Optimal Camera Configuration for Large-Scale Motion Capture Systems

Xiongming Dai
https://daixiongming.github.io/
Gerald Baumgartner
http://cse.lsu.edu/~gb/

Division of Computer Science and Engineering
Louisiana State University
Baton Rouge, LA 70803, USA

Abstract

Camera network designs and 3D marker-based motion capture systems are enabling high-quality real-time interaction for multiple users. For greater efficiency, the camera configuration is motivated by the need to achieve 3D realistic and dynamic effects. We convert each sensing requirement into the geometrical and optical constraints on sensor location, developing a binary integer programming model with an included occlusion culling factor, from which the 3D region of viewpoints that satisfies that constraint is computed by greedy heuristics with Riesz-particle scale optimization. The optimal camera configuration problem is $\mathcal{NP}$-hard. We prove that our performance ratio $H(k)$ grows at most logarithmically, under mild assumptions.

1 Introduction

Camera network design is widely used in computer vision, such as interactive virtual reality, video surveillance and immersive motion capture [6, 16, 20, 26, 31, 32, 33]. An effective design of a camera network for all tasks is a necessary process, and it may have a profound impact on the execution of subsequent tasks. For example, highly accurate 3D reconstruction requires a camera network with complete coverage, motion capture, recognition, and tracking require a high resolution, etc. Therefore, the optimal camera configuration problem is an important and practical topic.

The original studies of optimal camera configuration date back to the "art gallery problem" in computational theory [23, 24, 28], which theoretically posed the problem of determining the minimum number of point guards sufficient to cover the target modeled as a simple polygon P with n vertices. Vasek Chvátal has shown that $\lfloor \frac{n}{3} \rfloor$ guards ares sufficient in [7, 8]. It arises in real applications, such as robotics, motion planning, and camera network design [4, 5, 14, 18, 22, 27].

A great deal of prior works exists where the field of view, working volume, focus, visibility, resolution and occlusion are considered for formulations [7, 9, 19]. Current research in optimal camera configuration has focused on two main directions: how to formulate the problem and how to approximate the optimum. Most earlier studies discretize the space with grids to formulate the model [11, 13, 14, 16]. While most formulations are in $\mathcal{NP}$-hard, some approximation methods have been proposed, including greedy heuristics, semi-definite

programming, and simulated annealing [1, 2, 6, 15, 16, 26, 29, 33]. In practice, for motion capture, the marker should be tracked continuously, the motion of the corresponding point in the space should not be vulnerable to occlusions from either static or dynamic objects. Here we focus on the topic of maximizing the coverage with the occlusion culling factor under a continuous space of markers for large-scale motion capture systems.

In this paper, static occlusion culling is provided and extended into the dynamic scenario, where a probabilistic dynamic occlusion culling model is formulated for visibility analysis. We have built a new binary integer programming model incorporating occlusion culling factors, where the scene representation, camera model, visibility analysis, and geometric and optical constraints on sensor location are considered. Then, we present a greedy heuristic with a Riesz-particle scale optimization. The optimal camera configuration problem is $\mathcal{NP}$-hard. We analyze the complexity of the algorithm and prove that our performance ratio $H(k)$ grows at most logarithmically, under mild assumptions.

The remainder of this paper is organized as follows. In Section 2, we introduce visibility analysis and the formulation of the objective. In Section 3, we introduce the greedy heuristic with the Riesz-particle scale optimization. Compared with other heuristic algorithms, simulations and statistical analysis are conducted under the same scenarios and constraints for validation in Section 4. The summary of our contribution is outlined in Section 5 and the acknowledgement is provided in Section 6.

2 Visibility Analysis and Formulation of the Objective

2.1 Static Occlusion Culling

High-performance tracking depends on spatial points being visible simultaneously by at least two cameras and the triangulation accuracy. Theoretically, if the point can be visible simultaneously by at least two cameras, the point could be reconstructed. Practically, the radius of the marker is not negligible, we find that the target of working volume is visible from two cameras if the angle between them is less than or equal to $\frac{\pi}{2}$, which could be formulated as

$$\min(\theta_{i,j}) \leq \frac{\pi}{2} + \frac{2r}{L}, \forall i, j \in [1, M], j \neq i, \tag{1}$$

where $\theta_{i,j}$ denotes the included angle between camera i and camera j, r denotes the radius of the marker, M is the number of cameras and L is the effective range. This angle could be calculated by

$$\theta_{i,j} = \arccos \frac{([x_i\ y_i\ z_i]^t - T) \cdot ([x_j\ y_j\ z_j]^t - T)}{\|[x_i\ y_i\ z_i]^t - T\| \cdot \|[x_j\ y_j\ z_j]^t - T\|}, \tag{2}$$

where $[x_i'\ y_i'\ z_i']^t$ and $[x_j'\ y_j'\ z_j']^t$ are the ordinates of ith and jth camera placement, T is the markers in 3D working volume and $\|\cdot\|$ denotes the Euclidean norm of the vector.

