

METHOD AND SYSTEM FOR VERIFYING A TARGET IN AN IMAGE

FIELD OF THE INVENTION

[001] The present invention relates to the processing of image data. More
5 particularly, the present invention relates to a method and a system for verifying
a target in an image.

BACKGROUND OF THE INVENTION

[002] In high risk operations conducted by the police, military, fire brigade or
10 special forces worldwide, efficient communication between team members in
these operations is critical to the success of the operation. Typically, high risk
operations are time critical since human lives may be at stake. The need for
haste and accuracy can be appreciated, for example, by considering a scenario
where people who are trapped in a building on fire need to be located and
15 identified for rescue. Another scenario would be a case where snipers are
positioned around a predetermined location and are required to locate, within a
building, certain criminals who are holding hostages.

[003] With the advent of technology, the use of portable communication
20 devices such as wireless earpieces or other traditional means such as
handheld portable walkie-talkies for communication between team members in
such high risk operations are common. However, such critical information,
particularly if it relates to the identification of a specified target or object in an
area of interest, when communicated verbally, may be subject to errors, leading
25 to miscommunication and a failed mission.

[004] To circumvent the problem of miscommunication via verbal
communication, the sending of relevant images to team members through
image capturing means may be used. However, there may be a considerable

delay between the time when the scene is captured and the time when the information in the captured scene is relayed to other persons. Further, the captured scene or the target may not prove to be immediately evident in the image.

5

[005] There is, therefore, a need for a method and system for verifying a target in an image that can be communicated to others.

[006] Any discussion of documents, devices, acts or knowledge in this
10 specification is included to explain the context of the invention. It should not be taken as an admission that any of the material forms a part of the state of the art or the common general knowledge in the relevant art anywhere on or before the priority date of the disclosure and claims herein. All statements as to the date or representation as to the contents of these documents is based on the
15 information available to Applicants and does not constitute any admission as to the correctness of the date or contents of these documents.

OBJECTS OF THE INVENTION

[007] It is an object of the invention to provide a method to obtain a target
20 identified in a first image to be displayed on a second image.

[008] It is also an object of the invention to overcome, or at least substantially ameliorate, the disadvantages and shortcomings of the prior art.

25 [009] Other objects and advantages of the present invention will become apparent from the following description, taken in connection with the accompanying drawings, wherein, by way of illustration and example, an embodiment of the present invention is disclosed.

SUMMARY OF THE INVENTION

[010] According to the present invention, there is provided a system and method for verifying a target in an image. A first image input device receives a first image of a region of interest, where the first image input device is capable
5 of marking a target in the region of interest and sending the first image with the target marking, and a second image input device generates a second image of the region of interest. A processor processes the first and second images so that the target marked on the first image is identified on the second image. The respective fields of view are different.

10

[011] In one embodiment of the invention a system for identifying a target on an image comprises:

a first image input device for receiving a first image of a region of interest defining a first field of view, said first image input device being capable of
15 marking a target in the region of interest and sending the first image with the target marking;

a second image input device for generating a second image of the region of interest, said second image defining a second field of view; and

a processor for processing the first and second images so that the target
20 marked on the first image is identified on the second image,

wherein the first and second fields of view are different.

[012] In another embodiment of a system of the invention, the processor comprises:

25 an image processor for extracting a plurality of features of the region of interest from each of the first and second images and

an image matching processor for comparing and matching the plurality of features.

[013] In another embodiment of a system of the invention, the image processor extracts a first plurality of features of the first image and a second plurality of features of the second image.

5 [014] In another embodiment of a system of the invention, the image matching processor determines a plurality of pairs of features and wherein each pair of features includes a first feature, from the first plurality of features, and a second feature, from the second plurality of features, that matches the first feature.

10 [015] In another embodiment of a system of the invention, the image matching processor includes a matrix builder for determining a transformation matrix based on the plurality of pairs of features, whereby a first plane of the first image and a second plane of the second image are determined.

15 [016] In another embodiment of a system of the invention, the position coordinates of the marked target on the first plane of the first image are transformed to the second plane of the second image.

