A Networked Telerobotic Observatory for Collaborative Remote Observation of Avian Activity and Range Change [*]

Siamak Faridani [*] **Bryce Lee** [**] **Selma Glasscock** [***]
John Rappole [****] **Dezhen Song** [†] **Ken Goldberg** [‡]

[*] *IEOR Department, University of California, Berkeley, California*
[**] *EECS Department, University of California, Berkeley, California*
[***] *Welder Wildlife Foundation, Sinton, Texas*
[****] *Smithsonian National Zoological Park and Roger Tory Peterson Institute, Jamestown, New York*
[†] *Department of Computer Science and Engineering Texas A&M University, College Station, Texas*
[‡] *IEOR, EECS, and School of Information, University of California, Berkeley, California*

Abstract:
The scientific field study of wildlife often requires vigilant observation of detailed animal behavior over extended periods. In remote and inhospitable locations, observation can be an arduous, expensive, and dangerous experience for field scientists. We are developing a new class of networked teleoperated robotic "observatories" that allows "citizen scientists" and professional scientists to remotely observe, record, and index animal activity and behaviors via the internet. This paper describes CONE-Welder, installed at the Rob & Bessie Welder Wildlife Foundation in Texas to gather photographic and quantitative data for a biological study of avian activity and hypothesized range change for selected subtropical bird species. Since the system was deployed on 12 May 2008, over 600 users ("players") have participated online. Players have requested over 2.2 million camera frames and captured over 29,000 photographs. Within these photos, citizen scientists have classified 74 unique species, including eight avian species previously unknown to have breeding populations within the region. The collected dataset quantifies seasonal presence of birds of particular interest, e.g., the Green Jay (*Cyanocorax incas*). This paper describes the system architecture, the game interface that provides incentives for player participation, and initial data collected. CONE-Welder is available online at: http://cone.berkeley.edu/

Keywords: Networked Robot Cameras, Citizen Science, Frame Selection, Autonomous Camera

1. INTRODUCTION

To assist field biologists, we are developing Collaborative Observatories for Natural Environments (CONEs), a new class of networked teleoperated robotic "observatories" that allows "citizen scientists" and professional scientists to remotely observe, record, and index wildlife activity via the internet. Our broader goal is to advance understanding of automated and collaborative systems that combine sensors, actuators, and human input to observe and record natural behavior in remote settings.

This paper presents the latest in a series of field installations based on a joint effort among UC Berkeley, Texas A&M, the Smithsonian Institution, and the Rob & Bessie Welder Wildlife Foundation. CONE-Welder is deployed at the Rob & Bessie Welder Wildlife Refuge, 12 km NE of

Sinton, Texas (28E6'51.1" N, 97E25'2.2" W). The region in which the refuge is located has the highest diversity of bird species in North America outside of the tropics. Welder has detailed records of its avifauna, as well as many other aspects of its ecological communities, dating back to its establishment in 1954 [Blacklock (1984), Glasscock (2007)].

The field research objectives of CONE-Welder are as follows: 1) To collect data documenting daily and seasonal presence of subtropical bird species not previously known to breed as far north as the Welder Refuge (Fig 1); and 2) To record the daily and seasonal presence of individuals of some of these species that have been banded and color-marked allowing individual recognition by photo [Rappole et al. (2007)]. This study is relevant to larger questions regarding the proximate and ultimate causes for such shifts, which may include global effects such as climate change.

At Welder, scientists including co-authors Glasscock and Rappole designed, constructed, and maintain an avian

Fig. 1. Sample screenshot of CONE-Welder from an Internet browser. The interface (using Adobe Flash) allows users to share control of a networked telerobotic video camera to capture photos, and classify the photos taken by other players. This screenshot includes a Great Kiskadee (*Pitangus sulfuratus*) and a color-marked Green Jay (*Cyanocorax yncas*), two of the species of interest in this project.

feeding station. To document the presence of the new species, they capture and band individuals of these species, and to allow individual recognition, they color-band birds from these species. Rappole and colleagues have also undertaken two years of nest searches to locate nests and document breeding. The Welder Foundation and resident scientists provide additional field and laboratory support for the project and collect information on range change in these bird species.

The CONE-Welder networked robotic camera system engages citizens around the world, including students from local and non-local schools, to systematically photograph and collect data on the daily and seasonal occurrence of subtropical birds at the site.

