

# ***Transmission of Still and Moving Images Over Narrowband Channels***

## ***Application Report***

***Stefan Goss, Wilhelm Vogt, Rodolfo Mann Pelz, Dirk Lappe  
Communication Research Institute***

SPRA110  
February 1994



## **IMPORTANT NOTICE**

Texas Instruments (TI) reserves the right to make changes to its products or to discontinue any semiconductor product or service without notice, and advises its customers to obtain the latest version of relevant information to verify, before placing orders, that the information being relied on is current.

TI warrants performance of its semiconductor products and related software to the specifications applicable at the time of sale in accordance with TI's standard warranty. Testing and other quality control techniques are utilized to the extent TI deems necessary to support this warranty. Specific testing of all parameters of each device is not necessarily performed, except those mandated by government requirements.

Certain applications using semiconductor products may involve potential risks of death, personal injury, or severe property or environmental damage ("Critical Applications").

**TI SEMICONDUCTOR PRODUCTS ARE NOT DESIGNED, INTENDED, AUTHORIZED, OR WARRANTED TO BE SUITABLE FOR USE IN LIFE-SUPPORT APPLICATIONS, DEVICES OR SYSTEMS OR OTHER CRITICAL APPLICATIONS.**

Inclusion of TI products in such applications is understood to be fully at the risk of the customer. Use of TI products in such applications requires the written approval of an appropriate TI officer. Questions concerning potential risk applications should be directed to TI through a local SC sales office.

In order to minimize risks associated with the customer's applications, adequate design and operating safeguards should be provided by the customer to minimize inherent or procedural hazards.

TI assumes no liability for applications assistance, customer product design, software performance, or infringement of patents or services described herein. Nor does TI warrant or represent that any license, either express or implied, is granted under any patent right, copyright, mask work right, or other intellectual property right of TI covering or relating to any combination, machine, or process in which such semiconductor products or services might be or are used.

## Overview

The transmission of pictures over radio channels can be of great benefit to:

- Public authorities, such as the police and emergency services,
- Public transportation, such as railways, airlines and ships,
- Private citizens

Radio networks have a narrow bandwidth and therefore require low transmission rates. Moreover, radio channels are prone to interference, such as that caused by multipath propagation. This report describes channels and transmission methods in existing networks and shows how, by applying complex algorithms for coding images, you can develop a source codec on the basis of a multisignal processor system can be with the aim of achieving a source data rate as low as 8 kbps.

## Networks and Transmission Methods

### Nonpublic Land Mobile Telecommunication Network

The existing Nonpublic Land Mobile Telecommunication Network (NPLMTN) is characterized by narrow-band frequency modulation with a channel separation of 20 kHz in the 450-MHz region. For the considered low data rate video and speech transmission, the existing network structure can be used through application of available commercial equipment. For proper digital modulation methods with a constant envelope, consider the well-known variety of continuous phase modulation types. In this context, Gaussian minimum shift keying (GMSK) [1] is characterized by a relatively high bandwidth efficiency, which is achieved with a pulse shaping filter with Gaussian characteristics. A data rate of 16 kbps can be attained in this application for an ACI of  $-70$  dB, which is a general requirement for single-channel-per-carrier land mobile radio systems.

After the system inherent noncoherent demodulation (limiter + discriminator), a modified maximum likelihood sequence estimation (MLSE) with the Viterbi Algorithm (VA) is performed to obtain the transmitted data sequence. This method takes the effects of nonideal intermediate frequency filters at the demodulator output into account.

Due to the underlying narrow-band transmission — in other words, the signal bandwidth is less than the so called coherence bandwidth of the mobile channel (50–500 kHz) — the propagation system is characterized by time-selective fading. The received signal equals the product of the transmitted signal and a complex stochastic Gaussian process, which exhibits a Rayleigh or Rice distributed envelope and a uniform distributed phase [2].

### Public Switched Telephone Network (PSTN)

You can transmit video and speech in the analog telephone network through commercially available modems with data rates up to 24 kbps (Codex 326XFAST) in a synchronous mode. This high bandwidth efficiency is achieved with trellis coded  $m$ -ary QAM modulation [3]. The maximum likelihood detection is performed with the VA.

