(12) UK Patent Application (19) GB (11) 2 404 511 (13) A

(43) Date of A Publication 02.02.2005

(21) Application No: 0317541.1

(22) Date of Filing: 26.07.2003

(71) Applicant(s):
Hewlett-Packard Development Company L.P., 20555 S.H.249, Houston, Texas 77070, United States of America

(72) Inventor(s):
Maurizio Pilu

(74) Agent and/or Address for Service:
Hewlett-Packard Limited
Intellectual Property Section, Building 3, Filton Road, Stoke Gifford, BRISTOL, BS34 8QZ, United Kingdom

(51) INT CL^7:
H04N 5/232 , G03B 17/00

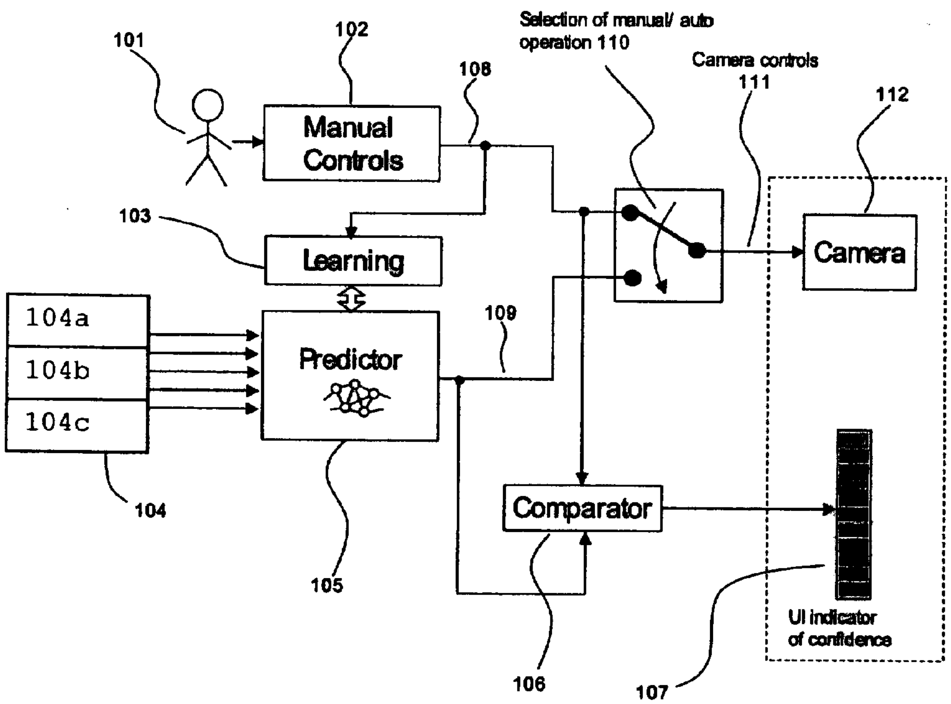
(52) UK CL (Edition X):
H4F FCCY

(56) Documents Cited:
JP 050137057 A **US 5634140 A**
US 5359385 A **US 5227835 A**

(58) Field of Search:
UK CL (Edition V) **H4F**
INT CL^7 **G03B, H04N**
Other: **Online: WPI, EPODOC, JAPIO**

