Increasing Efficiency in Airport Security Screening

A. Schwaninger
Department of Psychology, University of Zurich, Switzerland

Abstract

The prominence of aviation security has increased dramatically in recent years. As a reaction to the new threat situation large investments into modern security technology have been made. State-of-the-art x-ray screening equipment provides high resolution images and many image enhancement features. However, the most expensive equipment is of limited value if the humans who operate it are not selected and trained to perform their task accurately and efficiently. This article presents results of studies conducted over the last four years which clearly show that efficiency in x-ray screening can be increased substantially by investing into human factors technologies that are based on results of visual cognition, object recognition and psychophysics.

Keywords: Airport security, human factors, detection performance, computer-based training, pre-employment assessment

1 Introduction

X-ray screening of passenger bags is an essential component of aviation security. Baggage screening technologies have evolved substantially over the last decades and state-of-the art x-ray systems can provide high quality images in a good resolution. However, recognizing object shapes in x-ray images is still a very difficult problem for computers. Any serious computer vision scientist will tell you that there still is a long way to go in order to build an artificial visual recognition system that can detect objects as quickly and reliably as humans do (for recent advances in biologically motivated computer vision see for example Bülthoff, Lee, Poggio, & Wallraven [1]). Detecting all kinds of prohibited items quickly and reliably in x-ray images will continue to depend critically on human operators. Since there are large differences between people in terms of their

WIT Transactions on The Built Environment, Vol 82, © 2005 WIT Press
www.witpress.com, ISSN 1743-3509 (on-line)

visual abilities, a reliable object recognition test could be a critical component of a pre-employment assessment system (see next section). Moreover, screeners must know which objects are prohibited and what they look like in passenger bags. As will be shown in section 3, individually adaptive computer-based training can be a very effective and efficient tool in order to achieve and maintain a highly reliable and fast detection of threat items by human operators.

2 Image-based factors and screener ability

Several image-based effects influence how well threat items can be recognized in x-ray images (Schwaninger [2]; Schwaninger, Hardmeier, & Hofer [3]). As illustrated in Fig. 1a, viewpoint can strongly affect recognition performance. Systematic effects of viewpoint have been reported in many object recognition studies (for recent reviews see Graf, Schwaninger, Wallraven & Bülthoff [4]; Schwaninger [5]). In general, objects are difficult to recognize when depicted from an unusual viewpoint and when diagnostic features are not visible. Object recognition has also been shown to be dependent on exposure. When an object is not similar to a stored view, it often becomes difficult to recognize.

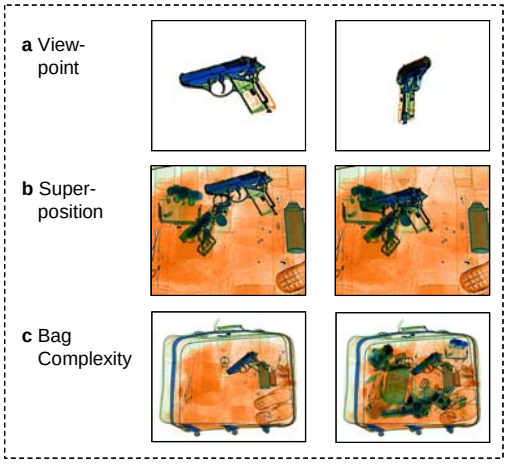

Figure 1: Image-based factors: a) effect of viewpoint (canonical vs. non-canonical), b) superpositition by other objects (low vs. high), c) bag complexity (low vs. high).

A second image-based factor is related to the fact that in x-ray images objects are often superimposed by other objects. As depicted in Fig. 1b this effect of superposition can impair detection performance substantially. A third factor is bag complexity, which is determined by the number and type of other objects in the bag (Fig. 1c). Since other objects distract attention, bag complexity can also affect detection performance.
In order to measure how well people can cope with these image-based factors, the x-ray Object Recognition Test (ORT) has been developed (Schwaninger [2]).

