
The Volca Project: A Multi-Sensory Camera for Augmented Reality

Daniel Buzzo

University of West of England

Bristol, UK

daniel.buzzo@uwe.ac.uk

ABSTRACT

The Volca project is a *speculative photography* research project that is in parts a proposal, a sensory experiment and also continual work in progress.

Volca is an experimental camera apparatus for recording what Flusser [8, 9] would term 'technical images'. Its design is to deliberately open the discussion of what it is to be photographer, or collaborator or even facilitator of a complex apparatus with a discrete system at its core. A system that extends beyond the physical body of the device and reaches into the heart of our human culture of creation, consumption, reflection, storytelling and the active shaping of our identity and self image.

The project has several separate aspects dealing with hardware, software, interaction design and media and photographic theory. This paper and the body of work it describes is the first collection of portraits and street photography produced with this new apparatus. The collection is arranged as a virtual reality photographic essay or artists photo-book. The making, distribution, display and consumption of these new *speculative photographs* is a vital part of the investigation of what can be considered the system of photography. A system where camera and photographer are only a small part of the philosophical whole.

CHI'18 Extended Abstracts, April 21–26, 2018, Montreal, QC, Canada

© 2018 Copyright held by the owner/author(s).

This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in *CHI'18 Extended Abstracts: CHI Conference on Human Factors in Computing Systems Extended Abstracts, April 21–26, 2018, Montreal, QC, Canada*, <https://doi.org/10.1145/3170427.3177767>.



Figure 1: Girl in Red Sweater. Still from 3D Volca Portrait rendered using Delaunay triangulation.

¹<https://developer.microsoft.com/en-us/windows/kinect>

²<http://stereolabs.com>

³<http://www.depthkit.tv>

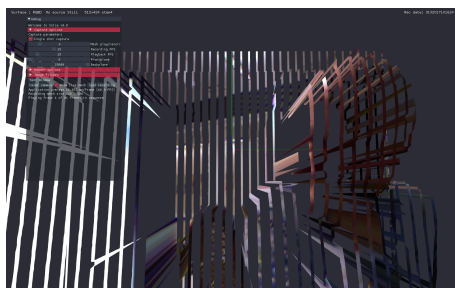


Figure 2: Realtime view of filtered volumetric mesh with surfacing and generated surface normals.

CCS CONCEPTS

- **Human-centered computing** → *Interaction design process and methods; Visualization techniques;*
- **Applied computing** → *Media arts;*

KEYWORDS

intelligent collaborative system, RGBD, immersive imaging, experimental photography, interaction design, street photography, augmented reality

ACM Reference Format:

Daniel Buzzo. 2018. The Volca Project: A Multi-Sensory Camera for Augmented Reality. In *CHI'18 Extended Abstracts: CHI Conference on Human Factors in Computing Systems Extended Abstracts, April 21–26, 2018, Montreal, QC, Canada*. ACM, New York, NY, USA, 6 pages. <https://doi.org/10.1145/3170427.3177767>

INTRODUCTION

Volca is a multi-sensory camera, built around a depth sensor and visual light sensor, capturing RGB and depth and meta data. Currently the system uses either a Microsoft Kinect¹ infra-red sensor or the StereoLabs² ZED sensor. The system is in continual development and currently has extra tools to capture additional sensory and meta-data. It is a novel experimental camera system that captures RGBD (RGB and Depth) images through live, visible light stereoscopy or through the use of infra-red structured light technology. Previous work in this area has investigated several applications such as robotic vision, video production and art installation. [11] [6] [12], the DepthKit 3D video project being most notable among them³. The Volca project is explicitly focussed on the arena that the conventional still image camera currently occupies and as a 'project vision' it imagines what a future multi-sensory camera for urban street photography might be.

The body of work described in this paper is a collection of augmented reality (AR) 3D portraiture and street photography shot using this experimental system. Organised as the Virtual Reality (VR) equivalent of a traditional photographic essay or monograph, the project represents one view of a speculative, virtual, immersive future where we not only consume new AR and VR content, we make it ourselves.

Speculative Photography

The work takes the position that the camera is a discrete apparatus that is pre-programmed with all the potential images that the camera can create, contained within it's in-built parameters. In this situation the 'photographer' is in actuality a facilitator of this apparatus. In this context it is important to consider, when investigating new means of photographic production, that the forms and means of dissemination and consumption of photographic images is implicit, and inextricably



Figure 3: Portable hand-built Volca experiment rig with USB button interface for eternal control, illustrating Kinect sensor array on tripod rather than hand held. with CPU, batteries controls .etc are carried in shoulder bag.

