RIVIDOC
A GAME FOR SIMULATING THE EXTRACTION OF RELEVANT INFORMATION FROM DOCUMENTARY VIDEOS

Grissa Vianney Maturana Gonzalez
Facultad de Minas, Universidad Nacional de Colombia,
gvmatura@unal.edu.co

Carlos Mario Zapata Jaramillo
Facultad de Minas, Universidad Nacional de Colombia,
cmzapata@unal.edu.co

Miguel David Rojas López
Facultad de Minas, Universidad Nacional de Colombia,
mdrojas@unal.edu.co

ABSTRACT

Currently, audiovisual documents constitute one of the main sources of information. Documentary videos are audiovisual documents aiming to capture reality by means of images, sounds, texts, and interviews. However, information contained in the documentary videos is not always relevant; in fact, much information is part of the context or the connection among image sequences of the video. Some authors try to identify the objects and actions in videos out of the background, while others employ algorithms to store and correlate video sequences for later search. However, few games are oriented to the information extraction from documentary videos and the information potentially learned from such videos. In this paper, we propose a game allowing for analysts to identify relevant information in documentary videos as a way to model a software system. In the game, players are encouraged to seek information by entering several scenarios of documentary videos represented by textual and audiovisual sequences.

INTRODUCTION

Currently, audiovisual documents constitute one of the main sources of information. Documentary videos are audiovisual documents aiming to capture the reality by means of images, sounds, texts, and interviews (Grant, 1998). However, information contained in the documentary videos is not always relevant; in fact, much information is part of the context or the connection among image sequences of the video.

Some authors try to identify the objects and actions in videos out of the background (Li et al., 2013), while others employ algorithms to store and correlate video sequences for later search (Adami et al., 2001); some others propose algorithms to identify the properties of the sequences of a video and store them in ontologies (Shirahama et al., 2007). In addition, some authors use the videos for reflecting reality and validating system software requirements (Carmichael et al., 2007). All of such studies lead us to identify the importance of information extraction and the usage of videos for creating knowledge bases for later use.

In this paper, we propose a game allowing for analysts to identify relevant information in documentary videos as a way to model a software system. In the game, players are encouraged to seek information by entering several scenarios of documentary videos—represented by textual and audiovisual sequences. Players should also model the information they have seen, heard, or read about the documentary video by using a pre-conceptual schema (Zapata et al., 2006). Thus, the aim of the game is reached when the players—analysts—are able to represent information provided from a particular domain, but in this case they use a documentary video as a source of information.

This paper is organized as follows: first, we present the fundamental concepts that support this work; then we present the structure of the game RIVIDOC; after that, we discuss some results of the activity; finally, we summarize conclusions and future work.

THEORETICAL FRAMEWORK

The documentary video is an audiovisual product whose base is the expression of reality. Rabiger, M. (1987) defines documentary videos as a "scrutiny of the organization of human life" and added that the documentary "focuses on the richness and ambiguity of life, as it is the really " aimed to detect individual and human values. Meanwhile Grant reaffirms that the documentary explores people and real situations. Thus, as we can interpret a documentary as a video recording of reality from its director's perspective.

Relevant information is information is important. Importance is defined by the Royal Spanish Academy (RAE) as the value that is given to someone or something, but, much of the information that is received daily has value according to the

purpose or goal that you have.

In the software engineering the actors responsible for determining the relevance of information received are Requirements Engineers, who face the task of understanding the domain of a stakeholder or client through different techniques such as interviews, brainstorming, and use cases; for solving a particular problem (Carrizo, 2012).

There are projects related to the use of audio-visual products to service engineering. One of them raises the problem of having an elderly user explain what issues they might be having when interacting with a software product, and proposes using video documentaries of the user's performance to better convey the message. (Carmichael et al., 2007). Another case is the use of video as a bridge between the end user and the system. The validation of the software application is achieved through feedback as presenting the current state of development of a system by actually operating the system as it exists at that point of time, through visual scenarios you can see the UI sketches, the way that it work and how to use the software (Stangl & Creighton, 2011). Some theorists, such as Creighton et al. (2006) use a Software Cinema Technique that basically employ a digital video in requirements engineering one example of using video for software development projects in general, the stakeholder activities are record in shots, video is always employed as a visual scenario that shows how the envisioned system may look, work, or be used. Xrave is a tool which maps the sequence of events from multiple views to recreate a UML model use case view.

