Visual detection of knives in security applications using Active Appearance Models

Andrzej Glowacz · Marcin Kmieć · Andrzej Dziech

Abstract In this paper, a novel application of Active Appearance Models to detecting knives in images is presented. In contrast to its popular applications in face segmentation and medical image analysis, we not only use this computer vision algorithm to locate an object that is known to exist in an analysed image, but–using an interest point typical of knives–also try to identify whether or not a knife exists in the image in question. We propose an entire detection scheme and examine its performance on a sample test set. The work presented in this paper aims to create a robust visual knife-detector to be used in security applications.

1 Introduction

One application of object detection in images is computer-aided video surveillance. It can be used both in security applications [15] and as legal evidence [14]. A CCTV operator usually monitors multiple video feeds at the same time; this is a complex and challenging task in terms of allocating attention effectively. One study suggests that detection rates for operators monitoring four, nine and 16 screens oscillate around 83 %, 84 % and 64 % respectively, and will drop significantly after an hour [23]. Therefore, the need to automate the process is obvious. There have been attempts at detecting suspicious events in video material [11] and recognising human activity in videos [13].

This study focuses on automatic detection of knives in images. Carrying knives in public is either forbidden or restricted in many countries; due to the fact that knives are both widely available and can be used as weapons, their detection is of high importance for security

A. Glowacz · M. Kmieć (✉) · A. Dziech
AGH University of Science and Technology, al. Mickiewicza 30, 30-059 Kraków, Poland
e-mail: kmiec@agh.edu.pl

personnel. For instance, the idea of software-based knife-detection has a practical application in surveillance of the public using CCTV. Should a knife be detected, an alarm is raised and the human operator can immediately focus their attention on that very scene, and either confirm or reject that detection. Although people will almost always outperform software algorithms for object detection in images [21], in the long run the computer could be of significant assistance to the human CCTV operator when it comes to dealing with tens of simultaneous video feeds for many hours a day. Another application of automatic knife-detection is computer-aided analysis of luggage x-ray scans. Visual detection in security applications approach is a new research area. Visual detectors designed to work in video surveillance or x-ray scanners are not widely available.

Knives are a very wide class of objects of immense diversity. Moreover, they easily reflect light, which reduces their visibility in video sequences; automatic knife-detection in images therefore represents a challenging task. In this paper, a novel application of the well-established Active Appearance Models (AAMs) is presented. So far, these have been extensively used for medical image interpretation [1] [25] [12], and for the matching and tracking of faces [10] [7]. Among many existing shape-modelling algorithms, such as the Active Contour Models (Snakes), we have focused on AAMs because they model not only the shape but also the appearance (that is, pixel intensities within the image region bounded by the shape). As the knife-blade typically possesses quite a uniform texture, modelling its appearance should contribute to the general resistance of the model so that it does not converge to objects that have a shape similar to that of the knife-blade.

The novelty of this work is twofold. Not only has there been (to the best knowledge of the authors) no other research on knife-detection, but also AAMs have so far not been used to detect objects belonging to a general class. They have been used in what is referred to as 'detection'–as in [12]–but that is not what is meant by 'detection' in computer vision, in the strict sense of the word. By detecting, for example, a face in an image, we mean answering the question of whether there is or there is not a face in the given image [22]. This process can be characterised by two parameters: the positive and the negative detection rates. As in the case of [12], before what is referred to as 'detection' is performed, an assumption exists that the object is somewhere in the image, and the task is to precisely locate it. For instance, given an image of a face, finding the nose is not a task of detection since we can assume that all faces have noses. It is, rather, the task of location, and can be characterised by the level of localisation accuracy, but not by positive and negative detection rates. In this case, the assumption is that there always is a face in the analysed image. Should the AAM be performed on a non-face image, it would converge to the parts of the image whose appearance is closest to its model. This is still theoretically correct, but makes no sense from a practical point of view. Moreover, AAMs are sensitive to the initial location of their landmark points in the analysed image. Even if there is a face somewhere in a large image, for the algorithm to correctly segment it into elements, the initial location of its landmark points needs to be roughly around the face region. A common technique for face segmentation with AAMs is the use of Viola and Jones's face detector [16] to initialise the AAM as in [6].

In this paper, a method for detection of objects, in this case knives, using AAMs is introduced. It aims to answer the question of whether or not there a knife exists in the given image.

2 Active Appearance Models

AAMs were introduced [4] in 1998 as a generalisation of the popular Active Contour Model (Snake) and Active Shape Model algorithms. They are a learning-based method, which was

originally applied to interpret face images. Due to their generic nature and high effectiveness in locating objects, numerous applications in medical image interpretation followed. The typical medical application is finding an object, usually an organ in a medical image of a specific body part, such as locating the bone in a magnetic resonance image (MRI) of the knee [5], the left and right ventricles in a cardiac MRI [18], or the heart in a chest radiograph [24].

