(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau

(43) International Publication Date
2 May 2013 (02.05.2013)

WIPO | PCT

(10) International Publication Number
**WO 2013/061337 A2**

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **SINHA, Aniruddha** [IN/IN]; Tata Consultancy Services, Bengal Intelligent Park, Building - D, Plot No. - A2 M2 & N2, Block -GP, Salt Lake Electronics, Complex, Sector -V Kolkata 700 091 West Bengal (IN). **SAHA, Arindam** [IN/IN]; Tata Consultancy Services, Bengal Intelligent Park, Building - D, Plot No. - A2 M2 & N2, Block -GP, Salt Lake Electronics, Complex, Sector -V Kolkata 700 091 West Bengal (IN). **PAL, Arpan** [IN/IN]; Tata Consultancy Services, Bengal Intelligent Park, Building - D, Plot No. - A2 M2 & N2, Block -GP, Salt Lake Electronics, Complex, Sector -V Kolkata 700 091 West Bengal (IN).

(54) Title: METHOD AND SYSTEM FOR EMBEDDING METADATA IN MULTIPLEXED ANALOG VIDEOS BROADCASTED THROUGH DIGITAL BROADCASTING MEDIUM
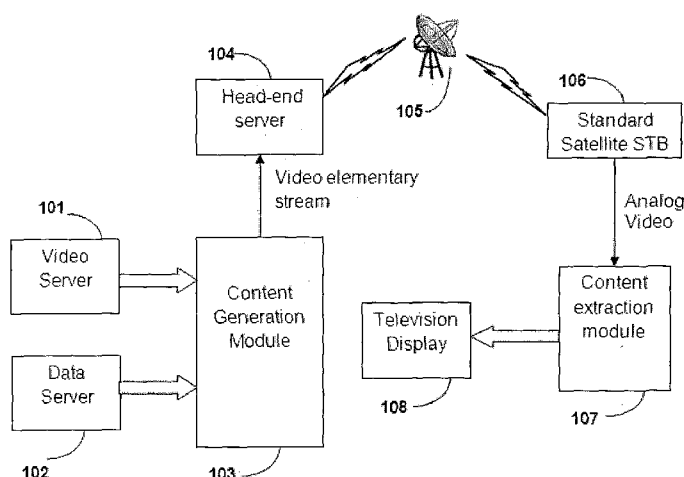
(57) **Abstract**: A method and system for broadcast of additional content such as metadata required for client specific interactive application in an analog domain along with conventional audio, video and PSI or SI data is disclosed. The present invention enables transmission of encoded audio data or EPG data, timestamp information required for audio video synchronization referred to as metadata by embedding such metadata in the pixels of video pixels and then encoding by the standard video encoder to generate an encoded stream. The encoded stream is decoded using the standard video decoder at the receiving station to generate a Composite Video Blanking and Sync (CVBS) analog video signal. From the CVBS signal, the RGB or YUV pixels of the videos are extracted. Finally a data extractor module retrieves the embedded metadata from the RGB or YUV pixels.

**(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, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, 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).

**Declarations under Rule 4.17**:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

**Published**:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

# METHOD AND SYSTEM FOR EMBEDDING METADATA IN MULTIPLEXED ANALOG VIDEOS BROADCASTED THROUGH DIGITAL BROADCASTING MEDIUM

## FIELD OF THE INVENTION

The invention generally relates to the field of television broadcasting. More particularly, the invention relates to a method and system for embedding metadata in multiplexed videos broadcasted in an analog domain through digital broadcasting medium.

## BACKGROUND OF THE INVENTION

Television broadcasting includes broadcasting of conventional video, audio, Program stream information (PSI) and System Information (SI) data etc through broadcasting stations such as satellite, Cable-TV and terrestrial stations. These stations transmit multiple audio-visual contents broadcasted by various broadcasting channels on the television sets located worldwide that displays the contents on TV display screen.

Considering the growth and scope of television broadcasting, there was a need to extend the television viewership to interactive applications. More particularly, there was a need for extending to develop interactive applications on the television such as online gaming, distance-learning and online chatting etc. These applications supported by the television broadcast have resulted in the concept of "Interactive TV".

The significance of "Interactive TV" has been accepted by all the sections of the society including the service providers, the customers and the government agencies. Therefore, there was a need in the art to enable a television broadcasting system that supports interactive applications as desired by the consumer at his or her home using the television as a medium for accessing various interactive applications.