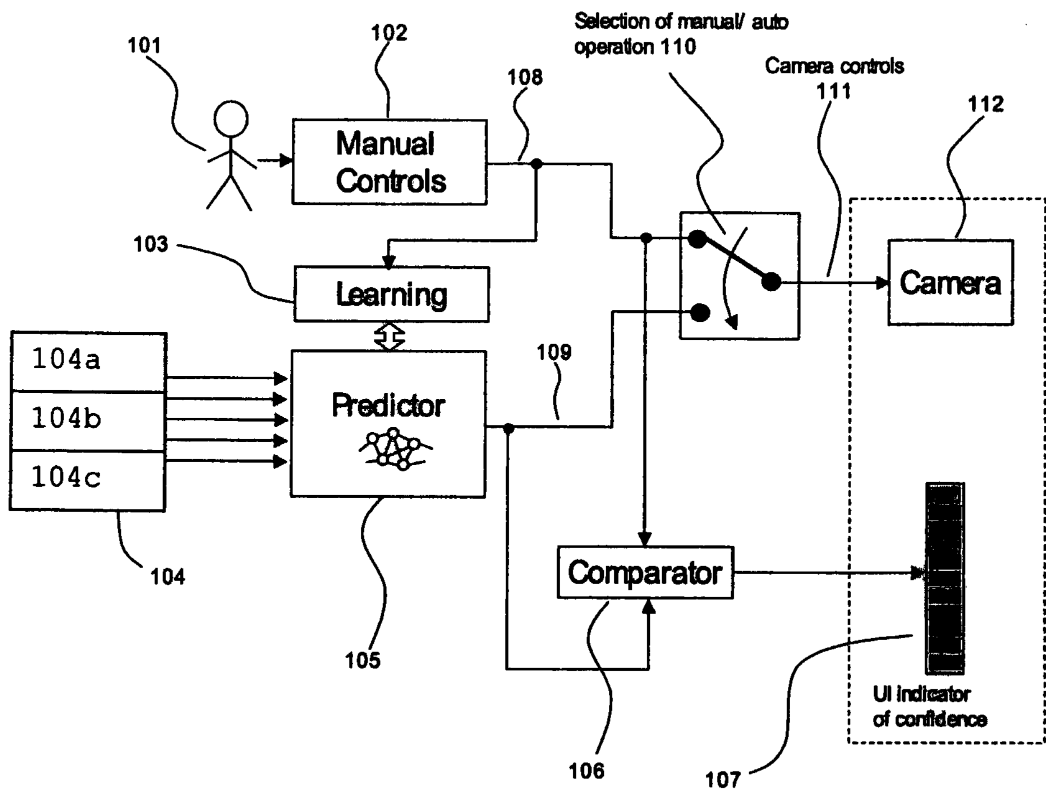
(54) Abstract Title: **Image capture device having a learning function**

(57) An image capture device 112 incorporates controls 102 for the operation of the image capture device and a learning/predictor element 103,105 , wherein the learning/predictor element is adapted to analyse use of the controls to learn a user's 101 image capture behaviour and to predict an image capture opportunity of interest to the user



GB 2 404 511 A

Original Printed on Recycled Paper



Selection of manual/ auto operation 110
Camera controls 111
101
102
108
112
Manual Controls
103
Learning
Camera
104a
104b
104c
Predictor
109
Comparator
104
105
UI indicator of confidence
106
107

Fig. 1

2404511

Image Capture Device having a Learning Function

This invention relates to an image capture device having a learning function and to a method of operating an image capture device.

It is an interest in the field of camera manufacture to look at ways to make the operation of cameras more and more automatic. Cameras are already widely available that detect various settings for a subject at which a camera is pointed to provide better focus or exposure by measuring light levels and also by determining which part of a field of view should be made the subject of camera focus.

Cameras also exist that use learning mechanisms that respond to sensorial inputs, such as an image in an image field of the camera. "Unsupervised clustering of ambulatory audio and video" by Clarkson and Pentland (1998), Proceedings of the International Conference of Acoustics, Speech and Signal Processing, Phoenix, Arizona 1999 describes a camera with a method that learns to cluster situations, such as supermarket or office, based upon audio and video clues. The training of this system uses a Hidden Markov Model (HMM) that is trained by labelling situations manually.

A further example is provided in "Context Awareness by Analysed Accelerometer Data" (Randell & Muller. Editors MacIntyre & Ianucci, The Fourth International Symposium on Wearable Computers, pp 175-176, IEEE Computer Society, October 2002) which describes user input being used to train a clustering based situation classification system.

