**Title:** VIDEO ENCODER AND VIDEO DECODER

A video encoder includes a frame rate downsampler for downsampling a high resolution sequence of frames and a video encoding module for encoding the low resolution sequence of frames. On the decoder-side this low resolution sequence of frames is used together with high resolution motion information for performing a motion information assisted frame interpolation. The high resolution motion information is derived on the encoder-side directly from the high resolution sequence before downsampling or is derived by analyzing the low resolution sequence of frames on the decoder-side by a subsequent scaling/interpolation of the low resolution motion information.
Published:
— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.
Video Encoder and Video Decoder

Field of the Invention

The present invention is related to video encoding/decoding and, particularly, to video processing in the context of low-capacity transmission channels.

Background of the Invention and Prior Art

In recent years industry saw an increased demand for digital distribution of visual content. Expectations and the requirements of the society greatly stressed the development of transport and storage mediums for an efficient storage and distribution of the digital content.

During the last two decades, methods and devices for efficient digital multimedia data transmission have emerged (MPEG) that efficiently employ some basic assumptions in digital signal processing and properties of the human central nervous system to efficiently reduce irrelevancy by means of lossy coding. Typical examples of these methods have been standardized in international bodies and they are widely known as MPEG-I, MPEG-2, H.261, H.263, H.264, etc...

Such modern video encoding algorithms heavily rely on motion detection and subsequent generation of motion patterns. Such a motion pattern or motion information is derived for a region of interest, which can be any irregular group of pixels determined by a certain motion detection algorithm, or which can be a macro block or a sub-macro block such as used in H.264 AVC. These motion information and, particularly, motion vectors are used for predictive coding so-called P-frames, so that in such modern encoding algorithms, a sequence of frames is transformed into an I-frame or "intra" frame, which is
self-contained. This means that such an I-frame is not predictively coded but is coded such that no additional information from different frames are required for decoding this picture.

Typical coding algorithms employ a block-wise Discrete Cosine Transform for luma samples, chroma samples, etc. Contrary thereto, so-called P-frames are encoded using prediction techniques. Particularly, a so-called motion-compensated prediction is applied, in which motion information is determined for a group of pixels. To this end, a group of pixels of an earlier picture is compared to the pixels of a later frame. A best match can be determined in that one can say that the group of pixels in a target picture is a moved representation of a group of pixels of the source picture. Such a moving vector describing this motion can be seen as a two-dimensional vector having a direction and a length and, therefore, points to a certain position in the target picture, which is different from a position in the source picture.

Then, the moved group of pixels is used as prediction data for the corresponding group of pixels in the target picture. Thus, the predicted group of pixels is subtracted from the actual group of pixels in the target picture. This subtraction results in residual pixel values, which are—in many cases—smaller than the original pixel values and are often equal to zero. Such small or zero pixel values can be efficiently encoded using a redundancy encoder such as a Huffman encoder or, as used in H.264 AVC, an arithmetic encoder.

Thus, for prediction frames, one has the motion vector information, the residual pixel information and, for pixels, which have not been subjected to a moving operation, i.e. so-called stationary pixels, residual values derived by subtracting the corresponding pixels from the preceding frame from the pixels of the actual frame.
Although all these modern video encoding algorithms result in a highly efficient data compression, the data rates are often too high for a required picture size, a given data transmission channel or a certain frame rate. In this context, frame rate means the number of frames per time unit such as per second or per minute. When one considers a transmission of an encoded sequence of frames, one will see that the data rate required increases with increasing picture size and, also, with increasing frame rate.

Stated in other words, to decrease a bit rate, one can reduce the size of the picture, i.e., the number of pixels within a frame, and/or one can reduce the frame rate, i.e., the number of frames in a unit time such as a second or a minute. Reducing picture size is often not tolerable, since increasing a reduced frame by pixel interpolation will result in a blurring of a picture, which is very annoying, when the viewer is very close to the display. However, it is exactly this situation which often happens in a wireless transmission scenario such as television broadcasting over the Internet or mobile phone resources. Even in non-transmission scenarios such as in the field of portable video players having a limited amount of storage resources, storing of a full resolution and full frame rate movie can quickly come in conflict with the storage resources of this portable device. In this situation it is also highly uncomfortable, when the frame resolution is reduced or when an interpolation is done from a reduced frame resolution to a higher frame resolution, since this interpolation results in blurring artefacts or in a kind of a granular picture where the viewer has the impression that she or he can distinguish each single pixel.

