

**Internal distribution code:**

- (A)  Publication in OJ  
(B)  To Chairmen and Members  
(C)  To Chairmen  
(D)  No distribution

**Datasheet for the decision  
of 5 November 2009**

**Case Number:** T 1349/05 - 3.5.04

**Application Number:** 02255507.2

**Publication Number:** 1283499

**IPC:** G06T 7/20

**Language of the proceedings:** EN

**Title of invention:**

Moving object tracking apparatus and method

**Applicant:**

SAMSUNG ELECTRONICS CO., LTD.

**Headword:**

-

**Relevant legal provisions:**

-

**Relevant legal provisions (EPC 1973):**

EPC Art. 56

**Keyword:**

"Inventive step (main request, first and second auxiliary requests - no)"

"Inventive step (third auxiliary request - yes)"

**Decisions cited:**

-

**Catchword:**

-



Case Number: T 1349/05 - 3.5.04

**D E C I S I O N**  
of the Technical Board of Appeal 3.5.04  
of 5 November 2009

**Appellant:** SAMSUNG ELECTRONICS CO., LTD.  
416 Maetan 3-dong  
Yeongtong-gu  
Suwon-si  
Gyeonggi-do 442-742 (KR)

**Representative:** Walaski, Jan Filip  
Venner Shipley LLP  
20 Little Britain  
London EC1A 7DH (GB)

**Decision under appeal:** Decision of the Examining Division of the  
European Patent Office posted 30 May 2005  
refusing European application No. 02255507.2  
pursuant to Article 97(1) EPC 1973.

**Composition of the Board:**

**Chairman:** F. Edlinger  
**Members:** M. Paci  
B. Müller

## Summary of Facts and Submissions

- I. The appeal is against the decision of the examining division refusing European patent application No. 02 255 507.2, which was published as EP 1 283 499 A1.
- II. The following document was cited as prior art in the decision under appeal:
- D1: US 5 867 584 A
- III. The decision under appeal was based on the ground that claim 1 did not involve an inventive step (Article 56 EPC 1973) having regard to the disclosure of D1.
- IV. With the statement of grounds of appeal the appellant (applicant) filed three sets of amended claims according to a main request, first auxiliary request and second auxiliary request, replacing all previous claims.
- V. In a communication accompanying the summons to oral proceedings the board referred to the following additional prior art documents:
- D2: KR 2001 0000107 A and  
D3: G. Halevy et al., "Motion of disturbances: detection and tracking of multi-body non-rigid motion", Machine Vision and Applications, 11(3):122-137, 1999.

Since D2 was written in Korean, the board also introduced the following machine translation and post-

published patent family member as evidence of the disclosure of D2:

D2a: Machine translation of D2 into English provided by the Korean Intellectual Property Office (KIPO) and  
D2b: WO 01/84844 A1.

In the communication accompanying the summons to oral proceedings the board expressed the preliminary opinion that the subject-matter of claim 1 according to each of the requests on file was obvious from the combined teachings of D1 and either D2 or D3.

- VI. With a fax dated 5 October 2009 the appellant filed further sets of claims according to a third, fourth and fifth auxiliary request, respectively.
- VII. During the oral proceedings held on 5 November 2009 before the board the appellant filed a set of amended claims replacing all previous claims according to the third auxiliary request and filed adapted description pages.
- VIII. The appellant's final requests are that the decision under appeal be set aside and that a patent be granted on the basis of the claims submitted in the appeal proceedings in the order of the main request to the fifth auxiliary request, namely the main request, first and second auxiliary requests as filed with the statement of grounds of appeal, the third auxiliary request as filed in the oral proceedings and the fourth and fifth auxiliary requests filed with the fax of 5 October 2009.

IX. Independent claim 1 according to the main request reads as follows:

"A method of tracking a moving object with a camera, the camera providing a plurality of image signals representing a scene at different respective times, the method comprising:

automatically acquiring a moving object to be tracked by detecting movement of the object within a scene from the plurality of image signals representing said scene, using a tracking window having a predetermined size;

adjusting the size of the tracking window in dependence on the size of the detected moving object;

predicting a new position for said moving object;

selecting a portion of a further, later image signal, representing said scene, in dependence on the size of said tracking window and said predicted new position;

analysing said selected portion to determine the correctness of said prediction; and

if said correctness does not meet a predetermined criterion, searching the whole of said further image signal for said moving object."

