Video augmentation to support video-based learning

Ilaria Torre
ilaria.torre@unige.it
University of Genoa
Genoa, Italy

Ilenia Galluccio
ilenia.galluccio@edu.unige.it
University of Genoa
Genoa, Italy

Mauro Coccoli
mauro.coccoli@unige.it
University of Genoa
Genoa, Italy

ABSTRACT

Multimedia content and video-based learning are expected to take a central role in the post-pandemic world. Thus, providing new advanced interfaces and services that further exploit their potential becomes of paramount importance. A challenging area deals with developing intelligent visual interfaces that integrate the knowledge extracted from multimedia materials into educational applications. In this respect, we designed a web-based video player that is aimed to support video consumption by exploiting the knowledge extracted from the video in terms of concepts explained in the video and prerequisite relations between them. This knowledge is used to augment the video lesson through visual feedback methods. Specifically, in this paper we investigate the use of two types of visual feedback, i.e. an augmented transcript and a dynamic concept map (map of concept's flow), to improve video comprehension in the first-watch learning context. Our preliminary findings suggest that both the methods help the learner to focus on the relevant concepts and their related contents. The augmented transcript has an higher impact on immediate comprehension compared to the map of concepts' flow, even though the latter is expected to be more powerful to support other tasks such as exploration and in-depth analysis of the concepts in the video.

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**; • **User interface design**;

KEYWORDS

intelligent user interfaces, visual feedback, hypervideos, MOOCs

ACM Reference Format:
Ilaria Torre, Ilenia Galluccio, and Mauro Coccoli. 2022. Video augmentation to support video-based learning. In *Proceedings of the 2022 International Conference on Advanced Visual Interfaces (AVI 2022), June 6–10, 2022, Frascati, Rome, Italy.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3531073.3531179

1 INTRODUCTION

Education, together with the whole society, is experiencing a leap towards virtualization, speeding up a process that was already under way. Online learning is thus becoming a usual practice and experience for millions of students, and we can expect that this will impact on future education and learning models. In this scenario, video-based learning is increasingly taking a central role, bringing with it the pros and cons that are already known after two decades of experience with Massive Open Online Course (MOOC). This poses new challenges with the goal of enhancing the learning experience and effectiveness and meanwhile overcoming drawbacks and limits of video-based learning. For instance, educational videos are often lengthy and thus hard to navigate when the user needs to recall concepts, moreover they lack a structure and tools for highlighting relevant contents and for exploring them. Concept maps have been long studied in education and have been shown to provide multiple benefits for students in general and also for students with learning disorders [8, 13]. This study represents the first step of a project that aims to automate the extraction of concepts and relations from video sources and the subsequent provision of visual feedback and hypervideo services. The motivation is that, while video augmentation services have been shown to be useful for learners [30, 33], they are hardly provided since their development is time-consuming and requires manual effort [39]. Specifically, in this paper we present two types of visual feedback —an *augmented transcript* and a dynamic concept map (*map of concept's flow*), and their evaluation, as methods to support video comprehension in the first-watch learning context.

AVI 2022, June 6–10, 2022, Frascati, Rome, Italy

ACM ISBN 978-1-4503-9719-3/22/06…$15.00
https://doi.org/10.1145/3531073.3531179

2 RELATED WORKS

If prior to COVID-19 pandemic, the digital transition highlighted a blended learning approach [11, 19] in the last two years, video production/consumption in the educational field has had an exponential growth and numerous related studies have been possible to also understand the perception, performance , and needs of students [31]. This means also opportunities and challenges for the development of intelligent visual interfaces that try to fill the shortcomings of a do-it-yourself education. Visual feedback and interactivity have been shown to improve the level of activity when studying alone [21] and offer new opportunities also for people with disabilities [15]. Technologies for video augmentation can be included in the aforementioned ecosystem, as a combination of digital video and hypertext, which draws largely upon audiovisual media as central parts of their structure. They consist of interconnected video scenes containing dynamic hyperlinks that are available to explore the video and that may refer to further information elements (such as, e.g. texts, photos and graphics) [42]. Video augmentation methods encompass different types of visual feedback, including video transcripts, concept maps, knowledge graphs and hypergraphs [5, 6, 27, 33, 36] that are specifically addressed in our project. Data driven investigations have explored possibilities for video augmentation through personalized fragment navigation [41], enhanced exploration of e-learning contents [28, 34], automatic display of