On the other hand, the reduction of the frame rate for reducing bit rate requirements or storage requirements can also not be satisfactory, since, on the decoder-side, the low frame rate may be too low so that the user does not
have the impression of a movie anymore but has the impression of a sequence of single frames which do not give the illusion of a continuous movement. On the other hand, when a frame interpolation is performed, which is based on straight-forward interpolation which is obtained, when a geometric or arithmetic mean value between the left and the right frame is calculated to generate an intermediate frame, the sharpness of the impression is highly degraded, since any motions within the sequence are smeared. This is due to the fact that, for calculating a certain pixel, the situation can come up that the left frame contributes a pixel of a moving object and the right frame contributes a picture of another stationary object. Thus, the motion is smeared, since the user does not receive a clear indication in the interpolated picture, whether a certain pixel belongs to a moving object or to a stationary object.

In view of this problem related to frame rate reduction, reducing the bit rate of an encoded video sequence is still problematic.

Summary of the Invention

It is an object of the present invention to provide an efficient but still high-quality coding/decoding scheme having low bit rate requirements and fulfilling high video quality demands.

This object is achieved by a video encoder for encoding a high resolution sequence of frames, comprising: a frame rate downsampler for downsampling the high resolution sequence so that a low resolution sequence of frames is obtained, the low resolution sequence of frames having a number of frames per time unit which is smaller than a number of frames per time unit of the high resolution sequence; and a video encoding module for encoding the low resolution sequence of frames.
In accordance with a further aspect of this invention, this object is achieved by a video decoder for decoding a low resolution sequence of frames, comprising: a video decoding module for decoding the low resolution sequence of frames to obtain a decoded low resolution sequence of frames; a motion information provider for providing high resolution motion information for a to be generated high resolution sequence of frames; and a frame interpolator for generating the high resolution sequence of frames using the decoded low resolution sequence of frames and the high resolution motion information.

Further aspects of the invention relate to computer programs implementing corresponding methods of encoding and decoding and also relate to an encoded video sequence generated by the video encoder and consumed by the video decoder, which can be transmitted on the fly via the Internet or which can be stored on a certain storage medium, or to a transmission system and a method of transmitting.

The present invention is based on the finding that reducing the frame rate is much more preferable over reducing the picture resolution or the picture size, when subjective quality assessments are conducted. The inventive method of frame rate downsampling is combined with an encoder-controlled or a decoder-controlled acquisition of high resolution motion information data which is, then, used within a decoder-side interpolator for creating synthetic frames, in which pixel groups or pixel blocks belonging to moving objects are reconstructed using the motion information, while remaining stationary pixels are generated using corresponding stationary pixels from an earlier or later frame.

In accordance with the present invention, a required amount of to be transmitted data is further reduced by exploiting
the inefficiency of the human central nervous system to
distinguish between real and artificially generated motion.
Furthermore, any (small) artefacts which might be incurred
by such an artificially generated motion due to the fact
that only motion information for the high resolution
sequence rather than residual values for the high resolution sequence are transmitted or generated on the
decoder-side, are much less annoying than artefacts
generated by reducing the resolution within a single frame.
In accordance with the present invention, a viewer always
has the impression that he is looking at a high resolution
full-size frame, and the user cannot distinguish between
artificially generated or real motion of objects within the
frame.

Brief Description of the Drawings

Subsequently, preferred embodiments of the present
invention are described in detail with reference to the
Figures, in which:

Fig. 1 illustrates a complete encoding/decoding system
in accordance with a first embodiment of the
present invention;

Fig. 2 illustrates a preferred embodiment of the decoder
or transmitter apparatus of the first embodiment;

Fig. 3 illustrates a preferred embodiment of a second
embodiment transmitter apparatus;

Fig. 4 illustrates a receiver or decoder apparatus being
associated to the first embodiment encoder
apparatus;

Fig. 5 illustrates a preferred decoder-embodiment
associated with the encoder apparatus in
accordance with the second embodiment of the present invention;

Fig. 6a a high resolution sequence of frames having a high resolution motion pattern;

Fig. 6b a low resolution sequence of frames having a low resolution motion pattern;

Fig. 7 a schematic representation of an inventive encoded video signal having an encoded low frame rate sequence preferably including low resolution motion patterns and, additionally, a high resolution motion information to be used by a decoder;

Fig. 8 illustrates a sequence of steps performed within an inventive decoder apparatus;

Fig. 9 illustrates a sequence of steps performed within a decoder-side frame interpolator in accordance with the present invention; and

Fig. 10 illustrates an example of creating a synthetic frame using motion information derived on the encoder-side or on the decoder-side.

Detailed Description of Preferred Embodiments

Before the inventive devices and methods are described in detail, a summary of features of preferred embodiments will be given.

Transmitter Side:

• Motion patterns analyzer 32
This block is responsible for analyzing the motion patterns between said amount of frames (typically equal, but not limited to downsampling factor) and making this data available to the other blocks.

Frame rate downsampler 22
• This block is responsible for reduction of the frame rate, so the video decoder could operate in more optimal conditions for a given limited transmission bandwidth conditions. The frame rate downsampler could re-use the data from the motion patterns analyzer in order to perform frame rate downsampling which is most satisfactory for the human central nervous system.

