



(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
16.02.2011 Bulletin 2011/07

(51) Int Cl.:
H04H 20/88 (2008.01) G10L 19/00 (2006.01)

(21) Application number: **09305680.2**

(22) Date of filing: **16.07.2009**

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO SE SI SK SM TR
Designated Extension States:
AL BA RS

- **Fadili, Moulay**
95000 Cergy (FR)
- **Moulehiawy, Abdelkrim**
75002 Paris (FR)

(74) Representative: **Hervouet, Sylvie et al**
Feray Lenne Conseil
Le Centralis
63, avenue du Général Leclerc
92340 Bourg-la-Reine (FR)

(71) Applicant: **Alcatel Lucent**
75008 Paris (FR)

(72) Inventors:
• **Zrihen, Jeremy**
95280 Jouy le Moutier (FR)

(54) **Method and apparatus for coding/decoding a stereo audio signal into a mono audio signal**

(57) A method for coding a stereo audio signal, composed of its two components comprising a first input audio signal (1) and a second input audio signal (2), into a single composite audio signal (3), comprising the steps of:
- averaging said first input audio signal (1) and said second input audio signal (2) into an averaged audio signal (4, 9),
- determining a first coding information (7) by comparing

said averaged audio signal (4, 9) to said first input audio signal (1),
- determining a second coding information (8) by comparing said averaged audio signal (4, 9) to said second input audio signal (2),
- embedding said first and second coding information (7, 8) into said averaged audio signal (4, 9) to obtain the composite audio signal (3).

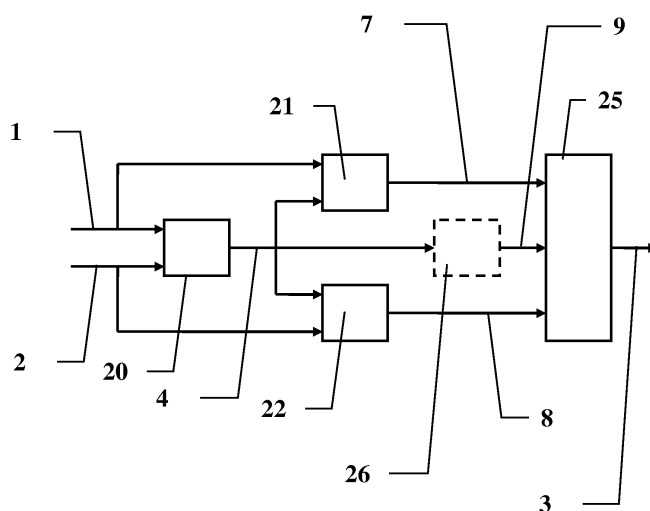


FIG. 1

Description

[0001] The technical domain of the invention is the domain of transmission of audio signal, and particularly stereo audio signal.

[0002] More particularly, the invention concerns a method to code and decode a stereo audio signal, comprising two separate audio signals into a single composite audio signal, in order to transmit it over a single link designed to transmit a mono audio signal.

[0003] During a communication through a communication system e.g. phone communication, be it RTC, TDM or IP, an audio signal, such as voice, music or any other sound, is generally transmitted in mono, over a single audio link between users. This leads to a loss of a real spatial dimension in which each user is moving and can be a source of involuntary weariness. Furthermore, a stereo audio signal can be a really interesting feature to enhance user hearing and for services such as e.g. vocal message, voice guide, music on hold, conferencing system and so on.

[0004] However, a stereo audio signal comprises two audio signals, and would then need twice the amount of capacities to be transmitted. Moreover, most of the existing systems are designed to manage and to transmit only single mono signals.

[0005] The idea of the invention is then to use an existing mono audio link to transmit a stereo audio signal. This needs a special coding of the stereo audio signal, composed of a first input audio signal and a second input audio signal, into a single composite audio signal.

[0006] The invention thus concerns a method for coding a stereo audio signal, composed of its two components comprising a first input audio signal and a second input audio signal, into a single composite audio signal, comprising the steps of:

- averaging said first input audio signal and said second input audio signal into an averaged audio signal,
- determining a first coding information by comparing said averaged audio signal to said first input audio signal,
- determining a second coding information by comparing said averaged audio signal to said second input audio signal,
- embedding said first and second coding information into said averaged audio signal to obtain a composite audio signal.

[0007] According to another feature of the invention, an adaptative filter may be used to determine said first and second coding information.

[0008] According to another feature of the invention, the signal may be digitalised applying G.711 codec and arranged into bytes using HDLC protocol, and the first and second coding information may be embedded by replacing the less significant bit of some of said bytes.

