INTERNATIONAL JOURNAL OF RESEARCH IN COMPUTER APPLICATION & MANAGEMENT

IJRCM

A Monthly Double-Blind Peer Reviewed (Refereed/Juried) Open Access International e-Journal - Included in the International Serial Directories

Indexed & Listed at:

Ulrich's Periodicals Directory ©, ProQuest, U.S.A., EBSCO Publishing, U.S.A., Cabell's Directories of Publishing Opportunities, U.S.A.,
Open J-Gage, India [link of the same is duly available at Inflibnet of University Grants Commission (U.G.C.)],
Index Copernicus Publishers Panel, Poland with IC Value of 5.09 & number of libraries all around the world.

Circulated all over the world & Google has verified that scholars of more than 3770 Cities in 175 countries/territories are visiting our journal on regular basis.

Ground Floor, Building No. 1041-C-1, Devi Bhawan Bazar, JAGADHRI – 135 003, Yamunanagar, Haryana, INDIA

http://ijrcm.org.in/

CONTENTS

CALL FOR MANUSCRIPTS

We invite unpublished novel, original, empirical and high quality research work pertaining to recent developments & practices in the areas of Computer Science & Applications; Commerce; Business; Finance; Marketing; Human Resource Management; General Management; Banking; Economics; Tourism Administration & Management; Education; Law; Library & Information Science; Defence & Strategic Studies; Electronic Science; Corporate Governance; Industrial Relations; and emerging paradigms in allied subjects like Accounting; Accounting Information Systems; Accounting Theory & Practice; Auditing; Behavioral Accounting; Behavioral Economics; Corporate Finance; Cost Accounting; Econometrics; Economic Development; Economic History; Financial Institutions & Markets; Financial Services; Fiscal Policy; Government & Non Profit Accounting; Industrial Organization; International Economics & Trade; International Finance; Macro Economics; Micro Economics; Rural Economics; Co-operation; Demography: Development Planning; Development Studies; Applied Economics; Development Economics; Business Economics; Monetary Policy; Public Policy Economics; Real Estate; Regional Economics; Political Science; Continuing Education; Labour Welfare; Philosophy; Psychology; Sociology; Tax Accounting; Advertising & Promotion Management; Management Information Systems (MIS); Business Law; Public Responsibility & Ethics; Communication; Direct Marketing; E-Commerce; Global Business; Health Care Administration; Labour Relations & Human Resource Management; Marketing Research; Marketing Theory & Applications; Non-Profit Organizations; Office Administration/Management; Operations Research/Statistics; Organizational Behavior & Theory; Organizational Development; Production/Operations; International Relations; Human Rights & Duties; Public Administration; Population Studies; Purchasing/Materials Management; Retailing; Sales/Selling; Services; Small Business Entrepreneurship; Strategic Management Policy; Technology/Innovation; Tourism & Hospitality; Transportation Distribution; Algorithms; Artificial Intelligence; Compilers & Translation; Computer Aided Design (CAD); Computer Aided Manufacturing; Computer Graphics; Computer Organization & Architecture; Database Structures & Systems; Discrete Structures; Internet; Management Information Systems; Modeling & Simulation; Neural Systems/Neural Networks; Numerical Analysis/Scientific Computing; Object Oriented Programming; Operating Systems; Programming Languages; Robotics; Symbolic & Formal Logic; Web Design and emerging paradigms in allied subjects.

Anybody can submit the **soft copy** of unpublished novel; original; empirical and high quality **research work/manuscript** ***anytime*** in ***M.S. Word format*** after preparing the same as per our **GUIDELINES FOR SUBMISSION**; at our email address i.e. infoijrcm@gmail.com or online by clicking the link **online submission** as given on our website (***FOR ONLINE SUBMISSION, CLICK HERE***).

GUIDELINES FOR SUBMISSION OF MANUSCRIPT

1. **COVERING LETTER FOR SUBMISSION:**

DATED: _____________

THE EDITOR
IJRCM

Subject: **SUBMISSION OF MANUSCRIPT IN THE AREA OF .**

(e.g. Finance/Marketing/HRM/General Management/Economics/Psychology/Law/Computer/IT/Engineering/Mathematics/other, please specify)

DEAR SIR/MADAM

Please find my submission of manuscript entitled '____' for possible publication in your journals.

I hereby affirm that the contents of this manuscript are original. Furthermore, it has neither been published elsewhere in any language fully or partly, nor is it under review for publication elsewhere.

I affirm that all the author (s) have seen and agreed to the submitted version of the manuscript and their inclusion of name (s) as co-author (s).

Also, if my/our manuscript is accepted, I/We agree to comply with the formalities as given on the website of the journal & you are free to publish our contribution in any of your journals.

NAME OF CORRESPONDING AUTHOR:
Designation:
Affiliation with full address, contact numbers & Pin Code:
Residential address with Pin Code:
Mobile Number (s):
Landline Number (s):
E-mail Address:
Alternate E-mail Address:

NOTES:
a) The whole manuscript is required to be in ***ONE MS WORD FILE*** only (pdf. version is liable to be rejected without any consideration), which will start from the covering letter, inside the manuscript.
b) The sender is required to mentionthe following in the **SUBJECT COLUMN** of the mail:
New Manuscript for Review in the area of (Finance/Marketing/HRM/General Management/Economics/Psychology/Law/Computer/IT/Engineering/Mathematics/other, please specify)
c) There is no need to give any text in the body of mail, except the cases where the author wishes to give any specific message w.r.t. to the manuscript.
d) The total size of the file containing the manuscript is required to be below **500 KB**.
e) Abstract alone will not be considered for review, and the author is required to submit the complete manuscript in the first instance.
f) The journal gives acknowledgement w.r.t. the receipt of every email and in case of non-receipt of acknowledgment from the journal, w.r.t. the submission of manuscript, within two days of submission, the corresponding author is required to demand for the same by sending separate mail to the journal.

2. **MANUSCRIPT TITLE**: The title of the paper should be in a 12 point Calibri Font. It should be bold typed, centered and fully capitalised.

3. **AUTHOR NAME (S) & AFFILIATIONS**: The author (s) **full name**, **designation**, **affiliation** (s), **address**, **mobile/landline numbers**, and **email/alternate email address** should be in italic & 11-point Calibri Font. It must be centered underneath the title.

4. **ABSTRACT**: Abstract should be in fully italicized text, not exceeding 250 words. The abstract must be informative and explain the background, aims, methods, results & conclusion in a single para. Abbreviations must be mentioned in full.

