

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
21 February 2008 (21.02.2008)

PCT

(10) International Publication Number
WO 2008/019846 A2

(51) International Patent Classification:
G06T 5/00 (2006.01)

B-1150 Woluwe St. Pierre (BE). **OSKARSSON, Magnus**
[SE/SE]; Lönngatan 52F, S-214 49 Malmö (SE).

(21) International Application Number:
PCT/EP2007/007229

(74) Agents: **BIRD, William, E.** et al.; Bird Goën & Co, Klein
Dalenstraat 42A, B-3020 Winksele (BE).

(22) International Filing Date: 16 August 2007 (16.08.2007)

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

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
06119041.9 16 August 2006 (16.08.2006) EP

(71) Applicants (for all designated States except US): **TOY-
OTA MOTOR EUROPE NV** [BE/BE]; Bourgetlaan 60,
B-1140 Brussel (BE). **LUND UNIVERSITY**; Universitet-
shuset (Chancellery), Paradisgaten 2, SE-22350 Lund (SE).

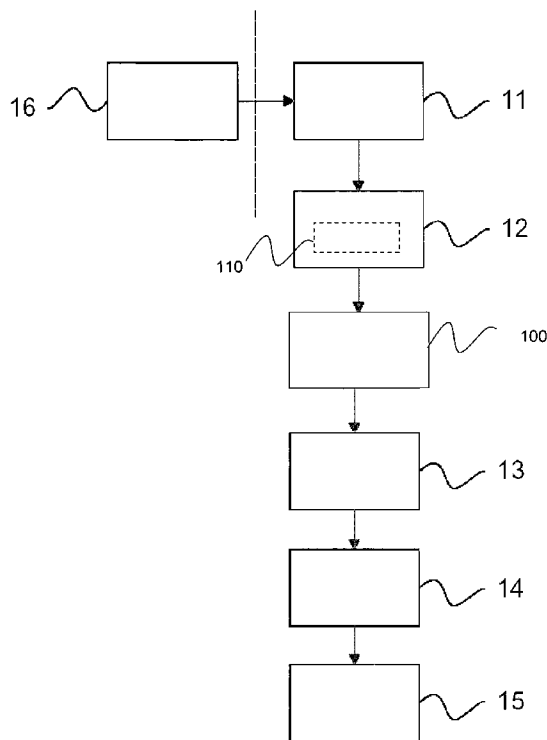
(72) Inventors; and

(75) Inventors/Applicants (for US only): **MALM, Henrik**
[SE/SE]; Slattervägen 7B, S-227 38 Lund (SE). **WAR-
RANT, Eric** [SE/SE]; Stora Fiskaregatan 15a, S-222 24
Lund (SE). **AMBECK-MADSEN, Jonas** [BE/BE]; Rue
Charles Martel 35, B-1000 Brussels (BE). **YANAGI-
HARA, Hiromichi** [JP/JP]; Avenue de l'Atlantique 57,

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

[Continued on next page]

(54) Title: A METHOD, AN APPARATUS AND A COMPUTER-READABLE MEDIUM FOR PROCESSING A NIGHT VI-
SION IMAGE DATASET



(57) Abstract: A method is disclosed for enhancing the quality of an image dataset, such as, e.g., reducing the noise in a noisy image data set and increasing the contrast in the image data set. The method may be used for processing a sequence of image datasets, e.g. night vision image datasets, wherein said sequence comprises at least two image datasets each having at least two pixels, and wherein each pixel has an intensity value. The method comprises calculating a structure tensor for each pixel in an image dataset comprised in the sequence of image datasets; calculating values in a summation kernel based on said structure tensor for each pixel in said image dataset; calculating a weighted intensity value for each pixel in said first image dataset, using as weights the values in said summation kernel; storing said weighted intensity value for each pixel in said image dataset as a processed intensity value for each corresponding pixel in a processed output image dataset; rotating a local coordinate system in which the summation kernel is described resulting in that the coordinate axes of said local coordinate system coincide with the directions of the eigenvectors of said structure tensor, where said eigenvectors are described in the global coordinate system of the image dataset, and scaling the coordinate axes of the local coordinate system in which the summation kernel is described by an amount related to the eigenvalues of the structure tensor via a width function $W(\lambda_i) = \sigma_i$, and wherein said eigenvalues depend on the amount of intensity variation in the direction of their corresponding eigenvectors, the width function being a decreasing function such that $w(0) = \sigma_{max}$ and $\lim_{a \rightarrow \infty} w = \sigma_{min}$. An apparatus and a computer readable medium are also provided.

WO 2008/019846 A2



Declaration under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

Published:

— *without international search report and to be republished upon receipt of that report*

**A METHOD, AN APPARATUS AND A COMPUTER-READABLE MEDIUM FOR
PROCESSING A NIGHT VISION IMAGE DATASET.**

Technical Field

5 The present invention generally relates to the field
of image processing. More particularly the present
invention relates to methods, apparatuses, computer
programs and computer-readable mediums for enhancing the
dynamic level, such as, e.g., reducing noise in a noisy
10 image dataset, particularly, but not limited to image
datasets with areas of low contrast such as image datasets
or image dataset sequences obtained in low light level
conditions, e.g. night vision datasets, or in other
conditions of relatively poor visibility.

15

Description of Related Art

In digital image processing, noise more or less
corresponds to visible grain present in an image dataset.
The term noise generally refers to the high frequency
20 random perturbations that are generally caused by
electronic noise in the input device sensor and circuitry
(e.g. scanner, digital camera). In dim light conditions,
e.g. during night vision, another source for noise is the
quantum nature of light. Since very few photons reach the
25 camera film or CCD detector the uncertainty in the true
intensity value is in general large. That is, the variance
of the intensity is large compared to the true value and
that leads to a grainy noisy image. The combination of this
large variance and the added internal electronic noise in
30 the imaging device may lead to disturbing levels of noise.
Noise reduction is a part of image analysis that strives to
reduce noise in an image dataset.

Many animals, especially those with a nocturnal

lifestyle, have excellent visual capabilities in dim light conditions and possess means of dealing with the inevitable noise. A feature of most animal visual systems, including the human, is the ability to sum the visual signal locally
5 in space (spatial summation) and time (temporal summation) to increase signal strength in each input channel and to improve the reliability of the intensity estimations. However, this comes at the cost of lower spatial and/or temporal resolution. For animals having to react to high-
10 speed events and/or that are moving with a high relative velocity, high temporal resolution is generally needed, and spatial summation is favored before temporal summation. For other animals that are mostly active in more static environments, temporal summation is instead favored in
15 order to obtain as high a spatial resolution as possible.

In the publication "Seeing better at night: life style, eye design and the optimum strategy of spatial and temporal summation", by E. Warrant, Vision Research, 39:1611-1630, 1999, a model is provided that attempts to
20 resemble visual signal processing in animal eyes that takes spatial and temporal summation into consideration. This model tries to find the optimal balance of spatial and temporal summation that maximizes visual performance. By studying the visual systems of a number of specific
25 animals, it is shown that the relative ratio of spatial and temporal summation of these animals fits well with their lifestyle and needs when it comes to temporal and spatial resolution. A drawback of the model is that it is restrictive when it comes to how the local neighborhood,
30 over which the summation is made, may be chosen. The neighborhood is only allowed to be adjusted along the axes in the spatiotemporal space, i.e. the neighborhood is symmetric around these axes. When dealing with a pixel in a

moving area of the visual field, only a small amount of temporal summation, following this model, will in general lead to motion blur. To reduce the high frequency noise, the model then instead dictates the use of a large amount of spatial blurring, which will generally lead to an unnecessary low spatial resolution. The model also assumes that a separate velocity estimation has been performed before it can be decided which amounts of spatial and temporal summation should be applied in each pixel.

When it comes to digital image processing, various attempts have been made for trying to estimate the motion trajectories corresponding to the moving areas of an image dataset through the spatiotemporal space and to reduce noise by filtering the sequence along these trajectories. Moving areas in an image dataset may occur e.g. by moving the camera while obtaining the image dataset and/or by obtaining an image dataset of one or more moving objects within the image dataset, using a stationary or moving camera. These noise reduction techniques are usually referred to as motion-compensated spatiotemporal filtering. In this way the motion blur is, ideally, avoided and the need for spatial filtering reduced. In some of these techniques weighted averages, depending on estimates of the noise level, are calculated in a small neighborhood along and around the motion trajectories. Other filtering techniques, such as so-called linear minimum mean square error filtering and Kalman and/or Wiener filtering, have been proposed to be applied along the motion trajectories. A drawback of the motion-compensating methods is that they rely rather heavily on a good estimation of the motion trajectory in every point being calculated separately before the filtering is applied. A slight error in the motion estimation will thus generally lead to excessive

blurring. The explicit estimation of the motion trajectories is especially complicated for sequences severely degraded by noise. Generally, the filtering is simply reduced in areas where the motion estimation is especially difficult, e.g. at object contours, generally leading to poor noise reduction in those areas.

Other noise reducing video processing methods use a cascade of directional filters where the filtering and smoothing is done primarily in the direction which corresponds to the filter which gives the highest output. These methods work well for directions that coincide with the fixed filter directions, but in general have a pronounced degradation in the output for directions in between the filter directions.

The publication "Nonlinear image filtering with edge and corner enhancement" by Nitzberg and Shiota, IEEE Trans. On Pattern Recognition and Machine Intelligence, 14:826-833,1992, discloses a method of nonlinear image filtering for noise reduction in a single 2D image dataset. According to this method, Gaussian shaped kernels are calculated that adapt to the local spatial intensity structure at each pixel. These kernels are used to calculate a weighted mean value in the spatial neighborhood around each pixel. A drawback of the method disclosed in this publication is that the model only takes into account spatial summation for individual 2D image datasets, and, hence, does not address spatiotemporal summation for image datasets, i.e. noise reduction in image dataset sequences, obtained e.g. during dim light conditions. Another drawback is that the function, which relates the spatial intensity structure to the shape of the Gaussian kernel, is fixed. For example, in the case of image datasets including both very dark areas with low intensity contrast and much brighter areas, the

intensity gradients that should influence the Gaussian kernels vary very much in magnitude between areas. The method presented in the above-mentioned publication has no means for dealing with this variation and will in general provide unnecessarily poor output resolution either in the dark or in the bright areas. Hence, an improved method for e.g. noise reduction in an image dataset comprising low contrast areas would be advantageous, allowing for increased image resolution, flexibility, and clarity.

10 G. Yang et al., in "Structure adaptive anisotropic image filtering", *Image and Vision Computing*, 14:135-145, 1996, describes structure sensitive anisotropic image filtering. The method computes a new image $f_{out}(x)$, by applying at each spatiotemporal point $x_0 = (x'_0, y'_0, t'_0)$, a kernel $k(x_0, x)$ to the original image $f_{in}(x)$ such that, for example for a 2D image:

$$f_{out}(x_0) = \frac{1}{\mu(x_0)} \int \int \int_{\Omega} k(x_0, x) f_{in}(x) dx$$

where

$$\mu(x_0) = \int \int \int_{\Omega} k(x_0, x) dx$$

20 is a normalizing factor. The normalization makes the sum of the kernel elements equal to 1 in all cases, so that the mean intensity in the image does not change. The area Ω over which the integration, or in the discrete case summation, is made is chosen as a finite neighborhood centered around x_0 . $x = (x', y', \dots)$ is an m-dimensional position vector, for example $x = (x', y')$ is a 2D position vector.

When function k is invariant over space, it uniformly smoothes out image noise without paying any attention to image details. The basic idea behind Yang's algorithm is

30

that the kernel $k(x_0, x)$ used should be made variable and allow to be shaped or scaled according to local features within the neighborhood Ω of x_0 . Since it is desired to adapt the filtering to the spatiotemporal intensity
5 structure at each point, in order to reduce blurring over spatial and temporal edges, a kernel $k(x_0, x)$ is calculated individually for each point x_0 . The kernels should be wide in directions of homogeneous intensity (weak gradient), so that uniform regions can be smoothed out, and narrow in
10 directions with important structural edges (strong gradient), so that edges and corners are preserved.

To find these directions, the intensity structure is analyzed by the so-called *structure tensor* or *second moment matrix*. This object has been developed and applied in image
15 analysis in numerous papers and books, see for example B. Jähne, "Spatio-temporal image processing", Springer, 1993.

The solution of Yang is said to work very well for images with low signal to noise ratios. However, Yang does not mention night vision images.

20

Summary of the Invention

Surprisingly, there is still a need for improved image processing for low contrast images, in particular night vision images. In particular for recognition
25 purposes, such as pedestrian detection, the output image quality preferably has to be improved.

It is an object of embodiments of the present invention to provide a method, apparatus and software for processing a sequence of night vision image datasets. In
30 particular, it is an advantage of embodiments of the present invention to provide a method for simultaneous noise reduction and sharpening of images in a night vision system.

Accordingly, the present invention preferably seeks to mitigate, alleviate or eliminate one or more of the above-identified deficiencies in the art and disadvantages singly or in any combination.

5 According to an aspect of the invention, a method for processing a sequence of image datasets, e.g. night vision image datasets, is provided. The sequence of image datasets may be provided for dynamic enhancement. The sequence comprises at least two image datasets each having at least
10 two pixels, wherein each pixel has an intensity value. The sequence of image datasets may e.g. comprise 2D, 3D, distance image or higher-dimensional monochrome or color image datasets.

The method comprises calculating, based on estimated
15 intensity derivatives in each pixel, a measure of an intensity structure, e.g. a structure tensor, for each pixel in an image dataset comprised in said sequence of image datasets and performing an eigen-decomposition of said measure of the intensity structure, e.g. structure
20 tensor; calculating values in a summation kernel based on said measure of the intensity structure, e.g. tensor structure, for each pixel in said image dataset; calculating a weighted intensity value for each pixel in said image dataset, using as weights the values in said
25 summation kernel; storing said weighted intensity value for each pixel in said image dataset as a processed intensity value for each corresponding pixel in a processed output image dataset; rotating a local coordinate system in which the summation kernel is described resulting in that the
30 coordinate axes of said local coordinate system coincide with the directions of the eigenvectors of said measure of the intensity structure, e.g. structure tensor, where said eigenvectors are described in the global coordinate system

of the image dataset, and scaling the coordinate axes of the local coordinate system in which the summation kernel is described by an amount related to the eigenvalues of the measure of the intensity structure, e.g. structure tensor, via a width function $w(\lambda_i) = \sigma_i$, and wherein said eigenvalues depend on the amount of intensity variation in the direction of their corresponding eigenvectors, the width function being a decreasing function depending on the noise level in each pixel such that $w(0) = \sigma_{\max}$ and $\lim_{a \rightarrow \infty} w(a) = \sigma_{\min}$.

In particular embodiments of the present invention, the width function may be defined by

$$w(\lambda_i, x_0) = \begin{cases} \Delta\sigma e^{-\frac{\lambda_i}{d(x_0)} + \frac{1}{2}} + \sigma_{\min}, & \lambda_i > \frac{d(x_0)}{2} \\ \sigma_{\max}, & \lambda_i \leq \frac{d(x_0)}{2} \end{cases} .$$

The width function may be adapted to a noise level of the image datasets. Adapting the width function to the noise level may comprise estimating the noise in the image datasets and scaling an estimation of the variance of the noise so that the result is of a magnitude that is suitable for insertion in the width function. As a less preferred option the noise level may be a signal-to-noise level.

Scaling an estimation of the variance may be performed by means of a fixed scaling factor. Alternatively, scaling an estimation of the variance may be performed by means of a scaling factor which is determined from the image dataset.

The method according to embodiments of the present invention may comprise repeating the above method steps for subsequent image datasets.

The step of calculating a measure of the intensity structure may comprise calculating a structure tensor based on estimated intensity derivatives in each pixel.

The step of calculating a measure of the intensity

structure may comprise calculating a structure tensor based on the mean value of the corresponding estimated intensity derivatives for each channel in each pixel when said image dataset is a multi-channel image dataset.

5 The method may further comprise calculating said structure tensor for each channel in said image dataset when said image dataset is a multi-channel image dataset.

 The method may further comprise calculating the elementwise mean structure tensor of said structure tensors
10 calculated for each channel in said image dataset when said image dataset is a multi-channel image dataset.

 The method may further comprise calculating said structure tensor using the mean value of the values in different channels or in each channel in said dataset when
15 said image dataset is a multi-channel image data set.

 The method may further comprise performing low-pass filtering of said image dataset before or during said step of calculating a measure of the intensity structure, e.g. the structure tensor.

20 The step of calculating of a measure of the intensity structure, e.g. the structure tensor, may further comprise calculating motion vectors for each pixel of each image dataset of said sequence of image datasets.

 The method may further comprise performing an
25 intensity transformation on said image dataset by an intensity transformation function to increase the intensity contrast in at least a part of said first image dataset.

 The method may further comprise performing a foreground-background segmentation.

