Extending AVF for Effective Road-surveying in the Web Environment

Srdjan Mihic
Faculty of Technical Sciences, University of Novi Sad, Republic of Serbia
smihic@uns.ac.rs

Dragan Ivetic
Faculty of Technical Sciences, University of Novi Sad, Republic of Serbia
ivetic@uns.ac.rs

ABSTRACT

This paper presents the results of the research made on adoption of Augmented Video Framework (AVF) to support effective road-surveying in the Web environment. The AVF augments road-condition state videos with non-perceptible data acquired by measuring devices (gyroscope, GPS receivers etc.). The videos and the data are provided by ROad Measurement and Data Acquisition System (ROMDAS). Current AVF implementation was created for offline viewing and was intended for internal use. The advantages of the AVF approach encouraged road engineers to request extension of the AVF to support road-surveying and management in the Web environment. Road-condition state augmented videos are now intended to be used for Web consulting purposes. Extending the AVF for Web consulting is not an easy task. In the Web environment various restrictions and challenges exist (e.g. video streaming, security, client-server responsibilities) that have to be taken into account when extending the AVF. Various factors and possible solutions have been discussed. Finally, necessary extensions and modifications of the AVF were presented.

Keywords: Web consulting, AVF, ROMDAS, augmented video, ontologies

1. Introduction

Two years ago we were kindly asked by the public institution Road Center of Vojvodina (Centar za puteve Vojvodine, CPV) to make the planning and maintenance of the road network maintained by CPV a more comfortable task. The road network consists of approximately 17.000 km of roads.

CPV is using ROad Measurement and Data Acquisition System (ROMDAS) [1] for road inspection and maintenance. The ROMDAS system consists of several measuring devices (gyroscope, GPS receivers etc.), a video camera mounted on a vehicle, and software to process the collected discrete data. The video camera captures video into AVI [2] format. The measuring devices capture the discrete data about the physical characteristics of the road-condition state such as: road roughness, transverse profile and rut depths, traffic density etc. After the completed survey run, the measured data are processed and analyzed.

Videos are stored separately from the corresponding data. Thus, road engineers have to search the videos manually in order to find details of interest provided by data analysis. Road engineers wanted to be able to modify the way the augmented data is presented in order to customize the presentation to their needs. Also, they wanted to be able to easily transfer their field-work to home and vice versa. Thus, we concluded that the augmented video should be a self-contained entity allowing the full data search according to data properties.

ROMDAS road-condition state videos were augmented with non-perceptible discrete data acquired by ROMDAS measuring devices. The augmented data were encapsulated in an object-oriented manner and integrated into the AVI file while maintaining compatibility [3].

The Augmented Video stream Framework (AVF) enables creation, search and playback of augmented AVI files for effective road surveying and maintenance [4].

The AVF architecture (Fig. 1) is based on three components: Serialization Management, Presentation and Search.

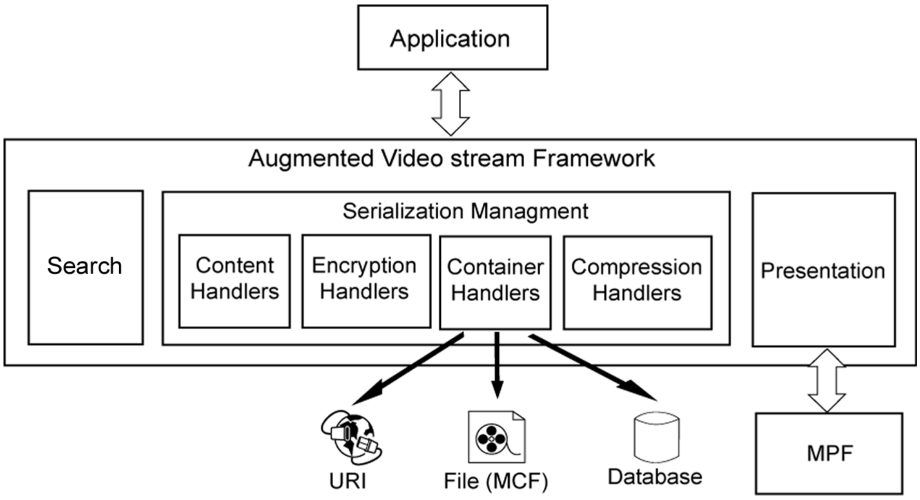


Fig. 1. The AVF architecture

The Serialization Management is dedicated to: data management, access rights management, encryption and compression of augmented data. All of its functionality is realized as plug-in components enabling usage of arbitrary algorithms that are most suitable for concrete situation.