2.2 Dynamic Occlusion Analysis

In this section, we develop a model to analyze the probability of visibility of markers, given a set of multi-camera configurations in the presence of random dynamic occluders.

Suppose we have a region R covered by M cameras. Let A_i denote the event that a target T in working volume $W(W \subset R)$ with orientation φ is visible from camera C_i. Thus, the

probability of the event that T can be covered by at least two cameras for triangulation could be formulated as [21, 25]:

$$P(\bigcup_{i=2}^{M} A_i) = \sum_{i=2}^{M} P(A_i) - \sum_{2<i<j\leq M} P(A_i \cap A_j) + \sum_{2\leq i<j<k\leq M} P(A_i \cap A_j \cap A_k) - \ldots + (-1)^{M-1} P(A_2 \cap \ldots \cap A_M). \tag{3}$$

Let m be the number of occluders in the region R. The target composed of infrared reflective markers would be obstructed by self or others in some region of occlusion denoted by R_i^t (Figure 1), and the volume of R_i^t is denoted by V_i^t. In order to estimate the probability of visibility for markers in the multiplayer motion capture system, any parts of the player should not be present in R_i^t.

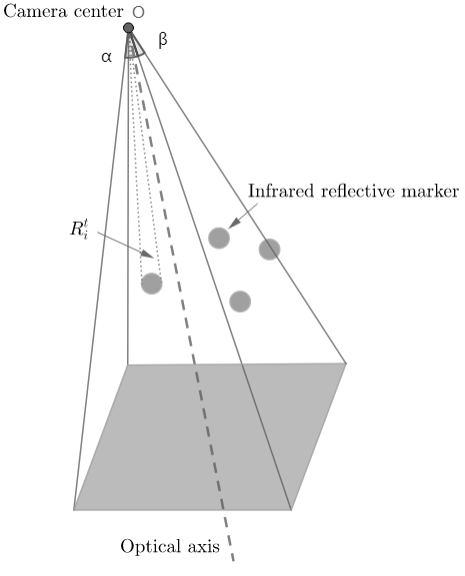

Figure 1: The skew pyramid model for cameras and infrared reflective marker in the region of interest R_i^t

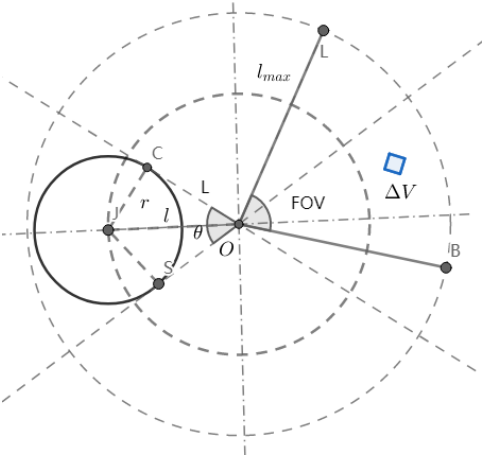

Figure 2: Dynamic occlusion analysis

Since people's positions and movements in the scene are independent, we can obtain the probability of the event that any parts of their bodies are not present in R_i^t is

$$P(A_i) = (\frac{V - V_i^t}{V})^m = (1 - \frac{V_i^t}{V})^{-\frac{V}{V_i^t}(-V_i^t \rho)}, \tag{4}$$

where $m = \rho V$, ρ is the density of occluders. From the limit theorem, we can approximately formulate the probability as

$$\lim_{V\to\infty} P(A_i) = e^{-\rho V_i^t}, \tag{5}$$

we discretize these occluders in the scene with sufficient small cubes. For V_i^t, we have

$\Delta V_i^t = \frac{\theta - 2\sin^{-1}(r/l)}{2\pi} d\Delta V$, $d\Delta V = \frac{1}{3}(2\pi)r(-\sin(\frac{\theta}{2}))(\frac{1}{2})\Delta l \Delta\theta = -\frac{1}{3}\pi r \sin(\frac{\theta}{2})\Delta l \Delta\theta$, and

$$\begin{aligned} V_i^t &= \int_0^{FOV} \int_{l_{\min}}^{l_{\max}} \frac{\theta - 2\sin^{-1}(r/l)}{2\pi} (-\frac{1}{3}\pi r \sin(\frac{\theta}{2})) dl d\theta \\ &= \int_0^{FOV} \int_{l_{\min}}^{l_{\max}} \frac{\theta - 2\sin^{-1}(r/l)}{6} (-r\sin(\frac{\theta}{2})) dl d\theta. \end{aligned} \tag{6}$$

If the radius of the marker $r \to 0$, the effective region of occlusion is a line of connection between the marker center and camera center, where $V_i^t \to 0$. Thus, the occlusion zone is negligible. For our case, the target has a constant value of radius r, and the event that it can be covered not only depends on whether it locates inside FoV but depends on whether occluders locate outside of the corresponding region of occlusion. We simulate the model with a circle of radius l to analyze all possible occlusions.

