Generation of Stereoscopic Interactive Learning Objects True to the Original Object

Diogo Roberto Olsen, Flávio de Almeida e Silva, Lucas Murbach Pierin, Aramis Hornung Moraes and Edson José Rodrigues Justino

Programa de Pós-graduação em Informática - PPGIa, Pontifícia Universidade Católica do Paraná - PUCPR,

Keywords: Learning Objects, Stereoscopy, Immersive Environments, Virtual Reality, Multidisciplinarity.

Abstract: Learning objects are used in many knowledge areas and may be aligned with technologies such as stereoscopy and ultra-high definition, instigating interactions and arousing interest in educational environments. Generating such objects from real pieces is a challenge because it requires computational resources to maintain quality and fidelity. Another challenge is to port these objects to different devices such as immersive theater, cell phones, virtual reality glasses, and televisions/projectors. Hence, we developed a framework capable of generating learning objects from real pieces with quality and fidelity of form, color, and texture. This article focuses on the generation of learning objects for these devices.

1 INTRODUCTION

In a teaching-learning process, the teacher can deploy elements that aim to arouse the students' interest regardless of their nature such as rocks, leaves, animals, videos, images, etc. These elements enrich the process by putting the students in contact with objects that bring learning closer to their reality.

However, not all of these elements are easily accessible to the teacher, although representations may be used to introduce them to the learners. Some of these elements may be seen as Learning Objects (LO), which, according to the Learning Technology Standards Committee (LTSC-IEEE), "Learning Objects are defined here as any entity, digital or non-digital, which can be used, re-used or referenced during technology supported learning"[1]. However, there is no consensus on this definition, as seen in (Wiley, 2002).

Interactivity, high resolution, stereoscopy, augmented reality and three-dimensional projection are present in movie theaters, advertising, televisions and cell phones. These techniques are available and may be used to support the teaching-learning process as a way of presenting and manipulating LOs. LOs need environments and repositories to be shared, as discussed in (Santiago and Raabe, 2010; Freire and Fernández-Manjón, 2016). There are LOs (Sinclair et al., 2013) that can be used multidisciplinarily, some can be customized(Méndez et al., 2016; Garrido and Onaindia, 2013). LOs can be evaluated by indicators themselves (Sanz-Rodriguez et al., 2010).

In order to create LOs compatible with these techniques, we developed a framework composed of a Full Frame Semi-spherical Scanner(F2S2) and processing software. The F2S2 digitizes, through photography, all the angles of an object, forming a hemisphere with these images; hence the term semi-spherical (Section 2). The terms full frame refer to the fact that each image represents the whole object for that viewing angle, unlike other technologies, such as the laser scanner that digitizes the object into parts to later reconstruct a given view. The LO consists of photos and navigation software, which allows for interactive visualization, i.e., you can visualize and manipulate the object from any desired angle. You can associate points of interest with hypermedia. For example, in a History class, one could pick up an archaeological object that, by its rarity, cannot be manipulated by the students. This piece can be scanned for Ultra High Definition (UHD) presentation, allowing for magnification and movement, showing details that even the naked eye would not see. Explanatory content (hypermedia) may be added at the teacher's points of interest to enrich knowledge. Figure 1 shows an example of a LO in UHD and with hypermedia.

Characteristic for LOs generated by F2S2 is that

[1]LTSC - http://grouper.ieee.org/groups/ltsc/wg12/

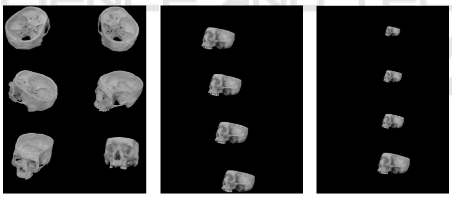
Roberto Olsen, D., de Almeida e Silva, F., Murbach Pierin, L., Hornung Moraes, A. and José Rodrigues Justino, E.
Generation of Stereoscopic Interactive Learning Objects True to the Original Object.
DOI: 10.5220/0006701402590266
In *Proceedings of the 10th International Conference on Computer Supported Education (CSEDU 2018)*, pages 259-266
ISBN: 978-989-758-291-2

