Algorithms and Systems for Shared Access to a Robotic Streaming Video Camera *

[Doctoral Symposium Paper]

Dezhen Song
IEOR Department
University of California, Berkeley
CA 94720-1777, USA

Categories and Subject Descriptors

H.5.2 [**INFORMATION INTERFACES AND PRESENTATION**]: Group and Organization Interfaces—*Web-based interaction*; H.4.3 [**Information System Applications**]: Communication Applications

1. INTRODUCTION

Robotic streaming video cameras with pan, tilt, and zoom controls are now commercially available and are being installed in hundreds of locations around the world[1]. Remote viewers can adjust camera parameters via the Internet to observe desired details in the scene. Current methods restrict control to one user at a time; users have to wait in a queue for their turn to operate the camera. In this thesis, we develop ShareCam, a new approach that eliminates the queue and allows many users to access and share control of the robotic camera simultaneously.

Since conflicting frame requests are made by users, a primary challenge is computing optimal camera parameters. We formalize the problem using a new metric, Intersection Over Maximum (IOM), to model the degree of satisfaction for each user, and seek to maximize total satisfaction for n users. We develop online algorithms to solve this optimization problem for cases where pan, tilt and zoom values can be either discrete or continuous variables. Recognizing that computing resources increases with the number of users, we also propose distributed algorithms to further improve computation speed by allocating computation to client computers. An implemented version of the system is online at http://tele-actor.net/sharecam/

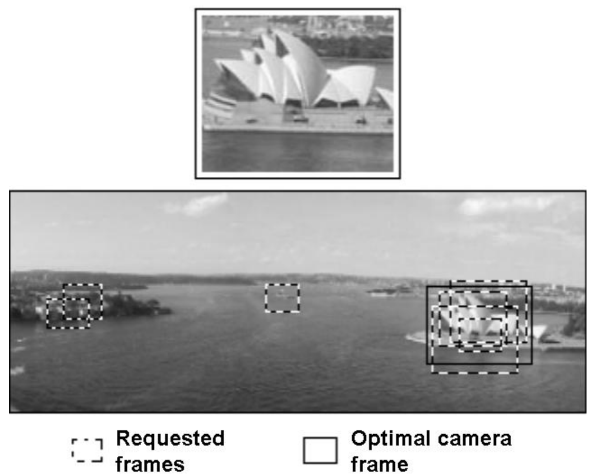

Figure 1: ShareCam user interface. Each Internet-based user sees two image windows. The upper window is live streaming video from the robotic camea. The lower window is a fixed "panorama" image of the camera's reachable range of view. Users request desired camera frames by positioning dashed rectangles in the lower window. Every 5-10 seconds, based on the frame requests from all current users, the system computes an optimal camera frame (shown with solid rectangle), moves the camera accordingly, and displays the resulting live video.

2. RELATED WORK

Kimber, Liu, Foote et al. describe a multi-user robot camera for videoconferencing in [7]. Similar to Sharecam, they formulate the robot parameter selection for multiple simultaneous users as an optimization problem based on position and area of overlap. To solve the problem, they propose an approximation based on the bounding boxes of all combinations of user frames. Our approach reduces running time from exponential to polynomial in the number of users and we provide formal bounds on the errors resulting from grid-based approximations.

*This work was supported in part by the National Science Foundation under IIS-0113147, by Intel Corporation, by Microsoft Corporation, and in part by UC Berkeley's Center for Information Technology Research in the Interest of Society (CITRIS). For more information please contact dzsong@ieor.berkeley.edu or goldberg@ieor.berkeley.edu.

[1]http://www.x-zone.canon.co.jp/WebView-E/index.htm

As shown in figure 1, the Sharecam interface allows users to control a robot camera with 3 degrees of freedom. In the taxonomy proposed by Tanie et al. [2], ShareCam is a Multiple Operator Single Robot (MOSR) system. Collaborative Telerobotics is motivated by applications such as education and journalism, where groups of users desire simultaneous access to a single robotic resource.

One precedent of an online MOSR system is described in McDonald, Cannon and colleagues [8]. For waste cleanup, several users assist in waste cleanup using Point-and-Direct (PAD) commands [1]. In this Internet-based MOSR system, collaboration is serial but pipelined, with overlapping plan and execution phases. The authors demonstrate that such collaboration improves overall execution time but do not address conflict resolution between users.

In [3] Goldberg and Chen analyze a formal model of collaborative control describe Internet-based MOSR system that averaged multiple human inputs to simultaneously control a single industrial robot arm. In [4, 5] we propose the "Spatial Dynamic Voting" (SDV) interface. The SDV collects, displays, and analyzes sets of spatial votes from multiple online operators using a Gaussian point clustering algorithm developed to guide the motion of a remote human "Tele-Actor".

