



Universitatea
Ștefan cel Mare
Suceava

ANALYSIS OF SOME ALGORITHMS FOR IMAGE RECOGNITION

dr.ing. Ovidiu-Ciprian UNGUREAN
prof.univ.dr. ing. Ștefan-Gheorghe PENTIUC

Sisteme Distribuite, decembrie 2015

1. OBIECTIVE

2. Identificarea clădirilor sau a obiectivelor turistice folosind tehnici de analiză și prelucrare a imaginilor

3. Algoritmi și metode

4. Experimentare

5. Contribuții

Suport financiar - Proiectul MappingBooks

Referinte



Universitatea
Ștefan cel Mare
Suceava

1. OBIECTIVE

Obiectivul principal

- ❑ Cercetare privind capacitățile de recunoaștere a unor forme de relief sau alte obiective în funcție de poziția și orientarea dispozitivului, cazuri de obstrucționare a perspectivei prin interpunerea de obiecte în imagine.

Obiectiv suplimentar

- ❑ Propunerea de algoritmi, proiectarea și realizarea unui sistem funcțional de recunoaștere a imaginilor eventual utilizând și informații de localizare.

2. IDENTIFICAREA CLĂDIRILOR SAU A OBIECTIVELOR TURISTICE FOLOSIND TEHNICI DE ANALIZĂ ȘI PRELUCRARE A IMAGINILOR

1) Sistemul prototip dezvoltat de SuperWise Technologies AG numit **eye-Phone** (eyePhone, 2008) are rolul de ghid de informații turistice.

Folosește

- localizarea GPS,
- recunoașterea de obiecte și
- informații specifice provenite de la conexiunea de internet

pentru a furniza detalii despre locația curentă a utilizatorului.

- 2) Sistemul **infoScope** (infoScope, 2001) este un sistem de tip AR ce folosește
- informații GPS,
 - realizând recunoașterea de imagini
- cu scopul de a suprapune în timp real detalii despre clădirile sau zonele din imaginea surprinsă de camera video.