EP 1,109,132 uses a user's opinion on the images presented, after they are captured, in a reinforcement learning framework to refine the ability of the system to predict what the user likes.

In one aspect, the invention provides a camera adapted to enable automatic prediction of user choices for camera control, the camera having: a camera controller; one or more inputs to the camera controller determined externally to the camera; one or more outputs from the camera controller relating to image capture functions of the camera; wherein the camera controller comprises a trainable computation system adapted to predict values for the one or more outputs from values of the one or more inputs upon training by receipt of user indications of suitable values of the one or more outputs for actual values of the one or more inputs.

In a further aspect, the invention provides a method of training a camera adapted to enable automatic prediction of user choices for camera control, the camera having a camera controller comprising a trainable computation system, one or more inputs to the camera controller determined externally to the camera and one or more outputs from the camera controller relating to image capture functions of the camera; providing a set of values for the one or more inputs, and indicating a preferred set of values for the one or more outputs for said set of values; repeating the preceding step at least once with at least one different set of values for the one or more inputs.

In a further aspect, the invention provides a method of image capture by a camera adapted to enable automatic prediction of a user choice for image capture, the camera having a camera controller comprising a trained computation system, one or more inputs to the camera controller determined externally to the camera and an output from the camera controller representing an image capture decision; wherein the trained computation system has been trained into a trained configuration by inputing sets of values for the one or more inputs with respective user image capture decisions, the method comprising the repeated steps of: the camera controller receiving an input set of values of the one or more inputs; the trained computation system of the camera controller calculating an image capture decision in response to the input set of values and the trained configuration of the trained computer system.

According to a further aspect of the present invention an image capture device incorporates controls for the operation of the image capture device and a learning/predictor element, wherein the learning/predictor element is adapted to analyse use of the controls to learn a user's image capture behaviour and to predict an image capture opportunity of interest to the user.

The reference to image capture behaviour is a reference to taking still images and moving images (e.g. video), which images may include audio capture also.

The learning/predictor element is preferably also adapted to predict settings of the manual controls for a given image capture opportunity.

The learning/predictor element is preferably adapted to analyse inputs from at least one sensor of the image capture device, which sensor(s) may be one or more of an image sensor, at least one motion sensor, and/or at least one biometric sensor.

The biometric sensor(s) may be one or more of a skin conductivity sensor, a brain activity sensor, a facial expression sensor and/or an eye movement sensor.

The learning/predictor element is preferably adapted to automatically control at least one, preferably a plurality, of the controls of the image capture device. Said automatic control may include an image capture control, such as a shutter release.

The image capture device preferably incorporates a prediction accuracy indicator, which may be a visual indicator. The prediction accuracy indicator is preferably adapted, in use, to compare operation of the controls with a prediction for use of the learning/predictor element. The automatic control by the learning/predictor element may be selectable by a user. The prediction accuracy indicator may indicate to a user a low level of prediction accuracy in the event that a user selects or attempts to select the automatic control of the image capture device by the learning/predictor element.

The prediction accuracy indicator may be adapted to provide an accuracy indication in conjunction with particular environmental factors, such as an indoor or outdoor location.

The learning/predictor element may be adapted to lower a threshold for an image capture decision in the event of a low level prediction accuracy. The image capture device thereby advantageously increases the likelihood of an acceptable image being obtained.

The learning/predictor element may be pre-trained with a standard set of predictions for given control/sensor inputs.

The learning/predictor element is adapted to be selectively enabled or disabled.

The learning/predictor element may comprise a learning unit and a predictor unit.

According to another aspect of the invention a method of operating an image capture device includes analysing use of controls of the image capture device with a learning/predictor element of the image capture device in order to learn a user's preferences and to predict an image capture opportunity of interest to the user.

The learning is preferably undertaken during normal use of the controls by the user.

The user may select an automatic or semi-automatic operation of the image capture device. The user's selection may be based on an output of a prediction accuracy indictor, which may give an indication of accuracy of prediction associated with a particular environment.

