Deep Visual Informatory Through Super-Resolution High Dynamic Range Imagery in Immersive Applications

Muhammad Jafni Jusof, Nik Nabilah Izzati Juhari and Nurul Nasmira Shafie

December 17, 2021

Deep Visual Informatory Through Super-Resolution High Dynamic Range Imagery In Immersive Applications

Muhammad Jafni Jusof[1]
Nik Nabilah Izzati Juhari[1]
Nurul Nasmira Binti Shafie[1]

[1]*Faculty of Computer, Media and Technology Management, TATIUC, 24000, Kemaman, Terengganu, Malaysia, Email:* muhammadjafni@uctati.edu.my, niknabilah@uctati.edu.my, nurulnasmira@uctati.edu.my

Abstract. Photography is probably the most common way of capturing light using a device called a camera; creating either artistic or scientific images, be it for personal or general use. Photographing has become a phenomenon due to its capacity to represent the reality that provides a new dimension of experience. This paper explores the use of the camera and photography techniques as to how it could be utilized to capture images beyond the capabilities of normal human vision. The super-resolution image is essential in revealing visual details of a captured scene while high dynamic range imaging deepens the details within details. A combination of super-resolution images and high dynamic range imaging should unveil a huge amount of information that could be analyzed or interpreted in many different ways. The process involves capturing a scene, construction by computational stitching, and the final communication output. While there are unlimited possibilities of applications from using the same technique to achieve super-resolution high dynamic range images, the focus of this study is towards immersive visualization purposes through the use of photographic equipment and techniques.

Keywords: Super-Resolution, High Dynamic Range, Immersive Application

1. Introduction

In its organic form, the eye is a fascinating creation itself that allows visual sensation to be experienced in an immersive space. The visual sense is the main and most used way to stimulate other senses without having to experience it in situ; feeding information to the brain that stimulates awareness and appreciation [1]. It permits multiple interpretations, construction of ideas and thoughts, but not without limitations. In that sense, related technology has come a long way in which seeing through artificial optical lenses could overcome the limitations up to an extent if not completely. Changing the lenses could simply alter a person's vision to experience a radical shift in perception; coupled with advanced electronic processing, opening new possibilities in seeing through a vibrant world. Human sight can be extended through a simple electronic imaging device such as a camera encompassing the lens, image sensor, and processor which is used in various applications throughout our daily lives to capture events in electronically generated images. It preserves time in a form that can be re-interpreted, analyzed, or revisited in the future. Visual records are an important part of observing, interacting, expressing, and most probably remembering [2]. The flexibility in digital photographic images in which can be easily manipulated, enhanced, stored, and distributed without being deteriorated over time leading towards a superior choice for visual preservation purposes. Therefore, the quality of the resulting image is necessary to make it viable for its purpose. While qualitatively image quality evaluation should differ subjectively between areas, a more common quantitative approach is preferable. Resolution is acknowledged as a common measure of image quality and has been accepted in many imaging-related fields [3]. Achieving higher and higher resolution is perceived to be the goal to cramp as much information as possible into a piece of imagery. The multimedia domain sees more of a need for super-resolution imagery especially in the immersive visualization application where other domains such as satellite imaging, microscopy, archeology, or astrology could benefit from in the form of visual content presentation [4][5][6]. This paper explores the method of using multiple low-resolution images which is more feasible to produce a higher resolution output, although there are other methods available that have been researched and published. Stitching multiple narrow-angle images to produce a greater image resolution exceeding the pixel counts of the sensor itself [7] could be the most basic and cheapest way to execute, and stacking images of the same composition with different exposure produces an even wider dynamic range of an image [8] for the use of immersive applications.

