The User Configuration of Microsoft Kinect

STS Research Paper
Presented to the Faculty of the
School of Engineering and Applied Science
University of Virginia

By

Rahat Maini

April 11, 2019

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

Signed: 

Approved: _____ Date _____
Benjamin J. Laugelli, Assistant Professor, Department of Engineering and Society

Introduction

In 2013 Microsoft was riding high on the coattails of its massively successful Xbox 360 console, setting sales records and bringing in a customer base of millions for a category of product the company had rarely experimented with in the past (Plant, 2018). However, this success did not translate over into the next generation of their product which received relatively poor sales, various customer concerns, and several decision reversals from corporate (Rigney, 2013)**.** Nearly all of these setbacks can be attributed to one technology, the Microsoft Kinect, a mandatory peripheral bundled with every brand new Xbox One that coupled an array of microphones and advanced camera sensors for home entertainment features that Microsoft claimed would be impossible otherwise. Today, the Kinect exists exclusively as a kit specially tailored for developers to utilize its potent array of technologies for purposes that are more ambitious than an upgraded living room controller (Microsoft Corp, n.d.**)**. Despite the Kinect's appeal and marketing as a gadget for all, its design indicates that certain biases concerning the user's identity were embedded within the product. Given the appeal and widespread use of its latest iteration in the academic world, the idea that the designers had intended for the Kinect to be less of a gaming peripheral for users of all ages and more of a tool for scientific research has some support.

By exploring how the designers' ideas about the Kinect's user base were embedded into the artifact, we gain a better understanding of the way unintentional and implicit bias can affect product design. Drawing on the science, technology, and society (STS) concept of User Configuration, I argue that the Kinect's current design resulted from a process of earlier failures in which the intended user of the Kinect was not the user that Microsoft had any interest in

selling in mass. Through user configuration, I argue that the designers unintentionally embedded biases about the users they were developing the product for into the technology's design. As a result, I argue that ideas about users' identities have been configured and embedded into the Kinect's design that led it to be vastly undervalued as a consumer electronics device and more successful as a device for academic research and experimentation.

Background

At the start of the decade that launched the 2010's, Microsoft was at a unique inflection point that was hinting more and more at a complete reinvention of the company. Microsoft became more interested in becoming more consumer-minded (Warren, 2013). This seems almost antithetical to the approach the company is taking today where a focus on developers, cloud infrastructure, and B2B transactions has led the company to an all-time high on the stock market (Farley, 2020)**.**

This shift began to percolate inside Microsoft as it saw an opportunity in a highly lucrative market, the TV and home entertainment segment, to have an iPhone-like hit product powered by Microsoft services and content. With a blockbuster gaming business in tow, a research arm that was flexing its skills in the fields of computer vision and human-computer interaction, and a hardware division producing innovative machines (Warren, 2012), Microsoft rightfully believed it was poised to develop the product that would reach beyond just gaming and infiltrate the entire living room experience which they controlled end-to-end (Patel, 2013), following a model of market success popularized by a familiar longtime rival, Apple.

Literature Review

While several scholars agree that the Microsoft Kinect is now an indispensable tool for academic research and application of concepts once confined to computer science laboratories, an analysis of what preconceived notions about a user's identity were embedded into the Kinect are missing. Scholars have not yet adequately considered that to understand the Kinect's current state as a developer marketed and designed tool, we must understand the biases of the designers that leaned heavily towards the Kinect serving users that more closely aligned with their goals. Once these biases are understood, scholars can begin to see the Kinect in a new light that sheds an understanding of its true strengths. To form a better understanding, I collated two scholarly sources: one concerning itself with a comparison between two generations of the Kinect sensors, and another exploring the Kinect sensor's extensive capabilities regarding full-body motion tracking.

