Understanding the safety-relevance of visual cue perception at a Surface Manager HMI

Lothar Meyer1, Michael Schultz1, Simon Schmidt-Roßleben2 and Hartmut Fricke1
Technische Universität Dresden
Chair of Air Transport Technology and Logistics
Dresden, Germany
1{meyer, schultz, fricke}@ifl.tu-dresden.de
2simon.schmidt-rossleben@mailbox.tu-dresden.de

Abstract— Current procedures in Air Traffic Management are characterized by activities of all involved operators to obtain information that is required for the decision making to control air traffic. In particular, pilots and ATCOs decisions depend on reliable information, which, in case of failed information perception, have the highest degrees of severity. Beside voice communication, activities to perceive visual information are one of the most important methods to establish a reliable and authentic picture of the air traffic. Focusing on the migration from today’s concept of operations to higher levels of automation e.g. through virtual control tower and A-SMGCS applications, the activities of operators visual information perception differ from conventionally proven patterns. A systematic identification of visual cues for safe decisions is needed for evaluation purposes of safety critical Air Navigation Systems. The availability of knowledge about the required visual information is considered as pivotal to assess novel strategies and philosophies of HMI design while mitigating risk to at least maintain the Accepted Level of Safety. This paper follows a novel approach to analyze risk of design-related visual cue perception by the use of Human In The Loop Simulations. It hence offers a predictive empirical approach to risk analysis during system design phase. This proof-of-concept study evaluates the proposed approach at the example of a novel A-SMGCS Surface Manager HMI and the probability of runway incursion. A conclusion on the most risk-inducing visual cues could be obtained exemplarily by the use of open and closed-end questionnaires and novices as test persons. The contribution of eye tracking data analysis will also be tackled.

Keywords-safety assessment; risk assessment; hazard identification; visual cue perception; air traffic control; aerodrome control services.

I. INTRODUCTION

Every decision made in ATM related domains relies on the appropriate provision of information. Air Traffic Controller (ATCO) and Pilots represent best operators of the ATM system that are continuously involved in monitoring and decision-making processes. Both ATCOs and Pilots historically make decisions by the use of the direct visual view, which is nowadays complemented by voice communication and assistance systems. Surveillance instruments shall complement the adequate provision of information that supports the operator in fulfilling all relevant tasks [1]. Nevertheless, 22% of occurred runway incursions were attributable to “failure to detect visual information” [2], which underlines the safety relevance of visual information perception. Approvingly, today’s flight deck and tower working environment is regarded as highly sophisticated. It is commonly agreed to be acceptable safe, expressed by an agreed Acceptable Level of Safety (ALoS) [3] resp. the Target Level of Safety (TLS) by ESARR 4 [4], of 1.55E-8 fatal accidents per flight hour in ATM. This very low TLS is the outcome of decades of evaluating training, procedures and instruments and contributes significantly to the fact that ATM features a very high safety level.

Since concepts are worked out in the scope of NextGen in U.S. and SESAR in the EU [5], numerous novel systems will become operable in ATM in the next decade. The substantial increase of start-ups of novel systems, as e.g. A-SMGCS, is a historically challenging situation and bears a number of paradigm changes that potentially impairs the proven patterns of operators’ activities significantly. The well-known design and safety evaluation principles based on elementary post-startup data might be harder to apply due to the overlaying effects of multiple parallel system changes. The retrospective identification of causes in the design of systems for evaluation purposes might be increasingly affected by a not yet known complex interaction between multiple novel systems, breaking the demanded Ceteris Paribus principles during startup phases. A method to predictively identify safety impacts will consequently become valuable when maintaining the safety level predictively during design phase.

The design of a sociotechnical system, such as the A-SMGCS Surface Movement Manager HMI [5], shall support the ATCO during all task related to the aerodrome control by means of assistance tools as well as semi and full automation. This new system provides traffic information visually, enabling the operator to act according to his procedures. This procedures can include take-off and landing clearances, both highly safety critical and potentially offering the highest degrees of severity. When proposing a novel sociotechnical system for startup to operation, the predictive identification of mandatory visual cues and the design dependent contribution to risk of perceptual errors, is most valuable for a pre-startup design evaluation and hence a proactive risk mitigation.

