جامعة الشرق الأوسط
MIDDLE EAST UNIVERSITY

Enhancing the Performance of the Hybrid Recovery Method by Defining Video Content Type

تحسين أداء طرق الاسترجاع الهجين من خلال تعريف محتوى الفيديو

By

Dalia Yousef AL-Jallad

Supervisor

Dr . Hebah H. O. Nasereddin

A Thesis Submitted in Partial Fulfillment of the Requirements for Master Degree in Computer Information Systems.

Department of Computer Information System

Faculty of Information Technology

Middle East University

January, 2016

المنارة للاستشارات

www.manaraa.com

Authorization Statement

I, Dalia Yousef AL-Jallad, authorize the Middle East University to supply a copy of my thesis to libraries, establishments or individuals upon their request.

Name : Dalia Yousef AL-Jallad

Date : 28 - 1 - 2016

Signature : ....

المنارة للاستشارات

www.manaraa.com

Examination Committee Decision

This is to certify that the thesis entitled "Enhancing the Performance of the Hybrid Recovery Method by Defining Video Content Type" was successfully defended and approved on 9rd January 2016.

Examination Committee Members	Signature
(Head of the Committee) **Dr. Sharefa Murad** Department of Computer Science Middle East University	28/1/2016
(Supervisor) **Dr . Hebah H. O. Nasereddin** Associate Professor Middle East University	28/1/2016
(External Committee Member) **DrMuhammad Shkukani** Department of Computer Science Applied Science University	28/1/2016

المنارة للاستشارات
www.manaraa.com

Acknowledgment

This thesis would not have been possible without the support of many people. Many thanks to my supervisor, Dr. Hebah Nasseraldin, who read my numerous. Also thanks to my colleagues, for their help and support. Thanks to the Middle East University for giving me the enhance to be one of its students. Finally, thanks to my parents, siblings and my numerous friends who endured this long process with me, always offering support and love.

المنارة للاستشارات
www.manaraa.com

Dedication

- To my parents for their support.
- To my Brothers for support and patience me.

المنارة للاستشارات
www.manaraa.com

Table of Contents

المنارة للاستشارات
www.manaraa.com

Chapter Four: Experimental Results

Chapter Five: Conclusion and Future Works

المنارة للاستشارات
www.manaraa.com

List of Figures

المنارة للاستشارات

www.manaraa.com

المنارة للاستشارات
www.manaraa.com

List of Tables

للاستشارات المنارة

www.manaraa.com

List of Abbreviations

Abbreviation	Meaning
ASR	Automatic Speech Recognition
CC	Connected Components
DCT	Discrete Cosine Transformation
FN	False Negative
FP	False Positive
GUI	Graphical User Interface
HMM	Hidden Markov Model
OCR	Optical Character Recognition
PDF	Probability Density Function
PNG	Portable Network Graphics
SVM	Support Vector Machine
TIE	Text Information Extraction
TN	True Negative
TP	True Positive
WWW	World Wide Web

www.manaraa.com

ABSTRACT

Enhancing the Performance of the Hybrid Recovery Method by Defining Video Content Type

Prepared By:

Dalia Yousef AL-Jallad

Supervised By:

Dr . Hebah H. O. Nasereddin

Online lecture video repositories brings many benefits for students and researchers in different fields of science such as reducing the time and the cost of reviewing lectures in their home or businesses. Indexing and archiving these lecture videos in such repositories are represented as one of the most important processes to ease the way for users to retrieve the target lecture videos from a thousands of stored videos. Thus, several approaches are available to ease the searching process such as searching for keywords in the lecture video title, description, or visitors' comments.

The main aim of this thesis is to study the possibility of classifying lecture video files based on its contents in order to find the suitable text detection algorithm and assign the suitable Optical Character Recognition (OCR) engine for text recognition that guarantees an acceptable performance for extracting the textual text from lecture video files in order to be used for indexing, archiving, and retrieving processes in online video repositories. To be able to meet this aim it has been outlined the different types of lecture video types, the different types of thresholding techniques, and the different types of OCR engines. The findings of literature review have been deployed to design

المنارة للاستشارات
www.manaraa.com

several experiments for this research. Three main experiments were conducted to detect the effects of lecture video type on the similarity between video frame images, the number of key frames, the estimated threshold value, and the performance in term of recall and precision for the OCR engine.

The results showed that the type of lecture video have an effect on the number of key frames and the similarity between them. Thus, the presentational lecture videos showed dissimilarity and discontinuity points between frame images which was responsible to generate a large number of key frames. Hence, the similarity in presentational lecture videos was less than 40% compared with the handwritten lecture videos which achieved a bounded similarity between 40% - 80% which were represented as a balanced similarity. Furthermore, the spoken text lecture videos achieved the highest percent of similarity that exceeds 80% between video frame images.

From performance point of view, the results of TP, FN, and FP was gathered to compute the precision and recall of using the selected OCR engines with the selected thresholding algorithms. Thus, the results showed that using the Tesseract OCR engine with the selected thresholding algorithms achieved the best results compared with the others. Hence, the recall results were 0.79, 0.9, and 0.92 respectively in term of iterative, OTSU, and the proposed thresholding techniques. From precision point of view, the results showed 0.89, 0.96, and 0.98 respectively.

Keywords : Online lecture video, optical character recognition , Otsu algorithm.

المنارة للاستشارات

www.manaraa.com

الملخص

تحسين أداء طرق الاسترجاع الهجين من خلال تعريف محتوى الفيديو

اعداد:

داليا يوسف الجلاد

اشراف:

د. هبة حسن ناصر الدين

جلبت مخازن ومنتديات محاضرات الفيديو على الانترنت العديد من الفوائد للطلبة والباحثين من شتى حقول العلوم كتقليل الوقت المستهلك والتكلفة المنفقة على مراجعة المحاضرات في بيوتهم أو حتى في أعمالهم. تعتبر فهرسة و أرشفة محاضرات الفيديو في مثل هذا النوع من المخازن الالكترونية من أهم العمليات التي يتم اجراؤها على هذه الملفات من أجل تسهيل طريقة استرجاع الزائرين للملفات المطلوبة من بين الاف ملفات محاضرات الفيديو المخزنة في مخازن الفيديو الالكترونية. لذا، توافرت مجموعة من المبادئ والطرق لتسهيل عمليات البحث عن الفيديو كمطابقة الكلمات المفتاحية مع عنوان ملف محاضرة الفيديو ، أو وصف الفيديو ، أو من خلال مطابقة الكلمات مع تعليقات الزوار على محاضرة الفيديو.

الهدف الرئيسي من هذه الرسالة هو دراسة امكانية تصنيف ملفات محاضرات الفيديو بناءا على محتواها من أجل استخدام الخوارزمية المناسبة للتحقق من وجود النصوص داخل الملف وبالتالي تعيين محرك (OCR) الانسب لاكتشاف واستخراج النصوص من ملفات محاضرات الفيديو بما يضمن أداءاً مناسبا بحيث يتم استخدام هذه النصوص في عمليات فهرسة وأرشفة محاضرات الفيديو في قواعد البيانات الخاصة بمستودعات محاضرات الفيديو الالكترونية. ولكي نقابل الهدف الرئيسي من الرسالة تم عرض أنواع محاضرات الفيديو المستخدمة، وعرض مختلف أنواع مخمنات درجة العتبة (Threshold) المستخدمة في هذا السياق ، بالاضافة الى عرض مجموعة من أنواع محركات استخراج النصوص (OCR). تم عقد ثلاثة تجارب أساسية من أجل اكتشاف تأثيرات نوع محاضرة الفيديو على التشابه بين اطارات ملف الفيديو بعد التقسيم، وعلى عدد الاطارات المفتاحية داخل ملف محاضرة الفيديو، وعلى قيمة درجة العتبة التي يتم تخمينها، وعلى الاداء من منظور قياس الدقة والاسترجاع لمحركات اكتشاف النصوص (OCR).

المنارة للاستشارات

www.manaraa.com

أظهرت نتائج الدراسة أن نوع محتوى محاضرة الفيديو له تأثير مباشر على عدد الاطارات المفتاحية بعد تقسيم ملف الفيديو وعلى التشابه بين الاطارات. لذا، محاضرات الفيديو المعتمدة على العروض التقديمية أظهرت نقاط عدم تماثل وانفصال كثيرة بين اطارات ملف الفيديو مما أدى الى توليد عدد كبير من الاطارات المفتاحية. ولهذا السبب أظهرت النتائج أن نسبة التشابه بين اطارات هذا النوع كانت أقل من 40%. وأما بالنسبة لفيديوهات الشرح المكتوبة فلقد أظهرت توازنا في مدى التشابه بين الاطارات فقد تباينت النتائج بين نسبة تشابه محدودة تزيد عن 40% وتقل بنفس الوقت عن 80%. وأما بالنسبة للفيديوهات التي تحوي محاضرات صوتية فلقد زادت نسبة التشابه بين الاطارت بما يزيد عن 80% وبالتالي الى عدد قليل من الاطارات المفتاحية في مثل هذا النوع من محاضرات الفيديو.

وأما من وجهة نظر الاداء، فقد تم تجميع نتائج (TP) و (FN) و (FP) لغاية حساب الدقة ومقياس الاعادة من خلال استخدام محركات الاكتشاف المختارة و خوارزميات تخمين قيمة العتبة المختارة أيضا في هذه الدراسة. لذا، أظهرت النتائج أن استخدام محرك اكتشاف النصوص (Tesseract) مع القيمة المخمنة من خلال الخوارزمية المحسنة في هذه الدراسة هو الافضل مقارنة مع تهجين الخوارزميات مع المحركات الاخرى المختارة في هذه الدراسة. ولهذا السبب كانت نتائج الاعادة 0.79، 0.9، 0.92 على الترتيب. وكانت نتائج الدقة 0.89، 0.96، و 0.98 على الترتيب أيضاً وهي الاعلى تقريبا بناءا على التجارب المقترحة في هذه الدراسة.

الكلمات المفتاحية : محاضرات الفيديو على الانترنت , محركات استخرج النصوص , الفيديوهات التعليمية .

المنـــارة للاستشارات
www.manaraa.com

CHAPTER ONE

Introduction

1.1. Overview

This chapter explains background about online digital video repositories, digital video processing algorithms, digital image processing, digital video segmentation, and performance evaluation technique. This chapter shows the problem statement of this research, author's contribution, and the outline of thesis chapters.

1.2. Online digital video repositories

Online digital video and imagery repositories are increasing nowadays due to the revolution in designing huge storage media capacities, the increasing of transmission bandwidth, and the new generation of compression algorithms.

Recently, many organizations own digital video archives, and these archives can be classified based on organization's interests such as commercial videos, entertainment videos, medical videos, and educational videos. Therefore, managing these video archives has the highest priority before publishing such online repositories on the internet.

The added value for any new online digital video repository should be related directly to the advances of compression, transmission, storage, archival, indexing, querying and browsing technologies (Tekinlap, S., and Altan, A., 2013). Therefore, visitors will be capable to find the needed video files in an easy and efficient manner.

The educational organizations have created their own repositories for online education. Hence, the education is accessible for a wide community of potential students (Martinze-Villaronga, et al., 2014). Therefore, the way of producing educational videos is by recording lectures during class room presentations, in order to share it over an online video repository to be accessible by students. Thus, a hundreds of lecture videos that are

المنارة للاستشارات
www.manaraa.com

covering several fields are stored and classified to be searchable materials for student's queries (Kamabathula, V., and Iyer, S., 2011).

Lecture videos are posted for many purposes in order to be versatile learning resources. Thus, it can be as additional material for conventional courses, primary material for distance learning coursework's, or a reference material for public community learning (Tuna, T., et al., 2011). In this context, thousands of complete course are posted which is evidence on the popularity of lecture videos.

Furthermore, several universities like Massachusetts Institute of Technology (MIT) developed and maintained an online lecture video portal that is responsible for aggregate and curate videos produced by institute's offices, laboratories, and centers. Thus, students will be capable to access these video files anytime in order to review their lectures. Therefore, they added a search property in the portal that is used to find the relevant video file based on student's query. The search tool in this web based system was built to compare student's queries with the video meta-data keywords. In contrast, the stored lecture videos was indexed and archived based on keywords of meta-data that describe video contents. Figure 1.1 shows the MIT video homepage.

المنارة للاستشارات

www.manaraa.com

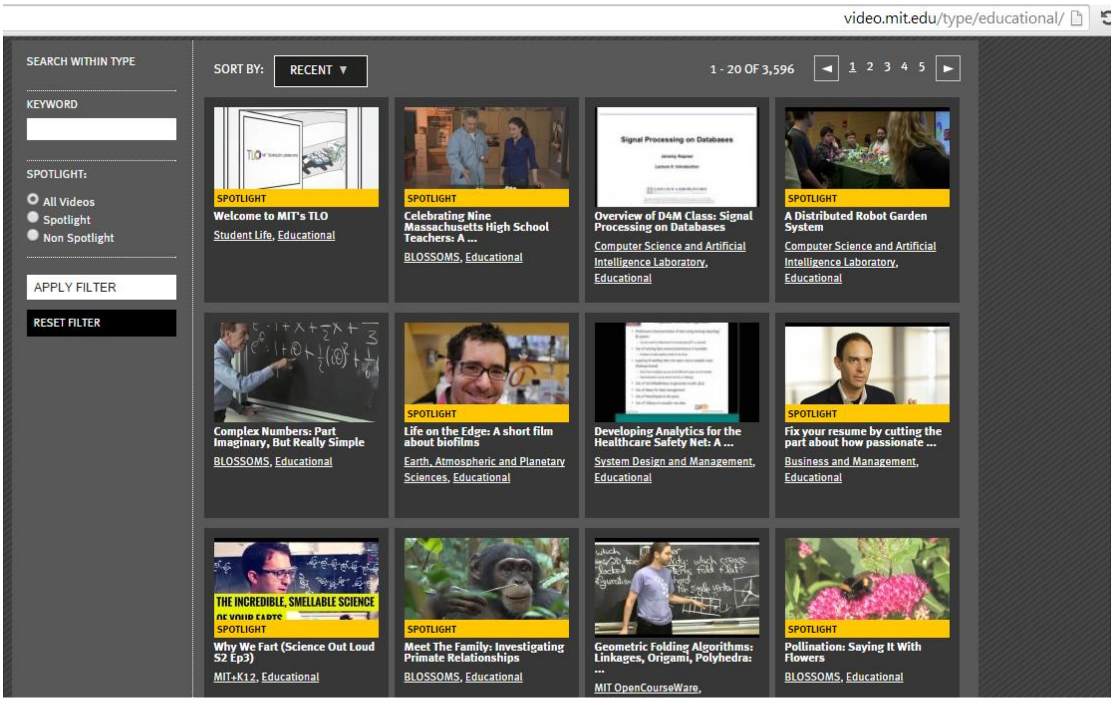


Figure 1.1 MIT Open Courseware video portal homepage (*video.mit.edu*)

COURSERA is another online lecture video repository which is represented as an educational platform that has partners with several universities and institutions all over the world. Moreover, this portal provides over 12000 of lecture videos in different fields of science. Figure 1.2 shows the COURSERA homepage.

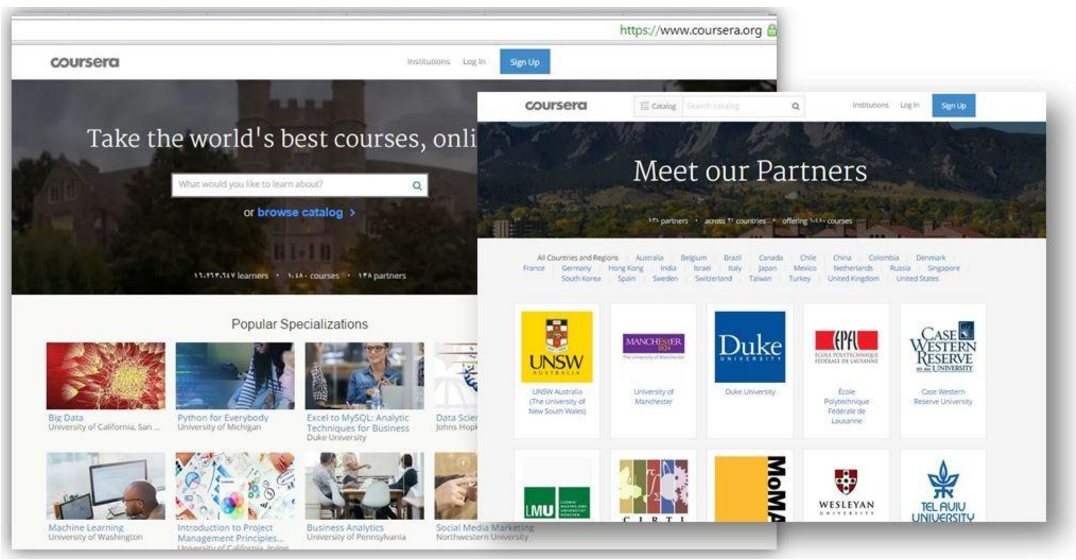


Figure 1.2 COURSERA lecture video portal (*coursera.org*).

المنارة للاستشارات
www.manaraa.com

1.3. Searching techniques in lecture videos

Online video repositories provide searching techniques based on video meta-data or lecture video contents (Richter, F., et al, 2014). In meta-data based technique, the search keywords will be matched with the meta-data that is related for each lecture video file. Meta-data is considered as the textual data that is used to describe lectures such as the title, the description, and visitor's comments (e.g. students, instructors, or public community) (Lienhart, R., and Effelsberg, W., 2000). In this technique the retrieved lecture video files are matched based on the search keywords without finding the location of these keywords in video contents. Figure 1.3 shows YouTube video meta-data.

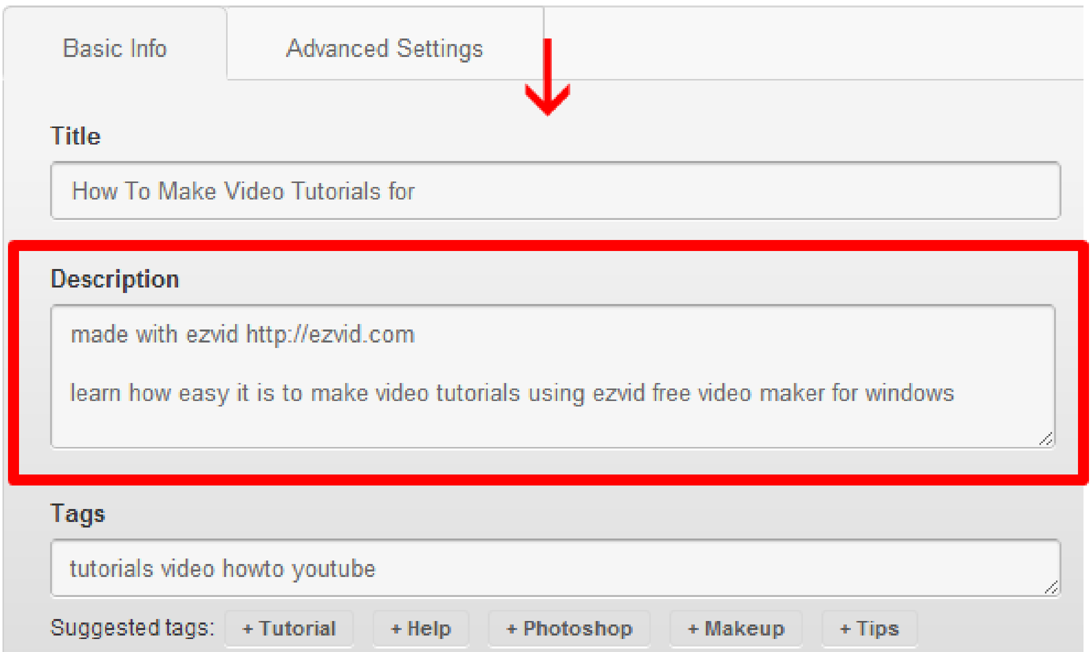


Figure 1.3 Tutorial video meta-data (Title, Description, and Tags) in YouTube (*youtube.com*)

In content based searching technique, the searching process is responsible to find the location of the search keywords in the contents of lecture videos (Patel, B., and Meshram, J., 2012). Hence, lecture video may contains a portion that shows lecturer

www.manaraa.com

speech, a portion that shows presentation current slide, or a portion that shows the hand written notes on the chalkboards or whiteboards. Figure 1.4 shows the presentation current slide and lecturer speech portions for a lecture in University of British Columbia (UBC) in Canada. This technique is capable to extract the meta-data from the appropriate video portion and matching each search keyword with the extracted information from video contents (Kamabathula, V., and Iyer, S., 2011). Thus, content based techniques still face some problems in time consumption, system performance, and accuracy. Furthermore, several techniques were proposed to automate the extraction of text like Optical Character Recognition (OCR) technology, Edge Detection technology, and Automatic Speech Recognition (ASR).

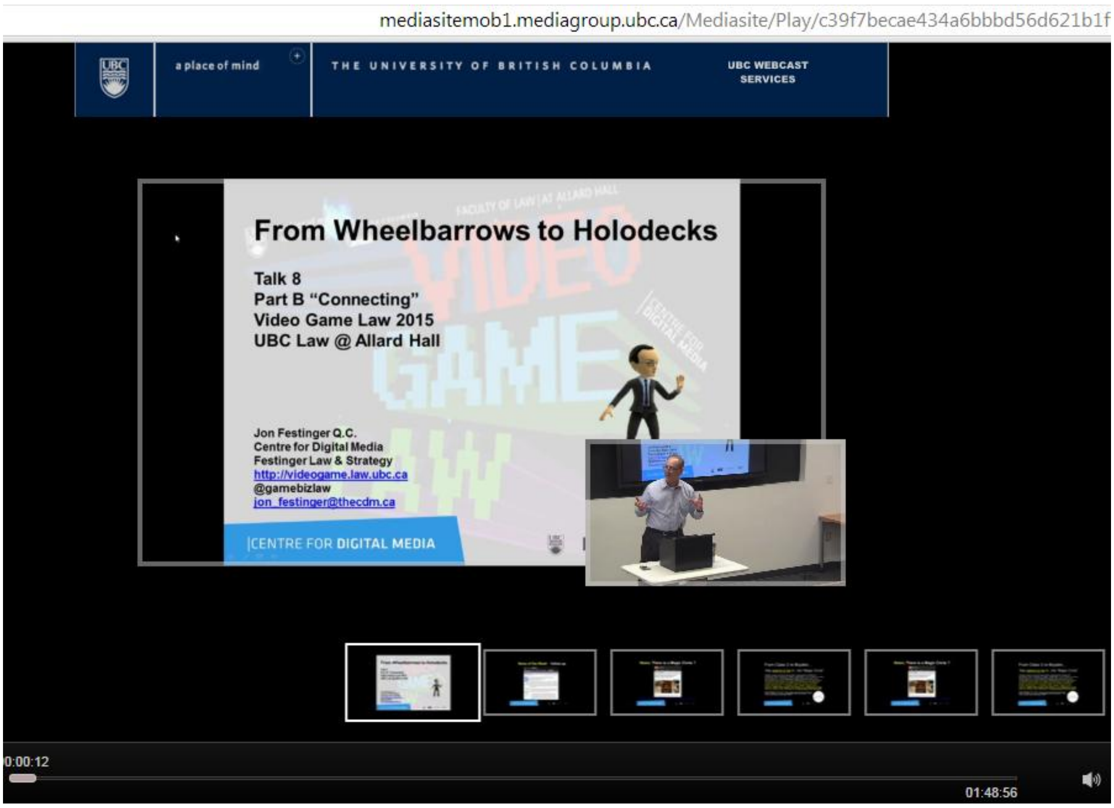


Figure 1.4 Presentation Slide and Lecturer Speech Portions (videogame.law.ubc.ca)

المنارة للاستشارات
www.manaraa.com

1.4. Video data structure

Ding, H. et al. imposed that the structure of video file contains frames, shots, scenes, and events. They showed that each video file is consisted into shots which are a set of similar and consequent frames, which could be recognized using segmentation techniques. In this context, each video file has two types of scenes which are simple and composite. Thus, simple scene can represent an event, but in composite one multiple scenes are capable to represent an event (Ding, H., et al, 2010).

Therefore, any video shot represented as a series of interrelated frames that was taken by a single camera to represent continues action in time and space. Hence we can say that each video shot represent a segment in video analysis. Figure 1.5 shows an illustration of video data structure.

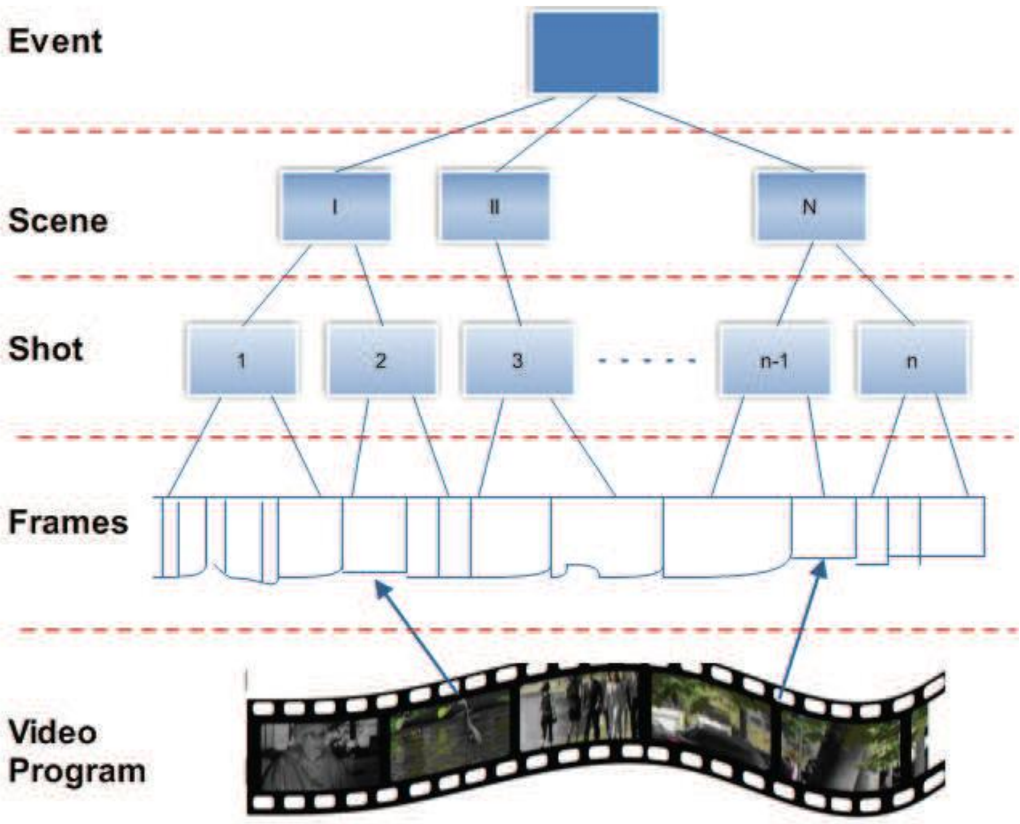


Figure 1.5 Video Data Structure (Ding, H., et al., 2010)

www.manaraa.com

1.5. Text Information Extraction (TIE)

Basically, the embedded text in videos is classified as a scene text and artificial text. The text that appears in video scenes and does not represent information about the content of video images called the scene text (Ye, Q., and Doermann, D., 2013). In contrast, the artificial text is laid over the image as a superimposed caption. Hence, artificial text represents a region in video has high contrast with high frequencies (Leon, M., et al, 2013). Therefore, it is often used as a successful key for video indexing and segmentation. Figure 1.6 illustrates the differences between scene and artificial texts in videos.

Figure 1.6 The differences between scene text and artificial text.

Jung, K. et al find out that the architecture of any TIE system contains four separated stages from importing video till exporting textual data from its contents. Thus, these stages are text detection, text localization, text extraction and enhancement, and text recognition (Jung, K., et al, 2004). Figure 1.7 shows the general structure of TIE system.

المنارة للاستشارات
www.manaraa.com

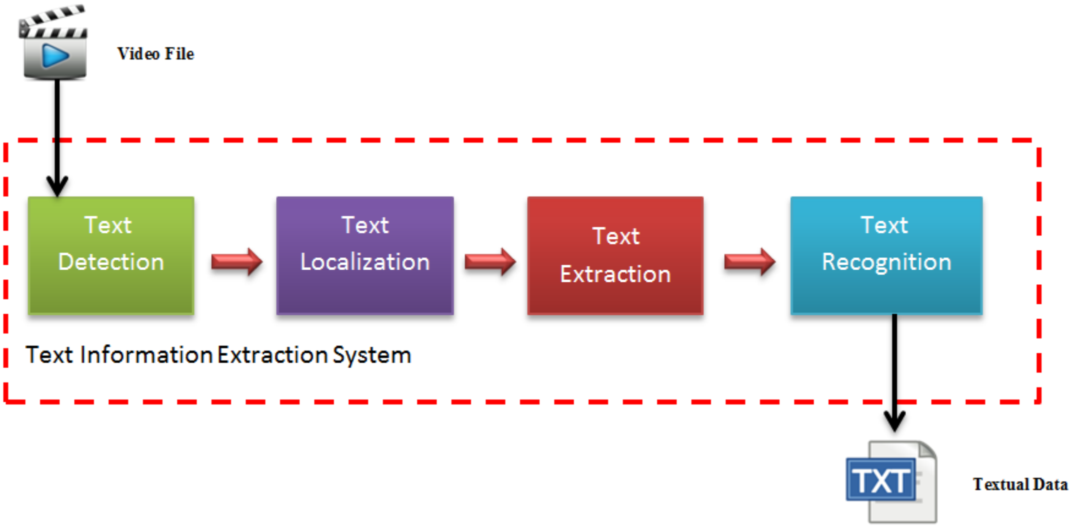


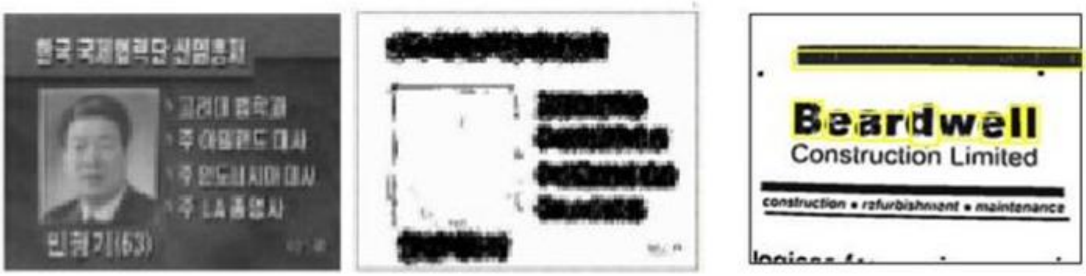
Figure 1.7 The general architecture of TIE system (Jung, K., et al, 2004)

Jung, K. et al explained the differences of TIE stages. Hence, they refer text detection stage for the process of determination the occurrence of text in specific frame. They defined text localization as the process of determining the location of text in image (i.e. the current image in sequence) and generating the bounding boxes around the text. In text extraction stage, they showed that the text components are segmented from background in order to feed the OCR engine in text recognition stage (Jung, K., et al, 2004).

Ngo, C., and Chan, C. show that video TIE systems should contains three primary stages which are text detection stage that is responsible to detect the regions of text in video, text segmentation stage that is responsible for segmenting the detected text regions in order to export the segments as binary images, and text recognition stage that is responsible to convert the text in the video frames into ASCII code (Ngo, C., and Chan, C., 2004).

المنارة للاستشارات
www.manaraa.com

In this context, the literature showed several techniques to implement text detection from videos. Thus, most of them were classified into spatial text detection, temporal text detection, image binarization-segmentation, and character recognition (Anthimopoulos, M., et al, 2010). Furthermore, many research works found out that the spatial text detection is the crucial stage between them. Thus, this approach was divided into two categories for text localization stage which are region based category and texture based category. Figure 1.8 illustrates the differences between texture based and region based approaches.



a) Texture Based Approach

b) Region Based Approach

Figure 1.8 Text localization stage categories a) texture based b) region based approaches

In texture based approach the video image is scanned at different scales using a sliding window. Thus, the video image later on is classified as a text or non-text based on texture-like feature. Kim, K., et al showed that the idea of texture based approach is to find the distinct textural properties from non-text regions (i.e. background) (Kim, K., et al, 2003).

In region based approach the properties of color or gray-scale in text region is considered as the core idea of finding the differences between them and the corresponding background. Hence, two primary techniques are used in this approach which is Connected Components (CC) that is responsible to group small components into successful larger components until text region is identified, and edge based technique which concentrates on the high contrast between text and background (Leon, M., et. al, 2013).

المنارة للاستشارات
www.manaraa.com

Anthimopoulos, M., et al. explained region based in video images as the process of grouping pixels that belongs to the same character based on the homogeneity of color, the strong edges between characters and background, and stroke filters (Anthimopoulos, M., et al, 2010).

