May 2008

Usability of Image Generation Platforms to Produce Oblique World Views

Michael T. Curtis
University of Central Florida

David Schuster
University of Central Florida, david.schuster@sjsu.edu

Florian Jentsch
University of Central Florida

R. Swanson
University of Central Florida

Michelle Harper-Sciarini
University of Central Florida

Follow this and additional works at: https://scholarworks.sjsu.edu/psych_pub

Recommended Citation
Usability of Image Generation Platforms to Produce Oblique World Views

Michael T. Curtis, David Schuster, Florian Jentsch, & Michelle Harper-Sciarini

University of Central Florida

Abstract

The interpretation of oblique aerial views in UAV operations is an important task, yet one that provides unique challenges. The human visual system is mal-adapted to visual perception from the air, especially when oblique views are involved. Hence, UAV operators have to be given training in the interpretation of oblique aerial visual stimuli. A novel way to train operators on how to accurately inspect and perceive oblique aerial images is to administer multiple trials of discrimination tasks for paired oblique images. In discrimination training, trainees view pairs of oblique aerial views and determine whether the images are taken from comparable positions, or contain similar information, etc. For such training, a number of images need to be developed. There are a number of ways that these training images can be generated: The methods differ in the amount of detail they provide and with respect to the resources they require for their development. This paper examines the feasibility and usability of different methods of image generation for aerial UAV imagery. Development of stimuli from sources such as aerial photographs, web-based models, and off-the-shelf simulations are discussed.

Introduction

Industry analysts predict that spending on unmanned aerial vehicles (UAV) will double in the next decade (Teal predicts, 2007). The bulk of the expenses is the result of the military’s demand for combat and surveillance aircraft whose operation does not endanger pilots. Further, the cost savings of UAVs to piloted aircraft are also appealing. At approximately $35 million per Global Hawk UAV (Bigelow, 2006) aircraft and $4.5 million for the smaller Predator drone (Herbert, 2003), there is considerable cost saving from the approximately $135 million per F-22 Raptor fighter jet currently in use (Military: F-22, n.d.), or the $300 million price tag on larger commercial jets.
The FAA is also under increased pressure, fueled by perceived industry benefit, to relax the stringent regulations on civil UAV operation (Jewell, 2005). The future benefit of allowing more UAV use, in civil airspace, has implications for improving industries as diverse as law enforcement and agriculture. The combination of industry spending, and increased interest in civil use, suggest an increasing ‘unmanned’ airspace. Based on the fundamentals of supply and demand, increased demand for UAV activity will breed an increased need for skilled operators. Given some factors surrounding the UAV industry, effective and efficient operator training is fundamental to maintaining safe airspace for the future; specifically, perceptual training.

Perception and UAV Training

Visual perception is a critical component of any aviation operation. In Manned Aerial Vehicle (MAV) operations, inexperience with flight can sometimes lead to visual misperception of the environment, which in turn, can cause actions that may lead to unstable or even unsafe flight procedures (Kern, 2002). The operational tasks of a UAV are similar to MAV operations, but UAV operators do not inhabit the aircraft, and they thus are not afforded the luxury of looking out the cockpit window or experiencing vestibular motion information. The lack of context and motion during UAV operations puts even more reliance on the operator’s visual system. As a result, the negative effect that inexperience has on visual perception during UAV operations may be worse than in MAV operations.

Visual perception in UAV operations is not only critical for basic operations. In some military tasks, for example, UAV operators must rely on what they perceive in order to make critical decisions about a detected target (McCarley & Wickens, 2004). Maximizing the UAV operator’s perceptual abilities is, therefore, an essential part of training.

Discrimination Training
Research on perceptual training for visual tasks that require image interpretation or target detection, such as baggage screening, has suggested that perceptual learning occurs through the deliberate practice of discrimination and differentiation (Hoffman & Fiore, 2007; Ward, 2008). One way to optimize training for these tasks is the use of discrimination training. In this type of training, operators must determine whether two images are, in some respect, same or different.

