

Vigle: A Visual Graphical Learning Module on Optical Character Recognition

Umesh Rao

Department of Education, University of Hawaii

urao@hawaii.edu

<http://www2.hawaii.edu/~urao/Vigle>

Abstract: Students of the Arts and Humanities use OCR to convert scanned images of old text (pre-1800 AD). They need to know how digital text is extracted from the scanned image. Thanks to cell phones and images captured with them, understanding this is useful for everybody. The processing steps employed by a typical OCR software are, in order, Binarization, Deskew, Segmentation, Character Segmentation and Character Recognition. In this research project, a standalone asynchronous visual graphical learning environment (VIGLE) on Optical Character Recognition (OCR) was developed. Constructivist learning strategy was employed. The learning module was integrated into a website that works on mobile. The project attempts to generalize the instruction so that it is useful for everybody. Latest web technology was used for the implementation to achieve one stop interface, browser compatibility, responsive window sizing and interactive visual content. Binarization, Deskew and Segmentation modules were implemented in the time available. The VIGLE consists of a graphical representation and a visual interface to the lessons. Results show that the participants found both the graphical representation and the visual interface helpful. They found the incomplete learning module on OCR at best moderately useful in helping them digitize text.

Statement of the Problem

Students of the Arts and Humanities use OCR to convert scanned images of pre-1800 text. They need to know how digital text is extracted from the scanned image. The processing steps employed by a typical OCR software is, in order, Binarization, Deskew, Segmentation, Character Segmentation and Character Recognition. The Students of Arts and Humanities utilize the open source software Ocropy for such extraction of text. Each processing step can be controlled using built-in parameters. One can find such old text for free download in repositories such as the Internet Archive and Google Books. However, scanning of text and optical character recognition is becoming more pervasive today, thanks to cell phones and images captured with them. So understanding the topic of Optical Character Recognition can be useful for everybody. Many people I spoke with

expressed interest, among them some are my classmates in the Learning Design and Technology graduate program.

Images from old text (pre-1800 AD) or Hawaiian newspapers was processed using Ocropy and serves as examples for illustration. Commands to run Ocropy was provided. It is known that an instruction module built using constructivist principles is very effective at teaching specialized skills (Payne, 2009). The go-to interface for this instructional module was a visual graphical interface such as was used by Moosavian (Moosavian, 2017). It displays the processing steps, such as is shown in the Figure 1 below. Following constructivist principles, a standalone asynchronous instruction module was built (Sommerville, 2004).

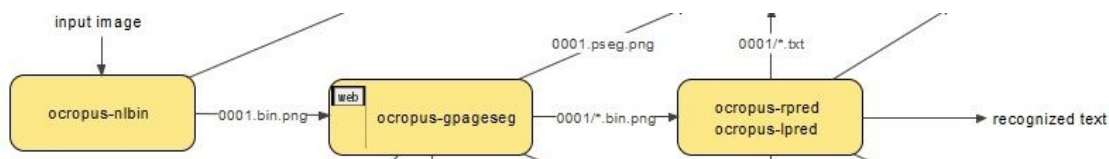


Figure 1: Processing steps in Ocropy

The purpose of this instructional design project was to design, develop and evaluate a stand-alone, asynchronous, visual and graphical learning module whose purpose was to help understand the topic of Optical Character Recognition for the students of Arts & Humanities and / or the students of Education, at the University of Hawaii.

Literature Review

It is advantageous to create a constructivist learning module at the workplace that is asynchronous (Sommerville, 2004). In this example, the guiding principles followed are, workers learn by doing and the purpose of leaders is to extend workers' interpretive understanding of, and relationship with, information. A study was done to find out the gaps in information among the workers. An organizational information system, called RISE, is then created. The constructivist form of learning is facilitated by an online discussion forum, an educational courseware stream and a knowledge base. To make tacit knowledge explicit, RISE workers initiate and sustain conversations among themselves. This data, information and knowledge are then captured in the form of formal representation in the information system. Finally, workers envision a more ideal information system and make changes to achieve that, on an ongoing basis.

E-learning that employs constructivist principles was effective at teaching employees highly specialized skill thought to require expert individuals and extensive prolonged training (Payne, 2009). The skill was identifying chromosomal abnormality in boars that causes reduced fertility in them. Results showed that trainees with no previous ability were able to learn the skilled task of chromosomal analysis in boars using this e-learning package. Trainees enjoyed the realism of the simulations. They also liked how the package was organized, as a series of problems to be solved, which they said kept them

engaged and motivated to learn. One observation is that the training material, the images of the chromosomes, is visual. This is similar to images in our project.

When it comes to high technology, systems are complex and interaction between systems are not easily understood. In such cases a graphical and interactive learning module (Moosavian, 2017) that gives the learners control of their learning, giving them time and opportunity to reflect, including sufficient examples that they can relate to and making the learning interactive will all be crucial. The Graphic syllabus was a graphical representation of the (text) syllabus, for clarity of communication. The interactive syllabus was an answer to students' different learning styles, offering a choice of media by which to study each title, as also a means for the learner to choose the order of progression. Another achievement of this work was the offering a "big picture" chart of Economics that was also interactive.

It is possible to develop critical thinking with constructivist teaching and learning (Topolovčan, 2017). The authors conclude that critical thinking apart from being an integral part of constructivist thinking, is also its most important dimension. Another conclusion was that the more frequent use of new media in class is connected to more frequent constructivist thinking and critical thinking. There is a long literature review at the beginning of this reference paper that goes over the latest publications on constructivist learning.

Video annotation tool incorporating in-video quizzes can predict learning outcomes and foster self regulated learning (Van Sebille, 2018). Self assessment strategies have long been known to aid self regulated learning. Specifically, the extent to which students utilize the available features of a video annotation technology called OVAL - Online Video Annotations for Learning, is explored. OVAL supports self regulation through socially shared annotations and through in-video quizzes. The research showed that interest and participation in the annotation feature declined as the course progressed.

Certain structural stimulus characteristics, such as novelty, surprisingness, complexity and ambiguity, lead to motivational states that result in curiosity and exploratory behavior (Berlyne, 1960 and 1974). Failure is known to be very effective in the learning process (Carroll, 1987, 1990) (Ivancic, 2000).

Ocropus is the optical character recognition software of choice for historical text from prior to 1800. Thomas Breuel (Breuel, 2008) is the original author of this software and is now open source on Github. Ocropus implements a Long Short Term Memory (LSTM) to achieve high performance (Breuel, 2013).

The learning solution of my project was the application of a constructivist form of learning to teach a specialized skill. The constructivist theory states that the occurrence of learning is by the learner, for the learner and in the learner. The learner interprets and interacts, based on his or her past history and knowledge, to process the images received through the senses to further build new knowledge, that is, binarization, segmentation,

etc. The module was asynchronous, visual, interactive and uses images of real text to teach a specialized skill. All the tests, pre-, post- and embedded, were self assessment tests, to promote self-regulated learning.

Methodology

The Visual Graphical Learning Module on Optical Character recognition can be taken by a student on their own time and assess for themselves their own performance, just like the graphical syllabus (Moosavian, 2017). It was also completely asynchronous like RISE (Sommerville, 2004). Finally, it was used to teach a specialized skill using the visual nature of images, similar to chromosomal analysis (Payne, 2009).

Research Questions/Goals.

The research questions of this learning module are:

- a) Does a learning module on OCR help students of the University of Hawaii, digitize text?
- b) How effective is the visual graphical constructivist format for learning about OCR?

The goals of the instruction module are that the participants will be able to:

- a) Explain each of the processing steps in the Optical Character Recognition of text, starting from a scanned image.
- b) Describe the effect of a parameter of a processing step on the image, in relation to OCR.

Content Analysis.

The domain of learning was primarily cognitive; Learning about OCR, the different processing steps and how the processing occurs using the different parameters. The intellectual skills are split into discrimination, concept, rule and problem solving, as shown in Appendix B. The visual and graphical nature of the user interface and the instruction falls under the affective domain of learning.

Performance objectives for each behavior were delineated (please see Appendix A). The terminal objectives of the instruction module are:

- Given a scanned image of a page of a document, the resulting binarized and segmented images, the University of Hawaii student will be able to explain the effect of binarization and segmentation of the image including black, white and grey colors, text sections, its corresponding segments and angle of text sections with respect to the border of the image.
- Given a scanned image of a page of a document, the University of Hawaii student will be able to determine at least two parameter values of the processing steps in Ocopy that will improve the final results of the processing such as reducing grey and reducing skew.

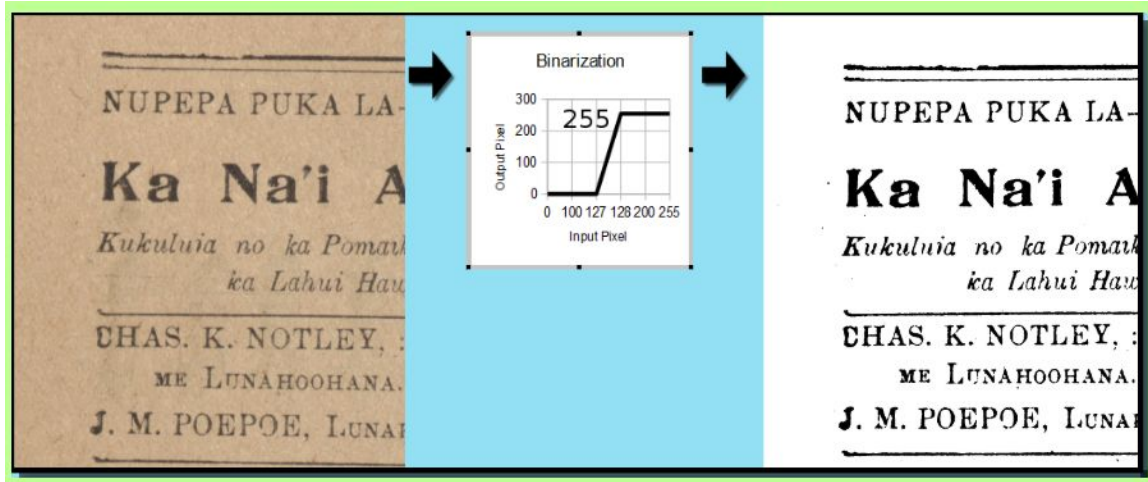


Figure 2: Visual Graphical Interactive Interface

Figure 3: VIGLE 2: The Visual Interactive interface

A visual graphical interface was developed that starts with a graphical representation of a processing step, such as for Binarization shown in Figure 2. On clicking the graphical representation, a visual interface to the learning module for the processing step was displayed, as shown in Figure 3.

There were three such processing steps that were implemented and a visual graphical interface developed for each. This is Binarization followed by Deskew followed by Segmentation. Each processing step displays links to pre-quest and post-assessment tests, as well as Learn and Apply sections.

More images of the graphical interface and the color scheme used can be found in Appendix G.

Participants.

The participants were initially envisaged to be undergraduate and graduate students of Arts and Humanities at the University of Hawaii. There is an initiative called the Digital Arts and Humanities Initiative where students embark on digitization of old text. The students of this initiative were to be invited to participate. Since the learning module does not encompass all the steps of Optical Character Recognition, it was decided not to solicit participation from the students of the Arts and Humanities. All of the participants were students and alumni of the Learning Design and Technology Program, Department of Education at the University of Hawaii.

The following information about the participants was gathered during the post survey. The participant population age was 25-45. They were in college because they were able to afford it. All of them had completed a Bachelor degree. They were not quite motivated to learn about OCR. An invitation for participation was sent by email, as shown in Appendix D. The consent form is shown in Appendix E.

Evaluation Instruments.

There were three processing steps illustrated in the graphical interface. They were Binarization, Deskew and Segmentation. Each processing step had pre-quest and post-assessment tests linked to it, as shown in Figure 3 earlier. Commands to apply binarization using Ocropy were provided.

The pre-quest for the Binarization step assessed the discriminatory behavior for this step, namely:

Given an image with a series of black, grey and white regions, the learner is able to correctly identify the black, grey and white regions within a time period of half an hour. The readings for this step was about pixels and the technology of scanning.

The post test for the Binarization step assessed the concept behavior for this step, namely: Given three binarized images hidden among four other non-binarized images, the learner is able to correctly analyze the images and identify the three that do not contain grey color within a time period of ten minutes.

The pre-quest for the deskew step assessed the discriminatory behavior for this step, namely:

Given clips from magazines that are deliberately text sections at different angles with the border of the image, the learner is able to identify the top three sections that exhibit the maximum angle with the border of the image.

