3D RANGE IMAGING CAMERA SENSING FOR ACTIVE SAFETY IN CONSTRUCTION

SUBMITTED: June 2007
REVISED: September 2007
PUBLISHED: April 2008
EDITORS: B. Akinci and C. Anumba

Jochen Teizer, Assistant Professor
Georgia Institute of Technology, Atlanta, Georgia, U.S.A.
email: teizer@gatech.edu http://www.rapids.gatech.edu

SUMMARY: *Accident reports in the United States' construction industry show over 1,000 fatalities for each of the past eleven years. Most of these accidents are related to falls and close contacts to equipment or other harmful substances. Although these accident statistics would have been significantly worse if safety training and best practices had not been provided to workers, the assessment of critical situations often involves the experience and judgement of field personnel, such as of safety coordinators. The manifold of accidents and their repetitive nature, however, allows the conclusion that proper inspection of construction sites is often not performed accurately or safety coordinators are not on hand when needed. For these reasons, this research utilizes emerging technologies to assist in the safety decision making process. This paper presents a review of the safety problem in construction. It then focuses on automated data collection devices and data processing algorithms which can become part in an active safety system for construction applications. A research approach is presented to use and develop emerging sensing technologies for accident avoidance. The purpose was to apply a 3D Range Imaging Camera as part of active sensing technologies that allow fast and accurate range measurements. The collected range data was used to generate real-time feedback about the location of objects in the field-of-view of the sensor. The experiments and data analysis performed demonstrate that construction safety can be improved by using emerging technologies such as 3D Range Imaging Cameras. In a workforce-material-machine dominant environment such as construction, more research is necessary to further validate the developed methods in long-term studies and in outdoor environments.*

KEYWORDS: *3D, range imaging, real-time data collection and processing, remote sensing, background subtraction, occupancy grid, safety.*

1. INTRODUCTION

The dangers of work are usually measured by the number of injuries or fatalities occurring to a group of workers, usually over a period of one year. Whether in manufacturing, construction, or other industries, based on historical recordings rates over the past decades have decreased because man-work-hours have increased (NIOSH, 2006). Although over 1,000 fatalities remain in the U.S. construction industry for each of the past eleven years, a few reasons can be mentioned that contributed to the fact that accident recordings did not increase simultaneously. These reasons can be named as industrialization and automation of work tasks, education and training, installation and use of safety tools and devices, etc. (CPWR 2006 and 2007). These actions have placed workforce outside of previously known hazardous work zones, provided them with adequate understanding and regulations of how to execute the task in a qualitative and safe manner, or prevented accidents by requiring workforce to wear or install so called *passive safety devices*, such as hardhats, gloves, goggles, trench boxes, and guard rails.

Passive safety devices can be defined as tools that can prevent accidents or fatalities, but once installed do not carry any additional function to prevent accidents actively (Teizer, 2007). A car airbag, for example, would fall into the category of passive safety devices, since it only inflates once an accident has occurred. *Active safety devices*, however, sense the environment for potential safety risks and take preventive action before any accident can occur. By providing the information necessary to make good safety decisions, actively monitoring the environment can prevent accidents and fatalities.

As a consequence from the remaining injury and fatality rates in construction and elsewhere, this research intends to reduce the likelihood of accidents by using active sensing technology. It is the purpose to detect and track construction resources (e.g., workforce, equipment, and materials) in real-time. This becomes ultimately necessary when contacts resulting from struck-by events of heavy equipment and many other equipment related accidents need to be avoided. Fig. 1 demonstrates a common problem in the construction/mining industry: Low operator visibility and missing appropriate sensing devices i.e. in the blind spot caused a dual fatal accident. For this particular application, the design of an active safety tool requires fast updates of machine location and its surrounding object position, e.g. workforce and materials. Technologies exist that can determine this positioning data.

FIG. 1: Fatality involving Heavy Construction Equipment and U.S. Safety Statistic (Left Image, Courtesy OSHA)

Until today and among others, numerous distance sensing systems exist such as Acoustical Distance Sensors, Thermal Range Detectors, Laser Detection and Ranging (LADAR), Light Detection and Ranging (LIDAR), laser scanners, and 3D Range Imaging Cameras. Due to the accurate, safe, fast, and easiness to handle, the use of optical range measuring has become the predominant method to collect range information in the construction and transportation sector (Shaw, 2006). In recent years the need for accurate and fast visualization, modelling, and simulation for in construction and transportation environments has increased the interest in optical three-dimensional (3D) imaging technologies. This demand has been creating the development of a variety of enabling range sensing systems based on light and laser range scanning approaches. The large number of existing and emerging range imaging technologies highly differs in their functional principle, their specifications, and lacks a standard terminology that easily allows understanding the advantages and disadvantages of each system.

This paper intends to classify optical range imaging techniques and investigates the differences in working methods and characteristics. In particular, 3D Range Imaging Cameras as state-of-the-art, large-field-of-view, high resolution, and high speed range sensing devices are put into perspective to existing 3D measurement approaches. Although experimental results and new developments in faster range image data processing methods are presented for indoor environments, this paper demonstrates that 3D Range Imaging Cameras have reached an accuracy level that allows them to be applied to some construction and transportation work tasks, e.g. to improve safety in heavy equipment operation.

2. BACKGROUND REVIEW

The literature briefly reviews existing research work in construction safety. Based on the current understanding this research justifies the need for automated active safety monitoring. It further classifies the terminology and working principle of potential range sensing technologies for active safety and their benefits and limitations.

2.1 Need for Automated Active Safety Monitoring

Many researchers have concentrated on construction safety and in finding reasons of why accidents happen and how to avoid them (Arboleda et al, 2002, Bernold and Ziadong, 1997, BLS, 2006, CPWR 2006, Cho et al, 2001, Hinze, 2005, Irizarry, 2002, Lee and Halpin 2003, NIOSH, 2006, OSHA, 2007). With these results significant contributions and changes in construction safety were successfully made or are underway. Several of these studies, e.g. Plog et al. (2006), recommend improvements in training and outreach on construction hazards, as well as increased regulatory actions and advances in technology. Advances in heavy equipment operation in trench work, for example, focuses mainly on improvements in lighter and stronger shields and on the promotion to use more trenchless technologies (Bernold, 1997). To have significant impact in trench safety and in construction safety in general, the presented research indicates that future innovations in safety can and need to go beyond the improvement of exiting technologies. In addition, this research demands clarification to better understand why these injuries and fatalities continue to occur and to take further action to prevent them.

2.2 Enabling Optical Range Sensing Technologies

This part of the literature review focuses on technologies that can make heavy equipment operation safer. Although other technologies offer technically feasible approaches for active safety, passive (light) and active (laser) optical range imaging systems and in particular emerging 3D Range Imaging Cameras are reviewed in more detail.