2. Super-Resolution Imagery

The human eye is made up of a complex structural construct of various cells and nerves [9] that allow visual information to be received and interpreted by the brain. It is often implied and compared to a camera which includes the main components of the lens, photoreceptors as the image sensor, and the brain as the processor [10]. It is hard to determine the exact resolution of human eyes even though numerous research and articles discussing this matter due to the nature of the eye itself in which it behaves more like a video camera instead of a photo camera. Visual information is being fed continuously through streams of frames per second at a certain field of view as the eyeball rotates around navigating a space. Given an approximately 17mm of focal length per eye, it was estimated that human eyes cover about 576 megapixels [11], other studies stated 130 megapixels; however, only the central cone receptors (about 6 megapixels) process the main image in focus (visual attention) within 55-degree angle, while the remaining is peripheral vision [10]. Although the numbers should vary, it considerably exceeds the highest resolution of the average consumer or professional camera to date.

The idea of super-resolution image reconstruction is to pass through a series of low-resolution images to produce a higher pixel density image [12]. With super-resolution, a low-resolution image is converted into a high-resolution image via complex algorithms and computational methods either using a single image that recovers information within a single frame [13] or multiple image-based systems which attain featured information from multiple frames [14]. However, these methods do have their limitation constraint by the lack of quality information needed for the process [15] and it is difficult to introduce small details which never been seen by the naked eye [16]. Increasing the number of sensors or even the number of pixels on an image sensor is not an easy task, instead increasing the number of images covering a smaller area of a larger intended scene will be more economical. Stitching multiple narrow-angle images to form the larger field of view is equivalent to how human eyes perceived information increasing the pixel density of the output image. The longer focal length of the lens equals a narrower field of view of the image on the sensor.

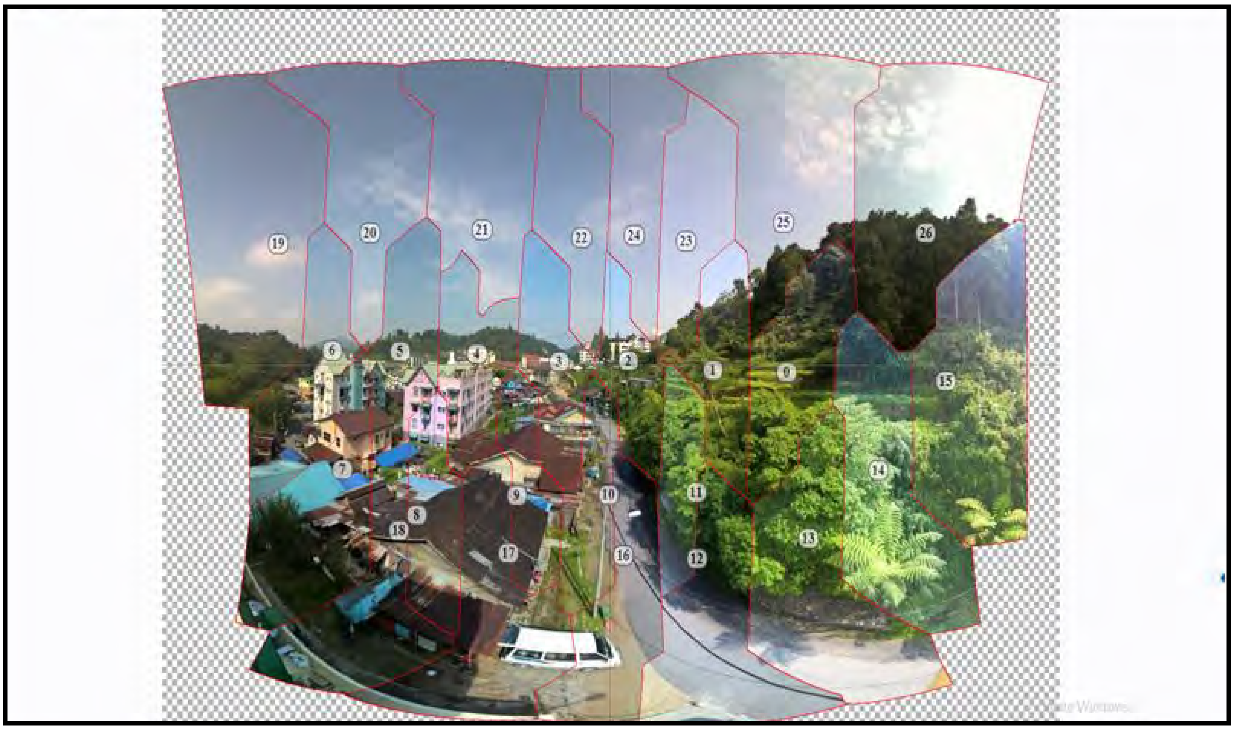

Figure 1. Manual stitching of randomly shot images.

