VidWiki: Enabling the Crowd to Improve the Legibility of Online Educational Videos

Andrew Cross, Mydhili Bayyapunedi[1], Dilip Ravindran[1], Edward Cutrell, and William Thies

Microsoft Research India
{t-across, cutrell, thies}@microsoft.com

ABSTRACT

Videos are becoming an increasingly popular medium for communicating information, especially for online education. Recent efforts by organizations like Coursera, edX, Udacity and Khan Academy have produced thousands of educational videos with hundreds of millions of views in their attempt to make high quality teaching available to the masses. As a medium, videos are time-consuming to produce and cannot be easily modified after release. As a result, errors or problems with legibility are common. While text-based information platforms like Wikipedia have benefitted enormously from crowdsourced contributions for the creation and improvement of content, the various limitations of video hinder the collaborative editing and improvement of educational videos. To address this issue, we present VidWiki, an online platform that enables students to iteratively improve the presentation quality and content of educational videos. Through the platform, users can improve the legibility of handwriting, correct errors, or translate text in videos by overlaying typeset content such as text, shapes, equations, or images. We conducted a small user study in which 13 novice users annotated and revised Khan Academy videos. Our results suggest that with only a small investment of time on the part of viewers, it may be possible to make meaningful improvements in online educational videos.

Author Keywords

Online education; massive open online course; crowdsourcing; wiki; video annotation

ACM Classification Keywords

H.5.2 [**Information Interfaces and Presentation**]: User interfaces; K.3.0 [**Computers and Education**]: General

[1] Mydhili Bayyapunedi and Dilip Ravindran were interns at Microsoft Research India during this work. They can be reached at mydhilib@gmail.com and diliprr@berkeley.edu.

CSCW'14, February 15 - 19 2014, Baltimore, MD, USA

ACM 978-1-4503-2540-0/14/02…$15.00.
http://dx.doi.org/10.1145/2531602.2531670

INTRODUCTION

Today we are seeing a paradigm shift in online education. In the last 5 years, organizations such as Coursera, edX, Khan Academy, and Udacity have released thousands of hours of video that have been viewed millions of times by learners. Proponents of Massive Open Online Courses (MOOCs) envision a world where videos serve as the primary delivery vehicle for educational content distributed to learners across the planet via the Web. Further, the concepts of "flipped classrooms" and blended learning are gaining traction in K-12 and university classrooms worldwide. These teaching techniques replace primary in-person instruction with video tutorials, reserving class time for more in-depth interaction. In short, online educational videos have the potential to revolutionize education in both living rooms and traditional classrooms throughout the world. As a result, high-quality, legible, up-to-date content that is also customized for local languages and idioms is of utmost importance for educational videos.

Despite the appeal of educational video content, it remains very time-consuming to create, update, and localize videos. Professors spend enormous quantities of time creating content for their classes; this time and expense is one of the main bottlenecks to the growth of MOOCs. This situation is not unlike prior production of encyclopedias – Britannica, Encarta, etc. – where a huge compendium of information was slow and expensive to create. However, time has shown that the crowd can greatly reduce these costs while also improving quality and timeliness. While not an absolute authority, Wikipedia has been shown to be accurate [10], covering a wide range of topics [12], and is available in numerous languages worldwide. With that in mind, we ask: Can the crowd play a similar role in improving educational videos? In particular, can we leverage the vast number of students who are closely viewing educational videos, enabling them to "give back" a small amount of effort that benefits current and future classmates?

In this work, we take a small step towards this vision by creating a "wiki" for educational videos. While such a platform could evolve to allow editing many aspects of video, in this paper we restrict our attention to two things: legibility of text and graphics, and language localization.

Legibility of text and graphics remains an unrecognized challenge in many educational videos. Prior research has shown that viewers of videos enjoy seeing handwriting in real-time as a form of engagement, drawing users into the

Figure 1. TypeRighting example. Typeface fades in to replace each handwritten phrase soon after it is drawn.

lesson despite being physically separate from the speaker [5]. However, for reviewing or skimming a video for specific content, typeface presentation is preferred for its legibility and visual appeal [5]. A recently-proposed style of presentation called TypeRighting, illustrated in Figure 1, combines the benefits of handwriting and typeface by morphing each handwritten phrase into a typewritten version shortly after it appears [5]. While TypeRighting has shown strong potential – about 80% of viewers prefer to watch Khan Academy videos in this style, compared to the original – until now it has been difficult to produce videos in this format.

Language localization for the content of videos has also received little attention to date. While there are subtitling platforms to convey audio in different languages, the text in the video itself remains untranslated. Until now, the main remedy for this was to record a new video (or just leave the handwriting in the original language).