Motion pattern quantizers and coder 34
• This block is responsible for efficient storing of the motion patterns. Methods for efficient storing are known to people skilled in the art - a typical embodiment could be based (but not limited to) vector quantization of the motion patterns, with optional variable accuracy depending of the importance of a specific region inside the picture.

Downsampled video signal encoder 22
• This block is responsible for encoding of the video signal with reduced frame rate, and thus with considerably less requirements for the transmission bandwidth. In the example embodiment, a typical state-of-the-art video encoder, such as MPEG-4 or H.264 could be used, but it should be understood that the invention is not limited to this particular arrangement.

Receiver Side:
Downsampled video signal decoder 42
- This block is responsible for decoding of the downsampled video signal and providing data to the frame interpolator 46.

Motion patterns decoder 44
- This block is responsible for decoding and inverse quantization of the motion patterns. After reconstruction, e.g. original motion patterns are available to the frame interpolator.

Frame interpolator 46
- This block is responsible for reconstruction of the missing frames that were omitted by the frame rate downsampler on the transmitter side. Motion patterns are used to help reconstructing of the original (high quality) video signal.

Example Simple Apparatus

On the transmitter side, input video signal is presented to the apparatus. Motion patterns analyzer block analyses the movement by the application of motion detection algorithms. A set of motion patterns is made available as the output result of this block. The frame rate downsampler reduces the number of frames per second of the input signal and feeds the downsampled video signal encoder (which is a MPEG-4 AVC encoder in this example apparatus) with the reduced frame rate signal, thus requiring up to 50% less of the bit rate before downsampling. At the same time, motion pattern quantizer and coder block performs quantization and entropy coding of the motion vectors and this signal is being combined with the output of the video signal encoder in the bitstream multiplexer. The final result is a bitstream with both downsampled video signal, and motion patterns.
This bitstream could be stored on a storage medium, or transmitted over a broadcast medium to the receiver.

On the receiver side, the signal is first fed to the bitstream demultiplexer, which separates coded motion patterns from the coded video signal, and feeds proper modules with their relevant signal (i.e. coded motion patterns are being sent to the motion patterns decoder and the video signal is being sent to the downsampled video signal decoder, which is an MPEG-4 AVC decoder in this simple encoder apparatus example). The motion patterns decoder performs decoding and inverse quantization and reconstructs the motion patterns, while the downsampled video signal decoder decodes the downsampled video bitstream to a downsampled decoded video signal. The results of these processes, i.e. motion patterns and the decoded video signal are sent to the frame interpolator, which reconstructs the missing frames by utilizing motion patterns and motion reconstruction algorithms.

Example Advanced Apparatus

An example advanced apparatus is identical to the example simple apparatus, but the transmitter employs more advanced technologies in order to generate better "hints" for the receiver. In the next paragraph, an advanced transmitter will be described.

On the transmitter side, the input video signal is presented to the apparatus. The motion patterns analyzer block analyzes the movement by the application of motion detection algorithms. By using psycho-visual properties of the human brain, Regions of Interest (RoIs) are found in the picture. These regions are analyzed with more or finer quantized resulting output movement patterns compared to areas which have been determined as psycho-visually less relevant.
The set of motion patterns is made available as the output result of this block. The frame rate downsampler reduces the number of frames per second of the input signal also by taking motion patterns into consideration for generation of the most psycho-visually acceptable downsampled signal. The downsampler feeds the downsampled video signal encoder (which is a MPEG-4 AVC encoder in this example apparatus) with the reduced frame rate signal, thus requiring up to 50% less of the bit rate. At the same time, motion pattern quantizer and coder blocks performs advanced quantization and entropy coding of the motion vectors by applying different precision depending on each motion pattern's relevant, as determined by the motion patterns analyzer. This signal is being combined by the output of the video signal encoder in the bitstream multiplexer. Final result is a bitstream with both downsampled video signal, and motion patterns.

Additional Possible Arrangements

Below are presented possible additional arrangements with reduced complexity compared to the Example Apparatus, but they still incorporate inventive features.

Transmitter Variation: The motion analyzer could be integrated in the video signal encoder, as the part of the motion estimation algorithms typically employed in modern codecs. Motion vectors could be used as a basis for the motion patterns, and with additional coding they could be stored in the bitstream, as described in Fig. 3.

Receiver Variation: Also, in the decoder - motion vectors already available as the result of the encoding process and stored in the video stream could be re-used as the basis for motion pattern detection. Decoded motion vectors could be sent to the frame rate interpolator, and the frame rate interpolator would estimate motion patterns for the interpolated frames based on motion vectors. Motion
patterns would then be used to regenerate missing frames similar to the Example Apparatus. This apparatus does not need any additional processing in the transmitter on top of typical video coding (i.e. MPEG-2, MPEG-4, MPEG-4 AVC), and it is described in Fig. 5.

