

(12) **CERERE DE BREVET DE INVENȚIE**

(21) Nr. cerere: **a 2022 00724**

(22) Data de depozit: **11/11/2022**

(41) Data publicării cererii:
30/08/2023 BOPI nr. **8/2023**

(71) Solicitant:
• **ASOCIAȚIA PENTRU DEZVOLTARE
URBANĂ, STR.ȘTIRBEI VODĂ, NR.158,
BL.23A, AP.7, SECTOR 1, BUCUREȘTI, B,
RO**

(72) Inventatori:
• **PATZELT DAN, STR.ȘTIRBEI VODĂ,
NR.158, BL.23A, AP.7, SECTOR 1,
BUCUREȘTI, B, RO**

(74) Mandatar:
**CABINET DE PROPRIETATE
INDUSTRIALĂ TUDOR ICLĂNZAN,
PIAȚA VICTORIEI NR.5, SC.D, AP.2,
TIMIȘOARA, TM**

(54) **METODĂ PENTRU GENERAREA DE REPREZENTĂRI
ALTERNATIVE ALE UNEI IMAGINI SURSĂ CE SUNT
ADAPTATE PERSOANELOR CU DIZABILITĂȚI SENZORIALE**

(57) Rezumat:

Invenția se referă la o metodă pentru generarea de reprezentări alternative ale unei imagini sursă ce sunt adaptate persoanelor cu deficiențe de vedere care urmăresc o scenă sau un detaliu al acesteia și în care persoană cu deficiențe de vedere dorește interpretarea și reprezentarea acesteia conform unui context particular. Metoda conform invenției cuprinde etapele de:

- achiziție a unei imagini digitale,
- segmentare și clasificare a imaginii achiziționate sub forma unei hărți cu zone ce delimitează obiectele din imagine și care au etichetate atribute descriptive sau etichete enumerative,
- construire a unui graf de relații pe baza vecinătății obiectelor și
- determinare a unui context de interogare al imaginii, precum și:
- transformare a grafului de relații într-o reprezentare alternativă în care un obiect din imaginea sursă este ignorat sau este transformat într-un obiect alternativ care poate fi: un mesaj text și/sau un mesaj audio și/sau un mesaj gestual și/sau un mesaj de feedback haptic și/sau un obiect suprapus peste o parte din imaginea inițială și/sau un obiect grafic poziționat într-o imagine tactilă și în care obiectele alternative sunt generate automat din graful de relații și sunt asamblate

conform contextului determinat într-o reprezentare alternativă ce poate fi o imagine tactilă și/sau o imagine augmentată și/sau un mesaj audio, textual sau haptic.

Revendicări: 5
Figuri: 8

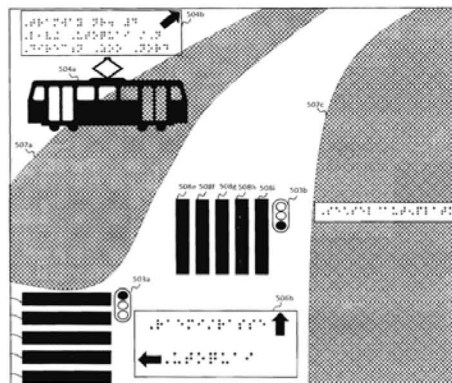


Fig. 5

Cu începere de la data publicării cererii de brevet, cererea asigură, în mod provizoriu, solicitantului, protecția conferită potrivit dispozițiilor art.32 din Legea nr.64/1991, cu excepția cazurilor în care cererea de brevet de invenție a fost respinsă, retrasă sau considerată ca fiind retrasă. Întinderea protecției conferite de cererea de brevet de invenție este determinată de revendicările conținute în cererea publicată în conformitate cu art.23 alin.(1) - (3).



87

**METODĂ PENTRU GENERAREA DE REPREZENTĂRI
ALTERNATIVE ALE UNEI IMAGINI SURSĂ CE SUNT ADAPTATE
PERSOANELOR CU DIZABILITĂȚI SENZORIALE**

Domeniul invenției este cel al transformării imaginilor într-un format alternativ adecvat persoanelor cu dizabilități senzoriale.

Așa cum este cunoscut din stadiul tehnicii, imaginile tactile sunt reprezentări schematice de obiecte și ansamble ale acestora ce sunt tipărite pe o suprafață netedă utilizând o cerneală vâscoasă care atunci când se usucă formează un relief ce poate fi pipăit ușor de o persoană. Aceste imagini ușor texturate în relief sunt o modalitate adecvată pentru învățarea nevăzătorilor despre obiectele pe care aceștia nu le pot vedea dar care îi înconjoară și pe care cu ajutorul simțului tactil le pot pipăii și despre care pot afla informații descriptive sau pot fi ghidați contextual de o persoană de sprijin alăturată sau cu ajutorul unui dispozitiv de supraveghere prevăzut cu cameră digitală și cu o aplicație software ce urmărește degetele nevăzătorilor amplasate pe imaginea tactilă, determină zona de interes atinsă de aceștia și oferă mesaje audio asociate zonei vizate din imagine. Un efect similar poate fi obținut și prin tipărirea pe imaginea tactilă a unor descrieri în limbaj Braille ce sunt asociate zonelor imaginii tactile dar care sunt statice, limitate ca dimensiune și nu oferă o experiență interactivă făcând învățarea dificilă.

Este cunoscută aplicația “Tactile Images” [TactileImages], o aplicație WEB accesibilă la adresa <https://editor.tactileimages.org> și care permite:

- Crearea și/sau alegerea de imagini tactile prin desenare pe calculator și care sunt imagini monocromatice cu linii și zone hașurate și,
- Crearea unui identificator unic al imaginii și o reprezentare vizuală a acestuia de tipul unui cod QR ce este suprapus imaginii tactile într-o parte

OFICIUL DE STAT PENTRU INVENȚII ȘI MĂRCI
Cerere de brevet de invenție
Nr. a 2022 00724
Data depozit 11-11-2022

86

liberă a acesteia și care poate fi citit prin scanare de o aplicație software ce permite identificarea imaginii pe baza codului și,

- Editarea unei hărți cu zone de interes suprapusă imaginii și asocierea de informații textuale sau audio asociate ce descriu zonele și care sunt redată sub o formă alternativă, preferabil sub forma unui mesaj audio, atunci când un utilizator le indică și,
- Salvarea imaginii digitale împreună cu meta datele descriptive și codul unic într-o bibliotecă de imagini sau într-un fișier numeric pe calculator pentru o utilizare ulterioară.

Avantajul aplicației menționate este că permite înlocuirea persoanei de sprijin printr-o aplicație software și deci permite paralelizarea procesului de învățare pentru un număr mare de persoane acesta fiind automatizat. Totuși, crearea imaginilor tactile deși este simplificată ea rămâne manuală.

Conform unui aspect cunoscut, imaginile tactile sunt simplificări statice limitate la câteva obiecte reprezentate schematic și de o complexitate redusă ce sunt tipărite în relief pe un suport neted. Reprezentarea este specifică unui context definit static iar descrierea, dacă este imprimată este la rândul său statică și dacă este reprezentată sub forma unui mesaj audio asociat poate fi dinamică adică diferită în funcție de contextul utilizatorului dat care este creat o dată cu crearea imaginilor tactile și în unele cazuri nu corespunde cu dorinței utilizatorului. Spre exemplu, un utilizator aflat într-un muzeu poate dori informații despre căile de acces spre un obiect de artă sau poate dori informații artistice sau istorice despre același obiect însă imaginea tactilă este optimizată doar pentru unul din scopurile enumerate mai sus. Constatăm că, în mod ideal, imaginea tactilă și descrierea asociată ar trebui să fie adaptate contextului dinamic al utilizatorului și mai precis utilizatorul ar trebui să primească cel puțin un mesaj diferit și preferabil și o imagine tactilă diferită în funcție de ce dorește să afle ca informații.

85

Crearea imaginilor tactile și obținerea de informații contextuale necesită editori umani care să creeze imaginile tactile și să adauge informațiile descriptive adecvate contextului de interes vizat. Numărul limitat de editori precum și dimensiunea lumii înconjurătoare face ca această strategie să fie prohibitivă din punct de vedere al resurselor chiar și dacă este simplificată printr-o aplicație software de editare.

Conform unui alt aspect dorit, în mod ideal, un utilizator dorește să obțină informații contextuale conform interesului său despre orice imagine sau obiect precum ar fi imaginea din locația în care se află. Similar, orice instituție publică dorește să poată genera și oferi imagini tactile despre obiectivele de interes locale precum și alte informații contextuale și care să fie generate preferabil în mod automat.