As shown in Figure 2, if the target draws near the camera, we can achieve the minimum distance of visible target $l_{\min}$ that is equal to $r/sin(\theta/2)$. θ denotes the smaller value of FoV. Similarly, if the target keeps away from the camera until the boundary of FoV, we can achieve the maximum distance of visible target $l_{\max}$ that is equal to $\sqrt{L^2 + r^2}$. Only the target located inside this circular region has the chance to be viewed given no overlap with other targets in the space.

For each marker inside V_i^t, the probability of targets that are visible, denoted by p, becomes a random value with respect to the distribution of occluders. Finding a closed-form $f(p)$ is quite difficult, but it is feasible to derive a more accurate informative model.

From (6), the boundary condition for the visible, we denote $P_m = e^{-\frac{4}{3}\pi(l_{\max}+r)^3 \frac{\theta}{2\pi}\rho} = e^{-\frac{2}{3}\theta(l_{\max}+r)^3\rho}$. P_m is a finite value, while $P(A_i)$ is the probability when obstacles locate outside of a circle centered at the camera with radius $l_{\max} + r$, shown in Figure 2, and it must be an impulse function at the boundary with amplitude P_m. To derive the edge condition of $P(A_i)$, we first extract a small portion near the boundary where $V_i^t = V_m - \Delta V$, where V_m denotes the volume in boundary situations and ΔV denotes an infinitely small volume. It satisfies when a target encroaches on an infinitely small volume ΔV from the region of occlusion. Since targets in the volume of ΔV have the probability of $sin^{-1}(r/l_m)/\pi$ to capture the target, $\Delta V_i^t = \Delta V \sin^{-1}(r/l_{\max})/\pi$, simultaneously, from (6), the probability that $V_i^t \in [V_m - \Delta V, V_m)$ is equal to $e^{-[\frac{4}{3}\pi(l_{\max}+r)^3 - \Delta V]\rho} - e^{-\frac{4}{3}\pi(l_{\max}+r)^3\rho} = \rho e^{-\frac{4}{3}\pi(l_{\max}+r)^3\rho}\Delta V$. Therefore, the limit near the boundary could be formulated as

$$f(p_m^-) = \lim_{\Delta V \to 0} \frac{\Delta P_m}{\Delta V_i^t} = \frac{\pi\rho e^{-\frac{4}{3}\pi(l_{\max}+r)^3\rho}}{\sin^{-1}(r/l_m)}. \tag{7}$$

Then, we can find that the probability density function is highly related to the density of obstacles, satisfying $f(p) \propto \rho e^{-\frac{4}{3}\pi(l_{\max}+r)^3\rho}$.

If we consider the situation that occluders cannot overlap each other, then the $(s+1)$-th obstacle has the remaining space of $V - sV_{ba}$, that equal to $\sum_{j=0}^{s-1} V_{ba}^j$, V_{ba}^j denotes the volume of the j-th obstacle, Hence, Equation (4) and (5) could be converted as $P(A_i) = \prod_{s=0}^{m-1}(1 - \frac{V_i^t}{V - sV_{ba}})$. Take the logarithm of both sides

$$\ln(P(A_i)) = \sum_{s=0}^{m-1} \ln(1 - \frac{V_i^t}{V - sV_{ba}}) \approx \int_0^m \ln(1 - \frac{V_i^t}{V - xV_{ba}}) dx. \tag{8}$$

As $V \to \infty$, based on the limit theorem, we can finally obtain the average dynamic visibility probability

$$P(A_i) \approx (1-\rho V_{ba})^{\frac{V_i^t}{V_{ba}}}, P(A_2 \cap ... \cap A_n) = \int_0^{\infty} (1-\rho V_{ba})^{\frac{V_i^t}{V_{ba}}} f(p^-)dp. \tag{9}$$

This probability spans the entire working space with respect to the locations of the target and occluders. If we know the distribution of V_{ba}, we can find the probability of the targets that are visible simultaneously by at least two cameras. In our simulation, we assume the location of V_{ba}, ρ and the occlusion areas V_i^t are normally distributed.

