

## DESCRIPTION

Title of the invention: Allocation method for low-rate link

[0001] The invention relates to a method for allocating video content on a limited-rate communication link. The invention is applicable for example in low-rate or limited-rate radio communication systems.

[0002] Rate constraints are highly stringent in terms of rate for channelling a real-time video on a low-rate (or limited-rate) radio link.

[0003] The prior art describes various rate distortion allocation techniques (rate distortion optimization or RDO).

[0004] Patent application US10116989 describes an RDO allocation technique that is implemented in the decoder before the display in order to compensate the time delay between the capture of the stream and the image, or "time to glass".

[0005] Patent CA3001672 describes an allocation method that changes resolution dynamically on the basis of the rate in order to meet rate constraints of the system.

[0006] Patent application US2015350654 discloses an allocation technique that changes frame rate (images/sec) on the basis of the capacities of the decoder of the receiver, and not the required rate.

[0007] These techniques are not always effective, for the following reasons:

- In the case of RDO, the resolution and the frame rate, i.e. the number of images to be compressed per second, are fixed,
- These solutions are additionally intrinsically implemented within the video compression module, which therefore means having to modify the structure of the compression component,
- The methods described in the prior art do not take into account stringent limited-rate constraints, notably in the field of radio communications.

[0008] In a conventional video compression system, each coded image may have a compressed rate that differs from one image to another, and reference is then made to an instantaneous rate. Reference is generally made to an average rate when calculating the average of these instantaneous rates over a time window (generally of a few seconds). In a "conventional" application in which latency is accommodated

with significant buffering (time between compression and sending, scheme found symmetrically on the receiver side: reception and decoding), it is possible to accept compressing the images in succession, assuming that it will be possible to compensate this by adapting the instantaneous rate of the following frames, and assuming that the packets will be kept in a memory and sent correctly by the transmission system.

[0009] The problem in a limited-rate system is that it is not adapted to support excessively high instantaneous rate "peaks" that may occur in a sequence of compressed images. The system does not manage to memorize (or buffer-store) the frames, and it risks losing packets that are however perfectly formed. These packet losses prior to sending are detrimental after transmission in terms of the decoding and the rendering of the transmitted images.

[0010] In conventional video over IP (Internet) systems such as VOD, there is always a memory large enough to support rate variations.

[0011] In a conventional MPEG compression system and, except for the frames called "Intra" (intrinsically coded), the images are compressed using the previous images in order to reduce the amount of information by coding, more or less, only the difference that exists between the content of the images and using predictive schemes (reference is made to "Inter" frames). When compressing an image, use is made of tools, offered by the compression standard, that will make it possible to depict the image using a minimum number of bits. These tools contain what are called predictive schemes that make it possible to approximate the pixels in the image and then to code the residuals (that is to say the difference between the prediction and the values of pixels under consideration, block by block), which may (or may not) be projected into what is called a frequency space. The projection is performed through a matrix computing operation known to a person skilled in the art, such as for example the discrete cosine transform, i.e. DCT. These coefficients (or the values of residuals if it has not been deemed beneficial to use the frequency transform) are then approximated through quantization, thereby making it possible to code only values having the smallest amplitude. The compression is thus more effective in terms of bits by the entropy coder, which is the last video compression step, since it creates the final "bitstream", which will be encapsulated in a packet associated with an image, and this is what is called a compressed image. Changing

the quantization step modifies the kept amount of information to be compressed. The greater the quantization step, the more the values under consideration are approximated, and the more distortion is created (the rendered quality will be worse). The more effective the compression, the lower the instantaneous rate will be. Reference is then made to a rate/distortion compromise. In a conventional scheme, modulating the quantization step is known to a person skilled in the art, since this is the parameter by way of which coders operate in order to meet an average rate setpoint for a compressed video sequence.

[0012] For "conventional" applications, coders assume that it is possible to be more "flexible" with regard to the instantaneous rate, and consider an average rate over a relatively large time interval (i.e. number of images considered to calculate the average rate is high), IP transmission systems supporting far greater latencies and memory capacities than systems that are highly limited in terms of rate and limited in terms of buffer memory size.

[0013] In the remainder of the description, the expressions "input video buffer" and "image buffer" denote the same module.

[0014] The method according to the invention consists notably in using a closed-loop control process to control the input of a video coder (predictive compression) receiving images from an input video buffer in order to adapt the images to the capacity of a radio buffer arranged downstream of the video buffer and before transmission in a communication system.