The Presentation component provides synchronized playback of augmented video using a commercial multimedia presentation framework (MPF). Similar to the Serialization Management component, it supports use of arbitrary MPF which are implemented as plug-ins. The consumers can turn off the augmented data presentation entirely or partially by specifying which classes and their attributes they would like to get presented. Also, they can change position, size and zoom concerning the augmented data presentations. Fig. 2 shows some examples of customization of augmented data presentations.

The Search component provides the ability to search augmented data which are structured in an object-oriented manner. The Search component does not advocate any particular search algorithm.

The data structures used for augmented data encapsulation and the AVF type system were presented in [3]. Modularity and extensibility of the AVF makes it applicable on a wide range of information systems.

Current AVF implementation uses Microsoft DirectShow [5] for synchronized playback of augmented videos. Microsoft DirectShow was chosen because it achieves the best performance on Windows (target) platform [6].

Currently, AVI is used as implementation MCF. The AVI format was chosen since ROMDAS provided videos in this format and it was adequate for internal use. A simple augmented video player was created for playback and search of augmented videos.

Current implementation was intended for internal use, thus it was optimized for offline

a) no augmented data are presented

b) partial augmented data presentation: GPS coordinates not presented, data presentations rearranged and resized

c) full augmented data presentation without customizations

Fig. 2. Examples of customization of augmented data presentation

viewing and manipulation. The advantages of the AVF approach stimulated the CPV to request extension of the AVF to support road-surveying and management in the Web environment. Road-condition state augmented videos are now intended to be used for Web consulting purposes. Road-condition state augmented videos can be used for consulting purposes of wide audience: from ordinary users and professional drivers interested in road-condition state, traffic density and possible traffic jams in order to plan their safest route, to road engineers to whom this approach gives more mobility and freedom in their work.

This paper presents the research that was conducted in order to extend the AVF to support road-surveying Web consulting.

2. WEB environment chalenges

Web environment poses various restrictions and demands various modifications on current AVF implementation. One must take into consideration various factors such as: video streaming, client-server responsibility, video compression, security, client consumption properties etc..

2.1 Augmented video streaming

The Web environment forces streaming on MCF, used for augmented video. A suitable MCF has to support streaming video playback over networks and interleaving of the augmented data with corresponding video and audio streams.

The AVI MCF was originally not intended for streaming. AVI was developed for playback of audio and video from hard disks and CD-ROMs on personal computers. It is adequate for downloading a video file from a remote site on the Internet for subsequent playback from the computer's hard drive. It is not well suited for real-time or streaming video playback over networks, because there is no mechanism to resynchronize audio and video streams [7].

A modification of AVI that supports streaming was presented in [7]. Unfortunately, both the original [2] and the AVI with streaming support [7] do not support interleaving of augmented data while maintaining compatibility. Thus, the augmented data have to be inserted at the end of the file [4].

Therefore, another MCF has to be chosen in order to support augmented video streaming in the Web environment.

A comprehensive study of commercial MCFs was conducted in order to find suitable MCF for augmented data encapsulation while maintaining compatibility. Following MCFs were examined: AVI, Matroska [8], QuickTime file format [9], Ogg file format [10], ASF [11], MP4 [12], RealMedia file format [13] and DivX media format [14]. MCFs were compared according to: supported platforms and codecs (COder/DECoder), suitability for augmented data encapsulation, interleaving and streaming capability. The support of arbitrary codec enables the use of the most suitable codec for the concrete situation. RealMedia and DivX MCF support only the use of their proprietary codecs. All the other MCFs support arbitrary codecs. The platform support includes ability to create and playback of MCF. Only AVI, Ogg and RealMedia have full support on all major platforms. This study showed that all of examined MCFs have similar features and almost all of them are suitable for augmented data encapsulation and storage. Support for augmented data encapsulation could not be determined for MP4, DivX, and RealMedia since full specifications were not publically available. All of the examined MCFs except AVI support streaming and interleaving of the augmented data. Table 1 shows comparison of examined MCFs [4].

