White Rose Research Online
eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of York.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/164402/

Version: Accepted Version

Proceedings Paper:
Vatavu, Radu Daniel, Saeghe, Pejman, Chambel, Teresa et al. (2020) Conceptualizing Augmented Reality Television for the Living Room. In: Proceedings of IMX '20: ACM International Conference on Interactive Media Experiences. 2020 Proceeding. ACM, New York, United States, pp. 1-12.

https://doi.org/10.1145/3391614.3393660

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

Conceptualizing Augmented Reality Television for the Living Room

Radu-Daniel Vatavu
MintViz Lab | MANSiD Center
University Ştefan cel Mare of Suceava
Suceava, Romania
radu.vatavu@usm.ro

Pejman Saeghe
Department of Computer Science
The University of Manchester
Manchester, United Kingdom
pejman.saeghe@manchester.ac.uk

Teresa Chambel
LASIGE, Faculdade de Ciências
Universidade de Lisboa
Lisbon, Portugal
mtchambel@ciencias.ulisboa.pt

Vinoba Vinayagamoorthy
BBC R&D
London, United Kingdom
vinoba.vinayagamoorthy@bbc.co.uk

Marian Florin Ursu
DC Labs, University of York
York, UK
marian.ursu@york.ac.uk

ABSTRACT

We examine the concept and characteristics of "Augmented Reality Television" (ARTV) using a four-step investigation method consisting of (1) an analysis of commonly-accepted perspectives on Augmented and Mixed Reality systems, (2) a literature survey of previous work on ARTV, (3) relevant connections with other areas of scientific investigation from TVX/IMX, such as Ambient Media, Interactive TV, and 3-D TV, and (4) by proposing a conceptual framework for ARTV called the "Augmented Reality Television Continuum." Our work comes at a moment when the excitement and hype about the potential of AR for home entertainment has overlooked rigorous analysis and clear-cut examinations of the concepts involved, which should be the hallmark of any exact science. With this work, our goal is to draw the community's attention toward fundamentals and first principles of ARTV and to tease out its salient qualities on solid foundations.

CCS CONCEPTS

• **Human-centered computing → Mixed / augmented reality**; **HCI theory, concepts and models**.

KEYWORDS

Augmented Reality; Mixed Reality; Interactive Television.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

Submitted for review to TVX/IMX 2020, ,

© 2020 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

ACM Reference Format:
anonymous author, anonymous author, anonymous author, anonymous author, and anonymous author. 2020. Conceptualizing Augmented Reality Television for the Living Room. In *Proceedings of (Submitted for review to TVX/IMX 2020).* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

New forms of computer-generated content, media, and platforms [4,34,44,60], anchored into and aligned with the physical world, continue to change how we perceive and interact with our surroundings. As display and sensing technology evolve, presentation of content that is photorealistic, adaptive, personalized and customizable, real-time responsive and interactive and, consequently, engaging for users becomes attainable. These developments impact how we experience various representations of reality, such as virtual [34], mediated [44], multimediated [45], amplified [20], alternate [15], augmented [1,4], augmediated [45], mixed [51, 74], blended [87], extended [46], and cross-reality [60].

For instance, applied to television, Augmented Reality (AR) can immerse viewers into an interactive storytelling space that enables fantasy worlds to "break out" of the confines of the physical TV frame as well as to "bring in" aspects of the physical world. As part of the televised show, the action can extend to the living room [35,36,80,81] to create a sense of "actually being there." Therefore, the combination of AR and television creates the premises for a hybrid medium that opens new horizons for storytelling and engagement with interactive media and digital content. For instance, imagine watching a live soccer game, where the formation of players in each team is symbolized on a miniaturized model of the game field that is rendered right in front of the TV screen, while statistics of the match are always visible right next to the TV set; see [78] (minute 02:05). Or, consider extending the field of view of the conventional TV set with synchronized video projections on the wall behind it [36,39], or putting up on the wall as many virtual TV screens as one wishes, and be able to adjust their location and size to match the architectural design of the room [80] with flexible control over which content is rendered where [81]. Moreover, imagine a detective TV series, where indispensable parts of the story are told through film with established editing conventions, but before the end of each episode, the viewers' living room is transformed into the crime scene, and viewers can continue to experience the plot of the story at a new level by searching for clues with their AR-enabled smart devices, before the final resolution of the episode.

By putting television and AR together in the form of ARTV (*read*: "Augmented Reality Television"), experiential rich scenarios can become the norm of living room TV-based entertainment [23,64,65]. However, while we are seeing more and more developments in AR for television presented and discussed at top-tier venues, such as CHI, UIST, TVX, and ISMAR [35,36,63–65,68,69,81,84,85], there is large heterogeneousness in the ARTV landscape as well as in the terminology used by researchers and practitioners, revealing concepts that are not crisply defined and fully understood. A more rigorous description of *what ARTV actually is* represents thus a requirement for our community to sustain growth in this area on solid foundations. A closer look at the larger context of AR research reveals the timely need for this requirement.

The Context in AR Research for ARTV

A key observation is that *it is not just ARTV and not just the TVX/IMX community that lack the desired level of rigorousness in properly specifying computer-generated augmented realities.* In fact, the terminology used in the scientific literature, industry, and media to refer to Augmented Reality (AR), Mixed Reality (MR), and Extended or Cross-Reality (XR) is split to the extent to which it has become difficult even for experts to define these concepts precisely and confidently [74].

There are several articles and blogs attempting to clarify the differences between AR, MR, VR, XR, and other "R" acronyms for the layman reader [27,33,46,71]. However, major industrial players refer to comparable technologies using different terms. For example, Google has adopted the "Augmented Reality" terminology when addressing its community of developers, speaking about "augmented reality experiences" enabled by the ARCore platform [54]; Microsoft promotes the term "Mixed Reality" with the Windows Mixed Reality platform and the HoloLens headset [49]; while other industry players, such as Qualcomm, speak directly about XR and envision *"the convergence of the smartphone, mobile VR headset, and AR glasses into a single XR wearable"* [66]. Although a few attempts have been made to clarify the terminology [79] and to demystify the VR landscape [32], empirical evidence suggests that the interpretation of relevant terminology by experts in the field remains varied [74].

The Responsibility of the TVX/IMX Community

As the "R" realities are not crisply defined, the term television itself is starting to become a troublesome one, in need for an updated definition. In this context, ARTV emphasis on new and must be properly characterized in a space where television meets computer-generated realities. Exactly because TV and "R" realities are converging, *it is our responsibility as a community to provide an informed answer for what is the new emerging concept of ARTV* to keep up with the expectations horizon created by a fast developing industry, a creative academia, and an excited public media [71,72,77,78].

To understand the hype and context in which our contribution for specifying ARTV is needed, we provide a few examples. In a January 2019 Visual Capitalist article piece, Jenny Scribani [71] noted that *"XR brings immersive experiences to the entertainment world, and offers consumers an opportunity to virtually experience live music and sporting events from the comfort of their VR headset"*; on their web page dedicated to XR, Qualcomm talks about how *"XR could replace all the other screens in your life, like that big TV in your living room"* [66]; and, during Facebook's F8 Developer Conference of 2017, Mark Zuckerberg touched on augmented television, among other examples of how AR technology could change users' lives: *"You want to watch TV? We could put a digital TV on that wall and instead of being a piece of hardware, it's a $1 app, instead of a $500 piece of equipment"*; see the video recording of the event at [77] (minute 4:35).

It seems though that excitement and hype about AR, MR, and XR technology for home entertainment, television included, has largely overpassed rigorous examination and understanding of the intrinsic concepts of Augmented Reality and television. However, to advance on scientific grounds, we need rigorous *conceptual formalization of what ARTV is.* This paper is an attempt toward such a formalization.

Contributions of This Work

Our practical contributions are as follows:

1. We conduct the first literature survey on AR for television in order to understand past efforts, critical ideas, and key projects. We examine various perspectives on AR/MR, from which we extract key characteristics for ARTV and identify areas of investigation from TVX/IMX related to ARTV, *e.g.*, Ambient Media, Interactive TV, and 3-D TV.
2. We introduce a conceptual framework for ARTV in the form of the "Augmented Reality Television Continuum," a 2-D representation space for ARTV concepts, devices, systems, and applications inspired by the 1-D Reality-Virtuality Continuum of Milgram and Kishino [51]. We also differentiate ARTV from AVTV (*read*: "Augmented Virtuality Television"), a complementary concept equally covered by our ARTV conceptual framework.

