Abstract—Sense and avoid systems for civilian unmanned air vehicles (UAVs) are essential in controlled airspace under visual flight rules (VFR). A prototype optical sensor accomplishes the task with attractive performance specifications. Key requirements include long-range detection (up to 10 km), wide field of view, discrimination of small threats against the background and tolerance of direct solar illumination. We demonstrate a prototype system based on a network of independent camera modules equipped with local processing. Availability of a fly-by-wire helicopter configured as a UAV emulator allows for realistic field tests with consumer components. Aspects of the design, implementation and evaluation of the prototype sensor are presented here, as are preliminary measurements to clarify the roles of platform motion, system optical point-spread, noise, direct sunlight and target highlighting.

I. INTRODUCTION

UAVs are extensively used in the military, government and private sectors. In order to operate in uncontrolled airspace, such aircraft should be able to autonomously “sense and react” to avoid other aircraft [1] [2]. The intent is to provide the level of safety equivalent to or exceeding that of an onboard pilot. With respect to collisions, this capability can be met by a “sense and avoid” (SAA) instrument. Numerous versions of the SAA instrument based on radar, LIDAR and passive-optical, among other technologies, have been envisaged [3].

Visual sensors are suitable for this application due to their mass, size, frame rate, power, angular resolution and coverage [4] [5] [6] [7]. Some of the shortcomings of passive optical technology include degraded performance under low atmospheric visibility, the indirect nature of range information, scene clutter and susceptibility to damage from sunlight. CMOS image sensors, however, overcome these shortcomings by being readily available, low cost, and rapidly evolving with advances in optical, computing and imaging technologies that are pushed by sundry applications [8].

II. BACKGROUND AND MOTIVATION

The critical factor in collision avoidance systems is the timely recognition of potential collisions. Sufficient time must be given to perform evasive actions that will maintain the minimum separation distance of 165 m (500 feet). A buffer of at least 12 seconds between detection and any collision is widely used [1]. Many sensor, target and environment parameters affect this time buffer, including target velocity, angular resolution, frame rate, optical field-of-view (FOV) in azimuth and elevation, atmospheric visibility, camera point-spread function, platform jitter and availability of processing capability.

Customized CMOS imagers can deliver dramatic improvements in frame rate, dynamic range, digital/analog integration, power dissipation and unit cost - but not necessarily simultaneously. Enunciation of the specifications for a custom imager needs abundant data on the sensitivity of system performance to these parameters. Suitable flight data is valuable in modeling realistic or worst-case scenarios, the acquisition of which is facilitated by the National Research Council of Canada (NRC) Bell 205 airborne UAV emulator. Use of the emulator relaxes the mass, power and size constraints of the prototype, allowing commercial off-the-shelf (COTS) components to be employed to collect data essential for the customized imager specification. It can be expected that some of the requirements will change with the UAV platform and the mission.

A common metric is required to simplify the system engineering of seemingly independent parameters. Relative importance and sensitivity of each parameter can be judged with this metric. The “range of first detection”, \( R_0 \), is an intuitive figure of merit that is platform-independent, and can be readily determined by ground-based measurements. Conversely, the time buffer \( t_{buf} \), although a more fundamental quantity, is scenario- and platform- dependent, where the link between the two is provided by modeling. Key parameters limit \( R_0 \) and therefore the performance of the system. They include the physical metrics of mass, power, size and cost, signal-to-noise ratio due to local contrast between background and target, optical performance due to point-spread function effects, image processing to raise local contrast and apparent height above the horizon. The end objective is the design and development of an optical sensor that satisfies the requirements of a SAA instrument. We present here the design decisions and preliminary measurements towards achieving this goal.

