Proceedings of the 2003 IEEE
International Conference on Robotics & Automation
Taipei, Taiwan, September 14-19, 2003

A Kite and Teleoperated Vision System for Acquiring Aerial Images

Paul Y. Oh and Bill Green
Mechanical Engineering & Mechanics, Drexel University, Philadelphia PA
Email: paul@coe.drexel.edu and weg22@drexel.edu

Abstract

In times of disaster acquiring aerial images is challenging. Runways may be crippled thus denying conventional aircraft in the area from taking off. Also the time required to schedule a satellite fly-by may delay first response efforts. Man backpackable aerial robots can be carried close to the disaster site and flown to capture aerial images. This paper integrates mechatronics, intelligent sensing, and mechanism synthesis in a teleoperable kite-mounted camera. Rapidly deployable, transportable by foot, easy to fly and affordable, our system can quickly acquire, process and distribute aerial images. Image mosaicing, edge detection, 3D reconstruction and geo-referencing resulting from images acquired by our aerial platform are also presented.

1 Introduction

For situational awareness and disaster mitigation, aerial images are essential for first responders, site commanders and decision makers. Such images are used to identify the extent of damages, structural integrity of buildings and bridges, ingress routes to the site and egress routes to nearby hospitals. There are a number of challenges in acquiring aerial images. First, disasters cripple runways, roads and bridges thus denying access to areas by conventional aircraft and land vehicles. Two, although satellites can provide detailed images, reprogramming a fly-by takes time and delays mitigation efforts. Third, unmanned aerial vehicles can also capture images but flying such aircraft demands highly skilled teleoperators. Fourth, a remote controlled aircraft can be employed and although flying is easier it demands the pilot keeps an eye on the model. This becomes very difficult when flying at night or in urban environments. The net effect is that in situational awareness, disaster mitigation and search-and-rescue operations there is a need for an aerial image acquisition system that's rapidly deployable, easily transportable and simple to fly.

Aerial robots have the potential to acquire aerial images but today's prototypes have limitations. Lighter-than-air aircraft like blimps [9] are difficult to deploy rapidly because helium must be backpacked and inflation time is required; airship volume is typically large for a buoyancy that robustly flies and lifts its sensor and instrumentation payload. Rotary wing aerial robots [2] including helicopters [6], [4] can also acquire aerial images but often require GPS to navigate autonomously. In urban environments however buildings and weather conditions can occlude the line-of-sight to satellites and limit the robot's ability to fly autonomously. Time pressures and risks to human life demand such robots become more reliable and robust before they can be field-ready.

Kites fit in backpacks, are deployable in minutes, easy to fly and affordable. Such characteristics make a kite an attractive platform for airborne instrumentation that quickly acquires aerial images. Kites can be designed to fly according to wind speed and payload mass. For example, a 3 m wingspan kite can easily fly in a 10 MPH wind [1] while carrying a 2 kg payload. A kite is tethered and hence easy to fly even at heights greater than a 1000 feet (approximately 70 building stories) and at night. This paper presents an aerial image acquisition system design we call *LEAP*: Low Elevation Aerial Photography. *LEAP* is a kite retrofitted with a teleoperated camera rig and attitude regulator. The design integrates mechanism synthesis, mechatronics, computer vision and wireless networking to acquire, process and distribute aerial images. *LEAP* was awarded the Philadelphia Port of Technology *Entrepreneurship in Technology* award for its potential in rapid disaster response. *LEAP* incorporates *partitioning*, a design philosophy that leverages dynamics when synthesizing a vision system [3] [5]. Section 2 presents both the kite's flight dynamics and the camera rig's underlying physics which enable camera orientation to be regulated despite motions of

[1] At 10 MPH winds, leaves are in motion and lightweight flags extend

the kite. Section 3 presents images that can be typically acquired by our system. Work with such images includes image mosaicing and $3D$ reconstruction. Section 4 concludes with our future disaster mitigation endeavors.

