UNIVERSITI PUTRA MALAYSIA

TEMPORAL VIDEO SEGMENTATION USING SQUARED FORM OF KRAWTCHOUK-TCHEBICHEF MOMENTS

SADIQ H. ABDULHUSSAIN

TEMPORAL VIDEO SEGMENTATION USING SQUARED FORM OF KRAWTCHOUK-TCHEBICHEF MOMENTS

By

SADIQ H. ABDULHUSSAIN

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Doctor of Philosophy

June 2018

DEDICATION

It broke my heart to lose you
But you did not go alone
A part of me went with you
The day god called you to heaven
A million times I've thought of you
A million times I've cried
May the winds of heaven blow softly and whisper in your ears
How much I love you and miss you.

To my beloved son ***Mohammed****, The bird in heaven.*

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

TEMPORAL VIDEO SEGMENTATION USING SQUARED FORM OF KRAWTCHOUK-TCHEBICHEF MOMENTS

By

SADIQ H. ABDULHUSSAIN

June 2018

Chairman: Associate Professor Abd Rahman Ramli, PhD
Faculty: Engineering

Rapid growth of multimedia data in cyberspace caused a swift rise in data transmission volume. This growth necessitates to look for superior techniques in processing data content. Video contains a lot of useful information; however, it consumes a vast storage space. Content based video indexing and retrieval (CBVIR) aims to automate the management, indexing and retrieval of video. Temporal Video Segmentation (TVS) process is the essential stage in CBVIR which aims to detect transitions between consecutive shots of videos. TVS algorithm design is still challenging because most of the recent methods are unable to achieve robust detection for different types of transitions: hard transition (HT) and soft transition (ST). In this regard, the aim of this study is to propose an efficient TVS algorithm with high precision and recall values, and low computation cost for detecting different types of transitions. In the first part of the proposed algorithm, unique moments coefficients (features) are extracted using a new hybrid set of orthogonal polynomials which is derived based on the modified forms of Krawtchouk and a Tchebichef polynomials. The extracted moments showed superior energy compaction and localization capabilities. For extracting moments, a mathematical model of block processing that requires low computational cost is proposed. Moreover, three different types of moments, namely smoothed moments, and moments of gradients in x and y directions, form the unique feature vectors using embedded image kernel model. In the proposed TVS, a modified candidate segment selection technique is initially employed to determine the candidate segments from the entire video. The Support Vector Machine (SVM) classifier is trained to detect transitions. Specifically, the HTs are detected by the trained SVM model and then refined to eliminate the false-alarm events. The fade transitions are detected based on the smoothed moments energy and the moments of gradients correlation for the candidate segments. In addition, the wipe and dissolve transitions are detected using the change-point detection technique, SVM model, and scale invariant feature transform (SIFT). For all TVS algorithm stages, the moments are computed only for region of interest. The proposed algorithm has been evaluated on four datasets: TRECVID 2001, TRECVID 2005, TRECVID 2006, and TRECVID 2007. The

performance of the proposed algorithm is compared to that of several state-of-the-art TVS algorithms. The improvement results of the proposed algorithm in terms of precision, recall, and F1-score are within the ranges (0.12 – 10.06), (1.65 – 8.25), and (0.88 – 13.85), respectively. Moreover, the proposed method shows low computation cost which is ~5.5% of real-time. The proposed method is found to be useful to tackle the limitations of the existing methods and serve TVS process efficiently.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

SEGMENTASI VIDEO TEMPORAL MENGGUNAKAN MOMEN *KRAWTCHOUK-TCHEBICHEF* BERBENTUK PERSEGI

Oleh

SADIQ H. ABDULHUSSAIN

Jun 2018

Pengerusi: Profesor Madya Abd Rahman Ramli, PhD
Fakulti: Kejuruteraan

Pertumbuhan pesat data multimedia di ruang siber menyebabkan pertambahan pantas dalam jumlah penghantaran data. Pertumbuhan ini memerlukan untuk mencari teknik yang lebih baik dalam memproses kandungan data. Video mengandungi banyak maklumat berguna; walau bagaimanapun, ia menggunakan ruang storan yang besar. Pengindeksan dan dapatan semula video berasaskan kandungan (CBVIR) bertujuan untuk mengautomasikan pengurusan, pengindeksan dan dapatan semula video. Proses Segmentasi Video Temporal (TVS) adalah tahap penting dalam CBVIR bertujuan untuk mengesan peralihan antara rakaman singkat (klip) video berturut-turut. Reka bentuk algoritma TVS masih mencabar kerana kebanyakan kaedah terkini tidak dapat mencapai pengesanan yang mantap untuk pelbagai jenis peralihan: peralihan keras (HT) dan peralihan lembut (ST). Dalam kaitan ini, matlamat kajian ini adalah untuk mencadangkan algoritma TVS yang cekap dengan nilai ketepatan dan panggilan balik yang tinggi, dan kos pengkomputan yang rendah untuk mengesan pelbagai jenis peralihan. Di bahagian pertama algoritma yang dicadangkan, pekali momen unik (ciri) disari menggunakan set hibrid ortogonal polinomial baru yang diterbitkan berdasarkan bentuk polinomial Krawtchouk dan Tchebichef yang diubahsuai. Momen yang disari menunjukkan pemadatan tenaga dan keupayaan penyetempatan yang lebih baik. Untuk menyari momen, model pemprosesan blok matematik yang memerlukan kos pengkomputan yang rendah dicadangkan. Selain itu, tiga jenis momen yang berbeza, iaitu momen terlicin, dan momen kecerunan dalam arah x dan y, membentuk vektor ciri unik menggunakan model kernel imej terbenam. Dalam TVS yang dicadangkan, teknik pemilihan calon segmen yang diubah suai pada mulanya digunakan untuk menentukan calon segmen daripada keseluruhan video. Pengelas Mesin Vektor Sokongan (SVM) dilatih untuk mengesan peralihan. Khususnya, HT dikesan oleh model SVM terlatih dan kemudian diperhalusi untuk menghapuskan peristiwa penggera palsu. Peralihan pudar dikesan berdasarkan tenaga momen terlicin dan momen korelasi kecerunan bagi calon segmen. Di samping itu, penghapusan dan pembubaran peralihan dikesan menggunakan teknik pengesanan titik perubahan, model SVM, dan penjelmaan ciri invarian skala (SIFT). Untuk semua peringkat algoritma TVS, momen dihitung hanya untuk kawasan berminat. Algoritma yang dicadangkan telah dinilai pada empat set data: TRECVID 2001, TRECVID 2005, TRECVID 2006, dan TRECVID 2007. Prestasi algoritma yang dicadangkan dibandingkan dengan beberapa algoritma TVS terkini. Peningkatan hasil algoritma yang dicadangkan dari segi ketepatan, panggilan balik, dan skor F1 adalah dalam julat (0.16 - 10.06), (1.65 - 8.25), dan (0.88 - 13.85), masing-masing. Selain itu, kaedah yang dicadangkan menunjukkan kos pengkomputan yang rendah iaitu ~ 5.5% daripada masa nyata. Kaedah yang dicadangkan didapati berguna untuk menangani batasan kaedah yang sedia ada dan dapat memberikan perkhidmatan TVS dengan cekap.

ACKNOWLEDGEMENTS

Firstly, I would like to express my special gratitude and appreciation to my supervisor Associate Professor Dr. Abd Rahman Ramli for his guidance throughout my Ph.D. study at UPM. His broad knowledge, interest and intuition in my study affected greatly of the way I do my research throughout my work. He is a kind person and made ever thing easy to me by his support and respectful comments during the years of my PhD study. He is a special person and loved by everyone. He supported me to be as independent researcher and gave the opportunity to be what I am today. Indeed, thanks a lot Prof. Ramli for your kindness, wisdom, respect, and support during my difficult moments in my work.

I would also like to thank my supervisory committee members: Associate Prof. Dr. Syed Abdul Rahman Al-Haddad, Prof. Dr. Iqbal M. Sariban and Associate Prof. Dr. Wissam A. Jassim for served as my best lecturers in my work field. They are a critical thinker and excellent academician who believe in high quality standard when analyzing different issues in their fields. And great thanks go to Prof. Dr. Wissam A. Jassim who had a really informative discussion in my defense and his significant ideas and suggestions which are very useful in improving and extending this work. He spent a lot of time to give me valuable comments and instruct me to develop and enhance this thesis during my study period.
Special thanks go to my wife, Basheera M. Mahmmod, for her great assistance and support in everything in all of these years of Ph.D work. Actually, I could not reach what I am today without her. She was very kindness, praiseworthy and kind hearted.

Last but not least, I would like to extend my heart felt appreciation to my beloved family, my children: Mariam, Jafar, and Toqa, for their believing in me and their understanding and encouragement was in the end what made this work possible. My big gratitude to my father and mother which are with me all days and nights of my work and for their asking and praying. And million thanks to my brothers Mosa and Bakir, Ahmed and my twin sisters Amna and Zahraa, for their continuous inspiration, support and trust me to continue this study.

Without any hesitation, I would like to express my thanks to the Ministry of Higher Education and Scientific Research in the Republic of Iraq for making my dream of acquiring higher education a reality. The scholarship offer came a time I needed it most due to my rising interest to pursue a PhD degree.
At the end, special appreciations go to UPM for providing me and all its students a peaceful, complete, and calm academic environment to study and finishing the scientific path.

