

Method for encoding streams of video data based on
groups of pictures (GOP)

5 The present invention deals with a method for encoding
video data streams with a view to enhancing the
reliability and the quality of reception of these
streams in a wireless digital network, governed by a
communication protocol for example of WIFI type. The
streams concerned are sent in multicast mode and
10 originate from at least one transmitter broadcasting
them to receiving devices, in this case situated within
a relatively restricted perimeter, such that they can
operate with a communication protocol as mentioned,
designed to operate in a more local environment. The
15 receiving devices are then provided with at least one
software application making it possible to process and
view the information received.

More particularly, the video data effected by the
20 invention are based on groups of pictures (GOP)
conventionally used in video coding (in particular in
the MPEG standards for the compression, the
decompression, the processing and video and audio
coding) to define the order of arrangement of the
25 images with internal coding and with predictive coding
allowing the generation of a visible image in a video
stream. The streams processed in the context of the
invention also comprise equally conventional
transmission error correction means, the correction
30 being done in this case by redundancy and being
performed via algorithms known as AL-FEC (application
level forward error correction) algorithms implementing
a matrix function.

35 Although the method of the invention can have many
practical applications, since it very generally
optimizes the transmission of streams of pictures in
real and/or quasi-real time, one of the applications

envisaged is the virtually instantaneous broadcasting of events of sports meeting or show type, within the perimeter of an enclosed space where said event takes place. The quality of the video streams transmitted to the client devices primarily allow the fluid and good quality viewing of the video pictures, but also make it possible to enrich the viewing through processes capable of conferring an added value on the streams received, for example a repetition and therefore a reviewing of certain sequences, if necessary in slow motion or even with a zoom effect. Such processes would be almost meaningless if the quality of the video streams was not assured.

The video streams are usually transmitted in the form of successive information packets which obey a predetermined subdivision for the sending thereof, these packets very generally comprising the pictures to be transmitted, encoded for the transmission, and additional information which, in case of reception problems, is used for reconstruction algorithms to try to restore the initial pictures.

As such, the method of the invention applies to video streams composed in practice of successive data packets G, each packet G comprising, as is known, a first set M containing native video data to be transmitted (the pictures of the video stream encoded) and a second set containing redundancy data obtained by a suitable algorithm. According to the invention, the first set of data is composed from all the data M of at least one group of pictures (GOP).

The subdivision of the information packets to be transmitted in video streams therefore relies, innovatively according to the invention, on particular groups of pictures developed and defined in the context of standards organizing the architecture of the video

data packets to be transmitted. The computer system charged with sending the video streams, and consequently the software driving same, consequently perform a subdividing of the information which is based
5 on these groups of pictures, or GOP, formatted in a way that is standard in the video coding world, identifying a specific architecture based on particular coded pictures. The data M may come under a single group of pictures GOP, of which they then contain all the data,
10 or two or more groups of pictures GOP, with this same condition of including all their data in one and the same packet.

Still according to the invention, the size of all the
15 pictures of the group of pictures is added to the data M of the native video stream, in an area juxtaposing each picture, which, for example occupies two data bytes per picture.

20 Furthermore, according to one possibility, an index table containing the absolute position of each picture of the group of the pictures can be added to the data M, said table being able to be located in an area for example situated at the start of the series of coded
25 pictures to be transmitted.

In practice, according to the invention, each packet G is then organized as K first data blocks of predetermined size L and N redundancy blocks of size L,
30 K being equal to M/L , the next integer value immediately higher than the result of the division M/L being assigned to K if M/L is not an integer number. The portion of block corresponding to the difference $(K \times L - M)$, which exists only if M/L is not an integer
35 number, is then filled by data identified as being stuffing data, for example zeros.

Preferably, according to the invention, for each of the blocks K and N, the information is added, comprising:

- the number of the group of pictures (GOP), incremented for each new group of pictures;
- 5 - the value of K, incremented from 1 to K for each block K;
- the value of N, beginning at K+1 and ending at N+K;
- the value of L; and
- 10 - the number of the block (from 1 to K+N).

The length of each of the blocks K and N then becomes L+ n bytes that the coding of the above information requires. The existence of all these additional data, combined with the subdivision into blocks K + N of the native video information, itself based on a standard of video coding (one or more GOP), makes it possible to considerably enhance the control of the transmission of the video streams. The reconstruction operations to be implemented given the assumption of a corrupted transmission of the video streams are clearly facilitated by the novel architecture proposed.

Moreover, the invention relates also to a video stream organized according to the method applied above and comprising a succession of coded data packets G each composed of a first set of native video data to be transmitted and of a second set of error correction redundant data, characterized in that the first set contains all the data of at least one group of pictures (GOP).

This type of video stream, structured according to the method of the invention, makes it possible to guarantee a reception in a quality that is sufficient first of all for the viewing of the streams transmitted to be simply comfortable, then also for any subsequent processes applied to the signals transmitted to be able

to be based on an information platform sufficient for the performance thereof.

The invention in reality culminates in organizing the information, that is to say the native video streams to be sent, so as to allow an optimization of the response of the system to defects in transmission of the video signal, based on a sufficiently structured and information-rich architecture.

10

It will now be described in more detail, with reference to the attached figures, representing an exemplary implementation of the method of the invention showing different possible steps for the generation of optimized video streams, for which:

15

- figure 1 represents a diagram schematically representing the organization of a group of pictures as used in the context of the invention;
- figure 2 illustrates, for a same data packet, the addition to the native video stream of the size of each coded picture of the group of pictures;
- figure 3 shows the size M of the picture packet when an index table is added to the data;
- figure 4 represents the subdivision into K blocks of length L of the data packet M , with the possible addition of stuffing when $K \times L > M$;
- figure 5 illustrates the addition of N blocks due to the encoding by redundancy (FEC); and
- figure 6 finally shows the addition of additional parameters enriching the information to be transmitted and facilitating the subsequent implementation of the data recovery, if appropriate.

