A Fuzzy Model for Coverage Evaluation of Cameras and Multi-Camera Networks

Aaron Mavrinac
mavrin1@uwindsor.ca

Jose L. Alarcon Herrera
alarconj@uwindsor.ca

Xiang Chen
xchen@uwindsor.ca

Department of Electrical and Computer Engineering
University of Windsor
401 Sunset Ave., Windsor, Ontario, Canada N9B 3P4

ABSTRACT

A comprehensive, intuitive, task-oriented three-dimensional coverage model for cameras and multi-camera networks using fuzzy sets is presented. The model captures the vagueness inherent in the concept of visual coverage. At present, the model can be used to evaluate, given a scene model and an objective, the coverage performance of a given camera or multi-camera network configuration, as a single numerical metric. Plans to use the model for optimal camera placement and other problems involving coverage are discussed. Examples of qualitative experimental validation of the coverage model are presented.

1. INTRODUCTION

In computer vision, the property of coverage of a given point in space by a camera or multi-camera network is vague. A point in space is normally considered covered, for a specific task, if useful information can be obtained for that task from the point, assuming such information exists in the scene. In most applications, visibility is not enough for a point to be covered; we require a certain level of “quality” in the imaging, which is generally imprecise. High-level computer vision applications are complex. Even when performance metrics for the final results are well-defined, it is usually impractical or impossible to trace these analytically, through the myriad factors imparted by various low-level algorithms and scene conditions, back to a set of requirements and/or optima for the parameters of the imaging system. As a result, these parameters are often designed by educated guessing, sometimes partially informed by application performance metrics, often involving much trial and error. There is no analytic indicator function for coverage.

A fuzzy representation, we contend, lends itself well to modeling coverage. The membership degree of a scene point in a fuzzy set representing a camera’s volume of coverage can concisely encapsulate the various underlying factors into a single metric, yielding a useful model for evaluating the coverage of such points. A number of these sets can then be transformed and combined in appropriate ways to yield an equivalent coverage model for a multi-camera network of a given type.

We present the fuzzy coverage model, a comprehensive three-dimensional coverage model for cameras and multi-camera networks, based on fuzzy sets, which is derived from well-studied parameters of the standard imaging model (obtained from internal and external calibration), and can be tuned for specific tasks via a small number of intuitive application parameters. The single-camera model can be seen as an adaptation of work on task-oriented, constraint-based modeling of camera coverage by Cowan and Kovesi [5], Stamos and Allen [17], Reed and Allen [16] and others to a more powerful and flexible representation with fuzzy sets, which itself is reminiscent of the attenuated disk coverage models found in the sensor network literature [19]. The extension to the multi-camera context has parallels to work by Mittal and Davis [14].

By providing a model of the desired coverage volume, a realistic and quantitative performance metric can be obtained. Besides evaluating the coverage of existing or proposed camera network deployments, the obvious application is as an objective model for optimal camera placement (sensor planning for multi-camera networks). Previous models used for this purpose by Erdem and Sclaroff [6], Hörster and Lienhart [9], Angella et al. [1], Zhao et al. [22], and others have used bivalent models (i.e., points are either covered or not covered) and do not account for the direction of observation. Our ambition is to develop optimization tools capable of taking advantage of the richer information in the fuzzy coverage model to achieve various goals. Since the model is rather complex and has not been explicitly designed for use in any specific optimization method, an analytic approach will involve reducing the model to various simpler forms focusing on the variables of interest. The scalar performance metric we propose would already allow for a generate-and-test approach similar to Yi et al. [20].

The primary contributions of this work are the concept of a coverage model based on fuzzy subsets of directional Euclidean space, a full parameterization of this model using well-studied models from computer vision and intuitive application parameters, and the use of the model to obtain a scalar performance metric of coverage in multi-camera networks. It is intended as a stepping stone toward the solution of a number of specific multi-camera network problems, as well as a better theoretical understanding of camera networks.
The remainder of this paper is organized thus. Section 2 introduces our geometric space and its properties, as well as some basic definitions and notation for fuzzy sets. Section 3 reviews the well-studied geometric and optical model of the camera, providing the basis for our realization of the single-camera coverage model. Section 4 investigates a few additional modeling concerns specific to the multi-camera paradigm. A full mathematical description of the fuzzy coverage model is given in Section 5; this is the main contribution of this work. In Section 6, we present some qualitative experimental validation of the model against real images. Section 7 discusses application particulars and directions for future work. Finally, we give some concluding remarks in Section 8.

