What Makes It Hard for Users to Follow Software Tutorial Videos?

Saelyne Yang
School of Computing, KAIST
saelyne@kaist.ac.kr

Juho Kim
School of Computing, KAIST
juhokim@kaist.ac.kr

Abstract

Software tutorial videos provide more detailed explanations than ordinary text-based tutorials as they contain much more visual information and can also express continuous information. However, despite the advantages of tutorial videos, people often aren't able to correctly follow the task shown in tutorials. In this research, we perform studies using Photoshop tutorials to examine the factors that influence how well a viewer can follow the tutorial video. Our analysis shows that there are generally four possible sources that affect the understanding of the tutorial: the user, the software (Photoshop), the tutorial, and the interaction between the user and the tutorial.

Keyword

Tutorial video; Video interaction

1. Introduction

Many people get help from tutorials to complete tasks. These tutorials range from 'how to cook a chocolate chip cookies' to 'how to remove an object from an image using Photoshop'. Tutorials provide procedural steps involved in the complex tasks.

Compared to text-based tutorials, video tutorials have an extra dimension of visualization making it possible for them to provide much more descriptive explanations [8, 15]. For example, by watching how-to videos to learn cooking, one can gain details that would not be available in text-based recipes. For such reasons, video tutorials have been watched by a great number of people [8, 13]. However, there are still cases where people try to follow a tutorial video but fail to successfully complete the task [9, 13]. This can be due to many factors such as the speed and quality of explanation that ultimately influence how well the viewer understands it. Online video tutorials are not customized for each learner but instead are completely static. Because the situation and requirement will be different for all learners, it is difficult for the content creator to consider all possibilities and cover all cases with one video.

In this research, we analyze in depth the factors that influence how well users understand a video tutorial and succeed in following it. We perform studies using Photoshop tutorials with 16 participants. We identify that there are generally four possible sources that raise confusion and difficulties when following the software tutorial videos: the user, the software (Photoshop), the tutorial, and the interaction between the user and the tutorial.

2. Related Work

2.1 Text, Image and Video-based Tutorials

Previous research has introduced a wide range of systems to improve online text or image-based tutorials. A number of approaches incorporated users' input on a tutorial by integrating community feedback [4] or automatically tracking users' traces when following a tutorial [1]. Other approaches enabled users to utilize the tutorials in more effective ways such as by analyzing multiple tutorials to support browsing of tutorials [3, 16] or embedding users' images in tutorials [2].

Nowadays, the form of tutorials has evolved to videos, where people benefit from detailed visual information. Researchers have introduced a number of approaches to improve the video tutorials such as labeling each step of how-to videos [5] or enabling various interaction techniques when navigating through videos [6].

Our work aims to understand the unique challenges that occur in following video tutorials, which provide benefits over text or image-based tutorials but also difficulties at the same time.

2.2 Software Tutorial Videos

Video demonstrations of software are popular sources of help [8, 13] as they provide detailed information on input and mouse movements in complex software. There have been efforts to make software tutorial videos more interactive by making videos responsive to users' click on the video [7], automatically pausing and playing the video in response to users' actions on applications [8], and identifying UI elements in videos so that users can directly interact with them [11]. Some approaches captured demonstrations of the community to provide alternative ways of completing tasks [12] and to recommend workflows [10]. Researchers also have explored ways to improve the use of videos, such as contextually presenting appropriate videos to reduce the loads of navigation [14] or segmenting screen-capture videos into steps using application commands and input events [15]. We aim to understand the challenges that occur in following software tutorial videos and thus provide insights for a better interface and design of the tutorial system.

3. Study Design

To examine the factors that affect how well learners follow a video tutorial, we recruited 16 participants and observed their behavior when following a Photoshop tutorial. Participants ranged from the age of 20 to 40 and with diverse levels of background knowledge of Photoshop [Table 1]. We first conducted a survey to collect their personal factors.

