



(51) International Patent Classification:

H04N 19/124 (2014.01) H04N 19/174 (2014.01)  
H04N 19/167 (2014.01) H04N 19/33 (2014.01)

(21) International Application Number:

PCT/GB2024/053051

(22) International Filing Date:

06 December 2024 (06.12.2024)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

2318815.4 08 December 2023 (08.12.2023) GB

(71) Applicant: **V-NOVA INTERNATIONAL LTD** [GB/G-B]; Level 2, 20 Eastbourne Terrace, Paddington, London W2 6LG (GB).

(72) Inventors: **MORGAN, Harry**; c/o V-Nova International Ltd, Level 2, 20 Eastbourne Terrace, Paddington, London Greater London W2 6LG (GB). **CICCARELLI, Lorenzo**; c/o V-Nova International Ltd, Level 2, 20 Eastbourne Terrace, Paddington, London Greater London W2 6LG (GB). **MOCKFORD, Kevin**; c/o V-Nova International Ltd, Level 2, 20 Eastbourne Terrace, Paddington, London Greater London W2 6LG (GB).

(74) Agent: **GILL JENNINGS & EVERY LLP**; The Broadgate Tower, 20 Primrose Street, London Greater London EC2A 2ES (GB).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— with international search report (Art. 21(3))

(54) Title: ENHANCEMENT CODING IMPROVEMENTS AND INTEGRATION

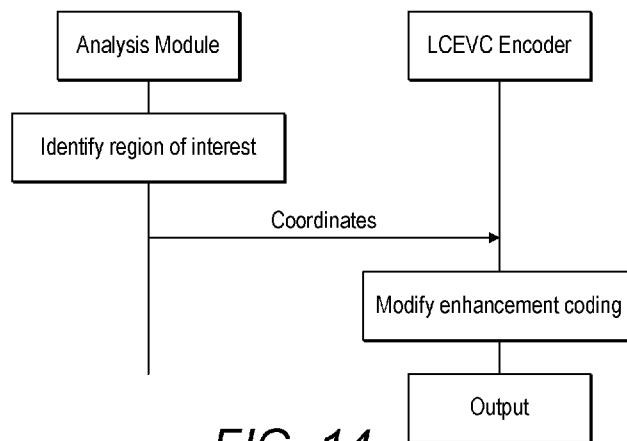


FIG. 14

(57) Abstract: The present disclosure relates to video encoding and particularly to enhancement encoding in a system on a chip, SoC, such as an SoC applied to a camera. There are provided systems on a chip, SoCs, comprising a plurality of enhancement encoders and a base codec configured to serve the plurality of enhancement encoders. There are provided methods for operating such SoCs; methods for providing an encoded video signal by determining a region of interest, RoI, in an input video signal; and methods for providing an RoI in a to-be-encoded input video signal. Further SoCs, methods, processing devices, computer programs, and the like are also provided.



## ENHANCEMENT CODING IMPROVEMENTS AND INTEGRATION

### BACKGROUND

Recent improvements in video coding technology have included the concept of hierarchical video coding. Examples include VC-6, standardised at SMPTE as ST  
5 2117, and LCEVC, standardised at MPEG as MPEG-5 Part II. Typically, these hierarchical encoding schemes use multiple resolution levels and an encoder (or encoding module) associated with each resolution level.

A standard specification for LCEVC is provided in the Text of ISO/IEC 23094-2 Ed  
1 Low Complexity Enhancement Video Coding published in November 2021, and  
10 many possible implementation details of LCEVC are described in patent publications WO 2020/188273 and WO 2020/188229. Each of these earlier documents is incorporated here by reference.

In these new coding schemes, efficiencies and optimisations are sought to reduce data size and/ or processing requirements while improving picture quality of the  
15 ultimately reconstructed image.

Broadly speaking, LCEVC enhances the reproduction fidelity of a decoded video after encoding and decoding using an existing codec. This is achieved by combining a base layer with an enhancement layer, where the base layer contains the video encoded using the existing codec, and the enhancement layer indicates  
20 a residual difference between the original video and an expected decoded video produced by decoding the base layer using the existing codec. The enhancement layer can be combined with the decoded base layer to more accurately reproduce the original video.

It is a flexible, adaptable, highly efficient and computationally inexpensive coding  
25 technology and format which combines a video coding format, a base codec, (e.g. AVC, HEVC, or any other present or future codec) with an enhancement level of coded data, encoded using a different technique. The technology uses a down-sampled source signal encoded using a base codec to form a base stream. An enhancement stream is formed using an encoded set of residuals which correct

or enhance the base stream for example by increasing resolution or by increasing frame rate. There may be multiple levels of enhancement data in a hierarchical structure.

5 Opportunities and improvements are sought to integrate these technologies in real-time scenarios. For example, the technologies allow practical improvements in contemporary video streaming usages.

### **SUMMARY**

The following sets out examples of integration of enhancement encoders into existing ecosystems, such as security cameras and other similar infrastructure. In  
10 many video scenarios, for example those not media oriented such as security cameras, most of the video feed is unimportant and only a minor part is important. However, in that important region, details are all important perhaps even essential. Examples include detecting faces for identification, detecting vehicle number  
15 plates for vehicle identification and tracking or detail in evidence-based scenarios where the activity needs to be recorded or monitored, but background is less relevant.

According to a first aspect of the invention, there may be provided a system on a chip, SoC, comprising a plurality of enhancement encoders and a base codec  
20 configured to serve the plurality of enhancement encoders, wherein: the SoC is configured to receive a plurality of input video streams at a first resolution; the base codec is configured to encode the plurality of input video streams to provide a plurality of base coded streams at a second resolution lower than the first  
25 resolution; and, each of the plurality of enhancement encoders is configured to encode a respective one of the plurality of base coded streams to provide a plurality of enhancement coded streams corresponding to the plurality of input video streams, the enhancement coded streams being at the first resolution.

In this way, the enhancement encoders may be configured to add enhancement to the base encoded video to create the encoded output video streams at the same resolution as the input video streams, while the base codec can sustain a

higher bitrate (i.e., throughput) at a relatively lower resolution without sacrificing the resulting quality of the output stream. In the implementations of the present disclosure, this provides improvements to the coding/processing capability per unit area of the chip over SoCs in the art. In other words, by configuring a base  
5 codec to serve a plurality of enhancement encoders in accordance with the present disclosure, the encoding/decoding hardware may handle more input streams, and higher resolution/frame rate streams, for a given chip area. In the implementations of the present disclosure, this also provides the opportunity for greater coding pipeline re-use, since encoding/decoding resources (i.e. parts of  
10 pipeline) may be reused for different tasks in order to optimise efficiency in a data compression workflow pipeline.

In an example of the first aspect, the SoC comprises a buffer configured to store one or more to-be-encoded input video streams during at least the encoding of one of the plurality of input video streams by the base codec. In this way, by  
15 buffering the raw data corresponding to one or more not-yet-encoded input streams, a base codec may be controlled to switch between encoding different streams at a rate where the base codec can encode at a quicker rate than the information being stored in the buffer.

In an example of the first aspect, the base codec is an AV1 codec or an AVC codec  
20 or an HEVC codec, and the plurality of enhancement encoders are LCEVC encoders. In another example, the base and enhancement layers may be arranged in a hierarchical structure. In another example, the SoC may be further configured to provide the base coded stream, such as in addition to the enhancement coded streams. In this way, enhancement encoding may be  
25 employed to provide enhanced compression to reduce bitrate, which is advantageous in high motion or otherwise challenging scenarios (e.g., in wearable cameras).

In an example of the first aspect, the plurality of enhancement encoders comprises  
30 three enhancement encoders or four enhancement encoders, and the plurality of input video streams comprises four input video streams. In another example, the plurality of enhancement encoders comprises three enhancement encoders, and

the plurality of input video streams comprises three input video streams. In another example, the number of enhancement encoders is at least equal to the number of input streams. In this way, the enhancement encoding implemented in the SoC may provide a high performance gain density compared to using a base encoder or several such base encoders alone, particularly when there are multiple  
5 input streams, as the base may be operated at a lower resolution.

In an example of the first aspect, the SoC is configured for use in one or more of a camera, a security camera, an automotive camera, and a body worn camera. In some examples, the SoC is configured to receive the plurality of input streams  
10 from a respective plurality of such cameras. In another example, the SoC may be configured for use in an edge device, wherein the edge device is configured to handle or otherwise receive each input stream of a plurality of input streams from each respective camera in a plurality of local cameras. In another example, the SoC may be configured to receive the plurality of input streams from a respective  
15 plurality of views captured by a multi-view camera. Optionally, different views may have different lens angles (e.g., a first view may be provided by a wide-angle lens, a second view may be provided by an ultra-wide angle lens, a third view may be provided by a fisheye lens, a fourth view may be provided by a narrow-angle lens, and so on). In this way, an enhancement encoder within a (video encoder) SoC  
20 may be integrated into challenging practical applications involving cameras, such as in a security scenario.

In an example of the first aspect, the base codec is configured to provide a set of coordinates corresponding to a region of interest, RoI, to at least one of the plurality of enhancement encoders; and, the RoI is usable to modify the encoding  
25 of the respective one of the plurality of the base coded streams by the at least one of the plurality of enhancement encoders to prioritise enhancement of the base coded stream at the RoI. In another example, the set of coordinates may have been received or otherwise provided to the base codec by an API or an AI object analysis model. In another example, the base codec, the API, or the AI object  
30 analysis model may insert the set of coordinates into the base coded stream such that the set may be extracted by one or more of the enhancement encoders.

Accordingly, the set of coordinates may be extractable by the enhancement encoding operation. In this way, the base coded stream may be adapted so as to permit RoI enhancement encoding by one or more enhancement encoders in the SoC, which in turn may provide residuals for particular areas of interest without  
5 needing to transmit residuals for the entire scene.

In an example of the first aspect, the SoC comprises a region of interest, RoI, identification module configured to provide a set of coordinates corresponding to the RoI to the base codec and/or at least one of the plurality of enhancement encoders; and, the RoI is usable to modify the encoding of the respective one of  
10 the plurality of the base coded streams by the at least one of the plurality of enhancement encoders to prioritise enhancement of the base coded stream at the RoI. In another example, the RoI identification module may be a separate module to the enhancement encoders and/or to the base codec. In another example, the RoI identification module may be the base codec, an API, an AI object analysis  
15 model, or another processing device or CPU module. In another example, the RoI identification module may insert the set of coordinates into the base coded stream such that the set may be extracted by one or more of the enhancement encoders. Accordingly, the set of coordinates may be extractable by the enhancement encoding operation. In this way, the one or more enhancement  
20 encoders may be provided with an indication of the area corresponding to the RoI by the RoI identification module, in order that the one or more enhancement encoders may provide encoded residuals in a specified area and therefore a higher resolution at the RoI.

In an example of the first aspect, at least one of the plurality of enhancement  
25 encoders is configured to determine a set of coordinates corresponding to a region of interest, RoI, based on one or more settings of a residual quantisation process of the base codec; and, the RoI is usable to modify the encoding of the respective one of the plurality of the base coded streams by the at least one of the plurality of enhancement encoders to prioritise enhancement of the base coded stream at  
30 the RoI. In another example, the settings of the residual quantisation process of the base codec may comprise one or more quantisation step widths, wherein

areas/portions of the base layer having a relatively lower quantisation step width may correspond to portions of the base coded stream corresponding to the Rol, and/or areas/portions of the base layer having a relatively higher quantisation step width may correspond to portions of the base coded stream not corresponding to the Rol. In another example, a step width applied during the encoding process may be modified (e.g. by a residual selection block or by applying a priority map during a residual quantisation process) to selectively encode (a) residual value(s) corresponding to the Rol. In this way, one or more enhancement encoders may be configured to provide (more) residuals in the area provided by the coordinates, and therefore a higher resolution for a particularly desired portion of the input video.

In another example, the Rol may be usable to modify an enhancement coding operation to prioritise enhancement of a base coded stream at the Rol by overriding a default enhancement coding operation according to the Rol to prioritise enhancement of a base video signal at the Rol. In some examples, the default enhancement coding operation may include an enhancement coding operation configured to encode residuals for an input video without reference to an Rol and/or without reference to a priority map. In some examples, the default enhancement coding operation may include an enhancement coding operation applied in a scenario where no object of interest is detected in the input video stream. In some examples, overriding the default enhancement coding operation may include varying, altering or updating the default encoding parameters of said operation to prioritise residual encoding of identified Rols. In some examples, overriding the default enhancement coding operation may include overriding a default enhancement coding operation according to the determined Rol to prioritise enhancement of a base coded stream at the Rol over the base coded stream not at the Rol.

In some examples, the modification of an enhancement coding operation to prioritise enhancement of a base video signal at the Rol may be implemented by an enhancement codec configured to perform the enhancement coding operation. In other words, the means for modifying the enhancement process according to

the present disclosure may be included as part of the (e.g. standalone) enhancement codec. In this way, enhancement modification may be independent of any given base codec and may be implemented without reference to, or without any requirement for, a base codec.

5 According to a further aspect of the invention, there may be provided a camera system comprising the SoC of the first aspect, wherein the camera system comprises a plurality of cameras, and each camera of the plurality of cameras is configured to: capture one of the plurality of input videos streams, and send the captured one of the plurality of input streams to the SoC. In an example, the  
10 camera system may comprise one or more downsampler modules configured to downsample (respective) one or more of the plurality of input video streams prior to the plurality of input video streams being passed to the base codec. In a further example, the camera system may comprise one or more upsampler modules configured to upsample an output of the base codec to generate an upsampled  
15 rendition of the output of the base codec to be input to one or more of the enhancement encoders. In some examples, the output of the base codec is a decoded rendition of a base encoded signal, wherein typically the output of the base codec is at a lower resolution than the input video signal(s).

According to a second aspect of the invention, there may be provided a multi-  
20 resolution camera comprising the SoC of the first aspect and any examples thereof, wherein the camera is configured in use to provide the plurality of input video streams to the SoC. In another example, the camera may be configured to output the plurality of enhancement coded streams corresponding to the plurality of input video streams, the enhancement coded streams being at the first  
25 resolution. In another example, the camera may be configured to output two video streams decodable to derive three video streams at respective three different resolutions, wherein the two video streams may include a first stream and a second stream, the first stream may include an enhancement coded stream at high resolution and a base coded stream at medium resolution, and the second  
30 stream may include a stream at low resolution.

In a specific example of utilising enhancement coding in a security camera environment, a camera system may be provided, the camera system being configured to output two video streams. Each output stream may be at a different level of quality and each output may be encoded. For example, the first quality  
5 may be (high definition) HD and the second may be UHD (ultra-high definition).

In a first example, a base codec may encode HD, then an LCEVC encoder may be configured to encode UHD. Within that UHD output, there may be configured to include two streams, an HD base and the HD+LCEVC enhancement. Alternatively, the base may encode standard definition (SD) with LCEVC encoding  
10 HD and UHD.

In a further alternative, the same base stream may be used, with two output streams being provided by a respective LCEVC encoder configured to output HD and UHD respectively. In this scenario, the base may be output at SD to be combined with HD from one encoding step and LCEVC from another encoding  
15 step. Of course the same LCEVC encoder can be used to encode each of the HD and UHD enhancements separately from the same base coder.

In a final alternative, the UHD output may be provided from an HD base (i.e. HD base + UHD enhancement) and the HD may be provided from an SD base (i.e. SD base + HD enhancement).

20 According to a third aspect of the invention, there may be provided a method for operating a system on a chip, SoC, the SoC comprising a plurality of enhancement encoders and a base codec configured to serve the plurality of enhancement encoders, and the method comprising: receiving, by the SoC, a plurality of input video streams at a first resolution; encoding, by the base codec, the plurality of  
25 input video streams to provide a plurality of base coded streams at a second resolution lower than the first resolution; and, encoding, by each of the plurality of enhancement encoders, a respective one of the plurality of base coded streams to provide a plurality of enhancement coded streams corresponding to the plurality of input video streams, the enhancement coded streams being at the first  
30 resolution. In this way, the enhancement encoders may be operated so as to add

enhancement to the base encoded video to create the encoded output video streams at the same resolution as the input video streams, while the base codec may be operated faster and at a relatively lower resolution without sacrificing output resolution.

- 5 In an example of the third aspect, the SoC comprises a buffer, and the method further comprises: storing, by the buffer, one or more to-be-encoded input video streams during at least the encoding of one of the plurality of input video streams by the base codec. In this way, the base codec may be controlled (e.g. by a CPU or module thereof, by a separate control module, or the like) to switch between  
10 encoding different streams such that a rate of encoding by the base codec is higher than the rate at which data is being stored into the buffer.