20 [017] In another embodiment of a system of the invention, the position coordinates of the marked target of the first image are multiplied by the transformation matrix to identify the target on the second image.

[018] In another embodiment of a system of the invention, the image processor uses a Speeded Up Robust Features algorithm.

25

[019] In another embodiment of a system of the invention, the image matching processor uses a Random Sample Consensus algorithm.

[020] In another embodiment of a system of the invention, the surface area of the target within the first image is not more than about 10% of the surface area of the first image.

- 5 [021] In another embodiment of a system of the invention, the first field of view of the region of interest defines a first viewing angle and the second field of view of the region of interest defines a second viewing angle, whereby the first and second viewing angles are spaced apart by less than about 120 degrees from each other.

10

[022] In another embodiment of a system of the invention, the first image is captured by a first mobile communication device having an image capturing means.

- 15 [023] In another embodiment of a system of the invention, the second image is captured by a second mobile communication device having an image capturing means.

- 20 [024] In another embodiment of a system of the invention, a system for comparing a first image and a second image of a region of interest where a target located in the first image will be identified on the second image, comprises:

an image input unit for receiving a first image of the region of interest defining a first field of view, said first image having the target in the region of
25 interest marked;

a detector for generating a second image of the region of interest, said second image having a second field of view;

an image processor for extracting a plurality of features of the first and second images; and

an image matching processor for matching the plurality of features of the first and second images,

wherein the first and second field of views are not identical.

- 5 [025] In another embodiment of the invention, a method for comparing a first image and a second image wherein a target located in the first image will be identified on the second image, comprises:

receiving a first image of a region of interest defining a first field of view, said first image having the target in the region of interest marked by a user;

- 10 generating a second image of the region of interest defining a second field of view; and

processing the first and second images such that the marked target in the first image is identified on the second image,

wherein the first and second field of views are not identical.

15

[026] In another embodiment of a method according to the invention, the surface area of the target within the region of interest of the first image is not more than about 10 per cent of the surface area of the first image.

- 20 [027] In another embodiment of a method of the invention, the first and second fields of view have a first angle of view and a second angle of view, respectively, with respect to the region of interest, whereby the first angle of view and the second angle of view are spaced apart from each other by less than about 120 degrees.

25

[028] In another embodiment of a method of the invention, processing the first and second images comprises:

(a) extracting a first plurality of features of the first image and a second plurality of features of the second image;

(b) determining a plurality of pairs of features, wherein each pair of features includes a first feature, from the first plurality of features, and a second
5 feature, from the second plurality of features, that matches the first feature;

(c) determining a transformation matrix using the plurality of pairs of features; and

(d) identifying the marked target in the second image based on the transformation matrix,

10 wherein the marked target in the second image is identified by multiplying the position coordinates of the marked target in the first image with the transformation matrix.

[029] In another embodiment of a method of the invention, steps (a) and (b)
15 comprise using a Speeded Up Robust Features algorithm.

[030] In another embodiment of a method of the invention, step (c) comprises using a Random Sample Consensus algorithm.

20 [031] In another embodiment of a method of the invention, the first image is captured by a first mobile communication device having an image capturing means.

[032] In another embodiment of a method of the invention, the second image
25 is captured by a second mobile communication device having an image capturing means.

[033] In another embodiment of a method of the invention, the target in the first image is marked by the user using touch screen technology.

[034] This invention may also be said broadly to consist in the parts, elements and features referred to or indicated in the specification of the application, individually or collectively, and any or all combinations of any two or more of said parts, elements or features, and where specific integers are mentioned herein which have known equivalents in the art to which this invention relates, such known equivalents are deemed to be incorporated herein as if individually set forth.