To address these problems, this paper describes the design, implementation, and usability evaluation of VidWiki: a tool that enables the crowd of students viewing educational videos to easily improve and localize those videos. While VidWiki bears some similarity to other tools for video annotation, none of these tools offer the features necessary for collaborative improvement of educational content. Using VidWiki, users can overlay text, shapes, images, equations, and hand-drawn content on top of a video to improve legibility and enable translation to multiple languages. Additionally, the annotations could be used to index textual content in the video, to convert the video to presentation slides (such as PowerPoint or similar), and potentially to save network bandwidth by loading only the audio and annotations, instead of the full video. Figure 2 illustrates an example video frame with overlain type in English and Hindi produced using VidWiki.

To explore the VidWiki user experience, we conducted a small user study asking 13 participants to annotate segments of Khan Academy videos. We measured their performance annotating videos from scratch, as well as incrementally improving on others' annotations. An expert administrator fine-tuned the annotations further to arrive at a final version of each video. On average, the annotating process required about 9.5 minutes of human effort for every 1 minute of video. When considering the size of the crowd viewing educational videos, the amount of work for a single user could be reduced drastically at scale. With just modest usage on the part of students, we believe that VidWiki could offer significant improvements to the overall viewing experience.

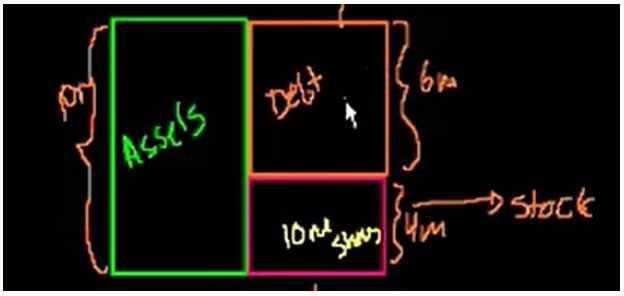


Figure 2a. Frame from original video

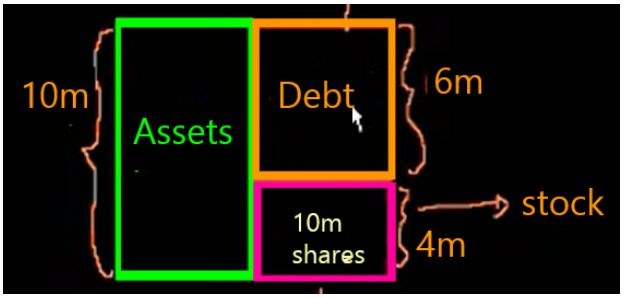


Figure 2b. Frame with annotations in English

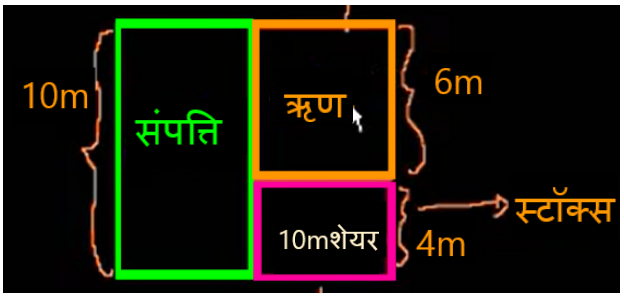


Figure 2c. Frame with annotations in Hindi

RELATED WORK

Prior tools for video annotation fall in three categories: annotating videos for indexing or object identification; adding captions; and altering videos by overlaying new content as a lens to highlight or obscure features in the video.

Several tools allow a user to create video annotations as a means to indexing videos or identifying objects in the video [16, 21, 25, 26]. These tools are all designed for marking up videos with metadata and do not have the capability to overlay text or other objects to obscure background content as part of seamless video playback. Annotations are used as additional content to build an index of objects or timings, rather than as a method to improve a video's visual quality.

Other captioning tools, such as YouTube Annotations [24] and Ambulant Captioner [18], allow users to add interactive features to videos with captions or links to other videos. However, these annotations do not fully obscure the text below, and do not allow custom shapes, equations, or images, making them inadequate to convert content to TypeRighting style.

A final group of annotation tools are designed to augment the video viewing experience by overlaying different kinds of content [11, 20, 22]. These tools are mostly domain-specific and are intended as supplemental content to draw attention to points in the video, rather than improve visual quality.

A recent release by Mozilla called PopcornMaker[1] combines several annotation concepts, allowing users to remix videos by adding additional content such as Wikipedia articles, online maps, or simple plain text. Although a powerful tool, remixes made using this platform are produced by a single user and not designed for collaborative editing. Produced videos are meant to be remixes, or collages of data, rather than an improved version of the original. The tool also lacks certain features necessary for annotating educational content such as panning events to shift annotations when an author scrolls the screen, or the ability to add equations for math or science videos.

Extensive research has examined the properties of collaborative crowdsourcing platforms, including volunteer incentives and editing behavior [1, 7], how most work is done by a minority of users [23], quantifying time spent editing or moderating an article [8], and optimizing the final quality while dealing with vandalism [2, 9, 13, 14]. This literature heavily influenced our design of VidWiki as a crowdsourcing platform.