This test uses only guns and knives, object shapes that are known well by novices. It contains 256 x-ray images, half of them with a threat item and half of them with harmless bags. The guns and knives are depicted in easy and difficult viewpoint, with little and much superposition in bags of low and high complexity. All x-ray images of the ORT are in black-and-white, as color is mainly diagnostic for the material of objects in the bag and thus is not very helpful for novices. The ORT is fully computer-based. Each image is only displayed for 4 seconds, simulating conditions of high passenger flow during which average inspection times often are in the range of 3-5 seconds. For each x-ray image, participants have to decide whether the bag contains a threat item or whether it is harmless.

Fig. 2 shows means and standard deviations from a study conducted with 284 candidates that applied for an x-ray screening job. Detection performance is indicated by A', which can be calculated from hit and false alarm rates (for details see Schwaninger, Hardmeier, & Hofer [3]; Hofer & Schwaninger [6]).

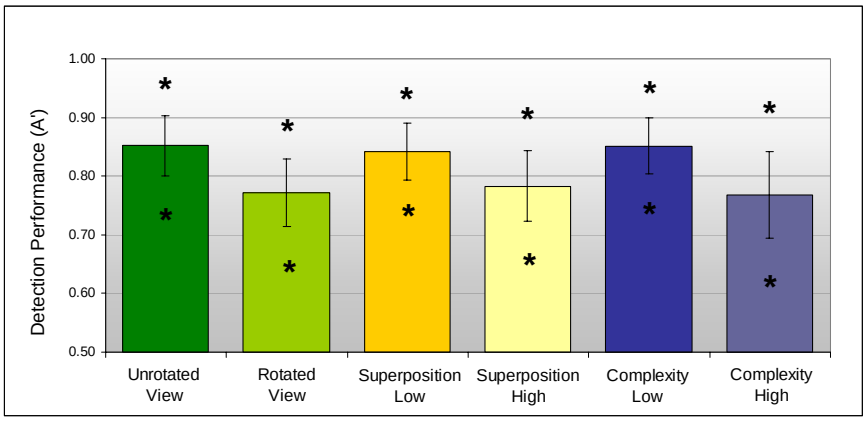

Figure 2: Effects of viewpoint, superposition, and bag complexity found in a study conducted with 284 novices that applied for an airport security screening job. Thin error bars represent standard deviations. Stars indicate the scores of the 2.5% best and 2.5% worst performing candidates.

There were systematic effects of view, superposition, and bag complexity. When the threat items were depicted in rotated view, recognition performance dropped substantially. When the guns and knives were superimposed by other objects, detection performance was also reduced significantly. Finally, showing threat items in complex bags, made them more difficult to recognize.

Really remarkable is the fact that there are very large differences between individuals with regard to how well they can cope with these image-based effects. The thin error bars in Fig. 2 depict standard deviations and the stars indicate the performance of the 2.5% best and 2.5% worst performing candidates. Similar main effects and large inter-individual differences have been consistently found for different participant groups involving novices and aviation

security screeners at different European airports. Screeners usually perform better, but we always found significant main effects of view, superposition and bag complexity. Most importantly, even for screeners, large differences can be found between people with regard to how well they can cope with effects of view, superposition and bag complexity. The results of a recent study conducted with 134 novices and 134 aviation security screeners suggest that these image-based effects are related to visual abilities. Schwaninger, Hardmeier, & Hofer [3] found that people differ substantially with regard to screener ability and this seems to remain relatively stable even after years of working as a screener. A recent analysis revealed that the results of the ORT are correlated with detection performance measured with CTI threat image projection on the job (correlations of $r > .5$). This further implies that a substantial increase in detection performance can be achieved if tests like the ORT are used as part of a pre-employment assessment system (which is the case at Zurich airport since 2004).

3 Computer-based training and screener competency

The previous section has shown that image-based factors such as view, superposition and bag complexity affect how well threat items can be detected. Even more important is the knowledge about which items are prohibited and what they look like in x-ray images of passenger bags. Consider the three x-ray images depicted in Fig. 3.

