The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

Placement Quality in Structured Light Systems

Nathaniel Bird and Nikolaos Papanikolopoulos
Department of Computer Science and Engineering
University of Minnesota
{bird,npapas}@cs.umn.edu

Abstract— **This paper presents a mathematical basis for judging the quality of camera and projector placement in 3D for structured light systems. Two important quality metrics are considered: visibility, which measures how much of the target object is visible; and scale, which measures the error in detecting the visible portions. A novel method for computing each of these metrics is presented. An example is discussed which demonstrates use of these two metrics. The proposed techniques have direct applicability to the task of monitoring patient safety for radiation therapy applications.**

I. Problem Description

In this paper, the problem of measuring the quality of camera and projector placement for structured light systems is examined. Placement of sensors is important for all detection tasks as clever algorithms can be defeated by poor sensor placement, but good placement can lead to acceptable results even from subpar algorithms. Camera placement for observation tasks in multiple camera systems is an area that has been explored. However, the interaction between projectors and cameras in structured light systems is significantly different than the interactions in multiple camera systems, leading to different placement criteria.

The quality of differing camera and projector placements is determined here by examining the physics of the problem. The spread of light from the projector into the scene and the uncertainty in the camera's detection of this light is modeled and accounted for. Taking this into consideration, judging the placement of cameras and projectors in such systems can be performed mathematically. One area where improved placement is applicable is tracking a patient's body position during radiation therapy, in which small body movements direct radiation where it is not intended. To detect such movements, the cameras and projectors in a structured light system must be placed such that the body can be reconstructed with the desired precision.

The paper is organized as follows. Section II gives a brief overview of the related literature. Section III then gives an intuitive description of the quality metrics used. The mathematics of these metrics are presented in Section IV. An example problem is analyzed in Section V. Finally, conclusions are presented in Section VI.

II. Related Work

Structured light is a vision-based method to perform 3D reconstruction of a given surface, much like stereo vision. The difference is that instead of using multiple cameras, a set of cameras and projectors are used instead. Mathematically, projectors act as cameras, only instead of detecting features inherent in the scene, they project artificial features out onto the scene for the actual cameras in the system to detect. Thus, while the correspondence problem is a major issue for stereo vision systems, it is significantly lessened for structured light systems. Batlle *et al.* [1] presents a survey of structured light methods and Salvi *et al.* [7] present a follow-up.

Examples of work dealing with optimal camera placement include the work of Bodor *et al.* [2], who present a method to determine the placement of a set of cameras for best visibility based upon an example distribution of trajectories. This work places the cameras by minimizing a cost function which penalizes foreshortening and poor resolution. The work of Mittal and Davis [6] deals with camera placement in the presence of randomly positioned occlusions in the environment for which an example distribution is known. Cameras are placed such that the probability of occlusion is minimized. Chen and Davis [3] present a camera placement parametrization which considers self-occlusions. Their work also provides a metric for analyzing error in the 3D position of a point seen by several cameras. Tekdas and Isler [9] present a method for optimally placing a stereoscopic camera pair to minimize the localization error for ground-based targets in 2D. Krause *et al.* [5] present a method utilizing Gaussian processes for 2D sensor placement.

The two dimensional error regions discussed by Kelly [4] provide justification for considering the real-world projection of detection error, as the area of this error changes for different sensor placements.

The work presented here seeks to expand upon this existing literature by expanding the placement question beyond merely cameras and into camera and projector systems.

III. Problem Formulation

The purpose of this paper is to classify different camera and projector placements around an example target object by how good they are for structured light-based reconstruction. There are two competing definitions for what constitutes good detection in a structured light system. The first metric, visibility, is concerned with how much of the target object is visible at any given point in time. The second metric, scale, is concerned with how precisely points on the target object can be detected. To examine the tradeoff between these metrics, consider the following limit cases. When the camera and projector are very far away from the object, the entire object is contained within the field-of-view, but points on the surface are not distinguishable from each other. Alternatively, when the camera and projector are as close as

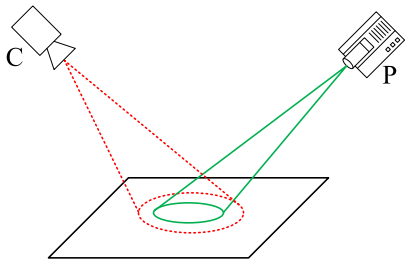

Fig. 1. The intuition behind the quality metric presented in this paper. Illuminating one point on the projector's image plane projects a cone of light into the scene. This cone of light intersects the scene in an ellipse shape (shown in green). The camera detects this illuminated ellipse imperfectly, making the real-world error bound the dotted red ellipse shown.

possible to the surface of the object, points on the surface are distinguishable, but not many of them are within the images.

