

Published:

- *with international search report (Art. 21(3))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

PRIORITY-BASED VIDEO ENCODING AND TRANSMISSION**BACKGROUND**

5 **[0001]** Virtual reality (VR) allows users to experience and/or interact with an immersive artificial environment, such that the user feels as if they were physically in that environment. For example, virtual reality systems may display stereoscopic scenes to users in order to create an illusion of depth, and a computer may adjust the scene content in real-time to provide the illusion of the user moving within the scene. When the user views images through a virtual reality system,
10 the user may thus feel as if they are moving within the scenes from a first-person point of view. Similarly, mixed reality (MR) combines computer generated information (referred to as virtual content) with real world images or a real world view to augment, or add content to, a user's view of the world, or alternatively combines virtual representations of real world objects with views of a three-dimensional (3D) virtual world. The simulated environments of virtual reality and/or the
15 mixed environments of mixed reality may thus be utilized to provide an interactive user experience for multiple applications.

SUMMARY

[0002] Various embodiments of a video encoding system are described that may encode high-
20 resolution video sources at low latencies for transmission over a communications link (e.g., a wireless link) to a device for decoding and display. Embodiments of the video encoding system may also provide graceful degradation of encoded video transmitted to the device to maintain a desired frame rate in varying conditions such as variations in the channel capacity of the communications link. An example application of the video encoding system is in virtual or mixed
25 reality systems in which video frames containing virtual content are rendered, encoded, and transmitted by a base station to a device (e.g., a notebook or laptop computer, pad or tablet device, smartphone, or head-mounted display (HMD) such as a headset, helmet, goggles, or glasses that may be worn by a user) for decoding and display. Various methods and apparatus may be implemented by the video encoding system to maintain the target frame rate through the wireless
30 link and to minimize the latency in frame rendering, transmittal, and display and to provide graceful degradation of encoded video transmitted to the device to maintain a desired frame rate in varying conditions such as variations in the channel capacity of the wireless connection.

[0003] The video encoding system may decompose the pixel data into frequency bands. In some embodiments, the video encoding system may perform a wavelet transform on the pixel data
35 prior to encoding to decompose the pixel data into frequency bands. In some embodiments, the

video encoding may perform a Fourier transform that decomposes the pixel data into frequency bands. The frequency bands are then organized into blocks that are provided to a block-based encoder for encoding/compression. The encoded frequency data is then sent to a wireless interface that packetizes the encoded frequency data and transmits the packets to the receiving device. On the receiving device, the encoded data is de-packetized and passed through a block-based decoder to recover the frequency bands. In some embodiments, wavelet synthesis is then performed on the recovered frequency bands to reconstruct the pixel data for display. In some embodiments, the encoder and wireless interface components of the video encoding system may leverage the decomposition of the data into frequency bands to optimize the compression and transmission of the data to, for example, provide graceful degradation over the wireless connection.

[0004] Decomposing the pixel data into frequency bands allows the frequency bands to be prioritized by the encoder and the wireless interface. Typically, in image and video transmission, the lower frequencies are more important, while the higher frequencies are less important. Higher frequencies usually correspond to details in the image, and thus can be considered as lower priority. The higher frequency bands contain a smaller percentage of the energy in the image. Most of the energy is contained in the lower frequency bands. Decomposing the pixel data into frequency bands thus provides a priority ordering to the data stream that can be leveraged by the encoder and the wireless interface when encoding and transmitting the data stream.

[0005] The priority ordering of the frequency bands may, for example, be leveraged to provide graceful degradation of the wireless connection in the system. Performance of the wireless connection can be monitored, and feedback from the device (e.g., from an HMD or from a handheld device) may be considered, to track performance of the wireless connection. If the system is falling behind for some reason, for example if the wireless connection degrades and bandwidth capacity of the wireless connection drops below a threshold, the encoder and wireless interface may prioritize the encoding and/or transmission of one or more of the lower frequency bands, and may reduce or drop the encoding and/or transmission of one or more of the frequency levels that have been assigned a lower priority level, for example one or more of the higher frequency bands.

[0006] In some embodiments, the pixel data is first sent through a pre-filter, for example an N-channel filter bank. In some embodiments, the pre-filter may reduce high frequency content in the areas of the image that are not in a foveated region so that high fidelity can be maintained in the foveated area and lower fidelity, which cannot be detected by the human visual system, is applied in the non-foveated (peripheral) region. The output of the pre-filter is then sent through a two-level wavelet decomposition to produce multiple frequency bands organized into frequency blocks. The frequency blocks for the frequency bands are then compressed by the encoder. The

encoder may categorize the encoded frequency bands into multiple priority levels or layers. The encoder may mark each frequency block with metadata indicating the frame/slice/tile/block of the frequency block, the frequency band represented in the block, the priority of the frequency band, and timing information. The timing information may, for example, include a time at which the frequency block was generated by the encoder, and a time (referred to as a “deadline”) at which the frequency block should be received at the device.

[0007] The wireless interface may packetize the encoded frequency blocks for transmission over the wireless connection. One or more encoded frequency blocks may be included in each packet. In some embodiments, the encoded frequency blocks are packetized according to priority levels or layers. The wireless interface of the base station may then transmit or drop packets according to the priority levels of the encoded frequency blocks in the packets and/or according to the deadlines of the frequency blocks in the packets.

[0008] In some embodiments, to achieve a hard latency requirement, the timing information for each packet may include a timestamp indicating when the frequency blocks in the packet were encoded by the encoder and/or a deadline time at which the packet must be received at the device. If the deadline is missed in the transmission from the wireless interface of the base station to the wireless interface of the device, then the packet may be dropped at the base station or at the device.

[0009] In some embodiments, the wireless interface may implement different policies for dropping packets based on the priorities and/or deadlines. As an example, in some embodiments, the wireless interface may drop packets strictly on a priority basis. As another example, in some embodiments the wireless interface can inspect the deadlines of each packet and drop packets according to both deadline and priorities. As yet another example, in some embodiments the wireless interface may decide to drop all lower priority packets once a packet of a slice with a specific priority is dropped, as the other packets with lower priority are likely not to be successfully transmitted, hence freeing up resources for transmission.

[0010] In some embodiments, in addition to dividing the frequency bands into priority levels, the frequency bands may also be marked with different priorities based on whether the corresponding pixels are in the foveated region or in the peripheral region of the frame.

[0011] In some embodiments, the wireless interface of the base station may optimize the transmission of packets by maximizing the minimum number of priority levels that are successfully transmitted. This may be accomplished by assigning different modulation coding schemes to each of the different priority levels depending upon channel bandwidth availability. In some embodiments, the encoder may signal to the wireless interface a preferred packet error rate that corresponds to a specific modulation coding scheme to maximize quality.

[0012] In some embodiments, the wireless interface of the base station may optimize transmission of the packets by transmitting packets of a lower priority from the end of a tile or slice if the deadline of the packets is closer to expiration than higher priority packets belonging to a next tile or slice.

5

BRIEF DESCRIPTION OF THE DRAWINGS

[0013] FIG. 1 is a block diagram illustrating a video encoding system that decomposes pixel data into frequency bands using a wavelet transform prior to encoding, according to at least some embodiments.

10 [0014] FIG. 2 illustrates a video encoding system that includes multiple encoders that process tiles from frames in parallel, according to at least some embodiments.

[0015] FIG. 3A illustrates an example frame divided into slices and tiles, according to at least some embodiments.

15 [0016] FIG. 3B illustrates an example tile divided into blocks, according to at least some embodiments.

[0017] FIG. 3C illustrates performing a wavelet transform of a pixel block that stores pixel data to generate frequency band data prior to encoding, according to at least some embodiments.

20 [0018] FIG. 4 is a high-level flowchart of a method of operation for VR/MR systems that include video encoding systems as illustrated in FIGS. 1 and 2, according to at least some embodiments.

[0019] FIG. 5 is a flowchart of a method of operation for a video encoding system as illustrated in FIG. 1, according to at least some embodiments.

[0020] FIG. 6 is a flowchart of a method of operation for a video encoding system as illustrated in FIG. 2, according to at least some embodiments.

25 [0021] FIG. 7 is a block diagram illustrating a video encoding system as illustrated in FIGS. 1 or 2 in which frequency bands are prioritized for transmission over a wireless connection, according to at least some embodiments.

[0022] FIG. 8A illustrates an example compressed block, according to at least some embodiments.

30 [0023] FIG. 8B illustrates an example packet, according to at least some embodiments.

[0024] FIG. 9 is a flowchart of a method of operation for a video encoding system as illustrated in FIG. 7, according to at least some embodiments.

[0025] FIG. 10 illustrates an example VR/MR system that may implement a video encoding system, according to at least some embodiments.

[0026] FIG. 11 is a block diagram illustrating components of a VR/MR system as illustrated in FIG. 10, according to at least some embodiments.

[0027] This specification includes references to “one embodiment” or “an embodiment.” The appearances of the phrases “in one embodiment” or “in an embodiment” do not necessarily refer to the same embodiment. Particular features, structures, or characteristics may be combined in any suitable manner consistent with this disclosure.

[0028] “Comprising.” This term is open-ended. As used in the claims, this term does not foreclose additional structure or steps. Consider a claim that recites: “An apparatus comprising one or more processor units ...” Such a claim does not foreclose the apparatus from including additional components (e.g., a network interface unit, graphics circuitry, etc.).

[0029] “Configured To.” Various units, circuits, or other components may be described or claimed as “configured to” perform a task or tasks. In such contexts, “configured to” is used to connote structure by indicating that the units/circuits/components include structure (e.g., circuitry) that performs those task or tasks during operation. As such, the unit/circuit/component can be said to be configured to perform the task even when the specified unit/circuit/component is not currently operational (e.g., is not on). The units/circuits/components used with the “configured to” language include hardware - for example, circuits, memory storing program instructions executable to implement the operation, etc. Reciting that a unit/circuit/component is “configured to” perform one or more tasks is expressly intended not to invoke 35 U.S.C. § 112, paragraph (f), for that unit/circuit/component. Additionally, “configured to” can include generic structure (e.g., generic circuitry) that is manipulated by software or firmware (e.g., an FPGA or a general-purpose processor executing software) to operate in manner that is capable of performing the task(s) at issue. “Configure to” may also include adapting a manufacturing process (e.g., a semiconductor fabrication facility) to fabricate devices (e.g., integrated circuits) that are adapted to implement or perform one or more tasks.

[0030] “First,” “Second,” etc. As used herein, these terms are used as labels for nouns that they precede, and do not imply any type of ordering (e.g., spatial, temporal, logical, etc.). For example, a buffer circuit may be described herein as performing write operations for “first” and “second” values. The terms “first” and “second” do not necessarily imply that the first value must be written before the second value.

[0031] “Based On” or “Dependent On.” As used herein, these terms are used to describe one or more factors that affect a determination. These terms do not foreclose additional factors that may affect a determination. That is, a determination may be solely based on those factors or based, at least in part, on those factors. Consider the phrase “determine A based on B.” While in this case,

B is a factor that affects the determination of A, such a phrase does not foreclose the determination of A from also being based on C. In other instances, A may be determined based solely on B.

[0032] “Or.” When used in the claims, the term “or” is used as an inclusive or and not as an exclusive or. For example, the phrase “at least one of x, y, or z” means any one of x, y, and z, as well as any combination thereof.

DETAILED DESCRIPTION

[0033] Various embodiments of a video encoding system are described. Embodiments of the video encoding system may encode high-resolution video sources at low latencies for transmission over a communications link (e.g., a wireless link) to a device for decoding and display. Embodiments of the video encoding system may also provide graceful degradation of encoded video transmitted to the device to maintain a desired frame rate in varying conditions such as variations in the channel capacity of the communications link.

[0034] An example application of the video encoding system is in virtual or mixed reality systems in which video frames containing virtual content are rendered, encoded, and transmitted to a device for decoding and display. Embodiments of a virtual or mixed reality system (referred to herein as a VR/MR system) are described in which embodiments of the video encoding system may be implemented. In some embodiments, the VR/MR system may include a device (e.g., a pad or tablet device, a smartphone, or a headset, helmet, goggles, or glasses worn by the user, referred to herein as a head-mounted display (HMD)), and a separate computing device, referred to herein as a base station. In some embodiments, the device and base station may each include wireless communications technology that allows the device and base station to communicate and exchange data via a wireless connection. In some embodiments, the device may include sensors that collect information about the user’s environment (e.g., video, depth information, lighting information, etc.) and information about the user (e.g., the user’s expressions, eye movement, hand gestures, etc.). The information collected by the sensors may be transmitted to the base station via the wireless connection. The base station may include software and hardware (e.g., processors (system on a chip (SOC), CPUs, image signal processors (ISPs), graphics processing units (GPUs), encoder/decoders (codecs), etc.), memory, etc.) configured to generate and render frames that include virtual content based at least in part on the sensor information received from the device via the wireless connection. The base station may also include an embodiment of the video encoding system as described herein that may pre-filter, compress and transmit the rendered frames to the device for display via the wireless connection.

[0035] In some embodiments, the VR/MR system may implement a proprietary wireless communications technology that provides a highly directional wireless link between the device

and the base station. In some embodiments, the directionality and bandwidth of the wireless communication technology may support multiple devices communicating with the base station at the same time to thus enable multiple users to use the system at the same time in a co-located environment. However, other commercial (e.g., Wi-Fi, Bluetooth, etc.) or proprietary wireless communications technologies may be supported in some embodiments.

[0036] Primary constraints to be considered on a wireless link include bandwidth and latency. A target of the VR/MR system is to provide a high resolution, wide field of view (FOV) virtual display (at a frame rate to provide the user with a high-quality VR/MR view. Another target is to minimize latency between the time a frame is rendered by the base station and the time the frame is displayed by the device. However, the channel capacity of the wireless link may vary with time, and the wireless link may thus support only a certain amount of information to be transmitted at any given time.

[0037] Various methods and apparatus are described herein that may be implemented by the video encoding system to maintain the target frame rate through the wireless link and to minimize the latency in frame rendering, transmittal, and display. In addition, the methods and apparatus may provide graceful degradation of encoded video transmitted to the device to maintain a desired frame rate in varying conditions such as variations in the channel capacity of the communications link.

[0038] In some embodiments, the video encoding system may perform a wavelet transform on the pixel data prior to encoding to decompose the pixel data into frequency bands. The frequency bands are then organized into blocks that are provided to a block-based encoder for encoding/compression. As an example, a frame may be divided into 128x128 blocks, and a two-level wavelet decomposition may be applied to each 128x128 block to generate 16 32x32 blocks of frequency data representing seven frequency bands that may then be sent to an encoder (e.g., a High Efficiency Video Coding (HEVC) encoder) to be encoded. The encoded frequency data is then sent to a wireless interface that packetizes the encoded frequency data and transmits the packets to the receiving device (e.g., an HMD, pad or tablet device, smartphone, etc.). On the receiving device, the encoded data is de-packetized and passed through a block-based decoder to recover the frequency bands. Wavelet synthesis is then performed on the recovered frequency bands to reconstruct the pixel data for display.

[0039] While embodiments are generally described in which the wavelet transform is a two-level wavelet decomposition applied to each pixel block from a video frame, in various embodiments the wavelet decomposition may be any number of levels (e.g., one level, two levels, three levels, four levels, etc.), and may be adjusted to trade-off quality of the encoded image vs. complexity of the blocks to be encoded.

[0040] In some embodiments, the video encoding system may perform slice-based rendering, encoding, and transmittal. Rendering, encoding, and transmitting entire frames may have a latency and memory impact as each frame needs to be completed, stored, and then transmitted to the next stage of the VR/MR system. In slice-based rendering, rather than rendering and encoding entire frames in the base station and transmitting the rendered frames to the device, the base station may render and encode parts of frames (referred to as slices) and transmit the encoded slices to the device as they are ready. A slice may, for example, be a row of 128x128 blocks, or two or more rows of blocks. Slice-based rendering and encoding may help to reduce latency, and may also reduce the amount of memory needed for buffering, which may reduce the memory footprint on the chip(s) or processor(s) as well as power requirements.

[0041] In some embodiments, the video encoding system may perform tile-based rendering, encoding, and transmittal. In tile-based rendering, encoding, and transmittal, each slice may be divided into multiple tiles (e.g., four tiles), and the base station may render and encode the tiles and transmit the encoded tiles to the device as they are ready.

