(19) **United States**

(12) **Patent Application Publication** (10) Pub. No.: **US 2009/0220173 A1**

Wajs (43) **Pub. Date:** **Sep. 3, 2009**

(54) **IMAGE PROCESSING METHOD AND SYSTEM**

(75) Inventor: **Andrew Augustine Wajs**, Haarlem (NL)

Correspondence Address:
NIXON PEABODY, LLP
401 9TH STREET, NW, SUITE 900
WASHINGTON, DC 20004-2128 (US)

(73) Assignee: **ACTIVE OPTICS PTY LIMITED**, Carlingford (AU)

Publication Classification

(57) **ABSTRACT**

In an image processing method a sequence of at least two image frames is provided and a motion vector representing relative movement of at least a part of said image frames is calculated. At least said part in the image frames is repositioned in accordance with the calculated motion vector, and the image frames are combined to form a combined image frame. The image frames are divided into blocks and a block motion vector representing relative movement of each block is calculated. Each block is repositioned in accordance with the calculated block motion vector and the blocks are combined to form combined blocks in order to obtain the combined image frame.

Fig. 1

Fig. 2

Capture Frames —22

Cache Image Frames —23

Alignment and Processing —24
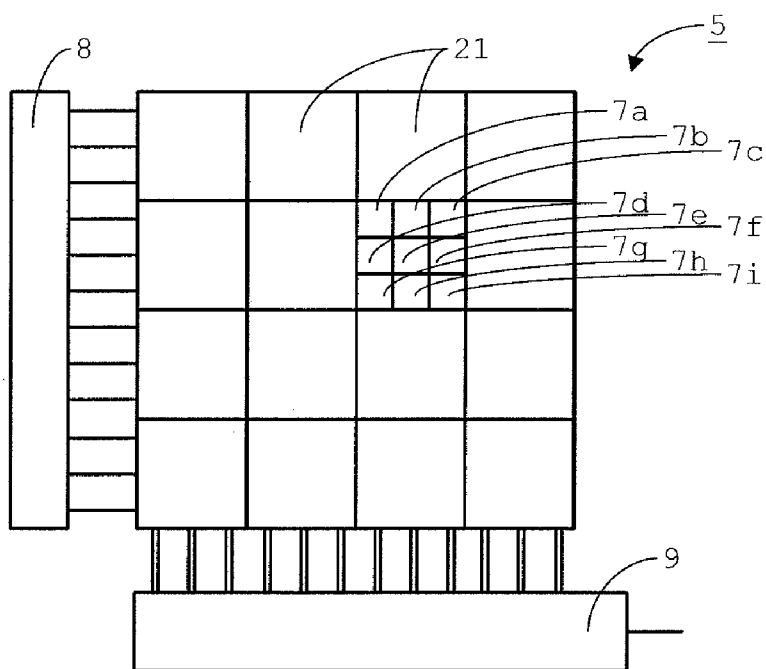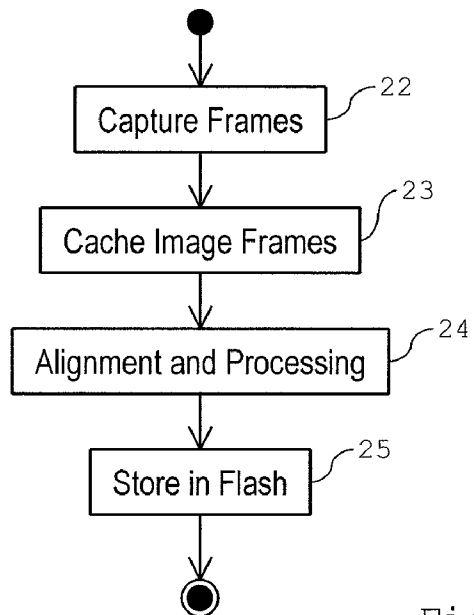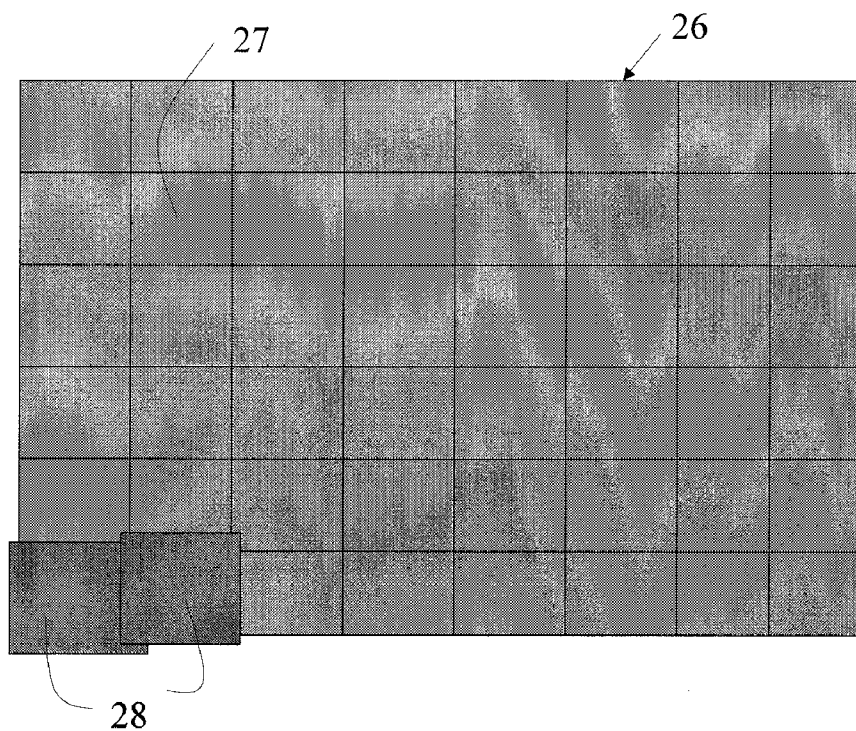
Store in Flash —25

