Sensors for Localization and Mapping

— or 30 dubious ways to find your robot without GPS

Simen Haugo, July 2018

CONTENTS

1 Basic Concepts **2**
1.1 Triangulation and trilateration 2

2 Sensors for localization and mapping **3**
2.1 Camera 3
2.2 Spectral Camera 4
2.3 Depth Camera 5
2.4 Event Camera 7
2.5 Light-field Camera 8
2.6 Laser 9
2.7 Radio 12
2.8 Sound 13

3 What sensor do I buy? **15**

INTRODUCTION

Finding our place in the universe was a central topic of discourse among our ancestors who gazed up at the heavens, and remains today on the minds of those who, from time to time, can spare a thought to that great mystery. Yet, we can't know our place without also knowing what our place looks like, as, to paraphrase a certain 17th century astronomer, *to be somewhere, you need somewhere to be.*

Think of it this way: to fly an airplane, you need to atleast know where you are relative to the airport. But really you need more details, since a world containing only your airplane and your destination, in the empty cartesian void of a MATLAB figure, would not do you well; for instance, the optimal path between the two would intersect with our planet, among other things of more or less importance. To plan your path, you need to know about the Earth and its mountains and valleys, cities and their skyscrapers, air traffic, whether the runway is covered with ice, the wind speed, and if there's a storm up ahead. Figuring out where you are (localization) is therefore often coupled with figuring out what's around you (mapping).

You can find many references on methods to tackle these problems, but in this short survey I wanted to address the practical problem of what sensor to buy. It's not a review of sensors, like phone reviews in a magazine, since that would be outdated by the time you read it. Instead it's a comparison of the trade-offs that are likely persist through the cycles of innovation—both in hardware and in software.

Along the way, you will find notes, such as this, that point to an example usage of a particular sensor. You can find the cited reference at the end of the article.

1.1 Triangulation and trilateration

The position of a thing can be split into a direction and a distance away from a reference point. If you are fortunate, both of these quantities are known. More often though you find yourself in the possession of one and lacking the other. Seafarers who navigated through night and along the shores used the following principles to guide them in those situations.

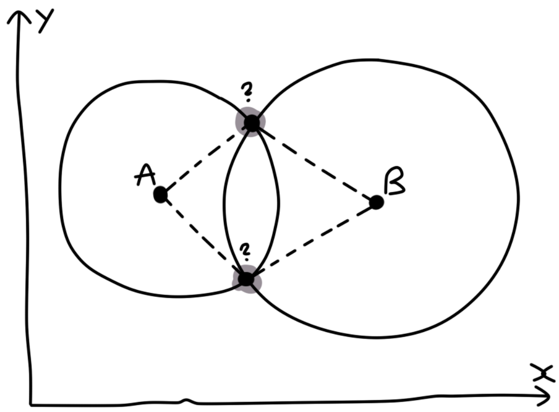


Figure 1: Trilateration in two dimensions: One distance measurement gives a circle of possible positions, two leaves two possible points, and a third is enough to select one or the other.

If you know the distance to a thing, you can be anywhere on a three-dimensional sphere around it. Knowing the distance to a second thing reduces that to a circle. With atleast four distances you can compute your position unambiguously. This is called trilateration.

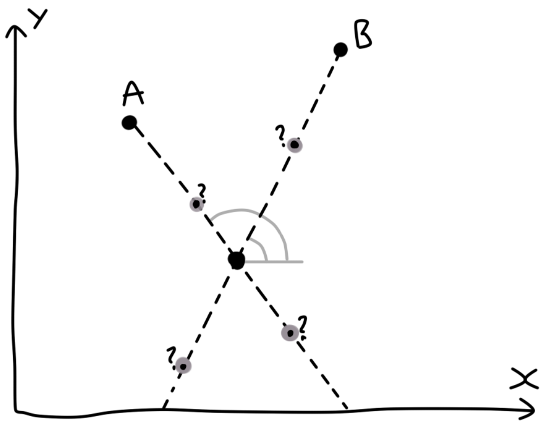


Figure 2: Triangulation in two dimensions

If you know the direction to a thing, you can be anywhere on a line. If you also know the direction to a second thing, you know you have to be at the intersection of the lines. This is called triangulation.

Although these are simple principles, the hard part is to actually measure the direction and distance to things. Over time, people have invented a myriad of strange devices exploiting physical phenomenon that, hopefully, has something to do with either direction or distance. Unfortunately, as we will see, no size fits all.

2 SENSORS FOR LOCALIZATION AND MAPPING

2.1 Camera

Cameras, like the one on your phone, can measure the direction to things: each pixel captures light from a particular direction, hence, each pixel defines a direction into the world. In that sense, cameras are direction sensors. But part of the trick is not just measuring the direction to something, but measuring the direction to a particular, reidentifiable thing. All you get from a camera is an array of colored pixels, and measuring the direction to the color red is not very helpful (unless you know there aren't other red things in the scene).

Figure 3: Artificial markers in the form of QR code patterns stuck to the walls. These patterns are easily detected with image processing libraries and have a relatively unique appearance.

An easy fix is to place *markers* in the environment, special objects crafted to be easily identifiable and detected with image processing. A marker can be all sorts of things: a QR code pattern (Fig. 3), a glowing LED [9] or a 3D-printed bear [20]; anything goes, as long as it stands out. You can also turn it around, placing markers on the robot and cameras in the environment.

[9] — a quadcopter fitted with a constellation of very bright LEDs, observed and triangulated from the camera of another quadcopter, across a canyon.
[20] — 3D-printed objects as visual landmarks.

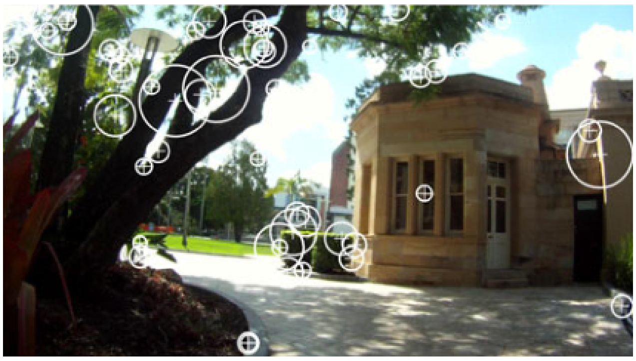


Figure 4: Natural markers in the form of SIFT descriptors. One of these is a block of pixels that has been processed and turned into a list of numbers that is surprisingly unique and consistent across different viewpoints and lighting changes.