[0015] The invention relates to a method for controlling the content of a video stream containing a plurality of images  $F_1, \dots, F_n$  on a low-rate transmission link in a device comprising an input video buffer, followed by a predictive video coder whose output is connected to an output video buffer, the output video buffer being connected to a radio buffer of known capacity, characterized in that it comprises at least the following steps:

- Estimating, in the output video buffer, a parameter  $P(Tr)$  associated with the occupancy level  $Tr$  of the radio buffer, and modifying the content of the input video buffer as follows:

- If the parameter  $P(Tr)$  associated with the estimated occupancy level  $Tr$  meets a predefined threshold, then transmitting the images contained in the input video buffer to the video coder on a rolling basis,
- If the estimated parameter does not meet the predefined threshold, then modifying the content of the input video buffer by modifying the image  $Fi+1$  to be compressed so as to adapt said content to the occupancy level  $Tr$  of the radio buffer, for example by replacing the image to be compressed with one or more images.

[0016] The value of the occupancy level  $Tr$  is estimated for example in real time, and this value is compared with a predefined occupancy level threshold value.

[0017] It is also possible to estimate the variation in the occupancy level  $Tr$  over a given time interval, and to modify the content of the input video buffer if the variation in the occupancy level  $Tr$  is continuous and exceeds a threshold value.

[0018] The content of the input buffer is modified for example by replacing, one or more times, an image to be coded with one or more previously coded images chosen from the images present in the image buffer.

[0019] According to one variant embodiment, the quality of the image present in the buffer is checked and the current image is replaced using an image whose contours are sharpest.

[0020] It is possible to replace the current image in the input video buffer using an image containing information useful for rendering in the decoder, such as an image containing critical information in the selected images in the video input buffer.

[0021] The image buffer may be used by regularly compressing different images in order to avoid a transmission outage.

[0022] The quantization step of the video coder may be modified for a replaced image in the buffer.

[0023] The method is used to control the content of a video stream in a radio system before transmission to the communication system.

[0024] The invention also relates to a system for controlling the content of a video stream within a device comprising a camera, a video coder taking, at input, images from an input video buffer, and generating compressed images in accordance with a

predictive model for an output video buffer, said output video buffer transmitting these compressed images to a radio buffer, characterized in that:

- The output video buffer is configured so as to quantize the number of packets from the compressed images and to transmit it to a video stream control device,
- The video stream control device is configured so as to estimate a parameter  $P(Tr)$  associated with the occupancy level  $Tr$  of the radio buffer and to send an order to the input video buffer so that this modifies the images to be transmitted if the estimated associated parameter does not meet a predefined threshold by performing the steps of the method according to the invention.

[0025] The input video buffer may comprise an image selection filter in order to select the images having information useful for rendering in the decoder.

[0026] Other features, details and advantages of the invention will become apparent from reading the description, which is given with reference to the appended drawings, which are given by way of illustrative and wholly non-limiting example and in which, respectively:

[0027] [Fig. 1] shows an illustration of a sequence of Intra frames and Predicted frames in a time/rate graph,

[0028] [Fig. 2] shows an example of a device according to the invention arranged upstream of a radio and of a communication system.

[0029] Figure 1 is a reminder of the intrinsic organization of the compression scheme using a video coder with prediction, with intrinsically coded Intra frames 10, the coding does not use the past and does not require a high rate and predicted frames 11 (GOP) whose coding uses the past and requires a lower rate in most cases.

[0030] Figure 2 illustrates an example of a system for modifying an input buffer for video transmission on a radio with a video coder with predicted compression.

[0031] A video camera 20 acquires images and transmits them to a first video buffer 21 whose output is connected to a video coder 22. The output of the video coder 22 is transmitted to the output video buffer 23, at the output of which the compressed frames  $F_c$  are transmitted to a radio buffer FIFO stack 25, and instantaneous rate information  $Q$  is transmitted to a rate controller 24. The rate controller 24 has an

output that transmits coding decisions firstly to the input video buffer and secondly to the video coder.

[0032] The FIFO stack 25 is connected to a packet creation device 26 before transmission to a modulator 27 for shaping with a waveform, before digital to analogue conversion 28 and transmission to a radio communication system 29.

[0033] The image input buffer (input video buffer) receives the stream of images  $F_i$  transmitted by the camera 20 and an order  $D_c$  from the video rate controller to adapt the content of the image input buffer on the basis notably of the fill level  $Tr$  of the radio buffer. The image input buffer may comprise a filter  $S$  known to a person skilled in the art, the role of which is notably to select the "best" image, for example the image whose contours are sharpest, and that will allow a better visual quality in terms of rendering (i.e. after decoding).

[0034] The video coder 22 receives fixed parameters, such as the resolution and the target average rate to be met. The video coder may be in accordance with any video coder or codec known to a person skilled in the art, such as MPEG-2, H.264, HEVC, or AV1. The structure of the coder is not modified for the basic implementation of the method according to the invention.