- In an example of the third aspect, the base codec is an AV1 codec or an AVC codec or an HEVC codec, and the plurality of enhancement encoders are LCEVC encoders. In another example, the base and enhancement layers may be  
15 arranged in a hierarchical structure. In another example, the method may further comprise providing the base coded stream, optionally in addition to the enhancement coded streams. In this way, LCEVC encoding (which may not use inter-frame prediction) may be employed to provide enhanced compression by reducing bitrates in high motion cases (e.g., video footage from wearable  
20 cameras).

- In an example of the third aspect, the plurality of enhancement encoders comprises three enhancement encoders or four enhancement encoders, and the plurality of input video streams comprises four input video streams. In another example, the plurality of enhancement encoders comprises three enhancement  
25 encoders, and the plurality of input video streams comprises three input video streams. In another example, the number of enhancement encoders is at least equal to the number of input streams. In this way, the enhancement coding operation may provide a high performance gain density over using a base encoder alone.

In an example of the third aspect, the plurality of input video streams are provided in use by one or more of a camera, a security camera, an automotive camera, and a body worn camera. In some examples, the SoC is configured to receive the plurality of input streams from a respective plurality of such cameras. In another  
5 example, the SoC may be configured for use in an edge device, wherein the edge device is configured to handle or otherwise receive each input stream of a plurality of input streams from each respective camera in a plurality of local cameras. In another example, the SoC may be configured to receive the plurality of input streams from a respective plurality of views captured by a multi-view camera.  
10 Optionally, different views may have different lens angles (e.g., a first view may be provided by a wide-angle lens, a second view may be provided by an ultra-wide angle lens, a third view may be provided by a fisheye lens, a fourth view may be provided by a narrow-angle lens, and so on). In this way, enhancement encoders within a video encoder SoC may be integrated into challenging practical  
15 applications involving cameras.

According to a fourth aspect of the invention, there may be provided a system on a chip, SoC, configured to be operated in use according to the method of the third aspect and any examples thereof.

According to a fifth aspect of the invention, there may be provided a method for  
20 providing an encoded video signal, the method comprising: determining a region of interest, RoI, in an input video signal; modifying an enhancement coding operation to prioritise enhancement of a base video signal at the RoI; and, encoding the input video signal using the modified enhancement coding operation to generate an encoded video signal. In this way, RoI enhancement encoding  
25 may focus residuals (and thereby prioritise enhancement) on areas of interest for high quality reconstruction at viable bitrates.

In an example of the fifth aspect, the determining a region of interest, RoI, in an input video signal comprises: analysing the input video signal to identify the RoI. In another example, the RoI may correspond to an area or areas of a frame of an  
30 image or video to be allocated residuals and/or to an area or areas of a frame of an image or video where an object of a particular class has been localised. In this

way, residuals corresponding to objects of different classes may be modified differently or not modified at all.

In an example of the fifth aspect, the analysing the input video signal to identify the Rol comprises: applying a machine learning (ML) model and/or a computer  
5 vision (CV) model to the input video signal. In another example, the ML and/or CV models may be a part of an artificial intelligence, AI, recognition model. In this way, different AI detection/recognition searches may be implemented at an appropriate quality level, resulting in higher efficiency and accuracy.

In another example, object detection may be implemented at the base video signal  
10 resolution, and/or object recognition may be implemented at full input video resolution, and/or decoding (i.e. partial decoding) may be implemented only on the identified Rol. In another example, a first object analysis task may be performed on a first version of an image or a video at a first level of quality, and a second object analysis task may be performed on a second version of the image  
15 or the video at a second level of quality, the second level of quality being higher than the first level of quality, wherein the first object analysis task may be an object detection task and the second object analysis task may be an object recognition task. In an another example, results of an object analysis task may comprise one or more confidence levels that a given object is or has been detected and/or  
20 recognised.

In an example of the fifth aspect, the determining a region of interest, Rol, in an input video comprises: receiving a set of coordinates corresponding to the Rol. In another example, the set of coordinates may be received by or otherwise provided to the enhancement coding operation from an API or from an AI object analysis  
25 model. In another example, the set of coordinates may be inserted into the base video signal, optionally by the API, the AI object analysis model, or by a base codec, such that the set may be extracted by the enhancement coding operation. In another example, the method may further comprise: providing AI-detected Rol coordinates to the enhancement coding operation to adjust a focus of the video  
30 according to the Rol coordinates, optionally on a per-frame basis. In this way, encoding the Rol enhancement based on the received Rol coordinates permits

prioritisation of residuals for known high priority areas, and permits optimisation of video quality, focus, and improvements in compression efficiency.

In an example of the fifth aspect, the determining a region of interest, RoI, in an input video comprises: analysing residual quantisation settings of a base codec configured to provide the base video signal. In another example, the settings of residual quantisation process of the base codec may comprise one or more quantisation step widths, wherein areas/portions of the base video signal having a relatively lower quantisation step width may correspond to parts of the base video signal corresponding to the RoI, and/or areas/portions of the base video signal having a relatively higher quantisation step width may correspond to parts of the base video signal not corresponding to the RoI.

In an example of the fifth aspect, the modifying an enhancement coding operation to prioritise enhancement of a base video signal at the RoI comprises: adjusting a priority map, the priority map comprising one or more weights corresponding to one or more residuals, wherein the one or more weights indicate to prioritise encoding residuals corresponding to the RoI. In this way, the priority map may guide decisions on which area or areas of a frame, scene or video should be allocated residuals to provide enhancement and thereby higher resolution.

In an example of the fifth aspect, the method further comprises: filtering the one or more residuals according to the priority map. In another example, filtering a residual may comprise selecting the residual for encoding where the residual corresponds at least in part to the RoI, or quantising the residual to zero where the residual corresponds to an area of the scene, frame or footage which is not in the RoI.

In an example of the fifth aspect, the modifying an enhancement coding operation to prioritise enhancement of a base video signal at the RoI comprises: encoding the input video signal using a base codec to provide the base video signal; and modifying a residual quantisation process of the base codec to selectively encode residuals corresponding to the RoI. In another example, non-ROI portions of the image/video may remain at base layer resolution.

In an example of the fifth aspect, the modifying a residual quantisation process of the enhancement coding operation to selectively encode residuals corresponding to the Rol comprises one or more of: adjusting a step width of the residual quantisation operation for the residuals corresponding to the Rol and/or for residuals not corresponding to the Rol; quantising residuals not corresponding to the Rol to zero. In this way, a step width applied during the encoding process may be modified to selectively encode (a) residual value(s) corresponding to the Rol.

In an example of the fifth aspect, the modifying an enhancement coding operation to prioritise enhancement of a base video signal at the Rol comprises: modifying the enhancement coding operation based on the Rol and a received indication of value of the Rol.

In some examples, the modifying an enhancement coding operation to prioritise enhancement of a base video signal at the Rol may be implemented by an enhancement codec configured to perform the enhancement coding operation. In other words, the means for modifying the enhancement process according to the present disclosure may be included as part of the (e.g. standalone) enhancement codec. In this way, enhancement modification may be independent of any given base codec and may be implemented without reference to, or without any requirement for, a base codec.

In an example of the fifth aspect, the modifying an enhancement coding operation to prioritise enhancement of a base video signal at the Rol comprises: overriding a default enhancement coding operation to prioritise enhancement of a base video signal at the Rol. In some examples, the default enhancement coding operation may include an enhancement coding operation configured to encode residuals for an input video without reference to an Rol and/or without reference to a priority map. In some examples, the default enhancement coding operation may include an enhancement coding operation applied in a scenario where no object of interest is detected in the input video stream. In some examples, overriding the default enhancement coding operation may include varying, altering or updating the default encoding parameters of said operation to prioritise residual encoding of identified Rols. In some examples, overriding the default enhancement coding

operation may include overriding a default enhancement coding operation according to the determined Rol to prioritise enhancement of a base video signal at the Rol over the base video signal not at the Rol.

5 In an example of the fifth aspect, the method further comprises: receiving the set of coordinates corresponding to the Rol and/or the indication of value of the Rol respectively from an application programming interface, API, or extracting the set of coordinates corresponding to the Rol and/or the indication of value of the Rol respectively from the base video signal by the enhancement coding operation.

10 In an example of the fifth aspect, before determining a region of interest, Rol, in an input video signal, the method comprises: receiving an input video signal from a camera, a security camera, an automotive camera, or a body worn camera. In this way, compared to methods known in the art where the motion of such cameras while employing typical compression techniques is ineffective due to the prevalence of inter-frame prediction, LCEVC does not use inter-frame prediction  
15 and so does not suffer such issues.

In an example of the fifth aspect, the Rol corresponds to an area of the input video signal where an object has been detected or recognised, or the Rol corresponds to an area of the base video signal where an object has been detected or recognised. In another example, the method may further comprise: determining  
20 the Rol as an area of one or more frames of the input video signal corresponding to an area in which an object has been detected or recognised, wherein the object is optionally of a particular class.

In an example of the fifth aspect, the method further comprises: transmitting or otherwise outputting the encoded video signal.

25 According to a sixth aspect of the invention, there may be provided a system on a chip, SoC, configured in use to perform the method of the fifth aspect and any examples thereof.

According to a seventh aspect of the invention, there may be provided a method for providing a region of interest, Rol, in a to-be-encoded input video signal, the

method comprising: analysing an input video signal to identify a region of interest, RoI, wherein the RoI is usable to modify an enhancement coding operation to prioritise enhancement of the input video signal at the RoI; and, transmitting or otherwise outputting a set of coordinates corresponding to the RoI. In this way, video quality may be optimised by encoding the RoI enhancement based on RoI coordinates and prioritising residuals in known high priority areas.

According to an eighth aspect of the invention, there may be provided a processing device configured in use to perform the method of the seventh aspect and any examples thereof.

10 In an example of the eighth aspect, the processing device is an RoI identification module and, optionally, the processing device is an RoI identification module configured for use in the system on a chip, SoC, of any of the first, fourth, and sixth aspects and any examples thereof. In another example, the RoI identification module may be a separate module to the enhancement coding operation (or the encoders/decoders/other modules thereof). In another example, the RoI identification module may insert the set of coordinates into the base coded stream such that the set may be extracted by the enhancement coding operation. In this way, the enhancement coding operation may be permitted to provide more residuals in the area indicated by the RoI identification module, and therefore a higher resolution for the particularly desired area of the scene.

25 According to a ninth aspect of the invention, there may be provided a camera system configured to be worn on a body, the camera system comprising a camera, a processor, and the SoC of the first, fourth, and sixth aspects and any examples thereof. In some examples, there may be provided a camera system comprising the multi-resolution camera of the second aspect and any examples thereof.

According to a tenth aspect of the invention, there may be provided a method for providing an encoded video signal, the method comprising: analysing a received input video signal to identify a region of interest; modifying an enhancement coding operation to prioritise enhancement of a base video signal at the region of

interest; and, encoding the received input signal using the enhancement coding operation to generate the encoded video signal.

In an example of the tenth aspect, the method further comprises: receiving an indication of value of the region of interest, and modifying the enhancement coding  
5 operation based on the region of interest and the indication of value. The indication of value may be received by an API or inserted into the base stream for extraction from the base stream by the enhancement coding operation.

In an example of the tenth aspect, modifying may comprise adjusting a priority  
10 map, the priority map being a set of residual weights (or a residual mask) corresponding to a residual or group of residuals. The weights may indicate to the encoder to prioritise the encoding of residuals according to the weights. In another example, the residuals may be filtered according to the weights. In another example, the residuals may be quantised differently according to the weights of the residuals. In another example, a step width applied during the encoding  
15 process may be modified to selectively encode a residual value corresponding to the region of interest.

In an example of the tenth aspect, analysing a received input video signal to identify a region of interest comprises: applying a machine learning and/or computer vision model to the input video signal.

20 In an example of the tenth aspect, the analysing and/or the modifying is provided by an enhancement decoder or may be received from a remote or external device or system.

In an example of the tenth aspect, the method further comprises: outputting and/or transmitting the encoded video signal.

25 According to an eleventh aspect of the invention, there may be provided a method for providing an encoded video signal, the method comprising: analysing a received input video signal to identify a region of interest; and, outputting a set of coordinates corresponding to a region of interest, the region of interest being for encoding enhancement data in that region.

In an example of the eleventh aspect, analysing a received input video signal to identify a region of interest comprises: applying a machine learning and/or computer vision model to the input video signal. In another example, the analysing may be provided by an enhancement decoder or may be received from a remote or external device or system.

According to a twelfth aspect of the invention, there may be provided a method for providing an encoded video signal, the method comprising: receiving a set of coordinates corresponding to a region of interest; modifying an enhancement coding operation to prioritise enhancement of a base video signal at the region of interest; and, encoding the received input signal using the enhancement coding operation to generate the encoded video signal.

In an example of the twelfth aspect, the method further comprises: receiving an indication of value of the region of interest, and modifying the enhancement coding operation based on the region of interest (e.g. the received coordinates thereof) and the indication of value. In another example, the set of coordinates and/or indication of value may be received by an API or inserted into the base stream for extraction from the base stream by the enhancement coding operation.

In an example of the twelfth aspect, modifying comprises adjusting a priority map, the priority map being a set of residual weights (or a residual mask) corresponding to a residual or group of residuals. The weights may indicate to the encoder to prioritise the encoding of residuals according to the weights. In another example, the residuals may be filtered according to the weights. In another example, the residuals may be quantised differently according to the weights of the residuals. In another example, a step width applied during the encoding process may be modified to selectively encode a residual value corresponding to the region of interest. In another example, the modifying may be provided by an enhancement decoder or may be received from a remote or external device or system.

In an example of the twelfth aspect, the method further comprises: outputting and/or transmitting the encoded video signal.

According to a thirteenth aspect of the invention, there may be provided a multi-resolution camera configured to output two video streams decodable to derive three video streams at three different resolutions, wherein the multi-resolution camera is configured to output a first encoded enhancement stream comprising an enhancement stream at a high resolution and a base stream at a medium resolution, and a second stream at a low resolution.

According to a fourteenth aspect of the invention, there may be provided an SoC for a camera, the SoC being configured to receive a plurality of input video streams and output a plurality of encoded output video streams, preferably at the same resolution. The SoC may comprise a base encoder for use in encoding a plurality of the input streams and a plurality of enhancement encoders configured to provide the plurality of encoded output video streams.

In an example of the fourteenth aspect, the SoC may comprise a (e.g. single) base encoder configured to encode a plurality (e.g. two, three, four, and so forth) of (e.g. different) video streams at a first resolution (e.g. resolution X), optionally wherein the plurality equals N. The SoC may comprise a plurality of (e.g. separate) enhancement encoders. Each enhancement encoder may be configured to encode a (e.g. single) video stream. The SoC may comprise N enhancement encoders. In other words, the SoC may comprise a number of enhancement encoders, such that the number of enhancement encoders equals the number of video streams that the base encoder is configured to process (e.g. encode). The video stream that each enhancement encoder processes may correspond to a rendition of one of the plurality of video streams that the base encoder processes (e.g. is input into the base encoder or is output by the base encoder). In other words, the SoC may comprise a (e.g. single) base encoder configured to process multiple video streams, and further configured to serve (e.g. distribute data to, be connected to, and so forth) multiple enhancement encoders.

According to a fifteenth aspect of the invention, there may be provided computer programs comprising respective instructions which, when executed by a respective computer, cause the computer to perform the method of the third, fifth, seventh, tenth, eleventh, and twelfth aspects respectively, and any examples

thereof. The computer programs of the fifteenth aspect may be stored as respective instructions on respective non-transitory computer-readable media or in respective signal transmissions.

In some examples, a base codec may be understood to include a base encoder  
5 and a base decoder, which may or may not be independent of each other. Herein, a base codec may alternatively be referred to as a base coding block. In some examples, the base codec may be provided as part of a general low complexity encoder, which may be configured to control an independent base encoder and decoder (e.g. as packaged as a base codec). In this way, the base encoder and  
10 decoder may be supplied as part of the low complexity encoder. In some examples, the low complexity encoder may be seen as a form of wrapper for the base codec, where the functionality of the base codec may be hidden from an entity implementing the low complexity encoder. In some examples, a base codec may be any other (e.g. non-enhancement) codec, such as AV1, AVC, HEVC,  
15 which typically operates at a lower resolution than the input signal, and the base codec may be provided together with one or more enhancement codecs as herein described.