[0042] In some embodiments, the video encoding system may perform tile-based encoding using a single encoder to process tiles from each slice. However, in some embodiments, the video encoding system may perform tile-based encoding using multiple encoders to process respective tiles from each slice. For example, in some embodiments, each slice may be divided into four tiles, each tile including multiple 128x128 blocks, and two encoders (e0 and e1) may operate on two tiles from each slice (e.g., e0 operates on t0 and t1; e1 operates on t2 and t3). Each encoder may multiplex the processing of blocks from different frequency bands between its two tiles to allow for 16 time units between the processing of blocks from the same frequency band. By multiplexing the processing of blocks between two tiles, dependencies between blocks in a frequency band may be handled appropriately. While embodiments are described in which each slice is divided into four tiles and two encoders operate on respective tiles from each slice, slices may be divided into more tiles (e.g., six or eight tiles) in some embodiments, and more encoders (e.g., three or four encoders) may be used in some embodiments.

[0043] In some embodiments, the video encoding system may perform pre-filtering of the pixel data in frames prior to the wavelet transform. Pre-filtering may, for example, reduce the resolution of the frames rendered by the base station prior to performing the wavelet transform, encoding, and transmission of the frames to the device over the wireless link, which may help in improving compression, and may reduce latency and bandwidth usage on the wireless link.

[0044] In some embodiments in which the device is an HMD, pre-filtering may include performing a lens warp on the frames on the base station prior to the wavelet transform. The lens warp is performed to correct for the distortion of the images introduced by the lenses on the HMD

that the images are viewed through, thus improving quality of the images. In some embodiments, the HMD may store lens warp data for the lenses, for example generated by a calibration process, and may provide the lens warp data to the base station over the wireless connection. The base station may then perform the lens warp on the frames based on the lens warp data for that HMD.

5 In conventional VR/MR systems, the lens warp is performed on the HMD after decoding and prior to display. Performing the lens warp on the base station in the pre-filter stage may reduce the resolution of the frames prior to performing the wavelet transform and encoding, which may help in improving compression, and may reduce latency and bandwidth usage on the wireless link. In addition, by performing the lens warp on the base station in the pre-filter stage rather than on the
10 HMD after decoding, filtering of the image data may only need to be performed once, as opposed to performing filtering on the base station to reduce resolution prior to encoding and then performing lens warp filtering on the HMD.

[0045] In some embodiments, pre-filtering may include filtering to reduce resolution in peripheral regions while maintaining higher resolution in foveated regions. In this method, gaze
15 tracking information obtained from the device may be used to identify the direction in which the user is currently looking. Human eyes can perceive higher resolution in the foveal region than in the peripheral region. Thus, a region of the frame that corresponds to the fovea (referred to as the foveated region) may be identified based at least in part on the determined gaze direction. In some
20 embodiments, the peripheral region (i.e. the portion of the frame outside the foveated region) may be pre-filtered to reduce information based on knowledge of the human vision system, for example by filtering high frequency information and/or increasing color compression. In some
25 embodiments, the amount of filtering applied to the peripheral region may increase extending towards the periphery of the frame. Pre-filtering of the peripheral region may help to provide improved compression of the frame.

[0046] FIG. 1 is a block diagram illustrating a video encoding system 120 that decomposes
30 pixel data into frequency bands using a wavelet transform prior to encoding, according to at least some embodiments. A VR/MR system 10 may include at least one device 150 (e.g., a pad or tablet device, a smartphone, or an HMD such as a headset, helmet, goggles, or glasses that may be worn by a user) and a computing device 100 (referred to herein as a base station). The base station 100
35 renders VR or MR frames including virtual content, encodes the frames, and transmits the encoded frames over a wireless connection 180 to the device 150 for decoding and display by the device 150.

[0047] In some embodiments, the device 150 may include sensors 160 that collect information
about the user 190's environment (e.g., video, depth information, lighting information, etc.) and
35 about the user 190 (e.g., the user's expressions, eye movement, gaze direction, hand gestures, etc.).

The device 150 may transmit at least some of the information collected by sensors 160 to the base station 100 via wireless connection 180. The base station 100 may render frames for display by the device 150 that include virtual content based at least in part on the various information obtained from the sensors 160, encode the frames, and transmit the encoded frames to the device 150 for decoding and display to the user via the wireless connection 180.

[0048] The base station 100 and device 150 may implement wireless communications technology that allows the base station 100 and device 150 to communicate and exchange data via a wireless connection 180. In some embodiments, the wireless connection 180 may be implemented according to a proprietary wireless communications technology that provides a highly directional wireless link between the device 150 and the base station 100. However, other commercial (e.g., Wi-Fi, Bluetooth, etc.) or proprietary wireless communications technologies may be used in some embodiments.

[0049] Primary constraints to be considered on the wireless connection 180 between the device 150 and the base station 100 in a VR/MR system 10 include bandwidth and latency. For example, in some embodiments, a target is to provide a high resolution, wide field of view (FOV) virtual display to the user at a frame rate that provides the user with a high-quality VR/MR view. Another target is to minimize latency between the time a video frame is captured by the device 150 and the time a rendered VR/MR frame based on the video frame is displayed by the device 150.

[0050] The base station 100 may include various hardware components for rendering, filtering, encoding, and transmitting video and/or images as described herein, for example various types of processors, integrated circuits (ICs), central processing units (CPUs), graphics processing units (GPUs), image signal processors (ISPs), encoder/decoders (codecs), etc. The base station 100 may include, but is not limited to, a GPU rendering 110 component, a wireless interface 130 component, and a video encoding system 120 that may include one or more hardware components that implement various methods that may help to maintain the target frame rate through the wireless connection 180 and to minimize the latency in frame rendering, encoding, transmittal, and display. The video encoding system 120 may include, but is not limited to, a pre-filter 122 component (e.g., an N-channel filter bank), a wavelet transform 124 component, and an encoder 126 component.

[0051] GPU rendering 110 may include one or more hardware components that may render frames for display by the device 150 that include virtual content based at least in part on the various information obtained from the sensors 160.

[0052] In some embodiments, the video encoding system 120 may include one or more hardware components that pre-filter 122 the pixel data in the rendered frames prior to performing a wavelet transform 124. Pre-filter 122 may, for example, reduce the resolution of the frames rendered on the base station 100 prior to performing the wavelet transform 124, encoding 126, and

transmission to the device 150 over the wireless connection 180, which may help in improving compression, and may reduce latency and bandwidth usage on the wireless connection 180.

[0053] In some embodiments, pre-filter 122 may perform a lens warp on the frames on the base station 100 prior to the wavelet transform 124. The lens warp is performed to correct for the distortion of the images introduced by the lenses on the device 150 that the images are viewed through, thus improving quality of the images. In some embodiments, the device 150 may store lens warp data for the lenses, for example generated by a calibration process, and may provide the lens warp data to the base station 100 over the wireless connection 180. The pre-filter 122 component of the video encoding system 120 may then perform the lens warp on the frames based on the lens warp data for that device 150. In conventional VR/MR systems, the lens warp is performed on the device 150 after decoding and prior to display. Performing the lens warp on the base station 100 in the pre-filter 122 stage may reduce the resolution of the frames prior to performing the wavelet transform 124 and encoding 126, which may help in improving compression, and may reduce latency and bandwidth usage on the wireless connection 180. In addition, by performing the lens warp on the base station 100 in the pre-filter 122 stage rather than on the device 150 after decoding, filtering of the image data may only need to be performed once, as opposed to performing filtering on the base station 100 to reduce resolution prior to encoding 126 and then performing lens warp filtering on the device 150.

[0054] In some embodiments, pre-filter 122 may also apply one or more filters to reduce resolution in peripheral regions while maintaining higher resolution in foveated regions. In this method, gaze tracking information obtained from the device 150 may be used to identify the direction in which the user is currently looking. Human eyes can perceive higher resolution in the foveal region than in the peripheral region. Thus, a region of the frame that corresponds to the fovea (referred to as the foveated region) may be identified based at least in part on the determined gaze direction. In some embodiments, the peripheral region (i.e. the portion of the frame outside the foveated region) may be pre-filtered to reduce information based on knowledge of the human vision system, for example by filtering high frequency information and/or increasing color compression. In some embodiments, the amount of filtering applied to the peripheral region may increase extending towards the periphery of the frame. Pre-filtering of the peripheral region may help to provide improved compression of the frame.

[0055] In some embodiments, a wavelet transform 124 component of the video encoding system 120 may include one or more hardware components (e.g., an N-channel filter bank) that perform a wavelet transform on the pixel data prior to encoding to decompose the pixel data into frequency bands. The frequency bands are then organized into blocks that are provided to a block-based encoder 126 for encoding/compression. As an example, as illustrated in FIGS. 3A through

3C, a frame may be divided into 128x128 blocks, and a two-level wavelet decomposition may be applied to each 128x128 block to generate 16 32x32 blocks of frequency data representing seven frequency bands that may then be sent to a block-based encoder (e.g., a High Efficiency Video Coding (HEVC) encoder) 126 to be encoded. The encoded frequency data is then sent to a wireless interface 130, implemented by one or more hardware components, that packetizes the data and transmits the packets to the device 150 over a wireless connection 180.

[0056] The device 150 may include various hardware components for decoding and displaying video and/or images as described herein, for example various types of processors, integrated circuits (ICs), central processing units (CPUs), graphics processing units (GPUs), image signal processors (ISPs), encoder/decoders (codecs), etc. The device 150 may include, but is not limited to, a wireless interface 152, a decoder 154 component (e.g., High Efficiency Video Coding (HEVC) decoder), a wavelet synthesis 156 component, and a display 158 component. On the device 150, the wireless interface 152 receives the packets that were transmitted over the wireless connection 180 by the base station 100. The encoded data is de-packetized and passed through a block-based decoder 154 (e.g., a High Efficiency Video Coding (HEVC) decoder) to recover the frequency bands. Wavelet synthesis 156 is then performed on the recovered frequency data to reconstruct the pixel data for display 158.

[0057] In some embodiments, the video encoding system 120 may perform slice-based rendering, encoding, and transmittal. Rendering, encoding, and transmitting entire frames may have a latency and memory impact as each frame needs to be completed, stored, and then transmitted to the next stage of the VR/MR system 10. In slice-based rendering, rather than rendering and encoding entire frames in the base station 100 and transmitting the rendered frames to the device 150, the base station 100 may render and encode parts of frames (referred to as slices) and transmit the encoded slices to the device 150 as they are ready. A slice may, for example, be a row of 128x128 blocks. Slice-based rendering and encoding may help to reduce latency, and may also reduce the amount of memory needed for buffering, which reduces the memory footprint on the chip(s) or processor(s) as well as power requirements.

[0058] In some embodiments, the video encoding system 120 may perform tile-based rendering, encoding, and transmittal. In tile-based rendering, encoding, and transmittal, each slice may be divided into multiple tiles (e.g., four tiles), and the base station 100 may render and encode the tiles and transmit the encoded tiles to the device 150 as they are ready.

[0059] In some embodiments, the video encoding system 120 may perform tile-based rendering, encoding, and transmittal using a single encoder 126 to process tiles from each slice. However, in some embodiments, the video encoding system 120 may perform tile-based encoding using multiple encoders 126 to process respective tiles from each slice. FIG. 2 illustrates a video

encoding system 220 that includes multiple encoders (two encoders 226A and 226B, in this example) that process tiles from rendered frames in parallel, according to at least some embodiments.

5 **[0060]** A GPU rendering 210 component of the base station 200 may include one or more GPUs and/or other components that render frames (or slices of frames) for display. A frame may be divided into slices, for example as illustrated in FIG. 3A. As illustrated in FIG. 3A, each slice may be divided into multiple tiles (four, in this example), each tile including multiple blocks. FIG. 3B illustrates an example tile that includes four 128x128 blocks. However, blocks of other sizes (e.g. 64x64, 32x32, etc.) may be used in some embodiments, and a tile may include more or fewer
10 blocks.

[0061] Pre-filter 222 and wavelet transform 224 components of the video encoding system 220 may then process each tile prior to encoding 226. In some embodiments, video encoding system 220 may include a separate pre-filter 222 component and wavelet transform 224 component for processing each tile. In this example, pre-filter 222A component and wavelet transform 224A
15 component process tile 0, pre-filter 222B component and wavelet transform 224B component process tile 1, pre-filter 222C component and wavelet transform 224C component process tile 2, and pre-filter 222D component and wavelet transform 224D component process tile 3. The pre-filter 222 components perform pre-filtering of the tiles as described herein, and the wavelet transform 224 components decompose the tiles into frequency bands as described herein.
20 However, in some embodiments, video encoding system 220 may include a single pre-filter 222 component and a single wavelet transform 224 component that process the tiles. In some embodiments, video encoding system 220 may include multiple (e.g., 2) pre-filter 222 components and multiple (e.g., 2) wavelet transform 224 components that each process multiple (e.g., 2) tiles.

[0062] Two encoders 226A and 226B may operate on two tiles from each slice (e.g., encoder
25 226A operates on tile 0 and tile 1; encoder 226B operates on tile 2 and tile 3). Each encoder 226 may multiplex the processing of blocks from different frequency bands (i.e., the 16 32x32 blocks illustrated in FIG. 3C) between its two tiles to allow for 16 time units between the processing of blocks from the same frequency band. By multiplexing the processing of blocks between two tiles, dependencies between blocks in the same frequency band may be handled appropriately.

30 **[0063]** While embodiments are described in which each slice is divided into four tiles and two encoders operate on respective tiles from each slice, slices may be divided into more tiles (e.g., six or eight tiles) in some embodiments, and more encoders (e.g., three, four, or more encoders) may be used in some embodiments.

[0064] FIG. 3C illustrates performing a wavelet transform of a pixel block that stores pixel data
35 to generate frequency band data prior to encoding, according to at least some embodiments. In

this example, a two-level wavelet decomposition is applied by the wavelet transform 324 component to a 128x128 pixel block 300 to generate sixteen 32x32 blocks 302 of frequency data representing seven frequency bands. The frequency blocks 302 are then provided to an encoder 326 for encoding. For example, the frequency blocks 302 may be written to a buffer by the wavelet transform 324 component, and read from the buffer by the encoder 326 component.

[0065] In the labels of the frequency blocks 302, the letter L represents a low pass filter, and the letter H represents a high pass filter. The blocks 302 labeled with two letters represent a one-level (2D) wavelet transform or decomposition. In the blocks 302 labeled with two letters (representing three of the seven frequency bands LH, HL, and HH), the first letter represents a vertical filter (either high or low) performed first, and the second letter represents a horizontal filter (either high or low) performed second. The blocks 302 labeled with four letters represent a two-level wavelet transform or decomposition. In the blocks 302 labeled with four letters, the first two letters (LL) indicate that there was first a vertical low pass filter followed by a horizontal low pass filter; the second two letters indicate that the resulting LL block was then filtered four ways, LL, LH, HL, and HH (thus generating four of the seven frequency bands (LLLL, LLLH, LLHL, and LLHH)).

[0066] Decomposing the pixel data into frequency bands as illustrated in FIG. 3C allows the frequency bands to be buffered and processed as separate streams by the encoder 326. Processing the frequency bands as separate streams allows the encoder 326 component to multiplex the processing of the independent streams. In block-based encoding methods such as HEVC encoding, blocks (referred to as coding tree units (CTUs)) are processed in a block processing pipeline at multiple stages; two or more blocks may be at different stages of the pipeline at a given clock cycle, and the blocks move through the pipeline as the clock cycles. The processing of a given block may have dependencies on one or more previously processed neighbor blocks, for example one or more blocks in the row above the given block and/or the block to the left of the given block. By multiplexing the processing of the streams of frequency band data, the encoder 326 spaces out the processing of the blocks in a given stream, thus providing additional clock cycles to process a neighbor block on which a given block has dependencies. For example, the block to the left of the given block may be several stages ahead of the given block in the pipeline when the given block reaches a stage that depends on the previously processed neighbor block. This allows the encoder 326 to better handle dependencies on previously processed blocks, and reduces or eliminates the need to wait for completion of processing of a neighbor block in the pipeline before processing the given block at a stage that depends on the neighbor block.