<u>Fig. 3</u>

<u>Fig. 4</u>

IMAGE PROCESSING METHOD AND SYSTEM

[0001] The invention relates to an image processing method, comprising, providing a sequence of at least two image frames, calculating a motion vector representing relative movement of at least a part of said image frames, repositioning at least said part in the image frames in accordance with the calculated motion vector, and combining the image frames to form a combined image frame. The invention further relates to an image processing system, a computer program configured to enable a programmable processing device to carry out the method and a digital camera comprising an image processing system. Such an imaging processing method and system are described in the international patent application PCT/EP04/05108. As described in this application, the image processing method and system can be used for example in a digital camera. Although the known method and system generally provide combined image frames of high quality, complex movement of the camera including rotation and translation can require significant processing to align the images and may require up scaling of the image which is very costly with respect to computation. Moreover, in case of complex movement including rotation and translation, the known method and system may result in misalignments which can not be compensated.

[0002] The invention aims to provide an improved image processing method and system of the above-mentioned type.

[0003] To this end the image processing method according to the invention is characterized by

[0004] dividing the image frames into blocks,

[0005] calculating a block motion vector representing relative movement of each block,

[0006] repositioning each block in accordance with the calculated block motion vector and combining the blocks to form combined blocks in order to obtain the combined image frame.

[0007] The image processing system for forming a combined image from a plurality of image frames comprises an arrangement for loading an array of pixel values of a first image and at least an array of pixel values of one further image, the image processing system being configured to perform the steps of calculating a motion vector representing relative movement of at least a part of said image frames, repositioning at least said part in the image frames in accordance with the calculated motion vector, and combining the image frames to form a combined image frame.

[0008] According to the invention the image processing system is further configured to divide the image frames into blocks, to calculate a block motion vector representing relative movement of each block, to reposition each block in accordance with the calculated block motion vector and to combine the blocks to form combined blocks in order to obtain the combined image frame.

[0009] In this manner a method and system are obtained, wherein the image is broken down into blocks and a block motion vector is calculated for each block. The blocks of at least a second image frame are moved to align with the blocks of a first image frame, whereafter the blocks are combined to obtain a combined image frame having a high image quality. The method of the invention allows simpler motion models to be computed in specific regions of the image rather than requiring a highly exact complex model to be computed for the entire image. Up scaling of an image frame fro alignment purposes is not necessary.

[0010] The invention will now be explained in further detail with reference to the accompanying drawings, in which

[0011] FIG. 1 shows schematically the layout of an exemplary digital camera;

[0012] FIG. 2 shows in very schematic fashion some components of an image capturing device in the camera;

[0013] FIG. 3 is a flow diagram illustrating a method of capturing image frames and forming a combined final image;

[0014] FIG. 4 schematically shows a first image divided into blocks and two blocks of a second image aligned with the blocks of the first image.