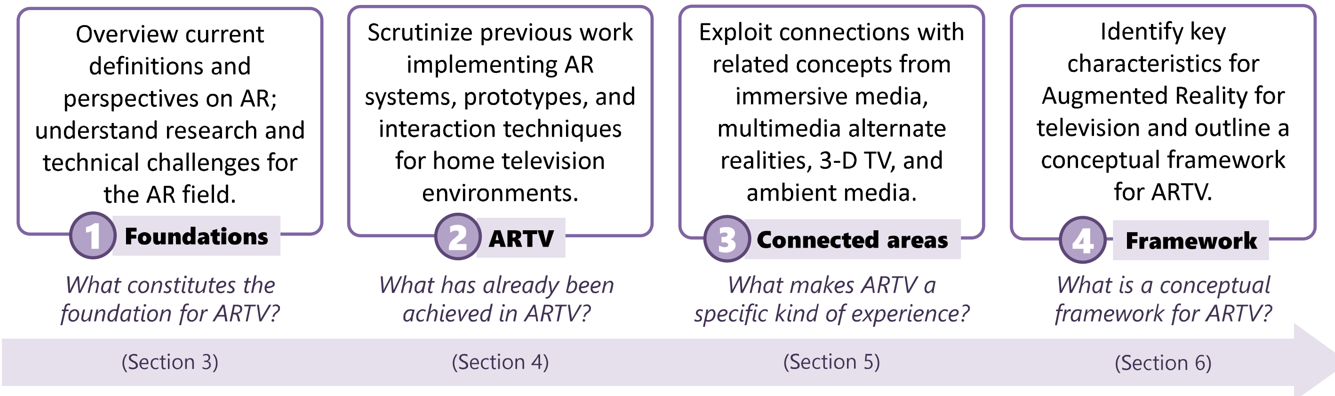

Figure 1: Our four-step method for specifying Augmented Reality for Television. Each step addresses a specific question and is covered by a distinct section of this paper, *e.g.*, Section 6 discusses *What is a conceptual framework for ARTV?*

2 THE INVESTIGATION METHOD OF THIS PAPER

Before we move on, we take a moment to describe the investigation method that we implement in this work to identify key characteristics of ARTV. Our method is composed of four steps, as follows (see Figure 1 for a visual illustration):

Step_1: We start from the perspectives and definitions of general-purpose AR and MR systems, examine current classifications, and overview research and technical challenges. This step enables us to provide an answer to *What are the foundations for ARTV?*

Step_2: We perform a targeted literature survey of previous work implementing AR for television. This step enables us to form an understanding of *What has already been achieved in ARTV?*

Step_3: Having established the foundations and understood the state-of-the-art, we proceed to identifying relevant connections between ARTV and other areas of investigation from TVX/IMX, such as Multimedia Alternate Realities, Ambient Media, Interactive TV, and 3-D TV. This step enables us to understand *What makes ARTV a specific kind of television experience?*

Step_4: We identify key components for ARTV, which we build into our new conceptual framework, the "Augmented Reality Television Continuum."

3 FOUNDATIONS: AUGMENTED, MIXED, AND VIRTUAL REALITIES

We overview in this section first principles of VR, AR, and MR systems and environments. While VR is commonly understood as fully-immersive environments that substitute real-world sensations with simulated cues [45,51,71], the distinction between AR and MR has been less obvious [74].

The transition from fully-immersive VR toward displays that combine elements from both the physical (real) world and the virtual (computer-generated) world was represented by the first prototypes of AR systems starting with Sutherland's [76] head-mounted 3-D display; see Azuma's [1] widely referenced survey of the state-of-the-art in AR in 1997 and Billinghurst *et al.*'s [10] comprehensive overview of the field photographed in 2015. However, since the introduction of the concept, AR has received many definitions in the scientific literature corresponding to various perspectives, which are relevant to scrutinize for the purpose of our investigation. This examination of the foundations of computer-generated realities is key to draw implications for ARTV. We start with Milgram and Kishino's [51] highly-influencing "Reality-Virtuality Continuum" that distinguished MR from AR.

The Reality-Virtuality Continuum

In 1994, Milgram and Kishino [51] introduced the "Virtuality" continuum, an imaginary line having the real and virtual worlds at its opposite ends. Later, Milgram *et al.* [50,52] referred to this line as the "Reality-Virtuality (RV) Continuum," the name in use today. As one moves along the RV Continuum, the degree of interpolation between the real and the virtual changes, leading to Augmented Reality (AR) and Augmented Virtuality (AV) world mixtures; see Figure 2. In fact, the primary environment or "substratum" that is augmented determines the distinction between AR and AV. Everything in the RV Continuum, except its ends, was defined as Mixed Reality (MR), a *"more encompassing term to supplement the existing definition of Augmented Reality (AR), which leads us to propose definitions of the associated concepts of Augmented Virtuality (AV) and then Mixed Reality (MR)"* [50].

Figure 2: Illustration of Milgram *et al.*'s [50–52] RV Continuum, which we reproduce in this work since it represents the foundation for our ARTV Continuum from Figure 4.

Definition of AR	*Perspective*	*Impacted by and/or impacts on*	*Implication for ARTV*
2019: Interface and gateway	**Integration**	**Cloud Computing: SaaS, PaaS, IaaS, XaaS & ARCloud**	Specific gateway and integration of TV genres
2017: Immersive experience, illusion	**Experiential**	**Human-Computer Interaction**	Specific type of immersion
2016: A new form of media	**Content consumption**	**Social & Interactive Media**	Specific form of media
1997: Combines the real and the virtual, interactive in real-time, and registered in 3-D	**Technology independence**	**Applied AR**	3-D registered and interactive in real time
1997: A specific form of intelligence amplification	**Task assistance**	**Human-Computer Interaction**	Specific services
1997: A variation of Virtual Reality	**Computer-generated realities**	**Computer Graphics & Virtual Reality**	Anchored in the physical, primary substratum of the living room environment
1994: The left-most part of the RV continuum			
1968: Information surrounding users in 3-D, enabled by see-through head-mounted displays	**Virtual Reality**	**Computer Graphics**	Informative and contextual

Figure 3: Evolving perspectives of Augmented Reality (left, shown in chronological order) and implications for ARTV (right).

According to the RV Continuum, MR encompasses AR, but excludes VR: *"the most straightforward way to view a Mixed Reality environment, therefore, is one in which real world and virtual world objects are presented together within a single display, that is, anywhere between the extrema of the virtuality continuum"* [51]. However, MR and AR have received other connotations in the recent years. For example, in a 2019 study, Speicher *et al.* [74] reported six definitions for MR from interviews conducted with AR/MR/VR experts from academia and industry as well as from a literature survey. They concluded that *"MR can be many things and its understanding is always based on one's context... there is no single definition of MR and it is highly unrealistic to expect one to appear in the future."* The authors also recognized that *"it is extremely important to be clear and consistent in terminology and communicate one's understanding of MR in order to avoid confusion and ensure constructive discussion"* [74] (p. 12). Milgram and Kishino's [51] definition of MR based on the RV Continuum was the most frequent perspective found in [74]. Thus, we also adopt this perspective in our work.

Augmented Reality *vs.* Augmented Virtuality

Milgram and Kishino's [51] formalization of the RV Continuum unveiled the concept of AV, where the virtual world is the primary environment that is augmented. By analogy, we differentiate between Augmented Reality TV (ARTV) and Augmented Virtuality TV (AVTV). In the former case, it is the real world (*e.g.*, the living room) that is augmented with virtual objects; in the later, it is a virtual world that is augmented with real objects, *e.g.*, a video feed of a physical TV set. In this work, we are interested in ARTV and, thus, we continue our discussion with an overview of definitions for AR. Section 6 resumes AVTV for the interested reader.

Perspectives on AR and Implications for ARTV

Several definitions have been proposed for AR in the scientific literature and it is essential for the purpose of this work to scrutinize them in order to understand the foundations of AR and to extract implications for ARTV. In the following, we present a chronological examination of evolving perspectives on AR, from which we identify key concepts, implications for ARTV, and connected areas of scientific investigation from TVX/IMX; see Figure 3 for an overview.

1968: AR is 3-D information surrounding the user, enabled by wearing see-through HMDs. Sutherland's [76] head-mounted display (HMD) represents the first instance of an AR device and system. Although there is no mention of the term "Augmented Reality" in Sutherland's paper, one quote characterizes well the ultimate goal of Sutherland's work: *"our objective in this project has been to surround the user with displayed three-dimensional information"* (p. 757). In this context, AR is visualization of 3-D spatial information enabled by a see-through HMD, a perspective that will dominate AR research for decades. For instance, in their 1994 paper introducing the concept of Mixed Reality and the RV Continuum, Milgram and Kishino [51] remarked that the prominent use of the term AR was limited at that time to what they called "Class 3 type displays," which are HMDs with see-through capabilities [51] (p. 1322). In another work, Milgram *et al.* [52] presented a definition of AR as *"a form of virtual reality where the participant's head-mounted display is transparent, allowing a clear view of the real world,"* which had originated from a call for papers of a representative journal special issue on AR. Pointing to the limitations of such definitions, Milgram and Kishino [51] argued that the AR concept can equally be extended to other types of displays, an observation that Milgram *et al.* [50,52] resumed in their follow-up work on real and virtual world display integration.