Personal factors include:

- OS of the device used
- Photoshop version
- language setting of Photoshop
- proficiency in Photoshop

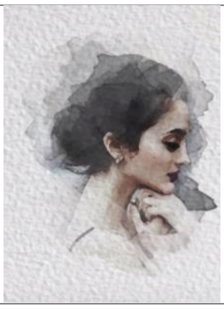
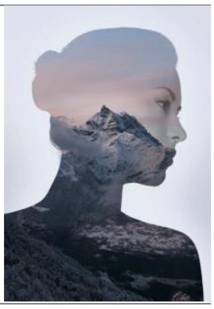
Based on the survey, we separated the participants into two groups: novice (12) and intermediate (4) users. This was based on their self-reports on their proficiency as well as how often they use Photoshop and 4 other questions related to Photoshop functionalities. Then, each participant was given a video tutorial of Photoshop which applies an effect on an image. Two different videos were used and each video was assigned to 8 users respectively [Table 2].

Participants were given about 45 minutes to follow the tutorial. They were encouraged to speak aloud their thoughts on how they perceive and follow the tutorial during the session, especially when they apply what they just saw from the tutorial. Participants' screens were recorded and a researcher took down notes on what they said. Once the participants completed the tutorial or the time went over 45 minutes, we did an interview for 15 minutes about their overall experience of the task and how we could improve the tutorial itself. Participants were compensated with ₩10,000 for their participation.

Table 1. Information of participants

		Novice (12)	Intermediate (4)
OS	Windows	9	2
	macOS	3	2
Version	Photoshop CC 2017	9	3
	Photoshop CC 2018	1	1
	Photoshop CS6	2	0
Language setting	English	6	2
	Korean	6	2

Table 2. Information of videos used in the study

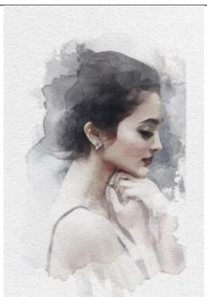
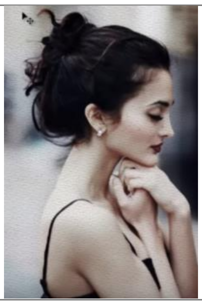
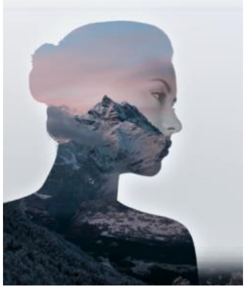
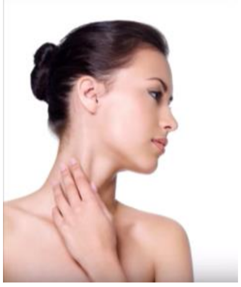
	Video 1 (v1)	Video 2 (v2)
Outcome		
OS / Version	Windows / Not found	macOS / Photoshop CC 2014
Language setting	English	English
Length	5m 52s	4m 52s

Purpose	Watercolor painting effect	Double exposure effect
Remark	Visual information on keystrokes	Descriptive audio explanation
Link	youtu.be/zJ8R3egJ_84	youtu.be/Mbf-QXCCXgM

4. Results

Participants generated diverse outcomes with varying time [Table 3]. For the first tutorial, 4 out of 8 successfully finished tutorial, while 6 out of 8 finished the second tutorial.

Table 3. Diverse outcomes with varying time

	Success	Failure
v1		
	31m 07s	Over 45m
v2		
	18m 48s	Over 45m

There were many variations in the performance of the participants: the outcome they produced, the time taken, and whether succeed in following the tutorial. By analyzing the screen recordings, the notes taken during the session and the interviews, we could see that there were 4 main sources where confusion could occur: Tutorial, User, Software, and Interaction [Figure 1].