The *visibility metric* quantifies how much of the target object is visible. When the target object is represented as a set of surface points, the visibility metric for a given camera and projector setup is then the percentage of these points visible to at least one camera and at least one projector. A point can be considered visible when its projection is within the width and height of the camera or projector's image plane.

The *scale metric* quantifies the accuracy of detection for each of the visible points. The intuition behind this metric is shown in Fig. 1. Consider a single pixel being illuminated on a projector's image plane. This point illuminates a cone which spreads out into the world from the projector. If this cone intersects an object that is locally planar about the area of intersection, an ellipse is illuminated on the surface of the object. This process is shown in green in Fig. 1. To perform reconstruction, a camera must then observe this illuminated ellipse. However, the camera cannot detect this perfectly–there is noise associated with the camera detecting the illuminated point, essentially enlarging the area in which the ideal point could be in the world, shown in red in Fig. 1. The error is additive due to there being only one sensor, a major difference between camera and projector systems and stereo camera systems.

Note that the model assumes the target object to be locally planar about the area where it intersects with the illumination cone. This is reasonable as this area is on the scale of a single pixel being illuminated.

For clarity, we will now walk through how the scale quality metric is computed for a given camera and projector position. A graphical depiction of this process is shown in Fig. 2. In this case, the camera and projector are arranged relative to a single point of interest, as shown in Fig. 2(a). The point of interest is illuminated by a single point on the projector image plane, shown in Fig. 2(b). This point of interest is also visible as a single point on the camera image plane, shown in Fig. 2(c).

The points in each of the projector and camera image planes have error circles around them with the diameter of a single pixel in their respective planes. These are shown in Fig. 2(d) and Fig. 2(e).

The target object is assumed to be planar in the area immediately surrounding the target point. Thus, the error circles on the camera and project image planes can then both be projected to ellipses on the target point tangent plane as shown in Fig. 2(f).

To find how well the camera can detect the illuminated ellipse from the projector, the underlying Gaussian distribution for each of the projected ellipses must be found, as shown in Fig. 2(g). These distributions are then convolved, giving the overall distribution shown in Fig. 2(h). The semimajor and semiminor axis lengths are then extracted from this combined distribution, providing real-world bounds on the scale of the error for detecting the target point from the camera and projector setup given.

Note that the propagation of the camera and projector error ellipses is fundamentally different than how such error propagates in multiple camera systems. In multiple camera systems, an intersection operation would be used on the error ellipses on the target point tangent plane (Fig. 2(f)). This is because multiple cameras are multiple sensors, and uncertainty decreases with multiple measurements. However, as only the camera in a camera and projector system is a sensor, only one actual measurement is taking place. Thus, the error from the projector adds to the error in sensing, so the total error increases instead of decreasing. Thus, the error ellipses must be convolved.

IV. Problem Mechanics

A. Camera Parameters

A camera can be defined by its intrinsic and extrinsic parameters. A camera's intrinsic parameters are described by the matrix $\mathbf{K}$, which represents the transformation from the camera's 3D frame to the image plane of the camera. The form of this matrix is shown in Eqn. (1). The camera's 3D frame has the z-axis perpendicular to the image plane, pointing out from the camera. The parameters of $\mathbf{K}$ are as follows. The scale factors along the x- and y-axis are α_x and α_y. The camera skew is s, which should be zero for most cameras. If s is not zero, it means that the x-axis and y-axis of the elements on the camera sensor are not perpendicular. Finally, the principal point of the image is $(x_0, y_0)^\top$, which is essentially the offset from $(0,0)^\top$ to the center of the image.

$$\mathbf{K} = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{1}$$

The camera's extrinsic parameters describe the camera's location within the greater world. This is represented as a rotation matrix $\mathbf{R}$ and a translation vector $\mathbf{t}$. Together, $\mathbf{R}$ and $\mathbf{t}$ transform from the frame of reference for the entire world to the frame of reference for the camera.

A projection matrix $\mathbf{P}$ can be defined which will map homogenous world points to homogenous points in the image, as shown in Eqn. (2), where the vector $\mathbf{0} = (0,0,0)^\top$.