Remarcăm că soluția propusă de aplicația "Tactile Images" [TactileImages] nu permite generarea automată de imagini tactile din imagini reale și nici descrierea conținutului acestora conform unui context al utilizatorului decât cel fixat la crearea lor și creat prin editare.

Tot din stadiul tehnicii sunt cunoscute rețelele neuronale și, în particular și spre exemplu rețelele de convoluție CNN adânci, rețelele neuronale Recurente RNN, rețelele neuronale cu memorie scurtă și lungă de tip LSTM, rețelele neuronale de tip Transformer și care sunt utilizate și pentru segmentarea și clasificarea imaginilor digitale în obiecte caracterizate prin proprietăți și în care, într-o etapă de segmentare, rețeaua de convoluție generează o hartă a imaginii reale analizate ce cuprinde zonele de delimitare ale obiectelor din imagine și în care, într-o etapă de clasificare ulterioară, imaginea fiecărui obiect din fiecare zonă segmentată este analizată de cel puțin o rețea neuronală de clasificare și care generează cel puțin o etichetă descriptivă a obiectului analizat. Spre exemplu, pot fi recunoscute nu doar granițele obiectului sau tipul acestuia dar și informații specifice precum culoare, orientarea, informații specifice etc.

gh

Tot din stadiul tehnicii este cunoscut domeniul subtitrării automate al imaginilor precum modelele firmei Microsoft VIVO [VIVO21] și OSCAR [OSCAR20] în care obiectele dintr-o imagine sunt identificate prin segmentare și clasificate cu etichete ce sunt folosite ca vectori de intrare într-o rețea neuronală de predicție și care generează descrieri textuale ale imaginii date.

Remarcăm că soluțiile de subtitrare a imaginilor amintite nu oferă posibilitatea interpretării imaginii diferit din punct de vedere contextual deși o aceeași imagine ar trebui să poată fii reprezentată și descrisă în mod diferit în funcție de locație sau de nevoile și interesului utilizatorului și care formează un context de interpretare.

Tot din stadiul tehnicii sunt cunoscute domeniile recunoașterii vocale, al sintezei vocale și cel al traducerii automate de text folosind rețele neuronale și care permit generarea unui text dintr-un mesaj audio, generarea unui mesaj audio vocal ce reprezintă un text dat sau care traduc un text dintr-o limba sau limbaj într-o altă limbă sau limbaj.

Remarcăm că deși sunt folosite, tehnologiile amintite nu sunt obiectul prezentei invenții ele fiind cunoscute și evidente pentru o persoană antrenată în domeniul aplicațiilor rețelelor neuronale numerice.

Problema tehnică a invenției este realizarea unui metode pentru generarea unei reprezentări alternative ale unei imagini și care este adaptată unui utilizator ce suferă de o dizabilitate senzorială dată și care sunt adaptate contextului de utilizare, metoda fiind configurată pentru:

- *Achiziția* unei imagini în format digital și,
- *Segmentarea semantică imaginii* și crearea unei hărți cu zonele ce delimitează obiecte constituate și,
- *Clasificarea* obiectelor din zonele delimitate prin etichete și atribute numerice și,



- *Construirea unui graf relațional* între obiectele identificate prin zonele de delimitare din harta de segmentare și claselor ce caracterizare și,
- *Determinarea unui context de interogare* și,
- *Transformarea grafului relațional* luând în considerare contextul de interogare într-o reprezentare alternativă asociată imaginii și contextului ce poate fii redată într-o etapă ulterioară,

metoda fiind caracterizată prin aceea că:

- graful relațional de obiecte este reprezentat numeric prin lista obiectelor clasificate împreună cu relațiile lui sau o reprezentare derivată din acestea și,
- contextul de interogare:
 - este compus din una sau mai multe părți și,
 - o parte a contextului de interogare este:
 - o locație geografică și/sau,
 - o dată în timp și/sau,
 - o zonă din imaginea achiziționată și care este preferabil zona unui obiect segmentat din imagine sau o arie circulară centrată pe un punct de urmărire numită și zonă de atenți și care cuprinde un număr de obiecte segmentate și/sau,
 - un cuvânt reprezentat numeric prin indexul dintr-un dicționar și care identifică o activitate sau interes al utilizatorului și/sau,
 - un cuvânt sinonim, o generalizare a acestuia sau o negare a unui antonim reprezentate tot printr-un index și un atribut boolean și/sau,
 - o întrebare scurtă a utilizatorului reprezentată prin secvența ordonată a indecșilor numerici ai cuvintelor ce o formează,

care au fost preferabil normalizați în clase sinonime, antonime negate și/sau generalizări, secvență ce conține și semnul ortografic de final ce este indexul unui tip enumerat și care poate avea una din valorile enumerării și/sau,

- o sumă de control de dimensiune fixă derivată din una din reprezentările de mai sus și care este preferabil obținută printr-o rețea neuronală de transformare de tip Encoder sau rețea neuronală RNN și,
- etapa de transformare:
 - este compusă din cel puțin o etapă de transformare simplă și care folosește ca intrare graful relațional, cel puțin o parte a contextului de interogare și rezultatul obținut de la o altă etapă de transformare conform compunerii și,

Alte avantaje sunt prezentate și vor deveni evidente din descrierea detaliată a invenției.

Se dă în continuare un exemplu de realizare a invenției în care oferim o descriere literară a funcționării unei implementări posibile ce trebuie înțeleasă larg și în care opțiunile de implementare alese nu sunt limitări ci exemple posibile și în legătură cu figurile care reprezintă:

Figura 1 ilustrează un exemplu o imagine fotografică capturată de o cameră fotografică digitală peste care este suprapusă o hartă de segmentare semi transparentă în care fiecare zonă colorată diferit delimitează un obiect sau o zonă a imaginii ce corespunde unei clasificări.

Figurile 2A-E ilustrează harta de segmentare din figura 1 în care sunt identificate și evidențiate obiecte și zone clasificate diferite.

Figura 3 prezintă graful relațiilor de proximitate dintre un obiect identificat și clasificat din figura 1 și care este vecin cu alte obiecte.

Figura 4 ilustrează un exemplu de reprezentare augmentată a imaginii din figura 1 adaptată pentru a fi percepută de o persoană cu deficiență de focalizare vizuală.

Figura 5 redă un exemplu de imagine tactilă generată automat ce corespunde imaginii din figura 1 și care o prezintă conform unui context.

Figura 6 este un exemplu de imagine tactilă reprezentând un elefant și un om în care liniile sunt tipărite în relief iar elefantul și omul cu texturi ușor de pipăit și diferite.

Figura 7 prezintă pașii metodei pentru generarea de informații contextuale adaptate persoanelor cu dizabilități.

Figura 8 prezintă un exemplu de filtrare al unui graf de relații de proximitate conform unui context de "navigare".

Este de la sine înțeles că desenele și descrierea detaliată ce urmează a fi prezentate în continuare sunt oferite ca exemple preferabile ele nelimitând spiritul invenției și trebuind interpretate larg. De asemenea este evident că, pentru o persoană antrenată în domeniul Inteligenței Artificiale, alte exemple și variante de implementări alternative sunt evidente și sunt ușor de identificat și extrapolat în spiritul invenției.

Prezenta invenție este ilustrată și descrisă în continuare. Aceleași referințe alfanumerice sunt utilizate pentru identificarea aceluiași element în diversele ilustrații. Referințele cu același prefix numeric urmat de un caracter sunt variante ale unei aceleiași entități.

Figura 1 ilustrează printr-o imagine alb-negru folosind texturi de dithering diferite pentru o imagine reală (100) din setul de date "The CityScapes dataset"

80

[CityScapes] și care reprezintă o intersecție din orașul Zurich din Elveția, imaginea reală este segmentată în zone ce delimitează părți ale obiectelor constituate și care sunt suprapuse peste aceasta. Pentru fiecare zonă sunt cunoscute obiectul aparținător, clasa de caracterizare adică tipul lui, și vecinii săi. Segmentarea imaginii este semantică în sensul că mai multe zone din imagine pot aparține unui același obiect cum este spre exemplu un vehicul poziționat în spatele unui copac sau un cărucior amplasat în spatele unui stâlp și care le separă în două zone care chiar dacă nu sunt adiacente totuși aparțin aceluiași obiect.

Figurile 2A-E detaliază pentru imaginea din figura 1 obiectele constituate. Fiecare obiect este identificat printr-un identificator alfanumeric în care prima cifră este numărul figurii. Instanțele de obiecte de același tip/clasă sunt identificate prin același identificator numeric și au sufixe diferite identificare printr-o literă.