3) Sistemul prezentat în (Mai et al., 2003) folosește

- datele de la GPS și
- de la senzorii de orientare

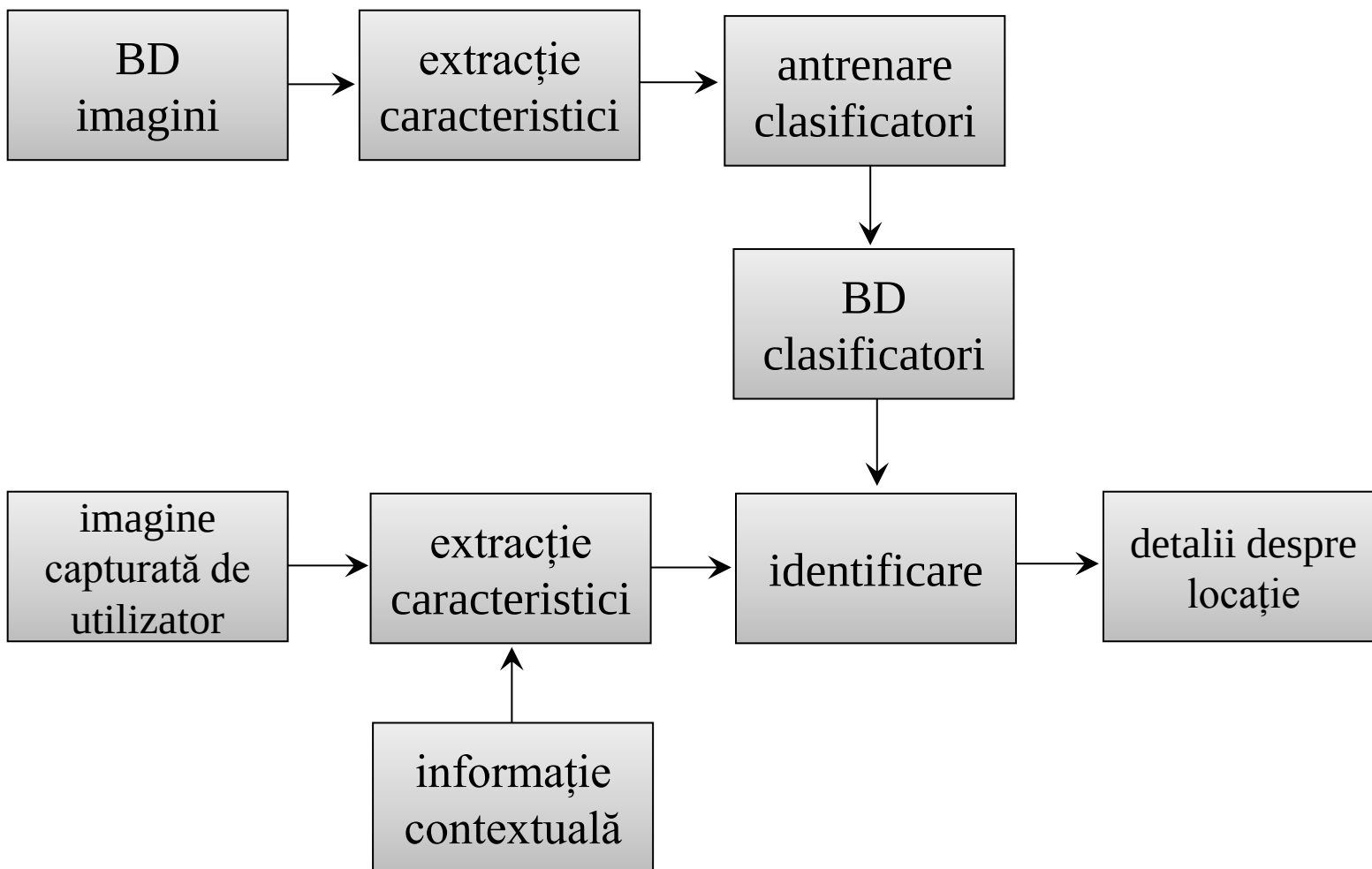
și caută să recunoască o imagine comparând-o cu alte imagini din baza de date folosind un **model 3D** pentru ca rezultatul să fie influențat cât mai puțin de faptul că unghiurile sub care au fost captate imaginile sunt diferite.

4) În sistemul Snap2Tell (Snap2Tell, 2007) detaliile despre

- locația utilizatorului puse la dispoziție de operatorul de telefonie mobilă

sunt folosite pentru a reduce spațiul de căutare în interiorul bazei de date.

Modelul mobil de recunoaștere ambientală (Chen, 2009)





Exemple de imagini folosite în ZuBuD (***Zurich Building Image Database***)

Baza de date Oxford Buildings (Philbin et al., 2007) conține 5,062 imagini culese din Flickr cu cele mai importante 11 obiective din Oxford.

Baza de date a fost adnotată manual fiind marcate 11 din cele mai importante obiective din Oxford. Sunt formate 55 de interogări ce pot fi folosite la evaluarea algoritmilor de identificare a locațiilor.

Fiecărei imagini îi este atribuită câte una din următoarele etichete privind gradul de vizibilitate al obiectivului:

- *Good*: Obiectivul se vede bine și clar în imagine.
- *OK*: Mai mult de 25% din obiectiv/clădire este prezent în imagine.
- *Bad*: obiectul nu este prezent în imagine
- *Junk*: mai puțin de 25% din obiectiv este vizibil sau există un grad foarte mare de zgomot sau suprapuneri ale altor obiecte peste obiectiv.

SCity (SCity, 2009) conține aproximativ 1.2 milioane de imagini

- realizate cu ajutorul unei mașini dotate cu GPS pe străzile orasului Seattle.
- au fost extrase și etichetate 20.000 de imagini pentru a forma *benchmark*-ul. Fiecare 6 imagini succesive sunt grupate ca scene.



