1

Introduction

When we travel by air or boat our baggage has to be checked because the safety of other passengers may be at stake. To preserve the passengers' privacy and because of the high cost of checking baggage manually, the preferred option is usually to use x-ray scanners. Figure 1.1 shows a typical x-ray image. Most objects are transparent to x-rays, but very dense metals, such as lead, completely absorb x-rays.

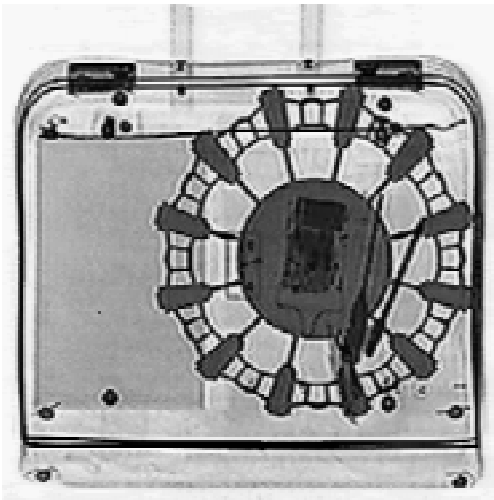

Figure 1.1 (**see colour figure on left cover flap**). Example of x-ray scan of suitcase containing a clock, a book and a tin opener. Colours (see Figure 8.2 on page 134) indicate the materials.

The baggage inspector has a hard job, as he has a large responsibility but little time: usually only about 6 seconds to decide whether a suitcase is safe or not. Guns and knives are relatively easy to detect, as they partly absorb and partly scatter x-rays and have a characteristic shape. But bombs are much harder to detect, as the typical parts of a bomb can be hidden easily: plastic explosives can have any shape, a delay mechanism can be extremely small, the detonator may be placed in such an orientation that it is hardly visible and very thin wires may not be visible at all. Furthermore, X-ray images get cluttered if many items are stacked on top of each other. Finally, some objects may completely absorb x-rays, thereby making things behind them invisible.

The Delft Virtual Window System (DVWS)

In order to unclutter stacked objects, to allow looking behind objects which absorb x-rays and to enhance image resolution, a spatial impression of the baggage contents may be given to the inspector instead of a single x-ray image. This thesis is about the optimization of baggage inspection by means of the Delft Virtual Window System (*DVWS*), a display technique that can give the observer a spatial impression of a scene. The DVWS was developed and patented by Smets, Overbeeke and Stratmann, and is described in detail in the patents (Smets, Stratmann and Overbeeke, 1988, 1990) and in the doctoral dissertation of Overbeeke and Stratmann (1988). Its potential for applications is described by Smets (1995).

The DVWS consists of a monitor, a viewpoint position tracker and a camera that looks at a scene (Figure 1.2). The system gives the observer a spatial impression of the scene on the monitor display by coupling the camera position to the viewing position of the observer. I will refer to an image belonging to a camera position as a *view*. For single pictures out of their 3D context I will use the word *image*, while I will use the word *view* for a single picture within the 3D context. To prevent the point of interest, the *fixation point*, from shifting off the screen as the observer moves, the camera rotates in such a way that it remains aimed at the fixation point. This coupling can be made for the left-right, up-down and forward-backward movements of the observer. In practice, the precise coupling between eye position and camera position differs slightly over implementations, and will be described for each experiment.

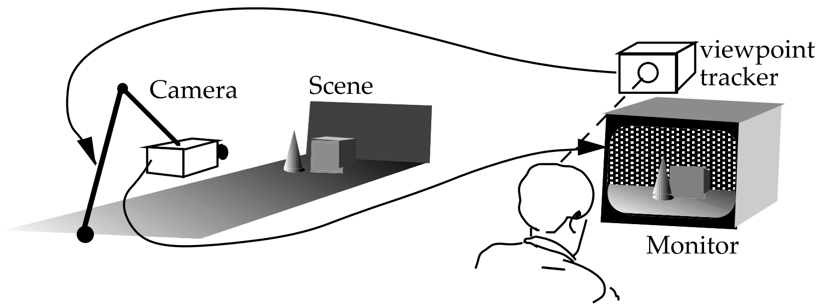

Figure 1.2. Principle of virtual window displays. The camera position and orientation are coupled to the observer's viewpoints. For the observer, the scene on the display becomes spatial.