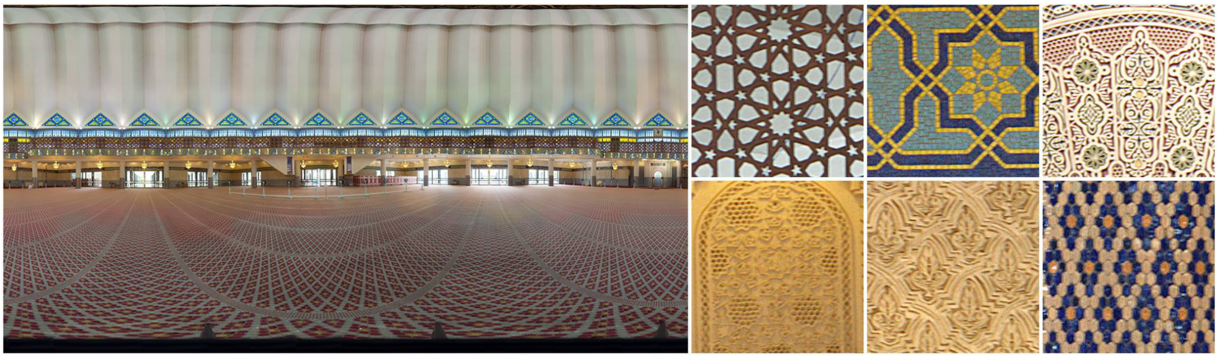

Figure 2. Smaller samples (right) taken from a super-resolution panoramic image.

Figure 3. An optical super-resolution photoscan of a Pua Kumbu cloth.

3. High Dynamic Range Imaging

Dynamic range refers to the margin value between the brightest bright and the darkest dark of an image; wider means more details can be captured or seen [2]. Normal human eyes are far better than currently, mass-produced image sensors. Average human eyes can adjust at least up to a million to one ratio, about 20EV (exposure value) [17] if not more. Although the high dynamic range (HDR) technique has existed for a long time, its purpose remains the same which is to compensate for the deficiency of the camera metering system in providing accurate exposure readings. This inadequacy results in compromise EV only for the subject in focus, thus producing a low dynamic range image, hiding detailed information within the brightest and darkest area of an image [17][18][19]. HDR imaging significantly broadens the capabilities by increasing the overall image quality and making it look closer to the original. If an image can represent a scene in a wide range of light, it should be close to what normal human eyes can see, if not exceeding it [20][21]. In some areas, such as heritage documentation, this is very crucial to have a record that is as close as possible to the original counterpart [22][23]. The quality of the visual content presented to the audience will determine the level of believability of realness perceived by the viewer [23], and thus, with careful adjustment and enhancement HDR can render images close to what the eyes can see.

A typical working principle is that a minimum of 3 exposures is required, consisting of under-exposed, normally-exposed and over-exposed images in order to stack them into one HDR image. The most under-exposed image should be fairly dark while the most over-exposed image should be very bright [24]. However, capturing multiple images of the same composition could possess some imperfection especially when involving moving subjects that lead to ghosting effects. Fortunately, as technology evolves, image sensors not only pack with more and more pixels, but the sensor itself increases in dynamic range, making it possible to capture just one image in RAW file format. The "one" file then goes into a process of exposure extraction; creating the same copies but with different exposure values [25]. These files can later be merged into an HDR image, eliminating defects from moving elements within a scene.

