Videopanorama Frame Rate Requirements Derived from Visual Discrimination of Deceleration during Simulated Aircraft Landing

Norbert Fürstenau and Stephen R. Ellis[1)]

German Aerospace Center (DLR), Institute of Flight Guidance, Braunschweig, Germany
norbert.fuerstenau@dlr.de
1)Ames Research Center, NASA, Moffett Field/Ca, USA
Stephen.R.Ellis@nasa.gov

Abstract In order to determine the required visual frame rate (FR) for minimizing prediction errors with 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 or whether a runway excursion would be expected. 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 using Bayes inference on prediction errors by linear FR-extrapolation of event probabilities conditional on predictions (stop, no-stop). Furthermore estimates were obtained from exponential model fits to the parametric and non-parametric perceptual discriminabilities d' and A (average area under ROC-curves) as dependent on FR. Decision errors are biased towards preference of overshoot and appear due to illusionary increase in speed at low frames rates. Both Bayes and A - extrapolations yield a framerate requirement of $35 < FR_{min} < 40$ Hz. When comparing with published results [12] on shooter game scores the model based d'(FR)-extrapolation exhibits the best agreement and indicates even higher $FR_{min} > 40$ Hz for minimizing decision errors. Definitive recommendations require further experiments with FR > 30 Hz.

1. Introduction

This chapter reviews a two-alternative decision experiment with simulated aircraft landing as dependent on video-framerate (FR) characteristics with the goal of determining the minimum framerate necessary for minimizing decision errors under Remote Tower working conditions. It collects results partially presented in previous publications [1][2][20] (see Chapters 8, 9, 12 of this book).
Recent proposals for decreasing cost of air-traffic control at small low-traffic airports have suggested that technology may remove the need for local control tow-

ers. Controllers could visually supervise aircraft from remote locations by video-links, allowing them to monitor many airports from a central point [3][4][14][15]. While many current towers on A-SMGCS-equipped airports, even some at busy airports like London-Heathrow, can continue to operate totally without controllers ever seeing controlled aircraft under contingency conditions, although with reduced capacity, it is clear from controller interviews that usually numerous out-the-window visual features are used for control purposes [5][6][7]. In fact, these visual features go beyond those required for aircraft detection, recognition, and identification [8].

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 and armed forces tasks have been studied, but most attention has been paid to pilot vision (e.g. [9]) and military tactical information transmission (e.g. [10]). Relatively little attention was 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 the very high resolution, wide field of view video systems needed to support human vision in virtual towers (see Chapters 8, 9).

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 [5] we have begun to investigate the effects of 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: $1 - \exp(-k/T)$. This kind of model captures the likely features that the rate of degradation of motion information increases with greater sample and hold delays T but that there is also an upper asymptote of discriminability 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.

We used two statistical analysis methods for deriving model based framerate requirement estimates via discriminability measurement: Bayes inference and signal detection theory (SDT) with parametric (ROC-isosensitivity-curve index d') as

well as non-parametric discriminability (A = average area under all proper ROC-curves). Bayes inference allows for concluding from the measured error probability conditional on the perceived world state, on the probability of this (unexpected) situation conditional on the measurement (see Appendix A2). Measuring these probabilities with different values of the independent variable (i.e. the framerate FR) allows for extrapolation to minimum FR for zero error probability. SDT as an alternative method has the advantage of separating the intrinsic subjective preference (tendency for more liberal or conserative (error avoidance) decisions) by simultaneously separating through the measurement of hit and false alarm rates (= probabilities conditional on the alternative experimental situations) from the decision criterion (or subjective decision bias) index c (for d') and b (for A) respectively).
Experimental Methods and results are provided in sections 2, 3. In section 4 the two alternative methods (Bayes inference and detection theory) are used for deriving from the measured response matrices the Bayes inference on risk of unexpected world state, and estimates of discriminabilities and decision criteria d', c and A, b respectively. These in turn are used to provide minimum framerate estimates for maximizing d' and A, and minimizing prediction error risk. We finish with a conclusion and outlook in section 5.

2. 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. Controllers from small, medium, and large German airports were approximately evenly distributed to the groups.

Apparatus

The experiment was conducted at a Remote Tower (RTO) videopanorama-console as part of the DLR Apron-and-Tower Simulator (ATS) of the Braunschweig DLR facility. This simulation system was used to generate 60 landings of a lightly loaded A319 transport at the Braunschweig airport with a 1680 m runway 08/86 (Figure 1, RWY was extended to 2500 m after this experiment). The simulated aircraft would first appear from E on the right most monitor while in the air at 300 m altitude 32 sec before touch down (Figure 2). 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.

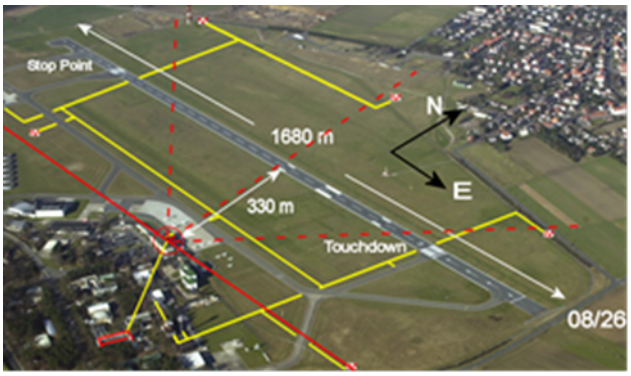

Figure 1. **Aerial view of Braunschweig airport showing the circled location of the simulated (and real) cameras, fields of view of the four cameras (radial sectors), and some dimensions and reference points.**

The simulator generated 60 1-minute landing scenarios with various dynamically realistic deceleration profiles of nominally 1, 2, or 3 m/s^2 maximum (initial) braking and frame rates of either 6, 12, and 24 fps emulating the video signals potentially coming from cameras mounted near the Braunschweig tower. Only the highest deceleration (3 m/s^2) was sufficient to cause the aircraft to stop near the stopping point (Figure 1) before the end of the runway (leftmost monitor in Figure 2).The video 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°. Viewing distance between operators and monitors (21" UXGA: 1600 x 1200 pixels with 4/3 format: 42 x 33 cm, luminousity sufficient for photopic office environment) was ca. 120 cm. An upper array of tiled monitors for a second airport was present but not used during the testing.