In general, an AAM can be described as a statistical model of the shape and pixel intensities (texture) across the object. The term 'appearance' means the combination of shape and texture, while the word 'active' describes the use of an algorithm that fits the shape and texture model in new images. In the training phase, objects of interest are manually labelled in images from the training set with so-called landmark points to define their shape. A set of three annotated knives has been depicted in Fig. 1.

Objects are thus defined by the pixel intensities within their shape polygon and the shape itself. Principle Component Analysis (PCA) is then employed to model variability between objects in images from the training set. Before that happens, all shape polygons need to be aligned to the normalised common mean. This is achieved by the means of Generalised Procrustes Analysis, a method widely used in statistical shape analysis. The algorithm outline can be summarised in four steps [20]:

1. Choose an initial reference shape (e.g. the first object's shape).
2. Align all other shapes to the reference shape.
3. Re-calculate the mean shape of the aligned shapes.
4. If the distance of the mean shape to the reference one is larger than a threshold, set mean shape as the reference shape and go back to point 2. Otherwise, return the mean shape.

The distance in point 4 is simply the square root of the sum of squared distances between corresponding points in two considered shapes, and is known as the 'Procrustes distance'. Superimposing two shapes includes translating, rotating and uniformly scaling objects, and is also a four step process:

1. Calculate the centres of gravity (COGs) both of the mean and the shape being superimposed to the mean.
2. Rescale the shape being superimposed so that its size equals that of the mean.
3. Align the position of the two COGs.
4. Rotate the shape being superimposed so that the Procrustes distance is minimal.

Fig. 1 Images of knives with landmark points annotated
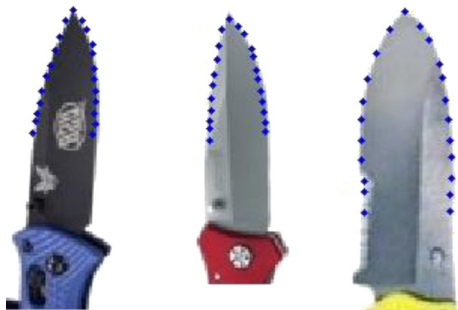
By the means of PCA performed on the shape data, we obtain the following formula, which can approximate any shape example:

$$x = \bar{x} + P_s b_s \qquad (1)$$

where $\bar{x}$ is the mean shape, $\boldsymbol{P}_s$ are the eigenvectors of the shape dispersion, and $\boldsymbol{b}_s$ is a shape parameters set. For the statistical model of appearance to be built, we need to warp all images in the training set so that their landmark points match the mean shape. Texture warped to match the mean shape is referred to as 'appearance' in the original paper [4]. Let us denote grey-level information from the shape-normalised image within the mean shape as $\boldsymbol{g}_m$. To reduce lighting variation, it is necessary to normalise pixel intensities with an offset $\boldsymbol{\Psi}$ and apply a scaling ξ:

$$g = (g_{im} - \psi) \big/ \xi \qquad (2)$$

The values of the two parameters are obtained in the course of a recursive process, details of which can be found in [4]. Once pixel intensities of samples in the training set have been normalised, a statistical model of pixel intensities can be defined:

$$g = \bar{g} + P_g b_g \qquad (3)$$

where $\bar{g}$ is the normalised mean pixel-intensity vector, $\boldsymbol{P}_g$ are the eigenvectors of the pixel-intensity dispersion and $\boldsymbol{b}_g$ is a parameter set. For each image in the training set, a concatenated vector **b** is generated:

$$b = \begin{bmatrix} W_s b_s \\ b_g \end{bmatrix} = \begin{bmatrix} W_s P_s^T (x - \bar{x}) \\ P_g^T (g - \bar{g}) \end{bmatrix} \qquad (4)$$

Now another PCA is applied on the concatenated vector, the result of which is as follows:

$$b = P_c c, \quad P_c = \begin{bmatrix} P_{c,s} \\ P_{c,g} \end{bmatrix} \qquad (5)$$

c is a parameter set that controls both the shape and the appearance, and $\boldsymbol{P}_g$ are the eigenvectors. New object instances can be generated by altering the $\boldsymbol{c}$ parameter, and shape and pixel intensities for a given $\boldsymbol{c}$ can be calculated as follows:

$$x = \bar{x} + P_s W_s P_{c,s} c, \quad g = \bar{g} + P_g P_{c,g} c \qquad (6)$$

W_s is a matrix of weights between pixel distances and their intensities.