Passive optical sensors work similarly to single-lens-reflex or film cameras. They do not transmit any form of energy into a scene, but use naturally present light to obtain range data in single shots or multi-frame grabs. They can be installed airborne, spaceborne, on-shore, or in sub-surface terrain. Examples of passive sensor systems are 3D color imagers, ultraviolet, or infrared noise-equivalent temperature difference sensors. Advantages are low power consumption, a high spatial resolution, and a multi-band capability. The visible spectrum and contrast limits their work best to daytime use (night-time requires sophisticated processing if possible at all due to low illumination power) (Shaw, 2006).

Instead of using a passive measurement principle that often requires extensive post data-processing to obtain range values, active sensors transmit some form of energy into a scene to receive a return signal that allows determining ranges. Examples of active sensors are laser scanners, LIDARs, LADARs, and 3D video range cameras. The following paragraphs are a synopsis to the physical limits of active vision based range imaging techniques.

The following technologies were identified to have potential benefits and limitations in developing active construction safety tools. This discussion focuses partially on technologies that have the potential to detect objects in close proximity to heavy equipment. Some of the reasoning may not be applied to other applications in construction. In general, reflectorless range measurement approaches based on light or laser sensing using the time-of-flight principle can be classified in five main categories (Teizer, 2006):

- (Stereo) Video Imaging or Time-Lapse Photography: Usual term for commercially available products being applied in various applications, e.g. machines in backing motion. Video Imaging requires stereo pictures to create distance information thus requiring preferably large distances of camera installations. Otherwise inexpensive and well known approach.
- Laser Radar or LADAR (Laser Detection and Ranging): Usual term for government-supported detection-related systems of hard targets, e.g. defense work. These laser sensors provide range images consisting of a set of point-measurements from one view-point by “moving” the laser beam using rotating mirrors/prisms. LADAR require complex, fine tuned components, and precise alignment of submitted and received light beams. The overall effort to collect range information is very expensive and needs extensive and time consuming post-processing of range data before a 3D model can be build (Roth, 2006).
- LIDAR (Light Detection and Ranging): Usual term for primarily commercial airborne mapping systems, typically measuring distributed scattering for environmental work. LIDAR has a similar working principle as LADAR.
- 3D Laser Scanner: Usual term for terrestrial/industrial use based on raster line scanning principle. 3D laser scanners have a similar working principle, advantages and limitations as LADAR. 2D and 3D laser scanners are larger than the size of a camcorder and generally expensive devices that rotate or pan the sensor to reconstruct an entire static scene through a complex mechanics. This raster scan principle adds one line after the other to a range array, but consumes time to synchronize the acquisition of single lines with the sensor motion. This can be very beneficial for building accurate 3D models since millions of range points can be provided. In conclusion, LADAR, LIDAR, 2D and 3D laser scanners can build high resolution range images of only static scenes. For safety applications this comes with high computational and purchase cost.
- Laser Rangefinder: Usual term for lower end commercial/industrial use, e.g. Total Station for sparse point data. The sparse point cloud approach focuses on selected points to avoid high computational costs of acquiring dense range point cloud data and therefore requires only a few

minutes to model a scene. Human intervention is needed to select object descriptive points (e.g. edges and corners) via a laser range finder (Teizer, 2007). As a result, the major limitation (and power) of the sparse point cloud approach is the requirement for human judgment and the focus on static objects in environments.

- 3D Video Range Camera aka. Flash LADAR: Usual term for emerging technology and prototypes useful for range imaging for real-time visualization and modeling of static and moving objects. The following paragraphs explain more performance details to the 3D Range Imaging Camera.

Other technologies exist that can be used in active safety. Since the following technologies do not use optical measurement principles, some of them might be limited in their application:

- RADAR (Radio Detection and Ranging): Usual term for military application requiring high and potentially unsafe radio signals. Commercial applications exist but have a very directional field-of-view.
- Ultrasound technologies: Usual term for commercially available products with short range of a very few meters.
- Thermal Imaging: Usual term for sensing objects and their differences on a temperature scale in a scene, e.g. human vs. material. Thermal images have found applications in the military and mining industry.
- RFID/Ultra-Wideband: Usual term for prototype technology that can measure distances to objects that are tagged based on using radio waves. Non-tagged object will not be detected.
- Global Positioning System (GPS): Usual term for commercially available technology that relies on satellite and base station signals. Requires tagging of resources and may not be available all times.

Optical 3D data acquisition of scenes is preferred over alternative methods such as radar or ultra-sonic since optical techniques allow fast and safe range acquisition of (untagged) objects at a high lateral resolution. While dense point cloud approaches such as LADAR, LIDAR, and laser scanner are precise but slow and expensive, the sparse point cloud approach generates 3D models at much lower cost and tends to be acquired faster, but contains only a few selected range points. A meaningful comparison of the sensing methods can be made on the basis of the following criteria (Teizer, 2007):

- Density of data used in modelling (a higher density offers a wider field of applications)
- Frequency of updating of the derived model (allows real-time modelling updates)
- Precision and accuracy (how well the model reproduces the actual scene)
- Richness of the derived model (information quantity and quality incorporated into the model).
- Data collection and processing method (real-time for static/dynamic objects, ambient conditions)

3. RANGE IMAGING TECHNOLOGY

Range imaging technology is an emerging technology that collects mainly range values to each image pixel, at a high update rate, of large field-of-views, and to distances of up to 50m. The following sections explain the technology in more detail:

3.1 Working Principle of 3D Range Imaging Camera

Emerging efficient 3D Range Imaging Cameras acquire and store range, amplitude, and intensity data in matrices of points (distance map) to entire scenes in real-time. They do not use scanning principles or stereo vision with complex filtering and correlation processing units that are needed for many real-time applications. The equipment itself does not require costly scanning components that require time and monetary investment. Instead, it uses Active-Sensor-Pixels (ASP) that acquire range, intensity, and amplitude data of entire scene field-of-view in one frame (Lange and Seitz, 2001).

