Journal of Intelligent and Robotic Systems (2005) 44: 139–159
DOI: 10.1007/s10846-005-9030-4

Design, Implementation and Testing of a Vision System for Small Unmanned Vertical Take off and Landing Vehicles with Strict Payload Limitations

M. KONTITSIS, R. D. GARCIA and K. P. VALAVANIS*
Department of Computer Science and Engineering, Center for Robot Assisted Search and Rescue and Center for Urban Transportation Research, University of South Florida, Tampa, FL, USA
e-mail: {mkontits, rdgarcia, kvalavan}@cse.usf.edu

(Received: 27 September 2005; in final form: 2 November 2005)

Abstract. Alternative designs of a simple, low cost and effective vision system for small, portable, unmanned aerial vertical take off and landing (VTOL) vehicles are presented. Design configurations follow the 'on-board' and 'on-the-ground' processing concept and they depend on very strict payload limitations and power supply restrictions. Hardware and software components for both designs are described; advantages and disadvantages of both alternatives are compared; computational complexity is calculated and trade offs are discussed. Implementations on a series of small unmanned VTOL vehicles as well as testing details are included and experimental results are presented.

Key words: on-board processing, on-the-ground processing, payload limitations, unmanned vehicles, vertical take off and landing.

1. Introduction

The central objective of this paper is to present design, implementation and testing details of a simple and effective vision system suitable for miniature and small unmanned aerial vertical take off and landing (VTOL) vehicles with very strict and limited payload capabilities, not exceeding 10 lb (less than 5 kg).

In general, when deciding about the aerial vehicle configuration and its sensors, several factors/constraints need be taken into account: Weight limitations dictate compromises with respect to on-board equipment. Increased payload ability may lead to using more advanced sensors at the possible expense of requiring more power to operate and a more powerful processing unit that may drain the power source (batteries) faster. On average, a small unmanned vehicle (non-electric helicopter) has enough fuel to operate for about 15–40 min. Extra batteries are needed to keep the on-board equipment operating for at least as long as the vehicle itself; however, the vision system component operational time

---

* Corresponding author.

varies depending on power consumption and performance in terms of processing power and communication range.

Weight and power constraints set upper bounds for the processing unit to be used on-board the vehicle. This, in turn, dictates restrictions to algorithms that are implemented; high performance, high complexity algorithms require considerable processing power to run in real-time or near real-time. On the other hand, pushing an unmanned vehicle to operate under maximum payload may result in unstable performance.

Considering trade offs between the above coupled factors, two alternative designs are presented, following the '*on-board processing*' and '*on-the-ground processing*' concept. Their advantages and disadvantages are stated and analyzed along with their associated computational complexity, while trying to balance requirements and limitations imposed by the unmanned VTOL vehicle itself. The presented vision system uses at a minimum a color camera to capture images. Captured images are then interpreted and processed either on-board the vehicle or on-the-ground after being transmitted via a communication channel.

Control of the pan/tilt mechanism is obtained either via teleoperation (operator on the ground using a remote controller), or autonomously. Under autonomous operation the camera tracks either the fastest moving object in an image, e.g., a car, or an object dictated by the ground operator by clicking the mouse on the computer screen, or focuses on a 'location' commanded by the ground operator. The unmanned vehicle may be either teleoperated or fly autonomously / semi autonomously regardless of how the pan/tilt functions. It is beyond the scope of this paper to describe the control configuration and controllers designed for small helicopters, but details may be found in [1, 2].

It is essential to emphasize that the presented research does not offer innovations in color vision or image processing; instead, the main contribution of the paper is related to the integrated pan/tilt vision–unmanned vehicle system, using mostly off the self components, satisfying very strict payload and power consumption limitations. Used hardware and software components are discussed, followed by the presentation of two design configurations for on-board and on-the-ground processing. Considering the design for on-board processing, the total weight of the vision system component 'box' is approximately 1.6 kg (~3.5 lb) excluding the weight of the batteries used to power the system; it is less for on-theground-processing. When compared to existing designs in terms of weight, the closest one is the *AVATAR* system designed by USC but implemented on a vehicle with 10 Kgr payload [3, 4] (see Section 2). Overall cost effectiveness has also been a factor, with the overall vehicle–sensors–vision component cost not exceeding $15 K, by far the lowest compared to all other existing systems.

As far as actual implementation is concerned, both versions of the vision system have been installed, integrated with and tested on the *Raptor 60, 70, 90 SE* helicopter models. In terms of applications, the system has been tested for mine detection, traffic monitoring, overlooking unmanned ground vehicles

(UGVs) surveying an area, and tasking/recruiting a UGV to navigate to a specific location for additional data collection, and fire detection. Experimental results have been obtained by flying the *Raptor 90 SE* vehicle; results are good despite the simplicity of the vision system.

In addition, since the designed vision system is simple and rather generic, as briefly stated in the last Section of the paper, it has also been installed and experimentally tested on an Army Research Lab (ARL) *E-MAXX RC* truck (unmanned ground vehicle) proving the flexibility, suitability and wide applicability of the proposed design.

The rest of the paper is organized as follows. Section 2 summarizes related work and approaches used to design vision systems for unmanned helicopters, while Section 3 describes the model helicopter used, the *Raptor 90 SE* along with its limitations. Section 4 discusses details of the designed and implemented machine vision system components, followed by remarks on the computational complexity of the used algorithms. Analysis of advantages and disadvantages of the proposed design alternatives is provided. The fifth Section is dedicated to communication issues, limitations and trade off. Section 6 concludes the paper, offering details of how the system functioned when installed on a UGV, presenting vision system enhancements under consideration and conclusions.

2. Related Work

Vision systems, techniques and algorithms suitable for unmanned aerial vehicles range in complexity from simple color segmentation to statistical pattern recognition. Published related work and proposed machine vision architectures indicate the use of both 'on-board' [3, 5, 18, 22, 25] and 'on-the-ground' processing setups [6–15], with the latter being preferred most of the times. For on-board vision systems, due to the limited processing power provided by the on-board computer, derived algorithms have the additional constraint to run in real-time, requiring reduction of the computational load by processing selected image windows instead of the entire image.

