Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 4, Issue. 5, May 2015, pg.1117 – 1123

RESEARCH ARTICLE

Arabic Text Extraction from Video Film

Prof. Dr. Kadhim Mahdi Al-Musawi, Noor Ali Al_helaly
Computer Science Department, Education College, Thi-Qar University, Thi-Qar, Iraq
kadhimmehdi63@gmail.com; nnnaaa90@gmail.com

Abstract*— *In recent years, Spread of research for treatment of Multimedia abstraction objects, tracking or recognition and including extracting texts research encompassing various international languages. The text including in video is an important thing for indexing and searching system. However, this text is difficult to detecting and recognizing because of the variability of its size, their low resolution characters and the complexity of the background. In this paper we offer an artificial approach to extract the Arabic text of the video. Our works on the using of image processing concepts to text detection then find the edges of the text and separated from the complex background. This is followed by tracing process through text in sequences frames, and then cutting the text to lines words and characters, to be ready for recognition using commercial OCR or program artificial neural networks.

Keywords: Pre-processing, Scene text, Video frame extraction, Text Detection, Text Tracing, Text segmentation and N.N.

1-Introduction:

Video is the most effective media for capturing the world around us. A large variety of video-based applications, such as video on demand, interactive TV, digital library, online distance learning, remote video surveillance, video conferencing, and so forth, has attracted much interest and attention. Efficient ways to analyze, annotate, browse, manipulate, and retrieve videos of interest based on their contents are becoming increasingly important and have attracted substantial research attention over the past decade accordingly [1][2].

Text objects embedded in video contain much semantic information related to the multimedia content.

Text in video images can be classified into Artificial text and Scene text. Artificial text is artificially overlaid on the video image at the time of editing such as: newscast titles, films subtitles, sport scores, persons of interest names, etc. (See Figure 1.a). Scene text is naturally existed in the field of view of the camera during video capture such as: text on street signs, text on trucks, writings on shirts, etc. Its usefulness is confined to ad hoc applications such as: navigation and surveillance systems (See Figure 1.b). Artificial text is more descriptive and meaningful to the video content than the scene text[3][4].

(a) Artificial texts, (b)Scene texts

Figure (1) : Artificial and Scene Texts

However, the variations of text due to differences in font, size, orientation, style, alignment, complex background, unknown layout makes the text extraction from video a challenging task.

Often, text information extraction (TIE) problem can be divided into five sub-problems: (i) detection, (ii) localization, (iii) tracking, (iv) extraction and enhancement, and (v) recognition [5].

In this paper we offer an artificial approach to extract the Arabic text from the video .our work on the using of image processing concepts to find the edges of the text and separated from the complex background, then followed by the process of tracing through text in sequences frames, and cutting the text to lines words and characters, To be ready for recognition using OCR commercial or program.

2. Arabic Language Overview:

Arabic language has a very rich vocabulary. More than 422 million people speak this language as their native speaking language, and over 1 billion people use it in several religion-related activities , making it one of the half dozen most popular languages in the world. Although Arabic texts have vast popularity, they have not received enough interest by researchers, especially in the field of text extraction from video images. Arabic alphabet is represented numerically by a standard communication interchange code approved by the Arab Standard of Metrology Organization (ASMO) which resembles the American Standard Code for Information Interchange (ASCII). Each character in the ASMO code is represented by one byte [6][7].

The main characteristics of Arabic text can be summarized as follows:

1. A character may have several strokes(Al fatha).
2. Arabic characters consist of strokes with horizontal, vertical and diagonal directions(al kaf , Al fatha).
3. Arabic letters can have more than one shape according to their position in the word: initial, middle, final, or standalone, Unlike English letters which have two shapes uppercase and lowercase.
4. In English, only two letters "i" and "j" in the lowercase representation have secondary character (complementary) that is located above the primary character which is called "dot or point", while in Arabic 15 letters out of 28 have dots which are located above, below or in the middle of the primary characters and 4 letters have another type of secondary character (complementary) that is located above or below the primary character which is called "Hamza".

5. Arabic characters are normally connected to each other to form meaningful word meaning that there are no gaps between most letters in a single word. Hence baseline of Arabic language is rich in pixels more than what is in baseline of English language.
6. Arabic language is horizontally alignment (i.e.: Arabic script is written and read from right to left).
7. Arabic letters are rich with angles or corners.