III. SCENARIO

The baseline scenario for the present discussion is operation under VFR conditions, no traffic information (e.g. transponder not operating) and non-cooperative targets [9]. A “platform”
houses the sensor – a task accomplished here by a fly-by-wire Bell 205 helicopter posing as a surrogate UAV [10]. For modeling convenience, both platform and target aircraft are assumed to be in level flight. Diffuse sunlight illuminates both the background, comprised of sky, clouds and ground, and the target. Detection needs to occur at ranges beyond a kilometre - perhaps as far as 10 km - in order to allow for evasive action by the platform. At such ranges, the angular size of the target is small and signal from the target is generally mixed with the background through the optical point-spread function of the sensor system. Reduced atmospheric visibility compounds this problem. Target aircraft fall under the FAR 23 general aviation standard; a Cessna 117 being a typical example with a wingspan of $\approx 10$ m and maximum airspeed of 250 km/h (70 m/s or 136 knots). Target altitudes and velocities are not expected to exceed 3000 m (10000 feet) and 360 km/h (100 m/s or 195 knots) respectively.

IV. INSTRUMENT DESCRIPTION

In the present work, a prototype collision avoidance sensor has been created to operate on the helicopter UAV emulator. Aside from the mass and power benefits, the helicopter has extensive image storage capability and a suite of navigation instruments (attitude, GPS, altimeter) that can be read by the sensor. The sensor was configured as a network of many camera modules [11], which allows the use of COTS optics for coverage of the field of view, reduces communication and processing bottlenecks and facilitates the introduction of new hardware and software components.

Several camera modules are clustered into a camera “group” that may have additional control and processing capability. A number of camera groups comprise the complete instrument (Fig. 1). This hierarchy means that most evaluations can be performed with a reduced instrument having only one or two camera groups. The total field of view needed for collision avoidance is $\pm 110^\circ$ azimuth by $\pm 15^\circ$ elevation [2], which can be achieved with about 20 camera groups (i.e. $10^2$ camera modules in total). Total dimensions of the instrument are not specified, but are expected to be near $60 \times 20 \times 10$ cm$^3$.

Fig. 1. Block diagram of the sensor system

Elements of a camera module are shown in block form within each module in Fig. 1. The optics consist of a lens, spectral filter and shutter. A typical detector array used here is a commercial digital camera module, including readout circuitry, digitizer, some pixel-level processing and a USB communications interface. The local computer receives the imagery in USB format, selects a region of interest and performs appropriate image processing. The image processing is intended to identify ‘events of interest’ that need to be characterized and reported up the hierarchy.

Fig. 2 is a photograph of a laboratory prototype of one camera group of the instrument. Standard C-mount lenses and processor boards dominate the photo. The camera modules (five in total) are mounted on a common baseplate to form a camera group with a total field of view of about $45^\circ$ azimuth by $8^\circ$ degrees elevation. A reduced-mass version of the camera module is the basis of the helicopter-mounted prototype. In the photograph, all of the camera modules have identical optics, detector arrays and preprocessors. Interchangeable $f/1.4$ lenses with variable aperture were used to measure the effect of aperture. A colour detector array was the most readily available, even though a monochrome sensor should improve the signal-to-noise ratio (SNR). Likewise, the local preprocessing computer was selected for low cost and rapid delivery. Ethernet links the camera modules to a camera-group computer that handles post-processing and communications. Imagery is usually processed only at the camera-module level. Table I lists selected specifications of a camera group. Mass and power are 10x over endpoint in this prototype.

V. DESIGN DISCUSSION

The development plan is to investigate range performance under various conditions (atmosphere, target and clutter) before sensor optimization. Table II lists representative features of the collision avoidance sensor/scenario. To illustrate the complexity of design tradeoffs, Fig. 3 shows three system parameters (mass, power and frame-rate), and the major subsystems (optics, detector array and processing), along with the approximate fractional contribution of each subsystem to these parameters. Various designs can lead to the same range of first detection. A tradeoff between power and mass, for instance, could keep the range constant while changing the processing and optics. The concentric circles indicate $R_0$ in a
qualitative way. Larger ranges generally require greater larger mass/power/frame-rate. The modular design of the instrument aids in the evaluation of design alternatives.

