

Determinarea celor mai similare 3 imagini pe baza descriptorului LBP si a distantei euclidiene

Proiect asignat in cadrul disciplinei Interfete Om-Masina

Coordonator stiintific: -

Student 1: Chisalescu Bogdan

Student 2: Chivu Gheorghe-Iulian

Student 3: Floricel Antonio-Stefan

date of today

1 Scopul lucrarii

Prezenta lucrare reprezinta o activitate de tip proiect din cadrul disciplinei Interfete Om-Masina. Aceasta urmareste dezvoltarea unei aplicatii software ce primeste ca date de intrare o imagine si o locatie in care se afla o multitudine de alte imagini, si returneaza dintre cele din urma 3 imagini ce sunt cat mai similare cu imaginea primita ca data de intrare. De asemenea, este necesar ca aplicatia sa prezinte utilizatorului o interfata grafica dezvoltata folosind API-ul cross-platform wxPython.

2 LBP si distanta euclidiană

2.1 Local binary pattern

In continuare vom descrie metodele LBP ce au fost folosite in actuala lucrare. Toate variantele de LBP prezentate opereaza pe imagini cu nivele de gri, in cazul in care imaginea de prelucrat nu este o imagine exclusiv cu nivele de gri, aceasta va fi convertita.

Local binary pattern sau in forma sa abreviata LBP este o operatie punctuala centrata pe o vecinatate a unei imagini in urma careia se poate forma un descriptor vizual. Aceasta operatie este caracterizata de extragerea unei texturi.

2.2 LBP fundamental

Operatia LBP fundamentala presupune parcurgerea imaginii pixel cu pixel si extragerea unei vecinatati

$$V_8 = \{(1, -1), (0, 1), (1, 1), (0, -1), (0, 0), (0, 1), (-1, -1), (-1, 0), (-1, 1)\}$$

cu coordonate relative fata de valoarea curent prelucrata. Pixelul curent aflat in prelucrare ce va avea in vecinatate coordonatele $(0,0)$ este considerat ca valoare de prag si sunt comparati cu acesta ceilalti pixeli apartinand vecinatatii. In urma comparatiei, daca valoarea din vecinatate este mai mare decat valoarea pixelului central atunci aceasta ia valoarea 1, altfel ii este asignata valoarea 0. In acest fel se creeaza un tipar de 1 si 0. Acest tipar este parcurs in sens orar in vecinatate, formadu-se cu el un numar binar ce mai apoi convertit in numar zecimal ne da valoarea descriptorului pentru acea vecinatate.

Aplicand LBP pe dimensiunea intregii imagini obtinem un descriptor de aceleasi dimensiuni cu imaginea initiala. In prezenta lucrare imaginile sunt parcurse astfel incat valorile marginale ce depasesc limitele imaginii sunt ignorate prin pozitionarea convenabila a punctului de start.

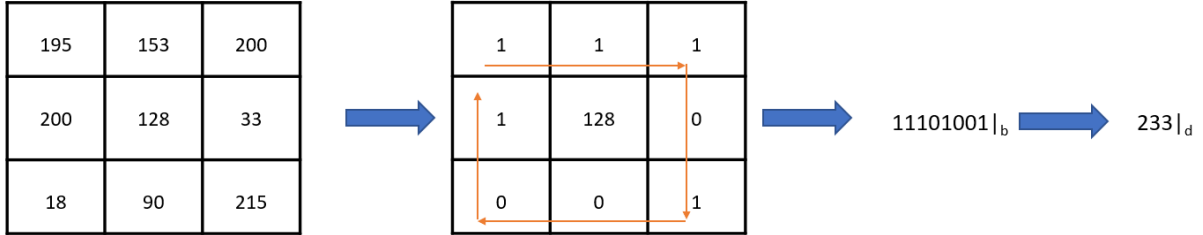


Figure 1: Schema de aplicare LBP fundamental pe vecinatatea V_8

2.3 LBP-mediu

Aceasta forma de LBP este identica cu cea fundamentala cu exceptia faptului ca, valoarea de prag este considerata media tuturor pixelilor ce apartin vecinatatii V_8 .