Applying the DVWS to baggage inspection

For x-ray baggage inspection, the setup of Figure 1.2 is not practical: for each new view a new x-ray photo has to be taken. The observer can take an infinite number of positions (Figure 1.3a), and therefore the number of views is potentially infinite, but the baggage contents may be damaged if more than about 25 x-ray photos are taken. In this thesis, an attempt is made to solve this problem by restricting the selectable views (A) to the horizontal arc (Figure 1.3b) and (B) to certain discrete positions along that arc (Figure 1.3c). As the number of selectable views is reduced, a mismatch is introduced between the actual observer's position and the displayed view. Either the distance between extreme viewpoints (*camera range*) will decrease or the angular distance between the available views (*angle between views*) will increase. This raises questions about the consequences of this reduction on the quality of the inspection.

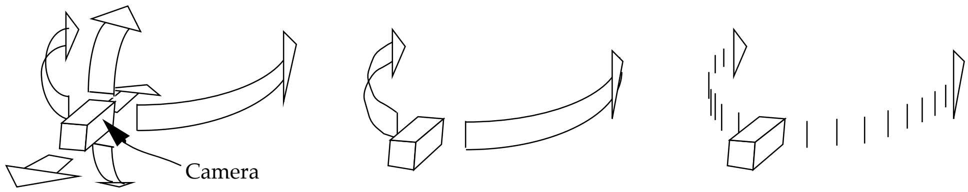

Figure 1.3a. In the unrestricted DVWS the number of available views is infinite in 3 dimensions.

Figure 1.3b. To limit the number of views, camera motion is restricted to the horizontal arc.

Figure 1.3c. To make the number of views finite, camera motion is made discrete.

In the building of a prototype system, other factors are also important, particularly the cost of the system and the best way of achieving interaction between the machine and the operator. In order to find the most cost-effective solution, Chapter 3 outlines a number of technical solutions that are able to record the required views. These solutions differ in the time required to record the required views, the number of x-ray cameras and x-ray sources, the size of the apparatus, the number of moving mechanical parts and, most important for the inspection quality, how the views are taken. For the man-machine interface, ergonomic aspects have to be considered. The ergonomic aspects of the eye position tracking mechanism need special attention, as it is not ergonomically acceptable to have operators moving around the display all day. Furthermore, baggage inspectors prefer not to wear special reflectors or mechanical tracking constructions on their head. An attempt was therefore made to replace viewpoint selection via eye position with viewpoint selection by a knob. As with restricting the number of available views, this may affect the quality of the inspection.

Previous work

The task of the inspector is different for inspecting hand baggage and inspecting hold baggage. These differences became clear after the first experiment (Chapter 4) and in the course of contacts with an airport. When inspecting hand baggage, the inspector tries to recognise all objects in order to detect the presence of dangerous items such as guns or knives. When inspecting hold baggage he tries to find parts of a bomb. For checking hand baggage, cheaper inspection machines are used that give little information to the inspector and thus necessitate manual inspection, because manually inspecting baggage of the passengers is part of security. A suitcase may be opened only in the presence of its owner, and therefore inspectors are not allowed to open hold baggage. For these reasons it was decided to concentrate on hold baggage in this thesis.

