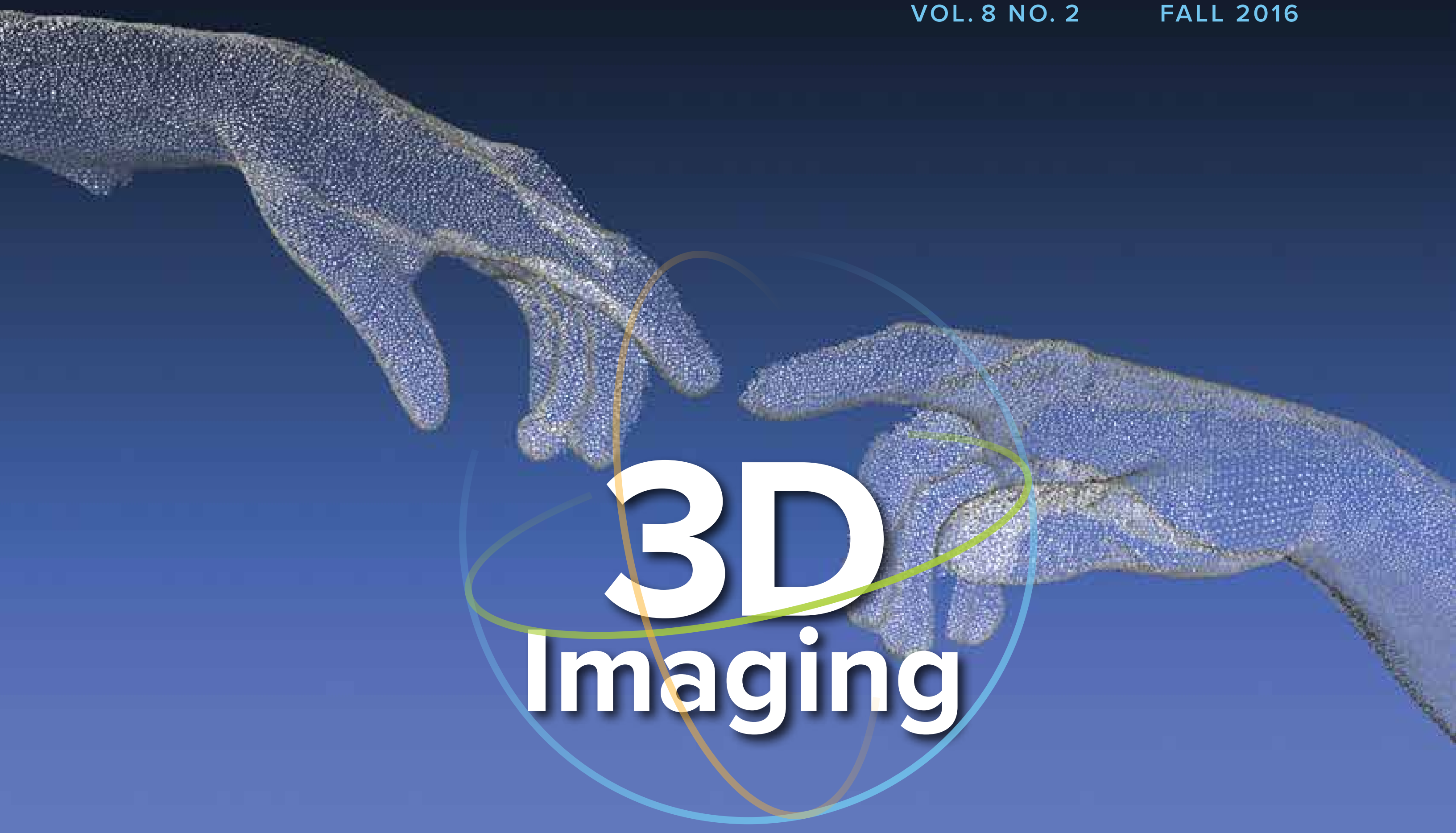
IQT QUARTERLY

VOL. 8 NO. 2 FALL 2016



3D Imaging

IQT IN-Q-TEL

IQT Quarterly is a publication of In-Q-Tel, Inc., the strategic investment organization that serves as a bridge between the U.S. Intelligence Community and venture-backed startup firms on the leading edge of technological innovation. *IQT Quarterly* advances the situational awareness component of the IQT mission, serving as a platform to debut, discuss, and debate issues of innovation in the areas of overlap between commercial potential and U.S. Intelligence Community needs. For comments or questions regarding IQT or this document, please visit www.iqt.org, write to iqtquarterly@iqt.org, or call 703-248-3000. The views expressed are those of the authors in their personal capacities and do not necessarily reflect the opinion of IQT, their employers, or the Government.

EDITORIAL

IQT Quarterly, published by In-Q-Tel, Inc.

Editor-in-Chief: Melissa Hayes

Theme Editor: Syd Ulvick

Contributing Editors: Carrie Sessine and Bobby Thorborg

Printed in the United States of America

CONTENTS

VOL. 8 NO. 2

ON OUR RADAR

What's the Buzz?

by Syd Ulvick

There is a tangible buzz around "3D imaging" today,[1] driven by emerging applications and corresponding investment opportunity. But what exactly is 3D imaging? This article provides a high-level foundation for understanding this robust and emerging sector.

Many application examples for 3D imaging come to mind, including augmented reality (AR), virtual reality (VR), computer-aided design, additive manufacturing, motion capture, medical diagnostics, photography, self-driving cars, and even surgery. The market for 3D imagers (standalone plus OEM[2] devices) is expected to reach $7.6 billion by 2020, with a compound annual growth rate (CAGR) of 39.4% from 2014-2020; $3.4 billion of that will be from the Asia-Pacific region alone.[3] (For perspective, consider that the global market for non-cellphone digital cameras is projected to hit $19.8 billion by 2020, but with a CAGR for North America of only 6.1% starting from 2013[4]). High profile acquisitions have also caught the public's attention: Apple purchased PrimeSense for $360 million, Facebook ponied up $2 billion for Oculus, and Google invested $542 million in Magic Leap.

An example of a significant emerging market for 3D imaging technology is OEM integration into cell phones. A well-known market driver for cell phones is the camera systems they contain, and 2D cameras have been improving dramatically to the point that the number of megapixels in an imaging array is no longer a differentiator. Cell phone producers are seeking new features to capture market share, and 3D imaging technology provides an opportunity; by 2018, analysts estimate that 80% of smartphones could be enabled with 3D imaging technology, with the 3D sensor elements and algorithms capturing $2.0 billion of the market's value.[5]

AR/VR is another emerging high-growth adopter of 3D imaging technology. Forecasts estimate that AR/VR could hit $150 billion in revenue by 2020, with AR alone accounting for $120 billion.[6] AR technology will likely become prolific where, like smart phones, consumers wear devices such as eyeglasses and other form factors, and the speed and quality of incorporated 3D imaging technology will play a significant role in market differentiation.

Motion tracking and gesture control applications represent another significant adopting market. For these applications it is necessary to eliminate background from the scene in order to more precisely determine and interpret foreground motion, and 3D imaging provides a solution to this challenge. The global market for gesture recognition is projected to reach $12.7 billion by 2020, with gaming consoles accounting for 53% of the market, desktop PC's 11%, TV's 10%, tablets 3%, smartphones 13%, and notebooks 9%.[7] The United States is presently the largest market, but Asia is predicted to have a CAGR of 53% between 2015 and 2020. These numbers represent a collection of emerging technologies, but 3D imaging, in particular, is gaining considerable traction in the market for gesture tracking.[8]

Self-driving cars and assisted driving systems utilize sensors for spatial awareness, and 3D imaging is particularly relevant not only for obstruction avoidance, but for object identification and anticipation. Future automobiles will have artificial intelligence that enable them to make decisions from the 3D identification of objects and a rapid determination of the likely response of those objects to the approach of the oncoming vehicle. These and other smart car technologies are projected to enjoy a market of $87 billion by 2030.[9]

An introduction to 3D imaging would be remiss without the mention of Project Tango, a platform developed by Google that has generated much press. Tango integrates motion tracking, area learning, and depth perception, the latter of which is a category of 3D imaging that detects distances, sizes, and surfaces in the environment. Tango is both hardware and software, and Lenovo recently announced their Phab2 Pro smartphone, "The World's First Tango-Enabled Smartphone". The Phab2 Pro contains four camera apertures plus a time-of-flight laser range finder; the significance of these enhancements are described in the next section.

What is 3D Imaging?

A Google search of 3D imaging produces a variety of definitions that contain the phrases "an illusion of depth in an image", "spatial information that trick a user's brain into believing and seeing depth in the images", "creating the optical illusion of depth", and "giving a three-dimensional appearance to the single image." These definitions emphasize the presentation of images in a manner that produces a perception of three-dimensionality. The emphasis of this issue of the IQT Quarterly is somewhat different: *3D imaging is the process of acquiring and producing a true three-dimensional model of an object or scene, where the model is sufficient for dimensional engineering applications.*

There are no one-size-fits-all technology solutions to this definition. Application requirements vary wildly with respect to accuracy, resolution, cost, and complexity. A solution that provides .003" accuracy and resolution for manufacturing reproduction would be woefully unsuitable for determining precise locations of windows on the side of a building. Conversely, a technology for providing virtual tours of real estate on a realtor's web site doesn't really require much dimensional accuracy at all.

How Does it Work?

A variety of 3D imaging devices are under development or are already in the marketplace, but with few exceptions they are themes or variations on the same foundational principles, either solely or in combination. A summary of these themes is presented in the following paragraphs.

Photometric Stereo Imaging

Photometric stereo imaging uses controlled illumination from different viewpoints to derive a surface shape from reflectance data. The technique is also known as "shape by shading", and is illustrated in Figure 1. Photometric stereo imaging provides good relative high-resolution image data, but cannot give absolute position data, and often poorly reproduces large-scale shapes.

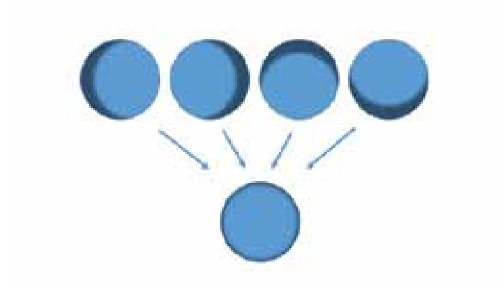


Figure 1 | Photometric Stereo Imaging, or "Shape by Shading". Input images of an object are taken using a single fixed camera position under various lighting conditions. If the locations of the light sources are known, the images can be recombined to produce a calibrated 3D image.

Geometric Stereo Imaging

In geometric stereo imaging, multiple camera apertures are simultaneously utilized at known positions (Figure 2). Geometric stereo imaging operates on the same principle that gives rise to human depth perception, where the parallax that occurs between different viewing positions (two eyes) is processed to produce a 3D image. Geometric stereo imaging provides good absolute 3D shape at larger scale. However at closer scale geometric complications begin to occur, particularly in regions of an object where equivalent images cannot be achieved for each camera.

Angular reflectance

This technique is best suited for near-field (close up) applications. Typically, a highly characterized incident light source (laser) reflects from the surfaces of the object to be imaged. Light detectors are arranged in a manner that permit the reflected light angles to be determined,

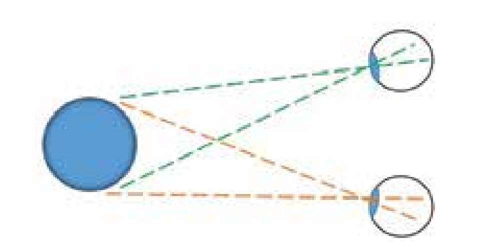


Figure 2 | Geometric Stereo Imaging. In geometric stereo imaging, multiple camera apertures are used to capture images from different viewing angles. A knowledge of the angles and the distance between the camera apertures are used to construct a 3D image.

from which a 3D model of the object can be constructed. Alternatively the light source may be moved, or both the light source and the detector can be moved together. Angular reflectance approaches can produce very high levels of resolution, but only at close-up ranges.

Shape from Polarization

Light that reflects from a surface does so with favored ranges of polarization, even from natural incident light. The reflected polarization and the degree to which the reflected light is polarized is a function of the surface angle, surface texture, and surface material. An analysis of the reflected light polarization can be used to generate a 3D model of an object. Techniques that rely upon properties of the object itself, like polarization, may be uniquely valuable for long-distance imaging.

LIDAR and Time of Flight

LIDAR (Light Detection and Ranging) and ToF (Time of Flight) techniques derive a 3D shape by using a laser to determine the relative distance between the laser source and the object (z-axis). By scanning the laser in the up/down and side/side directions across the surface (x and y axes), the shape of the entire object can be deduced.

There are two basic types of LIDAR. The first technique measures the phase shift between the laser source and the return signal at the detector, while the second technique measures the return time of a light pulse directed at the object. LIDAR techniques are very accurate in the z-direction but exhibit lesser resolution in the x and y axes.

Direct Contact

Direct contact imagers can produce very high resolution 3D images of objects for cases where the object can be directly touched by a contact sensor. The accuracy and resolution therefore depend upon the precision by which the position of an articulating arm that contains the sensor can be determined. The collection of points in three-dimensional space are then connected together to generate a 3D point cloud. While the accuracy and precision of this approach can be very high, the method is time consuming as the reader can imagine.

Structured Light Illumination

Oftentimes a featureless flat surface is difficult to image accurately. Also it can be difficult to discriminate foreground from background in a scene if the colors are all the same or if there are multiple similar objects. Structured light illumination offers a solution to these challenges, albeit at the expense of greater hardware complexity.

Structured light illumination is illustrated in Figure 3. A known pattern is projected upon the scene to be imaged, and an analysis of the illumination pattern as modified by the scene accommodates 3D reconstruction. Projected lines, square grids, light “points”, and a variety of complex patterns have all been used in combination with an infrared light source.

Single Aperture Approaches

There are a number of innovative approaches to combining one or more techniques into a single aperture, with the

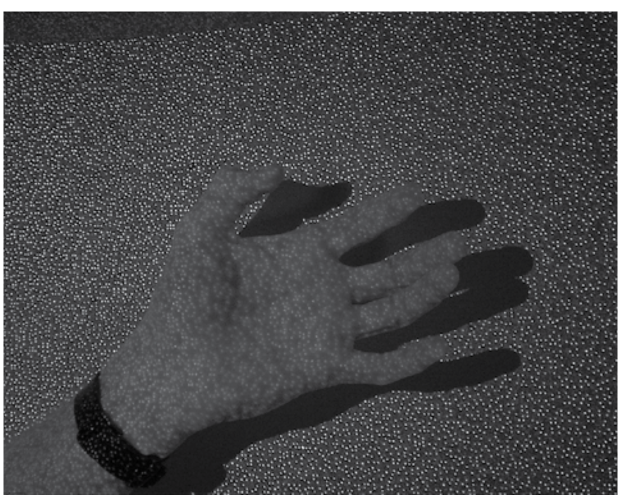


Figure 3 | Structured light pattern from an Xbox Kinect.

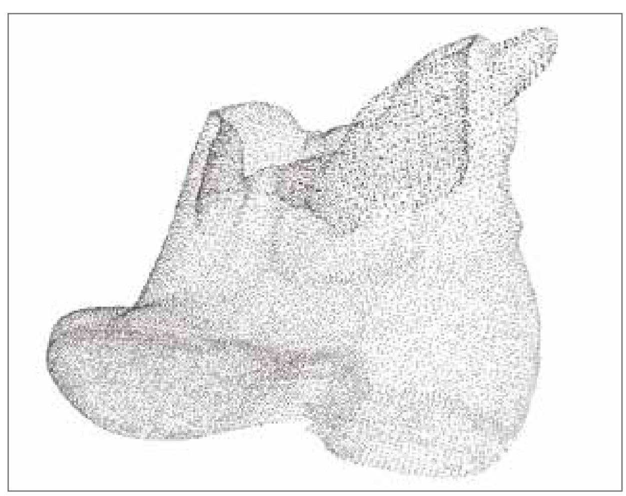


Figure 4 | Example of a point cloud.

goal of reducing size, cost, and complexity while maintaining system performance.

One approach utilizes a lenslet array, placed over a corresponding sensor array. Each lens in the lenslet array is precisely aligned over its own sensor array element. In this configuration, each array element samples the scene from a slightly different geometric perspective than the other array elements within the same sensor. If the angles for each element can be deduced, 3D information can be determined. Note that the refraction angle is a function of wavelength, thus a comparison of images by "color" will also provide depth information.

One method goes a step further and utilizes a single lens to analyze the difference in refraction between infrared light and visible light from a scene; when one wavelength of visible light is tightly focused by a lens onto a detector plane, other wavelengths of light will be out of focus by varying degrees that depend upon their wavelength. With a suitable detector array, infrared light can be imaged separately from visible light, after which the differences in image sharpness reveal depth information from which a 3D image can be generated.

Note that both of the above single aperture approaches produce information that can be recalculated later to enable a variety of interesting features. For example, an image may be "refocused" after the fact, to a different depth of field.

Approximation Approaches

Approximation methods may be sufficient in cases where the object to be imaged is a collection of geometric shapes, especially planes. Typically the user assigns specific data points to known features on a 2D image, after which a 3D representation is approximated. For example, the roof of a building may be known to be comprised of two rectangular planes, joined along one edge to form a ridge-line. In a 2D image of the roof, the rectangular planes will appear as parallelograms. If the user specifies that the parallelograms are really rectangles, an algorithm can determine the roof angles and true dimensions of the roof plane, once the four corners of the roof plane are designated as such by the operator.

Number Crunching

The literal presentation of a 3D image can take many forms, but in general they all first rely upon the generation of a "point cloud", which is a collection of known data points within a 3D coordinate system. An example of a point cloud appears in Figure 4. For full color presentations of a 3D image, each point in the point cloud might also have a color value associated with it.

> "Conversely the lack of a design win can leave a company scrambling to identify an alternative business strategy, especially in "winner take all" markets."

Real objects and scenes are not composed of finite points in space, but are a continuum. There are many steps required to transition from a point cloud to a realistic full color 3D representation, which typically first involves the generation of a mesh that connects the points in the point cloud, followed by algorithmic smoothing and correction. Color may be assigned to each geometric mesh element. The point cloud may also be a statistical compilation of overlapping images and the result of signal averaging processes. If multiple images are averaged, the system must also take into account physical motion of the imaging system while the data is being collected. Taken together, all but the smallest of images requires substantial amounts of memory and computational horsepower.

There are a variety of approaches to handling the computational resources for 3D imaging, ranging from novel algorithms that can perform processing on a smart phone, to operating the imager while being connected to a computer, to post-processing all data in the cloud. The approach to processing 3D imagery can be a market differentiator.

The "Design Win" Challenge

Another market differentiator is the "design win" challenge, particularly in the smart phone sector. As indicated at the beginning of this article, the payoffs can be substantial for an OEM (original equipment manufacturer) 3D imager that is incorporated into a smart device, which is the definition of a design win. Conversely the lack of a design win can leave a company scrambling to identify an alternative business strategy, especially in "winner take all" markets. Some technologies are very high performance, but do not lend themselves well to integration into consumer electronics; a standalone device may relieve the design win challenge, but may also be less attractive to investors. The approaches to market engagement by 3D imaging companies in response to these factors are diverse and interesting.

