10.1 Encoder Interfacing

It is anticipated that DTV encoders will contain the capabilities to encode video, audio, and ancillary data. DTVCC and NTSC captions are to be considered part of the ancillary data. DTV encoders should be able to encode the DTVCC and NTSC captions in the Picture User Data portion of the video elementary stream, and the Service Directory descriptors in the PMT and EIT.

It is also anticipated that a separate local or remote computing system (Caption Data Generator) will provide caption data and service descriptors to the DTV Encoder through a data interface port. Figure 18 depicts this arrangement:

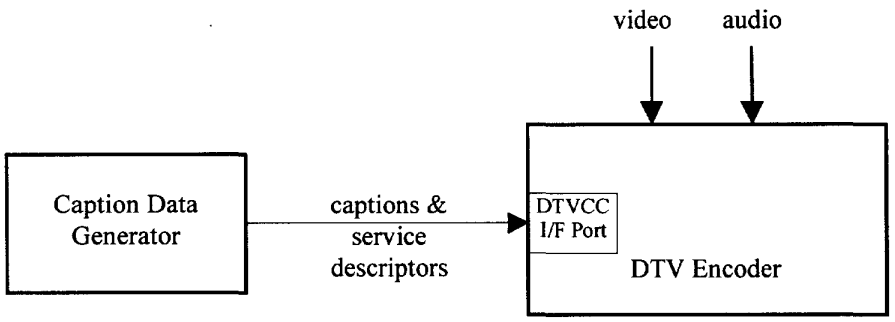

Figure 18 Caption Encoding Interface Example

The Caption Data Generator provides NTSC Captions, DTVCC Caption Channel Packets (Section 5) and Service Descriptors (Section 4.5) to the DTV Encoder.

It is recommended that DTV Encoders provide at least a basic byte-serial or byte-parallel connection for accepting closed-captioning data. Since the DTVCC Picture User Data rate for DTVCC and NTSC captions is 9600 bps, this port should provide a faster bit-rate interface (> 9600 bps) in order to overcome any port interface-protocol overhead required for handshaking, synchronization, check digits, and data demarcation.

10.1.1 Off-Line DTVCC Caption Encoding

Off-line encoding is the process of captioning program material prior to its airing. Caption providers prepare the captions for the programming before the encode process. The encode process usually entails sequencing the captions to the DTV encoder based upon timecode input which synchronizes the captions with specific portions of the program video. Caption providers attempt to have captions displayed at receivers as close as possible to specific (desired) frames of video.

Figure 19 shows an off-line encoding configuration where the Caption Generator is local to the DTV Encoder:

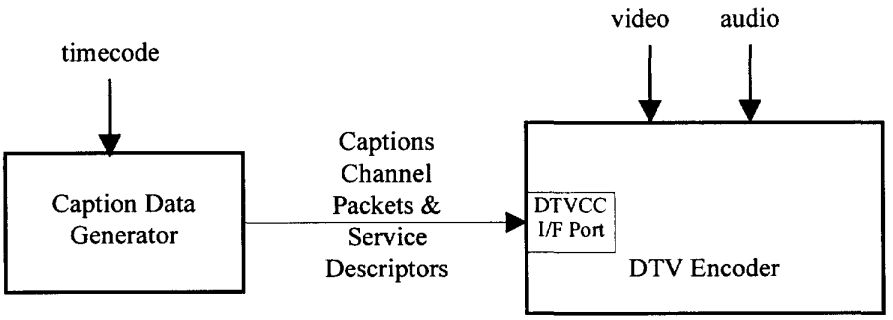

Figure 19 Off-line Caption Encoding Configuration Example

In order for captions to be displayed at the receiver at the desired video frames, the Caption Data Generator must take into account the delays inherent in the entire encode and decode process. DTV Encoders may not encode any bytes of a Caption Channel Packet until it receives the entire packet through the DTVCC I/F port.

For example, a 120-byte packet will take 500 milliseconds to be transmitted to an DTV Encoder with a 19.2K bps DTVCC I/F port interface. Once the packet is received, the DTVCC encoder must spread the packet across the available Picture User Data bits. For a 30Hz picture rate video, each of the Picture User Data constructs provides 40 bytes of caption data with 36 bytes available for DTVCC caption data, and 4 bytes for NTSC caption data (see Table 3). The User Data may therefore transport the packet in 3.333 (120 / 36) consecutive constructs, or within .111 seconds.

Since the Picture User Data are transmitted in Transport Order there may be a re-ordering delay during encoding. There may also be a re-ordering delay in the decoder in order to reconstruct the packet before it can be processed.

