Summary of results on optimal camera placement for boundary monitoring

Robert J. Holt[a]   Hong Man[b]   Rainer Martini[b]
Iraban Mukherjee[b]   Ravi Netravali[b]   Jing Wang[b]

[a] Department of Mathematics and Computer Science,
City University of New York, Bayside, New York, USA

[b] Department of Electrical and Computer Engineering,
Stevens Institute of Technology, Hoboken, New Jersey, USA

ABSTRACT

We consider the problems of placing cameras so that every point on a perimeter, that is not necessarily planar, is covered by at least one camera while using the smallest number of cameras. This is accomplished by aligning the edges of the cameras' fields of view with points on the boundary under surveillance. Taken into consideration are visibility concerns, where features such as mountains must not be allowed to come between a camera and a boundary point that would otherwise be in the camera's field of view. We provide a general algorithm that determines optimal camera placements and orientations.

Additionally, we consider double coverings, where every boundary point is seen by at least two cameras, with selected boundary points and cameras situated such that the average calibration errors between adjacent cameras is minimized. We describe an iterative algorithm that accomplishes these tasks. We also consider a joint optimization algorithm, which strikes a balance between minimizing calibration error and the number of cameras required to cover the boundary.

Keywords: Intrusion detection, border surveillance, camera placement, visibility

1. INTRODUCTION

Recently there has been much interest in the topic of intrusion detection and border surveillance. Every year sees hundreds of thousands of unauthorized national border crossings. While most of these crossings are by individuals seeking better employment, the possibility of entry by those with weapons of destruction is cause for concern.

Among solutions considered are physical barriers. These are useful in some cases, but frequently need to be augmented some kind of sensors. In this paper we investigate the placement of cameras on uneven terrains so that they cover a perimeter as efficiently as possible. Problems considered include single and double coverage, where in the latter every point on the boundary is seen by at least two cameras. In this case situation we may place cameras so that the average calibration error between consecutive cameras is minimized.

Papers concerned with intrusion detection, and the placement of cameras to monitor a perimeter, or its enclosed region in the case of a closed boundary, include[1–7] . In[8, 9] , surveys of many more papers on these topics are presented. While most of this work is concerned detecting changes in image sequences, the focus of this paper is the placement of cameras to provide an optimal collection of images, which is a major topic discussed in[4] .

(Send correspondence to R.J.H.)
R.J.H.: e-mail: rholt@qcc.cuny.edu
H.M.: e-mail: hman@stevens.edu
R.M.: e-mail: rmartini@stevens.edu
I.M.: e-mail: imukherj@stevens.edu
R.N.: e-mail: ran2290@gmail.com
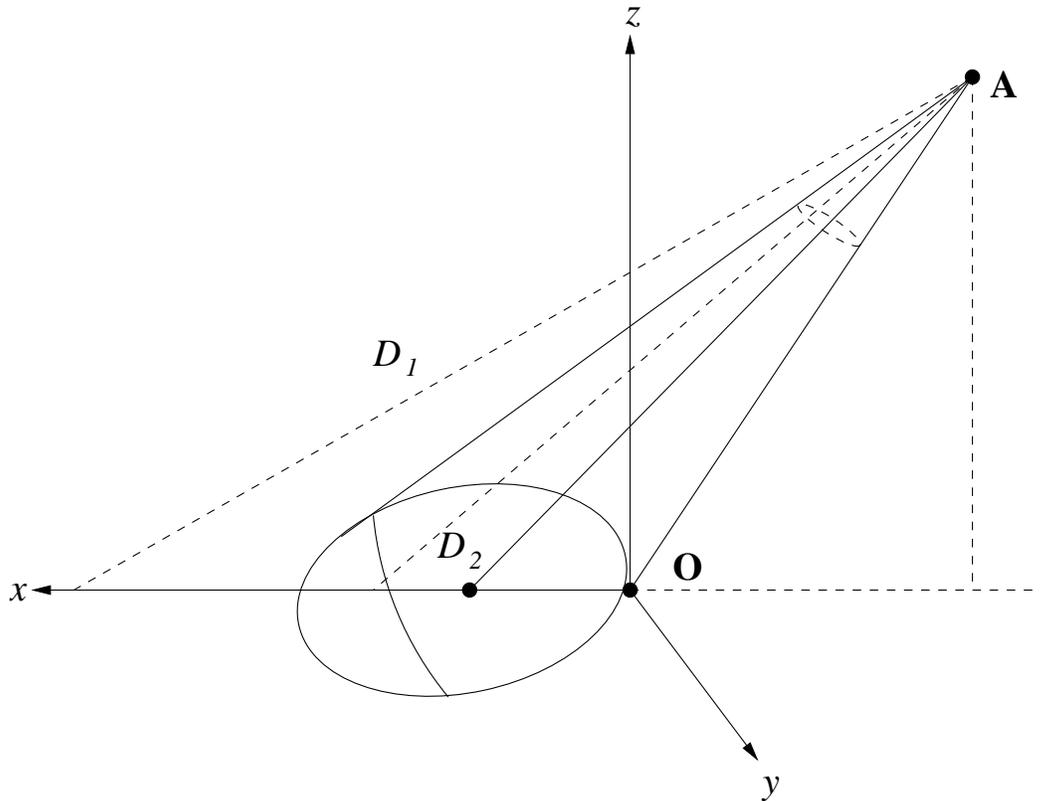J.W.: e-mail: jwang@stevens.edu

Figure 1. Geometry of monitoring camera with circular lens. The focal point is at A. A conical field of view intersects the ground plane in an ellipse. The range of the camera, denoted by D, indictates the maximun distance at which an object to be observed can be resolved. If D is less than the distance from the focus to the farthest point of the ellipse, as is the case when the distance is as small as D_2, then only features in that part of the ellipse before the circular arc may be resolved by the camera. IF D is as large as D_1, then any feature in the ellipse may be resolved.