Increased Decoder Complexity Variation: A simple apparatus could be built by employing frame rate reduction and with a frame-rate interpolator in the decoder that could predict motion patterns just by analyzing of decoded frames (during decoding or after decoding). In this embodiment, no additional motion patterns are transmitted, and the frame rate interpolator does not need any additional data in the process.

Subsequently, the inventive devices and methods are described in more detail. Regarding Fig. 1, a video encoder feeding a storage and/or transport medium 10, which is indicated by reference numeral 12, and a video decoder fed by the storage and/or transport medium 10, which is illustrated at reference numeral 14 are shown. On the encoder-side, there exists an input receiving a high resolution sequence of frames. The input is indicated by reference numeral 20. The high resolution sequence of frames is input into a frame rate downsampler 22. The downsampler downsamples the high resolution sequence so that a low resolution sequence of frames is obtained at frame rate downsampler output 24. This downsampled low resolution sequence of frames is also indicated as "downsampled signal" in Fig. 1. Particularly, the low resolution sequence of frames has a number of frames per time unit, which is smaller than a number of frames per time unit of the high resolution sequence input at 20.

Furthermore, the inventive video encoder includes a video encoding module 26 which is indicated as "video signal encoder" in Fig. 1. The video signal encoder 26 is operative to encode the low resolution sequence of frames
output at 24. The encoded high resolution sequence of frames is output on line 28 indicated as "coded video signal" and input into a bitstream multiplexer 30. In a first embodiment of the present invention, the video encoder 12 also includes a motion information analyzer for analyzing motion information of the high resolution sequence at 20. Particularly, the motion information analyzer includes (in the preferred embodiment of Fig. 1) a motion patterns analyzer 32 outputting high resolution motion patterns and a motion pattern quantizer and encoder 34 for outputting coded motion patterns. The frame rate of the reconstructed sequence of frames is substantially similar and, thus, preferably within a range of plus/minus 20 percent of the frame rate of the original sequence and even more preferably identical to the frame rate of the original sequence.

The coded motion patterns, i.e., the motion information as well as the coded low resolution sequence of frames can be input into the bitstream multiplexer 30 which forms a multiplexed bitstream having side information including the coded high resolution motion patterns and having main information including the low resolution coded video signal sequence. This encoded high resolution sequence of frames is input into the storage and/or transport medium 10 and is received from there by a bitstream demultiplexer 40 adapted to separate the coded motion patterns from the coded video signal. The coded video signal which is, of course, a low resolution sequence of frames is input into a video decoder 42, which can have a straight-forward appearance, since the low resolution encoded signal is as self-contained as the coded motion patterns derived from the high resolution video signal in a preferred embodiment.

While the coded video signal is input into the video decoder 42, the coded motion patterns are input into a motion patterns decoder 44 which outputs decoded motion patterns, which are, together with a downsampled signal,
input into a frame interpolator 46. Particularly, the frame interpolator is operative for generating the high resolution sequence of frames at an output 48 using the decoded low resolution sequence of frames generated by module 42 and the high resolution motion information generated by module 44. Thus, the motion patterns decoder 44 in Fig. 1 acts as a motion information provider for providing high resolution motion information for a to be generated high resolution sequence of frames. The motion information provider is, therefore, operative to extract the motion information from the bitstream received by the bitstream multiplexer 40.

As will be outlined later on, the motion patterns decoder 44 acting as the motion information provider can also be operative to indeed extract any high resolution motion information from the low resolution encoded sequence by, for example, scaling or "interpolating" of low resolution motion vectors. While any stationary pixels can indeed be synthesized using any pixel-wise interpolation from neighboring frames, motion-affected groups of pixels will, in accordance with the present invention, be generated using a high resolution motion information as will be discussed later.

Fig. 2 illustrates the transmitter or encoder-part of Fig. 1. Particularly, the motion information analyzer including the motion patterns analyzer 32 and the motion pattern quantizer encoder 32 is operative to analyze motion information for a motion of a block from a frame to an adjacent frame in the high resolution sequence at input 20. In a preferred embodiment, the motion information includes an identification information for identifying a group of pixels and further includes a motion indication on a direction of movement of the group of pixels from one frame to another frame of the high resolution sequence. Preferably, motion information is derived from one frame to the next frame of the high resolution sequence, although
any kind of time-staggered or overlapping motion vectors are, of course possible and even preferred in certain environments.

Depending on the way of downsampling, the frame rate downsampler 22 can simply delete, for example, every other frame to obtain an almost 50% rate reduction. Such a frame rate downsampler would not touch the remaining frames after the deletion of frames. In other embodiments, the frame rate downsampler 22 can be implemented for doing a kind of an interpolation using stationary pixels of adjacent frames and creating a kind of a "synthetic" movement having a smoothed moving path.