20

25

30

Referring to figure 1, the first set of data, forming the native video stream to be transmitted before encoding of the transmission error correction function, is composed from all the data of a group of pictures.

35

This data organization structure is well known in the field of video coding, since it is in particular defined by the MPEG standards. These groups of pictures, or GOP, define the nature and the order in which the pictures resulting from several particular types of codings are arranged. The grouping of these pictures in fact forms a particular series, this series being repeated periodically to constitute the encoding of a video stream. There are thus pictures with internal coding I and pictures with predictive coding, the latter being able to be of two types: the pictures P with predictive coding based on a past image, and the pictures B with bidirectional predictive coding. The pictures P contain information on difference, by motion prediction, with a picture I or a past image P. The pictures B contain information on difference with the past and future pictures I or P within a group of pictures.

For information, the pictures I or P can be used as reference pictures, although this is not generally the case with the pictures B. The visible pictures are, in short, generated from coded pictures contained in a group of pictures, which is also expressed as a number of coded pictures, which are constructed at the time of encoding so as to guarantee the video stream transfer rate.

As appears in particular in figure 1, the group of pictures N (GOP N) begins with a reference picture I_N , followed by pictures P and B in a particular order which defines said group of pictures. The video stream to be transmitted comprises this group of pictures N, followed by a group of pictures N+1 of which the initial picture I_{N+1} is the new reference picture, preceded by a group of pictures N-1 and so on. The set of these groups of pictures GOP_N forms the encoded video stream. For the present invention, the basis of the

subdivision of the video data packets to be sent consists of these groups of pictures. Figure 1 represents only one thereof, but there can be two or more thereof in the video data packets M.

5

Figure 2 illustrates another feature of the method of the invention, which enhances the coding of the video stream by adding the size of all the pictures I, P, B of the group of pictures in an area juxtaposing each of these pictures. The extra addition of an index table, appearing at the start of the video data packet in figure 3, further pads the encoded video information, the whole forming a first encoded set of size M.

15 This length or size M is the starting point of a calculation targeting a subdivision into K blocks of length L (even in figure 4), L being a constant and K an integer number. When the result of the division of M by L is not an integer, the immediately higher integer value is then assigned to K. In most cases, $K \times L > M$. Given this hypothesis, as represented in figure 4, a stuffing is applied to obtain K complete blocks of length L. This simply involves adding zeros to the block K, which are identified by the program as stuffing value.

Hitherto, the subject has been video coding, that is to say the coding of pictures aiming for the transmission of the video stream and ultimately allowing the generation of a visible image after decoding. Figures 5 and 6 add the concept of correction of any transmission errors. As mentioned previously, error correction is performed in this case by redundancy, by using an algorithm that is known per se of AL-FEC (application layer forward error correction) type. The error correction encoding implements a matrix function which culminates in adding N extra data blocks, also of

30
35

length L. These are the blocks K+1 to K+N appearing in these figures 5 and 6.

5 According to a final operation specific to the invention, aiming to further optimize the processing possibilities, information is added for each of the blocks K and N, comprising in particular but not exclusively:

- 10 - the number of the group, incremented for each new group of pictures (GOP);
- the value of K, incremented from 1 to K for each block K;
- the value of N, beginning at K+1 and ending at N+K;
- 15 - the value of L; and
- the number of the block (from 1 to K+N).

This complementary information in practice better organizes the information, that is to say the set of
20 video streams to be sent, including the redundancy blocks, so as to allow an optimization of the response of the system to video signal transmission defects. It in fact provides data on the blocks taken individually, in order to better trace the errors and locate them in
25 the data packets.

The invention is of course not limited to the examples described and explained with reference to the figures, but it encompasses the variants and versions which fall
30 within the scope of the claims.

CLAIMS

1. A method for encoding video data streams based on groups of pictures (GOP) used in video coding to
5 define the order of arrangement of the images with internal coding and with predictive coding allowing the generation of a visible image, including means for correcting transmission errors by redundancy, for transmission over a wireless
10 network governed by a local data communication protocol, said video stream being composed of successive data packets G, each packet G comprising a first set M containing the video data to be transmitted, and a second set containing
15 redundancy data obtained by a redundancy algorithm, characterized in that the first set of data M is composed from all the data of at least one group of pictures (GOP), and that the size of all the pictures of the group of pictures is added
20 to the data M, in an area juxtaposing each image.
2. The method for encoding video data streams as claimed in the preceding claim, characterized in
25 that an index table containing the absolute position of each picture of the group of pictures is added to the data M.
3. The method for encoding video data streams as
30 claimed in one of the preceding claims, characterized in that each packet G is organized as K first data blocks of predetermined size L and N redundancy blocks of size L, K being equal to M/L , the next integer value immediately higher
35 than the result of the division M/L being assigned to K if M/L is not an integer number.

4. The method for encoding video data streams as claimed in the preceding claim, characterized in that the portion of block corresponding to the difference $(K \times L - M)$, which exists if M/L is not an integer number, is filled by data identified as being stuffing data, for example zeros.
5. The method for encoding video data streams as claimed in one of claims 3 and 4, characterized in that, for each of the blocks K and N , the information is added, comprising:
- the number of the group of pictures (GOP), incremented for each new group of pictures;
 - the value of K , incremented from 1 to K for each block K ;
 - the value of N , beginning at $K+1$ and ending at $N+K$;
 - the value of L ; and
 - the number of the block (from 1 to $K+N$).