System and Algorithms for an Autonomous Observatory Assisting the Search for the Ivory-Billed Woodpecker

Dezhen Song, Ni Qin, Yiliang Xu, Chang Young Kim, David Luneau, and Ken Goldberg

Abstract— Ornithologists, conservationists, and millions of birdwatchers consider the Ivory-Billed Woodpecker (IBWO) the 'Holy Grail of Birds.' There have been hundreds of reports of sightings of this magnificent bird but no conclusive photo has been recorded since 1948. Under our broader research effort to develop networked autonomous "observatories" for natural environments, we have been working with the Cornell University researchers who have led a massive search initiated by eyewitness reports and low resolution video captured in the Cache River National Wildlife Refugee of Arkansas. In this paper we describe the two-camera autonomous observatory system we designed and installed in Arkansas that continuously scans the sky, recording high resolution video of candidate birds. We report on hardware, software, and algorithms based on 15 months of field experiments. Our image processing algorithm combines size filtering, nonparametric motion filtering, and temporal difference filtering to detect flying birds. Initial results suggest that we have met our four design goals: sensitivity, data reduction, accuracy, and robustness. Video segments with several bird species have been conclusively identified, video data has been consistently reduced by 99.9953%, the system is able to capture birds flying at a speed of up to 60km per hour with low false negative rates, and the system has held up to harsh field conditions. We plan for the system to be active for at least another year. As of submission, our system has recorded video of several Pileated Woodpeckers but none yet of the IBWO. For latest updates and samples of video, please see http://www.c-o-n-e.org/acone/.

I. INTRODUCTION

The Ivory-Billed Woodpecker (IBWO) is a magnificent creature that is of great interest to birdwatchers, ornithologists, and conservationists. The last confirmed U.S. sighting was in the early 1940s but a photo was taken in Cuba in 1948. In Feb. 2004, a credible eyewitness sighting was reported along Bayou DeView in eastern Arkansas, prompting a comprehensive and systematic search led by researchers at Cornell University and the Nature Conservancy. In Fall 2005, we joined the search effort by developing a high resolution robotic video system to observe the sky over an extended time period. Detailed high resolution video images are required to distinguish an IBWO from its cousin, the common Pileated Woodpecker. Our goal is to develop a robust autonomous system that detects when birds fly into

Fig. 1. Our autonomous observatory system installed along Bayou DeView, a bottomland forest near Brinkley, Arkansas. (a) The installation site. (b) A high resolution video frame of a red-tailed hawk captured by the system on Dec. 13, 2006. The red-tailed hawk has a body length of 22 inches and a wing span of 56 inches, close in size to the IBWO.

the field of view, keeping only the associated video segments. As illustrated in Fig. 1, the system has been installed in a clearing in the Bayou DeView. This project is part of our larger effort to develop autonomous and networked systems for collaborative observation of natural environments [1].

We began with the following four design goals:

a) Sensitivity: the ability to detect and record video sequences of sufficiently high resolution to clearly distinguish between the IBWO, the Pileated Woodpecker, and other species,

b) Data reduction: the system records 198GB of high resolution video data per day. Due to greatly reduced networking bandwidth in the wilderness, we want to discard at least 99% of this while maintaining criterion a),

c) Accuracy: the system should maintain a low false negative rate, which means the system should not miss an IBWO flying by the camera. However, it is acceptable if the system has a relatively high false positive rate as long as criterion (b) is satisfied, and

d) Robustness: the ability to operate autonomously in harsh conditions over long periods (i.e. mean time between maintenance > 2 years.)

We report our system and algorithm development progress including hardware design, software architecture, and a bird filter that combines size filtering, nonparametric motion filtering, and temporal difference filtering. Our system has been deployed in two locations: Texas A&M campus from May - Aug 2006 and Bayou DeView, a swampy bottomland forest near Brinkley, Arkansas from Oct 2006 until the present. Initial results suggest that our system has met the four design criteria. Fig. 1 shows the system as deployed in Arkansas and a captured high resolution image of a red-tailed hawk.