1.6. Problem statement

In the last years, the huge amount of educational video data on the World Wide Web (WWW) is growing rapidly due to the quick development of recording technology. Moreover, when users have found related video data, it is still difficult most of the time to judge whether a video is useful by only glancing at the title or description of video which is often brief. Therefore, it takes a long time for users to filter through all videos content until they find what they are looking for, which lead users to be unsatisfied with the video retrieval systems. Current video retrieval systems rely heavily on manually created text metadata includes a video title and video description if available; that leads researchers to find efficient video retrieval techniques to recover the videos content automatically. Previous studies results revealed that automatically recovered slide text and spoken text contain different content with varying error profiles. Experiments demonstrate that automatically extracted slide text enables higher precision video retrieval than automatically recovered spoken text and slide text, and because of the majority of the studies are using both text extracting algorithms (slide text, spoken text) which means more and more irrelative retrieved video. This research concentrate on applying several methods using different techniques in order to index lecture video based on its contents. Thus, finding the suitable recall and precision of the retrieved lecture video results based on visitor's queries. This research focuses on text recognition from lecture videos and compare the performance of several well-known algorithms.

للاستشارات المنارة

www.manaraa.com

1.7. Research questions

Problem accomplished by answering the following questions:

1. What are the issues that are affected with the lecture video type?
2. What are the issues that affect the thresholding value in text detection level?
3. How do we select the suitable text extraction engine based on lecture video type?
4. What are the issues that affect the recall and precision of the retrieved results?

1.8. Research objectives

The main aim of this research is to present a suitable solution to enhance the way of video retrieval. The proposed method in literature objectives to provide more efficient and accurate performance to extract the text from the lecture videos. The efficiency of these methods rise from its ability to use the most appropriate text extracting algorithm among OCR, ASR, and other algorithms, by segmenting the videos and using the specified segments of each video to test the accuracy of the retrieved text compared to the real text.

1.9. Motivation

Greater number of content-based video search methods has been proposed recently. Thus, the automated text detection methods that are applied on slide text and spoken text in lecture videos. Furthermore, the results of literature studies showed that slide text enables higher accurate video retrieval than spoken text that leads to large amount of irrelative retrieved video. Therefore, the researchers in the domain of text detection from video images; know the meta-data based techniques in video retrieval results may return irrelevant lecture videos in online digital video repositories as well as the results are returned without pointing to the search keywords in the contents of lecture video. The proposed research aspires to provide more efficient and accurate method to

المنارة للاستشارات

www.manaraa.com

extract the text from lecture videos. The novelty of this study is to recommend the most appropriate text extracting algorithm for any online lecture video repository.

1.10. Contribution

This thesis contributes the following issues:

- Detecting factors that affect the similarity of between lecture video frame images.
- Explaining the criteria that were used to identify lecture video type in term of its contents.
- Detecting the factors that affect the number of key frames in lecture videos.
- Detecting the factors that affect the quality of binary image in text detection level.
- Explaining the criteria for choosing the thresholding algorithm and text recognition engine based on lecture video type to achieve the suitable performance in term of recall and precision.

1.11. Methodology

The methodology that was used to develop the proposed method contains the following phases:

- Study and analysis phases.
- Design and implementation phase.
- Evaluation phase.

Study and analysis phase

In this phase the work was started based on the problem statement which was for classifying lecture videos based on its contents. Online video repositories provide several types of lecture videos such as presentational, handwritten, and spoken text lecture videos for students and visitors. These lecture video types had different properties as well as different techniques are available in the literature to detect and extract textual information from it. Therefore, finding the suitable thresholding and recognition techniques should

المنارة للاستشارات
www.manaraa.com

reserve the best performance for the whole TIE system. The acquired information and knowledge from this phase was as follows:

- Studying the specifications of each lecture videos type in any online lecture video repository.
- Understanding the effects of lecture video type on the similarity between lecture video frame images.
- Understanding the effects of lecture video type on the estimated thresholding value which affect the binary image and its quality.
- Studying the performance testing with regard to recall and precision in term of the automatically selected words with bounding boxes and the correctly recognized words.

Design and implementation phase

This research was carried out a case study which covers building three experiments; we decide to use VLC for splitting lecture videos into frame images, MATLAB for implementing the code of iterative threshold and OTSU algorithm. The following steps were used to meet thesis's goal:

- Splitting video into frame images.
- Extracting the color histogram for each frame image and compute the needed calculations (Chapter three).
- Extracting the number of key frames in lecture video and compute the needed calculations (Chapter three).
- Estimating the threshold values using the selected threshold algorithms (Chapter3).
- Computing the performance with regard to recall and precision for the selected OCR engines.

المنارة للاستشارات
www.manaraa.com

Evaluation phase

Three experiments have been designed to evaluate the similarity and the number of key frames in content analysis level, evaluate the threshold of iterative, OTSU, and the enhanced OTSU thresholding algorithms in the text detection level, and evaluating the performance of OCR-AD, Free-OCR, and Tesseract OCR engines in the text recognition level. The similarity in the first experiment was used to evaluate the identification of lecture video type, the number of missed pixels was used to evaluate the thresholding algorithm, and recall with precision performance metrics were used to evaluate the performance of text recognition engines.

1.12. Organization of the research

Chapter 1 – Introduction: This chapter provides an overview of the problem statement, contribution, and objectives to meet.

Chapter 2 – Literature review: This chapter aimed for explaining the previous studies, and the different technologies present in text detection and recognition in digital images.

Chapter 3 – The proposed method and experiments' design: This chapter describes the design of the test performed to measure the similarity between video frame images and the estimation of threshold value.

Chapter 4 – Experimental results: This chapter evaluates the results obtained during the implementation of the different experiments.

Chapter 5 – Conclusions and future research works: This chapter summarizes the entire study as well as it gives a critical point of view and some recommendations for future research works.

المنارة للاستشارات

www.manaraa.com

CHAPTER TWO

Literature Review and Background

This chapter shows a collection of the most relevant work in the literature that relate to the scope of this research. This literature review covers concepts that have been addressed in this research, namely, text detection in video lectures, OCR technology, ASR technology, and digital video processing. Finally, shows the software that have been used in this research.

- **Text detection**

Jeong, H. et al. developed a method to automatically detect slide changes in lecture videos. Their idea focused on capturing the regions of slide images that are identified by video frames. These regions are invariant to image scaling and rotation. Thus, they used these regions to compare the similarity between frames. In this context, they found out that if the similarity is smaller than the threshold, slide transition will be detected. For that purpose, they estimated the threshold based on mean and standard deviation of sample frames similarities. Consequently, their experimental results showed greater accuracy than template based and histogram based algorithms. As well as, their method had approximately 86% in detecting slide transitions (Jeong, H. et al., 2014).

The text which is displayed in lecture video files are closely related to the content of video file. As well as, it is used to index these videos especially in video portals. Yang, H. et al. found out an approach for automatic lecture video indexing based on OCR technology. They developed a novel video segmenter for automated slide video structure analysis and weight called Discrete Cosine Transformation (DCT), which was based on text detector. Hence, they used dynamic image contrast/brightness approach to enhance the image quality. They randomly choose 20 lecture videos from different

المنارة للاستشارات
www.manaraa.com

lecturers with varying in layout and font styles. Furthermore, they used 180 video frames which were varied between (640 x 480) pixels and (1024 x 768) pixels as a frame size. Their results of text detection were classified based on pixels and bounding boxes for both recall and precision. Therefore, their results for video segmenting and text detection and recognition algorithms are suitable for content based video indexing and retrieval (Yang, H. et al., 2011).

Leon, M. et al. proposed a method that take the advantage of texture and geometric features to detect the caption text. Hence, their technique was tested in two cases, one formed by news and sport event videos, and the other one by sport event videos only. Therefore, the result in the first case showed 249 caption text objects were extracted from a set of 150 images with a text different in size and color. The complexity in this case was in background textures. In contrast, the second case results showed 2063 caption text objects were extracted from 640 key frames. Hence, by comparing their experimental results in both cases they found that the number of false positive is very high due to the presence of advertising panels in images in the second case (Leon, M. et al., 2013).

Shivakumara, P. et al. proposed a method on the Lapacian in the frequency domain of video text detection. Hence, their method was able to handle text of arbitrary orientation. They supposed to make image filtering using Fourier-Lapacian for input image. Thus, k-means clustering was employed to identify the candidate text regions based on the maximum differences. They used text string straightness and edge density for false positive elimination. Hence, their experimental results showed that text extracting could be for both horizontal and non-horizontal orientation in lecture videos (Shivakumara, P. et al., 2011).

المنارة للاستشارات
www.manaraa.com

Sharma, N. et al. developed a method for arbitrarily- oriented text detection in video based on dominant text pixel selection, text representative and region growing. Their method focused on using gradient pixel direction and magnitude corresponding to SOBEL – which is edge detection engine in MATLAB – edge pixel of input frame to obtain dominant text pixels, and they pointed out that as a text representative. However, they eliminated the broken segments of each text representative in order to get candidate text representative. Hence, the perimeter of candidate text representative was growing alone the text component which they assign it as word patches. Moreover, the word paths were expanded in the same direction in SOBEL edge map to group neighboring word paths and to restore missing text information. The experimental results of their method outperformed in arbitrarily oriented, non-horizontal and horizontal text frames (Sharma, N. et al., 2012).

Zhao, X. et al. found out a corner based approach to detect text and caption from videos. Their hypothesis based on the observations of existence of corner points in character especially in text and caption. For that purpose, they used several discriminative features to describe the text region formed by the corner points. Their system was capable to detect video text with high precision and efficiency. Hence, discriminative features were used for text detection on the base corner points. Moreover, they designed a model to detect moving captions from video shots. Their experimental results showed that over 90% of detection ratio was attained (Zhao, X. et al., 2011).

Ngo, C. and Chan, C. found out an approach that was capable to detect and segment text in videos. Thus, their technique was capable to adopt appropriate operations for video frames of different modalities by classifying the background complexity. They tried to eliminate noise of images with high density by using repeated shifting

المنارة للاستشارات
www.manaraa.com

operations. In this manner, they used text enhancement technique through highlighting text regions of low contrast images. Hence, they applied a coarse-to-fine projection technique to extract text lines from video frames. Therefore, their system showed that for text detection the proposed scene-dependent analysis could balance both detection rate and false alarm rate. Furthermore, the experimental result of their research showed that commercial OCR package was capable to recognize 72% of the segmented character. In contrast, their approach showed better performance by exploiting the intra-frame relationship in videos (Ngo, C. and Chan, C., 2004).

- **Text recognition**

Cooper M. examined the relative utility of automatically recovered text from both visual and aural channels from lecture videos. He tried to apply video content analysis to detect slides and OCR to obtain their text. As well as, he used ASR to extract spoken text from recorded audio. He designed a controlled experiment with manually created ground truth for both slide and spoken text from more than 60 hours of lecture video. Therefore, he measured the accuracy based on comparing the extracted text with the ground truth text. Hence, his experiment results showed that automatically extracted slide text enable higher precision than automatically recovered spoken text (Cooper, M., 2013).

Matei, O. et al. proposed a method that combines two algorithms an artificial neural networks and K-nearest neighbor. Their approach was differing than other OCR systems because it was based on the angles of digits rather than the pixels. Their experimental results showed that by using neural networks approach for leveraging and k-nearest neighbor for confirmation of digit recognition in noisy environment was very successful. Hence, they found out high percent in accuracy rate of recognizing digital numbers which was 99.3% (Matei, O. et al., 2013).

المنارة للاستشارات

www.manaraa.com

Yang, H. et al. found that the textual data in lecture videos comes in different resolutions and with heterogeneous backgrounds. Therefore, it was difficult to apply OCR technology because its results. For that reason, they added a stage before applying OCR on video which was responsible to separate text from its background. Hence, they achieved this stage by suitable image binarization strategy. Thus, they analyzed the content distribution of image skeleton maps in order to approximate text gradient direction. Then they selected a text seed pixel by calculating the average of grayscale value of skeleton pixels. Then they applied automated seed region growing algorithm to obtain text pixels. Their hypotheses outperformed other reference methods for recognizing video text (Yang, H., 2012).

Peng, X. et al. described an approach to extract text from broadcast videos. They concentrated on a candidate block were detected based on edge detection results. They used corners and geometrical features in order to make initial classification which was built using Support Vector Machine (SVM). Their system focused on conditional random field based on method to locate the areas of text from video frames. The experimental results showed that their approach outperformed the sing SVM (Peng, X. et al., 2011).

- **Video indexing and recognition**

Reddy, P. and Devaraju, M. proposed an approach for content based lecture video indexing and retrieval from large lecture video archives. Thus, they applied visual and audio resources of lecture video for extracting content based meta-data. Therefore, they applied automatic video segmentation and key-frame detector to offer visual guidelines for content video recognition. They tried to extract textual metadata by applying OCR technology on key frames. As well as, they applied ASR on lecture video audio tracks.

المنارة للاستشارات
www.manaraa.com

Therefore, the transcript of OCR, ASR, and the detected slide text size were used to adopt keywords extraction (Reddy, P. and Devaraju, M., 2015).

Khan, M. et al. showed that it is important to have a way to index the huge volume of lecture videos in order to ease the searching operations on such video types. Therefore, the designed an experiment based on dividing video into segments and among these segments they identified key-frames in order to extract video contents for mining uses. They applied OCR technology on key-frames of video to extract textual data that was stored into digitized format and could be used as a keywords related to videos. Consequently, the core idea of their system was by developing a model which screenshots the in-between frames from a video. Thus, the snapped frames were arranged according to the replication characteristics. Then they applied OCR for mining text from snapped frames. As a result the extracted textual data were added as keywords for video that can be used for querying purposes in video portals (Khan, M. et al, 2014).

Tuna, T. et al. presented video indexing and keyword search that facilitate access to video content. Their experiment was based on dividing lecture videos into segments. Hence, any change of topics in the same video will be extracted based on the analysis of differences image from a pair of video frame. They proposed to emply bunary search with frame sampling to efficiently analyze long video files. Thus, they proposed to make image enhancement process before applying OCR engine, because of the diversity of colors, font sizes, and background in video key frames. The experimental results showed that the consumed time for extracting text from lecture video with one hour length was 14 minutes on typical desktop machine. The commercial OCR engines in the experiment were Tesserat, GOCR, and MODI. They demonstrated that auto method indexing and search framework are effective and efficient As well as, they found out that inversion and segmentation had a substantial impact on improving text

المنارة للاستشارات
www.manaraa.com

recognition capacity of each of the three OCR engines ever employed (Tuna, T., et al., 2011).

- **Software Tool**

Image processing is related to the operations that can be implemented on digital images to change the nature of an image for many reasons such as enhancing pictorial information for human interpretation purposes, or making digital images suitable for machine perception (Gonzalez, R. and Woods, R., 2008). Therefore, digital image processing involves changing the nature of digital images using computer applications. Hence, in this study MATLAB software tool was used to implement the proposed method and algorithms. Figure 2.1 shows the MATLAB release and version number that was used in this study.

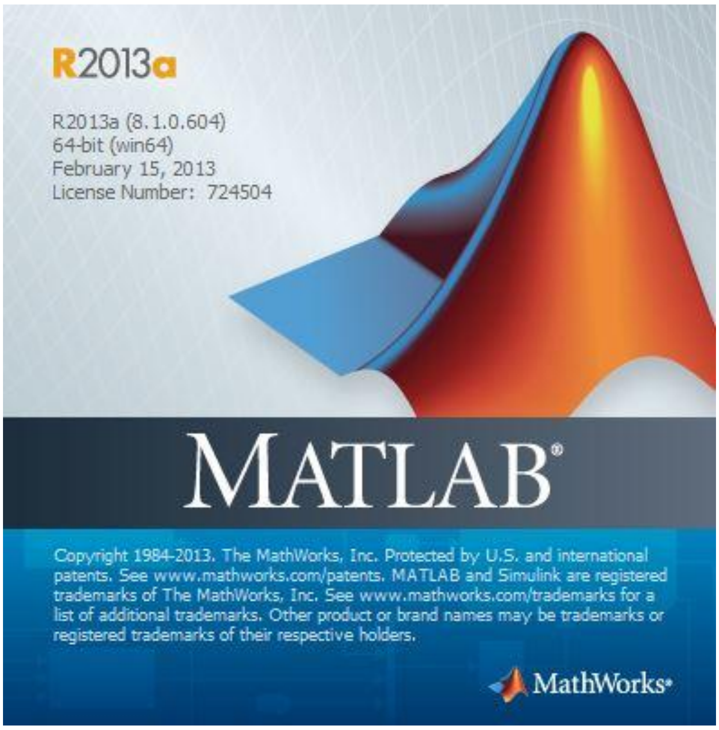


Figure 2.1 MATLAB release and version number for the software package

المنارة للاستشارات
www.manaraa.com

Solomon, C. defined MATLAB as a commercial product that is represented as a programming language that contains a set of expert toolboxes. Thus, it is used worldwide in the scientific, technical, and educational sectors (Solomon, C., 2011). Furthermore, McAndrew, A. explained the basic use of MATLAB as a visualization tool that is used for data analysis because of its capabilities in supporting matrices and matrix operations (McAndrew, A., 2004). The main design of MATLAB is to provide a set of programs for specific tasks and these programs called toolboxes. Hence, this study took into consideration MATLAB's image processing toolbox.

- **MATLAB functions**

MATLAB functions are keywords that take some parameters as input and produce some sort of results as output such as matrix, figure, or graph. In this section, a revision of the main functions was used in this study is discussed. Table 2.1 shows the MATLAB functions that were used in this study.

Table 2.1 The main MATLAB functions.

No.	Function Syntax	Task Description
1.	Imread(*filename*)	Read the image as a matrix of color values for the specified file name.
2.	Im2unit8(*filename*)	Convert the grayscale image to 8 bit unsigned integer
3.	Imhist(*filename*)	Calculate the histogram for the intensity image and display plot of the histogram.
4.	Sum (A)	Return the summation of element of array along the first dimension whose size does not equal 1.
5.	Mean(A)	Return the mean (i.e. average) of the elements of array (A) along the first array whose size does not equal 1.
6.	Std (A)	Return the standard deviation of the elements of array (A) along the first dimension array whose size does not equal 1.
7.	Cumsum(A)	Return the cumulative summation of array (A) starting at the beginning of first dimension array (A) whose size does not equal 1.
8.	Round(A)	Round each element (A) to the nearest integer.
9.	Makedist (*Dist_Type*)	Create a normal distribution using the parameter value.
10.	Pdf (*name, A*)	Return the probability distribution density for one parameter distribution family specified by (name) contains the value for the distribution (A).
11.	Max(A)	Return the largest element of (A).

المنارة للاستشارات
www.manaraa.com

MATLAB program contains three main parts which are the command window, the workspace folder for results, and the resources folder of execution. Figure 2.2 shows the Graphical User Interface (GUI) of the used MATLAB package that was used in this study with command line and resources folder.

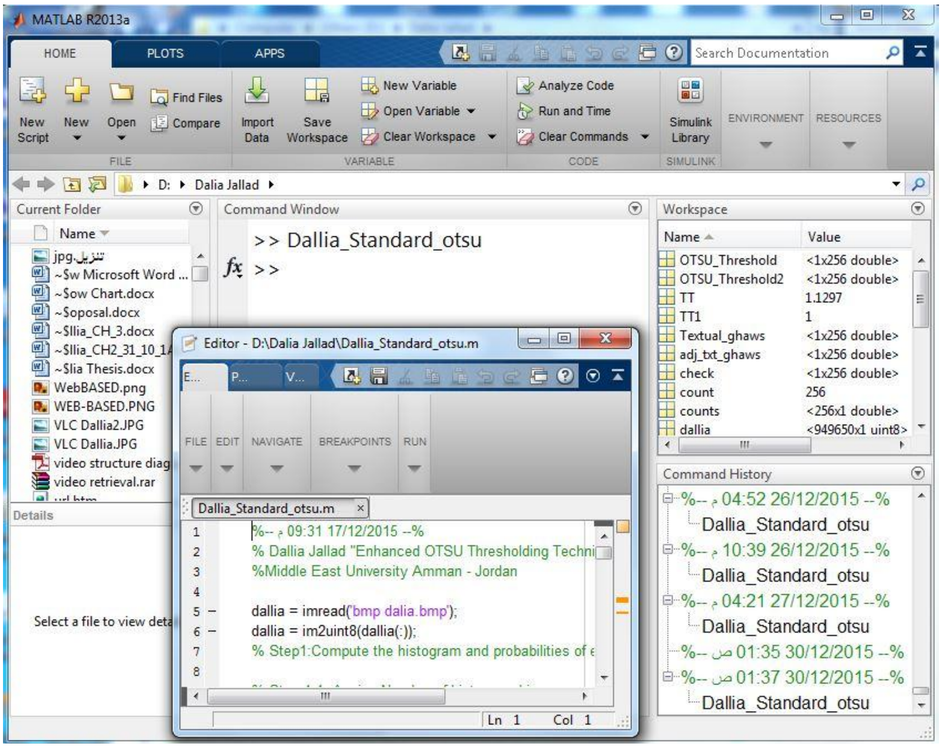


Figure 2.2 The main GUI of MATLAB software with the command window and resources folder and the workspace folder

www.manaraa.com

CHAPTER THREE

The Proposed Method and Experiment's Design

3.1. Overview

This chapter will present a detailed description of the proposed method as well as, discuss the proposed experiment's design, and finally define the evaluation criteria used for each research parameter.

3.2. The proposed method architecture

The main theme of this research is to find the suitable text recognition algorithm based on the contents of lecture video. To achieve this goal, several techniques were used to manage and design an appropriate lecture videos repository that is capable to retrieve lecture videos that match visitor's queries. This method was proposed to meet the best precision and recall measurements among several techniques that were proposed in literature. Hence, the proposed method was consisted to four levels which are lecture video content analysis level, text detection and localization level, text recognition level, and video archiving and indexing level. Figure 3.1 shows the general proposed method levels.

المنارة للاستشارات
www.manaraa.com

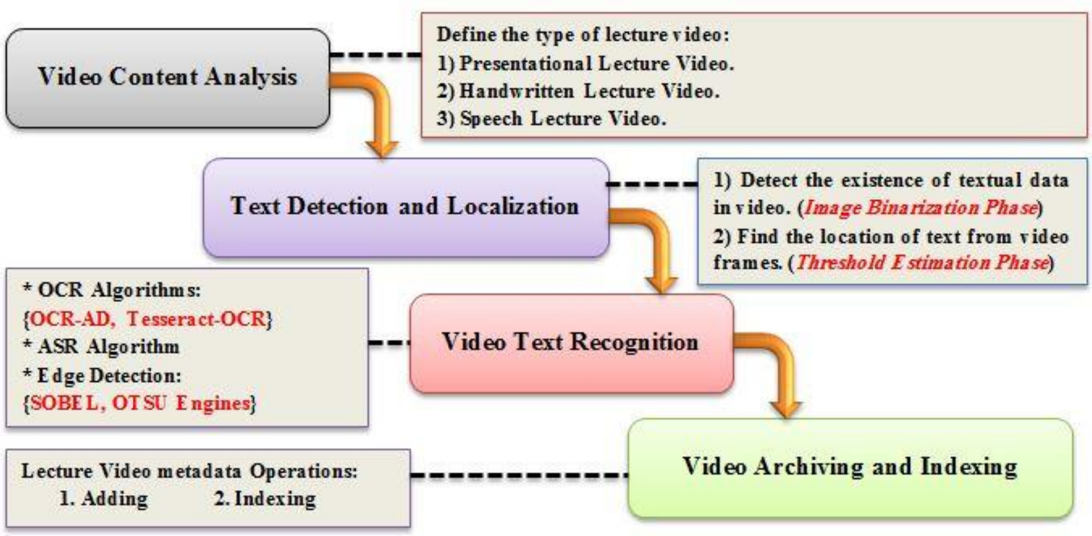


Figure 3.1 The general framework of the general proposed method.

In this method, the contents of lecture video were used to identify the technique of text recognition. Thus, the spoken lecture videos assigned to ASR engine, the presentation lecture videos assigned to two types of OCR engines (i.e. OCR-AD and Tesseract-OCR), and handwritten lecture videos assigned to two MATLAB edge detection engines (i.e. SOBEL and OTSU).

3.2.1 Video content analysis level

The main goal of this level is to classify the lecture video contents into presentational video, handwritten video or spoken lecture video. Thus, to achieve this goal, a technique will apply in order to evaluate the similarity between video frames, and extract video key frames. Several techniques found in the literature to evaluate visual contents in videos such as shot based criteria and color feature criteria. In shot based criteria the first frame is always selected as the first key frame but if we need to find another key frame we have to apply other criteria. In color feature based criteria,

للاستشارات المنارة
www.manaraa.com

the evaluation process is responsible to compare the current key frame with any new frame and if there is a significant change in color a new key frame is selected.

Therefore, the proposed method was designed to evaluate the similarity and extract key frames from videos based on the similarity of visual contents of video frames. Hence, L_n norm similarity metric have chosen to evaluate the similarity or discontinuity between video frames. This metric (equation 1) evaluate the color histogram (i.e. the graphical representation of color values and their occurrences in images) of each frame and compare the histogram of current frame H_c with the histogram of previous frame H_p for a given i number of frames where $F = \{1,2,3,\ldots i\}$, and for each histogram contains n number of color bins. Figure 3.2 shows the L_n norm for evaluating the similarity between video frames.

$$D_{L_{i,n}} = \sum_{F=1}^{i} \sum_{j=1}^{n} \left| H_c(j) - H_p(j) \right| \quad \text{........ (1)}$$

Where:

- ***D:*** *is the difference between color histogram frames.*
- ***H_c:*** *is the histogram of current frame.*
- ***H_p:*** *is the histogram of previous frame.*
- ***i:*** *the total number of video frames.*
- ***n:*** *the total number of color bins*

The normalized difference between any two frame histograms is between 0 and 1. Thus, the value that is close to 0 means that images are similar and those to 1 means that images are dissimilar which an evidence of occurrence of new key frame. In this research, the sample lecture video was downloaded from Washington university lecture

www.manaraa.com

videos website. The duration of sample lecture video was approximately 16 minutes as well as it covers a lecture in computer networks by Dr. David Wetherall. Hence, the process of splitting the sample video into frames was conducted using commercial VLC video player that provide scene filter function which allow user to identify the size, the color mode, and the format of the exported video frames. Figure 3.2 shows the sample video in VLC player.

Figure 3.2 The sample lecture video running in VLC player

Figure 3.3 shows the scene filter function in VLC video player and its properties. In our experiment the ratio of frame recording is based on 16 millisecond which is the threshold of changing images in video players, as well as, the format of the exported frames was chosen as PNG (Portable Network Graphics) format.

المنارة للاستشارات
www.manaraa.com

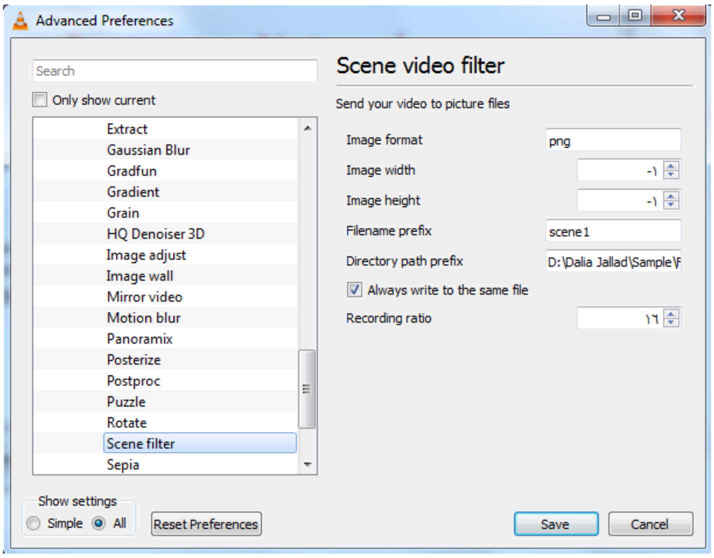


Figure 3.3 VLC scene filter function and its properties

The total number of frames was 193 frames. Figure 3.4 shows the folder that contained the resultant frames. The size of each frame was varied between 9.9 KB to 9.98 KB.

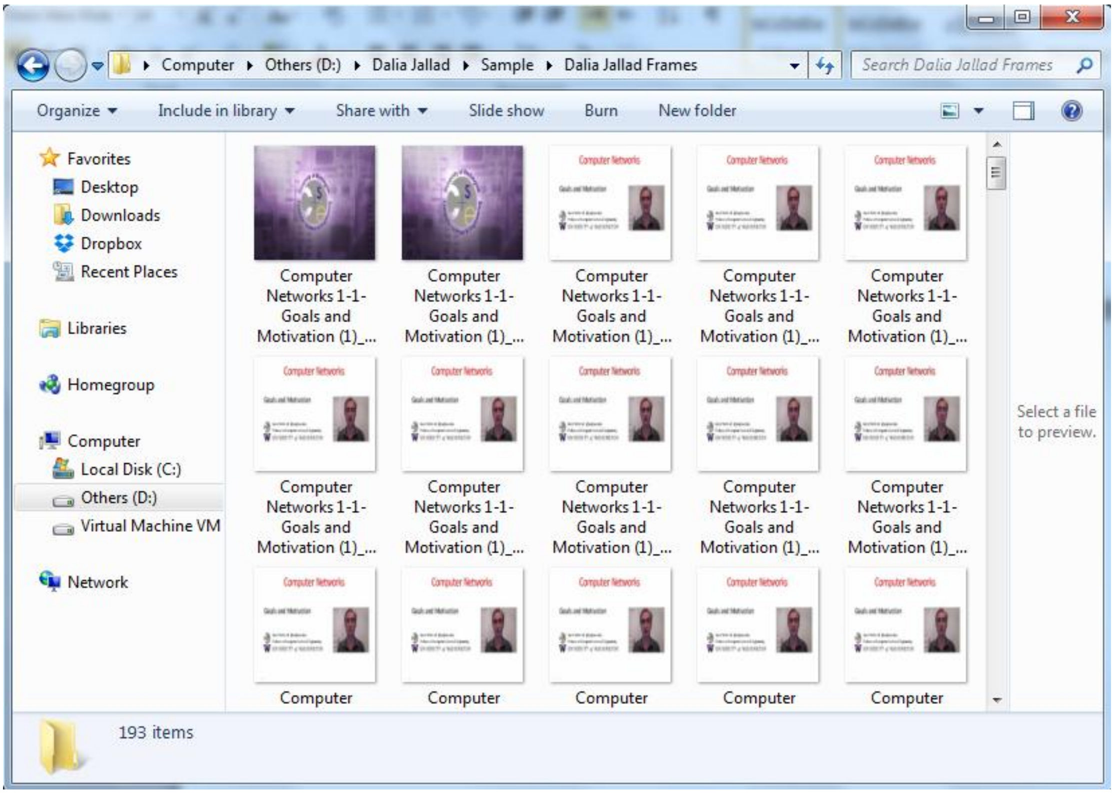


Figure 3.4 The exported frames folder

www.manaraa.com

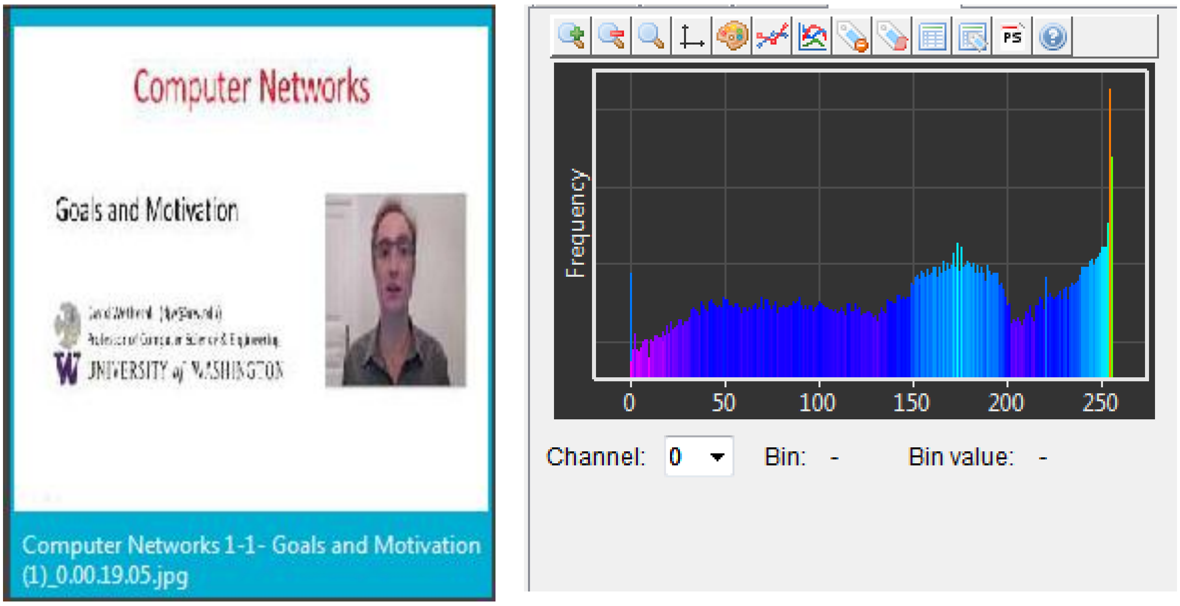
Furthermore, for each frame a color histogram was generated of the type *16 x 8 2D HS* using MATLAB software package in order to evaluate the normalized difference and extract the key frames of the sample lecture video. Figure 3.5 illustrate the preprocessing phase of extracting key frames and evaluating normalized differences. Thus, MATLAB support generating color histogram for HSV image type. Hence, the png image (i.e. frame image extension) was converted to HSV image in order to have three channels (i.e. H channel, S channel, and V channel).

```
% Middle East University - Master Thesis Propsed Program
% STEP 4: Generate  histogram  of H channel
numberOfBins = 16;
[hCounts hValues] = hist(H_Channel(:), numberOfBins);
subplot(3,3,7);
bar(hValues, hCounts);
title('Histogram of H Channel');

% STEP 5: Generate histogram of S channel
[sCounts sValues] = hist(S_Channel(:), numberOfBins);
subplot(3,3,8);
bar(sValues, sCounts);
title('Histogram of S Channel');

% STEP 6: Generate  histogram of V channel
[vCounts vValues] = hist(V_Channel(:), numberOfBins);
subplot(3,3,9);
bar(vValues, vCounts);
title('Histogram of V Channel');

% Construct 2D histogram with H along vertical and S along horizontal.
[rows cols numberOfChannels] = size(hsvImage);
maxH = max(max(H_Channel))
maxS = max(max(H_Channel))
hist2d = zeros(numberOfBins+1, numberOfBins+1);
for col = 1 : cols
for row = 1 : rows
r = int32(H_Channel(row, col) * numberOfBins) + 1;
c = int32(S_Channel(row, col) * numberOfBins) + 1;
hist2d(r, c) = hist2d(r, c) + 1;
end
end
subplot(3,3,3);
imshow(hist2d, []);
```

(a) The MATLAB Code was used to generate color histogram *16 x 8 2D HSV*

www.manaraa.com

The Original Frame Image ***The Generated MATLAB Color Histogram***



(b) The original frame image and its color histogram

Figure 3.5 The preprocessing phase of evaluating the key frames and normalized differences

Consequently, the key frames of sample video were extracted as well as the similarity between frames was evaluated based on normalized differences. Thus, the results in this level showed that presentational lecture videos contain the highest number of key frames compared with handwritten lecture videos. In contrast, the spoken lecture videos achieved the lowest number of key frames which reflects high similarity between video frames. Figure 3.6 shows sequential frames with its statistics and color histograms.