BRIEF DESCRIPTION OF THE DRAWINGS

[035] In order that the invention may be better understood and put into practical effect, reference will now be made to the accompanying drawings, in which:

[036] Figure 1 is a general block diagram of a system for receiving and processing the first and second images to verify a target in a second image according to an embodiment of the present invention;

20

[037] Figure 2 represents a process flow diagram of the method of verifying the marked target in the second image according to an embodiment of the present invention;

25 [038] Figure 3 illustrates the steps of the image matching processor for displaying the marked target in the second image according to an embodiment of the present invention;

[039] Figure 4 is a perspective view of an illustration of the method of identifying the target in the second image according to an embodiment of the present invention;

5 [040] Figure 5a is a view of a first image of the region of interest in a first field of view with a target marked according to an embodiment of the present invention;

[041] Figure 5b is a view of a second image of the region of interest in a
10 second field of view according to an embodiment of the present invention;

[042] Figures 6a and 6b are views of the first and second images, respectively, with the point interest detected using a SURF algorithm according to an embodiment of the present invention;

15

[043] Figures 7a and 7b are views of the first and second images, respectively, with the plurality of interest points matched according to an embodiment of the present invention;

20 [044] Figures 8a and 8b are views of the first and second images, respectively, with the plurality of matched features verified according to an embodiment of the present invention;

[045] Figures 9a and 9b are views of the first and second images,
25 respectively, showing the coordinates of the marked target in the first image appearing in the second image according to an embodiment of the present invention; and

[046] Figure 10 is a plan view of an illustration of two users capturing an image of a building according to an embodiment of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

5 [047] The present invention will now be described in detail in connection with preferred embodiments with reference to the accompanying drawings.

[048] Figure 1 is a block diagram illustrating a system 10 for receiving and processing a first image and a second image, whereby a target identified in the
10 first image can be displayed in the second image. The system 10 includes a detector 31, an image input unit 35, a processor 33, a memory 34, a display unit 36, a storage unit 32 and a communication interface 38. In a preferred embodiment of the present invention, system 10 may be embedded within a mobile communication device (not shown). Processor 33 includes an image
15 processor 37 for retrieving the first and second images and an image matching processor 39 for activating image matching algorithms, which will be explained in detail below.

[049] Display unit 36 can be any type of display for the output of information in
20 a human readable form, including, but not limited to, an LCD monitor, a mobile communication device display unit, a cathode ray tube monitor or the like. Processor 33 may be a printed circuit board or a processor executing a set of program instructions stored on a computer readable storage medium such as random access memory (RAM), read only memory (ROM), magnetic storage
25 medium or optical storage medium, etc.

[050] Detector 31 may comprise a device that is capable of capturing images, for example, a photographic device, such as a camera. Memory 34 can be any type of memory including random access memory (RAM), electronically

erasable memory (EPROM), or the like. Storage unit 32 can be any type of storage, including, but not limited to, magnetic or optical drives, local drive or network drive, or any storage for storing of information. A communication interface 38 may comprise any type of interface for connecting to a communication network, such as data or voice network, a land-line or wireless network, or the like. It is recognized that a person of ordinary skill in the art would understand the mode of operation of a communication interface and a detailed description of this would not be required for the scope of this invention.

10 [051] Figure 2 is an overview of the steps of the method of identifying a target in a second image from a marked target in a first image. With reference to Figures 4, 5a, and 5b, a mobile communication device 64 includes a detector 31 (not shown here) which captures a first image 72 of the region of interest from a first field of view (step 42). Mobile communication device 64 may include devices such as hand-held image recording devices, mobile telephones, palmtop computers, personal digital assistants (PDAs), and the like. A target 65 in first image 72 is marked by the user of mobile communication device 64 (step 44). Marking of the image may be enabled by touch sensor technology embedded in mobile communication devices which converts a touch on a screen or monitor to a marking on part of the image. The touch responsive surface of the touch screen has an electrical current or signal going through it. Once the screen is touched, the screen causes a voltage or a signal change which can be used to determine the location of the touch to the screen. For the purpose of this invention, it will be understood by a person of ordinary skill in the art that marking of an image on mobile communication devices may be made by current enabling technologies and will therefore not be described in detail.

30 [052] In a preferred form of the embodiment, the surface area of target 65 is preferably less than about 10 per cent of the size of the surface area of first image 72.