This paper bases on an approach that considers failed activities on the perception of visual information for risk analysis at the example of an experimental Surface Manager HMI application. Visual cues have been already subject of
research in ATC and flight crew operations [5] [6]. In the scope of a proof-of-concept study, a method is proposed that shall identify visual cues and its safety impact by means of a Human-In-The-Loop Simulation (HITLS). The chosen method consists of a questionnaire on the importance of visual cues and its probability to contribute to the occurrence of runway incursion in the case of failure occurrence. The questionnaires results and the causal identification will deliver findings about the system design related probability to cause runway incursion when the visual cue perception is compromised (failed perception).

II. METHODOLOGY

Our methodology was developed at the example of an A-SMGCS Surface Manager HMI in the context of tower ATCO tasks to control the surface movements of an airport. The methodology follows an empiric approach, concluding from the observable events on risk by conducting HITLS with systems that are not in operation yet. The term risk is used according to the following definition:

“Risk is defined as the probability that an accident occurs during a stated period of time” [5].

In the present context of aerodrome traffic control, the runway incursion is a precursor of the accident event and is selected as a risk indicating event, defined by ICAO Doc 4444 as following:

„Any occurrence at an aerodrome involving the incorrect presence of an aircraft vehicle or person on the protected area of a surface designated for the landing and take off of aircraft” [6]

The relation between the visual perception event and the runway incursion is modeled in figure 1, showing an event tree that divides the causal path from the initial decision situation to 3 possibilities cases, with the runway incursion among other cases. The principle objectives that are investigated within this paper are addressed as following:

1. How can safety critical visual information that might induce risk in the case of failed perception be identified (Hazard identification)? The chosen solution is a case-based closed-end questionnaire that is conducted during the simulation, triggered by the causation of the runway incursion by operators actions. The proposed visual cues for closed-end questionnaire are derived from a ‘Visual Cue Model’ that will be introduced in the following sections. The multi causal relationship, linking visual cue perception hazards and the runway incursion statistically, provides the detected failure rate of each specific visual cue.

2. How can the probability of the runway incursion be concluded, when perceiving a specific visual cue (situation dependent probability)? The dependent probability of runway incursion causation is assumed to be the failure rate of the specific visual cue weighted by the detected runway incursion rate.

(3) How can the overall contribution of a visual cue to the probability of runway incursion be concluded (risk evaluation)? This accounts for the rate of situations, in which a specific visual cue is of crucial relevance for the decision making. In order to take account of the situation dependent relevance of visual cues, the ‘importance’ shall be gathered by questionnaires during the HITLS, providing a factor that weights the dependent probability of runway incursion causation.

Following this approach, the failure questionnaire and the questionnaire for ‘importance’ are the databases for analysis.

![Event Tree Diagram](image-url)

**Figure 1.** The event tree model “visual perception” and the related cases 1 to 3.

A. Safety Assessment Methodology

Our methodological approach is based on the Safety Assessment Methodology (SAM) provided by Eurocontrol [10] to reflect the best practices for safety assessment of ANS. The SAM development in the context of the EATMP (European Air Traffic Management Programme) provides an aviation standard procedure to reflect the high potential of damage for both the aircraft and involved third parties. An obligation on the providers of Air Traffic Management services to ensure the safety of air traffic is demanded at the ICAO Annex 11 [3]. The “burden of proof” for ATC is to satisfactorily demonstrate safe procedures and systems by obtaining an ALos [29]. In prior research projects we gained sustainable experiences on safety assessment and we further proposed a model-based approach to improve the standard assessment methods to ensure, that safety relevant results will significantly enhance the system development process [20]. As the first step of the SAM, the Functional Hazard Assessment (FHA) outputs safety objectives for identified hazards respecting system definitions, functional requirements and the TLS during system definition phase.