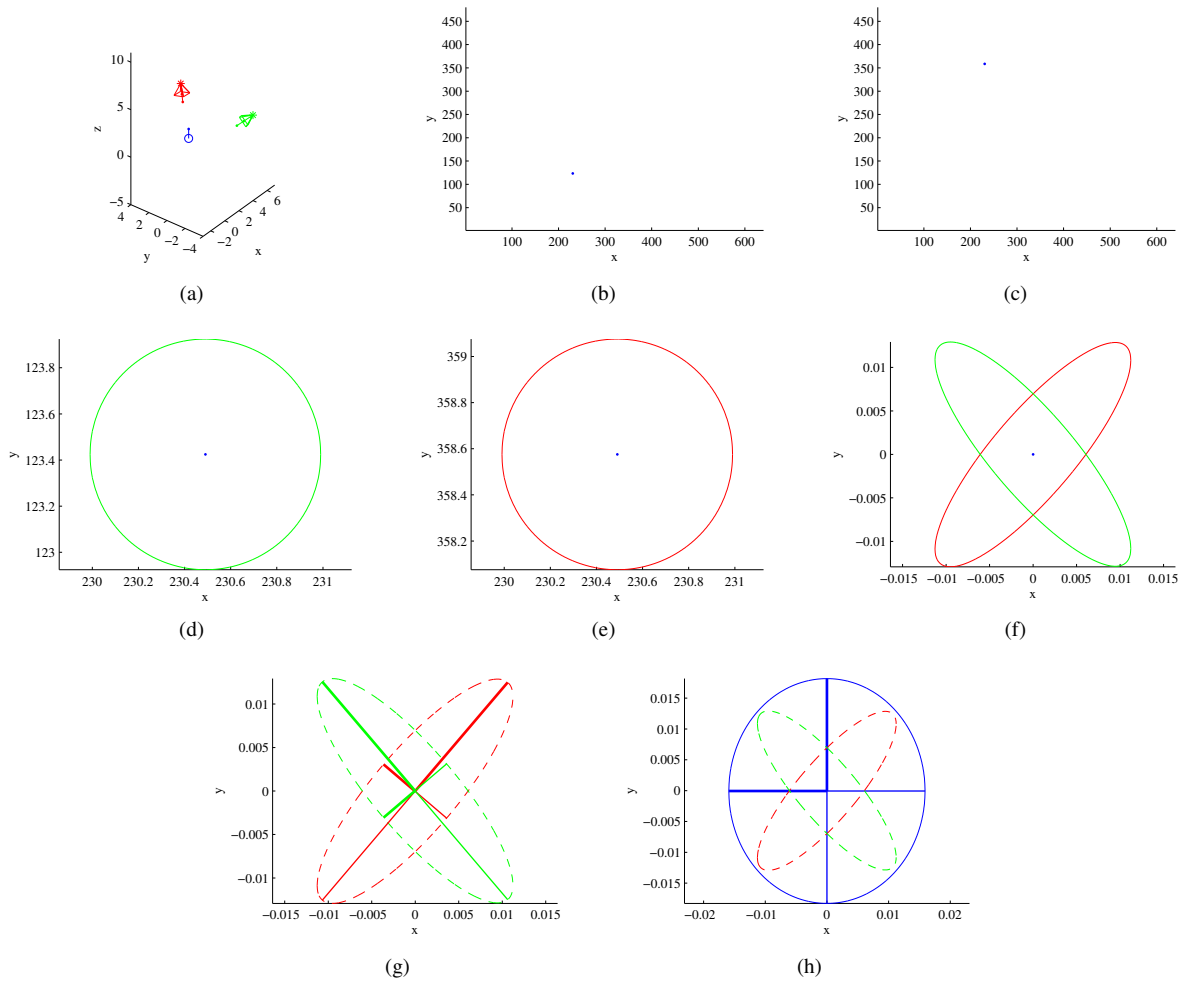

Fig. 2. (a) Setup of the camera and projector system. The projector is shown in green, the camera is shown in red, and the only target point is shown in blue with an arrow representing its normal vector. (b) Target point location on the projector image plane. (c) Target point location on the camera image plane. (d) Half-pixel width error ellipse around the target point location on the projector image plane. (e) Half-pixel width error ellipse around the target point location on the camera image plane. (f) Error ellipses from the projector (green) and camera (red) projected onto the target point tangent plane. (g) Gaussian distributions associated with the error ellipses on the target point tangent plane. (h) Final convolved Gaussian distribution on the target point tangent plane.

