

Description

This invention relates to analysis and processing of video images.

A video image (which will be understood to encompass frozen images such as facsimile images, in addition to moving images) will in general include at least one object which is of interest and a "background" of lesser interest (and hence of lesser importance).

To analyse the image, e.g. detect the presence/absence or position of a particular object of interest, is often desirable in a variety of applications.

In an image transmission system an improved picture quality might be achieved if data relating to important parts of the scene, i.e. objects of interest, is coded using relatively more bits than data relating to unimportant (i.e. background) parts. For example, in a videophone system a typical image comprises a head and shoulders against a background, and the face area of the head is visually the most important; it is thus desirable to be able to identify the head area from the shoulders and background so as to be able to process the head at a higher refreshment rate than the rest, so that the impression of smooth head motion is conveyed. The ability to locate a head within a head and shoulders scene can thus be used to modify the spatial allocation of video data, enabling a degree of visual importance to be attributed to blocks within the data.

Also, if the position of an object is accurately tracked with time it will be possible to predict its motion, thus allowing "motion compensated" DPCM.

One way of identifying different regions of an image is to utilise the method proposed by Nagao (M. Nagao - "Picture recognition and data structure", Graphic Languages - ed Make and Rossenfield, 1972). This method has been used in a videophone type system, on an image of a head and shoulders against a background. Some success was achieved in determining the sides of the head when the subject was clean shaven, but very little success was achieved in other cases; so this method is not considered reliable enough for the basis of an area identifying method.

Conventional coders, for instance hybrid discrete cosine transform coder, use no 'scene content' information to code the data within the scene, so each part of the scene is operated on as if it has the same visual importance as every other part.

Other image analysis applications are manifold (for example, in automated manufacturing systems).

It is also known to code video images for transmission using Vector Quantisation (VQ). In VQ coding, the image is represented initially by an array of digital data corresponding to the image frame. Blocks of array points ("Sub-arrays") are compared with vectors from a codebook, and the best-matching vector

selected using a "least squares" difference criterion. A code designating this vector is then transmitted to represent the sub-array. At the receiving end the indicated vector is selected from an identical codebook and displayed.

An example of such a coding technique is disclosed in an article titled "Image Coding using Vector Quantization" by A. Gersho and B. Ramamurthi, IEEE International Conference of Acoustics, Speech and Signal Processing, ICASSP 82, Proceedings; 3-5 May 1982, Paris, FR, Vol 1 of 3 IEEE (US) pgs 428-431.

An image is partitioned into cells of $p \times p$ pixels. Each cell is regarded as a vector of dimension p^2 and is encoded by searching through a codebook for a nearest matching representative vector. A binary word identifying the selected representative vector is assigned as the codeword to describe the original cell. The decoder uses this codeword to address a codebook. Each entry of the codebook contains a full precision digital representation of one of the N representative vectors. The codebook design is based on a clustering technique for vector quantizer design preceded by a classification of training cells into edge or shade cells. Each training vector (cell) is first classified as an edge or shade (non-edge) cell and a separate codebook design is performed for each of the two cell types. The final codebook is a concatenation of the two codebooks.

The underlying principle of the invention, however, is to use VQ as an identification (e.g. object location) method.

According to a first aspect of the present invention a method of analysing an image in which sub-arrays of the image are each compared with the members of a composite codebook of vectors to determine which of the vectors is most similar to each sub-array, the composite codebook of vectors comprising a concatenation of at least two codebooks, each codebook being derived from a respective training set of image data; characterised in that: one of the codebooks is a first codebook whose vectors have been derived from a set of image data derived from images of a first object; each of the vectors of the first codebook has associated with it a flag indicating that it is a member of the codebook; and each sub-array of the image which is determined to be most similar to a vector of the first codebook is labelled as being associated with the first object.

According to a second aspect of the present invention a method of encoding a series of images in which sub-arrays of each image are each compared with the members of a composite codebook of vectors to determine which of the vectors is most similar to each sub-array, the composite codebook of vectors comprising a concatenation of at least two codebooks, each codebook being derived from a respective training set of image data; characterised in that:

one of the codebooks is a first codebook whose vectors have been derived from a set of image data derived from images of a first object; each of the vectors of the first codebook has associated with it a flag indicating that it is a member of the codebook; each sub-array of an image which is determined to be most similar to a vector of the first codebook is labelled as being associated with the first object; and the images are encoded such that sub-arrays associated with an object are updated at a higher rate than the other sub-arrays.

