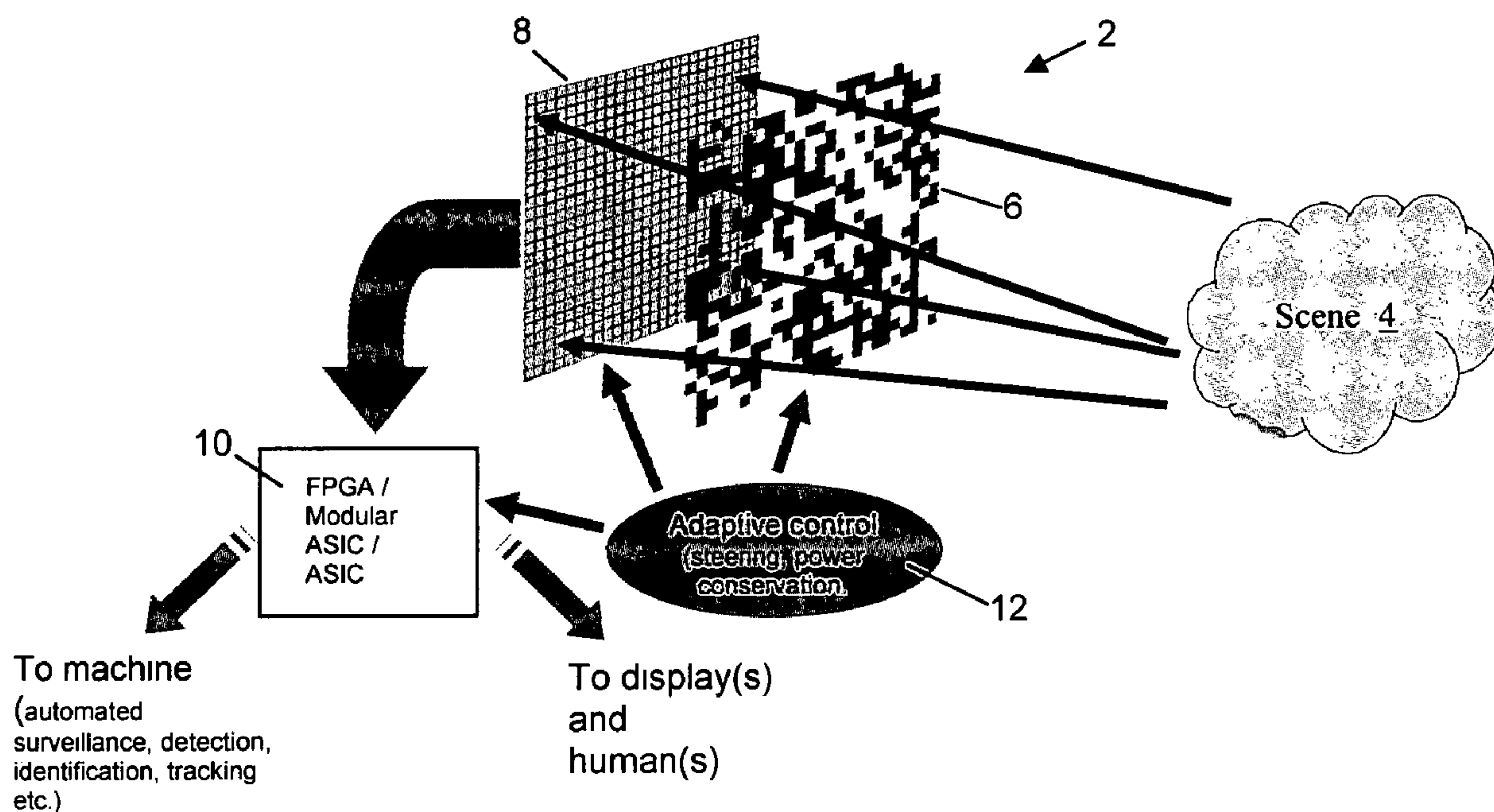




(86) Date de dépôt PCT/PCT Filing Date: 2007/07/27
 (87) Date publication PCT/PCT Publication Date: 2008/01/31
 (85) Entrée phase nationale/National Entry: 2009/01/19
 (86) N° demande PCT/PCT Application No.: GB 2007/002867
 (87) N° publication PCT/PCT Publication No.: 2008/012557
 (30) Priorité/Priority: 2006/07/28 (GB0615040.3)

(51) Cl.Int./Int.Cl. *G01T 1/29* (2006.01)
 (71) Demandeur/Applicant:
QINETIQ LIMITED, GB
 (72) Inventeur/Inventor:
STRENS, MALCOLM JOHN ALEXANDER, GB
 (74) Agent: FETHERSTONHAUGH & CO.

(54) Titre : PROCÉDE DE TRAITEMENT POUR CAPTEUR D'OUVERTURE CODEE
 (54) Title: PROCESSING METHOD FOR CODED APERTURE SENSOR



(57) Abrégé/Abstract:

This invention relates to a method of processing for a coded aperture imaging apparatus which is useful for target identification and tracking. The method uses a statistical scene model and, preferably using several frames of data, determines a likelihood of the position and/or velocity of one or more targets assumed to be in the scene. The method preferably applies a recursive Bayesian filter or Bayesian batch filter to determine a probability distribution of likely state parameters. The method of the present invention acts upon the acquired data directly without requiring any processing to form an image.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
31 January 2008 (31.01.2008)

PCT

(10) International Publication Number
WO 2008/012557 A3

(51) International Patent Classification:
G01T 1/29 (2006.01)

(21) International Application Number:
PCT/GB2007/002867

(22) International Filing Date: 27 July 2007 (27.07.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0615040.3 28 July 2006 (28.07.2006) GB

(71) Applicant (for all designated States except US): **QINETIQ LIMITED** [GB/GB]; 85 Buckingham Gate, London SW1E 6PD (GB).

(72) Inventor; and

(75) Inventor/Applicant (for US only): **STRENS, Malcolm, John, Alexander** [GB/GB]; QinetiQ Limited, Cody Technology Park, Ively Road, Farnborough, Hampshire GU14 0LX (GB).

(74) Agent: **DAVIES, Philip**; QinetiQ Limited, Intellectual Property, Malvern Technology Centre, St Andrews Road, Malvern, Worcestershire WR14 3PS (GB).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM,

AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:

— of inventorship (Rule 4.17(iv))

Published:

