

University of Dundee

## Graphic Novel Subtitles

Crabb, Michael; Gourlay, Amy

*Published in:*

ACM International Conference on Interactive Media Experiences (IMX '21)

*DOI:*

[10.1145/3452918.3465489](https://doi.org/10.1145/3452918.3465489)

*Publication date:*

2021

*Document Version*

Peer reviewed version

[Link to publication in Discovery Research Portal](#)

*Citation for published version (APA):*

Crabb, M., & Gourlay, A. (2021). Graphic Novel Subtitles: Requirement Elicitation and System Implementation. In *ACM International Conference on Interactive Media Experiences (IMX '21)*  
<https://doi.org/10.1145/3452918.3465489>

### General rights

Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from Discovery Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Graphic Novel Subtitles: Requirement Elicitation and System Implementation

Amy Gourlay  
University of Dundee  
Dundee, United Kingdom  
a.e.gourlay@dundee.ac.uk

Michael Crabb  
University of Dundee  
Dundee, United Kingdom  
m.z.crabb@dundee.ac.uk



Figure 1: An example comic created from our prototype system. Dr Who, The Power of Three (S07E04)

## ABSTRACT

Consuming subtitled video content relies on a viewers ability to match up and understand a number of visual inputs simultaneously. This can create challenges in immersion due to the overall readability of subtitles and the speed at which they are presented. In this paper we introduce Graphic Novel Subtitles as an alternative media consumption method that is based on combining video keyframes with subtitle text to create a comic-type experience. We carry out a requirement elicitation survey with 34 participants in order to explore this concept in more detail and identify key features that we present as system requirements. We then introduce a system that can automatically generate a graphic novel from video and subtitle files, and discuss our future evaluation plans.

## CCS CONCEPTS

• Human-centered computing → Accessibility.

## KEYWORDS

Accessibility, Graphic Novels, Subtitles, Survey.

### ACM Reference Format:

Amy Gourlay and Michael Crabb. 2021. Graphic Novel Subtitles: Requirement Elicitation and System Implementation. In *ACM International Conference on Interactive Media Experiences (IMX '21)*, June 21–23, 2021, Virtual Event, NY, USA. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3452918.3465489>

IMX '21, June 21–23, 2021, Virtual Event, NY, USA

© 2021 Copyright held by the owner/author(s).

This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in *ACM International Conference on Interactive Media Experiences (IMX '21)*, June 21–23, 2021, Virtual Event, NY, USA, <https://doi.org/10.1145/3452918.3465489>.

## 1 INTRODUCTION

The cognitive demands required to consume subtitled video content is challenging. A viewer must develop a basic understanding of the scene from the video itself, read the accompanying text, and then relate this text to the scene and specific characters. This complex action takes place repeatedly and all within a short time window. This set of actions can be made more difficult due to alterations in subtitle speed [35], readability [27], position [11], and the battle for attention between subtitles and the video itself [19].

The challenges that are faced when viewing time-dependant media such as a television show may be addressed by allowing people to consume content at a pace that is more suitable for their own circumstances. The simplest method to achieve this is to alter the playback speed of the media to make it faster/slower dependent on an individuals viewing preferences. However, temporal alterations such as this still require focused attention on multiple elements. An alternative media format that allows for the removal of time constraints during consumption is comic books and graphic novels. In this media format, the demand on following temporal pace is reduced and effort can be placed on reading comprehension and intake of images.

In this work we introduce Graphic Novel Subtitles. We present this as a summarisation of videos using keyframe identification to create a continuous pictorial narrative that is then combined with subtitle text: a graphic novel. To begin our investigations into this technique we carried out an online survey with 34 participants that aimed to understand the reasons surrounding why people watch subtitled videos and consume graphic novels. We also gather their initial feedback on a video file that had been manually turned into a graphic novel in order to elicit system design requirements.

## 2 RELATED WORK

Subtitles are a text version of the speech and non-speech audio information needed to understand media content [38]. This text is usually displayed at the bottom of the media being presented. Recent research in this area has highlighted the demand for subtitles to be presented in a different manner, with a focus on improving viewing experience for the user [17]. The presentation of subtitles within media follows best-practice guidance that ensures accessibility and familiarity across platforms [26]. This includes information relating to speech rate, text reduction, and positioning [9].