According to this study we have chosen Apple’s QuickTime file format. It was chosen because it supports streaming and interleaving of the augmented data, and the use of arbitrary codecs. It is supported on two major platforms and suitable for editing, as it is capable of importing and editing in place [9]. This provides means for easy modification of the augmented data without the need for data copying. Architecture of QuickTime file format is highly suitable for the discrete augmented data encapsulation, and provides support for interactive playback of augmented videos. Also, QuickTime file format is MPEG-4 compatible which provides interoperability with other MPEG-4 based systems.

Table 1. Comparison of MCFs

MCF	Supported codecs	Supported platform	Streaming and interleaving	Suitable for aug. data
AVI	All	Windows, Linux, Macintosh	No[1]	Yes
Matroska	All	Windows, Linux	Yes	Yes
QuickTime file format	All	Windows, Macintosh	Yes	Yes
Ogg file format	All[2]	Windows, Linux, Macintosh	Yes	No
ASF	All	Windows, Macintosh	Yes	Yes
MP4	All[3]	Windows, Macintosh	Yes	?[4]
Real Media format	RealVideo and RealAudio	Windows, Linux, Macintosh	Yes	?[4]
DivX media format	DivX codec ver. 6.x	Windows, Linux[5], Macintosh[5]	Yes	?[4]

[1]Implementations with streaming support exist [7]
[2]Codecs defined by Xiph.org are preferable
[3]Codecs defined by MPEG-4 standard are preferable
[4]It can not be determined because the specification is not publically available
[5]On Linux and Macintosh only playback is supported

2.2. Multimedia presentation framework's platform independence

The platform independence of MPF is desirable in the Web environment. Thus, the AVF presentation component should be platform independent or at least supported on several platforms. The disadvantage of this approach is that achievable performance is lower compared to MPFs that are created for one specific platform [6]. Current AVF implementation uses Microsoft DirectShow which is supported only on Windows platform, thus another MPF has to be chosen.

Alternatively, presentation component could be implemented for every supported platform. This approach achieves the highest performance because the best MPF for every platform is used. It requires only a slight modification of the current presentation component on Windows platform, but for every other platform it has to be implemented from scratch.

Several commercial multiplatform MPFs exist. Popular ones are: Sun's Java Multimedia Framework (JMF) [15], Apple's QuickTime [9], and GStreamer [16]. All of them possess similar features and architecture, and are extensible. They differ by their performance, supported MCFs and platforms. JMF is a MPF based on Java programming language. It supports playback of only a few MCFs. JMF is the only MPF supported on all major platforms (Windows, Linux/Unix and Macintosh). Sadly, JMF achieves the lowest performance, because it is dependant both on the underlying operating system and the Java Virtual Machine. GStreamer is an open source MPF designed for Linux/Unix platform. It was ported to other major platforms. Although GStreamer is fast and lean it is not stable API and supports only a few MCFs. QuickTime framework is supported on Macintosh and Windows platform and offers the best performance on Macintosh and the second best performance on Windows platform according to [6]. Also, only QuickTime MPF enables playback of all commercial MCFs. Table 2 shows comparison of presented MPFs.

Few factors have to be taken into account when choosing an MPF: platform independence, variety of supported MCFs, documentation and support, and performance. QuickTime MPF was chosen because it is multiplatform with the second best achievable performance, supports playback of chosen MCF and has excellent documentation and support.

Table 2. Comparison of MPFs

MPF	Supported MCF	Supported platform	Performance	Doc. and support
QuickTime	All	Windows, Macintosh	High	Excellent
Java Multimedia Framework	AVI, RealMedia, QuickTime	Windows, Linux, Macintosh	Low	Good
GStreamer	AVI, Ogg, Matroska	Linux, Windows[1], Macintosh[1]	High	Sufficient

[1]GStreamer was ported to Windows and Macintosh platforms

2.3. Client-server responsibilities

Depending on the nature and changeability of the augmented data, two strategies for augmented video storage are common: store-oriented and retrieval-oriented (Fig. 3). ANNODEX platform implements both strategies [17].

Store-oriented strategy (Fig. 3a) creates and plays augmented video on user request. When a client requests augmented video from server, the chosen video is augmented with the corresponding data and is then sent to the client. This strategy is suitable in situations where data frequently change. It requires substantial processing power on the server side to create the augmented video within an acceptable time interval. A large number of user requests is very demanding for following components of the media server: content directory, memory manager and the file system, and significantly degrades server's overall performance.