If you can't place anything, cameras or markers, in the environment, you can instead rely on what's already there, like you would use natural landmarks on a hiking trip. Researchers have used many kinds of landmarks: buildings, like the Eiffel tower [39, 31]; aerial terrain maps [22]; chairs and tables [42]; planes [41] and lines [43]; small 16x16 pixel image regions [36], or even 3x3 pixel regions [15].

[31] — recognizing where you are by comparing what you see with geo-referenced photographs...
[22] — or aerial maps.
[42] — using existing visual landmarks in the form of common household objects
[43] — or lines
[41] — or planes
[36] — or aggregations of pixels.

2.2 Spectral Camera

Spectral cameras are designed to perceive light of particularly selected wavelengths inside or outside the visible spectrum. Ordinary cameras are made to perceive light of particular wavelengths as well, but those are chosen so that the resulting image matches what a human would perceive, i.e. the visible spectrum, which might not be optimal for the localization problem. Ordinary cameras are sensitive to wavelengths outside the visible spectrum as well, but this is usually an undesired effect so they come with filters to block those wavelengths. For some scientific endeavours those wavelengths are exactly what we want.

Figure 5: Markers for infrared cameras. (Left) IR LEDs attached to a headset. (Middle) Retroreflective markers, intended to be lit by an infrared light. (Right) Dots of IR ink that absorb light of a particular wavelength and appear as black when lit by an infrared light. Source: iFixit, Creaform 3D and Narita et al. [37].

Like ordinary cameras, spectral cameras can be used with artificial markers, except that these markers can be made to absorb, reflect or emit light of the particular wavelengths susceptible by the camera. In other words, the markers can be identified by "color" outside the visible spectrum, such as infrared [32], which makes it easier to create markers that stand out against other stuff in the environment.

[37] — infrared ink dots painted on a sheet of paper and tracked from an infrared camera.
[32] — a quadcopter fitted with retroreflective markers, illuminated by an infrared light, and tracked by an infrared camera.

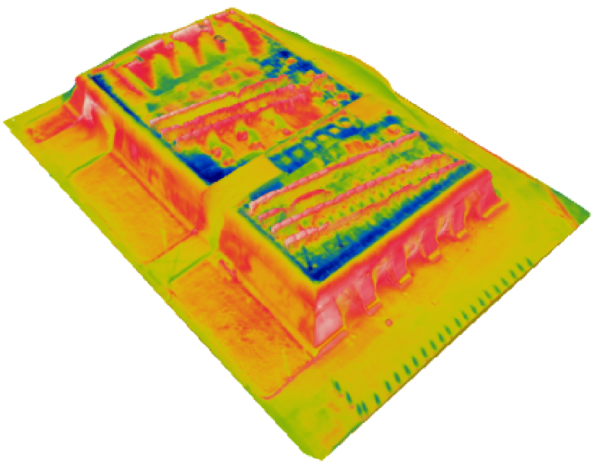


Figure 6: A 3D model of a building reconstructed solely from thermal infrared photographs using unmodified photogrammetry software, intended for color cameras. Source: [33]

Spectral cameras can also use the natural appearance of the environment, but, unlike ordinary cameras, they can use variations in "color" outside the visible spectrum as useful landmarks, or exploit the lack of variation of particular wavelengths: for example, using thermal infrared to see in the dark. But of course, like ordinary cameras, if there is not enough light, they too are blind. Some setups add a light source, emitting at the wavelengths of interest, to help.

[33] — using thermal infrared pixels as "visual" landmarks.

2.3 Depth Camera

The previous two cameras can measure direction. Depth cameras can also (or primarily) measure distance. Measuring distance solves some issues that ordinary cameras run into, like scale ambiguity (you don't know if a photo was taken in a doll house or a real house) or operating in flat-colored scenes (where it's hard to find good landmarks).

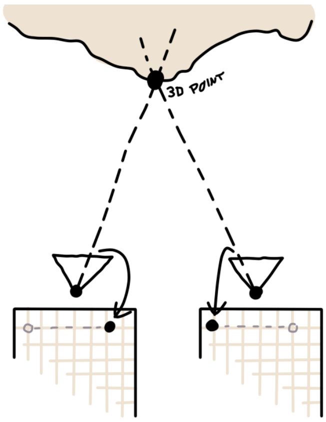


Figure 7: Stereo cameras compute distances by triangulation from two photographs, taken at a fixed distance apart.

One class of depth cameras are actually just two ordinary cameras placed (very carefully) next to each other (Fig. 7). These don't measure distance directly but triangulate it from two angles, like our eyes do, from the fact that a point in the scene appears at different coordinates in each camera based on the seperation between the cameras and the point's depth. If one of these is known, the other can be solved for.

[12] — a quadcopter carrying a stereo camera.

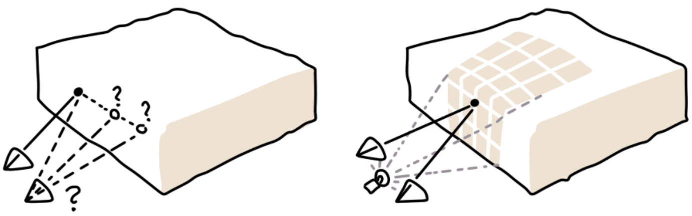


Figure 8: (Left) On poorly textured surfaces it becomes impossible to identify where a pixel in the left camera is observed in the right camera, and thus impossible to determine distance. (Right) Artificially introducing texture, e.g. with a light pattern projector, can help.

These devices don't work well if the scene lacks texture that makes it possible to identify which pixels in the two images correspond to the same point (Fig. 8). Some devices mitigate this by projecting a light pattern onto the scene to add texture [50].

[44] — a structured-light depth camera fixed to the ceiling and tracking a quadcopter from above.

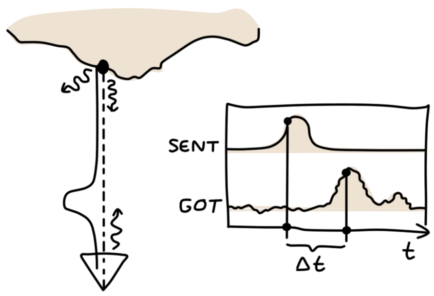


Figure 9: Time-of-flight depth cameras work by emitting signals of light and measuring how long the signal takes to return.

