**Εισαγωγή στην Επεξεργασία Εικόνας**

**2η ΕΡΓΑΣΙΑ**

**Θέμα:** “Τμηματοποίηση εικόνας κειμένου σε λέξεις”

**Φοιτήτρια:** Στεργίου – Καψάλη Βασιλική

A.M.: 180335

**Α’ Μέρος - Μέθοδος**

clc;

clear;

close all;

format compact;

more off;

% ------------------------------

% PART A

% ------------------------------

**Α1)** Διάβασμα και εμφάνιση της αρχικής εικόνας κειμένου που δίνεται.

% --- Step A1

% read the original RGB image

Filename='Troizina 1827.jpg';

I=imread(Filename);

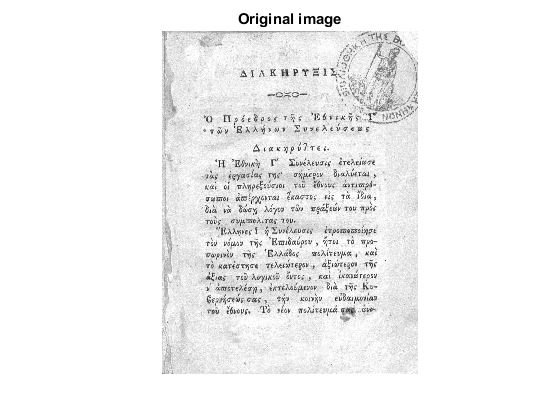
% show it

figure;

image(I);

axis image off;

title('Original image');



**Α2)** Μετατροπή της grayscale εικόνας και εμφάνιση της σε καινούργιο γράφημα. Για την μετατροπή χρησιμοποίηση της συνάρτησης **any\_image\_to\_grayscale\_func.m**.

% --- Step A2

% convert the image to grayscale

A=any\_image\_to\_grayscale\_func('Troizina 1827.jpg');

**Β5)** Εφαρμογή **gamma correction** στην grayscale εικόνα χρησιμοποιώντας την εντολή **imadjust**. Με τον τρόπο αυτό η εικόνα θα γίνει πιο «φωτεινή» ή πιο «σκοτεινή» αντίστοιχα που θα έχει άμεση επίπτωση στα επόμενα βήματα του Α’ μέρους και κατά συνέπεια και στα τελικά αποτελέσματα.

% --- Step B5

% apply gamma correction (a value of 1.0 doesn't change the image)

GammaValue=0.4;

A=imadjust(A,[],[],GammaValue);

% show the grayscale image

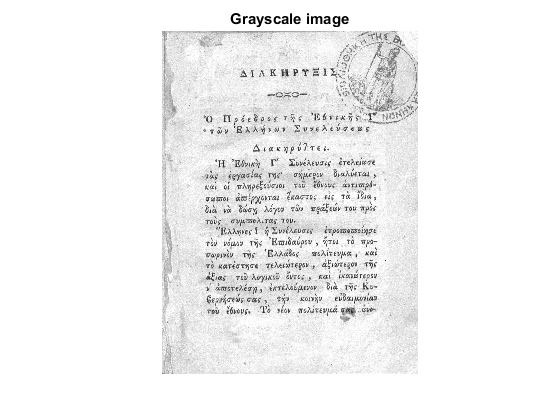
figure;

image(A);

colormap(gray(256));

axis image off;

title('Grayscale image');



**Α3)** Μετατροπή της grayscale εικόνας, σε δυαδική χρησιμοποιώντας την εντολή **graythresh** του Octave/Matlab η οποία βασίζεται στην μέθοδο του **Otsu.** Εμφάνιση δυαδικής εικόνας σε ένα καινούργιο γράφημα.

% --- Step A3

% convert the grayscale image to black-and-white

Threshold= graythresh(A);

BW = ~im2bw(A,Threshold);

Απεικόνιση **άρνησης** δυαδικής εικόνας (~BW), ώστε τα γράμματα να εμφανίζονται μαύρα τα γράμματα και το φόντο άσπρο.

