



- (51) **International Patent Classification:**
H04N 21/2662 (2011.01) *H04W 72/04* (2009.01)
H04B 7/005 (2006.01)
- (21) **International Application Number:**
PCT/SG2013/000124
- (22) **International Filing Date:**
28 March 2013 (28.03.2013)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
201202261-2 28 March 2012 (28.03.2012) SG
- (71) **Applicant: AGENCY FOR SCIENCE, TECHNOLOGY AND RESEARCH** [SG/SG]; 1 Fusionopolis Way, #20-10 Connexis, Singapore 138632 (SG).
- (72) **Inventors: ZHOU, Huan;** c/o PKM, Institute for Infocomm Research, 1 Fusionopolis Way #21-01, Connexis, South Tower, Singapore 138632 (SG). **SHU, Haiyan;** c/o PKM, Institute for Infocomm Research, 1 Fusionopolis Way #21-01, Connexis, South Tower, Singapore 138632 (SG). **YU, Rongshan;** c/o PKM, Institute for Infocomm Research, 1 Fusionopolis Way #21-01, Connexis, South Tower, Singapore 138632 (SG). **HUANG, Haibin;** c/o

PKM, Institute for Infocomm Research, 1 Fusionopolis Way #21-01, Connexis, South Tower, Singapore 138632 (SG). **RAHARDJA, Susanto;** c/o PKM, Institute for Infocomm Research, 1 Fusionopolis Way #21-01, Connexis, South Tower, Singapore 138632 (SG).

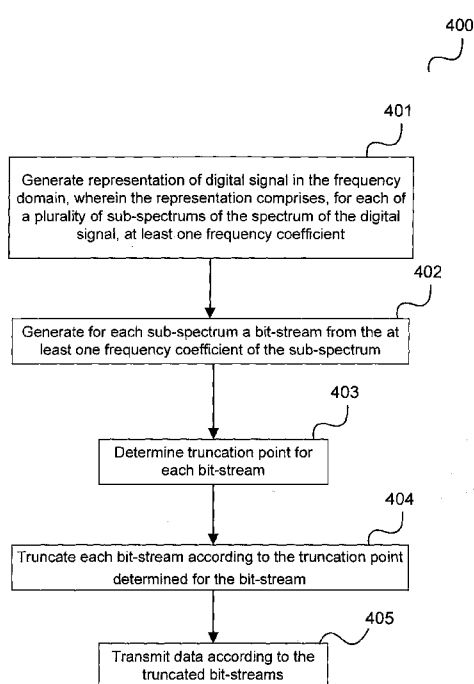
(74) **Agent: VIERING, JENTSCHURA & PARTNER LLP;** P.O. Box 1088, Rochor Post Office, Rochor Road, Singapore 911833 (SG).

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

(84) **Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,

[Continued on next page]

(54) **Title:** METHOD FOR TRANSMITTING A DIGITAL SIGNAL, METHOD FOR RECEIVING A DIGITAL SIGNAL, TRANSMISSION ARRANGEMENT AND COMMUNICATION DEVICE



(57) **Abstract:** A method for transmitting a digital signal is provided comprising generating a representation of the digital signal in the frequency domain, wherein the representation comprises, for each of a plurality of sub-spectrums of the spectrum of the digital signal, at least one frequency coefficient; generating, for each sub-spectrum, a bit-stream from the at least one frequency coefficient of the sub-spectrum; determining, for each bit-stream, a truncation point; truncating each bit-stream according to the truncation point determined for the bit-stream and transmitting data according to the truncated bit-streams.



EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, **Published:**

LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK,

SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,

GW, ML, MR, NE, SN, TD, TG).

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

**METHOD FOR TRANSMITTING A DIGITAL SIGNAL, METHOD FOR
RECEIVING A DIGITAL SIGNAL, TRANSMISSION ARRANGEMENT AND
COMMUNICATION DEVICE**

5 Embodiments of the invention generally relate to methods for transmitting a digital signal, methods for receiving a digital signal, transmission arrangements and communication devices.

10 To efficiently store and transmit audio content (or generally media content, e.g. video), an audio signal (or generally a media signal) may be compressed to save memory space and transmission bandwidth. There are two types of compression formats: non-scalable compression and scalable compression. Non-scalable compression is a compression format that is dedicated to a specified transmission bit rate. It provides the most efficient
15 compression in terms of rate-distortion performance. According to this compression, after the compression, the bit rate of, for example, each audio frame of an audio signal is fixed. This can be considered to be not network friendly since the network has to provide a fixed transmission bit rate. Further, to fulfil different requirements for different clients, the audio content has to be compressed into several non-scalable copies. This is illustrated in figure 1.

20 Figure 1 shows a communication system.

25 The communication system includes a non-scalable encoder 101 which provides a plurality of copies of the same audio signal (e.g. the same audio file) with different compression rates (and hence bit rates) to a server 103. The server 103 may chose, depending on the desired quality by a plurality of client devices 104, one of the copies 102 for each client 104 and transmit it to the client 104 via a communication network 105, e.g. the Internet.

30 This requires that the server 103 stores the various copies which may be seen to waste storage space of the server 103.

On the other hand, scalable compression can be seen as a network friendly coding format. It is illustrated in figure 2.

Figure 2 shows a communication system 200.

The communication system includes a scalable encoder 201 which provides an audio
5 signal (e.g. an audio file) in a single compressed version 202 to a server 203. The server
203 can truncate the compressed version 202 of the audio signal according to the desired
bit-rates of client devices 204 and send, to each client device 204, a truncated version 205
(i.e. a truncated bit-stream) of the audio file over a communication network 206, e.g. the
Internet.

10 In the compressed version 202, the audio signal is in the form of a scalable audio bit-
stream that can be truncated to adapt the bit rate of the audio signal, for example to the
actual network condition.

15 Accordingly, the server 203 can respond to the requirement from each end user and
provide the end users with the most suitable bit-stream by truncating the original scalable
audio. Although the scalable compression format has this benefit, it does require extra bits
to enable the scalability. Due to the introduction of these extra bits, the compression
efficiency is reduced.

20 The scalable compression format generally provides more flexibility in bit rate adaptation.
It is typically more suitable to be applied in real time streaming applications when
compared with the non-scalable compression formats. When the scalable format is used in
streaming applications, the scaling dimension and the coding priority order can
25 significantly influence the bit-stream quality. In the scalable format, the coding priority
order in the scalable bit-stream is typically fixed after the encoding process and cannot be
changed in the truncation process. This reduces the flexibility of the scalable format to
provide high quality output. An example is the scalable audio coding scheme MPEG-4
Scalable to Lossless Coding (SLS). In SLS, a core MPEG-4 AAC layer generates an AAC
30 compliant bit-stream and a lossless enhanced layer generates a scalable bit-stream which
scans and codes the residual bit-plane symbols from the Most Significant Bit (MSB) to the
Least Significant Bit (LSB). Both bit-streams are embedded in the final SLS bit-stream.
The resulting format is illustrated in figure 3.

Figure 3 shows an audio signal in encoded format 301 and truncated format 302.

For each frame 303, the encoded audio signal 301 includes core layer information 304 and enhancement layer information 305 wherein the enhancement layer information is in the form of a bit stream starting from the most significant bits 306 of the residual bit-plane symbols and ending with the least significant bits 307 which are further followed by low energy scale factor band information 308.

When the audio signal is to be delivered to an end user (e.g. a client device 104), the server 103 truncates bits from the enhancement layer bit-stream 305 from its end for each frame, which includes the least significant information part, as defined by the encoder. The more significant information part, which is located, relatively, in the front of the bit-stream 305, is kept (as illustrated in the truncated format 302).

In this case, the coding priority order of the encoded information in the enhancement layer bit-stream 305 is defined by the bit plane scanning order. Although the perceptual priority is considered and reflected on the bit-plane scanning order, it is pre-defined in the encoding process and cannot be changed after the scalable bit-stream 305 is generated, which leads to a fixed Rate-Distortion (R-D) relationship, typically optimized for the lossless case. This fixed R-D relationship cannot make the output quality of SLS bit-stream being optimized for all specified bit rate within a wide range.

Another scalable audio coding scheme is Bit-Sliced Arithmetic Coding (BSAC), which is used in MPEG-4 AAC. Its scalable bit-stream has the layered structure including one base layer and multiple enhanced layers with 1kbps enhancement step size. Like SLS, the coding priority order is fixed during its encoding process and cannot be changed during the truncation. Thus, only SNR (signal-to-noise ratio) scalability can be provided by truncating enhanced layers when bit-rate is insufficient.

Another scalable audio coding scheme is Scalable Sampling Rate coding (SSR), which was first published AAC in 1997. It is a hierarchical sampling-rate-scalable coding scheme and

provides sampling rate scalability with a very simple bit-stream splitter to remove upper frequency bands. However, it does not provide SNR scalability.

A higher flexibility of scalability is desirable.

5

According to one embodiment, a method for transmitting a digital signal is provided comprising generating a representation of the digital signal in the frequency domain, wherein the representation comprises, for each of a plurality of sub-spectrums of the spectrum of the digital signal, at least one frequency coefficient; generating, for each sub-spectrum, a bit-stream from the at least one frequency coefficient of the sub-spectrum; determining, for each bit-stream, a truncation point; truncating each bit-stream according to the truncation point determined for the bit-stream and transmitting data according to the truncated bit-streams.

10

15

According to a further embodiment, a method for receiving a digital signal is provided comprising receiving data including a plurality of bit-streams; generating, from each bit-stream, at least one frequency coefficient; aggregating the frequency coefficients to a frequency domain representation of a digital signal; and transforming the frequency domain representation to the time domain to reconstruct the digital signal.

20

Further, a transmission arrangement and a communication device according to the method for transmitting a digital signal and a method for receiving a digital signal are provided.