Table I presents a summary of machine vision techniques used by University research groups, the main processing unit (on-board, on-the-ground) and the unmanned VTOL vehicle platform they have been implemented on. Table II summarizes functionality and capabilities of existing fully operational vision systems, including techniques employed by each one of them.

It is essential to distinguish between 'miniature', 'small size' and 'large size' unmanned VTOL vehicles. While the former two classes impose strict payload limitations, the latter one has been proven to be extremely suitable for on-board vision systems. The most notable such platform is the *Yamaha Rmax*. It consists of a two-stroke horizontally opposed 246 cc engine mounted to a 3.63 m long frame [16] with payload capacity of approximately 30 kg allowing accommo-

Table I. Existing vision systems for VTOL vehicles.

Institution	Machine Vision techniques used	Processing unit	Vehicle
Berkeley University [6]	No details provided	No details provided	BEAR
Georgia Tech [18, 19]	Edge detectors, morphing, statistical pattern matching	On-board	Rmax by Yamaha
Standford University [10, 12]	YUV color segmentation, signum of Laplacian of Gaussian (sLoG)	On-the-ground	Hummingbird Aerospace Robotic Laboratory at Standford
MIT [21]	Template matching	On-the-ground	Black Star by TSK
Rose Hulman IT (RHIT) [22]	Template comparison	On-board	Bergen Twin
IT Berlin [15]	No details provided	On-the-ground	MARVIN by SSM Technik
University of Texas [13]	Edge linking matching	On-the-ground	XCell .60
Swiss Federal Institute of Technology (ETH) [23]	No details provided	On-board integrated in camera	Huner Technik
Carnegie Mellon University [24]	Template matching and RGB color	On-the-ground	Rmax by Yamaha
USC [3, 5]	Omnidirectional, optic flow	On-board	Bergen Twin
Southern Polytechnic State Univesity [14]	Stereo vision, Sobel egde detector	On-the-ground	Vario Robinson R22
Linkoping University, Sweden (WITAS) [25]	No details provided	On-board	Rmax by Yamaha

Table II. Summary of system characteristics and functionality.

Institution →		Berkeley University	Georgia Tech	Univ. of South California	COMETS* [26]	WITAS[+] [25]	CNRS[~] [27]
Exprimental setup	Dynamic observer	X	X	X	X	X	X
	Dynamic environment				X	X	
	Static / man-made environment	X		X			X
	Known landmarks	X	X	X			
	Natural landmarks				X		
	Calibrated cameras				X		
Capabilities	3D reconstruction / depth mapping		X	X			
	Object identification	X	X	X	X		
	Object tracking		X	X	X		
Methods used	Optic flow			X		X	X
	Motion estimation	X			X	X	X
	IMU data						X
	Template matching	X		X	X	X	

*COMETS is a multi-national effort supported by the European Commission.

+ Wallenberg laboratory for research on Information Technology and Autonomous Systems (WITAS).

~ Centre National de la Recherché Scientifique (CNRS) in France.

dation of vision systems containing multiple cameras, radar altimeter, and complete desktop size computer systems [17]. Such a system is capable of autonomous navigation and control as demonstrated in [18, 19].

However, quite recently there has been increased interest for on-board vision processing applied to smaller vehicles (less than 1/4 scale). The most notable system is the USC *AVATAR* system that incorporates three *firewire* cameras, a *Microstrain 3DM-G IMU*, and a stack of five *PC-104* boards [3]. It has been mounted on a *Bergen Industrial Twin helicopter* utilizing a 46 cc twin cylinder engine with a 10-kg payload capacity. *AVATAR* has been proven effective in object identification and assisted flight control [4, 8].

Under an ideal scenario, and assuming that technology advancements exist, the ultimate goal would be to design a full size on-board vision system for unmanned vehicles similar to the *SureSight Enhanced Vision System* developed by *CMC Electronics* [20]. This system is used in manned helicopters to improve the situational awareness of the crew, but its tremendous capabilities set a desirable target if UAVs reach the level of full autonomy.

3. The Unmanned VTOL Vehicle System

The utilized unmanned VTOL vehicle platform is the *Raptor 90 SE*. It has been selected for its combination of high power output and small size. It exhibits relatively low vibration and it is able to withstand winds up to 15 mph. Its low purchasing cost makes it a very attractive choice. The *Raptor 90 SE* has the following characteristics:

- Manufacturer: Thunder Tiger
- Rotor diameter: 710 mm (symmetrical)
- Dry weight: 5.8 kg
- Dimensions: 130 × 27 × 48 cm (w/o blades)
- Payload capacity: 4 kg
- Endurance: 18 min
- Battery: 4.8 V (2.6 A)
- Fuel: 30% nitrous
- Engine: OS 0.91 C-spec

The *Raptor 90 SE* components including added sensors are shown in Figure 1. They are: i) Power supply, providing electrical power to all on-board equipment. ii) GPS, IMU and Compass that provide data for vehicle navigation. The GPS provides the longitude and latitude of the current position, and the IMU complements that information with acceleration readings along and about the x, y and z axes of the vehicle; iii) On-board computer primarily used to acquire data from the GPS, IMU and Compass and output commands to the servo control board. It may also be used to acquire and interpret image data provided by the camera; iv) Servo control board receiving commands from the on-board