Pagliari and Pinto, in their paper compare the original Microsoft Kinect with its successor, the v2 in regards to depth measurement accuracy, image sensor precision, and sensor stability (Pagliari & Pinto, 2015). The primary focus of this paper was to establish the two sensors as primarily gaming-oriented peripherals that have shown great potential for research use, and then show markedly better results obtained with the v2 sensor compared to the original (Pagliari & Pinto, 2015). The main objective of the scholars here was to obtain scientific evidence to quantify the improvements that Microsoft had made, however, the undertones of gaming are prevalent here, which leaves a gap in scholarship since the authors fail to understand the motivations and design decisions that influenced the changes they are studying and hoping to make sense of.

Procházka, et al., in their paper argue for the potential use of the Microsoft Kinect for detection of gait disorders (Procházka, et al., 2015). In their experiment, using computer vision algorithms pre-built into the Kinect for skeletal tracking they can capture and classify data sets of individuals' gaits as healthy or unhealthy. The primary finding of their study was that specialized and expensive camera systems that house very specific sensors are no longer necessary for the rigorous scientific research teams such as themselves are carrying out. With 91.7% accuracy, the team was able to use the Microsoft Kinect to recognize Parkinson's disease within an individual (Procházka, et al., 2015). The team conducting this research have a strong understanding of the Kinect's current capabilities, but neglect to report much on the democratization of such advanced computer vision technology that Kinect brings. This would be particularly relevant since the paper does emphasize the low-cost nature of the sensor and the beginnings of their system setup, yet fails to provide insight into how the Kinect differs from standard, more expensive equipment beyond just the price tag. Such information would be highly valuable to research groups looking to duplicate or append efforts onto the project.

Pagliari and Pinto's publication overlooks the historical context of the Microsoft Kinect and what this implies for the further advancement of the technology by Microsoft. By attributing the peripheral as instrumental mainly to allowing "new interactions during games" and relegating its research potential to a secondary side-success, Palgiari and Pinto do not adequately form an understanding that would lead them to understand the Kinect's next iteration (Pagliari & Pinto, 2015). The identity of the user embedded within Kinect is effectively ignored in favor of a raw statistical understanding of the device. If Procházka, et al., had explored more the democratization aspect of the Kinect sensor, they would have arrived at a different list of tasks

for “further work” for their team. As of now, their paper exclusively sees Kinect as a low-cost alternative to pre-existing solutions, rather than a product developed with biases in mind that were fully realized in later generations of the product. In my analysis and research, I will advance understanding of the Kinect sensor by paying attention to the preconceived notions embedded within it regarding accessibility for all researchers, how this informs a new method of scientific research through computer vision, and why this is a worthwhile realization about the sensor itself.

Conceptual Framework

To frame my analysis of the Kinect’s public failure, I draw on Oudshoorn and Pinch’s concept of user configuration (Oudshoorn & Pinch, 2003). The STS concept of user configuration is mainly defined as when designers define the identity of the supposed users of their products (Oudshoorn & Pinch, 2003, p.8). This results in a technological artifact that carries certain values inherently within itself alongside the functions it is meant to perform.

The concept of user configuration proposes that specific ideas about user identity (e.g. gender, race, age, competency, values, socioeconomic status, etc.) are configured into a technology’s design (Oudshoorn & Pinch, 2003, p.6). The user is configured once their identity is defined and constraints are set upon their future actions regarding the engineered product they utilize. The reverse is also true, where designers of a product are configured by both users and their organizations, as they are constrained by organizational methods and testing of new products among people within their organization (Oudshoorn & Pinch, 2003, p.8).

The idea of user configuration is relevant to my research as it allows me to inspect the Kinect as a technological artifact beyond just its capabilities and functions. User configuration gives me insight into the design decisions behind the Kinect as well as the corporate priorities the product was assigned to fulfill. This understanding of who the user is meant to be from the viewpoints of different stakeholders divides my case study into individually addressable perspectives that I can then cite as the mismatch of user and product that resulted in the initial failure of Microsoft's Kinect. Following this, I can pinpoint Microsoft's lean into the original configured user identity for the Kinect sensor that led to its newfound success.