[0009] The invention also concerns a corresponding

method for decoding a stereo audio signal, from a single composite audio signal, into its two components comprising a first output audio signal and a second output audio signal, comprising the steps of:

- separating first and second coding information embedded into said composite audio signal out of said composite audio signal, to extract said first and second coding information and a central audio signal,
- synthesising said first output audio signal out of said first coding information and said central audio signal,
- synthesising said second output audio signal out of said second coding information and said central audio signal.

[0010] The invention also concerns a transmitter able to code and transmit such a stereo audio signal.

[0011] The invention also concerns a receiver able to receive and decode such a stereo audio signal.

[0012] The invention also concerns a transmitting system comprising said transmitter and said receiver.

[0013] The proposed solution advantageously offers the capability to transmit a stereo signal over a mono audio transmitting link.

[0014] Another advantage of the proposed solution is to offer transparency in that a classical mono receiver can receive said stereo signal coded as a composite audio signal and still proceed to retrieve an audio signal.

[0015] Others features, details and advantages of the invention will become more apparent from the detailed illustrating description given hereafter with respect to the drawings on which:

- figure 1 is a synoptic of the coding/emitting stage,
- figure 2 is a synoptic of the receiving/decoding stage,
- figure 3 is a synoptic detailing a particular embodiment of a part of the coding/emitting stage,
- figure 4 is a synoptic detailing a particular embodiment of a part of the receiving/decoding stage.

[0016] According to figure 1, presenting a synoptic of the coding stage, a stereo audio signal, initially composed of its two components comprising a first input audio signal 1 and a second input audio signal 2, is coded into a single composite audio signal 3. Said composite audio signal 3 is tailored to be transmitted over an audio link or channel able to transmit a mono audio signal. Said coding can be made by following several steps.

[0017] In a first step, the two input audio signals 1, 2 are averaged into an averaged audio signal 4. The averaged audio signal 4 is preferably an audio signal of the same kind as the two input audio signals 1, 2. For instance it is preferably an analogical signal if the input signals 1, 2 are analogical signals, and it is preferably a digital signal if the input signals 1, 2 are digital signals. It also preferably shares with the input signals 1, 2, the same bandwidth or coding scheme. Said averaged audio signal 4 may be obtained by adding the two signals 1, 2, and dividing the

result by two. Said averaging operation realised by averaging module 20 can be done analogically or digitally.

[0018] In a second step, each initial input audio signal 1, 2, is compared to said averaged audio signal 4. A first comparator/analyser 21 compares said averaged audio signal 4 to said first input audio signal 1. From that comparison it determines a first coding information 7, representative of the first input audio signal 1 or of the difference signal 5 obtained by subtracting the averaged signal 4 out of the first input audio signal 1. A second comparator/analyser 22 similarly, compares said averaged audio signal 4 to said second input audio signal 2. From that comparison it determines a second coding information 8, representative of the second input audio signal 2 or of the difference signal 6 obtained by subtracting the averaged signal 4 out of the second input audio signal 2.

[0019] Each of said first and second coding information 7, 8, may be either analogical or digital. However it is advantageously in a form that takes less occupancy/bandwidth/size, than the initial input audio signal 1, 2, in order to ease the following embedding step, but nonetheless containing the essential features of each respective original input audio signal 1, 2, so as to be able to synthesise said original signals out of said coding information 7, 8, during a further decoding stage. The comparing/analysing operations can be done either analogically or digitally.

[0020] In a third step, the now reduced coding information 7, 8, respectively representative of the input audio signals 1, 2, or of the two difference signals 5, 6, are embedded into the averaged audio signal 4. Said operation is done by an embedding module 25. The resulting signal 3 is the composite audio signal 3. Said composite audio signal 3 is preferably an audio signal of the same kind as the difference signals 5, 6 or the averaged audio signal 4. It can be either an analogical signal or a digital signal. Depending on the ingress signals 4, 7, 8, the embedding operation can be done either analogically or digitally.

[0021] Many ways can be used to analyse said input audio signals 1, 2, or said difference signals 5, 6, into first and second coding information 7, 8.

[0022] One preferred embodiment advantageously uses adaptative filtering and will be described with reference to figure 3, showing a zoomed view of the comparator/analyser modules 21 and 22 of figure 1. An upper first comparator/analyser 21 determines a first coding information 7 as output. It receives as input the first input audio signal 1 and the averaged audio signal 4. A first comparator 40 determines a first difference signal 5 by subtracting said first input audio signal 1 out of said averaged audio signal 4. Said first difference signal 5 is feedback into a first adaptative filter 42 to adapt its filtering. Said first adaptative filter 42 is used to filter the averaged audio signal 4. Said first filter 42 is self adapted to reach its optimum value to get its output signal very close to the first input audio signal 1, in order to get the first difference signal 5 very close to zero. The signal 7

obtained at output of the first comparator 40 is used as the first coding information 7.