المنارة للاستشارات
www.manaraa.com

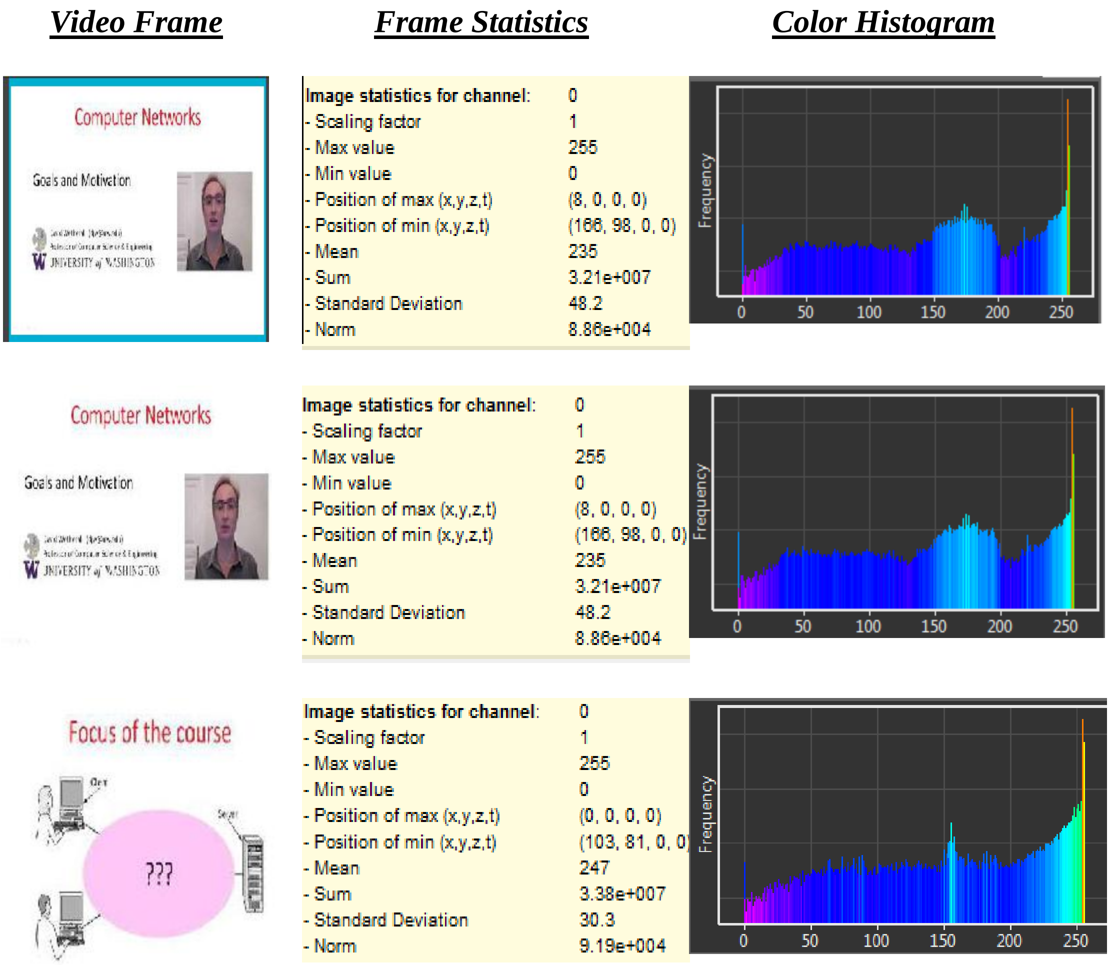


Figure 3.6 Evaluating the similarity between sequential video frames with the results of normalized differences and frame's color histogram

Therefore, the first frame in Figure 3.6 was taken as a key frame, but the second frame showed 99% of similarity which was ignored to be key frame. In contrast, the third frame reflects high dissimilarity based on its histogram (i.e. H_c) compared with the previous key frame histogram (i.e. H_p). Thus, the third frame was taken as a second key frame. In this level, the similarity of video frames have been classified into three categories light similarity, balanced similarity, and hard similarity based on the number of key frames and the normalized differences between color histograms. Figure 3.7 show the similarity classification criteria.

المنارة للاستشارات
www.manaraa.com

$$\text{Similarity}\ (F(n\text{-}1), F(n)) = \begin{cases} \text{Light} & : \quad D_L(F_{n-1}, F_n) < 0.4 \\ \text{Balanced} & : \quad 0.8 > D_L(F_{n-1}, F_n) >= 0.4 \\ \text{Hard} & : \quad 1 => D_L(F_{n-1}, F_n) >= 0.8 \end{cases}$$

Figure 3.7 The proposed similarity classification criteria

The hypothesis in this level showed the classification criteria between lecture videos types. Thus, light similarity criterion was assigned for presentational videos, balanced similarity criterion was assigned for handwritten lecture videos, and hard similarity criterion was assigned for spoken lecture videos. These results were used as input for the next level.

Basically, identifying the video's content type is important to identify the technique used for text recognition. Hence, the text in presentational lecture videos could be extracted using OCR engines, As well as, the text in handwritten lecture videos could be extracted using edge detection engines which is differ than the technique in presentational videos. On the other hand, the text in spoken lecture videos could be extracted using ASR engine. Figure 3.8 shows the flow chart of the operation was conducted in this level.

المنارة للاستشارات
www.manaraa.com

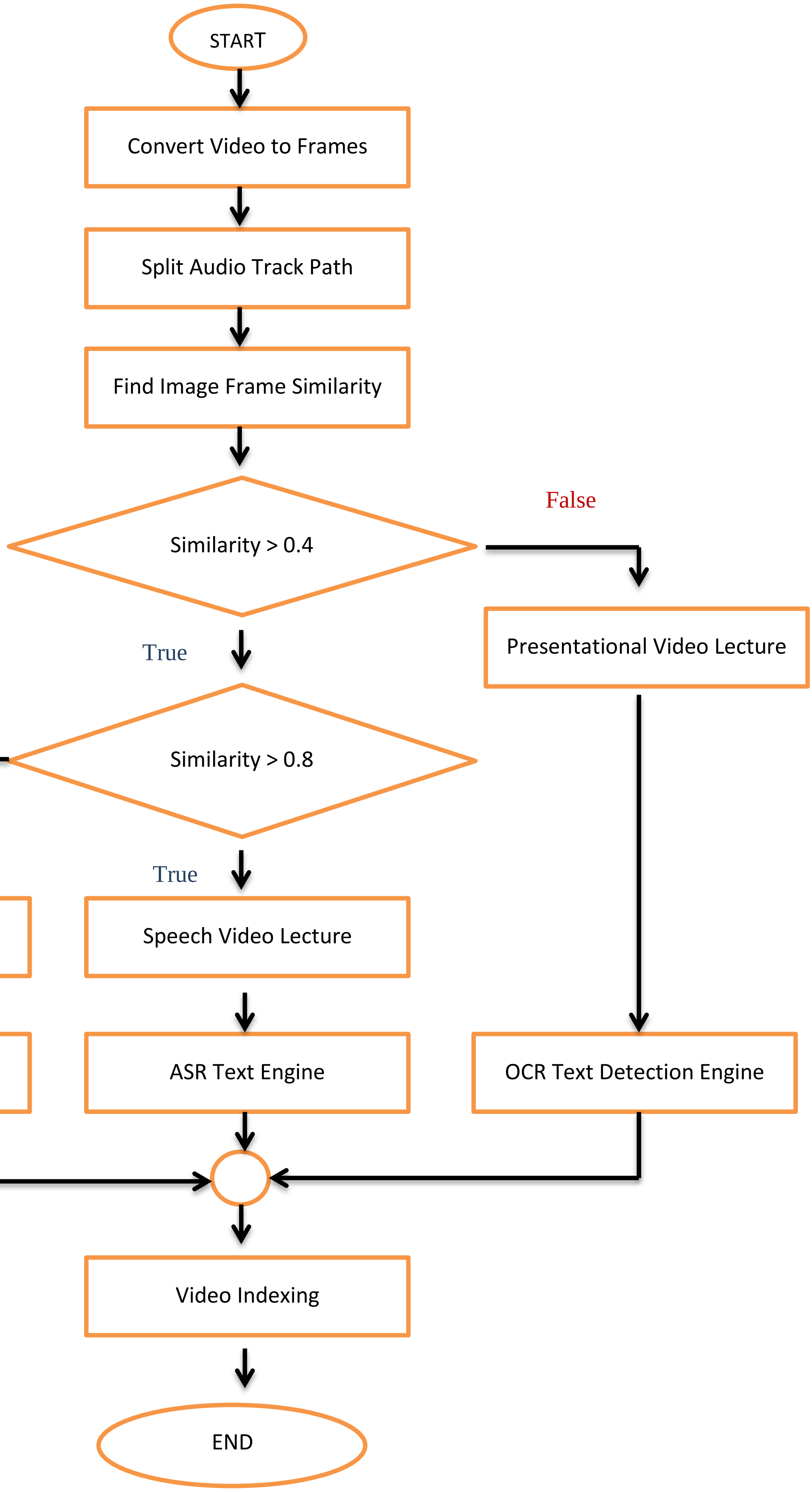


Figure 3.8 Flow chart of the proposed method in the video content analysis level

للاستشارات
www.manaraa.com

3.2.2 Text detection and localization level

This level aims to detect any textual data in lecture video frames as well as specify the regions that contain text in frames. Gallavata, J. et al. found out that the special characteristics of text in video frames are the text has properties of gray level (e.g. high level of contrast for a specific direction), morphological properties (e.g. the shape, the spatial distribution), geometrical properties (e.g. the size, alignment, orientation, stroke density, stroke statistics), and temporal properties (e.g. stability) (Gallavata J., et al., 2007). In our experiment, the text detection process was based on the geometrical properties of the text especially the stroke density and statistics as well as the text alignment.

In this context, LeBourgeios, F. found out that text characters form a regular texture that containing vertical strokes which are aligned horizontally (LeBourgeios, 1997). Therefore, text localization process was divided into two types based on the utilization of content features which were region based type, and texture based type. In region based type the properties of color/gray scale in a text region, or its variations with the frame image background were used to localize text regions. Furthermore, this type contains two approaches which are Connected Components (CC) and edge based approaches. In CC approach the idea is grouping the small parts into larger parts till all regions are identified in the image and guarantees a successful separation of text regions than non-text regions. Hence, the output of both approaches is moved to a Binarization process in order to determine the bounding boxes (i.e. the boxes that show text regions). In our experiment, CC and edge based approaches were employed to find out their effects on text recognition level and performance. Figure 3.9 illustrates the proposed experiments stages in this level.

المنارة للاستشارات
www.manaraa.com

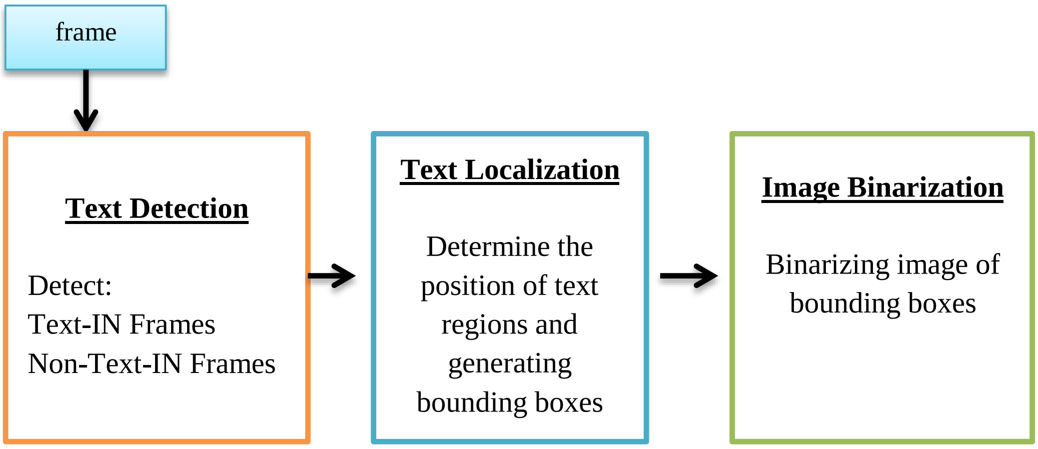


Figure 3.9 The main three stages of text detection and localization level in the proposed method

Basically, detecting text in frame images was implemented in this experiment by testing each pixel in the target image and compares it with a threshold. Thus, several alternatives found in the literature to generate and estimate threshold value. In this study, we had chosen the iterative and OTSU threshold algorithms. For the easy of computation, constancy, and effectiveness. Furthermore, the response time that is consumed using both algorithms is considerably less than other algorithms (Makkar, H. and Pundir, A., 2014). In iterative threshold technique, initial threshold is generated by calculating the mean of intensity regions ($T_{initial}$) in image. Consequently, the mean of pixels under and above the initial threshold are calculated (μ_{below}, μ_{above}) respectively which represents the new threshold (T_{new}), the process is repeated till the difference between any two thresholds does not exist which guarantee text region partition. Equation (2) hows the formulas used to generate threshold using iterative algorithm.

المنارة للاستشارات

www.manaraa.com

$$T_{new} = \frac{1}{2}(\mu_{below} + \mu_{above}) \quad \text{..................... (2)}$$

Where:

- μ_{below} : is the average of pixels below threshold.
- μ_{above} : is the average of pixels above threshold.

In OTSU threshold technique an optimal threshold can be calculated using a gray value histogram through using two Gaussian distribution in image. Therefore, class 0 represents non text pixels, and class 1 represents text pixels. Hence, the maximum inter-class variance is employed to find the threshold. Equation(3) Shows the formulas used to generate threshold using OTSU algorithm (Trier, O., and Jain, A., 1995).

$$\omega_{First} = \sum_{i=1}^{t} \frac{H_i}{N}$$

$$\omega_{Second} = \sum_{i=t+1}^{L} \frac{H_i}{N}$$

Where:

ω_{First} : First Gaussian distribution Class.

ω_{Second} : Second Gaussian distribution Class.

H: Gray Histogram
N: Number of pixels
L: Number of bin in histogram

$$t = AVGMAX\ (\omega_{First}\,\omega_{Second}\,(\mu_0 - \mu_1)^2) \quad \text{……………...….. (3)}$$

Consequently, the proposed enhancement concentrated on creating third class for the adjacent pixels. By calculated the ratio between the pixel's mean of the first Gaussian class with the overall mean of the first Gaussian class, in order to find the ratio of intensity of each pixel in the first class. Thus, the standard deviation of the generated class was calculated as well as the mean of its elements. Therefore, the normal distribution have been generated (i.e. Gaussian) for the values in the third class

المنارة للاستشارات
www.manaraa.com

in order to calculate the Probability Density Function (PDF). Figure 3.10 illustrates the binarization differences between the enhanced OTSU and traditional OTSU algorithm.

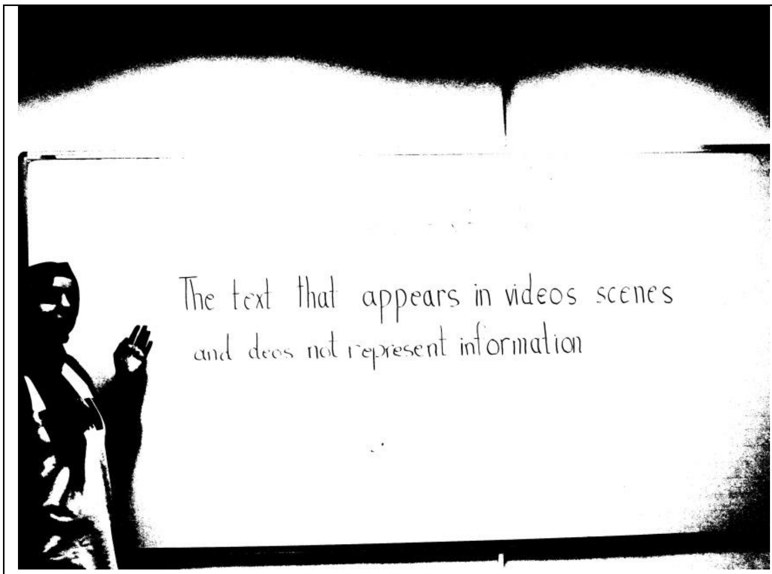
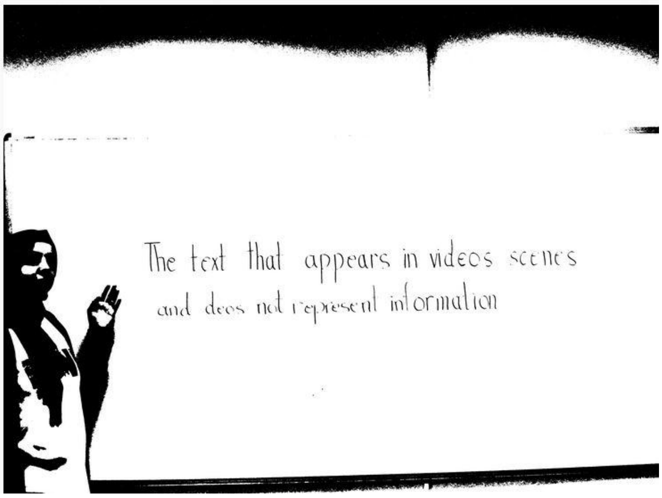
<table>
<tr>
<td>
</td>
<td>(A)
Sample Frame Image Binarization using the enhanced OTSU threshold technique. (e.g. Handwritten lecture video)</td>
</tr>
<tr>
<td>
</td>
<td>(B)
The Same Picture Using Traditional OTSU threshold technique.</td>
</tr>
</table>

Figure 3.10 The output binarization images in the proposed enhanced OTSU and traditional OTSU threshold techniques

Text localization process in this level is employed based on CC which was consisted into analyzing the geometrical arrangement of edges or the homogenous in color and gray scale which belongs to character. Hence, text localization in our

المنارة للاستشارات
www.manaraa.com

experiment focused on character features such as the density of character edge, the strength of character edge, and the horizontal distribution of characters. Thus, based on the estimated threshold using the proposed technique we slightly employed Gallavata, J. et. al algorithm to define the edge image of text region as well as determine the text region rectangle. Figure 3.11 illustrates a pixel with (x, y) coordinates and its edges coordinates from left and right sides (Gallavata, J., 2007).

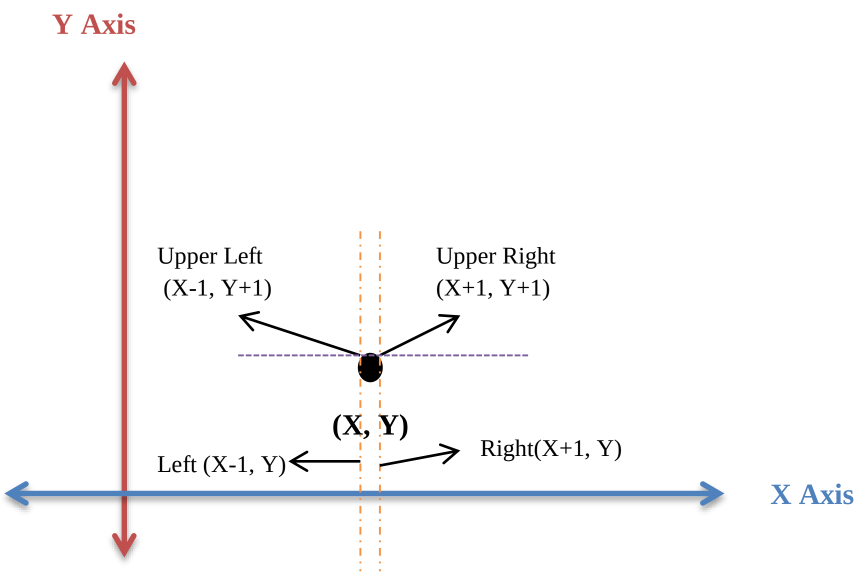


Figure 3.11 Textual pixel coordinates after binarization process and its edges coordinates from different sides

Gallavata, J., et al proposed a text localization algorithm that was responsible to compare image pixels with an estimated threshold in order to find the edges of text characters. Their algorithm focused on comparing each pixel with adjacent pixel from left (i.e. *(x-1, y)* coordinates), as well as, the adjacent pixel from the right (i.e. *(x+1 ,y)* coordinates). Thus, they found the normalized difference between the adjacent pixels and the target pixel. Their algorithm was working in a bottom up fashion since they

المنارة للاستشارات
www.manaraa.com

tried to find the upper left of the adjacent pixel by identifying *(x-1, y-1)* coordinates for upper left as well as *(x+1, y-1)* coordinates for the upper right (Gallavata, J., et. al., 2007). In this experiment, their algorithm was employed in top down fashion in order to check the top image pixel because the target video type in this research for lecture videos. Hence, the top of each video frame represents the title of each frame image which represents a good metadata for our hypothesis. Figure 3.12 illustrates the flow chart of the text localization algorithm in digital images.

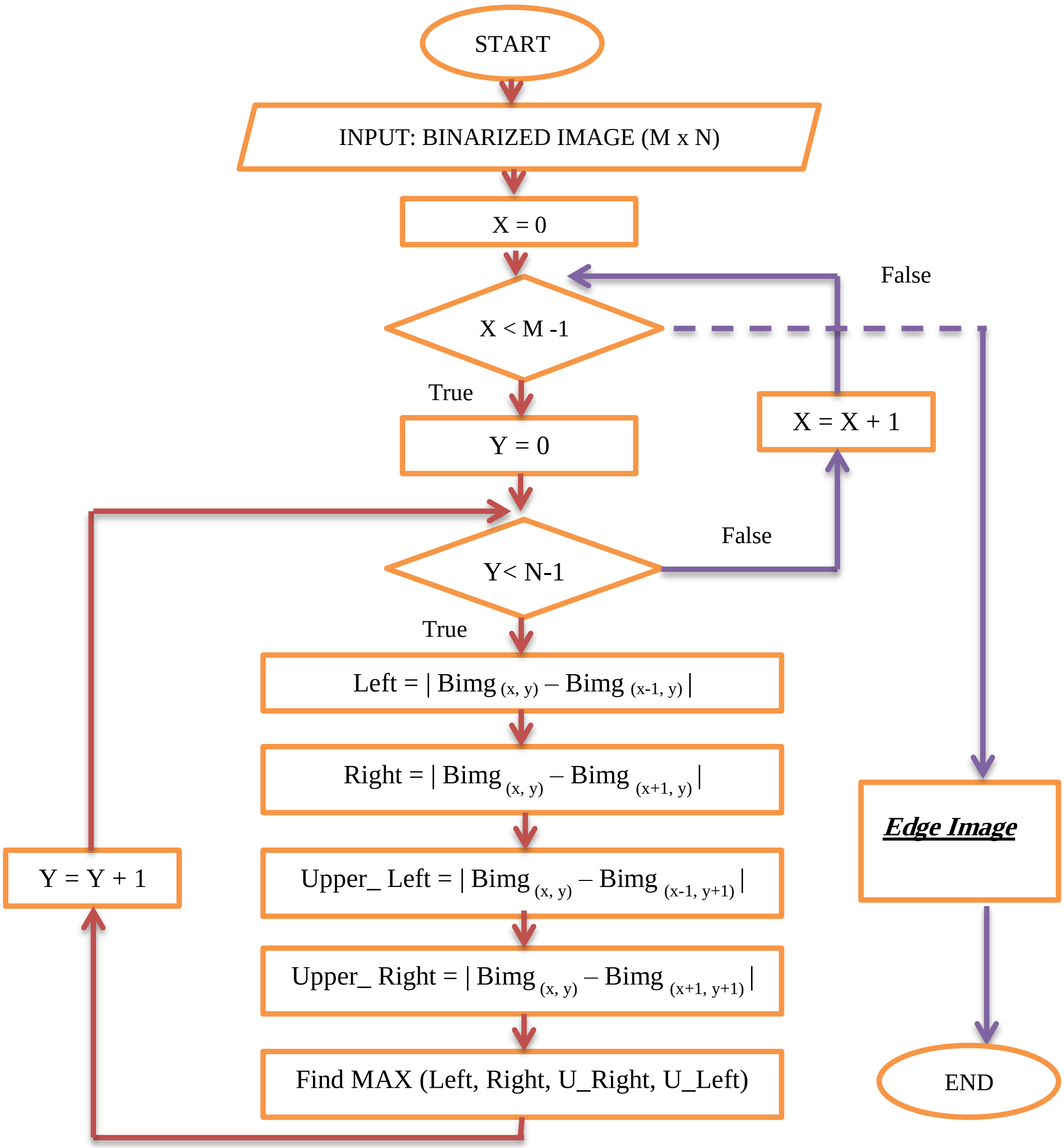
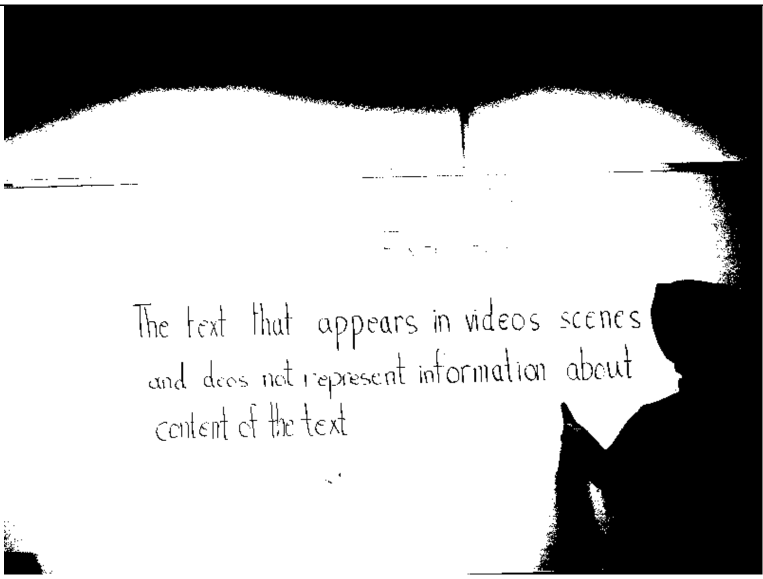


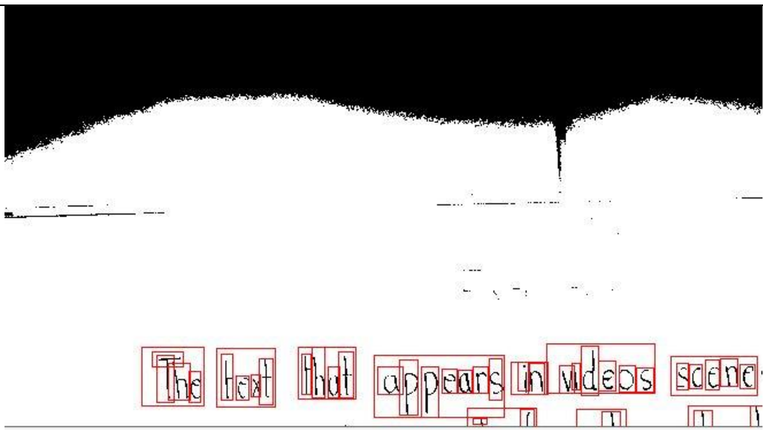
Figure 3.12 Flow Chart of Edge Detection of Character Region

www.manaraa.com

Figure 3.13 shows the results of edge detection of each character and text regions in our experiment.



(a) The image after binarization using our enhanced OTSU threshold technique



(b) The processed image after localizing text regions in image using bounding boxes

Figure 3.13 Text localization in the proposed experiment (a) The image after binarization (b) Text regions identification

www.manaraa.com

3.2.3 Video text recognition level

Basically, the main aim of this level is to discuss the used text recognition techniques in the experiment to implement the proposed method. Hence, different types of commercial open source OCR engines (e.g. Simple OCR, Tesseract OCR, and OCR-AD) were employed to extract text from sample presentational lecture videos and sample handwritten lecture videos. In this level, the input is the result (i.e. binarization image with bounding boxes for text regions) from the previous proposed levels. Hence, the comparison was assigned between these OCR engines based on the performance evaluation using precision and recall measures. Figure 3.14 illustrates the main processes of text recognition level.

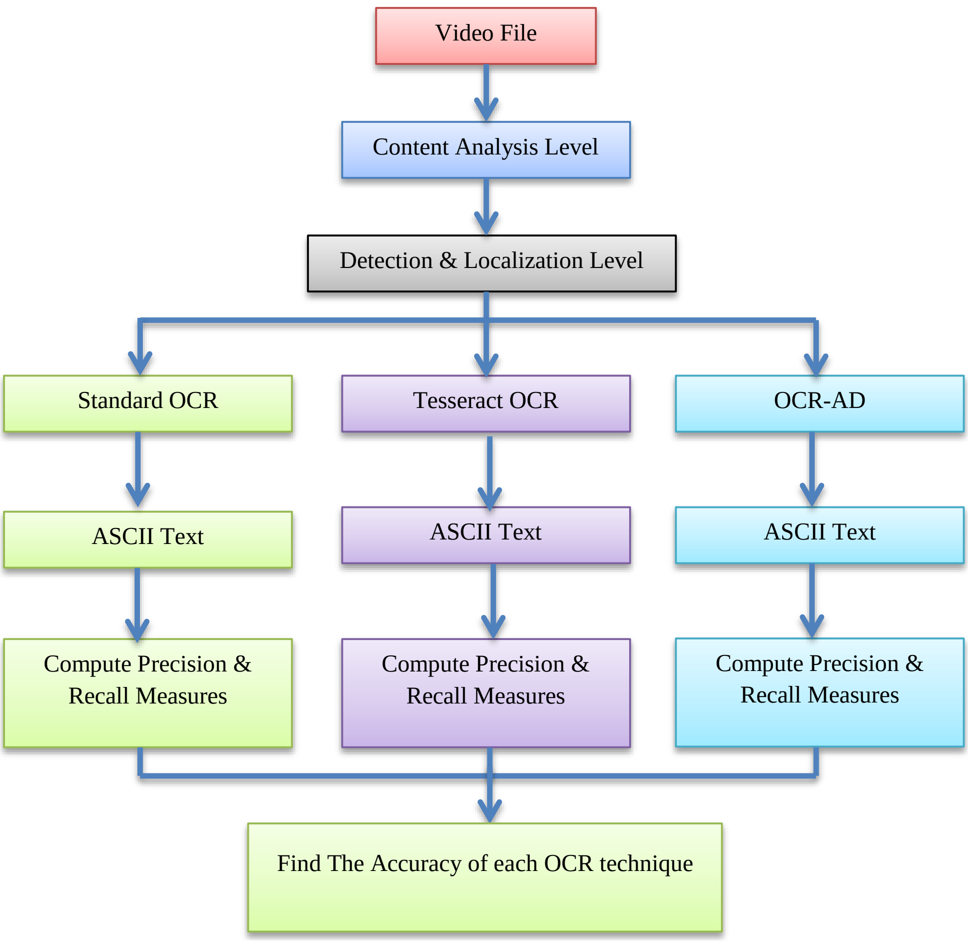


Figure 3.14 Text recognition level primary processes over several OCR techniques

المنارة للاستشارات
www.manaraa.com

Actually, performance evaluation step in this level concentrated on calculating the precision and recall of text results for each OCR technique was used. Thus, can use contingency table to make text classification of each OCR technique result. Table 3.1 shows the two by two contingency table of text classification in OCR techniques.