[0015] One example of an image processing system usable in the context of the method outlined herein is part of a digital camera 1. Other examples of application of the image processing system and method include a photocopier, scanning device and a personal computer or workstation.

[0016] The digital camera 1 comprises a lens system 2 for focussing on one or more objects in a scene. When a shutter 3 is opened, the scene is projected through an aperture 4 onto a photosensitive area 5 (FIG. 2) of an image-capturing device 6. The shutter time is controllable, as is the diameter of the aperture. As an alternative, or addition, to the shutter 3, the image capturing device could be electronically controlled to provide the same effect (electronic shutter). The image-capturing device 6 can be device implemented in Complementary Metal-Oxide Semiconductor (CMOS) technology, or a Charge-Coupled Device (CCD) sensor.

[0017] Referring to FIG. 2, the photosensitive area 5 is divided into areas occupied by pixel cells 7a-i, of which only nine are shown for clarity. Each pixel cell 7 includes a device for generating a signal indicative of the intensity of light to which the area that it occupies within the photosensitive area 5, is exposed. It is noted that the devices occupying the pixel cells 7a-i are generally provided as components of one integrated circuit. An integral of the signal generated by a device is formed during exposure, for example by accumulation of photocurrent in a capacitor. Subsequent to exposure of the photo-sensitive area 5 for the duration of an exposure time interval, the values of the integrals of the generated signals are read out by means of row selection circuit 8 and column selection and readout circuit 9.

[0018] It is noted that, for simplicity, this description will not focus on the way in which colour images are captured. It is merely observed that any known type of technology can be used, such as colour filters, a colour-sensitive variant of the image capturing device 6, etc. In this respect, it is also observed that the photosensitive area 5 need not be the surface area of an integrated circuit comprised in an image-capturing device, or at least not for all colour components. Furthermore, although in the present application, image frames will be said to be captured consecutively, this does not preclude embodiments, wherein image frames of different colour components are captured in order, so that 'consecutively' captured image frames detailing one colour component are alternated by those detailing other colour components.

[0019] The output of the column select and read-out circuit 9 is provided in the form of one or more analogue signals to an Analogue-to-Digital converter (A/D-converter) 10. The A/D-converter 10 samples and quantises the signals received from the image capturing device 6, i.e. records it on a scale with discrete levels, the number of which is determined by the

number of bits of resolution of the digital words provided as output by the A/D converter **10**. The A/D converter **10** provides as output an array of pixel values encoding a captured image frame.

[0020] A Digital Signal Processor (DSP) **11** performs such features as interpolation between pixels and optionally compression of the image. Each exposure of the image-capturing device during an exposure time interval results in at least one frame.

[0021] The digital camera **1** comprises a storage device **12** for storing the image data encoding the captured images or image frames. The storage device can be any usual type of storage device, e.g. built-in flash memory, inserted flash memory modules, a disk drive with a floppy disk, a PCMCIA-format hard disk, or an optical disk drive.

[0022] A microprocessor **13** controls the operation of the digital camera **1**, by executing instructions stored in non-volatile memory, in this example a Read-Only Memory (ROM) **14**. The instructions in ROM **14**, in some embodiments in combination with routines programmed for execution by DSP **11**, enable the digital camera **1** to execute the image processing and capturing methods outlined in the present application.

[0023] Advantageously, the microprocessor **13** communicates with a co-processor **15** in which at least part of an image compression algorithm is implemented in hardware. Algorithms to compress images in accordance with the JPEG-standard are usable, for example. As part of the compression algorithm, the able, for example. As part of the compression algorithm, the image data is transformed into the spatial frequency domain. The co-processor **15** executes at least this transformation, using a Discrete Cosine Transform (DCT) in most cases.

[0024] Indications of the operating conditions and settings of the digital camera **1** are provided on an output device **16**, for example a Liquid Crystal Display, possibly in combination with a sound-producing device (not illustrated separately).

[0025] An input device **17** is shown schematically as being representative of the controls by means of which the user of the digital camera provides commands. In addition, the digital camera **1** illustrated in FIG. **1** comprises a flash driver circuit **18** for providing appropriate driving signals to one or more sources of flash lighting. The illustrated digital camera **1** also comprises a motion sensor **19**, for providing a signal representative of the movement of the digital camera **1**, and thus of the image-capturing device **6**. Furthermore, the digital camera **1** comprises an exposure metering device **20**. The purpose of the exposure metering device **20** is to measure the strength of the ambient light, so that the microprocessor **13** can determine the intensity of light to be emitted by any connected flash, in combination with the correct values for the settings determining the exposure, which include the exposure time interval for each captured image frame, as will be elaborated on below.

