

An interactive Video System supporting E-Commere Product Placement

DIRK BALFANZ, MATTHIAS FINKE, CHRISTOPH JUNG, REINER WICHERT
Mobile Information Visualization
Computer Graphics Center
Rundeturmstrasse 6, 64283 Darmstadt
GERMANY

Abstract: - Based on the technological progress within the Internet, content like digital video has become more and more popular for any kind of user. Because of these new technologies a merging process of different media types has been started which leads to new kinds of applications. Within this paper such a new application is presented based on an interactive video system for the area of E-commerce. In order to be competitive as an e-shop provider on the market it is highly important to supply the customers with an recognizable added value concerning the presentation of the product line. The embedding of such a system for using an interactive video in the context of E-commerce is the main focus of this work.

Key-Words: - hyperlinked video, electronic commerce, online shopping, interactive catwalk, dynamic reassembled information, databases, PHP, SQL

1 Introduction

The Internet has a great impact on our everyday life. Since the beginning of the World Wide Web the number of Internet clients was rising tremendously and in private as at work Internet is used for a growing variety of purposes.

One specific area that is still discovering its potential is E-Commerce in its different flavors. For instance, more and more companies are going to establish new ways of product placement using the WWW. Having this technology companies can serve any customer despite of his or her location. In comparison to real-world shops it is easily possible to provide direct customer feedback together with additional services like 7x24 h service and extensive product information.

The advantage for a customer can be a much better quality of service or simply the wide range of internet shops and product offers he or she can choose from regardless where the shop is actually located in the real world. Of course this implies an operative logistic system behind any shop to supply the customers with the ordered products in time. However, because in virtual world spatial distances become negligible and different shops are only a mouse-click away a crucial point is customer loyalty. For an online-shop the customer base is highly important in order to be competitive between others. This means that one key-target of any online shop is to bind customers. If customers are satisfied with the service they will stay and come back next time respectively.

Concerning the products that are sold via the Internet, one could divide them into two main groups:

- services like selling theatre tickets and
- products as items, for instance a bike or clothes.

For the product group one can distinguish between new products and those well known and familiar to the customer. The latter case means customers are aware of what they are ordering because they already have all needed information about the item. These products might be repairable spares for bikes or cars, for instance. For the first sub-group this is not true and customers are highly dependent on the information that is given by the internet shop.

Therefore one challenge for E-Commerce is to find new ways of presentation techniques to provide the customer with the needed and useful information in order to sell their products, especially for items which are new to the customers. One possibility to overcome this challenge is the deployment of interactive video content containing *video-hyperlinks* in the field of E-commerce.

Where pure video presentations can already seize the customers attention, video interactivity allows for a quite natural way to acquire needed product information. Via extension of the familiar point-and-click paradigm to video presentation the shop user can be entertained and correctly informed at the same time.

This paper presents a system combining an interactive video presentation within an E-Commerce shop. The topics discussed will be system features, system architecture and process flow on the customers side as well as on the providers side (authoring).

2 System

2.1 Introduction

Besides textual and audio information, digital video has become more popular on the Internet. Higher bit rates, which means a faster transmission of data, new compression techniques and new transfer technologies, like the streaming technology, make it possible to transmit video content on the Internet.

The difference between a digital video and an interactive digital video is that the known VCR functionalities like play, stop, review, are extended by new options providing the user the ability to interact with the video content global instead of just watching it.

An interactive digital video consists of video objects, for instance, a mannequin on a catwalk. Once video objects are defined within a video sequence, which means the location of the objects of interest are determined at each time of their appearance, they can be selected. The selectable objects are defined as sensitive regions [1] within a video sequence and are used in order to retrieve additional information.

Additional information provides the user with more details about the object of interest, for instance the video object *mannequin* can be linked to a html page providing information on the clothes she is wearing combined with a ordering entry. The information can be, e.g., text, image, video or sound media, HTML pages or even another applications. An overview of the topic of digital video and hyperlinks is given by [2].