2.4 LBP circular

LBP-ul circular are la baza acelasi principiu ca si cel fundamental. Consideram imaginea de prelucrat $I(x,y)$ si coordonatele punctului curent relative la imagine x_p si y_p . Vecinatatea circulara este definita de cercul $C(I(x_p, y_p), R)$. Pe acest cerc sunt alese un numar de P puncte echidistante. Coordonatele punctelor sunt calculate dupa urmatoarele formule:

$$x_k = x_p + R \cos\left(\frac{2\pi k}{P}\right), k \in \{0, 1, \dots, P-1\} \quad (1)$$

$$y_k = y_p - R \sin\left(\frac{2\pi k}{P}\right), k \in \{0, 1, \dots, P-1\} \quad (2)$$

In cazul in care x_k si $y_k \notin \mathbb{Z}$ si deci nu se poate obtine direct $I(x_k, y_k)$, aceasta va fi interpolata biliniar.

Consideram un semnal bidimensional discret $f(x, y)$, si 2 indici x_k si y_k pentru care semnalul f nu este definit. Putem interpola biliniar valoarea semnalului f in punctul (x_k, y_k) considerand cele mai apropiate 4 puncte ale semnalului de (x_k, y_k) . Notand indicii acestor puncte considerandu-l pe cel din stanga jos origine, putem afla valoarea semnalului $f(x_k, y_k)$ ca fiind:

$$f(x_k, y_k) \cong f(0, 0)(1 - x)(1 - y) + f(1, 0)x(1 - y) + f(0, 1)(1 - x)y + f(1, 1)xy \quad (3)$$

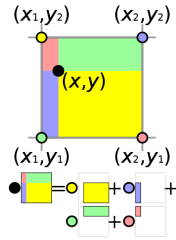


Figure 2: Interpretare vizuala a interpolarii biliniare

Avand la dispozitie valoarea intensitatii imaginii in punctul central $I(x_p, y_p)$ si valorile asociate punctelor echidistante $I(x_k, y_k), k \in \{0, 1, \dots, P - 1\}$ putem determina valoarea returnata de LBP circular pentru punctul (x_p, y_p) ca fiind:

$$LBP_{(P,R)} = \sum_{k=0}^{P-1} S(I(x_k, y_k) - I(x_p, y_p))2^k \quad (4)$$

unde $S(u)$ este definita ca:

$$S(u) = \begin{cases} 1, & u \geq 0 \\ 0, & u < 0 \end{cases} \quad (5)$$

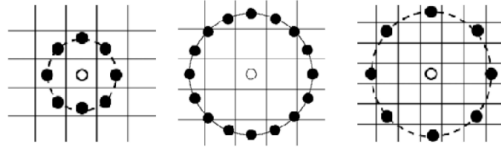


Figure 3: Vecinatatile circulare (8,1), (16,2), respectiv (8,3)

