www.ijecs.in
International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume 3 Issue 12 December 2014, Page No. 9742-9748

Metadata Construction Model for Web Videos: A Domain Specific Approach

Siddu P. Algur[1], Prashant Bhat[2*], Suraj Jain[3]

[1]School of Mathematics and Computing Sciences, Department of Computer Science
Rani Channamma University, Belagavi-591156, Karnataka, India
Siddu_p_algur@hotmail.com

[2*]School of Mathematics and Computing Sciences, Department of Computer Science
Rani Channamma University, Belagavi-591156, Karnataka, India
prashantrcu@gmail.com

[3]School of Mathematics and Computing Sciences, Department of Computer Science
Rani Channamma University, Belagavi-591156, Karnataka, India
Suraj_point@rediffmail.com

Abstract: ***The advances in computer and information technology together with the rapid evolution of multimedia data are resulted in the huge growth of the digital video. Due to the rapid growth of digital data and video database over the Internet, it is becoming very important to extract useful information from visual data. The scientific community has increased the amount of research into new technologies, with a view to improve the digital video utilization: its archiving, indexing, accessibility, acquisition, store and even its process and usability. All these parts of the video utilization entail the necessity of the extraction of all important information of a video, especially in cases of lack of metadata information. This paper describes importance of descriptive metadata of video, categorization of video contents based on video descriptive metadata, and high level metadata based web video modeling. The main goal of this paper is to construct high level descriptive metadata with and without timeline for all the category of videos and extraction of metadata from web videos such as YouTube and Face Book. By using the high level descriptive metadata information, a user is facilitated on the one hand to locate a specific video and on the other hand is able to comprehend rapidly the basic points and generally, the main concept of a video without the need to watch the whole of it.***

Keywords: Descriptive Metadata, Web Video, Metadata Construction Model.

1. INTRODUCTION

Over the last few years it has been observed an increasingly use of multimedia and web video data. This fact automatically entails that not only the transmission but also the store and the process of the relevant information are crucial. The digital web video could be considered as the most representative example of multimedia data. More particularly, every single day a quite large quantity of digital web videos is produced and therefore, the amount of information seems huge and uncontrollable. Thus, video web sites have become overcrowded, since they have to deal with this high amount of data and with the visiting users. In the year 2011, YouTube video-sharing website had more than 1 trillion views, while according to the official YouTube statistics, each month, more than 4 billion hours of video are watched and over 800 million unique users accessed the site [1].

Online media have a great number of advantages for social science researchers: they provide lasting and exacting accounts of discussions, they contain useful metadata (date, time, subject tags, user names, etc.), and they often contain comments, discussions, and so on. Unfortunately, this data tends to be difficult to retrieve from many sites. Further compounding the problem, this data is difficult to retrieve in different ways – creating a page scraper for one site doesn't help retrieve data from others, nor does it help when that site's design changes [2].

2. RELATED WORKS

In the last decades a lot of significant approaches on the video metadata field are observed. The main objective of these works is to facilitate the user to get significant metadata from the video data. In 2013, the authors Dim P. Papadopoulos, Vicky S. Kalogeiton, Savvas A. Chatzichristofis, Nikos Papamarkos [1] worked on the concept, Automatic summarization and annotation of videos with lack of metadata information. In this method, the authors proposed a new system that automatically generates and provides all the essential information both in visual and textual form of a video, in cases of lack of metadata information. This system consists of a key-frame-based video summarization method, responsible for the visual form of information, and of an innovative key-word-based video annotation method, inextricably linked to the textual information.

Video data personalization is increasingly viewed as an essential component of any front-end to a large information space. Video data personalization can achieve customization of the presentation of information. With respect to this aspect, W. Klippgen, T.D.C. Little, G. Ahanger, and D. Venkatesh [3], described techniques to facilitate this personalization, overviewed the unique characteristics of the video medium, and proposed a framework for personalized news delivery of video information in the news domain. The proposed technique is based upon the application and integration of existing techniques of video content and structure modeling, metadata collection, vector space analysis, personalization and filtering, metadata management, and video and audio composition. The core concept is the use of a set of vector-based and time-dependent filters for audio and video segment selection to generate formatted video-based compositions on-the-fly.