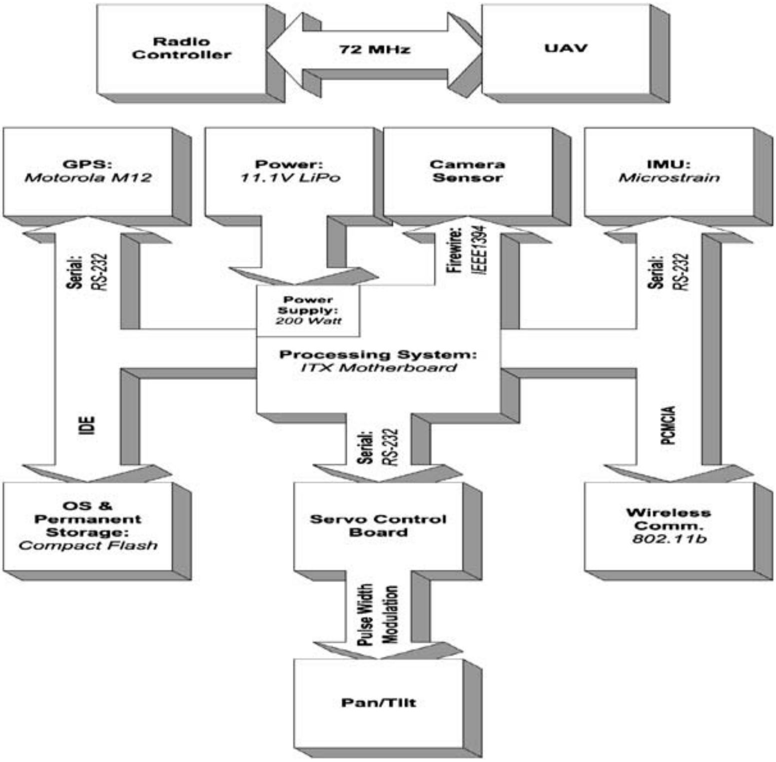

Figure 1. Equipment on-board a VTOL vehicle. *Arrows* indicate data flow.

computer, outputting signals to the servos; v) Vehicle servos that are attached to the control surfaces of the vehicle, dictating vehicle maneuvers; vi) Pat/tilt mechanism used for tracking and to direct the camera to a specific point in space regardless of the vehicle's heading. This allows for greater flexibility but it introduces the pan and tilt angles as two more values to be measured.

The vision system components are added to the above configuration. Details are provided below.

4. Proposed Vision System Designs

Two design alternatives, following on-board or on-the- ground processing are presented and compared. Advantages and disadvantages are related to their dependency on communications, available processing power, vehicle endurance and pertinent level of autonomy.

Placing the processing unit on the ground allows for utilization of a more powerful computer capable of running complex algorithms in real-time; the computer can be on the electrical power grid. Such advantages, however, come at the expense of increased reliance on the communication channel between the ground station and the VTOL vehicle. Without a reliable communication channel, there are either few or no images to be processed, rendering the latter useless in case of loss of communication. This has a profound impact on the vehicle's autonomy.

On-board processing results in greater vehicle autonomy. Image information is now transmitted over the communication channel for monitoring purposes only. Communication loss does not affect the function of the vision system or the vehicle, in general. The main disadvantage of the 'on-board' processing is that the processing unit on-board the VTOL must follow restrictions imposed by small payloads; further, limited on-board power supply cannot support power needs of a high performance computer. Durability is another limitation that cannot be overlooked.

4.1. ON-THE-GROUND PROCESSING DESIGN

The hardware configuration is shown in Figure 2, while Figure 3 shows the designed and attached on the vehicle platform with the pan/tilt mechanism, a flight computer and an RF video transmitter. This add on platform has already been integrated with the *RAPTOR 60, 70 and 90 SE* models. Equipment onboard the VTOL vehicle (see Table III for details), include: i) A color CCD camera, selected because it is lightweight and low power consuming; along with its power supply weighs 0.5 kg. ii) A 900-MHz transmitter/receiver set responsible for relaying image data back to the ground. The choice is made based on low weight and power consumption; total weight added by the transmitter and its power supply is 0.6 kg. It draws a current of 100 mA when operating. iii) A video converter acting as an interface device between the receiver and the firewire port of the computer. iv) Pentium 4 based laptop providing ample processing power to run machine vision algorithms. The operating system used is Linux.

Firewire interface (Figure 4): It provides access to actual image data. The open source libraries ***libdv***, ***libraw1394***, ***libavc1394***, ***libdc1394*** are utilized to access the firewire port and read data. Images are acquired in standard DV format with a resolution of *720x480* pixels and color depth of 24 bit. The *RGB* color space is used, with color depth of 8 bits. Down sampling follows to facilitate image transmission.
RGB to HSI conversion: The down sampled image undergoes a transformation from *RGB* to the more representative *HSI* space.
Threshold application module: This module is used to select pixel(s) that may belong to an object of interest. The object for which the vision system is searching is defined by a series of upper and lower bounds on Hue, Saturation and Intensity. The image is scanned and pixels that fall within those bounds are

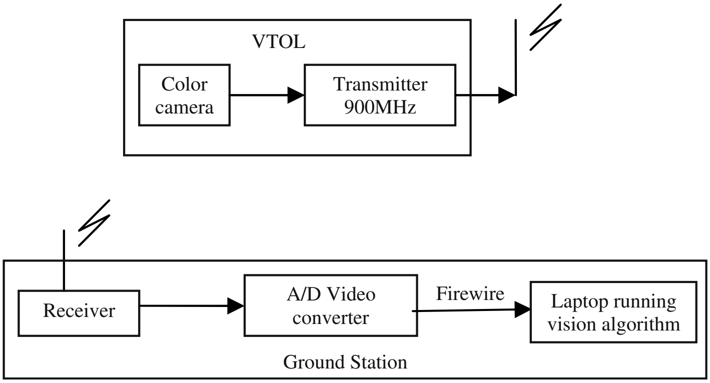

Figure 2. Hardware configuration for 'on-the-ground' processing.

selected as belonging to the object in question. The main implementation concern is simplicity.