GUIDING DESIGN PRINCIPLES

The purpose of VidWiki is fundamentally different from prior video annotation tools. While prior tools sought to *augment* videos with useful annotations, we seek to *replace* handwritten text with a typed overlay that exactly mirrors the content, style, and layout of the original video. Specifically, we aim to facilitate users' creation of TypeRighting annotations for handwritten content in videos. In the case of language translation, translated text entirely replaces the content underneath.

Design Goals

The most commonly-used tools influencing our initial design were PowerPoint animations and YouTube Annotations. However, there are several unique aspects of online educational videos that these tools are not well-equipped to address.

First, traditional caption annotations either persist throughout an entire video or frequently come and go (e.g., subtitles or YouTube Annotations). In contrast, in the majority of handwritten videos, objects such as text and drawings only disappear at two times: when the author clears the entire screen, or when they scroll the screen to reveal more blank space.

Second, because the added TypeRighting annotations aim to replace the original handwritten content, it is important for annotations to be opaque. That is, they should be filled with a background color that fully obscures the original content of the video. While this might seem like a simple feature, several existing tools support only transparent annotations. This makes them unfit for our purposes.

Third, given the vast audience of potential collaborators educational videos attract, the ability to crowdsource video annotations means that any annotation tool should support collaborative creation of content.

VidWiki is therefore a browser-based tool designed so that someone unfamiliar with video editing software can contribute. Annotations are modular, such that a video can be annotated all at once or revised in pieces at different times.

Constraints

There are several constraints that limit the capabilities of VidWiki. For example, one limitation is that it is difficult to programmatically access the pixel data of a video hosted in an outside domain, or frames within the video through a browser, and difficult for the server to process files without hosting or mirroring the videos ourselves. Because of this limitation, tools like an eyedropper tool to select color, or automatic recognition of on-screen text is difficult. Therefore, in its current incarnation, all annotations must be added manually. Future research will work toward automating steps that currently require human effort, for example, automatically generating some annotations that can later be checked, corrected or tweaked by annotators.

Since VidWiki is not a video editor, annotations are simply layers on top of the video that alter the appearance of the video without manipulating the actual pixel data. These layers are well-suited to our goals in this work, though they also have some limitations, such as at times obscuring recorded cursor or pointer movements in the video.

Other limitations affect the type of video that can be seamlessly annotated. When viewed as a tool to clean up legibility of text, VidWiki works particularly well with videos that maintain a stable background that primarily features text, like most Khan Academy videos and many other MOOC videos. However, videos that follow a speaker around or do not have a fixed background can be hard to annotate since the text location is unstable.

VIDWIKI

VidWiki allows users to draw annotations over a video canvas, altering the visual appearance of the video without modifying the video itself. A screenshot of the tool is shown in Figure 3, with various features highlighted. A video demonstration is also available online[2].

A user watching the video in a browser can toggle into the editing mode at any time. This mode reveals the extended toolset and properties shown in Figure 3. The various features are as follows:

