

3D Imager Design through Multiple Aperture Optimization

Sri Rama Prasanna Pavani, Jorge Moraleda, David G. Stork, and Kathrin Berkner

Ricoh Innovations Inc., 2882 Sand Hill Road Suite 115, Menlo Park, CA 94025-7057, USA
prasanna@rii.ricoh.com

Abstract: Three-dimensional (3D) imagers exhibit a tradeoff between device size and accuracy. We design compact and accurate 3D imagers by optimizing subsystem parameters using a multiple-aperture image simulator and an accuracy estimator operating on distorted views.

OCIS codes: (110.6880) Three-dimensional image acquisition; (110.1758) Computational Imaging

1. Introduction

Conventional optical systems such as cameras employ two-dimensional (2D) array sensors to detect images of 3D objects. This dimensionality reduction comes with significant loss of information in the depth (Z) dimension. While the transverse (X,Y) dimensions of the object space enjoy close to a one-to-one mapping with the image space, the depth dimension is often reduced to cues such as defocus, parallax, and perspective, all of which together complicate 3D position estimation. Existing 3D imaging techniques are either active or passive systems, depending on whether or not a predetermined optical field is projected on the object. Stereo imaging systems are prominent examples that fall in the passive category. By acquiring two different views of the same scene, these systems triangulate object points that appear in both views to estimate their 3D position.

2. Size vs. Performance Tradeoff

Typical stereo implementations have sizes that are at least of the order of several cubic centimeters. In these systems, any reduction in size is met with a penalty in performance. The depth uncertainty (δ) of a stereo system is, $\delta \approx \epsilon Z^2 / ib$, where ϵ is the uncertainty in disparity estimation, i is the lens-sensor distance, and b is the baseline distance between the two stereo lenses. From a system size point of view, i is the system thickness, b is the system length, and ϵ is inversely proportional to the system width (lens diameter). Reducing length, width, or thickness of a stereo system leads to a degradation of performance. This size vs. performance tradeoff makes it a challenge to design simultaneously accurate and compact system-on-chip stereo systems, whose dimensions are bounded by the dimensions of a CMOS or a CCD array sensor chip that is as small as a few square millimeters in surface area.

3. Multiple Aperture Design Method

Multiple-aperture designs consist of multiple imaging subsystems, all of which usually packed within the surface area of a single sensor array [1-5]. These designs are useful for 2D and 3D image acquisition with compact device sizes. Images formed by individual subsystems are often combined to form the final image, which in general reveals information not apparent in the individual subsystem images. Microlens arrays have frequently been employed in multiple-aperture designs because of their simplicity in design and fabrication. In this paper, we describe a method to design compact 3D imaging systems with high estimation accuracies using the multiple-aperture paradigm.

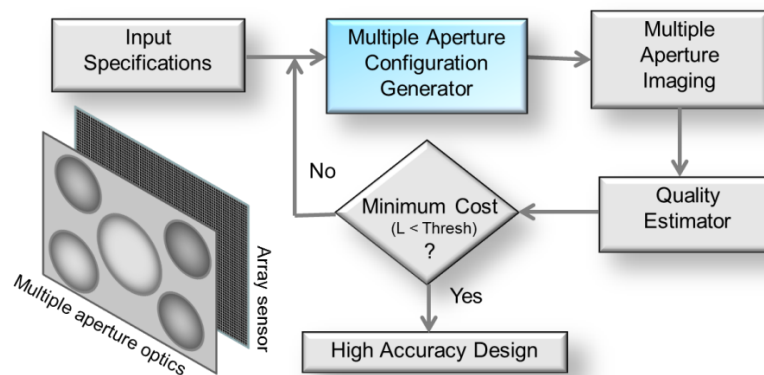


Fig. 1: Design method for compact and accurate 3D imagers. Multiple-aperture configurations offer more degrees of freedom than stereo.

JTuD4.pdf

In this paradigm, each subsystem records different information about the scene. The difference could be in viewing angle, which is related to the spatial positioning of the subsystem. Alternatively, the difference could also be in optical transfer function, polarization, or spectral properties of the subsystems. For a given system specification (sensor, field of view, object properties), we find a high-accuracy multiple-aperture subsystem configuration that minimizes a quality metric based on 3D position estimation uncertainties. We estimate such 3D uncertainties by estimating object positions from simulated multiple-aperture responses using a closed-form solution for position estimation and a Monte Carlo simulation for uncertainty estimation, after accounting for physical imaging effects such as diffraction, aberrations, distortion, quantization, and noise.