The post test for the deskew step assessed the concept behavior for this step, namely:

Given eight clips from magazines that are deliberately text sections from pages scanned at different angles with the border of the image, the learner is able to identify the skew angle of the parent image in at least six of the eight cases.

The pre-quest for the segmentation step assessed the discriminatory behavior for this step, namely:

Given sections of text from a newspaper or magazine and other sections that are not, the learner is able to correctly identify the text sections in all such cases.

The post test for the segmentation step assessed the concept behavior for this step, namely:

Given four images that has been segmented into its text sections, each a different number of sections, the learner is able to identify the text sections in the image prior to processing and the corresponding segments after the processing, for at least 95% of the segments.

An additional module level test is included here and is:

Given images of the individual steps in segmentation, the learner is able to identify the name of the step from four different choices, in all the cases.

The readings for this step consisted of an introduction to sections of text, segmentation and steps in the method to do segmentation. The complete chart of instructional objectives is shown in appendix A.

Project Design Strategies.

The project followed the constructivist learning strategy. The modules were illustrated with images of real text that have been manually edited or images that have been processed by Ocropy at different parameter values. The context was kept meaningful by choosing images of pre-1800 AD text, Hawaiian newspapers, local college brochures and magazines. An example image was provided for download and processing. The learner was led through a process of learning by discovery of images. Finally, the module was made massively open in that the learner had full control over their learning. The pre-assessment tests are called pre-quests signifying the quest of the learner to learn. It is up to the learner to take these tests and it is up to the learner to pace their learning.

Some of the tools mentioned by Moosavian for creating this graphical interface were Microsoft PowerPoint, Adobe Dreamweaver, Adobe Flash, ConceptDraw, Mindmap, Edge Diagrammer, etc. Free software that can also achieve this are HTML image map and H5P. On the other hand, the standalone module could perform better using self

regulated learning such as the use of the video annotation software called OVAL - Online Video Annotations for Learning (Van Sebille, 2018).

The LTEC faculty advised that once a learner is led outside of the user interface, the learner will lose attention on the learning and will not return back. This meant that a website was a better choice for this application. Initially I planned to develop the learning module with H5P and Learndash on the Wordpress content management system. These are well integrated and I have used them before. Unfortunately, with website development platforms such as Wordpress, Wix or Google sites, frames and image maps are not easily implemented. To implement a graphical interface for a learning module, the interface had to be static, preferably for all the modules. Since it was a website, it had to work on different browsers as well as on mobile systems.

The prototype was developed with frames and image maps. However, its pages were not responsive to the size of the display device. Also the graphical part of the visual interface was not showing clearly. So, the choice was made to develop the website using only client side programming. This was iframes, image maps using SVG, CSS, javascript and H5P. This was more work for me, but it achieved all the goals of one stop visual interface, static graphical interface for the OCR processing steps, interactive activities, browser compatibility and responsive scaling to different devices such as mobile. The website was not tested on only the Safari browser due to inaccessibility.

For visual design, information was kept simple and to the point to prevent extraneous processing, especially when it is an asynchronous module with little oversight (Mayer, 2011). Visual design elements and principles were used to enhance the visual appeal of the page (Hashimoto, 2002), such as fonts, lines, contrasting drop shadows, etc. Tetradic colors were chosen, as shown in Appendix H, where colors are used in the interface. Boxes in the visual interface were aligned to the Golden Section grid (Figure 4). Signaling was applied using colors to differentiate the header, the body and the content. Arrows were used to emphasize the serial nature of the processing.

The content created was made visual as much as possible. Videos and images were used to define concepts. Images were used for interactive activities. H5P was used to achieve the interactivity, such as course presentations, drag and drop, agamoto, etc. The LTEC lab was equipped with Ocropy installed on some of the Mac computers and arrangements made for participants to be able to schedule remote sessions and run the software. Appendix B shows the lesson plans delineated for the instruction module. Appendix C shows the lessons delineated.

Procedures.

The procedure followed for the project, in chronological order, was,

1. Wrote detailed project plan, documented IRB requirements, initiated solicitation for audience and collected information required for building the instruction module such as readings, technology, etc.

2. Continued drafting and revising project plan, began running simulations and generating images and began testing the technology for the creation of the graphical interface and the content.
3. Finalized project plans for approval, finalized the technology to use for the graphical interface and the content. Delineated the pre-quest and post-assess tests for each OCR processing step.
4. Upon IRB approval, began project implementation, populated the graphical interface with the pre-quest, post-assess and reading material. Sent email reminders to participants of upcoming mini coursework.
5. Completed project implementation, sent email with course details to participants, created a survey for student interviewing and started the final paper draft.
6. Conducted student interviews, analyzed data and completed final paper draft.
7. Created TCC Presentation Slides and conducted TCC Presentation.
8. Completed final paper.

The performance of the participant in the coursework was kept confidential and only shared with the participant. Private information was stripped before the data, such as performance, before it was used in reports. Reading material for the course was identified by searching the internet, the University of Hawaii Library website, etc. Information about the technology to use was collected by searching the internet. There were no pictures taken and no recordings made, of the participant. All responses were collected in text form.

There was an email sent to solicit participants for the learning module, it is shown in Appendix D. Their consent for participation and their permission for release of identifiable data was collected in the consent form, just before they began the course, please see Appendix E. The intent of the learning module was described in the same form. The timeline for the project is shown in Appendix J.

Results

Since the instruction module was an abridged version of the complete course, students of Arts and Humanities were not solicited as participants. Instead LTEC students and alumni were solicited as participants. This resulted in low turnout and lower participation. The initial one on one showed that the estimate of 2-3 hours per module was too conservative. This time duration estimate had turned some people away from participating. The time duration estimate was revised as 20-30 minutes per module and with advice from faculty, a second email was sent soliciting participation, including members of the cohort. Five participants completed the course. Two did not keep notes or scores. All participated in the post survey. Almost all participants gave feedback. The performance of the participants is illustrated in Figures 4 and 5 above. They performed worst with the Deskew module, then Binarization and Segmentation, in that order.

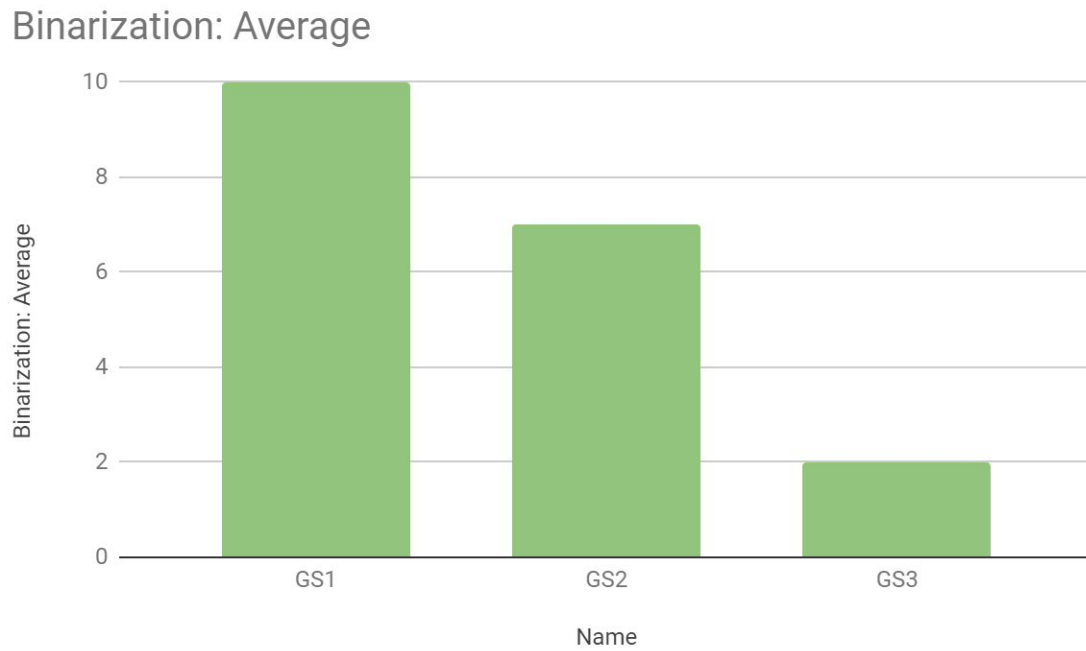


Figure 4: Average score of the participants in the Binarization module

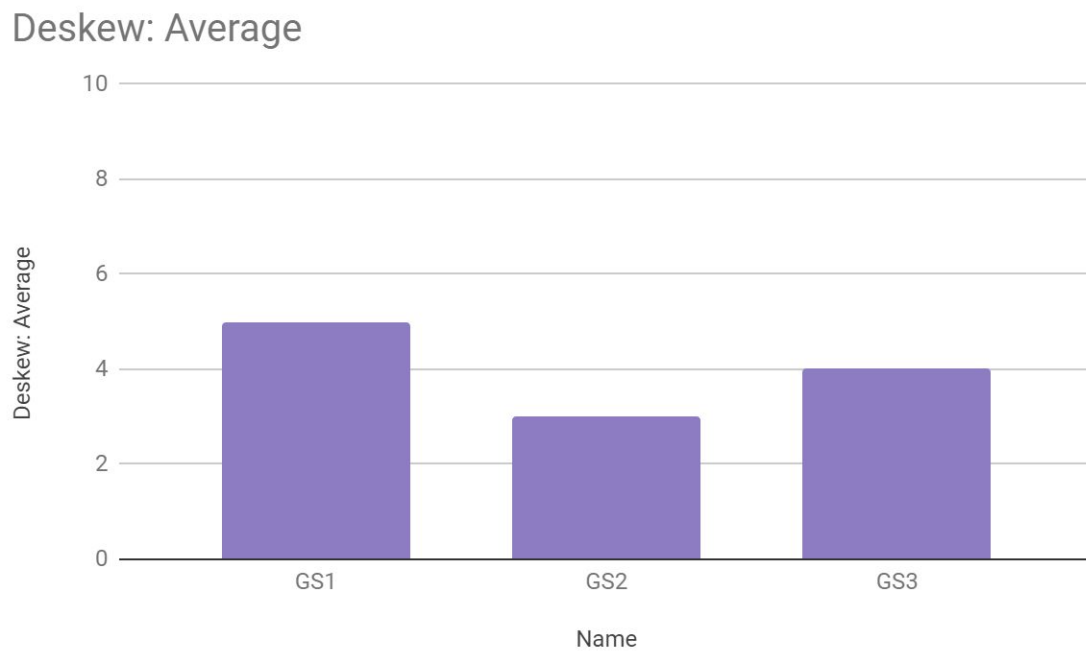


Figure 5: Average score of the participants in the Deskew module

Segmentation: Average

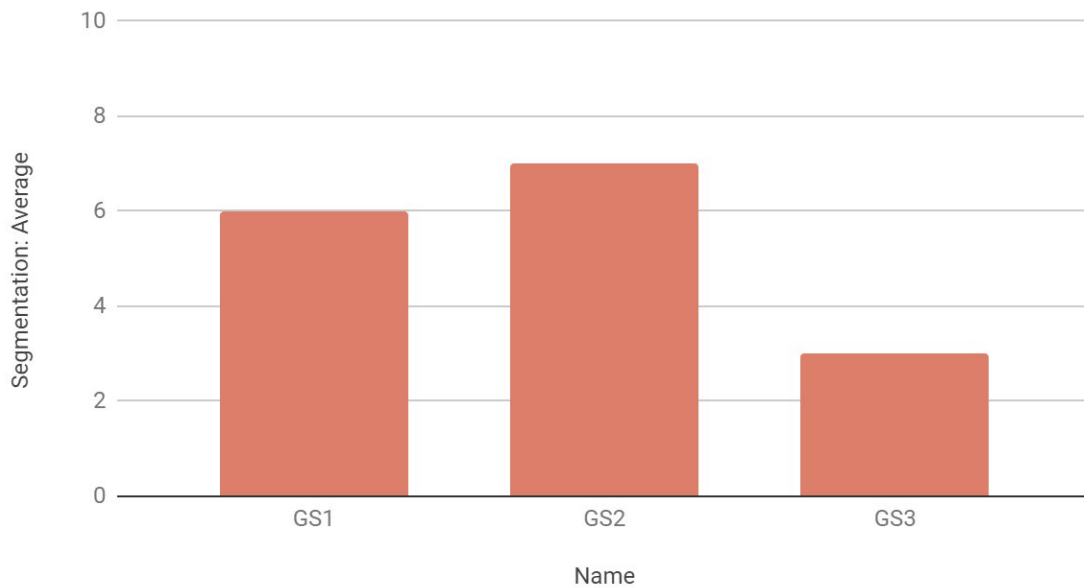


Figure 6: Average score of the participants in the Segmentation module

The overall module assessment showed the participants' performance was only average, as shown in the Figure 7 below.

