

Augmented Reality X-Ray Vision with Gesture Interaction

P. Geethan, P. Jithin, T. Naveen, K.V. Padminy, J. Shruthi Krithika and Shriram K. Vasudevan*

Department of Computer Science and Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham University, Coimbatore, India; kv_shriram@cb.amrita.edu

Abstract

Augmented reality is a new technology which is capable of presenting possibilities that are difficult for other technologies to offer and meet. AR will really alter the way individuals view the world. Augmented reality X-Ray Vision is an emerging concept. While AR deals with virtual and real objects coexisting in the same space, AR X-Ray Vision is a subdivision of the broad spectrum of AR, which provides a “see through” vision among real world objects. In this paper, we have thoroughly analysed the existing methodologies dealing with AR X-Ray Vision and we have come up with a convenient method that enables easy implementation.

This paper deals with creating a methodology to provide an X-Ray vision using the anaglyph technique and finally integrating it with the Leap Motion Controller to enable gesture interaction to move the window around through which the point of interest can be viewed. The limitations of the suggested methodology have also been discussed. This system enables the user to perceive depth between two regions with the help of just anaglyph glasses without the use of any head mounted display devices.

Can be extended to that of medical field, where X-Ray vision is of increasing importance to view the layers of skin and bones of a patient giving the doctors and surgeons an approximate depth perception.

Keywords: Anaglyph, Augmented Reality, Depth Perception, Gesture Interaction, Leap Motion Controller, X-Ray Vision

1. Introduction

The term “augmented reality” was coined in 1992 to refer to overlaying computer-presented material on top of the real world. From early 2004, Augmented Reality⁹ has experienced the evolution of a lot of promising computer vision algorithms which allow users, with the integration of powerful 3D render engines and other interactive tools¹², to experience this Augmented World in a better way. A few years back, it would have been fit to say that Augmented Reality was the next big thing. Now, Augmented Reality is becoming as common as that of mobile phones. The capabilities of new devices in last years have contributed to the spreading of Augmented Reality, not only in scientific research but also for development and commercial purposes. Every other technology has already or is

trying to integrate itself with Augmented Reality. One of the interesting extensions of the AR system is the idea of “X-Ray Vision”. Rather, we can call it the ability to “see through” one surface (the occluding¹¹ region) to view the other surface (the occluded¹¹ region).

We see two techniques, the edge overlay and tunnel cutout¹. The edge overlay visualization provides depth information to make hidden objects appear to be behind walls, rather than floating in front of them. It provides additional depth information and spatial awareness to users, by implementing edge detection technique on the occluding region and overlaying it on the occluded region. This translucent edge overlay method gives a sense of depth perception to the users rather than making the occluded object float on top of it. The limitations in this method are that if the background is too cluttered

*Author for correspondence

then too many edges are drawn and the remote scene can be difficult to see or if the background is too plain or lacks contrast, then few edges are drawn and little sense of depth is achieved. The tunnel cutout visualization method displays a representation of obstructing layers that may exist between the user and the remote scene by creating a tunnel like effect. The limitations in this method is that if there are too many occluding regions between the user and the occluded objects, then the tunnel may extend too long and much of screen space is wasted. In², the authors have tried to implement other salient features like hue, luminosity and motion apart from edges. This provides a richer context of the occlude object but does not impede users to select objects in the occluded area. The results points out that this method needs a higher level of adapting. Terms like occlusion, binocular disparity, stipple, virtual hole and motion parallax are described in³. It also talks about different depth perception protocols like verbal estimation, forced choice and perceptual matching among others. Anaglyph 3D reconstruction from an image using depth maps is discussed in⁴. In this paper, this technique is extended to that of videos and not just images. Implementation is done for each and every frame in the video captured. Recent trends and advances are covered in^{5,16} covering head mounted or see through displays, projection displays, visualization and perceptual problems. How head mounted displays are used to achieve Superman-like X-Ray vision are covered in^{10,15,17} and a survey of AR and the importance of Medical AR can be understood from^{10,14}. An evaluation of the Leap Motion Controller, its suitability for static and dynamic tracking is done in^{6,13}. This would give an understanding on the precision and reliability of the Leap Motion Controller and why it has been chosen for implementation of gesture interaction in this project.

2. Problem Presentation

Our view of the real world need not be limited by what our naked eye can see. The physical world can be enhanced to enjoy the benefits of digital media, by just augmenting the physical world with it. The potential for Augmented Reality X-Ray vision is undoubtedly endless. In this paper, we are creating a 3D based reconstruction of a remote scene based on the captured video feed, providing X-ray and depth perception⁸ for each of the frames. Finally, incorporating gesture interaction with this AR system, so that the user experiences a better visual perception.

