

An Interactive Tool for Sketch-Based Annotation

M. Antico¹, D. Avola¹, P. Bottoni¹, A. Hawash², K. Kanev³, and F. Parisi Presicce¹

¹Department of Computer Science, Sapienza University of Rome, Italy
{antico, avola, bottoni, parisi}@di.uniroma1.it

²Department of Computer Information Systems, An-Najah National University of Palestine
amjad@najah.edu

³Faculty of Informatics, Shizuoka University, Japan
kanev@rie.shizuoka.ac.jp

Annotation of Web content is becoming a widespread activity by which users appropriate and enrich information available on the Net. **MADCOW** (Multimedia Annotation of Digital Content Over the Web) is a system for annotating HTML pages that provides a uniform interactive paradigm to make annotations on text, images and video. **MADCOW** has also recently included novel features for groups and group annotations with reference to ontological domains, but interaction with the **MADCOW** client is currently limited to the use of common input devices (e.g., keyboard, mouse), requiring a precise selection of the portions to be annotated. In this paper we present a sketch-based interface, which can be used to annotate not only content but also aspects relative to the presentation of the information. While interacting with a standard Web browser, users can draw free-hand geometrical shapes (e.g., circles, rectangles, closed paths) for selecting specific parts of the Web pages to be annotated. The interaction mode depends on the adopted input device. For example, users interacting with touch-screen devices (e.g., smartphones, tablets) can draw shapes with their fingers, but in principle any device able to detect sketching gestures can be supported (e.g., graphic tablets, optical pens). The paper discusses interaction aspects together with an overview of the system architecture. Finally, preliminary experimental tests and some considerations on the usability are also reported.

1 Introduction

Currently, different techniques and methods for conducting collaborative learning and work are used. The ever-increasing availability of information over the Web, as well as the need to interact with this information in more natural and efficient ways, are promoting alternative modalities of human-computer interaction [1, 2]. In the last two decades speech, handwriting and sketch have proven to be three fundamental modalities for the creation of novel and highly interactive interfaces [3]. These interfaces are used more and more for interacting with different devices, favouring natural interaction through the expression of commands with graphical symbols [4]. In our context, sketch-based interfaces are a promising avenue to handle and annotate Web pages since, unlike the speech-based

interfaces, they can be used anywhere. Moreover, unlike handwriting-based interfaces, they are particularly suitable to select and highlight parts of digital documents and Web pages [5, 6, 7]. Finally, the diffusion of current touch-screen devices makes these interfaces a favourite tool for interaction, as the usage of simple sketched commands to scroll the phone book or visualize photos on our smartphones or tablets has become everyday experience.

Digital annotation is the process of adding information to the content of a multimedia document, enriching it with additional valuable information, without altering the original content. Reasons for annotating can vary: to create mementos or clarifications of interesting content; to entertain discussions with other users; to construct or integrate documents [8, 9]. On-demand training and education processes also use annotations [10, 11, 12].

MADCOW is a system for annotating HTML pages that provides a uniform interactive paradigm to make annotations on text, images and video. We have recently introduced a sketch-based interface that allows users to select in alternative way text, images, and in general areas in the page to be annotated, providing an additional interaction modality. More specifically, the sketch-based interface added to **MADCOW** enables users to draw common (e.g., squares, circles) or customized (e.g., closed paths) shapes around the portion of text or the portion of image to be annotated (including around the whole image). It is envisaged that this new interactive mode will increase **MADCOW** adoption, also favoring, intervention in discussion threads and collaborative work among **MADCOW** users.

In the rest of the paper, after giving some technical background on **MADCOW** in Section 2, Section 3 discusses related work on sketch-based interfaces for annotation tasks on Web resources. Section 4 presents a reference scenario, while Section 5 briefly shows the main technical aspects of the sketch-based interface and Section 6 reports preliminary experimental tests and some considerations on the usability. Finally, Section 7 concludes the paper.

2 Technical Background

As shown in Figure 1, **MADCOW** is a three-tiered client/server architecture Web-based system for annotation of HTML pages. The architecture follows the Model-View-Controller (MVC) pattern and a set of bookmarklets represent the core of the client side (the paper does not address the **MADCOW** architecture which is detailed in [13]).