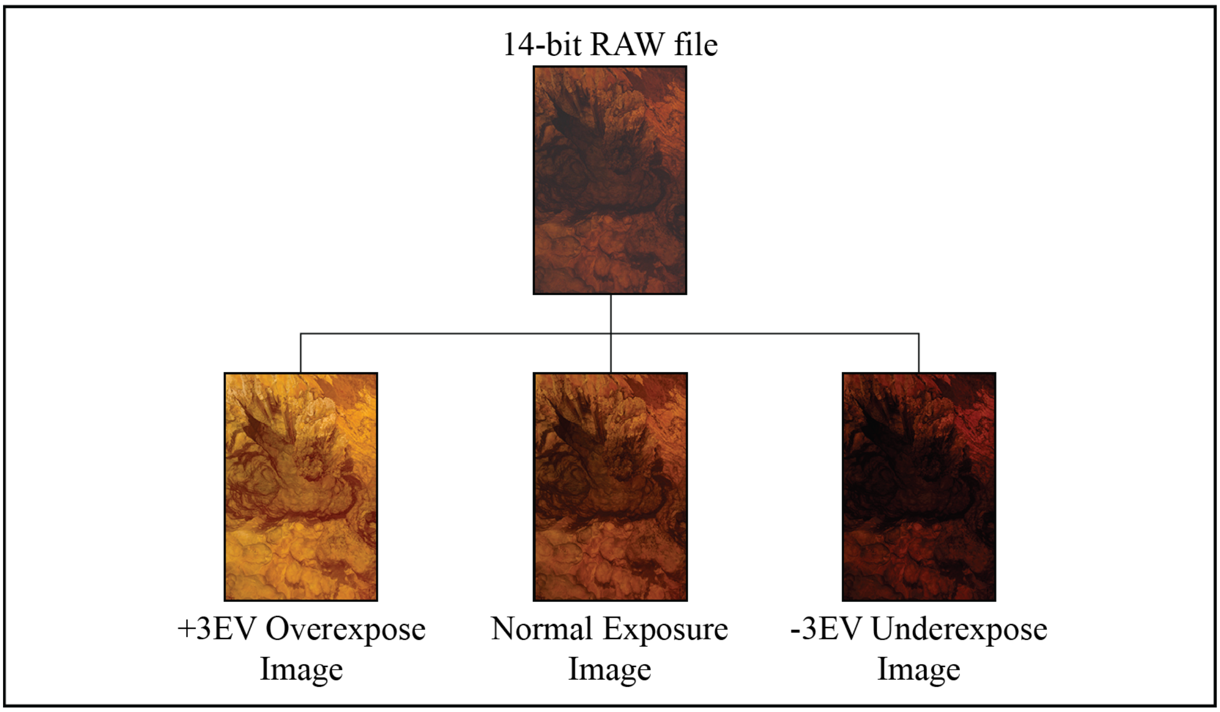

Figure 4. A single-shot HDR technique by extracting exposure from a single RAW file.

Figure 5. A sample of panoramic image in low dynamic range.

Figure 6. A sample of panoramic image in high dynamic range.