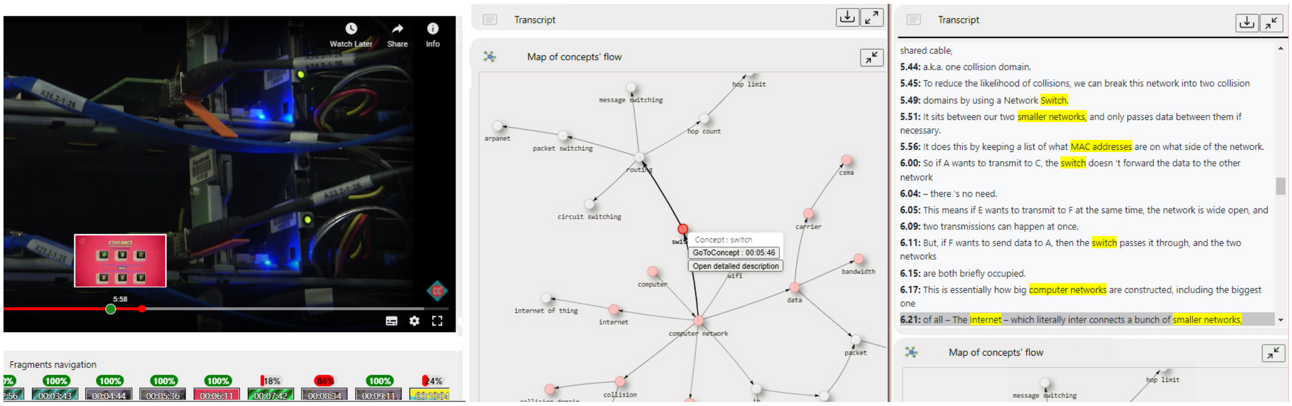

Figure 1: UI screenshot that displays the *Augmented transcript* (VF1) and the *Map of concept's flow* (VF2) as they appear when they are toggled on/off and the video is being played.

relevant frames and visual summaries representing points with high learner activity [20]. The advancements in machine learning and natural language processing supported the production of interactive transcript of videos and keyword summary [20], and text summarization [7, 20, 23]. From the pedagogical point of view, the application of concept maps [17] has been extensively tested in various scientific domains [9, 12, 35, 40] and at different levels of education (from primary school to university) [3, 4], proving a positive impact on learners, even in the context of students with special needs and specific learning disorders [8, 13]. However, the use of concepts maps and knowledge graphs to support video-based learning are less covered [26, 37, 38], likely due to the effort needed to build them using manual approaches and to the still low accuracy of fully automatic methods [39]. A trade-off to balance cost and utility is using a semi-automatic approach, that we also currently adopt when high accuracy is required. Semi-automathic concept map generation is used for instance by Hayama and Sato [18]. The authors propose a system that allows the creation of a concept map by the learner, not from scratch but with the support of a series of candidate components such as concept-labels and related words from lecture speech texts. Among the approaches for prerequisite extraction from videos, [10, 32] rely on Wikipedia and ontologies to identify domain concepts and relations, while very few works use unsupervised methods, one example is [2]. More research on prerequisite extraction can be found on educational texts, also applied to video transcript [1, 16, 25]. To the best of our knowledge, this paper is the first attempt to compare augmented transcript and concept map of prerequisites to support video-based learning, especially with the characterization of concept map as a dynamic concept map, that will be discussed in more detail in the next section. Conversely, the use of transcript together with video has already been shown to be effective in past research [22].

Video augmentation methods described in this paper have been designed by following the principles of instructional design, including the use of methods for structuring, highlighting and annotating content in video-based learning [5, 17, 36].

3 VISUAL FEEDBACK METHODS

The web interface we developed is aimed to provide a set of *functionalities* that are suited to support the learner in a specific phase, i.e., *learning context* [37] (first-watch, rewatch) or for a specific *goal* (exploring, searching, analyzing, annotating). To this aim, the user interface combines different canvases containing the video player and other visual tools and services that implement the designed functionalities. In this paper, we take into account specifically two Visual Feedback (VF) methods: the *augmented transcript* (VF1) and the *map of concepts' flow* (VF2). They will be described in this section and then tested to investigate their effect on the comprehension of a video lesson in the specific 'first-watch' *learning context*. However, it is worth noting that the two visual feedback methods include also features, particularly VF2, that have been designed to support also video exploration and in depth analysis, not addressed in this paper. As a common feature, both VF1 and VF2 are intended to make explicit and highlight to learners the knowledge embedded in the video. Thereby, the knowledge extracted from the video is used to augment the User Interface of the video player.