II. RELATED WORK

The IBWO is the third-largest woodpecker in the world. It has a distinctive ivory colored bill, white feathers under a black wing, and male birds have a red crest. A pair of birds may need 25km^2 or more of forest to feed. The loss of habitats due to the increasing human population and logging activities has greatly impacted the IBWO population in the past century. The last confirmed U.S. photos of IBWOs were taken by James Tanner in Louisiana in 1938. John Dennis took the last photos of this species in Cuba in April 1948.

Despite lack of conclusive evidence, the search for the legendary bird has never ceased. In 2005, the Cornell Laboratory of Ornithology and their colleagues reported the discovery of an IBWO in the Big Woods area of Arkansas [2] based primarily on a low-resolution video segment [3], so there is great interest in a high-resolution autonomous system.

Remote nature camera systems have been around since 1950s, Gysel and Davis [4] built an early video camera based on remote wildlife observation system to study rodents. Biologists use remote photography systems to observe nest predation, feeding behavior, species presence, and population parameters [5]–[10]. Commercial remote camera systems such as Trialmaster [5] and DeerCam have been developed since 1986 and have been widely used in wildlife observation. The Internet enables webcam systems that allow the general public to access remote nature cameras. Thousands of webcams have been installed around the world, for example, to observe elephants [11], tigers [12], bugs [13], birds/squirrels [14] [15], cranes [16], and swans [17]. Many other examples can be found at [18]. However, most of cameras perform simple time sampled recordings, and it is difficult or impossible for human experts to reliably review the tens of thousands of images recorded.

Song and Goldberg have developed systems and algorithms for networked cameras for a variety of applications such as construction monitoring [19], distance learning [20], and panorama construction [21].

Motion detection segments the moving objects from a video sequence. Existing motion detection techniques can be classified into three categories: background subtraction [22], [23], temporal differencing [24], [25], and optical flow [26], [27]. Background subtraction calculates the pixel-wise intensity difference between an input frame with a background reference model. To address the background noise, researchers propose many statistics-based background models such as temporal average [28], median absolute deviation (MAD) [29], adaptive Gaussian estimation [30], mixed Gaussian model, parameter estimation [31], non-parameter estimation [23], and Kalman filter compensation [32]. Temporal differencing calculates the pixel-wise intensity difference between two or three consecutive frames. Optical flow calculates the displacement flow vectors from a video sequence. A nature environment is noisy and unstructured. No single methodology can directly satisfy the four criteria in the IBWO search. During our system and software development, we carefully fine-tune the parameters to combine the strengths of nonparametric estimation, temporal differencing, and connectivity checking.
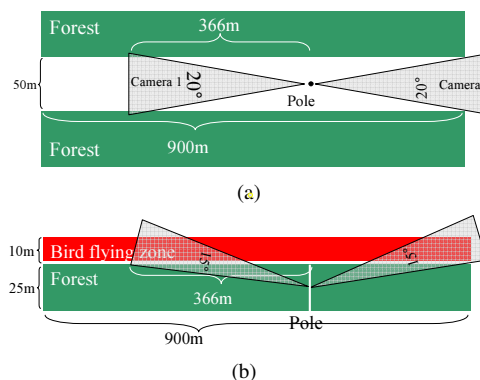
III. HARDWARE

Fig. 2. Schematic of the system installation site and camera coverage. The camera has a 20° horizontal field of view and a 15° vertical field of view. (a) top view of system coverage; (b) side view of system coverage.

Our system design was based on input from the Cornell ornithologists and the conditions of the installation site. As illustrated in Figs. 1 and 2, the system is installed in a clearing in the swampy forest that is flooded by Bayou DeView in Arkansas. The clearing is a narrow corridor that is about 900 meters long and 50 meters wide. It was formed when the forest was cut to allow a high voltage line to run through it. The system is installed on an electric pole in this power line cut. A bird flying across the power line cut is clearly exposed to the sky, which makes this an ideal location for installing the system. The site was carefully selected by the Cornell ornithologists.