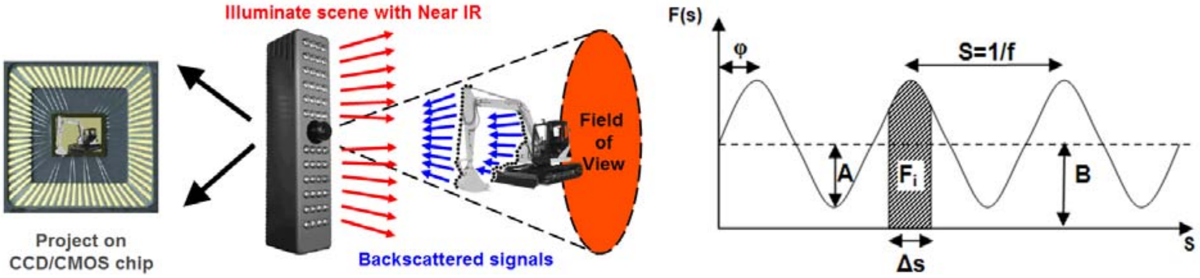


FIG. 2: Measurement Principle of 3D Range Imaging Camera (Teizer, 2006)

This research used a prototype 3D Range Imaging Camera (SR3000). Falling into the group of contactless distance measurement devices (see Fig. 2), it is based on the time-of-flight (TOF) principle using phase shift measurement. The camera uses 55 conventional light emitting diodes (LED's) to actively illuminate a scene by emitting sinusoidal modulated (spatial or temporal) near-infrared light (eye-safe, peak wavelength at 870nm) at a modulation frequency of 20MHz. The time the light needs to get to and to return from an impinged object in a scene back to the sensor is then measured using a practical synchronous sinusoidal demodulation. Focused through a lens and within the 3D range camera, a CMOS/CCD sensor chip is positioned to receive the incoming wave front. Each so called "lock-in-pixel" on the sensor chip is able to demodulate the incoming optical wave front in parallel and samples four discrete times within a period c(τi) (i=0,1,2,3), while each sample is delayed by a π/2 phase shift φ. It then calculates the amplitude, intensity, and range values based on the TOF principle (Lange and Seitz, 2001). Repeating this process for approximately 25,000 pixels on one chip makes real-time imaging possible. Both the detection and the complete demodulation are performed in the charge-domain using charge-coupled devices (CCD). That ensures almost noise free demodulation of the incoming light signal. The brightness information is the average of all four amplitude samples. The resolution (frame) refresh rate is currently limited to 50Hz. Experiments used 15.2Hz and a FOV of 41.7° horizontal and 44.6° vertical. A point at 7.5 m in the distance map represents a *voxel* (space volume also called VOlume piXEL) of about 4.6cm in each axis. Operating with only one modulation frequency f_{mod}=20MHz ultimately limits the unambiguous distance D to

$$D \leq \frac{\lambda_{\text{mod}}}{2} = \frac{c}{2 \cdot f_{\text{mod}}} \approx \frac{2.998 \cdot 10^{8} \frac{m}{s}}{2 \cdot 20 \cdot 10^{6} Hz} = 7.5m$$

Thus, applying a second source with different modulation frequency, period, Λ1 and Λ2, generates a so-called synthetic wavelength, that extends the non-ambiguous range, resulting in

$$\Lambda_{12} = \frac{\Lambda_1 - \Lambda_2}{|\Lambda_1 - \Lambda_2|}$$

Lowering the modulation frequency allows reaching multiple distances (SR3000 up to 22.5 m), but reduces the reflection of the emitted light from impinged objects to the extent that the range accuracy drops. Picking or sampling signals from the backscattered harmonic wave fronts is very hard to control since the length of sampling period is often too short to collect sufficient numbers of photons (photo charges) for reliable range data sampling. Thus, integrating the flux of particles (e.g. photons, electrons, ions, atoms or molecules) over time for certain intervals allows capturing relevant range information. Figure 3 illustrates in detail that during a sampling period Δs a flux of the returned photon wave front is measured over a minimum spatial area during the period S=1/f. Measured is a number of F_i particles (photons, electrons, etc.) for specific integration intervals. Four samples, each shifted S/2+nS with a length Δs are used to calculate amplitude A, offset B, and phase φ of the harmonic wave (Seitz, 2005). The phase shift is exploited for the extraction of the distance information by sampling. The received signal is off-set shifted by a mean optical power mainly due to additional background light and non-perfect demodulation.

$$A = \frac{\sqrt{(F_3 - F_1)^2 + (F_0 - F_2)^2}}{2}$$

$$B = \frac{F_0 + F_1 + F_2 + F_3}{4}$$

$$\varphi = \arctan\left(\frac{F_3 - F_1}{F_0 - F_2}\right)$$

Knowing the phase shift to each pixel on the sensor, the sensor directly measures the distance d_{Pixel} to the captured target:

$$d_{PIXEL} = \frac{c \cdot \varphi_{TOF}}{2 \cdot f \cdot 2\pi}$$

As an important side effect of this measurement principle, amplitude and offset can be used to determine accuracy levels of the measurement by calculating the standard deviation of the phase measurement σ

$$\sigma = \frac{\sqrt{B}}{\sqrt{2} \cdot A}$$

As a result, for most of these reasons, calibrating the 3D Range Imaging Camera is an important task in order to obtain accurate distance measurements.

3.2 Advantages and Limitations of 3D Range Imaging

Summarized in (Teizer, 2007), many researchers have found advantages and limitations of range imaging prototype sensors and in particular the influence of an ambient environment.

The biggest advantage of 3D Range Imaging Cameras is the ability to collect dense point cloud range data in real-time of a larger field-of-view. Especially in the application for detecting, tracking, and modelling moving objects, this technology is preferred to non-real-time range sensing methods like the previously mentioned sparse point or laser scanning approaches.

The ease of generating, manipulating, and detecting light is the reason why optical 3D sensing techniques have become the favourite approach in acquiring the 3D shape of our environment quantitatively. Continuously-modulated time-of-flight measurement has lower requirements to the sensing unit, since it (currently) operates on one bandwidth and at one modulation frequency. This allows reducing the manufacturing cost of the sensor. The biggest benefits from 3D Range Imaging Cameras are:

- Deliver range, amplitude, and intensity maps in one frame and at the same time
- Safe and very short data acquisition time with high frame update rate for immediate range feedback
- Wide field-of-view
- Capturing static and dynamic scenes and thus not conceivable to laser scanners
- Ease of use at day and during night
- Insensitivity to background light
- Handheld like small sized and compact devices
- Competitive prices

The applicability of optical 3D sensing techniques restricts its use to areas where line-of-sight is the preferred alternative. Light as a carrier wave to collect range data is sensitive to ambient environments. Physical effects deserve detailed consideration because they may also limit the performance of 3D vision methods. The main limitations to 3D Range Imaging Cameras are currently:

- Missing standardized calibration technique for laser range imaging systems and data processing algorithms (in general for all laser scanning systems)
- Ambient environment influencing measurements (e.g. atmospheric noise) requiring post data filtering
- Optics or physical camera effects (lens or detector) causing inaccuracies in distance measurement performance
- Non-optimal manufacturing of camera device and parts (unsymmetrical LED mount, pixel-saturation, dead pixels)

- Line of sight produces shadow effects (2½D image only)
- Diverted, extended, or (non-)reflected light beams (due to object colour, edges, corners)
- Ambient environment (background light and jitter noise)

Each of the optical range sensing methods has its own practical and theoretical difficulties and limitations, but all range imaging approaches are following the same functional relationship, ultimately limited by the quantum noise of the light generation and detection process (as one limiting factor to the accurate performance of optical range imaging systems). Optical diffraction, speckle phenomena, and physical effects deserve detailed consideration because they may also limit the performance of 3D vision methods.