For example, in discrimination training for baggage screening, trainees are given two X-ray images, either of which contains a target object (e.g., a gun). The target object seemingly appears in both images of the pair, but it is in different locations and orientations in each. Trainees must decide whether the target object is the same, or whether there are indeed two different target objects in the images. Similarly, in discrimination training to improve aircraft pilots’ ability to judge visual approaches to landing, the trainees must decide whether two aerial views of airports are taken from the same relative location to the runway, or from different ones (Figure 1). In this task, the trainees must learn to ignore differences in superficial characteristics, such as runway length, weather, visibility, and the topography surrounding the runway(s). Instead, trainees must become able to derive the relative locations of the images’ origins, compare them, and decide whether they are the same or different.
Studies, specifically focusing on discrimination training, have demonstrated improvements in target identification performance (Fiore, Scielzo, Jentsch, & Howard, 2006). Fiore et al., for example, have identified several important determinants of successful discrimination training, including, the presentation of difficult discrimination tasks in early training. In particular, completing difficult discrimination tasks early in training helped participants to make precise comparisons in their studies, whereas encountering easy discrimination tasks early in training caused participants to make less-precise comparisons. In other words, the difficult discrimination task let participants consider more features in the images before making a judgment.
Another determinant of successful discrimination training is the provision of multiple task trials. This holds true for both complex images and simple shapes (Sohn, Doane, & Garrison, 2004). The requirement for a large number of trials, however, can quickly become very costly. To offset this cost, Hoffman and Fiore (2007) suggested compressing time within training to maximize benefit. In other words, reduce the intervals between practice trials, and spend more time training for critical moments.

Summary

Above, we have discussed the importance of visual perception and discrimination perception training for the operation of UAVs. Furthermore, research that has established the key elements of effective discrimination training was described. Taking all of this into consideration, it seems logical that perceptual discrimination training can benefit UAV operators. For this type of training to be considered, however, one must first identify the images that should be used for the discrimination training.

Discrimination Training Images for UAV Operations

Views from aircraft are characterized by the obliqueness in which the world below is represented. An oblique viewpoint is defined as a viewpoint angle that is between the vertical and horizontal (cardinal) axes of orientation (Chown, Kaplan & Kortenkamp, 1995). Oblique viewpoints are not only prevalent in aviation, but have also been used in photography, 3-D modeling (Jizhou, Zongjian & Chengming, 2004), and as an aid in wayfinding tasks (Ohmi, 2000).

Terrestrial Vs. Aerial Viewpoints

Visual perspective can be categorized as terrestrial or aerial (Paine & Kiser, 2003). The terrestrial viewpoint is one for which the human visual system has adapted over thousands of
years. Physiologically speaking, our visual system is most heavily attuned to cardinal visual orientations. That is, objects that appear at or close to horizontal or vertical orientations are most easily perceived. Several studies have been conducted to examine prevalence of cardinal orientations in scene content (e.g., Baddeley & Hancock, 1991; Keil & Cristobal, 2000). Although scene content studies vary, the predominant outcome is either a horizontal or a vertical bias. A quick glance around should confirm this bias toward cardinal orientations for objects. Both behavioral and psychophysical research has suggested a reduction in a person’s ability to visually perceive objects that deviate from cardinal orientations. This effect is better known as the oblique effect (Appelle, 1972).

An aerial perspective viewpoint is much different from what humans are best adapted to. Among aerial views, there are two types of viewpoints of the earth’s surface; vertical and oblique. The vertical viewpoint is akin to terrestrial vision, in that it involves vision near or at a vertical orientation. In both aviation and UAV operations, however, an oblique viewpoint is more applicable. At oblique viewpoints, the horizontal and vertical biases of scene content from a terrestrial viewpoint are altered. Consequently, objects that appear on the cardinal axes, from the ground, may appear at oblique angles when seen from the air. The research on the oblique effect has demonstrated that it generally is more difficult to retrieve information from oblique images. Oblique viewpoints can further be categorized as high oblique and low oblique. High oblique viewpoints are those that include a visible horizon. Low oblique are viewpoints that are at angles not steep enough to include the horizon (Paine & Kiser, 2003). In aviation, pilots generally use high oblique viewpoints to maneuver the aircraft. This is because low oblique viewpoints are generally only present at very steep angles atypical of aircraft maneuvering, except in extreme flight maneuvers.
Given the relevance of oblique viewpoints in UAV operations and the results of research on the oblique effect, it becomes clear that perceptual discrimination training should include oblique images. These images can be used to familiarize trainees on the visual perceptual cues that differ between terrestrial and oblique viewpoints. In order to provide UAV operators with oblique viewpoint discrimination training, a series of static, oblique images, must be developed. Using these images, the training can highlight important cues and offer many practice trials in succession.