2. COMPARISON WITH STRAIGHTFORWARD TECHNIQUES

One straightforward way of placing cameras is to situate them evenly spaced around the boundary, at a fixed height and angle of the optic axis to the vertical. An improvement with regards to minimizing the number of cameras is achieved by placing them so that the field of views of consecutive cameras are barely overlapping. We make improvements by both allowing the angles to the vertical to be adjustable, and allowing the cameras to be situated a significant distance from the boundary in question.

We assume the boundary to be monitored is given, and take many sample points along it. The system of cameras will cover all of the sample points, and barely overlap in order to minimize the number required.

We shall need to determine the shape for the field of view of the cameras. With circular lenses, and mounted at some height looking at an angle downward toward the ground, a reasonable choice is that the area of ground coverage is an ellipse. This comes from the fact that the intersection of a cone, which represents the region of space visible from a pinhole camera, with a plane is an ellipse. A complication arises from the fact that cameras have finite range. That is, beyond a certain distance, an object such as a person or vehicle occupies too few pixels to determine what it is. Thus the range of the camera is bounded by a sphere about its focus, and the intersection of this sphere with the ground may or may not intersect the ellipse just described. In this case the field of view is an ellipse truncated by a circular arc.

In many situations, in particular those where the boundary under surveillance does not have long straight lines, the circular arcs might not intersect the ellipses. The possible situations are illustrated in Figure 1.

The actual choice of coverage area is not overwhelmingly important, as our algorithms readily adapt to other shapes, such as rectangles or sectors of circles.

3. CAMERA PLACEMENT FOR THREE-DIMENSIONAL TERRAINS

In many real-life situations the perimeter, for example a border between two regions, may be far from planar. In this case we can still sample the boundary, recording 3-D points. We seek to maximize the number of consecutive sample boundary points that lie within the cone of visibility of a camera, or a spherical cone. Say a camera has focal length F and radius r of the circular region of the image plane within the camera's field of view. This implies that the viewing angle of the camera is $2\tan^{-1}(r/F)$. Also suppose that D is the maximum range of the camera, which means that objects of interest, namely possible intruders, farther away than D from the focal point are too small to be successfully resolved.

We seek to find the position of a camera so that a typical sample boundary point lies on the edge of the cone of visibility. If the focal point is at the origin and the optic axis points straight down, then the spherical cone is described by

$$\frac{x^2 + y^2}{r^2} \leq \frac{z^2}{F^2} \ , \ \ x^2 + y^2 + z^2 \leq D^2 \ , \ \ z < 0 \ .$$

If this spherical cone is rotated through an angle α about the y-axis, bringing it toward the halfspace where $x > 0$, this set of inequalities becomes

$$\begin{aligned}
\frac{(x\cos\alpha + z\sin\alpha)^2 + y^2}{r^2} &\leq& \frac{(-x\sin\alpha + z\cos\alpha)^2}{F^2} \ , \\
x^2 + y^2 + z^2 &\leq& D^2 \ , \\
-x\sin\alpha + z\cos\alpha &<& 0 \ .
\end{aligned}$$

Next, if the spherical cone is translated so that its vertex is at the point $(-l, 0, h)$, raising the vertex to height h above the xy-plane, the set of inequalities becomes

$$\begin{aligned}
\frac{[(x+l)\cos\alpha + (z-h)\sin\alpha]^2 + y^2}{r^2} &\leq& \frac{[-(x+l)\sin\alpha + (z-h)\cos\alpha]^2}{F^2} \ , \\
(x+l)^2 + y^2 + (z-h)^2 &\leq& D^2 \ , \\
-(x+l)\sin\alpha + (z-h)\cos\alpha &<& 0 \ .
\end{aligned}$$

Then if the spherical cone is rotated through an angle θ around the z-axis, the set of inequalities is

$$\begin{aligned}
\frac{[(x\cos\theta + y\sin\theta + l)\cos\alpha + (z-h)\sin\alpha]^2 + (x\sin\theta - y\cos\theta)^2}{r^2} & \\
\leq \frac{[-(x\cos\theta + y\sin\theta + l)\sin\alpha + (z-h)\cos\alpha]^2}{F^2} \ , & \\
x^2 + y^2 + 2l(x\cos\theta + y\sin\theta) + l^2 + (z-h)^2 &\leq& D^2 \ , \\
-(x\cos\theta + y\sin\theta + l)\sin\alpha + (z-h)\cos\alpha &<& 0 \ .
\end{aligned} \quad (1)$$

Figure 2 shows the spherical cone after the rotation and translation described above, for $\theta = 0$.