Highlights of the design include the use of distributed camera modules, local image processing, instrument point-spread function and the hazards of direct sunlight illumination. System design enhancements, which are ongoing, generally address such issues by a combination of simulation and experiment. Only a few camera modules need to be modified for evaluations, and tests are first performed in the laboratory and on the ground.

A. Distributed Cameras

Although there can be mass/power and complexity penalties for using a distributed sensor network [11], the large and asymmetrical optical field of view of the collision avoidance sensor benefits greatly from partitioning of the optics into a large number of separate lenses with overlapping fields. Further benefits of a distributed sensor include enhanced reliability and parallel image processing capability. Custom CMOS fabrication will allow the imager and suitable preprocessing to be combined on a single chip. Finally, the cost of the sensor is linearly related to the total field of view (the field of view is augmented by simply adding more camera modules), whereas the single-sensor costs can rise quadratically and at the expense of frame rate.

Stray light increases when the sun is close to the field of view of a camera module. If the stray light is excessive or direct sunlight imminent, the affected detector arrays are powered down. Potential sensor damage is reduced by integrating mechanical aperture shut-off capability. Because the sensor has 10 - 100 camera modules (depending on field of view and other design considerations), the fractional loss of coverage by the (temporary) deactivation of a camera module is small, a boon compared to single-sensor devices. Spurious specular reflections can also supply direct-sunlight-equivalent irradiance on a subset of the detector array, but are initially considered rare and transient.

B. Image Processing

The point spread function (PSF) of the sensor system determines the mixing of point-target signals and the optical background. Contributions to the PSF include optical aberrations and diffraction, detector pixelation, platform motion and atmospheric scattering. Mitigating the effects of the PSF is the first task of image processing.

Reduction of signals that are not from the target of interest is important. The image of a small target is generally viewed against sky or other background (clouds, ground) whose per-pixel signal can be comparable to that of the target. Removal of this ‘background clutter’ is another task of the image processing. Further tasks include discrimination of signals of interest from random noise by a combination of noise estimation, signal deconvolution and thresholding [12]. Secondary image processing activities are expected as the instrument evolves.

### TABLE I

<table>
<thead>
<tr>
<th>Component</th>
<th>Parameter</th>
<th>Nominal Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Camera group</td>
<td>Number of modules</td>
<td>5</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Angular field coverage</td>
<td>$45 \times 8$</td>
<td>degrees</td>
</tr>
<tr>
<td></td>
<td>Module separation</td>
<td>10</td>
<td>degrees</td>
</tr>
<tr>
<td></td>
<td>Mass</td>
<td>4</td>
<td>kg</td>
</tr>
<tr>
<td></td>
<td>Power</td>
<td>30</td>
<td>W</td>
</tr>
<tr>
<td></td>
<td>Communications</td>
<td>1</td>
<td>Gbps</td>
</tr>
<tr>
<td></td>
<td>Resolution</td>
<td>$1600 \times 1200$</td>
<td>pixel</td>
</tr>
<tr>
<td></td>
<td>Dynamic Range</td>
<td>8</td>
<td>bit</td>
</tr>
<tr>
<td></td>
<td>Frame rate</td>
<td>5</td>
<td>fps</td>
</tr>
<tr>
<td></td>
<td>Pixel pitch</td>
<td>2.0</td>
<td>$\mu$m</td>
</tr>
<tr>
<td></td>
<td>Focal length ($f$)</td>
<td>25</td>
<td>mm</td>
</tr>
<tr>
<td></td>
<td>F-number ($f/#$)</td>
<td>1.4–16</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Effective FOV</td>
<td>$10 \times 8$</td>
<td>degrees</td>
</tr>
</tbody>
</table>