To provide good coverage of the region, we chose a two-camera system design with each camera facing upward in opposite directions along the corridor. We chose a camera lens with a 20° horizontal field of view and a 15° vertical field of view. Knowing that the bird often flies at tree-top height, which is about 10 meters above the tree, we setup the camera orientation to maximize coverage as illustrated.

The Cornell ornithologists advised us that to serve as conclusive evidence, a bird image should be at least 25×25 pixels. We chose Arecont Vision 3100 3Mega-pixel high resolution networked video cameras as the imaging device. As illustrated in Fig. 3, other major components of the system include a MiniITX computer with 1.4 GHz CPU and 1GB RAM, a LinkSys wireless access point, an AW900 long range wireless adaptor with a 900Mhz directional Yagi antenna, an external timer, an external USB hard disk, and a digital I/O box with a set of relays and an LED array. To deal with the harsh swampy environment, the whole system is protected by weatherproof and thermal-controlled enclosures.

There are two separate networks in the system. The internal network is managed by the LinkSys access point that is both a wireless router and a four-port wired switch that allows the MiniITX computer to talk to the two cameras
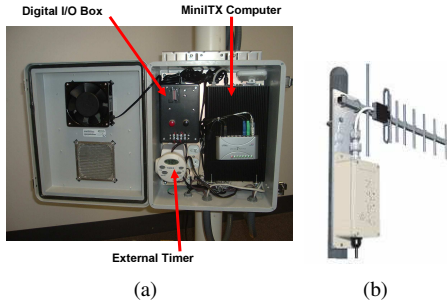
Fig. 3. System hardware configuration: (a) the MiniITX computer, the external timer, and the digital I/O box are protected in a weatherproof box; (b) the AW 900 long range outdoor wireless adaptor and a 15dBi AW15 Yagi Antenna.

via the T3 local ethernet. The local 2.4Ghz wireless service is used to facilitate *in-situ* system debugging. The external network bridges the computer to the Internet by the AW900 long range wireless adaptor. Running at 1.5Mbps and 900Mhz carrier frequency, the AW900 long range wireless adaptor can reach a maximum distance of 40 miles if equipped with a 15dBi Yagi directional antenna. Since there is no interesting activity at night, the external timer powers off the system each night. The external timer provides additional recoverability when the computer accidently crashes. Image data is stored in an external USB hard disk. Prior to the installation of the long range wireless network, one of us swapped the external hard disk every few weeks.

The customized digital I/O box has an LED array that displays the percentage of storage space left in the USB hard disk. The digital I/O box also controls a set of digital relays which can selectively power on or off individual cameras. This proves to be an important design choice because the camera firmware can crash and needs to be power-cycled from time to time. The digital I/O box is also equipped with a red push button that can power off the MiniITX, which has no either a keyboard or monitor. The simplified hardware interface makes it easy for non-experts to operate and maintain.

IV. SOFTWARE

To facilitate image acquisition, the MiniITX computer has a customized Microsoft Windows XP operating system. Due to the speed requirement, Microsoft Visual C++ has been chosen as the programming language in the development. As illustrated in Fig. 4, system software contains four main components: Bird Filter (BF), System Configuration Module (SCM), Service Module (SM), and Background Biometric (BB) filter. We will detail BF in the next section. The SCM is a configuration routine that allows us to adjust system parameters such as camera parameters, motion detection parameters, and on/off time on the field. The SM is a background process that monitors the whole system to detect if there is a software or hardware failure. The BB filter is still under development, it will be run offline to detect bird species automatically based on the biological information provided by the ornithologists.
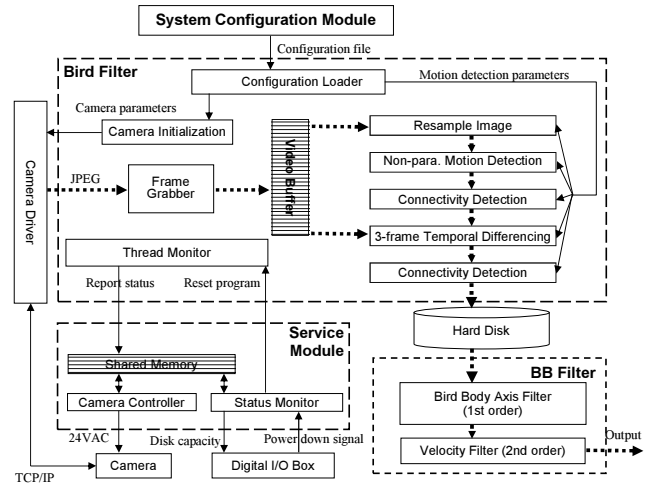