2.3 Formulation of the Objective

The camera configuration problem can be modeled as a Binary Integer Programming(BIP) model with an included occlusion culling factor, it can be formulated as

$$\begin{aligned}
\underset{x',y',z',\varphi}{\arg\max} \quad & \sum_{x=1}^{n_x}\sum_{y=1}^{n_y}\sum_{z=1}^{n_z} \lambda g_{xyz} \\
\text{s.t.} \quad & \sum_{x',y',z'} c_{x'y'z',\varphi,i} \geq 1, \;\; 1 - t_{xyz} \leq c_{x'y'z',\varphi,i} + g_{xyz} \leq 1 + t_{xyz}, \\
& \sum_{x=1}^{n_x}\sum_{y=1}^{n_y}\sum_{z=1}^{n_z} t_{xyz} \geq M, \;\; \sum_{i=1}^{M} c_i \cdot L_i \leq C_c, \;\; l_{i_{\min}} \leq \kappa \leq l_{i_{\max}}, \\
& x'_{\min} \leq x' \leq x'_{\max}, \;\; y'_{\min} \leq y' \leq y'_{\max}, \;\; z'_{\min} \leq z' \leq z'_{\max}, \\
& |\frac{x-x'}{z-z'}| \leq \tan(\frac{\alpha}{2}), \;\; |\frac{y-y'}{z-z'}| \leq \tan(\frac{\beta}{2}),
\end{aligned}$$

where $\lambda = P(\bigcup_{i=2}^{n} A_i)$, it can be derived from (3) and (9), g_{xyz} is a binary variable representing if the grid point (x,y,z) is covered by M cameras; $c_{x'y'z',\varphi,i}$ is a binary variable representing if the camera i is deployed at (x',y',z') with orientation φ; t_{xyz} denotes a binary variable at the grid (x,y,z) [30]. L_i denotes the price for the camera c_i, C_c denotes the total budget for the camera configurations. $l_{i_{\min}}$ and $l_{i_{\max}}$ are the minimum and maximum effective lengths of the ith camera. $\kappa = \frac{\|([x_i \; y_i \; z_i]^t - T)\cdot \mathbf{V}_o\|}{\|[x_i \; y_i \; z_i]^t - T\|\cdot\|\mathbf{V}_o\|}$, $\mathbf{V}_o$ is the direction of this point that is parallel to the optical axis. α and β denote the horizontal and the vertical FoV, respectively. n_x, n_y and n_z denote the number of grids in three directions, respectively. Our goal is to find the optimal x', y', z' and φ, to maximize the coverage of grid points x, y, z.

We will introduce greedy heuristics with the Riesz-particle scale to approximate the solution in section 3.

3 Greedy Heuristics with Riesz-Particle Scale

In this section, we will provide the procedure to approximate the solution to the objective.

Step I: If $\lambda(x,y,z)_j = 0$, for any $j = 1,2,...,n$, set $g(x,y,z)_j = 1$ and remove all constraints in which $g(x,y,z)_j$ appears with a coefficient of 1.

Step II: If $\lambda(x,y,z)_j > 0$, for any $j = 1,2,...,n$, and $g(x,y,z)_j$ does not appear with 1 as the coefficient in any of the remaining constants, set $g(x,y,z)_j = 0$;

Step III: For each of remaining variables, determine $\frac{\lambda(x,y,z)_j}{|\mathfrak{L}_j|}$, where $|\mathfrak{L}_j|$ is the number of constraints in which $g(x,y,z)_j$ appears with the coefficient 1, select the variable k' for which $\frac{\lambda(x,y,z)_j}{|\mathfrak{L}_j|}$ is minimum, set $g(x,y,z)_{k'}=1$ and remove all constants in which $g(x,y,z)_{k'}$ appears with the coefficient 1. Examine the resulting model.

Step IV: If there are no more constraints, set all the remaining variables $g(x,y,z)_j$ to 0 and stop, otherwise go to step I.

Step V: Based on the grid points (x,y,z) covered, to find the feasible camera location, where the visibility matrix from each camera i is related to the local grid points cluster $\mathfrak{C}(i)$ in the discretized space. We maximize the use of overlapping by

$$\mathbb{P}' \leftarrow \arg\max \left| \bigcup_{i=1}^{M} \bigcup_{j=1}^{M} (\mathfrak{C}(i)+\mathfrak{C}(j)-\mathfrak{C}(i)\cup\mathfrak{C}(j)) \right|. \tag{10}$$

Inspired by [10, 12], we introduce the Riesz energy to discretize rectifiable submanifolds of interest $\Omega \subset \mathfrak{C}$ via particle interaction(for grid points), where only a few samples, called Riesz particles, are required to scale the cardinality of $\mathfrak{C}(i)$.