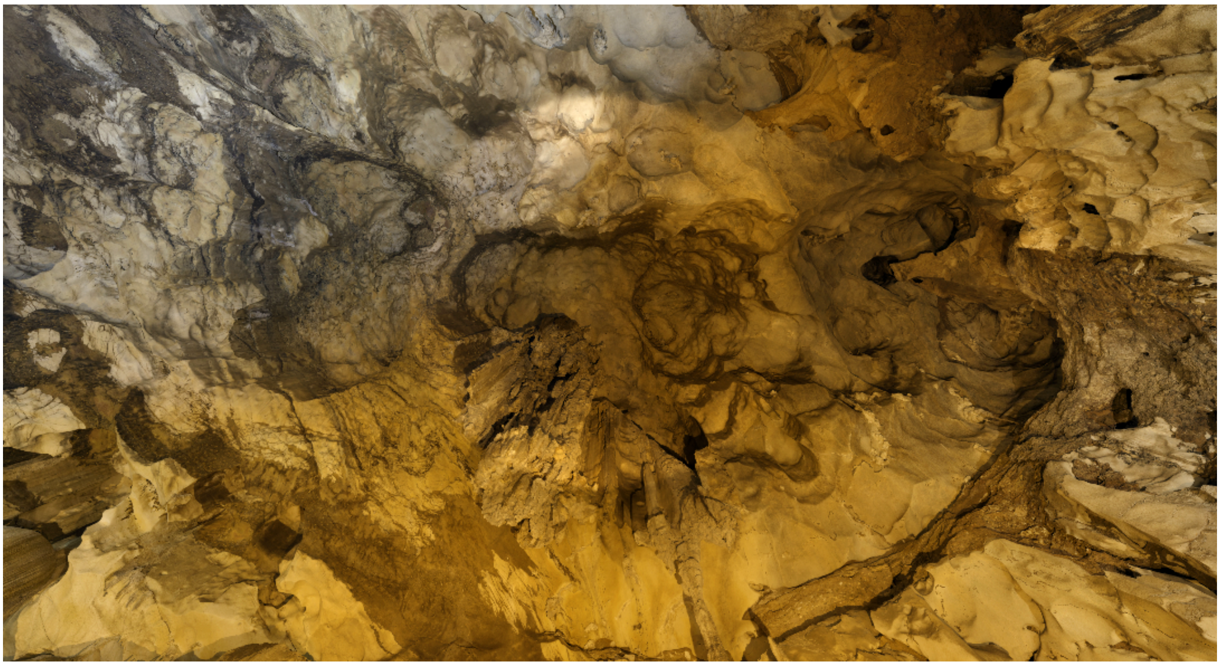

Figure 7. Lost details in shadow areas that could be recovered through the HDR process.

4. Immersive Imagery

It has been a tremendous challenge for visual creators to produce a highly realistic image [26]. For years many techniques have evolved and progressed to represent artificial scenes realistically but have yet close to the lifelike appearance. There is a huge interest in super-resolution HDR images in many applications due to the high pixel density and details that might be a critical need in specific applications. In a photo-based immersive environment that represents the realness of the visual, the viewer is wrapped in a surrounding image to simulate the experience of how a person sees the world [27]. Such visual details are essential in the sense that image resolution and low dynamic range tend to limit its states from appearing real in terms of visual perception [28]. Although immersive visualization could have many applications possibilities, throughout the years it is keener towards landscaping and tourism. An example study had proven that the tourism industry benefited the most from the immersive application with many respondents giving positive feedback [29]. Media such as sounds, videos images, augmented reality objects, and textual information can be overlaid on top of in an immersive environment to create multimedia-rich content. This would be more engaging for viewers to experience rather than static forms of presentation. Furthermore, the interactive exploration feature of the details and elements embedded within is a good means of encouraging viewers to learn and communicate [30].