> **Key AR concepts:** information, HMD, 3-D.
> **Implications for ARTV:** Augmented television renders relevant information in the 3-D space surrounding the viewer and, if physically present, the TV set, enabled by some dedicated visualization device.
> **Connected TVX/IMX area:** 3-D TV.

***1994: The real environment is "augmented" by means of virtual (computer graphic) objects*.** This perspective was adopted by Milgram and Kishino [51] as an operational definition for AR, because it was encompassing in terms of classes of displays compared to the previous, HMD-based approach. However, the generality of this definition generated terminology problems for Milgram and Kishino in relation to their Class 5 and Class 6 displays, *i.e.*, completely graphic display environments, that made the authors ponder about the nature of the primary environment that is actually augmented. The result was the distinction between AR and AV and, eventually, the introduction of the term MR to cover all types of mixtures between physical and virtual.

> **Key AR concepts:** virtual objects, computer graphics, real environment.
> **Implications for ARTV:** Augmented television superimposes virtual content onto the real environment, which is the primary substratum that is augmented.
> **Connected TVX/IMX area:** 3-D TV.

***1997: A variation of VR*.** According to this definition and perspective from Azuma [1], *"in contrast [to VR], AR allows the user to see the real world, with virtual objects superimposed upon or composited with the real world. Therefore, AR supplements reality, rather than completely replacing it."* Eight years later, in their book on spatial Augmented Reality (sAR), Bimber and Raskar [11] resumed this definition to specify the characteristics of AR by contrasting them with VR: *"The fact is that in contrast to traditional VR, in AR the real environment is not completely suppressed; instead it plays a dominant role. Rather than immersing a person into a completely synthetic world, AR attempts to embed synthetic supplements into the real environment (or into a live video of the real environment)."* In this antinomy, AR is meant to supplement the physical reality instead of completely replacing it [1], while focusing on embedding synthetic supplements into the physical world [11]. Like Milgram *et al.* [50–52] before them, Bimber and Raskar recognized the need to go beyond the technology of eye-worn and HMD displays, and defined sAR as *"new display paradigms [that] exploit large spatially-aligned optical elements, such as mirror beam combiners, transparent screens, or holograms, as well as video projectors"* [11].

> **Key AR concepts:** superimposed, composited, supplemented reality.
> **Implications for ARTV:** Augmented television supplements the content of the primary TV broadcast.
> **Connected TVX/IMX areas:** 3-D TV, Interactive TV (iTV), Ambient Media.

***1997: A specific form of "intelligence amplification"*.** In his 1997 survey of AR systems, applications, and developments, Azuma [1] provides two other perspectives on AR besides the antinomy with VR. One of them, more general, capitalizes on the utility of computers to assist users in their tasks, which can be interpreted as access to a higher level of human intelligence. From this perspective, AR implements a specific case of intelligence amplification [1]. The second perspective is more structured and systematic; see next.

> **Key AR concepts:** intelligence amplification, task assistance, the computer as a tool.
> **Implications for ARTV:** Augmented television assists viewers by providing new services.
> **Connected TVX/IMX areas:** Ambient Media, iTV.

***1997: A system that combines the real and the virtual, is interactive in real time, and registered in 3-D*.** Azuma [1] recognized the large influence of the HMD perspective on AR research and, to avoid limiting the field to a specific form of technology, proposed three essential properties of any AR system. These properties constituted into a structured and systematic definition of what AR is, which has been largely adopted since;[1] see, for example, Billinghurst *et al.*'s [10] survey of AR that cites Azuma's three characteristics right from the first paragraph of their chapter defining AR.

> **Key AR concepts:** real-time interactivity, 3-D.
> **Implications for ARTV:** Augmented television is registered in 3-D and interactive in real time.
> **Connected TVX/IMX area:** iTV, 3-D TV.

***2016: A new form of media*.** This perspective builds on a specific challenge that has been identified for AR systems. According to Azuma [2], *"The ultimate and most important challenge facing AR [...] is experiential in nature: How will we establish Augmented Reality as a new form of media, enabling new types of experiences that differ from established media? If AR is to become ubiquitous in consumer usage, then we must [...] [develop] new types of experiences that are engaging and compelling in different ways than traditional media such as books, movies, and even Virtual Reality."* By adopting the experiential perspective, AR systems can be seen

[1] Azuma's [1] 1997 paper has been cited over 10,000 times, https://scholar.google.com/scholar?um=1&ie=UTF-8&lr&cites=17196017931627326366

as implementers of a specific kind of (ambient) media with characteristics from both the real and the virtual world.

> **Key AR concepts:** new media, experience.
> **Implications for ARTV:** Augmented television is a new medium for a new digital media experience.
> **Connected TVX/IMX area:** Ambient Media.

2017: An immersive experience and an illusion. As a direct effect of superimposing virtual 3-D objects on top of the users' direct view of the physical world, AR generates the illusion of an immersive mix-world experience [3].

> **Key AR concepts:** immersion, experience, illusion.
> **Implications for ARTV:** Augmented television is an immersive experience.
> **Connected TVX/IMX areas:** Multimedia Alternate Realities.

2019: The interface and gateway to a 1:1 correspondence between the digital and the real world. This recent perspective from Azuma [4] converges the Internet (as the World Wide Web, cloud computing, and access to remote repositories of content and information) with AR, and contours the vision of an "AR Cloud" representing the gateway access to persistent virtual content attached to real locations.

> **Key AR concepts:** convergence, gateway, interface.
> **Implications for ARTV:** Augmented television is a gateway to cloud-based digital content and corresponding services for television.
> **Connected TVX/IMX area:** on-demand TV, iTV.

All these perspectives on AR highlight distinct concepts, such as *immersion*, *information*, *interface*, *experience*, useful to derive implications for ARTV, as we have been doing in this section. They are also useful to identify areas of scientific investigation from TVX/IMX that connect to ARTV; we relate to these areas in Section 5. For now, we continue our examination of ARTV with an overview of the field.

4 A LITERATURE SURVEY OF AR FOR TELEVISION

We overview prior work in AR for television and discuss concepts and/or implementations for AR and the living room.

Method

We conducted a targeted literature survey to locate relevant papers for ARTV. We identified a total number of 17 papers (from 338 candidates) describing ARTV systems or concepts (see Table 1) by running the following platform-compatible query against the ACM Digital Library (N=108 papers), Scopus (N=192 papers), and IEEE Xplore (N=38 papers):

```
(‘‘Augmented Reality’’) AND (Television OR TV)
```

We considered papers published in conference proceedings and journals during the last 20 years (2000 to 2019) that explicitly contained our key terms in their title, abstract, or keywords.[2] Based on our experience, we also considered for inclusion other papers that were not picked up by the query, from which we selected [35] as an extension of [36].

Next, we review these papers and classify them according to Milgram & Kishino's [51] three dimensions of the taxonomy for merging real and virtual worlds, as follows:

- *Extent of World Knowledge (EWK)* refers to how much about the world being modeled is known to the system.
- *Reproduction Fidelity (RF)* refers to the quality with which the synthesizing display reproduces the actual or intended images of the objects being displayed.
- *Extent of Presence Metaphor (EPM)* is the extent to which the viewer is intended to feel present within the scene.

We also identify for each paper the classes of the implemented AR displays, according to [51]: monitor-based video displays (Class 1), HMDs (Class 2), HDMs with see-through capability (Class 3), video-based see-through HMDs (Class 4), completely graphic display environments (Class 5), and completely graphic, but partially immersive environments, such as large displays, in which physical objects play a role in the virtual world (Class 6). Furthermore, inspired by previous initiatives to specify contributions in HCI [42,89], we group prior work according to their contributions to ARTV using the following categories: (1) new technology, (2) application, (3) method, (4) user study, and (5) design recommendations. Table 1 presents a summary of our literature survey.

Window-on-the-World Displays (Class 1)

Many of the papers identified in our literature survey implemented the "window-on-the-world" metaphor for augmented television using hand-held devices, such as smartphones and tablets. For example, Gómez *et al.*'s [26] system enabled users to browse a tree of resources linked to a physical object using AR on a hand-held device; as a use case, they augmented a map of the world displayed on a TV set, enabling the viewer to access extra information about specific map regions based on the viewer's distance from the TV. Kawakita and Nakagawa [37] created a system where a 3-D virtual TV character appeared to come out of the TV screen when viewed on a hand-held display. Revelle *et al.* [67] proposed a transmedia game for children to learn new words. Kawamura and Otsuki [38] presented an imaginary world on the TV screen which was equally viewable in 3-D free-viewpoint AR on a hand-held display. And Zimmer *et al.* [91] immersed viewers into a story by eliminating the TV set and delivering the content using AR exclusively.

[2] In cases where the same authors published evolving work in different venues over time, such as [63–65], we kept just the latest publication.