$$\begin{aligned} &\min_{\mathbf{x_i},\mathbf{x_j}\in\Omega} \varepsilon_{\beta'}(\Omega,N) = \min_{\mathbf{x_i},\mathbf{x_j}} \left\{ \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{\omega(\mathbf{x_i},\mathbf{x_j})}{\| \mathbf{x_i}-\mathbf{x_j} \|^m} \right\}^{\frac{1}{m}}, \\ &\omega(\mathbf{x_i},\mathbf{x_j}) \propto e^{\left[\alpha'\cdot\gamma(\mathbf{x_i})\gamma(\mathbf{x_j})+\beta'\cdot\|\mathbf{x_i}-\mathbf{x_j}\|\right]^{-\frac{m}{2d}}}. \end{aligned} \tag{11}$$

β' is the local discrepancy coefficient and is positive to balance off the local conflict with the distributed points when short-range interaction between points is the dominant effect. As $m\to\infty$, the formulation is convex under mild conditions, the denominator approximates $\| \mathbf{x_i}-\mathbf{x_j} \|$, thus, our criterion inherits the properties of Riesz energy, termed as weighted Riesz β'-energy criterion. To obtain a finite collection of point sets that are distributed according to a specified non-uniform density such as might be used as points for weighted integration or design of complex surfaces where more points are required in regions with higher curvature. We introduce $\omega(\mathbf{x_i},\mathbf{x_j})$ in (11), where $\gamma(\mathbf{x}) \propto -\ln f(\mathbf{x})$, $\| \mathbf{x_i}-\mathbf{x_j} \|$ is contained so as to be locally bounded for $\alpha'=-1$. Thus, given a proper distribution $f(\mathbf{x})$, we can use $\varepsilon_{\beta'}(\Omega,N)$ to draw a sequence of N-point configurations that are "well-separated" and have asymptotic distribution $f(\mathbf{x})$.

To approximate the optimum, the procedure of a maximum overlapping coverage method is provided in Algorithm 1, where for each discretized subcover, we find the most use of the overlapping by $i \leftarrow \arg\max \frac{|\cup_{j=1}^{M}((S'\cap\mathfrak{C}(j))|}{|S'\cap\mathfrak{C}^*|}$, this will ensure that the current grid can be covered by as many cameras as possible. As the grid is sample-based, the cardinality of $\mathfrak{C}(i)$ has been large-scaled, then the time cost that can be measured by $\mathcal{O}(\sum_{i=1}^{M}|\mathfrak{C}(i)|)$ will be decreased greatly. The performance ratio $H(k)$ is the worst-case ratio of the size of the optimal configuration to the size of the approximate one with the largest overlap S.

Theorem 1 For $\max|\mathfrak{C}(i)| \le k, i=1,2,...,M, M, k \in \mathbb{R}^+$ and for all sufficient large $n \gg k$,

$$\sum_{j=1}^{k}(\frac{1}{j}) \le H(k) \le 1+\ln(k) \le \sum_{j=1}^{k}(\frac{1}{j})+\frac{1}{2}. \tag{12}$$

Algorithm 1: Set Covering for finding that subcover which has the maximum overlapping coverage

Input: Covered grid points $g(x,y,z)$ by $\mathbb{P}'$, $\{g(x,y,z)\} \subset \mathfrak{G}$, $\mathbb{P}' \subset \mathbb{P}$, $\mathfrak{G}$, $\mathfrak{C} = \bigcup_{i=1}^{n} \mathfrak{C}(i)$.
Output: The position and orientation of the cameras: $\mathbb{P}'$.
$\mathbb{P}' \leftarrow \varnothing$;
$\mathfrak{C}^* \leftarrow \varnothing$;
while $\{g(x,y,z)\} \nsubseteq \mathfrak{C}^* \& \mathfrak{G} \nsubseteq \varnothing$ **do**
  $\forall S' \in \mathfrak{C}, i \leftarrow \arg\max \frac{\left|\bigcup_{j=1}^{M}((S' \cap \mathfrak{C}(j))\right|}{|S' \cap \mathfrak{C}^*|}$;
  $\mathfrak{G} \leftarrow \mathfrak{G} \setminus \mathfrak{C}(i)$;
  $\mathfrak{C}^* \leftarrow \mathfrak{C}^* \cup \mathfrak{C}(i)$;
  $\mathbb{P}' \leftarrow \mathbb{P}' \cup \{P_{c_i}\}, P_{c_i} \in \mathbb{P}$;
end

4 Simulation and Results Analysis

We use three types of cameras, Vantage 8 [1], Vantage 5 [1] and RTS 4000 [2] in 15m × 15m × 3.7m. The assembled space of support frames fixed on the steel pipe occupies a size of 50cm × 50cm × 32cm so that the camera can be installed in any direction($[0, 180°]$). This also gives us more flexibility to deal with points that are obscured by static and dynamic factors.