3 Corner detection

It is a characteristic of all knives that the tip of the knife is distinct, lies at the intersection of two edges, and can be regarded as a corner that is a point that has two different, dominant-edge directions in its neighbourhood [3]. Multiple corner detectors are used in computer vision; for the purpose of tip-of-the-knife detection, we have used the Harris corner-detection algorithm [9]. Such local features have been shown to be well suited to matching and recognition [17] [19]. This popular interest-point detector is invariant to geometric transformation and, to a large extent, to illumination changes–as well as being resistant to

image noise. It is based on the simple principle that edges and corners change noticeably more than other elements of the image, with a window shifted a little in all directions. The difference in pixel intensities caused by a window shift (Δ_x, Δ_y) is measured by the means of the local auto-correlation function [9]:

$$c(x,y) = \sum_W \left[I(x_i, y_i) - I(x_i + \Delta_x, y_i + \Delta_y) \right]^2 \tag{7}$$

where $I(x_i, y_i)$ denotes greyscale pixel-intensity at the given location in the Gaussian window W centred on (x, y). By approximating the shifted image with a Taylor series, it can be proved [9] that expression (7) is equal to:

$$c(x,y) = \begin{bmatrix} \Delta_x & \Delta_y \end{bmatrix} C(x,y) \begin{bmatrix} \Delta_x \\ \Delta_y \end{bmatrix} \tag{8}$$

where is C a matrix that summarises the predominant directions of the gradient in the specified neighbourhood of the point (x, y). The eigenvalues of the matrix C form a description of the pixel in question; if they are both high, they indicate a corner.

In Fig. 2, three knife images with corners detected are shown. The number of detected corners depends on the threshold value used. What we are interested in is detection of the tip of the knife. Therefore, the threshold needs to be set reasonably low to ensure that the tip of the knife is detected as a corner, even at the expense of more corners being detected. The Harris corner detector implemented in OpenCV 2.4 [2], which we used for corner detection, has a threshold for detected corners in the range 0 to 255. The higher the threshold, the fewer corners are detected. In order to determine the threshold at a level that, on one hand, will guarantee that no knife-tip will be left undesignated as a corner and that, on the other, will not generate too many false positives (in the sense of knife-tip detections), experiments were performed on a test set of 40 images containing one knife each. Three of these images are shown in Fig. 2. The test set is described in depth in the next section. The lowest threshold at which all knife tips were detected was 204, and the average border value of the threshold at which the knife-tip is designated as a corner for all images was 217. The border value for a given image generated on average 2.1 corners, which is in line with the assumption that the tip is a strong corner. The lowest threshold of 204 generated on average 7.8 additional corners for each image. Due to the fact that each unnecessary corner, i.e. a corner that is not a knife-tip, leads to unnecessary computational effort and therefore wastes time in the detection scheme described in the next section, it is crucial to limit the number of detected corners. The threshold value for the Harris corner detector that is employed in our approach is set so that only three corners are detected. Since the knife-tip is likely to be a strong corner, this approach reduces calculation time for the whole detection scheme.

**Fig. 2** Interest points in knife images detected by the Harris corner detector

4 Overcoming the AAM's variance to rotation and scale change for detection purposes

AAMs are not invariant to rotation. An AAM will converge only to objects that are oriented at the same angle as the objects in the images it was trained for. If we want it to converge to an object of the class it was trained for, but placed at an arbitrary angle in an image, one obvious workaround is to rotate that image until the object considered is placed at a right angle. This is a plausible solution, but a drawback to this approach is the computational effort needed to rotate the input image. For example, rotating a 640×480 pixel image by 15° takes some 10 ms on a modern PC (Intel(R) Core(TM) i5-3450 CPU @ 3.10GHz, 8GB RAM), so a whole rotation by 360° with a 15-° angle step will take some 240 ms.

Instead of rotating the input image by 360° at a fixed angle step, different AAMs could be trained for different knife orientations. From a computational point of view, this approach has a significant computational advantage over the former, as it requires just as many AAM single searches, while it does not require rotation of the input image. Hence, this approach of using multiple AAM trained to locate objects in different orientations, i.e. at different angles, has been chosen.

Of course, AAMs are not invariant to scale change. Experiments [10] [7] suggest that AAMs will correctly converge to objects of varying size, as parts of the face may differ in size between different people, but that this requires proper AAM initialisation, for instance with a face detector of a different kind [16], as mentioned in the intro-duction. In the case of knives, the effects of scale change can, however, be neglected, due to the specific shape of the knife-blade. The actual shape does not change dramatically with scale. This has been illustrated in Fig. 3. In the picture, we can see two knife-blade outlines; the second is twice as big dimension-wise as the first. The depicted situation illustrates applying the same AAM to the two objects. In the case of the smaller outline, the AAM roughly encompasses half of the contour's length, whereas in the case of the larger outline, the same AAM encompasses less than a third of the blade's length, yet theoretically still converges to it. The main advantage of this approach is the lack of need of subsampling the input image, which basically eradicates all the inconvenience associated with it. First, it reduces the overall computational cost and, second, there is no need to interpret positive results from slightly varying scales. It may well happen that the same object in slightly different scales will yield positive-detection results and these will need to be interpreted, because it may be unclear whether there is just one actual object in the analysed image, or in fact more. A sophisticated solution to interpreting such results utilising an adaptive mean-shift algorithm with kernel density estimation has been proposed in [8], and could well be adopted in the case of AAMs applied to object detection, if image subsampling was used. But since that is not the case, it illustrates only how much computational effort has been spared.