3.1 Augmented transcript

Video transcript is augmented by highlighting the concepts explained in the video and by linking them to the points of the video where the concept is explained. The effectiveness of transcripts to support comprehension has already been studied in previous research, among which, recently [22], while the automatic extraction of keywords from videos to help users understand and learn terminologies has been addressed for example in [39]. In the system we developed, the augmentation of the transcript is performed by using state-of-the-art libraries for keyword extraction and natural language processing and concepts are represented using SKOS data model [29]. The right part of Figure 1 shows an example of the augmented transcript.

3.2 Map of concepts' flow

The second method is a temporal graph whose nodes are the concepts explained in the video, connected through edges to other

concepts that are prerequisites of that concept. Following the principles of hypervideos, concepts are linked to points where they occur in the video, moreover, the progress bar of the video player is enhanced with some markers of different size that show where the concept is explained and where it is just mentioned.
The major contribution with respect to the literature is the dynamic nature of the map, that reflects the way how concepts are explained in the video: since concepts evolve together with the video flow (as their explanation goes deeper), they are initially presented with lower complexity, resulting in a map that shows a simple graph. Later on in the video, the same concepts may be deepened with other notions. Thus, concepts do not have a static representation during the whole video and the map evolves dynamically and gradually according with the video flow. This is the reason why we do not refer to this temporal graph simply as a concept map, but as *map of concept's flow*, as shown in Figure 1. The graph is built as a view over a semantic RDF graph, extracted from the video and annotated using the W3C Web Annotation Vocabulary[1]. The prerequisite relations in the graph are estimated using an algorithm designed for textbooks [1, 24], that we adapted for the extraction from the transcript. Basically, the approach uses a Hidden Markov Model to identify bursting intervals of keywords [24] that reveal the focus of the stream on a concept, then uses temporal algebra to model the relations between concepts showing bursting appearance along the video, in order to identify prerequisite relations between them.

4 EXPERIMENT

We conducted a controlled experiment to empirically study the effectiveness and limitations of the Visual Feedback methods to support learners in their *first-watch* of a video lesson.

Our *hypothesis* is that both the *augmented transcript* (VF1) and the *map of concepts' flow* (VF2) can increase the learner comprehension, helping the learner focus on the relevant concepts and their related concepts and contents. While for the *augmented transcript* there are already evidences in the literature for its effectiveness, for the *map of concepts' flow*, as defined above, the literature is scarce and the risk for learners to get distracted by the dynamic map is high [14], with the effect of potential reduced attention and comprehension.

The independent variable of our experiment is thereby the Visualization Feedback method, with VF1 and VF2 as the experiment conditions. In addition to them, for the experiment we also added a third condition that is the simple video player with subtitles switched on, used as baseline. Thus, the 3 conditions used in the experiment are: VF1 (augmented transcript), VF2 (map of concepts' flow), VF3 (baseline video player).

User Interface setting. For this experiment, participants are required to use only the *video player*, the *video transcript* and the *map of concept's flow*. The frames containing the transcript and the map can be toggled on and off in order to view them separately.
The video used in the experiment is a video lesson on computer networks, available on YouTube.[2]

[1]https://www.w3.org/TR/annotation-vocab/
[2]https://www.youtube.com/watch?v=3QhU9jd03a0

4.1 Participants

For the experiment we recruited 12 university students of foreign languages, 9 females and 3 males, in an age range between 21 and 32 years old (AVG=24.6, SD=2.7).

Only subjects that had in their career an introductory computer course were considered for recruitment. This to ensure that all of the participants could have the base knowledge required to understand the video content but not too much to be able to answer to a questionnaire by using only their background knowledge. None of the participants was affected from learning difficulties or eyesight problems (self-reported in an anonymized questionnaire).