% show the black-and-white image

figure;

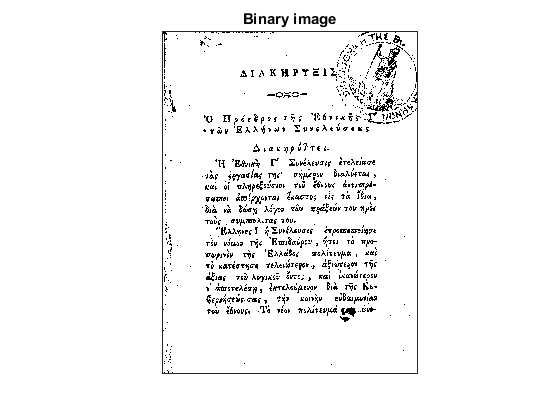
image(~BW);

colormap(gray(2));

axis image;

set(gca,'xtick',[],'ytick',[]);

title('Binary image');



**Α4)** Ακολουθία από μορφολογικές πράξεις για απαλοιφή θορύβου και των μεγάλων συνδεδεμένων στοιχείων που δεν αντιστοιχούν σε λέξεις της εικόνας.

% --- Step A4

% make morphological operations to clean the image

Επιλογή δομικών στοιχείων για μορφολογικές πράξεις

A=strel('rectangle', [3 10]);

B=strel('diamond', 1);

Διαστολή, με την εντολή indilate με χρήση του δομικού στοιχείου rectangle.

C=imdilate(BW, A);

Χρήση της imclearborder για αφαίρεση σφραγίδας και θορύβου από εικόνα.

C=imclearborder(C);

Δυαδική πράξη AND μεταξύ της διεσταλμένης- καθαρισμένης εικόνας και της ασπρόμαυρης, για επαναφορά γραμμάτων.

stamp=and(BW, C);

% show the cleared from stamp, image

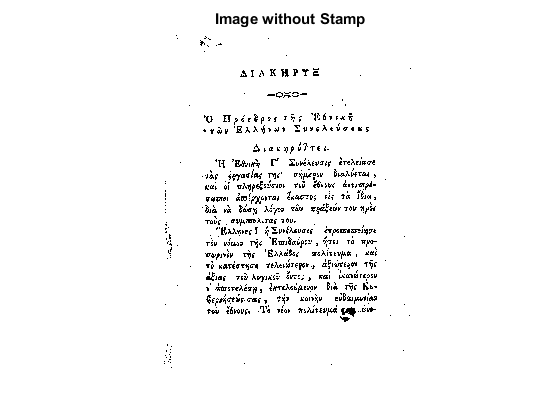
figure;

image(~stamp);

colormap(gray(2));

axis image off;

title('Image without Stamp');



Χρήση του δομικού στοιχείου diamond και εντολής imopen για αφαίρεση στιγμάτων από την εικόνα.

imopen : Άνοιγμα εικόνας με σκοπό το σπάσιμο μικρών συνδέσμων ανάμεσα στις λέξεις.

% open image and clear noise

stamp=imopen(stamp, B);

% show cleaned image

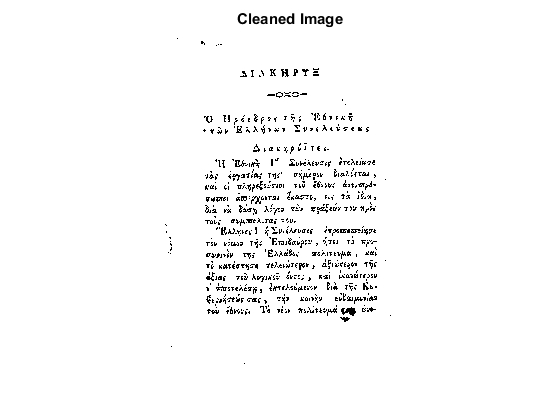
figure;

image(~stamp);

colormap(gray(2));

axis image off;

title('Cleaned Image');



ΕΡΩΤΗΣΗ 1: Με ποιο σκεπτικό και ποια σειρά επιλέξατε τις διάφορες μορφολογικές και λογικές πράξεις στο βήμα Α4 με σκοπό τον καθαρισμό της εικόνας και την αφαίρεση της σφραγίδας; Αιτιολογήστε αναλυτικά την απάντησή σας.