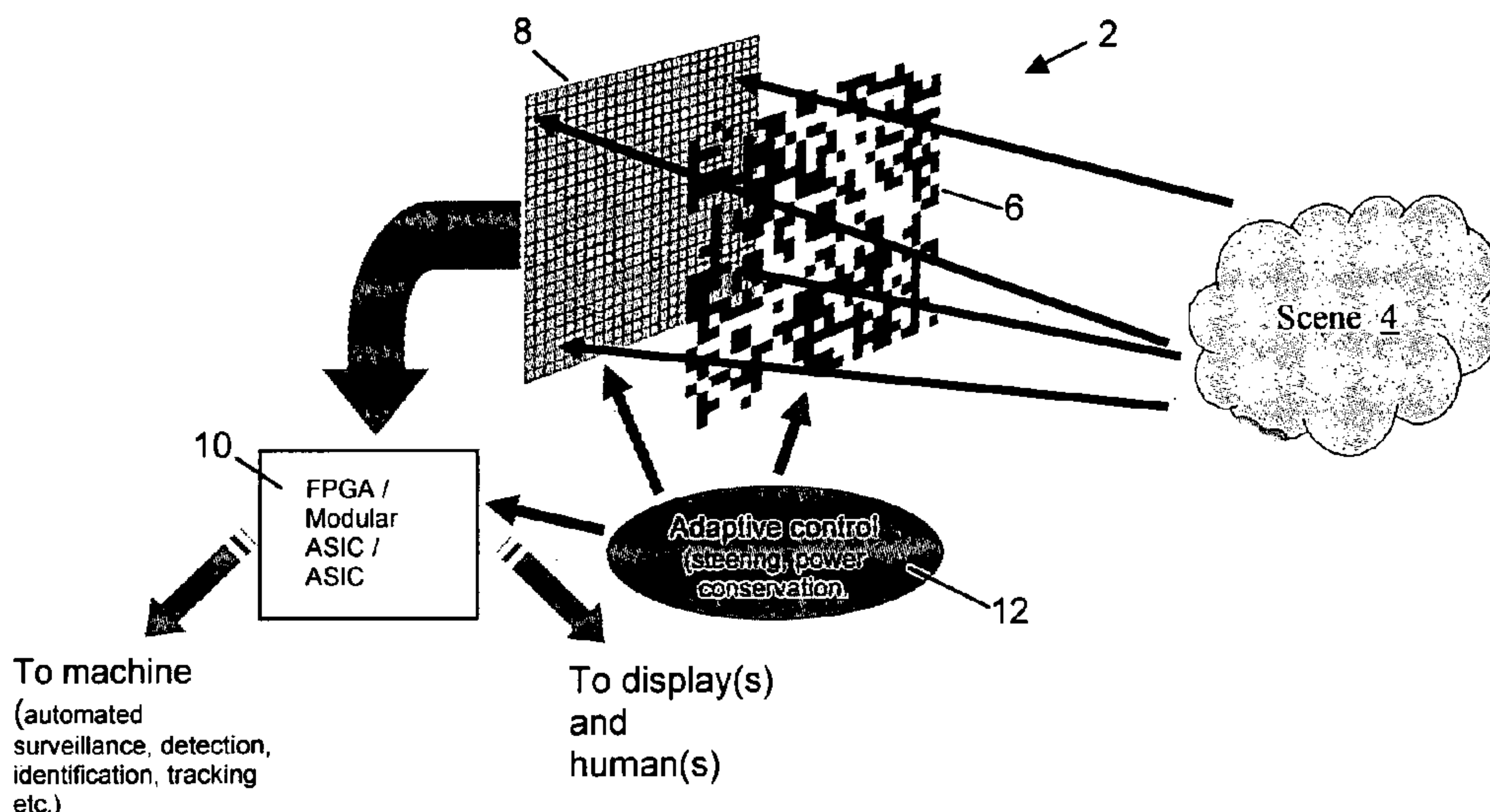
— with international search report

— before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

(88) Date of publication of the international search report:

24 April 2008

(54) Title: PROCESSING METHOD FOR CODED APERTURE SENSOR



(57) Abstract: This invention relates to a method of processing for a coded aperture imaging apparatus which is useful for target identification and tracking. The method uses a statistical scene model and, preferably using several frames of data, determines a likelihood of the position and/or velocity of one or more targets assumed to be in the scene. The method preferably applies a recursive Bayesian filter or Bayesian batch filter to determine a probability distribution of likely state parameters. The method of the present invention acts upon the acquired data directly without requiring any processing to form an image.

Processing Method for Coded Aperture Sensor

This invention relates to the processing of signals obtained using a coded aperture apparatus, especially to a method of detecting and/or tracking one or more targets using
5 data received from a detector array in a coded aperture apparatus.

Coded aperture imaging is a known imaging technique which is primarily used in high energy imaging such as X-ray or γ -ray imaging where suitable lens materials do not generally exist, see for instance E. Fenimore and T.M. Cannon, "Coded aperture imaging
10 with uniformly redundant arrays", Applied Optics, Vol. 17, No. 3, pages 337 – 347, 1 February 1978. It has also been proposed for three dimensional imaging, see for instance "Tomographical imaging using uniformly redundant arrays" Cannon TM, Fenimore EE, Applied Optics **18**, no.7, p. 1052-1057 (1979)

15 Coded aperture imaging exploits the same principles as a pinhole camera but instead of having a single small aperture uses a coded aperture mask having an array of apertures. The small size of the apertures results in a high angular resolution but increasing the number of apertures increases the radiation arriving at the detector thus increasing the signal to noise ratio. Each aperture passes an image of the scene to the detector array
20 and so the pattern at the detector array is the interference pattern from an overlapping series of images and is not recognisable as the scene. Processing is needed to reconstruct the original scene image from the recorded data.

The reconstruction process requires knowledge of the aperture array used and system
25 configuration and the aperture array chosen is often coded to allow subsequent good quality image reconstruction. The processing is performed using a mathematical model of the particular array at a set location.

Coded aperture imaging can be thought of as a geometric imaging technique and for the
30 applications it is usually used for, e.g. astronomy, diffraction is negligible.

Recently, see our co-pending UK patent application GB0510470.8, the present inventors have proposed using a reconfigurable coded aperture imager having a reconfigurable coded aperture mask means. The use of a reconfigurable coded aperture mask means
35 allows different coded aperture masks to be displayed at different times. This allows, for example, the direction and FOV of the imaging system to be altered without requiring any

moving parts. Further the resolution of the imaging system can also be altered by changing the coded aperture mask displayed on the coded aperture mask means.

5 The pattern displayed on the coded aperture mask means is a coded aperture mask and at least part of the coded aperture mask is a coded aperture array. That is either the whole pattern displayed on the mask means is a coded aperture array or only part of the pattern is a coded aperture array. For the avoidance of doubt the term aperture used herein does not imply a physical hole in the mask means but merely an area of the pattern which allows increased radiation to reach the detector.

10

As mentioned the reconfigurable mask means can display a variety of coded aperture masks having a variety of coded aperture arrays in different positions on the mask means. The field of view of the imaging system can be altered by changing the position of the coded aperture mask on the array relative to the detector and the resolution can be altered by changing the size of the coded aperture array. Knowledge of the particular array displayed and its location is used in reconstructing the scene image in the same way as for a fixed coded aperture.

15

GB0510470.8 teaches a versatile and lightweight imaging system that can be rapidly configured to have different fields of view or resolution without any moving parts. It eliminates the need for conventional optics, gives conformal imaging capability, can have an infinite depth of field and gives inherent power free encryption since decoding of the image requires knowledge of the coded aperture array used. The imaging apparatus described therein is particularly suitable for several imaging and surveillance applications in the visible, infrared or ultraviolet wavebands.

20

However, high resolution imaging requires small aperture sizes and a longer optical path from the detector to the mask, which increases the effects of diffraction. Diffraction causes a blurring of the pattern formed by the mask on the detector array, which can reduce the signal to noise level and affect image quality, especially with conventional processing schemes.

30

It is an object of the present invention to provide a method for processing the signal from a coded aperture apparatus to provide accurate target location and/or tracking within a scene. As used in this specification the term target means any desired object which it is wished to detect or track in a scene, for example this could be people for a security

35

surveillance system or vehicles in a traffic monitoring system. In a search and rescue environment the targets may be people or boats/other vehicles.

Thus according to the present invention there is provided a method of locating/detecting
5 a target in a scene comprising the steps of taking data acquired from a detector array arranged to receive radiation from the scene via a coded aperture array and determining a likelihood for the location of one or more targets assumed to be present in the scene using the acquired data and a statistical scene model.

10 The present invention therefore operates using the acquired data directly, i.e. the data received at the detector array in the coded aperture system, and the likelihood for the location of a target is determined using the acquired data which has not been processed to form an image. As mentioned above the signal received at the detector array can be thought of as a series of overlapping images, one from each aperture in the array. The
15 total detector signal therefore has a spatial intensity pattern which is dependent on the scene being imaged and the coded aperture array used but which is not an image of the scene.

In conventional coded aperture imaging the detected intensity pattern is processed to
20 form an image, for instance by applying a suitable deconvolution. Once an image has been formed standard image processing techniques can be applied to the processed image data to detect image features etc.

The present invention however uses the signal as recorded by the detector array and
25 determines the likelihood of the locations of one or more targets assumed to be present without processing the data to form an image. The method of the present invention uses a statistical model of the scene to infer parameters (often referred to as a state) that describe the scene, including the position and/or velocity of any targets in the scene. In simple terms the method models possible target positions and/or velocities and
30 determines the likelihood that the signal received at the detector array was due to a target being located at a particular part of the scene.

The present invention exploits the fact that processing the information recorded at the detector array to form an image will generally result in a loss of information. A loss of
35 information occurs because the resolution (angle of the scene subtended by a pixel) of an image that is formed by conventional processing is limited by the amount of memory

and computational resources that are available. Also, conventional processing requires the choice of “regularisation” parameters that balance noise cancellation/smoothing against sharpness; making a choice about the value of these parameters implies a loss of information in the resultant image when compared with the information available in the measured signal. The accuracy of target localisation in the formed image is unlikely to be much better than the pixel spacing. Working directly with the signal recorded by the detector array avoids this loss of information in target detection and tracking and potentially in many other sensing functions.

