Graphic Novel Subtitling

Exploring the translation of videos into graphic novels through a web-based application

Amy Gourlay
University of Dundee
Dundee, UK

Supervised by Micheal Crabb

ABSTRACT

For some people, the language and cognitive demands of watching subtitled content is challenging. Common issues surrounding traditional subtitling are explored in this paper. These include but are not limited to; subtitle speed [39], subtitle readability [29], subtitle position [12] and the constant battle for attention between moving subtitles on a moving picture [21]. Related research has suggested that current subtitling standards do not always meet the needs of the individual [17]. This paper explores the concept of Graphic Novel Subtitling, which is the conversion of a video into a captioned graphic novel to help with the language and cognitive demands of watching subtitled content.

Feedback on this subtitling style compared to traditional subtitling has been gathered in the form of a questionnaire. 76% of participants gave positive or neutral feed-back compared to 17% of participants who had a negative response to the medium. Written feedback included responses relating to the readability of the medium, stating it was clear and easy to understand. A user-focused, web-based application has been proposed for the conversion of videos into graphic novels. Specific features of this application will be updated as the project progresses as the complexity of some features has not yet been determined. This application will be developed using key-frame identification and subtitle segmentation according to timestamps. There will be a focus on user-oriented design, will the goal of a self-explanatory interface that is easy to use.

A questionnaire has been carried out to support the background research in this area. Results indicated that although participants were more familiar with traditional subtitling methods, there are also a number of issues present. These include pace, font size, background colour, obstruction of the video, and the fact that reading subtitles can distract from the scene. When exposed to graphic novel subtitling, written feedback included responses relating to the readability of the medium, stating it was clear and easy to understand. Enjoyment was a also a prominent factor in the feedback with participants mentioning that it was a fun new opportunity for media consumption. There was a divide in the opinions on immersion into the material, with some participants stating that it takes away from the atmosphere and pacing of the scene in comparison to others who thought they were able to get more emotion from the graphic novel compared to regular subtitling.

From this research an outline for the proposed application has been created: A web-based application that supports the conversion of videos into graphic novels to allow for the enjoyment and engagement of media content for users. This will make use of file loading, key-frame detection and subtitle segmentation through the use of timestamps. Further features will be explored as the project progresses and the scope becomes more clear.

The project will follow a combination of waterfall and agile methodologies. The waterfall method was used for the research and planning phases of this project. An agile approach will be adopted for the design and implementation phases of the project.

1 INTRODUCTION

Approximately 10% of the television audience use subtitles (captions) to support their viewing experience. Subtitles enable viewers to participate in an experience that is often taken for granted by the general audience [11]. Subtitles are generally displayed at the bottom of the video media being presented and little research has examined alternatives to this. The purpose of this project is to investigate an alternative presentation of subtitled video content – Graphic Novel Subtitles.

Previous research involving subtitles has highlighted that current standards do not always suit the needs of the individual, leading to a poor viewing experience. In this work, there is a focus on improving comprehension of and immersion into subtitled media. Current subtitling involves moving text on a moving background which can often be tiresome and distracting for viewers [21]. In contrast to this, graphic novel subtitling can enhance immersion in the content as it allows people to have more freedom over the length of time they have to process it.

2 RELATED WORK

In order to support the development of a graphic novel subtitling application three different research areas have been explored: subtitling, video summarisation and comic creation. Automatic generation of a graphic novel from video pulls from each of these areas.

Subtitles are a text version of the speech and non-speech audio information needed to understand media content [43]. This text is usually displayed at the bottom of media content. Recent research in this area has highlighted the demand for these captions to be presented in a different manner, with a focus on improving viewing experience for the user [17].

Graphic novel subtitling also relates to the summarisation of videos using key-frame identification to create a continuous pictorial narrative [15]. The combination of subtitling and key-frame detection leads to automated comic creation, which has also been explored further in other contexts.

2.1 Personalisation of subtitling

The presentation of subtitles within media follows best-practice guidance that ensures accessibility and familiarity across different platforms [43] [28]. The BBC published a Subtitle Guidelines document in 2018 that describes the best practice for authoring subtitles [10]. It focuses on a number of areas such as rate of speech, text reduction, synchronisation and position of subtitles.