However, as can be observed, the present broadcasting scenario is limited to transmit conventional audio, video, PSI and SI data only. Further, the data type associated with interactive applications may not be enabled to transfer using existing broadcasting infrastructure. Moreover, if the data type is other than the conventional audio, video, PSI and SI data then the data of such data type may not be transmitted using the existing broadcasting infrastructure. Hence, there is a need to enable system for transferring any other data type along with the conventional data type supported for transmission.

In the background art, a PCM Adaptor is known to enable embedding of raw audio data into a video frame. Also, there are systems known in the art embedding teletext and closed captions in the lines of vertical blanking interval (VBI) as disclosed in the CEA-608 standard.

The PCM adaptor as disclosed in the background art is limited to sending raw audio data as video frame and hence is not feasible to transmit data of any other type in the existing broadcasting scenario. Moreover, the audio analog data is not digitally encoded and transmitted through a digital broadcast network. Further, the payload size capacity for the teletext and the closed caption embedded in the VBI lines is very less.

Further, in the background art, there exist various systems to transmit digitally encoded bits in the VBI (vertical blanking interval) wherein, 7175 bits per second per line can be transmitted. However, there is a limit for the number of VBI lines which is of the order of 20-40 lines.

Moreover, at the receiver end, for decoding the embedded data of different types using the existing systems requires the support for specialized hardware devices. The VBI lines contain teletext, closed captions and electronic program guide (EPG). Hence, there is a limit for number of free lines available to send new type of data.

Therefore, as can be appreciated the existing systems and methods in the background art suffers from following limitations:

- The existing systems and methods lacks in transmitting data of any data type using existing broadcast infrastructure and limits in supporting few data-types to be transmitted.

- The existing systems and methods implemented for transmitting data of any type lacks in transmitting large volume of data using the existing broadcast infrastructures.

- The existing systems and methods implemented for transmitting data of any type requires specialized hardware devices for decoding such data at the receiver end.

In view of the above lacunae observed in the background art, there is a long-felt need for a method and system enabling transfer of any type of data in large volume along with the conventional data using existing broadcast infrastructure. More specifically, there is a need in the art for a system and method for transmitting data of any type in a format retrieved by the receiver terminal without any requirement of specialized hardware device.

## OBJECT OF THE INVENTION

The principal object of the invention is to enable a method and system for embedding metadata in a video frame transmitted trough digital broadcasting medium.

Yet another object of the invention is to enable a method and system for generating data symbols and pilot symbols related to the said metadata.

Yet another object of the invention is to enable a method and system for embedding the generated data and pilot symbols in the video pixels of the video frame to generate data pixels.

Yet another object of the invention is to enable a method and system for extraction of data symbols and video pixels at the receiver.

Still another object of the invention is enabling a method and system for displaying the metadata embedded in the video pixel corresponding to the extracted data symbols and the video pixels as video data on the display screen of the receiver.

## SUMMARY OF THE INVENTION

Before the present methods, systems, and hardware enablement are described, it is to be understood that this invention is not limited to the particular systems, and methodologies described, as there can be multiple possible embodiments of the present invention which are not expressly illustrated in the present disclosure. It is also to be understood that the terminology used in the description is for the purpose of describing the particular versions or embodiments only, and is not intended to limit the scope of the present invention.

The present invention enables a method for transmission and reception of information of any data type using the existing broadcasting infrastructure. The information is transmitted in such format that is easily decoded by the receiving device for display on the screen of the display device along with the video signals.

The present invention enables generation of data symbols for the information to be transmitted along with other video signals. The generated data symbols along with pilot symbols are inserted in the video pixel region of the analog video frame. The pilot symbols are detected by the match filter at the receiving station and accordingly the data symbols are extracted from the video pixels.

The present invention further enables rendering of metadata embedded in the video pixels derived from the extracted data symbols and the videos from the video pixels on the display screen.

The present invention enables a system comprising a content generation module that generates pilot symbols and data symbols representing the metadata to be transmitted along with video data signals. Further, the content generation module embeds the generated pilot and data symbols in the pixel region of video frames to generate data pixels that are multiplexed with video pixels. The multiplexed video pixel is generated and encoded to form a digital transport stream.

At the receiving station, the transport stream is decoded to generate analog CVBS signal. A content extraction module extracts the data symbols and the pilot symbols embedded in the video pixel region along with the video pixels of the videos embedding the data symbols from the CVBS signal. The metadata corresponding to data symbols extracted and the videos corresponding to video pixels are then displayed on the television screen.