### TABLE II

<table>
<thead>
<tr>
<th>Features</th>
<th>Nominal Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range of First Detection</td>
<td>10</td>
<td>km</td>
</tr>
<tr>
<td>Time to evade</td>
<td>12</td>
<td>s</td>
</tr>
<tr>
<td>Angular Resolution</td>
<td>$0.2 (\simeq 1)$</td>
<td>mrad (pixel)</td>
</tr>
<tr>
<td>Azimuth Field-of-regard</td>
<td>$\pm 110$</td>
<td>degrees</td>
</tr>
<tr>
<td>Elevation Field-of-regard</td>
<td>$\pm 15$</td>
<td>degrees</td>
</tr>
<tr>
<td>Overall dimensions</td>
<td>$60 \times 20 \times 10$</td>
<td>cm$^3$</td>
</tr>
<tr>
<td>Scene Dynamic Range</td>
<td>$10^5$</td>
<td></td>
</tr>
<tr>
<td>Cost</td>
<td>$10^4$</td>
<td>$</td>
</tr>
<tr>
<td>Mass</td>
<td>2</td>
<td>kg</td>
</tr>
<tr>
<td>Power</td>
<td>50</td>
<td>W</td>
</tr>
<tr>
<td>Frame rate</td>
<td>10</td>
<td>Hz</td>
</tr>
</tbody>
</table>
VI. SIMULATIONS AND EXPERIMENTS

Measurements were compared to MATLAB models embedded in a system-level simulation that could reproduce experimental results and predict sensor performance. Initially, a library of in-flight imagery was acquired from the surrogate UAV platform. Image sequences were generated as a series of fly-bys with target and platform at level altitudes and with varying trajectories. Most images did not contain targets and were used to create realistic backgrounds in the raytracing simulator. Both measured and synthetic pointlike targets were then combined with the background within the simulator.

Performance predictions for selected laboratory conditions were used to validate the simulator. Field tests were used for simulator validation of realistic PSF and platform vibration. The simulator was able to predict performance under extreme field conditions as well, such as low visibility (due to inclement weather) and head-on collisions.

A. SNR Variation with Target Range

Each pixel of the detector array projects out onto a “virtual area” or “footprint”, $A_{foot}$, at the range of the target, as shown schematically in Fig. 4. Light from the scene elements within the footprint can be imaged onto the pixel, while light from outside the footprint hits other pixels. For this discussion, it is assumed that light diffusely reflected from the target has a clear line of sight to the sensor, and that light from the background scene can be considered to arise from a ‘virtual scene’ that is at the same range as the target but is partially obscured.

In the presence of optical point spread, the footprint boundaries of each detector element blur and expand. This causes scene features in $A_{foot}$ to spread among nearby pixels, and background to bleed into $A_{foot}$. Fig. 5 shows the variation of SNR for a single pixel as a function of target range for different background reflectivities, as per Eqs. 1 and 2. The images used were generated via raytracing synthetic targets at realistic distances embedded into flight imagery. Ambient lighting was controlled to simulate different background intensities.

\begin{equation}
SNR = \frac{|dn_{sig} - dn_{back}|}{dn_{noise}}
\end{equation}

Here, $dn_{sig}$ and $dn_{back}$ refer to target and background intensities in digital number respectively. The noise is modeled as a combination of shot and read noise measured in electrons, and converted to digital number by the conversion gain $cv$.

Returning to Fig. 5, the noise is primarily the shot noise from the combined signal-plus-background. The SNR values require the selection of numerous instrument parameters (integration time, aperture, quantum efficiency and so forth), and so can best be considered as relative. These parameters affect both signal and background equally. Note that the signal can be brighter or darker than the background. When signal and background have the same reflectivity, the signal disappears. In this situation, the use of spectral or polarization features of the target and background may be important.

An effective target area of 1.0 m$^2$ has been used. At sufficiently long range, the target is small compared to the pixel footprint ($A_{foot}$), and most of the signal (and noise) comes from the background. For nearby targets, the entire pixel is filled with light from the target. At intermediate ranges, the signal depends on the difference in reflectivity between target and background. The variation of signal with range arises from the increase in occlusion of the background by the target, until $A_{foot}$ is completely filled by the target.