Fig. 3 Blade-shape similarity in different scales
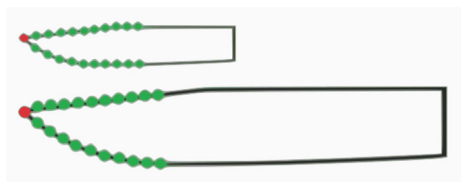
5 Detection scheme and results

As stated in the introduction, AAMs have been used to locate objects in images and–in our application–to detect objects. In this section, the procedure of object detection utilising AAM is described. It is based on the assumption that if a knife exists in the analysed image, its tip will be designated as a corner by the Harris corner detector. All the designated points will be used to initialise AAMs trained to locate knives. The results of running the Harris corner detector on knife images have been presented in Fig. 2. We can see that the tip of the knife is highly likely to be designated as a corner.

The general principle of multiscale image search is that at each corner, AAMs trained for all knife orientations are run, and if at least one of them converges to the knife orientation it was trained for, from the initial location designated by the corner and from slightly varying locations, the detection result is positive. We have trained 24 AAMs for objects rotated by a 15-degree angle step to cover a full 360-degree rotation. Each AAM is composed of 25 landmark points defining the shape polygon.

Each AAM was trained using the same six images of knives rotated by a respective angle. Adding additional images to the training set brought little (if any) improvement to the AAM performance. We believe that this is due to the rather simple shape we are dealing with, in contrast to, for example, the shape of the human face.

A general outline of our knife-detection scheme is presented below:

1. Put the points designated by the Harris corner detector in the knife-tip candidate set. Choose the first point and remove it from the set. Choose the first AAM of the 24 AAMs trained for different object orientations.
2. Set the AAM's landmark corresponding to the knife-tip at the chosen point.
3. Run the chosen AAM, calculate the minimum-area bounding box of the landmark points that have now converged to a new location. The box's longer symmetry axis is considered as the knife-candidate's pseudo symmetry axis.
4. Calculate the percentage of landmark points that lie on the edges found in the image undergoing detection by the Canny edge detector.
5. Set the AAM's landmark point corresponding to the knife-tip to the following three locations: three pixels to the left of, above and to the right of the initial knife-tip, and perform steps 3 and 4 on them.
6. The result of detection is positive if the following conditions have been met:

    I.   In all four cases (the original corner and the three points in its neighbourhood), landmark points of the model converge to the same location.
    II.   In all four cases, the knife-candidate's pseudo symmetry axes have a similar skew-angle.
    III. Most of the landmark points lie on edges detected by the Canny edge detector.
          If the result is positive, stop the detection procedure. If at least one of the above conditions is not met, the result of detection is negative. In this case, continue to the next point.
7. Choose the next AAM, i.e. the AAM trained for the orientation increased by 15° in relation to the current AAM, and go to point 3. If all orientations have been covered, go to the next point. Theoretically multiple AAM searches could be performed in parallel in order to speed up the detection process.
8. If the knife-tip candidate set is not empty, choose the next point, remove it from the set and go back to point 2.
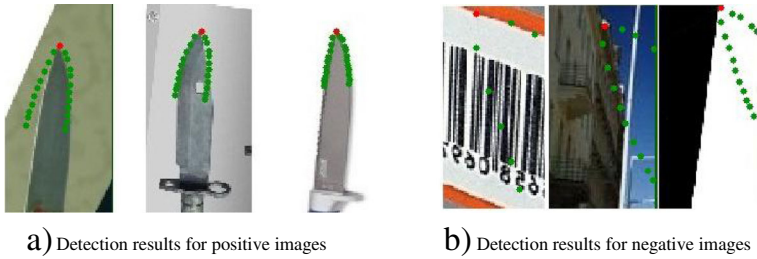
a) Detection results for positive images   b) Detection results for negative images

Fig. 4 Sample detection results

The idea behind this detection scheme is that if the appearance (and so the shape) that the AAM has been trained to locate exists in the image, the model will converge to it if reasonably initialised, even from slightly varying locations (see point 5). In other words, the AAM will converge to the same location, which is evaluated with the minimum-area bounding box and skew-angle. If the appearance that the AAM is searching for does not exist, the model will converge to some random locations. Moreover, to make the positive-detection condition even stronger, we assume that most of the landmark points lie on an edge, to eliminate cases where there are no edges visible (and therefore no knife exists). The exact number of landmarks that actually lie on an edge can be chosen only by a heuristic rule. We have chosen this threshold to be 70 % of the points on each of the two edges of the knife-blade.