[0026] It will be noted that the density of the areas occupied by the pixel cells **7***a-i* determines the maximum attainable spatial resolution of a captured image frame. The readout time depends on the number of pixel cells. It can be relatively long in embodiments such as the one illustrated in FIG. **2**, because each row is selected in turn using row selection circuit **8**, whereupon the column selection and readout circuit **9** senses the values of the accumulated photocharge stored in the photodevices in the pixel cells in that row. To reduce the

total time involved in repeatedly exposing the photosensitive area and capturing an image frame, the spatial resolution can be set to a different value between exposures.

[0027] In one embodiment, the microprocessor **13** controls the image-capturing device **6** in such a manner that the one pixel value read out per cluster area **21** represents an integral of the signal generated in one of the pixel cells **7** that lie within the cluster area **21**. This embodiment has the virtue that it can be used with any type of image-capturing device **6**.

[0028] In a preferred embodiment, the image-capturing device **6** has the capability to "bin" the outputs of multiple pixel cells. In this embodiment, the microprocessor **13** directs the image-capturing device **6** to generate an array of pixel values (each value being associated with one of the defined cluster areas **21**) in such a manner that each pixel value is representative of the integral of the sum of the signals generated by at least two device in pixel cells that occupy the same defined cluster area **21**. In this shown embodiment, this could mean that the pixel value for one cluster area **21** is the sum, or alternatively the average, of the integrals of the signal generated by all nine of the shown pixel cells **7***a-7i*. This embodiment is preferred, because it increases the sensitivity. Effectively, each pixel value represents the amount of light that fell on the whole of a defined cluster area **21**, instead of just on the area occupied by one pixel cell **7**. Thus, smaller light fluxes are detectable. Furthermore, binning decreases the amount of noise, i.e. leads to a low resolution image with a higher Signal-to-Noise-Ratio (SNR). As the binning capability is a function of the image-capturing device that is implemented in hardware, binning does not add appreciably to the read out time. As binning results in fewer pixels to be read out, the read out time is actually reduced. Preferably, the number of image frames that are captured at the highest resolution is equal to, more preferably lower than, the number of image frames captured at lower spatial resolutions. A combined final image formed on the basis of such a series of image frames will have a good SNR.

[0029] In an embodiment, upon receiving a command from a user to capture an image, the microprocessor **13** controls the digital camera **1** to carry out a series of steps **22-25**. A user of the digital camera **1** may input a desired exposure time for a combined final image, together with settings determining the amount of flash light, the diameter of aperture **4** and the sensitivity of the photodevices in the pixel cells **7**. In alternative embodiments, the microprocessor determines one or more of these values automatically, using a signal output by the exposure metering device **20**, and possibly one or more pre-defined combinations of values. Subsequently, the microprocessor **13**, upon receiving a command actually to capture the combined final image, executes a first step **22** of capturing a number of image frames. This step **22** comprises retrieving the desired exposure time for the combined final image, determining the number of image frames to be captured and, for each image frame, calculating exposure settings determining an exposure level applicable to the image frame. The settings include the exposure time interval for the frame. Preferably, the other settings are determined such as to result in exposure time intervals for the image frames that, together, are shorter than the desired exposure time for the combined final image. It is noted that the embodiment in which "binning" is carried out allows a reduction in the exposure time interval applicable to the image frames, because binning increases the sensitivity. Effectively, 'binning' results in the introduction of an extra amplification of the photo-electric signal. The microproces-

3

sor **13** advantageously takes account of this. It calculates the length of the exposure time interval applicable to the image frame at a lower spatial resolution value in dependence on the spatial resolution value, i.e. the amount of 'binning'.

[0030] When calculating the settings determining the exposure levels applicable to the image frames, the microprocessor **13** preferably implements one or more of the methods outlined in international patent application PCT/EP04/051080. That is, they are calculated such that the total exposure level that is determined as desirable for the combined final image is unevenly distributed over the image frames.