10 The method of the present invention takes the data directly acquired by the detector array and interprets it under a statistical scene model. The scene model is a generative statistical model of the signal at the detector array due to the objects (and surfaces etc) that are assumed (or are known) to form the scene, i.e. it is a model of the signal at the detector array from modelled entities within the scene. A scene model consists of two parts. The first part is the set of possible states. The state is a set of parameters describing objects in the scene, their motion, or other attributes or properties of the scene as a whole such as parameters of a geometric transformation describing its motion. The second part of the scene model is a function that, given any detector signal and any state, provides a probability that the particular signal would be obtained conditional upon the particular state.

The model may include elements modelling the propagation of radiation from the entities in the scene to the coded aperture array and propagation through the coded aperture array to the detector array. It may also include a model of the detector. The model therefore accounts for atmospheric attenuation, diffraction effects, spectral effects, detector performance etc.

For instance imagine the data is acquired from a coded aperture system having a thermal infrared detector array with a fixed field of view arranged to view the sky for aircraft (targets). The sky is generally a uniform cold background whereas the targets are heat sources, i.e. are high intensity sources to the infrared detector array. The state will describe the possible positions of targets in the scene and optionally also possible target attributes such as size, shape, orientations etc., which may have prior probabilities based on assumptions about known targets and target behaviour. The first part of the statistical scene model will therefore be the set of possible states. If a target of known size, shape, orientation and intensity is assumed to be at a particular part of the scene it

is possible to estimate the signal that would be expected from that target, given knowledge of the particular coded aperture array used when the data was acquired. Therefore the second part of the scene model calculates the probability that the particular signal was obtained from each possible state. The statistical model must allow
5 for various forms of uncertainty including thermal noise at the detector and any unmodelled signal effects which may include clutter in the scene, diffraction and reflection. (These effects instead may be modelled, in which case they do not need to be 'explained away' as noise.)

10 The invention works by evaluating the probability of the signal under the scene model at multiple states. The application of Bayes rule allows probabilities of *signal given state* to be interpreted as probabilities of *state given signal* (up to some unknown normalising constant). Therefore a probability distribution can be obtained over the chosen set of states. The state with maximal probability is the maximum likelihood (most likely) one,
15 but the distribution as a whole is useful as an input to higher level processes or for the information update stage of a tracking process.

The invention can apply Bayesian inference to a single frame of data and the set of states will include one or more possible targets or possible parts of the scene. The most
20 likely state determined by the method will therefore indicate the most likely position or positions of one or more targets assumed to be in the scene. The method of the present invention can however combine several frames of data to improve accuracy, especially with a target which is moving relative to the scene. Thus the method can apply recursive Bayesian filtering or Bayesian batch processing to a sequence of frames. The skilled
25 person will be aware that recursive Bayesian filters have been used for tracking applications previously. For instance recursive Bayesian filters have been applied to the data from conventional imaging systems (and so applied to image data) to track targets.

The basic idea of using a recursive Bayesian filter, a sequential form of Bayesian
30 inference, is to model possible evolving state (e.g. target attributes including velocity) and to update the uncertainty distribution over states as more information is acquired, i.e. more frames of data are captured, so as to reduce the uncertainty in state (e.g. increase the accuracy of target location). In the case of Bayesian filtering, an additional (third part) of the scene model is required. This is a dynamics model which accounts for the
35 evolution of state over time: for example the way a moving target's position changes as a function of its velocity. This dynamical model can be stochastic. i.e. the transition from

6

one state to the next is a probabilistic process. The method described therefore determines the most likely state parameters, e.g. the position and motion of a target assumed to be in the scene. To determine whether a target is actually present or not would require an additional step. For instance a measure of signal strength could be used to determine whether there is indeed a target present; effectively the confidence in the most likely state is thresholded. Other rules could be applied however such as an indication of motion relative to the scene background being a good indicator that a target is present. Multiple target tracking is achieved by extending the state to contain the parameters for multiple targets and to ensure there is some competition or data association mechanism to encourage each parameter group to converge onto a different target. These data association and competition mechanisms are well explored in conventional target tracking literature.

The present invention therefore offers a direct method of processing data obtained using a coded aperture system without requiring an image to be formed. As the method works on the direct signal the accuracy of tracking and target localisation can improved as compared to methods operating on the image processed data.

It should be noted that one could process the data from the coded aperture sensor to form images and then apply conventional Bayesian detection/processing to the images. Prior to the present invention indeed this is how one would have applied Bayesian processing to a coded aperture sensor. As mentioned above however processing the images will result in a loss of information which the present invention, by working with the direct data, avoids. Furthermore applying a "full" Bayesian image based approach would be infeasible because it would require an initial processing stage to generate a large sample of super-resolved images rather than a single estimated image, which would have significant impact on memory (space) and processing (time) requirements. The present invention avoids the image processing overheads and can provide more accurate localisation.

The method can estimate the position of one or more targets from a single frame of data, and track the position and motion of one or more targets from multiple frames using the direct signal. If the presence/absence of a target is not determined until information from multiple frames has been taken into account, then a track before detect process will be obtained. This allows detection of targets that are difficult to detect because they have low contrast or are located in cluttered areas of the scene.

The present invention does not constrain the contents of the scene state which may include other information alongside or instead of the parameters of targets.

- 5 For instance part of the state in the scene model may be a representation of the scene background (rather than only the objects (targets) of particular interest). In other words the scene model may include information about entities which appear in the scene which are not targets – the general background. The obtained background estimate can be exploited for several purposes such as eliminating clutter from the target detection
10 process or forming a composite image of the scene (a mosaic).

For instance imagine that the data is acquired from a detector array fixed relative to a particular scene and the scene has a high degree of clutter, i.e. they may be buildings, vegetation etc. in the scene with a varying degree of contrast. If the method is looking
15 for targets which are highly contrasting to the scene background the contribution from the background could be represented as general noise in the scene model and the targets identified as described – especially if they are moving relative to the scene background. However if the targets are of low/negative contrast and/or there is target like clutter in the scene background, it may be desirable to include known elements of the scene
20 background in the state of the scene model. In effect the known contribution of the modelled background elements will be taken into account in determining the probability that a particular signal which has been received was due to a particular state. In this way the effect of clutter in the scene can be reduced.

25 Where background is included in the model there may in effect be two parts to the model, a background scene model which is used to determine a background probability for each part of the scene, i.e. a probability that a particular part of the scene corresponds to the background and a target probability model which determines the likelihood of a target. This approach improves the chance of detecting a target because
30 the background model will 'explain away' clutter, preventing it from interfering with the target model's inference process. Effectively, the relative probability between target and background is computed and used to maximise the ability to discriminate between the two. With this method, it is feasible to detect low and negative contrast targets that have relative motion across the scene background.

The part of the scene model which represents scene background may be constructed in various ways. For instance one or preferably more frames of data could be processed to form an actual image of the scene. Key objects, i.e. areas of contrast, in the scene could then be identified and added to the state of the scene model. Note that although image
5 processing may therefore be performed on some data acquired from the detector array this is purely for the purposes of improving the scene model by incorporating background elements – there is no target detection. The method of detecting and/or locating a target of the present invention would then take the unprocessed acquired data and use the scene model to determine the most likely state without doing any image processing.

10

If the detector array is fixed relative to the scene whilst the data is acquired there will be no apparent motion of the scene. Thus movement of targets relative to the background can be easily detected and, if the scene model includes any background data it can be assumed that the modelled background elements are not moving from frame to frame.
15 However if the detector is moving relative to the scene then the movement of a target will be a function of both the target's actual motion and the relative movement of the detector and the scene and so is a more complicated function. Also, if one includes background elements in the scene model then these will have apparent motion. It may be desirable therefore to determine the relative movement of the detector array to the scene. This
20 could be done by forming an image of the scene from each frame of data and using standard image processing techniques to identify motion of the background, e.g. linear optic flow operators could be applied to the scene images. However the present invention also allows background scene motion to be inferred directly from the data received at the detector array. If part of the state in the scene model describes the
25 motion of the scene background (but not its contents) in the form of parameters for a geometric transformation from one frame to the next, then these parameters may be directly inferred without forming an image. The benefit may be the ability to track the background at a much higher accuracy than would be obtained if an image were formed, for the same reason as a higher target position accuracy is obtained. The estimation of
30 background motion (also known as image registration) is particularly important if the sensor is itself moving, because it allows the effects of this motion to be removed/reduced in subsequent processing stages.