30 The method may further comprise rendering each of said first output image dataset for presentation on a display.

 The method according to embodiments of the present

invention may furthermore comprise sharpening edges of contours in the image. The sharpening may be obtained by convolution of the summation kernel with a high-boost filter giving the summation kernel \bar{S}_H . Alternatively, the sharpening may be obtained by using a summation function S_h obtained by approximating a discrete high-boost filter with a continuous function h and calculating a new summation function analytically by the convolution of the original summation kernel S with the function h , and discretizing the summation function S_h , giving the summation kernel \bar{S}_h .

According to another aspect of the invention, a computer program product comprises program instructions for causing a computer system to perform the method according to the embodiments of the present invention when said program instructions are run on a computer system. The computer program product may be embodied on a record medium, stored in a computer memory, embodied in a read-only memory, or carried on an electrical carrier signal.

According to yet another aspect of the invention, a computer-readable medium having embodied thereon a computer program for processing by a processor of a sequence of image datasets is provided. The sequence comprises at least two image datasets, each image dataset having at least two pixels, wherein each pixel has an intensity value. The computer program comprises a first calculation code segment for calculating, based on estimated intensity derivatives in each pixel, a measure of an intensity structure, e.g. a structure tensor, for each pixel in an image dataset comprised in said sequence of image datasets, and for performing an eigen-decomposition of said measure of the intensity structure, e.g. structure tensor; a second calculation code segment for calculating values in a summation kernel based on said measure of the intensity

structure, e.g. structure tensor, for each pixel in said image dataset; a third calculation code segment for calculating a weighted intensity value for each pixel in said image dataset, using as weights the values in said summation kernel; a storage code segment for storing said weighted intensity value for each pixel in said image dataset as a processed intensity value for each corresponding pixel in a processed output image dataset; a further calculation code segment for rotating a local coordinate system in which the summation kernel is described resulting in that the coordinate axes of said local coordinate system coincide with the directions of the eigenvectors of said measure of the intensity structure, e.g. structure tensor, where said eigenvectors are described in the global coordinate system of the image dataset; and a scaling code segment for scaling the coordinate axes of the local coordinate system in which the summation kernel is described by an amount related to the eigenvalues of the measure of the intensity structure, e.g. structure tensor, via a width function $w(\lambda_i) = \sigma_i$, and wherein said eigenvalues depend on the amount of intensity variation in the direction of their corresponding eigenvectors, the width function being a decreasing function depending on the noise level in each pixel such that $w(0) = \sigma_{max}$ and $\lim_{a \rightarrow \infty} w(a) = \sigma_{min}$. The calculation code segment for rotating and the scaling code segment may be part of the second calculation code segment.

The width function may be a decreasing function defined by

$$w(\lambda_i, x_0) = \begin{cases} \Delta\sigma e^{-\frac{\lambda_i}{d(x_0)} + \frac{1}{2}} + \sigma_{min}, & \lambda_i > \frac{d(x_0)}{2} \\ \sigma_{max}, & \lambda_i \leq \frac{d(x_0)}{2} \end{cases}$$

The computer program may further comprise a repeating code segment for repeating the above steps for subsequent image datasets.

According to another aspect of the invention, an apparatus comprising units configured to perform the method according to the embodiments of the invention is provided.

According to a further aspect of the invention, an apparatus for processing a sequence of image datasets is provided. The sequence comprises at least two image datasets each having at least two pixels, wherein each pixel has an intensity value.

The apparatus comprises a first calculation unit for calculating, based on estimated intensity derivatives in each pixel, a measure of an intensity structure, e.g. a structure tensor, for each pixel in an image dataset comprised in said sequence of image datasets and performing an eigen-decomposition of said measure of the intensity structure, e.g. structure tensor; a second calculation unit for calculating values in a summation kernel based on said measure of the intensity structure, e.g. structure tensor, for each pixel in said image dataset; a third calculation unit for calculating a weighted intensity value for each pixel in said image dataset, using as weights the values in said summation kernel; and a storage unit for storing said weighted intensity value for each pixel in said image dataset as a processed intensity value for each corresponding pixel in a processed output image dataset. The apparatus further comprises a calculation unit for rotating a local coordinate system in which the summation kernel is described resulting in that the coordinate axes of said local coordinate system coincide with the directions of the eigenvectors of said measure of the intensity structure, e.g. structure tensor, where said

eigenvectors are described in the global coordinate system of the image dataset, and a scaling unit for scaling the coordinate axes of the local coordinate system in which the summation kernel is described by an amount related to the eigenvalues of the measure of the intensity structure, e.g. structure tensor, via a width function $w(\lambda_i) = \sigma_i$, and wherein said eigenvalues depend on the amount of intensity variation in the direction of their corresponding eigenvectors, the width function being a decreasing function depending on the noise level in each pixel such that $w(0) = \sigma_{max}$ and $\lim_{a \rightarrow \infty} w(a) = \sigma_{min}$. The calculation unit for rotating and the scaling unit for scaling may be part of the second calculation unit.

In particular embodiments, the width function may be a decreasing function defined by

$$w(\lambda_i, x_0) = \begin{cases} \Delta\sigma e^{-\frac{\lambda_i}{d(x_0)} + \frac{1}{2}} + \sigma_{min}, & \lambda_i > \frac{d(x_0)}{2} \\ \sigma_{max}, & \lambda_i \leq \frac{d(x_0)}{2} \end{cases}.$$

The apparatus according to embodiments of the present invention may further comprise a repeating unit for repeating the above method steps for subsequent image datasets.

According to another aspect, a method is provided for processing a sequence of night vision image datasets, wherein said sequence comprises at least two image datasets each having at least two pixels, and wherein each pixel has an intensity value. The method comprises calculating, based on estimated intensity derivatives in each pixel, a measure of an intensity structure, e.g. structure tensor, for each pixel in an image dataset comprised in said sequence of image datasets; calculating values in a summation kernel based on said measure of the intensity structure, e.g.

structure tensor, for each pixel in said image dataset;
calculating a weighted intensity value for each pixel in
said image dataset, using as weights the values in said
summation kernel; storing said weighted intensity value for
5 each pixel in said image dataset as a processed intensity
value for each corresponding pixel in a processed output
image dataset. Calculating the values in the summation
kernel may be performed so that, when applying the
summation kernel to an input image, sharpening of contours
10 in the input image is obtained at the most important
spatial and temporal edges while smoothing is applied in
the appropriate directions in the input image data.

Performing sharpening while smoothing may be obtained
by using a kernel \bar{S}_H obtained by convolving an anisotropic
15 filter kernel with a high-boost filter. Alternatively,
performing sharpening while smoothing may be obtained by
using a kernel \bar{S}_h obtained by analytically convolving a
continuous approximation of a discrete high-boost filter
with an anisotropic filter function and discretizing the
20 resulting filter function.

In yet another aspect of the present invention, an
apparatus is provided for processing a sequence of night
vision image datasets, said sequence comprising at least
two image datasets each having at least two pixels, wherein
25 each pixel has an intensity value. The apparatus comprises
a first calculation unit for calculating, based on
estimated intensity derivatives in each pixel, a measure of
an intensity structure, e.g. a structure tensor, for each
pixel in an image dataset comprised in said sequence of
30 image datasets, and performing an eigen-decomposition of
said measure of the intensity structure, e.g. structure
tensor; a second calculation unit for calculating values in
a summation kernel based on said measure of the intensity

structure, e.g. structure tensor, for each pixel in said image dataset; a third calculation unit for calculating a weighted intensity value for each pixel in said image dataset, using as weights the values in said summation
5 kernel; a storage unit for storing said weighted intensity value for each pixel in said image dataset as a processed intensity value for each corresponding pixel in a processed output image dataset. The second calculation unit is adapted for calculating values in the summation kernel so
10 that, when applying the summation kernel to an input image, sharpening of contours in the input image is obtained at the most important spatial and temporal edges while smoothing is applied in the appropriate directions in the input image data.

15 According to embodiments of the present invention, the second calculation unit may be adapted for performing sharpening by using a kernel \bar{S}_H obtained by convolving an anisotropic filter kernel with a high-boost filter. According to alternative embodiments, the second
20 calculation unit may be adapted for performing sharpening by using a kernel \bar{S}_h obtained by analytically convolving a continuous approximation of a discrete high-boost filter with an anisotropic filter function and discretizing the resulting filter function.

25 According to yet another aspect of the present invention, a computer-readable medium is provided having embodied thereon a computer program for processing by a processor of a sequence of image datasets, wherein said sequence comprises at least two image datasets, each image
30 dataset having at least two pixels, wherein each pixel has an intensity value. The computer program comprises a first calculation code segment for calculating, based on estimated intensity derivatives in each pixel, a measure of

an intensity structure, e.g. a structure tensor, for each pixel in an image dataset comprised in said sequence of image datasets, and for performing an eigen-decomposition of said measure of an intensity structure, e.g. structure tensor; a second calculation code segment for calculating values in a summation kernel based on said measure of the intensity structure, e.g. structure tensor, for each pixel in said image dataset; a third calculation code segment for calculating a weighted intensity value for each pixel in said image dataset, using as weights the values in said summation kernel; a storage code segment for storing said weighted intensity value for each pixel in said image dataset as a processed intensity value for each corresponding pixel in a processed output image dataset.

15 The second calculation code segment is adapted for calculating values in the summation kernel so that, when applying the summation kernel to an input image, sharpening of contours in the input image is obtained at the most important spatial and temporal edges while smoothing is applied in the appropriate directions in the input image data.

In a computer-readable medium according to embodiments of the present invention, the second calculation code segment may comprise code for performing sharpening by using a kernel \bar{S}_H obtained by convolving an anisotropic filter kernel with a high-boost filter.

25 According to alternative embodiments, the second calculation code segment may comprise code for performing sharpening by using a kernel \bar{S}_h obtained by analytically convolving a continuous approximation of a discrete high-boost filter with an anisotropic filter function and discretizing the resulting filter function.

30 According to another aspect of the invention, the

method, apparatus, computer program product or computer-readable medium according to any one of the embodiments of the invention is used in a vehicle, robot, video camera, surveillance system or a manufacturing system, etcetera.

5 According to another aspect of the invention, the method, apparatus, computer program product or computer-readable medium according to any one of the embodiments of the invention is used for detecting and/or analyzing objects, such as e.g. pedestrians and/or traffic signs, in
10 the vicinity of a vehicle. The vehicle may e.g. be a vehicle selected from the group comprising: a car, a van, a bus, a truck, an articulated hauler, a boat, etc.. Furthermore, the method, apparatus, computer program product or computer-readable medium according to any one of
15 the embodiments of the invention may be used for a night vision system of a vehicle.

Further embodiments of the invention are defined in the dependent claims.

Some embodiments of the invention allow for enhancing
20 the dynamic level, such as e.g. reducing noise in a noisy image dataset, particularly, but not limited to image datasets with areas of low contrast such as image datasets or image dataset sequences obtained in low light level conditions or other conditions of poor visibility.

25 It is an advantage of some embodiments of the invention that they allow for noise reduction in an image dataset comprising low contrast areas, thus allowing for increased image resolution, flexibility, and clarity.

30 **Brief Description of the Drawings**

Further objects, features and advantages of embodiments of the present invention will appear from the following detailed description of embodiments of the

invention, reference being made to the accompanying drawings, in which:

Fig. 1 is a schematic illustration of a method according to an embodiment of the present invention;

5 Fig. 2 is an illustration of the appearance of a typical width function according to embodiments of the present invention;

10 Fig. 3 is an illustration showing the spatiotemporal intensity structure of a grey object moving from left to right in an image dataset sequence according to an embodiment of the present invention;

Fig. 4 is a schematic view of a method according to an embodiment of the present invention;

15 Fig. 5 is an illustration showing an increasing intensity transformation function according to an embodiment of the present invention;

Fig. 6 is a schematic view of a method according to an embodiment of the present invention;

20 Fig. 7 is a schematic view of an apparatus according to an embodiment of the present invention; and

Fig. 8 is a schematic view of a computer-readable medium according to an embodiment of the present invention.

25 Fig. 9 illustrates an estimation of a Gamma distribution from the histogram of the second largest eigenvalues. The dots represent the eigenvalues and the dashed line the least squares fitted Gamma distribution.

Detailed Description of Embodiments

30 Several embodiments of the present invention will be described in more detail below with reference to the accompanying drawings in order for those skilled in the art to be able to carry out the invention. The invention may, however, be embodied in many different forms and should not

be construed as limited to the embodiments set forth herein. Rather, these embodiments are provided so that this disclosure will be thorough and complete, and will fully convey the scope of the invention to those skilled in the art. The embodiments do not necessarily limit the invention. Furthermore, the terminology used in the detailed description of the particular embodiments illustrated in the accompanying drawings is not intended to be limiting the invention.

Embodiments of the present invention provide methods for enhancing the dynamic level, such as, e.g., reducing noise in a noisy image dataset, and/or enhancing the sharpness of object contours in an image dataset, particularly, but not limited to image datasets with areas of low contrast such as image datasets or image dataset sequences obtained in low light level conditions or other conditions of poor visibility that reduces the intensity contrast. The image dataset may, for example, be a 2D image dataset, sequences of 2D image datasets, a 3D image dataset, such as e.g. Magnetic Resonance Imaging and Computer Tomography imaging image dataset or sequences of 3D images or even higher-dimensional image datasets. The image datasets may in particular be night vision datasets.

Embodiments of the present invention provide methods for increasing sharpness in image datasets. Increasing sharpness may be performed per se, or in combination with dynamic level enhancement.

Embodiments of the present invention provide a digital image processing technique that utilizes summation (such as weighted averaging) to improve the visual quality of an image dataset, and particularly for image datasets obtained under conditions that have resulted in images comprising low contrast areas, e.g. areas suffering from

poor lighting conditions or other poor visibility conditions, such as haze and fog etc. The image datasets under consideration could also be, e.g. low contrast image datasets obtained by a medical imaging systems and other
5 imaging techniques resulting in low contrast image data. Alternatively, the low contrast image datasets may be night vision datasets.

The following description focuses on embodiments of the present invention applicable to an image dataset or a
10 sequence of image datasets and in particular to noise reduction and/or enhancement of sharpness of an object contour in an image dataset or a sequence of image datasets comprising low contrast areas, e.g. in a night vision image dataset.

15 In an embodiment of the present invention, according to Fig. 1, a method for noise reduction in a sequence of low contrast, e.g. night vision, image datasets is provided, wherein each image dataset has at least two pixels, wherein each pixel has an intensity value. The
20 method shown in Fig. 1 comprises:

- the calculation 11, based on estimated intensity derivatives in each pixel of an image dataset, of a measure of an intensity structure, e.g. a structure tensor, in a neighborhood around each pixel in an image dataset
25 comprised in the sequence of image datasets. A structure tensor or second-moment matrix is a matrix representation of partial derivative information in each pixel of in image. It is typically used to represent information about gradients or edges present in an image. Eigen-decomposition
30 may be applied to a structure tensor matrix to form the eigenvalues and eigenvectors, respectively.

- the calculation 12 of values in a summation kernel based on the measure of the intensity structure, e.g.

structure tensor, for each pixel in the image dataset.

- the calculation 13 of a weighted intensity value for each pixel in the image dataset, using as weights the values in the summation kernel.

5 - the storage 14 of the weighted intensity value for each pixel in said image dataset as a processed intensity value for each corresponding pixel in a processed output image dataset.

A method according to embodiments of the present
10 invention may furthermore comprise rotating a local co-ordinate system in which the summation kernel is described, resulting in that the co-ordinate axes of said local co-ordinate system coincide with the directions of the
eigenvectors of said structure tensor, where said
15 eigenvectors are described in the global co-ordinate system of the image dataset. The method may furthermore comprise scaling the co-ordinate axes of the local co-ordinate system in which the summation kernel is described by an amount related to the eigenvalues of the structure tensor
20 via a width function $w(\lambda_i) = \sigma_i$, the width function being a decreasing function such that $w(0) = \sigma_{\max}$ and $\lim_{a \rightarrow \infty} w(a) = \sigma_{\min}$, and wherein said eigenvalues depend on the amount of intensity variation in the direction of their corresponding
eigenvectors. For calculating the width σ_i of the summation
25 kernel in the direction of eigenvector e_i , i.e. for calculating the scaling of the axes in this direction, the width function depends on the eigenvalue λ_i that corresponds to the eigenvector e_i , and preferably does not depend on the eigenvalues corresponding to the other
30 eigenvector directions. Hence, the width function need only depend on the eigenvalue λ_i that corresponds to the eigenvector e_i .

A method according to embodiments of the present invention may include the repetition of the above steps for a subsequent image dataset until all image datasets of the sequence of image datasets have been processed.

5

IMAGE DATASET TYPE

In an embodiment of the present invention, said image dataset comprises a single 2D, 3D, or higher-dimensional monochrome image or a single multi-channel image, such as a
10 single color image.

In an embodiment of the present invention, said sequence of image datasets comprises 2D, 3D, or higher-dimensional monochrome image datasets or multi-channel image datasets, such as color image datasets.

15 An example of a multi-channel color image dataset is an RGB image dataset comprising three channels, e.g. one channel for red, one channel for green and one channel for blue. The special considerations that have to be taken when dealing with color images and other multi-channel image
20 datasets are referred to below.