(a) (b)

Figure 1: LO: (a) Original, (b) with Hypermedia content.

they can be acquired and reproduced in UHD, which, as mentioned, enables enlargement without loss of image quality, allowing the visualization of details as well as maintaining the fidelity of shape, color, and texture of the original object. LOs may also be stereoscopic, allowing visualization with the notion of depth. The stereoscopy is formed with stereoscopic lenses or by the difference of angles in the photos obtained by the F2S2, as will be seen in section 2.

Interactivity, as seen in (Koukopoulos and Koukopoulos, 2017; Hung et al., 2017), allows the object to be manipulated in three ways: scaling (as already mentioned), rotation and translation. Rotation allows you to change the viewing point to any desired angle. Translation allows a displacement of the image within the projection, for example, horizontal, vertical, and diagonal. Figure 2 shows the three components of interactivity. In this interactive process, it is possible to select points of interest and add a presentation of contents, as seen in Figure 1.



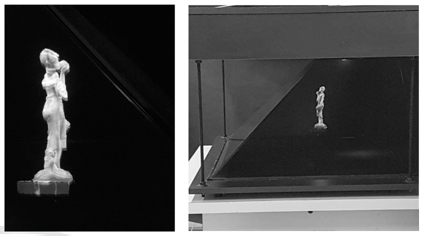
(a) (b) (c)

Figure 2: Interactivity components: (a) rotation, (b) translation, and (c) scaling.

The content may be shown in different environments, such as a television (TV) or multimedia projector, which are common access in schools and universities and bring the benefit of being stereoscopic and/or 4K, and UHD, respectively. The characteristics of these environments allow visualization with the original quality. In addition, interactivity may be provided using input devices. Figure 3 shows an example of an LO shown on TV.

An environment that is currently gaining prominence, especially in advertising, is the Holographic Pyramid (Sidharta et al., 2007; Rossi, 2015). Its operation consists of a monitor/TV/projector that projects the image on a screen, usually made of glass, causing the sensation of three-dimensionality. The name pyramid does not fit all forms of this environment because in some cases it shows only one face for visualization, in others, there are three or four faces. Figure 4 shows an LO in the holographic pyramid.

Figure 3: LO on TV.



(a) (b)

Figure 4: LO in a Holographic Pyramid: (a) Lateral Projection, (b) Pyramid.

For teaching-learning purposes using the objects generated by F2S2, the pyramid can be interactive, as is the case of the TV. Besides presenting tridimensionality, each face of the pyramid offers a different point of view.

Virtual reality (VR) glasses and mobile phones (Hassan et al., 2013) may also be used in teaching-learning since they are naturally stereoscopic as they project an image in each eye. An advantage of RV glasses are features such as the accelerometer and the gyroscope, which allow the user's movement to be identified and applied to the LO, giving a spatial notion. Figure 5 shows the RV Glasses and an LO on a mobile phone.

Figure 5: LO on VR glasses.

Another environment that offers interactive and stereoscopic features is the immersive theater, like

Fulldome (Lantz, 2007). In this immersive environment, the projection occurs through different projectors, some of which are polarized to the right eye and others to the left eye. Thus, stereoscopic images are formed with the aid of glasses. In addition, there is the possibility of interaction with controls present in each chair, so that the teacher can launch questions or challenges and the student responds directly my means of these controls. Figure 6 shows the Immersive Theater inside[2].

Figure 6: Inside the Immersive Theater.

The aim of this article is to present examples of LOs generated by F2S2 and the application possibilities in different knowledge areas. So the remaining part of the article is divided into three sections: in section 2, the methodology, in section 3, the applications, and in section 4, the conclusions.