Decision mechanism: This module is responsible for the final decision regarding the presence of a target. It raises or lowers an alarm signal indicating that something of interest may be present in the image. In detail, if the number of the selected pixels exceeds a threshold, then the decision mechanism classifies the frame as containing an object and increases a counter by a constant value. In the case that nothing is detected the same counter is decreased by a quantity relative to the exponential of its current value. When the value of the counter is greater than a certain threshold then the alarm is raised. This can be viewed as a 'leaky bucket' that fills gradually every time a frame is found to contain an

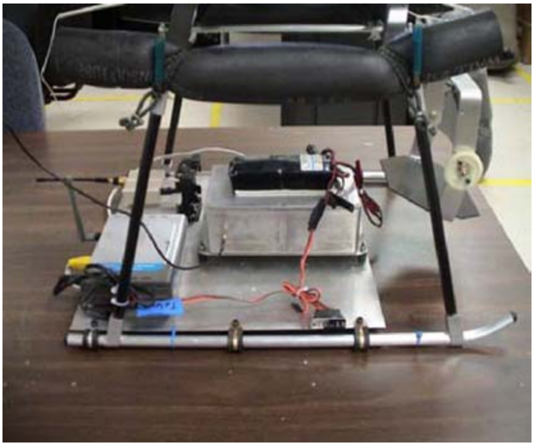

Figure 3. Add-on platform.

Table III. On-board equipment-payload limitations considered.

Device	Specifications		
	Weight	Power Consumption	% of payload
Color CCD camera	0.028 Kg	1.32 W	1.03
Transmitter	0.020 Kg	1.08 W	0.73
GPS	0.091 Kg	<1 W	3.34
IMU	0.090 Kg	<1 W	3.30
Pan – tilt A	1.224 Kg	Approx 1 W	44.88
Pan – tilt B	0.2 Kg	Approx 0.5 W	7.33
Flight control computer	0.567 Kg	Approx 2 W	20.79
Power supply	0.363 Kg		13.31

object of interest and drains rapidly when an object is not present. The operation of the decision making module may be described in pseudo-code as:

```
IF object == detected THEN counter = $counter + C1$
    Else counter = $counter - \exp(counter/C2)$
IF counter > activation threshold THEN alarm = ON
        Else alarm = OFF.
```

The constant *C1* is related to the rate at which the counter is increased with each ‘detection,’ while *C2* controls the descent of the counter’s value when an object is not present. By selecting those two constants it is possible to tune the behavior of the decision making mechanism in terms of its tendency to raise the alarm. For the same threshold value a larger value for *C1* will result in an easier activation of the alarm since the counter will be increased by a larger amount. Similarly, a smaller value of *C2* will lead to a steeper descent of the counter

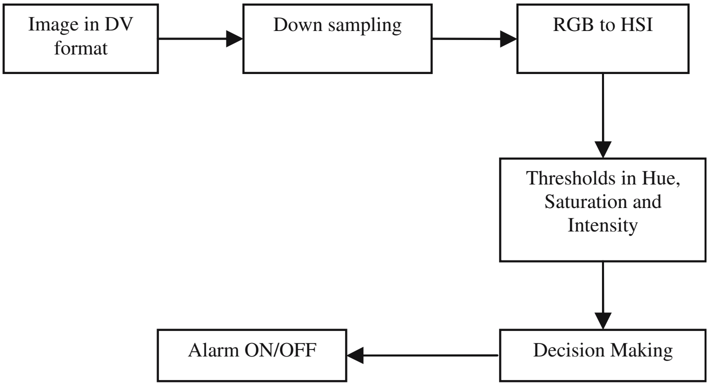

Figure 4. Vision system software modules.

when an object is not present resulting in a faster deactivation. Typical values for *C1* and *C2*, found after some experimentation, are 2 and 40, respectively.

To avoid extremely high counter values that would make proper deactivation of the alarm almost impossible, an upper bound, typically 100, is introduced and the counter is not allowed to exceed that bound even if the object of interest is continuously present in the image. There is also a lower bound, usually 0 to 20, which the counter never goes below in order to avoid very low values that would prevent the alarm to be activated properly. In this way a scale from 0 to 100 is created for the values of counter with higher ones corresponding to a higher number of recent frames containing an object of interest.

4.2. ON-BOARD PROCESSING

The alternative to have the image processing unit on-board the VTOL allows for greater autonomy for the vehicle since it doesn't rely on a ground based computer. The role of the communication channel also becomes secondary as far as the operation of the system is concerned. It is used only to send images for monitoring by on ground personnel. The payload in this configuration includes (Figure 5):

- 1.2-GHz EPIA processor
- Via-embedded motherboard
- Unibrain firewire camera
- Microstrain 3DM-G IMU
- 1 gig 266 MHz RAM
- 1 gig compact flash
- Compact flash to IDE adapter
- Motorola M12+ GPS receiver
- 8-channel servo controller
- 200-W power supply
- 11.1-V LiPo battery
- 802.11B cardbus

This configuration has been chosen because of its high computational capabilities, various I/O ports, size, low heat emission, and cost. The vision system is packaged into a 32 × 19 × 5 cm basswood box, Figure 6, mounted on a lightweight aluminum sheet. This sheet is mounted directly on the *Raptor's* skids via rubber insulated pipe clamps. The slim design of the enclosure allows mounting of the system without modification to the standard carbon fiber skids. The box is coated with a gas proof heat shrunk plastic typically used to coat model airplanes.

Basswood was chosen for the enclosure due to its lightweight nature and its lack of electrical conductance. The system's camera is mounted directly to a

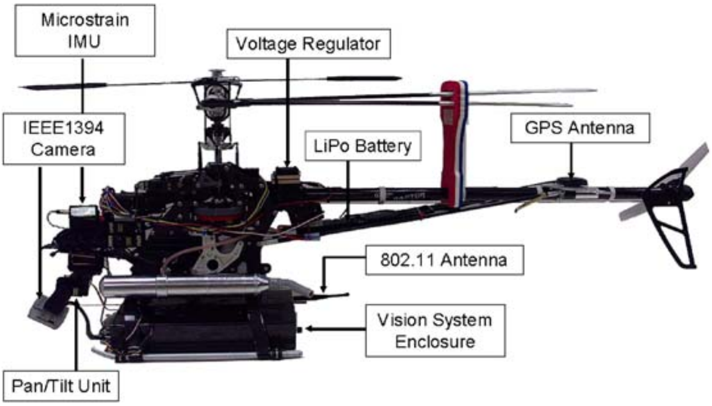

Figure 5. The equipment used in the ‘on-board’ configuration.