A three-dimensional imaging system such as 3D Range Imaging Cameras can rapidly measure the range, amplitude and intensity image including thousands of points to objects or scenes at high update rates. Once these dense "point clouds" are accurate enough they can be processed to generate e.g. 3D models. Significant deviations between the measured data and the reality occur due to several reasons. Moreover, due to the lack of existing standardized terminology for 3D imaging systems, the National Institute of Standards and Technology (NIST) has pre-defined the following terms (NIST, 2006):

- *Calibration* as a set of operations that allows correcting the differences determined by a quantitative measurement of the relationship between instrument and its corresponding output values. Influencing parameters on raw data are determined as follows: Temperature of environment, reflectivity of material, and distance to object.
- *Accuracy* as the closeness between the measured value and the true measurand value.
- *Precision* as the closeness of agreement between independent results.
- *Errors* as differences in *true value* and actual *measurement* (e.g. *systematic errors* as reproducible differences of the measuring instrument and *random errors* as unrepeatable differences in measurement and true value).

4. ALGORITHMS, PERFORMANCE, AND RESULTS

Although many application areas may have particular needs and requirements to adapt technology, most of them demand fast data collection and data processing. To be helpful in safety applications, for example, technology is required to work at high update rates because situations on a construction site may change instantly. This makes hardware and software development challenging since it needs to collect data and process the data fast enough to provide meaningful information for any warning or obstacle avoidance system.

In any active safety system, "background subtraction" offers one solution to find the location of objects within the field-of-view of the range camera as quickly as possible. The Range Imaging Camera provides range, intensity, and amplitude values to each pixel in a frame generally at update rates higher than 10Hz. Fast data processing becomes essential to combine sensor, data collection, data processing, information, and feedback to a successful system for safety applications (see Fig. 3). Whereas the amplitude in the process of data acquisition determines the accuracy of the measurements, the range and intensity values can be further explored to segment objects and provide information to an active safety system, e.g. issuance of an alarm. Generally active safety systems rely on rapid updates in order to react quickly to events that may endanger human lives. Thus processing the collected data as quickly as possible is essential in the feedback loop.

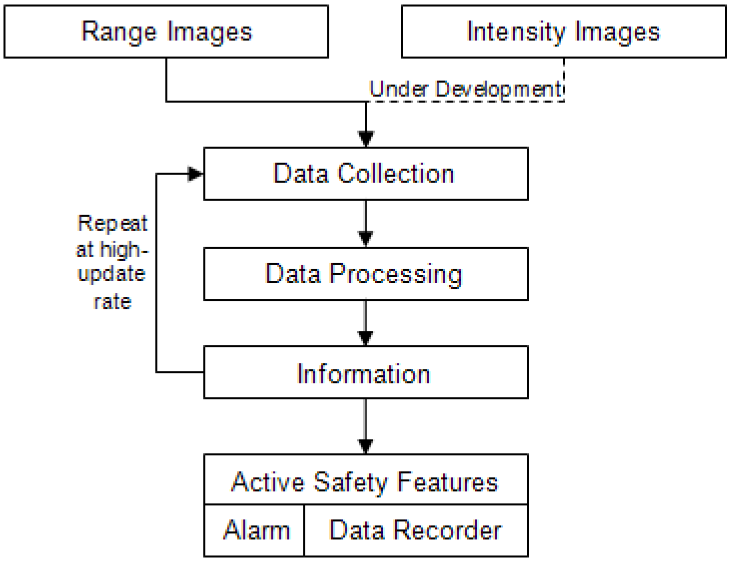


FIG. 3: Flowchart of High Level View of Active Safety Algorithm

The main aspects to the 3D range imaging technology have been discussed above. The following sections present three different algorithms that were developed in this research: 1) Background subtraction from a static 3D Range Imaging Camera using a variable threshold approach, 2) Background subtraction from a static 3D Range Imaging Camera using a probabilistic threshold approach, and 3) Occupancy Grid Algorithm for static and dynamic 3D Range Imaging Camera movements.

4.1 Background Subtraction using a Variable Threshold Approach

A total of 33 experiments were performed indoors where a smaller hallway branched off from a larger one. The 3D Range Imaging Camera was installed on a tripod and pointed towards the smaller hallway (see Fig. 4). The first 100 range images were collected with an empty hallway to determine the background image. All tests had a person walking up and down the hallway, in a straight or zigzag pattern, or hiding temporarily outside the sensors line-of-sight in an alcove. All tests were repeated for different camera settings including signal amplitude values of 20, 40, 60, 80, and 100. Each experiment lasted in between 5 and 15 seconds and subsequently range data frames of a minimum of 100 to a maximum of 200 range frames were collected. Final experiments included the person wearing a reflective safety vest and helmet.

The first range data processing algorithm based on background subtraction was developed in MATLAB® and involves several steps. First, 100 range (and intensity) frames without any target in the scene were recorded. These 100 frames build the background image based on averaging the corresponding pixels of all background frames. With this method, error readings, local minima and maxima became less weight on the next processing steps. Significant random range measurements ("salt and pepper noise") throughout the entire range image were filtered from the background images. Each range frame is acquired and stored in separate three-dimensional arrays. The dimensions correspond to horizontal image size, vertical image size, and number of frames. Next, the algorithm calculates the time average of the set of background images (i.e., for each pixel, it calculates the average value over all frames), subtracts it from the range images with object, and takes the absolute value. By doing so, an array containing the absolute difference between each test image and the background is created. Finally, the algorithm applies a threshold (calculated by multiplying the single maximum difference pixel value by a user-supplied parameter, in percent) to create a mask, zeroing all background pixels. A threshold of 0.2, for example, means that all range points that have range difference lower than 20% between background image and range image with object are filtered out. Generally the higher the threshold, the more data is filtered out. In a final image, the remaining points are range points only. However, noise can still appear in the remaining range image. A final step is calculating and plotting the trajectory of the object. The trajectory is created by calculating the centroid of each frame in the result data and plotting the x (left-right position) and y (distance from camera) portions of it.