2 METHODS

Three phases are stipulated for the generation of objects: acquisition, processing, and visualization. In the acquisition process, the F2S2 is prepared, choosing the background color, setting the camera, setting the number of Frames and Streams, and lighting. In the processing stage, lens correction, photo context reduction, segmentation, and the organization of the Streams are accomplished. Finally, the visualization stage consists of adapting the Streams to the desired projection formats and the configuration of the navigation software. Next, these three phases are explained in more detail, starting with the acquisition.

2.1 Aquisition

The F2S2 is controlled by a RaspBerry PI[3] and an Arduino[4], in addition to other electronic components and motors. A communication protocol has been developed that transmits instructions for moving and controlling the camera. The protocol is loaded into these controllers. The object to be scanned is placed in an infinite background that may be covered with different materials and colors in order to facilitate segmentation (as described in section 2.2). The camera and lighting settings are set manually as well as the background color, which depends on the object. This is the only manual step in the entire process.

Acquisitions occur by moving a camera over an arc-shaped path based on the size of the object. Figure 7 shows the semi-sphere formed by the arc during the scanning process.

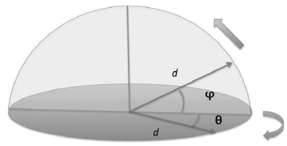
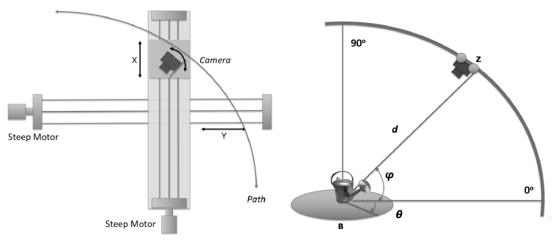


Figure 7: Semi-spherical movement.

The angular variation θ represents the object's rotation over its own axis and φ represents the camera's angle relative to the object. The variable d represents the distance between the camera and the center of the object, that is, the radius of the semi-sphere, which is constant throughout the movement.

To perform the movement, the F2S2 uses four degrees of freedom, of which two are the X and $Y - axes$, responsible for the arc displacement of the camera. These axes start from an initial position determined as angle zero (φ), traversing a path until reaching its endpoint: the angle of ninety degrees relative to the base of the object.

The $X - axis$ is responsible for the vertical movement of the arm that supports the camera. The $Y - axis$ horizontally moves the arm of the $X - axis$, keeping the camera in the arc motion at a distance d from the object. The other two degrees of freedom are the Z and $B - axes$, represented respectively by φ and θ in Figure 7. Figure 8 shows these axes.



(a) (b)

Figure 8: Axes: (a) X and Y-axes, (b) Z and B-axes.

The image acquisition process starts with all axes set to the zero-angle position. The object sits on a turntable in the center of the $B - axis$. The acquisition takes place through up to 360 shots determined by the angle (θ) of the turn of the dish, that is, one can take a

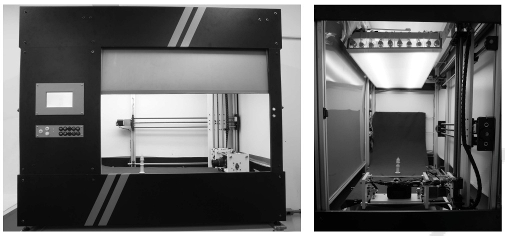
[2]Screenshot of Digital Arena http://www.ftddigitalarena.com.br/wp-content/360tour/index.html

[3]https://www.raspberrypi.org/

[4]https://www.arduino.cc/

photo for each degree. The set of photos for the same camera position is called Stream and each photo is called Frame.

When the Stream is acquired at a certain angle, the camera is repositioned one degree up in the arc, i.e., the X and $Y-axes$ are moved. In addition, to keep the object in the focal center of the camera, the $Z-axis$ is moved to an angle determined by φ. The maximum of Streams that can be acquired is ninety, i.e., a Stream for each degree in the arc. Hence, it is possible to acquire 360 Frames per Stream and ninety Streams, giving a total of 32,400 photos, representing each possible view degree by degree of the object. Figure 9 shows two views of F2S2.



