Revisiting Trends in Augmented Reality Research:
A Review of the 2nd Decade of ISMAR (2008–2017)

Kangsoo Kim, *Student Member, IEEE*, Mark Billinghurst, *Senior Member, IEEE*,
Gerd Bruder, *Member, IEEE*, Henry Been-Lirn Duh, *Senior Member, IEEE*,
and Gregory F. Welch, Senior Member, IEEE

Abstract—In 2008, Zhou et al. presented a survey paper summarizing the previous ten years of ISMAR publications, which provided invaluable insights into the research challenges and trends associated with that time period. Ten years later, we review the research that has been presented at ISMAR conferences since the survey of Zhou et al., at a time when both academia and the AR industry are enjoying dramatic technological changes. Here we consider the research results and trends of the last decade of ISMAR by carefully reviewing the ISMAR publications from the period of 2008–2017, in the context of the first ten years. The numbers of papers for different research topics and their impacts by citations were analyzed while reviewing them—which reveals that there is a sharp increase in AR evaluation and rendering research. Based on this review we offer some observations related to potential future research areas or trends, which could be helpful to AR researchers and industry members looking ahead.

Index Terms—Augmented reality, mixed reality, survey, trends.

✦

1 INTRODUCTION

In 2008, authors Zhou, Duh, and Billinghurst published an article summarizing the previous ten years of Augmented Reality (AR) research presented at the International Symposium on Mixed and Augmented Reality (ISMAR), the leading academic conference for AR [152]. From hundreds of ISMAR papers published during that period, they identified the key research areas of Tracking, Interaction, and Displays, and discussed important directions for future research in these areas. Their survey was historically valuable, and significant because it helped to highlight key themes of AR research, and topics that appeared to be ripe for new research. For example, five of the top ten papers published at ISMAR during their review period were about tracking techniques, whereas there were few papers on collaborative AR systems.

A wide range of compelling research has been presented at ISMAR in the ten years since the survey by Zhou et al. Inspired by their work we provide an updated review using the same approach. We began with the same topic categories including "Tracking", "Interaction", and "Display", but have added four more emerging topics, which we identified while reviewing the published papers such as "Reconstruction" and "Perception." We compare the trends we have observed during the 2nd decade of ISMAR to those of the 1st decade, and identify possible future research areas and trends. Overall the goal is to provide a substantive and useful review perspective on an exciting period of AR research.

We hope this paper will be helpful for new researchers and students in academia in summarizing the current research trends and finding interesting research topics that they want to focus on. However, it will also be useful for senior researchers and people from industry to help them see a big picture of AR research trends, particularly based on the trends of ISMAR. Overall, we hope that this research will help AR to be absorbed into our daily lives and influence humans positively in many different ways—e.g., the way that we think, feel, behave, or communicate in the future.

In the rest of this paper, we first describe our review methodology in Section 2, then we present a high level description of the ISMAR research topics in Section 3, which is followed by a meta review of ISMAR publications in Section 4. In Section 5, we present a review of the research on some of the major research topics that Zhou et al. covered ten years ago, i.e., "Tracking", "Interaction", and "Display", along with "Application" by reviewing the highest-cited or awarded ISMAR publications. In Section 6, we present new trends that were observed from our reviews of the last ten years of ISMAR publications, e.g., "Evaluation" and "Rendering". Finally some of our insights and future directions that we anticipate are presented in Section 7, before we conclude the paper in Section 8.

2 METHOD

In this survey paper, we follow the same method as used in Zhou et al.'s paper [152], namely reviewing the published conference papers and other related material from the conference proceedings of ISMAR'08 to ISMAR'17. This includes Full and Short papers up until 2014, and then ISMAR Conference and IEEE Transactions on Visualization and Computer Graphics (TVCG) Journal papers until 2017. Of course there are other venues for publishing Augmented/Mixed Reality (AR/MR) research, but we hope that the review of ISMAR can provide an excellent snapshot of the research taking place in the overall AR research field since ISMAR is the premier academic conference in the field. Our goal was also to compare and contrast the most recent decade of research to that summarized by Zhou et al.

Like Zhou et al., we focused particularly on the ISMAR Science & Technology track, providing an interesting snapshot of emerging research trends in AR/MR over the last ten years. We excluded posters from the review, as these are typically shorter and not normally reviewed as rigorously, and we also did not include papers from other tracks, such as the Arts and Humanities publications, which have a different scope from the papers covered by Zhou et al., and would have made direct comparisons difficult. Overall this left a total of 264 ISMAR papers that needed to be read and reviewed.

We collected all ISMAR papers over the past ten years and formalized our process to classify and review the papers. Specifically, our method to analyze the ISMAR papers included the (a) classification of all papers based on their major contributions to the ISMAR research topics described in Section 3, (b) collection of recent citation counts for all papers from Google Scholar, and (c) in-depth reviews of the most highly cited papers in the topic areas as well as those papers that received Best Paper or Honorable Mention Awards. Parts (a) and (b)

- *Kangsoo Kim, Gerd Bruder, and Gregory F. Welch are with the University of Central Florida. E-mail: kskim@knights.ucf.edu, {bruder, welch}@ucf.edu.*
- *Mark Billinghurst is with the University of South Australia. E-mail: mark.billinghurst@unisa.edu.au.*
- *Henry Been-Lirn Duh is with La Trobe University. E-mail: b.duh@latrobe.edu.au.*

were performed in a first step of the classification and review process. In the second step, we divided the topic areas among the first three authors based on their experience and we reviewed the papers in depth with a view on impactful research.

In reviewing these papers we sought to answer the following key questions:

(1) Have the research trends changed from the first decade of ISMAR?

(2) Are there any new research areas that have been introduced?

(3) What are the key developments and challenges in the research areas?

3 ISMAR RESEARCH TOPICS

In Zhou et al.'s survey [152], they grouped all of the ISMAR publications into 11 AR research topics, including primary and secondary topics. Zhou et al. based their selection of topics on their own research experience and related surveys such as [5, 8]. They further found that the five primary topics Tracking Techniques, Interaction Techniques, Calibration and Registration, AR Applications, and Display Techniques are the core AR technology areas needed to deliver an AR application. The remaining topics reflected more emerging research interests. In order to be consistent, we used the same topic categories in our work. The topic categories are:

(a) Tracking Techniques: methods for tracking a target object/environment via cameras and sensors, and estimating viewpoint poses.

(b) Interaction Techniques and User Interfaces: techniques and interfaces for interacting with virtual content.

(c) Calibration and Registration: geometric, or photometric calibration methods, and methods to align multiple coordinate frames.

(d) AR Applications: research on AR systems in application domains such as medicine, manufacturing, or military, among others.

(e) Display Techniques: research on display hardware to present virtual content in AR, including headworn, handheld, and projected displays.

(f) Evaluation/Testing: research focusing on human-subject studies evaluating AR techniques or systems.

(g) Mobile/Wearable AR: research on AR applications and techniques for wearable and mobile platforms, such as tablets and smartphones.

(h) AR Authoring: research on methods, techniques, and systems for authoring virtual content in AR.

(i) Visualization: research into methods that use AR to make complex 2D/3D data easier to navigate through and understand.

(j) Multimodal AR: research combining different input (and output) modalities together, such as combined speech and gesture interfaces.

(k) Rendering: research into techniques for computer graphics rendering; and other sensory modalities, such as sound and haptics.

While conducting the review of the ISMAR papers, we found that several further topics related to emerging research interests appeared over the past ten years, which extend Zhou et al.'s list. We formalized the process to decide which new categories to add by collecting potential keywords, assigning them to and cross-checking them among the first three authors, before we finally decided on the most common and impactful keywords as new categories. These categories are:

(l) Perception: research investigating human perception/cognition in AR.

(m) Collaboration/Social: research on interactive collaborative systems for multiple remote or co-located users.

(n) Reconstruction: research on methods that automatically generate 3D virtual environments/objects based on images or other forms of data collected from the real environment/objects.

(o) Modeling: research on methods for creating virtual content via virtual primitives and tools with a human user's involvements.

4 META-REVIEW ANALYSIS OF ISMAR PROCEEDINGS

In this section, we present the results from a high level meta-review based on the number of papers and citations over the last ten years for each category, which we defined in Section 3. The citation counts are based on Google Citation Index on February 2, 2018. As we addressed, we have a total of 15 ISMAR research topics that we want to discuss in this paper including 11 original categories and 4 additional ones. First, we evaluated the number of papers for each category and their percentage over the total number of classifications (see Table 1). We should note that most papers are not limited to a single topic, but generally cover multiple topics; thus, the total classification count is much larger than the number of published papers—439 classifications when only the original 11 categories are considered and 526 when the new categories are included among 264 published papers.

In 2008, the most frequent five research topics were Tracking (20.1% over 313 classified categories), Interaction (14.7%), Applications (14.4%), Calibration (14.1%), and Display (11.8%). However, as seen in Table 1, there were some changes in this by 2018. One major change is the increase of Evaluation research (to 16.4% over 439 classified categories), although the most frequent topic is still Tracking (19.4%). One reason behind the dramatic increase of Evaluation research might be because AR technology is getting mature and close to real users, so new methods or systems need to be evaluated with real users. Also, with the maturity of the AR research, ISMAR community requires more rigorous evaluation when they accept papers to publish. The Applications category (12.5%) is still one of the most popular research areas, and Rendering research (12.5%) has also dramatically increased in popularity. One reason for the growth in Rendering papers could be the recent trend in acquiring information from the real scene with advanced sensing devices and using it for more believable graphical renderings in AR. Finally the frequency of Interaction research (11.4%) dropped down a bit from the third in 2008 to the fifth in 2018, but it is still one of major research topics. Interestingly, Calibration disappeared from the top five list, which might be because many calibration issues have been resolved and the topic is less popular than some emerging topics. The overall research trend changes by year can be observed in Figure 1.

Regarding the research impacts, we compared the same metrics that Zhou et al. used—the proportion of papers that have more than five citations in each category over all the papers that have more than five citations per year (see Table 2), but also added a new graph for the average citation counts per year for each category (see Figure 2). The most cited paper [99], which was classified as Tracking and Reconstruction research, was excluded as an outlier from Figure 2 because it has an exceptionally high number of citations per year (301.43/year) compared to the other papers (e.g., the citation count for the second highest paper [73] was 55.56/year).

In 2018, the most cited five research topics are Mobile, Reconstruction, Tracking, AR Applications, and Evaluation according to the average cites per year. Generally the impact reflects the publication frequency as we see Tracking, Applications, and Evaluation research are included in the top five research topics with the highest cites per year. One interesting observation is how influential Reconstruction and Mobile research has become, with high citation numbers. This could be because over the last ten years mobile devices with high computational power and sensing modules have become affordable and ubiquitous. Many AR researchers are interested in how to achieve AR on mobile platforms and how to generate 3D models reconstructed from arbitrary environments using these devices. We will discuss more details about this throughout the paper, particularly in Section 5.1 and Section 7.1.

5 THE UPDATES ON PREVIOUS TRENDS

As described in the previous section, the research and publication areas of ISMAR have changed between the first and second decades of the ISMAR conference. Some research topics have received less attention than before, and new research topics have emerged. In this section, we first present the recent updates on the research areas that Zhou et al. identified: Tracking, Interaction, and Display. We also cover AR Applications presented for the second decade, and will then

Table 1. Research topic classification results—paper counts and percentage of each category.

Year	'08	'09	'10	'11	'12	'13	'14	'15	'16	'17	Total	% (w/ New Category)	%	Zhou08 (%)
# of Papers	26	24	24	26	28	26	33	22	24	31	264	-	-	276
Category														
Tracking	7	8	9	13	10	10	6	7	7	8	85	16.2	19.4	63 (20.1)
Interaction	6	7	6	4	6	3	6	3	5	4	50	9.5	11.4	46 (14.7)
Calibration	1	0	0	0	2	3	6	4	5	7	28	5.3	6.4	44 (14.1)
AR App.	6	4	2	8	3	7	7	6	7	5	55	10.5	12.5	45 (14.4)
Display	2	1	1	0	0	1	1	3	3	2	14	2.7	3.2	37 (11.8)
Evaluation	10	5	8	2	9	5	9	5	6	13	72	13.7	16.4	18 (5.8)
Mobile	6	5	1	5	8	2	3	3	3	4	40	7.6	9.1	19 (6.1)
Authoring	3	2	0	0	1	2	1	1	0	0	10	1.9	2.3	12 (3.8)
Visualization	2	3	2	2	3	1	4	1	0	3	21	4.0	4.8	15 (4.8)
Multimodal	0	2	0	0	0	0	2	1	2	2	9	1.7	2.1	8 (2.6)
Rendering	4	3	3	3	7	6	9	3	5	12	55	10.5	12.5	6 (1.9)
Total	47	40	32	37	49	40	54	37	43	60	439	-	100.0	313 (100.0)
New Category														
Perception	2	2	3	0	3	2	9	4	2	11	38	7.2	-	-
Collaboration	0	3	1	1	0	0	2	0	2	0	9	1.7	-	-
Reconstruction	0	1	1	5	4	4	5	3	2	4	29	5.5	-	-
Modeling	2	1	4	0	0	2	1	0	1	0	11	2.1	-	-
Grand Total	51	47	41	43	56	48	71	44	50	75	526	100.0	-	-

Table 2. Proportion of highly cited papers.

