Holodeck: Immersive 3D Displays Using Swarms of Flying Light Specks[*]

Extended Version

Shahram Ghandeharizadeh
University of Southern California
Los Angeles, California, USA
shahram@usc.edu

ABSTRACT

Unmanned Aerial Vehicles (UAVs) have moved beyond a platform for hobbyists to enable environmental monitoring, journalism, film industry, search and rescue, package delivery, and entertainment. This paper describes 3D displays using swarms of flying light specks, FLSs. An FLS is a small (hundreds of micrometers in size) UAV with one or more light sources to generate different colors and textures with adjustable brightness. A synchronized swarm of FLSs renders an illumination in a pre-specified 3D volume, an FLS display. An FLS display provides true depth, enabling a user to perceive a scene more completely by analyzing its illumination from different angles.

An FLS display may either be non-immersive or immersive. Both will support 3D acoustics. Non-immersive FLS displays may be the size of a 1980's computer monitor, enabling a surgical team to observe and control micro robots performing heart surgery inside a patient's body. Immersive FLS displays may be the size of a room, enabling users to interact with objects, e.g., a rock, a teapot. An object with behavior will be constructed using FLS-matters. FLS-matter will enable a user to touch and manipulate an object, e.g., a user may pick up a teapot or throw a rock. An immersive and interactive FLS display will approximate Star Trek's Holodeck.

A successful realization of the research ideas presented in this paper will provide fundamental insights into implementing a Holodeck using swarms of FLSs. A Holodeck will transform the future of human communication and perception, and how we interact with information and data. It will revolutionize the future of how we work, learn, play and entertain, receive medical care, and socialize.

1 INTRODUCTION

This brave new ideas paper describes a next generation multimedia display realized using *Flying Light Specks*, FLSs. An FLS is a miniature (hundreds of micrometer in size) Unmanned Aerial Vehicle (UAV) or[1] drone with either a reflective surface, one or more light sources, or both. The intensity and brightness of both the light sources and the reflective surface may be adjusted to display different colors and textures. A display may consist of a large swarm (millions, billions, and trillions) of FLSs that render a three dimensional (3D) depth scene that may include motion and 3D acoustics. We use the term *illuminations*, *luminations* for short, to refer to such renderings and prevent confusion with today's 2D images, e.g., JPEG [46], and motion graphics, e.g., MPEG [48].

FLS displays will complement today's flat screens and may incorporate them for certain application use cases; see the gaming application of Section 1.1. Advantages of FLS displays compared with today's flat screens will be multi-fold. First, their rendered luminations will capture reality more accurately. A lumination will be rendered by controlling position of FLSs in a 3D volume that has depth. Second, their 3D luminations provide depth to the naked eye. Moreover, a user looking at an FLS display from different angles may observe different content. Third, FLS displays may be immersive by occupying an arbitrary space, transforming a wall, a ceiling, a room, a concert hall into a 3D depth display.

With immersive displays, we introduce FLS-matter, a swarm of FLSs that will realize a geometric shape in 2D (e.g., square) and 3D (e.g., a pyramid). Its lumination will scale to render its shape in different sizes. FLS-matter is the building block of an object with behavior. As detailed in Section 4, it will enable a user to interact with objects rendered as luminations. An interactive immersive FLS display will manipulate both the user's depth perception and spatial awareness, approximating Star Trek's Holodeck.

Table 1: Number of FLSs for different sized spheres.

Sphere (radius in mm)	Surface area	Number of FLSs
Ping-pong (20)	5,027	19,040
Tennis (34.29)	14,776	55,968
Basketball (120)	180,956	685,438
Room size (1524)	29,186,351	110,554,359

To maintain the same resolution, the number of FLSs required by a lumination increases as a function of its size. Assuming today's 0.264 millimeter pixels, Table 1 shows the number of FLSs required to display a sphere as a starting point for a 3D Earth. As size is increased from a ping-pong to a basketball, the number of FLSs increases to hundreds of thousands. With a room sized (10 feet by 10 feet) lumination, the number of FLSs is hundreds of millions.

1.1 Applications

It is tempting to motivate FLS displays in the context of a visionary system such as Memex [11]. We side-step this temptation and motivate FLS displays using several short term application areas that would benefit from them if they were available today.

Health-care: FLS displays will be ideal interfaces for today's 3-D scanners such as the Magnetic Resonance Imaging [101] (MRI) and surgical devices such as the da Vinci Robot [12]. Today's MRI scanners show 3D volumes as a sequence of 2D images. They are ripe to use a true 3D depth FLS display. A typical da Vinci deployment provides its surgeon with a 3D video (using color depth) but

[*]A shorter version of this paper appeared in ACM-Multimedia Asia 2021 [35].
[1]This paper uses the terms UAV and drone synonymously.

not the assistants, even though they must collaborate. There have been studies that quantify the benefits of providing assistants with their own 3D video monitor using color depth [103]. A large FLS display will enable the surgeon and assistants to view the same 3D lumination to collaborate effectively.

Entertainment and gaming: FLS displays will transform entertainment and gaming applications. An example is multi-player fight games such as Mortal Kombat, Street Fighter, Fortnite and others. FLS displays enable players to view the fight scene from different angles. These luminations might be on a game board that sits on a table. It may incorporate today's flat screens as its bottom. For example, in a scene that shows multiple characters standing on a wooden flooring, a flat screen will render the wooden flooring while FLSs delineate the 3D fighting characters. Players may participate either remotely or in-person by either sitting or standing around the game board. Players may use either a traditional game controller or wear body sensors that translate their movement into luminations. With the latter, when the players are in the same room, they may see the posture of their opponent(s) and anticipate their next move, introducing new gaming experiences. Players may also get a good cardiovascular exercise from playing such games.

Design and manufacturing: Display of a 3D design using flat 2D displays is not natural and may result in human errors. 3D printing empowers a designer to evaluate a final design prior to manufacturing it. An FLS display will fill the gap between design and 3D print by enabling a designer to see their 3D designs and examine them from different perspectives in real-time. With this application, an FLS display maybe the size of a game board that renders small designs (knots and bolts of an airplane seat) or the size of a room for larger designs (the airplane). Both will enable a designer to zoom in and out to scrutinize components. This may be done using a traditional keyboard and mouse, a gesture based interface, or more novel interfaces (see Section 8). FLS displays may be an integral part of design. For example, they may render luminations of a simulation that stresses a design to analyze its failure thresholds. This may include lumination renderings from real environments with manufactured components, e.g., flight recorder data from a crashed airplane. These concepts apply to manufacturing where an FLS display may render the machine tools on a factory floor and how they manufacture a product, e.g., an automobile.

1.2 Contributions

The contributions of this paper are:

(1) The idea of 3D displays using swarms of FLSs.
(2) A variant of the first idea, immersive and interactive displays using FLS-matter, approximating Star Trek's Holodeck.
(3) Requirement of FLS displays formulated as research challenges, including a survey of the relevant literature.

To the best of our knowledge, the first two ideas are novel and not described elsewhere. See [1, 21, 84] for a survey of applications using UAVs and swarms.

The rest of this paper is organized as follows. Section 2 describes related work. Sections 3-9 detail challenges of FLS displays. Section 10 concludes this paper.

2 RELATED WORK

Our idea, a 3D display using swarm of FLSs, is inspired by today's drones used in light shows at night. These outdoor and indoor shows are performed using illuminated, synchronized, and choreographed groups of drones arranged into various aerial formations. Almost any image and motion graphic can be recreated using a computer program that controls flight path and lighting of drones. An FLS display is similar because each FLS is a drone and a lumination is realized by synchronizing FLSs.

With outdoor shows, a drone's GPS enables it to be at the right position and follow the correct path in the night sky to illuminate its light for the desired effect. Millimeter drones of an FLS display must be positioned with micrometer precision. Even if GPS was available, it would be too coarse [58]. The best some GPS devices claim is 3 meter accuracy 95% of the time. The United States government currently claims 7.8 meter with 95% confidence for horizontal accuracy for civilian (SPS) GPS. Real-Time Kinematic (RTK) GPS provides centimeter-level accuracy using additional base stations. Vertical accuracy is worse with GPS solutions. Thus, an FLS display requires novel positioning systems as described in Section 3.1.

There are other significant differences between show drones and FLS displays. First, the number of drones in FLS displays will be significantly larger. According to the Guinness World Record, the most unmanned aerial vehicles in simultaneous flight for a light show is 3,281 for tens of minutes spread over a kilometre in Shanghai, China on March 29, 2021 [114]. The number of drones in an FLS display will be several orders of magnitude larger. Second, the area used by an FLS display may be significantly smaller. An FLS display may be as small as a standard game board. This is orders of magnitude smaller than the night sky of outdoor drone shows and concert halls [83] of indoor shows. Third, each FLS is significantly smaller than show drones. This enables FLSs to operate in tightly constrained environments in tight formations with higher accelerations, providing higher stability and more rapid response to changes in luminations [53].

