# Modelling and Design of VAML

**Congrong LIU[1] and Jesse S. JIN[1,2]**

[1]School of Information Technologies
University of Sydney, NSW 2006
[2]School of Computer Science & Engineering
University of NSW, Sydney 2052

crl@it.usyd.edu.au

## Abstract

Non-linear interactive access to video is very important in video information management. The term non-linear refers to that there are more than one video streams delivering to users. To access video in a non-linear interactive way, it will involve issues such as: hyper-link to the external text source, synchronization between audio and video, hyper-link to external video, parameter passing, buffers for multiple streams of video, segmentation of video objects, representation of video sequences and the organization of source data. It is significant to study these issues clearly and develop a video annotation mark-up language (VAML). This paper introduces the main idea of VAML.

*Keywords*: non-linear interactive access, annotation, segmentation, embedding marks, identify marks, video function, parsing, browser, VAML.

## 1 Introduction

Video as an entertainment or information source in consumer, military, and broadcast television applications is widespread. Typically however, the video is simply presented to the viewer, with only minimal manipulation. Examples include chroma keying (often used in news and weather broadcasts) where specific colour components are detected and used to control the video source. In the past few years, the advent of digital video and increases in computational power has meant that more complex manipulation can be performed. An effective and efficient visual data management system is highly desired. Recent technology in digital video processing has moved to "content-based" storage and retrieval. Nitsuwat and Jin (1999) consider the contents of a video and investigate an algorithm for sub-shot extraction and key-frame selection to represent the whole video in panoramic mosaic-based representation form. Nitsuwat et al. (1999) present a scheme for extracting moving objects. Keith et al. (1998) present some highlights in annotating video by aligning features extracted from the video to a reference set of features. Wilcox and Boreczky (1998) describe a method for indexing and retrieval of multimedia data based on annotation and segmentation. Miyamori and Iisaku (2000) propose an automatic annotation method of sports video for content-based retrieval. Sudhir and Lee (1996) present an approach to automatic annotation of video sequences by dominant camera motion interpretation. Harrison and Baecker (1992) discuss user requirements for video analysis. Adjeroh and Lee (1995) present some mechanisms to ensure synchronization in distributed multimedia presentations, in the presence of multiple remote sources, limited network bandwidth, statistical network delays, and asynchronous user interaction. Kozuch et al. (1996) present the first analyses of video server performance under various non-linear video application loads. Liao and Li (1997) describe a new protocol, called Split and Merge (SAM), which offers true VOD (Video-on –Demand) services while allowing multiple users to share the same video stream. Park and Ryou (1998) propose a scheme that can support an interactive VCR for all users requesting the same video stream with batching. Christel et al. (2000) propose interactive maps for a digital video library to improve library access. Majumder et al. (1997) propose an algorithm to create panoramic frames for a teleconference session that are both geometrically registered and intensity blended, and so on.

Although there is plenty of research work on the video processing, there still exist many problems for non-linear interactive access to video. That is why I choose this project theoretically.

## 2 Applications

The applications of this project can be described as following:

Firstly, It can be used in the distance education. For distance education, we can annotate the video to include the intra-relation of video frames and inter-relation between videos and other media (such as audio, text, 3d animation, etc). So, the students can only select the specific subject but not the rest of video and pace themselves, and avoid learning knowledge passively.

Secondly, it can be used as internet video transfer. Navigating/browsing through a large collection of video titles requires a larger degree of non-linear video service. To browse and navigate content-based digital video collections, non-linear and non-sequential access into video documents is essential. After segment video into a group of video streams and annotate them, we can transfer the video streams that we want only.

Thirdly, it can be used in immersive video. In order to create a single seamless panorama about an object through limited video camera, we can interpolate the

whole image through some existing information, and even zoom in or zoom out to get different image.
Finally it can be even used in business, such as fashion shop. After click different parts of model's wearing that represents the goods of fashion shop possess, we can get the relevant information, which we want to know.