B. Effects of the point-spread function (PSF)

The SNR of a pixel containing the target image is shown in Fig. 6, with various point-spread functions added. Here, $\sigma$ represents the standard deviation of the PSF in units of pixels. The impact of PSF on instrument performance can be seen from its effect on $R_0$ (Fig. 7), with the detection threshold from Fig. 6. There is a large reduction of $R_0$ for large PSF width, which can be caused by reduced atmospheric visibility [13], providing, as expected, a strong link between visibility and first detection. Some over-design of the sensor under normal conditions (i.e., $R_0$ greater than 10 km) would...
allow for adequate performance at the low-visibility end of VFR.

Image processing is expected to recover some of the performance reduction caused by the PSF. In Fig. 8, the SNR versus range is shown for low atmospheric visibility, for the uncorrected and corrected cases. Also shown is the curve with negligible PSF. Here, a PSF $\sigma = 10$ [pixel] was used, as the worst-case combined PSF as per the full-width at half maximum value of 7.5 [pixel] estimated in Fig. 9. The correction was performed by applying a simple Laplacian of Gaussian (LoG) based point-detection algorithm [12]. In this case, $R_0$ has increased by 40% due to image processing i.e. the range of detection shifts from 8 km out to nearly 12 km. Advanced processing techniques will improve the performance.

Fig. 9 displays the representative point spread functions. The PSF of the optics/detector has a full width at half maximum of 2 pixels, which includes diffraction, aberrations, camera focus and detector array pixelation. As mentioned previously, atmospheric scattering broadens the PSF by an amount related to the atmospheric visibility. In addition, the motion of the sensor platform adversely affects the PSF at long integration times. The motion broadening in Fig. 9 has been determined using the measured jitter in the yaw axis of the surrogate UAV in level flight (speed of 90 knots) using recorded accelerometer, altimeter and GPS data. An integration time of 75 ms was used, which is at the upper end of expected integration times and so provides a worst-case approximation of platform-motion jitter.

C. Horizon modeling

Apparent target height above the horizon, $h_{tgt}$, is a key identifier of targets on a collision course. Note that $\Delta h_{tgt}$ can be measured on each image that contains a target, since the location of the horizon line can be estimated from prior measurements (from the same and different camera modules) or modeling if it is not present on the image.

The utility of $h_{tgt}$ can be best illustrated by a ‘fence’ scenario, which assumes that the first detection of the target occurs at the outer fence boundary, placed at $R_0$. Targets
piercing the fence are nominally detectable and, assuming level flight, follow a predictable trajectory. Calibration and other systematic errors are not expected to be significant. Fig. 10 shows a number of near-collision trajectories. Preliminary flights with a Bell 205 platform and a Harvard Trainer aircraft target validate Fig. 10 for selected (safe) trajectories. It is expected that targets on a collision course will maintain an altitude coincident with that of the platform, and therefore exhibit a ‘stuck pixel’ profile in the imager, barring noise. Platform motion is common to target and horizon. Such motions may shift the target position, but they also shift the horizon (or any other scene feature) by an equal amount. Hence, measurements of $\Delta h_{tgt}$ can be averaged over many frames, regardless of the platform motion.

For targets on a collision course i.e. matching altitude (100 m in the figure), the horizon-height profile is relatively constant. For all other target elevations, $\Delta h_{tgt}$ grows non-linearly with decreasing closing distance. In order to track $h_{tgt}$, the angular positioning precision of the sensor needs to be less than about 1.0 mrad. Because there is only a single target, angular resolution, in the sense of the Rayleigh criterion [14] and the PSF, is not important. For the modeled sensor choices, target miss distance is discernable out to at least 7.5 km. Although the variation may be within the platform PSF, appropriate image processing to deconvolve the point-spread is expected to reveal the target at this range.

VII. CONCLUSION

In conclusion, we present a system design informed by simulation and real-life image and platform data. The range of first detection $R_0$ was used as a key experimental figure of merit, and preliminary models and measurements show the importance of local background illumination, system and atmospheric PSF, image processing, and apparent horizon position for the long-range collision detection scenario.

ACKNOWLEDGMENT

The authors would like to thank the National Research Council (NRC) of Canada and the Ontario Centres of Excellence (OCE) for their financial support.

REFERENCES