Systems such as Northrop Grumman's Virtual Immersive Portable Environment (VIPE) Holodeck surrounds a user with 2D monitors [65]. VIPE uses Kinect to detect a user action such as crawling and jumping to display the appropriate training content on the 2D monitors. Our proposed immersive 3D FLS display may replace VIPE's 2D monitors, manipulating both the trainee's depth perception and spatial awareness to provide a more realistic experience.

Virtual Reality (VR) [34, 52, 91] is a computer generated simulation of a 3D image or environment. A user may interact with this environment using equipment such as a head-mounted display with a screen inside or gloves fitted with sensors. Augmented Reality (AR) [29, 32, 91] uses technology such as a mobile phone, tablet, or AR glasses to present the user with an enhanced version of the real physical world. The enhancements may include digital visual elements, sound, or other sensory stimuli. An immersive FLS display is different than both by providing the human sense of touch with no gloves and true depth with no mobile phone or head mounts.

Fast 3D printing has been suggested as a mechanism to realize 3D displays [26]. The idea is for a 3D object to arise out of a puddle in real-time to show a 3D image, and move and shift shape (by growing parts) around the display to show motion graphics. This idea is inspired by the T-1000 robot from the Terminator 2 movie. To realize such a display, 3D printing must become orders of magnitude faster and polymer chemistry has been suggested as a solution [26]. An FLS display employs small drones with one or more light/reflective sources to generate color and texture. At the same time, an FLS display may incorporate fast 3D printing to represent objects in a lumination that are stationary enough to match the speed of 3D printing. This hybrid display requires lighting of the 3D printed objects. Alternative solutions may include backlighting or FLSs that project light onto these printed objects. The rest of this paper focuses on challenges of FLS displays and defers hybrids using 3D printing to future work.

2.1 Recording of luminations

Similar to today's cameras and camcorders that record a scene in 2D that is subsequently displayed using today's monitors and TVs, 3D depth sensing cameras and camcorders will record luminations. We use the term *depthcam* to refer to these devices.

Today's depthcams can be classified into those that use either Time-of-Flight (ToF), Triangulation, or both. For example, Microsoft's Kinect One uses a flash of light to illuminate a scene and thousands of sensors measure the return time from different parts of the scene. This pixelated ToF device measures depth of a million points per second. A similar technology is employed by iPhone X and its dot projector [43] to create a 3D map of the user's facial structure [62, 86]. The LiDAR Intel RealSense depthcam uses a variant of the same technique by employing Lidar used in some automobiles. This depthcam uses a proprietary MEMS micro scanning technology [112] to acquire depth data for indoor applications.

The D4xx series of Intel RealSense [112] depthcams use triangulation technique via stereoscopic imaging. It uses two cameras to record a scene and compare image patches from one camera with image patches in the other camera to collect depth data. The basic idea is that those objects that are closer to the camera are displaced more in the horizontal axis. A limitation of this technique is with low texture scenes that cause the image patches from the different cameras to become ambiguous, e.g., a white wall with no texture. In this scenario, a ToF technique using infrared (IR) light may be used to acquire depth data. This technique is used by Intel RealSense D43x and D45x depthcams.

A depthcam may include Inertial Measurement Unit (IMU) to capture 6 degrees of freedom data[2], e.g., Intel RealSense D435i. The structure from motion of the depthcam may be used to capture depth data [24]. Moreover, this data may be fused with data from other sensors to refine depth data and make it more accurate.

Today's depthcam technologies are used in games, to authenticate users, and to navigate a robot in an obstacle course. We foresee their evolution as recording devices for FLS displays.

3 CONTINUOUS DISPLAY OF MOTION LUMINATIONS

A rendering of a motion lumination must be continuous and respect the spatio-temporal requirements of its motion. This means FLSs must be at the right (x,y,z) coordinate of the display as a function of time and illuminate either (a) light at the pre-specified color, brightness, and texture or (b) provide a reflective surface for an external light source to generate color and texture.

Figure 1 shows the block diagram for continuous display adapted from the autonomous vehicles studies [14, 73]. Perception will transform streams of sensory data and the meta-data describing a lumination to a contextual understanding of the environment around each FLS and its goal. Planning refers to the decisions that must be made by each FLS to execute its goal from a given start position. It will generate streams of decisions that dictate the velocity, flight path, color/texture and brightness of light generated by each FLS. These streams may also control external and back lighting of a scene. A continuous display of a lumination is the ability of FLSs to execute the collision-free flight decisions in a timely manner.

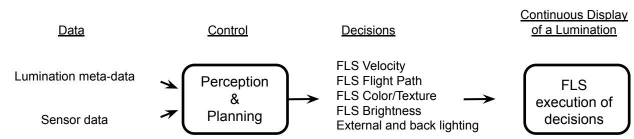

Figure 1: Continuous display using stream processing.

To illustrate, consider a lumination of a running person. If the pipeline of Figure 1 is processed too slow then the display may show walking instead of running. At its worst, the rendering may result in a distorted collection of lights in a 3D space. A continuous display avoids these undesirable luminations by requiring each FLS or grouping of FLSs to be at the right coordinate at the right time to render the right color.

Continuous display is challenging for several reasons. First, flight of FLSs must avoid collisions. This challenge is exacerbated by complex scenes that may require a large number of FLSs that must render their color in close proximity to one another. This is because no positioning system guarantees 100% accuracy and reported position of FLSs may be erroneous. Second, the scene to be displayed may not be known in advance. An example is a real-time stream generated by a RealSense depthcam. Third, failure of FLSs is the norm rather than the exception, see Section 6. A continuous display technique must render a lumination even though its FLSs are failing frequently. Fourth, battery lifetime of FLSs is finite and power is scarce. Fifth, the network bandwidth is finite and there is a latency for one FLS to communicate with another.

These challenges motivate algorithms and frameworks in support of a positioning system, a planner to compute flight path and lighting of each FLS as a function of time and space, a communication network for FLSs, and a system to charge FLS batteries. Below, we describe these components in turn.

3.1 Positioning system

An indoor positioning system is required for an FLS to position itself at the right coordinate to render a motion lumination. Even with a static lumination that has no motion, a positioning system is required to prevent FLSs from unpredictable drifting [75] that may cause crashes. Outdoor drones use the global positioning system, GPS, to remain stationary. GPS may not work indoors

[2]Six degrees of freedom refers to the freedom of movement of a depthcam in three-dimensional space: Forward/back, up/down, left/right, pitch, yaw, and roll.

due to lack of line of sight with orbiting satellites[3]. A 3D FLS display requires a fine-grained positioning system to enable FLSs to maintain and control their position with hundreds of micrometer accuracy and 95% confidence. Indoor Location Systems (ILS) have been proposed based on Magnetic signals, Vision-based, Audible sound, Infrared (IR), Ultra-sound (UPS), Radio Frequency (RF) including RFID and WLAN, Bluetooth, Ultra-Wideband (UWB) and Ultra High Frequency (UHF), or a hybrid of these technologies. See [38, 44, 60, 63, 113] for a survey of these technologies including their strengths and limitations.

Several studies report optical motion capture systems enable millimeter-level position tracking using multiple infrared cameras [53, 75, 80]. This solution requires a unique marker arrangement for each drone. The small size of an FLS may challenge the applicability of this solution. For example, [75] reports that it was impossible to form unique markers for 49 drones with each measuring 92 millimeters diagonally. They develop a method initialized with known positions that is subsequently updated using frame-by-frame tracking. While such techniques may be a good starting point, we believe decentralized positioning systems that include active participation of FLSs will be developed. In its simplest, FLSs may detect and either repel one another [55, 90] or maintain a fixed geometry. This will minimize the likelihood of FLS crashes when a positioning system provides inaccurate positions to FLSs.

3.2 Perception and Planning

A planner computes flight path and lighting of each FLS as a function of time and space. It may maneuver multiple FLSs in close proximity of one another to render a static lumination.

Several studiesdetail path planning algorithms for UAVs, e.g., outdoor light shows [89], search and rescue [42], services in an urban area [105], and environmental sensing [85]. A challenge is to move from a given initial positions to a set of predefined targets while avoiding collisions with obstacles as well as other UAVs [4, 5, 7, 13, 15, 19, 20, 30, 33, 39, 53, 68–70, 74, 75, 87, 92, 97–100, 108]. Some studies consider *downwash*, a region of instability caused by the flight of one UAV that adversely impacts other UAVs entering this region [4, 7, 33, 75, 100, 108], e.g., loss of control or unpredictable behavior. This effect is considered in collision avoidance by modelling each UAV as axis-aligned ellipsoids [4, 7, 75] or cylinders [33, 108] that results in a larger separation along the Z-axis. The latter is appropriate for quadrotors.