Universitatea
Ștefan cel Mare
Suceava

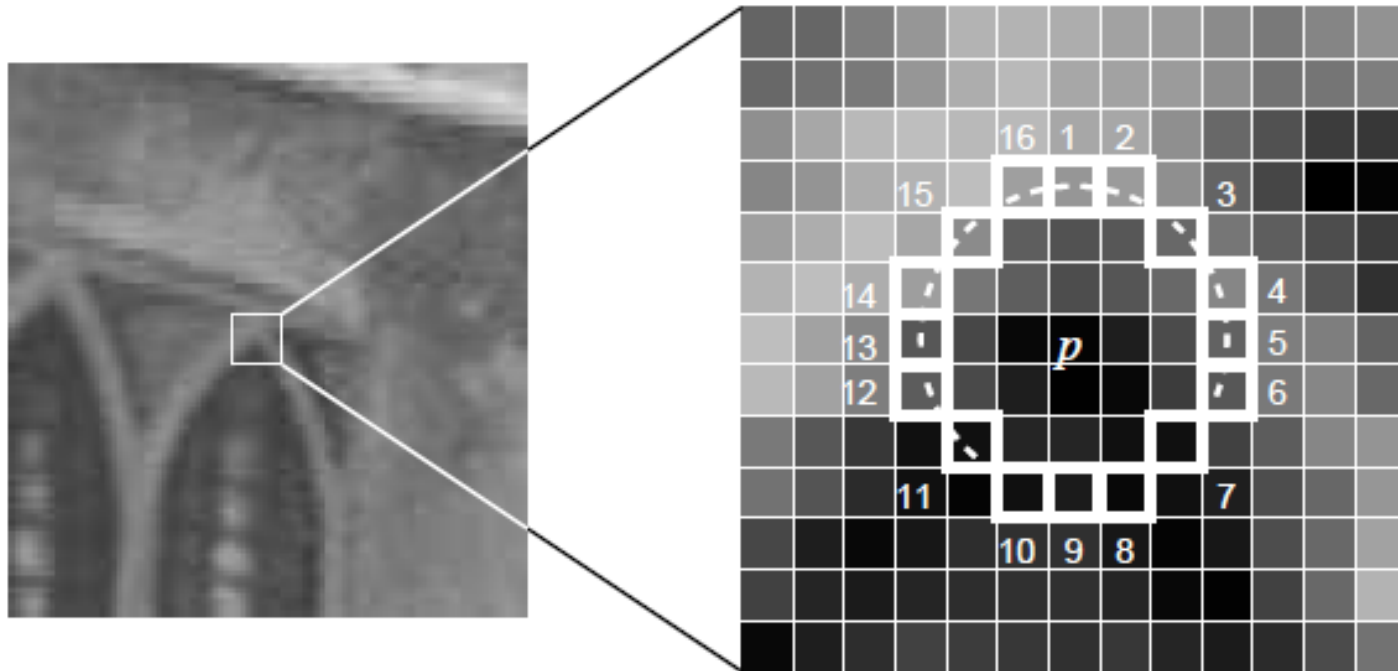
3. ALGORITMI ȘI METODE

- În literatura de specialitate, metoda preponderentă de identificare a descriptorilor este SIFT (Lowe, 1999),
- denumire ce provine de la *Scale-Invariant Feature Transform*,
- Metodă protejată prin US Patent și de aici probleme legate de drepturi de autor și de utilizare a algoritmului.

- Am preferat să folosim metoda **ORB (*Oriented FAST and Rotated BRIEF*)**
- prezentată în lucrarea (Rubblee et al., 2011)
- Această metodă utilizează o variantă îmbunătățită a descriptorului BRIEF (Calonder et al., 2010) și a metodei de identificare a punctelor cheie FAST .
- În cele ce urmează se va prezenta pe scurt metoda ORB precum și BRIEF pentru a se justifica decizia luată.

- **se construiește o bază de date cu imagini cunoscute;** imaginile sunt analizate, și pentru fiecare imagine se identifică un set de **puncte cheie**, împreună cu **descriptorii** asociați acestora;
- **imaginea de recunoscut este analizată în mod similar:** se identifică punctele cheie și se generează descriptorii asociați acestora;
- **se caută în baza de date imaginea care are cel mai mare număr de puncte cheie similare** cu imaginea de recunoscut; similaritatea a două puncte cheie este evaluată cu ajutorul descriptorilor asociați acestora;
- **se compară numărul de puncte cheie similare cu o anumită valoare de prag.** Dacă este mai mare => se consideră că imaginea de recunoscut reprezintă același obiectiv (clădire, vârf de munte etc.) cu imaginea din baza de date.