Απαλοιφή θορύβου και των μεγάλων συνδεδεμένων στοιχείων που δεν αντιστοιχούν σε λέξεις της εικόνας, μέσω διαστολής (dilation), με την εντολή indilate με χρήση του δομικού στοιχείου rectangle, ώστε να πετύχουμε διόγκωση του θορύβου.

Χρήση της imclearborder στην διεσταλμένη εικόνα, για αφαίρεση σφραγίδας και θορύβου.

Δυαδική πράξη AND μεταξύ της διεσταλμένης-καθαρισμένης εικόνας και της ασπρόμαυρης για την εμφάνιση των γραμμάτων.

Χρήση του δομικού στοιχείου diamond και εντολής imopen για αφαίρεση στιγμάτων από την εικόνα και άνοιγμα εικόνας, με σκοπό το σπάσιμο μικρών συνδέσμων ανάμεσα στις λέξεις..

**Α5)** Συνένωση συνδεδεμένων στοιχείων σε μεγαλύτερα με σκοπό κάθε συνδεδεμένο στοιχείο να αντιστοιχεί (όσο αυτό είναι δυνατό) σε κάποια **λέξη** του κειμένου, με χρήση δομικού στοιχείου line, με length 18 κατόπιν δοκιμών. Διαστολή για ένωση των γραμμάτων των λέξεων.

Έπειτα, εφαρμογή τις **bwlabel** και **label2rgb** ώστε να εμφανιστεί με διαφορετικό χρώμα το κάθε συνδεδεμένο στοιχείο. Για τον σκοπό αυτό χρήση της εντολής **bwlabel** του OCTAVE/MATLAB με συνδεσιμότητα 4 σημείων (4-connectivity).

% --- Step A5

% make morphological operations for word segmentation

% find the right length

D = strel('line', 18, 0);

C=imdilate(stamp, D);

% find the connected components (using bwlabel)

% colors

[C, count]=bwlabel(C, 4); % 4-connectivity

%[ C, count]=bwlabel(C, 8); % 8-connectivity

%show word segmentation

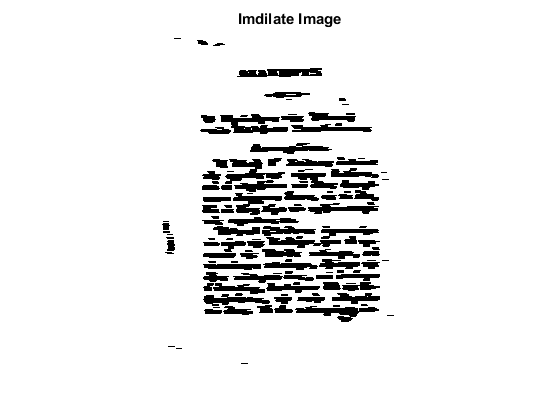
figure;

image(~C);

colormap(gray(2));

axis image off;

title('Imdilate se Image');



Χρωματισμός συνδεδεμένων στοιχείων

RGB = label2rgb(C, 'lines');

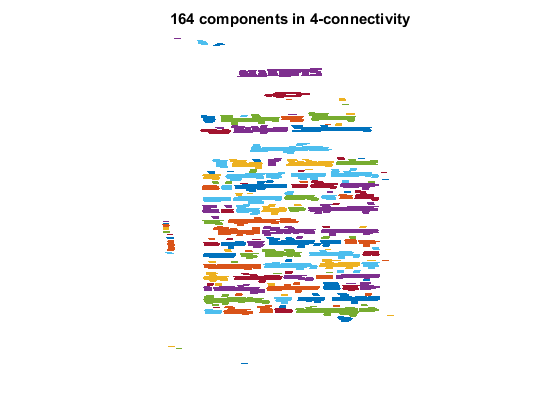
figure;

image(RGB);

colormap(gray(2));

axis image off;

title(sprintf('%g components in %d-connectivity',count,4));



ΕΡΩΤΗΣΗ 2: Με ποιο σκεπτικό και ποια σειρά επιλέξατε τις διάφορες μορφολογικές και λογικές πράξεις στο βήμα 5 με σκοπό τον προσδιορισμό των λέξεων; Αιτιολογήστε την απάντησή σας.