3- The proposed algorithm for text extraction:

Text extraction isa process by which we convert Printed document/scanned page or image in which text are available to regions character that a computer can recognized it
After convert video to sequences of frames apply three main tasks to extract text they are:

1-pre-processing .
2- tracking
3- segmentation

3:1- pre-processing:

The first stage in pre-processing is video frames extraction. In this stage, the video containing the text is split into frames,by using some concepts of image processing ,to get enhancement image for candidate region include text by applied convert colour image to grey by using Matlab's instruction , The characteristics (1, 2, 4, 5, 6 and 7) of Arabic language mentioned in (Section 2) provide more pixels in text regions than background regions. In order to highlight those pixels over smooth background we apply Laplacian of Gaussian detector.
Laplacian of Gaussian (LOG) is an edge detection operator in which Gaussian low-pass filter (LPF) smoothing is performed prior to the application of the Laplacian edge detector because Laplacian is extremely sensitive to noise [8].
Luminance component of a predetermined Region of Interest(ROI) is filtered by 5x5 Laplacian of Gaussian mask with Sigma,(σ=0.52). Laplacian generates “double edges,” that is, positive and negative values for each edge.
In the flat regions and along the ramp, the Laplacian is zero. Large positive values of the Laplacian will occur in the transition region from the low plateau to the ramp; large negative values will be produced in the transition from the ramp to the high plateau [8] , in general , the transition in our case is the transition between text and background.
(Figure 2) shows two examples of using Laplacian of Gaussian filter, one can notice that LOG operator is efficient for revealing the objects with high gradient magnitudes and discards others.

(A)original image

(B) grey image

(C)edge image

(D)final edge image after enhancement and remove noise

Figure (2): pre-processing

3:2Tracking :

Tracking process we use to get a frames that contains the text without recurrence ,many frames not contain text in order that applied tracking process using matching and some geometric conditions

Figure (3):video without tracking

3:3 Segmentation :

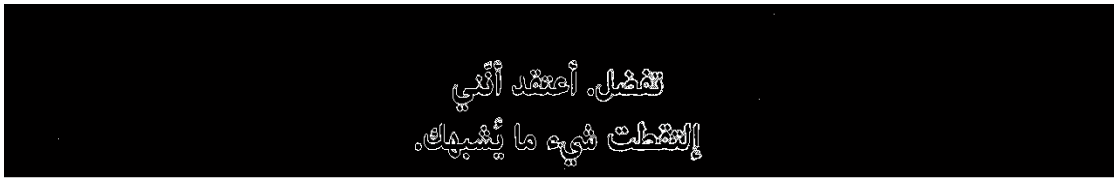
The frames that getting by tracking may be contain more than one text's line In this process we segmented frame to get only text's line we now apply various techniques for segmentation of document lines, words and characters. The process of segmentation mainly follows the following pattern[9]:

1- Identify the text lines in the page.
2- Identify the words in individual lines.
3- Finally identify individual characters in each word.

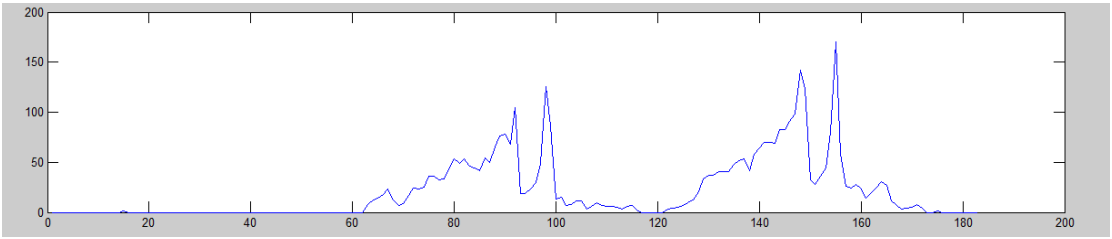
3:3:1 Line Segmentation.

The global horizontal projection method is used to compute sum of all white pixels on every rowand construct corresponding histogram. The steps for line segmentation algorithm are :

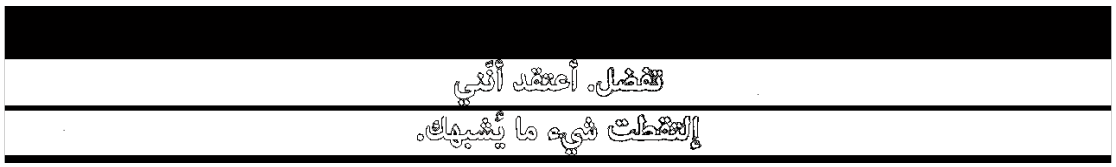
step 1:Construct the Horizontal Histogram for the image .
step 2:Count the white pixel in each row.
Step 3 :Using the Histogram, find the rows containing no white Pixels .
Step 4:Replace all such row by 1 .
Step 5 :Invert the image to make empty rows as 0 and text lines Will have original pixels.
Step 6 :Mark the Bounding Box for text lines using standard
Matlab functions (region props and rectangle).
Step 7 :Copy the pixels in Bounding Box and save in separate file.
You can see application of algorithm in figure 4