[0067] In addition, decomposing the pixel data into frequency bands as illustrated in FIG. 3C allows the frequency bands to be prioritized by the encoder 326 and the wireless interface.

Typically, in image and video transmission, the lower frequencies are more important, while the higher frequencies are less important. Higher frequencies usually correspond to details in the image, and thus can be considered as lower priority. The higher frequency bands contain a smaller percentage of the energy in the image. Most of the energy is contained in the lower frequency bands. Decomposing the pixel data into frequency bands thus provides a priority ordering to the data stream that can be leveraged by the encoder 326 and the wireless interface when encoding and transmitting the data stream. For example, in some embodiments, different compression techniques may be used on the different frequency bands, with more aggressive compression applied to the lower priority bands, and more conservative compression applied to the higher priority bands. As another example, the priority ordering of the frequency bands may help in providing graceful degradation of the VR/MR system. Performance of the wireless connection can be monitored, and feedback from the device (e.g., from an HMD) may be considered, to track performance of the overall system. If the system is falling behind for some reason, for example if the wireless connection degrades and bandwidth capacity of the wireless connection drops below a threshold, the encoder 326 and wireless interface may prioritize the encoding and transmission of one or more of the lower frequency bands, and may reduce or drop the encoding and/or transmission of one or more of the frequency levels that have been assigned a lower priority level, for example one or more of the higher frequency bands.

[0068] As described above, the wavelet transform decomposes an image into frequency bands. In some embodiments, this may be leveraged to send the same signal to displays of varying resolution. As an example, suppose that a two-level wavelet decomposition is applied to decompose the signal into seven bands. If four of the bands are sent (LLLL, LLLH, LLHL and LLHH), the bands may be reconstructed to the original intended resolution at less visual quality. As an alternative, the bands may also be reconstructed at 1/4th resolution (1/2 in each dimension) which may be suitable for a display panel with smaller display resolution.

[0069] FIG. 4 is a high-level flowchart of a method of operation for VR/MR systems that include video encoding systems as illustrated in FIGS. 1 and 2, according to at least some embodiments. As indicated at 400, the device sends data to the base station over the wireless connection. As indicated at 410, the base station renders frames including virtual content based at least in part on the data received from the device. As indicated at 420, the base station compresses the rendered data and sends the compressed data to the device over the wireless connection. As indicated at 430, the device decompresses and displays the virtual content to generate a 3D virtual view for viewing by the user. As indicated by the arrow returning from 430 to 400, the method continues as long as the user is using the VR/MR system.

[0070] In some embodiments, rather than rendering and encoding entire frames in the base station and transmitting the rendered frames to the device, the base station may render and encode parts of frames (referred to as slices) and transmit the encoded slices to the device as they are ready. A slice may, for example, be a row of 128x128 blocks. In some embodiments, the video encoding system may perform tile-based rendering, encoding, and transmittal. In tile-based rendering, encoding, and transmittal, each slice may be divided into multiple tiles each including one or more blocks (e.g., four tiles, each including four blocks), and the base station may render and encode the tiles and transmit the encoded tiles to the device as they are ready.

[0071] FIG. 5 is a flowchart of a method of operation for a video encoding system as illustrated in FIG. 1, according to at least some embodiments. The method of FIG. 5 may, for example, be performed at 420 of FIG. 4. The method of FIG. 5 assumes slice-based encoding and transmission is being performed. However, in some embodiments, tile-based encoding and transmission may be performed.

[0072] As indicated at 510, the pre-filter component applies lens warp and/or foveation filters to pixel blocks in a slice of the frame. In some embodiments, pre-filtering may include performing a lens warp on the frames on the base station prior to the wavelet transform. The lens warp is performed to correct for the distortion of the images introduced by the lenses on the device that the images are viewed through, thus improving quality of the images. In some embodiments, the device may store lens warp data for the lenses, for example generated by a calibration process, and may provide the lens warp data to the base station over the wireless connection. The base station may then perform the lens warp on the frames based on the lens warp data for that device. Performing the lens warp on the base station in the pre-filter stage may reduce the resolution of the frames prior to performing the wavelet transform and encoding, which may help in improving compression, and may reduce latency and bandwidth usage on the wireless link. In addition, by performing the lens warp on the base station in the pre-filter stage rather than on the device after decoding, filtering of the image data may only need to be performed once, as opposed to performing filtering on the base station to reduce resolution prior to encoding and then performing lens warp filtering on the device.

[0073] In some embodiments, pre-filtering at 510 may also include filtering to reduce resolution in peripheral regions while maintaining higher resolution in foveated regions. In some embodiments, gaze tracking information obtained from the device may be used to identify the direction in which the user is currently looking. A region of the frame that corresponds to the fovea (referred to as the foveated region) may be identified based at least in part on the determined gaze direction. The peripheral region (i.e. the portion of the frame outside the foveated region) may be pre-filtered to reduce information based on knowledge of the human vision system, for

example by filtering high frequency information and/or increasing color compression. Pre-filtering of the peripheral region may help to provide improved compression of the frame.

[0074] As indicated at 520, the wavelet transform component applies a wavelet transform technique to the pixel blocks to decompose the pixel data into N (e.g., 7) frequency bands. The frequency bands are then organized into blocks that are provided to a block-based encoder for encoding/compression. As an example, a frame may be divided into 128x128 blocks, and a two-level wavelet decomposition may be applied to each 128x128 block to generate 16 32x32 blocks of frequency data representing seven frequency bands, for example as illustrated in FIG. 3C.

[0075] As indicated at 530, the encoder applies an encoding technique to the frequency bands in the blocks to compress the data. The encoder may, for example, be a High Efficiency Video Coding (HEVC) encoder. However, other encoding techniques may be used in some embodiments. Decomposing the pixel data into frequency bands as indicated at element 520 allows the frequency bands to be buffered and processed as separate streams by the encoder. Processing the frequency bands as separate streams allows the encoder component to multiplex the processing of the independent streams. In block-based encoding methods such as HEVC encoding, blocks (referred to as coding tree units (CTUs)) are processed in a pipeline at multiple stages; two or more blocks may be at different stages of the pipeline at a given clock cycle, and the blocks move through the pipeline as the clock cycles. The processing of a given block may have dependencies on one or more previously processed neighbor blocks, for example one or more blocks in the row above the given block and/or the block to the left of the given block. By multiplexing the processing of the streams, the encoder spaces out the processing of the blocks in a given stream, thus providing additional clock cycles to process a neighbor block on which a given block has dependencies. For example, the block to the left of the given block may be several stages ahead of the given block in the pipeline when the given block reaches a stage that depends on the previously processed neighbor block. This allows the encoder to better handle dependencies on previously processed blocks, and reduces or eliminates the need to wait for completion of processing of a neighbor block in the pipeline before processing the given block at a stage that depends on the neighbor block.

[0076] As indicated at 540, the wireless interface packetizes the compressed data and sends the packets to the device over the wireless connection.

[0077] Decomposing the pixel data into frequency bands as indicated at element 520 allows the frequency bands to be prioritized by the encoder at element 530 and the wireless interface at element 540. Typically, in image and video transmission, the lower frequencies are more important, while the higher frequencies are less important. Higher frequencies usually correspond to details in the image, and thus can be considered as lower priority. The higher frequency bands

contain a smaller percentage of the energy in the image. Most of the energy is contained in the lower frequency bands. Decomposing the pixel data into frequency bands thus provides a priority ordering to the data stream that can be leveraged by the encoder and the wireless interface when encoding and transmitting the data stream. For example, in some embodiments, different compression techniques may be used on the different frequency bands, with more aggressive compression applied to the lower priority bands, and more conservative compression applied to the higher priority bands. As another example, the priority ordering of the frequency bands may help in providing graceful degradation of the VR/MR system. Performance of the wireless connection can be monitored, and feedback from the device (e.g., from an HMD) may be considered, to track performance of the overall system. If the system is falling behind for some reason, for example if the wireless connection degrades and bandwidth capacity of the wireless connection drops below a threshold, the encoder and wireless interface may prioritize the encoding and transmission of one or more of the lower frequency bands, and may reduce or drop the encoding and/or transmission of one or more of the frequency levels that have been assigned a lower priority level, for example one or more of the higher frequency bands.

[0078] At 550, if there are more slices to be encoded and transmitted, the method returns to element 510 to process the next slice. Otherwise, at 560, if there are more frames to be encoded and transmitted, the method returns to element 510 to begin processing the next frame.

[0079] FIG. 6 is a flowchart of a method of operation for a video encoding system as illustrated in FIG. 2, according to at least some embodiments. The method of FIG. 6 may, for example, be performed at 420 of FIG. 4. In the method of FIG. 6, the video encoding system may perform tile-based encoding using multiple encoders to process respective tiles from each slice.

[0080] As indicated at 600, a rendering engine renders a slice including multiple tiles (four tiles, in this example), each tiles including multiple pixel blocks (four 128x128 pixel blocks, in this example).

[0081] As indicated at 610, the pre-filter component applies lens warp and/or foveation filters to the slice. In some embodiments, pre-filtering may include performing a lens warp on the frames on the base station prior to the wavelet transform. The lens warp is performed to correct for the distortion of the images introduced by the lenses on the device that the images are viewed through, thus improving quality of the images. In some embodiments, the device may store lens warp data for the lenses, for example generated by a calibration process, and may provide the lens warp data to the base station over the wireless connection. The base station may then perform the lens warp on the frames based on the lens warp data for that device. Performing the lens warp on the base station in the pre-filter stage may reduce the resolution of the frames prior to performing the wavelet transform and encoding, which may help in improving compression, and may reduce

latency and bandwidth usage on the wireless link. In addition, by performing the lens warp on the base station in the pre-filter stage rather than on the device after decoding, filtering of the image data may only need to be performed once, as opposed to performing filtering on the base station to reduce resolution prior to encoding and then performing lens warp filtering on the device.

5 **[0082]** In some embodiments, pre-filtering at 610 may also include filtering to reduce resolution in peripheral regions while maintaining higher resolution in foveated regions. In some embodiments, gaze tracking information obtained by the device may be used to identify the direction in which the user is currently looking. A region of the frame that corresponds to the fovea (referred to as the foveated region) may be identified based at least in part on the determined
10 gaze direction. The peripheral region (i.e. the portion of the frame outside the foveated region) may be pre-filtered to reduce information based on knowledge of the human vision system, for example by filtering high frequency information and/or increasing color compression. Pre-filtering of the peripheral region may help to provide improved compression of the frame.

[0083] In some embodiments, the video encoding system may include a single pre-filter
15 component that process the tiles. In some embodiments, the video encoding system may include a separate pre-filter component for processing each tile. In some embodiments, the video encoding system may include multiple (e.g., 2) pre-filter components that each process multiple (e.g., 2) tiles.

[0084] As indicated at 620, the wavelet transform component applies a wavelet transform
20 technique to the pixel blocks in the slice to decompose the pixel data into N (e.g., 7) frequency bands. The frequency bands are then organized into blocks (e.g., CTUs) that can be provided to a block-based encoder for encoding/compression. As an example, a frame may be divided into 128x128 blocks, and a two-level wavelet decomposition may be applied to each 128x128 block to generate 16 32x32 blocks of frequency data representing seven frequency bands, for example as
25 illustrated in FIG. 3C.

[0085] In some embodiments, the video encoding system may include a single wavelet
transform component that process the tiles. In some embodiments, the video encoding system may include a separate wavelet transform component for processing each tile. In some embodiments, the video encoding system may include multiple (e.g., 2) wavelet transform components that each
30 process multiple (e.g., 2) tiles.

[0086] In the method of FIG. 6, an example video encoding system includes two encoders
configured to encode the blocks of frequency data from different ones of the slices that are generated at element 620 in parallel. For example, a first encoder may be configured to encode blocks from slices 0 and 1, and a second encoder may be configured to encode blocks from slices
35 2 and 3. As indicated at 630A, the first encoder applies an encoding technique to the frequency

bands in the blocks of tiles 0 and 1, multiplexing the processing of blocks from different frequency bands between the two different tiles. As indicated at 630B, a second encoder applies an encoding technique to the frequency bands in the blocks of tiles 2 and 3, multiplexing the processing of blocks from different frequency bands between the two different tiles.

5 **[0087]** The encoders may, for example, be a High Efficiency Video Coding (HEVC) encoders. However, other encoding techniques may be used in some embodiments. Decomposing the pixel data into frequency bands as indicated at element 620 allows the frequency bands to be buffered and processed as separate streams by the encoders at elements 630A and 630B. Processing the frequency bands as separate streams allows the encoders to multiplex the processing of the independent streams. In block-based encoding methods such as HEVC encoding, blocks (referred to as coding tree units (CTUs)) are processed in a pipeline at multiple stages; two or more blocks may be at different stages of the pipeline at a given clock cycle, and the blocks move through the pipeline as the clock cycles. The processing of a given block may have dependencies on one or more previously processed neighbor blocks, for example one or more blocks in the row above the given block and/or the block to the left of the given block. By multiplexing the processing of the streams, the encoder spaces out the processing of the blocks in a given stream, thus providing additional clock cycles to process a neighbor block on which a given block has dependencies. For example, the block to the left of the given block may be several stages ahead of the given block in the pipeline when the given block reaches a stage that depends on the previously processed neighbor block. This allows the encoder to better handle dependencies on previously processed blocks, and reduces or eliminates the need to wait for completion of processing of a neighbor block in the pipeline before processing the given block at a stage that depends on the neighbor block.

15 **[0088]** As indicated at 640, the wireless interface packetizes the compressed data generated by the encoders at element 530 and sends the packets to the device over the wireless connection. As indicated by the arrow returning from 650 to 600, the method continues as long as the user is using the VR/MR system.

20 **[0089]** Decomposing the pixel data into frequency bands as indicated at element 620 allows the frequency bands to be prioritized by the encoders at elements 630A and 630B and the wireless interface at element 640. Typically, in image and video transmission, the lower frequencies are more important, while the higher frequencies are less important. Higher frequencies usually correspond to details in the image, and thus can be considered as lower priority. The higher frequency bands contain a smaller percentage of the energy in the image. Most of the energy is contained in the lower frequency bands. Decomposing the pixel data into frequency bands thus provides a priority ordering to the data stream that can be leveraged by the encoder and the wireless interface when encoding and transmitting the data stream. For example, in some embodiments,

35

different compression techniques may be used on the different frequency bands, with more aggressive compression applied to the lower priority bands, and more conservative compression applied to the higher priority bands. As another example, the priority ordering of the frequency bands may help in providing graceful degradation of the VR/MR system. Performance of the wireless connection can be monitored, and feedback from the device (e.g., from an HMD) may be considered, to track performance of the overall system. If the system is falling behind for some reason, for example if the wireless connection degrades and bandwidth capacity of the wireless connection drops below a threshold, the encoder and wireless interface may prioritize the encoding and transmission of one or more of the lower frequency bands, and may reduce or drop the encoding and/or transmission of one or more of the frequency levels that have been assigned a lower priority level, for example one or more of the higher frequency bands.

[0090] While embodiments are described in which each slice is divided into four tiles and two encoders operate on respective tiles from each slice, slices may be divided into more tiles (e.g., six or eight tiles) in some embodiments, and more encoders (e.g., three or four encoders) may be used in some embodiments.

Prioritized frequency band transmission

[0091] FIGS. 1 through 6 illustrate a video encoding system in which a pre-filter is performed on pixel data, followed by a wavelet transform to decompose the pre-filtered pixel data into frequency bands, followed by encoding by a block-based encoder that interfaces with a wireless interface that transmits the encoded frequency data to a receiving device (e.g., an HMD) over a low latency wireless connection. In some embodiments, the encoder and wireless interface components of the video encoding system may leverage the decomposition of the data into frequency bands to optimize the compression and transmission of the data to, for example, provide graceful degradation over the wireless connection.

[0092] Decomposing the pixel data into frequency bands as illustrated in FIG. 3C allows the frequency bands to be prioritized by the encoder and the wireless interface. Typically, in image and video transmission, the lower frequencies are more important, while the higher frequencies are less important. Higher frequencies usually correspond to details in the image, and thus can be considered as lower priority. The higher frequency bands contain a smaller percentage of the energy in the image. Most of the energy is contained in the lower frequency bands. Decomposing the pixel data into frequency bands thus provides a priority ordering to the data stream that can be leveraged by the encoder and the wireless interface when encoding and transmitting the data stream.