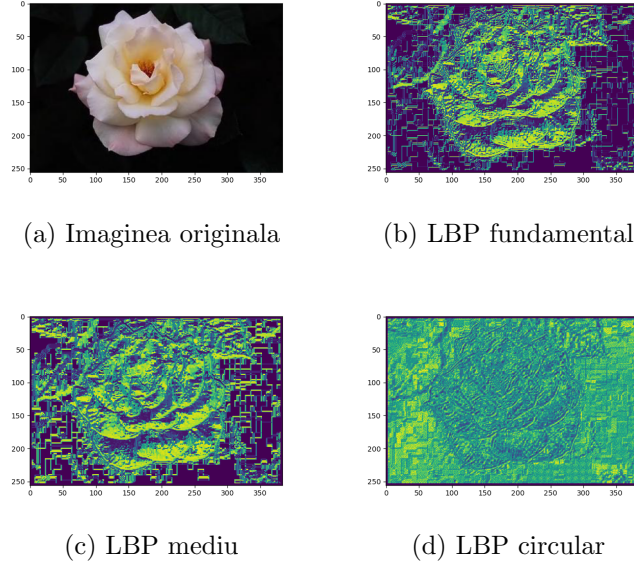


Figure 4: Descriptorii LBP

2.5 Distanța euclidiană

Pentru a putea determina cele mai "apropiate" imagini este necesar să dispunem de un mod prin care să măsurăm distanța între descriptorii lbp; pentru aceasta lucrare a fost aleasă distanța euclidiană. Definim distanța euclidiană pentru două semnale bidimensionale discrete ca fiind $D_e(p, q) = \sqrt{\sum_{i=1}^M \sum_{j=1}^N (p_{ij} - q_{ij})^2}$, important de menționat fiind faptul că este necesar ca cele 2 semnale să aibă aceleași dimensiuni ale suportului. În cazul nostru M și N sunt numărul de linii respectiv de coloane al descriptorilor LBP.

3 Interfața grafică

La proiectarea interfeței grafice au fost luate în considerare noțiunile teoretice cu privire la dezvoltarea interfețelor vizuale, prezentate în cursul 2 al disciplinei Interfețe Om-Mășină. Au fost respectate în acest mod: simetria, echilibrul, regularitatea, secvențialitatea, unitatea, proporționalitatea, și gruparea.

Atunci când este lansată în execuție aplicația, utilizatorul este întâmpinat de primul stadiu al interfeței grafice.

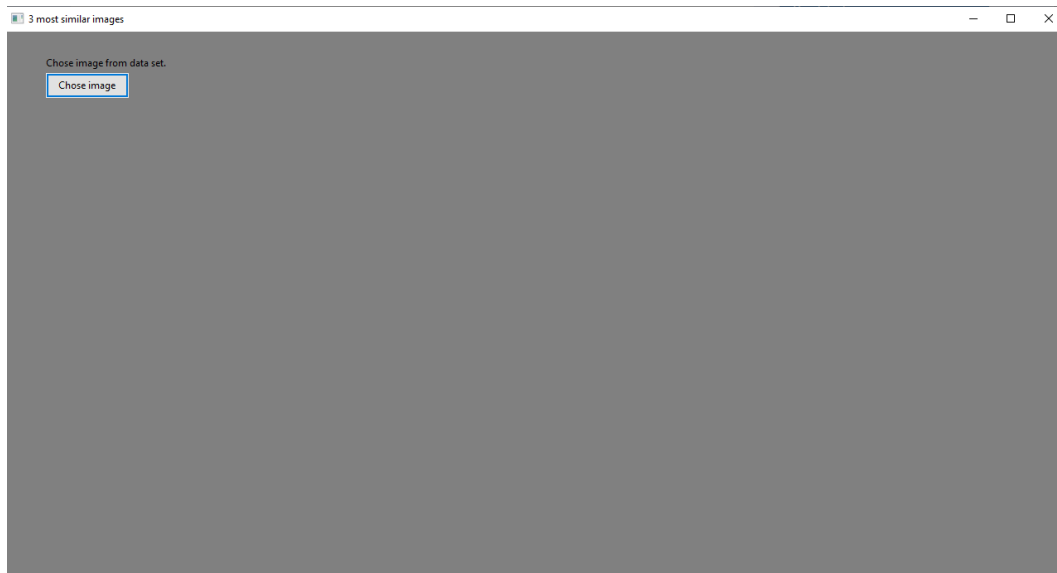


Figure 5: Primul stadiu al interfetei grafice

Pentru a respecta principiul secventialitaii, utilizatorul are la dispozitie performarea unei singure actiuni si anume incarcarea unei imagini, imagine ce reprezinta imaginea sursa pentru care se vor identifica mai tarziu cele 3 alte imagini similare. Deasupra butonului este plasat un text informational ce arata starea curenta a aplicatiei sau actiunea ce este curent intretinuta, de asemenea functionalitatea butonului este inscrisa pe suprafata sa.

La apasarea butonului, o fereastra de explorare a fisierelor se va deschide si va fi necesara selectarea imaginii sursa. Dupa ce este selectata, imaginea se va incarca in fereastra interfetei.

