The effect of image enhancement functions on x-ray detection performance

Stefan Michel
University of Zurich, Department of Psychology, Visual Cognition Research Group (VICOREG), Switzerland
Saskia Koller, Markus Ruh and Adrian Schwaninger
University of Zurich, Department of Psychology, Visual Cognition Research Group (VICOREG), Switzerland

Abstract

State-of-the art x-ray screening systems offer a variety of so-called "image enhancement" functions (IEFs). Examples are color inversion, edge-enhancement, organic only, metal only etc. IEFs are often promoted because they would bring out detail that is obscured or highlight certain features, such as for example organic content. In this study, we investigated the usefulness of IEFs for cabin baggage screening (CBS) and hold baggage screening (HBS). For CBS, the standard image and nine IEFs have been tested with 443 x-ray screeners that completed a standardized x-ray image interpretation test using Smiths-Heimann HISCAN 6040i x-ray imagery (X-Ray CAT, Koller & Schwaninger, 2006). The results showed that the standard image provided the best detection performance. Some IEFs impaired detection performance substantially if images were displayed only with one IEF, which was also dependent on threat type (guns, knives, improvised explosive devices, other threat items). A second experiment with 83 HBS screeners using a bomb detection test, Smiths-Heimann HISCAN 10080 2i imagery, and 5 IEFs showed similar results. Together with previous work conducted with Rapiscan images (Klock, 2005), these results highlight the importance of systematically studying the usefulness of IEFs in order to optimize human-computer interaction in x-ray screening.

Introduction

In recent years, the importance of baggage x-ray screening at airports has increased dramatically. The image quality of older x-ray screening equipment was sometimes in need of improvement. For example an early version of a colouring algorithm as enhancement function did not serve the purpose of increasing detection performance of threat objects, actually it impaired it. This was due to the occlusion of object parts by the opaque colouring algorithm (Schwaninger, 2005a; Schwaninger, 2005c). But there was a large technological progress in the last years, especially for x-ray screening machines, which nowadays provide high image quality and various image enhancement functions (IEFs). The main objective of such functions is to process an image so that the result is more suitable than the original image for a specific application as for example x-ray screening at airports (Gonzalez & Woods, 2002). In x-ray images, the image enhancements might increase the visibility of objects within the bag and remove background noise. The aim of this study is to investigate if IEFs help increasing the detection of threat objects in passenger bags.

The X-Ray Competency Assessment Test (X-Ray CAT) was used in this study. It contains threat images of the four categories guns, knives, improvised explosive devices (IEDs) and other threat objects (see Koller & Schwaninger, 2006 for details). Hold baggage screening focuses on the detection of bombs as other threat items as guns or knives do not pose a threat in hold baggage. In this study they conducted a Bomb Detection Test (BDT) which contains only IEDs.

This is a reliable way to prove if these enhancement functions actually are of use (i.e. effectuate a higher or at least equally high threat detection performance as with original x-ray images). These findings can furthermore help to optimize the human-machine interaction.

Background - Image Enhancement Filters (IEFs)

The nine image enhancement filters used in this study can be applied to the x-ray images on the screen while working at an x-ray machine. Each pixel in the image format used in these x-ray machines has a material and a luminance value. To show the images on a screen, the pixel values are color coded using red for organic, blue for metallic and green for mixed organic/metallic material. The luminance value defines the luminance of the pixel.