All of the features disclosed herein may be combined with any of the above aspects, in any combination.

For a better understanding of the invention, and to show how embodiments of the same may be brought into effect, reference will now be made, by way of example, to the accompanying diagrammatic drawing, in which:

Figure 1 is a schematic flow diagram of an image capture device having a learning function.

A camera 112 has manual controls 102 and a predictor of user interest 105, both of which produce camera control parameter signals 108, 109 respectively, such as shutter and exposure and other usual controls of a camera.

The camera 112 can be controlled manually by a user 101 using manual controls 102 to provide control signals 108 to the camera 112. The camera 112 can also be controlled automatically by the predictor 105 by signals 109 provided to the camera 112. Alternatively, the camera 112 can be operated by a combination of manual controls 102 operated by the user and by signals 109 from the predictor 105.

A logical switch 110 selects the mode of operation, which may be either fully manual using the manual controls 102 or fully automatic using the predictor 105, or a mixture of the two.

The manual controls 102 are operated by the user 101 in a conventional way. For example, zoom setting, shutter operation, time lapse, video capture (including length of video capture), video snapshot capture, audio clip capture and the usual options for a camera, be it a stills camera or a video camera.

The predictor 105 could be implemented using a neural network, a Hidden Markov Model, or another trainable method that can produce outputs in response to inputs according to parameters at least partially determined by a training process. Both neural networks and Hidden Markov Models are well known concepts for which implementations are well known. A Hidden Markov Model implementation involves the establishment of a number of Hidden Markov Models which are identified (in pre-training or by a user) against input data as representing a desired output value (or set of output values) or a null value. In the simplest case of the invention, there may be one (or more) Hidden Markov Models representing a picture taking event (or a specific type of picture taking event) and a null Model indicating that no picture taking event should occur. The predictor would then match the input data that it receives against the set of Hidden Markov Models to determine the best match, and in preferred embodiments the degree of confidence attached to that match.

In preferred embodiments, the predictor 105 makes use of relevance feedback, such as that described in "Relevance Feedback: A Power Tool in Interactive Content-based Image Retrieval," IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Segmentation, Description and Retrieval of Video Content, vol. 8 pp. 644-655, Sep. 1998, see Y. Rui, T. Huang, M. Ortega, and S. Mehrotra. Furthermore, EP 444,685 "Camera having

Learning Function" discloses a suitable learning algorithm for this implementation.

The predictor has inputs from sensors 104 which include motion sensors 104a, which are operable to detect motion of the camera 112, such as left to right motion or up/down motion, as well as camera shake. A motion sensor may be a miniature gyroscopic device such as a Murata Enc 03J, which is a piezoelectric type device.

The predictor 105 also takes input from image analysis inputs, an example of which is described in "Unsupervised clustering of ambulatory audio and video" by Clarkson and Pentland (1998), Proceedings of the International Conference of Acoustics, Speech and Signal Processing, Phoenix, Arizona 1999.

The predictor 105 also takes input from biometric sensors 104c. Various biometric sensors 104c may be used, such as those detecting skin conductivity, perhaps of a user's finger on a shutter button. The skin conductivity of a user changes when he is interested in a particular subject. A suitable implementation of the use of skin conductivity and showing interest may be found in "StartleCam: A Cybernetic Wearable Camera", in Proceedings of the International Symposium on Wearable Computers, pages 42-49, 1998, J. Healey and R.W. Picard.

The biometric sensors 104c may also include the use of brain activity measurements to measure a user's interest. Details for implementation of such a sensor are provided in "Summarising Wearable Video", IEEE Int. Conf. on Image

Processing, III:398-401, Thessaloniki, Greece, 2001, K. Aizawa, K.-I. Ishijima, M. Shiina.