3. Proposed Methodology

Our approach involves the following steps:

3.1 Initial Setup

Live video feeds of the occluding region and the occluded region are captured by placing the cameras at appropriate distances in front of these regions. The system design is shown in (Figure 1).

3.2 Merging the Video Feeds

Video Feeds(Figure 2) shows how the two videos are merged, with the larger frame denoting the occluding region and the smaller window surrounded by a black border, displaying the occluded region. The video feed of the occluding region is obtained and the window size to view the occluded region is estimated. The pixels of the occluding region, enclosed by the window with coordinates: (x,y) , $(x+w,y)$, $(x+w,y+w)$, $(x,y+w)$, are set to white, by initializing the RGB components to 255, where (x,y) denotes the top left corner of the window and w is the width of the window. A similar operation is performed with the video feed of the occluded region as well. Here, the pixels of the occluded region excluding the ones that are enclosed by the window are set to white. Now, an AND operation is performed between the pixels of the two video feeds for each frame. This will result in an overlaid visualization of the two videos as shown in the (Figure 2). For adding depth perception, the following method is used.

3.3 Implementing Anaglyph

As discussed, the anaglyph methodology has been applied to every frame of the video feed. A copy of each frame is made. (Figure 3) shows the red filter applied to one copy

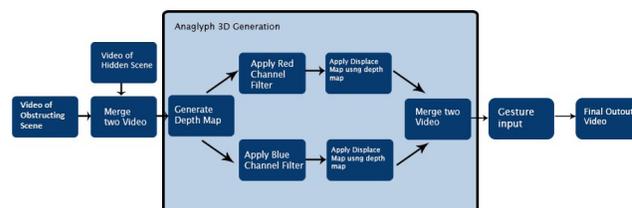


Figure 1. The system design: shows the extraction of video feed from occluding and occluded region, implementation of anaglyph and gesture interaction, final output.

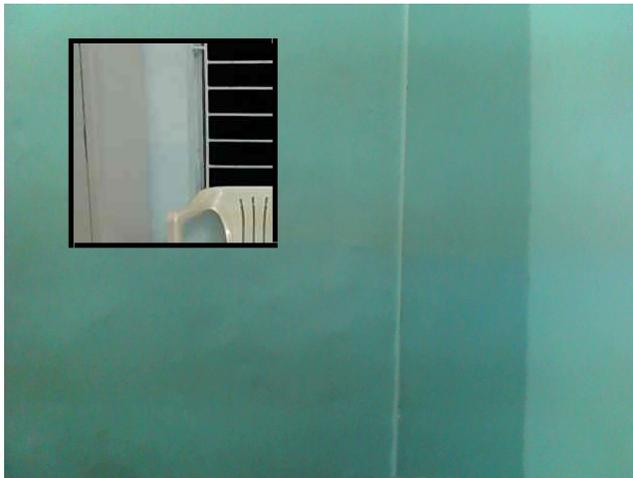


Figure 2. Frame at time t , of merged video feeds from the occluding and the occluded regions.



Figure 3. Red filter applied to both occluding and occluded regions.

of the frame. This is obtained by considering only the 'R' component of the RGB values of every pixel in the frame. Similarly the blue filter (Figure 4) is applied to the other copy of the frame by obtaining the 'B' component for each pixel leaving out the 'R' and 'G' component. These two frames with channel filters applied, are superimposed, while making use of the depth map⁷ and the displacement map with corresponding offsets. The super imposed frame is shown in (Figure 5).

3.4 Gesture Interaction to Move the Window

(Figure 6) shows the window being moved to a different position, with the corresponding view of the occluded



Figure 4. Blue filter applied to both occluding and the occluded regions.



Figure 5. Final anaglyph generated using depth and displacement maps, at a particular time t .



Figure 6. Anaglyph generated for the same frame at a different window position.

region at that position. This is achieved by integrating the system with the Leap Motion Controller. This enables gesture interaction with the system, whereby one can move the window to the desired location within the frame. An arbitrary gesture can be used. A ‘forward pointed fore-finger’, has been used in this case, to move the window around.

4. Results and Analysis

Initially, we made the whole system run as single threaded. This resulted in crashing of the system within seconds of running it as it included processing every pixel in each frame of both the video feeds. Later improvisations were done to avoid this glitch by running the system in a multi-threaded environment. Each task was run on a separate thread. This helped in improving the efficiency and effectiveness of the AR X-Ray system. Adopting anaglyph methodology posed to be an easy alternative over the other existing methods like the Edge overlay and the Tunnel cut out, but didn't seem very effective in providing accurate depth perception of the occluded region. Anaglyph methodology provides a stereoscopic vision but does not give the exact depth perception for the user. However, implementing this method is a faster and a simple way to render video formats while considering the complexity of the system. Using Leap Motion posed to be an advantage because of its high precision and sensitivity, giving the user the feel of looking through one region from another with the help of a rectangular hole made in the occluded region, while enabling them to move the window around just by using the finger tips.