Reference	Contributions made					Extent of World Knowledge	Reproduction Fidelity	Extent of Presence Metaphor	Display class
	New tech	Application	Method	User study	Design				
Stauder and Robert (2002) [75]	✓	✓	✓	–	–	where	video / video	n.a.	1
Vatavu (2012) [80]	✓	✓	✓	✓	–	where	HD video / Projected video	Real time imaging	6
Jones *et al.* (2013) [36]	✓	✓	–	✓	–	where	HD video / Real-time 3D animation	Real time imaging	6
Vatavu (2013) [81]	✓	✓	✓	✓	✓	where	HD video / Projected video	Real time imaging	6
Jones *et al.* (2014) [35]	–	–	–	–	–	world completely modeled	HD video / Real-time 3D animation	Real time imaging	6
Gómez *et al.* (2014) [26]	✓	✓	✓	✓	–	where + what	HD video / simple graphics	Monoscopic imaging	1
Kawakita and Nakagawa (2014) [37]	✓	✓	✓	–	–	where + what	HD video / Real-time 3D animation	Real time imaging	1
Revelle *et al.* (2015) [67]	✓	✓	–	✓	–	where + what	HD video / Real-time 3D animation	Real time imaging	1
Vatavu (2015) [82]	✓	✓	✓	✓	✓	where	HD video / Visible surface imaging	Monoscopic imaging	1
Baillard *et al.* (2017) [5]	–	✓	–	–	–	where + what	HD video / Real-time 3D animation	Real time imaging	1 & 3
Sotelo *et al.* (2017) [73]	✓	✓	✓	–	–	n.a.	HD video / Real-time 3D animation	n.a.	1
Kawamura and Otsuki (2018) [38]	✓	✓	✓	–	–	world completely modeled	HD video / Real-time 3D animation	Real time imaging	1
Kimura and Rekimoto (2018) [39]	✓	✓	✓	✓	–	world unmodeled	HD video / Generated context images	n.a.	5
Zimmer *et al.* (2018) [91]	–	✓	✓	✓	✓	where	HD video / 3D animation	Monoscopic imaging	1
Popovici and Vatavu (2019) [65]	–	–	–	✓	–	n.a.	n.a.	n.a.	n.a.
Geerts *et al.* (2019) [24]	–	–	✓	✓	–	where	n.a.	n.a.	6
Saeghe *et al.* (2019) [69]	–	✓	–	✓	✓	where	HD video / Real-time 3D animation	Real time imaging	3
Vinayagamoorthy *et al.* (2019) [85]	–	✓	–	✓	✓	where	HD video / Holographic video	Monoscopic imaging	3

Table 1: Classification of ARTV-related papers according to their contributions, Milgram and Kishino's [51] display classes and dimensions, such as Extent of World Knowledge and Presence Metaphor. *Note:* papers are listed in chronological order.

Besides hand-held devices, Class 1 displays have been implemented with augmentations on the TV screen itself. For example, Stauder and Roberts [75] proposed a method to ensure photometric realism when virtual artifacts were overlaid on top of the TV content, and Sotelo *et al.* [73] described a system enabling the viewer to interact with a 3-D model overlaid on broadcast video. Vatavu's [82] "Audience Silhouettes" prototype overlaid live 3-D representations of viewers' body movements directly on top of the TV broadcast.

Head-Mounted Displays (Classes 2, 3, and 4)

A few systems have implemented ARTV using HMDs. One example is Baillard *et al.* [5] that created a multi-user system where AR content was displayed on the periphery of the physical TV set, viewable using both a hand-held and an HMD. Saeghe *et al.* [69] displayed virtual animated content related to the broadcast outside the TV frame. Vinayagamoorthy *et al.* [85] described a prototype where a sign language interpreter was presented to the viewer next to the TV set.

Graphic Display Environments (Classes 5 and 6)

In 2012, Vatavu [80] proposed an interactive home entertainment system where multiple virtual TV screens were video projected on the living room wall and independently controlled. In a follow-up paper, Vatavu [81] introduced "AroundTV," a video projection-based system for the area surrounding the physical TV set for interacting with virtual TV screens and graphical user interface widgets. At the same time, Jones *et al.* [36] introduced "IllumiRoom" for computer games, a system that projected visualizations of the game in the periphery of the TV screen. The follow-up "RoomAlive" [35] system and installation was meant to transform the entire room into an immersive, augmented space. To address peripheral vision, Kimura and Rekimoto [39] proposed "ExtVision," a system that generated and displayed context images in the area surrounding the TV set.

User Study Contributions in ARTV

In our survey, we found that nine papers implemented the window-on-the-world metaphor (Class 1 displays), three papers used see-through HMDs (Class 3), and six implemented graphic environments of Class 5 or 6; see Table 1. Some papers did not present actual implementations, since they focused on user studies. For example, Geerts *et al.* [24] used a co-design approach to generate a scenario where extra content was displayed in the viewers' environment. Popovici and Vatavu [64] proposed an agenda for researchers to consider when designing for ARTV. In two follow-up papers, they elicited user preferences regarding ARTV [65] and used the findings to consolidate the original research agenda [63].

World Knowledge, Presence, and Fidelity

We used Milgram & Kishino's [51] dimensions to classify the ARTV literature. We found that all system papers used information about the location (*where*) of the modeled world, four papers modeled both the *where* and *what*, while only two papers implemented a complete model of the world [35,38]. As the spatial understanding capability of AR advances, we can expect systems to approach complex models of the displayed world. While six papers described a real-time hi-fidelity 3-D animation, one paper used basic graphics [26] and another [85] employed holographic video via chroma-keying techniques. Regarding the Extent of Presence Metaphor dimension, four papers [26,82,85,91] used AR to deliver a monoscopic image, while the rest delivered real-time images with AR artifacts viable from multiple points of view.

5 TVX/IMX AREAS CONNECTED TO ARTV

Section 3 highlighted key concepts for ARTV, which led to connections to several areas of scientific investigation representative for the TVX/IMX community. We discuss these areas in this section to unveil further aspects of ARTV.

Interactive Television

Broadcasters are trying to engage TV audiences through the addition of data services on top of traditional television to increase participation and feedback [83]. This has resulted in the TV evolving from a purely audiovisual platform to one with in-built interactive services, such as teletext, electronic program guides, or red-button services. Users are increasingly opting to purchase smart TVs with IP connectivity that run applications and integrate with connected devices in the home, such as conversational user interfaces [16,88]. Besides the attractive prospect of personalizing the ways in which users might control their connected TVs, there have been ventures to personalize and augment the viewing experience itself [22,41,83] through synchronized companion experiences based on audio watermarking, fingerprinting, and HbbTV 2.0 [86]. The potential to personalize television experiences connects interactive TV with ARTV, where virtual objects augment viewers' personal TV watching experience.

On the big screen itself, the move from broadcast to streaming over IP enables more interactive storytelling. The go-to obvious format has been play along quizzes and voting, but the interactive potential can go further [31]. For example, Netflix Bandersnatch [55] is a non-linear branching narrative that progresses the stories off depending on the choice the viewer makes. Also, as content is increasingly consumed over IP, content creators know more about their audiences through the data they can collect. Unlike branching narratives, object-based media (OBM) allows content to change according to the requirements of each individual audience member [30]. This affords a versatile manner in which the story is remixed according to the audience [7,17,30,40], connecting to the potential of ARTV to render new ways for the virtual story to play out. Examples of OBM experiences include immersive audio [6,21], branching narratives [12], personalized documentary [29], and personality quizzes [8]. In ARTV, future media offerings will more easily personalize to the audience, which means adding interaction and augmentation to engage the audience in the storytelling process.

3-D Television

3-D photography, cinema, and TV have a long history from the first examples of passive stereoscopic 3-D cinema in the 1990s to manufacturers involved in marketing autostereoscopic 3-D TVs in the late 2000s [48,59]. The 3-D TV display systems use a combination of a 2-D image and a depth map [28] (*i.e.*, depth image-based rendering) to synthesize new virtual views and, consequently, to augment the original image from the 2D-plus-depth data [19]. These advances in 3-D TV technology are relevant to support developments in ARTV since, according to Azuma [1], 3-D registration is one of the three key characteristics of any AR system.

Ambient Media

Research in Ambient Media has unveiled a new type of media, conceptually different from television, print, and digital media, that define the media environment of smart spaces. According to Lugmayr *et al.* [43], *"Ambient media in a larger scale define the media environment and the communication of information in ubiquitous and pervasive environments."* Among its characteristics, ambient media is subtle, unmonopolizing, and addressing peripheral awareness [70], while it can morph and manifest in various ways [61,62]. These properties make ambient media especially relevant for ARTV where, according to a recent perspective [2], AR technology itself can be qualified as a specific form of new media.