In an embodiment of the present invention, the image dataset or sequence of image datasets is generated by a video camera, a still image camera, or other imaging devices for the purpose of e.g. medical imaging,
25 surveillance, industrial inspection, industrial and scientific measurements, and image-based navigation of different kinds of vehicles. The above mentioned imaging devices comprise in addition to video cameras and still image cameras for example microscopic 2D or 3D imaging
30 devices such as fluorescence microscopes, different kinds of image scanning devices, image digitizers and medical imaging devices such as Magnetic Resonance Imaging systems, Computed Tomography systems or Ultrasound systems. The

image datasets considered could also have been obtained using any kind of multi-spectral, image forming, measuring technique.

In embodiments of the present invention, the image
5 datasets or sequences of image datasets are night vision image datasets or sequences of night vision image datasets. The imaging devices for obtaining such night vision image datasets may be night vision cameras or night vision systems of a vehicle or a robot, for example.

10 In an embodiment of the present invention, the image dataset or sequence of datasets is generated by a system constructing a distance image of a scene, where the intensity in each pixel in the said distance image relates to the distance to an object in the corresponding direction
15 in the scene. This system could for example be comprised of a 3D camera or a set of multiple cameras.

STRUCTURE TENSOR

In an embodiment of the present invention, the
20 intensity structure depends on the first image dataset and/or preceding image datasets and/or subsequent image datasets of the sequence of image datasets.

In an embodiment of the present invention, the calculation of a measure of the intensity structure
25 comprises calculating a structure tensor J in each pixel, based on measures of the rate of intensity variation along the different axes in the spatiotemporal space in the current pixel. These measures may, e.g., be based on estimation of intensity derivatives in each pixel and/or
30 more generally based on weighted intensity values.

The structure tensor J may be used to analyze the local intensity structure in the neighborhood of each pixel and may be calculated using image intensity derivatives. An

example of the structure tensor is given in Eqn. 1. The dimensions in this example relate to an image dataset sequence of 2D image datasets where x and y represent the spatial dimensions and t represents the temporal dimension
 5 The image data is accordingly treated as 3D intensity data. The structure tensor, J , in the point (x,y,t) in the 3D intensity data may then generally be calculated according to Eqn. 1

$$J(x,y,t) = M(\nabla I(x,y,t)^T \nabla I(x,y,t)) = M \left(\begin{matrix} \left[\frac{\partial I}{\partial x} \right] \\ \left[\frac{\partial I}{\partial y} \right] \\ \left[\frac{\partial I}{\partial t} \right] \end{matrix} \begin{bmatrix} \frac{\partial I}{\partial x} & \frac{\partial I}{\partial y} & \frac{\partial I}{\partial t} \end{bmatrix} \right) \quad (1)$$

10 wherein $\frac{\partial I}{\partial x}$ denote the measure of the rate of intensity variation (intensity derivative estimation) along the x axis and respectively for the other axes, and wherein M denotes calculating, for each element in the matrix $\nabla I^T \nabla I$, a weighted intensity value (e.g. mean value) in a
 15 neighborhood around the currently processed pixel in the spatial, or in this case, spatiotemporal space. This weighted intensity value may be implemented by element-wise convolution with a bell-shaped function ,e.g. Gaussian kernel. The weighting kernel may be isotropic, i.e. of
 20 equal width in relation to all axes, and symmetric around each axis, i.e. if W is a discrete function representing the weighting kernel in 3D, $W(x,y,t)=W(-x,y,t)$, and respectively for the other axes . The weighted intensity value averages the gradient data ∇I around each pixel \bar{x} in
 25 the processed image dataset. The resulting structure tensor J is a matrix whose eigenvector, corresponding to the largest eigenvalue, points approximately in the gradient

direction and the other two eigenvectors are orthogonal to this vector. From eqn. 1 it may be observed that the structure tensor is constructed from the image intensity derivatives in the directions of the axes of the different
5 dimensions in the currently processed image dataset.

Before or during the calculation of the intensity derivatives, the image dataset may be low-pass filtered or smoothed, e.g. by convolving the image dataset with a symmetric isotropic bell-shaped, e.g. Gaussian, kernel.
10 Thus, high frequency content in the image dataset may be reduced or even removed. The low-pass filtering may thus be used to make the derivative estimations less sensitive to high frequency noise in the image dataset.

Input parameters for the structure tensor calculation
15 may be the maximum size of the neighborhood N used for the calculations of the measures of the rate of intensity variation and for the low-pass filtering in order to reduce or remove the high frequency noise. The maximum size of N may be set by parameters describing the extent of the
20 neighborhood in the spatial, or spatial and temporal, dimensions. One parameter may be the maximum extent in the spatial dimensions, n_s , indicating for example in a 2D image dataset a maximum spatial extent as being $n_s \times n_s$ pixels. For the temporal dimension, in the case of an image
25 dataset sequence, the parameters might be one parameter for the number of image datasets used preceding the first image dataset in time, n_b , and one parameter for the number of image datasets used succeeding the first image dataset in time, n_f . These parameters may be set as global
30 parameters.

If the image dataset is a multi-channel image dataset, for example a color dataset comprising different

color channels, there are different options on how to calculate the measure of the intensity structure. Hereinbelow, four different options are explained.

In an embodiment of the present invention, wherein
5 said image dataset is a multi-channel image dataset, the calculation 11 comprises calculating a separate structure tensor $J_i(\bar{x})$ for each pixel \bar{x} and each channel i following the procedure above and keeping all structure tensors for forthcoming stages in the processing. This means that for
10 an RGB image dataset there will be one structure tensor for each of the R, G, and B channel and that each channel is treated as a separate monochromatic image dataset, similar to the procedure for monochromatic sequences according to embodiments of the present invention.

15 In another embodiment of the present invention, wherein said image dataset is a multi-channel image dataset, the calculation 11 comprises calculating a separate structure tensor $J_i(\bar{x})$ for each pixel \bar{x} and each channel i and then averaging the corresponding structure
20 tensor elements over the different channels into one structure tensor $J(\bar{x})$ for pixel \bar{x} . This structure tensor will then represent all channels at the current pixel in the following stages of the method according to some embodiments of the present invention. This means that for a
25 RGB image dataset there will be one single structure tensor for the R, G, and B channels to be used in the following stages.

In an embodiment of the present invention, wherein
30 the image dataset is a multi-channel image dataset, the calculation 11 comprises calculating a mean value for the respective measures of the rate of intensity variation (intensity derivative estimations) over the different

channels and based on these mean values calculating a single structure tensor $J(\bar{x})$ to be used in the current pixel \bar{x} in the following stages.

In an embodiment of the present invention, wherein
5 the image dataset is a multi-channel image dataset, the calculation 11 comprises calculating a mean value for the intensity value in R, G and B channels and to calculate measures of the rate of intensity variation based on this mean intensity value. These measures of the rate of
10 intensity variation are then used to calculate a single structure tensor $J(\bar{x})$ to be used in the current pixel \bar{x} in the following stages.

SUMMATION KERNEL

15 The calculation 12 of the summation kernel may comprise calculating the eigenvectors and eigenvalues of the structure tensor for the currently processed pixel \bar{x} and subsequently calculating a summation function S based on said eigenvalues and eigenvectors. The summation
20 function is then discretized, forming a finite block of pixels called the summation kernel \bar{S} . Using said eigenvalues and eigenvectors for the calculation of the summation function gives a flexible method where the influence of the intensity levels and the magnitudes of the
25 intensity gradients on the shape of the summation function may be adapted to the current image dataset or to different local areas in the current image dataset.

In an embodiment of the present invention, the calculation 12 results in rotating the local coordinate
30 system in which the summation kernel is described so that the coordinate axes of said coordinate system coincides with the directions of the eigenvectors of the structure

tensor, where said eigenvectors are described in the global coordinate system of the image dataset. Further, the width of the summation function along the directions of said eigenvectors depends on the corresponding eigenvalues via a function called the width function. The magnitude of the eigenvalues corresponds to the amount of intensity variation in the direction of the corresponding eigenvectors. The summation function in this way adapts itself to the intensity structure in the neighborhood around the each pixel currently being processed.

In another embodiment of the present invention, the calculation 12 comprises diagonalizing the structure tensor in the currently processed pixel x according to the following equation, known in the field of mathematics and image processing,

$$J = RLR^T \quad (2)$$

where R is an orthogonal rotation matrix that has the normalized eigenvectors of J as column vectors. The matrix L is a diagonal matrix that has the corresponding eigenvalues of J as diagonal elements. As an example in 3D,

$$L = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \quad (3)$$

where λ_1 is the eigenvalue corresponding to the eigenvector in this first column of R etc.

25

The calculation 12 also comprises calculating the widths σ_i of the summation function in the directions of the corresponding eigenvectors using the eigenvalues λ_i . This calculation is managed by the width function $w(\lambda_i) = \sigma_i$.

The width function w is a decreasing function such that $w(0) = \sigma_{max}$ and $\lim_{a \rightarrow \infty} w(a) = \sigma_{min}$.

In particular embodiments of the present invention, the width function w may be defined as an exponential
 5 function which depends directly on the eigenvalue λ_i in the current eigenvector direction e_i , for example in the following way:

$$w(\lambda_i, x_0) = \begin{cases} \Delta\sigma e^{-\frac{\lambda_i}{d(x_0)} + \frac{1}{2}} + \sigma_{min}, & \lambda_i > \frac{d(x_0)}{2} \\ \sigma_{max}, & \lambda_i \leq \frac{d(x_0)}{2} \end{cases}$$

where $\Delta\sigma = (\sigma_{max} - \sigma_{min})$, so that $w(\lambda_i) = \sigma_i$ attains its
 10 maximum σ_{max} below $\lambda = \frac{d}{2}$ and asymptotically approaches its minimum σ_{min} when $\lambda \rightarrow \infty$. The quantity $d(x_0)$ is introduced which is a measure of the variance of the noise in the point x_0 , as explained below.

Since structure tensor J is a positive semi-definite
 15 matrix, all $\lambda_i \geq 0$ and the preset constants $\sigma_{max} > \sigma_{min} \geq 0$ consequently correspond to the maximum and minimum widths of the kernel.

When the widths $w(\lambda_i, x_0)$ have been calculated, a summation function may subsequently be constructed for
 20 every point x_0 according to eqn. 4:

$$S(x) = e^{-\frac{1}{2}(x-x_0)^T R \Sigma^2 R^T (x-x_0)} \tag{4}$$

The basic form of the summation function S that is constructed at each point x_0 is that of a bell-shaped function, e.g. a Gaussian function, including a rotation
 25 matrix R and a scaling matrix Σ . The rotation matrix is constructed from the eigenvectors e_i of structure tensor J :

$$R = [e_1 \quad e_2 \quad e_3]$$

while the scaling matrix has the following form:

$$\Sigma = \begin{bmatrix} \frac{1}{\sigma_1} & 0 & 0 \\ 0 & \frac{1}{\sigma_2} & 0 \\ 0 & 0 & \frac{1}{\sigma_3} \end{bmatrix}.$$

After discretizing the summation function, the resulting summation kernel may be applied to an original image f_{in} at each spatiotemporal point x_0 so as to calculate the output intensity f_{out} of the smoothing stage in the current pixel x_0 .

Fig. 2 illustrates the appearance of a width function. This choice of the width function gives a large width of the summation function in the direction of an eigenvector corresponding to a small eigenvalue, i.e. in directions with small intensity variation, and a small width in the direction of an eigenvector corresponding to a large eigenvalue and consequently large intensity variation.

Moreover, the calculation may comprise constructing the summation function S for the current point using an auxiliary function G with the following characteristics. The characteristics are presented in the case of 2D, for higher dimensions the analogous characteristics should be considered.

- 1) G should attain its maximum value in the origin, i.e. $G(0,0) \geq G(x,y)$ for all x and y .
- 2) G should be a decreasing function in all directions, i.e. $G(x_2,y) \geq G(x_1,y)$ if $|x_2| > |x_1|$ and $G(x,y_2) \geq G(x,y_1)$ if $|y_2| > |y_1|$.
- 3) G should be isotropic in relation to the coordinate axes, i.e. $G(x,y) = G(y,x)$ for all x and y .
- 4) G should be symmetric in relation to each coordinate axis, i.e. $G(x,y) = G(-x,y)$ and $G(x,y) = G(x,-y)$ for all x and y .

y .

The auxiliary function G could for example be a bell-shaped function, such as for example the Gaussian function

5

$$G(x,y) = e^{-\frac{1}{2}(x^2+y^2)} \tag{5}$$

The calculation 12 may further comprise scaling and/or rotating the local coordinates \hat{x} , corresponding to a local coordinate system in the neighborhood N centered on the currently processed pixel \bar{x} . This is accomplished using the matrix R and a width matrix Σ . Let, in the case of three dimensions,

15

$$\Sigma = \begin{bmatrix} \frac{1}{\sigma_1} & 0 & 0 \\ 0 & \frac{1}{\sigma_2} & 0 \\ 0 & 0 & \frac{1}{\sigma_3} \end{bmatrix}$$

wherein $w(\lambda_i) = \sigma_i$ are the widths calculated from the eigenvalues λ_i using the width function w . The summation function S is then constructed from G in the following way

$$S(\hat{x}) = G(\Sigma R^T \hat{x})$$

20 In the case of a Gaussian function in a 3D setting, the resulting summation function will be

$$S(x,y,t) = e^{-\frac{1}{2}(R^T \hat{x})^T \begin{bmatrix} \frac{1}{\sigma_1^2} & 0 & 0 \\ 0 & \frac{1}{\sigma_2^2} & 0 \\ 0 & 0 & \frac{1}{\sigma_3^2} \end{bmatrix} (R^T \hat{x})} \tag{6}$$

25

In an embodiment of the present invention, the

calculation 12 further comprises discretizing the first summation function S so that it constitutes a finite 2D, 3D or 4D block according to the dimension of the currently processed image dataset, with boundaries defined by the numbers n_s , n_b and n_f , wherein n_s is a predetermined maximal number of pixels along each spatial dimension, n_b is a predetermined maximal number of preceding image datasets, and n_f is a predetermined maximal number of subsequent image datasets comprised in the sequence of image datasets. The block of pixels may be referred to as a summation kernel \bar{S} hereinafter.

The calculation 12 may moreover comprise normalizing the values of the summation kernel so that the total sum of all pixel values within the summation kernel is 1.

Through the rotation and scaling described above, the summation function S may be adjusted according to the intensity variation in the image dataset and, hence, the summation function may be automatically elongated along intensity streaks in the spatiotemporal space that corresponds to motion trajectories in an image sequence. Thus, this method does not utilize an explicit motion estimation to estimate motion trajectories like some prior art methods, but obtains such trajectories more or less automatically. Since the method does not distinguish between axes that represent spatial dimensions and the axis that represents a possible temporal dimension, the shape of the kernel adjusts to all intensity structures in a similar way, without considering whether it represents a motion trajectory or not.

In an embodiment of the present invention, wherein the image dataset is a multi-channel image dataset, e.g. a color image data set, the calculation 12 comprises

calculating one summation kernel \bar{S}_i , based on each structure tensor $J_i(\bar{x})$ calculated in the calculation 11 step for each pixel \bar{x} .

In another embodiment of the present invention,
5 wherein the image dataset is a multi-channel image dataset, e.g. a color image dataset, the calculation 12 comprises calculating the summation kernel \bar{S} based on the structure tensor $J(\bar{x})$ common for all channels calculated in the calculating 11 step for each pixel \bar{x} .

10 In a further embodiment of the present invention, wherein the image dataset is a multi-channel image dataset, the calculating 12 comprises calculating a separate summation kernel for all channels, and the calculating 13 comprises summation with a separate summation kernel for
15 all channels.

Fig. 3 is an illustration showing the spatiotemporal intensity structure of a grey object moving from left to right in an image dataset sequence wherein t indicates a temporal dimension and x indicates a spatial axis. At
20 point A and point B, the summation kernel 31 is elongated in the t and the x direction, respectively. At point C the summation kernel 31 is elongated along the grey intensity structure, which would be the result of the present embodiments.

25 It may be noted that a restriction of only being able to change the width of the neighborhood over which the summation is performed, symmetrically around the spatial and temporal axes will in general result in including intensities in the, in this case, white background of the
30 images, which will lead to an image where contours and edges are blurred. In the present method, the summation kernel is automatically rotated and shaped as to fit the

general intensity structure. In the case of an image sequence this may for example lead to a summation that is performed mainly along a direction that is adjusted to the direction of motion of a moving object. This means that the method enables applying summation using a summation kernel that is extended over a larger area while keeping the amount of blurring to a minimum in the resulting noise reduced output image dataset.

10 LOCAL ADAPTATION BASED ON LOCAL NOISE - NOISE MEASURE

The amount of noise in an image sequence changes depending on the brightness level. Since, according to embodiments of the present invention, it is desired that the algorithm adapts to changing light levels, both spatially within the same image and/or temporally in the image sequence, the above width function, in embodiments of the present invention, preferably depends on the noise level step 110 in Fig. 1. This adaptation is governed in accordance with embodiments of the present invention by the measure d . A less preferred option for the noise level is that the width function is adapted to the SNR.