Figura 2A prezintă zone ale imaginii segmentate din figura 1 în care sunt evidențiate semafoare (202a), (202b), panouri de circulație (202a)(202b)(202c), mai multe automobile (201a)(201b)(201f) și respectiv zonele (201c-e) ce ilustrează părți ale unui același automobil care este parțial acoperit în prim plan de o persoană.

Figura 2B prezintă zone ale imaginii segmentate din figura 1 în care sunt evidențiate un tramvai (204) ce are două pantografe (205a) (205b), două străzi (206a)(206b), două pavaje pietonale identificate prin zonele (207a-b)(207c-f) și diferite zone cu semne carosabile (208a-h) din care două delimitări pietonale (208a)(208d), o zebra formată marcaje pietonale (208e-h) și două marcaje de sens (208b)(208c). Remarcăm că deși zonele cu semne carosabile (208a-h) sunt similare fiind zone ce delimitează linii albe, ele au o însemnătate diferită dată de geometria și locația lor relativă la străzi și la semafoarele identificate în figura 2.

Figura 2C prezintă zone ale imaginii segmentate din figura 1 în care sunt evidențiate mai multe persoane (209a-z), două cărucioare cu copii delimitate de zonele (210a) și respectiv (210b-c) și, zone ale cerului (211a-d).

Figura 2E ilustrează zone ale imaginii segmentate din figura 1 în care sunt evidențiate zonele ce delimitează stâlpi (212a-w), un bec de iluminare stradală (213), ghivece ornamentale cu flori atașate stâlpilor (214a)(214b) și mai multe steaguri (215a-c).

Figura 2F ilustrează zone ale imaginii segmentate din figura 1 în care sunt evidențiate zone de coronament verde (216a-m), fațade ale unor clădiri (217a-j) acestea fiind separate în zone disjuncte fiind parțial acoperite de alte zone ce se află în prim plan.

Conform un unui aspect cunoscut din stadiul tehnicii sunt cunoscute metodele de segmentare și etichetare [Shervin2020] a imaginilor digitale bazate preponderent pe utilizarea rețelelor neuronale de convoluție adânci CNN precum modelele AlexNet, VGG-16, GoogLeNet, ResNet și care sunt folosite spre exemplu într-o arhitectură de tip encoder-decodor și care pot fi utilizate pentru segmentarea imaginilor în obiecte caracterizate prin proprietăți.

Conform un unui aspect cunoscut din stadiul tehnicii, segmentarea imaginilor este preferabil o segmentare semantică în sensul că obiectele obstrucționate de alte obiect aflate în prim plan și care sunt separate în imagine în mai multe părți disjuncte sunt fuzionate împreună formând un obiect al cărui arie este interpolată din cea a părților. În acest caz regiunile hărții de segmentare nu mai sunt disjuncte ci suprapuse în mai multe planuri în care obiecte diferite aflate în planuri diferite pot ocupa părți comune dar suprapuse ale imaginii.

Figura 3 prezintă o parte a grafului relațional (300) al obiectelor identificate în figurile 2A-E prin zonele enumerate în prealabil și care prezintă obiectul (301af) ce este un autoturism marca BMW și care este acoperit de zonele (201a)

și (201f) fiind caracterizat prin atribute de conținut (i.e. detalii ale conținutului zonelor) ce sunt obținute prin clasificarea zonelor (201a) și respectiv (201f) în conjuncție și, atribute de locație (i.e. poziționarea obiectului în scena imaginii) și care are relații de proximitate directă (306c), (306d), (306e), (307c), (308d), (309f), (309g), (309h), (309i), (309j), (310c), (312p), (312k), (316g) și (316h) cu zonele vecine (206c), (206d), (206e), (207c), (208d), (209f), (209g), (209h), (209i), (209j), (210c), (212p), (212k), (216g), (216h) cu care partajează o graniță într-o direcție specificată.

Conform unui aspect al acestei invenții graficul relațional (300) este direct obținut din imaginea segmentată și clasificată.

Figura 4 exemplifică o imagine augmentată (400) a imaginii reale din figura 1 în care zonele de segmentare nu sunt ilustrate iar semafoarele (203a) și (203b) din figura 2 sunt augmentate de o suprafață circulară (403a) și (403b) ce are culoarea semaforului respectiv dar pe o suprafață mărită a imaginii permițând unei persoane ce are o dizabilitate vizuală de focalizare sau de iluminare să vadă culoarea semaforului chiar dacă, pentru ea, nu ar fii posibil să distingă culoarea acestuia în cazul real când dimensiunea și depărtarea semaforului sau condițiile de iluminare nu i-ar permite aceasta.

Conform unui aspect și obiectiv al acestei invenții se dorește generarea unor imagini augmentate de tipul cele din figura 4 ce pot fii afișate pe un dispozitiv mobil cu cameră și ecran sau ochelari cu cameră și vedere augmentată plecând de la o imagine capturată de acesta, segmentată și clasificată precum cea ilustrată în figura 1 și care este reprezentată printr-un graf relațional precum cel din figura 3.

Figura 5 ilustrează o imagine tactilă (500) ce abstractizare a imaginii din figura 1 conform unui context utilizator și care poate fii citită de o persoană cu dizabilități de vedere prin pipăire. În imagine, strada (206b) din figura 2 și care

conform locației este la intersecție cu strada „Utoquai” este reprezentată prin masajul în Braille (506b) și care conține denumirea străzilor și sensul de deplasare. Tot în imagine, pavajele pietonale identificare prin zonele (207a-b) și (207c-f) din figura 2B sunt abstractizate prin zonele hașurate (507a) și respectiv (507c) în care zona hașurată (507c) conține și un mesaj în Braille care precizează locația și anume piața „Sechseläuten platz”. În imagine, delimitarea pietonală (208a) și marcajele pietonale (208e-h) din figura 2B precum și semafoarele (203a) și respectiv (203b) din figura 2A și care sunt în vecinătatea lor sunt abstractizate prin zonele de zebra cu linii (508a1-a5) și respectiv (508e-i) și respectiv ideogramele de semafor (503a) și respectiv (503b) care sunt configurate conform cu starea semafoarelor (203a) și respectiv (203b) din figura 2A. Tot în figură, tramvaiul (204) ilustrat în figura 2B este reprezentat prin ideograma (504a) și mesajul Braille (504b) asociat care indică detalii despre tramvai și direcția acestuia de deplasare.

Conform unui aspect și obiectiv al acestei invenții se dorește generarea unor imagini tactile de tipul celei din figura 5 ce pot fi imprimate în relief pe un suport fix.

Conform unui aspect al aceste invenții, imaginea tactilă (500) este obținută prin transformarea grafului relațional (300) al imaginii segmentate (100) ilustrate în figura 1 și care este transformat conform unui context de interes al utilizatorului. Astfel, o aceeași imagine (100) poate fi abstractizată printr-o imagine tactilă (500) în mod diferit și specific unui context fix al utilizator. Spre exemplu, imaginea tactilă (500) din în figura 5 este o hartă de „navigare” sau „deplasare” scopul utilizatorului fiind acela de a „naviga” sau de a se „deplasarea”.

Conform unui aspect al aceste invenții, imaginea tactilă (500) are și o componentă dinamică determinată de utilizator atunci când citește harta și care, spre exemplu indică ideograma (504a) asociată tramvaiului (204) din figura 2A.

26

Spre exemplu, mesajul în Braille (504b) poate fii înlocuit printr-un mesaj audio și care în funcție de context poate fii diferit.

Figura 6a exemplifică imagina tactilă (600) a unui elefant ce poate fii utilizată în scop pedagogic. Aceiași figură este reprezentată augmentat în figura 6b așa cum este ea editată cu ajutorul unui program de editare a imaginilor tactile [TactileImages] și care conține un identificator (601) sub forma unui cod de bare și un număr de zone de interes (602a-h) create de editor și pentru care este precizat un mesaj de descriere audio asociat care, atunci când utilizatorul indică o parte a imaginii (600) este redat în mod dinamic de către un sistem informatic adaptat interpretării ei conform contextului utilizatorului și anume, spre exemplu, conform cu poziția indexului arătător pe imagine.

Figura 7 ilustrează conform unei implementări preferabile pașii unei metode automate de generare a unei reprezentări alternative a unei imagini reale adaptată pentru a fii interpretată de o persoană cu dizabilități senzoriale conform unui context.

Conform metodei ilustrate în figura 7, pasul de achiziție al imaginii (702) este configurat să capteze o imagine digitală a unei scene ce conține obiecte precum cea ilustrată în figura 1.

