**3. Arhitectura aplicației**

Proiectul recunoaște poziția corpului uman la un moment dat prin supravegherea unui individ cu un dispozitiv Kinect și își propune atât recunoașterea poziției individului, cât și a mișcării în curs.

O schemă a arhitecturii sistemului este prezentată în cele ce urmează, în Fig. 5:

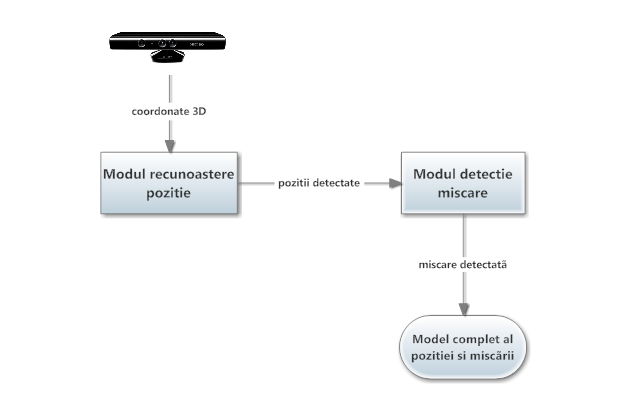


Fig. 5 – Arhitectura generală a aplicației

Aplicația primește ca date de intrare coordonate ale scheletului corpului uman. Acestea sunt furnizate de către dispozitivul Kinect. Alături de senzorii de culoare, cameră și cele 4 microfoane ale dispozitivului, se află senzorul de adâncime infraroșu, care se dovedește a fi foarte important în cadrul întregii aplicații. În acest mod, este posibilă preluarea informațiilor articulațiilor scheletului uman în spatiu tridimensional.

Modulul de recunoastere a poziției corpului uman, preia datele furnizate de dispozitiv și construiește câte un vector de caracteristici pentru fiecare parte componenta a scheletului corpului uman:

❖ trunchi

❖ membru superior stâng

❖ membru superior drept

❖ membru inferior stâng

❖ membru inferior drept

Cu ajutorul valorilor din vectorul de caracteristici, se construiesc arbori de decizie pentru fiecare componentă a corpului în parte. Astfel, pentru determinarea poziției corpului uman la un moment de timp se interoghează în mod paralel pe câte un fir de execuție fiecare din acești arbori de decizie și se furnizeaza o poziție pentru fiecare parte componentă. Pentru fiecare parte componentă în parte din cele cinci, se construiește un arbore de decizie. Algoritmul de construcție de la baza acestora este un algoritm ID3 cu extinderi C4.5 pentru atributele numerice (unghiuri între articulatții). Cei cinci arbori diferă prin atributele și clasele din componența acestora.

Alături de modulul complex de determinare a poziției statice a individului, se află modulul de recunoaștere a mișcării. O mișcare se poate descrie folosind pozițiile intermediare, prin înlănțuire. Astfel, pentru detecția mișcării corpului uman este necesară evaluarea poziției curent recunoscute în raport cu poziția sau pozițiile anterioare (în cazul unei mișcări complexe, ce necesita parcurgerea mai multor stări până la întâlnirea unei stări finale).

Modulul de detecție a mișcării folosește rezultatul furnizat de primul modul și împarte poziția completă rezultată în cinci poziții corespunzătoare fiecărei părti componente a corpului uman. În funcție de poziția primită, se reactualizează starea curentă din lantul Markov. Următoarea stare depinde de tranziția efectuată. Fiecare din această tranziție este caracterizată de o pondere (probabilitate de efectuare a tranziției). Aceasta își va mări valoarea, în timp ce restul tranzițiilor de ieșire din starea trecută își vor micșora probabilitatea.

**3.1 Modulul de recunoaștere a poziției**

Modulul de recunoaștere a poziției corpului uman folosește ca date de intrare coordonate tridimensionale furnizate de dispozitivul Kinect. Deoarece în cadrul aplicației de recunoaștere a pozițiilor și mișcărilor în timp real resursa de timp reprezintă un factor foarte important, antrenarea și testarea se realizează în mod paralel pentru fiecare arbore în parte – atât construcția, cât și parcurgerea lor - pentru obținerea rezultatului final într-un timp cât mai scurt. Datele de intrare obținute de la dispozitiv sunt folosite atât în etapa de învățare, cât și în etapa de testare:

❖ etapa de învățare – se dorește construirea a 5 arbori de decizie pentru cele 5 părți componente ale corpului uman pe baza atributelor obținute din prelucrarea coordonatelor 3D obținute de la dispozitivul Kinect: unghiurile formate între diferite articulații. Clasele posibile din cadrul algoritmului de clasificare sunt în numar aproximativ de 4 pentru fiecare componentă (excepție făcând trunchiul, unde se pot identifica 5 poziții), ceea ce înseamnă că pot fi detectate maxim 4\*4\*4\*4\*5= 1280 de poziții (numărul real de poziții identificate este mai mic deoarece sunt eliminate cazurile în care ambele membre inferioare sunt ridicate – lateral, frontal - pentru oricare din pozițiile trunchiului și a membrelor superioare). Setul total de poziții identificate este anexat lucrării.