2. DEFINITIONS AND CONVENTIONS

2.1 Directional Space

We express positions and orientations in three-dimensional Euclidean space in a right-handed Cartesian coordinate system, with the axes of its basis denoted \( x, y, \) and \( z, \) and fixed-axis rotation angles about these axes denoted \( \theta, \phi, \) and \( \psi, \) respectively. For every equivalence class on angles \( \bar{x}_2 \) we consider the value \( \theta \in [0, 2\pi). \)

For direction – separate from the “direction” of the vector in \( \mathbb{R}^3 \) – we also employ a two-angle orientation defined by inclination angle \( \rho \in [0, \pi], \) measured from the \( z \)-axis zenith, and azimuth angle \( \eta \in [0, 2\pi], \) measured from the \( x \)-axis in the \( x-y \) plane (in the direction of \( \psi). \) Note that \( (0, \eta_1) \sim (0, \eta_2) \) and \( (\pi, \eta_1) \sim (\pi, \eta_2) \) for any \( \eta_1 \) and \( \eta_2. \)

**Definition 1.** The directional space \( \mathbb{D} = \mathbb{R}^3 \times [0, \pi] \times [0, 2\pi) \) consists of three-dimensional Euclidean space plus direction, with elements of the form \( (x, y, z, \rho, \eta). \)

For convenience, we denote the spatial component \( p_s = (x, y, z) \) and the directional component \( p_d = (\rho, \eta). \)

A standard 3D pose \( P : \mathbb{R}^3 \to \mathbb{R}^3, \) consisting of rotation matrix \( R \) and translation vector \( T, \) may be applied to \( P \in \mathbb{D}. \) The spatial component is transformed as usual, i.e. \( P_p = Rp_s + T. \) The direction is transformed as follows. If \( d \) is the unit vector in the direction of \( p_d, \) then \( P(d) = (\arccos(R_{d,i}p_s), \arctan(2(R_{d,j}R_{d,k}))) \), where \( \arctan(2) \) is an arc tangent variant returning the angle from the \( x \)-axis to a vector in the full range \( [0, 2\pi). \)

2.2 Fuzzy Sets

A fuzzy set [21] is a pair \( \langle S, \mu \rangle, \) where \( S \) is a set (called the *universal set*) and \( \mu : S \to [0, 1] \) is a membership function indicating the grade of membership of elements in \( S \) to the fuzzy set. For a fuzzy set \( A = \langle \mu_A \rangle, \) the set \( \operatorname{supp}(A) = \{x \in S_A | \mu_A(x) > 0\} \) is called the support of \( A, \) and the set \( \operatorname{ker}(A) = \{x \in S_A | \mu_A(x) = 1\} \) is called its kernel.

The standard fuzzy union operation is defined for fuzzy sets \( A \) and \( B \) as \( A \cup B = \langle \mu_A \cup \mu_B \rangle, \) where \( \mu_{A \cup B}(x) = \max(\mu_A(x), \mu_B(x)) \) for \( x \in S_A \cup S_B. \) Similarly, the standard fuzzy intersection operation is defined as \( A \cap B = \langle \mu_A \cap \mu_B \rangle, \) where \( \mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x)) \) for \( x \in S_A \cap S_B. \) Other t-norms and t-norms yield different union and intersection operations [11]; in particular, we also use the algebraic product t-norm for intersection, for which \( \mu_{A \cup B}(x) = \mu_A(x) \cdot \mu_B(x). \)

In modeling the coverage of cameras and multi-camera networks, we deal with fuzzy subsets of \( \mathbb{R}^3 \) and \( \mathbb{D}. \)