First, the noise is estimated in each 2D image. This is done by subtracting a smoothed image e.g. an isotropically smoothed version of the input image from the original image. Typically, a Gaussian kernel with a suitable standard deviation, e.g. 2 is used for this operation but the invention is not limited thereto. For example any bell-shaped function can be used or any function that decreases with distance from the pixel under consideration. This provides a sequence of noise images $f_n(x)$. 2D images are used in this case since motion will give a response in the noise estimation if isotropic spatiotemporal smoothing is applied. The measure d is then

a scaled estimation of the variance ν of the noise f_n in a spatiotemporal neighborhood Ω centered at x_0 , e.g.:

$$d(x_0) = \rho\nu(x_0) = \frac{\rho}{N-1} \sum_{\Omega} (f_n(x) - \bar{f}_n(x_0))$$

where N is the number of pixels and $\bar{f}_n(x_0)$ is the mean of
 5 the noise in the neighborhood Ω centered at x_0 . The scaling factor ρ is a constant that is multiplied with the variance $\nu(x_0)$ so that the measure $d(x_0)$ will be of a magnitude that is suitable for insertion in the width function w . A suitable value for ρ can be found that works for most image
 10 sequences. In experiments, $\rho = \frac{1}{64}$ has been generally used. However, hereinafter a procedure is suggested for the estimation of ρ .

As an alternative to working with a pre-determined fixed value for the scaling factor ρ , e.g. $\rho = \frac{1}{64}$, a value
 15 for the scaling factor ρ may be estimated from real images, e.g. on the fly, as explained hereinbelow. If the image sequence is modeled as a signal with added independent identically distributed (i.i.d.) Gaussian noise of zero mean, the following can be said about the noise in the
 20 estimated eigenvalues.

First of all the sequence is filtered with Gaussian filters as well as differentiation filters to estimate the gradients in the sequence at every point. The noise in the estimated gradients will hence also be Gaussian but with a
 25 different variance. The structure tensor is calculated from the estimated gradients. As the structure tensor is made up of the outer product of the gradients, and the tensor is then filtered elementwise, the tensor elements will be a sum of squares of Gaussian distributed variables. A sum of
 30 n squared Gaussian distributed variables, with variance ν ,

will be Gamma distributed with mean $n\nu$ and variance $(n\nu)^2$. The noise in the filtered tensor elements will hence be Gamma distributed and so will the noise in the eigenvalues.

In figure 9 a histogram of the second largest
 5 eigenvalue of the structure tensors for an image is shown in dots, curve 90. A Gamma probability density function (PDF), which depends on two independent parameters, can be fitted to such a histogram by any suitable regression, e.g. in a least squares manner. This gives us the mean $\bar{\lambda}$ and
 10 standard deviation $S\lambda$ of the Gamma distribution. The resulting Gamma distribution is shown as the dashed line, curve 91. Experiments have shown that the histogram of the eigenvalues generally follow a Gamma distribution well. More generally a parameter function may be used to fit the
 15 histogram, whereby the parameters may be used in similar way as the mean $\bar{\lambda}$ and standard deviation $S\lambda$ of the Gamma distribution described. From the parameter values obtained with the fit, the overall measure b then may be derived based on a correlation between the parameters and the
 20 overall measure. The number of parameters used may be two or more.

An overall measure b of the noise level in an eigenvalue λ_i can then be estimated as the mean plus a number n times the standard deviation, i.e. $b = \bar{\lambda} + ns\lambda$. Any
 25 of the eigenvalues λ_i could be used to calculate b . In the spatio-temporal case with a 3x3 structure tensors the second largest eigenvalue may often be used for the calculation of b .

An estimate of the global noise variance ν in the
 30 image can be calculated as

$$\nu = \frac{1}{N-1} \sum_{\Omega} (f_n(x) - \bar{f}_n)$$

where the summation is taken over the whole image, i.e. Ω refers to the whole image, N is the number of pixels in this image and \bar{f}_n is the global mean over the whole image.

5 The measure b is of the same magnitude in relation to the global variance ν as the magnitude of the local measure $d(x_0)$ in relation to the local variance $\nu(x_0)$. This gives a way of calculating the scaling factor ρ , $\rho = b/\nu$ where b and ν are estimated from the images.

10

SHARPENING OF OBJECT CONTOURS

A structure tensor is surprisingly good at finding the direction of motion in very noisy input sequences. However, in some cases the elongation of the constructed summation kernels can be slightly misaligned with the motion direction. In these cases the contours of the moving objects can be somewhat blurred. In embodiments of the present invention the addition of some high-boost filtering is proposed to sharpen up these contours. The high-boost filter H is defined as a 3x3x3 tensor with the value -1 in all elements except the center element which has the value 27A - 1, as described by R. Gonzales and R. Woods in "Digital Image Processing", Addison-Wesley, 1992. The best results have been attained using $A = 1.2$. If the filter is applied after the complete smoothing summation process, problems of negative values and rescaling of intensities will be encountered, which will give a global intensity flickering in the image sequence.

25
30 According to embodiments of the present invention, it is instead proposed that, for each point in the image, the

constructed summation kernel \bar{S} is filtered with a highboost filter, step 100 in Fig. 1, thus yielding a new kernel \bar{S}_H . This kernel is then normalized so that its coefficients add up to 1. In this way, application of the constructed kernels \bar{S}_H in the summation step will add some sharpening at the most important spatial and temporal edges together with smoothing in other appropriate directions.

In another embodiment of the invention, not illustrated in the drawings, the discrete high-boost filter H can be approximated by a continuous function h , for example

$$h = e^{-\frac{1}{2\sigma^2}|\hat{x}|^2} \left(1 - \frac{\alpha^2}{2}|\hat{x}|^2\right) \tag{7}$$

By using h a summation function S_h with sharpening functionality can be calculated by performing the convolution between h and S analytically,

$$S_h = S * h$$

Using the example function in equation (7) the summation function S_h is given by

$$S_h = e^{-0.5(\hat{x}_{hb} \cdot \hat{x}_{hb})} \left(1 - \frac{3\alpha^2\sigma^2}{2} - \frac{\alpha^2\sigma^4}{2}((\hat{x}_{hb2} \cdot \hat{x}_{hb2}) - (s \cdot s))\right)$$

where

$$\begin{aligned} \hat{x}_{hb} &= \left[\frac{x_r}{\sqrt{\sigma_x^2 + \sigma^2}} \quad \frac{y_r}{\sqrt{\sigma_y^2 + \sigma^2}} \quad \frac{t_r}{\sqrt{\sigma_t^2 + \sigma^2}} \right], \\ \hat{x}_{hb2} &= \left[\frac{x_r}{\sigma_x^2 + \sigma^2} \quad \frac{y_r}{\sigma_y^2 + \sigma^2} \quad \frac{t_r}{\sigma_t^2 + \sigma^2} \right], \\ s &= \left[\frac{1}{\sqrt{\sigma_x^2 + \sigma^2}} \quad \frac{1}{\sqrt{\sigma_y^2 + \sigma^2}} \quad \frac{1}{\sqrt{\sigma_t^2 + \sigma^2}} \right]. \end{aligned}$$

Here x_r , y_r and t_r are the local coordinates rotated by the previously described rotation matrix R , i.e.

$$\begin{bmatrix} x_r \\ y_r \\ t_r \end{bmatrix} = R^T \begin{bmatrix} x \\ y \\ t \end{bmatrix} = R^T \hat{x}$$

where R is obtained from the eigenvectors of the structure tensor. The widths σ_x , σ_y and σ_t are calculated as usual using the width function w and the eigenvalues of the structure tensor. The parameters σ and ρ are inherited from the function h . The general form of the summation function S_h is used globally and the parameters σ and ρ are also set globally.

Other summation functions S_h can similarly be constructed using other choices of h which approximates a discrete sharpening filter. After obtaining S_h , the function is normalized and discretized before performing the filtering summation, giving the discrete summation kernel \bar{S}_h .

15

SUMMATION

In an embodiment of the present invention, the calculation 13 comprises calculating a weighted intensity value of the intensities in a neighborhood N of each pixel in the currently processed image dataset by using:

20

$$O(\bar{x}) = \sum_{\hat{x} \in N} \bar{S}(\hat{x}) I(\bar{x} + \hat{x}) \tag{8}$$

wherein $O(\bar{x})$ defines the intensity in pixel \bar{x} in the first noise reduced output image dataset, \bar{S} is the summation kernel calculated for pixel \bar{x} in the calculating 12 step and I is the input image. In embodiments including sharpening, e.g. as illustrated by step 100, the summation kernel \bar{S} can be replaced by \bar{S}_H or \bar{S}_h . The coordinate \hat{x} refers to a local coordinate system in the neighborhood N .

25

This step of calculating the weighted average is hereinafter also referred to as summation. The size of the neighborhood N used in the summation is determined by the parameters n_s , n_f and n_b defined above.

5 In an embodiment of the present invention, wherein the image dataset is a multi-channel image dataset, e.g. an RGB color image data set, and the calculation 12 comprises the calculation of one summation kernel \bar{S}_i for each channel i , e.g. one for each of the R, G and B channels, then the
10 weighted intensity value may be calculated by using:

$$O_i(\bar{x}) = \sum_{\hat{x} \in N} \bar{S}_i(\hat{x}) I_i(\bar{x} + \hat{x}) \quad (9)$$

wherein I_i represents the image dataset in channel i in the first multi-channel input image dataset I and O_i
15 represents the corresponding channel in the first noise reduced multi-channel output image dataset O .

In an embodiment of the present invention, wherein the image dataset is a multi-channel image dataset, e.g. an RGB color image data set, and the calculation 12 comprises
20 the calculation of one common summation kernel \bar{S} for all channels i in the first multi-channel input image dataset I , then the weighted intensity value may be calculated by using

$$25 \quad O_i(\bar{x}) = \sum_{\hat{x} \in N} \bar{S}(\hat{x}) I_i(\bar{x} + \hat{x}) \quad (10)$$

RENDERING

In an embodiment of the present invention, the method
30 further comprises rendering 15 of the created output image

dataset for presentation on a display.

The rendering 15 may comprise transmitting each pixel in each first noise reduced output image dataset to an output unit, such as a display for presentation.

5

STORING

In an embodiment of the present invention, the storing 14 comprises storing each first noise reduced output image dataset on a storage device. Optionally, each
10 first noise reduced output image dataset may be stored on a storage device, such as a hard drive.

IMAGE STABILIZATION

In some cases the currently processed image dataset
15 may comprise motion disturbances. This may be the case for an image dataset when the obtaining unit, such as a video camera, obtaining the image dataset is shaking or vibrating while obtaining the consecutive image datasets. In these cases, the image dataset may be processed so that the
20 influence of this disturbance is reduced to a minimum. A commonly known technique of reducing this motion disturbance is image stabilization, which is implemented in many commercial cameras. Image processing techniques for image stabilization by comparing consecutive images is
25 known in the field and may be implemented by a person ordinary skilled in the art.

In an embodiment of the present invention, wherein the currently processed image dataset comprises motion disturbances, the method further comprises performing image
30 stabilization, step 16 in Fig. 1. The order in which the step of performing image stabilization step and the intensity transformation step (see hereinbelow) are performed may be changed.

INTENSITY TRANSFORMATION

In some cases, the currently processed image dataset, two or more of the image datasets of the sequence of image datasets, or the entire sequence of image datasets comprise
5 low contrast areas, in particular e.g. in night vision image datasets. This may, e.g., be the result of obtaining the image datasets in low light level conditions, or more generally low contrast conditions. In these cases, the
10 intensities may be transformed using an intensity transformation. The term intensity transformation is known to a person of ordinary skill in the art.

Performing an intensity transformation on an image dataset having low contrast areas generally results in
15 increasing the dynamic range of the intensities in the low contrast areas in the image dataset. For example, in image datasets containing very dark areas, e.g. areas having a very low mean intensity, the intensities may be amplified heavily, while in brighter areas having a higher mean
20 intensity the amplification may be significantly smaller. After the intensity transformation, the image dataset will in general have higher contrast and the depicted structures will be more visible. However, because of high uncertainty in the true intensity values when the image dataset is
25 obtained under poor visibility conditions (e.g. low light level conditions such as night vision conditions) the image data may contain large amounts of noise. This noise may be amplified during the intensity transformation, resulting in the noise being even more visible and disturbing in the
30 image dataset.

In an embodiment of the present invention, according to Fig. 4, the method further comprises:
transformation 41 (411, 412, 413) of the intensities in the

image dataset.

The calculating 13 step of the method in general reduces the noise in the image dataset, either after the intensity transformation, according to the transforming 411
5 step and the transforming 412 step, or before the intensity transformation, according to the transforming 413 step.

The transformation 41 is either performed before calculating 11 step, calculating 12 step and calculating 13
10 step, as indicated by 411 in Fig. 4, after calculating 11 step but before the calculating 12 step and calculating 13 step, as indicated by 412 in Fig. 4, or after calculating 11, calculating 12 step and calculating 13 step, as indicated by 413 in Fig. 4.

The decision of when to perform the intensity
15 transformation might affect the calculating 12 step of the summation kernel. For example, performing the structure tensor calculation in the calculating 11 step before the transforming 41 step, as indicated by 412 and 413 in Fig. 4, results in the eigenvalues of the structure tensor
20 generally being much smaller than if the transforming 41 step is performed before the calculating 11 step of the structure tensor, as indicated by 411 in Fig. 4. If the same structures in the image dataset should give the same widths of the summation function in the two cases, it might
25 be necessary to change the steepness of the decrease of the width function w , used in the calculating 12 step. This change of the width function w can be performed automatically using a technique as described in the section LOCAL ADAPTATION BASED ON LOCAL NOISE - NOISE MEASURE, step
30 110 in Fig 1.

As explained above, the intensity transforming 41 step of the method may transform the original intensity values of the image dataset to new intensity values, so

that the contrast is increased in areas with low intensity contrast and a small dynamic range.

In an embodiment of the present invention the intensity transforming 41 step comprises utilizing an increasing function T such that $I_T(\bar{x})=T(I(\bar{x}))$, where I_T is the transformed image dataset and I is the original image dataset. The function T may be constructed such that the range of intensities constituting the dark areas, or, generally, the low contrast areas, is being mapped to a wider dynamic range in the resulting image dataset. An example of a transformation function T is illustrated as the smooth curve in Fig. 5. A way of constructing such an intensity transformation function T is by using histogram equalization, which is a commonly known technique within the field of image analysis. The construction of the function T used for the intensity transforming 41 step of the currently processed image dataset is in some embodiments of the invention either performed using only the currently processed image dataset or, in the case of an image dataset sequence, performed using the currently processed image dataset together with a number of preceding ($\max n_b$) and succeeding ($\max n_f$) image datasets. In some embodiments of the present invention, the image dataset may be divided into small tiles where the intensity transformation is performed locally within each tile. These tiles will generally have the same number of dimensions as the complete image dataset, i.e. for an image dataset sequence comprising 2D image datasets the tiles may be three-dimensional and include, apart from the currently processed image dataset also preceding ($\max n_b$) and succeeding ($\max n_f$) image datasets. This type of contrast enhancing intensity transformations and histogram

equalization are known techniques and may be implemented by a skilled person within the field of image analysis.

Different techniques for intensity transformation and histogram equalization on multi-channel image datasets, such as RGB image datasets, has been previously proposed and may be applied at the intensity transforming step 41 of the method according to an embodiment of the present invention. One suitable technique for intensity transformation on a RGB image dataset comprises transforming the image dataset to a representation where one of the channels represent the intensity of the color, such as the HSV (Hue, Saturation and Value) representation where the V (value) channel represents the intensity, and then apply the intensity transformation on this channel of the image data and leave the other channels unaffected.

In an embodiment of the present invention the method further comprises performing an intensity transformation on an image dataset by an intensity transformation function to increase the intensity values of at least a part of the image dataset.

OPTICAL FLOW

In the case of image dataset sequences obtained by a video camera that is moving, e.g. a night vision camera, all the pixels in the image dataset sequence are in general occupied by moving structures. In this case the calculation comprises calculating a measure of the intensity structure by, for each pixel in the image dataset, finding a vector, commonly called a motion vector, that connects the current pixel with the pixel in the succeeding image that corresponds most accurately to the same object point in the depicted scene as the current pixel. This can for example be done using optical flow estimations resulting in

an optical flow field. However, there are other methods for calculating the above mentioned vectors, all of which may be implemented by embodiments of the present invention. In the following, it should be appreciated that some form of optical flow estimation is applied. The calculation 12 then comprises calculating a summation function based on the optical flow field.

The terms optical flow field and optical flow estimation are known for a person ordinary skilled in the art. Any method for calculating the optical flow may be implemented by embodiment of the present invention. Such methods may include, but is not limited to, applying image derivative approximations which are used together with the so-called optical flow constraint equation or so-called block matching methods. Block matching methods generally identify the area in the succeeding image dataset that best matches the neighborhood around the currently processed pixel in the currently processed first image dataset. The vector connecting the currently processed pixel and the central pixel of the area with the best match is defined as the optical flow vector for the currently processed pixel.