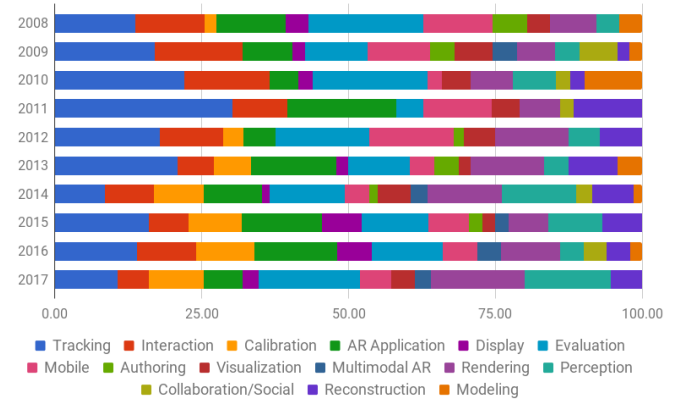| Category	% Papers	% Citations	Zhou08 % Papers	% Citations
Tracking	19.4 (16.2)	25.2 (20.9)	20.1	32.1
Interaction	11.4 (9.5)	10.6 (8.8)	14.7	12.5
Calibration	6.4 (5.3)	0.8 (0.7)	14.1	12.5
AR App.	12.5 (10.5)	10.6 (8.8)	14.4	12.5
Display	3.2 (2.7)	3.3 (2.7)	11.8	5.4
Evaluation	16.4 (13.7)	15.4 (12.8)	5.8	1.8
Mobile	9.1 (7.6)	14.6 (12.2)	6.1	7.1
Authoring	2.3 (1.9)	2.4 (2.0)	3.8	8.9
Visualization	4.8 (4.0)	5.7 (4.7)	4.8	5.4
Multimodal	2.1 (1.7)	0.8 (0.7)	2.6	0.0
Rendering	12.5 (10.5)	10.6 (8.8)	1.9	1.8
Total	100.0	100.0	100.0	100.0
Perception	(7.2)	(3.4)	-	-
Collaboration	(1.7)	(1.4)	-	-
Reconstruction	(5.5)	(9.5)	-	-
Modeling	(2.1)	(2.7)	-	-
Grand Total	(100.0)	(100.0)	-	-

Fig. 1. Trends of ISMAR research topics within ISMAR 2008–2017.

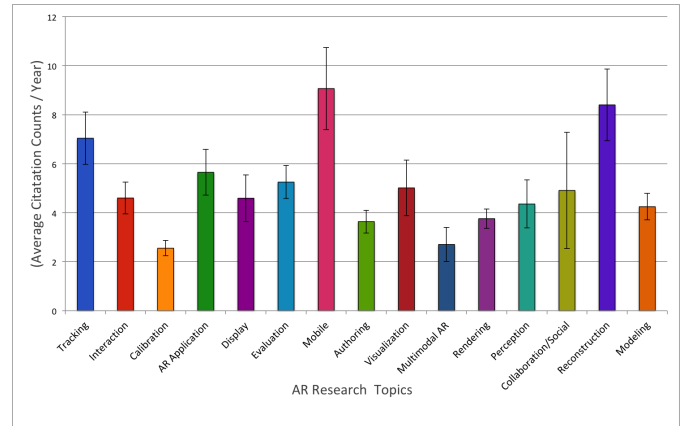Fig. 2. Research impact comparison by average citation counts per year. The error bars indicate the standard error of the mean (SEM).

introduce new trends that we found from the second decade of ISMAR publications in Section 6.

In addition to the highly cited papers and the awarded papers for the second decade of ISMAR, to cover the transition between the first and the second decade, we also review the papers from ISMAR'07. Many of these were covered by Zhou et al. but they might not have received much attention in their survey due to the fact that they had only recently been published and did not have many citations at the time. Half of the papers (4 out of 8) that Zhou et al. reviewed from ISMAR'07 have been highly cited over the previous decade, indicating a high impact on the research field. In this section, we resume these papers briefly and add seven additional highly impactful papers that were not considered by Zhou et al., including three awarded papers.

5.1 Tracking Techniques

Tracking is the most popular topic in the second decade of ISMAR as it was in the first decade. The most highly cited paper is a tracking paper [99] with over 300 citations per year, and five of the top ten

ISMAR papers with the highest citation rates are tracking papers. The reason behind this is that tracking is one of the most fundamental techniques enabling AR technologies, but it is still challenging to

achieve low-latency tracking with high precision and accuracy in a small footprint. Advances in this field can positively impact a wide range of AR applications, so there are many AR researchers still trying to create new tracking methods.

In this section, we present the recent trends in research on tracking technologies within the past ten years of ISMAR publications. We further compare the papers of the second decade of ISMAR with those of the first decade by reviewing the highest-cited papers. Overall, we found 31 papers that have more than five citations per year, including two tracking dataset and evaluation papers [39, 83]. Among the 31 highest-cited papers, we distilled them into three research aspects: (1) Simultaneous Localization and Mapping, (2) RGB-D Data and Reconstruction, and (3) Hybrid Tracking and Mobile Platforms.

5.1.1 Simultaneous Localization and Mapping

Originally proposed in the field of robotics, Simultaneous Localization and Mapping (SLAM) denotes the computational technique that creates and updates a map of an unknown space where a robot agent is located, while simultaneously tracking the agent's location in it. SLAM is known as a great alternative to traditional AR tracking approaches because it avoids the necessity for prior information, such as reference images or 3D models. SLAM has overcome many limitations and has been enhanced in terms of the robustness of tracking over the past ten years—typical limitations for SLAM are a high computational cost to deal with tracking and mapping simultaneously and tracking loss caused by fast camera motions.

Reitmayr et al. [119] employed SLAM to develop an AR annotation technique while tracking unknown environments. In this way, the system does not require pre-defined 3D models to track the real objects in the environment, which virtual annotations are superimposed on. Also, the users do not need to manually specify the full 3D pose of annotations. In 2009, Klein and Murray [73] first showed the ability to run SLAM-based tracking on a mobile platform with low computational power, such as a smartphone, by optimizing and migrating their previous Parallel Tracking and Mapping (PTAM) system [71]. This paper received the Best Paper Award in 2007 and the ISMAR Science & Technology Impact Paper Award in 2017 recognizing its huge impact in AR tracking research.

Since then, SLAM has been widely used as a tracking method for AR in many ISMAR papers. For example, Gauglitz et al. [35] received the Best Paper Award in 2012 for a method that alternatively used a panorama mapping and tracking technique and a keyframe-based SLAM technique depending on the camera movement, so that it could be robust to different camera motions, e.g., parallax-inducing or rotation-only motions. Tan et al. [137] presented a SLAM system that could detect changed features and adaptively update the map to achieve robust tracking and mapping in dynamic environments, and compared its performance to PTAM. Arth et al. [4] utilized existing untextured 2.5D city map models with initial estimates of the camera pose given a single image frame and sensor data (e.g., GPS, compass, and accelerometer) on mobile devices to reliably initialize and extend a 3D SLAM map. Liu et al. [85] proposed a monocular SLAM system that was reliable to fast camera motion with strong rotation by a novel multi-homography-based feature tracking method, which extracted multiple homographies to track 3D points with the current image frame assuming the piecewise planarity of the scene. Overall, these papers show that large strides have been made at ISMAR over the past ten years to improve SLAM-based tracking.

5.1.2 RGB-D Data and Reconstruction

In many AR applications, having a 3D model of the surrounding environment and objects is beneficial for providing a more realistic experience with plausible occlusions and interactions between the virtual content and the real environment. Although SLAM-based tracking systems maintain an environmental map that is updated while the sensor (e.g., a camera) is moving, the map is relatively sparse because the task is focused on tracking and localizing the sensor in the map rather than reconstructing a dense 3D model of the environment. The incorporation of commodity RGB-D camera sensors, such as Microsoft

Kinect, Intel RealSense, and ASUS Xtion Pro, with computationally powerful devices and sophisticated reconstruction algorithms could resolve this problem by utilizing the dense depth maps provided by the sensors.

Adopting this new approach, Newcombe et al. [99] wrote the most highly cited paper over the past ten years of ISMAR publications, which also received the Best Paper Award in 2011. They presented the *KinectFusion* system, which could map and track complex indoor environment accurately in real time with only a low-cost RGB-D camera, e.g., the Microsoft Kinect. This was shown to work robustly in arbitrary indoor environments with different lighting conditions while providing a dense reconstruction model of the environment. This paper showed the potential of RGB-D cameras and had a very strong influence not only on research of tracking and reconstruction techniques, but also on research areas such as modeling, interaction, and applications in AR.

After KinectFusion, the use of RGB-D cameras and data became more popular, and many variations have been developed to improve the quality of tracking and reconstruction. For example, McIlroy et al. [94] presented *Kinectrack*, which decoupled the dot pattern emitter and IR camera of the Kinect for real-time low-cost pose estimation. Glocker et al. [36] introduced an efficient re-localization method for RGB-D cameras in real-time 3D reconstruction. They used randomized ferns [105] and simple binary feature tests to encode whole image frames, and so could achieve fast recovery from tracking failures, with seamless continuation of mapping. Dou et al. [28] presented a system that could build 3D reconstructions of dynamic objects (e.g., human bodies) using RGB-D cameras. They collected the 3D data of moving human bodies over time, and aligned them to build noise-and-hole-free 3D human models using a non-rigid matching algorithm based on both geometry and texture measurements. Salas-Moreno et al. [124] proposed a method for using higher-level entities, such as planes and surface elements extracted from depth images, to build dense structures of the environment in real-time 3D SLAM systems. They showed that their method was intuitive and beneficial for AR applications with planar surfaces in the scene (e.g., walls).

Although RGB-D cameras are cheap and widely accessible, depth information can also be extracted from normal RGB cameras by stereo vision triangulation. Many research works primarily employed this approach to escape from heavy dependency on RGB-D sensors, which are not applicable in certain environments, such as outdoors. Pradeep et al. [110] was able to build real-time dense 3D reconstructions of the environment based on dense depth maps extracted from a single RGB camera. This used efficient stereo matching between the input frame and a key frame while the camera was exploring the environment. In mobile platforms, which are often used outdoors, depth extraction methods from RGB images are common. Schöps et al. [130], which received the Best Short Paper Award in 2014, extracted a semi-dense depth map from the environment using the whole image intensities to build a 3D reconstruction model on a mobile phone. However, these dense volumetric reconstructions have a high computational cost, which is normally not affordable on mobile devices. Kähler et al. [61] and Ondruska et al. [102] achieved dense and photo-realistic 3D reconstructions on tablet computers and mobile phones by highly optimizing the computational pipeline.

In addition, researchers tried to achieve robust tracking and reconstruction of deformable objects to broaden the use of AR in different situations. For example, Haouchine et al. [48] demonstrated an efficient real-time method to capture and augment highly elastic objects from a single view while avoiding restrictive assumptions, e.g., smoothness or geometric constraints.

5.1.3 Hybrid Tracking and Mobile Platforms

Previously, both *sensor-based tracking* methods, based on magnetic, acoustic, inertial, optical and mechanical sensors, and *vision-based tracking* methods, which use visual features from images for tracking, were popularly used for AR. However, there are many situations in which a single tracking method is not enough to guarantee reliable and robust tracking for AR. For example, vision-based tracking methods might not be useful in texture-less environments, while sensor-based

tracking methods may not provide accurate registration. Hybrid tracking techniques are a promising alternative since they can combine multiple sources of data to improve the tracking quality, e.g., images from RGB and depth cameras, and inputs from different sensors such as the Global Positioning System (GPS) and an Inertial Measurement Unit (IMU), to improve the tracking quality. Such techniques are particularly useful in mobile platforms, which various sensors are inherently embedded.

In 2007, Reitmayr and Drummond [118] presented hybrid tracking techniques employing both GPS and visual data for more accurate localization of AR devices while avoiding undesired additional user inputs for re-initialization when the tracking failed. Over the past ten years, hybrid tracking methods have become more popular in ISMAR publications. While all 31 highly cited tracking papers use vision-based tracking in some ways, 16 of them are about hybrid tracking methods. The presented hybrid tracking approaches fuse different sensor data, typically from GPS and IMUs, with image data from cameras. For example, Schall et al. [128] presented a system that refined and employed the fusion of GPS, inertial and vision measurements using a Kalman filter to increase the robustness and accuracy of camera pose estimates. Oskiper et al. [104] received the Best Paper Award in 2013 for an AR binocular system that used two wide and narrow field of view cameras in combination with an IMU and GPS to reduce jitter and drift for high-precision augmentation of telescopic imagery with virtual objects and effects. Arth et al. [3] used GPS data and a panoramic view of the user's environment to localize the mobile phone that the user is holding and to estimate its six degrees of freedom (DoF) pose. His later work [4], showed how outdoor localization and SLAM initialization could be performed with the use of 2.5D Maps, and was honored with the Best Paper Award in 2015. Ventura and Höllerer [141] proposed a system that achieved self-localization and 6 DoF pose estimation using both video stream and IMU data in real time, along with a pre-reconstructed point cloud model in an arbitrary outdoor environment. Kurz and Benhimane [79] presented a vision-based tracking method that incorporated the direction of gravity measured with IMUs to improve the detection and description of visual feature points for the quality of planar object tracking in mobile AR applications. Several years later Sweeney et al. [136] won the ISMAR 2015 Best Short Paper Award, for work that used a mobile phone IMU to measure the vertical direction and then find the absolute pose of a single or multi-camera system using computer vision methods. Kurz and Benhimane's method [79] performed planar template tracking, while Sweeney's work [136] provided full 3D tracking.