X. Independent claim 1 according to the first auxiliary request reads as follows:

"A method of tracking a moving object with a camera, the method comprising:

automatically detecting a moving object within a scene from a plurality of image signals representing said scene at respective different times from a camera, using a tracking window having a predetermined size;

adjusting the size of the tracking window in dependence on the size of the detected moving object;  
predicting a new position for said moving object;  
selecting a portion of a further, later image signal, representing said scene, said portion corresponding to the size-adjusted tracking window centred on said predicted new position;

processing only an image signal corresponding to the moving object within said size adjusted tracking window to determine the correctness of said prediction;  
and

if said correctness does not meet a predetermined criterion, searching the whole of said further image signal for said moving object."

XI. Independent claim 1 according to the second auxiliary request reads as follows:

"A method of tracking a moving object, comprising:

filming a monitored area using a camera (10);  
generating a binary disturbance image signal from an input image signal acquired from the camera (10) representing a sequence of frames;

acquiring from the binary disturbance image signal information about the moving object using a moving window, having a predetermined size, as an initial tracking window;

adjusting the size of the initial tracking window so that the binary disturbance image signal contains the moving object;

predicting information about the location of the moving object in a subsequent frame to which the centre of the moving object is to move based on currently

acquired information and previously acquired information about the centre of the moving object;  
moving the centre of the moving window to the location where the centre of the moving object is predicted to move to;  
acquiring actual information about an actual centre of the moving object in the moving window and the size of the moving window; and  
comparing the actual information about the moving object with the predicted information about the moving object, and determining the tracking status of the moving object based on a resultant error range of the predicted and actual information, the actual information about the moving object being acquired from the subsequent frame."

XII. Independent claims 1 and 12 according to the third auxiliary request read as follows:

"1. A method of tracking a moving object, comprising:  
filming a monitored area using a camera (10);  
generating a binary disturbance image signal from an input image signal acquired from the camera (10) representing a sequence of frames;  
acquiring from the binary disturbance image signal information about the moving object using a moving window, having a predetermined size, as an initial tracking window;  
adjusting the size of the initial tracking window so that the binary disturbance image signal contains the moving object;  
predicting information about the location of the moving object in a subsequent frame to which the centre of the moving object is to move based on currently

acquired information and previously acquired information about the centre of the moving object;

moving the centre of the tracking window to the location where the centre of the moving object is predicted to move to;

acquiring actual information from the subsequent frame about an actual centre of the moving object in the tracking window and the size of the tracking window; and

comparing the actual information about the moving object with the predicted information about the moving object, and determining the tracking status of the moving object based on a resultant error range of the predicted and actual information, further comprising:

performing zoom operations on the camera so that the size of the tracking window located at the predicted location and the size of the moving object acquired from the subsequent frame are maintained at a certain ratio."

"12. A tracking system comprising a camera (10) and processing means (120, 130, 140, 150, 160) configured to process image signals from the camera (10), **characterised in that** the processing means (120, 130, 140, 150, 160) is configured for causing the system to perform a method according to any preceding claim."

Claims 2 to 11, 13 and 14 according to the third auxiliary request are dependent on either claim 1 or claim 12.

XIII. The examining division's reasoning in the decision under appeal with respect to claim 1 then on file can be summarised as follows.



The method of claim 1 differs from the tracking method disclosed in D1 only in that:

- (a) the method is for tracking objects with a camera  
and
- (b) if no match is found the whole image is searched.

As to feature (a), the expression "tracking a moving object with a camera" is not considered to imply that the camera can follow the target, only that the camera is used for generating the claimed image signals. It is considered obvious to the skilled person that the method of D1 can be applied to image signals coming from a camera. Besides, no features relating to camera control are set out in claim 1.

As to feature (b), D1 does not disclose that if the criterion is not met, the whole picture is searched. However it discloses that if the system fails to track the object in a frame, the system will warn the user. Alternatively, a user can set a very low matching threshold or respecify the object window. Thus it is considered to be within the capabilities of the skilled person aware of D1 to respecify the window so as to look elsewhere in the picture, possibly searching the whole picture if tracking is lost.

Hence the subject-matter of claim 1 does not involve an inventive step (Article 56 EPC 1973).