Εφαρμόζοντας διαστολή και του δομικού στοιχείου line με length 18, πετύχαμε συνένωση γραμμάτων και διαχωρισμό λέξεων. Το πλήθος των bits του δομικού στοιχείου line δίνεται από τον χρήστη της εφαρμογής, κατόπιν δοκιμών. Έπειτα, εφαρμογή τις **bwlabel** και **label2rgb** ώστε να εμφανιστεί με διαφορετικό χρώμα το κάθε συνδεδεμένο στοιχείο. Για τον σκοπό αυτό χρήση της εντολής **bwlabel** του OCTAVE/MATLAB με συνδεσιμότητα 4 σημείων (4-connectivity).

ΕΡΩΤΗΣΗ 3: Πόσα συνολικά στοιχεία (components) βρέθηκαν στην εικόνα; Αν χρησιμοποιήσετε συνδεσιμότητα 8 σημείων (8-connectivity) σαν δεύτερη παράμετρο στην **bwlabel** πόσα συνολικά components δημιουργούνται; Πού οφείλεται η διαφορά στις δύο αυτές τιμές; Ποιος τύπος συνδεσιμότητας είναι καλύτερος κατά την γνώμη σας και γιατί;

164 στοιχεία με συνδεσιμότητα 4.

163 στοιχεία με συνδεσιμότητα 8.

Η συνδεσιμότητα 8 σημείων, καλύπτει μεγαλύτερη επιφάνεια και ενώνει περισσότερα αντικείμενα σε ένα.

Προτίμηση του τύπου συνδεσιμότητας 4 σημείων (4-connectivity), όπου και εντοπίζει περισσότερα στοιχεία, διότι οι λέξης έχουν γειτονικά σημεία μόνο στην οριζόντια και κάθετη κατεύθυνση και όχι διαγώνια.

**A6)** Ορισμός περιεγραμμένου ορθογώνιου.

Εμφάνιση αρχικής εικόνας του κειμένου και πάνω της σχεδιασμός ορθογωνίων των λέξεων που εντοπίστηκαν.

% --- Step A6

% borders

for i = 1:count

[row, col]=find(C == i);

for j = [row col]

C(row, col)=i;

end

end

RGB = label2rgb(C, 'lines');

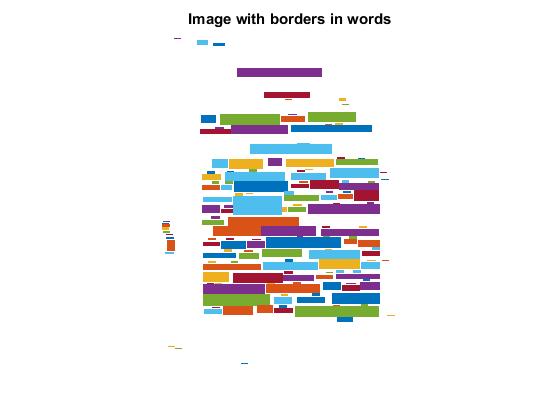
figure;

image(RGB);

colormap(gray(2));

axis image off;

title('Image with borders in words');



% extract bounding boxes properties

% bwconncomp(BW) returns the connected components CC found in the binary image BW

% regionprops measure properties of image regions

CC = bwconncomp(C, 4);

s = regionprops(CC, 'BoundingBox');

% create a matrix nX4. n = bounding box total number and 4 properties of bounding box (x,y,width,height)

bbox = vertcat(s.BoundingBox);

% show the original image

figure;

image(I);

axis image off;

title('Final results');



% draw bounding boxes edges on the original image

for idx = 1 : numel(s)

rectangle('Position', bbox (idx,:), 'edgecolor', [rand rand rand], 'LineWidth', 2);

end

ΕΡΩΤΗΣΗ 4: Προτείνετε μια παραλλαγή στην ακολουθία που χρησιμοποιείτε στο βήμα 6 ώστε το ορθογώνιο που περιγράφει κάθε λέξη να συμπεριλαμβάνει (κατά το δυνατόν) και τα **σημεία στίξης** της εκάστοτε λέξης (τόνο, περισπωμένη κλπ). Τι αλλαγή κάνατε;

Δημιουργία πλαισίων για τα στοιχεία που έχουν μέγεθος πάνω από 1 pixel (όσο περίπου πρέπει να είναι ένα τόνος).