Hybrid tracking approaches are often combined with SLAM, called visual-inertial SLAM. For example, Liu et al. [85] proposed a more robust SLAM system that dealt with fast camera motion with severe rotation using IMU data for camera pose optimization, and demonstrated the effectiveness of the system. Li et al. [82] presented a hybrid visual-inertial tracking approach to fuse camera and IMU measurements for metric distance estimation and localization of a mobile AR device. Given the prior work addressed above, most hybrid tracking examples with various sensing modules are targeting mobile platforms for AR.

5.2 Interaction Techniques and User Interfaces

The usefulness of AR as perceived by practitioners and end-users in application domains is to a large degree dependent on effective and efficient user interfaces in AR. Hence, an important research direction is the development of interaction techniques with virtual content in AR.

In total, there have been 50 papers published on the topic of interaction techniques and user interfaces at ISMAR over the last ten years, which made up 11.4% of the conference, but only 13 of these papers had an average rate of more than 5 citations per year. These numbers indicate a slight decline in publications on this topic compared to 14.7% in Zhou et al. [152].

5.2.1 Tangible AR

After the seminal concept of *tangible user interfaces* (TUIs) was presented by Ishii and Ullmer [56] in 1997, most early interaction research

at ISMAR was focused on demonstrating the benefits of using TUIs in AR, and this approach have since become more and more widespread. The trend to move toward TUIs continued over the last ten years and many AR research prototypes included physical objects as an intuitive way to interact with virtual content. In total, 16 papers published at ISMAR from 2008 to 2017 used TUIs.

A noteworthy advance in the field of TUIs consisted of the use of complex physical objects and deformable surfaces instead of rigid objects for interaction in AR. A high-impact paper in this field focused on an approach that allows users to create complex surfaces by interacting with physical objects and then mapping virtual content onto their physical construction [60]. These *surface particles* then allow programmed content to be created independently of the display object and to be reused on different surfaces. Another high-impact paper covered the design and development of mixed reality books [37], which combine the natural look and feel of a real book with projected virtual 2D or 3D content that is overlaid over the real pages. The deformable pages of the book thus act as generic proxies for virtual information and can be supplemented with additional affordances such as AR elements in the book that can be touched and interacted with. The authors discussed in detail their considerations in the fields of human factors and user experience, and they mapped out the design space for allowing users to push, pull, roll, and press different interactive features presented in a mixed reality book. While the authors focused their research on books, many of their ideas can be transferred to other tangible interfaces in AR as well, which made this a compelling and influential work.

5.2.2 Mid-Air Interaction

As an intangible alternative to TUIs, research on mid-air interaction has received wide-spread attention over the last ten years. Depictions of natural user interaction with intangible virtual content in AR have become more and more prominent in science fiction movies such as *Iron Man* (2008), which shows a 3D user interface in AR designed by John Underkoffler. Moreover, advances have been made in mid-air hand tracking and gesture recognition technology such as the *Leap Motion Controller* (2013). However, compared to this wide-spread interest in this field, only a low number of 5 papers at ISMAR focused on mid-air interaction with intangible virtual content over the last ten years.

The most impactful paper has been presented by Ha et al. [46], who described a user interface called *WeARHand*, which used a head-mounted display and depth sensors to track the user's hand in mid-air and included gestures to select and manipulate virtual 3D objects. The described method thus required no environmentally tethered tracking devices or gloves. The authors further included visual shadows and occlusion feedback with a semi-transparent virtual proxy-hand, thus addressing one persistent challenge in mid-air interaction with stereoscopic displays, the problem of users misperceiving depth. In an effort to tie such intangible mid-air interfaces to TUIs, Petersen and Stricker [108] presented considerations about a general design space from tangible real objects to virtualized objects which they called *continuous natural user interfaces* (CNUIs).

While mid-air user interfaces are considered too tiring for long-term use in some application domains, they have been prominently used in the field of motor dysfunction assessment and rehabilitation. For instance, Cidota et al. [25] used a mid-air AR gaming interface with free hand and body tracking to assess motor dysfunctions in a patient's upper extremities. Further inspiring research in mid-air interaction was performed by Regenbrecht et al. [117], who re-mapped the perceived position of a user's hands in an AR gaming user interface. They visually amplified a user's hand movements such that small actual hand movements lead to perceived larger movements, which was found to be useful for post-stroke rehabilitation.

5.2.3 Mobile Device User Interfaces

Papers that focused on research with mobile AR devices also often included some form of user interface for interaction with the virtual content. Research on such user interfaces is important, which has been clearly shown in an online survey on mobile AR applications presented

by Olsson and Salo [101]. They found that 25 out of 90 participants in the survey commented on bad usability in general, a poor-looking user interface, or a lack of user interface feedback in mobile AR applications they had experience with.

Multiple papers over the last ten years presented user interfaces for mobile AR devices with improvements related to ergonomics and human factors. For instance, Veas and Kruijff [140] as well as Schall et al. [127] proposed improvements to the ergonomic form factor of mobile AR devices, e.g., using one or two joystick handles attached to the AR device, which allow holding it for extended periods of time. Moreover, Baričević et al. [13] suggested providing a 3 DoF pointer for the user's dominant hand to interact with virtual objects while the non-dominant hand holds the mobile device showing an AR magic lens view. For 2D GUI interactions in an AR maintenance scenario, Henderson and Feiner [52] demonstrated the use of an ergonomic wrist-worn mobile device.

5.2.4 Collaborative User Interfaces

While many of the user interfaces presented at ISMAR over the last ten years are designed to be used by one user at the same time, a small but increasing number of papers explicitly focused on ways to improve collaboration in AR or took the potential collaborative scenarios into account. Although there are few highly cited papers in this category, Nilsson et al. [100] received the Honorable Mention Award for presenting a collaborative AR tool for the police or military personnel for use in joint planning tasks to deal with catastrophic situations. They evaluated the tool with representatives from different organizations, and the results showed the benefits of the proposed AR collaborative tool in task performance compared to the traditional tools.

Unlike co-located collaborative scenarios, a particular challenge in remote collaborative AR systems is the physical absence of the remote users in the local collaboration space, i.e., at least one user is not in the same physical space as the other users. A high-impact paper in this scope described a collaborative telepresence system based on *shader lamps* avatars [84]. The paper described the system and pipeline for capturing a user's head in real time, streaming it over the network, and then projecting it onto a robotic physical head, to give the illusion of seeing the remote user's head to multiple viewers. The paper further discussed how the remote user should see the world, e.g., via a 360-degree camera mounted on the robotic head, and how gaze directions and eye contact can be matched between multiple users.

Multiple papers have investigated eye contact during collaboration in AR and shown that eye contact is important but challenging when using head-mounted displays (HMD) in co-located AR [112] or remote collaboration in AR [45].

5.2.5 Authoring and Modeling

Multiple research papers have been presented in the fields of AR authoring and modeling, although similar numbers of papers have been focused on these topics in the last ten years and the decade before.

In the field of AR authoring, the most often cited papers focused on sketching user interfaces. A paper by Bergig et al. [16] described a framework in which a user can sketch objects on a sheet of paper, which are then acquired by a webcam, and turned into a 3D virtual scene that is then augmented and simulated on top of the physical sketch. The described system stands out as it combines AR with sketching and allows users to sketch the objects in-place and modify them by editing the sketch itself. Printed sketches can be combined with hand sketches to form a scene. Complementing this framework is an approach described by Magnenat et al. [89] honored with the Honorable Mention Award, in which users can colorize and texture sketches and drawings in a coloring book to change the appearance of corresponding 3D objects presented in AR. Their paper focuses on colorizing 3D characters to capture the imagination of children and provide early opportunities for creative expression, but the approach could be extended to other application fields as well.

For AR modeling, the most highly cited paper was presented by van den Hengel et al. [139], which described an interactive process for generating 3D texture-mapped models of real objects within an AR system using an in-situ image-based modeling approach. The described system combined a real-time camera tracking system and automated image analysis with a user interface consisting of a range of modeling interactions called *Jiim*. This paper was presented in advance of later developments such as KinectFusion [99] and MonoFusion [110], which automated the process of generating 3D models of real objects.

5.2.6 Multimodal Interaction

Finally, since 2008, only 2.1% of the papers investigated multimodal interaction in AR, which has slightly decreased from the 2.6% papers that were published on this topic over the previous ten years. The low contributions to this field at ISMAR indicate that visual feedback and hand-based input are still dominant in the field, with multimodal AR remaining a niche research topic.

There was only one ISMAR multimodal paper that received more than 5 citations per year. The paper was presented by Piumsomboon et al. [109], who demonstrated that hand gestures and speech commands can be combined for effective interaction in AR. This multimodal interaction technique was named Gesture-Speech, and was rated significantly higher in usability for certain tasks such as uniform resizing of virtual objects than other interaction techniques compared against.

5.3 Display Techniques

Multiple papers have been presented at ISMAR in which novel AR display technologies were described. Zhou et al. [152] showed a high percentage of papers published on display technologies (11.8%) and calibration techniques (14.1%). Over the previous ten years, however, these numbers have changed. We now see a significant decline in papers published on display technologies (3.2%) and calibration techniques (6.4%). This may be because display technology has been more commoditized, and also the growing number of display focused conferences that academics can publish in. While fewer in number, some of these display papers have addressed major challenges and have had a high impact on research in the field.

A long-standing challenge in the field of stereoscopic HMDs is the ability to correctly simulate the naturally coupled accommodation and convergence cues. Most existing HMDs present a pair of stereoscopic images at a fixed focal distance, which causes a conflict between convergence and accommodation cues and has been linked to different adverse effects compared to normal viewing in the real world. A highly cited paper by Liu et al. [86], which received the Best Student Paper Award in 2008, addressed this issue and presented a novel liquid lens design that could actively adjust the focal distance of a HMD from infinity to the near point of the eye. A monocular optical see-through HMD prototype was presented and evaluated, which showed great potential for future developments toward correct accommodation and convergence cues. Five years later, another high-impact paper on this topic was presented by Maimone and Fuchs [91], who described a novel optical see-through HMD design with multiple simultaneous focal depths inspired by multi-layered desktop 3D display designs [143]. The presented design created focused images via a stack of spatial light modulators positioned closer than the eye accommodation distance. An early experimental prototype was presented in the paper, which proved the concept but revealed challenges such as the low quality of the imagery and high computational requirements. However, practical displays based on this design could support a wide field of view, selective occlusion, and could be constructed in a compact eyeglasses-like form factor.

Another persistent challenge in the field of AR displays is latency, which can induce "swimming" artifacts in AR views, particularly when using optical see-through HMDs. In order to reduce the latency of AR displays, Zheng et al. [151] presented a novel image generation approach based on direct control of the internal display technology instead of going through the usual video interfaces, such as HDMI. They presented a proof-of-concept prototype of an optical see-through display based on a projector with an imaging chip that was directly controlled by a computer, similar to the way that random access memory is controlled. The authors showed that their prototype could reduce

latency and provided benefits for AR overlays on a moving object compared to a conventional video interface.

While recent developments in AR HMDs, such as Microsoft's HoloLens, have made HMDs more and more popular, they are not the only type of AR displays that can be used effectively. In the field of telepresence, a high-impact paper by Lincoln et al. [84] presented a creative combination of a projector-based display with a robotic pan-tilt unit. The display surface was shaped like a human head and it was projected upon using the *Shader Lamps* technique [116] to give it the appearance of a real head. The projected imagery was generated in real time by capturing the head of a real person. The human head shape was mounted on a robotic platform that would pan and tilt the head, matching the corresponding real head rotations of the person being projected. The authors presented a proof-of-concept prototype and promising results of the effectiveness of this telepresence approach.

5.4 Applications

AR technology has been used in many different types of applications, and the AR Application category was among the top five research topics throughout the first and second decades of ISMAR conferences. The focus of previous survey by Zhou et al. [152] was more concentrated on Tracking, Interaction, and Display techniques, but we would like to include AR Applications in our review to cover more general trends of AR research in ISMAR conferences and to describe what circumstances AR could be practically useful.

We start with a review of the papers related to AR Applications from ISMAR'07 with high citation counts. Following that, reviews of 14 highly-cited papers within ISMAR'08 and ISMAR'17 are described in different application topics. There are various domains that use AR including education, marketing, and simulation, among others, but we focus on some of highly cited impactful application examples, such as maintenance and medical applications.

5.4.1 Industrial/Military Maintenance

One of the big areas that could benefit from AR is maintenance and training in industrial or military contexts. Previously in 2007, Pentenrieder et al. [107] presented a continuous developing process of industrial AR application in factory planning context. They employed various platforms from web-based PC to AR HMD, and considered different quality measures to develop the application, such as usability and accuracy.

For the recent decade, there were also several impactful papers in this domain. For example, Schall et al. [127] developed an outdoor AR application in a handheld mobile platform that shows virtual redlining, which is illustration of the underground infrastructure such as electricity or telecommunication lines, for field workers of utility companies. They performed a series of field trials and interviews with actual field workers for the system evaluation, and found that a palette of pre-defined symbols and the spatial interaction technique using point-and-shoot metaphor were preferred. Henderson and Feiner [52], which received the Best Paper Award in 2009, prototyped an AR maintenance application for military mechanics to see visual guidelines in AR while dealing with an armored vehicle turret. The prototype AR tool allowed the mechanics to improve task performance, such as faster identification of tasks. They also used an AR maintenance tool with a procedural assembly task to investigate the user's task performance and preference [53], which we will describe in more detail in Section 6.1.2. Fite-Georgel [31] presented a comprehensive survey of industrial AR applications with a taxonomy of the use cases, such as product design, manufacturing, and inspection and maintenance. They pointed out the lack of actual products using AR despite the advanced technical achievements in display, tracking, and rendering for AR.