[0035] The output video buffer 23 contains the compressed bitstream of each image with a view to being encoded using an IP/RTP (Internet/Protocol Real-Time Transport Protocol) protocol, for example.

[0036] The video rate controller 24 has the following data: the type of images to be coded, Intra or Predicted, the features of the radio buffer, such as the average rate, the memory capacity of the buffer at input thereof and the transmission frequency within the radio.

[0037] The idea of the method according to the invention is notably to modify the content of the image buffer, i.e. the images to be compressed by the video coder, in order to "control" the stream of images compressed by the video coder and transmitted to the radio buffer, thus avoiding this buffer being saturated or, by contrast, exhausting this buffer, where an insufficient number of compressed images to be transmitted would lead to an outage in the stream.

[0038] A video coder codes the images with a given frequency, generally 25 or 30 images per second. The first image buffer therefore has to regularly provide the

images to be coded to the coder. If the instantaneous rate risks leading to the capacities of the occupancy level of the radio buffer being exceeded, rather than compressing a new image, it is possible to choose, from the images present in the image buffer, an image that it will be desired to compress several times in a row, but not necessarily the one that is the first in the list. Without going into detail, this is tantamount to simulating reducing the number of images per second (frame rate) artificially at input, and compressing the same number of images per second in the coder.

[0039] The image buffer contains the  $N$  images to be compressed. These images originate from the camera, and they are transmitted at a given frequency, for example thirty images per second. The transmission frequency is set by the radio part of the system. This information is initially known to the rate control device.

[0040] The method consists notably in estimating the fill level  $Tr$  of the radio buffer, or an associated parameter  $P(Tr)$ , in order to decide on the modification of the content of the image video buffer. If the radio buffer is approaching saturation (i.e. it exceeds a threshold able to be set by the user on the basis of the radio system that is used), the video rate controller 24 will transmit an order  $D_c$  to the input video buffer or image buffer. The order  $D_c$  that is given consists for example in duplicating an already compressed image and replacing, in the input video buffer, a new image to be compressed with this duplicated image in order to facilitate the compression, i.e. use less rate. This thus meets the rate constraint imposed by the radio buffer. In the opposite case, if the radio buffer is "empty" or not very full, the video rate controller may on the other hand ask to use the image buffer regularly and automatically by compressing different images, or even to reduce the quantization step used by the compression in order to avoid an insufficient amount of bitstream data to be transmitted and risking a communication outage.

[0041] To control the modification of the input buffer, the method estimates, in the output video buffer, the occupancy level  $Tr$  of the radio buffer or a parameter  $P(Tr)$  linked to the occupancy level. To this end, the output buffer may count the number of packets and the size thereof in order to predict a value of the occupancy level  $Tr$  for the radio buffer. The output video buffer transmits this information to the video rate controller in real time. Said video rate controller knows the intrinsic capacities of the radio system (i.e. the size of the radio buffer and the transmission frequency of the

radio data), and retains in memory the value or values of the occupancy level determined at a given time or over a time interval. It is therefore capable of deciding whether or not it is necessary to modify the compression parameters and to transmit the images without modification, or else to momentarily duplicate an image, etc.

[0042] When the value of the occupancy level  $Tr$  at the current time exceeds a threshold value,  $V_{threshold}$ , then the rate controller will give the instruction to the image input buffer not to transmit the new current image  $I_{curr}$  to the video coder, but to transmit a duplicated image (the image that has already been coded by the video coder) to the video coder. The coding of this duplicated image requires fewer resources in terms of rate.

[0043] The estimation is performed using a closed-loop control operation, and the radio buffer is incremented over time until reaching a saturation threshold.

[0044] On the other hand, when the radio buffer contains practically no more packets to be transmitted, below another threshold, this may lead to a packet transmission outage.

[0045] The threshold value  $V_{threshold}$  is determined for example on the basis of the features of the radio buffer and of the waveform.

[0046] It is also possible for the rate controller to monitor the variation in the occupancy level of the radio buffer measured in real time. When the rate controller observes a significant increase or a linear increase in the occupancy level, it may give the order to the input buffer to modify its content to be transmitted to the video coder.

[0047] If the instantaneous rate is consistent with the capacities and the sending rate of the radio, then the following image is transmitted to the video coder, as in a conventional scheme, and if not the following image is erased from the input buffer and replaced with the previous one, as explained above.

[0048] According to one embodiment, it is possible, in the input video buffer, to check the quality of an image present in the buffer and to use the "sharpest" image in the conventional video sense, whose contours are sharpest. This image will then be used as substitute image for the current image. The system makes it possible to code images that are visually more "significant" in terms of decoding. This check is

performed using a filter known to a person skilled in the art and will therefore not be described since it is not the subject of the invention.