Grayscale	The Grayscale filter removes the material information from the image and shows only the luminance value.
Luminance High	In this filter, the luminance of the image is increased by applying a gamma correction (Pratt, 2001) to the luminance value. This allows the screeners to see details in dark areas of x-ray images, but as a consequence the visibility of details in light areas of the images is reduced.
Luminance Low	As the opposite of the Luminance High filter, the luminance of the image is decreased. Details in light areas of the image become more visible, dark areas lose the details.
Luminance Negative	In the Luminance Negative filter, the luminance of the image is inverted. The material value and therefore the hue of each pixel remains the same.
Metal Only	Here, only the metallic parts of the image are shown in color. The organic parts are transformed to light gray with low contrast. The organic parts of the mixed organic/metallic pixels are removed as well, giving them a blue color similar to the all-metallic parts. Applying this filter allows the screeners to concentrate on the metallic objects perhaps leading to less search time for such objects.
Metal Stripping	The Metal Stripping filter removes the metal from the image. Metallic parts are transformed to light gray and from the mixed organic/metallic pixels the metallic part is removed. As some mixed organic-metallic parts originate from metallic objects laying upon organic objects, this removal of metal sometimes shows the complete organic object without distracting metallic parts.
Organic Only	The Organic Only filter shows the organic parts of the image in color, while the metallic pixels are set to gray. The mixed organic/metallic pixels are assigned to the metallic or organic parts depending on the proportion of metallic and organic material. The difference to the Metallic Stripping filter is that less of the image remains visible and that the remaining mixed organic/metallic pixels are still green.
Organic Stripping	As the opposite to the Organic Only filter, the metallic parts of the image remain colored and the organic parts are shown in light gray with low contrast. The resulting image is similar to the Metal Only image, except that in this filter the mixed organic/metallic pixels are still green.
Super Enhancement	The Super Enhancement filter adaptively adjusts the contrast of the image. Similar to a Local Histogram Equalization (Gonzalez & Woods, 2002) or an Adaptive Contrast Enhancement (Stark, 2000), the luminance of each pixel is adjusted to the luminance of its surrounding pixels. In the resulting image each area has a medium average luminance which in contrast to Luminance Low and Luminance High makes details in dark and light areas visible at the same time.

Experiment 1

Participants

A total of 443 airport security screeners of the Cabin Baggage Screening (CBS) at a European airport conducted the X-Ray CAT twice. They were assigned to one of ten image enhancement groups in a way to get equal groups in terms of gender, age, days on job. Furthermore, the ten groups showed an equal average of detection performance A', which was calculated from data of a separate test conducted prior to this study.

Materials

For this study the X-Ray CAT has been used. It is composed of 256 images, 128 images of passenger bags once containing a threat object and once not. Threat objects were chosen from four categories, guns, knives, improvised explosive devices (IEDs) and other prohibited items. The nine predefined conventional image enhancement algorithms (see table above) were applied to the x-ray images. The task is to visually inspect the images and to judge whether they are OK (contain no prohibited item) or NOT OK (contain prohibited item). In this study, images disappeared after 10 seconds. Additionally, screeners had to indicate the perceived difficulty of each image on a 100 point scale (difficulty rating; 1=easy, 100=difficult). All responses are given by using the mouse to press buttons and use a slider control on the screen.

Design

Participants first conducted the X-Ray CAT with one of the nine image enhancement filters and a second time with the unaltered x-ray images. A control group conducted the test twice with the unaltered x-ray images.

Results and Discussion

Detection performance was measured using A', a measure derived from hit and false alarm rates (Pollack & Norman, 1964, Hofer & Schwaninger, 2004). Detection performance values are calculated generally for the complete test as well as for each threat category separately. Detection performance is worse for all image enhancement filters compared to the original images. This applies to the hit rate as well as to the false alarm rate and therefore also to A'. Figure 1 shows means and standard deviations of A' scores broken up by image enhancement function and pooled across threat categories. An analysis of variance (ANOVA) with the between-participants factor enhancement was carried out on individual A' scores averaged per screener across threat category. There was a main effect of