In an embodiment of the present invention, the calculation 11 comprises calculating optical flow vectors for each point \bar{x} in the currently processed image dataset.

Since the image sequence may be very noisy the high frequency content of the currently processed image may be reduced before or during the optical flow estimation, e.g. low-pass filtering or smoothing may be applied. This can, for example, be performed by convolution with an isotropic symmetric Gaussian kernel of the proper number of dimensions (3D Gaussian kernel for a 2D image sequence). Generally, a standard deviation in the range of 1 and 2 pixels will be suitable for the Gaussian kernel in this

processing step.

In this embodiment, the calculating 12 comprises constructing an auxiliary function G according to the characteristics of the function G as explained
 5 hereinabove;

constructing a rotation matrix R . The rotation matrix may be constructed in the following way. In the case of a 2D optical flow vector in pixel \bar{x} being obtained from calculation (11), the vector is first extended to 3D by
 10 interpreting the time difference between the currently processed image and the succeeding image as one pixel step along the temporal axis. This vector, after normalization to length 1, is then used as the first column vector in the matrix R . The other two column vectors are then chosen as
 15 a pair of orthogonal vectors of length 1 that are orthogonal to the first column vector. The vectors should be chosen such that the three column vectors relative orientation matches the relative orientation of the axes in the spatiotemporal space in which the coordinates of the
 20 image dataset is represented.

Furthermore, in this embodiment the calculation 12 further comprises

constructing a width matrix Σ such that

$$\Sigma = \begin{bmatrix} \frac{1}{\sigma_{max}} & 0 & 0 \\ 0 & \frac{1}{\sigma_{min}} & 0 \\ 0 & 0 & \frac{1}{\sigma_{min}} \end{bmatrix}$$

25 where the parameters $\sigma_{max} > \sigma_{min} \geq 0$ are preset constants corresponding to the maximum and minimum widths of the kernel,

constructing a summation function S for each pixel \bar{x}
 30 in the currently processed image dataset using

$$S(\hat{x}) = G(\Sigma R^T \hat{x})$$

wherein \hat{x} are the coordinates of a point in a local coordinate system in a neighborhood N , centered on the currently processed pixel \bar{x} ; and

5 calculating the discrete summation kernel \bar{S} by discretizing the summation function S so that it constitutes a 2D, 3D or 4D block according to the dimension of the currently processed image dataset, with boundaries defined by the numbers n_s , n_b and n_f , wherein n_s is a
10 predetermined maximum number of pixels along each spatial dimension, n_b is a predetermined maximum number of preceding image datasets, and n_f is a predetermined maximum number of subsequent image datasets comprised in the sequence of image datasets.

15 In an embodiment of the present invention the motion of the camera is a pure translational motion or an almost pure translational motion. In this embodiment the optical flow method is modified so that the output flow field is constrained to being a translational flow field, e.g. an
20 optical flow field that is the result of a translational motion. An extra option is to also preset the length of the flow vectors (motion vectors). Following this, only two parameters for the complete flow field in the image dataset, the so-called focus of expansion, needs to be
25 calculated. The focus of expansion is a pixel in the image dataset plane that corresponds to the projection of the direction of translation of the camera onto the image plane. If data regarding the speed by which the camera is moving, a model of the camera (a model of how the object
30 points are projected onto the image plane), and a reasonable approximation of the depth of the scene in the different parts of the image plane is available, a useful

pre-estimation of the length of the flow vectors may be performed. If data regarding the translation direction of the camera is available this information may additionally be used to calculate an improved estimation of the translational flow field.

In another embodiment of the present invention the camera is being mounted in a car and the car is being equipped with a computer, e.g. an onboard computer, with capability to store the speed and driving direction of the car. In this embodiment the speed data for the estimation of the length of the optical flow vectors, according to the last paragraph, may be obtained from the computer. The speed and the driving direction data, together with a camera model and a depth estimation, may optionally also be used directly to estimate a reasonable complete (or almost complete) flow field, with or without constraints on being a translational flow field.

In an embodiment of the present invention, the calculation of a measure of the intensity structure further comprises calculating the optical flow vectors (motion vectors) for each pixel of each image dataset of the sequence of image datasets.

FOREGROUND BACKGROUND SEGMENTATION

In the case when the image dataset sequence to be processed is obtained by a static video camera, e.g. a video camera that does not move between obtaining consecutive image datasets, an optional procedure for processing is provided.

In an embodiment of the present invention, according to Fig. 6, the method further comprises:

performing a low-pass filtering or a smoothing, on the first image dataset, e.g. by performing a convolution

using an isotropic, symmetric Gaussian kernel, with the number of dimensions matching that of the image dataset (3D Gaussian kernel for a 2D image dataset sequence). This will in general reduce the high frequency noise in the first
5 image dataset and prevent that noise is classified as foreground objects, i.e. relevant moving objects, in the succeeding foreground-background segmentation of the first image dataset.

In an embodiment of the present invention, the method
10 further comprises

performing 62 foreground-background segmentation, which identifies pixels that are occupied by a moving object, e.g. a pedestrian, and classifies those pixels as foreground pixels. Pixels that are identified as not being
15 occupied by a moving object are classified as background pixels. There are currently various known methods that may be used for foreground-background segmentation in this embodiment. The foreground-background segmentation method is allowed to use the same number of image datasets
20 preceding and succeeding the currently processed image, n_b and n_f , as the calculation 11 step and the calculation 13 step.

In an embodiment of the present invention the method further comprises

25 calculating 63 a summation kernel \bar{S} using a summation function S . The summation function S is constructed from an auxiliary function G which in turn is constructed according to the characteristics of G listed in the SUMMATION KERNEL section. Further, a width matrix Σ
30 is constructed as

$$\Sigma = \begin{bmatrix} \frac{1}{\sigma_{max}} & 0 & 0 \\ 0 & \frac{1}{\sigma_{min}} & 0 \\ 0 & 0 & \frac{1}{\sigma_{min}} \end{bmatrix}$$

in case the temporal coordinate amounts to the last coordinate in the coordinate system of the spatiotemporal space. The parameters $\sigma_{max} > \sigma_{min} \geq 0$ are preset constants corresponding to the maximum and minimum widths of the kernel.

The summation function is then constructed using

$$S(\hat{x}) = G(\Sigma\hat{x})$$

i.e. without any rotation. The letter \hat{x} represents the coordinates of a point in a local coordinate system in a neighborhood N , centered on the currently processed pixel \bar{x} .

Furthermore this embodiment comprises calculating a discrete summation kernel \bar{S} by discretizing the summation function S according to the corresponding procedure in calculating 12 step.

The method further comprises calculating a weighted intensity value for each pixel in the first image dataset, using

$$O(\bar{x}) = \sum_{\hat{x} \in N} \bar{S}(\hat{x}) I(\bar{x} + \hat{x})$$

wherein $O(\bar{x})$ defines the intensity in pixel \bar{x} in the first noise reduced output image dataset, \bar{S} is the summation calculated for pixel \bar{x} using the procedure detailed above and I is the input image. Since the same summation kernel \bar{S} is used for each pixel \bar{x} in the entire currently processed image dataset I the summation is easily and effectively performed using linear convolution.

After performing foreground-background segmentation

the classified foreground pixels are processed in the calculation 11 step, the calculation 12 step, and the calculation 13 step. Additionally, the calculation 100 step for sharpening and the calculation step 110 for adapting the kernel to the local noise level may be included in the processing, as illustrated in Fig 6. In an embodiment of the present invention, the storing 14 step comprises storing the entire or almost the entire processed image from the summation 63 as the noise reduced output image and then for every foreground pixel processed in the calculation 11 step, the calculation 12 step, and the calculation 13 step, store the output value in the corresponding pixel in the noise reduced first output image dataset. This means that the summation in the background pixels may be performed by the performing 64 step while the summation in the foreground pixels is performed by the calculation 11 step, the calculation 12 step, and the calculation 13 step. The advantage of this embodiment is inter alia that the background pixels, which most commonly constitute the larger area in a static image dataset sequence, are not put into the calculation 11 step, the calculation 12 step, and the calculation 13 step, which reduces the processing time of the method significantly without decreasing the resolution of the resulting noise reduced output image dataset.

In an embodiment of the present invention according to Fig. 7, an apparatus 70 for e.g. noise reduction in a sequence of image datasets is provided, for example in a sequence of night vision image datasets, wherein each image dataset has at least two pixels, wherein each pixel has an intensity value. The apparatus comprises

a first calculation unit 71 for calculating a measure of the intensity structure for each pixel in a first image

dataset comprised in the sequence of image datasets,

a second calculation unit 72 for calculating a first summation kernel \bar{S} based on the measure of the intensity structure in each pixel in the first image dataset,

5 a third calculation unit 73 for calculating a weighted intensity value for each pixel in the first image dataset, using as the weight each corresponding pixel in the first summation kernel \bar{S} , and

10 a storage unit 74 for storing the weighted intensity value for each pixel in the first image dataset as the intensity value for each corresponding pixel in a noise reduced first output image dataset.

In an embodiment of the present invention, the apparatus further comprises a setting unit 75 for setting 15 an image dataset of the sequence subsequent to said first image dataset as said first image dataset.

In an embodiment of the present invention, the apparatus further comprises a repeating unit 76 for 20 repeating the tasks of the first calculation unit 71, the second calculation unit 72, the third calculation unit 73 and the storage unit 74 until all image datasets of the sequence of image datasets have been processed.

In an embodiment of the present invention the apparatus comprises a first memory 771 capable of storing 25 and deleting a sequence of image datasets or individual image datasets.

In some embodiments of the present invention, when an image dataset is no longer needed for the forthcoming processing from the first memory, e.g. when the image 30 dataset precedes the currently processed image dataset by more than n_b image datasets, the image dataset may be deleted from the first memory 771. The processing of each

image dataset thus requires maximum n_f image datasets succeeding and n_b image datasets preceding the currently processes image dataset.

In an embodiment of the present invention the
5 apparatus further comprises an image stabilization unit 772 for performing image stabilization if motion disturbance is present in the image dataset.

In an embodiment of the present invention the
apparatus further comprises an intensity transformation
10 unit 773 for performing intensity transformation 7731, 7732, 7733 of the currently processed image dataset.

In an embodiment of the present invention the
apparatus further comprises render unit 774 for rendering
the created output image dataset for visualization on a
15 display 775.

The unit(s) may e.g. be in hardware, such as a processor with a memory. The processor could be any of variety of processors, such as Intel or AMD processors, CPUs, microprocessors, Programmable Intelligent Computer
20 (PIC) microcontrollers, Digital Signal Processors (DSP), etc. However, the scope of the invention is not limited to these specific processors. The memory may be any memory capable of storing information, such as Random Access Memories (RAM) such as, Double Density RAM (DDR, DDR2),
25 Single Density RAM (SDRAM), Static RAM (SRAM), Dynamic RAM (DRAM), Video RAM (VRAM), etc. The memory may also be a FLASH memory such as a USB, Compact Flash, SmartMedia, MMC memory, MemoryStick, SD Card, MiniSD, MicroSD, xD Card, TransFlash, and MicroDrive memory etc. However, the scope
30 of the invention is not limited to these specific memories.

In an embodiment of the present invention the
apparatus 70 is used in a vehicle for detection of moving

objects around the vehicle by means of a camera, in particular e.g. by means of a night vision camera. As an example the camera may be placed in the forward direction of the vehicle to detect objects located in the vehicle path. The object may be animals or any moving or static
5 object, which is difficult to discover using the human eye in dim light conditions such as during night vision. The camera is connected to the apparatus and a display for presentation. The great advantage of this embodiment is
10 that animals or people present in front of the vehicle may be detected before the driver or passengers have detected them visually.

This embodiment could also be used as a part of a system for automatic detection of different objects such as
15 pedestrians, animals, traffic signs, line markers etc using automatic image analysis techniques. The noise reduction method would then work as a pre-processing step for the succeeding automatic image analysis, making this analysis much more robust, efficient and flexible.

20 In an embodiment of the present invention the apparatus 70 comprises units arranged for performing all of the method steps defined in any one of the embodiments. For example, in an embodiment of the present invention a calculation unit for sharpening 700 and/or a calculation
25 unit for adapting the kernel 710 to the local noise level may be included or may be incorporated in one of the calculation units.

In an embodiment of the present invention, according to Fig. 8, a computer-readable medium 80 is provided having
30 embodied thereon a computer program for processing by a computer for e.g. noise reduction in a sequence of image datasets, wherein each image dataset has at least two pixels, wherein each pixel has an intensity value. The

computer program comprises

a first calculation code segment 81 for calculating, based on estimated intensity derivatives in each pixel, a measure of the intensity structure, e.g. a structure
5 tensor, for each pixel in a first image dataset comprised in the sequence of image datasets, and for performing an eigen-decomposition of said measure of the intensity structure, e.g. structure tensor,

a second calculation code segment 82 for calculating
10 a first summation kernel \bar{S} based on the measure of the intensity structure, e.g. structure tensor, in each pixel in the first image dataset,

a third calculation code segment 83 for calculating a weighted intensity value for each pixel in the first image
15 dataset, using as the weight each corresponding pixel in the first summation kernel \bar{S} ,

a storage code segment 84 for storing the weighted intensity value for each pixel in the first image dataset as the intensity value for each corresponding pixel in a
20 noise reduced first output image dataset.

In embodiments of the present invention, the computer program comprises a further calculation code segment for rotating a local coordinate system in which the summation kernel is described resulting in that the coordinate axes
25 of said local coordinate system coincide with the directions of the eigenvectors of said structure tensor, where said eigenvectors are described in the global coordinate system of the image dataset, and

a scaling code segment for scaling the coordinate
30 axes of the local coordinate system in which the summation kernel is described by an amount related to the eigenvalues of the structure tensor via a width function $w(\lambda_i) = \sigma_i$, and

wherein said eigenvalues depend on the amount of intensity variation in the direction of their corresponding eigenvectors,

the width function being a decreasing function
 5 depending on the noise level in each pixel $w(0) = \sigma_{max}$ and $\lim_{a \rightarrow \infty} w(a) = \sigma_{min}$. As an example, the width function may be defined by

$$w(\lambda_i, x_0) = \begin{cases} \Delta\sigma e^{-\frac{\lambda_i}{d(x_0)} + \frac{1}{2}} + \sigma_{min}, & \lambda_i > \frac{d(x_0)}{2} \\ \sigma_{max}, & \lambda_i \leq \frac{d(x_0)}{2} \end{cases}$$

In an embodiment of the present invention, the
 10 computer program further comprises a sharpening code segment 800, for constructing summation kernels \bar{S}_H or \bar{S}_h that provides sharpening at the most important spatial and temporal edges together with smoothing in other appropriate directions.

15 In an embodiment of the present invention, the computer program further comprises a calculation code segment 810 for adapting the kernel to the local noise level may be included in the processing, as illustrated in Fig 8.

20 In an embodiment of the present invention, the computer program further comprises a setting code segment 85 for setting an image dataset of the sequence subsequent to said first image dataset as said first image dataset.

In an embodiment of the present invention, the
 25 computer program further comprises a repeating code segment 86 for repeating the tasks of the first calculation code segment 81, second calculation code segment 82, third calculation code segment 83 and storage code segment 84 until all image datasets of the sequence of image datasets
 30 have been processed.

In an embodiment of the present invention, the computer-readable medium further comprises an image stabilization code segment 871 for performing image stabilization if motion disturbance is present in the image dataset data.

In an embodiment of the present invention, the computer-readable medium further comprises an intensity transformation code segment 872 for performing intensity transformation 8721, 8722, 8723 of the currently processed image dataset.

In an embodiment of the present invention, the computer-readable medium further comprises render code segment 873 for rendering the created output image dataset for visualization on a display.

In an embodiment of the present invention, the computer-readable medium comprises code segments arranged, when run by an apparatus having computer processing properties, for performing all of the method steps defined in any one of the embodiments.

The invention may be implemented in any suitable form including hardware, software, firmware or any combination of hardware, software and firmware. The elements and components of an embodiment of the invention may be physically, functionally and logically implemented in any suitable way. Indeed, the functionality may be implemented in a single unit, in a plurality of units or as part of other functional units. As such, the invention may be implemented in a single unit, or may be physically and functionally distributed between different units and processors.

Combinations and modifications of the above-mentioned embodiments of the present invention should be able to be implemented by a person ordinary skilled in the art to

which this invention belongs.

Applications and use of the above-described embodiments according to the invention are various and include all fields wherein dynamic enhancement, e.g. noise
5 reduction, of an image dataset comprising low contrast areas is used. A noisy low contrast image dataset is e.g. obtained using a video camera or a still image camera in low-light conditions such as e.g. night vision conditions and/or in foggy, hazy, dusty and snowy conditions. The
10 camera could for example be a consumer camera obtaining images for private use, a camera used for surveillance purposes, an imaging device used for industrial inspection purposes, an imaging device used for navigation of different kinds of vehicles or an imaging device used for
15 different kinds of industrial and scientific measurements. In particular the camera could be a night vision camera. The above-mentioned imaging devices do not only comprise 2D cameras, but also microscopic 2D or 3D imaging devices, such as fluorescence microscopes where the resulting images
20 generally are of low-contrast and very noisy or medical imaging device such as different kinds of scanning devices or tomographic devices such as MR imaging, computer tomography and x-ray imaging. The imaging devices may also comprise measuring equipment, which outputs multi-channel
25 image datasets.