Identificarea unui punct cheie



Pixelul p este înconjurat de 12 pixeli mai luminoși decât el (diferența mai mare decât o valoare de prag)

Setul de descriptori asociați unei imagini este generat astfel:

- se identifică punctele cheie din imagine; cel mai comun punct cheie este cel reprezentat de un colț (vîrfurile unghiurilor ascuțite);
- se analizează pixelii dintr-o zonă de eșantionare de dimensiuni fixe din jurul fiecărui punct cheie; rezultatul acestei analize, de regulă o combinație de biți de lungime fixă, este descriptorul asociat punctului cheie.

BRIEF - Binary Robust Independent Elementary Features

- Sunt parcurse un număr n de perechi de pixeli ($p1$, $p2$) din acea zonă.
- Dacă intensitatea pixelului $p1$ este mai mare decât cea a pixelului $p2$ se reține 1, altfel 0 pentru respectiva poziție din șirul binar.
- Rezultă un descriptor binar de lungime n , o combinație de n biți ($n = 128, 256, 512 \dots$) care reprezintă zona de eșantionare de dimensiune $S \times S$ din jurul unui punct cheie.

Probleme

- Influența zgomotelor – filtrare gaussiană
- Stabilirea locațiilor perechilor de pixeli ($p1$, $p2$)
- În (Calonder et al., 2010) s-au analizat 5 metode de determinare a vectorilor X și Y (figura 1) din punctul de vedere al influenței lor asupra ratei de recunoaștere:
- GI) X și Y au o distribuție uniformă în zona de eșantionare
-

- GII) X și Y sunt aleși în mod aleator folosind o distribuție de tip gaussian izotropică, centrată în origine care este considerată în centrul zonei de eșantionare;
- GIII) X și Y sunt aleși în mod aleator, mai întâi X_i folosind o distribuție gaussiană centrată în origine și o deviație standard de $0.04 \cdot S^2$, iar apoi sunt aleși y_i cu o distribuție gaussiană centrată în x_i și cu o deviație standard de $0.01 \cdot S^2$;
- GIV) X și Y sunt aleși în mod aleator dintr-un careu grosier în coordonate polare (*coarse polar grid*);
- GV) pentru fiecare i , x_i are valoarea $(0, 0)$, iar y_i primește toate valorile posibile asigurând acoperirea uniformă, sub 30 de unghiuri egale cu 12 grade.