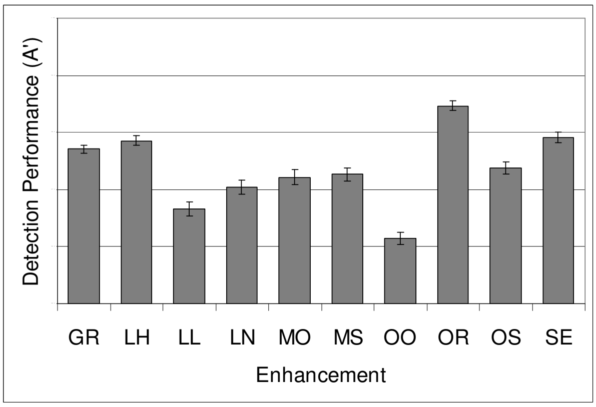

Figure 1. Detection performance Experiment 1 pooled across threat categories

Figure 2a. Detection performance for knives Experiment 1

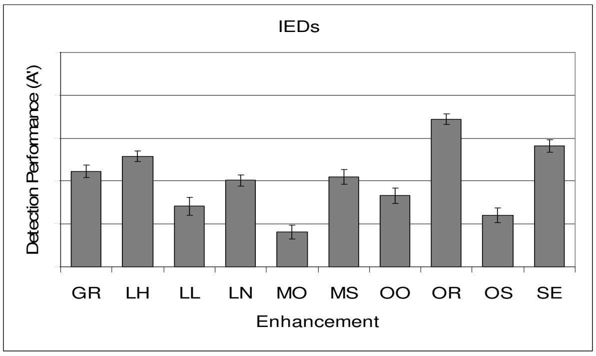

Figure 2b. Detection performance for IEDs Experiment 1

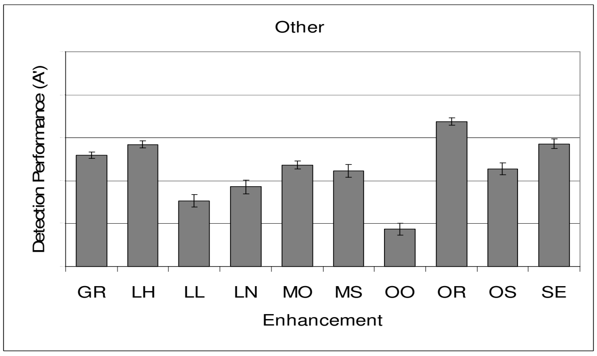

Figure 2c. Detection performance for other threat items Experiment 1

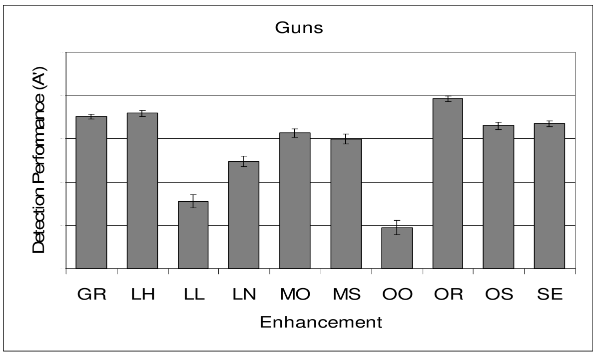

Figure 2d. Detection performance for guns Experiment 1

enhancement with an effect size of $\eta 2 = .46$, $F(9, 433) = 41.67$ and $p < .001$. An ANOVA on A' with the between-participants factor enhancement and the within-participants factor threat category (guns, IEDs, knives, other threat objects) showed significant main effects of threat type with an effect size of $\eta 2 = .30$, $F(3, 1299) = 180.84$ and $p < .001$, and of enhancement with an effect size of $\eta 2 = .48$, $F(9, 433) = 43.66$ and $p < .001$. The interaction of threat category x enhancement was also significant with $\eta 2 = .32$, $F(27, 1299) = 22.91$ and $p < .001$. This result implies that the effect of image enhancement filters on detection performance varies depending on the threat category. This is consistent with the results depicted in Figure 2. One-way ANOVAs on A' scores calculated separately for each threat category showed a main effect of enhancement for guns with an effect size of $\eta 2 = .64$, $F(9, 433) = 86.09$ and $p < .001$, for IEDs $\eta 2 = .32$, $F(9, 433) = 22.38$ and $p < .001$, for knives $\eta 2 = .32$, $F(9, 433) = 23.10$ and $p < .001$ as well as for other prohibited items with an effect size of $\eta 2 = .43$, $F(9, 433) = 36.27$ and $p < .001$.
The fact that IEFs affected the detection of threat items differently depending on threat category can at least partly be explained by the materials of the threat objects. The Organic Only filter for example is effective for the detection of improvised explosive devices (IEDs) as these are mostly constructed of organic material, whereas this filter has no effect for the detection of knives or guns because knives and guns usually consist of metallic material. In contrast, the Metal Only filter helps detecting knives and guns as only metallic material is shown and it consequently has no effect for the detection of IEDs. The data also shows that the Metal Stripping filter might be a better option than the Organic Only filter. This is because the Metal Stripping filter includes information about organic material hidden behind metallic parts, whereas the Organic Only filter simply removes these parts from the image. The removal of the color-coded material information by the grayscale filter does impair the threat detection, but the luminance information seems