Immersive Media and Multimedia Alternate Realities

Immersion, Interactivity, and Imagination (I^3) [13] are central concepts associated with the sense of presence in VR, AR, and MR with strong roots in computer-generated graphics and content. Immersion is influenced by sensory and perceptual modalities associated with "presence," *i.e.*, the feeling of being inside the computer-generated reality due to realistic feedback, participation, and social immersion [14]. Developments in immersive media in the context of interactive television in specialized communities and events, such as TVX/IMX (and, before it, EuroITV), ACM Multimedia, IEEE QoMEX, or general CHI/ISMAR venues, have focused on aspects such as audiovisual immersion, 3-D and panoramic 360° multi-view and holographic video, spatial and stereoscopic audio, perceptual immersion and multi-sensory interaction, interactive and immersive television, film, and cinema [14,15].

Chambel *et al.* [14,15] introduced the concept of "Multimedia Alternate Realities" (MMARs) as *"different spaces, times or situations that can be entered thanks to multimedia contents and systems, that coexist with our current reality, and are sometimes so vivid and engaging that we feel we are living in them ... immersive experiences that may involve the user in a different or augmented world, as an alternate reality"* [15]. To properly characterize such realities, a taxonomy with eight dimensions was proposed [15] consisting of: the alternate (*e.g.*, different space, time, context), the virtual-augmented spectrum, the real-fictional spectrum, the level of interactivity, the level of immersion (*e.g.*, presence and belonging, imagination, and engagement), the multisensorial (the media and modalities involved), the personal (adaptation to preferences and contexts), and the social dimension (individualized or shared realities). MMARs go beyond the focus of VR, AR, and MR by addressing new media and immersive experiences. In the context of ARTV, these dimensions become relevant when television is part of the audiovisual content delivered inside the MMAR. Of those, the virtual-augmented and real-fictional spectra connect directly to ARTV.

6 A CONCEPTUAL FRAMEWORK FOR ARTV

Based on our findings and discussion so far, we introduce a conceptualization for ARTV. To this end, we adopt the following three principles regarding AR technology, the real-world environment of the living room, and end users:

P_1. *Independence of AR rendering technology.* Just like Azuma [1] for AR, Bimber and Raskar [11] for spatial AR, and Milgram *et al.* [50–52] for MR, we believe that the technology to render computer-generated graphics in the living room environment should not drive the concept of ARTV. From this perspective, smartphones and tablets [37, 38,67], smart glasses and HMDs [84,85], video projections [80,81], and wall- and room-sized projections [35, 36,39] are implicitly encompassed by our framework. All that matters is that the TV experience is augmented.

P_2. *Focus on the living room.* We restrict our discussion of ARTV to the context of the living room environment. Thus, we exclude mobile television [47,58], for which a proper investigation will need the context of mobile AR [18]. Our focus on the living room is motivated by traditional TV being predominant in the TV landscape, *e.g.*, 93% of streamers watch traditional TV on a typical day [56] as the growth of mobile media levels out [90].

P_3. *Focus on the viewers' side of ARTV.* AR has two applications in TV broadcasting: (1) TV production, which aims to increase productivity and/or reduce costs, often referred to as the "virtual studio" [25], and (2) application on the viewers' side, which aims to create novel experiences in the viewers' environment during television watching. In this work, our focus has been on the latter.

The ARTV Continuum

Our exploration from the previous sections revealed many key characteristics of ARTV. We have seen that the RV Continuum [51] (Figure 2) represents the most accepted, go-to source for defining AR and to distinguish MR from AR [74]. In the following, we adopt the RV Continuum as the basis for our conceptualization of ARTV. But, while this continuum can be successfully employed to characterize the degree of mixture between real and virtual, its 1-D design is insufficient for our purpose. By considering the TV viewing experience where a *physical TV set* is placed in a *physical living room*, we recognize the fact that each of these two entities, world and TV, can be independently augmented. By adopting this perspective, we propose and introduce a 2-D conceptual framework for ARTV; see Figure 4.

The horizontal axis of Figure 4, going from a completely real to a completely virtual environment is Milgram and Kishino's [51] RV Continuum that applies to the living room environment (the world). The vertical axis, going from a completely physical TV device to fully virtual televised content is

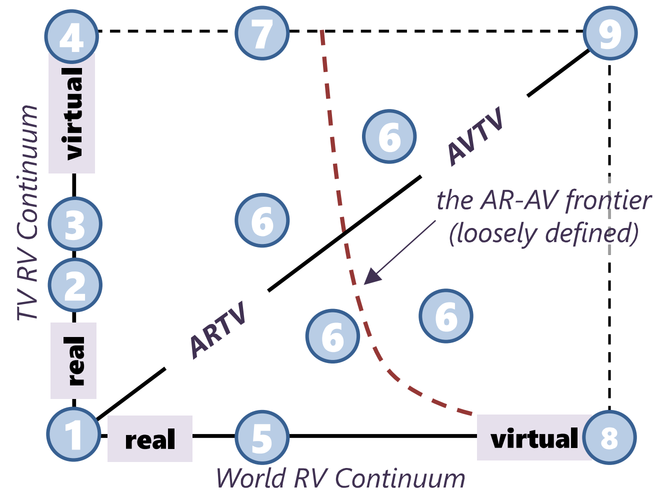

Figure 4: The ARTV Continuum. *Note:* the red dotted line loosely marks a delineation between ARTV and AVTV.

our adaptation of Milgram and Kishino's RV Continuum for television. Together, these two orthogonal axes characterize the various ways in which a television viewing experience can be augmented, *e.g.*, in terms of the world, the televised content, or both. In this conceptual framework, content can flexibly flow between the living room and the TV set, while the living room and the TV set can independently flow across their respective continua from real to virtual. We call this space the "Augmented Reality Television Continuum."

Examples of Using the ARTV Continuum

To demonstrate the utility of the ARTV conceptual framework, we enumerate various types of ARTV generated from the mixture of the two axes of Figure 4. For each category, we note examples from the literature where available and, where not, we highlight open areas for future work.

❶ *Conventional world/conventional TV.* This category represents the conventional TV viewing experience where neither the TV nor the room are augmented. The corresponding position in the ARTV Continuum is the bottom-left of Figure 4. Similar to how the RV Continuum captures the real world [51], the conventional TV viewing experience is equally captured by our framework.

❷ *Conventional world/conventional TV with on-TV augmentation.* In this category, contextual augmentations are shown on the TV screen. Examples include LinkedCulture [57], CollaboraTV [53], and Audience Silhouettes [82]. Teletext and Picture-in-Picture can also be included.

❸ *Conventional world/conventional TV with off-TV augmentation.* Different from the previous category, augmentation is now perceived to be off the TV screen, *e.g.*, on the wall behind it, in front or next to the TV set. Implementations may involve hand-held devices [37,38,67], HMDs [69,85], or video projection-based technology [35,36,39,81]. This category is also located on the vertical axis in Figure 4.

❹ *Conventional world/virtual TV.* A physical TV is not needed for this category of ARTV, where televised content is presented either in a virtual TV frame or even without a TV frame altogether. Examples include TV projected on the wall [77,80], and when the storyteller appears to be sitting in the living room [91] in the viewers' space.

❺ *Augmented world/conventional TV.* In contrast to the previous categories, the aim in this case is to augment the living room directly. This often requires spatial understanding of the surfaces present in the living room for meaningful alignment of the computer-generated graphics. Viewers experience an augmented world, but watch a conventional TV screen. Nevertheless, the augmented world can offer affordances with respect to television watching, *e.g.*, the possibility to record the TV show, open a mid-air browser window with a fact sheet about the current televised content, have Skype calls with friends watching the same broadcast, etc. Opposed to the previous categories, the living room is now the substratum that is augmented, but the TV is still conventional, which places this category on the horizontal axis; see Figure 4.

❻ *Augmented world/conventional TV with augmentation.* The world is augmented and so is the TV set, either by means of on-TV or off-TV augmentation. The degree of augmentation of each component, world and TV, positions implementations of this category at various locations in the ARTV Continuum illustrated in Figure 4.

❼ *Augmented world/virtual TV.* This category is achieved when the world is augmented (as in ❻), but the physical TV is virtual (as in ❹). These characteristics position category ❼ at the top of Figure 4. For instance, in the RoomAlive [35] prototype, the room is modeled and content is projected on its surfaces directly without using a physical TV set. The home entertainment prototype of Vatavu [80] implemented virtual TV screens exclusively.

❽ *Virtual world/conventional TV.* This category resembles the Augmented Virtuality of the RV Continuum [51]: a virtual world is augmented by a physical TV set. As a use-case, imagine watching your favorite TV show as a live video feed of your physical TV screen, while wearing a VR headset but still being physically present in the same room with friends watching the same physical TV screen.

❾ *Virtual world/virtual TV.* Similar to the previous category, with the exception that the physical TV set is replaced by one or more virtual screens. This includes the scenario where a virtual TV is aligned to a virtual model of a conventional TV set in order to recreate a familiar TV viewing experience in a virtual space. This category is located at the top-right of Figure 4.