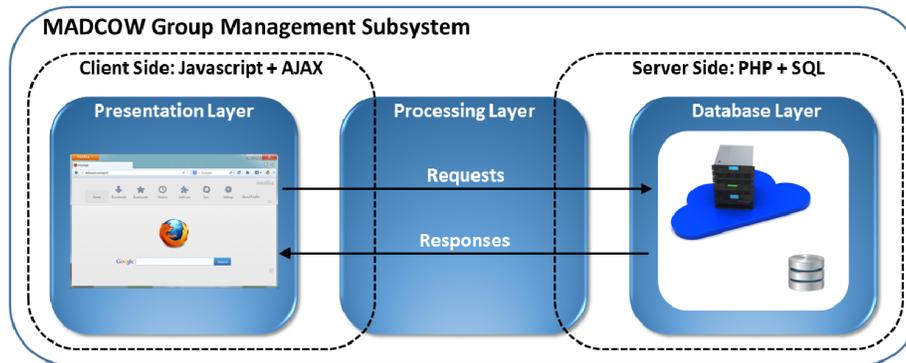


Fig. 1. MADCOW three-tiered architecture. The Presentation Layer deals with the interaction between users and MADCOW and manages requests to the server side. The Processing Layer, distributed between client and server, deals with information processing and the Database Layer manages information storage and responses to the client side.

MADCOW supports the annotation of (portions of) text, image or video, with textual content, links to other resources, and user-defined tags [8]. Annotations can be published in one of the following three ways:

- (1) *public*: viewable by any user;
- (2) *private*: viewable by their submitters only;
- (3) *group-related*: viewable only by members of the group to which they are posted.

After selecting the Web content to be annotated, a pop-up window appears containing data related to the selected part and offering the possibility of adding some textual comment, and possibly links to other documents and a set of tags, which will constitute the annotation. In addition, it allows users to set the annotation type, the annotation visibility (private, public, or group) and, in case of group annotation, it enables users to select the target groups from a drop down menu (see Figure 2).

One of the main objectives of MADCOW is to provide simple ways for users to create annotations and discussion threads [14]. The uniform style of interaction [15] and the advanced group features [16, 17] support users in exploring and enriching visited pages.

3 Related Work

Some systems, from literature or commercial tools, adopt sketch-based interfaces for freehand drawing of shapes on Web resources on which to post annotations or include additional material. Some of these systems enable users to draw shapes without linking them to specific parts of the Web pages. In other cases, users

can link their annotations but only to a single type of resource (typically text). Some other systems take snapshots of pages and add annotations by means of offline applications, in some cases allowing online annotation of digital resources. **MADCOW** allows the annotation of portions of images, as well of a resource as a whole. The novel sketch-based interface acting on text and images enable users to annotate Web pages maintaining the same paradigm and same functionalities of the manual selection.

TagCaptcha [18] introduces a method of annotating images for use in a retrieval setting by exploiting the need for CAPTCHAs (Completely Automated Public Turing test to tell Computers and Humans Apart) online. TagCaptcha presents the user with a number of images that must be correctly labelled in order to pass the test. Images are divided into two subsets: a control or verification set for which annotations are known, and an unknown set for which no verified annotations exist. The verification set is used to control against the tags provided for the unknown set. If the user provides correct verification tags, the tags for the unknown set are promoted. Usability results from a small user study as well as sample user tags from the online demonstration system were reported. In [19], the authors emphasise the collaborative design review as an important part of architectural design work. The Space Pen system supports annotation and drawing on (and inside) 3D VRML/Java models using a regular Web browser to exchange text and sketched annotations for review. The work of [20] presents a sketch-based two-layer representation, called Scene Structure Graph (SSG), to facilitate the video authoring process. One layer in SSG uses sketches as a concise form with which the visualization of scene information is easily understood and the other layer uses a graph to represent and edit the narrative structure in the authoring process. With SSG, authoring occurs in two stages. First, various sketch forms, such as symbols and hand-drawing illustrations, are used as basic primitives to annotate the video clips and the hyperlinks encoding spatio-temporal relations are established in SSG. Then, sketches in SSGs are modified and a new SSG is composed.

Tools for sketching Web sites are mainly adopted for conducting asynchronous sessions of e-learning or collaborative work. With Awesome ScreenShot¹ users can upload images and annotate them by drawing different shapes, starting discussion threads on shared images. Although some kind of collaboration can be conducted in this way, online sketches cannot be drawn over existing Web sites.