## BRIEF DESCRIPTION OF DRAWINGS

The foregoing summary, as well as the following detailed description of preferred embodiments, is better understood when read in conjunction with the appended drawings. For the purpose of illustrating the invention, there is shown in the drawings exemplary constructions of the invention; however, the invention is not limited to the specific methods and architecture disclosed in the drawings:

Figure 1 is a system architecture diagram **100** illustrating various hardware elements enabled to broadcast any type of information using existing broadcast infrastructure in accordance to an exemplary embodiment of the invention.

Figure 2 is a block diagram **200** illustrating generation of data symbols and embedding of the said symbols in video pixels in accordance to an exemplary embodiment of the invention.

Figure 3 is a block diagram **300** illustrating extraction of data symbols and video pixels to display video and metadata in accordance to an exemplary embodiment of the invention.

Figure 4 is a block diagram **400** illustrating the generation of the data symbols and embedding them in the video pixels of the analog video frames in accordance to an exemplary embodiment of the invention.

## DETAILED DESCRIPTION:

The description has been presented with reference to an exemplary embodiment of the invention. Persons skilled in the art and technology to which this invention pertains will appreciate that alterations and changes in the described method and system of operation can be practiced without meaningfully departing from the principle, spirit and scope of this invention.

Referring to figure 1 is a system architecture diagram **100** illustrating various hardware elements performing the task of broadcasting any type of data using existing broadcast infrastructure in accordance to an exemplary embodiment of the invention.

In an exemplary embodiment, as illustrated in figure 1, the system architecture **100** comprises a video server **101** storing plurality of videos and a data server **102** storing data used for interactive television applications referred to herein as metadata collectively to be transmitted using existing digital broadcast infrastructure.

Further, the system comprises a content generation module **103**, a head-end server **104**, and a broadcasting station **105** as illustrated in figure 1. In this embodiment, the said broadcasting station is illustrated as a satellite broadcasting station. However, in alternative embodiments, the broadcasting station may be a cable-TV broadcasting station, a terrestrial broadcasting station or combinations thereof.

As illustrated in figure 1, at a receiving station, a standard satellite set-top-box (STB) **106** is shown receiving the videos and data transmitted by the broadcasting station. Further, a content extraction module **107** is shown coupled to the STB performing the task of extracting the videos and the data to be displayed on the display screen of the television **108** in accordance to the exemplary embodiment. In alternative

embodiments, the content extraction module **107** may display the videos and data on the display or monitor screen of a computer, a laptop, a mobile phone, Smartphone or combinations thereof.

According to an exemplary embodiment, the videos captured and stored at the video servers are in analog domain. These multiple videos may be carried by multiple broadcasting channels. The analog video frames are multiplexed in spatial and temporal domain to transmit multiple videos through one or two broadcasting channels as disclosed in the patent application 1894/MUM/2011.

According to an exemplary embodiment, the present invention enables transmission of any other data required for interactive applications such as the metadata to be displayed on the television along with the analog videos broadcasted by multiple broadcasting channels. In order to facilitate this, the metadata representing information of any data type that is required for interactive applications, more specifically for "Interactive TV" purposes is embedded into multiple analog video frames in a format that is decoded by the STB receiver at the receiving station to display on the television display.

In an exemplary embodiment, such metadata stored in data server **102** is transferred as text data or binary data. Further, the binary or text data is inserted in the video frames in a format that is generated by the said content extraction module **103** in such a way that it is able to be extracted by the content extract module **107** at the receiving station to be displayed on the television screen **108**.

In an exemplary embodiment, the content generation module **103** generates data and pilot symbols and then inserts the generated pilot symbols followed by the data symbols in the video frames carrying the analog videos to be transmitted trough digital broadcasting medium.

According to an exemplary embodiment of the present invention, the generation and insertion of data and pilot symbols is implemented as illustrated in figure 2.

As illustrated in figure 2, the metadata referring to data required for broadcast of interactive applications on television is transmitted in the form of data bits **201**. Further, pilot bits **202** are used as reference bits that are used for identification of the data bits **201** to be transmitted.

According to an exemplary embodiment, each of the data bits **201** and the pilot bits **202** are first converted into data symbols **203**. The data symbols **203** generated are then inserted in the pixel region of the analog videos to form data pixels. The analog videos comprise plurality of video pixels. Therefore, few of the plurality of video pixels in each of the analog videos are utilized for embedding the generated data symbols **203**. More specifically, the data symbols are embedded in the 'Y' and/or 'U' and/or 'V' component of the each of the video pixels of the said analog videos. Further, as illustrated in figure 2, the remaining video pixels **204** of the multiple analog videos to be broadcasted are multiplexed with the formed data pixels using a multiplexer **205** implemented in the content extraction module to generate multiplexed video frames **206** comprising video pixels with embedded metadata.