The use of a background scene model and the use of optic flow are compatible within the
35 Bayesian framework: it is possible to estimate scene motion using optic flow. Knowledge

of the scene motion supports estimate/updating of the background scene model from a moving sensor system.

- As mentioned the state that is inferred can represent various attributes of a target assumed to be in the scene, other entities in the scene or motion of the scene or entities in the scene. As mentioned the state could infer one or more of the position of a target in the scene, the motion of a target and other attributes of a target that are unknown such as size, shape, spectral signature etc. and the state may relate to these attributes for one or more targets in the scene. The state can also describe aspects of the scene such as one or more of the motion of the scene itself, the scene background or the parameters of other entities in the scene or aspects of the scene such as straight edges, corners, surfaces, illumination state, camera parameters etc. These may be included in the state describing a target for target detection/tracking as described above. However the method may be used purely to determine some aspect of the scene without any target tracking or detection. This could be performed to infer some parameters prior to performing a target detection/tracking step or it may be a stand alone process to determine some information about the scene without needing to form an image. Thus generally the method of the present invention relates to directly inferring scene attributes using Bayesian inference directly on the data received by the detector array. Hence in another aspect of the invention there is provided a method of inferring scene attributes comprising the step of taking data acquired from a detector array arranged to receive radiation from the scene via a coded aperture array and performing Bayesian inference on said data directly to infer said attributes. As with the first aspect of the invention the Bayesian inference is applied to data which has not been processed to form an image. The attributes to be inferred may include one or more of parameters describing one or more targets in the scene, motion of the scene, the scene background, parameters of other entities in the scene, aspects of the scene (such as straight edges, corners, surfaces, illumination state, camera parameters).
- The Bayesian inference process is performed as described above with respect to the first aspect of the invention, i.e. a statistical model of the scene including uncertainty in said attributes is formed and the probability of state given signal is inferred from the probability of signal given state.
- As mentioned the method described above according to the present invention offers a useful detection and tracking method for targets in a scene. However the direct

Bayesian method of the present invention is also applicable to accurate tracking even if targets have been detected via another method. For instance an external image based track before detect process can be used to give target detection with relatively coarse localisation and motion information. From any such confirmed detections a dedicated
5 direct tracking filter could be initiated.

Therefore in another aspect of the present invention there is provided a method of tracking an identified target in a scene comprising the steps of taking several frames of data acquired from a detector array arranged to receive radiation from the scene via a
10 coded aperture array and applying a recursive Bayesian filter to the acquired data.

As with the first aspect of the invention the recursive Bayesian filter is applied to the acquired data which has not been processed to form an image of the target or entity of interest. All of the embodiments and advantages described above with respect to the
15 first aspect of the invention are also applicable to this aspect of the invention.

The invention will now be described by way of example only with respect to the following drawings of which;

Figure 1 shows a coded aperture imaging apparatus,

5

Figure 2 shows a coded aperture system and illustrates different mask patterns having different fields of view, and

Figure 3 shows a simulation of a scene, a coded aperture array, the resulting intensity pattern at the detector array and a reconstructed image of the scene,

10

Figure 4 shows the determined positional error over time, in detector pixels, as the method of the present invention was applied to simulated data, and

Figure 5 shows the determined velocity error over time, in pixels per step, as the method of the present invention was applied to simulated data.

15

Coded aperture imaging (CAI) is based on the same principles as a pinhole camera. In a pinhole camera, images free from chromatic aberration are formed at all distances away from the pinhole, allowing the prospect of more compact imaging systems, with a much larger depth of field. However, the major penalty is the poor intensity throughput, which results from the small light gathering characteristics of the pinhole. Nevertheless, the camera is still able to produce images with a resolution determined by the diameter of the pinhole, although diffraction effects have to be considered. The light throughput of the system can be increased by several orders of magnitude, while preserving angular resolution, by using an array of pinholes. Each detector element sees the result of the summation of contributions from the various pinholes, corresponding to each viewpoint of the scene.

25

30

Another way of understanding the operating principle of conventional CAI is to observe that this is a purely geometric imaging technique. Light from every point in a scene within the field of regard (FOR) of the system casts a shadow of the coded aperture onto the detector array. The detector measures the intensity sum of these shadows. The coded aperture is specially designed such that its autocorrelation function is sharp with very low sidelobes. Typically pseudorandom or uniformly redundant arrays (URA) (such as

35

described in E. Fenimore and T.M. Cannon, "Coded aperture imaging with uniformly redundant arrays", Applied Optics, Vol. 17, No. 3, pages 337 – 347, 1 February 1978) are used where a deconvolution or decorrelation of the detector intensity pattern can yield a good approximation to the point distribution in the scene.

5

Figure 1 shows schematically an example of a coded aperture imaging system, generally indicated 2. Rays of light from points in the scene 4 fall onto a particular coded aperture array 6. The coded aperture array acts as a shadow mask and therefore a series of overlapping coded images are produced on the detector array 8. At each pixel on the
10 detector array, the intensities from the overlapping, coded images are summed. The output from the detector array 8 is passed to a processor 10 where image of the scene can be subsequently decoded from the detector signals using a variety of digital signal processing techniques.

15 Recently GB0510470.8 has proposed using a reconfigurable mask means 6 to provide reconfigurable coded aperture array. The coded aperture mask means 6 is controlled by a controller 12 which controls the reconfigurable mask means to display different coded aperture masks. If only part of the coded aperture mask means displays a coded
20 aperture array, the rest of the mask preventing radiation from reaching the detector, then the field of view of the apparatus is determined by the location and size of the coded aperture array relative to the detector, changing its position or size changes the field of view and/or resolution of the imager. This is illustrated in Figure 2. In figure 2a the whole of the mask is used as a coded aperture array and the apparatus in this case is receiving radiation from the whole field of regard of the system. In figure 2b however the
25 majority of the mask is opaque, blocking radiation from reaching the detector array 8. The coded aperture array only forms a small part of the mask and hence only radiation from a restricted field of reaches the detector array. In this mode the apparatus is operating with a narrower field of view but higher resolution.

30 CAI therefore offers the ability to provide a compact and lightweight imaging apparatus which has a large depth of field and has a changeable field of view without requiring any moving parts.

Note that conventional CAI devices do not generally include any elements with optical
35 power, i.e. lenses, in the optical path. In some applications a CAI imager is being used because a suitable lens material is not available and hence the use of a lens is not

possible. The present invention is applicable to a CAI apparatus which does include an element with optical power, such as a lens, in the optical path. The purpose of the lens is not to focus the incident radiation to form a focussed image of the scene – otherwise no CAI processing would be necessary – but simply to provide some degree of optical concentration so that the diffraction pattern is less spread out. As used in this
5 specification therefore the term coded aperture imager should be taken to include an apparatus which has a detector array receiving radiation from a scene via a coded aperture array and also having an optical element such as a lens in the optical path but where the optical element does not form a substantially focussed image of the scene on
10 the detector array. In other words a device which has a coded aperture array in the optical path to provide coding of the signal at the detector array which, using a pattern based on the coded aperture pattern, can be processed to form an unobstructed image of the scene.

15 As mentioned above then, in any coded aperture imager, the intensity pattern formed at the detector array 8 is a series of overlapping images of the scene created by the various apertures. This intensity pattern requires decoding to construct a recognisable image of the scene. Figure 3 shows a scene 50 and a representative coded aperture array 52. As can be seen by the simulated intensity pattern at the detector array 54, the signal
20 received by the detector is an encoded version of the scene and is not directly recognisable. It has previously been thought that in order to extract useful information from the signal it is necessary to decode the signal recorded at the detector array, usually by means of a deconvolution, and to reconstruct an image of the scene. Image 56 is a reconstructed scene image.