Table 1

<table>
<thead>
<tr>
<th>Characteristic Category</th>
<th>Characteristic</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implementation</td>
<td>Cost</td>
<td>The monetary and time investment required to produce one or many images</td>
</tr>
<tr>
<td></td>
<td>Availability</td>
<td>Ease of access to the image generation method and target</td>
</tr>
<tr>
<td></td>
<td>Replicability</td>
<td>The difficulty of obtaining more images beyond the first</td>
</tr>
<tr>
<td></td>
<td>Inputs</td>
<td>The information required to set up the image generation method</td>
</tr>
<tr>
<td>Outcome</td>
<td>Specificity</td>
<td>The level of precision in the definition of the image</td>
</tr>
<tr>
<td></td>
<td>Resolution</td>
<td>Highest level of detail represented by the image</td>
</tr>
</tbody>
</table>

Oblique Image Generation

In this paper, we examine methods of generating oblique images and assess them from a general usability standpoint. There are several ways oblique imagery can be obtained. The preferable method depends on the intended use of the images. For perceptual training applications, the angle to, and distance from the targets, as well as the field of view, are important qualities of the image. Thus, images are needed which match very precise parameters.
Consequently, the image generation method must be very accurate and result in images of high fidelity.

For other applications, the absence or presence of a target may be the only important feature of the image. For identification of vehicles, as an example, a series of images, each containing a different type of vehicle, may be all that is needed. It is helpful to begin with a description of the image requirements and to keep them in mind when comparing image generation methods, as the methods vary in their output.

As we will describe, some image generation methods require more complex input data for image creation. In this section, we will compare several methods of generating images. For classification and assessment, we identified six criteria to consider when selecting a method of image generation (Table 1). Each of the six characteristics can be categorized as either an implementation characteristic or an output characteristic.

**Implementation**

Implementation characteristics are those that deal with the creation of the image. The four implementation characteristics define (a) the monetary and time demands of the method (cost), (b) the effort involved in obtaining the necessary equipment (availability), (c) the setup of the method (input), and (d) the ability of the method to produce consistent images (replicability). Together, these implementation characteristics represent the investment into the project.

**Cost.** The cost of generating an image or a set of images is determined by both financial and time considerations. Effective low-cost solutions are desirable. Given that rental costs of a helicopter can reach $400-$600 per hour, aerial photography is not a catchall solution (Angel & Hickman, 2002). Off-the-shelf PC-based simulations are modestly priced when compared to full-
scale simulator time or custom development costs. Other applications requiring more advanced technology are now possible at greatly reduced costs from years past (i.e., GPS).

**Availability.** Availability refers to the ease at which the images can be generated. Specifically, it is important that images of the desired targets can be obtained. Some of the image generation methods are limited in what content they provide. In many cases, it is desirable to train individuals in real world settings or with full-scale simulators. However, the great demand for time in these training environments restricts their availability.

**Input.** A component of availability is determined by the input data that is required to generate the image. As more specific images are needed, more data is required to utilize the image generation method. For example, obtaining any oblique image of a target requires the general location of the target. Compare this to the need for an image at a specific angle, size, and field of view. In the latter case, the exact coordinates are needed, as are the relative coordinates and position of the “camera.” Indeed, additional equipment or tools may be required to obtain all the necessary input data. In short, the input information that is needed depends on the method, and for oblique viewpoints, some input information is hard to determine if not measured by an instrument (i.e., heading and angle of the aircraft).