Although the present invention has been described above with reference to specific embodiments, it is not intended to be limited to the specific form set forth herein. Rather, the invention is limited only by the
30 accompanying claims.

Some embodiments of the present invention may allow for combining an intensity transformation stage and a noise reduction stage. In the noise reduction stage an optimal or

almost optimal spatiotemporal summation kernel may be found which may be used to calculate a weighted average (spatiotemporal summation) around each point. The shape of this summation kernel in general depends on the general
5 intensity structure of the spatiotemporal space, without reference to which axis being the temporal axis and which axes being the spatial axes. The spatiotemporal intensity structure may be made up of the spatial structures in the scene and their motion in the image sequence.

10 An advantage of some embodiments of the present invention is that it, using a generally applicable methodology, may find an optimal or almost optimal summation kernel in the spatiotemporal space that may be used to calculate a weighted intensity value for each pixel
15 in the image dataset. The calculation of this weighted intensity value may effectively reduce the noise and minimize the amount of spatial blur and motion blur and even sharpen some contours and edges in the image sequence, i.e. maximize the resolution in the output.

20 Another advantage with some embodiments of the present invention is that it is possible to reduce the noise in image datasets comprising low contrast, e.g. image datasets taken during very dim light conditions, especially during night vision conditions, thereby making it possible
25 to improve the visibility in such image datasets.

Yet another advantage of some embodiments of the present invention is that the true colors of the objects in the image datasets can be preserved and the output image dataset thus contains all or almost all colors of the
30 objects in the image dataset, as embodiments of the present invention may process multi-channel image data, such as triple channel RGB color data. The image dataset or sequence of image datasets may thus be taken by an ordinary

camera detecting white light (400nm-700nm) in dim light conditions, e.g. night vision conditions, and thus have a clear advantage over IR detectors, which only detect light in the infrared interval above 700nm, and which are
5 currently used for dim light conditions such as night vision conditions.

In this specification and claims, the term "comprises/comprising" does not exclude the presence of other elements or steps. Furthermore, although individually
10 listed, a plurality of means, elements or method steps may be implemented by e.g. a single unit or processor. Additionally, although individual features may be included in different claims, these may possibly advantageously be combined, and the inclusion in different claims does not
15 imply that a combination of features is not feasible and/or advantageous. In addition, singular references do not exclude a plurality. The terms "a", "an", "first", "second" etc do not preclude a plurality. Reference signs in the claims are provided merely as a clarifying example and
20 shall not be construed as limiting the scope of the claims in any way.

CLAIMS

1. A method for processing a sequence of night vision image datasets, wherein said sequence comprises at least two image datasets each having at least two pixels, wherein each pixel has an intensity value, said method comprising:

5 each pixel has an intensity value, said method comprising:
calculating (11), based on estimated intensity derivatives in each pixel, a structure tensor for each pixel in an image dataset comprised in said sequence of image datasets, and performing an eigen-decomposition of
10 said structure tensor,
calculating (12) values in a summation kernel based on said structure tensor for each pixel in said image dataset,
calculating (13) a weighted intensity value for each
15 pixel in said image dataset, using as weights the values in said summation kernel,
storing (14) said weighted intensity value for each pixel in said image dataset as a processed intensity value for each corresponding pixel in a processed output image
20 dataset,
the method furthermore comprising:
rotating a local coordinate system in which the summation kernel is described resulting in that the coordinate axes of said local coordinate system coincide
25 with the directions of the eigenvectors of said structure tensor, where said eigenvectors are described in the global coordinate system of the image dataset, and
scaling the coordinate axes of the local coordinate system in which the summation kernel is described by an
30 amount related to the eigenvalues of the structure tensor via a width function $w(\lambda_i) = \sigma_i$, and wherein said eigenvalues depend on the amount of intensity variation in the direction of their corresponding eigenvectors,

the width function being a decreasing function depending on the noise level in each pixel such that $w(0) = \sigma_{\max}$ and $\lim_{a \rightarrow \infty} w(a) = \sigma_{\min}$.

5 2. The method of claim 1 wherein the width function is defined by

$$w(\lambda_i, x_0) = \begin{cases} \Delta\sigma e^{-\frac{\lambda_i}{d(x_0)} + \frac{1}{2}} + \sigma_{\min}, & \lambda_i > \frac{d(x_0)}{2} \\ \sigma_{\max}, & \lambda_i \leq \frac{d(x_0)}{2} \end{cases}.$$

10 3. The method of any of the previous claims, wherein the width function is adapted to a noise level of the image datasets.

15 4. The method according to claim 3, wherein adapting the width function to the noise level comprises estimating the noise in the image datasets and scaling an estimation of the variance of the noise so that the result is of a magnitude that is suitable for insertion in the width function.

20 5. The method according to claim 4, wherein scaling an estimation of the variance is performed by means of a fixed scaling factor.

25 6. The method according to claim 4, wherein scaling an estimation of the variance is performed by means of a scaling factor which is determined from the image dataset.

30 7. The method according to any of the previous claims, comprising:
repeating the method steps for subsequent image datasets.

8. The method according to any of the previous claims, wherein said sequence of image datasets comprises 2D, 3D, distance image or higher-dimensional monochrome or
5 color image datasets.

9. The method according to any of the previous claims, wherein said step of calculating (11) a structure tensor comprises:

10 calculating a structure tensor based on the mean of the corresponding estimated intensity derivatives for each channel in each pixel when said image dataset is a multi-channel image dataset.

15 10. The method according to any of the previous claims, further comprising:

calculating said structure tensor for each channel in said image dataset when said image dataset is a multi-channel image dataset.

20

11. The method according to any of the previous claims, further comprising:

calculating the elementwise mean structure tensor of said structure tensors calculated for each channel in said
25 image dataset when said image dataset is a multi-channel image dataset.

12. The method according to any of the previous claims, further comprising:

30 calculating said structure tensor using the mean value of the values in each channel in said dataset when said image dataset is a multi-channel image data set.

13. The method according to any of the previous claims, further comprising:

performing low-pass filtering of said image dataset before or during said step of calculating the structure
5 tensor.

14. The method according to any of the previous claims, wherein said step of calculating (11) of the structure tensor further comprises:

10 calculating the motion vectors for each pixel of each image dataset of said sequence of image datasets.

15. The method according to any of the previous claims, further comprising:

15 performing an intensity transformation on said image dataset by an intensity transformation function to increase the intensity contrast in at least a part of said first image dataset.

20 16. The method according to any of the previous claims, further comprising:

rendering (15) each of said first output image dataset for presentation on a display.

25 17. The method according to any of previous claims, further comprising performing a foreground-background segmentation.

18. The method according to any of the previous
30 claims, furthermore comprising sharpening of the image.

19. The method according to claim 18, wherein the sharpening is obtained by convolution of the summation

kernel with a high-boost filter giving the summation kernel \bar{S}_H .

20. The method according to claim 18, wherein the
5 sharpening is obtained by using a summation function S_h
obtained by approximating a discrete high-boost filter with
a continuous function h and calculating a new summation
function analytically by the convolution of the original
summation kernel S with the function h , and discretizing
10 the summation function S_h , giving the summation kernel \bar{S}_h .

21. An apparatus (70) for processing a sequence of
night vision image datasets, said sequence comprising at
least two image datasets each having at least two pixels,
15 wherein each pixel has an intensity value, said apparatus
comprising:

a first calculation unit (71) for calculating, based
on estimated intensity derivatives in each pixel, a
structure tensor for each pixel in an image dataset
20 comprised in said sequence of image datasets, and
performing an eigen-decomposition of said structure tensor,

a second calculation unit (72) for calculating values
in a summation kernel based on said structure tensor for
each pixel in said image dataset,

25 a third calculation unit (73) for calculating a
weighted intensity value for each pixel in said image
dataset, using as weights the values in said summation
kernel,

a storage unit (74) for storing said weighted
30 intensity value for each pixel in said image dataset as a
processed intensity value for each corresponding pixel in a
processed output image dataset,

wherein the apparatus further comprises:

a calculation unit for rotating a local coordinate system in which the summation kernel is described resulting in that the coordinate axes of said local coordinate system coincide with the directions of the eigenvectors of said structure tensor, where said eigenvectors are described in
 5 the global coordinate system of the image dataset, and

a scaling unit for scaling the coordinate axes of the local coordinate system in which the summation kernel is described by an amount related to the eigenvalues of the structure tensor via a width function $w(\lambda_i) = \sigma_i$, and wherein
 10 said eigenvalues depend on the amount of intensity variation in the direction of their corresponding eigenvectors,

the width function being a decreasing function such
 15 that $w(0) = \sigma_{\max}$ and $\lim_{a \rightarrow \infty} w(a) = \sigma_{\min}$.

22. The apparatus of Claim 21, the width function being a decreasing function defined by

$$w(\lambda_i, x_0) = \begin{cases} \Delta\sigma e^{-\frac{\lambda_i}{d(x_0)} + \frac{1}{2}} + \sigma_{\min}, & \lambda_i > \frac{d(x_0)}{2} \\ \sigma_{\max}, & \lambda_i \leq \frac{d(x_0)}{2} \end{cases}.$$

20

23. The apparatus according to any of claims 21 or 22, further comprising a repeating unit (76) for repeating the above steps for subsequent image datasets.

24. A computer-readable medium (80) having embodied thereon a computer program for processing by a processor of a sequence of image datasets, wherein said sequence comprises at least two image datasets, each image dataset having at least two pixels, wherein each pixel has an
 30 intensity value, said computer program comprising

a first calculation code segment (81) for calculating, based on estimated intensity derivatives in each pixel, a structure tensor for each pixel in an image dataset comprised in said sequence of image datasets, and
5 for performing an eigen-decomposition of said structure tensor,

a second calculation code segment (82) for calculating values in a summation kernel based on said structure tensor for each pixel in said image dataset,

10 a third calculation code segment (83) for calculating a weighted intensity value for each pixel in said image dataset, using as weights the values in said summation kernel,

a storage code segment (84) for storing said weighted
15 intensity value for each pixel in said image dataset as a processed intensity value for each corresponding pixel in a processed output image dataset,

a further calculation code segment for rotating a local coordinate system in which the summation kernel is
20 described resulting in that the coordinate axes of said local coordinate system coincide with the directions of the eigenvectors of said structure tensor, where said eigenvectors are described in the global coordinate system of the image dataset, and

25 a scaling code segment for scaling the coordinate axes of the local coordinate system in which the summation kernel is described by an amount related to the eigenvalues of the structure tensor via a width function $w(\lambda_i) = \sigma_i$, and wherein said eigenvalues depend on the amount of intensity
30 variation in the direction of their corresponding eigenvectors,

the width function being a decreasing function such

that $w(0) = \sigma_{\max}$ and $\lim_{a \rightarrow \infty} w(a) = \sigma_{\min}$.

25. A computer-readable medium according to claim 24, the width function being a decreasing function defined by

5

$$w(\lambda_i, x_0) = \begin{cases} \Delta\sigma e^{-\frac{\lambda_i}{d(x_0)} + \frac{1}{2}} + \sigma_{\min}, & \lambda_i > \frac{d(x_0)}{2} \\ \sigma_{\max}, & \lambda_i \leq \frac{d(x_0)}{2} \end{cases} .$$

26. The computer-readable medium according to any of claims 24 or 25, further comprising a repeating code
10 segment (86) for repeating the above steps for subsequent image datasets.

27. A method for processing a sequence of night vision image datasets, wherein said sequence comprises at
15 least two image datasets each having at least two pixels, wherein each pixel has an intensity value, said method comprising:

calculating (11), based on estimated intensity derivatives in each pixel, a structure tensor for each
20 pixel in an image dataset comprised in said sequence of image datasets,

calculating (12) values in a summation kernel based on said structure tensor for each pixel in said image dataset,

25 calculating (13) a weighted intensity value for each pixel in said image dataset, using as weights the values in said summation kernel,

storing (14) said weighted intensity value for each pixel in said image dataset as a processed intensity value
30 for each corresponding pixel in a processed output image dataset,

wherein the method furthermore comprises calculating values in the summation kernel so that, when applying the summation kernel to an input image, sharpening of the input image is obtained at the most important spatial and
5 temporal edges while smoothing is applied in the appropriate directions in the input image data.

28. A method according to claim 27, wherein performing sharpening while smoothing is obtained by using
10 a kernel \bar{S}_H obtained by convolving an anisotropic summation filter kernel with a high-boost filter.

29. A method according to claim 27, wherein performing sharpening during smoothing is obtained by using
15 a kernel \bar{S}_h obtained by analytically convolving a continuous approximation of a discrete high-boost filter with an anisotropic summation filter function and discretizing the resulting summation filter function.

20 30. An apparatus (70) for processing a sequence of night vision image datasets, said sequence comprising at least two image datasets each having at least two pixels, wherein each pixel has an intensity value, said apparatus comprising:

25 a first calculation unit (71) for calculating, based on estimated intensity derivatives in each pixel, a structure tensor for each pixel in an image dataset comprised in said sequence of image datasets, and performing an eigen-decomposition of said structure tensor,

30 a second calculation unit (72) for calculating values in a summation kernel based on said structure tensor for each pixel in said image dataset,

a third calculation unit (73) for calculating a

weighted intensity value for each pixel in said image dataset, using as weights the values in said summation kernel,

a storage unit (74) for storing said weighted
5 intensity value for each pixel in said image dataset as a processed intensity value for each corresponding pixel in a processed output image dataset,

wherein the second calculation unit is adapted for calculating values in the summation kernel so that, when
10 applying the summation kernel to an input image, sharpening of the input image is obtained at the most important spatial and temporal edges while smoothing is applied in the appropriate directions in the input image data.

15 31. An apparatus according to claim 30, wherein the second calculation unit is adapted for performing sharpening by using a kernel \bar{S}_H obtained by convolving an anisotropic summation filter kernel with a high-boost filter.

20

32. An apparatus according to claim 30, wherein the second calculation unit is adapted for performing
sharpening by using a kernel \bar{S}_h obtained by analytically convolving a continuous approximation of a discrete high-
25 boost filter with an anisotropic summation filter function and discretizing the resulting summation filter function.

33. A method for processing a sequence of night vision image datasets, wherein said sequence comprises at
30 least two image datasets each having at least two pixels, wherein each pixel has an intensity value, said method comprising:

calculating (11), based on estimated intensity

derivatives in each pixel, a structure tensor for each pixel in an image dataset comprised in said sequence of image datasets, and performing an eigen-decomposition of said structure tensor,

5 calculating (12) values in a summation kernel based on said structure tensor for each pixel in said image dataset,

 calculating (13) a weighted intensity value for each pixel in said image dataset, using as weights the values in
10 said summation kernel,

 storing (14) said weighted intensity value for each pixel in said image dataset as a processed intensity value for each corresponding pixel in a processed output image dataset,

15 the method furthermore comprising:

 rotating a local coordinate system in which the summation kernel is described resulting in that the coordinate axes of said local coordinate system coincide with the directions of the eigenvectors of said structure
20 tensor, where said eigenvectors are described in the global coordinate system of the image dataset, and

 scaling the coordinate axes of the local coordinate system in which the summation kernel is described by an amount related to the eigenvalues of the structure tensor
25 via a width function $w(\lambda_i) = \sigma_i$, and wherein said eigenvalues depend on the amount of intensity variation in the direction of their corresponding eigenvectors,

 and wherein the width function is adapted to a noise level of the image datasets.

30

34. The method according to claim 33, wherein adapting the width function to the noise level comprises estimating the noise in the image datasets and scaling an

estimation of the variance of the noise so that the result is of a magnitude that is suitable for insertion in the width function.

5 35. The method according to claim 34, wherein scaling an estimation of the variance is performed by means of a fixed scaling factor.

10 36. The method according to claim 34, wherein scaling an estimation of the variance is performed by means of a scaling factor which is determined from the image dataset.

15 37. An apparatus (70) for processing a sequence of night vision image datasets, said sequence comprising at least two image datasets each having at least two pixels, wherein each pixel has an intensity value, said apparatus comprising:

20 a first calculation unit (71) for calculating, based on estimated intensity derivatives in each pixel, a structure tensor for each pixel in an image dataset comprised in said sequence of image datasets, and performing an eigen-decomposition of said structure tensor,

25 a second calculation unit (72) for calculating values in a summation kernel based on said structure tensor for each pixel in said image dataset,

 a third calculation unit (73) for calculating a weighted intensity value for each pixel in said image dataset, using as weights the values in said summation kernel,

30 a storage unit (74) for storing said weighted intensity value for each pixel in said image dataset as a processed intensity value for each corresponding pixel in a processed output image dataset,

wherein the apparatus further comprises:

a calculation unit for rotating a local coordinate system in which the summation kernel is described resulting in that the coordinate axes of said local coordinate system
5 coincide with the directions of the eigenvectors of said structure tensor, where said eigenvectors are described in the global coordinate system of the image dataset, and

a scaling unit for scaling the coordinate axes of the local coordinate system in which the summation kernel is
10 described by an amount related to the eigenvalues of the structure tensor via a width function $w(\lambda_i) = \sigma_i$, and wherein said eigenvalues depend on the amount of intensity variation in the direction of their corresponding eigenvectors, and wherein the width function is adapted to
15 a noise level of the image datasets.