This standard of subtitles does not always suit every individual's needs. A common issue associated with subtitling standards are subtitles being too slow or too fast for different viewers which can hinder the comprehension of the content [37] [36]. One study, carried out in 2008, explored the effect of caption rate and text reduction on comprehension of captions by people who are deaf or hard of hearing [3]. The BBC currently recommends a subtitle speed of between 160-180 words-per-minute [10]. The 2008 study tested caption rates of 130, 180 and 230 WPM. Results showed a high discrepancy between users comprehension at each level and identified reading proficiency as a dominant factor. The research recommended that users would benefit from an option to select caption rates and text reduction methods that suit the majority of the viewers under the majority of circumstances or, to provide individual tailoring of caption delivery.

Comprehension of the material is an important factor in subtitling standards. Eye-tracking technologies have allowed researchers to study, with reasonable accuracy, where a person is looking while viewing a subtitles video [32]. This can show the competition between different sources of information (i.e. non-static subtitles on a moving background) [21]. Dynamic subtitles have been proposed as a potential solution where varying positions are used for subtitles, according to underlying video content (i.e. position of the speaker) [2].

There is significant diversity in the needs and preferences of frequent subtitle users which points to the need for personalisation in the way subtitles are displayed [1] [39]. One mainstream method of subtitle customisation is the user-controlled resizing of subtitle text and subtitle viewing area [29]. Another method is adaptive subtitling which takes an object-based approach to personalisation of subtitles that creates opportunities for adaptation surrounding words (e.g., names, locations), phrases (eg.catchphrases), speakers/characters, accents, audio-descriptive elements, and scene composition [17].

2.2 Video Summarisation and Key Frame Detection

Video Summarisation is often used to enable easier browsing and navigation of videos [24]. The aim is to reduce the amount of data that must be examined in order to retrieve key information from a video file [15]. Video summarisation is commonly used in large video databases (e.g. YouTube, Netflix, Vimeo) and so automatic indexing, search and retrieval of relevant material is crucial [4].

There are three main methods used to support video summarisation. Techniques used include keyframe, transcript, and metadata based techniques [30]. Keyframes can be used to create a series of representative images that relate to certain points in a video and allow for more efficient browsing or scrubbing [6]. Transcript-based summarisation involves segmenting videos into sections and providing short text summaries and thumbnails for each section [31]. This method is commonly used alongside video lectures where clear sections can be identified. Metadata for frames can be analysed to determine metadata-based prioritization data for a video that can then be used to automatically generate a textual summary for that video [25].

Keyframe detection is the most widely used method of video summarisation. Common algorithms used for this process take into account motion, shot length and colour properties of shots [35]. In the context of summarisation, a series of key frames can be extracted to represent key events while preserving the overall contents of the video [15]. This is achieved using key frame detection that identifies sequences of keyframes that are maximally distinct and individually carry the most information [41]. Keyframe detection in video summarisation is currently used for efficient visual browsing and navigating of a video [41] [16] and dynamic timeline control [33].

Novel methods of using keyframe summarisation have included a focus on the creation of comic-style artifacts [14]. By combining these keyframes with a subtitling file or script, comic books or graphic novels can be created as a storybook representation of a film or tv series [20]. CORVIS is an image to comic generator that takes images and text and adds comic book elements to them in a semi-automated manner (i.e. motion lines, speech bubbles) [20]. This system could be combined with automated video keyframe detection for a more efficient conversion process.

2.3 Automated Comic Creation

Comic books are a graphic art form combining text and images to effectively represent a story [40]. They originated during the end of the nineteenth centuary in the USA as humorous panel drawings in newspapers [44]. Today, comics are published in almost every country in the world, with each region having its own comics industry that each have different stylistic approaches to comic creation [44]. The development of technology in recent years has prompted the importance of digitalisation of comics and graphic novels to offer a wider outreach for the medium [19].

Comic narrative is constructed specifically in the comic book medium [23]. A common approach is the four-point tetrahedral model, the four points in question are the 'medium', the 'reader', the 'verbal narrative', and the 'visual narrative' [26]. There are many different elements involved in the creation of comics, the most prominent being; panel layout [5] and speech balloon placement [13]. Panels must be kept in sequential order from top left to bottom right and can be a variety of different shapes. Due to these irregular shapes, images must be cropped appropriately considering important regions such as faces and prominent objects [27]. Speech balloon contents, type and placement are also a pivotal element of comic creation.

It is important that the dialogue associated with each panel matches the image contained there. This is achieved by considering the number of words and emotion embedded in subtitles, audio energy corresponding to the targeted frame, varying speech balloons sizes so appropriate shapes can then be generated [7]. Research into the integration of text and image has shown how annotations can

contribute to the expression and flow of a comic. Different types of speech balloons (i.e. smooth edges, jagged edges) can convey different emotions [9].

