Generation of Engineering Research Directions through Artistic Process

Marco Pinter[1], Angus Graeme Forbes[2], Danny Bazo[1], and George Legrady[1]

[1]Media Arts and Technology Program, University of California, Santa Barbara
[2]Department of Computer Science, University of Illinois at Chicago
{pinter,dannybazo,legrady}@mat.ucsb.com,aforbes@uic.edu

Abstract. This paper describes the creation of *SwarmVision*, a system of autonomous robotic cameras that functions both as the basis for an art installation and as an instrument for generating novel engineering research. Through a series of interviews with engineers who examined the project, we illustrate how experimentation with the instrument could lead to potential new research directions in computer vision, machine learning, swarm robotics, remote collaboration, and visualization. This suggests that an unstructured and aesthetic approach to research can inform and inspire new engineering research directions, within and beyond the scope of this particular project.

Key words: Media arts methodologies, computational photography, swarm robotics, autonomous robotic cameras, arts and engineering.

1 Introduction

SwarmVision is a project with both artistic and engineering outputs, exploring novel uses of autonomous robotic cameras, computer vision techniques, and computational photography. It invites the public to reflect on how automation and intelligent spaces are transforming our relationship with the built environment. The project's automated system translates human photographic vision into rules that govern machine vision. A key component of the project has been to position it into a public context through installations in public museums, conferences, and other exhibition venues. The project integrates both arts concerns (such as philosophy of aesthetics, installation design, and dramaturgy for the behavior of the robots) and the exploration of engineering topics (such as computer vision, control theory and signal processing). The fundamentally interdisciplinary nature of the project has brought together artists and engineers with a variety of specialties to work toward a common goal. Development and experimentation with the system was conducted in an unstructured, exploratory fashion, often focused on aesthetics. As the project evolved, it became apparent that this process could provide significant benefit to informing and generating novel research directions in engineering.

While creative projects sometimes promote, test, or augment engineering research, they are less often used as a means to generate new research directions [3].

Fig. 1. *SwarmVision* installation at SIGGRAPH 2013 in Anaheim, CA.

However, some artists and researchers have sought to situate arts projects as a space where potential "blind spots" in scientific practice can be acknowledged and resolved. Because there are no obvious or standardized ways to represent probabilistic, fleeting, or otherwise unintuitive data, this necessitates the attention of an artistic sensibility that is trained in thinking about issues of representation [11].

The use of *generation* as an effective methodological approach to interdisciplinary projects has been proposed in [4]. Generation is characterized by the following activities:

- Initially implementing ideas without clear hypotheses about outcomes;
- Clarifying limitations of personal expertise and the boundaries of a line of inquiry or a discipline;
- Bringing existing ideas into new contexts;
- Switching between aesthetic, research, and engineering roles.

An artistic project is thus seen as a mechanism to explore the boundaries of a problem and, more specifically, to trace the contours of what is already known about a particular topic, and to provide a bottom-up approach to creating new technologies. This requires the expertise to switch between speculative roles and engineering roles, or the ability to find and communicate with those with appropriate expertise. Thus, it becomes important as well to have the tools to *mediate* between disciplines. Some recent artistic projects that explicitly made use of generation as a methodological tool to investigate new research topics include *Fluid Automata* [6], *Cell Tango* [5], *Annular Genealogy* [7], and the *New Dunites* project [1]. In the following sections we explore methods of generation in relation to *SwarmVision*. A further contribution of this paper is our use of open-ended interviews as a way to evaluate and solicit feedback about the effectiveness of the project in generating novel research ideas.

2 Design and Implementation of a Robotically-Actuated Multi-Camera Instrument

The multi-camera instrument for the *SwarmVision* installation has been described previously by the authors [9]. In this section we provide a detailed summary of the features of the project relevant to this discussion. Each camera system is referred to as a "robot," and consists of a pan-tilt-zoom camera that translates along a wall or other support structure and analyzes the scene via a particular computer vision algorithm. The platform layout is flexible and allows multiple configurations. Each robot outputs both the raw image in its view as well as a score of the visual interest of the subject matter currently in its view. The robots study incoming video frames from their respective cameras and compute features of interest, then re-assign their target gaze locations to the sub-region with the most interesting features.