Fig. 1 shows our design procedure. In the *Input Specifications* block, we specify operational parameters that every multi-aperture configuration must satisfy: Sensor specifications, field of view, object properties, and disparity estimation algorithm.

The *Multiple Aperture Configuration Generator* generates a multiple-aperture design based on input specifications and feedback obtained from the previous iteration. Once a design is generated, this unit lists subsystem x, y, and z positions, dimensions, optical transfer function, spectral, and polarization properties. By nature of the specified problem, this unit confronts a multi-dimensional unknown search space. However, constraints can be imposed on the design space from physical system properties (such as non-overlapping lenses) and parameter interdependencies.

The *Multiple Aperture Imaging* unit simulates the response of the multiple-aperture system designed by the configuration generator. Calibration objects are positioned at 3D locations where the system needs to be optimized, and the subsystem images are simulated after accounting for distortion, diffraction, aberrations, detector sampling, noise, and quantization.

The *Quality Estimator* unit uses the simulated images from the multiple-aperture imaging unit to develop a quality metric for the system generated by the configuration generator. First, a disparity matching algorithm such as Normalized cross-correlation (NCC), sum of absolute differences (SAD), or sum of squared differences (SSD) is used in combination with a centroid estimator for sub-pixel block localization for each subsystem. These localization estimates are then fed into a closed-form analytical solution that minimizes a least-squares error function (eqn. 1) to estimate the 3D coordinates of an object point.

$$\text{Find } (X, Y, Z) \text{ that minimizes } \sum_{i=1}^K \left\{ \left[\frac{X - l_x^i}{Z - l_z^i} - f_x^i(d_x^i, d_y^i) \right]^2 + \left[\frac{Y - l_y^i}{Z - l_z^i} - f_y^i(d_x^i, d_y^i) \right]^2 \right\}, \quad (1)$$

where K refers to the number of sub-systems, l_x^i is the X position of subsystem i , d_x^i is the image position in the X dimension of subsystem i , and f_x^i is a non-linear subsystem function that relates image space to object space. Parameters for the y dimension are similarly defined. By using a Monte Carlo method, the quality estimator finds the dimensions (δ_x , δ_y , δ_z) of an uncertainty cloud in X, Y, and Z dimensions that exists around the mean position of the least-squares estimate. The cost function (L) is a function of the size of this uncertainty cloud. For example, a cost function for 3D imaging could be $(\delta_x^2 + \delta_y^2 + \delta_z^2)^{1/2}$. A multiple-aperture design that minimizes the cost function is chosen as the output of the optimization algorithm.

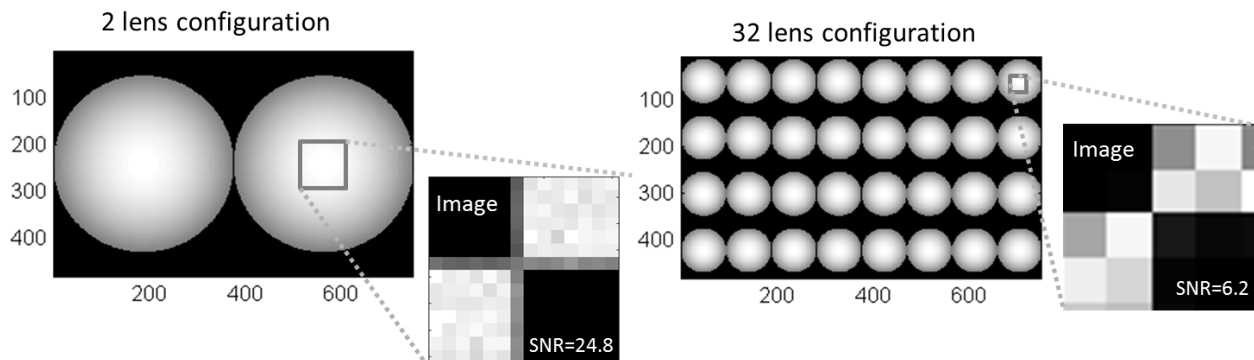


Fig. 2. A configuration generator showing a stereo configuration and a multiple-aperture configuration with 32 identical subsystems. Insets: Simulated images of a calibration object.