to be more important than the material information. When inserting a threat object into a bag, the part of the bag with the object inside normally becomes darker than its surrounding. Therefore most threat objects can be detected by just looking at the darker parts of the image. But the removal of material information can conceal objects with the same luminance but different material than its surrounding. A similar problem appears when using the Super Enhancement filter. There the material information stays the same, but the luminance contrast is slightly reduced which has a negative influence on the threat object detection. The Luminance High filter allows better threat object detection than the Luminance Low filter. With the Luminance Low filter, most objects inside the bag have a luminance close to black, which generally reduces the differentiation of these objects. The impairment of threat detection when using the Luminance Negative filter shows that screeners perform better with a dark object on a light background. The material and luminance information remains in the image, the luminance is simply inverted. In summary, these results challenge the idea that IEFs improve the image and enhance detection. In fact, some of the IEFs impaired detection substantially.

Experiment 2

Participants

Data of 83 aviation security screeners of the Hold Baggage Screening (HBS) was analyzed. All screeners were assigned to six equal groups concerning gender, age, days on job and detection performance A', which was calculated from training data. The procedure used to create the two groups was the same as for CBS screeners (see Experiment 1).

Materials

For the HBS screeners stimuli containing IEDs were used. The Bomb Detection Test (BDT) is composed of 200 x-ray images, 100 images of passenger bags once containing an IED and once not. In order to still have enough participants in each group the number of investigated enhancements for HBS had to be limited to five, namely Black-White (corresponds to the Grayscale filter in Experiment 1), Luminance Negative, Organic Only, Organic Stripping and Super Enhancement.

Design

Participants first conducted the BDT with one of the five image enhancement filters and a second time with the unaltered x-ray images. A control group conducted the test twice with the unaltered x-ray images.

Results and Discussion

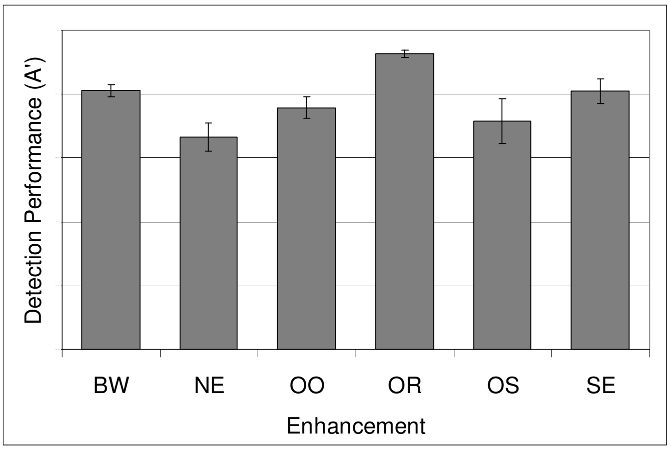

Figure 3. Detection performance Experiment 2

Analyses were the same as in Experiment 1 but there was only one threat category, i.e. IEDs. Figure 3 shows means and standard deviations of A' scores broken up by image enhancement function. There was a main effect of enhancement with an effect size of $\eta 2 = .26$, $F(5, 77) = 5.29$, $p < .001$.