2.2 Hyperlinked Video (MOVieGoer)

The MOVieGoer system developed by the Computer Graphic Center in Darmstadt Germany [10] is based on the system architecture shown in figure 1. It consists of a client-server architecture which allows global interaction. Based on data component separation, the system is open for customizations and the introduction of personalization.

The MOVieGoer client component (player) is written completely as a Java applet. Using Java, the application can be updated for the use on MHP [3] (Multimedia Home Platform) compliant devices, for instance.

In order to retrieve additional information from a digital video selectable objects are defined. The defined objects are annotated with video hyperlinks which have a beginning, a duration and an end within a sequence. This is a major difference between a hyperlink in a hypertext and a video hyperlink in an interactive video. The time a object is visible in a video sequence determines the maximum time the video hyperlink can exist

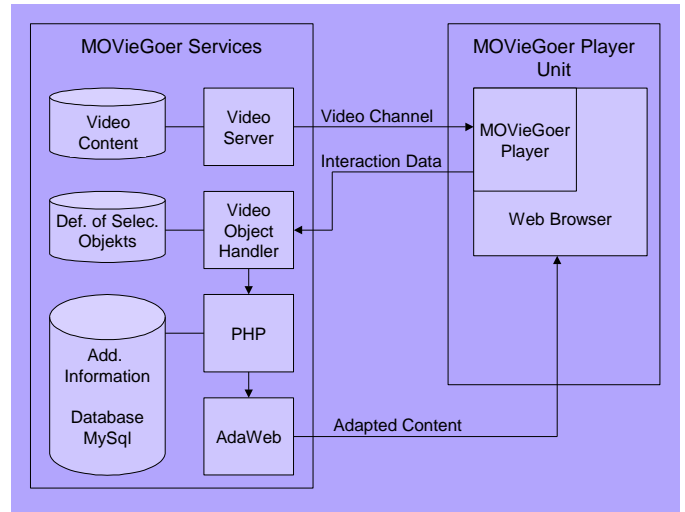


Figure 1: MOVieGoer System Architecture

The objects in a digital video typical change their location within a sequence and therefore tracking is required. Tracking the objects remains one of the main difficulties in authoring an interactive video. Tracking an object manually in each frame of a sequence is insufficient and time consuming since the object might be visible in hundreds of frames. On a detailed discussion on object tracking and a video tracking based solution see [4].

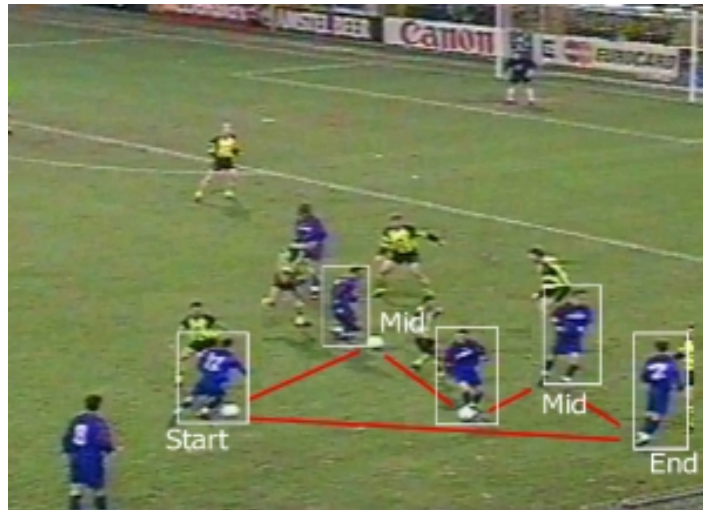


Figure 2: Definition of Key-Frame within a Video Scene

So far the MOVieGoer system supports one half automatic (key-frame method [5]) and two full automatic tracking methods (Camshift Tracker [6] and Kanade-Lucas-Tomasi Feature Tracker [7]) in order to define the selectable objects within a video sequence.

The definition on the selectable objects is stored in a separate file as indicated in figure 1. The file includes the following specification of an object in a video sequence:

- Object ID,
- Start Frame (IN), MID Frame (MID) or End Frame (OUT),
- Shape of key frame,
- Interpolation mode,
- Co-ordinates of the defining rectangle.