Fig. 2. *SwarmVision* robot camera visualizations.

A robot's motion is guided initially by its current target position (which likely contains "interesting" features) and where that target location lies relative to the robot's current position on the rail, along with its camera's pan, tilt and zoom levels. Additionally, zoom level is adjusted incrementally from wide-angle to telephoto, following a maximization scheme. The nature of the features being detected can be highly dependent on scale, e.g. curves in a wide view of a room versus curves on a single zoomed-in flower, and thus the robots tend to achieve a zoom level that maximizes the number of features it finds interesting. Finally, a robot that finds particularly salient features in a region can command the other robots to turn their gaze to that area and proceed with their own particular algorithmic study of the region.

3 Implementation and Distribution of the Artwork

This instrument was utilized as the basis of an art installation, designed in the Experimental Visualization Laboratory at UC Santa Barbara and subsequently exhibited internationally. *SwarmVision* was premiered at SIGGRAPH 2013 in Anaheim [10], and subsequently featured at the Vox Gallery at Mois de la Photo in Montreal [2], the Run Run Shaw Creative Media Center in Hong Kong, and the Miami Art Fair.

SwarmVision explores the translation of rules of human photographic behavior to machine language. Viewers entering the room and watching the cameras become instantly aware that the robots are studying and analyzing the scene. Through projected visualizations, viewers become cognizant of the robots' autonomy and their search for subjects of interest based on each robots unique perceptual interests. Spectators may reflect on the technological and cultural roles of autonomous systems.

The installation space contains two screens or projections on which are depicted four visualizations. The first screen represents the perceptions of each of the three cameras, in other words a depiction of what each camera is seeing with its unique computer vision algorithm. The second screen represents a non-literal representation of the three cameras explorations, composed for aesthetic interest. Images from the robots cameras are placed at locations based on the orientation, focus and zoom of the cameras at the moment in time when the images were captured. The non-literal 3D space becomes filled with hundreds of stacked photographs, creating evolving volumetric forms of new and fading images. This process results in the natural emergence of unexpected patterns.

4 An Artistic Approach to Inspiring Research Directions

SwarmVision engages in the study of images, guided by an approach based in aesthetics and visual language, and enabled by computer vision algorithms and computational photography. The system is designed to facilitate illumination of current research while inspiring further engineering development. More specifically, we aimed to inform scientific research in computer vision and computational photography through an aesthetics-based and semiotics-based approach to camera systems engineering, and to enable engineering researchers to visualize computational photography and image construction in new ways that provoke new insight.

While engineering research directs itself to enhancing signal quality, artistic practice tends to embrace noisy systems [8]. For instance, whereas a tool such as Microsoft's Photosynth [12] will normalize and blend images so that they seamlessly stitch together into an apparently single image, our system is based on *stacking* images as they are acquired, thereby exposing the underlying autonomous investigations and illustrating the noisy, but meaningful, intermediate moments. Through the automated process of exploring, centering, zooming and maximizing saliency, the system results in a dense layering of continuous images. Studying the resultant clustered composition enables a researcher to see both the scene and its *process* of automated construction, simultaneously, in a manner not possible with a literal engineering reconstruction. With this process, we dwell into a perceptional situation that sidesteps standard engineering work, and addresses a messy and realistic scenario.

Fig. 3. Camera view of autonomous image capture process viewed in the 3D virtual scene overview.

Put another way, artists may be initially inspired by nature or an artwork or some other stimulus. Once inspired, a critical part of the artistic process is active unstructured exploration, creating accidents and unexpected circumstances and watching that until a clear definition of the problem (the artwork) emerges. Finally the artist creates or constructs their piece. Looking instead at scientists and engineers, they also may be inspired by nature or art or other stimuli. However, typically missing from the process is any stage of unstructured exploration before the problem is defined. In other words, the moment of inspiration is followed immediately by the careful design and implementation of the end product, followed by testing and evaluation in a controlled way that can be graphed and charted. We posit that introducing an unstructured, exploratory step may facilitate creative inspiration and "optimize" the concept/hypothesis *before* structured implementation commences.