Bounce² application from **ZURB** is a very easy to use annotation tool. It lets people add feedback to any interface and share it with other people, creating an image of a visited page, on which users can add annotations. Users will also be able to share their annotated content via a unique URL. Markup.io³ is a bookmarklet based tool that users can drag and drop onto their browser toolbar. Once installed users, can adopt freehand draw to post annotations on the Web pages. Moreover, the tool allows users to share annotated Web pages via

¹<http://www.awesomecreenshot.com/>

²<http://bounceapp.com/>

³<http://markup.io/>

a generated URL. Marquee⁴ is designed to annotate images anywhere. User can simply draw a rectangular box on the image and leave a note. Users can drag and drop files straight from their computers, or even from another browser window, then invite people to participate in the discussion by their email addresses, and finally point to areas in dragged image to have discussions with others by using a freehand tool to draw one's ideas. ScreenDraw⁵ is an extension of Firefox that allows users to annotate their selected Web pages. Users can circle elements, type in notes, or draw on any Web page. Finally, they can save the annotated work as PNG or JPG files. Screenshot⁶ enables users to capture all or part of a Web page and save it or share it with co-workers, friends, or family. The editing tool allows users to mark up images for highlight or blackout before sending them to others. The ability for the tool to work on all platforms gives it strength and widens its worldwide base of users. Jing⁷ allows users to create annotated screen capture images. Once a user takes a screenshot of a Web page, he or she can add text, arrows, draw on it and highlight sections of text.

The reported tools have various affinities with the novel interaction modality introduced into **MADCOW**. They enable users in producing annotations by using a sketch-based interface and most of them are designed to work in online mode. Moreover, many of them allow users to annotate both the whole item (e.g. an image) and components within it. Despite this, most of these tools are designed to treat only a single type of item (e.g., images), whereas **MADCOW** can treat different types of item in a Web page, and support private, public and group-based annotations [21, 22] with this novel modality of interaction.

4 A Reference Scenario

Antonio is a radiology student who utilizes the **MADCOW** annotation system for collaborating with other students. Antonio is the owner and moderator of the RadiologyGroup, a group that collects textual annotations related to a selected set of thematic Web pages regarding his studies. Currently he is studying the procedure overview related to X-Rays of the limbs. In particular, he is focused on the X-Ray of the elbow and needs to discuss several issues with his colleagues. Antonio has found an interesting Web site related to his object of study and posts two annotations. The first one (annotation on text, Figure 2 - bottom) reports some considerations about the steps related to the X-Ray procedure of the elbow. He knows that some diagnostic investigations require several different positions of the extremity but he also knows that a forced pose of a damaged extremity may provoke pain and further damages. The second one (annotation on image, Figure 2 - top) gives information about the bones that compose the elbow; he knows that some colleagues had issues with the details of the joint.

⁴<http://www.marqueed.com/>

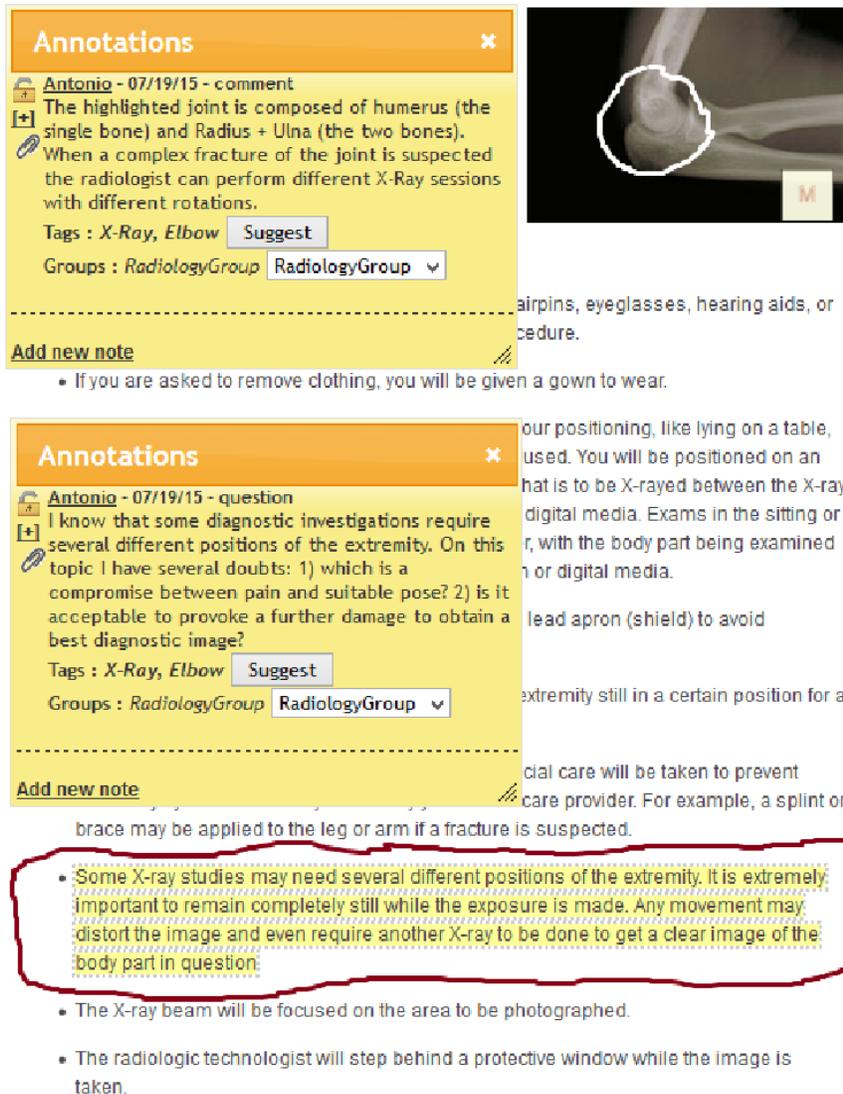
⁵<https://addons.mozilla.org/en-US/firefox/addon/screen-draw/>