XIV. The appellant essentially argued as follows before the board of appeal regarding the main request and first, second and third auxiliary requests.

*Main request*

Claim 1 according to the main request differs from that of the appealed decision essentially in that it has been clarified that the objects to be tracked are automatically acquired by detecting their movement.

D1 is not the closest prior art for the method of claim 1 because it discloses a method of tracking an object in a pre-recorded video sequence, not a method of tracking a moving object with a camera. Instead, D2 should be the starting point for the assessment of inventive step.

If D1 is nevertheless regarded as the closest prior art, the method of claim 1 differs from the method of D1 at least by the following features identified by the board in the communication accompanying the summons to oral proceedings:

- (a) a moving object is tracked with a camera which provides the plurality of image signals;
- (b) a moving object to be tracked is acquired by detecting movement of the object within the scene;
- (c) if the correctness does not meet a predetermined criterion, the whole of said further image signal is searched for said moving object.

In addition to these differences D1 also does not disclose the step of adjusting the size of the tracking window in dependence upon the size of the detected moving object.

These distinguishing features are not obvious from D1. In D1 the objects to be tracked are identified either

manually by the user or automatically by pattern recognition. There is no suggestion in D1 to automatically identify objects to be tracked by the fact that they are moving. Equally, there is nothing in D1 to suggest the solution of searching the whole picture if tracking is lost. The examining division's reasoning with respect to these features is thus based on hindsight.

Furthermore the skilled person would have no reason to combine the teaching of D1 with those of D2 or D3. Even if he/she had, D2 would still not disclose searching the whole of said further image signal when a tracked object is lost. In D2 when a tracked object stops moving, a template matching method is used for tracking it further, thereby leading away from performing a search in the whole image.

Accordingly, the method of claim 1 would not have been obvious to the skilled person.

*First auxiliary request*

Claim 1 according to the first auxiliary request differs from claim 1 considered in the appealed decision essentially in that the size-adjusted tracking window is centred on said predicted position and that the processing of the image signal corresponding to the moving object is performed within this size-adjusted tracking window in order to determine the correctness of said prediction.

The method of claim 1 allows a quicker decision to be made about the correctness of the prediction than the method of D1, which relies on a set of several tracking windows to check the correctness of the prediction.

*Second auxiliary request*

Claim 1 according to the second auxiliary request specifies, in particular, that the moving objects are acquired using a binary disturbance signal. This feature is not derivable from D1.

*Third auxiliary request*

Claim 1 according to the third auxiliary request further specifies that the method comprises the step of performing zoom operations on the camera so that the size of the tracking window located at the predicted location and the size of the moving object acquired from the subsequent frame are maintained at a certain ratio.

No camera is mentioned in D1 and no zooming either. Since the video sequence in D1 is pre-recorded, no zooming operation can be performed in reaction to the tracking operation.

**Reasons for the Decision**

1. The appeal is admissible.

*Main request*

2. Amendments

Claim 1 according to the main request differs from that of the appealed decision essentially in that it has been clarified that the objects to be tracked are automatically acquired by detecting their movement.

3. Inventive step (Article 56 EPC 1973)

3.1 Closest prior art

The appellant has disputed that D1 represents the closest prior art because it is not a method of tracking a moving object with a camera but a method of tracking a moving object in a pre-recorded video sequence.

The board does not share the appellant's conclusion. Although D1 does not explicitly mention a camera, the image signals must be generated somewhere and D1 gives several examples of video sequences which must have been generated by a camera (see "Pro Football's greatest games" on column 1, lines 51 to 53, and "marine video" on column 4, lines 46 to 49). The appellant acknowledged during the oral proceedings before the board that tracking of a moving object in claim 1 according to the main request could be performed with a fixed camera whose only role in the method of claim 1 would be to generate the image signals. Hence the method steps of claim 1 could be performed with pre-recorded camera image signals. The

board thus sees no reason to exclude D1 as a starting point for the assessment of inventive step.

As to D2, it should be noted that, since its disclosure is in Korean, the board will also refer to D2a (a machine translation of D2 into English provided by the Korean Intellectual Property Office) and to D2b (a patent family member of D2 having a disclosure apparently closely matching that of D2 but published after the priority date of the present application). The appellant has not disputed these facts. The relevant disclosure of D2 is also described in paragraphs [0005] to [0011] of the present published application.