[0049] It is also possible to use, as criterion for replacing the current image, images that contain the most information useful for rendering in the decoder. For example, it will be possible to choose the images that contain the greatest amount of face for video surveillance, or those containing critical information in the images selected in the image buffer.

[0050] In some cases, the video rate controller will send an order to the video coder so that this modifies the quantization step associated with this image, if necessary. This will make it possible to reduce (or increase) the amount of information to be compressed and reduce (or increase, respectively) the average rate.

[0051] By way of example, it is considered that the input video buffer contains 5 images F0, F1, F2, F3, F4, corresponding respectively to the times: k, k+1, k+2, k+3 and k+4.

[0052] When the value of the instantaneous rate for an image at the time k-1 has exceeded the bit setpoint significantly by virtue of exceeding a provided threshold, and there is a risk of momentarily greatly exceeding the rate capacities on the transmission link (risk of losses before sending), the method will replace, in the buffer, the images F0, F1, F2, F3, F4, corresponding respectively to the times: k, k+1, k+2, k+3 and k+4, with duplicates of the image F0 at the time k, for the five images, or else replace, in the buffer, the three first images with the image F0 and the two last ones using F3 at the time k+3. The choice will depend on the detected rate exceedance.

[0053] The method according to the invention notably offers the following advantages:

[0054] Making it possible to artificially modify the number of ("different") frames to be coded per second,

[0055] Reducing the amount of (visual) information to be coded,

[0056] Not requiring any a priori information on the captured content.

## CLAIMS

1. Method for controlling the content of a video stream containing a plurality of images  $F_1, \dots, F_n$  on a low-rate transmission link in a device comprising an input video buffer (21), followed by a predictive video coder (22) whose output is connected to an output video buffer (23), the output video buffer being connected to a radio buffer (25) of known capacity, characterized in that it comprises at least the following steps:
  - Estimating, in the output video buffer (23), a parameter  $P(Tr)$  associated with the occupancy level  $Tr$  of the radio buffer (25), and modifying the content of the input video buffer (21) as follows:
    - If the parameter  $P(Tr)$  associated with the estimated occupancy level  $Tr$  meets a predefined threshold, then transmitting the images contained in the input video buffer to the video coder on a rolling basis,
    - If the estimated parameter does not meet the predefined threshold, then modifying the content of the input video buffer by replacing the image  $F_{i+1}$  to be compressed with one or more images so as to adapt said content to the occupancy level  $Tr$  of the radio buffer.
2. Method according to Claim 1, characterized in that the value of the occupancy level  $Tr$  is estimated in real time, and this value is compared with a predefined occupancy level threshold value.
3. Method according to either of Claims 1 and 2, characterized in that the variation in the occupancy level  $Tr$  is estimated over a given time interval, and the content of the input video buffer is modified if the variation in the occupancy level  $Tr$  is continuous and exceeds a threshold value.
4. Method according to one of Claims 1 to 3, characterized in that the content of the input buffer is modified by replacing, one or more times, an image to be coded with a previously coded image from the images present in the image buffer.
5. Method according to one of Claims 1 to 3, characterized in that the quality of the image present in the buffer is checked and the current image is replaced using an image whose contours are sharpest.
6. Method according to one of Claims 1 to 3, characterized in that the current image in the input video buffer is replaced using an image containing information useful for rendering in the decoder, such as an image containing critical information in the selected images in the video input buffer.
7. Method according to one of Claims 1 to 3, characterized in that the image buffer is used by regularly compressing the different images in order to avoid a transmission outage.

8. Method according to one of Claims 1 to 6, characterized in that the quantization step of the video coder is modified for a replaced image in the buffer.
9. Method according to one of the preceding claims, characterized in that the content of a video stream in a radio system before transmission to the communication system is controlled.
10. System for controlling the content of a video stream within a device comprising a camera (20), a video coder (22) taking, at input, images from an input video buffer (21), and generating compressed images in accordance with a predictive model for an output video buffer (23), said output video buffer (23) transmitting these compressed images to a radio buffer (25), characterized in that:
  - The output video buffer (23) is configured so as to quantize the number of packets from the compressed images and to transmit it to a video stream control device (24),
  - The video stream control device (24) is configured so as to estimate a parameter  $P(Tr)$  associated with the occupancy level  $Tr$  of the radio buffer and to send an order to the input video buffer so that this modifies the images to be transmitted if the estimated associated parameter does not meet a predefined threshold by performing the steps of the method according to Claim 1.
11. System according to Claim 10, characterized in that the input video buffer (21) comprises an image selection filter in order to select the images having information useful for rendering in the decoder.