Furthermore, other sensors 104 may be incorporated as input devices for the predictor 105, a face expression may be detected by one of the sensors 104 to give an indication to the camera of when an image may be of interest to the user 101 - see for example Automatic Analysis of Facial Expressions: The State of the Art, Maja Pantic, Leon J.M. Rothkrantz, IEEE Transactions of Pattern Analysis and Machine Intelligence, December 2000 (Vol. 22, No. 12),pp. 1424-1445. Also, an eye movement sensor may be provided.

As the user 101 operates the camera 112 the manually set control parameters 108 (for example when the user is taking shots or turning audio capture on) are used as examples to train the predictor 105 through a learning module 103. Thus, the predictor 105 predicts user interest against the actual input measured by the sensors 104. In this way, the camera 112 learns the user's image capture behaviour, i.e. how the user 101 is using the camera 112 and in which situations the user is taking pictures. The learning is achieved during normal use of the camera by the user 101.

A comparator 106 is also provided, which comparator receives inputs from the predictor 105 and also from the manual controls 102 via the manual control signals 108. The comparator 106 compares the predictions of the predictor 105 with the actual control signals 108 inputted by the user 101. With this comparison a confidence level is derived. The confidence level is passed for display to a graphical indicator of user confidence 107. The system may alternatively indicate confidence by an audio signal.

The user will know that the camera can be trusted to operate automatically in some situations when the predictions of the predictor 105 are consistently accurate, according to some, such as a percentage of success. A prediction failure may be logged if the predictor does not predict an image capture opportunity of interest to the user within one second of the user operating the camera controls 102. Alternatively, the camera could refuse to be set in automatic mode, for instance by omitting an audio tone, to indicate that the comparator 106 has determined that the predictor 105 is providing predictions 109 of insufficient accuracy.

The camera described herein is particularly advantageous in incorporating a learning module 103 which controls the predictor 105 to adjust the parameters output by the predictor 105 relating to the user's interest.

Furthermore, manual use of the camera 112 manual control signals 108 is used to train, either from scratch, or by fine tuning, the predictor 105 with the learning module 103 by making use of the current inputs 108 to the camera 112.

The camera 112 by use of the predictor 105 can be used to automatically operate at least some of the camera functions, by predicting when the user is interested in a particular view. The automatic operation can be to take a camera shot automatically, based upon the current and passed manually inputs 108 from the user 101.

The camera also provides a useful comparison between predicted camera operation from the predictor 105 and the actual operation (resulting from use of the manual controls 102). The comparison is indicated through the confidence graphical indicator 107, which gives the level of confidence of the camera in predicting the user preferences and operation of the camera 112.

The comparator 106 is operable to log failures as well as successes when the camera is operating in manual mode via the manual controls 102. The camera could be reporting where/when it failed, for example indoors, or when there are too many subjects for image capture, by means of analysis of the inputs from the sensors 104. In this way, the user 101 can judge by means of the confidence graphical indicator 107 whether the camera has learnt sufficiently to operate in the particular conditions. The graphical indicator 107 may include an indication of environment type associated with the level of confidence to provide this information, such as a high level for indoor situations, or a low confidence level for outdoor situations.

In a situation where the switch 110 is put to automatic mode when the predictor 105 has a poor confidence level, the statistics resulting from the failure could be stored to modify how the predictor 105 operates. For example, if a situation has been detected where the predictor 105 is likely to fail, then the camera 112 will lower a threshold for an image capture decision to thereby operate by taking more shots than would otherwise normally be required, to compensate for poor predicted performance.

The camera 112 may be "pre-trained" based on an average user before being provided to an end user, in which situation it will be necessary for an end user 101 only to fine tune the predictor 105 parameters by normal use. Another option for the camera 112 is to allow the user 101 to switch off the training mode via the graphical indicator 107 which graphical indicator doubles as a user interface.

The camera described herein can learn when to take pictures and how those pictures should be taken based on manual inputs.