Other scenarios can be imagined in our conceptual space, depending on the interpolation between real and virtual on both the world and the TV axes; see the multiple instances of ARTV category ❻ shown in Figure 4. We note that many of the categories that we discussed in this section haven't been proposed yet in the literature, which reveals the generative power [9] of our conceptual framework for ARTV.

It is also important to note that Figure 4 accommodates AVTV as well, represented by the region located at the intersection of the Augmented Virtuality part of the world continuum with the TV axis. However, just like in Milgram and Kishino's [50] case, the distinction between ARTV and AVTV can only be defined in loose terms, *e.g.*, *"As we venture away from the poles of the RV continuum towards the centre, we also eventually begin to encounter the problem of deciding whether in fact what we are doing is augmenting a real world with virtual graphic objects, or whether we are modifying a virtual environment by augmenting it with real data,"* and *"it it is not always [...] simple [...] to distinguish between AR and AV"* [50], which is equally true for our ARTV space.

7 CONCLUSION

We examined in this paper the concept of ARTV by following a four-step method starting from understanding the foundations of ARTV in AR to highlighting key concepts, performing a literature survey of ARTV and connected areas and, finally, to proposing a continuum for conceptualizing and structuring current and future work in ARTV.

We found that ARTV can be many things and that prior work has implemented it in various ways, from on-TV augmentations to off-TV content visualized via AR-enabled handheld devices, HMDs, wall- and room-sized projections, and holograms. By drawing from the various perspectives and key properties that we examined, we can conclude that *ARTV reveals itself as a specific type of experience, immersion, media, service, and gateway for televised content.* While the specific implementation may vary (and, in the years to come, we are to see more innovations in this regard), what is important at this moment is to have a rigorous basis to structure future research and developments, to be consistent in our terminology, and to communicate our understanding of ARTV precisely by relating to proper frameworks. Our ARTV Continuum is an attempt toward such a systematization and toward providing the community with a common vocabulary for possible categories of ARTV systems to enable better understanding and communication of advances in ARTV.

Our goal in this paper has been to provide researchers and practitioners with a conceptual framework in which to think about, talk about, and position their ARTV work and prototypes. We look forward to seeing how the community will employ our conceptual framework to inform the design of novel interactive augmented experiences for television.

REFERENCES

[1] Ronald T. Azuma. 1997. A Survey of Augmented Reality. *Presence: Teleoper. Virtual Environ.* 6, 4 (Aug. 1997), 355–385. https://doi.org/10.1162/pres.1997.6.4.355

[2] Ronald T. Azuma. 2016. The Most Important Challenge Facing Augmented Reality. *Presence: Teleoper. Virtual Environ.* 25, 3 (Dec. 2016), 234–238. https://doi.org/10.1162/PRES_a_00264

[3] Ronald T. Azuma. 2017. Making Augmented Reality a Reality. *Imaging and Applied Optics 2017 (3D, AIO, COSI, IS, MATH, pcAOP)*, JTu1F.1. http://www.osapublishing.org/abstract.cfm?URI=AIO-2017-JTu1F.1

[4] Ronald T. Azuma. 2019. The road to ubiquitous consumer augmented reality systems. *Human Behavior and Emerging Technologies* 1, 1 (2019), 26–32. https://doi.org/10.1002/hbe2.113

[5] Caroline Baillard, Matthieu Fradet, Vincent Alleaume, Pierrick Jouet, and Anthony Laurent. 2017. Multi-device Mixed Reality TV: A Collaborative Experience with Joint Use of a Tablet and a Headset. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology (VRST '17)*. ACM, New York, NY, USA, Article 67, 2 pages. https://doi.org/10.1145/3139131.3141196

[6] Chris Baume. 2017. The Mermaid's Tears. https://www.bbc.co.uk/rd/blog/2017-09-mermaids-tears-object-based-audio

[7] Rik Bauwens, Hendrik Lievens, Maarten Wijnants, Chris Pike, Iris Jennes, and Werner Bailer. 2019. Interactive Radio Experiences. In *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '19)*. Association for Computing Machinery, New York, NY, USA, 273–278. https://doi.org/10.1145/3317697.3323347

[8] BBC. 2019. His Dark Materials: Discover your Daemon. https://www.bbc.co.uk/rd/blog/2019-07-personalised-documentary-data-instagramifaction

[9] Michel Beaudouin-Lafon. 2004. Designing Interaction, Not Interfaces. In *Proceedings of the Working Conference on Advanced Visual Interfaces (AVI '04)*. Association for Computing Machinery, New York, NY, USA, 15–22. https://doi.org/10.1145/989863.989865

[10] Mark Billinghurst, Adrian Clark, and Gun Lee. 2015. A Survey of Augmented Reality. *Foundations and Trends in Human-Computer Interaction* 8, 2–3 (March 2015), 73–272. https://doi.org/10.1561/1100000049

[11] Oliver Bimber and Ramesh Raskar. 2005. *Spatial Augmented Reality: Merging Real and Virtual Worlds*. A. K. Peters, Ltd., Natick, MA, USA.

[12] Matthew Brooks. 2019. Creating a Personalised and Interactive Episode of BBC Click with StoryKit. https://www.bbc.co.uk/rd/blog/2019-04-object-based-media-click-interactivity-tv

[13] Grigore C. Burdea and Philippe Coiffet. 2003. *Virtual Reality Technology*. John Wiley & Sons.

[14] Teresa Chambel, V. Michael Bove, Sharon Strover, Paula Viana, and Graham Thomas. 2013. Immersive Media Experiences: Immersiveme 2013 Workshop at ACM Multimedia. In *Proceedings of the 21st ACM International Conference on Multimedia (MM '13)*. Association for Computing Machinery, New York, NY, USA, 1095–1096. https://doi.org/10.1145/2502081.2503831

[15] Teresa Chambel, Rene Kaiser, Omar A. Niamut, Wei Tsang Ooi, and Judith A. Redi. 2016. AltMM '16: Proceedings of the 1st International Workshop on Multimedia Alternate Realities. Association for Computing Machinery, New York, NY, USA. https://dl.acm.org/doi/proceedings/10.1145/2983298

[16] John Corpuz. 2019. 20 Remote Control Apps for Android Devices. https://www.tomsguide.com/uk/us/pictures-story/494-android-tv-remote-apps.html

[17] Jasmine Cox, Rhianne Jones, Chris Northwood, Jonathan Tutcher, and Ben Robinson. 2017. Object-Based Production: A Personalised Interactive Cooking Application. In *Adjunct Publication of the 2017 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '17 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 79–80. https://doi.org/10.1145/3084289.3089912

[18] Marco de Sá and Elizabeth F. Churchill. 2013. Mobile Augmented Reality: A Design Perspective. In *Human Factors in Augmented Reality Environments*, Huang W., Alem L., and Livingston M. (Eds.). Springer, New York, NY, USA. https://doi.org/10.1007/978-1-4614-4205-9_6

[19] F. de Sorbier, Y. Takaya, Y. Uematsu, I. Daribo, and H. Saito. 2010. Augmented reality for 3D TV using depth camera input. In *Proceedings of the 16th International Conference on Virtual Systems and Multimedia*. 117–123. https://doi.org/10.1109/VSMM.2010.5665956

[20] Jennica Falk, Johan Redström, and Staffan Björk. 1999. Amplifying Reality. In *Proceedings of the 1st International Symposium on Handheld and Ubiquitous Computing (HUC '99)*. Springer-Verlag, Berlin, Heidelberg, 274–280. https://doi.org/10.5555/647985.743879

[21] Jon Francombe. 2018. Vostok-K Incident: Immersive Audio Drama on Personal Devices. https://www.bbc.co.uk/rd/blog/2018-10-multi-speaker-immersive-audio-metadata

[22] David Geerts, Rinze Leenheer, Dirk De Grooff, Joost Negenman, and Susanne Heijstraten. 2014. In Front of and behind the Second Screen: Viewer and Producer Perspectives on a Companion App. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video (TVX '14)*. Association for Computing Machinery, New York, NY, USA, 95–102. https://doi.org/10.1145/2602299.2602312

[23] David Geerts, Evert van Beek, and Fernanda Chocron Miranda. 2019. Viewers' Visions of the Future. In *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '19)*. Association for Computing Machinery, New York, NY, USA, 59–69. https://doi.org/10.1145/3317697.3323356

[24] David Geerts, Evert van Beek, and Fernanda Chocron Miranda. 2019. Viewers' Visions of the Future. In *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '19)*. ACM, New York, NY, USA, 59–69. https://doi.org/10.1145/3317697.3323356

[25] Simon Gibbs, Constantin Arapis, Christian Breiteneder, Vali Lalioti, Sina Mostafawy, and Josef Speier. 1998. Virtual studios: An overview. *IEEE multimedia* 5, 1 (1998), 18–35. https://doi.org/10.1109/93.664740