4.2 Experimental design

4.2.1 Procedure. The experiment was conducted as an online study.
1) Participants were first asked to fill in a quick demographic survey including the informed consent and the questions about learning and possible eyesight problems.
2) Then they have been assigned to one of 4 groups of 3 people (details in Sec. 4.2.2) and, separately for each group, they have been briefed about the experiment, and then invited to try the user interface and the VFs for 3 minutes.
3) After this short training period, participants were instructed about the tasks they had to perform: all of them had to watch the provided video under 3 different visual conditions: with the *augmented transcript* toggled on and the map off (VF1), with the *map of concepts' flow* toggled on and the transcript off (VF2), with both the transcript and the map toggled off (VF3). To avoid bias, the sequence of the 3 conditions for the 3 participants in the group was different, following a schema explained in Sec. 4.2.2. As a result, the session is split in 3 sessions for each participant, i.e. one for each VF condition. After each session, participants had to answer a couple of questions (Q) aimed to Qa) assess the comprehension of the portion of the video watched in that session and Qb) collect subjective feedback about the VF condition. In each session, participants were not allowed to pause or rewind the video and neither use the interactive features available on the user interface. Each session had a start and end point at given times and covered a specific topic.
4) At the end, participants were asked to compare VF1 and VF2 methods considering different learning contexts and goals. They were given three more minutes to explore the interactive features of the tool which they could not use during the experiment. The overall time required to each participant for the experiment was about 40 minutes.

4.2.2 VF conditions and groups. The presentation order of the three VF conditions was varied between participants, according to a double Latin square (3*3) with three experimental conditions and three experimental periods, i.e., sessions. All the possible combinations of the VF conditions are covered through 2 groups of 3 participants and 3 sessions. We replicated this schema 2 times, for a total of 4 groups and 36 trials.

4.3 Results

4.3.1 Qa results: comprehension. After each session, participants were asked to answer multiple choice questions about the content

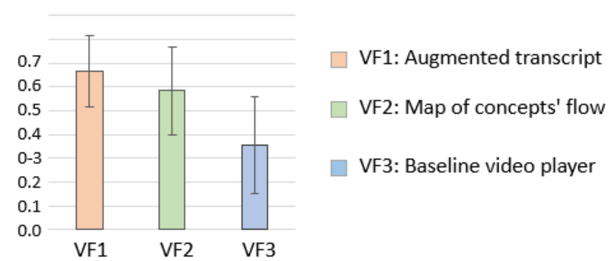

Figure 2: Qa results for video comprehension after VF1, V2 and VF3 sessions.

		5-point Likert scale (1 disagree – 5 agree)	VF1 AVG	VF1 SD	VF2 AVG	VF2 SD
Qb1	Positive	The VF helped me to understand the content	3.33	0.78	2.83	0.94
Qb2	Positive	The VF helped me to focus on the important concepts	3.83	1.03	3.67	0.89
Qb3	Positive	The VF was in a right position	4.25	0.75	3.25	0.87
Qb4	Negative	The VF distracted me	1.75	0.75	2.42	0.79
Qb5	Negative	The VF didn't add any value to my watching experience	2.00	0.74	1.92	0.79
Qb6	Negative	I would prefer watching the video without the VF on	1.42	0.67	1.83	0.72

Figure 3: Qb results on subjective feedback after VF1 and VF2 sessions.

explained in the portion of the video watched in that session. In order to investigate if there was any effect of the *augmented transcript* (VF1) or the *map of concepts' flow* (VF2) on the *first-watch* comprehension results, we computed the average score obtained for each VF condition considering all the trials. Each question was scored in a range between 0-1. The results, shown in Figure 2 are: VF1 (AVG=0.67 SD=0.27), VF2 (AVG=0.58 SD=0.33), VF3 (AVG=0.35 SD=0.36). We used t-test to investigate if the better results of VF1 and VF2 compared to VF3 baseline were significant and we found that only VF1 results are significantly better than VF3 ($p=0.02$). We also checked for significance the difference between VF1 and VF2 and we obtained that VF1 results were not significantly better than VF2 results.

4.3.2 Qb results: subjective feedback. In order to collect the personal opinion of participants about VF1 and VF2, after each session under VF1 and VF2 conditions, participants had also to rate their agreement with 6 statement items (three positive Qb1-Qb3, and three negative Qb4-Qb6) on a 5-point Likert scale (1=disagree, 5=agree). Fig. 3 shows higher appreciation for VF1 than VF2 (Qb1-Qb3) and less disliking for VF1 than VF2 (Qb4-Qb6). Their difference is significant only for the group of positive statements ($p=0.047$). Considering the single items, we observe that only for Qb3 and Qb4 the difference is significant ($p=0.006$, $p=0.046$ respectively).