25

In the drawings, like reference characters generally refer to the same parts throughout the different views. The drawings are not necessarily to scale, emphasis instead generally being placed upon illustrating the principles of the invention. In the following description, various embodiments of the invention are described with reference to the drawings, in which:

30

Figure 1 shows a communication system.

Figure 2 shows a communication system.

Figure 3 shows an audio signal in encoded format and truncated format.

Figure 4 shows a flow diagram.

5 Figure 5 shows a transmission arrangement.

Figure 6 shows a flow diagram.

Figure 7 shows a communication device.

10

Figure 8 shows a communication arrangement.

Figure 9 shows a bit plane representation for an audio signal and a bit-stream.

15 Figure 10 shows an encoder.

Figure 11 shows a truncator.

Figure 12 illustrates a truncation process.

20

Figure 13 shows a receiver.

Figure 14 shows a performance diagram.

25 According to one embodiment, a scalable coding scheme is provided that can provide two dimensional compression scalability: frequency scalability and SNR scalability. A method for transmitting a digital signal according to one embodiment is illustrated in figure 4.

Figure 4 shows a flow diagram 400.

30

The flow diagram 400 illustrates a method for transmitting a digital signal.

In 401, a representation of the digital signal in the frequency domain is generated, wherein the representation comprises, for each of a plurality of sub-spectrums of the spectrum of the digital signal, at least one frequency coefficient.

- 5 In 402, for each sub-spectrum, a bit-stream is generated from the at least one frequency coefficient of the sub-spectrum.

In 403, for each bit-stream, a truncation point is determined.

- 10 In 404 each bit-stream is truncated according to the truncation point determined for the bit-stream.

In 405, data according to the truncated bit-streams (e.g. data including the information from the truncated bit-streams) are transmitted.

15

In other words, the frequency spectrum is segmented and for each spectrum segment (or sub-spectrum) an individual bit-stream (e.g. enhancement layer bit-stream) is generated and individually truncated (which may include cutting the bit-stream completely, i.e. omitting the bit-stream). Thus, in addition to the scalability in terms of bit rate a scalability in terms of frequency is achieved since, for example, bit-streams for certain sub-spectrums may be truncated more strongly than bit-streams for other sub-spectrums.

20

According to one embodiment, the coding priority order in a bit-stream can be seen to be flexibly changed in the truncation process. For example, a scalable encoder generates a plurality of bit-streams, with each bit-stream being associated with a subset of spectral information of the signal and the bit-streams may be ("intelligently") truncated, e.g. by a server independently, e.g. by completely dropping out some bit-streams (to provide frequency scalability) or truncating those bit-streams (to provide SNR scalability).

25

- 30 Compared to SLS and BSAC the coding scheme described above provides additional frequency scalability. In term of scalability capability, the coding scheme described above provides both frequency and SNR scalability simultaneously.

By introducing the two-dimensional scalability, the trade-off between the two scalability can be adjusted at the truncator which allows generating an R-D optimized output under a given network bandwidth.

- 5 The method may further comprise dividing the spectrum into the plurality of sub-spectrums.

The digital signal is for example a media signal.

- 10 For example, the digital signal is an audio signal or a video signal.

According to one embodiment, generating a bit-stream from the at least one frequency coefficient of the sub-spectrum includes bit-plane scanning the at least one frequency coefficient of the sub-spectrum.

15

The method may further comprise multiplexing the truncated bit-streams to an overall bit-stream wherein transmitting data according to the truncated bit-streams includes transmitting the overall bit-stream.

- 20 The method may further comprise transmitting, for each truncated bit-stream, information about the truncation point. The truncation point specifies at which point a bit-stream is truncated and may for example be expressed by a bit position (e.g. the last bit that is not truncated) or a length of the truncated bit-stream.

- 25 According to one embodiment, determining the truncation point includes determining a truncation length and truncating the bit-stream according to the truncation point includes truncating the bit-stream according to the truncation length.

- 30 The data are for example transmitted to a communication device via a communication channel and the truncation point is for example determined based on the capacity of the communication channel.

According to one embodiment, determining a truncation point for a bit-stream includes deciding whether the bit-stream should be omitted from the data and, if it is decided that the bit-stream should be omitted, setting the truncation point such that truncating the bit-stream according to the truncation point leads to the bit-stream being omitted from the data. For example, omitting a bit-stream may be expressed by a truncation length of zero.

The method illustrated in figure 4 is for example carried out by a transmission arrangement as illustrated in figure 5.

Figure 5 shows a transmission arrangement 500.

The transmission arrangement 500 comprises an encoder 501 configured to generate a representation of the digital signal in the frequency domain, wherein the representation comprises, for each of a plurality of sub-spectrums of the spectrum of the digital signal, at least one frequency coefficient and to generate, for each sub-spectrum, a bit-stream from the at least one frequency coefficient of the sub-spectrum.

Further, the transmission arrangement 500 comprises a truncator 502 configured to determine, for each bit-stream, a truncation point and to truncate each bit-stream according to the truncation point determined for the bit-stream.

The transmission arrangement 500 further comprises a transmitter 503 configured to transmit data according to the truncated bit-streams.

The transmitted truncated bit streams are for example processed according to a method as illustrated in figure 6.

Figure 6 shows a flow diagram 600.

The flow diagram 600 illustrates a method for receiving a digital signal.

In 601, data including a plurality of bit-streams is received.

In 602, at least one frequency coefficient is determined from each bit-stream,

In 603, the frequency coefficients are aggregated to a frequency domain representation of a digital signal.

5

In 604, the frequency domain representation is transformed to the time domain to reconstruct the digital signal.

For example, the data includes an overall bit-stream and the method comprises
10 demultiplexing the overall bit-stream to generate the plurality of bit-streams.

The method illustrated in figure 6 is for example carried out by a communication device as illustrated in figure 7.

15 Figure 7 shows a communication device 700.

The communication device 700 comprises a receiver 701 configured to receive data including a plurality of bit-streams.

20 The communication device 700 further comprises a decoder 702 configured to generate, from each bit-stream, at least one frequency coefficient, aggregate the frequency coefficients to a frequency domain representation of a digital signal and transform the frequency domain representation to the time domain to reconstruct the digital signal.

25 It should be noted that embodiments described in context of the method illustrated in figure 4 are analogously valid for the transmission arrangement 500, the method illustrated in figure 6 and the communication device 700 and vice versa.

The components of the transmission arrangement 500 and the communication device (e.g.
30 the encoder, the truncator, the decoder, the transmitter and the receiver) may for example be implemented by one or more circuits. A "circuit" may be understood as any kind of a logic implementing entity, which may be special purpose circuitry or a processor executing software stored in a memory, firmware, or any combination thereof. Thus a "circuit" may

be a hard-wired logic circuit or a programmable logic circuit such as a programmable processor, e.g. a microprocessor (e.g. a Complex Instruction Set Computer (CISC) processor or a Reduced Instruction Set Computer (RISC) processor). A "circuit" may also be a processor executing software, e.g. any kind of computer program, e.g. a computer
5 program using a virtual machine code such as e.g. Java. Any other kind of implementation of the respective functions which will be described in more detail below may also be understood as a "circuit".

In the following, examples are described in greater detail.

10 Figure 8 shows a communication arrangement 800.

The communication arrangement 800 includes a signal source 801, a transmission arrangement 802 which includes an encoder 803, a truncator 804 and a transmitter 805, a
15 communication network 806 and a receiving communication device 807.

The signal source provides a audio signal 808, (or more generally a digital signal representing media content, e.g. an audio signal or a video signal) to the transmission arrangement 802. The transmission arrangement 802 may be implemented using a single
20 device, e.g. a server computer, or a plurality of components.

The transmission arrangement 802 generates an encoded and truncated audio signal 809 from the audio signal 808 and transmits it via the communication network 806 to the receiving communication device 807.

25 In this example, before the actual encoding process, the encoder 803 divides each audio spectral frame into m segments with each one covering a specified frequency range. Then the encoder 803 carries out a scalable encoding process for each segment. For each segment, the encoder 803 packs the produced coding information into a segmental bit-stream in an order of segmental side information, segmental data information (from the
30 most perceptually important information to the least important information). This is illustrated in figure 9.

Figure 9 shows a bit plane representation 901 for an audio signal and a bit-stream 907.

The bit plane representation includes a frequency coefficient for each of a plurality of frequencies (e.g. for each of a plurality of scale factor bands) for one audio frame. The frequency coefficients are grouped according to a segmentation of the whole frequency range into sub-spectrums 902. The frequency coefficients for example represent a residual audio signal (e.g. after core layer encoding).

For each sub-spectrum, a segmental bit-stream 903 is generated by bit-plane scanning the frequency coefficients. Each segmental bit-stream 903 includes side information 904, followed by the bits of the frequency coefficients in the order from most significant bits 905 to least significant bits 906.

The bits of all segmental bit-stream 903, in this example concatenated in the order of the frequency segments, forms the overall (encoded) bit-stream 907.

The final overall bit-stream 907 in this example contains the segmental bit-streams 903 in an order of low frequency segment to high frequency segment.

This bit-stream structure allows frequency scalability and SNR scalability since segmental bit-streams 903 (e.g. high frequency segmental bit-streams) may be omitted (i.e. cut off from the bit stream 907) and segmental bit-streams 903 may be truncated individually at different lengths.

The encoder 803 provides the final overall bit-stream 907 to the transmitter 805 which transmits it over the network 806 to the receiving communication device 807.

In the following, examples for the encoder 803, the truncator 804 and an decoder included in the receiving communication device 807 are described with reference to figures 10, 11, 12 and 13.

Figure 10 shows an encoder 1000.

The encoder 1000 comprises a time/frequency transform block 1001, an optional psychoacoustic analysis block 1002, a spectrum segmentation block 1003, a scalable encoder 1004 and a bit-stream multiplexer 1005. The various components and blocks may be implemented by one or more circuits as described above.