Some authors focus their work on the automated extraction of information from audio and video. According to Lui & Wang (2001) a new approach for automatically generating a list of major casts for video based on both audio and visual information, each major cast is characterized by two attributes: face and speech. The detection procedure is to find corresponding face occurrences and speech segments by analyzing video at two levels; Audio and visual information is utilized separately at low level, and at high level cues from different modalities are combined. Other algorithm is propose by Adami et al. (2001), Hidden Markov Model (HMM) which seeks to select the information to translate index sequence. For this, the authors divide the audio and video and identify different types of scenes and dialogues in order to classify them by the algorithm (e.g., relevant or not), set index and descriptions to facilitate analysis. In another case, the information obtained from videos using semantic pattern is a combination of low-level features (e.g. color; motion and audio) associated with events of a certain kind. An event is modeled to have 4 dimensions of semantic contents (e.g. Action, Location, Time and Shooting-technique). Accordingly the events in the video we can be classified in one of four defined dimensions. The last two are intended to establish semantic relations between events from a video and the possible genres that can belong according to the characteristics obtained (Shirahama et al., 2007).

Some requirements engineers employ the pre-conceptual schemas (Zapata et al., 2006) as semiformal representation that allows modeling the natural language of the stakeholder and translate their relationships, concepts and constraints in a schema that easily understand and validate the information for everyone. The basic symbolism of pre-conceptual schema is presented in exhibit 1.

The concepts refer to nouns and noun phrases included in the speech of the person concerned. The structural relationships represent the ongoing relationship between concepts through verbs "has" and "is". Dynamic relationships refer to verbs of activity associated with an actor and an action. The conditional represent a causality or domain constraints. The implication is used as a cause-effect relationship between dynamic relationships. The connection used to connect concepts with relationships and vice versa. The notes are the set of values that can take a concept and concept-note connections are connections that allow attaching a note with the concept (Zapata et al., 2006).

Likewise, some games are related to the world of audiovisual documents like Hollywood-blockbuster®—aiming to train the best film with the best actors, producers and scenarios (knizia, 2000)—and Scene it?®—a board game and Xbox 360 videogame aiming to search for the most movie-buff participant by asking questions, anagrams, clips, and images related to the movie world (Mattel, 2008). Also there are some exercises that show the uses of video in the classroom, Boscia & Stinkney (2010) propose one exercise using video clips for demonstrate and understand of course concepts outside of the classroom, the students search, watch and discussing videos,

EXHIBIT 1
NOTATION USED IN THE PRE-CONCEPTUAL SCHEMA

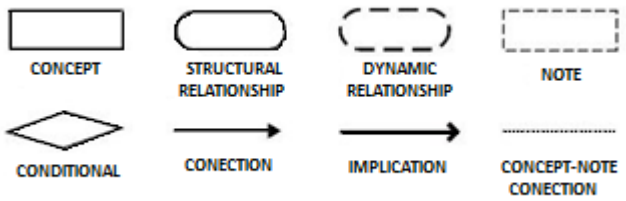


EXHIBIT 2
METHODOLOGY USED FOR THE GAME DESIGN.

Steps	Description
1	Identify the theme of the game
2	Establish the purpose of the game
3	Identify the instructional objectives
4	Identify and define general concepts of the theme
5	Select candidate techniques
6	Select the appropriate technique (s) according to characterization
7	Incorporating of specific knowledge
8	Development of initial test
9	Consolidate of the final version
10	Elaborate an evaluation survey

after they try to provide a description of the concept using they skills. Heimman & Pittenger (2000) claims that the videos have different source of communication, for example the silence, eye communication, and voice dynamics. However, few games are oriented to the information extraction from documentary videos and the information potentially learned from such videos.

RIVIDOC

RIVIDOC is a game that was designed with the methodology proposed by Gómez. M. (2010) consists of ten key steps are used. See exhibit 2.

The purpose of the game RIVIDOC is encourage the participants that select the relevant information pertaining to a trivial domain. Also, strengthen the interpretation of the information received and their different sources of procurement, text, image and audio. Finally other objective is oriented in the use the information that is relevant for modeling a system. The principal concepts are: documentary videos, information, audio, image, text. In the beginning of the game, we select three possible games: ¿Guess Who?, Pictionary and Clue. Using the methodology proposed by Gomez, M. (2010), Clue obtained the best score.