Figure 2. **Participant at a simulation console judging the outcome of a landing aircraft just after touchdown (2nd monitor from left). Approach on the rightmost monitor, touchdown is on the left side of second monitor from the right. Reconstructed panorama compressing the 180°-tower view to ca. 120° for subjects at the RTO-console.**

Experimental Design and Task

The three matched subject groups were used in an independent groups, randomized block design in which the 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 equation of motion used for the post-processing of logged simulation data assumed that the only braking force (deceleration) after touchdown is given by:

$$\ddot{x} = -b_{min} - (b_0 - b_{min})e^{-t/\tau} \qquad (1)$$

with $d^2x/dt^2(t=0) = -b_0$, i.e. braking acceleration is assumed to consist of a constant and an exponentially decreasing part. Of course this is a strongly simplified model which neglects e.g. friction and different external forces like braking via the weels and reverse thrust. Parameter values as obtained from exponential fits to the logged simulation data are listed in Table 1. Also listed are the stop times t_{stop} = t(v = 0), v(t=0) = v_0 = 70 m/s and positions x_{Stop} as calculated from the solution to (1). The table verifies that only the highest nominal deceleration avoids runway excursion (stop for x < ca. 1500 m).

Table 1. **Deceleration Profiles by fitting equation (1) to logged deceleration data.**

	Landing Braking Parameters		
Nominal value m/s^2	1.0	2.0	3.0
b_0 / m/s^2	1.33	1.76	2.39
b_{min} / m/s^2	0.45	1.01	1.64
τ / s	41.3	22.0	12.0
t_{stop} / s	85.1	54.4	37.4
x_{stop} / m	2544	1748	1238

Braking acceleration profiles (decelerations) according to the equation of motion (1) with parameters in Table 1 are shown in Figure 3. Calculations refer to runway coordinates with x || RWY, rotated by +4.1 degree with regard to (E, N, up)-coordinates; x = 0 at ARP. Touchdown is at x = + 520 m. Closest distance from observation point to runway is d_{TWR} = 330 m at x = +245 m

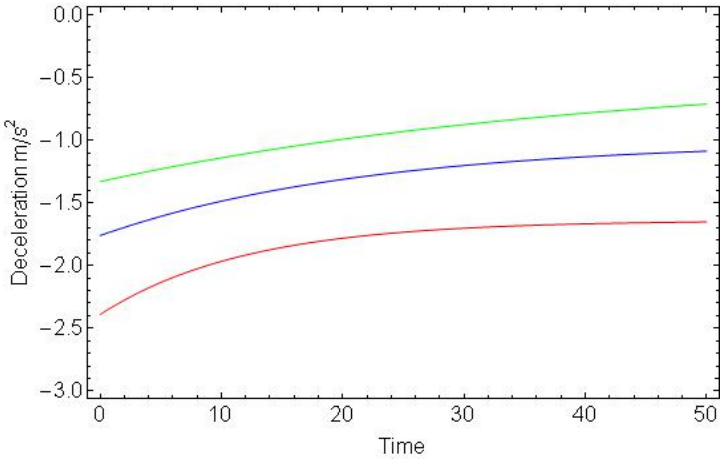

Figure 3. **Deceleration profiles (= decrease of braking acceleration) as obtained by fitting logged simulator data using equ. (1) for the three nominal braking values 1, 2, 3 m/s^2.**

The participants' task was to report as soon as possible whether the landing aircraft would stop before the end of the runway (stop event S2 (high deceleration), no-stop event S1 (runway excursion due to low deceleration)), with response time measured by pressing the space bar. 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.

In addition to the objective data, we recorded participants' subjective certainty regarding each of their decisions on a 0-3 Likert-like scale presented after each landing (0-total guess, 3-total certainty).

3. Results

Errors, reaction times and estimates of judgment certainty were subjected to planned Two-Way independent groups ANOVA's based on a mixed design with Subjects nested within Update rate condition but crossed with Repetition which was quantized into 8 Experimental Blocks of 10 landings each, the period of randomization of the deceleration condition. Decision errors appeared to show a learning effect as can be seen in Figure 4.

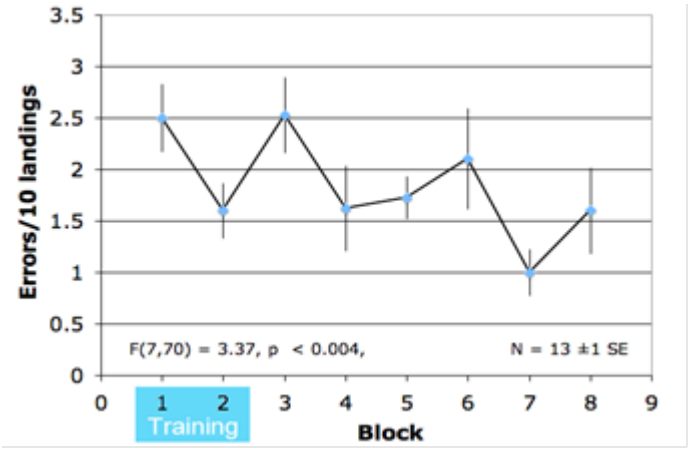

Figure 4. Error Rate as a function of repetition Block

But once the training blocks were removed and the remaining blocks grouped into two categories First three (3,4,5) and last three (5,6,7) the statistically significant effect proved unreliable and disappeared (F(1,10)=1.52, ns).

[1] eye movements will not be discussed in this Chapter. For analysis of eye movements see Chapter 5 and references therein.

Response Times

Figure 5 shows the measured response times plotted into a graphic of the airport layout, as measured by participants pushing of the keyboard space-bar at the operator console (see Figure 2). The space bar pressing with yes-answer (=stop predicted) or no-answer (= overshoot predicted) occurs typically at RT = 10 – 11 s after observed touchdown. RT corresponds to A/C positions between 700 and 900 m behind the threshold.

We achieved the goal of approximately equal response times in the different Frame Rate conditions ($F(2,8) = 0.864$, ns). Response times after training remained approximately constant across Blocks with a statistically significant variation ($F(5,40) = 3.91$, $p < 0.006$) of less than ±2.5% when the training blocks were excluded.

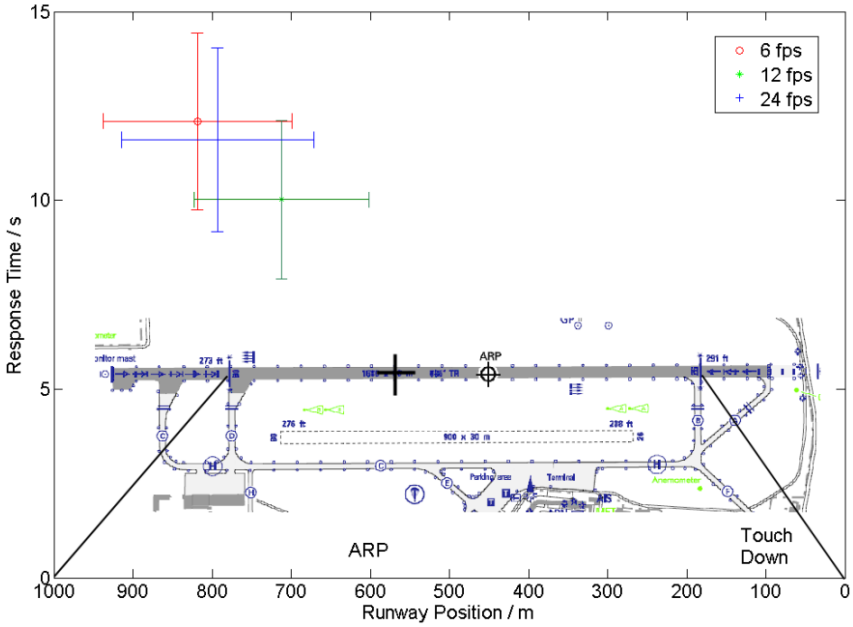

Figure 5. **Airport layout (inset projected to abszissa via solid black lines) with response times (ordinate) typically 10 - 11 s after touchdown, and with A/C typically around 800 behind threshold (black cross), separated for the three framerates and averaged over all landings (decelerations) and participants. ARP = Airport reference point at 600 m.**