2 *LEAP* Design

Like Benjamin Franklin, many people have lifted payloads off the ground with a kite. However there are a number of challenges that cannot be overcome *ad hoc*. First, a kite is airborne when wind pressure, rope and tail tensions and gravity are in dynamic balance. The kite wingspan, center of pressure, bridle point and tail length must be designed in concert to handle expected wind speeds and lift the desired payload mass. Second, aerial image acquisition requires a means to remotely both adjust camera orientation and monitor field-of-view and resolution from the ground. Without such means, capturing desired images is hapchance. Lastly, as the kite responds to changes in wind speed and direction, a rig suspended off the kite will swing. This introduces torques that can destabilize dynamic balance and cause the kite to crash. More importantly, controlling camera orientation is nearly impossible when the rig swings. An analytic approach was used to integrate kite flight dynamics, camera mechatronics and the rig mechanism's kinematic synthesis to overcome these challenges.

2.1 Kite Flight Dynamics

Analyzing the forces acting on a kite can yield wingspan dimensions needed to airlift a desired payload mass. *Ad hoc* oversizing of a wingspan or trial-and-error flights is not attractive; larger wingspan kites demand stronger rope and are harder to keep under control in high winds. Essentially, kites remain airborne and stationary when kite weight w, rope b and tail t tensions and wind force p are in balance as shown in Figure 1. The direction lines of force meet at one common point C called the concurrency point. Ground anchor A, bridle B and tail T are fixed points and cannot change. The center of pressure P and mass center M are points within or near the kite. For angles of attack α between 15 to 40 degrees relative to the airstream, the location of P will not change much and can be assumed to be fixed.

Wind force is proportional to the square of wind speed. Assuming constant kite weight and an unstretchable rope, changes in wind speed will result in a force imbalance thus prompting kite and/or tail

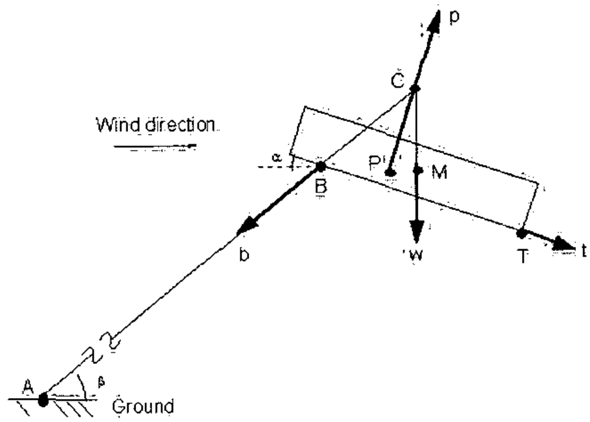

Figure 1: Simple kite model with forces in balance

movement. For most kites, the location of the bridle point B is often a small distance away from the kite's sail and hence movement about B will be small. With tail force typically being small, the only remaining movement possible is about A. Aerodynamically, to compensate for force increase arising from higher wind speed, the kite must decrease angle of attack. This is a counter-clockwise arching in Figure 1 and is the marvel of kite flight dynamics; in this non-linear dynamic balancing act, the kite flies into new states of stability.

Kite wing span needed to successfully airlift a payload in expected wind speed can be calculated. The underlying physics can be appreciated by assuming a kite sail that is square with sidelength L. The result is a wind force acting on an effective area of $A = L^2$.

Wind force at the center of pressure P (see Figure 1) is proportional to the kite's effective area and the square of wind speed V^2. The balance of forces p, w and b discussed above dictates that $w \approx AV^2$ or

$$V^2 \approx \frac{w}{A} \tag{1}$$

In other words, wind speed squared is proportional to kite/payload weight divided by the kite sail area. Independent of wind speed is buoyancy which dictates a constant mass ratio μ (mass of air displaced versus mass of kite). Kite mass is proportional to its weight w. The displaced air mass, being a volume, must be proportional to another volume, namely L^3. This yields

$$\mu \approx \frac{L^3}{w} \tag{2}$$

In other words with a constant mass ratio, upscaling a stable kite to a sidelength $L' = XL$ where $X > 1$ will result in a new effective area $A' = X^2 A$ and from Equation 2, the new kite weight is $w' = X^3 w$. From Equation 1, the new wind speed required to be airborne is $V' = VX^{1/2}$. Such upscaling results in weight growing with volume, loss of stability at higher wind speeds and more wind is needed to remain airborne.