The angle α is determined by h and l. From Figure 2 we see that

$$\tan\left(\alpha - \tan^{-1}\frac{r}{F}\right) = \frac{l}{h} \ ,$$

from which

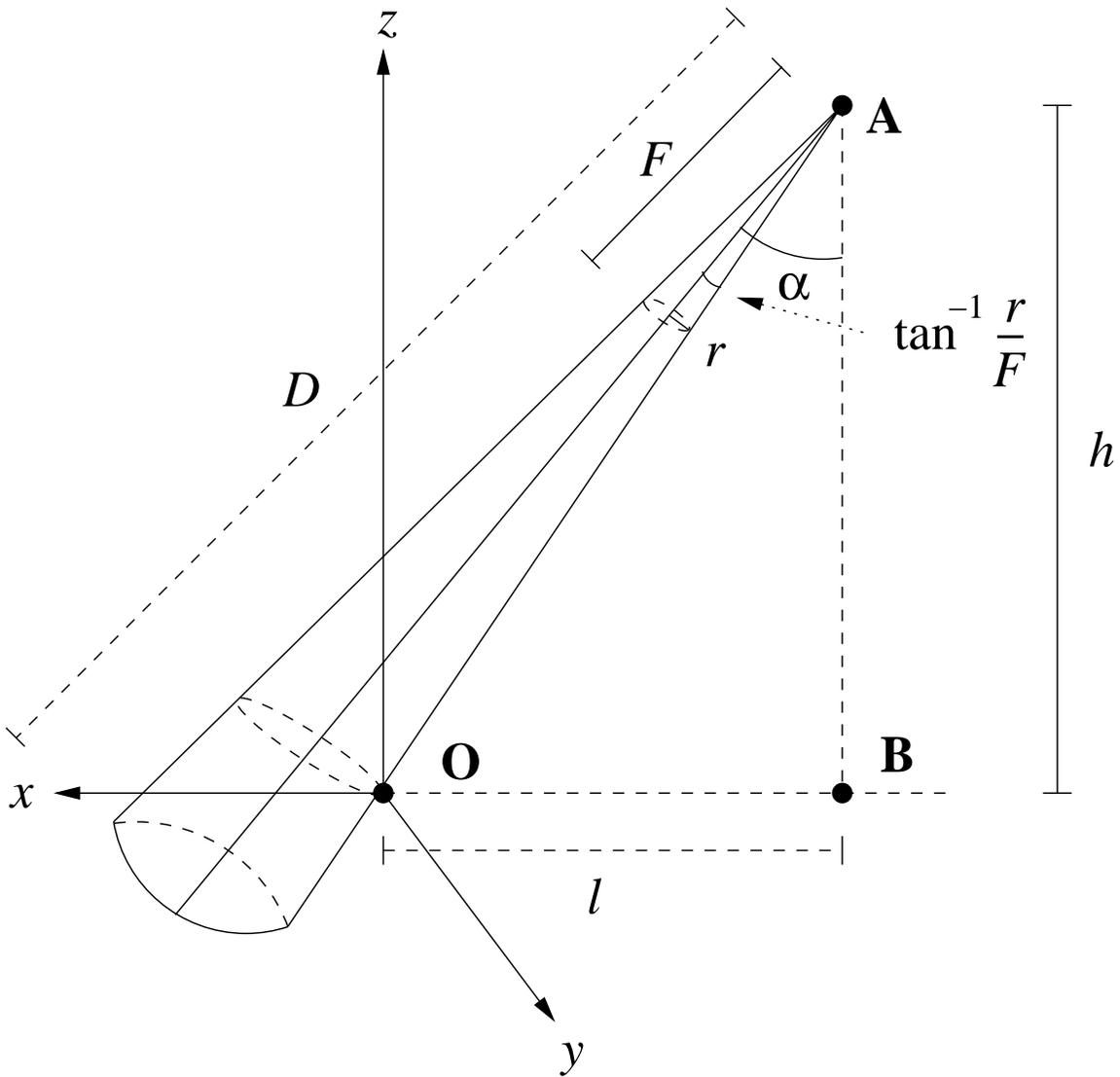$$\alpha = \tan^{-1}\frac{r}{F} + \tan^{-1}\frac{l}{h} = \tan^{-1}\frac{Fl + rh}{Fh - rl} \ . \quad (2)$$

Figure 2. Spherical cone of visibility for a pinhole camera. The optic axis is at an angle α from vertical, and the camera is placed so that the generator segment that makes the smallest angle with the z-axis touches the x-axis (segment $\overline{AO}$).

The intersection of the spherical cone with the xy-plane $z = 0$ is the ellipse

$$
\begin{aligned}
&[(F^2 cos^2\alpha - r^2\sin^2\alpha)\cos^2\theta + F^2\sin^2\theta]x^2 \\
&- 2(F^2 + r^2)sin^2\alpha\sin\theta\cos\theta\, xy \\
&+ [(F^2 cos^2\alpha - r^2\sin^2\alpha)\sin^2\theta + F^2\cos^2\theta]y^2 \\
&- 2\cos\theta[(F^2 + r^2)h\sin\alpha\cos\alpha - l(F^2\cos^2\alpha - r^2\sin^2\alpha)]x \\
&- 2\sin\theta[(F^2 + r^2)h\sin\alpha\cos\alpha - l(F^2\cos^2\alpha - r^2\sin^2\alpha)]y \\
&+ F^2(h\sin\alpha - l\cos\alpha)^2 - r^2(h\cos\alpha - l\sin\alpha)^2 = 0 \ .
\end{aligned}
$$