⁶<https://chrome.google.com/webstore/detail/webpage-screenshot/bdfnieppndfdhcgbmhfdldjgegkolk>

⁷<https://www.techsmith.com/jing.html>

In both cases, Antonio uses a tablet and employs its touch-sensitive surface to draw a selection shape around text and the image detail.

During the procedure



Annotations ✕

Antonio - 07/19/15 - comment
The highlighted joint is composed of humerus (the single bone) and Radius + Ulna (the two bones).
When a complex fracture of the joint is suspected the radiologist can perform different X-Ray sessions with different rotations.
Tags : X-Ray, Elbow
Groups : RadiologyGroup

Annotations ✕

Antonio - 07/19/15 - question
I know that some diagnostic investigations require several different positions of the extremity. On this topic I have several doubts: 1) which is a compromise between pain and suitable pose? 2) Is it acceptable to provoke a further damage to obtain a best diagnostic image?
Tags : X-Ray, Elbow
Groups : RadiologyGroup

Add new note

- If you are asked to remove clothing, you will be given a gown to wear.

• Some X-ray studies may need several different positions of the extremity. It is extremely important to remain completely still while the exposure is made. Any movement may distort the image and even require another X-ray to be done to get a clear image of the body part in question.

- The X-ray beam will be focused on the area to be photographed.
- The radiologic technologist will step behind a protective window while the image is taken.

Fig. 2. Sketch-based interface of MADCOW. The first popup (top) is related to the image (the M on the image highlights the presence of an annotation). The second popup (bottom) is related to the text (highlighted in yellow to indicate the presence of an annotation). Different colours can be selected for sketching on HTML pages and images.

5 Sketch-Based Interface: An Overview

When a user chooses to use the sketch-based interface, an invisible and drawable canvas is layered on the Web page. The interface works by mapping the shapes drawn by user with the related items of the Document Object Model (DOM) (and CSS) for the document currently presented by the browser. The sketch-based interface considers closed shapes composed of only one stroke (i.e., a single gesture). When the user closes the path and lifts the pointing device (which might be the finger) from the screen, the popup appears. The following pseudocode simplifies the process:

```
function sketch(){
    canvas = system.createCanvas();
    shape = canvas.drawShape();
    system.map(shape, website.DOM.resources);
    onMouseRelease(system.showPopUp());
}
```

The sequence diagram presented in Figure 3 depicts the main steps executed by **MADCOW** to implement the sketch-based interface. When a **MADCOW** user selects on the browser the sketch-based modality, an environmental object, **mdctrl**, executes the function **createCanvas()** which equips the whole Web page, **document**, with a drawable canvas. The coordinates of the set of pixels (i.e., shape) drawn by user, **drawShape()**, are stored into a suitable data structure whose reference is returned to the **mdctrl** object. The whole mechanism exploits a set of functions and sub-functions injected by the system (server side) into the Web page when a user selects the sketch-based interface. The **mdctrl** object executes the function **assignShapeToDom()** that assigns the shape to the appropriate resources in the DOM tree of the current Web site. Afterwards, the execution of the function **open()** shows the new pop-up window with all components needed to create a new annotation and prepare it for submission. After the user fills the necessary data in the pop-up window, the annotation can be submitted, **submit()**, to the server side of **MADCOW**. Note that, the internal processes of request and response (i.e., **submit()** function, as well as **php** and **db** objects) of **MADCOW** are independent from the specific interaction modality.