[0031] Following the first step **22** in which the image frames are captured, the arrays of pixel values encoding the image frames are cached in a second step **23**. Following the second step **23**, they are aligned and processed in a third step **24**. The combined final image resulting from the third step is stored in storage device **12** in a final step **25**. Although the present description will now continue on the assumption that the digital camera **1** carries out all of the steps **22-25**, the third and fourth steps **24,25** could be carried out in a separate image processing system, for example a personal computer or workstation. In that case, the second step would involve committing the generated arrays of pixel values to storage in the storage device **12** or transferring them to the computer via a data link (not shown).

[0032] The present invention specifically relates to the alignment carried out in the third step **24**. In case of complex movements of the digital camera **1**, including rotation and translation, an alignment of a sequence of image frames by means of a global motion vector as described in international patent application PCT/EP04/051080 may not provide an accurate alignment over the full image. The image processing method of the invention will be described starting from a sequence of two image frames, wherein preferably a first image frame is a high resolution image frame and the second image frame is a low resolution image frame. However, the sequence of image frames used may comprise any other number of image frames and the image frames may be of the same resolution.

[0033] FIG. **4** very schematically shows a first image frame **26** divided into blocks **27** having an initial size of for example 16×16 pixels. A second image frame is divided into blocks **28** of the same size. In FIG. **4**, two blocks **28** of the second image frame are shown, which blocks are aligned with the corresponding blocks **27** of the first image frame **26** using the method of the invention. In the alignment process described, the blocks **27** of the high resolution image frame **26** remain fixed relative to each other. The blocks **28** of the low resolution image frame are moved to align with the blocks **27** of the high resolution image frame. To determine the amount of movement of the blocks **28**, block motion vectors each representing relative movement of each of the blocks **28**, are calculated after dividing the image frames into the blocks **27**, **28**. Each block **28** is moved or repositioned in accordance with the corresponding calculated block motion vector and the blocks **27**, **28** are combined to form combined blocks which provide the combined image frame.

[0034] The size of the blocks depends on the amount of variation that requires compensating. In a first embodiment the variation or distribution about the mean of the motion vector across the entire image can be examined. If there is a sizeable variation in the motion vector across the image then the block size can be reduced. A global motion vector can be determined using the method described in the above-men-

tioned international patent application or by an algorithm known as RANSAC or any other suitable algorithm. The variation in the motion vector is determined by taking sample blocks as reference, which are available as part of the RANSAC algorithm and examining the value of the vector for each sample block with respect to the global motion vector. If the deviation is to large, the size of the blocks is reduced.

[0035] In an alternative embodiment a block motion vectors can be determined and if the vector for block A is significantly different than the vector for block B then the block size is reduced to allow for more refined alignment.

[0036] In another variant of the invention the magnitude of the differences between the two frames across the image is checked. This can be measured from block to block by comparing the pixel values after alignment of the blocks. If the magnitude of the differences for a block is greater than a particular threshold the blocks may be divided into smaller blocks and the alignment recalculated for the smaller blocks.

[0037] It is noted that in case the amount of variation shows that the size of the blocks needs to be reduced, it is possible to reduce the size of all blocks across the entire image or to reduce only the size of a part of the blocks where the variation between the blocks exceeds the threshold used. It is further noted that any combination of the different methods for checking the amount of variation can also be used.

[0038] When it has been determined that the amount of variation can be compensated using the initial size of the blocks **27**, **28** or a reduced size, the block motion vectors are calculated, if they were not yet determined to check the amount of variation. According to a first embodiment the block motion vectors can be calculated using the single or global motion vector and the position of the respective blocks **27**, **28** with respect to the point of the image to which the single motion vector relates. Each block motion vector is interpolated from the global motion vector by calculating the offset of each block caused by the rotation of the image frame. A rotation by a certain angle results in an offset in the x and y domain that is determined by the angle of rotation and the distance of the block from the point of rotation of the image frame.

[0039] In a second method a block motion vector is calculated for each of the blocks **28** by using a method according the above-mentioned international patent application or a known method, such a RANSAC or a least mean squares algorithm.

[0040] As schematically shown in FIG. **4**, repositioning of the blocks **28** may result in an overlap of the blocks **28** at their boundaries. Problems which may occur due to misalignment between the pixels in the boundary areas can be handled in several ways.

[0041] As a first option, especially in cases wherein the second and possible further images have a low resolution, it can be assumed that there will be no visible effects of the misalignment of the pixels in the boundaries. The amount of misalignment can be checked by comparing the pixel values of the pixels in the boundary areas. If the amount of misalignment exceeds a threshold value, it can be decided to reduce the size of the blocks to ensure that the transitions are not visible.