4.3.3 VFs comparison results. Since the *augmented transcript* and, mostly, the *map of concepts' flow* have not been designed to support only the first-watch of a video but also other phases and tasks of the learning process, we asked participants to provide a score on a range 0-3 to VF1 and VF2, considering also other possible uses (ref. Section 3). Results confirm that VF1 is preferred over VF2 for first-watch (AVG VF1=2.67, VF2=1.42), while for rewatch they gain the same score (AVG=2.75). Conversely, for the exploring goal, participants see potential benefits from VF2 more than VF1 (AVG VF1=2.14, VF2=2.54) and the same for in depth-analysis (AVG VF1=0.43, VF2=2.23) where the difference is significant. While the last result is unsurprising, the potentials of augmented transcript also for exploration was not in the expected results.

5 DISCUSSION AND LIMITATIONS

The overall finding from this preliminary evaluation suggest that both the *augmented transcript* (VF1) and the *map of concepts' flow* (VF2) can increase the learner immediate comprehension (first-watch learning context), helping the learner focus on the relevant concepts and their related concepts and contents. However, we observe that the increased performance, compared to using a simple video player, is statistically significant only for the *augmented transcript*. This confirms the results already available in the literature [22] about the effectiveness of combining video and transcript. In our approach, we use an augmented transcript thus providing further support for focusing the learner attention to the relevant concepts in the video lesson. This emerges also from the results to Qb questions about the subjective feedback to each visual condition. The *map of concepts' flow* performed well overall but results seem to suggest that while it can support the comprehension in the first-watch phase, it is expected to provide more support for other learning tasks that are not specifically investigated in this paper but that seem to emerge from the answers to the third group of questions about VF comparison in different learning contexts and for different goals. Participants seem to appreciate the possibility for exploration and particularly for in-depth analysis that the map of concepts' flow can provide when interacting with it.

Anyway, limiting the discussion to the objectives of this experiment, we observe that both the visual feedback methods have been designed with the goal of improving video-based learning by providing a bit of structure to video content. By increasing the immediate understanding –as shown in this experiment–, we could expect a general improvement in efficiency (less time to acquire the knowledge) of the learning process. However, further research is needed to make any statement about that.

Considering the results of Qb questions, another useful finding is about the risk of distraction that can be brought by the presence of visual feedback. This is a relevant topic in information visualization, and a risk we acknowledge, especially for the *map of concepts' flow*. This risk is confirmed by results that show a difference that is statistically significant about distraction produced by the map compared to the transcript (Qb4). This could also be due to the specific layout of the user interface that does not allow to resize or move the panels of the transcript and the map. Results to Qb3 about the position of the VF seem to support this hypothesis. This is a limit of the current user interface that we need to investigate with specific usability tests. As next steps, we plan on the one hand to make some changes to the user interface, in order to reduce the problems that the subjective feedback already highlighted, and on the other hand to perform a user study where participants are asked to use the tool for a longer period, at least one month, in order to investigate the improvements required and the real potential of the tool as a support for different learning contexts and goals.

REFERENCES

[1] Giovanni Adorni, Chiara Alzetta, Frosina Koceva, Samuele Passalacqua, and Ilaria Torre. 2019. Towards the identification of propaedeutic relations in textbooks. In *International Conference on Artificial Intelligence in Education*. Springer, 1–13.

[2] Fareedah ALSaad, Assma Boughoula, Chase Geigle, Hari Sundaram, and ChengXiang Zhai. 2018. Mining MOOC Lecture Transcripts to Construct Concept Dependency Graphs. *International Educational Data Mining Society* (2018).

[3] Amparo Bes-Piá, Blasco Tamarit Encarna, and María José Muñoz-Portero. 2011. Different applications of concept maps in Higher Education. *Journal of Industrial Engineering and Management (JIEM)* 4, 1 (2011), 81–102.

[4] Maria Birbili. 2006. Mapping Knowledge: Concept Maps in Early Childhood Education. *Early Childhood Research & Practice* 8, 2 (2006), n2.

[5] Alberto AP Cattaneo, Hans van der Meij, Carmela Aprea, Florinda Sauli, and Carmen Zahn. 2019. A model for designing hypervideo-based instructional scenarios. *Interactive learning environments* 27, 4 (2019), 508–529.