3. SINGLE-CAMERA MODEL

3.1 Coordinate Systems

The camera frame is the coordinate system of the 3D scene from the camera’s point of view, which projects directly to the image. The positive \( z \)-axis extends out along the principal axis, and the \( x \) and \( y \) axes are parallel to the \( u \) and \( v \) axes, respectively, of the image plane. The origin is situated in the principal plane spanned by the \( x \) and \( y \) axes. The image plane is the 2D projected image of the scene. The point where the principal axis intersects this plane is called the principal point \( p_p. \) A world frame coordinate system, defined by a rigid 3D Euclidean transformation (the camera extrinsics) to the camera frame, is often considered. Here, this will be deferred to the multi-camera context.

We assume that internally calibrated cameras rectify the image in the \( u-v \) plane according to e.g. Brown’s lens distortion model [3].

3.2 Field of View

According to the pinhole camera model [7], if \( p = (x, y, z) \) is a point in the scene, in camera frame coordinates, its image plane projection \( p' = (u, v) \) is given by

\[
\begin{align*}
    u &= \frac{f}{s_u} x + o_u \\
    v &= \frac{f}{s_v} y + o_v
\end{align*}
\]

where \( s_u \) and \( s_v \) are the effective pixel dimensions in units of distance, and \( o_u \) and \( o_v \) are the pixel coordinates of \( p_p. \)

For a rectilinear projected image, the angles of view are given by

\[
\begin{align}
    \alpha_{hl} &= 2 \arctan \frac{o_u s_u}{2f} \\
    \alpha_{hr} &= 2 \arctan \frac{(w - o_u)s_u}{2f} \\
    \alpha_{vl} &= 2 \arctan \frac{o_v s_v}{2f} \\
    \alpha_{vr} &= 2 \arctan \frac{(h - o_v)s_v}{2f}
\end{align}
\]

where \( w \) and \( h \) are the image (sensor) width and height in pixels. A scene point \( p = (x, y, z) \) is within the field of view if \( z \sin \alpha_{hl} < x < z \sin \alpha_{hr} \) and \( z \sin \alpha_{vl} < y < z \sin \alpha_{vr} \) (implying \( z > 0. \)) For convenience, we define \( \alpha_h = \alpha_{hl} + \alpha_{hr} \) and \( \alpha_v = \alpha_{vl} + \alpha_{vr}. \)

3.3 Resolution

*Pixel resolution* (number of pixels per unit distance) is a function of the distance along the principal axis. The horizontal and vertical resolutions for a given scene depth \( z \) are given by

\[
\begin{align}
    R_h(z) &= \frac{1}{z} \frac{w}{2 \sin(\alpha_h/2)} \\
    R_v(z) &= \frac{1}{z} \frac{h}{2 \sin(\alpha_v/2)}
\end{align}
\]

The largest distance \( z_R \) before which resolution falls below a given \( R \) can then be computed:

\[
    z_R = \frac{1}{R} \min \left[ \frac{w}{2 \sin(\alpha_h/2)}, \frac{h}{2 \sin(\alpha_v/2)} \right]
\]
4.2 Occlusion parameters

In contrast, it is relatively easy to determine whether a particular \( p \in \mathbb{R}^3 \) is occluded, from the point of view of a camera whose parameters are known, given a scene model \( S \) consisting of a set of plane segments (which represent opaque surfaces in the scene). The point of intersection \( p_0 \) between the line from \( p \) to \( p_b \) of the camera and a plane can be found via well-known means; \( p \) is occluded iff \( p_b \) exists, is unique, is not \( p \), and lies within the plane segment, for any plane segment in \( S \).

4.3 Feature Matching

A significant proportion of multi-camera network applications, particularly those involving 3D reconstruction, rely heavily on matching local image features. A variety of feature detection and description algorithms exist, but a major common limitation is degradation of performance over large rotational transformations of the viewpoint [2, 8, 13, 15]. Accordingly, directional visibility (discussed in Section 3.5) has implications beyond accounting for self-occlusion in such applications.