[0023] Similarly, a lower second comparator/analyser 22 determines a second coding information 8 as output. It receives as input the second input audio signal 2 and the averaged audio signal 4. A second comparator determines a second difference signal 6 by subtracting said second input audio signal 2 out of said averaged audio signal 4. Said second difference signal 6 is feedback into a second similar adaptative filter 43 to adapt its filtering. Said second adaptative filter 43 is used to filter the averaged audio signal 4. Said second filter 43 is self adapted to reach its optimum value to get its output signal very close to the second input audio signal 2, in order to get the second difference signal 6 very close to zero. The signal 8 obtained at output of the second comparator is used as the second coding information 8. In a digital embodiment, said adaptative filters 42, 43, are advantageously initialised to zero at start-up. Said adaptative filters 42, 43, may be based either on least mean square, LMS, family or on recursive least square, RLS, family kind of adaptative filters.

[0024] Said first comparator/analyser module 21, respectively second comparator/analyser module 22, may also comprises, between the output of said comparator 40 and the branching of a feedback line 44, an additional quantifier 41. The aim of said quantifier 41 is to drastically reduce the occupancy of its output signal with respect to its input signal, by compressing the first difference signal 5, respectively the second difference signal 6, into the first coding information 7, respectively the second coding information 8. This may be done by analogically or digitally compressing and/or digitalising said signal.

[0025] In a digital embodiment, said quantifier 41 is used to reduce the number of bits used to code the coding signal 7, 8. This is possible because the signal should be close to zero. The coding signal 7, 8, may be coded using very few bits. The convergence time of the adaptative filters 42, 43, depends on the number of bits used to code the coding signal 7, 8. 2x40 bits can then be enough to code the coding signals 7, 8, of two channels of a stereo audio signal, as shown in a later described digital embodiment.

[0026] Another possible embodiment of the determining/analysing step comprises spectral analysis of the difference signal 5, 6, thus producing a decomposition over a given set of frequencies. Since said frequencies are given or predefined in accordance with the decoding stage, the coefficients of said decomposition may be sufficient to wholly define the difference signal 5, 6. Said coefficients may be retained as the coding information 7, 8.

[0027] The first and second coding information 7, 8, whatever their shape or kind, are then embedded into the averaged audio signal 4, by replacing a removed part of said averaged audio signal 4. Said removed or replaced part of the averaged audio signal 4 is advantageously chosen among a least significant part of said

averaged audio signal 4.

[0028] Due to that feature, the removed, and modified by the insertion of coding information 7, 8, part of the averaged signal 4, thus lacking at the decoding stage, would not impact to much the signal when decoded and rendered.

[0029] Another important advantage of said feature is that the eventual decoding of a so coded signal by a classical, e.g. existing prior art mono, receiver remains possible without too much distorting the audio signal. This provides a transparency of the method which allows a progressive deployment of the invention, where the old existing mono receivers can remain in service.

[0030] The embedding step can be done analogically by including a signal coding said first and second coding information 7, 8, into the averaged signal 4. In the analogical case, the least significant part of the signal may be e.g. a frequency not used by the signal, such as a non audible frequency, however contained in the bandwidth of the transmitting/receiving link.

[0031] It is important to notice that the invention may be carried over either analogically or digitally. When referring to figure 1, digitalisation of signal can be done as early as at input stage, applying to input audio signals 1, 2, or as late as at output stage, applying to composite audio signal 3, or at any stage in between. For instance, on figure 1, the digitalisation is applied by a digitalisation module 26 on an averaged audio signal 4 to produce a digitalised averaged audio signal 9.

[0032] The embedding step can also be done digitally by including some bits coding the first and second coding information 7, 8, into the bits coding the averaged signal 9. In the digital case, the least significant part of the signal 9 is some of the least significant bits.

[0033] According to a preferred embodiment of the invention, the digitalisation is made using G.711 codec, and before the transmission the digitalised signal 9 is arranged according to HDLC protocol into bytes. Since the G.711 codec uses a rather large bandwidth, the least significant bit of each of said bytes is quite not relevant to the so coded audio signal. In a HDLC frame, said least significant bits, may then be considered a least significant part of the signal and may be used to code the first and second coding information 7, 8. The resulting composite audio signal 3 comprising the averaged audio signal 9 and so embedded first and second coding information 7, 8, is thus not modified in its size/type and can be transmitted as it previously used to be over one single mono transmitting link. Moreover, it is quite unaffected audio-phonically when decoded and rendered.