5.4.2 Medical Applications

Another big and popular application domain using AR is medical simulation and training. Applications in various medical or clinic contexts have continuously appeared in ISMAR. For example, Bichlmeier et al. [17] presented a medical AR application that superimposed 3D medical imaging data on the patient's body in real time so that surgeons could see patient data without having to look away at separate displays, and also have natural and intuitive perception of 3D medical imagery. They demonstrated different medical situations with a cadaver or a real human patient employing the prototype system to explore the effectiveness of AR in medical scenarios. Haouchine et al. [49] superimposed pre-operative Computed Tomography (CT) images on the laparoscopic view for surgery guidance, and demonstrated the benefits of real-time augmented information even in the case of liver or other anatomical structures for minimally invasive surgery. In the same manner, Collins et al. [26] overlaid a pre-operative Magnetic Resonance Image (MRI) on the laparoscopic view in uterine laparosurgery. They evaluated the system qualitatively via a preliminary study with a patient suffering two myomas, and found that the surgeon reported that the myomas appeared to be localized well through the AR system. Recently Chen et al. [24] surveyed literature related to medical MR over the last two decades, and analyzed 1,403 relevant papers adapting a text mining method. Based on the results, they presented a taxonomy of medical MR applications and technologies while reporting a gradual increase of training and education applications in the medical domain.

5.4.3 Games

Entertainment and gaming is also a huge application domain for AR technology as we saw from the success of Pokémon Go [50]. Although there were a few papers introducing or specifically targeting game applications for the second decade of ISMAR publications, we could not find highly cited papers among them. However, here we introduce a couple of AR game applications from ISMAR'07, which have high citation counts. For example, Schmalstieg and Wagner [129] introduced location-based museum game applications using handheld AR devices and fiducial markers, and discussed users' positive feedback indicating the practical value of the applications. Chekhlov et al. [23] proposed an automatic plane discovery method using a SLAM-based technique to build a game where a virtual character could jump around flat surfaces in the real world. Unlike typical AR applications that used a pre-defined real environment, this approach enabled the users to dynamically generate the planar surface based on the real game environment.

5.4.4 Tour Guide and AR Browsers

A combination of AR technologies with tour guides, e.g., based on virtual agents, has been shown for museums, and cultural heritage sites, etc. For instance, Miyashita et al. [98] presented a highly cited paper that described their experiences with an AR-based museum guide as an extension of existing audio guides. Another highly cited paper by Haugstvedt and Krogstie [51] described a mobile AR application with old photographs and interesting information about a historical street.

In the domain of cultural heritage or city-size tours, AR browsers are a useful tool to visualize AR content. MacIntyre et al. [88] illustrated a city tour project that benefited from their prototype web-based mobile AR framework, which leveraged the WWW ecosystem, e.g. Google Earth. Grasset et al. [38] also investigated visualization techniques for AR browser applications, which we review in more detail in Section 6.2.2

5.4.5 Telecommunication and Broadcasting

AR could be also a good medium to replace or improve conventional telecommunication and broadcasting methods, such as phones, video conferencing, and televisions, because of the immersive experience that AR can provide. For example, Fuchs and his colleagues developed telepresence systems using animatronic Shader Lamps Avatars [84] and commodity depth cameras like Kinect sensors [90] for remote collaboration and telecommunication. Similarly, for a broadcasting scenario, Grundhöfer et al. [43] described an approach employing an imperceptible coded image projected on a screen to achieve AR without harming the visual naturalness caused by ad-hoc marker patterns. They demonstrated the potential of this approach by adapting it for television studio situations.

6 THE DEVELOPMENT OF NEW TRENDS

Over the previous ten years, there have been significant updates and changes in ISMAR research trends. Some of the new research trends and observations were already introduced in the previous section, such as Reconstruction and Collaboration. Here we explicitly present two important changes in AR research that we identified from the ISMAR 2008–2017 publications, i.e., sharp increases of Evaluation and Rendering research.

6.1 Evaluation

User evaluation and feedback has become one of the main categories for research presented at ISMAR. From the papers published at ISMAR over the previous ten years, 16.4% of them related to evaluation and testing and 19 had an average citation rate of more than 5 citations per year. In contrast, Zhou et al. [152] found only 5.8% of papers focusing on evaluation, showing a significant increase in user evaluation research in recent years. This is particularly true for the highly cited papers with 15.4% of the highly cited papers in recent years being related to evaluation, compared to only 1.8% in the decade to 2008. Overall, not only has the number of evaluation papers increased, but the number of evaluations throughout all papers has also increased. This is in line with the expectation that most ISMAR papers should include some form of evaluation when a new input or interaction method is introduced, or an application is presented. This has not always been the case in the early days of the conference.

The most highly cited 19 Evaluation papers fall into three main categories: (1) Survey papers — that survey people's response to technology or provide an overview of important areas of AR research, (2) User evaluation—papers which report on an interesting user interface and evaluation with a number of users, and (3) Perception—papers related to fundamental perceptual studies. In the rest of this section we review highly cited ISMAR papers from each of these areas.

6.1.1 Survey Papers

Survey papers are among the most highly cited of the evaluation papers. These typically pick an area of AR and provide a comprehensive review of the related work in this area and directions for future work. For example, the highest cited paper was a survey of Perceptual Issues in Augmented Reality by Kruijff et al. [77], a follow on to an earlier paper by Drascic and Milgram [29]. Kruijff et al.'s paper provided an overview of perceptual issues in AR related to the environment where the AR application is being used, the capture process, the augmentation process, the display device and the users. They provided a summary of over 60 papers in these areas and identified directions for future research. Readers are able to use this survey paper to quickly understand key issues associated with perception in AR, and start conducting their own research in the area.

Another interesting example is the work of Olsson and Salo [101] who conducted an online survey of people who had experience with current consumer AR applications. They focused on mobile AR applications and in particular AR browsers, and AR recognition applications, collecting responses from over 90 people to set of questions focused on user acceptance and experience. Their paper provides useful guidance for people wanting to measure the technology acceptance of consumer AR applications, as well as useful guidelines for developing successful applications, such as curiosity and the novelty being among the main motivators for installing mobile AR applications.

These papers provide an example of how surveys can be developed that provide a summary of previous research in the field, guidelines for applying this research, and directions for future work.

6.1.2 User Evaluation

A significant number of the highly cited evaluation papers (10 of 19) reported on the development of an interesting AR interface and then evaluation of it with a number of sample users. For example, Sandor et al. [125] reported on the effect of different saliency cues in an AR X-ray interface. Schwerdtfeger and Klinker [131] described an experiment using an AR interface for stock picking. Piumsomboon et al. [109] compared multimodal input techniques for gesture interaction with

an AR scene. In most of these papers a mixture of qualitative and quantitative experimental measures were used, such as performance time and accuracy (quantitative), and subjective surveys (qualitative). Interestingly none of the published evaluation papers with the highest impact reported on any collaborative user studies although there were some papers in the collaboration topic, e.g., Nilsson et al. [100].

Some evaluation papers reported on user evaluations performed on novel AR interfaces. For example, in the WeARHand interface (Ha et al. [46]) two RGB-D cameras were attached to the an AR see-through head mounted display to enable the user to interact with virtual content using their bare hands. They proposed a novel gesture-based two-handed interaction method and then evaluated it in a simple block manipulation tasks. The data collected included quantitative measures (performance time, placement error) and qualitative measures (user preferences, subjective feedback). In related work, Piumsomboon et al. [109] explored how speech could be combined with gesture input in a multimodal AR interface. They evaluated the proposed system with a similar set of quantitative and qualitative cues.

Other types of evaluation relate to how AR can be used to improve industrial tasks, where the experiment compares performance with AR to a more traditional system. For example, Henderson and Feiner [53] reported on how AR could be used for procedural assembly tasks. In one condition the user saw AR cues overlaid on the real objects via a see-through HMD, while in the second condition the visual cues were presented on a monitor separate from the physical workspace. They found that there were significant improvements in overall time and accuracy in the AR condition compared to the non-AR condition. In a similar way, Schwerdtfeger and Klinker [131] evaluated how AR could be used to improve stock picking in warehouses. They conducted several experiments comparing between different visualization styles presented on an optical see-through HMD for guiding the user to the correct picking location. They measured performance time, errors and experimenter subjective observations, and found that a virtual frame around the target object produced significantly better performance than the other options.

6.1.3 Perception

There were a few perception studies related to visualization techniques to present virtual content. For example, Kalkofen et al. [62], which received the Best Student Paper Award in 2007, drew the edges of the physical target object, which virtual data should be registered on, while rendering the virtual data. This allowed users to not only focus on the virtual imagery but also understand the surrounding context, e.g., depth perception by occlusions among physical and virtual objects. White et al. [145] conducted a pilot study comparing seven different "visual hint" presentations, such as text, ghost effect, and animation, and found that the participants tended to like ghost effects with text or diagram cues. Similarly, Robertson and MacIntyre [120] explored the effect of AR content registration error on a user's task performance. In their study, they compared AR representations with three different registration errors, i.e., no error, fixed error, and random error, in a block placement task. They found that some registration error actually helped the participant perform the task more effectively in a shorter time with fewer mistakes. The paper received the Honorable Mention Award and was published in the TVCG journal later in 2009.

Over the past decade there have only been a few papers published at ISMAR that focus on evaluating user perception of AR cues. Aside from the survey paper of Kruijff et al. [77], the next two most highly cited perception papers related to haptic perception (Punpongsanon et al. [113]) and non-photorealistic rendering in AR (Steptoe et al. [133]). In both cases the authors describe an AR system that is designed to create the perceptual illusion of something that is not there, and then evaluate the system to see how realistic the perception is.

In [113] Punpongsanon et al. present a system that projects AR cues onto a user's hand that change color as the user presses harder onto a soft physical object. They explore how the perception of softness can change by using multi-sensory integration of visual and haptic sensations in the human brain, in this case creating the illusion that the hand is pressing with more force than it really is. In their user study

they were able to show that a projected visual effect that changes the appearance of a user's hand and a touched object can significantly influence the perception of softness. They showed that the effect was most noticeable when virtual cues were projected on the user's fingertips. After the user study they also provided several example applications where people were able to use the method to perceive different levels of softness with a variety of objects.

Steptoe et al. [133] addressed the issue of discernibility in AR applications, which is where the virtual AR content can be easily discernible from the image of the real world. It is typically very difficult to have AR content blend into the real world due to the need to match shadows, lighting effects, tracking, occlusion and other factors. In this case they explored what happened when a visual effect was applied to the entire AR view and if this reduced the discernibility of the AR content. They found that applying a stylized (edge-enhancement) effect made the AR content very hard to discern from the real world background video, compared to an unmodified view, or a virtualized view. This was tested using a typical perceptual evaluation where they had people look at views of different scenes with multiple objects and judge which ones are virtual or real.

6.2 Rendering and Visualization

Rendering is another important emerging area of research. Previously in the first decade, a low percentage of ISMAR papers focused on rendering (1.9%); however, there has been a significant increase (to 12.5%) for the second decade, with a total of 55 papers. We review and discuss the most impactful papers in the fields of rendering technologies along with visualization, which has received the same amount of attention (4.8%) as over the previous ten years (4.8%).

6.2.1 Rendering

There are many different papers focused on rendering techniques for AR, such as for creating realistic lighting and shadows. For the seamless augmentation of real scenes with virtual objects, a major challenge lies in realizing believable global lighting effects, which has been addressed by multiple high-impact papers at ISMAR. For instance, Klein and Murray [72], which received the Honorable Mention Award in 2008, presented a compositing rendering method that could generate virtual imagery reflecting the visual characteristics of the real imagery captured from low-cost commercial cameras, such as distortions and noise. In this way, they could achieve more realistic and visually seamless integration of virtual and real imagery. Knecht et al. [74] received the Best Paper Award in 2010 for their novel plausible realistic rendering method, which combines Instant Radiosity and Differential Rendering to reduce the artificial look of virtual objects superimposed onto real scenes. Lensing and Broll [81] described an approach based on real-time image-space global illumination and an RGB-D camera to simulate indirect illumination without the need for pre-computations. Gruber et al. [40] proposed an approach using arbitrary scene geometry as a light probe for the measurement of real-world lighting, also known as photometric registration. Jachnik et al. [59] described an algorithm for real-time surface light-field capture from a single hand-held camera, which was able to capture dense illumination information for general specular surfaces. Meilland et al. [95] showed how to use an RGB-D camera as a dynamic light-field sensor. Kan and Kaufman [64] presented a ray-tracing based rendering system, which demonstrated high-quality specular effects in AR such as caustics, refraction, and reflection, which had not been presented in real-time AR before. Menk and Koch [96] received the Best Student Paper Award in 2011 for their work on truthful color reproduction in spatial augmented reality (SAR), which included a physically-based technique that computes the influences of ambient light, material, pose, and color model of the projector to adjust the RGB values of the projected imagery. Lastly, the impressive work of Rohmer et al. [121] showed how photorealistic augmentation could be achieved on mobile devices, by using a differential illumination method and sharing the computation between a stationary PC and the mobile. They did this in a way that was able to run in real time on a consumer level tablet, and were awarded the Best Paper of ISMAR 2014 for this research.

AR rendering approaches can not only be used to augment the real world with believable virtual content, but also for diminished reality where real objects can be removed from view. In this direction, two high-impact papers were presented by the authors Herling and Broll [54,55]. The first paper introduced an approach that first identified the real objects to be removed and then used an image completion and synthesis algorithm to fill the areas in a live video stream. The second paper focused on improving the performance and image quality when removing objects in front of non-trivial near-planar backgrounds.