Αύξηση ύψους πλαισίων κατά 1 pixel προς τα πάνω, ώστε να περιλαμβάνει και τους τόνους.

**A7)** Αποθήκευση ορθογωνίων σε ένα αρχείο κειμένου «**results.txt**».

% --- Step A7

% Convert from the [x y width height] bounding box format to the [xmin ymin xmax ymax] format

% Calculate pixel coordinates and extract to a results.txt with dlmwrite function

xmin = results(:,1)+0.5;

ymin = results(:,2)+0.5;

xmax = xmin + results(:,3) - 1;

ymax = ymin + results(:,4) - 1;

dlmwrite('results.txt',[xmin,ymin,xmax,ymax],'delimiter','\t');

fprintf(Values extracted to file "results.txt".\n');

**Β’ Μέρος - Αποτίμηση και συγκριτικές δοκιμές**

% ------------------------------

% PART B

% ------------------------------

Β1) Διάβασμα αρχείου των αποτελεσμάτων της μεθόδου **results.txt** που δημιουργήθηκε στο Α’ μέρος βήμα Α7. Για κάθε προσδιορισμένη θέση λέξης, υπολογισμός του 𝐼𝑂𝑈 με όλες τις λέξεις του αρχείου **Troizina 1827\_ground\_truth.txt**.

% --- Step B1

% load the ground truth

GT=dlmread('Troizina 1827\_ground\_truth.txt');

% load our results

R=dlmread('results.txt');

% --- Step B2

% calculate IOU for all the results ...

overlapRatio = bboxOverlapRatio(R, GT);

%max value per colunm

IOU=max(overlapRatio);

% --- Step B3

% calculate the Score for this pair of (GammaValue,IOUThreshold) ...

IOUThreshold=0.5; % or 0.3 or 0.7

score=(IOU>=IOUThreshold);

score5=(sum(score)/size(R,1))\*100;

fprintf('IOU score is %0.2f%%', score5);

Score = 68.29%

Όσο ανεβαίνει το κατώφλι τόσες λιγότερές λέξεις το ικανοποιούν.

ΕΡΩΤΗΣΗ 6: Παρουσιάστε και σχολιάστε τα αποτελέσματα που πίνακα. Ποιός συνδυασμός (gamma value – IOU threshold) δίνει το καλύτερο 𝑺𝒄𝒐𝒓𝒆; Ποια από τις τρεις τιμές IOU threshold επηρεάζεται περισσότερο (έχει μεγαλύτερη διασπορά τιμών) από τις διάφορες τιμές gamma value; Αιτιολογείστε τις απαντήσεις σας.

|  |  |  |  |
| --- | --- | --- | --- |
| Πίνακας με 𝑆𝑐𝑜𝑟𝑒𝑠 | | | |
|  | **IOU threshold** | | |
| **gamma value** | 0.3 | 0.5 | 0.7 |
| 0.6 | 64,93 | 64,18 | 61,19 |
| 0.8 | 67,97 | 65,63 | 60,16 |
| 1.0 | 70,73 | 68,29 | 60,98 |
| 1.2 | 73,73 | 72,03 | 61,86 |
| 1.4 | 79,09 | 77,27 | 66,36 |

Ο καλύτερος συνδυασμός (gamma value – IOU threshold) που δίνει το καλύτερο 𝑺𝒄𝒐𝒓𝒆 είναι gamma value = 1,4 και IOU threshold = 0.3.

Τη μεγαλύτερη διασπορά τιμών παρουσιάζει IOU threshold = 0,3 από τις διάφορες τιμές gamma value.