According to a third aspect of the present invention a coder for encoding a video image comprises; a vector quantization means for comparing sub-arrays of the image with the members of a composite codebook of vectors to determine which of the vectors is most similar to each sub-array, the composite codebook of vectors comprising a concatenation of at least two codebooks, each codebook being derived from a respective training set of image data; characterised in that: one of the codebooks is a first codebook whose vectors have been derived from a set of image data derived from images of a first object; each of the vectors of the first codebook has associated with it a flag indicating that it is a member of the codebook; the coder includes labelling means for labelling as being associated with the first object each sub-array of the image which is determined to be most similar to a vector of the first codebook; and an image coding means which encodes updates the sub-arrays associated with an object are updated at a higher rate than the other sub-arrays.

An embodiment of the invention will now be described by way of non-limitative example concerned with the identification of a head in a head and shoulders against a background scene, with reference to the accompanying drawings in which:

figure 1 is a block diagram illustrating the initial stages of operation of parts of a coder embodying the invention;

figures 2a-g show schematically various stages in the training sequence used to derive the codebook;

figure 3 is a block diagram illustrating the operation of a coder embodying the invention;

figure 4a shows schematically a frame to be analysed;

figure 4b illustrates the sub-array blocks used in vector quantising figure 4a;

figure 4c shows the state of flags corresponding to the vector quantised image of figure 4b;

figure 4d shows schematically the result of analysing the frame of figure 4a according to the invention; and

figure 5 shows schematically a coder embodying the invention.

To enable the invention to operate, it is necessary to have provided a composite codebook which in-

cludes vectors flagged as being "head". Preferably others are flagged as being "background". It is possible to derive a "standard" codebook for either an average or a given speaker, but to allow flexibility and greater accuracy of identification, this codebook is derived at the start in an initial "training" sequence. A preferred way of implementing such a sequence will now be described.

To generate "head" and "background" parts of the codebook, it is necessary to unambiguously obtain some "head only" data and "background only" data; a crude initial head detection algorithm is required.

Referring to figures 1 and 2, in order to detect the head, digital data representing several contiguous frames of the head and shoulders image are captured; for instance in a store 1. One of these frames is depicted in figure 2a. This data does not need to be extremely accurate, but rather representative.

On the assumption that the prime moving areas within the data sequence are directly associated with the head area, frame differencing 2 is applied to the data representing each adjacent pair of frames. This process typically yields a set of difference data for each adjacent pair representing moving areas together with random noise across the whole image area.

For all picture elements (pels) represented by each set of difference data, each pel above a given threshold value of intensity is set to maximum intensity (255) and each pel below the threshold is set to minimum intensity (0). This 'thresholding' 3 removes a large quantity of the random noise and some of the moving areas.

Median filtering 4 is next applied to each set of difference data, which very effectively removes most of the remaining random noise, but erodes only small amounts of the moving areas.

The image represented by each set of data at this stage will rarely provide a clear outline of the head; unless the head to background contrast is very high and the movement of the head between adjacent frames is more than one pel. Often only one side and the top of the head may be depicted as shown in figure 2b.

Generally, the moving areas will be clustered in regions around the head area, but some isolated clusters may arise due to motion in other areas of the image. A clustering process 5 is used to remove some of the isolated clusters: two orthogonal histograms are generated, one representing the number of 'moving' pels in the columns of the image represented by the data and one representing the number of moving pels in the rows of the image represented by the data. The first order moments are calculated and the 'centre of gravity' of the image determined, as shown in figure 2c. A rectangle is then generated, centred on these co-ordinates, of such dimensions that a given percentage of moving area is included within it, see figure 2d. The pels remaining outside this rectangle

are set to zero intensity, figure 2e. By a suitable choice of rectangle isolated clusters are removed by this process.

Constraints are imposed on the selection of the rectangles in order to reduce the occurrence of incorrect rectangles. Since a very small movement of the head between one frame and the next may produce a very small rectangle, the rate of change of size of the rectangle from one set of data to the next is restricted: either each of the boundary lines of the rectangle are constrained to lie within a small distance of the corresponding boundary in the immediately preceding set of data; or the maximum rate of change of the size of the rectangle is linked to the frame difference energy (eg. the square of the difference data), so if the difference energy is small the change is kept small, but if the difference energy is large the rate of change may be greater.