The project builds on our past installations that have developed new models for collaborative observation drawing on computational geometry, stochastic modeling, and optimization [Dahl (2007)].

CONE-Welder introduces several new features:

- Remote environment with extreme bird diversity
- Professionally designed feeding stations
- Lights for observation at night
- Flash interface for cross-browser compatibility
- Zone based image classification
- Multi-dimensional image and classification scoring metrics

1.1 Related Work and Previous Experience

Shortly after the introduction of the World Wide Web, Goldberg and his colleagues developed the Mercury (1994) and the Telegarden (1995-2004) networked telerobotic systems. In the Telegarden, users could plant and water seeds remotely over a nine-year period [Goldberg et al. (1995)]. These and many subsequent Internet-based telerobotic systems are surveyed in [Goldberg and Siegwart (2002)]. Goldberg and colleagues later explored multi-user control of telerobots and human "Tele-Actors" [Goldberg et al. (2003)]. For recent examples see Kim et al. (2002), Schiff et al. (2007) and Song et al. (2008b). In other work,

Kimber, Liu, Foote et al. developed a multi-user robotic camera for video conferencing [Kimber et al. (2002) and Liu et al. (2002)].

The "frame selection problem" for a shared networked telerobotic camera was defined in [Song et al. (2006)]. They study the problem of controlling a single, online, robotic camera based on simultaneous frame requests from n users. The initial algorithm based on grouping and sorting of virtual corners had time complexity $O(n^2m)$ for n users and m zoom levels. Har-Peled et al. improved this to $O(mn^{3/2}log^3n)$, and proposed a near linear ϵ-approximation algorithm. Song et al. describe the approximate and distributed algorithms for solving the frame selection problem [Song and Goldberg (2007)]. They show that with approximation bound ϵ, their algorithms run in $O\left(n/\epsilon^3\right)$ time. They also show that their algorithm can be distributed to run with time complexity $O\left(n/\epsilon^3\right)$ at each client and in $O\left(n+1/\epsilon^3\right)$ at the server. See Song (2009) for recent results.

In the context of avian observation, Chen and colleagues developed the Bird-Watching Learning system (BWL) [Chen et al. (2003)]. BWL participants must be physically present at the observation site where they use camera-equipped PDAs. BWL explores how the network can be used to classify the resulting images.

CONE-Welder's zoning and classification game was inspired by von Ahn's PeekaBoom [Von Ahn et al. (2006)] game where remote users collaborate to identify and label photos. In CONE-Welder, users are restricted to a finite set of image classification labels.

1.2 System Architecture

CONE-Welder has three major components: a robotic camera, a multi-purpose network server, and a cross-platform client interface. As shown in Fig. 2, channeling communication through an intermediate server improves
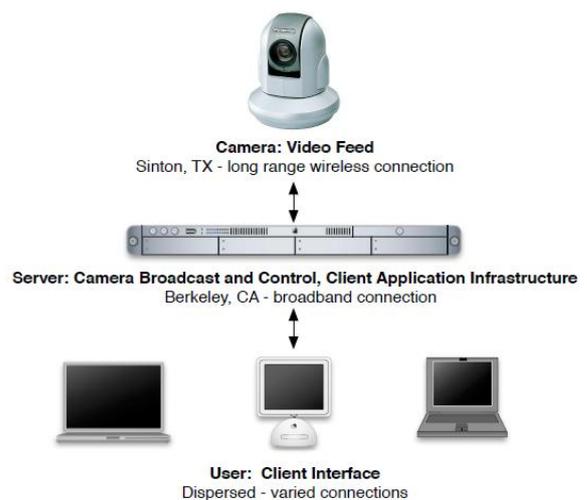

Fig. 2. Block diagram of the CONE system. The camera is communicated with using a relay server, and users receive video stream from the server. The server processes the user requests and runs the frame selection algorithm.
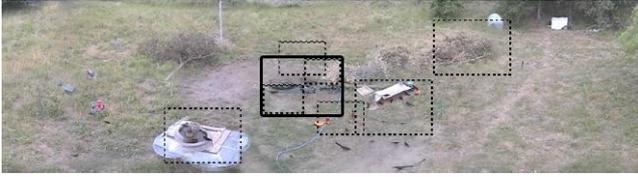
Fig. 3. Camera frame selection among multiple requests, adapted from [Dahl (2007)]. Dotted rectangles represent user frame requests r_i. The solid bordered rectangle indicates the frame f chosen by the camera.

performance and ease of maintenance. This layout also allows us to modify content and functionality at a single location, promoting thin-client development.