Decision Statistics: Response Matrix

The experimental results of this two-alternative decision experiment concerning decision errors as dependent on video framerate are summarized in the stimulus-

response matrices of Table 2. It shows group averages of measured probability estimates, with standard errors of mean (), of correct rejection C = P(no|S1), false alarm FA = P(yes|S1), miss M = P(no|S2), and hit H = P(yes|S2). S1 = stimulus with runway excursion, S2 = stimulus with stop on the runway, yes = stop predicted (high deceleration perceived), no = no stop predicted (low deceleration perceived). Probabilities in horizontal rows (constant stimulus) sum up to 1.

Table 2. **Response matrices (measured H, M; C, FA rates) for the three framerates.**

Alternative Stimuli	Response for 3 Video Framerates: Probability Estimates				
	No-stop predicted			Stop predicted	
Low Deceleration. No-stop Stimulus S1. n(S1) = 40	p(no|S1) = C	6	0.86 (0.02)	p(yes|S1) = FA	0.14 (0.02)
		12	0.89 (0.03)		0.11 (0.03)
		24	0.94 (0.01)		0.06 (0.01)
High Deceleration. Stop Stimulus S2 n(S2) = 20	p(no|S2) = M	6	0.55 (0.06)	p(yes|S2) = H	0.45 (0.06)
		12	0.45 (0.05)		0.55 (0.05)
		24	0.22 (0.07)		0.78 (0.07)

These results may be presented in the form of Venn-diagrams as depicted in Figure 6, that clarifies the character of the measured rates H, M, CR, FA as conditional probabilities and their base sets with regard to situations (world states) S1 = no stop and S2 = stop event.

The different areas (width) of the two columns representing situations (or alternatives) S1, S2 reflect different numbers of experimental no-stop (n(S1)) and stop rates (n(S2)) respectively to be observed by the subjects, and of corresponding a-priori probabilities p(S1), p(S2): n(S1) + n(S2) = 60 with n(S2)/n(S1) = ½ (see also .

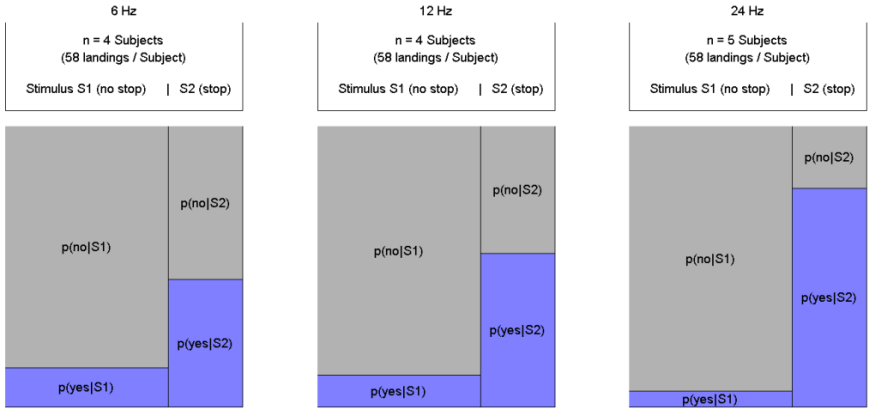

Figure 6. **Venn diagrams representing measured rates of correct (H = p(y|S2), CR = p(no|S1)) and false decisions (M = p(no|S2), FA = p(yes|S1)) for the two given world states (situations, events) S1 (= no stop on RWY, insufficient braking braking, alternative 1 or "noise", in terms of SDT, see below) and S2 (stop on RWY, sufficient braking, alternative 2 or "signal + noise", in terms of SDT).**

As a preliminary analysis of the results Figure 7 does show a significant effect of frame rate on the average error numbers per 10 landings and invites discussion. Extrapolation indicates a minimum framerate > 30 Hz for minimizing decision errors.

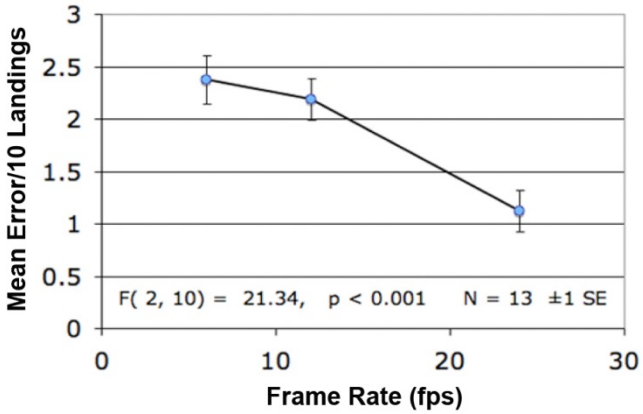

Figure 7. **Error Rate as a function of Frame Rate**

Also it can be seen in Table 2 that like in the averaged error plot of Figure 7 the measured probability estimates indicate a trend dependent on framerate (FR): the hit rate H = p(yes|S2) increases with framerate whereas the false alarm rate FA = p(yes|S1) decreases. We will show in the following data analysis and discussion section how the measured probabilities in the response matrix can be used for deriving a (Bayes) inference on risk probabilities for safety critical decisions, de-

pendent on the video framerate as system parameter (risk for a world state different from the predicted event, i.e. risk of surprise situation) by using the a priori knowledge on relative frequencies of the planned experimental situation alternatives S1, S2.

Besides the Bayes inference the conditional probabilitities of the detailed response matrix (Table 2, Figure 6) will be used to derive a theoretically grounded data analysis for narrowing down the quantitative framerate requirements. Specifically the measured estimates of response probabilities conditional on the priori knowledge of experimental conditions (p(S1), p(S2)), suggests the use of signal detection theory (SDT) to derive a quantification of the detection sensitivity (discriminability) as the basis for estimating FR_{min}. This SDT-discriminability is free of a subjective criterion, i.e. free of a tendency towards more conservative (avoiding false alarms) or more liberal (avoiding misses) decision. For extrapolating towards a minimum required framrate we will provide an initial hypothesis of a perceptual model to be used for fitting our data. A model based data analysis would also provides guidelines for future experiments with the potential to generate further evidence supporting the conclusion.