[26] David Gómez, Ana M Bernardos, and José R Casar. 2014. A system to enable level-of-detail mobile interaction with augmented media objects. In *Proceedings of the 8th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*. IEEE, 346–351. https://doi.org/10.1109/IMIS.2014.47

[27] Lauren Goode. 2019. Get Ready to Hear a Lot More About 'XR'. https://www.wired.com/story/what-is-xr/

[28] O. Grau and V. Vinayagamoorthy. 2010. Stereoscopic 3D Sports Content without Stereo Rigs. *SMPTE Motion Imaging Journal* 119, 7 (Oct 2010), 51–55. https://doi.org/10.5594/J17314

[29] Nick Hanson. 2019. Creating a Personalised and Interactive Episode of BBC Click with StoryKit. https://www.bbc.co.uk/rd/blog/2019-07-personalised-documentary-data-instagramifaction

[30] Nick Hanson. 2019. Storytelling of the Future. https://www.bbc.co.uk/rd/blog/2019-02-storytelling-interactive-digital-drama

[31] Nick Hanson. 2019. Where Next For Interactive Stories? https://www.bbc.co.uk/rd/blog/2019-01-interactive-drama-stories-branching-narrative

[32] Intel. [n.d.]. Virtual Reality vs. Augmented Reality vs. Mixed Reality. https://www.intel.co.uk/content/www/uk/en/tech-tips-and-tricks/virtual-reality-vs-augmented-reality.html

[33] Kaitlyn Irvine. 2017. XR: VR, AR, MR - What's the Difference? https://www.viget.com/articles/xr-vr-ar-mr-whats-the-difference/

[34] Aric Jenkins. 2019. The Fall and Rise of VR: The Struggle to Make Virtual Reality Get Real. https://fortune.com/longform/virtual-reality-struggle-hope-vr/

[35] Brett Jones, Rajinder Sodhi, Michael Murdock, Ravish Mehra, Hrvoje Benko, Andrew Wilson, Eyal Ofek, Blair MacIntyre, Nikunj Raghuvanshi, and Lior Shapira. 2014. RoomAlive: Magical Experiences Enabled by Scalable, Adaptive Projector-Camera Units. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14)*. Association for Computing Machinery, New York, NY, USA, 637–644. https://doi.org/10.1145/2642918.2647383

[36] Brett R. Jones, Hrvoje Benko, Eyal Ofek, and Andrew D. Wilson. 2013. IllumiRoom: Peripheral Projected Illusions for Interactive Experiences. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 869–878. https://doi.org/10.1145/2470654.2466112

[37] Hiroyuki Kawakita and Toshio Nakagawa. 2014. Augmented TV: An augmented reality system for TV programs beyond the TV screen. In *Proceedings of the 2014 International Conference on Multimedia Computing and Systems (ICMCS)*. IEEE, 955–960. https://doi.org/10.1109/ICMCS.2014.6911158

[38] Yuki Kawamura and Kazuhiro Otsuki. 2018. 3D Free-Viewpoint Augmented Reality on a Second Display Synchronized with a Broadcast Program. In *Proceedings of the 4th International Conference on Communication and Information Processing (ICCIP '18)*. Association for Computing Machinery, New York, NY, USA, 273–277. https://doi.org/10.1145/3290420.3290449

[39] Naoki Kimura and Jun Rekimoto. 2018. ExtVision: Augmentation of Visual Experiences with Generation of Context Images for a Peripheral Vision Using Deep Neural Network. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 427, 10 pages. https://doi.org/10.1145/3173574.3174001

[40] Jie Li, Zhiyuan Zheng, Britta Meixner, Thomas Röggla, Maxine Glancy, and Pablo Cesar. 2018. Designing an Object-Based Preproduction Tool for Multiscreen TV Viewing. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (CHI EA '18)*. Association for Computing Machinery, New York, NY, USA, Article Paper LBW600, 6 pages. https://doi.org/10.1145/3170427.3188658

[41] Valentin Lohmüller and Christian Wolff. 2019. Towards a Comprehensive Definition of Second Screen. In *Proceedings of Mensch Und Computer 2019 (MuC'19)*. Association for Computing Machinery, New York, NY, USA, 167–177. https://doi.org/10.1145/3340764.3340781

[42] Víctor Manuel López Jaquero, Radu-Daniel Vatavu, Jose Ignacio Panach, Oscar Pastor, and Jean Vanderdonckt. 2019. A Newcomer's Guide to EICS, the Engineering Interactive Computing Systems Community. *Proc. ACM Hum.-Comput. Interact.* 3, EICS, Article Article 1 (June 2019), 9 pages. https://doi.org/10.1145/3300960

[43] Artur Lugmayr, Bjoern Stockleben, Thomas Risse, Juha Kaario, and Bogdan Pogorelc. 2013. New Business, Design and Models to Create Semantic Ambient Media Experiences. *Multimedia Tools Appl.* 66, 1 (Sept. 2013), 1–5. https://doi.org/10.1007/s11042-012-1239-1

[44] Steve Mann. 1999. Mediated Reality. *Linux J.* 1999, 59es (March 1999), 1. https://doi.org/10.5555/327697.327702

[45] Steve Mann, Tom Furness, Yu Yuan, Jay Iorio, and Zixin Wang. 2018. All Reality: Virtual, Augmented, Mixed (X), Mediated (X,Y), and Multimediated Reality. arXiv:cs.HC/1804.08386 https://arxiv.org/abs/1804.08386

[46] Bernard Marr. 2019. What Is Extended Reality Technology? A Simple Explanation For Anyone. https://www.forbes.com/sites/bernardmarr/2019/08/12/what-is-extended-reality-technology-a-simple-explanation-for-anyone/#6ffa4ba37249

[47] Jennifer McNally and Beth Harrington. 2017. How Millennials and Teens Consume Mobile Video. In *Proceedings of the 2017 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '17)*. Association for Computing Machinery, New York, NY, USA, 31–39. https://doi.org/10.1145/3077548.3077555

[48] L. M. J. Meesters, W. A. IJsselsteijn, and P. J. H. Seuntiens. 2004. A survey of perceptual evaluations and requirements of three-dimensional TV. *IEEE Transactions on Circuits and Systems for Video Technology* 14, 3 (March 2004), 381–391. https://doi.org/10.1109/TCSVT.2004.823398

[49] Microsoft. [n.d.]. Windows Mixed Reality. https://www.microsoft.com/en-gb/windows/windows-mixed-reality

[50] Paul Milgram and Herman Colquhoun Jr. 1999. *A Taxonomy of Real and Virtual World Display Integration.* Springer-Verlag, Berlin, Heidelberg.

[51] Paul Milgram and Fumio Kishino. 1994. A Taxonomy of Mixed Reality Visual Displays. *IEICE Transactions on Information and Systems* E77-D, 12 (December 1994), 1321–1329. https://search.ieice.org/bin/summary.php?id=e77-d_12_1321

[52] Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. 1995. Augmented Reality: A Class of Displays on the Reality-Virtuality Continuum. In *Proceedings of the Society of Photo-Optical Instrumentation Engineers 2351, Telemanipulator and Telepresence Technologies*, Vol. 2351. https://doi.org/10.1117/12.197321

[53] Mukesh Nathan, Chris Harrison, Svetlana Yarosh, Loren Terveen, Larry Stead, and Brian Amento. 2008. CollaboraTV: Making Television Viewing Social Again. In *Proceedings of the 1st International Conference on Designing Interactive User Experiences for TV and Video (UXTV '08)*. Association for Computing Machinery, New York, NY, USA, 85–94. https://doi.org/10.1145/1453805.1453824

[54] n.d. [n.d.]. ARCore Overview | Google Developers. https://developers.google.com/ar/discover

[55] Netflix. 2018. Black Mirror: Bandersnatch. https://www.netflix.com/gb/title/80988062

[56] Nielsen. 2018. Streamers Show Strong Ties with Traditional TV. https://www.nielsen.com/us/en/insights/news/2018/streamers-showstrong-ties-with-traditional-tv.html

[57] Lyndon Nixon, Lotte Belice Baltussen, and Johan Oomen. 2015. LinkedCulture: Browsing Related Europeana Objects While Watching a Cultural Heritage TV Programme. In *Proceedings of the 8th International Conference on Personalized Access to Cultural Heritage - Volume 1352 (PATCH'15)*. CEUR-WS.org, Aachen, DEU, 37–40.