Fig. 4. System software diagram

V. BIRD FILTER

The Bird Filter (BF) makes use of information about the IBWO provided by the Cornell ornithologists:

Assumption 1: **[Known facts about the IBWO]**

1) *An adult IBWO has a body length of 48cm.*
2) *An IBWO can fly at $30 \sim 60km/hr$.*
3) *It takes a minimum size of 25×25 pixels to clearly distinguish the IBWO from the common pileated woodpecker.*

A. Input and Output

The BF is a multi-threaded process that performs filtering on the acquired image in real time. The process decides whether to keep the video on the hard disk or to delete it. The filter makes the decision by filtering out images without motion and images with noisy motions. The noisy motions include the motions caused by vibrations of tree branches, moving clouds, sun positions, water reflections, dropping tree leaves, flying insects, and any moving objects smaller than 25×25 pixels in the image. As illustrated in Fig. 4, the BF acquires frames using the frame grabber thread. The frames are stored in a video buffer. Therefore, the input to the BF are image frames captured by the cameras and the output of the BF are image frames that contain fast-moving objects that are larger than 25×25 pixels.

B. Parameters

When the BF starts, it loads the configuration parameters such as camera parameters, regions of interest, and object size to initialize each relevant module. Camera parameters refer to camera auto iris gain that enables the camera to adapt itself to different lighting conditions in the outdoor environment. The image resolution is set to 1600×1200 pixels to ensure a good balance between frame rate and resolution. At this resolution, the Arecont vision camera runs at 11 frames per second (fps). Two cameras provide a total of 22fps to the system. To ensure the imaging of a fast-moving object, the camera exposure time is set to be less than

1/100 of a second. The regions of interest refer to where we perform bird detection on the image. It is stored as a binary map that can be defined at the installation site to facilitate the quick installation of the system.

C. Spatiotemporal Downsampling

Since the two cameras combined provide 22 fps at a resolution of 1600×1200 pixels each, it is impractical and unnecessary to analyze every image in real time. Therefore, we downsample video frame sequence spatiotemporally. We partition the continuous video sequence into sequential 7-frame video sequences. Define F to be a frame, the ith video sequence defined is,

$$F_i = \{F_{i1}, F_{i2}, ..., F_{ij}, ..., F_{i7}\}. \qquad (1)$$

For each segment, we process its 4th frame F_{i4}, $i = 1, ..., \infty$, at a resolution of 400×300 with motion detection. In the downsampled image, we are interested in capturing motion objects that are bigger than 6×6 pixels, which is equivalent to the 25×25 pixels in the original size. There is a possibility that a bird might be missed due to the temporal downsampling. It takes a bird about 1 second to fly cross the power line cut, which should be sufficient time for the camera to capture 11 frames. However, there is a small chance that a bird might not appear on the 4th frame of the video sequence, and we could miss the bird completely. However, this is the natural limit imposed by the computation power and camera field of view. The downsampling operation can reduce noisy motions and increases computation speed.

D. Nonparametric Motion Filtering

To eliminate periodical noisy motions caused by vibrating tree branches and their shadows, we adopt the nonparametric background subtraction algorithm proposed by Elgammal et.al [23].