Comic creation is a time consuming art form and prior work has investigated the automation of this process. As previously mentioned, key frames can be automatically extracted and placed in a comic-style layout. Researchers have also looked into the automatic placement of speech balloons by considering the position of the speaker while maintaining the reading order of dialogues by applying a cartographical labelling algorithm [8]. Face-detection algorithms have been explored to allow for automatic placement of speech balloons that do not obscure characters faces [34]. This is accomplished using standard image processing which can also assist with ensuring speech balloons do not overlap each other.

Web-based comic creation has been explored to allow for user creation and customisation of comics. Creation of comics or graphic novels is often difficult for beginners to produce to a high quality as it requires a lot of experience and domain knowledge [5]. Comic building applications [40] [22] [38] have been developed to allow users to create their own comics by using artwork supplied by the developer.

2.4 Research Questions and / or Aims

This literature review has identified previous research in personalised subtitling models for live video that focuses on the experience of the user. It has also explored video to comic conversion as an industry tool, with a focus on automation of the laborious nature of comic creation. A gap has been identified in the combination of these areas: a user-centred approach to conversion of video to graphic novel as an access service to help with the language and cognitive demands of watching subtitled content. This static subtitling method may allow for a less demanding viewing experience as users will have more freedom over the length of time they can spend immersed in the material.

This justifies the creation of an easy-to-use web-based application for automated conversion of video to graphic novel. Two research questions have been identified:

RQ1 What specific challenges exist in automatic creation of graphic novels from subtitled media, and what methods can be used to overcome these?

RQ2 How is user experience affected by graphic novel subtitling compared to traditional subtitling?

3 STUDY 1: QUESTIONNAIRE ON COMPARISON OF SUBTITLING METHODS

3.1 Design

The questionnaire included 9 questions across three sections. The first section gathered information on participants background with usage of subtitles and graphic novels. It focused on how regularly the participants used both of these mediums and why they used them. The second section asked questions relating to a subtitled video, focusing on participants thoughts about the subtitling technique and also if they faced any issues with it. The third section presents participants with a graphic novel based on the previous subtitled video. They are asked for their thoughts on this medium and what could make it better.

3.2 Method

Ethical approval for the questionnaire was obtained from the University of Dundee ethics board. The questionnaire was created in Jisc Online Surveys and distributed on social media (e.g. Facebook) and Reddit (r/samplesize). Participants were entirely anonymous, participant numbers were assigned based on their response number (i.e. in order of completion).

3.3 Results

Participants were asked how regularly they use subtitles for reasons other than language translation: Regularly(11), Sometimes(3), Rarely(11), Never(9). Participants who responded with 'Regularly', 'Sometimes' or 'Rarely' were then asked why they use subtitles in an open-ended question. Four themes were identified within the responses:

1) Too loud an environment: 8 participants mentioned background noise in their response, for example *"Oftentimes I watch TV while cooking, which can be somewhat loud so I may not be able to properly hear the show"* (P3) and *"If other people are talking and I can't hear."* (P7).

2) Helps to understand the content: There were 9 participants who described reasons relating to a better understanding of the content stating it was *"Easier for concentration"* (P6) and *"To follow dialogue/storyline closer"* (P18).

3) Poor audio quality: 6 participants mentioned the quality of the audio as their reason for using subtitles, this category also includes participants who required volume to be low. *"To clarify what is being said if the audio is unclear."* (P10), *"Only required where verbal language is not clear or unable to have sound while watching"* (P17).

4)Trouble understanding accents (3): 3 participants described issues with understanding some accents or dialects stating *"(When) someone says something in a dialect of English that I am unfamiliar with"* (P4).

Participants were also asked how regularly they read graphic novels or comics: Regularly(2), Sometimes(5), Rarely(8), Never(19). Participants who responded with 'Regularly', 'Sometimes' or 'Rarely' were then asked why they read graphic novels or comics in an open-ended question. Three themes were identified within the responses:

1) Entertainment purposes: 7 participants gave responses relating to enjoyment of the medium, stating simply *"I like them"* (P2), *"It's a fun storytelling medium"* (P3) and *"Mostly for my own entertainment."* (P4).

2) For the artwork: 3 participants mention the visual components of a graphic novel or comic: *"I love the artwork."* (P8), *"I like the artwork and it feels more immersive"* (P18) and simply *"Cool aesthetic"* (P34).

