Determination of Frame Rate Requirements for Video-panorama-based Virtual Towers using Visual Discrimination of Deceleration during Simulated Aircraft Landing: alternative analysis

Stephen R. Ellis[1], Norbert Fürstenau[2] and Monika Mittendorf[2]

Keywords: virtual tower, video-panorama, framerate, visual discrimination, speed perception

Abstract

In order to determine the required visual frame rate for out-the-window video displays at remote/virtual airport towers, thirteen active air traffic controllers viewed high dynamic fidelity simulations of landing aircraft and decided whether aircraft would stop as if to be able to make a turnoff. The viewing conditions and simulation dynamics replicated visual rates and environments of transport aircraft landing at small commercial airports. The required frame rate was estimated from a model fit to the perceptual discriminability of the condition in which the aircraft would stop. Asymptotic performance appears to require rates from 30-80 fps. Analysis based on the sensitivity parameter A points to the lower end of this range, but definitive recommendations require further testing at frame rates of 45-60 fps. Errors appear due to illusory increase in speed at low frames rates.

Introduction

Recent proposals for new air traffic control have suggested that technology may remove the need for air traffic controllers to be present in airport towers. (JPDO, 2007). Controllers could therefore visually supervise aircraft from remote locations by videolinks, allowing them to monitor many airports from a central point (Hadden, Lee, Geyer, Sheridan, Francis,Woods, Malonson, 2008)). While many current towers, even some at busy airports like San Francisco, can continue to operate totally without controllers ever seeing controlled aircraft, it is clear from controller interviews that numerous out-the-window visual features are used for control purposes (Ellis & Liston, 2010; Van Schaik, Lindqvist & Roessingh, 2010; Ellis & Liston, 2011). In fact, these visual features go beyond those required by the FAA which typically only include requirements for aircraft detection, recognition, and identification. (FAA, 2006; Watson, Ramierez & Salud, 2009).

Potentially important additional visual features identified by controllers in interviews involve subtle aircraft motion. These could be degraded by low dynamic quality of remote visual displays of the airport environment. In fact, the dynamic visual requirements for many aerospace tasks have been studied, but most attention has been paid to pilot vision (e.g. Grunwald & Kohn, 1994)) with relatively little attention paid to the unique aspects of controller vision which, for example, emphasize relative motion cues. Consequently, there is a need to study some of these visual motion cues to understand how their use may be affected by degraded dynamic fidelity, e.g. low visual frame rates. Such low rates could be due to typically low rates of aircraft surveillance systems, e.g. 1-4 Hz, or to image processing loads arising from of the very high resolution, wide field of view video systems needed to support human vision in virtual towers.

Since preliminary investigation of the role of visual features in tower operations has shown that their principal function is to support *anticipated separation* by allowing controllers to predict future aircraft positions (Ellis & Liston, 2011), we have begun to investigate the effects of

[1] NASA Ames Research Center, Moffett Field, CA USA
[2] German Aerospace Center (DLR), Braunschweig, Germany

frame rates on the deceleration cues used to anticipate whether a landing aircraft will be able to brake on a runway, as if to make a turn off before the runway end.

Our specific hypothesis is that the disturbance due to low frame rate affects the immediate visual memory of image motion within the video frame. Memory processes classically have an exponential decay. Accordingly, one might expect discriminability of the visual motion associated with aircraft deceleration to reflect this feature, degrading only a bit for higher frame rates but more rapidly for the longer period, lower frame rate conditions. A possible descriptive function could be of the form: $1 - e^{-k/t}$. This kind of model captures the likely features that the rate of degradation of motion information increases with greater sample and hold delays but that there is also an upper asymptote corresponding to continuous viewing which is determined by the inherent task difficulty. Significantly, fitting such a model to the drop off in detection performance provides a theoretically based method to estimate that frame rate required to match visual performance out the tower window.

This frame rate for asymptotic visual performance is useful to know for design since it directly impacts the required communication bandwidth for remote or virtual towers that depend on relaying image information from the tower to the centralized control room. High bandwidth for the very high-resolution imagery required to match human visual resolution and luminance range could be an expensive and recurring cost of a remote tower. Prototype designs suggest it could easily exceed several 100 MB/sec! (Fürstenau et.al. 2009) A quantitative performance measure is needed to understand the trade-offs affecting user performance as cheaper, lower communication bandwidths are examined during design analysis. This is one goal of the following study.