5

Given an audio input 1006 (i.e. a digital audio signal), the transform block 1001 converts the audio input 1006 into its spectral representation, (e.g. using integer MDCT (Modified Discrete Cosine Transform)). The psychoacoustic analysis block 1002 may optionally perform a psychoacoustic analysis on the audio input 1006. For example, under low
10 network bandwidth scenario, to achieve better perceptual quality, a general psychoacoustic model of the signal can be generated which can be used to guide the subsequent scalable encoding process so that the perceptually important information can be encoded with higher priority. However, under other scenarios, the psychoacoustic analysis can be omitted to reduce the encoder complexity.

15

After the transformation by the transform block 1001, the spectrum of the transformed signal is divided into several segments by the spectrum segmentation block 1003 (e.g. into spectrum segments 902 as explained with reference to figure 9), indexed from the lowest frequency of the spectrum of the transformed signal to the highest frequency of the
20 transformed spectrum. The segmentation boundaries could be fixed (e.g. the spectrum may be divided evenly) or adaptive to the psychoacoustic analysis results. For example, the spectrum segmentation may be adaptive to the signal sampling rate via pre-defined lookup tables. Exemplary values for typical 48kHz audio sampling rate is listed in table 1 below.

Segment Index i	Frequency Range w_i (kHz)
1	[0 3.75)
2	[3.75 6.20)
3	[6.20 8.25)
4	[8.25 12.0)
5	[12.0 14.25)
6	[14.25 15.0)
7	[15.0 20.25)
8	[20.25 24.0)

Table 1: An example of segmentation boundaries for 48kHz audio

For the case of adaptive segmentation of the spectrum, the corresponding boundary information is for example provided to the truncator 804 as part of metadata 1007. For example, the truncator 804 and the transmitter 805 are part of a server while the encoder 803 is separate from the server and the encoder 803 provides the metadata 1007 to the server.

According to the segmentation of the spectrum, the transformed digital signal (i.e. the frequency coefficients) are grouped.

These groups of frequency coefficients are independently encoded from each other by the scalable encoder 1004 to generate segmental bit-streams (e.g. corresponding to segmental bit-streams 1003). The encoding process can for example be carried out like the bit-plane coding adopted in SLS, where the spectral coefficients are converted to binary symbols, the latter are sequentially scanned and encoded from the MSB to LSB, from low frequency to high frequency, as long as the coefficient to be encoded is significant for the present coding bit-plane.

Alternatively, the encoding process can be guided by some external information, like the perceptually enhanced bit-plane coding where the spectral coefficients are firstly pre-processed based on the corresponding psychoacoustic information, and then the processed results are coded following SLS bit-plane coding technique.

Lastly, the bit-stream multiplexer 1005 packs the segmental bit-streams in the order of segmental index, and the resulting final overall scalable bit-stream 1008, associated with the possible metadata 1007 (like the segmental boundary information), is sent to the truncator 804 (e.g. to a server including truncator 804 and transmitter 805 and which has received a request for the audio signal from the receiving communication device 807).

As illustrated in figure 9, the information for one audio frame corresponds to m independent segmental bit-streams 903. For each segmental bit-stream 903, side information 904 and perceptually important bits are located in the front of segmental bit-stream 903, and less important bits are located at the rear of segmental bit-stream 903.

Figure 11 shows a truncator 1100.

When the truncator 1100 (or e.g. a server housing the truncator 1100) receives a streaming requirement, according to the available network bandwidth, it can drop (omit) some segments from the scalable bit-stream 1101 received from the encoder 803 (along with meta data 1102) to realize frequency scalability, and truncate any segmental bit-stream to realize SNR scalability. The result is a truncated scalable bit-stream 1103.

The truncation is illustrated in figure 12.

Figure 12 illustrates a truncation process.

As explained above, for each audio frame, the scalable bit-stream 1201 includes m segmental bit-streams 1202. The truncator 1203 for example receives information about bandwidth available in the communication network 806 for communication with the receiving communication device 807 and the metadata 1204 from the encoder 803 and individually truncates the segmental bit-streams 1202 or even omits segmental bit-stream 1202 such that the truncated bit-stream 1205 includes n truncated segmental bit streams 1206.

The (optional) metadata 1204 can for example include the segmentation boundary information if the spectral segmentation is adaptive.

As an example, it is assumed that the bandwidth (i.e. the bit-rate, for example) of the whole input bit-stream 1101 is B_0 , under a given available network bandwidth B such that the truncator decides that only a subset of the segmental bit-streams 1202 is kept, namely those
 5 corresponding to low frequency segments. The number n of the segmental bit-streams 1202 that are kept may for example be set to be proportional to the ratio between B_0 and B , for example according to

$$n = \min \left(\max \left(1, \left\lfloor \frac{B}{B_0} \cdot m \right\rfloor \right), m \right)$$

Accordingly, the truncator 1100 drops the remaining $m - n$ high frequency segments to realize frequency scalability. To realize SNR scalability, the remaining low frequency segments are further truncated, e.g. subject to the constraint that the length summation of truncated segmental bit-streams 1206 should approximately be equal to the available bit
 15 budget for each audio frame. The individual segmental truncation length, for example, can be simply decided based on the relative segmental bandwidth information, e.g. according to the following mathematic expressions:

$$\sum_{i=1}^n (r_i + p_i) \approx \bar{B}$$

$$\frac{r_i + p_i}{r_j + p_j} = w_i / w_j \quad (i, j \in [1, \dots, n])$$

where $\bar{B}$ denotes the average bit budget for each frame, which is proportional to the network bandwidth, frame length, and inversely proportional to the signal sampling frequency; r_i is the truncated bit-stream length of the i^{th} truncated segmental bit-stream 1206 (which may be zero if the bit-stream is omitted); p_i is the corresponding length of
 25 segmental side information and w_i is the corresponding segmental bandwidth, which may be fixed or may be obtained by the truncator from the metadata information 1204.

The truncation information (i.e. the information on how the segmental bit-streams 1202 have been truncated, e.g. the length of the truncated segmental bit-streams 1206) is transmitted to the receiving communication device 807 as additional side information. For example, the segmental side information 904 as illustrated in figure 9 is updated after the
5 truncation. The side information may be directly represented by the truncated lengths r_i , or its corresponding bit-plane scanning numbers for more efficiency. In one embodiment, no matter which representation method is adopted, the length of segmental side information p_i is updated to include the truncation information before the truncation process is carried out.

10 The truncator 804 provides the truncated scalable bit-stream 1103 to the transmitter 805 which transmits it to the receiving communication device 807 via the network 806.

Figure 13 shows a receiver 1300.

15 The receiver 1300 includes a bit-stream multiplexer 1301, a scalable decoder 1302 and a frequency to time transform block 1303 which may be implemented by one or more circuits as described above.

The receiver 1300 receives the truncated scalable bit-stream 1304.

20

The demultiplexer 1301 demultiplexes the truncated scalable bit-stream 1304 into the truncated segmental bit-streams 1206.

The scalable decoder 1302 independently decodes the truncated segmental bit-stream 1206
25 to generate a decoded spectral information (i.e. reconstructed frequency coefficients). The F/T transform block 1303 transforms the decoded spectral information back to the time domain to reconstruct the audio signal in the time domain and outputs the reconstructed audio signal 1305. The frequency-to-time domain transformation of the transform block 1303 is complementary to the time-to-frequency transformation performed by the
30 transform block 1001 on the encoding side.

The overall rate-distortion performances of various coding schemes are illustrated in figure 14.

Figure 14 shows a performance diagram 1400.

5 Bit rate increases from left to right along a bit rate axis 1401 and quality of the reconstructed signals increases from bottom to top along a quality axis 1402.

The streaming quality of a non-scalable coding scheme is shown as a staircase curve 1403 since the server can only store limited copies of non-scalable audio.

10 A dash-dotted curve 1404 illustrates the performance of a traditional scalable coding scheme (such as SLS). This can only reach sub-optimal streaming quality due to the high initial side information (located at the beginning of a scalable bit-stream) and fixed scanning order.

15 The performance of the flexible scalable scheme as described above is illustrated by a solid curve 1405. It can make up for the shortcomings of the traditional scalable coding scheme due to its two dimensional scalable capability.

20 Comparing to the non-scalable scheme, the flexible scalable scheme as described above has fine-grain scalability merit and might be more advantageous at the low bit-rate scenario.

25 Comparing to the traditional scalable coding scheme, the flexible scalable scheme as described above can provide better streaming quality at low, or even at middle, bit-rate. By introducing the frequency scalability, the initial side information is divided into several parts and dispatched to different positions within a scalable bit-stream, so much less side information are left at the beginning of a scalable bit-stream. Apart from it, by introducing the intelligent truncator, a more optimal coding priority order for a given bit-rate can be provided by taking advantage of the trade-off between SNR and frequency scalability.

30 Both two factors lead to streaming quality improvement at low to middle bit-rate. At high bit-rate, due to insignificant difference of side information amount, the flexible scalable scheme as described above can provide almost same high quality as traditional scalable coding scheme.

Various embodiments as for example described above introduce a scalable audio coding scheme that can provide 2-D compression scalability: frequency scalability and SNR scalability. It allows the coding priority order in a bit-stream to be flexibly changed in the truncation process, characterized by (i) Spectral segmentation (ii) Intelligent truncation. The spectral segmentation introduces 2-D compression scalability, and the Intelligent truncation provides an R-D optimized output under a given network bandwidth.

In summary, a scalable audio coding scheme for providing 2-D compression scalability is provided according to an embodiment comprising the steps of (i) transforming the input signal into frequency domain to generate the signal spectrum; (ii) dividing the signal spectrum into several segments; (iii) scalably encoding the spectral information of the signal of each segment to generate a plurality of segmental bit-streams; and (iv) truncating the segmental bit-streams to generate truncated bit-streams.

On the decoding side, the scheme may further comprise scalably decoding the truncated bit-streams to generate decoded spectral information and transforming the decoded spectrum into the time domain.

