Nazlena Mohamad Ali¹ and Alan F. Smeaton^{2,1}

¹Centre for Digital Video Processing and ²CLARITY: Centre for Sensor Web Technologies Dublin City University Ireland {nmohamadali,asmeaton}@computing.dcu.ie

Abstract. This paper describes our work in examining the question of whether providing a visual informatics application in an educational scenario, in particular, providing video content analysis, does actually yield real benefit in practice. We provide a new software tool in the domain of movie content analysis technologies for use by students of film studies students at Dublin City University, and we try to address the research question of measuring the 'benefit' from the use of these technologies to students. We examine their real practices in studying for the module using our advanced application as compared to using conventional DVD browsing of movie content. In carrying out this experiment, we found that students have better essay outcomes, higher satisfactions levels and the mean time spent on movie analyzing is longer with the new technologies.

Keywords: Video browsing, Film studies, usage study, educational technologies.

1 Introduction

With the increase in the current research agenda in visual informatics, teaching and learning is an application that can benefit from this. Numerous applications can be applied at school either in primary or secondary levels. In educational applications, the strategy of using add-on features and other visual informatics that incorporate technologies in particular video content analysis often strengthens the overall appeal of the video medium, and can provide richer interactions with very large data sets or archives for example in navigating and browsing a video sequence. People are now creating, editing, storing, indexing, searching, browsing and playing video directly and indirectly more easily then before. With the growth in management tools for digital video and its potential valuable usage as a learning tool, digital video can offer exciting ways for students to study better, especially in the context of film studies.

2 Emergence of Video Content Analysis Technologies

A video element describes sequences of moving pictures. At a logical level, a video document can be divided into a set of basic components such as episode, scene, shot, and frame. Video data can be retrieved using a number of approaches such as by using metadata and browsing by keyframe, text transcript search, keyframe matching, semantic feature filtering, object matching and combinations of these techniques [3]. There is a complexity in video elements, which need proper organization as compared to when dealing with only text elements. In a huge video database, necessary indicators on the audio, visual and textual elements will help the video material to be more searchable and browsable [4].

In content analysis systems, video data is typically structured automatically into temporal shots, which represent basic access and retrieval units. This processing step is called shot boundary detection (SBD). In most approaches, SBD is based on a measure of similarity between two or more adjacent frames. Usually a shot change is indicated by large changes in the similarity value [2]. Colour histograms [7, 12, 2], edge detection [12, 2] and using macroblocks [2] are among the many features that can be used for similarity comparisons.

Among the possible video content analysis techniques is segmenting scenes into several event types such as *Exciting*, *Montage* and *Dialogue*, based on film grammar [5]. The key multimedia techniques used in this work are:

- Scene detection automatically segment a movie into a number of scenes. Camera shot boundary detection is used first and segmented shots are clustered back together by considering their visual similarity and temporal distance;
- Scene classification automatically classify the nature of a scene into *Action*, *Dialogue* and *Montage*. Within-scene shots are analyzed in terms of the amount of motion (in the case of Action), in terms of alternating shots (in the case of Dialogue), and in terms of motion speed and existence of music spanning multiple shots (in the case of Montage). Movie shooting and editing conventions are also used as heuristics for the classification.
- Keyframe extraction automatically select a most representative still image from a sequence of video. For each scene and shot, the most average (common) visual frame is determined and selected as the keyframe.

These are among the active research areas in the field of multimedia at the moment and are steadily improving their accuracy and robustness. Taking advantage of these automatic content-based techniques, numerous possible application scenarios can be imagined and a large number of novel demonstration systems have been built. Our work incorporates these technologies into a system we called MOVIEBROWSER2.

Other work related to movie video browsing includes work at INRIA [14], and the *Virtual Screening Room* [17]. These tools provide novel movie content browsing and searching features. A project at the INRIA laboratory is related to the development of an integrated tool for watching, browsing and searching a movie, synchronized with its scripts. It was demonstrated with the movie "The Wizard of Oz". The *Virtual Screening Room* project is an informative browser for playing a movie with many useful features such as clip searching and various in-depth information representations. These projects are very useful and effective educational applications

for students of Film Studies in understanding, appreciating, teaching and learning movies.