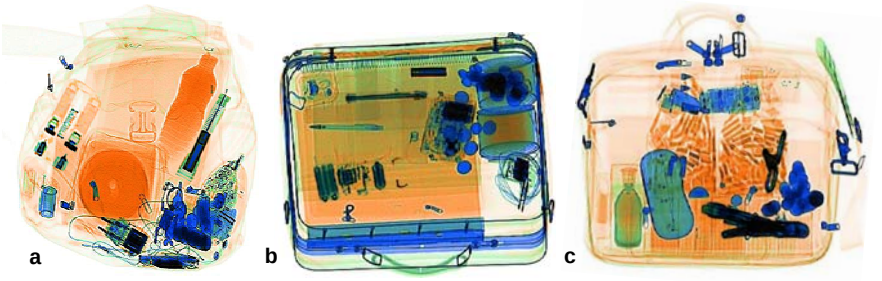

Figure 3: Threat detection in x-ray images is strongly dependent on training and expertise. a) two threat items that are not well known from everyday experience (self-defense gas spray and torch light with shooting mechanism), b) threat item that looks very different in the X-ray image (taser) vs. threat item that looks like an innocent object (switchblade knife looking like a pen), c) two threat items that are difficult to recognize without training because they are depicted in unusual views (gun and scissors).

The bag on the left of Fig. 3 contains two threat items that are difficult to recognize for novices because most people are not exposed to them in reality. The bag in the middle contains a threat item that is quite difficult to identify without training because the x-ray image looks very different than the real object. In contrast, the other threat item in Fig. 3b is difficult to recognize

because it looks similar to a pen, a well known everyday object. The bag on the right illustrates another problem that was already mentioned in the previous section. When objects are depicted in an unusual view, they become difficult to recognize if view-invariant diagnostic features are not available (for recent reviews see Graf, Schwaninger, Wallraven & Bülthoff [4]; Schwaninger [5]). Since this is often the case in x-ray images, a large image library in which objects are depicted in different views is necessary in order to achieve reliable threat detection despite changes in viewpoint.

3.1 Multiple views TIP library

Essentially, object recognition means to match visual information from stimuli (e.g. an x-ray image) to visual representations stored in memory. The content of visual memory is dependent on exposure. Shapes are difficult to recognize if they are not similar to objects that have been encountered earlier and were stored in visual memory. This was illustrated in Fig. 3 at the beginning of section 3.

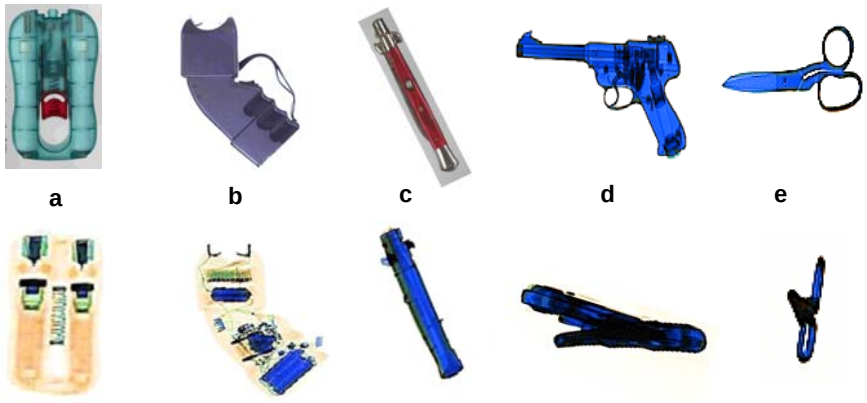

Figure 4: Illustration of different knowledge-based factors using different types of threat items. a) Forbidden object with an unusual shape (self-defense gas spray), b) prohibited item that looks very different in the x-ray image (taser), c) threat object that looks like a harmless object (switchblade knife looking like a pen), d) and e) effects of viewpoint illustrated with a gun and scissors.