Lynxmotion pan/tilt unit (noted as pan-tilt B in Table III). This unit is, in turn, mounted directly to the underside of the *Raptor's* servo tray. The pan/tilt system consists of two Futaba S3004 servos that are interconnected by 1/3 cm laser cut Lexan. This setup allows the camera to pan and tilt up to 90°. Servo commands are issued by the eight channel servo control board located within the enclosure. Power for the vision system is supplied via the 14.8 V 4Ah Lithium Polymer (LiPo) battery mounted on the lower front section of the boom. LiPo's were selected based on their high amperage, low weight, and small packaging. Voltage from the battery is regulated to 12 V by a LT1084CP-12 regulator mounted to the rear of the main shaft. Power distribution to the system is controlled by the 200-

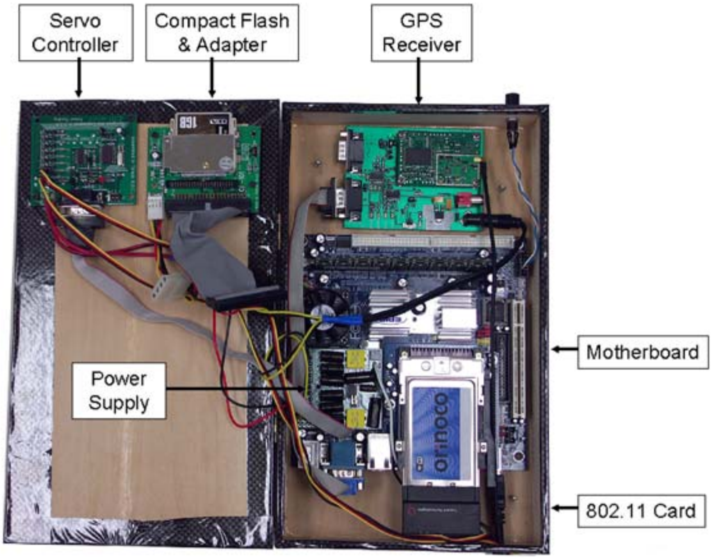

Figure 6. The computer on-board the VTOL vehicle.

W ATX power supply. The power supply plugs directly into the motherboard allowing the unit to add nothing to the physical dimensions of the system. The median for all peripherals of the system is the EPIA VIA M2 motherboard. This 1.2 GHz ATX motherboard provides multiple I/O interfaces, RAM, and CPU on a single board.

This ATX board has a distinct advantage over typical PC-104 boards that require distinct boards for processor, ram, interfaces, etc. The ATX form motherboard also allows for an extremely thin designed enclosure where PC-104 boards are limited to a stack type configuration.

The two implemented systems share most of the software modules, but there exist differences too as shown in Figure 7. The on-board system has been enriched with a region of interest selection mechanism that allows for faster processing since only a portion of each frame's pixels is processed instead of the whole image. Particularly, only regions that have been previously classified as containing an object are selected for further processing. To allow for the introduction of new regions as candidates for selection, the algorithm scans a frame in its entirety every other 15 frames.

4.3. COMPLEXITY

During the design of both systems, simplicity has been the primary constraint. This has resulted in an algorithm having to apply a threshold on the pixels of the image, making it of order $O(n^2)$, where n is the dimension of a square image. The main computational burden is posed by the conversion of the image into the HSI color space. More specifically, the calculation of the *hue* component of the image includes a call to the inverse cosine and the square root function. The decision making module has also been slightly modified and it now relies on a majority vote among the last 15 frames in order to classify a region as one that contains an object of interest. This is not preventing the system from achieving near real-time

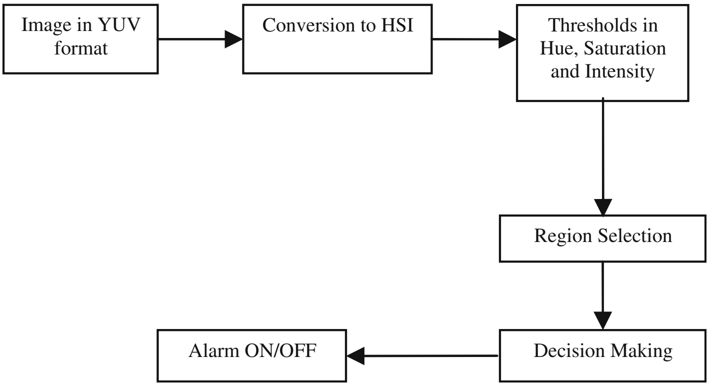

Figure 7. Vision system modules for the on-board system.

performance. Especially with the use of the down sampling module, the number of pixels to be evaluated is significantly decreased and therefore adding to the speed with which the system operates. The computational cost of the down sampling is also $O(n^2)$, but it is completed much faster since it requires nothing but a rearrangement of the existing pixels.

With the incorporation of the region selection mechanism this number is drastically decreased which allows the on-board vision system, despite having less computational power than the ground based laptop, to achieve a processing rate of 30 to 80 fps. The region selection algorithm consists of a series of iterations, each of which expands a bounding box around a given pixel. The computational time that it needs depends on the number and the size of the targets present as well as the size of the image. The worst case is again $O(n^2)$, where n the size of the nxn image. The common case though is to have a small number of regions, usually one or two. Since every region is not allowed to exceed a certain size, if it is to be considered a valid object, it can be said that the computational time is bounded by a constant. The Decision making module is a simple equality test in the on-board system and a little more complex leaky-bucket mechanism for the system with the processing unit on the ground. In both cases the delay is negligible and independent of the size of the image. The evaluation of both systems in terms of complexity is summarized in Table IV.