25 The present invention lies at least partly however in the realisation that for the purposes of automated target detection and/or tracking it is possible to work directly with the signal received at the detector array. Further as no information is lost, unlike in a deconvolution process, target location and tracking using the direct signal can offer improved accuracy
30 over that based on the processed image data.

The invention uses Bayesian inference to determine the location of targets in the scene. The skilled person will be familiar with Bayesian inference but a simplistic description will be given to aid understanding.

Consider that the signal on the detector array, denoted by the random variable \mathbf{S} , has a particular form, \mathbf{s} . One then is seeking to determine what is the probability distribution across the scene that the signal \mathbf{s} arises from a target at that point in the scene, i.e.

$$5 \quad P(\mathbf{X} = \mathbf{x} \mid \mathbf{S} = \mathbf{s}) \quad \text{Eqn (1)}$$

where \mathbf{X} is a random variable for the state vector representing the target position – and possibly other attributes of the target such as a velocity, size, brightness etc – and \mathbf{x} denotes a particular realisation of \mathbf{X} . If multiple targets (or other objects of interest including the scene background) are assumed/known to be present, \mathbf{x} must also encode these.

Following Bayes theorem

$$15 \quad P(\mathbf{X} = \mathbf{x} \mid \mathbf{S} = \mathbf{s}) = \frac{P(\mathbf{S} = \mathbf{s} \mid \mathbf{X} = \mathbf{x}) P(\mathbf{X} = \mathbf{x})}{P(\mathbf{S} = \mathbf{s})} \quad \text{Eqn (2)}$$

The marginal probability $P(\mathbf{S} = \mathbf{s})$ is an unknown quantity but is constant (assuming it is non zero) and thus does not affect the relative probability distribution. The prior probability density over states is $P(\mathbf{X} = \mathbf{x})$. In a single frame detection process, this encodes any *prior knowledge* about the likely position (and motion) with which a target could appear. As much prior information as possible about desired targets can be incorporated to improve accuracy.

Thus based on a generative model of the signal resulting from possible target states, $P(\mathbf{S} = \mathbf{s} \mid \mathbf{X} = \mathbf{x})$, a likelihood model can be generated, $P(\mathbf{X} = \mathbf{x} \mid \mathbf{S} = \mathbf{s})$.

The generative model can have multiple stages consisting of:

- model of the reflection/emission of photons by the object
- model of atmospheric attenuation
- 30 - model of propagation through the coded aperture (mask), including diffraction effects
- model of the conversion by the detector of the photons into a signal.

If the state is target position, this likelihood model gives a probability distribution for the target's position in the scene, given the observed signal. This probability distribution can be used itself as an indicator of possible targets and their location in the scene.

A greater accuracy can be obtained by combining several frames of data. The invention can therefore employ a recursive Bayesian filter with several frames of data. Here an estimate of how the scene will change state over time (from \mathbf{x} at time t to \mathbf{x}' at time $t+1$) is given as a dynamics model $P(\mathbf{X}_{t+1} = \mathbf{x}' | \mathbf{X}_t = \mathbf{x})$, which could for example model the motion of a target. The skilled person will be aware of Bayesian tracking filters applied in other tracking applications and how such a filter could be implemented here. The state of the filter is a low-dimensional parameterisation of the target's position, motion and possibly other attributes. The filter works as for a conventional Bayesian tracker (such as a particle filter or Kalman filter) but with a likelihood model that directly interprets the signal at the detector array rather than the reconstructed image. The information update stage of the tracker is therefore identical to equation (1) except that $P(\mathbf{X} = \mathbf{x})$ is replaced by a predictive distribution calculated by application of the dynamics model to the posterior distribution from the preceding time steps. This novel approach may provide much better localisation accuracy than could be obtained by working with a reconstructed image, because its accuracy is not limited by any particular choice of image resolution (made in the design of the reconstruction algorithm), which is limited by the memory and processing resources that are available in operation of the sensor. Generally the tracker will be a *multi-target* tracker that directly fits a *mixture* of these likelihood models to the observed signal.

20

For targets that do not have a very high contrast to the local scene background it is desirable to address the contribution of the scene background to the signal as much as possible. In such instances therefore the scene model used within the likelihood computation preferably includes as much information as possible about the scene background.

25

Additionally or alternatively where there is relative motion between the detector array and the scene it may be advantageous to identify the relative motion of the scene between frames. In conventional image processing this is known as image registration. The idea behind image registration is to find an estimate of the geometric transformation (which may include translation, rotation, expansion, or other effects) of the background between two successive frames. In conventional processing, this transformation (and its uncertainty) can be estimated by application of linear 'optic flow' operators to the sensor image. The *direct* approach proposed here works directly with the signal received at the detector array rather than an image reconstructed from it. Specifically, given a parameterised family of geometric transformations of the scene and an assumption or

30

35

estimate for the depth variation in the scene, it is possible to relate the rate of change of each parameter in the transformation to the rate of change of the signal at each detector element. This enables the most likely geometric transformation between two successive observations of the scene to be estimated without ever forming an image. Furthermore,
5 higher-order derivatives can be exploited to provide an estimate for the uncertainty in the geometric transformation.

The method of the present invention has been tested in simulations to prove its validity. One embodiment of the method of the present invention was implemented using Matlab
10 TM.

To verify the method an exemplar diffraction pattern was computed using a one-dimensional model of a system containing a random mask pattern and an optical element. The internal structure in the diffraction pattern means that although it is spread
15 across multiple detector pixels, a small spatial shift will cause large fluctuations in the received signal. This property is referred to as shift sensitivity and it is desirable for accurate tracking. A two dimensional diffraction pattern was then synthesised from the one-dimensional pattern by assuming linear separability between the 2 axes. (Note that in general this separability property does not hold, but it is expected that the spatial
20 structure in our synthesised 2D pattern is similar to that which would be observed from a full 2D optical model). Test sequences of detector signal frames were synthesised using shifted and sampled copies of this diffraction pattern, according to the target's motion, and with added independent and identically distributed noise in each detector pixel. The noise level was chosen so that the signal level of the correlation filter response was
25 equal to the noise level (i.e. a signal to noise ratio of 0 dB.) These sequences were provided as input to the tracker, and the achieved tracking accuracy was estimated from multiple runs using robust statistics.

The results shown in Figure 4 and Figure 5 show that the tracker, over a sequence of
30 observations, builds up confidence in the position and velocity of the target. Figure 4 shows the root-mean-square positional error of determined location of the target - measured in detector pixels - as a function of the successive steps (frames processed). Figure 5 shows the root-mean-square error in target velocity as a function of successive steps - measured in detector pixels per step. It can be seen that the measured error
35 drops over time until the determined positional error is less than one pixel and the determined velocity error is less than 0.01 pixel per step. The target has some unknown

acceleration and this causes the tracker to reach an equilibrium level of accuracy rather than converge to arbitrary accuracy. These results illustrate that direct target tracking without forming an image is achievable with the CAI. (Note that the inclusion of an optical element in the modelled system affects the spread of the diffraction pattern and hence
5 the signal to noise ratio, but its omission would not change the structure of the tracking algorithm.)

Claims

1. A method of locating/detecting a target in a scene comprising the steps of taking data acquired from a detector array arranged to receive radiation from the scene via a coded aperture array, and determining a likelihood for the position of a target assumed to be located in the scene using the acquired data and a statistical scene model.
2. A method as claimed in claim 1 wherein the likelihood of a target being located in the scene is determined using the acquired data which has not been processed to form an image.
3. A method as claimed in claim 1 or claim 2 wherein the scene model is a generative statistical model of the signal at the detector array due to one or more modelled entities within the scene.
4. A method as claimed in any preceding claim wherein the scene model comprises a set of possible states and a function that determines a probability that the particular signal would be obtained conditional upon the particular state.
5. A method as claimed in claim 4 wherein the step of determining a likelihood for the position of a target in the scene comprises the step of determining a probability distribution over the set of states.
6. A method as claimed in claim 5 wherein the method also determines an estimate of motion of a target in the scene.
7. A method as claimed in any preceding claim wherein the step of determining a likelihood of a target in the scene combines data from several frames.
8. A method as claimed in claim 6 wherein a recursive Bayesian filter or Bayesian batch filter is applied to several frames of data.
9. A method as claimed in claim 8 wherein the scene model comprises a dynamics model which accounts for the evolution of a state over time.