$$\mathbf{P} = \begin{bmatrix} \mathbf{K} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix}. \tag{2}$$

B. Projector Parameters

A projector can be described by the same parameters used to describe a camera. A projector has an intrinsic parameter matrix $\mathbf{K}$, as well as extrinsic parameters described by a rotation matrix $\mathbf{R}$ and translation vector $\mathbf{t}$. Keep in mind, however, that while the equations are the same, the physical process works in reverse for the projector. Instead of the scene being projected upon the image plane, as is the case with a camera, the image plane is being projected onto the scene. This distinction is important for determining the measurement error for a given structured light system.

C. Target Point Parameters

It is assumed that a representative target is given, which can be used to determine the quality of the placement of the camera(s) and projector(s). In this case, the target is recorded as a set of 3D points in the world coordinate system, $\{p_i\}$, and a set of surface normal vectors associated with those points, $\{n_i\}$. All points must lie on a convex surface.

D. Determining Visibility of Target Points

For any of the points p_i, there are two steps to finding out if it is visible to a single camera or projector (the process is identical). The first is to ensure that the point in the world projects to a point within the bounds of the camera/projector image plane. The projection of p_i onto the image plane of the camera/projector with projection matrix $\mathbf{P}$ is the point p_{Ii}. This point can be found as shown in Eqn. (3). It is within the bounds of the image plane when $1 \leq \frac{p_{Ii}(1)}{p_{Ii}(3)} \leq \alpha_x$ and $1 \leq \frac{p_{Ii}(2)}{p_{Ii}(3)} \leq \alpha_y$.

$$p_{Ii} = \mathbf{P} \begin{pmatrix} p_i \\ 1 \end{pmatrix}. \tag{3}$$

The angle of the normal vector of the target point and the camera must be checked to ensure visibility. If l_i is the location of the camera/projector in the world, we find $v_i = l_i - p_i$. Then, the angle β can be found as shown in Eqn. (4). The point is visible if $\beta < \frac{\pi}{2}$. This ensures that the normal of the target point is not directed away from the camera, which would mean the surface at that point is not visible.

$$\beta = \cos^{-1}\left(\frac{n_i \cdot v_i}{\|n_i\|_2 \|v_i\|_2}\right). \tag{4}$$

Note that for a point p_i to be considered visible by a coded structured light system, it must be visible by at least one camera *and* at least one projector. Also note that self-occlusions are not considered by this visibility model. If the target points lie on a non-convex surface, a more robust method must be used to determine point visibility.

E. Visibility Quality Metric

The visibility quality metric is the ratio of visible points ($n_{visible}$) on the target object to the total number of points on the object (n_{total}). This ratio is shown in Eqn. (5). In this case, a higher ratio is better, but it will always be less than one.

$$q_{visible} = \frac{n_{visible}}{n_{total}}. \tag{5}$$

The quality measure $q_{visible}$ records the percentage of the points on the object that are visible. However, it does not encode how finely those points are detected.

F. Homography Matrix

A homography matrix between the plane around a target point and a camera can be found as shown in Eqn. (6). The target point is p_i and its surface normal is n_i. Also, $\mathbf{R_T}$ is the rotation matrix between the target plane's frame of reference and that of the world frame. The matrices $\mathbf{K}$, $\mathbf{R}$, and the vector $\mathbf{t}$ are the camera's parameters.

$$\mathbf{H} = \begin{bmatrix}\mathbf{K} & \mathbf{0}\end{bmatrix} \begin{bmatrix}\mathbf{R} & \mathbf{t} \\ \mathbf{0}^\top & 1\end{bmatrix} \begin{bmatrix}\mathbf{R_T} & p_i \\ \mathbf{0}^\top & 1\end{bmatrix} \begin{bmatrix}1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1\end{bmatrix}. \tag{6}$$

The homography matrix $\mathbf{H}$ maps from the plane around p_i to the image plane of the camera, as shown in Eqn. (7). Here, p_T is a point in the target plane and p_I is the corresponding point in the camera's image plane.

$$p_I = \mathbf{H} p_T. \tag{7}$$

G. Ellipses

Ellipses can be represented two ways. The first is the parametric form, in which five parameters are used to describe the ellipse: the ellipse center (h, k), the semimajor axis length a, the semiminor axis length b, and the angle with the x-axis ϕ (Fig. 3).

