Measurement Formulae for Image-Based Factors in X-ray Imagery

Bolfing, A. & Schwaninger, A.

University of Zurich
Visual Cognition Research Group (VICOREG)
Technical Report
November 26, 2007
Point of Contact: Dr. Adrian Schwaninger, a.schwaninger@psychologie.uzh.ch

Image Based Factors

Schwaninger (2003b) and Schwaninger, Hardmeier, and Hofer (2005) have identified three image based factors which affect threat detection in x-ray screening significantly: View Difficulty, Superposition and Bag Complexity. For the development of a statistical model for automatic image difficulty estimation these concepts were implemented mathematically by image measurement formulae (Schwaninger, Michel, & Bolfing, 2007). Such a model is very valuable for x-ray image difficulty estimation to analyze detection performance of computer-based training (CBT) and threat image projection (TIP) data. Automated image measurements of view difficulty, superposition and bag complexity are also essential for the development of effective and efficient CBT algorithms (e.g. Schwaninger 2004b, 2005c). Moreover, such measurements are important for the development of reliable computer-based tests for measuring x-ray image interpretation competeny (e.g. Koller, Hardmeier, Michel, & Schwaninger, 2007, in press) and are very useful for the development of pre-employment assessment tests (e.g. Hardmeier, Hofer, & Schwaninger, 2006a).The following sections describe the theoretical concepts of the image based factors and updated formulae on how they can be measured automatically.

FTI View Difficulty (FtiVD)

Even with the aid of 3D volumetric models, it is not (yet) possible to satisfyingly determine the degree of a 3-dimensional rotation of a physical threat item automatically from its 2-dimensional x-ray image. Difficulties regarding image segmentation arise from the very heterogeneous backgrounds of x-ray images. Therefore, this image based factor is not (yet) being calculated directly by image measurements, but statistically from TIP detection performance. FTI View Difficulty as it is conceptualized and implemented here reflects image difficulty attributes unique to each exemplar view of an FTI and not solely the degree of rotation.

$$\mathrm{FtiVD}_{OV_j} = \frac{\sum_{i=1, j \neq i}^{N_{OV}} \left(max(\mathrm{DetPerf}) - \mathrm{DetPerf}_{OV_i} \right)}{N_{OV} - 1}$$

The general FTI View Difficulty formula above describes a slight modification of the mean of the inverted detection performance value (DetPerf). Inverted means here, that the measured detection performance is subtracted from the theoretically maximal detection performance (e.g.: 1 for hit rates; 4.65 for d', for details on d' please consult Green & Swets (1966)). The slight modification refers to the averaging over all items (index N_{OV}) containing the same FTI object (subindex *O*) in the same view (subindex *V*) as does the item in question whose FTI View Difficulty is calculated. The detection performance value of the item in question is not included in this averaging. The single detection performance values themselves are mean values averaged across subjects. The exclusion of the inverted detection performance of the item in question from the averaging is needed to circumvent a circular argument when using the formula to predict detection performance of human operators (e.g. airport security screeners).

For analyzing TIP data the formula should be modified. First, because usualy only SN trials (bag images containing a threat item; signal-plus-noise) are recorded, the detection performance measure is the hit rate. As described in the previous paragraph, its inverted performance measure is the miss rate (miss rate = 1 – hit rate). Unlike when using the hit rate, the miss rate reflects difficulty: the more an item was missed the more difficult the item ought to be. If a large TIP data set is used, the exclusion of the item in question from the averaging can be abandoned due to its very small weight (as for example in the *Study on the effect of cabin bag size on the effectiveness of screening*).

$$\mathrm{FtiVD}_{OV} = \frac{\sum_{i=1}^{N_{OV}} \mathrm{MissRate}_{OVi}}{N_{OV}}$$

Exemplary illustration: The FTI View Difficulty formula shows the calculation of the image based factor FTI View Difficulty. The TIP library used in this study consisted of 376 threat objects, whereof each was depicted in up to 72 views. This number results from 6 basic views x 4 flip versions (no flip, horizontal flip, vertical flip, horizontal and vertical flip) x 3 plane rotations. The six basic views consist of the most typical/canonical view itself, rotated by 45˚ and 85˚ around the horizontal and the vertical axis and a combination of the two 45˚ rotations. Some of the threat items are rotational solids, which can be depicted only in up to 36 views (3 basic views x 4 flip versions x 3 plane rotations = 36 views). The effects of flipping and plane rotation were not analyzed separately.