Traditional subtitle display does not always suit every individual's needs [17]. A common issue associated with subtitling relates to presentation speed being too slow or too fast for different viewers which can hinder comprehension of the content [32, 33]. It is recommended that subtitles are shown at a speed of between 160–180 words-per-minute [9], however it has also been suggested that due to individual user characteristics, options to select caption rate on a personal level would be beneficial [3].

There is significant diversity in the needs and preferences of frequent subtitle users which suggests personalisation is required for altering subtitle display [2, 35]. One mainstream method of subtitle customisation is the user-controlled resizing of subtitle text and subtitle viewing area [27]. Another method is adaptive subtitling which takes an object-based approach [1] to personalisation of subtitles that creates opportunities for adaptation surrounding words (e.g., names, locations), phrases (e.g. catchphrases), characters, audio-descriptive elements, and scene composition [17].

### 2.1 Comics as a Media Format

Comic books and graphic novels are art forms that combine text and images to effectively represent full (graphic novel) or serialised (comic book) stories [36]. Comics originated during the end of the nineteenth century in the USA as humorous panel drawings in newspapers [39]. 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 [39]. The development of technology in recent years has prompted the importance of digitisation of comics and graphic novels to offer a wider outreach for the medium [18].

Comic narrative is constructed specifically in the comic book medium [21]. 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*' [24]. There are many different elements involved in comic creation, the most prominent visual narrative aspects being; panel layout [5] and speech balloon placement [12]. 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 regions such as faces and prominent objects [25].

It is important that the dialogue associated with each panel matches the image contained within. Factors such as emotional presentation, length of the text being presented, and any graphical alterations to text are all areas that are considered. Once the text itself has been formatted, varying speech balloon sizes and shapes can be generated to house the resultant dialogue within the correct location [7] and to convey different emotions [8].

### 2.2 Video Summarisation

Comic creation is a time consuming art form and prior work has investigated the automation of this process. 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]. The barrier to entry for creating comics can be lowered through the use of pre-existing graphic material [36]. Comic building applications have been developed which allow users to create their own comics based on artist curated material [20, 34].

One method that can be used to assist in creating comics from existing video media is in the usage of video summarisation and key frame identification techniques. Video summarisation is often used to enable easier browsing and navigation of videos [22]. 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. Netflix); automatic indexing, search, and retrieval of material is therefore crucial [4].

There are three main methods used to support video summarisation. This includes keyframe, transcript, and metadata based techniques [28]. 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 [29]. Transcript-based summarisation is commonly used alongside video lectures where clear sections can be identified. Metadata for frames can be analysed to determine prioritisation data for a video that can then be used to automatically generate a textual summary [23].

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 [31]. In the context of summarisation, a series of keyframes can be extracted to represent key events while preserving the overall contents of the video [15]. This is achieved using keyframe detection that identifies sequences of keyframes that are maximally distinct and individually carry the most information [37]. Keyframe detection in video summarisation is used for efficient visual browsing and navigating of a video [16, 37] and dynamic timeline control [30].

### 2.3 Challenges in Media Consumption

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 may potentially be created as a storybook representation of a film or TV series. It is unclear, however, whether this method will assist in alleviating the language, cognitive, and temporal demands of watching subtitled content. We introduce graphic novel subtitling as a summarisation of videos using keyframe identification to create a continuous pictorial narrative that is then combined with text from subtitle files. We position graphic novel subtitles as creating a comic like media experience from existing video. 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. However, despite the potential that presenting subtitled videos in this manner creates, little is known about design requirements for this type of media.

### 3 REQUIREMENT ELICITATION SURVEY

In order to better understand opportunities for our graphic novel subtitling approach we carried out an online survey with 34 participants. This included 9 questions across three sections that presented the same media content across 3 different conditions. The first section gathered information on participant background with usage of subtitles and graphic novels. It focused on how regularly 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 traditional subtitling. The third section presented participants with a curated graphic novel based on the previous subtitled video. Participants were asked for their thoughts on this medium as a method of eliciting requirements for a future prototype system.

#### 3.1 Traditional Subtitling

Participants were 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 (based on [9]). Participants were asked to comment on the style of subtitling that was used through an open ended question. 15 participants gave a positive response, 10 participants gave a negative response and 7 participants were indifferent to the method. Key positive themes included:

**1) Improved Understanding:** Compared to the initial video, 5 participants mentioned an improved understanding of the scene content: *"Made it much easier to understand the scene"* (P13).