To simplify video editing using metadata, the authors, Juan Casares, A. Chris Long, Brad A. Myers, Rishi Bhatnagar, Scott M. Stevens, Laura Dabbish, Dan Yocum, and Albert Corbett [4], developed a system called Silver. The proposed system Silver shows that, video metadata can be used to create an advanced user interface by providing an editable transcript view; coordinating the selection across all views including partial selections and different selections in audio and video; and smart editing through smart snap, smart selection and smart cut and paste. Silver uses video and metadata from the Informedia Digital Video Library.

There have been observed that, there exists insufficient and inappropriate for content description of complex data types such as videos, which require more detailed relational models. For this, the authors D. M. Shotton, A. Rodriguez, N. Guil and O. Trelles [5], proposed a metadata classification schema for the characterization of items and events in videos that permits subsequent query by content. With respect to MPEG-7 nomenclature, metadata intrinsic to the information content of the video are defined as either structural or semantic, where structural metadata are numerical feature primitives produced by analyzing the color, shape, texture, structure and motion within the video frames, whereas semantic metadata describe the locations and timings of individual items and particular actions or events in the video, and are thus of higher information value. In this aspect, the semantic metadata required to describe the visual information content of videos are defined and classified into four distinct classes: Media Entities; Content Items; Events; and Supplementary Items, and three types of property tables are defined: Identity Tables; Spatio-Temporal Position Tables; and Event Tables, in which these metadata may be stored in a relational database.

The authors, F. Fernández-Martínez, A. Hernández García and F. Díaz de María [6], presented a computational model capable to predict the viewer perception of car advertisements videos by using a set of low-level video descriptors as metadata. This proposed method relies on the hypothesis that these descriptors could reflect the aesthetic value of the videos and, in turn, their viewers' perception. To that effect, and as a novel approach to this problem, automatically annotate video corpus, downloaded from YouTube, by applying an unsupervised clustering algorithm to the retrieved metadata linked to the viewers' assessments of the videos.

One of the challenge today is to go for a automated precise alignment of high-level semantic features gained from the analysis of complementary sources and the low/mid-level features extracted by audio-video analysis, supporting at the end a real cross-media search through large archives or through the web [7]. In this regard, a work has been done by the authors T. Declerck, P. Buitelaar, M. Alcantara, M. Labský and V. Svátek [7], is adding semantic metadata to multimedia material, on the base of the results of the automatic analysis applied to associated language material, being speech transcripts or various types of textual documents related to video/image material.

As it is essential aspect to concentrate on metadata requirements while analyzing video contents, the authors Werner Bailer and Martin [8], reviewed the metadata requirements for multi-view video content and analyze how well these requirements are covered in existing metadata standards, both in terms of the coverage of metadata elements and the capabilities to structurally describe multi-view video content.

3. METADATA FOR WEB VIDEOS

In recent times there has been an explosion in the number of online videos. With the gradually increasing multimedia content, the task of efficient query-based video retrieval has become important. The proper genre or category identification of the video is essential for this purpose. The automatic genre identification of videos has been traditionally posed as a supervised classification task of the features derived from the audio, visual content and textual features, whereas some works focus on classifying the video based on the metadata (text) provided by the uploader, other works attempt to extract low-level features by analyzing the frames, signals, audio etc, along with textual features. There have been some recent advances in incorporating new features for classification like the social content comprising of the user connectivity, comments, interest etc. Therefore, to classify or categorize the web video effectively and efficiently, it is essential to analyze the descriptive metadata given by the uploader, website metadata like comments, tags, etc, and technical metadata. In this section web video metadata classes and repository model is proposed, which represents classification of metadata for web video as- technical metadata, web metadata and descriptive metadata. The proposed model is shown in Figure 1.