[1] https://popcorn.webmaker.org/

[2] http://youtu.be/qtkymHwsRzo

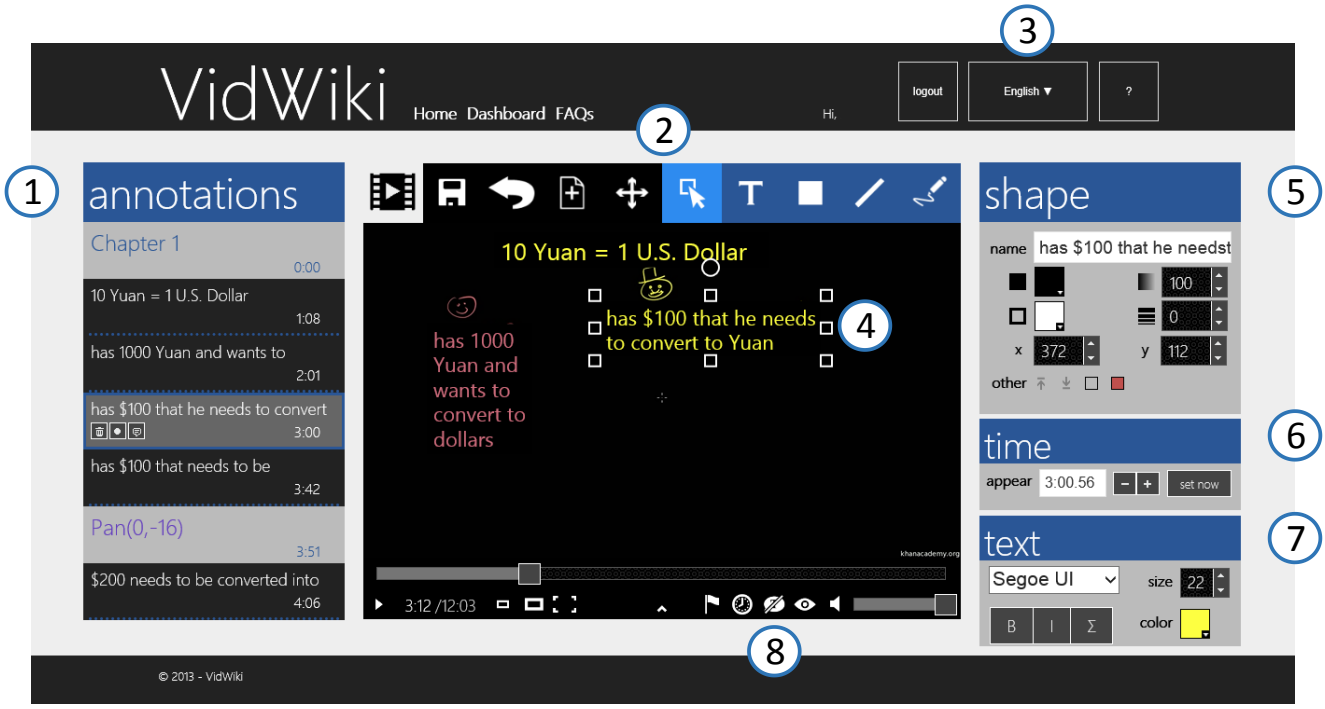


Figure 3. VidWiki and several salient features (see text for discussion).

1. An annotation list shows the current annotations and the time at which they appear.
2. The toolbar above the video allows a user to add annotations. The most commonly used tools are: a text tool, allowing entry of plaintext or LaTeX equations which are rendered over the video; a shape tool, which draws common shapes and also allows users to add images by their URL; a pen tool, which allows users to draw custom content; a chapter tool, which mimics a 'slide change' and clears all visible annotations; and a pan tool, which scrolls all visible annotations by a given pixel count.
3. A language selection menu. One usage for VidWiki is to create annotations in languages other than the original text language.
4. The annotations are drawn over the video, and are selectable to adjust sizing or position.
5. Shape properties including fill and stroke color, opacity, stroke width, and position.
6. Timing properties for when the annotation should appear. Annotations can only disappear with a chapter or pan event.
7. Text properties for color, font, size, decoration, and for rendering text as a LaTeX equation.
8. Video properties to show or hide annotations or subtitles. Users can also flag inappropriate content, or adjust the timing of all annotations relative to what the annotator specified. For example, one could introduce a delay in all the annotations, causing handwritten content to remain visible for a longer time before being obscured by typeface.

ADDING AN ANNOTATION

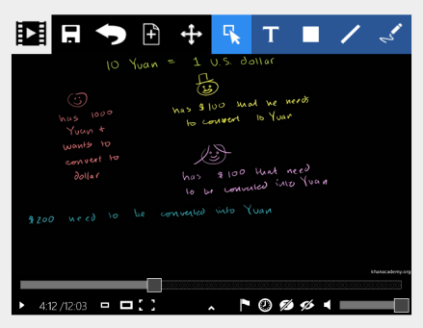
To add an annotation at a given time, a user follows the steps outlined in Figure 4. While in edit mode, the user selects the desired tool and draws the annotation directly over the video canvas. For text content, the user types the text, adjusts the font and color, and finally moves it into position over the original handwriting. The timing of the annotation can be tweaked in the timing panel, or set to the current video time.

Other annotations such as equations, images, shapes, or free-drawn content can be added in a similar way. Chapter and pan events, which clear and scroll the screen, respectively, are added by clicking on corresponding tool icons. Each annotation is a distinct unit, and any user can move, resize, delete, or adjust the timing.

Getting Annotations from the Crowd

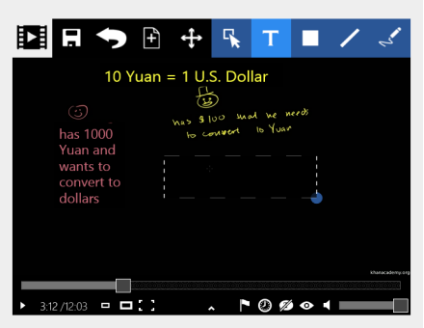
While there is high potential to automate many of the features described above (e.g., using OCR), in the current incarnation, annotations are created manually by members of the crowd. They can be added piecemeal, or all at once; they can be edited, tweaked, or changed by other users of the platform.

Any platform that draws participation from the crowd faces challenges of merging conflicting annotations or vandalism. Many of these issues have been sorted out by projects such as Wikipedia. For now, VidWiki allows users to flag inappropriate content for a moderator to consider and take action. The history of revisions is available to system administrators, but is not yet exposed to users.



1. Original scene from a Khan Academy video before adding any annotations.

2. Video paused when user wants to add an annotation.



3. Using the text tool, user draws annotation box over the video at the given time.

4. User types the text for the annotation.

5. User adjusts font to match the video and moves annotation into place over the text.

6. Scene after user has added multiple annotations over the text.

Figure 4. Adding an annotation in VidWiki.