[58] Virpi Oksman, Elina Noppari, Antti Tammela, Maarit Mäkinen, and Ville Ollikainen. 2007. Mobile TV in Everyday Life Contexts: Individual Entertainment or Shared Experiences?. In *Proceedings of the 5th European Conference on Interactive TV and Video (EuroITV '07)*. Springer-Verlag, Berlin, Heidelberg, 215–225. https://doi.org/10.1007/978-3-540-72559-6_23

[59] H. M. Ozaktas and L. Onural. 2010. *Three-Dimensional Television: Capture, Transmission, Display.* Springer, Heidelberg, Germany. http://dx.doi.org/10.1007/978-3-540-72532-9

[60] J. A. Paradiso and J. A. Landay. 2009. Guest Editors' Introduction: Cross-Reality Environments. *IEEE Pervasive Computing* 8, 3 (July 2009), 14–15. https://doi.org/10.1109/MPRV.2009.47

[61] Bogdan Pogorelc, Artur Lugmayr, Björn Stockleben, Radu-Daniel Vatavu, Nina Tahmasebi, Estefanía Serral, Emilija Stojmenova, Bojan Imperl, Thomas Risse, Gideon Zenz, and et al. 2013. Ambient Bloom: New Business, Content, Design and Models to Increase the Semantic Ambient Media Experience. *Multimedia Tools Appl.* 66, 1 (Sept. 2013), 7–32. https://doi.org/10.1007/s11042-012-1228-4

[62] Bogdan Pogorelc, Radu-Daniel Vatavu, Artur Lugmayr, Björn Stockleben, Thomas Risse, Juha Kaario, Estefania Constanza Lomonaco, and Matjaž Gams. 2012. Semantic Ambient Media: From Ambient Advertising to Ambient-Assisted Living. *Multimedia Tools Appl.* 58, 2 (May 2012), 399–425. https://doi.org/10.1007/s11042-011-0917-8

[63] Irina Popovici and Radu-Daniel Vatavu. 2019. Consolidating the Research Agenda of Augmented Reality Television with Insights from Potential End-Users. In *Proceedings of 2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct) (ISMAR '19)*.

73–74. https://doi.org/10.1109/ISMAR-Adjunct.2019.00033

[64] Irina Popovici and Radu-Daniel Vatavu. 2019. Towards Visual Augmentation of the Television Watching Experience: Manifesto and Agenda. In *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '19)*. ACM, New York, NY, USA, 199–204. https://doi.org/10.1145/3317697.3325121

[65] Irina Popovici and Radu-Daniel Vatavu. 2019. Understanding Users' Preferences for Augmented Reality Television. In *Proceedings of the 2019 International Symposium on Mixed and Augmented Reality (ISMAR '19)*. 269–278. https://doi.org/10.1109/ISMAR.2019.00024

[66] Qualcomm. [n.d.]. Extended Reality XR | Immersive VR | Qualcomm. https://www.qualcomm.com/invention/extended-reality

[67] Glenda Revelle, Emily Reardon, Kristin Cook, Lori Takeuchi, Rafael Ballagas, Koichi Mori, Hiroshi Horii, Hayes Raffle, Maria Sandberg, and Mirjana Spasojevic. 2015. Electric Agents: Combining Collaborative Mobile Augmented Reality and Web-Based Video to Reinvent Interactive Television. *Comput. Entertain.* 12, 3, Article 1 (Feb. 2015), 21 pages. https://doi.org/10.1145/2702109.2633413

[68] Asreen Rostami, Chiara Rossitto, and Annika Waern. 2018. Frictional Realities: Enabling Immersion in Mixed-Reality Performances. In *Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '18)*. Association for Computing Machinery, New York, NY, USA, 15–27. https://doi.org/10.1145/3210825.3210827

[69] Pejman Saeghe, Sarah Clinch, Bruce Weir, Maxine Glancy, Vinoba Vinayagamoorthy, Ollie Pattinson, Stephen Robert Pettifer, and Robert Stevens. 2019. Augmenting Television With Augmented Reality. In *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '19)*. ACM, New York, NY, USA, 255–261. https://doi.org/10.1145/3317697.3325129

[70] Albrecht Schmidt, Hans-Werner Gellersen, and Michael Beigl. 1999. Matching Information and Ambient Media. In *Proceedings of the Second International Workshop on Cooperative Buildings, Integrating Information, Organization, and Architecture (CoBuild '99)*. Springer-Verlag, Berlin, Heidelberg, 140–149. https://doi.org/10.1007/10705432_13

[71] Jenny Scribani. 2019. What is Extended Reality (XR)? https://www.visualcapitalist.com/extended-reality-xr/

[72] José Somolinos. 2019. Augmented Reality disrupts Television. https://medium.com/@josesomolinos/augmented-reality-disrupts-television-9d79ed01f27

[73] Rafael Sotelo, Jose Joskowicz, and Nicolás Rondan. 2017. Experiences on hybrid television and augmented reality on ISDB-T. In *2017 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*. IEEE, 1–5. https://doi.org/10.1109/BMSB.2017.7986231

[74] Maximilian Speicher, Brian D. Hall, and Michael Nebeling. 2019. What is Mixed Reality?. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, Article Paper 537, 15 pages. https://doi.org/10.1145/3290605.3300767

[75] J. Stauder and P. Robert. 2002. Let the sunshine on your screen: introducing augmented reality into interactive television. In *Proceedings. IEEE International Conference on Multimedia and Expo*, Vol. 1. 837–840 vol.1. https://doi.org/10.1109/ICME.2002.1035912

[76] Ivan E. Sutherland. 1968. A Head-Mounted Three Dimensional Display. In *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I (AFIPS '68 (Fall, part I))*. Association for Computing Machinery, New York, NY, USA, 757–764. https://doi.org/10.1145/1476589.1476686

[77] TIME. 2017. Facebook Reveals Augmented Reality, Virtual Reality and More at Annual F8 Conference. https://www.youtube.com/watch?v=YtUye84PuFY

[78] TWiT Tech Podcast Network. 2018. Augmented Reality on Your Television. https://www.youtube.com/watch?v=lpj90PkfPJ8

[79] Unity. [n.d.]. https://unity3d.com/what-is-xr-glossary

[80] Radu-Daniel Vatavu. 2012. Point & Click Mediated Interactions for Large Home Entertainment Displays. *Multimedia Tools and Applications* 59, 1 (2012), 113–128. https://doi.org/10.1007/s11042-010-0698-5

[81] Radu-Daniel Vatavu. 2013. There's a World Outside Your TV: Exploring Interactions Beyond the Physical TV Screen. In *Proceedings of the 11th European Conference on Interactive TV and Video (EuroITV '13)*. ACM, New York, NY, USA, 143–152. https://doi.org/10.1145/2465958.2465972

[82] Radu-Daniel Vatavu. 2015. Audience Silhouettes: Peripheral Awareness of Synchronous Audience Kinesics for Social Television. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video (TVX '15)*. ACM, New York, NY, USA, 13–22. https://doi.org/10.1145/2745197.2745207

[83] Vinoba Vinayagamoorthy, Penelope Allen, Matt Hammond, and Michael Evans. 2012. Researching the User Experience for Connected Tv: A Case Study. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems (CHI EA '12)*. Association for Computing Machinery, New York, NY, USA, 589–604. https://doi.org/10.1145/2212776.2212832

[84] Vinoba Vinayagamoorthy, Maxine Glancy, Paul Debenham, Alastair Bruce, Christoph Ziegler, and Richard Schäffer. 2018. Personalising the TV Experience with Augmented Reality Technology: Synchronised Sign Language Interpretation. In *Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '18)*. Association for Computing Machinery, New York, NY, USA, 179–184. https://doi.org/10.1145/3210825.3213562

[85] Vinoba Vinayagamoorthy, Maxine Glancy, Christoph Ziegler, and Richard Schäffer. 2019. Personalising the TV Experience Using Augmented Reality: An Exploratory Study on Delivering Synchronised Sign Language Interpretation. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 532, 12 pages. https://doi.org/10.1145/3290605.3300762

[86] Vinoba Vinayagamoorthy, Rajiv Ramdhany, and Matt Hammond. 2016. Enabling Frame-Accurate Synchronised Companion Screen Experiences. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video (TVX '16)*. Association for Computing Machinery, New York, NY, USA, 83–92. https://doi.org/10.1145/2932206.2932214

[87] John Waterworth and Kei Hoshi. 2016. *Human-Experiential Design of Presence in Everyday Blended Reality*. Springer, Switzerland. https://www.springer.com/gp/book/9783319303321

[88] Brian Westover. 2019. How to connect Alexa to your Samsung Smart TV. https://www.tomsguide.com/uk/us/samsung-smart-tv-alexa-how-to,news-29557.html

[89] Jacob O. Wobbrock and Julie A. Kientz. 2016. Research Contributions in Human-Computer Interaction. *Interactions* 23, 3 (April 2016), 38–44. https://doi.org/10.1145/2907069

[90] Zenith Media. [n.d.]. Media Consumption Forecasts 2018. https://www.zenithmedia.com/product/media-consumption-forecasts2018

[91] C. Zimmer, N. Ratz, M. Bertram, and C. Geiger. 2018. War Children: Using AR in a Documentary Context. In *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 390–394.