3. EXPECTED CONTRIBUTIONS

In [11], we formulate the ShareCam problem. Let c be a vector of camera parameters that users can control, in this case a camera frame $[x, y, z]$, where x, y specify the center point of the frame, which is corresponding to pan and tilt, and z specifies size of the frame, which corresponds to zoom level. c defines a rectangular camera frame (the camera has a fixed aspect ratio of 4:3). User i requests a desired frame r_i. Given requests from n users, the system computes a single global frame c^* that will best satisfy the set of requests.

We define a Generalized Intersection Over Maximum (GIOM) metric for user "satisfaction" $s(c, r_i)$ based on how the user's requested frame r_i compares with a candidate camera frame c. Each of n users submits a request. Let

$$s(c) = \sum_{i=1}^{n} s_i(r_i, c) \tag{1}$$

In the memoryless frame selection model, we want to find c^*, the value of c that maximizes $s(c)$ based only on the current set of requests: $\max_c s(c)$. In each motion cycle, we servo the camera to this frame. In [11], we reported initial results on exact algorithms for discrete zoom levels: for n users and m zoom levels, the exact algorithm runs in $O(n^2 m)$ time. Har-Peled et al. [6] improved the exact algorithm to $O(mn^{3/2} \log^3 n)$ and proposed a near linear $\epsilon-$approximation algorithm.

In [10], we expand the geometric problem for continuous pan, tilt, and zoom and propose a grid-based $\epsilon-$ approximation algorithm in $O(n/\epsilon^3)$. In [9] we describe interface, system architecture, and implementation of a preliminary system.

We are currently working on exact algorithms for continuous pan, tilt, and zoom. The preliminary result show that we can do it in $O(n^3)$. We will also explore advanced data structures and algorithmic technique to further improve algorithm performance. We are also working on automated panorama image generation and calibration techniques that ensure the precise correspondence between the panorama frame and the associated live streaming video.

4. ACKNOWLEDGMENTS

We are grateful to Frank van der Stappen, Sariel Har-Peled, Vladlen Koltun, Dana Plautz, Anatoly Pashkevich, Cory Cox, Yong Rui, J. Foote, Q. Lui, and Eric Paulos for insightful discussions and feedback. Thanks to M. Faldu and W. Guan for their contributions to the ShareCam system, and I.Y. Song, K. "Gopal" Gopalakrishnan , Ron Alterovitz, A. Levandowski, M. Mckelvin, A. Ho, and J. Vidales for helpful suggestions.

5. REFERENCES

[1] D. J. Cannon. *Point-And-Direct Telerobotics: Object Level Strategic Supervisory Control in Unstructured Interactive Human-Machine System Environments.* PhD thesis, Stanford Mechanical Engineering, June 1992.

[2] N. Chong, T. Kotoku, K. Ohba, K. Komoriya, N. Matsuhira, and K. Tanie. Remote coordinated controls in multiple telerobot cooperation. In *IEEE International Conference on Robotics and Automation*, volume 4, pages 3138–3343, April 2000.

[3] K. Goldberg and B. Chen. Collaborative control of robot motion: Robustness to error. In *International Conference on Intelligent Robots and Systems (IROS)*, volume 2, pages 655–660, October 2001.

[4] K. Goldberg, D. Song, Y. Khor, D. Pescovitz, A. Levandowski, J. Himmelstein, J. Shih, A. Ho, E. Paulos, and J. Donath. Collaborative online teleoperation with spatial dynamic voting and a human "tele-actor". In *IEEE International Conference on Robotics and Automation (ICRA)*, volume 2, pages 1179–84, May 2002.

[5] K. Goldberg, D. Song, and A. Levandowski. Collaborative teleoperation using networked spatial dynamic voting. *The Proceedings of The IEEE*, 91(3):430–439, March 2003.

[6] S. Har-Peled, V. Koltun, D. Song, and K. Goldberg. Efficient algorithms for shared camera control. In *19th ACM Symposium on Computational Geometry, San Diego, CA*, June 2003.

[7] Q. Liu, D. Kimber, L. Wilcox, M. Cooper, J. Foote, and J. Boreczky. Managing a camera system to serve different video requests. In *Proceedings of IEEE International Conference on Multimedia and Expo (ICME), Lausanne, Switzerland*, volume 2, pages 13–16, Aug. 2002.

[8] M. McDonald, D. Small, C. Graves, and D. Cannon. Virtual collaborative control to improve intelligent robotic system efficiency and quality. In *IEEE International Conference on Robotics and Automation*, volume 1, pages 418–424, April 1997.

[9] D. Song and K. Goldberg. Sharecam part I: Interface, system architecture, and implementation of a collaboratively controlled robotic webcam. In *IEEE/RSJ International Conference on Intelligent Robots (IROS)*, Nov. 2003.

[10] D. Song, A. Pashkevich, and K. Goldberg. Sharecam part II: Approximate and distributed algorithms for a collaboratively controlled robotic webcam. In *IEEE/RSJ International Conference on Intelligent Robots (IROS)*, Nov. 2003.

[11] D. Song, A. F. van der Stappen, and K. Goldberg. Exact and distributed algorithms for collaborative camera control. In *The Workshop on Algorithmic Foundations of Robotics, December*, Dec. 2002.