Translation of On-screen Text

Many MOOCs aspire to reach a global audience, including viewers who do not speak English as their primary language. Given the amount of time and effort required to create high-quality videos, it would be highly desirable to translate content from one language to another without having to re-record it from scratch. While tools such as Amara[3] enable subtitling of videos, they focus only on the audio track and do not translate the text content in the video itself.

VidWiki extends the capabilities of video translators by using annotations to display translations of on-screen text. The tool currently supports user-generated translations, where the author indicates the language in which text is typed, and the viewer can select among all of the available languages. In the future, machine translation can be used to produce the first pass of translated annotations, and a human can later make adjustments where needed. This model has previously been explored in the context of Wikipedia [17] but has not been applied to video content until now.

USER STUDY

Study Goals

There were two primary goals to our user study. The first was to evaluate the usability of VidWiki by novice users. As VidWiki introduces several novel concepts in video annotation, we sought to understand users' first time experiences with the tool.

The second goal was to explore the dynamics of annotating a video as a member of the crowd, including review and improvement of other users' annotations. In short, we wanted to explore the quality of annotations produced by novice users, how much time it took to annotate a set of videos, and how users managed this process in a collaborative setting.

The user study did not explore the translation aspect of VidWiki because we believe the authoring experience would be similar, provided the correct keyboard tools are available to type in a given language. Also, this study is not a substitute for a large-scale field trial of VidWiki, which (as discussed later) will be an important focus for future work.

[3] http://www.amara.org/

Participants

We conducted a usability evaluation of VidWiki with 13 students. All students were pursuing undergraduate or graduate studies in India, and 12 of the 13 specialized in computer science or a related technical field. None of them had ever used the VidWiki tool before. These students are typical of many viewers of online educational videos.

Methodology

Before initiating the task, each participant was given a 5-minute tutorial on using VidWiki. This in-person tutorial introduced the notion of annotating a video in the TypeRighting style, and demonstrated the various tools available for annotation in VidWiki. This demonstration used a previously annotated video different from those participants would work with later. We suggested that participants use large text when possible, keep the color scheme as similar to the video as possible, and have the annotations appear roughly three seconds after the author had finished handwriting a word, phrase, or number.

Each participant was given two separate Khan Academy videos to convert from handwritten style to TypeRighting style. The videos came from two economics playlists, and one statistics playlist from the Khan Academy website; they covered the topics of financial and capital markets, currency exchange, and an introduction to probability and statistics. The first video given to each participant was unannotated, and the participant's task was to annotate the first four minutes of the video from scratch. After completing the first task, each participant filled out a short survey on their experience using the tool and was directed to the second video.

The second video had been previously annotated by another participant, and the participant's task was to review the existing annotations and to fix or adjust where the annotations were inadequate with regards to content, color, position, or timing. If the current annotations were perfect, users could re-submit them without any changes. To bootstrap this process, the first participant revised a video prepared by the researchers. Following this, participants filled out a second survey on their experience to complete the trial.

We asked participants to annotate only the first four minutes of each video to keep the experiment short and because the videos varied in total duration. Therefore, 13 study participants annotated 52 minutes of video content, and revised another 52 minutes of video content. VidWiki was instrumented to track user actions and timing throughout the study. Though the trials were unassisted, a researcher was available nearby if a participant had any questions. Early questions were due to bugs in the initial system, or were about how to handle certain situations in videos such as when graphics and text were very close, requiring the text annotation to obscure the graphics as well. We extended our tutorial to cover these common issues and the amount of assistance towards the end of the experiment was minimal.

All timing data and survey results were anonymous and analyzed in aggregate. Both surveys used 5-point Likert-scale questions to assess the usability of VidWiki, the perceived quality of annotations, and the participants' willingness to annotate videos based on the final output and potential audience benefitting from their work. Participants were also given two free-response questions asking about the positive and negative aspects of the tool.

Results

To estimate the feasibility of using VidWiki to crowdsource video annotations, we wanted to understand both the usability of the system as well as how long it takes users to annotate videos. The participants annotated 52 minutes of video in a cumulative 321 minutes of work (M = 24.7, SD = 7.14 per four-minute video); participants revised 52 minutes of video in 121 minutes of work (M = 9.28, SD = 3.39 per video). Therefore, in total, it took a combined 8.5 minutes of novice work to annotate and revise each minute of the original video. Variability in annotation time was a product of both the participant's speed and the complexity of the video, as some videos had more text to annotate than others.

While our study aimed to examine usage by novices, there is obvious potential for users to become more efficient as they gain experience with the tool. As an example measure of how much efficiency could improve, a person familiar with the tool annotated each of the 13 videos from scratch, annotating 52 minutes of video in 101 minutes (M = 8.38, SD = 3.68 per video) corresponding to 1.93 minutes of annotation work for every minute of video.

Post-task surveys suggest that participants felt that 1) the video quality improved after each round of annotation and revision; 2) the tool was usable (also including many suggestions for further improvements); and 3) that participants felt the amount of time spent was reasonable if a large audience would benefit from their work. However, as participants were answering questions about the experimenter's technology, we acknowledge that the results may be biased by demand characteristics [6]. Summaries of survey responses on a 5-point Likert-scale (1 – Strongly Disagree, 5 – Strongly Agree) are shown in Figures 5 and 6.