In another rendering approach, Magnenat et al. [89] presented a texturing process that applied the captured texture from a 2D colored drawing, which children made, onto a 3D virtual character in real time, while employing their deformable surface tracking method. Tomioka et al. [138] rendered a virtual scene geometrically consistent with the real scene from user's perspective using a video see-through (VST) mobile display to improve the user's visibility.

6.2.2 Visualization

AR visualization techniques have been used to enable users to easily understand virtual content or enhance their perception of the virtual information or the surrounding environment, e.g. providing occlusion among virtual and physical objects or depth perception. For example, White et al. [145] employed "visual hints" in AR to provide the information about potential actions that users could perform with tangible AR objects. Kalkofen et al. [62] presented an interactive visualization technique that rendered AR content along with contextual imagery of the physical object which the content should be overlaid on, and showed the improved depth perception with the proposed technique. Kalkofen et al. [63] continued the investigation of visualization effects on perception while presenting a method that could adaptively vary the visual saliency within the virtual representation using ghost effects. Grasset et al. [38] proposed a visualization technique to optimize the positions of virtual labels in the AR view by employing an image-based approach that could identify visually salient regions on the image.

Three high-impact papers presented in the field of AR visualization, focused on the challenge of enabling users to view occluded points of interest using *X-ray* visualizations in AR. The first paper described a novel approach to create an illusion of seeing moving objects through occluding surfaces using a camera viewing the occluded area [14]. The proposed technique could create video views of dynamic occluded objects in real time. The authors approximated the scene as piece-wise planar in 3D such that the process did not require any 3D reconstruction. The second paper described an improvement of X-ray visualization techniques in AR with image-based ghostings [153], which can help to understand the relationship between hidden information and the real view. The authors analyzed camera images of the real view and grouped pixels into regions, which were then used to determine whether a region should preserve the real view or include synthetic overlays. The third paper focused on facilitating correct occlusions between occluders and occluded objects in AR X-ray visualization [125]. The system aimed to preserve the context of occluder objects such as important visual landmarks through saliency maps.

7 RESEARCH DIRECTIONS

Previously in Zhou et al. [152], the authors identified limitations that current AR systems were facing at the time, and suggested future directions that could be predicted from their review. In this section we reiterate upon the previous points from their paper, and compare them with the actual trends over the previous ten years of ISMAR publications to see if those limitations have been addressed. In addition, based on the developments over the recent ten years, we identify new potential future research directions. This was done by identifying some of the most popular common themes and keywords among the high impact paper reviews. The numbers of papers in each of these areas and some of the key papers are mentioned in the remainder of this section. Finally, we share several insights that arose while reviewing and analyzing the 2nd decade of ISMAR research.

7.1 Tracking

In terms of AR tracking, Zhou et al. [152] identified three limitations for real time 3D tracking:

(1) Complexity of the scene (finding target objects in a cluttered scene),

(2) Finding distinguishable features for robust outdoor tracking,

(3) The limitations of each tracking method for robust tracking.

In the decade since then, these difficulties have been addressed and possible solutions presented in many of the ISMAR tracking papers. For example, Park et al. [106] showed how to track multiple occluding 3D objects, and Roussos et al. [123] presented a method for simultaneous segmentation, motion estimation and dense 3D reconstruction of dynamic scenes on a mobile phone. There have been many papers on suitable features for outdoor tracking, such as the work of Arth et al. [4] who use the edges of buildings and map information for robust city scale tracking. Finally, a number of hybrid tracking approaches have been developed, including [128] which combines input from GPS, gyroscopes, magnetometers and accelerometers in a mobile platform.

Previously Zhou et al. [152] also suggested a few anticipated future directions in AR tracking research:

(1) Recognition systems that can acquire a reference representation of the real world for tracking, and robust feature detectors (21 papers),

(2) Tracking without known features (prior knowledge, or maps, model-free), including SLAM based systems (20 papers),

(3) Ubiquitous tracking and pervasive middleware combining data from multiple hybrid sensors into a seamless tracking framework (2 papers).

The list above includes the number of ISMAR papers published in each of these areas in parentheses. As we discussed above in Section 5.1, there has been a significant amount of research presented in recognition systems and feature descriptors, and SLAM-based tracking. However, the results above show that there has been almost no research at ISMAR in Ubiquitous Tracking and the middleware required to provide pervasive AR tracking services. A large number of papers have presented novel systems for feature recognition and description, especially on mobile devices, where computational resources are limited. For example, one of the highly cited papers that also received the Best Paper Award in 2008 was Wagner et al. [142], which optimized and extended computer vision feature descriptors to work on mobile phones. Yang and Cheng [148] developed an ultra-fast feature descriptor for scalable AR on handheld devices. Others have researched novel representations and techniques, such as shape recognition from hand drawn sketches [47] or tracking from the deformable surface of coloring book pages [89].

The bulk of the tracking papers have addressed the identified area of tracking without known features. In this case, SLAM has become a dominant tracking method resolving the issue with necessary prior knowledge for tracking. Beginning with Klein and Murray's highly cited PTAM paper [73] there have been many papers on SLAM methods for mobile phones. Hybrid approaches employing vision-based SLAM and complementary sensors (GPS, inertial sensors, etc.) have been actively being pursued to achieve more robust and higher quality tracking [85]. In addition, SLAM approaches with depth information have been developed to provide dense 3D reconstruction models that can be utilized for AR tracking. For example, Newcombe et al.'s highly cited work on KinectFusion [99], which inspired many researchers developing tracking systems using depth sensors (e.g. [147]). Nowadays, wearable AR headsets such as the Microsoft HoloLens, the Meta Meta 2, and DAQRI Smart Glasses, exploit SLAM methods with depth cameras/data and complementary internal sensors.

One interesting approach to tracking without known features is to use multiple rolling shutter cameras as dynamic 1D sensors. These act as line scanners and by combining multiple inputs together a high-speed tracking system can be developed. Bapat et al. [12] showed how this could work with 10 Go-Pro cameras arranged in 5 stereo-pairs, and were able to achieve room scale tracking at 120 fps with a display pixel

error of less than 0.5 pixels. They were awarded the ISMAR 2016 Best Paper Award for this work.

However, AR tracking still has a long way to go before it is truly ubiquitous and only two ISMAR papers presented research in ubiquitous tracking. The first of these [114] presented a method for enabling wide area tracking using sensor fusion from a dynamic combination of mobile and stationary trackers. The second explored how smartphones could be combined with networked infrastructure and fixed sensors to provide large scale tracking [149]. Other systems have also been developed for wide area tracking (e.g. [4]), but without a focus on providing a ubiquitous tracking service. This shows that there is considerable opportunity for research in this area. For example, further research is needed at least related to the means for collecting, and retrieving such data at the scale and with the robustness envisioned for ubiquitous AR.

In terms of other directions for future tracking research, as the boundary between the virtual and real/physical spaces is increasingly blurred, opportunities for interaction between those two spaces increase. For example, the future of AR "tracking" should move beyond mere object tracking and pose estimation, to enable broader context-aware AR, and implicit understanding of the real–virtual spaces. The convergence of AR with the Internet of Things (IoT) technology, could be a further avenue for achieving robust ubiquitous AR tracking, by sharing useful information about the physical environment with AR systems via local IoT devices [22]. Rambach et al. [115] received the Best Poster Award for presenting the concept of "Augmented Things," in which physical objects can encapsulate and carry all the necessary tracking and augmentation information required for AR applications.

With respect to the trend toward context-awareness or understanding [42], AR tracking could include more advanced computer vision and sensor fusion techniques, including peer-to-peer collaborative sensing, that use semantic understanding of the environmental/situational context. Steed and Julier [132] proposed a method to align the coordinates of multiple tracking systems using human user's behavior without disruptive procedures. Also, exploiting Artificial Intelligence (AI) and Deep Learning (DL) for AR tracking should be emerging. In fact, in 2017 there were already at least three papers that used DL methods for tracking and rendering [32, 34, 92], and research might be able to employ DL techniques to further recognize, track, and understand the user's behavioral/emotional status.

7.2 Interaction Techniques and User Interfaces

In terms of interaction techniques and user interfaces, Zhou et al. [152] identified three areas of limitations with current AR systems:

(1) Problems with using physical objects and gesture as methods for interacting with virtual content,

(2) The lack of good Human Factors design for the physical form factor and comfort of AR interface technology,

(3) Poor interaction design from a cognitive perspective that makes it complex to use the AR systems.

In the years since 2008, these limitations have been addressed to some extent, but work still remains. For example, there have been some excellent papers presented at ISMAR on gesture tracking [46] and how to use gesture to interact with AR content [109], and commercial AR HMDs such as the HoloLens and Meta 2 both have reliable gesture input. However the problems identified with showing the state of digital data associated with physical tools in Tangible Interfaces remain. Similarly, many of the Human Factors issues have been addressed and AR systems are easier to use than ever before, especially handheld AR applications. In this case the physical difficulty of using AR is reduced as most people are familiar with mobile phone interaction styles, although the need for good interaction design from a cognitive perspective is still needed as users deal with AR HMDs in a wide variety of form factors.

Zhou et al. [152] also predicted that the following research areas would be increasingly important in the years to come:

(1) New interaction techniques based on ubiquitous computing (4 papers),

(2) Interaction based on a Tangible User Interface approach (12 papers),

(3) Researching the social, cultural and psychological phenomena behind AR (6 papers).

The list above includes the number of ISMAR papers published in each of these areas, and overall there has been a relatively modest amount of research in these topics, especially in applying a Ubiquitous Computing approach to interaction. The small amount of research associated with Ubiquitous Computing may be due to the relative difficulty in creating robust wide area tracking, but some research has been conducted in developing a robust outdoor visualization tool [127]. Other research has shown how Ubiquitous Computing sensors can be added to physical objects to create a smart AR training experience [75].

Tangible User Interfaces (TUI) have become one of the most popular ways of interacting with AR applications, as shown by a steady stream of ISMAR papers in this area. TUI techniques have been applied to AR coloring [89], interacting with AR books [37] and maps [93]. There has also been research into how blocks can be used as AR input devices for children with autism [10] and for game play [146]. Tangible UI methods have also been applied to manipulating handheld AR devices themselves for input, such as shaking to select from a menu [144], and also to projected spatial AR interfaces [60].

In the rest of this section, we outline three more directions where we perceive a strong trend toward AR user interface research. First, even though only a low number of papers at ISMAR focused on *multimodal interaction* in the last ten years (see Section 5.2.6), we believe that it will become more and more important in the near future. While traditional considerations of AR user interfaces mainly focused on basic forms of interactions with virtual content, namely selection and manipulation [19, 97], where the benefits of multimodal interaction are questionable, we now see AR user interfaces becoming popular in more diverse fields. For instance, this includes natural human interaction with intelligent virtual agents, such as embodied AR humans, where interaction relies on all human input and output modalities. A series of research studies performed by Kim et al. [65–67] presented how the interaction with embodied AR humans could influence human perception and behaviors, such as the sense of social/co-presence, the perceived trust or confidence in such technology, and avoidance behavior. Due to the unique characteristics of AR where virtual and real things co-exist, the research on how the virtual content behaves or interacts with the surrounding environment is important [68]. In this direction, the AR/MR community and the Intelligent Virtual Agents (IVA) community should influence and interact with each other in the future to investigate the benefits or characteristics of AR agents and the interaction with them, while trying to understand the human perception of and behavior with the agents in social contexts. As AR is pushing forward in the field of social interaction [11], we see multimodal *social user interfaces* as a major challenge for the next ten years.

Second, much previous research focused on *natural user interfaces*, such as tangible [56] or intangible "touching the void" [20] user interfaces, trying to replicate real-world interaction with virtual content in AR. While natural interaction is a powerful paradigm for AR user interfaces, it introduces a number of challenges. Such user interfaces are mainly characterized by natural input, i.e., users can touch, grasp, or move an object and see virtual content being affected interactively. However, current-state interfaces lack bi-directional interaction, i.e., the ability of the virtual to affect the physical state of the real world. As Sutherland [134] stated in his *Ultimate Display*, the computer should be able to control the existence of matter during interaction. We consider this a long-term challenge beyond the next ten years. However, there may be some near-term approaches that could give the virtual some control over the real. An example are 3D- or 4D-printers, which are currently too slow to be used effectively as part of an AR user interface, but future research in this direction and real-time *printing* of virtual objects could close the loop between real and virtual interaction.

Third, when we look into recent advances in the field of VR, we see that a large number of papers was focused on novel perceptually-based user interfaces which were inspired by natural interaction but were not constrained by physical limitations. As AR technologies become more and more common in our society, we believe that the use of natural user interfaces will decline, whereas the use of *magical* and *augmented* interaction beyond the confines of real-world interaction will be introduced and increase. For example, people will be able to communicate and interact with each other or computers via brain–computer interfaces and affective computing technology. Research in this direction is not common yet, but we see a future trend in adopting some of these interaction methods to AR.

7.3 Displays

Regarding display techniques, Zhou et al. [152] identified limitations and suggested research directions in the following areas:

(1) Problems with HMDs, especially their FOV, wearability, and how they distort the images shown, among other things,

(2) The lack of mobility with projection based AR, and limited numbers of handheld projected AR systems,

(3) Limited tracking range for mobile AR, and lack of hybrid tracking.