3 Film Studies Work Context

Our application domain area in this experiment is Film Studies. Film Studies is a study about film history, theory, and criticism. The typical tasks of Film Studies students are to read and analyze movies. Reading a movie in their context refers to the process of understanding and analyzing movie content closely, looking for different levels of meaning and critique from different elements like framing, depth of field, plot, shots, camera angle, lighting and so on. On a broader level it also involves an understanding of the generic conventions and narrative structure of individual movies [11].

The conventional approach for students to work is to get DVDs from the University library or rental shop. The restrictions of getting only a one-day loan and a lack of DVD resources (i.e. old production year, VHS format) contributes to the main problems in their studies.

Online technology enables movie content information to be accessed easily with the existence of the Internet, thus providing resourceful information such as the IMDB (imdb.com), the Internet Archive (www.archive.org), and other online web hosting video. Video material can be accessed by Film Studies students more easily than before.

4 Experiment Protocol

Our main experimental question to be explored is "Do students who use our newlyintroduced technology (the software tool) get more 'benefit' compared to the conventional way of performing their work task?". We address this main question by dividing into sub-questions as the following:

- Q1. Do students make use of the alternative access features afforded by MOVIEBROWSER2 and spend less time in completing the essay-writing task?
- Q2. Given the same amount of time, is the 'outcome' for those using MOVIEBROWSER2 better or worse than those who use only a standard DVD media player interface for browsing movies?
- Q3. Do the students who use MOVIEBROWSER2 have a higher level of satisfaction than those using only a standard DVD media player?

We recruited all 7 students from CM524 *Film Theory and History* (Masters Class module). The task given is to 'read' small sequences of a film. Each student needs to produce a small piece of writing based on reading movie sequences and we will compare how students perform this task using MOVIEBROWSER2 versions and using only a standard movie DVD player as shown in Fig.1.

The first task was about reading and analysing a dinner conversation scene in the movie *American Beauty*, while the second task was comparing and contrasting a few

related scenes from the same movie that had musical elements in them. Participants need to browse and analyse relevant scenes or sequences as described in the task by using the assigned interface systems as described below:

- Task 1 Find the scene and perform some analysis of its content. Analyse the techniques used by the director to make the scene more tense between all the family members in the fighting dinner scene where *Lester* (the father) throws away the food plate onto the wall. [Hint: characters smiling to/at each other in the beginning and there are some recognizable changes in dress etc.]
- Task 2 Find the scenes and perform some analysis. Analyse the director's use of music to express the characters' emotion more cinematically. [Hint: you could contrast these scenes a scene where *Carolyn* (the mother) is driving and singing in the car on her way back home and a scene where *Lester* (the father) and *Angela* (*Jane's* friend) are together in the bath]



Fig. 1. (a): DVD player interface screenshot, (b) MOVIEBROWSER2 screenshot

We choose the movie, *American Beauty (1999)* and participants were brought to a 2-hour screening of the movie before the experiments were carried out. In order to reduce and control the learning effect during the task performance with a smaller number of participants, we choose a within-groups design [15, 1] approach.

For the MOVIEBROWSER2, the interaction logs were captured automatically by the system while for the DVD interface, they were observed by volunteer researchers from our research centre who took notes on what users clicked and their overall behaviour in accomplishing the task for example writing notes on paper then typing essay material into WORD. Some guidelines were given to colleague researchers before the session begin especially on what they need to observe such as the number of clicks participants do on the player interface (e.g. play, slider bar, pause, stop, etc.) and reporting overall user actions in accomplishing the tasks in a pre-designed form. Each task needs to be completed within 30 minutes. CCTVs were used (with participants' permission) to record the session and the CCTV footage was used to verify the observations made by colleague researchers.

5 Results

In answering Q1, we examined the user log data from both system variants (DVD and MOVIEBROWSER2 interface) on each task. Data were collected from MOVIEBROWSER2 on the interaction clicks and from the observation sheets for the DVD interface where colleague researchers manually recorded the number of interaction clicks made and the button pressed (i.e. slider bar, stop and etc). The sessions were recorded (with student permission) using CCTV as part of the experimental protocol. CCTV content was then used for verification in this process and in determining participant activities during the session.