[053] First image 72 containing marked target 65 is transmitted via a communication interface 38 to a second mobile communication device 66 (step 46). Second mobile communication device 66 receives first image 72 via image input unit 35 (step 48) and stores it into memory 34. A second user who
5 receives the first image would be unable to verify marked target 65 from first image 72 as his or her field of view or viewing angle is not the same as that of the first user. Although the second user may be able to identify marked target 65 based on the first image, this may take a considerable amount of time. In time critical missions or situations, the second user may not be able to afford
10 the time to analyze the first image and identify the target quickly.

[054] In order to display the marked target 65 in the second user's field of view quickly, the second user will capture a second image 74 of the region of interest from his field of view (step 41). In a preferred embodiment of the
15 present invention, the approximate time to display marked target 65 in the second image upon capturing the second image is from about 0.5 to about 5 sec. Second image 74 will be compared with first image 72 stored in memory 34 of second mobile communication device 66 via image matching processor 39 (step 43). Image matching processor 39 will be activated, which will be
20 explained in further detail below.

[055] Image matching processor 39 will compare the first and second images (72 and 74) by extracting, matching, verifying, and transforming the features or interest points obtained from the first and second images to display the marked
25 target in second image 74. Second image 74 will then appear to the user with the target marked for identification (step 47).

[056] Figure 3 is a diagram of the steps involved in image matching processor 39 according to a preferred embodiment of the invention. Each of the steps will
30 be explained in detail with reference to the first and second images in Figures

5a to 9b. The present invention employs a Speeded Up Robust Features (SURF) algorithm to extract relevant features from the first and second images 72, 74. The SURF algorithm was first presented by Herbert Bay et al. in 2006, "SURF: Speeded Up Robust Features", *Proceedings of the Ninth European Conference on Computer Vision*, May 2006, incorporated herein by reference. The SURF algorithm is a robust image detector and descriptor which includes an interest point detection and feature descriptor scheme.

[057] Unlike many traditional corresponding feature detection methods, SURF features are computed based on the sum of approximated 2D Haar wavelet, using integral images. This allows the computation of rectangular box filters to be near constant time, therefore leading to much faster computing performance than many other methods without compromising performance.

[058] In order to find corresponding matching features between two similar images, a SURF (Speeded Up Robust Features) algorithm is used. It is a scale- and rotation-invariant interest point detector and descriptor.

[059] First, interest points (82 and 84) are detected at distinctive locations in the image, such as corners, blobs, and T-junctions. For the purpose of discussion, the terms "interest points" and "features" are used interchangeably. The detector (Fast-Hessian detector) is based on the basic approximation of the $H(x, \sigma)$ Hessian matrix and makes an efficient use of integral images. Given a point $x = (x, y)$ in an image I , the Hessian matrix in x at scale σ is defined as follows:

$$H(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix}$$

where $L_{xx}(x, \sigma)$ is the convolution of the Gaussian second order

derivative $\frac{\partial^2}{\partial x^2} g(\sigma)$ with the image I in point x , and similarly for $L_{xy}(x, \sigma)$ and $L_{yy}(x, \sigma)$.

5

[060] With reference to Figure 3, a feature extractor 52 extracts features 82 and 84 (or interest points) from first image 72 and second image 74 of the same region of interest, both taken from different fields of view or viewing angles. In an embodiment of the invention as shown in Figures 5a and 5b, first
 10 image 72 may be taken from a left side of the building (Figure 5a) and second image 74 may be taken from the right side of the same building (Figure 5b). Feature extractor 52 may perform more than one function, such as, for example, interest point detection and feature description. Interest point detection detects interest points in an image according to conditions preferably
 15 reproducible under image variations such as variations in brightness and perspective. The neighborhood of each interest point is described as a feature vector. These feature vectors are represented by feature descriptors and are preferably distinctive.

20 [061] Figures 6a and 6b show an example of the SURF algorithm detecting the interest points (82 and 84) using the "Fast-Hessian" detector. In a preferred embodiment of the present invention, the SURF feature algorithm may be used to extract a first plurality of features 82 (or interest points) of the first image (Figure 6a) and a second plurality of features 84 of the second image (Figure
 25 6b). Figure 6a shows examples of the types of features that may be detected from the image. For example, these may be the corners of the buildings, the corners of each of the windows, or the trees.