Conform metodei ilustrate în figura 7, pasul de segmentare și clasificare (703) analizează imaginea pixel cu pixel și creează o hartă cu zone obiectelor din imagine precum și un set de proprietăți sau atribute asociate fiecăruia.

Conform metodei ilustrate în figura 7, pasul de construire a grafului de relații (704) reprezintă harta obiectelor segmentate printr-un graf al vecinătăților în care obiectele din zone adiacente sunt vecine și sunt conectate direct și formează un graf descriptiv al imaginii cum este cel ilustrat în figura 3.

Așa cum o persoană antrenată în domeniul segmentării semantice a imaginilor poate remarca, pașii de achiziție (702) segmentare-clasificare (703)

45

și cel de construire al grafului de relații (704) sunt cunoscuți și, deși sunt necesari pentru realizarea metodei, ei nu sunt considerați noi sau inventivi fiind folosite implementări existente evidente.

Tot conform metodei ilustrate în figura 7, pasul de determinare al contextului (705) determină un context e interpretare al grafului de relații obținut din pasul de construire a grafului de relații (704) și care, spre exemplu, este cel din figura 3 și în care obiectul (301af) este central în imagine.

Conform unui aspect al acestei invenții și într-o implementare de bază, contextul determinat în pasul de determinare al contextului (705) cuprinde:

- *atribute* noi ce decorează graful de obiecte cu informații specifice contextuale și/sau,
- *coeficienți de importanță* ce au valori subunitare pozitive și care asociază obiectelor și/sau atributelor acestora și care precizează un grad de importanță al obiectului și/sau al atributului asociat conform contextului ales și în care valoarea extremă zero precizează că atributul sau obiectul nu este reprezentativ pentru contextul iar valoare unitară precizează că obiectul sau atributul acestuia are importanță majoră.

Conform unui aspect al acestei invenții și într-o implementare de bază, coeficienți asociază obiectelor și atributelor din contextul determinat au valoarea zero când obiectul sau atributul nu este important în contextul dat sau valoarea unu când obiectul sau atributul asociat este important în contextul determinat.

Conform unui aspect al acestei invenții și într-o implementare de bază, coeficienții de importanță sunt precizați într-un tabel predefinit și în care pentru fiecare clasă de obiect și atribut și pentru fiecare context este precizată o valoare de zero sau unu.

74

Conform unui aspect al acestei invenții și într-o implementare de bază, contextul determinat nu cuprinde atribute noi ci doar ansamblul de coeficienți asociați obiectelor și care, spre exemplu, într-un caz sunt zero cu excepția:

- obiectului din centrul imaginii sau,
- a regiunii centrale din imagine formată din obiectul central împreună cu vecinii săi sau,
- a unei părți din graful relațional și care conform contextului este importantă.

Spre exemplu, pentru un context de “navigație”, cu excepția obiectelor din graful relațional care sunt căi de acces, semafoare, marcaje pietonale sau semne de orientare, toate celelalte obiecte sunt asociate unor coeficienți nuli.

Conform unui aspect al acestei invenții și într-o implementare preferabilă ilustrată spre exemplu în figura 8, în pasul de augmentare-filtrare (706) din figura 7 contextul determinat în pasul de determinare al contextului (705) și care este spre exemplu un context de “navigare” ales de utilizator ce marchează ca importante doar obiectele ce sunt vecine semaforului (801a) care este obiectul central din imaginea achiziționată și care, conform tabelului de importanță predefinit pentru contextul de “navigare” determină că doar obiectele (801a), (802a), (805a), (807a-809a) și (812a) ce sunt semafoare, marcaje pietonale, pavaje pietonale, indicatoare rutiere, drumuri, sau vehicule ce sunt importante în contextul ales și în consecință au coeficienții de importanță unitari iar restul obiectelor din graful relațional (801a-812a) sunt ignorați având coeficienții de importanță zero.

Tot din figura 8 și conform unui aspect al acestei invenții, pasul de augmentare-filtrare (706) din figura 7 aplicat grafului obiectele (801a-812a) produce un graf nou format din obiectele (801b), (802b), (805b), (807b-809b) și

43

(812b) ce sunt copii ale obiectelor (801a), (802a), (805a), (807a-809a) și (812a) și care erau considerate importante pentru contextul de navigare.

Conform unui aspect al acestei invenții ilustrat în figura 8, în pasul de augmentare-filtrare (706) din figura 7 prin eliminarea stâlpului (804a), semaforul (801b) devine vecin cu indicatorul (805b) ne existând noduri ne conectate. Prin eliminarea unor obiecte ne importante din graf nu există riscul apariției de sub grafuri ne conectate deoarece vecinii nodului ne important ce este eliminat devin automat vecini.

Conform unui aspect al acestei invenții și într-o implementare de bază, în pasul de augmentare-filtrare (706), pe lângă coeficienții de importanță asociați se pot augmenta obiectele grafului cu atribute noi de context precum data și ora curentă, locația geografică a imaginii sau alte atribute contextuale specifice obiectelor, scenei, locației curente sau utilizatorului și care permit interpretarea grafului de relații în mod diferit și specializat. Spre exemplu, dacă poziția curentă unde imaginea este achiziționată este un punct de interes bine cunoscut precum un muzeu atunci un atribut de locație și/sau identificatorul punctului de interes este adăugat fiecărui obiect din graful de relații făcându-l specific contextului.

Conform unui aspect al acestei invenții ilustrat în figura 7, pasul de transformare (707) poate fi compus și cuprinde și sub pași de transformare (707) ce sunt composabili conform unui graf de transformare predefinit.

Conform unui aspect al acestei invenții, poate exista un sub pas de transformare care agregă anumite atributele ale unui obiect și le transformă în alte atribute conform unei funcții de agregare predefinite. Spre exemplu, automobilul (301af) din figura 3 posedă un număr de atribute de locație relativă precum: "Distanța: 12m", "Unghi 37 grade", "Elevație: 0", "Orientare: 100 grade relativ" și care sunt corelate cu poziția GPS și orientarea curentă

72

împreună cu harta locală fiind transformate într-o locație și direcție de deplasare absolute și anume: ”Locația: Intersecția Sechseläuten platz cu strada Utoquai”, ”Orientare: Nord”, ”Direcția: Înainte”. Acest gen de transformare este binecunoscut în stadiul artei și este folosit în aplicațiile de navigare rutieră pentru îndrumarea deplasării pe un drum plecând din locația curentă.

Conform unui aspect al acestei invenții și într-o implementare preferabilă, poate exista un sub pas de transformare (707) ce transformă un obiect împreună cu atributele sale într-o propoziție conform unui șablon predefinit pentru tipul obiectului și în care subiectul este clasa obiectului iar atributele sunt adjective ale acestuia. Spre exemplu, pentru obiectul (301af) din figura 3 ce este un automobil este generată o propoziția pe baza șablonului predefinit de forma: ”Un [automobil] [sedan] cu marca [BMW] de culoare [albastră] cu numărul de înmatriculare [TM 33 BMW] aflat în [Intersecția Sechseläuten platz cu strada Utoquai] se deplasează [înainte] pe direcția [Nord]”. Șablonul este predefinit și ales astfel încât importanța acumulată a atributelor să fie maximă adică se caută șablonul care acoperă un număr maxim de atribute ponderate de importanța lor și care este un coeficient subunitar predefinit. Spre un alt exemplu, semaforul (203a) din figura 2 și care are atributele ”Stare: Roșu”, ”Locație: Intersecția Sechseläuten platz cu strada Utoquai”, ”Direcția: înainte” este transformat conform unui șablon într-o propoziție de forma ”Stai! Nu traversa! Semaforul este pe [Roșu!]”.

Conform unui aspect al acestei invenții și într-o implementare preferabilă, o propoziție textuală poate fi transformată într-un mesaj audio printr-o metodă de sinteză vocală bine-cunoscută din stadiul artei.

Conform unui aspect al acestei invenții și într-o implementare preferabilă, rezultatul transformării este înregistrat în format digital sub forma unei imagini tactile anotate zonal cu mesaje textuale, în braille, mesaje vocale sintetizate sau cu alte anotări suprapuse peste zona obiectului sau obiectelor descrisă.



Conform unui aspect al acestei invenții și într-o implementare preferabilă, contextul poate fii un obiect sau o zonă din imagine ce conține o vecinătate de obiecte.

Conform unui aspect al acestei invenții și într-o implementare preferabilă, un sub pas de transformare (707) transformă graful relațional într-o reprezentare serială în care obiectele sunt ordonate într-o secvență conform produsului dintre importanța lor și al distanței față de obiectul central care, spre exemplu este inversul numărului de legături ce trebuie parcurse pentru a ajunge de la obiectul principal la obiectul curent. Alte metrice de ordonare și distanță sunt posibile și evidente pentru o persoană antrenată în teoria grafurilor.