[0093] The priority ordering of the frequency bands may, for example, be leveraged to provide graceful degradation of the wireless connection in the VR/MR system. Performance of the wireless connection can be monitored, and feedback from the device (e.g., an HMD) may be considered, to track performance of the wireless connection. If the system is falling behind for some reason, for example if the wireless connection degrades and bandwidth capacity of the wireless connection drops below a threshold, the encoder and wireless interface may prioritize the encoding and/or transmission of one or more of the lower frequency bands, and may reduce or drop the encoding and/or transmission of one or more of the frequency levels that have been assigned a lower priority level, for example one or more of the higher frequency bands.

[0094] In some embodiments, the pixel data is first sent through a pre-filter, for example an N-channel filter bank. In some embodiments, the pre-filter may reduce high frequency content in the areas of the image that are not in a foveated region so that high fidelity can be maintained in the foveated area and lower fidelity, which cannot be detected by the human visual system, is applied in the non-foveated (peripheral) region. The output of the pre-filter is then sent through a two-level wavelet decomposition to produce multiple (e.g., seven) frequency bands, for example as frequency blocks as illustrated in FIG. 3C.

[0095] The frequency blocks for the frequency bands are then compressed by the encoder. The encoder may categorize the encoded frequency bands into multiple (e.g., five) priority levels. The encoder may mark each frequency block with metadata indicating the frame/slice/tile/block of the frequency block, the frequency band represented in the block, the priority of the frequency band, and timing information. The timing information may, for example, include a time at which the frequency block was generated by the encoder, and a time (referred to as a “deadline”) at which the frequency block should be received at the device (e.g., an HMD).

[0096] The wireless interface may packetize the encoded frequency blocks for transmission over the wireless connection. One or more encoded frequency blocks may be included in each packet. In some embodiments, the encoded frequency blocks are packetized according to priority levels. The wireless interface of the base station may then transmit or drop packets according to the priority levels of the encoded frequency blocks in the packets and/or according to the deadlines of the frequency blocks in the packets.

[0097] In some embodiments, the wireless interface may implement different policies for dropping packets based on the priorities and/or deadlines. As an example, in some embodiments, the wireless interface may drop packets strictly on a priority basis. As another example, in some embodiments the wireless interface can inspect the deadlines of each packet and drop packets according to both deadline and priorities. As yet another example, in some embodiments the

wireless interface may decide to drop all lower priority packets once a packet of a slice with a specific priority is dropped, as the other packets with lower priority are likely not to be successfully transmitted, hence freeing up resources for transmission.

[0098] In some embodiments, to achieve a hard latency requirement, the timing information for each packet may include a timestamp indicating when the frequency blocks in the packet were encoded by the encoder and/or a deadline time at which the packet must be received at the device. If the deadline is missed in the transmission from the wireless interface of the base station to the wireless interface of the device, then the packet may be dropped at the base station or at the device.

[0099] In some embodiments, in addition to dividing the frequency bands into priority levels, the frequency bands may also be marked with different priorities based on whether the corresponding pixels are in the foveated region or in the peripheral region of the frame. As an example, referring to FIG. 3C, the LLLL frequency band may be typically marked as priority 1, but in the peripheral region the LLLL frequency band may instead be marked with priority 2 or 3. As another example, referring to FIG. 3C, the HH frequency band may be typically marked as priority 5, but in the foveated region the HH frequency band may be marked with priority 4 or 3.

[0100] In some embodiments, the wireless interface of the base station may optimize the transmission of packets by maximizing the minimum number of priority levels that are successfully transmitted. This may be accomplished by assigning different modulation coding schemes to each of the different priority levels depending upon channel bandwidth availability. In some embodiments, the encoder may signal to the wireless interface a preferred packet error rate that corresponds to a specific modulation coding scheme to maximize quality.

[0101] In some embodiments, the wireless interface of the base station may optimize transmission of the packets by transmitting packets of a lower priority from the end of a tile or slice if the deadline of the packets is closer to expiration than higher priority packets belonging to a next tile or slice.

[0102] Embodiments are generally described as applying a wavelet transform to divide an image into different frequency bands that can then be prioritized. However, in some embodiments, other methods may be used to divide an image into frequency bands. For example, in some embodiments, a Fourier transform may be applied that decomposes an image into different frequency bands. The frequency bands may be divided into different priorities after the Fourier transform, and the priority-based video encoding methods may be applied to the prioritized frequency bands as described herein.

[0103] FIG. 7 is a block diagram illustrating a video encoding system as illustrated in FIGS. 1 or 2 in which frequency bands are prioritized for transmission over a wireless connection, according to at least some embodiments. A wavelet transform 706 component performs two-level

wavelet decomposition on pre-filtered pixel blocks (e.g., 128x128 blocks from a tile or slice) to decompose each pixel block into frequency blocks representing frequency bands (sixteen 32x32 frequency blocks representing seven frequency bands B0-B6, in this example). A block-based encoder 708 component may then multiplex the encoding of the frequency blocks (e.g., 32x32 CTUs) corresponding to the seven frequency bands. The encoder 708 may output the compressed blocks categorized into five priority levels P1-P5, from highest priority to lowest priority. In this example, band 0 (LLLL) is output as P1 (highest priority), band 1 (LLLH) and band 2 (LLHL) are output as P2, band 3 (LLHH) are output as P3, band 4 (LH) and band 5 (HL) are output as P4, and band 6 (HH) is output as P5 (lowest priority).

10 **[0104]** FIG. 8A illustrates an example compressed block, according to at least some embodiments. In some embodiments, the encoder 708 may tag each compressed block with metadata indicating the frame/slice/tile/block of the block, the frequency band of the compressed data in the block (e.g., B0-B6), the priority of the frequency band (e.g., P1-P5), and timing information. The timing information may, for example, include a time at which the compressed block was generated by the encoder 708, and a time (referred to as a “deadline”) at which the compressed block should be received at the device.

15 **[0105]** The wireless interface 710 may packetize the encoded frequency blocks for transmission over the wireless connection to the device. FIG. 8B illustrates an example packet, according to at least some embodiments. The packet may include a packet header that may, for example, include information about the packet data payload. One or more compressed blocks may be included in each packet. In some embodiments, the compressed blocks may be packetized according to the priority levels, from highest priority (P1) to lowest priority (P5). The wireless interface 710 of the base station may then transmit or drop packets according to the priority levels of the encoded frequency blocks in the packets and/or according to the deadlines of the encoded frequency blocks in the packets. In some embodiments, the wireless interface may implement different policies for dropping packets based on the priorities and/or deadlines. As an example, in some embodiments, the wireless interface may drop packets strictly on a priority basis. As another example, in some embodiments the wireless interface can inspect the deadlines of each packet and drop packets according to both deadline and priorities. As yet another example, in some

20
25
30

35 **[0106]** In some embodiments, to achieve a hard latency requirement, the timing information for each packet may include a timestamp indicating when the frequency blocks in the packet were encoded by the encoder 708 and/or a deadline time at which the packet must be received at the

device. If the deadline is missed in the transmission from the wireless interface 710 of the base station to the wireless interface of the device, then the packet may be dropped at the wireless interface of the base station or at the device.

[0107] In some embodiments, in addition to categorizing the frequency bands into priority levels as illustrated in FIG. 7, the frequency bands may also be categorized into different priority levels based on whether the corresponding pixels are in the foveated region or in the peripheral region of the frame. As an example, the LLLL frequency band (band 0) may be typically marked as priority 1 (P1), but in the peripheral region the LLLL frequency band may instead be categorized into P2 or P3. As another example, the HH frequency band (band 6) may be typically marked as priority 5 (P5), but in the foveated region the HH frequency band may instead be categorized into P4 or P3.

[0108] FIG. 9 is a flowchart of a method of operation for a video encoding system as illustrated in FIG. 7, according to at least some embodiments. As indicated at 900, a wavelet transform component of the video encoding system decomposes the pixel data into N (e.g., 7) frequency bands. For example, in some embodiments, the pixel data may be decomposed into seven frequency bands by a two-level wavelet decomposition as illustrated in FIG. 3C.

[0109] As indicated at 910, an encoder component of the video encoding system applies an encoding technique to the frequency bands to compress the data. As indicated at 920, the encoder tags the compressed frequency band data with priority and timing information and provides the tagged data to the wireless interface of the base station. In some embodiments, the encoder may categorize the encoded frequency bands into multiple (e.g., five) priority levels P1-P5, from highest priority to lowest priority, as illustrated in FIG. 7. In some embodiments, in addition to categorizing the frequency bands into priority levels as illustrated in FIG. 7, the frequency bands may also be categorized into different priority levels based on whether the corresponding pixels are in the foveated region or in the peripheral region of the frame. In some embodiments, the encoder may tag each compressed block with metadata indicating the frame/slice/tile/block of the block, the frequency band of the compressed data in the block (e.g., B0-B6), the priority of the frequency band (e.g., P1-P5), and timing information as illustrated in FIG. 8A. The timing information may, for example, include a time at which the compressed block was generated by the encoder, and a time (referred to as a “deadline”) at which the compressed block should be received at the device.

[0110] As indicated at 930, the wireless interface of the base station packetizes and sends the compressed data to the device over the wireless connection based on the priority and timing information. In some embodiments, the compressed blocks may be packetized according to the priority levels, from highest priority (e.g., P1) to lowest priority (e.g., P5). The wireless interface

of the base station may then transmit or drop packets according to the priority levels of the encoded frequency blocks in the packets. In some embodiments, the wireless interface may implement different policies for dropping packets based on the priorities and/or deadlines. As an example, in some embodiments, the wireless interface may drop packets strictly on a priority basis. As another
5 example, in some embodiments the wireless interface can inspect the deadlines of each packet and drop packets according to both deadline and priorities. As yet another example, in some embodiments the wireless interface may decide to drop all lower priority packets once a packet of a slice with a specific priority is dropped, as the other packets with lower priority are likely not to be successfully transmitted, hence freeing up resources for transmission.

10 **[0111]** In some embodiments, the priority ordering of the frequency bands may help in providing graceful degradation of the VR/MR system. Performance of the wireless connection can be monitored, and feedback from the device may be considered, to track performance of the overall system. If the system is falling behind for some reason, for example if the wireless connection degrades and bandwidth capacity of the wireless connection drops below a threshold, the wireless
15 interface may prioritize the transmission of one or more of the lower frequency, higher priority bands (e.g., P1 through P3, as illustrated in FIG. 7), and reduce, delay, or drop the transmission of one or more of the higher frequency, lower priority bands (e.g., P4 and P5, as illustrated in FIG. 7). In some embodiments, the wireless interface may implement different policies for dropping packets based on the priorities and/or deadlines. As an example, in some embodiments, the
20 wireless interface may drop packets strictly on a priority basis. As another example, in some embodiments the wireless interface can inspect the deadlines of each packet and drop packets according to both deadline and priorities. As yet another example, in some embodiments the wireless interface may decide to drop all lower priority packets once a packet of a slice with a specific priority is dropped, as the other packets with lower priority are likely not to be successfully
25 transmitted, hence freeing up resources for transmission.

[0112] In some embodiments, to achieve a hard latency requirement, the timing information for each packet may include a timestamp indicating when the frequency blocks in the packet were encoded by the encoder and/or a deadline time at which the packet must be received at the device. If the deadline is missed in the transmission from the wireless interface of the base station to the
30 wireless interface of the device, then the packet may be dropped at the wireless interface of the base station or at the device.

[0113] As indicated by the arrow returning from element 940 to element 900, the method may continue as long as there is data to be transmitted to the device.

Example VR/MR system

[0114] FIG. 10 illustrates an example VR/MR system 2000 that may implement a video encoding system, according to at least some embodiments. A VR/MR system 2000 may include at least one device 2150 (e.g., a notebook or laptop computer, pad or tablet device, smartphone, hand-held computing device or an HMD such as a headset, helmet, goggles, or glasses that may be worn by a user) and a computing device 2100 (referred to herein as a base station). The base station 2100 renders VR or MR frames including virtual content, encodes the frames, and transmits the encoded frames over a wireless connection 2180 to the device 2150 for decoding and display by the device 2150.

[0115] The base station 2100 and device 2150 may each include wireless communications technology that allows the base station 2100 and device 2150 to communicate and exchange data via the wireless connection 2180. In some embodiments, the wireless connection 2180 may be implemented according to a proprietary wireless communications technology that provides a highly directional wireless link between the device 2150 and the base station 2100. However, other commercial (e.g., Wi-Fi, Bluetooth, etc.) or proprietary wireless communications technologies may be used in some embodiments.

[0116] In some embodiments, the device 2150 may include sensors that collect information about the user's environment (e.g., video, depth information, lighting information, etc.) and/or about the user (e.g., the user's expressions, eye movement, gaze direction, hand gestures, etc.). The device 2150 may transmit at least some of the information collected by sensors to the base station 2100 via wireless connection 2180. The base station 2100 may render frames for display by the device 2150 that include virtual content based at least in part on the various information obtained from the sensors, encode the frames, and transmit the encoded frames to the device 2150 for decoding and display to the user via the wireless connection 2180. To encode and transmit the frames, the base station 2100 may implement a video encoding system as illustrated in FIGS. 1 through 9.

[0117] FIG. 11 is a block diagram illustrating functional components of and processing in an example VR/MR system as illustrated in FIG. 10, according to some embodiments. Device 2150 may be, but is not limited to, a notebook or laptop computer, pad or tablet device, smartphone, hand-held computing device or an HMD such as a headset, helmet, goggles, or glasses that may be worn by a user. Device 2150 may include a display 2156 component or subsystem that may implement any of various types of virtual or augmented reality display technologies. For example, an HMD device 2150 may include a near-eye system that displays left and right images on screens in front of the user's eyes that are viewed by a subject, such as DLP (digital light processing), LCD

(liquid crystal display) and LCoS (liquid crystal on silicon) technology VR systems. As another example, an HMD device 2150 may include a direct retinal projector system that scans left and right images, pixel by pixel, to the subject's eyes. To scan the images, left and right projectors generate beams that are directed to left and right reflective components (e.g., ellipsoid mirrors) located in front of the user's eyes; the reflective components reflect the beams to the user's eyes. To create a three-dimensional (3D) effect, virtual content at different depths or distances in the 3D virtual view are shifted left or right in the two images as a function of the triangulation of distance, with nearer objects shifted more than more distant objects.

[0118] Device 2150 may also include a controller 2154 configured to implement device-side functionality of the VR/MR system 2000 as described herein. In some embodiments, device 2150 may also include memory 2170 configured to store software (code 2172) of the device 2150 component of the VR/MR system 2000 that is executable by the controller 2154, as well as data 2174 that may be used by the software when executing on the controller 2154. In various embodiments, the controller 2154 may be a uniprocessor system including one processor, or a multiprocessor system including several processors (e.g., two, four, eight, or another suitable number). The controller 2154 may include central processing units (CPUs) configured to implement any suitable instruction set architecture, and may be configured to execute instructions defined in that instruction set architecture. For example, in various embodiments the controller 2154 may include general-purpose or embedded processors implementing any of a variety of instruction set architectures (ISAs), such as the x86, PowerPC, SPARC, RISC, or MIPS ISAs, or any other suitable ISA. In multiprocessor systems, each of the processors may commonly, but not necessarily, implement the same ISA. The controller 2154 may employ any microarchitecture, including scalar, superscalar, pipelined, superpipelined, out of order, in order, speculative, non-speculative, etc., or combinations thereof. The controller 2154 may include circuitry to implement microcoding techniques. The controller 2154 may include one or more processing cores each configured to execute instructions. The controller 2154 may include one or more levels of caches, which may employ any size and any configuration (set associative, direct mapped, etc.). In some embodiments, the controller 2154 may include at least one graphics processing unit (GPU), which may include any suitable graphics processing circuitry. Generally, a GPU may be configured to render objects to be displayed into a frame buffer (e.g., one that includes pixel data for an entire frame). A GPU may include one or more graphics processors that may execute graphics software to perform a part or all of the graphics operation, or hardware acceleration of certain graphics operations. In some embodiments, the controller 2154 may include one or more other components for processing and rendering video and/or images, for example image signal processors (ISPs),

encoder/decoders (codecs), etc. In some embodiments, controller 2154 may include at least one system on a chip (SOC).