As shown in Table 1, highest interactions were found on the slider-bar when the tasks were carried out using the DVD interface. While interactions on MOVIEBROWSER2 were no longer at the slider-bar but has been shifted into 'event-keyframe' that referred to playing the scene from the keyframe view and also from the respective timeline bar (either Dialogue or Montage).

	TASK 1		TASK 2			
	DVD	MOVIEBROWSER2	DVD	MOVIEBROWSER2		
Dialogue	0 (0%)	8 (25%)	0 (0%)	2 (3%)		
Exciting	0 (0%)	2 (6%)	0 (0%)	0 (0%)		
Montage	0 (0%)	2 (6%)	0 (0%)	24 (37%)		
event-keyframe	0 (0%)	14 (44%)	0 (0%)	27 (42%)		
slider-bar	52 (58%)	1 (3%)	59 (72%)	2 (3%)		
pause button	16 (18%)	0 (0%)	11 (13%)	0 (0%)		
play button	13 (15%)	0 (0%)	11 (13%)	2 (3%)		
slow button	6 (7%)	0 (0%)	1 (1%)	1 (2%)		
faster button	1 (1%)	0 (0%)	0 (0%)	2 (3%)		
fullscreen button	1 (1%)	1 (3%)	0 (0%)	0 (0%)		
stop button	0 (0%)	4 (13%)	0 (0%)	4 (6%)		
volumn-up button	0 (0%)	0 (0%)	0 (0%)	1 (2%)		
Total	89	32	82	65		

Table 1. Interaction logs for a given tasks. Notes: Dialogue, Exciting and Montage referred to timeline bar; event-keyframe referred to play scene from keyframe view

We also determined participants' time completion for each task. The completion time when using MOVIEBROWSER2 is longer or almost to the maximum (30

minutes) for most of participants as compared to standard DVD player. Only one participant out of 7, finished the task with MOVIEBROWSER2 earlier, at 17 minutes.

These results only answer the research question in part. In a conventional DVD interface, usage on the slider bar seems higher and is used as the main interaction during navigation of movie content. This appears to be the 'only way' to accomplish the task, working with other support buttons such as pause/play/stop etc. This feature is useful for blind seeking of movie content. In the MOVIEBROWSER2 interface, participants make use of the alternative access features afforded by the technologies in completing their task. The process of navigating, and playing clips can be seen from the interaction log captured on the added features and not utilizing the standard buttons such as slider bar. For example, in looking for music background to appear in the scene, a montage timeline bar is used and further browsing is enhanced using the particular shot keyframe view instead of using a slider-bar. The time taken to accomplish the task when using the MOVIEBROWSER2 interface was relatively longer than when using the conventional standard player even though it was predicted to be faster initially.

In order to answer Q2, we performed a qualitative analysis and looked at the quality of the essay results from students who completed an assignment as part of this module. This resulting outcome was chosen for analysis mainly to provide a general overview of participants' performance as a result of having used either a DVD or MOVIEBROWSER2 interface. Participants' essays were marked based on a known, predefined answer scheme and were categorized into three quality groups, and these categories were formally approved by the module lecturer (i.e. 'Very Good', 'Good' and 'Basic').

	Standard Player	MovieBrowser2
U1	Not much more to scene description. Basic	Solid enough but would like have liked a bit more detail [specific of lyrics and music]. Good
U2	Very good comparisons between the two scenes would have liked a little more description of Lester's dream state as define by the 'off key' music. Very Good	Lots of good stuff from different perspective with excellent conclusion. Very Good
U3	Relatively basic scenarios description. Basic	Solid and good analysis [good on detail of music instrument and effect]. Good
U4	Well done/describes.	Use of several techniques to express this [i.e. camera cut/editing pace] solid description and compare other parts of the film. Very Good
U5	Describes the scene and lacking of explanation. Basic	Good re string instruments used to express sexual interest solid and good contrast of scenes. Good
U6	Lack of details. Basic	Basic description but use of camera techniques in the explanation such as shots mid/long and editing pace. Basic
U7	Good detail but not really well structured. Good	Good contrast and very good drawn conclusion – describes the music effect. Very Good