5. **KEYWORDS**: Abstract must be followed by a list of keywords, subject to the maximum of five. These should be arranged in alphabetic order separated by commas and full stops at the end.

6. **MANUSCRIPT**: Manuscript must be in ***BRITISH ENGLISH*** prepared on a standard A4 size ***PORTRAIT SETTING PAPER***. It must be prepared on a single space and single column with 1" margin set for top, bottom, left and right. It should be typed in 8 point Calibri Font with page numbers at the bottom and centre of every page. It should be free from grammatical, spelling and punctuation errors and must be thoroughly edited.

7. **HEADINGS**: All the headings should be in a 10 point Calibri Font. These must be bold-faced, aligned left and fully capitalised. Leave a blank line before each heading.

8. **SUB-HEADINGS**: All the sub-headings should be in a 8 point Calibri Font. These must be bold-faced, aligned left and fully capitalised.

9. **MAIN TEXT**: The main text should follow the following sequence:

   INTRODUCTION

   REVIEW OF LITERATURE

   NEED/IMPORTANCE OF THE STUDY

   STATEMENT OF THE PROBLEM

   OBJECTIVES

   HYPOTHESES

   RESEARCH METHODOLOGY

   RESULTS & DISCUSSION

   FINDINGS

   RECOMMENDATIONS/SUGGESTIONS

   CONCLUSIONS

   SCOPE FOR FURTHER RESEARCH

   ACKNOWLEDGMENTS

   REFERENCES

   APPENDIX/ANNEXURE

   It should be in a 8 point Calibri Font, single spaced and justified. The manuscript should preferably not exceed ***5000 WORDS***.

10. **FIGURES &TABLES**: These should be simple, crystal clear, centered, separately numbered &self explained, and **titles must be above the table/figure**. **Sources of data should be mentioned below the table/figure**. It should be ensured that the tables/figures are referred to from the main text.

11. **EQUATIONS**:These should be consecutively numbered in parentheses, horizontally centered with equation number placed at the right.

12. **REFERENCES**: The list of all references should be alphabetically arranged. The author (s) should mention only the actually utilised references in the preparation of manuscript and they are supposed to follow **Harvard Style of Referencing**. The author (s) are supposed to follow the references as per the following:

- All works cited in the text (including sources for tables and figures) should be listed alphabetically.
- Use (**ed.**) for one editor, and (**ed.s)** for multiple editors.
- When listing two or more works by one author, use --- (20xx), such as after Kohl (1997), use --- (2001), etc, in chronologically ascending order.
- Indicate (opening and closing) page numbers for articles in journals and for chapters in books.
- The title of books and journals should be in italics. Double quotation marks are used for titles of journal articles, book chapters, dissertations, reports, working papers, unpublished material, etc.
- For titles in a language other than English, provide an English translation in parentheses.
- The location of endnotes within the text should be indicated by superscript numbers.

PLEASE USE THE FOLLOWING FOR STYLE AND PUNCTUATION IN REFERENCES:

BOOKS

- Bowersox, Donald J., Closs, David J., (1996), "Logistical Management." Tata McGraw, Hill, New Delhi.
- Hunker, H.L. and A.J. Wright (1963), "Factors of Industrial Location in Ohio" Ohio State University, Nigeria.

CONTRIBUTIONS TO BOOKS

- Sharma T., Kwatra, G. (2008) Effectiveness of Social Advertising: A Study of Selected Campaigns, Corporate Social Responsibility, Edited by David Crowther & Nicholas Capaldi, Ashgate Research Companion to Corporate Social Responsibility, Chapter 15, pp 287-303.

JOURNAL AND OTHER ARTICLES

- Schemenner, R.W., Huber, J.C. and Cook, R.L. (1987), "Geographic Differences and the Location of New Manufacturing Facilities," Journal of Urban Economics, Vol. 21, No. 1, pp. 83-104.

CONFERENCE PAPERS

- Garg, Sambhav (2011): "Business Ethics" Paper presented at the Annual International Conference for the All India Management Association, New Delhi, India, 19–22 June.

UNPUBLISHED DISSERTATIONS AND THESES

- Kumar S. (2011): "Customer Value: A Comparative Study of Rural and Urban Customers," Thesis, Kurukshetra University, Kurukshetra.

ONLINE RESOURCES

- Always indicate the date that the source was accessed, as online resources are frequently updated or removed.

WEBSITES

- Garg, Bhavet (2011): Towards a New Natural Gas Policy, Political Weekly, Viewed on January 01, 2012 http://epw.in/user/viewabstract.jsp

IMAGE RETRIEVAL USING SHOT BOUNDARY DETECTION AND KEY FRAME EXTRACTION BASED TECHNIQUE FOR VIDEO SUMMARIZATION

ASHWINI P
ASST. PROFESSOR
VIDYA VIKAS INSTITUTE OF ENGINEERING & TECHNOLOGY
MYSURU

ABSTRACT

A browsing facility that provides an information oriented summary for selection of actual content is a necessity. This has lead to an increasing demand of efficient techniques to store, retrieve, index and summarize the video content. In general, management of such activities over large collections of video requires knowledge of the "content" of the video. A video thus contains multimodal information like speech, audio, text and picture information. The pictorial information in video is considered to be the series of frames which are images, sequential over specific speed to make it as video. Further, a video can be processed with the objective of extracting the information about the content conveyed in it. The algorithms developed for this purpose, referred as "video content analysis" algorithms. The definition of "content" is highly application dependent but there are a number of commonalities in the applications of content analysis. This increased demand for intelligent processing and analysis of multimedia information has led to the development of different approaches for intelligent video management.

KEYWORDS
HistogramCalculation, Keyframes, Sufficient Content ,TREC VID dataset ,Video Analysis.

1. INTRODUCTION

The important issues in the field of video summarization are ***Video Analysis, Video representation, video browsing and Video Retrieval.*** Video analysis is the basis for later video processing. It includes ***shot boundary detection*** and ***key frame extraction*** *as the* ***important mechanisms.*** Considering that each video frame is a 2D object and the temporal axis makes up the third dimension, a video stream spans a 3D space. Video representation is the mapping from the 3D space to the 2D view screen. Different mapping functions characterize different video representation techniques like representation based on sequential key frames, representation based on groups, representation based on scenes. Video browsing and retrieval are the two ultimate goals of a video access system and they are closely related to video analysis and representation. Video retrieval is concerned with how to return similar video clips (or scenes, shots, and frames) to a user given a video query. There are two major categories of video retrieval. One is to first extract key frames from the video data, then use image retrieval techniques to obtain the video data *indirectly*. Although easy to implement, it has the obvious problem of losing the temporal dimension. The other technique incorporates motion information (sometimes object tracking) into the retrieval process. Although this is a better technique, it requires the computationally expensive task of motion analysis. If object trajectories are to be supported, then this becomes more difficult.