While specific aspects have been described, it should be understood by those skilled in the art that various changes in form and detail may be made therein without departing from the spirit and scope of the aspects of this disclosure as defined by the appended claims. The scope is thus indicated by the appended claims and all changes which come within the meaning and range of equivalency of the claims are therefore intended to be embraced.

Claims

1. A method for transmitting a digital signal comprising:
generating a representation of the digital signal in the frequency domain, wherein
5 the representation comprises, for each of a plurality of sub-spectrums of the
spectrum of the digital signal, at least one frequency coefficient;
generating, for each sub-spectrum, a bit-stream from the at least one frequency
coefficient of the sub-spectrum;
determining, for each bit-stream, a truncation point;
10 truncate each bit-stream according to the truncation point determined for the bit-
stream;
transmitting data according to the truncated bit-streams.
2. The method according to claim 1, further comprising dividing the spectrum into the
15 plurality of sub-spectrums.
3. The method according to claim 1 or 2, wherein the digital signal is a media signal.
4. The method according to any one of claims 1 to 3, wherein the digital signal is an
20 audio signal or a video signal.
5. The method according to any one of claims 1 to 4, wherein generating a bit-stream
from the at least one frequency coefficient of the sub-spectrum includes bit-plane
scanning the at least one frequency coefficient of the sub-spectrum.
25
6. The method according to any one of claims 1 to 5, further comprising multiplexing
the truncated bit-streams to an overall bit-stream wherein transmitting data
according to the truncated bit-streams includes transmitting the overall bit-stream.
- 30 7. The method according to any one of claims 1 to 6, further comprising transmitting,
for each truncated bit-stream, information about the truncation point.

8. The method according to any one of claims 1 to 7, wherein determining the truncation point includes determining a truncation length and truncating the bit-stream according to the truncation point includes truncating the bit-stream according to the truncation length.

5

9. The method according to any one of claims 1 to 8, wherein the data are transmitted to a communication device via a communication channel and the truncation point is determined based on the capacity of the communication channel.

10

10. The method according to any one of claims 1 to 9, wherein determining a truncation point for a bit-stream includes deciding whether the bit-stream should be omitted from the data and, if it is decided that the bit-stream should be omitted, setting the truncation point such that truncating the bit-stream according to the truncation point leads to the bit-stream being omitted from the data.

15

11. A transmission arrangement comprising
an encoder configured to generate a representation of the digital signal in the frequency domain, wherein the representation comprises, for each of a plurality of sub-spectrums of the spectrum of the digital signal, at least one frequency
coefficient and to generate , for each sub-spectrum, a bit-stream from the at least one frequency coefficient of the sub-spectrum;
a truncator configured to determine, for each bit-stream, a truncation point and to truncate each bit-stream according to the truncation point determined for the bit-stream; and
a transmitter configured to transmit data according to the truncated bit-streams.

20

25

12. A method for receiving a digital signal comprising:
receiving data including a plurality of bit-streams;
generating, from each bit-stream, at least one frequency coefficient;
aggregating the frequency coefficients to a frequency domain representation of a digital signal; and
transforming the frequency domain representation to the time domain to reconstruct the digital signal.

30

13. The method according to claim 12, wherein the data includes an overall bit-stream and the method comprises demultiplexing the overall bit-stream to generate the plurality of bit-streams.

5

14. The method according to claim 13, wherein the data includes, for each bit-stream, information about a truncation point of the bit-stream and demultiplexing the overall bit-stream is based on the information about the truncation points.

10

15. A communication device comprising:
a receiver configured to receive data including a plurality of bit-streams; and
a decoder configured to generate, from each bit-stream, at least one frequency coefficient, aggregate the frequency coefficients to a frequency domain representation of a digital signal and transform the frequency domain representation to the time domain to reconstruct the digital signal.

15

1 / 14

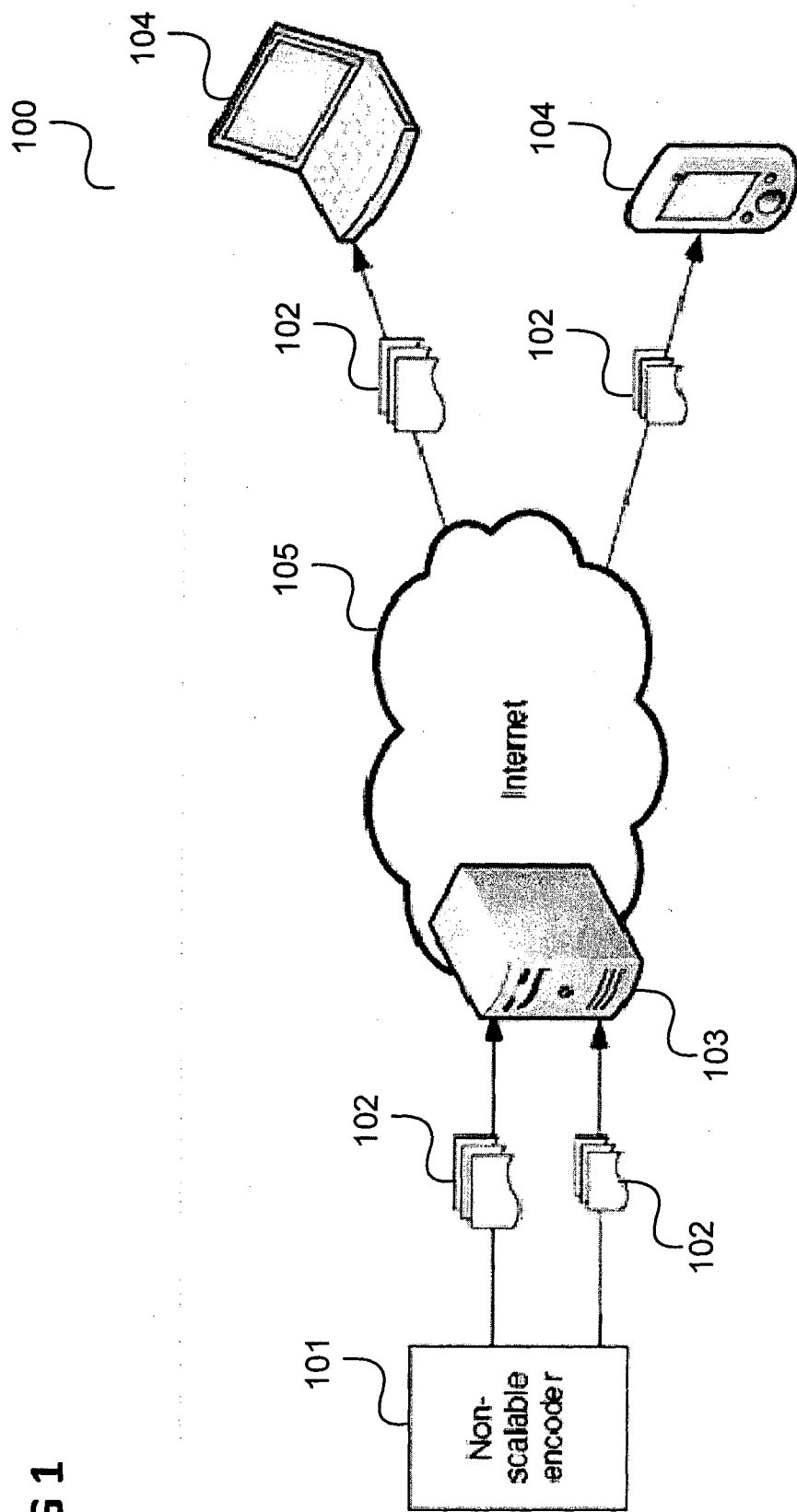
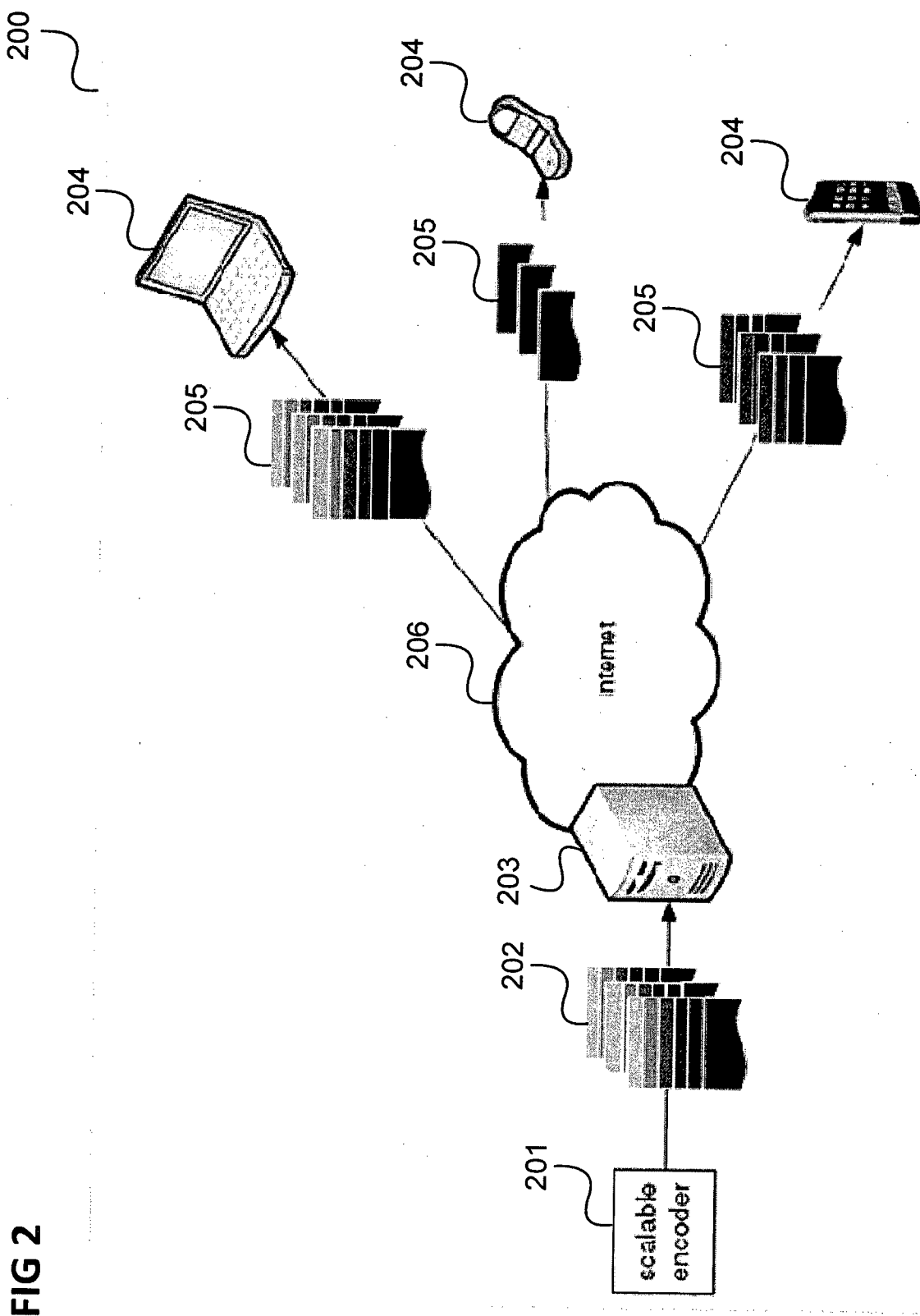