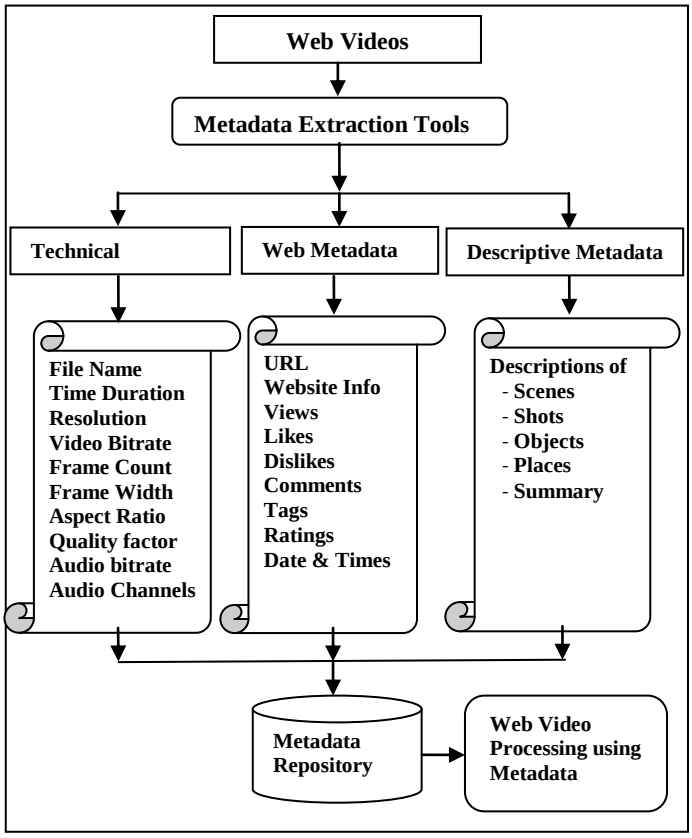

Figure 1: Web video metadata classes and repository

3.1 Technical metadata for web videos

The technical metadata of web videos allows for detailing properties of internal information of video data such as:

- Physical format for media dependent materials (e.g., stock brand, base and binder for magnetic tapes, disc surface type for disc media)
- Dimensions for media dependent materials (e.g., diameter, gauge, height, thickness, width, length)
- Generation for analog media.
- Frame description (e.g., pixel or line counts horizontal and vertical, frame rate)
- Data rate (may be expressed in terms of maximum, minimum, nominal, mode)
- Format in terms of broadcast standards (e.g., NTSC, PAL, SECAM)
- Formatting description for digital materials (e.g., name of creating application, commercial name of format, profile, version)
- Bits per sample and sampling description in terms of chrominance and luminance (e.g., 4:2:2, 4:2:0

3.1.1 Experiments with technical metadata extraction

Various experiments are made to extract effective technical metadata from web videos using different available tools such as MediaInfo, MooO, etc. Observations have been made through different experiments and extracted technical metadata of the web video titled "Comedy Nights with Kapil - Kangana Ranaut - 6th October 2013" from YouTube official website URL - http://www.youtube.com/watch?v=KxBKyI-7aPc and following technical metadata are extracted and shown in Table 1, which are stored in metadata repository for future use.

Table 1: Technical metadata for web videos

File Name: Comedy Nights with Kapil - Kangana Ranaut - 6th October 2013 - Full Episode (HD)-KxBKyI-7aPc.mp4
File Size (Bytes): 1,096,249,273
[Format]
Format(Short Name): mov,mp4,m4a,3gp,3g2,mj2
Format (Long Name): QuickTime / MOV
Duration: 00:57:57.65
Duration (Microseconds): 3,477,651,667
Bit Rate (bits/sec): 2,521,815
Number of Streams: 2
[Stream #0]
Type: Video
Codec (Short Name): h264
Codec (Long Name): H.264 / AVC / MPEG-4 AVC / MPEG-4 part 10
Codec Time Base: 1/50
Time Base: 1/50
Real Base Frame Rate: 25/1
Average Frame Rate: 25/1
Duration: 00:00:00.65
Duration (Microseconds): 173,882
Bit Rate (bits/sec): 2,327,360
Number of Frames: 86,941
Width: 1,280
Height: 720
Pixel Aspect Ratio: 0:1

3.2 Web metadata

The web metadata includes, information of URL of the video, website information, comments given by the users, total number of views, number of likes and dislikes, rating information, tags, date and time of uploading, subscription information etc. An attempt is made to extract the web metadata using InfoExtractor tool for the YouTube video URL: http://www.youtube.com/ watch?v= ZqgZFyYrYRQ.