### 3.2 Disclosure of D1

D1 discloses a tracking method with the following steps:

- tracking a moving object in a plurality of image signals representing a scene at different respective times (*see column 2, lines 43 to 48*);
- automatically acquiring a moving object to be tracked using a tracking window having a predetermined size (*see column 4, lines 3 to 6 and 43 to 51*);
- adjusting the size of the tracking window in dependence on the size of the detected moving object (*see column 4, lines 6 to 10 and 44 to 49*);
- predicting a new position for said moving object (*see column 5, lines 1 to 3*);
- selecting a portion of a further, later image signal, representing said scene, in dependence on the size of said tracking window and said predicted new position (*see column 5, lines 40 to 49*);

- analysing said selected portion to determine the correctness of said prediction (*see from column 5, line 50, to column 6, line 20*).

### 3.3 Distinguishing features

The method of claim 1 therefore differs from the method of D1 in that:

- (a) a moving object is tracked with a camera which provides the plurality of image signals;
- (b) a moving object to be tracked is acquired by detecting movement of the object within the scene;
- (c) if the correctness does not meet a predetermined criterion, the whole of said further image signal is searched for said moving object.

### 3.4 Objective technical problem

The appellant stated during the oral proceedings that the objective technical problem was to improve the efficiency of the tracking method of D1. The board has no objection to this general formulation of the objective problem.

### 3.5 Obviousness

Regarding feature (a), the appellant confirmed during the oral proceedings that the camera mentioned in this feature could be a fixed camera, a situation which would have been obvious, if not implicit, in the system of D1 for the reasons presented in section 3.1, second paragraph, *supra*.

As to feature (b), the board agrees with the appellant that in D1 the objects to be tracked are acquired either manually by the user (see column 4, lines 3 to 8) or automatically by pattern recognition (see column 4, lines 44 to 49), but not automatically based on the detection of the movement of the object within a scene. However the skilled person would have considered other known acquisition methods as alternatives depending on the circumstances of the intended usage. The present application acknowledges that various types of tracking systems for automatically detecting and tracking objects were known at the priority date, such as object detection and tracking methods based on movement detection using a "disturbance map" (see paragraphs [0002] and [0003] of the present published application). The advantages and disadvantages of the various methods and the types of scenes in which they are most efficiently used were well known in the art, as discussed, for instance, in D2 and D3 (see D2b, page 9, lines 11 to 17, and D2a, section 2.2, second paragraph, and D3, Abstract and figure 1). The skilled person was therefore well aware that motion detection algorithms, in particular those using a disturbance map, were very effective at automatically acquiring objects in video sequences in which the objects of interest are moving. It would thus have been desirable for the skilled person to use such an algorithm for automatically detecting moving objects, as in D1, instead of the pattern recognition algorithm mentioned in column 4, lines 43 to 51, of D1. For the above reasons, the skilled person would have arrived at feature (b) in an obvious manner.



With regard to feature (c), when an object is lost during tracking in D1, i.e. when the correctness of the prediction of the new position of the object does not meet a predetermined criterion, the system displays a warning on the screen and awaits further commands from the user (see column 6, lines 35 to 39). Such a course of action is logical because in this embodiment of D1 the object to be tracked must first be manually selected by the user (see column 4, lines 6 to 10). Thus, when the system has lost the object to be tracked, it must ask the user to find it again. However in a system in which the object is automatically identified it would be desirable to also automatically track the moving object as long as possible and to extend the search to the whole of the further image when the tracking performed in the (small) tracking window has lost the object, for example by selecting a set of test windows which covers the whole of the frame (see D1, column 5, lines 43 to 48). If the moving object cannot be found in the further image signal it would become necessary to repeat the initial automatic object (motion based) identification over the whole of the scene. This is the equivalent of the user scanning the whole scene with his/her eyes to find the object.

For the above reasons, the method of claim 1 according to the main request does not involve an inventive step in view of D1 and the teaching of either D2 or D3.

### 3.6 Appellant's arguments

The appellant argued that D1 did not disclose the step of adjusting the size of the tracking window in dependence upon the size of the detected moving object.