Key Frames are the still images that can summarize the video in a rapid and very compact way so that the users can grasp the overall video content more quickly than by watching a set of video sequences. A shot detection based key frame selection will yield a summary that maintains temporal order but at the cost of increased redundancy. A clustering approach to video summarization results in a set of key frames which may not preserve temporal order but eliminates redundancy. The significant challenge for key frame-based video summarization is to devise methods for selecting key frames that represent the essence of the video content and not just form a random collection of frames.

Thus the importance of key frame extraction in the video summarization has motivated to take up the task of devising a methodology for efficient extraction of the key frames without losing the information content of the original video.

1.1 VIDEO SUMMARIZATION

One of the major applications of key frame extraction is *Video summarization*. Video summarization is a novel technology of content based video compression, which efficiently finds significant information from video and eliminates redundant data. The key frames should resultantly abstract the original video without omitting the important information.

For an application like video summarization, the frames of interest can be extracted based on one or more statistical features which are then used to represent the summary of the entire video. The result of generating such video summaries can range from just a collection of key frames representing the essence of a video to generating a video clip summarizing the essential content of the video by keeping temporal order intact. The simplest of the techniques for this type of temporal video summary is variable speed fast forwarding using time compression, which is to play the full video at a higher speed but still be intelligible to the viewer about the video contents. A second popular technique for generating temporal video summary is the video skimming technique which is also called gisting, where video summaries are generated by incorporating both audio and video information from the source video and played using a browsing tool. This type of summary generation is enabled by using key frame-based approach. The summaries vary from a static pictorial overview of the video content to a collection that maintains temporal order and therefore, conveys time sequenced content.

2. LITERATURE SURVEY

2.1 OVERVIEW

The problem of high quality key frame extraction and video summarization is addressed in [1]. The key frames are derived from QR decomposition where in key frame extraction system needs two presumptions. First one of which is a criterion to measure the video dynamicity for detecting the number of key frames in each shot needed to produce a summary with a predefined length. The second presumption is an accurate method that detects the independent key frames within shots.

A novel framework for video summarization was proposed in [5] in which multiple index features obtained from video frames and were combined to describe the frame difference between consecutive frames. It was observed that certain frame difference features have more influence in generating a representative frame difference measure. Moreover, some features are more relevant than others in different video genres. Therefore, for each video genre, the weights of different features were pre-determined at training phase by indirectly utilizing the Relevance Feedback Mechanism. The idea of key frames extraction from single shots in video sequences is presented in [3]. The method was implemented by an efficient two-step algorithm, which is classified neither to clustering nor to temporal variations based techniques. In the first step, an MST (minimal spanning tree) graph is constructed, where each node is associated to a single frame of the shot. In the second step, key frames were extracted based on the principle of their maximum spread. The number of the selected key frames was controlled by an adaptively defined threshold, while the validity of the results was evaluated by the fidelity measure.

A method for video summarization using a graph theoretical approach was proposed in [7] .This thesis aims comparative analysis of the state of the art shot boundary detection algorithms. The major methods that have been used for shot boundary detection such as pixel intensity based, histogram-based, edge-based, and motion vectors based, were implemented and analyzed. A recent method which utilizes "graph partition model" together with the support vector machine classifier as a shot boundary detection algorithm was also implemented and analyzed. Moreover, a novel graph theoretic concept, "dominant sets", was also successfully applied to the shot boundary detection problem as a contribution to the solution domain.

Edge-based approaches measure the similarity or dissimilarity as change in intensity of edges was presented in [4] gave a method of extracting the key frames by Accumulated Histogram Intersection Measure. It stated that, first extract key frames by matching of DC image sequence constructed from the MPEG video sequence. Then using the region segmentation-based projective histogram and its moments as database indices for video retrieval. Results clearly validated and dominated the method and that in conjunction with other indexing techniques, such as color, can provide a power lid framework for video indexing and retrieval.

Efficient key frame extraction, using local semantics in form of a region thesaurus specifically, works on certain MPEG-7 color and texture features are locally extracted from key frame regions. Then, using a hierarchical clustering approach a local region vocabulary is constructed to facilitate the description of each frame in terms of higher semantic features. The mechanism that was proposed in [6] presents that thesaurus consists of the most common region types that are encountered within the video shot, along with their synonyms. These region types carry semantic information. Each keyframe was represented by a vector consisting of the degrees of confidence of the existence of all region types within this shot. Using this keyframe representation, the most representative keyframe was then selected for each shot. Where a single keyframe was not adequate, using the same algorithm and exploiting the presence of the region types of the visual thesaurus, more keyframes are extracted.

Key Frames can be extracted by analyzing the similarity between two consecutive frames of a video sequence, the algorithm determines the complexity of the sequence in terms of changes in the visual content expressed by different frame descriptor features. This methodology has been proposed in the paper [11] which escapes the complexity of existing methods based, for example, on clustering or optimization strategies, dynamically and rapidly selects a variable number of key frames within each sequence. The key frames were extracted by detecting curvature points within the curve of the cumulative frame differences. Another advantage was that it can extract the key frames on the fly: curvature points can be determined while computing the frame differences and the key frames can be extracted as soon as a second high curvature point has been detected.

A Method of Independent Component Analysis (ICA) [9] suggests that projecting video frames from illumination-invariant raw feature space into low dimensional ICA subspace, each video frame is represented by a two-dimensional compact feature vector. An iterative clustering algorithm based on adaptive thresholding is developed to detect cuts and gradual transitions simultaneously in ICA subspace. An iterative clustering algorithm based on adaptive thresholding is developed to detect cuts and gradual transitions simultaneously in ICA subspace. Experimental results successfully validate the new method and show that it can effectively detect both abrupt transitions and gradual transitions.

An algorithm that concentrates mainly on the athletics video was given in [10]. It proposed a simple but effective method to extract key frames from an athletic sport sequence. The extracted features are the start and finish race frame and the type of competition. Specifically they adopted camera motion detection technique for both race detection and features extraction. An effective rule has been developed for segmenting start/race/finish from motion vector sequences. Proposed algorithm is based on motion vector analysis, and camera motion detection.