Blocks 32 and 26 can even be connected via information line 50. This makes sense, when the video signal encoder 26 is based on a motion-compensated prediction, and when this device actually can use any motion information already-derived by the high resolution motion patterns analyzer. This would result in a more efficient encoder-device. However, there are other embodiments, in which the self-contained operations of the motion patterns analyzer 32 on the one hand and the video signal encoder 26 on the other hand are highly preferred, since such solutions can easily by implemented even when a video signal encoder is already in use and cannot be changed.

Motion information includes motion patterns, moving vectors, moving vector differences, absolute signaling of motion information any other way of signaling motion of an object via, for example, a parameterized trajectory.

Fig. 4 illustrates the receiver 14 from Fig. 1. The video decoder for decoding a low resolution sequence of frames receives, as an input, the coded bitstream having a coded representation of the low resolution sequence of frames. The signal on line 60 is input into the bitstream multiplexer 40 and the bitstream multiplexer distributes
the coded motion patterns to the motion patterns decoder 44. Thus, the bitstream multiplexer 40 has the functionality of a motion information provider for providing high resolution motion information for generating a high resolution sequence of frames output at line 62. Furthermore, the inventive decoder includes a frame interpolator for generating the high resolution sequence of frames using a decoded low resolution sequence of frames at line 64, and using the high resolution motion information at line 66.

Fig. 3 illustrates a second embodiment of the present invention, in which only the coded video signal is transmitted without separate high resolution motion information. Nevertheless, the motion patterns analyzer 32 can be provided to support the video signal encoder encoding the low resolution sequence output by a block 22. However, it is even preferred to not use at all the motion patterns analyzer 32 so that the Fig. 3 embodiment only includes a frame rate downsampler and a straight-forward video signal encoder which inputs a video signal into the bitstream multiplexer without any additional high resolution motion information. Then, the whole computational requirements are on the decoder-side, since such a decoder will have to derive high resolution motion information from the decoded or encoded low resolution sequence.

This can be seen in the Fig. 5 embodiment, which is a decoder matching to the Fig. 3 encoder not having the motion patterns analyzer 32. Thus, the Fig. 5 decoder additionally includes the motion information provider 70, which derives high resolution motion information from a decoded low resolution sequence or an encoded low resolution sequence, which preferably has a low resolution motion information in explicit form. The high resolution motion vectors and the low resolution sequence indicated as "downsampled signal" are input into the frame interpolator
46 so that the frame interpolator 46 can generate the reconstructed high quality signal.

Subsequently, preferred embodiments of the inventive methods will be described in more detail. Fig. 6a illustrates a high resolution sequence of frames having a high frame rate. Frames are indicated at 81, 82, 83, 84, 85, 86, 87. Exemplarily, it is indicated that frames 81, 82, 83 contain a specific group of pixels 90 moving from a certain place in frame 81 to another place in frame 82 and finally arriving at a final place in frame 83. Movement of this area 90 is indicated by motion information 92. Particularly, motion information 92 gives the motion from frame 81 to frame 82. Additionally, further motion information 94 is given illustrating the movement of area 90 from frame 82 to frame 83. The motion information items 92 and 94 are a high resolution motion pattern.

When Fig. 6b is considered, it can be seen that the downsampler has performed a downsampling operation, in which every other frame has been deleted. Now, this low resolution sequence of frames can be processed using a straight-forward video decoder. Naturally, such a decoder includes a motion information analyzer. Alternatively, an additionally self-contained motion information analyzer can be provided which can then control any specific encoder-related tasks. Such a motion pattern analysis will result in a low resolution motion pattern illustrated at 96 at Fig. 6b. It becomes clear that the motion vector 96 does not give any information where to place area 90 in a to be reconstructed block between blocks 81 and 83.

Low resolution motion pattern information 96 is used for video-encoding the low resolution sequence of frames. This encoded low frame rate sequence is illustrated at 100 in Fig. 7. Item 100 is accompanied by item 102 having the high resolution motion information such as 92 and 94 of Fig. 6a. Fig. 8 illustrates a general sequence of steps to be taken
when decoding in accordance with the present invention. First of all, the decoding of the received low frame rate sequence is performed as illustrated at 110 in Fig. 8. Then, high resolution motion information is acquired (112), which will be done by the motion information provider which can, for the first embodiment, extract high resolution motion information from the bitstream, i.e., item 102 from Fig. 7 or which can process low resolution motion pattern 96 included in item 100 or which can process decoded frames to obtain low resolution motion pattern information, which can then be used to obtain high resolution information. Finally, a synthetic frame is created between two frames of a low frame rate sequence as indicated in step 114, when a low resolution sequence is considered, which was decimated by deleting every other frame. When, however, an even more dramatic downsampling was performed, i.e., that the downsampler deleted every second and every third frame, the deleted third frame can be artificially created using information from the artificially created second frame and the low resolution fourth frame. Naturally, one could also synthesize less frames than were deleted so that a medium frame rate between the original high frame rate or the transmitted low frame rate is recreated on the decoder-side.