Referring to figure 4 is a block diagram **400** illustrating the generation of the data symbols and embedding in the video pixel in accordance to an exemplary embodiment of the invention. In this exemplary embodiment, a 3 byte symbol is generated for each 1 bit of data. However, in an alternative embodiment of the invention, N byte symbol can be generated for each one bit of data, where N is an integer or fraction.

As illustrated in figure 4, a data byte **401** indicating metadata to be transferred is shifted using shift register **402** to extract one bit data **403**. In the embodiment of the present invention, if the data bit **403** extracted is '0', then a data symbol 0x801111 **404** is generated containing three byte data. On the other hand, if the data bit extracted **403** is '1', then a data symbol 0x80EBEB **405** is generated.

Further, as illustrated in figure 4, the generated data symbols **404** and **405** are inserted in the video pixel region **406** of the analog video frames. In this

embodiment, the data symbols are embedded in the Y, U and V component of the video pixel region. Each of the video pixels in Y, U and V is of 8 bits or 1 byte size.

Now, referring to figure 1, the metadata inserted multiplexed video frames generated are encoded using any of the standard video encoders such as MPEG2, MPEG4, and H264 etc or combinations thereof as known in the art. The video encoder encoding the multiplexed video frames may be implemented in the content generation module 103. Thus, the output of the content generation module 103 contains a digital video elementary stream comprising the encoded video frames containing video pixels with embedded metadata as data symbols.

As illustrated in figure 1, the head-end server 104 receives the digital video elementary stream for further processing and implementing head-end functionalities. The head-end server 104 comprises a transport stream generation module that generates a digital transport stream comprising various analog video frames multiplexed in spatial and temporal domain captured from different video servers, wherein the video pixel region of the said analog video frames comprises data related to interactive television applications captured from different data servers as illustrated in figure 1.

Further, the generated transport stream is transmitted through the digital broadcasting medium. Thus, the invention enables transmission of multiplexed video frames in an analog domain with embedded metadata through existing digital broadcasting structure without any requirement of specialized hardware circuits. As illustrated, the transport stream is transmitted to the receiving station using the satellite broadcasting antenna 105 acting as the broadcasting station.

The transport stream is received by the standard satellite set-top-box (STB) 106 implemented at the receiving station as illustrated in figure 1. The STB 106 comprises a video decoder that decodes the received transport stream to generate the video pixels from the stream. Further, the STB is configured to generate the

Composite Video Blanking and Sync (CVBS) analog video signal from the decoded video pixels from the transport stream.

The CVBS video signal is further received by the content extraction module 107 as illustrated in figure 1. The task of the content extraction module 107 is to extract the embedded data symbols from the video pixel region of the analog video frames. Figure 3 illustrates extraction of the video pixels with data bits from the CVBS analog video signal in accordance to an exemplary embodiment of the present invention.

As illustrated in figure 3, the CVBS analog video signal 301 is received at a matched filter implemented in the content extraction module 107 illustrated in the figure 1. The function of the match filter is to extract the pilot symbols from the video signal for identifying the data symbols representing the metadata to be displayed.

In an exemplary embodiment, the extraction of pilot pattern is implemented by using a matched filter. The matched filter is designed using the similar pilot symbols pattern inserted while generating the symbols.

The content extraction module 107, after capturing a video frame from the CVBS analog video signal derives a threshold value of detecting pilot pattern for that particular frame using auto correlation value of similar pilot symbol pattern generated for that particular video frame with zero shifts.

In an exemplary embodiment, threshold value is determined to be 80% of the auto correlation value of the pilot symbols. Further, the correlation values between the matched filter and the received data from the captured frame are collected at the content extraction module. In an exemplary embodiment, if the position provides that the maximum values of correlation are greater than the threshold value then that position is determined as the starting location of pilot pattern. Further, a sharp peak in correlation value is observed that detects the pilot symbols pattern.

Now, again referring to figure 3, as illustrated, as-soon-as the pilot symbols **302** are detected from the analog video CVBS signal **301**, the corresponding data symbols **303** following the pilot symbols **302** representing metadata for interactive applications are extracted by the content extraction module **107** illustrated in figure 1. Further, as illustrated in figure 3, the data bits **304** corresponding to data symbols **303** are extracted. Also, the video pixels **305** corresponding different analog videos multiplexed by different broadcasting channels are extracted from the analog video CVBS signal **301**.