**Replicability.** Real world settings and high-end simulators are also not very scalable. That is, as the training need grows, supplying training to a large number of individuals becomes cost-prohibitive. A method is replicable if it can be repeated with minimal additional time and expense and can produce images consistently. For training development, replicability is critical in maintaining reliability. If creating an image twice produces different results, it will be difficult to standardize training. If, in a discrimination training task, the key critical perceptual cue was
viewpoint angle, an inconsistent generation of viewpoint angle would give an unreliable comparison.

Outcome

Outcome characteristics are the qualities of the generated image. The degree that these are important depends on the application of the images. In situations where recognition of a specific target (e.g., a tank or building) is the goal, matching image quality to the resolution that real world application will present would be important. In contrast, a general perceptual concept such as change in visual angle would not require as much image quality as specific target training.

Specificity. Specificity refers to the ability to obtain the exact image desired. The most specific image is one that views a geo-referenced object at an exact angle and from an exact distance away. As described previously, very specific images tend to require very specific input data. The highly specific perceptual training example will need similarly specific input data about target location, desired viewing angle, and the like. The image generation method should support the input of this specific location information. If the image generation method required exact coordinates in our vehicle recognition scenario, however, it would be a hindrance.

Resolution. The resolution of the image refers to the quality of the image that is produced from a given method. Resolution is the highest level of contrast that can be represented by the image. In a high resolution image, fine lines can be very close together and still be distinguished. In many cases this can affect the level of perceived realism (Duh, Lin, Kenyon, Parker & Furness, 2002). Resolution is often expressed in pixels or lines per unit of length. Increasing the size of an image will lower its resolution, as long as the number of pixels or lines
Generation of Oblique Aerial Images

remains constant. Thus, higher resolution images can be enlarged more before the effect of the pixels or lines is detectable.

Platform comparisons

Considering the importance of the aforementioned characteristics, we now examine specific methods of generating oblique images. By highlighting the strengths and weaknesses of each method, we hope to provide a clear picture of the capabilities of each method, while recognizing that the importance of the strengths and weaknesses will vary depending on the application.

Four methods are discussed in descending order of photorealism. Two methods utilize existing or new photographs. The third uses Google Earth, a tool that mixes various satellite photos with rendered objects. Finally, we examine Microsoft Flight Simulator, which uses exclusively rendered (computer-generated) images.

Preexisting Photography

One method of oblique image generation is to use preexisting photographs. Preexisting photography is widely distributed through a number of sources. The most prominent of these are Internet sources where a simple keyword search can turn up countless images on many topics. The limitless potential for the use of preexisting images is promising, but it comes at the cost of precision which is typically reduced, be it only because the images are not accurately labeled with information about when and where they were taken.

Implementation. In terms of implementation, preexisting photographs are generally agreeable to restricted budgets. Preexisting photographs can be accessed for little (e.g., Pictometry, or Air Images Ltd.) to no charge (e.g., Google Image Search). The time to locate images will depend on the availability and specificity of the desired images though. The
availability of preexisting images is less predictable. For example, a search for images of the Atlanta Hartsfield Jackson International Airport resulted in a large number of images, from a wide range of locations. More specific inquiries (e.g., along specific runways) can even result in a more precise selection of images. At the same time, however, a similar search for a dirt airstrip in the Saudi Arabian desert would prove substantially less fruitful.

If images actually are available for a particular location, there are no additional input requirements to develop the image. In this case, replication of images is merely a product of availability. If a large number of images are available, there is a better chance a replicable image exists. In many cases, especially for developing images for approach and landing viewpoints, the available images are severely limited. This could reduce the training benefit.

Outcome. Preexisting photographs of oblique images lack the benefit of specificity. Although there may be a wealth of runway pictures at different angles of approach, and different runways, without specificity the consistency of images in a perceptual training program will be limited. Unless measured beforehand, specifics like coordinates, angle of approach, or heading may not be in a usable format. For example, one may not have enough data to determine the attitude of the camera or the geo-referenced coordinates of the target. This limits the analysis of the objects in the photo. Due to the preexisting nature of this method, the resolution can be variable. Despite the potential variation, in most cases the photos will be high resolution pictures.

Using preexisting photography has its greatest utility for target recognition training. An example of this would be for distinguishing between friendly and enemy tanks. In this case, the specific viewpoint coordinates are not as important as just having different perspectives of the same target.