An innovative approach to the selection of representative frames of a video shot for video summarization by analyzing the differences between two consecutive frames of a video sequence [8]. The algorithm determined the complexity of the sequence in terms of visual content changes. Three descriptors were used to express the frame's visual content: a color histogram, wavelet statistics and an edge direction histogram. Similarity measures were computed for each descriptor and combined to form a frame difference measure. The use of multiple descriptors provides a more precise representation, capturing even small variations in the frame sequence. This method can dynamically and rapidly select a variable number of key frames within each shot, and does not exhibit the complexity of existing methods based on clustering algorithm strategies.

The three popular approaches to shot boundary detection, namely statistical difference, histogram comparison, and pixel difference were compared in [15]. Instead of using common performance measure as recall and precision, it was proposed that new performance measure as Retrieval Success Index (RSI) was introduced.

A review of certain basic information extraction operations that can be performed on video is presented in [13] [14]. These operations were particularly useful as a foundation for further semantic processing. Specifically, the review focuses on shot boundary detection and condensed video representation (also called summarization and abstraction). Shot boundary detection is the complete segmentation of a video into continuously imaged temporal video segments. Condensed video representation is the extraction of video frames or short clips that are semantically representative of the corresponding video. Both tasks are very significant for the organization of video data into more manageable forms.

3. KEY FRAME EXTRACTION

The fast evolution of the digital video technology has opened new areas of research. The most important aspect will be to develop algorithms to perform video cataloguing, indexing and retrieval. The basic step is to find a way for video abstraction, as this will help us more for browsing a large set of video data with sufficient content representation. One such basic step in video abstraction is key - frame extraction. Hence key frame extraction has been an important technique in video summarization, browsing, searching and understanding.

A video sequence normally contains a large number of frames. In order to ensure that humans do not perceive any discontinuity in the video stream, a frame rate of at least 25fps is required, that is, 90,000 images for one hour of video content. This sheer volume of video data is a barrier to many practical applications, and therefore, there is a strong demand for a mechanism that allows the user to gain certain perspectives of a video document without watching/addressing the video in its entirety. This mechanism is termed video abstracting.

There are two basic forms of video abstracts:

- **Key-frames:** These are also called representative frames, R-frames, still-image abstracts or static storyboard and a set consists of a collection of salient images extracted from the underlying video source. Hence, the keyframe set *R* is defined as follows:

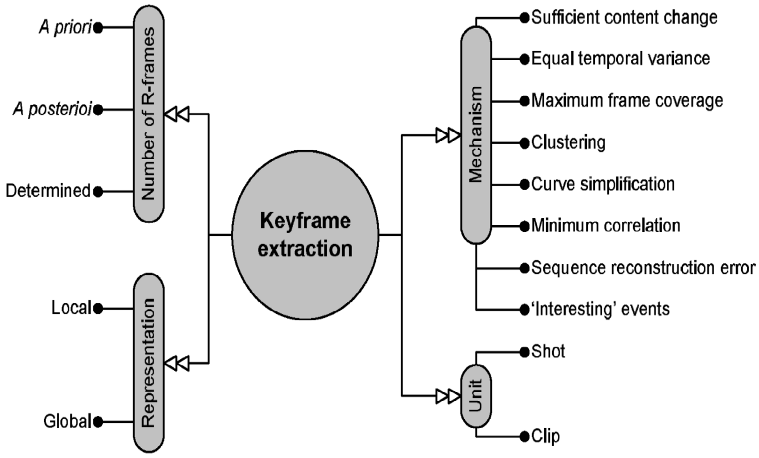
  $$\mathcal{R} = \mathcal{A}_{Keyframe}(\mathbf{V}) = \{f_{r_1}, f_{r_2}, \ldots, f_{r_\mathbf{k}}\}, \quad (3.1)$$

  Where A keyframe denotes the keyframe extraction procedure.
- **Video skimming:** It is method to extract the significant audio and video information and create a "video skim" which represents a very short synopsis of the original. The goal of video skim is to show the utility of integrating language and image understanding techniques for video skimming by extraction of significant information, such as specific objects, audio keywords and relevant video structure. This is also called a moving-image abstract, moving storyboard, or summary sequence. This type of abstract consists of a collection of video segments (and corresponding audio) extracted from the original video. These segments are joined by either a cut or a gradual effect (e.g., fades, dissolve, wipe). It is itself a video clip, but of significantly shorter duration. One popular kind of video skim in practice is the movie trailer. The video skim K is defined as follows:

  $$\mathcal{K} = \mathcal{A}_{Skim}(\mathbf{V}) = \mathbf{E}_{i_1} \odot \mathbf{E}_{i_2} \odot \ldots \odot \mathbf{E}_{i_\mathbf{k}}, \quad (3.2)$$

Where A_{Skim} denotes the skim generation procedure. $E_i \sqsubset V$ is the i^{th} excerpt to be included in the skim, and $\odot$ is the excerpt assembly and integration operation (e.g., cut, fade, dissolve, wipe). Note that the preceding representation is applicable to the video stream only, whereas the audio stream of a video skim can be re-sequenced, edited, and does not necessarily synchronize with the original video.

FIGURE 3.1: ATTRIBUTES OF KEYFRAME EXTRACTION TECHNIQUES



One advantage of a video skim over a keyframe set is the ability to include audio and motion elements that potentially enhance both the expressiveness and information of the abstract. In addition, it is often more entertaining and interesting to watch a skim than a slide show of keyframes. On the other hand, since they are not restricted by any timing or synchronization issues, once keyframes are extracted, there are further possibilities of organizing them for browsing and navigation purposes, rather than the strict sequential display of video skims. Hence keyframes are used in this project because of the easy of extraction and organizing the keyframes.

Note that although video skims and keyframes are often generated differently, these two forms of video abstract can be transformed from one to the other. Video skims can be created from keyframes by joining fixed-size segments, sub shots, or the whole shots that enclose them.

3.1 KEYFRAMES

Keyframes are a set of salient images extracted from video sequences. They provide a simple yet effective way of summarizing the content of videos for browsing and retrieval and are also widely used in video abstraction due to their compactness. Much research has been conducted in the past few years in understanding the problem of key-frame extraction and developing effective algorithms.