5. THE FUZZY COVERAGE MODEL

5.1 Single Camera

Consider a fuzzy subset \( C \subset \mathbb{D} \), with \( \mathbb{D} \) in the camera frame basis, where membership grade \( \mu_C(p) \) indicates coverage at point \( p \) from direction \( p_b \).

Our realization of \( C \) is entirely determined from nine intrinsic parameters of the imaging system \( (A, f, s_u, s_v, o_u, o_v, w, h, z_s) \) and five intuitive application parameters \( (\gamma, R_1, R_2, c_{max}, \varsigma) \) for a total of 14 parameters. The intrinsic parameters are normally obtained from camera specifications and calibration, and the application parameters are set based on task requirements. Four individual factors – visibility, resolution, focus, and direction – yield component fuzzy subsets of \( \mathbb{D} \), individually defined in Sections 5.1.1 through 5.1.4.

These components are combined via algebraic product fuzzy intersection:

\[
C = C_V \cap C_R \cap C_F \cap C_D
\]

Conveniently, if it is not desirable to account for a component in a certain application, it can simply be left out of (13). In the particular case where directional visibility is not relevant (thus \( C = C_V \cap C_R \cap C_F \)), \( C \) can be treated as a fuzzy subset of \( \mathbb{R}^3 \) rather than of \( \mathbb{D} \) for practical purposes, since, as will be seen, the directional components have no effect.

5.1.1 Visibility

Visibility is a bivalent condition depending whether \( p \) is within the field of view, a function of \( x, y \), and \( z \). The indicator functions are discussed in Section 3.2. However, points near the edges may not be considered “fully visible” for an application’s purposes, so a general-purpose application parameter \( \gamma \) is introduced to fuzzify visibility into a trapezoidal membership function. The value of \( \gamma \) is simply the number of pixels inward from the edge of the image at which features are considered to be fully within the field of view.

Figures 1 and 2 show one-dimensional cross-sections of the fuzzy subset for visibility \( C_V \), in terms of \( x \) and \( y \), respectively, normalized with respect to \( z \). The set has kern(\( C_V \) =
{p \in \mathbb{R} | -\sin(\alpha_{hl})+\gamma_h \leq p_x \leq \sin(\alpha_{hl})-\gamma_h, -\sin(\alpha_{vr}) - \gamma_v \leq p_x \leq \sin(\alpha_{vr}) + \gamma_v, p_x > 0} \text{ and } \text{supp}(C_R) = \{p \in \mathbb{R} | -\sin(\alpha_{vl}) \leq p_x \leq \sin(\alpha_{vl}), -\sin(\alpha_{vl}) \leq p_y \leq \sin(\alpha_{vl}), p_x > 0\}, \text{ where } \gamma_h \text{ and } \gamma_v \text{ are calculated from } \gamma \text{ as follows:}

\begin{align*}
\gamma_h &= \frac{2\gamma}{w} \sin \left( \frac{\alpha_h}{2} \right) \quad (14) \\
\gamma_v &= \frac{2\gamma}{h} \sin \left( \frac{\alpha_v}{2} \right) \quad (15)
\end{align*}

**5.1.2 Resolution**

Resolution is inversely proportional to the z-coordinate of p. Figure 3 shows a one-dimensional cross-section of the fuzzy subset for resolution C_R, in terms of z. The set has kern(C_R) = \{p \in \mathbb{R} | 0 \leq p_z \leq z_1\} and supp(C_R) = \{p \in \mathbb{R} | 0 \leq p_z \leq z_2\}.

**5.1.3 Focus**

Focus is also a function of z. Figure 4 shows a one-dimensional cross-section of the fuzzy subset for focus C_F, in terms of z. The set has kern(C_F) = \{p \in \mathbb{R} | z_2 \leq p_z \leq z_0\} (see (11) in Section 3.4) and supp(C_F) = \{p \in \mathbb{R} | z_n \leq p_z \leq z_f\}.