Fig. 5 Forty knife images used for evaluation of the detector

Fig. 6 Wrongly classified images
from the positive test set
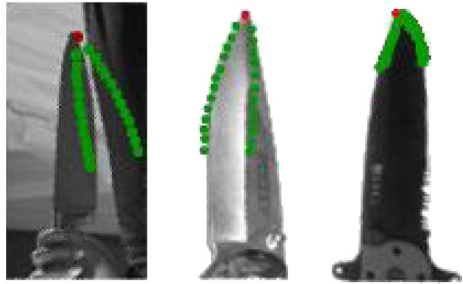
The positive test set consisted of 40 images in normalised 64×128-pixel format that depicted knives vertically centred with the tip facing up. These images are presented in Fig. 5. Three correctly classified images, along with AAM landmarks that have correctly converged to the blade, are presented in Fig. 4a. Of those 40 images, 37 have been correctly labelled as containing a knife, whereas in three cases the result of detection was negative. These are the last three images in Fig. 5. The negative test set consisted of 40 images randomly cropped from images that did not contain knives; samples from this set are shown in Fig. 4b. None of the images has been falsely labelled as containing a knife.

It is worth taking a closer look at the three positive images that have been wrongly classified as not containing knives. The proposed approach failed to detect the knife in the first of these images because the AAM converged to the contour to the right of the actual knife. The knife in the second image has a non-uniform texture with distinct contours. Initialised from slightly varying locations, the model converged to different locations. In the case of the third false negative, due to the uncommon shape of the blade, the AAM converged only to the very top part of the blade. These three cases are presented in Fig. 6. They demonstrate the limitations of the proposed detection scheme, but in our view the detection results are nonetheless satisfactory.

The results of the tests show reasonable classification accuracy for knife/non-knife classification of the proposed detection approach. The AAM failed to correctly converge in the case of a very strong edge in the neighbourhood of the blade. Moreover, blades of an unusual shape or texture also pose a challenge for the detector. The detection accuracy and the false-positive rates are summarised in Table 1.

6 Application in baggage-scanning systems

The proposed knife-detection scheme can be applied to interpreting images produced by baggage-scanning x-ray systems. Modern baggage-scanners produce images of excellent

Table 1 Characteristics of the proposed AAM-based detection scheme

Number of images in positive test set	Number of correctly classified positive images	Classification accuracy	Number of images in negative test set	Number of wrongly classified negative images	False-positive rate
40	37	92.5 %	40	0	0 %

quality and are deployed at numerous locations, including airports, courts of law, government offices and other venues where the danger of bringing in a dangerous tool or explosive exists.

The state-of the art technology utilised in these devices allows for a relatively quick interpretation of the content of the analysed piece of luggage by specially trained staff. Typically, a baggage scanner consists of a screening tunnel, a heavy-duty conveyor belt and an operator interface. The sophisticated x-ray scanners not only produce high-quality x-ray images of luggage contents but also distinguish between substances based on their atomic z-number, as materials are displayed in different colours according to a specific range of atomic z-numbers. This technique is based on dual-energy x-ray absorptiometry (DXA), and has been originally used in medical applications. In addition to determining atomic z-numbers, modern baggage-scanners can distinguish between materials of different densities. For the purposes of our study, we have compiled a set of nine images from baggage x-ray scanning systems, found through a popular image search engine. Three of these are depicted in Fig. 7. Although they come from various scanners where apparently different colours were assigned to similar materials, it is possible to see the clear distinction between objects, even against complicated backgrounds.

However, we were able to find only three x-ray scans containing a knife that were of reasonable quality. Therefore, we have composed an additional set of nine sample images by pasting knives into the non-knife x-ray scans. Due to the fact that metal objects can be displayed in different colours according to the scanner's configuration and make, all images were converted to greyscale. Figure 8 illustrates the searching procedure on three sample images, shown in separate columns. First, corners are found using the Harris corner detector, as in a). Then, at each corner, multiple AAMs are initialised, and trained to model knives at orientations varying by 15°; in total, 24 different AAMs cover a whole rotation. In b), we see AAMs that match the actual orientation of the blade initialised at the tip designated by a corner. In c), we can see the same AAMs that have fully converged to the blade. AAMs initialised not at a knife-tip or with a wrong orientation are demonstrated in d), and e) shows how such AAMs diverge.