In an exemplary embodiment, the metadata corresponding to the extracted data bits **304** and the analog videos corresponding to extracted video pixels **304** are rendered on the display screen of the television **108** coupled to the said content extraction module **107** as illustrated in figure 1.

Thus, the present invention enables to generate content for broadcast with not only conventional audio, video and PSI (Program-specific information) or SI data but also any type of data that is required for client for specific application, more particularly for "Interactive TV" applications in the analog domain. The present invention can support transmit of even encoded audio data or EPG data, or some timestamp information for audio video synchronization or some metadata information to create an interactive client etc using the existing digital broadcast infrastructure.

**BEST ENABLED MODE:**

Consider a distance-learning education environment realized in accordance to an exemplary embodiment of the present invention. In the distant learning environment, the tutorials are in the form of audio visual contents and are need to be broadcasted with the support of local language and curriculum guidelines.

In an exemplary embodiment, consider the analog videos to be transmitted through represent tutorial videos in the distance learning scenario. Let the QA (Question-

Answer) series related to the tutorial videos represent the metadata to be embedded in the video pixel region of the videos.

In an exemplary embodiment of the present invention, the first task is to generate pilot and data symbols for the metadata to be transmitted along with the video frames of the tutorials. The metadata to be transmitted is in the form of hexadecimal format or data bits.

Each of the data bits are initially converted into data symbols. The data symbol length is of three bytes and is represented as: A symbol representing data bit 0 is X00 and a symbol representing data bit 1 is X11, wherein the value of X is equivalent to 0.5. In this manner, the pilot and data symbols are generated.

The next task is to embed the generated pilot and data symbols in the video pixel region of each of the video frames. In an exemplary embodiment, the symbols are embedded in the 'Y' component (luminance) of the video pixel region. The maximum value of Y is 0xEB and represented as data bit 0 while the minimum value of Y is 0x11 represented as data bit 1. Also, in this embodiment, the center value of Y is represented as X having value 0x80.

Considering the above Y values, the data symbols generated for data bits 0 and 1 are represented as 0x80EBEB and 0x801111 respectively, both of 3 bytes in length. In an alternative embodiment, the generation of symbols can be implemented wherein a symbol can be of two byte in length containing bit information representing data to be transmitted.

In an exemplary embodiment, consider a pilot pattern to be embedded in the video pixel comprises 48 bits representing the start of actual metadata embedded in the video pixel region of the video frame. Let the metadata representing QA series is inserted between 51st pixel and 670th pixel in a line of the video frame. Now, as each bit is implemented as three byte symbols, the pilot pattern symbols are inserted

between 51$^{st}$ and 194$^{th}$ pixel. In an exemplary embodiment, the possible pilot pattern symbols embedded in the pixel video region may be as follows:

10101010110011100011110000111100011100110101010101

OR

00001111000111001101010101010100110001110000111 1

In an exemplary embodiment, after the pilot symbols are embedded the audio data for the tutorial representing QA series are embedded as data symbols in the video frames. The audio data is inserted as part of each frame. The pilot symbols in each of the frames is followed by a frame number (8-bit), a tutorial ID (32-bits) and the payload length (16-bit) indicating audio data symbols to be followed after the tutorial ID. These data is then followed by audio symbols referred herein as metadata of the tutorials embedded in each of the video frames.

In an exemplary embodiment, for one second of audio data for one tutorial, the length of data is approximately 5000 bits for one language. Therefore, the audio symbols embedded are around 5000 symbols those occupy 15000 video pixels in the video frame. If the audio data is inserted only in alternate lines, then it will occupy 50 lines approximately. The lines to be inserted is altered every frame if used in successive frames.

In accordance to preferred embodiment of the present invention, the embedding of audio symbols is implemented as follows:

Frame number having audio is = 1 + 3*I+J, where I is the tutorial ID and J is the language ID.

Further, assuming 0 < 8 <= I and 0 < J <= 1, then the frame numbers 1, 4, 7, 10, 13, 16, 19 and 22 contains all the 8 tutorials for one language. Note that frame number 0 is the special frame and length of GOP is 25.

The tutorial video frames along with special frame carrying audio symbols are then transmitted by the broadcasting station to the receiving station. On the receiver station, the content extraction module extracts the audio symbols embedded in the tutorial video frames and the tutorial video pixels to be displayed on the display screen of the television. At the receiver end, the video frame is retrieved from the decoded transport stream to generate a CVBS video signal.