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

We examined 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, because it would seem to be going too fast. Aircraft in fact not stopping would not be subject to a bias since they would merely seem to be overshooting the end of the runway in any case. Thus, we would expect subjects to be more likely to incorrectly identify a stopping aircraft (S2) as non-stopping versus one that is not stopping (S1) as stopping. Details of this analysis are also presented in the following discussion (section 4)

4. Data Analysis and Discussion

The present analysis will start with the simulation results of the movement/braking dynamics as obtained by integration of equ. (1) using the parameter values of Table 1 with decelleration profiles of Figure 3. It provides an impression of the requirements on perceptual discrimination during the experiments. The second subsection provides derivation of the Bayes-inference on risk of unexpected world states by using likelihood values based on the response matrix of Table 2. The

Bayes risks in turn are used for estimating via linear regression the minimum frame rate requirement that minimizes the risk of predicting the false world state. This result will be compared to the frame rate extrapolations of maximum discriminability based on a hypothetized exponential discriminability decrease as obtained obtained from sensitivity index d' and nonparametric discriminability A (= average area under the ROC-curves). Also the associated response bias will be discussed in more detail.

Simulation of Movement after Touchdown

The integration of the simplified equation of motion (1) for the braking dynamics with accelerations shown in Figure 3 yields the observed angular movement at the simulated control tower / camera position after transformation into the corresponding reference frame. The result for the velocity dependence on runway position before the transformation is shown in Figure 8.

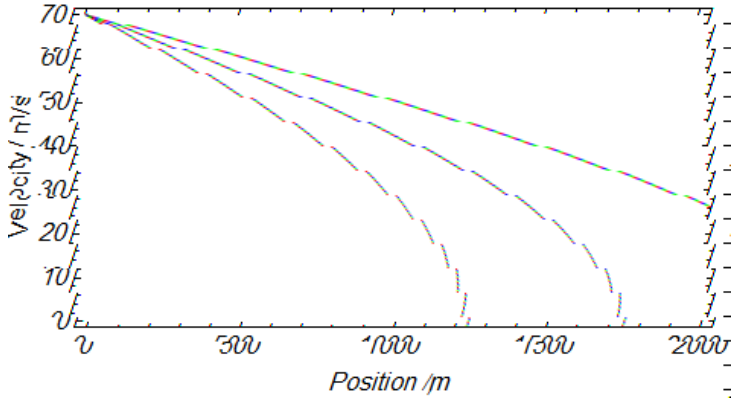

Figure 8. **Phase or state space diagram depicting simulated velocity (integration of equation of movement (1)) vs. position.**

This phase- (ore state-) space diagram velocity v(position x) confirms that in fact only the highest deceleration value (red line) leads to a stop at 1200 m) before the runway end (at 1650 m). The medium braking results in a slight overshoot whereas the lowest deceleration leads to a dramatic runway excursion. The following Figure 9 shows how this result translates into the viewing angle coordinates of an observer at the tower position.

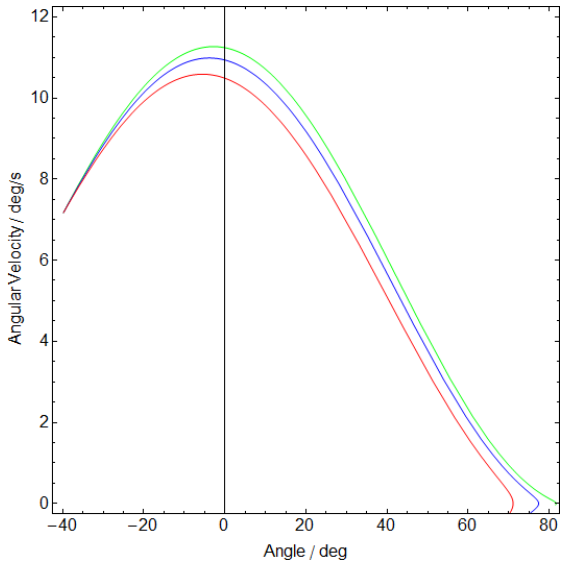

Figure 9: **Simulated angular velocity vs. observation angle phase space after transformation of integrated equation of movement into observer coordinates at tower position. Highest angular speed near the normal from TWR to the RWY. R = 10 – 11 s is at 44 - 48 deg.**

The participants prediction about stop/no stop or sufficient/insufficient braking is done some time after passing the 0-angle point at ca. 44 – 48 degrees, corresponding to the 10 – 11 s response time and 700 – 900 m distance from touchdown. In fact the decision seems to depend on subtle differences between trajectories in angular state space at decision time considering the fact that the real 180°-panorama view is compressed to ca 120° in the RTO-CWP panorama reconstruction. It was unclear during the preparation phase of the experiment if these small differences were large enough for discriminating at all between sufficient (stop event) and insuffient braking (no-stop event).

Bayes Inference: Risk of unexpected world state

The Bayes inference probabilities, with standard errors of mean (), about unexpected event S1 (runway excursion with predicted stop) and unexpected situation S2 (stop occuring no stop predicted) as calculated via Bayes law using the measured likelihoods (yes, or no predictions conditional on situations S1 and S2 respectively) are summarized in Table 3. Here the probabilities (for the same FR) of the columns add to 1.

Table 3. **Bayes Inference matrix for probabilities of actual world states (situations) conditional on decisions based on perceived evidence (likelihood x a priori knowledge).**

Event Alternatives	Bayes Inference on Event Probabilities conditional on Prediction				
	No stop predicted (no-response)			Stop predicted (yes-response)	
Low Deceleration	p(S1|no)	6	0.78 (0.02)	p(S1|yes)	0.40 (0.03)
		12	0.81 (0.02)		0.30 (0.04)
No stop event S1		24	0.91 (0.02)		0.13 (0.02)
High Deceleration	p(S2|no)	6	0.22 (0.02)	p(S2|yes)	0.60 (0.03)
		12	0.19 (0.02)		0.70 (0.04)
Stop event S2		24	0.09 (0.02)		0.87 (0.02)

The runway overshoot probability conditional on stop predicted (Bayes inference on the probability of world state S1 different from prediction "stop" based on perceived evidence) is given by

$$p(S1|\ yes) = p(yes|S1)\ p(S1) / p(yes) \quad (2)$$

with a priori knowledge of no-stop stimulus probability $p(S1) = n(S1) /(n(S1) + n(S2)$, according to the ratio of the Venn diagram areas and $p(yes) = p(yes|S1)p(S1) + p(yes|S2)p(S2)$. Equation (2) quantifies the risk of an overshoot occurring when predicting a stop, i.e. a surprising unexpected world state. It is proportional to the likelihood of missing a planned overrun $p(yes|S1)/p(yes)$ (for a brief introduction on Bayes inference and references see Appendix 2).