Analysis

The Microsoft Kinect was vastly undervalued as a consumer electronics device, as seen by its lackluster sales figures (Rigney, 2013). Instead, the peripheral found its success when reincarnated as a device for academic research and experimentation after earlier iterations had proven to gain traction in those fields rather than gaming (Microsoft Corp, n.d.**)**. In this section, I will analyze how the designers of the Kinect had configured a specific user identity at odds with the user Microsoft was targeting, and continued to do so until Microsoft followed suit by matching the user identity with the market of sale. I argue why the framework of user configuration offers us new understanding that had previously been missed during the analysis of the Kinect.

Unusual advancements from v1 to v2

In Kinect v2, we see Microsoft double down on the vision aspects of the Kinect sensor in terms of quality and accuracy. As Pagliari and Pinto discovered, the newer sensor has a wider angle lens, greater frame rate, and lower sensor noise in extreme conditions (Pagliari & Pinto, 2015). Also notable was the v2's ability to acquire data in the range of 0.50-4.50 meters, where the v1 sensor can only calculate depth in the range of 0.80-4.00 meters (Pagliari & Pinto, 2015). During the development and release process of the Kinect v2 however, whistleblower Edward Snowden detailed exactly how the National Security Agency had been essentially spying on the citizens of the United States, sending the country into panic and reflection (Sottek & Kopfstein, 2013). As a result, the national conversation became very much about how consumer electronics had become privacy-invasive and infiltrated people's pockets, homes, and even wrists with smartwatches. Despite these concerns, Microsoft unwaveringly continued development with minimal attention to features that the industry was adopting such as end-to-end encryption to assuage feelings of unrest among privacy-conscious individuals (Metz, 2018). As expected, the release of an always-on camera that has infrared capabilities to see users in complete darkness and microphones listening continuously in the middle of a living room was not exactly well-received (Keilman, 2016).

What is easily derivable from this situation is that the designers of the Kinect had a user identity in mind that was not the layman using an Xbox in their living room. Pagliari and Pinto's research reduces Kinect to mainly a gaming peripheral that Microsoft kept advancing for a richer Xbox experience, and the research community benefits as a byproduct of these efforts (Pagliari & Pinto, 2015). However it becomes clear that the improvements made to Kinect that wholly disregarded the market and privacy-conscious climate were made to benefit a user that disregards

the societal concern of privacy in favor of a tool that is unfiltered in its capabilities of environment sensing and capturing data. As Pagliari and Pinto were unable to recognize this user configuration, they missed out on properly understanding the Kinect as an artifact, which would have lent them a better understanding of the goals and aspirations of the device and its designers. Such is evident now, years later when Microsoft has released a v3 of the Kinect sensor that is unapologetically focused on mass data collection with no privacy measures built-in by design as it is meant exclusively to be a research tool rather than a living room computer (Lardinois, 2019). This evolution across three generations of the Kinect is gapped in understanding that is best explained through the lens of user configuration.

Alternate Viewpoint and Response

I have just argued that Microsoft made some questionable business decisions when upgrading their Kinect sensor from v1 to v2 and did so in favor of advancing their product as a research tool. However, it may be argued that Microsoft was seeking to improve the experience of Kinect gaming by advancing the technical specifications of the product rather than improving the experience for researchers. It can be stated that Microsoft was aiming to bring higher fidelity technologies to the Kinect sensor as an effort to position the Kinect as a more valuable gaming peripheral.

However, this viewpoint would be flawed for several reasons, chief among them being that the Kinect detracted computing resources from the Xbox console when in use, resulting in lower fidelity gaming. Microsoft confirmed that the Xbox sees a 10% performance decline when the Kinect is in use, which would lead to an inferior gaming experience in terms of frame rate as

well as resolution (Statt, 2014). Had Microsoft truly valued the gaming experience when using Kinect, the device would not throttle the processing power of the Xbox console. As Microsoft should have expected consequently, developers simply did not produce many games for the device (Hanson, 2018). Identifying the configured user of the Kinect as someone other than a gamer is an idea with strong evidential support, and is indicative of pure advancement of the Kinect as a research tool.