Although simple and computationally efficient sampling-based methods may produce no keyframes for a shot, yet semantically producing too many keyframes with identical content to represent a long static segment thus failing to effectively represent the actual video content. The review of keyframe extraction techniques discussed will focus only on techniques that take into account the underlying dynamics, to different degrees and from varying viewpoints, of the video sequence. The discussion of fundamental aspects of the current approaches in keyframe extraction, as depicted in Figure 3.1. These aspects are: the size of the keyframe set, the base unit, representation scope, the underlying computational mechanisms, and visualization method for extracted keyframes. For each aspect, a categorization is provided of how they are addressed in existing works and describe their pros and cons.

THE SIZE OF THE KEYFRAME SET

There are different options for determining the number of keyframes in an automatic keyframe extraction process, and they strongly shape the underlying formulation of the optimal keyframe set. The size of the keyframe set can be fixed as a known priori, left as an unknown posteriori, or determined internally within the abstraction process. Most of the techniques only offer one option for the size of keyframe set. However, if the algorithm produces a number of keyframes progressively, two options are available: It can stop when the number of keyframes reaches a priori value or when certain criteria are satisfied (i.e., a posteriori). Further for applications where efficient utilization of network bandwidth and storage is crucial, the size of keyframe set should be measured by the total number of bits required to code the summary with a given coding strategy.

- **A Priori:** The number of keyframes is decided beforehand and given as a constraint to the extraction algorithm. It can be assigned as a specific number or ratio over the length of the video that may vary according to user knowledge of the video content. Also called rate constraint keyframe extraction, this approach is suitable and often required in mobile device systems where available resources are limited. For these systems the number of keyframes is distributed differently, depending on the transmission bandwidth, storage capacity, or display size of the receiving terminal. A special yet common case is when one keyframe is selected per shot, which is often the first frame, middle frame, or that frame closest to the average content of the shot. This technique has a disadvantage in that it does not ensure that all important segments in a video contain at least one keyframe.

Ideally, the keyframe extraction problem with a priori size k can be formulated as an optimization problem of finding the frame set R = { fi1 , fi2 , . . . , fik} that is differs least from the video sequence with respect to a certain summarization perspective:

$\{r1, r2, \ldots, \mathbf{rk}\} = \arg\min \{D(R, \mathbf{V}, \rho) \mid 1 \le ri \le \mathbf{n}\}$ (3.3)

Where n is the number of frames in the original video sequence, ρ is the summarization perspective that the user is interested in, and D is a dissimilarity measure.

Most current keyframe extraction techniques have ρ as "visual coverage," which aims to cover as much visual content with as few frames as possible; however, ρ can also be the number of objects, number of faces, etc. It can also represent a viewpoint on what constitutes the optimal keyframe set, and shape the underlying computational solution.

- **A Posteriori:** In this approach, user does not know the number of extracted keyframes until the process finishes. The number of keyframes is often determined by the level of visual change, itself. More keyframes are needed to represent a video sequence with a lot of action and movements than required for a static one. However, an excessively large number of keyframes may be produced for highly dynamic scenes, causing inconvenience during interactive tasks. The formulation of the keyframe extraction problem with no specified size requires a dissimilarity tolerance ε, also called the fidelity level. This can generally be formulated as

$\{r1, r2, \mathbf{rk}\} = \arg\min \{\mathbf{k} \mid D(R, \mathbf{V}, \rho) < \varepsilon, 1 \le ri \le \mathbf{n}\}$ (3.4)

Technically, with a significant increase in computational cost, this approach can be converted in to the first approach (number of keyframes is a priori) by iteratively reducing the fidelity level until the number of keyframes produced approximates a predefined value, and vice versa.

- **Determined:** This is in essence the a posteriori approach; however, there is need to acknowledge the approaches that attempt to determine the appropriate number of keyframes before the full extraction is executed. For example, with cluster-based methods, cluster validation can be performed to determine the number of clusters before or within the actual clustering process on the other hand, consider the number of keyframes allocated for each shot as being proportional to its accumulative content change.

REPRESENTATION SCOPE OF INDIVIDUAL KEYFRAMES

This aspect of the keyframe extraction algorithm relates to whether an extracted keyframe is representative of its local neighborhood segment or represents noncontiguous segments of a video clips. The other approach tends to produce a smaller keyframe set, while the first preserves the temporal visual progression, which may be important for understanding the action and events in a summary.

UNDERLYING COMPUTATIONAL MECHANISMS

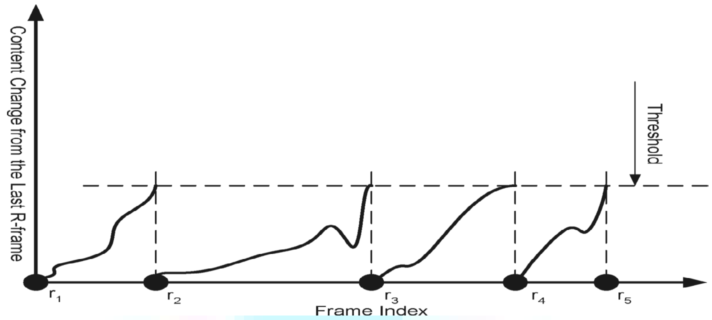
The underlying computational mechanisms for keyframe extraction can be categorized into eight somewhat overlapping classes, as depicted in Figure 3.1. The same are briefly discussed below.

(i) Sufficient Content Change: This method proceeds sequentially and only requires knowledge about the video sequence until the current temporal position, selecting a frame as the keyframe only if its visual content significantly differs from previously extracted keyframes as shown in Figure 3.2. In the other words, in its basic form, the sufficient content change method selects the next keyframe f_{ri+1} based on the most recently selected keyframe f_{ri} as follows:

fr_{i+1} = arg min {C (ft, fri) > ε, i < t ≤ **n**} (3.5)

with the first frame of the base unit, for example, the shot, often selected as the first keyframe, namely, $r_1 = 1$.

FIGURE 3.2: ILLUSTRATION OF THE SUFFICIENT CONTENT CHANGE METHOD



(ii) Equal Temporal Variance: This is somewhat a variant of the sufficient content change method, which requires that the number of keyframes is specified a priori, and assumes that a good keyframe set should have individual keyframes represent video segments (bi , bi+1−1) of equal temporal variance. In this approach, the determination of {bi} and {ri} is performed separately in two steps. Theoretically, segment boundaries {bi} are selected so that

V (b_1, b_2) = V (b_2, b_3) = ·· · = V (b_i, b_{i+1}) (3.6)

Where V (b_i, b_{i+1}) denotes the temporal variance of the segment (b_i, b_{i+1-1}), which has f_{ri} (selected keyframe) as its representative frame. Since the exact equality of temporal variance is often not achieved, thus it is an optimization problem.