Module Assessment

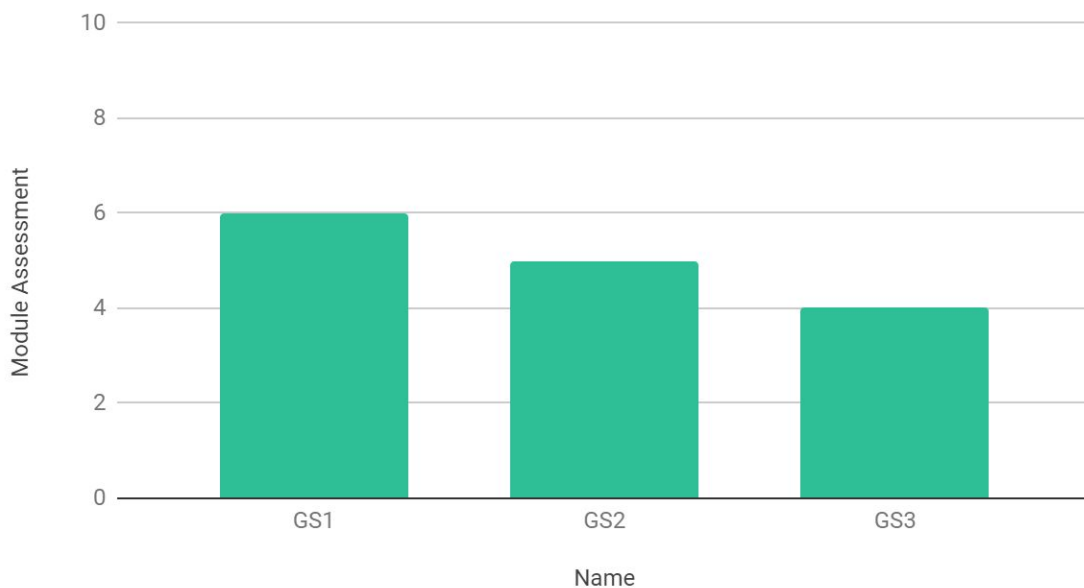


Figure 7: Performance of the participants in overall module assessment

Some of the non participants provided reasons for their non participation. Some of the participants who signed up did so without fully reading the consent form which had crucial details about the project. One was expecting personal touch. On reading the details one participant decided not being the right target audience for the course. One confused the lab for the course. All of them signed up, but did not participate.

One of the participants reported not being able to get the website to work on the Safari web browser. Several participants reported finding it inconvenient having to click the “Back” button of the browser. One participant additionally found by clicking the back button, coming across blank space and having to scroll the window. Two of the participants expressed a need to refer back to the content while answering the test questions, without explicitly clicking the back button. One wanted pause, forward and reverse controls on the video.

The audience's reaction was measured using a post-survey. An outline of this survey is shown in Appendix F. The participants reported low or medium utility for a course on OCR, in their lives, please see Appendix I for detailed results. For all of them, this learning module was their introduction to OCR.

The module on Binarization earned an average score of 65% on clarity. The Deskew module performed the worst with an average score of 45%. The average score was 60% for the segmentation module. Overall, the participants were ambivalent that the learning module on OCR motivated them to apply OCR to an image, at an average score of 50%. This was not enough to motivate them to try running Ocropy on an image.

6. The graphical interfaces such as the one in the figure below, for Binarization

5 responses

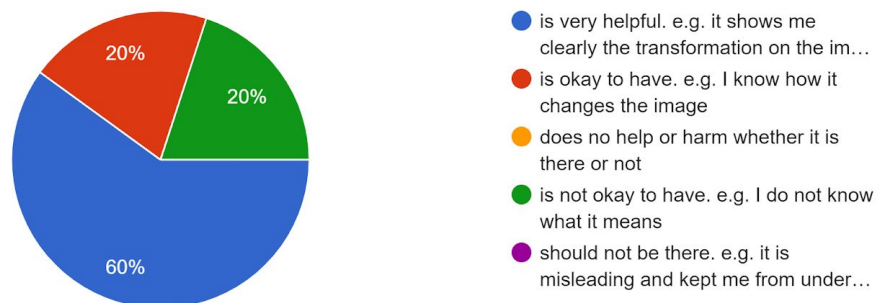


Figure 8: Participant rating of the Graphical Interface

The graphical representation, the visual interface and the H5P interactive activities was scored highly by the participants, as shown in the figures 8, 9 and 10 respectively. One participant spoke very highly of the H5P interactive activities, in particular.

7. The Visual interface to the learning modules, such as the one in the figure below, for the segmentation module

5 responses

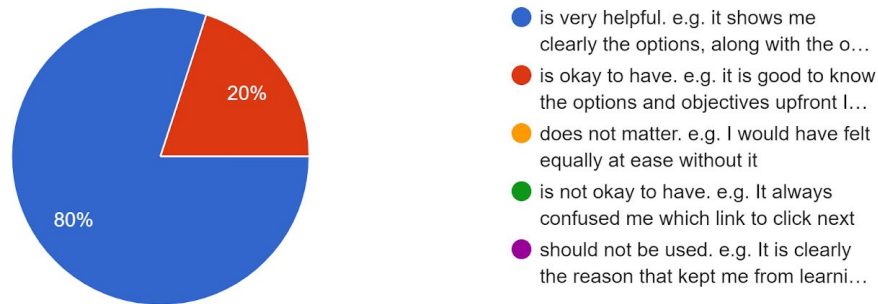


Figure 9: Participant rating of the Visual Interface

8. The H5P interactive activities, such as the one shown in the figure below

5 responses

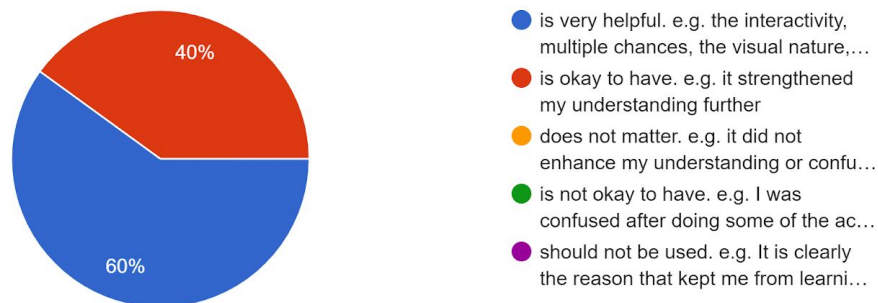


Figure 10: Participant rating of the H5P interactive activities

Discussion

As for limitations of the technology, the lesson I learnt during the choosing of the development platform is not to discount what is on hand. After searching different CMS, LMS and website tools, it turned out simple HTML5, iFrames, CSS, H5P and a bit of javascript meets the needs.

As it exists, the instruction module is only moderately effective at helping students digitize text and learning about OCR. The content was not tailored for LTEC and should

be simplified further. For example, guessing the skew number for a text section was too cumbersome. One the skew of one of the text sections was 35 degrees, which meant up to 36 guesses for the answer. Another exercise involved matching text section clip to the image it originated from. This necessitated going back and forth between the clip and the original image.

Participants found the videos and content interesting, but had difficulty answering the assessment questions. The participants reported difficulty with the Binarization and Deskew module more than the Segmentation module. The rigor involved in defining a section of text can be reduced. Its content was therefore too technical for first time OCR instruction.

The graphical interface worked as well as it did for Moosavian's graphical syllabus. The visual interface did as well as was discovered by Sommerville and Payne. What was missing from the module was the peer to peer interaction where some peers become teachers after they participate in the learning. The visual nature of the instruction helped as was its interactive nature, as was found by Payne as well.

Conclusion

When I took up a part time job with the Digital Arts and Humanities Initiative, University of Hawaii, one of the first few sentences its director told me is that there is little understanding among the students and faculty about how Optical Character Recognition works. Since they do not have a technical background there is a dire need for training. Additionally, Optical Character Recognition is becoming pervasive in everyday life, fueled by, images taken with cell phones. Building an instructional module that is useful for the former also helps the latter.

An asynchronous standalone visual graphical learning module on Optical Character Recognition was built and found to just meet this instructional need. The constructivist nature of the course was found to fit well for this need. This means it can serve as a continuing or supplementary education. Simplifying the content further, removing the browsing inconveniences and improving the signalling further will help. Supporting student-teacher-student interaction needs to be added for such a standalone course. Using H5P Application Programming Interface will allow student monitoring and providing feedback to the learner.

References

Berlyne, D.E. (1960). Conflict, arousal and curiosity. New York. Grove.

Berlyne, D.E. (1974). Novelty, complexity and interestingness. In *Studies in new experimental aesthetics*. 175-80. New York. Wiley.

- Breuel, T. (2008). The OCRopus open source OCR system. *Proceedings of SPIE 6815, Document Recognition and Retrieval XV, 68150F*.
- Breuel, T.M., Ul-Hasan, A., Al Azawi, M., Shafait, F. (2013). High-Performance OCR for Printed English and Fraktur using LSTM Networks. *12th International Conference on Document Analysis and Recognition*.
- Carroll, J.M., 1987. Minimalist design for active users. In: Baecher, R.M., Buxton, W.A.S. (Eds.), *Readings in Human-Computer Interaction. A Multidisciplinary Approach*. Morgan Kaufman, Los Altos, CA.
- Carroll, J.M., 1990. *The Nurnberg Funnel: Designing Minimalist Instruction for Practical Computer Skill*. MIT Press, Cambridge, MA.
- Ivancic, K., Hesketh, B., 2000. Learning from errors in a driving simulation: effects on driving skill and self-confidence. *Ergonomics 43 (12)*, 1966–1984.
- Hashimoto, A, 2002. Design Elements and principles of design. *Visual Design Fundamentals*, 1-50. Charles River Media.
- Mayer, R. E. (2011). Applying the Science of Learning to Multimedia Instruction. *Psychology of Learning and Motivation, 55*, pp. 77-108.
- Moosavian, S. (2017). Using The Interactive Graphic Syllabus In The Teaching Of Economics. *American Journal of Business Education – Second Quarter 2017 10 (2)*, 45-63. Retrieved February 4, 2018.
- Payne, A. M., Stephenson, J. E., Morris, W. B., Tempest, H. G., Milleham, A., & Griffin, D. K. (2009). The use of an e-learning constructivist solution in workplace learning. *International Journal of Industrial Ergonomics, 39*, 548-553. Retrieved February 4, 2018.
- Sommerville, M. M., & Vazquez, F. (2004). Constructivist Workplace Learning: An Idealized Design Project. *Lifelong Learning Conference*, 330-335. Retrieved February 4, 2018 from <https://eprints.qut.edu.au/109403/1/109403.pdf>.
- Topolovcan, T., Matijevic, M. (2017). Critical Thinking as a Dimension of Constructivist Learning: Some of the Characteristics of Students of Lower Secondary Education in Croatia. *Center for Educational Policy Studies Journal, 7(3)*, 47-66.
- Van Sebille, Y., Joksimovic, S., Kovanovic, V., Mirriahi, N., Laura Stansborough, R., Dawson, S. (2018). Extending video interactions to support self-regulated learning in an online course. *35th International Conference on Innovation, Practice and Research in the Use of Educational Technologies in Tertiary Education (ASCILITE 2018)*, Geelong, Australia.