Using this definition, different object shapes and interpolation methods can be used. Furthermore, it is possible with the separation of the definition of the selectable objects and the video data itself to have more than one definition file, e.g., to support different user groups.

User interaction on video data can be performed with the mouse pointer. The user selects an object inside the video frame of his interest by simply "clicking" on it.

For personalization the system offers him at starting time of the video presentation, the possibility to specify his preferences, which are then used to adapt the additional information that is presented. If no preferences are specified, a default setting is used. Within the system, each user is assigned a Profile ID.

When a user interacts with the video content, interaction data is generated on the client side, consisting of the following parameters [8]:

- co-ordinates of the mouse click
- frame number of the video
- the user profile ID
- the movie ID

This data is transmitted to the server side. The video object handler on the server side determines by means of the interaction data and the definition of the selectable objects whether any object at that particular frame number, including the x-, y-co-ordinates, exists. If the result is positive, the video object handler contacts the adaptation component AdaWeb [9] (see figure 1) by transferring the Profile ID, the user preferences, and a reference pointing to the additional information. AdaWeb then adapts the additional information according to the user preferences and sends it to the client side where it is displayed in addition to the video.

Before the AdaWeb component influences this process the PHP component (see chapter 2.2) reassembles the information stored in the database.

Updating additional information can easily be realized, since the content is located on a specific server and not on many unknown clients. This option offers great flexibility and provides the system with a controlled update functionality. With the update functionality, the interactive video is realized once and then sent to the clients with additional information that can be changed at any time in order to keep the presentation up-to-date. This is especially important for E-commerce applications since the information about products can

change rather fast concerning time-limited offers, for instance.

2.3 E-Commerce

The MOVieGoer system has been integrated into a standard e-shop to demonstrate the potential of the interactive video technology as a special feature for customers of an e-shop.

The resulting enhanced e-shop is based upon WAMP and embeds the interactive video seamlessly. WAMP is a short form of "Windows Apache MySQL PHP", a combination often used by developers for the creation of dynamic web sides, online applications and for the usage in an electronic commerce environment. PHP [13] is available on a lot of web servers and has become an important new standard in web programming languages.

The source code of PHP scripts is "hidden", i.e. the PHP scripts generates HTML pages on the fly and PHP code is never sent by the web server (Apache [11]) to the customers browser although PHP code can be a content of HTML pages. Therefore the customer or user of an PHP based online application can't copy important parts of the source code to use it in own applications. This is an advantage compared with other scripting languages like JavaScript. An important feature of PHP is the broad range of supported data base management systems (DBMS) which are an essential part of every e-commerce application. The enhanced e-shop described in this paper is based upon the fast and reliable MySQL server [12], a DBMS using the standardized Structured Query Language (SQL) to support very large data bases. The multithreaded MySQL server has growing market shares especially in web applications.

A digital video has to be edited to create the necessary files [see chapter 2.1] before it is able to serve as a hyperlinked video in a product presentation.

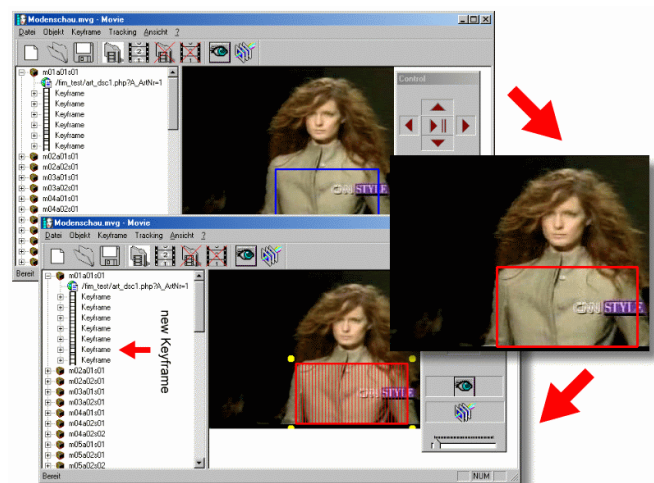


Figure 3: MOVieEditor

An authoring tool called MOVieEditor is used to define the sensitive regions (see figure 3), which represents all of the selectable objects within a video sequence, i.e. the editor determines the number of the available information in the final presentation. The result of the authoring process are several additional files and the unchanged video. These files consist of special information for the server side of the MOVieGoer system which are necessary for the handling of the user interaction and the resulting presentation of the requested information. The files contain the definition of the selectable objects in the digital video and the path (URL) to the additional information represented by each object, for instance.