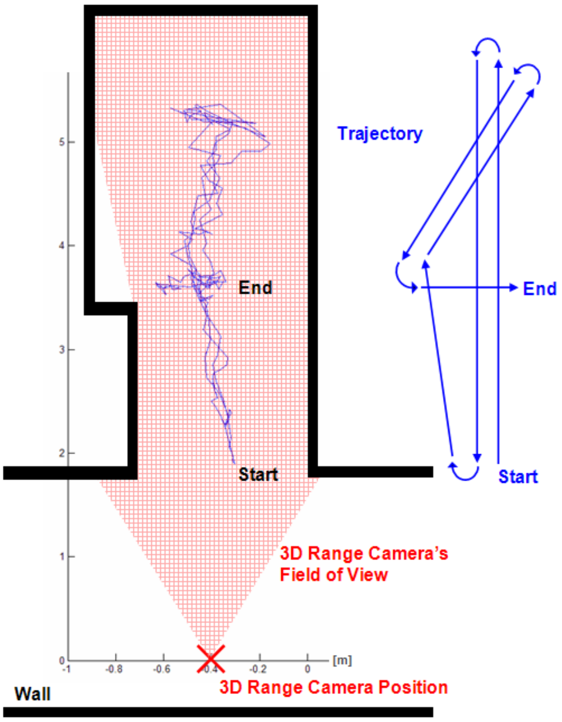


FIG. 4: Trajectory in Plan View with Overlaid Building Hallway Layout

In order to test the minimum and maximum performance of the algorithm the process of data collection and processing was separated. All range frames were recorded at an update rate of 15.2Hz. Once recorded and stored on the computer hard drive, the algorithm had optimal access to the data. Table 1 displays the results of lowest, average, and highest performance of the algorithm (times needed to process range data and times needed to visually plot data). All 33 experiments were evaluated using four different thresholds, 0.05, 0.1, 0.2, and 0.5. Example using a Dell Optiplex 745/Intel Core 2 Duo Processor/2GB RAM: To process 100 range frames at a threshold value of 0.1 took 250.9 milliseconds or 0.2509 seconds and 0.0668 seconds to display its trajectory. Using the same threshold, upper and lower boundaries of processing times were 0.23 and 0.27 seconds. The graphical display of the trajectory took in between 0.046 and 0.089 seconds for a total of 100 range frames. Based on all experiments and all thresholds, this correlates to an average frequency to process range frames of 401Hz and to an average frequency of 1566Hz to display the information on a screen. The achieved performance exceeds by far the update rate raw range data is provided from the 3D Range Imaging Camera. Even the lowest possible experienced performance of the developed data processing and display algorithm still outperformed the update rate of the camera (15.2Hz) by far (202Hz for processing range data and 909Hz for displaying the trajectory) (see Table 1 and Table 2).

TABLE 1: Algorithm Performance – Impact of Thresholds on Data Processing and Plotting Time

33 different experiments	Processing Range Data of 100 Frames [milliseconds]			Display of Trajectory of 100 Frames [milliseconds]		
Threshold	Lowest	Average	Highest	Lowest	Average	Highest
0.05	240	260.2	320	55	74.4	110
0.1	230	250.9	270	46	66.8	89
0.2	225	247.4	495	45	62.9	99
0.5	230	233.8	270	37	49.4	64

TABLE 2: Algorithm Performance – Minimum, Average, Maximum Possible Update Rates

33 different experiments	Processing Range Data			Graphical Display of Trajectory		
	Lowest	Average	Highest	Lowest	Average	Highest
Update Rates [Hz]	202	401	444	909	1566	2703

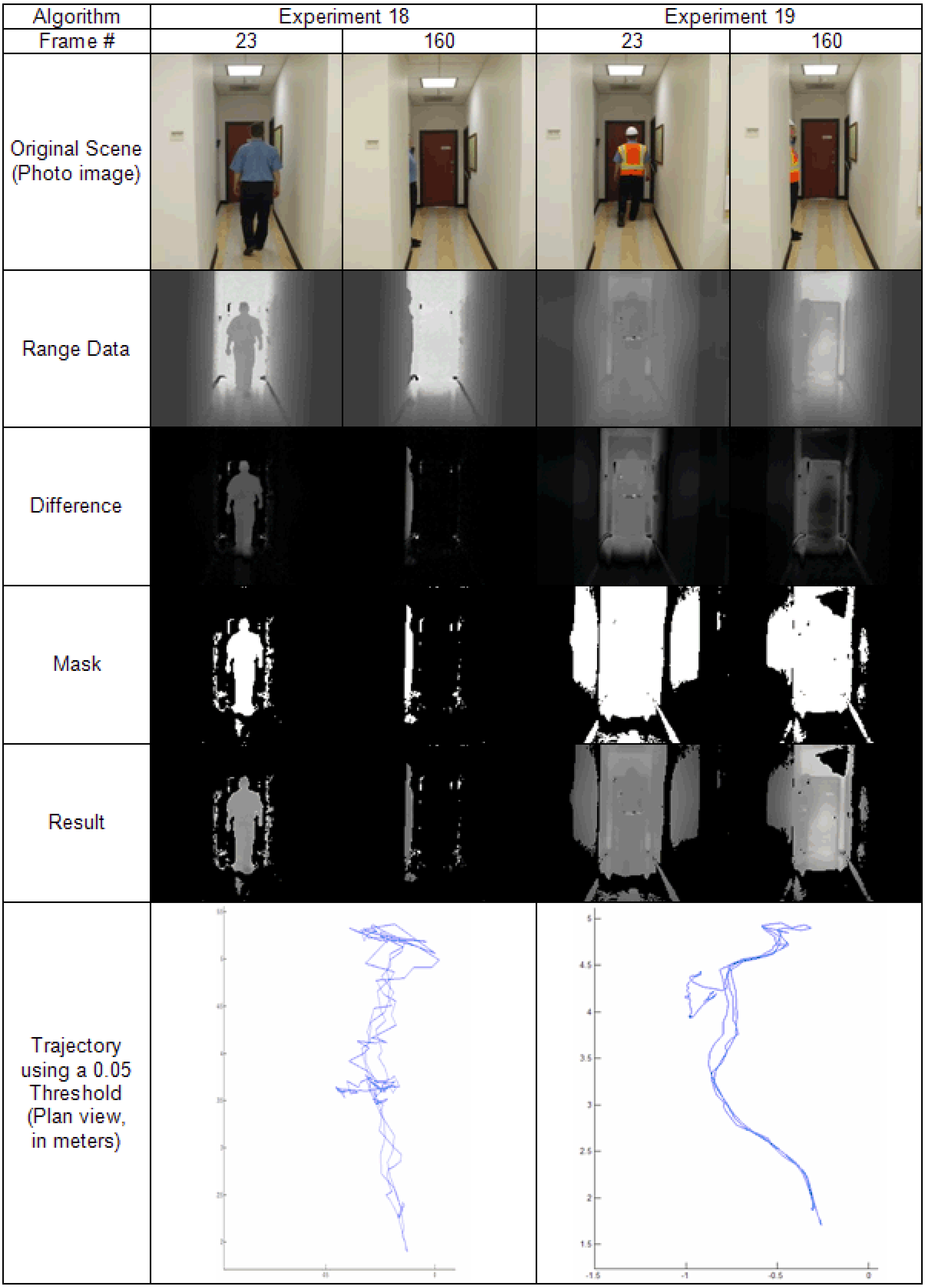


FIG. 5: Trajectory of Experiment With and Without Safety Vest using Background Subtraction at Threshold 0.05

In Fig. 5, experiments were recorded that show a difference of two setups. The person in experiment #18 is wearing no safety vest, where as in experiment #19 the person does wear a safety vest. Safety vests are made out of reflective fluorescent material that makes it hard for any camera to track an object due to the bright reflected illumination. Although, a similar observation was recorded using the 3D Range Imaging Camera, the intent of the following experiment, to detect and track the location of a person wearing a safety vest using a 3D Range Imaging Camera, was successful. The second column in Fig. 5 demonstrates the result of two individual frames and a final trajectory of the path of the person. The influence of the safety vest on creating the trajectory can be seen by comparing it to the trajectory without safety vest (left column, experiment #18).