Quality of Annotations

To assess the quality of participants' annotations, the first author served as a moderator, taking the revised annotations and fixing or improving them if necessary. Like moderators in online forums, this individual was familiar with the tool as well as the expected style and quality of the final annotations. It took the moderator 54.4 minutes to revise the existing 52 minutes of annotations (M = 4.18, SD = 1.46 per video). This amounts to only slightly more time than watching the videos to ensure correctness, a reasonable expectation of moderators on this and related platforms.

Adding this moderation time to the annotation and revision phases, it took a total of 496 minutes for novice workers to annotate, and a more experienced user to moderate the 52

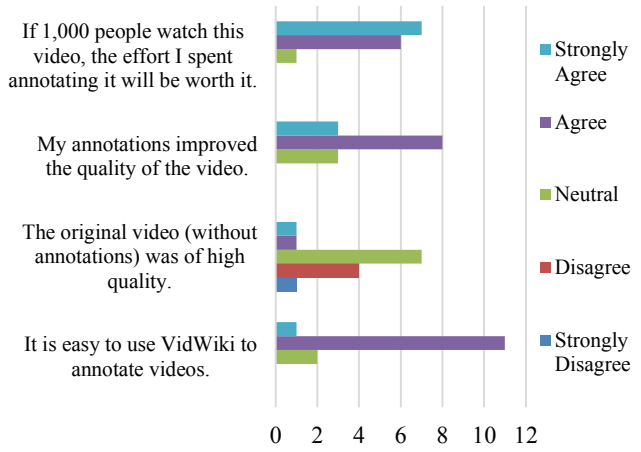


Figure 5. Results of survey after annotating a video.

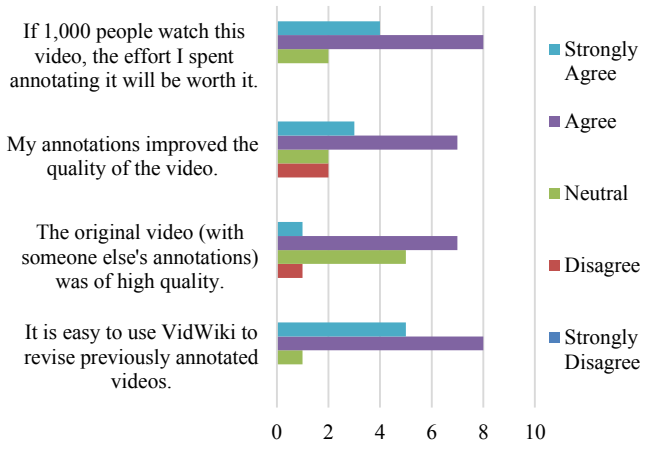


Figure 6. Results of survey after revising a video.

Scene from original video | First user's annotations | Second user's revisions of first user's annotations | Final version after moderator's revisions

Figure 7. Iterative improvement of annotations during user study.

minutes of video; this corresponds to 9.5 minutes of combined novice and moderator work to annotate each minute of the original video.

Figure 7 shows a sample progression of a frame from the original video to the annotated final product. Figure 8 shows three additional videos before and after annotation, demonstrating the improvement in visual quality.

Common struggles, or situations that took the most time for users, involved perfecting the timing, location, or color of an annotation. Many of the tweaks in revisions were slight adjustments to these three annotation properties. Two videos included panning in the first four minutes which required more effort than adding text annotations, but participants were able to correctly add pan events.

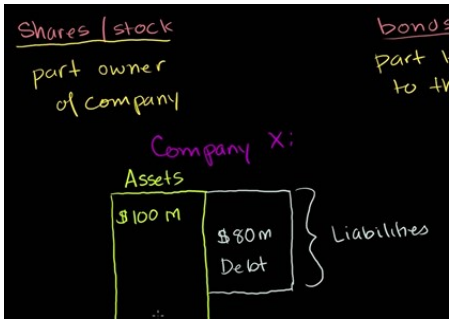
After the annotations were revised, the videos had an average of 17 total annotations per video. Revisers changed the text content in 7% of annotations. Most changes were minor such as altering the text spacing by adding new lines or capitalizing words. Annotation timing was changed in 19% of annotations, where the average time adjustment was 6.9 seconds. Annotation text or shape color was changed in 9% of annotations. Some original annotators had not bothered to match color schemes at all, in which case the revisions added color. In other cases, revisers adjusted the colors to better match the background video.