(a)original image



(b)Image histogram



(c) Blank space between the lines

(d)Lines segmentation

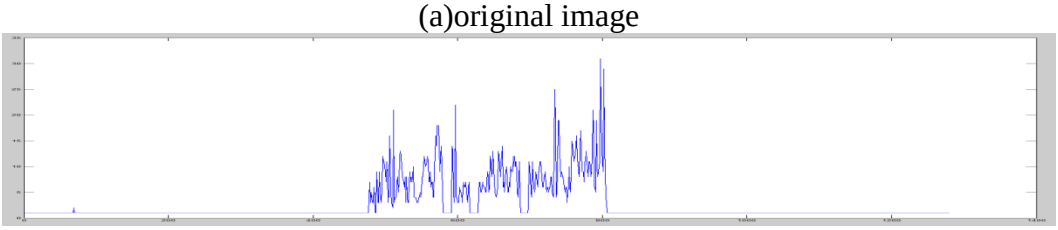
Figure 4 :line segmentation

3:3:2 Word Segmentation

The global horizontal projection method is used here to compute sum of all white pixels on every column and construct corresponding histogram. The steps for line segmentation algorithm are :

Step 1: Construct the Vertical Histogram for the image.
Step 2:Count the white pixel in each column.
step 3:Using the Histogram, find the columns containing no white pixel.
step 4 :Replace all such columns by 1
Step 5 :Invert the image to make empty rows as 0 and text words will have original pixels.
Step 6 :Mark the Bounding Box for word.
Step 7:Copy the pixels in the Bounding Box and save in separate file.
You can see application of algorithm in figure 5

(a)original image



(b)Image histogram

(c) Blank space between the words

(d) word segmentation

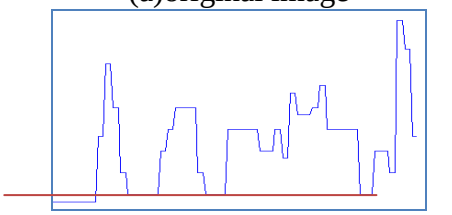
Figure 5 :word segmentation

3:3:4 Character Segmentation

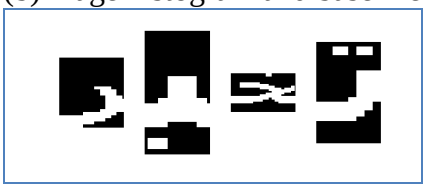
The steps for characters segmentation algorithm are :

Step 1: Count the baseline (more row contains the white pixel of word image
Step 2: Count the white pixel in each column.
Step 3: Find the position of char segment in which value equal baseline.
Step 4: Copy the pixels in the bounding box and saving in separate file.
You can see application of algorithm in figure 6

(a)original image



(b)image histogram and baseline



(c)character segmentation

Figure (6): character segmentation

4. Conclusion:

In this paper, we propose an efficient method to deal with background complexity in Arabic text that present in subtitle of video film. We using image processing concepts and project profile to segmentation lines ,words and character this method we can produce quite clear text for Arabic text we can recognize it by any OCR method .

References:

[1] Zhang Y., "Advances in Image and Video Segmentation", IRM Press, ISBN 1-59140-753-2, 2006.
[2]Michael A. Smith and Takeo Kanade,"Multimodal VideoCharacterization and Summarization ",Kluwer Academic Publishers,2005.

[3] Abbas H. AL-Asadi and Thika Ali H. Subber,"Arabic text Extraction from video images",Journal of Basrah Researches ((Sciences)) Vol.(39). No.(4) A, 2013 B vvv,nncd
[4] Jung K., Kim K. I., and Jain A. K., "Text Information Extraction in Imagesand Video :A Survey", Elsevier Ltd Pattern Recognition Society, Vol. 37,pp. 977–997, 2004.
[5] Hassin A. H., Tang X., and Liu J., "Printed Arabic Character Recognition Using HMM" Journal of Computer Science and Technology, V19(4), pp. 538-543, 2004.
[6] Dr. Gheith A. Abandah and Prof. Mohammed ZekiKhedher,"Printed and Handwritten Arabic Optical Character Recognition Initial Study", The University of Jordan,2004.
[7] Pratt W. K., "Digital Image Processing", WILEY, ISBN: 978-0-471-76777-0, Printed in the United States of America, 2007.
[8]Dr J.-P. Thiran, Prof. H. Bourlard,"text detection and recognition in images and video sequences", Lausanne, EPFL 2003