Alternatively, changing mass ratio, a heavier kite for instance, can increase stability. For a specific wind speed V, an upscaling with $L' = XL$ where $X > 1$ will increase effective area $A' = X^2 A$ and from Equation 1 results in $\frac{w'}{A'} = \frac{w}{A}$. Thus kite weight grows with area and yields $\mu' = X\mu$.

The net effect is that given a kite that flies stably at a specific wind speed or defined mass ratio, the necessary changes in wing span can be calculated.

2.2 Camera Rig Mechatronics

A rig, Figure 2, to be suspended off the kite's rope, was designed with two RC (radio-controlled) servos that pan and tilt a lightweight 2.4 GHz camera. With respect to the rig, the servos permit control over camera orientation from the ground. Live video is wirelessly transmitted to a ground-based receiver. The receiver is connected to a portable camcorder which allows the camera's field-of-view to be monitored live and recorded from the ground.

Using the camcorder's IEEE1394 firewire interface, video can be digitized out in the field with a laptop computer. Equipped with a 802.11b wireless network card, base stations in the field permit the laptop to web stream video and/or upload image stills to remotely located web servers. The net effect is a portable system that gives users with Internet access, an "eye" in the sky. *LEAP* can rapidly acquire, process and distribute aerial images.

2.3 Picavet

The rig's RC servos permit fixating the kite-mounted camera. When the rig is tied to the kite's rope, it will sway when the kite moves in the wind. This swaying makes radio-controlling camera orientation to fixate on ground subjects very difficult. Swaying also introduces a torque which can destabilize the kite and cause it to crash. A mechanism based on an elliptical pendulum, known as a Picavet linkage, can be kinematically synthesized to keep the camera rig attitude constant despite changes in kite orientation.

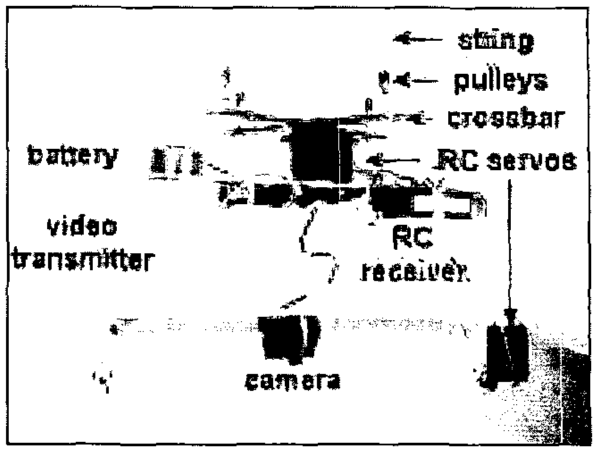

Figure 2: Mechatronic camera rig designed with RC servos, 1/4 mile range 4 mm lens focal length video transmitter and battery pack. Pan and tilt range are both $\pm 30^o$

The net effect is that the camera's image plane can be stabilized, and therefore controlled. The Picavet linkage, Figure 3, consists of a crossbar utilizing four pulleys on each end as attachment points, one continuous rope, two brackets that are fixed to the kite line and a ring used to constrain the two innermost lines as they cross.

As the kite increases/decreases its angle of attack, the rope glides effortlessly through the pulleys on the cross, keeping the rig undisturbed. Figure 3 is used to simulate a change in kite orientation. The left photo shows the Picavet's initial orientation with the rig's attitude being parallel to the ground. The right photo depicts that while there is a change in the kite's orientation, the rig maintains its initial position. Furthermore, if the linkage is initially positioned such that the cross is at some angle relative to the ground – it would also maintain its original orientation.

Figure 4 depicts a small body force diagram. All tensions are the same because a single rope loops through all pulleys at the crossbar tie-points. The x-axis goes through the midpoint of the bar and is perpendicular to it. This demands that ϕ_{1a}, ϕ_{1b}, θ_{1a} and θ_{1b} are all less than 90^o. Furthermore