The advantageous provision of a method by which the camera can indicate to a user 101 when it is ready for automatic operation, via the comparator and display 107 has significant advantages.

CLAIMS:

1. A camera adapted to enable automatic prediction of user choices for camera control, the camera having:

a camera controller;

one or more inputs to the camera controller determined externally to the camera;

one or more outputs from the camera controller relating to image capture functions of the camera;

wherein the camera controller comprises a trainable computation system adapted to predict values for the one or more outputs from values of the one or more inputs upon training by receipt of user indications of suitable values of the one or more outputs for actual values of the one or more inputs.

2. A camera as claimed in claim 1, wherein one of said one or more outputs is an image capture command signal.

3. A camera as claimed in claim 1 or claim 2, wherein one or more of said one or more outputs are parameters for image capture.

4. A camera as claimed in any preceding claim, wherein the camera is adapted such that the one or more outputs are used to determine image capture.

5. A camera as claimed in claim 4, further comprising a user interface by which a user can enable or disable use of the one or more outputs to determine image capture.

6. A camera as claimed in any preceding claim, wherein one or more of the one or more inputs are signals sensed from a body of a user.

7. A camera as claimed in any preceding claim, wherein one or more of the one or more inputs are signals determined by analysis of images received by the camera.

8. A camera as claimed in any preceding claim, wherein the trainable computation system has been pre-trained with reference to standard data.

9. A camera as claimed in any preceding claim, wherein the camera is adapted to be trained by the user by user confirmation of appropriate output values.

10. A camera as claimed in any preceding claim, further comprising means to display to a user a confidence value in respect of the one or more outputs.

11. A camera as claimed in any preceding claim, wherein the trainable computation system uses a plurality of Hidden Markov Models.

12. A method of training a camera adapted to enable automatic prediction of user choices for camera control, the camera having a camera controller comprising a trainable computation system, one or more inputs to the camera controller determined externally to the camera and one or more outputs from the camera controller relating to image capture functions of the camera;

providing a set of values for the one or more inputs, and indicating a preferred set of values for the one or more outputs for said set of values;

repeating the preceding step at least once with at least one different set of values for the one or more inputs.

13. A method as claimed in claim 12, wherein at least one of the sets of values comprises standard data with which the camera is pre-trained before provision to a user.

14. A method as claimed in claim 12 or claim 13, wherein one of said one or more outputs is an image capture command signal.

15. A method as claimed in any of claims 12 to 14, wherein one or more of said one or more outputs are parameters for image capture.

16. A method as claimed in any of claims 12 to 15, wherein one or more of the one or more inputs are signals sensed from a body of a user.

17. A method as claimed in any of claims 12 to 16, wherein one or more of the one or more inputs are signals determined by analysis of images received by the camera.

18. A method of image capture by a camera adapted to enable automatic prediction of a user choice for image capture, the camera having a camera controller comprising a trained computation system, one or more inputs to the camera controller determined externally to the camera and

an output from the camera controller representing an image capture decision;

wherein the trained computation system has been trained into a trained configuration by inputing sets of values for the one or more inputs with respective user image capture decisions, the method comprising the repeated steps of:

the camera controller receiving an input set of values of the one or more inputs;

the trained computation system of the camera controller calculating an image capture decision in response to the input set of values and the trained configuration of the trained computer system.

19. A method as claimed in claim 18, wherein the image capture decision comprises a decision to capture an image.

20. A method as claimed in claim 18 or claim 19, wherein the image capture decision comprises a decision concerning parameters for capture of an image.

21. A method as claimed in any of claims 18 to 20, wherein the method comprises using the image capture decision to determine image capture.

22. A method as claimed in any of claims 18 to 20, further comprising a user enabling or disabling user of the image capture decision to determine image capture.

23. A method as claimed in any of claims 18 to 22, wherein one or more of the one or more inputs are signals sensed from a body of a user.

24. A method as claimed in any of claims 18 to 23, wherein one or more of the one or more inputs are signals determined by analysis of images received by the camera.