The other representation of an ellipse is the general form, as shown in Eqn. (8). Note this is the general form for a conic section.

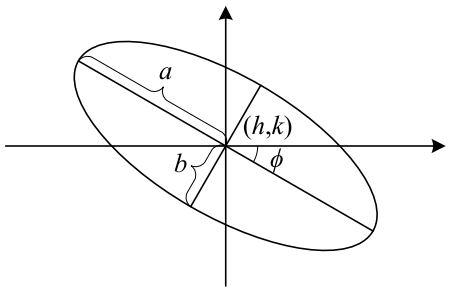

Fig. 3. Depiction of an ellipse with its center (h, k), semimajor axis length a, semiminor axis length b, and angle ϕ labeled.

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0. \tag{8}$$

1) Parametric Form to General Form: One possible set of equations for the transformation from the set of ellipse parameters to the general form is shown below. Space is limited for these equations, so $c\phi = \cos\phi$, and $s\phi = \sin\phi$.

$$A = \frac{c^2\phi}{a^2} + \frac{s^2\phi}{b^2} \tag{9}$$

$$B = 2\left(\frac{1}{b^2} - \frac{1}{a^2}\right) c\phi s\phi \tag{10}$$

$$C = \frac{s^2\phi}{a^2} + \frac{c^2\phi}{b^2} \tag{11}$$

$$D = \left(\frac{2k}{a^2} - \frac{2k}{b^2}\right) c\phi s\phi - 2h\left(\frac{c^2\phi}{a^2} + \frac{s^2\phi}{b^2}\right) \tag{12}$$

$$E = 2h\left(\frac{1}{a^2} - \frac{1}{b^2}\right) - 2k\left(\frac{s^2\phi}{a^2} + \frac{c^2\phi}{b^2}\right) \tag{13}$$

$$F = \left(\frac{h^2}{a^2} + \frac{k^2}{b^2}\right) c^2\phi + 2hk\left(\frac{1}{b^2} - \frac{1}{a^2}\right) c\phi s\phi + \left(\frac{k^2}{a^2} + \frac{h^2}{b^2}\right) s^2\phi - 1. \tag{14}$$

2) General Form to Parametric Form: The set of equations for the transformation from the general form of an ellipse to the parametric form is shown below. This is for the case in which $B \neq 0$.

$$\phi = \frac{1}{2}\tan^{-1}\left(\frac{B}{C - A}\right) \tag{15}$$

$$a = \sqrt{\frac{2|F|\sin(2\phi)}{(A + C)\sin(2\phi) - B}} \tag{16}$$

$$b = \sqrt{\frac{2|F|\sin(2\phi)}{(A + C)\sin(2\phi) + B}} \tag{17}$$

$$h = \frac{BE - 2CD}{4AC - B^2} \tag{18}$$

$$k = \frac{2AE - DB}{B^2 - 4AC}. \tag{19}$$

If $B = 0$, then the following equations for a and b must be used instead.

$$a = \frac{1}{\sqrt{A}} \tag{20}$$

$$b = \frac{1}{\sqrt{C}}. \tag{21}$$

H. Discussion of Gaussian Distributions

An ellipse can be considered a level set of a Gaussian distribution. Thus, we can find the Gaussian distribution associated with a given ellipse, and vice-versa.

1) Ellipse Parameters to Gaussian Distribution: A fixed Gaussian distribution is represented by a vector mean μ and a covariance matrix $\mathbf{\Sigma}$. These can be found from the parameters as follows. Note here that $c\phi = \cos\phi$, and $s\phi = \sin\phi$.

$$\mu = \begin{pmatrix} h \\ k \end{pmatrix} \tag{22}$$

$$\mathbf{\Sigma} = \begin{bmatrix} c\phi & s\phi \\ -s\phi & s\phi \end{bmatrix} \begin{bmatrix} a^2 & 0 \\ 0 & b^2 \end{bmatrix} \begin{bmatrix} c\phi & -s\phi \\ s\phi & c\phi \end{bmatrix}. \tag{23}$$

2) Gaussian Distribution to Ellipse Parameters: The parameters ϕ, a, and b require us to find the eigenvalues and eigenvectors of $\mathbf{\Sigma}$. The eigenvalues are λ_1 and λ_2, where $\lambda_1 \geq \lambda_2$. The associated eigenvectors are $\mathbf{e}_1$ and $\mathbf{e}_2$.