4.1 Interviews

The *SwarmVision* creators engaged in lengthy interviews with four UC Santa Barbara engineering researchers regarding potential research directions inspired from the artwork, and also, more generally, taking an unstructured approach to generating engineering research directions. These individuals interviewed included: (a) a faculty member focusing on computer graphics, imaging, visualization, and computer vision; (b) a faculty member focusing on image/video analysis, feature extraction, steganography and signal/image processing; (c) a PhD researcher focused on computational photography and computer vision; and (d) a PhD Candidate in computer vision focused on augmented reality. In framing the research, we presented the work as not a problem-based project or an optimization-based project, but rather as an instrument whose use suggests new possibilities, reframing our own perceptions and allows us to intermingle imagination and experience in the inspiration of new engineering research directions and solutions. The use of the instrument in this context requires the engineer to, temporarily, stop acting as a problem solver and enter a role of *exploration.*

Some interviewees embraced the concept of using an unstructured, aesthetic approach to engineering exploration, and were open to seeing what this approach could uncover. These conversations led to interesting potential engineering research directions, discussed in more detail in the next section. This was especially true of engineers involved in computer graphics, computational photography, augmented reality and HCI, in which, according to one researcher, the nature of a human viewer "in the loop" already makes the research slightly less about engineering optimization and slightly more in the direction of aesthetics.

Other engineering researchers were resistant to embracing the utility of the unstructured research and to taking an intuitive, exploratory approach. The lack of a clearly defined problem, hypothesis and evaluation process was antithetical to their process and what they perceived as a proper engineering approach employing hypotheses, goals, structured research, testing and evaluation according to specific criteria.

4.2 Engineering Research Directions

The interviews discussed above yielded a number of promising engineering research directions that could potentially utilize the instrument described herein. Interviewees pointed out some specific uses in research as well general advantages of the system to ongoing research. The interviewees were presented in advance with live or video documentation of the robotic instrument in process, and visualizations of the aggregated and messy image data. As discussed previously, the visualization provides a single dynamic representation of both the content of the scene and how that content was assembled (i.e. through movement of multiple robotic cameras.) This is a viewpoint that would not be apparent though a literal engineering reconstruction of the scene, nor by graphs of the robotic camera motion. We posit that in some cases, this dualistic representation facilitated brainstorming by the interviewees on research directions they would not have arrived at otherwise. A list of these potential research directions, suggested by the interviewees, follows below.

Telepresence, Collaboration and Education In the domain of remote collaboration and education, there is often the requirement of seeing a remote object or scene. Choosing a best view in these cases is an ill-defined task. A robotic camera instrument such as that designed here, in combination with a simple and intuitive user interface, could provide an intelligently-assisted piloting of multiple degree-of-freedom cameras to provide a useful aggregated view. Using ongoing video feeds as guidance, users could move the cameras to optimize the video orientations and get a sense of the space, as one would if walking around in person. This process could be live during remote consultations, or set up in advance for local viewpoint optimization. Research could also focus on where in the range of assistance to autonomy the system would function most optimally.

Understanding Human-Based Viewing and Machine Learning of Viewing Rules Experimentally, subjects could be brought into a space and use the instrument to pick views which are "best," either from the point of view of aesthetics or understanding content. Subjects could be divided into pools of amateurs from the general population versus photographers and visual artists. The results may be studied for identification of systematic elements or rules that artists (or the general public) look for when analyzing scenes. While much of this analysis often involves intuition, and operates at an implicit level, a study of the above results may point to some number of contributing rules. These rules may serve as insights for automating such behavior using computer vision and graphics techniques. Thus machine learning methods for computer vision and compositional framing may be significantly enhanced by methods gleaned from studies on human subjects with this instrument.