B. Visual Cue Model

The initial proposed visual model of tower ATCOs tasks includes demanded visual cues captured of the task analysis [16, 19] and a study of the manual of operations of the German
ANSP DFS [31]. Taking into account the cognitive information processing of human, visual cues can be composed and derived of lower level visual cues, e.g. the distance between two a/c on a radar screen is a conclusion of two single position cues. Visual cues can also be the result of an interpretation by operator’s experiences e.g. optical indicators for wind shearing in the final approach airspace.

Summarizing, a visual cue tree can be found representing a cue hierarchy that fundamentally bases on “elementary cues” in the lowest level of each branch. The proposed demand model is illustrated in figure 2 at the example of issuing a take-off clearance.

![Visual cue tree](image)

**Figure 2.** Model for the case “take-off clearance”

### III. EXPERIMENTAL DESIGN

#### A. Experimental Environment

The primary objective of the experimental design is to assess the proposed model, to identify cues that need to be considered and to exclude cues that are not demanded. Secondly, cues that were contributory to the occurrence of runway incursion are to be identified. The chosen HITLS consists of test persons that operate a Surface Manager HMI as the primary working device. The device complies with the Eurocontrol A-SMGCS Implementation level 3 [32], with the functional exception of a missing automat that prevents runway incursion (Runway Incursion Prevention and Alerting Systems, RIPAS). Tasks to be performed by the test persons are defined by ICAO Annex 11 [3, pp. 3-1] and ICAO PANS-ATM doc 4444 [33] for tower and ground control services. The Surface Manager HMI allows for the selection of target a/c by pen strokes, as well as granting pushback, taxi, lineup or take-off clearances on an airport surface surveillance radar screen presenting the entire traffic situation of airport Frankfurt a. M. (figure 3).

As one part of the experimental environment the test persons wear a mobile eye tracking device. These glasses allow for recording the individual eye movement during the experiment. Mapping the data form the included scene camera with the display information results in reliable statements about visual areas of interest.

#### B. Test scenarios

The generated traffic consists of inbound and outbound a/c traffic movements at Frankfurt a. M. airport (ICAO code: EDDF) on the four active runways in direction 25, operating 25L and 25R as landing only runways, 25C and 18 as take-off only runways. Runway dependencies are present are between 18 and 25C, as well as 18 and 25L. The dependency between 25C and 25L has been disregarded due to simulation function limitations. The random traffic generator distributes 42 movements over 45 minutes per execution run with uniformly distributed destination or departure gates (including north and south area stands) and runway. The 42 routes of the movements are initialized by a database providing 42 predefined route proposals that ensures similar taskload for all experimental executions. The simulated a/c agents are capable to separate on taxiways and to solve taxi obstruction conflicts as well as taxi crossing situations autonomously.

The execution scenario assumes one controller for both ground and tower controller tasks controlling the whole airport and inducing higher taskload than realistic scenarios would do. The scenarios are consciously adjusted to induce a task overload during a peak time in the scenario provoking sufficient samples of safety critical events to record and to conclude from; this is called the overload scenario.

#### C. Test persons and tasks

For experiment execution, there were 6 student novices available for test persons with each one conducting the experiment 10 hours divided in 4 days. With increasing hours of training, the behavior during decision situations is assumed to become more constant, which will contribute to the quality and reproducibility of observed visual activities and questionnaire results. The test persons are students in the 4th year of diploma course of studies Transport Engineering at the Technische Universität Dresden. The students had been instructed according to the tasks of tower and ground control services, as well as the provided A-SMGCS functions of the
HMI. Head on conflicts were to be prevented by the test person through routing and guiding anticipatory. The most safety critical task is assumedly clearing movements for take-off on 25C and 18 due to the runway dependencies involving movements on 25C, 25L and 18 at a time.

