Machine Learning Image Segmentation to Improve Object Recognition in Mixed Reality

by **GUILLERMO FERNANDO ESQUIVEL TABARES**

Through the implementation of an existing object recognition model along with image enhancement techniques, the current thesis aims at facilitating the navigation experience in a Mixed Reality environment by recognizing more efficiently those objects that provide users with useful information to navigate properly.

WHAT IS MIXED REALITY?

MR is thought of as the combination of the virtual and the real in a bid to yield a digital world where elements from both dimensions co-exist. Although there is no universal consensus over these terms, MR spans Augmented Reality (AR) and Virtual Reality (VR), being considered advanced techniques to help users navigate through digital models (Tadros and Franklin, 2019)

IMAGE SEGMENTATION

Image segmentation is the process of yielding clustered subregions to identify sections of interest in imagery, which are used in further image processing stages to simplify the analysis and understanding of the whole scene (Zhu et al., 2020).

Object recognition model

Object recognition is defined as a set of subtasks that provide a semantic understanding of digital images and video recordings, aiming to ascertain the shape along with the identity of elements based on known predefined labels (Yang, 2009).

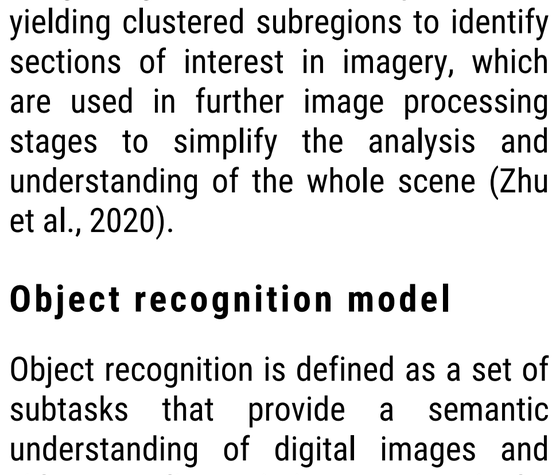

Figure 1. Instance segmentation of objects in imagery employing YOLACT (You Only Look At Coefficient).

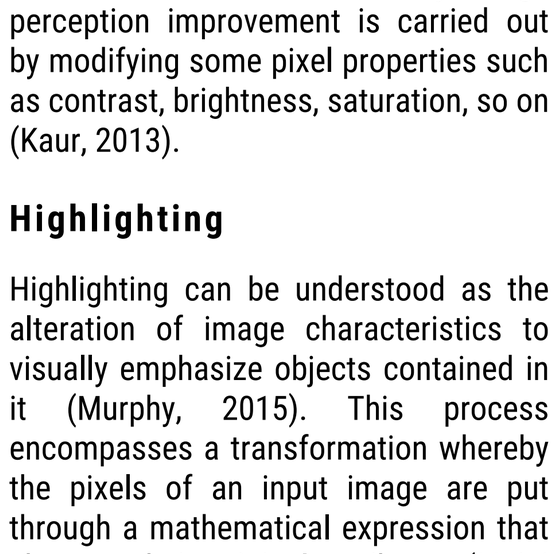

Figure 2. Techniques to highlight object features: a) detected object b) brightness enhancement c) contrast enhancement d) saturation enhancement.