[0119] The memory 2170 may include any type of memory, such as dynamic random access memory (DRAM), synchronous DRAM (SDRAM), double data rate (DDR, DDR2, DDR3, etc.) SDRAM (including mobile versions of the SDRAMs such as mDDR3, etc., or low power versions of the SDRAMs such as LPDDR2, etc.), RAMBUS DRAM (RDRAM), static RAM (SRAM), etc. In some embodiments, one or more memory devices may be coupled onto a circuit board to form memory modules such as single inline memory modules (SIMMs), dual inline memory modules (DIMMs), etc. Alternatively, the memory devices may be mounted with an integrated circuit implementing system in a chip-on-chip configuration, a package-on-package configuration, or a multi-chip module configuration.

[0120] In some embodiments, sensors 2160 may include, but are not limited to, one or more gaze tracking sensors (e.g., IR cameras with an IR illumination source) that may be used to track position and movement of the user's eyes. In some embodiments, there may be two gaze tracking sensors, with each gaze tracking sensor tracking a respective eye. In some embodiments, the information collected by the gaze tracking sensors may be used to adjust the rendering of images by the base station 2100, and/or to adjust the projection of the images by the projection system of the device 2150, based on the direction and angle at which the user's eyes are looking. For example, in some embodiments, content of the images in a region around the location at which the user's eyes are currently looking may be rendered with more detail and at a higher resolution than content in regions at which the user is not looking, which allows available processing time for image data to be spent on content viewed by the foveal regions of the eyes rather than on content viewed by the peripheral regions of the eyes. Similarly, content of images in regions at which the user is not looking may be compressed more than content of the region around the point at which the user is currently looking. In some embodiments there may be two gaze tracking sensors located on an inner surface of the device 2150 at positions such that the sensors have views of respective ones of the user's eyes. However, in various embodiments, more or fewer gaze tracking sensors may be used, and gaze tracking sensors may be positioned at other locations. In an example non-limiting embodiment, each gaze tracking sensor may include an IR light source and IR camera, for example a 400x400 pixel count camera with a frame rate of 120 FPS or greater, HFOV of 70 degrees, and with a working distance of 10 millimeters (mm) to 80mm.

[0121] In some embodiments, the device 2150 may include at least one inertial-measurement unit (IMU) 2162 configured to detect position, orientation, and/or motion of the device 2150, and to provide the detected position, orientation, and/or motion data to the controller 2154 of the device 2150 and/or to the base station 2100.

[0122] Device 2150 may also include a wireless interface 2152 configured to communicate with an external base station 2100 via a wireless connection 2180 to send sensor inputs to the base station 2100 and to receive compressed rendered frames, slices, or tiles from the base station 2100. In some embodiments, the wireless interface 2152 may implement a proprietary wireless communications technology that provides a highly directional wireless link between the device 2150 and the base station 2100. However, other commercial (e.g., Wi-Fi, Bluetooth, etc.) or proprietary wireless communications technologies may be used in some embodiments.

[0123] The base station 2100 may be an external device (e.g., a computing system, game console, etc.) that is communicatively coupled to device 2150 via a wireless interface 2180. The base station 2100 may include one or more of various types of processors (e.g., SOCs, CPUs, ISPs, GPUs, codecs, and/or other components) for rendering, filtering, encoding, and transmitting video and/or images. The base station 2100 may render frames (each frame including a left and right image) that include virtual content based at least in part on the various inputs obtained from the sensors 2160 via the wireless connection 2180, filter and compress the rendered frames (or slices of the frames) using a video encoding system as described herein, and transmit the compressed frames or slices to the device 2150 for display.

[0124] Base station 2100 may be or may include any type of computing system or computing device, such as a desktop computer, notebook or laptop computer, pad or tablet device, smartphone, hand-held computing device, game controller, game system, and so on. Base station 2100 may include a controller 2110 comprising one or more processors that implement base-side functionality of the VR/MR system 2000 including the video encoding system as described herein. Base station 2100 may also include memory 2120 configured to store software (code 2122) of the base station component of the VR/MR system 2000 that is executable by the base station controller 2110, as well as data 2124 that may be used by the software when executing on the controller 2110.

[0125] In various embodiments, the base station controller 2110 may be a uniprocessor system including one processor, or a multiprocessor system including several processors (e.g., two, four, eight, or another suitable number). The controller 2110 may include central processing units (CPUs) configured to implement any suitable instruction set architecture and may be configured to execute instructions defined in that instruction set architecture. For example, in various embodiments the controller 2110 may include general-purpose or embedded processors implementing any of a variety of instruction set architectures (ISAs), such as the x86, PowerPC, SPARC, RISC, or MIPS ISAs, or any other suitable ISA. In multiprocessor systems, each of the processors may commonly, but not necessarily, implement the same ISA. The controller 2110 may employ any microarchitecture, including scalar, superscalar, pipelined, superpipelined, out of

order, in order, speculative, non-speculative, etc., or combinations thereof. Controller 2110 may include circuitry to implement microcoding techniques. The controller 2110 may include one or more processing cores each configured to execute instructions. The controller 2110 may include one or more levels of caches, which may employ any size and any configuration (set associative, direct mapped, etc.). In some embodiments, the controller 2110 may include at least one graphics processing unit (GPU), which may include any suitable graphics processing circuitry. Generally, a GPU may be configured to render objects to be displayed into a frame buffer (e.g., one that includes pixel data for an entire frame). A GPU may include one or more graphics processors that may execute graphics software to perform a part or all of the graphics operation, or hardware acceleration of certain graphics operations. In some embodiments, the controller 2110 may include one or more other components for processing, rendering, filtering, and encoding video and/or images as described herein, for example one or more of various types of integrated circuits (ICs), image signal processors (ISPs), encoder/decoders (codecs), etc. In some embodiments, the controller 2110 may include at least one system on a chip (SOC).

[0126] The base station memory 2120 may include any type of memory, such as dynamic random access memory (DRAM), synchronous DRAM (SDRAM), double data rate (DDR, DDR2, DDR3, etc.) SDRAM (including mobile versions of the SDRAMs such as mDDR3, etc., or low power versions of the SDRAMs such as LPDDR2, etc.), RAMBUS DRAM (RDRAM), static RAM (SRAM), etc. In some embodiments, one or more memory devices may be coupled onto a circuit board to form memory modules such as single inline memory modules (SIMMs), dual inline memory modules (DIMMs), etc. Alternatively, the devices may be mounted with an integrated circuit implementing system in a chip-on-chip configuration, a package-on-package configuration, or a multi-chip module configuration.

[0127] Base station 2100 may also include one or more wireless technology interfaces 2130 configured to communicate with device 2150 via a wireless connection 2180 to receive sensor inputs from the device 2150 and send compressed frames, slices, or tiles from the base station 2100 to the device 2150. In some embodiments, a wireless technology interface 2130 may implement a proprietary wireless communications technology that provides a highly directional wireless link between the device 2150 and the base station 2100. In some embodiments, the directionality and band width of the wireless communication technology may support multiple devices 2150 communicating with the base station 2100 at the same time to thus enable multiple users to use the system 2000 at the same time in a co-located environment. However, other commercial (e.g., Wi-Fi, Bluetooth, etc.) or proprietary wireless communications technologies may be used in some embodiments.

[0128] In some embodiments, the base station 2100 may be configured to render and transmit frames to the device 2150 to provide a 3D virtual view for the user based at least in part on sensor 2160 inputs received from the device 2150. In some embodiments, the virtual view may include renderings of the user's environment, including renderings of real objects in the user's environment, based on video captured by one or more scene cameras (e.g., RGB (visible light) video cameras) that capture high-quality, high-resolution video of the user's environment in real time for display. In some embodiments, the virtual view may also include virtual content (e.g., virtual objects, virtual tags for real objects, avatars of the user, etc.) rendered and composited with the projected 3D view of the user's real environment by the base station 2100.

[0129] While not shown in FIGS. 10 and 11, in some embodiments the VR/MR system 2000 may include one or more other components. For example, the system may include a cursor control device (e.g., mouse) for moving a virtual cursor in the 3D virtual view to interact with virtual content. While FIGS. 10 and 11 show a single device 2150, in some embodiments the VR/MR system 2000 may support multiple devices 2150 communicating with the base station 2100 at the same time to thus enable multiple users to use the system at the same time in a co-located environment.

[0130] The methods described herein may be implemented in software, hardware, or a combination thereof, in different embodiments. In addition, the order of the blocks of the methods may be changed, and various elements may be added, reordered, combined, omitted, modified, etc. Various modifications and changes may be made as would be obvious to a person skilled in the art having the benefit of this disclosure. The various embodiments described herein are meant to be illustrative and not limiting. Many variations, modifications, additions, and improvements are possible. Accordingly, plural instances may be provided for components described herein as a single instance. Boundaries between various components, operations and data stores are somewhat arbitrary, and particular operations are illustrated in the context of specific illustrative configurations. Other allocations of functionality are envisioned and may fall within the scope of claims that follow. Finally, structures and functionality presented as discrete components in the example configurations may be implemented as a combined structure or component. These and other variations, modifications, additions, and improvements may fall within the scope of embodiments as defined in the claims that follow.

CLAIMS**WHAT IS CLAIMED IS:**

- 5 1. A video encoding system, comprising:
- one or more processors configured to:
- decompose pixel blocks from a video frame into a plurality of frequency bands; and
- organize the frequency bands from the pixel blocks into a plurality of frequency
- 10 blocks each including frequency data for one of the frequency bands;
- an encoder configured to:
- apply an encoding technique to the frequency blocks to compress the frequency
- data for each frequency band; and
- 15 assign a priority level to the compressed frequency data for each frequency band,
- wherein at least two of the plurality of frequency bands are assigned a
- different priority level; and
- a wireless interface configured to prioritize transmission of the compressed frequency data
- 20 for the frequency bands over a wireless connection to a receiving device according
- to the priority levels assigned to the frequency bands.
2. The video encoding system as recited in claim 1, wherein the one or more
- processors are configured to apply a wavelet transform to the pixel blocks from the video frame to
- 25 decompose the pixel blocks into the plurality of frequency bands.
3. The video encoding system as recited in claim 2, wherein the wavelet transform is
- a two-level wavelet decomposition applied to each pixel block from the video frame.
- 30 4. The video encoding system as recited in claim 1, wherein lower frequency bands
- are assigned a higher priority level, and higher frequency bands are assigned a lower priority level.
5. The video encoding system as recited in claim 1, wherein, to prioritize transmission
- of the compressed frequency data for the frequency bands over a wireless connection to a receiving

device according to the priority levels assigned to the frequency bands, the wireless interface is configured to:

track performance of the wireless connection; and

transmit one or more higher priority frequency bands and not transmit one or more lower
5 priority frequency bands upon detecting that performance of the wireless
connection has dropped below a threshold.

6. The video encoding system as recited in claim 1, wherein the encoder is further
configured to:

10 determine whether a pixel block is in a foveated region or in a peripheral region of the
video frame; and

assign a lower priority level to the compressed frequency data for at least one frequency
band in the pixel block upon determining that the pixel block is in the peripheral
region of the video frame.

15 7. The video encoding system as recited in claim 1, wherein, to assign a priority level
to the compressed frequency data for each frequency band, the encoder is further configured to tag
metadata for the compressed frequency data for each frequency band with an indication of the
priority level assigned to the frequency band.

20 8. The video encoding system as recited in claim 1, wherein the encoder is further
configured to tag metadata for the compressed frequency data for each frequency band with timing
information for the compressed frequency data for the frequency band, wherein the timing
information includes a deadline for receiving the compressed frequency data for the frequency
25 band at the receiving device.

9. The video encoding system as recited in claim 8, wherein the video frame is divided
into a plurality of slices, each slice including at least one row of pixel blocks, and wherein the
wireless interface is further configured to:

30 determine that the deadline for receiving the compressed frequency data for a lower priority
frequency band from a first slice at the receiving device is closer than the deadline
for receiving the compressed frequency data for a higher priority frequency band
from a second slice at the receiving device; and

transmit the compressed frequency data for the lower priority frequency band from the first slice to the receiving device prior to transmitting the compressed frequency data for the higher priority frequency band from the second slice.

5 10. A method, comprising:

one or more processors configured to implement:

decomposing pixel blocks from a video frame into a plurality of frequency bands;

organizing the frequency bands from the pixel blocks into a plurality of frequency

10 blocks each including frequency data for one of the frequency bands;

applying an encoding technique to the frequency blocks to compress the frequency data for each frequency band;

assigning a priority level to the compressed frequency data for each frequency band,

15 wherein at least two of the plurality of frequency bands are assigned a different priority level; and

prioritizing transmission of the compressed frequency data for the frequency bands over a wireless connection to a receiving device according to the priority levels assigned to the frequency bands.

20 11. The method as recited in claim 10, wherein prioritizing transmission of the compressed frequency data for the frequency bands over a wireless connection to a receiving device according to the priority levels assigned to the frequency bands comprises:

tracking performance of the wireless connection; and

25 transmitting one or more higher priority frequency bands and not transmitting one or more lower priority frequency bands upon detecting that performance of the wireless connection has dropped below a threshold.

12. The method as recited in claim 10, further comprising:

30 determining whether a pixel block is in a foveated region or in a peripheral region of the video frame; and

assigning a lower priority level to the compressed frequency data for at least one frequency band in the pixel block upon determining that the pixel block is in the peripheral region of the video frame.

13. The method as recited in claim 10, wherein assigning a priority level to the compressed frequency data for each frequency band comprises tagging metadata for the compressed frequency data for each frequency band with an indication of the priority level assigned to the frequency band.

5

14. The method as recited in claim 10, further comprising tagging metadata for the compressed frequency data for each frequency band with timing information for the compressed frequency data for the frequency band, wherein the timing information includes a deadline for receiving the compressed frequency data for the frequency band at the receiving device.

10

15. The method as recited in claim 14, wherein the video frame is divided into a plurality of slices, each slice including at least one row of pixel blocks, and wherein the method further comprises:

determining that the deadline for receiving the compressed frequency data for a lower
15 priority frequency band from a first slice at the receiving device is closer than the
deadline for receiving the compressed frequency data for a higher priority
frequency band from a second slice at the receiving device; and

transmitting the compressed frequency data for the lower priority frequency band from the
first slice to the receiving device prior to transmitting the compressed frequency
20 data for the higher priority frequency band from the second slice.

20

16. A system, comprising:

a device comprising one or more processors and a display subsystem;

25

a base station comprising one or more processors configured to:

decompose pixel blocks from a video frame into a plurality of frequency bands and
organize the frequency bands from the pixel blocks into a plurality of
frequency blocks each including frequency data for one of the frequency
bands;

30

apply an encoding technique to the frequency blocks to compress the frequency
data for each frequency band;

assign a priority level to the compressed frequency data for each frequency band,
wherein at least two of the plurality of frequency bands are assigned a
different priority level; and

prioritize transmission of the compressed frequency data for the frequency bands over a wireless connection to the HMD according to the priority levels assigned to the frequency bands;

wherein the one or more processors of the device are configured to:

- 5 decompress the compressed frequency bands received from the base station according to the priority levels assigned to the frequency bands;
perform wavelet synthesis on the decompressed frequency bands to reconstruct the pixel data for the video frame; and
provide the reconstructed pixel data to the display subsystem for display.

10 17. The system as recited in claim 16, wherein, to prioritize transmission of the compressed frequency data for the frequency bands over a wireless connection to a receiving device according to the priority levels assigned to the frequency bands the one or more processors of the base station are further configured to:

- 15 track performance of the wireless connection; and
transmit one or more higher priority frequency bands and not transmit one or more lower priority frequency bands upon detecting that performance of the wireless connection has dropped below a threshold.

20 18. The system as recited in claim 16, wherein the one or more processors of the base station are further configured to:

- determine whether a pixel block is in a foveated region or in a peripheral region of the video frame; and
assign a lower priority level to the compressed frequency data for at least one frequency
25 band in the pixel block upon determining that the pixel block is in the peripheral region of the video frame.