3) Continuous narrative: 2 participants gave reasons relating to the continuous story lines present in some comics: *"Sometimes to keep up to date with the DC and Marvel universe."* (P7) and that comics can be *"Further along in the story/ more content than the adaptations."* (P21).

3.4 Subtitling Feedback

The participants were first shown a 40 second clip from the British sci-fi TV series "Doctor Who" (Series 7, Episode 4) without subtitles. This clip was then shown again with no sound and traditional subtitling. Participants were asked what they thought about this style of subtitling in an open-ended question.

1) Positive responses: 15 participants responded that this was a good style of subtitling, it should be noted, however, that 12 out of the 15 participants also reported at least one issue with this style of subtitling. Responses included *"It seems accurate"* (P2), *"The subtitles are nice and easy to understand."* (P4), *"It helped with the delivery and momentum of the story"* (P29). There were also 2 responses that stated the subtitles had allowed them to pick up on parts they had missed in the original video: *"I now know what he says at the start-was quiet and rushed so I didn't catch it without subtitles."* (P5), *"I felt I understood everything said in the version with subtitles there were moments during the video without I couldn't tell you what was said as it was muttered or said too quietly"* (P14).

2) Negative responses: 10 participants gave responses with negative connotations. Most responses involved issues that are covered in the next section some examples include: *"Not clear who is speaking at times"* (P17), *"A bit fast at times"* (P23) and *"Not effective, loses impact from the visuals."* (P28).

3) Indifferent responses: 7 participants were indifferent to this technique either stating facts instead of opinions or mentioning it was "fine" or "okay" whilst listing what they also disliked about it. *"It's OK. It gets the words across but doesn't convey the emotion as well as sound."* (15), *"Same as subtitles I've read for foreign films."* (P20).

Participants were then asked to report issues with this style of subtitling, selecting all options that applied. Findings of this question are summarised in Table 1.

Table 1: Issues participants found with traditional subtitling

Issue	Count
Reading subtitles can distract from the scene	21
Subtitles can be too fast	13
Font size can be too small	12
Font or background colour can make subtitles hard to read	11

Participants had the option to describe any other issues they had with this subtitling method. Some responses repeated the issues mentioned above and so were not included.

1) Speaker Identification: 5 participants mentioned that they were unsure what character was speaking at certain points. *"It wasn't always clear when a new character was speaking"* (P23), *"Sometimes it's difficult to tell who's saying what"* (P33).

2) Subtitle Position: 2 participants described issues with where the subtitles were placed stating *"They block what's happening"* (P24), *"They can be under the pause bar etc, meaning you have to wait for the bar to disappear before you see the subtitles."* (P32).

3.5 Graphic Novel Subtitling Feedback

Participants were then shown the video in the form of a graphic novel. The graphic novel was created by taking key-frames from certain points in the video taking into account shots, motion within shots and dialogue per shot. The dialogue was added in speech balloons with tails leading towards the speaking character. The result was a 2 page graphic novel consisting of 16 panels.

Participants were asked what they thought about presenting the video as a graphic novel in an open-ended question. 16 participants gave positive responses, 10 participants were indifferent to the medium and 6 participants disliked the graphic novel format. The opinions in these responses contained 3 themes:

1) Clear and easy to understand: 12 participants gave answers that related to an easy understanding, including those who weren't that fond of the medium: *"I don't find it as interesting, however it is very easy to follow and read."* (P9), *"Easier to understand but not as enjoyable"* (P23). 5 participants also noted that it was much easier to see who was speaking which was an issue raised in the traditional subtitling section. *"Surprisingly fun to read. Conveys the emotion much better and lets me go through at my own pace. It also shows who's saying what, which the previous subtitles didn't make clear in hindsight."* (P15), *"It makes it clearer who is speaking at a time and the timing of speech."* (P16), *"It is clear, accessible and it is obvious who is speaking. There is also clear emphasis on certain words."* (P18).

2) Enjoyment: 6 participants specifically mentioned that they enjoyed this medium stating *"I think it looks fun"* (P2), *"It is cool, able to get across the context and tone of the scene"* (P7).

3) Less immersive: 6 participants thought that the graphic novel style takes away from the atmosphere and pacing of the scene. *"does not allow characters to come to life therefore loses context"* (P28), *"Takes away the atmosphere/differing moods and attitudes. Like reading a text message, no context of what expressions are meant to be."* (P30). 4 participants contradict these statements saying they got more emotion from the graphic novel than the regular subtitles on a silent video. *"It captures more emotion in the scene without being distracting"* (P25).