Other devices also emit light, but use it to calculate distance by measuring the time it takes to return. These are similar to lidar sensors (see Sec. 2.6), but make a trade-off to have shorter range, especially in sunlight, but higher resolution and lower price [17].

[25] — a quadcopter carrying a time-of-flight depth camera, tracking its position and mapping the environment.

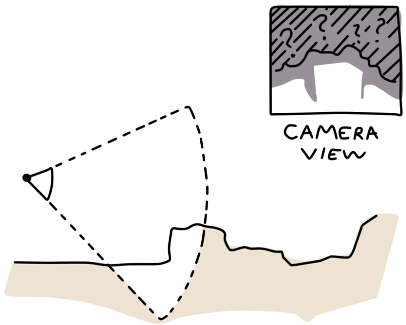


Figure 10: The high resolution of depth cameras lets you perceive thin structures, but the short range can be problematic for localization if there aren't many landmarks nearby.

The range, especially outdoors, is limited by the emission power permitted by law and ambient lighting like the sun. The short range is problematic if an algorithm relies on a consistently large coverage of distance measurements (Fig. 10).

[52] — a quadcopter carrying a short-range depth camera and a long-range lidar.

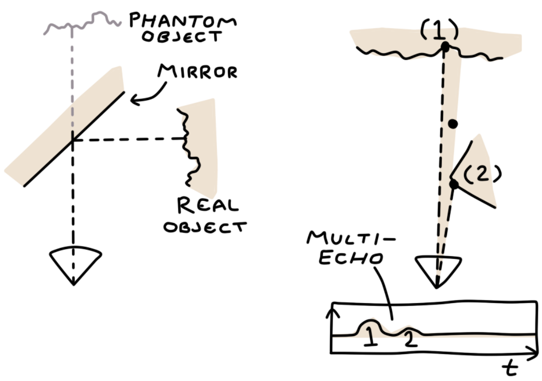


Figure 11: (Left) Specular reflections can cause objects to appear behind walls. (Right) Light can bounce from both background (1) and foreground (2) to the same detector, which has to seperate them. Some devices are unable to distinguish these and guess the distance is somewhere in-between.

Also, specular or polished surfaces can reflect the light, so if any light does make it back to the sensor, it is unlikely to produce a correct distance. Smoke and dust may likewise scatter the light elsewhere, and confuse the sensor. Another issue is that the light detectors do not receive light from an infinitesimal line, but from a cone of non-zero width, so that at an edge multiple echoes of light may return to the same detector (Fig. 11). Despite these drawbacks, active depth cameras can be useful indoors, where surfaces are often poorly textured and distances are suitably short.

[4] — a quadcopter carrying a depth camera and localizing itself in a 3D model of the environment, scanned beforehand.

2.4 Event Camera

Event cameras have only been commercially available since 2008. Unlike ordinary cameras, they do not output images, but a stream of individually timestamped pixel events: if the brightness seen by any one pixel changes significantly from the last stored value (e.g. by 10-15%), you get an event saying if it got brighter or darker along with (very precisely) when it occurred (Fig. 12).

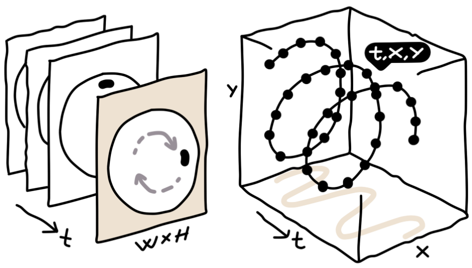


Figure 12: Imagine a spinning disk with a dot. An ordinary camera (left) would output frames of images, capturing the disk in its entirety at a fixed rate. An event camera (right) produces an output only for the pixels that change.

This quirky mode of operation gives it a number of advantages: a microsecond timestamp resolution, low power consumption and an ability to discern brightness changes on sunlit or shaded surfaces alike, has attracted the attention of those fed up with the poor capabilities of ordinary cameras, during fast motion or in high dynamic range scenes. Indeed, there have been attempts at extending the localization algorithms used for ordinary cameras to event cameras [40, 19], but the asynchronous nature of its output makes it difficult to apply algorithms that have so far formed the backbone of computer vision.

[40] — a quadcopter carrying an event camera.

One way to take advantage of its unique properties is to discriminate markers in time frequency, rather than spatial- (texture) or spectral frequency (color): e.g. by attaching a plate of blinking LEDs to a robot, or in the environment [35, 8]. The camera can also be made into an active depth sensor, by placing an IR filter in front of the camera and scanning a laser across the field of view [10]. The high time resolution makes it possible to get scans of useful density even at 250 Hz, and the dynamic range makes it work in scenes containing both very bright and very dark things.

[35] — a quadcopter carrying two event cameras.
[8] — a quadcopter with blinking LEDs, tracked from a fixed event camera.

2.5 Light-field Camera

A photograph captures light as seen from a single point in the world, but this is a meager description of *all* the light that is present, which passes through many different points. Such a richer description of light is usually referred to as a *light-field* and a simple light-field, or *plenoptic*, camera can be built by placing many cameras next to each other and synchronizing them to take photographs at the same time. This is similar to a passive stereo camera, but instead of two cameras and only one baseline, there are many cameras and multiple baselines in multiple directions. They can therefore be used in similar ways for localization and mapping [49, 11] and also share some drawbacks, such as not being able to see in darkness or estimate depth from poorly textured surfaces—but plenoptic cameras can offer some advantages.

[13] — a track-wheel robot carrying a plenoptic camera, mapping the environment and tracking its position.

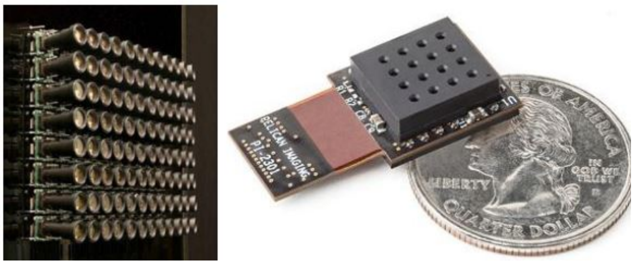


Figure 13: (Left) A light-field camera built by stacking many ordinary cameras in a grid. (Right) A coin-size 4x4 light-field camera. Source: [49].