The rectangle - rectangles are used because they require very few bits of data to define - is then shrunk if necessary, at 6 in figure 1, and as shown in figure 2f, to become the smallest rectangle that can be placed around the data to enclose all the remaining non-zero pels. This rectangle is assumed to represent an approximate model of the head.

A border is then created, at 7 in figure 1, around the final rectangle, as shown in figure 2g. This border defines an exclusion zone from where no data will later be taken. This ensures that when the border is applied to the respective frame of the original image the data outside the border will be exclusively head data and the data outside the border will be exclusively background data.

If five frames of data are initially captured in the store 1, then four adjacent pairs of frames are analysed and four sets of data result. After the four borders have been set 7 the head area data and the background area data are extracted from the first four frames of the original image respectively and the Linde-Buzo-Grey algorithm is applied to generate a VQ codebook for each area 8, for example, a 9 bit background codebook and 10 bit head codebook (i.e. codebooks containing respectively 2^9 and 2^{10} entries). The two codebooks are then combined 9 to form one codebook, each entry of which has an associated flag indicating its origin.

Referring now to Figures 3 and 4a-d, after this training sequence is completed, the composite codebook is used to locate the head in successive image frames. The VQ coder operates just as it would in a prior art system using VQ as the transmission coding, but for each block of pels coded 10, the code generated will include a flag (for example, the first digit) indicating whether that block is "head" or "background" so that the head position is known for each frame.

It will of course be appreciated that when the codebook is derived at the coder as indicated above, VQ cannot be used as the transmission code (unless this

codebook is made known to the decoder first by transmitting an indication of the vectors).

Since the quantisation process is inherently approximate, it will be appreciated that occasionally blocks from the head part of the image may best match a vector from the "background" part of the codebook, or vice versa. The actual identification of the head will thus usually evolve ignoring isolated "head" blocks using erosion and clustering 11, 12 (for example, as described above), or designating the area with the highest concentration of "head" blocks as the actual head.

Another method involves detecting isolated "head" blocks and then examining the error between the block and the "head" vector, and that between the block and the best-matching "background" vector, and if the two scores are similar (i.e. there is ambiguity about whether the block is "head" or "background"), reflagging the block to "background" instead.

If the head blocks are too scattered, it may be that the codebook is insufficient to characterise the head. In this case, a retraining sequence may be employed to regenerate the codebook.

This retraining sequence may either simply be a further sequence of the kind described above, or it may attempt to improve (rather than simply redefine) the codebook. For example, a count may be kept of the number of "incorrect" (i.e. scattered) as opposed to "correct" (i.e. concentrated in the head area) occurrences of each vector, and the scatter may this be reduced by rejecting from the codebook vectors which occur incorrectly too often.

Or, alternatively, the approximate head location derived by locating the greatest concentration of "head" blocks may be used, in the same manner as described above, as an area for generating a new "head" codebook.

These latter approaches, in which VQ coder "learns" from each retraining sequence, are preferred on grounds of accuracy.

Figure 5 shows a block diagram of a video coding apparatus (eg for a video telephone) embodying the invention. Video signals are fed from an output 20 to a frame store 21 where individual picture element values are recorded in respective store locations so that desired sub-arrays of pels are accessible for further processing. The sub-array sizes may typically be 8×8 . In an initial, training, phase of the apparatus a training control unit 22 - which may for example be a suitably programmed microprocessor system - carries out the codebook generation method described above, and enters the vectors (and flags) in a VQ codebook store 23. It will be understood that the VQ process involves matching 8×8 sub-array to the nearest one of the stored vectors, viz. a number of 8×8 patterns which are consistently fewer in number than the maximum possible number (2^{64}) of such patterns.

In the coding phase of the apparatus, the matching is carried out by VQ control logic 24 which receives successive sub-arrays from the frame store 21 and compares each of these with all the vectors in the codebook store. The simplest form of comparison would be to compute the mean square difference between the two; the vector giving the lowest result being deemed to be the best match. The output from the VQ control logic is the sequence of flags associated with the vectors thus identified.

The actual logic is carried out in this example by an inter frame differential coder 25, in which an inter-frame difference (in subtractor 26) is taken (in conventional manner) between pels from the frame store 21 and a previous frame predictor 27. As is usual in such systems, a quantizer 28 and an output buffer 29 (to match the irregular rate of data generation to a transmission link operating at a constant rate) are shown. A receiver (not shown) uses the difference-information to update a reconstructed image in a frame store. The flag output from the VQ control logic 24 is connected (if required, via erode/cluster circuits 30) to the differential coder 25. When the flag indicates that "head" information is being processed, the coder operates normally. If however "background" is indicated, then the generation of difference information is carried out less frequently (eg on alternate frames only). This operation is illustrated by a switch 31 which, when the flag indicates "background", breaks the coding loop on alternate frames.