FIG 1

2/14



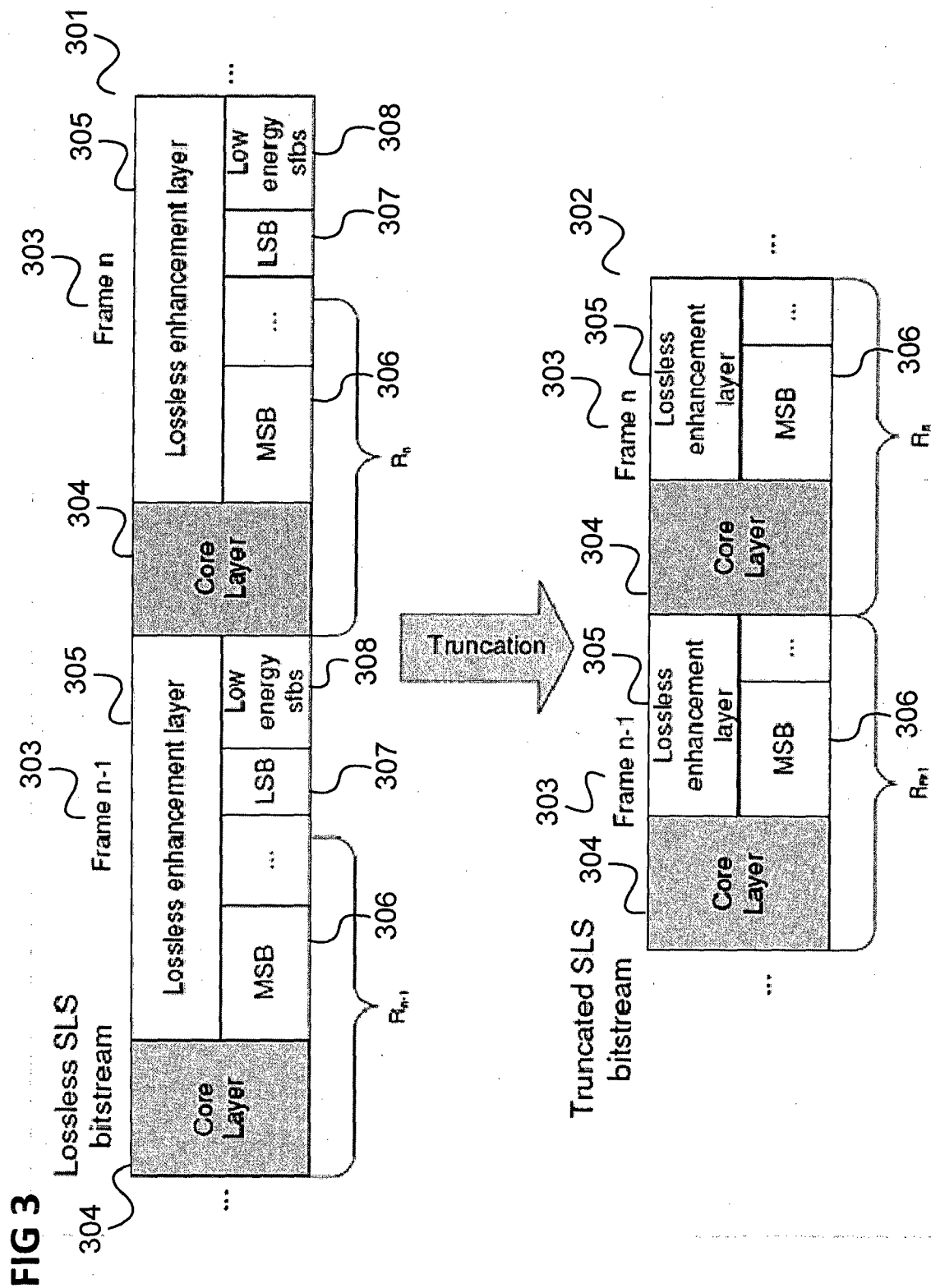


FIG 4

4/14

400

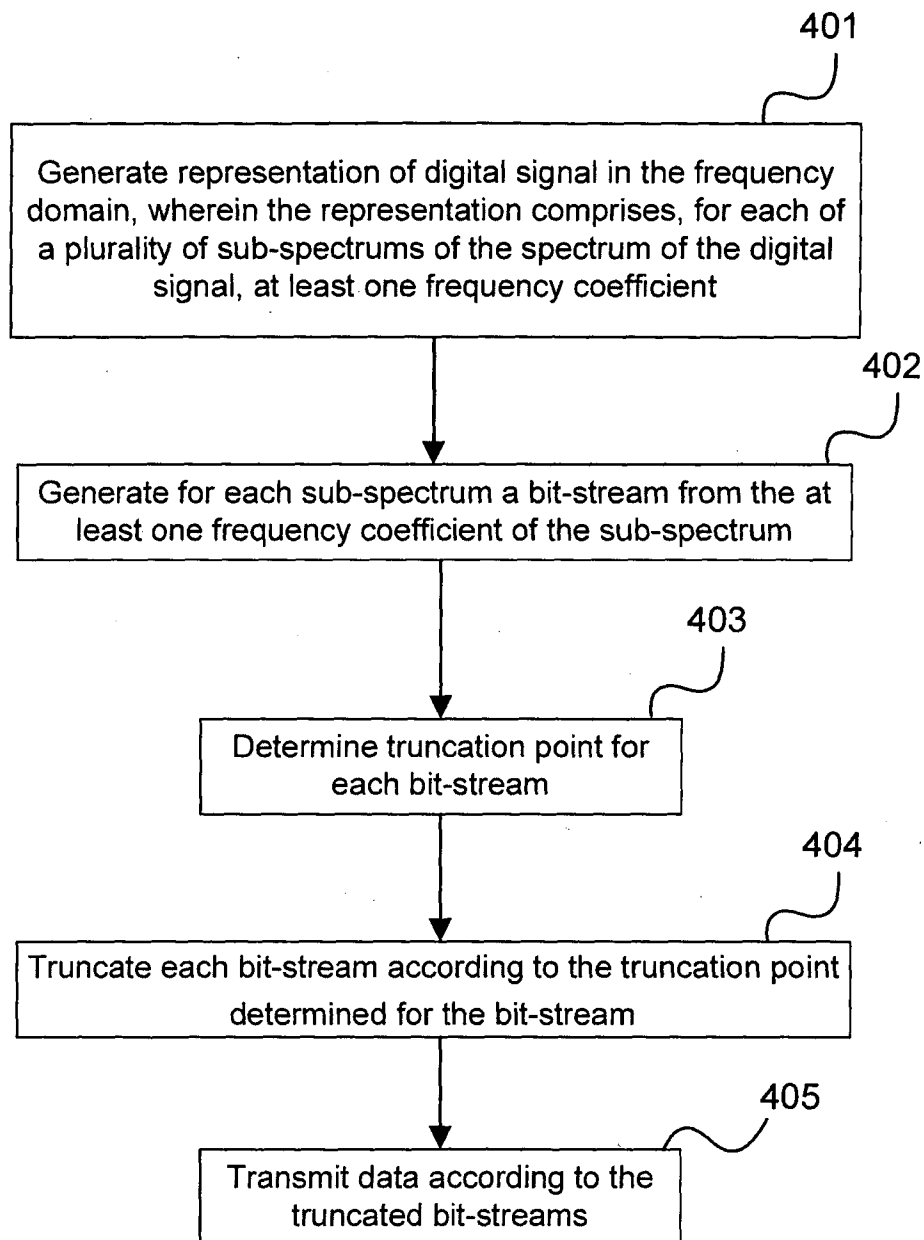


FIG 5

500

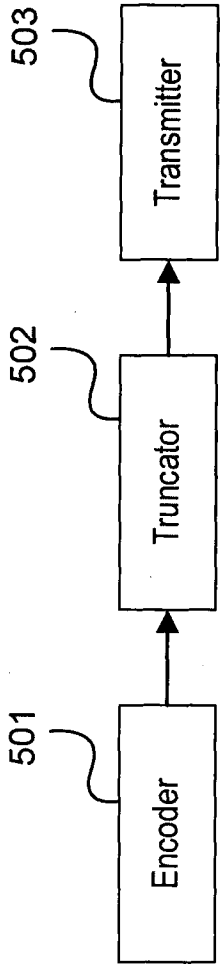


FIG 6