Photography
Generating images using aerial photographs differs from using preexisting imagery in that the images are generated after predetermining desired viewpoints. Images can then be captured using photographic or video image capturing software.

**Implementation.** Taking oblique images of a target can provide a great deal of freedom in target selection and very specific data. However, there is a high time and resource cost, as it requires the use of a camera and a mount for the camera, typically on an aerial vehicle. For some applications, especially when a series of images is desired at specific coordinates and angles, it is problematic that the image measurements are obtained, not set. In other words, equipment can precisely describe the geo-referenced coordinates of the image, but if this image is taken from an aerial vehicle, it will be difficult to obtain an image from a precise location in space. When the target is more important than the field of vision, such as in instantaneous vehicle speed measurements, aerial photography can provide a sufficient level of specificity (Angel & Hickman, 2002). A GPS unit provides the X, Y, and Z coordinates of the photograph. To determine the angle at which the photograph is taken, an Inertial Measurement Unit (IMU) is needed. An IMU records the attitude, or position in space, of the aircraft.

**Outcome.** Obtaining and using actual photographs provides the highest level of coverage and resolution of all the methods we examined. Images can be obtained for any location that the camera can go. The camera can be focused on any target within the focal power of the lens. The resulting images are of the real objects, and because they are taken when needed, changes over time can be observed.

In short, when a limited number of high-resolution images of a target are needed, and the resources are available to obtain them, photography is preferable. For developing training
stimuli, quantity and control may be more valuable than resolution. In that case, generating images using computer software is a better solution.

**Google Earth**

There are a number of programs available via retail and on-line that provide geo-global positioning images. Of these, Google Earth is one of the most widely used. Google Earth imagery was developed using satellite imagery of the earth’s surface.

**Implementation.** Google Earth is widely available as a download from the Google website and comes in three different versions (Which version, n.d.). The basic version is free, and the advanced versions cost between $20 and $400 per year, so the cost is comparatively low, especially when compared to the use of actual aerial photography where, depending upon the platform, hourly costs of $400 and more are typical.

The implementation strengths of Google Earth include ease-of-use and fast image generation. One can enter geo-referenced coordinates, giving an exact top-down view of a geographical point. However, the angle of elevation must be adjusted manually, using a mouse-based slider control, and no angle is displayed. Rather, one must calculate the angle using the elevation and distance. Starting from a top-down view at a particular elevation (labeled Eye alt in Google Earth), one then tilts the map using the slider until one reaches the calculated distance away from the target. This is further complicated by the fact that Google Earth does not tilt the map smoothly. Although one cannot see it from the interface, the zoom changes slightly as the camera pans. The result is only approximate; one obtains an image close to the correct angle of elevation, but with an off-center target. This method can be repeated to generate many images, each taking only a few minutes to generate.
Outcome. Google Earth generates oblique views from satellite photography of the Earth. Thus, the majority of the images are photographs. Some 3-D rendered buildings in major cities are available as an option. The detail, and availability of an image, varies by location. Most of the United States is thoroughly imaged, especially cities.

Google Earth falls between using preexisting photography and using a computer simulation in terms of image quality. Regardless of the detail of the satellite photo, the final image will be limited in resolution. Any image displayed on a computer monitor will be limited to the maximum number of pixels per inch that are displayed. On a typical display, there are 96 dots per inch. Printed materials typically have a resolution near 600 dots per inch (dpi). The $400 per year Google Earth Plus is capable of outputting higher resolution images up to 4800 x 4800 pixels in size. To maintain 600 dpi, the image must be 8 inches x 8 inches or less. This is much lower quality than the photographic method described above, but the benefit is a much larger variety of images.

The variety of images is a specific benefit of satellite imagery. Using the zoom function, many views can be generated beyond the range of aerial photography. The images are updated when new satellite photos are taken. The tool does not display changes over shorter periods of time, and one typically cannot access past images, so it is only useful for stationary targets.

Microsoft Flight Simulator

Commercial-off-the-shelf (COTS) software has steadily gained acceptance as a low cost alternative to training using real world applications (i.e., aviation and military applications). For the generation of oblique aerial images, there are several available options. Although there are differences between different COTS software that can be used for oblique image generation, the available products are generally very similar. In light of this, we selected a widely available well-
known software package Microsoft Flight Simulator X (Microsoft, 2006) to represent available COTS image generation platforms.