[0042] As a second option, the pixel values in the boundary areas can be calculated as follows. Each pixel in the boundary areas will have a first value predicted by the one block and a second value predicted by the other block. The final pixel value is calculated by taking the average of the two pixel

values. At the corners of four blocks, the pixels may have four predicted values, each corresponding to one of the blocks meeting at the corner. Again the final pixel value is calculated as an average of the four predicted pixel values.

[0043] As an improvement of this averaging option, it is possible to weight the average depending on the location of the pixel within the block. For example, a pixel in the boundary between blocks A and B, closer to five pixels into block A will be 90% weighted to the pixel values associated with block A. This will reduce down to 50% at the boundary with block B and then down to 10% five pixels into block B.

[0044] In general, an algorithm could be used according to which pixels that are equidistant from the centres of two adjacent blocks can be assigned the average value of the two predicted values. As a pixel lies closer to the centre of one block and further from the centre of the other block, the average is weighted according to the ratio of the distance.

[0045] The method described allows simpler motion models to be computed in specific regions of the image rather than requiring a highly exact complex model to be computed for the entire image. Further, if there is motion in the image of a subject, then this motion can be compensated using the method described, as the motion for each individual block can be made different to compensate for the subject motion.

[0046] It is noted that the wording motion vector is to be understood as any combination of variables describing motion of an array of one or more pixels.

[0047] The invention is not restricted to the above-described embodiments which can be varied in a number of ways within the scope of the invention.

1. Image processing method, comprising
providing a sequence of at least two image frames
calculating a motion vector representing relative movement of at least a part of said image frames,
repositioning at least said part in the image frames in accordance with the calculated motion vector,
combining the image frames to form a combined image frame,
dividing the image frames into blocks,
calculating a block motion vector representing relative movement of each block,
repositioning each block in accordance with the calculated block motion vector and
combining the blocks to form combined blocks in order to obtain the combined image frame.

2. Method according to claim 1, wherein the size of the blocks is determined according to the amount of the variation in the image frames.

3. Method according to claim 2, wherein the amount of variation is determined by determining the deviation from the mean of a global motion vector, wherein the size of the blocks is reduced with respect to an initial size if the amount of variation exceeds a threshold value.

4. Method according to claim 2, wherein the image frames are divided into blocks of an initial size, wherein the block motion vector is determined for at least a plurality of blocks of the initial size and the difference between the block motion vectors is used as a representation of the amount of variation, wherein the image frames are divided into blocks of a reduced size if the amount of variation exceeds a threshold value.

5. Method according to claim 2, wherein the pixel values of corresponding blocks of the sequence of image frames are compared to determine the amount of variation, wherein the image frames are divided into blocks of a reduced size if the amount of variation exceeds a threshold value.

6. Method according to claim 2, wherein a global motion vector is determined for the image frames, wherein a block motion vector for each of the blocks is determined by means of the global motion vector and the position of the block in the image frames.

7. Method according to claim 2, wherein pixel values in the boundary areas of the repositioned blocks are checked for mismatch, wherein if the mismatch between the pixel values exceeds a threshold value the size of the blocks is reduced and the steps to obtain the combined image frame are repeated.

8. Method according to claim 2, wherein the pixel values in the boundary areas of the repositioned blocks are averaged.

9. Method according to claim 8, wherein the pixel values in the boundary areas of the repositioned blocks are averaged according to a weighting algorithm.

10. Method according to claim 2, wherein a high resolution image is used as a first image, wherein the blocks of each further image are repositioned with respect to the blocks of the high resolution image.

11. Image processing system for forming a combined image frame from a plurality of image frames, which the image processing system comprises an arrangement for loading an array of pixel values of a first image and at least an array of pixel values of one further image, the image processing system being configured to perform the steps of
calculating a motion vector representing relative movement of at least a part of said image frames,
repositioning at least said part in the image frames in accordance with the calculated motion vector,
combining the image frames to form a combined image frame, wherein said system is further configured
to divide the image frames into blocks,
to calculate a block motion vector representing relative movement of each block,
to reposition each block in accordance with the calculated block motion vector and
to combine the blocks to form combined blocks in order to obtain the combined image frame.

12. Computer program configured, when loaded into a programmable processing device (1) to enable the programmable processing device to carry out a method according to claim 1.

13. Digital camera comprising an image processing system according to claim 11.

* * * * *