Herein, an input signal may be received from or otherwise originate from such a device as a camera, security camera, automotive camera, body worn camera,  
20 wearable camera, multi-resolution camera, camera SoC, camcorder, sensor. Herein, any method may further comprise a step of output the signal resulting from, provided by, or otherwise obtained according to that method. Herein, any method may be performed at least at a camera, by an SoC, and/or by a camera SoC; and any method may be applied at least to operate a camera, an SoC, and/or  
25 a camera SoC.

Though the description of each aspect has its own emphasis, for any example, feature, or advantage that is not described in relation to one aspect, reference may be made to related descriptions of any other aspects. In this way, the examples herein described are not limited in their implementation to one particular  
30 aspect and may be implemented in other aspects.

**BRIEF DESCRIPTION OF THE DRAWINGS**

Exemplary implementations of the invention may be described with reference to the accompanying drawings, in which:

- 5 Figure 1 illustrates an image including Rols identified according to the present disclosure;
- Figure 2 illustrates an image including an Rol identified according to the present disclosure, and a base encoded version thereof;
- Figure 3 illustrates an image including an Rol at high resolution compared to the non-Rol portions of the image at low resolution;
- 10 Figure 4 illustrates an image including an Rol identified according to the present disclosure, and a base encoded version thereof wherein only the Rol is enhanced according to the present disclosure;
- Figure 5 illustrates an image including an Rol identified according to the present disclosure, and a base encoded version thereof;
- 15 Figure 6 illustrates an image including an Rol identified according to the present disclosure, and a base encoded version thereof wherein only the Rol is enhanced according to the present disclosure;
- Figure 7 illustrates an image including an Rol identified according to the present disclosure, and a base encoded version thereof wherein only the Rol is enhanced according to the present disclosure;
- 20 Figure 8 illustrates an image including Rols identified according to the present disclosure, and a base encoded version thereof wherein one Rol is enhanced according to the present disclosure;
- Figure 9 illustrates examples of per frame bitrate as a percentage of total bitrate in an example of the present disclosure;
- 25 Figure 10 illustrates a camera SoC in the art and a camera SoC having an enhancement coding operation according to the present disclosure;
- Figure 11 illustrates an exemplary encoding process according to the present disclosure;
- 30 Figure 12 illustrates an exemplary encoding process according to the present disclosure;

Figure 13 illustrates an exemplary encoding process according to the present disclosure;

Figure 14 illustrates an exemplary process of providing an indication of an RoI according to the present disclosure;

5 Figure 15 illustrates an exemplary process of providing an indication of an RoI according to the present disclosure.

## DETAILED DESCRIPTION

### Bandwidth Savings

LCEVC encoders may provide particular savings in storage and streaming bitrate.

10 In an initial implementation example, a camera is configured to output multiple streams, each of different resolutions. This may be referred to as a typical multi-resolution scenario and may for example output 3 streams from each camera. The streams may be low quality, medium quality (HD) and high quality (UHD) depending on the camera and processing capabilities.

15 It is proposed that a single LCEVC stream can replace the top two qualities, by using the base layer for the medium quality (HD) and the enhanced stream for high quality (UHD), unlocking significant storage savings by removing a standalone UHD stream.

20 LCEVC may also provide enhanced compression, reducing bitrates especially in challenging/high motion cases (e.g., wearables). By wearable, we mean a body worn camera. The motion of such cameras means that typical coding techniques of modern compression algorithms may be particularly ineffective since they achieve efficiency gains by predicting motion from one frame to the next. LCEVC does not use inter-frame prediction and so does not suffer such issues.

25 It is proposed to integrate an LCEVC encoder into a body worn camera system comprising a camera, a processor and a video encoder SoC. However, an LCEVC encoder could equally be integrated into any kind of security camera.

Rate control in LCEVC allows optimal base vs. enhancement bitrate allocation, reducing bitrates spikes in peak-motion scenes, by backing-up to a lower resolution.

It is proposed that a rate control method be adapted or calibrated to smooth  
5 bitrates spikes by allowing a base resolution to be used where a high-motion scene requires a large amount of bits to encode and transmit the scene.

### **Quality**

LCEVC encoders may provide enhanced quality and more details for forensic analysis.

10 In examples, Region of Interest (RoI) enhancement encoding allows to focus LCEVC residuals on areas of interest (e.g., faces, plates) for pristine quality reconstruction at viable bitrates.

It is proposed that a set of coordinates, and preferably a value parameter, are provided to an LCEVC encoder. The LCEVC encoder may be configured to  
15 provide more residuals in the area provided by the coordinates.

LCEVC compression efficiency allows more details in high motion scenarios. This is a benefit of using LCEVC in security and automotive camera devices and systems.

The indication of an area (e.g. the set of coordinates) may be provided by a base  
20 encoder to the enhancement encoder (e.g. LCEVC encoder).

The indication of an area (e.g. the set of coordinates) may be provided by a RoI identification module to one or more of the enhancement encoder (e.g. LCEVC encoder) and the base encoder. The RoI identification module may be a separate module to one or more of the enhancement encoder and the base encoder.

25 In one example, the enhancement encoder determines an indicated area by analysing the quantisation settings of the base. For example, if a particular

area/portion of the base layer has a small quantisation step width, then this may be determined by the enhancement encoder to be an RoI because the base layer is of a higher quality at this region.

### **Analytics**

- 5 LCEVC encoding tools can uniquely leverage Analytics at camera level to optimise quality, for instance by encoding the RoI enhancement based on detected RoI coordinates and prioritising residuals on known high priority areas.

LCEVC encoding tools may be integrated with and are amenable for AI image processing. Higher quality details, such as in RoIs, improve AI recognition  
10 accuracy.

The native dual-layer format allows different AI searches at the most appropriate quality level, resulting in higher efficiency and accuracy (e.g., object detection on base layer, object recognition at full resolution, partial decoding only on RoI).

To elaborate, a first object analysis task may be performed on a first version of an  
15 image at a first level of quality. Then, a second object analysis task may be performed on a second version of the image at a second level of quality, the second levels of quality being higher than the first level of quality.

The object analysis tasks can be categorised into two broad categories. A first type of object analysis may be object detection, which detects the presence of an  
20 object in an image. The position of the object in the image may also be localised.

In other words, object detection relates to finding one or more instances of one or more objects of one or more particular classes and localizing the one or more objects within the representation. Object detection may therefore relate to detecting all objects belonging to certain classes for which the object analysis  
25 element has been trained, as well as localizing them within the image. An object analysis element, for example, may have been trained to detect human faces, vehicles, signs, and animals. If such an object analysis element 370 detects one or more such objects, the specific location of each such detected object is

returned, for example via a bounding box. For example, a bounding box may be provided in relation to each human face and each vehicle detected in an image. A result of such object detection may be a level of confidence that one or more objects have been detected and localized.

- 5 A second type of object analysis task is object recognition. Object recognition relates to identifying an object that has been detected. For example, object recognition may relate to determining a class label, such as “car”, “lorry” “motorbike” etc, to which a detected vehicle belongs. In the case of a human face, object recognition may relate to recognising the identity of the particular person  
10 whose face has been detected. A result of such object recognition may, for example, be a level of confidence that a given detected vehicle is a car. For example, a result of such object recognition may correspond to an 80% level of confidence that a car has been recognised and a 20% level of confidence that a car has not been recognised.

### 15 **Density & efficiency**

- LCEVC provides a >2x performance density gain over AV1 alone (and similarly over AVC/HEVC chipsets), thanks to the base encoder only needing to support 1/4 resolution video. In an example multiviewer system from monitoring cameras, it is possible to reduce coding complexity by avoiding multiple UHD decoding in  
20 mosaic view by just coding 1/4 resolution base layer. The full resolution LCEVC-enhancement stream could be coded only when needed, e.g. in response to a detection (at the low resolution layer) of a particular object.

### **Priority Map**

- LCEVC has a “Priority Map” feature that highlights priority levels within each  
25 frame, guiding decisions on where to allocate the enhancement first. It is proposed to calibrate LCEVC Priority Map according to their needs, selectively prioritising high-quality reconstruction in important areas (e.g., faces).

Figure 1 illustrates an example of an image, highlighting areas (i.e., regions) of interest of a priority map. A priority map is a map configured to provide weights

(or a residual mask) to a set of residuals, guiding decisions on which area or areas of a frame should be allocated residuals to provide enhancement.

In other words, the residuals are modified (i.e. selected for encoding or quantized to 0) based on the indications in the map corresponding to the residuals.

- 5 In accordance with the above discussion on object analysis, the area or areas of the frame to be allocated residuals could correspond to areas where an object of a particular class has been localised. Further, residual data corresponding to objects belonging to different classes can be modified in different ways according to their class. For example, residuals relating to a detected sign could be modified
- 10 differently to residuals relating to a detected vehicle. In other examples, residuals relating to objects in a particular class may not be modified at all, for example where greater fidelity is needed.

An example of a priority map is provided in patent publication WO2020/188229, the contents of which are incorporated by reference.

- 15 Figure 2 illustrates an image with a region of interest highlighted, showing the image encoded using x264 and LCEVC encoding of an x264 base, without residuals being encoded.

### **Region of Interest (ROI) enhancement**

- 20 Figure 3 illustrates a region of interest at 2160p resolution compared to a base layer image at 1080p resolution.

The LCEVC format allows flexibility on where to place the enhancement and the extend to which part of an image is to be enhanced. Residuals can be placed exclusively in areas (i.e., regions) of interest (Rols) for forensic purposes (e.g., faces, plates, text, etc) which are targeted for high quality reconstruction, while

25 non-Rol portions of the frame remain with base layer quality.

The enhancement may be placed by modifying the residuals of the enhancement layers or selectively encoding them, for example, killing or not encoding the not

selected residuals. This may be thought of as ranking or filtering, or weighting the residuals.

The modification may be performed by modifying (or selectively modifying) a step width of a quantisation process, for example, to increase a dead zone which  
5 quantizes the residuals to 0. Examples of modification may be found in WO2023/187308, GB2313070.1 and WO2020/188229, the contents of which are incorporated by reference.

For instance, AI-detected Roi coordinates can be fed to an LCEVC encoder to adjust areas of focus on a per-frame basis.

10 That is, Rols may be determined in accordance with a result of an object analysis, such as object detection or object recognition task. Put simply, a region of interest may correspond to an area of a frame where an object has been detected or recognised.

In an example the coordinates may be fed by an API to the enhancement coding  
15 operation or via the base stream where they are passed to the enhancement coding operation or extracted from it.

Examples of the passing of such information to an enhancement operation are given in patent publication WO2020/212701 which is incorporated by reference.

The bitrate cost is affordable, as Rols often represent a small portion of the frame  
20 (1/100th as in the example), e.g., within 1-3 Mbps.

Figures 4-8 illustrate examples of region of interest encoding using different encoders and parameters.

Region of interest coding examples are provided in PCT/GB2023/052755, WO2018/015764 and WO2020/165575, which are incorporated by reference.

### **LCEVC adaptive base reconfiguration**

LCEVC rate control allows automatic per-frame base vs. enhancement bitrate adaptation. In peak complexity scenes, it would allocate majority of the bitrate to the base layer, turning to a more pleasant lower resolution image and massively  
5 reducing bitrate spikes. In normal/ low motion scenarios the base would require less bitrate for a good quality, hence more bitrate can be allocated to the enhancement.

Figure 9 illustrates examples of per frame bitrate as a percentage of total bitrate in a gaming example.

10 Examples of rate control in LCEVC are provided in WO2020/188230 and GB2308476.7, both of which are incorporated by reference.

### **Encoding density benefits due to LCEVC low complexity**

According to an example illustrated in Figure 10, there may be provided an SoC for a Camera, configured to receive four input video streams and output 4 encoded  
15 output video streams, preferably at the same resolution. The SoC may comprise one base encoder, e.g. AV1 and a plurality of LCEVC encoders, e.g. 3. The encoders may be configured such that the base encoder encodes a reduced resolution of the input video (e.g.  $\frac{1}{4}$ ) and the LCEVC encoders are configured to add enhancement to that base encoded video to create the encoded output video  
20 streams at the same resolution as the input video streams.

Since the base encoder is operating at only a quarter of the resolution, it may be possible to output multiple streams (e.g. 4) at the resolution using only one base encoder. For example, if the base encoder has a throughput of X bits/s, then without utilising an enhancement stream, the base encoder could output a single  
25 stream (of X bits/s). However, if the base encoder is operating at only a quarter of the resolution, then the base encoder will, generally speaking, take a quarter of the time to process a stream. This means that the base encoder can process (e.g. encode) a first stream for a portion of time (where the portion of time is a quarter of the time it would have spent encoding the full resolution base), process (e.g.

encode) a second stream for a second portion of time (where the portion of time is a quarter of the time it would have spent encoding the full resolution base), process (e.g. encode) a third stream for a third portion of time (where the portion of time is a quarter of the time it would have spent encoding the full resolution base), process (e.g. encode) a fourth stream for a fourth portion of time (where the portion of time is a quarter of the time it would have spent encoding the full resolution base). The base encoder may utilise a buffer to achieve this, for example, the 'raw' data corresponding to the second stream may be stored in a buffer during the first portion of time, then when the first stream has been encoded, the base encoder can retrieve the buffered second stream and begin processing it. In this way, a base encoder can switch between encoding different streams, at a rate where the base encoder can encode at a quicker rate than the information being put into a buffer.

This provides >2x encoding density benefits due to LCEVC low complexity, more density per single chip (4x LCEVC base layers) and supports 8Kp60.

LCEVC provides a ~2.3x performance density gain over AV1 alone. This is based on the base encoder only needing to support  $\frac{1}{4}$  resolution video. In a worked example:

1 x AV1 4kP60 encoder (660 MHz clock) = 1x N mm<sup>2</sup>

1 x LCEVC 4kP60 encoder = 0.25 x N mm<sup>2</sup>

1 x AV1 4kP60 encoder + 3-4 x LCEVC 4kP60 encoder = 1.75 x N mm<sup>2</sup>

It can encode: 4 x 4kP60 (880 MHz clock); 3 x 4kP60 (660 MHz clock).

Similar performance density gains are possible with H.264 and HEVC.

### Further Examples

There may be provided a method for providing an encoded video signal, comprising: analysing a received input video signal to identify a region of interest;

modifying an enhancement coding operation to prioritise enhancement of a base video signal at the region of interest; and, encoding the received input signal using the enhancement coding operation to generate the encoded video signal.

There may be provided a method for providing an encoded video signal, comprising: analysing a received input video signal to identify a region of interest; outputting a set of coordinates corresponding to a region of interest, the region of interest being for encoding enhancement data in that region.

There may be provided a method for providing an encoded video signal, comprising: receiving a set of coordinates corresponding to a region of interest; modifying an enhancement coding operation to prioritise enhancement of a base video signal at the region of interest; and, encoding the received input signal using the enhancement coding operation to generate the encoded video signal.

Modifying may comprise adjusting a priority map, the priority map being a set of residual weights (or a residual mask) corresponding to a residual or group of residuals. The weights indicate to the encoder to prioritise the encoding of residuals according to the weights. The residuals may be filtered according to the weights. Alternatively, the residuals may be quantised differently according to the weights of the residuals.

A step width applied during the encoding process may be modified to selectively encode a residual value corresponding to the region of interest.

The input signal may be received from a camera, security camera, automotive camera or body worn camera.

The method may comprise outputting and/or transmitting the encoded video signal.

The method may comprise receiving an indication of value of the region of interest and modifying the enhancement coding operation based on the region of interest and the indication of value.

The set of coordinates and/or indication of value may be received by an API or inserted into the base stream for extraction from the base stream by the enhancement encoding operation.

5 The analysing a received input video signal to identify a region of interest may comprise applying a machine learning and/or computer vision model to the input video signal.

The analysing and modification may be provided by an enhancement decoder or may be received from a remote or external device or system.

10 There may be provided a multi-resolution camera configured to output two video streams decodable to derive three video streams at three different resolutions. The multi-resolution camera is configured to output a first encoded enhancement stream comprising an enhancement stream at a high resolution and a base stream at a medium resolution, and a second stream at a low resolution.

15 There may be provided an SoC for a Camera, configured to receive a plurality of input video streams and output a plurality of encoded output video streams, preferably at the same resolution. The SoC may comprise a base encoder for use in encoding a plurality of the input streams and a plurality of enhancement encoders configured to provide the plurality of encoded output video streams.