This ellipse is centered at

$$
(x,y) = (-\cos\theta, -\sin\theta)\left[l - \frac{(F^2 + r^2)h\sin\alpha\cos\alpha}{F^2\cos^2\alpha - r^2\sin^2\alpha}\right] \ ,
$$

and has semimajor and semiminor axes of

$$
a = \frac{Frh}{F^2\cos^2\alpha - r^2\sin^2\alpha} \qquad \text{and} \qquad b = \frac{rh}{\sqrt{F^2\cos^2\alpha - r^2\sin^2\alpha}} \ ,
$$

and eccentricity

$$
e = \frac{(F^2 + r^2)\sin^2\alpha}{F^2} \ .
$$

A sphere of radius D, the range of a camera, mounted at a height h above the ground, intersects the horizontal plane through the base of the mount in a circle of radius $\sqrt{D^2 - h^2}$. This will typically intersect the ellipse described above, specifically when $D^2 - h^2 < l + 2a$. Thus the planar field of view of a camera may be described as an elliptic sector. Examples of these elliptic sectors are shown in Figure 3 show several of these possible elliptic sectors. Specifically, these are the sectors obtained when $r/F = 0.2$, $\alpha = 70°$, and h ranges from $0.1D$ to $0.9D$. The eccentricity of the ellipses may be expressed in terms of D as

$$
e = \sqrt{1 - \frac{h^2}{D^2}} - \frac{r}{F} = \frac{F\sqrt{D^2 - h^2} - hr}{FD} \ .
$$

Over an uneven terrain these fields of view resemble warped elliptic sectors.

3.1. Algorithm

The algorithm works as follows. Any sample point is selected to begin with, and corresponds to the origin **O** in Figure 2. Then a three-parameter family of spherical cones is tested to determine which contains the largest number of consecutive boundary points. These parameters are coordinates of the prospective focal point positions **A** of the camera. Trial points are selected from a solid hemisphere of radius D above and about the sample point. With a change of coordinates to bring the sample point to the origin, the inequalities (1) are tested. The test point for which the greatest number of consecutive sample boundary points satisfy all the inequalities is chosen as the position of the first camera. Ties may be broken arbitrarily.

The last boundary point covered is chosen as the starting vertex for the second spherical cone. This second spherical cone is constructed in the same manner as the first, subject to the condition that the first sample boundary point missed by the first spherical cone must be contained in the second. Naturally the sample points should be sufficiently dense so that several of them lie within each spherical cone. This process is repeated until all of the sample boundary points are covered.

The process is repeated with different starting points for the first spherical cone. It is possible that one more or less cone is required for different starting points.
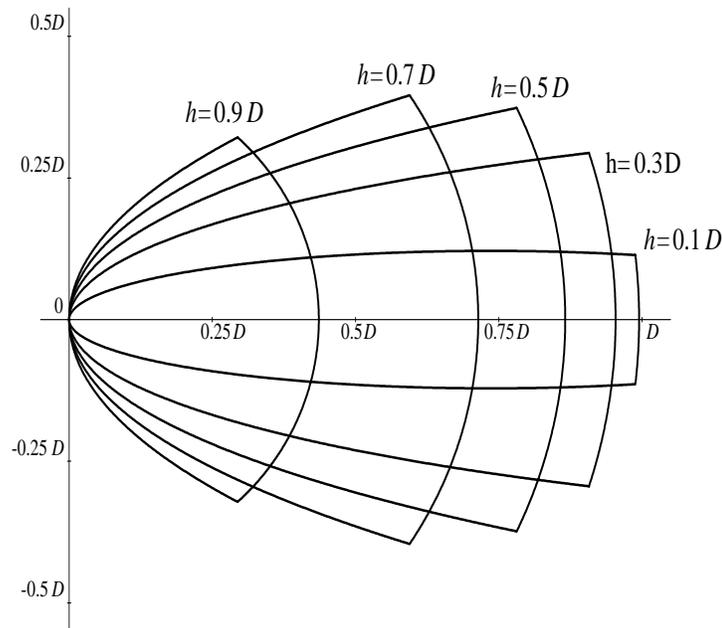
Figure 3. Possible elliptic sectors for varying camera heights. The parameters here are $r/F = 0.2$, $\alpha = 70°$, and h, the height of the camera above the ground, ranging from $0.1D$ to $0.9D$.

3.2. Double coverage

One frequently wants every point on the boundary to be covered by two cameras, in case of failure of one device, or to permit feature-based calibration. To accomplish this, one can start at a middle point of the first elliptic sector in one covering, and work one's way around the boundary as before.

Another possibility is to incorporate the fact that often there will be a tie among several candidate camera locations for covering the maximum number of consecutive sample boundary points. In this situation one should select camera positions that minimize the calibration error, with optic axes as close to right angles as possible.

Experiments indicated that better results are obtained when one traverses the boundary for the second covering in the opposite direction as was used for the first covering. This remained true for both open and closed boundary curves. When one goes in the same direction for both covers, there often is a tendency for the cameras in the two sets to bunch together with the result that the last several pairs of cameras cover the same points.

3.3. Coverings of modified regions

In some circumstances the number of cameras required to monitor a planar non-convex region may be reduced by monitoring its convex hull, or portions of it, since that will have a shorter perimeter. In some situations this may not be possible, such as when the exterior of the region is property belonging to someone else, or contains terrain unsuitable for the mounting of cameras.

A space curve does not have a convex hull in the same sense as a planar curve, but one may proceed as follows. Project the boundary space curve onto a plane, typically a horizontal one. Then intersect the cylinder of this projection with the actual terrain. When practical, this may result in requiring the monitoring of a significantly shorter boundary.

Due to various circumstances, it may not be possible to place a camera at a location determined by the algorithm, such as on the wrong side of a border. For a closed boundary, a way to circumvent this possibility is to first start with a curve that is within the original boundary, a distance equal to the range D of the cameras, and apply the algorithm to that curve. For an open boundary, one can first start with a curve that is "parallel"

to the original boundary, at a distance D from the original toward the side on which cameras may be placed. Of course nothing sensitive should be in the region between these two curves in either case. This region could be considered a buffer zone.