In contrast, in retrieval-oriented strategy (Fig. 3b) augmented video is created offline and inserted into the database. On user request the chosen augmented video is sent to the client. This strategy is suitable for applications where data are rarely changed and therefore the augmentation can be performed offline. Here, required processing power and retrieval time are significantly reduced compared to store-oriented strategy. High compression ratio is achievable because the augmentation is performed offline. In store-oriented strategy compression has to be performed within an acceptable time interval and thus high compression ratio can not be achieved. Thus, retrieval-oriented strategy should potentially create smaller augmented video files to be transferred over the network.

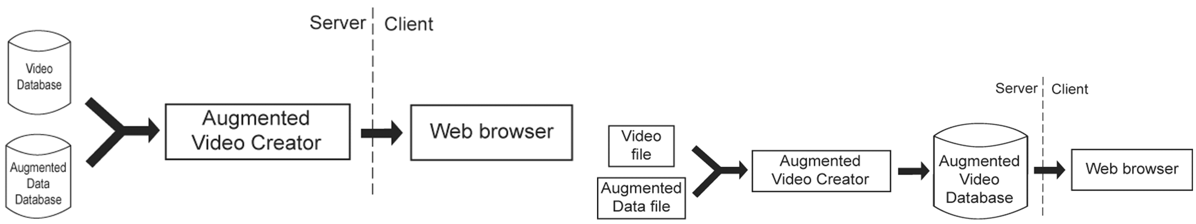
Retrieval-oriented strategy was chosen because the ROMDAS data are unchangeable; they are acquired at the time of video capture. It provides a more effective and responsive system since the client's request is served directly from the database.

In current AVF implementation (stand-alone system) consumers can turn off the augmented data presentation entirely or partially by specifying which classes and their attributes they would like to get presented – users can filter augmented data presentations. Also, they can customize the augmented data presentation to suit their needs by changing position, size, and zoom of the augmented data presentations during the playback. Customizaiton of augmented data presentation during the playback is hardly achievable in the Web environment, because of the many user requests that have to be realized in acceptable time interval. Therefore, we decided not to support live customization of the augmented data presentations.

2.4 Augmented data security and user authorization

For security reasons, it has been required to restrict access to augmented data while maintaining compatibility. Security and user authorization in current AVF implementation are based on augmented data encryption. A user who wants to access the augmented data has to provide adequate key in order to decrypt the augmented data. Authorized users can watch, search and change the way the augmented data are presented during the video playback. On the other hand, unauthorized users can watch the video only.

Our previous solution based authorization on encryption, but in the Web environment this could be easily transferred into server's jurisdiction. Since client-server architecture is used, users can authorize themselves at the server and the augmented data encryption can be seen as another security measure. Users can be classified as follows by their access rights to the augmented data:



a) Store-oriented strategy b) Retrieval-oriented strategy

Fig. 3. Augmented video client-server strategies

- Unauthorized users – who can watch only the video without ability to watch the augmented data.
- Authorized users with limited rights – who can watch augmented data and customize augmented data presentations during the playback.
- Advanced users - who have all the rights of restricted users and the ability to issue interval queries for the augmented data.

Another form of user classification can be realized according to semantic rights (e.g. user can watch augmented videos of one particular region).

User classification provides that all users can watch the video, and only the authorized ones can enjoy the benefits of augmented videos.

2.5 Augmented video ontology

Current implementation of AVF data structures was focusing on effective data encapsulation in an object-oriented manner [3]. The inclusion of ontologies would be interesting in the Web environment. Ontologies do not only describe the data, but also establish relationships between the data that would enhance search capabilities [18]. Ontology is one of the major means in achieving the vision of the Semantic Web [19] – the next generation of the Web, intended to significantly improve experience of web application users [20]. The use of ontologies would enable establishment of complex relationships between the video and the augmented data and even between similar videos (e.g. road-condition state videos where vertical signalization shows trend toward constant damage over the last five years). Ontologies allow definition of constraints in relations and events. Some simple forms of relations already exist in the current implementation (e.g. successor, predecessor, hasAttribute, isClass…). Unfortunately, current implementation lacks more complex relationships.

Ontology can be defined as an explicit specification of conceptualization [20]. This means that ontology is a formal representation of a set of concepts within a domain and the relationships between those concepts. It is used to reason about the properties of that domain, and may be used to define the domain. Ontology usually consists of: individuals (instances), concepts (classes), attributes, relationships and events. Instances, classes, and attributes denote similar concepts in object-oriented programming. Typically, a relationship is an attribute whose value is another object in the ontology [21].