APPENDIX A

Instruction Analysis

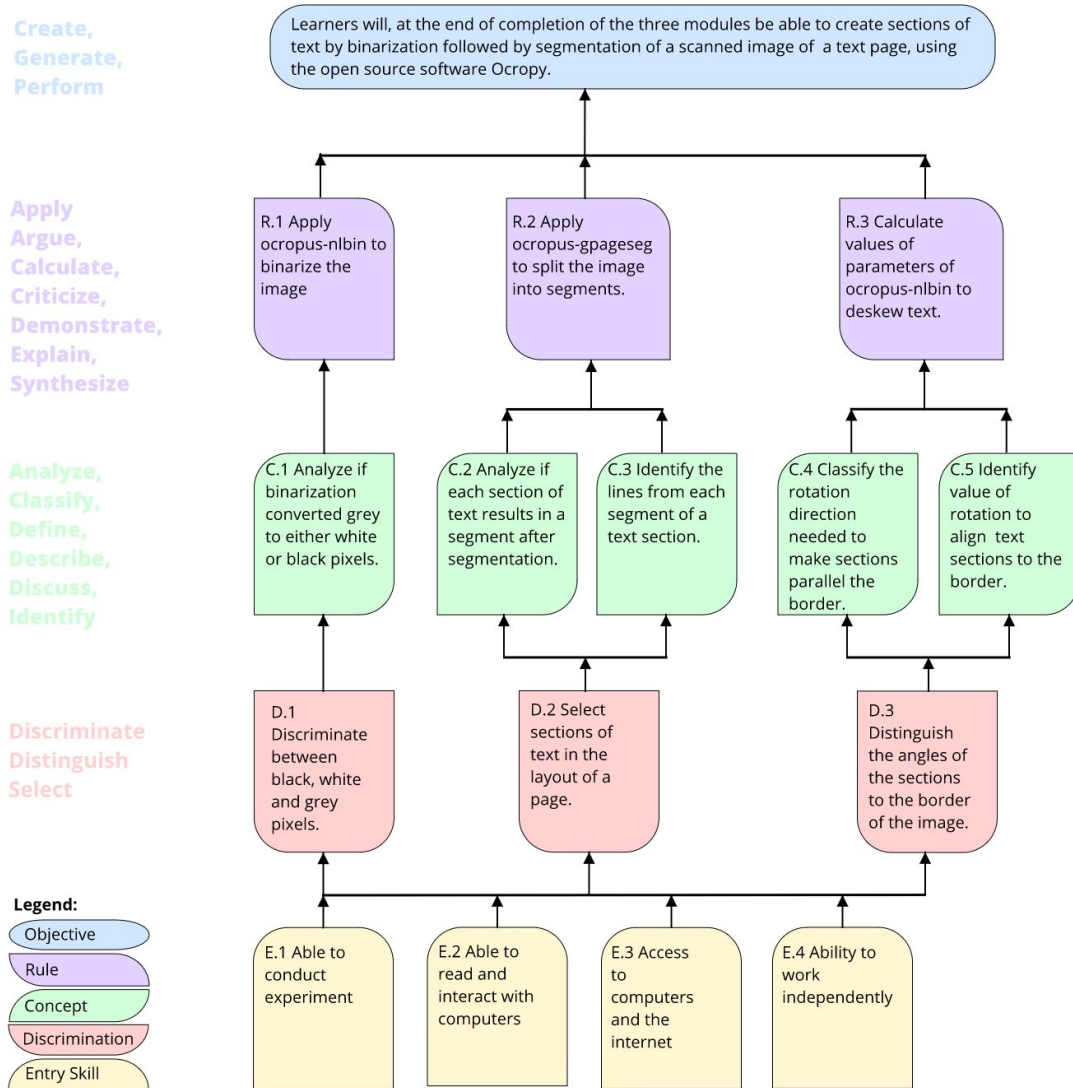
	Behavior	Performance Objectives
Entry Level		
E1.	Able to conduct experiment.	Given a computer keyboard that is not hooked to the computer, the learner is able to hook it up properly to the correct port on the computer and get it up and working within a time period of 5 minutes.
E2.	Able to read and interact with computers.	Given a working computer connected to the internet, the learner is able to open up a browser and search for the topic of Ocropy within a time period of 5 minutes.
E3.	Able to access computers and the internet.	Given the place where the learner lives, he or she is able to find a computer that is connected to the computer and equipped with an internet browser within an hour of requesting it.
E4.	Able to work independently.	Given the parts of a computer such as the monitor, the keyboard and the mouse, the learner is able to assemble and boot up the computer within an hour of requesting it and without needing supervision.
Discrimination		
D1.	Discriminate between black, white and grey.	Given an image with a series of black, grey and white regions, the learner is able to correctly identify the black, grey and white regions within a time period of half an hour.
D2.	Distinguish the angles of the sections to the border of the page.	Given clips from magazines that are deliberately text sections at different angles with the border of the image, the learner is able to identify the top three sections that exhibit the maximum angle with the border of the image.
D3.	Select sections of text in the layout of a page.	Given sections of text from a newspaper or magazine and other sections that are not, the learner is able to correctly identify the text sections in all such cases.

	Behavior	Performance Objectives
Concept		
C1.	Analyze if binarization converted grey to either black or white.	Given three binarized images hidden among four other non-binarized images, the learner is able to correctly analyze the images and identify the three that do not contain grey color within a time period of ten minutes.
C2.	Analyze if each section of text from the image becomes a segment after segmentation.	Given three images that has been segmented into its text sections, each a different number of sections, the learner is able to identify the text sections in the image prior to processing and the corresponding segments after the processing, for at least 95% of the segments.
C3.	Identify the lines from each segment of a text section.	Given an image and four other images that are results from segmentation using Ocropy, the learner is able to identify correctly which resulting image corresponds to which one of the five images in all the four cases.
C4.	Classify the rotation direction to make sections parallel the border of the image.	Given eight clips from magazines that are deliberately text sections from pages scanned at different angles with the border of the image, the learner is able to identify the skew, both angle and direction of the parent image in at least six of the eight cases.
C5.	Identify value of rotation to align text sections to the border of the image.	Given four images that are clips of text sections at different values to rotation, the learner is able to identify correctly which resulting text section corresponds to what value of rotation in all four cases.
	Behavior	Performance Objectives
Rule		

R1.	Apply Ocropus-nlbin to binarize the image. -> Describe the effect of binarization on the image.	Given an image and the binarized image, the learner is able to demonstrate the effect of binarization including black, white and grey colors.
R2.	Apply Ocropus-gpageseg to split the image into segments. -> Describe the effect of segmentation on the text sections of the image.	Given an image and the image after segmentation, the learner is able to describe and demonstrate the effect of segmentation including text sections and its corresponding segments.
R3.	Calculate values of parameters of Ocropus-nlbin to deskew text. -> Describe the effect of binarization on the angle of the text sections of the image.	Given an image and the image after binarization, the learner is able to describe and demonstrate the effect of binarization including angle of text sections with respect to the border of the image and rotation of the text.
Terminal Objective		
	Learners will, at the end of completion of the three modules be able to create	Given a scanned image of a page from a document, the resulting binarized and segmented images, the UH A&H student will be able to explain the effect of binarization and segmentation on the image including black, white and grey colors, text sections, its corresponding segments and angle of text sections with respect to the border of the image.
	sections of text by binarization followed by segmentation of a scanned image of a text page, using the open source software Ocropy.	Given a scanned image of a page from a document, the UH A&H student will be able to determine at least one parameter value (rotation) of the processing steps in Ocropy that will improve the image segmentation results.

APPENDIX B

Instruction Chart and Lesson Plans



Here is the lesson plan for the module on binarization.

Graphical Learning Module on Optical Character Recognition			
Lesson Plan			
Module: Binarization			
Learning Objectives: Upon completion of this module the learner will be able to:			
1. Discriminate between black, white and gray colors,			
2. Analyze the effect of binarization on a black, white or gray pixel,			
3. Describe the effect of binarization on an image, including the black, white and gray pixels,			
4. Apply ocropus-nlbin to binarize an image.			
Topics	Learning Stage	Exercises / Activities	Keyword
	Pre quest - Engage		
Pixels, black white and Gray	1. Brief on pixels 2. Brief on scanning 3. Describe interactively what makes Black, white and Gray pixels	Zoomed in images Click on a pixel and choose color	Pixel Pixel value Black, white and gray
	Learn-Explain		
Thresholding	1. What is a threshold? 2. Describe interactively effect of thresholding	Interactive threshold Choose a threshold	Threshold
	Post - Assessment		
	Learner demonstrates the effect of binarization including black, white and gray pixels	Drag and drop images with different pixels	
	Post learn – Explore		
Ocropus-nlbin	Apply Ocropus-nlbin	Run ocropus-nlbin	

Here is the lesson plan for the deskew module..

Graphical Learning Module on Optical Character Recognition			
Lesson Plan			
Module: Deskew			
Learning Objectives: Upon completion of this module the learner will be able to: 1. Distinguish the angles of the sections to the border of the page, 2. Classify the rotation needed to make sections parallel the border of the image, 3. Demonstrate the effect of binarization on the angle of the text sections of the image, 4. Apply Ocropus-nlbin to deskew the image.			
Topics	Learning Stage	Exercises / Activities	Keyword
	Pre quest - Engage		
Text sections	1. What is an angle of text section? 2. Describe interactively what makes sections of text in parallel with the border of the page	Zoomed in images Click on a angle of a section of text	Angles Alignment Skew
	Learn-Explain		
Segmentation	1. What is deskew? 2. Describe interactively effect of deskew on the angle of the text sections of the image	Interactive deskew Choose a parameter value	
	Post - Assessment		
	Learner is able to demonstrate the effect of deskew including the angle of the text sections with respect to the border of the image	Drag and drop images with text sections at different angles	
	Post learn – Explore		
Ocropus-gpageseg	Apply Ocropus-nlbin	Run ocropus-nlbin	

Here is the lesson plan for the segmentation module.

Graphical Learning Module on Optical Character Recognition			
Lesson Plan			
Module: Segmentation			
Learning Objectives: Upon completion of this module the learner will be able to:			
1. Select sections of text in the layout of a page,			
2. Analyze if each section of text from the image becomes a segment after segmentation,			
3. Describe the effect of segmentation on the text sections of the image,			
4. Apply Ocropus-gpageseg to split the image into segments.			
Topics	Learning Stage	Exercises / Activities	Keyword
	Pre quest - Engage		
Text sections	1. What is a section of text? 2. Describe interactively what makes sections of text with example images	Zoomed in images Click on a section of text	Text Borders Sections
	Learn-Explain		
Segmentation	1. What is segmentation? 2. Describe interactively effect of segmentation on the text sections of the image	Interactive segmentation Choose a parameter value	
	Post - Assessment		
	Learner is able to demonstrate the effect of segmentation including text sections and its corresponding segments	Drag and drop images, segmented and non segmented	
	Post learn – Explore		
Ocropus-gpageseg	Apply Ocropus-gpageseg	Run ocropus-gpageseg	

APPENDIX C**Lessons****Module Binarization****Pre-quest**

The pre-quest for the Binarization step assesses the discrimination behavior for this step, namely

- Given an image with a series of black, grey and white regions, the learner is able to correctly identify the black, grey and white regions within a time period of half an hour.

Pre-Learn / Engage:**Objective: Discriminate between black, white and gray pixels**

Brief on scanning. Introduction to the scanner. From document to images.
Images made of pixels.

Brief on Pixels:

<https://www.youtube.com/watch?v=m8c1CAT2zEI>

How a flatbed scanner works:

<https://www.youtube.com/watch?v=shnYIjG20i4>

Describe what makes Black, white and Gray pixels with example images

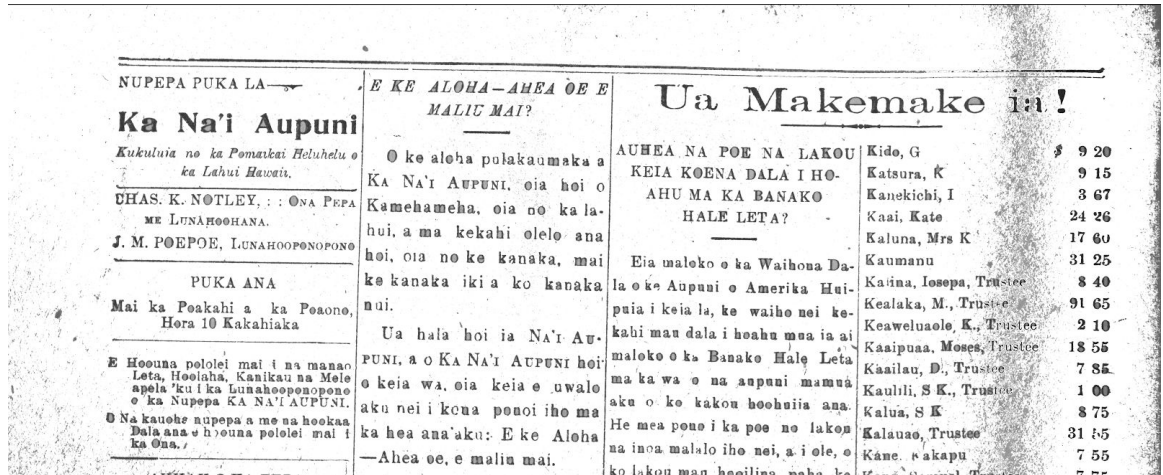


Figure: An image with black, white and gray pixels



Figure: Zooming into the word “rus” shows the pixels.