	Correct Text	**Not Correct Text**
Selected Text	True Positive (TP)	False Positive (FP)
Not Selected Text	False Negative (FN)	True Negative (TN)

Table 3.1 The 2 x 2 contingency table of the extracted text using OCR techniques

For instance, if the extracted text was not correct but in same time was from the selected text (i.e. from the localized text) the result will be FP. Hence, contingency table will be used to identify TP, FN, FP, and TN from the extracted text results. In this context, the recall measure represents the percent of the correct extracted text that was selected. In contrast, precision measure represents the percent of the selected text (i.e. localized text) that was correct. (4) and (5) shows the formulas that were used to compute recall and precision in this research.

$$\text{Recall} = \frac{TP}{TP + FN} \qquad \ldots\ldots\ldots\ldots\ldots\ldots (4)$$

Where:

- TP : is True Positive
- FN : is False Negative

$$\text{Precision} = \frac{TP}{TP + FP} \qquad \ldots\ldots\ldots\ldots\ldots\ldots (5)$$

Where:

- TP : is True Positive
- FP : is False Positive

المنارة للاستشارات
www.manaraa.com

In this research, the ASR algorithm - that is usually used to recognize textual data from spoken lecture videos – is used to handle the textual data extracted using OCR techniques. Furthermore, our experiment took into consideration that aggregating results of OCR and ASR techniques leads for better results. Hence, the audio track which was splied in the content analysis level in order to extract text using a well-known ASR technique. Therefore, in this experiment, an open source HTK toolkit that was built in C++ programming language was used, as well as, it is available for academic and research purposes (1). Figure 3.15 shows the spoken lecture video main text extraction processes.

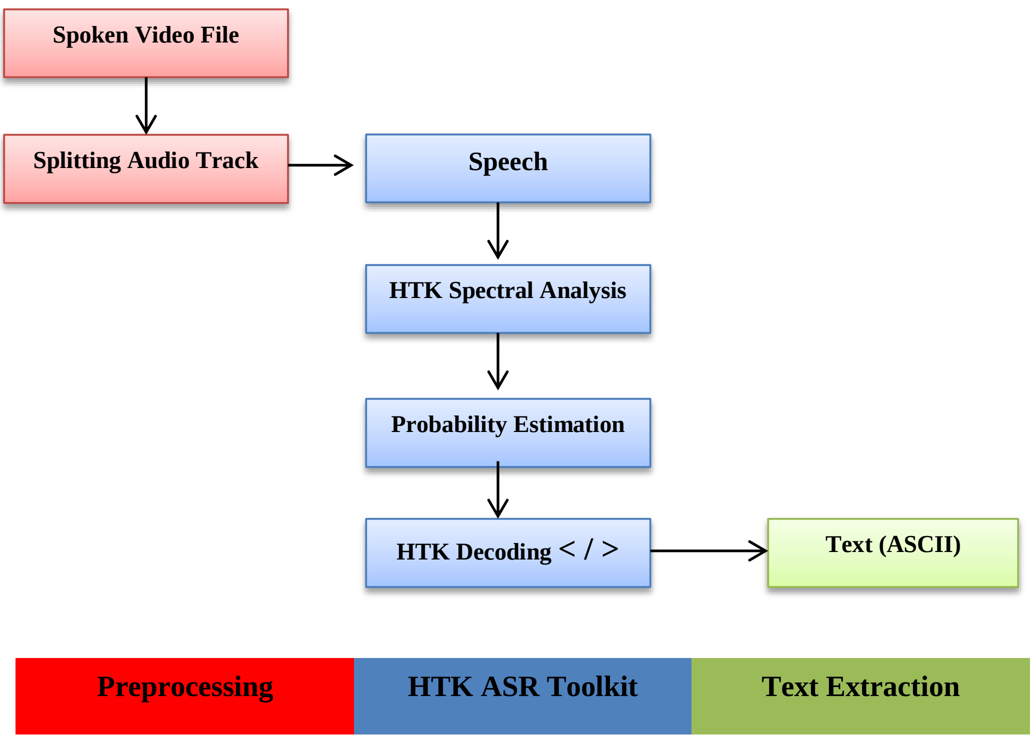


Figure 3.15 The general framework of ASR technology (Lawate, A, and Wankhade, M., 2015)

---

(1) Retrieved from http://htk.eng.cam.ac.uk/ ,"*Hidden Markov Model Toolkit*", Date Accessed: 2nd December 2015.

المنارة للاستشارات
www.manaraa.com

3.2.4 Video archiving and indexing level

The main goal of this level is to calculate the accuracy of retrieving video from sample lecture video database based on the extracted text which represents the meta-data of the target lecture video, as well as, used as a description for the target video. Hence, the balanced F1 measure was employed to evaluate the accuracy of retrieving results in the proposed system. Thus, F1 measure were represented as the ratio between recall and precision metrics Figure 3.16 illustrates the general framework of the proposed system in this level.

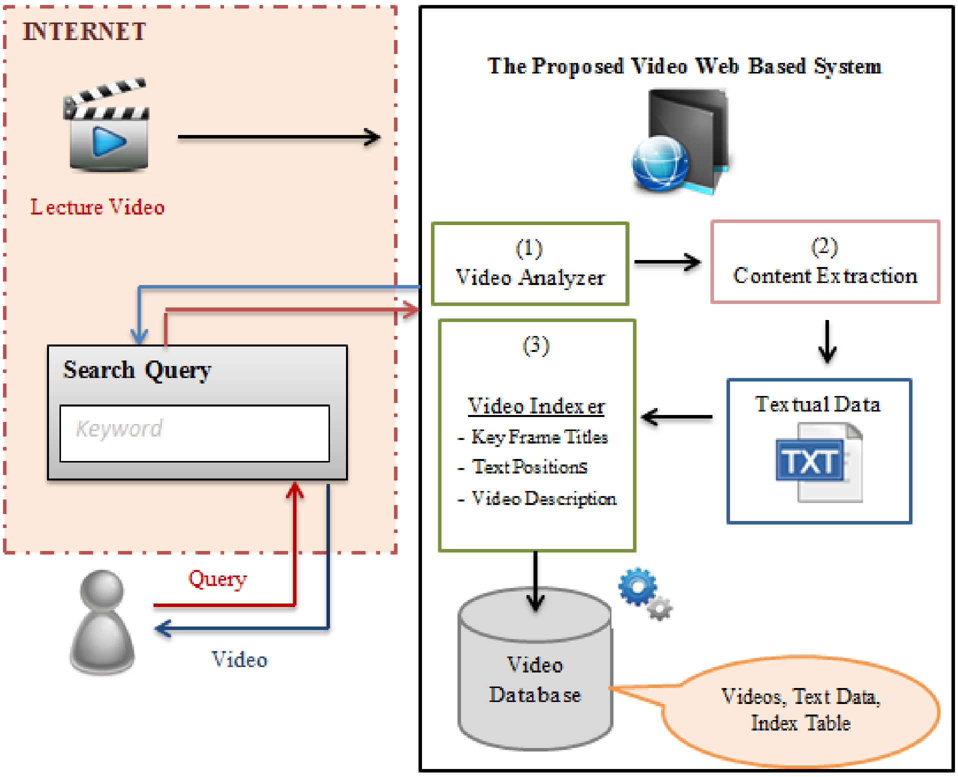


Figure 3.16 The general framework of the proposed archiving and indexing system

The extracted textual data are stored in a database for each video. Thus, each video text scene is connected with its contents textual data in order to ease extracting the video with the position of the text that match the search keywords.

المنارة للاستشارات
www.manaraa.com

CHAPTER FOUR

The Experimental Results

4.1. Overview

This chapter discussed the results of the proposed experiments which were discussed in chapter three. In this chapter, the performance of lecture video text recognition was discussed, as well as the accuracy of retrieving lecture videos from online video repositories was discussed too. Furthermore, the results discussion in this chapter covered the proposed levels in this study. Therefore, the flow of results in this chapter was discussed as the following:

- The experimental results of the content analysis level were discussed in term of normalized difference, video's key frames generation, and calculating the frames similarities.
- The experimental results of text detection and localization in term of generating and estimating the target threshold based on using the iterative threshold, OTSU algorithm, and our enhanced algorithm.
- The experimental results of text recognition level from performance perspective based on calculating recall, precision, and calculating the accuracy.

4.2. The experimental results of content analysis level

This section discussed the results of the proposed experiment which was responsible to identify the contents of lecture videos. Therefore, this proposed experiment took into consideration the three types of lecture videos which are the presentational lecture videos, the handwritten lecture videos, and the spoken lecture videos. Hence, the proposed experiment in this level was conducted on a sample lecture videos from different content types. Thus, the number of videos in this sample was 56

www.manaraa.com

lecture videos - of the type MP4 - were downloaded from Coursera, MIT CourseWare, and YouTube official websites. Furthermore, the length of the sample videos was varied between (4- 25) minutes with different storage sizes.

Basically, the results in this level was computed over three stages which were generating the color histogram for each video frame as well as finding the values of color for each pixel in the histogram, calculating the normalized differences between frame histograms using formula (1) in order to find the key frames, and calculating the similarity between the frames histograms in order to classify the lecture video contents based on the similarity classification criteria in chapter three.

4.2.1 Presentational lecture video analysis

This section discussed the results of analyzing presentational lecture video type. Furthermore, we conducted the proposed experiment on a sample video that was splitted into 193 frames. Then generated a color histogram for each video frame. And extracted each pixel's color value, as well as, applied the proposed computations on these histogram data (See Appendix A for a table showing all color values in the generated color histogram for selected three sequential frames in a sample video).

Consequently, calculated the normalized differences between the color values of each pixel in two sequential frames as well as calculated the summation for color values of each frame image and the summation of the normalized differences values (See Appendix B for a table showing the normalized differences and the corresponding calculations). Table 4.1 shows the calculation of the main variable in this level.

المنارة للاستشارات
www.manaraa.com

Table 4.1 The results were used to classify the lecture video type based on the mean of the key frame's color histogram's similarities for a presentational sample lecture video

Frame # / Variables	Summation	Difference	Norm. Diff.	Similarity	Key Frame
Frame 1	136797	0	0	1	
Frame 2	136800	37571	0.274641813	0.725358	FALSE
Frame 3	134260	111997	0.834179949	0.16582	TRUE
Frame 4	136749	38259	0.279775355	0.720225	FALSE
Frame 5	273370	158474	0.579705162	0.420295	TRUE
Frame 6	136741	33254	0.24318968	0.75681	FALSE
Frame 7	273323	32715	0.119693549	0.880306	FALSE
Frame 8	268520	33587	0.125081931	0.874918	FALSE
Frame 9	136727	30794	0.225222524	0.774777	FALSE
Frame 10	150683	92193	0.611834115	0.388166	TRUE
AVERAGE	178397	56884.4	0.329332408	0.670668	FALSE
Mean Key Frames	0.276992968	***Lecture Video Type***	Presentational		

Basically, the similarity in this study was computed based on two main factors which were the summation of color pixels in the colored histogram, and the absolute difference between the current histogram and the previous one. Furthermore, the normalized difference have been calculated by finding the ratio of the difference to the summation of color values. Hence, the complement of the normalized difference was assigned as the similarity. Figure 4.1 shows a graph of selected ten sequential frames in order to represent the similarities between video's frames.

المنارة للاستشارات
www.manaraa.com

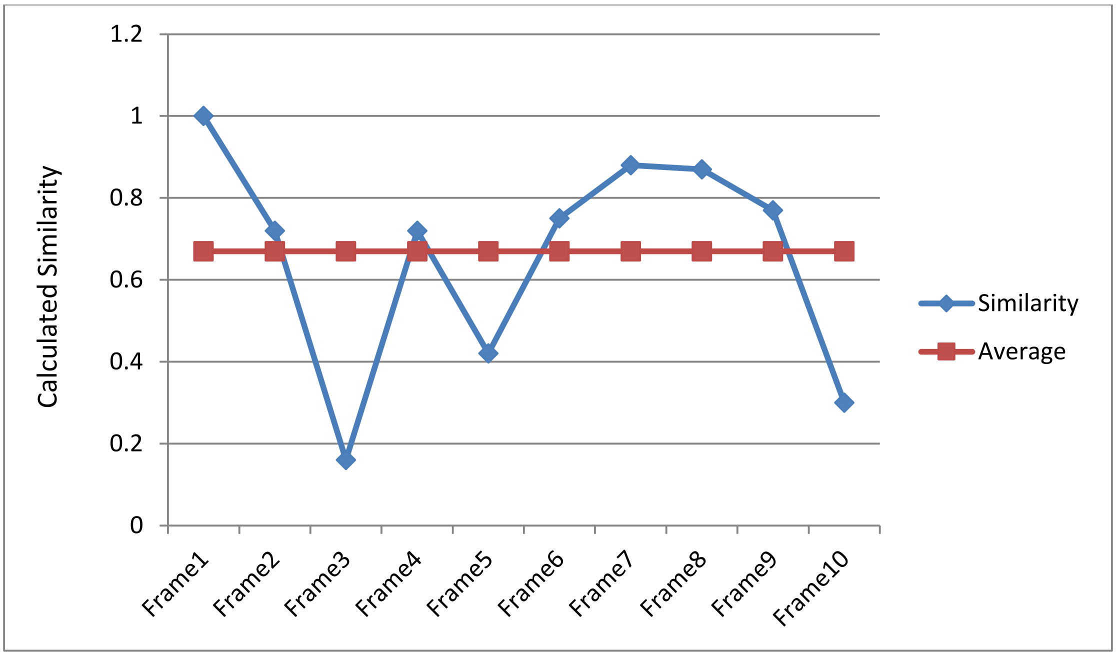


Figure 4.1 The similarity results of ten selected sequential video frames in term of color histogram pixel's values

Consequently, the results showed that the proposed experiment successfully detected three key frames out of ten sequential frames which were frame 3, frame 5 and frame 10 respectively. Hence, the extraction process was done based on the similarity results compared with the mean of the frame's similarities. Thus, based on the study's hypothesis the key frames was chose from the frames that were below the average. For instance, in the selection of frame 3 as a key frame, the similarity between frame 2 and the current (i.e. frame number 3) was 0.16 due to the large normalized difference between them which was 0.8 (See Appendix C for table that shows the original images of the frames and it's color histograms).

In this context, the results showed that the summation of pixel's color value for each video frame's histogram does not have a direct effect on identifying the key frames. On the other hand, the absolute difference had a direct effect on frame's similarity. Thus,

المنارة للاستشارات
www.manaraa.com

the results found out that the similarities between the extracted key frames had a direct effect on identifying the type of lecture video. Therefore, in this experiment the mean of the similarities of the key frames was calculated in the sample video in order to assign the similarity of the sample video as light, balanced, or hard similarities. Hence, the presentational lecture video samples achieved light similarities and the results was varied between (0.27 – 0.33) due to the differences between the text and objects embedded to the slide in lectures. Table 4.2 shows the similarities results of the extracted key frames from 82 sequential frames out of 193 frames in the sample lecture video.

Table 4.2 The similarities results of the extracted key frames from 82 frames in the sample video

Key Frame	**The computed similarities**
Frame 3	0.16
Frame 5	0.4
Frame 10	0.38
Frame 22	0.19
Frame 25	0.11
Frame 34	0.5
Frame 50	0.32
Frame 69	0.36
Average	***0.3025***

4.2.2 Handwritten lecture video analysis

This section discussed the results of conducting the proposed experiment on handwritten lecture videos. The sample videos in this level were created by the author in MP4 format, the time of these lecture videos was varied between (3 – 7) minutes. Furthermore, we applied the same experiment processes alike presentational video analysis stage. Figure 4.2 shows two snapshots from two different sample videos of the author.

المنارة للاستشارات

www.manaraa.com

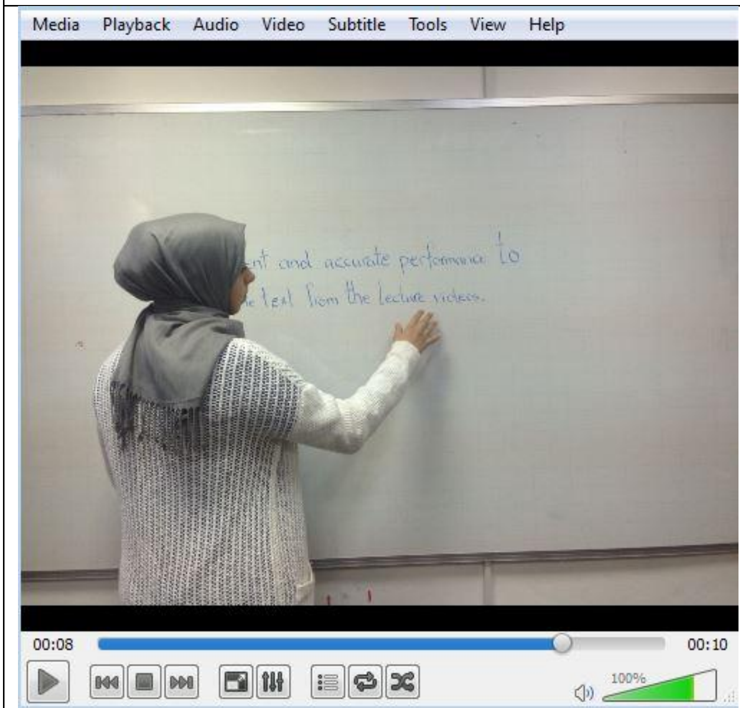
<table>
<tr>
<td>
</td>
<td>(A)
Text Information
Extraction Lecture Video
By: Dalia Jallad</td>
</tr>
<tr>
<td>
</td>
<td>(B)
The type of text in lecture
videos
By: Dalia Jallad</td>
</tr>
</table>

Figure 4.2 Snapshot from two handwritten lecture videos in image processing and text recognition techniques

Basically, the results of this case have been compared with results of presentational lecture video analysis, which lead to found out that the effects of the

www.manaraa.com

main variable in this level in order to differentiate the handwritten videos type than the presentational ones. Furthermore we conducted the experiment on several handwritten lecture videos to make sure that the computed results semantically the same. Table 4.3 shows the results of the handwritten lecture videos in term of the primary variables in this study.

Table 4.3 The similarity results of ten selected sequential video frames in term of color histogram pixel's values for a handwritten sample lecture video

Frame # / Variables	Summation	Difference	Norm. Diff.	Similarity	Key Frame
Frame 1	12763683	0	0	1	
Frame 2	12023560	6990655	0.581413076	0.4185869	TRUE
Frame 3	11985591	4239274	0.353697536	0.6463025	FALSE
Frame 4	12162516	3315971	0.272638572	0.7273614	FALSE
Frame 5	12203279	3518194	0.288299071	0.7117009	FALSE
Frame 6	11324561	3987653	0.35212429	0.6478757	FALSE
Frame 7	12750201	4425657	0.347104881	0.6528951	FALSE
Frame 8	12365241	7172330	0.580039645	0.4199604	TRUE
Frame 9	11875632	4686329	0.39461723	0.6053828	TRUE
Frame 10	12649877	2936447	0.232132455	0.7678675	FALSE
AVERAGE	**12210414.1**		**0.340206676**	**0.621993**	FALSE
Mean Key Frames	**0.41927364**	***Lecture Video Type***	Handwritten		

The result showed that the similarity - in case of handwritten lecture videos - between each two frames higher than the similarities in the presentational video frames. Thus, a small number of key frames can be detected as well as the similarities between the key frames were higher in handwritten video's frames. Therefore, we assigned the handwritten frame's similarity as a balanced similarity. Hence, the similarity that was calculated over the test lecture video samples was varied between (0.4-0.8). Figure 4.3 shows a graph of the calculated similarities between frames in a handwritten video sample.

المنارة للاستشارات
www.manaraa.com

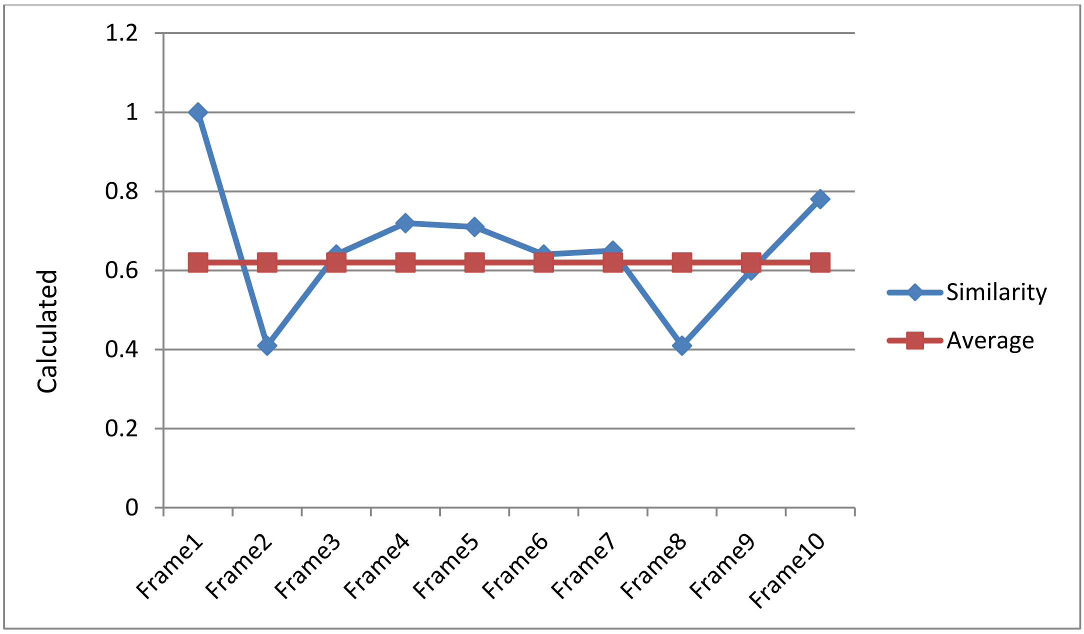


Figure 4.3 The similarity results of ten selected sequential video frames in term of color histogram pixel's values in case of handwritten lecture videos

For more explanation, figure 4.4 shows that the results of similarities were chosen as a key frames were frame 2 and frame 8. Hence, their similarities were below the average of frame's similarities. However, in this case result found that the similarities gaps for above average similarities were intangible compared with the case of presentational lecture videos. Furthermore, the gaps between the extracted key frames were also intangible in term of the proposed similarity measurement. The average of number of key frames was extracted in 7 minutes from handwritten sample videos was varied between three to five key frames which supports that the similarity between frames in this case was high for many reasons such as:

- The similarity in the colors values of frame images.
- The similarity of closed object in the frame images (e.g. the instructor's shape, the board, the lecture room properties… etc.).
- The static position of capturing camera used in recording process.

المنارة للاستشارات
www.manaraa.com

4.2.3 Spoken lecture videos analysis

This section discussed the analysis results that were extracted after studying the spoken lecture videos in the proposed experiment. Therefore, the proposed experiment in this case was conducted on several spoken lecture videos were downloaded from several educational websites. Thus, the sample spoken video files were MP4 format and its time duration was varied between 4 to 10 minutes. This section embedded the results of a sample spoken video downloaded from LetsTalk official website which was 8 minutes duration. Table 4.4 shows the results that were extracted with the main calculations to evaluate the main variables in this study.

Table 4.4 The similarity results of twenty of sequential frames that cover two minutes from the sample spoken video

Frame # / Variables	Summation	Difference	Norm. Diff.	Similarity	Key Frame
Frame 1	305630	0	0	1	
Frame 2	305498	15319	0.050144354	0.949855646	False
Frame 3	305562	15253	0.049917856	0.950082144	False
Frame 4	304988	15125	0.049592115	0.950407885	False
Frame 5	404432	40950	0.101253115	0.898746885	True
Frame 6	402561	14267	0.035440592	0.964559408	False
Frame 7	402495	15253	0.037896123	0.962103877	False
Frame 8	403160	17485	0.043369878	0.956630122	False
Frame 9	255356	21742	0.085143878	0.914856122	False
Frame 10	245308	25834	0.105312505	0.894687495	False
Frame 11	246287	25100	0.101913621	0.898086379	False
Frame 12	249147	28879	0.11591149	0.88408851	False
Frame 13	246101	25425	0.103311242	0.896688758	False
Frame 14	348070	98215	0.282170253	0.717829747	True
Frame 15	348112	13252	0.038068208	0.961931792	False
Frame 16	342974	12459	0.036326369	0.963673631	False
Frame 17	341093	14156	0.041501878	0.958498122	False
Frame 18	338742	12654	0.037355864	0.962644136	False
Frame 19	336794	13569	0.040288722	0.959711278	False
Frame 20	341002	10125	0.029691908	0.970308092	False
Average	**323665.6**	**21753.1**	**0.069230499**	**0.800769501**	
Mean KFrames	**0.808288316**	***Lecture video Type***		***Spoken Text***	

www.manaraa.com

Consequently, the similarity results showed that the frames of spoken lecture video contain a high similarity compared with the presentational and handwritten lecture videos. Thus, the results showed that two key frames were extracted frame 5 and frame 14 respectively. Furthermore, we assigned the similarity between the key frames as a hard similarity because of the low normalized difference between them. Figure 4.4 illustrates the graph of spoken lecture video sample frames similarities in term of color histogram.

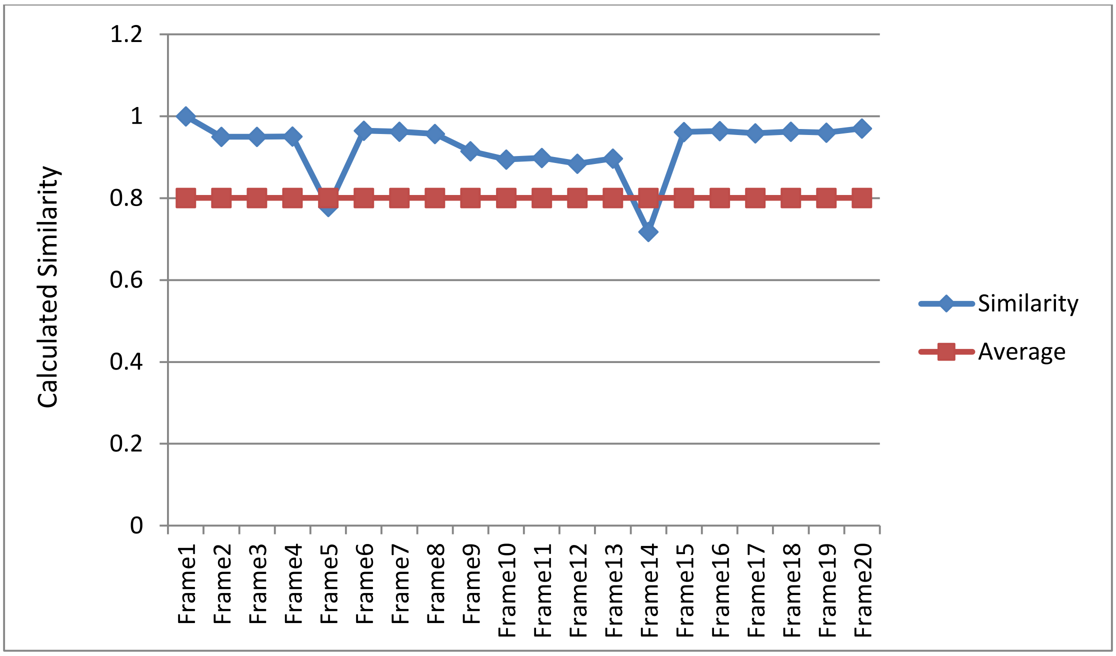


Figure 4.4 The graph of spoken lecture video sample frames similarities in term of color histogram

The results in this case showed that an absence of any drastic gaps between frame's similarities. Hence, the two frames existed below the average line and were assigned as key frames. Thus, the difference between frame 5 and the other frames was intangible. However, in frame 14 result found out a tangible decreasing in frame's similarity.

www.manaraa.com

Discussion

This section discussed the results from video content analysis level of presentational videos case, handwritten videos case, and spoken text videos case. Therefore, the results showed that the frame's similarity was light in case of presentational lecture videos because of the diversity in the objects embedded in the presentation slides which lead for different color and textures. In the handwritten lecture video case, the results showed a balanced similarity between its frames due to the identical environment and objects that is repeated at each frame image. In the spoken text lecture videos case, the hard similarity was existed and identified due to the same background effects since the concentration of the spoken text of the presenter speech.

The proposed classification criteria successfully identified the content type of each lecture video in the sample set of videos using the proposed experiments. Therefore, the results showed that the key frames that were extracted in each case varied based on the type of the lecture video. Hence, result found out that the presentational lecture videos contained the highest number of key frame. In contrast, the spoken text lecture videos achieved the lowest number of key frames.

In this context, the results reflected a direct effect of the video content type on the similarity between video frames. Thus, the impact of presentational lecture video contents was translated in two points of view. The first is the total number of key frames, and the second is the drastic gaps between video's frames in term of similarity.

4.3. The experimental results of text detection and localization level

In this level, the discussion of the results was in term of finding the number of words and the number of character which were detected in image frames. Hence, the

المنارة للاستشارات
www.manaraa.com

comparison is between the iterative threshold algorithm, OTSU threshold algorithm, and enhanced algorithm on the standard OTSU.

4.3.1 Iterative threshold algorithm analysis

In this section, the iterative threshold algorithm was implemented using MATLAB code in order to evaluate the image threshold. Hence, the mean of the intensity colored pixels taken as the initial threshold. As well as, the estimation of the new threshold was done by comparing the pixels values with the initial threshold in two cases the above and below cases. Thus, the summation of the above mean was computed also the summation for the below mean was computed. Therefore, the new threshold was estimated and this process should be repeated to eliminate the difference between the generated thresholds which guarantees text detection. Figure 4.5 shows the used MATLAB code.

المنارة للاستشارات
www.manaraa.com

```
%-- 18/12/2015 09:31 م --%
% Dalia Jallad "Iterative Thresholding Technique"
%Middle East University Amman - Jordan

clear all;
close all;
clc
dalia = imread ('bmp dalia.bmp');
level = graythresh (dalia);
BW = im2bw (dalia, level);
imshow (BW);
dalia = im2uint8 (dalia(:) );
T=mean( dalia );
% STEP 1: Compute mean intensity of image from histogram
[counts,N] = imhist ( dalia );
i=1;
pixel_mean = cumsum( counts );
T( i ) = ( sum(N.*counts) ) / pixel_mean( end );
T( i )= round( T( i ) );
% STEP 2: compute Mean above Threshold and
% Mean below Threshold using T from step 1
pixel_mean_below = cumsum( counts( 1:T( i ) ) );
Mean_Below_Thresh = sum(N( 1:T(i) ).*counts( 1:T(i) ) )/pixel_mean_below( end );
pixel_mean_above = cumsum( counts(T( i ):end ) );
Mean_Above_Thresh = sum( N(T(i):end ).*counts( T(i):end ) )/ pixel_mean_above( end );
i=i+1;
T(i) = round( ( Mean_Above_Thresh + Mean_Below_Thresh ) / 2 );
% STEP 3 to n: Repeat step 2 till eliminating the difference
While abs ( T(i) - T(i-1) )>=1
pixel_mean_below = cumsum( counts( 1:T(i) ) );
Mean_Below_Thresh = sum( N( 1:T(i) ).*counts( 1:T(i) ) )/pixel_mean_below( end );
pixel_mean_above = cumsum ( counts( T(i):end ) );
Mean_Above_Thresh = sum (N(T(i):end).*counts(T(i):end))/pixel_mean_above(end);
i=i+1;
T(i)=round((Mean_Above_Thresh+Mean_Below_Thresh)/2);
Threshold=T ( i );
End
level = (Threshold - 1) / (N(end) - 1);
```

Figure 4.5 Threshold estimation code in the iterative threshold algorithm using MATLAB

Threshold estimation is the most important step in the iterative threshold algorithm. Hence, we assigned the array of the target image *bmp dalia.bmp* to the variable *dalia* in order to assign an array with the size 650 x 480 x 3. Thus, we computed the mean of the target image *T* as well as we assigned the image histogram into an array

المنارة للاستشارات
www.manaraa.com

count x n where *n* is the color contrast in histogram (0 - 255) and *count* is the pixels that contains *n* color's contrast. Using the cumulative summation for each array *row (n)*, the pixel mean was calculated. Therefore, the second threshold was computed by dividing the summation of *count* values on the pixel mean. At that case we were able to classify pixels based on the mean above threshold *Mean_Above_Thresh* and the mean *Mean_Below_Thresh*. Thus, the mean of pixels above was calculated and the mean of pixels below was also computed. By dividing the summation of both mean below / above thresholds over 2 the new Threshold *T(i)* was calculated. Consequently, we repeated these steps until the absolute difference between *T (i)* and *T (i-1)* would be less than 1. Table 4.5 shows the results of threshold generation using the iterative threshold algorithm (See Appendix D for a table shows the whole results of iterative experiment).