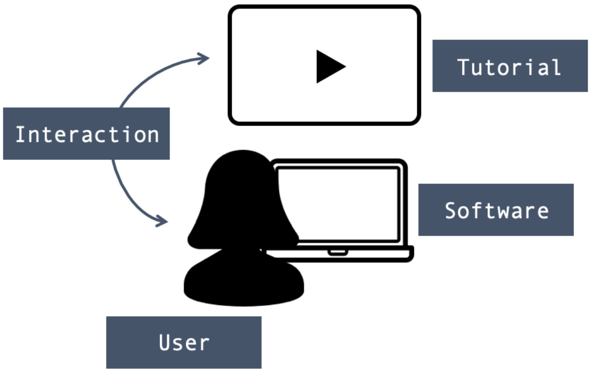


Figure 1 Sources where confusion occurs

4.1 Tutorials

Visual

Software like Photoshop or Premiere Pro has lots of icons and menus in them. However, because software tutorials often capture the whole screen, it is unlikely that small text or icons in it are visible clearly. We found that users often wanted parts of the screen to be enlarged in cases such as when clicking icons or menus, or typing a parameter value. Subtitles could be an additional aid here.

Also, there are many important features in the borders of software such as the top menus or icons at the bottom. However, video interfaces like titles or the progress bar often hide them so that users are not able to identify the important features well.

Explanation

There were some contents that were not explained by the creators of the tutorials. These include factual contents like the shortcuts they used or their customized setting, and also explanatory descriptions like how to fill the background with white.

Users also pointed out that the lack of information about each step made them confused. Being not aware of the purpose of each step made it harder for them to find alternative ways. Users wanted to know what the subgoal of each step was as well as exactly where they are at in the whole process.

4.2 User

Behavior

Users sometimes behaved differently to the tutorial without knowing it. They skipped some parts, usually the fast parts, and didn't know what caused problems. Also, they were unlikely to notice small differences such as the order of layers or lock/visible options of layer, which made them think that something else has caused a problem.

Background knowledge

Creators of the tutorial sometimes assume that users know something. However, the lack of background knowledge or basic concepts makes it hard for users to understand the task and figure out what caused the problem. Users often wanted more information about basic concepts such as what a layer is or what a particular function does.

Table 5. Results of users with their success status

	success	failure
Novice (12)	6	6
Intermediate (4)	4	0

Not surprisingly, intermediate users were all able to succeed, while half of the novice users failed in following the tutorial completely [Table 5]. This shows that prior or background knowledge influences how well one follows a tutorial.

The above points – Behavior and Background knowledge of users – can be considered as a problem of tutorials (4.1), since lack of information in tutorials can lead to users' misbehavior and misunderstanding. However, instead of distinguishing users' points from tutorials' points, we examine how difference in users' behavior and background knowledge can affect how well they follow the tutorials.

4.3 Software

OS

Tutorials often cover only certain operating systems (OS), while users watching the tutorial might use different types of OS. Different OS can have a different user interface of the software as well as different ways to execute a function. In the study, OS difference was a problem when an author of the tutorial used shortcut keys by telling which keys they used. For example, "command and click" only works on macOS while it should be done with "ctrl and click" in Windows.

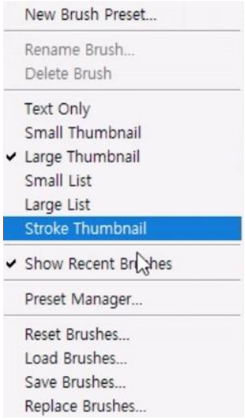
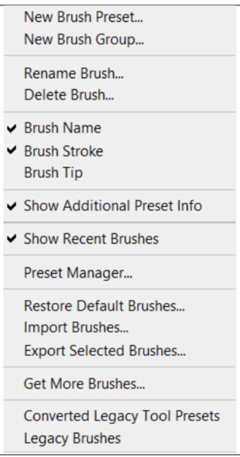
Version

Usually, a different version of the software has a different user interface or different functionalities. In the study, there was a case where a different version had a different interface. When customizing brush setting, the initial way the brush is shown and the menus of setting were different [Table 6].