Today's collision avoidance techniques may be categorized into centralized [5, 15, 19, 39, 53, 70, 70, 74, 75, 89, 92] and decentralized [4, 7, 13, 20, 33, 68, 69, 87, 92, 97–100, 108]. A centralized algorithm executes on a CPU with abundant amount of memory and storage. A decentralized algorithm may use the limited processing and storage capability of each FLS, requiring multiple of them to collaborate by communicating with one another. Centralized algorithms may further be divided into online [15, 19, 75] or offline [5, 39, 53, 70, 70, 74, 89, 92]. Offline techniques assume a static environment with known obstacles, a known scene, and reliable UAVs. Online techniques minimize the number of such assumptions.

To illustrate, [89] presents a centralized, offline algorithm to compute a flight/lighting plan for outdoor light show performances. This algorithm requires drones to be placed in a field and in a specific arrangement. Each drone is provided with a pre-computed flight path and timed light display. All drones take off in a specific order, fly to the performance area, spread out and turn into the first performance pattern, change the formation into the next pattern, and repeat until all pre-specified performance patterns are complete. Finally, the drones return into a formation and fly back to the takeoff position. This algorithm is centralized because it assumes complete information of the field, intended display patterns and their duration, obstacles, and participating drones. It is offline because it executes once, producing a plan for each drone and communicating it to that drone. If during the light show, a new obstacle is introduced (e.g., a bird), an online algorithm is required to compute a new plan for the impacted drone(s). This online algorithm is more complex than the offline algorithm because it must monitor the environment, detect new obstacles and the impacted drones, compute a new plan with a new flight path for each drone, identify the drones whose flight paths have changed, and communicate the new flight path to these drones. An offline algorithm does not have these steps. It may not even require drones to communicate because the drones are provided with a flight path once at the beginning.

It is important to note that centralized offline algorithms are of value because they may guarantee optimality in terms of metrics such as minimum flight time, time to realize a lumination, or amount of power consumed. They can be used as yardsticks to evaluate online algorithms (both centralized and decentralized) that are typically heuristic based.

Collision avoidance techniques may further be categorized into those that evaluate their technique using a simulation [4, 20, 39, 68, 70, 75, 87, 98], an implementation using either outdoor UAVs, robots, or indoor quadrotors [5, 19, 92], or both [7, 13, 15, 33, 53, 69, 74, 89, 97, 99, 100, 108].

Existing techniques and algorithms are a great starting point to develop an online planning component in support of continuous display. An ideal algorithm or framework should provide a scalable execution time as a function of computing resources. This means a centralized algorithm should execute faster as a function of the number of CPU/GPU cores used to execute the algorithm. Similarly, a decentralized algorithm should execute faster as a function of the number of participating FLSs. Both concepts are challenging to realize in practice.

The concept of scalability extends beyond algorithms to include the display itself. For example, a scalable positioning system will enable players of a game to stretch their 13 inch game board to 40 inches. To maintain the resolution of luminations, users may add new FLSs. It is interesting to note that with a decentralized algorithm for planning, the additional FLSs will increase the processing capability of the system to compute spatio-temporal plans automatically. A centralized algorithm may require increased processing capability (e.g., additional cores) to schedule the new FLSs. This may require a shut-down and restart of a game board. An elastic decentralized algorithm will incorporate the new FLSs and continue the game without starving for processing capability.

[3] Even if GPSs worked indoors, its accuracy is too coarse for an FLS display, see Section 2.

3.3 Communication

FLSs must communicate to either receive a plan from a centralized planner or collaborate to compute a plan using a decentralized planner. The communication will almost certainly be wireless. The emerging 5 GHz 802.11n and 802.11ac offering real world bandwidths ranging from hundreds of megabits per second to more than 1 gigabit per second are promising. This is because a spatio-temporal plan consists of meta-data that describes positions as a function of time with lighting specifications. This data can be compressed. Moreover, a plan may apply to a group of FLSs and broadcasted to all FLSs once.

With a very large number of FLS, it is possible that FLSs exhaust the available wireless bandwidth. There are networking technologies such as the power efficient ZigBee (802.15.4, 2.4 GHz) that provide for mesh communication between FLSs that are in close proximity of one another. It may motivate use of multi-hop ad-hoc networks such as CAN [79] and Chord [88] that route information to an FLS in a decentralized manner. ZigBee is designed primarily for low-mobility devices such as those found in a home, e.g., smart bulbs and power outlets. It is a great starting point to implement a network with mobility standards for FLS displays. This network will enable FLSs flying in a tight formation (see FLS-Matter of Section 4) to communicate locally while traveling at high speeds. This local communication should be independent of the global communication, providing for a large number of independent communications that do not interfere with one another. In essence, the number of independent communications should scale as a function of the display size assuming fix sized FLSs.

3.4 Power Conscious Display

Continuous display algorithms must minimize the amount of power consumed to render a lumination. This maximizes battery life of the participating FLSs. For example, a lumination of a scene may consist of stationary objects that serve as its background, e.g., a building or a tree. Depending on the duration of the scene, the stationary objects may be realized using FLSs that are docked into one another either horizontally and/or vertically. These FLSs may use the base or sides of a display for support to stop flying altogether to save power. The base or sides of the display may serve as a power source to charge the batteries of these FLSs. These structures may serve as charging stations for other FLSs in their close vicinity.

The emerging wireless chargers by companies such as Aeterlink, Motorola, Xiaomi, Oppo and others [104] may be used to charge FLS batteries. During CES 2021, Motorola demonstrated a charging technology that charged several phones at up to 1 meter away [72]. Aeterlink's Airplug claims to power devices up to 20 meters away [9, 10]. These may be deployed in FLS game boards and non-immersive displays to charge FLS batteries continuously. A challenge is for this technology is to charge a large number of FLSs simultaneously.

Wireless charging technology may not be appropriate for immersive FLS displays as wireless chargers emit EMF radiation which has been shown to be harmful to the human body [6]. This motivates a framework that recycles FLSs continuously. This framework will replenish those FLSs with exhausted batteries with new ones. It will charge their batteries and recycle them as new. Note that this framework must be in the background and not interfere with continuous display of motion luminations.

With a motion lumination, an algorithm may pre-stage dark FLSs that light up to facilitate continuous motion. Pre-staging FLSs will consume power. However, its use may be more power efficient than flying FLSs around to render the lumination, e.g., by reducing the total distance travelled by FLSs.

There will be a tradeoff between the quality of a lumination and its consumed power. A display may sacrifice the quality of a lumination by using fewer FLSs, skipping positioning of certain FLSs all together, and controlling the brightness of light produced by FLSs. These tradeoffs will have thresholds. For example, a bright lumination will consume more power and will most likely result in a higher quality display up to a certain threshold. Beyond this threshold, increasing brightness will only consume power without providing a tangible human perceived improvement in quality. This perception will most likely be impacted by environmental factors external to a display, e.g., lightning of the room. Hence, an algorithms must identify the relevant environmental factors and incorporate them when trading power for quality of a lumination.

A technique may save power by not rendering those portions of a lumination that are not visible to a user. It may employ an eye tracker to compute the user's angle of view and identify regions of a 3D lumination that are hidden from the user. An algorithm will use this information to minimize the number of FLSs with no impact on the user perceived quality of lumination, saving power.

As FLS displays are constructed, additional factors will almost certainly be identified that impact the tradeoff between power and quality of display.

4 FLS-MATTER

FLS-matter is a pre-specified swarm of FLSs for 2D (e.g., square) and 3D (e.g., a pyramid) geometric shapes. It is the building block of objects with behavior. Example objects include a rock, water, a teapot, a sword, etc. The behavior associated with an object is inherited by its FLS-matters. This behavior may further be specialized at the granularity of each FLS-matter.

FLS-matter detects human touch (force) when it is perturbed by having its FLSs pushed out of their expected position. The default behavior will be for the FLS-matter to stop rendering its lumination, i.e., its FLSs go dark and scatter. An object may over-ride this behavior. For example, it may model gravity to disintegrate its FLS-matters more naturally by having them fall to the bottom of the display prior to scattering and going dark.

An interesting research direction is how to model a swarm adjusting its flight pattern to exert force back at the human touch without breaking FLSs or causing injury to the user. This is more natural, requiring the human to exert enough force (specified as a threshold by the object behavior) to justify an object breaking. It requires models that quantify the amount of force exerted by a swarm based on its speed, flight pattern, mass of its FLSs, and the number of FLSs used at the perturbation point. For example, if an object is a rock and the user picks it up (exerts sufficient force against gravity) then the lumination must move up. If the user stops exerting force then the lumination must fall down at a speed dictated by the gravity pull and object mass. Upon impact to the display floor, different objects may behave differently. For example, with a rock, if the impact force exceeds a threshold then its lumination may fragment into smaller luminations (rock pieces). These are a few example behaviors assigned to an object and inherited by its FLS-matter(s). Linden Lab's Second Life provides many example objects and avatars in its market place [54]. An immersive FLS display will be able to render these objects and avatars, providing users with an interactive 3D experience that is significantly richer than today's 2D monitors and screens of mobile devices.