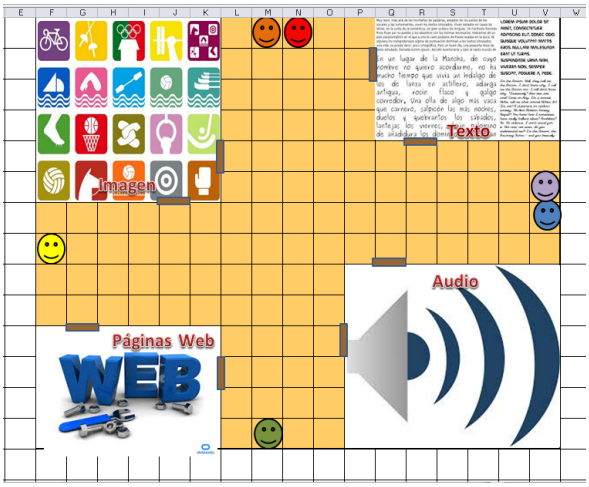
After selecting Clue like the best option, the board of the original game was modified to incorporate scenarios knowledge as the three fundamental components of documentary videos: text, image and audio, in addiction we include websites as a complementary source of information. See exhibit 3.

Also as elements of the game have video clips identified with a type (audio, video, and image) representing some of the information that can be extracted to visualize, hear and read. Additionally you have text files that represent additional conceptual information. See exhibit 4

Finally we have an incomplete pre-conceptual schema that should be filled with concepts and relationships own dynamic domain of video being displayed.

As participants move around the board are getting smaller segments of video or text that help to complete the pre-conceptual schema.

EXHIBIT 3
BOARD OF GAME RIVIDOC

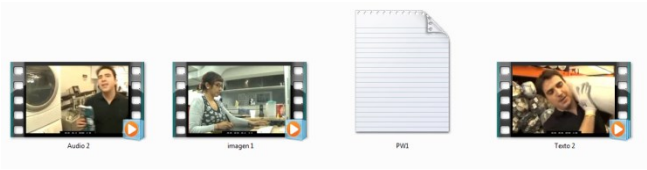


Now, we present a description of rules and actions during the game.

(1) First, we give a brief an explanation of the components of a pre-conceptual schema, together with walkthroughs and practical example is given.
(2) Each team is given the pre-conceptual schema, headphones, video clips, a plug audio and Tablet. They have 10 minutes to visualize and evaluate the pre-conceptual schema given. We can see the preconceptual schema in the appendix 1 with the words to require for complete it.
(3) Select the team that starts. Turns advance clockwise. Each team selects a player who will be responsible to share clips with other players, a player who will be in charge of throwing the dice and moving their analyst on the board and enter scenarios. The other two analysts must complete pre-conceptual schema information.
(4) Analysts must enter different scenarios on the board (audio, image, text and web pages) for video clips of the other players, entering one of the scenarios, the analyst chooses one of the panels to shows one of the clips that are part of that scenario. For example, if the analyst number one enters in the text scenario, you can choose which team (maybe the team number three) to show the clip of text that has in its possession. If the team number three doesn't have a clip identified as text type, the right team must show your clip type text, and so on if need be. The idea is to display the clips of the other players.
(5) When a team shows a video clip to another team, must number the clip that is showing, e.g. text 1, this clip should be allowed to play only twice. If you want to see the clip again the team must re-enter the stage and request it.
(6) If one team has two clips of the same type (example: audio 1, audio 3) shall ask what you want to see, so that the opposing team can ask the other clip on another occasion.
(7) Text files representing a web page should be left display for 15 seconds.
(8) Clips may have information that can be extracted from audio, video or image, in addition to its type, i.e., if the clip is text type, indicates that text information is relevant, however may have visual or auditory information that help to solve the pre-conceptual schema.

The winner is the team that after seeing the clips of the other players, first and correctly fill out the pre-conceptual schema.