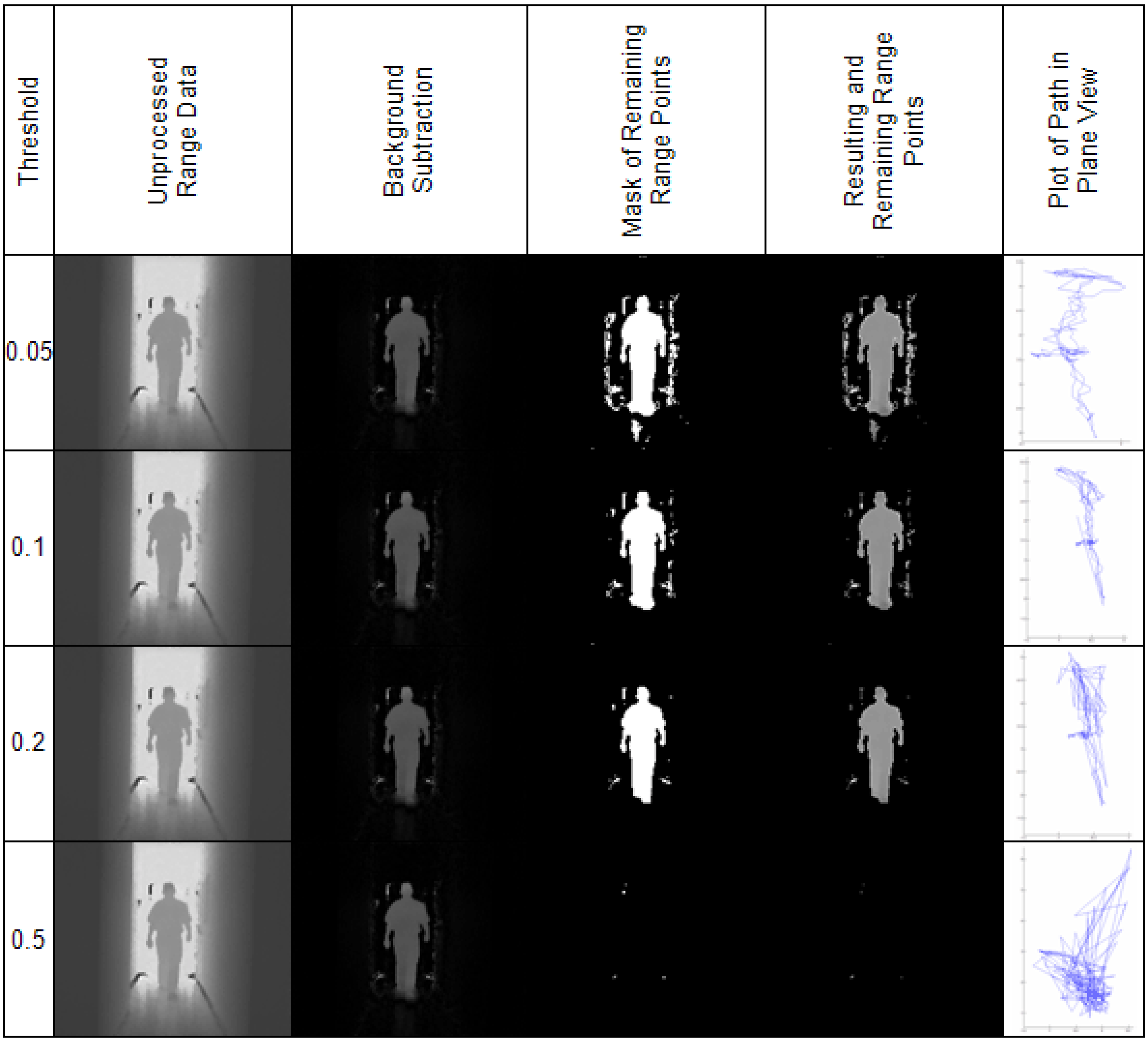


FIG. 6: Comparison of Threshold Levels of Experiment 18, Range Frame 23

General comments to the results are:

- The average/min/max performance of the algorithm was determined by tracking a single object in the path of the 3D range imaging camera.
- The object detection read rate was 100% in the field-of-view of 3D Range Imaging Camera, including experiments with a safety vest. The quantification of the difference in paths with and without safety vest needs to be researched in more detail.

As demonstrated in Fig. 4, the trajectory with overlaid building layout is illustrated. Based on distance data previously analyzed for range measurement accuracy, the expected location accuracy of a single range point is within a few centimetres in horizontal, vertical, and depth orientation at maximum object to camera distances of 7.5 meters (Teizer, 2006). The analysis of the experienced trajectories validates these findings. Although no absolute location measurements can be given from a human walking a hallway (curves are expected), trajectory results of the start points, end points, and location of the person within the niche were compared to the dimensions in reality. For example, the vertical distance from 3D Range Camera to niche was 3.5m compared to the approximate measurement of 3.6m in the trajectory image (Experiment 18).