[062] Next, each interest point is represented by the distribution of Haar-wavelet responses within its neighborhood. Here, a square region centered
 30 around the interest point is firstly constructed. Then, the region is split up

regularly into smaller 4 x 4 square sub-regions. For each sub-region, a four-dimensional descriptor vector v is built as follows:

$$v = (\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|)$$

where d_x and d_y are wavelets responses.

- 5 This results in a descriptor vector for all 4 x 4 sub-regions of length 64. So each interest point is represented by a 64-dimensional vector $\mathbf{D}_{1 \times 64}$. For the marked target 65 in Figure 6a in accordance with an embodiment of this invention, its SURF feature vector $\mathbf{D}_{1 \times 64} = [0.091, 0.307, -0.404, -0.196, \dots, 0.523, -0.118, 0.335, -0.125]$.

10

[063] Once the first and second plurality of features (82 and 84) are extracted from the first and second images (72 and 74, respectively), these plurality of features are matched by a feature matcher 54. The plurality of features is represented by the descriptor vectors obtained from the SURF algorithm.

15

- [064] The descriptor vectors are matched between the two images. The matching is often based on a distance between the vectors, for example, the Mahalanobis or Euclidean distance. Here, the Nearest Neighbor Distance Ratio matching strategy is used. For an interest point 82 in the first image, its
20 corresponding interest point 84 in the second image is found if

$$\frac{\|\mathbf{D}_A - \mathbf{D}_B\|}{\|\mathbf{D}_A - \mathbf{D}_C\|} < \tau$$

- 25 where $\mathbf{D}_B$ is the first and $\mathbf{D}_C$ is the second nearest neighbor to $\mathbf{D}_A$, and τ is a pre-defined threshold.

[065] Figures 7a and 7b show the found plurality of pairs of matching features 86 with $\tau = 0.88$ in accordance with one embodiment of the invention.

[066] Once the pairs of matching features 86 from the first and second images are identified, the plurality of matched descriptor vectors are used to determine or compute a transformation matrix **M**. The transformation matrix **M** between the first image **I_S** and the second image **I_R**, is determined as follows:

$$\mathbf{I}_R = \mathbf{M} \times \mathbf{I}_S$$

[067] A transformation matrix **M** is determined for the purpose of projectively transforming a first plane of the first image into the second plane of the second image. Once a transformation matrix is determined (step 56), the marked target in the first image lying on the first plane will be transformed into the second plane of the second image (step 58). This allows the position of the marked target 65 in the first image 72 to be displayed as the same marked target on the second image 74 even if the viewing angles of the region of interest in the first and second images are different (step 57).

[068] In accordance with a preferred embodiment of the invention, the transformation between first image 72 and second image 74 is assumed as a 3 x 3 transformation matrix. To compute the transformation matrix **M_{3x3}**, a RANSAC algorithm is used in a preferred embodiment of this invention. Inliers, i.e., the plurality of pairs of matched features, will be retained. Outliers, the plurality of pairs of features which are unable to be matched, are then recognized and removed.

[069] The RANdom SAmple Consensus (RANSAC) algorithm is employed to verify the plurality of matched features and is an estimating algorithm that is able to work with large sets of putative correspondences. The best way to determine the transformation matrix **M_{3x3}** is to compute **M** for all possible combinations, verify every solution, and then use the correspondence with the best verification. The verification procedures may be done in different ways. The RANSAC algorithm is a hypothesis and verifies algorithm, but works in a

different way. Instead of systematically working through itself through the possible plurality of matched features, it selects its correspondence points randomly and then computes the transformation matrix and performs the verifications. The RANSAC algorithm repeats this procedure for a certain amount of times and decides to use the correspondence set with the best verification. Other such algorithms that may be used are least mean squares (LMS) and least-median-of-squares (LMedS) methods.