The proposed knife-detection scheme has been evaluated on a test set composed of three luggage x-ray scans found through a popular internet search engine, as well as nine artificially composed images with knives manually pasted into different x-ray scans of luggage. The resulting images were of high quality and in all nine cases knives were therefore correctly detected. Figure 9 illustrates the results of running our searching procedure on three genuine images containing knives. Two of these images, designated as a) and b), were correctly classified as containing knives. The result of detection in the case of image
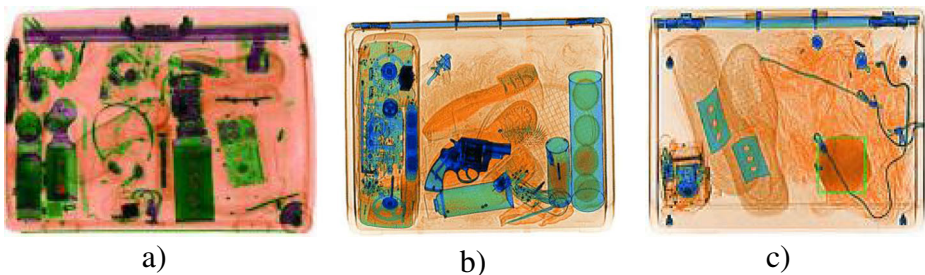

a)   b)   c)

Fig. 7 Modern x-ray baggage-scans (**a**) Source: http://science.howstuffworks.com/transport/flight/modern/airport-security4.htm (**b**) Source: http://www.smithsdetection.com/1025_6460.php (**c**) Source: http://www.smithsdetection.com/x-ray_inspection_baggage_freight.php
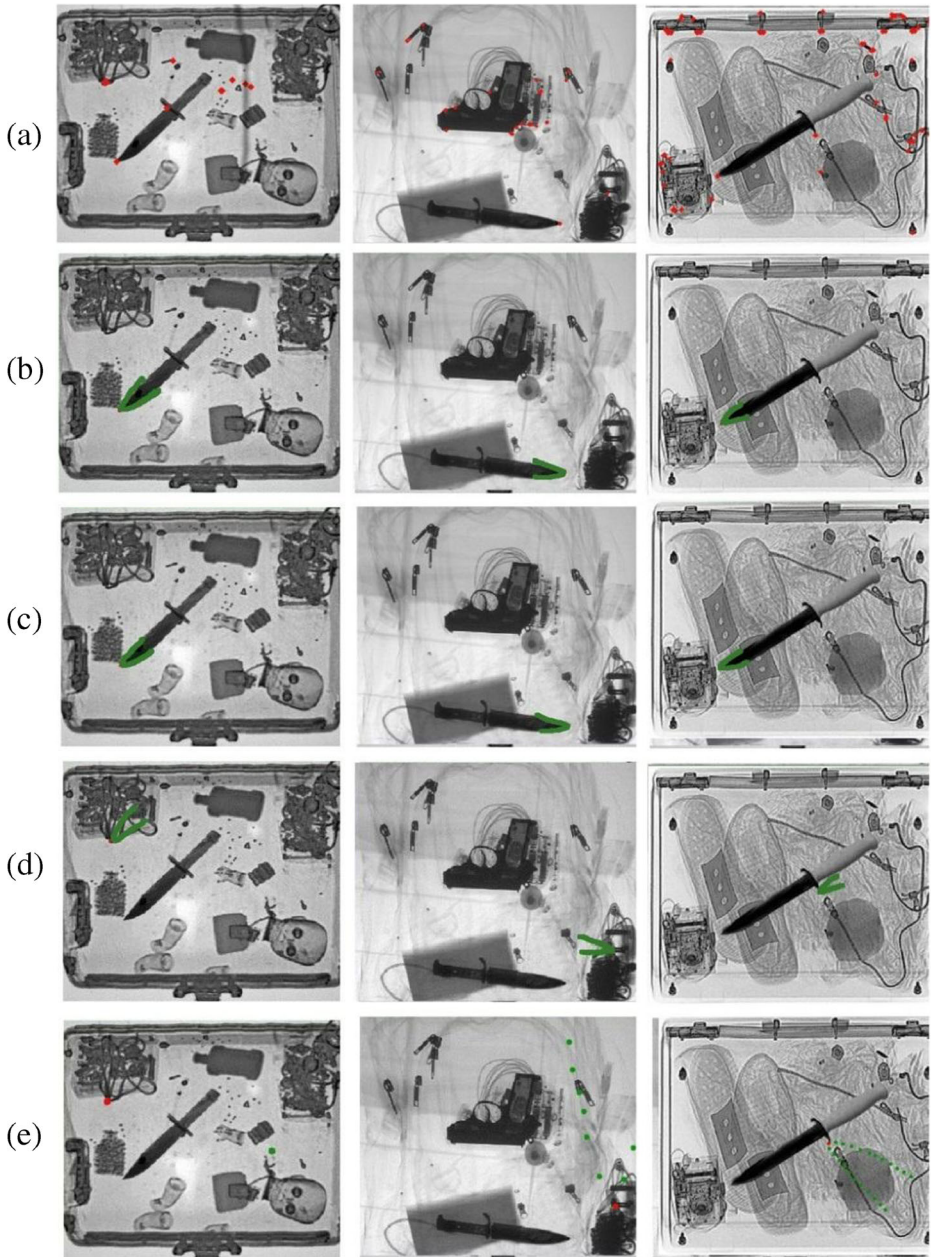
Fig. 8 The results of running AAMs on three sample x-ray images of luggage shown in separate columns: **a**) detected corners, **b**) AAM at initial location designated by a corner, **c**) AAM converged to the blade, **d**) AAM at a corner, and **e**) AAM diverged

