System Design of Time-of-Flight Range Camera for Car Park Assist and Backup Application

Sunil Acharya
Canesta, Inc.
440 N. Wolfe Road, Sunnyvale, CA 94085
acharya@canesta.com

Colin Tracey, Abbas Rafii
Canesta, Inc.
440 N. Wolfe Road, Sunnyvale, CA 94085
ctracey@canesta.com, arafii@canesta.com

Abstract

External sensing applications in automobiles are enabled by Microsystems that can sense and analyze 3D environment. In this paper, we describe the system design of a time-of-flight (TOF) 3D range camera for parking assist and backup application. The camera operates based on measuring the time delay of modulated infrared light from an active emitter, using a single detector chip fabricated on standard CMOS process. The system analysis approach involves a step-by-step process that maps the application requirements to a series of camera design decisions producing the specification for a camera solution that is optimized for size, cost, power and application performance. The analysis process uses the empirical and theoretical performance properties of a 3D TOF sensor and the camera optics. It defines the profile of an optimal light source that delivers the necessary depth resolution for the given volume of interest in the application.

1. Introduction

The application that is used in the analysis is motivated by the tragic incidences that involve a car hitting a child when the driver is backing up. In the latest data from 2007 [2], 725 incidents were reported in which 942 children were involved and 231 fatalities were caused. A 3D camera that can perform detection function may be used to perhaps warn the driver of the impeding accident.

In this work, we develop our system analysis with Canesta's [1, 3, 4] Electronic Perception (EP) 3D TOF camera technology. To capture the depth, each camera pixel measures the time delay of modulated infrared light from an active emitter. The camera contains a light source constructed from a bank of infrared LEDs (or laser diodes), a lens system for the detector chip, a detector chip with phase-sensitive pixels fabricated on standard CMOS process, and an embedded CPU for application processing. The theory of operation of such a camera is described in [3]. The camera can also operate under strong sunlight conditions using Canesta's Sunshield^{TM} technology. The pixel has the ability to substantially cancel the effect of ambient light at the expense of producing a slightly higher noise.

With a large field of view, a 3D camera can replace several of the existing ultra-sound sensors that are presently used in the bumper of the cars. An ultra-sound camera works by time stamping the sound echo from the closest object without providing any information about the shape of the objects and surfaces.

It is sometimes proposed to perform the detection tasks using conventional 2D imaging sensors and analysis software. But in practice, achieving cost-effective and reliable performance during all vehicular usage scenarios is a formidable challenge. The appearance of objects in a 2D image varies greatly, depending on illumination conditions, surface materials, and object orientation. These variations in the image complicate the task of software that must interpret the scene. Whereas, the 3D shape of objects is invariant to those confounding effects.

RADAR and scanning LIDAR technologies are also similarly proposed. But they have difficulty discriminating objects due to limited temporal or angular resolution or have mechanical parts that are difficult to maintain.

Ultrasonic sensors have been used for measuring distances [7]. In a back-over application, the timestamp from the first echo is used. The other returned signals are typically ignored. Despite the use of multiple sensors to achieve good coverage, the U.S.' National Highway Traffic Safety Administration tests show blind spots that are not covered by the sensors.

For this application, a TOF 3D camera has several advantages over other 3D data capture methods relating to cost and performance. For instance, unlike stereo cameras that require two imaging optics, a 3D TOF camera uses a single lens and thus has a lower optics cost and less reliance on precision mechanical tolerances. Since the 3D TOF camera does not require a baseline, the 3D performance of the camera is relatively independent of the camera size and range of operation. The performance of the depth camera is characterized by how much a noisy instantaneous 3D point measurement deviates from its time-average mean [5]. In this paper depth uncertainty, depth resolution, depth jitter or noise are used interchangeably.

There is a fundamental relationship between the performance of the EP 3D camera and the amount of reflected active light that reaches the pixel. We have characterized this relationship using both theoretical calculations and verified it through empirical measurements. The amount of emitted active light that reaches the sensor is governed by the power of the active light, field of view of camera, reflectivity of the target and the transmission characteristics of the sensor lenses. A substantial part of the power budget of the camera is consumed in producing the light. Therefore, it is important to design the light source power and its projection profile in such a manner that the optical energy is used judiciously. On the other hand, the proper operation of the application's image processing and detection algorithm imposes certain upper bound in the noise threshold of the depth measurement in different parts of field of view. Therefore, the design of a complete solution is an exercise in achieving an optimal balance between the overall size, cost and power of the camera while meeting the requirements of the application.