Table 2: Web metadata for YouTube Video

Video data from Comedy Nights with Kapil - Sonakshi Sinha & Ranvir Singh - 29th June 2013 - Full Episode (HD) (http://www.youtube.com/watch?v= ZqgZFyYrYRQ&feature =youtube_gdata_player)	
URL: http://www.youtube.com/watch?v=ZqgZFyYrYRQ&feature =youtube_gdata_player	
Title: Comedy Nights with Kapil - Sonakshi Sinha & Ranvir Singh - 29th June 2013 - Full Episode (HD)	
Author: Colors TV	Avg. rating: 4.4669695
Duration: 3182	Comments: 561
Views: 4315747	Category: Shows
Ratings: 8810	Keywords:

The Table 2 shows the web metadata extracted for the above mentioned video. This information can be used for further analysis and even for the construction of the descriptive metadata.

3.3 Descriptive metadata

Descriptive metadata is different for different videos. The web video metadata helps users to search and personalize video data in efficient ways. Therefore it is necessary to give sufficient metadata to a video while uploading to the web. The one, who uploads the video to the web, has to give descriptive metadata clearly with complete detail. The descriptive metadata includes general descriptive metadata such as, Title, Alternative Title (optional), Creator, Contributor, Publisher, Date Created, Language, Abstract/Summary, Physical Description/ Sequence, Geographic Base (optional), Subjects, Genre, Table of Contents, Note, Identifier, Location and Rights, and event specific metadata and object specific metadata. To illustrate, the URL- http://www.youtube.com/watch?v=h5fW1zodoCg has selected from YouTube for descriptive metadata. Table 3 shows descriptive metadata extracted from the selected video.

Table 3: Descriptive Metadata for YouTube Video

Video data from Cute Babies Belgaum (http://www.youtube.com/watch?v= h5fW1zodoCg&feature= youtube_gdata_player) URL: http://www.youtube.com/watch?v=h5fW1zodoCg&feature=youtube_gdata_player	
Title: Cute Babies Belgaum **Description:** Author: Rashi B Published:2014-10-15T04:40: 48.000Z	Category: People Keywords: Duration:69: http://i1.ytimg.com/vi/h5fW1zodoCg/0.jpg

It is observed from the extracted metadata that, there is no sufficient descriptive metadata for the selected web videos. In this case, based on technical metadata, web metadata and descriptive metadata it is very difficult to conclude the category of the video. Further, it is difficult, to personalize, customize the video, to retrieve specific event, object of the video using metadata and for research purpose. Therefore we strongly recommend that, the video uploader has to give sufficient descriptive metadata. A number of professional industry and consortia standardization efforts are in progress to provide more detailed video descriptors. The new member of the MPEG family, Multimedia Content Description Interface, or MPEG-7 [11], aims at providing standardized core technologies allowing description of audiovisual data content in multimedia environments. Also it has been observed that there exists poor descriptive metadata for web videos as majority of the web videos doesn't have rich descriptive metadata. To overcome from this problem, to enrich descriptive metadata of web videos, we propose new descriptive metadata model which has 3 modules described in section 4.

4. PROPOSED DESCRIPTIVE METADATA MODEL

An effective and efficient construction model for descriptive metadata for web videos is proposed in this section. The proposed model is shown in the Figure 2 and it consists of three modules namely- General Descriptive Metadata, Event Specific Metadata and Object Specific Metadata.

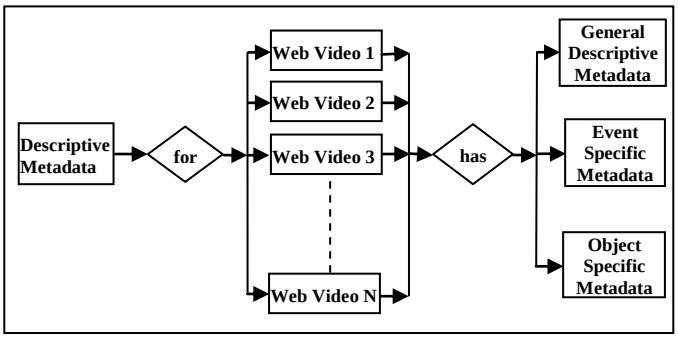

Figure 2: Proposed descriptive metadata model

4.1 Construction of General descriptive Metadata

The general descriptive metadata includes the following information.