c) was incorrect, because the shape of the blade was different than the shape our AAMs were trained on. More precisely, the blade becomes significantly wider at one
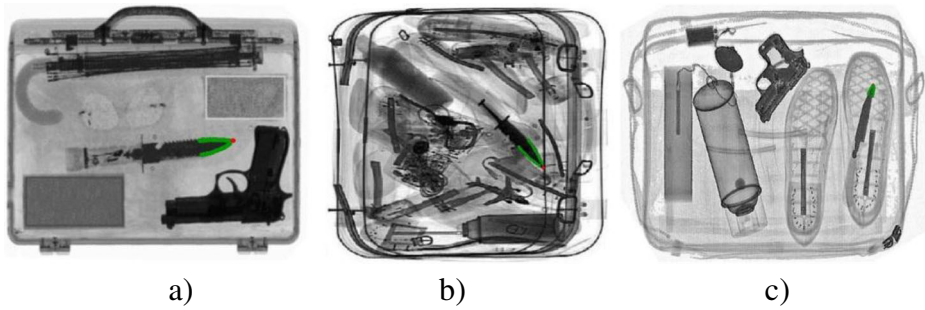
Fig. 9 The results of running AAMs on genuine x-ray scans of luggage (**a**) Source: http://www.deceptology.com/2010/05/could-airport-screeners-find-water.html (**b**) Source: http://aks.rutgers.edu/eyetrack/Real/world.html (**c**) Source: http://www.dpl-surveillance-equipment.com/2500023125.html

point along its length. An AAM initialised exactly at the tip converged to the blade, while AAMs initialised from slightly varying locations converged to other locations and their bounding boxes did not match.

An undoubted advantage of utilising the AAM in knife-detection is its observed zero false-positive detection rate. If there is no knife in an image, AAMs initialised from slightly varying locations will either diverge or converge to different locations. Theoretically it could be the case that an object of a similar shape to the knife-blade will be falsely recognised as a knife. When it comes to application in x-ray baggage-scan analysis, a low non-zero false-positive rate should present little problem in a practical setting, since such images need always to be assessed by a human operator. On the other hand, the condition that the knife-tip is clearly visible in order to properly initialise AAMs may limit the number of possible applications, but most often should be met in baggage x-ray scans. Moreover, due to the technical capabilities of x-ray scanners, objects made from different materials are clearly distinct in images, which in addition to the fact that the size of a knife is relatively large in relation to the size of the suitcase makes our approach of utilising AAMs especially well suited to application in baggage x-ray scan analysis.

7 Summary and conclusions

Applying AAMs to the problem of object detection is a novel approach. So far, they have been used to locate an object in images that were known to contain it, such as medical images containing a body organ–or images of faces, where the task was to locate particular facial elements. In our work, we have used the fact that the knife-blade has a very specific interest point, which can be easily detected as a corner, i.e. its tip. This point can be used to initialise the AAM. The rule that an AAM trained for the orientation close to that of the knife in the image initialised from similar locations should converge to the same location allows us to decide whether the object in question is in fact a knife. The presented approach is well suited to applications where the knife-tip is clearly visible, such as in baggage-scanning systems. If it is not the case, then combining the AAM with a detector of a different kind, due to the AAM's theoretical zero false-positive detection rate, will surely create a robust knife-detector; this is the subject of ongoing research.