❖ etapa de testare – se interoghează cei 5 arbori construiți în etapa precedentă și se obține o poziție totală (ce caracterizează întregul corp) prin reuniunea celor 5 poziții ale părților componente: trunchi și cele 4 membre.

Arhitectura acestui modul este prezentată în Fig. 6 :

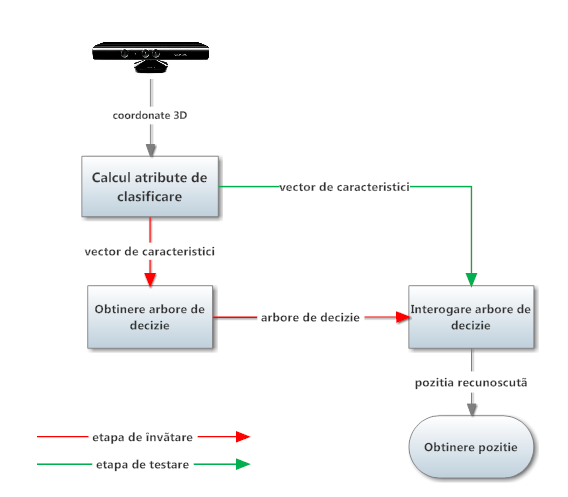


Fig. 6 – Arhitectura modulului de recunoaștere a poziției

**Calcul vector de caracteristici**

Pentru a construi cei 5 arbori, este necesară determinarea vectorului de caracteristici - calculul atributelor ce realizează clasificarea în arbori pentru fiecare intrare din setul de test. Setul de test este constituit din coordonatele tridimensionale a 20 de articulații, ilustrate in Fig. 7.

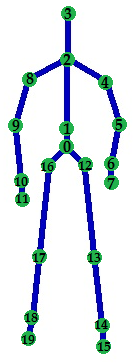


Fig. 7 – Cele 20 de coordonate 3D furnizate de dispozitivul Kinect

Cu ajutorul acestor coordonate, se calculează unghiurile necesare fiecărei pozitii ale corpului, pentru o ulterioară clasificare:

❖ arborele caracteristic al trunchiului :

∘ ⊀ (2-1-13) – unghiul format de articulațiile 2, 1 și 13

∘ ⊀ (2-1-17) – unghiul format de articulațiile 2, 1 și 17

∘ ⊀ ((1,2), yOz) – unghiul format de dreapta articulațiilor 1 și 2 cu planul yOz

∘ ⊀ ((1,2), xOy) – unghiul format de dreapta articulațiilor 1 și 2 cu planul xOy

❖ arborii caracteristici (2) ai membrelor superioare :

∘ ⊀ ((8-9), yOz), respectiv ⊀ ((4-5), yOz) – unghiul format de dreapta articulațiilor 8 și 9, respectiv 4 și 5 cu planul yOz

∘ ⊀ ((8-9), xOy), respectiv ⊀ ((4-5), xOy) – unghiul format de dreapta articulațiilor 8 și 9,respectiv 4 și 5 cu planul xOy

∘ d(8y, 9y), respectiv, d(4y, 5y) – distanța dintre articulațiile 8 și 9, respectiv 4 și 5

❖ arborii caracteristici (2) ai membrelor inferioare :

∘ ⊀ (16-17-18), respectiv ⊀ (12-13-14) – unghiul format de articulațiile 16, 17 și 18, respectiv 12, 13 și 14

∘ ⊀ ((17-18), xOy), respectiv ⊀ ((13-14), xOy) – unghiul format de dreapta articulațiilor 17 și 18, respectiv 13 și 14 cu planul xOy

∘ ⊀ ((17-18), yOz), respectiv ⊀ ((13-14), yOz) – unghiul format de dreapta articulațiilor 17 și 18, respectiv 13 și 14 cu planul yOz

Folosind ca atribute unghiurile formate dintre un plan și diferite drepte formate din două puncte de articulație, pentru o aceeași poziție, dar cu rotație diferită față de dispozitivul Kinect, rezultatul furnizat va fi același. Cu alte cuvinte, poziția rezultată este invariantă față de rotația în jurul celor 3 axe de coordonate.