Finally, participants were asked if they could think of anything that would make this medium better in an open-ended question. Four themes were identified within these responses:

1) Expressive Text: 5 participants mentioned that using different fonts, punctuation and capitalisation would help make the dialogue more expressive: *"The font can be bolder when the actor is being more expressive/abrupt, eg when Karen Gillan says "fine be busy" etc."* (P26), *"Similarly to graphic novels you could capitalise if they're shouting / exclaiming something. Like when she says "FINE!""* (P33).

2) Varying panel size/shape: 3 participants thought that a more varied panel size to highlight more important frames. For example: *"Panel sizes are too consistent, varying sizes help with effect."* (P21), *"Making the panel sizes vary depending on the importance of what's being said in each one"* (P34).

3) Number of panels: 3 participants mentioned number of panels, 2 of whom thought there should be more frames per shot to better capture the motion and 1 who thought there were too many panels per page (Likely due to a device issue).

3.6 Summary of Survey Findings

Overall, the survey has demonstrated that there are issues with current, mainstream subtitling methods and although people are more familiar with the former, many are willing to explore other

mediums for media consumption. Participants have highlighted any issues present in the proposed graphic novel subtitling which will be kept in mind for the implementation of this application. Suggestions for improvement will also be considered depending on their complexity.

4 THE SYSTEM

From the research into related work in this area and taking into account feedback from Study 1, an outline of the proposed system has been created.

4.1 Application

A web-based application will be created that allows users to convert videos into graphic novels. Features of this application will be updated as the project progresses as the complexity of some features has not yet been determined. Two lists have been created to represent baseline features and additional features. Baseline features have a higher priority and must be achieved. Once these features are functioning, additional features may be explored. These are outlined below:

1) Baseline Features: A web-based application where users can upload video and subtitle files which are converted into an online graphic novel.

1.1 File loading (i.e. video files, subtitle files)

1.2 Key-frame identification using difference between frames

1.3 Frames displayed sequentially in panels

1.4 Subtitle segmentation (i.e. dialog matching panel content, displayed at bottom)

2) Additional Features: Users can edit the graphic novel by placing speech balloons with panel dialog, user can select non-uniform panel sizes and the finished graphic novel can be downloaded by the user.

2.1 Downloadable output file

2.2 User modification of subtitle placement (i.e. adding speech balloons)

2.3 Varying panel sizes

The specific frameworks and technologies to be used for implementation will be researched in Sprint 1.

4.2 Design

The Related Work section of this report highlighted a gap in user-oriented design of similar applications. So, user experience is the main focus in the design of the proposed application with the goal of a self-explanatory interface that is easy to use. Prototyping will be carried out in AdobeXD with adherence to the Web Content Accessibility Guidelines (WCAG 2.1) [42].

5 THE PLAN

This project will follow a combination of Waterfall and Agile methodologies. The waterfall method was used for the research and planning phases of this project. This decision was made due to the serial nature of the work and the time commitments to other modules during semester 1. An agile approach will be adopted for the design and implementation phases of the project. Weekly sprints with set goals will be crucial for creating a timeline and understanding the scope of the project.

5.1 Progress so far

The first half of this project has been heavily research-based. A literature review was created to organise the research and condense the volume of information that needed to be examined. The initial focus was on subtitling and through this research two further areas were identified: video summarisation and automated comic creation. The literature review was organised into sub-categories which allowed for easy look-up when writing the Related Work section of this report.

Graphic Novel subtitling is a new research area and so personal research had to be undertaken alongside the literature review. A questionnaire was created to gather participant opinion on the topic and output of the proposed system. The results highlighted issues with current subtitling and feedback on a graphic novel subtitling file. The combination of these research tasks has created the groundwork for a user-oriented design and implementation.

5.2 Things still to do

The next phase of this project will follow an agile approach with sprint 1 starting the week beginning 1/2/21. A rough timeline for agile development has been layed out in the Gantt Chart displayed in Figure 1. Trello will be used for the product and sprint backlogs, project boards have already been set-up.

The design phase will be carried out in sprint 1 where wireframes and an interactive prototype will be created using AdobeXD with a focus on usability. Research has been undertaken for this area in the form of the semester 1 User Experience module where the focus was on design sprints and the importance of user-oriented design.