Thus, in principal, one can artificially create as many intermediate frames as required, although the deviation of the artificially created frames to the original frames will increase with an increasing number of frames to be synthesized on the decoder-side.

Fig. 9 illustrates a preferred sequence of steps to be used for the inventive embodiments. It is to be noted that Fig. 9 does not depend on whether the high resolution motion information was acquired by extracting from the bitstream, or whether high resolution motion information was generated by analyzing the received low resolution signal without any additional high resolution information.
First of all, areas in a frame, for which motion information exists, are determined as indicated at 120. Preferably, this frame is a low frame rate sequence frame. Alternatively, this frame can also be a synthetic frame generated before processing the actual frame.

Then, in step 122, the detected areas are moved to target positions indicated by the motion information so that pixels in a synthetic frame are generated. Straight-forward video decoders now retrieve the residual signal which can then be combined with the predicted signal, i.e., the result of block 122, to obtain the reconstructed moved areas in the synthetic frame. In accordance with the present invention, and for bit rate savings reasons, only motion information are present and residual pixels are not present. Therefore, the pixels generated by block 122 will have some errors, which is, however, much less annoying compared to a simple straight-forward interpolation without motion information.

Stationary pixels are then filled in the synthetic frame as indicated at 124. This filling of stationary pixels has to occur for those pixels, for which it was not possible to detect motion information in the data provided by the motion information provider. These stationary pixels can be taken from the source or originating frame or can even be taken from the next low resolution frame, i.e., the frame following the synthetic frame, when a 50 % decimation as indicated in Fig. 6b compared to Fig. 6a has been performed.

The output of step 124 is a synthetic frame having moving areas and stationary pixels. The only pixels which can be missing in the synthetic frame generated by step 124 are the pixels which are uncovered by moving a certain area in the left frame to a certain position in the synthetic frame. To fill this gap, pixel values for synthetic frame
pixels corresponding to pixels of the detected areas are acquired (126) so that a complete synthetic frame has been established. Preferably but not necessarily, the complete synthetic frame output by step 126 is subjected to a smoothing operation as indicated at 128 in Fig. 9 so that a smoothed synthetic frame is obtained.

This smoothing is, preferably, not performed in time, i.e., over the frames, but is preferably performed within the frame to eliminate any blocking artifacts, i.e., any visible boarders within the synthetic frame, which are not caused by the picture content but which were caused during building the synthetic frame.

Fig. 10 illustrates a more detailed example how the synthetic frame is generated in a preferred embodiment. Furthermore, the Fig. 6b embodiment is applied, in which there are low resolution sequence frames 81 and 83, and a high resolution sequence frame 82 between those frames 81 and 83 was deleted by an encoder-side downsampler. The task of the process in Fig. 9 is to generate the synthetic frame 82'. To this end, explicitly signaled high resolution motion information 92 is used to move the group of pixels 90 to a different place from frame 81 to 82'. Information where to move the group of pixels 90 can, however, also be derived from the low resolution motion pattern 96 by scaling the low resolution motion vector. When one synthetic frame has two low resolution frames as direct neighbors, motion vectors can be scaled to half length. Then, the scaled motion vector points to the right position for the synthetic frame.

Nevertheless, by moving block 90 away from its position in the left low resolution frame rate frame there will occur an uncovered area 150, which cannot be filled with pixels from the low resolution frame 81, which was the source of the moving block. Although, all other stationary pixels, i.e., pixels which are not within the moving area and which
are not within the uncovered area can be taken from frame 81, frame 81 does not have pixels to fill in the uncovered area 150. However, the right low resolution frame rate frame 83 may have stationary pixels at the position corresponding to the uncovered area 150 in synthetic frame 82' so that the pixels can be taken from the right low resolution frame rate frame 83.

Naturally, stationary pixels in the synthetic frame 82' can also be calculated by interpolating between corresponding pixels in low resolution sequence frames 81 and 83. When an interpolation is performed which calculates an arithmetic mean value, and when the pixels belonging to a moving area 90 in the low resolution frame are set to zero for this kind of interpolation, then filling of the uncovered area by the pixels of the right low resolution frame rate frame automatically takes place.

When it is, for any reason, not possible to use the right low resolution sequence frame rate frame pixels for the uncovered area 150, then the pixel for the uncovered area can also be generated by using corresponding pixels from frames which are not direct neighbors in the sequence or by using neighboring stationary pixels within the same frame or can even be randomly generated and subsequently smoothed.

Depending on certain implementation requirements of the inventive methods, the inventive methods can be implemented in hardware or in software. The implementation can be performed using a digital storage medium, in particular a disk, DVD or a CD having electronically readable control signals stored thereon, which cooperate with a programmable computer system such that the inventive methods are performed. Generally, the present invention is, therefore, a computer program product with a program code stored on a machine readable carrier, the program code being operative for performing the inventive methods when the computer
program product runs on a computer. In other words, the inventive methods are, therefore, a computer program having a program code for performing at least one of the inventive methods when the computer program runs on a computer.