Language setting

The difference in the language setting was a big problem for the users. When the language of the software in the tutorial is different from the language in the users' software, they tended to map the menus by their position but this didn't work in some cases. Some menus were sorted in alphabetical order, so then users tried to find the most similar word, but some of them struggled to translate and find the most appropriate word. Then they searched for it on the internet or used a dictionary. However, there were still cases where users felt that the translated word was awkward.

Table 6. Version difference makes UI difference

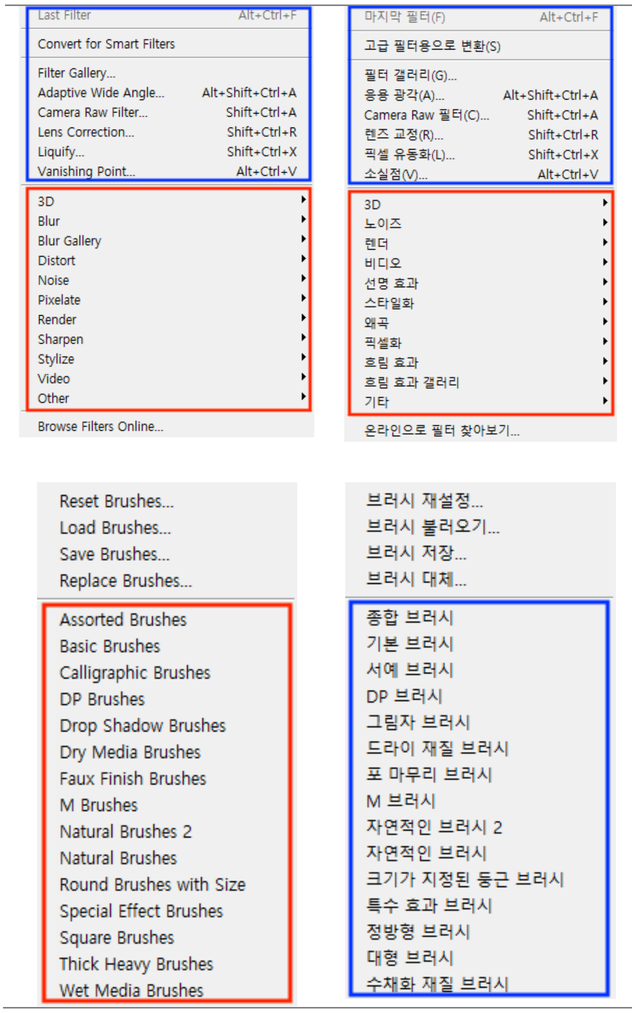
Photoshop CC 2017 (Windows)	Photoshop CC 2018 (Windows)
	

New Brush Preset...
Rename Brush...
Delete Brush
Text Only
Small Thumbnail
✓ Large Thumbnail
Small List
Large List
Stroke Thumbnail
✓ Show Recent Brushes
Preset Manager...
Reset Brushes...
Load Brushes...
Save Brushes...
Replace Brushes...

New Brush Preset...
New Brush Group...
Rename Brush...
Delete Brush...
✓ Brush Name
✓ Brush Stroke
Brush Tip
✓ Show Additional Preset Info
✓ Show Recent Brushes
Preset Manager...
Restore Default Brushes...
Import Brushes...
Export Selected Brushes...
Get More Brushes...
Converted Legacy Tool Presets
Legacy Brushes