For every pixel at time t, Elgammal's algorithm updates a Gaussian model $N(\mathbf{0}, \Sigma)$ from its intensity values from the corresponding pixels in previous frames F_{i4}, $i = 1, ..., t$, where $\Sigma = \text{diag}\{\sigma_r^2, \sigma_b^2, \sigma_g^2\}$ is the variance-covariance matrix for three color channels. The Gaussian distribution updates itself as a new sample comes in. Therefore, for a periodic noise, the Gaussian model can characterize the periodic intensity change in its variance if the algorithm has enough samples. The algorithm then predicts if a pixel is a foreground pixel based on probability thresholding. After extensive tests, we set the thresholding point to be the 98th percentile.

This method has been proven to be robust in dealing with periodic noise. In our field test conducted on the Texas A&M campus, this method successfully filtered out the noisy background motions introduced by a rotating radio antenna. The output of nonparametric motions filtering is a binary map with white pixels as motion pixels, which is defined as B_{i4} for frame F_{i4}.

E. Connectivity Check

Unfortunately, the nonparametric filter cannot effectively filter out non-periodic noises such as moving clouds or dropping leaves. Further filtering is needed. We first perform a connectivity check to determine the size of the region that triggers the motion. Recall the required size in Assumption 1, we only keep the images with big moving objects. Recall that B_{i4} is a downsampled image. A size of 6×6 pixels is equivalent to the 25×25 pixels in the original image. If a B_{i4} contains a moving object that is bigger than 6×6 pixels, we proceed to the next step. Otherwise, we discard the entire segment F_i.

F. Temporal Differencing

Since a moving cloud can take on any shape or size, the downsampling and the nonparametric motion tracking cannot get rid of the false alarms triggered by moving clouds. On a cloudy day, the system might accumulate huge amounts of video data containing only moving clouds.

Observing the data, we notice that the velocity of a moving cloud is still relatively slow if compared with that of a flying bird. In adjacent frames, the displacement of a moving cloud is negligible if compared with the displacement of a flying bird. Therefore, for each motion frame F_{i4} detected by the nonparametric motion detector, we combine the motion frame with two immediate adjacent frames F_{i2} and F_{i3} to judge the velocity difference. We know that motion on frame F_{i4} is detected using the previous frames $F_{i-1,4}$, $F_{i-2,4}$, $F_{i-3,4}$, ... $F_{1,4}$. For a slow moving object such as a cloud, although there exists an intensity difference $|F_{i4} - F_{i-1,4}|$ for the motion to be detected, the intensity difference between adjacent frames $|F_{i4} - F_{i3}|$ and $|F_{i5} - F_{i4}|$ should be much smaller than those of a fast moving object. Therefore the sum of $|F_{i4} - F_{i3}|$ and $|F_{i5} - F_{i4}|$ is a good thresholding function to judge if the moving speed of the object is fast enough. In our experiment, the threshold point is 30. We name it 3-frame temporal differencing. It is capable of filtering out objects that are significantly slower than the IBWO.

G. Algorithm

To summarize the BF, we present Algorithm 1. For each frame with n pixels, the algorithm runs in time linear to n.

VI. EXPERIMENTS AND RESULTS

Two field tests have been conducted for the autonomous observation system. The system had been installed on the Texas A&M campus from May 2006 to Oct. 2006 for the initial test. After 5 months' testing and tuning, the system has been moved to Brinkley, AR to assist in the search for the IBWO since Oct. 26, 2006.

A. Sensitivity

Fig. 5 illustrates four species of birds imaged by our system in Arkansas. Among the samples, Fig. 5(a) is the closest cousin of the IBWO. Although the image is blurred, Cornell and U. Arkansas at Little Rock ornithologists were able to verify that it is a Pileated Woodpecker. A Pileated Woodpecker has a body length of 40cm, which is just slightly smaller than that of the IBWO. The Northern Flicker in Fig. 5(b) is a smaller kind of woodpecker that has a size