**Implementation.** Flight Simulator is readily available for $50 and can be installed on modern personal computers with a dedicated graphics card. Its wide availability, coupled with its accuracy, makes it an attractive option for generating training stimuli. In order to generate coordinate-specific images external inputs are required. Web-based tools such as the TPL Where Are You (TPL-WAY) calculator can simplify this process. Using TPL-WAY (Figure X), the necessary glideslope or altitude of the camera can be calculated. The user supplies the target location as geo-referenced coordinates, includes the desired ground rotation of the target in degrees, and inputs either the altitude or the glideslope. TPL-WAY will then compute the remaining values (e.g., heading, altitude, and coordinates) for entry into Microsoft Flight Simulator.

![TPL-WAY calculator](image)

**Figure 2.** Screenshot of TPL-WAY calculator.

**Outcome.** Microsoft Flight Simulator trades photo-realism for high image-generation accuracy. Using the cockpit windshield as a camera lens, an image can be generated for any
location the plane can be “placed”. Flight Simulator does not have the accuracy limitations of
Google Earth. However, for other purposes, the benefit of the tool’s precision and speed is not as
important as image quality. In visual discrimination training for pilots, for example, more precise
viewpoint angles are needed.

This is the largest limitation of the software; all terrain is rendered, and the environment
does not change over time. Although the software does include extensive terrain and feature
detail, and additional terrain and features can be installed, the potential targets are limited to
what the designers of the environment have chosen to depict. The rendered buildings and plants
do not have as high of resolution as a photograph. Furthermore, Flight Simulator does not have
an image-file generation function, so images must be taken as screenshots using the screen
capture command built into Windows. Consequently, the maximum image resolution is limited
to that of the computer display, typically 96 dots per inch.

Summary

There are a number of available options for the generation of oblique aerial images. In
order to optimize the utility of image generation, consideration should be taken on the function
of the perceptual training that will result from these images (Table 2). For less location specific
training, photographic imagery will likely provide the best option. On the other hand, training
that requires location precision will sacrifice image quality for more exact specification of the
image.

Table 2

<table>
<thead>
<tr>
<th>Method</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>Best For</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preexisting photography</td>
<td>Low cost</td>
<td>Low control over targets, not very specific</td>
<td>Images of a common target</td>
</tr>
<tr>
<td></td>
<td>High availability</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low input data</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>High resolution</td>
<td>High cost</td>
<td>A limited number of high-resolution images</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------</td>
<td>-----------</td>
<td>-------------------------------------------</td>
</tr>
<tr>
<td>Photography</td>
<td>Imaging of any target High resolution</td>
<td>Difficult to obtain an image from exact coordinates</td>
<td></td>
</tr>
<tr>
<td>Google Earth</td>
<td>Low cost Easy to use Can generate multiple views of the same target</td>
<td>Low resolution Content varies by location Not very accurate</td>
<td>Photo-realistic images of a large, stationary target</td>
</tr>
<tr>
<td>Microsoft Flight Simulator</td>
<td>Low cost High accuracy Fast generation of many images using TPL-WAY</td>
<td>Rendered images are not photorealistic</td>
<td>Many specific images; Training stimuli</td>
</tr>
</tbody>
</table>

**Conclusion**

Based on previous perceptual training research, using static oblique images can provide an effective approach for training visual perceptual tasks in UAV operation. By considering the requirements and limitations of a particular training program, developers can consider which implementation and outcome characteristics are most important to optimize training objectives.

Perceptual discrimination training represents an easily implemented, scalable method of training for UAV operators. Given the increased demand for effective training, alternatives to high fidelity simulation training will almost certainly be required. Although discrimination training represents only a portion of skills relevant to UAV operation, if implemented appropriately, it can improve operator performance by reducing UAV accidents and improving the ability of an operator to identify targets from an aerial viewpoint. This alone will reduce the number of civilian casualties in military engagements. These improvements in current UAV training will lead to an easier transition into civilian airspace.
References