The values of z_n and z_f are the near and far limits of depth of field for the maximum acceptable circle of confusion diameter, application parameter \(\epsilon_{\text{max}}\), as given by (10).

**5.1.4 Direction**

As shown in Section 3.5, directional visibility is a function of inclination angle \(\rho\). An application parameter \(\zeta\) is introduced to fuzzify directional visibility into a trapezoidal membership function, which can be used to account for the quality of coverage of an object face dependent on viewing angle, or for the ability to match features (see Section 4.3).

The fuzzy subset for direction C_D has kern(C_D) = \{p \in \mathbb{R} | p_x \geq \pi/2+\Theta+\zeta\} and supp(C_D) = \{p \in \mathbb{R} | p_y \geq \pi/2+\Theta\}, where

\[
\Theta = \left( \frac{p_y}{r} \sin p_x + \frac{p_x}{r} \cos p_x \right) \arctan \left( \frac{r}{p_z} \right) \quad (16)
\]

with \(r = \sqrt{p_x^2 + p_y^2}\) and for \(p_z > 0\) as per (12).

The value of \(\zeta\) is in the range \([0, \pi/2]\), and reflects the angle of the surface normal of a feature, relative to the principal axis, at which application performance begins to degrade.

**5.2 Simple Multi-Camera Network**

The simple multi-camera network is the case where the network’s coverage is defined simply as the combined coverage of the constituent cameras. This is the appropriate model for applications which perform no 3D reconstruction among the cameras, and therefore do not specifically require pairwise overlap.

**5.2.1 In-Scene Single-Camera Model**

The in-scene model for a single camera, \(C^s\), is simply the camera model \(C\) transformed to the world frame. Thus, \(C^s\) has six additional parameters – the pose, or extrinsic parameters, of the camera, defined by \(x, y, z, \theta, \phi,\) and \(\psi\) – for a total of 20 parameters.

**5.2.2 Adding the Occlusion Constraint**

Given the scene model \(S\), if \(V \subset \mathbb{R}^3\) is the set of all points visible to the camera – the complement of the set of all occluded points as defined in Section 4.2 – then the in-scene model with occlusion \(C^s_o = C^s \cap V\).

Network coverage for a simple multi-camera network is \(C_N = \bigcup_{k \in \mathbb{N}} C^o_k\).

**5.3 3D Multi-Camera Network**

The 3D multi-camera network is a more complex case where coverage depends on at least two cameras imaging the same point (including direction). This is the appropriate model for applications which perform 3D reconstruction or other pairwise processing of shared scene.

**5.3.1 Pairwise Coverage**

Because the 3D multi-camera network performs pairwise processing, we must consider the coverage models of pairs of cameras. The pairwise coverage model for cameras \(k\) and \(l\) is \(C^o_{kl} = C^o_k \cap C^o_l\), where \(C^o_k\) and \(C^o_l\) are defined as in the simple multi-camera network case (see Section 5.2.2).

Network coverage for a 3D multi-camera network is \(C_N = \bigcup_{k, l \in \mathbb{N}^2} C^o_{kl}\) (for \(k \neq l\).
5.3.2 Direction and Feature Matching

Conceptually, the membership degree of $p \in D$ in $C_N^p$ will depend heavily on the rotation component of the relative pose $P_{kl}$: the larger the angle between the principal axes of $k$ and $l$, the smaller the range of directions covered in common. The application parameter $\zeta$ thus plays an important role for applications using methods such as feature matching (see Section 4.3).

5.4 Coverage Performance

In order to evaluate scene coverage, a model $D$ of the desired coverage (the objective) is necessary. This takes the same form as $C_N$ – a fuzzy subset of $D$ – but rather than indicating the actual coverage, membership grade $\mu_D(p)$ indicates the importance of coverage at $p$. We define a scalar performance metric

$$m(C_N, D) = \frac{|C_N \cap \hat{D}|}{|D|}$$

(17)

where the form $\hat{A}$ denotes a discrete fuzzy set sampled from a continuous fuzzy set $A$ (i.e. $\mu_A(x) = \mu_A(x)$ for any $x \in \hat{A}$), and $|\hat{A}| = \sum_{x \in S} \mu_A(x)$ is the scalar cardinality of $\hat{A}$. In (17), $S_{C_N} = S_{D}$; that is, $C_N$ and $D$ are sampled on the same discrete "grid" of points in $D$.