References

1. Beichel R, Bischof H, Leberl F, Sonka M (2005) Robust active appearance models and their application to medical image analysis. Medical Imaging, IEEE Transactions on 24(9):1151–1169
2. Bradski G (2000) The OpenCV Library. Dr. Dobb's Journal of Software Tools
3. Bradski, Rost G, Kaehler A (2008) Learning opencv, 1st edition, O'Reilly Media, Inc.
4. Cootes TF, Edwards GJ, Taylor CJ (1998) Active appearance models. In Proceedings: European Conference on Computer Vision 2:484–498
5. Cootes TF, Taylor CJ (2001) Statistical models of appearance for computer vision. Tech. report. University of Manchester
6. Cristinacce D, Cootes TF (2004) A comparison of shape constrained facial feature detectors. Automatic face and gesture recognition. In Proceedings:. Sixth IEEE International Conference, pp 375–380
7. Daijin Kim, Jaewon Sung (2006) A real-time face tracking using the stereo active appearance model. In Proceedings: Image Processing, 2006 IEEE International Conference, pp 2833–2836
8. Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In Proceedings: Computer Vision and Pattern Recognition (CVPR): pp. 886–893
9. Derpanis KG (2004) The Harris corner detector. Accessed 10 April 2013 http://www.cse.yorku.ca/~kosta/CompVis_Notes/harris_detector.pdf
10. Edwards GJ, Taylor CJ, Cootes TF (1998) Interpreting face images using active appearance models. Automatic face and gesture recognition. In: Proceedings: Third IEEE International Conference, pp 300–305
11. Lavee G, Khan L, Thuraisingham B (2007) A framework for a video analysis tool for suspicious event detection. Multimedia Tools and Applications 35(1):109–123
12. Leung KYE, van Stralen M, van Burken G, van der Steen AFW, de Jong N, Bosch JG Automatic (2010) 3D left ventricular border detection using active appearance models. In Proceedings: Ultrasonics Symposium (IUS) pp 197–200
13. Liang Y-M, Shih S-W, Shih AC-C (2011) Human action segmentation and classification based on the Isomap algorithm. Multimedia Tools and Applications: pp 1–20
14. Maalouf A, Larabi M-C, Nicholson D (2012) Offline quality monitoring for legal evidence images in video-surveillance applications. Multimedia Tools and Applications: pp 1–30
15. Mikolajczyk K, Schmid C (2004) Scale & affine invariant interest point detectors. International Journal of Computer Vision 60(1):63–86
16. Mitchell S, Lelieveldt B, Geest R, Schaap J, Reiber J, Sonka M (2000) Segmentation of cardiac MR images: An active appearance model approach. In Proceedings: Image Processing, San Diego CA, SPIE 1:224–234
17. Nam Y, Rho S, Park JH (2012) Intelligent video surveillance system: 3-tier context-aware surveillance system with metadata. Multimedia Tools and Applications 57(2):315–334
18. Paul Viola and Michael Jones (2001) Rapid Object Detection Using a Boosted Cascade of Simple Features. In Proceedings: IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Kauai, USA
19. Schmid C, Mohr R, Bauckhage C (2000) Evaluation of interest point detectors. International Journal of Computer Vision 37(2):151–172
20. Stegmann MB, Gomez DD (2002) A brief introduction to statistical shape analysis. Technical report. In: Informatics and mathematical modelling, Technical University of Denmark, DTU
21. Tadeusiewicz R (2011) How intelligent should Be system for image analysis? In: Kwasnicka H, Jain LC (eds) Innovations in intelligent image analysis. Studies in computational intelligence, vol. 339. Springer Verlag, Berlin
22. Tadeusiewicz R, Introduction to Intelligent Systems (2011) In: Wilamowski BM, Irvin JD (Eds.) The industrial electronics handbook—intelligent systems. CRC Press, Boca Raton. pp 1–1 – 1–12.
23. Tickner AH, Poulton EC (1973) Monitoring up to 16 synthetic television pictures showing a great deal of movement. Ergonomics 16(4):381–401

24. van Ginneken B, Stegmann MB, Loog M (2004). Segmentation of anatomical structures in chest radiographs using supervised methods: A comparative study on a public database. Medical Image Analysis
25. Yifeng J, Zhijun Z, Feng C, Hung Tat T, Tze Kin L (2004) An enhanced appearance model for ultrasound image segmentation. Pattern Recognition 3:802–805

Andrzej Glowacz received his Ph.D. in telecommunications from the AGH University of Science and Technology in 2007. He is currently an Assistant Professor at the AGH University. Andrzej Glowacz has been working on numerous commercial projects, and EU research projects including INSIGMA (as Deputy Project Coordinator), INDECT, GAMA, OASIS Archive, CARMEN, DAIDALOS, DAIDALOS 2, EuroNGI, and EuroFGI. His main professional areas of interest are intelligent information systems, multimedia systems, pattern recognition, wireless QoS, modern transport protocols, and advanced systems programming. He is the author of over eighty scientific papers and technical reports; he also serves as a reviewer of international journals and conferences.

Marcin Kmieć completed his MSc in Control Engineering and Robotics at the AGH University of Science and Technology in 2009. Currently he is a PhD student at the AGH University Department of Electrical Engineering, Automatics, Computer Science and Electronics. His research interests include object recognition, natural scene understanding and using machine learning techniques for video sequence analysis. He has worked for ABB Corporate Research Center in Krakow and taken part in INSIGMA research project.

Professor Andrzej Dziech, Ph.D. Hab received his M.Sc. and Ph.D in telecommunications from the Electro-technical Institute in Leningrad, Faculty of Automation and Computer Science, in 1970 and 1973. He received postdoctoral degree in engineering, Science Tech, from the Poznan University of Technology, Faculty of Electrical Engineering in 1978. He is full professor from 1986. His fields of interest are related to digital communication, image and data processing, data compression, information and coding theory, random signals, computer communications networks and signal processing. He has worked in number of foreign universities, most recently in 2001–2003 he worked as a visiting professor at the University of Wuppertal in Germany. He co-authored 180 publications including 5 books. He promoted 18 Ph.D. students and approx. 100 M.Sc. He was awarded 4 times by Ministry of Education of Poland for his research achievements. He serves as a co-ordinator of FP7 project INDECT.