**2) Presentation:** 2 participants agreed that the subtitles were displayed well on the screen: *"Dark border makes it easy to read. Text box are an appropriate size."* (P21).

Following up on this question, all 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 / background colour makes subtitles hard to read	11

These results highlight that despite positive participant response to subtitling, issues still remain in how they are used to support the viewing experience. It also identified key areas that could be improved in our own implementation. Participants had the option to describe any other issues they had with this subtitling method. Some responses repeated the issues that were 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).

#### 3.2 Graphic Novel Subtitling

Participants were then shown the video in the form of a graphic novel. The graphic novel was created by taking keyframes from the video and taking into account shots, motion within shots, and dialogue per shot. 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. The 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: *"It is very easy to follow and read. It is clear who the speaker is, and flows well."* (P9), *"Interesting method, easy to read and understand."* (P11). 5 participants also noted that it was much easier to see who was speaking which was an issue raised regarding traditional subtitling. *"It shows who's saying what, which previous subtitles didn't make clear."* (P15).

**2) Positive impact on experience:** 17 participants reported a positive experience with the medium with a focus on accuracy and enjoyment: *"Surprisingly very effective,"* (P10), *"I think it looks fun"* (P2), *"It is cool, can 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."* (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 improve this medium. Three themes were identified within 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).

#### 3.3 Summary of Survey Findings

Our survey findings demonstrated that whilst people are familiar with traditional subtitling techniques, many are willing to explore other options for media consumption. Our participants found benefits in how graphic novel subtitling lets them have a direct input



into overall pace of consuming media and some found the experience to be “*surprisingly fun*” (P15). Despite this, some participants commented that this initial curated mockup was less immersive than the video media, but suggestions were made relating to the presentation of panels and textual elements that may improve this.

## 4 SYSTEM IMPLEMENTATION

In this section we introduce design requirements for a system that will be capable of creating a graphic novel from a video file and accompanying subtitle data. Based on previous literature and our own survey findings, we established three requirements for our system:

- (1) **Created not Curated:** The system should allow for graphic novels to be created automatically from a video and subtitle file with very little user intervention.
- (2) **Stories via Subtitles:** The system should use the subtitles as the key point for leading the story forward. These should be easy for the reader to locate and may also give additional expression to phrases by altering the text display.
- (3) **Immersion via Illustration:** The final created comic should promote immersion in the content. The system should use varying panel sizes/shapes and the number of panels per page may also be altered in order to add to the emotion of specific scenes.

We created our system as a web-based application built using React and Node.js. When a video and subtitle file are uploaded, keyframes are extracted and the subtitle file is parsed with text assigned to corresponding keyframes based on timestamp information. Frames are then styled according to user preferences.

### 4.1 Keyframe Detection

High quality keyframe detection must consider motion, shot length and colour properties of shots [40]. To create a graphic novel from a video, key frames also must represent all key events while preserving the overall contents of the video. In order to achieve this we used FFmpeg.wasm<sup>1</sup> to create the algorithm for key frame extraction with the `-force_key_frames` command used to create keyframes within specific time intervals. This was combined with `-skip_frame` and `nokey` to allow keyframes to be extracted from the uploaded video file. Through this method, accurate timestamps could be identified for the allocation of subtitles.

### 4.2 Assigning Subtitles

SRT files are a common file format for adding closed captions or subtitles to videos. An SRT file is a plain-text file that contains critical information regarding subtitles, including the start and end timestamp of the text to ensure the subtitles match the audio of a video [13]. For our application, an SRT file is uploaded alongside a corresponding video file to create text for the graphic novel panels. The uploaded SRT file is parsed in order to retrieve the timestamp and text. We use the start timestamp to assign text frames to our graphic novel as subtitles will not be split between panels. Subtitles are assigned to a panel using an algorithm that matches the subtitle start timestamp to the closest keyframe timestamp.

<sup>1</sup><https://www.npmjs.com/package/@ffmpeg/ffmpeg>

## 4.3 Graphic Novel Styling

The combined keyframe and subtitle data is displayed in a graphic novel style layout with panels and text. In line with the design recommendations that were derived from our initial survey we created the ability for panel layouts and text positioning to be altered according to user preferences. Users are able to flick through pages using arrows to mimic a real-life graphic novel experience. Once the user has chosen their desired layout, the graphic novel can be downloaded for offline viewing and consumption.