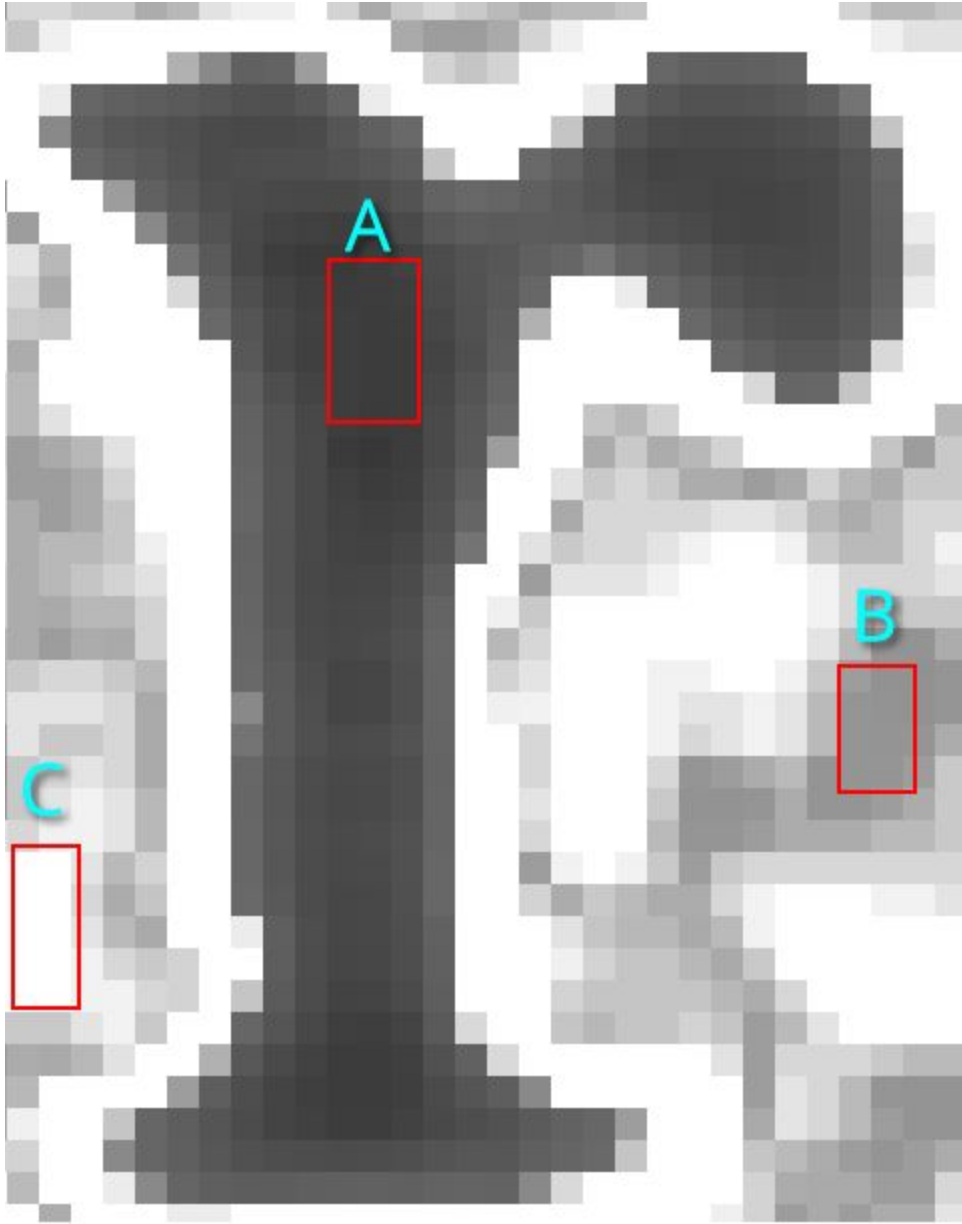


Figure: Interactive r

Q. In the figure Interactive r, which region is made up of black pixels?

1. A
2. B
3. C

Exercises

1. Zoomed in images

Q. What is your estimate for the value of the pixel in region A in the previous figure, also reproduced below?



1. 0 to 5
2. 6 to 50
3. 51 to 100
4. 100 to 255

Answer: 3. 51 to 100 (actual value is 60).

2. Click on a pixel and choose color

Learn - Explain:

Objective: Analyze the effect of binarization on black, white and gray pixels

Objective: Describe the effect of binarization including the black, white and gray pixels

What is a threshold

<https://www.youtube.com/watch?v=6pX3II2eVs0>

Describe the effect of thresholding.

1. Show snippet before and after threshold = 1
2. Show a snippet before and after threshold = 100
3. Show a snippet before and after threshold = 200

Exercises

Interactive Threshold

Choose a threshold

Post-learn / Explore: Apply Ocropus-nlbin to binarize an image

Apply Ocropus-nlbin using the command provided

Exercises

Run Ocropus-nlbin

Post-Assess

The post test for the Binarization step assesses the concept behavior for this step, namely

- Given three binarized images hidden among four other non-binarized images, the learner is able to correctly analyze the images and identify the three that do not contain grey color within a time period of ten minutes.

Module Deskew.**Pre-quest**

The pre-quest for the deskew step will assess the discrimination behavior for this step, namely,

- Given clips from magazines that are deliberately text sections at different angles with the border of the image, the learner is able to identify the top three sections that exhibit the maximum angle with the border of the image.

Learn / Engage:

Objective: Distinguish the angles of the text sections to the border of the page.

What is a angle of sections of text?

Describe interactively what makes sections of text parallel with the border of the page, with example images.

Learn - Explain:

Objective: Classify the rotation direction to make text sections parallel to the border of the image.

Objective: Describe the effect of binarization on text sections of an image.

Describe interactively what is deskew?

Describe interactively the effect of binarization on the angle of text sections of the image.

Post-learn / Explain:

Objective: Apply Ocropus-nlbin to de-skew the image.

Apply Ocropus-nlbin to deskew the image, using the command provided.

Exercises

Run Ocropus-nlbin to deskew the image.

Post-assess

The post test for the deskew step assesses the concept behavior for this step, namely,

- Given eight clips from magazines that are deliberately text sections from pages scanned at different angles with the border of the image, the learner is able to identify the skew angle of the parent image in at least six of the eight cases.

Module Segmentation.

Pre-quest

The pre-quest for the segmentation step will assess the discrimination behavior for this step, namely

- Given sections of text from a newspaper or magazine and other sections that are not, the learner is able to correctly identify the text sections in all such cases.

Pre-learn / Engage:

Objective: Select sections of text in the layout of a page

What is a section of text?

Describe interactively what makes sections of text with example images.

Learn - Explain:

Objective: Analyze if each section of text from the image becomes a segment after segmentation.

Objective: Describe the effect of segmentation on text sections of an image.

What is segmentation?

Describe interactively the effect of segmentation on the text sections of the image.

Post-learn / Explain:

Objective: Apply Ocropus-gpageseg to split the image into sections of text.

Apply Ocropus-gpageseg using the command provided.

Exercises

Run Ocropus-gpageseg.

Post-asses

The post test for the segmentation step assesses the concept behavior for this step, namely

- Given four images that has been segmented into its text sections, each a different number of sections, the learner is able to identify the text sections in the image prior to processing and the corresponding segments after the processing, for at least 95% of the segments.

APPENDIX D**Solicitation**

Kitty: Please forward to LTEC mailing lists.

My dear LTEC colleagues / alumni,

My name is Umesh Rao. I am a final semester graduate student of Learning Design and Technology. I have created or am creating a standalone asynchronous learning module on Optical Character Recognition (OCR) that is visual, graphical and implements constructivist principles. I would like to invite you to be part of the audience for my project which means you get to try out the learning module and provide feedback.

The consent form, link appended below, has details about the project and answers to questions you may have. If it seems of interest and or value to you, please fill in the consent form attached and I will add your name and email address to my audience list.
https://docs.google.com/forms/d/e/1FAIpQLSeyTCyRMg_O9TwfVCLbNsW_2YTDozbIXyqSAiNAcrwEgee3A/viewform

Mahalo,
Umesh Rao
urao@hawaii.edu

APPENDIX E**Consent Form****UNIVERSITY OF HAWAII
CONSENT FORM FOR RESEARCH PARTICIPATION**

Study Title: Visual Graphical Learning (ViGLe) Module on Optical Character Recognition

Principal Investigator: Dr. Curtis Ho

Student Researcher: Umesh Rao

I am a graduate student of the Learning Design and Technology Program at the University of Hawaii, Department of Education. I have developed or will be developing an instructional module which I invite you to take part in as a student. This form has important information about the reason for this instructional module, what we will ask you to do if you decide to be a student of this course, and the way we would like to use information about you if you choose to be a student of the course.

You may contact the UH Human Studies Program at (808) 956-5007 or uhirb@hawaii.edu. to discuss problems, concerns and questions; obtain information; or offer input with an informed individual who is unaffiliated with the specific research protocol. Please visit <http://go.hawaii.edu/jRd> for more information on your rights as a research participant.

Why is this research being done?

Scanning of text and optical character recognition is becoming more pervasive today, thanks to cell phones and images captured with them. Typically, every time such an image is received, it runs through an OCR software, the letters is recognized and stored in a file.

The objectives of the instruction module that you are consenting to be a student of are:

Given a scanned image of a page from a document, the resulting binarized and segmented images, the University of Hawaii student will be able to explain the effect of binarization and segmentation on the image including black, white and grey colors, text sections, its corresponding segments and angle of text sections with respect to the border of the image.

Given a scanned image of a page from a document, the University of Hawaii student will be able to determine at least two parameter values of the processing steps in the software called Ocopy, that will improve the final results of the processing.

What will I do if I choose to participate as a student of this course?

If you chose to participate in this learning module as a student, you will be asked to discover your way through the learning module and attempt the post-assessment tests. You will have three individual modules in the course, each module will require 20-30 minutes of your time, for a total of about 1.5 hours. You can take the course wherever you like and however many times you like. You have to note down your scores and send it to me. Here is the worksheet for that:

https://docs.google.com/spreadsheets/d/1YyFCAN_kowS3v7sLYYnI7T-7DbKyLS5_RaAqc4BD9ik/edit?usp=sharing

I may quote your remarks in presentations or articles resulting from this work. A pseudonym will be used to protect your identity, unless you specifically request that you be identified by your true name.

What are the possible risks or discomforts?

To the best of our knowledge, the things you will be doing have no more risk of harm than you would experience in everyday life. As with all research, there is a chance that confidentiality of the information we collect from you could be breached – we will take steps to minimize this risk, as discussed in more detail below in this form.

What are the possible benefits for me or others?

The possible benefits to you from this coursework includes a better understanding of the topic of Optical Character Recognition and its processing steps.

What will the researcher do with my information and the other information collected?

Results of this study may be used in publications and presentations. Your study data will be handled as confidentially as possible. If results of this study are published or presented, individual names and other personally identifiable information will not be used. To minimize the risks to confidentiality, we will keep all data regarding the coursework participants within my account, such as your demographic data and your performance record. We may share the data we collect from you for use in future research studies or with other researchers – if we share the data that we collect about you, we will remove any information that could identify you before we share it.

Will the research cost me anything?

Participation in this coursework will involve no cost to you. You will not be paid for participating in this coursework.

What are my rights as a research participant?

Participation in this study is voluntary. You do not have to answer any question you do not want to answer. You may withdraw from this study at any time, and you will not be penalized in any way for deciding to stop participation.

If you decide to withdraw from this study, the researchers will ask you if the information already collected from you can be used.

Who can I contact if I have questions or concerns about this research study?

If you have questions, you are free to ask them now. If you have questions later, you may contact the researchers at curtis@hawaii.edu or urao@hawaii.edu.

If you have any questions about your rights as a participant in this research, you can contact the following office at the University of Hawaii:

Human Studies Program
University of Hawaii
2425 Campus Road, Sinclair 1
Honolulu, HI 96822

Consent

I have read this form and the research study has been explained to me. I have been given the opportunity to ask questions and my questions have been answered. If I have additional questions, I have been told whom to contact. I understand that by filling the form below and clicking on submit, I am agreeing to participate in the research study described above.