In this paper, we describe the requirement of a backup and park assist application and extract the parameters that are relevant to the design of the camera. We describe the design elements of a 3D TOF. Using the size of smallest obstacles that must be detected at different distances behind the car, we determine the array size (number of pixels) of a 3D chip needed to detect the object. Since this application must use a wide angle lens, we describe how much distortion [6] is produced by the lens and how the distortion affects the illumination pattern of the light source. We then show how the light source intensity profile is designed such that the upper limit of the noise threshold required by the application is achieved in all parts of the camera field of view.

2. Application Description and Requirements

The backup and park assist application uses the images obtained from a 3D TOF camera that is mounted on the back of the car and detects obstacles behind the car. The detection is performed using a computer vision algorithm that estimates the ground plane behind the car, and detects the objects that are taller than certain height threshold above the ground and wider than certain width threshold, and are in the trajectory of the car as it backs up [4].

The importance of this application is evident in the statistics that KIDS and Cars Organization [2] maintains. Table 1 is quoted from their web site. It is obvious that the number of children involved in the accidents and the fatality rate of these accidents are staggering.

YEAR	INCIDENTS	CHILDREN INVOLVED	FATALITIES
2007 (as of 12/17/07)	725	942	231
2006	598	742	219
2005	454	553	226
2004	502	607	174
2003	610	762	189
2002	435	602	133
2001	403	524	124
2000	353	452	92

Table 1: Fatalities reported by Kids and Cars

Obstacle detection is a key requirement for a backup system. The system needs to detect, isolate, measure, locate, recognize, and track objects such as people, traffic, objects and other roadside features.

A single high frame rate focal-plane-array 3D sensor is desirable because it can serve multiple safety and convenience functions simultaneously, allowing applications to jointly exploit shape and appearance information in a dynamic scene. The output of the sensor should be a sequence of 2D arrays of pixel values, where each pixel value describes the brightness and Cartesian X, Y, Z coordinates of a 3D point on the surfaces of the scene.

The application requirement defines a detection zone behind the car. The application must detect obstacles and determine their location relative the car in this region. The detection zone may be between 150 to 300 cm long (from the bumper) and 180 cm wide. The height of the detection zone is about 60-80cm above the ground. These dimensions of course vary from one design to another design (Figure 1).

The mounting location of the camera is dictated by the car design guidelines and the ability of the camera to image the detection zone. A typical mounting location for the camera is near the rear license plate and about 105cm above the ground. This location is generally available because of the requirements to display the license plate and availability of power to illuminate the license.

The detection zone size and the mounting location of the camera are generally input parameters for the design process.

The volume of interest (VOI) of the camera is essentially the 3D volume intersection of the field of view of the camera and the detection zone.

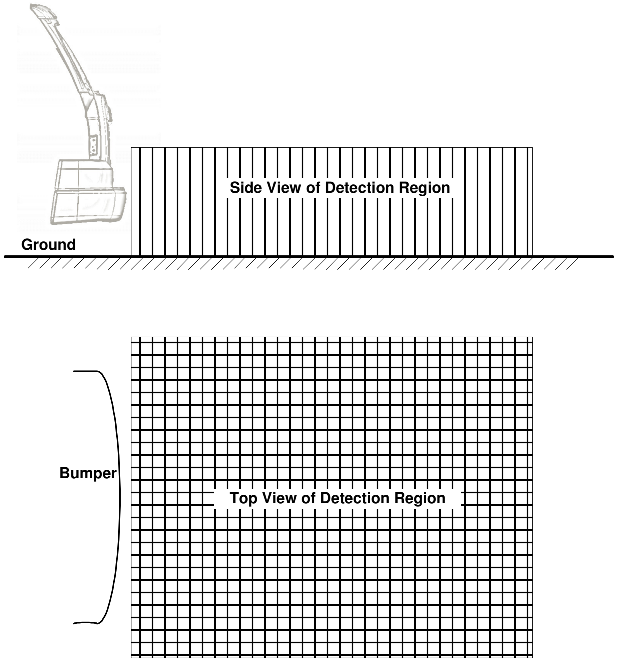

Figure 1: Detection zone of the application

The camera must be able to range objects from 10cm to as far as the length of the detection zone. Given that the camera is mounted in a recessed position relative to the bumper, any object closer than the minimum range is presumable is contact with the bumper.

The test detection object (known as ISO pole) is cylinder of 7.5cm diameter. This pole must be detected in any location and in any angle in the detection zone. The criterion for detection is determined by the ability of the application to determine the range of the pole. At the minimum, the pole must be imaged by at least one depth pixel width. The detection is of course is assisted by other pixels along the length of the pole.