DISCUSSION

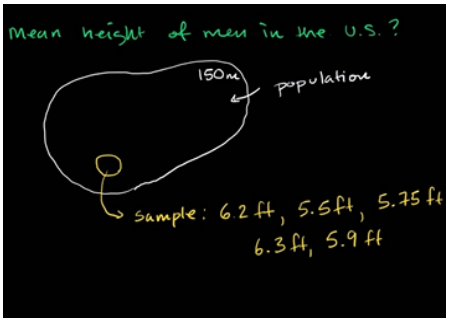
In this study, novice users were able to use the VidWiki tool to successfully annotate and revise 52 minutes of content. Though it may seem expensive to invest 9.5 minutes of work for every minute of video and to involve three different users in the process, dividing the annotation tasks across thousands of students in a common MOOC context may be quite reasonable.

As an example, consider edX's inaugural course (MIT 6.002x), that enrolled almost 155,000 students. Even with the high attrition rates typical to MOOCs, over 7,000 people passed the course [19]. These students are highly motivated and invested, and, similar to contributors to Wikipedia, they represent the top percentage and those most likely to contribute video annotations. The 14-week course contained roughly 28 hours of lecture material, requiring 15,960 minutes of work to completely annotate given our experience in the user study. If each of those motivated 7,000 students invested about 2.3 minutes – either as volunteer work, an assignment, or even paid work – they could annotate the entire course. Since most MOOC video lectures are intended to be reused year after year, the cumulative effort of several consecutive classes of students could author the annotations and continuously revise them as needed. Khan Academy videos, while targeted at a younger, less-experienced audience, attract equally high numbers of viewers and potential annotators.

Video 1: Original

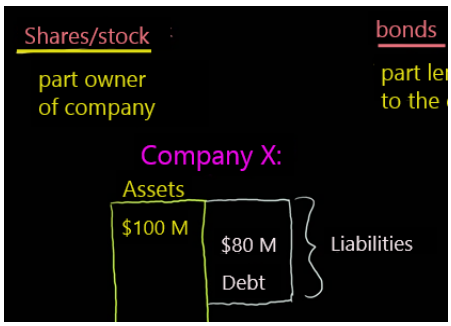


Video 2: Original

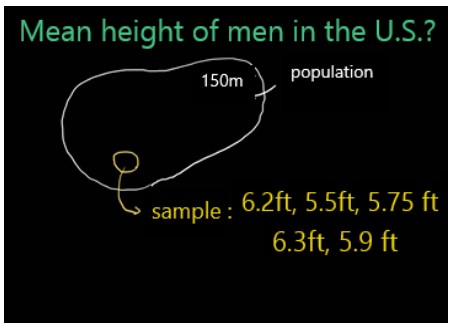


Video 3: Original

Video 1: Final with annotations



Video 2: Final with annotations



Video 3: Final with annotations

Figure 8. Frames from original and final versions of videos during user study.

A study on Wikipedia contributors estimates that the average session length, the time a user spends editing an article, is 10 minutes with some more expert users spending hours or entire days editing articles [8]. While annotating Wikipedia articles is predominately text editing, and annotating videos introduces other facets such as visual quality and timing, such findings suggest that it is feasible to ask each viewer to contribute 2.3 minutes of time to enhance an entire course.

As another point of comparison, a wiki on Coursera discusses the time required for contributors to add subtitles for the audio track and sync them with the video [5]. It suggests that it is not unreasonable to expect non-professionals to spend 15 minutes subtitling a single minute of video, which is about 50% longer than the time required for VidWiki annotations.

As an initial study into the usability of VidWiki, this paper does not observe patterns of real-world use, which will only be evident upon releasing the tool to a larger audience. A future deployment at scale may reveal interesting new facets of collaborative annotating that could promote or hinder VidWiki's success.

Motivation to Contribute

There are open questions regarding students' motivation and willingness to contribute annotations, either voluntarily or in response to incentives. The survey responses show that while annotating videos takes time, students are willing to contribute knowing that their work can benefit a large number of peers. Augmenting this intrinsic motivation with the right incentive structure could be key for the future success of the platform.

There are several models that might prove fruitful, such as assigning annotation work as homework, leveraging a teaching assistant (or similar role) to help contribute, or paying workers to annotate. Any of these schemes could tap into the large audience already watching educational videos; this is the population most familiar with the content and, arguably, most motivated to want to improve it. In contrast to more complicated video editing tasks that require experts, VidWiki allows the annotation process to be broken up into smaller tasks appropriate for volunteers with little training. Khan Academy has an integrated "points" system whereby learners earn credit for watching videos or answering questions; annotating could be another way to earn points. More formal MOOC classes have certificates based on credit, so annotating could even become a requirement for receiving credit.

Subsets of users do contribute to forums and Wikipedia articles without any direct benefits such as course credit or financial incentives. Prior research has explored how to maximize the likelihood that a reader or viewer of crowdsourced content might contribute [15]; a similar exploration for VidWiki could evaluate related incentives.

CONCLUSIONS AND FUTURE WORK

We present VidWiki, an online browser-based tool to facilitate on-screen annotation of educational videos to improve their visual quality and localize their language. The tool allows novice users to add layers on top of a video, overlaying handwritten content with more legible typeface, or extending a video by adding other annotations like equations, shapes, images, or free-drawn custom content.