It will be apparent from the foregoing that any visually distinctive object or objects may be accurately detected, recognised or located using methods according to the invention.

Claims

1. A method of analysing an image in which sub-arrays of the image are each compared with the members of a composite codebook of vectors to determine which of the vectors is most similar to each sub-array, the composite codebook (9) of vectors comprising a concatenation of at least two codebooks, (8) each codebook being derived from a respective training set of image data; characterised in that:
 - one of the codebooks is a first codebook whose vectors have been derived from a set of image data derived from images of a first object; each of the vectors of the first codebook has associated with it a flag indicating that it is a member of the codebook; and
 - each sub-array of the image which is determined to be most similar to a vector of the first codebook is labelled as being associated with the first object.

2. A method as claimed claim 1 in which one of the codebooks is derived from a set of image data (2a) derived from images of human heads.
3. A method of encoding a series of images in which sub-arrays of each image are each compared with the members of a composite codebook of vectors to determine which of the vectors is most similar to each sub-array, the composite codebook of vectors comprising a concatenation of at least two codebooks, each codebook being derived from a respective training set of image data; characterised in that:
 - one of the codebooks is a first codebook whose vectors have been derived from a set of image data derived from images of a first object; each of the vectors of the first codebook has associated with it a flag indicating that it is a member of the codebook;
 - each sub-array of an image which is determined to be most similar to a vector of the first codebook is labelled as being associated with the first object; and
 - the images are encoded such that sub-arrays associated with an object are updated at a higher rate than the other sub-arrays.
4. A coder for encoding a video image comprising; a vector quantization means for comparing sub-arrays of the image with the members of a composite codebook of vectors to determine which of the vectors is most similar to each sub-array, the composite codebook of vectors comprising a concatenation of at least two codebooks, each codebook being derived from a respective training set of image data; characterised in that;
 - one of the codebooks is a first codebook whose vectors have been derived from a set of image data derived from images of a first object; each of the vectors of the first codebook has associated with it a flag indicating that it is a member of the codebook;
 - the coder includes labelling means for labelling as being associated with the first object each sub-array of the image which is determined to be most similar to a vector of the first codebook; and
 - an image coding means which updates the sub-arrays associated with an object at a higher rate than the other sub-arrays.
5. A coder as claimed in claim 4 in which the coding means is a discrete cosine transform coder.
6. A coder as claimed in claim 4 in which the coding means is a motion compensated DPCM coder.



In the coding phase of the apparatus, the matching is carried out by VQ control logic 24 which receives successive sub-arrays from the frame store 21 and compares each of these with all the vectors in the codebook store. The simplest form of comparison would be to compute the mean square difference between the two; the vector giving the lowest result being deemed to be the best match. The output from the VQ control logic is the sequence of flags associated with the vectors thus identified.

The actual logic is carried out in this example by an inter frame differential coder 25, in which an inter-frame difference (in subtractor 26) is taken (in conventional manner) between pels from the frame store 21 and a previous frame predictor 27. As is usual in such systems, a quantizer 28 and an output buffer 29 (to match the irregular rate of data generation to a transmission link operating at a constant rate) are shown. A receiver (not shown) uses the difference-information to update a reconstructed image in a frame store. The flag output from the VQ control logic 24 is connected (if required, via erode/cluster circuits 30) to the differential coder 25. When the flag indicates that "head" information is being processed, the coder operates normally. If however "background" is indicated, then the generation of difference information is carried out less frequently (eg on alternate frames only). This operation is illustrated by a switch 31 which, when the flag indicates "background", breaks the coding loop on alternate frames.

It will be apparent from the foregoing that any visually distinctive object or objects may be accurately detected, recognised or located using methods according to the invention.

