

3D Sensing/Gesture Recognition and Lumentum



How 3D Sensing/Gesture Recognition Works

The first generation of 3D sensing systems worked much like human 3D recognition in nature. A light source such as the sun bathes an object in a full spectrum of light. The eye senses reflected light, but only a limited portion of the spectrum. The brain compares a series of these reflections and computes movement and relative location.

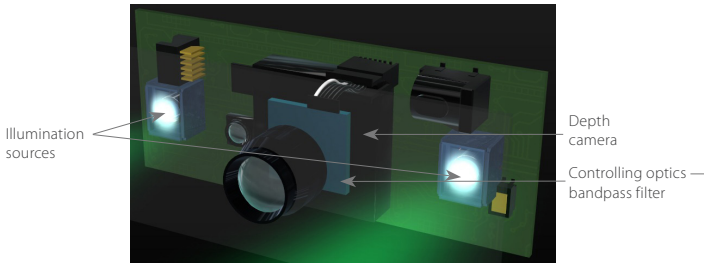
In a computer game where a player swings a golf club in front of 3D sensor, illumination sources in the device flood the player and surrounding area with invisible, near-infrared light. The light bounces off the player and reflects back to the device. Optical filters screen out spurious and ambient light, letting only the near-infrared spectra through to the light sensor. Interpreting differences in the light bouncing back from different parts of the play, firmware creates an electronic 3D map of the player and sends it to the computer game. The result is a very realistic gaming experience with a minimum of accessories such as remotes.



Basic 3D Sensing System Components

Despite the number of different technologies that support 3D sensing systems, they all share a basic component list:

- **Illumination Sources** – LEDs or laser diodes typically generating infrared or near-infrared light. This light isn't normally noticeable to users and is often optically modulated to improve the resolution performance of the system. Lumentum is a major supplier of illumination sources.
- **Controlling Optics** – optical lenses help optimally illuminate the environment and focus reflected light onto the detector surface. A bandpass filter lets only reflected light that matches the illuminating light frequency reach the light sensor, eliminating ambient and other stray light that would degrade performance.
- **Depth camera** – a high performance optical receiver detects the reflected, filtered light, turning it into an electrical signal for processing by the firmware.
- **Firmware** – very-high-speed ASIC or DSP chips process the received information and turn it into a format which can be understood by the end-user application such as video game software.



Illumination Sources

These have been important gating items in the proliferation of gesture-recognition devices. Due to their inherent spectral precision and efficiency, infrared diode lasers are a preferred option, particularly for high-volume consumer electronic applications used in environments with significant ambient lighting. These applications are characterized by a limited source of electrical power and a high density of components, factors that drive a need to minimize dissipated thermal power.

The lasers often work with other, wavelength-sensitive optical components such as filters and detectors that require tight wavelength control over a wide temperature range. And, for high data-rate systems such as 3D sensing, these lasers must operate with very low failure rates and with minimal degradation over time.

Safe, low-cost, low-power lasers have been used for decades in scanners, DVDs, and other consumer applications. 3D sensing applications typically use diffusion elements, reducing the amount of laser light to a small fraction of that of a light bulb. Customizable lasers can generate wavelengths of 800/1000/1310/1550 nm, at powers from 1 mW to 10 W, with speeds up to 1 GHz, and with typical 55% efficiency. This contrasts starkly with LED sources, which can be very inefficient and spectrally imprecise.

A gating issue with these light sources has been manufacturability. With any high-volume, low-margin consumer electronic component, the component cost profile can make the difference between an inventor's fantasy and a revenue-generating reality. Job shops and boutique contract manufacturers can effectively produce high-performance proof-of-concept or prototype parts in very low volumes. However, producing the parts in the millions to very tight performance standards (Lumentum manufacturing has a return rate of less than 2 ppm/40M) is a completely different challenge, and the margins on consumer electronics are extremely thin. Globally, the number of qualified manufacturers for high-volume 3D sensing components is very low.

Lumentum has decades of experience manufacturing optical components for telecom, submarine, and space applications. With its wafer-level passivation capabilities, short-and long-term reliability prediction techniques, proprietary method for incorporating multiple lasers on a single chip, and over 40M units in the field, Lumentum is the world's premier suppliers of high-volume, high-quality illumination sources such as those used in 3D sensing systems.

Controlling Optics

Optical filters are critical components in controlling optics for 3D sensing. Typically, these are narrow bandpass near-infrared filters with very low signal-to-noise ratios in the desired band and thorough blocking elsewhere. Limiting the light that gets to the sensor eliminates unnecessary data unrelated to the gesture-recognition task at hand. Combined with noise-reducing software algorithms, this dramatically reduces the processing load on the firmware.

Depth Camera

The depth camera used for 3D sensing is typically a CMOS or CCD chip similar to those used in cell phones. There are clear differences in how these two types of technologies capture images. However, as the technologies mature, their capabilities and applicability are merging. Each can be used for high-quality imaging and each can be manufactured in high volumes with good quality at relatively low price points. Unlike the situation with illumination sources and optical filters, a number of manufacturers have the ability to effectively meet demand for these components.