$$\tan\phi_{1a} = \tan\phi_{1b} = \frac{r}{l_{\text{bar}}/2} \quad (3)$$

So $\tan\phi_{1a} = \tan\phi_{1b}$ and because ϕ_{1a} and ϕ_{1b} are less

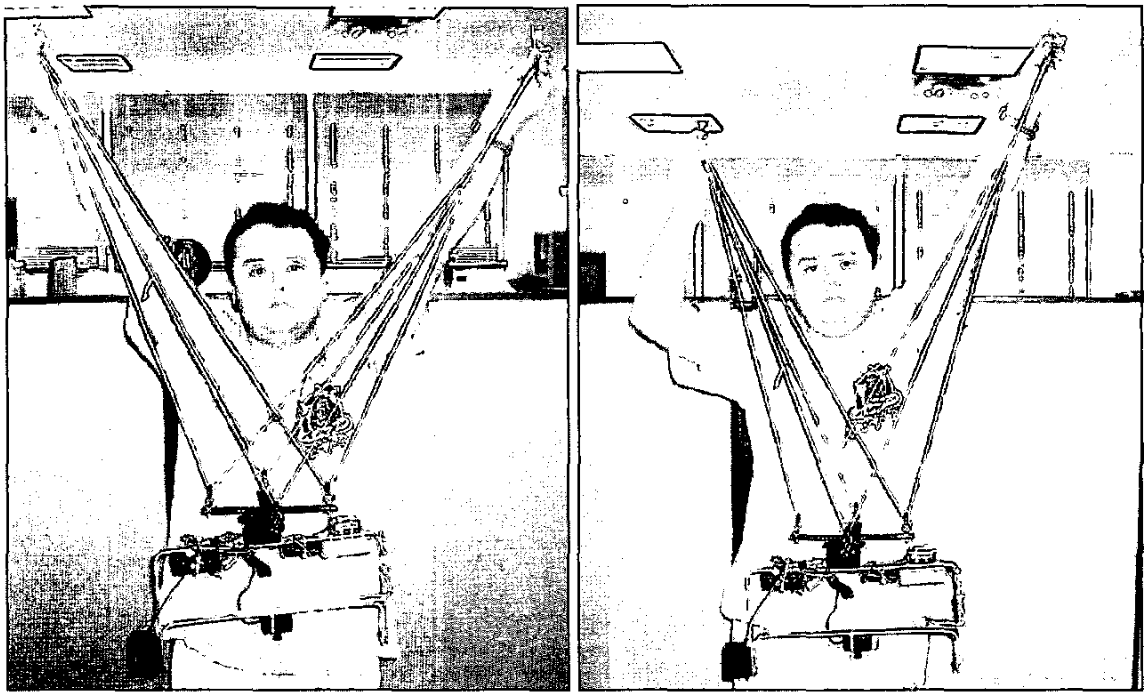

Figure 3: Left: Picavet initial attitude. Right: Attitute is unchanged despite pendulum sway

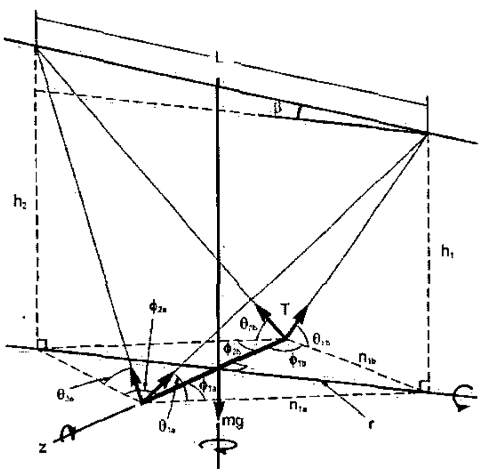

Figure 4: Small body force diagram for a Picavet

than 90^o we have $\phi_{1a} = \phi_{1b}$. Thus $n_{1a} = n_{1b}$ and

$$\tan\theta_{1a} = \frac{h_1}{n_{1a}} \quad \text{and} \quad \tan\theta_{1b} = \frac{h_1}{n_{1b}} \tag{4}$$