Concluding Thoughts

- There are numerous compelling markets associated with 3D imaging, each of which represent an investment opportunity;
- 3D imaging technology can enable differentiation for existing markets;
- There are a variety of 3D imaging approaches, and there are no one-size-fits-all 3D imaging technology solutions;
- OEM design wins can drive company success; there are interesting examples of alternative approaches.
- 3D imaging greatly leverages (and relies upon) advances in computing power and memory.
- While there are niche applications, 3D imaging is not a niche sector and is an important emerging technology area. **Q**

Dr. Sydney Ulvick *has more than 28 years of start-up, technology development, and strategic venture investing experience in high technology companies. He has been at IQT since 2003, where he is presently a Senior Vice President in the Field Technologies Practice. He has served in a number of management capacities (VP, SVP, President) in medium and small companies, including start-ups, and has hands-on experience evaluating solutions in the field. His past and present efforts have involved biotechnology, materials science, nanotechnology, electronics and components, internet of things, sensors, chemical detection, lasers, optics, power, and imaging, applied to medical diagnostics, security, specialty chemicals, consumer electronics, alternative energy, manufacturing, defense, and intelligence.*

References

1. 3D = Three Dimensional
2. Original Equipment Manufacturer
3. https://www.alliedmarketresearch.com/press-release/global-3D-camera-market-is-expected-to-reach-7-6-billion-by-2020-allied-market-research.html
4. https://www.linkedin.com/pulse/digital-camera-market-size-worth-usd-1977-billion-2020-pawar-6046907829464346624
5. https://www.alliedmarketresearch.com/press-release/global-3D-camera-market-is-expected-to-reach-7-6-billion-by-2020-allied-market-research.html
6. Digi-Capital, http://www.digi-capital.com/news/2015/04/augmentedvirtual-reality-to-hit-150-billion-disrupting-mobile-by-2020/#.V5km_2fdW6I
7. Global Industry Analysts, Inc.: http://www.strategyr.com/pressMCP-7897.asp
8. Global Industry Analysts, Inc.: http://www.strategyr.com/pressMCP-7897.asp
9. http://spectrum.ieee.org/cars-that-think/transportation/self-driving/self-driving-cars-market-2030

A Look Inside

This issue of the *IQT Quarterly* dives into the evolving world of 3D sensing and imaging, by examining the emerging technology as well as the many real world and technical applications.

Syd Ulvick of IQT's Field Technologies Practice opens this issue with an introduction to the "ins and outs" of the 3D imaging industry, which includes a look at both the various technical aspects as well as the opportunities for real world adoption. He discusses the various approaches to this technology, and in turn, the large market space in which it fits.

Jason Trachewsky with Aquifi explores the intersection between machine learning and 3D imaging as he explores the process of producing 3D images with low-cost, two-dimensional sensors.

We then take a look at the innovative use of 3D imaging systems to capture big data, as Christopher Machut of Netarus explores the use and findings of unmanned 3D imaging systems.

Henry Kapteyn and Margaret Murnane, co-founders of Kapteyn-Murnane Laboratories Inc., out of the University of Colorado, Boulder, go on to explore the possibilities and methods of 3D imaging at the nanoscale.

Next, Amy Lessner and Thomas Burnet of Rattan Software provide an exciting look into Zebra Imaging, producers of the first ever light-field display table and the benefits and uses of 3D light-file visualizations.

Scott Ackerson and Don Meagher of Quidient take us through an exploration of Intrinsic Imaging, a new, powerful 3D imaging technology, utilizing image polarimetry and advanced scene reconstruction technology.

We then change directions and get an inside look into life at a start-up in the 3D sensing field. Richard Neumann of 2r1y gives us an inside look into the quest for augmented illumination and all of the bumps in the road on the way there.

Finally, we close the issue with a technology overview from IQT portfolio company Fuel3D, whose unique, high-speed 3D scanning technology is changing the way we capture 3D images. **Q**

Making Machines See in 3D Using Inexpensive 2D Image Sensors

by Jason A. Trachewsky

Cameras are now embedded in a vast array of devices, from smartphones to game consoles, automobiles, security systems, quality control equipment, and medical devices. The two-dimensional view of the world they provide not only creates a record of the past but also enables real-time detection of hazards, improvements to health care, and reduction of the cost of making and shipping goods among many other applications. More recently, sensors that provide a three-dimensional view of their environment have expanded these capabilities, accurately inferring the shape, size and location of nearby objects. Some of the recent work in 3D sensing has built upon the large investment made in producing small, low-cost 2D image sensors to produce a compact 3D scanner for small accessories, including 3D imaging smartphone cases, as well as possible future integration into phones, AR and VR devices and drones.

Introduction to 3D scanning

The market for 3D scanning products is in the early stages of development. The most profitable of the usages typically have been industrial, and the number of deployed devices have been relatively small due to the cost of the scanning technology. These devices have mostly been used for product inspection and reverse engineering. In the past five years, the continued reduction of the cost of computation and emergence of new technologies like those used in the Microsoft Kinect have significantly dropped the price of 3D scanning, enabling a broader range of industrial applications as well as a number of consumer usages including entertainment.

The goal of a 3D scanner is to create a model of the boundary of the object, which typically is represented as a mesh comprising connected polygonal faces, each polygon typically being a triangle. The polygons may also have associated colors or may simply represent the local geometry of the object. This mesh is fit to a set of unordered points, a "point cloud", that is estimated using one of

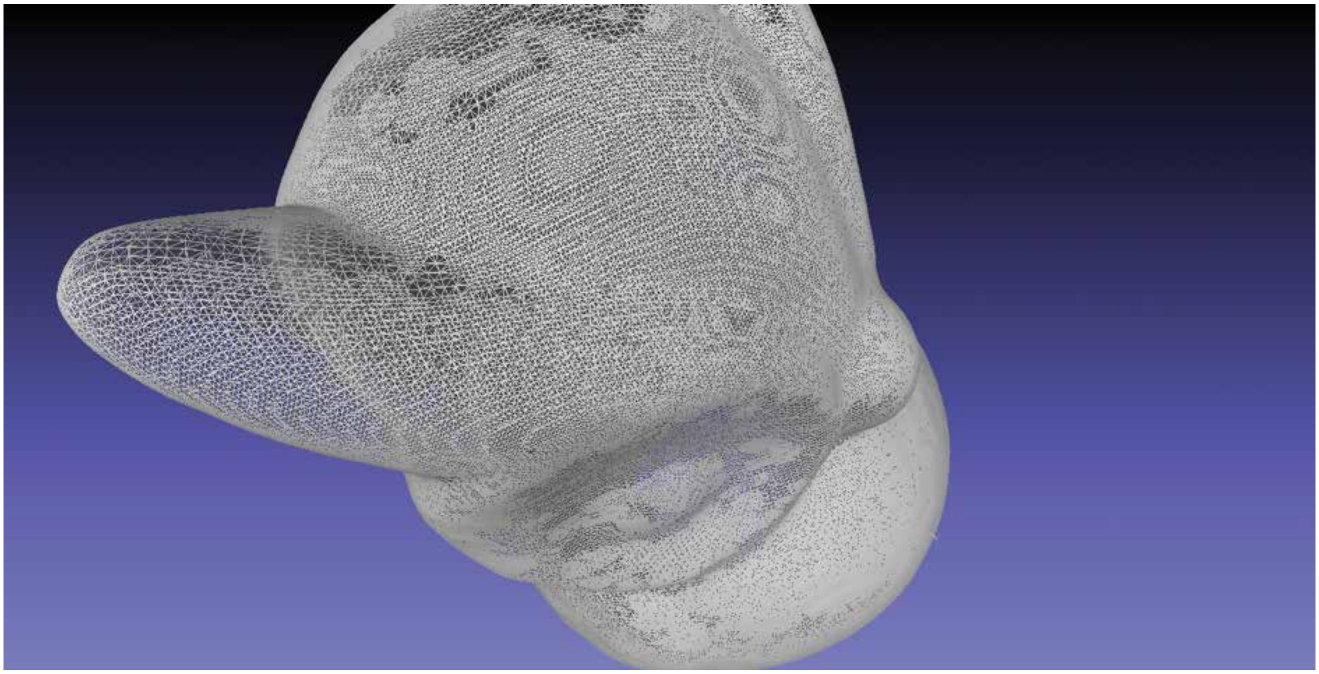


Figure 1 | Triangular mesh of figurine boundary, generated by Aquifi 3D scanner

several techniques. Figure 1 shows an example of a triangular mesh with roughly 70,000 faces estimated from a scan of a small figurine.

The depth camera: a key component of low-cost 3D scanners

The first step in creating a 3D model with accurate dimensions generally involves a depth camera. While it is possible to use features like edges and corners extracted from several images captured by a single moving 2D camera to reconstruct a 3D point cloud, the resulting model will not have accurate dimensions without the use of many scale cues, like objects of known dimension, in several parts of the scene visible at different depths from the camera. Gyroscopes and accelerometers may also be used with a single camera to produce a higher-quality point cloud, but these tend to introduce distortions in the dimensions of objects in the scene due to biases and drift in the output(s) of the underlying sensors. The highest-quality results are always produced by sensor packages that estimate accurate 3D information about surfaces at each sampling instant rather than over several contiguous samples in time.

Depth cameras are like ordinary 2D cameras, except that they produce a monochrome image or map in which each pixel represents a distance from the camera or, equivalently, an (x, y, z) position in the real world. Figure 2 shows an example depth map, in which warm colors like red and orange map to surfaces that are closer to the camera and deeper blue colors represent objects at greater distance from the camera.

The majority of commercial depth cameras broadly are either geometric sensors, which include stereoscopic, "structured light", and laser line scanners, or sensors that estimate distance from the propagation time of very short light pulses or the phase shift of a continuous wave reflected from a surface. The latter sensors are subclasses of LiDARs

Figure 2 | Depth map, one snapshot in time

with multiple receivers (pixels) and the ones using short pulses of light are often called "Time of Flight" (ToF) sensors. ToF and phase-shift detection sensors without a total optical output power limit have an absolute depth estimation error that can be roughly constant with distance but suffer from multipath interference from nearby objects and bias due to harmonics. They can require several Watts of electrical power even at relatively short maximum distances to targets due to the need to uniformly illuminate a large patch of solid angle. Laser line scanners use a laser with a diffractive optical element and a sweeping mechanism to produce a horizontal or vertical line that is moved across the scene and one or more cameras sensitive to the projected wavelength; they can accurately estimate disparity and hence depth at lower optical output power but are very slow. Structured light scanners like the first Microsoft Kinect project a two-dimensional pattern at a particular wavelength and measure the distortion of the pattern (or 2D code) due to the variable depths of nearest surfaces using a single camera to more quickly estimate disparity. The projector and camera are calibrated together to enable depth estimation. Fundamentally, the performance of structured light scanners is very similar to that of the active stereoscopic cameras described below, but structured light scanners are more sensitive to blurring or distortion of the projected pattern due to process and temperature variation, requiring careful heat sinking, and are totally unusable outdoors in sunlight due to its interference with the projected code.

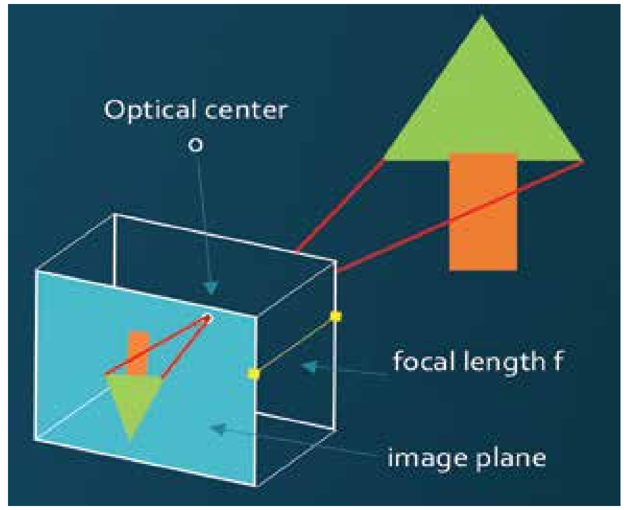


Figure 3 | Pinhole camera model

Depth from stereo image pairs

A pair of eyes enables the human brain to generate depth cues based upon the disparity between the apparent locations of features like edges and corners in the image from the left vs. the right eye. A pair of camera sensors behaves in the same way.

When two cameras have equal focal length f and are separated by a physical distance B (called the "baseline"), the estimated depth of any point from the camera can be determined from a distortionless pinhole camera model (Figure 3) as:

$$\hat{Z} = \frac{B \cdot f}{d}$$

where d is the disparity or distance along a straight line from a feature's position in the image plane of the left camera, relative to a boundary of the image, to the same feature's image position in the image plane of the right camera, relative to the same boundary of the right camera; disparity increases as the feature point nears the observer. Estimating a point's disparity from the pair of 2D images, however, is a challenge. Features of some sort must be identified in one image and then again, in a different position, in the other. A best set of matches between points in the left and right images must be found, usually through the use of local features like edges and corners. This correspondence (point matching) problem is illustrated in Figure 4 and may involve searching the entire space of

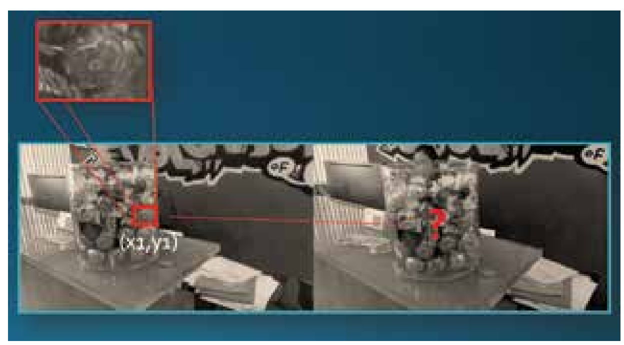


Figure 4 | Finding correspondences by matching features

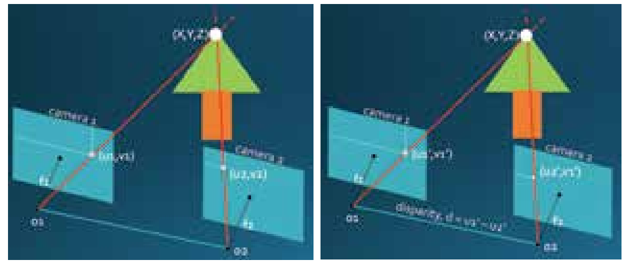


Figure 5 | Image planes showing top of tree in each for the general un-calibrated case (left), and same diagram with properly-rectified images (right)

possible matches, perhaps with smoothness constraints to mitigate sensor noise, interference and low reflectivity of some regions, or through a simpler search over a local region in the right image determined by the detected feature location in the left image.

With two or more cameras, even carefully oriented relative to each other, there will be some misalignment either during or after manufacturing, and a computationally-expensive two-dimensional search is required to find these left vs. right image feature matches. A shortcut involves geometric calibration of the multi-camera system in software so that the search may be performed only along a single dimension, as shown in the right image in Figure 5. This calibration usually requires a set of known calibration targets like checkerboards with squares of known dimensions and is performed in the factory. Keeping the camera system in calibration under mechanical stresses like drops and torsion (twisting) and temperature changes, however, is a challenge, and a dynamic calibration system that uses images normally collected by the camera system to update estimates of the calibration parameters during operation generally is required.

When only ambient light is used by the cameras, the depth camera is considered a “passive stereoscopic system”; when light is projected onto the scene, the depth camera is an “active stereoscopic system”. In the latter system, a pseudo-random 2D pattern or code is typically transmitted to facilitate disparity estimation on surfaces without visible features, like white walls. Aquifi has implemented both types of systems for different applications, and both are based upon the same fundamental concepts and algorithms.

One of the existing software implementations of a passive stereo depth camera using Aquifi’s techniques is found in the Dell Venue 8 7800 Android tablet family. In the tablet’s rear camera system, three visible light color (RGB) cameras are used in a “trinocular” configuration to generate an estimate of the (x, y, z) location of the closest opaque surface to the camera at each image pixel. The tablet uses one autofocus 8 Mpixel camera and two fixed-focus 720p cameras in an isosceles triangle configuration. Figure 6 shows the back side of the tablet and the location of the cameras. Depth estimates are generated through the use of an efficient global disparity search algorithm, and depth is computed from disparity using the geometric calibration parameters for all three cameras. The stated operational range for the 3D camera is 30 centimeters to

Figure 6 | Dell Venue 8 7840 tablet showing trinocular camera system: two of the cameras are at opposite ends of the narrow black strip and one higher-resolution camera is at the left edge of the tablet

10 meters, though objects with sufficient texture at greater than 10 meters can be ranged in typical indoor and outdoor ambient lighting. The dimensions of objects in the field of view may be measured using a simple touch-based application.

This device also implements dynamic calibration, which re-computes many of the intrinsic parameters (like focal lengths and principal points) as well as the extrinsic geometric parameters (like relative rotations and translations) of the camera system in the field using only a set of natural RGB images without calibration standards like checkerboards. Figure 7 shows an example measurement of a planar surface. The device had been factory calibrated, then submitted to both drop and torsion tests. Before dynamic calibration was enabled, the linear measurements were mostly incorrect; after dynamic calibration was enabled, measurements were accurate to within one inch at the target distance.

While the passive stereo depth sensor is quite effective in many environments, to address some additional applications, like scanning spaces with plain white walls, the passive stereo system was augmented with an infrared

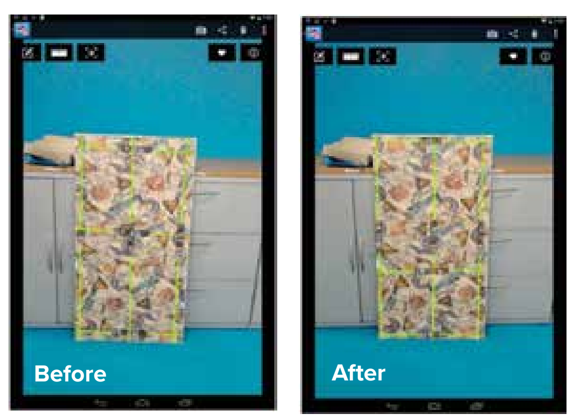


Figure 7 | Benefit of dynamic calibration (should be 4’ high, 2’ wide)

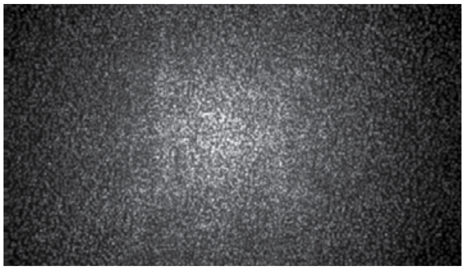


Figure 8 | IR image showing projected pseudo-random 2D pattern at 850 nm

(IR) pattern projector based on a laser diode and diffractive optical element (DOE). An example of the projected pseudo-random IR pattern is shown in Figure 8.

One implementation of the active stereo depth camera is shown in Figure 9. In this prototype design, two synchronized global shutter monochrome camera sensors with IR-pass filters and one RGB camera sensor are used to estimate depth and then to derive the shapes and sizes of imaged objects. Because two IR sensors and a RGB sensor are used together, the depth sensor is able to function indoors or outdoors, in total darkness or sunlight. Another version of the same sensor will also be implemented as a rugged smartphone case that also protects against shock and vibration.

An example of the pattern projected by this depth camera onto an object and its corresponding disparity map is shown in Figure 10. Again, warm colors in the depth map indicate points closer to the depth camera.

Generating a 3D model from a depth camera

An active stereo depth camera may be used in conjunction with a software pipeline running on a low-cost mobile application processor in our 3D scanning accessory or on a smartphone (Figure 11) to generate a complete 3D object model by moving the 3D sensor around the object. The interactive, real-time pipeline enables viewing of the point cloud updates as scanning progresses, and a non-interactive, off-line pipeline may either compute the textured mesh immediately following the interactive scan or at some convenient point in the future. Computation may be easily shifted from device to cloud or vice versa, depending on the need of the application.