10. A method as claimed in any of claims 7 – 9 wherein the method comprises a track-before-detect process.
11. A method as claimed in any preceding claim wherein part of the scene model is an estimation of scene background.
12. A method as claimed in any preceding claim further comprising the step of estimating scene motion effects.
13. A method as claimed in any preceding claim wherein the step of estimating scene motion effects comprises estimating any motion effects directly from the acquired data.
14. A method as claimed in claim 13 wherein the step of estimating scene motion effects comprises the step of comparing rate of change of the signal at each detector element in successive frames with the rate of change of each parameter in a family of parameterised geometric transformations representing possible motions.
15. A method of inferring scene attributes comprising the step of taking data acquired from a detector array arranged to receive radiation from the scene via a coded aperture array and performing Bayesian inference on said data directly to infer said attributes.
16. A method as claimed in claim 15 wherein the Bayesian inference is applied to data which has not been processed to form an image.
17. A method as claimed in claim 15 or claim 16 wherein the attributes to be inferred comprise one or more of parameters describing one or more targets in the scene, motion of the scene, the scene background, parameters of other entities in the scene or aspects of the scene.
18. A method of tracking a target in a scene that has been detected by an external process comprising the steps of taking several frames of data acquired from a detector array arranged to receive radiation from the scene via a coded aperture array and applying a recursive Bayesian filter to the acquired data.

19. A method of tracking as claimed in claim 18 wherein the Bayesian recursive filter or Bayesian batch filter is applied to the acquired data which has not been processed to form an image.