When the object is detected in the detection zone, an audible warning pulsating tone is played. The pulse rate increases as the obstacle gets closer to the car. Alternatively, visual displays can be used to show the location and even the color-coded 3D cloud shape of the object relative to the car bumper. For instance, red color can tag close objects and green color can be used to paint far objects.

The application also imposes a minimum latency requirement. The latency is defined as the delay between the appearance of the obstacle and the object detection and issuance of warning to the driver. Considering the driver reaction time, the latency should be less than 100msec.

Of course the expectation is that the depth sensor accurately measures the distance to the object. The *accuracy error* requirement of 5% to 15% is typical.

The other measurement requirement is the repeatability of the measurements. In the current ultra-sound solutions ±12cm of repeatability is typical. The repeatability requirement states that the warning for an obstacle shall be consistent to within the repeatability range. That is, the warning provided for an obstacle placed at the same position at a different time will result in the same warning as initially provided. We assume a repeatability value of ±5 cm for our analysis. That is, if the object must be moved more than ±5 cm (translating to 1σ) to achieve the same warning tone then the system fails the repeatability requirement. This information is summarized in Table 2.

Item	**Requirement**
Range	150-300 cm
Minimum object size	7.5cm spherical object
Latency	< 100msec
Accuracy error	5%-15% of measuring range
Repeatability error	± 5cm

Table 2: Summary application requirement

3. TOF Performance Drivers

One of the parameters that are important in the performance of a backup application is the amount of jitter in the depth measurement. In EP TOF camera the level of measurement uncertainty is relatively independent of distance and a function of the active brightness level at the pixel. The following relationship is empirically observed:

$jitter \approx K \times b^{-0.4}$ (in the absence of strong ambient light)

Where depth *jitter* or uncertainty is defined as how much the instantaneous 3D point measurement deviates from its time average mean, *b* is the amount of active light received by the pixel (range determined by analog to digital converter resolution), and K is a constant. Therefore, by adjusting the value b through varying the shutter or adjusting the light source power, a desired level of upper depth jitter threshold for the application can be achieved.

The modulation frequency is another important performance factor of an EP TOF camera. The improvement in the depth measurement uncertainty has almost a linear relationship with the modulation frequency. Therefore, for a given amount of active light source power, it is more advantageous to run the camera at a higher modulation frequency to achieve a given level of depth uncertainty performance.

4. Camera Design

The camera combines the TOF sensor chip, sensor optics, light source and light source electronics, into a compact and cost effective package suitable for the application. The components of the camera module and the significance of each component for an optimal design are now described.

1. A printed circuit board (PCB) which carries the sensor chip, its support circuits and the electronic drive circuits for the light source. The TOF chip and its lens could be placed on this PCB.
2. A heat sink and thermal management strategy to maintain the components within their specified operating range through the environmental operating conditions specified for the module. This is particularly an important design parameter because of the extreme temperatures that the unit must operate.
3. Electrical connectors which meet automotive standards for temperature, vibration and electro-magnetic interference (EMI), connecting the module to the system processor and power supply.
4. Sensor optics including lens that defines the required field of view (FOV) which collects light and focuses the image on the sensor.
5. A narrow band pass filter matching the wavelength of the active light source system specially optimized to minimize sunlight entry to the sensor and to maximize transmission of light from the active light source.
6. Active light source consisting of LEDs or Lasers to deliver the required light output power. The power of the active light source must meet the performance requirement of the application while not being over-designed to raise module cost and size, and cause heat management problem.
7. Light Source Optics to shape the LED light into the beam profile required for system performance and maximize efficiency of use of the light source generated power. The projection pattern of the light source must be carefully designed to deliver the necessary performance in every part of detection zone and application VOI.

For a large volume production, it makes economic sense to optimize each component for the specific application and installation.

5. Lens and Chip Design

In this part of the analysis, the goal is to determine the number of horizontal and vertical pixels needed in the TOF depth sensor chip to be able to detect the ISO pole in every part of the VOI. This is done by an iterative optical simulation that will calculate the maximum allowable barrel lens distortion [6] to meet the requirements (i.e. one pixel thick imaging of the ISO pole) given a lens focal length.

There are a couple of practical restrictions that should be taken into account. The analysis must make some assumption about the pixel size of TOF chip. Due to the pixel design, features and fabrication process, there are limitations on the smallest pixel size that can be manufactured. Using the pixel size of a currently available chip is a good starting point. The other input assumption for this step of analysis is the choice of a sensor lens focal length. There are also practical limitations on the choice of the focal length that relates to the lens format size and manufacturability issues.