$$\phi = -\tan^{-1}\left(\frac{\mathbf{e}_1(1)}{\mathbf{e}_1(2)}\right) \tag{24}$$

$$a = \sqrt{\lambda_1} \tag{25}$$

$$b = \sqrt{\lambda_2} \tag{26}$$

$$h = \mu(1) \tag{27}$$

$$k = \mu(2). \tag{28}$$

3) Convolution of Gaussian Distributions: Convolving Gaussian distributions is very useful, and can be accomplished with the following equations. The convolved Gaussian distribution is represented by mean μ_c and covariance $\mathbf{\Sigma}_c$.

$$\mu_c = \mu_1 - \mu_2 \tag{29}$$

$$\mathbf{\Sigma}_c = \mathbf{\Sigma}_1 + \mathbf{\Sigma}_2. \tag{30}$$

I. Projection of Ellipses

Taking an ellipse in the form of Eqn. (8), a matrix representation of the ellipse, $\mathbf{C}$, can be created as shown in Eqn. (31).

$$\mathbf{C} = \begin{bmatrix} A & B/2 & D/2 \\ B/2 & C & E/2 \\ D/2 & E/2 & F \end{bmatrix}. \tag{31}$$

Once the ellipse is in this matrix form, the projection of the ellipse onto another plane can be found using the homography matrix calculated in Eqn. (6). For instance, Eqn. (32) shows the projection of an ellipse in the camera image plane ($\mathbf{C}_I$) onto the target point tangent plane ($\mathbf{C}_T$).

$$\mathbf{C}_T = \mathbf{H}^\top \mathbf{C}_I \mathbf{H}. \tag{32}$$

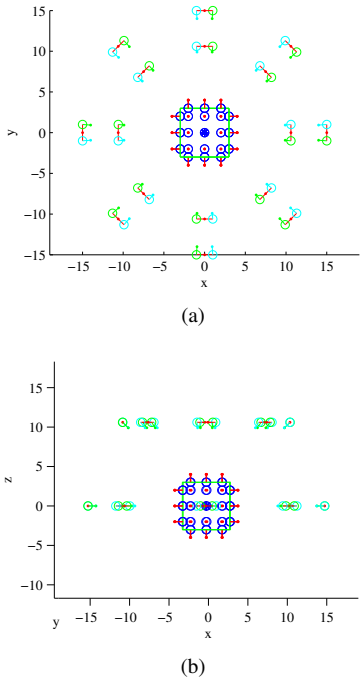

Fig. 4. Two views of the set of camera/ projector placements tested, seen around the cubic target object.

J. Scale Quality Metric

The scale quality metric encodes the scale at which the points that are visible is detected. Here, i iterates over all visible points. The area of an ellipse is $ab\pi$, where a and b are the semimajor and semiminor axis lengths. Thus, if a_c and b_c are the semimajor and semiminor axis of the convolved Gaussian distribution, we get the equation for q_{scale} shown in Eqn. (33).

$$q_{scale} = \frac{\pi}{n_{visible}} \sum_{i=1}^{n_{visible}} a_{ci} b_{ci}. \tag{33}$$

The quality metric q_{scale} is in world units. Thus, it can be used if resolution at a desired scale can be observed and tracked by the system. The best camera and projector placement is the one in which q_{scale} is minimized.

K. Multiple Cameras and/or Projectors

For situations where multiple cameras and/or projectors are utilized, it should be generally assumed that the worst-case scale and visibility metrics calculated be used for each point on the target object. This is because it is not possible to tell just from the projected patterns and corresponding camera images in the real system which image or pattern is better than another, especially if all projectors and cameras operate simultaneously. This being the case, the error has to be assumed to be the worst possible when scoring a particular placement setup.

V. Example

The quality of a set of camera and projector pair placements around a cubic target object are examined in this

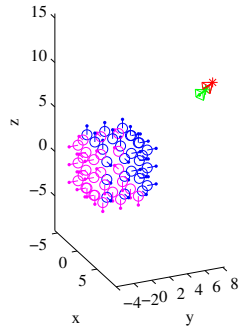

Fig. 5. The camera/ projector placement tested with the best $q_{visible}$ score. Visible target points are shown in blue, nonvisible points in magenta. Note that three faces (50%) of the cube are visible.