The board disagrees. D1 discloses that when the objects are detected by the user, the user uses the pointer to define a tracking window around the object so that the tracking window envelops the object (see *column 4, lines 6 to 10*). When the object is automatically detected by the system, instead of manually, the tracking window is specified by the system (see *column 4, lines 44 to 49*). Although D1 does not explicitly repeat that in this case too the tracking window is defined "around the object so that the tracking window envelops the object", this feature is regarded as implicit - and in any case obvious - because it is the same tracking window which is referred to.

The appellant also argued that, when a tracked object is lost in D2 because it has stopped moving, a template matching method is used for tracking it further, thereby leading away from performing a search in the whole image.

The board is not convinced by this argument. Claim 1 does not specify how the whole of the further image signal is searched for moving objects. A template matching algorithm can be expected to find stopped objects but will likely fail to track moving objects which have suddenly rotated and changed direction. Hence template matching does not remove the need for searching the whole image for moving objects.

4. Accordingly, the main request is not allowable.

*First auxiliary request*

5. Amendments

Claim 1 according to the first auxiliary request differs from claim 1 considered in the appealed decision in that it further specifies that the portion selected in the later image signal corresponds to the size-adjusted tracking window centred on the predicted new position and that only an image signal corresponding to the moving object within said size-adjusted tracking window is processed to determine the correctness of the prediction. These amendments are based on page 7, lines 9 to 11, of the application as filed (page 4, lines 16 and 17, of the published application).

6. Inventive step (Article 56 EPC 1973)

In D1 the system creates a set of test windows composed of every possible window of the same size and shape as the tracking window and having a centre pixel less than a predefined distance from the predicted centre point of the tracked object in the next frame (see column 5, lines 43 to 48). The intensity distance between the predicted tracking window and each of the test windows is then calculated and the test window with the lowest intensity distance (i.e. the highest correctness) is selected as the best match window (see from column 5, line 50, to column 6, line 20).

In other words, the method of D1, like the method of claim 1 of the first auxiliary request, comprises the step of determining the correctness of the prediction

for the tracking window centred on the predicted centre in the next frame. However the method of D1 goes further than in claim 1 in that it also checks the correctness of the prediction for all the other windows (the "test windows") whose centre is offset from the predicted centre by up to a predefined distance. This last step allows better matching windows to be found (and thus the tracking to be improved), but obviously at the cost of some additional processing. The skilled person would have been aware of this trade-off between tracking accuracy and computing speed in the method of D1. The board thus considers that, for tracking systems in which the computing speed is more important than the fine-tuning of the tracking accuracy, the skilled person would have adapted the method of D1 to have only one test window, i.e. the tracking window centred on the predicted centre, if the prediction is sufficiently good and to extend the processing only to other areas of the whole image signal if the correctness does not meet a predetermined criterion.

For the above reasons, the method of claim 1 according to the first auxiliary request does not involve an inventive step.

7. Accordingly, the first auxiliary request is not allowable.

*Second auxiliary request*

8. Amendments

The method of claim 1 according to the second auxiliary request is based on the method of claim 17 of the

application as filed. Its steps, except for the missing last step, generally correspond to those of claim 1 of the appealed decision, albeit worded in a more limited manner with explicit references to a binary disturbance image signal and to the centre of the moving object and the centre of the tracking window.

9. Inventive step (Article 56 EPC 1973)

The board regards the method of claim 1 according to the second auxiliary request as obvious from the combined teachings of D1 and either D2 or D3 for the reasons given in section 3 *supra* and because the methods of D2 and D3 are based on a disturbance map and the method of D1 uses the centres of the moving object and of the tracking window for tracking the object (see D1, column 5, lines 1 to 3 and 40 to 49). The feature that the disturbance map is "binary" (i.e. the disturbance map shows the moving objects in black and the fixed background in white, or vice versa) is known from D2 (see section 2.3 of D2a, the paragraph bridging pages 9 and 10 in D2b and black and white images in figure 5 of D2a and D2b which represent "binary disturbance maps") and is regarded as obvious from D3 (see page 129, first paragraph, the appearance/disappearance of a disturbance being detected by comparison to a threshold level, thus yielding a binary result which could be represented as a binary disturbance map).