To a large extent these limitations have been addressed in both academia and industry in the past ten years, and recent commercial products tried to overcome some of the limitations, e.g., optical see-through HMDs such as the Meta 2 have a wider FOV of 90 degrees, and the Epson Moverio can be worn for many hours at a time. There have been some ISMAR papers addressing these concerns as well. Handheld projectors have now become common place, and when coupled with vision-based tracking, they overcome the lack of mobility of previous systems. Finally, there has been a huge amount of research in tracking for handheld AR, and companies such as Google (with ARCore) and Apple (with ARKit) have released hybrid tracking solutions that largely cover the limitations identified.

Zhou et al. [152] also identify the following key research topics:

(1) Better HMDs with higher resolution, wider FOV, etc. (4 papers),

(2) Projection based AR, and spatial AR, handheld projectors (16 papers),

(3) Handheld displays and new handheld interaction methods (36 papers).

In this list we show the actual numbers of ISMAR papers which have been published in each of these topics areas. As can be seen, there have been only four papers which presented research into new types of HMD designs. For example, Maimone and Fuchs [91] presented a highly cited paper with a novel design for a wide field of view optical see through HMD, while Liu et al. [86] developed an HMD with variable focal planes that could support viewing at different focal lengths. There is still a need for more research in this area.

As predicted, projection based AR has become increasingly popular at ISMAR, with 16 papers published overall, 11 of them in the last three years alone. These papers cover a wide range of topics, with calibration, image quality, interaction methods, and tracking being particularly popular. For example, one of the most highly cited is the work of Jones et al. [60] who describe how spatial AR techniques can be used to turn almost any surface into an interactive area. Other papers such as Kitajima et al. [70] describe new projector tracking approaches that are suitable for handheld projectors.

Of these three areas, the one with the most publications is in handheld displays, although almost half of this research has been about tracking (16 papers). However, there has been a number of interesting papers on novel interaction methods for handheld devices, such as using a thermal camera with a mobile phone to make any surface interactive [78], or on using mobile devices for reconstruction [111].

Many research challenges in the area of displays remain. The release of the Microsoft HoloLens in 2016 has shown integrated solutions to many challenges in the field of Optical See-Through (OST) HMD designs and rekindled the interest in this field. At the same time, it emphasizes the open challenges with such designs. There are several important research directions for HMDs for the next ten years:

- Field of view: A small augmented visual field with a large unaugmented periphery can induce attention tunneling and have negative effects on user behavior [80]. It will remain a challenging goal to build displays that can cover the entire human visual field.

- Resolution: Even current high-resolution displays do not reach the full spatial and temporal acuity of the human eyes. In order to reach that goal in the next ten years, variable resolution displays and *foveated rendering* [44] are promising approaches.
- Focus distance: While early displays with variable focus distances and correctly simulated convergence and accommodation cues have been presented at ISMAR (see Section 5.3), it will remain a challenging research and development task to build practical prototypes.
- Filtering light: Most of the current-state OST display designs only can add light, not take it away, which makes it challenging to show dark virtual imagery where there is light in the real background. A major challenge is building displays that can selectively filter light entering the eyes.

One important area that was not well covered in the reivewed ISMAR papers is Light Field Displays. In a recent review of how to create "True AR" Sandor et al. [126] make a case for Light Field Displays, and how they are necessary to display photorealistic virtual content to the user. They also talk about the importance of being able to capture Light Fields using light field sensors. There have been a few ISMAR papers in this area e.g., Orlosky et al. [103], most notably Itoh's work on optical see-through head mounted displays [57, 58], and we expect to see significantly more papers in this area in the future.

7.4 Applications

There have been many survey papers trying to capture comprehensive knowledge of existing literature about AR applications in different domains. Some of them targeted specific application domains, such as medical [24], education [9], AR browser [41], while the others generally included a broader scope of AR applications [18, 76], covering mobile AR applications [22].

Although certain application scenarios are targeted and described to show the benefits of AR in the AR application research field, researchers have been believing AR will be a new media paradigm to replace or enhance the conventional media platforms and communication tools for decades [87]. In retrospect, however, Azuma [6] discusses that AR applications have been developed primarily for very specific users, e.g., researchers or professionals, in specific domains, such as maintenance and repair of complex equipment and medical visualization, due to the expensive effort to build the AR systems.

Recent advances of AR technology and promising commercial products for AR enable ordinary consumers, without an AR research or professional background, to become familiar with the concept of AR and understand the potential of this technology for their lives. Thus, we expect that AR application research will not only try to find specific "killer applications" for AR in the future, but also pursue general use cases as a social medium beyond domain-specific applications. Along with the advent of various consumer-level applications, user-based studies will be also emphasized more as we saw the trend of increased evaluation research in this review of ISMAR, considering the perceptual and cognitive aspects as well as the task performance and the quality of interaction and communication via AR [135], which we will discuss in the following section.

7.5 Evaluation

The evaluation papers published broadly align with Swan and Gabbard's survey work [135] which identifies Interaction, Perception and Collaboration as the three main areas that AR user studies are conducted in. This was based on a broader survey of all AR papers published until 2005. More recently, Dünser et al. [30] added system usability as a fourth area for user evaluation, and Dey et al. [27] also provided more application areas in the most comprehensive review of AR evaluation studies to date, adding Education and Medical as areas with a high number of publications with user studies.

This highlights several directions for future work. First, there is a need for more ISMAR publications on collaborative systems. This is an important area for AR user studies (from [27, 30, 135]), but few of these papers have appeared in ISMAR. We could identify only nine ISMAR papers on collaborative telepresence AR systems (less than 2% of all papers), only two of which were highly cited, Lincoln et al. [84] and Maimone et al. [90].

Secondly, as AR enters the mainstream there is an opportunity for more evaluation studies with commercial systems, or ethnographic studies of AR users in real world settings. Olsson and Salo's work [101] is one of the few ISMAR papers that evaluates user experience with commercial AR applications. Outside ISMAR there are a number of publications that provide good examples of how to do this, such as an evaluation of the Pokémon Go AR game [2].

Thirdly, there has been relatively little exploration in ISMAR publications of the social, cultural and psychological phenomena behind AR. An example of the type of psychological research that could be done is the recent paper of Baumeister et al. [15] who compared the cognitive load of users performing a task with three different types of AR displays, such as Spatial AR, Optical See-Through HMD, and Video See-Through HMD. Another example is the paper by Bruder et al. [21], who evaluated and manipulated motion perception in AR. The social and cultural aspects of AR can be understood through studies such as the work of Harborth and Pape [50] who collected feedback about AR from 683 Pokémon Go users. There could be a lot more research in this area in the future.

AR is an effective method to influence human perception, physiology, and cognition via the immersive experience in the mixed environment, where virtual and real things co-exist. Human perceptual sense could be amplified, manipulated, or replaced by the virtual stimuli in AR; thus, the evaluation of human perception and cognition of virtual things in AR should be intuitively an interesting and important research domain considering the second and third points addressed above.

Finally, apart from Yu et al. [150] there is little work on novel evaluation methods. AR has some unique characteristics compared to non-AR applications, such as the tight connection between real and virtual world elements, or the need to provide correct perceptual cues. This means that there may be new evaluation methods that could capture more accurately the user experience in AR. A good example of how a novel methodology can be developed is found in the VR literature, where Gabbard et al. [33] outline a usability framework for evaluating VR systems. There is a need for similar frameworks and approaches to be developed for AR systems.

7.6 Rendering and Visualization

Rendering and visualization research has become a major research area in the recent ten years. Considering the technical processes for achieving AR, this research area is considered as the final step—visually presenting virtual content to the users. Thus, the increase of rendering/visualization research could show that AR technology is getting close to the end users and real life. Here we suggest potential directions for future research.

Firstly, rendering/visualization has become tightly associated with tracking techniques, which acquire the information from the scene, to create realistic computer graphics seamlessly blended with the scene as reviewed in Section 6.2. Reconstruction research (Section 5.1.2) could be also considered to create better rendering as well. Thus, the scene acquisition and plausible physical–virtual relationship will be more emphasized and continue improving the quality of visual appearance, such as discussed by Kim et al. [69].

Secondly, the range of rendering/visualization will be expanded to different modalities beyond the visible graphics, e.g., olfactory, tactile, audio rendering. One example of how audio rendering could be beneficial is the *SonifEye* system of Roodaki et al. [122], which was awarded the ISMAR 2017 Best Paper Award. SonifEye uses audio only cues to provide feedback to users trying to touch an object at a precise angle, such as inserting a needle into a highly viscous liquid. In a user study they compared audio-only feedback to visual-only AR cues, and a combination of both audio and visual cues. They found that in the audio-only case the tracking error was halved compared to conditions with visual cues and produced more correct object touches compared to the visual-only condition. This research aligns with the recent movement of "augmented humans" [1]—an attempt to overcome the current limitations of the human body using advanced technological approaches, and mapping user input into different sensory output channels.

8 CONCLUSIONS

There are many factors that influence trends and accomplishments in the ongoing evolution of AR research. While reflection on the past might not be essential, it can be useful for us as a community to step back periodically and reflect on where we have been, and where we are going, to identify successes, challenges, and opportunities. We think that ten years (a decade) as considered by Zhou et al. [152] in 2008 is a useful period of time over which to reflect. Also, one could reflect on a previous decade independent of preceding trends and accomplishments, comparing the ISMAR research trends of this decade with the categories and predictions offered by Zhou et al. in 2008. In our effort to do so we particularly sought to answer three key questions presented in Section 2: (1) Are there any changes in existing research trends?; (2) Are there any new research areas?; and (3) What are the key developments and challenges? We have found some significant changes in research trends in the eleven research topics used by Zhou et al., for example the popularity of calibration and display research has subsided some, while rendering and evaluation have risen to become major research areas in this decade of ISMAR compared to the previous. We have also introduced four new research categories, that we considered dominant or at least important in the last decade. Finally we have also offered some of our own insights and suggestions for the future directions of AR research based on our analysis of the recent developments and challenges in ISMAR research.

The present survey paper still has limitations, e.g., we might miss insightful and potentially impactful research while concentrating on the analysis based on the average citation counts, and the scope of our survey based on ISMAR publications could not cover influential research works from other venues, which also consider AR/MR mainly or carefully. To reduce the limitations, we reviewed all the awarded papers, which already showed their values in the year, whether they have high citation counts or not.

There are still many challenges in human factors as well as technical aspects while we are moving forward to the general use of AR technology as a new form of media [6, 7]. However, we strongly believe that the AR research and user community will continue to grow in line with the dramatic increase of commercial interest in AR/MR we saw in the past years, and that AR technology will continue to develop even more dynamically and effectively over the next ten years, toward the vision of pervasive presence in our daily lives [42]. We hope that our efforts here are helpful for AR researchers who are interested in, and have passion for, helping to create this future.

REFERENCES

[1] Augmented Human International Conferences website. `http://www.augmented-human.com`. [Online; accessed 10-July-2018].

[2] T. Althoff, R. W. White, and E. Horvitz. Influence of Pokémon Go on Physical Activity: Study and Implications. *Journal of Medical Internet Research*, 18(12):e315, 2016.

[3] C. Arth, M. Klopschitz, G. Reitmayr, and D. Schmalstieg. Real-time self-localization from panoramic images on mobile devices. *IEEE International Symposium on Mixed and Augmented Reality*, pages 37–46, 2011.

[4] C. Arth, C. Pirchheim, J. Ventura, D. Schmalstieg, and V. Lepetit. Instant Outdoor Localization and SLAM Initialization from 2.5D Maps. *IEEE Transactions on Visualization and Computer Graphics*, 21(11):1309–1318, 2015.

[5] R. T. Azuma. A Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments*, 6(4):355–385, 1997.

[6] R. T. Azuma. The Most Important Challenge Facing Augmented Reality. *Presence: Teleoperators and Virtual Environments*, 25(3):234–238, 2016.

[7] R. T. Azuma. Making augmented reality a reality. In *Imaging and Applied Optics 2017 (3D, AIO, COSI, IS, MATH, pcAOP)*, page JTu1F.1. Optical Society of America, 2017.

[8] R. T. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre. Recent Advances in Augmented Reality. *IEEE Computer Graphics & Applications*, 21(6):34–47, 2001.

[9] J. Bacca, R. Fabregat, S. Baldiris, S. Graf, and Kinshuk. Augmented reality trends in education: A systematic review of research and applications. *Educational Technology & Society*, 17:133–149, 2014.

[10] Z. Bai, A. F. Blackwell, and G. Coulouris. Through the looking glass: Pretend play for children with autism. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 49–58, 2013.

[11] J. N. Bailenson. *Experience on Demand: What Virtual Reality Is, How It Works, and What It Can Do*. W.W. Norton, New York, USA, 2018.

[12] A. Bapat, E. Dunn, and J.-M. Frahm. Towards Kilo-Hertz 6-DoF Visual Tracking Using an Egocentric Cluster of Rolling Shutter Cameras. *IEEE Transactions on Visualization and Computer Graphics*, 22(11):2358–2367, 2016.

[13] D. Baričević, L. Cha, M. Turk, T. Höllerer, and D. A. Bowman. Handheld AR Magic Lenses with User-Perspective Rendering. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 197–206, 2012.

[14] P. Barnum, Y. Sheikh, A. Datta, and T. Kanade. Dynamic Seethroughs: Synthesizing Hidden Views of Moving Objects. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 111–114, 2009.

[15] J. Baumeister, S. Y. Ssin, N. A. M. ElSayed, J. Dorrian, D. P. Webb, J. A. Walsh, T. M. Simon, A. Irlitti, R. T. Smith, M. Kohler, and B. H. Thomas. Cognitive Cost of Using Augmented Reality Displays. *IEEE Transactions on Visualization and Computer Graphics*, 23(11):2378–2388, 2017.