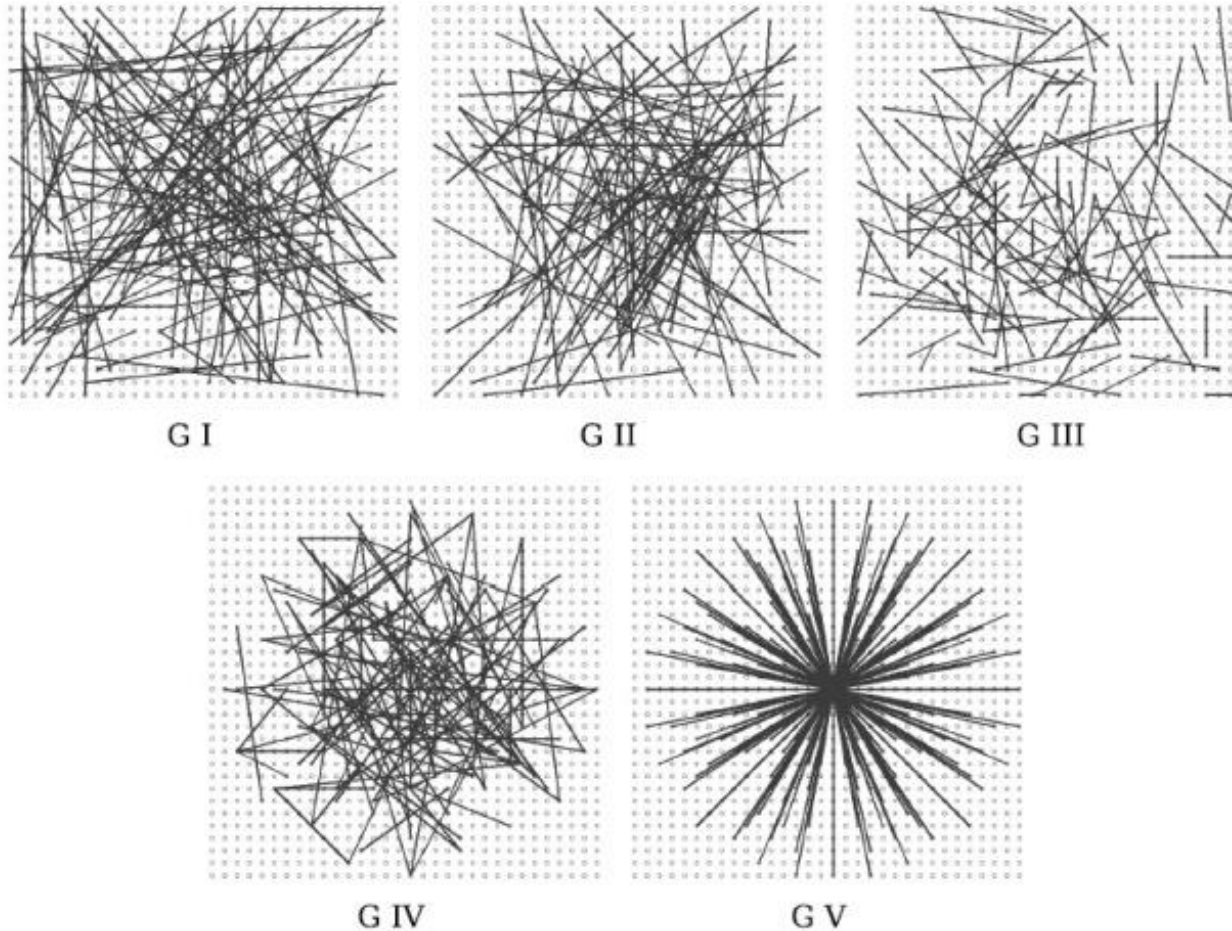


Figura 1. Cele 5 șabloane de eșantionare, conform (Calonder et al., 2010)

Descriptorul ORB (Rublee et al., 2011)

Reprezintă o combinație

1.a unei variante de descriptor BRIEF mai puțin sensibil la rotații în planul imaginii

2.și a metodei de identificare a punctelor cheie FAST prezentate în (Rosten and Drummond, 2006) și (Rosten et al., 2010),

dar îmbunătățită prin adăugarea unei componente rapide și precise de orientare.

Îmbunătățire esențială adusă de ORB

suportul de recunoaștere a descriptorilor în condiții de variație la rotație.

Pentru a determina unghiul descriptorului,

- **se determină centrul de greutate** (centroidul) C al intensității luminoase a pixelilor punctului cheie.
- fie originea sistemului de coordonate O ca fiind chiar centrul de simetrie al zonei, **se determină orientarea vectorului OC** conform (Rosin, 1999).
- Fiind cunoscut acest unghi **zona de eșantionare poate fi rotită spre o direcție canonică**, iar descriptorul asociat punctului cheie va fi calculat astfel încât să devină invariant la rotații în planul XOY .

Pentru a obține o rată de recunoaștere cât mai bună este foarte important ca perechile de eșantionare să îndeplinescă următoarele două condiții:

- lipsa de corelații (*uncorrelation*): se dorește ca perechile de eșantionare să fie cât mai diferite, astfel încât fiecare pereche să aducă un plus de informație rezultatului analizei descriptorilor, maximizându-se astfel cantitatea de informații înglobată de către descriptor;
- - variabilitate mare a perechilor (*high variance of the pairs*) în jurul mediei de 0.5 care face ca o caracteristică să aibă un potențial discriminativ mai mare deoarece va fi cât mai diferită relativ la intrări diferite.