According to the inspectors, a ‘standard’ bomb consists of a battery, a delay mechanism, a detonator, explosives and wires connecting these parts. The operator tries to match these components to the actual baggage contents, and if they do not match the suitcase may be safe. This is quite a ‘fuzzy’ process, as the components of a bomb have no standard shape, and because some components are not essential. Therefore there is a large variation among inspectors. There is similar uncertainty about what makes a good inspector for medical x-ray screening (Bass and Chiles, 1990). Nodine, Kundel, Lauver and Toto (1996) have shown that in the case of medical x-ray images it takes a huge number of trials before radiologists know what a normal situation looks like. For baggage inspection, much more variation exists in normal situations, thus suggesting an even longer learning curve. Nevertheless, I assume that wires between objects provide an important clue. In 1983, the FAA required scanners to be able to display wires of 0.5mm diameter (Tsacoumis, 1983), and at that time most x-ray equipment was able to display wires with a diameter of 0.16 mm diameter (Dorey, 1983). Currently, most machines are able to display wires down to 0.1 mm diameter.

Much research on x-ray baggage inspection is classified, as depriving potential terrorists from information about the inspection process is part of the security. This gives problems for the precise description of the baggage inspection task. Furthermore, the inspection task itself seems to be not strictly defined (*operationalized*) either. It is not clear precisely what baggage inspectors are looking for: the baggage inspection task is not well *operationalized*. Most baggage inspectors will explain that they are looking for parts of a bomb, but often items resembling such parts are present, while the inspectors do not mark them as parts of bombs. This problem with the operationalisation of the task seemed quite irrelevant at the beginning of the project. However, this problem makes the implications of research on tasks that seem related to baggage inspection uncertain, and as the project progressed it became clear that it was of crucial importance. Given this knowledge, an analysis of the baggage inspection task would be an appropriate first step, but as observed above the importance of this question was underestimated.

Below, previous work that set the direction for the research of this thesis is discussed. The work in question is split into two parts: a technological part that describes how x-ray images can be made, and a perceptual part that describes the aspects of the display system that were expected to affect the performance of the inspector. The way the images are made does not dictate the way they are presented. In principle, a conventional scanner can be adapted in a straighforward way to provide views for presentation with the DVWS, by rotating the baggage in the required orientation and storing the acquired views. The important questions here are about the trade-off between scanner price, the amount of exposure of the baggage to x-rays, and the inspectors' performance.

Previous work - technological

The technological aspects of x-ray baggage inspection can be split into three categories: the technology to record the required number of selectable views, the technology for recognising aspects of the material in the baggage and the technology that analyses the visual cues in the images.

Currently, conventional x-ray baggage scanners use a fan-shaped x-ray beam shining on a row of x-ray sensors: the *sensor line* (Figure 1.4a). The baggage is pulled through the fan-shaped beam, and the baggage is scanned line by line. These lines are put together on the display to form the x-ray image. To present the inspector a spatial image of the baggage, a conventional scanner can be adapted in a straightforward way (Evans, Godber and Robinson, 1994). Figure 1.4b shows a scanner that can record two images for *stereoscopic* viewing (where images taken from slightly different viewpoints are presented to the left and the right eye).

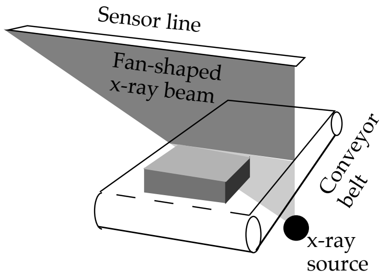

Figure 1.4a. Conventional x-ray baggage scanner. The baggage is pulled through a fan-shaped x-ray beam, and the scan lines are collected to form the x-ray image.

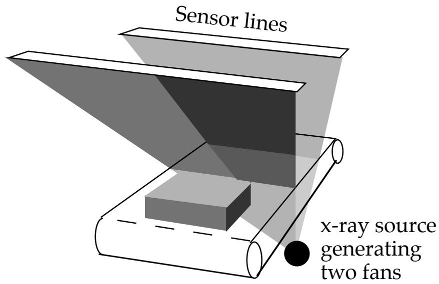

Figure 1.4b. As Figure 1.4a, but two images are scanned simultaneously.