6/14

600

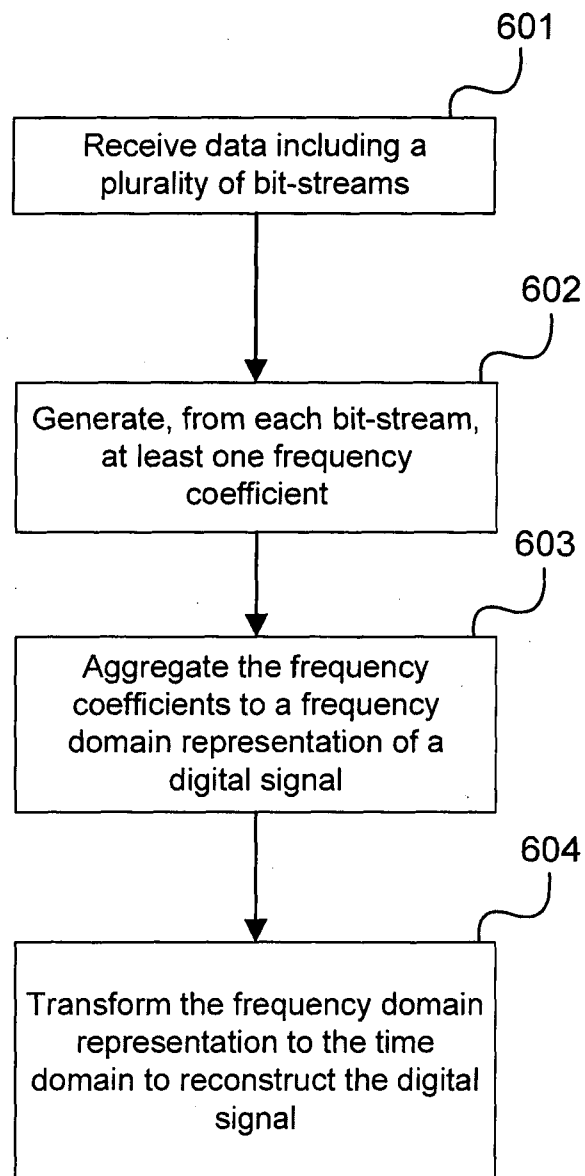
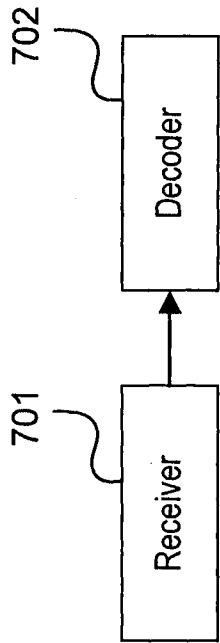
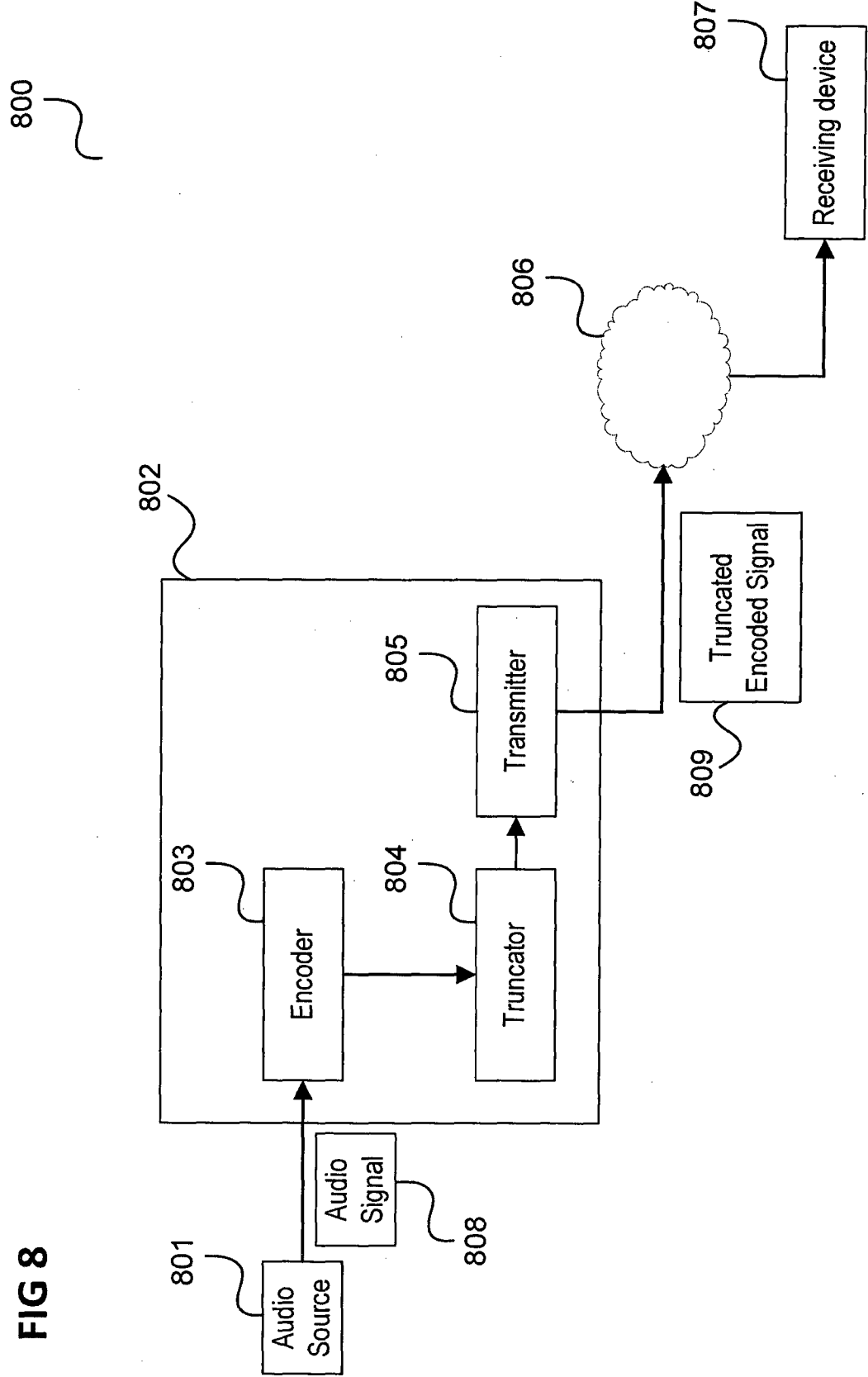
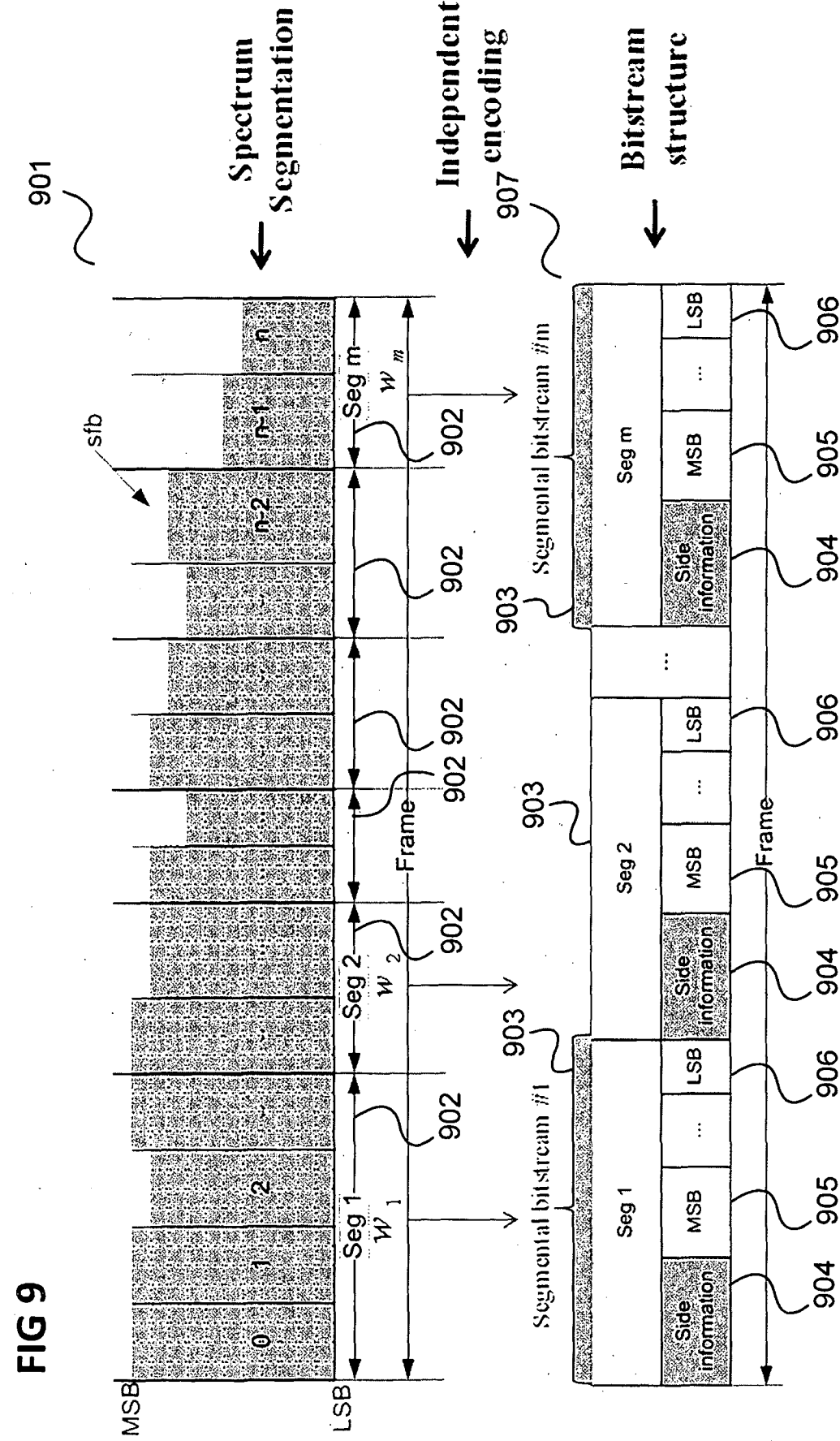