Sadiq H. Abdulhussain

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

Abdul Rahman b. Ramli, PhD
Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Chairman)

M. Iqbal b. Saripan, PhD
Professor
Faculty of Engineering
Universiti Putra Malaysia
(Member)

Syed Abdul Rahman Al-Haddad, PhD
Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Member)

Wissam A. Jassim, PhD
Research Fellow
Faculty of Engineering
University of Deblin, Ireland
(Member)

____

ROBIAH BINTI YUNUS, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

Declaration by graduate student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any other institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and Innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software.

Signature:____________________________ Date:_________________

Name and Matric No.: Sadiq H. Abdulhussain, GS40252

Declaration by Members of Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature: ______________________________

Name of Chairman of Supervisory Committee: Associate Professor Dr. Abdul Rahman b. Ramli

Signature: ______________________________

Name of Member Of Supervisory Committee: Professor Dr. M. Iqbal b. Saripan

Signature: ______________________________

Name of Member Of Supervisory Committee: Associate Professor Dr. Syed Abdul Rahman Al-Haddad

Signature: ______________________________

Name of Member Of Supervisory Committee: Dr. Wissam A. Jassim

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

1D	One Dimension
2D	Two Dimension
3D	Three Dimension
CBVIR	Content Based Video Indexing and Retrieval
CDSS	Construction of Dissimilarity/Similarity Signal
CHP	Change Point
CLDS	Classification of CDSS
COD	Compressed Domain
COM	Continuous Orthogonal Moment
DCT	Discrete Cosine Transform
DFT	Discrete Fourier Transform
DKT	Discrete Krawtchouk Transform
DKTT	Discrete Krawtchouk-Tchebichef Transform
DS	Dissimilarity Signal
DTKT	Discrete Tchebichef Krawtchouk Transform
DTT	Discrete Tchebichef Transform
DWT	Discrete Wavelet Transform
EBA	Edge-Based Approach
EC	Energy Compaction
ECR	Edge Change Ratio
FAM	False Alarm Signals
FAR	False Alarm Rate
FOI	Fade Out-In
GED	Generalized Eigenvalue Decomposition
GM	Geometric Moments
HBA	Histogram Based Approach
HDM	Histogram Difference Metric
HT	Hard Transition
IDKTT	Inverse Discrete Krawtchouk-Tchebichef Transform
KP	Krawtchouk Polynomial
KPC	Krawtchouk Polynomial Coefficients
LKP	Local Key-Point
MOG	Moment of Gradient
MSD	Miss Detect
NN	Neural Network
NT	Non-Transition
OCM	Object/Camera Motion
OM	Orthogonal Moment
OP	Orthogonal Polynomial
PBA	Pixel Based Approach

PC	Polynomial Coefficients
PELT	Pruned Exact Linear Time
RMSE	Root Mean Square Error
ROI	Region of Interest
ROVI	Representation of Visual Information
SBRA	Bi-Recursive Algorithm
SKTP	Squared Krawtchouk Tchebichef Polynomial
SKTT	Squared Krawtchouk Tchebichef Transform
SS	Similarity Signal
ST	Soft Transition
STA	Statistical Based Approach
SVM	Support Vector Machine
TBA	Transform Based Approach
TC	Transform Coefficient
TP	Tchebichef Polynomial
TPC	Tchebichef Polynomial Coefficients
TR	Transition
TV	Television
TVS	Temporal Video Segmentation
UCD	Uncompressed Domain
VEP	Video Editing Process
VFL	Video Frame Level
VPP	Video Production Process
VSW	Video Sharing Websites
WDT	Wipe and Dissolve Transition
OPBP	Orthogonal Polynomial Based Block Processing
OPEIK	Orthogonal Polynomial embedded image kernel

CHAPTER 1

INTRODUCTION

This chapter presents an introduction of the work carried out. A background on multimedia data over the cyberspace is first presented with a focus on their importance in real applications. Then, the need for automated video content analysis is explained. Afterward, temporal video segmentation (TVS) with its limitations are presented, which explain how the problem statements have been formed as a result of the drawbacks found in the existing methods. Subsequently, the research problem statements are listed in detail before presenting the research aim and objectives. In addition, the brief methodology and research scope are described. The organization of this thesis is provided at the end of this chapter.

1.1 Background

The immense development of computer performance, availableness of storage media, and multimedia technologies during the past decades resulted in the dominance of multimedia data in the cyberspace. The rapid increase of multimedia data over the cyberspace through the past two decades have led to swift rise in the data transmission volume and repository size [1]. However, video, among multimedia data, is considered the most consumed in terms of storage space and information [2]. Companies and individuals have been sharing their media through video sharing websites (VSW) such as YouTube, VIMEO, and Dailymotion, to broaden their audience. The size of these VSWs is dramatically increasing annually. For example YouTube, one of the popular VSW globally [3], 72 hours of videos every minute were uploaded in 2014 and 4 billion hours watched every day [1]. Videos are uploaded and viewed in inconceivable rate at VSWs and social networks [4].

Video repetition is quite occurred in many forms, for example, downloading and uploading a video as it is, after inserting a logo, and covering (hiding) copyrights by replacing video features. Moreover, various video editing software packages are readily available on computers and portable devices which provide the ability to the users to combine two or more videos into one video, omit certain video parts, and alter video by other video contents. Portable devices such as smart-phones, motivate individuals to upload their videos easily to a VSW or social network. Therefore, video consumption is growing up very rapidly. This rapid increment in video data has invoked the need for an efficient management of video that can process and store that sheer volume of data [5], [6]. Indexing and retrieval of multimedia information are utilized to store, depict, and arrange multimedia data appropriately and swiftly [7]. Decades ago, multimedia databases of videos have been comparatively small in which the annotation was done manually on keywords. Conversely, nowadays, these databases become enormous in size and video content information, which activate the need for automated video structure analysis without human involvement [7]–[10]. Video structure analysis involves content based video indexing, and retrieval (CBVIR), that aims to automate the management, indexing, and retrieval of video [8] with respect to their spatiotemporal, visual and semantic contents [11]. CBVIR have various and wide applications. For example, browsing video folders, news event analysis, digital museums, intelligent management in VSW, and video surveillance [7].

A Shot is considered the basic entity of the video sequence [1], [12]. It is defined as a contiguous sequence of frames that have temporal and locale connection. Shot frames are acquired and logged by a single camera [8], [13], [14]. In video production process, shots are aggregated together to form a scene and scenes concatenated together to form the entire video. The aggregation between video shots is known as a shot transition which can be divided into two types: hard transition (HT) and soft transition (ST). The HT is the editing process of concatenating two shots side by side. While, the ST is the editing process of involving multiple frames in the transition and have many forms such as fade, wipe, and dissolve. Generally, frames involved in a transition are not preferred for video indexing or summarization processes because it have low information content [10].

TVS, named also shot boundary detection, partitions the video into its basic units (shots) to be forwarded to the CBVIR for further analysis [15], [16]. In other words, TVS is utilized as an initial and substantial stage in CBVIR; where, its performance affects the results of the next CBVIR stages [10], [17], [18].

Detection of transitions in TVS algorithms is performed by the statistical machine learning-based and/or rules-based techniques. The machine learning-based technique includes supervised and unsupervised learning [19]. Feature extraction is considered the first step in TVS algorithms which aims to obtain significant representation of visual information. Feature extraction can be categorized based on the processed domain into: compressed domain (COD) and uncompressed domain (UCD) based algorithms [20]. TVS algorithms are primarily centered on the UCD, for instance, pixel-based approach [21]. Then they are developed to encompasses other approaches such as: histogram-based [22], edge-based [23], transform-based [24], and local keypoint-based [1] approaches. In transform-based TVS algorithms, various studies used discrete transforms such as Fourier transform, Wavelet transform, and Walsh-Hadamard transform as a feature extraction tool. These methods exhibits a good performance in detecting video shot transitions [25]; however, their computational cost is considered high [19]. In addition, improvement in terms of the detection accuracy for HTs and STs is still demanded [19].

In general, the efficient performance of TVS is based on its ability to detect the shot transition (shot boundary) in the video scene. That is, TVS performance can be measured by its accuracy in detecting correct transition. Where, TVS accuracy mostly depends on how effectively the visual content of video frames are extracted and represented [25].

The existing TVS algorithms either give good recall at the expense of much higher false detection rate, i.e. low precision, or very low false detection rate at the expenses of very low recall [25]. The other significant factor influences TVS performance is the computational cost of the algorithm that always needs to be reduced, where algorithm speed is increased. Note that, in the shot, frames are very similar in terms of their visual content. Therefore, when transition occurs, a change in similarity values will appear. In HT, the rate of change is very high, but for ST it is not so apparent [26]. In addition to that, there are some special effects that appear in the video scene such as; flashlights or light variations, object/camera motion (camera/object motion highly disturbs the accuracy of detecting shot transitions correctly), and camera operation (such as zooming, panning, and tilting). These effects impact greatly on TVS performance. To fulfill the maximum efficiency, TVS should be able to detect shot transitions between two consecutive shots by minimizing both false alarm signals (FAM), i.e. false positives, within a shot (intra-shot frames), and miss detects (MSD), i.e. false negatives, between two consecutive shots (inter-shot frames) during transition detection process. Accordingly, design of a TVS algorithm, which can combine the solutions to these problems, becomes a necessity.