Figure 1

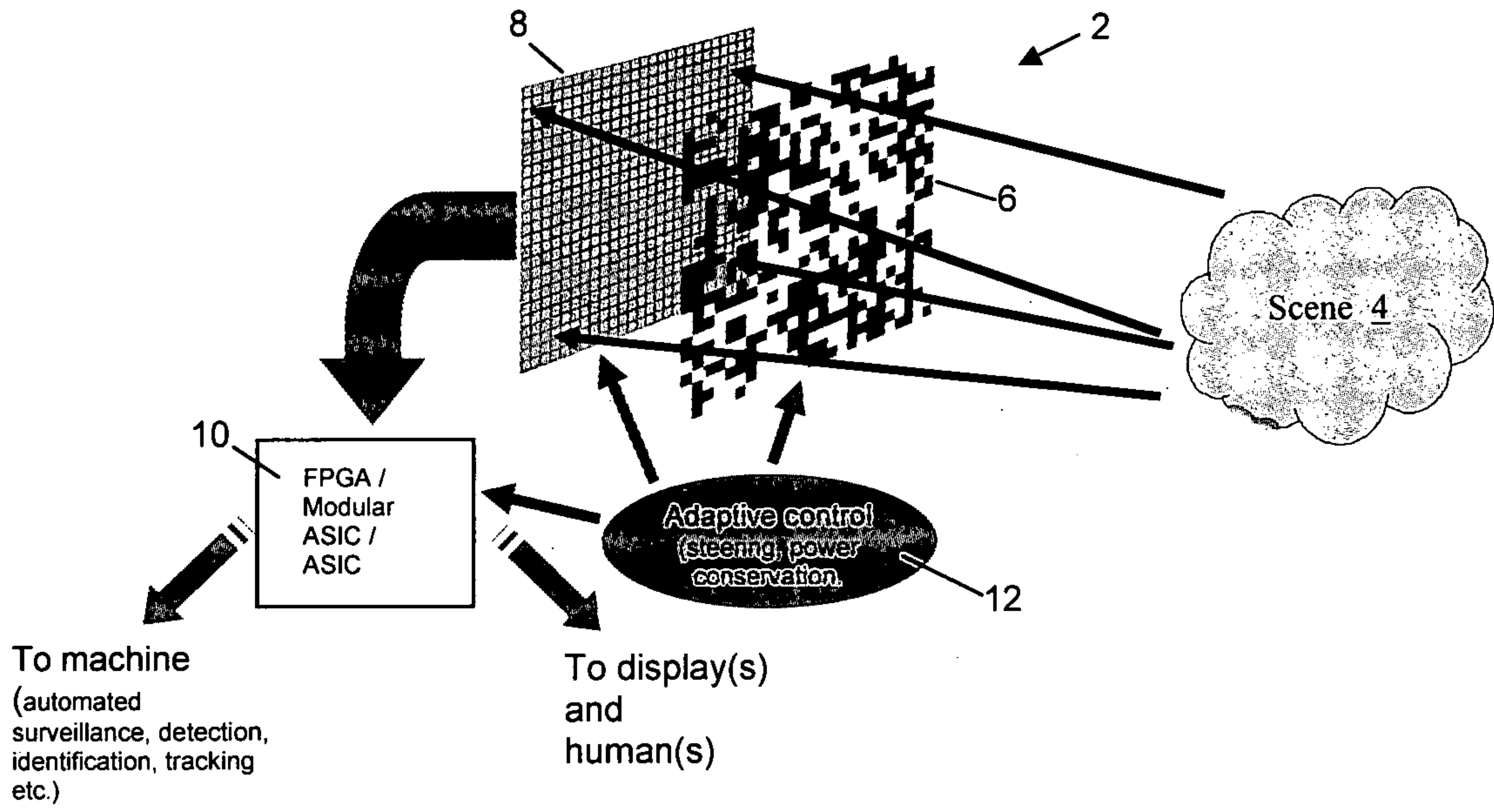


Figure 2

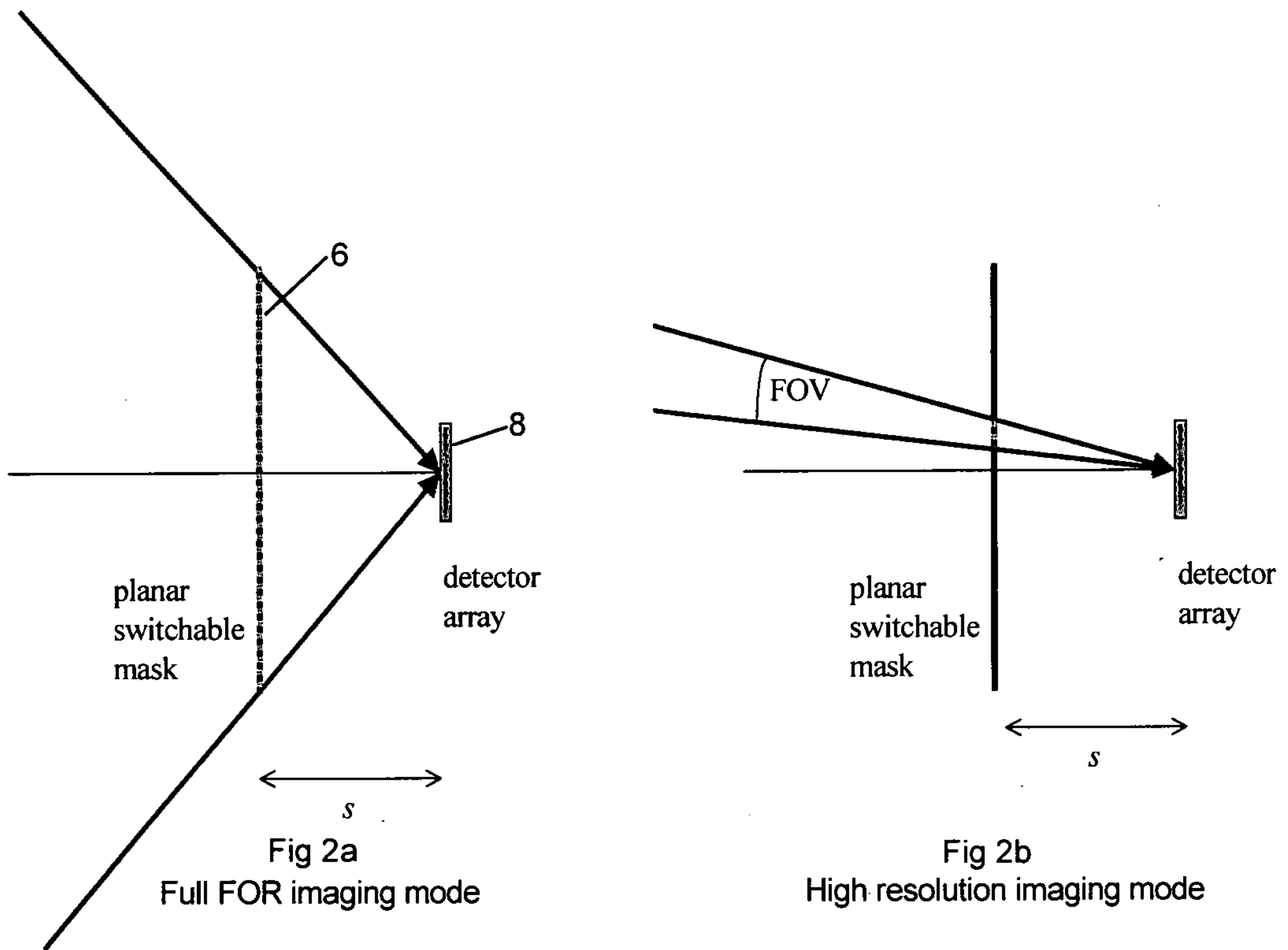


Figure 3

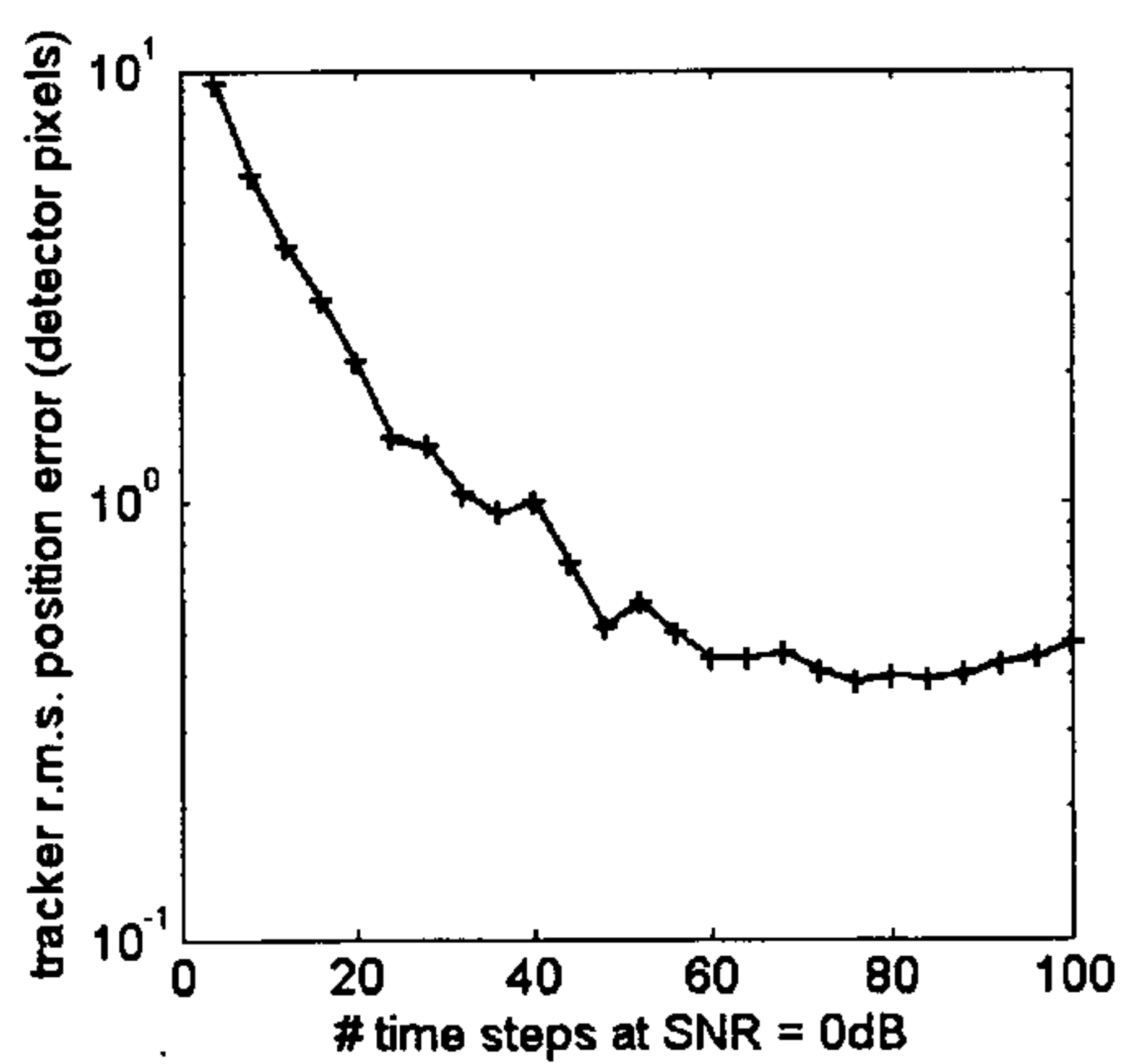
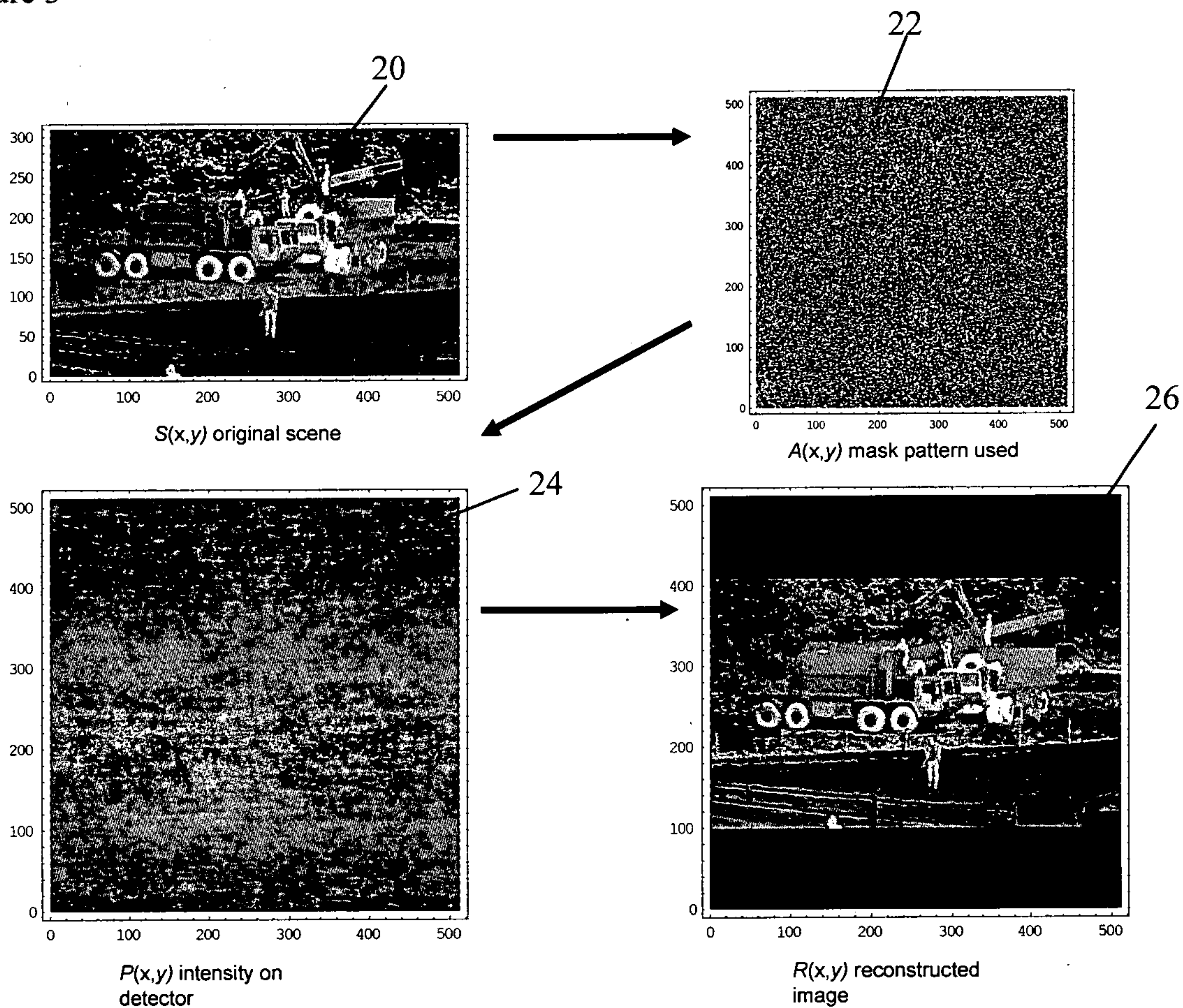