Table 4.5 The generated thresholds during the execution of the iterative threshold program in MATLAB and the calculated *Mean_ Below_ Thresh* and *Mean_Above_Thresh*

Loop (i)	Threshold (T)
1	152
2	143
3	137
4	133
5	130
6	128
7	127
8	127
Mean Below Threshold	85.80894379
Mean Above Threshold	167.3627977
Gray Level Threshold	0.496062992

Basically, after generating the target threshold we made an application in C#.NET programming language to binarize the target image based on the computed target threshold as well as generating the bounding boxes on the detected text. Hence, the application reads user's threshold (i.e. the calculated iterative threshold) in order to binarize the image. In this context, the binirized image should contain the highest

المنارة للاستشارات
www.manaraa.com

intensity pixels. Figure 4.6 shows the Graphical User Interface (GUI) of the iterative threshold application.

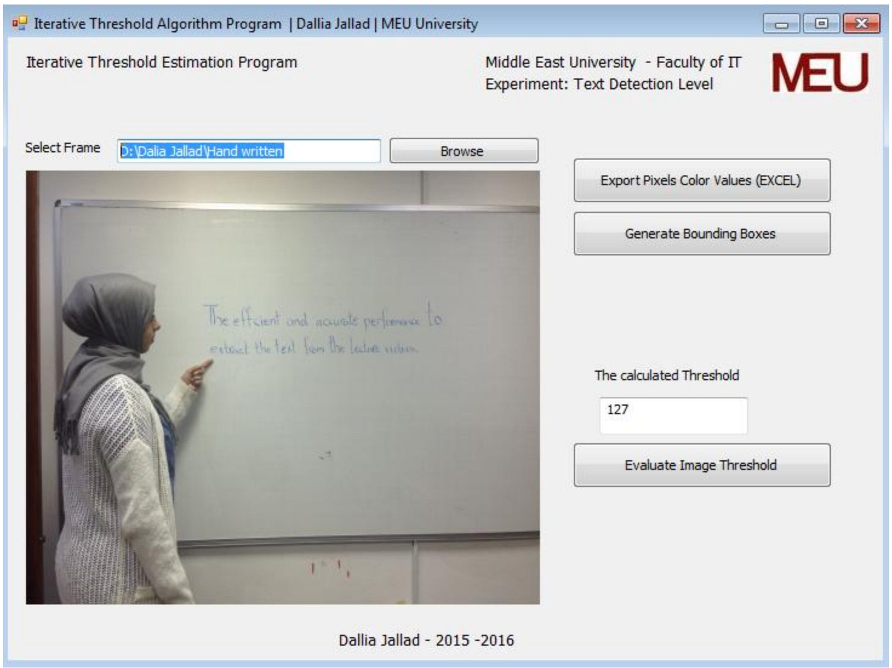


Figure 4.6 The GUI of the iterative threshold application

The results of the iterative threshold approaches showed that the binarized image successfully detected the text area. On the other hand, the shape of the text was not accurate. Figure 4.7 shows the binarized image and the generated bounding boxes.

www.manaraa.com

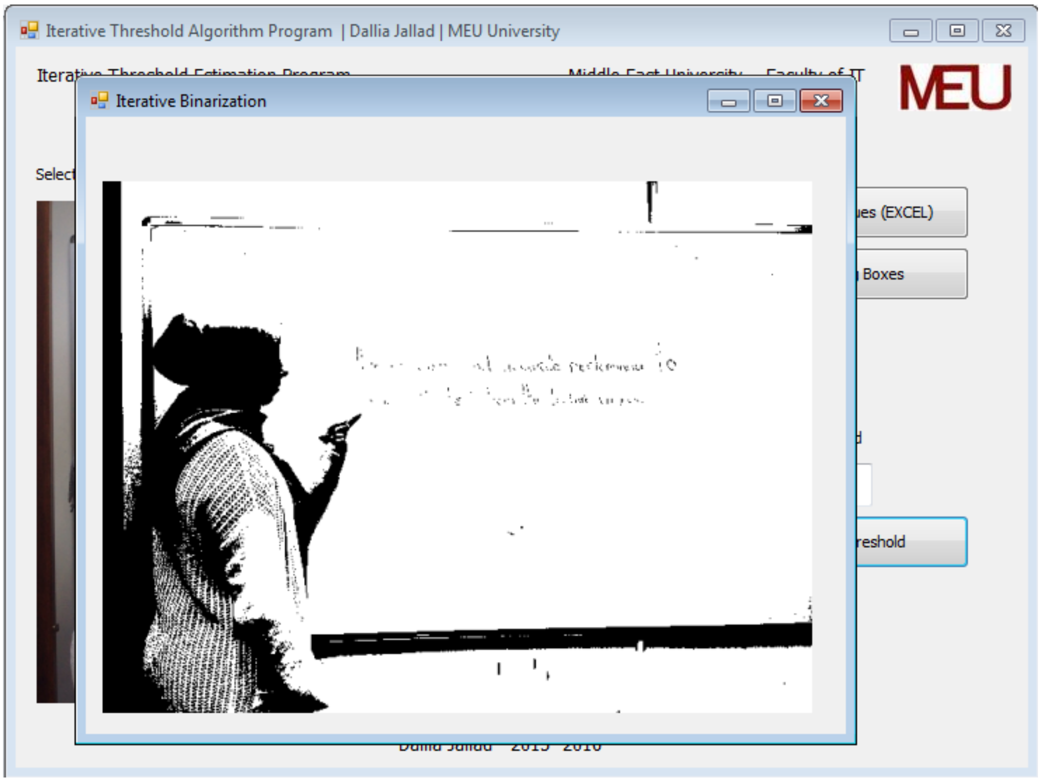


Figure 4.7 Frame image binarization results using the iterative threshold technique

Thus, the text was detected in the image frame and the bounding box was generated to localize the text in the frame image. Figure 4.8 shows the frame image with the bounding boxes after binarization using the iterative technique.

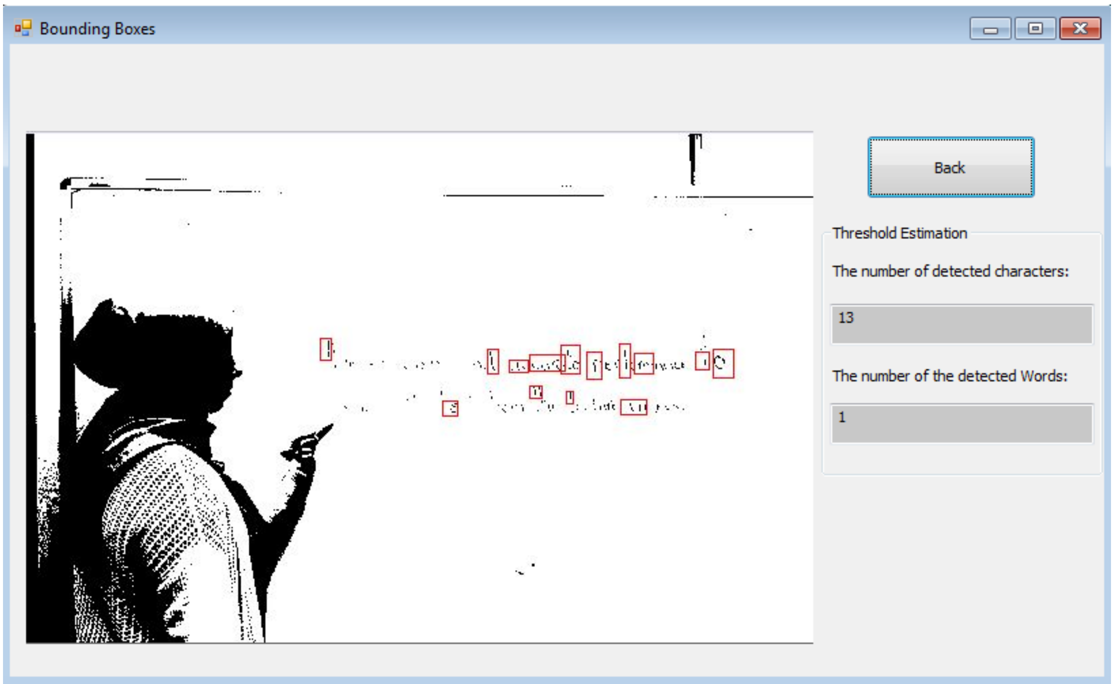


Figure 4.8 The generated bounding boxes for the detected text in frame image using the iterative technique.

المنارة للاستشارات
www.manaraa.com

Standard OTSU's threshold algorithm analysis

This section discussed the standard OTSU thresholding algorithm in term of the binarized image and the bounding boxes. Therefore, the results in this section concentrated on two variables which were the estimated threshold number and the number of detected characters as well as the words. For this purpose, the code of OTSU thresholding algorithm was implemented using MATLAB. Figure 4.9 shows the MATLAB implementation code for OTSU in term of generating the target threshold.

```
%-- 17/12/2015 09:31 م --%
% Dalia Jallad "OTSU Thresholding Technique"
%Middle East University Amman - Jordan

Dalia = imread('bmp dalia.bmp');
Dalia = im2uint8(Dalia(:));
% Step1: Compute the histogram and probabilities of each intensity level

% Step 1.1: Assign Number of histogram bins
hist_nbins = 256;

% Step 1.2: Generating image histogram (0-256) based on pixel intensity
counts  = imhist(Dalia,hist_nbins);

% Step 1.3: Computing The Percentage of Pixels Matrix
%              To Evaluate Pixel's Percent to the Summation of Pixel's values
p = counts / sum(counts);

%Step2: Setting up the initialization Matrices
% Step 2.1: Non_Textual_ghaws(0) and its mean mu_1(0)
non_txt_ghaws = 0;
mu1 = 0;
% Step 2.2: Textual_ghaws(0) and its mean mu_2(0)
Textual_ghaws = 1;
mu2 = mean(Dalia(:));

% Step 3: Step through all possible thresholds from 0 to maximum intensity (255)
% Step 3.1: Update Ghaussian CLasses and Means mu_i
% Step 3.2: Compute The Target Threshold

for t = 1:hist_nbins
   non_txt_ghaws(t) = sum(p(1:t));
   Textual_ghaws(t) = sum(p(t+1:end));
   mu1(t) = sum(p(1:t).*(1:t)');
   mu2(t) = sum(p(t+1:end).*(t+1:hist_nbins)');
end
% Step 4: OTSU Threshold Estimation
OTSU_Threshold = non_txt_ghaws .* Textual_ghaws .* (mu2-mu1).^2;

%Step 5: The desired threshold is the maximum of estimated Threshold
[~,threshold] = max(OTSU_Threshold);
```

Figure 4.9 The OTSU algorithm implementation code in MATLAB

المنارة للاستشارات
www.manaraa.com

Basically, the implementation of OTSU algorithm was achieved in five basic sequential steps in order to estimate the target threshold. In the first step, generate the grayscale histogram with a 256 color bins as well as it was assigned to *count* matrix in order to calculate the ratio of each pixel to the summation of all pixel's values and it was assigned to *p* matrix. In the second step, initialize two matrices the *non_txt_ghaws* that was responsible on representing the first Gaussian class, and *Textual_ghaws* that was responsible on representing the second Gaussian class. Hence, the mean of the first matrix *mu1* was initialized with zero as well as the mean of the second matrix *mu2* was initialized with the mean of the grayscale image. In the third step, the first Gaussian class matrix was updated due to the summation of the pixels till the current. In contrast, the second Gaussian class was updated due to the summation of current till the end of pixels without the pixels that was calculated in the first class matrix. Thus, the mean of each class matrix was updated *mu1* and *mu2* respectively. In the fourth step, the OTSU threshold matrix was estimated based on the formula in chapter 3. In the fifth step, the target threshold was picked as the maximum threshold in the OTSU threshold matrix (See Appendix E that shows the table for the generated matrices and our calculations).

Consequently, the estimated target threshold which was 163 was used in an application developed in C#.NET to find the binarized image and generating the bounding boxes. Figure 4.10 shows the binarized image with 163 as a threshold. The resultant binary image showed an enhanced text shape than the result of the iterative thresholding in the previous section. However, some of the text region shapes was inaccurate. On the other hand, the target threshold detected the text regions successfully.

المنارة للاستشارات
www.manaraa.com

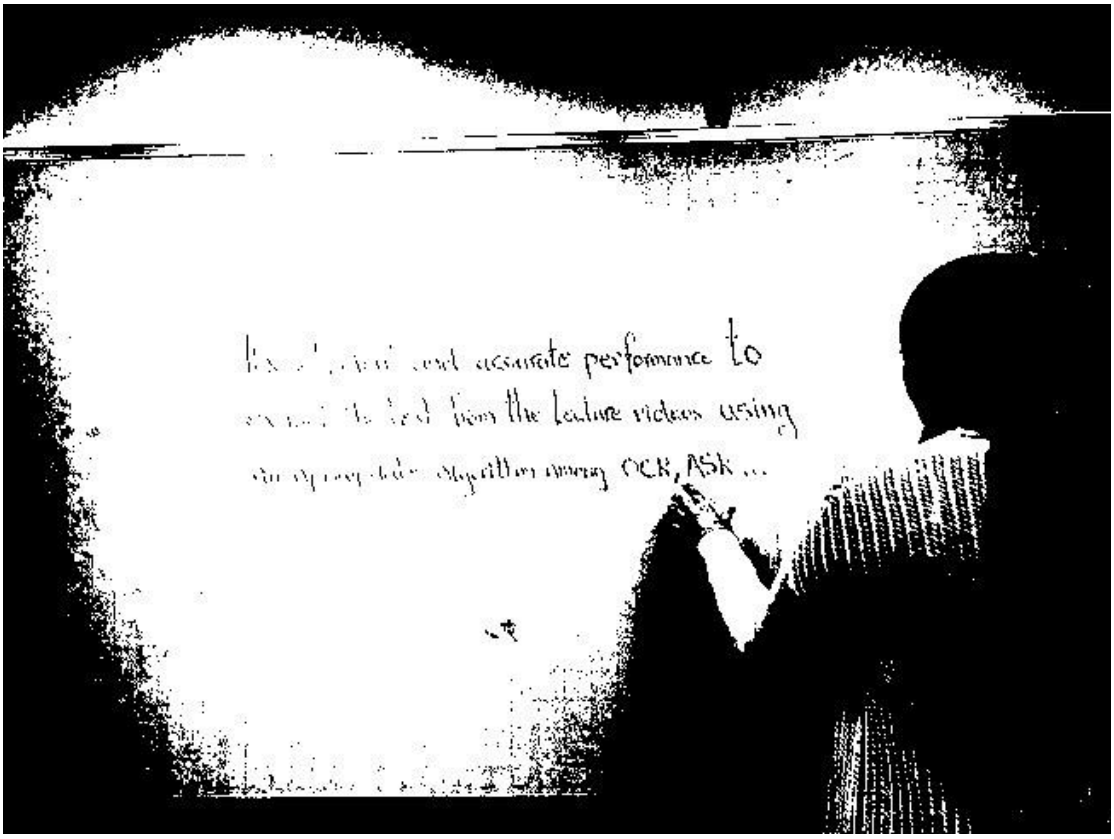
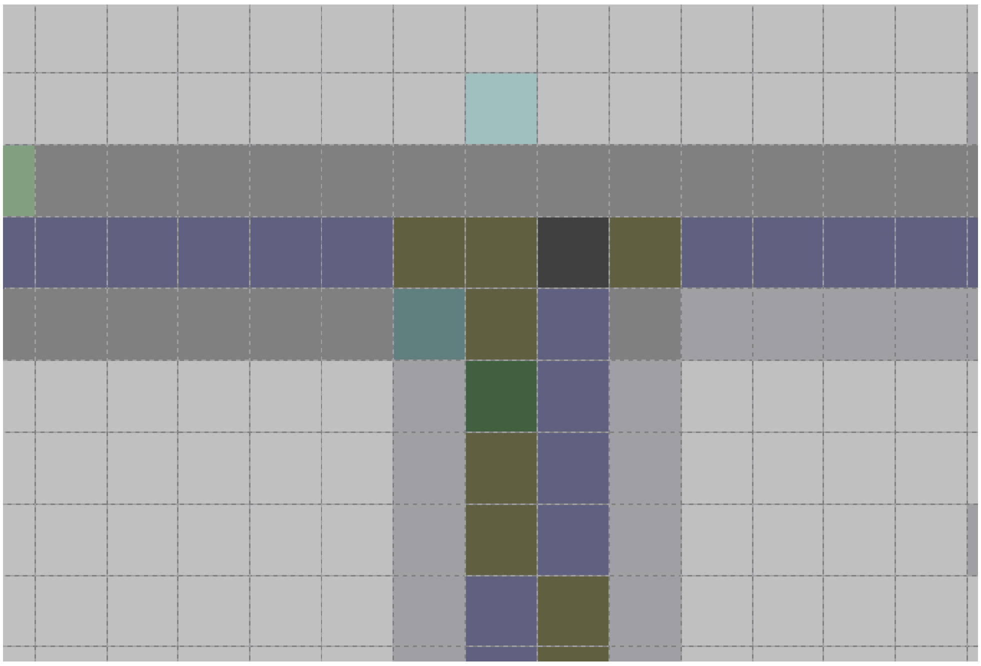


Figure 4.10 The binarized image after implementing OTSU's threshold

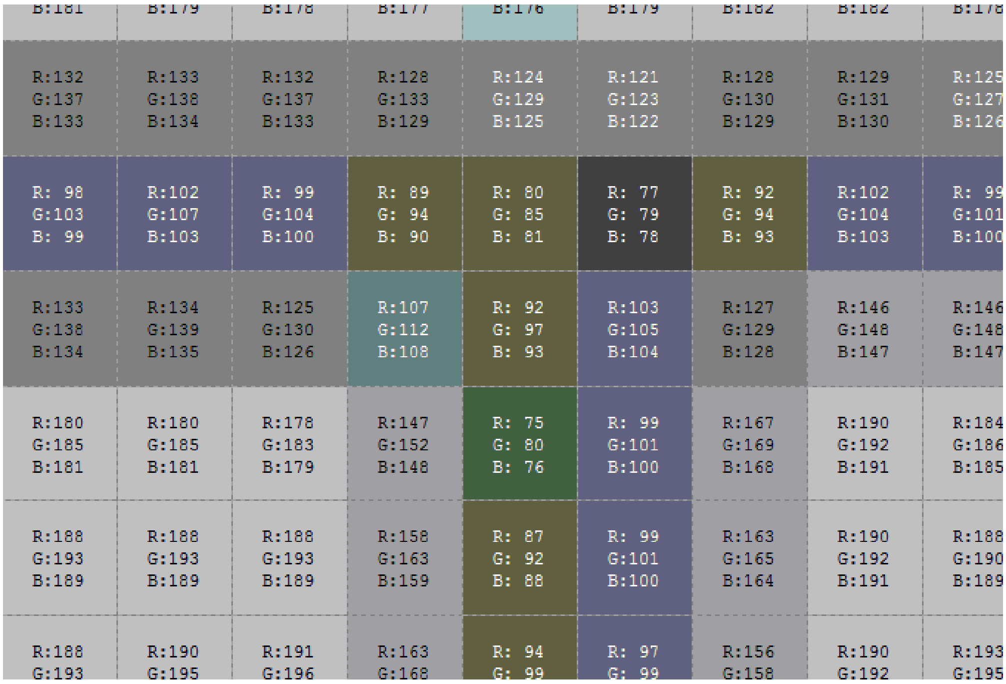
Author's enhancement on OTSU's Threshold estimation

This section discussed author's enhancement on estimating the image's threshold using the standard OTSU algorithm. Actually, the primary idea in OTSU algorithm is to use the grayscale image to minimize the intra class variance between the black and white boxes. Thus, the estimated threshold represents the criteria of the pixel's occurrence in the binarized image. Therefore, our enhancement took into consideration the intensity of the pixel's values that represent the adjacent pixels of the textual pixels. Thus, the intensity of the adjacent textual pixels showed small differences compared with the textual pixel's intensity. Thus, these pixels could be taken into consideration to enhance the binarized image. Figure 4.11 shows the intensity of textual pixels and their adjacent pixels.

المنارة للاستشارات
www.manaraa.com



(a) Snapshot of textual pixels with their adjacent pixels



(b) Textual pixel's values with adjacent pixel's values

Figure 4.11 Sample video frame image shows the same image with/without pixel's values

Consequently, the proposed enhancement concentrated on creating third class for the adjacent pixels called *intensity_class*. Thus, we calculated the ratio between the pixel's mean of the first Gaussian class *mu1* with the overall mean of the first Gaussian

www.manaraa.com

class, in order to find the ratio of intensity of each pixel in the first class. Thus, we calculated the standard deviation of the generated *intensity_class* as well as the mean of its elements. Therefore, the normal distribution (i.e. Gaussian) was generated for the values in the third class (i.e. *intensity_class*) in order to calculate the Probability Density Function (PDF). Figure 4.12 shows a graph of the probability density with the pixel values of the third class.

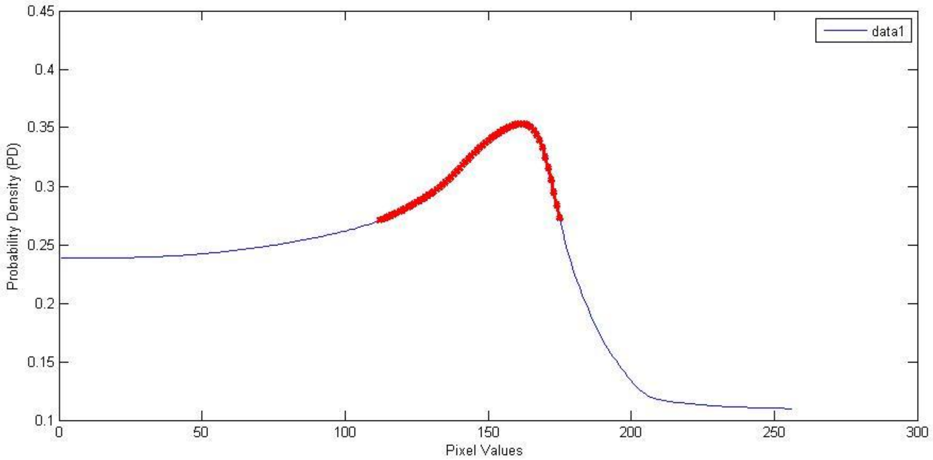


Figure 4.12 Test histogram of the probability density with pixels' values in the proposed intensity class

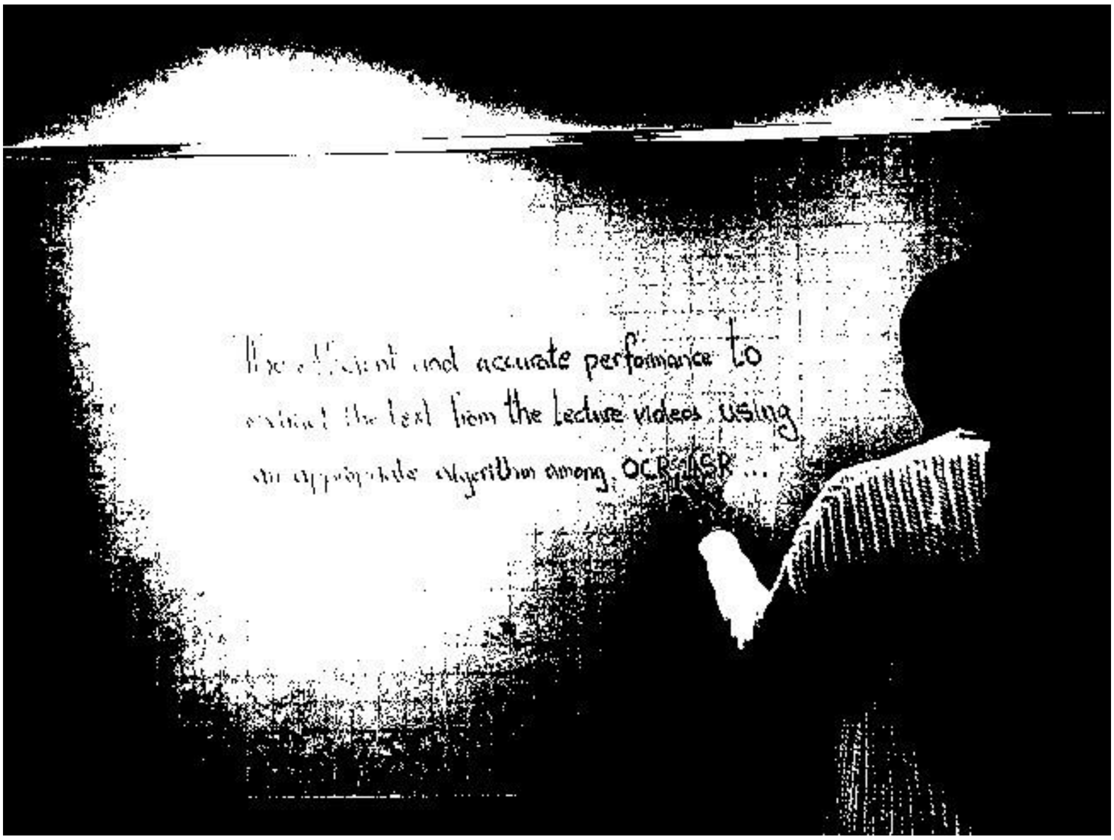
The probability of density for the adjacent pixels were increasing drastically from 120 pixel value till the peak in 163 that means the adjacent pixels were detected first till detecting the textual pixels. In contrast, the probability density were decreasing drastically too after achieving the peak. The probability density for the textual pixels and their adjacent pixels were varied between 2.6 and 3.8.Therefore, the mean of the adjacent class array was computed as well as the OTSU threshold equation was multiplied with the new adjacent array values with subtracting the mean of the adjacent class array mu3 from the basic two OTSU Gaussian classes. Figure 4.13 shows the MATLAB code for the enhanced OTSU algorithm.

www.manaraa.com

```
%Generate the intensity class by finding the ratio of each pixel's mean with the overall mean

for count=1:hist_nbins
intensity_class(count)=mu1(count)/mean(mu1);
end
%Compute the standard deviation for the first class (non_txt_ghaws)
TT=std(intensity_class);
%Compute the mean for the first class (non_txt_ghaws)
TT1=mean(intensity_class);
% Define the Normal Distribution of the pixels in non_txt_ghaws
pd=makedist('Normal',TT1,TT);
%Generate the third Class
adj_txt_ghaws=pdf(pd,intensity_class);
%Compute the mean of the adjcent class
mu3=mean(adj_txt_ghaws);

OTSU_Threshold2 = adj_txt_ghaws.*non_txt_ghaws .* Textual_ghaws .* ((mu2-mu3)-(mu1-mu3)).^2;

% The desired threshold is the maximum of estimated Threshold
 [~,threshold2] = max(OTSU_Threshold2);
```

Figure 4.13 The MATLAB code for the proposed enhancement on OTSU threshold

The generated threshold after the enhancment was 172. The binarized image showed more dark areas as well as more text shapes. Figure 4.14 shows the binarized image after the proposed enhancement.



المنارة للاستشارات
www.manaraa.com

Figure 4.14 The binarized image of the proposed enhanced OTSU threshold algorithm

4.4. The experimental results of text recognition level

In this section, the results of text recognition level were discussed in term of performance based on the precision and recall measures. Therefore, the results was classified based on the type of OCR was used and the technique used to binarize the input frame image. The recall and precision was calculated by computing the correctly selected texts which was represented as TP, the correctly detected text but without selection that was represented as FN, the not correctly recognized text and selected that was represented as FP (See Appendix F for a table shows the results of recognition for 194 image frames).

CPU Transaction Time (ms)				
		The used Threshold Algorithm		
Text Recognition	Tools	Iterative	Standard OTSU	Enhanced OTSU
	Free OCR	8155.3	7423.9	6993.4
	Tesseract OCR	6467.4	5795.4	4767.8
	OCR-AD	5665.2	4387.3	4107.3
	Average	**6762.63**	**5868.87**	**5289.5**

Table 4.6 The average of averages of CPU response time for each text extraction in each video

The comparison in this level took into consideration three types of open source commercial OCR applications which were Tesseract OCR, OCR-AD, and Free-OCR. On the other hand, the binarized images were processed using three techniques the iterative thresholding, the standard OTSU, and the proposed enhancement on OTSU algorithm. Table 4.7 the calculated precision and recall for the variables in the comparison for this section.

المنارة للاستشارات
www.manaraa.com

Table 4.7 The results of precision and recall for the used thresholding techniques with several OCR software programs (Tesseract, OCR-AD, and Free-OCR).

The Used Technique	The used Recognition S/W	Recall	Precision	F 1 Measure
Iterative	*OCR*	0.77	0.85	0.905
Basic OTSU	*OCR*	0.782	0.879	0.88
Proposed Threshold	*OCR*	0.7905	0.8977	0.9
Average		**0.7808333**	**0.875567**	
Iterative	*OCR - Tesseract*	0.798	0.891	0.895
Basic OTSU	*OCR - Tesseract*	0.9	0.96	0.937
Proposed Threshold	*OCR - Tesseract*	0.92	0.98	0.938
Average		**0.8726667**	**0.943667**	
Iterative	*OCR - AD*	0.654	0.702	0.931
Basic OTSU	*OCR - AD*	0.639	0.7211	0.886
Proposed Threshold	*OCR - AD*	0.735	0.7811	0.94
Average		**0.676**	**0.7347**	

The results showed that the proposed enhancement had a positive effect in precision and recall performance metrics for all text recognition applications. Thus, the iterative thresholding binarized images achieved the lowest results in precision and recall. Furthermore, no drastic gaps between the enhanced OTSU and the standard OTSU thresholding algorithms. Furthermore, after calculating the F1 measure that was used to represent the accuracy of methods were selected in this study as shown in Table 4.7 the our enhancement OTSU method showed higher accuracy rates compared with iterative and Standard OTSU methods. Therefore, in case of using Free OCR based on iterative, standard OTSU, and Our enhanced OTSU the results were 0.865, 0.937, and 0.938 respectively.

From text recognition applications point of view, the results showed that the Tesseract OCR achieved the best precision and recall metrics. In contrast, the OCR-AD text recognition application had a negative effect on recall and precision. Furthermore, Free-OCR application achieved balanced recall and precision.

المنارة للاستشارات
www.manaraa.com

Consequently, the total system's performance was measured by calculating the accuracy of the whole system in term of finding the ratio between the correctly classified to the total number of videos were taken in the experiments (See Appendix G for a table shows the full details of the videos were taken in the experiments and the results of its classification based on the criteria was proposed in this study).

$$Accuracy = \frac{Correctly\_Classified}{Number\_\text{of\_Videos}} \times 100\%$$

$$Error\_Rate = \frac{In\_Correctly\_Classified}{Number\_\text{of\_Videos}} \times 100\%$$

$$Accuracy = \frac{49}{56} \times 100\% \qquad Error\_Rate = \frac{7}{56} \times 100\%$$

$$Accuracy = 87.5\% \qquad Error\ Rate = 12.5\ \%$$

المنارة للاستشارات

www.manaraa.com

CHAPTER FIVE

Conclusions and Future Research Works

5.1. Overview

This chapter summarizes the conclusions of this study. Thus, the main theme of this study was to focus on the way of identifying the lecture video type based on the similarity between video frames. Several experiments were conducted to extract the effects of lecture video contents on the similarities between video frames, to show the effects of thresholding techniques on the detecting text from frame images, and to find the performance in term of recall and precision of several OCR engines. Thus, this chapter was organized as section 5.2 to discuss the main conclusions, section 5.3 to preview the limitations of this study, and section 5.4 to preview the future research works in the field of lecture video classification and textual data recognition.

5.2 Conclusions

According to study goals and the experimental results, the similarity between video frames had a direct effect on classifying the lecture video type. Therefore, the empirical finding showed that the similarity mean of the detected key frames was varied depending on lecture video type. Hence, result found out that the mean of presentational lecture video key frames was less than 0.4, in case of handwritten lecture video the mean of similarity between video key frames was bounded between 0.4 and 0.8, and the spoken text lecture videos achieved the highest similarity which their similarity exceeded 0.8 in all sample videos.

المنارة للاستشارات
www.manaraa.com

Furthermore, this study took into consideration several common thresholding algorithms that are used in detecting textual data in video frame images in order to find the effects of threshold value on the binarized image. Thus, we proposed to compare the iterative thresholding and the standard OTSU thresholding algorithms that are widely referenced in the literature. For comparison purposes, we assigned some enhancement on the OTSU algorithm in term of using a third class that detecting the adjacent textual pixels. Consequently, the empirical findings showed that the threshold value in the iterative algorithm was the lowest value that caused a missing of large number of pixels in the binarized image as well as inaccurate text shapes. In contrast, using OTSU algorithm the estimated threshold value generate better binary image than in iterative algorithm. Furthermore, the enhancement on OTSU algorithm had a positive effect on the binary image in term of the number of missing pixels which was the lowest compared with the other algorithms in this study, as well as, the text shape was more accrue which increase the chance to be recognized using OCR engines. In contrast, the binary image after enhancing the OTSU contained more dark areas which may be a negative indicator in other fields such as noise detection or image closed object detection.