1.3 Collaborative Camera Control

Since the camera can only stream video from one viewpoint at a time, we developed an algorithm to satisfy multiple, simultaneous user requests. Given a set of n simultaneous frame requests $R = \{r_1...r_n\}$, we must select an optimal single frame f (Fig. 3). We formalize and then minimize user dissatisfaction.

Specifically, we are interested in reducing the time taken to fulfill a request r_i or "time-dissatisfaction". Requests that have been waiting longer get weighted more. This metric prevents starvation, where frame requests are left unfulfilled indefinitely.

The CONE frame-selection algorithm minimizes both the mean and the variance of time-dissatisfaction across all requests r_i. Mean is a measure of general satisfaction and variance is a measure of fairness. The memory-less algorithm [Song et al. (2008a)] only minimizes user dissatisfaction for a single frame. It does not take into account satisfaction over time. To solve this problem the time-dissatisfaction model is used in CONE [Dahl (2007)]. In this model users expects to view their request for some amount of time T_{SAT}. Ideally, in a single user camera, a user's request is satisfied immediately. In a shared camera, the resulting frame may only cover a portion of a requested frame. In the time-dissatisfaction model, the "coverage-time" is also taken into account and a frame that is partially covered receives some satisfaction.

We define request fulfillment as how well the current response satisfies the original request and use intersection over maximum (IOM) metric:

$$S(f, r_i) = \frac{P_i}{max\left(A_f, A_{r_i}\right)} \qquad (1)$$

where f is a candidate frame, r_i is user i's frame selection, A_{r_i} is the area enclosed by r_i, A_f is the area enclosed by f, and P_i is the intersection of A_f and A_{r_i}. The maximum value is achieved when the requested frame and selected frame are identical, resulting in a complete intersection. This maximum considers both frame location and zoom level, since each requester is expecting a level of detail with their coverage. Fulfillment also depends on coverage time. Ideally, we would provide coverage as long as desired. Since this is often not possible, we define a set duration of time that a request can receive attention before being considered fulfilled.

Schmidt and Dahl implement these metrics in the "d-weighted" frame selector, developed for CONE Sutro Forest [Dahl (2007)]. They used the time-satisfaction model as an input to the partial-frame satisfaction [Song et al. (2003)]. The goal is to maximize the total satisfaction per request weighted by time-dissatisfaction (eq. 2).

$$\sum_{r_i \in R} S(f, r_i) D_{r_i}(t) \qquad (2)$$

The dissatisfaction value ($D_{r_i}(t)$) is defined recursively:

$$D_{r_i}(t) =$$
$$\begin{cases} 1 & t = 0 \\ D_{r_i}(t-1)\left(1 - \dfrac{S(f, r_i)}{T_{SAT} - E(r_i, t)}\right) & S(f, r_i) > 0, t \neq 0 \\ D_{r_i}(t-1) + 1 & S(f, r_i) = 0, t \neq 0 \end{cases} \quad (3)$$

$E(r_i, t)$ is the cumulative coverage function. $E(r_i, t)$ is also calculated recursively:

$$E(r_i, t) = \begin{cases} 0 & t = 0 \\ E_{r_i, t-1} + S(f, r_i) & t > 0 \end{cases} \qquad (4)$$

When $E(r_i, t) = T_{SAT}$ the request is satisfied. A requests dissatisfaction will decrease during fulfillment, but incrementally increase for each frame selection where the requested frame is neglected.

2. CLIENT INTERFACE

The CONE-Welder interface shown in (Fig. 1) is designed to look like a "bird lookout". Figure 4 is an example of zone classification where users can define a bounding box (zone) around the bird in the picture, and propose the name of the species. Submitted classifications help establish zone identifications through a voting scheme. When the total number of classifications exceeds a predefined threshold (currently 3), the zone is classified based on the majority vote (if there is one). If these conditions are not met, the zone remains unclassified.

Figure 5 is a snapshot of the online classification report. It uses the timestamps of classified photographs to plot the number of daily sightings for each species. A general sighting plot is also available to help identify trends.