Table 7. Different language setting has an inconsistent position rationale

English setting	Korean setting


Last Filter Alt+Ctrl+F
Convert for Smart Filters
Filter Gallery...
Adaptive Wide Angle... Alt+Shift+Ctrl+A
Camera Raw Filter... Shift+Ctrl+A
Lens Correction... Shift+Ctrl+R
Liquify... Shift+Ctrl+X
Vanishing Point... Alt+Ctrl+V
3D
Blur
Blur Gallery
Distort
Noise
Pixelate
Render
Sharpen
Stylize
Video
Other
Browse Filters Online...
마지막 필터(F) Alt+Ctrl+F
고급 필터용으로 변환(S)
필터 갤러리(G)...
응용 광각(A)... Alt+Shift+Ctrl+A
Camera Raw 필터(C)... Shift+Ctrl+A
렌즈 교정(R)... Shift+Ctrl+R
픽셀 유동화(L)... Shift+Ctrl+X
소실점(V)... Alt+Ctrl+V
3D
노이즈
렌더
비디오
선명 효과
스타일화
왜곡
픽셀화
흐림 효과
흐림 효과 갤러리
기타
온라인으로 필터 찾아보기...
Reset Brushes...
Load Brushes...
Save Brushes...
Replace Brushes...
Assorted Brushes
Basic Brushes
Calligraphic Brushes
DP Brushes
Drop Shadow Brushes
Dry Media Brushes
Faux Finish Brushes
M Brushes
Natural Brushes 2
Natural Brushes
Round Brushes with Size
Special Effect Brushes
Square Brushes
Thick Heavy Brushes
Wet Media Brushes
브러시 재설정...
브러시 불러오기...
브러시 저장...
브러시 대체...
종합 브러시
기본 브러시
서예 브러시
DP 브러시
그림자 브러시
드라이 재질 브러시
포 마무리 브러시
M 브러시
자연적인 브러시 2
자연적인 브러시
크기가 지정된 둥근 브러시
특수 효과 브러시
정방형 브러시
대형 브러시
수채화 재질 브러시

[Table 7] shows an example of the inconsistency in translating words. In the blue box in the first row, the position of each menu item is maintained when the items are translated, while in the red box the items are in alphabetical order and so their Korean counterparts changed positions to maintain the Korean alphabetical order. However, in the second row, while the names of brushes are in alphabetical order in English, their positions remain the same after they are translated into Korean.

Customized setting

Customized settings of the user interface or shortcuts might also cause conflicts when following tutorials. If there are any customized settings in the software used, users should be aware of those settings and see if there are any conflicts between their settings and the tutorials. Customized setting of the tutorial itself is also an important factor that users should know, which is mentioned in Tutorial–Explanation.

4.4 Interaction

Types of information from video tutorials

Video tutorial has many types of information such as subtitles, audio explanation, visual cues, and mouse track. Users felt looking at all information simultaneously was difficult. At the same time, they also felt that subtitles and audio explanations are both needed. They preferred audio for some explanations like what/why we are doing, and subtitles for key information.

Applying tutorial to software

Users had to apply what they learn from the tutorial to the actual software they are using. In this process, they found switching between the video tutorial and Photoshop inconvenient. Some of them opened two windows in one screen while others opened them at two separate screens. They felt there were too many pauses and plays interaction. Also, users using separate screens sometimes forgot what they just saw so that they had to replay the same part many times.

5. Discussion & Conclusion

We have analyzed in detail the factors that influence how well users follow a video tutorial: Tutorial, User, Software, and Interaction. With the identified factors, researchers and practitioners could come up with a design that tackles the source(s) of confusion. For example, a 'Video Tree' can be suggested, which is a system where the video tutorial is segmented into a sequence of steps and users can mark confusion points and suggest improvements for the explanations in each segmented video. This approach can improve the explanation part in Tutorial as well as the Background knowledge part in User. Other factors that are mentioned in this study could also be considered. A new design to tackle the problem of language differences in tutorials or a new interface that supports better interactions between the user and the tutorial can be proposed.

Because there still exist lots of limitations on the effectiveness of video tutorials, we believe working on solving each of the challenges will be meaningful for the improvement of human-video interactions.

Acknowledgement

This work was supported by KAIST Undergraduate Research Program. We would like to thank our study participants for their time.

