

+

o

Recunoaștere de text în imagini

Procesare de imagini

Bozdog Raluca – Delia

CTI română, seria A, grupa 30233

An universitar 2021 - 2022

Îndrumător:

Asist. Dr. Inginer Robert Varga



Contextul problemei

- Pornind de la o imagine color despre care se știe că are în compoziție doar text, să se extragă textul sub formă de caractere

+

•

- **Simplificare a problemei:**

- s-a ales un singur font care să fie recunoscut (Times New Roman)
- Imaginile furnizate vor conține text orientat aproape orizontal

- **Utilitate practică:**

- E-books
- Semne de circulație – mașini autonome
- Traduceri
- Persoane cu probleme de vedere

○

Manual de utilizare

- Instrucțiuni în consolă
- Meniu continuu (după executarea unei rulări există mereu posibilitatea unei noi execuții)
- Utilizatorul

- +
 - Introduce de la tastatură numele imaginii în care dorește să identifice textul
 - Numele furnizat trebuie să conțină și extensia imaginii
 - Imaginea trebuie să se găsească în folder-ul *Images* al proiectului

○

- Se afișează imaginea aleasă sau un mesaj de atenționare în caz de eroare
- Se închide imaginea pentru a debloca execuția
- Pe măsură ce este identificat, textul este afișat în consolă
- Simultan se face și scrierea într-un fișier de ieșire
- Fișierul de ieșire este generat automat pe baza numărului de ordine al imaginii curente
- Numele fișierului de ieșire îi este indicat utilizatorului la finalul rulării

```
Enter image name and extension to start looking for text
>>Image:      14.png
Close image to continue

Found text:

abcdefghiiklmnopqrs
tuvwxyzABCDEFGG
HIJKLMNOPQRST
UVWXYZ0 1 2345 67
89,<.>/?; : ,,,[ {} ]\~!+-
-)(*&^%$#@! `~

You can find the text in text_from_image_1.txt

Enter image name and extension to start looking for text
>>Image:      -
```

Funcții definite

➤ Funcții dezvoltate la laborator

- ❑ `Mat_<uchar> colorToGrayscale(Mat_<Vec3b> img)`
- ❑ `Mat_<uchar> grayscaleToBinary(Mat_<uchar> img, int threshold)`
- ❑ `int* histogram(Mat_<uchar>img)`
- ❑ `float* fdp(Mat_<uchar> img)`
- ❑ `Mat_<uchar> autoBinary(Mat_<uchar> img, float err)`
- ❑ `structTagImage tagBinaryImageBFS(Mat_<uchar> img, int nbNeigh)`
- ❑ `Mat_<Vec3b> tagImageToRGBImage(Mat_<int> img)`
- ❑ `Mat_<uchar> boundingBox(Mat_<uchar> img)`

Funcții definite

➤ Funcții dezvoltate „de la zero”

- ❑ `void iterateOverDirectory(void)` - preprocesarea imaginilor șablon
- ❑ `int getMatchPercentage(Mat_<uchar> img, Mat_<uchar> pattern)` - găsirea potrivirii unei imagini cu un șablon dat
- ❑ `structMatch getBestMatch(Mat_<uchar> img)` - cea mai bună potrivire
- ❑ `structBoundingBox* getBoundingBoxes(structTagImage tag)`
- ❑ `int* occupiedRows(Mat_<uchar> img) & int* occupiedCols(Mat_<uchar> img, int rmin, int rmax)` - care rânduri și coloane din imagine sunt ocupate
- ❑ `structRow* boundRows(Mat_<uchar> img, int* rowNb)` - delimitarea randurilor de text
- ❑ `void isolateForMatching(structBoundingBox* boundingBoxes, int nbLabels, Mat_<uchar> img, FILE* out_file)` - tratare caz particular (to be discussed)
- ❑ `void showRows(Mat_<uchar> img, FILE* out_file)` - procesarea succesivă a fiecărui rând de text

Descrierea metodei

❑ Idei inițiale (și limitările lor):