6. EXPERIMENTAL VALIDATION

6.1 Apparatus

6.1.1 Software Implementation

The fuzzy coverage model and fuzzy vision graph have been implemented in an object-oriented software stack using Python. First, we have developed FuzzPy, a generic open-source Python library for fuzzy sets and graphs. Using this functionality, we have developed various classes for the fuzzy sets used in the model. The Camera class, initialized with the 14 model parameters, returns the $\mu$-value of any spatial point in camera coordinates using continuous trapezoidal fuzzy sets to implement $C$. The MultiCameraSimple and MultiCamera3D classes build their respective discrete fuzzy sets from Camera objects and a supplied scene model $S$. Coverage performance $m$ can be estimated given a discrete fuzzy set $D$.

6.1.2 Cameras

Prosilica EC-1350 cameras, with a sensor resolution of 1360 x 1024 pixels and square pixel size of 4.65 $\mu$m, are employed. These are fitted with Computar M3Z1228C-MP manual varifocal lenses, with a focal length range of 12 mm to 36 mm and a maximum aperture ratio of 1:2.8. The intrinsic and extrinsic camera parameters are found using HALCON.

6.2 Single-Camera Model Validation

We demonstrate that the fuzzy coverage model $C$, for various cameras and vision tasks, yields a useful and accurate metric for coverage. Quantitative validation would require a representative cross-section of possible single camera tasks – for example, various feature detection, object recognition, tracking, and motion analysis algorithms – to be tested for performance, and compared to predictions from a model $C$, built from application parameters based entirely on task requirements (see Section 7.1 for further discussion). This is beyond our scope at this stage; here, we demonstrate simple qualitative relationships between the fuzzy coverage model and the appearance of features in the associated image, showing that the model appears to reflect reality at least reasonably well.
Table 1: Estimated Coverage for Points in Fig. 7

<table>
<thead>
<tr>
<th>3D Point (x, y, z)</th>
<th>(p, η)</th>
<th>Model</th>
<th>Human</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0, 0)</td>
<td>(0, 0)</td>
<td>0.82</td>
<td>0.73</td>
</tr>
<tr>
<td>(12.5, -8.5, 265.0)</td>
<td>(π/4, π)</td>
<td>0.56</td>
<td>0.56</td>
</tr>
<tr>
<td>(67.0, -8.5, 398.0)</td>
<td>(π/4, 0)</td>
<td>0.43</td>
<td>0.37</td>
</tr>
<tr>
<td>(−114.0, -8.5, 80.0)</td>
<td>(π/4, π)</td>
<td>0.55</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>(π/4, 0)</td>
<td>0.65</td>
<td>0.68</td>
</tr>
</tbody>
</table>

