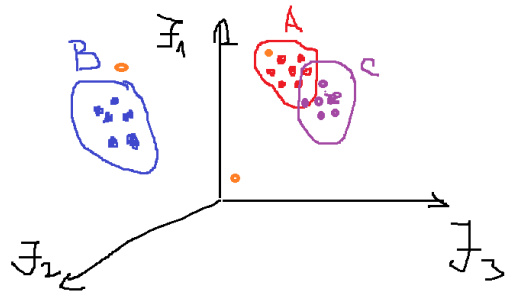
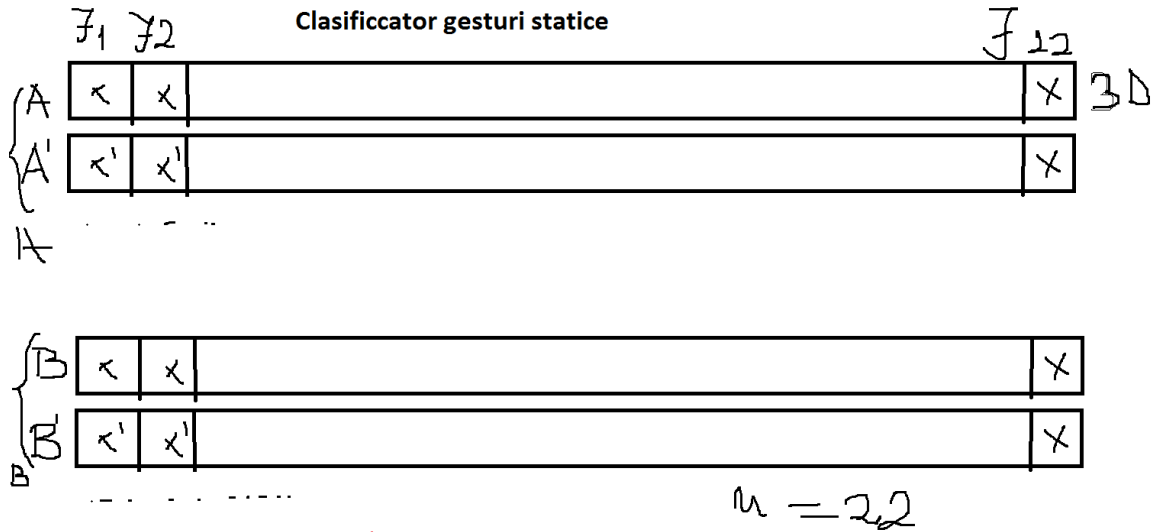
A joint's rotation refers to the angle of the joint's bone, which is the bone that ends at the joint and starts at the joint's parent. The hand module expresses joint rotation in two ways:

- **Local rotation** - the angle between the joint's bone and its parent's bone. This value is stored in the localRotation property of [JointData](#)
- **Global rotation** - the angle between the joint's bone and the camera coordinate system (Y axis). This value is stored in the globalOrientation property of [JointData](#)





Recunoasterea de gesturi complexe (statice / dinamice)



Detectie gesturi dinamice

V_g

V_t

DTW (Dynamic Time Warping)



Bibliografie

- [1] E. Visky, Interfețe om-calculator: detecția mâinii și a gesturilor prin viziune monoculară, Lucrare de diploma, Catedra de Calculatoare, UTCN, 2007.
- [2] C. Bara, Detectarea și interpretarea gesturilor mâinii din achiziția video în timp real, Lucrare de diploma, Catedra de Calculatoare, UTCN, 2009.
- [3] Trigo T. R., Pellegrino S. R. M., An analysis of features for hand-gesture classification. 17th International Conference on Systems, Signals and Image Processing (IWSSIP 2010).
- [4] E. Iclodean, Sistem de detecție și interpretare a gesturilor bazat pe markere de culoare, Lucrare de diploma, Catedra de Calculatoare, UTCN, 2010.
- [5] Theodoridis, K. Koutroumbas, Pattern Recognition, 2-nd Edition, Academic Press, 2003.