The client side consists of the MOVieGoer video player embedded into the front-end of the e-shop. It sends data to the server side when a customer clicks on the video with the intention to retrieve additional information about an object, i.e. information about an article like a dress. The e-shop is completely operational without the feature of a hyperlinked video. This feature has an own entry in the menu called I-catwalk (interactive catwalk). If a customer doesn't want to watch the presentation of products via video he or she can navigate through the shop and buy a product like the customer is used to do it from a traditional e-shop.

The server side contains the back-end of the e-shop, the digital video, the definition of the selectable objects and the additional information [see chapter 2.1]. Each of the last three elements is stored separately. The e-shop consists of several PHP scripts and some static HTML pages. The interaction data (x-, y- co-ordinates, the movie title and the frame number) transmitted from the client side is processed by using the implemented content filters. If an object is existing at the selected location in the video, an URL for the additional information is sent back to the customers browser. In case of an existing product, this URL consists of the address of a PHP script and an article number. The requested information will be retrieved by the PHP script from the MySQL server and integrated into a HTML template. The resulting reassembled HTML page will be transmitted to the client side.

Each request for information about an identical article can result in different HTML pages if the content of the data base is changed on the server during the queries. The important advantage concerning the combination of MySQL and PHP is that the requested additional information can be changed at any time in order to keep the presentation up to date.

The connection between client and server in the enhanced e-shop depends on three data channels, i.e. two forward channels and one back channel. The first forward channel is used to transmit the digital video towards the client (browser) and the second forward channel is used to transport additional information from the server to the client. The back channel is used by the client in order to send the interaction data from the customers interaction to the server.

3 Scenario

The following example is a demonstration for the online ordering process in the enhanced e-shop supported by the embedded MOVieGoer system. It is focused on the new functions concerning the added value, although a customer is still able to use the conventional category groups of an online shop like article catalog, search, bestseller or service.

The technology demo has been conceived as an interactive and moderated presentation of a fashion shop, i.e. a presenter who has acquainted himself with the MOVieGoer system and the enhanced e-shop, demonstrates the system to a circle of interested observers. Figure 4 shows the relevant process flow.

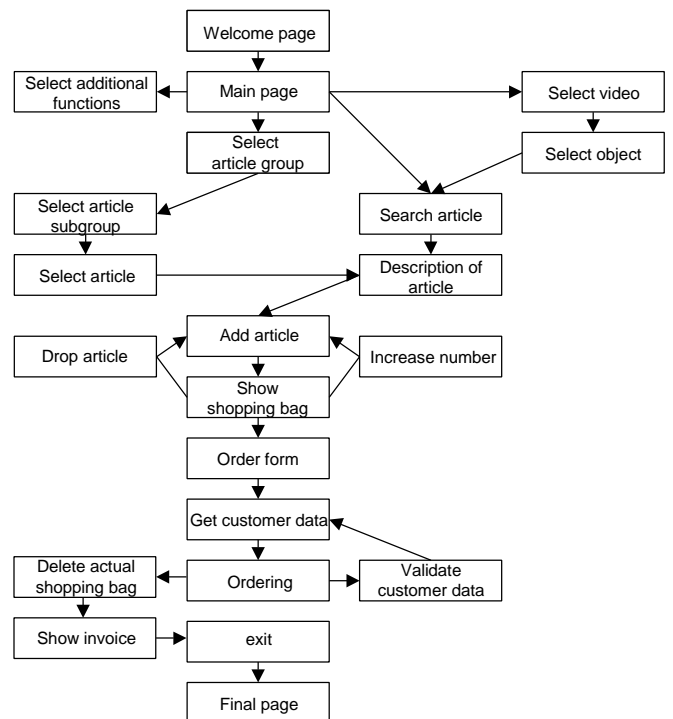


Figure 4: Process Flow of the Enhanced E-Shop

The preferred way of presenting the client-server architecture is the utilization of two computers. The first one acts as a server with the complete installed system, i.e. MOVieGoer system, e-shop and all media data. The second computer (client) only requires an installed

modern browser like the Internet Explorer from Microsoft.