EXHIBIT 4
EXAMPLE OF VIDEO CLIPS AND WEB SOURCE



RESULTS

The first interaction of the game was developed in the postgraduate course "Management Games" offered in the Universidad de Colombia, School of Mines. In a group of sixteen students divided into four groups of four students. In this interaction the participants had the list of terms. After the luck factor did not favor some, the team number two won the game. A group survey indicated that the game have a level difficulty of as "fair", similarly the level of fun was evident like "good", on the other hand, the students say that they learned "the importance of paying attention to the details of the videos" also "listen and understand business processes". Additionally all students was agreed on the closeness of the game to reality because the video detailing a real business process.

Likewise, we count the number of words that were placed correctly in the pre-conceptual schema. The words were divided into dynamic relationships and concepts, this correct location depended on the understanding of information seen, heard or read.

Exhibit 5 shows the results obtained for each team, considering that in total have 26 words to locate within the pre-conceptual schema; its distribution was as follows:

From these data we can see the percentage of concepts and dynamic relationships that were placed correctly (See exhibit 6.), thus it can be seen that the dynamic relationships were easily identified while the location of the concepts granted greater difficulty for the teams.

On average concepts equipment was twelve, indicating that it is possible to recognize important concepts from a video, on the other hand, the average dynamic relationships is five of seven dynamic relationships to locate, therefore we can say that it is identify possible actions and sequence of actions from the images and audio found in various videos.

In the second interaction the game was developed in the undergraduate course "requirements engineering" offered in the Universidad de Colombia, School of Mines. In this interaction, the video clips were public for all groups, no further word list were presented. Each team entering a stage was entitled to ask three words, which were validated by assistant. Then students should place the concepts and dynamic relationships correctly.

In conducting the survey, students awarded a "very good" degree of fun, considered equally important to consider the processes and process details. Additionally, having a monitor validating the information extracted from the video the students stated that the level of the video was actually something they could understand the manufacturing process through the video and could validate the information the stakeholder (played by assistant)

The results obtained are represented in the exhibit 7.

In this case there are eighteen concepts and seven dynamic relationships for been locate correctly in the pre-conceptual schema. Likewise the graphics representing the percentage of dynamic relationships and concepts performed correctly located. Because this interaction neither team managed to complete the preconceptual schema because the identification of the right words more complicated, but the closest team for win was the team number three.

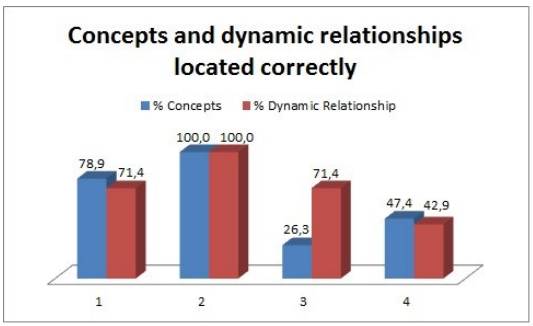
CONCLUSIONS AND FUTURE WORK

Nowadays, students are faced with multiple sources of information, including documentary videos. During the game the students had the opportunity to appreciate this source of information, identify your main components (text, audio and

EXHIBIT 5
NUMBER OF WORDS LOCATED BY TEAM – FIRST INTERACTION

Team	Concepts	Dynamic Relationship	Total
1	15	5	20
2	19	7	26
3	5	5	10
4	9	3	12

EXHIBIT 6
CONCEPTS, DYNAMIC RELATIONSHIPS AND AVERAGES. FIRST INTERACTION



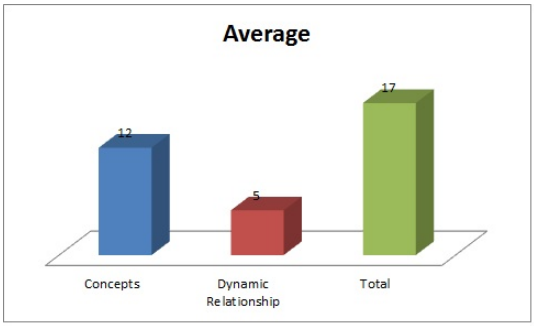


image). Additionally they learned to give importance to details, actions and own video players.

In pursuing these videos as a source of information and encourage oriented games recognition processes. Also seeks to give clues about the activities that are confronted the requirements engineer when starting a new recognition systems.