38. A computer-readable medium (80) having embodied thereon a computer program for processing by a processor of
20 a sequence of image datasets, wherein said sequence comprises at least two image datasets, each image dataset having at least two pixels, wherein each pixel has an intensity value, said computer program comprising

a first calculation code segment (81) for
25 calculating, based on estimated intensity derivatives in each pixel, a structure tensor for each pixel in an image dataset comprised in said sequence of image datasets, and for performing an eigen-decomposition of said structure tensor,

30 a second calculation code segment (82) for calculating values in a summation kernel based on said structure tensor for each pixel in said image dataset,

a third calculation code segment (83) for calculating

a weighted intensity value for each pixel in said image dataset, using as weights the values in said summation kernel,

5 a storage code segment (84) for storing said weighted intensity value for each pixel in said image dataset as a processed intensity value for each corresponding pixel in a processed output image dataset,

wherein the second calculation code segment is adapted for calculating values in the summation kernel so that, when applying the summation kernel to an input image, sharpening of the input image is obtained at the most important spatial and temporal edges while smoothing is applied in the appropriate directions in the input image data.

15

39. A computer-readable medium (80) according to claim 38, wherein the second calculation code segment comprises code for performing sharpening by using a kernel \bar{S}_H obtained by convolving an anisotropic summation filter kernel with a high-boost filter.

40. A computer-readable medium (80) according to claim 38, wherein the second calculation code segment comprises code for performing sharpening by using a kernel \bar{S}_h obtained by analytically convolving a continuous approximation of a discrete high-boost filter with an anisotropic summation filter function and discretizing the resulting filter function.

41. A computer-readable medium (80) having embodied thereon a computer program for processing by a processor of a sequence of image datasets, wherein said sequence comprises at least two image datasets, each image dataset having at least two pixels, wherein each pixel has an

30

intensity value, said computer program comprising

a first calculation code segment (81) for calculating, based on estimated intensity derivatives in each pixel, a structure tensor for each pixel in an image dataset comprised in said sequence of image datasets, and
5 for performing an eigen-decomposition of said structure tensor,

a second calculation code segment (82) for calculating values in a summation kernel based on said structure tensor
10 for each pixel in said image dataset,

a third calculation code segment (83) for calculating a weighted intensity value for each pixel in said image dataset, using as weights the values in said summation kernel,

15 a storage code segment (84) for storing said weighted intensity value for each pixel in said image dataset as a processed intensity value for each corresponding pixel in a processed output image dataset,

a further calculation code segment for rotating a
20 local coordinate system in which the summation kernel is described resulting in that the coordinate axes of said local coordinate system coincide with the directions of the eigenvectors of said structure tensor, where said eigenvectors are described in the global coordinate system
25 of the image dataset, and

a scaling code segment for scaling the coordinate axes of the local coordinate system in which the summation kernel is described by an amount related to the eigenvalues of the structure tensor via a width function $w(\lambda_i) = \sigma_i$, and
30 wherein said eigenvalues depend on the amount of intensity variation in the direction of their corresponding eigenvectors, and wherein the width function is adapted to a noise level ratio of the image datasets.

42. Use of the method, apparatus, or computer-readable medium according to any one of the previous claims in a vehicle, robot, video camera, surveillance system or a
5 manufacturing system.

43. Use of the method, apparatus, or computer-readable medium according to any one of claims 1 to 41 for a night vision system of a vehicle.
10

44. Use of the method, apparatus, or computer-readable medium according to any one of claims 1 to 41 for detecting and/or analyzing objects in the vicinity of a vehicle.