Methods

Subjects

Thirteen active German tower controllers were recruited as volunteer subjects for the experiment. The participants' ages ranged from 25 – 59 yrs. and were divided into 3 experimental groups of 4, 4, 5.

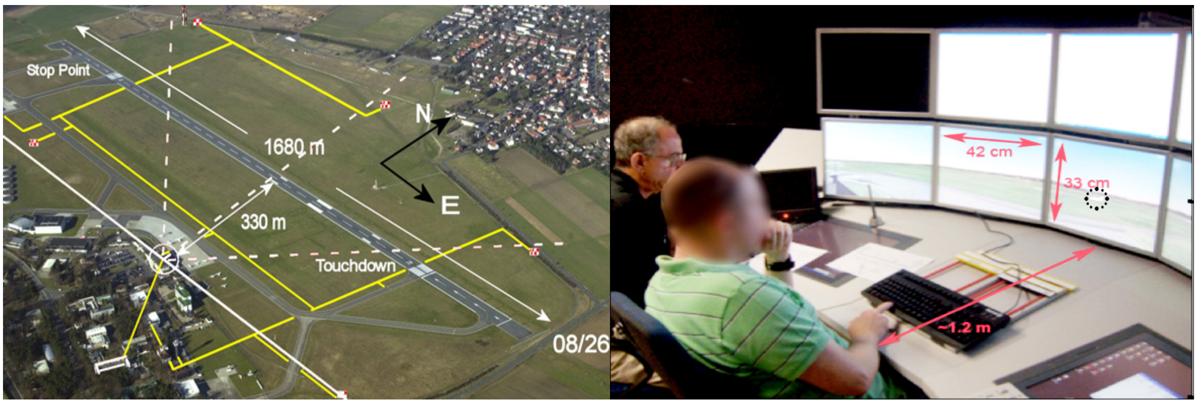

Figure 1: a) (left) Aerial view of Braunschweig airport showing the circled location of the simulated cameras, fields of view of the four cameras (radial sectors), and some dimensions and reference points. b) (right) Participant at a simulation console judging the outcome of a landing aircraft just after touchdown. The circle indicating touchdown has been added for clarity in this particular figure.

Apparatus

The experiment was conducted at a Remote Tower (RTO) console as part of the DLR Apron-and-Tower Simulator of the Braunschweig DLR facility. This simulation system was used to generate 60 landings of a lightly loaded A319 transport at the Braunschweig airport (Fig.1a).

The simulated aircraft would first appear on the right most monitor while in the air at 300 m altitude 32 sec before touch down (Fig.1b). Then it would fly to touch down seen on the next monitor to the left. Thereafter, it would either roll through to the end of the runway or stop 250 m before the runway end.

The simulator generated 60 1-minute landing scenarios with various dynamically realistic deceleration profiles and frame rates of either 6, 12, and 24 fps emulating the video signals potentially coming from cameras mounted near the Braunschweig tower. These files were then used in turn as input simulating the actual cameras so the participants could use the video console as if it were connected to actual cameras on the airfield. They present approximately a 180° view as seen from airport tower but compress it to an approximately 120°. An upper array of tiled monitors was present but not used during the testing.

Experimental Design and Task

The three matched subject groups were used in an independent groups, randomized block design in which three different landing deceleration profiles were used to produce 60 landings to the west on the Braunschweig airport's Runway 26. Each group was assigned to one of the three video frame rate conditions. The approaches were all equivalent nominal approaches for an A319 aircraft but varied in the amount of deceleration after touchdown.

The participants' task was to report as soon as possible whether the landing aircraft would stop before the end of the runway. In all cases they were then allowed to watch the actual outcome and use a certainty level compatible with actual operations. The three different deceleration profiles were randomized to produce a sequence of 30 landings in 3 blocks of 10. The three blocks were repeated once to provide the 60 landings in the experimental phase used for each of the independent groups. The experimental phase was preceded by a training phase during which the subjects were given familiarity practice with 20 landings similar to those used experimentally. This approach gave participants a chance to learn the task and adapt to a head mounted video-based eye tracker that they wore during the experiment[1]. Including instructions, the experiment required 1.5-2 hr per subject.

Further details regarding the subjects and experiment may be found in Ellis et.al. (2011).