Consequently, other criteria took place in this study which was to detect the effect of the binary image on the performance of the used OCR engine. For this purpose, three commonly OCR engines have been used which were Free-OCR, OCR-AD, and Tesseract OCR engines. Thus, we conducted several experiments to calculate the performance in term of recall and precision. Therefore, the empirical finding s showed that the Tesseract OCR engine has the highest performance using the enhanced binary image. Furthermore, the OCR-AD achieved the lowest performance compared with the

المنارة للاستشارات

www.manaraa.com

other OCR engines. Hence, the binary image had a direct effect on the performance of OCR engine. Thus, the result found out that using the iterative threshold binary image the performance was decreased drastically. On the other hand, the result found out that using the enhanced OTSU algorithm binary image the performance was increased.

5.3. Research questions discussion

In this section, I first answer the research questions posed in Chapter 1. After that I give a discussion on how I achieved the research goal of this thesis.

5.3.1. Research Question 1: What are the issues are affected with lecture video type?

Through my research and implementation of various lecture video types, I have found that a combination of issues were affected based on the type lecture video such as the similarity of lecture video frames, the number of lecture video key frames, and the color depth of lecture video frames.

5.3.2. Research Question 2: What are the issues that affect the thresholding value in text detection level?

Through my research and implementation of various digital video processing algorithms (i.e. methods), threshold value is affected by the type of lecture video and the duration of lecture videos. Thus, threshold value was varied by changing the period of lecture video which reflects another factor that can be studied in future works.

المنارة للاستشارات
www.manaraa.com

5.3.3. Research Question 3: How do we select the suitable text extraction engine based on lecture video type?

Precision and recall (and accuracy) was chosen as the metrics to evaluate the different approaches. The decision was based on what other researchers had used to evaluate their approaches, as well as my own experience in using these metrics successfully in previous projects. Precision and recall goes well together in describing the performance of a classification system. Therefore, choosing the suitable text extraction system was by applying edge detection for handwritten lecture videos, ASR (e.g. HMM HTK speech recognition) for spoken lecture videos because of the highest margin of frame similarity between its frames, and applying OCR engines for presentational lecture videos.

5.4 Future research works

This research opens the door for finding the suitable text detection algorithm with suitable OCR engine to achieve the best results based on identifying the type of lecture video. Thus, the results of this study can be used to design an automatic text information recognition system that has intelligent features in order to increase the performance of indexing and archiving processes in online video repositories. This research work opens the door to evaluate the accuracy using F1-measure to cover more thresholding techniques in the literature. Covering more domains using the proposed test experiment will enable more and more domains to be evaluated. Also, achieving more accurate results is still topic of continuous and constant research.

المنارة للاستشارات

www.manaraa.com

References

- Anthimopoulos, M., Gatos, B., and Pratikatis, I., (2010). *A Two Stage Scheme for Text Detection in Video Image*. Image Vision Comput., PP. (1413-1426), January, 2010. Vol. 28, Issue. 9, DOI: 10.1016/j.imavis.

- Cooper, M., (2013). *Presentation Video Retrieval Using Automatically Recovered Slide and Spoken Speech text*. Proceedings SPIE Multimedia Content and Mobile Devices, March, 2013, Volume 8667, DOI: 10.1117/12.2008433.
- Ding, H., Pan, W., Guan, Y., (2010), *Video Analysis and Indexing*. Digital Video, Floriano De Rango (Ed), PP. (360-380), InTech, ISBN: 978-953-7619-70-1.
- Gallavat, J., Qeli, E., and Friesleben, L., (2007). *Detecting Text in Video Using Fuzzy Ensembles*. Eight IEEE International Symposiums in Multimedia. PP. (283-290), DOI: 10-1109/ISM.2006.60, ISBN: 0-7695-2746-9.
- Gonzalez, R. and Woods, R., (2008). *Digital Image Processing*. 2008 Pearson Education Inc., Upper Saddle River, New Jersey. ISBN: 978-0-13-168728-8.
- Jeong, H., Kim, T., and Kin, H., (2015). *Automatic Detection of Slide Transition in Lecture Videos*. Multimedia Tools and Application Conference, January, 2015, Volume 74, Issue 18, PP. 7537-7554, DOI: 10.1007/s11402-014-19906.
- Jung, K., Kim, K., and Jain, A., (2004). *Text Information Extraction in Images and Videos: A Survey*. Elsevier Ltd, PP. (977-997), DOI: 10.1016/j.patog.2003.10.12.

المنارة للاستشارات

www.manaraa.com

- Kamabathula, V., and Iyer, S., (2011). *Automated Tagging To Enable Fine-Grained Browsing of Lecture Videos*. IEEE International Conference on Technology for Education, T4E 2011, Tamil Nadu, India. PP. (96-102), DOI: 10.1109/T4E.2011.23, ISBN: 978-0-7695-4534-9.
- Khan, M., Lanote, S., and Solariki, S. (2014). *A Review: Video Searching Using Speech and Video Text Information*. International Journal of Electrical, Electronics, and Computer Engineering, PP. 169-173, Volume 3, Issue 2, DOI: 169-173 (2014), ISBN: 2277-2626.
- Kim, K., Jung, K., and Kim, J., (2003). *Texture-Based Approach fir Text Detection in Images Using Support Vector Machine and Continuously Adoptive Mean Shift Algorithm*. Transaction on Pattern Analysis and Machine Intelligence, PP. (1631-1639). Vol. 25, Issue. 12, IEEE Computer Society.
- Lawate, A., and Wankkhade, M., (2015). *Automatic Speech Recognition System for Lecture Videos*. International Journal of Engineering Science and Research Technology. July, 2015. Impact Factor: 3.785. PP. (911-916), Vol. 4, Issue. 7, ISSN: 2277-9655.
- Lebourgeois, F., (1997). *Robust multi-font OCR system from gray level images in Document Analysis and Recognition.* Proceedings of the Fourth International Conference on, vol.1, no., pp.1-5 vol.1, 18-20 Aug 1997. DOI: 10.1109/ICDAR.1997.619803
- Leon, M., Vilaplana, V., Gasull, A., and Marques, F. (2013). *Region Based Caption Text Extraction*. Analysis, Retrieval and Delivery of Multimedia Content, New York, 2013, PP. 21-36, DOI: 10.1007/978-1-4614-3831-1-2, Springer Science + Business Media.

www.manaraa.com

- Lienhart, R., and Effelsberg, W., (2000). *Automatic Image Segmentation and Text Recognition for Video Indexing*. PP. (69-81), Vol. 8, Issue 1, Multimedia Systems, DOI: 10.1007/s005300050006.
- Martinez-Villaronga, A., Agua, M., Cerd, J., Ferrer, J., and Juan, A. (2014), *Language Model Adaption for Lecture Transcript by Document Retrieval*. Advances in Speech Language Technologies for Liberian Languages Second International Conference, Lecture Notes in Computer Science 8854, Springer, 2014. PP.(129-137), DOI: 10.1007, ISBN: 978-3-319—13622-6.
- McAndrew, A., (2004). *An Introduction to Digital Image Processing with MATLAB*, Victoria University of Technology, USA.
- Metai, O., Pop, P., and Valean, H., (2013). *Optical Character Recognition in Real Environments Using Neural Networks and K-nearest Neighbour*. Applied Intelligence, July, 2013, PP. 739-748, Volume 39, Issue 4, DOI: 10.1007/s10489-013-0456-2.
- Makkar, H. and Pundir, A. (2014). *Image Analysis Using Improved OTSU's Thresholding Method.* International Journal on Recent and Innovation Trends in Computing and Communication, vol. 2, no., 8, pp. 2122-2126. ISSN: 2321-8169.
- Ngo, C., and Chan, C., (2004). *Video Text Detection and Segmentation for Optical Character Recognition*. Multimedia Systems, PP. 261-272, Volume 10, DOI: 10.1007/s00530-004-0157-0.

المنارة للاستشارات

www.manaraa.com

- Patel, B., and Meshram, J., (2012). *Content Based Video Retrival Systems*. International Journal of UbuComp. (IJU), PP. (13-30), Vol. 3, Issue 2, April. 2012, DOI: 10.5121/iju.2012.3202.
- Peng, X., Cao, H., Parasad, R., and Natrajan, P., (2011). *Text Extraction from video using Conditional Random effect fields*. IEEE International Conference on Document Analysis and Recognition, PP. 1029-1033, DOI: 10.1109/CDAR.2011.
- Reddy, P., and Devaraju, M. (2015). *Optical Character Recognition Based on Video Retrieval Application for High End Applications*. International Journal and Magazine of Engineering Technology Management and Research, August, 2015. PP. 1202-1208, Volume 2, Issue 2, ISBN: 2348-4845.
- Richter, F., Ries, C., Romberg, S., and Lienhart, R., (2014). *Partial Contour Matching for Document Pieces With Content Based Prior*. PP. (1-6), IEEE International Conference on Multimedia and Expo (ICME 2014), Chengdu, China. DOI: 10.1109/ICME.2014, ISBN: 978-4799-4761-4.
- Sharma, N., Shivkumara, P., Pal, U., Blumenstein, M., and Tan, C. (2012). *A New Method for Arbitrarily Oriented Text Detection in Video*. International Workshop on Document Analysis Systems, Queen Lands, Australia, March, 2012, PP. 74-78, IEEE Computer Society, DOI: 10.1109/DAS.2012.6, ISBN: 978-1-4673-0868-7.
- Shivakumara, P., Phan, T., and Tan, C., (2011). *A Lapalacian Approach to Multi-Oriented Text Detection in Video*. IEEE Transaction on Pattern Analysis and Machine Intelligence, PP. 413-419, Volume 33, Issue 2, DOI: 10.1109/TPAMI.2010.166, ISBN: 0162-8828/11.

المنارة للاستشارات

www.manaraa.com

- Solomon, C., (2011). *Fundametal of Digital Image Processing*. Wiely-Blackwell 2011 by John Wiley & Sons Ltd. West Sussex, UK. ISBN: 978-0-470-844724.
- Tekinlap, S., and Altan, A., (2013). *Utilization of texture, contrast, and color homogeneity for detecting and recognizing text from video frames*. In proceedings of the 2013 International Conference on Image Processing (ICIP) Barcelona, Catalonia, Spain. PP. (505-508), DOI: 10.1109/ICIP.2013, ISBN: 12472092.
- Trier, O., and Jain, A., (1995). *Goal Directed Evaluation of Binarization Method*. IEEE Transaction on Pattern Analysis and Machine Intelligence. PP. (1191-1201), United Kingdom.
- Tuna, T., Subhlok, J., and Shah, S., (2011). *Indexing and Keyword Search to Ease Navigation in Lecture Videos*. IEEE Applied Imagery Pattern Recognition Workshop AIPR, Washington DC, USA, October, 2011, PP. 1-8, DOI: 10.1109/AIPR.2011.6176364.
- Yang, H., Quehl, B., and Sack, H. (2012). *A Skeleton Based Binarization Approach For Video Text Recognition*. The 13th International Workshop on Image Analysis for Multimedia, 2012, Dublin, Ireland, PP. 595-608, DOI: 10.1109/WIAMIS.2012.6226754.
- Yang, H., Siebert, M., Lutine, P., Sack, H., and Meniel, C., (2011). *Automatic Lecture Video Indexing Using OCR Technology*. IEEE International Symposium on Multimedia, PP. 54-61 , DOI: 10.1109/ISM.2011.26, ISBN: 978-0-7695-4589-9/11.

المنارة للاستشارات
www.manaraa.com

- Ye, Q., and Doermann, D., (2013). *Scene Text Detection Via Integrated Discrimination of Component Appearance and Consensus*. Camera-Based Document Analysis and Recognition 5th International Workshop, CBDAR,2013, Washington, USA. PP. (47-59), Lecture Notes in Computer Science, Springer, 2014. DOI: 10.101007. ISBN: 978-3-319-05166-6.
- Zhao, X., Lin, K., Fu, Y., Liu, Y., and Hury, T., (2011). *Text From Corners: A Novel Approach to Detect Text and Caption in Video*. IEEE Transaction on Image Processing, March, 2013, PP. 790-799, Volume 20, Issue 3, DOI: 10.1109/TIP.2010.2068553.
-

المنارة للاستشارات

www.manaraa.com

Appendices

Appendix A

Histogram color values for each pixel for three sequenced frames.

FRAME #1		FRAME #2		FRAME #3	
Pixel, Color Value	Color Value	Pixel, Color Value	Color Value	Pixel, Color Value	Color Value
0.5, 45	**45**	0.5, 211	**211**	0.5, 0	**0**
1.5, 3	**3**	1.5, 12	**12**	1.5, 0	**0**
2.5, 0	**0**	2.5, 7	**7**	2.5, 0	**0**
3.5, 4	**4**	3.5, 23	**23**	3.5, 0	**0**
4.5, 4	**4**	4.5, 24	**24**	4.5, 0	**0**
5.5, 7	**7**	5.5, 18	**18**	5.5, 0	**0**
6.5, 5	**5**	6.5, 12	**12**	6.5, 0	**0**
7.5, 2	**2**	7.5, 19	**19**	7.5, 1	**1**
8.5, 4	**4**	8.5, 22	**22**	8.5, 2	**2**
9.5, 8	**8**	9.5, 20	**20**	9.5, 0	**0**
10.5, 9	**9**	10.5, 16	**16**	10.5, 1	**1**
11.5, 10	**10**	11.5, 20	**20**	11.5, 0	**0**
12.5, 6	**6**	12.5, 22	**22**	12.5, 1	**1**
13.5, 2	**2**	13.5, 29	**29**	13.5, 0	**0**
14.5, 8	**8**	14.5, 30	**30**	14.5, 0	**0**
15.5, 7	**7**	15.5, 18	**18**	15.5, 0	**0**
16.5, 9	**9**	16.5, 23	**23**	16.5, 0	**0**
17.5, 3	**3**	17.5, 26	**26**	17.5, 1	**1**
18.5, 13	**13**	18.5, 26	**26**	18.5, 2	**2**
19.5, 8	**8**	19.5, 23	**23**	19.5, 1	**1**
20.5, 8	**8**	20.5, 37	**37**	20.5, 1	**1**
21.5, 9	**9**	21.5, 23	**23**	21.5, 0	**0**
22.5, 15	**15**	22.5, 31	**31**	22.5, 0	**0**
23.5, 12	**12**	23.5, 31	**31**	23.5, 0	**0**
24.5, 9	**9**	24.5, 45	**45**	24.5, 1	**1**
25.5, 16	**16**	25.5, 34	**34**	25.5, 2	**2**
26.5, 6	**6**	26.5, 30	**30**	26.5, 2	**2**
27.5, 13	**13**	27.5, 42	**42**	27.5, 1	**1**
28.5, 10	**10**	28.5, 36	**36**	28.5, 3	**3**
29.5, 13	**13**	29.5, 42	**42**	29.5, 2	**2**
30.5, 15	**15**	30.5, 38	**38**	30.5, 6	**6**
31.5, 12	**12**	31.5, 48	**48**	31.5, 25	**25**
32.5, 12	**12**	32.5, 58	**58**	32.5, 20	**20**
33.5, 15	**15**	33.5, 67	**67**	33.5, 70	**70**

www.manaraa.com

34.5, 18	**18**	34.5, 66	**66**	34.5, 116	**116**
35.5, 17	**17**	35.5, 59	**59**	35.5, 271	**271**
36.5, 9	**9**	36.5, 72	**72**	36.5, 187	**187**
37.5, 16	**16**	37.5, 76	**76**	37.5, 357	**357**
38.5, 14	**14**	38.5, 96	**96**	38.5, 324	**324**
39.5, 9	**9**	39.5, 59	**59**	39.5, 405	**405**
40.5, 15	**15**	40.5, 61	**61**	40.5, 381	**381**
41.5, 28	**28**	41.5, 72	**72**	41.5, 458	**458**
42.5, 21	**21**	42.5, 71	**71**	42.5, 393	**393**
43.5, 21	**21**	43.5, 83	**83**	43.5, 322	**322**
44.5, 22	**22**	44.5, 63	**63**	44.5, 644	**644**
45.5, 18	**18**	45.5, 67	**67**	45.5, 453	**453**
46.5, 27	**27**	46.5, 73	**73**	46.5, 948	**948**
47.5, 28	**28**	47.5, 61	**61**	47.5, 410	**410**
48.5, 31	**31**	48.5, 57	**57**	48.5, 1256	**1256**
49.5, 24	**24**	49.5, 102	**102**	49.5, 564	**564**
50.5, 20	**20**	50.5, 82	**82**	50.5, 2189	**2189**
51.5, 17	**17**	51.5, 76	**76**	51.5, 573	**573**
52.5, 21	**21**	52.5, 72	**72**	52.5, 1197	**1197**
53.5, 21	**21**	53.5, 84	**84**	53.5, 507	**207**
54.5, 26	**26**	54.5, 69	**69**	54.5, 1229	**1229**
55.5, 22	**22**	55.5, 63	**63**	55.5, 475	**475**
56.5, 29	**29**	56.5, 87	**87**	56.5, 1484	**1484**
57.5, 30	**30**	57.5, 71	**71**	57.5, 572	**572**
58.5, 21	**21**	58.5, 74	**74**	58.5, 1549	**1549**
59.5, 25	**25**	59.5, 67	**67**	59.5, 730	**730**
60.5, 27	**24**	60.5, 73	**73**	60.5, 1114	**114**
61.5, 22	**22**	61.5, 83	**83**	61.5, 751	**751**
62.5, 16	**16**	62.5, 66	**66**	62.5, 1182	**1182**
63.5, 38	**38**	63.5, 66	**66**	63.5, 690	**690**
64.5, 42	**42**	64.5, 57	**57**	64.5, 1223	**1223**
65.5, 40	**40**	65.5, 68	**68**	65.5, 590	**590**
66.5, 24	**24**	66.5, 67	**67**	66.5, 1033	**1033**
67.5, 27	**27**	67.5, 60	**60**	67.5, 612	**612**
68.5, 26	**26**	68.5, 68	**68**	68.5, 1080	**1080**
69.5, 43	**43**	69.5, 64	**64**	69.5, 642	**642**
70.5, 32	**32**	70.5, 67	**67**	70.5, 922	**922**
71.5, 28	**28**	71.5, 60	**60**	71.5, 538	**538**
72.5, 28	**28**	72.5, 69	**69**	72.5, 683	**683**
73.5, 31	**31**	73.5, 80	**80**	73.5, 549	**549**
74.5, 37	**37**	74.5, 88	**88**	74.5, 619	**619**
75.5, 26	**26**	75.5, 74	**74**	75.5, 513	**513**
76.5, 39	**39**	76.5, 84	**84**	76.5, 724	**724**
77.5, 29	**29**	77.5, 71	**71**	77.5, 623	**623**
78.5, 24	**24**	78.5, 55	**55**	78.5, 610	**610**

المنارة للاستشارات
www.manaraa.com

79.5, 30	**30**	79.5, 75	**75**	79.5, 607	**607**
80.5, 43	**43**	80.5, 65	**65**	80.5, 902	**902**
81.5, 24	**24**	81.5, 76	**76**	81.5, 538	**538**
82.5, 18	**18**	82.5, 63	**63**	82.5, 531	**531**
83.5, 37	**37**	83.5, 64	**64**	83.5, 459	**459**
84.5, 34	**34**	84.5, 63	**63**	84.5, 550	**550**
85.5, 28	**28**	85.5, 68	**68**	85.5, 435	**435**
86.5, 29	**29**	86.5, 72	**72**	86.5, 543	**543**
87.5, 51	**51**	87.5, 78	**78**	87.5, 508	**508**
88.5, 24	**24**	88.5, 72	**72**	88.5, 702	**702**
89.5, 74	**74**	89.5, 62	**62**	89.5, 511	**511**
90.5, 50	**50**	90.5, 65	**65**	90.5, 590	**590**
91.5, 26	**26**	91.5, 67	**67**	91.5, 577	**577**
92.5, 29	**29**	92.5, 72	**72**	92.5, 446	**446**
93.5, 26	**26**	93.5, 68	**68**	93.5, 475	**475**
94.5, 22	**22**	94.5, 63	**63**	94.5, 485	**485**
95.5, 28	**28**	95.5, 80	**80**	95.5, 585	**585**
96.5, 20	**20**	96.5, 78	**78**	96.5, 562	**562**
97.5, 27	**27**	97.5, 78	**78**	97.5, 542	**542**
98.5, 23	**23**	98.5, 87	**87**	98.5, 549	**549**
99.5, 35	**35**	99.5, 76	**76**	99.5, 487	**487**
100.5, 25	**25**	100.5, 80	**80**	100.5, 528	**528**
101.5, 34	**34**	101.5, 75	**75**	101.5, 535	**535**
102.5, 37	**37**	102.5, 74	**74**	102.5, 545	**545**
103.5, 33	**33**	103.5, 76	**76**	103.5, 532	**532**
104.5, 30	**30**	104.5, 66	**66**	104.5, 494	**494**
105.5, 42	**42**	105.5, 65	**65**	105.5, 570	**570**
106.5, 39	**39**	106.5, 63	**63**	106.5, 583	**582**
107.5, 42	**42**	107.5, 74	**74**	107.5, 693	**693**
108.5, 29	**29**	108.5, 67	**67**	108.5, 631	**631**
109.5, 33	**33**	109.5, 51	**51**	109.5, 576	**576**
110.5, 26	**26**	110.5, 56	**56**	110.5, 634	**634**
111.5, 33	**33**	111.5, 75	**75**	111.5, 624	**624**
112.5, 24	**24**	112.5, 53	**53**	112.5, 594	**594**
113.5, 26	**26**	113.5, 69	**69**	113.5, 653	**653**
114.5, 38	**38**	114.5, 70	**70**	114.5, 630	**630**
115.5, 32	**32**	115.5, 86	**86**	115.5, 740	**740**
116.5, 31	**31**	116.5, 61	**61**	116.5, 558	**558**
117.5, 30	**30**	117.5, 54	**54**	117.5, 659	**659**
118.5, 24	**24**	118.5, 81	**81**	118.5, 636	**636**
119.5, 26	**26**	119.5, 50	**50**	119.5, 673	**673**
120.5, 30	**30**	120.5, 52	**52**	120.5, 668	**668**
121.5, 28	**28**	121.5, 57	**57**	121.5, 684	**684**
122.5, 37	**37**	122.5, 63	**63**	122.5, 641	**641**
123.5, 35	**35**	123.5, 48	**48**	123.5, 667	**667**

المنارة للاستشارات
www.manaraa.com

124.5, 39	**39**	124.5, 45	**45**	124.5, 618	**618**
125.5, 43	**43**	125.5, 63	**63**	125.5, 619	**619**
126.5, 27	**27**	126.5, 46	**46**	126.5, 621	**621**
127.5, 31	**31**	127.5, 47	**47**	127.5, 611	**611**
128.5, 30	**30**	128.5, 50	**50**	128.5, 597	**597**
129.5, 33	**33**	129.5, 38	**38**	129.5, 658	**658**
130.5, 36	**36**	130.5, 41	**41**	130.5, 658	**658**
131.5, 44	**44**	131.5, 44	**44**	131.5, 750	**750**
132.5, 38	**38**	132.5, 46	**46**	132.5, 678	**678**
133.5, 36	**36**	133.5, 56	**56**	133.5, 716	**716**
134.5, 35	**35**	134.5, 58	**58**	134.5, 726	**726**
135.5, 36	**36**	135.5, 60	**60**	135.5, 704	**704**
136.5, 20	**20**	136.5, 57	**57**	136.5, 682	**682**
137.5, 38	**38**	137.5, 88	**88**	137.5, 675	**675**
138.5, 35	**35**	138.5, 72	**72**	138.5, 695	**695**
139.5, 44	**44**	139.5, 87	**87**	139.5, 735	**75**
140.5, 24	**24**	140.5, 70	**70**	140.5, 613	**613**
141.5, 29	**29**	141.5, 90	**90**	141.5, 783	**783**
142.5, 24	**24**	142.5, 91	**91**	142.5, 780	**780**
143.5, 40	**40**	143.5, 85	**85**	143.5, 780	**780**
144.5, 31	**31**	144.5, 79	**79**	144.5, 776	**776**
145.5, 39	**39**	145.5, 88	**88**	145.5, 781	**781**
146.5, 28	**28**	146.5, 97	**97**	146.5, 769	**769**
147.5, 35	**35**	147.5, 111	**111**	147.5, 743	**743**
148.5, 46	**46**	148.5, 88	**88**	148.5, 726	**726**
149.5, 49	**49**	149.5, 147	**147**	149.5, 933	**933**
150.5, 94	**94**	150.5, 117	**117**	150.5, 802	**802**
151.5, 26	**26**	151.5, 157	**157**	151.5, 864	**864**
152.5, 43	**43**	152.5, 181	**181**	152.5, 891	**891**
153.5, 36	**36**	153.5, 183	**183**	153.5, 985	**985**
154.5, 134	**134**	154.5, 163	**163**	154.5, 2126	**2126**
155.5, 143	**143**	155.5, 171	**171**	155.5, 1108	**1108**
156.5, 390	**390**	156.5, 189	**189**	156.5, 1472	**1472**
157.5, 77	**77**	157.5, 216	**216**	157.5, 2428	**2428**
158.5, 145	**145**	158.5, 233	**233**	158.5, 1032	**1032**
159.5, 51	**51**	159.5, 215	**215**	159.5, 2008	**2008**
160.5, 107	**107**	160.5, 205	**205**	160.5, 1068	**1068**
161.5, 49	**49**	161.5, 291	**291**	161.5, 891	**891**
162.5, 41	**41**	162.5, 252	**252**	162.5, 838	**838**
163.5, 69	**69**	163.5, 193	**193**	163.5, 725	**725**
164.5, 49	**49**	164.5, 234	**234**	164.5, 755	**755**
165.5, 41	**41**	165.5, 217	**217**	165.5, 701	**701**
166.5, 34	**34**	166.5, 234	**234**	166.5, 674	**674**
167.5, 37	**37**	167.5, 236	**236**	167.5, 726	**726**
168.5, 69	**69**	168.5, 297	**297**	168.5, 659	**659**

المنارة للاستشارات
www.manaraa.com

169.5, 39	**39**	169.5, 268	**268**	169.5, 709	**709**
170.5, 40	**40**	170.5, 263	**263**	170.5, 607	**607**
171.5, 35	**35**	171.5, 275	**275**	171.5, 632	**632**
172.5, 40	**40**	172.5, 415	**415**	172.5, 590	**590**
173.5, 38	**38**	173.5, 238	**238**	173.5, 584	**584**
174.5, 32	**32**	174.5, 499	**499**	174.5, 554	**554**
175.5, 36	**36**	175.5, 285	**285**	175.5, 564	**564**
176.5, 39	**39**	176.5, 474	**474**	176.5, 499	**564**
177.5, 40	**40**	177.5, 245	**245**	177.5, 519	**499**
178.5, 40	**40**	178.5, 284	**284**	178.5, 543	**519**
179.5, 40	**40**	179.5, 318	**318**	179.5, 561	**561**
180.5, 63	**63**	180.5, 315	**315**	180.5, 524	**524**
181.5, 49	**49**	181.5, 263	**263**	181.5, 483	**483**
182.5, 91	**91**	182.5, 330	**330**	182.5, 435	**435**
183.5, 43	**43**	183.5, 185	**185**	183.5, 444	**444**
184.5, 49	**49**	184.5, 234	**234**	184.5, 437	**437**
185.5, 46	**46**	185.5, 213	**213**	185.5, 406	**406**
186.5, 47	**47**	186.5, 210	**210**	186.5, 420	**420**
187.5, 53	**53**	187.5, 213	**213**	187.5, 428	**428**
188.5, 30	**30**	188.5, 213	**213**	188.5, 501	**501**
189.5, 64	**64**	189.5, 218	**218**	189.5, 447	**447**
190.5, 36	**36**	190.5, 230	**230**	190.5, 439	**439**
191.5, 49	**49**	191.5, 163	**163**	191.5, 417	**417**
192.5, 27	**27**	192.5, 176	**176**	192.5, 442	**442**
193.5, 29	**29**	193.5, 213	**213**	193.5, 414	**414**
194.5, 43	**43**	194.5, 245	**245**	194.5, 388	**388**
195.5, 43	**43**	195.5, 210	**210**	195.5, 384	**384**
196.5, 30	**30**	196.5, 159	**159**	196.5, 418	**418**
197.5, 33	**33**	197.5, 147	**147**	197.5, 370	**370**
198.5, 27	**27**	198.5, 91	**91**	198.5, 585	**585**
199.5, 38	**38**	199.5, 122	**122**	199.5, 400	**400**
200.5, 45	**45**	200.5, 60	**60**	200.5, 428	**428**
201.5, 29	**29**	201.5, 52	**52**	201.5, 450	**450**
202.5, 43	**43**	202.5, 63	**63**	202.5, 509	**509**
203.5, 49	**49**	203.5, 45	**45**	203.5, 517	**517**
204.5, 35	**35**	204.5, 50	**50**	204.5, 631	**631**
205.5, 38	**38**	205.5, 52	**52**	205.5, 567	**267**
206.5, 45	**45**	206.5, 51	**51**	206.5, 698	**698**
207.5, 44	**44**	207.5, 55	**55**	207.5, 649	**649**
208.5, 51	**51**	208.5, 34	**34**	208.5, 870	**870**
209.5, 52	**52**	209.5, 59	**59**	209.5, 593	**593**
210.5, 73	**73**	210.5, 47	**47**	210.5, 1101	**1101**
211.5, 43	**43**	211.5, 49	**49**	211.5, 617	**617**
212.5, 46	**46**	212.5, 54	**54**	212.5, 947	**647**
213.5, 58	**58**	213.5, 44	**44**	213.5, 666	**666**

المنارة للاستشارات
www.manaraa.com

214.5, 54	**54**	214.5, 47	**47**	214.5, 1222	**1222**
215.5, 53	**53**	215.5, 62	**62**	215.5, 696	**696**
216.5, 67	**67**	216.5, 89	**89**	216.5, 1008	**1008**
217.5, 79	**79**	217.5, 82	**82**	217.5, 608	**608**
218.5, 56	**56**	218.5, 63	**63**	218.5, 807	**807**
219.5, 65	**65**	219.5, 80	**80**	219.5, 424	**424**
220.5, 82	**82**	220.5, 178	**178**	220.5, 621	**621**
221.5, 88	**88**	221.5, 101	**101**	221.5, 401	**401**
222.5, 60	**60**	222.5, 80	**80**	222.5, 576	**576**
223.5, 77	**77**	223.5, 109	**109**	223.5, 344	**344**
224.5, 95	**95**	224.5, 79	**79**	224.5, 359	**359**
225.5, 81	**81**	225.5, 79	**79**	225.5, 291	**291**
226.5, 99	**99**	226.5, 107	**107**	226.5, 360	**360**
227.5, 101	**101**	227.5, 108	**108**	227.5, 272	**272**
228.5, 117	**117**	228.5, 101	**101**	228.5, 329	**329**
229.5, 97	**97**	229.5, 97	**97**	229.5, 228	**228**
230.5, 118	**118**	230.5, 113	**113**	230.5, 234	**234**
231.5, 140	**140**	231.5, 122	**122**	231.5, 288	**288**
232.5, 127	**127**	232.5, 101	**101**	232.5, 246	**246**
233.5, 159	**159**	233.5, 127	**127**	233.5, 208	**208**
234.5, 169	**169**	234.5, 137	**137**	234.5, 189	**189**
235.5, 172	**172**	235.5, 151	**151**	235.5, 172	**172**
236.5, 181	**181**	236.5, 148	**148**	236.5, 285	**285**
237.5, 221	**221**	237.5, 174	**174**	237.5, 135	**135**
238.5, 226	**226**	238.5, 217	**217**	238.5, 313	**313**
239.5, 218	**218**	239.5, 176	**176**	239.5, 153	**153**
240.5, 272	**272**	240.5, 243	**243**	240.5, 157	**157**
241.5, 293	**293**	241.5, 262	**262**	241.5, 126	**126**
242.5, 282	**282**	242.5, 266	**266**	242.5, 135	**135**
243.5, 299	**299**	243.5, 236	**236**	243.5, 129	**129**
244.5, 378	**378**	244.5, 275	**275**	244.5, 146	**146**
245.5, 377	**377**	245.5, 302	**302**	245.5, 125	**125**
246.5, 401	**401**	246.5, 278	**278**	246.5, 157	**157**
247.5, 552	**552**	247.5, 301	**301**	247.5, 118	**118**
248.5, 732	**732**	248.5, 388	**388**	248.5, 158	**158**
249.5, 475	**475**	249.5, 399	**399**	249.5, 160	**160**
250.5, 727	**727**	250.5, 499	**499**	250.5, 158	**158**
251.5, 931	**931**	251.5, 487	**487**	251.5, 156	**156**
252.5, 642	**642**	252.5, 533	**533**	252.5, 182	**182**
253.5, 1061	**1061**	253.5, 1084	**1084**	253.5, 136	**136**
254.5, 92726	**92726**	254.5, 96698	**96698**	254.5, 406	**406**
255.5, 26252	**26252**	255.5, 9940	**9940**	255.5, 849	**849**

المنارة للاستشارات

www.manaraa.com

Appendix B

The normalized difference between selected four sequential presentational video frames (i.e. Frame 1, Frame 2, Frame 3, and Frame 4) in the proposed experiment.