(a) (b)

Figure 9: Views of F2S2: (a) lateral view, (b) internal view.

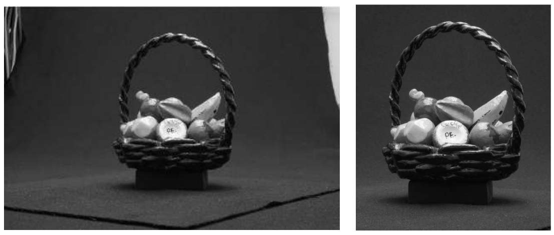
It takes about 30 hours to complete the acquisition of an object.

2.2 Processing

Camera lenses cause deformation in the acquired images, which is inherent to all lenses and cameras. Despite the little importance of it in amateur photos, in professional/scientific processes, it requires correction. There is specialized software that can perform the process automatically on a set of images from a configuration chosen by the user. An example of free software that accomplishes this job is Darktable[5]. Thus, processing starts with correcting deformities in all Frames.

Depending on the size and shape of the object, it may be impossible to frame it so that it occupies the entire space in the picture. To reduce the amount of space required for disk storage and eliminate unwanted parts, such as parts of the scanner that may appear in the photos, context reduction is applied by selecting a region where the object appears integrally in all Frames. The rest of the image is discarded. Figure 10(a) shows a complete image, and Figure 10(b) shows the same image with context reduction.

The segmentation step (Pedro et al., 2013; Chauhan et al., 2014)removes the background of the



(a) (b)

Figure 10: Context Reduction: (a) original image, (b) image with context reduction.

image, keeping only the object of interest. There is no ideal solution in the literature for the segmentation problem since each image has its own characteristics. In addition, the acquisition generates a high volume of images. Hence, the process has to be automated.

In order to bypass these problems, we have tested several algorithms, such as Support Vector Machine (SVM)(Huang et al., 2012; Tsai et al., 2006), K-means (Xu et al., 2017), K-Nearest Neighbor (KNN)(Xu and Wunsch, 2005), Sobel (Jo and Lee, 2012), Canny (Chen et al., 2014), Graph Cut (Boykov and Jolly, 2001), and Watershed (Tian and Yu, 2016). The algorithm with the best results, considering processing time and quality, was the SVM. Figure 11 shows the result of a segmentation performed on a Darth Vader doll on a blue background.

(a) (b)

Figure 11: Segmentation: (a) Image with context reduction, (b) Segmented image.

The Streams are organized during the segmentation of the images when they receive a filename containing the identification of the Stream to which they belong and the Frame; for instance, the first image of the first Stream is named S001F0001.

The equipment used for the job was an Avell Fullrange W1713 Pro CL notebook with an Intel Core I7-4910MQ 2.9GHz processor, 32GB RAM, NVIDIA Quadro K3100M and Intel HD Graphics 4600 graphics card, and Windows 10 PRO 64bits operating system. The Streams context reduction, segmentation, and organization system was developed using Matlab

[5]www.darktable.org

Toolbox [6] v.R2015b, and OpenCV-python 3.0 v.3.6.2 of Python[7].

The processing time for a set of 32,400 images was 12 hours and 30 minutes, occupying 1.2 Gigabytes of memory, which is acceptable considering that the acquisition time is longer than the processing time.

2.3 Visualization

As each device has its specificity, it is necessary to prepare the images for each of them. Stereoscopic devices have their own format for visualization, such as anaglyph, polarized and stereo pair, among others (Oh et al., 2017; Liu et al., 2017; Lei et al., 2017). The anaglyph consists of two superimposed images of the same object with different color patterns - usually red and blue or red and green - each pattern having a small displacement with respect to one another. Special eyeglasses are used to filter each pattern for one eye. The notion of depth is generated by the differences in patterns and the small displacement.