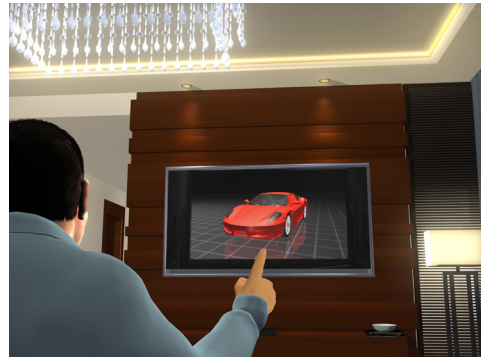


Firmware

In current applications, processing 3D sensing data is a joint task of embedded firmware in the 3D sensing device and software in the gaming console or computer. This aspect of 3D sensing is highly proprietary and developers maintain a considerable portion of their equity in 3D sensing algorithms. The important takeaway here is a paradox: 3D sensing devices are extraordinarily sophisticated in many respects, particularly in the manufacturing techniques used to build them. At the same time, 3D sensing devices are comparatively simple and low cost because of the sophistication of their design along with the availability of singular manufacturing expertise. There are no moving parts in a gesture-recognition device, and the components are all technologically mature. 3D sensing devices will rapidly get smaller, less expensive, and even more capable than they are today.

The Role of 3D Sensing in Natural User Interfaces

The essence of a natural user interface is invisibility(UI): a user interacts with technology in a way that hides the number of steps between intent and result. For example, to turn on a light, a complex interface means walking to a wall, finding a switch, and flipping a toggle. A natural interface means merely saying "light" or flicking one's hand from anywhere in a room. To play an interactive video sports game, a complex interface means using a joystick with lots of buttons. A natural interface means having the game understand body and hand gestures directly.



Letting the Machine Use Initiative

Users tend to think of a user interface in one direction: they give commands for a machine to perform. The advantage of 3D sensing as part of a natural user interface is that it is a powerful way to let machines respond to needs before they are expressed without converting the needs into commands. The human-computer interaction (HCI) works in two directions and the machine can initiate an action independently, without prompting from a human.

A physical-fitness video game enabled with 3D sensing could recognize that a user was using their left arm less than their right arm. The game could then cue the user to address the imbalance, or it could provide a specific set of game conditions that would naturally balance the arm usage. All this would take place without the user's knowing: the machine would be proactively interacting with the human in an invisible way.

Adding Functionality

3D sensing will add functionality to existing interfaces rather than eliminate them. An effective user interface provides many different ways to accomplish the same task. For consumer applications, this can be problematic; some users would prefer a smaller feature set with a simpler UI to a richer feature set with a more complex UI. Most mission-critical applications, however, need to operate under diverse working conditions and this argues for robust command systems.

For example, a home alarm system typically uses a keyboard as an authentication device to screen entry. If the user knows the code, they're in. A 3D sensing system can recognize a face, a gait, and even clothes and let a user in automatically. But rather than eliminate the keyboard and just use 3D sensing, it makes more sense to combine the two devices. With 3D sensing as an add-on to the keyboard, the door will unlock automatically as a user approaches the door; and, if the user has given the door code to a friend, the door will unlock after he or she enters the code on the keyboard.

Combining Capabilities

Similarly, 3D sensing alone can make for a rich gaming experience, but adding additional capabilities with physical devices can heighten the experience exponentially. For example, a gamer could shoot a ray gun by pointing a finger, and a physical sensor could adjust the firing rate based on the tightness of finger pressure against the fist. The natural user interface does not preclude complexity: adding 3D sensing simply magnifies the scope and richness of HCI.



The Future

Several developments will characterize future HCI:

- 3D sensing devices will be everywhere; experienced, high-quality suppliers such as Lumentum can manufacture components in high volumes at low costs
- 3D sensing will let machines recognize human needs before they are articulated as commands; it will enable greater two-way HCI
- 3D sensing will be a ubiquitous add-on capability; integrated with other technologies, it will enhance rather than replace user-interface components

3D sensing enabled natural user interfaces will soon be part of virtually all HCI. The technologies are sufficiently mature and the devices themselves are inexpensive. The first mass-market implementation was the Microsoft® Kinect™, a consumer gaming application, with over eight million units sold within 60 days of its introduction¹. Software development kits from manufacturers promote 3D sensing capabilities, and they drive an increasing number of sophisticated applications including facial recognition, background substitution, and 3D modeling.



1 Consumer Electronics Show kick-off speech, Steve Ballmer, April 5, 2011.



North America
Toll Free: 844 810 LITE (5483)

Outside North America
Toll Free: 800 000 LITE (5483)

China
Toll Free: 400 120 LITE (5483)

© 2016 Lumentum Operations LLC
Product specifications and descriptions in this document are subject to change without notice.