[6] Alberto AP Cattaneo, Hans van der Meij, and Florinda Sauli. 2018. An empirical test of three instructional scenarios for hypervideo use in a vocational education lesson. *Computers in the Schools* 35, 4 (2018), 249–267.

[7] Sangwoo Cho, Franck Dernoncourt, Tim Ganter, Trung Bui, Nedim Lipka, Walter Chang, Hailin Jin, Jonathan Brandt, Hassan Foroosh, and Fei Liu. 2021. StreamHover: Livestream Transcript Summarization and Annotation. *arXiv preprint arXiv:2109.05160* (2021).

[8] Stephen Ciullo, Terry S Falcomata, Kathleen Pfannenstiel, and Glenna Billingsley. 2015. Improving learning with science and social studies text using computer-based concept maps for students with disabilities. *Behavior Modification* 39, 1 (2015), 117–135.

[9] Barbara J Daley and Dario M Torre. 2010. Concept maps in medical education: an analytical literature review. *Medical education* 44, 5 (2010), 440–448.

[10] Furong Dang, Jintao Tang, and Shasha Li. 2019. MOOC-KG: a MOOC knowledge graph for cross-platform online learning resources. In *2019 IEEE 9th International Conference on Electronics Information and Emergency Communication (ICEIEC)*. IEEE, 1–8.

[11] Charles Dziuban, Charles R Graham, Patsy D Moskal, Anders Norberg, and Nicole Sicilia. 2018. Blended learning: the new normal and emerging technologies. *International journal of educational technology in Higher education* 15, 1 (2018), 1–16.

[12] Glenn W Ellis, Al Rudnitsky, and Becky Silverstein. 2004. Using concept maps to enhance understanding in engineering education. *International Journal of Engineering Education* 20, 6 (2004), 1012–1021.

[13] Marion Fesmire, Martha CP Lisner, Patricia R Forrest, and William H Evans. 2003. Concept maps: A practical solution for completing functional behavior assessments. *Education and Treatment of Children* (2003), 89–103.

[14] Olivia Foulds. 2022. Investigating How Word Clutter and Colour Impact Upon Learning. In *Orchestration of Learning Environments in the Digital World*. Springer, 135–151.

[15] Peng Gang, Jiang Hui, S Stirenko, Yu Gordienko, T Shemsedinov, O Alienin, Yu Kochura, N Gordienko, A Rojbi, JR López Benito, et al. 2018. User-driven intelligent interface on the basis of multimodal augmented reality and brain-computer interaction for people with functional disabilities. In *Future of Information and Communication Conference*. Springer, 612–631.

[16] Jonathan Gordon, Linhong Zhu, Aram Galstyan, Prem Natarajan, and Gully Burns. 2016. Modeling concept dependencies in a scientific corpus. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 866–875.

[17] Nuno Guimarães, Teresa Chambel, José Bidarra, et al. 2000. From cognitive maps to hypervideo: Supporting flexible and rich learner-centred environments. *Interactive Multimedia Electronic Journal of Computer-Enhanced Learning* 2, 2 (2000), 1–7.

[18] Tessai Hayama and Shuma Sato. 2020. Supporting online video e-learning with semi-automatic concept-map generation. In *International Conference on Human-Computer Interaction*. Springer, 64–76.

[19] Stefan Hrastinski. 2019. What do we mean by blended learning? *TechTrends* 63, 5 (2019), 564–569.

[20] Juho Kim, Philip J Guo, Carrie J Cai, Shang-Wen Li, Krzysztof Z Gajos, and Robert C Miller. 2014. Data-driven interaction techniques for improving navigation of educational videos. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*. 563–572.

[21] Elena E Kotova and Ivan A Pisarev. 2021. Expansion of the Students Educational Indicators Activity Interface in the Moodle Environment by Means of Intelligent Agents. In *2021 IV International Conference on Control in Technical Systems (CTS)*. IEEE, 158–161.

[22] Charlotte Kramer, Johannes König, Sarah Strauss, and Kai Kaspar. 2020. Classroom videos or transcripts? A quasi-experimental study to assess the effects of media-based learning on pre-service teachers' situation-specific skills of classroom management. *International Journal of Educational Research* 103 (2020), 101624.

[23] Krishna Kulkarni and Rushikesh Padaki. 2021. Video Based Transcript Summarizer for Online Courses using Natural Language Processing. In *2021 IEEE International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS)*. IEEE, 1–5.