Hence $\tan\theta_{1a} = \tan\theta_{1b}$ and since θ_{1a} and θ_{1b} are both less than 90^o, we have $\theta_{1a} = \theta_{1b}$. Similarly

$$\phi_{2a} = \phi_{2b} \tag{5}$$

$$\theta_{2a} = \theta_{2b} \tag{6}$$

Proving the sum of the moments about the center of mass of the bar are equal to zero will prove that it remains in its initial position. The moments about the y-axis can be disregarded because rotation about the y-axis will not change the bar's orientation relative to the ground. $\sum M_x = 0$ can be shown. Assuming counterclockwise as the positive direction yields

$$\sum M_x = \begin{array}{c} -\frac{1}{2}Tl_{\text{bar}}\sin\theta_{1a} - \frac{1}{2}Tl_{\text{bar}}\sin\theta_{2a} \\ +\frac{1}{2}Tl_{\text{bar}}\sin\theta_{1b} + \frac{1}{2}Tl_{\text{bar}}\sin\theta_{2b} \end{array} \tag{7}$$

Assuming that the rope is attached at the centerline of the bar, the perpendicular distance between the bar and any resulting horizontal tension forces will be zero. This makes the moment about the z-axis zero so $\sum M_z = 0$

3 Data

Raw aerial images can appear extremely convoluted to the untrained eye. Therefore, software is needed to process these images in order to facilitate the interpretation of data for site commanders and decision makers. A single image still has a very limited field of view, making it impossible to comprehend what is happening in the surrounding areas. However, the field-of-view can be significantly increased by stitching together several image stills to form a mosaic. Furthermore, any image taken from the air will encompass a large amount of useless information. Edge detection is a fundamental technique that will filter the image while still preserving its structural properties. Additionally, software can be used to construct a 3D model of the scene, making it possible to navigate through streets and around buildings. Finally, geo-location referencing will give an accurate approximation of a scale in an aerial photograph.

Image Mosaicing is the process of stitching several images together to yield a single larger image. Because a stationary camera typically has a field-of-view of only 50^o, a moving camera can be used to capture more slices of an area. These multiple image slices can be mosaiced together to give an entire view of a scene. For example, in Figure 5, the surrounding three images were acquired by the aforementioned mechatronic camera rig, and used to generate the resulting mosaic (center). Looking at a single mosaiced image rather than six or seven separate images showing equivalent information will be time efficient. The method of mosaicing [7] identifies and relates common points among two or more images. Any two common points are related to each other by a translation and a rotation. Four common points, two in each image, are needed to generate the transformation matrix, M. Once, the transformation matrix is known, the points in the original images, u, can be mapped algorithmically to the mosaiced image, u', by the relation $u' = Mu$.

3D Reconstruction: is a technique that generates 3D models from 2D images [8]. Figure 6 (bottom) shows the *3D* reconstructed model of the image still (top) using a commercial version of Facade [1] called Canoma. Once created, the user can rotate, pan, zoom the *3D* model via a VRML-enabled web browser. [2] This creates a very easy-to-use interface that will facilitate collecting and analyzing data.

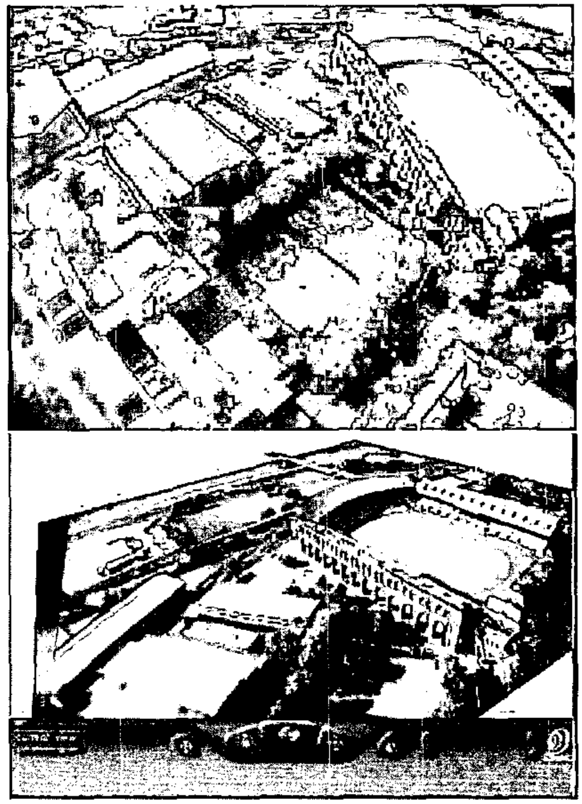

Figure 6: Aerial photo of urban ground scene at 1000 feet (top) and 3D reconstructed model (bottom).