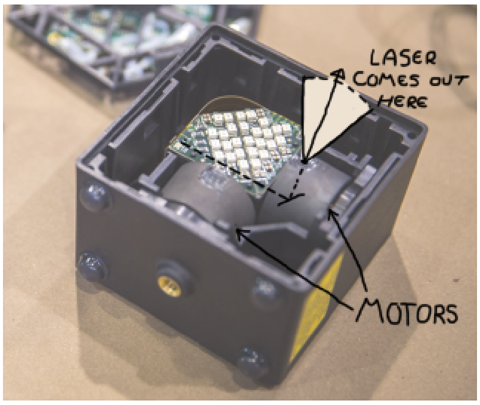
For example, light-field cameras in use (at the time of writing) are much smaller, even fitting inside a phone. These are made with microscopic arrays of lenses, that interlace images from different viewpoints onto a single sensor: trading off the number of views (angular resolution) with the number of pixels available for each view (spatial resolution). Other cameras, based on interlacing in frequency, have also been tried, but are less popular [49] than micro lens array cameras, which have even made it to smartphones. For example, Google's Pixel 2 [3] has micro lenses that subdivides the sensor into two: each pixel has two photodiodes that see slightly shifted viewpoints. Despite the short baseline of less than 1 millimeter, it is possible to obtain useful depth estimates up-close, although the method used is different from traditional methods designed for longer baselines [3, 6].

Figure 14: Two photos taken at a short vertical baseline: the edge of the white surface at the top would be ambiguous in a horizontal baseline, but can be more easily triangulated vertically.

With a smaller form factor, micro lens array cameras can be advantageous over wider baseline stereo cameras, and the added number of views can improve robustness on self-similar textures (Fig. 14). However, the usefulness may be limited to close ranges or for rough depth estimates because hey, short baseline [51].

2.6 Laser

Lasers are narrow beams of light that can be directed in precise angles, for example by rotating a mirror that it shines at with a precise motor (like those in spinning hard disks) or with small electronically controlled (MEMS) mirrors.



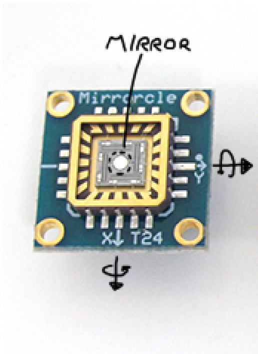
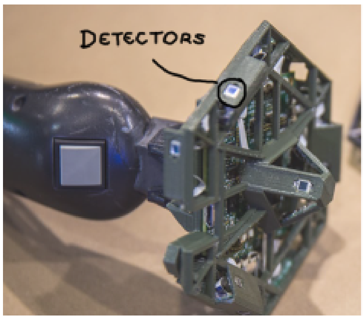


Figure 15: Two ways to steer lasers. (Left) A "Lighthouse base station" housing two IR laser emitters that shine into spinning drums which reflect and spread the lasers into vertical and horizontal sheets [1]. (Right) A MEMS mirror that can be rotated in two axes [2].

Being able to direct light very precisely means you can measure angles very precisely, which is useful for triangulation. For example, you can put a marker on a robot and triangulate its position by tracking it with lasers in the environment. If you put a bunch of markers in a fixed constellation you can recover its orientation as well [28].

[1] — photodetectors on hand-held virtual reality controllers, triangulated from spinning lasers.
[26] — a quadcopter carrying a large retroreflective ball, tracked by a steerable laser



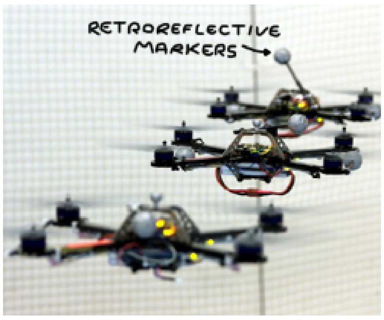


Figure 16: Types of markers. (Left) The HTC Vive hand controller with photodetectors distributed on its surface, each wired to a central hub [1]. (Right) Retroreflective balls attached to quadcopters [32].

What are these markers then? Some are made of a retroreflective material that reflects all light back the direction it came from, such that a photoreceptor placed near the laser can detect the returning signal. The markers could also be photoreceptors themselves responsible for detecting when they are hit by a laser [26].

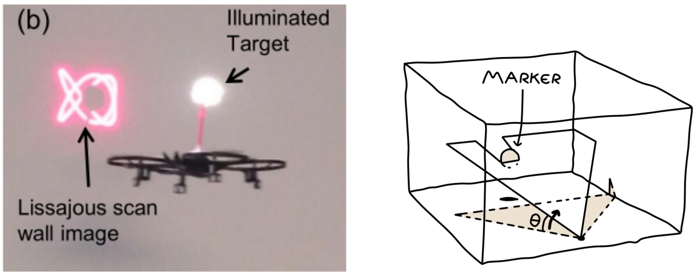


Figure 17: (Left) A laser point oscillating in a Lissajous curve, generated by a sinusoidal voltage on the axes of a MEMS mirror, tracking a retroreflective ball attached to a quadcopter [26]. (Right) A laser beam split into a horizontal line being swept vertically to measure a marker's vertical angle.

To steer a laser towards a marker, you can oscillate the beam in a pattern and adjust its center direction based on where in the pattern the marker is detected [26]. Alternatively, the laser can be spread into a line that is swept over the entire space: to obtain a 3D direction, you only need to sweep in two directions, one for each angle [28].

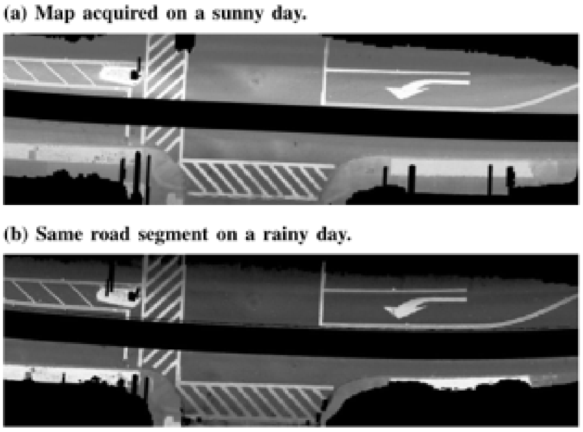


Figure 18: The return signal strength of a laser is a measure of the surface reflectivity for the wavelength of the laser. Here is shown the resulting image when a laser is swept across a road, where the road markings have higher reflectivity than the asphalt. Source: [30].