Polarized stereoscopy consists of two overlapping images, but it requires the use of a pair of glasses in which the lenses are prepared to filter the polarized light so that each eye receives each of the images. The stereoscopic pair consists of two images, one of which is horizontally polarized and the other vertically. Thus, the device emits each image in a different polarization but simultaneously. When the glasses receive these images, the lenses filter each one for each eye, causing the brain to interpret the image with a notion of depth.

F2S2 is able to acquire stereoscopic images using stereoscopic lenses developed by the manufacturers of the cameras used, but it is also capable of generating stereoscopy through lenses without this functionality. For the latter, the formation of stereoscopic vision occurs by joining the acquired images at different angles, for example, one image at angle zero and another at the angle of 4 degrees from zero. This value is related to the distance between the camera and the object, which, in turn, is related to Parallax and Disparity, as can be seen in (Kramida, 2016).

The Stereo Pair, on the other hand, is based on two images, each of which is displayed for only one eye, which has the same effect as polarized stereoscopy. The technique of the Stereo Pair, which has been known since the mid-nineteenth century, consists of placing side by side two images acquired at different angles, as cited above.

In the case of the images generated by the F2S2, they require specific processing to transform them into stereoscopic images. This processing is the junction of the images of certain angles related to the parallax and disparity, the average distance between the retinas, which varies from 4 to 7 centimeter.

As far as navigation is concerned, two ways have been developed: if the device has enough memory to support all images, they are loaded into memory. The change in visualization takes place according to the interaction of the user. If he presses the left or right arrow key, for example, the picture changes to the Frame immediately to the right or left of the current frame. This change of Frames is related to the change in the angles of axis B. Changes up or down, using the respective arrow keys, imply changes in the Streams, which relates to the angles in the arc as defined in subsection 2.1.

The same process occurs when using VR glasses, but instead of pressing keys, it reads the accelerometer and gyro signals to carry out the change of Frames or Streams. In addition, joysticks may be used, which allow interaction by pressing their buttons.

The interactions also allow zooming in and out, giving the sensation of approaching or distancing from the object, and translation, i.e., the displacement of the object in the projection.

However, if there is not enough memory to load all the images, navigation takes place through a moving windows. It works by loading a subset of the images closest to the current view. When moving within the window, which means to change Frame or Stream, the closest images are found, unloading those Frames that, according to the new position, have become more distant and loading the new closer ones.

The size of the windows, that is, the number of Frames loaded, is configurable depending on the equipment (memory and processor) used. Not all equipment will be able to run the browser, given the size and quantity of images. Navigation itself occurs in a similar way to that described in the first case.

3 APPLICATIONS

In this section, we present some possibilities of using the LOs in the Immersive Theater (Fulldome), where we used human skulls to study anatomy, the Pyramid, where we used a bronze bust to study the arts, Virtual Reality Glasses, using the inside of a car (starter motor) for studies in the area of mechanics, and the TV / Projector, where we used a replica of a car for

[6] Matlab Imaging Processing Toolbox: https://www.mathworks.com/products/image.html?s_tid=srchtitle

[7] OpenCV-Python: https://pypi.python.org/pypi/opencv-python

the study in design. This section has been divided into subsections that address each environment with its LO[8].

3.1 Immersive Theater - LO Skulls

The immersive theater used has four polarized projectors to generate stereoscopy, with each pair projecting an image for each eye. The chairs are equipped with controls that allow interaction. The exhibition room has 120 seats and seats for wheelchair users. As the projection screen format is semi-spherical and has a diameter of 14 meters, the viewers should be slightly inclined to have a full view.

Two human skulls, one with the skull box open and the other with it closed, were made available by the medical course. They were digitalized obtaining two LOs in UHD, which allows the visualization of anatomical details. The LOs present fidelity of shape, color, and texture to the original pieces. There are, of course, images of skulls created in 3D drawing and modeling programs to which a texturing has been applied; this is called rendering. In comparison to the O2 generated by F2S2, however, these rendered pieces do not have the three fidelity characteristics, which can be seen in Figure 12.