The launched server waits for the requests from the client side. The browser displays a well designed welcome page as a teaser which leads to the main page of the online shop. The welcome page is just for demonstration purpose, i.e. a commercial shop wouldn't hold up the customer with such a page.

The structure of the main page consists of three parts: header, body and footer. The static header is subdivided into a frame for the corporate identity (logo, quick search, restart, exit) and a frame for the customer centric upper navigation bar. The dynamic area inside the body contains content like output of PHP scripts, hyperlinked videos or static HTML pages. The logo of the fictitious shop owner in the upper left corner of the header represents the products and the added value for the customers. F.I.M. is a short form for "fashion in motion", a first hint concerning the integrated interactive video system. The third and last frame of the main page is the footer which is used for the visitor centric lower navigation bar.

Important for the further course of the demonstration are only the upper navigation bar and the area in the middle of the screen (footer). The third icon on the upper navigation bar shows a black cat on the screen of a TV (see figure 5). The cat is a second hint for the added value and represents the interactive catwalk (i-catwalk) which is launched by a click on the icon.

A MySQL query on the server side fetches the available fashion videos and a PHP script generates the entries to choose from. Each hyperlinked entry is labeled with the title of the corresponding video.

If a customer has selected an interesting video the presentation environment for the video is set up, i.e. the area of the body is subdivided into three frames. The two frames on the left side are used to show the video with its corresponding control elements and a short description of the video. The frame on the right side is needed to present the information about an object, the content of the shopping bag and additional advices. The following screenshot shows the existent information in the data base about an existing object.



Figure 5: presenting additional information of an object

The clothes and accessoires (objects) presented by a model in a fashion video can be selected with the mouse which initiates a transmission of interaction data to the server side. If an object with a corresponding article is defined at this position a query is sent to the MySQL server. The resulting information (picture, name, price, article number etc.) is integrated into a template as mentioned before and the final HTML page is transmitted to the client side. The customer can add the article of interest to the shopping bag with a simple click on a corresponding button.

If an article is added to the shopping bag current content is shown on the right side of the screen. The video is still visible and running so that other articles are available for selection. The content of a shopping bag consist of information (picture, article number, sum etc.) which the customer expects in an typical online shop, see figure 6.

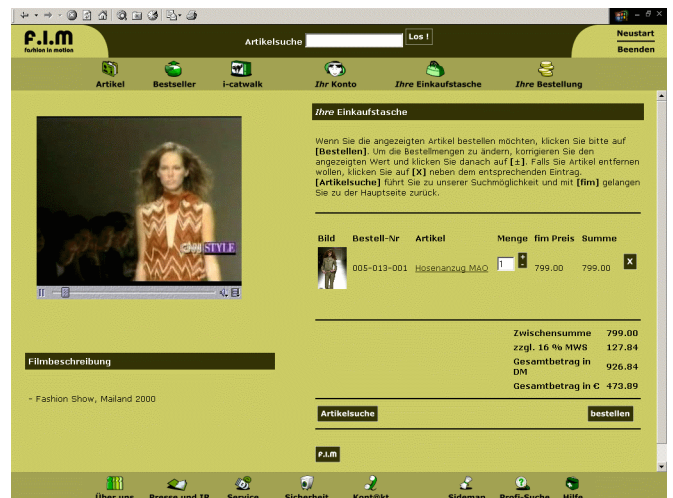


Figure 6 : video and actual shopping bag

The customer is able to change the number of an article, the articles can be dropped, and the ordering process can be initiated. If customer data is correct and validated the

invoice will be presented otherwise hints concerning the missing or wrong input is shown so that the customer can change incorrect parts of his input.

The demonstration ends with a final page which enables the presenter to restart the presentation again.