Figure 9 | 3D active stereoscopic sensor protype from Aquifi with one RGB and two monochrome camera sensors

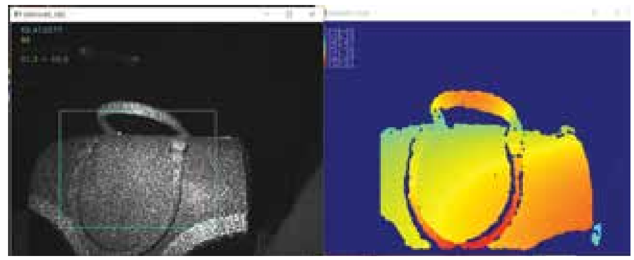


Figure 10 | IR image with projected pattern (left) and disparity map from Aquifi 3D sensor (right)

Figure 12 shows the point cloud from a purse scanned by this pipeline, with green dots indicating high-confidence object contour data and red points indicating either low-confidence points or the table top on which the purse has been placed. The current generation of 3D scanner from Aquifi typically estimates a final mesh with a r.m.s. surface position error of between 1 and 1.5 mm when an object the size of a purse is scanned at close range (e.g., 30-50 centimeters distance). Success can be assessed during the scanning process, and the same pipeline also may be used to scan the boundaries of rooms at better than 1 cm accuracy.

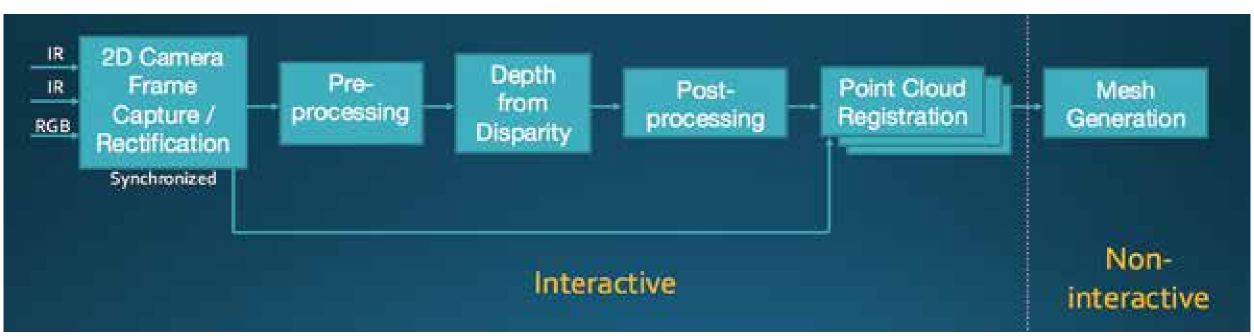


Figure 11 | 3D scanning pipeline at a high level

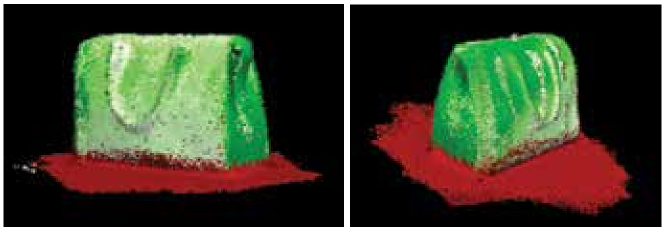


Figure 12 | Point cloud of purse

Applications of 3D scanning

Photorealistic 3D scale models of products that can be navigated in a browser or app on a 2D display or navigated in a 3D AR or VR display to guide purchase decisions will be commonplace within the next five years. Information generated from 3D scans about surface normal vectors will enable dynamic relighting of a scene including the scanned objects to see how products look together in a purchaser's home or place of work at different times of day. Scale information for each scanned object will enable determinations of how products fit together, e.g., how various products like tables and chairs might fit into a space.

Another key application area for the technology is automated inspection and quality control. Parts or goods to be packed at a distribution center may be automatically inspected for defects and identified as one of a number of possible objects and checked against a packing list. The depth of tire tread may be checked for wear, and puncture damage may be identified. This measurement may be performed automatically on trucks with low-cost scanners built into the vehicles or at measurement stations to reduce the likelihood of a costly and potentially-deadly tire blowout on the road.

There are also numerous forensic applications of a portable, low-cost 3D scanner. Recording damage from automobile accidents and automatically generating estimates, crime scene analysis, and determining the exact dimensions of a space or how it has been modified by a building contractor are all possible with a 3D scanner based on low-cost stereo sensors. Since the stereoscopic 3D scanning algorithms may be made computationally-efficient, real-time scanning on drone platforms will be possible as the markets for both drones and 3D scanners grow. By increasing the accessibility and robustness of 3D scanning technology, new solutions to many old problems will become possible. **Q**

Figure 13 | Truck tire 2D photo (top) and 3D model from which tread wear may be estimated (bottom)

Jason Trachewsky *is currently the Vice President of Engineering at Aquifi, a developer of 3D sensing technology. Prior to Aquifi, Jason was the Chief Operating Officer and a key technology developer at Passif Semiconductor, a VC-backed startup developing extremely low-power wireless technology. Passif Semiconductor was acquired by Apple, Inc. in April 2013. Jason came to Passif Semiconductor from Broadcom Corporation, where he founded the WiFi Business Unit and spent 11 years leading engineering. Jason entered Broadcom through its acquisition of Epigram, a VC-backed developer of home networking integrated circuits in May 1999. Jason has authored over 100 patents and holds a BSEE from Stanford University.*

The Importance of Expediting the Analytics Process

by Christopher Machut

On January 18, 2012, PBS aired a program, "3D Spies of World War II." The introduction of this program says it all when it comes to the importance of expediting the "process" of 3D imaging and other forms of big data interpretation. *"It is the early days of World War II. Hitler's armies storm through Europe, crushing everything in their path. Western Europe is on the brink of defeat. Allies turn to the iconic Spitfire (Britain's most advanced fighter plane) to help change the course of the war. Now the fighter plane is not just armed with weapons that kill, but with one that will keep them one step ahead of the Nazi war machine: a surveillance camera… The photo interpreters have a secret weapon to expose German secrets. Using stereoscopes, they bring the Nazi's world to life in 3D."*

These brave reconnaissance pilots took tens of millions of 3D images during the war for photo interpreters. By all accounts, this bravery and skilled interpretation, not only helped to defeat Hitler and his Nazi rule, but also changed what it meant to be strategic using analytics.

At Netarus, we recognize and highly appreciate this level of innovation and are on the cusp of expediting the analytics process with a new, first-of-its-kind solution.

A new solution to be more strategic

3D imaging and big data platforms have been evolving for decades so that industries of all kinds are more strategic. Analytics programs are the engines for this growth. According to International Data Corporation's forecast, released on November 9, 2015, spending in big data technology and services will reach $48.6 billion by year-end of 2019. Capturing this data with unmanned systems also is a multi-billion dollar market within itself.

Capturing 3D images and other big data with unmanned systems is not new nor are analytics systems. At Netarus, we saw an opportunity to use this platform with unmanned systems to enable any analytics program to be scaled and parallelized across multiple computing environments, which would significantly expedite processing times to near real-time. We figured out that the key to this expediting process is to leverage the scalability of the cloud.

This opportunity came to light after perfecting our two signature products, HoistCam™, a durable wireless camera system, and HoistCam Director™, an unmanned sensor data capture and storage system, to create a new product, SiteTrax™, an automated process of third-party analytic applications using data captured from unmanned systems in the cloud.

Background and SiteTrax™ development

When using an unmanned system to capture 3D images and other big data, the process consists of data collection then interpretation with an analytics program. While practical, this approach is using an unmanned system as a novelty, which limits the usefulness with an enterprise environment where multiple times and locations are required. Current technology for unmanned systems requires more than 24 hours to capture, store, analyze, and interpret big data produced by analytics programs. Figure 1 depicts the workflow of current data collection from unmanned systems.

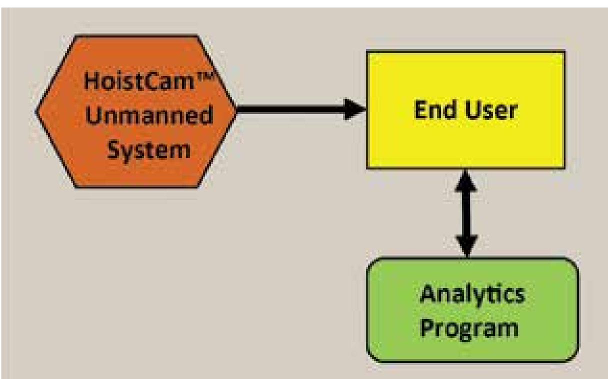


Figure 1 | Flow diagram of current data collection from unmanned system

The development of our unmanned sensor data capture and storage system, HoistCam Director™ (www.HoistCamDirector.com), improves the current data capture and storage framework by enabling 3D images and other big data to be stored in a Mobile Digital Video Recorder and pushed into the Centralized Video and Image Storage repository in the cloud almost instantly. This process indexes the data based on when the unmanned system was used and the time data was collected. When data is passed into an analytics program, this indexing positions data to be easily searched, reducing the process time of capturing, storing, and analyzing data from 24 hours to approximately 10 hours. Figure 2 shows the workflow of using our in-the-cloud unmanned sensor data capture storage system.

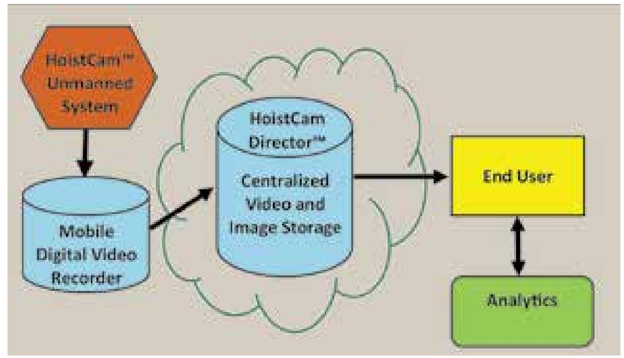


Figure 2 | Flow diagram of HoistCam Director™, Netarus' in-the-cloud unmanned sensor data capture storage system

While HoistCam Director™ expedited the capture, store, and analyze process, we saw an opportunity to expedite the process even more to near real time by migrating and automating the analyzing process in the cloud. This new solution, currently in development, is called SiteTrax™ (www.SiteTrax.io). Automating the process of third-party analytics applications, along with the data captured from the unmanned systems in the cloud, will create a first-of-its-kind analytics solution. SiteTrax™ is a data integration platform where video and images captured by HoistCam Director™ are analyzed to produce orthomosaics, 3D point clouds, and digital surface maps. This unique platform will reduce the amount of time to capture, store, analyze, and interpret data from unmanned systems to under 30 minutes. Figure 3 illustrates the workflow of our new in-the-cloud data integration platform.

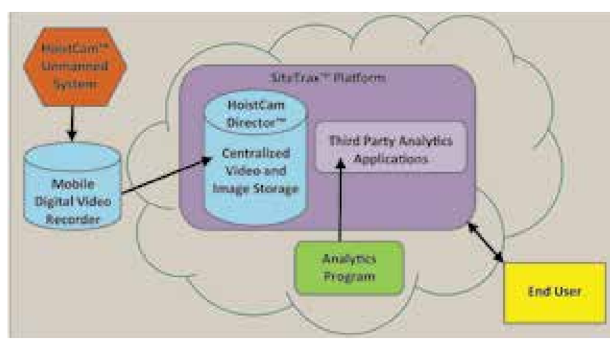
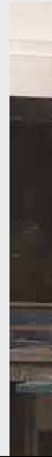


Figure 3 | Flow diagram of SiteTrax™, Netarus' in-the-cloud new solution to migrate and automate the analyzing process in near real time

SiteTrax™ is in the beta phase of development, but thanks to significant investments from regional and state-wide seed funds, the SiteTrax™ web and mobile interfaces are advancing quickly. Using the existing MySQL application

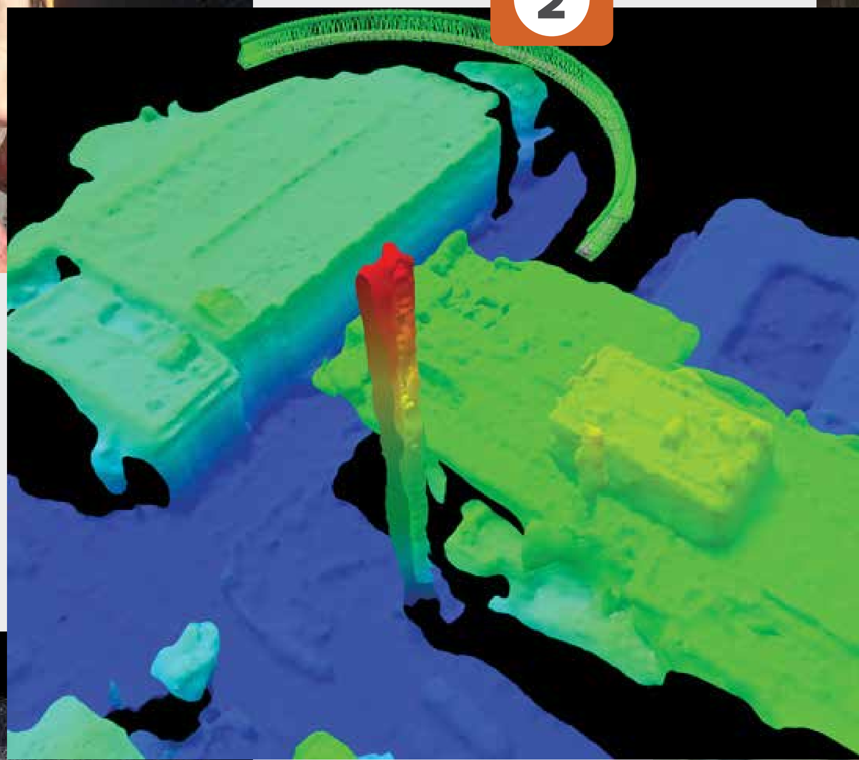
1

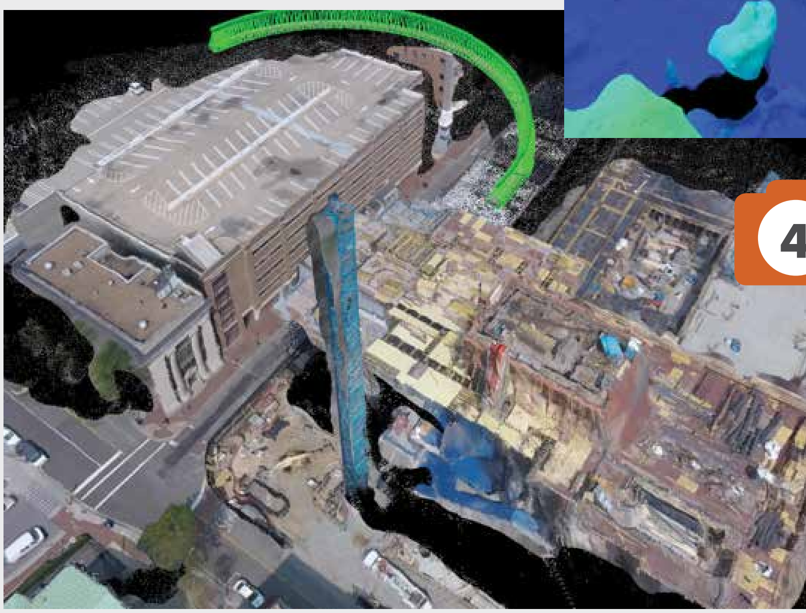
Rigger attaching a HoistCam rapidly deployable wireless camera to hook block of tower crane with four 100lb magnets and a safety lanyard.



Multiple video and image frames captured from HoistCam camera system and HoistCam Director processed into a Digital Surface Model (DSM).

2





4

Multiple video and image frames captured from HoistCam camera system and HoistCam Director processed into a 3D point cloud.

View from wireless HoistCam rapidly deployable wireless camera attached to the hook block of a tower crane.

5

3

Building under construction in Norfolk, Virginia. Hook block from tower crane on the right with HoistCam (yellow) box looking down.

program interface, we also have nearly completed development of an administrator interface to manage the communications with various analytics tools, such as Pix4D, including relevant systems and pre-condition verifications. By early 2017, we anticipate that backend event scheduling, third-party plugin, and analytics delivery systems will be completed, and interfaces will have been thoroughly field tested, positioning us to go to commercialization not long after.

Our ultimate goal of SiteTrax™ is to provide the future proof architecture for any air, land, sea, or space unmanned system data capture system whereas any third-party analytics program can analyze this stored data and generate information that can be interpreted in real time. This will allow any industry, including the U.S. intelligence community, to be more strategic with their decision-making, including their situational analysis and awareness.

Progress leading to innovation

Innovation has come a long way since WWII and the incredible use of stereoscopes. So many of us in this business treasure learning about those moments in time. This is what drives us at Netarus: the possibility to create and deliver a technology solution that can change business as usual for a single customer, industry, or the world. Q

Christopher Machut *is Netarus' chief technology officer, principal investigator, and system architect. He is the founder of Netarus, LLC, and designer of its signature products, HoistCam™, HoistCam Director™ and SiteTrax™ platforms. A serial entrepreneur, Christopher has more than 20 years working with distributed systems to tie together disparate software and hardware networks.*

Quantitative 3D Nanoscale Imaging: New Capabilities in X-ray Microscopy

by Henry Kapteyn and Margaret Murnane

INTRODUCTION

Nanotechnology has become a major driver for our economy as electronic devices, energy harvesting systems, and new materials make use of engineered structures with ≈10nm dimensions – far smaller than the wavelength of visible light. However, this rapid progress in nanotechnology presents major challenges. Typically, we can no longer directly *see* what we can *make*, and established imaging techniques such as electron microscopy are often destructive, making it impossible to image function. Traditional techniques are also typically limited to fields of view orders of magnitude smaller than devices of interest that are covered in nanostructures. In the biosciences, advanced optical imaging techniques such as *Stochastic Optical Reconstruction Microscopy* (STORM), *Stimulated Emission Depletion* (STED), and *Photoactivated Localization Microscopy* (PALM) use fluorescent labeling to image functioning biosystems with sub-diffraction limited resolution; this capability is so revolutionary that these techniques were recognized by the 2014 Nobel Prizes even at a very early stage of development. However, since these techniques require labeling, they generally cannot be used outside biology, and to date no technique can effectively capture function

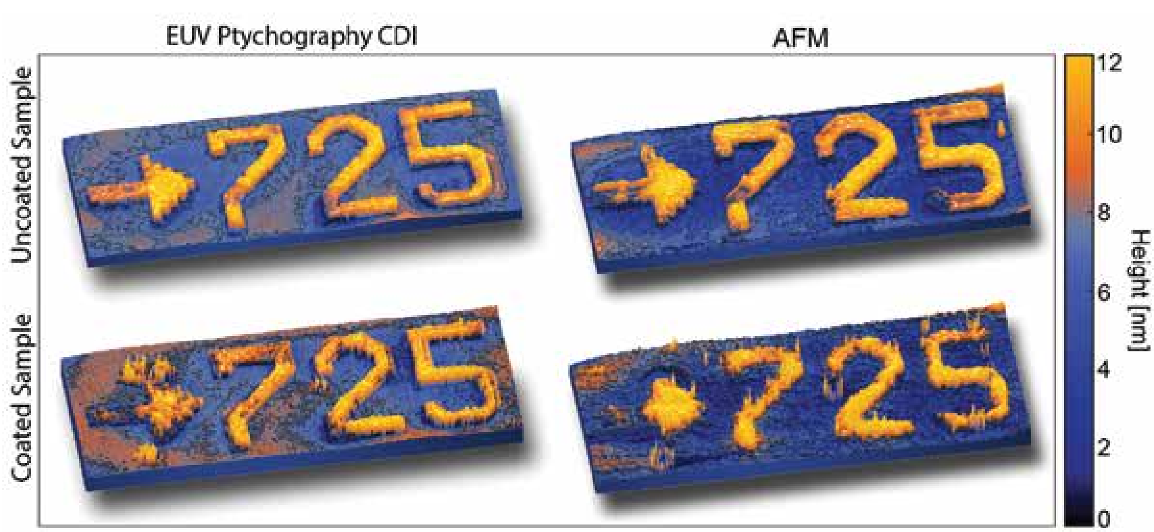


Figure 1 | (adapted from[15]) Height maps showing the surface topography of both an uncoated and Al coated sample demonstrating nanometer axial resolution and good agreement with AFM images.