Results

Statistical analyses of errors, decision times, and subjective judgment certainty have been reported earlier in Ellis et.al. (2011). The present report will focus on a not previously considered comparison of d' and A, two alternative parameters used to estimate user sensitivity of detection that an aircraft will stop. d' is the usual measure for this purpose but A has recently been suggested as an improved "nonparametric" alternative which requires fewer statistical assumptions (Zhang and Mueller, 2005).

For Figure 2b) the uncertainties of the estimated exponential model parameters $d'_{max} = 3.28$ (= asymptote, ± 1.7, 95% confidence) and frame-rate-constant = 1/0.052 = 19.4 (± 17) were obtained with Matlab "nlpredci" which uses the results of "nlinfit". An envelope of 95% of possible results are shown as a shaded region based on a Monte Carlo simulation using the measured standard errors of the means for each group. (see Ellis et.al. 2011).

For Figure 3a) the uncertainties of the estimated exponential model alternative parameters A_{max} and the frame-rate-constant were obtained with the same software. Since the A parameter does not require the usual assumptions of Signal Detection Theory (SDT), i.e., statistical independence of signal and noise, normality of both the signal and signal + noise distributions, it may be considered to provide a better estimate of the frame rate at which participants' perfor-

[1] These eye movements have not been analyzed yet and will not be discussed in this report.

mance asymptotes. From Figure 3a) this value seems to be about 40 fps, a value close to that estimated from extrapolation of errors as reported in Ellis et.al. (2011). The corresponding asymptote from the d' analysis could be as high as 80 fps

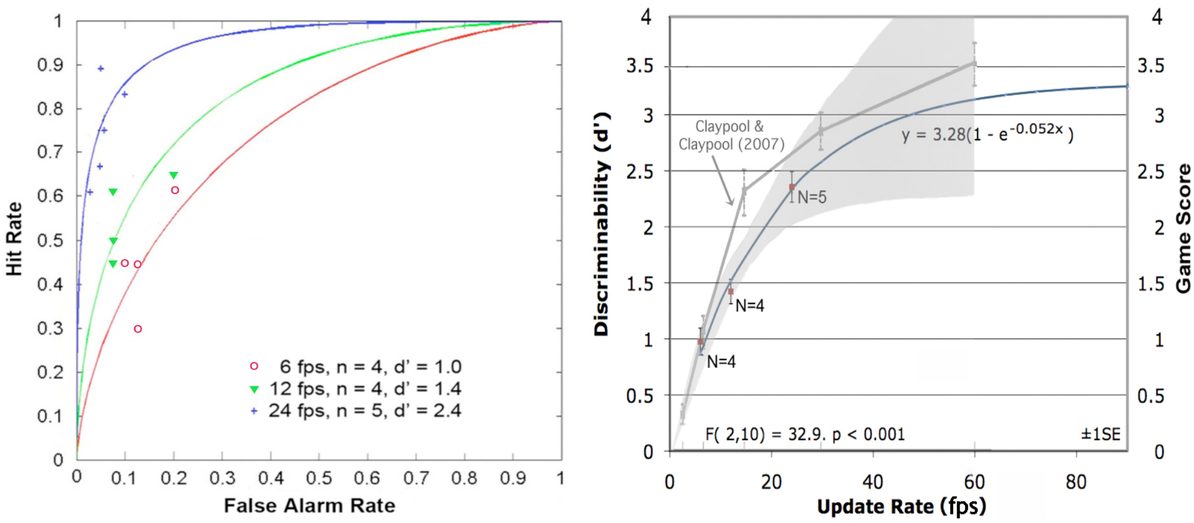

Figure 2: a) (left) ROC curves for each of the groups using different frame rates show that SDT analysis nicely separates the groups with different frame rates and that controllers were biased towards low false alarm rates, i.e. they were biased against falsely reporting that an aircraft would stop on the runway. b) (right) Regression model for d' (darkest solid trace) of the discriminability of landings with stopping aircraft. The lighter grey trace plots comparative data from Claypool & Claypool (2007). Shaded area shows the 95% regression confidence band based on Monte Carlo simulation.