D. Eye tracking

The data gathered from the Tobii Glasses mobile eye tracking system consists of the scene video data, eye fixation point data and microphone data. A static cue code map of the surrounding working environment assisted to map fixation coordinates to instrument information that were probably percepted (e.g. [15, 27]). Following this strategy, the coordinate transformation from fixation coordinates to screen coordinates promises to deliver “elementary cues” of the Surface Manager HMI that were probably detected (figure 2).

To map the recorded scene video to the displayed scenario the optical flow of the background (experimental environment, e.g. monitors) and foreground (individual objects, e.g. hand, pen) have to be calculated and separated. The optical flow of the backgrounds provides information about the head movements and will be used to classify the position of the frame in the global coordinate system. The initial frame sequence is exemplarily shown at figure 4 followed by figure 5, where the frames are relocated considering the individual head movement.

Using the final aggregate picture of the scene camera, the local coordinates of tracked eye movements will be transformed the global coordinate system and mapped to the displayed scenario. During the eye tracking process each 33 ms the location of the focal point is determine and rated as a fixation or not. These fixations are aggregated, where the radius of the display location correlates with the frequency of the stored fixations (figure 6).

E. Questionnaire method

The questionnaire method is designed to freeze the simulation run immediately after randomly selected clearance situations with 6 till 8 interceptions per execution run (figure 1). The monitors are turned full black during the brake (black freeze) and the test person shall fill out the questionnaires form, namely the “regular questionnaire”. Related task was estimate the importance of the given set of visual cues according to the model figure 2 that were used for the last clearance decision on a scale from 1 (no importance at all) to 4 (very important). The questionnaires form is limited to elementary cues (cf. visual cue model) to assure comparability to the eye tracking data.

IV. RESULTS

A. Visual cue questionnaire

The execution did finalize with the assessment of the demanded visual cues sampling 153 taxi clearances and 74 take-off clearances by regular questionnaires. The test persons used the open part of the questionnaire to propose visual cues that were probable useful to complement the situational picture during all or individual situations.

The results of the closed end questionnaire show a varying degree of importance of a visual cue depending on the type. The error bar chart, figure 7, shows selected results of the regular questionnaire for the taxi clearance situation. The
information about the selected route is assumed as the most important one, allowing the operator to forecast potential head on conflicts with surrounding a/c and contribution to traffic congestions at the airport bottlenecks. In contrast, all test persons agreed that the altitude of the target a/c is never taken into account when deciding for a taxi clearance and is hence identified as a candidate for exclusion from the visual cue quantity.

The relative large standard deviation of visual cues addressing information of other a/c is attributable to individual varying habits of the test person, who monitor the environment more or less intense, which is indictable by lower standard deviation when analyzing test person wise. An individual bias in the estimation is also observable. The experimental environment supported only the clearance by the pen stroke and no voice communication. In turn to real operations, there were also no flight strips to be handled in which the callsign could have been relevant (callsign of target a/c M=1.61; SD=0.88, callsign of other a/c M=1.27; SD=0.44).

Compared to the taxi clearance results, the take-off clearance showed clear differences concerning the importance of information describing the surrounding traffic (figure 8). Other Position (M=3.24; SD=0.89), Speed (M=1.75; SD=1.05), Altitude (M=1.62; SD=1.05) became more important by the circumstance to adjust the take-off clearance with the traffic of dependent runways. The importance of the route, both target and other, decreased by nearly the half. Obviously, there are no routing activities to be done for an a/c waiting on the take-off position. The test person assumed the point of time “when to clear”, involving dependencies to other runways and expected time of a/c unblocking the runway for a non-conflicted take-off, as the most important cues. A significant difference can be quantified on other a/c altitude when separately analyzing the take-off on runway 18 (M=2.1) and 25C (M=1.3). The test person used the altitude and position information of the landing aircraft to anticipate the expected occupancy of 25L.

The open part questionnaire identified a runway coloring function, indicating the runway occupancy, to be a proper function shall support the test person during take-off decision making.

**B. Failed visual perception detection**

The runway incursion event was triggered 43 times in total during take-off and lineup decisions.