3. GAME MODEL FOR USERS

As an incentive for online users to gather data and generate useful information for field biologists, we designed a game where users gain points for taking and classifying photos. Leading players are recognized on leader boards. The system is designed to be self-organizing, in that scores are assigned automatically based on activity of all users. For example, there is a *Daily Parrot Award* for the most useful commentary of the day; *Cumulative Photographer Award* for the user who generates interest in another user's photo; and *Cumulative Primary Classification Award* given to the first person to classify a new zone correctly. There are currently seven different types of awards in the system. As of April 2009 there were over 97,000 awards with total value of over 125,000 points. A diagram of the total daily values of awards is shown in (Fig. 6).

3.1 Rating-based Awards

User interest of a photograph can be related to the presence of novel events that are valuable to researchers. We attempt to capture this interest through the "star" rating system (Fig. 4). Aggregate calculations over the data can help identify these interesting pictures. We reward the photographers of these pictures with the daily "Eagle Eye" award. To determine the winner, we consider all photographs taken on the target day. Valid candidates within this set must also have at least one rating from the target day. Each candidate photograph p_i is assigned a value $V(p_i)$ according to the following formula:

$$V(p_i) = \sum_{r \in R_i} \left(r - \frac{S_{max} + S_{min}}{2} \right), S_{min} \leq r \leq S_{max} \quad (5)$$

where R_i is the set of submitted ratings for p_i, S_{min} is the least possible stars given by a rating (In the current
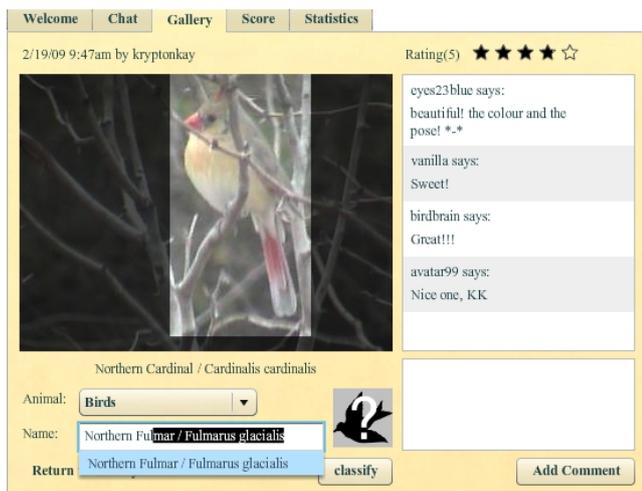
Fig. 4. The box drawn around the bird is an example of Zone Classification. On this image, the photographer and at least three other participants have classified the bird in the Zone as a Northern Cardinal. Users can also rate each photo by assigning stars to each picture (top right)
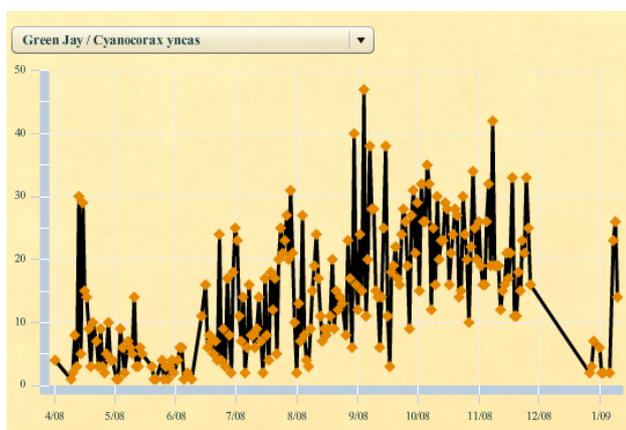
Graph of Green Jay sightings over time.

Fig. 5. Avian data visualizations on the CONE-Welder site.
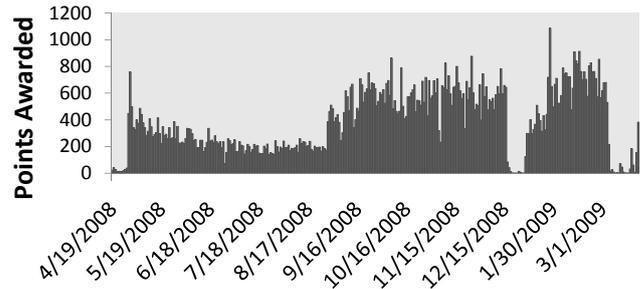
Fig. 6. Daily value of game points as of 6 April 2009. There are two maintenance periods in the diagram during which no points were allocated.