Figure 10 depicts the Bayes probability estimates for unexpected (surprise) world states dependent on framerate, i.e. a) unexpected runway excursion (S1) conditional on erroneous perception of a high braking deceleration (answer "yes": stop predicted) and b) the probability $p(S2|no\text{-}stop) = p(n|S2)\ p(S2)/p(n)$, that an unexpected stop occurs when predicting no-stop. Both surprise events suggest a linear fit to the three framerate data as most simple model. As expected the p(S1|yes)-graph (upper three data points) shows that for decreasing frame rates (FR -> 0) the conditional probability for a runway excursion occuring when a stop is predicted rises to chance (0.48 ± 0.01).

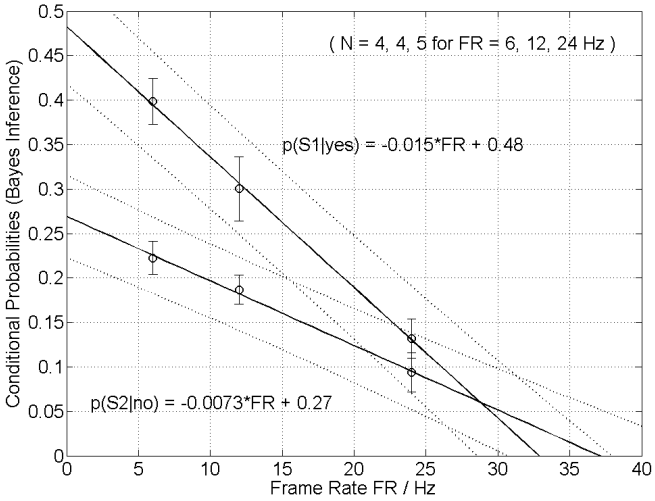

Figure 10. **Bayes inference for the three framerates (Abscissa) on probability of a) (upper data points and fit) unexpected situation S1 "a/c will not stop before RWY-end" (braking acceleration < threshold), given the alternative (false) stop-prediction, as calculated from measured likelihoods of subjects predicting "stop on RWY " conditional on S1 (= FA); and b) (lower data points and fit) of world state S2 "a/c will stop before RWY-end" (braking acceleration > threshold) as calculated from measured probabilities (likelihood) of subjects predicting "overshoot", conditional on S2 (a priori knowledge). Ordinate: mean (with stderr of mean) of probability for (unexpected) situation S_i conditional on prediction/decision d_i , averaged for all subjects within each FR-group. Straight line = linear fit with 95% confidence intervals (dotted).**

Comparing both graphs one immediately recognizes a bias of the lower one, with p(S2|no) -> 0.27 for FR -> 0 Hz, indicating a significantly reduced number of unexpected stop events conditional on the false "no" response, as would be expected by chance for lim FR → 0. As mentioned above the S2/S1 imbalance of 1/3 stop events and 2/3 no-stop partly explains this bias: the extrapolation to FR = 0 (no movement information available), yields p(S2|n) = 0.27 and p(S1|n) = 0.73 for the complimentary case so that for low FR with large position jumping p(S2|n)/p(S1|n) ≈ 0.4 reflects the S2/S1 imbalance of 1/2. The decrease of the p(S2|n)-bias and decision bias p(n|S2) (tendency for false overshoot prediction under S2) with increasing FR goes in parallel with the decreasing overall decision error. So the Bayes analysis confirms the previously reported decision bias [1][2] as quantified by M – FA = p(n|S2) – p(y|S1) which also decreases with increasing framerate (see Figure 11 below). Within the 95% confidence interval of the linear fit to the data also p(S2|no) predicts zero bias and 100% correct response for frame rates > 35 Hz, which is compatible with the FR-limit of zero-error prediction obtained with the "unexpected stop"- probability. The linear extrapolation of the Bayes analysis narrows the initial estimate of FR_{min} > 30 Hz as depicted in Figure 7, to ca. 30 – 45 Hz in Figure 10.

The hypothetical visual memory effect mentioned above would suggest an exponential approach to a minimum error probability with increasing FR instead of a linear behavior. The exponential fit to our data, however yields a significantly reduced goodness ($F = 140$, $p = 0.054$) as compared to the linear case ($F = 645$, $p = 0.025$), which demonstrates the necessity of experimental data at higher framerates.

The Bayes analysis also confirms the observation reported before in [1][2] (see also below, Figure 11) that the error bias appears exclusively connected with the preference of no-stop decisions, i.e. unexpected stop situations with a lower than chance error probability at FR = 0, because the false-stop prediction errors, as expected yield a chance Bayes probability $p(S1|yes) = 0.5$ for FR -> 0 (see Fig.10). The same is true for the complementary case $p(S2|yes)$. The observation of a significant bias of the unexpected-stop event inference ($p(S2|no)$ 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 starting point for such analysis of the speed perception error could be the spatio-temporal aliasing artifacts that introduce higher temporal frequency information into the moving images.

The measured probabilities of Table 2 used for calculating the Bayes inference are based on error statistics composed of intrinsic discriminability and subjective criteria, i.e. it includes a decision bias or subjective preference for positive or negative decisions. In what follows parametric and nonparametric variants of signal detection theory (SDT) are used for quantitatively separating both contributions and comparing the resulting FR_{min}-estimates with those of the Bayes inference.

Response Bias

From the results described above we would expect subjects to be more likely to incorrectly identify a stopping aircraft versus one that is not stopping. Indeeed when we compared the likelihood of erroneously identifying an overshoot versus that of erroneously identifying a stop (Table 2) $M - FA = p(n|S2) – p(y|S1)$, 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 ($p(S1) = 2\ p(S2)$) and subjects quickly sense this bias during the experiment. What is interesting, however, is that the bias is a decreasing function of the frame rate as depicted in Figure 11.