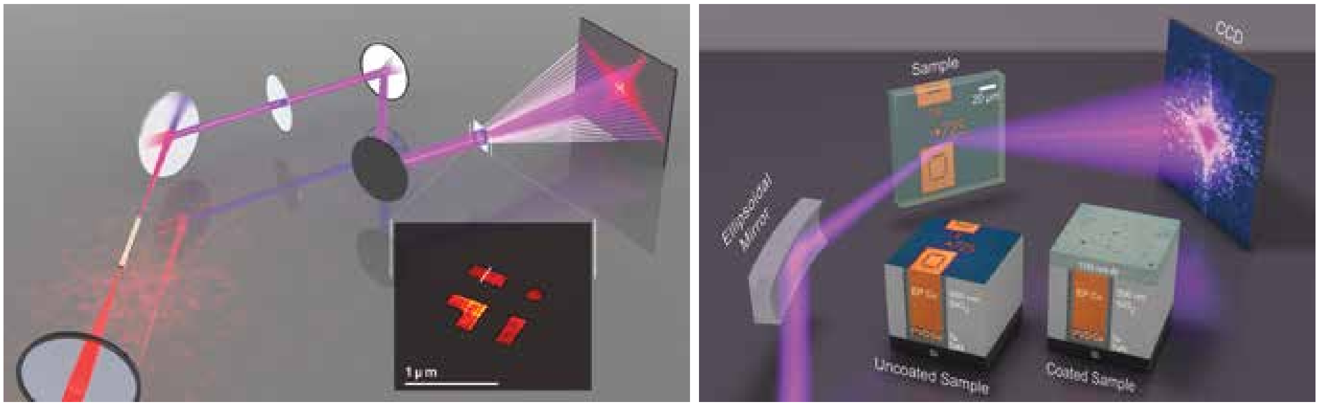


Figure 2 | (adapted from [3] and [14]) Schematic of typical CDI microscopes. An ultrashort near-infrared pulse is focused into a hollow capillary where high harmonic generation produces 13 nm or 30 nm light. The driving laser light is stripped away with the combination of silicon rejector mirrors (set at Brewster's angle for the fundamental) and thin metal filters. The light is spectrally filtered by multilayer mirrors and focused onto the sample using a spherical curved mirror or off-axis ellipsoidal mirror. The scattered light from the sample is collected on a CCD detector.

in (often visibly opaque) 3D nanosystems. This significantly slows iterative design and optimization, manufacturing process inspection, as well as presents difficulties, for example, in identifying counterfeit microelectronics.

Despite these challenges, one promising approach that has proven its ability to capture dynamic response at the nanoscale, with potential for full 3D imaging of functioning nanosystems, is short wavelength (EUV and x-ray) microscopy. [EUV stands for *extreme ultraviolet*, and is the spectral region between ultraviolet and x-ray wavelengths from ≈10-100 nm.] Recently we have demonstrated the use of coherent EUV and x-ray light to capture 3D images of semiconductor structures at the nanoscale, even beneath interfering layers such as aluminum (Figure 1). Although x-ray microscopy has a long history, the 21st century has brought revolutionary advances in coherent light sources and in the imaging techniques employed to use them. In particular, coherent high-harmonic generation (HHG) sources, x ray free electron lasers (XFELs) and coherent diffraction imaging (CDI) have given rise to a new generation of powerful nanoscale imaging techniques that have major advantages compared with other approaches: the ability to image with amplitude and phase contrast in 3D at the *wavelength limit,* revealing both quantitative height and material composition information. The combination of HHG sources and CDI has provided imaging with fidelity limited only by fundamental quantum noise; inherent elemental and compositional contrast; the ability to image into visibly opaque structures; and the ability to capture charge, heat, and spin flow in new materials and dynamically functioning nanosystems. These capabilities promise to contribute significantly to faster and smaller devices for consumers in coming years.

A revolution in coherent imaging and light source technologies

Over the past decade or so, *coherent* computational imaging techniques have been developed for both x-ray and electron microscopy, simplifying the optics required for imaging and consequently circumventing previous resolution limits. These CDI techniques make it possible to learn essentially everything possible using illumination at the light (or electron) wavelength employed in the microscope, by obtaining a full description of how the object scatters light[1-4]. The tradeoff to make use of CDI is the need for coherent, i.e., laser-like, illumination of the object, and the need for computer reconstruction of the image.

However, simultaneous with the development of CDI and advances in computing power has been the emergence of a new class of bright, tabletop-scale, coherent, short-wavelength light sources—essentially the first tabletop "x-ray lasers."[5-7] These new sources, based on the process of high-order harmonic generation (HHG), have proven themselves an ideal illumination source, and their combination with CDI has proven truly transformative. This potential has been recognized through DARPA support of this work (*Program in Ultrafast Laser Science Engineering* [PULSE] program), the recent establishment of the NSF Science and Technology Center on real-time functional imaging (STROBE) at the University of Colorado at Boulder, as well as through investment for KMLabs Inc. of Boulder, Colorado through Intel Capital and others for development of the coherent sources.

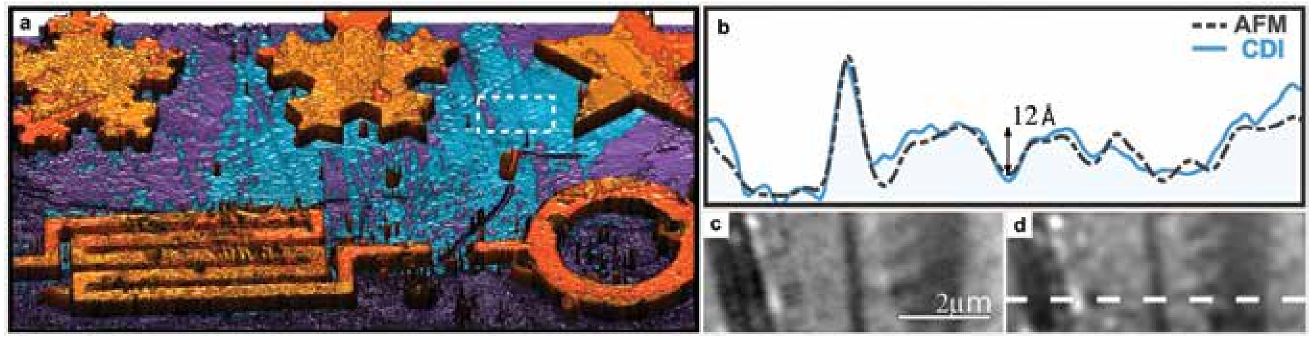


Figure 3 | Quantitative comparison between CDI using 29nm illumination and AFM. (a) 3D rendering of the CDI height map generated from a ptychography phase reconstruction. The white rectangular region highlighted in (a) is zoomed in and compared to an AFM image of the same region (c,d). The profiles plotted in (b) agree to within (6Å 95% confidence interval). The AFM image (d) was smoothed with a Gaussian point spread function to mitigate the lateral resolution mismatch between (c) & (d).

HHG is the most extreme version of "nonlinear" optics yet observed. By taking an intense femtosecond laser pulse (~10-14 sec in duration) and focusing it into a gas, it is possible to coherently upconvert this light to much shorter wavelengths, spanning the entire electromagnetic spectrum from the ultraviolet to soft x-rays—12 octaves in bandwidth to date.[8] In the HHG process, photons from the driving laser pulse are combined coherently into much higher energy (i.e., shorter wavelength) photons, through a process that is in essence a coherent, laser-driven version of the x-ray tube. The HHG process was first discovered in the late 1980s[9,10] and techniques for efficiently implementing it have developed rapidly, especially in the 2010s.[7,8] Until recently, it was thought that useful sources could only be implemented to photon energies below ~100 eV (i.e., λ>10 nm). Recently, we showed how to generate coherent x-rays on a tabletop with λ<1 nm (photon energy >1 keV), well into the x-ray region of the spectrum,[8] and to generate light with controlled linear or circular polarization often essential for obtaining enhanced material contrast in images.[11,12] Moreover, HHG is perfectly synchronized to the driving laser, making it an ideal source for stroboscopic ultrafast imaging with x-rays. HHG sources are complementary to large scale synchrotrons and x-ray free electron laser (XFEL) sources. These are also used for coherent imaging and can generate light well into the hard x-ray spectral region, but are orders of magnitude larger and more expensive, and less naturally suited for very high time resolution.

CDI is a computational imaging technique, often called *lensless imaging*.[4] In CDI, rather than creating an image of an object on a detector using a lens system, one simply illuminates an object with a coherent illumination source and captures the light scattered from it directly using a detector. An advanced "phase retrieval" algorithm then substitutes for the lens to quantitatively solve for the object that is consistent with the scattered light. From these data, one can reconstruct an image. CDI generally requires coherent (i.e. laser-like) illumination, but because it does not need any image-forming lenses, it is ideal for X-rays. Traditional X-ray microscopes have been limited in resolution solely because of technical limitations of the image-forming optics, with a best resolution of ≈ 25 nm, at least an order of magnitude worse than the theoretical resolution limit. By eliminating these optics in favor of computational image reconstruction, CDI allows diffraction-limited resolution, making full use of short-wavelength HHG sources.

CDI is a very general imaging technique and can use any coherent illumination source, including electrons. Furthermore, because it fully describes how light scatters from an object, it inherently yields 3D information. It is the most photon-efficient imaging method, since every photon scattered from the object can, in principle, be captured and detected using highly efficient CCDs with no lossy optics. Figure 2 shows the very simple setup required for coherent diffraction microscopes, which can easily work in either transmission or reflection.[3,13-16]

Ultrahigh Resolution, 3D Applications of Next Generation Coherent Microscopes

CDI with EUV light offers complementary information to well-established surface imaging techniques such as scanning electron microscopy (SEM) and atomic force microscopy (AFM). It provides chemical specificity, nanometer resolution, long working distances compared to AFM, and 3D information where most traditional modalities are limited to surface imaging. It is a reference-free, simultaneous phase and amplitude imaging modality that, in a reflection geometry, naturally yields chemical sensitive contrast in amplitude images and both chemical and height information in phase images. In the transverse direction, CDI yields diffraction-limited resolution at the wavelength limit, and ultrahigh sub-nanometer axial resolution enabling atomic scale profilometry and visualization of surface topography. We have demonstrated 6 Å axial resolution, allowing the generation of height maps from phase images with about one atomic layer precision (shown in Figure 3).

The class of algorithm that makes possible robust experimental use of CDI is called ptychography, a powerful, advanced technique that has gained popularity both at synchrotrons and with tabletop HHG sources in the last decade.[17-19] In ptychography, a beam is scanned over an area to capture the scattered light patterns from several overlapping regions. The redundancy in these data allows for a powerful reconstruction algorithm that takes into account both the measured data and the constraint that the object reconstructed has to be consistent with all of the diffraction patterns that share information from the same area on the sample. This technique simultaneously yields amplitude and phase reconstructions of both the object and the illumination, and has proven robust to noise, positioning errors, and other etc., to be of practical use in

> Recently we have demonstrated the use of coherent EUV and x-ray light to capture 3D images of semiconductor structures at the nanoscale, even beneath interfering layers such as aluminum.

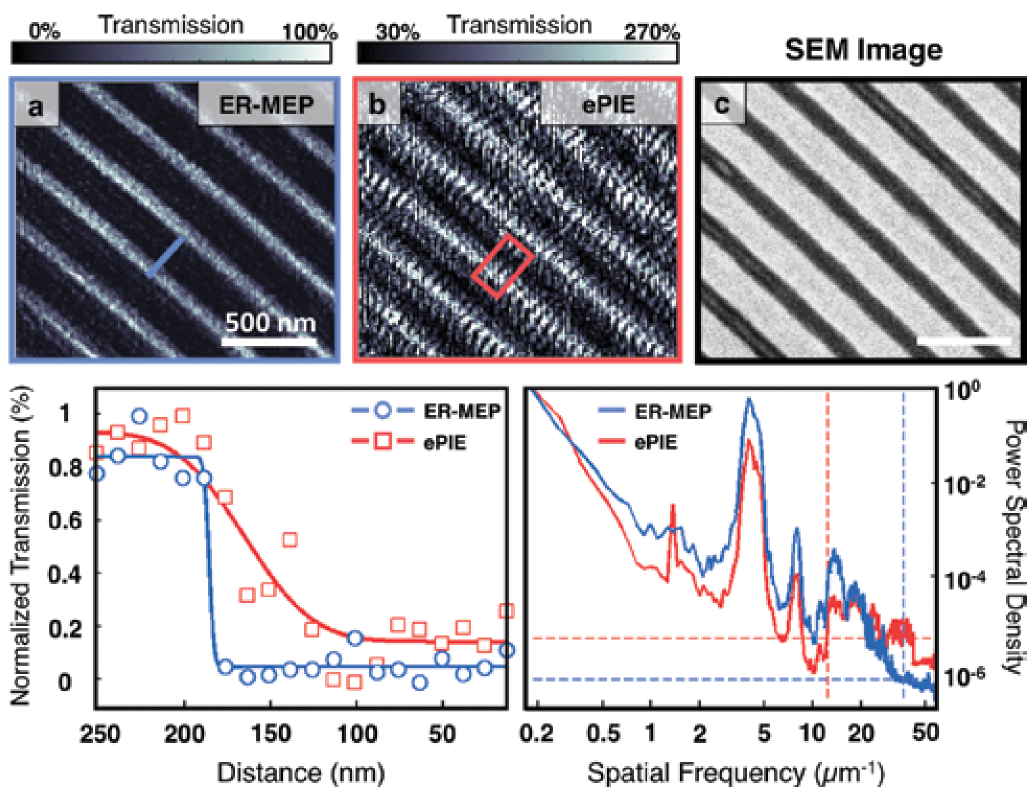


Figure 4 | (adapted from [13]) (a) 14 nm transverse resolution ptychography CDI image of a Fresnel zone plate acquired in a transmission geometry and reconstructed using the modulus enforced probe ptychography technique. (b) Traditional ptychography CDI reconstruction using the ePIE algorithm of the same data set as in (a) that suffers from artifacts. (c) SEM image of the same region, with a contrast flip compared to (a) and (b) because the SEM operates in reflection as opposed to transmission.

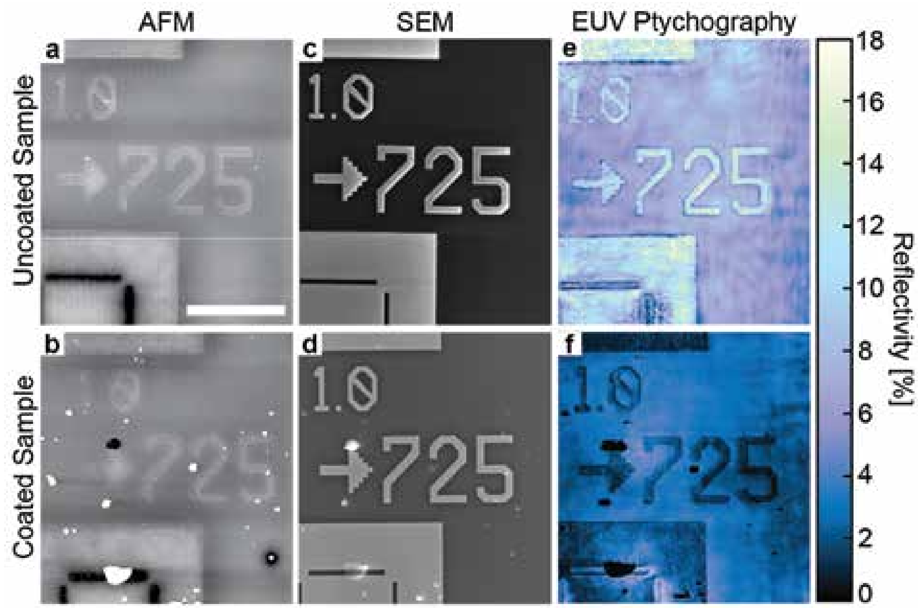


Figure 5 | (adapted from [14]) Comparison of AFM, SEM, and EUV CDI images of two identical damascene-style samples consisting of Cu nanostructures inlaid in SiO2 and polished flat, one of which is coated with 100 nm of Al. Scale bar is 20 μm and shared between all images. The contrast flip between the features and substrate from image (e) to (f) is explained by formation of an Al_2O_3 oxide layer at the Al—Cu interface and interdiffusion between the Al, Cu, and Al2O3. Contrast in (b) is due to small height variations on the surface of the Al coated sample, and contrast in (d) is due to a combination of surface height variations and chemical contrast that cannot be separated. Only (f) purely images the composition of the buried interface in a quantitative manner that allows for subsurface chemical changes to be identified.

the laboratory. Surprisingly, ptychography is even powerful enough to tolerate various sources of incoherence, including multiple spatial and spectral modes, or even propagation through optically thick, scattering media.[20-23] Furthermore, extensions of the algorithm provide calibrated, artifact-free reconstructions of the absolute reflectivity or transmissivity of the sample.[15,16] This enables truly quantitative CDI for the first time, which allows for a host of new applications in science and technology.

With tabletop microscopes employing these new, flexible ptychography techniques, CDI is just beginning to reveal its immense power and utility. Using the *Modulus enforced probe* (MEP) ptychography technique, we recently demonstrated 14nm transverse resolution using 13.5 nm illumination, a world record for any 13 nm light source.[16] (Figure 4) Acquiring images using 13 nm illumination with nanometer scale transverse resolution is crucial for semiconductor industry applications in which actinic inspection at the wavelength used to fabricate devices enables unparalleled defect inspection and quality control.

Another natural advantage of EUV light is that it is partially penetrating through a wide variety of materials, while still allowing reflection-mode imaging. Recently, we showed that it is possible to use HHG CDI to non-destructively image subsurface phenomena occurring at deeply buried interfaces.[15] We obtained highly sensitive, quantitative information about reactions and diffusion at buried interfaces with few nanometer precision using absolute reflectivity images and height maps generated by ptychography CDI. As shown in Figure 5, by observing a contrast flip between copper nanostructures and a surrounding SiO2 substrate when the structures were coated with a thick layer of aluminum, we were able to quantify oxidation reactions and diffusion at the buried interface. We also generated height maps showing the surface topography of our semiconductor samples that agreed with AFM to nanometer precision, as shown in Figure 1.

In the near future, rotation of the sample and camera coupled with HHG CDI will allow full 3D tomographic imaging of samples in transmission and reflection, as well as EUV reflectometry that will provide highly sensitive, spatially localized, non-destructive quantification of nanoscale samples structure and composition. Such

Figure 6 | Tabletop-scale coherent "x-ray laser" light source, the KMLabs Wyvern™ Ultrafast laser driving the KMLabs $XUUS_4$ high-harmonic light source. The EUV beamline can be seen in the upper right. (www.kmlabs.com)

quantification will make possible, for example, unprecedented measurements of trace elements buried deeply within semiconductor samples. When combined with stroboscopic, pump-probe imaging that makes use of the ultrafast pulses naturally provided by HHG sources, we will be able to image femtosecond dynamics of nanoscale charge, heat, and spin transfer in 3D, functioning, real-world devices of interest.