1.2 Problem Statement

The management and search of video data for specific events from large video database are considered difficult [19]. Therefore, an effective TVS algorithm is required as an essential step in CBVIR [25]. Although there is much attention on TVS in the last two decades, there are still no favorable algorithm for detecting all transition types between shots due to the randomness and size of video raw data [27]. Therefore, it is necessary to develop a robust TVS algorithm that has accurate feature extraction with suitable dissimilarity measure and classification. On the other hand, recent TVS based applications increase the demand for increasing detection accuracy and decreasing computational cost concurrently. These issues are rarely implemented in the same work [28]. Moreover, optimization TVS algorithm performance via a forceful discrete transform must also be considered. In particular, the following problems have been addressed throughout this study.

1- An efficient algorithm with constant quality for detecting HTs and STs have not been found yet [29]. The representation of visual content and extraction of compacted features are the significant steps in TVS algorithms [20] to handle multiple types of transitions. Therefore, the accurate feature extraction from video frames or subset of it, called a region of interest (ROI), is the most significant step [30]. Essentially, accurate features should have a compact representation (high energy compaction) of frame's visual content [31]. Besides, the minor changes in the content would have marginal impact on the descriptor [26]. Moreover, the values of visual compacted features must be invariant throughout a shot. This poses a challenge when searching for accurate features that satisfy the previous properties [32].
2- One of the essential properties of high TVS performance is the fast computation [10]. Increasing the speed of TVS computation becomes a bottleneck for application in real-time systems [1]. Detection accuracy can be improved by

extracting more features, however, this needs additional processing on every video frames feature which increase the computational time [19]. Besides, the efficiency of TVS algorithm is increased by local features extraction which is more tolerant of illumination changes and small movements than global features [19]. However, this lead to increase in the computational cost of TVS algorithm. Essentially, small block size increases computational cost and large block size lead to poor representation [25]. Therefore, accelerating the computation of shot boundaries detection must be studied and improved extensively.

3- The detection accuracy of different shot transition types is fundamental in determining the robustness of TVS algorithm. The TVS algorithm robustness depends essentially on the ability of the algorithm to distinguish between disturbance factors within shot and transition between shots which is still an open issue [32]. In other words, the robustness of transition detection stills a challenging task in TVS algorithms due to some disturbances caused by rapid movements and advances in video editing technologies [19]. Specifically, accurate identification of the duration and points of ST in a video scene are significant challenges for the researchers [33]. Moreover, robust TVS algorithm needs to treat all frames to determine the occurrence of shot boundary exactly between consecutive video shots; that is considered time-consuming. Many studies make an attempt to increase the robustness of HT and ST detection, however, this issue remains a controversial topic.

1.3 Aim and Objectives

To fulfill the maximum efficiency of TVS trends, the performance of the proposed algorithm should be specified, such that transitions will be detected perfectly. However, in any transition detection process, TVS should be able to detect shot transitions between two consecutive shots by minimizing false alarm signals (FAMs) and miss detects (MSDs) with low computational cost. Therefore, this research aims to obtain an optimum detection process by increasing detection accuracy and reducing computational load without sacrificing the quality of the detection performance for any types of shot transition. Moreover, managing the disturbance factors of video shot in various scenarios is vital. Consequently, the objectives of this thesis consist of the following specific points:

1- To investigate the use of a new discrete transform that displays a robust energy compaction (EC) and localization properties to extract accurate features that represent the visual content of the frames.
2- To develop methods for feature extraction of the desired video frames that show fast computation, i.e. reducing the computational cost, without degrading other TVS algorithm performances.
3- To design a robust TVS algorithm that improves the detection performance for HTs and STs with a constant quality of detection accuracy and minimizes the computational cost.
4- To evaluate the performance of the proposed TVS algorithm in terms of accuracy and computational cost for different types of transitions.

1.4 Main Contributions of the Thesis

The main contributions of this thesis are presented as follows.

A. Developing a new recurrence algorithm for TP
A new recurrence algorithm to compute the Tchebichef polynomial coefficients (TPCs) values based on the two traditional recurrence algorithms is derived to deal with signals of sizes up to 6144.

B. Developing a new recurrence algorithm for KP
A new recurrence algorithm to compute the Krawtchouk polynomial coefficients (KPCs) values based on a new mathematical model is derived to deal with signals with the largest value of 3440.

C. Developing a new hybrid form of OP
A new hybrid form of OP, called SKTP, is derived from the new orthogonal polynomials (KP and TP). SKTP is used to transform signal into moment domain for information representation. The signal representation exhibits: high EC, localization properties, and dominated signal distribution.

D. Building a new orthogonal polynomial-embedded image kernel method
One of the main contributions of this work is building a new OPEIK method for TVS algorithms that efficiently reduce computation time.

E. Building a new block processing method
This work proposes a new OPBP method that efficiently reduces the computation time for TVS algorithms.

F. Introducing a developed candidate segment technique.
This work proposes a developed candidate segment technique as preprocessing step to reduce the computational cost by discarding static segments.

G. Introducing a new TVS algorithm
This work contributes an accurate TVS algorithm which has low computation cost and can detect efficiently different types of shot boundaries. Moreover, this algorithm used a new technique for representing visual content of video frame termed as frame active area to reduce the effect of persistent and variable visual materials.

1.5 Thesis Scope

The scope of this research focuses on the detection of HT and STs (fade, dissolve, and wipe) based on discrete OP. This thesis is directed toward accurately detecting different transition between shots with low computational cost. Four well-known video datasets, including HTs and STs, are selected to be implemented in this work. These dataset are TRECVID 2001, 2005, 2006, and 2007 [34]. TRECVID was established to evaluate and benchmark TVS tasks [34], and it has contributed to the improvement of TVS algorithms [20]. These different datasets are carefully selected because they include the aforementioned types of transitions with different genres and transition intervals. In addition, these datasets include multiple types of object/camera motion that make the evaluation of TVS algorithm robust. The experiments carried out on a laptop (HP dv6) with 2.20 GHz CPU and 8 GB RAM.

The flow of the proposed study has been based on transform-based approach and local key-point (LKP) in the uncompressed domain for feature extraction. The global and local feature are extracted. The extracted features are performed for a selected video frames to reduce the number of processed frames.

1.6 Thesis Organization

This thesis is divided into five chapters including this chapter. Chapter Two presents a comprehensive review of existing TVS, different approaches of TVS algorithms, various techniques that solve different problems in TVS, theoretical background is also included and discussed in this chapter. Chapter Two ends by highlighting the main gaps in recent research that should be considered when proposing a TVS algorithm with high specifications.

Chapter Three provides a complete description of the research methodology steps. The work flow of this study is divided into multiple sections. The first section presents the derivation of the new discrete transform that facilitates and positively improves the detection process. Then, OPEIK and OPBP are presented. Afterward, in the second section, the proposed TVS with all its stages are demonstrated.

Chapter Four presents the results and discussion of the proposed algorithm. A comparison is performed based on two aspects. In the first aspect, each stages of the proposed algorithm are evaluated individually. In the second aspect, the entire proposed algorithm is evaluated and compared with other state-of-the-art algorithms. Different measurements are used in the comparison assessment. Moreover, experimental tests have been presented to provide comprehensive explanations of the remarkable results and the robustness of the proposed TVS. This thesis ends with a summary and conclusion in Chapter Five. Potential ideas for future work are also suggested.

REFERENCES

[1] M. Birinci and S. Kiranyaz, "A perceptual scheme for fully automatic video shot boundary detection," *Signal Process. Image Commun.*, vol. 29, no. 3, pp. 410–423, 2014.

[2] C. Liu, D. Wang, J. Zhu, and B. Zhang, "Learning a Contextual Multi-Thread Model for Movie/TV Scene Segmentation," *IEEE Trans. Multimed.*, vol. 15, no. 4, pp. 884–897, Jun. 2013.

[3] 2016. [Online]. Available: http://www.alexa.com/topsites.

[4] M. Del Fabro and L. Böszörmenyi, "State-of-the-art and future challenges in video scene detection: a survey," *Multimed. Syst.*, vol. 19, no. 5, pp. 427–454, 2013.

[5] I. Gonzalez-Diaz, T. Martinez-Cortes, A. Gallardo-Antolin, and F. Diaz-de-Maria, "Temporal segmentation and keyframe selection methods for user-generated video search-based annotation," *Expert Syst. Appl.*, vol. 42, no. 1, pp. 488–502, 2015.

[6] M. B. Fayk, H. A. El Nemr, and M. M. Moussa, "Particle swarm optimisation based video abstraction," *J. Adv. Res.*, vol. 1, no. 2, pp. 163–167, 2010.

[7] W. Hu, N. Xie, L. Li, X. Zeng, and S. Maybank, "A Survey on Visual Content-Based Video Indexing and Retrieval," *IEEE Trans. Syst. Man, Cybern. Part C (Applications Rev.*, vol. 41, no. 6, pp. 797–819, 2011.

[8] R. Priya and T. N. Shanmugam, "A comprehensive review of significant researches on content based indexing and retrieval of visual information," *Front. Comput. Sci.*, vol. 7, no. 5, pp. 782–799, 2013.

[9] K. Choroś, "Improved Video Scene Detection Using Player Detection Methods in Temporally Aggregated TV Sports News," in *International Conference on Computational Collective Intelligence*, pp. 633–643, 2014.