Claims

1. A method of analysing an image in which sub-arrays of the image are each compared with the members of a composite codebook of vectors to determine which of the vectors is most similar to each sub-array, the composite codebook (9) of vectors comprising a concatenation of at least two codebooks, (8) each codebook being derived from a respective training set of image data; characterised in that:
 - one of the codebooks is a first codebook whose vectors have been derived from a set of image data derived from images of a first object; each of the vectors of the first codebook has associated with it a flag indicating that it is a member of the codebook; and
 - each sub-array of the image which is determined to be most similar to a vector of the first codebook is labelled as being associated with the first object.
2. A method as claimed claim 1 in which one of the codebooks is derived from a set of image data (2a) derived from images of human heads.
3. A method of encoding a series of images in which sub-arrays of each image are each compared with the members of a composite codebook of vectors to determine which of the vectors is most similar to each sub-array, the composite codebook of vectors comprising a concatenation of at least two codebooks, each codebook being derived from a respective training set of image data; characterised in that:
 - one of the codebooks is a first codebook whose vectors have been derived from a set of image data derived from images of a first object; each of the vectors of the first codebook has associated with it a flag indicating that it is a member of the codebook;
 - each sub-array of an image which is determined to be most similar to a vector of the first codebook is labelled as being associated with the first object; and
 - the images are encoded such that sub-arrays associated with an object are updated at a higher rate than the other sub-arrays.
4. A coder for encoding a video image comprising; a vector quantization means for comparing sub-arrays of the image with the members of a composite codebook of vectors to determine which of the vectors is most similar to each sub-array, the composite codebook of vectors comprising a concatenation of at least two codebooks, each codebook being derived from a respective training set of image data; characterised in that;
 - one of the codebooks is a first codebook whose vectors have been derived from a set of image data derived from images of a first object; each of the vectors of the first codebook has associated with it a flag indicating that it is a member of the codebook;
 - the coder includes labelling means for labelling as being associated with the first object each sub-array of the image which is determined to be most similar to a vector of the first codebook; and
 - an image coding means which updates the sub-arrays associated with an object at a higher rate than the other sub-arrays.
5. A coder as claimed in claim 4 in which the coding means is a discrete cosine transform coder.
6. A coder as claimed in claim 4 in which the coding means is a motion compensated DPCM coder.



7. A coder as claimed in any one of claims 4, 5 and 6 further including composite codebook generating means to generate the composite codebook of vectors including: identification means (1, 2, 3, 4, 5, 6, 7) for identifying an area of the image corresponding to an object; generating means for generating a first codebook from sub-arrays within that area and for generating a further codebook from sub-arrays outside that area; and combining means for combining the first and further codebook (9) to form a composite codebook (9).
8. A coder as claimed in claim 7 in which the identification means is arranged to analyse the difference between a pair of temporally separated image frames to identify areas of the image corresponding to the object.
9. A coder as claimed in claim 8, in which the identification means (1, 2, 3, 4, 5, 6, 7) comprises:
 - (a) means (2,3) for generating from a pair of frames (2a) of the image array elements within a notional two dimensional field, the value of each position within the field indicating whether the difference between the luminance levels of the picture elements at corresponding positions in the two frames lies above or below a threshold;
 - (b) clustering means (5) for determining the centre of gravity within the said field of those array elements indicating a difference above the threshold; and
 - rectilinear delimiting means (6, 7) for determining the position of a boundary about the centre of gravity which encloses a given proportion of those array elements, whereby all picture elements lying within a boundary so determined are identified as belonging to an area of the image corresponding to the object.
10. A coder as claimed in claim 9 which means further comprises filtering means (4) for median filtering the array elements within the notional two dimensional field prior to the determination of the centre of gravity.
11. A coder as claimed in either one of claims 9 and 10, wherein said boundary about the centre of gravity is of a finite number of elements in thickness.
12. A coder as claimed in any one of claims 9, 10 and 11 wherein said boundary about the centre of gravity is of rectangular shape.
13. A coder as claimed in claim 12, wherein the rectangular shaped boundary is centred upon the centre of gravity and each side of the rectangle is

moved inward, if possible, until it abuts at least one of those array elements indicating a difference above the threshold.

Patentansprüche

1. Verfahren zum Analysieren eines Bildes, wobei Untermatrizen des Bildes jeweils mit den Elementen eines zusammengesetzten Codebuchs von Vektoren verglichen werden, um zu bestimmen, welcher der Vektoren jeder Untermatrix am ähnlichsten ist, wobei das zusammengesetzte Codebuch (9) von Vektoren eine Aneinanderkettung von zumindest zwei Codebüchern (8) aufweist, wobei jedes Codebuch von einem jeweiligen Übungssatz von Bilddaten abgeleitet wird;

dadurch gekennzeichnet, daß

eines der Codebücher ein erstes Codebuch ist, dessen Vektoren von einem Satz Bilddaten abgeleitet sind, welche von Bildern eines ersten Gegenstandes abgeleitet sind;