If placing markers in the environment or on the robot is not an option, the other option is to measure what's already there. For example, emitting laser pulses and measuring the strength of the returning light gives a measure of the surface reflectivity for that particular wavelength (and angle of incidence)—kinda like a spectral camera. Lasers can be particularly useful because you can choose the wavelength to capture "images" (Fig. 18) that are more resilient to common variations in visible light images, like shadows, time of day, sunlit or overcast [30].

[30] — laser used to capture the appearance of street asphalt for a very specific wavelength.

The second, and arguably more popular use, is to measure distance to things. When laser is used for this, it is called *lidar*, or *light detection and ranging*. Some lidars spin a beam in a circle very quickly, emit a pulse of light at regular intervals and measure how long it takes to return; giving what resembles a row of pixels in a depth camera, but with much wider field of view in one direction (e.g. 360 degrees) and more narrow field of view in the other (e.g. 0.1 degrees).

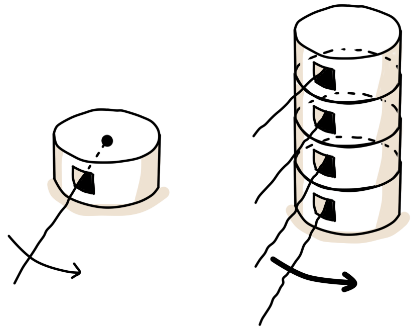


Figure 19: (Left) A one-axis scanning lidar. (Right) A two-axis scanning lidar, where the second axis is made by stacking layers of emitter/receivers.

[53] — a quadcopter carrying a two-axis spinning lidar.

To increase the resolution, some devices stack layers of emitter-detector pairs [48], and spin the stack around in circle. Other setups, especially the DIY kind, spin a single pair around in a sphere with a second motor [46]. A third kind, called "solid state" or "scannerless" lidar spread the light beam to cover a larger portion of the scene and observe the returning light from multiple detectors [24], instead of using just one laser per detector. The difference between this third kind and depth cameras can get blurry, but in the market, lidar are usually rated as having low resolution but long range and wide viewing angles, even in sunlight; while depth cameras have high resolution, but short range, *especially* in sunlight [24].

Lidar does face some challenges though: distinguishing multiple echoes from a single pulse; specular reflections from polished surfaces and grazing angles; attenuation or scattering in fog and smoke; and the long range could make lidar more prone to interfere with other laser devices [34].

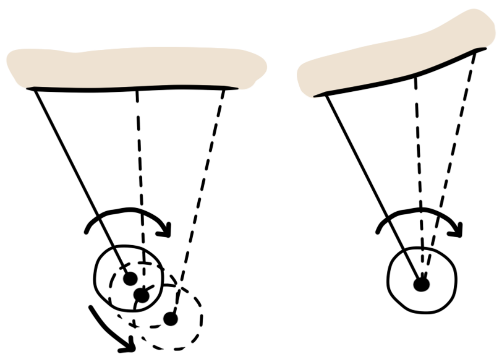


Figure 20: (Left) As the lidar moves away from the object during a scan, the measured distance will increase. (Right) If the lidar is assumed to be stationary during the scan, this increasing distance will result in a perceived deformation of the object: e.g. a straight wall is perceived to be slanted.

Another issue can be caused by motion. Whereas depth cameras can measure depth at each pixel simultanously, for lidar, even during a single planar scan, not to mention two-axis scanning, there will be a time difference between the first and the last range measurement. If either sensor or object moves during this time, the computed points will be deformed. There are ways to mitigate this effect, but their success depends on the environment and type of motion [53].

2.7 Radio

Along with light, radio waves are also part of the electromagnetic spectrum, but its interaction with matter is qualitatively differently. One difference from light is that radio waves can more easily penetrate materials and participating media. Since the inability to see through fog and smoke is a crux of lidar, replacing it with radio waves is an appealing idea. Indeed, *radar* is the lidar of radio waves: emit a radio wave, and measure how long it takes to bounce back.

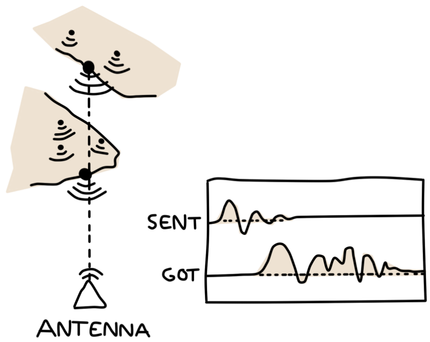


Figure 21: As a radio wave interacts with matter, some of it may be reflected back, and some of it will continue. Compared with a laser, the returned signal contains many more echoes, and requires more extensive filtering to seperate into distinct surface reflections.

Unfortunately (but also fortunately, as we will remark upon soon), we can't make narrow "beams" of radio waves as easily as we can with lasers, which reduces the angular precision with which you can send a signal or determine where it came from. Also, the ability to penetrate, not only fog, but walls and other objects, means that multiple echoes become more prominent (which may or may not be desired).

Angular precision can be improved by placing an array of radio elements next to each other and comparing the slightly different arrival times, or phase differences, of a received signal. A similar array can also emit waves in a narrower direction, by carefully adding delays to create constructive interference in the desired direction, and destructive interference elsewhere. But, to reach similar angular precision as lidar, these arrays have to be big [34], which makes radio less suited as a full replacement for lidar, unless the environment has easily distinguishable reflecting features [45],

[45] — a quadcopter carrying a radar.

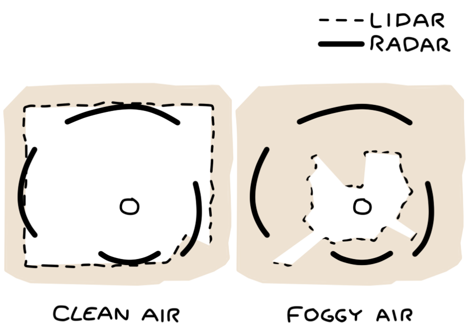


Figure 22: Radio waves' ability to pass through stuff can complement lidar in the presence of fog or smoke, which attenuate or scatter light elsewhere. On the other hand, poor angular precision makes it less capable as a full replacement for mapping.

On the other hand, poor angular precision is also one of its strong points: sound is omnidirectional, but propagates slowly and is easily attenuated; light is fast, but is likewise hard to detect when spread out. Radio waves are fast, preserve a high signal/noise ratio over long distances and can also diffract around occlusions. While these properties makes radio a good complement for lidar in the presence of thick visual obscurants, like fog [18], they also make radio very interesting for trilateration. So interesting, in fact, that we use it for GPS.