The payload field in the video frame indicates the presence of audio data symbols embedded in the video frame. Then, the pilot symbols patterns are extracted followed by the extraction of the audio data symbols following the pilot symbols in the video frame. Then, the extraction module extracts the video pixels of the tutorial videos from the CVBS signal.

Finally, the tutorial videos and the metadata extracted are rendered on the display screen of the television.

## ADVANTAGES OF THE INVENTION

The present invention has following advantages:

- The present invention enables to transmit any type of client specific information in large volume, as part of the analog video frame, utilized for different purpose of user interactive applications using the existing broadcasted medium.

- The present invention does not require any specialized hardware for enabling transfer of data of any type using existing broadcast infrastructures.

- The present invention enables transmitting multiple videos with embedded metadata through available channel bandwidth at the transmitting station by multiplexing the said videos and metadata in a spatial and temporal domain.

## CLAIMS:

1. A system for broadcasting a metadata along with at least one analog video carried by at least one analog video frame characterized by inserting the said metadata in the said analog video frame, the system comprising:

   a) a content generation module (103) to generate at least one set of data pixels comprising at least one data symbol and at least one pilot symbol related to the said metadata embedded into at least one first set of video pixels of the said video frame;

   b) a multiplexer in the said content generation module (103) to multiplex the said data pixels with a second set of video pixels of the video frame to generate at least one set of multiplexed video pixels;

   c) a match filter at the receiver (106) to detect the said pilot symbol in the multiplexed video pixels by matching the stored threshold auto-correlated value of the said pilot symbol; and

   d) a content extraction module (107) to extract the said data symbol corresponding to the detected pilot symbol from the multiplexed video pixels while identifying the said second set of video pixels as pure video pixels.

2. A system of claim 1, wherein the said metadata may include encoded audio data, timestamp information related to audio-video synchronization associated to the said analog video, Electronic Program Guide (EPG) data or combinations thereof.

3. A system of claim 1, further comprising a standard video encoder to encode the said first set of multiplexed video pixels to generate a video elementary stream.

4.  A system of claim 1, further comprising a head-end server to convert the said transport stream and transmit it through digital broadcasting medium.

5.  A system of claim 1, further comprising a standard video decoder at the receiver to decode the said multiplexed video pixels from the received transport stream and to generate an analog video Composite Video Blanking and Sync (CVBS) signal.

6.  A system of claim 1, further comprising a display device (108) to display the metadata corresponding to the said extracted data symbol and an analog video data corresponding to the identified pure video pixels from the multiplexed video pixels.

7.  A method for broadcasting a metadata along with at least one analog video carried by at least one analog video frame characterized by inserting the said metadata in the said analog video frame, the said method comprises the steps of:

    a)  generating at least one data bit (201) and at least one pilot bit (202) related to the said metadata;

    b)  converting the said data bit into at least one data symbol (203) and the said pilot bit into at least one pilot symbol (203);

    c)  embedding (204) the said data symbol and the said pilot symbol into at least one first set of video pixels of the said analog video frame to form at least one first set of data pixels;

    d)  multiplexing (205) the said first set of data pixels with at least one second set of video pixels of the analog video frame to generate at least one set of multiplexed video pixels;

    e)  detecting (302) the said pilot symbol in the said multiplexed video pixels at receiver by matching the stored threshold auto-correlated value of the said pilot symbol using a matched filter;

f)  extracting (303) the said data symbol corresponding to the detected pilot symbol from the said first set of data pixels in the multiplexed video pixels;

g)  retrieving the said data bit (304) from the said extracted data symbol;

h)  obtaining the metadata corresponding to the retrieved data bit; and

(i)  retrieving the said second set of video pixels (305) from the multiplexed video pixels.

8.  A method of claim 7, wherein the said data bit represents the actual metadata to be broadcasted along with at least one analog video of the said analog video frame.

9.  A method of claim 7, wherein the said pilot bit represents reference data indicating the start of the actual metadata to be broadcasted along with at least one analog video of the said analog video frame.

10.  A method of claim 7, wherein the said data symbol and the said pilot symbol are embedded into Y, U or V component of the said first set of video pixels.

11.  A method of claim 7, further comprising the step of encoding and transmitting the said multiplexed video pixels by a standard video encoder and then generating the transport stream and then transmitting as transport stream to the said receiver.

12.  A method of claim 11, wherein the said multiplexed video pixels in the transport stream are decoded by a standard video decoder and an analog video Composite Video Blanking and Sync (CVBS) signal is generated.