jeder der Vektoren des ersten Codebuchs eine dazu zugeordnete Kennmarke hat, welche anzeigt, daß er ein Element des Codebuchs ist; und

jede Untermatrix des Bildes, von dem man bestimmt, daß es einem Vector des ersten Codebuchs am ähnlichsten ist, als dem ersten Gegenstand zugeordnet gekennzeichnet wird.
2. Verfahren nach Anspruch 1, wobei eines der Codebücher von einem Satz Bilddaten (2a) abgeleitet ist, welche von Bildern menschlicher Köpfe abgeleitet sind.
3. Verfahren zum Codieren einer Serie von Bildern, wobei Untermatrizen jedes Bildes jeweils mit den Elementen eines zusammengesetzten Codebuchs von Vektoren verglichen werden, um zu bestimmen, welcher der Vektoren jeder Untermatrix am ähnlichsten ist, wobei das zusammengesetzte Codebuch von Vektoren eine Aneinanderkettung von zumindest zwei Codebüchern aufweist, wobei jedes Codebuch von einem jeweiligen Übungssatz von Bilddaten abgeleitet ist;

dadurch gekennzeichnet, daß

eines der Codebücher ein erstes Codebuch ist, dessen Vektoren von einem Satz Bilddaten abgeleitet worden sind, welche von Bildern eines ersten Gegenstandes abgeleitet sind;

jeder der Vektoren des ersten Codebuchs dazu ein Kennzeichen zugeordnet hat, welches anzeigt, daß er ein Element des Codebuchs ist;

jede Untermatrix eines Bildes, von dem man bestimmt, daß es einem Vektor des ersten Codebuchs am ähnlichsten ist, als mit dem er-



sten Gegenstand verbunden gekennzeichnet wird; und

die Bilder so codiert sind, daß Untermatrizen, welche einem Gegenstand zugeordnet sind, mit einer höheren Geschwindigkeit auf den neuesten Stand gebracht werden als die anderen Untermatrizen.

4. Codiergerät zum Codieren eines Videobildes, welches aufweist:

eine Vektorquantisiereinrichtung zum Vergleichen von Untermatrizen des Bildes mit den Elementen eines zusammengesetzten Codebuchs von Vektoren, um zu bestimmen, welcher der Vektoren jeder Untermatrix am ähnlichsten ist, wobei das zusammengesetzte Codebuch von Vektoren eine Aneinanderkettung von zumindest zwei Codebüchern aufweist, wobei jedes Codebuch von einem entsprechenden Übungssatz von Bilddaten abgeleitet ist;

dadurch gekennzeichnet, daß

eines der Codebücher ein erstes Codebuch ist, dessen Vektoren von einem Satz Bilddaten abgeleitet worden sind, welche von Bildern eines ersten Gegenstands abgeleitet sind;

jeder der Vektoren des ersten Codebuchs ein ihm zugeordnetes Kennzeichen hat, welches anzeigt, daß er ein Element des Codebuchs ist;

das Codiergerät eine Kennzeichnungseinrichtung beinhaltet zum als dem ersten Gegenstand zugeordnete Kennzeichnen jeder Untermatrix des Bildes, von dem bestimmt wird, daß es einem Vektor des ersten Codebuchs am ähnlichsten ist; und

eine Bildcodiereinrichtung, welche die einem Gegenstand zugeordneten Untermatrizen schneller auf den neuesten Stand bringt als die anderen Untermatrizen.

5. Codiergerät nach Anspruch 4, wobei die Codiereinrichtung ein diskretes Kosinustransformations-Codiergerät ist.

6. Codiergerät nach Anspruch 4, wobei die Codiereinrichtung ein bewegungskompensiertes DPCM-Codiergerät ist.

7. Codiergerät nach einem der Ansprüche 4, 5 und 6, welches weiterhin zusammengesetzte Codebuch-Erzeugungseinrichtungen aufweist, um das zusammengesetzte Codebuch von Vektoren zu erzeugen, einschließlich: einer Identifizierungseinrichtung (1, 2, 3, 4, 5, 6, 7) zum Identifizieren einer einem Gegenstand entsprechenden Fläche des Bildes; einer Erzeugungseinrichtung zum Erzeugen eines ersten Codebuchs aus Untermatrizen innerhalb der Fläche und zum Erzeugen eines weiteren Codebuchs aus Untermatri-

zen außerhalb der Fläche; und eine Kombinationseinrichtung zum Kombinieren des ersten und weiteren Codebuchs (9), um ein zusammengesetztes Codebuch (9) zu bilden.