Visualizing Multiple Viewpoints in Meaningful and Useful Ways More generally, engineering researchers may employ the instrument to explore new directions in visualizing multiple perspectives simultaneously, and/or capturing a time element in visualizing a scene or perspective motion over time. The instrument facilitates rapid play and exploration of perspective representation in a way that can yield unexpected results. While a standard engineering approach to this type of visualization might be inspired by a single concept from the mind of the engineer followed by months of research and implementation on that single approach, an *unstructured* exploration may uncover multiple approaches before a single optimal approach is chosen for implementation.

Additional Research Possibilities The interviewees also suggested that, after further study, the instrument may possibly contribute to the following research areas: Multi-agent systems, swarm behavior, detecting occlusion, object tracking, activity recognition, intelligent data capture, and scene reconstruction, among others.

5 Conclusion and Future Directions

This paper describes a successful collaborative interdisciplinary arts project and also presents an example of an effective methodological approach to generate new research with both artistic *and* engineering outputs. We have examined how an arts project may lead to new research, technology, and applications. The multi-camera instrument created for *SwarmVision* could support a number of additional features, as described in [9]. Once such a revised instrument is available, engineering partners will be brought in to brainstorm additional novel research. Future work will involve collaborating with the engineering partners who were inspired to suggest these directions, and facilitating laboratory access to their PhD students to pursue them. Finally, new instruments may be developed, perhaps in conjunction with artistic installation, again generated via an unstructured and aesthetic approach.

Acknowledgments Special thanks to our colleagues Daniel Vaquero, Ahana Reddy, Yuan-Yi Fan for contributions of computer vision and image processing algorithms, literature review, and for the development of a stable controller for robotic diagonal movement. This project has been funded by the Robert W. Deutsch Foundation and National Science Foundation IIS #1149001 grant.

References

[1] Burbano, A., Bazo, D., DiCicco, S., and Forbes, A. G. (2012). The new dunites. In *Proceedings of the ACM International Conference on Multimedia (MM)*, pages 1501–1502, Nara, Japan.

[2] ExpVisLab (2014). Swarm vision. In Wombell, P., editor, *Drone: The Automated Image— Le Mois de la Photo à Montréal*, pages 66–69. Karber Verlag.

[3] Forbes, A. G. (2013). Media arts roles in art-science collaborations. In *Proceedings of the Re-new Digital Arts Festival*, pages 276–282, Copenhagen, Denmark.

[4] Forbes, A. G. (2014). *Emerging Methodologies for Interdisciplinary Research Practice*. PhD thesis, Media Arts & Technology Program, University of California, Santa Barbara.

[5] Forbes, A. G., Höllerer, T., and Legrady, G. (2010). Behaviorism: A framework for dynamic data visualization. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1164–1171.

[6] Forbes, A. G., Höllerer, T., and Legrady, G. (2013). Generative fluid profiles for interactive media arts projects. In *Proceedings of the International Symposium on Computational Aesthetics in Graphics, Visualization, and Imaging (CAe)*, pages 123–129, Anaheim, California.

[7] Forbes, A. G. and Odai, K. (2012). Iterative synaesthetic composing with multimedia signals. In *Proceedings of the International Computer Music Conference (ICMC)*, pages 573–578, Ljubjiana, Slovenia.

[8] Legrady, G. (2006). Perspectives on collaborative research and education in media arts. *Leonardo*, 39(3):215–218.

[9] Legrady, G., Bazo, D., and Pinter, M. (2013a). SwarmVision: Autonomous aesthetic multi-camera interaction. In *Proceedings of the ACM International Conference on Multimedia (MM)*, pages 737–740, Barcelona, Spain.

[10] Legrady, G., Pinter, M., and Bazo, D. (2013b). Swarm vision. *Leonardo*, 46(4):408–409.

[11] Malina, R. (2011). Alt. art-sci: We need new ways of linking arts and sciences. *Leonardo*, 44(1):2.

[12] Snavely, N., Seitz, S. M., and Szeliski, R. (2006). Photo tourism: Exploring photo collections in 3d. *ACM Transactions on Graphics (TOG)*, 25(3):835–846.