Once a clear design has been established, implementation can begin. The scope of the implementation phase will become more accurate as the project progresses. Research into the required technologies will be carried out to make this more clear. A proposed timeline is outlined in the Gantt chart below.

Due to the exploratory nature of this project, it is easy to become over-ambitious. Two subsections of requirements have been developed to reduce this risk. This is why a set of baseline features has been created followed by a set of additional features that can be explored if the time allotted allows it. The product and sprint backlogs will help with the planning and organising of the timeline for these features.

The sprint documentation will the kept up-to-date on a weekly basis. A sprint retrospective will be written at the end of each sprint to measure the success of the sprint and, identify goals and priorities for the next sprint, this will also aid the writing of the final report.

User testing will be carried out at the end of the implementation phase, this will involve user testing of the actual application. The basis of this testing will be outlined later in the development process.

The report will be a constant focus but formal writing will commence towards the end of the implementation phase. The mid-term report, sprint retrospectives and user testing analysis will all assist with the report writing.

Weekly project advisor meetings will assist in the overall development of this of this project.

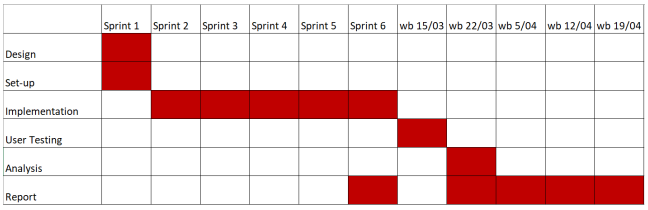

Figure 1: Gantt Chart

6 REFLECTIONS

The project is progressing as planned. The research phase has been highly successful with key research areas being identified. The questionnaire was a follow-on that has supported conclusions identified the related research section. This in-depth research has created a good groundwork for implementation.

A steep learning curve that has been identified so far would be the process of academic writing. The literature review, along with the semester 2 research frontiers module has assisted with this due to the high number of research papers that were required to be read. This analysis of papers has identified common themes in the structure of academic writing which has helped with this report.

In writing this report, it has been identified that the scope of the project could be more accurate at this point in time. This should hopefully become more clear in the coming weeks with the beginning of the implementation phase and further research into the required technologies.

7 CONCLUSION

For some people, the language and cognitive demands of watching subtitled content is challenging. Issues surrounding traditional subtitling have been explored through background research and a user study. Common themes in these issues have arisen including; subtitle speed [39], subtitle readability [29], subtitle position [12] and the constant battle for attention between moving subtitles on a moving picture [21].

Related research has been carried out to study the possibility of presenting subtitles in a different manner in order to improve the viewing experience. This includes subtitle personalisation (i.e. for speed and size) [3] [29], dynamic placement of subtitles [2] and adaptive subtitling [17].

An application has been proposed that mitigates many of the issues surrounding current subtitling methods. A graphic novel representation of a video allows the user to process the information at their own speed, which can result in a more immersive experience. Study 1 explored participant opinion on this proposed alternative subtitling format. 76% of participants gave positive or neutral feedback compared to 17% of participants who had a negative response to the medium. Written feedback included responses relating to the readability of the medium, stating it was clear and easy to understand. Enjoyment was a also a prominent factor in the feedback with participants mentioning that it was a fun new opportunity for media consumption. There was a divide in the opinions on immersion into the material, with some participants stating that it takes away from the atmosphere and pacing of the scene in comparison to others who thought they were able to get more emotion from the graphic novel compared to regular subtitling.

7.1 Limitations

The questionnaire on opinions on subtitling, focused only on the perception of graphic novel subtitling. The proposed application is a semi-automated conversion that will require some user input. The nature of the questionnaire did not allow for exploration into the opinion on user creation of a graphic novel, only opinion on the graphic novel itself. User testing will be carried out as the final application is implemented to mitigate this user opinion barrier.

The Application will require a subtitling file along with the video. This places limitations on the video content that users can convert. For example, if a user wants to create a graphic novel for a family video, they would have to create a subtitling file to go alongside it. A solution to this is not in the scope of this project but it could be explored in future work.

Automatic placement of speech balloons has been explored in the past and would be a good addition to this project. However, due to the complexity of the feature and the time constraints of the project, it will not be included in this application. User placement of speech balloons may be explored depending on progress level.

7.2 Generalisations and Future Work

As mentioned in the limitations section, this work could be expanded to include functionality for conversion of standalone videos without a subtitling file. Automatic generation of subtitles using voice recognition software [18] could be explored to allow for more casual video to graphic novel conversion (i.e. for scrap-booking purposes). The focus of this project is very much on accessible media. Automatic generation of subtitles is generally frowned upon in subtitling guidelines [43] so this would not be a suitable feature for the research area of this particular application.