4 Conclusion

In time critical situations there are challenges to acquiring aerial images with conventional aircraft, satellites, unmanned aerial vehicles and radio-controlled or robotic blimps, airplanes and helicopters. Challenges like crippled runways, damaged roads, delays, transportability, deployment delays, user difficulty and expense must be all considered when designing a system to augment mitigation efforts. Presented in this paper is *LEAP*: Low Elevation Aerial Photography system that overcomes these challenges. *LEAP* is a mechatronic teleoperated kite retrofitted with a camera rig and a ground-based monitoring and network system that permits rapid acquisition, processing and distribution of aerial images. The off-the-shelf kite presented in the paper is man-backpackable, affordable, easy to deploy even at night and can quickly reach

[2] http://prism.mem.drexel.edu/projects/kite/index.html hosts the VRML model where one can virtually fly through an urban area near West Philadelphia

Figure 5: Image mosaic created from aerial images acquired by *LEAP*

flying altitudes a thousand feet or greater.

The flight dynamics that relate payload weight, wind speed and kite wingspan were presented. In addition, the attitude-regulating ability of the Picavet camera rig was analyzed. Elliptical pendulum literature is widespread but to the best of our knowledge, this paper is the first to present the Picavet's underlying mathematics. The intent of this was to assist researchers, interested in airborne instrumentation, in assembling their own kite-based rig. The image mosaicing, edge-detection and $3D$ reconstruction results leveraging the acquired aerial images illustrate some potential tasks. The marriage of mechatronics, intelligent sensing and mechanism synthesis follows our philosophy [3] [5] to holistically design computer vision systems. This resulted in an affordable, rapidly deployable, transportable and easy-to-fly system. In applying *LEAP* to disaster mitigation, we are currently investigating ingress/egress route generation. We hope to reference *LEAP*'s aerial images with private and public geospatial databases to automate driving directions for emergency medical teams moving between the site and hospitals.

References

[1] Debevec, P., Taylor, C.J., Malik,J., "Modeling and Rendering Architecture from Photographs: A Hybrid Geometry-and-Image-based Approach", *Proc. SIGGRAPH 96 Computer Graphics Proceedings*, Annual Conference Series, New Orleans, pp. 11-20, Aug. 1996.

[2] Hamel, T.; Mahony, R., Chriette, A., "Visual Servo Trajectory Tracking for a Four Rotor VTOL Aerial Vehicle", *IEEE International Conference on Robotics and Automation (ICRA)*, Washington, D.C., pp. 2781-2786, 2002.

[3] Oh, P.Y., Allen, P.K., "Visual Servoing by Partitioning Degrees of Freedom," *IEEE Transactions on Robotics and Automation*, V17, N1, pp. 1-17, February 2001

[4] Sharp, C.S., Shakernia, O., Sastry, S.S., "A Vision System For Landing an Unmanned Aerial Vehicle *IEEE International Conference on Robotics and Automation (ICRA)*, Seoul, Korea, pp. 1720-1727, 2001.

[5] Stanciu, R., Oh, P.Y., "Designing Visually Servoed Tracking to Augment Camera Teleoperators", *IEEE Conference on Intelligent Robotics and Systems (IROS)*, Lausanne Switzerland, Oct. 2002.

[6] Saripalli, S., Montgomery, J.F., Sukhatme, G.S., "Vision-based Autonomous Landing of an Unmanned Aerial Vehicle", *IEEE International Conference on Robotics and Automation (ICRA)*, Washington, D.C., pp. 2799-2804, 2002.

[7] Szeliski, R., "Video Mosaics for Virtual Environments", *IEEE Computer Graphics and Applications*, Vol. 16, No. 2, pp. 22-30, 1996

[8] Taylor, C., Kriegman, D., "Structure and Motion From Line Segments in Multiple Images," *IEEE Trans Pattern Analysis and Machine Intelligence (PAMI)*, V17 N11 pp. 1021-1032, 1995

[9] Zhang, H., Ostrowski, J.P., "Visual Servoing With Dynamics: Control of an Unmanned Blimp", *IEEE International Conference on Robotics and Automation (ICRA)*, Detroit, pp. 618-623, 1999.