[070] Figures 8a and 8b represent the manner in which the transformation matrix **M** is computed in a preferred embodiment of the invention. Having determined the plurality of features from the SURF algorithm, the RANSAC algorithm will determine inliers, i.e., the plurality of pairs of matched features 88, as indicated in dotted lines in Figures 8a and 8b. The outliers, i.e., the plurality of pairs of features which either do not match or are unable to be matched, will be recognized and removed. Using Figures 8a and 8b as a preferred embodiment of the present invention, the transformation matrix **M** was determined to be:

$$\mathbf{M}_{3 \times 3} = \begin{bmatrix} 1.003 & -0.007 & -1.781 \\ 0.015 & 0.994 & 3.633 \\ 0 & 0 & 1 \end{bmatrix}$$

[071] Finally, once the transformation matrix $\mathbf{M}_{3 \times 3}$ has been determined, the marked coordinate **P** (x, y) (represented by the star) on the first image **I_S**, its corresponding coordinates on **P** (x', y') on the second image, **I_R** may be obtained by the following formula:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \mathbf{M}_{3 \times 3} \times \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

[072] Figure 4 represents a situation where the method and system according to a preferred embodiment of the present invention may be employed. A

building 62 on fire is shown where some of the occupants have been trapped within the building. A rescue operation team or firemen may be at the scene. Typically, it is not unusual for firemen to be dispersed away from each other as they each have different roles in emergency situations. A fireman who spots
5 trapped occupants on a certain floor within a certain unit will wish to communicate this information to other teams. To prevent miscommunication and the need for speed in this situation, he may capture an image of the building with his mobile communication device 64 and mark the target, i.e. the window of the building where the trapped occupants are, on the building and
10 send it to colleagues. A colleague receives the image, captures a second image 74 of the building from where he or she is located with a mobile communication device 66 enabled with the system of the present invention, and an image processor 37 sends the second image 74 with the marked target 65 displayed on his or her mobile communication device 66.

15

[073] Figure 5a is an example of the first image 72 captured by the detector showing a first field of view of the region of interest. Figure 5b is an example of the second image 74 captured by a second detector showing a second field of view of the region of interest. Figure 10 represents a plan view of the locations
20 of the first detector and the second detector where the first and second images, respectively, are captured. Preferably, the first and second viewing angle A should not exceed 120 degrees.

[074] There may be instances where a user, upon receiving the first image in
25 his or her mobile communication device, does not know to which building the sender of the first image is referring. This could be due to more than one building being on fire in the preferred embodiment of this invention. In such an instance, the system preferably is equipped with a global positioning system (GPS) (not shown) to estimate the location of the building in the first image. It
30 may also include a magnetometer or a range finder (not shown) to estimate the direction at which the detector is pointing.

[075] It will be recognized that the terms “field of view” and “viewing angle” are used interchangeably in the context of the description of this invention. Field of view in the context of this invention is defined as the view of the region of interest as captured in an image with respect to the detector or image capturing means.

[076] The present invention has been described above with the aid of functional building blocks illustrating the implementation of specific functions and relationships thereof. The boundaries of these functional building blocks have been arbitrarily defined herein for the convenience of the detailed description. Alternative boundaries can be defined so long as the specific functions and relationships thereof are appropriately performed.

[077] Although the invention has been herein shown and described in what is conceived to be the most practical and preferred embodiment, it is recognized that departures can be made within the scope of the invention, which is not to be limited to the details described herein but is to be accorded the full scope of the appended claims so as to embrace any and all equivalent devices and apparatus.

[078] “Comprises/comprising” when used in this specification is taken to specify the presence of stated features, integers, steps or components but does not preclude the presence or addition of one or more other features, integers, steps, components or groups thereof.