In general, the equal temporal variance method is more computationally expensive than the sufficient content change method. However, the produced keyframe set is globally “optimal” and independent of any processing order.

(iii) Maximum Frame Coverage: It is referred to as the fidelity based approach, this method attempts to use the representation coverage (i.e., the list of frames that a given frame can represent) of individual frames in the sequence as the starting point. Let $\mathbf{C}_i$ (ε) denote the set of frames in **V** that can use f_i as their representative with respect to a tolerance or fidelity value ε. Under this framework, the optimal set of keyframes from **V** with no rate constraint can be extracted as

{r1, r2, **rk**} = arg min {**k**|**C**r1 (ε) ∪ **C**r2 (ε) ∪ ·· ·∪**C**rk (ε) = **V**} (3.7)

When a given number of keyframes is specified as a constraint, the problem can be interpreted as either finding the minimum fidelity value ε such that all frames can be covered by at least one keyframe or finding the set of frames that can represent as many frames as possible.

{r1, r2, **rk**} = arg max {|**C**r1 (ε) ∪ **C**r2 (ε) ∪ · · · ∪ **C**rk (ε)|} (3.8)

Two crucial components in this approach are the determination of coverage for a frame and the optimization procedure. The advantage of the maximum frame coverage formulation is that it does not require the representation coverage of a frame to be a contiguous segment as in the sufficient content change and equal temporal variance methods, and therefore, can generate a more concise keyframe set. However, the dissimilarity needs to be computed on all frame pairs, making it computationally more expensive, and thus unsuitable for real-time and/or online applications. It also requires a true visual similarity metric to fully realize its optimality.

(iv) Clustering: This approach treats video frames as points in the feature space (e.g., color histogram) and works on the assumption that the representative points of clusters formed in this space can be used as keyframes for the entire video sequence. It does not rely on any explicit modeling, while making use of generic techniques developed for data clustering. Clustering can be performed as clip-based or shot-based, and often involves the four following steps. Existing methods are distinguished based on how these steps are implemented:

- Preprocessing the data - Preprocessing the input data is aimed at improving the effectiveness and efficiency of the clustering process.
- Clustering the data - Standard clustering techniques or their domain-dependent variants now can be applied to the data to identify potential clusters.
- Filtering clusters - Since clustering outputs may be noisy and/or the cluster itself not sufficiently significant, keyframes are not assigned to all clusters.
- Extracting the representative points of each cluster - The most common and intuitive solution is to select that point closest to the cluster centroid as the representative point of the cluster, which eventually forms the keyframe set of the video sequence

(v) Minimum Correlation Among Keyframe: Techniques in this class often deal with the rate constrained keyframe extraction problem and work on the basis that the keyframe set should have minimal correlation among its elements (sometimes only the correlation among successive elements is considered). It aims to select frames that are dissimilar to each other. The optimal extraction of keyframes via the minimal correlation criterion can be formulated as

{r1, r2, **rk**} = arg min {Carr (fr1, fr2, . . . , fo**rk**)} (3.9)

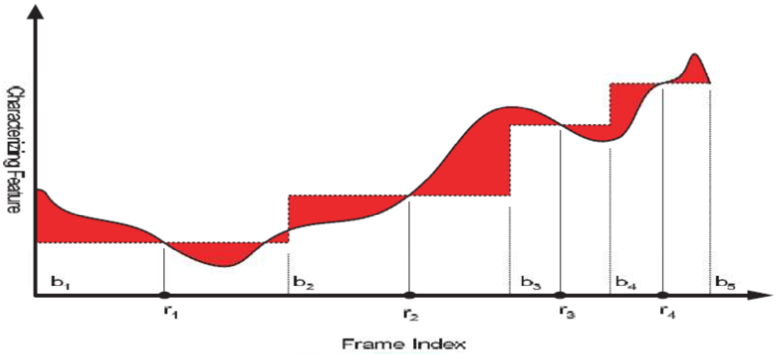
Where, Corr (.) is a correlation measure.

(vi) Sequence Reconstruction Error (SRE): This approach is proposed in Liu and Kender, and is based on a metric called the Shot Reconstruction Degree (SRE) score. This measures the capability of the keyframe set for reconstructing the original video sequence/shot. It is useful when the number of keyframes is set as a priori and the temporal progression is crucial to the designated application. Assume that a frame interpolation function I (t, R) that calculate the full image or some characterizing feature at time t in the video sequence from the keyframe set. The SRE score ε (**V**, R) of the keyframe set is defined as

$$\mathcal{E}(\mathbf{V}, \mathcal{R}) = \sum_{i=1}^{\mathbf{n}} \mathcal{D}(f_i, \mathcal{I}(i, \mathcal{R})), \quad (3.8)$$

Where D (.) computes the difference between two image frames.

FIGURE 3.3: ILLUSTRATION OF A SEQUENCE RECONSTRUCTION ERROR



(vii) Curve Simplification: The curve simplification approach is related to clustering-based methods in that it also treats each frame in the video sequence as a point in a multidimensional feature space. However, rather than looking for clusters, this approach searches for a set of points such that removal of the remaining points least changes the shape of the curve connecting all points through their temporal ordering. In this respect, it is thus related to sequence reconstruction error methods. The difference is that points on the curve are not necessarily equally spaced by the frame index and an explicit modeling of the error between the final curve and original frame trajectory is not required. Instead, this approach uses standard curve simplification algorithms such as the binary curve splitting algorithm and discrete contour evolution

(viii) "Interesting" Events: In contrast to the preceding approaches, which focus on maximizing the extent and balance of visual coverage of the keyframe set, methods based on interesting events attempt to identify frames that are semantically important. Most work in this category assumes an association between the "interestingness" of a frame, and the motion patterns around the frame, as well as its content characteristics (e.g., containing human faces or having high spatial complexity).