There is also the video buffer processing delay described in Section 4.4.4. However, this delay is not of concern since the video follows the same delay.

Therefore, the total delay from Caption Data Generator to the display of captions on the decoder is as follows:

TotalDelay = dSerialTransmit + dUserDataEncode + dDisplayOrderEncode + dDisplayOrderDecode

where:

dSerialTransmit = the packet transmit time to the encoder

dUserDataEncode = the delay for waiting to break the packet up into the Picture User Data constructs in the encoder.

dDisplayOrderEncode = the delay for Display Order processing of the User Data in the encoder.

dDisplayOrderDecode = the delay for Display Order processing of the User Data in the decoder.

The TotalDelay may significantly be shortened by implementing the following recommendations:

1. Increase the DTVCC I/F port speed to the highest practical baud rate.
2. Minimize the DTVCC I/F port interface protocol overhead.
3. Keep the Caption Channel Packets the same size as the bandwidth available in the Picture User Data constructs. For example, a 30Hz frame rate provides 36 bytes of DTVCC data per User Data construct. If the packet size is also 36 bytes, then dUserDataEncode is reduced to a single frame delay time - packets do not have to be split across multiple User Data frames. Individual packets do not

have to be re-ordered in the encoder and decoder, but there is a possibility that packets will be transported out of sequence. Also, the serial transmit time is reduced, and thus, the DTVCC Encoder can begin encoding the packets sooner.

10.1.2 Real-Time DTVCC Caption Encoding

Real-time encoding is the process of captioning live program material (e.g., broadcast news and sporting events) as it is airing. See Figure 20. Caption providers employ real-time court-reporting techniques where the live audio is monitored by a stenographer and entered into a computing system. The stenograph input is translated into caption data. In most cases, the caption provider will be remote to the DTV encoding system and the captions will be transmitted to the encoder via telephone modems.

The caption provider may be listening to the program audio from a local DTV receiver, or they may have a dedicated audio feed from the broadcaster which is tapped into the broadcast audio prior to transmission. Audio being monitored from an DTV receiver will lag the pre-transmitter audio by approximately 0.5 seconds (see Section 4.4.4). For real-time captioning, the pre-transmitter audio feed is preferred since the caption delay will be reduced, and the captions will better match the audio at DTV receivers.

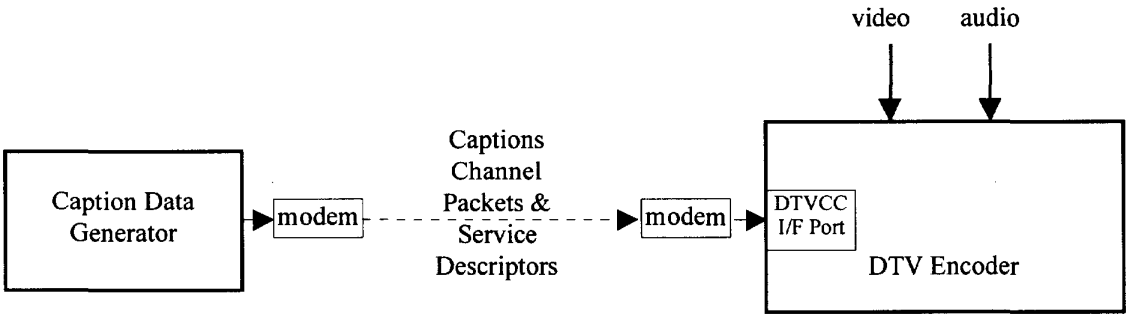

Figure 20 Real-Time Caption Encoding Configuration Example

The delays described in the previous section also apply to real-time captioning. Thus, the recommendations for delay reduction previously described are likewise applicable (e.g., high speed modems and small packets).

10.1.3 Multi-Purpose DTVCC Interface Port

Although DTV Encoder manufacturers may be tempted to implement separate ports for DTVCC Captions, NTSC Captions, and Service Directory encoding; it is strongly recommended that a single port be provided for all three data entities. In this way, all captioning related material can be inserted into the transport stream from a single computer system interface. The number of physical connections is minimized for both local and remote captioning.

DTV Encoder manufacturers should implement an interface protocol which allows each of these three data sources to be distinguished from each other. Similarly, Caption Data Computer manufacturers should support this same protocol when providing caption-related material to the DTV Encoder.

It would serve the industry best if a standard DTVCC Interface Port protocol were developed. This would minimize the development costs for the Caption Data Generators.

10.1.4 Service Directory Insertion

DTV Encoder manufacturers should provide commands within the DTVCC Interface Port protocol to allow the Caption Data Generator to add Service Descriptors (see Section 4.5) to the PMT and EIT. When receiving a descriptor for a service that is already encoded in the PMT and EIT, the old descriptor is to be replaced with the new descriptor.