Table 2: Estimated Coverage for Points in Fig. 8

<table>
<thead>
<tr>
<th>N</th>
<th>P_1</th>
<th>P_2</th>
<th>P_3</th>
<th>P_4</th>
<th>P_5</th>
<th>P_6</th>
<th>P_7</th>
<th>P_8</th>
<th>P_9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.64</td>
<td>0.93</td>
<td>0.63</td>
<td>0.92</td>
<td>0.90</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.93</td>
</tr>
<tr>
<td>2</td>
<td>0.54</td>
<td>0.00</td>
<td>0.00</td>
<td>0.92</td>
<td>0.94</td>
<td>0.95</td>
<td>0.95</td>
<td>0.94</td>
<td>0.71</td>
</tr>
<tr>
<td>3</td>
<td>0.41</td>
<td>0.00</td>
<td>0.94</td>
<td>0.93</td>
<td>0.00</td>
<td>0.95</td>
<td>0.95</td>
<td>0.94</td>
<td>0.00</td>
</tr>
<tr>
<td>4</td>
<td>0.41</td>
<td>0.94</td>
<td>0.95</td>
<td>0.00</td>
<td>0.00</td>
<td>0.93</td>
<td>0.00</td>
<td>0.95</td>
<td>0.00</td>
</tr>
<tr>
<td>5</td>
<td>0.49</td>
<td>0.94</td>
<td>0.93</td>
<td>0.93</td>
<td>0.00</td>
<td>0.00</td>
<td>0.88</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>6</td>
<td>0.57</td>
<td>0.00</td>
<td>0.91</td>
<td>0.00</td>
<td>0.92</td>
<td>0.93</td>
<td>0.93</td>
<td>0.94</td>
<td>0.92</td>
</tr>
<tr>
<td>A</td>
<td>0.84</td>
<td>0.94</td>
<td>0.95</td>
<td>0.92</td>
<td>0.93</td>
<td>0.90</td>
<td>0.95</td>
<td>0.95</td>
<td>0.93</td>
</tr>
<tr>
<td>B</td>
<td>0.74</td>
<td>0.93</td>
<td>0.43</td>
<td>0.92</td>
<td>0.94</td>
<td>0.95</td>
<td>0.95</td>
<td>0.94</td>
<td>0.93</td>
</tr>
<tr>
<td>C</td>
<td>0.66</td>
<td>0.94</td>
<td>0.95</td>
<td>0.00</td>
<td>0.00</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.91</td>
</tr>
<tr>
<td>D</td>
<td>0.71</td>
<td>0.94</td>
<td>0.93</td>
<td>0.93</td>
<td>0.93</td>
<td>0.93</td>
<td>0.93</td>
<td>0.94</td>
<td>0.92</td>
</tr>
</tbody>
</table>

6.3 Multi-Camera Model Validation

We demonstrate that the relationships between coverage performance values of various network configurations are sound and useful. As in the single-camera experiments, we take a qualitative approach at present, judging the perceived coverage in images from a variety of configurations and scenarios.

Figure 8 shows a typical set of images from our qualitative multi-camera model validation experiments. In this case, the non-directional coverage of eight small objects, placed at points p_1 through p_8 around the periphery of an occluding box, is tested for various combinations of up to six cameras. Table 2 shows the simple multi-camera network results, using γ = 20, R_1 = 3.0, R_2 = 0.5, c_{max} = 0.0048, and neglecting direction (no C_D component). D is specified as having μ_D = 1 in a region extending 200 mm outward from the box in x and y and the full height of the box in z (and 0 otherwise). The numbered networks 1 - 6 are the individual cameras, A is the network of all 6 cameras, and the remainder are defined as F = {1, 2}, G = {3, 4, 5}, and H = {5, 6}.

Some properties of the model are immediately clear from these results. First, the coverage degrees of p_1 through p_8 in A, F, G, and H are limited to the maximum of that of their constituent cameras. Second, in all cases where N_i ⊆ N_j for two networks N_i and N_j, m(N_i, D) ≤ m(N_j, D). One can also see clearly that despite having 3 cameras, G performs more poorly than the 2-camera networks F and H; this is consistent with the m values computed for the individual cameras from these networks. In general, the model gives a clear, intuitive, and accurate indication of coverage performance for the specified region.

7. DISCUSSION

7.1 Setting Application Parameters

In order to accurately evaluate coverage for a given task, the application parameters need to be chosen appropriately. An inherent advantage of models based on fuzzy sets is that such parameters need not be particularly precise, since the relative structure of results tends to be quite robust [11]. Though the importance and effects of the parameters will vary depending on the task, possible considerations include:
• γ: feature detector window size, target size
• R1, R2: pixel resolution requirements, target size
• c_{\text{max}}: required sharpness, robustness against blurring
• ζ: performance of feature detector over rotation

Multi-camera network applications often perform a number of distinct subtasks with disparate requirements. One may either take advantage of the aforementioned robustness of the model and find a reasonable ‘lowest common denominator’ set of application parameters, or else evaluate distinct models for each subtask.