[10] H. Bhaumik, S. Bhattacharyya, M. Das Nath, and S. Chakraborty, "Hybrid soft computing approaches to content based video retrieval: A brief review," *Appl. Soft Comput.*, vol. 46, pp. 1008–1029, 2016.

[11] G. C. Chaves, "Video Content Analysis by Active Learning," PhD Thesis, Federal University of Minas Gerais, 2007.

[12] N. J. Janwe and K. K. Bhoyar, "Video shot boundary detection based on JND color histogram," in *Image Information Processing (ICIIP), 2013 IEEE Second International Conference on*, pp. 476–480, 2013.

[13] U. Gargi, R. Kasturi, and S. H. Strayer, "Performance Characterization of Video-Shot-Change Detection Methods," *IEEE Trans. Circuits Syst.*, vol. 8, no. 10, pp. 4761–4766, 2000.

[14] Z.-M. Lu and Y. Shi, "Fast video shot boundary detection based on SVD and pattern matching," *Image Process. IEEE Trans.*, vol. 22, no. 12, pp. 5136–5145, 2013.

[15] K. Choroś, "False and miss detections in temporal segmentation of TV sports

news videos–causes and remedies," in *New Research in Multimedia and Internet Systems*, pp. 35–46, 2015.

[16] L. H. Iwan and J. A. Thom, "Temporal video segmentation: detecting the end-of-act in circus performance videos," *Multimed. Tools Appl.*, vol. 76, no. 1, pp. 1379–1401, 2017.

[17] G. Gao and H. Ma, "To accelerate shot boundary detection by reducing detection region and scope," *Multimed. Tools Appl.*, vol. 71, no. 3, pp. 1749–1770, 2014.

[18] M. Tavassolipour, M. Karimian, and S. Kasaei, "Event Detection and Summarization in Soccer Videos Using Bayesian Network and Copula," *Ieee Trans. Circuits Syst. Video Technol.*, vol. 24, no. 2, pp. 291–304, 2014.

[19] S. Tippaya, S. Sitjongsataporn, T. Tan, M. M. Khan, and K. Chamnongthai, "Multi-Modal Visual Features-Based Video Shot Boundary Detection," *IEEE Access*, vol. 5, pp. 12563–12575, 2017.

[20] A. Amiri and M. Fathy, "Video shot boundary detection using QR-decomposition and gaussian transition detection," *EURASIP J. Adv. Signal Process.*, vol. 2009, no. 1, pp. 1–12, 2010.

[21] C. W. Ngo, T. C. Pong, and R. T. Chin, "Video partitioning by temporal slice coherency," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 8, pp. 941–953, 2001.

[22] M. Tkalcic and J. F. Tasic, "Colour spaces: perceptual, historical and applicational background," in *The IEEE Region 8 EUROCON 2003. Computer as a Tool.*, vol. 1, pp. 304–308 vol.1, 2003.

[23] R. Zabih, J. Miller, and K. Mai, "A feature-based algorithm for detecting and classifying production effects," *Multimed. Syst.*, vol. 7, no. 2, pp. 119–128, 1999.

[24] L. G. G. Priya and S. Domnic, "Walsh – Hadamard Transform Kernel-Based Feature Vector for Shot Boundary Detection," *IEEE Trans. IMAGE Process.*, vol. 23, no. 12, pp. 5187–5197, 2014.

[25] J. Mondal, M. K. Kundu, S. Das, and M. Chowdhury, "Video shot boundary detection using multiscale geometric analysis of nsct and least squares support vector machine," *Multimed. Tools Appl.*, vol. 77, no. 7, pp. 8139–8161, Apr. 2017.

[26] D. Dutta, S. K. Saha, and B. Chanda, "A shot detection technique using linear regression of shot transition pattern," *Multimed. Tools Appl.*, vol. 75, no. 1, pp. 93–113, 2016.

[27] G. Pal, D. Rudrapaul, S. Acharjee, R. Ray, S. Chakraborty, and N. Dey, "Video Shot Boundary Detection: A Review," in *Emerging ICT for Bridging the Future - Proceedings of the 49th Annual Convention of the Computer Society of India CSI Volume 2*, pp. 119–127, 2015.

[28] A. Ahmed, "Video Representation and Processing for Multimedia Data Mining," in *Semantic Mining Technologies for Multimedia Databases*, IGI Global, 2009, pp. 1–31.

[29] A. SenGupta, D. M. Thounaojam, K. M. Singh, and S. Roy, “Video shot boundary detection: A review,” in *Electrical, Computer and Communication Technologies (ICECCT), 2015 IEEE International Conference on*, pp. 1–6, 2015.

[30] C. Cotsaces, N. Nikolaidis, and I. Pitas, “Video shot boundary detection and condensed representation: a review,” *IEEE Signal Process. Mag.*, vol. 23, no. 2, pp. 28–37, 2006.

[31] W. Zheng, J. Yuan, H. Wang, F. Lin, and B. Zhang, “A novel shot boundary detection framework,” in *Visual Communications and Image Processing 2005*, p. 596018, 2005.

[32] A. Amiri and M. Fathy, “Video Shot Boundary Detection Using Generalized Eigenvalue Decomposition And Gaussian Transition Detection,” *Comput. Informatics*, vol. 30, pp. 595–619, 2011.

[33] H. Bhaumik, M. Chakraborty, S. Bhattacharyya, and S. Chakraborty, “Detection of Gradual Transition in Videos: Approaches and Applications,” in *Intelligent Analysis of Multimedia Information*, no. December, 2017, pp. 282–318.

[34] TRECVID, 2018. [Online]. Available: http://trecvid.nist.gov/.

[35] J. Yuan *et al.*, “A formal study of shot boundary detection,” *Circuits Syst. Video Technol. IEEE Trans.*, vol. 17, no. 2, pp. 168–186, 2007.

[36] M. N. Asghar, F. Hussain, and R. Manton, “Video indexing: a survey,” *Int. J. Comput. Inf. Technol.*, vol. 3, no. 01, pp. 148–169, 2014.

[37] L.-Y. Duan, M. Xu, Q. Tian, C.-S. Xu, and J. S. Jin, “A unified framework for semantic shot classification in sports video,” *IEEE Trans. Multimed.*, vol. 7, no. 6, pp. 1066–1083, Dec. 2005.

[38] W. Ren and M. Sharma, “Automated video segmentation,” in *International Conference on Information, Communication, and Signal Processing*, pp. 1–11, 2001.

[39] Y. Chen, Y. Deng, Y. Guo, W. Wang, Y. Zou, and K. Wang, “A Temporal Video Segmentation and Summary Generation Method Based on Shots’ Abrupt and Gradual Transition Boundary Detecting,” in *Communication Software and Networks, 2010. ICCSN ’10. Second International Conference on*, pp. 271–275, 2010.

[40] W. Tong, L. Song, X. Yang, H. Qu, R. Xie, and Ieee, “CNN-Based Shot Boundary Detection and Video Annotation,” in *2015 Ieee International Symposium on Broadband Multimedia Systems and Broadcasting*, 2015.

[41] A. Kowdle and T. Chen, “Learning to Segment a Video to Clips Based on Scene and Camera Motion,” in *Computer Vision-ECCV 2012*, pp. 272–286, 2012.

[42] J. Bescós, G. Cisneros, J. M. Martínez, J. M. Menéndez, and J. Cabrera, “A unified model for techniques on video-shot transition detection,” *Multimedia, IEEE Trans.*, vol. 7, no. 2, pp. 293–307, 2005.

[43] R. W. Lienhart, “Comparison of automatic shot boundary detection

algorithms," *Proc. SPIE Storage Retr. Image Video Databases VII*, vol. 3656, no. SPIE 3656, pp. 290–301, 1998.

[44] Y. N. Li, Z. M. Lu, and X. M. Niu, "Fast video shot boundary detection framework employing pre-processing techniques," *IET Image Process.*, vol. 3, no. 3, pp. 121–134, 2009.

[45] J. Baber, N. Afzulpurkar, and S. Satoh, "A framework for video segmentation using global and local features," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 27, no. 05, p. 1355007_1-1355007_29, Aug. 2013.

[46] A. Hanjalic, "Shot-boundary detection: Unraveled and resolved?," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 2, pp. 90–105, 2002.

[47] X. Jiang, T. Sun, J. Liu, J. Chao, and W. Zhang, "An adaptive video shot segmentation scheme based on dual-detection model," *Neurocomputing*, vol. 116, pp. 102–111, 2013.

[48] J. Cao and A. Cai, "A robust shot transition detection method based on support vector machine in compressed domain," *Pattern Recognit. Lett.*, vol. 28, no. 12, pp. 1534–1540, 2007.

[49] X. Ling, O. Yuanxin, L. Huan, and X. Zhang, "A Method for Fast Shot Boundary Detection Based on SVM," in *Image and Signal Processing, 2008. CISP '08. Congress on*, vol. 2, pp. 445–449, 2008.

[50] H. Fang, J. Jiang, and Y. Feng, "A fuzzy logic approach for detection of video shot boundaries," *Pattern Recognit.*, vol. 39, no. 11, pp. 2092–2100, 2006.

[51] K. Choroś and M. Gonet, "Effectiveness of video segmentation techniques for different categories of videos," *New Trends Multimed. Netw. Inf. Syst.*, vol. 181, p. 34, 2008.