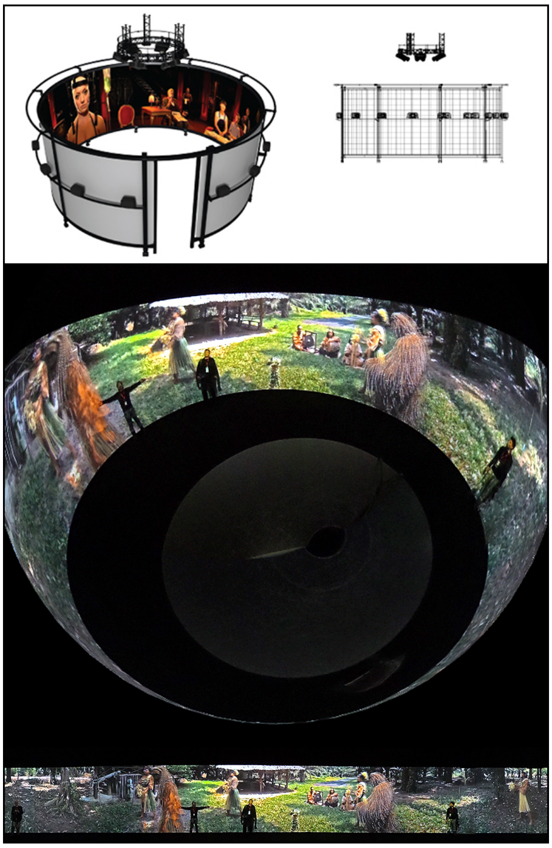

Figure 8. 360 Panoramic Theatre at School of Creative Media, City University Hong Kong.

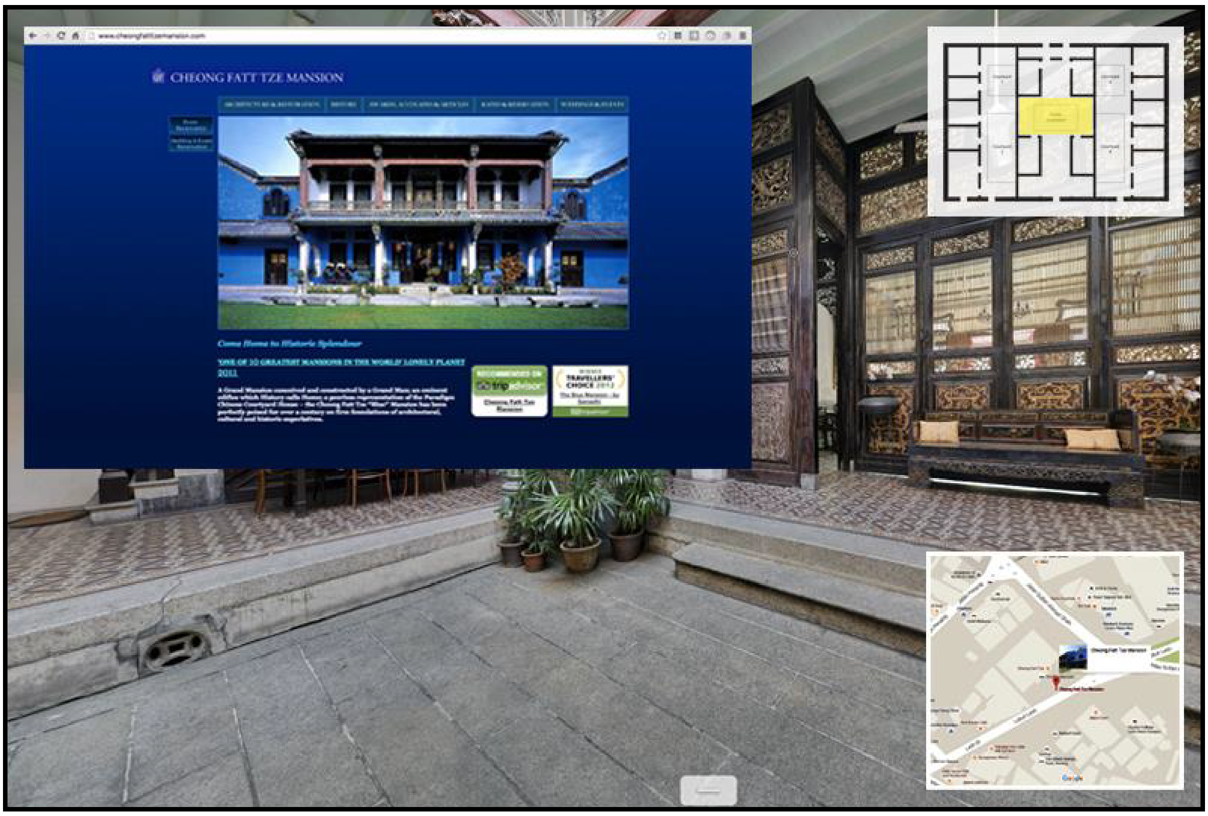

Figure 9. Multimedia HDR super-resolution panoramic projection.