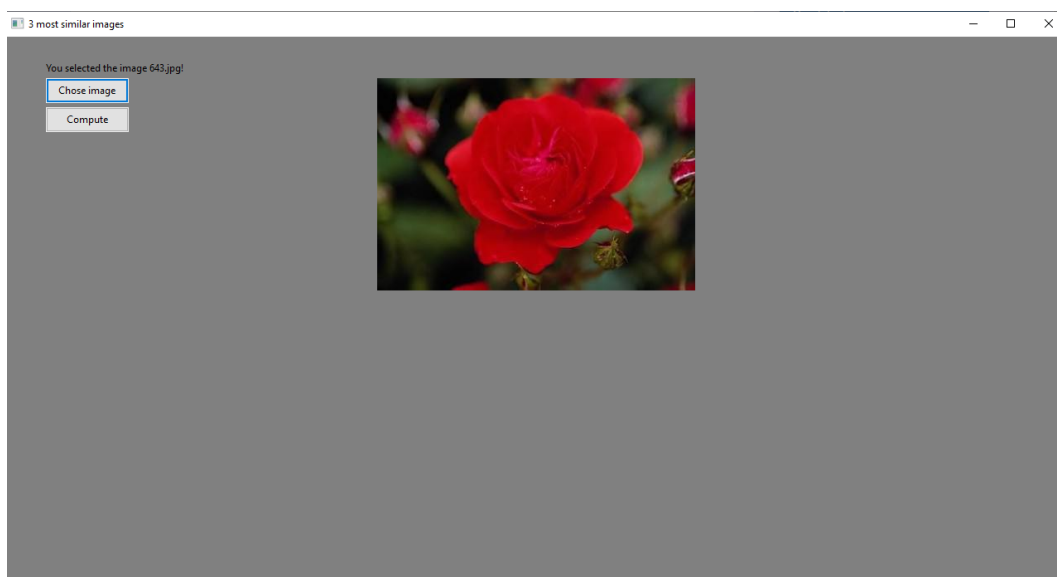


Figure 6: Stadiul intermediar al interfetei grafice

În stadiul intermediar este prezentată imaginea sursă încărcată, și în textul informațional este printat numele acestei imagini așa cum apare în fișier. Tot în această etapă, respectând principiul grupării, sub primul buton este afișat un altul ce are ca funcționalitate pornirea algoritmului, ce constă în determinarea setului de trasaturi ale tuturor imaginilor ce se află în același director cu imaginea sursă. După calcularea descriptorilor, pe baza distanței euclidiene sunt selectate cele mai similare 3 imagini și afișate conform principiului regularității sub imaginea sursă. Interfața a fost proiectată cu ajutorul modului wx. Plecând de la structura clasică a unei implementări de interfață wxPython (clasa Frame, clasa Panel) au fost introduse funcționalitățile dorite, și anume: butoanele proiectate folosind constructorul `wx.Button(...)`, textul folosind `wx.StaticText(...)`, atasarea unor evenimente butoanelor cu metoda `wx.Bind(...)` și afișarea imaginilor folosind metoda `static.Bitmap(...)`.