The above-described embodiments are merely illustrative for the principles of the present invention. It is understood that modifications and variations of the arrangements and the details described herein will be apparent to others skilled in the art. It is the intent, therefore, to be limited only by the scope of the impending patent claims and not by the specific details presented by way of description and explanation of the embodiments herein.
What is claimed is:

1. Video encoder for encoding a high resolution sequence of frames, comprising:

5 a frame rate downsampler for downsampling the high resolution sequence so that a low resolution sequence of frames is obtained, the low resolution sequence of frames having a number of frames per time unit which is smaller than a number of frames per time unit of the high resolution sequence; and

a video encoding module for encoding the low resolution sequence of frames.

2. Video encoder of claim 1, further comprising:

a motion information analyzer for analyzing motion information of the high resolution sequence.

3. Video encoder of claim 2, wherein the motion information analyzer is operative to analyze a motion information for a motion of a block from a frame to an adjacent frame in the high resolution sequence.

4. Video encoder of claim 2, wherein the motion information analyzer is operative to generate motion information for a group of pixels, the motion information including an indication for identifying the group of pixels and information on a direction of movement of the group of pixels from one frame to another frame of the high resolution sequence.

5. Video encoder of claim 2,

wherein the motion information analyzer is operative to generate motion information for a group of pixels, the motion information indicating a movement of a
group from one frame to the next frame in the sequence.

6. Video encoder of claim 1, in which the frame rate downsampler is operative to delete frames of the high resolution sequence.

7. Video encoder of claim 6, in which the frame rate downsampler is operative to not change frames of the high resolution sequence which remain after deletion of frames.

8. Video encoder of claim 2, further comprising an output interface for outputting the encoded low resolution sequence of frames and the analyzed motion information of the high resolution sequence of frames.

9. Video encoder of claim 2, in which the frame rate downsampler is operative to use the analyzed motion information for generating the low resolution sequence of frames, wherein frames of the low resolution sequence of frames are different from corresponding frames in the high resolution sequence of frames.

10. Video encoder of claim 9, in which the frame rate downsampler is operative to analyze the high resolution motion information and to smooth motion of areas from one frame to the next frame of the low resolution sequence.

11. Video encoder of claim 9, in which the frame rate downsampler is operative to perform a frame rate downsampling operation, which is most satisfactory for the human central nervous system.

12. Video encoder of claim 2, in which the motion information analyzer is operative to compress the high resolution motion information.
13. Video encoder of claim 12, in which the motion information analyzer is operative to identify regions of higher importance and regions of a comparatively smaller importance for a visual impression; and in which the motion information analyzer is further operative to quantize motion information on the more important region finer than motion information on a less important region.

14. Video encoder of claim 2, in which the video encoding module is operative to perform a motion compensation during encoding and in which the video encoding module is operative to use motion information generated by the motion information analyzer during the motion compensation.

15. Video decoder for decoding a low resolution sequence of frames, comprising:

- a video decoding module for decoding the low resolution sequence of frames to obtain a decoded low resolution sequence of frames;
- a motion information provider for providing high resolution motion information for a to be generated high resolution sequence of frames; and
- a frame interpolator for generating the high resolution sequence of frames using the decoded low resolution sequence of frames and the high resolution motion information.

16. Video decoder of claim 15, in which in input signal further includes the high resolution motion information, and
wherein the motion information provider is operative to extract the high resolution motion information from the input signal.

17. Video decoder of claim 16,

wherein the high resolution motion information is quantized and encoded, and

wherein the motion information provider is operative to decode and dequantize the high resolution motion information.

18. Video decoder of claim 15,

wherein the motion information provider is operative to analyze the encoded or decoded low resolution sequence of frames for a decoder-side generation of high resolution motion information.

19. Video decoder of claim 18, in which the motion information provider is operative to extract low resolution motion information and to derive high resolution motion information from the low resolution motion information.

20. Video decoder of claim 19, in which the motion information provider is operative to derive the high resolution motion information by interpolation or scaling according to a number of synthetic frames to be inserted between two adjacent low resolution sequence frames, the number being greater than or equal to 1.

21. Video decoder of claim 18,
in which the frame interpolator is operative to generate a synthetic frame positioned with respect to time between two adjacent frames,

by detecting a group of pixels in a frame for which high resolution motion information exists,

by placing the group of pixels at a position in the synthetic frame indicated by the high resolution motion information,

by deriving stationary pixels in the synthetic frame using corresponding pixels of at least an adjacent frame, and

by determining remaining pixels of the synthetic frame at the position of the group of pixels in the frame in which the group of pixels was detected.