## 5 FUTURE WORK

Our work in this area so far has focused on the design and build of a system that can automatically create a graphic novel from a video and accompany subtitle file. The design requirements of our system was based on previous research relating to video summarisation techniques, comic creation, and our own user survey where we elicited additional requirements. Our future work aims to investigate the opportunities and challenges that exist when consuming longer media artefacts that are created using our system.

Our initial outputs for video files of 5 minutes in length have created graphic novels with between 150-175 panels. We aim to investigate the challenges that exist in consuming media of this length. Our investigation will focus on key factors relating to the user experience of subtitled media [10] and how these are impacted by this change in presentation. We aim to understand the impact that this media presentation format has on viewer involvement, overall endurability of the experience, the attention that viewers have towards this medium, and the overall temporal differences in consumption when compared to traditional subtitled video content. Future work should also examine the effectiveness of graphic novel subtitles for varying genres of media. Our initial investigations have used a combination of short-form drama and documentary type programmes, however we have also created graphic novels from quiz shows and childrens entertainment that we believe will also be valuable.

## 6 CONCLUSION

Engaging in the consumption of subtitled video media requires high levels of concentration from the viewer in order to focus on the text presented and also the video being shown. This in turn can lead to a reduction in viewer immersion in what is being watched. We introduce the concept of Graphic Novel Subtitles where we propose using video stills and subtitle text to create an alternative viewing experience.

In this paper we present findings from an initial user survey where we illustrate that whilst participants believe that traditional subtitles can assist with the delivery of a show, they result in a loss of impact from visuals and can lead to confusion around who is speaking. We then obtain initial reactions towards a curated version of a graphic novel styled implementation of the same media. We use our findings from this survey to create three design requirements for a system to automatically create graphic novel subtitled content and present our future plans to evaluate this media.

## ACKNOWLEDGMENTS

We would like to thank all participants for taking part in this work.