⁴<http://openframeworks.cc>

⁵<https://opencv.org>

⁶<https://www.w3.org/XML/>

⁷<https://www.exif.org>

encoded, within the mechanisms and systems of the camera apparatus itself. Therefore, *new means of production of images sit hand in hand with new means of dissemination and of consumption* of these 'images'. It is these new means of dissemination and consumption, rather than means of creation, that we concentrate on with this collection of 'portraits' presented within this work.

The project development is one of *Speculative Photography*, using Research through Design (RtD) [16] techniques and hardware and software making to investigate what a camera of the future *could* be. Guided by some of the notions of 'camera as a discrete apparatus' from Flusser [8] and ideas of visualising the imperceptible [4] it must be borne in mind when approaching photographic apparatus that a photographic apparatus;

- is a sensor capturing data
- is a mechanism operating over time
- is a machine that generates images
- despite how it is commonly imagined does not 'see' in any human like way,
- is a system where all possible outputs are already encoded within it
- is explicitly a product of the means of consumption of the images it creates

The current apparatus hardware configuration is linked to and controlled by custom software written in C++ using the OpenFrameWorks library/toolkit ⁴. It incorporates Computer Vision processing capabilities, via the openCV project libraries ⁵ to analyse and manipulate colour image data and volumetric depth data. The system filters colour and depth data in semi-passive ways to smooth, define and aesthetically manipulate the source material. The system also applies rudimentary meta or semantic filtering to generate knowledge of the contents captured by the apparatus. The meta-data is currently formatted as XML⁶, saved alongside image and depth data and experimentally extends the current Design rule for Camera File system (DCF) Exchangeable Image File Format (EXIF) image meta-data standard⁷. This standard is common among contemporary manufacturers of camera systems and image processing software but is principally geared toward 2D RGB data from a single sensor source. The processes and results of extending this standard to incorporate multi-sensor data and contextual awareness will be the subject of a future paper.

The system also incorporates a variety of image analysis routines to assemble a rudimentary 'understanding' of the incoming sensor data. Among them is a Haar Cascade Classifier for face detection and other simple computer vision techniques for object and boundary analysis. This system identifies edges, contours, movement, objects, using algorithmic processing from sources such as the openCV library[15] to identify not just Regions of Interest (ROI) but semantic *Things of Interest* (TOI). The proposition is that combined with contextual understanding of day, date and geo-location the system is envisaged to, one day, begin to understand *things* in their place, and their relationship to each other.



Figure 4: Painted and unsurfaced three dimensional mesh of butcher and butcher's stall.

The future addition of machine learning and advanced vision techniques is intended to bring a higher degree of semantic understanding inside the system of the apparatus. The intention is that the system will not only define *'things of interest'* but identify them semantically. Eg beyond what it currently can do as *hand, face, eye, smile*, but as *car, bicycle, cat*. This enhanced understanding by the system would enable entering into a higher level of dialog with the human operator of the apparatus and the machine and human consumers of the images produced. Previous work has investigated this possible collaboration with intelligent machines[3], and the nature of time and representations of time[1, 2, 4, 5].

SYSTEM DESIGN

The conceptual approach to the system design is rooted in ideas of conversation theory from cybernetician Gordon Pask [10] and the user conceptual modelling approaches of Bill Moggridge [13, 14] et al and built against a backdrop of the 'speculative framework' (in the speculative sense of Dunne and Raby et al) [7]. The interaction with the apparatus is thought about as a collaboration or dialog with an intelligent system, where the human operator/facilitator is not the dominant, powerful, privileged partner in the interaction, but the outside interloper.

As the system develops complexity and functionality the form of both hardware and software is being refined and improving in reliability and accessibility. The interface has two main modalities, facilitating image capture and reviewing, processing or visually investigating images. The interface uses an iPac interface board linked to trigger buttons on the sensor head with audio feedback, for main functions, to enable screenless shooting. In experimental phase currently is the addition of a ChalkElectronics HD multitouch HD screen⁸ and porting software to the nVidia line of Jetson GPU accelerated computing modules⁹.

⁸<https://www.chalk-elec.com>

⁹<https://developer.nvidia.com/embedded/buy/jetson-tx2>

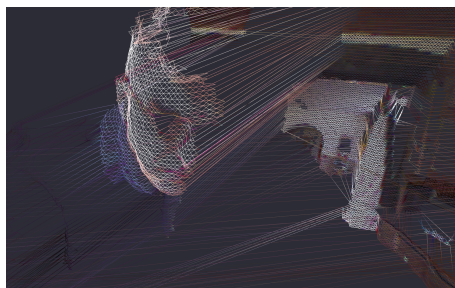


Figure 5: Painted and unsurfaced three dimensional mesh portrait.