The six threat items in Fig. 3 are very difficult to recognize due to different reasons (see also Fig. 4). First, many threat objects are not known from everyday experience. Examples are the self-defense gas spray and the torch light with shooting mechanism that are contained in the x-ray image depicted in Fig. 3a and Fig. 4a. Second, certain threat items look quite different than the real object. An example is the taser weapon shown in Fig. 3b and Fig. 4b. Moreover, certain other threat items look very similar to harmless objects. For example the

switchblade knife in Fig. 3b could be easily mistaken with a pen (see also Fig. 4c). A third very important but often neglected problem are effects of rotation. As mentioned earlier, when objects are depicted in unusual viewpoints, they can become very difficult to recognize. This was illustrated for a gun and scissors in Fig. 3c. The upper row in Fig. 4d and Fig. 4e depicts these two threat items in canonical view, i.e. the view that is easiest to recognize. As you can see in the bottom row, effects of viewpoint can indeed be detrimental to recognition performance.

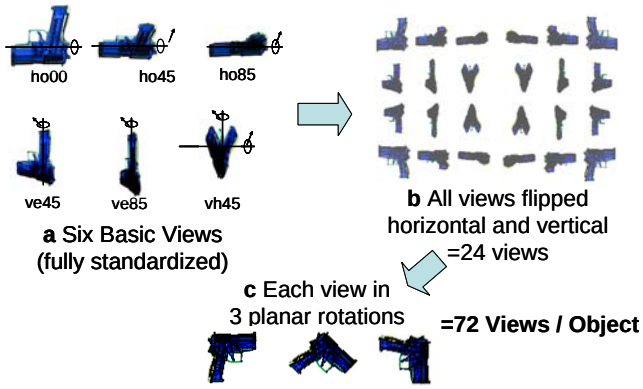

Figure 5: Basic principles of the multiple views x-ray image library.

Based on a close collaboration between vision scientists and aviation security experts we have built a multiple views library that currently contains 36'000 x-ray images of fictional threat items (FTIs). Object recognition research indicates that for most objects six views are sufficient to capture the qualitative differences resulting from viewpoint changes. These six views constitute the basic view set (Fig. 5a). The easiest view is referred to as the canonical view. It can be determined objectively by analyzing detection performance data from CBT or TIP. Alternatively, subjective rating methods can also provide a good estimate of viewpoint difficulty, at least on an ordinal scale. The remaining views of the basic view set can be created by rotating the object around the vertical axis and the horizontal axis. The method depicted in Fig. 5a uses 45° and 85° rotations around either the horizontal or the vertical axis and 45° rotations around both the horizontal and the vertical axis. Using computer algorithms mirror-reversed versions can be created as depicted in Fig. 5b. Since rotations in the plane can also have an effect on recognition performance it is advisable to create plane rotated versions (Fig. 5c).

The current version of our view-based library contains 500 prohibited items and each of them can be displayed in 6 basic views x 4 mirror reversals x 3 plane rotations = 72 views. This results in a total of 500 prohibited items x 72 views = 36'000 x-ray images. Of course it is not necessary to show each screener all these images. But an individually adaptive training system should determine for each screener which views are difficult and adapt the training so that the trainee

becomes able to detect threat items reliably even if they are substantially rotated away from the easiest view.