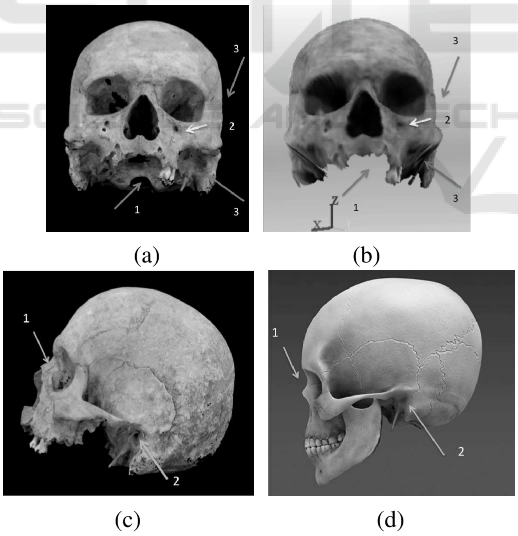


Figure 12: Comparison between Skulls: (a) Skull generated by F2S2, (b) Skull generated by 123D, (c) Skull Generated by F2S2, and (d) Synthetic Skull.

Figures 12(a) and 12(b) show the difference between an image generated by F2S2 and one generated by the AutoDesk 123D Catch[9] software. The latter presents fidelity problems in the rendering process, which does not occur in the LO generated by F2S2. Figures 14 (c) and 14 (d) show the difference between the LO generated by F2S2 and a synthetic 3D model. As can be seen in Figure 12(d), the synthetic model does not represent the ear hole.

The skull with the open skull box was used for the generation of an LO for the immersive theater, with the main objective being the anatomical study in the medical area; however, the same LO may be used in different courses (Dragon et al., 2013) such as anthropology, physiotherapy, nursing, and even in high school. The main advantage, in courses that do not require manipulating corpses in laboratories, which is unhealthy, is that the fidelity allows a detailed visualization of the piece. Figure 13 shows the preparation of the LO for immersive theater projection.

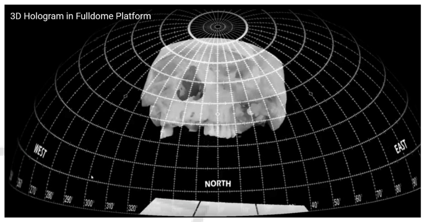


Figure 13: LO immersive theater preparation.

3.2 Pyramid - Bronze Bust

The pyramid used has three faces, made of glass for a better reflection of the object. It is 1.05 meters at the base, 48 cm high and 67 cm deep. For this experiment, we used a bronze bust. The purpose of this LO is its use in areas such as arts, history and even to create a virtual museum. This LO allows the study of bronze sculpture techniques, the analysis of the characteristics of the piece (art history), and even the study of the character represented in the piece (history).

This LO projected on the pyramid conveys the feeling that the piece is physically inside it, but it is only a visual representation. There is also the feeling of three-dimensionality. The teacher can place points of interest that he can access through keyboard, mouse or joystick interaction. Figure 14 shows the pyramid with the bronze bust.

3.3 VR Glasses - Starter Motor

Another relevant purpose for F2S2 is the generation of LOs for inspection. In this context, a starter motor, a piece that starts the rotation process of the car's engine, was used for the LO. The inspection process can also be performed on other mechanical pieces or on electronic equipment and components.

[8]We present the videos of the LOs in project's channel in youtube https://www.youtube.com/channel/UCXEFzyZGrGlLANNrCawNgWQ - Anonymous.

[9]https://www.autodesk.com.br/

Figure 14: LO Pyramid preparation.