20 In other words, the system on a chip (SoC) may comprise a (e.g. single) base encoder configured to encode a plurality (e.g. two, three, four, and so forth) of (e.g. different) video streams at a first resolution (e.g. resolution X), optionally wherein the plurality equals N. The system on a chip (SoC) may comprise a plurality of (e.g. separate) enhancement encoders. Each enhancement encoder may be configured to encode a (e.g. single) video stream. The SoC may comprise N  
25 enhancement encoders. In other words, the SoC may comprise a number of enhancement encoders, such that the number of enhancement encoders equals the number of video streams that the base encoder is configured to process (e.g. encode). The video stream that each enhancement encoder processes may correspond to a rendition of one of the plurality of video streams that the base

encoder processes (e.g. is input into the base encoder or is output by the base encoder). In other words, the SoC may comprise a (e.g. single) base encoder configured to process multiple video streams, and further configured to serve (e.g. distribute data to, be connected to, and so forth) multiple enhancement encoders.

- 5 Each of these methods may be performed at a camera or on a camera SoC.

### **Applying modification**

Figure 11 shows an example of applying residual processing to select, de-select or otherwise modify residuals based a region of interest.

An input full resolution video 100 is processed to generate various encoded  
10 streams 101, 102, 103. A first encoded stream (encoded base stream) is produced  
by feeding a base codec (e.g., AVC, HEVC, or any other codec) with a down-  
sampled version of the input video. The encoded base stream may be referred to  
as the base layer or base level. A second encoded stream (encoded level 1  
stream) is produced by processing the residuals obtained by taking the difference  
15 between a reconstructed base codec video and the down-sampled version of the  
input video. A third encoded stream (encoded level 2 stream) is produced by  
processing the residuals obtained by taking the difference between an up-sampled  
version of a corrected version of the reconstructed base coded video and the input  
video. In certain cases, the components of may provide a general low complexity  
20 encoder. In certain cases, the enhancement streams may be generated by  
encoding processes that form part of the low complexity encoder and the low  
complexity encoder may be configured to control an independent base encoder  
and decoder (e.g. as packaged as a base codec). In other cases, the base  
encoder and decoder may be supplied as part of the low complexity encoder. In  
25 one case, the low complexity encoder may be seen as a form of wrapper for the  
base codec, where the functionality of the base codec may be hidden from an  
entity implementing the low complexity encoder.

A down-sampling operation illustrated by downsampling component 105 may be  
applied to the input video to produce a down-sampled video to be encoded by a

base encoder 113 of a base codec. The down-sampling can be done either in both vertical and horizontal directions, or alternatively only in the horizontal direction. The base encoder 113 and a base decoder 114 may be implemented by a base codec (e.g. as different functions of a common codec). The base codec, and/or  
5 one or more of the base encoder 113 and the base decoder 114 may comprise suitably configured electronic circuitry (e.g. a hardware encoder/decoder) and/or computer program code that is executed by a processor.

Each enhancement stream encoding process may not necessarily include an up-sampling step. In this figure for example, the first enhancement stream is  
10 conceptually a correction stream while the second enhancement stream is up-sampled to provide a level of enhancement.

Looking at the process of generating the enhancement streams in more detail, to generate the encoded Level 1 stream, the encoded base stream is decoded by the base decoder 114 (i.e. a decoding operation is applied to the encoded base  
15 stream to generate a decoded base stream). Decoding may be performed by a decoding function or mode of a base codec. The difference between the decoded base stream and the down-sampled input video is then created at a level 1 comparator 110 (i.e. a subtraction operation is applied to the down-sampled input video and the decoded base stream to generate a first set of residuals). The output  
20 of the comparator 110 may be referred to as a first set of residuals, e.g. a surface or frame of residual data, where a residual value is determined for each picture element at the resolution of the base encoder 113, the base decoder 114 and the output of the downsampling block 105.

The difference is then encoded by a first encoder 115 (i.e. a level 1 encoder) to  
25 generate the encoded Level 1 stream 102 (i.e. an encoding operation is applied to the first set of residuals to generate a first enhancement stream).

As noted above, the enhancement stream may comprise a first level of enhancement 102 and a second level of enhancement 103. The first level of enhancement 102 may be considered to be a corrected stream, e.g. a stream that  
30 provides a level of correction to the base encoded/decoded video signal at a lower

resolution than the input video 100. The second level of enhancement 103 may be considered to be a further level of enhancement that converts the corrected stream to the original input video 100, e.g. that applies a level of enhancement or correction to a signal that is reconstructed from the corrected stream.

5 In the example, the second level of enhancement 103 is created by encoding a further set of residuals. The further set of residuals are generated by a level 2 comparator 119. The level 2 comparator 119 determines a difference between an up-sampled version of a decoded level 1 stream, e.g. the output of an upsampling component 117, and the input video 100. The input to the upsampling component  
10 117 is generated by applying a first decoder (i.e. a level 1 decoder) to the output of the first encoder 115. This generates a decoded set of level 1 residuals. These are then combined with the output of the base decoder 114 at summation component 120. This effectively applies the level 1 residuals to the output of the base decoder 114. It allows for losses in the level 1 encoding and decoding  
15 process to be corrected by the level 2 residuals. The output of summation component 120 may be seen as a simulated signal that represents an output of applying level 1 processing to the encoded base stream 101 and the encoded level 1 stream 102 at a decoder.

As noted, an up-sampled stream is compared to the input video which creates a  
20 further set of residuals (i.e. a difference operation is applied to the up-sampled re-created stream to generate a further set of residuals). The further set of residuals are then encoded by a second encoder 121 (i.e. a level 2 encoder) as the encoded Level 2 enhancement stream (i.e. an encoding operation is then applied to the further set of residuals to generate an encoded further enhancement stream).

25 Thus, as illustrated and described above, the output of the encoding process is a base stream 101 and one or more enhancement streams 102, 103 which preferably comprise a first level of enhancement and a further level of enhancement. The three streams 101, 102 and 103 may be combined, with or without additional information such as control headers, to generate a combined  
30 stream for the video encoding framework that represents the input video 100. It should be noted that the components shown in Figure 1 may operate on blocks or

coding units of data, e.g. corresponding to 2x2 or 4x4 portions of a frame at a particular level of resolution. The components operate without any inter-block dependencies, hence they may be applied in parallel to multiple blocks or coding units within a frame. This differs from comparative video encoding schemes wherein there are dependencies between blocks (e.g. either spatial dependencies or temporal dependencies). The dependencies of comparative video encoding schemes limit the level of parallelism and require a much higher complexity.

Figure 11 illustrates a residual selection block 140. Residuals are processed (i.e. modified and/or ranked and selected) in order to determine which residuals should be transformed and encoded, i.e. which residuals are to be processed by the first and/or second encoders 115 and 121. Preferably this processing is performed prior to entropy encoding. Residual selection 140 is an optional step that may configure or activate processing or modification of residuals i.e. residual processing is performed according to a selected mode. For example, the “residual mode” may correspond to a residual pre-processing mode, wherein residuals for enhancement layers are pre-processed prior to encoding. This mode may be turned on and off depending on requirements. For example, the residual mode may be configured via one or more control headers or fields. In alternative embodiments, the residuals may always be modified (i.e. pre-processed) and so selection of a mode is not required. In this case, residual pre-processing may be hard-coded. Examples of residuals processing will be described in detail below. The residual mode, if selected, may act to filter residuals within one or more of the level 1 and level 2 encoding operations, preferably at a stage prior to the encoding sub-components. As shown and described above, the residual processing may be indicated or embedded in the base stream or passed by an API to the enhancement encoding process.

Figure 12 is a flow diagram illustrating an exemplary encoding scheme according to the present disclosure. The scheme comprises receiving input data at the input 201, such as image or video data from a camera or other sensor data from another sensor. In some examples, the input data may be in the form of ‘raw’ or unencoded data. In other examples, the input data may be in the form of base coded data.

The scheme further comprises running the input data through an object analysis module 202 to identify one or more objects of interest in the input data. The object analysis module 202 may comprise or otherwise employ an AI object analysis model to perform this identification. Analysing the input data to identify one or more objects of interest in the input data may include analysing the input data to detect the presence of an object and/or to localise an object within the input data, wherein in some examples the object may be an object belonging to a particular object class. Analysing the input data as such may further include recognising the detected and/or localised object as being an object belonging to that particular class, which in some examples may include determining a confidence level that an object detected in the input data is an object of said class. Once one or more objects are identified (i.e., at least localised or otherwise detected), the scheme further comprises determining one or more Rols corresponding to the one or more objects. The step of determining the Rol may be performed by an Rol identification module 203 as described above. The Rol may correspond to the area in the input data (or the areas in the frames, scenes, streams etc. of the input data) corresponding to where the object is identified. Following the determination of one or more Rols, generally speaking the scheme involves utilising the Rol(s) to influence the enhancement encoding process and, in particular, to influence the encoding decisions (such as residual encoding decisions) made by, for example, one or more enhancement encoders, a quantisation process, or a residual selection module.

However, optionally the encoding scheme may further comprise a pre-analysis step performed by a pre-analysis module 204. Pre-analysis may involve conducting a preliminary analysis on the input data (or the base coded data) prior to (e.g. enhancement) encoding in order to generate header data that can be integrated with the encoded streams. The header data may be usable to modify or control enhancement by the enhancement encoder. The header data may be usable to assist a decoder which subsequently decodes the output data. The header data may comprise encoder configuration parameters including one or more of: resolution settings, quantisation parameters, temporal and/or spatial prediction parameters, bitrate control settings, an indicator indicating a type of

codec. The pre-analysis module may operate locally with the encoder or remotely (e.g. via a network). Such parameters may be applicable at different levels of granularity i.e. per block (e.g. coding unit), per frame, or per group of frames. Further examples of pre-analysis are provided in patent publication WO  
5 2023/187308, the contents of which is incorporated here by reference.

Following Rol detection (and optionally following pre-analysis), the scheme comprises determining, by a priority mapping module 205, one or more priority maps according to the Rol(s) detected in the input data. Determining a priority map may optionally comprise determining the priority map based on the detected  
10 Rol(s) and one or more parameters provided by the pre-analysis module 204. The priority map(s), or the metadata signalling the priority map(s), may be comprised in the encoder configuration parameters, optionally together with the encoder configuration parameters resulting from the optional pre-analysis step. Each priority map may comprise a residual mask, which in turn may comprise  
15 weightings to be applied to residuals at varying granularities (e.g. blocks, frames, etc.) and/or weightings to be applied to the raw input data. The weightings may be configured according to the detected Rols. High priority data may be indicated in the priority map by being assigned a priority (and an associated weighting) above a certain threshold. High priority data may correspond to data in an Rol.  
20 Low priority data may be indicated in the map by being assigned a priority (and an associated weighting) below a certain threshold. Low priority data may correspond to data in regions that are not determined to be an Rol. High priority data may be prioritised for enhancement encoding over low priority data. For example, the Rol data may be enhancement encoded, with residuals corresponding to the Rol being generated and thereafter used in temporal prediction, transformed, quantised, entropy coded, and transmitted. The non-Rol data may similarly be enhancement encoded, with the corresponding residuals being generated but thereafter discarded or quantised to zero prior to transmission. Alternatively, the non-Rol data may not be enhancement encoded  
25 and, optionally, may be base encoded only. In this way, the priority map indicates which residuals (i.e. the residuals corresponding to the Rol) are more important to  
30

encode, and thereby permits adjustment of the enhancement encoding for visually significant areas.

Figure 13 is a flow diagram illustrating an exemplary encoding scheme according to the present disclosure. The scheme comprises inputting data supplied by a camera or by a sensor at an input 301. The input data may comprise image or video data or other sensor data, and it may be in the form of 'raw', unencoded data or in the form of base coded data. The scheme further comprises a prioritisation step performed by a prioritisation module 302. The prioritisation module 302 may be separate to the modules of the encoding operation, for example prioritisation module 302 may be implemented externally and/or by an API. The purpose of the prioritisation step is to determine portions of the input data which are to be prioritised for enhancement encoding over other portions of the input data, to thereby influence the priority map created by the priority mapping module 303. The prioritisation step may include a step of performing object analysis as described above and/or a step of performing RoI detection as described above. Where object analysis is performed on the input data, the resulting object(s) localised in the input data by an object analysis module of the module 302 may be used directly to inform the creation of the priority map. For example, the prioritisation module 302 may configure in-stream metadata (such as by assigning one or more value parameters) according to a particular portion of a frame of the input data (such as a block, coding unit, or other subdivision of the frame) where one or more objects are detected and/or recognised. The object analysis module may incorporate an AI object analysis model and, optionally, an ML and/or CV model thereof. Where object analysis is performed on the input data, the resulting indication of an object localised in the input data may be input to the RoI detection module of the prioritisation module 302 to determine RoI according to the methods described above. Likewise, RoI detection may be performed on the input data by an RoI detection module of the prioritisation module 302 without necessarily performing object analysis beforehand. For example, the RoI detection module may be supplied with one or more sets of coordinates in the input data in which there is expected to be one or more objects of interest, or in which it is already known that there are objects of interest, without

the scheme necessarily requiring additional real-time object analysis. In this way, the Rol detection module may determine one or more Rols which it can thereafter provide to the priority mapping module 303. The resulting prioritisation information or prioritisation metadata provided by the object analysis and/or Rol detection is usable by the priority mapping module 303 to determine the relative priorities of different portions of the input data. Once the priority mapping module 303 receives this information/metadata, it may proceed to determine a priority map, such as by configuring a residual mask to be applied during a subsequent residual quantisation process. The weights of the residual mask may be comprised in the encoder configuration data, optionally together with other prioritisation metadata provided by the prioritisation module 302 and which may be embedded in the input stream(s) to the enhancement coding operation. As described above, the priority map may indicate to prioritise enhancement encoding in order to encode more (e.g. LCEVC) residuals in the prioritised portions of the input video and thereby encode more details for said portions. Alternatively or additionally, the priority map may indicate to de-prioritise enhancement encoding for input data having a lower priority in order to encode less (e.g. LCEVC) residuals and therefore less details for such portions. This may be applicable in such cases as CCTV imagery, where areas of the footage in which faces are identified may correspond to low or below-threshold priority portions of the input data in order to avoid privacy restrictions on the CCTV footage. In this way, the prioritisation information is translated into the enhancement encoded stream via the priority map.

Figure 14 provides an example process of providing an indication of an Rol to an enhancement encoder, such as an LCEVC encoder. The process comprises identifying, by an analysis module, an Rol, as described above. Identifying the Rol may further comprise determining an indication of the Rol usable (such as by an encoder or a CPU module) to modify an enhancement encoding operation. An indication of an Rol may include a set of coordinates (e.g. coordinates in or otherwise according to the frame of the input video) corresponding to the Rol, or an indication of an Rol may include an indication of value (or 'value parameter'), as described above.

Using the example of Rol coordinates and LCEVC, the analysis module may be configured to provide the set of coordinates for the Rol directly to the LCEVC encoder. Once the Rol coordinates are received at the LCEVC encoder, the enhancement coding operation may be modified according to the Rol coordinates  
5 according to the present disclosure.

In the example of Figure 14, the modification of the enhancement coding operation to prioritise enhancement at the Rol may be implemented by an enhancement encoder configured to perform the enhancement coding operation. In this way, enhancement modification may be independent of any given base codec and may  
10 be implemented without reference to, or without any requirement for, a base codec.

Figure 15 provides an example process of providing an indication of an Rol to an enhancement encoder, such as an LCEVC encoder. The process comprises identifying, by an analysis module, an Rol, as described above. Identifying the  
15 Rol may further comprise determining an indication of the Rol usable (such as by an encoder or a CPU module) to modify an enhancement coding operation. An indication of an Rol may include a set of coordinates (e.g. coordinates in or otherwise according to the frame of the input video) corresponding to the Rol, or an indication of an Rol may include an indication of value (or 'value parameter'),  
20 as described above.

Using the example of Rol coordinates, the analysis module may be configured to provide the set of coordinates for the Rol to the base codec. The analysis module typically is also configured to provide the input data signal to the base codec in addition to the set of coordinates. For example, the analysis module may be  
25 configured to embed the coordinates as metadata (such as header data) in the input data stream such that said coordinates may be extractable by an encoder, a decoder, and/or another module (e.g. a residual selection module, a quantisation module etc.) further along in the encoding process. The input data signal may alternatively be provided to the base codec separately, for example the input data  
30 signal may be provided directly to the base codec from the input source, in which case the analysis module may be configured to provide the coordinates to the

base codec as a part of a separate data stream configured between the analysis module and the base codec. Once the input video signal and the coordinates are received at the base codec, the base codec may encode the input signal according to the particular encoding scheme of that base codec (e.g., AVC, HEVC, AV1 etc.).