[18] — a radar compensating the reduced visibility of lidar in fog and smoke.

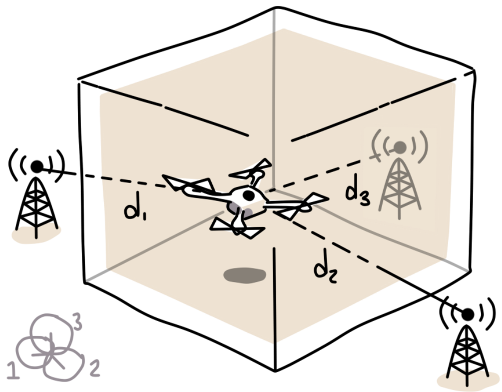


Figure 23: Radio can be used to measure distance by broadcasting a signal. If the receiver knows when the signal was sent, for example by having the sender encode a timestamp in the signal, they can calculate how long it took to get there, and thus how far it travelled. With enough measurements, a robot can compute its own position using the trilateration principle.

The wide reach of radio means that sender and receiver do not need a narrow line of sight to communicate: e.g. to measure distance [5, 21]. By placing radio nodes in the environment and measuring distances, a robot can work out where it is [29], and possibly where the nodes are [38], without worrying too badly about where it is pointing or if there's stuff in the way.

[29] — locating a robot carrying a radio receiver from pulses emitted by transmitters in the environment.

Radio-based localization approaches have traditionally been deemed as expensive, inaccurate and/or noisy, but changes in regulations and the internet of things is causing a trend of smaller and cheaper sensors. This can potentially be used to build more accurate systems, or serve as a cheap and flexible alternative to other sensors [29, 38].

2.8 Sound

While our hearing ability is not as sophisticated as some other creatures, our brain can do some tricks and work out the angle to stuff that are audible to us. This complements our vision, as sound can make its way through dust, smoke or physical occluders and, unlike our eyes, our ears work in the dark. These properties makes acoustic localization an attractive contender.

Using sound for localization dates back to the discovery of piezoelectricity and its application to producing and measuring ultrasonic waves, which later became *sonar*: the acoustic analog of *lidar* and *radar*. The motivation for using ultrasonic waves is that, compared to audible waves, ultrasonic waves (20 kHz and above) can be made to spread much less, which in turn lets you more precisely determine its direction; a crucial bit of information for detecting submarines. Although sonar could in principle be used like a lidar, there are a number of challenges due to the physics of sound (see [27] for more details).

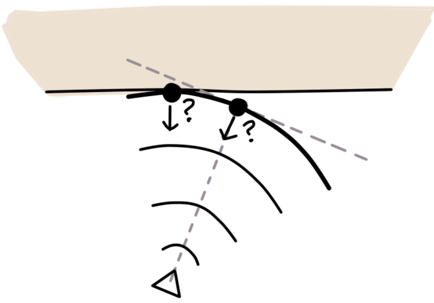


Figure 24: As a sound wave propagates through space, the first echo does not necessarily come from the direction the speaker is pointing, but from the first point of contact between an object and the arc that sound makes as it spreads.

First, it is hard to make a sound wave as narrow as a laser, which reduces how well you can resolve the direction (Fig. 24). As with radio, you can improve this by placing an array of microphones next to each other and comparing the slightly different arrival times of the received signal [27]. You can also use an array to emit sound in a narrow direction, by carefully delaying each signal to create an interference pattern. However, the trade-off between the size of such an array and its angular precision can make them quite large [34].

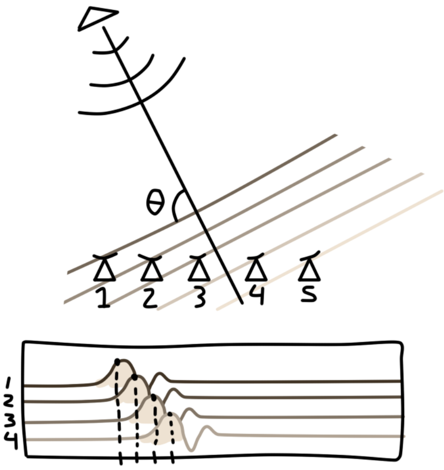


Figure 25: Phase-array microphones are packs of microphones put next to each other, maybe in a line, a circle or a volumetric grid, so that a sound wave will reach the individual elements of the array at different times depending on its direction. The name comes from the resulting time delay (phase shift) of the measured signals.

Second, the roughness for most indoor surfaces tends to be smaller than the wavelength (e.g. 6.6 millimeter at 50 kHz), which causes the sound to be reflected; to the sound wave, every wall in your house looks like a perfect mirror. This means you can only reliably measure the range at perpendicular angles to things, otherwise you might see phantom objects appearing behind walls.

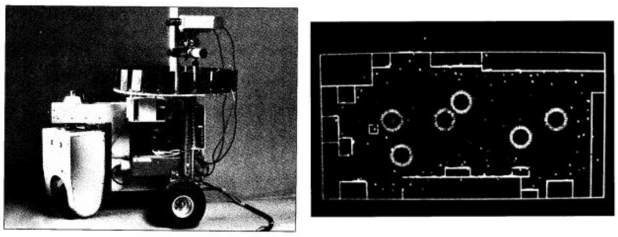


Figure 26: A spinning sonar attached to a wheeled robot (left), which was capable of creating the map seen on the right [14].

These setbacks did not stop researchers from trying and they got some pretty good results by making assumptions on the type of geometry in the environment [47]. But outside corridor-like environments, sonar is most often used as a "last-resort" to avoid collision with nearby objects. The slow speed of sound also makes it difficult to use sonar on continuously moving platforms: for example, a round-trip to a wall ten meters away takes as much as 60 milliseconds.

[47] — a wheeled robot carrying a one-axis spinning sonar, tracking walls in an office-like environment.

A more popular approach has been to perform triangulation, for example, by placing microphone arrays at known locations in the environment [23]. A target emitting sounds (natural or artificially generated) can be triangulated from the set of angles measured by each microphone. This setup can also be reversed with sound sources in the environment and a microphone array on the target [16]. However, the number of sources may be orders of magnitudes fewer than e.g. visual landmarks for camera setups, which makes the problem harder, especially in three dimensions, where the added difficulty of resolving direction will add measurement uncertainty.