Algorithm 1: Bird Filter Algorithm

input : $F_i, i = 1, ..., t$
output: motion video sequences
Use F_0 to initialize the Gaussian model in the
nonparametric filter; $O(n)$
for *each F_i* **do**
    downsample F_{i4} to 400×300; $O(n)$
    for *each pixel p in frame f* **do**
        compute probability of being motion pixel using
        intensity Gaussian model; $O(n)$
        update pixel intensity Gaussian model using
        pixel p intensity value; $O(n)$
        if *pixel p is motion pixel* **then**
            update B_{i4}

    check B_{i4} for the motion object; $O(n)$
    if *find motion object larger than 6x6 pixels* **then**
        Reset B_{i4} to black; $O(1)$
        downsample F_{i3} and F_{i5} to 400×300; $O(n)$
        for *each pixel p in F_{i4}* **do**
            compute intensity difference with pixel in
            F_{i3} and F_{i5}; $O(n)$
            if *larger than threshold* **then**
                update B_{i4}
        check B_{i4} for motion object larger than 6x6
        pixel; $O(n)$
        if *find motion object* **then**
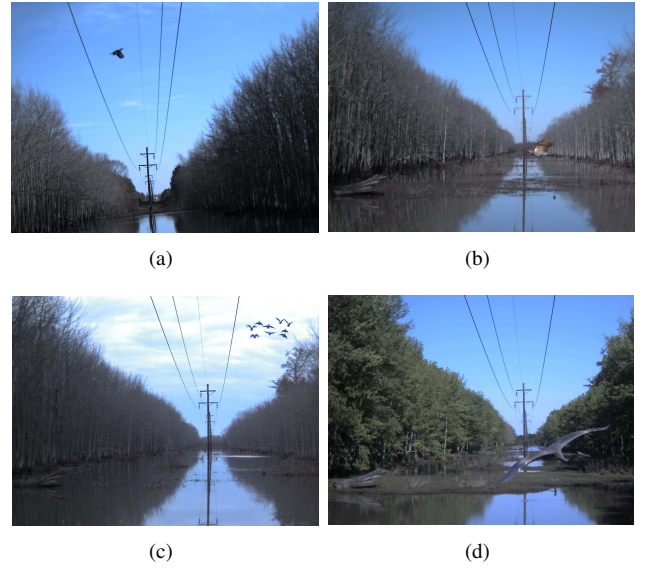 F_i is motion sequence;

Fig. 5. Sample birds imaged by the system. (a) A Pileated Woodpecker (02/16/2007). (b) A Northern Flicker Woodpecker (02/27/2007). (c) A flock of Canada Geese (10/28/2006). (d) A Great Blue Heron (04/28/2007).

of 28-31 cm and a wingspan of 42-51cm. Fig. 5(c) show a flock of Canada Geese caught by the system. Fig. 5(d) is a Great Blue Heron with a wingspan of close to 2m. Birds caught by the system can be either bigger or smaller than the IBWO and fly either faster or slower than the IBWO. This suggests that our system is capable of capturing conclusive images of an IBWO.

B. Data Reduction

As of Sep. 4, 2007, the system has collected over 25GB of images. A total of 113,836 images have been captured by the BF. Considering that there were a total of 245,520,000 images captured by the two cameras during the 310 days, the BF reduced the data by 99.9953%.

C. Accuracy

We consider both false negative and false positive rates. A false negative means that the system fails to detect when a bird flies by. Again, we tested the system using the data from both the Texas A&M campus and Brinkley, AR.

To test the false negative rate, we turn on the recording mode of the camera and sample every frame. Then we manually count the number of images containing a flying bird that is bigger than 25×25 pixels. Comparing those with the algorithm output, we then get the false negative rate. A total of 80,000 image frames were collected over a 2-hour

period on campus. There were three birds flying across the camera field of view in this 2-hour period and all have been detected by the BF. As mentioned earlier, the only reason a bird is missed by the system is the fact that it does not appear in F_{i4}, which is possible if the bird's flying trajectory is very close to the boundary of the camera field of view. The false negative test is actually the test of how many birds do not fly close to the center of the camera field of view. In the test data set, none of the birds fly close to the boundary of camera field view. We believe it could be less than perfect in the long run. Since the boundary of camera field of view is much smaller in comparison to overall field of view, the false negative rate should be a small value ($< 20\%$). We are testing the false negative rate using the data from AR and will report the result in future revisions of this paper.