În cazul unor articulații obturate de către o altă parte a corpului sau de către un alt obiect din mediul înconjurator, pozitia respectivei articulații va fi furnizată de către Kinect, cu ajutorul unor estimări. Evident, cea mai bună soluție în acest caz este folosirea în paralel a două dispozitive Kinect astfel încât ambele rezultatele furnizate pentru aceeași poziție vor fi interpolate, rezultând o poziție finală, fără vreo articulație obturată.

Așa cum s-a mentionat în paragraful anterior, dispozitivul Kinect furnizează cele 20 de coordonate în mod necondiționat. Coordonatele 3D ale articulațiilor sunt însoțite de o etichetă ce poate lua una din următoarele 3 valori:

❖ „tracked” - în cazul în care coordonata a fost detectată cu succes,

❖ „not tracked” - în cazul în care coordonata nu a fost detectată (scheletul nu este afisat),

❖ „inferred” - în cazul în care poziția articulației nu a fost detectată, dar a putut fi estimată de către dispozitiv.

Pentru etapa de antrenare din cadrul modulului de recunoaștere a poziției sunt folosite doar coordonatele articulațiilor detectate cu succes sau estimate de către dispozitiv.

**Obținere arbore de decizie**

Valorile atributelor obținute în cadrul submodulului anterior sunt utilizate în cadrul procesului de construcție a arborelui de decizie după un algoritm de bază ID3 cu extinderi pentru atributele numerice. În fiecare nod din arbore se realizează o partiționare în funcție de unul din atributele calculate la pasul precedent, cu excepția cazurilor în care lista atributelor după care se realizează partiționarea devine vidă sau toate datele de test din nod corespund aceleeași clase. Un astfel de nod reprezintă o frunză în cadrul arborelui de decizie.

**Interogarea arborelui de decizie**

Pentru obținerea unei poziții rezultat, se parcurge arborele de decizie corespunzător, îndeplinind condițiile fiecărei muchii, până la întâlnirea unei frunze. În funcție de arborele interogat, clasele rezultate pot fi:

❖ arborele caracteristic al trunchiului :

∘ drept

∘ aplecat lateral

∘ aplecat frontal

∘ așezat

∘ culcat

❖ arborii caracteristici (2) ai membrelor superioare :

∘ lângă corp

∘ ridicat sus

∘ ridicat lateral

∘ ridicat frontal

❖ arborii caracteristici (2) ai membrelor inferioare :

∘ drept

∘ ridicat frontal

∘ ridicat lateral

∘ îndoit

**3.2 Modulul de detecție a mișcării**

Modulul de detecție a mișcării folosește poziția curent detectată de către modulul de recunoaștere a pozitiei corpului. Din această poziție generală a corpului se obțin pozițiile particularizate ale fiecărei componente a corpului în parte: trunchi, membru superior st\ng, membru superior drept, membru inferior stâng, membru inferior drept. Lanțul Markov modifică starea curentă în funcție de tranziția efectuată. Tranziția este determinată de noua poziție a trunchiului sa membrului corespunzător lanțului. Dupa efectuarea acestei tranziții, se actualizează atât ponderea ce caracterizează tranziția efectuată, cât și celelalte tranziții de ieșire disjuncte ale stării precedente.

Arhitectura modulului de detecție a mișcării este ilustrată în Fig. 8:

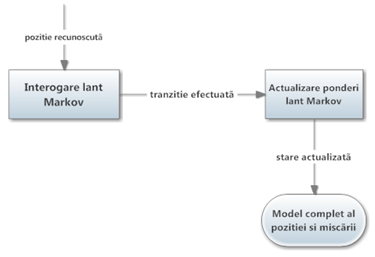


Fig. 8 – Arhitectura modulului de detecție a mișcării

**Interogare lanț Markov**

Submodulul de interogare a lanțului Markov folosește ca dată de intrare poziția recunoscută în cadrul modulului precedent. În momentul în care poziția curentă a cel puțin uneia din părțile componente ale corpului uman s-a modificat, se realizează o tranziție din starea curentă într-o nouă stare incidentă. Actualizarea acestei stări depinde de ultima poziție recunoscută. Dacă noua stare este finală, se obține rezultatul mișcării detectate.

**Actualizare ponderi**

Indiferent de tipul noii stări (stare inițială, stare intermediară, stare finală), tranziția activată își actualizează ponderea – probabilitatea de apariție a evenimentului tranziției crește, în timp ce probabilitățile de apariție a evenimentelor disjuncte acestei tranziții scad.

Datorită faptului că probabilitatea de apariție a zgomotului în cadrul datelor de testare a mișcării este mare, tranzițiile nu vor fi efectuate de la un cadru înregistrat la altul. Soluția acestei probleme vizează o determinare a poziției următoare în funcție de ultimele cadre. Frecvența cu care se realizează o tranziție este determinată empiric. Poziția următoare va fi considerată poziția cu rata de apariție cea mai mare în cadrul ultimelor cadre.