4.4. EXPERIMENTAL RESULTS

Several experiments have been conducted to evaluate the system's performance. Experiments were conducted using the *RAPTOR 90 SE*. The 'add on' platform was successfully integrated with the helicopter, without exceeding maximum payload capacity. The VTOL vehicle executed flights under a teleoperated mode; the pan/tilt mechanism with the camera was also controlled from the ground.

The vision system that used the 'on-the-ground configuration' processes 15 frames per second.

This was mainly due to the limits set by the firewire drivers that the system used to capture the digitized video in DV format. The on-board system however, despite being less powerful in terms of processing power achieved 30 fps. This is

Table IV. Computational complexity of the software modules.

System configuration	Modules			
	Downsampling	Conversion to HIS	Region selection	Decision-making
On-board system	N/A	$O(n^2)$	Constant	Constant
On-the-ground system	$O(n^2)$	$O(n^2)$	N/A	Constant

Table V. Comparison of the two approaches.

System configuration	Properties			
	Dependence on comms	Available CPU power	Autonomy	Endurance
On-board system	Low	Low	High	40 min
On-the-ground system	High	High	Low	11/2 h on DC Indefinite (AC)

due to the drivers that allowed capturing images at lower resolutions. The performance can also be accredited to the region of interest mechanism which selects only part of the image for further processing. The evaluation of each approach is presented in Table V.

The vehicle flew in an outside environment with several artificial landmarks that needed be identified. A set of landmarks were black, almost spherical, some half-buried in the ground, imitating mine types. Figures 8, 9 and 10 show input images, objects of interest, captured by the camera while the vehicle was flying, as well as the resulting system output. In Figures 8, 9 and 10, pixels belonging to the object are painted white, while all others are black. Shadows and other disturbances had little effect on the overall system performance, since only few frames were misclassified as containing objects of interest. Therefore, the final output of the decision making module was not affected. The angle at which the camera was directed towards the ground had no effect on the system as illustrated in Figure 9 which shows a view of the object from directly above. Figure 11 demonstrates a case where the system erroneously classified some image pixels as belonging to the object of interest. The object of interest was intentionally placed in the shadow of a tree in an attempt to test the system's

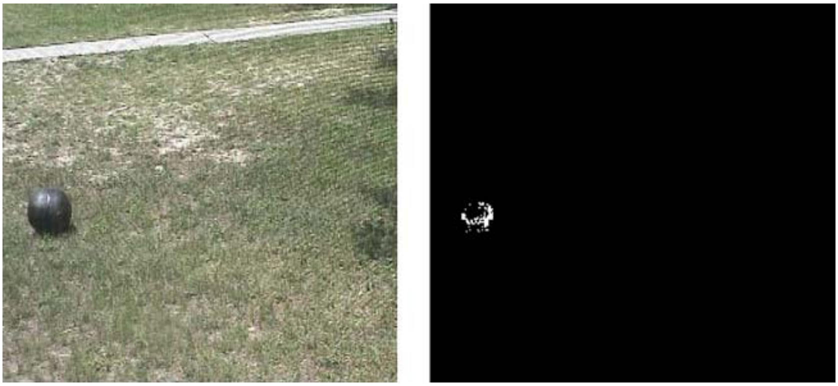

Figure 8. First pair: Input image and system response. Object of interest pixels classified correctly.

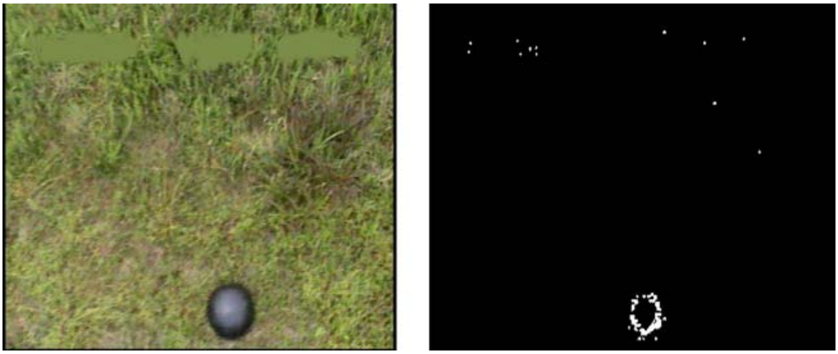

Figure 9. Second pair: Some pixels have been misclassified as belonging to the object of interest.

sensitivity to changes in lighting conditions. This can be remedied by the introduction of a module in the algorithm that will compensate for the various lighting conditions. In Figure 12 another test under different lighting conditions is demonstrated. The minimum bounding rectangle containing the object is superimposed over the actual input image. More results demonstrating the operation of the current version of the vision system may be found in: http://www.prldev.csee.usf.edu/public/vtolVision/.

5. Communication Issues

Communications and communication range play a very important role in relaying telemetry or video data back and forth between the ground station and the vehicle. Video may be transmitted either over a wireless Ethernet network or over a separate dedicated channel. The trade off is reliability and bandwidth; although a separate channel is preferred, one should also consider weight and

Figure 10. Third pair: Pixels belonging to the object of interest are classified correctly.

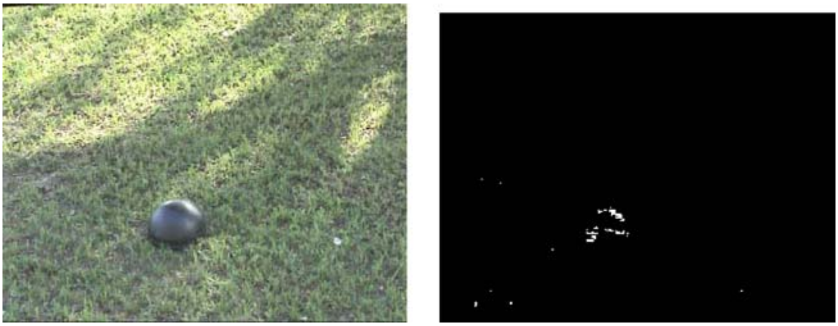

Figure 11. Fourth pair: Illustration of performance in different lighting conditions. Some pixels are misclassified due to the presence of shadow.

power consumption saved by using only wireless Ethernet for all data transmissions.