Theoretically expected characteristics regarding co-variation with human detection performance: FTI View Difficulty varies negatively with detection performance and positively with image difficulty.

Superposition (SP)

This image based factor refers to how much the pixel intensities at the location of the FTI in the threat bag image differ from the pixel intensities at the same location in the same bag without the FTI (non-threat bag image). The Superposition formula is shown below. $I_{SN}(x,y)$ denotes the pixel intensities of a threat image (SN: signal-plus-noise) and $I_N(x,y)$ denotes the pixel intensities of the corresponding harmless bag (N: noise). Pixel intensity values as described in this formula are transformed into values between 0 and 1 before squaring and retransformed into values between 0 and 255 again afterwards. The constant term *C* is freely selectable and has only aesthetic character by means of getting positive values if *C* is set to a value larger than the maximum Superposition value. For statistical analyses it can be set to zero.

$$\mathrm{SP} = C - \sqrt{\sum_{x,y} \left(I_{SN}(x,y) - I_N(x,y)\right)^2}$$

In short: Superposition equals the negative Euclidian distance between the SN images (signal plus noise or threat) and N images (noise or non-threat images) regarding pixel intensity values.

Note: This mathematical definition of Superposition is dependent on the size of the FTI in the bag. For further development of the computational model it is conceivable to split up superposition and the FTI Size of the threat item into two separate image based factors.

Theoretically expected characteristics regarding co-variation with human detection performance: Superposition varies negatively with the detection performance and positively with image difficulty.

Bag Complexity (BC)

As explained in the introduction, the original image based factors View Difficulty, Superposition and Bag Complexity were proposed by Schwaninger (2003b) and Schwaninger et al. (2005). In order to implement this concept technically, Schwaninger et al. (2007) decided to increase implementability by split up the factor Bag Complexity into the factors Opacity and Clutter. Both Opacity and Clutter represent different aspects of the original concept Bag Complexity.

Please note that the name of the image based factor Transparency known from earlier studies was changed into the name Opacity. It is exactly the same measure, but the factor's meaning is inverted in order to facilitate data interpretation.

Opacity (OP)

The image based factor Opacity reflects the extent to which x-rays are able to penetrate objects in a bag. This depends on the specific material density of these objects. These attributes are represented in x-ray images as different degrees of luminosity. Heavy metallic materials such as lead are known to be very hard to be penetrated by x-rays and therefore appear as dark areas on the x-ray images.

$$\mathrm{OP} = \frac{\sum_{x,y} \left(I_N(x,y) < 64\right)}{\mathrm{BS}}$$

The formula above shows the image measurement formula for Opacity. $I_N(x,y)$ denotes the pixel intensities of the harmless bag. **64** is the pixel intensity threshold beneath which the pixels are counted. The implementation of the image measurement for the image based factor Opacity is simply achieved by counting the number of pixels being darker than a certain threshold (e.g. < 64) relative to the bag's overall size. **BS** represents the formula of the image based factor Bag Size described further down. It is used here to standardize the Opacity value on bag size.

Theoretically expected characteristics regarding co-variation with human detection performance: Opacity varies negatively with detection performance and positively with image difficulty.

The used Opacity intensity threshold **64** was evaluated by maximizing the correlation between the Opacity value and d' (a signal detection performance measure by Green and Swets (1966)) on an X-Ray ORT data set from twelve undergraduate students. For more information on the X-Ray Object Recognition Test (X-Ray ORT) please consult Schwaninger, Hardmeier, and Hofer (2005))

Clutter (CL)

This image based factor is designed to express bag item properties like textural unsteadiness, disarrangement, chaos or just clutter. In terms of the bag images, this factor is closely related to the amount of items in the bag as well as to their structures in terms of complexity and fineness. The method used in this study is based on the assumption, that such texture unsteadiness can be described mathematically in terms of the amount of high frequency regions.

$$\mathrm{CL} = \frac{\sum_{x,y} I_{hp}(x,y)}{\mathrm{BS}}$$

$$\text{where} \quad I_{hp}(x,y) = I_N * \mathcal{F}^{-1}(hp(f_x, f_y)) = \mathcal{F}^{-1}(\mathcal{F}(I_N \cdot hp(f_x, f_y))$$

$$\text{and} \quad hp(f_x, f_y) = 1 - \frac{1}{1 + \left(\frac{\sqrt{f_x^2 + f_y^2}}{f}\right)^d}$$