Participant's Name

Participant's Email Address

Date

APPENDIX F**Post Survey**

Likert 5-point scale

1-Strongly Disagree, 2-Agree, 3-Neutral, 4-Disagree, 5-Strongly Disagree

Categories

a) Demographics

- Name
- Age
- Department / Program
- Relevance of OCR in your life

b) Graphical Interface

- Ease of Use; signaling, extraneous processing
 - The order of image processing is clear to me.
 - There is no information that confuses me.
- Aesthetics; looks, colors, layout

c) Website

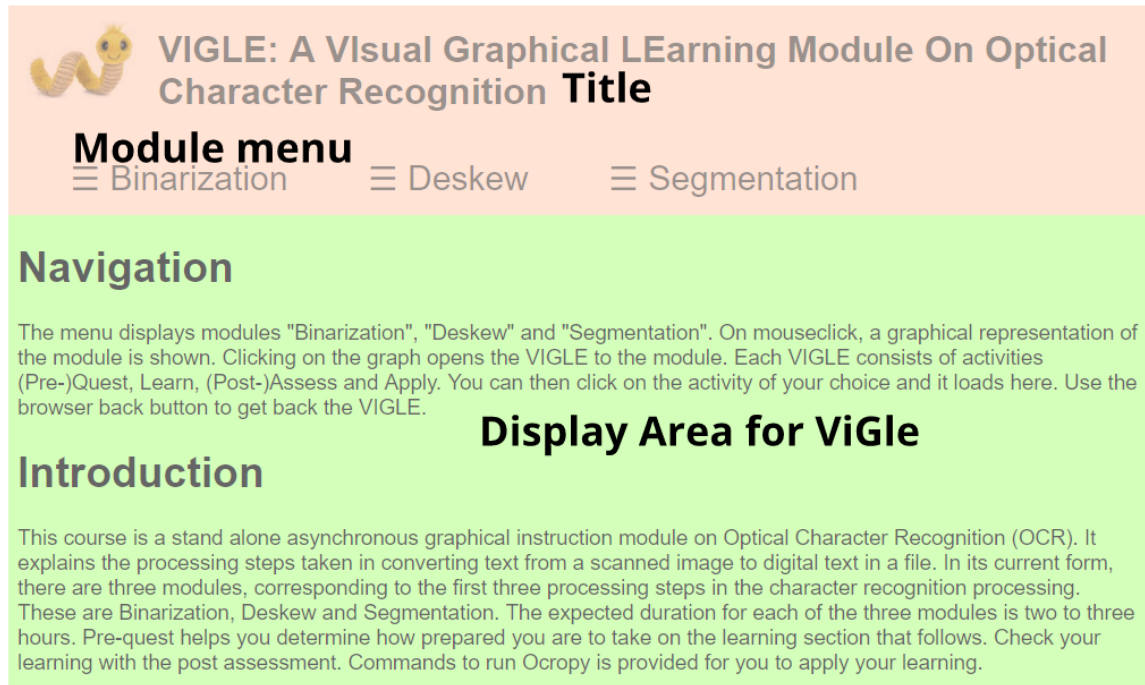
- Ease of use; navigation, extraneous processing
- Aesthetics; looks, colors, layout

d) Content

- Usefulness; helps work, helps study, helps personal life
 - helps me at work
 - helps me at study
 - helps me in my personal life
- Presentation; organization, relevance of material
 - helps understand what character recognition means
 - helps understand how digitization occurs
 - helps understand challenges of character recognition
- Format; asynchronous, Graphical
- Comments or Suggestions

APPENDIX G

Learning Module Pages



ViGLE: A Visual Graphical Learning Module On Optical Character Recognition Title

Module menu
 ≡ Binarization ≡ Deskew ≡ Segmentation

Navigation

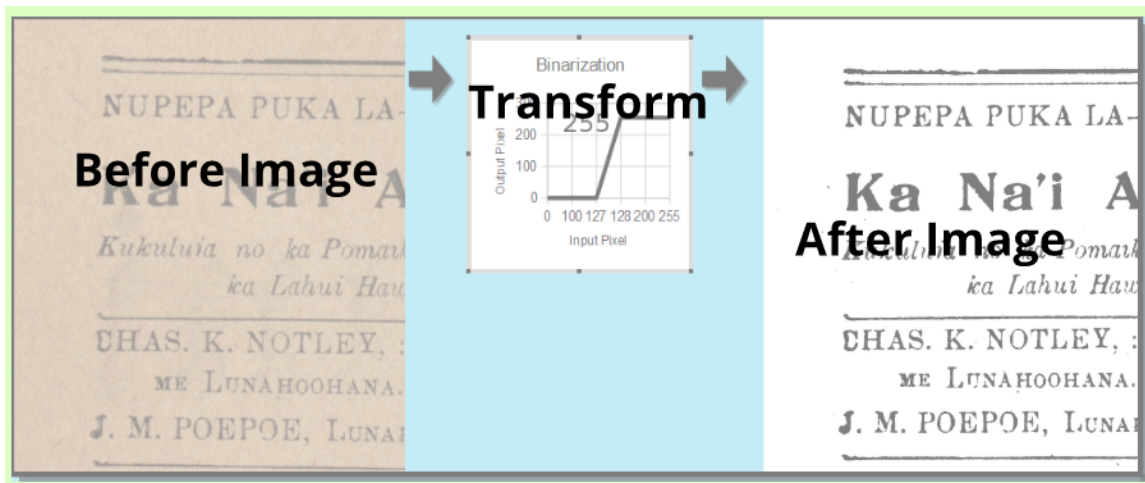
The menu displays modules "Binarization", "Deskew" and "Segmentation". On mouseclick, a graphical representation of the module is shown. Clicking on the graph opens the ViGLE to the module. Each ViGLE consists of activities (Pre-)Quest, Learn, (Post-)Assess and Apply. You can then click on the activity of your choice and it loads here. Use the browser back button to get back the ViGLE.

Display Area for ViGLE

Introduction

This course is a stand alone asynchronous graphical instruction module on Optical Character Recognition (OCR). It explains the processing steps taken in converting text from a scanned image to digital text in a file. In its current form, there are three modules, corresponding to the first three processing steps in the character recognition processing. These are Binarization, Deskew and Segmentation. The expected duration for each of the three modules is two to three hours. Pre-quest helps you determine how prepared you are to take on the learning section that follows. Check your learning with the post assessment. Commands to run Ocropy is provided for you to apply your learning.

Figure: Home page of the website



Before Image → **Transform** → **After Image**

Binarization

Input Pixel	Output Pixel
0	0
100	0
127	0
128	255
200	255
255	255

Figure: Graphical Interface to the module

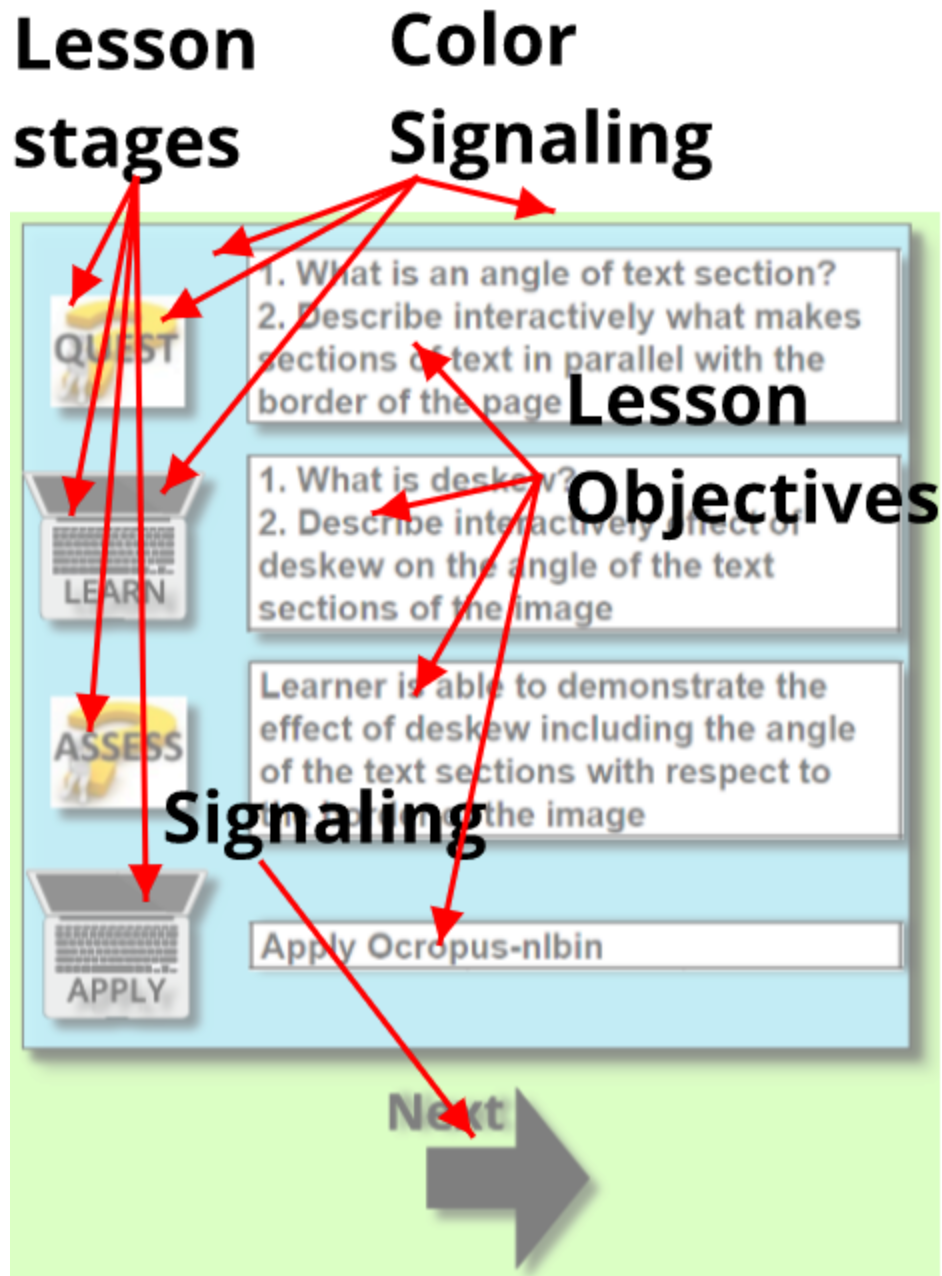


Figure: Visual Interface to the module

APPENDIX H





Color scheme

Your Color Scheme

Woah, I'm loving this.

You used the Color Calculator to select the following color(s). Print or share to save your color wheel chart as hex, RGB, and CMYK values.



HEX	RGB			CMYK			
 HEX: #bbff8e	R: 187	G: 255	B: 142	C: 27	M: 0	Y: 44	K: 0
 HEX: #ff8ebb	R: 255	G: 142	B: 187	C: 0	M: 44	Y: 27	K: 0
 HEX: #ffb08e	R: 255	G: 176	B: 142	C: 0	M: 31	Y: 44	K: 0
 HEX: #8ee8ff	R: 142	G: 232	B: 255	C: 44	M: 9	Y: 0	K: 0

About Sessions College

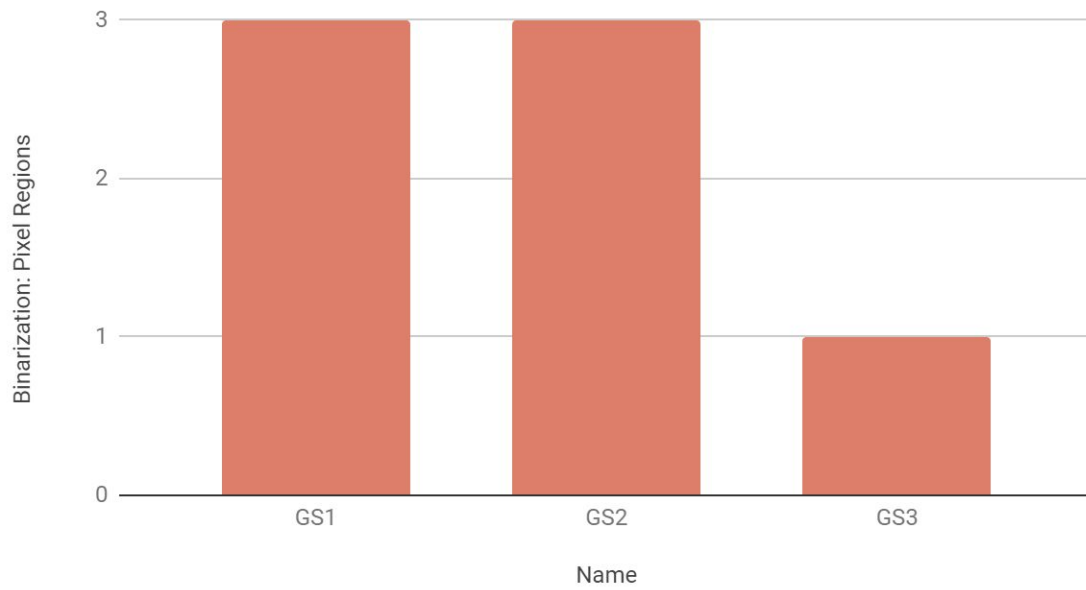
The Color Calculator was brought to you by Sessions College, a fully online school of visual arts. Flexible and affordable learning options include a range of degree and certificate programs in [Graphic Design](#), [Web Design](#), [Illustration](#), [Photography](#), and more.