[52] K. Choroś, "Reduction of faulty detected shot cuts and cross dissolve effects in video segmentation process of different categories of digital videos," in *Transactions on Computational Collective Intelligence V*, pp. 124–139, 2011.

[53] Z. Černeková, C. Kotropoulos, and I. Pitas, "Video shot-boundary detection using singular-value decomposition and statistical tests," *J. Electron. Imaging*, vol. 16, no. 4, pp. 43012–43013, 2007.

[54] R. A. Joyce and B. Liu, "Temporal segmentation of video using frame and histogram space," *IEEE Trans. Multimed.*, vol. 8, no. 1, pp. 130–140, Feb. 2006.

[55] Z. Černeková, I. Pitas, and C. Nikou, "Information theory-based shot cut/fade detection and video summarization," *Circuits Syst. Video Technol. IEEE Trans.*, vol. 16, no. 1, pp. 82–91, 2006.

[56] S. V. Porter, "Video segmentation and indexing using motion estimation," PhD Thesis, University of Bristol, 2004.

[57] T. Barbu, "Novel automatic video cut detection technique using Gabor filtering," *Comput. Electr. Eng.*, vol. 35, no. 5, pp. 712–721, 2009.

[58] Y. Kawai, H. Sumiyoshi, and N. Yagi, "Shot boundary detection at TRECVID 2007," in *TRECVID*, 2007.

[59] A. Hameed, “A novel framework of shot boundary detection for uncompressed videos,” in *Emerging Technologies, 2009. ICET 2009. International Conference on*, pp. 274–279, 2009.

[60] B. H. Shekar and K. P. Uma, “Kirsch Directional Derivatives Based Shot Boundary Detection: An Efficient and Accurate Method,” in *Second International Symposium on Computer Vision and the Internet (VisionNet’15), Procedia Computer Science*, vol. 58, pp. 565–571, 2015.

[61] P. Over, T. Ianeva, W. Kraaij, A. F. Smeaton, and U. De Val, “TRECVID 2005 - An Overview,” *NIST*, pp. 1–27, 2005.

[62] M. Cooper, T. Liu, and E. Rieffel, “Video Segmentation via Temporal Pattern Classification,” *IEEE Trans. Multimed.*, vol. 9, no. 3, pp. 610–618, 2007.

[63] S. Lefèvre and N. Vincent, “Efficient and robust shot change detection,” *J. Real-Time Image Process.*, vol. 2, no. 1, pp. 23–34, 2007.

[64] C. Grana and R. Cucchiara, “Linear transition detection as a unified shot detection approach,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 4, pp. 483–489, 2007.

[65] S. Aryal, K. M. Ting, T. Washio, and G. Haffari, “Data-dependent dissimilarity measure: an effective alternative to geometric distance measures,” *Knowl. Inf. Syst.*, pp. 1–28, 2017.

[66] D.-D. Le, S. Satoh, T. D. Ngo, and D. A. Duong, “A text segmentation based approach to video shot boundary detection,” in *2008 IEEE 10th Workshop on Multimedia Signal Processing*, pp. 702–706, 2008.

[67] G. Camara-Chavez, F. Precioso, M. Cord, S. Phillip-Foliguet, and A. de A. Araujo, “Shot Boundary Detection by a Hierarchical Supervised Approach,” in *2007 14th International Workshop on Systems, Signals and Image Processing and 6th EURASIP Conference focused on Speech and Image Processing, Multimedia Communications and Services*, pp. 197–200, 2007.

[68] F. Pacheco, M. Cerrada, R.-V. Sánchez, D. Cabrera, C. Li, and J. V. de Oliveira, “Attribute clustering using rough set theory for feature selection in fault severity classification of rotating machinery,” *Expert Syst. Appl.*, vol. 71, pp. 69–86, 2017.

[69] M.-S. Lee, Y.-M. Yang, and S.-W. Lee, “Automatic video parsing using shot boundary detection and camera operation analysis,” *Pattern Recognit.*, vol. 34, no. 3, pp. 711–719, 2001.

[70] T. Kikukawa and S. Kawafuchi, “Development of an automatic summary editing system for the audio-visual resources,” *Trans. Inst. Electron., Inform., Commun. Eng*, vol. 75, pp. 204–212, 1992.

[71] A. Nagasaka and Y. Tanaka, “Automatic video indexing and full-video search for object appearances,” in *Visual Database Systems II*, pp. 113–127, 1992.

[72] H. Zhang, A. Kankanhalli, and S. W. Smoliar, “Automatic partitioning of full-motion video,” *Multimed. Syst.*, vol. 1, no. 1, pp. 10–28, 1993.

[73] B. Shahraray, “Scene change detection and content-based sampling of video sequences,” in *IS&T/SPIE’s Symposium on Electronic Imaging: Science &*

Technology, pp. 2–13, 1995.

[74] G. Lupatini, C. Saraceno, and R. Leonardi, “Scene break detection: a comparison,” in *Research Issues In Data Engineering, 1998.’Continuous-Media Databases and Applications’. Proceedings., Eighth International Workshop on*, pp. 34–41, 1998.

[75] R. W. Lienhart, “Reliable transition detection in videos: A survey and practitioner’s guide,” *Int. J. Image Graph.*, vol. 1, no. 3, pp. 469–486, 2001.

[76] M. G. Chung, H. Kim, and S. M. H. Song, “A scene boundary detection method,” in *Image Processing, 2000. Proceedings. 2000 International Conference on*, vol. 3, pp. 933–936 vol.3, 2000.

[77] A. M. Ferman and A. M. Tekalp, “Efficient filtering and clustering methods for temporal video segmentation and visual summarization,” *J. Vis. Commun. Image Represent.*, vol. 9, no. 4, pp. 336–351, 1998.

[78] J. Meng, Y. Juan, and S.-F. Chang, “Scene Change Detection in a MPEG compressed Video Sequence,” *Proc. IS&T/SPIE Int. Symp. Electron. Imaging*, vol. 2417, no. February, pp. 14–25, 1995.

[79] H. Fang, Y. Yin, P. Norhashimah, and J. Jiang, “A hybrid scheme for temporal video segmentation,” in *Third IEEE International Workshop on Electronic Design, Test and Applications (DELTA’06)*, p. 6 pp.-, 2006.

[80] Z. H. Z. Huan, L. X. L. Xiuhuan, and Y. L. Y. Lilei, “Shot Boundary Detection Based on Mutual Information and Canny Edge Detector,” *2008 Int. Conf. Comput. Sci. Softw. Eng.*, vol. 2, pp. 1124–1128, 2008.

[81] I. Koprinska and S. Carrato, “Temporal video segmentation: A survey,” *Signal Process. Image Commun.*, vol. 16, no. 5, pp. 477–500, 2001.

[82] R. Tapu and T. Zaharia, “Video Segmentation and Structuring for Indexing Applications,” *Int. J. Multimed. Data Eng. Manag.*, vol. 2, no. 4, pp. 38–58, Oct. 2011.

[83] G. Ciocca and R. Schettini, “Dynamic storyboards for video content summarization,” *Proc. 8th ACM Int. Work. Multimed. Inf. Retr. - MIR ’06*, p. 259, 2006.

[84] S. V Porter, M. Mirmehdi, and B. T. Thomas, “Video cut detection using frequency domain correlation,” in *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, vol. 3, pp. 409–412, 2000.

[85] J. S. Boreczky and L. a Rowe, “Comparison of video shot boundary detection techniques,” *J. Electron. Imaging*, vol. 5, no. April, pp. 122–128, 1996.

[86] J. Mas and G. Fernandez, “Video shot boundary detection based on color histogram,” *Noteb. Pap. TRECVID2003, Gaithersburg, Maryland, NIST*, 2003.

[87] M. Verma and B. Raman, “A Hierarchical Shot Boundary Detection Algorithm Using Global and Local Features,” in *Proceedings of International Conference on Computer Vision and Image Processing: CVIP 2016, Volume 2*, pp. 389–397, 2017.

[88] M. Ahmed, A. Karmouch, and S. Abu-Hakima, “Key frame extraction and

indexing for multimedia databases," in *Vision Interface '99*, vol. 99, pp. 506–511, 1999.

[89] M. Ahmed and A. Karmouch, "Video segmentation using an opportunistic approach," *Multimed. Model. Ottawa, Canada*, pp. 389–405, 1999.

[90] D. M. Thounaojam, T. Khelchandra, K. M. Singh, and S. Roy, "A Genetic Algorithm and Fuzzy Logic Approach for Video Shot Boundary Detection," *Comput. Intell. Neurosci.*, vol. 2016, no. Article ID 8469428, p. 14, 2016.

[91] X. Qian, G. Liu, and R. Su, "Effective Fades and Flashlight Detection Based on Accumulating Histogram Difference," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 10, pp. 1245–1258, Oct. 2006.

[92] Q. G. Ji, J. W. Feng, J. Zhao, and Z. M. Lu, "Effective Dissolve Detection Based on Accumulating Histogram Difference and the Support Point," in *Pervasive Computing Signal Processing and Applications (PCSPA), 2010 First International Conference on*, pp. 273–276, 2010.

[93] O. Küçüktunç, U. Güdükbay, and Ö. Ulusoy, "Fuzzy color histogram-based video segmentation," *Comput. Vis. Image Underst.*, vol. 114, no. 1, pp. 125–134, 2010.