The biggest contribution to the occurrence of a hazard can be allocated to the position of other a/c (figure 9). This addresses in particular the position of departing and landing a/c when clearing for take-off. Both in the importance and the failure rate, the contribution of the visual cue other a/c plays a dominant rule, which is obvious by its safety relevance to represent environmental traffic information.

**C. Eye Tracking**

To evaluate the potential of visual cues and their relevance to prevent hazardous situations, the application of the eye tracking equipment will be demonstrated at the following example (figure 10). To give the take-off clearance to a/c #1, the test person has to check three different areas of interests to be clear of other aircraft:

- Area 1: inside runway protection zone and climb-out area of runway 25C,
- Area 2: facing intersection of runway 25C and runway 18,
- Area 3: departing from runway 25C.
The progress of two scenarios table 1 and table 2 showed exemplarily the situations instantly before a runway incursion occurs.

**TABLE I. SCENARIO A**

<table>
<thead>
<tr>
<th>Action of test person</th>
<th>Time [s]</th>
</tr>
</thead>
<tbody>
<tr>
<td>notices a/c #1 requesting clearance</td>
<td>0.0</td>
</tr>
<tr>
<td>checks for other a/c lined up for take-off on runway 25C</td>
<td>0.5</td>
</tr>
<tr>
<td>checks for departing a/c on the radar, no other departing a/c detected on radar for landing on 25L</td>
<td>1.0</td>
</tr>
<tr>
<td>clicks clearance symbol at Surface Manager HMI</td>
<td>2.5</td>
</tr>
<tr>
<td>confirms take-off clearance for a/c #1 (failure mode D)</td>
<td>3.7</td>
</tr>
<tr>
<td>causes runway incursion, a/c #1 and a/c #2</td>
<td>4.2</td>
</tr>
</tbody>
</table>

**TABLE II. SCENARIO B**

<table>
<thead>
<tr>
<th>Action of test person</th>
<th>Time [s]</th>
</tr>
</thead>
<tbody>
<tr>
<td>notices a/c #1 requesting clearance</td>
<td>0.0</td>
</tr>
<tr>
<td>checks climb-out area of runway 25C, no departing a/c detected</td>
<td>1.1</td>
</tr>
<tr>
<td>scans runway 25C for other a/c, a/c #2 not perceived at all (failure mode C)</td>
<td>1.9</td>
</tr>
<tr>
<td>checks Radar for departing a/c, no other departing a/c detected on Radar</td>
<td>3.0</td>
</tr>
<tr>
<td>clicks clearance symbol</td>
<td>3.5</td>
</tr>
<tr>
<td>scans end of runway 25C for other a/c</td>
<td>5.2</td>
</tr>
<tr>
<td>confirms take-off clearance for a/c #1</td>
<td>5.6</td>
</tr>
<tr>
<td>causes runway incursion, a/c #1 and a/c #2</td>
<td>6.7</td>
</tr>
</tbody>
</table>

The following table 3 summarizes the progress regarding to the essential checks for giving the take-off clearance and the performed checks. The eye tracking data clearly indicates the lack of information, since the test person has not focused the relevant areas of interests at the Surface Manager HMI and radar environment.

**TABLE III. OVERVIEW OF VISUAL CUES TO CHECK**

<table>
<thead>
<tr>
<th>Area of interests</th>
<th>Source</th>
<th>Scenario A</th>
<th>Scenario B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Surface Manager</td>
<td>NO</td>
<td>YES, a/c not detected</td>
</tr>
<tr>
<td>2</td>
<td>Surface Manager</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>3</td>
<td>Radar</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