Difference (Frame1, Frame 2)	*Difference (Frame2, Frame 3)*	*Difference (Frame3, Frame 4)*
166	211	45
9	12	3
7	7	0
19	23	4
20	24	4
11	18	7
7	12	5
17	18	1
18	20	2
12	20	8
7	15	8
10	20	10
16	21	5
27	29	2
22	30	8
11	18	7
14	23	9
23	25	2
13	24	11
15	22	7
29	36	7
14	23	9
16	31	15
19	31	12
36	44	8
18	32	14
24	28	4
29	41	12
26	33	7
29	40	11
23	32	9
36	23	13
46	38	8
52	3	55

المنارة للاستشارات
www.manaraa.com

48	50	98
42	212	254
63	115	178
60	281	341
82	228	310
50	346	396
46	320	366
44	386	430
50	322	372
62	239	301
41	581	622
49	386	435
46	875	921
33	349	382
26	1199	1225
78	462	540
62	2107	2169
59	497	556
51	1125	1176
63	123	186
43	1160	1203
41	412	453
58	1397	1455
41	501	542
53	1475	1528
42	663	705
49	41	90
61	668	729
50	1116	1166
28	624	652
15	1166	1181
28	522	550
43	966	1009
33	552	585
42	1012	1054
21	578	599
35	855	890
32	478	510
41	614	655
49	469	518
51	531	582
48	439	487
45	640	685
42	552	594
31	555	586

المنارة للاستشارات

www.manaraa.com

45	532	577
22	837	859
52	462	514
45	468	513
27	395	422
29	487	516
40	367	407
43	471	514
27	430	457
48	630	678
12	449	437
15	525	540
41	510	551
43	374	417
42	407	449
41	422	463
52	505	557
58	484	542
51	464	515
64	462	526
41	411	452
55	448	503
41	460	501
37	471	508
43	456	499
36	428	464
23	505	528
24	519	543
32	619	651
38	564	602
18	525	543
30	578	608
42	549	591
29	541	570
43	584	627
32	560	592
54	654	708
30	497	527
24	605	629
57	555	612
24	623	647
22	616	638
29	627	656
26	578	604
13	619	632

المنارة للاستشارات
www.manaraa.com

6	573	579
20	556	576
19	575	594
16	564	580
20	547	567
5	620	625
5	617	622
0	706	706
8	632	640
20	660	680
23	668	691
24	644	668
37	625	662
50	587	637
37	623	660
43	12	31
46	543	589
61	693	754
67	689	756
45	695	740
48	697	745
49	693	742
69	672	741
76	632	708
42	638	680
98	786	884
23	685	708
131	707	838
138	710	848
147	802	949
29	1963	1992
28	937	965
201	1283	1082
139	2212	2351
88	799	887
164	1793	1957
98	863	961
242	600	842
211	586	797
124	532	656
185	521	706
176	484	660
200	440	640
199	490	689
228	362	590

المنارة للاستشارات
www.manaraa.com

229	441	670
223	344	567
240	357	597
375	175	550
200	346	546
467	55	522
249	279	528
435	90	525
205	254	459
244	235	479
278	243	521
252	209	461
214	220	434
239	105	344
142	259	401
185	203	388
167	193	360
163	210	373
160	215	375
183	288	471
154	229	383
194	209	403
114	254	368
149	266	415
184	201	385
202	143	345
167	174	341
129	259	388
114	223	337
64	494	558
84	278	362
15	368	383
23	398	421
20	446	466
4	472	468
15	581	596
14	215	229
6	647	653
11	594	605
17	836	819
7	534	541
26	1054	1028
6	568	574
8	593	601
14	622	608

المنارة للاستشارات

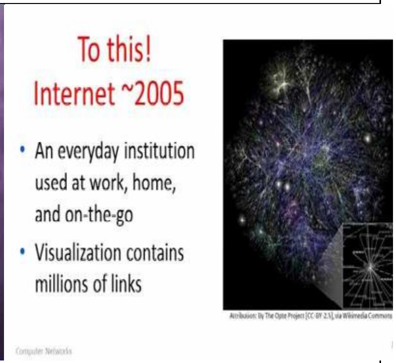
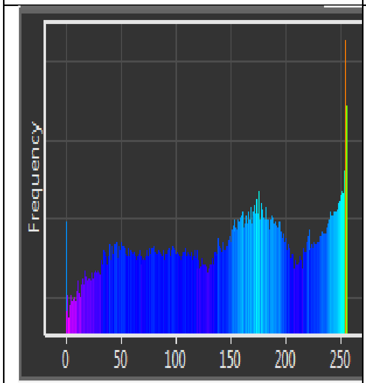
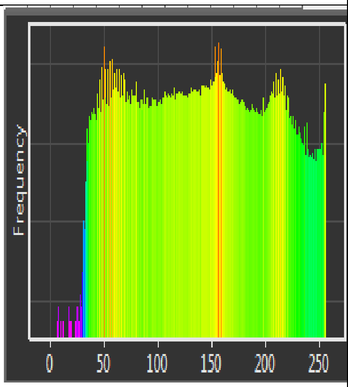
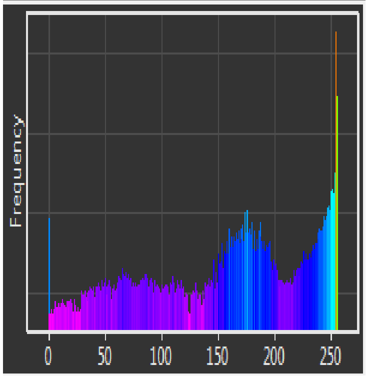
www.manaraa.com

7	1175	1168
9	634	643
22	919	941
3	526	529
7	744	751
15	344	359
96	443	539
13	300	313
20	496	516
32	235	267
16	280	264
2	212	210
8	253	261
7	164	171
16	228	212
0	131	131
5	121	116
18	166	148
26	145	119
32	81	49
32	52	20
21	21	0
33	137	104
47	39	86
9	96	87
42	23	65
29	86	115
31	136	167
16	131	147
63	107	170
103	129	232
75	177	252
123	121	244
251	183	434
344	230	574
76	239	315
228	341	569
444	331	775
109	351	460
23	948	925
3972	96292	92320
16312	9091	25403

المنارة للاستشارات
www.manaraa.com

Appendix C

The extracted key frames original images and its color histograms.

Frame 3	Frame 5	Frame 10
		
Color Histogram (Frame 3)	**Color Histogram (Frame 5)**	**Color Histogram (Frame 10)**
		

www.manaraa.com

Appendix D

The color histogram of each pixel in spoken text lecture videos for five key frame

Frame 1	Frame 2	Frame 3	Frame 4	Frame 5	Diff (F1 and F2)	Diff (F2 and F3)	Diff (F3 and F4)	Diff (F4 and F5)
4399	5066	16815	13164	8777	667	11749	3651	4387
704	820	587	1047	826	116	233	460	221
231	386	17070	14418	1000	155	16684	2652	13418
351	260	219	181	1007	91	41	38	826
458	428	1410	822	1011	30	982	588	189
322	278	173	213	1013	44	105	40	800
16	1878	655	482	1024	1862	1223	173	542
651	677	612	896	1032	26	65	284	136
490	688	2253	1531	1033	198	1565	722	498
595	677	119	93	1037	82	558	26	944
2030	1035	439	537	1038	995	596	98	501
1820	1313	165	119	1038	507	1148	46	919
583	785	207	126	1041	202	578	81	915
911	719	895	757	1049	192	176	138	292
1136	1353	129	149	1051	217	1224	20	902
576	646	1094	947	1067	70	448	147	120
401	389	563	1865	1070	12	174	1302	795
1413	1062	383	267	1073	351	679	116	806
602	1098	1106	513	1075	496	8	593	562
1589	830	270	163	1084	759	560	107	921
554	812	736	149	1088	258	76	587	939
555	847	292	168	1105	292	555	124	937
1223	965	327	144	1129	258	638	183	985
1321	1447	333	195	1134	126	1114	138	939
454	429	84	78	1139	25	345	6	1061
558	546	669	149	1141	12	123	520	992
760	746	593	383	1142	14	153	210	759
573	962	101	77	1163	389	861	24	1086
514	455	2143	1412	1169	59	1688	731	243
1461	96645	116	125	1170	95184	96529	9	1045
588	641	240	396	1183	53	401	156	787
833	945	99	97	1183	112	846	2	1086
1717	1045	114	75	1184	672	931	39	1109
537	801	449	276	1187	264	352	173	911
1511	1761	173	207	1191	250	1588	34	984
641	953	118	87	1202	312	835	31	1115
648	707	616	911	1218	59	91	295	307
569	618	325	502	1251	49	293	177	749
292	250	262	329	1259	42	12	67	930

المنارة للاستشارات
www.manaraa.com

260	397	698	505	1267	137	301	193	762
1075	622	202	133	1291	453	420	69	1158
371	287	189	189	1302	84	98	0	1113
219	519	721	1311	1312	300	202	590	1
1016	867	79	240	1314	149	788	161	1074
1508	945	78	97	1323	563	867	19	1226
1653	664	65	85	1328	989	599	20	1243
564	1127	86	67	1352	563	1041	19	1285
385	391	113	87	1496	6	278	26	1409
257	259	1561	2561	1636	2	1302	1000	925
388	280	909	861	1726	108	629	48	865
421	470	5585	4566	1728	49	5115	1019	2838
408	431	207	81	1742	23	224	126	1661
426	400	69	100	1784	26	331	31	1684
654	1554	709	429	1840	900	845	280	1411
799	1170	61	93	1873	371	1109	32	1780
1046	1103	64	69	2486	57	1039	5	2417
1486	756	81	63	2549	730	675	18	2486
649	976	69	69	2606	327	907	0	2537
400	405	325	201	2654	5	80	124	2453
1798	801	64	69	3630	997	737	5	3561
479	823	91	71	3751	344	732	20	3680
835	918	136	62	4146	83	782	74	4084
1525	774	65	74	8268	751	709	9	8194
369	1582	71	91	179	1213	1511	20	88
345	1732	89	113	183	1387	1643	24	70
788	759	86	104	190	29	673	18	86
1217	824	95	108	197	393	729	13	89
294	278	758	1646	207	16	480	888	1439
514	542	89	125	208	28	453	36	83
306	547	93	93	210	241	454	0	117
1207	760	100	105	214	447	660	5	109
363	474	148	89	217	111	326	59	128
1168	796	81	71	217	372	715	10	146
621	454	151	140	218	167	303	11	78
291	326	760	549	223	35	434	211	326
399	535	124	186	225	136	411	62	39
344	486	94	87	229	142	392	7	142
482	527	93	99	232	45	434	6	133
555	407	116	166	236	148	291	50	70
381	473	152	149	238	92	321	3	89
344	404	147	95	239	60	257	52	144
546	404	92	132	243	142	312	40	111
498	523	116	237	243	25	407	121	6
306	330	109	133	243	24	221	24	110
537	511	104	118	245	26	407	14	127
604	643	137	322	245	39	506	185	77

www.manaraa.com

362	315	123	127	246	47	192	4	119
375	316	134	142	249	59	182	8	107
370	452	100	132	250	82	352	32	118
829	282	127	138	251	547	155	11	113
364	350	135	201	256	14	215	66	55
507	518	713	725	257	11	195	12	468
563	541	130	136	258	22	411	6	122
563	622	107	260	258	59	515	153	2
656	1374	135	124	259	718	1239	11	135
457	451	81	85	260	6	370	4	175
451	465	83	81	260	14	382	2	179
457	456	91	67	261	1	365	24	194
391	419	80	245	263	28	339	165	18
2013	2259	114	156	264	246	2145	42	108
2595	2713	126	146	266	118	2587	20	120
358	332	150	173	266	26	182	23	93
537	548	118	149	268	11	430	31	119
356	378	80	73	275	22	298	7	202
333	354	131	214	275	21	223	83	61
446	443	83	70	279	3	360	13	209
774	775	119	131	279	1	656	12	148
2305	2599	126	195	281	294	2473	69	86
671	681	105	135	286	10	576	30	151
1742	496	75	69	290	1246	421	6	221
829	798	106	208	290	31	692	102	82
3116	3511	152	128	291	395	3359	24	163
63	275	164	125	294	212	111	39	169
3961	5197	122	112	295	1236	5075	10	183
371	325	124	722	295	46	201	598	427
528	579	123	168	296	51	456	45	128
300	361	652	540	301	61	291	112	239
1172	1234	105	113	302	62	1129	8	189
439	410	93	100	307	29	317	7	207
1513	1883	101	168	307	370	1782	67	139
480	582	128	324	310	102	454	196	14
1852	2021	132	135	316	169	1889	3	181
1456	1496	111	118	319	40	1385	7	201
917	1005	91	272	327	88	914	181	55
699	1638	175	143	331	939	1463	32	188
3388	2706	97	175	332	682	2609	78	157
1396	1570	126	99	333	174	1444	27	234
460	429	270	97	336	31	159	173	239
1880	2073	142	143	337	193	1931	1	194
2922	2424	115	112	341	498	2309	3	229
728	804	372	235	346	76	432	137	111
871	909	328	175	347	38	581	153	172
733	1594	157	151	347	861	1437	6	196

المنارة للاستشارات
www.manaraa.com

526	539	155	158	350	13	384	3	192
1056	1045	114	140	351	11	931	26	211
4654	4999	107	130	359	345	4892	23	229
459	973	145	126	361	514	828	19	235
693	652	126	135	364	41	526	9	229
1458	671	156	135	368	787	515	21	233
396	406	169	66	374	10	237	103	308
1175	902	133	107	379	273	769	26	272
311	500	782	515	379	189	282	267	136
7951	7353	117	107	382	598	7236	10	275
657	1466	708	1258	384	809	758	550	874
810	825	197	238	386	15	628	41	148
261	292	272	269	394	31	20	3	125
245	270	253	289	396	25	17	36	107
406	245	693	526	400	161	448	167	126
1002	1077	187	152	401	75	890	35	249
1075	831	90	108	401	244	741	18	293
3741	3850	113	127	403	109	3737	14	276
569	564	220	207	406	5	344	13	199
778	803	199	421	408	25	604	222	13
306	288	674	858	413	18	386	184	445
858	1014	160	161	417	156	854	1	256
639	702	363	371	422	63	339	8	51
595	321	771	483	423	274	450	288	60
1728	1633	114	202	427	95	1519	88	225
547	597	215	232	428	50	382	17	196
354	279	144	143	432	75	135	1	289
1384	1288	106	134	438	96	1182	28	304
469	464	1245	1168	438	5	781	77	730
705	766	289	580	447	61	477	291	133
683	663	883	483	448	20	220	400	35
625	715	97	118	450	90	618	21	332
281	310	225	203	453	29	85	22	250
262	282	277	283	461	20	5	6	178
300	275	246	278	464	25	29	32	186
680	799	483	353	467	119	316	130	114
256	272	284	334	469	16	12	50	135
836	1480	111	108	473	644	1369	3	365
277	253	212	250	474	24	41	38	224
1223	1516	120	101	489	293	1396	19	388
332	261	236	227	492	71	25	9	265
3279	3217	105	113	501	62	3112	8	388
18415	16570	208	234	502	1845	16362	26	268
616	690	72	91	507	74	618	19	416
13618	14203	118	107	507	585	14085	11	400
922	581	341	548	519	341	240	207	29
5612	4822	145	120	523	790	4677	25	403

المنارة للاستشارات
www.manaraa.com

259	288	223	238	525	29	65	15	287
683	795	138	116	534	112	657	22	418
1471	1754	150	209	537	283	1604	59	328
328	287	195	193	547	41	92	2	354
1369	1609	120	155	556	240	1489	35	401
1491	2262	345	389	562	771	1917	44	173
3201	3999	122	505	563	798	3877	383	58
262	258	302	358	587	4	44	56	229
1359	1450	137	355	588	91	1313	218	233
1016	1133	144	177	590	117	989	33	413
79	793	170	148	590	714	623	22	442
352	417	25524	14182	591	65	25107	11342	13591
236	335	385	490	595	99	50	105	105
358	279	422	510	613	79	143	88	103
259	226	344	471	616	33	118	127	145
341	550	1193	1639	624	209	643	446	1015
392	342	13816	11634	626	50	13474	2182	11008
231	256	485	615	628	25	229	130	13
4822	4552	123	133	637	270	4429	10	504
361	275	167	173	643	86	108	6	470
382	521	2748	1639	643	139	2227	1109	996
53608	52292	126	111	646	1316	52166	15	535
207	243	3046	4292	647	36	2803	1246	3645
238	242	337	345	649	4	95	8	304
400	410	226	71	651	10	184	155	580
278	404	720	500	653	126	316	220	153
276	453	1190	495	657	177	737	695	162
213	237	482	1585	663	24	245	1103	922
402	244	442	527	684	158	198	85	157
223	269	574	1563	685	46	305	989	878
280	249	244	266	694	31	5	22	428
1183	743	69	75	708	440	674	6	633
287	301	1082	479	710	14	781	603	231
1034	1167	171	159	718	133	996	12	559
1251	1345	139	140	781	94	1206	1	641
413	430	1105	1036	848	17	675	69	188
385	424	1647	3211	875	39	1223	1564	2336
214	259	1638	1078	879	45	1379	560	199
1975	1072	112	89	884	903	960	23	795
702	1131	838	108	892	429	293	730	784
733	236	190	151	894	497	46	39	743
411	523	66581	82272	899	112	66058	15691	81373
378	438	1818	1168	900	60	1380	650	268
382	481	1317	1318	901	99	836	1	417
2320	1133	544	472	901	1187	589	72	429
2173	2361	122	112	902	188	2239	10	790
452	505	2946	1939	904	53	2441	1007	1035

المنارة للاستشارات

www.manaraa.com

427	414	3301	1211	906	13	2887	2090	305
413	432	1127	712	910	19	695	415	198
737	786	143	90	911	49	643	53	821
717	938	140	106	915	221	798	34	809
430	431	2790	1532	916	1	2359	1258	616
413	389	1513	1372	926	24	1124	141	446
400	450	77	276	932	50	373	199	656
418	418	3665	3159	936	0	3247	506	2223
452	460	1830	1060	946	8	1370	770	114
1353	1384	136	91	947	31	1248	45	856
933	735	553	82	948	198	182	471	866
382	441	1003	721	950	59	562	282	229
197	263	557	959	951	66	294	402	8
386	302	1972	1769	952	84	1670	203	817
408	407	124	77	959	1	283	47	882
611	562	2807	1490	960	49	2245	1317	530
421	456	985	1414	965	35	529	429	449
406	460	1577	1124	965	54	1117	453	159
1762	817	146	320	972	945	671	174	652
435	474	1043	919	977	39	569	124	58
625	703	770	332	977	78	67	438	645
858	1202	92	90	982	344	1110	2	892
442	531	3330	3220	983	89	2799	110	2237
454	355	163	716	986	99	192	553	270
483	478	72	91	988	5	406	19	897
379	437	86	93	11913	58	351	7	11820
666	650	834	1065	15607	16	184	231	14542
1573	708	73	106	41498	865	635	33	41392
517	670	70	95	78816	153	600	25	78721
305630	**404432**	**255356**	**245308**	**348070**	**153194**	**525800**	**79834**	**412108**
1193.867188	**1579.8125**	**997.484375**	**958.234375**	**1359.648438**	**598.4140625**	**2053.90625**	**311.8515625**	**1609.796875**

المنارة للاستشارات

www.manaraa.com

Appendix D

The iterative thresholding technique test data after executing the target threshold in MATLAB.

N	Pixel Mean	Pixel Mean Above	Pixel Mean Below
0	20	3751	20
1	22	7627	22
2	25	11596	25
3	25	15825	25
4	25	20231	25
5	27	24977	27
6	28	29895	28
7	42	35339	42
8	53	40874	53
9	66	46708	66
10	88	53068	88
11	150	59802	150
12	232	66844	232
13	336	74452	336
14	485	81949	485
15	666	89623	666
16	883	97422	883
17	1163	105130	1163
18	1458	113114	1458
19	1817	121135	1817
20	2221	129055	2221
21	2698	137124	2698
22	3267	145144	3267
23	3895	153121	3895
24	4563	161229	4563
25	5293	169256	5293
26	6058	177530	6058
27	6831	185891	6831
28	7642	194559	7642
29	8423	203422	8423
30	9194	212917	9194
31	9988	222748	9988
32	10796	233053	10796
33	11593	244063	11593
34	12402	255715	12402
35	13266	267858	13266

Loop (i)	Threshold (T)
1	152
2	143
3	137
4	133
5	130
6	128
7	127
8	127
Mean Below Threshold	85.80894379
Mean Above Threshold	167.3627977
Gray Level Threshold	0.496062992

المنارة للاستشارات
www.manaraa.com

36	14129	281183	14129
37	14988	295625	14988
38	15892	311404	15892
39	16781	328391	16781
40	17728	346964	17728
41	18691	367054	18691
42	19658	388451	19658
43	20692	411388	20692
44	21700	434519	21700
45	22773	457027	22773
46	23840	478938	23840
47	24970	499723	24970
48	26161	519553	26161
49	27365	538412	27365
50	28617	555348	28617
51	29876	570480	29876
52	31192	583619	31192
53	32529	595399	32529
54	33828	605992	33828
55	35148	615890	35148
56	36503	625234	36503
57	37816	634396	37816
58	39108	643414	39108
59	40446	652226	40446
60	41732	660728	41732
61	42992	668962	42992
62	44304	676919	44304
63	45653	684087	45653
64	46998	690899	46998
65	48360	697148	48360
66	49696	702816	49696
67	51021	708245	51021
68	52412	713466	52412
69	53925	718720	53925
70	55297	723904	55297
71	56772	729211	56772
72	58234	734307	58234
73	59682	739290	59682
74	61169	744161	61169
75	62676	748581	62676
76	64189	752510	64189
77	65711	755855	65711
78	67170	758802	67170
79	68724	761162	68724

المنارة للاستشارات
www.manaraa.com

80	70297	763014	70297
81	71889	764352	71889
82	73463	765437	73463
83	75107	766332	75107
84	76650	767115	76650
85	78210	767768	78210
86	79822	768355	79822
87	81467	768900	81467
88	83032	769424	83032
89	84746	769897	84746
90	86376	770314	86376
91	88043	770674	88043
92	89719	771030	89719
93	91486	771392	91486
94	93139	771735	93139
95	94818	772079	94818
96	96484	772443	96484
97	98255	772790	98255
98	100038	773127	100038
99	101724	773462	101724
100	103499	773769	103499
101	105287	774060	105287
102	107162	774345	107162
103	109128	774658	109128
104	111063	774928	111063
105	113148	775184	113148
106	115335	775400	115335
107	117605	775585	117605
108	119911	775777	119911
109	122260	775929	122260
110	124727	776062	124727
111	127329	776198	127329
112	129897	776315	129897
113	132527	776443	132527
114	135300	776561	135300
115	138201	776671	138201
116	141153	776784	141153
117	144264	776892	144264
118	147376	777031	147376
119	150603	777170	150603
120	153787	777297	153787
121	157152	777434	157152
122	160517	777541	160517
123	164122	777625	164122

المنارة للاستشارات
www.manaraa.com

124	167725	777675	167725
125	171365	777727	171365
126	175116	777781	175116
127	178992	777854	
128	182961	777934	
129	187190	778285	
130	191596		
131	196342		
132	201260		
133	206704		
134	212239		
135	218073		
136	224433		
137	231167		
138	238209		
139	245817		
140	253314		
141	260988		
142	268787		
143	276495		
144	284479		
145	292500		
146	300420		
147	308489		
148	316509		
149	324486		
150	332594		
151	340621		
152	348895		
153	357256		
154	365924		
155	374787		
156	384282		
157	394113		
158	404418		
159	415428		
160	427080		
161	439223		
162	452548		
163	466990		
164	482769		
165	499756		
166	518329		
167	538419		

المنارة للاستشارات
www.manaraa.com

168	559816		
169	582753		
170	605884		
171	628392		
172	650303		
173	671088		
174	690918		
175	709777		
176	726713		
177	741845		
178	754984		
179	766764		
180	777357		
181	787255		
182	796599		
183	805761		
184	814779		
185	823591		
186	832093		
187	840327		
188	848284		
189	855452		
190	862264		
191	868513		
192	874181		
193	879610		
194	884831		
195	890085		
196	895269		
197	900576		
198	905672		
199	910655		
200	915526		
201	919946		
202	923875		
203	927220		
204	930167		
205	932527		
206	934379		
207	935717		
208	936802		
209	937697		
210	938480		
211	939133		

المنارة للاستشارات
www.manaraa.com

212	939720		
213	940265		
214	940789		
215	941262		
216	941679		
217	942039		
218	942395		
219	942757		
220	943100		
221	943444		
222	943808		
223	944155		
224	944492		
225	944827		
226	945134		
227	945425		
228	945710		
229	946023		
230	946293		
231	946549		
232	946765		
233	946950		
234	947142		
235	947294		
236	947427		
237	947563		
238	947680		
239	947808		
240	947926		
241	948036		
242	948149		
243	948257		
244	948396		
245	948535		
246	948662		
247	948799		
248	948906		
249	948990		
250	949040		
251	949092		
252	949146		
253	949219		
254	949299		
255	949650		

المنارة للاستشارات
www.manaraa.com

Appendix E

The results of standard OTSU algorithm data sets.

t	count	Non Textual Gaussian	Textual Gaussian	p	Mean Non_txt_Gh (mu1)	Mean_txt_Gh(mu2)
1	20	2.11E-05	1.00E+00	2.11E-05	2.11E-05	1.53E+02
2	2	2.32E-05	1.00E+00	2.11E-06	2.53E-05	1.53E+02
3	3	2.63E-05	1.00E+00	3.16E-06	3.47E-05	1.53E+02
4	0	2.63E-05	1.00E+00	0	3.47E-05	1.53E+02
5	0	2.63E-05	1.00E+00	0	3.47E-05	1.53E+02
6	2	2.84E-05	1.00E+00	2.11E-06	4.74E-05	1.53E+02
7	1	2.95E-05	1.00E+00	1.05E-06	5.48E-05	1.53E+02
8	14	4.42E-05	1.00E+00	1.47E-05	1.73E-04	1.53E+02
9	11	5.58E-05	0.9999442	1.16E-05	2.77E-04	1.53E+02
10	13	6.95E-05	0.9999305	1.37E-05	4.14E-04	1.53E+02
11	22	9.27E-05	0.9999073	2.32E-05	6.69E-04	1.53E+02
12	62	0.000157953	0.999842	6.53E-05	1.45E-03	1.53E+02
13	82	0.000244301	0.9997557	8.63E-05	2.57E-03	1.53E+02
14	104	0.000353815	0.9996462	0.00011	4.11E-03	1.53E+02
15	149	0.000510714	0.9994893	0.00016	6.46E-03	1.53E+02
16	181	0.000701311	0.9992987	0.00019	9.51E-03	1.53E+02
17	217	0.000929816	0.9990702	0.00023	1.34E-02	1.53E+02
18	280	0.001224662	0.9987753	0.00029	1.87E-02	1.53E+02
19	295	0.001535302	0.9984647	0.00031	2.46E-02	1.53E+02
20	359	0.001913336	0.9980867	0.00038	3.22E-02	1.53E+02
21	404	0.002338756	0.9976612	0.00043	4.11E-02	1.53E+02
22	477	0.002841047	0.997159	0.0005	5.21E-02	1.53E+02
23	569	0.003440215	0.9965598	0.0006	6.59E-02	1.53E+02
24	628	0.004101511	0.9958985	0.00066	8.18E-02	1.53E+02
25	668	0.004804928	0.9951951	0.0007	9.94E-02	1.53E+02
26	730	0.005573632	0.9944264	0.00077	1.19E-01	1.53E+02
27	765	0.006379192	0.9936208	0.00081	1.41E-01	1.53E+02
28	773	0.007193176	0.9928068	0.00081	1.64E-01	1.53E+02
29	811	0.008047175	0.9919528	0.00085	1.89E-01	1.53E+02
30	781	0.008869584	0.9911304	0.00082	2.13E-01	1.53E+02
31	771	0.009681462	0.9903185	0.00081	2.39E-01	1.53E+02
32	794	0.010517559	0.9894824	0.00084	2.65E-01	1.53E+02
33	808	0.011368399	0.9886316	0.00085	2.93E-01	1.53E+02
34	797	0.012207655	0.9877923	0.00084	3.22E-01	1.53E+02
35	809	0.013059548	0.9869405	0.00085	3.52E-01	1.53E+02
36	864	0.013969357	0.9860306	0.00091	3.84E-01	1.53E+02
37	863	0.014878113	0.9851219	0.00091	4.18E-01	1.53E+02
38	859	0.015782657	0.9842173	0.0009	4.52E-01	1.53E+02
39	904	0.016734586	0.9832654	0.00095	4.90E-01	1.53E+02