Conform unui aspect al acestei invenții și într-o implementare preferabilă, reprezentarea serială a grafului relațional este introdusă la intrarea unei rețele neuronale recurente RNN cunoscute din stadiul artei și pre-antrenate pentru generarea de descrieri textuale pornind de la o propoziție sămânță și pentru care coeficienții rețelei sunt fixați la modelul standard preluat din stadiul artei cu excepția celor de pe ultimele două sau trei straturi de neuroni și care sunt ajustați prin antrenare prin exemple de imagini descrise prin grafuri relaționale serialitate fiind folosite aceleași tehnici precum [OSCAR20] și [VIVO21] care nu țineau cont de context.



Glosar de termeni:

Imagini Tactile	Reprezentări schematice amprentare în relief sau printate în relief pe un suport neted și care pot fii citite tactil cu degetele prin pipăire de către persoanele nevăzătoare
CNN	Rețele Neuronale de Convoluție
RNN	Rețele Neuronale Recurente

Referințe:

[TactileImages] <https://tactileimages.org/>

[CityScapes] The CityScapes dataset: <https://www.cityscapes-dataset.com/>

[Shervin2020] „Image Segmentation Using Deep Learning: A Survey” Shervin Minaee, Yuri Boykov, Fatih Porikli, Antonio Plaza, Nasser Kehtarnavaz, Demetri Terzopoulos, „IEEE Trans. Pattern Anal. Mach. Intell”, Vol. 44, Nb. 7, 2022

[VIVO21] „VIVO: Visual Vocabulary Pre-Training for Novel Object Captioning”, Xiaowei Hu, Xi Yin, Kevin Lin, Lijuan Wang, Lei Zhang, Jianfeng Gao, Zicheng Liu, Microsoft Corporation, „Thirty-Third Conference on Innovative Applications of Artificial Intelligence” IAAI 2021

[OSCAR20] „Oscar: Object-Semantics Aligned Pre-training for Vision-Language Tasks”, Xiujun Li, Xi Yin, Chunyuan Li, Pengchuan Zhang, Xiaowei Hu, Lei Zhang, Lijuan Wang, Houdong Hu, Li Dong, Furu Wei, Yejin Choi, Jianfeng Gao, „European Conference for Computer Vision” ECCV 2020.

REVENDICĂRI

1. O metodă pentru generare de reprezentări alternative ale unei imagini sursă ce sunt adaptate persoanelor cu dizabilități senzoriale care urmăresc o scenă sau un detaliu al acesteia și în care persoana cu dizabilități dorește interpretarea și reprezentarea acesteia conform unui context **caracterizată prin aceea că** cuprinde etapele de:

- **Achiziție** a unei imagini digitale și,
- **Segmentare și clasificare** a imagini achiziționate sub forma unei hărți cu zone ce delimitează obiectele din imagine și care au etichetate și atribute descriptive care au valori numerice sau sunt etichete enumerative și,
- **Construire a unui graf de relații** pe baza vecinătăților obiectelor și,
- **Determinare a unui context de interogare** și care cuprinde:
 - un punct sau o zonă din imagine predeterminată sau aleasă de persoana cu dizabilități și/sau,
 - locația geografică a persoanei cu dizabilități și/sau,
 - locația geografică sau punctul de interes al locației imaginii și/sau,
 - orientarea persoanei cu dizabilități relativ la scena urmărită în imagine și/sau,
 - orientarea imaginii și/sau,
 - data și ora când imaginea a fost făcută și/sau,
 - data și ora când persoana cu dizabilități face observația și/sau,
 - metadate ce caracterizează imaginea,

metoda fiind caracterizată prin aceea că mai cuprinde și etapele de:

- **Augmentarea – filtrarea** grafului de relații conform contextului în care graful este:

- simplificat prin eliminarea obiectelor fără interes și claselor și atributelor descriptive ce nu sunt importante pentru contextul determinat și în care obiectele eliminate sunt înlocuite prin legături de interconectare a vecinilor lor și/sau,
- augmentat cu o parte a atributelor de context și,
- **Transformare** ce transformă graful de relații simplificat și anotat într-o reprezentare alternativă adaptată pentru a fi percepută de către persoana cu dizabilități și care este:
 - un mesaj textual și/sau,
 - un mesaj audio sau vocal de sinteză și/sau,
 - un mesaj gestual și/sau,
 - un mesaj de feedback haptic și/sau,
 - un obiect grafic sau ideogramă suprapă peste o parte din imaginea inițială și,

în care reprezentarea alternativă este:

- o imagine tactilă augmentată și/sau,
- o imagine augmentată și/sau,
- un mesaj audio, textual sau haptic.

2. Metodă conform revendicării 1, **caracterizată prin aceea că**:

- **etapa de transformare** este compusă din sub etape de transformare organizate într-un graf aciclic de transformare.

3. Metodă conform revendicării 1, **caracterizată prin aceea că**, pentru o persoană nevăzătoare:

- în etapa de transformare, segmentele imaginii sau grupuri ale acestora sunt înlocuite printr-o ideogramă tactilă standard acompaniată de

GA

meta date descriptive ce sunt mesaje textuale, braille, mesaje audio de sinteză vocală sau ideograme suprapuse imaginii.

4. Metodă conform revendicării 1, **caracterizată prin aceea că**, pentru o persoană cu dificultăți de focalizare,

- în etapa de transformare, segmente sau grupuri de segmente din imagine sunt:
 - mărite sau micșorate în mod intenționat pentru a putea fi vizualizate prin ochelari adaptabili sau ochelari pentru realitatea augmentată și/sau,
 - înlocuite cu de meta datele descriptive obținute din clasificarea segmentelor ce pot fi redate printr-un canal senzorial secundar și,

5. Metodă conform revendicării 1, **caracterizată prin aceea că**, pentru o persoană cu dificultăți de auz:

- în etapa de transformare, segmente sau grupuri de segmente din imagine sunt înlocuite printr-un clip video ce conține un mesaj gestual ce poate fi reprodus printr-un clip video augmentat peste imaginea urmărită.