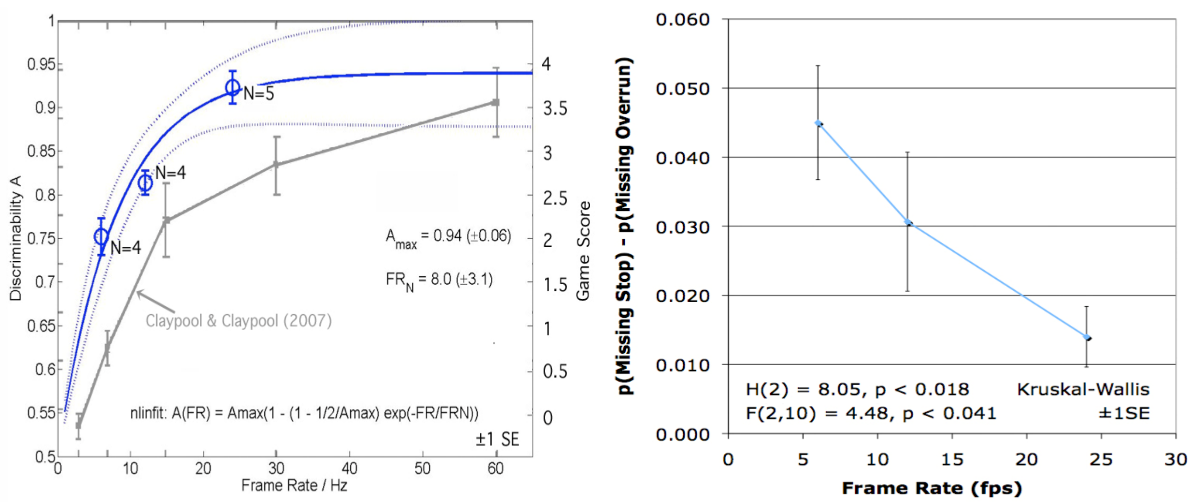

Figure 3: a) (left) Regression model for A (darkest solid trace) of the discriminability of landings with stopping aircraft. 95% regression confidence intervals flanks the model fit. Lighter grey trace shows comparative data from Claypool & Claypool (2007) . b) (right) Bias towards reporting a runway overrun increases the likelihood of missing a planned stop over missing a planned overrun.

Discussion

The principal result of the study, shown in Figures 2a, 2b, and 3a) suggests that relatively high frame rates will be required for imagery in virtual or remote towers if controllers working in them are expected to perform the kinds of subtle visual motion discrimination currently made in physical towers. The usefulness of the signal detection approach to estimate group-wide d' can be seen in Figure 2a which shows clear experimental group separation. Figures 2b and 3a show our model fit to estimate asymptotic performance. They also include for reference a replotted result from Claypool & Claypool (2007) examining the effect of change in frame rate on video game shooting score. Their overlaid data empirically support our theoretical supposition that the users' performance at higher and higher frame rates may be modeled by an exponentially approached limit. It is especially interesting that their report of the effect of frame rate on video game score in a first-person-shooter game resembles our results since their task and response measure was so different.

In contrast to their count of shots on target, our analysis of d' is particularly useful since it can be argued to be bias-free, independent of user criteria and primarily a function of the task requirements and perceptual estimation noise. It can additionally be cross checked with extrapolation of the error data (see Ellis et.al. 2011). But an extrapolation for errors is harder to justify theoretically without a computational error model.

We are generally happy with the level of difficulty of the task as we were able to avoid basement or ceiling effects that could have prevented a useful comparison of frame rates. The roll off of d' at a level of about 3 shows that the task difficulty was set at a reasonable level for controllers not very familiar with the airport and who are monitoring it using a virtual tower. One would expect controllers familiar with the airport working in its tower could operate with a d' of 4 or higher because of extensive safety training. But it is, nevertheless, clear at least one more data point at a higher frame rate is needed to confirm the suggested frame rate requirement.

Interestingly, during debriefings after the experiment, subjects in the lower two frame rate groups reported that they felt that the aircraft were moving "too fast" and that it was this extra apparent speed that made discrimination hard. By "too fast" the controllers meant to refer to the apparent ground speed of the aircraft compared to what they would expect to see from a tower.

We can examine this possibility by looking at a response bias that could arise from aircraft appearing to move "too fast." Such a bias would lead subjects to underestimate whether an aircraft actually coming to a stop would in fact stop. This underestimation would occur because the aircraft would seem to be going too fast. Aircraft that are in fact not stopping would not be subject to a bias since they would merely seem to be simply more likely to overshoot the end of the runway. This outcome is one which would occur in any case.