3.4. Visibility

It is possible that mountains, buildings, or other features obstruct some of the sample boundary points from the camera positions as computed above. To overcome this problem, we can triangulate the surface being monitored. Each wall and roof of a rectangular building can be split into two triangles, while hills and mountains may be modeled as pyramids. The triangles of the triangulation are stored in a table. When a boundary point is tested to see whether it is within the range of a candidate camera position, the line segment connecting the focal point to the boundary point is tested against the triangles in the table. If the line segment is found to intersect a triangle, the boundary point is deemed not visible for that camera given the position and orientation in question. Consequently the algorithm can be modified to include this condition in addition to (1) when determining whether a point is visible.

To determine whether the line segment between the focal point an a boundary point intersects a triangle, the plane of the triangle and a normal vector is determined by taking cross products of two of the sides. If the two points are on the same side of the plane, which occurs when their dot products with the normal vector have the same sign, then the segment does not intersect the triangle. If the points are on opposite sides, then the point of intersection $\mathbf{p}$ of the segment and the plane of the triangle is found. If the vertices of the triangle are $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$, then $\mathbf{p}$ is in the interior of $\Delta\mathbf{p}_1\mathbf{p}_2\mathbf{p}_3$ if and only if $\mid \mathbf{p} \quad \mathbf{p}_1 \quad \mathbf{p}_2 \mid$, $\mid \mathbf{p} \quad \mathbf{p}_2 \quad \mathbf{p}_3 \mid$, and $\mid \mathbf{p} \quad \mathbf{p}_3 \quad \mathbf{p}_1 \mid$ all have the same sign.

4. JOINT AND SEPARATE OPTIMIZATION

Another problem of interest is placing cameras so that the maximum number of sample boundary points is covered while simultaneously minimizing the average calibration error. Joint optimization is the process of working toward those goals simultaneously. The coordinate systems of two cameras are related by a rotation followed by a translation, and the calibration error is a measure of how far off the computed values of these transformation parameters are from the actual values.

We determine the sample points in the intersection of each consecutive pair of cameras' field of view as obtained by the double coverage algorithm. This number must be at least two in order for calibration to be possible. As more information than just the coordinates of two or three feature points is required for calibration, other constraints on the cameras may be incorporated. These include the distance from the focal point of a camera to a feature point, the distance between two cameras, and the angle formed by a feature point and two cameras. Numerical tests indicate that calibration error tends to be the smallest when the optic axes of two cameras are perpendicular, and greatest when they are nearly parallel.

One can start with a double covering as in Section 3.2. This partitions the set of sample boundary points into subsets that are covered by two consecutive cameras. In order to maintain the minimum number of cameras required for the double coverage, in the optimization problem the cameras must be required to cover the points in the original double cover. For example, suppose camera 1 covers boundary points $\mathbf{p}_1$, $\mathbf{p}_2$, $\mathbf{p}_3$, and $\mathbf{p}_4$, while camera 2 covers points $\mathbf{p}_3$, $\mathbf{p}_4$, $\mathbf{p}_5$, and $\mathbf{p}_6$, so that $\mathbf{p}_3$ and $\mathbf{p}_4$ are in common to both. Then cameras 1 and 2 should be placed so that the calibration error, using $\mathbf{p}_3$ and $\mathbf{p}_4$ as the feature points, subject to the conditions that $\mathbf{p}_1$ and $\mathbf{p}_2$ are in the field of view of the new position of the first camera, and $\mathbf{p}_5$ and $\mathbf{p}_6$ are in the field of view of the new location of the second camera.

In separate optimization, one first places all the cameras so that their fields of view overlap sufficiently. Then the camera positions are adjusted so as to minimize calibration error while still keeping the same sample points in their fields of view. In joint optimization, once the first two cameras are in place, subsequent cameras are positioned immediately so that the calibration error is minimized. For example, the third camera is placed so that the error is minimized subject to the conditions that all of the points original seen by the third camera are seen by it in its new location, and that the second camera is already placed by the optimization of the first two cameras. If a maximum threshold for the allowable calibration error cannot be avoided with just two feature

points in common to the fields of view of two cameras, then the latter camera is placed so that it has three points in common with the former, and so on. Different weights may be placed on the number of cameras and the calibration error, allowing the user to determine the relative importance of each.

Another possibility is to try to optimize the placement of all the cameras at once. If possible, this would provide the best results, but the size of the problem is far too large for a moderate number of cameras.

One might also consider the regions covered in a single coverage, placing two cameras at a time to cover the points seen by one camera. The drawback to this method is that in order for two cameras to cover the same set of points, they will often be very near each other, particularly when two sample boundary points are near the vertices of the ellipse obtained in the single coverage.

5. EXPERIMENTAL RESULTS

The authors wrote programs in Java incorporating the algorithms above.