Table 2. Essay quality remarks

Table 2 represents the results and remarks of the essay quality evaluation. Participants [i.e. U2, U4, U6] who used MOVIEBROWSER2, and then the standard DVD player stay with the same essay quality. Participants [i.e. U1, U3, U5, U7], who used the standard DVD player, followed by MOVIEBROWSER2 get a better essay outcome. This results shows there is some increased performance relatively either from 'Basic' to 'Good' or 'Good' to 'Very Good'. In order to provide more supporting evidence, we look further at the essay quality in a different way. The essay remarks also show interesting feedback given from the module lecturer. Most of the remarks given on essays after using MOVIEBROWSER2, from the lecturer's point of view show that participants' give more deep and critical essay analysis. In the essays, we are able to find these expressions or statements from a variety of analysis which makes the essay quality better as for example:

- Point of view: ("It might suggest/reflecting . . . /I think . . . /noted that . . . /The director succeeds in given . . . /tells us . . . ")
- Compare and contrast scene: ("Elsewhere in the film . . . /The scene is preceded by . . . /As the scene progress . . . /Directly contrast the earlier scene . . . ")
- Use of different techniques: ("The camera cut relatively fast . . . /Music helps to distinguish . . . /The pace of the scene is faster, hinting an explosion to come . . . /Non-descript colours (grey, brown etc) that complimented with music . . . ")

Participants who completed the task using MOVIEBROWSER2 generally have solid remarks that represent richness or variability in the written essay. On the other hand, when using the standard player, the remarks are basically only basic explanation details on sequences. An essay that was produced using the standard player generally has the basic type of description of scenes; the essay quality was limited in terms of variety of expressions, points of view and lack of detailed explanations. We are not comparing on each individual participant in this experiment, but referring to the individually produced outcomes from a given task.

By analyzing the essay outcome and the remarks given by the lecturer in a qualitative way, we answered the experimental question and reveal that the outcome (essay quality) of those using MOVIEBROWSER2 are better or at least stay at the same level (not worse) than participants' current performance. In relation to this as the findings found in the previous question, most participants who get better essay outcomes take a longer time to complete the task in MOVIEBROWSER2.

For Q3, we captured participants' responses to using both interfaces using semantic differentials and Likert point-scales of 1–7 (the higher values the better) for a variety of questions from post-task questionnaires. We used a paired sample t-test [8], for the statistically significant differences of means scores with $p \le 0.05$. However, in this explanation, we would not emphasize a significant difference due to the small number of sample users. We reported the means and standard deviations (SD) on each system variant. Participants perceived satisfaction levels on the system variants are more towards MOVIEBROWSER2 with generally higher mean scores in all aspects (Table 3).

	System	Ā	Ã	\bar{B}	\tilde{B}	р
Differential	easy to use	5.43	1.27	6.29	0.76	0.017
	easy to learn	6.00	1.15	6.29	0.95	-
	effective	5.14	1.68	6.43	0.53	0.049
	satisfying	5.14	1.57	5.29	1.50	-
	stimulating	5.14	1.57	6.00	1.15	-
	novel	4.57	1.72	6.43	0.53	-
	intuitive	5.00	1.41	6.00	0.82	-
	helpful	5.14	1.77	6.29	0.76	-
Likert	access movie content	4.14	1.77	6.14	0.69	
	analyse movie efficiently	4.00	2.31	6.29	0.49	0.022
	explore movie scenes efficiently	4.00	2.24	6.57	0.53	0.019
	complete task quickly	4.71	2.06	5.86	0.69	-
	enhance reading ability	3.86	2.19	5.71	0.95	0.045
	better outcome/essay	4.57	1.27	6.00	0.82	0.008
	discover new ideas	3.71	2.36	5.29	1.38	0.042
	focus	3.71	2.50	5.29	1.38	0.033

Table 3. Post-task results. Note: A - DVD, B - MOVIEBROWSER2

In affirming the overall satisfaction results shown in Table 3, we asked participants how they felt about the task they performed in the experiment using both system variants. Responses from students favoured MOVIEBROWSER2, as can be seen from some extracts below:

"If you are looking for a particular scene (i.e. dialogue), you can narrow the search parameter by selecting a category, making it quicker to complete the search". [U2] "MOVIEBROWSER2 is more useful because you could browse frames of the film together making it easier to contrast scenes. The DVD media player took longer to find scenes— it encouraged more random watching". [U4]

"MOVIEBROWSER2 is easier to use and (provides) more information".[U6]

Finally, we asked participants the question of which system they would prefer for their film reading task as part of their studies. All of the 7 participants chose MOVIEBROWSER2 as the favoured tool to be used. This question was asked mainly to gauge participants' future intentions.