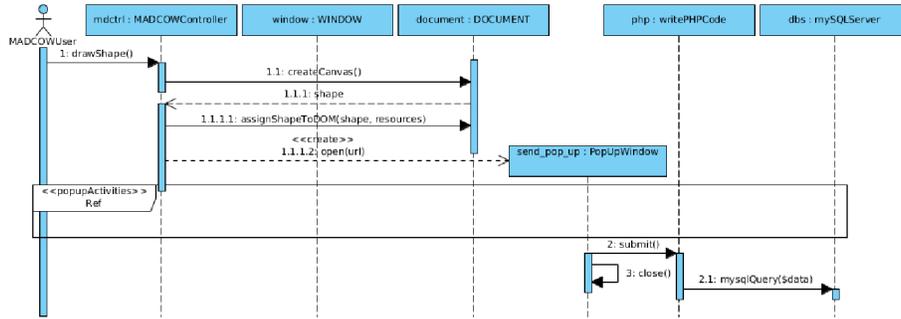


Fig. 3. Sketch-based Annotation Sequence Diagram.

6 Preliminary Experimental Tests

During the implementation of the main functions of the sketch-based interface, we have performed some functional tests to check their work and discover possible issues. The current section briefly reports all these observations and considerations. Table 1 summarizes the different kind of Web sites used to test the sketch-based interface. In particular, we have observed that the interface works quite well on Web sites (Wiki and Information) with a low level of dynamic items, while its performance is poorer on Web sites having a high level of these items (social). These issues are due to the amount of changes inside the DOM; in fact social Web sites, are prone to dynamic changes due to real-time user posts, system spots, and other. Another issue is represented by those sites (E-Commerce) that embed third part elements (e.g., PDF file) within the presentation of the HTML page. Finally, based on our experience, we claim that sketch-based interfaces are powerful tool since a wide range of users already use them to control smartphones, tablets, car commands, and so on.

Table 1. Preliminary Observations and Considerations.

Kind of Web Site	Text/Images/Video	Dynamic Items	Success Rate
Wiki	Text(70%)/Images(25%)/Video(5%)	Low	High
Social	Text(50%)/Images(25%)/Video(25%)	High	Low
E-Commerce	Text(60%)/Images(35%)/Video(5%)	High	Medium
E-Learning	Text(60%)/Images(25%)/Video(15%)	Medium/High	Medium
Information	Text(80%)/Images(10%)/Video(10%)	Low/Medium	High

7 Conclusions and Future Work

We have enriched **MADCOW** with an alternative, sketch-based, interaction mode, for selecting specific parts of the Web pages to be annotated. This mode

enables users to interact with **MADCOW** everywhere using it with a wide range of novel touch-screen devices. Future avenues of experimentation might include the use of advanced devices of interaction (e.g., LEAP Motion Controller) to handle **MADCOW** functions and the integration of 3D technologies (e.g., Oculus Rift) to be able to annotate presentations of virtual reality.