66

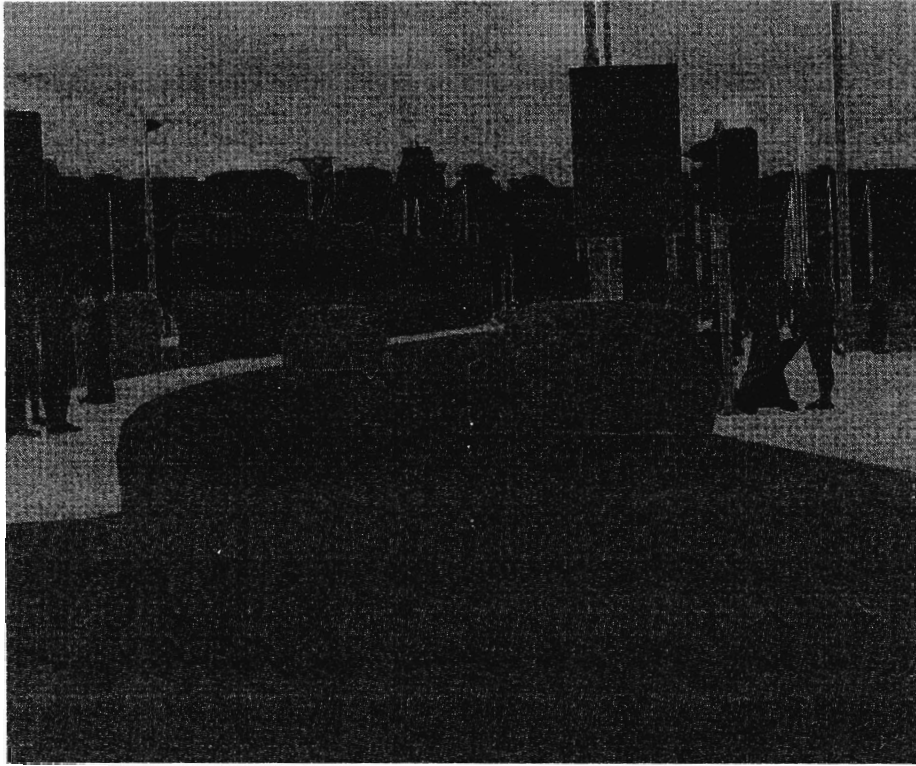


Figura 1

es

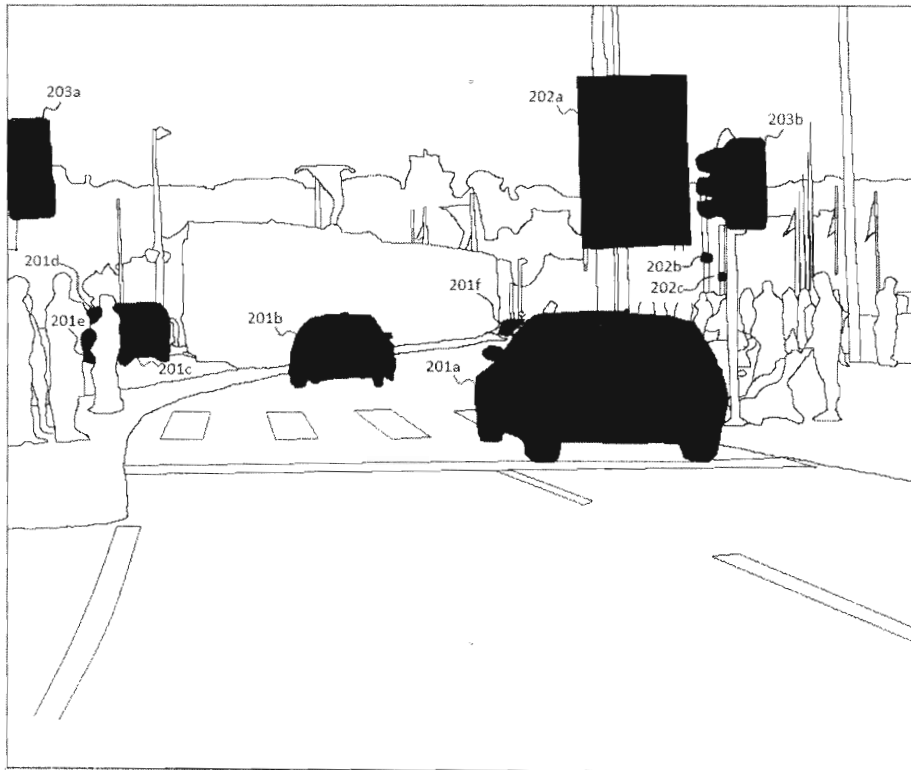


Figura 2A

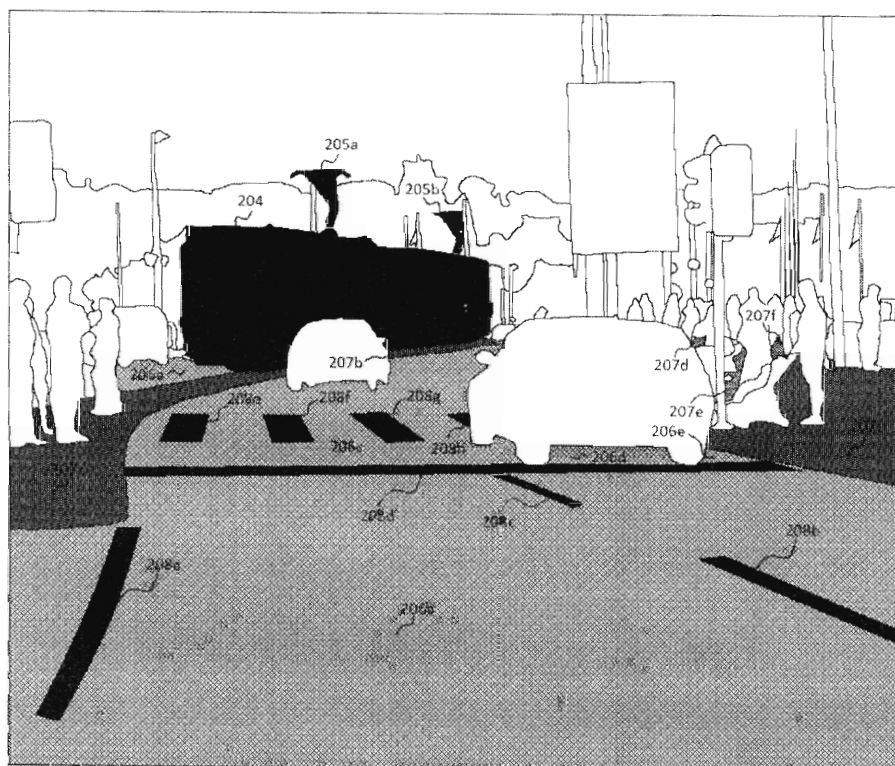


Figura 2B

64

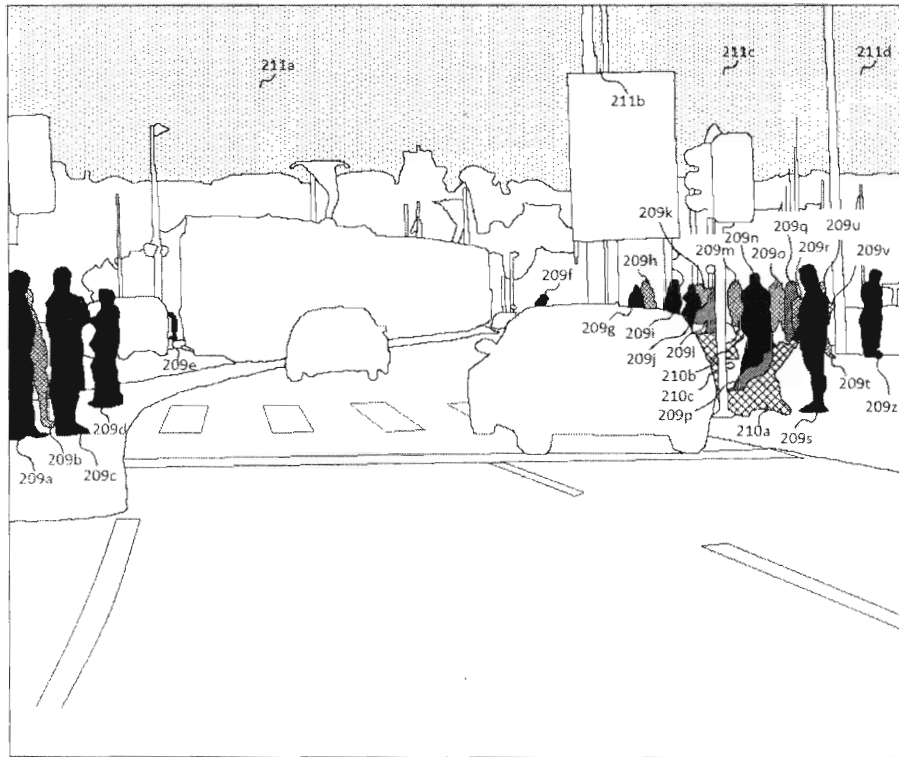