Future work is aimed at improving the game, its board and integration of video clips of the board. Additionally, it is intended that the game have more terminological offer more options for the identification of concepts and dynamic relationships, allowing some synonymous.

REFERENCES

Adami, N; Bugatti,A; Migliorati, P; Branze, V. (2001). *Low Level Processing Of Audio And Video Information For Extracting The Semantics Of Content. IEEE Fourth Workshop on Multimedia Signal Processing,* (pp 607-612). Cannes.

Boscia. M. W., Stickney, L. T. (2010). Moving Outside the Classroom: An Experiential Exercise Involving Video Clips. *Developments in Business Simulation & Experiential Learning*. 37, 228-232.

Carmichael, A., Newell, A.F., Morgan, M. (2007). *The Efficacy of narrative video for raising awareness in ITC designers about older users' requirements. Interacting with Computers*, 19, 587- 596.

Carrizo, D. (2012). *Comparación de efectividad de las técnicas de educción de requisitos software: visión novel y experta. Ingeniare. Revista chilena de ingeniería,*20 (3), 386-397.

Grant, B. (1998). *Documenting the Documentary: Close Readings of Documentary Film and Video.* Detroit, Michigan: Wayne State University Press.

Gómez, M. C. (2010). *Definición de un método para el diseño de juegos orientados al desarrollo de habilidades gerenciales como estrategia de entrenamiento empresarial*. Universidad Nacional de Colombia. Retrieved from http://www.bdigital.unal.edu.co/1968/

Heimann, B. A. & Pittenger, K. K. S. (2000). *Incorporating Video As A Teaching Strategy In Interpersonal Communication. Developments in Business Simulation & Experiential Learning*, 27, 11-14. Reprinted in the Bernie Keys Library, 9th edition. (Available from http://ABSEL.org).

Knizia. R. (2000). Hollywood Blockbuster (Dream Factory). Retrieved from http:// www.boardgamereviewsbyjosh.com/2013/12/ hollywood-blockbuster-dream-factory.html

Mattel. (2008) Scene it?. Retrieved from http:// service.mattel.com/instruction_sheets/M1266-0920.pdf

LI, W; Chang, H; Lien, K, Chang, H; Wang, C. (2013) *Exploring Visual and Motion Saliency for Automatic Video Object Extraction. IEEE Transactions on image processing: a publication of the IEEE Signal Processing Society*. 22(7), 2600 – 2610.

Liu, Z; Wang, Y. (2001). *Major Cast Detection In Video Using Both Audio and Visual Information. IEEE International Conference on Acoustics, Speech, and Signal Processing,* 3, 1413 – 1416. Salt Lake City, UT.

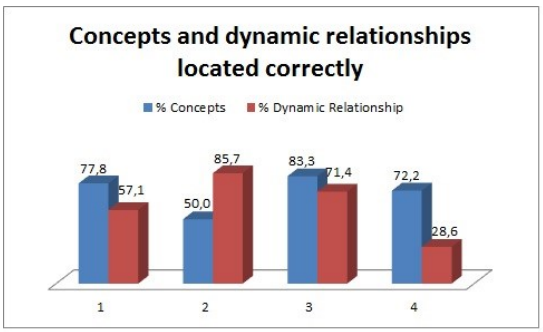
Rabiger, M. (1987) *"Dirección de documentales"*. España: Instituo Oficial de Radio y Televisión.Radio Televisión Española.

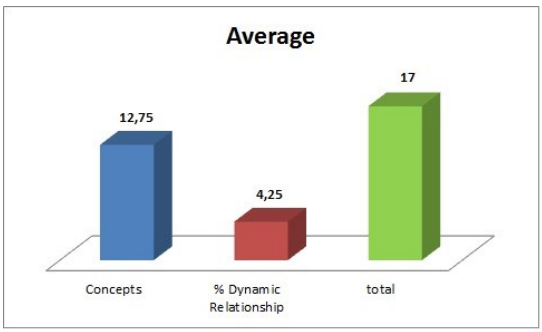
Real Academia Española. (2012). *Diccionario de la lengua española (22th ed.),* Retrieved from http://www.rae.es/