5 3D ACOUSTICS

It is important for FLS displays to address acoustics from the start instead of an afterthought. FLS displays require noise reduction [45] to minimize and possibly eliminate the impact of unwanted noise attributed to their mechanical components, e.g., the buzzing sound of their rotors. In addition, they require sound synthesis, propagation, and rendering techniques. These techniques enhance the sense of realism and presence [56, 61] for FLS displays. This is especially true for immersive FLS displays that require spatialized audio rendering. We describe each in turn.

Synthesis generates sound for an object realized using FLS-matter. This may be aerodynamic sound such as that produced by an FLS-matter sword when waved in the air [27], blades of a fan or wind turbine rotating [107], and instruments that produce sound through air vibration [2, 96]. Other examples [61] include flame sound [16, 47, 64] and water bubble sound [25, 28].

Propagation describes what a human ear perceives as sound transmitted from its source travels and changes due to factors such as scattering, reflection, refraction, diffraction and others. A recent survey [61] categorizes propagation techniques into wave-based [17, 31, 66, 77, 115], geometric [3, 18, 23, 51, 59, 78], diffraction [8, 50, 82, 94], and their hybrids [37, 81, 110]. We describe a wave-based propagation technique in Section 5.1. For an overview of the other techniques see [61].

Rendering is the pipeline to display complex scenes with hundreds of sounds [49, 76, 102]. Sounds may be associated with moving objects [95]. A pipeline will include devices to deliver sound to a user, e.g., speakers, ear phones, etc. With gaming, an FLS displays will almost certainly use an audio middleware such as FMOD. This middleware adds sound or music to a game by providing APIs that receive events from the game engine. It provides a workflow independent of the visuals that is integrated with the visuals [61].

5.1 Wave-based sound propagation

This section presents a wave-based propagation technique [66, 77] for immersive FLS displays. Subsequently, it describes the role of machine learning [93] and generative adversarial networks [36] to expedite this technique.

An object-oriented approach to audio wave sound propagation models each object with an acoustic behavior characterized by a function [66, 77]. This function maps an arbitrary incident field on the object to a scattered field. Inter-object relationships map the outgoing scatter field of one object to the incident field on an inter-connected object. An FLS display will compute the acoustic response of a lumination to a sound source by solving a global linear system that considers inter-object wave propagation. It performs a summation for all outgoing sources across all objects at the listener location for each ear. The listener may wear ear phones to hear the 3D stereoscopic acoustics.

A recent study approximates the acoustic scatting field of an object using neural networks for interactive sound propagation in dynamic scenes [93]. This technique requires an approximate training dataset and long training time. Moreover, it has no error bounds on arbitrary objects.

A hybrid approach that combines both the wave sound propagation and the neural network learning approach may be more suitable for FLS displays. The idea is to use generative adversarial network [36], GANs, to simultaneously train two models: a generative model G that captures the data distribution and a discriminative model D that estimates the probability that a sample came from the training data. The training procedures for G is to maximize the probability of D making a mistake. G can be thought of as producing incorrect sound at the listener location. D can be thought of as detecting the incorrect sound. Training of the models is complete once G recovers the training data distribution and D equals to $\frac{1}{2}$ everywhere. The strength of this approach is that the object oriented approach to audio wave sound propagation may generate training data specific to an application. As the training database grows in size, the adversarial network trains itself to produce accurate results. Representation of this data to capture the key elements of an application will have a significant impact on the correctness of the network. It will constitute a main focus of FLS displays.

6 FAILURES

FLSs will fail and must be garbage collected. Some failed FLSs will recover and become operational after sometime similar to today's commodity servers. An FLS display must detect these and recycle them as operational.

FLSs will fail for several reasons. First, each FLS is a special-purpose computing system (similar to a cell phone) that may reboot. Its reboot may cause it to fall to the base of a display and break. Second, each FLS is a mechanical device with a mean time to failure, MTTF. Given a large number of FLSs and assuming independent failures, the likelihood of *an* FLS failing is proportional to the number of FLSs[4]. Third, the positioning system (see Section 3.1) will most likely not be 100% accurate. Similar to today's GPS, it will provide positions with some confidence. When reported positions are erroneous, FLSs may crash into one another or the bottom of the display and break. Fourth, the heuristics that compute the fly path for individual FLSs may not be 100% accurate, causing FLSs crashes for complex scenes. These broken FLSs must be garbage collected. A display may include redundant FLSs that substitute for the broken FLSs. Overtime, a display may exhaust its redundant FLSs. At that point, it must either be replaced with a new display or repaired by removing its garbage FLSs and replenishing it with new operational FLSs.

It is possible for an FLS that rebooted to fall without breaking. In this case, once its recovers, it will most likely identify itself as a functional FLS to the display and be re-incorporated.

[4]This is similar to a disk failing frequently in a thousand disk RAID even though each disk is highly reliable, i.e., has a high MTTF [71].

An interesting research direction will be to design robust algorithms that use fewer FLSs than required to provide the highest quality lumination. Another is for a failed FLS to repel other FLSs as it falls to the ground along an arbitrary path. A solution will enable FLSs to move away from the failed FLS in different contexts, e.g., rendering a complex scene, FLS-matter, etc.

An FLS display is comparable to today's data centers with a large number of servers. Some of the errors and failures seen in data centers may manifest themselves in an FLS display. It is important to identify lessons learnt from these errors and incorporate them in FLS displays to enhance their reliability.

7 DISPLAY OF MISSING DATA

A file containing a lumination may lack the 3D depth, color, texture, and brightness of FLSs from all possible view angles. With an FLS display, a user looking at a lumination from one of these angles with no data may observe empty gaps. A simple solution is to fill the gaps with FLSs that transition color from the foreground object to the background objects gradually. Such a rendering must consider the volume occupied by each object. Its limitation is that it may produce un-realistic illuminations. For example, with a picture of people standing in front of a car, there is a clear gap between people and the car. Filling the gap to look continuous is not realistic. It will look more un-realistic as the distance between the people and the car increases.

An alternative is to infer the 3D geometry and structure of objects in the scene and render them to fill the empty gaps in a lumination. There exist techniques that infer 3D geometry from one or multiple 2D images [22, 41, 57, 106, 109, 111].See [40] for a survey of deep learning techniques. A key challenge with today's techniques is that they must estimate depth from the 2D images. This estimation will be significantly more accurate with a depthcam (see Section 2.1). Instead of using 2D images, an FLS display will use a database of objects and their 3D geometry to fill the empty gaps. It may include a user in the loop, expanding the system's database with the missing objects that are personal to that user, e.g., a user's pet.

Certain applications may require a hybrid of solutions. For example, with a medical application, it may be best to show the empty gaps to a physician to inform them of the missing data. The physician may then request the FLS display to fill the gaps by inferring the 3D geometry and structure using a database of the human anatomy or a deep learning technique trained using this database.

8 NOVEL USER INTERFACES

Applications of Section 1.1 highlight the use of keyboard and mouse, gaming consoles, and gesture control. While it is important for FLS displays to be backward compatible with traditional user interfaces, they may enable novel user interfaces. For example, a user wearing a haptic glove may reach into an FLS display with their hand and interact with objects that constitute a lumination. This is similar to today's touch screen but in 3D. This enables a surgeon to reach into an FLS display and separate human organs rendered as one lumination into different luminations and examine each individually. Immersive room-size (and stadium-size) FLS displays will enable multiple users wearing special outfits to interact with virtual objects and one another. We believe applications of FLS displays will motivate novel user interfaces that are more intuitive, effective, and relevant than today's user interfaces [67].

9 STANDARDS AND MARKET FORCES

Each FLS must cost less than 1 cent to realize an economically viable 3D display. Consider a 3D display with one million FLSs. At 1 cent per FLS, this display costs $10,000. Even with one hundred thousand FLSs, the display still costs $1,000. Still expensive given that we are not considering the cost associated with the other display components, e.g., positioning system, power, casing, etc. Past trends in consumer electronic goods show a significant drop in costs as they are adopted widely. We believe mass production of FLS displays will reduce the cost of each FLS to lower than 1 cent.

Standards will play an important role as FLS displays are developed. From a micro perspective, these standards will identify functionality implemented internal and external to a display. These standards will enable an FLS display to inter-operate with diverse devices, e.g., gaming consoles by different manufacturers, a set-top box for 3D shows, a computer server, a mobile device such as a smart phone, etc. They will detail formatting of a lumination and interfaces for middle-wares to enable different vendors to develop software and content for FLS displays. They will also ensure the display is backward compatible with today's 2D images and videos. From a macro perspective, standards will produce displays with features not available in today's displays. One example is scalable displays that may adjust their size. A user with a small (15 inch front) display will be able to decrease or increase the display size (say to 13 or 24 inches), see Section 3.