[0034] For instance, a 20 ms G.711 frame comprises a payload of 160 bytes. This could provide at most 160 least significant bits that could be used to embed first and second coding information 7, 8. When using HDLC protocol, this number decreases to 126 due to protocol bytes. This is rather enough to embed first and second coding information 7, 8, comprising 8 coefficients with 5 bits of precision each for two difference signals 5, 6, that is:

$8 \times 5 \times 2 = 80$ bits.

[0035] Said preferred embodiment is particularly advantageous since the G.711 codec is used for public switched telephone network, PSTN, for integrated services data network, ISDN and also for voice transport over IP networks with little compression. The invention may then directly be applied to these applications.

[0036] Said coding scheme may obviously be implemented in a transmitter for transmitting said stereo audio signal, before a transmitting module that only needs to be able to transmit a single mono audio signal, since the composite audio signal 3 presents all the characteristics of a mono audio signal.

[0037] Said coding stage may be advantageously used in conjunction with a corresponding decoding stage described now with reference to figure 2.

[0038] When receiving a so coded composite audio signal 3, the aim is to retrieve the two components of the initial stereo audio signal. This can be done applying the following steps.

[0039] In a first step, the first and second coding information 7, 8, embedded into the composite audio signal 3 are separated out of said composite audio signal 3. Said separation step, which may be realised by a separating module 30, extracts first and second coding information 7, 8, on one side, and a central audio signal 10, 11, on the other side. The method used here to separate depends on the corresponding method used to embed, and is its corresponding reverse method. Said central audio signal 10, 11, generally compares to the averaged audio signal 4, 9, of the coding stage.

[0040] In a second step, e.g. respectively realized by a first synthesising module 34 and a second synthesising module 35, a first output audio signal 14 and a second output audio signal 15 are synthesised respectively out of first and second coding information 7, 8. Said synthesising step is made applying a process reverse from the one uses in the comparing/analysing step of the coding stage.

[0041] The thus obtained first output audio signal 14 and second output audio signal 15, normally compare respectively to the first input audio signal 1 and to the second input audio signal 2 inputted in the coding stage.

[0042] As noted for the coding stage, all steps of the decoding stage may indifferently apply to either analogical signals of digital signals, and may be done either analogically or digitally.

[0043] For instance, figure 2 illustrates a digital to analogical converter 31 for converting a digital central audio signal 10 into an analogical central audio signal 11.

[0044] Obviously, the steps of the decoding stage must correspond to those of the coding stage. If first and second coding information 7, 8, embedded into the composite audio signal 3, were coefficients coded from adaptive filtering, said coefficients are retrieved during the separating step of the decoding stage, and are then used accordingly, in synthesise step, to synthesise the two output audio signals 14, 15, out of said coefficients 7, 8. The

synthesising step corresponding to the analysing step of the coding process described with reference to figure 3, will now be described with reference to figure 4.

[0045] Figure 4 shows in detail the synthesising modules 34 and 35 of figure 2, in an embodiment adapted to decode the first and second coding information 7, 8, when said first and second coding information 7, 8, were obtained, at coding stage, through adaptative filtering, as previously described.

[0046] An upper first synthesiser 34 determines a first output audio signal 14 as output. It receives as input the first coding information 7 and the central audio signal 10, 11. A first adaptative filter 45 filters said central audio signal 10, 11, the first adaptative filter 45 being adapted by said first coding information 7. Since the coding process produced said first coding information 7 as previously described, said filtering adapted by said first coding information 7, synthesises at its output a first output audio signal 14. At start-up the adaptative filter 45 is initialised to zero. Since it uses the same central audio signal 10, 11 as input and the same signal 7 for adaptation, the resulting output signal 14 compares to the first input audio signal 1, when convergence of filters are good.

[0047] Similarly, a lower second synthesiser 35 determines a second output audio signal 15 as output. It receives as input the second coding information 8 and the central audio signal 10, 11. A second adaptative filter 46 filters said central audio signal 10, 11, the second adaptative filter 46 being adapted by said second coding information 8. Since the coding process produced said second coding information 8 as previously described, said filtering adapted by said second coding information 8, synthesises at its output a second output audio signal 15. At start-up the adaptative filter 46 is initialised to zero. Since it uses the same signal 10, 11 as input and the same signal 8 for adaptation, the resulting output signal 15 compares to second input audio signal 2, when convergence of filters are good.