Outlook

The combination of high harmonic generation and ptychographic coherent diffraction imaging is enabling an unprecedented view of the nano-world. The ability to non-destructively image thick objects with chemical specificity and sub-nanometer depth resolution, generate quantitative height and reflectivity/transmissivity images that provide complete information about samples' topography and composition, and probe dynamic phenomena with femtosecond time resolution in three-dimensions gives HHG-CDI an exciting future. CDI stands alone as a non-destructive, long working distance, high-contrast, diffraction limited resolution imaging technique that offers unique and complementary information to well established techniques, with a host of potential applications. As the technology develops with increasingly robust illumination sources (see Figure 6) and reconstruction algorithms, as well as an ever-decreasing tabletop form-factor, commercial systems are quickly approaching. **Q**

Henry Kapteyn *and* ***Margaret Murnane*** *are Professors in the Department of Physics at the University of Colorado, Boulder, and are also fellows of JILA and co-founders of Kapteyn-Murnane Laboratories Inc. (KMLabs). Henry was educated at Harvey Mudd College (BS), Princeton University (MA), and UC Berkeley (PhD 1989), while Margaret received BS and MS degrees from University College Cork, Ireland, and PhD (1989) from UC Berkeley. Christina Porter received her BA degree from Reed College, and is currently a PhD student in the Kapteyn/Murnane research group. Dennis Gardner, Elizabeth Shanblatt, and Michael Tanksalvala are students in the K/M group, while Drs. Daniel Adams and Giulia Mancini are Scientists in the K/M group. Dr. Xiaoshi Zhang is the product manager and leader for the XUUS product at KMLabs Inc..*

References

1. JN Cederquist, JR Fienup, JC Marron, and RG Paxman, "PHASE RETRIEVAL FROM EXPERIMENTAL FAR-FIELD SPECKLE DATA," Optics Letters 13(8), 619-621 (1988). dx.doi.org/10.1364/ol.13.000619
2. JW Miao, P Charalambous, J Kirz, and D Sayre, "Extending the methodology of X-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens," Nature 400(6742), 342-344 (1999). dx.doi.org/10.1038/22498
3. MD Seaberg, DE Adams, EL Townsend, DA Raymondson, WF Schlotter, Y Liu, CS Menoni, L Rong, C-C Chen, J Miao, HC Kapteyn, and MM Murnane, "Ultrahigh 22 nm resolution coherent diffractive imaging using a desktop 13 nm high harmonic source," Optics Express 19(23), 22470-22479 (2011). dx.doi.org/10.1364/oe.19.022470
4. J Miao, T Ishikawa, IK Robinson, and MM Murnane, "Beyond crystallography: Diffractive imaging using coherent x-ray light sources," Science 348(6234), 530-535 (2015). dx.doi.org/10.1126/science.aaa1394
5. A Rundquist, CG Durfee, ZH Chang, C Herne, S Backus, MM Murnane, and HC Kapteyn, "Phase-matched generation of coherent soft X-rays," Science 280(5368), 1412-1415 (1998).
6. HC Kapteyn, MM Murnane, and IP Christov, "Extreme Nonlinear Optics: Coherent X Rays from Lasers," Physics Today 58(3), 39-44 (2005).
7. T Popmintchev, MC Chen, P Arpin, MM Murnane, and HC Kapteyn, "The attosecond nonlinear optics of bright coherent X-ray generation," Nature Photonics 4(12), 822-832 (2010). dx.doi.org/10.1038/nphoton.2010.256
8. T Popmintchev, M-C Chen, D Popmintchev, P Arpin, S Brown, S Ališauskas, G Andriukaitis, T Balčiunas, OD Mücke, A Pugzlys, A Baltuška, B Shim, SE Schrauth, A Gaeta, C Hernández-García, L Plaja, A Becker, A Jaron-Becker, MM Murnane, and HC Kapteyn, "Bright Coherent Ultrahigh Harmonics in the keV X-ray Regime from Mid-Infrared Femtosecond Lasers," Science 336(6086), 1287-1291 (2012). dx.doi.org/10.1126/science.1218497
9. A McPherson, G Gibson, H Jara, U Johann, TS Luk, IA McIntyre, K Boyer, and CK Rhodes, "Studies of multiphoton production of vacuum-ultraviolet radiation in the rare gasses," Journal of the Optical Society of America B 4, 595-601 (1987).
10. M Ferray, A Lhuillier, XF Li, LA Lompre, G Mainfray, and C Manus, "Multiple-Harmonic Conversion of 1064-Nm Radiation in Rare-Gases," Journal of Physics B-Atomic Molecular and Optical Physics 21(3), L31-L35 (1988).
11. O Kfir, P Grychtol, E Turgut, R Knut, D Zusin, D Popmintchev, T Popmintchev, H Nembach, JM Shaw, A Fleischer, H Kapteyn, M Murnane, and O Cohen, "Generation of bright phase-matched circularly-polarized extreme ultraviolet high harmonics," Nature Photonics 9(2), 99-105 (2015). dx.doi.org/10.1038/nphoton.2014.293
12. T Fan, P Grychtol, R Knut, C Hernandez-Garcia, DD Hickstein, D Zusin, C Gentry, FJ Dollar, CA Mancuso, CW Hogle, O Kfir, D Legut, K Carva, JL Ellis, KM Dorney, C Chen, OG Shpyrko, EE Fullerton, O Cohen, PM Oppeneer, DB Milosevic, A Becker, AA Jaron-Becker, T Popmintchev, MM Murnane, and HC Kapteyn, "Bright circularly polarized soft X-ray high harmonics for X-ray magnetic circular dichroism," Proceedings of the National Academy of Sciences of the United States of America 112(46), 14206-14211 (2015).
13. RL Sandberg, A Paul, DA Raymondson, S Hadrich, DM Gaudiosi, J Holtsnider, RaI Tobey, O Cohen, MM Murnane, HC Kapteyn, C Song, J Miao, Y Liu, and F Salmassi, "Lensless Diffractive Imaging Using Tabletop Coherent High-Harmonic Soft-X-Ray Beams," Physical Review Letters 99(9), 098103-098104 (2007).
14. B Zhang, DF Gardner, MD Seaberg, ER Shanblatt, HC Kapteyn, MM Murnane, and DE Adams, "High contrast 3D imaging of surfaces near the wavelength limit using tabletop EUV ptychography," Ultramicroscopy 158, 98-104 (2015). dx.doi.org/10.1016/j.ultramic.2015.07.006
15. ER Shanblatt, CL Porter, DF Gardner, GF Mancini, RM Karl, MD Tanksalvala, CS Bevis, VH Vartanian, HC Kapteyn, DE Adams, and MM Murnane, "Quantitative Chemically-Specific Coherent Diffractive Imaging of Reactions at Buried Interfaces with Few-Nanometer Precision," Nano Letters 10.1021/acs.nanolett.6b01864(2016). dx.doi.org/10.1021/acs.nanolett.6b01864
16. M Tanksalvala, DF Gardner, GF Mancini, ER Shanblatt, X Zhang, BR Galloway, CL Porter, and R Karl, "Coherent ptychographic imaging microscope with 17.5nm spatial resolution employing 13.5 nm high harmonic light," Microscopy and Microanalysis 22(Supplement 3), 88-89 (2016).
17. JM Rodenburg and HML Faulkner, "A phase retrieval algorithm for shifting illumination," Applied Physics Letters 85(20), 4795-4797 (2004). dx.doi.org/10.1063/1.1823034
18. AM Maiden and J. M. Rodenburg, "An improved ptychographical phase retrieval algorithm for diffractive imaging," Ultramicroscopy 109, 1256–1262 (2009).
19. P Thibault, M Dierolf, O Bunk, A Menzel, and F Pfeiffer, "Probe retrieval in ptychographic coherent diffractive imaging," Ultramicroscopy 109(4), 338-343 (2009). dx.doi.org/10.1016/j.ultramic.2008.12.011
20. F Zhang, I Peterson, J Vila-Comamala, ADF Berenguer, R Bean, B Chen, A Menzel, IK Robinson, and JM Rodenburg, "Translation position determination in ptychographic coherent diffraction imaging," Optics Express 21(11), 13592-13606 (2013). dx.doi.org/10.1364/oe.21.013592
21. P Thibault and A Menzel, "Reconstructing state mixtures from diffraction measurements," Nature 494(7435), 68-71 (2013). dx.doi.org/10.1038/nature11806
22. DJ Batey, D Claus, and JM Rodenburg, "Information multiplexing in ptychography," Ultramicroscopy 138, 13-21 (2014). dx.doi.org/10.1016/j.ultramic.2013.12.003
23. TM Godden, R Suman, MJ Humphry, JM Rodenburg, and AM Maiden, "Ptychographic microscope for three-dimensional imaging," Optics Express 22(10), 12513-12523 (2014). dx.doi.org/10.1364/oe.22.012513

Light-field Display for Collaboration and Situational Awareness

by Amy Lessner and Thomas Burnet

Industry has developed software, hardware, and sensor technology to capture 3D scenes and to provide for a complete and accurate representation of the information. Now there is a growing requirement to improve the ability for natural, collaborative visualization, and interaction with the data. Improved human-machine interfaces must allow for the fusion of various types of information (software, sensor, etc.) to create one holistic view of 3D information to improve situational awareness, reduce cognitive load, and increase trust in the data being displayed. The ability to rapidly sort through and intuitively visualize information in 3D will improve the decision maker's capability to act with precision, certainty, and expediency.

Human performance has long been associated with the ability to visualize and grasp 3D concepts. Human binocular vision and acuity, and the accompanying 3D retinal processing of the human eye and brain are specifically designed to understand and perceive the natural 3D world. The ability to resolve depth within a scene, whether natural or artificial, improves our spatial understanding of that environment.

Computers and 3D rendering onto 2D displays have greatly enhanced the human ability to visualize complex scenes from a single perspective; however, when 3D information is projected in 2D, many of the natural depth cues are lost or incorrect. Occlusion, specular highlights, and gradient shading must be perspective correct for each eye for the human visual system to work naturally, otherwise the scene is in conflict with what the human visual system expects.

A light-field display projects 3D imagery that is visible to the unaided eye (without glasses or head tracking) and allows for perspective correct visualization within the display's projection volume. A light-field table provides full 360° viewing with precise spatial registration, independent of the location of multiple simultaneous viewers. All of the expected depth cues are provided to recreate the 3D visual as one would see in the natural real-world. As a result, the viewers quickly and naturally understand complex spatial relationships between objects. This allows for faster, richer comprehension of the data. This type of display has been portrayed for decades in movies such as Star Wars, Prometheus, and Avatar. It has long been the centerpiece of war rooms, command centers, medical facilities, laboratories, and entertainment arenas

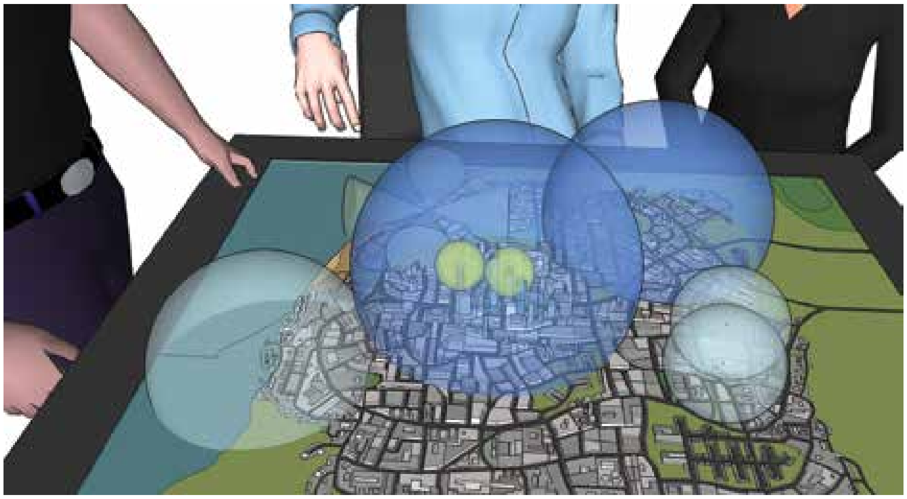


Figure 1 | Light-field Situational Awareness Display

in the world of science fiction; however, the light-field display is no longer just a science fiction concept. An Austin-based company, Zebra Imaging, has produced the first ever light-field display table to bring the sci-fi dreams into a reality.

There are a variety of approaches and considerations in developing a light-field display. Zebra Imaging's architecture is based on four major subsystems, that when combined, produce a convincing dynamic light-field image that enables natural perception and interaction. The computation and distribution subsystems are responsible for converting the 3D data (point clouds, triangles, etc.) to pixel information which is sent to the modulation subsystem. Here the information is converted to light rays which are angularly distributed through a micro-lens array to fill a 90° light-field visualization frustum.[1]

Light-field Visualization Performance Studies

Studies have been conducted in recent years exemplifying the benefits of 3D light-field visualizations for mission planning, training, and rehearsal. In 2008, an experiment was conducted by the Air Force Research Laboratory to assess human performance in identifying patterns or subtle targets, comparing the effectiveness of 2D and 3D static

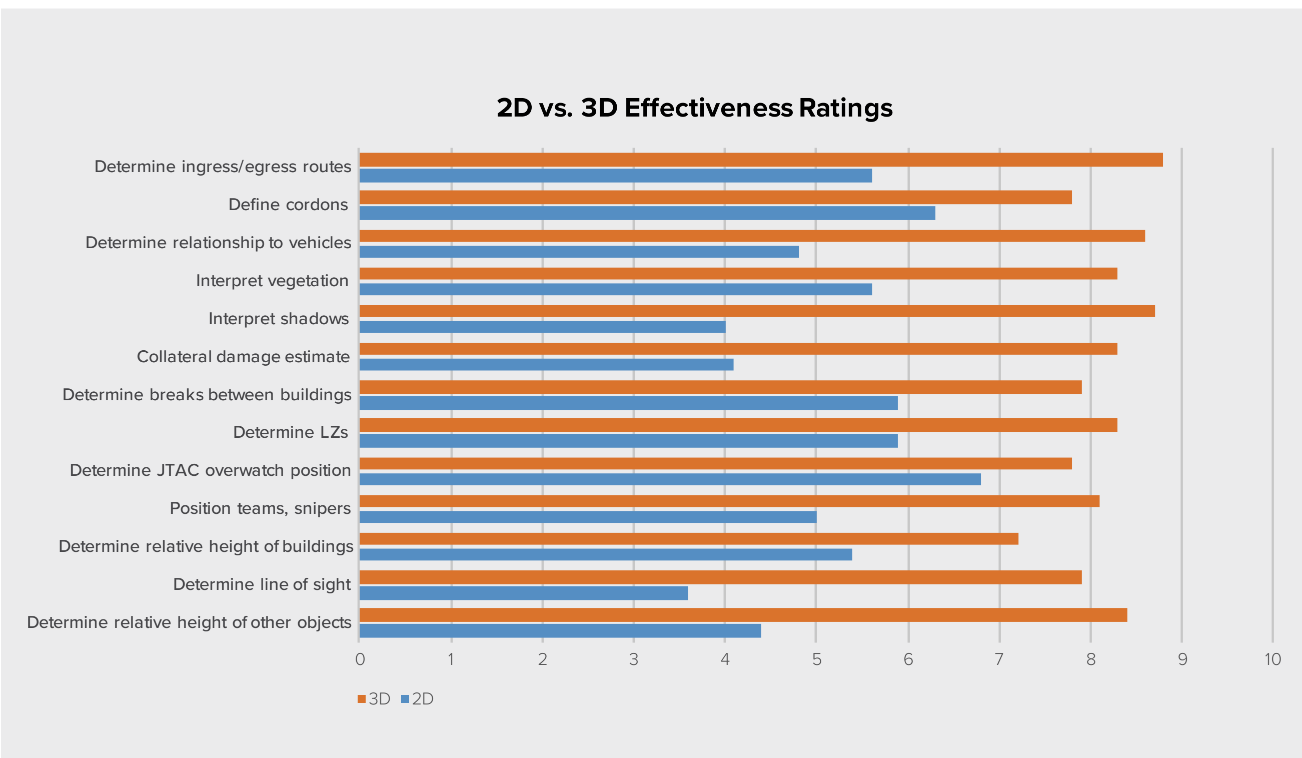


Figure 2 | 2D vs. 3D Effectiveness Rating by JTAC visual stimuli based on identical example data sets. The experiments were conducted using high resolution 3D light-field hologram prints provided by Zebra Imaging. These static print holograms are a reflective pixel grating imaged into a photopolymer. By shining a point light source onto the photopolymer a 3D visualization is projected.

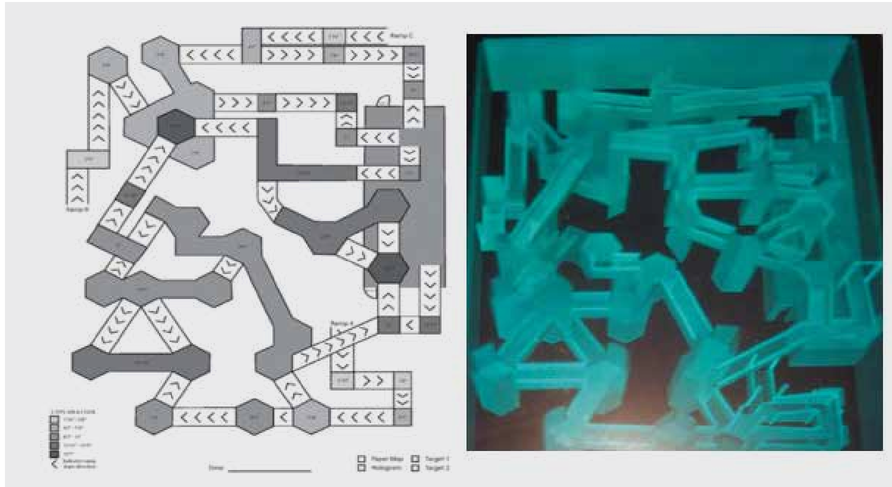


Figure 3 | L. 2D representation of the AOI. R. Hologram of same AOI.

Nine experienced Joint Terminal Attack Controllers (JTACs) participated in the evaluation. They were provided a written scenario similar to what is used both in combat and in training. The JTACs were shown conventional 2D imagery and light-field images relating to the scenario. A detailed survey of the nine JTACS rated the 3D light-field images as more effective than the 2D photos for all mission planning and execution tasks and subtasks. The most frequently cited benefits of the print holograms were relative height information, enhanced collateral damage estimation, and determining lines of sight and lines of fire (Figure 4).

In the paper, *Investigating Geospatial Holograms for Special Weapons and Tactics Teams* (Sven Fuhrmann, *et al.*), the researchers investigated memory retention and mission success by comparing the accomplishments of two SWAT teams engaged in a hostage rescue scenario. One team developed a mission profile and rehearsed the mission plan using standard imagery and topographic maps. The second team developed their mission using holographic topographic maps. The two teams were ranked on the time required to accomplish their mission objectives. The research found that mission planning with the 3D light-field visuals increased mission performance and wayfinding by roughly 30%.