Figura 2C

63

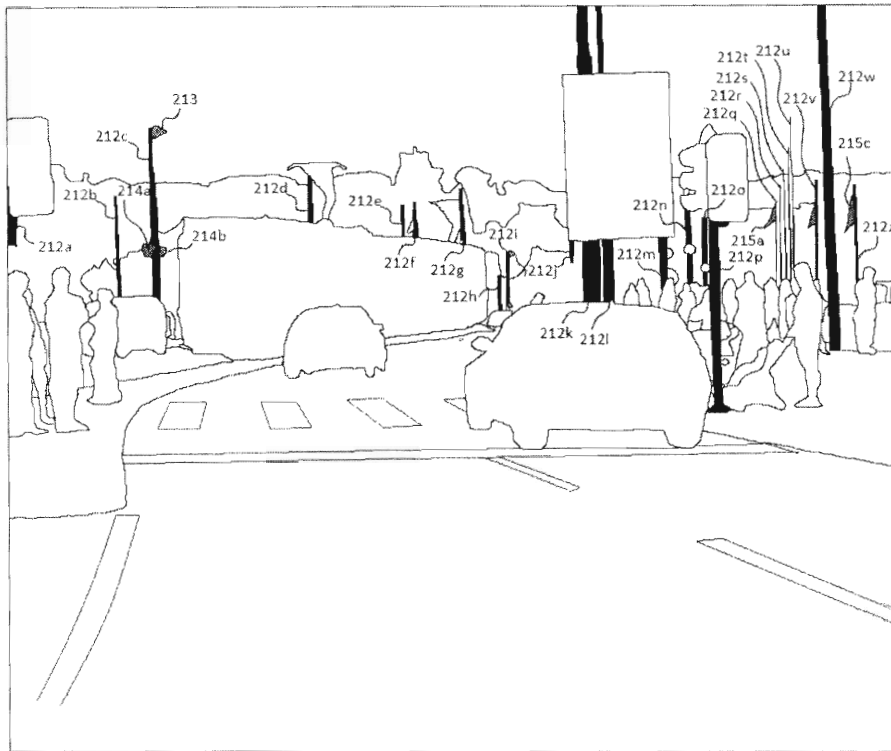


Figura 2D

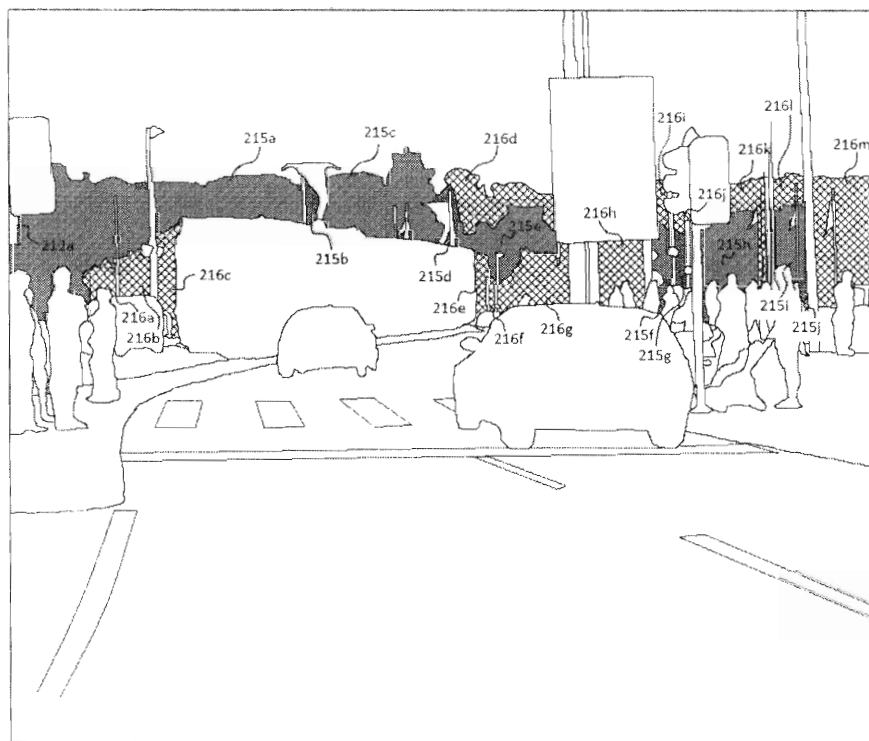


Figura 2E

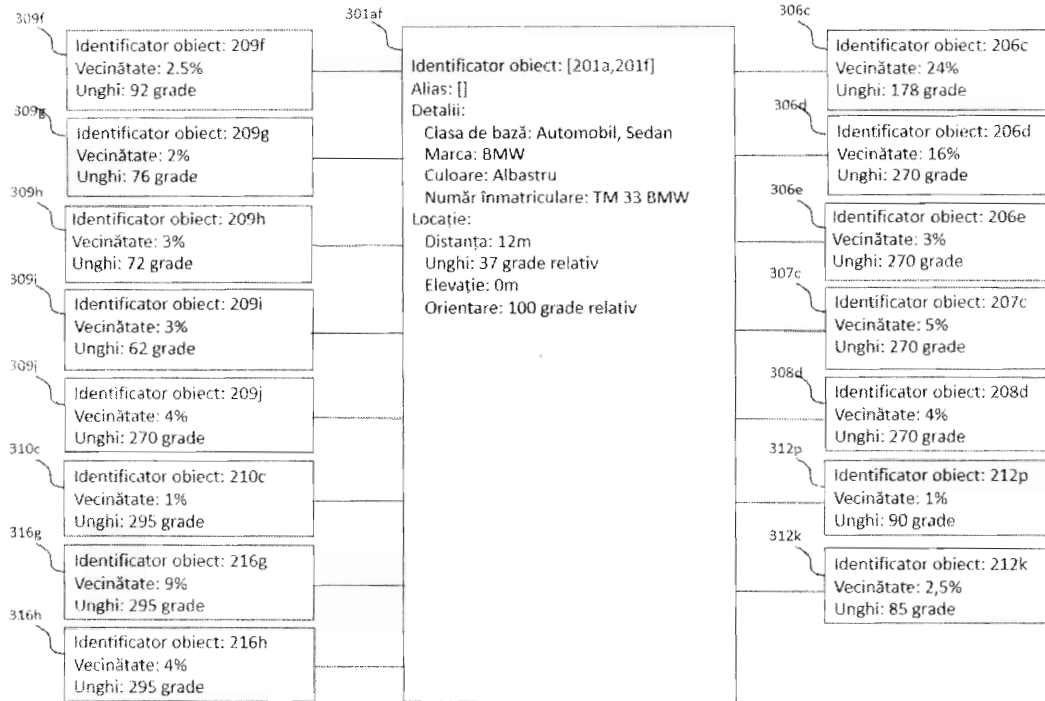


Figura 3

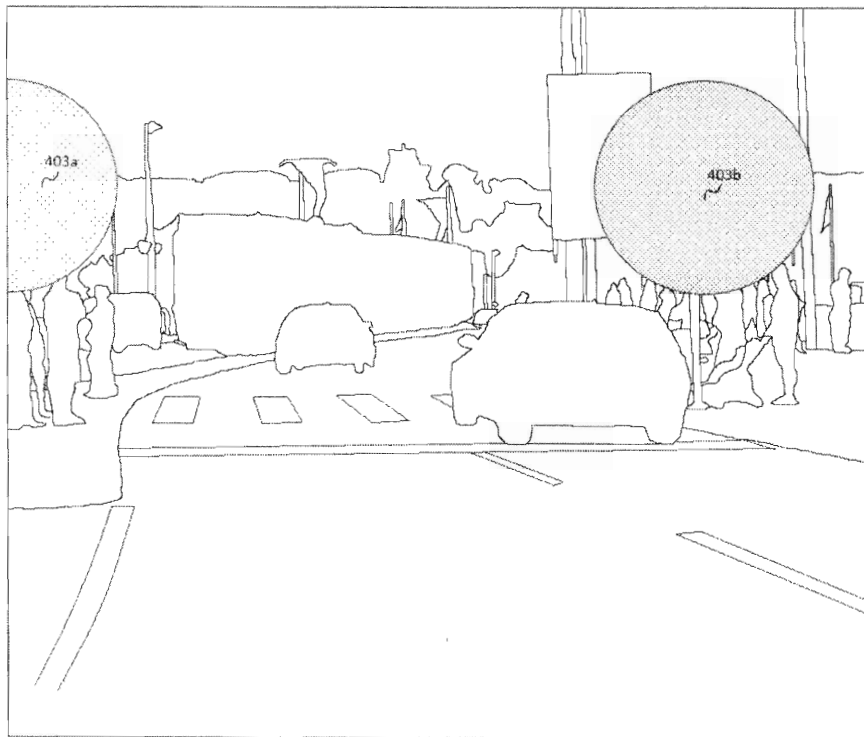