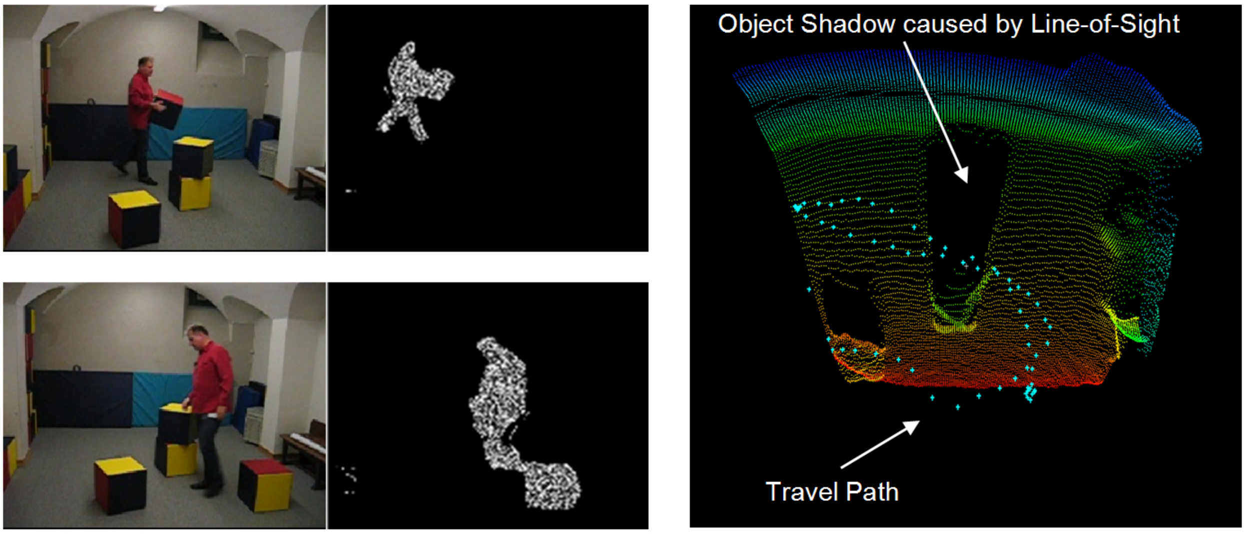
The effect of the threshold values on the range image data processing can be seen in Fig. 6. For threshold values, 0.05, 0.1, 0.2, and 0.5 the same range frame (#23) was evaluated. All threshold values except 0.5 allowed the background subtraction algorithm to track the path of one single object in the range camera's field-of-view. Increasing the threshold value mainly reduces "salt and pepper" noise (local and random extreme range measurements on a few distinct pixel points). Setting the threshold value too high (e.g., 0.5) excludes critical information of the object in the scene. Recommendation is given to use smaller thresholds, e.g. 0.05, or 0.1. Both threshold accurately reflect the actual movement in the scene. However, all experiments and all thresholds recorded the appearance of an object within the scene at all times. This detection rate is important for safety applications, since even a small error percentage could cause an accident or fatality. Future research, on the other hand, must clearly indicate and differentiate false alarms from real alarms.

The achieved performances demonstrate that developing active safety features, tools, and instruments using 3D Range Imaging Cameras is plausible. Based on the update rate of television screens (about 30Hz), the achieved minimal frame processing and plotting rate of 202 Hz is sufficient for real-time feedback to operators or workers (visual through screen monitoring in machine driver's cabin, audio, or vibration alarms on helmets).

4.2 Background Subtraction using a Probabilistic Approach

Kahlman and Ingensand (2006) and Teizer and Kahlman (2007) presented a probabilistic approach to detect and track resources (person and box). In this background subtraction method, a range frame of the original scene is taken and subtracted from all succeeding frames. In a static field of view of the range camera, only those range points remain in the scene that changed their value. As a result, objects can be detected and tracked. In an example experiment shown in Fig. 7, a moving person changes the environment by putting a box in the range sensor's field-of-view.

The first few 3D range frames containing the basic information of the environment were used as background. In order to find changes in the field-of-view like the appearance of persons, in every captured frame the distance towards the neighboring points in pixel space was calculated. If a distance limit exceeds, the point is marked as an obstacle point. In the upper part of Fig. 7a the person with the box is identified as a single object. After the cube has been laid down and the person turns away, both are regarded and tracked as separate objects.



(a) Original image and 3D extraction in range image (b) Detection of static and track of moving objects

FIG. 7: Trajectory in Plan View with Overlaid Building Hallway Layout (Teizer and Kahlman, 2007)

Fig. 7b displays the plan view of the 3D space of the same indoor scene. The lighter points show the person's trajectory. For simplicity reasons the centroid was used to track the volume of objects remaining after the background subtraction. Further developments should match a representation of the object (cube, cylinder, complex surfaces) to gain even better results.

Although the background subtraction algorithms achieve high update rates it is ultimately limited to fixed or slowly moving camera locations. Outdoor construction environments such as heavy machine movements where sensors are required to work from non-stationary platforms may require other methods of processing range data. Continuously changing background images ask for different data processing solutions that can adapt to changing environments quickly and reliably. Occupancy Grid Algorithms offer a potential solution. In the next paragraph a solution is presented.

4.3 Occupancy Grid Algorithm

A shortcoming of the background subtraction approach is its unfeasibility to reference range frames in environments that consistently change their background. Thus, a different approach is based on “Occupancy Grids” that allow allocating the collected range data from a static or moving sensor platform (e.g. vehicle) into a predefined fixed voxel system. Occupancy Grids have the advantage that no a-priori data is required to determine the number of all clusters (Moravec and Elfes, 1985). More details to the developed Occupancy Grid System algorithm, experiments, and accuracies are explained in Teizer (2007) and Teizer et al. (2007).

Range information is ideal to find and track objects in three-dimensional space. To reduce computational cost, once individual voxels reach a threshold number of range points, each voxel is filled with the number of corresponding range points falling into the voxel. Comparing the fill factor of voxel neighbors and their distance to others allows to group objects (voxel groups). Results to one experiment are illustrated in Fig. 8. This experiment shows the feasibility of detecting static and moving objects from a static or moving sensor position in real-time. Moreover, the accuracy of the modeling approach to position, dimension, direction, and velocity was determined by comparing the processed range data of the 3D video range imaging camera to reality measurements using a Total Station (position, dimension, and direction verification) and camcorder (velocity). Results demonstrate sufficient location accuracies of less than 10cm in all axis directions for the object centroid. These preliminary results are considered to be sufficient for applications in construction safety.

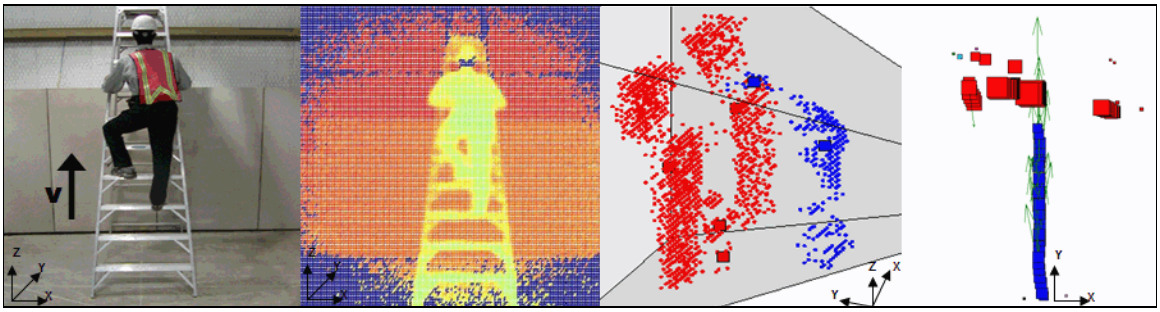


FIG. 8: Occupancy Grid Algorithm – Scene, Range Data, Processed Voxel Image, Trajectory (Teizer, 2007)

4.4 Summary

In conclusion, the developed methods to detect and track static and moving objects and the experienced accuracies of position, dimension, direction, and velocity values achieved satisfactory results. The research objective to demonstrate that emerging 3D Range Imaging Cameras can be applied in various applications including active safety for construction has been successful.