8. Codiergerät nach Anspruch 7, wobei die Identifizierungseinrichtung angeordnet ist, um den Unterschied zwischen einem Paar zeitlich getrennter Bilddatenblöcke zu analysieren, um dem Gegenstand entsprechende Flächen des Bildes zu identifizieren.

9. Codiergerät nach Anspruch 8, wobei die Identifizierungseinrichtung (1, 2, 3, 4, 5, 6, 7) aufweist:

(a) eine Einrichtung (2, 3), mit der man aus einem Paar von Datenblöcken (2a) der Bildmatrixelemente innerhalb eines konzeptionellen zweidimensionalen Feldes den Wert jeder Position innerhalb des Feldes erzeugt, welche anzeigt, ob der Unterschied zwischen den Luminanzniveaus der Bildelemente bei entsprechenden Positionen in den beiden Datenblöcken über oder unter einem Schwellwert liegt.

(b) Eine Zusammenballungseinrichtung (5) zum Bestimmen des Schwerpunktes innerhalb des Feldes der Matrixelemente, wodurch ein Unterschied über dem Schwellwert angezeigt wird; und

eine geradlinige Begrenzungseinrichtung (6, 7) zum Bestimmen der Position einer Grenze um den Schwerpunkt, welche einen gegebenen Anteil der Matrixelemente umschließt, wobei alle innerhalb einer so bestimmten Grenze liegenden Bildelemente identifiziert werden als zu einer Fläche des Bildes gehörend, welches dem Gegenstand entspricht.

10. Codiergerät nach Anspruch 9, dessen Einrichtung weiterhin eine Filtereinrichtung (4) aufweist zur Mittelungsfilterung der Matrixelemente innerhalb des konzeptionellen zweidimensionalen Feldes, bevor der Schwerpunkt bestimmt wird.

11. Codiergerät nach einem der Ansprüche 9 und 10, wobei die Grenze um den Schwerpunkt in ihrer Dicke aus einer endlichen Anzahl von Elementen besteht.

12. Codiergerät nach einem der Ansprüche 9, 10 und 11, wobei die Grenze um den Schwerpunkt rechteckförmig ist.

13. Codiergerät nach Anspruch 12, wobei die rechteckförmige Grenze um den Schwerpunkt zentriert ist und jede Seite der Rechtecks nach innen bewegt wird, wenn möglich, bis sie an zumindest eines der Matrixelemente stößt, welches einen



Unterschied oberhalb des Schwellwertes anzeigt.

Revendications

1. Un procédé d'analyse d'une image dans lequel des sous-réseaux de l'image sont comparés chacun aux éléments d'un livre de codes composite de vecteurs pour déterminer celui des vecteurs qui est le plus semblable à chaque sous-réseau, le livre de codes composite (9) de vecteurs comprenant un enchaînement d'au moins deux livres de codes (8), chaque livre de codes étant dérivé d'un jeu respectif d'apprentissage de données d'images;
 - caractérisé en ce que :
 - l'un des livres de codes est un premier livre de codes dont les vecteurs ont été dérivés à partir d'un jeu de données d'images dérivées d'images d'un premier objet;
 - à chacun des vecteurs du premier livre de codes est associé un drapeau indiquant que c'est un élément du livre de codes; et
 - chaque sous-réseau de l'image qui est déterminé comme le plus semblable à un vecteur du premier livre de codes est étiqueté comme étant associé au premier objet.
2. Un procédé selon la revendication 1 dans lequel l'un des livres de codes est dérivé d'un jeu de données d'images (2a) dérivées d'images de têtes humaines.
3. Un procédé d'encodage d'une série d'images dans lequel des sous-réseaux de chaque image sont comparés chacun aux éléments d'un livre de codes composite de vecteurs pour déterminer celui des vecteurs qui est le plus semblable à chaque sous-réseau, le livre de codes composite de vecteurs comprenant un enchaînement d'au moins deux livres de codes, chaque livre de codes étant dérivé d'un jeu respectif d'apprentissage de données d'images;
 - caractérisé en ce que :
 - l'un des livres de codes est un premier livre de codes dont les vecteurs ont été dérivés à partir d'un jeu de données d'images dérivées d'images d'un premier objet;
 - à chacun des vecteurs du premier livre de codes est associé un drapeau indiquant que c'est un élément du livre de codes; et
 - les images sont encodées d'une manière telle que les sous-réseaux associés à un objet sont mis à jour à une fréquence plus élevée que les autres sous-réseaux.
4. Un codeur destiné à encoder une image vidéo