[0048] Said adaptative filters 45, 46, may be based either on LMS family or on RLS family kind of adaptative filters.

[0049] It will be apparent to those skilled in the art, that the process would benefit of said first adaptative filter 45 used for decoding being identical to the first adaptative filter 42 used to code the first input audio signal 1 into said composite audio signal 3, and of said second adaptative filter 46 used for decoding being identical to the second adaptative filter 43 used to code the second input audio signal 2 into said composite audio signal 3.

[0050] All of said previously described decoding schemes may obviously be implemented in a receiver of said stereo audio signal, after a receiving module, able to receive the composite audio signal 3, as a single mono audio signal, decoding it so as to provide the first output audio signal 14 and the second output audio signal 15, components of said stereo audio signal.

[0051] Said two output audio signals 14, 15, may then be audiophonically rendered by any known stereo player

such as a pair of loudspeakers or a stereo headset generally reproducing the initial stereo signal composed of the two input audio signals 1, 2. In fact, the only differences are only caused by the process losses.

Claims

1. A method for coding a stereo audio signal, composed of its two components comprising a first input audio signal (1) and a second input audio signal (2), into a single composite audio signal (3), **characterised in that** it comprises the steps of:

- averaging said first input audio signal (1) and said second input audio signal (2) into an averaged audio signal (4, 9),
- determining a first coding information (7) by comparing said averaged audio signal (4, 9) to said first input audio signal (1),
- determining a second coding information (8) by comparing said averaged audio signal (4, 9) to said second input audio signal (2),
- embedding said first and second coding information (7, 8) into said averaged audio signal (4, 9) to obtain the composite audio signal (3).

2. The method of claim 1, wherein said determining a first coding information (7) step adaptatively filters said averaged audio signal (4, 9) through a first adaptative filter (42) in order to zeroing a first difference signal (5), obtained by subtracting said first input audio signal (1) out of said averaged audio signal (4, 9), the first coding information (7) being said first difference signal (5), and wherein said determining a second coding information (8) step adaptatively filters said averaged audio signal (4, 9) through a second adaptative filter (43) in order to zeroing a second difference signal (5), obtained by subtracting said second input audio signal (2) out of said averaged audio signal (4, 9), the second coding information (8) being said second difference signal (8).

3. The method of claim 2, wherein each determining step respectively further comprises:

- quantifying said first (5), respectively second (6), difference signal in order to compress said difference signal (5, 6), the first (7), respectively second, coding information (8), being said first (5), respectively second, difference signal (6) compressed.

4. The method of any one of claims 1 to 3, wherein said embedding step comprises replacing into said averaged audio signal (4, 9) a least significant part of the signal by said first and second coding information (7, 8).

5. The method of claim 4, wherein the averaged audio signal (4) is digitalised, and wherein said least significant part of the signal comprises some less significant bits of said digitalised signal (9).
6. The method of claim 5, wherein said digitalisation applies G.711 codec and arranges the digitalised signal (9) into bytes using HDLC protocol, and wherein said least significant part of the signal comprises the less significant bit of some of said bytes.
7. A composite audio signal (3) **characterised in that** it is produced by the method for coding according to any one of claims 1 to 6.
8. A method for decoding a stereo audio signal, from a single composite audio signal (3), into its two components comprising a first output audio signal (14) and a second output audio signal (15), **characterised in that** it comprises the steps of:
- separating first and second coding information (7, 8) embedded into said composite audio signal (3) out of said composite audio signal (3), to extract said first and second coding information (7, 8) and a central audio signal (10, 11),
 - synthesising said first output audio signal (14) out of said first coding information (7) and said central audio signal (10, 11),
 - synthesising said second output audio signal (15) out of said second coding information (8) and said central audio signal (10, 11).
9. The method of claim 8, wherein synthesizing said first output audio signal (14) step adaptatively filters said central audio signal (10, 11) through a first adaptive filter (45) adapted by said first coding information (7), and wherein synthesizing said second output audio signal (15) step adaptatively filters said central audio signal (10, 11) through a second adaptive filter (46) adapted by said second coding information (8).
10. The method of claim 9, wherein said first adaptive filter (45) is identical to a first adaptive filter (42) used to code a first input audio signal (1) into said composite audio signal (3), and wherein said second adaptive filter (46) is identical to a second adaptive filter (43) used to code a second input audio signal (2) into said composite audio signal (3).
11. A transmitter able to transmit a stereo audio signal, composed of a first input audio signal (1) and a second input audio signal (2), as a single composite audio signal (3), **characterised in that** it comprises means for implementing the steps of the coding method of any one of claims 1 to 6 and transmitting means for transmitting said composite audio signal (3).
12. A receiver able to receive a stereo audio signal, transmitted as a single composite audio signal (3), and to deliver its components comprising a first output audio signal (14) and a second output audio signal (15), **characterised in that** it comprises receiving means for receiving said composite audio signal (3) and means for implementing the steps of the decoding method of any one of claims 8 to 10.
13. A communication system able to transmit a stereo audio signal, composed of a first input audio signal (1) and a second input audio signal (2), as a single composite audio signal (3), and to deliver its components comprising a first output audio signal (14) and a second output audio signal (15), a transmitter according to claim 11 and a receiver according to claim 12.