5. APPLICATIONS AND FUTURE WORK

As many applications can be considered for 3D range imaging, in construction the surveillance of construction sites and the surrounding of heavy equipment like dozers, excavators, cranes, and trucks can be named to increase security and safety. Range imaging features fast and reliable capturing of objects in a larger field-of-view, at high update rates, and at day or night work. Thus, the use of range imaging may be preferred over video cameras that require a secondary light source to illuminate the scene. Research studies using this and other emerging sensing technologies are under way at the Real-time Automated Project Information Decision Systems (RAPIDS) laboratory at the Georgia Institute of Technology and include the following aspects:

- Validate existing results in long-term studies, e.g. record rate of near misses on outdoor construction sites
- Long range and wide field-of-view measurements using multiple 3D Range Imaging Cameras
- Monitor workforce behaviour and location

This research demonstrated that active safety in construction is technically feasible. Many automated obstacle avoidance algorithms exist in the field of robotics, but intelligent data processing methods are still missing and need to be developed that can rapidly interpret images. Foremost the segmentation and classification of objects in range images is non trivial and requires vast amounts of data processing resources (capacity and speed). Targeted hardware developments are planned to assist workforce and operators, e.g., the use of handheld sensors, or the installation of temporary or fixed intelligent safety devices on jobsites, e.g., 3D Range Imaging Cameras on a fixed pole or a backhoe. Mass manufactured and implemented 3D range imaging devices would be relatively inexpensive to find adaptation in construction.

6. CONCLUSION

The research goal was to apply emerging technologies to collect and process safety relevant data from construction sites to make site work safer. This paper introduced the terminology to an emerging three-dimensional imaging system, called 3D Range Imaging Camera. The 3D Range Imaging Camera allows collecting thousands of range points from objects or entire scenes at update rates equal to video. The working principle of 3D Range Imaging Cameras was explained in detail and distinguished from existing optical range imaging systems. Thoughts were explained that can assist the development of accurate calibration methods for a 3D Range Imaging Camera. Results to experiments were demonstrated and showed the advantages and limitations of this emerging range imaging technology. Potential applications with high impact on construction for accurate, high frame rate, and wide field-of-view range sensing were offered. The demonstrated approach is believed to be in particular useful for applications in safety and night operations where static and dynamic objects can hardly be recognized by existing intensity or colour based sensing systems. Once applied in the field, 3D Range Imaging Cameras can offer valuable data to create new training and education tools in regards to safety in construction (e.g., so called “black box recording” for simulation events). In summary, in combination with fast data processing algorithms 3D Range Imaging Cameras are a promising emerging range sensing technology that can help reduce accidents and fatalities.

7. REFERENCES

Arboleda C.A., C.A., Abraham D.M., Wirahadikusumah, R.D., and Irizarry, J. (2002). Trench-related Fatalities in Construction: An Analysis of Fatality Assessment and Control Evaluation (FACE) Records. *Proceedings of the 1st International Conference on Construction in the 21st Century*: Challenges and Opportunities in Management and Technology, p. 277-282.

Bernold L. and Xiaodong H. (1997). CAD-Integrated Real-Time Control for Robotic Excavation and Pipe Laying: Development and Testing. *<http://www.azfms.com/DocReviews/Nov97/art15.htm> (*Sept 30, 2006).

BLS, Bureau of Labor Statistics (2006). *Census of fatal occupational injuries*. U.S. Department of Labor. Washington, DC. <http://www.bls.gov/iif/home.htm> (Sept 30, 2006).

CPWR, Center to Protect Worker’s Rights (2006). *Safe work in Trenches*. <http://www.cpwr.com> (Sept, 30, 2006).

CPWR, Center to Protect Worker’s Rights (2007). *Trenching-Related Injuries and Deaths.* The Center to Protect the Worker’s Rights. <www.cpwr.com> (Jan. 17, 2007).

Cho Y.-K., Haas C.T., Liapi K.A., Sreenivasan S.V. (2001). Rapid Visualization of Geometric Information in a Construction Environment. *Fifth International Conference on Information Visualization*, p. 31.

Hinze J.W. (2005). The use of trench boxes for worker protection. *Journal of Construction Engineering and Management*, ASCE, 131(4), 494-500.

Irizarry J., Abraham D.M., Wirahadikusumah R.D., and Arboleda C. (2002). Analysis of Safety Issues in Trenching Operations. *10th Annual Symposium on Construction Innovation and Global Competitiveness*, <https://engineering.purdue.edu/CSA/ publications/trenching03> (Jan.16, 2007)

Kahlmann T. and Ingensand H., (2006). Calibration of the Fast Range Imaging Camera SwissRanger for Use in the Surveillance of the Environment, *SPIE*, Vol.6396-5.

Lange R. and Seitz P., (2001). Solid-state, Time-Of-Flight Range Camera. I*EEE Journal of Quantum Electronics*, Vol.37, pp.390-397.

Lee S., and Halpin, D.W. (2003). Predictive tool for estimating accident risk. *Journal of Construction Engineering and Management*, ASCE, 129(4), 431-36.

Moravec H.P. and Elfes A.E., (1985). High-Resolution Maps from Wide-Angle Sonar. *IEEE International Conference on Robotics and Automation, pp.116-121.*

NIOSH (2006). Fatality Assessment and Control Evaluation Program. Department of Labor. <*http://www.cdc.gov/niosh/face/faceweb.html*> (March 20, 2006).

NIST (2006). 3rd Workshop on Performance Evaluation of 3D Imaging Systems. *http://www.bfrl.nist.gov /861/CMAG/LADAR_workshop/Terminology_pre-standard_v10.pdf.* Accessed June 20, 2006.

OSHA, Georgia Tech, and ASSE (2007). “Trench Safety Task Force”. <*http://www.osha.gov/dcsp/alliances/regional/region_4.html.*> Jan. 10, 2007.

Plog B.A., Materna B., Vannoy J., Gillen M., (2006). *Strategies to Prevent Trenching-Related Injuries and Deaths.* The Center to Protect Worker’s Rights. <http://www.cpwr.com> (Sept. 30, 2006).

Roth M (2006). 3D Visualization Techniques for Laser Radar. *SPIE*, SC717.

Seitz P. (2005). Unified analysis of the performance and physical limitations of optical range-imaging techniques. *First Range Imaging Research Day*, pp.9-17.

Shaw J. (2006). Introduction to Optical and Infrared Sensor Systems. *SPIE*, SC789.

Teizer J. (2006). *Real-time 3D Modeling to Detect and Track Resources on Construction Sites.* Dissertation, University of Texas at Austin.

Teizer J. and Kahlmann T. (2007), Range Imaging as an Emerging Optical 3D Measurement Technology, *Transport. Research. Board*, Washington D.C., (in press).

Teizer J., Caldas C.H., and Haas C.T. (2007). “Real-Time Three-Dimensional Occupancy Grid Modeling for the Detection and Tracking of Construction Resources,” *ASCE Journal of Construction Engineering and Management*, (in print).