comportant :

un moyen de quantification de vecteurs pour comparer des sous-réseaux de l'image avec les éléments d'un livre de codes composite de vecteurs pour déterminer celui des vecteurs qui est le plus semblable à chaque sous-réseau, le livre de codes composite de vecteurs comprenant un enchaînement d'au moins deux livres de codes, chaque livre de codes étant dérivé d'un jeu respectif d'apprentissage de données d'images;

caractérisé en ce que :

l'un des livres de codes est un premier livre de codes dont les vecteurs ont été dérivés à partir d'un jeu de données d'images dérivées d'images d'un premier objet;

à chacun des vecteurs du premier livre de codes est associé un drapeau indiquant que c'est un élément du livre de codes;

le codeur inclut des moyens d'étiquetage pour étiqueter comme étant associé au premier objet chaque sous-réseau de l'image qui est déterminé comme le plus semblable à un vecteur du premier livre de codes; et

un moyen de codage d'image qui met à jour les sous-réseaux associés à un objet à une fréquence plus élevée que les autres sous-réseaux.

5. Un codeur selon la revendication 4 dans lequel le moyen de codage est un codeur à transformée discrète par cosinus.
6. Un codeur selon la revendication 4 dans lequel le moyen de codage est un codeur par modulation différentielle par impulsion codée, (ou DPCM) compensé en mouvement.
7. Un codeur selon l'une quelconque des revendications 4, 5 et 6 comprenant en outre un moyen générateur de livre de codes composite afin d'engendrer le livre de codes composite de vecteurs incluant: un moyen d'identification (1, 2, 3, 4, 5, 6, 7) prévu pour identifier une zone de l'image correspondant à un objet; un moyen générateur prévu pour engendrer un premier livre de codes à partir de sous-réseaux situés à l'intérieur de cette zone et pour engendrer un autre livre de codes à partir de sous-réseaux situés à l'extérieur de cette zone; et un moyen de combinaison prévu pour combiner le premier et l'autre livre de codes (9) pour former un livre de codes composite (9).
8. Un codeur selon la revendication 7 dans lequel le moyen d'identification est agencé pour analyser les différences entre une paire de trames temporellement séparées d'image afin d'identifier des zones de l'image correspondant à l'objet.



9. Un codeur selon la revendication 8, dans lequel le moyen d'identification (1, 2, 3, 4, 5, 6, 7) comprend :
- (a) un moyen (2, 3) de génération, à partir d'une paire de trames (2a) du réseau d'image, des éléments situés à l'intérieur d'une trame notionnelle à deux dimensions, la valeur de chaque position située à l'intérieur de la trame indiquant si la différence entre les niveaux de luminance des éléments de représentation à des positions correspondantes dans les deux trames est situé au-dessus ou au-dessous d'un seuil; 5
 - (b) des moyens de groupage (5) prévus pour déterminer le centre de gravité, à l'intérieur de ladite trame, des éléments de réseaux qui indiquent une différence supérieure au seuil; et 10
 - (c) un moyen de délimitation rectiligne (6, 7) prévu pour déterminer la position d'une limite, entourant le centre de gravité, qui enferme une proportion donnée de ces éléments de réseau, grâce à quoi tous les éléments de représentation situés à l'intérieur d'une limite ainsi déterminée sont identifiés comme appartenant à une zone de l'image correspondant à l'objet. 20
10. Un codeur selon la revendication 9 qui comprend en outre un moyen de filtrage (4) prévu pour un filtrage médian, avant la détermination du centre de gravité, des éléments du réseau situés à l'intérieur de la trame notionnelle à deux dimensions. 25
11. Un codeur selon la revendication 9 et 10, dans lequel ladite limite entourant le centre de gravité est d'une épaisseur à nombre fini d'éléments. 30
12. Un codeur selon l'une quelconque des revendications 9, 10 et 11 dans lequel ladite limite entourant le centre de gravité est de forme rectangulaire. 35
13. Un codeur selon la revendication 12, dans lequel la limite à forme rectangulaire est centrée sur le centre de gravité et chaque côté du rectangle est déplacé vers l'intérieur, si possible, jusqu'à ce qu'il vienne en butée sur au moins l'un des éléments du réseau qui indique une différence supérieure au seuil. 40

45

50

55