[23] — a robot carrying an ultrasonic speaker, emitting pulses at regular intervals that are picked up by microphones attached to the ceiling.
[16] — a robot carrying a phase-array microphone, measuring the direction to sound sources in the environment.

3 WHAT SENSOR DO I BUY?

You might be tempted to give up trying to consider all these alternatives and just buy the latest iPhone, which has presumably got localization and mapping nailed down by now. However, off-the-shelf devices are often aimed at a wide range of use cases, and might not be reliable enough in your particular scenario. Many of the papers I read while studying this subject rarely use a single sensor; instead, they complement cameras with inertial sensors, lidar, radio beacons or ultrasonic sensors such that, where one mode of sensing fails, another can fill in the gaps. This article hopefully gives an overview of what the strengths and drawbacks are, independent of particular brands or models, so that you may have a better clue of what sensors you should search for.

The best solution comes from considering your particular problem, taking advantage of domain knowledge and assumptions you can get away with, and complementing sensors with other sensors. For example, if you have access to a reliable map of an area, it makes little sense to assume nothing a priori. Or, knowing that the environment is "corridor-like", with vertical walls from a flat floor, can greatly simplify the necessary sensors and algorithms [7].

REFERENCES

[1] Lighthouse tracking examined. `http://doc-ok.org/?p=1478`. Accessed: 2018-04-17.

[2] Mems mirrors. `http://mirrorcletech.com/devices.html`. Accessed: 2018-04-17.

[3] Portrait mode on the pixel 2 and pixel 2 xl smartphones. `https://research.googleblog.com/2017/10/portrait-mode-on-pixel-2-and-pixel-2-xl.html`. Accessed: 2018-04-16.

[4] Project tango - visual slam development. `https://www-robotics.jpl.nasa.gov/tasks/showTask.cfm?FuseAction=ShowTask&TaskID=256&tdaID=700069`. Accessed: 2018-04-16.

[5] Abdulrahman Alarifi, AbdulMalik Al-Salman, Mansour Alsaleh, Ahmad Alnafessah, Suheer Al-Hadhrami, Mai A Al-Ammar, and Hend S Al-Khalifa. Ultra wideband indoor positioning technologies: Analysis and recent advances. *Sensors*, 16(5):707, 2016.

[6] Robert Anderson, David Gallup, Jonathan T Barron, Janne Kontkanen, Noah Snavely, Carlos Hernández, Sameer Agarwal, and Steven M Seitz. Jump: virtual reality video. *ACM Transactions on Graphics (TOG)*, 35(6):198, 2016.

[7] Joydeep Biswas and Manuela M Veloso. Localization and navigation of the cobots over long-term deployments. *The International Journal of Robotics Research*, 32(14):1679–1694, 2013.

[8] Andrea Censi, Jonas Strubel, Christian Brandli, Tobi Delbruck, and Davide Scaramuzza. Low-latency localization by active led markers tracking using a dynamic vision sensor. In *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*, pages 891–898. IEEE, 2013.

[9] E Cledat and DA Cucci. Mapping gnss restricted environments with a drone tandem and indirect position control. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 4:1, 2017.

[10] Oliver S Cossairt, Nathan Matsuda, and Mohit Gupta. Motion contrast 3d scanning. In *Computational Optical Sensing and Imaging*, pages CT2E–1. Optical Society of America, 2015.

[11] Donald G Dansereau, Ian Mahon, Oscar Pizarro, and Stefan B Williams. Plenoptic flow: Closed-form visual odometry for light field cameras. In *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*, pages 4455–4462. IEEE, 2011.

[12] A Deris, I Trigonis, A Aravanis, and EK Stathopoulou. Depth cameras on uavs: A first approach. *Proceedings of the ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Nafplio, Greece*, pages 1–3, 2017.

[13] Fengchun Dong, Sio-Hoi Ieng, Xavier Savatier, Ralph Etienne-Cummings, and Ryad Benosman. Plenoptic cameras in real-time robotics. *The International Journal of Robotics Research*, 32(2):206–217, 2013.

[14] Alberto Elfes. Using occupancy grids for mobile robot perception and navigation. *Computer*, 22(6):46–57, 1989.

[15] Jakob Engel, Vladlen Koltun, and Daniel Cremers. Direct sparse odometry. *IEEE transactions on pattern analysis and machine intelligence*, 2017.

[16] Christine Evers, Alastair H Moore, and Patrick A Naylor. Localization of moving microphone arrays from moving sound sources for robot audition. In *Signal Processing Conference (EUSIPCO), 2016 24th European*, pages 1008–1012. IEEE, 2016.

[17] Péter Fankhauser, Michael Bloesch, Diego Rodriguez, Ralf Kaestner, Marco Hutter, and Roland Siegwart. Kinect v2 for mobile robot navigation: Evaluation and modeling. In *Advanced Robotics (ICAR), 2015 International Conference on*, pages 388–394. IEEE, 2015.

[18] Paul Fritsche, Simon Kueppers, Gunnar Briese, and Bernardo Wagner. Radar and lidar sensorfusion in low visibility environments. In *ICINCO (2)*, pages 30–36, 2016.

[19] Guillermo Gallego, Jon EA Lund, Elias Mueggler, Henri Rebecq, Tobi Delbruck, and Davide Scaramuzza. Event-based, 6-dof camera tracking from photometric depth maps. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017.

[20] Dorian Gálvez-López, Marta Salas, Juan D Tardós, and JMM Montiel. Real-time monocular object slam. *Robotics and Autonomous Systems*, 75:435–449, 2016.

[21] Javier González, Jose-Luis Blanco, Cipriano Galindo, A Ortiz-de Galisteo, Juan-Antonio Fernández-Madrigal, Francisco Angel Moreno, and Jorge L Martínez. Mobile robot localization based on ultra-wide-band ranging: A particle filter approach. *Robotics and autonomous systems*, 57(5):496–507, 2009.

[22] Christina Grönwall, Joakim Rydell, Michael Tulldahl, Erik Zhang, Fredrik Bissmarck, and Erika Bilock. Two imaging systems for positioning and navigation. In *Research, Education and Development of Unmanned Aerial Systems (RED-UAS), 2017 Workshop on*, pages 120–125. IEEE, 2017.

[23] Mike Hazas and Andy Hopper. Broadband ultrasonic location systems for improved indoor positioning. *IEEE Transactions on mobile Computing*, 5(5):536–547, 2006.