- Information describing the title of the video/video file name.
- Information about the storage features of the content (storage format/encoding) .
- Information of releasing years and publishing year on the web.
- Creation and production processes of the content (director, short feature movie/video)
- Information related to the usage of the content and copy right information.
- Structural information of the content of various scenes.
- Summary of the video content.

The general descriptive metadata for a domain specific video, such as entertainment movie is shown in Table 4.

Table 4: Construction of General Descriptive Metadata

Metadata	**Data Type**	**Description**
Title	String	File name (Movie name)
Storage Format	String	Storage format of the movie (Ex- MP4,AVI)
Year	Date	Releasing Year
Director	String	Director of the movie
Actors	String	Actors information
Theme	String	Theme of the movie (Ex- Action, comedy etc)
Language	String	Movie language
Songs	Number	Number of songs in the movie
Fights	Number	Number of fights in the movie
Emotional	Number	Number of emotional scenes in the movie

Comedy	Number	Number of comedy scenes
Romance	Number	Number of romantic scenes
Summary	String	Brief summary of the video

4.2 Construction of Event Specific Metadata

To personalize web videos in depth for specific purpose, the descriptive event specific metadata can be constructed with timeline. With the help of timeline based descriptive metadata, users can extract specific part (in duration) or specific events (for ex- songs, fights etc) from the web video as it includes the following information.

- Conceptual information of the events.
- Information about collections of events.
- Information about category of the events.
- Summary of the events
- Timeline and frame information with respect to events.

The descriptive event specific metadata for a domain specific video, such as entertainment movie is shown in Table 5.

Table 5: Descriptive timeline metadata for an event in a video clip

Metadata	**Description**	**Value**
Event category	Category of the event (Ex- songs, discussion, fights, horror, funny etc)	String
Time	Total time duration of the song	Time Value
Starting time	In hh:mm:ss	Time Value
Ending time	In hh:mm:ss	Time Value
Starting Frame	Frame Number	Number Value
Ending Frame	Frame Number	Number
Summary	Brief summary of the event	String

In the similar way, it is possible to give descriptive timeline metadata for fighting events, comedy events, emotional events etc.

4.3 Construction of Object Specific Metadata

It has been observed that, in the existing web videos object description model, the object specific metadata is limited to very few properties. To enrich object specific descriptive metadata, we propose additional metadata for object specific descriptive metadata which helps to know about specific object and its properties as well and includes the following information.

- Conceptual information of the objects (interactions among objects)
- Information about collections of objects.
- Object caption, ID, motion information and brief summary as description of objects.

The descriptive object specific metadata for a domain specific video, such as entertainment movie is shown in Table 6.

Table 6: Construction of Object Specific Metadata

Metadata	**Description**	**Value**
Caption	Object name	String
ID	Object ID	Number
Category	Object category (Ex, animate, inanimate etc)	String
Motion	Motion of the object	Boolean
Shape	Shape of the object	String
Color	Color of the object	Color
Summary	Brief description of the object	Number

5. EXPERIMENTS AND RESULT DISCUSSION

The aim is to construct and enrich descriptive metadata and extraction of all the three type of metadata ie, technical metadata, web metadata and descriptive metadata from web video, the proposed method used two tools to extract the metadata namely, MooO Video Info which extracts technical metadata from a video (offline) and InfoExtractor [9] which extracts web metadata and descriptive metadata from web video. The MooO, a video information extraction tool is used to extract technical video metadata such as duration, resolution, length, frame information etc. The screen shot of the experimental result is shown in the Figure 3.