[16] O. Bergig, N. Hagbi, J. El-Sana, and M. Billinghurst. In-place 3D sketching for authoring and augmenting mechanical systems. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 87–94, 2009.

[17] C. Bichlmeier, S. M. Heining, M. Rustaee, and N. Navab. Laparoscopic Virtual Mirror for Understanding Vessel Structure: Evaluation Study by Twelve Surgeons. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 1–4, 2007.

[18] M. Billinghurst, A. Clark, and G. Lee. A Survey of Augmented Reality. *Foundations and Trends® in Human-Computer Interaction*, 8(2-3):73–272, 2014.

[19] D. A. Bowman, E. Kruijff, J. J. LaViola, and I. Poupyrev. *3D User Interfaces: Theory and Practice*. Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA, 2004.

[20] G. Bruder, F. Steinicke, and W. Stuerzlinger. Touching the Void Revisited: Analyses of Touch Behavior on and above Tabletop Surfaces. In P. Kotzé, G. Marsden, G. Lindgaard, J. Wesson, and M. Winckler, editors, *Human-Computer Interaction – INTERACT 2013*, pages 278–296, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.

[21] G. Bruder, P. Wieland, B. Bolte, M. Lappe, and F. Steinicke. Going With the Flow: Modifying Self-Motion Perception with Computer-Mediated Optic Flow. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 67–74, 2013.

[22] D. Chatzopoulos, C. Bermejo, Z. Huang, and P. Hui. Mobile Augmented Reality Survey: From Where We Are to Where We Go. *IEEE Access*, 5:6917–6950, 2017.

[23] D. Chekhlov, A. Gee, A. Calway, and W. Mayol-Cuevas. Ninja on a Plane: Automatic Discovery of Physical Planes for Augmented Reality Using Visual SLAM. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 1–4, 2007.

[24] L. Chen, T. Day, W. Tang, and N. W. John. Recent Developments and Future Challenges in Medical Mixed Reality. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 123–135, 2017.

[25] M. A. Cidota, P. J. M. Bank, P. W. Ouwehand, and S. G. Lukosch. Assessing Upper Extremity Motor Dysfunction Using an Augmented Reality Game. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 144–154, 2017.

[26] T. Collins, D. Pizarro, A. Bartoli, M. Canis, and N. Bourdel. Computer-

Assisted Laparoscopic myomectomy by augmenting the uterus with pre-operative MRI data. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 243–248, 2014.

[27] A. Dey, M. Billinghurst, R. W. Lindeman, and J. E. Swan. A Systematic Review of 10 Years of Augmented Reality Usability Studies: 2005 to 2014. *Frontiers in Robotics and AI*, 5(37):1–28, 2018.

[28] M. Dou, H. Fuchs, and J. M. Frahm. Scanning and tracking dynamic objects with commodity depth cameras. *IEEE International Symposium on Mixed and Augmented Reality*, pages 99–106, 2013.

[29] D. Drascic and P. Milgram. Perceptual Issues in Augmented Reality David. *Proc. SPIE 2653, Stereoscopic Displays and Virtual Reality Systems III*, 2653:123–134, 1996.

[30] A. Dünser, R. Grasset, and M. Billinghurst. A survey of evaluation techniques used in augmented reality studies. Technical report, HIT Lab NZ, University of Canterbury, 2008.

[31] P. Fite-Georgel. Is there a Reality in Industrial Augmented Reality? In *IEEE International Symposium on Mixed and Augmented Reality*, pages 201–210, 2011.

[32] A. Fond, M.-O. Berger, and G. Simon. Facade Proposals for Urban Augmented Reality. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 32–41, 2017.

[33] J. Gabbard, D. Hix, and J. E. II, Swan. User-Centered Design and Evaluation of Virtual Environments. *IEEE Computer Graphics and Applications*, 19(6):51–59, 1999.

[34] M. Garon and J.-F. Lalonde. Deep 6-DOF Tracking. *IEEE Transactions on Visualization and Computer Graphics*, 23(11):2410–2418, 2017.

[35] S. Gauglitz, C. Sweeney, J. Ventura, M. Turk, and T. Hollerer. Live Tracking and Mapping from Both General and Rotation-Only Camera Motion. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 13–22, 2012.

[36] B. Glocker, S. Izadi, J. Shotton, and A. Criminisi. Real-time RGB-D camera relocalization. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 173–179, 2013.

[37] R. Grasset, A. Dünser, and M. Billinghurst. The design of a mixed-reality book: Is it still a real book? In *IEEE International Symposium on Mixed and Augmented Reality*, pages 99–102, 2008.

[38] R. Grasset, T. Langlotz, D. Kalkofen, M. Tatzgern, and D. Schmalstieg. Image-Driven View Management for Augmented Reality Browsers. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 177–186, 2012.

[39] L. Gruber, S. Gauglitz, J. Ventura, S. Zollmann, M. Huber, M. Schlegel, G. Klinker, D. Schmalstieg, and T. Höllerer. The City of Sights: Designing an Augmented Reality Stage Set. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 157–163, 2010.

[40] L. Gruber, T. Richter-Trummer, and D. Schmalstieg. Real-time photometric registration from arbitrary geometry. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 119–128, 2012.

[41] J. Grubert, T. Langlotz, and R. Grasset. Augmented Reality Browser Survey. Technical report, Institute for Computer Graphics and Vision, Graz University of Technology, Austria, 2011.

[42] J. Grubert, T. Langlotz, S. Zollmann, and H. Regenbrecht. Towards Pervasive Augmented Reality: Context-Awareness in Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics*, PP(99):1–1, 2016.

[43] A. Grundhöfer, M. Seeger, F. Haentsch, and O. Bimber. Dynamic Adaptation of Projected Imperceptible Codes. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 1–10, 2007.

[44] B. Guenter, M. Finch, S. Drucker, D. Tan, and J. Snyder. Foveated 3D graphics. *ACM Transactions on Graphics*, 31(6):1, 2012.

[45] K. Gupta, G. A. Lee, and M. Billinghurst. Do You See What I See? The Effect of Gaze Tracking on Task Space Remote Collaboration. *IEEE Transactions on Visualization and Computer Graphics*, 22(11):2413–2422, 2016.

[46] T. Ha, S. Feiner, and W. Woo. WeARHand: Head-Worn, RGB-D Camera-Based, Bare-Hand User Interface with Visually Enhanced Depth Perception. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 219–228, 2014.

[47] N. Hagbi, O. Bergig, J. El-Sana, and M. Billinghurst. Shape Recognition and Pose Estimation for Mobile Augmented Reality. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 65–71, 2009.

[48] N. Haouchine, J. Dequidt, M. O. Berger, and S. Cotin. Single View Augmentation of 3D Elastic Objects. In *IEEE International Symposium*

[49] N. Haouchine, J. Dequidt, I. Peterlik, E. Kerrien, M.-O. Berger, and S. Cotin. Image-guided Simulation of Heterogeneous Tissue Deformation for Augmented Reality During Hepatic Surgery. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 199–208, 2013.

[50] D. Harborth and S. Pape. Exploring the hype: Investigating technology acceptance factors of Pokémon Go. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 155–168, 2017.

[51] A.-C. Haugstvedt and J. Krogstie. Mobile Augmented Reality for Cultural Heritage: A Technology Acceptance Study. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 247–255, 2012.

[52] S. J. Henderson and S. Feiner. Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 135–144, 2009.

[53] S. J. Henderson and S. K. Feiner. Augmented reality in the psychomotor phase of a procedural task. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 191–200, 2011.

[54] J. Herling and W. Broll. Advanced self-contained object removal for realizing real-time diminished reality in unconstrained environments. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 207–212, 2010.

[55] J. Herling and W. Broll. PixMix: A real-time approach to high-quality Diminished Reality. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 141–150, 2012.

[56] H. Ishii and B. Ullmer. Tangible bits: towards seamless interfaces between people, bits, and atoms. *ACM SIGCHI Conference on Human Factors in Computing Systems*, pages 234–241, 1997.

[57] Y. Itoh, T. Amano, D. Iwai, and G. Klinker. Gaussian Light Field: Estimation of Viewpoint-Dependent Blur for Optical See-Through Head-Mounted Displays. *IEEE Transactions on Visualization and Computer Graphics*, 22(11):2368–2376, 2016.

[58] Y. Itoh and G. Klinker. Light-Field Correction for Spatial Calibration of Optical See-Through Head-Mounted Displays. *IEEE Transactions on Visualization and Computer Graphics*, 21(4):471–480, 2015.

[59] J. Jachnik, R. A. Newcombe, and A. J. Davison. Real-time surface light-field capture for augmentation of planar specular surfaces. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 91–97, 2012.

[60] B. R. Jones, R. Sodhi, R. H. Campbell, G. Garnett, and B. P. Bailey. Build Your World and Play In It—Interacting with Surface Particles on Complex Objects.pdf. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 165–174, 2010.

[61] O. Kähler, V. A. Prisacariu, C. Y. Ren, X. Sun, P. Torr, and D. Murray. Very High Frame Rate Volumetric Integration of Depth Images on Mobile Devices. *IEEE Transactions on Visualization and Computer Graphics*, 21(11):1241–1250, 2015.

[62] D. Kalkofen, E. Mendez, and D. Schmalstieg. Interactive Focus and Context Visualization for Augmented Reality. In *IEEE and ACM International Symposium on Mixed and Augmented Reality*, pages 1–10, 2007.

[63] D. Kalkofen, E. Veas, S. Zollmann, M. Steinberger, and D. Schmalstieg. Adaptive Ghosted Views for Augmented Reality. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 1–9, 2013.

[64] P. Kán and H. Kaufmann. High-quality reflections, refractions, and caustics in Augmented Reality and their contribution to visual coherence. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 99–108, 2012.

[65] K. Kim, L. Boelling, S. Haesler, J. N. Bailenson, G. Bruder, and G. Welch. Does a Digital Assistant Need a Body? The Influence of Visual Embodiment and Social Behavior on the Perception of Intelligent Virtual Agents in AR. In *IEEE International Symposium on Mixed and Augmented Reality*, 2018. in press.

[66] K. Kim, G. Bruder, and G. F. Welch. Exploring the Effects of Observed Physicality Conflicts on Real-virtual Human Interaction in Augmented Reality. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, pages 31:1–31:7, 2017.

[67] K. Kim, D. Maloney, G. Bruder, J. N. Bailenson, and G. F. Welch. The Effects of Virtual Human's Spatial and Behavioral Coherence with Physical Objects on Social Presence in AR. *Computer Animation and Virtual Worlds*, 28(3-4):e1771, 2017.

[68] K. Kim and G. Welch. Maintaining and Enhancing Human-Surrogate Presence in Augmented Reality. In *Proceedings of the IEEE Interna-*

tional Symposium on Mixed and Augmented Reality Workshop on Human Perception and Psychology in Augmented Reality, pages 15–19, 2015.

[69] Y. Kim, H. Park, S. Bang, and S. H. Lee. Retargeting human-object interaction to virtual avatars. *IEEE Transactions on Visualization and Computer Graphics*, 22(11):2405–2412, 2016.

[70] Y. Kitajima, D. Iwai, and K. Sato. Simultaneous Projection and Positioning of Laser Projector Pixels. *IEEE Transactions on Visualization and Computer Graphics*, 23(11):2419–2429, 2017.

[71] G. Klein and D. Murray. Parallel tracking and mapping for small AR workspaces. In *IEEE/ACM International Symposium on Mixed and Augmented Reality*, pages 225–234, 2007.

[72] G. Klein and D. Murray. Compositing for Small Cameras. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 57–60, 2008.

[73] G. Klein and D. Murray. Parallel Tracking and Mapping on a Camera Phone. *IEEE International Symposium on Mixed and Augmented Reality*, pages 83–86, 2009.

[74] M. Knecht, C. Traxler, O. Mattausch, W. Purgathofer, and M. Wimmer. Differential Instant Radiosity for Mixed Reality. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 99–107, 2010.

[75] A. Kotranza, D. S. Lind, C. M. Pugh, and B. Lok. Real-time in-situ visual feedback of task performance in mixed environments for learning joint psychomotor-cognitive tasks. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 125–134, 2009.

[76] D. W. F. V. Krevelen and R. Poelman. A Survey of Augmented Reality Technologies, Applications and Limitations. *The International Journal of Virtual Reality*, 9(2):1–20, 2010.

[77] E. Kruijff, J. E. Swan, and S. Feiner. Perceptual Issues in Augmented Reality Revisited. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 3–12, 2010.

[78] D. Kurz. Thermal touch: Thermography-enabled everywhere touch interfaces for mobile augmented reality applications. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 353–354, 2014.

[79] D. Kurz and S. Benhimane. Gravity-Aware Handheld Augmented Reality. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 111–120, 2011.

[80] M. Lee, G. Bruder, T. Höllerer, and G. Welch. Effects of Unaugmented Periphery and Vibrotactile Feedback on Proxemics with Virtual Humans in AR. *IEEE Transactions on Visualization and Computer Graphics*, 24(4):1525–1534, 2018.

[81] P. Lensing and W. Broll. Instant indirect illumination for dynamic mixed reality scenes. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 109–118, 2012.

[82] P. Li, T. Qin, B. Hu, F. Zhu, and S. Shen. Monocular Visual-Inertial State Estimation for Mobile Augmented Reality. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 13–22, 2017.

[83] S. Lieberknecht, S. Benhimane, P. Meier, and N. Navab. A Dataset and Evaluation Methodology for Template-based Tracking Algorithms. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 145–151, 2009.