10 CONCLUSIONS

3D displays using swarms of FLSs is a transformative idea that will pioneer a new era in multimedia systems. Immersive and interactive 3D displays will resemble Star Trek's Holodeck. Their realization will revolutionize the future of how we work, learn, play and entertain, receive medical care, communicate, and socialize.

While the technical challenges of FLS displays may appear daunting, it is important to remember the humble beginnings of today's displays as televisions with cathode ray tubes and a mechanical scanning system invented in[5] 1907. Computer monitors using CRT technology with a monochrome display were first introduced in 1973. While LED display technology was developed in 1977, it was not readily available for purchase until about 30 years later with release of one in late 2009. LCD displays [6] were introduced in the mid 1990s and outsold CRT monitors for the first time in 2003. LCD displays became dominant starting 2007.

Non-immersive FLS displays may observe an evolution similar to today's monitors to become a consumer electronic good. We are hopeful that their introduction and adoption will be faster.

Immersive FLS displays raise many interesting research challenges, see discussion of FLS-matter in Section 4. It must consider safety of users and may require a longer development, deployment, and adoption life cycle.

[5] The first electronic television with no mechanical scanning was invented in 1927.

[6] The difference between a LED and a LCD display is the backlighting. The first LCD monitors used CCFL instead of LEDs to illuminate the screen.

11 ACKNOWLEDGMENTS

We thank ACM-Multimedia Asia's anonymous reviewers (Brave New Ideas track) for their valuable comments.

REFERENCES

[1] Yousef Alghamdi, Arslan Munir, and H. La. 2021. Architecture, Classification, and Applications of Contemporary Unmanned Aerial Vehicles. *IEEE Consumer Electronics Magazine* (2021), 1–10.

[2] Andrew Allen and Nikunj Raghuvanshi. 2015. Aerophones in Flatland: Interactive Wave Simulation of Wind Instruments. *ACM Trans. Graph.* 34, 4, Article 134 (July 2015), 11 pages. https://doi.org/10.1145/2767001

[3] Lakulish Antani, Anish Chandak, Lauri Savioja, and Dinesh Manocha. 2012. Interactive Sound Propagation Using Compact Acoustic Transfer Operators. *ACM Trans. Graph.* 31, 1, Article 7 (Feb. 2012), 12 pages. https://doi.org/10.1145/2077341.2077348

[4] Senthil Hariharan Arul and D. Manocha. 2020. DCAD: Decentralized Collision Avoidance With Dynamics Constraints for Agile Quadrotor Swarms. *IEEE Robotics and Automation Letters* 5 (2020), 1191–1198. https://doi.org/10.1109/LRA.2020.2967281

[5] Federico Augugliaro, Angela Schoellig, and Raffaello D'Andrea. 2012. Generation of Collision-Free Trajectories for a Quadrocopter Fleet: A Sequential Convex Programming Approach. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE/RSJ International Conference on Intelligent Robots and Systems.* 1917–1922. https://doi.org/10.1109/IROS.2012.6385823

[6] Azab Azabl, Khalat Abdurazag, Ebrahim Shaban, and Albasha Mohamed. 2018. Electromagnetic Fields and Its Harmful Effects on the Male Reproductive System. *AIS Bioscience and Bioengineering* 4, 1 (2018), 1–13.

[7] Daman Bareiss and Joran van den Berg. 2013. Reciprocal Collision Avoidance for Robots with Linear Dynamics using LQR-Obstacles. In *Proceedings - IEEE International Conference on Robotics and Automation.* 3847–3853. https://doi.org/10.1109/ICRA.2013.6631118

[8] M. A. Biot and I. Tolstoy. 1957. Formulation of Wave Propagation in Infinite Media by Normal Coordinates with an Application to Diffraction. *The Journal of the Acoustical Society of America* 29, 3 (1957), 381–391. https://doi.org/10.1121/1.1908899 arXiv:https://doi.org/10.1121/1.1908899

[9] Sascha Brodsky. Lifewire, May 18, 2021, https://www.lifewire.com/your-smartphone-could-soon-charge-over-the-air-5185196. Your Smartphone Could Soon Charge Over the Air.

[10] Dalvin Brown. The Washington Post, March 5, 2021, https://www.washingtonpost.com/technology/2021/03/05/wireless-charging-over-the-air/. One Innovation We Won't be Seeing Soon: Over-The-Air Charging.

[11] Vannevar Bush. 1945. As We May Think. *Atlantic Monthly* 176 (1945), 101–108.

[12] John C. Byrn, Stefanie Schluender, Celia M. Divino, John Conrad, Brooke Gurland, Edward Shlasko, and Amir Szold. 2007. Three-Dimensional Imaging Improves Surgical Performance for both Novice and Experienced Operators using the da Vinci Robot System. *The American Journal of Surgery* 193, 4 (2007), 519–522. https://doi.org/10.1016/j.amjsurg.2006.06.042

[13] Tomáš Báča, Daniel Hert, Giuseppe Loianno, Martin Saska, and Vijay Kumar. 2018. Model Predictive Trajectory Tracking and Collision Avoidance for Reliable Outdoor Deployment of Unmanned Aerial Vehicles. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS).* 6753–6760. https://doi.org/10.1109/IROS.2018.8594266

[14] Mitch Campion, Prakash Ranganathan, and Saleh Faruque. 2019. UAV Swarm Communication and Control Architectures: A Review. *Journal of Unmanned Vehicle Systems* 7, 2 (2019), 93–106. https://doi.org/10.1139/juvs-2018-0009

[15] D. Cappello, S. Garcin, Z. Mao, M. Sassano, A. Paranjape, and T. Mylvaganam. 2020. A Hybrid Controller for Multi-Agent Collision Avoidance via a Differential Game Formulation. *IEEE Transactions on Control Systems Technology* PP (07 2020), 1–8. https://doi.org/10.1109/TCST.2020.3005602

[16] Jeffrey N. Chadwick and Doug L. James. 2011. Animating Fire with Sound. *ACM Trans. Graph.* 30, 4, Article 84 (July 2011), 8 pages. https://doi.org/10.1145/2010324.1964979

[17] Chakravarty R. Alla Chaitanya, Nikunj Raghuvanshi, Keith W. Godin, Zechen Zhang, Derek Nowrouzezahrai, and John M. Snyder. 2020. Directional Sources and Listeners in Interactive Sound Propagation Using Reciprocal Wave Field Coding. *ACM Trans. Graph.* 39, 4, Article 44 (July 2020), 14 pages. https://doi.org/10.1145/3386569.3392459

[18] Anish Chandak, Christian Lauterbach, Micah Taylor, Zhimin Ren, and Dinesh Manocha. 2008. AD-Frustum: Adaptive Frustum Tracing for Interactive Sound Propagation. *IEEE Transactions on Visualization and Computer Graphics* 14, 6 (2008), 1707–1722. https://doi.org/10.1109/TVCG.2008.111

[19] Yu Fan Chen, Mark Cutler, and Jonathan How. 2015. Decoupled Multiagent Path Planning via Incremental Sequential Convex Programming. *Proceedings - IEEE International Conference on Robotics and Automation* 2015 (06 2015), 5954–5961. https://doi.org/10.1109/ICRA.2015.7140034

[20] Hui Cheng, Q. Zhu, Z. Liu, Tianye Xu, and Liang Lin. 2017. Decentralized Navigation of Multiple Agents Based on ORCA and Model Predictive Control. *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (2017), 3446–3451. https://doi.org/10.1109/IROS.2017.8206184

[21] Soon-Jo Chung, A. Paranjape, P. Dames, S. Shen, and Vijay R. Kumar. 2018. A Survey on Aerial Swarm Robotics. *IEEE Transactions on Robotics* 34 (2018), 837–855.

[22] Brian Curless and Marc Levoy. 1996. A Volumetric Method for Building Complex Models from Range Images. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '96).* Association for Computing Machinery, New York, NY, USA, 303–312. https://doi.org/10.1145/237170.237269

[23] Norm Dadoun, David G. Kirkpatrick, and John P. Walsh. 1985. The Geometry of Beam Tracing. In *Proceedings of the First Annual Symposium on Computational Geometry* (Baltimore, Maryland, USA) *(SCG '85).* Association for Computing Machinery, New York, NY, USA, 55–61. https://doi.org/10.1145/323233.323241

[24] Andrew J. Davison. 2003. Real-Time Simultaneous Localisation and Mapping with a Single Camera. In *Proceedings of the Ninth IEEE International Conference on Computer Vision - Volume 2 (ICCV '03).* IEEE Computer Society, USA, 1403.

[25] Grant B. Deane. 2013. Determining the Bubble Cap Film Thickness of Bursting Bubbles from their Acoustic Emissions. *The Journal of the Acoustical Society of America* 133 2 (2013), EL69–75.

[26] Joseph DeSimone. Onstage at TED2015. See https://www.ted.com/talks/joseph_desimone_what_3d_printing_was_100x_faster?language=en#t-20146. What if 3D Printing was 100x Faster?