المنارة للاستشارات

www.manaraa.com

40	889	0.017670721	0.9823293	0.00094	5.27E-01	1.53E+02
41	947	0.01866793	0.9813321	0.001	5.68E-01	1.53E+02
42	963	0.019681988	0.980318	0.00101	6.11E-01	1.53E+02
43	967	0.020700258	0.9792997	0.00102	6.54E-01	1.53E+02
44	1034	0.02178908	0.9782109	0.00109	7.02E-01	1.53E+02
45	1008	0.022850524	0.9771495	0.00106	7.50E-01	1.53E+02
46	1073	0.023980414	0.9760196	0.00113	8.02E-01	1.53E+02
47	1067	0.025103986	0.974896	0.00112	8.55E-01	1.53E+02
48	1130	0.026293898	0.9737061	0.00119	9.12E-01	1.53E+02
49	1191	0.027548044	0.972452	0.00125	9.73E-01	1.53E+02
50	1204	0.02881588	0.9711841	0.00127	1.04E+00	1.52E+02
51	1252	0.03013426	0.9698657	0.00132	1.10E+00	1.52E+02
52	1259	0.031460012	0.96854	0.00133	1.17E+00	1.52E+02
53	1316	0.032845785	0.9671542	0.00139	1.25E+00	1.52E+02
54	1337	0.034253672	0.9657463	0.00141	1.32E+00	1.52E+02
55	1299	0.035621545	0.9643785	0.00137	1.40E+00	1.52E+02
56	1320	0.037011531	0.9629885	0.00139	1.48E+00	1.52E+02
57	1355	0.038438372	0.9615616	0.00143	1.56E+00	1.52E+02
58	1313	0.039820987	0.960179	0.00138	1.64E+00	1.52E+02
59	1292	0.041181488	0.9588185	0.00136	1.72E+00	1.52E+02
60	1338	0.042590428	0.9574096	0.00141	1.80E+00	1.52E+02
61	1286	0.043944611	0.9560554	0.00135	1.88E+00	1.52E+02
62	1260	0.045271416	0.9547286	0.00133	1.97E+00	1.52E+02
63	1312	0.046652977	0.953347	0.00138	2.05E+00	1.51E+02
64	1349	0.048073501	0.9519265	0.00142	2.14E+00	1.51E+02
65	1345	0.049489812	0.9505102	0.00142	2.24E+00	1.51E+02
66	1362	0.050924025	0.949076	0.00143	2.33E+00	1.51E+02
67	1336	0.052330859	0.9476691	0.00141	2.43E+00	1.51E+02
68	1325	0.05372611	0.9462739	0.0014	2.52E+00	1.51E+02
69	1391	0.05519086	0.9448091	0.00146	2.62E+00	1.51E+02
70	1513	0.056784078	0.9432159	0.00159	2.73E+00	1.51E+02
71	1372	0.058228821	0.9417712	0.00144	2.84E+00	1.51E+02
72	1475	0.059782025	0.940218	0.00155	2.95E+00	1.51E+02
73	1462	0.06132154	0.9386785	0.00154	3.06E+00	1.50E+02
74	1448	0.062846312	0.9371537	0.00152	3.17E+00	1.50E+02
75	1487	0.064412152	0.9355878	0.00157	3.29E+00	1.50E+02
76	1507	0.065999052	0.9340009	0.00159	3.41E+00	1.50E+02
77	1513	0.067592271	0.9324077	0.00159	3.53E+00	1.50E+02
78	1522	0.069194967	0.930805	0.0016	3.66E+00	1.50E+02
79	1459	0.070731322	0.9292687	0.00154	3.78E+00	1.50E+02
80	1554	0.072367714	0.9276323	0.00164	3.91E+00	1.50E+02
81	1573	0.074024114	0.9259759	0.00166	4.04E+00	1.49E+02
82	1592	0.075700521	0.9242995	0.00168	4.18E+00	1.49E+02
83	1574	0.077357974	0.922642	0.00166	4.32E+00	1.49E+02
84	1644	0.079089138	0.9209109	0.00173	4.47E+00	1.49E+02
85	1543	0.080713947	0.9192861	0.00162	4.60E+00	1.49E+02
86	1560	0.082356658	0.9176433	0.00164	4.74E+00	1.49E+02

www.manaraa.com

87	1612	0.084054125	0.9159459	0.0017	4.89E+00	1.49E+02
88	1645	0.085786342	0.9142137	0.00173	5.04E+00	1.48E+02
89	1565	0.087434318	0.9125657	0.00165	5.19E+00	1.48E+02
90	1714	0.089239193	0.9107608	0.0018	5.35E+00	1.48E+02
91	1630	0.090955615	0.9090444	0.00172	5.51E+00	1.48E+02
92	1667	0.092710999	0.907289	0.00176	5.67E+00	1.48E+02
93	1676	0.09447586	0.9055241	0.00176	5.84E+00	1.48E+02
94	1767	0.096336545	0.9036635	0.00186	6.01E+00	1.47E+02
95	1653	0.098077186	0.9019228	0.00174	6.18E+00	1.47E+02
96	1679	0.099845206	0.9001548	0.00177	6.35E+00	1.47E+02
97	1666	0.101599537	0.8984005	0.00175	6.52E+00	1.47E+02
98	1771	0.103464434	0.8965356	0.00186	6.70E+00	1.47E+02
99	1783	0.105341968	0.894658	0.00188	6.88E+00	1.47E+02
100	1686	0.107117359	0.8928826	0.00178	7.06E+00	1.46E+02
101	1775	0.108986469	0.8910135	0.00187	7.25E+00	1.46E+02
102	1788	0.110869268	0.8891307	0.00188	7.44E+00	1.46E+02
103	1875	0.112843679	0.8871563	0.00197	7.65E+00	1.46E+02
104	1966	0.114913916	0.8850861	0.00207	7.86E+00	1.46E+02
105	1935	0.116951508	0.8830485	0.00204	8.08E+00	1.45E+02
106	2085	0.119147054	0.8808529	0.0022	8.31E+00	1.45E+02
107	2187	0.121450008	0.87855	0.0023	8.55E+00	1.45E+02
108	2270	0.123840362	0.8761596	0.00239	8.81E+00	1.45E+02
109	2306	0.126268625	0.8737314	0.00243	9.08E+00	1.44E+02
110	2349	0.128742168	0.8712578	0.00247	9.35E+00	1.44E+02
111	2467	0.131339967	0.86866	0.0026	9.64E+00	1.44E+02
112	2602	0.134079924	0.8659201	0.00274	9.94E+00	1.44E+02
113	2568	0.136784078	0.8632159	0.0027	1.03E+01	1.43E+02
114	2630	0.13955352	0.8604465	0.00277	1.06E+01	1.43E+02
115	2773	0.142473543	0.8575265	0.00292	1.09E+01	1.43E+02
116	2901	0.145528353	0.8544716	0.00305	1.13E+01	1.42E+02
117	2952	0.148636866	0.8513631	0.00311	1.16E+01	1.42E+02
118	3111	0.15191281	0.8480872	0.00328	1.20E+01	1.41E+02
119	3112	0.155189807	0.8448102	0.00328	1.24E+01	1.41E+02
120	3227	0.158587901	0.8414121	0.0034	1.28E+01	1.41E+02
121	3184	0.161940715	0.8380593	0.00335	1.32E+01	1.40E+02
122	3365	0.165484126	0.8345159	0.00354	1.36E+01	1.40E+02
123	3365	0.169027536	0.8309725	0.00354	1.41E+01	1.39E+02
124	3605	0.172823672	0.8271763	0.0038	1.45E+01	1.39E+02
125	3603	0.176617701	0.8233823	0.00379	1.50E+01	1.38E+02
126	3640	0.180450692	0.8195493	0.00383	1.55E+01	1.38E+02
127	3751	0.184400569	0.8155994	0.00395	1.60E+01	1.37E+02
128	3876	0.188482072	0.8115179	0.00408	1.65E+01	1.37E+02
129	3969	0.192661507	0.8073385	0.00418	1.71E+01	1.36E+02
130	4229	0.197114726	0.8028853	0.00445	1.76E+01	1.36E+02
131	4406	0.201754331	0.7982457	0.00464	1.83E+01	1.35E+02
132	4746	0.206751961	0.793248	0.005	1.89E+01	1.35E+02
133	4918	0.211930711	0.7880693	0.00518	1.96E+01	1.34E+02

المنارة للاستشارات

www.manaraa.com

134	5444	0.21766335	0.7823367	0.00573	2.04E+01	1.33E+02
135	5535	0.223491813	0.7765082	0.00583	2.12E+01	1.32E+02
136	5834	0.229635129	0.7703649	0.00614	2.20E+01	1.31E+02
137	6360	0.236332333	0.7636677	0.0067	2.29E+01	1.31E+02
138	6734	0.243423367	0.7565766	0.00709	2.39E+01	1.30E+02
139	7042	0.25083873	0.7491613	0.00742	2.49E+01	1.29E+02
140	7608	0.258850103	0.7411499	0.00801	2.60E+01	1.27E+02
141	7497	0.26674459	0.7332554	0.00789	2.72E+01	1.26E+02
142	7674	0.274825462	0.7251745	0.00808	2.83E+01	1.25E+02
143	7799	0.283037961	0.716962	0.00821	2.95E+01	1.24E+02
144	7708	0.291154636	0.7088454	0.00812	3.06E+01	1.23E+02
145	7984	0.299561944	0.7004381	0.00841	3.19E+01	1.22E+02
146	8021	0.308008214	0.6919918	0.00845	3.31E+01	1.20E+02
147	7920	0.316348128	0.6836519	0.00834	3.43E+01	1.19E+02
148	8069	0.324844943	0.6751551	0.0085	3.56E+01	1.18E+02
149	8020	0.33329016	0.6667098	0.00845	3.68E+01	1.17E+02
150	7977	0.341690096	0.6583099	0.0084	3.81E+01	1.15E+02
151	8108	0.350227979	0.649772	0.00854	3.94E+01	1.14E+02
152	8027	0.358680567	0.6413194	0.00845	4.07E+01	1.13E+02
153	8274	0.36739325	0.6326067	0.00871	4.20E+01	1.11E+02
154	8361	0.376197546	0.6238025	0.0088	4.34E+01	1.10E+02
155	8668	0.38532512	0.6146749	0.00913	4.48E+01	1.09E+02
156	8863	0.394658032	0.605342	0.00933	4.62E+01	1.07E+02
157	9495	0.404656452	0.5953435	0.01	4.78E+01	1.06E+02
158	9831	0.415008687	0.5849913	0.01035	4.94E+01	1.04E+02
159	10305	0.425860054	0.5741399	0.01085	5.12E+01	1.02E+02
160	11010	0.437453799	0.5625462	0.01159	5.30E+01	1.00E+02
161	11652	0.449723582	0.5502764	0.01227	5.50E+01	9.85E+01
162	12143	0.462510399	0.5374896	0.01279	5.71E+01	9.64E+01
163	13325	0.476541884	0.5234581	0.01403	5.94E+01	9.41E+01
164	14442	0.491749592	0.5082504	0.01521	6.18E+01	9.16E+01
165	15779	0.508365187	0.4916348	0.01662	6.46E+01	8.89E+01
166	16987	0.52625283	0.4737472	0.01789	6.76E+01	8.59E+01
167	18573	0.545810562	0.4541894	0.01956	7.08E+01	8.27E+01
168	20090	0.566965724	0.4330343	0.02116	7.44E+01	7.91E+01
169	21397	0.589497183	0.4105028	0.02253	7.82E+01	7.53E+01
170	22937	0.613650292	0.3863497	0.02415	8.23E+01	7.12E+01
171	23131	0.638007687	0.3619923	0.02436	8.65E+01	6.70E+01
172	22508	0.661709051	0.3382909	0.0237	9.05E+01	6.30E+01
173	21911	0.684781762	0.3152182	0.02307	9.45E+01	5.90E+01
174	20785	0.706668773	0.2933312	0.02189	9.83E+01	5.52E+01
175	19830	0.72755015	0.2724498	0.02088	1.02E+02	5.15E+01
176	18859	0.747409045	0.252591	0.01986	1.05E+02	4.80E+01
177	16936	0.765242984	0.234757	0.01783	1.09E+02	4.48E+01
178	15132	0.781177276	0.2188227	0.01593	1.11E+02	4.20E+01
179	13139	0.795012899	0.2049871	0.01384	1.14E+02	3.95E+01
180	11780	0.80741747	0.1925825	0.0124	1.16E+02	3.73E+01

المنارة للاستشارات
www.manaraa.com

181	10593	0.818572106	0.1814279	0.01115	1.18E+02	3.53E+01
182	9898	0.828994893	0.1710051	0.01042	1.20E+02	3.34E+01
183	9344	0.838834307	0.1611657	0.00984	1.22E+02	3.16E+01
184	9162	0.848482072	0.1515179	0.00965	1.24E+02	2.98E+01
185	9018	0.857978202	0.1420218	0.0095	1.25E+02	2.81E+01
186	8812	0.867257411	0.1327426	0.00928	1.27E+02	2.63E+01
187	8502	0.876210183	0.1237898	0.00895	1.29E+02	2.47E+01
188	8234	0.884880746	0.1151193	0.00867	1.30E+02	2.30E+01
189	7957	0.893259622	0.1067404	0.00838	1.32E+02	2.14E+01
190	7168	0.900807666	0.0991923	0.00755	1.33E+02	2.00E+01
191	6812	0.907980835	0.0920192	0.00717	1.35E+02	1.86E+01
192	6249	0.914561154	0.0854388	0.00658	1.36E+02	1.74E+01
193	5668	0.920529669	0.0794703	0.00597	1.37E+02	1.62E+01
194	5429	0.926246512	0.0737535	0.00572	1.38E+02	1.51E+01
195	5221	0.931744327	0.0682557	0.0055	1.39E+02	1.40E+01
196	5254	0.937276891	0.0627231	0.00553	1.41E+02	1.30E+01
197	5184	0.942735745	0.0572643	0.00546	1.42E+02	1.19E+01
198	5307	0.948324119	0.0516759	0.00559	1.43E+02	1.08E+01
199	5096	0.953690307	0.0463097	0.00537	1.44E+02	9.70E+00
200	4983	0.958937503	0.0410625	0.00525	1.45E+02	8.65E+00
201	4871	0.964066761	0.0359332	0.00513	1.46E+02	7.62E+00
202	4420	0.968721108	0.0312789	0.00465	1.47E+02	6.68E+00
203	3929	0.972858422	0.0271416	0.00414	1.48E+02	5.84E+00
204	3345	0.976380772	0.0236192	0.00352	1.48E+02	5.12E+00
205	2947	0.97948402	0.020516	0.0031	1.49E+02	4.49E+00
206	2360	0.981969147	0.0180309	0.00249	1.50E+02	3.98E+00
207	1852	0.983919339	0.0160807	0.00195	1.50E+02	3.57E+00
208	1338	0.985328279	0.0146717	0.00141	1.50E+02	3.28E+00
209	1085	0.986470805	0.0135292	0.00114	1.50E+02	3.04E+00
210	895	0.987413258	0.0125867	0.00094	1.51E+02	2.84E+00
211	783	0.988237772	0.0117622	0.00082	1.51E+02	2.67E+00
212	653	0.988925394	0.0110746	0.00069	1.51E+02	2.52E+00
213	587	0.989543516	0.0104565	0.00062	1.51E+02	2.39E+00
214	545	0.990117412	0.0098826	0.00057	1.51E+02	2.27E+00
215	524	0.990669194	0.0093308	0.00055	1.51E+02	2.15E+00
216	473	0.991167272	0.0088327	0.0005	1.51E+02	2.04E+00
217	417	0.991606381	0.0083936	0.00044	1.52E+02	1.95E+00
218	360	0.991985468	0.0080145	0.00038	1.52E+02	1.86E+00
219	356	0.992360343	0.0076397	0.00037	1.52E+02	1.78E+00
220	362	0.992741536	0.0072585	0.00038	1.52E+02	1.70E+00
221	343	0.993102722	0.0068973	0.00036	1.52E+02	1.62E+00
222	344	0.993464961	0.006535	0.00036	1.52E+02	1.54E+00
223	364	0.99384826	0.0061517	0.00038	1.52E+02	1.45E+00
224	347	0.994213658	0.0057863	0.00037	1.52E+02	1.37E+00
225	337	0.994568525	0.0054315	0.00035	1.52E+02	1.29E+00
226	335	0.994921287	0.0050787	0.00035	1.52E+02	1.21E+00
227	307	0.995244564	0.0047554	0.00032	1.52E+02	1.14E+00

المنارة للاستشارات
www.manaraa.com

228	291	0.995550992	0.004449	0.00031	1.52E+02	1.07E+00
229	285	0.995851103	0.0041489	0.0003	1.52E+02	9.99E-01
230	313	0.996180698	0.0038193	0.00033	1.53E+02	9.24E-01
231	270	0.996465013	0.003535	0.00028	1.53E+02	8.58E-01
232	256	0.996734586	0.0032654	0.00027	1.53E+02	7.95E-01
233	216	0.996962039	0.003038	0.00023	1.53E+02	7.42E-01
234	185	0.997156847	0.0028432	0.00019	1.53E+02	6.97E-01
235	192	0.997359027	0.002641	0.0002	1.53E+02	6.49E-01
236	152	0.997519086	0.0024809	0.00016	1.53E+02	6.12E-01
237	133	0.997659138	0.0023409	0.00014	1.53E+02	5.78E-01
238	136	0.997802348	0.0021977	0.00014	1.53E+02	5.44E-01
239	117	0.997925552	0.0020744	0.00012	1.53E+02	5.15E-01
240	128	0.998060338	0.0019397	0.00013	1.53E+02	4.83E-01
241	118	0.998184594	0.0018154	0.00012	1.53E+02	4.53E-01
242	110	0.998300426	0.0016996	0.00012	1.53E+02	4.25E-01
243	113	0.998419418	0.0015806	0.00012	1.53E+02	3.96E-01
244	108	0.998533144	0.0014669	0.00011	1.53E+02	3.68E-01
245	139	0.998679514	0.0013205	0.00015	1.53E+02	3.32E-01
246	139	0.998825883	0.0011741	0.00015	1.53E+02	2.96E-01
247	127	0.998959617	0.0010404	0.00013	1.53E+02	2.63E-01
248	137	0.99910388	0.0008961	0.00014	1.53E+02	2.27E-01
249	107	0.999216553	0.0007834	0.00011	1.53E+02	1.99E-01
250	84	0.999305007	0.000695	8.85E-05	1.53E+02	1.77E-01
251	50	0.999357658	0.0006423	5.27E-05	1.53E+02	1.64E-01
252	52	0.999412415	0.0005876	5.48E-05	1.53E+02	1.50E-01
253	54	0.999469278	0.0005307	5.69E-05	1.53E+02	1.36E-01
254	73	0.999546149	0.0004539	7.69E-05	1.53E+02	1.16E-01
255	80	0.99963039	0.0003696	8.42E-05	1.53E+02	9.46E-02
256	351	1	0	0.00037	1.53E+02	0.00E+00

المنارة للاستشارات
www.manaraa.com

Appendix F

Table shows the recognition results that were used to compute the precision and recall performance measurements.

The automatic recognition system					Ground truth values
Frame #	Selected Text	Not Selected	Correct Rec.	Not Correct Rec.	# Words
Frame 1	8	5	5	3	13
Frame 2	8	5	5	3	13
Frame 3	7	6	5	2	13
Frame 4	8	5	5	3	13
Frame 5	5	3	4	1	8
Frame 6	5	3	3	2	8
Frame 7	5	3	3	2	8
Frame 8	5	3	3	2	8
Frame 9	5	3	3	2	8
Frame 10	9	4	7	2	13
Frame 11	9	4	7	2	13
Frame 12	9	4	7	2	13
Frame 13	9	4	7	2	13
Frame 14	9	4	8	1	13
Frame 15	9	4	7	2	13
Frame 16	9	4	7	2	13
Frame 17	9	4	7	2	13
Frame 18	2	4	1	1	6
Frame 19	2	4	1	1	6
Frame 20	2	4	1	1	6
Frame 21	2	4	1	1	6
Frame 22	5	4	4	1	9
Frame 23	5	4	3	2	9
Frame 24	5	4	3	2	9
Frame 25	4	2	2	2	6
Frame 26	4	2	2	2	6
Frame 27	4	2	2	2	6
Frame 28	4	2	2	2	6
Frame 29	4	2	2	2	6
Frame 30	4	2	2	2	6
Frame 31	4	2	2	2	6
Frame 32	4	2	2	2	6
Frame 33	4	2	2	2	6
Frame 34	12	4	8	4	16
Frame 35	12	4	8	4	16
Frame 36	11	5	8	3	16
Frame 37	10	6	7	3	16
Frame 38	12	4	9	3	16

المنارة للاستشارات
www.manaraa.com

Frame 39	12	4	8	4	16
Frame 40	12	4	8	4	16
Frame 41	12	4	8	4	16
Frame 42	12	4	8	4	16
Frame 43	15	2	10	5	17
Frame 44	15	2	10	5	17
Frame 45	15	2	10	5	17
Frame 46	15	2	10	5	17
Frame 47	15	2	10	5	17
Frame 48	11	7	5	6	18
Frame 49	11	7	5	6	18
Frame 50	11	7	5	6	18
Frame 51	25	12	14	11	37
Frame 52	25	12	14	11	37
Frame 53	25	12	14	11	37
Frame 54	25	12	14	11	37
Frame 55	25	12	14	11	37
Frame 56	25	12	14	11	37
Frame 57	25	12	14	11	37
Frame 58	25	12	14	11	37
Frame 59	25	12	14	11	37
Frame 60	25	12	14	11	37
Frame 61	24	13	19	5	37
Frame 62	25	12	19	6	37
Frame 63	25	12	19	6	37
Frame 64	25	12	19	6	37
Frame 65	25	12	19	6	37
Frame 66	25	12	19	6	37
Frame 67	25	12	19	6	37
Frame 68	25	12	19	6	37
Frame 69	25	12	19	6	37
Frame 70	7	6	3	4	13
Frame 71	7	6	3	4	13
Frame 72	7	6	3	4	13
Frame 73	7	6	3	4	13
Frame 74	7	6	3	4	13
Frame 75	7	6	3	4	13
Frame 76	7	6	3	4	13
Frame 77	7	6	3	4	13
Frame 78	7	6	3	4	13
Frame 79	7	6	3	4	13
Frame 80	9	1	6	3	10
Frame 81	9	1	6	3	10
Frame 82	9	1	6	3	10
Frame 83	8	2	6	2	10
Frame 84	8	2	5	3	10
Frame 85	8	2	5	3	10

المنارة للاستشارات
www.manaraa.com

Frame 86	14	6	9	5	20
Frame 87	14	6	9	5	20
Frame 88	14	6	9	5	20
Frame 89	14	6	9	5	20
Frame 90	14	6	9	5	20
Frame 91	14	6	9	5	20
Frame 92	14	6	11	3	20
Frame 93	23	2	18	5	25
Frame 94	23	2	18	5	25
Frame 95	23	2	18	5	25
Frame 96	23	2	18	5	25
Frame 97	23	2	18	5	25
Frame 98	23	2	18	5	25
Frame 99	23	2	18	5	25
Frame 100	23	2	18	5	25
Frame 101	23	2	18	5	25
Frame 102	23	2	18	5	25
Frame 103	23	2	18	5	25
Frame 104	8	1	5	3	9
Frame 105	7	2	4	3	9
Frame 106	7	2	4	3	9
Frame 107	8	1	6	2	9
Frame 108	7	2	4	3	9
Frame 109	8	1	3	5	9
Frame 110	6	3	5	1	9
Frame 111	8	1	6	2	9
Frame 112	11	3	5	6	14
Frame 113	11	3	5	6	14
Frame 114	11	3	5	6	14
Frame 115	11	3	5	6	14
Frame 116	12	2	7	5	14
Frame 117	8	6	7	1	14
Frame 118	10	4	8	2	14
Frame 119	12	2	6	6	14
Frame 120	14	3	9	5	17
Frame 121	13	4	8	5	17
Frame 122	13	4	8	5	17
Frame 123	13	4	8	5	17
Frame 124	13	4	8	5	17
Frame 125	13	4	8	5	17
Frame 126	13	4	8	5	17
Frame 127	15	2	13	2	17
Frame 128	28	5	19	9	33
Frame 129	28	5	19	9	33
Frame 130	28	5	19	9	33
Frame 131	28	5	19	9	33
Frame 132	28	5	19	9	33

المنارة للاستشارات
www.manaraa.com

Frame 133	28	5	19	9	33
Frame 134	27	7	18	9	34
Frame 135	27	7	18	9	34
Frame 136	27	7	18	9	34
Frame 137	27	7	19	8	34
Frame 138	16	12	14	2	28
Frame 139	16	12	15	1	28
Frame 140	17	11	10	7	28
Frame 141	17	11	10	7	28
Frame 142	17	11	10	7	28
Frame 143	17	11	10	7	28
Frame 144	11	4	7	4	15
Frame 145	11	4	4	7	15
Frame 146	11	4	4	7	15
Frame 147	11	4	5	6	15
Frame 148	11	4	5	6	15
Frame 149	11	4	5	6	15
Frame 150	12	3	8	4	15
Frame 151	12	3	8	4	15
Frame 152	12	3	8	4	15
Frame 153	12	3	8	4	15
Frame 154	12	3	8	4	15
Frame 155	12	3	8	4	15
Frame 156	3	4	1	2	7
Frame 157	3	4	1	2	7
Frame 158	3	4	1	2	7
Frame 159	5	3	4	1	8
Frame 160	5	3	4	1	8
Frame 161	5	3	3	2	8
Frame 162	5	3	3	2	8
Frame 163	9	2	5	4	11
Frame 164	9	2	5	4	11
Frame 165	9	2	6	3	11
Frame 166	9	2	6	3	11
Frame 167	9	2	6	3	11
Frame 168	9	2	6	3	11
Frame 169	7	1	5	2	8
Frame 170	6	2	4	2	8
Frame 171	6	2	2	4	8
Frame 172	6	2	2	4	8
Frame 173	6	2	2	4	8
Frame 174	6	2	2	4	8
Frame 175	6	2	2	4	8
Frame 176	8	1	5	3	9
Frame 177	8	1	5	3	9
Frame 178	7	2	6	1	9
Frame 179	8	1	7	1	9

المنارة للاستشارات

www.manaraa.com

Frame 180	8	1	5	3	9
Frame 181	13	4	8	5	17
Frame 182	13	4	8	5	17
Frame 183	13	4	8	5	17
Frame 184	13	4	8	5	17
Frame 185	13	4	8	5	17
Frame 186	13	4	8	5	17
Frame 187	13	4	8	5	17
Frame 188	13	4	8	5	17
Frame 189	13	4	8	5	17
Frame 190	12	5	10	2	17
Frame 191	12	5	11	1	17
Frame 192	13	4	8	5	17
Frame 193	13	4	8	5	17
Frame 194	13	4	8	5	17

www.manaraa.com

Appendix G

Table shows the final video classification for the target sample videos in the proposed experiment.

Video #	Video Name	Course Name	Video Source	Video Duration	Number of Key Frames	Avg. Similarity of Key frames	Extracted Video Type	Actual Video Type
1	Threat Models	Security	MIT	[15:00]	54	53.23%	Handwritten	Handwritten Video
2	Control Hijacking Attacks	Security	MIT	[14:26]	49	60.70%	Handwritten	Handwritten Video
3	Web security Model	Security	MIT	[13:11]	55	79.31%	Handwritten	Handwritten Video
4	Network Security	Security	MIT	[20:18]	69	81.30%	Spoken Text	Handwritten Video
5	SSL and HTTPs	Security	MIT	[06:09]	52	51.90%	Handwritten	Handwritten Video
6	Data Tracking	Security	MIT	[13:25]	77	63.50%	Handwritten	Handwritten Video
7	Introduction and Scope	Artificial Intelligence	MIT	[14:21]	86	45.93%	Handwritten	Handwritten Video
8	Searching : Hill Climbing	Artificial Intelligence	MIT	[03:20]	35	68.10%	Handwritten	Handwritten Video
9	Search Constraints	Artificial Intelligence	MIT	[08:11]	48	65.40%	Handwritten	Handwritten Video
10	Learning Identification	Artificial Intelligence	MIT	[02:26]	32	69.80%	Handwritten	Handwritten Video
11	Learning Genetic Algorithms	Artificial Intelligence	MIT	[03:51]	58	84.60%	Spoken Text	Handwritten Video
12	Learning Boosting	Artificial Intelligence	MIT	[04:15]	67	39.69%	Presentational	Handwritten Video
13	Probabilistic Inference I	Artificial Intelligence	MIT	[02:29]	41	48.01%	Handwritten	Handwritten Video
14	Model Merging	Artificial Intelligence	MIT	[09:00]	80	22.10%	Presentational	Handwritten Video
15	Software Product Management	Project Management	Coursera	[04:30]	39	51.60%	Handwritten	Presentational Video
16	Client Needs and S/W Req.	S/W Engineering	Coursera	[10:07]	72	39%	Presentational	Presentational Video
17	Agile Planning for S/W Products	S/W Engineering	Coursera	[06:10]	59	42.10%	Handwritten	Presentational Video
18	S/W Improvement Metrics	S/W Engineering	Coursera	[12:35]	116	17.10%	Presentational	Presentational Video
19	Principles of Computing	Fundamentals of Computing Specialization	Coursera	[20:21]	163	10.40%	Presentational	Presentational Video
20	Algorithmic Thinking	Fundamentals of Computing Specialization	Coursera	[04:00]	68	16.80%	Presentational	Presentational Video
21	HTML, CSS and JavaScript	Full Stack Web Development	Coursera	[15:26]	205	17.64%	Presentational	Presentational Video
22	Front-End JavaScript	Full Stack Web Development	Coursera	[03:59]	48	35.80%	Presentational	Presentational Video
23	Server Side Development	Full Stack Web Development	Coursera	[19:00]	176	24.50%	Presentational	Presentational Video
24	Network Components	Computer Network	YouTube	[15:59]	199	26.78%	Presentational	Presentational Video
25	Physical Layer	Computer Network	YouTube	[14:00]	137	8.30%	Presentational	Presentational Video
26	Error Detection	Computer Network	YouTube	[05:14]	80	19.90%	Presentational	Presentational Video
27	Introduction and Scope	Algorithms	YouTube	[17:20]	193	14.05%	Presentational	Presentational Video
28	Integer Multiplication	Data Structure	YouTube	[07:47]	85	6.30%	Presentational	Presentational Video
29	Merge Sort	Data Structure	YouTube	[06:22]	58	25.39%	Presentational	Presentational Video
30	Big Oh Notation	Algorithms	YouTube	[03:10]	43	40.10%	Handwritten	Presentational Video
31	Graph Search Overview	Data Structure	YouTube	[20:03]	304	11.60%	Presentational	Presentational Video
32	Python - Selection	Algorithms	YouTube	[02:41]	68	18.96%	Presentational	Presentational Video
33	AVL Trees	Algorithms	YouTube	[11:38]	29	56.3	Handwritten	Handwritten Video
34	Introduction to data structures	Data Structure	YouTube	[07:55]	96	17.04%	Presentational	Presentational Video

المنارة للاستشارات
www.manaraa.com

35	Hashing Algorithms	Data Structure	YouTube	[02:19]	63	23.80%	Presentational	Presentational Video
36	MATLAB Basics	MATLAB Course	YouTube	[15:12]	152	15.80%	Presentational	Presentational Video
37	Plotting Functions	MATLAB Course	YouTube	[06:10]	82	11.30%	Presentational	Presentational Video
38	Vectors	MATLAB Course	YouTube	[07:41]	99	26.10%	Presentational	Presentational Video
39	Array Manipulation	MATLAB Course	YouTube	[06:00]	72	29.68%	Presentational	Presentational Video
40	Monte Carlo Simulation	MATLAB Course	YouTube	[10:50]	101	20.90%	Presentational	Presentational Video
41	What is OOP?	Object Oriented Programming	YouTube	[08:23]	118	13%	Presentational	Presentational Video
42	Introduction to Solid Principles	Object Oriented Programming	YouTube	[12:05]	8	81.30%	Spoken Text	Spoken Text Video
43	Symbolic Execution	Web application	MIT	[05:39]	3	94.60%	Spoken Text	Spoken Text Video
44	Medical Software	Software Engineering	MIT	[14:19]	6	91.23%	Spoken Text	Spoken Text Video
45	صفات المدير الناجح	ادارة الاعمال	YouTube	[03:08]	2	83.70%	Spoken Text	Spoken Text Video
46	Creating Conceptual Model	Object Oriented Programming	YouTube	[10:42]	4	82.30%	Spoken Text	Spoken Text Video
47	Designing Impromptu Speeches	Introduction to public Speaking	YouTube	[22:09]	7	93.90%	Spoken Text	Spoken Text Video
48	Designing Informative Speeches	Introduction to public Speaking	YouTube	[18:36]	7	92.89%	Spoken Text	Spoken Text Video
49	Lecture Transcript	Introduction to public Speaking	Coursera	[09:00]	5	85.87%	Spoken Text	Spoken Text Video
50	القيادة ومهارات الاتصال داخل المؤسسة	الموارد البشرية	YouTube	[02:54]	3	96.10%	Spoken Text	Spoken Text Video
51	مقابلة العمل	الموارد البشرية	YouTube	[04:26]	2	95.18%	Spoken Text	Spoken Text Video
52	Job Interview tips	Human Resources	YouTube	[02:33]	3	89.72%	Spoken Text	Spoken Text Video
53	كيف تحب عملك	الموارد البشرية	YouTube	[11:08]	3	90.55%	Spoken Text	Spoken Text Video
54	International Relations	Political Science	YouTube	[20:13]	12	79.30%	Handwritten	Spoken Text Video
55	Public Law	Political Science	YouTube	[15:41]	9	88.16%	Spoken Text	Spoken Text Video
56	Introduction to Politics	Politics Principles	YouTube	[06:30]	4	90.09%	Spoken Text	Spoken Text Video

System Performance:

$$Accuracy = \frac{Correctly\_Classified}{Number\_of\_Videos} \times 100\%$$

$$Error\_Rate = \frac{In\_Correctly\_Classified}{Number\_of\_Videos} \times 100\%$$

$$Accuracy = \frac{49}{56} \times 100\%$$

$$Error\_Rate = \frac{7}{56} \times 100\%$$

Accuracy = 87.5% *Error Rate* = 12.5 %

www.manaraa.com

Appendix H

Table shows the CPU response time for each lecture video text extraction process using the iterative threshold algorithm.

CPU Response Time (ms)			
Video #	**Iterative / Free OCR**	**Iterative / Tesseract**	**Iterative / OCR AD**
Video 1	8445	6853	6115
Video 2	9647	8453	8025
Video 3	5853	3402	1996
Video 4	8746	7253	6593
Video 5	8242	6583	5793
Video 6	6734	4576	3397
Video 7	7545	5654	4685
Video 8	8264	6612	5828
Video 9	11264	10605	10594
Video 10	10514	9607	9403
Video 11	6695	4524	3335
Video 12	10760	9935	9794
Video 13	8619	7085	6392
Video 14	7888	6112	5231
Video 15	8609	7071	6376
Video 16	7756	5937	5022
Video 17	7831	6036	5141
Video 18	9327	8028	7518
Video 19	9253	7928	7399
Video 20	9260	7938	7411
Video 21	8569	7018	6313
Video 22	6827	4699	3545
Video 23	8611	7075	6380
Video 24	9458	8201	7725
Video 25	8400	6793	6044
Video 26	8906	7466	6848
Video 27	8620	7086	6394
Video 28	7665	5815	4877
Video 29	8219	6552	5756
Video 30	7217	5218	4164
Video 31	8770	7286	6632
Video 32	6883	4774	3634
Video 33	6956	4871	3749
Video 34	7196	5191	4131
Video 35	5217	2556	987
Video 36	9280	7964	7442
Video 37	8248	6591	5803
Video 38	7247	5258	4211
Video 39	9217	7880	7342
Video 40	6360	4077	2802

المنارة للاستشارات
www.manaraa.com

Video 41	7852	6063	5173
Video 42	7723	5892	4968
Video 43	8242	6583	5794
Video 44	8237	6576	5784
Video 45	7145	5122	4050
Video 46	7919	6152	5279
Video 47	7794	5986	5081
Video 48	8528	6964	6248
Video 49	8960	7539	6934
Video 50	8975	7558	6957
Video 51	7146	5124	4051
Video 52	8018	6285	5437
Video 53	6821	4691	3535
Video 54	6914	4815	3683
Video 55	7940	6181	5313
Video 56	9367	8081	7581
Average	**8155.339286**	**6467.410714**	**5655.267857**

المنارة للاستشارات
www.manaraa.com