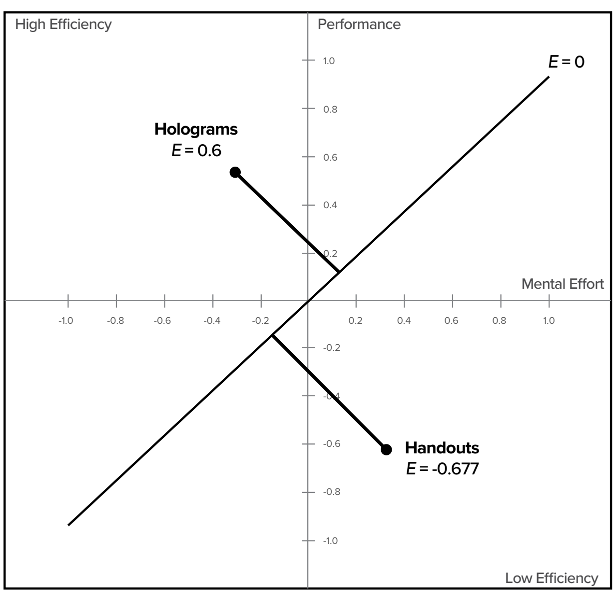


Figure 4 | Efficiency of Medical Hologram and Textbook Handout Conditions (Hackett, 2013)

A similar study was conducted by Matthew Hackett at the Army Simulation and Training Technology Center in Orlando, Florida. Mr. Hackett investigated the effectiveness of training medical students using 3D holograms over traditional textbooks. The study involved measuring the memory retention and congnitive load between the two groups of medical students. The conclusion of the study showed a significant performance improvement using light-field holograms over traditional textbook imagery (Figure 4). One of the reasons cited in the paper was the reduced congnitive load required to conceptualize complex 3D relationships (Hackett, 2013).

Light-field Display Development History

Inspired by the holographic topographic light-field prints, the Defense Advanced Research Projects Agency (DARPA) launched the *Urban Photonic Sandtable Display (UPSD)* program with the goal of creating a scalable, interactive, glasses-free holographic display to assist team-based mission planning, visualization, and interpretation of complex 3D data such as intelligence and medical imagery.

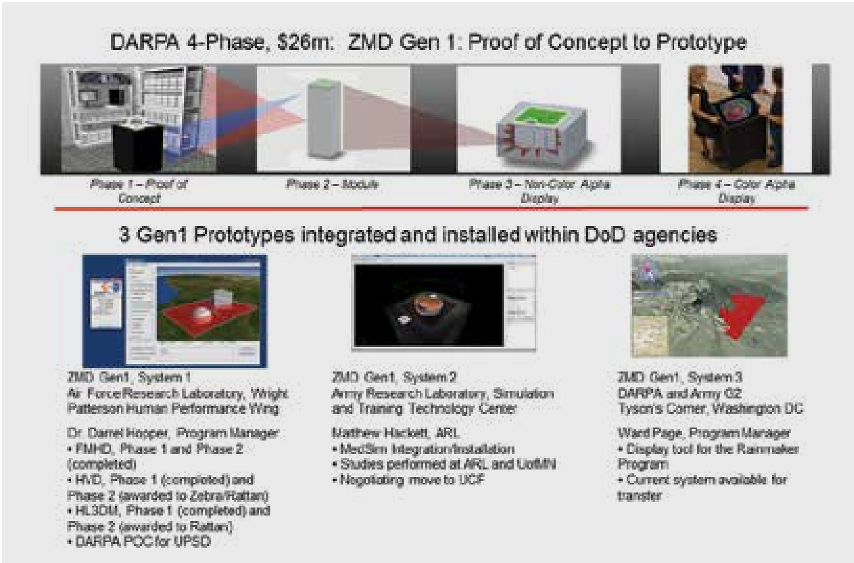


Figure 5 | DARPA 4-Phase Program and Prototype Insertion program

This program was awarded to Zebra Imaging who completed the four-phase, five-year effort successfully and created four 21" prototype displays with three systems transitioned to agencies within the Department of Defense.

The Air Force Research Laboratory Program, *Full Multiplex Holographic Display* (FMHD), integrated one of the UPSD prototype displays with NASA's World Wind, a Google Earth type of application for visualizing terrains in 3D. The second display transitioned to the Army Research Laboratory under the program, *Hologram Imaging Technology (HIT) Applied to Medical Modeling and Simulation*, with the objective of investigating the use of holographic and 3D imaging for medical simulation and training. This effort resulted in the above-referenced paper, *Medical Holography for Basic Anatomy Training* by Matthew Hackett, which was awarded Best Paper at the 2013 Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC). His paper concluded that the use of perspective correct holographic images of an anatomical heart model resulted in higher memory retention for medical students. The third system was transitioned to the Army G2, integrated with *Zones of Protection* (ZoP), and installed in a Sensitive Compartmented Information Facility in Washington, D.C. ZoP integrates real-time, real-world intelligence, and operational data with a precisely accurate synthetic environment. The UPSD prototype was a natural visualization tool for this application as collaboration was an inherent feature of ZoP.

Next Generation Light-field Display Development

Zebra Imaging has created a separate business unit, Rattan Software, with the sole purpose of developing the next generation Light-field Display. Built on the proven architecture of the UPSD prototype system, Rattan Software's roadmap supports the development of three display systems: monochrome light-field display, color light-field display, and the compact color light-field display system with integrated custom computation. The three strategic off-ramps allow for integration and socialization beginning in 2018 with a monochromatic system designed for early adopters such as the military, medical institutions, and training centers. Aggressive research and development at the component level continues to produce significant results including a breakthrough on the optical layer that takes the system unit cost down by a factor of almost 10x, making it potentially ideal for commercial deployment, and insights on alternative processing chip architecture will be valuable not just for light field displays, but also for the VR and AR industries as well.

Rattan is currently building relationships within the first two verticals poised for early adoption: the military and the medical/educational community. Early involvement with these two groups helps transition the pre-production prototypes to manufacturing while ensuring each vertical's requirements are captured and addressed. The system is designed for manufacturability, and that vendor and component sourcing drives down the overall price of the system.

There is active participation from many agencies within the DoD with the common objective of creating the best, most intuitive visualization capability for their area of interest. Small Business Innovation Research funds from the Army, Navy, and Air Force, and are augmenting the development of Rattan's LfD roadmap which was restarted after the initial DARPA investment with private seed funding.

Zebra Imaging and Rattan Software are in active conversations with commercial partners in the healthcare market to accelerate the development schedule and implement a robust go-to-market plan. These discussions also help to validate the requirements for such applications as pre-operative planning and post-operative review; education, training and simulation; medical device training and sales, and eventually diagnostics. In the majority of these use cases the 3D data is already available, such as computed tomography, magnetic resonance imaging, echocardiography, ultrasound, etc. With a light-field display, the ability to visualize patient specific data to understand the uniqueness of a patient's blood vessels, heart valves, and the troughs and crests of the human brain could significantly reduce surgical time, recovery time, and unnecessary procedures.

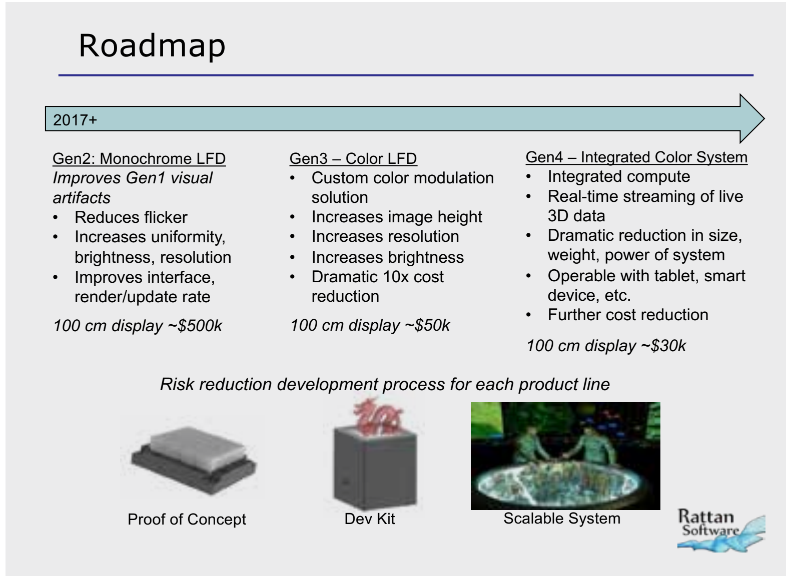


Figure 6 | Rattan Software's Light-filed Display Roadmap

The light-field display table is the future. Initially, it will be the centerpiece of every joint intelligence operations center providing the common view of the dynamic environment. It will be in research and teaching hospitals to facilitate a more intuitive understanding of the human body. As price is driven down and capability increased a light-field table will provide a way to validate a design before incurring the expense and time associated with advanced manufacturing. The oil and gas community can visualize seismic data to reduce the millions of dollars it costs to operate a platform in the search for resources. And, in time, families and friends will gather around a light-field display in their home or at a bar or restaurant to watch sporting events, concerts, movies/plays, and current events as they unfold. The information is there, all that's left is a display system that brings people together and promotes collaboration. **Q**

Amy Lessner *is the Director of Operations at Rattan Software, where she directs the next generation light-field display roadmap. Previously, Amy was the Zebra Imaging Program Administrator for the DARPA Urban Photonic Sandtable Display Program and the Program Manager for Zebra's prototype transitions to the Air Force Research Laboratory and the U.S. Army Research Laboratory.*

Thomas Burnett *is Rattan Software's Chief Technology Officer and is the Principal Investigator for the next generation light-field display. Thomas was also the software architect of the first ever light-field display at Zebra Imaging, where he developed the software that created the Tactical Digital Holograms under the US Army's Tactical Battlefield Visualization Program. This effort launched the DARPA Urban Photonic Sandtable Display Program.*

Reference

1. A frustum is the portion of a cone or pyramid that remains after its upper part has been cut off by a plane parallel to its base, or that is intercepted between two such planes.

Intrinsic Imaging™

By Scott Ackerson and Don Meagher

Imagine being able to acquire a true 3D (volumetric) image using natural illumination, from a single view and from long distances (meters to kilometers). Imagine being able to capture volumetric images of real scenes that include shiny and translucent objects. By "understanding" how a polarized lightfield physically interacts with the objects in a scene, Intrinsic Imaging is turning these imaginings into reality. This article describes Intrinsic Imaging, a powerful new approach to 3D imaging, including problems that it can resolve and applications that it can address.

Background

3D Imagers are devices that use images to reconstruct digital 3D models[1] of objects and scenes (hereinafter "scenes"). They are comprised of sensors and Scene Reconstruction Engines ("Engines"). Sensors are classified as time-of-flight, stereoscopic, stereo correspondence, coherent or polarimetric. Engines are classified by denseness (how sparse or dense is the reconstructed model) and isotropy (Lambertian or non-Lambertian — see Terminology for additional explanation of these terms).

Problem

The spaces that people occupy in the course of their daily lives ("Quotidian Scenes") have the following properties:

- They are volumetrically dense (e.g., they contain bushes, windows, occluding objects, etc.).
- They extend "as far as the eye can see" (e.g., a room with a door opening to a backyard and to mountains in the distance, is a quotidian scene).
- They include glossy, shiny, translucent (highly "non-Lambertian") objects.
- They include featureless surfaces (such as large white walls).
- They include participative media (such as water, glass and fog).

Terminology

Isotropic means the same in all directions. Anisotropic means the opposite.

Lambertian means matte or diffusely reflecting, as in a "matte" surface. Non-Lambertian means the opposite. An unfinished block of wood is roughly Lambertian. The same block, finished with polyurethane, is non-Lambertian.

Light means electromagnetic energy in visible, infrared and ultraviolet wavelengths.

Lightfield means light flowing in every direction through every point in a scene.

Media means matter occupying a volumetric element of space through which light propagates or with which light interacts. Media can be homogeneous, heterogeneous or void (devoid of matter).

Quotidian means "everyday." Quotidian scenes are spaces that people occupy in the course of their daily lives, including kitchens, offices, jungles and space capsules. Quotidian scenes (i) extend "as far as the eye can see," (ii) include non-Lambertian (shiny, translucent) objects, and (iii) are generally cluttered with objects that are occluded (blocked) in certain views.

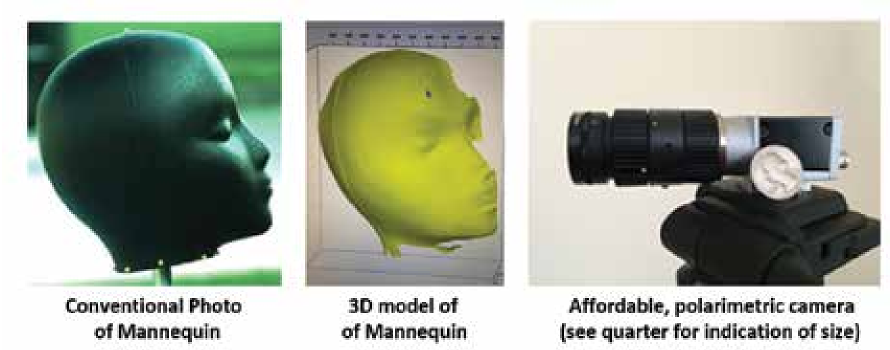


Figure 1 | Actual Intrinsic Image of Mannequin (Passive, Single View, Outdoors, Long Distance)

Yet, today's 3D Imagers typically do not handle these properties well, virtually guaranteeing that they will not "see" very well in three-dimensions.

Solution

Quidient developers are building 3D imaging systems based on Intrinsic Imaging technology that address the problem stated above. Sensors used in the systems include polarimetric cameras. The Scene Reconstruction Engines (called Quidient Locale Engines™) are polarimetrically-intensified, dense volumetric and non-Lambertian. The systems are completely passive. They are theoretically capable of reconstructing dense volumetric models of quotidian scenes to high accuracies. Quidient Locale has already been used with polarimetric cameras (including cameras from a preferred sensor partner, Photon-X, Inc.) to reconstruct simple scenes. For example, Figure 1 shows the volumetric reconstruction of a mannequin performed using a single image from a polarimetric camera. Note that the dark surface of the mannequin is intentionally featureless, which presents a challenge for many other 3D imaging approaches. The mannequin was located outdoors in natural sunlight at a distance of about 4 meters. The model was created automatically (e.g., no human editing/cleanup) and is globally accurate to a few millimeters.

Technology

Quidient Scene Recognition Engines employ three key technologies:

Intrinsic SLAM™. Simultaneous Localization and Mapping ("SLAM") is an algorithmic method for building a digital model of a scene from images, while simultaneously determining the location of (localizing) moving cameras used to acquire the images. (While Quidient Engines reconstruct 3D surfaces from single polarimetric images, multi-view imaging is used to create more complex scene reconstructions.) Most SLAM methods today solve for sparse (point-like) features that are Lambertian and are located in Lambertian regions. In contrast, Intrinsic SLAM solves for media elements ("mediels") that are non-Lambertian and are located in non Lambertian regions illuminated by a non-isotropic scene lightfield. By modeling the scattering characteristics of non-Lambertian mediels in a polarimetric lightfield, Intrinsic SLAM is made more solvable and tractable than conventional SLAM.

Referring to Figure 2, which is a top view of a theoretical scene, a polarimetric camera samples the polarimetric lightfield as it moves through the default media (e.g. air, fog, water, space) in the scene. The camera path is indicated by the black vectors depicted in the figure. Default media is represented by the pink gradient. The green and red (fuzzy) vectors on the path represent the lightfield captured by the camera as it progresses along the path. Mediels, including surface elements ("Surfels"), are reconstructed in observed regions as the camera moves through the scene. Objects of Interest ("OOIs") can be identified by the operator for particular focus of attention. Unexplained regions (for which no object model explains observations) can be effectively rendered by retaining lightfield information, but cannot be dimensioned.

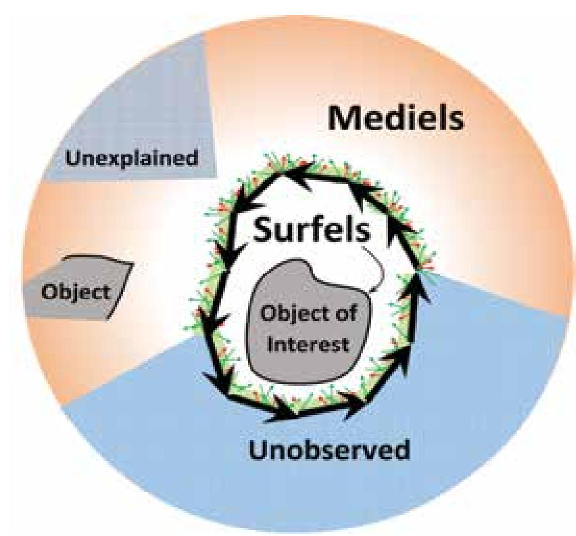


Figure 2 | Intrinsic Imaging

Spatial Processing. Modeling the media comprising a quotidian scene from an image requires a substantial amount of memory and computing. The challenge is compounded when anisotropic lightfields and non-Lambertian light-media interaction functions are also modeled. Extracting and updating all this information based on a series of polarimetric images requires even more computational power and resources. Special volumetric and multi-resolution data structures and high-speed algorithms have been developed to make such computations tractable.

Along another computational dimension, some applications place a high value on throughput and responsiveness. Real-time or near real-time operation is a major goal while maintaining real-world complexity. To achieve this within cost, power, and other constraints, methods have been derived for performing the required computations over a large number of simple hardware processors. Quidient Locale Engines make extensive use of various hardware acceleration methods including multi-core processing, Field-Programmable Gate Array and Graphical Processing Unit technology to accelerate computations.

Imaging Polarimetry. Depending on the application, Quidient Scene Recognition Engines operate with various types of polarimetric imagers, including conventional, multi-spectral, visible and infrared cameras.

To further explain the nature and value of imaging polarimeters, there are three basic characteristics of natural light: frequency, intensity and polarization. Humans see the first two characteristics as color and brightness. Since humans cannot see polarization (the way that the transverse fields of propagating light are oriented) with their unaided eyes, we tend to overlook the information it carries. But, it turns out that when light reflects from a surface, its polarization is affected in such a way that the shape of an object can be directly sensed. Shape from polarization, while complex, can be understood by considering the simple diffuse case.

Referring to Figure 3, when natural background light enters a diffuse surface, it "bounces around" in the subsurface molecules, "forgets where it came from" (becomes depolarized) and then exits the surface. When it exits the surface, it is partially polarized as a function of material and surface orientation in accordance with Fresnel's Law. This effect on polarization is represented by the violet ellipses in Figure 3, where the shape and orientation of

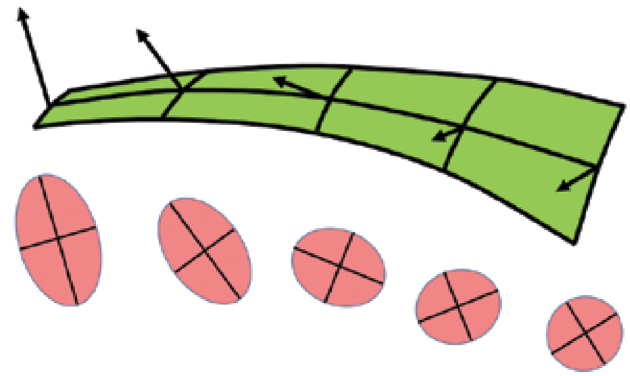


Figure 3 | Shape from Polarization, Diffuse Case

the ellipses represents the magnitude and orientation of the linearly polarized light reflecting from the surface relative to the viewing axis of the camera. In the figure, it can be seen that at larger angles between the surface normal and the camera axis, the light is more strongly polarized than the smaller angles that are more nearly perpendicular to the camera axis. A measurement of the polarization angles, as viewed by the camera, will yield the orientation of surface elements and therefore the shape of the surface can be directly measured.