Preconceived ideas regarding use as a tool for research

Microsoft opted to sell the Kinect as a standalone device direct to consumers, foregoing any business licensing or complex deals that many advanced researching tools adopt (Protalinski, 2014). Along with their decision to sell the sensor as they did, was the decision to release the SDK (software development kit) for the device for free, a very welcome move in the eyes of researchers clamoring for low-cost software tools, as this often results in more community adoption and support. For a corporation such as Microsoft, the decision to not monetize the Kinect sensor software through licenses and B2B dealings speaks strongly of the culture that the company wanted to instill within Kinect of open development and research.

Procházka, et al., in their paper fail to recognize the embedding of this user identity and continue to use Kinect as they would a standard tool of research, in the silo of their lab where they conduct the data collection, algorithm development, and result classification (Procházka, et al., 2015). The understanding gained from viewing Kinect as inexpensive hardware democratized through free software enables researchers to use the device as a sort of lower accuracy tool that fills in the gaps with mass data collection instead. Again, the user identity configured by the

designers of the Kinect lends strong support to the idea that the device was never meant to be a living room computer that the layman was comfortable with. The treatment of the SDK is more evidence of this user identity, as the designers of the Kinect enforce the idea of open collaboration with little regard to consumer privacy (rightfully so, as the objective of a data-gathering research tool should be usability and results rather than data security). The release of the SDK is indicative of a larger trend in computer science, where algorithms as a result of mass data collection can downplay hardware deficiencies (Shankland, n.d.). This effort would not be possible in the utility of the Kinect as a gaming peripheral and would serve no benefit to a user of the Kinect to blast IR waves in order to turn on their television. In recognizing the configured user as an academic with access to a network of data, the Kinect is realized as a device intended by its designers to be most accessible and powerful to scientific researchers. This analysis is missing from the scholarship of Procházka, et al., who are unable to view the Kinect as such a device. In doing so, they miss out on research that would change their study in the future considerably.

Conclusion

In this paper, I have argued that the Microsoft Kinect today in its well-received form is a product that has had a specific user identity configured within it since its inception as a gaming peripheral. This embedded user identity is instrumental in understanding the Kinect's growth arc and future trajectory. After enough failures in the entertainment market, Microsoft realized the device was best suited to be sold and marketed directly to the user it had configured for the device (D'Angelo, 2019).

It is important to recognize how impactful the configuration of user values was in this case in shaping the technology I have analyzed, evident by my use of user configuration as the lens through which I viewed the Microsoft Kinect. This argument should matter to those who claim the Kinect would naturally settle into its current place and position, as they are missing a key insight behind the reasoning for the Kinect's evolution. Lacking understanding of the designer's biases in the development of technology is a significant drawback to any technological analysis, and dangerously overlooks an important aspect of sociotechnical research: the designer inadvertently influences a product just as they do intentionally. Beyond those involved in the field of the technology being discussed, this argument is significant to any creative engineer or curious-minded individual seeking an understanding that is deeper than the surface-level of a particular product. The case study of the Microsoft Kinect can be applied to a plethora of other artifacts in our society, as the story of how they came to be is one where the biases of the designers often intensify through generations until finally taking grasp of the product entirely. Having this lens of analysis is a valuable inspective tool towards understanding the foundations of the tools, devices, and technologies we utilize daily around us.

Reference List

Branscombe, M. (2018, October 9). How Microsoft is making its most sensitive HoloLens depth sensor yet. Retrieved from https://www.zdnet.com/article/how-microsoft-is-making-its-most-sensitive-hololens-depth-sensor-yet/

D'Angelo, W. (2019, April 1). Switch vs PS4 vs Xbox One Global Lifetime Sales January 2019. Retrieved from https://www.vgchartz.com/article/436094/switch-vs-ps4-vs-xbox-one-global-lifetime-salesjanuary-2019/

Farley, A. (2020, January 30). Microsoft Stock At All-Time High After Another Strong Quarter. Retrieved from https://www.investopedia.com/microsoft-stock-at-all-time-high-after-strong-quarter-4783544