APPENDIX I

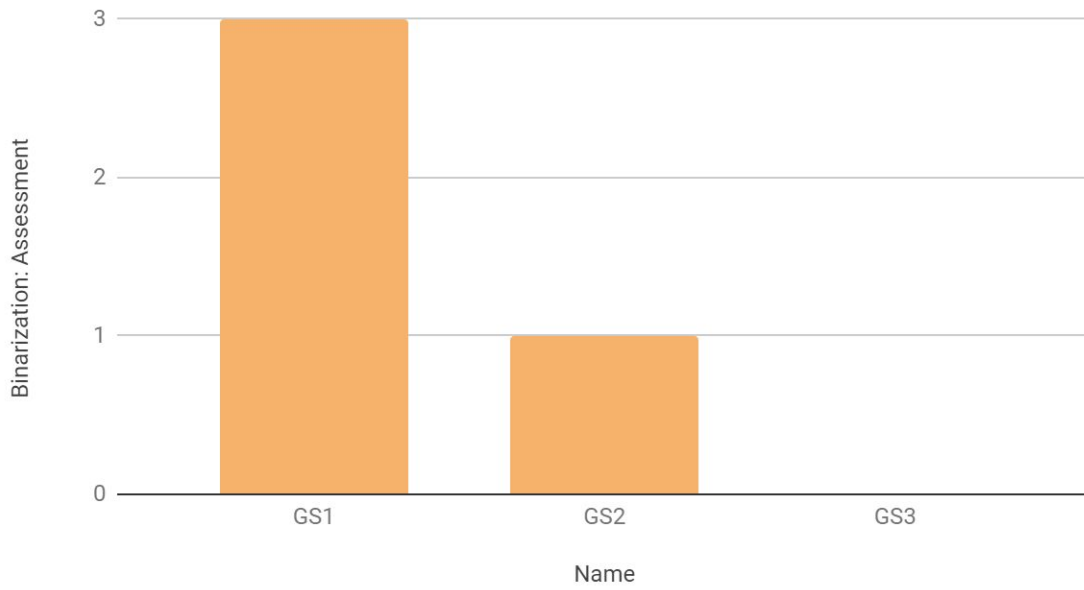
Results

Participant Scores

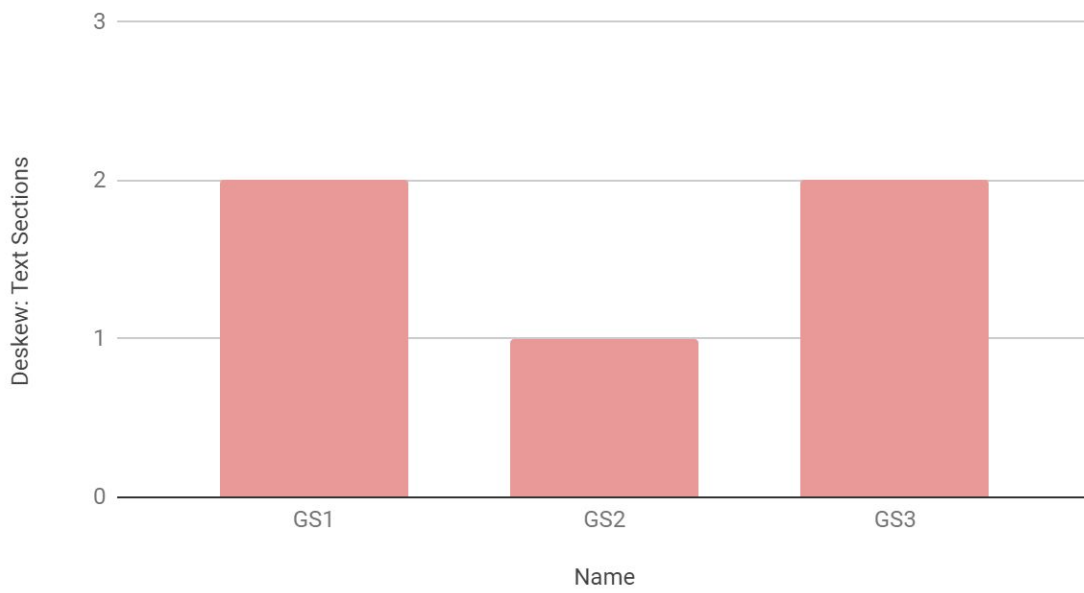
Binarization: Pixel Regions



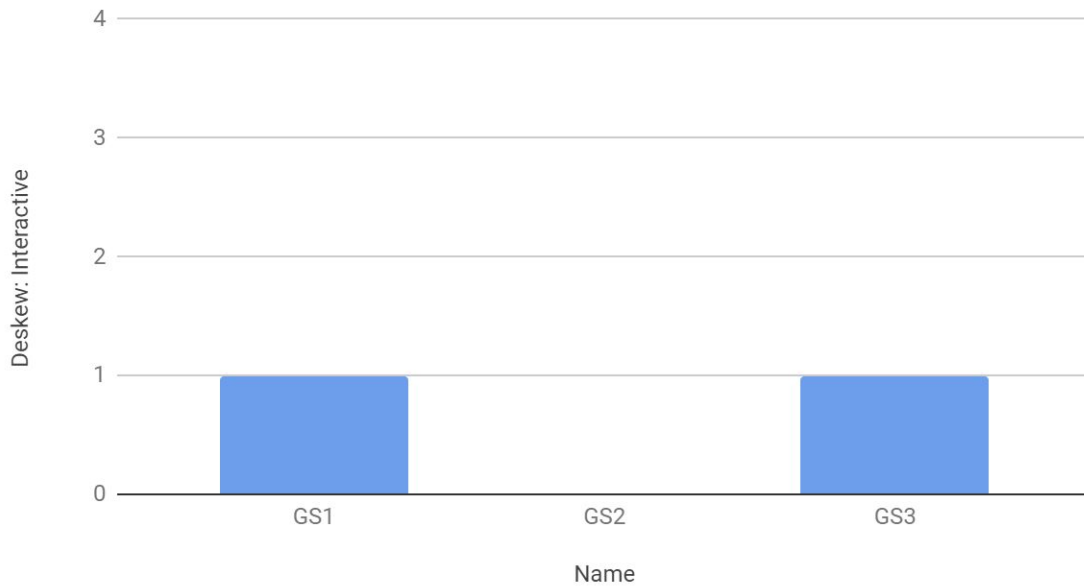
Binarization: Assessment



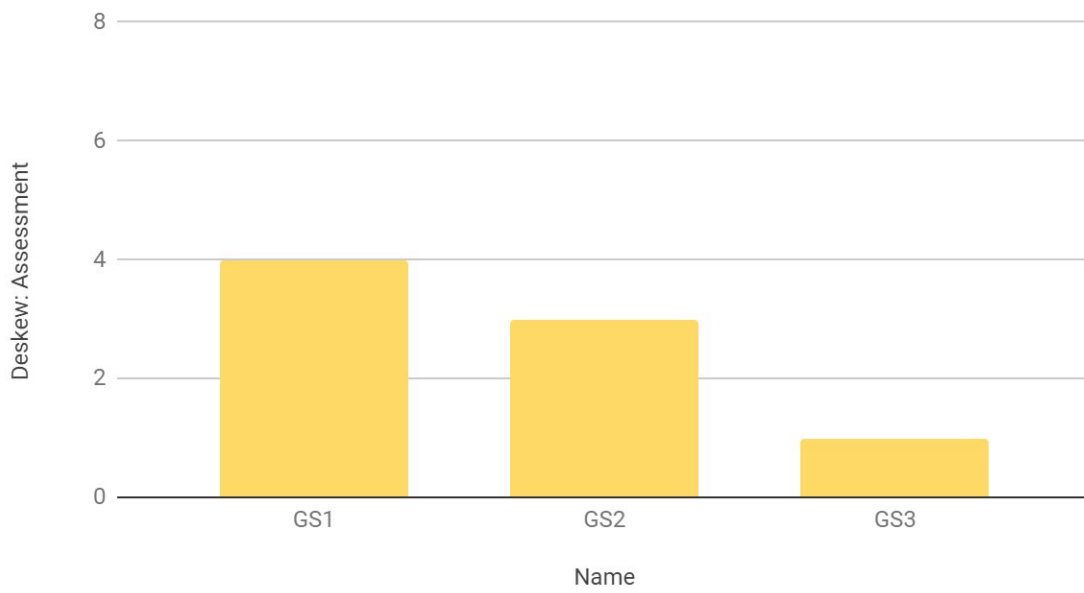
Deskew: Text Sections



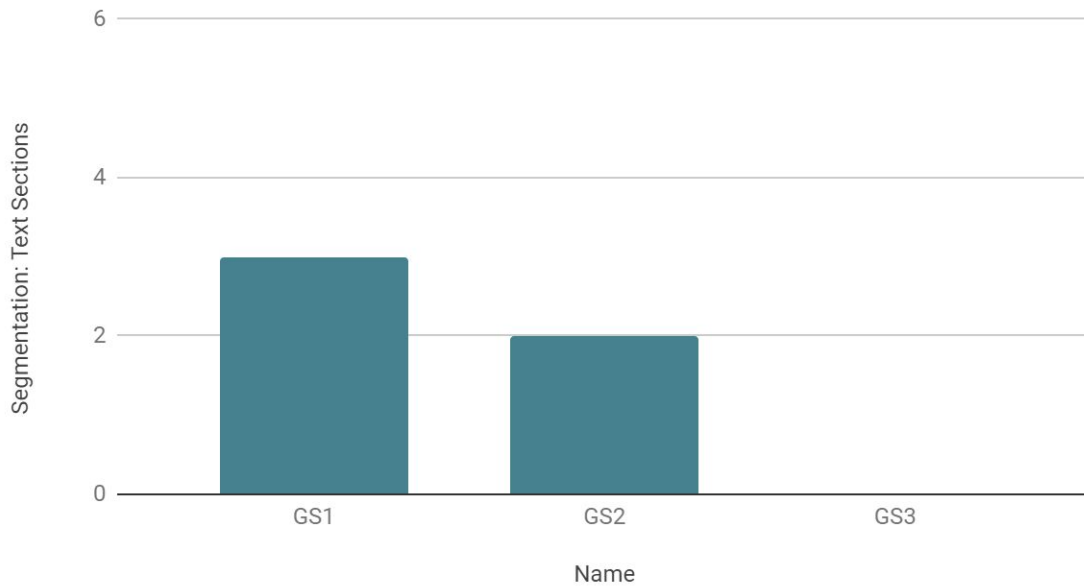
Deskew: Interactive



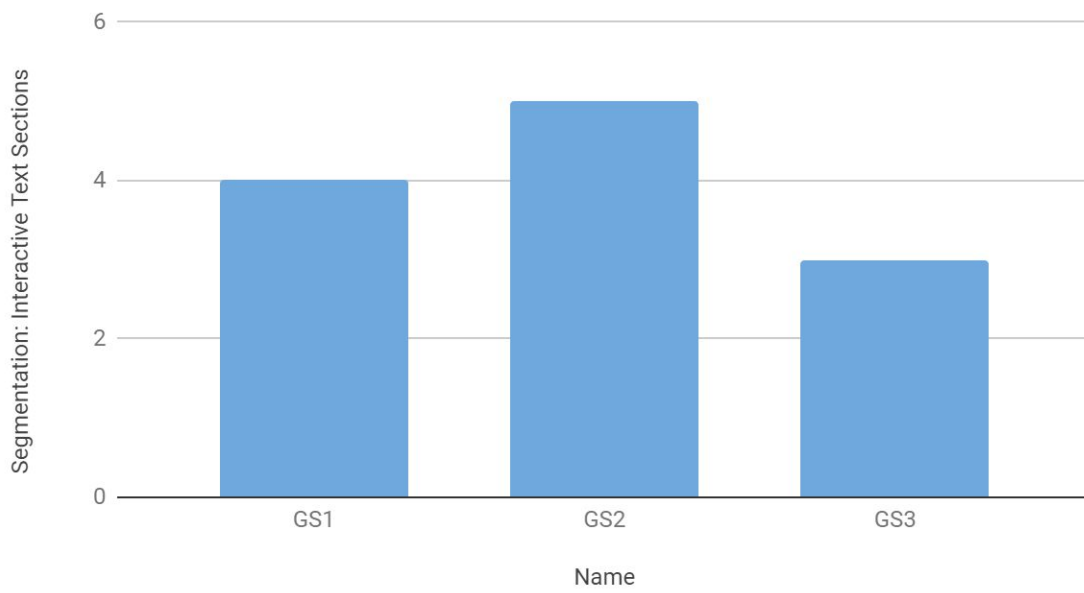
Deskew: Assessment



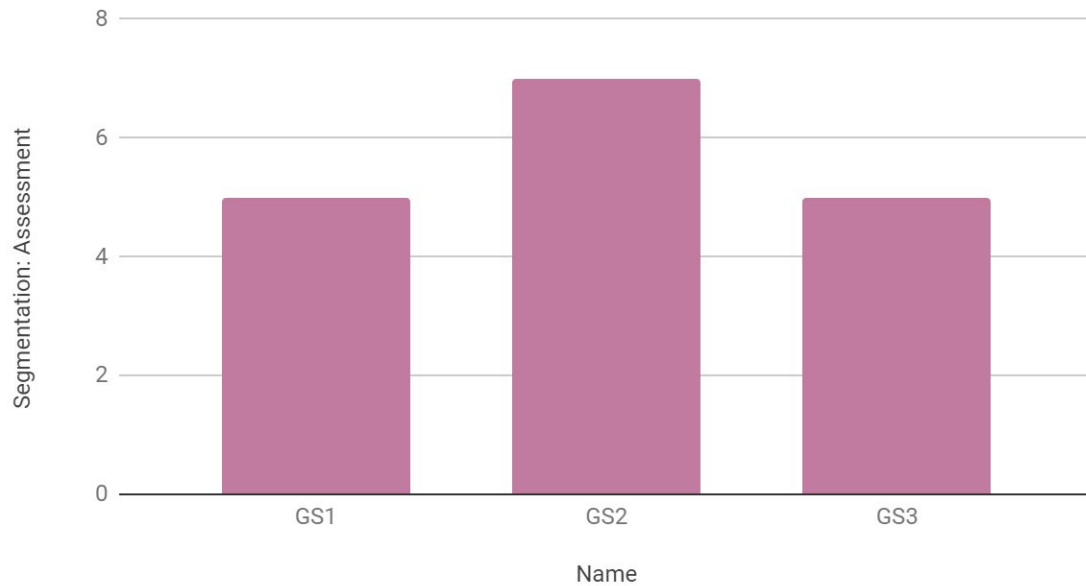
Segmentation: Text Sections



Segmentation: Interactive Text Sections



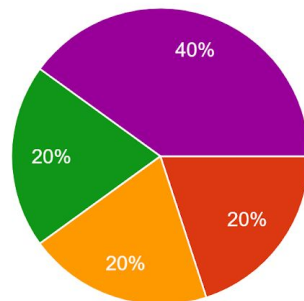
Segmentation: Assessment



Post-Survey

1. OCR for me

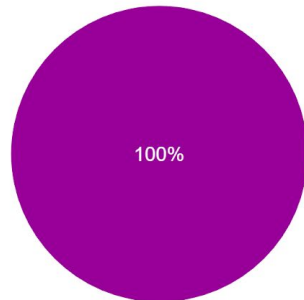
5 responses



- is very useful both in my personal and professional life
- is sometimes useful either in personal or in professional life
- could be useful the few times I scan a document
- is not at all useful, I am just learning out of interest
- is not at all useful, I am in this for the learning design or other reasons

2. I have used OCR and / or OCR software

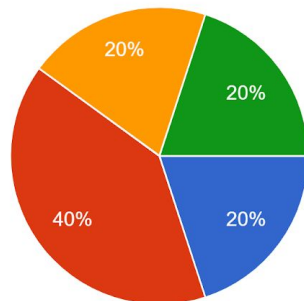
5 responses



- extensively in the past and know its concepts very well
- moderately in the past and know some basics about it
- just used it maybe and know little about how it works
- seldom in the past and am only learning about it now
- never in the past, this learning module is my introduction

3. The learning module on Binarization

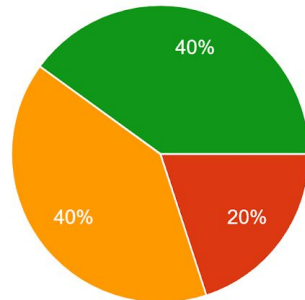
5 responses



- is very clear. e.g. I know what it does to the image and how to control its r...
- is moderately clear. e.g. I know what it does to the image, but not how to c...
- is neither clear or unclear. e.g. I can tell for some images and not for oth...
- is not clear. e.g. I think I know it, but do not know how it works
- is very confusing. e.g. Every time I see a binarized image I still wonder...

4. The learning module on Deskew

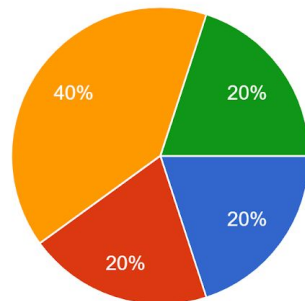
5 responses



- is very clear. e.g. I know what it does to the image and how to control its r...
- is moderately clear. e.g. I know what it does to the image, but not how to c...
- is neither clear or unclear. e.g. I can tell for some images and not for oth...
- is not clear. e.g. I think I know it, but do not know how it works
- is very confusing. e.g. Every time I see a deskewed image I still wonde...

5. The learning module on segmentation

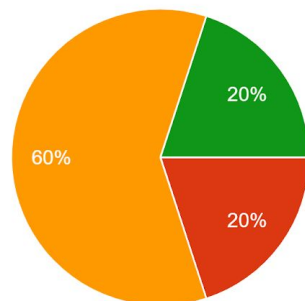
5 responses



- is very clear. e.g. I know what it does to the image and how to control its r...
- is moderately clear. e.g. I know what it does to the image, but not how to c...
- is neither clear or unclear. e.g. I can tell for some images and not for oth...
- is not clear. e.g. I think I know it, but do not know how it works
- is very confusing. e.g. Every time I see a segmented image I still wond...

9. The ViGLE learning module motivated me to apply Ocropy

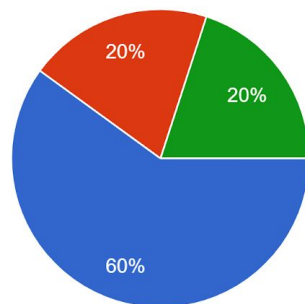
5 responses



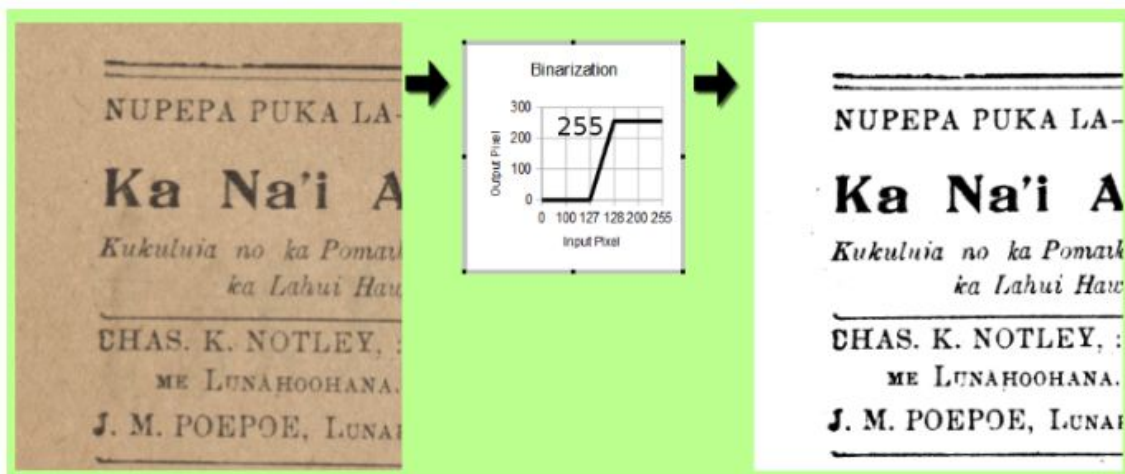
- Strongly agree. e.g. I have taken or plan to have a remote session to ru...
- Agree. e.g. I want to run Ocropy on this image I have
- Neither agree or disagree. e.g. I know I could run Ocropy if I find somethin...
- Disagree. e.g. I am confused about using Ocropy, I am not trying it
- Strongly disagree. e.g. I am confused about using Ocropy and the steps i...

6. The graphical interfaces such as the one in the figure below, for Binarization

5 responses

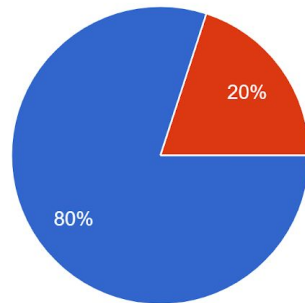


- is very helpful. e.g. it shows me clearly the transformation on the im...
- is okay to have. e.g. I know how it changes the image
- does no help or harm whether it is there or not