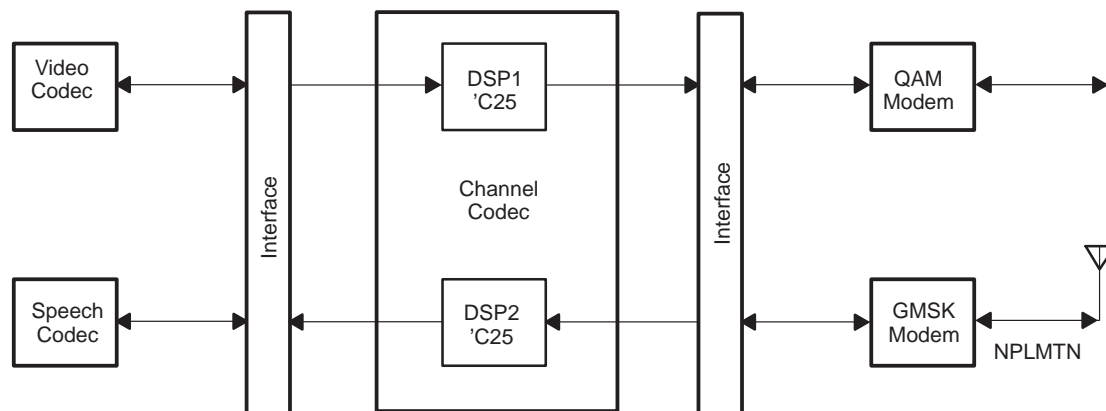
An exact characterization of the underlying propagation medium is a difficult task because a typical telephone channel cannot be defined. A simple model assumes a band-limited nonideal bandpass system and additive white Gaussian noise (AWGN). The linear distortion of the transmitted signal results in intersymbol interference (ISI), in which the error patterns are characterized by error bursts. Figure 1 depicts the system developed for video and speech transmission in the PSTN and NPLMTN.

Other applications of video and speech transmission are the analog and digital cordless telephone systems CT1 and DECT, and the analog and digital mobile radio telephone systems C and D (GSM) [4]. A current field of research includes the future public land mobile telecommunication system (FPLMTS), or the European version UMTS.

### Forward Error Correction (FEC)

The effects of radio channels on data transmission can be compensated for to some extent through application of FEC. Due to the limited resources (e.g., finite data rate), an efficient channel coding is a primary goal. Furthermore, due to the different sensitivity of individual symbols or symbol groups in the source coded data sequences (video, speech) to channel errors, you should devise an unequal error protection. In general, you can do this with block codes, but the application of convolutional codes allows ML decoding with soft decisions and channel state information. Rate compatible punctured convolutional (RCPC) codes [5] support dynamic allocation of redundancy with one encoder and one decoder. In the case of channels with memory, like the mobile channel, an additional interleaver and deinterleaver must be considered. Figure 1 shows two TMS320C25 digital signal processors (DSPs) implementing the corresponding algorithms for channel encoding.