5. Conclusion

In today's demand for high fidelity imagery, detail and visual clarity are a major concern amongst visual enthusiasts in every aspect from content creation to content delivery. Electronic device manufacturers in the consumer level up to the professional segments had been pushing for higher resolution and HDR imaging for the past 10 to 15 years. Furious competition can be seen since the debut of High Definition Television (HDTV) in early 2000. The race had then been picking up speed with 3D visuals, 4K, and now 8K resolution. This has indirectly opened a new leaf in driving high-resolution visual needs covering computer games, digital cameras, filmmaking, advertising, multimedia, and all sorts, not to mention virtual reality. Immersive technologies had found their way into the market and have highly been accepted by many; from panoramic imaging on phones to virtual reality (VR) headsets; from augmented reality (AR) to mixed reality (MR), and now we are at the brink of extended reality (XR) which will push immersive technologies and applications further. Television programs around the world are adopting immersive technologies with virtual sets, and museums have been keen on presenting their exhibition materials immersively. Some medical-related sectors also had been using immersive technology to study the human brain. Immersive technologies could also be the key for virtual tourism and exploration especially in times such as during pandemics. This shows that there are virtually unlimited applications of immersive technologies in various aspects of our lives. However, development in hardware possesses setbacks as a huge investment in research and technical difficulties need to be addressed. More refined solutions such as exploration of less hardware-oriented techniques and methods are most welcome; techniques that can be learned and adopted by the masses to fill in the gap of hardware deficiencies and reduction of owning or operating cost of specific hardware. Apart from empowering techniques and methods, various hardware and software can be explored in streamlining the production of super-resolution HDR images or producing various forms of output for other purposes that are more specific to a particular field. However, many technical challenges remain in order to further the limits in imaging technologies in handling and creating such large-sized high detailed output.

References

[1] Adel, E., Elmogy, M., & Elbakry, H. (2014, August). Image Stitching based on Feature Extraction Techniques: A Survey. *International Journal of Computer Applications*, *99*(6), 8. DOI: 10.5120/17374-7818

[2] Arnold, D. (2008). Digital Artefacts: Possibilities and Purpose. In M. Greengrass & L. Hughes, *The virtual representation of the past* (1st ed.). Aldershot, Hants, England: Ashgate.

[3] Btz, M., Koloda, J., Eichenseer, A., and Kaup, A. (2016). Multi-image super-resolution using a locally adaptive denoising-based refinement. In 2016 IEEE 18th International Workshop on Multimedia Signal Processing (MMSP), pages 1–6.

[4] Baker, S., & Kanade, T. (n.d.). Limits on Super-Resolution and How to Break Them. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

[5] Betancourt, D., & del Río, C. (2006, July). Study of the human eye working principle: an impressive high angular resolution system with simple array detectors. *Fourth IEEE Workshop on Sensor Array and Multichannel Processing, 2006.* 10.1109/SAM.2006.1706098

[6] Brown, G.J. (2006). RPS Journal.

[7] Coe, C., & Weston, C. (2009). *Creative DSLR Photography: The ultimate creative workflow guide*. Focal Press.

[8] Cicala, R. (2012, November 17). *TheCameraVersustheHumanEye*. Retrieved July 20, 2021, from https://petapixel.com/2012/11/17/the-camera-versus-the-human-eye/

[9] Clark, R. N. (2005, January). *Notes on the Resolution and Other Details of the Human Eye*. ClarkVision.com. Retrieved July 3, 2021, from https://clarkvision.com/imagedetail/eye-resolution.html

[10] Cheng, P., Qiu, Y., Wang, X., and Zhao, K. (2017). A new single image super-resolution method based on the infinite mixture model. IEEE Access, 5:2228–2240.

[11] Davis, H. (2012). Creating HDR Photo: The Complete Guide to High Dynamic Range Photography. Amphoto Book.