V. CONCLUSION

The main purpose of this paper was to investigate a method that identifies safety critical visual cues and quantifies its causal relation to the probability of runway incursion in the scope of a risk analysis. A candidate method has been nominated that consists of a HITLS and an overload scenario, promising best conclusions on findings about runway incursion and its causative failed perception. The questionnaire and eye tracking data recorded the critical moment during the decision-making of the novice test person whether a safe clearance can be granted or not. Therein, the regular questionnaire was designed to gather the test person’s estimations of proposed visual cues concerning its importance for decision making. The results showed a diverse mix of the importance with the position information of the `target` and `other a/c` as the most important cues. Following the approach to converge quantitatively to a true demand of visual cues, few visual cues could be obtained by an open part of the questionnaire. This is due to the fact that novices use the provided information by the solution uncritically and without any experienced reference. Visual cues that are to be excluded from the proposed amount were identified according to our expectations (e.g. `altitude of taxiing target a/c`). It was clearly observable that novices used the application impartial with an observable high degree to learn from the situation. Even simulation bugs were taken into account for clearance decision-making during training. It is assumable that the pattern of activities converges to be stable rapidly. The major drawback of novices is the missing awareness regarding consequences of their decisions. The use of novices as test persons in an overload scenario provoked a large amount of sampled runway incursion events. This is beneficial to collect a surrounding set of failed perception that covers a large bandwidth of cases for analysis for causing pattern that might be symptomatic for the situational development of a runway incursion. The disadvantage of the novices test persons were the use of information without any reference to real operations.

The failure questionnaire (figure 9) showed a similar profile to the results of the regular questionnaires (figure 8), reflecting the same relations between `target a/c` and `other a/c` information within the take-off clearances. Weighting the failure rate of `position other a/c` by its importance, it is the most risk-inducing visual cue identified during take-off clearance decision making.

The eye tracking camera provided valuable information about visual scan pattern of the test person on the HMI surface. The order of picking up cues during decision making gives detailed information how information was misperceived. The eye tracking camera alone didn’t allow conclusions about the identification of visual cues, since the pure fixation data provides multiple possible cues at a time by one coordinate. A combination with questionnaire-based visual cue identification allows analyzing a specific situation.

References


**AUTHOR BIOGRAPHY**

**Lothar Meyer** (born in Hanover, Germany, 1981) studied electrical engineering at Technische Universität Dresden (TUD) from 2002-2008. He gained experiences in the field of automation systems and programming during internships at Bosch Rexroth, Volkswagen Nutzfahrzeuge and Airbus. After diploma at the institute of Aerospace Engineering and Laboratory of Control Theory at TUD in 2008 he changed to the chair of Air Transport Technology and Logistics at TUD. There, he works as a lecturer (flight characteristics and safety) and gained substantial experience in the field of Safety Assessment, Virtual Control Tower and A-SMGCS during the research project iPort/ LuFoIV.

**Michael Schultz** (born in Rostock, Germany, 1976) studied business and engineering at Technische Universität Dresden (2002) and holds an PhD degree in Aviation Technologies (2010). During several internships at Siemens Financial Services and the BMW Research Center he gained experiences in the field of quality engineering and system design. After two years employment at the automotive industry, he changed to the chair of Air Transport Technology and Logistics at TUD. His academic and industrial research projects particularly focus on passenger dynamics, future air traffic management procedures, airport performance, and reliable system design.

**Simon Schmidt-Roßleben** (born in Bonn, Germany, 1988) graduated at “Einstein” secondary school in Potsdam in 2007. He gains work experience as an Aircraft ground handler at PortGround GmbH at Dresden airport since 2010 while studying in the diploma course of studies transport engineering at Technische Universität Dresden since 2008.

**Hartmut Fricke** (born in Berlin, Germany in 1967) studied Aeronautics and Astronautics at Technische Universität (TU) Berlin from 1985-1991. From 1991 to 1995 he was a research fellow in Flight Operations, Airport Planning, and ATM at TU Berlin, where he completed his doctor thesis in ATM (ATC-ATFM Interface). In 2001 he finished his Habilitation including HIL experiments with an A340 full flight simulator in cooperation with Eurocontrol Experimental Center (EEC). Since December 2001 he has been Head of the Institute of Logistics and Aviation, and Professor for Aviation Technologies and Logistics at TU Dresden.