For the Riesz energy model, the parameters that we use are $\beta' = 4, m = 40$. The number of particles that are unique is 250 for each camera-based cluster, here we provide a 20-camera configuration, with 12 V8 cameras and 8 V5 cameras, $\alpha \in [-60°, 60°]$ and $\beta \in [25°, 35°]$.

4.1 Comparison of the Different Approaches

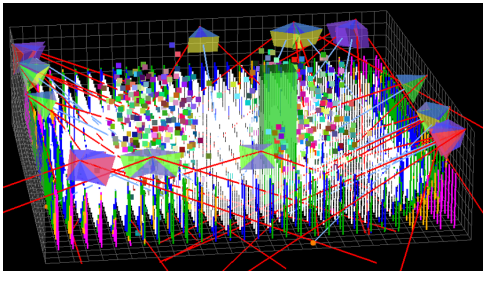

Figure 3: Camera configuration with greedy algorithm [16].

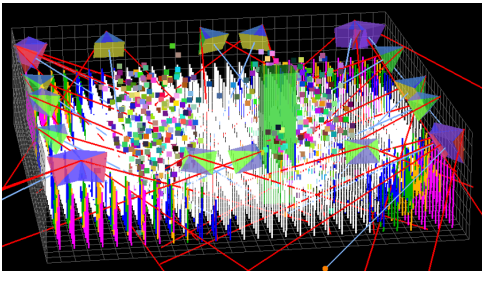

Figure 4: Camera configuration with multiple objectives[3].

In order to obtain the globally optimal camera configuration, we need to solve the whole BIP model point by point. However, for large-scale motion capture systems, the number of these grid points (34235) is huge and a large number of auxiliary variables will also be

[1] https://www.vicon.com/
[2] www.realis-e.com

introduced in the solution process, which will require sufficient memory and computational effort, while it is not applicable in the specific experimental simulation process.

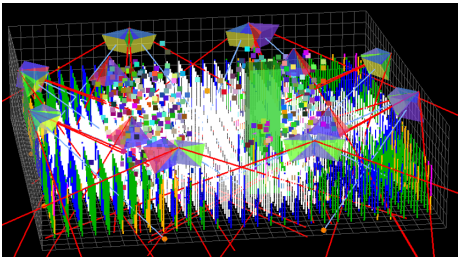

Figure 5: Camera configuration with our method.

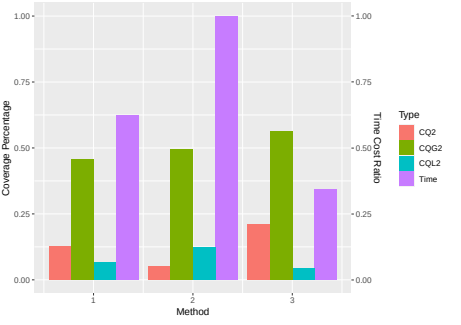

Figure 6: Static occlusion analysis for the greedy algorithm [16], multiple objectives method [3] and ours, respectively.

Instead, we focus more on how to evaluate the quality of the proposed heuristics and how flexible the proposed solution approximates the global optimum in the application. We sampled 60 sets of camera positions, each of which is uniformly distributed over the entire 3D contour. To ensure the consistency of populations $\mathfrak{C}$ and the diversity of individuals $\mathfrak{C}(i)$, we sample with the posterior α and β by evolutionary algorithms to derive the set of coverage points $\mathfrak{C}(i)$ of the camera in the current state where the fitness is reachable. We use EOlib [17], with the following parameters: the size for tournament selection $t' = 4$; the offspring is a linear combination of parents with a rate $c' = 0.8$; the range for real uniform mutation $\varepsilon' = 0.1$ with the rate $m' = 0.5$; we use real-valued fitness by minimizing the total penalty score to evaluate the offspring, for white and blue points covered by at least 3 cameras, the score is 1; for green points covered by 2 cameras, the score is 30; for pink points covered by only 1 camera, the score is 1000; for red points that are invisible, the score is 10000.

We compared our proposal with two other approximation algorithms. Hörster and Lienhart [16] presented an iterative that places one camera during each iteration, while we update all the cameras' locations from the candidate sets instead, then compute a list of grid points that are adequately covered for every camera position, pose and type combination, in the greedy traverse, the objective is to search for the position-orientation-type combination with the highest rank. Chen et al. [3] presented the optimization of coverage and sufficient overlapped areas in a continuous space to improve the camera hand-off success.