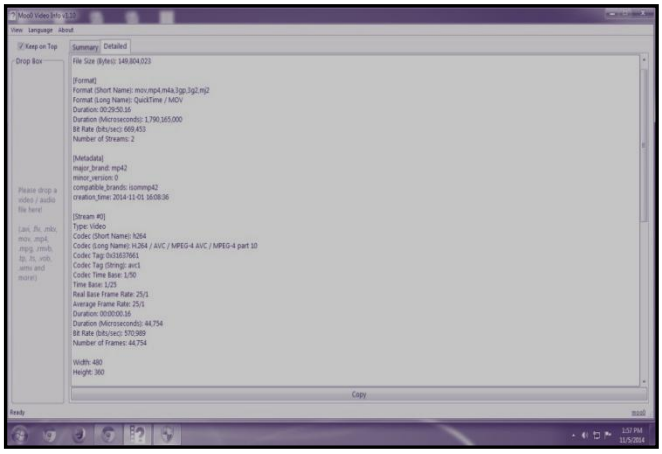

Figure 3: Technical metadata extraction using MooO

InfoExtractor was designed to permit social science researchers to retrieve data and metadata in structured formats from popular online media sites. It was designed in such a way that it would not be dependent on page scraping, and would free users from having to perform such scraping and copying themselves. It is designed to be maximally flexible in the data formats it provides, and as such is well-suited to any number of quantitative research approaches and methods. InfoExtractor provides important capabilities to social science researchers interested in studying online media. It can provide structured data that can be easily manipulated and analyzed [9]. Based on the proposed descriptive metadata construction model, descriptive metadata are constructed for videos and are uploaded to the YouTube.

To demonstrate the model, the metadata are extracted using InfoExtractor tool for constructed descriptive metadata and the experimental screen shots are shown in Figure 4, 5 6 and 7.

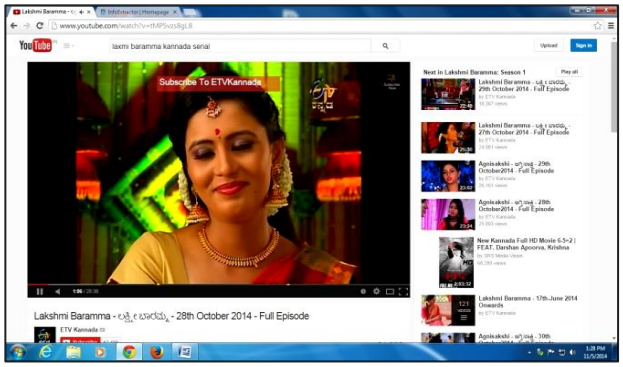

Figure 4: The metadata is constructed video is uploaded to the YouTube

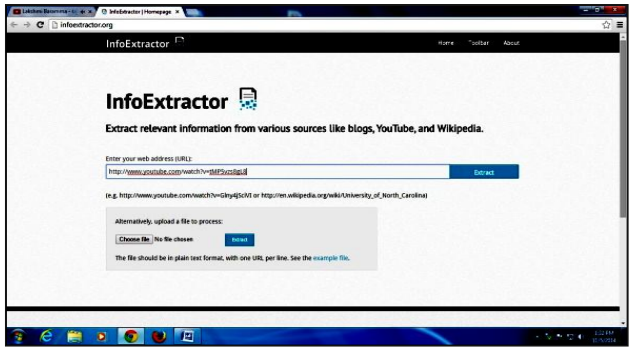

Figure 5: The metadata constructed video is used as input to the InfoExtractor.