Applications

Applications for Intrinsic Imaging can be classified under the categories of 3D Visualization or 3D Scanning (everything other than video). The purpose of 3D Visualization applications is to create notions of a 3D scene in the mind of a person viewing a display. The purpose of 3D Scanning applications is to extract information about a scene for purposes such as navigation, situational awareness and metrology.

People often think of 3D Visualization applications when they hear the term "3D Imaging" from seeing 3D movies like Avatar, but it is nearly certain that it will be the 3D Scanning applications that enrich the lives of people the most over the next decade. Life-improving, new-category applications such as "shape-based shopping" (scan your feet to order shoes), "Cognitive Inspection™" (scan your roof to assess storm damage by tossing a disposable UAV into the air), "Situational Awareness" (my child just fell in my pool), and "3D Endoscopy" (immersive 3D in surgical scenes through a thin endoscopes) will soon be available to everyone.

It is theoretically possible that Intrinsic Imaging will unify 3D Visualization and 3D scanning applications.

First "3D Images"

3D imaging (creating 3D models of a scene by sensing light) and 3D display (creating notions of a real scene in the mind of a viewer by projecting light into the viewer's eyes) are often confused. 3D displays can be conventional (Lambertian screen), holographic (non-Lambertian screen) or volumetric. Conventional displays can create strong notions of three-dimensionality by using depth cues such as perspective, light and shadowing. One can argue that the first people on earth to create 3D images (loosely defined – see Footnote 1) were Renaissance painters, who used cues to create strong notions of depth in their paintings.

Platforms

Scene Recognition Engines can be embedded on physical platforms such as UAVs, automobiles, people, mobile devices, digital cameras, and appliances.

Figure 4 | Intrinsic Imaging Platforms

Products

Referring to Figure 5, Quidient's flagship product is its Locale Engine (lower left), which is a Scene Reconstruction Engine based on its Intrinsic Imaging technology. In a networked 3D Imaging System, a master app (top center) is typically provided along with one or more sensors (lower right). The preferred sensor inventory includes one or more polarimetric cameras. The app can be Quidient Locale Pro (depicted in Figure 6), or a customized app developed from Pro, or it can be written from scratch for a specific application. Files associated with Intrinsic Imaging

Market vertical	Key Applications
Defense	Serious gaming, simulation, planning, target recognition.
Digital Cameras	3DTV, 3D film, 3D photography, visualization.
Education	Communication, documentation, distance learning, affective computing.
Entertainment	3DTV, 3D film, 3D photography, animation, motion capture, toys-to-life, sports performance, social networking, virtual and augmented reality.
Healthcare	Remote medicine, prosthesis, orthotics, dental, plastic surgery, sleep apnea masks, endoscopy, mental health (digital nursing), training.
Infrastructure	Home walkthrough, bricolage, construction, building management, casualty inspection.
Manufacturing	Cognitive inspection™, robotics, process automation, quality control, freight logistics, design, corrosion control, food processing, training.
Mobile	Communication, documentation, virtual and augmented reality, gesture recognition.
Retail	Mass customization ("Personalization") of clothing, catalog websites, 3D body dimension sharing websites, digital signage.
Security	ATM access, biometrics based on 3D handprints and faces, forensics, soldier faces (documentation), child alert, vehicle access, passport control, border control, gesture recognition.
Visually Cognitive Devices	Appliances, cell phones, computers, tablets.

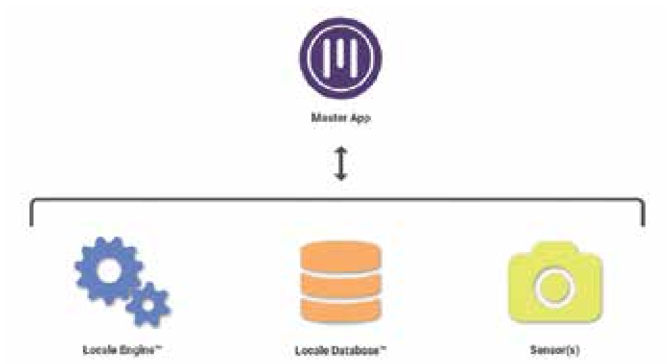


Figure 5 | Intrinsic Imaging Network Architecture

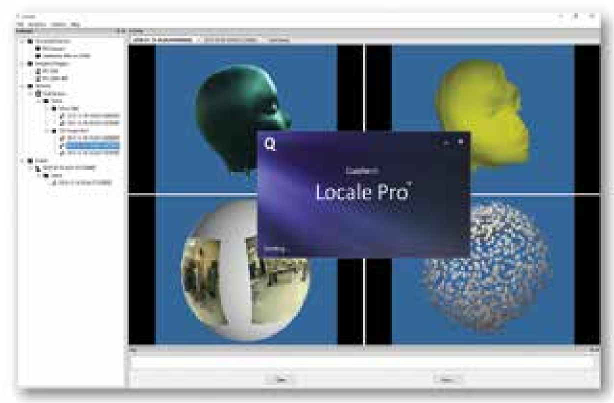


Figure 6 | Quidient Locale Pro™ Application

sessions are true databases (depicted bottom center of Figure 5) that include sophisticated spatial information.

We hope that Quidient partners will ultimately write hundreds of apps, some with very simple user interfaces designed to run on mobile phones or wearable devices. In the near-term, Locale Engines can be made available on private and public networks using an Engine-as-a-Service (EaaS) business model. Longer-term, Quidient expects to license Intrinsic Imaging technology for use under a System on a Chip (SoC) model.

Key features of Quidient Intrinsic Imaging Systems include:

- 3D via single aperture
- Dense volumetric models of quotidian scenes
- Natural, (i.e., passive) illumination
- Operation at all distances
- Capable of highly miniaturized form factor
- Conventional "camera-like" operation

Conclusion

Intrinsic Imaging is a powerful new 3D Imaging technology based on imaging polarimetry and advanced Scene Reconstruction Engine technology. Intrinsic Imaging systems may uniquely satisfy unserved needs for devices that "see well" in three dimensions, unifying 3D visualization and 3D scanning, and fueling continued growth of the global 3D Ecosystem. **Q**

Scott Ackerson *is Quidient's Chief Executive. Scott is an entrepreneur and inventor with deep experience in 3D imaging and metrology. Quidient is Scott's third product company (the others being SMX Corporation and Dimensional Photonics, Inc., which had successful exits with FARO Technologies (NASDAQ) and Danaher (NYSE), respectively). Scott holds ten patents. He has a master's degree in Business from Wharton/Penn and a bachelor's degree in Systems Engineering from the U.S. Naval Academy. Early in his career, Scott served as an Officer in the U.S. Submarine force.*

Donald Meagher, Ph.D. *is Quidient's Director of Spatial Modeling. Don has extensive experience in the development of efficient algorithms for handling, processing and exploiting extremely large spatial datasets and their implementation in software and specialized hardware for use in intelligence, medicine, biometrics, industrial laser scanning and other areas. He developed and patented the first commercial 3D medical workstation (later licensed to General Electric). Don attended Rensselaer Polytechnic Institute (RPI) and UC Berkeley, and holds a bachelor's degree, master's degree and Ph.D. in Electrical Engineering (EE) and a master's degree in Computer and Systems Engineering. His doctoral thesis was the derivation of "octree" technology to improve the efficiency of 3D processing.*

Reference

1. People sometime refer to "stereoscopic images" and "depth-cued conventional images" as "3D Images." But, they are more properly thought of as display methods used to create notions of 3D scenes in the minds of viewers. Stereoscopic images are 2x2D (two conventional images, one for each eye). Depth-cued images are 2D. To avoid confusion, we will use the term "volumetric" to mean "true" 3D.

2r1y High Speed High Density 3D Imaging

The Evolution of a Technology

By Richard Neumann

Every startup begins by trying to solve a problem. More often than not what they set out to solve does not end up becoming the final technology or product. Oftentimes, the process of getting there is more fascinating than the end result. As Laozi wrote in Tao Te Ching, "… a journey of a thousand miles begins with a single step", even the longest and most difficult ventures have a starting point.

I feel it's important to set the stage for what 2r1y is and what it isn't; 2r1y stands for 2 Russians and 1 Yank. The reality is they aren't really Russians, nor do they like being called that. Serge is a former Ukrainian and Vanda was born in Odessa, the daughter of an Armenian physicist. Both escaped the former Soviet Union in the late 1970s during the height of the Cold War. Their families arrived in the United States with nothing and have managed to achieve the American dream by working hard and building successful lives for themselves. I, the Yank, am a native Californian, the son of a rocket scientist and a Hollywood actress. Later joining the team is Igor, a coding genius; James, an optical guru; and Greg an FPGA master. 2r1y is entirely self-funded by the three founders. There is no large or small corporate affiliation, no grant money, angel or VC investment in our company. We operate out of a converted room in the back of Vanda's house in San Francisco.

The problem we set out to solve was how to provide augmented illumination to support gesture recognition in Smart TVs. At the time (2012), "smart" TVs were the hot item and gesture recognition was the biggest challenge. People generally watch TV in a darkened room. To capture human motion you need a high- resolution image, 720p or higher, and at a frame rate of at least 60fps. An economically viable consumer electronic solution requires the use of a commercial, off-the-shelf, complementary metal oxide semiconductor sensor. Getting that level of performance out of an image sensor requires a lot of light. Getting people to change their TV habits, watching with all the lights on, is not a feasible solution. Nor is putting a bright light source on top of the TV. Staring at a bright light would be a less than desirable entertainment experience.

The most obvious solution is to use non-visible light near-infrared (NIR) to provide augmented illumination. This solution comes with some significant human safety issues. With no ability to see IR, humans won't instinctively blink and rapid and irrevocable damage including blindness can occur. CMOS sensitivity over 700nm drops off rapidly, in most cases starting at 10% or less and falling to zero at 1000nm. The amount of augmented NIR illumination necessary in a living room sufficient to capture 720p @60fps is approximately 3,000 watts — far above the safe limit for humans. The challenge began as how to utilize a small, safe amount of NIR to capture human gestures.

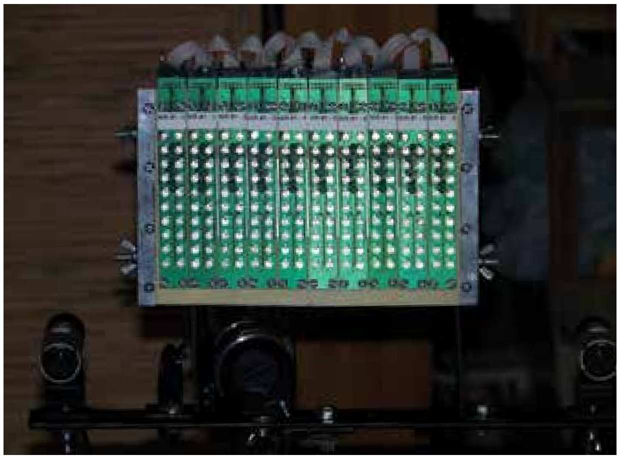


Figure 1 | 2r1y NIR LED regional illumination.

Figure 2 | 2r1y Initial regional illumination test, HD @30fps.

If you examine the operating environment, the region of interest (ROI) for gesture recognition software is only a person or smaller, like a hand or a face. This represents a fraction of the entire field of view (FOV). The problem was now becoming more specific: how to place light only on the ROI. To further complicate the problem, there can be multiple ROIs and those ROIs can and will move frame to frame.

Our first experiment was to see if we could break the FOV into regions and only illuminate those regions where the ROIs were found. The setup consisted of a bank of LEDs mounted on a semi-circular back plane. The LEDs were controlled by an Arduino processor. By pulsing the LEDs individually within the exposure time of a single frame, we could vary the intensity by region. This provided lower illumination for the entire FOV and maximum illumination to the ROI.

The results were crude, but encouraging. We could produce sufficient illumination to capture human motion at 30fps with a minimal amount of NIR.

We needed two things to improve the performance: The first was a more precise and faster way of placing light exactly where and when we needed it. The second was an image sensor of sufficient resolution, sensitivity, and frame rate to accurately capture the illumination.

Pointing the light in hindsight was actually the easier of the two processes. We found an off-the-shelf device which used a dual axis micro-electro mechanical system (MEMS) mirror and provided all of the interfaces we needed. Repurposing it seemed simple enough. All we had to do is replace their green laser with an NIR laser. As it turned out, it was a bit more complicated. The device used a Corning G-1000 frequency doubling laser to produce green, and unfortunately emitted more NIR than green light. Every step of their optical path was engineered to block or dump IR. With a lot of clever engineering, we bypassed all of the IR blocks to get an unobstructed circuitous path from the laser to the MEMS.

A much bigger challenge was the camera required to capture the illumination. We searched every supplier we could find and nowhere was there an open source camera that gave us access to the imaging processes that we needed. Naively we set out to build one ourselves, and based it on an OmniVision OV9714 sensor which has a 3 um architecture and a 28.7° well angle. This design produced about 2X to 3X the sensitivity to NIR as other COTS, red green green blue (RGGB) sensors we evaluated.

We hired a consulting firm to design the camera and the mezzanine circuitry that would control both illumination and image capture. For 18 months we had courted the top consumer electronics companies and convinced them to meet with us for a live demo. A few days before we were scheduled to give our first demo the consultants showed up with a condominium of evaluation and development boards that was nearly a foot cubed.

We realized the system wasn't going to work. No matter what we did, there was no way to get the camera working. The consultants informed us that it would take another week, maybe two to get the software working at a cost of $40,000 per week.

Rather than throw good money after bad, we tapped Igor's expertise. He planted himself cross-legged on the floor ignoring the confusion around him and began coding furiously while simultaneously Skyping with his wife in Germany and in the midst of a WOW (World of Warcraft) raid. He looked up and said, "It's Tuesday…it's Wednesday, it's Friday. Done!"

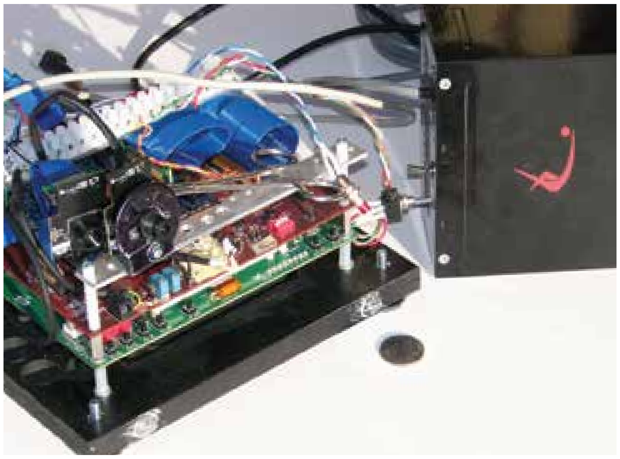


Figure 3 | Consultant-delivered "Mobile Phone" sized hardware, with a quarter for perspective. Repositioned camera and power supply (missing illumination device).

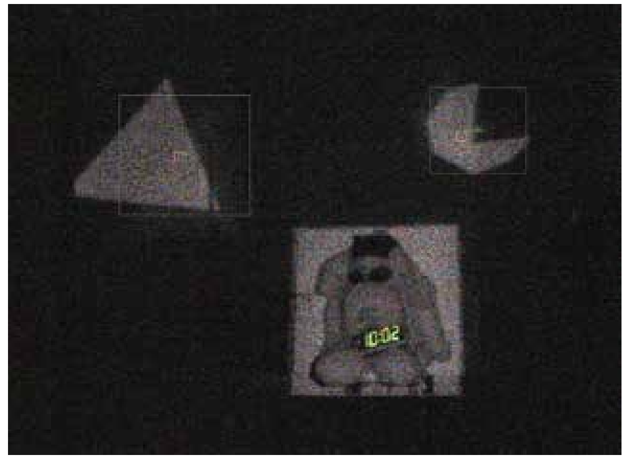


Figure 4 | Initial results of multiple ROI tracked illumination.

In the few hours remaining before our first demo, we were able to successfully demonstrate that we could use a small, safe level of NIR and illuminate very specific ROIs, conform the illumination to the ROI, and follow multiple ROI's independently. However, the live demo didn't fare as well. When we set up and turned on our system, there was a faint, audible click and a tiny whiff of smoke.

It was a memorable low point. Yet, we didn't give up.

There was enough to hint that we were on the right track. We could essentially paint any object in our system's FOV with a precise amount of IR illumination. And if the object moved, we could follow it and the illumination could morph to match changes in the object(s).

It was back to the whiteboard. The first thing we had to do was get a working camera, something less than a monumental Rube Goldberg conglomeration of dev kits, eval boards, and a custom mezzanine board which we had little clue as to what it actually did. And, a camera that wasn't pointing at the desktop would be a plus.

We fortunately found Greg, an FPGA/imaging master who informed us that our existing system was unsalvageable. He built a working camera platform in a few months, and coupled with Igor's code, we had a somewhat stable and repeatable platform with a high level of precision.

We were still fighting synchronization and our scanning illumination combined with the rolling shutter of the camera made it impossible to capture an image without rolling horizontal blackouts. At least we had the original mezzanine board and that had a synchronization circuit on it that we thought worked. After countless hours of effort, we had to finally concede that the last thing we thought was salvageable from the consultants was actually an expensive useless mess.

James had some prior experience with synch circuit design and stepped up to the challenge of designing a new circuit. A few weeks later we had a very nice compact and reliable circuit that kept the illumination and image sensor in perfect synchronization. The circuit picked up the saw tooth signal from the MEMS driver and produced a clean square wave which triggered each frame. The only way to access a synch signal was to solder a wire to a hair fine trace on a flex circuit. It took days of practicing under a microscope to make one connection.

With a synchronized image capture and illumination we had a stable system that ran at 60fps. The next thing we needed was to get the overall size down. We gave Greg the task of designing and building a small compact version of the camera — roughly a 45mm cube. We started with a blank whiteboard, and listed everything we had hoped a development camera would have had. We even included wild wishes like dual onboard microphones – something missing from scientific cameras. It only took 88 days to go from a blank whiteboard to a finished camera. We used a 720p image sensor capable of up to 120fps, synchronized and streaming uncompressed data over USB3.0. With all of the right pieces finally in place we built three "princesses", lab based systems that were stable, reliable, and repeatable.

We invited the top consumer electronic companies and tech investors to observe live demonstrations. They came and showed interest, but our key market had waned. The smart TVs had moved from the front of the booth at CES to the back.

Figure 5 | Synchronized MEMs signal and camera triggering square wave

Figure 6 | The Princess – illumination module (left), Gen 2 camera (right), control and synch circuity (base).

Seems consumers didn't want to wave or talk to their TV. Customers weren't going to surrender their remote.

Another low point, but not defeat. We had developed a technology that precisely places a 1mm spot of light into an environment. And, we could do that nearly half a million times in a given frame of exposure. Furthermore, we could give each point an 8-bit gradient in intensity, grey scale. So what could we do with this?

We realized we could produce scalable sinusoidal wave patterns. Which meant – we could use our technology to produce structured light, the basis for generating a point cloud.

We found open source code, that Igor later dubbed abandoned ware. It was basic and relatively stable, but drastically unfinished. Kyle McDonald had developed the OFX code as part of a tech-art project. They had abandoned the project when they failed to solve two key challenges, using non visible light, and synchronization, which are the two of the problems we had solved.

After several months we began to produce relatively high density point clouds. At 1280 x 800 @60fps, we beat PrimeSense/Microsoft's Kinect's 320 x 240 and we did it at twice the speed. And, because our system was dynamic, meaning we could adjust and scale the structured light to the ROI, we eliminated the problems with calibration.