We used RV spectacles for the presentation of this LO. The characteristic of these glasses is that they are intrinsically stereoscopic, since there is a division in the device itself and, consequently, it separates an image for each eye. The advantage is that, as it uses the accelerometer and gyroscope, the visualization of the LO is natural, that is, it does not require keyboard and mouse for interaction. One is free to move around in the environment without losing the focus of the LO.

A disadvantage is the reduction of the quality and quantity of images, since cell phones, for example, have limitations in computational capacity. Although there are fixed-screen glasses with a cable connection to a computer, this disadvantages the user's free movement, limited by the length of the cable. Figure 15 shows the use of LO in RV glasses.

Figure 15: LO on VR Glasses.

3.4 TV/Projector - Miniature Replica of a Car

The TV/projector environment is the one that has greater availability of use since schools and universities have easy access. For this experiment, a replica of a Corvette car was chosen, aiming at studying its design, which may be of interest to industrial design and design courses, for instance.

The advantage of the LO on TVs is that TVs may have a UHD resolution of 4K and 8K, allowing a detailed view of the images provided as the images themselves also have a 4K and 8K definition. Another feature of the TV is that it can be stereoscopic, using the technique of polarized stereoscopy. Figure 16 shows the LO displayed on a 4K TV.

Figure 16: LO on a 4K TV.

4 CONCLUSIONS

The use of F2S2 in educational environments instigates multidisciplinarity and interdisciplinarity since the same Learning Object may be reused in different areas of knowledge and addressed in different ways in the same area.

An example of this is the Bronze Bust, which may be used in History for the study of the character or the analysis of the artistic movement at the time in which the piece was created. It may also serve different disciplines, such as those in the arts course.

Thus, we have shown the feasibility of using the learning objects generated by F2S2, even as a way to promote interaction and dynamism in educational environments. This is mainly due to the approximation of technologies that are currently present in areas such as film and advertising. We believe that these technologies may arouse interest, since people also have contact with some of these technologies in their daily lives, outside the school environment.

The comprehensiveness of the learning objects generated by F2S2 goes beyond disciplinary boundaries and knowledge areas since the same object may be used in both the human and the exact sciences.

In the future, we plan to accomplish the scanning of other objects and make the object bases as well as the visualization software available. We also aim to evaluate the learning objects with students, in order to verify interaction and aroused interest.

ACKNOWLEDGEMENTS

We gratefully acknowledge the support of IFPR, UTFPR, CAPES and Nvidia GPU Grant Program for the donation of the TITAN XP GPU.

REFERENCES

Boykov, Y. Y. and Jolly, M. P. (2001). Interactive graph cuts for optimal boundary amp; region segmentation of objects in n-d images. In *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, volume 1, pages 105–112 vol.1.

Chauhan, A. S., Silakari, S., and Dixit, M. (2014). Image segmentation methods: A survey approach. In *2014 Fourth International Conference on Communication Systems and Network Technologies*, pages 929–933.

Chen, H., Ding, H., He, X., and Zhuang, H. (2014). Color image segmentation based on seeded region growing with canny edge detection. In *12th International Conference on Signal Processing (ICSP)*, pages 683–686.

Dragon, T., Mavrikis, M., McLaren, B. M., Harrer, A., Kynigos, C., Wegerif, R., and Yang, Y. (2013). Metafora: A web-based platform for learning to learn together in science and mathematics. *IEEE Transactions on Learning Technologies*, 6(3):197–207.

Freire, M. and Fernández-Manjón, B. (2016). Metadata for serious games in learning object repositories. *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje*, 11(2):95–100.

Garrido, A. and Onaindia, E. (2013). Assembling learning objects for personalized learning: An ai planning perspective. *IEEE Intelligent Systems*, 28(2):64–73.

Hassan, H., Martinez-Rubio, J. M., Perles, A., Capella, J. V., Dominguez, C., and Albaladejo, J. (2013). Smartphone-based industrial informatics projects and laboratories. *IEEE Transactions on Industrial Informatics*, 9(1):557–566.