With the machine made by Isorad (Isorad, 1988) the baggage can be rotated during scanning. This enables the inspector to look around x-ray absorbing objects and to select a view that provides sufficient cues to decide whether a suitcase is safe. This system is rarely used, probably because it is designed for careful inspection of a few suitcases and because it exposes the baggage to a continuous beam of x-rays. In current practice, sometimes the baggage is rescanned in a rotated orientation with a conventional scanner by placing a piece of foam under the baggage, but the amount of rotation is limited by the height of the scanning tunnel. Rescanning usually takes too much time, and therefore a scanning machine exists (EG&G Astrophysics, 1996) that always takes two images: one front view and one at a 60° angle (see Figure 3.7 on page 43). Finally, a complete spatial reconstruction of the baggage can be made with a CT scanner (Henderson, 1990; Invision, 1997). With such a reconstruction any desired view or cross-section of the baggage can be inspected. However these machines are rare, as they are expensive, bulky and slow, and the scanning tunnel is too small to fit all baggage (Henderson, 1990). Currently, Invision and EG&G Astrophysics are cooperating (Dotzler, 1996; InVision & EG&G, 1997) to improve the scanning speed of the CT scanners, and hope to attain the same speed as the fastest line scanners (1500 bags/hour).

The second category is the way the properties of the materials in the baggage are derived. The density of the material can be derived from the amount of x-ray absorbance and scattering, and suspicious densities (near the density of nitrogen, the principal constituent of explosives) can be detected. By scanning with two x-rays with different energy levels, organic, inorganic and metallic materials can be distinguished. This material information is usually merged into the image by using pseudo-colours, where each colour represents a particular class of material (Figure 1.1, see also Figure 8.2 on page 134).

The third category is automatic object recognition. Especially for bomb detection it is useful to detect automatically the presence of objects resembling detonators and batteries. Such information is usually merged into the image by adding arrows or other signs pointing to the suspicious items, and flashing a warning sign or sounding a beep.

Usually, baggage inspection is not done solely based on x-ray images. For detecting explosives, nitrogen-detection can be carried out more precisely with neutron scanners or gamma ray scanners than with x-ray scanners. Neutron scanners determine the way neutrons are absorbed and what radiation is generated when the baggage is exposed to neutrons. Gamma ray machines expose the baggage to gamma rays with an energy that is absorbed highly by nitrogen. These machines are expensive because of the necessary shielding and short-lived neutron- or gamma-ray sources. Another important technique is vapour detection. Even a few evaporated atoms from explosives can be detected, but explosives evaporate very little. Finally, in high-risk flights the owners of the baggage are screened and judged on the basis of an interview.

Previous work - perceptual

I am looking for the combination of viewpoints and image quality that gives optimal inspector performance. This section discusses relevant literature to sketch our major expectations and gaps in our knowledge that have to be filled.

Image resolution and number of grey levels are known to have an effect on the response time when manipulating remote objects (Ranadivé, 1979). Observers controlling the camera motion can cope with much lower resolution than observers not having this control (Smets and Overbeeke, 1995). These results suggest an interaction between image resolution, number of grey levels and amount of control over the camera. For several inspection tasks, performance has been shown to improve with increasing numbers of available views. However, Kersten and Bülthoff (1991) reported that transparency can interfere with the human bias to see objects as being rigid, and therefore the results mentioned above may not hold for transparent scenes such as x-ray images.