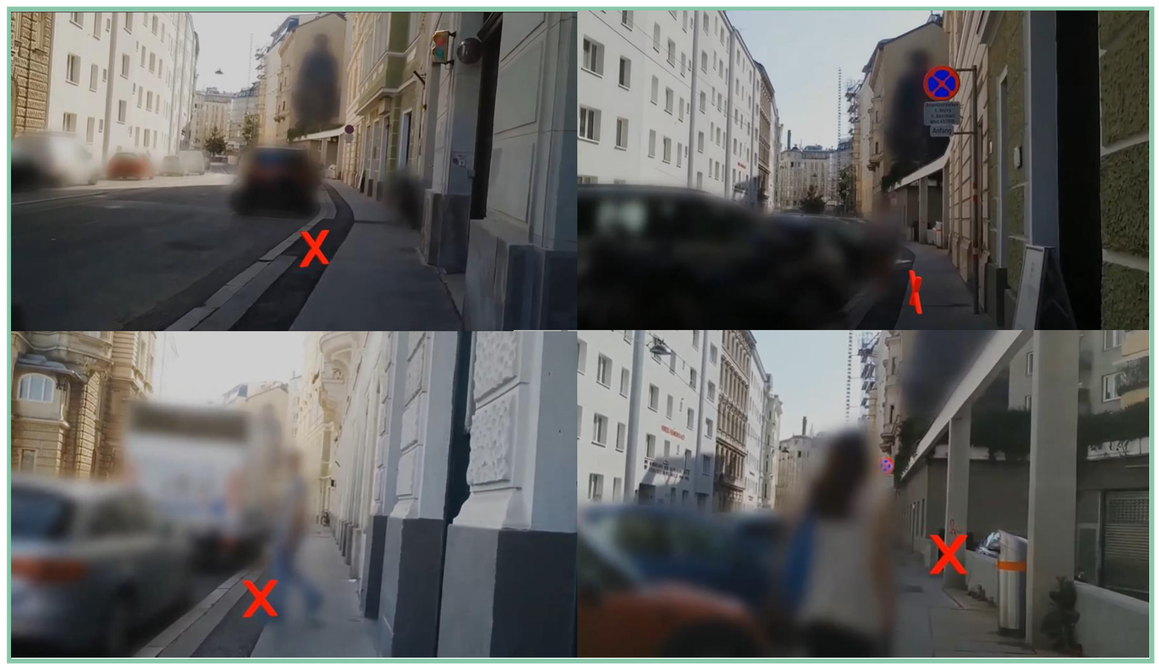

Figure 4. Integration of the object recognition model along with the highlighting and downgrading techniques. The red X-shaped rotating three-dimensional object traces the route to follow through the pedestrian area, making up a context of mixed reality.

IMAGE ENHANCEMENT

As one of the areas belonging to image processing, the image enhancement oversees the process of making certain features of interest in an image more obvious to the observer. Usually, image perception improvement is carried out by modifying some pixel properties such as contrast, brightness, saturation, so on (Kaur, 2013).

Highlighting

Highlighting can be understood as the alteration of image characteristics to visually emphasize objects contained in it (Murphy, 2015). This process encompasses a transformation whereby the pixels of an input image are put through a mathematical expression that changes their original attributes (Maini and Aggarwal, 2010).

Downgrading

Another approach to highlight portions of the image is made by reducing the role of those elements that might attract the user's attention. This operation allows redirecting concentration to parts of the image where the user needs to be focused on. In this case, the downgrading method is applied over the object to be debased, so their attributes are harder to identify and perceived as part of the background (Murphy, 2015).

Figure 3. Techniques to downgrade object features: a) detected object b) light cover c) dark cover d) blurring

IMPLEMENTATION

The object recognition model along with the image enhancement techniques were applied to a video recording of one-minute composed of 1,818 frames. Additionally, an X-shaped rotating three-dimensional object tracing the trajectory of the route was added to create the Mixed Reality context (see Figure 4).

CONCLUSIONS

This thesis aimed at facilitating the navigation experience in Mixed Reality by recognizing more efficiently those objects that provide relevant information to navigate. By conducting two user tests, it was possible to determine that the highlighted objects were striking and evident to the users. However, a significantly opposite effect did not occur when the characteristics of less relevant objects for the navigation task were de-emphasized.

THESIS CONDUCTED AT

Chair of Cartography
Department of Aerospace and Geodesy
Technische Universität München

Technische Universität München

THESIS ASSESSMENT BOARD

Chair Professor: Prof. Dr.-Ing. Liqiu Meng, Technische Universität München

Supervisor: Dr.-Ing. Christian Murphy, Technische Universität München

Reviewer: Dr. Rania Kounadi, University of Twente

YEAR

2020

KEYWORDS

Navigation, image segmentation, object recognition, machine learning, image enhancement.

REFERENCES

[1] Maini, R., and Aggarwal, H. (2010). A Comprehensive Review of Image Enhancement Techniques. Journal of Computing, 2(3), 8–13.

[2] Kaur, S. (2013). Contrast Enhancement Techniques for Images-A Visual Analysis. International Journal of Computer Applications, 64(17), 20–25.

[3] Murphy, C. (2015). Intellectual highlighting of remote sensing imagery for better image map design. Proceedings of the 27th International Cartographic Conference.

[4] Tadros, M., & Franklin, J. (2019). Navigation methods for three dimensional (3D) models in mixed reality (MR) environments. U.S. Patent No. 10,180,735. Patent and Trademark Office.

[5] Yang, M. (2009). Object Recognition. In L. Liu & M. T. Özsu (Eds.), Encyclopedia of Database Systems.

[6] Zhu, Z., Li, D., Hu, Y., Li, J., Liu, D., & Li, J. (2020). Indoor scene segmentation algorithm based on full convolutional neural network. Neural Computing and Applications.

This master thesis was created within the Cartography M.Sc. programme – proudly co-funded by the Erasmus+ Programme of the European Union.

Technische Universität München · TU WIEN TECHNISCHE UNIVERSITÄT WIEN · TECHNISCHE UNIVERSITÄT DRESDEN · ITC UNIVERSITY OF TWENTE. · Erasmus+ · cartographymaster.eu