5. EXPERIMENTATION AND RESULTS

The software tool used for the proposed work is MATLAB 7.11 version. MATLAB is a programming environment for algorithm development, data analysis, visualization, and numerical computation. The algorithm has been implemented on *windows 7* operating system installed on a machine with 2 Gb of primary memory and Intel I_3 second generation processor. The proposed model accepts input video in the ".avi" or ".mpg" format size ranging from 3 to 10 Mb. 3GP video converter software is used to convert other files from you tube (".flv", ".Mp4") to ".avi" format. The experimental dataset used in this project are from the youtube action dataset ,the Open Video Project and some are from trecvid dataset. Since the multimedia information retrieval community does not yet have more number of standard collections of video dataset to be used for research purposes other than TRECVID, which needs more copyright formalities for download. The purpose of the Open Video Project is to collect and make available a repository of digitized video content for the digital video, multimedia retrieval, digital library, and other research communities.

AN OVERVIEW OF IMPLEMENTATION IS GIVEN BELOW

1. The input video will be converted with the help of a script *vid2frm*.
2. The script in turn calls a function *copyFrames* and the converted frames will be stored into a specified folder. Further processing needs the frames from this location.
3. The script *HistogramCalculation* is called which will accept the size of the video segment in terms of frames from the user. In the first part of the script, the frames will be divided into number of segments or groups based on the size entered.
4. In the second part, Histogram differences will be calculated and key frames are selected from individual segments, and stored in a common place for further processing.
5. The above script calls a function called *substr* in order to get the substring of a string for processing.
6. And the same script will calculate the compact ratio for the generated key frame set

5.1 OVERVIEW OF VIDEO SAMPLES

We have used different kind of videos for testing and analysis of the algorithm. For example sports videos from youtube action, open videos downloaded from Open Video Project clips, Historical clips downloaded from TREC VID dataset. But more emphasis is given to the sports clips.

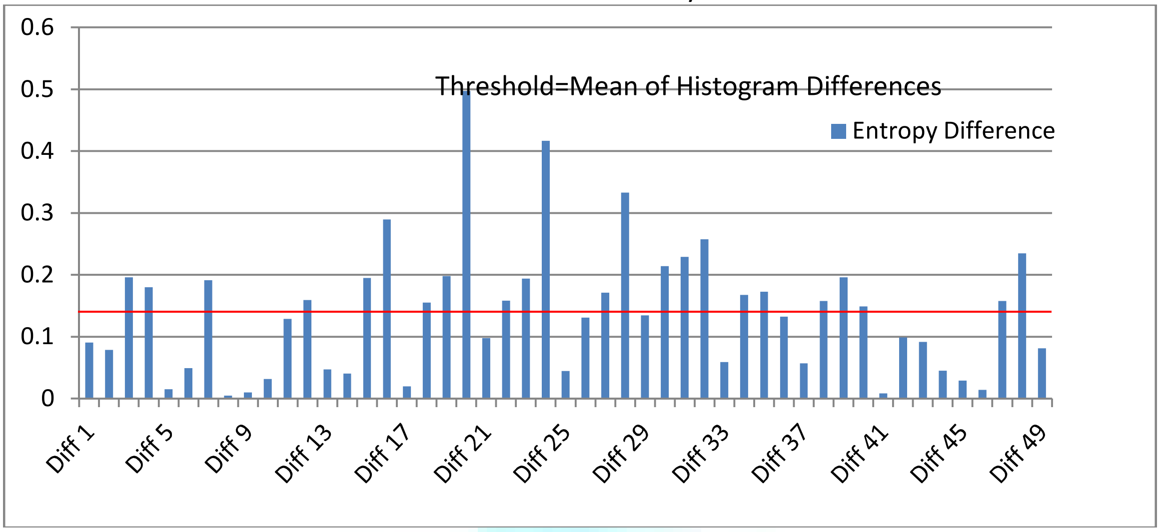
The descriptions about some of the sample video clips used for the testing are given below.

Sports video (.avi file)

FIGURE 5.1: TEST (5).AVI

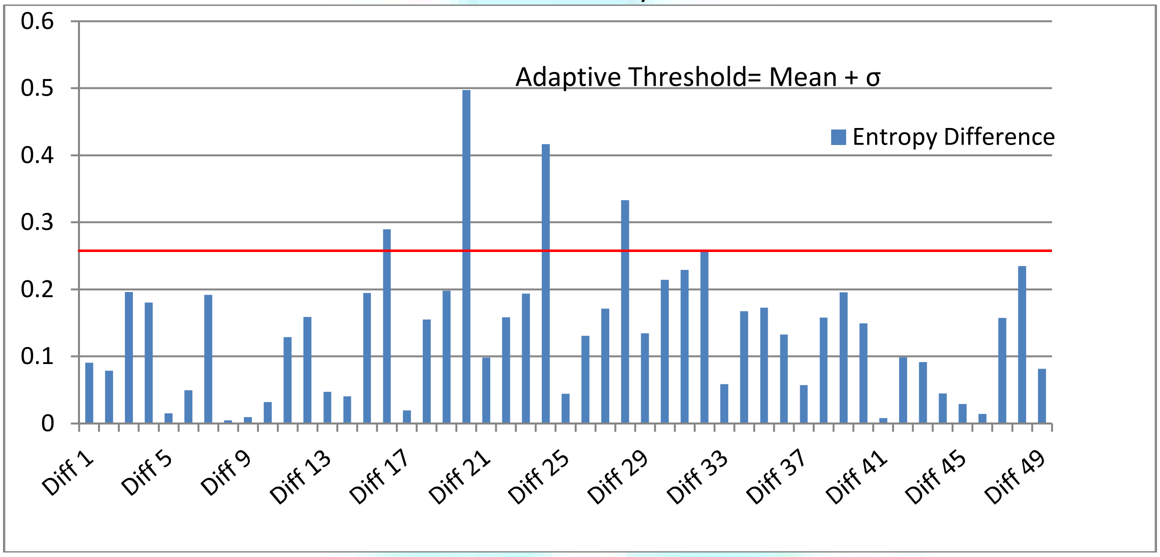
The video clip which is shown above demonstrates a sports event in which a sportsman was injured and he was given a first aid in the ground itself. Then the game was continued normally. The size of the input video file is 1.13 Mb. The duration of the file is 36 seconds at the frame rate of 25 frames/second. For the input video Test (5).avi file there are totally 908 frames extracted. Corresponding to the input file 58 key frames was extracted. The adaptive thresholding mechanism is described in the figure 5.2. The following chart demonstrates that, there are more than 40% of the frame differences values are above the threshold. Hence the value of threshold is adjusted so that the number of key frames that fall above the threshold will not exceed 10% of the group size. The mechanism is called adaptive thresholding and depicted in the chart 5.2.