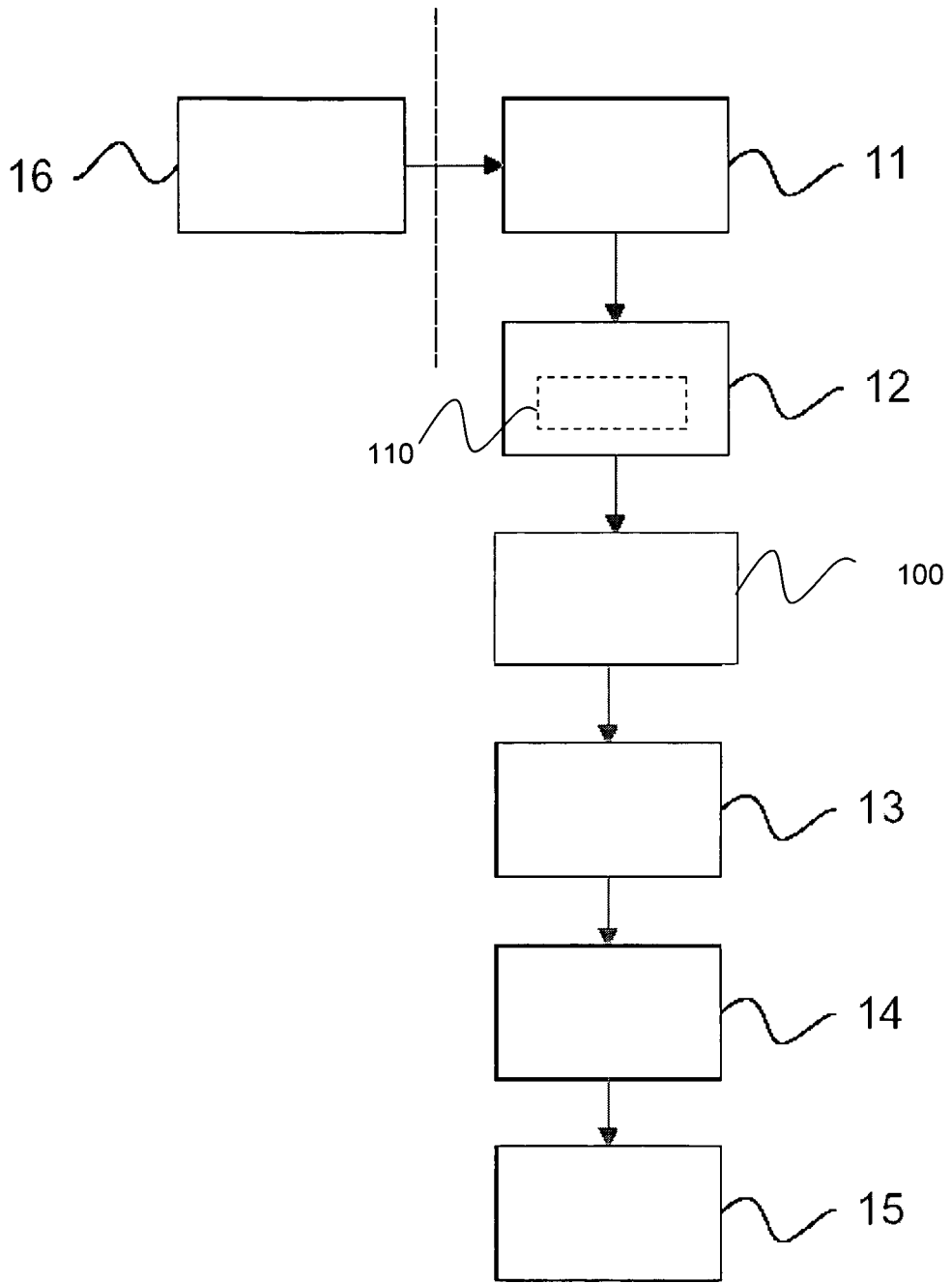


Fig. 1

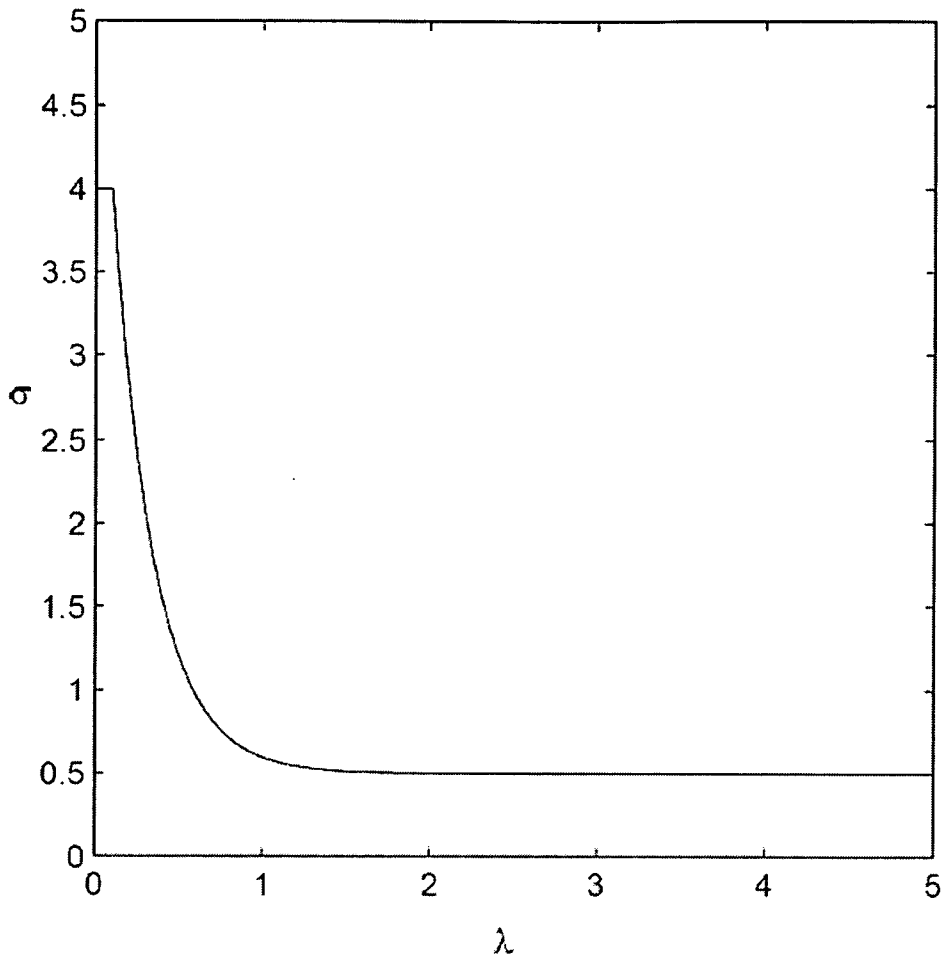


Fig. 2

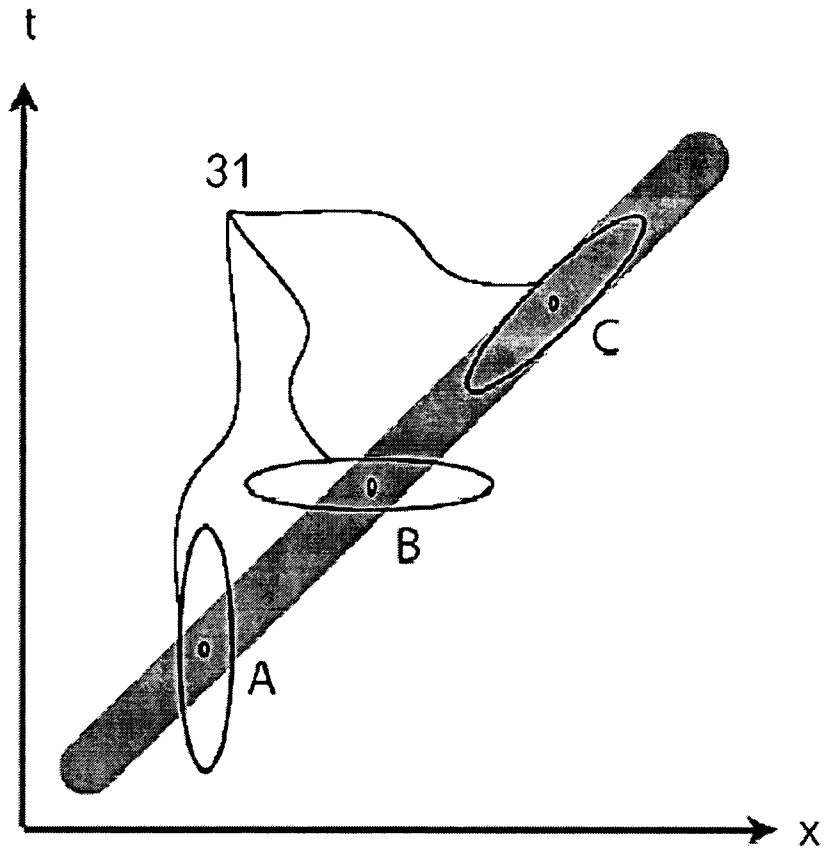


Fig. 3

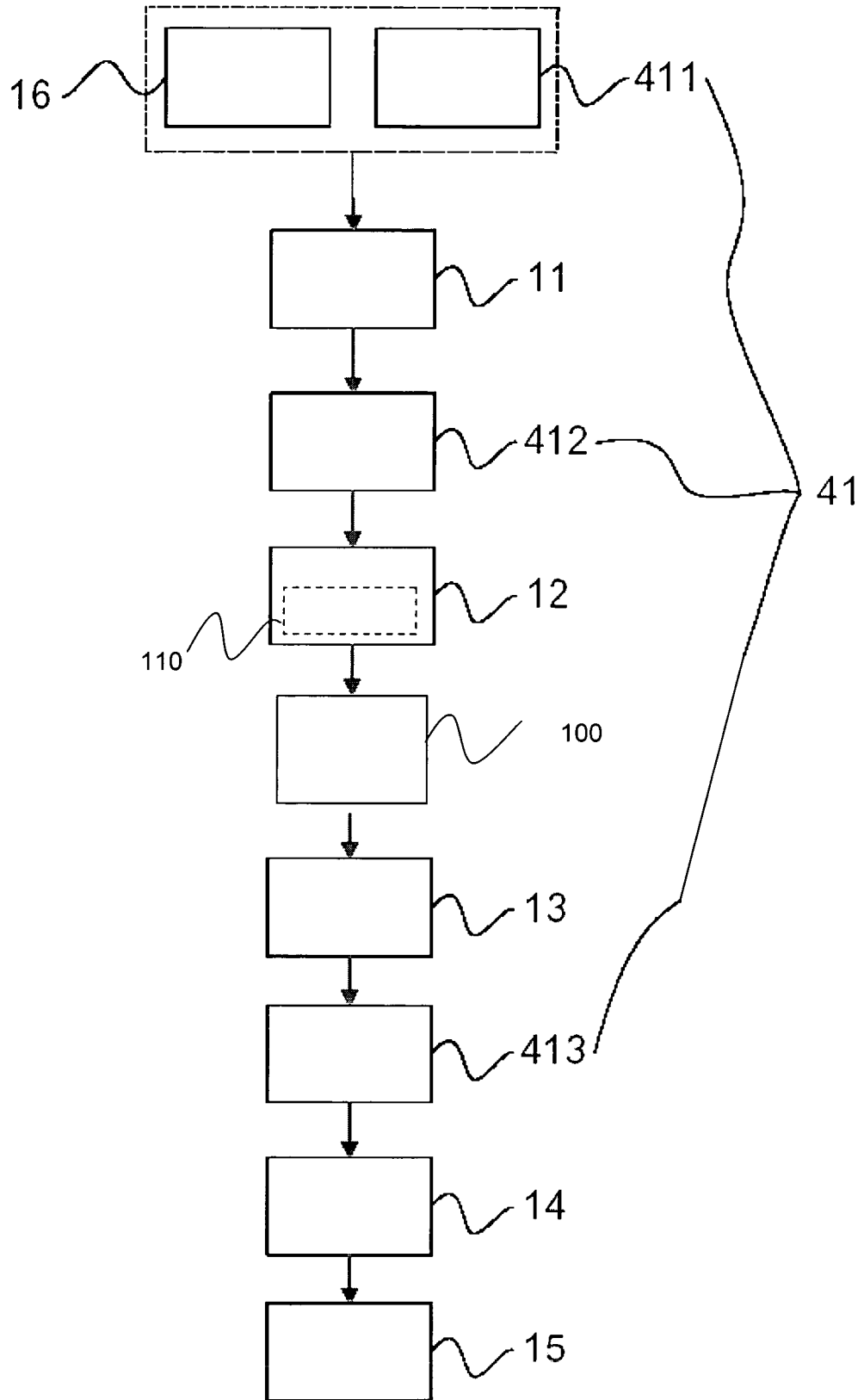


Fig. 4

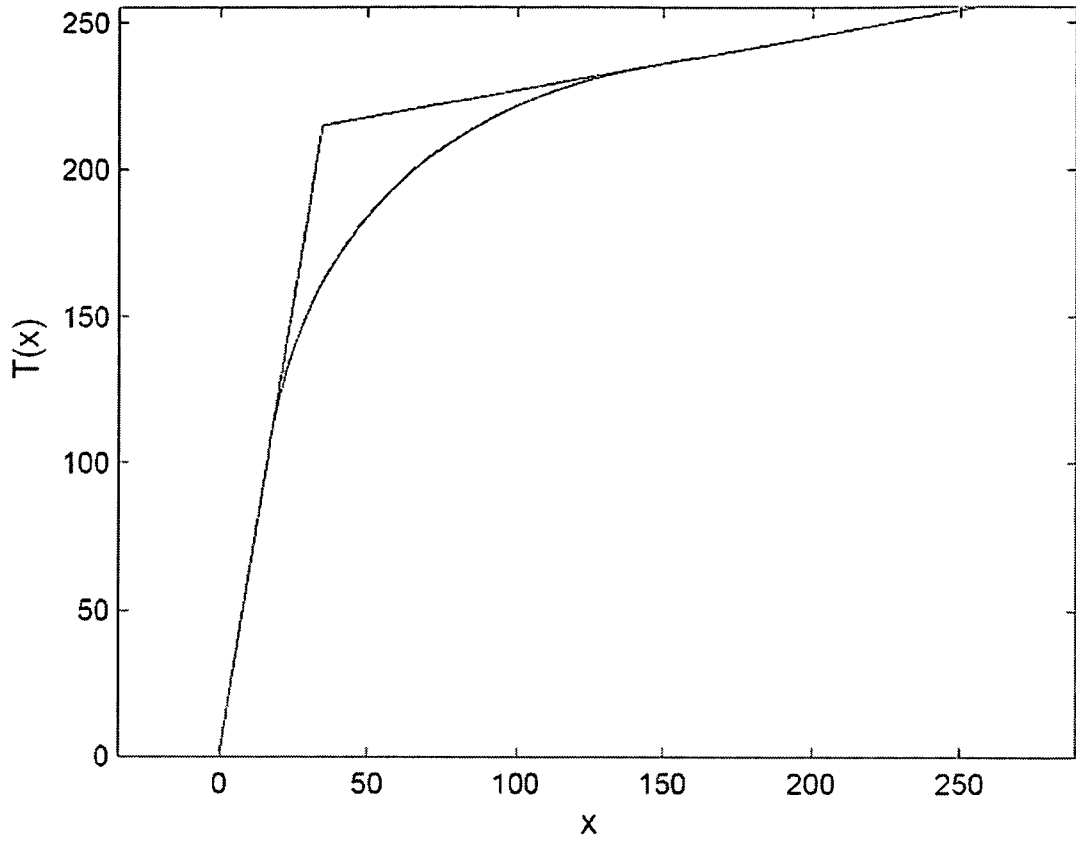


Fig. 5

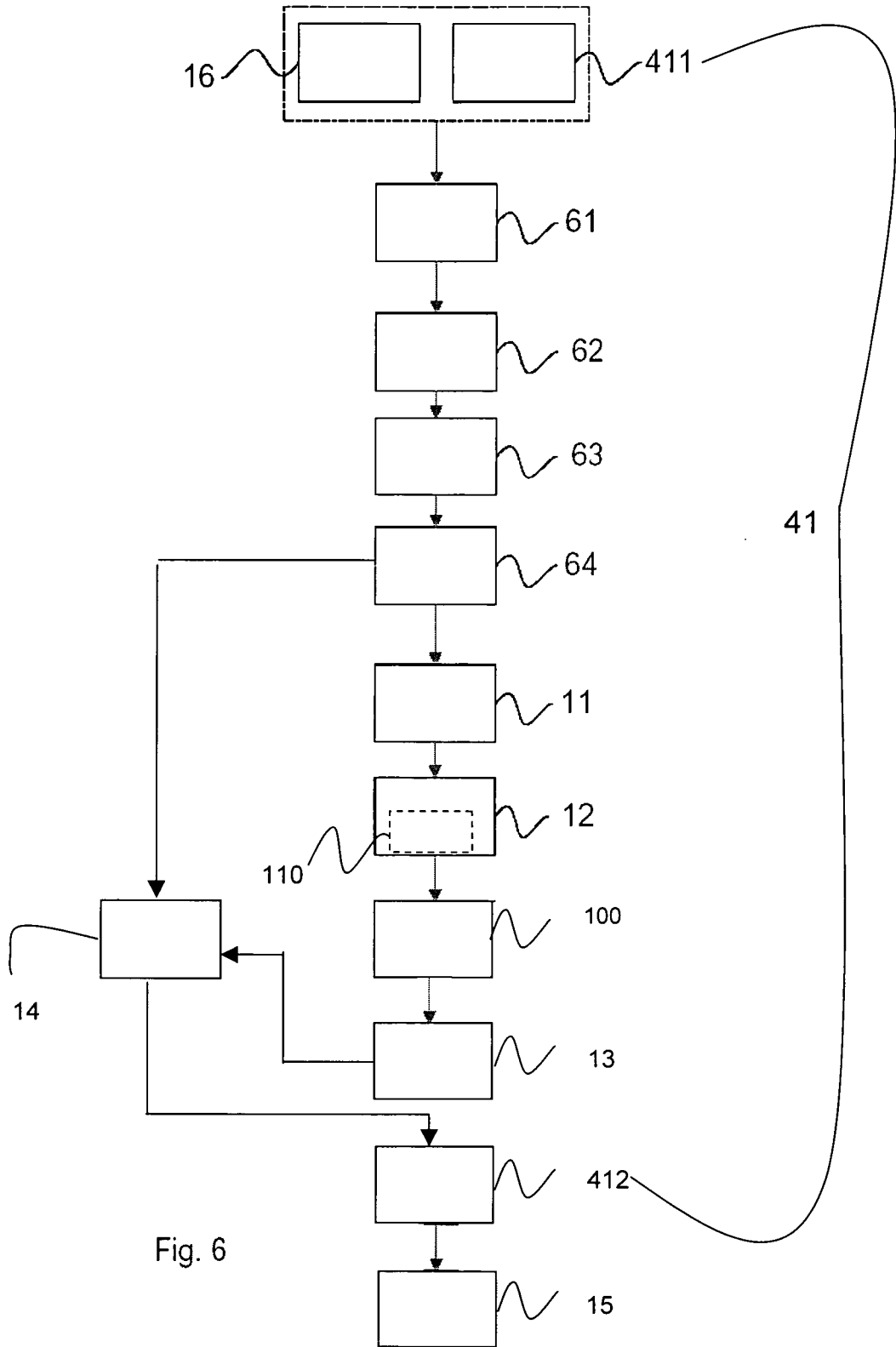


Fig. 6

7/9

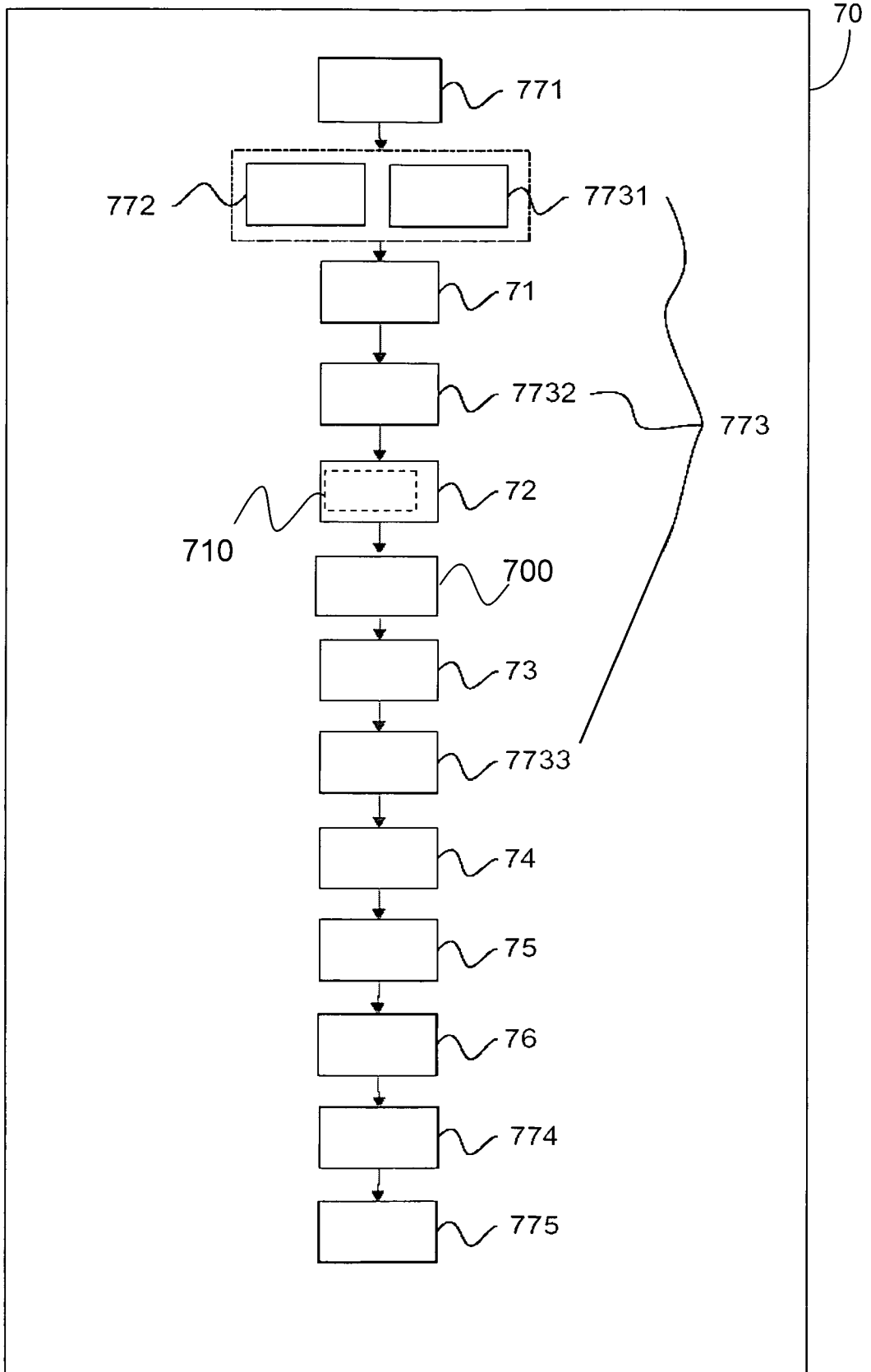


Fig. 7

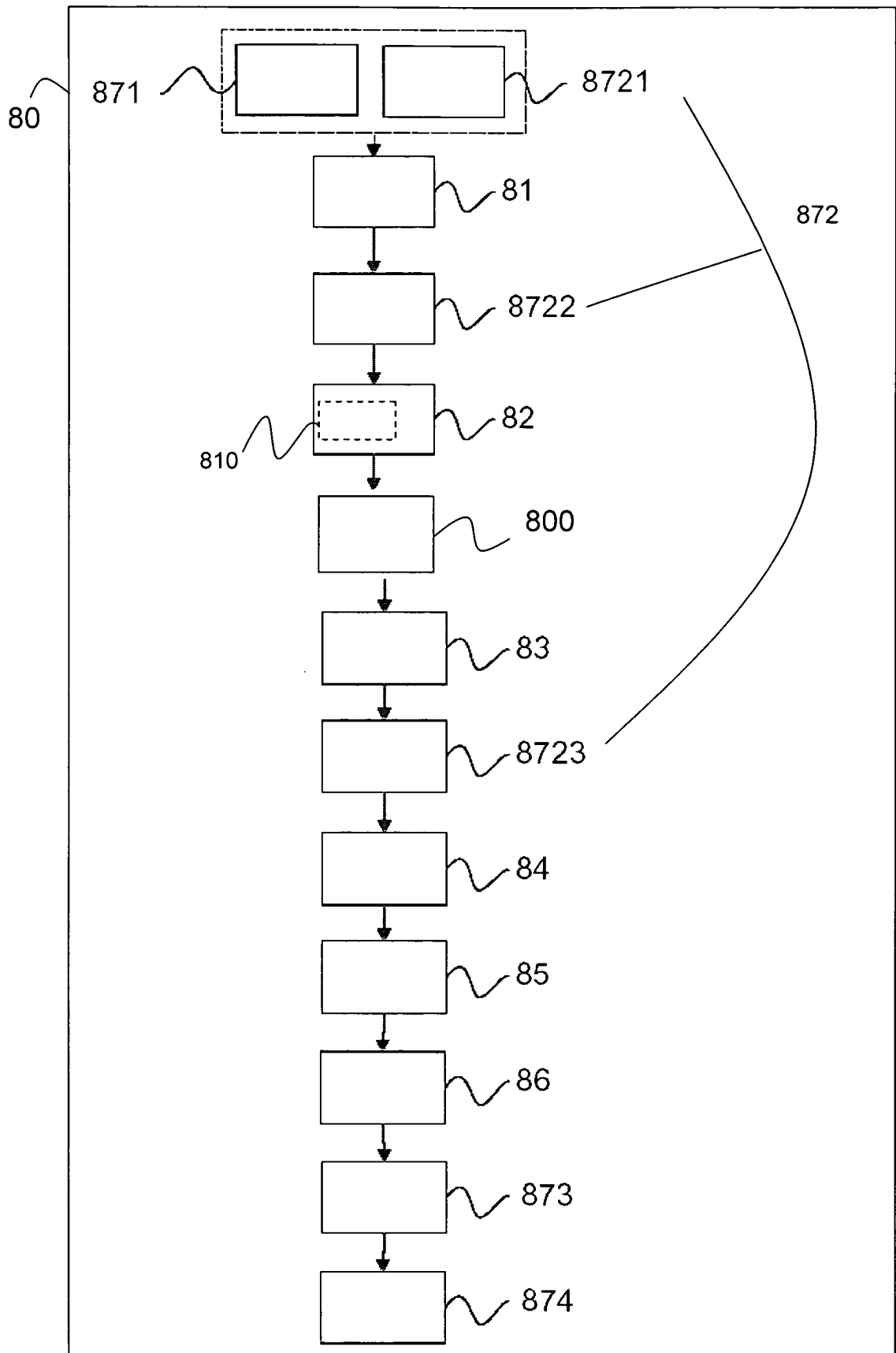


Fig.8

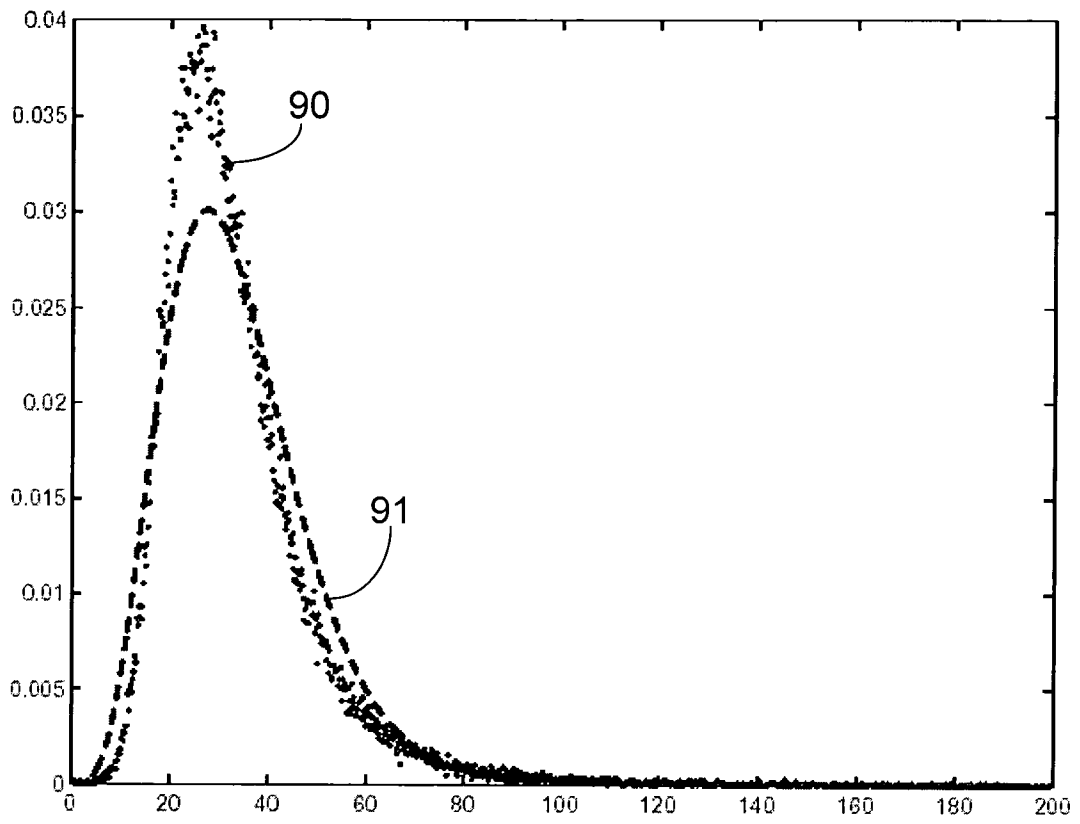


Fig. 9