[27] Yoshinori Dobashi, Tsuyoshi Yamamoto, and Tomoyuki Nishita. 2003. Real-Time Rendering of Aerodynamic Sound Using Sound Textures Based on Computational Fluid Dynamics. *ACM Trans. Graph.* 22, 3 (July 2003), 732–740. https://doi.org/10.1145/882262.882339

[28] Kees van den Doel. 2005. Physically Based Models for Liquid Sounds. *ACM Trans. Appl. Percept.* 2, 4 (Oct. 2005), 534–546. https://doi.org/10.1145/1101530.1101554

[29] Ruofei Du, Eric Turner, Maksym Dzitsiuk, Luca Prasso, Ivo Duarte, Jason Dourgarian, João Afonso, Jose Pascoal, Josh Gladstone, Nuno Cruces, Shahram Izadi, Adarsh Kowdle, Konstantine Tsotsos, and David Kim. 2020. DepthLab: Real-time 3D Interaction with Depth Maps for Mobile Augmented Reality. In *UIST '20: The 33rd Annual ACM Symposium on User Interface Software and Technology, Virtual Event, USA, October 20-23, 2020*, Shamsi T. Iqbal, Karon E. MacLean, Fanny Chevalier, and Stefanie Mueller (Eds.). ACM, 829–843. https://doi.org/10.1145/3379337.3415881

[30] T. Engelhardt, T. Konrad, Bjorn E. Schafer, and D. Abel. 2016. Flatness-Based Control for a Quadrotor Camera Helicopter using Model Predictive Control Trajectory Generation. *24th Mediterranean Conference on Control and Automation (MED)* (2016), 852–859.

[31] Graeme Fairweather, Andreas Karageorghis, and P.A. Martin. 2003. The Method of Fundamental Solutions for Scattering and Radiation Problems. *Engineering Analysis with Boundary Elements* 27, 7 (2003), 759–769. https://doi.org/10.1016/S0955-7997(03)00017-1 Special issue on Acoustics.

[32] Steven Feiner, Blair Macintyre, Tobias Höllerer, and Anthony Webster. 1997. A Touring Machine: Prototyping 3D Mobile Augmented Reality Systems for Exploring the Urban Environment. *Personal Technologies* 1, 74–81. https://doi.org/10.1007/BF01682023

[33] Eduardo Ferrera, Alfonso Alcántara, J. Capitán, Á. R. Castaño, P. Marrón, and A. Ollero. 2018. Decentralized 3D Collision Avoidance for Multiple UAVs in Outdoor Environments. *Sensors (Basel, Switzerland)* 18 (2018).

[34] Scott S. Fisher. 2016. The Nasa Ames Viewlab Project-a Brief History. *Presence: Teleoper. Virtual Environ.* 24, 4 (Dec. 2016), 339–348. https://doi.org/10.1162/PRES_a_00277

[35] Shahram Ghandeharizadeh. 2021. Holodeck: Immersive 3D Displays Using Swarms of Flying Light Specks. In *ACM Multimedia Asia* (Gold Coast, Australia). https://doi.org/10.1145/3469877.3493698

[36] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron C. Courville, and Yoshua Bengio. 2014. Generative Adversarial Nets. In *Advances in Neural Information Processing Systems 27: Annual Conference on Neural Information Processing Systems 2014, December 8-13 2014, Montreal, Quebec, Canada*, Zoubin Ghahramani, Max Welling, Corinna Cortes, Neil D. Lawrence, and Kilian Q. Weinberger (Eds.). 2672–2680. https://proceedings.neurips.cc/paper/2014/hash/5ca3e9b122f61f8f06494c97b1afccf3-Abstract.html

[37] Emmanuel Granier, Mendel Kleiner, Bengt-Inge Dalenbäck, and Peter Svensson. 1996. Experimental Auralization of Car Audio Installations. *Journal of The Audio Engineering Society* 44 (1996), 835–849.

[38] Yanying Gu, Anthony C. C. Lo, and Ignas G. Niemegeers. 2009. A Survey of Indoor Positioning Systems for Wireless Personal Networks. *IEEE Commun. Surv. Tutorials* 11, 1 (2009), 13–32. https://doi.org/10.1109/SURV.2009.090103

[39] Michael Hamer, Lino Widmer, and Raffaello Drandrea. 2018. Fast Generation of Collision-Free Trajectories for Robot Swarms Using GPU Acceleration. *IEEE*

Access PP (12 2018), 1–1. https://doi.org/10.1109/ACCESS.2018.2889533

[40] X. Han, H. Laga, and M. Bennamoun. 2021. Image-Based 3D Object Reconstruction: State-of-the-Art and Trends in the Deep Learning Era. *IEEE Transactions on Pattern Analysis & Machine Intelligence* 43, 05 (may 2021), 1578–1604. https://doi.org/10.1109/TPAMI.2019.2954885

[41] Richard Hartley and Andrew Zisserman. 2003. *Multiple View Geometry in Computer Vision* (2 ed.). Cambridge University Press, USA.

[42] Samira Hayat, Evsen Yanmaz, Christian Bettstetter, and Timothy Brown. 2020. Multi-objective Drone Path Planning for Search and Rescue with Quality-of-Service Requirements. *Autonomous Robots* 44 (09 2020). https://doi.org/10.1007/s10514-020-09926-9

[43] Kristofer Henderson, Xiaomeng Liu, Justin Folden, Brevin Tilmon, Suren Jayasuriya, and Sanjeev Koppal. 2020. Design and Calibration of a Fast Flying-Dot Projector for Dynamic Light Transport Acquisition. *IEEE Transactions on Computational Imaging* 6 (2020), 529–543. https://doi.org/10.1109/TCI.2020.2964246

[44] Joon Heo, Mu Wook Pyeon, Jung Whan Kim, and Hong-Gyoo Sohn. 2007. Towards the Optimal Design of an RFID-Based Positioning System for the Ubiquitous Computing Environment. In *Proceedings of the 2nd International Conference on Rough Sets and Knowledge Technology* (Toronto, Canada) *(RSKT'07)*. Springer-Verlag, Berlin, Heidelberg, 331–338.

[45] W. Herkes, R. Olsen, and S. Uellenberg. 2012. The Quiet Technology Demonstrator Program: Flight Validation of Airplane Noise-Reduction Concepts. In *12th AIAA/CEAS Aeroacoustics Conference (27th AIAA Aeroacoustics Conference)*. https://doi.org/10.2514/6.2006-2720 arXiv:https://arc.aiaa.org/doi/pdf/10.2514/6.2006-2720

[46] Graham Hudson, Alain Léger, Birger Niss, Istvan Sebestyén, and Jørgen Vaaben. 2018. JPEG-1 Standard 25 Years: Past, Present, and Future Reasons for a Success. *J. Electronic Imaging* 27, 04 (2018). https://doi.org/10.1117/1.JEI.27.4.040901

[47] Matthias Ihme, Heinz Pitsch, and Daniel Bodony. 2009. Radiation of Noise in Turbulent Non-Premixed Flames. *Proceedings of the Combustion Institute* 32, 1 (2009), 1545–1553. https://doi.org/10.1016/j.proci.2008.06.137

[48] ISO/IEC JTC 1/SC 29, SC 29/WG 11 2001. *Coding of Moving Pictures and Audio*. Standard. International Organization for Standardization, Tokyo, Japan.

[49] Jean-Marc Jot and Antoine Chaigne. 1991. Digital Delay Networks for Designing Artificial Reverberators. *Journal of The Audio Engineering Society* (1991).

[50] R.G. Kouyoumjian and P.H. Pathak. 1974. A Uniform Geometrical Theory of Diffraction for an Edge in a Perfectly Conducting Surface. *Proc. IEEE* 62, 11 (1974), 1448–1461. https://doi.org/10.1109/PROC.1974.9651

[51] A. Krokstad, S. Strom, and S. Sørsdal. 1968. Calculating the Acoustical Room Response by the Use of a Ray Tracing Technique. *Journal of Sound and Vibration* 8, 1 (1968), 118–125. https://doi.org/10.1016/0022-460X(68)90198-3

[52] Logan Kugler. 2021. The State of Virtual Reality Hardware. *Commun. ACM* 64, 2 (Jan. 2021), 15–16. https://doi.org/10.1145/3441290

[53] Alex Kushleyev, Daniel Mellinger, Caitlin Powers, and Vijay Kumar. 2013. Towards a Swarm of Agile Micro Quadrotors. *Autonomous Robots* 35 (11 2013), 573–7527. https://doi.org/10.1007/s10514-013-9349-9

[54] Linden Labs. 2021. Second Life Marketplace, Oct 28, 2021. See https://marketplace.secondlife.com/.

[55] Emmett Lalish. 2009. Distributed Reactive Collision Avoidance, Ph.D. Disertation, University of Washington.