It follows that a minimum sensor array of 61x111 pixels, not centered on the optical axis, is needed to capture the image of the whole VOI. Figure 2 : Showing ISO pole (blue) in vertical and horizontal orientations in the VOI (green), superimposed on the pixel array (red) and the lens outline (black), the ISO poles (blue boxes) placed vertically or horizontally throughout the VOI (green 3D outline) project to the sensor array (red rectangle outline) such that the pole images are always one pixel or more thick.

From this step of analysis, it also follows that the FOV of the sensor lens is 142° in diagonal direction at the minimum.

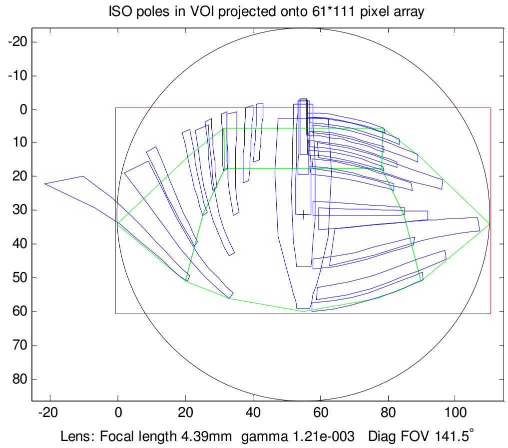

Figure 2 : Showing ISO pole (blue) in vertical and horizontal orientations in the VOI (green), superimposed on the pixel array (red) and the lens outline (black)

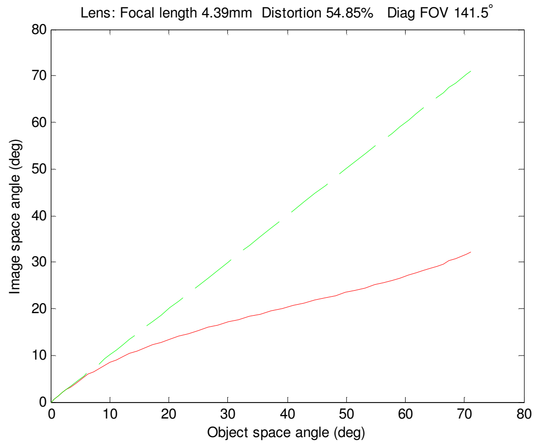

Figure 3: Distortion characteristic of lens in degrees off optical axis

The distortion characteristic of the maximally distorted lens is shown in Figure 3. The horizontal axis shows the object space angle in degrees and the vertical axis gives the image space angle. This information is used by the lens manufacturer to design the correct lens for this application.

6. Light Source Design

The light source design is based on the transmission profile determined from the lens and the VOI geometry. The lens profile (LP_1) (Figure 4) is a function describing the relative light intensity needed to compensate for the lens transmission losses. The profile is typically found by fitting a polynomial model to measurement data, or to a lens manufacturer's specification.

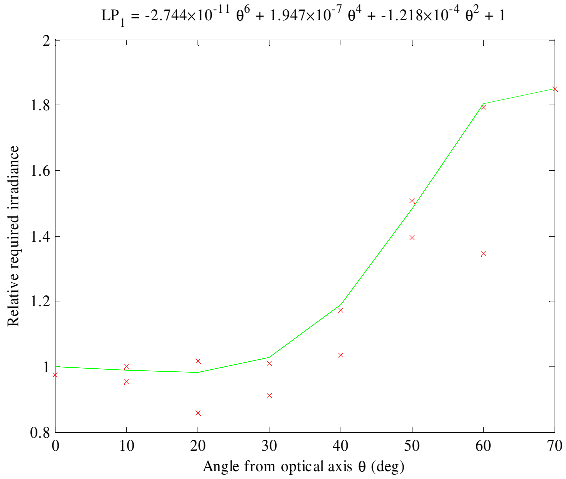

Figure 4: Relative required intensity profile to compensate for lens transmission losses

Assuming that all surfaces are Lambertian and oriented to face the light source, the radiant intensity (W/sr) of the light source must vary proportionally to the product of LP_1 and the inverse square of distance to the farthest point in the VOI in each direction. A depth elevation map (DEM) of the furthest distances within the VOI is created from the known camera location, viewing angle, and the VOI geometry. The product of the inverse squared DEM and LP1 generates a light scatter pattern which is provided to a manufacturer for design of a custom light diffuser (Figure 5).