[94] A. Adnan and M. Ali, "Shot boundary detection using sorted color histogram polynomial curve," *Life Sci. J.*, vol. 10, no. 4, pp. 1965–1972, 2013.

[95] Z. Li, X. Liu, and S. Zhang, "Shot Boundary Detection based on Multilevel Difference of Colour Histograms," in *2016 First International Conference on Multimedia and Image Processing (ICMIP)*, pp. 15–22, 2016.

[96] R. Dadashi and H. R. Kanan, "AVCD-FRA:A novel solution to automatic video cut detection using fuzzy-rule-based approach," *Comput. Vis. Image Underst.*, vol. 117, no. 7, pp. 807–817, 2013.

[97] L. Krulikovska, J. Pavlovic, J. Polec, and Z. Cernekova, "Abrupt cut detection based on mutual information and motion prediction," in *ELMAR, 2010 PROCEEDINGS*, pp. 89–92, 2010.

[98] T.-Y. Shih, "The reversibility of six geometric color spaces," *Photogramm. Eng. Remote Sensing*, vol. 61, no. 10, pp. 1223–1232, 1995.

[99] S. Park, J. Son, and S. J. Kim, "Study on the effect of frame size and color histogram bins on the shot boundary detection performance," in *2016 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia)*, pp. 1–2, 2016.

[100] D. Pye, N. J. Hollinghurst, T. J. Mills, and K. R. Wood, "Audio-visual segmentation for content-based retrieval.," in *The International Conference on Spoken Language Processing (ICSLP'98)*, 1998.

[101] R. Zabih, J. Miller, and K. Mai, "A Feature-Based Algorithm for Detecting and Classifying Scene Breaks," *Proc. third ACM Int. Conf. Multimed. Multimed. 95*, vol. 95, pp. 189–200, 1995.

[102] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, no. 6, pp. 679–698, 1986.

[103] J. Nam and A. H. Tewfik, “Combined audio and visual streams analysis for video sequence segmentation,” in *Acoustics, Speech, and Signal Processing, 1997. ICASSP-97., 1997 IEEE International Conference on*, vol. 4, pp. 2665–2668, 1997.

[104] R. W. Lienhart, “Reliable dissolve detection,” in *Photonics West 2001-Electronic Imaging*, pp. 219–230, 2001.

[105] W. J. Heng and K. N. Ngan, “Integrated shot boundary detection using object-based technique,” in *Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on*, vol. 3, pp. 289–293, 1999.

[106] W. J. Heng and K. N. Ngan, “An Object-Based Shot Boundary Detection Using Edge Tracing and Tracking,” *J. Vis. Commun. Image Represent.*, vol. 12, no. 3, pp. 217–239, 2001.

[107] J. Zheng, F. Zou, and M. Shi, “An efficient algorithm for video shot boundary detection,” in *Intelligent Multimedia, Video and Speech Processing, 2004. Proceedings of 2004 International Symposium on*, pp. 266–269, 2004.

[108] L. G. Roberts, “Machine perception of three-dimensional soups,” PhD Thesis, Massachusetts Institute of Technology, 1963.

[109] H.-W. Yoo, H.-J. Ryoo, and D.-S. Jang, “Gradual shot boundary detection using localized edge blocks,” *Multimed. Tools Appl.*, vol. 28, no. 3, pp. 283–300, 2006.

[110] I. Sobel and G. Feldman, “A 3x3 isotropic gradient operator for image processing,” *a talk Stanford Artif. Proj.*, pp. 271–272, 1968.

[111] C. Chan and A. Wong, “Shot Boundary Detection Using Genetic Algorithm Optimization,” in *Multimedia (ISM), 2011 IEEE International Symposium on*, pp. 327–332, 2011.

[112] A. Dailianas, R. B. Allen, and P. England, “Comparison of automatic video segmentation algorithms,” in *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 2615, 1996, pp. 2–16.

[113] W. J. Heng and K. N. Ngan, “High accuracy flashlight scene determination for shot boundary detection,” *Signal Process. Image Commun.*, vol. 18, no. 3, pp. 203–219, 2003.

[114] S. H. Kim and R.-H. Park, “Robust video indexing for video sequences with complex brightness variations,” in *Proceedings of lnt. Conf. Signal and Image Processing*, pp. 410–414, 2002.

[115] R. Mukundan, P. Raveendran, and W. A. Jassim, “New orthogonal polynomials for speech signal and image processing,” *IET Signal Process.*, vol. 6, no. 8, pp. 713–723, Oct. 2012.

[116] B. M. Mahmmod, A. R. bin Ramli, S. H. Abdulhussain, S. A. R. Al-Haddad, and W. A. Jassim, “Signal compression and enhancement using a new orthogonal-polynomial-based discrete transform,” *IET Signal Process.*, vol. 12, no. 1, pp. 129–142, Aug. 2018.

[117] G. L. Priya and S. Domnic, “Edge Strength Extraction using Orthogonal Vectors for Shot Boundary Detection,” *Procedia Technol.*, vol. 6, pp. 247–254,

2012.

[118] T. Vlachos, “Cut detection in video sequences using phase correlation,” *IEEE Signal Process. Lett.*, vol. 7, no. 7, pp. 173–175, 2000.

[119] O. Urhan, M. K. Gullu, and S. Erturk, “Modified phase-correlation based robust hard-cut detection with application to archive film,” *IEEE Trans. circuits Syst. video Technol.*, vol. 16, no. 6, pp. 753–770, 2006.

[120] S. Porter, M. Mirmehdi, and B. Thomas, “Temporal video segmentation and classification of edit effects,” *Image Vis. Comput.*, vol. 21, no. 13, pp. 1097–1106, 2003.

[121] A. Barjatya, “Block matching algorithms for motion estimation,” *IEEE Trans. Evol. Comput.*, vol. 8, no. 3, pp. 225–239, 2004.

[122] F. Dufaux and J. Konrad, “Efficient, robust, and fast global motion estimation for video coding,” *IEEE Trans. image Process.*, vol. 9, no. 3, pp. 497–501, 2000.

[123] E. Bruno and D. Pellerin, “Video shot detection based on linear prediction of motion,” in *Multimedia and Expo, 2002. ICME’02. Proceedings. 2002 IEEE International Conference on*, vol. 1, pp. 289–292, 2002.

[124] I. A. Zedan, K. M. Elsayed, and E. Emary, “Abrupt Cut Detection in News Videos Using Dominant Colors Representation,” in *Proceedings of the International Conference on Advanced Intelligent Systems and Informatics 2016*, A. E. Hassanien, K. Shaalan, T. Gaber, A. T. Azar, and M. F. Tolba, Eds. Cham: Springer International Publishing, 2016, pp. 320–331.

[125] R. Jain, R. Kasturi, and B. G. Schunck, “Dynamic vision,” in *Computer Vision: Principles*, IEEE Computer Society Press, 1991, pp. 469–480.

[126] A. M. Alattar, “Detecting And Compressing Dissolve Regions In Video Sequences With A DVI Multimedia Image Compression Algorithm,” in *Circuits and Systems, 1993., ISCAS ’93, 1993 IEEE International Symposium on*, pp. 13–16, 1993.

[127] B. T. Truong, C. Dorai, and S. Venkatesh, “New Enhancements to Cut, Fade, and Dissolve Detection Processes in Video Segmentation,” in *Proceedings of the Eighth ACM International Conference on Multimedia*, pp. 219–227, 2000.

[128] J. Z. Miadowicz, “Story tracking in video news broadcasts,” PhD Thesis, University of Kansas, 2004.

[129] E. Ribnick, S. Atev, O. Masoud, N. Papanikolopoulos, and R. Voyles, “Real-time detection of camera tampering,” in *Video and Signal Based Surveillance, 2006. AVSS’06. IEEE International Conference on*, p. 10, 2006.

[130] H. Bay, T. Tuytelaars, and L. Van Gool, “SURF: Speeded up robust features,” in *European conference on computer vision*, pp. 404–417, 2006.

[131] D. G. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints,” *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.

[132] C. Harris and M. Stephens, “A combined corner and edge detector,” in *In Proc. of Fourth Alvey Vision Conference*, pp. 147–151, 1988.

[133] M.-H. Park, R.-H. Park, and S. W. Lee, “Shot boundary detection using scale invariant feature matching,” *Proc. SPIE*, vol. 6077. p. 60771N–60771N–9, 2006.

[134] C. R. Huang, H. P. Lee, and C. S. Chen, “Shot Change Detection via Local Keypoint Matching,” *IEEE Trans. Multimed.*, vol. 10, no. 6, pp. 1097–1108, Oct. 2008.

[135] M. Birinci, S. Kiranyaz, and M. Gabbouj, “Video shot boundary detection by structural analysis of local image features,” in *WIAMIS 2011: 12th International Workshop on Image Analysis for Multimedia Interactive Services, Delft, The Netherlands, April 13-15, 2011*, 2011.

[136] G. Liu, X. Wen, W. Zheng, and P. He, “Shot Boundary Detection and Keyframe Extraction Based on Scale Invariant Feature Transform,” in *Computer and Information Science, 2009. ICIS 2009. Eighth IEEE/ACIS International Conference on*, pp. 1126–1130, 2009.