CHART 5.1: HISTOGRAM DIFFERENCE V/S THRESHOLD



In the following chart it is demonstrated that the value of threshold was adjusted in number of iterations by increasing the threshold by a scalar factor equal to standard deviation.

CHART 5.2: HISTOGRAM DIFFERENCE V/S ADAPTIVE THRESHOLD



Some of the sample key frames extracted by the algorithm are listed below:

FIGURE 5.2 SAMPLE KEY FRAMES OF TEST (5).AV

CONCLUSION AND FUTURE SCOPE

The objective of the proposed work an attempt to generate a key frame set of a large video by segmenting it into semantic shots. The key frame extraction technique relies on the fact that there will be considerable changes in the information content of the frames in the video sequence. The mechanism has to be supported by global or adaptive thresholding of discontinuity measures between consecutive frames. The strategy of key frame extraction will decide the representativeness of every shot into the key frame set generated. The proposed method has been tested on different video segments of different genres like sports, add clips, educational, historical and news clips.

The performance of the video summarization algorithm is evaluated by the criteria like compression ratio and fidelity. Finally the quality of the key frame set is evaluated by representative observations by a group of human observers with the help of criteria informativeness.

REFERENCES

1. Ali Amiri, Mahmood Fathy, Atusa Naseri , *Computer Engineering Department, Iran University of Science and Technology, Tehran, Iran* " **Keyframe extraction and video summarization using QR- Decomposition**", IEEE 2007
2. Naveed Ejaz and Sung Wook Baik , College of Electronics and Information Engineering, Sejong University, Seoul, Republic of Korea. *"***Weighting low level frame difference features for key frame extraction using Fuzzy comprehensive evaluation and indirect feedback relevance mechanism By***".*
3. D. Besiris, F. Fotopoulou, N. Laskaris, G. Economou, *Department of Physics, Electronics Laboratory, University of Patras* "**Key frame extraction in video sequences: a vantage points approach**", IEEE 2007
4. EMRAH AŞAN, M.S, Department of Electrical and Electronics Engineering, *The Graduate School Of Natural And Applied Sciences of Middle East Technical University* "**Video Shot Boundary Detection by Graph Theoretic Approaches**", September 2008
5. Eung Kwan Kang, Sung Joo Kim, and Joon So0 Choi, Dept.of Physics Chung-Ang University, 221 Huksuk-Dong, "**VIDEO RETRIEVAL BASED ON SCENE CHANGE DETECTION IN COMPRESSED STREAMS",** IIT assessment
6. Evaggelos Spyrou and Yannis Avrithis, *National Technical University of Athens Image, Video and Multimedia Laboratory Zographou, 15773 Athens, Greece*. "**Keyframe Extraction using Local Visual Semantics in the form of a Region Thesaurus**", IEEE 2007
7. Frkdkric Dufaux, *Compaq Computer Corp., Cambridge Research Lab.,"* **Key frame selection to represent a video",** 0-7803-6297-7/00/ IEEE 2000
8. G. Ciocca1, R. Schettini *Dipartimento di Informatica Sistemistica e Comunicazione (DISCo) Università degli studi di Milano-Bicocca "***An innovative algorithm for key frame extraction in video summarization**".
9. Gentao Liu Xiangming Wen Wei Zheng Peizhou *He School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, China*, **"Shot Boundary Detection and Keyframe Extraction based on Scale Invariant Feature Transform",** Eigth IEEE/ACIS International Conference on Computer and Information Science, 2009
10. Giuseppe Caccia, Rosa Lancini*, CEFRIEL - Politecnico di Milano Milan, Italy***, " Algorithm for Summarization and Ketframes extraction in Athletic video" 0**-7803-713-3/02 IEEE 2002
11. Gianluigi ciocca , Raimondo Schettini *,Imaging and vision laboratory dipartmento di informatica sistemistica e communicazione univerita degli study di Milano bicocca, archimboldi 8, 20126 , Milano Italy "* **Dynamic key frame extraction for video summarization".**
12. Yueting Zhuang; Ruogui Xiao; *Fei Wu Institute of Artificial Intelligenc of Zhejiang University Zhejiang University, Hangzhou, P. R. China, 310027* **" Key Issues in Video Summarization and Its Application".** ICICS-PCM-2003 December.
13. Maheshkumar H. Kolekar, Somnath Sengupta *Electronics and Electrical Communication Engineering Department, "***VIDEO SHOT BOUNDARY DETECTION: A COMPARATIVE STUDY OF THREE POPULAR APPROACHES"** IIT Karagpur.
14. Costas Cotsaces, *Student Member, IEEE,* Nikos Nikolaidis, *Member, IEEE,* and Ioannis Pitas, *Senior Member, IEEE* **"Video Shot Boundary Detection and Condensed Representation: A review".** IEEE
15. Chong-Wah Ngo, Yu-Fei Ma‡ & Hong-Jiang Zhang‡ Department of Computer Science. City University of Hong Kong. Microsoft Research Asia Haidian District, Beijing, PRC **"Automatic Video Summarization by Graph Modeling".**

REQUEST FOR FEEDBACK

Dear Readers

At the very outset, International Journal of Research in Computer Application & Management (IJRCM) acknowledges & appreciates your efforts in showing interest in our present issue under your kind perusal.

I would like to request you tosupply your critical comments and suggestions about the material published in this issue as well as on the journal as a whole, on our E-mail**infoijrcm@gmail.com** for further improvements in the interest of research.

If youhave any queries please feel free to contact us on our E-mail **infoijrcm@gmail.com**.

I am sure that your feedback and deliberations would make future issues better – a result of our joint effort.

Looking forward an appropriate consideration.

With sincere regards

Thanking you profoundly

Academically yours

Sd/-
Co-ordinator

DISCLAIMER

ABOUT THE JOURNAL

In this age of Commerce, Economics, Computer, I.T. & Management and cut throat competition, a group of intellectuals felt the need to have some platform, where young and budding managers and academicians could express their views and discuss the problems among their peers. This journal was conceived with this noble intention in view. This journal has been introduced to give an opportunity for expressing refined and innovative ideas in this field. It is our humble endeavour to provide a springboard to the upcoming specialists and give a chance to know about the latest in the sphere of research and knowledge. We have taken a small step and we hope that with the active co-operation of like-minded scholars, we shall be able to serve the society with our humble efforts.

Our Other Journals