version of CONE $S_{min} = 0$), and S_{max} is the greatest possible stars given by a rating (currently $S_{max} = 5$). This formula is based on the idea that not every submitted rating is positive. It adjusts the rating system so that a mid-range rating is considered neutral. The photograph with the highest, positive, calculated value receives the award. In the event of a tie, all winners receive the award, but points are equally distributed among them. Low-rated photographs could potentially win this award when the candidate set is small. We account for differing levels of competition by rewarding points as a function of the candidate size. Each award a has a point value $P(a)$, calculated as follows:

$$P(a) = B_A \left(\frac{C_a}{D_A} \right), a \in A \quad (6)$$

where A is the award type of a, B_A is the base point value for A, C_a is size of the candidate photograph set for a, and D_A is a scaling factor for A. A maximum and minimum value is also imposed to restrict the possible point range.

3.2 Time-based Awards

To encourage users to use the system throughout a 24-hour period, we initiated the "Early Bird" and "Night Owl" awards. We consider these conditions in our scoring model by comparing the winning photograph to the runner up. The awarded point value is determined by an exponential decay function. The function awards points based on how close the winning photograph is time-wise to the runner up photograph. We designed this function to have a maximum value of 10 points and a half-life of one hour. It helps prevent users from targeting set times by awarding a negligible amount of points for solitary camera operation.

Table 1. Summary of statistics as of 6 April 2009

| Case	Amount
Frames requested by users	2,294,535
Frames selected by the system	2,014,196
Subselections	33,110
Comments	15,609
Ratings	15,362
Awards distributed	97,326
Total value of awards	125,375
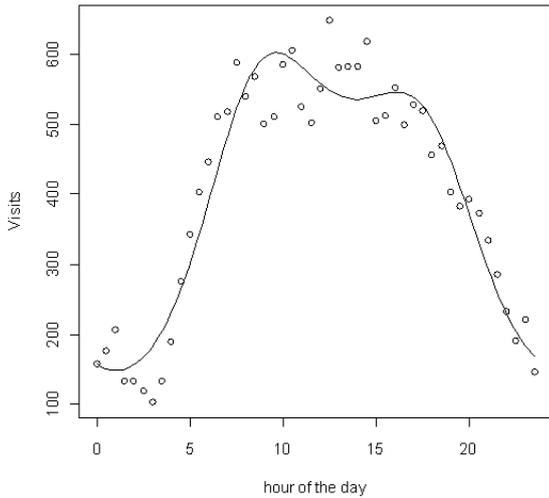
Fig. 7. Poisson regression on the number of visits per hour of the day. Midnight is represented as time t=0.

4. DATA

Table 1 summarizes system usage. These data are summarized in Fig. 7 where the number of log-ins to the system is modeled as a non-homogenous Poisson process.

Over 460 users have logged in to the system between April 18, 2008 to April 6, 2009. Of these, 256 users have contributed to classification and zoning. CONE has a highly dedicated community of active users. The 30 most active users account for 120,838 score points, 96.4% of the total. A histogram of the number of snapshots is shown in Fig. 8

4.1 Image Classification

Image classifications are useful to researchers to help document new species and track previously seen species. Participants defined zones, each a classification opportunity, on 93 percent of all photographs. Among these zoned photographs, 73 percent had at least one zone with a consensus classification. Furthermore, consensus classifications were established with an average of 4.5 votes. Users have identified a total of 74 unique species to date, shown in Fig. 9. These results confirm that Welder Refuge is extraordinarily diverse, and also confirm the the presence of eight species whose range did not extend to the Welder refuge 30 years ago [Rappole et al. (2007)].
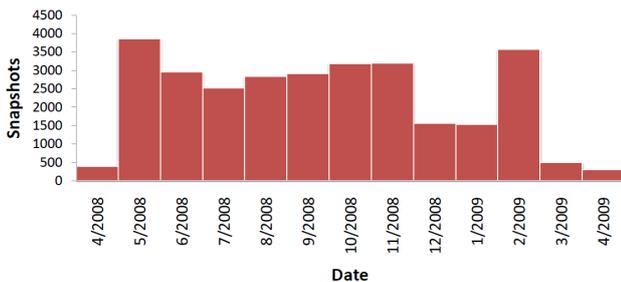

Fig. 8. Histogram of the number of snapshots taken by users.