I expect that providing multiple views improves the spatial impression of the baggage. Geometrically, it is possible to reconstruct a scene from two distinct views (Longuet-Higgins, 1981) although some assumptions may be required (Ullman, 1979). With perspective assumptions even a single view may contain sufficient spatial cues to reconstruct the scene up to a scaling factor (Halloran, 1989). However, experiments testing human performance with sparse spatial scenes (e.g., Braunstein, Hoffman, Shapiro, Andersen and Bennett, 1987) suggest that human accuracy in reconstructing the 3D scene from distinct views will increase with more distinct views, more points and with increased constraints, even when fewer or more viewpoints are provided than are geometrically necessary. Furthermore, providing multiple views may enhance scene rigidity (Todd, Akerstrom, Reichel and Hayes, 1988; Todd and Bressan, 1990).

For tasks with real scenes, an increasing number of available views can also increase observer performance. For instance, multiple images may resolve ambiguities in cluttered parts of the image such as the so-called 'camouflage effects' (Nodine and Kundel, 1987; Vyborny, 1997). For breast cancer screening with x-ray images, Wald, Murphy, Major, Parkes, Townsend and Frost (1995) showed that, compared with a single view, two mammographs increase the chances of detecting a cancer. Evidence exists that humans remember a number of views of a spatial object, and use them for recognition (Perrett, Harries and Loker, 1992). Bülthoff and Edelman (1992) showed an increase in recognition errors as the available view gets further away from the learned view. Making multiple views available will increase the chance that a view close to such a learned view is available, and may therefore improve observer performance. Multiple views are expected to compensate for low image resolution (Smets and Overbeeke, 1995). When presenting multiple views, an intuitive way to select a desired view seems essential to improve performance (e.g., Diner and Venema, 1989), and a low number of fixed viewpoints always seems to be suboptimal (Gaver, Sellen, Heath and Luff, 1993).

The accuracy of the coupling between the camera position and the actual viewing position of the observer may also have an effect on performance. If these positions do not match, the apparent sizes in the scene may not match the sizes as might be measured with a measuring rod (*distortion*). For static images, Lumsden (1980) showed that the apparent layout is affected by the viewing distance. Halloran (1989) showed that viewpoint displacements parallel to the display may also cause distortion. In the case of film, Meister (1966) made an analysis of places with 'acceptable' distortion, but he bases his results on geometry and not on experimental results. To prevent such distortions caused by an inappropriate viewpoint, Gibson (1971) indicated that perspective pictures should be looked at with one eye, and that a reduction screen should be placed in front of the display to hide the rest of the world. All these results implicitly assume that human perception relies strongly on the geometry of the scene. However, perspective is only one of the many available depth cues, and distortion will not depend on geometry alone. In searching the literature on distortion I found no studies where the picture's perspective is coupled to movements of the observer (as is the case with the DVWS). Chapter 7 decribes an experiment which tested whether and to what extent perspective distortion occurs in virtual window systems.

The aim of this thesis is to find the requirements for enhancing x-ray baggage inspection with the DVWS, and to propose technical solutions that fulfill the requirements. There are a lot of cross-relations between the technical possibilities, technical solutions and investigated research questions. In order to arrive at an overview, it may be useful to read the abstract at the back of this thesis.

The effects of static image quality and the number of available views on the observer performance will be investigated for transparent scenes in experiments described in Chapters 4 to 6. The effects of the accuracy of the coupling between the camera position and the actual viewing position, and other parameters that might introduce distortion, were investigated in the experiment described in Chapter 7. Because looking with one eye, as suggested by Gibson (1971), may meet resistance from baggage inspectors, the implications of looking with both eyes was also investigated in the experiment described in Chapter 6. For similar reasons, replacement of viewpoint selection via the eye position with manual viewpoint selection was investigated in the experiment described in Chapter 5.

With the results of the experiments of Chapters 4 to 7 and the technical discussion of Chapter 3, technical choices were made for a prototype system for baggage x-ray inspection with the DVWS. X-ray photos of real baggage were made in accordance with these choices. Chapter 8 describes an experimental test of the prototype system. Chapter 9 concludes with a discussion of the relevance of the results for x-ray baggage inspection, virtual window displays, Industrial Design Engineering and perceptual theories.