FIG 7

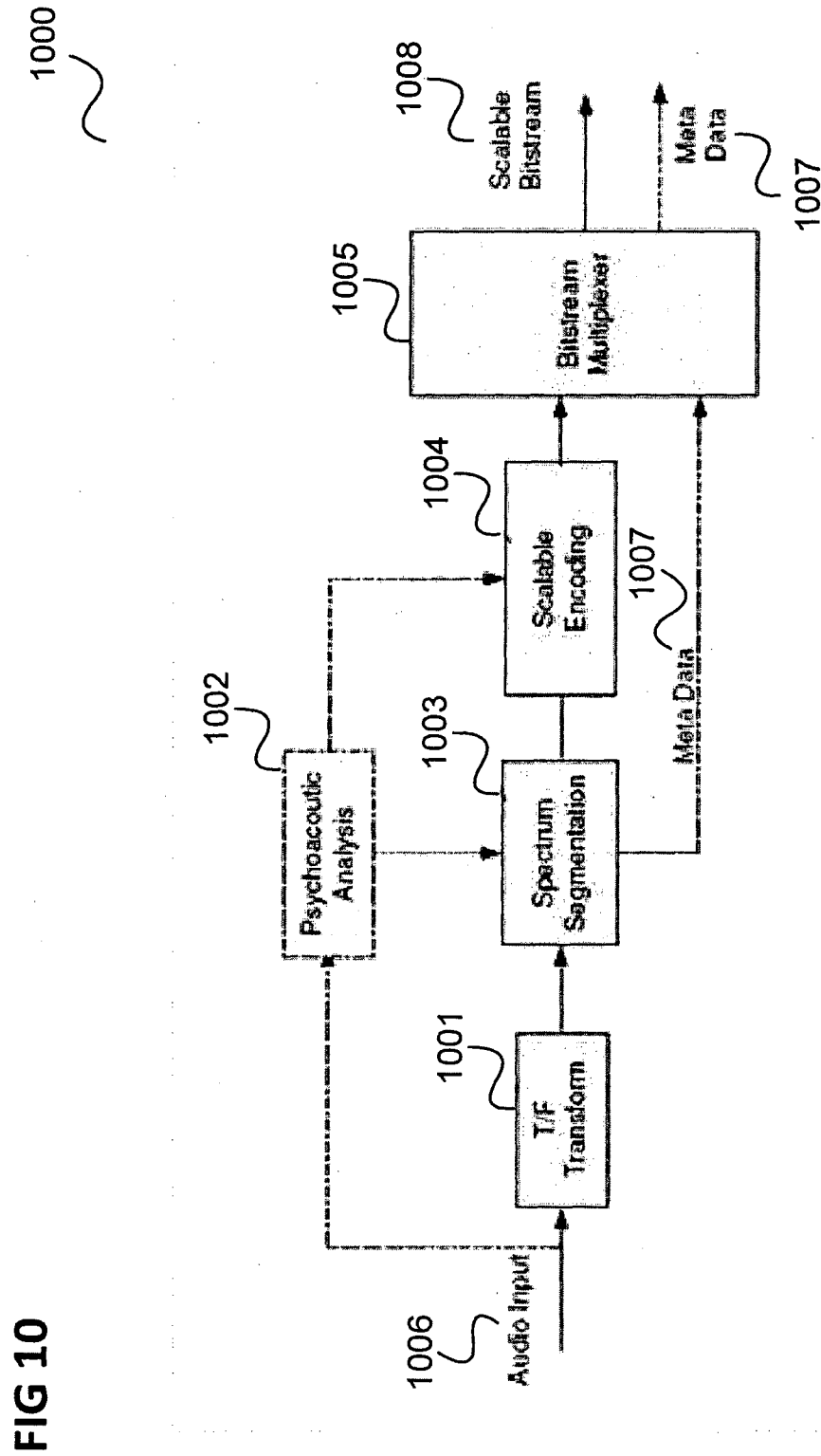
700







10/14



11/14

FIG 11

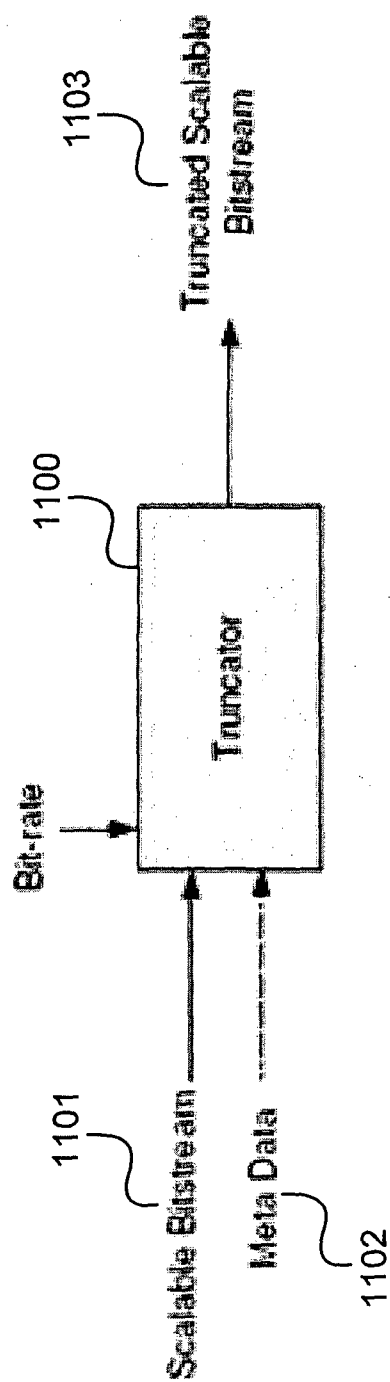


FIG 12

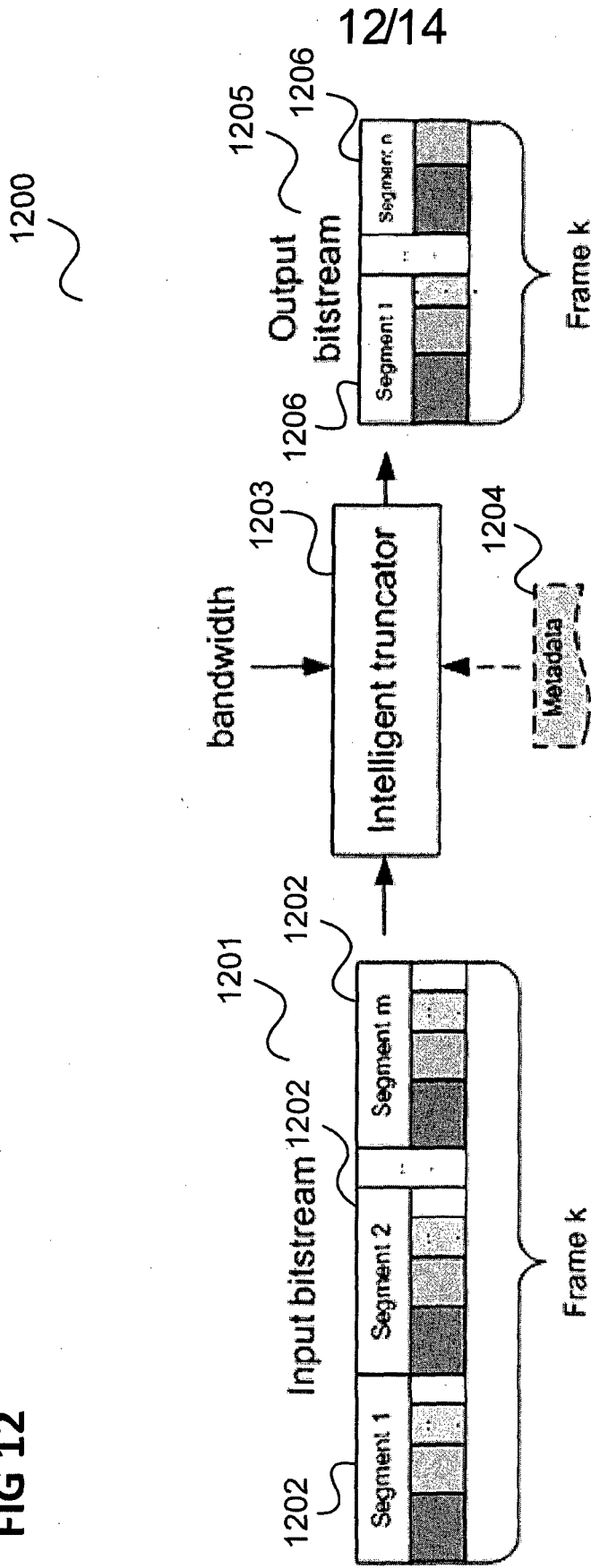
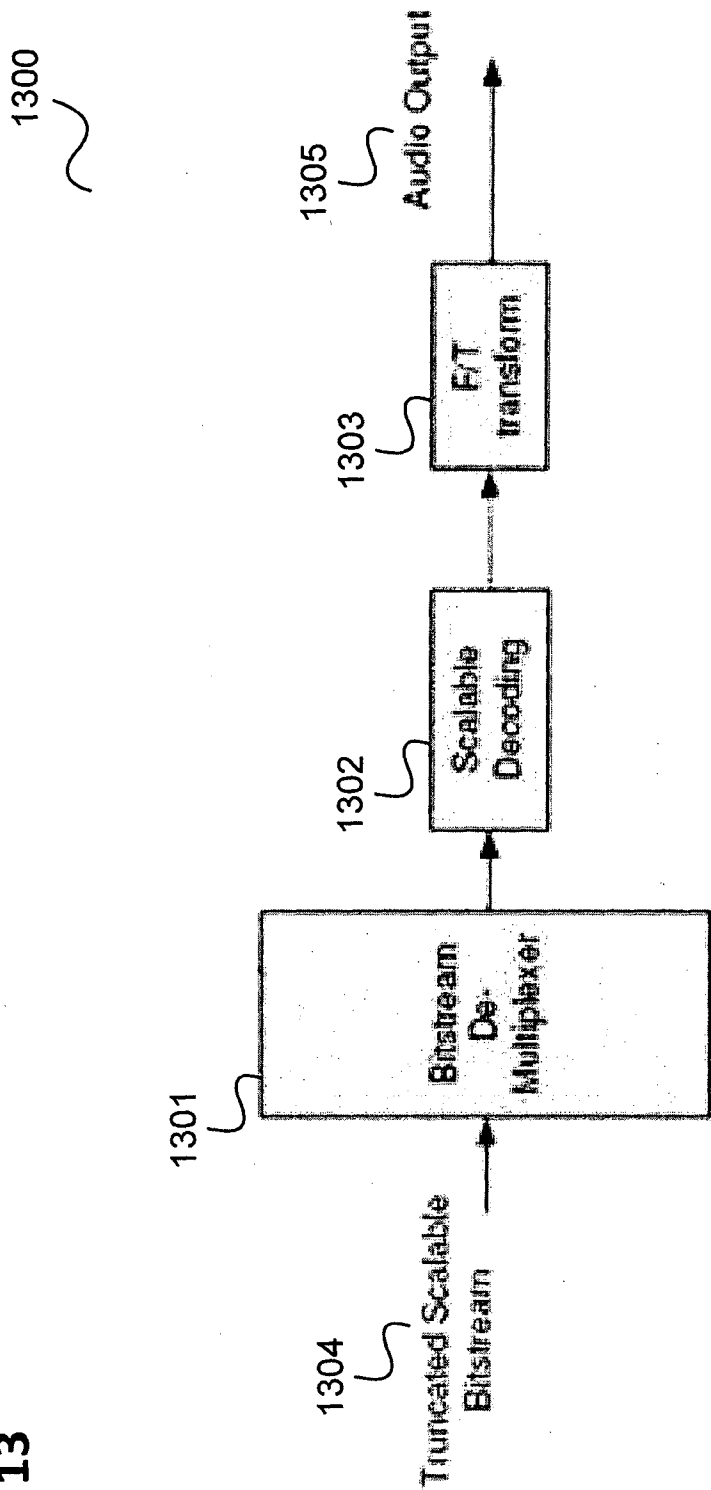
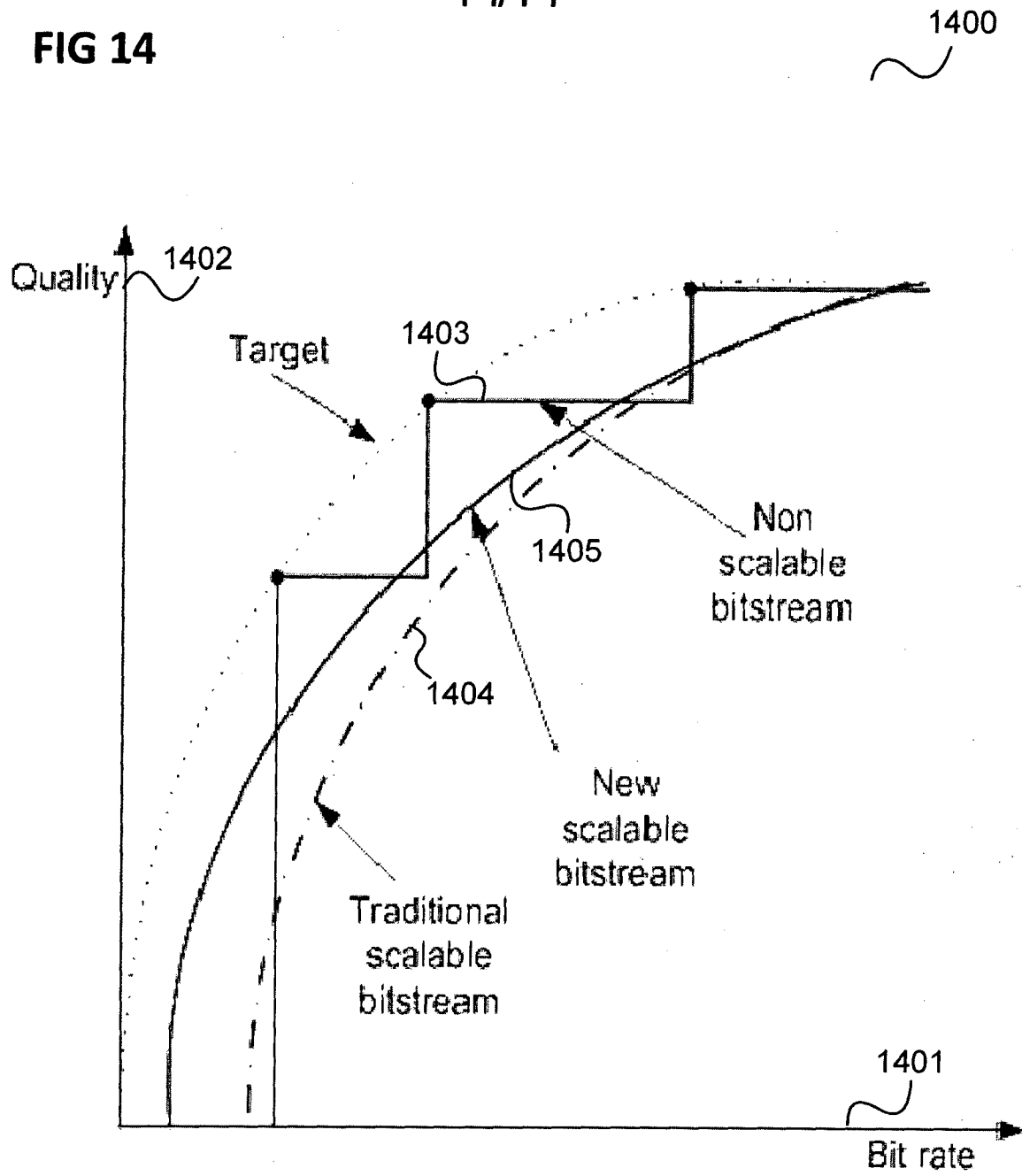


FIG 13



14/14

FIG 14



INTERNATIONAL SEARCH REPORT

International application No.
PCT/SG2013/000124

A. CLASSIFICATION OF SUBJECT MATTER

H04N 21/2662 (2011.01) H04B 7/005 (2006.01) H04W 72/04 (2009.01)

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)

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)

EPODOC, WPI, INSPEC with keywords: digital signal, scalable compression, bit stream and similar terms.

Google Patents and Google Scholar searched with similar terms and in addition using truncate, and sub spectrum.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	Documents are listed in the continuation of Box C	



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	"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
"E" earlier application or patent but published on or after the international filing date	"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
"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)	"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
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search
4 June 2013

Date of mailing of the international search report
04 June 2013

Name and mailing address of the ISA/AU

AUSTRALIAN PATENT OFFICE
PO BOX 200, WODEN ACT 2606, AUSTRALIA
Email address: pct@ipaustalia.gov.au
Facsimile No.: +61 2 6283 7999

Authorised officer

Dr. Rasika Perera
AUSTRALIAN PATENT OFFICE
(ISO 9001 Quality Certified Service)
Telephone No. 0262833116