The project produced a collection of bird images with more than 29,000 individual photos. This collection can serve as a training set for bird detection algorithms. The data set is available at (http://cone.berkeley.edu/).

4.2 Avian Range Change Data

There are fifteen Subtropical or Balconian (central Texas) species that now occur at Welder during the breeding period that were not there as breeders 30 years ago [Oberholser et al. (1974), Blacklock (1984), Rappole and Blacklock (1985)]. We have documented the presence of eight of these species by photos taken at the CONE-Welder site (Table 2). Photographs of a newly-fledged Green Jay and a juvenile Bronzed Cowbird being fed by a Northern Cardinal confirm breeding by those species. A juvenile Eastern Bluebird was photographed in July 2008. In addition, we obtained photos of color-banded Green Jays from every month of the year, demonstrating year-round residency for this species at Welder.

5. CONCLUSION AND FUTURE WORK

Our three-year NSF project is drawing to a close; this paper describes our latest Collaborative Observatory for a Natural Environment to date. CONE-Welder is installed at the Welder Wildlife Foundation in Texas to gather photographic and quantitative data for a biological study of avian distribution and activity. We described the system architecture, the game interface that provides incentives for player participation and initial data collected.
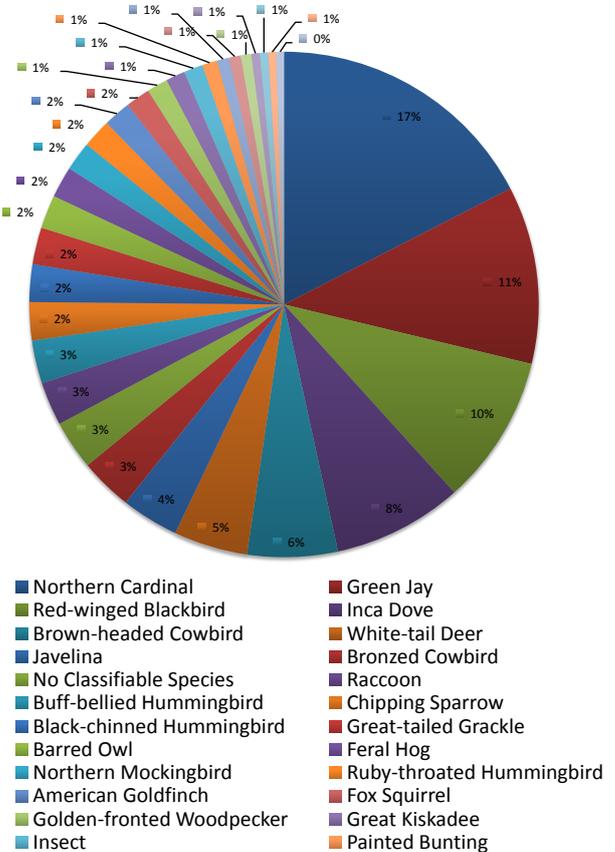

Fig. 9. Species classification totals: 28 April 2008 - 6 April 2009

Table 2. Subtropical or Balconian (central Texas) species that now occur at Welder during the breeding period that were not there as breeders 30 years ago

Species	Photos
Green Jay (*Cyanocorax yncas*)	3659
Bronzed Cowbird (*Molothrus aeneus*)	1710
Buff-bellied Hummingbird (*Amazilia yucatanensis*)	1671
Black-chinned Hummingbird (*Archilochus alexandri*)	768
Great Kiskadee (*Pitangus sulphuratus*)	516
Eastern Bluebird (*Sialia sialis*)	144
Audubon's Oriole (*Icterus graduacauda*)	28
Couch's Kingbird (*Tyrannus couchii*)	12

In future work, to collect data when no users are online and to provide an ongoing stream of interesting video for passive viewers, we are developing an 'autonomous agent" to control the camera. It will be based on statistical learning of patterns and desired frames from the first year of user/frame/timing data that we collected. We are working on software to automatically detect if a bird is visible in the video frame using machine vision. We are interested in determining if it is possible to automate the "search" for birds in the camera's workspace.