- is not okay to have. e.g. I do not know what it means
- should not be there. e.g. it is misleading and kept me from under...



7. The Visual interface to the learning modules, such as the one in the figure below, for the segmentation module

5 responses



- is very helpful. e.g. it shows me clearly the options, along with the o...
- is okay to have. e.g. it is good to know the options and objectives upfront I...
- does not matter. e.g. I would have felt equally at ease without it
- is not okay to have. e.g. It always confused me which link to click next
- should not be used. e.g. It is clearly the reason that kept me from learni...

The interface is a vertical light blue panel with four stages, each with an icon on the left and a text box on the right:

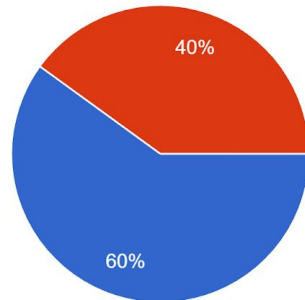
- QUEST** (yellow question mark icon):
 1. What is a section of text?
 2. Describe interactively what makes sections of text with example images
- LEARN** (laptop icon):
 1. What is segmentation?
 2. Describe interactively effect of segmentation on the text sections of the image
- ASSESS** (yellow question mark icon):

Learner is able to demonstrate the effect of segmentation including text sections and its corresponding segments
- APPLY** (laptop icon):

Apply Ocropus-gpageseg

8. The H5P interactive activities, such as the one shown in the figure below

5 responses




- is very helpful. e.g. the interactivity, multiple chances, the visual nature,...
- is okay to have. e.g. it strengthened my understanding further
- does not matter. e.g. it did not enhance my understanding or confu...
- is not okay to have. e.g. I was confused after doing some of the ac...
- should not be used. e.g. It is clearly the reason that kept me from learni...

Segmentation Assessment

Below is a presentation that you can use to assess your understanding of Segmentation and overall.

Multiple choice quiz.



Which of these options represents the image on the left?

- Column separators.
- Inverted original image.
- Image cleaned of punctuation.
- Segmented image.

✔ Check

◀ 12 / 12 ▶

APPENDIX J

Timeline

Date	Task
October 2018	<ul style="list-style-type: none"> ● Began writing detailed project plan. ● Began documenting the IRB requirements. ● Initiated solicitations for the audience. ● Collectd information required for the instruction module such as readings, technology, etc.
November 2018	<ul style="list-style-type: none"> ● Continued drafting and revising the project plan. ● Began running simulations and generating images. ● Began testing the technology for the creation of the graphical interface and the content.
December 2018	<ul style="list-style-type: none"> ● Finalized project plans for approval ● Finalized technology to use for the graphical interface and the content. ● Completed delineating the pre-quest and post-assess tests for each OCR processing step.
January 2019	<ul style="list-style-type: none"> ● Upon IRB approval, began project implementation. ● Completed populating the graphical interface with the pre-quest, post-assess and learning content. ● Sent email reminder to participants of upcoming mini coursework.
February 2019	<ul style="list-style-type: none"> ● Completed project implementation ● Sent email with course details to participants. ● Surveyed student post course completion. ● Started the final paper draft.
March 2019	<ul style="list-style-type: none"> ● Conducted student interviews ● Analyzed data ● Created TCC Presentation Slides
April 2019	<ul style="list-style-type: none"> ● Complete final paper draft ● Conduct TCC Presentation
May 2019	<ul style="list-style-type: none"> ● Complete final paper

Table 1: Instruction Design Project Goals and Timeline