3.2 Individually adaptive training

After analysing the results of the first studies conducted at Zurich Airport in 2000 and 2001 we were looking for a CBT system that has a large image library containing thousands of images depicting threat items in different viewpoints. Moreover, such a system should be efficient and effective by applying individually adaptive training algorithms that calculate for each screener which items and views need to be trained particularly. The results of our first survey were that all existing CBT systems had only small image libraries and they did not feature individually adaptive training algorithms. Instead, the commercially available systems focussed on providing many image processing features such as negative image, black and white, edge detection etc. It actually sometimes seemed that the quality of a CBT would be dependent on providing an exact replica of the x-ray machine with as many image processing and enhancement features as possible. However, the goal of the training system we had in mind was to train screeners to detect threat items reliably within the first 3-5 seconds of inspection time *before* image processing features might be applied. In order to achieve this, a scientifically based training system was needed that features a large image library, sophisticated adaptive training algorithms and feedback mechanisms that are based on results of perceptual learning. Since nothing like that existed, we decided to create such a system on our own. In a close collaboration between psychologists, vision and computer scientists as well as aviation security experts the first version of X-Ray Tutor was built in 2002 and installed at Zurich airport. It is now used by some of the largest security organizations in the world. The United States Transportation Security Administration is using X-Ray Tutor at airports across the country since spring 2004 as part of its recurrent training and professional development program for airport security screeners. The Canadian Air Transport Security Authority has started using X-Ray Tutor at several airports as part of a research project with the University of Zurich on x-ray image tutoring technology for aviation security screeners. X-Ray Tutor is operational in Switzerland since 2002, at 17 airports in Germany since 2003, and since 2004 in Belgium and other European countries.

At the heart of X-Ray Tutor are results from scientific studies on how the human brain processes visual information in order to recognize objects from different viewpoints (Schwaninger [7], [8]; for recent overviews on object recognition see Graf, Schwaninger, Wallraven & Bülthoff [4]; Schwaninger [5]). The system uses a sophisticated algorithm to create training sessions for each individual student based on detection performance, response time and x-ray image difficulty ratings. The individually adaptive algorithm of X-Ray Tutor starts with easy x-ray images. Then, view difficulty is increased and trainees learn to detect threat items in different viewpoints. X-Ray Tutor uses the view-based image library described above, which allows presenting each threat item in 72 different viewpoints. As mentioned before, it is not necessary to show each item in all viewpoints. The adaptive algorithm of X-Ray Tutor shows for each trainee

preferentially those views, which need to be trained especially. Once a threat item can be recognized despite changes of viewpoint, it is shown in more complex bags with more superposition. The individually adaptive algorithm uses objective measures of view difficulty, superposition and bag complexity. These measures have been shown to be correlated with human detection performance and they are based on several years of research involving vision scientists, image processing specialists, statisticians, and aviation security experts. As students progress the algorithm monitors students' performance and custom designs each subsequent training session to target areas in need of improvement.

Students learn more when given the opportunity to receive immediate feedback about their performance. X-Ray Tutor provides detailed feedback during a training session about the student's performance. Students can review any x-ray image to learn where the threat item is located and what it looks like both in x-ray and standard photographic view (see Fig. 6). At the end of each training session, students are given feedback on their performance, so they can visually see their improvement over time.

Figure 6: Screenshot of X-Ray Tutor CBS. Immediate feedback is provided after each trial and trainees can display a threat information window showing the threat item as x-ray image and as photograph.

X-Ray Tutor can also be used for several research purposes. Using the data of training trials is complicated by the fact, that the training system adapts to each individual and displays more difficult images the better a screener becomes in detecting threats. In order to achieve a more reliable measurement of detection performance across time, a certain number of "test images" can be shown within training sessions. These images are the same for all screeners providing a more reliable basis for standardized detection performance measurement. The test images can be shown periodically at criterion-based or time-based intervals so that an evaluation of training effectiveness across time becomes possible. When X-Ray Tutor is used in a research project, the data from testing trials is stored in

reports than can be imported easily into applications for further analysis such as standard office software or analysis tools for large databases. Using the reporting feature, detection performance can be analyzed and visualized across time, for individual screeners, groups of screeners, specific FTIs and groups of FTIs. This data can then be used in order to answer several important questions such as for example:
- How does detection performance increase over time?
- Which threat items are difficult to recognize?
- How does viewpoint affect threat detection performance?
- How reliable are individual performance measurements?