General Discussion

The aim of this study was to investigate the effect of image enhancement filters on x-ray detection performance of airport security screeners. Experiment 1 was conducted with CBS screeners. It was found that the unaltered image resulted in the best performance, whereas some IEFs resulted in substantial impairments of detection performance. These effects varied across threat categories, which is dependent on the material composition of the threat categories. For instance, IEDs contain a lot of organic material (explosive). Applying the Metal Only filter or the Organic Stripping filter therefore results in a massive impairment of detection, whereas the impairment is marginal with the Organic Only filter. In contrast, guns and knives contain mainly metallic parts. It is therefore not surprising that detection performance is substantially impaired when the Organic Only filter is applied, while the Metal Only or the Organic Stripping filters result in a smaller impairment. Experiment 2 was conducted with HBS screeners and again the unaltered x-ray images resulted in the best performance, while some IEFs impaired performance substantially. Thus this study provided converging evidence that the unaltered x-ray image is the best, at least for the equipment tested here and as an overall result by averaging across different exemplars of four categories (guns, IEDs, knives, and other threat items). It should be pointed out that this study does not necessarily imply that IEFs are of no value. For instance, it could be that certain IEFs are useful for certain types of images. By averaging across images a potential value of IEFs might has been averaged out. Moreover, this study did not consider the effect of a sequence of enhancements. Screeners saw the images with only one filter applied. It can not be ruled out that detection performance may be better when particular enhancement filters are applied sequentially to images. At the x-ray screening machines screeners have the possibility to switch between filters and look at an image with more than one filter sequentially. This procedure could help detecting particular objects and thus the enhancement filters, applied sequentially, could actually provide support for the detection and not, as found in this study, impair detection performance. Furthermore, there may be a learning effect; if a screener regularly sees x-ray images with a certain filter, detection could become better due to a customization to the filter and therefore improvement of searching performance. Finally, there could be stable inter-individual differences regarding the processing of visual information resulting in a preference of certain IEFs for some vs. other screeners. Should this be the case, a potential benefit of the IEFs might have been averaged out by calculating mean values across screeners.

In summary, this study has provided converging evidence that for the equipment tested in Experiments 1 and 2, the unaltered image provides the best image quality regarding detection performance averaged across screeners and across different exemplars of the four threat categories used (guns, IEDs, knives, others). However, additional research is needed to completely rule out any potential benefits of image enhancements for certain images and/or screeners.

References

Gonzalez, R.C., & Woods, R.E. (2002). *Digital Image Processing (2nd ed.)*. Upper Saddle River, NJ: Prentice Hall.

Hofer, F., & Schwaninger, A. (2004). Reliable and valid measures of threat detection performance in X-ray screening. *IEEE ICCST Proceedings*, 303-308.

Klock, B. A. (2005). Test and evaluation report for X-ray detection of threats using different X-ray functions. *IEEE ICCST Proceedings,* 182-184.

Koller, S., & Schwaninger, A. (2006). Accessing X-ray image interpretation competency using the X-Ray Tutor Competency Assessment Test X-Ray CAT. *Proceedings of the 2nd International Conference on Research in Air Transportation, ICRAT 2006, Belgrade, Serbia and Montenegro, June 24-28, 2006,* 399-402.

Pollack, I. & Norman, D.A. (1964). A non-parametric analysis of recognition experiments. *Psychonomic Science*, 1, 125-126.

Pratt, W. K. (2001). *Digital Image Processing: PIKS Inside, 3rd Edition*. John Wiley & Sons, Inc., New York, NY

Schwaninger, A. (2005a). Objekterkennung und Signaldetektion: Anwendungen in der Praxis. In B. Kersten & M. Groner (Hrsg.), *Praxisfelder der Wahrnehmungspsychologie* (pp. 106-132). Bern: Huber.

Schwaninger, A. (2005c). X-ray Imagery: enhancing the value of the pixels. *Aviation Security International*, Oct 2005, 16-21.

Schwaninger, A. & Hofer, F. (2004). Evaluation of CBT for increasing threat detection performance in X-ray screening. In: K. Morgan and M. J. Spector (Eds.), *The Internet Society 2004, Advances in Learning, Commerce and Security* (pp. 147-156). Wessex : WIT Press.

Stark, J.A. (2000). Adaptive image contrast enhancement using generalizations of histogram equalization. *IEEE Transactions on Image Processing*, 9, 5, 889 - 896.