Thus, we would expect subjects to be more likely to incorrectly identify a stopping aircraft versus one that is not stopping as a function of frame rate. Indeed, when we compare the likelihood of erroneously identifying a stop versus that of erroneously identifying an overshoot, all 13 subjects showed this bias. (sign-test, $p < 0.001$). This general bias towards identifying an aircraft as not stopping, however, is not surprising since approximately twice as many aircraft observed in fact do not stop versus those that do. Subjects quickly sense this bias during the experiment. What is interesting, however, is that the bias *is* a decreasing function of the frame rate. (Figure 3b). This effect confirms several anecdotal observations from the participants in the low frame rate conditions and suggests the need for counter measures, perhaps temporal filtering to smooth out the discontinuities. Such an approach would undoubtedly benefit from a computational model of speed perception. One input for such modeling of the speed perception error could be the spatio-temporal aliasing artifacts that introduce higher temporal frequency information into a user's "window of visibility" (Watson, Ahumada & Farrell, 1986).

In closing it is useful to note that degraded performance due to reduced visual frame rates affects interfaces in other types of aerospace user environments, such as that of UAV sensor suite operators where low frame rates (< 12 fps) are also often encountered. Consequently, the current findings could be useful outside of air traffic control (see Kempster, 2000).

Acknowledgement

We wish to thank DLR personnel F. Morlang, M. Schmidt, T. Schindler for technical assistance operating the DLR Apron and Tower simulator (ATS) and the preparation of video files. A. Papenfuss and C. Möhlenbrink provided key assistance throughout the course of actual experimentation.

Literature

Claypool, K. T. and Claypool, (2007). *On frame rate and player performance in first person shooter games.* Multimedia Systems, 13, 3–17.

Ellis, S. R. and Liston, D. B. (2011). *Static and motion-based visual features used by airport tower controllers.* NASA TM-2011-216427, Ames Research Center, Moffett Field, CA.

Ellis, S. R., Fürstenau. N. and Mittendorf M.(2011). *Frame Rate Effects on Visual Discrimination of Landing Aircraft Deceleration: Implications for Virtual Tower Design and Speed Perception.* Proceedings of the Human Factors and Ergonomics Society, Sept. 19-23, 2011, Las Vegas, NV USA.

Ellis, S. R. and Liston, D. B. (2010). *Visual features involving motion seen from airport control towers.* Proc.11th IFAC/-IFIP/IFORS/IEA Symposium. on Analysis, Design, and Evaluation of Human-Machine Systems 9/31-10/3, 2010, Valenciennes, France.

Fürstenau. N., Schmidt. M., Rudolph, M., Möhlenbrink, M., Papenfuß, A., and Kaltenhäuser, S. (2009). *Steps Towards the Virtual Tower: Remote Airport Traffic Control Center (RaiCe).* Proceedings of the ENRI International Workshop on ATM/CNS. Tokyo, Japan, 67-76.

Grunwald, A. J. and Kohn, S. (1994). *Visual field information in low-altitude visual flight by line-of-sight slaved head mounted displays*. IEEE Systems Man and Cybernetics, *24*, 1, Jan. 1994, 120-134.

Hannon, D., Lee, J., Geyer, T., M., Sheridan, T., Francis, M., Woods, S., Malonson, M. (2008). *Feasibility evaluation of a staffed virtual tower.* The Journal of Air Traffic Control, Winter 2008, 27-39.

JPDO (Joint Planning and Development Office) (2007). *Concept of operations for the Next Generation Air Transportation System Version 2.0*, 13 June 2007, pp. 2-27--2-36.

Kempster, K. A. (2000). *Frame rate effect on human spatial interpretation of visual intelligence.* Masters Thesis, NPS, Monterrey CA.

Van Schaik, F.J., Lindqvist, G., and Roessingh, H.J.M. (2010). *Assessment of visual cues by tower controllers.* Proc. 11th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design, and Evaluation of Human-Machine Systems August 31 - September 3, 2010, Valenciennes, France.

Watson, A.B. , Ahumada Jr., A.J. and Farrell, J. (1986). *Window of visibility: Psychophysical theory of fidelity in time-sampled visual motion displays*. Journal of the Optical Society of America A, 3, pp. 300–307.

Watson, A.B., Ramirez, C.V. and Salud, E. (2009). *Predicting visibility of aircraft.* PLoS ONE. 2009; 4(5): e5594. Published online 2009 May 20. doi: 10.13.

Zhang, J. & Mueller, S. T. (2005). *A note on ROC analysis and nonparametric estimate of sensitivity*. Psychometrika, 70, (1) 203-212.