Our results were vetted by VanGogh Imaging, and the initial results showed we had achieved a less than 2mm resolution point cloud in under 50ms. With the assistance of Laurence Hassebrook, Ph.D., Center for Visualization & Virtual Environments, University of Kentucky, we pushed that to less than 1mm resolution in under 400ms.

We continued to improve the process. Due to the unique use of a Bayer filtered CMOS sensor, we were also capturing

Figure 7 | Simple sinusoidal structured light pattern.

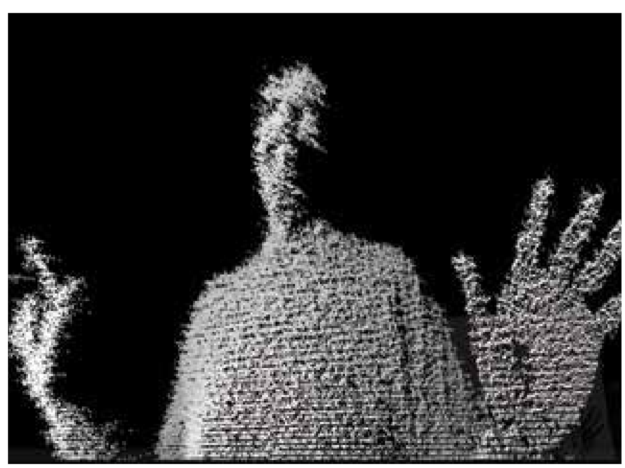


Figure 8 | First successful point cloud, range 2 meters, acquisition time 50ms.

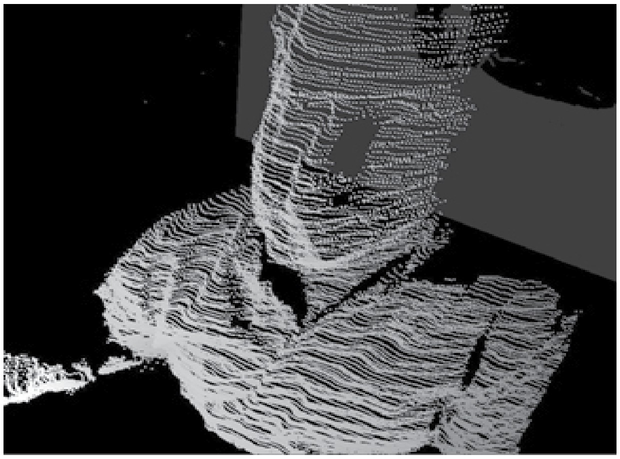


Figure 9 | Refined point cloud, note details on the shirt collar, resolution <2mm.

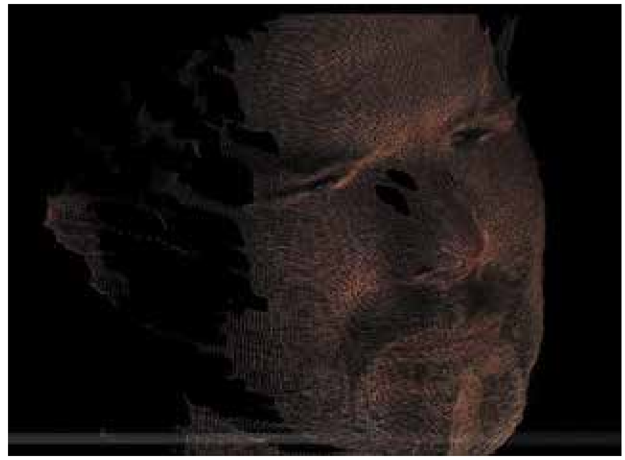


Figure 10 | SUB2r sub mm resolution, acquisition 400ms, RGB and Z.

color. When the IR cutout filter is removed from the lens, Bayer sensors oversaturate with ambient IR which results in a pinkish image that is difficult to correct. We added a narrow bandpass filter which cut the portion of IR between 700nm and our laser. We could now create a streaming point cloud in full color at 60fps.

One challenge still existed: point clouds are relative locations in space. There is no true dimension in a point cloud unless it is either calibrated or a separate mechanism is utilized to provide a measured distance. For a system to be truly adaptive and dynamic and keep to a single source of illumination with a single image sensor, there had to be a way of generating a true measured distance.

Back to the white board we went. We were able to find the exact distance to a single point projected into an environment. With considerable effort we extended that to three simultaneously. When we placed more spots into the FOV, the question quickly became: which spot was the camera seeing? Was it one that was originally targeted in an upper corner or was it one that was targeted closer and missed the ROI? We needed to measure not three, but thousands of points. This drove us to develop a methodology we call RAP – ray angle of probability. This methodology allows us to associate a projected spot with a captured spot. We also have the ability to correspond a measured spot within a point cloud. When these two methods are combined into a hybrid, the result is a high density point cloud which captures a measured Z11 and RGB12 at the same time. This brought our point cloud to a resolution of 1.3MP. Furthermore, it is theoretically scalable, which means that given a 5 or 10MP sensor and illumination source we could return a 10MP point cloud with measured Z and RGB.

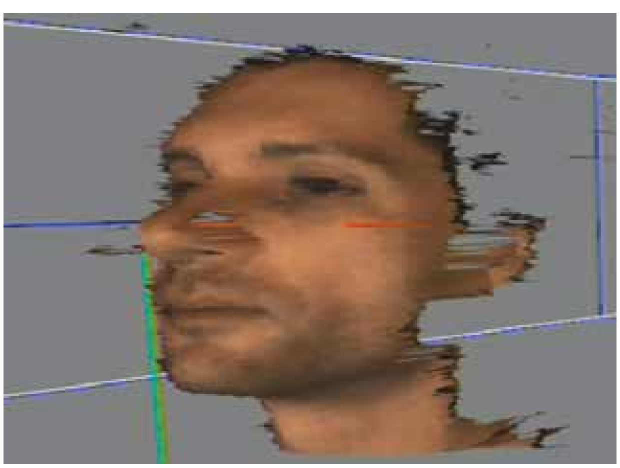


Figure 11 | Igor RGB point cloud, streaming at 60fps.

Figure 12 | Mia, full scan RGB point cloud.

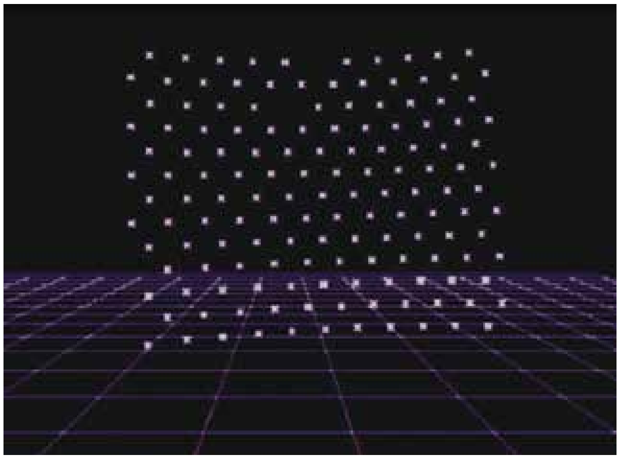


Figure 13 | Initial results of RAP.

Figure 14 | SUB2r open architecture camera platform.

And here is where the development hit its greatest challenge, funding. Without the resources to develop this into a commercialized technology further development would be limited by our ability to personally fund the project out of pocket.

It is the unexpected path that presents itself which oftentimes leads to the most interesting discoveries. The little camera we built had unique features such as good low light sensitivity, high frame rate, and the ability to access chip level registers. From time to time we posted videos of what the camera could do, both with our core technology and some other interesting applications. These applications ranged from sports analysis, to mounting it on a telescope, to integrating it with the virtual avatar streaming software FaceRig. As with all things on the internet, post something interesting on YouTube and sooner or later people will find you. And find us they did, by the thousands. We soon began getting requests for the camera. Finally, we decided we could no longer turn away potential revenue.

In 2015 we spun the camera off into its own company called SUB2r. We went back to the white board and redesigned the camera from scratch. Our goal was to build on our camera and develop an open architecture digital imaging platform. Key features we added were interchangeable sensor boards, interchangeable optics, 20 channels of expandable I/O on the camera board, onboard programmable FPGA, expandable memory, interlaced onboard and external audio, access to the AUX I/O, output over USB3 or GigE, onboard user controlled compression, and fully UVC 1.5 compliance and access to the on-chip registers.

We began with the goal of providing simple augmented illumination. We ended up with a hybrid, dynamic, high-density, high-speed method for depth mapping. And, we developed the first open architecture digital camera platform. We are continuing to develop the core tech while we commercialize the camera. We can't wait to see where we go from here. **Q**

__Richard Neumann__ is the CEO and Founder of 2r1y. Rich has over 30 years of experience in the high tech sector. He is formerly the founder/President of TriLumina, a VCSEL technology company, co-founder of Ravix Financial, a CFO consulting firm, and he has held senior management roles in Fibex (acquired by Cisco), Zaffire (acquired by Center Point), Micro Power Systems, Spectra Physics, and Amdahl. Rich holds a degree in business from San Jose State University.

3D Scanning and the Human Form

by Andrew Larkins

How well do your glasses fit you? Are the lenses in the correct position to maximize performance? This is a good example of an application where custom made products can really make a difference to the user.

Modern manufacturing techniques allow automated manufacture of custom products made to fit the individual. To support the automated design of these products it can be advantageous to use 3D imaging of the human form to generate high resolution 3D models.

Many of today's 3D imaging or scanning technologies were originally developed for applications in industrial inspection and are not well suited to scanning people. They were optimized to scan machined metal parts, injection molding and similar materials. People can be far more challenging to scan than many manufactured parts. For example, humans are living breathing objects who are continuously moving so a slow 3D scanning process cannot be used.

At Fuel3D, we have developed a unique innovation: a high-speed 3D scanning technology that provides the ability to create cost-effective, high-quality scanning products for a wide range of applications, including those that capture the human form. The innovation is unique, patented and presents multiple market sectors with the opportunity to integrate 3D images and data into their operations.

Our technology delivers 3D capture speeds below 0.1 seconds, enabling the fast and accurate capture of high resolution 3D data to a level that has not been possible before, and at a fraction of established price points.

Application areas and technological challenges

To understand the applications and challenges associated with scanning the human form, it is helpful to consider the various components of the human body for which custom solutions are frequently designed.

The head

Let's first consider a relatively simple example. Crash helmets are designed to help absorb energy in the event of an accident by spreading the load of the impact across the head. For a helmet to work well it must be a close fit to the head and, to get the best performance, the helmet should be custom fit to the individual. This offers both improved performance and improved comfort.

For such an application, a 3D scanning system that can capture a full 360 degree model all around the head is

Figure 1 | Photo of the Scanbrella product concept.

required. Typically, the required resolution for development of a customized crash helmet is 1 to 3 mm. One of the main challenges with scanning a subject for a helmet is factoring in the wearer's hair, which can come in many styles, but will be flattened against the head while in the helmet. But thanks to the soft nature of hair, a simple balaclava can be used to crush hair to the position it will be within the helmet, ensuring an accurate scan can be taken. As an added benefit, if the wearer's face is scanned at the same time — this data could be used to help design alignment of helmet mounted optics to the eyes.

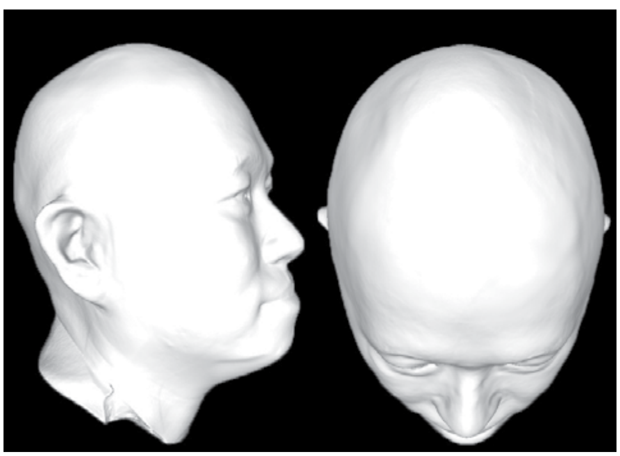


Figure 2 | Photo of a Scanbrella head scan.

The feet

At the other end of the body, it can be useful to scan feet for shoe fitting. In its simplest form, 3D scan data can be used to assess size in terms of length and width, as is already common in retail applications. But using 3D scanning, this can be taken a stage further by analyzing arch height. This allows better fitting footwear to be selected from an off-the-shelf range providing greater comfort.

But the main application of high performance 3D foot scanning, however, is in the manufacture of custom made orthotics. These are used for both medical and sports applications to improve comfort, posture, or physical performance. The typical required resolution is around 0.5 to 2mm.

First-generation foot scanning systems were based on mechanical flatbed scanning technology, similar to photocopiers. Such systems had the disadvantage of a 3 to 15 second scan time, which is sufficient for movement to cause distortion in the 3D model. One of the challenges for this application is the ability to provide a weight bearing scan with the person standing on glass.

Figure 3 | Image of the Cryos product. The Cryos has a capture time of less than 0.1 seconds.

The face

3D scanning of the face can be used for a wide range of different applications, from analyzing the impact of cosmetic wrinkle reduction creams to producing lifelike facial avatars for use in VR applications.

A growing application area in the next few years will be in the eyewear sector. For glasses to work well, the lenses must be correctly positioned in front of the eyes. Error of position of even 1 to 2mm can significantly degrade the performance, particularly of varifocals. Currently, achieving an accurate, comfortable fit on the bridge of the nose and ears is often done by manual mechanical adjustment, but this has limitations.

Using 3D scanning to build custom made eyewear allows this application to be optimized. This may, for example, include maximizing the field of view and enhancing protection from projectiles. Both of these characteristics rely on being able to design frames so that they are close to, but do not touch the cheeks, forehead and eyebrows. For good performance, resolution of 3D scanning technology for eyewear applications needs to be around 0.25 mm.

Figure 4 | Image of a Mirror scanner. The mirror allows natural user interaction.

3D models may not only be used for the design and fitting of eyewear. A 3D scanning system with good color capture allows a person to be visualized in 3D with the eyewear in place. This form of virtual-try-on allows a customer to wear their current glasses, but see how they would look with new frames. The key to successful virtual-try-on is a high quality, high resolution color 3D scan, so that the viewer concentrates on the glasses rather than any defects in the scan, which can be distracting.

One of the greatest challenges is to provide a scan with realistic looking eyes. The lens of the eye is transparent and so does not scan well with most optical scanning systems. This can be distracting to the viewer.

In addition, many optical scanning technologies do not successfully scan hair, as it does not present a single surface and often leaves voids in the 3D model. This does not look good for visualization and does not allow for checking the potential of the eyewear touching facial hair. Our scanning technology provides a continuous surface with an approximation to the hair's texture. This can be particularly useful for eyewear fitting around eyebrows.

Figure 5 | Color image of cropped face scan.

Some facial scanning applications require an even higher level of detail. For example, cosmetic testing requires pore-level imaging at a resolution of 0.1mm or less, and both color information and 3D geometry to show wrinkles. This level of geometry detail is better than what is used by many movie studios in feature films.

The full body

Some of the greatest challenges in scanning the body come from the soft nature of the body. It is very difficult to use 3D scanning to measure the waist sufficiently and accurately to provide good fitting jeans. Scanning to 20mm accuracy is not limited by the scanning technology, but by the body. The human body changes shape significantly depending on the clothes being worn. It is therefore necessary to scan for the waist measurement with clothes and a belt or waist band in place. The clothes however often obstruct the view making measurements particularly challenging.

3d scanning of the human body can also be used for a number of medical applications, from the monitoring of tumor growth to burns. The scanning technology can help change the design and manufacture of custom fit prosthetics from a craft industry to a more scientific, repeatable process. Accuracy requirements typically are in the range of 0.25 to 1mm.

Some medical applications present particular challenges for optical scanning technology. For example, wound imaging can be difficult because the surface of the wound is often wet. This produces strong specular reflections which can confuse many scanning systems.

Scanning technology for the human body

3D scanning technology is a bit like distance measuring technology, in that no one solution suits all applications; you would not use a micrometer to measure the length of a road. The choice of technology needs to be based on a detailed understanding of the requirements.

For industrial 3D scanning applications there has been a long held view that you should choose your solution based on any two of the three characteristics:

- Speed of 3D model generation
- Level of resolution
- Cost

When your application involves 3D scanning of the human form it is also important to consider additional factors, such as:

Movement

People move all the time. Even when asked to sit still, there is still some movement, so fast capture is very important. The blink of an eye is approximately 0.2 seconds, so a capture time of 0.2 seconds or less is highly desirable.

Systems that require the scanner to be moved around a person to form a 3D model will often suffer from motion artifacts. Snapshot capture from multiple viewpoints addresses these issues.

Resolution

For applications requiring visualization of the face, it is important to have high resolution as people are particularly sensitive to errors in facial characteristics. For many applications, an ideal resolution is to match that of the unaided eye, which is approximately 0.1 mm for an object that is close by.

Color

For many 3D body scanning applications, color is not essential, as all that is required is the geometry for custom fit product design. However for visualization many people are not familiar with seeing a monochrome 3D model and find it much easier to interact if information is presented in color.

Summary

3D scanning the human body can facilitate many new applications and enhance existing ones. Commodity, off-the-shelf scanning technology can be used for some basic scanning applications. For example, consumer products for gesture recognition such as the Kinect for Xbox can be used for limited scanning of the human form. However for high-performance business or mission-critical applications, it is beneficial to have 3D scanning solutions built and optimized to meet the specific requirements of the particular application. **Q**

Andrew Larkins *was originally trained as a Physicist at Imperial College, London and obtained an MBA through the Open University. He has 25 years' experience in commercialization of new technologies in fields ranging from semiconductor to medical devices. Currently he is Chief Technical Officer of Fuel3D, overseeing the development of the company's patented 3D scanning technology for a range of commercial, industrial and medical applications.*

From the IQT Portfolio

The *IQT Quarterly* examines trends and advances in technology. IQT has made a number of investments in innovative technologies, and several companies in the IQT portfolio are garnering attention for their unique solutions.

Fuel3D

Fuel3D is a developer of advanced 3D scanning systems and solutions. Originally developed for the medical imaging sector, Fuel3D technology has been adapted for the broader 3D market, with the goal of bringing the benefits of point-and-shoot 3D imaging to consumers, professionals and businesses. The technology combines photometric stereo imaging with stereoscopic imaging to produce a single 3D image. Fuel 3D has been an IQT portfolio company since December 2014.

www.fuel-3d.com

IkeGPS

IkeGPS delivers mobile geospatial solutions, and has quickly become a global player in the utility and defense segments with measurement solutions based on its smart laser-camera GPS technology and 3D modeling software. IkeGPS' products can determine height, width, area, and location, and are based on commercial smartphones and tablets. The company is located in Wellington, NZ, and has been an IQT portfolio company since December 2011.

www.ikeGPS.com

Pelican Imaging

Pelican Imaging has developed a computational imaging technology to provide depth mapping at every pixel, giving users the freedom to refocus after the fact, focus on multiple subjects, segment objects, take linear depth measurements, apply filters, change backgrounds, and create 3D models. Pelican's depth-sensing array technology is significantly thinner than many existing mobile cameras. Pelican Imaging is located in Santa Clara, CA, and has been an IQT portfolio company since March 2010.

www.pelicanimaging.com

Evolv

Evolv is developing low-cost physical threat detection technology that provides flexibility across a wide range of threat profiles and operational scenarios. The technology further identifies the location of threats on a person as they pass by the Evolv scanners. Evolv is located in Waltham, MA, and has been an IQT portfolio company since March, 2015.

www.evolvtechnology.com

IQT®

IN-Q-TEL