Ontologies can be classified as: domain and upper ontologies. A domain ontology (or domain-specific ontology) models a specific domain. It represents the particular meanings of terms in that domain (e.g. bug in computer domain has meaning of error, but in biology denotes a class of living organisms). Common issues in domain ontologies include incompatibility, very hard or even impossible extensibility and manual process of creation. Different ontologies in the same domain can also arise due to different perceptions of the domain based on cultural background, education, ideology, or because a different representation language was chosen [20]. An upper ontology (or foundation ontology) is a model of the common objects that are generally applicable across a wide range of domain ontologies. It contains a core glossary in which terms the domain objects can be described. Upper ontology can serve as base for building domain-specific ontologies. Ontology classification permits ontologies that belong in both classes (e.g. The Gellish ontology [21]). Variety of ontology languages was published. Most popular are: OWL [22] and Gellish [21].

Various video systems based on ontologies were published [23, 24]. These systems use ontologies to describe the video content. They are not suitable for description of augmented vides because they provide means only to describe the video. We plan to use ontology for description of both the video and the augmented data. Thus, a novel ontology (domain and upper) has to be designed.

3. Conclusion

In this paper we presented the results of the research made on adoption of Augmented Video Framework (AVF) to support effective road-surveying in the Web environment. Current AVF implementation was created for offline viewing and manipulation of augmented videos. The advantages of the AVF approach encouraged road engineers to request extension of the AVF to support road-surveying and management in the Web environment. Road-condition state augmented videos are now intended to be used for Web consulting purposes.

Extending the AVF for Web consulting is not an easy task. In the Web environment various restrictions and challenges exist that have to be taken into account when extending the AVF.

The necessity for video streaming and platform independence caused the change in storage and presentation components. Currently, AVI multimedia container format (MCF) is used for augmented data encapsulation and storage, but it does not support streaming or interleaving of the augmented data. Therefore, another MCF was chosen based on our comprehensive study of MCFs. QuickTime file format was chosen because of following reasons:

- support for streaming and interleaving of the augmented data,
- support of arbitrary codecs,
- a MPEG-4 compliant,
- editing in place,
- well documented and supported on two major platforms.

Due the lack of platform independence, Microsoft DirectShow, the current multimedia presentation framework (MPF), had to be changed with QuickTime MPF. QuickTime MPF has shown as better than the other MPFs that we considered. Its advantages are:

- the best possible performance on supported platforms,
- chosen MCF supported,
- well documented and supported.

We have chosen retrieval-oriented strategy since the augmented data is unchangeable (they are acquired at the time of video capture).

We had to restrict the interaction to filtering of the augmented data presentations in order to achieve effective interactivity with the augmented data presentations in the Web environment.

Security was enhanced by the use of user authorization on server and encryption of the augmented data. Users were classified both by their access rights and semantic rights.

Use of ontologies was intended for better description and more effective search of augmented videos. Existing video ontologies are not suitable since they only allow the description of the video. Therefore, it was decided to design a new ontology that would provide means to describe the video and the corresponding augmented data, as well as the relationships with other augmented videos.

Our approach has several advantages:

1. Augmented video is self-contained document, referenced by a single URL. This natural and permanent association prevents mismatch between the video and the corresponding data and maintains the binding even when streaming.
2. Use of ontologies provides a formal system for description of both the video and the augmented data and complex relationships among them.
3. Users can interact with the augmented data presentations and adapt them to their needs. Users can turn off the augmented data presentation entirely or partially by specifying which classes and their attributes they would like to get presented.

This paper presented challenges and possible solutions in order to adapt the AVF for the road-surveying Web consulting.

4. Acknowledgment

This research has been supported by IT project: “WEB portals for data analysis and consulting”, No.13013, financed by the Government of Republic of Serbia.

References:

[1] C. R. Bennett, A. Chamorro, C. Chen, H. De Solminihac, and G. W. Flintisch, “Data Collection Technologies for Road Management”, The World Bank, East Asia Pacific Transport Unit, Washington, D.C. Tech. Rep. TRN-30, Feb. 2007.

[2] CORPORATE Microsoft Corp. “Microsoft Windows multimedia programmer's reference”, Microsoft Press, WA: Redmond, USA, 1991, ISBN: 1-55615-389-9.