5 Thereafter, the base codec is configured to provide in effect two output streams, as further described with reference to Figure 11. A first output stream is a first base encoded stream output for transmission or storage, wherein the first base encoded stream is, once decoded, configured for recombination with one or more enhancement encoded streams to reconstruct the input signal. A second output

10 stream is a second base encoded stream output for Rol-modified enhancement encoding at the one or more enhancement encoders according to the present disclosure. In the example of Figure 15, only one such enhancement encoder is shown; however, this is not limited herein, and several such enhancement encoders may be provided as part of an enhancement coding operation described

15 above. Once again, the coordinates corresponding to the Rol may be provided to the enhancement encoder by the base codec. Providing the coordinates as such may include embedding the coordinates in the base coded video stream or may include providing the coordinates in a separate stream between the base codec and the enhancement encoder. Once the base coded signal and the Rol

20 coordinates are received at the enhancement encoder, the enhancement encoding of the base coded signal may be modified according to the Rol coordinates according to the present disclosure.

**CLAIMS**

1. A system on a chip, SoC, comprising a plurality of enhancement encoders and a base codec configured to serve the plurality of enhancement encoders, wherein:
  - 5           the SoC is configured to receive a plurality of input video streams at a first resolution;  
            the base codec is configured to encode the plurality of input video streams to provide a plurality of base coded streams at a second resolution lower than the first resolution; and,
  - 10           each of the plurality of enhancement encoders is configured to encode a respective one of the plurality of base coded streams to provide a plurality of enhancement coded streams corresponding to the plurality of input video streams, the enhancement coded streams being at the first resolution.
- 15 2. The SoC of claim 1, wherein the SoC comprises a buffer configured to store one or more to-be-encoded input video streams during at least the encoding of one of the plurality of input video streams by the base codec.
3. The SoC of claim 1 or claim 2, wherein the base codec is an AV1 codec or an AVC codec or an HEVC codec, and the plurality of enhancement encoders are  
20 LCEVC encoders.
4. The SoC of any of claims 1 to 3, wherein the plurality of enhancement encoders comprises three enhancement encoders or four enhancement encoders, and the plurality of input video streams comprises four input video streams.
- 25 5. The SoC of any of claims 1 to 4, wherein the SoC is configured for use in one or more of a camera, a security camera, an automotive camera, and a body worn camera.

6. The SoC of any of claims 1 to 5, wherein:  
the base codec is configured to provide a set of coordinates corresponding to a region of interest, Rol, to at least one of the plurality of enhancement encoders; and,  
5 the Rol is usable to modify the encoding of the respective one of the plurality of the base coded streams by the at least one of the plurality of enhancement encoders to prioritise enhancement of the base coded stream at the Rol.
- 10 7. The SoC of any of claims 1 to 5, wherein:  
the SoC comprises a region of interest, Rol, identification module configured to provide a set of coordinates corresponding to the Rol to the base codec and/or at least one of the plurality of enhancement encoders; and,  
the Rol is usable to modify the encoding of the respective one of the  
15 plurality of the base coded streams by the at least one of the plurality of enhancement encoders to prioritise enhancement of the base coded stream at the Rol.
- 20 8. The SoC of any of claims 1 to 5, wherein:  
at least one of the plurality of enhancement encoders is configured to determine a set of coordinates corresponding to a region of interest, Rol, based on one or more settings of a residual quantisation process of the base codec; and,  
the Rol is usable to modify the encoding of the respective one of the  
25 plurality of the base coded streams by the at least one of the plurality of enhancement encoders to prioritise enhancement of the base coded stream at the Rol.
- 30 9. A multi-resolution camera comprising the SoC of any of claims 1 to 8, wherein the camera is configured in use to provide the plurality of input video streams to the SoC.

10. A camera system configured to be worn on a body, the camera system comprising a camera, a processor, and the SoC of any of claims 1 to 8.
11. A method for operating a system on a chip, SoC, the SoC comprising a plurality  
5 of enhancement encoders and a base codec configured to serve the plurality of enhancement encoders, and the method comprising:  
receiving, by the SoC, a plurality of input video streams at a first resolution;  
encoding, by the base codec, the plurality of input video streams to  
10 provide a plurality of base coded streams at a second resolution lower than the first resolution; and,  
encoding, by each of the plurality of enhancement encoders, a respective one of the plurality of base coded streams to provide a plurality of enhancement coded streams corresponding to the plurality of input video  
15 streams, the enhancement coded streams being at the first resolution.
12. The method of claim 11, wherein the SoC comprises a buffer, and the method further comprises:  
storing, by the buffer, one or more to-be-encoded input video streams  
20 during at least the encoding of one of the plurality of input video streams by the base codec.
13. The method of claim 11 or claim 12, wherein the base codec is an AV1 codec or an AVC codec or an HEVC codec, and the plurality of enhancement  
25 encoders are LCEVC encoders.
14. The method of any of claims 11 to 13, wherein the plurality of enhancement encoders comprises three enhancement encoders or four enhancement encoders, and the plurality of input video streams comprises four input video streams.

15. The method of any of claims 11 to 14, wherein the plurality of input video streams are provided in use by one or more of a camera, a security camera, an automotive camera, and a body worn camera.
- 5 16. A method for providing an encoded video signal, the method comprising:  
determining a region of interest, RoI, in an input video signal;  
modifying an enhancement coding operation to prioritise  
enhancement of a base video signal at the RoI; and,  
encoding the input video signal using the modified enhancement  
10 coding operation to generate an encoded video signal.
17. The method of claim 16, wherein the determining a region of interest, RoI, in an input video signal comprises:  
analysing the input video signal to identify the RoI.
- 15 18. The method of claim 17, wherein the analysing the input video signal to identify the RoI comprises:  
applying a machine learning model and/or a computer vision model to  
the input video signal.
- 20 19. The method of claim 16, wherein the determining a region of interest, RoI, in an input video comprises:  
receiving a set of coordinates corresponding to the RoI.
- 25 20. The method of claim 16, wherein the determining a region of interest, RoI, in an input video comprises:  
analysing residual quantisation settings of a base codec configured to  
provide the base video signal.

21. The method of any of claims 16 to 20, wherein the modifying an enhancement coding operation to prioritise enhancement of a base video signal at the RoI comprises:
- 5            adjusting a priority map, the priority map comprising one or more weights corresponding to one or more residuals, wherein the one or more weights indicate to prioritise encoding residuals corresponding to the RoI.
22. The method of claim 21, wherein the method further comprises:
- 10            filtering the one or more residuals according to the priority map.
23. The method of any of claims 16 to 22, wherein the modifying an enhancement coding operation to prioritise enhancement of a base video signal at the RoI comprises:
- 15            encoding the input video signal using a base codec to provide the base video signal; and
- modifying a residual quantisation process of the base codec to selectively encode residuals corresponding to the RoI.
24. The method of claim 23, wherein the modifying a residual quantisation process of the enhancement coding operation to selectively encode residuals corresponding to the RoI comprises one or more of:
- 20            adjusting a step width of the residual quantisation operation for the residuals corresponding to the RoI and/or for residuals not corresponding to the RoI;
- 25            quantising residuals not corresponding to the RoI to zero.
25. The method of any of claims 16 to 20, wherein the modifying an enhancement coding operation to prioritise enhancement of a base video signal at the RoI comprises:
- 30            modifying the enhancement coding operation based on the RoI and a received indication of value of the RoI.

26. The method of claim 19 or claim 25, wherein the method further comprises:  
receiving the set of coordinates corresponding to the RoI and/or the  
indication of value of the RoI respectively from an application programming  
interface, API, or  
5 extracting the set of coordinates corresponding to the RoI and/or the  
indication of value of the RoI respectively from the base video signal by the  
enhancement coding operation.
27. The method of any of claims 16 to 26, wherein, before determining a region of  
10 interest, RoI, in an input video signal, the method comprises:  
receiving an input video signal from a camera, a security camera, an  
automotive camera, or a body worn camera.
28. The method of any of claims 16 to 27, wherein the RoI corresponds to an area  
15 of the input video signal where an object has been detected or recognised, or  
the RoI corresponds to an area of the base video signal where an object has  
been detected or recognised.
29. The method of any of claims 16 to 28, wherein the method further comprises:  
20 outputting the encoded video signal.
30. A first system on a chip, SoC, configured to be operated in use according to  
the method of any of claims 11 to 15.
- 25 31. A second system on a chip, SoC, configured in use to perform the method of  
any of claims 16 to 29.
- 30 32. A camera system comprising the SoC of any of claims 1 to 8, 30, or 31, wherein  
the camera system comprises a plurality of cameras, and each camera of the  
plurality of cameras is configured to: capture one of the plurality of input videos  
streams, and send the captured one of the plurality of input streams to the  
SoC.

33. The camera system of claim 32, wherein the camera system comprises one or more downsampler modules configured to downsample respective one or more of the plurality of input video streams prior to the plurality of input video streams being passed to the base codec

5

34. The camera system of claim 33, wherein the camera system comprises an upsampler module configured to upsample an output of the base codec to generate an upsampled rendition of the output of the base codec to be input to one or more of the enhancement encoders.

10

35. The camera system of claim 34, wherein the output of the base codec is a decoded rendition of a base encoded signal, wherein the output of the base codec is at a lower resolution than the input video streams.

15 36. A method for providing a region of interest, Rol, in a to-be-encoded input video signal, the method comprising:

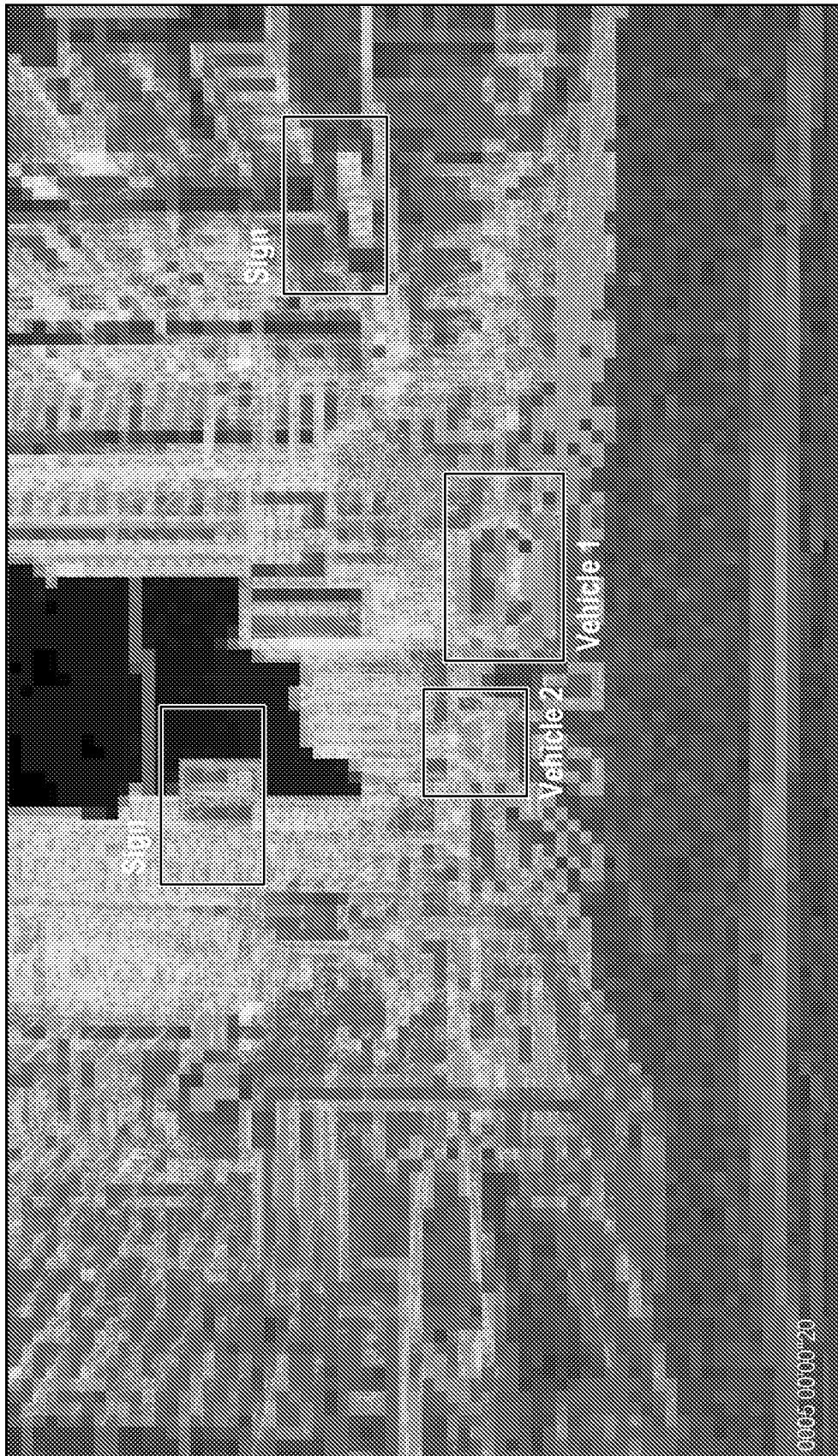
analysing an input video signal to identify a region of interest, Rol, wherein the Rol is usable to modify an enhancement coding operation to prioritise enhancement of the input video signal at the Rol; and,

20

outputting a set of coordinates corresponding to the Rol.

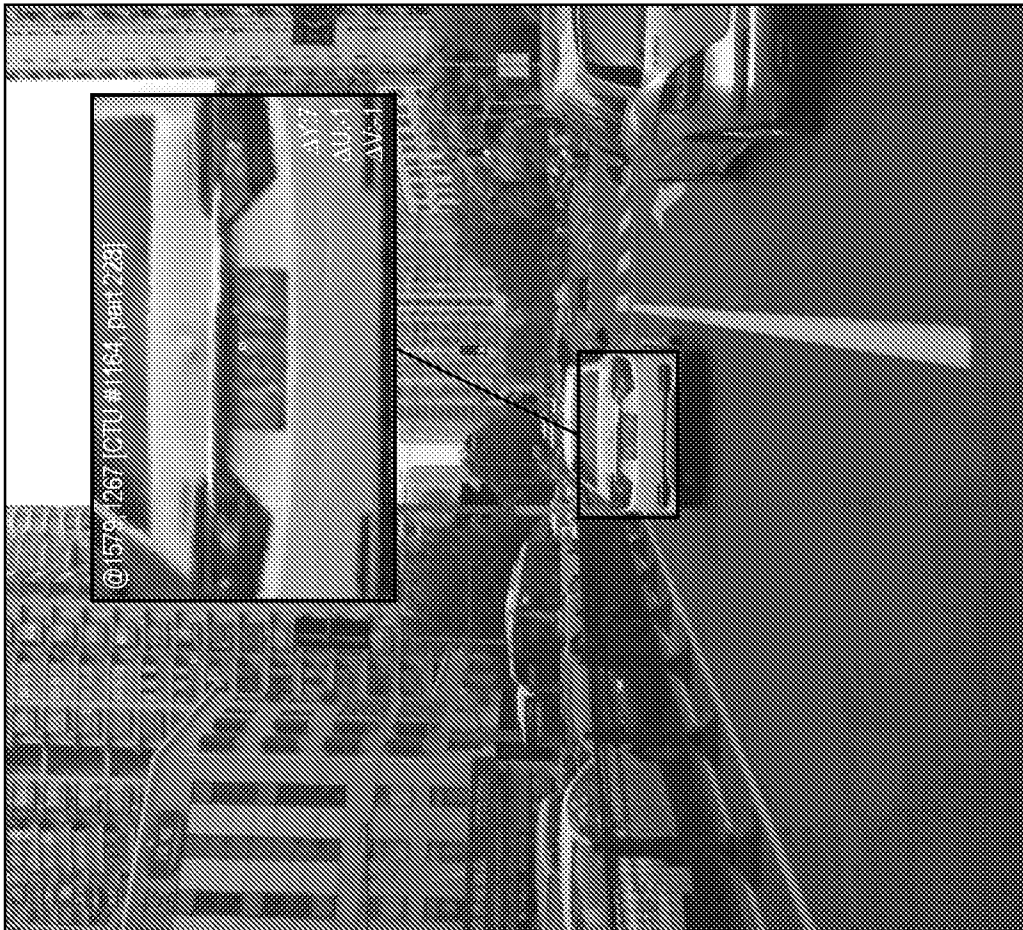
37. An Rol identification module configured in use to perform the method of claim 36.