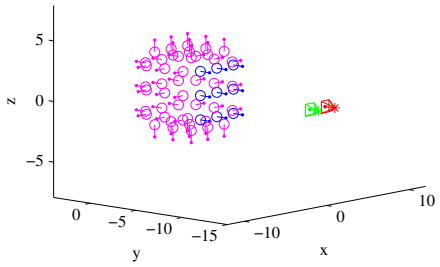

Fig. 6. The camera/ projector placement tested with the best q_{scale} score. Visible target points are shown in blue, nonvisible points in magenta. Here, only one face of the cube is visible, as this minimizes the size of the error projections.

example. The target points chosen are regularly distributed across the faces of the cube. For all camera and projector placements, the center point between the camera and projector is a constant distance from the center of the cube and the camera and projector are a constant distance from each other. The camera and projector point towards the center of the cube. The purpose of this example is to show which of the camera and projector pair positions has the best $q_{visible}$ score and which has the best q_{scale} score.

Two views of the set of examined camera/projector pairs around the target object are shown in Fig. 4. Each of these views was examined to determine which view was best in terms of each of the quality metrics.

The camera/projector pair location with the best $q_{visible}$ score is shown in Fig. 5. Note that three faces of the cube are visible to both the camera and projector in this setup. This placement makes 50% of the cube visible, the maximum amount possible. Thus, this camera and projector placement maximizes the visibility metric.

Finally, the camera/projector pair location with the best q_{scale} score is shown in Fig. 6. Here, only one face is visible to the camera and projector. This is reasonable because if multiple faces are visible, the error ellipses on the faces are longer due to the grazing angles between the cube faces and the camera/projector, leading to larger error bounds. Thus, a view of a single face reduces the average error per point.

VI. Conclusions

In this paper, a mathematical basis for judging the quality of camera and projector placement for structured light systems was presented. Two metrics of quality must be considered: visibility, which measures how much of the target object is visible, and scale, which measures how well the parts that are visible can be seen. Methods for computing each of these metrics were presented. An example was shown which demonstrates these quality metrics.

The method presented here is useable in domains containing multiple cameras and multiple projectors. In addition, differing resolutions between the cameras and projectors are taken into account. An application where placement quality is important is patient body tracking, in which cameras and projectors must be placed such that a patient's body can be reconstructed to a desired precision.

Future directions include addressing self-occlusion in the target object, which requires a more robust visibility model. This may be possible using something similar to the visibility regions discussed by Tarabanis *et al.* [8]. In addition, while it is expected that the relative importance of the two quality metrics is dependent on application, the optimal trade-off between them must still be examined in depth.

Using the method presented here, the quality of competing camera and project placements in structured light systems can now be mathematically compared. No longer must guesswork be used in placing cameras and projectors for structured light systems.

VII. Acknowledgements

This work was supported by NSF #CNS-0821474, by the Medical Devices Center at the University of Minnesota, and by the Digital Technology Center at the University of Minnesota.

References

[1] J. Batlle, E. M. Mouaddib, and J. Salvi. Recent progress in coded structured light as a technique to solve the correspondence problem. *Pattern Recognition*, 31(7):963–982, 1998.

[2] R. Bodor, A. Drenner, P. Schrater, and N. Papanikolopoulos. Optimal camera placement for automated surveillance tasks. *Journal of Intelligent and Robotic Systems*, 50:257–295, 2007.

[3] X. Chen and J. Davis. Camera placement considering occlusion for robust motion capture. Technical Report CS-TR-2000-07, Stanford Department of Computer Science, 2007.

[4] A. Kelly. Precision dilution in triangulation-based mobile robot position estimation. In *Intelligent Autonomous Systems*, 2003.

[5] A. Krause, A. Singh, and C. Guestrin. Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies. *Journal of Machine Learning Research*, 9:235–284, February 2008.

[6] A. Mittal and L. S. Davis. A general method for sensor planning in multi-sensor systems: Extension to random occlusion. *International Journal of Computer Vision*, 76(1):31–52, 2008.

[7] J. Salvi, J. Pagés, and J. Batlle. Pattern codification strategies in structured light systems. *Pattern Recognition*, 37(4):827–849, April 2004.

[8] K. Tarabanis, R. Tsai, and A. Kaul. Computing occlusion-free viewpoints. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(3):279–292, March 1996.

[9] O. Tekdas and V. Isler. Sensor placement algorithms for triangulation based localization. In *IEEE International Conference on Robotics and Automation*, pages 4448–4453, April 2007.