We hope CONE-Welder can be a useful model for observatories in other locations. The system has been remarkably robust, remaining online 24 hours a day for approximately one year so far, with short periods of downtime due to network interruptions and maintenance. It has attracted over 600 "citizen scientists", with over 30 dedicated regular (daily) players, who have collectively identified over 70 unique species, including eight unexpected bird species whose breeding range previously did not include the Welder Refuge. A large corpus of photographic, taxonomic, and timing data have been collected. In many photos, colored leg bands can be clearly distinguished. This set also provides valuable machine learning information for researchers in the field of computer vision. We look forward to analyzing these data, and will report results in a future publication.

ACKNOWLEDGEMENTS

REFERENCES

Blacklock, G. (1984). *Checklist of Birds of the Welder Wildlife Refuge*. Welder Wildlife Foundation, Sinton, Texas.

Chen, Y., Kao, T., and Sheu, J. (2003). A mobile learning system for scaffolding bird watching learning. *Journal of Computer Assisted Learning*, 19, 347–359.

Dahl, A. (2007). Implementation of a collaborative observatory for natural environments. Technical Report UCB/EECS-2007-71, EECS Department, University of California, Berkeley.

Glasscock, S. (2007). *Birds of the Welder Wildlife Refuge, San Patricio County, Texas*. Rob & Bessie Welder Wildlife Foundation, Sinton, Texas.

Goldberg, K., Mascha, M., Gentner, S., Rothenberg, N., Sutter, C., and Wiegley, J. (1995). Desktop teleoperation via the world wide web. In *1995 IEEE International Conference on Robotics and Automation, 1995. Proceedings.*, volume 1.

Goldberg, K. and Siegwart, R. (2002). *Beyond webcams: an introduction to online robots*. MIT Press.

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

Kim, J., Choi, B., Park, S., Kim, K., and Ko, S. (2002). Remote control system using real-time mpeg-4 streaming technology for mobile robot. In *Consumer Electronics, 2002. ICCE. 2002 Digest of Technical Papers. International Conference on*, 200–201.

Kimber, D., Liu, Q., Foote, J., and Wilcox, L. (2002). Capturing and presenting shared multi-resolution video. In *SPIE ITCOM 2002. Proceeding of SPIE, Boston*, volume 4862, 261–271.

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

Oberholser, H., Fuertes, L., Kincaid, E., Winckler, S., and Rowlett, J. (1974). *The bird life of Texas*. University of Texas Press.

Rappole, J.H., Blacklock, G.W., and Norwine, J. (2007). Apparent rapid range change in south Texas birds: response to climate change? *In The Changing Climate of South Texas, 1900-2100: Problems and Prospects, Impacts and Implications. (J. Norwine, and K. John, Eds.)*, 133–143.

Rappole, J. and Blacklock, G. (1985). *Birds of the Texas Coastal Bend: abundance and distribution*. Texas A&M University Press.

Schiff, J., Meingast, M., Mulligan, D., Sastry, S., and Goldberg, K. (2007). Respectful cameras: Detecting visual markers in real-time to address privacy concerns. In *International Conference on Intelligent Robots and Systems (IROS)*, 971–978.

Song, D. and Goldberg, K. (2007). Approximate algorithms for a collaboratively controlled robotic camera. *IEEE Transactions on Robotics*, 23(5), 1061–1070.

Song, D., Pashkevich, A., and Goldberg, K. (2003). Sharecam part II: Approximate and distributed algorithms for a collaboratively controlled robotic webcam. In *2003 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2003.(IROS 2003). Proceedings*, volume 2.

Song, D., Qin, N., and Goldberg, K. (2008a). Systems, control models, and codec for collaborative observation of remote environments with an autonomous networked robotic camera. *Autonomous Robots*, 24(4), 435–449.

Song, D., van der Stappen, A., and Goldberg, K. (2006). Exact algorithms for single frame selection on multiaxis Satellites. *IEEE Transactions on Automation Science and Engineering*, 3(1), 16–28.

Song, D. (2009). *Sharing a Vision: Systems and Algorithms for Collaboratively-Teleoperated Robotic Cameras (Springer Tracts in Advanced Robotics)*. Springer.

Song, D., Qin, N., Xu, Y., Kim, C.Y., Luneau, D., and Goldberg, K. (2008b). System and algorithms for an autonomous observatory assisting the search for the Ivory-Billed Woodpecker. In *IEEE International Conference on Automation Science and Engineering (CASE), Washington DC, August, 2008*.

Von Ahn, L., Liu, R., and Blum, M. (2006). Peekaboom: a game for locating objects in images. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, 55–64. ACM New York, NY, USA.