6 Discussion

The findings from the essay outcome revealed that there are slightly improving or better results which is also supported by the remarks from the module lecturer that shows students have more variability (more opinions, expressions) in their written essay when using MOVIEBROWSER2. This is a better outcome from the lecturer's point of view. It would be better if this perception could be further explored to more detail on the improvement from the student's point of view particularly in their learning. We understand that in order to measure performance of a movie browser interface, many factors need to be considered. The link between 'cognitive' and 'pedagogic' learning and any measurement of the deepening of students' perceptions of the audio-visual stimulus, using technical or structured approaches, are of course difficult to prove in terms of any direct benefits. Any improvements also so much depend on the aptitude and linguistic dexterity of individual students and their motivation. These issues will give another interesting perspective and motivation for future work.

Our findings also show that satisfaction levels are higher after using the newly introduced tool with higher mean scores in all aspects of statements given as compared to when using a conventional standard player. User preferences and subjective opinions reflected from the comments favoured MOVIEBROWSER2. All 7 students intended to use the software application in their future learning. In discussing user perceived satisfaction in an experimental setup, especially from the perspective of student learning, we might have expected mixed satisfaction levels as the outcome does not always mean higher satisfaction when a new technology is introduced. In some cases, even though a new system with new technology is

introduced, users still prefer their old ways as their tendency or practice does not change easily. Many factors influence such results. For example, previous studies in e-learning and work carried out in [9, 13] reported that students are less satisfied with the use of a technology in learning due to them already being accustomed to their traditional ways or experiencing frustration or worry during online learning. On the other hand, other studies as in [16, 6] found a higher level of subjective satisfaction from students with a new e-learning tool compared to the traditional classroom learning.

Besides expected findings from the lab setting experiment in terms of better performance in the essay outcome and students' perceived satisfaction levels, we also saw surprising findings that show the time taken to complete the task when using the newly-introduced software application is longer than the conventional way for the majority of students. The results of longer time taken to complete the tasks might indicate that students are engaged with the activity being performed using the newly introduced tool. From the learning perspective, the longer time spent represents the meaning of engagement (similarly to lecturer's feedback/remarks that show engagement, thus it was used as a 'measurement' in the essay outcome). Looking from this perspective, we cannot say that by having used the newly-introduced tool that afford extra features for navigating and browsing movie scenes in carrying essay analysis, the task can be completed faster than the conventional way. In this experiment, we interpret this phenomenon from another domain perspective, the subarea of psychology, which is flow theory that reflects engagement in carrying out an activity. Flow Theory according to Csikszentmihalyi [10] is "The state in which people are so involved in the activity that nothing else seems to matter; the experience itself is so enjoyable that people will do it even at great cost, for the sheer sake of doing it".

We cannot claim from this finding yet that this phenomenon shows our design influenced the engagement or playing element which then brought into a sense of feeling intrinsic motivation in completing their assignment from this experiment. We reasoned that they are immersed with the tool from the feedback we collected either from written comments or verbally informal responses. Some students did request permission to use the tool after the experiment and are very interested to learn more. Our assumption here relates to the learning theory of engagement and flow as described in the psychology field. Having completed the task longer than the conventional way may look as a 'failure' in terms of scientific computing where it should be faster to complete, it does show interesting and good evidence from the perspective of the learning process for a newly-introduced educational technology.

We consider some limitation in this experiment, which is only 7 users, used in a pre-defined lab setting and limited time frame. 30 minutes duration might not be enough to evaluate essay performance. Future research should include more student cohorts and the experiment should be carried out over a longer time frame (i.e. a longitudinal study). Our interesting findings related to engagement factors were considered as surprising or unexpected, and we can suggest another research perspective of 'measuring' the flow or engagement level among students for example.