13.  A method of claim 6, wherein the said pilot symbol is detected from the said decoded transport stream to extract the said corresponding data symbol.

14. A method of claim 13, wherein the said data symbol is extracted from the Y, U or V component of the said first set of video pixels from the said multiplexed video pixels.

15. A method of claim 7, wherein the said metadata is selected from a group comprising of an encoded audio data, timestamp information related to audio-video synchronization associated to the said analog video, Electronic Program Guide (EPG) data or combinations thereof.

16. A method of claim 7, further comprising the step of rendering the obtained metadata and the retrieved second set of video pixels on a display device.
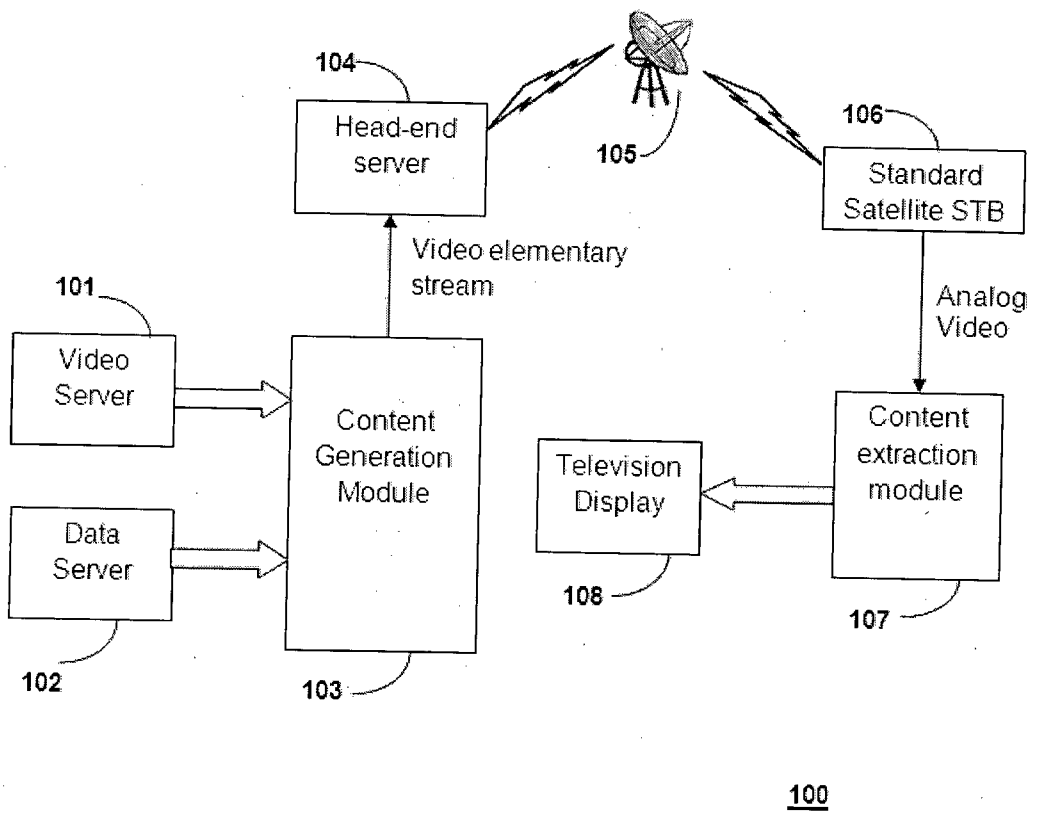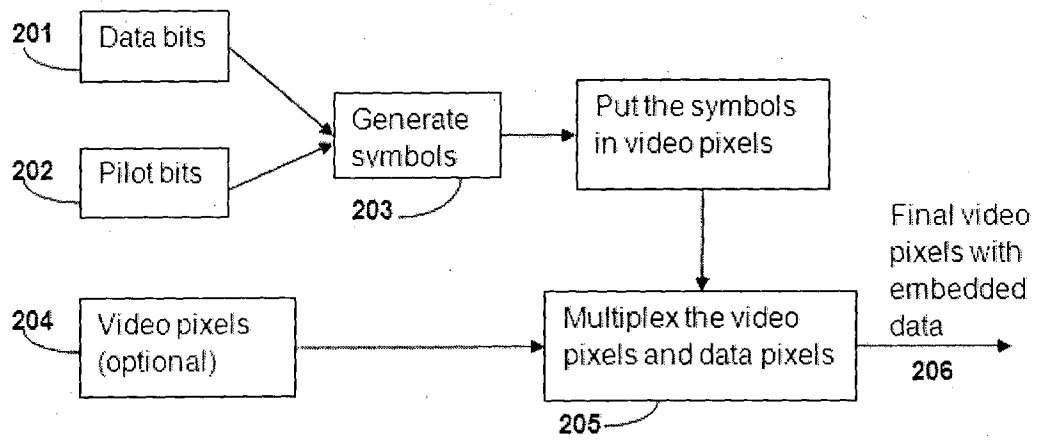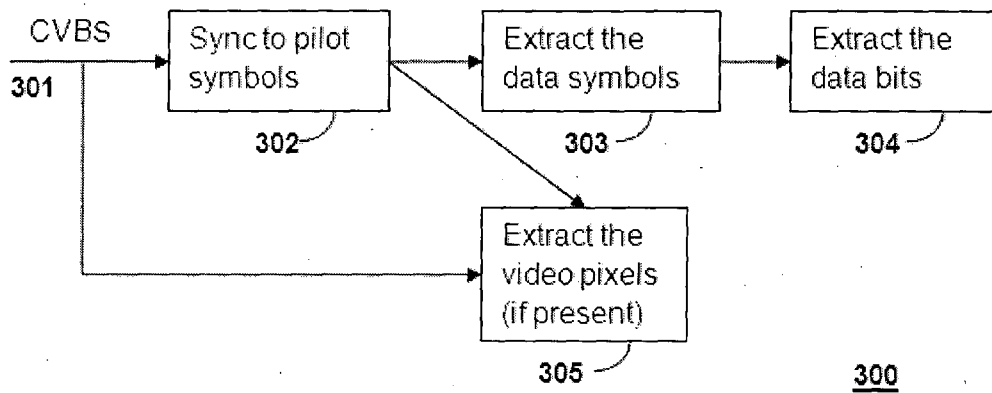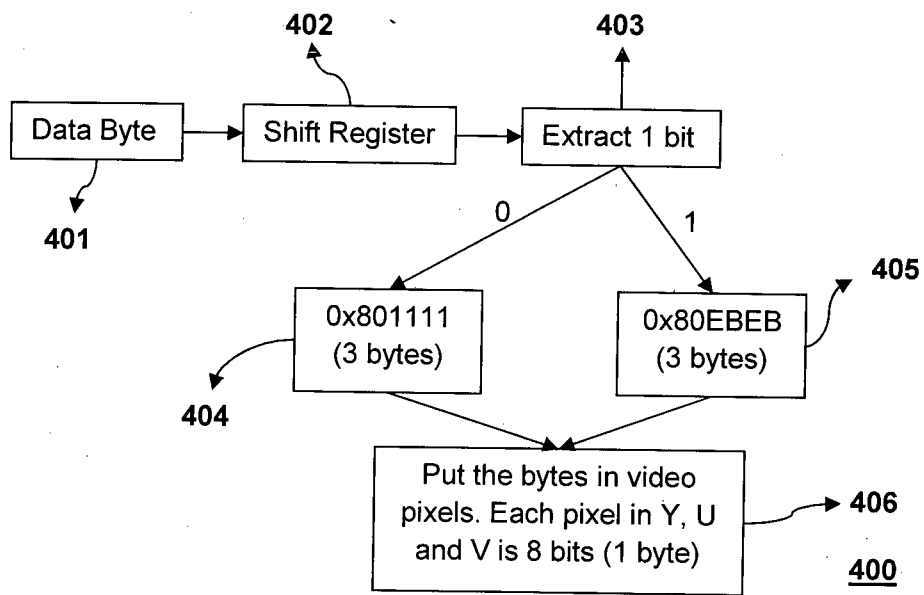
Figure 1

2/4



Figure 2

200

3/4



Figure 3

**402**

**403**

```
┌──────────┐      ┌──────────────┐      ┌──────────────┐
│ Data Byte│ ───▶ │Shift Register│ ───▶ │ Extract 1 bit│
└──────────┘      └──────────────┘      └──────────────┘
```

**401**

0          1

**405**

```
┌──────────────┐              ┌──────────────┐
│  0x801111    │              │  0x80EBEB    │
│  (3 bytes)   │              │  (3 bytes)   │
└──────────────┘              └──────────────┘
```

**404**

```
┌──────────────────────┐
│ Put the bytes in video│
│ pixels. Each pixel in Y, U│
│ and V is 8 bits (1 byte)│
└──────────────────────┘
```

**406**

**400**

**Figure 4**