Bibliography

- [1] M. Chen, D. Zhang, and L. Zhou, “Empowering Collaborative Commerce with Web Services Enabled Business Process Management Systems”. *Decision Support Systems*, Vol. 43, No. 2, pp. 530-546, 2007.
- [2] M. Addisu, D. Avola, P. Bianchi, P. Bottoni, S. Levialdi, and E. Panizzi, “Annotating Significant Relations on Multimedia Web Documents”. *Multimedia Information Extraction: Advances in Video, Audio, and Imagery Analysis for Search, Data Mining, Surveillance, and Authoring*, pp. 401-417, 2012.
- [3] P. Grifoni, F. Ferri, M.C. Caschera, A. D’Ulizia, and M. Mazzei, “MIS: Multimodal Interaction Services in a Cloud Perspective.”. *Journal of Next Generation Information Technology*, pp. 1-10, 2014.
- [4] D. Avola, P. Bottoni, A. Dafinei, and A. Labella, “FcBD: An Agent-Based Architecture to Support Sketch Recognition Interfaces”. *Proc. DMS 2011*, pp. 295-300, 2011.
- [5] J.J. Shah, N. Vargas-Hernandez, J.D. Summers, and S. Kulkarni, “Collaborative Sketching (C-Sketch) An Idea Generation Technique for Engineering Design”. *The Journal of Creative Behavior*, Vol. 35, No. 3, pp. 168-198, 2001.
- [6] D. Avola, F. Ferri, P. Grifoni, and S. Paolozzi, “A Framework for Designing and Recognizing Sketch-Based Libraries for Pervasive Systems”. *Proc. UNIS-COM 2008*, Vol. 5, pp. 405-416, 2008.
- [7] P.R. Cohen, E.C. Kaiser, M.C. Buchanan, S. Lind, M.J. Corrigan, and R.M. Wesson, “Sketch-Thru-Plan: A Multimodal Interface for Command and Control”. *Communications of the ACM*, Vol. 58, No. 4, pp. 56-65, 2015.
- [8] P. Bottoni, R. Civica, S. Levialdi, L. Orso, E. Panizzi, and R. Trinchesi, “MADCOW: A Multimedia Digital Annotation System”. *Proc. AVI 2004*, pp. 55-62, 2004.
- [9] R. Kawase, E. Herder, G. Papadakis, and W. Nejdl, “In-Context Annotations for Refinding and Sharing”. *Proc. WEBIST 2010*, Vol. 75, pp. 85-100, 2010.
- [10] D. Bargerion, J. Grudin, A. Gupta, E. Sanocki, F. Li, and S. Leetiernan, “Asynchronous Collaboration Around Multimedia Applied to On-Demand Education”. *Journal of Management Information Systems*, Vol. 18, No. 4, pp. 117-145, 2002.
- [11] A. Sakar, and G. Ercetin, “Effectiveness of Hypermedia Annotations for Foreign Language Reading”. *Journal of Computer Assisted Learning*, Vol. 21, No. 1, pp. 28-38, 2005.
- [12] Y.-S. Lai, H.-H. Tsai, and P.-T. Yu, “Integrating Annotations into a Dual-Slide Powerpoint Presentation for Classroom Learning”. *Educational Technology & Society*, Vol. 14, No. 2, pp. 43-57, 2011.
- [13] D. Avola, P. Bottoni, and R. Genzone, “Light-Weight Composition of Personal Documents from Distributed Information”. *Proc. IS-EUD 2011*, Vol. 6654, pp. 221-226, 2011.
- [14] D. Avola, P. Bottoni, S. Levialdi, and E. Panizzi, “Annotation Threads in MADCOW 2.0”. *Proc. HC 2010*, pp. 159-166, 2010.

- [15] D. Avola, P. Bottoni, S. Leviardi, and E. Panizzi, "Interacting Annotations in MADCOW 2.0". *Proc. AVI 2010*, pp. 389-390, 2010.
- [16] A. Hawash, "Introducing Groups to an Annotation System". In *Proc. CHI-taly 2013*, pp. 1-10, 2013.
- [17] D. Avola, P. Bottoni, and A. Hawash, "Relevance Measures for the Creation of Groups in an Annotation System". *Journal of Visual Languages & Computing*, Vol. 25, No. 6, pp. 695-702, 2014.
- [18] D. Morrison, S. Marchand-Maillet, and . Bruno, "TagCaptcha: Annotating Images with CAPTCHAs". *Proc. MM 2010*, pp. 1557-1558, 2010.
- [19] T. Jung, M.D. Gross, and E. Yi-Luen Do. 2002. "Sketching Annotations in a 3D Web Environment". *Proc. CHI-EA 2002*, pp. 618-619, 2002.
- [20] M. Cui-Xia, L. Yong-Jin, W. Hong-An, T. Dong-Xing, and D. Guo-Zhong, "Sketch-Based Annotation and Visualization in Video Authoring". *IEEE Transactions on in Multimedia*, Vol.14, No.4, pp. 1153-1165, 2012.
- [21] D. Avola, P. Bottoni, and A. Hawash, "Users-Groups Matching in an Annotation System: Ontological and URL Relevance Measures". *Proc. CSIT 2014*, pp. 100-109, 2014.
- [22] D. Avola, P. Bottoni, and A. Hawash, "Using Ontologies for Users-Groups Matching in an Annotation System". *Proc. CSIT 2013*, pp. 38-44, 2013.