5. Conclusion

This paper explains the implementation of anaglyph methodology for depth perception. While there are other methods existing for augmented reality x-ray vision that are more accurate, those have been mostly theories or methods that are complex to implement. The method that we have suggested is easier to implement and is simple. But this does not provide accurate depth perception. On analyzing the effectiveness of this system, we have come to a conclusion that it has to be improved in its qualitative aspects. Nevertheless, this system is not affected by any external factors like luminosity, as these factors do not have an effect over calculating the depth map.

All it requires is for the video feed to be of considerable quality, as that would directly determine the quality of the anaglyph generated. This method with further improvisations can be extended to the medical field where it can be used to improve the visual saliency⁸ for the doctors.

6. References

1. Avery B, Sandor C, Thomas BH. Improving spatial perception for augmented reality X-Ray vision. Proceedings IEEE Virtual Reality International Conference. Louisiana. 2009; p. 79–82.
2. Sandor C, Cunningham A, Dey A, Mattila VV. An augmented reality X-Ray system based on visual saliency. IEEE International Symposium on Mixed and Augmented Reality (ISMAR). 2010 Oct; p. 27–36.
3. Livingston MA, Dey A, Sandor C, Thomas BH. Human factors in augmented reality environments: pursuit of X-ray vision for augmented reality. Springer; 2012.
4. Niloufar M, Ng KW, Mahdi B. Depth-based 3D anaglyph image modeling. Internat J Scient Knowled. (IJSK). 2014 Mar; 4:9–14.
5. Azuma R, Baillet Y, Behringer B, Feiner S, Julier S, MacIntyre B. Recent advances in augmented reality. IEEE Comput Graph Appl Mag. 2001 Nov; 21:34–47.
6. Weichert F, Bachmann D, Rudak B, Fisseler D. Analysis of the accuracy and robustness of the leap motion controller. 2013, May; PubMed Central Archive. [Internet]. 13(5):6380–93. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3690061/>
7. Saxena A, Chung SH, Ng AY. 3-d depth reconstruction from a single still image. Internat J Comput Vision (IJCV). 2007 Jun; 76:53–69.
8. Mendez E. On the usage of context for augmented reality visualization [PhD dissertation], Department Comp Sci, Graz Univ; Graz, AU: 2010.
9. Silva R, Oliveira JC, Giraldo GA. Introduction to augmented reality. LNCC Nat Lab for Scientific Computation; 2003. Rio de Janeiro: Br, Tech Rep. ISSN: 0101 6113.
10. Blum T, Stauder R, Euler E, Navab N. Superman-like X-ray vision: towards brain-computer interfaces for medical augmented reality. Proceeding IEEE International Symposium on Mixed and Augmented Reality, Atlanta. 2012; p. 271–72.
11. Furmanski C, Azuma R, Daily M. Augmented-reality visualizations guided by cognition: perceptual heuristics for combining visible and obscured information. Proceedings IEEE International Symposium on Mixed and Augmented Reality, Darmstadt, De: 2002; p. 215–320.
12. Miranda FR, Tori R, Bueno CES, Trias LP. Designing and implementing an spatial augmented reality X-Ray. RITA.

- The J Theoret Appl Informat. 2008; [Internet] 15(3):47–74. Available from: http://www.seer.ufgrs.br/rita/article/view/rita_v15_n3_p47-74
13. Guna J, Jakus G, Pogacnik M, Tomazic S, Sodnik J. An analysis of the precision and reliability of the leap motion sensor and its suitability for static and dynamic tracking. PubMed Central Archive. 2014 Feb; [Internet]. 14(2): 3702–20. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3958287/>
 14. Azuma RT. A survey of augmented reality. IEEE Comput Graph Appl Mag. 2001 Dec; 21(6):34–47.
 15. Lin C, Tseng H, Lay Y, Chen C, Yeh M. Design and reliability analysis of image type head gesture interface in distributed control system. IndJST. 2012 Aug; 5(8):3150–56.
 16. Lee D, Shon J, Kim Y. Design and implementation of OSMD based learning management system for mobile learning. Indian J Sci technol (INDJST). 2015 Jan; 8(S1):154–60.
 17. Baghyasree T, Janakiraman K, Parkavi D. Efficient remote video monitoring for industrial environment based on LPC2148 and GSM. Indian J Sci technol (INDJST). 2014 Sep; 7(9):1333–41.