Example 1

This is a synthetic example, a space curve given by $(x, y, z) = (2(2 + \cos 3\theta) \cos \theta, 2(2 + \cos 3\theta) \sin \theta, \sin 5\theta)$, with the sample boundary points taken for $\theta = 9k$ degrees, $k = 0, 1, \ldots, 39$. The viewing angle of the cameras is 22.6 degrees, and the range of each camera is 11 units. The camera locations, denoted by the gray squares, lie on the major axes of their resective elliptic sectors. For example, the square at the top, toward the left side, is the camera whose field of view is the elliptic sector at the center of the left edge of the diagram.
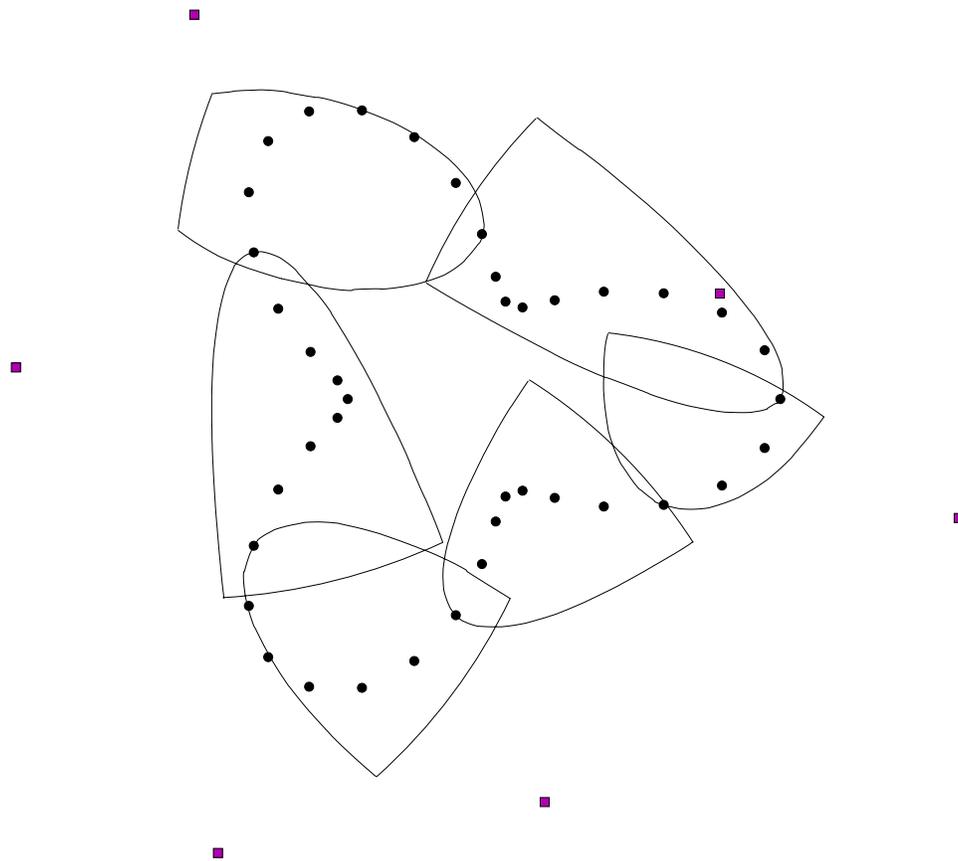

Figure 4. Synthetic example of hilly terrain. The black circles are border points, and the gray squares are the camera locations, and lie on the major axes of their respective elliptic sectors.

Example 2

For comparison purposes, Figure 5 is the same example with the all the cameras at a fixed height above the ground. Even with the optimal height an additional camera is required to cover the entire boundary.
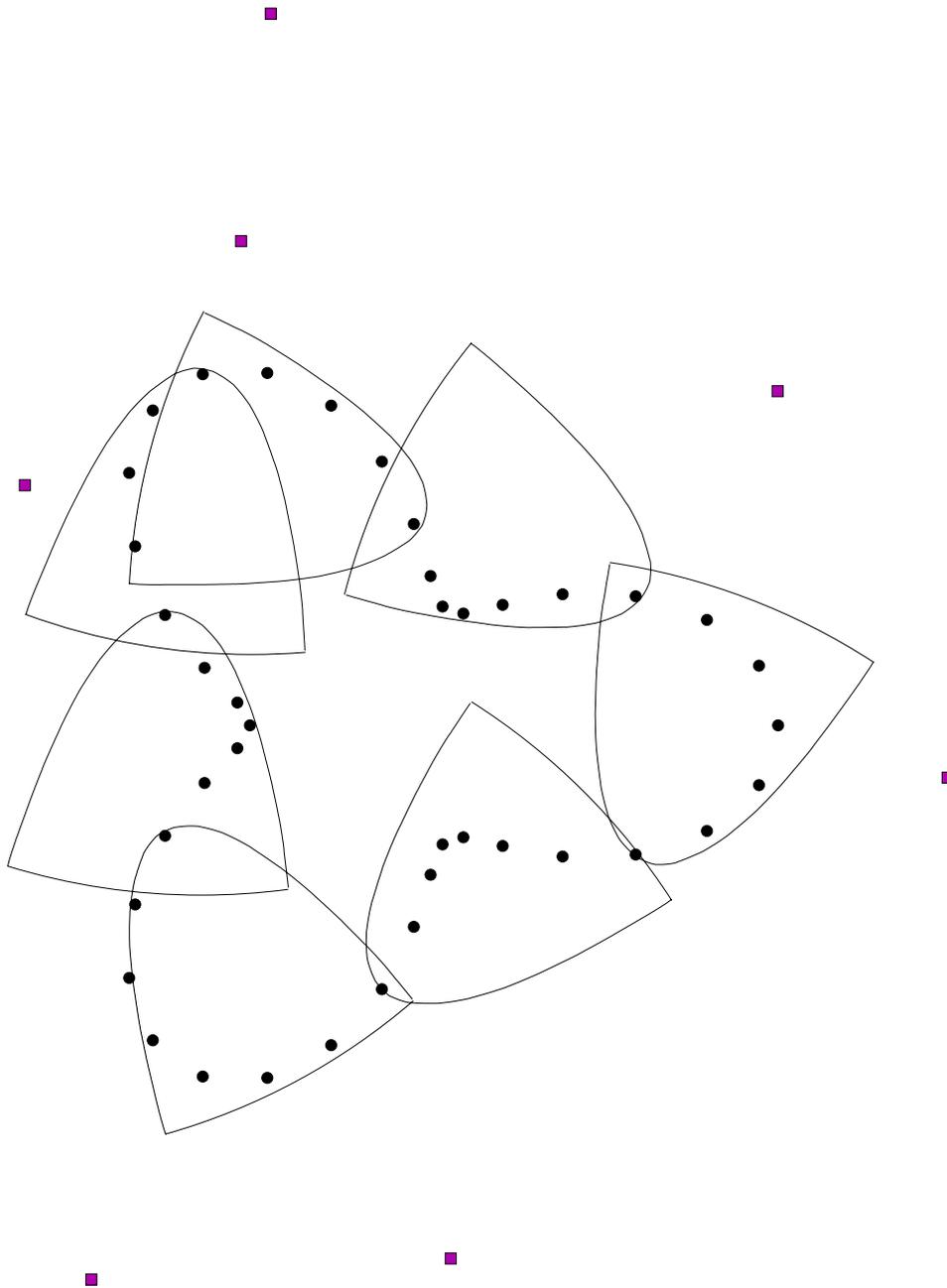