Huang, G. B., Zhou, H., Ding, X., and Zhang, R. (2012). Extreme learning machine for regression and multiclass classification. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 42(2):513–529.

Hung, I.-C., Kinshuk, and Chen, N.-S. (2017). Embodied interactive video lectures for improving learning comprehension and retention. *Computers and Education*.

Jo, J. H. and Lee, S. G. (2012). Sobel mask operations using shared memory in cuda environment. In *2012 6th International Conference on New Trends in Information Science, Service Science and Data Mining (ISSDM2012)*, pages 289–292.

Koukopoulos, Z. and Koukopoulos, D. (2017). Integrating educational theories into a feasible digital environment. *Applied Computing and Informatics*.

Kramida, G. (2016). Resolving the vergence-accommodation conflict in head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics*, 22(7):1912–1931.

Lantz, E. (2007). A survey of large-scale immersive displays. In *Proceedings of the 2007 Workshop on Emerging Displays Technologies: Images and Beyond: The Future of Displays and Interacton*, EDT '07, New York, NY, USA. ACM.

Lei, J., Wu, M., Zhang, C., Wu, F., Ling, N., and Hou, C. (2017). Depth-preserving stereo image retargeting based on pixel fusion. *IEEE Transactions on Multimedia*, 19(7):1442–1453.

Liu, Z., Yang, C., Rho, S., Liu, S., and Jiang, F. (2017). Structured entropy of primitive: big data-based stereoscopic image quality assessment. *IET Image Processing*, 11(10):854–860.

Méndez, N. D. D., Morales, V. T., and Vicari, R. M. (2016). Learning object metadata mapping with learning styles as a strategy for improving usability of educational resource repositories. *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje*, 11(2):101–106.

Oh, H., Kim, J., Kim, J., Kim, T., Lee, S., and Bovik, A. C. (2017). Enhancement of visual comfort and sense of presence on stereoscopic 3d images. *IEEE Transactions on Image Processing*, 26(8):3789–3801.

Pedro, R. W. D., Nunes, F. L. S., and Machado-Lima, A. (2013). Using grammars for pattern recognition in images: A systematic review. *ACM Comput. Surv.*, 46(2):26:1–26:34.

Rossi, D. (2015). A hand-held 3d-printed box projector study for a souvenir from a mixed-reality experience. In *2015 Digital Heritage*, volume 1, pages 313–316.

Santiago, R. and Raabe, A. (2010). Architecture for learning objects sharing among learning institutions 2014 lop2p. *IEEE Transactions on Learning Technologies*, 3(2):91–95.

Sanz-Rodriguez, J., Dodero, J. M. M., and Sanchez-Alonso, S. (2010). Ranking learning objects through integration of different quality indicators. *IEEE Transactions on Learning Technologies*, 3(4):358–363.

Sidharta, R., Hiyama, A., Tanikawa, T., and Hirose, M. (2007). Volumetric display for augmented reality. In *17th International Conference on Artificial Reality and Telexistence (ICAT 2007)*, pages 55–62.

Sinclair, J., Joy, M., Yau, J. Y.-K., and Hagan, S. (2013). A practice-oriented review of learning objects. *IEEE Transactions on Learning Technologies*, 6(2):177–192.

Tian, X. and Yu, W. (2016). Color image segmentation based on watershed transform and feature clustering. In *2016 IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, pages 1830–1833.

Tsai, C.-F., McGarry, K., and Tait, J. (2006). Claire: A modular support vector image indexing and classification system. *ACM Trans. Inf. Syst.*, 24(3):353–379.

Wiley, D. A. (2002). The instructional use of learning objects: Online version. http://www.test.org/doe/ acessed in 2017/10/15.

Xu, J., Han, J., Nie, F., and Li, X. (2017). Re-weighted discriminatively embedded k -means for multi-view clustering. *IEEE Transactions on Image Processing*, 26(6):3016–3027.

Xu, R. and Wunsch, D. (2005). Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, 16(3):645–678.