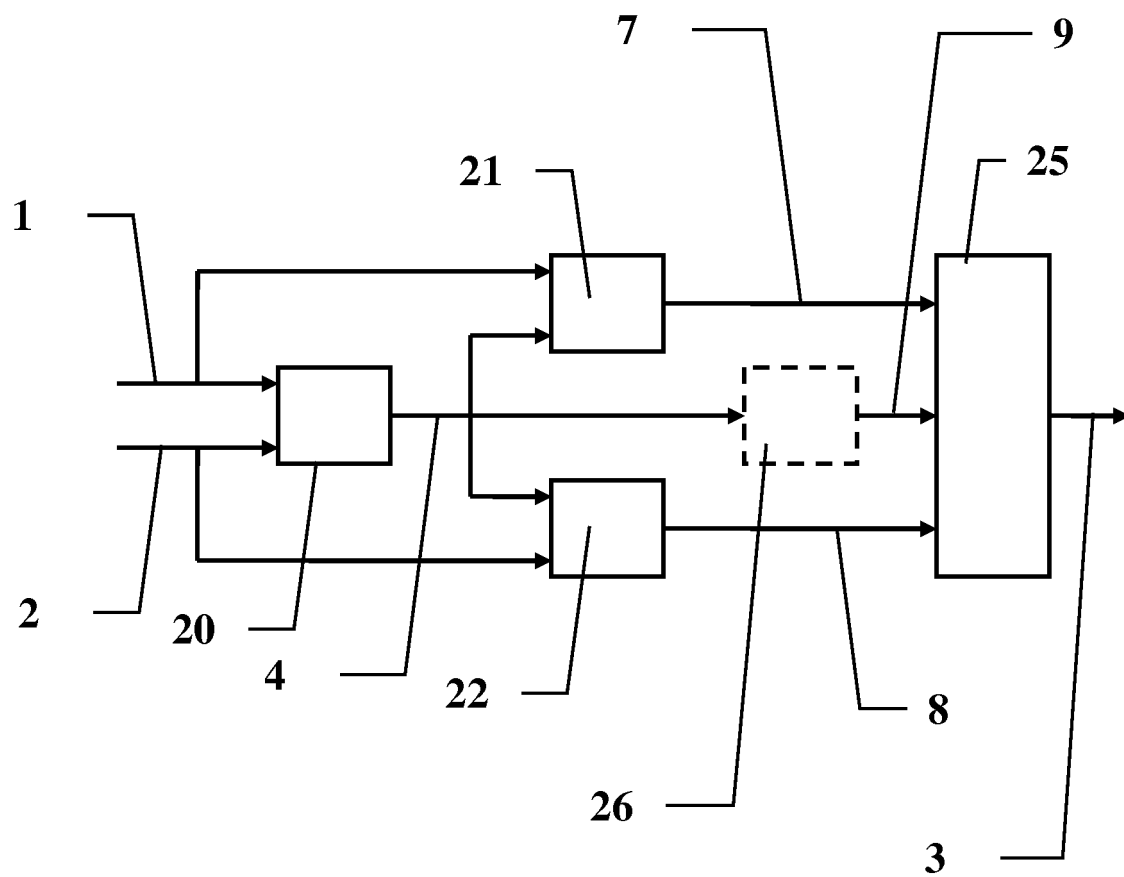


FIG. 1

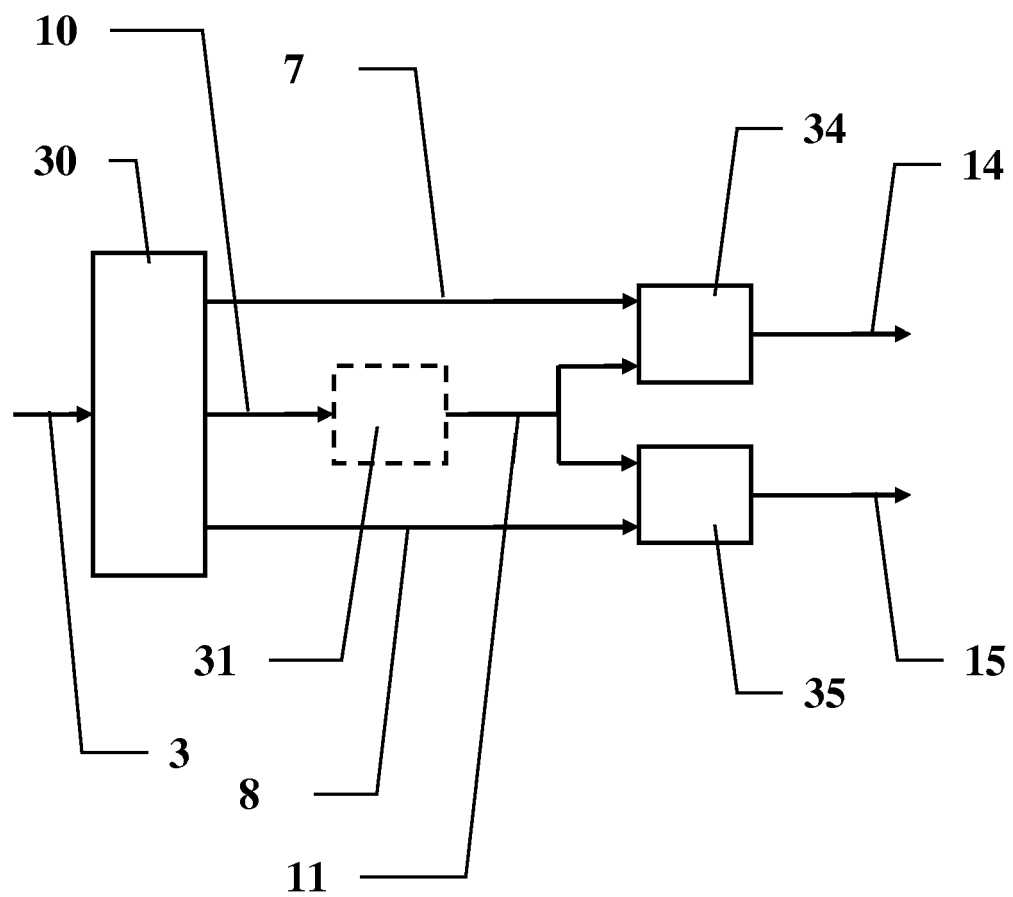


FIG. 2

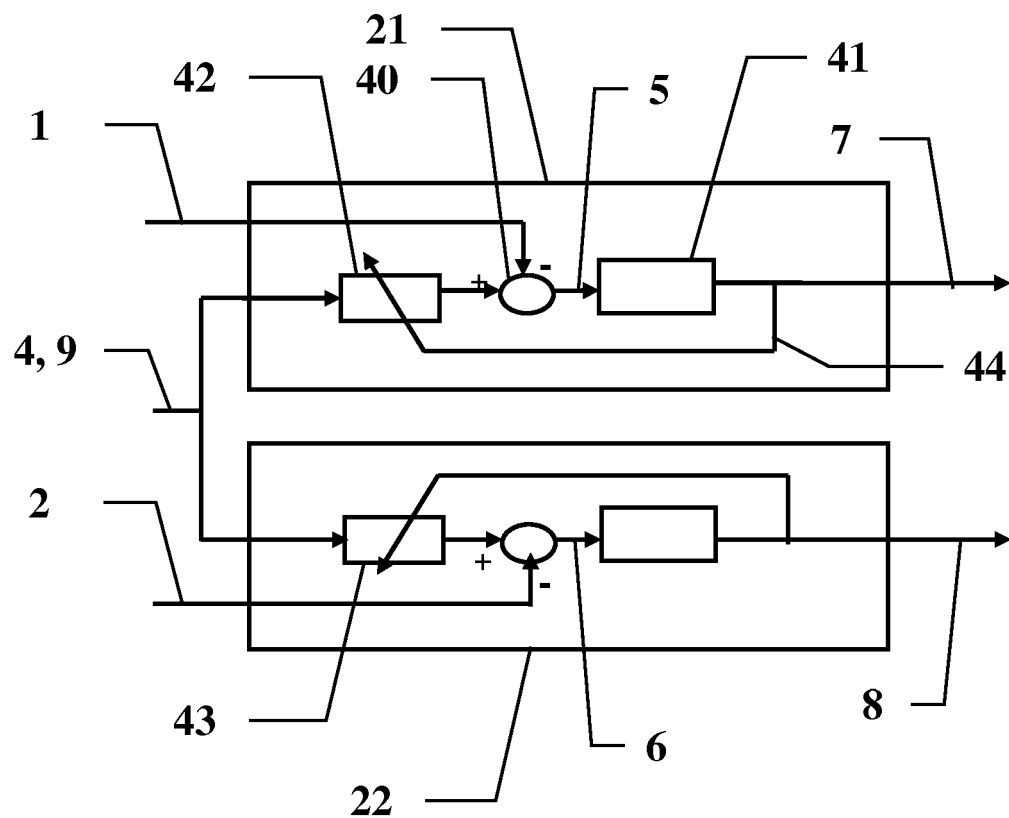


FIG. 3

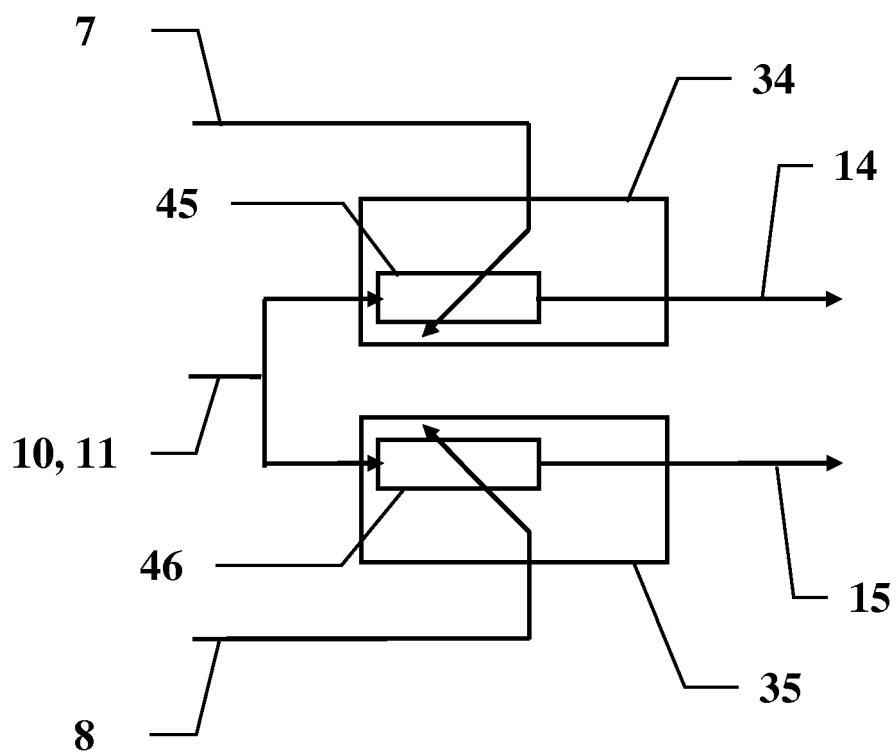


FIG. 4



EUROPEAN SEARCH REPORT

Application Number
EP 09 30 5680

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
X	US 2007/053521 A1 (FUCHIGAMI NORIHIKO [JP] ET AL) 8 March 2007 (2007-03-08)	1,7,8, 11-13	INV. H04H20/88
A	* the whole document *	2-6,9,10	G10L19/00
X	EP 1 926 082 A1 (DEUTSCHE TELEKOM AG [DE]) 28 May 2008 (2008-05-28)	1,7,8, 11-13	
A	* the whole document *	2-6,9,10	
A	US 2003/220800 A1 (BUDNIKOV DMITRY N [RU]) 27 November 2003 (2003-11-27)	1-13	
	* the whole document *		
			TECHNICAL FIELDS SEARCHED (IPC)
			H04H G10L
The present search report has been drawn up for all claims			
Place of search Munich		Date of completion of the search 10 August 2010	Examiner Horn, Ralph
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document			

6

EPO FORM 1503 03.82 (P04C01)

**ANNEX TO THE EUROPEAN SEARCH REPORT
ON EUROPEAN PATENT APPLICATION NO.**

EP 09 30 5680

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

10-08-2010

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2007053521 A1	08-03-2007	NONE	
EP 1926082 A1	28-05-2008	DE 102006055737 A1 US 2008136686 A1	29-05-2008 12-06-2008
US 2003220800 A1	27-11-2003	NONE	