Figure 5. Synthetic example with all cameras at a fixed height above the ground. An additional camera is needed compared to the example in Figure 4

.

Example 3

Next is the same example with double coverage. Two sets of cameras, the solid squares with the solid-line bordered elliptic sectors, and the hollow squares with the dashed-line borders, are positioned so that each boundary point is covered by at least one camera in each set.
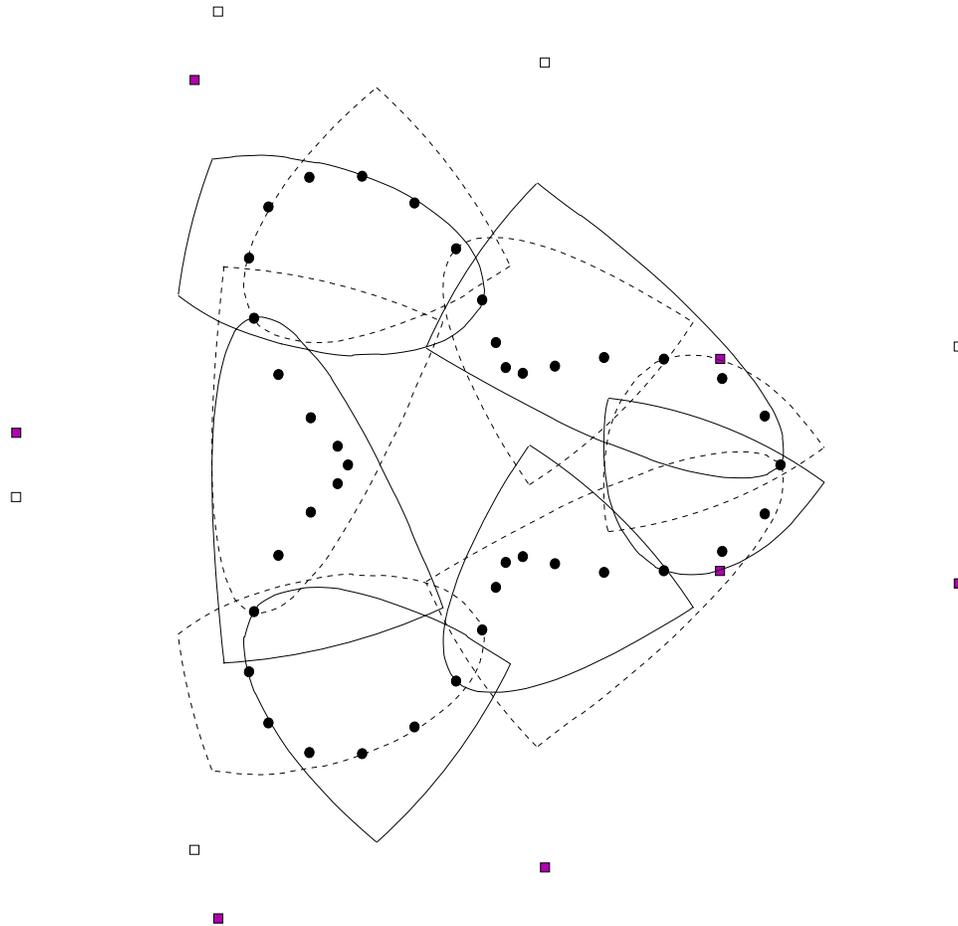

Figure 6. Double coverage. Each point on the boundary is covered by at least one elliptic sector with a solid boudnary and another with a dashed-line boundary. The hollow squares are the camera locations for the dashed-line sectors.