[137] B. Youssef, E. Fedwa, A. Driss, and S. Ahmed, “Shot boundary detection via adaptive low rank and svd-updating,” *Comput. Vis. Image Underst.*, vol. 161, no. Supplement C, pp. 20–28, 2017.

[138] J. Xu, L. Song, and R. Xie, “Shot boundary detection using convolutional neural networks,” in *2016 Visual Communications and Image Processing (VCIP)*, pp. 1–4, 2016.

[139] S. Lian, “Automatic video temporal segmentation based on multiple features,” *Soft Comput.*, vol. 15, no. 3, pp. 469–482, 2011.

[140] P. P. Mohanta, S. K. Saha, and B. Chanda, “A model-based shot boundary detection technique using frame transition parameters,” *IEEE Trans. Multimed.*, vol. 14, no. 1, pp. 223–233, 2012.

[141] J. Lankinen and J.-K. Kämäräinen, “Video Shot Boundary Detection using Visual Bag-of-Words,” in *VISAPP (1)*, pp. 778–791, 2013.

[142] ImageNet, [Online]. Available: http://www.image-net.org/.

[143] P. S. A. Bhalotra and B. D. Patil, “Shot boundary detection using radon projection method,” *Int. J. Signal Image Process.*, vol. 4, no. 3, p. 60, 2013.

[144] M. Parmar and M. C. Angelides, “MAC-REALM: A Video Content Feature Extraction and Modelling Framework,” *Comput. J.*, vol. 58, no. 9, pp. 2135–2170, 2015.

[145] J. Chen, J. Ren, and J. Jiang, “Modelling of content-aware indicators for effective determination of shot boundaries in compressed MPEG videos,” *Multimed. Tools Appl.*, vol. 54, no. 2, pp. 219–239, 2011.

[146] S. J. F. Guimaraes, Z. K. G. do Patrocinio, K. J. F. Souza, and H. B. de Paula, “Gradual transition detection based on bipartite graph matching approach,” in *Multimedia Signal Processing, 2009. MMSP’09. IEEE International Workshop on*, pp. 1–6, 2009.

[147] M. Tsujitani and Y. Tanaka, “Cross-validation, bootstrap, and support vector machines,” *Adv. Artif. Neural Syst.*, vol. 2011, 2011.

[148] Y. Zhang and L. Wu, "Classification of fruits using computer vision and a multiclass support vector machine," *sensors*, vol. 12, no. 9, pp. 12489–12505, 2012.

[149] M. Achirul Nanda, K. Boro Seminar, D. Nandika, and A. Maddu, "A Comparison Study of Kernel Functions in the Support Vector Machine and Its Application for Termite Detection," *Information*, vol. 9, no. 1, p. 5, 2018.

[150] W. A. Jassim, "A New Discrete Orthogonal Function Based on Tchebichef and Krawtchouk Polynomials and Its Applications to Speech and Image Analysis," Universiti Malaya, 2012.

[151] D. Wilk-Kolodziejczyk, K. Regulski, and G. Gumienny, "Comparative analysis of the properties of the nodular cast iron with carbides and the austempered ductile iron with use of the machine learning and the support vector machine," *Int. J. Adv. Manuf. Technol.*, vol. 87, no. 1–4, pp. 1077–1093, 2016.

[152] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Min. Knowl. Discov.*, vol. 2, no. 2, pp. 121–167, 1998.

[153] H.-T. Lin and C.-J. Lin, "A study on sigmoid kernels for SVM and the training of non-PSD kernels by SMO-type methods," *Submitt. to Neural Comput.*, vol. 3, pp. 1–32, 2003.

[154] C.-W. Hsu, C.-C. Chang, C.-J. Lin, and others, "A practical guide to support vector classification," *Dep. Comput. Sci. Inf. Eng.*, 2003.

[155] D. G. Lowe, "Object recognition from local scale-invariant features," in *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, vol. 2, pp. 1150–1157, 1999.

[156] A. P. Winkin, "Scale-Space Filtering," in *International Joint Conference on Artificial Intelligence*, pp. 1019–1022, 1983.

[157] T. Lindeberg, "Scale-space theory: A basic tool for analyzing structures at different scales," *J. Appl. Stat.*, vol. 21, no. 1–2, pp. 225–270, 1994.

[158] P. Toharia, O. D. Robles, R. Suarez, J. Luis Bosque, and L. Pastor, "Shot boundary detection using Zernike moments in multi-GPU multi-CPU architectures," *J. Parallel Distrib. Comput.*, vol. 72, no. 9, pp. 1127–1133, 2012.

[159] Y. S. Abu-Mostafa and D. Psaltis, "Recognitive Aspects of Moment Invariants," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 698–706, Nov. 1984.

[160] Ming-Kuei Hu, "Visual pattern recognition by moment invariants," *IEEE Trans. Inf. Theory*, vol. 8, no. 2, pp. 179–187, Feb. 1962.

[161] C.-H. Teh and R. T. Chin, "On image analysis by the methods of moments," *Pattern Anal. Mach. Intell. IEEE Trans.*, vol. 10, no. 4, pp. 496–513, 1988.

[162] C.-W. Chong, P. Raveendran, and R. Mukundan, "Translation and scale invariants of Legendre moments," *Pattern Recognit.*, vol. 37, no. 1, pp. 119–129, Jan. 2004.

[163] A. Khotanzad and Y. H. Hong, "Invariant image recognition by Zernike moments," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 5, pp. 489–497, May 1990.

[164] Y. Sheng and L. Shen, "Orthogonal Fourier–Mellin moments for invariant pattern recognition," *J. Opt. Soc. Am. A*, vol. 11, no. 6, p. 1748, Jun. 1994.

[165] R. Mukundan, S. H. Ong, and P. A. Lee, "Image analysis by Tchebichef moments," *IEEE Trans. Image Process.*, vol. 10, no. 9, pp. 1357–1364, 2001.

[166] Pew-Thian Yap, R. Paramesran, and Seng-Huat Ong, "Image analysis by krawtchouk moments," *IEEE Trans. Image Process.*, vol. 12, no. 11, pp. 1367–1377, Nov. 2003.

[167] H. Zhu, M. Liu, H. Shu, H. Zhang, and L. Luo, "General form for obtaining discrete orthogonal moments," *IET Image Process.*, vol. 4, no. 5, p. 335, 2010.

[168] R. Koekoek, P. A. Lesky, and R. F. Swarttouw, *Hypergeometric orthogonal polynomials and their q-analogues*. Springer Science & Business Media, 2010.

[169] J. J. Foncannon, "Irresistible integrals: symbolics, analysis and experiments in the evaluation of integrals," *Math. Intell.*, vol. 28, no. 3, pp. 65–68, 2006.

[170] F. Jan, S. Tomáš, and Z. Barbara, "Moments and Moment Invariants in Pattern Recognition," *Chippenham, UK Wiley Sons Ltd*, 2009.

[171] K. Nakagaki and R. Mukundan, "A Fast 4 x 4 Forward Discrete Tchebichef Transform Algorithm," *IEEE Signal Process. Lett.*, vol. 14, no. 10, pp. 684–687, Oct. 2007.

[172] R. Mukundan, "Some Computational Aspects of Discrete Orthonormal Moments," *IEEE Trans. Image Process.*, vol. 13, no. 8, pp. 1055–1059, Aug. 2004.

[173] A. F. Nikiforov, V. B. Uvarov, and S. K. Suslov, "Classical orthogonal polynomials of a discrete variable," in *Classical Orthogonal Polynomials of a Discrete Variable*, Springer, 1991, pp. 18–54.

[174] M. E. H. Ismail, "Classical and quantum orthogonal polynomials in one variable Cambridge Univ," *Pres, Cambridge, UK*, 2005.

[175] G. Zhang *et al.*, "A symmetry and bi-recursive algorithm of accurately computing Krawtchouk moments," *Pattern Recognit. Lett.*, vol. 31, no. 7, pp. 548–554, May 2010.

[176] S. Park, J. Son, and S. J. Kim, "Effect of adaptive thresholding on shot boundary detection performance," in *2016 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia)*, pp. 1–2, 2016.

[177] S. D. Thepade and A. A. Tonge, "An optimized key frame extraction for detection of near duplicates in content based video retrieval," in *Communications and Signal Processing (ICCSP), 2014 International Conference on*, pp. 1087–1091, 2014.

[178] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*. New York: Dover Publications, 1964.

[179] L. Chihara and D. Stanton, "Zeros of generalized Krawtchouk polynomials," *J. Approx. Theory*, vol. 60, no. 1, pp. 43–57, 1990.

[180] W. A. Jassim and P. Raveendran, "Face Recognition Using Discrete Tchebichef-Krawtchouk Transform," in *Multimedia (ISM), 2012 IEEE International Symposium on*, pp. 120–127, 2012.

[181] K.-H. Thung, R. Paramesran, and C.-L. Lim, "Content-based image quality metric using similarity measure of moment vectors," *Pattern Recognit.*, vol. 45, no. 6, pp. 2193–2204, Jun. 2012.

[182] W. A. Jassim, R. Paramesran, and M. S. A. Zilany, "Enhancing noisy speech signals using orthogonal moments," *IET Signal Process.*, vol. 8, no. 8, pp. 891–905, 2014.