We tested the tool with 13 users unfamiliar with the system who annotated several Khan Academy videos, requiring 9.5 minutes of work for each minute of the video to be annotated. Extending this to the online scenario where Khan Academy videos have over 200,000 viewers, and MOOC courses have anywhere from 10,000 to 160,000 students on each iteration of the course, such a tool could enable large crowds of students to collaboratively annotate and improve lecture videos through various incentive schemes.

In the future, one could extend VidWiki to include audio recording and real-time capture of typed or penned annotations, so that an author could record entire lectures through the browser. This would further lower the expertise required to author an instructional video. Once a video or course is authored in the tool, the annotations are already inserted as modular objects, ready for translation, resizing, rearrangement, recoloring, etc., by authors or collaborators who seek to modify, update, or extend the video.

ACKNOWLEDGMENTS

We are grateful to Jessica Herron for her contributions to the design and user experience of VidWiki, and to Steven Drucker for his support for the project. We thank Nakull Gupta for preparing VidWiki for a broader release, and Vidya Natampally and the Microsoft Research Connections team for their ongoing support. We also thank the participants and the anonymous reviewers for their thoughtful feedback.

REFERENCES

1. Antin, J., Cheshire, C., & Nov, O. Technology-mediated contributions: editing behaviors among new Wikipedians. *CSCW,* (2012).
2. Arazy, O., & Nov, O. (2010). Determinants of wikipedia quality: the roles of global and local contribution inequality. *CSCW,* (2010).
3. Ayache, S., Quénot, G., & Tseng, A. The ligvid system for video retrieval and concept annotation. *CBMI,* (2010).
4. Coursera: Video Subtitles, https://share.coursera.org/wiki/index.php/Video_Subtitles
5. Cross, A., Bayyapunedi, M., Cutrell, E., Agarwal, A., & Thies, W. TypeRighting : Combining the Benefits of Handwriting and Typeface in Online Educational Videos. *CHI,* (2013).
6. Dell, N., Vaidyanathan, V., Medhi, I., Cutrell, E., and Thies, W., "Yours is Better!" Participant response bias in HCI. *CHI*, (2012).
7. Dow, S., Kulkarni, A., Klemmer, S., & Hartmann, B. Shepherding the crowd yields better work. *CSCW,* (2012).
8. Geiger, R., & Halfaker, A. Using edit sessions to measure participation in Wikipedia. *CSCW,* (2013).
9. Geiger, R., & Ribes, D. The work of sustaining order in Wikipedia: the banning of a vandal. *CSCW,* (2010).
10. Giles, G. Internet encyclopaedias go head to head. Nature, 438 (2005), 900-901.
11. Goldman, D., & Gonterman, C. Video object annotation, navigation, and composition. *UIST,* (2008).
12. Halavais, A., & Lackaff, D. An analysis of topical coverage of Wikipedia. *Journal of Computer-Mediated Communication*, (2008).
13. Halfaker, A., Keyes, O., & Taraborelli, D. Making peripheral participation legitimate: reader engagement experiments in Wikipedia. *CSCW,* (2013).
14. Hansen, D., Schone, P., & Corey, D. (2013). Quality control mechanisms for crowdsourcing: peer review, arbitration, & expertise at family search indexing. *CSCW,* (2013).
15. Hoffmann, R., Amershi, S., Patel, K., Wu, F., Fogarty, J., & Weld, D. S. (2009). Amplifying community content creation with mixed initiative information extraction. *CHI,* (2009).
16. Kipp, M. (2008). Spatiotemporal coding in anvil. *LREC,* (2008).
17. Kumaran, A., Saravanan, K., & Maurice, S. (2008). wikiBABEL: community creation of multilingual data. *WikiSym,* (2008).
18. Laiola Guimarães, R., Cesar, P., & Bulterman, D. Creating and sharing personalized time-based annotations of videos on the web. *DocEng,* (2010).
19. Lessons Learned from MITx's prototype course, (2012). http://web.mit.edu/press/2012/mitx-edx-first-course-recap.html.
20. Lux, M., & Riegler, M. (2013). Annotation of endoscopic videos on mobile devices: a bottom-up approach. *MMSys,* (2013).
21. Müller, S., Miller, G., & Fels, S. Using temporal video annotation as a navigational aid for video browsing. *UIST,* (2010).
22. Singh, V., Latulipe, C., Carroll, E., & Lottridge, D. The choreographer's notebook: a video annotation system for dancers and choreographers. *ACM C&C,* (2011).
23. Stuckman, J., & Purtilo, J. (2009). Measuring the wikisphere. *WikiSym,* (2009).
24. Yuen, J., & Russell, B. Labelme video: Building a video database with human annotations. *IEEE Conference on Computer Vision*, (2009).
25. YouTube Annotations, http://www.youtube.com/t/annotations_about
26. Zimmermann, T., & Weber, M. CoVidA: pen-based collaborative video annotation. *VIGTA,* (2012).