Se aplică un algoritm de tip **greedy** ce constă în următoarele:

1. se execută testul pe toate zonele elementare de eșantionare;
2. se ordonează testele în funcție de distanța față de media 0.5 obținându-se vectorul T ;
3. se execută o căutare de tip **greedy**:
 - primul test este adăugat în vectorul R și scos din vectorul T ;
 - următorul test din T este comparat cu toate testele din R . Dacă corelația absolută a acestuia este mai mare decât un anumit prag, atunci nu este adăugat în R ;
 - pasul anterior este repetat până când lungimea lui R devine 256;
4. Dacă R are mai puțin de 256 de elemente atunci se mărește valoarea pragului de la pasul anterior și se reia execuția pasului 3.

Rezultă un vector cu 256 de perechi necorelate cu o foarte mare variație.

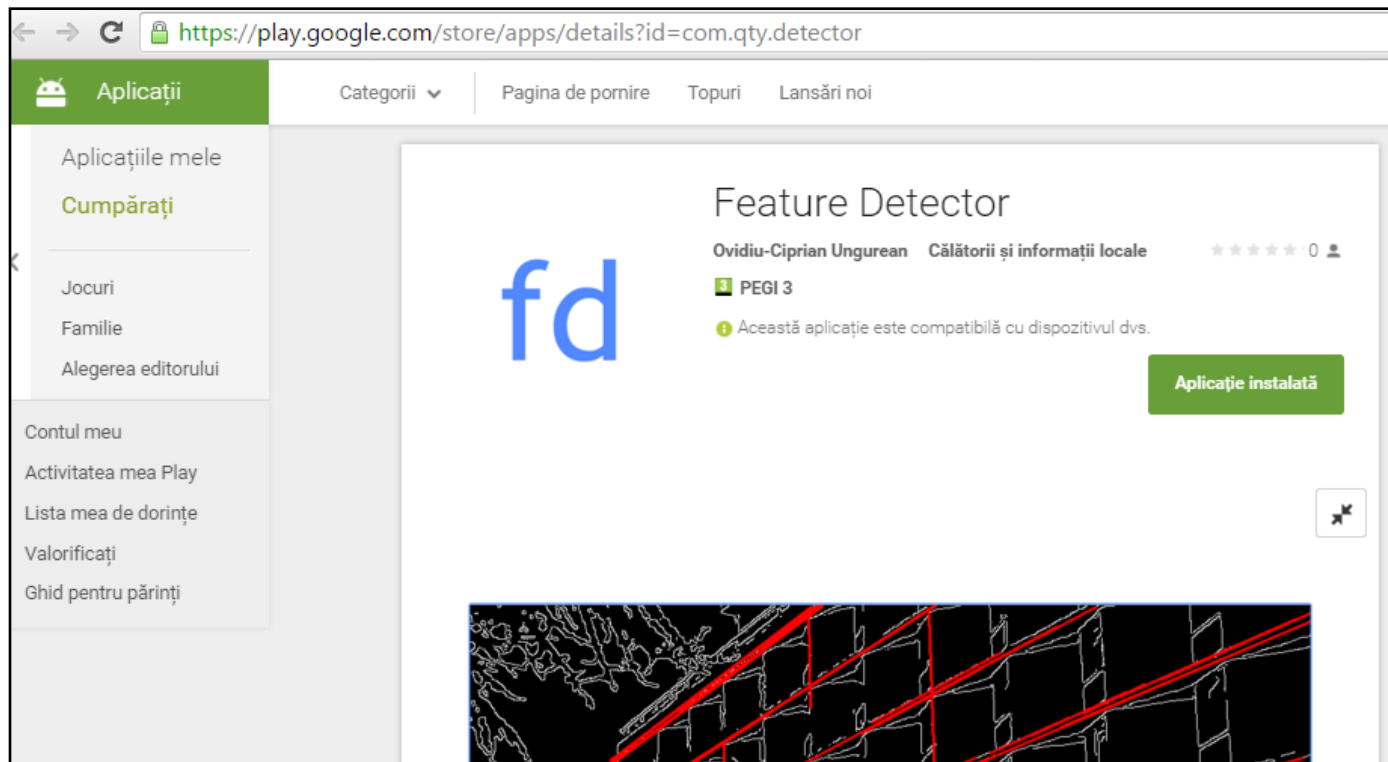
Potrivirea descriptorilor ORB





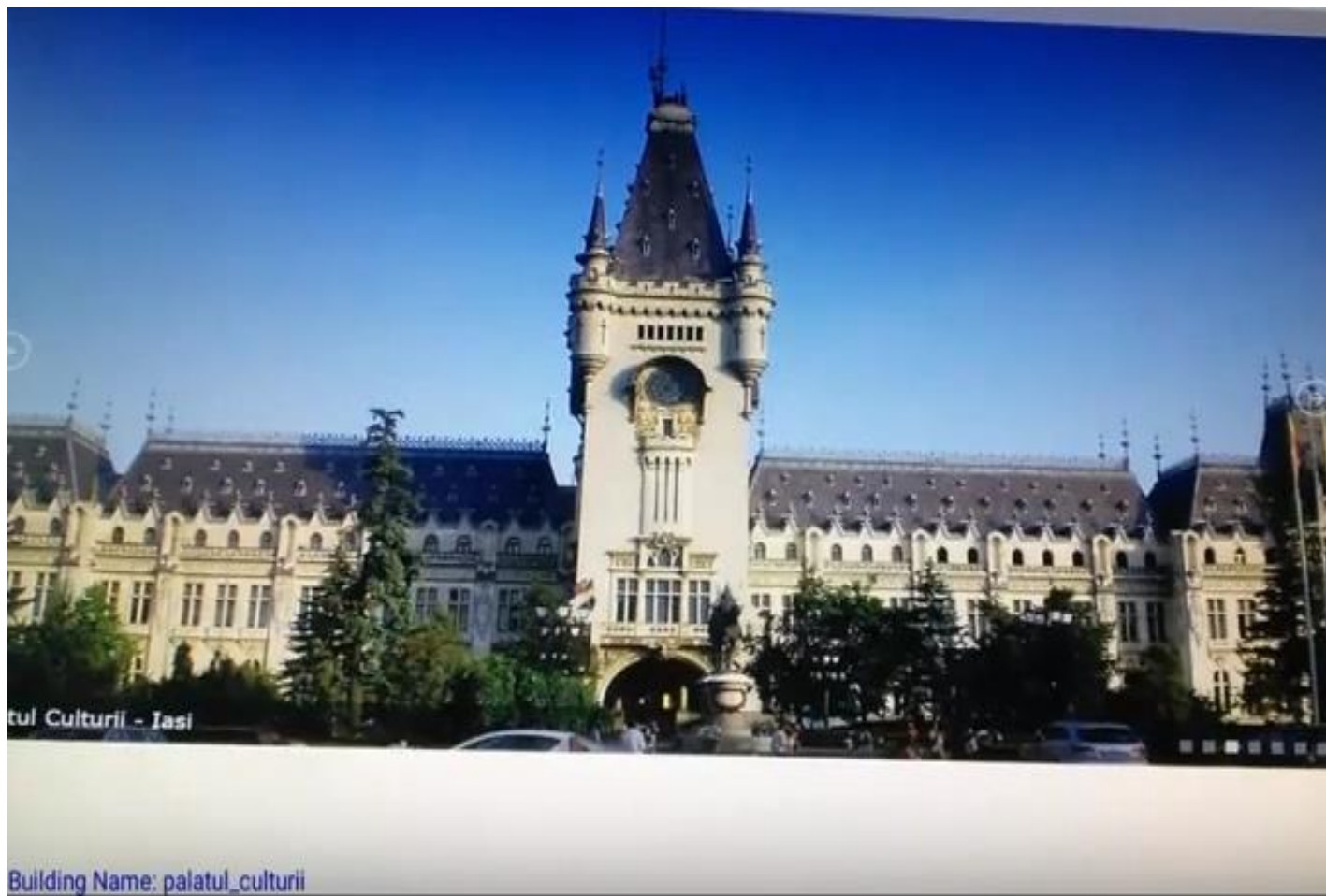
Universitatea
Ștefan cel Mare
Suceava

4. EXPERIMENTARE













Universitatea
Ștefan cel Mare
Suceava

5.CONTRIBUȚII

1. Stabilirea orientării zonei elementare pe baza metodei regresiei liniare
 - descriptorii vor fi supuși la o translație și o rotație înaintea memorării lor în baza de date
2. Metoda de hashing bazată pe clasificarea ierarhică a descriptorilor
 - funcția hash = funcția discriminant
 - creșterea numărului de descriptori determină coborârea cu m nivele în dendrogramă