Example 4

This example is a portion of the United States-Mexico border. The latitude, longitude, and elevation of thirty points were determined from a web site. The results in Figure 7 arise when the cameras have a viewing angle of 20 degrees.
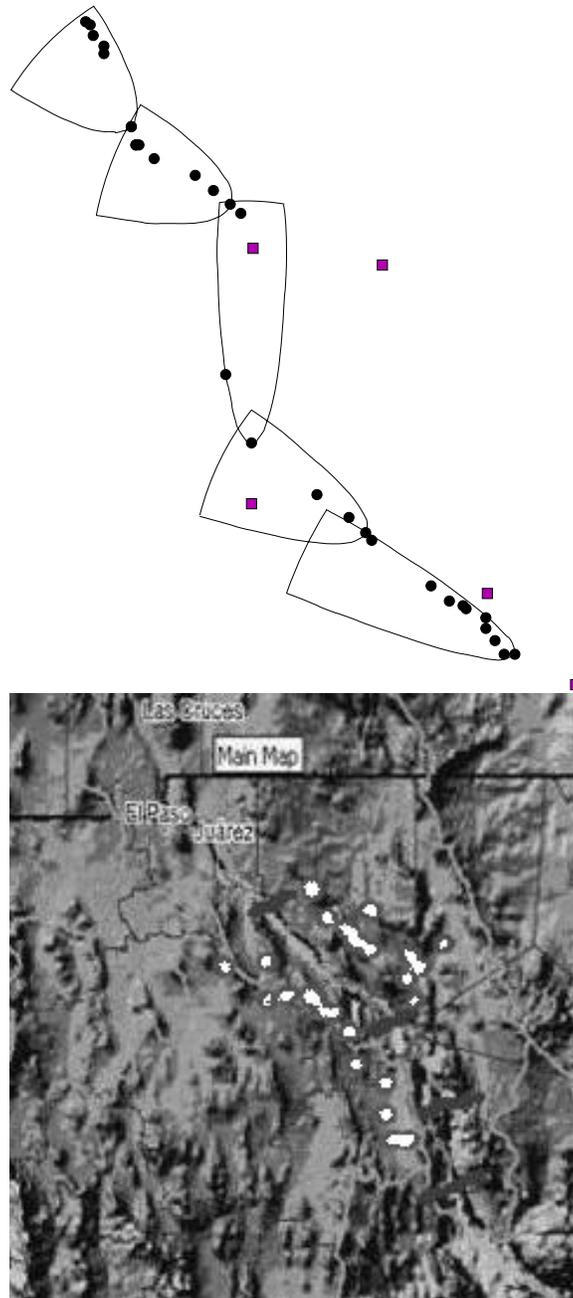

Figure 7. Optimal camera arrangement over portion of US-Mexico border. The parameters here are $F = 1$, $r = 0.2$, $D = 200$ pixels. The black circles are border points, and the gray squares are the camera locations, and lie on the major axes of their respective elliptic sectors. The bottom picture is the source, where the lower set of white dots constitute the boundary points used.

6. CONCLUSIONS AND DISCUSSION

In this paper we described an algorithm for the placement and orientation of cameras over uneven terrains so that they cover a boundary as efficiently as possible. By aligning edges of the cameras' fields of visions, we can guarantee complete coverage of a boundary in an efficient manner. The algorithm can be modified to accommodate obstacles such as hills or buildings. When physical constraints, such as cameras required to be on one side of a border, the algorithms may be applied to similar curves in the allowable region.

Double coverings are achievable by minor modifications, and can be obtained so that the calibration errors between successive cameras is minimized. A joint optimization process can be used to strike a balance between the number of cameras and maximum allowable calibration error.

An open question is how close to the absolute minimal coverings do these algorithms come. Our algorithms have the feature that the vertices of the elliptical sectors always lie on the boundary to be monitored. This guarantees coverage, but in many cases it may be possible to adapt the curve of the boundary to the shape of the field of view to allow coverings with fewer cameras. Naturally this involves a great increase in computational complexity.

Another approach is to take a representative sample of the shapes of fields of view, and move them so that they cover maximal portions of the boundary. With the elliptic sectors, it may not always be the case that the vertex lies along the boundary, but again the search space would be higher dimensional, with a corresponding computational cost.

ACKNOWLEDGMENTS

REFERENCES

1. M. J. Black, D. J. Fleet, and Y. Yacoob, "Robustly estimating changes in image appearance," *Computer Vision and Image Understanding* **78**, pp. 8–31, 2000.
2. R. T. Collins, A. J. Lipton, H. Fujiyoshi, and T. Kanade, "Algorithms for cooperative multisensor surveillance," *Proceedings of the IEEE* **89**, pp. 1456–1477, 2001.
3. A. J. Lipton, C. H. Heartwell, N. Haering, and D. Madden, "Automated video protection, monitoring & detection," *IEEE AESS Systems Magazine* **18**, pp. 3–18, 2003.
4. I. Pavlidis, V. Morellas, P. Tsiamyrtzis, and S. Harp, "Urban surveillance systems: From the laboratory to the commercial world," *Proceedings of the IEEE* **89**, pp. 1478–1497, 2001.
5. C. Stauffer and W. E. L. Grissom, "Learning patterns of activity using real-time tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**, pp. 747–757, 2000.
6. C. Stauffer and K. Tieu, "Automated multi-camera planar tracking correspondence modeling," in *Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1*, pp. 259–266, (Madison, Wisconsin), July 2003.
7. S. Young, M. Forshaw, and M. Hodgetts, "Image comparison methods for perimeter surveillance," in *Seventh International Conference on Image Processing and its Applications*, pp. 799–802, (Manchester, UK), July 1999.
8. R. J. Radke, S. Andra, O. A-Kofahi, and B. Roysam, "Image change detection algorithms: A systematic survey," *IEEE Transactions on Image Processing* **14**, pp. 294–307, 2005.
9. M. Valera and S. A. Velastin, "Intelligent distributed surveillance systems: A review," *IEE Proceedings - Vision, Image, and Signal Processing* **152**, pp. 192–204, 2005.