The false positive rate indicates the percentage of the images stored that are not triggered by bird motions. Since we perform motion detection computations on only the 4th frame of every 7-frame video segment, we collect the statistics only on the frame in which motion detection is performed. For the 1205 captured motion image files from the Texas A&M campus over a 6-day test period, the false positive rate is 32.9%. The false positive rate is 96% for the nine months of data collected in AR. The high false positive rate in AR is expected because we are more conservative in parameter settings. For example, our probability threshold in nonparametric motion filter is 99.9% for the experiment on the Texas A&M campus and is 98% for the experiment in AR. We purposefully lower the probability threshold to increase sensitivity. Also there are large numbers of insects in the forest that can trigger false alarms when they fly close to the lens. As long as the size of the files is not too big to be transferred, this false positive rate is acceptable.

D. Robustness

After ten months in the Arkansas wilderness, the system has run continuously except for occasional power outages. The system has survived very large temperature variations from winter to summer, severe weather changes, and has worked under high humidity conditions.

VII. Conclusion and Future Work

This paper reports our system and algorithm development for an autonomous observatory to assist the search for the IBWO. Data collected thus far suggests that the system achieves four design criteria: sensitivity, data reduction, accuracy, and robustness.

In the future, we will improve filter efficiency by developing more powerful filters that combine bird specific biological information such as size, color, velocity, and flying pattern. We will also activate the long range wireless link to allow images to be automatically uploaded to an online imaging database. After post-filtering, we will invite the public to annotate the images. We are also working on systems and algorithms for distributed and networked robotic actuation that can actively modify the environment to improve observation.

References

[1] D. Song, N. Qin, and K. Goldberg, "Systems, control models, and codec for collaborative observation of remote environments with an autonomous networked robotic camera," *Autonomous Robots*, vol. 24, no. 4, pp. 435–449, May. 2008.

[2] J. W. Fitzpatrick, M. Lammertink, D. Luneau, T. W. Gallagher, B. R. Harrison, G. M. Sparling, K. V. Rosenberg, R. W. Rohrbaugh, E. C. H. Swarthout, P. H. Wrege, S. B. Swarthout, M. S. Dantzker, R. A. Charif, T. R. Barksdale, J. V. Remsen, S. D. Simon, and D. Zollner, "Ivory-billed woodpecker (campephilus principalis) persists in continental north america," *Science*, vol. 308, no. 5727, pp. 1460–1462, June 2005.

[3] F. E. Hayes and W. K. Hayes, "The great ivory-billed woodpecker debate: Perceptions of the evidence," *Birding Magazine*, vol. 39, no. 2, pp. 36–41, March/April 2007.

[4] L. W. Gysel and E. M. J. Davis, "A simple automatic photographic unit for wildlife research," *Journal of wildlife management*, vol. 20, pp. 451–453, 1956.

[5] T. E. Kucera and R. H. Barrett, "The trailmaster camera system for detecting wildlife," *Wildlife Society Bulletin*, vol. 21, pp. 505–508, 1993.

[6] K. Iida, R. Takahashi, Y. Tang, T. Mukai, and M. Sato, "Observation of marine animals using underwater acoustic camera," *Japanese Journal of Applied Physics*, vol. 45, no. 5B, pp. 4875–4881, 2006.

[7] T. L. Cutler and D. E. Swann, "Using remote photography in wildlife ecology: a review," *Wildlife Society Bulletin*, vol. 27, no. 3, pp. 571–581, 1999.

[8] M. D. Sanders and R. F. Maloney, "Causes of mortality at nests of ground-nesting birds in the upper waitaki basin, south island, new zealand: a 5-year video study," *Biological Conservation*, vol. 106, no. 2, pp. 225–236, 2002.

[9] M. M. Stake and D. A. Cimprich, "Using video to monitor predation at black-capped vireo nests," *BioOne*, vol. 105, no. 2, pp. 348–357, 2003.

[10] S. B. Lewis, P. DeSlmone, K. Titus, and M. R. Fuller, "A video surveillance system for monitoring raptor nests in a temperate rainforest environment," *Northwest Science*, vol. 78, no. 1, pp. 70–74, 2004.