[12] Economou, M., & Tost, L. (2008). Educational Tool Or Expensive Toy? Evaluating VE Evaluation And Its Relevance For Virtual Heritage. In Y. E. Kalay, T. Kvan, & J. Affleck, New Heritage: New Media and Cultural Heritage. London and New York: Routledge.

[13] Fitzgerald, B. (January, 2008). A Brief on HDRI. The Digital Media Club (http://www.dmcscad.com)

[14] Gallo, O., Gelfand, N., Chen, W.-C., Tico, M., & Pulli, K. (2010). Artifact-free High Dynamic Range imaging. In *2009 IEEE International Conference on Computational Photography (ICCP)*. IEEE. 10.1109/ICCPHOT.2009.5559003
[15] Gaidhani, P. (n.d.). Super-resolution. *Education Information UK, CVonline*. https://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/AV1011/Super_Resolution_CVonline.pdf
[16] Gulbins, J., & Gulbins, R. (2010). Photographic Multishot Technique. Rockynook.
[17] Hirsch, R. (2012). *Light and Lens: Photography in the Digital Age* (Second ed.). Focal Press.
[18] Hui Khoon, L. (2010). A Study of Impacts of Photography And Immersive Imaging On Tourism. Universiti Malaysia Sarawak.
[19] Larson, M.D, et. al. 1997. A Visibility Matching Tone Reproduction operator for High Dynamic Range Scenes. Visualization and Computer Graphics, IEEE Transactions. 3(4), 291-306.
[20] Malczewski K., Stasiński R. (2009) Super Resolution for Multimedia, Image, and Video Processing Applications. In: Grgic M., Delac K., Ghanbari M. (eds) *Recent Advances in Multimedia Signal Processing and Communications. Studies in Computational Intelligenc*e, vol 231. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-02900-4_8
[21] Masaoka, K., Nishida, Y., Sugawara, M., Nakasu, E., & Nojiri, Y. (2013, March). Sensation of Realness From High-Resolution Images of Real Objects. *IEEE Transactions On Broadcasting*, *59*(1). 10.1109/TBC.2012.2232491
[22] McCollough, F. (2008). Complete Guide to High Dynamic Range Digital Photography. Lark Books.
[23] Mohamed, R. S., & Yousif, G. A. (2010, October). Input resolution and its effect of the printed image quality on digital toner printing systems (case study – Sinai, Egypt). *The Egyptian Journal of Remote Sensing and SpaceSciences*, *13*(1), 75-80. doi:10.1016/j.ejrs.2010.07.009
[24] Murphy, P. (1999). *Psychology of VR Panorama*. Retrieved July 2, 2016, from http://www.culture.com.au/virtual/optic.html
[25] Roussou, M. & Slater, M. (2005). A Virtual Playground For The Study of The Role Interactivity In Virtual Learning Environments. In *The Eight Annual International Workshop on Presence. PRESENCE.*
[26] Roussou, M. (2008). The Components of Engagement In Virtual Heritage Environments. In Y. Kalay, T. Kvan & J. Affleck, New *Heritage: New Media and Cultural Heritage* (1st ed.). London: Routledge.
[27] Singh, A., & Sidhu, a. S. (2016, September). Super Resolution Applications in Modern Digital Image Processing. *International Journal of Computer Applications*, *150*(2). https://www.ijcaonline.org/archives/volume150/number2/singh-2016-ijca-911458.pdf
[28] Tylor, D. (2012). Understanding HDR Photography: The Expanded Guide. Ammonite.
[29] Yue, L., Shen, H., Li, J., Yuan, Q., Zhang, H., & Zhang, L. (2016). Image super-resolution: The techniques, applications, and future. *Signal Processing*, (128), 389-408.
[30] Zhu, Jie; Zhang, Ellean; and Del Rio-Tsonis, Katia (November 2012) Eye Anatomy. In: eLS. John Wiley & Sons, Ltd: Chichester. DOI: 10.1002/9780470015902.a0000108.pub2