The above formula shows the image measurement formula for Clutter. It represents a convolution of the empty bag image (N for noise) with the convolution kernel derived from a high-pass filter in Fourier space. I_N denotes the pixel intensities of the harmless bag image. F^{-1} denotes the inverse Fourier transformation. $hp(f_x,f_y)$ represents a high-pass filter in Fourier space. **BS** represents Bag Size as described in the next section. Cut-off frequency f and transition d (the filter's order) were set to $f = 0.03$ and $d = 11$. The pixel summation on the high pass filtered image was restricted to the bag's area.

Theoretically expected characteristics regarding co-variation with human detection performance: Clutter varies negatively with the detection performance and positively with image difficulty.

Bag Size (BS)

The Bag Size formula below shows is applicable to grayscale images represented by pixel intensity values between 0 (black) and 255 (white). All pixels with pixel intensities lower than 254 (near white) are counted and summed up.

$$\mathrm{BS} = \sum_{x,y} (I_N(x,y) < 254)$$

The process of how the intensity threshold was determined is described in detail in the appendix at the end of this technical report.

Theoretically expected characteristics regarding co-variation with human detection performance: Bag Size varies negatively with the detection performance and positively with image difficulty (if there is a significant relationship).

General Remarks

It should be noted that the theoretically expected characteristics at the end of each section have been empirically confirmed for FTI View Difficulty, Superposition, and Opacity (e.g. Schwaninger, et al., 2005, 2007; Hardmeier et al. 2006b). However, for Clutter ambiguous weak effects found and the effect of Bag Size is currently being investigated. Additionally, earlier implementations of Clutter were not standardized by Bag Size, so that earlier results cannot be compared directly with the results we are expecting for future studies.

References

Green, D.M., & Swets, J.A. (1966). *Signal detection theory and psychophysics*. New York: Wiley, 1966.

Hardmeier, D., Hofer, F., & Schwaninger, A. (2006a). Increased detection performance in airport security screening using the X-Ray ORT as pre-employment assessment tool. Proceedings of the *2nd International Conference on Research in Air Transportation, ICRAT 2006,* Belgrade, Serbia and Montenegro, June 24-28, 2006, 393-397.

Koller, S.M., Hardmeier, D., Michel, S., & Schwaninger, A. (2007). Investigating training and transfer effects resulting from recurrent CBT of x-ray image interpretation. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th Annual Cognitive Science Society* (pp. 1181-1186). Austin, TX: Cognitive Science Society.

Koller, S.M., Hardmeier, D., Michel, S., & Schwaninger, A. (in press). Investigating training, transfer, and viewpoint effects resulting from recurrent CBT of x-ray image interpretation. Journal of Transportation Security, in press.

Schwaninger, A. (2003b). Evaluation and selection of airport security screeners. *AIRPORT*, 02/2003, 14-15.

Schwaninger, A. (2004b). Computer based training: a powerful tool to the enhancement of human factors. *Aviation Security International*, FEB/2004, 31-36.

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

Schwaninger, A., Hardmeier, D., & Hofer, F. (2005). Aviation security screeners visual abilities & visual knowledge measurement. *IEEE Aerospace and Electronic Systems, 20(6),* 29-35.

Schwaninger, A., Michel, S., & Bolfing, A. (2007). A statistical approach for image difficulty estimation in x-ray screening using image measurements. *Proceedings of the 4th Symposium on Applied Perception in Graphics and Visualization, ACM Press, New York, USA*, 123-130.

Appendix

Determination of the pixel intensity threshold for differentiation between bags and their backgrounds

Due to some minor artifacts in x-ray image capturing, differentiation of bags from their background cannot be simply calculated by assuming the bags to be simply non-white ($I<255$; $I\neq255$). The following figure shows some sample images of three different bags. The first row contains the actual bag images, while each of the other rows corresponds to one of the (strict) thresholds 234 (which is equal to 0.92 on a scale between 0 and 1), 253, 254 and 255, respectively. For our implementation we have chosen the threshold of 254 pixels, which is marked by a red rectangle.