For on-the-ground processing, the communication channel is of most importance. If it fails or becomes unavailable, the whole vision system comes to a hold due to the lack of input. Therefore in the implementation that follows the ‘processing on the ground’ configuration a separate video link exists that utilizes a pair of RF receiver-transmitter. The transmission rate achieved is 30 fps.

The on-board processing system utilized wireless Ethernet to transmit images and data. Since the channel exists for monitoring purposes only, it is less crucial for the vision system functioning. In fact, image data processing is completely independent of the communication channel. However, the ground station must provide the human operator with some visual input. Image data are relayed through a socket using UDP protocol. Initially the whole image was passed through a buffer across the socket. This proved to be inefficient in terms of bandwidth utilization (see Figure 13). After some experimentation it was found that by braking down the image into parts that can fit in the 1,472 bytes of the

Figure 12. Test image under different lighting conditions. A red minimum bounding rectangle superimposed over the area that the algorithm detected the object.

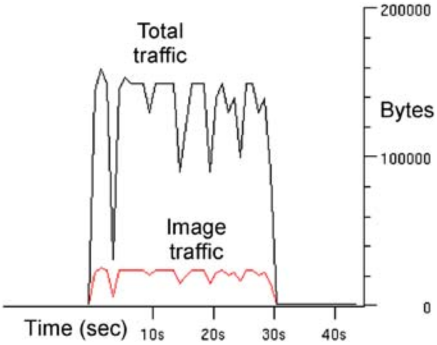

Figure 13. Actual bandwidth utilization-image is passed as a single buffer through the socket.

UDP data packet the portion of the network bandwidth that is used to carry image data is very close to 95% (see Figure 14). This leads to a better utilization of the available network bandwidth. It allows for a transfer rate of approx 15 fps.

6. Concluding remarks and discussion

Two different approaches regarding VTOL vehicle specific vision systems have been presented and compared. Given the comparative study results, a system was designed and implemented considering payload capabilities and limitations of the *RAPTOR* series unmanned helicopters. Experiments and evaluation tests included cases where the system was called to find objects of interest under different lighting conditions in outdoor environments. Performance has been shown to be quite promising in terms of accuracy.

Despite the fact that the vision system was originally designed with an aerial platform in mind, it can also be used on board an unmanned ground vehicle

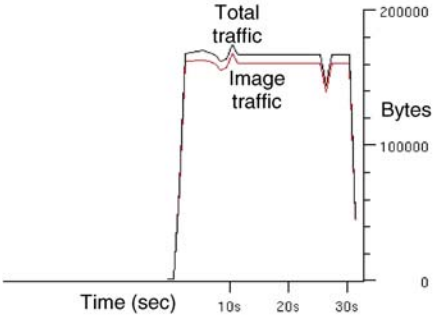

Figure 14. Actual bandwidth utilization-image passed through the socket after broken down into a series of smaller parts.

(UGV). The utilized UGV platform is an EMAXX RC truck, shown in Figure 15, with the following characteristics:

- Manufacturer: Traxxas Corporation
- Max speed: 30 mph
- Drive system: Shaft-drive 4-WD
- Dry weight: 3.8 kg
- Dimensions: 49 × 41 × 27 cm
- Endurance: 40 min
- Battery: Dual: 7.2V LiPo
- Motor: Dual Titan™ 550
- Speed controller: EVX FWD/REV electronic

This platform was chosen due to its rugged nature, wide wheel base, and adjustable suspensions system. Although the dynamics of the UGV are not as sensitive as the VTOL’s, special attention must be taken to assure that platform is resilient to rollovers, high centering, and ground strikes, see Figure 15. To prevent rollovers the vision system is mounted as close to the platforms natural Center of Gravity (CG) as physically possible and the stock shock mounts are moved away from the CG to increase the wheel base of the platform. To prevent high centering and ground strikes spacers were inserted into the stock springs of the suspension system. This forces the shocks to become stiff causing the system to react more aggressively to vertical forces. With only minor modifications regarding primarily the mounting of the vision box on the UGV, the vision system can be considered as a versatile processing unit suitable for a variety of small unmanned vehicles.

Regardless of the carrying platform, an important issue that needs be addressed is automatic system adaptability to extreme lighting variations. The current version of the designed system relies on fixed thresholds to identify (describe) objects of interest. This, although proven functional, lacks adaptabil-

Figure 15. E-MAXX RC truck (UGV) with the on-board vision system.

ity. Further, when the mission objective changes, the vision system must be re-initiated to accommodate such changes.

Future versions of the system, currently under design, will include ability to learn from experience (to make the system even more robust). The system will be able not only to respond to a broader spectrum of input images but also to accumulate ‘experience’ as it operates. This would help minimize chances of false alarms while maintaining a high correct identification rate. An integral part of any future development of the system will be the incorporation of a tracking mechanism. This will allow the system to track and follow objects in its field of view either automatically or as directed by a human operator..

Acknowledgements

This research has been partially supported by ONR Grants, N00014-03-01-786 and N00014-04-10-487, and a DOT Grant through the USF CUTR, 2132-1054.