JTUD4.pdf

4. Results

In a proof-of-concept implementation of our design method, we chose to follow a checker board configuration of lenses (Fig. 2), with 2^N (N is an integer) identical lenses. The lenses are packed on a 752x480 pixel sensor with $6\mu\text{m}$ pixel width and 10-bit dynamic range. The system was assumed to operate in the shot noise limit. In any given configuration, the 2^N lenses are positioned within the area of the sensor without any overlap. Therefore, an increase in N results in a decrease in lens diameter, consequently a decrease in SNR (Fig. 3a) and resolution, and a decrease in the number of underlying pixels. In each configuration, the field of view is maintained to be 90 degrees for all subsystems, which causes any increase in N to reduce focal length and magnification of the subsystems. The image space $F\#$, however, remains unaffected by N . So are the aberrations, which are fixed to be $\lambda/10$ spherical and $\lambda/2$ defocus ($\lambda=500\text{nm}$). In our checker-board configuration of lens arrangement, successive configurations used different number of pixels on the sensor (Fig. 3c). The cost function was set to be $L=\delta_z$ for best depth accuracy.

As we increase N , some factors contribute to increasing depth error, and some contribute to decreasing the error. An increase in N comes with a decrease in SNR and resolution. This increases the disparity error (Fig. 3b), which in turn increases the depth error. However, the increase in N also results in the object being imaged from more view angles, which reduces the depth error. The effect of increasing N on depth error is shown in Fig. 3d. An interplay between view angles, SNR, resolution, and sensor usage is responsible for the complex dependence of depth error on N . In the 32 lens case, the additional views provided enough depth information to significantly overcome the effects of SNR and resolution (Fig. 3d), resulting in a depth error that is 2x lower than stereo. Note that the 32 lens configuration used the same number of sensor pixels as stereo.

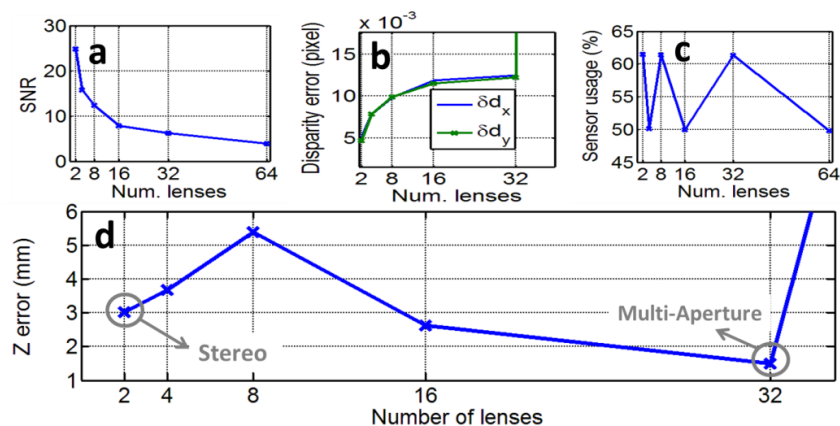


Fig. 3. In the shot noise limit, a 32 lens multiple-aperture system performs twice as good as a stereo system for depth estimation. This shows that multiple low resolution views can carry more depth information than two higher resolution views.

5. Conclusion

We introduced a multiple-aperture optimization method to design compact and accurate 3D imagers, and showed that multiple-aperture systems carry more depth information than stereo despite reduction in SNR and resolution. This design method allows designers to evaluate the performance of a variety of configurations before prototyping.

6. References

- [1] Jun Tanida, Tomoya Kumagai, Kenji Yamada, Shigehiro Miyatake, Kouichi Ishida, Takashi Morimoto, Noriyuki Kondou, Daisuke Miyazaki, and Yoshiki Ichioka, "Thin Observation Module by Bound Optics (TOMBO): Concept and Experimental Verification," *Appl. Opt.* 40, 1806-1813 (2001).
- [2] R. Horisaki, S. Irie, Y. Ogura, and J. Tanida, "Three-dimensional information acquisition using a compound imaging system," *Opt. Rev.* 14, 347-350 (2007).
- [3] Ryoichi Horisaki, Yoshizumi Nakao, Takashi Toyoda, Keiichiro Kagawa, Yasuo Masaki, and Jun Tanida, "A compound-eye imaging system with irregular lens-array arrangement," *Proc. SPIE* 7072, 70720G (2008).
- [4] A. Portnoy, N. Pitsianis, X. Sun, D. Brady, R. Gibbons, A. Silver, R. Te Kolste, C. Chen, T. Dillon, and D. Prather, "Design and characterization of thin multiple aperture infrared cameras," *Appl. Opt.* 48, 2115-2126 (2009).
- [5] Ryoichi Horisaki, Keiichiro Kagawa, Yoshizumi Nakao, Takashi Toyoda, Yasuo Masaki, and Jun Tanida, "Irregular Lens Arrangement Design to Improve Imaging Performance of Compound-Eye Imaging Systems," *Applied Physics Express* 3, 022501 (2010).