4 Conclusion and Future Work

E-shopping can have clear advantages for a customer as for example additional services and information, direct feedback or simply the wide range of easy accessible offers of all kinds. One particular challenge for e-shop operators is how fast and easy customers can move to an competitive offer. This requires extra effort to bind an potential customer at first glance with a persuasive presentation and a good level of service.

In this paper we presented the approach to get customers attention via video interaction. The specific win of this approach is the eye-catching quality of video presentation for itself in combination with the very familiar way of handling the interaction part. As point-and-click is well-known to all Internet users there is immediate capability to access the information potential behind the new presentation form of interactive video.

The implementation of our approach is based upon the MOVie system. MOVieEditor allows to define selectable objects and offers as well partly automated support to follow moving objects via different tracking methods. MOVieGoer based on a client-server architecture comprises the server side with video content and additional information stored separately and the client side for visualization and interaction. Based on data component separation, the system is open for customizations and the introduction of personalization.

Used technologies are standards such as Java applets, PHP or MySQL which facilitates future use for example on MHP [3] (Multimedia Home Platform). Separated storage of video data on the one hand and object information as well as the additional information that is linked to the video stream on the other hand enables easy updating and authoring.

From our experience with the shown test system, attractiveness of the regarding site is increased remarkably. Extended information availability in combination with the eye-catching effect of video presentations as well as the fun using the interactivity reaches the goal to attract the customer attention.

Future extensions to this concept may take place in different key areas. The most promising from our point of view will be real time video hyperlinking.

Different to the given scenario, where the fashion show was pre-captured and authored offline one can imagine

the same setting as a life event. Added value for potential customers would be the flavor of entering more or less a real shop and getting interaction facilities to persons and not only to systems, e.g. via a chat extension.

Technological challenges are here to define the objects to be hyperlinked and especially to track them completely automated during the event.

Current research efforts for life tracking often based on active or passive markers which are recognized by image processing or trough additional equipment at the set.

Future applications shall support marker-free tracking of objects.

References:

- [1] V. Burrill, T. Kirste, and J. Weiss, Time-varying Sensitive Regions in dynamic multimedia objects: A pragmatic approach to content-based retrieval from video, *Information and Software*, 36(4), 1994
- [2] N. Sawhney, D. Balcom, and I. Smith, HyperCafe: narrative and aesthetic properties of hypervideo, *Proceedings of the 7th ACM conference on HYPERTEXT*, Bethesda, MD USA, 1996.
- [3] DVB, Multimedia Home Platform (MHP), TS 101 812, <http://www.dvb.org/standards/index.html>
- [4] V.M. Bove, Jr., J. Dakss, S. Agamanolis, and E. Chalom, Adding Hyperlinks to Digital Television, *SMPTE Journal*, Nov. 1999, pp. 795-801
- [5] Finke M., An Architecture of a Personalised Dynamic Interactive Video System, Digital Content Creation Conference, Bradford, Britain, 2000
- [6] Gary R Bradski, Computer Vision Face Tracking for use in a Perceptual User Interface, *Intel Technology Journal* Q2, 1998
- [7] Jean-Yves Bouguet: Pyramidial Implementation of the Lucas Kanade Feature Tracker, Description of the algorithm, *Intel Research Lab*, 1999
- [8] A. Höhl, WWW Application for Interactive Video, *Diploma Thesis*, TH-Darmstadt, 1996
- [9] C. Freytag, C. Kumpf, C., and L. Neumann, Utilization of device and user profiles for mobile WWW-Access, *5th Int. Workshop on Mobile Multimedia Communications (MoMuC'98)*, Berlin 1998
- [10] Computer Graphic Center, ZGDV, <http://www.zgdv.de/miv>, Darmstadt, Germany
- [11] Apache Software Foundation, Apache Server, <http://www.apache.org>
- [12] MySQL AB, MySQL Server, <http://www.mysql.com>.
- [13] PHP Group, PHP Scripting Language, <http://www.php.net>
- [14] phpWizard, phpMyAdmin, <http://www.phpwizard.net>
- [15] E. Schicker, *Datenbanken und SQL*, Teubner, 1999