INTERNATIONAL SEARCH REPORT C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		International application No. PCT/SG2013/000124
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2000/049570 A1 (UNISEARCH LIMITED et al.) 24 August 2000 See the whole document in particular the abstract, figure 1, claims 1, 2, 15-16, page 3 lines 23-27, page 4 line 26 to page 5 line 2, page 8 line 11 to page 9 line 13, page 10 line 32 to page 11 line 23.	1-15
X	EP 1730725 B1 (MICROSOFT CORPORATION) 09 December 2009 See the whole document in particular the abstract, figures 1-6, claim 1, paragraphs 0001, 0009, 0011, 0018-0020, 0025-0034, 0037-0040, and 0043-0053.	1-15

Form PCT/ISA/210 (fifth sheet) (July 2009)

INTERNATIONAL SEARCH REPORT Information on patent family members		International application No. PCT/SG2013/000124	
This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.			
Patent Document/s Cited in Search Report		Patent Family Member/s	
Publication Number	Publication Date	Publication Number	Publication Date
WO 2000/049570 A1	24 Aug 2000	AU 776355 B2	02 Sep 2004
		AU 2784200 A	04 Sep 2000
		US 6760482 B1	06 Jul 2004
		WO 0049570 A1	24 Aug 2000
EP 1730725 B1	09 Dec 2009	CN 1813286 A	02 Aug 2006
		CN 1813286 B	24 Nov 2010
		EP 1730725 A1	13 Dec 2006
		EP 1730725 B1	09 Dec 2009
		JP 2007532934 A	15 Nov 2007
		JP 4745986 B2	10 Aug 2011
		JP 2011186479 A	22 Sep 2011
		KR 20110042137 A	22 Apr 2011
		KR 101130355 B1	27 Mar 2012
		KR 20060121655 A	29 Nov 2006
		KR 20110093953 A	18 Aug 2011
		US 2005165611 A1	28 Jul 2005
		US 7460990 B2	02 Dec 2008
		US 2009083046 A1	26 Mar 2009
		WO 2005076260 A1	18 Aug 2005
End of Annex			
Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001. Form PCT/ISA/210 (Family Annex)(July 2009)			