[24] Radu Horaud, Miles Hansard, Georgios Evangelidis, and Clément Ménier. An overview of depth cameras and range scanners based on time-of-flight technologies. *Machine Vision and Applications*, 27(7):1005–1020, 2016.

[25] Marco Karrer, Mina Kamel, Roland Siegwart, and Margarita Chli. Real-time dense surface reconstruction for aerial manipulation. In *Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on*, pages 1601–1608. IEEE, 2016.

[26] Abhishek Kasturi, Veljko Milanovic, Bryan H Atwood, and James Yang. Uav-borne lidar with mems mirror-based scanning capability. In *Laser Radar Technology and Applications XXI*, volume 9832, page 98320M. International Society for Optics and Photonics, 2016.

[27] Lindsay Kleeman and Roman Kuc. Sonar sensing. In *Springer handbook of robotics*, pages 753–782. Springer, 2016.

[28] Steven LaValle. Virtual reality. 2016.

[29] Anton Ledergerber, Michael Hamer, and Raffaello D'Andrea. A robot self-localization system using one-way ultra-wideband communication. In *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*, pages 3131–3137. IEEE, 2015.

[30] Jesse Levinson, Michael Montemerlo, and Sebastian Thrun. Map-based precision vehicle localization in urban environments. In *Robotics: Science and Systems*, volume 4, page 1. Citeseer, 2007.

[31] Stephanie Lowry, Niko Sünderhauf, Paul Newman, John J Leonard, David Cox, Peter Corke, and Michael J Milford. Visual place recognition: A survey. *IEEE Transactions on Robotics*, 32(1):1–19, 2016.

[32] Sergei Lupashin, Markus Hehn, Mark W Mueller, Angela P Schoellig, Michael Sherback, and Raffaello D'Andrea. A platform for aerial robotics research and demonstration: The flying machine arena. *Mechatronics*, 24(1):41–54, 2014.

[33] E Maset, A Fusiello, F Crosilla, R Toldo, and D Zorzetto. Photogrammetric 3d building reconstruction from thermal images. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 4:25, 2017.

[34] Meysam Moallem. A micromachined millimeter-wave radar technology for indoor navigation. 2014.

[35] Elias Mueggler, Nathan Baumli, Flavio Fontana, and Davide Scaramuzza. Towards evasive maneuvers with quadrotors using dynamic vision sensors. In *Mobile Robots (ECMR), 2015 European Conference on*, pages 1–8. IEEE, 2015.

[36] Raul Mur-Artal, Jose Maria Martinez Montiel, and Juan D Tardos. Orb-slam: a versatile and accurate monocular slam system. *IEEE Transactions on Robotics*, 31(5):1147–1163, 2015.

[37] Gaku Narita, Yoshihiro Watanabe, and Masatoshi Ishikawa. Dynamic projection mapping onto deforming non-rigid surface using deformable dot cluster marker. *IEEE transactions on visualization and computer graphics*, 23(3):1235–1248, 2017.

[38] FJ Perez-Grau, F Caballero, L Merino, and A Viguria. Multi-modal mapping and localization of unmanned aerial robots based on ultra-wideband and rgb-d sensing.

[39] Nathan Piasco, Désiré Sidibé, Cédric Demonceaux, and Valérie Gouet-Brunet. A survey on visual-based localization: On the benefit of heterogeneous data. *Pattern Recognition*, 74:90–109, 2018.

[40] Henri Rebecq, Guillermo Gallego, Elias Mueggler, and Davide Scaramuzza. Emvs: Event-based multi-view stereo—3d reconstruction with an event camera in real-time. *International Journal of Computer Vision*, pages 1–21, 2017.

[41] Renato F Salas-Moreno, Ben Glocken, Paul HJ Kelly, and Andrew J Davison. Dense planar slam. In *Mixed and Augmented Reality (ISMAR), 2014 IEEE International Symposium on*, pages 157–164. IEEE, 2014.

[42] Renato F Salas-Moreno, Richard A Newcombe, Hauke Strasdat, Paul HJ Kelly, and Andrew J Davison. Slam++: Simultaneous localisation and mapping at the level of objects. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 1352–1359. IEEE, 2013.

[43] Yohann Salaün, Renaud Marlet, and Pascal Monasse. Line-based robust sfm with little image overlap. In *3DV 2017 International Conference on 3D Vision 2017*, 2017.

[44] Milton CP Santos, Lucas V Santana, Alexandre S Brandão, Mário Sarcinelli-Filho, and Ricardo Carelli. Indoor low-cost localization system for controlling aerial robots. *Control Engineering Practice*, 61:93–111, 2017.

[45] AF Scannapieco, A Renga, G Fasano, and A Moccia. Ultralight radar for small and micro-uav navigation. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 333–338, 2017.

[46] Stefan Schubert, Peer Neubert, and Peter Protzel. How to build and customize a high-resolution 3d laserscanner using off-the-shelf components. In *Conference Towards Autonomous Robotic Systems*, pages 314–326. Springer, 2016.

[47] Juan D Tardós, José Neira, Paul M Newman, and John J Leonard. Robust mapping and localization in indoor environments using sonar data. *The International Journal of Robotics Research*, 21(4):311–330, 2002.

[48] Charles Toth and Grzegorz Jóźków. Remote sensing platforms and sensors: A survey. *ISPRS Journal of Photogrammetry and Remote Sensing*, 115:22–36, 2016.

[49] Gaochang Wu, Belen Masia, Adrian Jarabo, Yuchen Zhang, Liangyong Wang, Qionghai Dai, Tianyou Chai, and Yebin Liu. Light field image processing: An overview. *IEEE Journal of Selected Topics in Signal Processing*, 11(7):926–954, 2017.

[50] Pietro Zanuttigh, Giulio Marin, Carlo Dal Mutto, Fabio Dominio, Ludovico Minto, and Guido Maria Cortelazzo. *Operating Principles of Structured Light Depth Cameras*, pages 43–79. Springer International Publishing, Cham, 2016.

[51] N Zeller, F Quint, and U Stilla. Narrow field-of-view visual odometry based on a focused plenoptic camera. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2(3):285, 2015.

[52] Ji Zhang, Michael Kaess, and Sanjiv Singh. A real-time method for depth enhanced visual odometry. *Autonomous Robots*, 41(1):31–43, 2017.

[53] Ji Zhang and Sanjiv Singh. Low-drift and real-time lidar odometry and mapping. *Autonomous Robots*, 41(2):401–416, 2017.