These questions are currently being investigated in research projects conducted in our lab. The following section summarizes the results of a study that was conducted in 2002/2003 in order to measure training effectiveness and investigate perceptual learning mechanisms. A detailed description of this study can be found in Schwaninger and Hofer [9].

3.3 Measuring Effectiveness and investigating perceptual learning

The study was conducted in the period of December 2002 – May 2003 with 72 participants. None of them had received computer-based training before. Each week, 1-2 training sessions of 20 minutes were completed during the six months period. X-Ray Tutor HBS for training detection of improvised explosive devices (IEDs) was used. The research questions were as follows:
- Does IED detection performance increase over time as a result of individually adaptive training?
- Are there large differences in detection performance when images are displayed 4 vs. 8 seconds?
- Does training result in a decrease of response time? If so, is that decrease different for threat images vs. harmless bag images?

Four tests were conducted in which new IEDs were used that had not been shown previously during training. Display durations at test were 4 and 8 seconds (during training, images were displayed for a maximum of 12 seconds). The four tests were of equal difficulty as determined in a pilot study. Signal detection theory was used to analyze the data (Green & Swets [10]). Detection performance was measured using d' = z(H) – z(FA), whereas H represents the hit rate, FA the false alarm rate and z the z-transformation. Similar results were obtained using non-parametric A' instead of d' as performance measure (for a comprehensive introduction on signal detection and x-ray screening see Schwaninger [11]; for more detailed information on different detection measures of x-ray detection performance see Hofer and Schwaninger [6]). The results of the four tests are displayed in Fig 7. There was only a small effect of display duration, i.e. detection performance was similar for display durations of 4 and 8 seconds. This is consistent with the view that recognition is a fast process occurring within the first few seconds of image inspection. More important is the knowledge about threat items, which is dependent on training. Indeed, large

increases of detection performance were observed across the four months of training (Fig. 7a). In order to assess training effectiveness we calculated % increase as compared to baseline measurement (first test results). As you can see in Fig. 7b relative detection performance was increased by about 71%. This is a remarkable effect if it is taken into account that on average screeners took only 28 training sessions during the six months period (*SD* = 10 training sessions). Moreover, for a subgroup of 52 screeners, who on average took 31 training sessions (*SD* = 8 training sessions), the training effect was even more pronounced. For this group relative detection performance was increased by 84%!

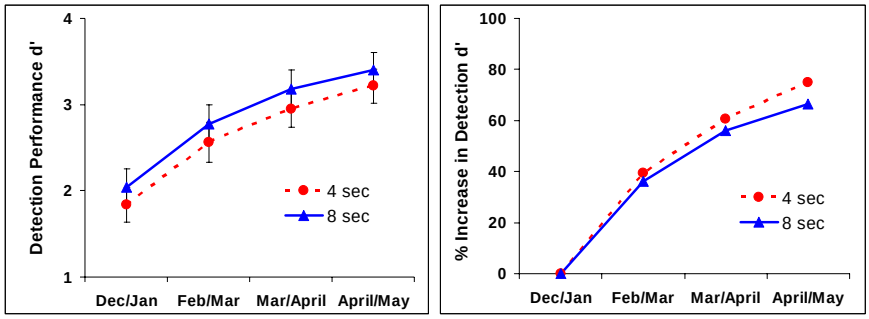

Figure 7: a) Absolute detection performance measured by d', b) detection performance increase expressed in percent relative to the first measurement. Error bars represent SEM.

The analysis of response times revealed interesting findings with regard to efficiency. As can be seen in Fig. 8, training with X-Ray Tutor resulted in a much faster detection of threat items. Correct response times for threat images dropped from about 5 seconds before training to about 3.5 after six months of training.