References

1. Castillo-Effen, M., Alvis, W., Castillo, C., Valavanis, K. P. and Moreno, W.: Modeling and Visualization of Multiple Autonomous Heterogeneous Vehicles, *Proceedings, IEEE International Conference on SMC*, October 2005.
2. Castillo, C., Alvis, W., Castillo-Effen, M., Valavanis, K. and Moreno, W.: Small Scale Helicopter Analysis and Controller Design for Non-Aggressive Flights, *Proceedings, IEEE International Conference on SMC*, October 2005.
3. Vision Hardware, University of Southern Calafornia, [online] 2004, http://www.robotics.usc.edu/~avatar/vision_hw.htm (Accessed: 27 January 2005).
4. Mejias, L., Saripalli, S., Sukhatme, G. and Cervera, P.: Detection and Tracking of External Features in an Urban Environment Using an Autonomous Helicopter, *IEEE International Conference on Robotics and Automation*, pp 3983–3988, 2005.
5. Hrabar, S. and Sukhatme, G. S.: A Comparison of Two Camera Configurations For Optic-Flow Based Navigation of a UAV Through Urban Canyons, *Proceedings, IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sep 2004, pp. 2673–2680.
6. Sastry, V.: Vision based detection of autonomous vehicles for pursuit evasion games, *15th Triennial World Congress*, Barcelona, Spain, 2002.
7. Sattigeri, R. and Calise, A. J.: An adaptive approach to vision based information control, *Guidance, Navigation and Control Conference*, No. AIAA-2003-5727, Austin, Texas, August 2003.
8. Hrabar, S. and Sukhatme, G. S.: Omnidirectional Vision for an Autonomous Helicopter, *Proceedings, IEEE International Conference on Robotics and Automation*, pp 3602–3609, 2003.
9. Saripalli1, S., Montgomery, J. F. and Sukhatme, G. S.: Vision-based Autonomous Landing of an Unmanned Aerial Vehicle, *Proceedings, IEEE International Conference on Robotics and Automation*, pp. 2799–2804, May 2002.
10. Woodley, B. R., Jones, H. L., LeMaster, E. A., Frew, E. W. and Rock, S. M.: Carrier Phase GPS and Computer Vision for Control of an Autonomous Helicopter, *ION GPS-96*, Kansas City, Missouri, September 1996.

11. Debitetto, J.: Modeling and simulation for small autonomous helicopter development, *AIAA-1997-3511 Modeling and Simulation Technologies Conference*, New Orleans, Louisiana, August 1997.
12. Woodley, B., Jones, H., Frew, E. and LeMaster, E.: Dr Stephen Rock "A Contestant in the 1997 International Aerial Robotics Competition'', http://sun-valley.stanford.edu/papers/WoodleyJFLR:97.pdf.
13. Holifield, N., Lallinger, J. and Underwood, G.: International Aerial Robotics Competition 2004, University of Texas at Austin IEEE Robot Team Aerial Robotics Project June 1, 2004 http://iarc1.ece.utexas.edu/~lynca/final_documentation/utiarc2004.pdf.
14. Burleson, W., Salhany, W. and Hudak, J.: Southern polytechnic State University Autonomous Remote Reconnaissance System, http://a-robotics.spsu.edu/SPSU_paper2005.pdf.
15. Musial, M., Brandenburg, U. W. and Hommel, G.: MARVIN – Technische Universität Berlin's flying robot for the IARC Millennial Event, in: *Proceedings Symposium of the Association for Unmanned Vehicle Systems 2000*, Orlando, Florida, USA, 2000.
16. RMAX- http://www.yamahamotor.co.jp/global/business/sky/lineup/rmax/index.html (Accessed: 27 January 2005).
17. Johnson, E. and Mishra, S.: Flight Simulation for the Development of an Experimental UAV, *AIAA Modeling and Simulation Technologies Conference and Exhibit*, 5–8 August 2002.
18. Proctor, A., Gwin, B., Kannan, S., Koller, A., Christophersen, H. and Johnson E.: Ongoing Development of an Autonomous Aerial Reconnaissance System at Georgia Tech, http://controls.ae.gatech.edu/gtar/iarcpapers/git2004.pdf.
19. Johnson, E. N., Calise, A. J., Tannenbaum, A. R., Soatto, S., Hovakimyan, N. and Yezzi, A. J.: Active-Vision Control Systems For Complex Adversarial 3-D Environments, A tutorial, Proceedings of the American Control Conference, 2005.
20. Enhanced Vision System (EVS) Overview, CMC Electronics, [online] 2001, http://www.cmcelectronics.ca/En/Prodserv/Commav/commav_evs_overview_en.html (Accessed: 22 January 2005).
21. Johnson, E., DeBitetto, P., Trott, C. and Bosse, M.: The 1996 MIT/Boston University/Draper Laboratory Autonomous Helicopter System, *Proceedings of the 15th Digital Avionics Systems Conference*, 1996.
22. Groven, J., Holk, E., Humbert, C., Krall, J. and Schue, D.: Rose-Hulman Institute of Technology Autonomous Helicopter for the 2004 International Aerial Robotics Competition.
23. Chapuis, J., Eck, C., Geering, H. P., Mudra, R., Schneuwly, B. and Sommerhalder, R.: The Swiss Entry into the 1996 International Aerial Robotics Competition, *AUVSI '96 Proceedings*, pp. 947–953, Orlando, Florida, July 1996.
24. Amidi, O.: An Autonomous Vision-Guided Helicopter, MA Thesis, Carnegie Mellon University, 1996.
25. Nordberg, K., Doherty, P., Farneback, G., Forssen, P.-E., Granlund, G., Moe A. and Wiklund, J.: Vision for a UAV helicopter, *IEEE/RSJ International Conference on Intelligent Robots and Systems, Proceedings, Workshop WS6 Aerial Robotics*, pp 29–34, Lausanne, 2002.
26. Ollero, A., Ferruz, J., Caballero, F., Hurtado, S. and Merino, L.: Motion compensation and object detection for autonomous helicopter visual navigation in the COMETS system, *Proceedings, IEEE International Conference on Robotics and Automation*, New Orleans, Louisiana (USA), April 2004.
27. Ruffier; F. and Franceschini, N.: Visually guided micro-aerial vehicle: automatic take off, terrain following, landing and wind reaction, *Proceedings IEEE International Conference on Robotics and Automation*, New Orleans, Louisiana, April 2004.
28. Gonzalez, R. and Woods, R.: *Digital Image Processing*, Addison Wesley, 1992.