Fig. 4

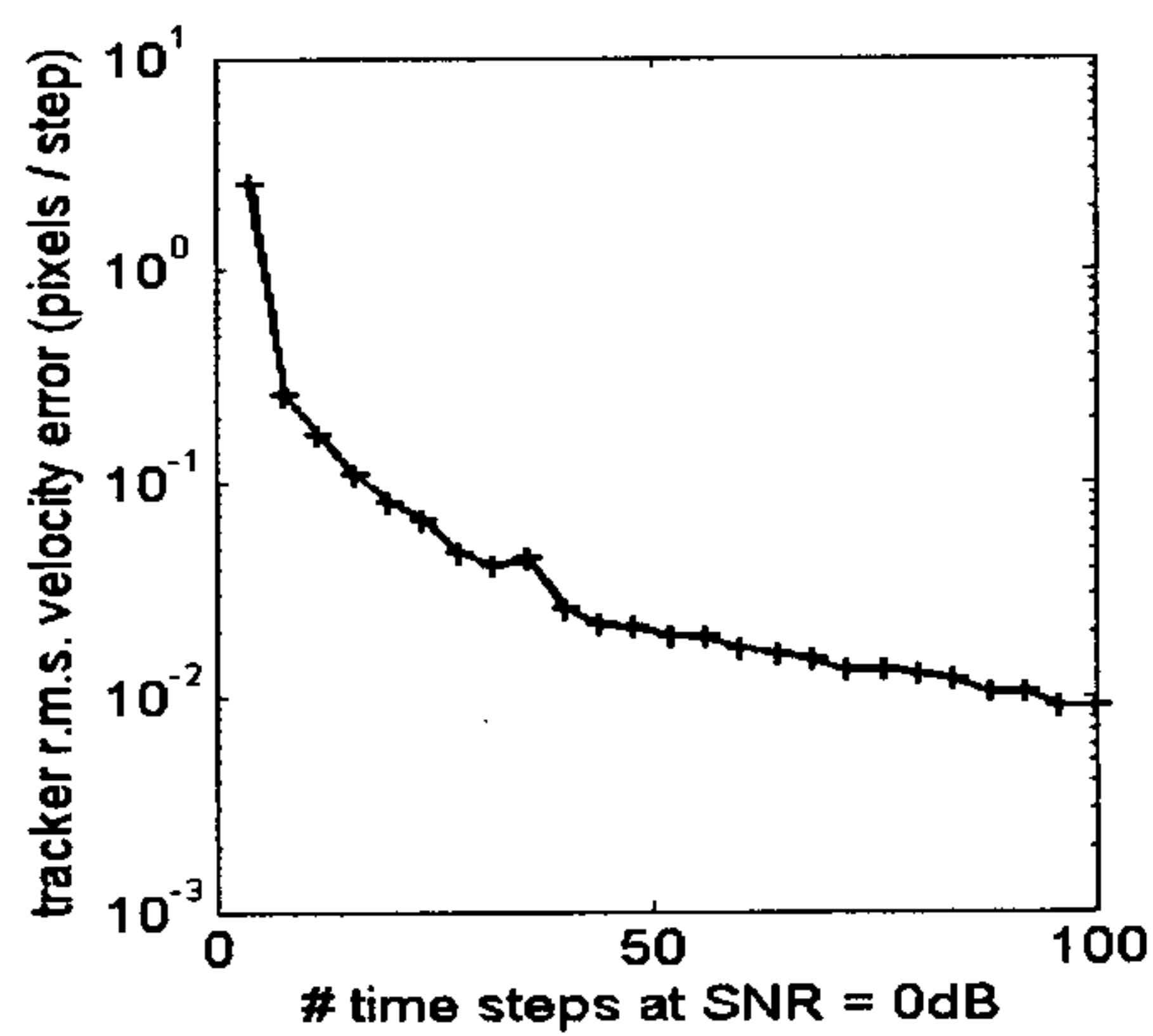
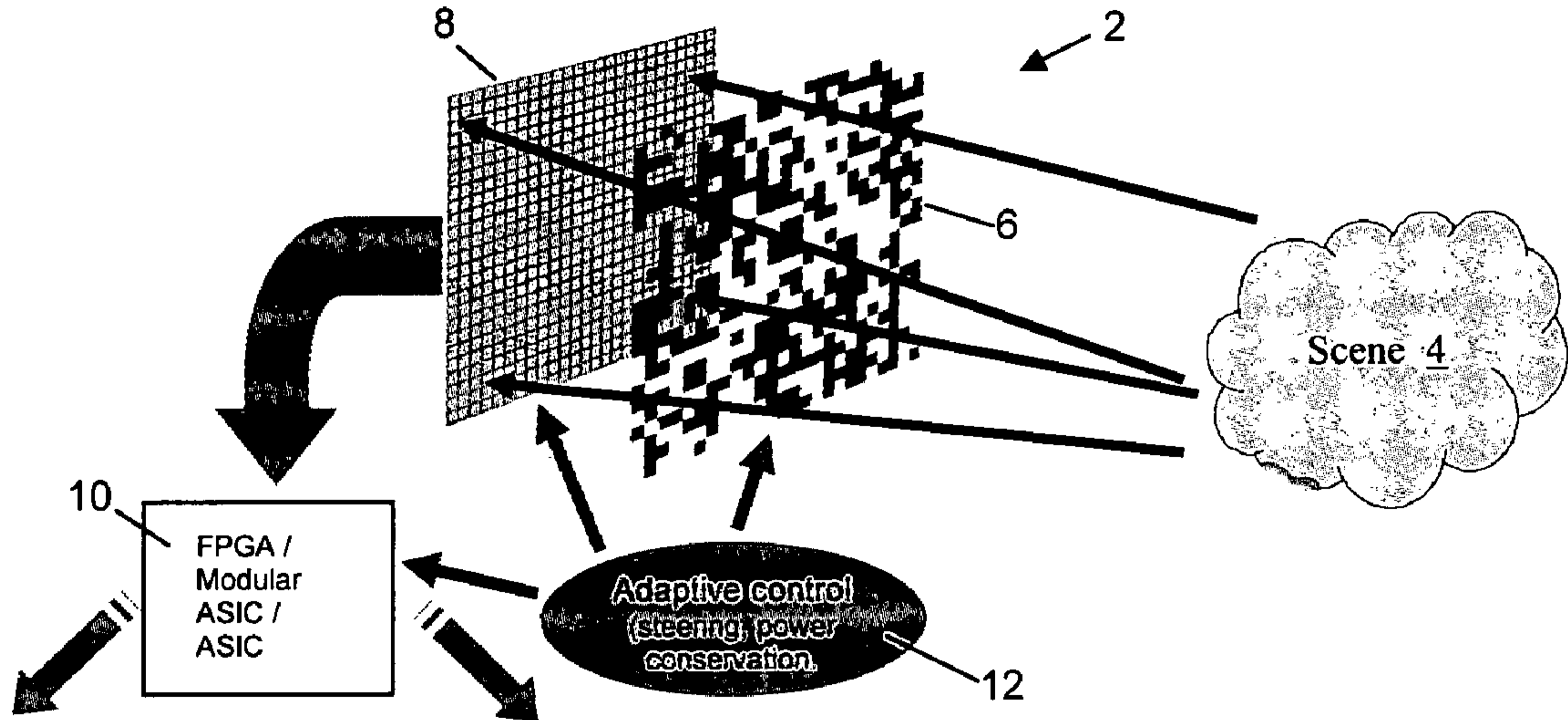


Fig. 5



To machine
(automated surveillance, detection, identification, tracking etc.)

To display(s)
and human(s)