Universitatea
Ștefan cel Mare
Suceava

SUPPORT FINANCIAR

Rezultatele cercetărilor prezentate în această lucrare au fost obținute în cadrul proiectului

MappingBooks - Intra în carte!

finanțat prin Programul PARTENERIATE -
Proiecte colaborative de cercetare aplicativă –
Competiția 2013 (PCCA 2013),

Consortiul de realizare fiind format din
Universitatea “Alexandru Ioan Cuza” din Iași,
SC SIVECO Romania SA și
Universitatea „Ștefan cel Mare” din Suceava.



Universitatea
Ștefan cel Mare
Suceava

REFERINȚE

(Calonder et al., 2010) Calonder, Michael, et al. "Brief: Binary robust independent elementary features." Computer Vision–ECCV 2010. Springer Berlin Heidelberg, 2010. 778-792.

(Chen 2009) Chen, Tao, et al. "A survey on mobile landmark recognition for information retrieval." Mobile Data Management: Systems, Services and Middleware, 2009. MDM'09. Tenth International Conference on. IEEE, 2009.

(eye-Phone 2008) -

http://www.esa.int/Our_Activities/Space_Engineering_Technology/TTP2/Tourist_information_wherever_you_are

(Infoscope 2001) I. Haritaoglu. "Infoscope: Link from real world to digital information space". în G. Abowd, B. Brumitt, and S. Shafer, editors, Ubicomp 2001: Ubiquitous Computing: Third International Conference Atlanta, Georgia, USA, vol.2201/2001 of Lecture Notes în Computer Science, pp.247–255. Springer-Verlag Heidelberg, September 30 - October 2 2001.

(Lowe 1999) Lowe, D.G., "Object recognition from local scale-invariant features," în Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on , vol.2, no., pp.1150-1157 vol.2, 1999, doi: 10.1109/ICCV.1999.790410

- (Mai et al. 2003) W. Mai, G. Dodds, and C. Tweed, “A pda-based system for recognizing buildings from user-supplied images”, Mobile and Ubiquitous Information Access: Mobile HCI International Workshop, vol.2954, pp. 143–157. Lecture Notes în Computer Science, Springer-Verlag Heidelberg, Sept 2003.
- (OpenCV, 2015) OpenCV Manager - <https://play.google.com/store/apps/details?id=org.opencv.engine&hl=ro>
- (Rosin 1999) Rosin, Paul L. “Measuring corner properties.” Computer Vision and Image Understanding 73.2 (1999): 291-307.
- (Ungurean 2015a) Ungurean Ovidiu-Ciprian, Aplicatia de evaluare Feature Detection (FD) - <https://play.google.com/store/apps/details?id=com.qty.detector>
- (Ungurean 2015b) Ungurean Ovidiu-Ciprian, Clip cu testele efectuate local – <https://www.youtube.com/watch?v=Lw0BkrnMbwc>
- (Ungurean 2015c) Ungurean Ovidiu-Ciprian, Imaginile de antrenare folosite pentru teste - <http://fd.deproba.ro/>
- (VOC 2006) M. Everingham. The PASCAL Visual Object Classes Challenge 2006 (VOC2006) Results. <http://host.robots.ox.ac.uk/pascal/VOC/index.html>
- (ZuBuD 2003) H. Shao, T. Svoboda, and L. Van Gool. 2003. ZuBuD: Zurich Buildings Database for Image Based Recognition. Technical Report.



Universitatea
Ștefan cel Mare
Suceava

MULȚUMESC!!