19. The system as recited in claim 16, wherein the one or more processors of the base station are further configured to:

- 30 determine that the deadline for receiving the compressed frequency data for a lower priority frequency band from a first pixel block of the frame at the device is closer than the deadline for receiving the compressed frequency data for a higher priority frequency band from a second pixel block of the frame at the HMD; and

transmit the compressed frequency data for the lower priority frequency band from the first pixel block to the device prior to transmitting the compressed frequency data for the higher priority frequency band from the second pixel block.

- 5 20. The system as recited in claim 16, wherein the device is a head-mounted display (HMD).

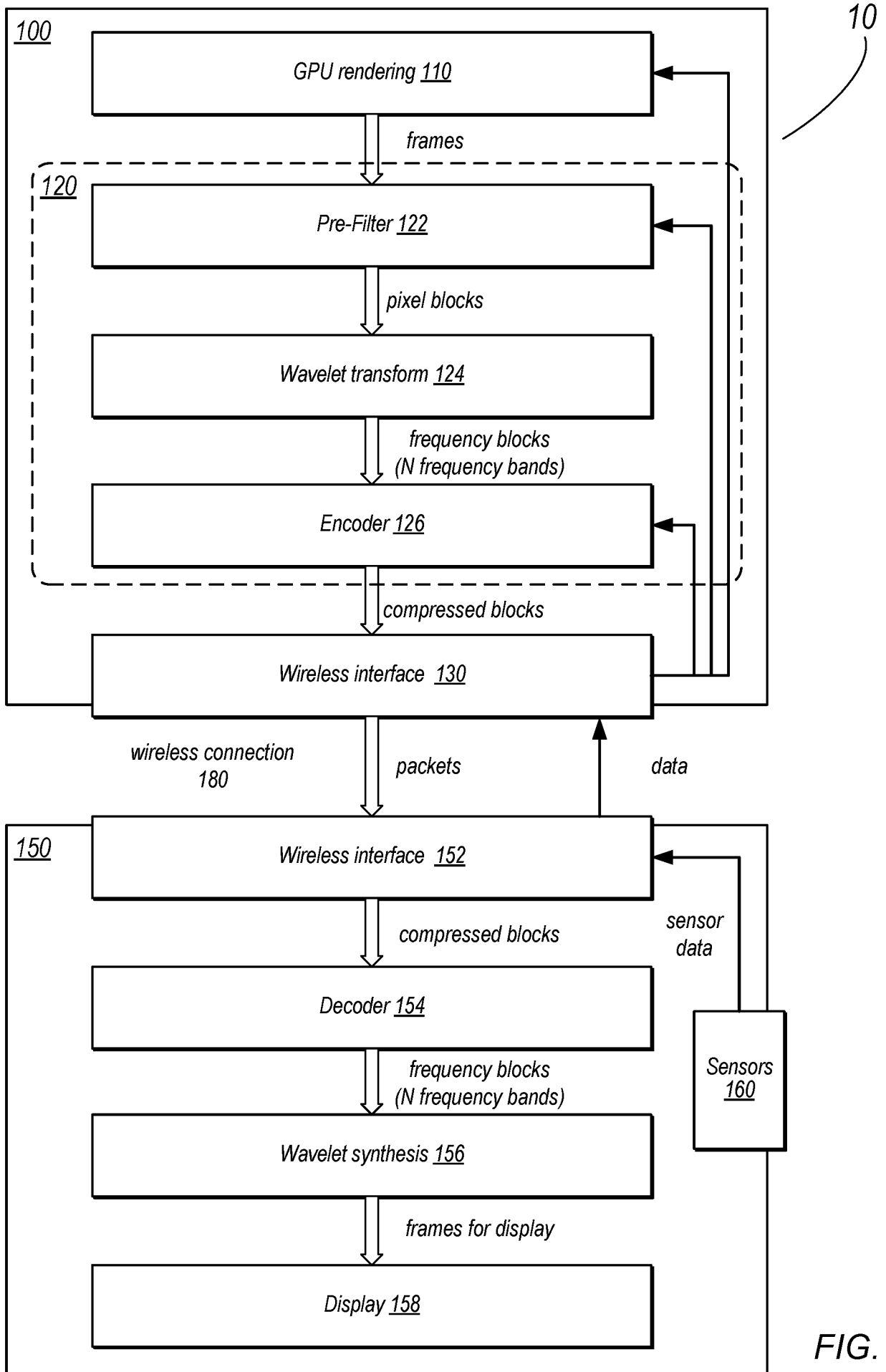


FIG. 1

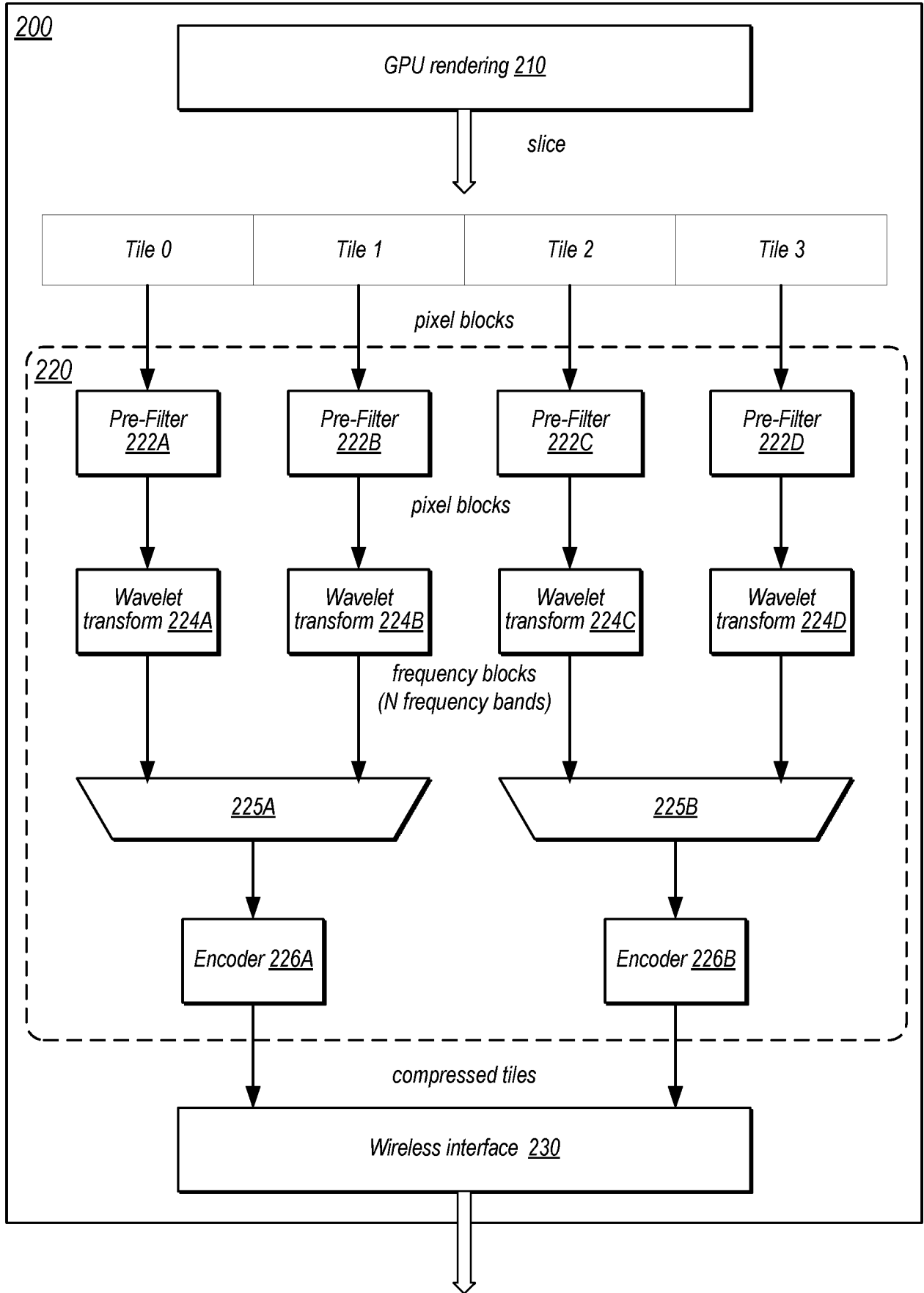


FIG. 2

3 / 12

Frame

<i>slice 0</i>	<i>Tile 0</i>	<i>Tile 1</i>	<i>Tile 2</i>	<i>Tile 3</i>
<i>slice 1</i>	<i>Tile 0</i>	<i>Tile 1</i>	<i>Tile 2</i>	<i>Tile 3</i>
<i>slice 2</i>	<i>Tile 0</i>	<i>Tile 1</i>	<i>Tile 2</i>	<i>Tile 3</i>
<i>slice 3</i>	<i>Tile 0</i>	<i>Tile 1</i>	<i>Tile 2</i>	<i>Tile 3</i>