Hanson, K. (2018, March 8). Xbox One's Kinect was Ahead of its Time, so why did it Fail? Retrieved from https://attackofthefanboy.com/articles/xbox-ones-kinect-was-ahead-of-its-time-so-why-did-it-fail/

Keilman, J. (2016, June 16). Is my Xbox spying on me? Retrieved from https://www.chicagotribune.com/news/ct-toys-online-spying-keilman-hf-0106-20160101-column.html

Lardinois, F. (2019, July 11). Microsoft's $399 Azure Kinect AI camera is now shipping in the US and China. Retrieved from https://techcrunch.com/2019/07/11/microsofts-399-azure-kinect-ai-camera-is-now-shipping-in-the-u-s-and-china/

Metz, C. (2018, January 10). Forget Apple vs. the FBI: WhatsApp Just Switched on Encryption for a Billion People. Retrieved from https://www.wired.com/2016/04/forget-apple-vs-fbi-whatsapp-just-switched-encryption-billion-people/

Microsoft Corp (2011). BUILD 2011 Keynote. Retrieved from https://channel9.msdn.com/Events/BUILD/BUILD2011/KEY-0001

Microsoft Corp (n.d.). Azure Kinect DK – Develop AI Models: Microsoft Azure. Retrieved from https://azure.microsoft.com/en-us/services/kinect-dk/

Oudshoorn, N. & Pinch, T. (2003). Introduction: how users and non-users matter. In N. Oudshoorn and T. Pinch (Eds.), How users matter (pp. 1-25).

Pagliari, D., & Pinto, L. (2015). Calibration of Kinect for Xbox One and Comparison between the Two Generations of Microsoft Sensors. Sensors (14248220), 15(11), 27569–27589. https://doi.org/10.3390/s151127569

Patel, N. (2013, November 8). Xbox, watch TV: inside Microsoft's audacious plan to take over the living room. Retrieved from https://www.theverge.com/2013/11/8/5075216/xbox-one-tv-microsofts-plan-to-take-over-the-living-room

Plant, M. (2018). Top 10 best-selling videogame consoles. Retrieved from https://guinnessworldrecords.com/news/2018/12/top-10-best-selling-videogame-consoles-551938/

Procházka, A., Vyšata, O., Vališ, M., Ťupa, O., Schätz, M., & Mařík, V. (2015). Use of the image and depth sensors of the Microsoft Kinect for the detection of gait disorders. Neural Computing & Applications, 26(7), 1621–1629. https://doi.org/10.1007/s00521-015-1827-x

Protalinski, E. (2014, July 15). Kinect for Windows v2 Sensor and SDK 2.0 Preview Released. Retrieved from https://thenextweb.com/microsoft/2014/07/15/microsoft-starts-shipping-kinect-windows-v2-sensor-releases-free-sdk-2-0-public-preview/

Rigney, R. (2013). Xbox 180: Microsoft Fully Reverses Xbox One's DRM Policies. Retrieved from https://www.wired.com/2013/06/xbox-one-drm/

Shankland, S. (n.d.). Why iPhone and Pixel cameras are so damn good: Computational photography. Retrieved from https://www.cnet.com/news/computational-photography-why-google-says-its-pixel-4-camera-will-be-so-damn-good/

Sottek, T. C., & Kopfstein, J. (2013, July 17). Everything you need to know about PRISM. Retrieved from https://www.theverge.com/2013/7/17/4517480/nsa-spying-prism-surveillance-cheat-sheet

Statt, N. (2014, June 5). Microsoft: Xbox One is 10 percent more powerful without Kinect. Retrieved from https://www.cnet.com/news/microsoft-xbox-one-is-10-percent-more-powerful-without-kinect/

Warren, T. (2013, September 20). Microsoft looks to Apple and Google to reinvent itself. Retrieved from https://www.theverge.com/2013/9/20/4749574/microsoft-business-model-mix-of-google-and-apple

Warren, T. (2012). 10.6-inch Microsoft Surface tablets announced, powered by Windows 8. Retrieved from https://www.theverge.com/2012/6/18/3094157/new-microsoft-surface-windows-tablet