Here, we translate it into the statistical analysis of the average included angle for the points of interest. Both the algorithms assume that the camera's position and poses are provided in advance, although our proposal has relaxed this assumption, we still follow it to maintain consistency for comparison. The performance of the algorithms has been tested separately, we obtained average results by conducting multiple simulations, the rendering results are shown in Figure 3, Figure 4 and Figure 5, respectively. The statistical analysis of the static coverage and the time consumed are normalized and both of them are shown in Figure 6, where the percentage of points obtained by our method that are covered simultaneously by at least 2 cameras is the largest (for the legends, "CQ2" means the camera quantity is 2, "CQG2" means the camera quantity is greater than 2 and "CQL2" means the camera quantity is less than 2). The fourth column is the time's cost ratio: 723338ms for

Approaches	Static Visibility	Dynamic Visibility	Total Penalty Score
Greedy algorithm [16]	88.44%	53.34%	173245495
Multiple objectives [3]	87.71%	73.71%	336689528
Ours	**95.67%**	**88.56%**	**36471524**

Table 1: Comparison of the different approaches for optimal camera configurations.

the greedy algorithm [16], 435832*ms* for multiple objectives method [3] and 243963*ms* for ours, respectively. The result shows that our approach significantly outperforms the other two competing techniques. For a single point, if the range of the included angle is larger, more cameras will be able to cover it, then, it is less likely to be dynamically occluded. If we define the dynamic visibility $\lambda' = \frac{\sum_{i=1}^{n'} \delta(\Delta\tilde{\theta}_i \geq \pi/2)}{n'}, \Delta\tilde{\theta}_i = \max\tilde{\theta}_i - \min\tilde{\theta}_i, \tilde{\theta}_i$ denotes the included angle for the point P_i and the static visibility $\lambda'' = \frac{\sum_{i=1}^{n} \delta(g_{xyz} \geq 2)}{n}$, $\delta(\cdot)$ is the impulse function, we obtain the largest ratio for both static visibility and dynamic visibility with the lowest score penalty which is shown in Table 1.

5 Conclusion

We have proposed a probabilistic dynamic occlusion model that reflects the target self and mutual occlusion behavior which is commonly found in the feature-based motion capture system. Then, we develop a BIP with the occlusion culling factor for the optimal camera configuration and translate it as the set cover problem. We further introduce the greedy heuristic with the Riesz-particle Scale to approximate the solution. The optimal camera configuration problem is $\mathcal{NP}$-hard. The computation of the constrained optimization is sample-based, we prove that our performance ratio $H(k)$ grows at most logarithmically, under mild assumptions.

6 Acknowledgements

We thank Tomohiro Nagasaka (Shenzhen Realis Multimedia Technology Co., Ltd.) for providing partial code for the experiments. We also thank Qiuzi Xu (Shenzhen Realis Multimedia Technology Co., Ltd.) for providing cameras and real test sites.

References

[1] Mohammad Al Hasan, Krishna K Ramachandran, and John E Mitchell. Optimal placement of stereo sensors. *Optimization Letters*, 2(1):99–111, 2008.

[2] Krishnendu Chakrabarty, S Sitharama Iyengar, Hairong Qi, and Eungchun Cho. Grid coverage for surveillance and target location in distributed sensor networks. *IEEE Transactions on Computers*, 51(12):1448–1453, 2002.

[3] Chung-Hao Chen, Yi Yao, Wei-Wen Hsu, Andreas Koschan, and Mongi Abidi. Continuous camera placement using multiple objective optimisation process. *IET Computer Vision*, 9(3):340–353, 2015.

[4] Xing Chen and James Davis. Camera placement considering occlusion for robust motion capture. *Computer Graphics Laboratory, Stanford University, Tech. Rep*, 2(2.2):2, 2000.

[5] Xing Chen and James Davis. An occlusion metric for selecting robust camera configurations. *Machine Vision & Applications*, 19(4), 2008.

[6] Dimitrios Chrysostomou and Antonios Gasteratos. Optimum multi-camera arrangement using a bee colony algorithm. In *2012 IEEE International Conference on Imaging Systems and Techniques Proceedings*, pages 387–392. IEEE, 2012.

[7] Vasek Chvátal. A combinatorial theorem in plane geometry. *Journal of Combinatorial Theory, Series B*, 18(1):39–41, 1975.

[8] Vasek Chvatal. A greedy heuristic for the set-covering problem. *Mathematics of operations research*, 4(3):233–235, 1979.

[9] Cregg K. Cowan and Peter D Kovesi. Automatic sensor placement from vision task requirements. *IEEE Transactions on Pattern Analysis and machine intelligence*, 10(3): 407–416, 1988.