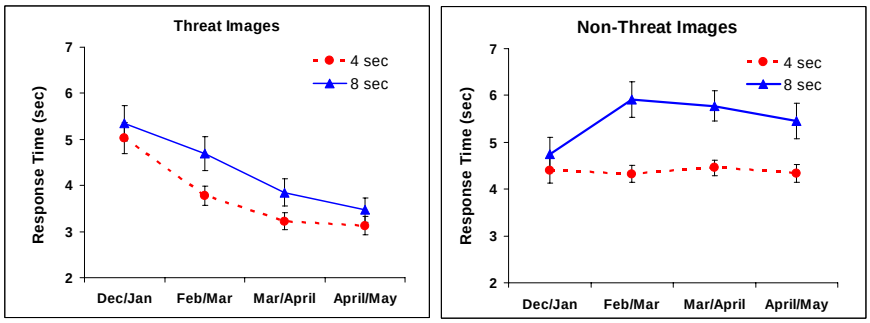

Figure 8: Response time data as a function of training. a) Response times for correct decisions on threat images, b) response times for correct decisions on non-threat images. Error bars represent SEM.

There was also an effect of display duration, although much smaller in effect size than the overall training effect. For harmless bags average response times remained constant at about 5 seconds. This finding is consistent with the assumption that individually adaptive CBT increases the number and strength of representations in visual memory. Because the participants have learnt what IEDs look like they became able to detect them reliably within a few seconds of inspection time. However, the acquired knowledge does not help if there is no threat item in the bag. Since no substantial reduction in response time was observed for harmless bag images, the learning effect seems rather related to visual memory representations than to increased general visual processing capacities.

4 Conclusion

These results show that pre-employment assessment and individually adaptive training can be very powerful tools in order to increase efficiency in aviation security screening. Moreover, the last study also provides a nice example for the value of X-Ray Tutor as a research tool. For measuring detection performance and training effectiveness, as well as gaining a better understanding of perceptual learning mechanisms and object recognition.

5 Acknowledgements

I am thankful to airport and police authorities who provided data and funding for conducting the studies reported in this paper. Thanks to Franziska Hofer and Judith Riegelnig for their help in preparing the manuscript.

6 References

[1] Bülthoff, H.H., Lee, S.-W., Poggio, A., & Wallraven, C., *Biologically motivated computer vision, Lecture Notes in Computer Science, 2525*, Berlin: Springer, 2002.
[2] Schwaninger, A., Evaluation and selection of airport security screeners. *AIRPORT*, 02/2003, 14-15, 2003.
[3] Schwaninger, A., Hardmeier, D., & Hofer, F., Measuring visual abilities and visual knowledge of aviation security screeners. *IEEE ICCST Proceedi*ngs, *38*, 258-264, 2004.
[4] Graf, M., Schwaninger, A., Wallraven, C., & Bülthoff, H.H., Psychophysical results from experiments on recognition & categorisation. *Information Society Technologies (IST) programme,* Cognitive Vision Systems - CogVis (IST-2000-29375), 2002.
[5] Schwaninger, A., Object recognition and signal detection. In: B. Kersten & M.T. Groner (Eds.), *Praxisfelder der Wahrnehmungspsychologie*. Bern: Huber, 2004.

[6] Hofer, F., & Schwaninger, A., Reliable and valid measures of threat detection performance in x-ray screening, *IEEE ICCST Proceedings, 38, 303-308*, 2004.
[7] Schwaninger, A., Training of airport security screeners. *AIRPORT*, 05/2003, 11-13, 2003.
[8] Schwaninger, A., Computer based training: a powerful tool to the enhancement of human factors. *Aviation Security International*, FEB/2004, 31-36, 2004.
[9] Schwaninger, A. & Hofer, F. Evaluation of CBT for increasing threat detection performance in x-ray screening. In: K. Morgan and M. J. Spector, *The Internet Society 2004, Advances in Learning, Commerce and Security (pp. 147-156)*. Wessex: WIT Press, 2004.
[10] Green, D. M. & Swets, J. A., *Signal detection theory and psychophysics*, Wiley: New York, 1996.
[11] Schwaninger, A., Reliable measurements of threat detection. *AIRPORT*, 01/2003, 22-23, 2003.