## 3 Design Issues

The term non-linear refers that there is more than one typical ordering of video frames delivered to users. To access video in non-linear interactive way, it will involve such issues: hyper-link to the externa text source, synchronization between audio and video, hyper-link to externa video, hyper-link to 3D animation, parameter passing, buffers for multiple streams of video, segmentation of video objects, representation of video sequences and the organization of source data. So it is significant to study these issues clearly and develop a video annotation mark-up language, which include: algorithms for embedding marks, algorithms for identify marks, and modelling necessary video functions.

## 4 Implementation

After investigating these design issues, which was mentioned before, we will design a complete set of marks and develop a video annotation mark-up language, which is implemented as:
Firstly, segmentation and annotation. Segmentation and annotation of video sequence are important for digital video libraries, content-based video browsing and video data mining. Video segmentation and annotation require manipulation of the video stream to composite synthetic imagery and information with real video imagery. The manipulation involves only the 2D image space or the 3D scene space. So the key problems to be solved are: proposing algorithms for embedding marks, algorithms for identify marks, and modelling necessary video functions.
Secondly, design parsing for encoding and decoding video data.
Thirdly, design web browser for browsing and navigating video flexibly.
Finally, experimental results showing the significance of the proposed method will be provided, such as distance education.

## 5 References


NITSUWAT S., JIN J.(1999): Video segmentation based on the presence and/or absence of moving objects. *Proceedings of the SPIE, Vol 3846, Eds.* C C Jay Kuo; Shih Fu Chang; Sethuraman Panchanathan, *Boston United State of America, 18-23 Sep., SPIE, Boston, 1999, pp35-45*.

NITSUWAT S., JIN J., HUDSON H. M.(1999): Motion-based video segmentation using fuzzy C-prototypes and classical mixture model. *Proceeding of the Pan-Sydney Area workshop on Visual Information Processing, 10 November , p69-76.*

KEITH J. HANNA, HARPREET S. SAWHNEY, RAKESH KUMAR, Y. GUO AND S. SAMARASEKARA (1998): Annotation of video by alignment to reference imagery. P*roceedings of the IEEE International Conference on Multimedia Computing and Systems Volume I.*

LYNN WILCOX, JOHN BORECZKY (1998): Annotation and segmentation for multimedia indexing and retrieval. *Proceedings of the 31st Hawaii International Conference on System Sciences (HICSS'98), published by the IEEE Computer Society.*

HISASHI MIYAMORI AND SHU-ICHI IISAKU (2000): Video annotation for content-based retrieval using human behaviour analysis and domain knowledge. *Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition* .

SUDHIR G., LEE J. C.M (1996): Video annotation by motion interpretation using optical flow streams. *Journal of Visual Communication and Image Representation, v7, n4, Dec, p354-368.*

HARRISON B. L., BAECKER R.M.(1992): Designing video annotation and analysis systems. *Proceedings of Graphics Interface'92, May 1992.*

DONALD A. ADJEROH AND M. C. LEE: Synchronization mechanisms for distributed multimedia presentation systems. *1995 International Workshop on Multi-Media Database Management Systems (MMDBMS 95).*

MICHAEL KOZUCH, WAYNE WOLF, AND ANDREW WOLFE: New challenges for video servers: Performance of Non-Linear Applications under User Choice. *Proceedings of the 1996 International Conference on Computer Design (ICCD '96).*

LIAO W., LI V. O. K. (1997): The split and merge (SAM) protocol for interactive video_on _demand systems. *Proceedings of the INFOCOM'97.*

PARK H. K., RYOU H. B.(19998) : Multicast delivery for interactive video-on demand service. *Proceedings of the 13th International Conference on Information Networking (ICOIN'98).*

CHRISTEL M. G., OLLIGSCHLAEGER A. M., HUANG C.(2000): Interactive maps for a digital video library. *Multimedia, Jan.—Mar. 2000,vol.7, No.1, pp.60-67.*

MAJUMDER A., GOPI M., SEALES W. B., AND FUCHS H.(1999): Immersive teleconferencing: a new algorithm to generate seamless panoramic video imagery. *Proceedings of the International Conference on Multimedia 1999.*