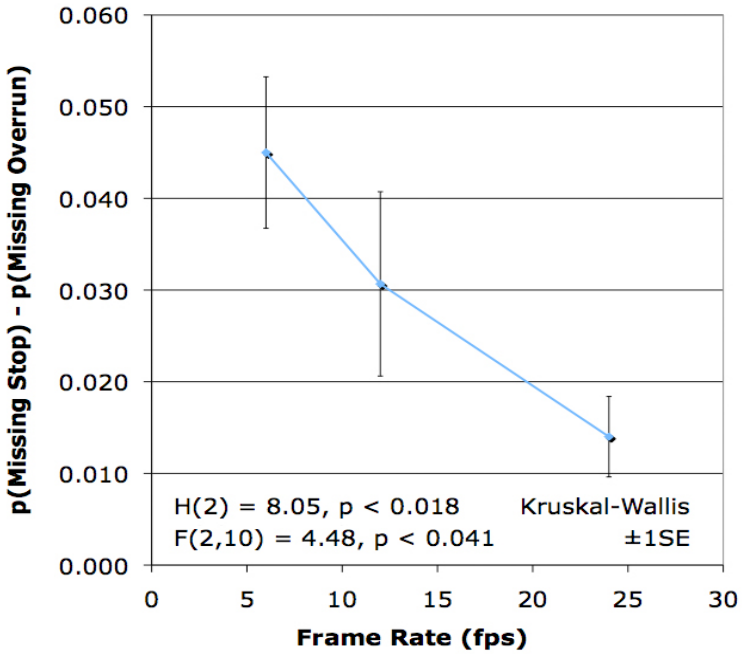

Figure 11. **Error bias (M – FA) towards reporting a runway overrun increases the likelihood of missing a planned stop over missing a planned overrun. Effect decreases with FR.**

The significance of this result, however needs support based on theoretical considerations and on alternative analysis. The detection bias is clearly reflected by the Bayes analysis as performed above (Figure 10). Like the error difference it exhibits a lower than chance probability for p(S2|no) with lim FR → 0, yielding p(S1|yes)/p(S2|no) ≈ ½, that reflects the p(S1)/p(S2)-ratio and like the above error difference converges to zero with increasing FR.

Of particular practical interest is the inferred risk of missing a high speed turnoff or of a runway excursion occurring when a stop is predicted, i.e. the conditional probability of overshoot p(S1|yes) (S1 = no stop event) due to low or abnormal braking when evidence suggests normal braking (stop prediction).

SDT Discriminability d' and Decision bias c

The principal result of data analysis using signal detection theory (SDT) is shown in Figure 12 and Figure 13. It confirms the Bayes analysis and suggests that relatively high update rates FR_{min} > 30 Hz 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. Figure 12depicts the experimental results of Table 1 in ROC-space (receiver operating

characteristics) H vs. FA. Plotted are the measured hit and false alarm rates for the 13 participants and the three framerates together with the respective averages (black crosses) and the ROC isosensitivity- and isobias-curves, parametrized by discriminability d' and criterion value c respectively. d' and c are calculated according to:

$$d' = 0.5\,(z(H) - z(FA)) \quad (3)$$

$$c = -\,(z(H) + z(FA)) \quad (4)$$

with z = z-score of cumulative Gaussian densities of the S1-, S2-familiarity distributions (see also Appendix A2).

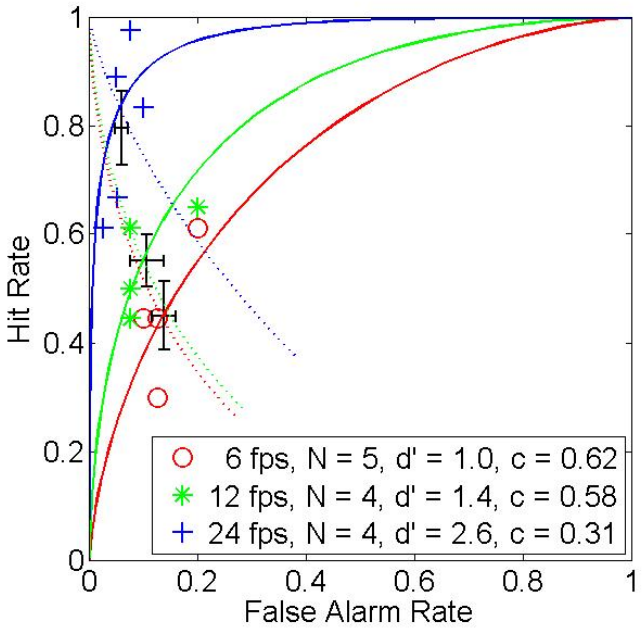

Figure 12. **ROC curve pairs parametrized (d', solid curves, c, dotted curves) for each of the three frame rates based on Hit and False Alarm rates for each subject. Crosses are the averages for each framerate subgroup of participants. ROC-curves d'(z(H), z(FA)) and c(z(H), z(FA)) are calculated with the d' and c subgroup-averages of the 13 participants.**

The positive criterion values indicate the controllers tendency to make conservative decisions, i.e. avoiding false alarms, incrasing misses and trying to be certain about their decisions, according to their work ethics and the written instructions of the experiment. The decrease of this effect is consistent with the decreasing error bias M – FA with increase of FR as reported above.

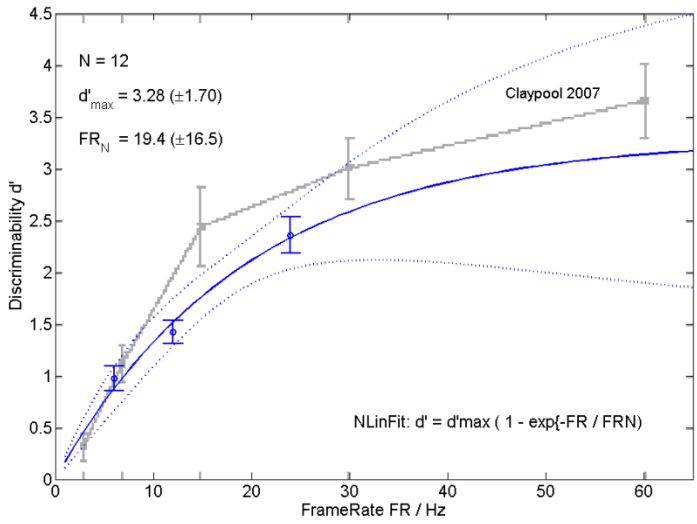

Figure 13. **Group averages (N = 12 subjects) ofxperimental discriminability values d' and exponential regression model (blue solid trace) for the stop/no-stop discriminability of landing aircraft. The lighter grey trace plots comparative data from Claypool & Claypool (2007). Dotted lines shows the 95% regression confidence range**

In Figure 13 we have also replotted a result from Claypool & Claypool (2007) examining the effect of change in frame rate on video game shooting score. These 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 certainly 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 particular, their use of shooting score does not capture the interplay of shooting frequency and hits in a way analogous to that of correct detections and false alarms in our experiment.

Our analysis of d' is in contrast to their count of shots on target and it 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 shown in Figure 4 and the Bayes inference in Figure 10, but this extrapolation for errors is harder to justify theoretically without a computational error model. A linear extrapolation which likely underestimates the value, however, suggests a ~40 fps would be needed for a vanishingly small error rate. Based on our exponential memory (sample-and-hold) decay hypothesis the asymptote of the d'(FR)-analysis, like the Claypool(2007) data indicates a higher FR_{min} value, more towards 60 Hz..