[10] Xiongming Dai and Gerald Baumgartner. Weighted Riesz particles. Submitted to the 38th Annual AAAI Conference on Artificial Intelligence, 2023.

[11] Xiongming Dai and Gerald Baumgartner. Operator-free equilibrium on the sphere. https://doi.org/10.48550/arXiv.2310.00012, 2023.

[12] Xiongming Dai and Gerald Baumgartner. Chebyshev particles. https://doi.org/10.48550/arXiv.2309.06373, 2023.

[13] Xiongming Dai and Gerald Baumgartner. Variance reduction of resampling for sequential Monte Carlo. https://doi.org/10.48550/arXiv.2309.08620, 2023.

[14] Uğur Murat Erdem and Stan Sclaroff. Automated camera layout to satisfy task-specific and floor plan-specific coverage requirements. *Computer Vision and Image Understanding*, 103(3):156–169, 2006.

[15] Jose-Joel Gonzalez-Barbosa, Teresa Garcia-Ramirez, Joaquin Salas, Juan-Bautista Hurtado-Ramos, et al. Optimal camera placement for total coverage. In *2009 IEEE International Conference on Robotics and Automation*, pages 844–848. IEEE, 2009.

[16] Eva Hörster and Rainer Lienhart. On the optimal placement of multiple visual sensors. In *Proceedings of the 4th ACM international workshop on Video surveillance and sensor networks*, pages 111–120, 2006.

[17] Maarten Keijzer, Juan J Merelo, Gustavo Romero, and Marc Schoenauer. Evolving objects: A general purpose evolutionary computation library. In *International Conference on Artificial Evolution (Evolution Artificielle)*, pages 231–242. Springer, 2001.

[18] Brenden Keyes, Robert Casey, Holly A Yanco, Bruce A Maxwell, and Yavor Georgiev. Camera placement and multi-camera fusion for remote robot operation. In *Proceedings of the IEEE international workshop on safety, security and rescue robotics*, pages 22–24. National Institute of Standards and Technology Gaithersburg, MD, 2006.

[19] Scott Mason. Heuristic reasoning strategy for automated sensor placement. *Photogrammetric engineering and remote sensing*, 63(9):1093–1102, 1997.

[20] Anurag Mittal and Larry S Davis. Visibility analysis and sensor planning in dynamic environments. In *European conference on computer vision*, pages 175–189. Springer, 2004.

[21] Anurag Mittal and Larry 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.

[22] Dennis Nieuwenhuisen and Mark H Overmars. Motion planning for camera movements. In *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA'04. 2004*, volume 4, pages 3870–3876. IEEE, 2004.

[23] Joseph O'Rourke. An alternate proof of the rectilinear art gallery theorem. *Journal of Geometry*, 21(1):118–130, 1983.

[24] Joseph O'rourke. *Art gallery theorems and algorithms*, volume 57. Oxford New York, NY, USA, 1987.

[25] Cheng Qian and Hairong Qi. Coverage estimation in the presence of occlusions for visual sensor networks. In *International Conference on Distributed Computing in Sensor Systems*, pages 346–356. Springer, 2008.

[26] Pooya Rahimian and Joseph K Kearney. Optimal camera placement for motion capture systems. *IEEE transactions on visualization and computer graphics*, 23(3):1209–1221, 2016.

[27] Bill Triggs and Christian Laugier. Automatic camera placement for robot vision tasks. In *Proceedings of 1995 Ieee International Conference on Robotics and Automation*, volume 2, pages 1732–1737. IEEE, 1995.

[28] Jorge Urrutia. Art gallery and illumination problems. In *Handbook of computational geometry*, pages 973–1027. Elsevier, 2000.

[29] Xiaohui Wang, Hao Zhang, and Haoran Gu. Solving optimal camera placement problems in IoT using LH-RPSO. *IEEE Access*, 8:40881–40891, 2019.

[30] H Paul Williams. *Model building in mathematical programming*. John Wiley & Sons, 2013.

[31] Yi Yao, Chung-Hao Chen, BESMA ABIDI, DAVID PAGE, ANDREAS KOSCHAN, and MONGI ABIDI. Multi-camera positioning for automated tracking systems in dynamic environments. *International Journal of Information Acquisition*, 7(03):225–242, 2010.

[32] Jian Zhao, Sen-Ching Cheung, and Thinh Nguyen. Optimal camera network configurations for visual tagging. *IEEE Journal of Selected Topics in Signal Processing*, 2(4): 464–479, 2008.

[33] Jian Zhao, Ruriko Yoshida, Sen-ching Samson Cheung, and David Haws. Approximate techniques in solving optimal camera placement problems. *International Journal of Distributed Sensor Networks*, 9(11):241913, 2013.