REFERENCES

[1] Mike Armstrong, Andy Brown, Michael Crabb, Chris J Hughes, Rhianne Jones, and James Sandford. 2015. Understanding the diverse needs of subtitle users in a rapidly evolving media landscape. (2015).

[2] Andy Brown, Rhia Jones, Mike Crabb, James Sandford, Matthew Brooks, Mike Armstrong, and Caroline Jay. 2015. Dynamic subtitles: the user experience. In *Proceedings of the ACM international conference on interactive experiences for TV and online video*. 103–112.

[3] Denis Burnham, Greg Leigh, William Noble, Caroline Jones, Michael Tyler, Leonid Grebennikov, and Alex Varley. 2008. Parameters in television captioning for deaf and hard-of-hearing adults: Effects of caption rate versus text reduction on comprehension. *Journal of deaf studies and deaf education* 13, 3 (2008), 391–404.

[4] Janko Calic and E Izuierdo. 2002. Efficient key-frame extraction and video analysis. In *Proceedings. International Conference on Information Technology: Coding and Computing*. IEEE, 28–33.

[5] Ying Cao, Antoni B Chan, and Rynson WH Lau. 2012. Automatic stylistic manga layout. *ACM Transactions on Graphics (TOG)* 31, 6 (2012), 1–10.

[6] Wei-Ta Chu and Hsing-Han Wang. 2012. Enabling portable animation browsing by transforming animations into comics. In *Proceedings of the 2nd ACM international workshop on Interactive multimedia on mobile and portable devices*. 3–8.

[7] Wei-Ta Chu and Chia-Hsiang Yu. 2013. Optimized speech balloon placement for automatic comics generation. In *Proceedings of the 3rd ACM international workshop on Interactive multimedia on mobile & portable devices*. 1–6.

[8] Bong-Kyung Chun, Dong-Sung Ryu, Won-Il Hwang, and Hwan-Gue Cho. 2006. An automated procedure for word balloon placement in cinema comics. In *International Symposium on Visual Computing*. Springer, 576–585.

[9] Neil Cohn. 2013. Beyond speech balloons and thought bubbles: The integration of text and image. *Semiotica* 2013, 197 (2013), 35–63.

[10] British Broadcasting Corporation. 2018. BBC Subtitle Guidelines. https://bbc.github.io/subtitle-guidelines/

[11] Michael Crabb, Rhianne Jones, and Mike Armstrong. 2015. The Development of a Framework for Understanding the UX of Subtitles. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility*. 347–348.

[12] Michael Crabb, Rhianne Jones, Mike Armstrong, and Chris J Hughes. 2015. Online news videos: the UX of subtitle position. In *Proceedings of the 17th international ACM SIGACCESS conference on Computers & accessibility*. 215–222.

[13] Charles Forceville, Tony Veale, Kurt Feyaerts, et al. 2010. Balloonics: The visuals of balloons in comics. *The rise and reason of comics and graphic literature: Critical essays on the form* (2010), 56–73.

[14] Tsukasa Fukusato, Tatsunori Hirai, Shunya Kawamura, and Shigeo Morishima. 2016. Computational cartoonist: A comic-style video summarization system for anime films. In *International Conference on Multimedia Modeling*. Springer, 42–50.

[15] Ciocca Gianluigi and Schettini Raimondo. 2006. An innovative algorithm for key frame extraction in video summarization. *Journal of Real-Time Image Processing* 1, 1 (2006), 69–88.

[16] Andreas Girgensohn, John Boreczky, and Lynn Wilcox. 2001. Keyframe-based user interfaces for digital video. *Computer* 34, 9 (2001), 61–67.

[17] Benjamin Gorman, M Crabb, and M Armstrong. 2021. Adaptive Subtitles: Preferences and Trade-Offs in Real-TimeMedia Adaption. (2021).

[18] Boris Guenebaut. 2009. Automatic subtitle generation for sound in videos.

[19] Anh Khoi Ngo Ho, Jean-Christophe Burie, and Jean-Marc Ogier. 2012. Panel and speech balloon extraction from comic books. In *2012 10th IAPR International Workshop on Document Analysis Systems*. IEEE, 424–428.