25



Example of encode priority map surface, 540p base layer

**FIG. 1**



LCEVC x264, 2160p, 6 Mbps base  
- NO residuals (base only)

x264, 2160p, 8 Mbps  
Frame 157

FIG. 2

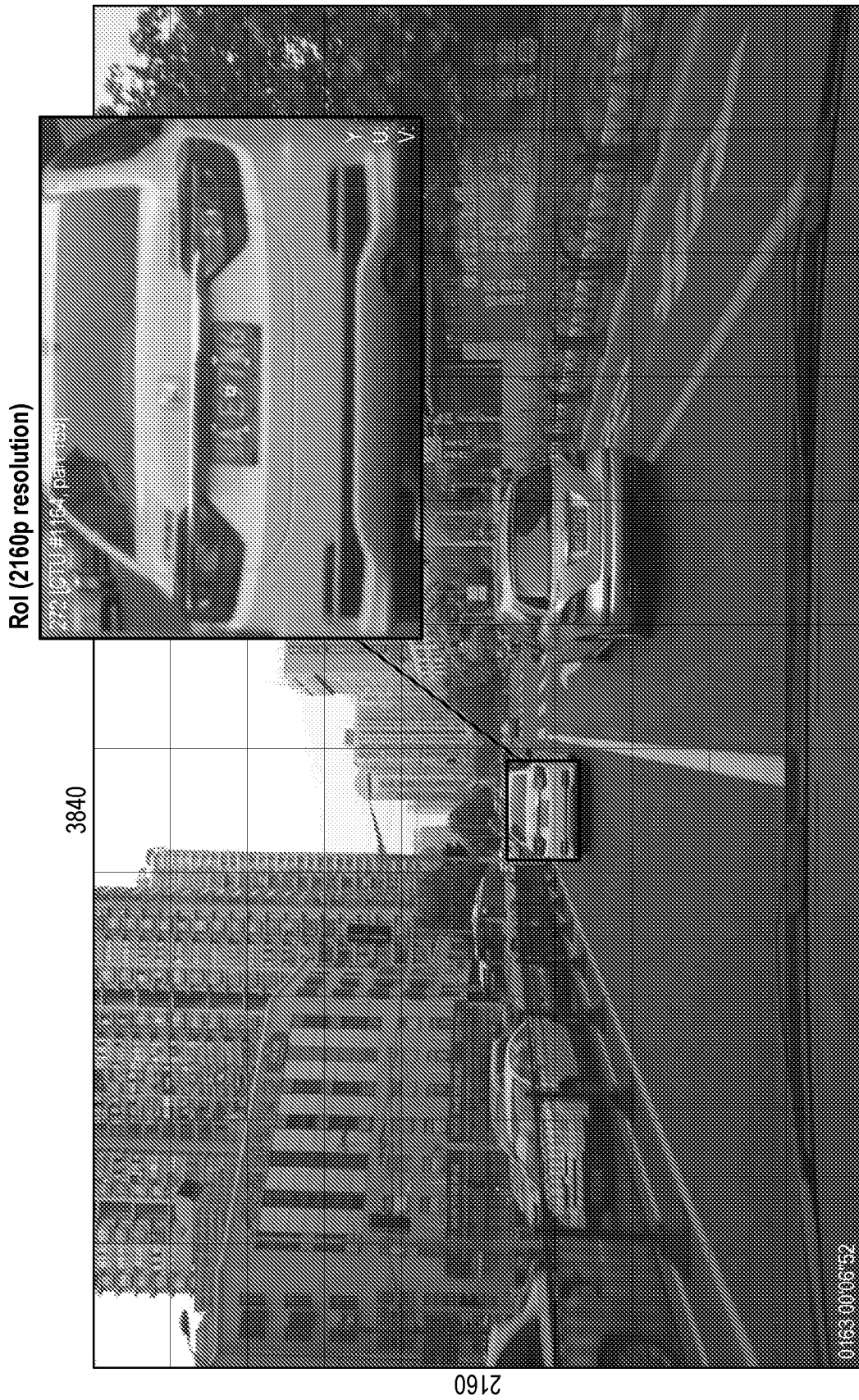
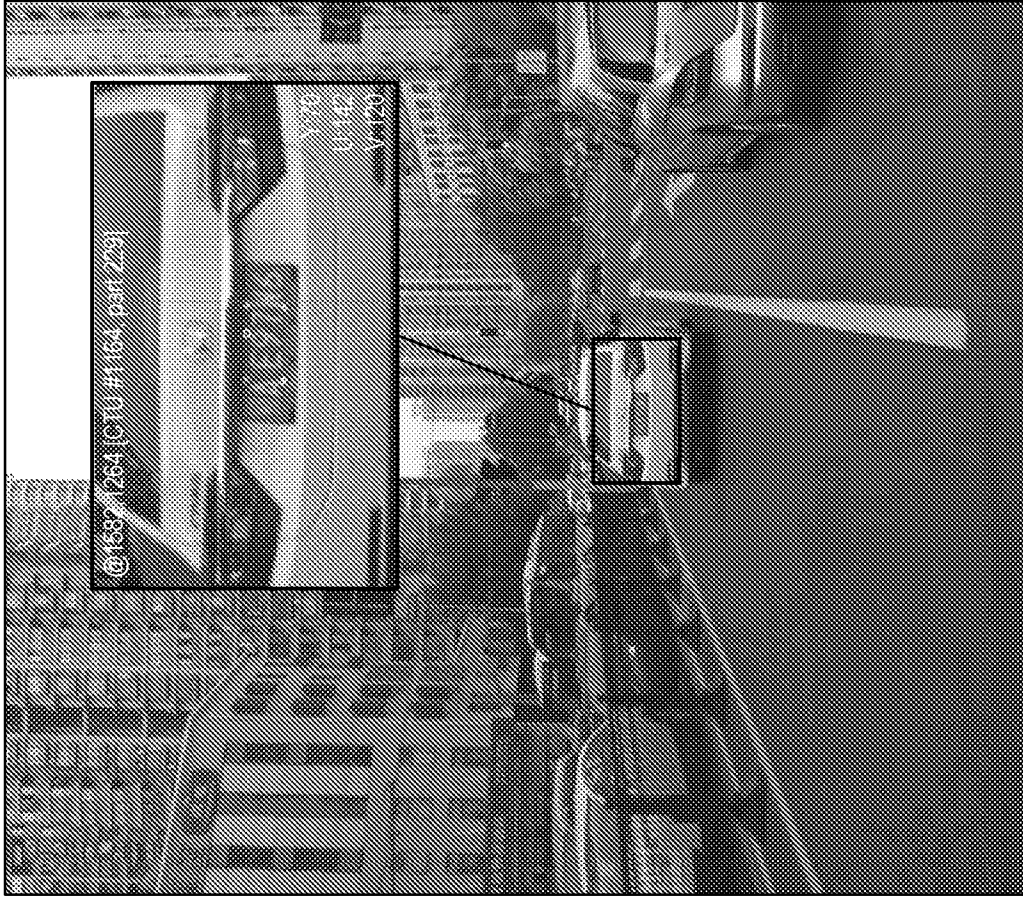
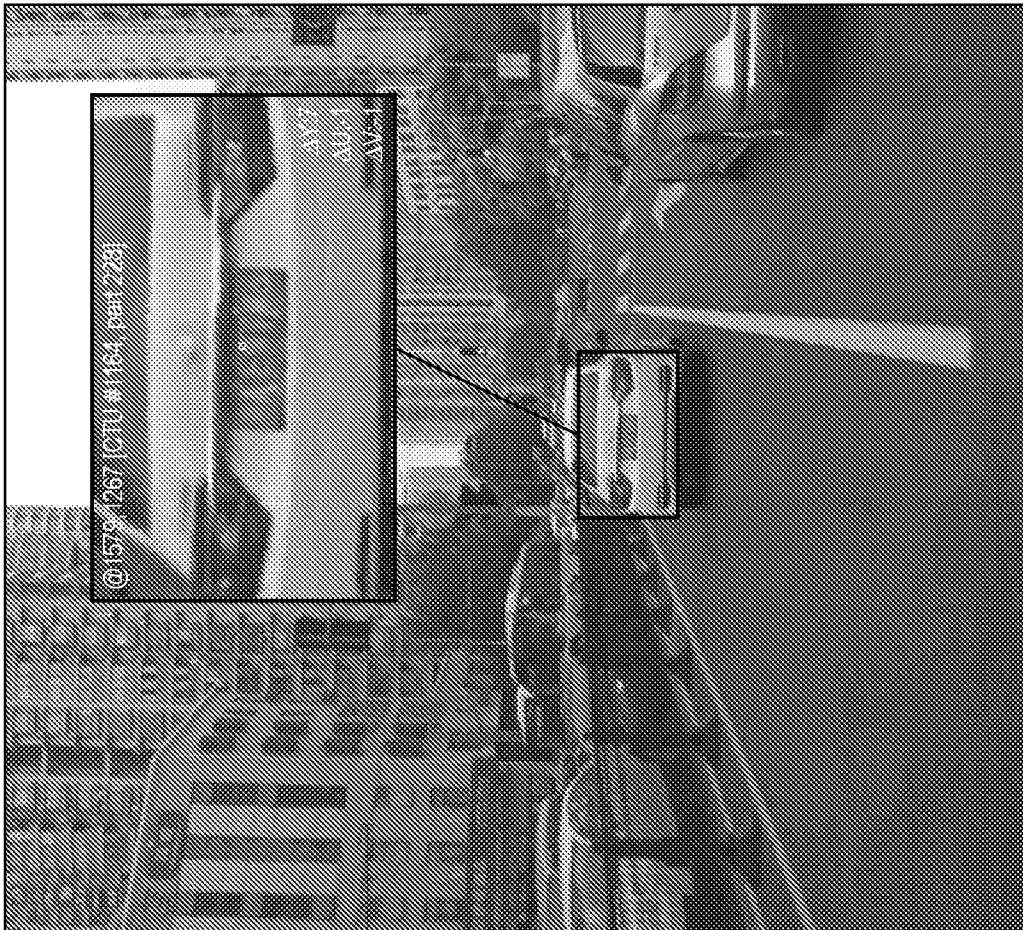


FIG. 3

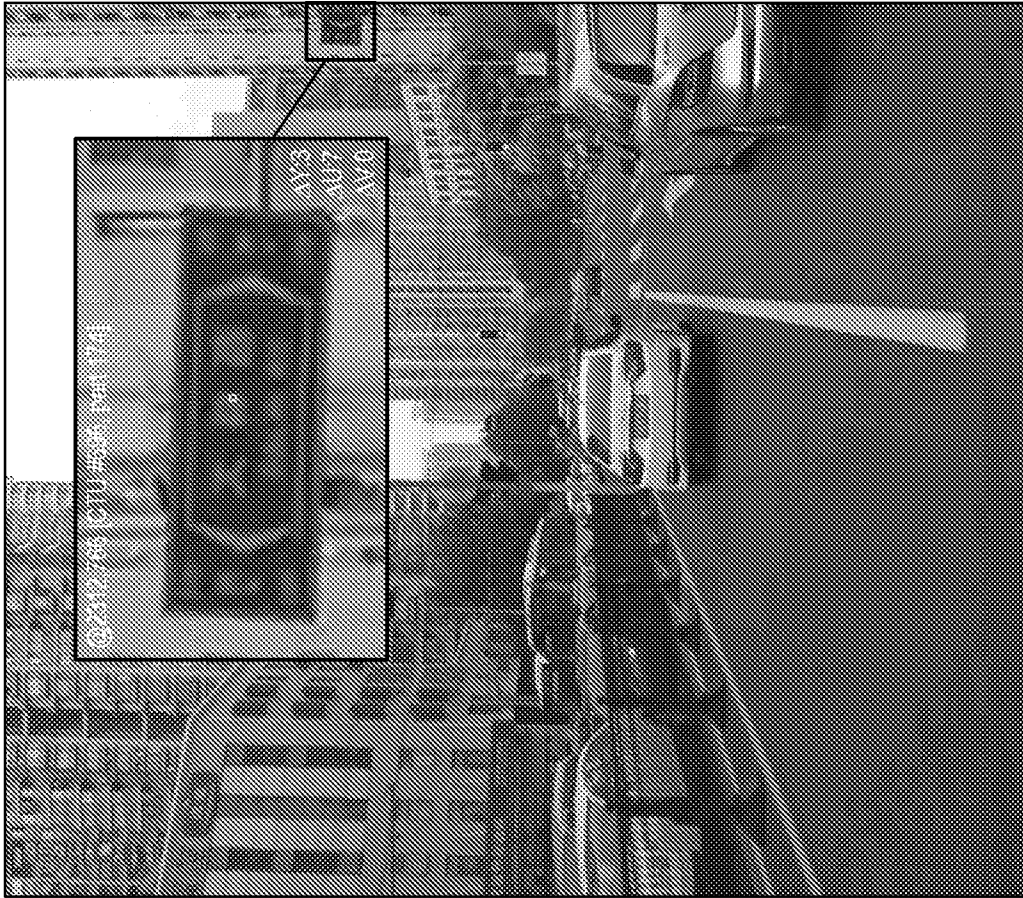


LCEVC x264, 2160p, 6 Mbps base,  
residuals applied on RoI only

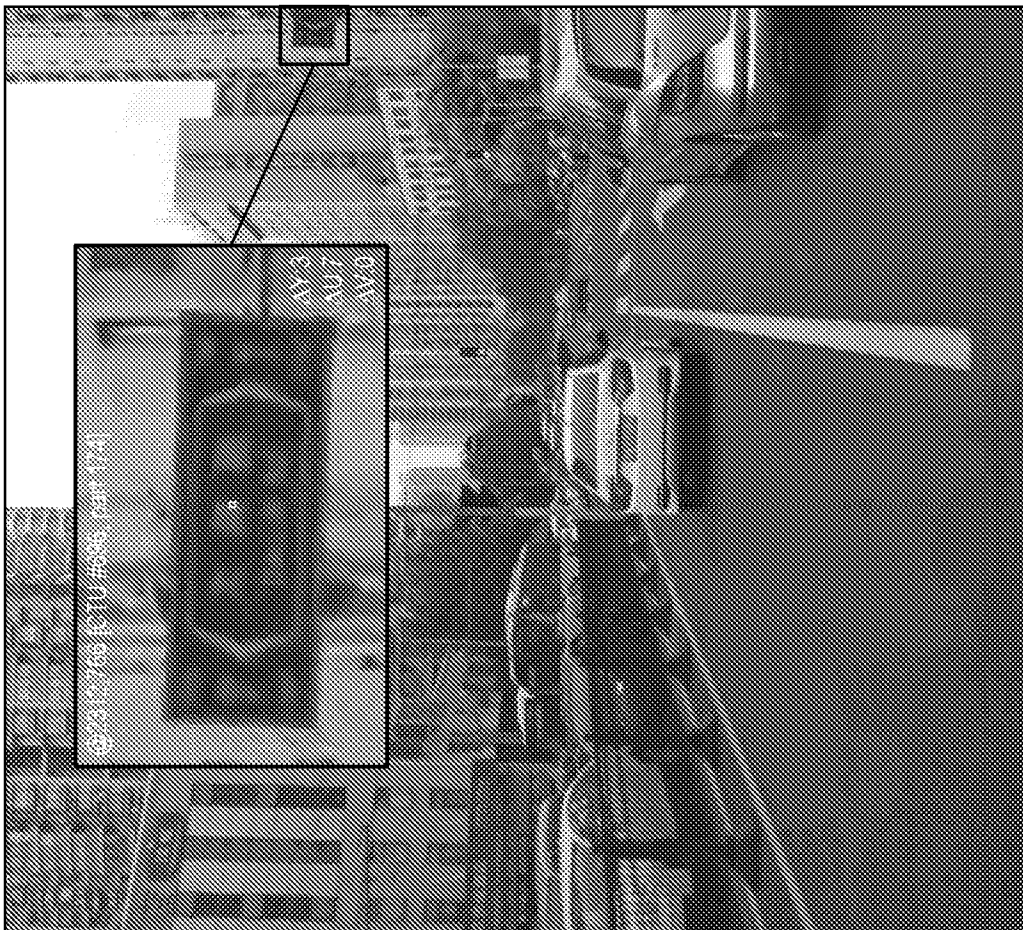


x264, 2160p, 8 Mbps  
Frame 157

**FIG. 4**



LCEVC x264, 2160p, 6 Mbps base -  
NO residuals (base only)