[56] Pontus Larsson and Mendel Kleiner. 2002. Better Presence and Performance in Virtual Environments by Improved Binaural Sound Rendering. In *Virtual, Synthetic, and Entertainment Audio conference*. AES, Espoo, Finland.

[57] A. Laurentini. 1994. The Visual Hull Concept for Silhouette-Based Image Understanding. *IEEE Trans. Pattern Anal. Mach. Intell.* 16, 2 (Feb. 1994), 150–162. https://doi.org/10.1109/34.273735

[58] Leong Lee, Matthew Jones, Gregory S. Ridenour, Sam J. Bennett, Arisha C. Majors, Bianca L. Melito, and Michael J. Wilson. 2016. Comparison of Accuracy and Precision of GPS-Enabled Mobile Devices. In *2016 IEEE International Conference on Computer and Information Technology (CIT)*. IEEE, 73–82. https://doi.org/10.1109/CIT.2016.94

[59] T. Lewers. 1993. A Combined Beam Tracing and Radiant Exchange Computer Model of Room Acoustics. *Applied Acoustics* 38, 2 (1993), 161–178. https://doi.org/10.1016/0003-682X(93)90049-C

[60] Hui Liu, Houshang Darabi, Pat P. Banerjee, and Jing Liu. 2007. Survey of Wireless Indoor Positioning Techniques and Systems. *IEEE Trans. Syst. Man Cybern. Part C* 37, 6 (2007), 1067–1080. https://doi.org/10.1109/TSMCC.2007.905750

[61] Shiguang Liu and Dinesh Manocha. 2020. Sound Synthesis, Propagation, and Rendering: A Survey. *CoRR* abs/2011.05538 (2020). arXiv:2011.05538 https://arxiv.org/abs/2011.05538

[62] Keith Lyon and Arnaud Laflaquiere. Apple Inc., U.S. Patent 2020/0251882, 2020. Vertical Emitters with Integral Microlenses.

[63] Luca Mainetti, Luigi Patrono, and Ilaria Sergi. 2014. A Survey on Indoor Positioning Systems. In *22nd International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*. IEEE, 111–120. https://doi.org/10.1109/SOFTCOM.2014.7039067

[64] Damián Marelli, Mitsuko Aramaki, Richard Kronland-Martinet, and Charles Verron. 2010. Time-Frequency Synthesis of Noisy Sounds with Narrow Spectral Components. *Trans. Audio, Speech and Lang. Proc.* 18, 8 (Nov. 2010), 1929–1940.

[65] Allen McDuffee. 2014. A Holodeck Videogame Designed to Train Soldiers, Jan, 2014. See https://www.wired.com/2014/01/holodeck/.

[66] Mehra, Ravish and Raghuvanshi, Nikunj and Antani, Lakulish and Chandak, Anish and Curtis, Sean and Manocha, Dinesh. 2013. Wave-Based Sound Propagation in Large Open Scenes Using an Equivalent Source Formulation. *ACM Trans. Graph.* 32, 2, Article 19 (April 2013), 13 pages. https://doi.org/10.1145/2451236.2451245

[67] Silvia Mirri, Catia Prandi, and Paola Salomoni. 2019. Human-Drone Interaction: State of the Art, Open Issues and Challenges. In *Proceedings of the ACM SIGCOMM 2019 Workshop on Mobile AirGround Edge Computing, Systems, Networks, and Applications* (Beijing, China) *(MAGESys'19)*. Association for Computing Machinery, New York, NY, USA, 43–48. https://doi.org/10.1145/3341568.3342111

[68] Daniel Morgan, Soon-Jo Chung, and Fred Hadaegh. 2013. Decentralized Model Predictive Control of Swarms of Spacecraft Using Sequential Convex Programming. In *Advances in the Astronautical Sciences*.

[69] Daniel Morgan, Soon-Jo Chung, and Fred Y. Hadaegh. 2014. Model Predictive Control of Swarms of Spacecraft Using Sequential Convex Programming. *Journal of Guidance, Control, and Dynamics* 37, 6 (2014), 1725–1740. https://doi.org/10.2514/1.G000218

[70] Thulasi Mylvaganam, Mario Sassano, and Alessandro Astolfi. 2017. A Differential Game Approach to Multi-Agent Collision Avoidance. *IEEE Trans. Automat. Control* PP (04 2017), 4229–4235. https://doi.org/10.1109/TAC.2017.2680602

[71] David A. Patterson, Garth Gibson, and Randy H. Katz. 1988. A Case for Redundant Arrays of Inexpensive Disks (RAID). In *Proceedings of the 1988 ACM SIGMOD International Conference on Management of Data* (Chicago, Illinois, USA) *(SIGMOD '88)*. Association for Computing Machinery, New York, NY, USA, 109–116. https://doi.org/10.1145/50202.50214

[72] P. Peter. 2021. Motorola Demos Wireless Charger that Works at 3m, Handles Four Phones Simultaneously, GSMARENA, September 8, 2021, https://www.gsmarena.com/motorola_demonstrates _a_wireless_charger_that _works_at_up_to_3m_can_handle_four_phones_simultaneousl-news-50850.php.

[73] Siby Jose Plathottam and P. Ranganathan. 2018. Next Generation Distributed and Networked Autonomous Vehicles: Review. *2018 10th International Conference on Communication Systems & Networks (COMSNETS)* (2018), 577–582.

[74] James Preiss, Wolfgang Hoenig, Nora Ayanian, and Gaurav Sukhatme. 2017. Downwash-Aware Trajectory Planning for Large Quadcopter Teams. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (04 2017), 8.

[75] James Preiss, Wolfgang Honig, Gaurav Sukhatme, and Nora Ayanian. 2017. Crazyswarm: A Large Nano-Quadcopter Swarm. In *IEEE International Conference on Robotics and Automation (ICRA)*. 3299–3304. https://doi.org/10.1109/ICRA.2017.7989376

[76] Nikunj Raghuvanshi and John Snyder. 2014. Parametric Wave Field Coding for Precomputed Sound Propagation. *ACM Trans. Graph.* 33, 4, Article 38 (July 2014), 11 pages. https://doi.org/10.1145/2601097.2601184

[77] Nikunj Raghuvanshi and John Snyder. 2018. Parametric Directional Coding for Precomputed Sound Propagation. *ACM Trans. Graph.* 37, 4, Article 108 (July 2018), 14 pages. https://doi.org/10.1145/3197517.3201339

[78] Richard Raspet, André L'Espérance, and Gilles A. Daigle. 1995. The Effect of Realistic Ground Impedance on the Accuracy of Ray Tracing. *The Journal of the Acoustical Society of America* 97, 1 (1995), 154–158. https://doi.org/10.1121/1.412333 arXiv:https://doi.org/10.1121/1.412333

[79] Sylvia Ratnasamy, Paul Francis, Mark Handley, Richard Karp, and Scott Shenker. 2001. A Scalable Content-Addressable Network. *SIGCOMM Comput. Commun. Rev.* 31, 4 (Aug. 2001), 161–172. https://doi.org/10.1145/964723.383072

[80] Robin Ritz, Mark W. Müller, Markus Hehn, and Raffaello D'Andrea. 2012. Cooperative Quadrocopter Ball Throwing and Catching. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 4972–4978. https://doi.org/10.1109/IROS.2012.6385963

[81] A. Rungta, C. Schissler, N. Rewkowski, R. Mehra, and D. Manocha. 2018. Diffraction Kernels for Interactive Sound Propagation in Dynamic Environments. *IEEE Transactions on Visualization & Computer Graphics* 24, 04 (apr 2018), 1613–1622. https://doi.org/10.1109/TVCG.2018.2794098

[82] Carl Schissler, Ravish Mehra, and Dinesh Manocha. 2014. High-Order Diffraction and Diffuse Reflections for Interactive Sound Propagation in Large Environments. *ACM Trans. Graph.* 33, 4, Article 39 (July 2014), 12 pages. https://doi.org/10.1145/2601097.2601216

[83] Maximilian Schulz, Federico Augugliaro, Robin Ritz, and Raffaello D'Andrea. 2015. High-speed, Steady Flight with a Quadrocopter in a Confined Environment using a Tether. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 1279–1284. https://doi.org/10.1109/IROS.2015.7353533

[84] Hazim Shakhatreh, A. Sawalmeh, Ala Al-Fuqaha, Zuochao Dou, Eyad K. Almaita, Issa M. Khalil, Noor Shamsiah Othman, A. Khreishah, and M. Guizani. 2019. Unmanned Aerial Vehicles (UAVs): A Survey on Civil Applications and Key Research Challenges. *IEEE Access* 7 (2019), 48572–48634.