[183] Y. Chang, Y. Zi, J. Zhao, Z. Yang, W. He, and H. Sun, "An adaptive sparse deconvolution method for distinguishing the overlapping echoes of ultrasonic guided waves for pipeline crack inspection," *Meas. Sci. Technol.*, vol. 28, no. 3, p. 35002, 2017.

[184] R. M. Gray, "Toeplitz and Circulant Matrices: A Review," *Found. Trends® Commun. Inf. Theory*, vol. 2, no. 3, pp. 155–239, 2006.

[185] W. Zhou, L. Yu, W. Qiu, Y. Zhou, and M. Wu, "Local gradient patterns (LGP): An effective local-statistical-feature extraction scheme for no-reference image quality assessment," *Inf. Sci. (Ny).*, vol. 397, no. Supplement C, pp. 1–14, 2017.

[186] M. S. Islam, "Local gradient pattern-A novel feature representation for facial expression recognition," *J. AI Data Min.*, vol. 2, no. 1, pp. 33–38, 2014.

[187] L. Liu, Y. Hua, Q. Zhao, H. Huang, and A. C. Bovik, "Blind image quality assessment by relative gradient statistics and adaboosting neural network," *Signal Process. Image Commun.*, vol. 40, no. Supplement C, pp. 1–15, 2016.

[188] T. Kar and P. Kanungo, "A motion and illumination resilient framework for automatic shot boundary detection," *Signal, Image Video Process.*, vol. 11, no. 7, pp. 1237–1244, 2017.

[189] H. Liu, H. Lu, and X. Xue, "A Segmentation and Graph-Based Video Sequence Matching Method for Video Copy Detection," *IEEE Trans. Knowl. Data Eng.*, vol. 25, no. 8, pp. 1706–1718, Aug. 2013.

[190] J. Yang, B. Jiang, B. Li, K. Tian, and Z. Lv, "A Fast Image Retrieval Method Designed for Network Big Data," *IEEE Trans. Ind. Informatics*, vol. 13, no. 5, pp. 2350–2359, Oct. 2017.

[191] H. Zhu, H. Shu, J. Liang, L. Luo, and J.-L. Coatrieux, "Image analysis by discrete orthogonal Racah moments," *Signal Processing*, vol. 87, no. 4, pp. 687–708, 2007.

[192] J. S. RANI and D. DEVARAJ, "Face recognition using Krawtchouk moment," *Sadhana*, vol. 37, no. 4, pp. 441–460, 2012.

[193] J. Xu, Y. Y. Tang, B. Zou, Z. Xu, L. Li, and Y. Lu, "The Generalization Ability of Online SVM Classification Based on Markov Sampling," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 26, no. 3, pp. 628–639, Mar. 2015.

[194] W. S. Noble, "What is a support vector machine?," *Nat. Biotechnol.*, vol. 24, no. 12, pp. 1565–1567, Dec. 2006.

[195] C. Seiffert, T. M. Khoshgoftaar, J. Van Hulse, and A. Napolitano, "RUSBoost: A hybrid approach to alleviating class imbalance," *IEEE Trans. Syst. Man, Cybern. A Syst. Humans*, vol. 40, no. 1, pp. 185–197, 2010.

[196] R. Batuwita and V. Palade, "Class imbalance learning methods for support vector machines," in *Imbalanced Learning: Foundations, Algorithms, and Applications*, H. He and Y. Ma, Eds. Wiley, 2013, pp. 83–96.

[197] V. López, A. Fernández, S. Garc\'\ia, V. Palade, and F. Herrera, "An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics," *Inf. Sci. (Ny).*, vol. 250, pp. 113–141, 2013.

[198] Q. Wang, Z. Luo, J. Huang, Y. Feng, and Z. Liu, "A novel ensemble method for imbalanced data learning: bagging of extrapolation-SMOTE SVM," *Comput. Intell. Neurosci.*, vol. 2017, 2017.

[199] D. Furundzic, S. Stankovic, S. Jovicic, S. Punisic, and M. Subotic, "Distance based resampling of imbalanced classes: With an application example of speech quality assessment," *Eng. Appl. Artif. Intell.*, vol. 64, no. Supplement C, pp. 440–461, 2017.

[200] B. X. Wang and N. Japkowicz, "Boosting support vector machines for imbalanced data sets," *Knowl. Inf. Syst.*, vol. 25, no. 1, pp. 1–20, 2010.

[201] M. Vij, V. Naik, and V. Gunturi, "Using smartphone-based accelerometer to detect travel by metro train," 2016.

[202] and C.-J. L. Chih-Wei Hsu, Chih-Chung Chang, "A Practical Guide to Support Vector Classification," *BJU Int.*, vol. 101, no. 1, pp. 1396–400, 2003.

[203] K. K. Sergios Theodoridis, *Pattern Recognition*, 4th ed. Academic Press, 2008.

[204] H. Zhang, R. Hu, and L. Song, "A shot boundary detection method based on color feature," in *Computer Science and Network Technology (ICCSNT), 2011 International Conference on*, vol. 4, pp. 2541–2544, 2011.

[205] X. Gao, J. Li, and Y. Shi, "A Video Shot Boundary Detection Algorithm Based on Feature Tracking," in *Rough Sets and Knowledge Technology: First International Conference, RSKT 2006, Chongquing, China, July 24-26, 2006. Proceedings*, G.-Y. Wang, J. F. Peters, A. Skowron, and Y. Yao, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 651–658.

[206] C.-W. Su, H. Y. M. Liao, H.-R. Tyan, K.-C. Fan, and L.-H. Chen, "A motion-tolerant dissolve detection algorithm," *IEEE Trans. Multimed.*, vol. 7, no. 6, pp. 1106–1113, Dec. 2005.

[207] R. Killick, P. Fearnhead, and I. A. Eckley, "Optimal Detection of Changepoints With a Linear Computational Cost," *J. Am. Stat. Assoc.*, vol. 107, no. 500, pp. 1590–1598, 2012.

[208] P. Medentzidou and C. Kotropoulos, "Video summarization based on shot boundary detection with penalized contrasts," in *2015 9th International Symposium on Image and Signal Processing and Analysis (ISPA)*, pp. 199–

203, 2015.

[209] D. Potapov, M. Douze, Z. Harchaoui, and C. Schmid, "Category-Specific Video Summarization," in *Computer Vision -- ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part VI*, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds. Cham: Springer International Publishing, 2014, pp. 540–555.

[210] R. Ram, A. Shetty, and S. Chaudhuri, "Optimal Shot Detection and Recognition Using Shiryaev-Roberts Statistics," in *Proceedings of the Seventh Indian Conference on Computer Vision, Graphics and Image Processing*, pp. 33–39, 2010.

[211] M. Pollak, "Optimal detection of a change in distribution," *Ann. Stat.*, pp. 206–227, 1985.

[212] P. Rubin-Delanchy, D. J. Lawson, M. J. Turcotte, N. Heard, and N. M. Adams, "Three Statistical Approaches to Sessionizing Network Flow Data," in *2014 IEEE Joint Intelligence and Security Informatics Conference*, pp. 244–247, 2014.

[213] S. Smith, *Digital signal processing: a practical guide for engineers and scientists*. Newnes, 2013.

[214] D. M. Thounaojam, V. S. Bhadouria, S. Roy, and K. M. Singh, "Shot boundary detection using perceptual and semantic information," *Int. J. Multimed. Inf. Retr.*, vol. 6, no. 2, pp. 167–174, 2017.

[215] T. J. Ross, *Fuzzy Logic with Engineering Applications*, Third Edit. John Wiley & Sons, Ltd, 2010.

[216] A. Vedaldi and B. Fulkerson, "Vlfeat: An Open and Portable Library of Computer Vision Algorithms," in *Proceedings of the 18th ACM International Conference on Multimedia*, pp. 1469–1472, 2010.

[217] P. H. S. Torr and A. Zisserman, "MLESAC: A New Robust Estimator with Application to Estimating Image Geometry," *Comput. Vis. Image Underst.*, vol. 78, no. 1, pp. 138–156, 2000.

[218] C. J. Van Rijsbergen, *Information retrieval*. London: Butterworths, 1979.

[219] J. Makhoul, F. Kubala, R. Schwartz, R. Weischedel, and others, "Performance measures for information extraction," in *Proceedings of DARPA broadcast news workshop*, pp. 249–252, 1999.

[220] O. V. Project, 2016. [Online]. Available: https://open-video.org.

[221] A. L. Da Cunha, J. Zhou, and M. N. Do, "The nonsubsampled contourlet transform: theory, design, and applications," *IEEE Trans. image Process.*, vol. 15, no. 10, pp. 3089–3101, 2006.

[222] J. R. Fienup, "Invariant error metrics for image reconstruction," *Appl. Opt.*, vol. 36, no. 32, pp. 8352–8357, 1997.

[223] A. K. Jain, *Fundamentals of digital image processing*. Prentice-Hall, Inc., 1989.

[224] D. Salomon, *Data compression: the complete reference*. Springer Science &

Business Media, 2004.

[225] C.-C. Chang and C.-J. Lin, “LIBSVM,” *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 1–27, Apr. 2011.

[226] P. Over, T. Ianeva, W. Kraaij, and A. F. Smeaton, “TRECVID 2006 - An overview,” *NIST*, pp. 1–29, 2006.

[227] P. Over, G. Awad, W. Kraaij, and A. F. Smeaton, “TRECVID 2007-Overview,” *NIST*, pp. 1–27, 2014.