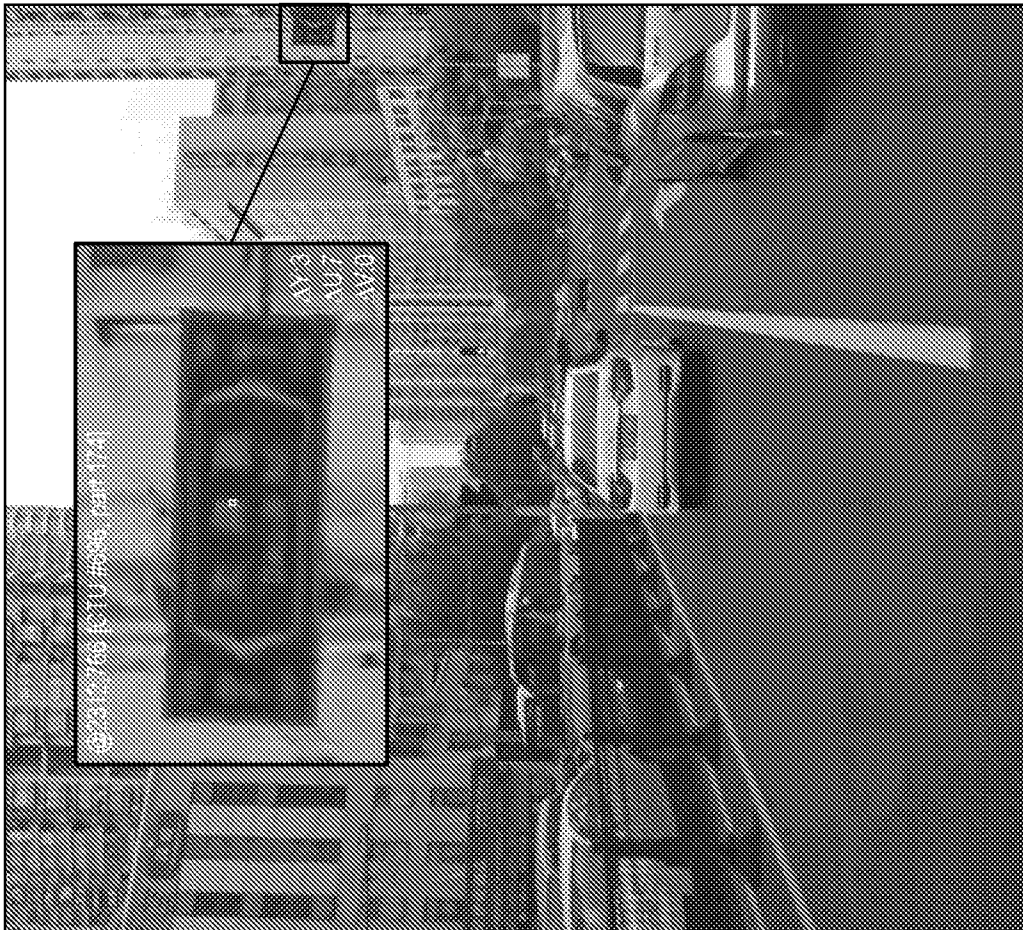


x264, 2160p, 8 Mbps  
Frame 157

FIG. 5



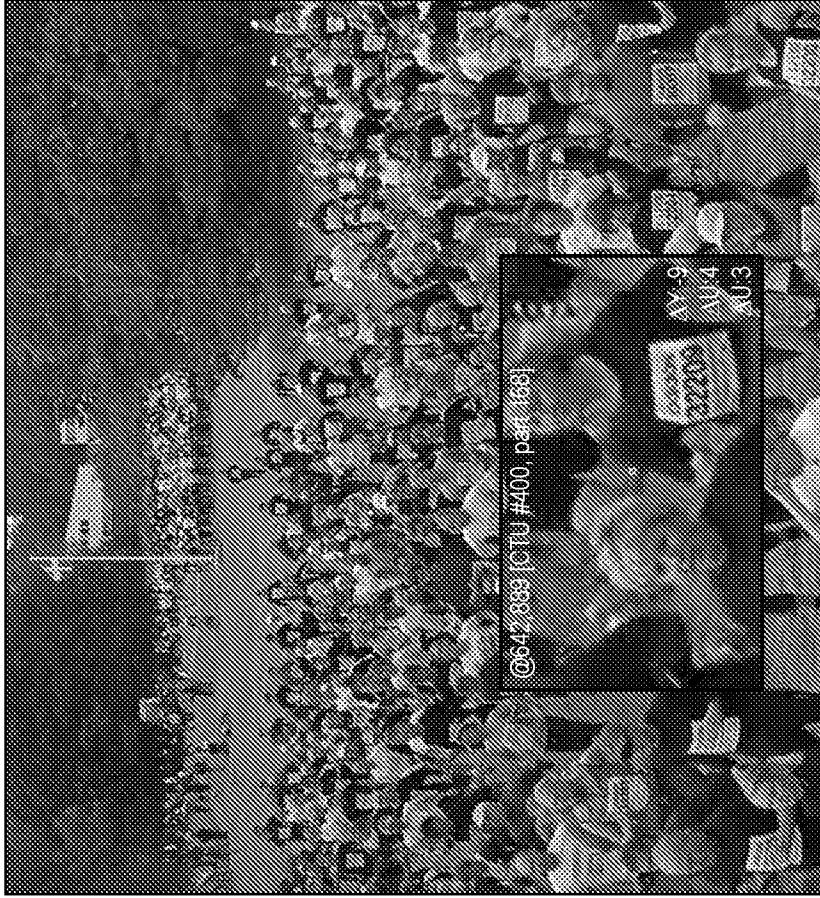
LCEVC x264, 2160p, 6 Mbps base, residuals applied on RoI only