22. Video decoder of claim 21, in which the frame interpolator is operative to determine the remaining pixels of the synthetic frame by using corresponding pixels of a frame not used in the step of detecting, by using neighboring pixels from a synthetic frame or by generating random pixel values or by using a combination of the above-mentioned methods.

23. Video decoder of claim 21, wherein the step of detecting is performed on a low resolution frame or a preceding synthetic frame, when at least two synthetic frames are generated between two adjacent low resolution frames.

24. Video decoder of claim 21, in which the frame interpolator is operative to smooth the synthetic frame to obtain a smoothed synthetic frame to be placed into the high resolution sequence of frames.
25. Video decoder of claim 15, in which the high resolution motion information includes motion vectors or motion vector differences associated with areas in a low resolution frame and indicating motions of the areas to an adjacent to be generated synthetic frame or to a further to be generated synthetic frame, when at least two synthetic frames are to be generated between two adjacent low resolution frames.

26. Method of encoding a high resolution sequence of frames, comprising:

downsampling the high resolution sequence so that a low resolution sequence of frames is obtained, the low resolution sequence of frames having a number of frames per time unit which is smaller than a number of frames per time unit of the high resolution sequence; and

encoding the low resolution sequence of frames.

27. Method of decoding a low resolution sequence of frames, comprising:

decoding the low resolution sequence of frames to obtain a decoded low resolution sequence of frames;

providing high resolution motion information for a to be generated high resolution sequence of frames; and

generating the high resolution sequence of frames using the decoded low resolution sequence of frames and the high resolution motion information.

28. Encoded video signal having an encoded low resolution sequence of frames and, additionally, high resolution motion information to be used for generating a high resolution sequence of frames.
29. Computer program having a program code for performing the method of claims 27 or 28, when running on a computer.

30. Transmission system comprising:

   a video encoder for encoding an original high resolution sequence of frames, the video encoder comprising:

   a frame rate downsampler for downsampling the high resolution sequence of frames so that a low resolution sequence of frames is obtained, the low resolution sequence of frames having a number of frames per time unit which is smaller than a number of frames per time unit of the high resolution sequence; and

   a video encoding module for encoding the low resolution sequence of frames; and

   a video decoder for decoding the low resolution sequence of frames, comprising:

   a video decoding module for decoding the low resolution sequence of frames to obtain a decoded low resolution sequence of frames;

   a motion information provider for providing high resolution motion information for a to be generated high resolution sequence of frames having a frame rate substantially similar to the frame rate of the original high resolution sequence of frames; and

   a frame interpolator for generating the to be generated high resolution sequence of frames
using the decoded low resolution sequence of frames and the high resolution motion information.

31. Method of transmitting video information, comprising:

encoding an original high resolution sequence of frames, the step of encoding comprising:

- downsampling the high resolution sequence so that a low resolution sequence of frames is obtained, the low resolution sequence of frames having a number of frames per time unit which is smaller than a number of frames per time unit of the high resolution sequence; and

- encoding the low resolution sequence of frames; and

decoding the low resolution sequence of frames, the step of decoding comprising:

- decoding the low resolution sequence of frames to obtain a decoded low resolution sequence of frames;

- providing high resolution motion information for a to be generated high resolution sequence of frames having a frame rate substantially similar to the frame rate of the original high resolution sequence of frames; and

- generating the to be generated high resolution sequence of frames using the decoded low resolution sequence of frames and the high resolution motion information.
FIG 2

Motion Patterns Analyser

Frame Rate Downsampler

Downsampled Signal

Motion Pattern Quantizer & Coder

Video Signal Encoder

Coded Video Signal

Bit Stream Multiplexer

FIG 3

Motion Patterns Analyser

Frame Rate Downsampler

Motion Vectors

Downsampled Signal

Video Signal Encoder

Coded Video Signal

Bit Stream Multiplexer
FIG 6A

high resolution motion pattern

sequence of frames (high frame rate)

FIG 6B

low resolution motion pattern

sequence of frames (low frame rate)

FIG 7

high resolution motion information 90,94

encoded low frame rate sequence (MPEG4,H264)
- low resolution motion pattern 96,
- encoded I-,P- frames etc.

FIG 8

decoding of low frame rate sequence 110

acquiring of high resolution motion information 112

creating frame between two frames of low frame rate sequence 114

SUBSTITUTE SHEET (RULE 26)
FIG 9

determining areas in low frame rate frame, for which motion information exists

moving the detected areas to target positions indicated by the motion information in synthetic frame

synthetic frame only having moving areas

filling of stationary pixels using a low frame rate frame

synthetic frame having moving areas and stationary pixels (not having "uncovered areas")

acquiring pixel values for synthetic frame pixels corresponding to pixel of the detected areas

complete synthetic frame

smoothing of the synthetic frame

smoothed synthetic frame
- stationary pixels are taken from left frame or from right frame or are calculated as a mean value of left and right frames;

- uncovered area is filled using pixels from the frame, from which the detected areas do NOT originate (here: right frame);