[11] AfricaWebCams, "http://www.zulucam.org/," 2005.

[12] tigherhomes.org, "http://www.tigerhomes.org/ animal/ web-cams.cfm," 2005.

[13] D. E. Weber, B. Grosser, and G. Fried, "http://bugscope.beckman.uiuc.edu/," 2005.

[14] newyorkwild.org, "http://www.newyorkwild.org/ webcams/ web-cams.htm," 2005.

[15] JamesReserve, "http://www.jamesreserve.edu/webcamsphp.lasso," 2005.

[16] NationalGeographic, "http://magma.nationalgeographic.com/ ngm/ cranecam/cam.html," 2005.

[17] SwanCam, "http://www.osage.net/ mccb/trumpeterswan.html," 2005.

[18] wildcam.com, "http://www.wildcam.com/," 2005.

[19] D. Song, Q. Hu, N. Qin, and K. Goldberg, "Automating high resolution panoramic inspection and documentation of construction using a robotic camera," in *IEEE Conference on Automation Science and Engineering, Edmonton, Canada*, Aug. 2005.

[20] K. Goldberg, D. Song, and A. Levandowski, "Collaborative teleoperation using networked spatial dynamic voting," *The Proceedings of The IEEE*, vol. 91, no. 3, pp. 430–439, March 2003.

[21] N. Qin, D. Song, and K. Goldberg, "Aligning windows of live video from an imprecise pan-tilt-zoom robotic camera into a remote panoramic display," in *IEEE International Conference on Robotics and Automation (ICRA), Orlando, Florida*, May 2006.

[22] P. L. Rosin, "Thresholding for change detection," in *Sixth International Conference on Computer Vision*, Jan 1998, pp. 274–279.

[23] A. Elgammal, D. Harwood, and L. Davis, "Non-parametric model for background subtraction," in *6th European Conference on Computer Vision. Dublin, Ireland*, June 2000, pp. 751–767.

[24] C. Anderson, P. Burt, and G. van der Wal, "Change detection and tracking using pyramid transformation techniques," in *Proceeding of SPIE - Intelligent Robots and Computer Vision*, vol. 579, 1985, pp. 72–78.

[25] A. J. Lipton, H. Fujiyoshi, and R. S. Patil, "Moving target classification and tracking from real-time video," in *Proceeding of IEEE Workshop Applications of Computer Vission*, 1998, pp. 8–14.

[26] I. Cohen and G. Medioni, "Detecting and tracking moving objects for video surveillance," in *IEEE Computer Society Conference Computer Vision and Pattern Recognition in Fort Collins, CO*, vol. 2, 1999, p. 325.

[27] J. Barron, D. Fleet, and S. Beauchemin, "Performance of optical flow techniques," *International Journal of Computer Vision*, vol. 12, no. 1, pp. 42–77, 1994.

[28] N. Friedman and S. Russell, "Image segmentation in video sequences: a probabilistic approach," in *Proceedings of 13th Conference on Uncertainty in Artifical Intelligence*, 1997, pp. 1–3.

[29] Y. H. Yang and M. D. Levine, "The background primal sketch: an approach for tracking moving objects," in *Machine Vision and Applications*, vol. 5, 1992, pp. 17–34.

[30] M. Kohle, d. Merkl, and J. Kastner, "Clinical gait analysis by neural networks: Issues and experiences," in *Proceedings of IEEE Symposium on Computer-Based Medical Systems*, vol. 5, 1997, pp. 138–134.

[31] W. E. L. Grimson, C. Stauffer, R. Romano, and L. Lee, "Using adaptive tracking to classify and monitor activies in a site," in *Proceedings of IEEE Symposium on Computer-Based Medical Systems*, vol. 5, 1997, pp. 138–134.

[32] C. Ridder, O. Munkelt, and H. Kirchner, "Adaptive background estimation and foreground detection using kalman-filtering," in *Proceedings of International Conference on Recent Advances in Mechatronics*, 1995, pp. 193–199.