The scene model \( S \) is also often required. This can potentially be generated from CAD layouts of the scene, e.g., for surveillance networks in a building, if fixed camera locations are known and can be matched to the poses in the model. Alternatively, static background estimation and 3D reconstruction can be used to automatically generate \( S \) from image data. Lobaton et al. [12] discuss obtaining a similar scene model for their topological coverage model.

7.2 Optimal Camera Placement

For a given camera type \( k \) of a set of camera types \( K \), a fixed set of approximate typical intrinsic parameters (perhaps an experimental average) yield type-specific single-camera coverage \( C_{k} \). Each such camera type also has an associated cost \( c_{k} \). It is assumed that the set of camera types is finite and static. The network coverage \( C_{N} \) can be obtained from \( C_{k,i} = C_{k} \) for \( k \in K \) and \( i \in \{1, 2, \ldots, n_{k}\} \), the pose of each camera \( P_{k,i} \), and the scene model \( S \). The total cost of the network is \( E = \sum_{k} n_{k} c_{k} \).

The objective is either to maximize \( m \), or else to minimize \( E \). Particular problem cases vary in terms of how they constrain minimum coverage degree \( m \), maximum total cost \( E \), the number of cameras of each type \( n_{k} \), and the set of allowed poses of cameras \( P_{k,i} \).

A wide variety of problem cases can be envisioned. Hörster and Lienhart [9] examine four good examples, translated here to our terminology and notation:

1. Given a fixed number of cameras \( n_{1} \) of a single type \( (|K| = 1) \), determine \( P_{1,i} \) such that \( m \) is maximized.
2. Given a variable number of cameras \( n_{k} \) each of one or more types, and a maximum total cost \( E_{\text{max}} \), determine \( n_{k} \) for \( k \in K \) and \( P_{k,i} \) such that \( m \) is maximized.
3. Given a fixed number of cameras in fixed locations (i.e., the translation component of \( P_{k,i} \) is fixed), determine their orientations (rotation component of \( P_{k,i} \) ) such that \( m \) is maximized.
4. Given one or more types of cameras \( (|K| \geq 1) \) and a minimum coverage performance \( m_{\text{min}} \), determine \( n_{k} \) for \( k \in K \) and \( P_{k,i} \) such that \( E \) is minimized.

Problems like these, phrased in the language of the general model, might be approached in a variety of ways computationally. A principal topic for future investigation is how to reduce or constrain this general model into various simpler forms for specific problems, which can then be attacked with numerical optimization techniques. One possible approach involves reduction to a fuzzy set-coverage problem [10]. Other possibilities include binarization via alpha cuts for binary integer programming and other similar methods, and topological representations based on fuzzy graphs for combinatorial optimization.

8. CONCLUSIONS

The fuzzy coverage model presented is a comprehensive and intuitive model of camera and camera network coverage. Based on qualitative validation, the model appears to conform very closely to reality, and is able to evaluate coverage with a single, intuitive metric.

By itself, the model allows for comparison of the coverage performance of most types of camera or camera network configurations, for a vast array of conceivable applications. This allows for generate-and-test approaches to planning camera placement. A goal we are actively pursuing for future work is automatic optimization of camera placement, which will use the numerical metrics of the fuzzy coverage model (or a variety of possible simplified derivative models) as a performance indicator.

Additionally, the general concept of the model – namely, using parameterized fuzzy subsets of \( \mathbb{R}^{3} \) and \( \mathbb{D} \) for sensor coverage – may be generalized to various non-directional and directional sensing modalities. Graph (classic or fuzzy) formalisms derived from the model, using coverage overlap or dissimilarity metrics such as fuzzy Hausdorff distance [4] for membership valuation, may provide additional opportunities for optimization in the network context. In particular, it is hoped that this line of research will benefit existing multi-camera applications such as calibration and tracking. More generally, we hope that exploration of interrelationships through these models will yield insights toward a more fundamental theoretical understanding of camera networks.

Acknowledgment

This research was supported in part by the Natural Sciences and Engineering Research Council of Canada.

9. REFERENCES