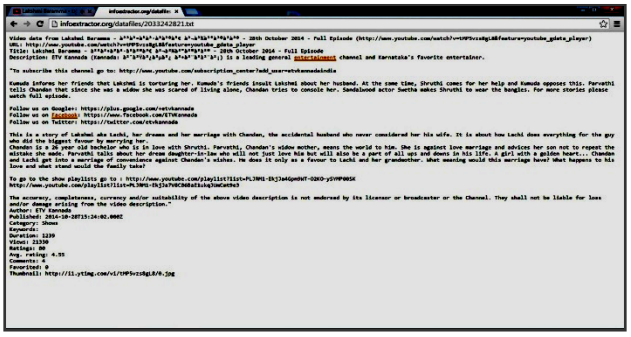

Figure 6: Web metadata and descriptive metadata are extracted in text format

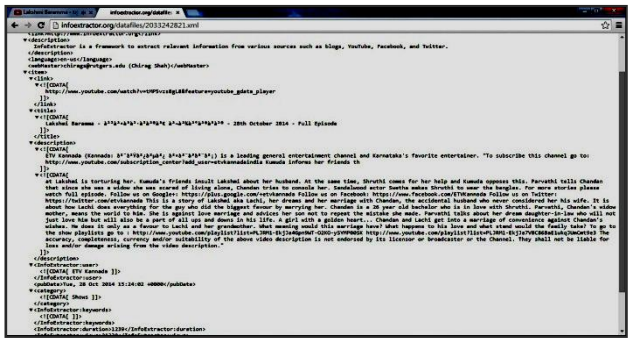

Figure 7: Web metadata and descriptive metadata are extracted in XML format

This demonstrates that, there is sufficient web metadata and descriptive metadata; it helps in categorizing the video from the web as one of the application of proposed descriptive metadata model. Further, all the three metadata – technical metadata, web metadata and descriptive metadata are combined and stored in metadata database to process the web video based on metadata, to personalize, customize the video, to retrieve specific event, object of the video using metadata. For example, in case of classification/categorization of web videos, web metadata and descriptive metadata plays important role. For instance, based on viewer's comments from web metadata and description of an event from descriptive metadata, it is possible to predict a video category/class. If comments of a web video contains the word "Cricket", "Batsman", "Runs" etc, and event descriptions contains descriptions about an innings, wicket fall, sixes etc, then the proposed model can predict that video class as sports category. The Figure 8 shows the classification/categorization of web videos as, Class 1 for entertainment videos, Class 2 for sports videos, and Class 3 for News/Discussion videos, and Class 4 for commercial videos and so on. All videos like movies, songs, dance, reality shows etc are fall into Class-1, all sports videos like cricket, hockey, football etc are fall into Class-2, whereas news, discussion, reports etc are fall into Class-3. Advertisement, business / commercial videos fall into Class-4 and so on.

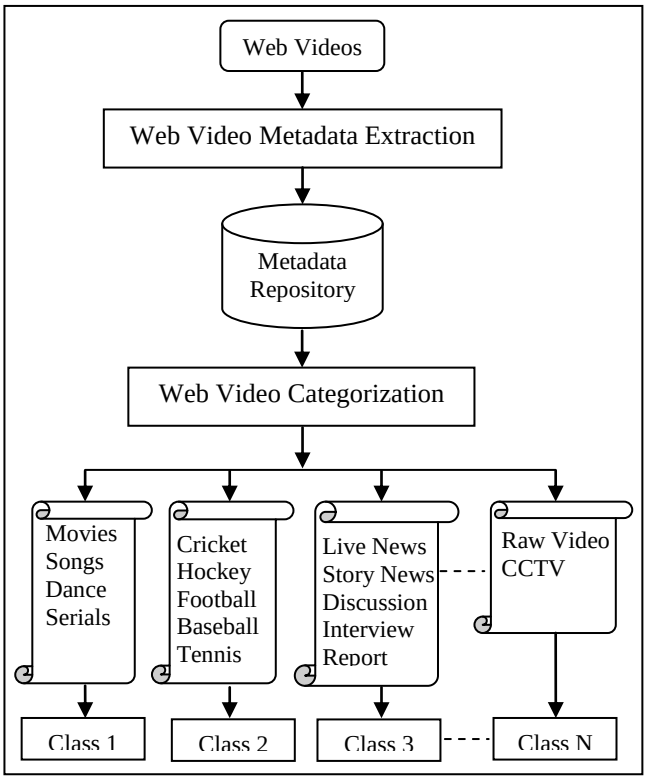

Figure 8: Categorization of web videos

6. CONCLUSION AND FUTURE WORK

In this paper, we described significance of web video descriptive metadata, presented an effective and efficient method for construction and extraction of web video descriptive metadata. The proposed method demonstrated the effectiveness of constructing the descriptive metadata with timeline for a domain specific web video. The papers also suggested the construction of event specific and object specific metadata and which are considered to be very effective. Using proposed descriptive metadata model, users may process the video contents effectively and efficiently. The future work includes, classification, clustering and predictive analysis of web videos based on descriptive metadata.