[3] S. Mihic, D. Ivetic. (2008). Data Structures for Road Condition Avi File Video Augmentation (2008). 0857-0858, Annals of DAAAM for 2008 & Proceedings of the 19th International DAAAM Symposium, ISBN 978-3-901509-68-1, ISSN 1726-9679, pp 429, Editor B. Katalinic, Published by DAAAM International, Vienna, Austria 2008.

[4] S. Mihic. “Augmented Video stream Framework”. MSc thesis (in Serbian),

Faculty of Technical Sciences, Serbia: Novi Sad, 2007.
[5] D. M. Pesce. “Programming Microsoft DirectShow for Digital Video and Television“, Microsoft Press, WA: Redmond, USA, 2003, ISBN: 9780735618213.
[6] R. Kroeger. (2004). Admission Control for Independently-authored Realtime Applications, PhD thesis, [Online]. Available: http://www.cgl.uwaterloo.ca/~rjkroege/robert_kroeger_phd.pdf
[7] R. Zimmermann, “Streaming of DivX AVI movies”, in *Proc. of the 2003 ACM Symposium on Applied Computing*, 2003, pp. 979-982.
[8] Matroska Official Homepage, [Online]. Available: http://www.matroska.org
[9] T. Monroe. “QuickTime Toolkit, Volume Two: Advanced Movie Playback and Media Types”, 1st ed., Morgan-Kaufmann Publishers Inc., CA: San Francisco, USA, ISBN: 0-12-088402-X
[10] S. Pfeiffer. “The Ogg Encapsulation Format Version 0”, RFC Editor, USA, 2003
[11] S. McEvoy. “Fundamentals of Programming the Microsoft Windows Media Platform“, Microsoft Press, WA: Redmond, USA, 2003, ISBN:0735619115.
[12] F. C. Pereira, and T. Ebrahimi. “The Mpeg-4 Book“. Prentice Hall PTR., NJ: Upper Saddle River, USA, 2002, ISBN:0130616214
[13] J. Angel. “Realmedia Complete: Streaming Audio and Video Over the Web“. McGraw-Hill, Inc., NY: New York, USA, 1998
[14] U. Bruegmann. “DivX R.T.F.M. - DivX 6“, Lulu.com, 2006, ISBN:1847286763
[15] L. deCarmo. “Core Java Media Framework”. Prentice Hall PTR, NJ: Upper Saddle River, USA, 1999, ISBN: 0130115193
[16] E. Walthinsen. “GStreamer - GNOME Goes Multimedia”, Tech. Rep. GUADEC 2001, April 2001.
[17] S. Pfeiffer, C. Parker and C. Schremmer. ”Annodex: A Simple Architecture to Enable Hyperlinking”, in *Proc. of the 5th ACM SIGMM international Workshop on Multimedia information Retrieval*. 2003. pp. 87-93.
[18] M. Uschold and M. Gruninger. „Ontologies: principles, methods, and applications“, *Knowledge Engineering Review*, vol. 11 no. 2, pp. 93-155, 1996.
[19] G. Antoniou and F. Harmelen. “A Semantic Web Primer“, 2nd ed., The MIT Press., MA: Cambridge, USA, ISBN:0262012421
[20] M. Korotkiy. “On the Effect of Ontologies on Web Application Development Effort“, in *Proc. of the Knowledge Engineering and Software Engineering workshop*, 2005.
[21] A. Van Renssen. “Gellish: A generic Extensible Ontological Language Design and Application of a Universal Data Structure“, IOS Press/Delft University Press, Amsterdam, The Netherlands, 2005, ISBN: 978-90-407-2597-5
[22] I. Horrocks and P. F. Patel-Schneider. „Reducing OWL entailment to description logic satisfiability“, in *Proc. of the 2nd International Semantic Web Conference (ISWC)*, 2003.
[23] M. G. Christel and R. M. Conescu. “Addressing the challenge of visual information access from digital image and video libraries“, in *Proc. of the 5th ACM/IEEE-CS Joint Conference on Digital Libraries*, 2005, pp. 69-78.
[24] A. R. Francois, R. Nevatia, J. Hobbs and R. C. Bolles. “VERL: An Ontology Framework for Representing and Annotating Video Events“, *IEEE MultiMedia*, vol. 12 no. 4, pp. 76-86, Oct. 2005.