10. Accordingly, the second auxiliary request is not allowable.

*Third auxiliary request*

11. Admissibility of the third auxiliary request

Claim 1 according to the third auxiliary request differs from claim 1 according to the second auxiliary request essentially by the addition of the step of performing zoom operations on the camera so that the size of the tracking window located at the predicted location and the size of the moving object acquired from the subsequent frame are maintained at a certain ratio.

The remaining claims 2 to 14 are either the same as those in the second auxiliary request or contain some minor clarifying amendments. Description pages were also filed to adapt the description to the claims and to acknowledge D3.

The board considered that this request, filed one month before the oral proceedings and amended for clarification during the oral proceedings, was a limitation of the subject-matter for which protection was sought by the second auxiliary request filed with the statement of grounds of appeal and did not add undue complexity to the case. Moreover the request was *prima facie* likely to overcome the outstanding objections. For these reasons, the board decided to exercise its discretion under Article 13(1) RPBA in admitting the third auxiliary request into the proceedings.

12. *Amendments (Article 123(2) EPC)*

The amendments made to claim 1 are derivable from claims 17 and 26 and page 12, paragraph 2, of the application as filed. The board is therefore satisfied that the requirements of Article 123(2) EPC are met.

13. Inventive step (Article 56 EPC 1973)

*Claim 1*

Claim 1 according to this request includes the step of performing zoom operations on the camera so that the size of the tracking window located at the predicted location and the size of the moving object acquired from the subsequent frame are maintained at a certain ratio. This step, in combination with the other steps of the method of claim 1, provides at least the following advantages:

- (a) The tracking method can easily adapt to processing very small objects (see paragraph [0045] of the published application).
- (b) The size of the object remains constant from one frame to the next even if the object is travelling towards the camera or away from it, which can be very useful in circumstances where a numeral or a character is to be checked (see paragraph [0048] of the published application).
- (c) Maintaining the ratio allows the moving object to be clearly recognized in a monitored area (see paragraph [0052] of the published application).

All the examples of video sequences given in D1 are apparently pre-recorded, thus leaving plenty of time to the user or the system to perform the manual or automatic detection of the objects and the subsequent tracking. There is no suggestion in D1 that the detection and tracking process could be performed in real time on a video sequence of a scene currently being filmed. Moreover, in the examples given (Pro Football's greatest games or an educational programme on marine life), there is no control chain from the moving object tracking system to the camera. As a consequence, the system of D1 does not perform zoom operations on the camera to change the content of the next frame because the next frame has already been filmed when the current frame is being analysed by the system. In other words, the step of performing zoom operations in the method of claim 1 is simply impossible in the system of D1, unless this system is made to work on real time video sequences from a local camera, a situation not contemplated in D1.

D2 discloses the presence of a zoom on the camera only on page 6, lines 26 to 28, and on page 26, lines 6 to 19. However D2 neither mentions nor suggests performing the zoom operation in such a way that the ratio of the size of the object to the size of the tracking window is maintained at a certain value from one frame to the next. Moreover, for the reasons set out in the previous paragraph, zooming operations during the tracking operation are not possible in D1. Thus without knowledge of the present invention any teaching in this direction in D2 would in any case not be applicable to the method of D1.



D3 does not mention any zoom operation.

For the above reasons, the method of claim 1 according to the third auxiliary request is not rendered obvious by the combined teachings of D1, D2 and D3.

*Claim 12*

The tracking system of claim 12 according to the third auxiliary request comprises processing means configured to cause the system to perform a method according to claim 1. This system is therefore also not rendered obvious by D1, D2 and D3 for the reasons set out above with respect to claim 1.

*Claims 2 to 11, 13 and 14*

These claims are dependent either on claim 1 or claim 12. Hence their subject-matter is also not suggested by D1, D2 and D3.

14. For the above reasons the board concludes that the decision under appeal has to be set aside and that a patent shall be granted on the basis of the appellant's third auxiliary request.

## Order

### For these reasons it is decided that:

1. The decision under appeal is set aside.
  
2. The case is remitted to the first instance with the order to grant a patent in the following version:  
Description:  
Pages 1, 3, and 4 filed in the oral proceedings  
Page 2 filed with the letter of 9 July 2004  
Pages 5 to 14 as originally filed  
Claims:  
No. 1 to 14 filed in the oral proceedings  
Drawings:  
Sheets 1 to 9 as originally filed.

The Registrar:

The Chairman:

L. Fernández Gómez

F. Edlinger