REFERENCES

[1]. Dim P. Papadopoulos , Vicky S. Kalogeiton, Savvas A. Chatzichristofis, Nikos Papamarkos, "Automatic Summarization And Annotation Of Videos With Lack Of Metadata Information", 2013 Elsevier, Expert Systems With Applications 40 (2013) 5765–5778.

[2]. "Data Dictionary For Preservation Metadata", Final Report Of The PREMIS Working Group, May 2005.

[3]. W. Klippgen, T.D.C. Little, G. Ahanger, And D. Venkatesh, "The Use Of Metadata For The Rendering Of Personalized Video Delivery", Multimedia Communications Laboratory, 1996.

[4]. Juan Casares, A. Chris Long, Brad A. Myers, Rishi Bhatnagar, Scott M. Stevens, Laura Dabbish, Dan Yocum, And Albert Corbett, "Simplifying Video Editing Using Metadata", DIS2002, London © Copyright 2002

[5]. D. M. Shotton, A. Rodríguez, N. Guil And O. Trelles, "A Metadata Classification Schema For Semantic Content Analysis Of Videos", Journal Of Microscopy, Vol. 205, Pt 1 January 2002

[6]. F. Fernández-Martínez , A. Hernández García, F. Díaz De María, "Succeeding Metadata Based Annotation Scheme And Visual Tips For The Automatic Assessment Of Video Aesthetic Quality In Car Commercials", 0957-4174/_ 2014 Elsevier Ltd.

[7]. T. Declerck, P. Buitelaar, M. Alcantara, M. Labský, V. Svátek, "Adding Semantic Metadata To Audio-Video Material By Automatic Analysis Of Complementary Sources"

[8]. Werner Bailer And Martin H.O_Ernig, "Metadata For Creation And Distribution Of Multi-View Video Content"

[9]. Chirag Shah, Charles File, "Infoextractor – A Tool For Social Media Data Mining", JITP 2011.

[10]. Jurandy Almeida , Neucimar J. Leite, Ricardo Da S. Torres, " Online Video Summarization On Compressed Domain", Elsevier , J. Vis. Commun. Image R. 24 (2013) 729–738

[11]. Website Reference: Http://Www.Lib.Utexas.Edu/Schema /Video_Metadata_Guidelines_V1.Pdf

Author Profile

Dr. Siddu P. Algur is working as Professor, Dept. of Computer Science, Rani Channamma University, Belagavi, Karnataka, India. He received B.E. degree in Electrical and Electronics from Mysore University, Karnataka, India, in 1986. He received his M.E. degree in from NIT, Allahabad, India, in 1991.He obtained Ph.D. degree from the Department of P.G. Studies and Research in Computer Science at Gulbarga University, Gulbarga. He worked as Lecturer in the department of E & E at KLE Society's College of Engineering and Technology from September 1986 to October 1992.

From October 1992 to August 2007 he worked as Lecturer and Assistant Professor in the Department of Computer Science and Engineering at SDM College of Engineering and Technology, Dharwad. His research interest includes Data Mining, Web Mining, and Information Retrieval from the web and Knowledge discovery techniques. He has published more than 40 research papers in peer reviewed International Journals and chaired the sessions in many International conferences.

Mr. Prashant Bhat is pursuing Ph.D programme in Computer Science at Rani Channamma University Belagavi, Karnataka, India. He received B.Sc and M.Sc (Computer Science) degrees from Karnatak University, Dharwad, Karnataka, India, in 2010 and 2012 respectively. His research interest includes Data Mining, Web Mining, web multimedia mining, and Information Retrieval from the web and Knowledge discovery techniques.

He has attended and participated in International and National Conferences and Workshops in his research field.

Mr. Suraj Jain is pursuing Ph.D programme in Computer Science at Rani Channamma University Belagavi, Karnataka, India. He received BCA and MCA degrees from Indira Gandhi National Open University, New-Delhi, India, in 2005 and 2007 respectively. His research interest includes Data Mining, Web Mining, and web multimedia mining. He is also working as lecturer and PG coordinator, JSS College Dharwad. He has 13 years of professional experience in training and academic background.