## REFERENCES

- [1] Mike Armstrong, Matthew Brooks, Anthony Churnside, Michael Evans, Frank Melchior, and Matthew Shotton. 2014. Object-based broadcasting-curation, responsiveness and user experience. (2014). <https://doi.org/10.1049/ib.2014.0038>
- [2] 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). <https://doi.org/10.5594/JMI.2016.2614919>
- [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. <https://doi.org/10.1093/deafed/enn003>
- [4] Janko Calic and E Izquierdo. 2002. Efficient key-frame extraction and video analysis. In *Proceedings. International Conference on Information Technology: Coding and Computing*. IEEE, 28–33. <https://doi.org/10.1109/ITCC.2002.1000355>
- [5] Ying Cao, Antoni B. Chan, and Rynson W. H. Lau. 2012. Automatic Stylistic Manga Layout. *ACM Trans. Graph.* 31, 6, Article 141 (Nov. 2012), 10 pages. <https://doi.org/10.1145/2366145.2366160>
- [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* (Nara, Japan) (IMMPD '12). Association for Computing Machinery, New York, NY, USA, 3–8. <https://doi.org/10.1145/2390821.2390825>
- [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* (Barcelona, Spain) (IMMPD '13). Association for Computing Machinery, New York, NY, USA, 1–6. <https://doi.org/10.1145/2505483.2505486>
- [8] Neil Cohn. 2013. Beyond speech balloons and thought bubbles: The integration of text and image. *Semiotica* 2013, 197 (2013), 35–63. <https://doi.org/10.1515/sem-2013-0079>
- [9] British Broadcasting Corporation. 2018. BBC Subtitle Guidelines. <https://bbc.github.io/subtitle-guidelines/>
- [10] 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* (Lisbon, Portugal) (ASSETS '15). Association for Computing Machinery, New York, NY, USA, 347–348. <https://doi.org/10.1145/2700648.2811372>
- [11] 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* (Lisbon, Portugal) (ASSETS '15). Association for Computing Machinery, New York, NY, USA, 215–222. <https://doi.org/10.1145/2700648.2809866>
- [12] Charles Forceville, Tony Veale, Kurt Feytaerts, 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.
- [13] Caroline Forsey. 2021. What Is an SRT File, & Why Is It So Important for Video? <https://blog.hubspot.com/marketing/srt-file>
- [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. [https://doi.org/10.1007/978-3-319-27671-7\\_4](https://doi.org/10.1007/978-3-319-27671-7_4)
- [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. <https://doi.org/10.1007/s11554-006-0001-1>
- [16] Andreas Girgensohn, John Boreczky, and Lynn Wilcox. 2001. Keyframe-Based User Interfaces for Digital Video. *Computer* 34, 9 (Sept. 2001), 61–67. <https://doi.org/10.1109/2.947093>
- [17] Benjamin M. Gorman, Michael Crabb, and Michael Armstrong. 2021. Adaptive Subtitles: Preferences and Trade-Offs in Real-Time Media Adaption. , Article 733 (2021), 11 pages. <https://doi.org/10.1145/3411764.3445509>
- [18] 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. <https://doi.org/10.1109/DAS.2012.66>
- [19] 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.
- [20] 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. <https://doi.org/10.1.1.149.2728>
- [21] A. David Lewis. 2010. The shape of comic book reading. *Studies in Comics* 1, 1 (2010), 71–81. <https://doi.org/doi:10.1386/stic.1.1.71/1>
- [22] 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. <https://doi.org/10.1109/FSKD.2012.6233777>
- [23] 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.
- [24] Scott McCloud. 1994. Understanding Comics: The Invisible Art. 1993. *Reprint*. New York: HarperPerennial (1994).
- [25] 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* (Scottsdale, Arizona, USA) (MM '11). Association for Computing Machinery, New York, NY, USA, 795–796. <https://doi.org/10.1145/2072298.2072461>
- [26] Netflix. 2020. English Timed Text Style Guide. <https://partnerhelp.netflixstudios.com/hc/en-us/articles/217350977-English-Timed-Text-Style-Guide>
- [27] Jakob Nielsen. 1998. Method and system for resizing the subtitles of a video. US Patent 5,805,153.
- [28] Amy Pavel. 2019. *Navigating Video Using Structured Text*. eScholarship, University of California.
- [29] Amy Pavel, Colorado Reed, Björn Hartmann, and Maneesh Agrawala. 2014. Video Digests: A Browsable, Skimmable Format for Informational Lecture Videos. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology* (Honolulu, Hawaii, USA) (UIST '14). Association for Computing Machinery, New York, NY, USA, 573–582. <https://doi.org/10.1145/2642918.2647400>
- [30] Suporn Pongnunkul, Jue Wang, Gonzalo Ramos, and Michael Cohen. 2010. Content-Aware Dynamic Timeline for Video Browsing. In *Proceedings of the 23rd Annual ACM Symposium on User Interface Software and Technology* (New York, New York, USA) (UIST '10). Association for Computing Machinery, New York, NY, USA, 139–142. <https://doi.org/10.1145/1866029.1866053>
- [31] Z. Rasheed and M. 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. II–343. <https://doi.org/10.1109/CVPR.2003.1211489>
- [32] James Sandford. 2015. The Impact of Subtitle Display Rate on Enjoyment Under Normal Television Viewing Conditions. <https://www.bbc.co.uk/rd/publications/whitepaper306>
- [33] James Sandford. 2016. How fast should subtitles be? <https://www.bbc.co.uk/rd/blog/2015-09-how-fast-should-subtitles-be>
- [34] Yasuyuki Sumi, Ryuuki Sakamoto, Keiko Nakao, and Kenji Mase. 2002. ComicDiary: Representing individual experiences in a comics style. In *International Conference on Ubiquitous Computing*. Springer, 16–32. <https://doi.org/10.1.1.521.2347>
- [35] Agnieszka Szarkowska and Olivia Gerber-Morón. 2018. Viewers can keep up with fast subtitles: Evidence from eye movements. *PLoS one* 13, 6 (2018), e0199331. <https://doi.org/10.1371/journal.pone.0199331>
- [36] 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* (Roma, Italy) (AVI '10). Association for Computing Machinery, New York, NY, USA, 281–288. <https://doi.org/10.1145/1842993.1843043>
- [37] Jaco Vermaak, Patrick Perez, Michel Gangnet, and Andrew Blake. 2002. Rapid Summarisation and Browsing of Video Sequences.. In *BMVC*. 1–10.
- [38] w3. 2020. Captions/Subtitles in Making Audio and Video Media Accessible. <https://www.w3.org/WAI/media/av/captions/>
- [39] Federico Zanettin. 2008. Comics in translation: An overview. *Comics in translation* (2008), 1–32.
- [40] Hong Zhang, Bo Li, and Dan Yang. 2010. Keyframe detection for appearance-based visual SLAM. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2071–2076. <https://doi.org/10.1109/IROS.2010.5650625>