- Caracter = obiect
- Caracter = zonă bounding-box discretă

+

•

○

fg

ij

ab

❑ Soluția propusă – combinarea celor două metode inițiale

abcdefghijklmnopqrstuvwxyz
rstuvwxyzABCDE
FGHIJKLMNO PQ
RSTUVWXYZ012
3456789,<.>/?:;'"'[
{]}\\|=+ -
_)(*&^%\$#@!`~

Descrierea metodei



- Crearea unei „baze de date” cu imagini șablon bounding-box
- Transformarea imaginii colorate de intrare în imagine grayscale
- Binarizarea automată a imaginii
- +
 - ○ **Proiecție pe verticală** => rândurile din imagine
 - Pentru fiecare rând – **proiecție pe orizontală** = > bounding-boxes pentru caractere /
- grupări de caractere care au „umbre” suprapuse
- Calcul procent matching între imaginea bounding-box și fiecare șablon
- Alegere șablon cu procent maxim de potrivire
- Dacă potrivirea este mai slabă decât un prag dat, se presupune că este un caz de „glued characters” = mai multe caractere care împart același bounding-box
- Se realizează extragerea obiectelor din imaginea bounding-box prin **etichetare**

Descrierea metodei – cazuri limită

- Tratarea cazurilor speciale $i = j$:

- Mijlocul bounding-box-ului unui obiect proiectat pe axa orizontală face parte din proiecția pe orizontală a unui alt obiect

- Vector de intersecții

- 0 dacă nu există intersecții
- $n > 0$ pentru indicii x și k dacă obiectele cu aceiași indici se suprapun

- Vector auxiliar de structuri **structBoundingBox** – va deține noul bounding box al obiectelor ce sunt, de fapt, un singur caracter (au același număr la indicele lor în vectorul de intersecții)

- La parcurgerea finală

- dacă în vectorul de intersecții se găsește 0, se găsește best match pentru bounding-box-ul inițial (nemodificat)
- dacă se găsește un alt număr, mai mare decât 0, se găsește best match pentru noul bounding-box (cel comun, calculat în funcție) și se marchează cu -1 în vectorul de intersecții
- Dacă se găsește -1, caracterul a fost deja introdus anterior = > ignorare

Algoritmi folositi – modificări etichetare

- Pentru a servi aplicației particulare de recunoaștere de text, s-a modificat ordinea de parcurgere a imaginii pentru etichetare:

- Se preferă parcurgerea mai întâi crescător pe coloane și apoi descrescător pe linii

```
for (int j = 0; j < img.cols; j++)  
    for (int i = img.rows - 1; i >= 0; i--)
```

- Astfel se asigură **păstrarea ordinii caracterelor** din textul imaginii cu cele din textul de ieșire al programului

Rezultate finale

Chapter 1

It was a bright cold day in April, and the clocks were striking thirteen. Winston Smith, his chin nuzzled into his breast in an effort to escape the vile wind, slipped quickly through the glass doors of Victory Mansions, though not quickly enough to prevent a swirl of gritty dust from entering along with him.

Chapter 1

It was a bright cold day in April, and the clocks were striking thirteen- Winston Smith, his chin nuzzled into his breast in an effort to escape the vile wind, slipped quickly through the glass doors of Victory Mansions, though not quickly enough to prevent a swirl of gritty dust from entering along with him-

Rezultate finale

Sometimes people are
beautiful not in looks,
not in what they say,
just in what they are.

Markus Zusak –

The Book Thief

Sometim+es people are

beautiml not in looks,
not in what they say,
just in what they are.

Markus zusak -

The Book Thief

Rezultate finale

Wow, chiar merge
surprinzator de bine
si daca textul e rotit

wow, chiar merge
surprinzator de bine
si daca textu1 e rotit

Bibliografie

- <https://learnopencv.com/image-resizing-with-opencv/>
- https://docs.opencv.org/4.x/d9/df8/tutorial_root.html
- <https://codeburst.io/optical-character-recognition-recognizing-text-to-labels-on-an-android-platform-4c20bddc9175>

+

•

○

Grupa 30233 semigrupa 1
e cea mai tare