x264, 2160p, 8 Mbps Frame 157

FIG. 6

LCEVC x264, 1080p, 5 Mbps base,  
residuals applied on RoI only



x264, 2160p, 6 Mbps  
Frame 236



FIG. 7



x264, 1080p, 8 Mbps  
Frame 236



LCEVC x264, 1080p, 6 Mbps base +  
residuals applied on Rol only

FIG. 8

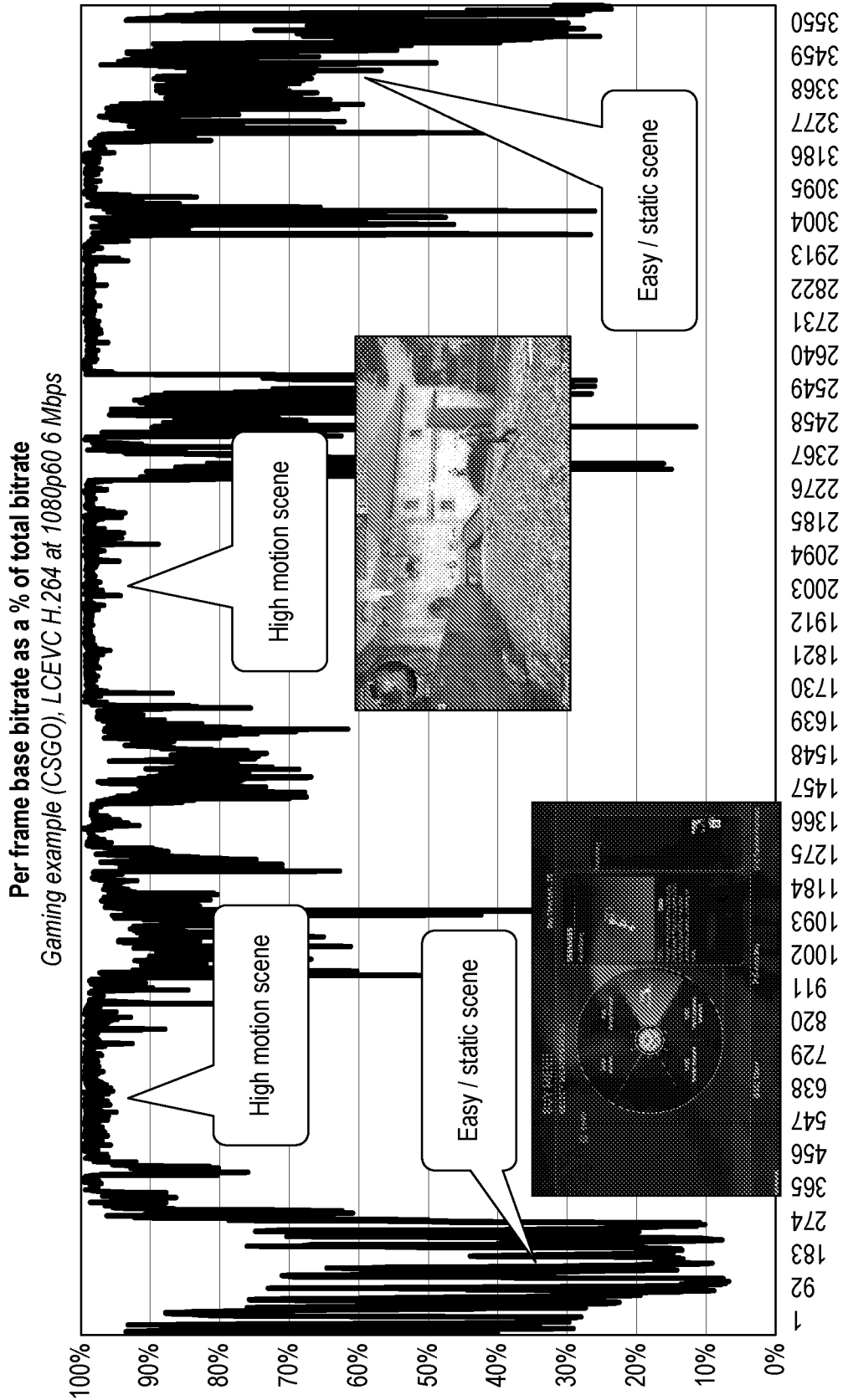


FIG. 9

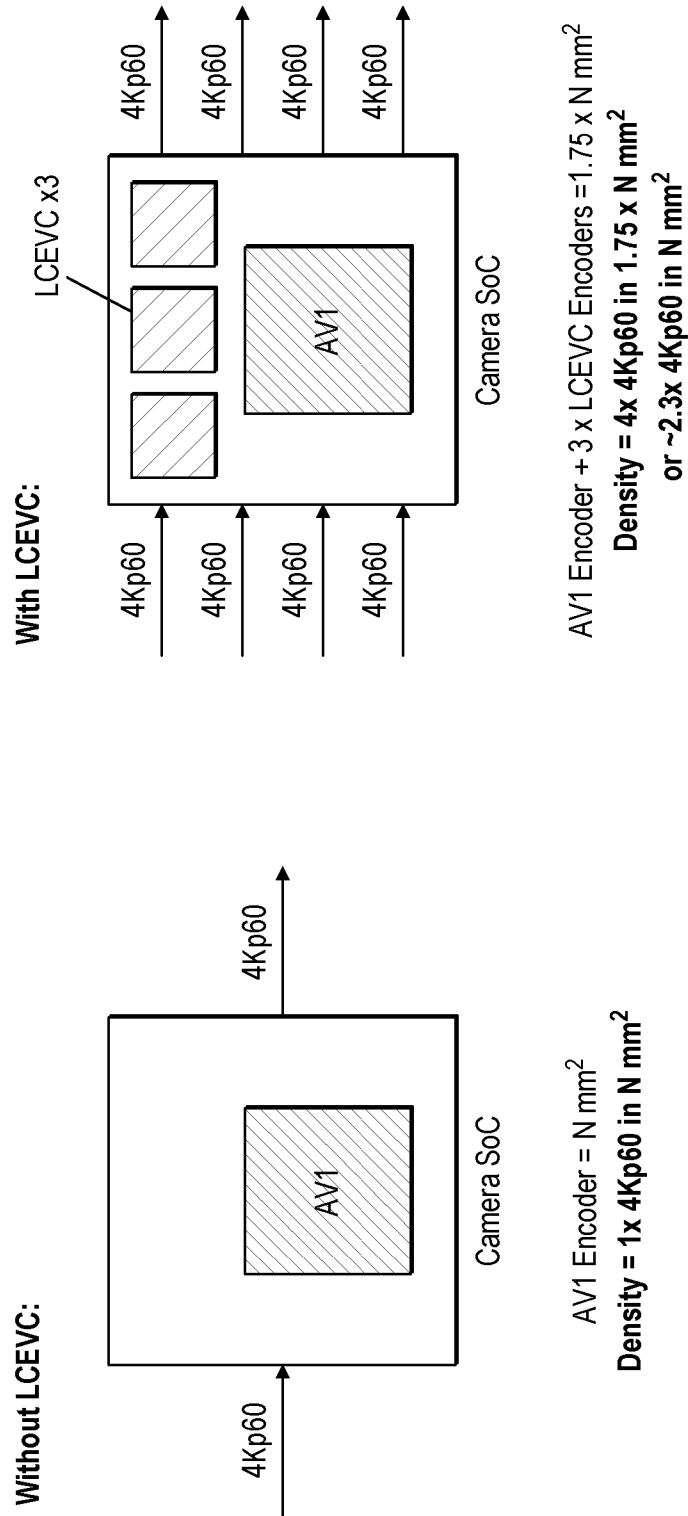


FIG. 10

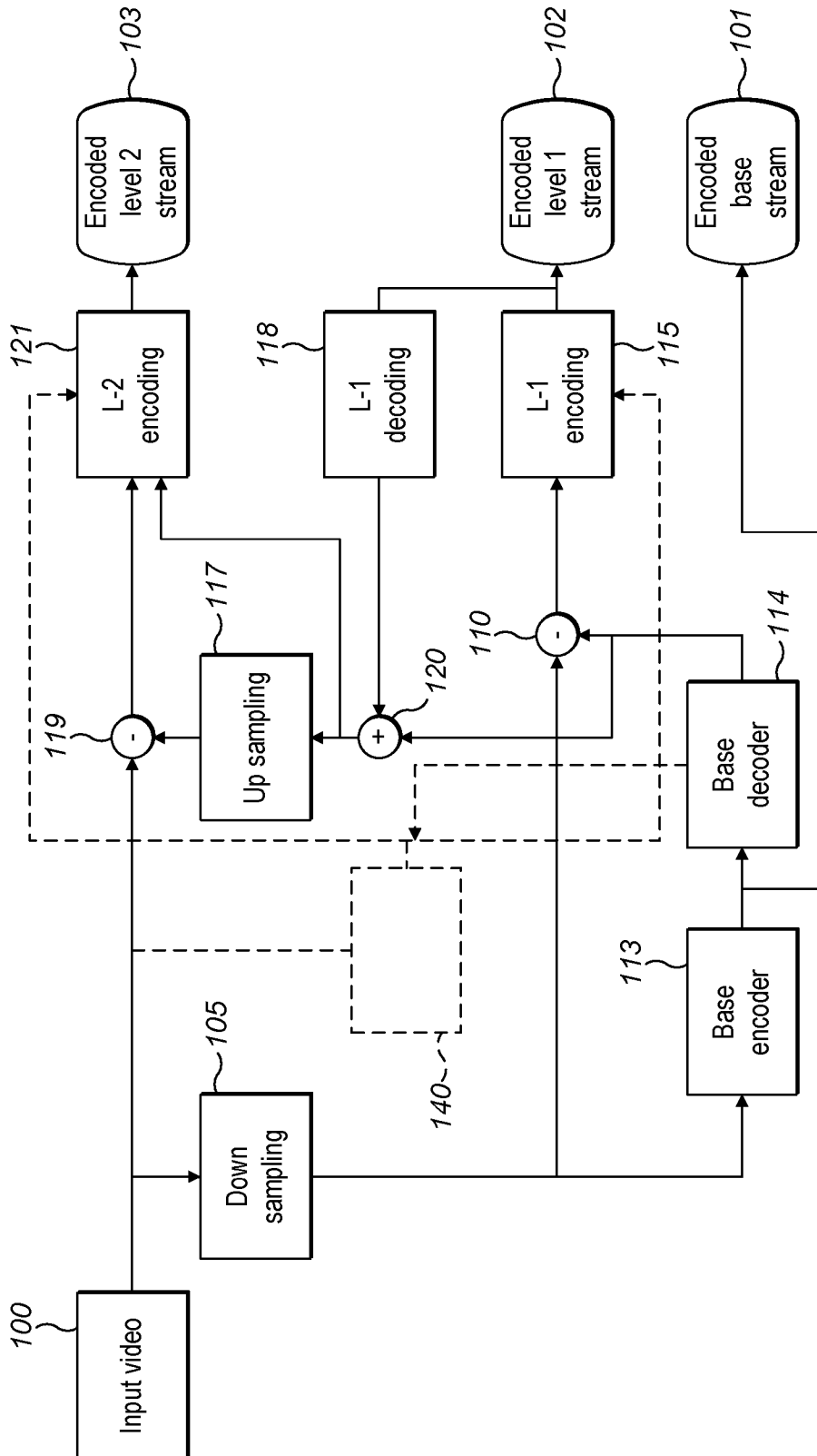


FIG. 11

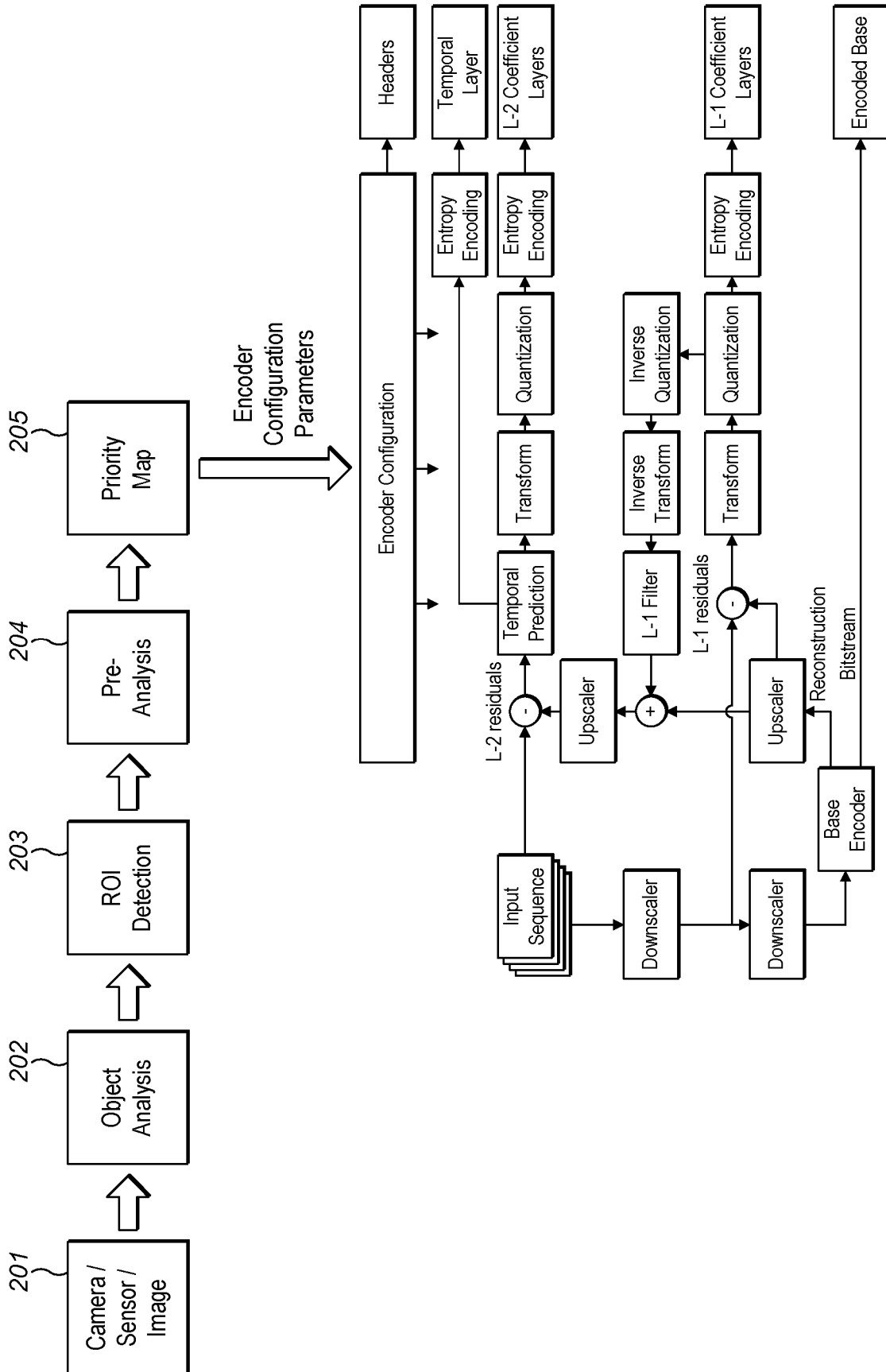


FIG. 12

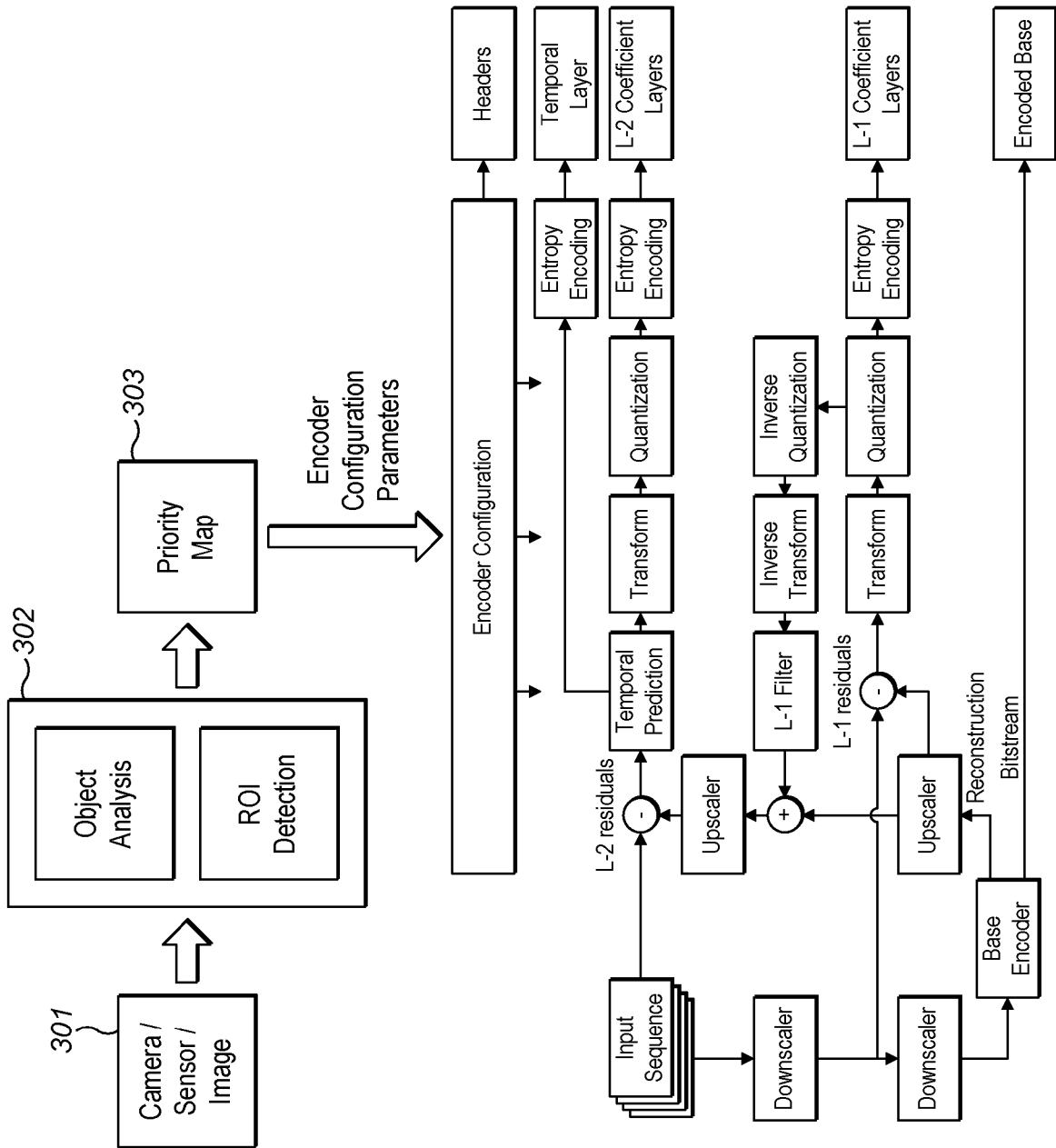
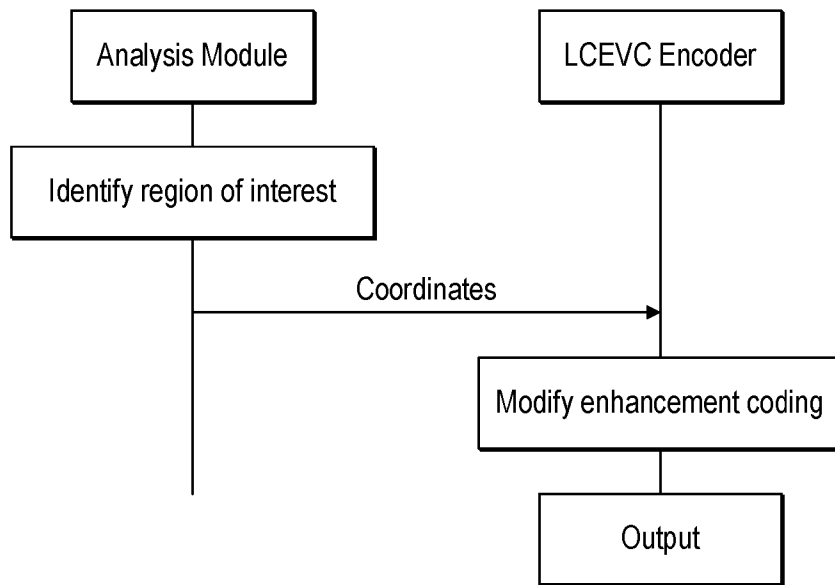


FIG. 13



**FIG. 14**

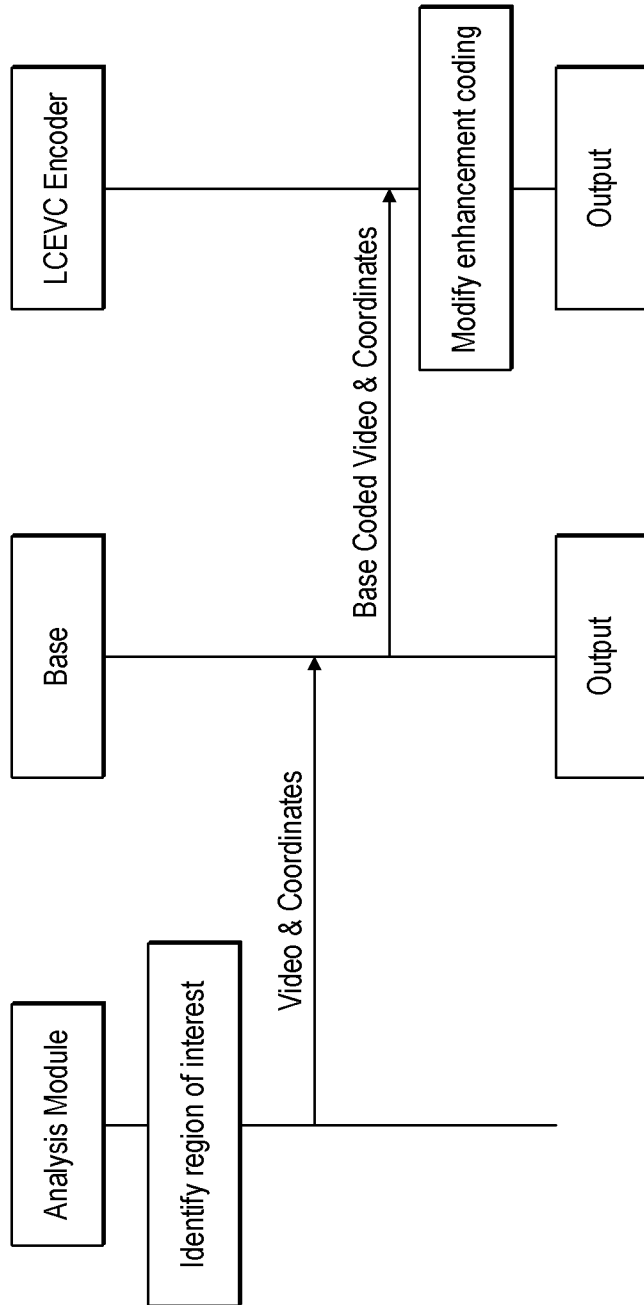


FIG. 15

# INTERNATIONAL SEARCH REPORT

International application No PCT/GB2024/053051
---

<b>A. CLASSIFICATION OF SUBJECT MATTER</b>				
INV. H04N19/124 H04N19/167 H04N19/174 H04N19/33				
ADD.				
According to International Patent Classification (IPC) or to both national classification and IPC				
<b>B. FIELDS SEARCHED</b>				
Minimum documentation searched (classification system followed by classification symbols) H04N				
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched				
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO- Internal				
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>				
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.		
X	<p>FERRARA S ET AL: "Deployment Status of the LCEVC standard", 144. MPEG MEETING; 20231016 - 20231020; HANNOVER; (MOTION PICTURE EXPERT GROUP OR ISO/IEC JTC1/SC29/WG11),</p> <p>,</p> <p>no. m65265</p> <p>9 October 2023 (2023-10-09), XP030312795, Retrieved from the Internet: URL:https://dms.mpeg.expert/doc_end_user/documents/144_Hannover/wg11/m65265-v1-m65265-DeploymentstatusoftheLCEVCstandard.zip</p> <p>m65265 - Deployment status of the LCEVC standard.docx</p> <p>[retrieved on 2023-10-09]</p> <p>page 1 - page 2</p> <p style="text-align: center;">-----</p> <p style="text-align: center;">- / - -</p>	1,9-11, 16,30, 31,36,37		
<table style="width: 100%; border: none;"> <tr> <td style="width: 50%; border: none;"><input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C.</td> <td style="width: 50%; border: none;"><input checked="" type="checkbox"/> See patent family annex.</td> </tr> </table>			<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C.	<input checked="" type="checkbox"/> See patent family annex.
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C.	<input checked="" type="checkbox"/> See patent family annex.			
<p>* Special categories of cited documents :</p> <table style="width: 100%; border: none;"> <tr> <td style="width: 50%; border: none;"> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier application or patent but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> </td> <td style="width: 50%; border: none;"> <p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>"&amp;" document member of the same patent family</p> </td> </tr> </table>			<p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier application or patent but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p>	<p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>"&amp;" document member of the same patent family</p>
<p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier application or patent but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p>	<p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>"&amp;" document member of the same patent family</p>			
Date of the actual completion of the international search		Date of mailing of the international search report		
28 January 2025		06/02/2025		
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Authorized officer  Oelbaum, Tobias		

# INTERNATIONAL SEARCH REPORT

International application No PCT/GB2024/053051
---

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2023/067541 A1 (BOYCE JILL, [US] ET AL) 2 March 2023 (2023-03-02) paragraphs [0031], [0039] - [0043], [0058], [0082]; figures 1,3,8 -----	1-37
X	US 11 172 208 B2 (NOKIA TECHNOLOGIES OY [FI]) 9 November 2021 (2021-11-09) column 7, line 57 column 16 -----	1-37
X	TIMMERER CHRISTIAN ET AL: "Special issue on Open Media Compression: Overview, Design Criteria, and Outlook on Emerging Standards", PROCEEDINGS OF THE IEEE, IEEE. NEW YORK, US, vol. 109, no. 9, 19 August 2021 (2021-08-19), pages 1423-1434, XP011873259, ISSN: 0018-9219, DOI: 10.1109/JPROC.2021.3098048 [retrieved on 2021-08-19] section G -----	1,9-11, 30

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/GB2024/053051

## Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2.  Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
  
3.  Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

**see additional sheet**

1.  As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2.  As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.
3.  As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4.  No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims;; it is covered by claims Nos.:

### Remark on Protest

- The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- No protest accompanied the payment of additional search fees.

**FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210**

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-37

Region of interest-based scalable coding

1.1. claims: 1-5, 9-15, 21-25, 27-30, 32-35

Implementation of region of interest-based scalable coding

1.2. claims: 6-8, 16-20, 26, 31, 36, 37

Determining a region of interest for scalable coding

---

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No PCT/GB2024/053051
---

Patent document cited in search report	Publication date	Patent family member(s)	Publication date	
US 2023067541	A1	02-03-2023	CN 115315954 A	08-11-2022
			EP 4136848 A1	22-02-2023
			JP 2023521553 A	25-05-2023
			KR 20230002318 A	05-01-2023
			US 2023067541 A1	02-03-2023
			WO 2021211884 A1	21-10-2021
-----				
US 11172208	B2	09-11-2021	EP 3590257 A1	08-01-2020
			US 2018249163 A1	30-08-2018
			WO 2018158496 A1	07-09-2018
-----				