Nonparametric Discriminability A and Decision Bias b

Detectability A and likelihood bias parameter b were suggested as improved "nonparametric" alternatives of the conventional discriminability d' and criterion c because it requires fewer statistical assumptions (in its final form it was presented by Zhang and Mueller in 2005 [13]). In [2] we compared A with d' to estimate user sensitivity of detection that an aircraft will stop. Discriminability A and b are independent of the distributional assumptions required for deriving the conventional d' and c parameters for detectability and bias (see Appendix A2). The Zhang & Mueller formulas yield the average area A under all possible proper ROC curves (i.e. all concave curves within the range (0,0) – (1,1)) with non-increasing slope, obtained from the measured hit (H) and false alarm rates (FA). The constant A-isopleths cut the constant b-isopleths at the group mean (<FA>, <H>) coordinates which are used for calculating the A and b-ROC-curves: A:= A_{mean}(H,FA) and b:= b_{mean}(H,FA) for the three different framerate conditions according to the Zhang & Mueller equations (see Appendix 2).
Figure 14 shows the measured hit rates versus false alarm rates for all subjects together with their means (black crosses, as given in Table 1) and isopleths parametrized by constant discriminability A(FR) and constant decision bias b(FR).

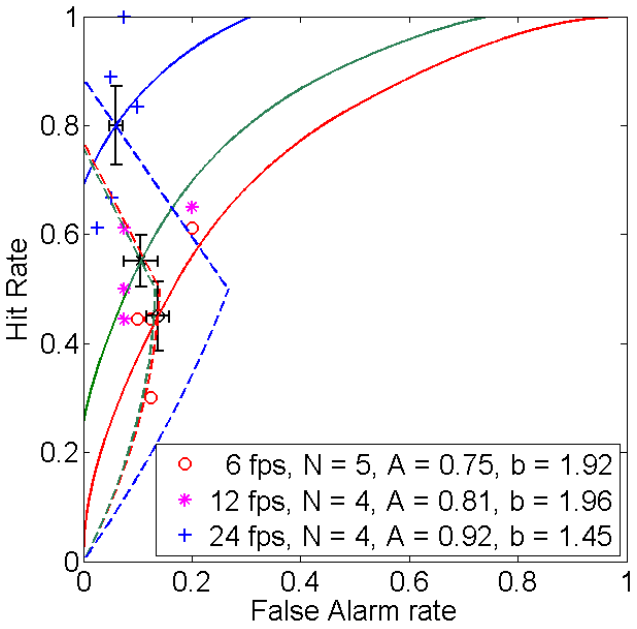

Figure 14. **Measured hit vs. false alarm rates (H, FA) for all 13 subjects and the three group averages with standard errors (crosses) and with ROC-curves for the three framerates. Straight lines = constant sensitivity A-isopleths; dotted lines = constant bias (likelihood ratio) b-isopleths.**

Individual hit rates (relative frequencies) are scattered between 0.3 and 1, whereas false alarms rates concentrate in the low probability range < 0.2, indicating conservative decisions, as would be expected for trained air traffic controllers. Circles, stars and crosses represent individual measurements (Hit, False Alarm) for FR = 6, 12, 24 Hz respectively, as obtained from the 13 subjects with repeated measurements (60 landings per subject). Black crosses with error bars show the group mean values of the individually measured (F,H)-values and the standard errors of means for the three different framerates. Solid curves represent the isopleths parametrized with the group mean A-values via equations (15) in Appendix 2. The three dotted curves represent the decision bias b, obtained from the parametric representation given in Appendix 2. b apparently decreases with sufficiently high framerate FR towards the neutral criterion value b = 1 which confirms the Bayes inference result in Figure 10 that the overestimation of speed (error bias in favor of misses, decreasing FA) decreases with framerate: the criterion shifts to more liberal values.

The three (group-average) discriminability parameters A(FR) are depicted in Figure 15 together with an exponential fit and 95% confidence intervals (using Matlab "Nlinfit").

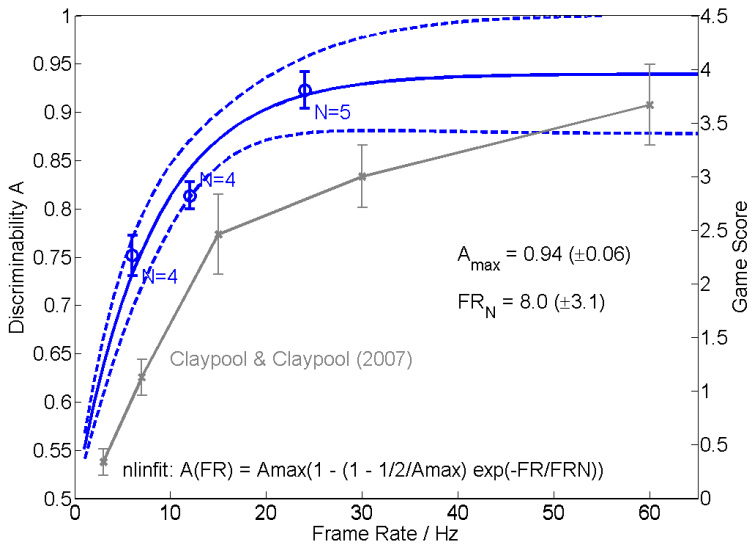

Figure 15. **Group averages (13 subjects) and exponential regression model for A (darkest solid trace) of the discriminability of landings with stopping vs non-stopping aircraft. 95% regression confidence intervals flanks the model fit. Lighter grey trace shows redrawn comparative data from [12]**

Again, like in the d'(FR)-analysis the exponential model fit to our three data points is based on the hypothesis that low framerates might disturb the visual short term memory so that with increasing visual discontinuity the speed estimate

or sequential sampling of the speed information up to the decision time becomes biased. Since the A parameter unlike the classical d', does not require the usual assumptions of Signal Detection Theory (SDT), e.g., normality of both the signal and noise distributions, it may be considered to provide a better estimate of the frame rate at which participants' performance asymptotes as provided in Ellis et al. [1]. From Figure 15 this value seems to be in the range 30 - 40 fps, a result close to the Bayes analysis, whereas the parametric SDT analysis d'(FR) appears to asymptote at a significantly larger value.

Alternatively and for the sake of parsimony our three data points, like with the Bayes analysis may be fitted with a straight line, yielding an extrapolation to ca. 31 Hz for A = 1 (maximum discriminability), which lies at the lower end of the Bayes fit confidence intervals.

Like in the d'(FR) analysis our results are compared with the (re-drawn) published results of Claypool & Claypool [12]. The latter were obtained with subject scores in a shooter game under different framerates. As mentioned above they suggest a significantly higher asymptotic FR-value for maximizing shooter scores as compared to our extrapolation in Figure 15, apparently more consistent with our d'(FR)-extrapolation.

Clearly, additional experiments with FR > 30 Hz are needed, if possible supported by a well founded theoretical model, in order to clarify this discrepancy between the different data analysis approaches.