Figura 4

61

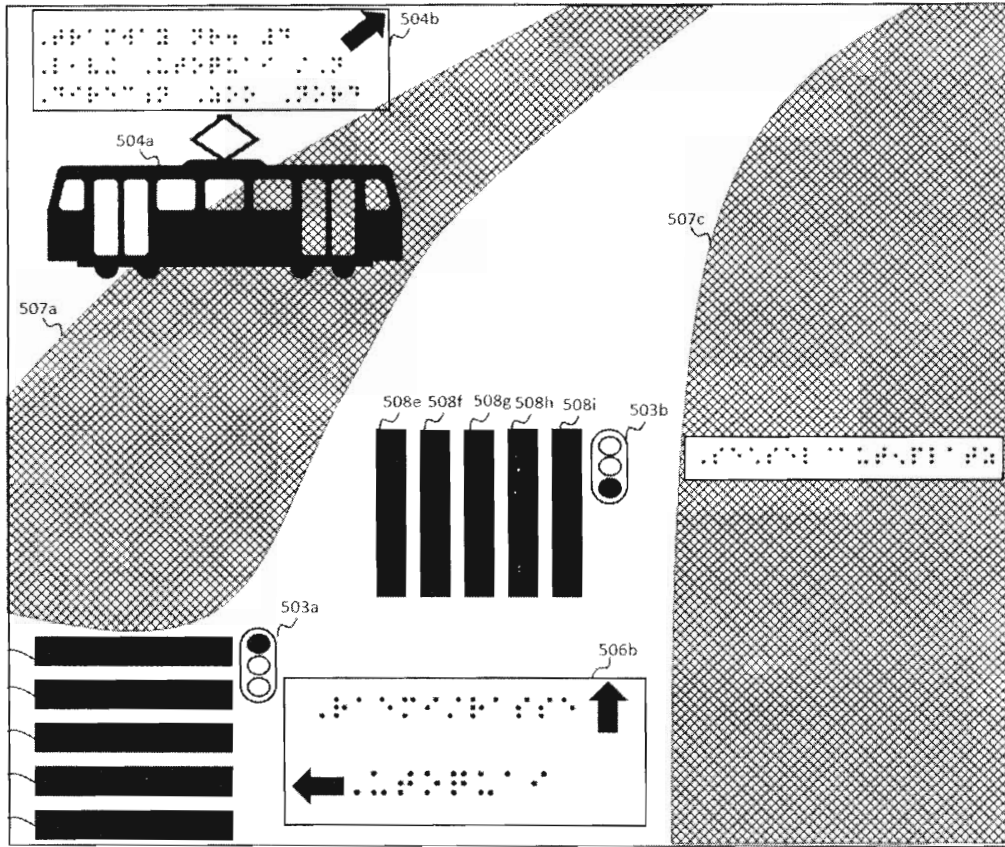


Figura 5

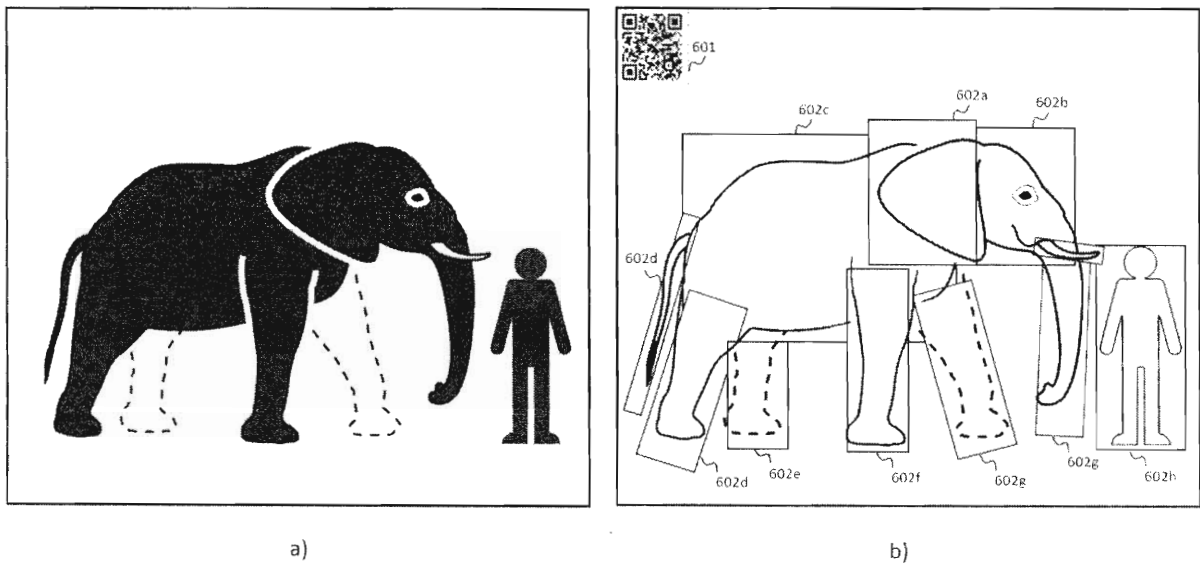


Figura 6

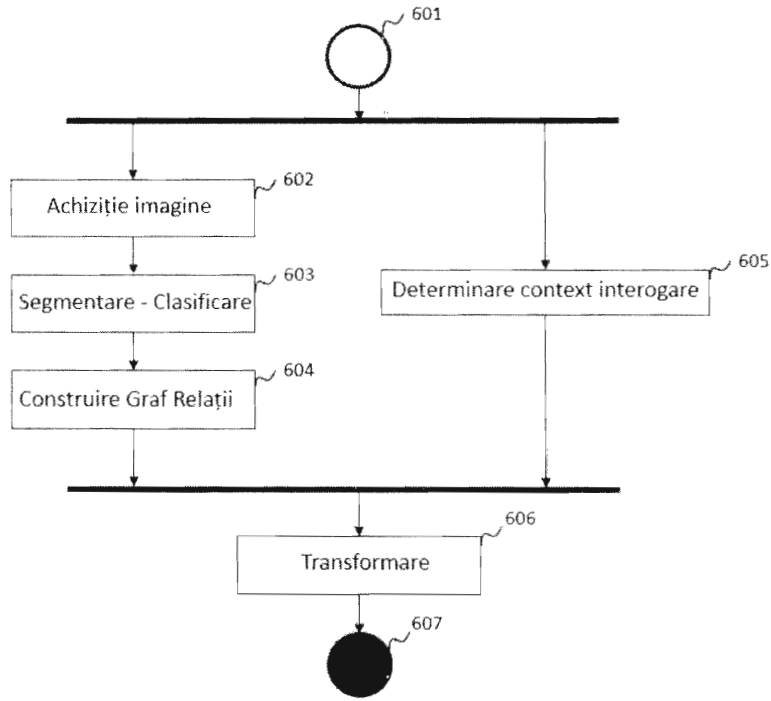


Figura 7

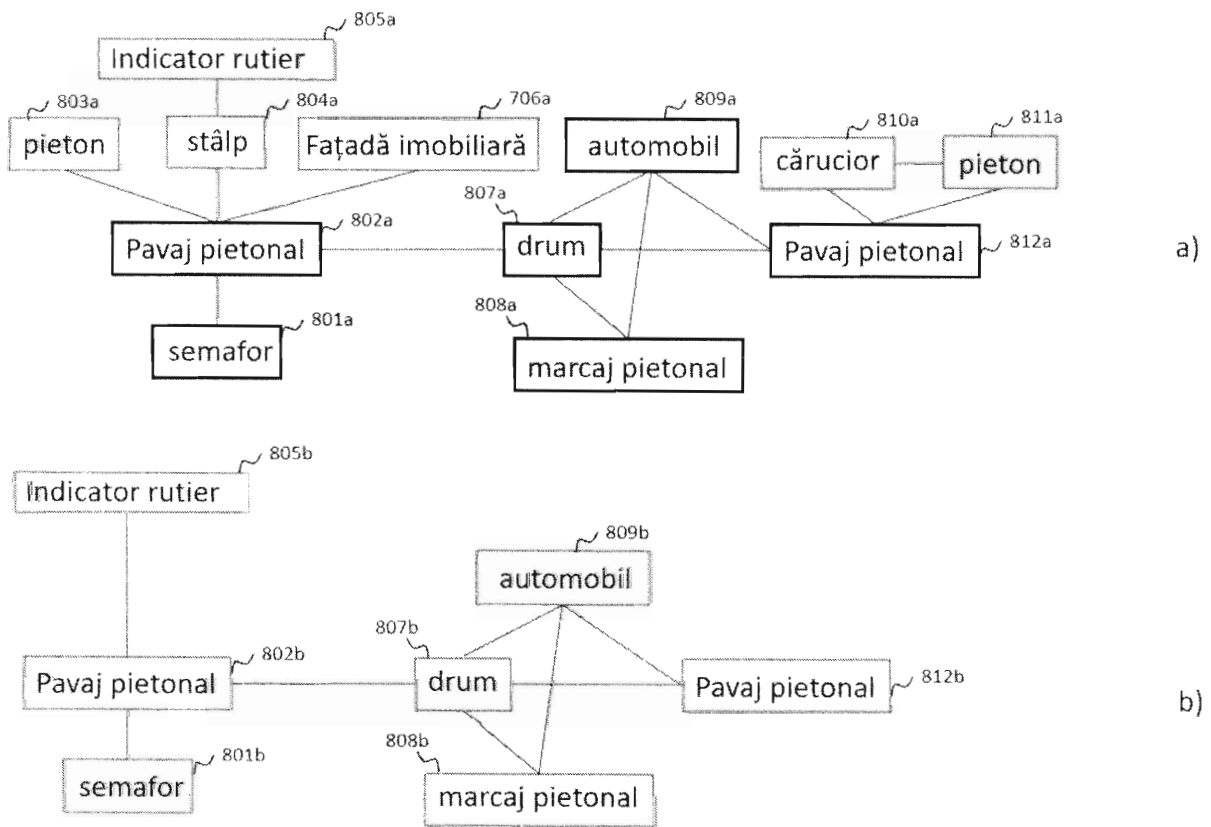


Figura 8