[20] Won-Il Hwang, Pyung-Jun Lee, Bong-Kyung Chun, Dong-Sung Ryu, and Hwan-Gue Cho. 2006. Cinema comics: Cartoon generation from video stream. *GRAPP* 6 (2006), 299–304.

[21] Jan-Louis Kruger, Agnieszka Szarkowska, and Izabela Krejtz. 2015. Subtitles on the moving image: An overview of eye tracking studies. *Refractory: A Journal of Entertainment Media* 25 (2015), 1–14.

[22] David Kurlander, Tim Skelly, and David Salesin. 1996. Comic chat. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*. 225–236.

[23] A. David Lewis. 2010. The shape of comic book reading. *Studies in Comics* 1, 1 (2010).

[24] Huayong Liu, Wenting Meng, and Zhi Liu. 2012. Key frame extraction of online video based on optimized frame difference. In *2012 9th International Conference on Fuzzy Systems and Knowledge Discovery*. IEEE, 1238–1242.

[25] Shujie Liu, Ke Zhang, ZHOU Xiaosong, Hsi-jung Wu, Chris Y Chung, James O Normile, Douglas S Price, and Albert E Keinath. 2019. Video analysis techniques for improved editing, navigation, and summarization. US Patent 10,452,713.

[26] Scott McCloud. 1994. Understanding Comics: The Invisible Art. 1993. *Reprint. New York: HarperPerennial* (1994).

[27] Emi Myodo, Satoshi Ueno, Koichi Takagi, and Shigeyuki Sakazawa. 2011. Automatic comic-like image layout system preserving image order and important regions. In *Proceedings of the 19th ACM international conference on Multimedia*. 795–796.

[28] Netflix. 2020. English Timed Text Style Guide. https://partnerhelp.netflixstudios.com/hc/en-us/articles/217350977-English-Timed-Text-Style-Guide

[29] Jakob Nielsen. 1998. Method and system for resizing the subtitles of a video. US Patent 5,805,153.

[30] Amy Pavel. 2019. *Navigating Video Using Structured Text*. eScholarship, University of California.

[31] Amy Pavel, Colorado Reed, Björn Hartmann, and Maneesh Agrawala. 2014. Video digests: a browsable, skimmable format for informational lecture videos.. In *UIST*, Vol. 10. Citeseer, 2642918–2647400.

[32] Elisa Perego. 2012. Eye tracking in audiovisual translation. Aracne.

[33] Suporn Pongnumkul, Jue Wang, Gonzalo Ramos, and Michael Cohen. 2010. Content-aware dynamic timeline for video browsing. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology*. 139–142.

[34] Jacqueline Preuß and Jörn Loviscach. 2007. From movie to comic, informed by the screenplay.. In *SIGGRAPH Posters*. 99.

[35] Zeeshan Rasheed and Mubarak Shah. 2003. Scene detection in Hollywood movies and TV shows. In *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.*, Vol. 2. IEEE, II–343.

[36] James Sandford. 2015. The Impact of Subtitle Display Rate on Enjoyment Under Normal Television Viewing Conditions. https://www.bbc.co.uk/rd/publications/whitepaper306

[37] James Sandford. 2016. How fast should subtitles be? https://www.bbc.co.uk/rd/blog/2015-09-how-fast-should-subtitles-be

[38] Yasuyuki Sumi, Ryuuki Sakamoto, Keiko Nakao, and Kenji Mase. 2002. Comic-Diary: Representing individual experiences in a comics style. In *International Conference on Ubiquitous Computing*. Springer, 16–32.

[39] Agnieszka Szarkowska and Olivia Gerber-Morón. 2018. Viewers can keep up with fast subtitles: Evidence from eye movements. *PloS one* 13, 6 (2018), e0199331.

[40] Hiroaki Tobita. 2010. Comic engine: interactive system for creating and browsing comic books with attention cuing. In *Proceedings of the International Conference on Advanced Visual Interfaces*. 281–288.

[41] Jaco Vermaak, Patrick Perez, Michel Gangnet, and Andrew Blake. 2002. Rapid Summarisation and Browsing of Video Sequences.. In *BMVC*. 1–10.

[42] w3. 2018. Web Content Accessibility Guidelines (WCAG). https://www.w3.org/WAI/standards-guidelines/wcag/

[43] w3. 2020. Captions/Subtitles in Making Audio and Video Media Accessible. https://www.w3.org/WAI/media/av/captions/

[44] Federico Zanettin. 2008. Comics in translation: An overview. *Comics in translation* (2008), 1–32.