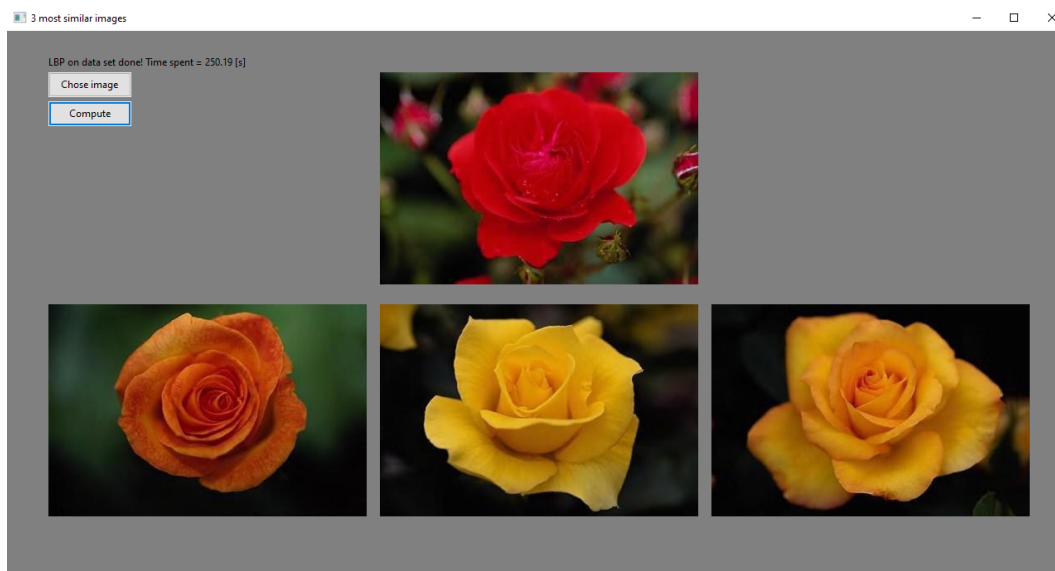


Figure 7: Stadiul final al interfetei grafice

În stadiul final, interfața da posibilitatea utilizatorului de a alege o altă imagine sursă, dar și de a calcula din nou descriptorii LBP și distanțele aferente, cât și selecția celor mai similare 3 imagini, însă datorită faptului că în algoritmul de determinare al imaginilor similare și al descriptorilor nu intervin procese stochastice rezultatele vor fi aceleași pentru aceeași imagine sursă.

4 Performante

Pentru a analiza performanța și pentru a alege varianta de LBP optimă va trebui să definim o metrică sau un mod prin care se va putea discerne între rezultatele obținute. Având în vedere faptul că, algoritmul determină și alege cele mai similare imagini dintr-un

mediu controlat in totalitate de utilizator, rezultatele vor fi intr-o mare masura dependente de setul de date, adica de baza de imagini pusa la dispozitie. In acest sens am considerat ca pot aparea doua tipare distincte destul de probabile in setul de date: o situatie in care imaginile apartin mai multor clase si exista mai mult de o imagine din fiecare clasa si o situatie in care imaginile apartin mai multor clase dar nu exista mai mult de o imagine din fiecare clasa. Pentru primul caz am definit criteriul de performanta asemanator cu modul in care un subiect uman ar alege cele mai similare imagini, deci in cazul in care exista o imagine sursa din clasa C ar fi alese ca cele mai similare tot imagini apartinand clasei C ; astfel am impartit setul de date in 5 clase de imagini: monumente, dinozauri, flori, cai si masini. Alegandu-se o imagine din aceste 5 clase ca imagine sursa, vor fi considerate ca rezultate valide doar imaginile din aceeaasi clasa cu aceasta, altfel orice alta imagine dintr-o clasa distincta va fi considerata ca identificata gresit. In acest mod scorul maxim pentru o imagine sursa este de 3/3 imagini identificate corect. Pentru al doilea caz, in care nu exista mai multe imagini din aceeaasi clasa, nu am putut defini o metrica deoarece criteriul uman de selectie al imaginilor similare in acest caz este mult prea subiectiv. Am considerat pentru o mai buna aproximare scorul mediu pe intreg setul de date ca indice al performantei, deoarece ne intereseaza performantele algoritmului aplicat pe un numar cat mai mare de clase.

Tipul de LBP	Scorul mediu
LBP fundamental	1.56
LBP mediu	1.14
LBP circular (8,1)	1.26
LBP circular (16,2)	1.08
LBP circular (8,3)	0.9

Table 1: Scorul mediu pentru fiecare varianta de LBP

Din datele obtinute putem observa ca cea mai buna performanta o are varianta de LBP fundamental cu 1.56 imagini identificate corect in medie pentru o anumita imagine sursa.

5 Bibliografie

References

- [1] Cursul de Interfete Om-Maisna, Andreea Griparis, Facultatea de Electronica, Telecomunicatii si Tehnologia Informatiei.
- [2] Cursul de Prelucrarea Imaginilor, Constantin Vertan, Facultatea de Electronica, Telecomunicatii si Tehnologia Informatiei.

- [3] Kopka, H., Daly P.W., *A Guide to LaTeX*, Addison-Wesley, Reading, MA, 1999.
- [4] Pan, D., “A Tutorial on MPEG/Audio Compression,” *IEEE Multimedia*, Vol.2, pp.60-74, Summer 1998.