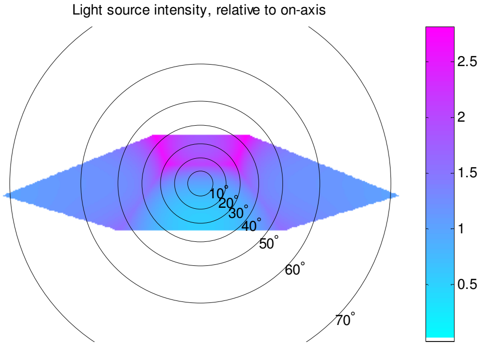

Figure 5: Radiant intensity pattern for the design of a custom light diffuser

The pixel surface is assumed to be Lambertian, facing parallel to the image plane, and therefore at a glancing angle relative to the incident light ray passing through the lens. To account for this Lambertian effect, the foregoing intensity profile is multiplied at each point by a factor of $\cos^3(\theta)$. The sum total of this relative profile is the multiplier factor relating light power illuminating just the surface area seen by the on-axis pixel to the total light power needed for the whole scene. This multiplier assumes the light distribution can be shaped exactly as desired. In this application, the maximum light should be aimed towards the distant back of the VOI, not along the lens optical axis.

Earlier we mentioned that the 1σ uncertainty requirement for this application was about ±5 cm. We also noted that there is well defined relationship between the active light power on the pixel and the depth measurement uncertainty. To meet a goal of 3σ uncertainty of 1.5cm or better, the active brightness at each pixel must be at least 70.63 units assuming the pixel structure operating at 44 MHz

An 808nm laser light source is chosen for the design. The sensitivity of the pixel at this wavelength is taken into account. Then the target area corresponding to the on-axis pixel must be illuminated by 0.05mW peak IR power during 15ms exposure. Applying the profile-based multiplier and allowing for miscellaneous losses, the total peak optical IR power over the FOV is 1.7W. For that shutter at 30 Hz frame rate, the average power for an optimal light source is 0.4W considering the duty cycle of the light source power.

The results are summarized in Table 3.

No.	Light Source	Peak Power (Optical Output)	Average Power (Optical Output)
1	Laser – 808nm (Infrared)	1.68W	0.38W

Table 3: Power analysis results

7. Conclusions

In this paper we have described a systematic method to determine the critical operational specification of a cost-effective TOF camera that meets the requirement of a car backup and parking assist application. The performance specification of a TOF camera is governed by a number of well-understood physical properties such as the number of depth pixels in chip, the power of active light and camera optics. For detecting an object and deciding that it constitutes an obstacle, the application requires a minimum number of pixels to image the object with certain level of depth data resolution. Choosing each one of these parameters in isolation from the others will results in a design that may not be optimal in terms of size, power and cost of the camera, and more importantly, may not meet the key requirements of the application. By posing the problem as a complete system of hardware and software, we were able to optimize the final solution across multiple dimensions.

We used the requirement of the application to define the field of view of the camera and used the repeatability factor to establish a maximum depth uncertainty in 3D data that can be tolerated by the application. We defined the components of the camera and determined the components that are drivers for the cost and performance. We used the detectability requirement of the application, to determine the effective FOV, lens distortion and minimum horizontal and vertical number of pixels in the chip. Using the relationship between the active brightness and depth uncertainty data, we defined the required illumination at furthest corners and sides of the detection zone. Finally, using the illumination requirement in the detection zone, we established the active light source pattern and its minimum power to illuminate the detection zone resulting in a system solution that is optimized for cost, size and power.

8. Acknowledgement

We would like to acknowledge the initial work of our colleague Steve Hsu on this project.

9. References

[1] www.canesta.com
[2] www.kidsandcars.org
[3] S.B. Gokturk, H. Yalcin, C. Bamji, “A Time-of-Flight Depth Sensor—System Description, Issues, and Solutions," in Proc. IEEE Workshop Real-Time 3D Sensors, 2004
[4] S Hsu, S. Acharya, A. Rafii and R. New “Performance of a Time-of-Flight Range Camera for Intelligent Vehicle Safety Applications”
[5] A. O. Allen, “Probability, Statistics and Queuing Theory,” Academic Press, 1978.
[6] Francis A. Jenkins and Harvey E. White, “Fundamentals of Optics,” 4th ed., McGraw-Hill (1976).
[7] A. Carullo and M. Parvis,An, " Ultrasonic Sensor for Distance Measurement in Automotive Applications", in IEEE Sensors J., Vol. 1, No. 2, (2001)