[84] P. Lincoln, G. Welch, A. Nashel, A. Ilie, A. State, and H. Fuchs. Animatronic shader lamps avatars. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 27–33, 2009.

[85] H. Liu, G. Zhang, and H. Bao. Robust Keyframe-Based Monocular SLAM for Augmented Reality. *IEEE International Symposium on Mixed and Augmented Reality*, pages 1–10, 2016.

[86] S. Liu, D. Cheng, and H. Hua. An optical see-through head mounted display with addressable focal planes. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 33–42, 2008.

[87] B. MacIntyre, J. D. Bolter, E. Moreno, and B. Hannigan. Augmented Reality as a New Media Experience. In *IEEE and ACM International Symposium on Augmented Reality*, pages 197–206, 2001.

[88] B. MacIntyre, A. Hill, H. Rouzati, M. Gandy, and B. Davidson. The Argon AR Web Browser and standards-based AR application environment. In *IEEE International Symposium on Mixed and Augmented Reality,*, pages 65–74, 2011.

[89] S. Magnenat, D. T. Ngo, F. Zünd, M. Ryffel, G. Noris, G. Rothlin, A. Marra, M. Nitti, P. Fua, M. Gross, and R. W. Sumner. Live Texturing of Augmented Reality Characters from Colored Drawings. *IEEE Transactions on Visualization and Computer Graphics*, 21(11):1201–1210, 2015.

[90] A. Maimone and H. Fuchs. Encumbrance-Free Telepresence System with Real-Time 3D Capture and Display using Commodity Depth Cameras. In

[91] A. Maimone and H. Fuchs. Computational augmented reality eyeglasses. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 29–38, 2013.

[92] D. Mandl, K. M. Yi, P. Mohr, P. M. Roth, P. Fua, V. Lepetit, D. Schmalstieg, and D. Kalkofen. Learning lightprobes for mixed reality illumination. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 82–89, 2017.

[93] S. Martedi, H. Uchiyama, G. Enriquez, H. Saito, T. Miyashita, and T. Hara. Foldable Augmented Maps. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 65–72, 2010.

[94] P. McIlroy, S. Izadi, and A. Fitzgibbon. Kinectrack: Agile 6-DoF Tracking Using a Projected Dot Pattern. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 23–29, 2012.

[95] M. Meilland, C. Barat, and A. Comport. 3D Flioh Dynamic Range Dense Visual SLAM and Its Application to Real-time Object Re-lighting. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 143–152, 2013.

[96] C. Menk and R. Koch. Interactive Visualization Technique for Truthful Color Reproduction in Spatial Augmented Reality Applications. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 157–164, 2011.

[97] M. R. Mine. Virtual environment interaction techniques. Technical report, Chapel Hill, NC, USA, 1995.

[98] T. Miyashita, P. G. Meier, T. Tachikawa, S. Orlic, T. Eble, V. Scholz, A. Gapel, O. Gerl, S. Arnaudov, and S. Lieberknecht. An Augmented Reality Museum Guide. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 65–72, 2008.

[99] R. A. Newcombe, A. J. Davison, S. Izadi, P. Kohli, O. Hilliges, J. Shotton, D. Molyneaux, S. Hodges, D. Kim, and A. Fitzgibbon. KinectFusion: Real-time dense surface mapping and tracking. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 127–136, 2011.

[100] S. Nilsson, B. Johansson, and A. Jönsson. Using AR to support cross-organisational collaboration in dynamic tasks. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 3–12, 2009.

[101] T. Olsson and M. Salo. Online User Survey on Current Mobile Augmented Reality Applications. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 75–84, 2011.

[102] P. Ondruska, P. Kohli, and S. Izadi. MobileFusion: Real-Time Volumetric Surface Reconstruction and Dense Tracking on Mobile Phones. *IEEE Transactions on Visualization and Computer Graphics*, 21(11):1251–1258, 2015.

[103] J. Orlosky, P. Kim, K. Kiyokawa, T. Mashita, P. Ratsamee, Y. Uranishi, and H. Takemura. VisMerge: Light Adaptive Vision Augmentation via Spectral and Temporal Fusion of Non-visible Light. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 22–31, 2017.

[104] T. Oskiper, M. Sizintsev, V. Branzoi, S. Samarasekera, and R. Kumar. Augmented Reality Binoculars. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 219–228, 2013.

[105] M. Ozuysal, M. Calonder, V. Lepetit, and P. Fua. Fast keypoint recognition using random ferns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(3):448–461, 2010.

[106] Y. Park, V. Lepetit, and W. Woo. Multiple 3D Object tracking for augmented reality. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 117–120, 2008.

[107] K. Pentenrieder, C. Bade, F. Doil, and P. Meier. Augmented Reality-based Factory Planning - An Application Tailored to Industrial Needs. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 1–9, 2007.

[108] N. Petersen and D. Stricker. Continuous natural user interface: Reducing the gap between real and digital world. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 23–26, 2009.

[109] T. Piumsomboon, D. Altimira, H. Kim, A. Clark, G. Lee, and M. Billinghurst. Grasp-Shell vs Gesture-Speech: A Comparison of Direct and Indirect Natural Interaction Techniques in Augmented Reality. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 73–82, 2014.

[110] V. Pradeep, C. Rhemann, S. Izadi, C. Zach, M. Bleyer, and S. Bathiche. MonoFusion: Real-time 3D reconstruction of small scenes with a single web camera. *IEEE International Symposium on Mixed and Augmented Reality*, pages 83–88, 2013.

[111] V. A. Prisacariu, O. Kahler, D. W. Murray, and I. D. Reid. Real-Time 3D Tracking and Reconstruction on Mobile Phones. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 89–98, 2013.

[112] E. Prytz, S. Nilsson, and A. Jönsson. The importance of eye-contact for collaboration in AR system. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 119–126, 2010.

[113] P. Punpongsanon, D. Iwai, and K. Sato. SoftAR: Visually Manipulating Haptic Softness Perception in Spatial Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics*, 21(11):1279–1288, 2015.

[114] D. Pustka and G. Klinker. Dynamic gyroscope fusion in ubiquitous tracking environments. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 13–20, 2008.

[115] J. Rambach, A. Pagani, and D. Stricker. Augmented Things: Enhancing AR Applications leveraging the Internet of Things and Universal 3D Object Tracking. In *Adjunct Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*, pages 103–108, 2017.

[116] R. Raskar, G. Welch, K.-L. Low, and D. Bandyopadhyay. Shader Lamps: Animating Real Objects With Image-Based Illumination. In *Eurographics Workshop on Rendering*, pages 89–102, 2001.

[117] H. Regenbrecht, G. McGregor, C. Ott, S. Hoermann, T. Schubert, L. Hale, J. Hoermann, B. Dixon, and E. Franz. Out of reach? - A novel AR interface approach for motor rehabilitation. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 219–228, 2011.

[118] G. Reitmayr and T. W. Drummond. Initialisation for Visual Tracking in Urban Environments. In *IEEE and ACM International Symposium on Mixed and Augmented Reality*, pages 1–9, 2007.

[119] G. Reitmayr, E. Eade, and T. W. Drummond. Semi-automatic Annotations in Unknown Environments. In *IEEE and ACM International Symposium on Mixed and Augmented Reality*, pages 1–4, 2007.

[120] C. M. Robertson and B. MacIntyre. An Evaluation of Graphical Context as a Means for Ameliorating the Effects of Registration Error. In *IEEE and ACM International Symposium on Mixed and Augmented Reality*, pages 1–14, 2007.

[121] K. Rohmer, W. Buschel, R. Dachselt, and T. Grosch. Interactive Near-Field Illumination for Photorealistic Augmented Reality on Mobile Devices. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 29–38, 2014.

[122] H. Roodaki, N. Navab, A. Eslami, C. Stapleton, and N. Navab. Sonif-Eye: Sonification of Visual Information using Physical Modeling Sound Synthesis. *IEEE Transactions on Visualization and Computer Graphics*, 23(11):2366–2371, 2017.

[123] A. Roussos, C. Russell, R. Garg, and L. Agapito. Dense Multibody Motion Estimation and Reconstruction from a Handheld Camera. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 31–40, 2012.

[124] R. F. Salas-moreno, B. Glocker, P. H. J. Kelly, and A. J. Davison. Dense Planar SLAM. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 157–164, 2014.

[125] C. Sandor, A. Cunningham, A. Dey, and V.-V. Mattila. An Augmented Reality X-Ray System Based on Visual Saliency. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 27–36, 2010.

[126] C. Sandor, M. Fuchs, A. Cassinelli, H. Li, R. Newcombe, G. Yamamoto, and S. Feiner. Breaking the Barriers to True Augmented Reality. *ArXiv e-prints*, Dec. 2015.

[127] G. Schall, E. Mendez, and D. Schmalstieg. Virtual redlining for civil engineering in real Environments. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 95–98, 2008.

[128] G. Schall, D. Wagner, G. Reitmayr, E. Taichmann, M. Wieser, D. Schmalstieg, and B. Hofmann-Wellenhof. Global pose estimation using multi-sensor fusion for outdoor Augmented Reality. *IEEE International Symposium on Mixed and Augmented Reality*, pages 153–162, 2009.

[129] D. Schmalstieg and D. Wagner. Experiences with Handheld Augmented Reality. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 1–13, 2007.

[130] T. Schöps, J. Engel, and D. Cremers. Semi-Dense Visual Odometry for AR on a Smartphone Feature Based. In *IEEE International Symposium on Mixed and Augmented Reality*, number September, pages 145–150, 2014.

[131] B. Schwerdtfeger and G. Klinker. Supporting Order Picking with Augmented Reality. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 91–94, 2008.

[132] A. Steed and S. Julier. Behaviour-Aware Sensor Fusion : Continuously Inferring the Alignment of Coordinate Systems from User Behaviour. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 163–172, 2013.

[133] W. Steptoe, S. Julier, and A. Steed. Presence and Discernability in Conventional and Non-Photorealistic Immersive Augmented Reality. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 213–218, 2014.

[134] I. E. Sutherland. The ultimate display. In *Proceedings of the Congress of the Internation Federation of Information Processing (IFIP)*, pages 506–508, 1965.

[135] J. E. Swan and J. L. Gabbard. Survey of User-Based Experimentation in Augmented Reality. In *International Conference on Virtual Reality*, pages 1–9, 2005.

[136] C. Sweeney, J. Flynn, B. Nuernberger, M. Turk, and T. Hollerer. Efficient Computation of Absolute Pose for Gravity-Aware Augmented Reality. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 19–24, 2015.

[137] W. Tan, H. Liu, Z. Dong, G. Zhang, and H. Bao. Robust monocular SLAM in dynamic environments. *IEEE International Symposium on Mixed and Augmented Reality*, pages 209–218, 2013.

[138] M. Tomioka, S. Ikeda, and K. Sato. Approximated User-Perspective Rendering in Tablet-Based Augmented Reality. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 21–28, 2013.

[139] A. van den Hengel, R. Hill, B. Ward, and A. Dick. In Situ Image-based Modeling. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 107–110, 2009.

[140] E. Veas and E. Kruijff. Vesp'R: Design and evaluation of a handheld AR device. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 43–52, 2008.

[141] J. Ventura and T. Höllerer. Wide-area scene mapping for mobile visual tracking. *IEEE International Symposium on Mixed and Augmented Reality*, pages 3–12, 2012.

[142] D. Wagner, G. Reitmayr, A. Mulloni, T. Drummond, and D. Schmalstieg. Pose tracking from natural features on mobile phones. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 125–134, 2008.

[143] G. Wetzstein, D. Lanman, M. Hirsch, and R. Raskar. Tensor displays: compressive light field synthesis using multilayer displays with directional backlighting. *ACM Transactions on Graphics*, 31(4):1–11, 2012.

[144] S. White, D. Feng, and S. Feiner. Interaction and presentation techniques for shake menus in tangible augmented reality. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 39–48, 2009.

[145] S. White, L. Lister, and S. Feiner. Visual Hints for Tangible Gestures in Augmented Reality. In *IEEE and ACM International Symposium on Mixed and Augmented Reality*, pages 1–4, 2007.

[146] S. Willi and A. Grundhofer. Spatio-temporal point path analysis and optimization of a galvanoscopic scanning laser projector. *IEEE Transactions on Visualization and Computer Graphics*, 22(11):2377–2384, 2016.

[147] Z. Yan, M. Ye, and L. Ren. Dense Visual SLAM with Probabilistic Surfel Map. *IEEE Transactions on Visualization and Computer Graphics*, 23(11):2389–2398, 2017.

[148] X. Yang and K.-T. Cheng. LDB: An ultra-fast feature for scalable Augmented Reality on mobile devices. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 49–57, 2012.

[149] W. Yii, W. H. Li, and T. Drummond. Distributed visual processing for augmented reality. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 41–48, 2012.

[150] M. Yu, H. Lakshman, and B. Girod. A Framework to Evaluate Omnidirectional Video Coding Schemes. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 31–36, 2015.

[151] F. Zheng, T. Whitted, A. Lastra, P. Lincoln, A. State, A. Maimone, and H. Fuchs. Minimizing Latency for Augmented Reality Displays: Frames Considered Harmful. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 195–200, 2014.

[152] F. Zhou, H. B. L. Duh, and M. Billinghurst. Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR. In *Proceedings of the 7th IEEE International Symposium on Mixed and Augmented Reality*, pages 193–202, 2008.

[153] S. Zollmann, D. Kalkofen, E. Mendez, and G. Reitmayr. Image-based ghostings for single layer occlusions in augmented reality. In *IEEE International Symposium on Mixed and Augmented Reality*, pages 19–26, 2010.