CLAIMS:

1. A system for identifying a target on an image, which comprises:
 - a first image input device for receiving a first image of a region of interest
 - 5 defining a first field of view, said first image input device being capable of marking a target in the region of interest and sending the first image with the target marking;
 - ... a second image input device for generating a second image of the region of interest, said second image defining a second field of view; and
 - 10 a processor for processing the first and second images so that the target marked on the first image is identified on the second image,
 - wherein the first and second fields of view are different.
2. The system according to claim 1, wherein the processor comprises:
 - 15 an image processor for extracting a plurality of features of the region of interest from each of the first and second images and
 - an image matching processor for comparing and matching the plurality of features.
- 20 3. The system according to claim 2, wherein the image processor extracts a first plurality of features of the first image and a second plurality of features of the second image.
4. The system according to claim 3, wherein the image matching processor
 - 25 determines a plurality of pairs of features and wherein each pair of features includes a first feature, from the first plurality of features, and a second feature, from the second plurality of features, that matches the first feature.

5. The system according to claim 4, wherein the image matching processor includes a matrix builder for determining a transformation matrix based on the plurality of pairs of features, whereby a first plane of the first image and a second plane of the second image are determined.
- 5
6. The system according to claim 5, wherein position coordinates of the marked target on the first plane of the first image are transformed to the second plane of the second image.
- 10 7. The system according to claim 6, wherein the position coordinates of the marked target of the first image are multiplied by the transformation matrix to identify the target on the second image.
- 15 8. The system according to claim 3, wherein the image processor uses a Speeded Up Robust Features algorithm.
9. The system according to claim 4, wherein the image matching processor uses a Random Sample Consensus algorithm.
- 20 10. The system according to claim 1, wherein the surface area of the target within the first image is not more than about 10% of the surface area of the first image.
- 25 11. The system according to claim 1, wherein the first field of view of the region of interest defines a first viewing angle and the second field of view of the region of interest defines a second viewing angle, whereby the first and second viewing angles are spaced apart by less than about 120 degrees from each other.

12. The system according to claim 1, wherein the first image is captured by a first mobile communication device having an image capturing means.

13. The system according to claim 12, wherein the second image is
5 captured by a second mobile communication device having an image capturing means.

14. A system for comparing a first image and a second image of a region of interest where a target located in the first image will be identified on the second
10 image, comprising:

an image input unit for receiving a first image of the region of interest defining a first field of view, said first image having the target in the region of interest marked;

a detector for generating a second image of the region of interest, said
15 second image having a second field of view;

an image processor for extracting a plurality of features of the first and second images; and

an image matching processor for matching the plurality of features of the first and second images,

20 wherein the first and second field of views are not identical.

15. A method for comparing a first image and a second image wherein a target located in the first image will be identified on the second image, comprising:

25 receiving a first image of a region of interest defining a first field of view, said first image having the target in the region of interest marked by a user;

generating a second image of the region of interest defining a second field of view; and

processing the first and second images such that the marked target in the first image is identified on the second image,

wherein the first and second field of views are not identical.

5 16. The method according to claim 15, wherein the surface area of the target within the region of interest of the first image is not more than about 10 per cent of the surface area of the first image.

10 17. The method according to claim 16, wherein the first and second fields of view have a first angle of view and a second angle of view, respectively, with respect to the region of interest, whereby the first angle of view and the second angle of view are spaced apart from each other by less than about 120 degrees.

15 18. The method according to claim 15, wherein processing the first and second images comprises:

(a) extracting a first plurality of features of the first image and a second plurality of features of the second image;

20 (b) determining a plurality of pairs of features, wherein each pair of features includes a first feature, from the first plurality of features, and a second feature, from the second plurality of features, that matches the first feature;

(c) determining a transformation matrix using the plurality of pairs of features; and

25 (d) identifying the marked target in the second image based on the transformation matrix,

wherein the marked target in the second image is identified by multiplying the position coordinates of the marked target in the first image with the transformation matrix.

19. The method according to claim 18, wherein steps (a) and (b) comprise using a Speeded Up Robust Features algorithm.
20. The method according to claim 18, wherein step (c) comprises using a
5 Random Sample Consensus algorithm.
21. The method according to claim 15, wherein the first image is captured by a first mobile communication device having an image capturing means.
- 10 22. The method according to claim 21, wherein the second image is captured by a second mobile communication device having an image capturing means.
23. The method according to claim 21, wherein the target in the first image
15 is marked by the user using touch screen technology.