> NOTE--The new descriptor may be larger or smaller than the old descriptor because of the optional Service Title field in the descriptor.

DTV Encoder manufacturers should also provide commands to allow Service Descriptors to be deleted from the PMT and EIT. The gap left by the deleted descriptor in the PMT or EIT should be closed by moving any subsequent descriptors up in the Service Directory construct in the PMT or EIT, respectively.

Caption Data Generators should use these service descriptor add/delete commands during the captioning session. A "descriptor add" command should be issued at the beginning of the session; a "descriptor delete" command should be issued at the end of the session to remove the descriptor from the PMTand EIT (so that viewers will not see the descriptor in the on-screen directory for services that are no longer present).

10.2 Downstream Encoding

DTV Encoder manufacturers should provide the capability to add new services to program streams that already contain caption services. Caption Channel Packets received for the new service should be inserted in the earliest-available free bytes in the Picture User Data. If a gap in the User Data is not large enough to contain the packet, then it is recommended that subsequent packets be removed from the User Data and buffered, allowing the new packet to be inserted. The buffered packets would be let back into the User Data after the inserted packet. It is up to the DTV Encoder manufacturer to implement this scheme in an efficient manner as is economically possible. The intent is to maximize the utilization of the DTVCC bandwidth.

BIBLIOGRAPHY

ATSC A/54, Guide to the Use of the ATSC Digital Television Standard (1995)

- Advanced Television Systems Committee (ATSC), 1750 K Street N.W., Suite 1200, Washington, DC 20006; Phone 202-828-3130; Fax 202-828-3131; Internet http://www.atsc.org/stan&rps.html

Annex A (Informative)

Figure 21 illustrates one example of a decoder implementation for processing the PMT, EIT and User Data in the DTVCC Transport Stream. It shows the separate paths of the Service Descriptors in the PMT and EIT and the service data in the Picture User Data bits.

NOTE--Different block diagrams may exist for other implementations.

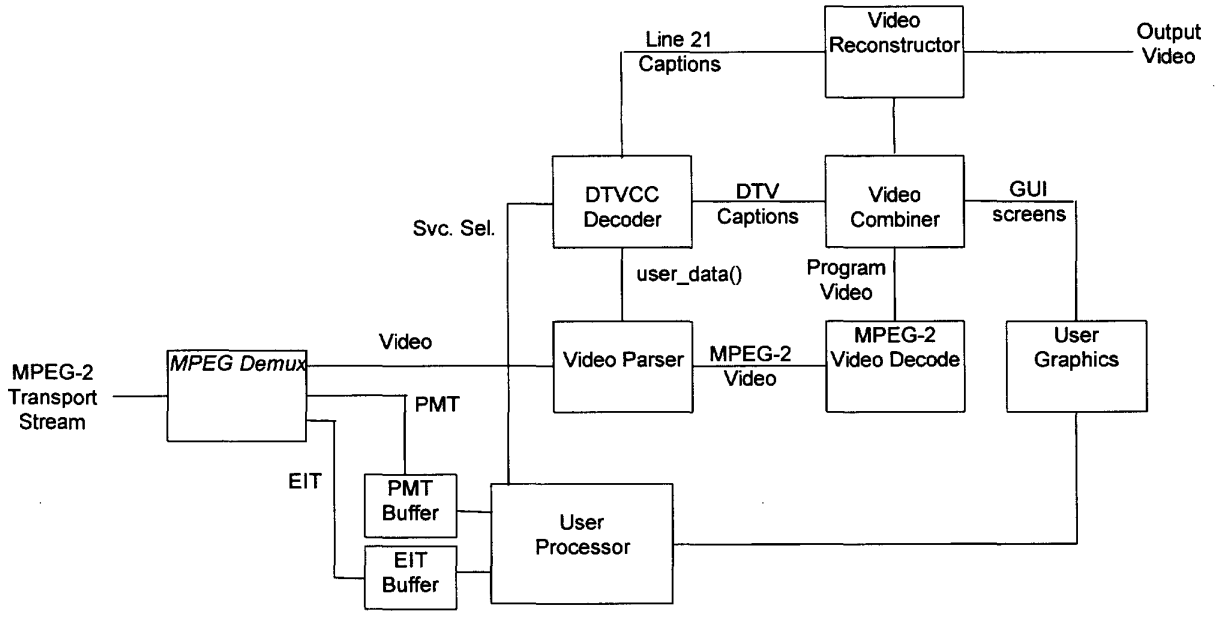

Figure 21 DTVCC Transport Stream Decoder