⋮

FIG. 3A

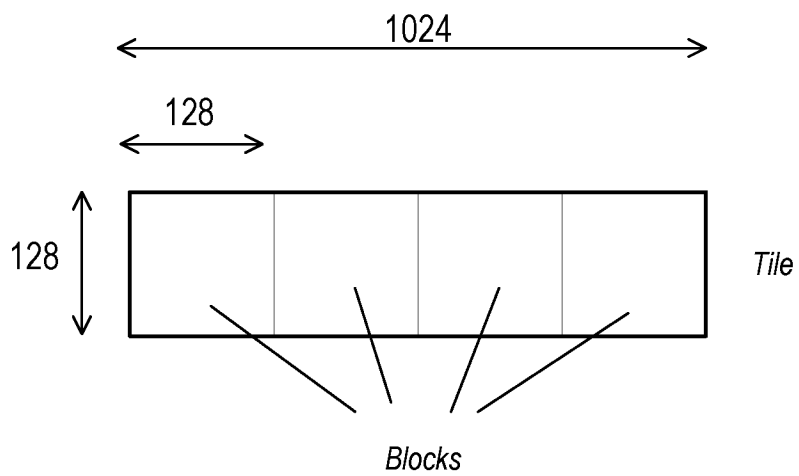


FIG. 3B

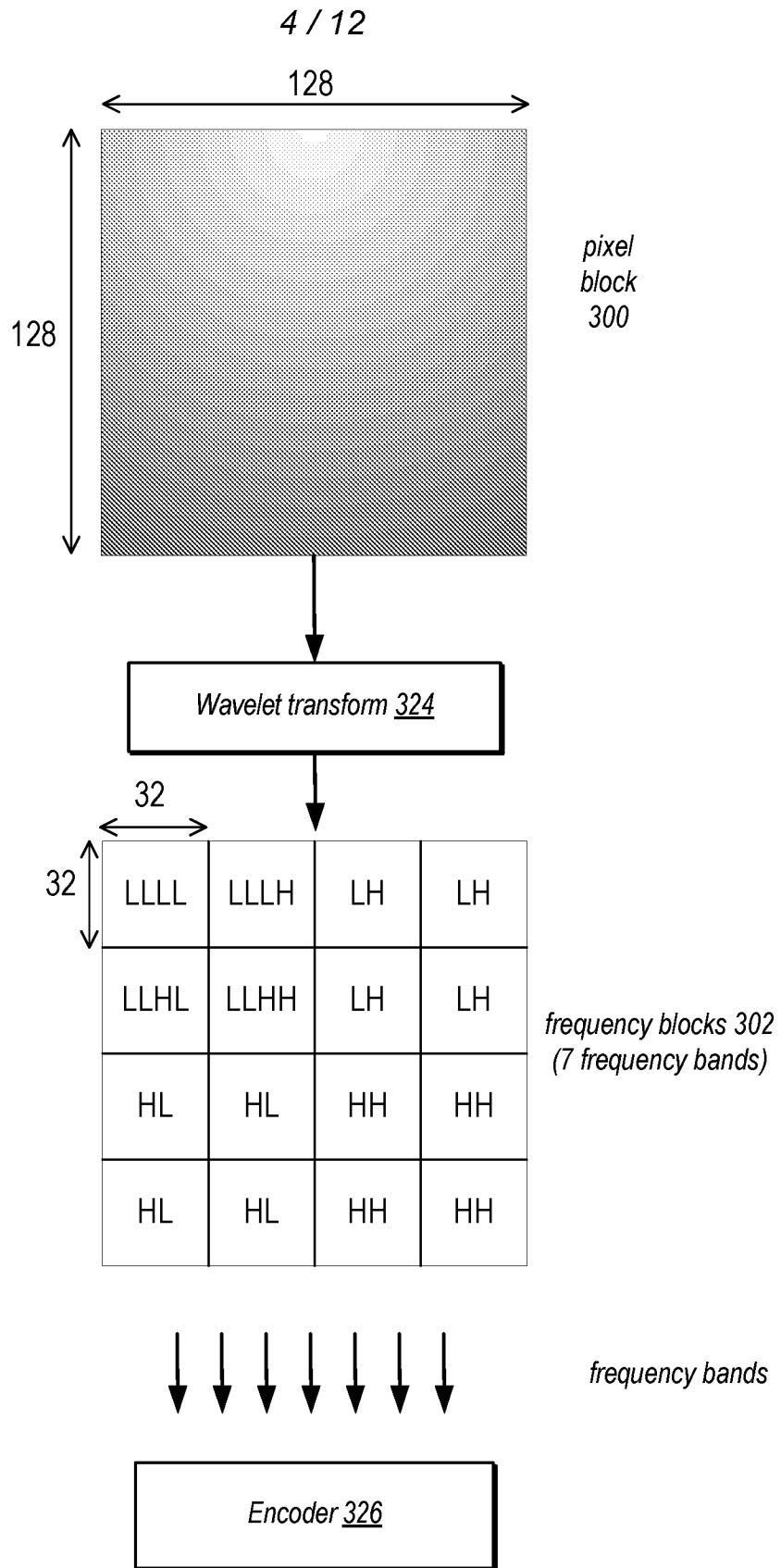


FIG. 3C

5 / 12

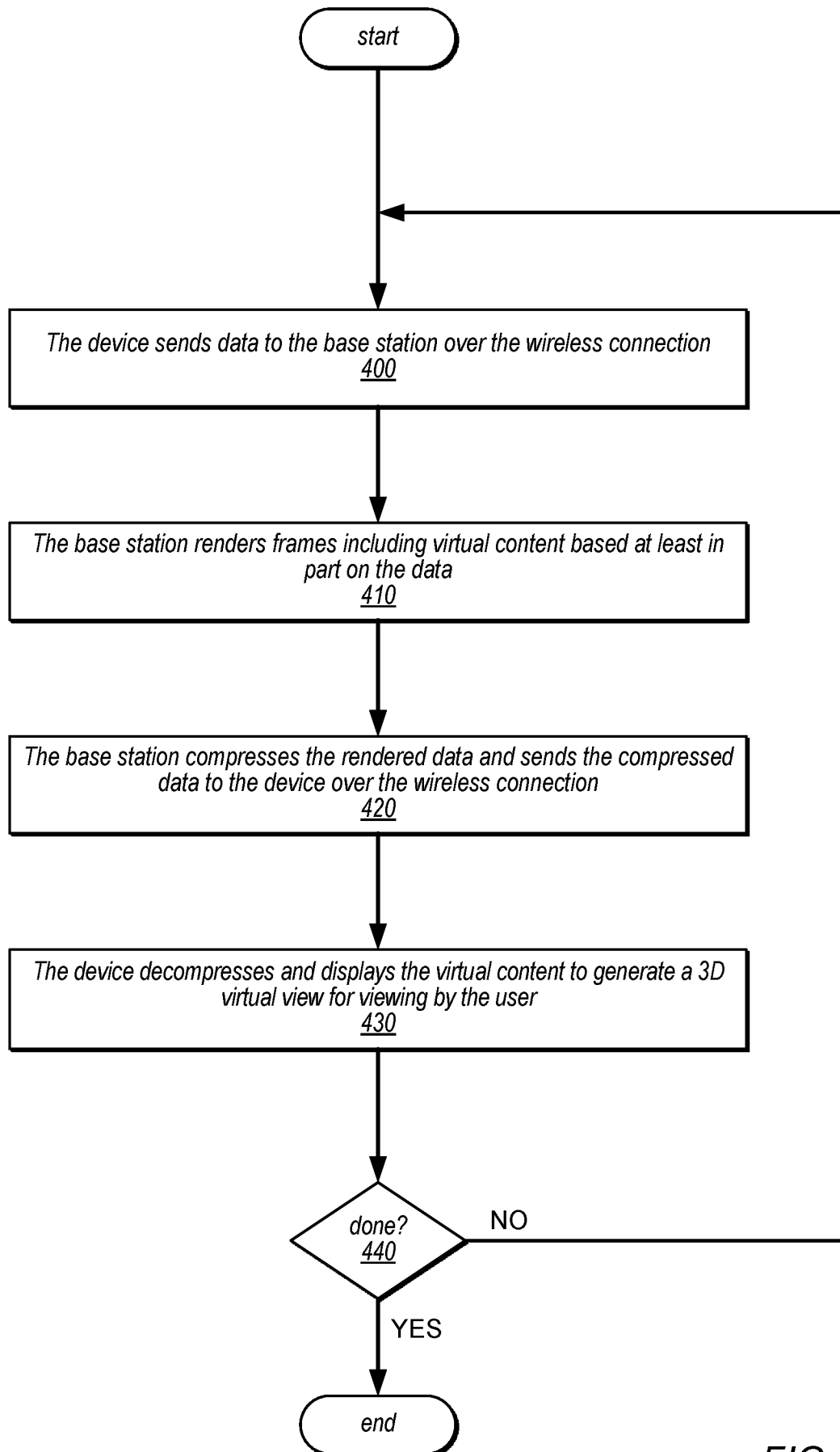


FIG. 4

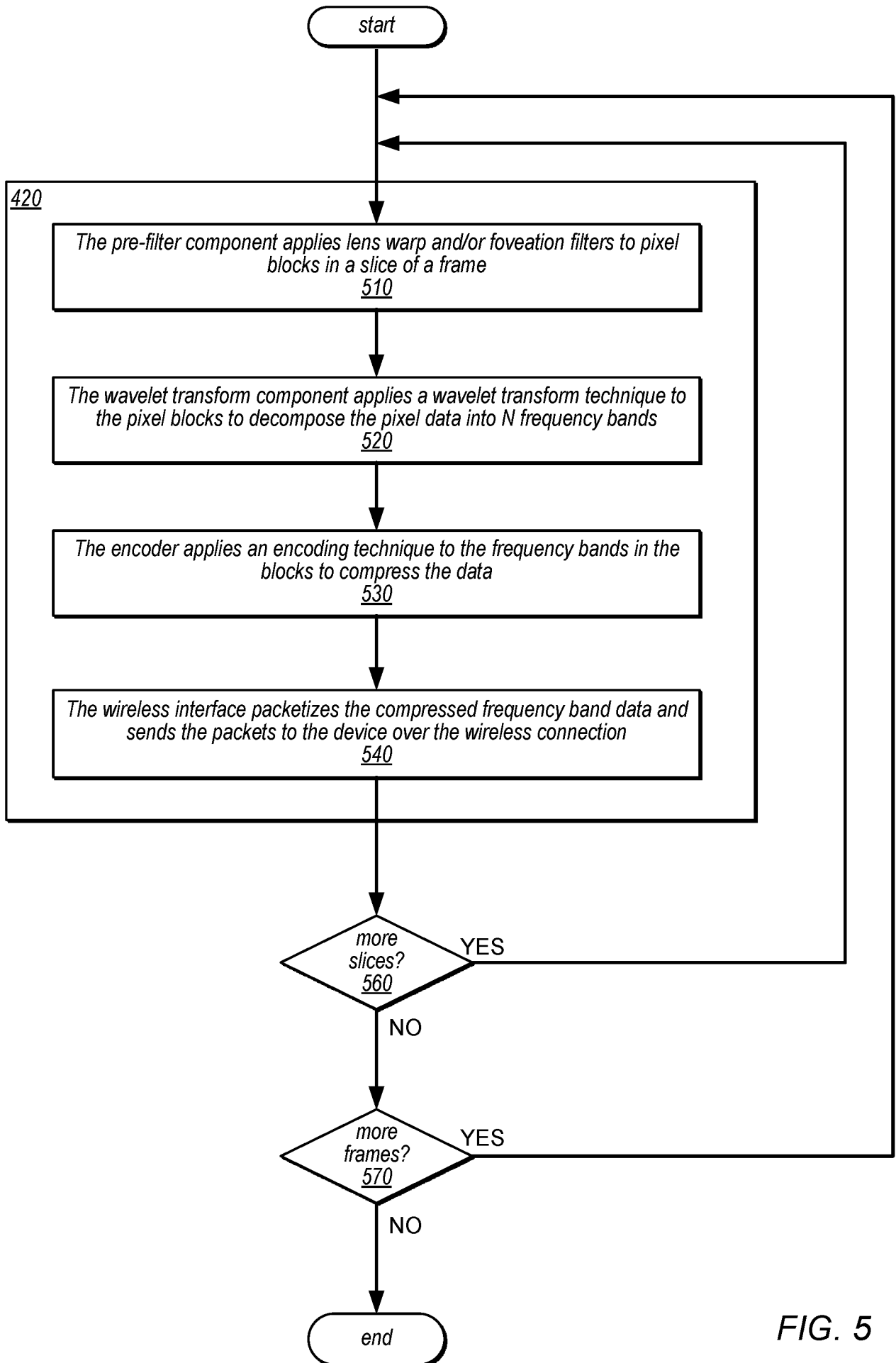


FIG. 5

7 / 12

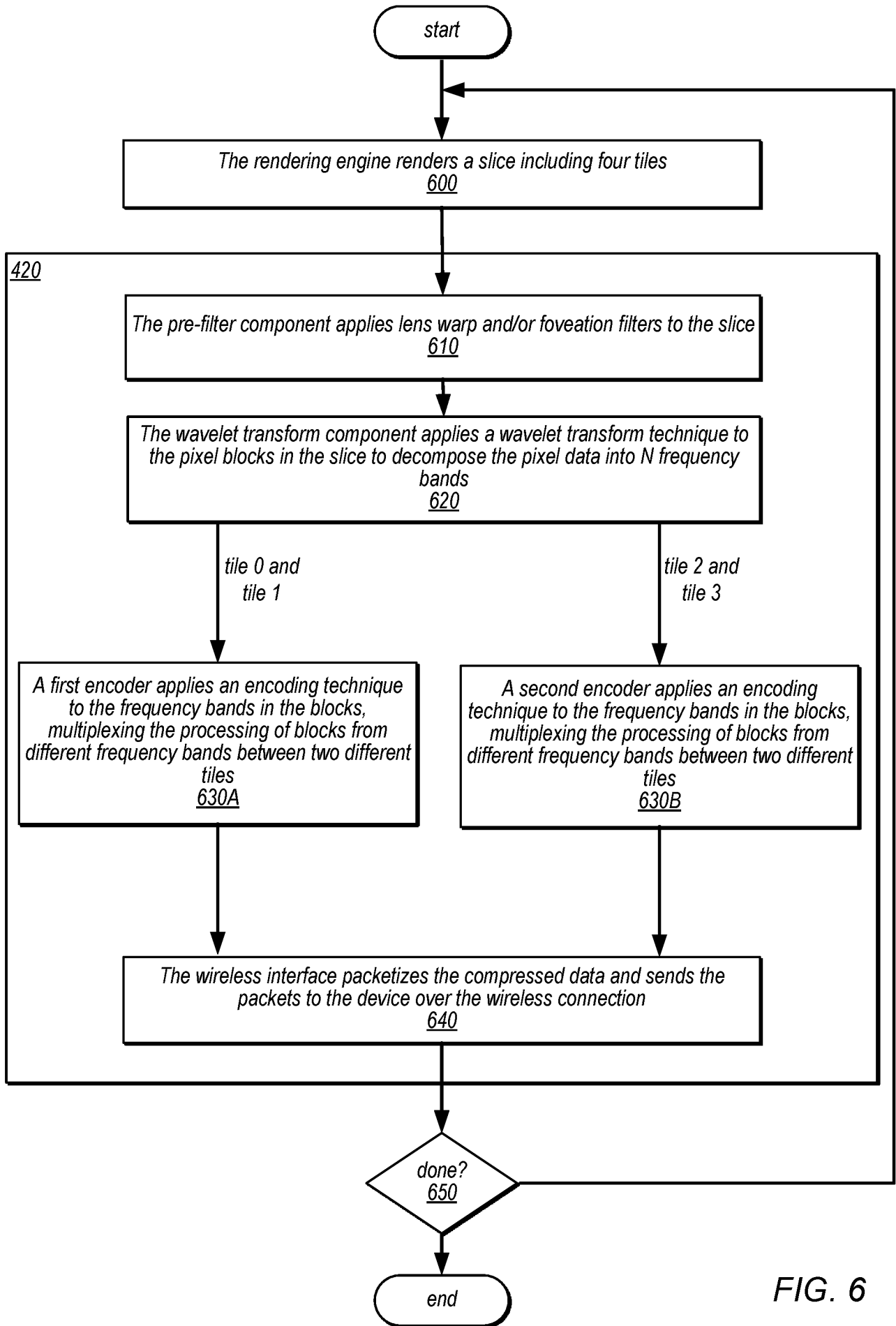


FIG. 6

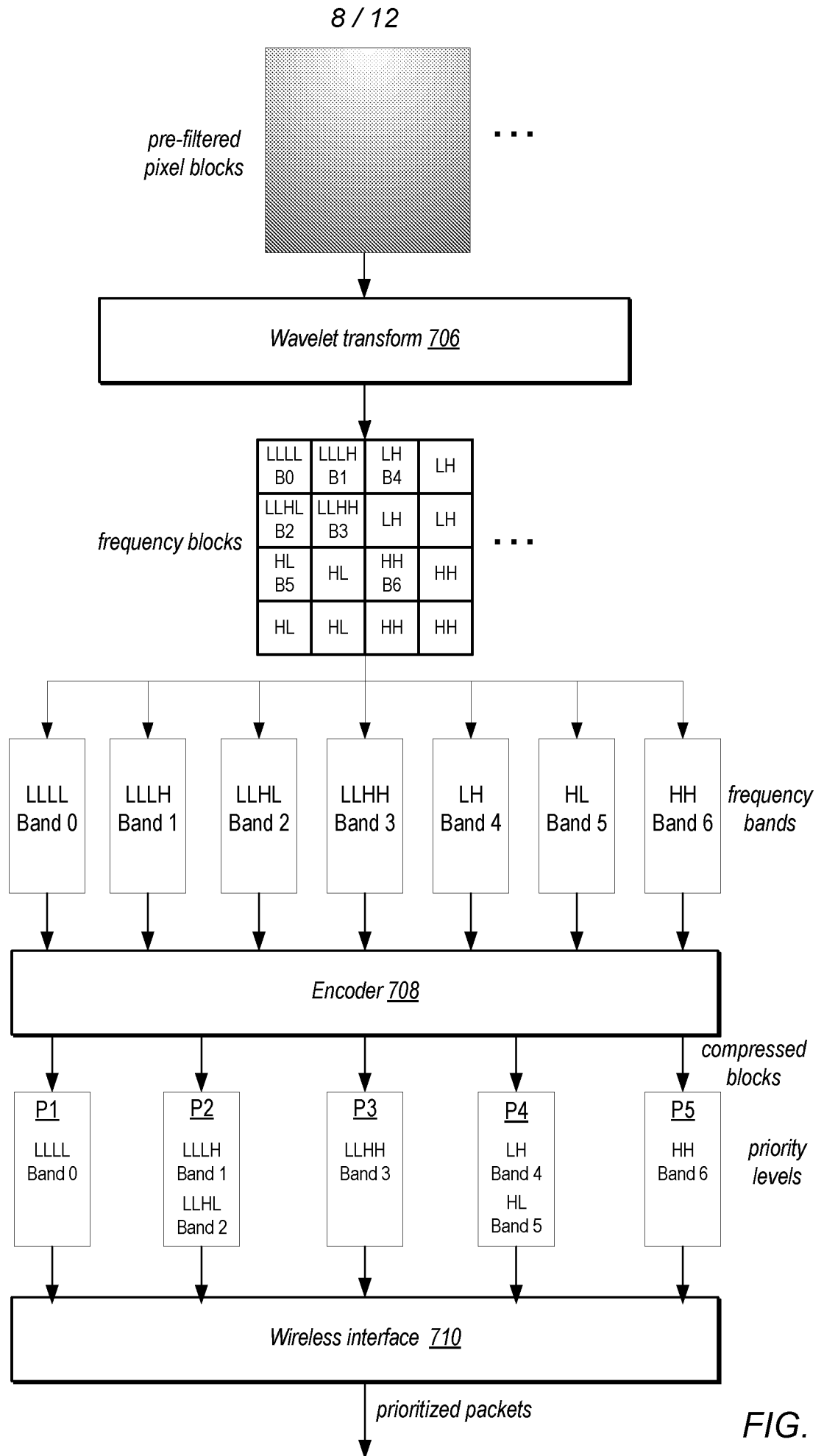


FIG. 7

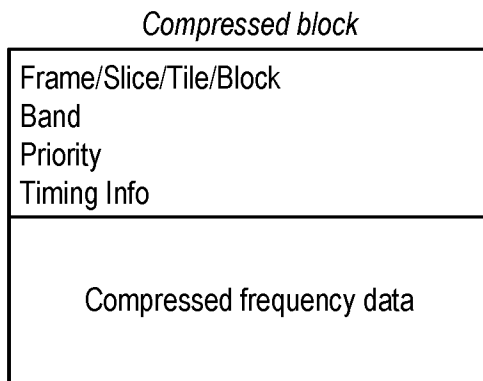


FIG. 8A

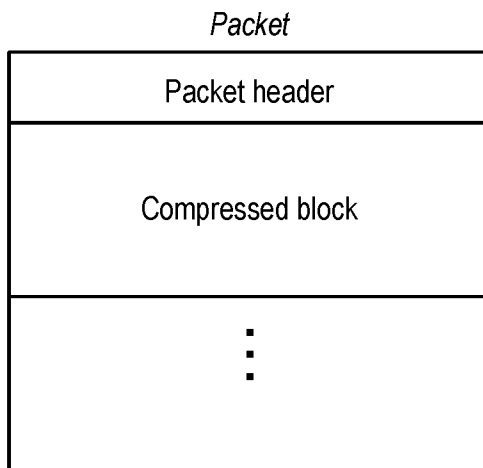


FIG. 8B

10 / 12

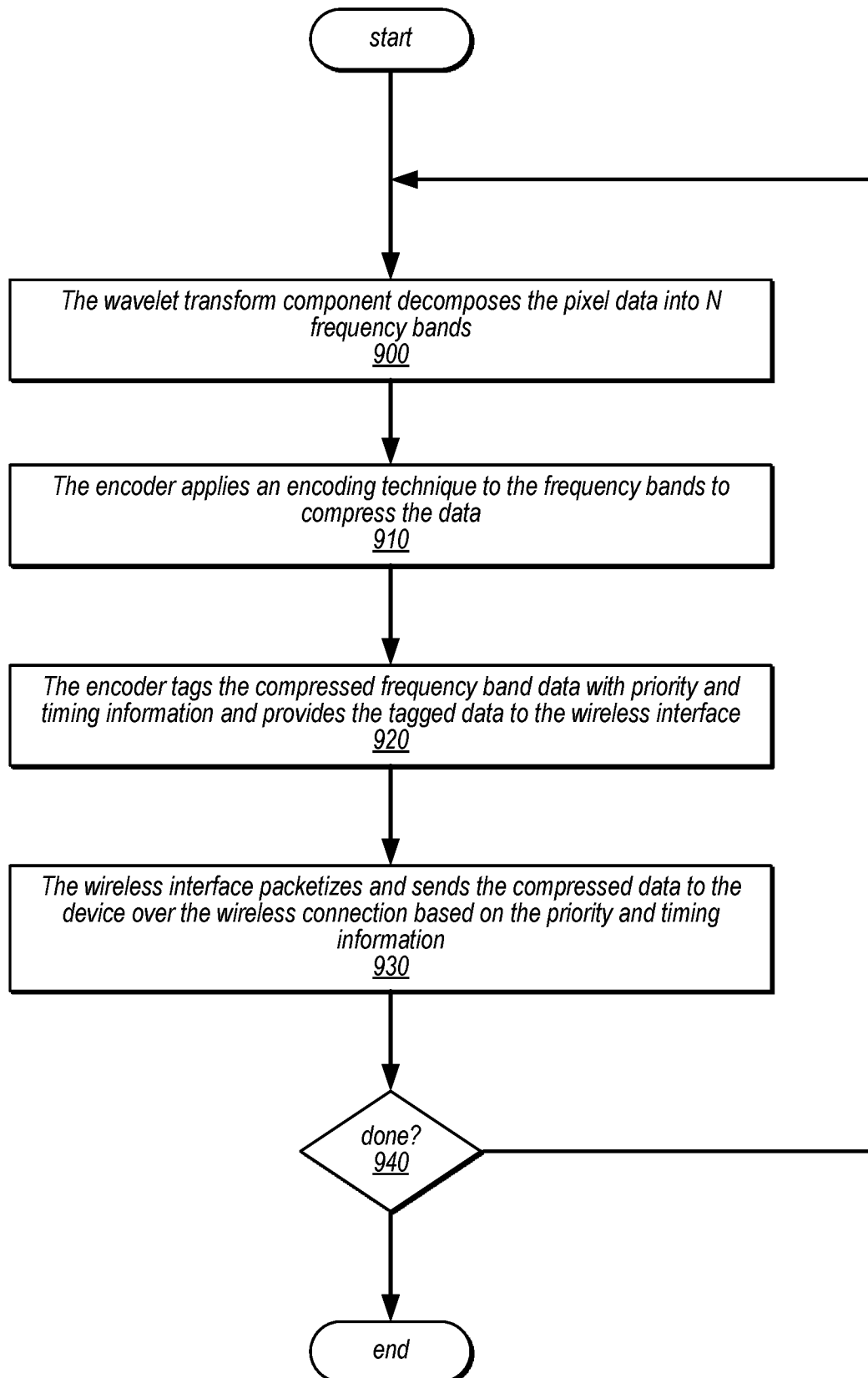


FIG. 9

2000

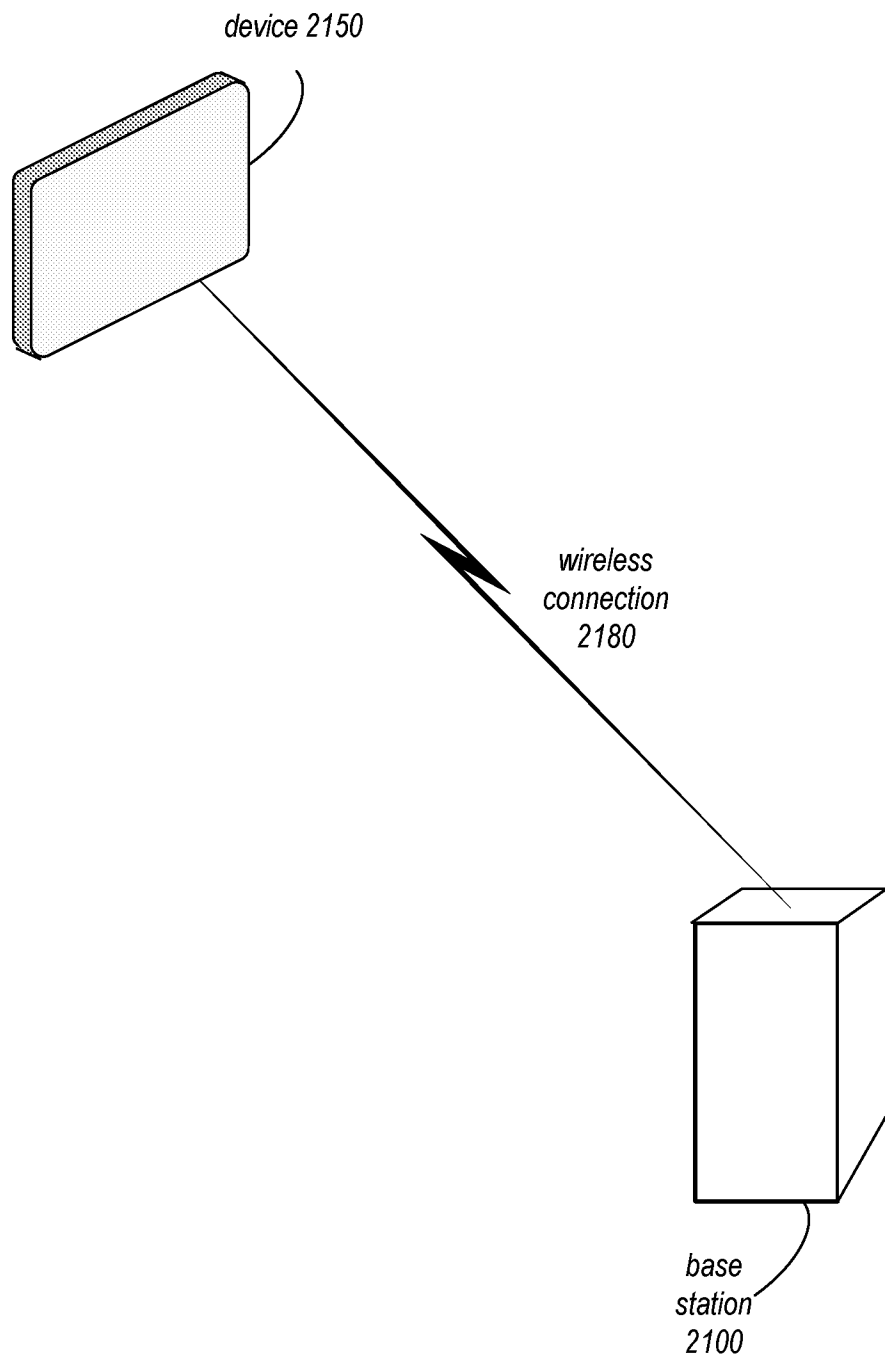


FIG. 10

(WORLD)

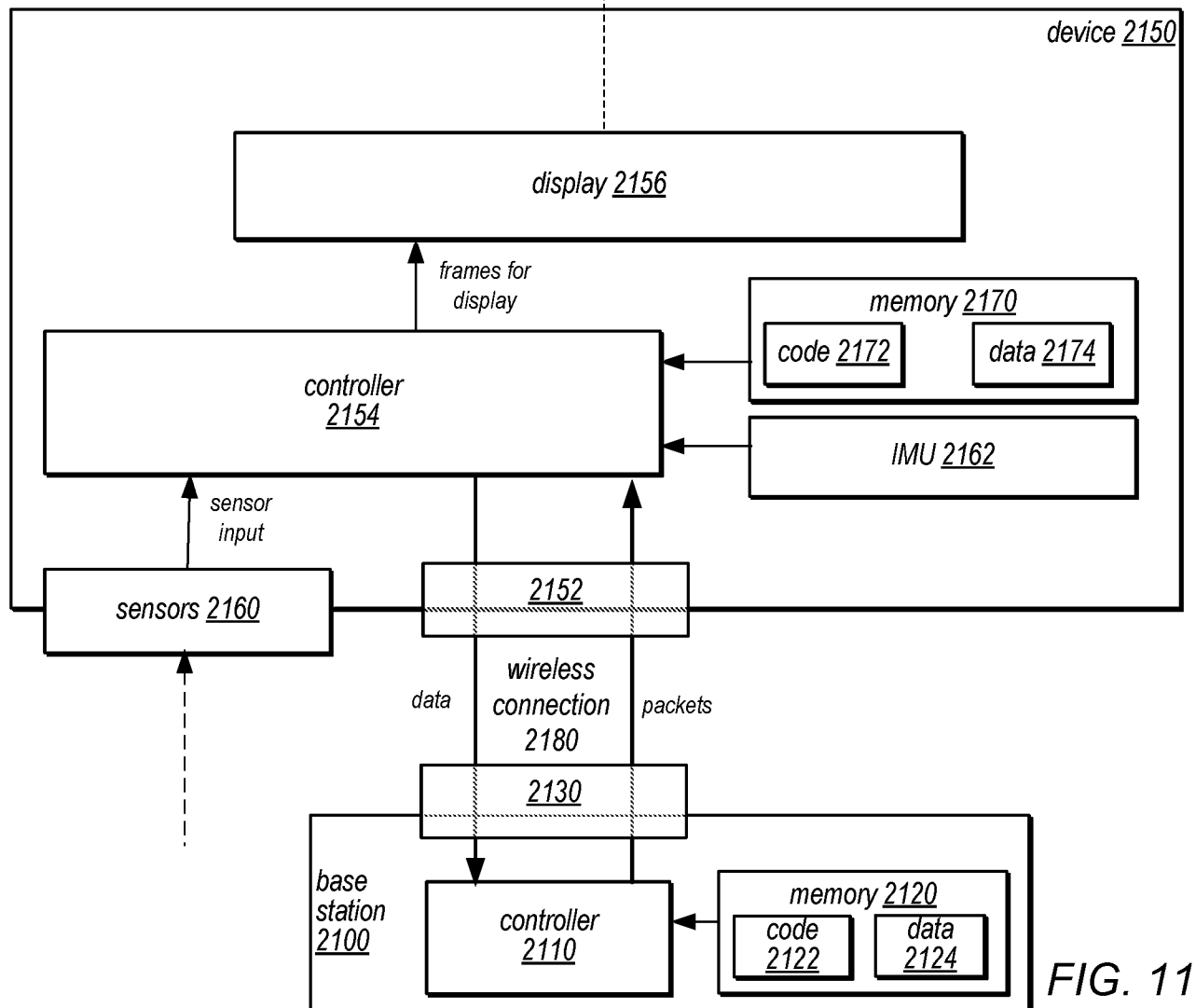
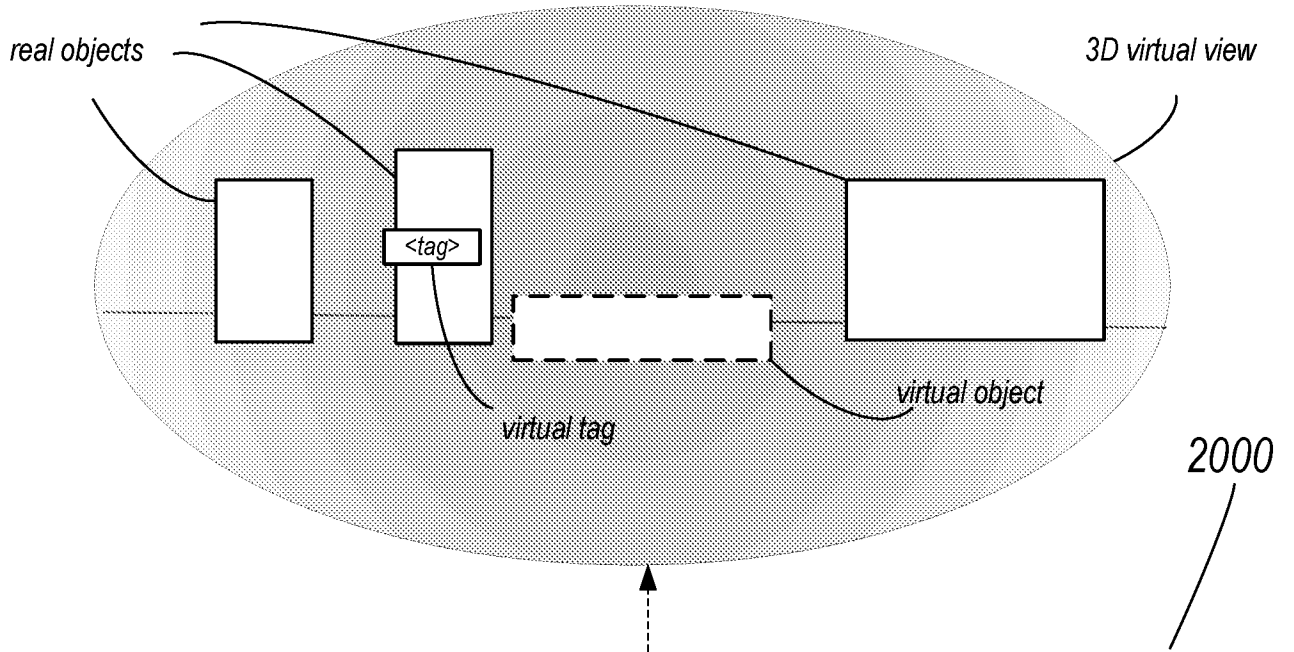


FIG. 11

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2019/039591

A. CLASSIFICATION OF SUBJECT MATTER
 INV. H04N19/132 H04N19/164 H04N19/167 H04N19/17 H04N19/169
 H04N19/37 H04N19/645 H04N19/63
 ADD.
 According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED
 Minimum documentation searched (classification system followed by classification symbols)
 H04N A63F

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, WPI Data, COMPENDEX, INSPEC, IBM-TDB

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 1 301 021 A2 (EASTMAN KODAK CO [US]) 9 April 2003 (2003-04-09)	1-4,6,7, 10,12, 13,16, 18,20
Y	paragraph [0012] - paragraph [0018]; figure 1 paragraph [0038] - paragraph [0053]; figures 3-5	5,8,9, 11,14, 15,17,19
X	----- US 9 955 176 B2 (INTEL CORP [US]) 24 April 2018 (2018-04-24) column 13, line 12 - line 42; figure 4 column 20, line 58 - column 22, line 22; claim 1; figures 8A, 8B column 42, line 25 - column 44, line 29; figures 23-24B column 57, line 45 - line 59 ----- -/--	1-3,10, 16,20

Further documents are listed in the continuation of Box C. See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
 "E" earlier application or patent but published on or after the international filing date
 "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)
 "O" document referring to an oral disclosure, use, exhibition or other means