5. Conclusion

It is clear from controller interviews that numerous out-the-windows visual features are used for control purposes [5][6][7] (see also Chapters 2, 3), which in fact go beyond those required for aircraft detection, recognition, and identification [8]. In the present work, for analyzing frame rate effects on prediction errors we focused on the landing phase of aircraft because we expected any perceptual degradation to be most pronounced in this highly dynamic situation.

Our preliminary results on the minimum framerate for minimizing prediction errors (FR_{min} > 30 Hz) show that a definitive recommendation of a minimum video framerate and a confirmation of our initial hypothesis of visual short-term memory effects resulting in the proposed asymptotic characteristic requires a further experiment with FR > 30 Hz. This high-FR experiment was not possible with the video replays used in the described experiments for technical reasons. Obviously the presented experimental data are not sufficient to decide in favor of the visual short term memory hypothesis versus a heuristic decision basis, e.g. sequential sampling or comparison of time dependent aircraft position with landmarks for thresholding. One alternative approach might be some variant of a relative judgement or diffusion model of two-alternative decision making (e.g.[16]).

A formal model for predicting the hypothetical visual memory effects would also be of great help. Recent studies which might be of use for this purpose investigate neural models for image velocity estimation (e.g.[17]) and quantify the temporal dynamics of visual working memory by measuring the recall precision under periodic display presentations between 20 ms and 1 s [18][19].
Also more detailed tower controller work analysis would be useful to clarify the operational relevance of increased framerate for decision error reduction with dynamic events in the airport environment.

Acknowledgments

Our special thanks is due to Monika Mittendorf for excellent support in the data preparation and evaluation. Furthermore we wish to thank DLR personnel Frank Morlang, Markus Schmidt, and Tristan Schindler for technical assistance in the operation of the DLR Apron-and Tower simulator (ATS) and the preparation of video files. Anne Papenfuss and Christoph Möhlenbrink organized the framework for this experiment that was part of a larger human-in-the loop RTO-simulation trial, and we are indebted to them for valuable assistance in the conduct of the experiment.

References

[1] S.R. Ellis, N. Fürstenau. N., M. Mittendorf, “Frame Rate Effects on Visual Discrimination of Landing Aircraft Deceleration: Implications for Virtual Tower Design and Speed Perception”,. Proceedings Human Factors and Ergonomics Society, 55th Annual Meeting, Sept. 19-23, 2011, Las Vegas, NV USA, pp. 71-75.

[2] S. R. Ellis, N. Fürstenau, M. Mittendorf, „Determination of Frame Rate Requirements of Video-panorama-based Virtual Towers using Visual Discrimination of Landing Aircraft Deceleration during simulated Aircraft Landing. Fortschritt-Berichte VDI, vol. 22, no. 33, 2011, pp.519-524.

[3] M. Schmidt, M. Rudolph, B. Werther, N. Fürstenau, “Development of an Augmented Vision Videopanorama Human-Machine Interface for Remote Airport Tower Operation”, Proc. HCII2007 Beijing, Springer Lecture Notes Computer Science 4558, 2007, 1119-1128.

[4] D.Hannon, J. T.Lee, M. Geyer, T., Sheridan, M. Francis, S. Woods, M.Malonson, “Feasibility evaluation of a staffed virtual tower". The Journal of Air Traffic Control, Winter 2008, 27-39.

[5] S. R. Ellis, and D. B. Liston, “Visual features involving motion seen from airport control towers”, Proc.11th IFAC/-IFIP/IFORS/IEA Symp. on Analysis, De¬sign, and Evaluation of Human-Machine Systems 9/31-10/3, 2010, Valenciennes, France.

[6] F.J.Van Schaik, G.Lindqvist, H.J.M. Roessingh, “Assessment of visual cues by tower controllers”, Proceedings 11th IFAC/IFIP/IFORS/IEA Symp. on Analysis, Design, and Evaluation of Human-Machine Systems August 31-September 3, 2010, Valenciennes, France

[7] S. R. Ellis, D. B. Liston, “ Static and motion-based visual features used by airport tower controllers”, NASA TM-2011-216427, Ames Research Center, Moffett Field, CA.

[8] A.B. Watson, C.V. Ramirez, and E.Salud, “Predicting visibility of aircraft”, PLoS ONE. 2009; 4(5): e5594. Pub-lished online 2009 May 20. doi: 10.1371/journal.pone.0005594

[9] A. J. Grunwald, and Kohn, S., “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.

[10] K. A. Kempster, “Frame rate effect on human spatial interpretation of visual intelligence”, MastersThesis, NPS, Monterrey CA. (2000)

[11] N. Fürstenau, M.,Schmidt, M.Rudolph, C.Möhlenbrink, A.Papenfuß, S. Kaltenhäuser, “Steps Towards the Virtual Tower: Remote Airport Traffic Control Center (RAiCe)”, Proc. EIWAC 2009, ENRI Int. Workshop on ATM & CNS, Tokyo, 5.-6.3.2009, pp. 67-76

[12] K. T. Claypool, and M. Claypool, “On frame rate and player performance in first person shooter games”, Multimedia Systems, 13, 2007, 3–17.

[13] J.Zhang, S.T.Mueller, “A note on ROC analysis and non-parametric estimate of sensitivity”, Psychometrica, vol. 70, no.1, 2005, pp.203-212

[14] SESAR-JU Project 06.09.03, “Remote Provision of ATS to a Single Aerodrome - Validation Report”, edition 00.01.02, www.sesarju.eu

[15] F.J. van Scheijk, J.J.M. Roessingh, J. Bengtsson, G. Lindqvist, K. Fält, “Advanced remote tower project validation results”, Proceedings 11th IFAC/IFIP/IFORS/IEA Symp. on Analysis, Design, and Evaluation of Human-Machine Systems August 31-September 3, 2010, Valenciennes, France, DOI 10.3182/20100831-4-FR-2021.00025

[16] F.G. Ashby, “A biased random walk model for two choice reaction times”, J. Mathematical Psychology, 27, 1983, 277-297

[17] J.A. Perrone, “A visual motion sensor based on the properties of V1 and MT neurons”, Vision Research 44, 2004, 1733-1755

[18] P.M. Bays, N. Gorgoraptis, N. Wee, L. Marshall, M. Husain, “Temporal dynamics of encoding, storage, and reallocation of visual working memory”, J. of Vision (2011) 11(10):6, 1-15

[19] D.E.Anderson, E.K.Vogel, E.Awh, “Precision in visual working memory reaches a stable plateau when individual item limits are exceeded”, J. Neuroscience, 31, 2011, 1128-1138

[20] N. Fürstenau, M. Mittendorf, S. Ellis,: *Remote Towers: Videopanorama Frame Rate Requirements derived from Visual Discrimination of Deceleration during Simulated Aircraft Landing.* Proceedings 2nd SESAR Innovation Days, Braunschweig 2012