7 Conclusion

In conclusion, from carrying out this experiment, we found that students spent more time and use the alternative features afforded by new technologies based on visual analytics in completing their tasks, written essays resulted in various aspect of critical analysis and better outcome/remarks from the lecturer, positive quantitative and qualitative feedback with higher satisfaction levels in terms of interest and educational effectiveness.

We believe that, by implementing innovative technologies such as this need further work as consideration on issues in learning particularly in individual learning style, background and skill. The software application we developed could provide potential in learning for students developing the core skills of textual analysis, within the film studies domain, thus contribute in pushing the trend (i.e. starting to recognize the importance of user experience in multimedia) further to that direction. Integrating both knowledge from the technology on automatic movie content analysis, with the practice from the realistic contextual end-users, could provide a bridge to establish the Film Studies discipline and in the setting of Film Studies, it is shown to be useful.

Acknowledgements. The work was supported by the Ministry of Higher Education and University Kebangsaan Malaysia and by Science Foundation Ireland as part of the CLARITY CSET (07/CE/I1147).

References

- 1. A. Dix, J. Finlay, G.D. Abowd, and R. Beale: Human-Computer Interaction. Prentice Hall, 2004.
- A.F. Smeaton: Indexing, Browsing, and Searching of Digital Video. Annual Review of Information Science and Technology, 38:371–410, 2004.
- A.F. Smeaton: Techniques Used and Open Challenges to the Analysis, Indexing and Retrieval of Digital Video. Information Systems, 32(4):545–559, 2007.
- A. Haubold, P. Dutta, and J.R. Kender: Evaluation of Video Browser Features and User Interaction with VAST MM. In MM '08: Proceeding of the 16th ACM International Conference on Multimedia, pages 449–458, Vancouver, British Columbia, Canada, 2008.
- 5. B. Lehane: Automatic Indexing of Video Content via the Detection of Semantic Events. PhD Thesis, Centre for Digital Video Processing, Dublin City University, 2006.
- F. Amir, S.M. Iqbal, and M. Yasin: Effectiveness of Cyber-learning. In Frontiers in Education Conference, 1999. FIE'99. 29th Annual, volume 2, 1999.
- J. Mas and G. Fernandez: Video Shot Boundary Detection Based on Color Histogram. In Proceedings of the TRECVid Workshop, Gaithersburg, Maryland USA, 2003.
- 8 J. Pallant: A Step by Step Guide to Data Analysis using SPSS version 12. Second Edition. Open University Press, 2005.
- J.C. Rivera and M.L. Rice: A Comparison of Student Outcomes and Satisfaction Between Traditional and Web Based Course Offerings. Online Journal of Distance Learning Administration, 5(3), 2002.
- 10. M. Csikszentmihalyi: Flow: The Psychology of Optimal Experience. Harper Perennial

New York, 1991.

- P. Brereton: How to write textual analysis of a film, CM135 Analyzing Media Content lecture notes. School of Communications, Dublin City University, 2008.
- P. Browne, A.F. Smeaton, N. Murphy, N.E. O'Connor, S. Marlow, and C. Berrut: Evaluating and Combining Digital Video Shot Boundary Detection Algorithms. In Proc. Irish Machine Vision and Image Processing Conference (IMVIP2000), 2000.
- R.H. Maki, W.S. Maki, M. Patterson, and P.D. Whittaker: Evaluation of a Web based Introductory Psychology Course: I. Learning and satisfaction in on-line versus lecture courses. Behavior Research Methods, Instruments and Computer, 32(2):230– 239, 2000.
- 14. R. Ronfard: Reading Movies: an Integrated DVD Player for Browsing Movies and Their Scripts. In Proceedings of the 12th Annual ACM International Conference on Multimedia, pages 740–741, 2004.
- 15. S. Smith-Atakan: Human-computer Interaction. Thomson, 2006.
- S.R. Hiltz and B. Wellman: Asynchronous Learning Networks as a Virtual Classroom. Communications ACM, 40(9):44–49, 1997.
- 17. MIT Center for Educational Computing Initiatives. The Virtual Screening Room. Available from http://caes.mit.edu/projects/virtual screening room/index.htm [Accessed 21 November 2008].