 "P" document published prior to the international filing date but later than the priority date claimed

"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
 "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
 "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
 "&" document member of the same patent family

Date of the actual completion of the international search 10 October 2019	Date of mailing of the international search report 23/10/2019
--	--

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 Georgiou, Georgia
--	---

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2019/039591

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	EP 1 073 275 A2 (SAMSUNG ELECTRONICS CO LTD [KR]) 31 January 2001 (2001-01-31) paragraph [0012] - paragraph [0019]; figures 2-3b -----	5,8,9, 11,14, 15,17,19
Y	US 2016/227435 A1 (FREEMAN DUDLEY C [US] ET AL) 4 August 2016 (2016-08-04) paragraph [0062] - paragraph [0063] -----	5,11,17
A	US 2009/189981 A1 (SIANN JON [US] ET AL) 30 July 2009 (2009-07-30) paragraph [0096] - paragraph [0106]; figure 1A paragraph [0241] -----	1-20

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No PCT/US2019/039591

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 1301021	A2	09-04-2003	EP 1301021 A2 09-04-2003
			JP 2003202850 A 18-07-2003
			US 2003067476 A1 10-04-2003
			US 2004207632 A1 21-10-2004
			US 2004207635 A1 21-10-2004

US 9955176	B2	24-04-2018	CN 108293138 A 17-07-2018
			US 2017155906 A1 01-06-2017
			WO 2017095491 A1 08-06-2017

EP 1073275	A2	31-01-2001	CN 1277517 A 20-12-2000
			EP 1073275 A2 31-01-2001
			JP 3870009 B2 17-01-2007
			JP 2001028752 A 30-01-2001
			KR 20010002218 A 15-01-2001
			US 6744738 B1 01-06-2004

US 2016227435	A1	04-08-2016	US 2016227432 A1 04-08-2016
			US 2016227435 A1 04-08-2016
			US 2016227546 A1 04-08-2016
			US 2016227551 A1 04-08-2016
			US 2018249368 A1 30-08-2018
			WO 2016126735 A2 11-08-2016

US 2009189981	A1	30-07-2009	EP 2238758 A2 13-10-2010
			US 2009189981 A1 30-07-2009
			US 2011096168 A1 28-04-2011
			US 2016173833 A1 16-06-2016
			WO 2009094591 A2 30-07-2009