Reference

1. Alok Mysore and Philip J. Guo. Porta: Profiling Software Tutorials Using Operating-System-Wide Activity Tracing. In Proc. UIST 2018, ACM Press (2008), 201-212.
2. Gierad Laput, Eytan Adar, Mira Dontcheva, and Wilmot Li. Tutorial-based interfaces for cloud-enabled applications. In Proc. UIST 2012, ACM Press (2012), 113-122.
3. Pavel, A., Berthouzoz, F., Hartmann, B., Agrawala, M. Browsing and Analyzing the Command-Level Structure of Large Collections of Image Manipulation Tutorials. Technical Report No. UCB/EECS-2013-167
4. Andrea Bunt, Patrick Dubois, Ben Lafreniere, Michael A. Terry, and David T. Cormack. TaggedComments: promoting and integrating user comments in online application tutorials. In Proc. CHI 2014, ACM Press (2014), 4037-4046.
5. Sarah Weir, Juho Kim, Krzysztof Z. Gajos, and Robert C. Miller. Learnersourcing Subgoal Labels for How-to Videos. In Proc. CSCW 2015, ACM Press (2015), 405-416.
6. Juho Kim, Philip J. Guo, Carrie J. Cai, Shang-Wen (Daniel) Li, Krzysztof Z. Gajos, and Robert C. Miller. Data-driven interaction techniques for improving navigation of educational videos. In Proc. UIST 2014, ACM Press (2014), 563-572.
7. Cuong Nguyen and Feng Liu. Making Software Tutorial Video Responsive. In Proc. CHI 2015, ACM Press (2015), 1565-1568.
8. Suporn Pongnumkul, Mira Dontcheva, Wilmot Li, Jue Wang, Lubomir Bourdev, Shai Avidan, and Michael F. Cohen. Pause-and-play: automatically linking screencast video tutorials with applications. In Proc. UIST 2011, ACM Press (2011), 135-144.
9. C. Ailie Fraser, Mira Dontcheva, and Scott Klemmer. Software videos: Rich content and learning potential, but a challenge for sensemaking. Ext. Abstracts CHI 2018.
10. Xu Wang, Benjamin Lafreniere, and Tovi Grossman. Leveraging Community-Generated Videos and Command Logs to Classify and Recommend Software Workflows. In Proc. CHI 2018, ACM Press (2018), Paper 285, 13 pages.
11. Nikola Banovic, Tovi Grossman, Justin Matejka, and George Fitzmaurice. Waken: reverse engineering usage information and interface structure from software videos. In Proc. UIST 2012, ACM Press (2012), 83-92.
12. Benjamin Lafreniere, Tovi Grossman, and George Fitzmaurice. Community enhanced tutorials: improving tutorials with multiple demonstrations In Proc. CHI 2013, ACM Press (2013), 1779-1788.
13. Kimia Kiani, George Cui, Andrea Bunt, Joanna McGrenere, and Parmit K. Chilana. Beyond "One-Size-Fits-All": Understanding the Diversity in How Software Newcomers Discover and Make Use of Help Resources. In Proc. CHI 2019, ACM Press (2019), Paper 340, 14 pages.
14. C. Ailie Fraser, Tricia J. Ngoon, Mira Dontcheva, and Scott Klemmer. RePlay: Contextually Presenting Learning Videos Across Software Applications. In Proc. CHI 2019, ACM Press (2019), Paper 297, 13 pages.
15. Pei-Yu Chi, Sally Ahn, Amanda Ren, Mira Dontcheva, Wilmot Li, and Björn Hartmann. MixT: Automatic generation of step-by-step mixed media tutorials. In Proc. UIST 2012, ACM Press (2012), 93.
16. Minsuk Chang, Leonore V. Guillain, Hyeungshik Jung, Vivian M. Hare, Juho Kim, and Maneesh Agrawala. RecipeScape: An Interactive Tool for Analyzing Cooking Instructions at Scale. In Proc. CHI 2018, ACM Press (2018), Paper 451, 12 pages