PORTRAIT PORTFOLIO

This work illustrates a selection of images from the output from the project, using the apparatus in real world situations. The portrait collection has been assembled from many thousands of images shot in several countries over the course of the last 12 months of development. The images in the portfolio are viewed piece by piece in an immersive 3D gallery of individual 'shots'. Each portrait view is capable of being manipulated and seen from any direction in the virtual environment. The portfolio is available as an executable file for mac OSX and Windows computers or can be supplied as a series of individual models and meshes for inclusion into a compendium of other works in 3D. File types and formats can be arranged to suit specific applications, subject to technical considerations. Viewing is suitable for a variety immersive situations including Oculus Rift, Samsung Gear VR, Google cardboard etc. The work is also suitable for 2D, flat screen based viewing of the individual shots provided that suitable interfaces for interaction allow viewers to investigate the pieces. Though devices such as touch screen, mouse or joystick.

This flexibility of dissemination and of consumption is a deliberate part of the Volca project's investigation. Through the active, practical investigation of how these new types of *speculative photographs* may be shared and viewed the Volca camera itself will subsequently develop. In part being shaped by the system of distribution and consumption of images as much as the possibilities of originating them.

You can see more of the project online at <http://www.buzzo.com> and see the current source-code at <https://github.com/danbz/volume-camera>.

CONCLUSION

This first collection of *speculative photographs* are intended to investigate what a future photograph may be. Bringing together the first outputs from one version of what a *speculative camera* might be. Taking photographs in a future where imaging technologies that perceive beyond the human sensorium are common-place and new systems of dissemination, consumption, and of artistic judgement are established.

REFERENCES

- [1] Daniel Buzzo. 2013. Changing Perspectives of Time in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Extended Abstracts*. ACM - Association for Computing Machinery, Paris, France.
- [2] Daniel Buzzo. 2014. Time travel: Time dilation. In *Electronic Visualisation and the Arts*. British Computer Society, 170–176. <https://doi.org/10.14236/ewic/eva2014.21>
- [3] Daniel Buzzo. 2015. Collaborating with intelligent machines. In *CHI EA '15 Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. ACM - Association for Computing Machinery, Seoul, Korea, 1–2.
- [4] Daniel Buzzo and David Jonas. 2015. Designing For The Impossible : Creating A Mobile Application For Time Dilation. In *Electronic Visualisation and the Arts*. British Computer Society, London.
- [5] Daniel Buzzo and Nicolo Merendino. 2015. Not all Days are Equal : Investigating the Meaning in the Digital Calendar. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press, Seoul, Korea.
- [6] Leandro Cruz, Djalma Lucio, and Luiz Velho. 2012. Kinect and RGBD images: Challenges and applications. *Proceedings: 25th SIBGRAP - Conference on Graphics, Patterns and Images Tutorials, SIBGRAP-T 2012* June 2017 (2012), 36–49. <https://doi.org/10.1109/SIBGRAP-T.2012.13>
- [7] Anthony Dunne and Fiona Raby. 2013. Speculative everything: Design, fiction and social dreaming. *The MIT Press* III, 1 (2013). <https://doi.org/10.1093/jdh/epv001>
- [8] VilÅlm Flusser. 1983. Towards a philosophy of photography. *English* (1983), 1–96. [https://doi.org/10.1016/S0031-9406\(10\)62747-2](https://doi.org/10.1016/S0031-9406(10)62747-2)
- [9] VilÅlm Flusser. 2011. *Into the Universe of Technical Images*. Vol. 224. University of Minnesota Press. 225 pages.
- [10] Usman Haque. 2007. Distinguishing concepts: Lexicons of interactive art and architecture. *Architectural Design* 77, 4 (2007), 24–31. <https://doi.org/10.1002/ad.484>
- [11] Jiajun Lu, Hrvoje Benko, and Andrew D Wilson. 2017. Hybrid HFR Depth : Fusing Commodity Depth and Color Cameras to Achieve High Frame Rate , Low Latency Depth Camera Interactions. (2017). <https://doi.org/10.1145/3025453.3025478>
- [12] Djalma Lucio, Leandro Cruz, and Luiz Velho. 2012. RGBD Camera Effects. *Decom.Ufop.Br* June 2017 (2012), 1–4. <http://www.decom.ufop.br/sibgrapi2012/eproceedings/wivis/00002.pdf>
- [13] Bill Moggridge. 1999. Design: Expressing experiences in design. *Interactions* 6, 4 (1999), 17–25. <https://doi.org/10.1145/306412.306430>
- [14] Bill Moggridge. 2006. Multisensory and Multimedia. *Designing Interactions* (2006), 513–586.
- [15] Vadim Pisarevsky. 2010. Introduction to OpenCV. *Agenda* 42 (2010), 433–434. <https://doi.org/10.1016/j.jenrad.2010.04.011>
- [16] John Zimmerman, Erik Stolterman, and Jodi Forlizzi. 2010. An Analysis and Critique of Research through Design: towards a formalization of a research approach. *Proc. DIS 2010* (2010), 310–319. <https://doi.org/10.1145/1858171.1858228>