[24] Seulki Lee, Youkyoung Park, and Wan C Yoon. 2015. Burst analysis for automatic concept map creation with a single document. *Expert systems with applications* 42, 22 (2015), 8817–8829.

[25] Chen Liang, Jianbo Ye, Shuting Wang, Bart Pursel, and C Lee Giles. 2018. Investigating active learning for concept prerequisite learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 32.

[26] Ching Liu, Juho Kim, and Hao-Chuan Wang. 2018. ConceptScape: Collaborative concept mapping for video learning. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–12.

[27] Craig Locatis, James Charuhas, and Richard Banvard. 1990. Hypervideo. *Educational Technology Research and Development* 38, 2 (1990), 41–49.

[28] Cui-Xia Ma, Yang Guo, and Hong-An Wang. 2016. VideoMap: An interactive and scalable visualization for exploring video content. *Computational Visual Media* 2, 3 (2016), 291–304.

[29] Alistair Miles and Sean Bechhofer. 2009. SKOS simple knowledge organization system reference. (2009). https://www.w3.org/TR/skos-reference/

[30] Xiaojie Niu, Jingjing Zhang, Kate M Xu, and Xuan Wang. 2021. The Impact of Productive Failure on Learning Performance and Cognitive Load: Using Hypervideo to Facilitate Online Interactions. In *2021 International Conference on Advanced Learning Technologies (ICALT)*. IEEE, 30–32.

[31] Debajyoti Pal and Syamal Patra. 2021. University students' perception of video-based learning in times of COVID-19: A TAM/TTF perspective. *International Journal of Human–Computer Interaction* 37, 10 (2021), 903–921.

[32] Liangming Pan, Chengjiang Li, Juanzi Li, and Jie Tang. 2017. Prerequisite relation learning for concepts in moocs. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 1447–1456.

[33] Marco Perini, Alberto AP Cattaneo, and Giuseppe Tacconi. 2019. Using Hypervideo to support undergraduate students' reflection on work practices: a qualitative study. *International Journal of Educational Technology in Higher Education* 16, 1 (2019), 1–16.

[34] Stefan Rabiger, Tuğberk Dalkılıç, Alperen Doğan, Buket Karakaş, Berk Türetken, and Yücel Saygın. 2020. Exploration of video e-learning content with smartphones. (2020).

[35] Alberto Regis, Pier Giorgio Albertazzi, and Ezio Roletto. 1996. Concept maps in chemistry education. *Journal of Chemical Education* 73, 11 (1996), 1084.

[36] Florinda Sauli, Alberto Cattaneo, and Hans van der Meij. 2018. Hypervideo for educational purposes: a literature review on a multifaceted technological tool. *Technology, pedagogy and education* 27, 1 (2018), 115–134.

[37] Kyoungwon Seo, Samuel Dodson, Negar M Harandi, Nathan Roberson, Sidney Fels, and Ido Roll. 2021. Active learning with online video: The impact of learning context on engagement. *Computers & Education* 165 (2021), 104132.

[38] Gorla Shree Shanmukhaa, Sruthi Keerthi Nandita, and M Vamsee Krishna Kiran. 2020. Construction of knowledge graphs for video lectures. In *2020 6th international conference on advanced computing and communication systems (ICACCS)*. IEEE, 127–131.

[39] Lijie Shao, Fuwei Zhang, Ruomei Wang, Fan Zhou, and Shujin Lin. 2018. An Efficient Expansion Word Extraction Algorithm for Educational Video. In *2018 7th International Conference on Digital Home (ICDH)*. IEEE, 131–136.

[40] Els Van Zele, Josephina Lenaerts, and Willem Wieme. 2004. Improving the usefulness of concept maps as a research tool for science education. *International Journal of Science Education* 26, 9 (2004), 1043–1064.

[41] Gaurav Verma, Trikay Nalamada, Keerti Harpavat, Pranav Goel, Aman Mishra, and Balaji Vasan Srinivasan. 2021. Non-Linear Consumption of Videos Using a Sequence of Personalized Multimodal Fragments. In *26th International Conference on Intelligent User Interfaces*. 249–259.

[42] Carmen Zahn, Beatriz Barquero, and Stephan Schwan. 2004. Learning with hyperlinked videos—design criteria and efficient strategies for using audiovisual hypermedia. *Learning and Instruction* 14, 3 (2004), 275–291.