[85] Rutuja Shivgan and Ziqian Dong. 2020. Energy-Efficient Drone Coverage Path Planning using Genetic Algorithm. In *IEEE 21st International Conference on High*

Performance Switching and Routing (HPSR). IEEE, 1–6. https://doi.org/10.1109/HPSR48589.2020.9098989
[86] Alexander Shpunt, Gerard Medioni, Daniel Cohen, Erez Szli, and Ronen Deitch. Apple Inc., U.S. Patent 9,582,889 B2, 2017. Depth Mapping Based on Pattern Matching and Stereoscopic Information.
[87] Martin Stephan, Javier Alonso-Mora, and Roland Siegwart. 2013. Reciprocal Collision Avoidance With Motion Continuity Constraints. *IEEE Transactions on Robotics* 29 (08 2013), 899–912. https://doi.org/10.1109/TRO.2013.2258733
[88] Ion Stoica, Robert Morris, David Karger, M. Frans Kaashoek, and Hari Balakrishnan. 2001. Chord: A Scalable Peer-to-Peer Lookup Service for Internet Applications. *SIGCOMM Comput. Commun. Rev.* 31, 4 (Aug. 2001), 149–160. https://doi.org/10.1145/964723.383071
[89] Hang Sun, Juntong Qi, Chong Wu, and Mingming Wang. 2020. Path Planning for Dense Drone Formation Based on Modified Artificial Potential Fields. *39th Chinese Control Conference (CCC)* (2020), 4658–4664.
[90] Jiayi Sun, Jun Tang, and Songyang Lao. 2017. Collision Avoidance for Cooperative UAVs With Optimized Artificial Potential Field Algorithm. *IEEE Access* PP (08 2017), 18382–18390. https://doi.org/10.1109/ACCESS.2017.2746752
[91] Ivan E. Sutherland. 1968. A Head-Mounted Three Dimensional Display. In *AFIPS '68 (Fall, part I)*.
[92] Camilla Tabasso, Venanzio Cichella, Syed Mehdi, Thiago Marinho, and Naira Hovakimyan. 2021. Time Coordination and Collision Avoidance Using Leader-Follower Strategies in Multi-Vehicle Missions. *Robotics* 10 (02 2021), 34. https://doi.org/10.3390/robotics10010034
[93] Zhenyu Tang, Hsien-Yu Meng, and Dinesh Manocha. 2021. Learning Acoustic Scattering Fields for Highly Dynamic Interactive Sound Propagation. In *IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE, 826–836.
[94] Nicolas Tsingos, Thomas Funkhouser, Addy Ngan, and Ingrid Carlbom. 2001. Modeling Acoustics in Virtual Environments Using the Uniform Theory of Diffraction. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '01)*. Association for Computing Machinery, New York, NY, USA, 545–552. https://doi.org/10.1145/383259.383323
[95] Nicolas Tsingos, Emmanuel Gallo, and George Drettakis. 2004. Perceptual Audio Rendering of Complex Virtual Environments. *ACM Trans. Graph.* 23, 3 (Aug. 2004), 249–258. https://doi.org/10.1145/1015706.1015710
[96] Nobuyuki Umetani, Athina Panotopoulou, Ryan Schmidt, and Emily Whiting. 2016. Printone: Interactive Resonance Simulation for Free-Form Print-Wind Instrument Design. *ACM Trans. Graph.* 35, 6, Article 184 (Nov. 2016), 14 pages. https://doi.org/10.1145/2980179.2980250
[97] Jur van den Berg, Stephen Guy, Ming Lin, and Dinesh Manocha. 2011. *Reciprocal n-Body Collision Avoidance*. Vol. 70. 3–19. https://doi.org/10.1007/978-3-642-19457-3_1
[98] Jur van den Berg, Ming Lin, and Dinesh Manocha. 2008. Reciprocal Velocity Obstacles for Real-Time Multi-agent Navigation. *ICRA*, 1928–1935. https://doi.org/10.1109/ROBOT.2008.4543489
[99] Jur van den Berg, Jamie Snape, Stephen Guy, and Dinesh Manocha. 2011. Reciprocal Collision Avoidance with Acceleration-Velocity Obstacles. 3475–3482. https://doi.org/10.1109/ICRA.2011.5980408
[100] Jur van den Berg, David Wilkie, Stephen Guy, Marc Niethammer, and Dinesh Manocha. 2012. LQG-obstacles: Feedback Control with Collision Avoidance for Mobile Robots with Motion and Sensing Uncertainty. *Proceedings - IEEE International Conference on Robotics and Automation* (05 2012), 346–353. https://doi.org/10.1109/ICRA.2012.6224648
[101] Marinus T. Vlaardinger and Jacques A. Broek. 1999. *Magnetic Resonance Imaging Theory and Practice* (2 ed.). Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/978-3-662-03800-0
[102] M. Wand and W. Straßer. 2004. Multi-Resolution Sound Rendering. In *Proceedings of the First Eurographics Conference on Point-Based Graphics* (Switzerland) *(SPBG'04)*. Eurographics Association, Goslar, DEU, 3–11.
[103] Fokko P Wieringa, Henri Boum, Pieter T Eendebak, Jean-Paul A van Basten, Harrie P Beerlage, Geert A. Smits, and Jelte E Bos. 2004. Improved Depth Perception with Three-dimensional Auxiliary Display and Computer Generated Three-dimensional Panoramic Overviews in Robot-assisted Laparoscopy. *Journal of Medical Imaging (Bellingham, Wash.)* 1, 1 (April 2004).
[104] Stephen Williams. 2018. For Electric Cars Without a Plug, Thank Tesla (the Scientist), The New York Times, May 31, 2018, https://www.nytimes.com/2018/05/31/business/electric-cars-wireless-charging.html.
[105] Yu Wu, Kin Huat Low, Bizhao Pang, and Qingyu Tan. 2021. Swarm-Based 4D Path Planning For Drone Operations in Urban Environments. *IEEE Transactions on Vehicular Technology* 70, 8 (2021), 7464–7479. https://doi.org/10.1109/TVT.2021.3093318
[106] Zhirong Wu, Shuran Song, Aditya Khosla, Linguang Zhang, Xiaoou Tang, and Jianxiong Xiao. 2015. 3D ShapeNets: A Deep Representation for Volumetric Shape Modeling. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, Boston, USA, 1912–1920.
[107] Siqi Xu and Shiguang Liu. 2020. Physics-Guided Sound Synthesis for Rotating Blades. In *Advances in Computer Graphics*. 233–244. https://doi.org/10.1007/978-3-030-61864-3_20
[108] Yang Xu, Shupeng Lai, Jiaxin Li, Delin Luo, and Yancheng You. 2019. Concurrent Optimal Trajectory Planning for Indoor Quadrotor Formation Switching. *Journal of Intelligent & Robotic Systems* 94 (05 2019). https://doi.org/10.1007/s10846-018-0813-9
[109] Xinchen Yan, Jimei Yang, Ersin Yumer, Yijie Guo, and Honglak Lee. 2016. Perspective Transformer Nets: Learning Single-View 3D Object Reconstruction without 3D Supervision. In *Proceedings of the 30th International Conference on Neural Information Processing Systems* (Barcelona, Spain) *(NIPS'16)*. Curran Associates Inc., Red Hook, NY, USA, 1704–1712.
[110] Hengchin Yeh, Ravish Mehra, Zhimin Ren, Lakulish Antani, Dinesh Manocha, and Ming Lin. 2013. Wave-Ray Coupling for Interactive Sound Propagation in Large Complex Scenes. *ACM Trans. Graph.* 32, 6, Article 165 (Nov. 2013), 11 pages. https://doi.org/10.1145/2508363.2508420
[111] Quanzeng You and Hao Jiang. 2020. Real-time 3D Deep Multi-Camera Tracking. *CoRR* abs/2003.11753 (2020), 17 pages. arXiv:2003.11753 https://arxiv.org/abs/2003.11753
[112] Aviad Zabatani, Vitaly Surazhsky, Erez Sperling, Sagi Ben Moshe, Ohad Menashe, David H. Silver, Zachi Karni, Alexander M. Bronstein, Michael M. Bronstein, and Ron Kimmel. 2020. Intel® RealSense™ SR300 Coded Light Depth Camera. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 42, 10 (2020), 2333–2345. https://doi.org/10.1109/TPAMI.2019.2915841
[113] Faheem Zafari, Athanasios Gkelias, and Kin K. Leung. 2019. A Survey of Indoor Localization Systems and Technologies. *IEEE Communications Surveys and Tutorials* 21, 3 (2019), 2568–2599. https://doi.org/10.1109/COMST.2019.2911558
[114] Echo Zhan. 2021. 3,281 Drones Break Dazzling Record for Most Airborne Simultaneously, May 17, 2021. See https://www.guinnessworldrecords.com/news/commercial/2021/5/3281-drones-break-dazzling-record-for-most-airborne-simultaneously-655062.
[115] Yipin Zhou, Zhaowen Wang, Chen Fang, Trung Bui, and Tamara Berg. 2017. Visual to Sound: Generating Natural Sound for Videos in the Wild. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 3550–3558.