25. An image capture device incorporates controls for the operation of the image capture device and a learning/predictor element, wherein the learning/predictor element is adapted to analyse use of the controls to learn a user's image capture behaviour and to predict an image capture opportunity of interest to the user.

26. An image capture device as claimed in claim 25, which incorporates a prediction accuracy indicator.

27. An image capture device as claimed in either claim 25 or claim 26, in which the learning/predictor element is adapted to predict settings of the manual controls for a given image capture opportunity.

28. An image capture device as claimed in any of claims 25 to 27, in which the learning/predictor element is adapted to analyse inputs from at least one sensor of the image capture device.

29. An image capture device as claimed in claim 28, in which said sensors may be one or more of an image sensor, at least one motion sensor, and/or at least one biometric sensor.

30. An image capture device as claimed in claim 29, in which the or each biometric sensor may be one or more of a skin conductivity sensor, a brain activity sensor, a facial expression sensor and/or an eye movement sensor.

31. An image capture device as claimed in any of claims 25 to 30, in which the learning/predictor element is adapted to automatically control at least one of the controls of the image capture device.

32. An image capture device as claimed in any of of claims 26 to 31, in which the prediction accuracy indicator is adapted, in use, to compare operation of the controls with a prediction for use of the learning/predictor element.

33. An image capture device as claimed in claim 31, in which the automatic control by the learning/predictor element is selectable by a user.

34. An image capture device as claimed in claim 33, in which the prediction accuracy indicator indicates to a user a low level of prediction accuracy in the event that a user selects or attempts to select the automatic control of the image capture device by the learning/predictor element.

35. An image capture device as claimed in any one of claims 26 to 34, in which the prediction accuracy indicator is adapted to provide an accuracy indication in conjunction with particular environmental factors.

36. An image capture device as claimed in any one of claims 26 to 35, in which the learning/predictor element is adapted to lower a threshold for an image capture decision in the event of a low level prediction accuracy.

37. An image capture device as claimed in any one of claims 28 to 36, in which the learning/predictor element is pre-trained with a standard set of predictions for given control/sensor inputs.

38. An image capture device as claimed in any of claims 25 to 37, in which the learning/predictor element is adapted to be selectively enabled or disabled.

39. A method of operating an image capture device includes analysing use of controls of the image capture device with a learning/predictor element of the image capture device in order to learn a user's image capture behaviour and to predict an image capture opportunity of interest to the user.

40. A method as claimed in claim 39, in which the learning is undertaken during normal use of the controls by the user.

41. A method as claimed in either claim 39 or claim 40, in which the user selects an automatic or semi-automatic operation of the image capture device.

42. An image capture device substantially as described herein with reference to the accompanying drawing.

43. A method of operating an image capture device substantially as described herein with reference to the accompanying drawing.

Application No: GB 0317541.1
Claims searched: 1 - 24
Examiner: Richard Baines
Date of search: 8 December 2003

Patents Act 1977 : Search Report under Section 17

Documents considered to be relevant:

Category	Relevant to claims	Identity of document and passage or figure of particular relevance
X	1 - 5, 7, 12, 14, 15, 17 - 22 & 24	US 5,227,835 (KODAK) - whole document
X	1 - 5, 7 - 9, 12 - 15, 17 - 22 & 24	US 5,359,385 (MINOLTA) - whole document
X	1, 3, 4, 7, 8, 12, 13, 15, 17, 18, 20, 21 &24	US 5,634,140 (MINOLTA) - whole document
A	-	JP 5,137,057 (HITACHI)

Categories:

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.

Field of Search:

Search of GB, EP, WO & US patent documents classified in the following areas of the UKC[V]:

H4F

Worldwide search of patent documents classified in the following areas of the IPC[7]:

H04N, G03B

The following online and other databases have been used in the preparation of this search report:

Keywords in EPODOC, WPI, JAPIO, INTERNET