EXHIBIT 7
NUMBER OF WORDS LOCATED BY TEAM – SECOND INTERACTION

Team	Concepts	Dynamic Relationship	Total
1	14	4	18
2	9	6	15
3	15	5	20
4	13	2	15

EXHIBIT 8
CONCEPTS, DYNAMIC RELATIONSHIPS AND AVERAGES. SECOND INTERACTION



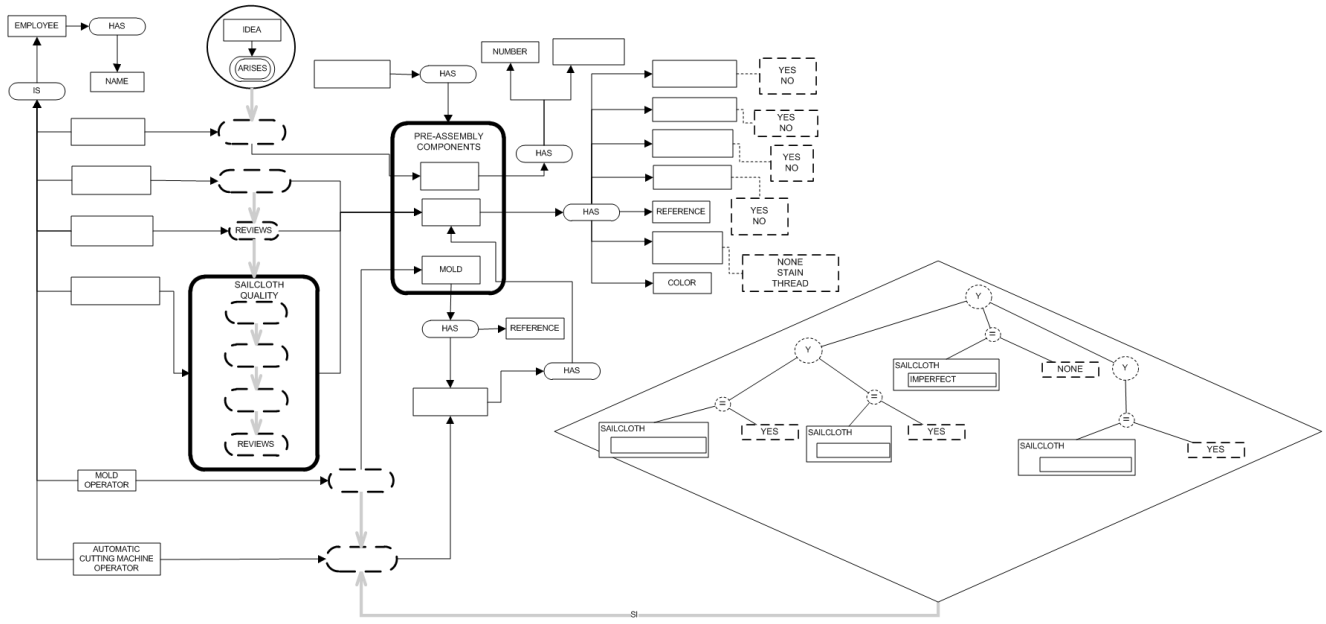


Shirahama, K., Otaka, K., Uehara K. (2007). *Content-Based Video Retrieval Using Video Ontology. Ninth IEEE International Symposium on Multimedia Workshops*, 283-289. Taichung, Taiwan.

Stangl, H., Creighton O. (2011). *Continuous Demonstration. Fourth International Workshop on Multimedia and Enjoyable Requirements Engineering - Beyond Mere Descriptions and with More Fun and Games (MERE),* 38-41. Trento.

Zapata, C., Gelbukh, A., Isaza, F. (2006) *"Pre-conceptual schema: A conceptual-graph-like knowledge representation for requirements elicitation"*. Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 27-37.

APPENDIX A
Pre-concetual schema



WORDS FOR COMPLETE THE PRECONCEPTUAL-SCHEMA

Dynamic Relationship	Concepts
CUTS	DESIGNER
DRAWS	DESING
EXPOSES	EMPLOYEE
RUB	ERGONOMY
PRINTS	IMPERFECT
WASHES	IMPERMEABILITY
TRANSPORTS	OPERATOR OF IMPERFECTS
	SAILCLOTH
	BACKPACK
	SAILCLOTH OPERATOR
	CUTTING PART
	QUALITY OPERATOR
	WASH RESISTANCE
	PHYSICAL RESISTANCE
	UV RESISTANCE