**Figure 1. Digital Transmission System**

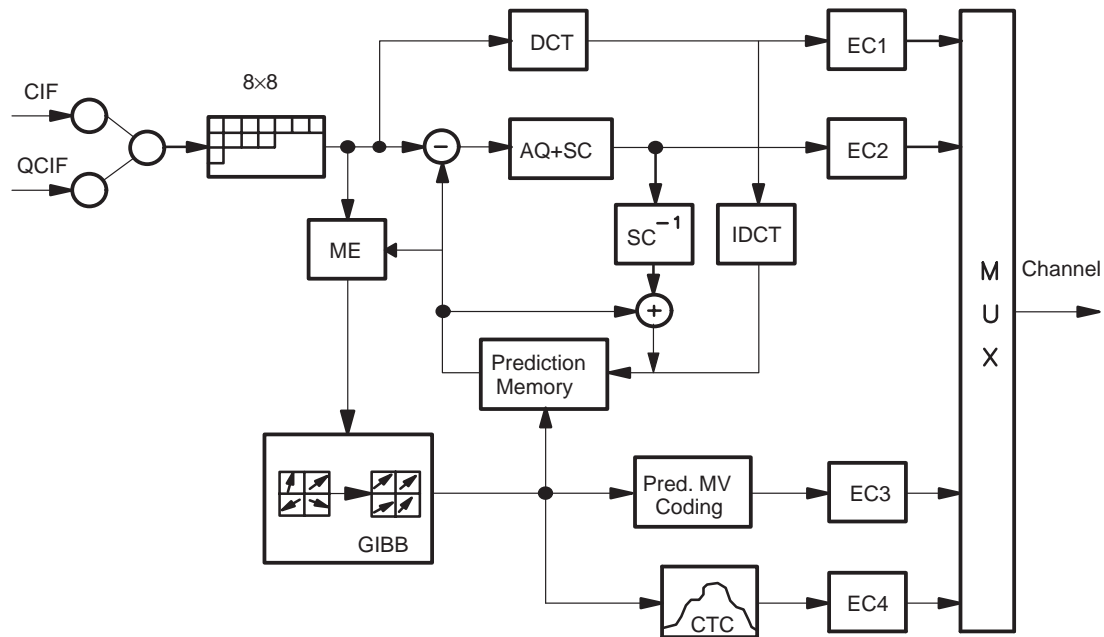


## Image Source Coding

### A Hybrid Codec for Moving Pictures

The source codec for moving pictures in an ISDN environment is committed and standardized as an H.261 recommendation. The data rate is fixed with regard to one B-channel (64 kbps). For data rates between 8 and 16 kbps, more efficient algorithms are necessary. A hybrid source codec for  $p \times 8$  kbps is shown in Figure 2. The following text describes this in detail and addresses the main differences with respect to the H.261 codec.

Figure 2. Hybrid Codec for  $p \times 8$  kbps



The input image format is QCIF (quarter common intermediate format) with a spatial resolution of  $180 \times 144$  pixels for the luminance signal (Y) and  $90 \times 72$  pixels for the color different components (U and V). The temporal resolution is reduced from 50 Hz to 6.25 Hz (factor 8). These operations are carried out in an additional preprocessing stage.

Like an ISDN codec, the hybrid codec is split into the motion estimation part and the coding stage of the prediction error. In the example shown in Figure 2, the QCIF image is the input information for a block generation stage that divides the input image into 396 blocks of  $8 \times 8$  pixels. In the next step, a motion estimation (ME) for every block is performed by calculating a mean square error (MSE) between luminance blocks of the input image (also called original image) and the last coded and decoded image (prediction image). Every block of the original image is matched in a window of  $40 \times 40$  pixels in the prediction image. The window is centered with regard to the block position in the original image. For a fixed number of dedicated positions in the window, the mean square error between the original and the predicted block is computed. The result is the motion vector for the minimum of mean square errors.

For typical videophone applications, the size of moving objects (i. e., persons) is higher than the block size. The minimization of the MSE per block leads to nonhomogeneous vector fields inside objects. The additional postprocessing stage (GIBB) smoothes the vector field with a model-based algorithm [6]. In this

case, the maximum probability of the moving direction of objects is computed, starting with the vectors of the ME. This operation leads to a much more homogeneous vector field and to a subjectively better reconstruction and motion compensation in the prediction memory on the coder and the decoder side.

Another advantage is the reduced number of bits for the differential coded motion vectors. This reduction is greater than the additional bits required for coding the larger prediction error.

The coder control stage has the observed value of the bit consumption per frame. The necessary bits for a set of attributes and the coded motion vectors are subtracted from the total number of bits per frame. The remaining bits are used for coding the prediction errors of the blocks.

A DPCM Loop (differential pulse code modulation) performs the coding. The MSE for every block of the original image is sorted and compared with a fixed threshold to obtain an intraframe/interframe decision. The blocks with an MSE above this threshold are intraframe coded. As in an H.261 Codec, the block will be transformed in the frequency domain by a DCT (discrete cosine transformation). A linear quantization of the nonzero coefficients and a run-length coding of zero coefficients reduces the number of bits for the underlying block.

Blocks with an MSE below the intraframe/interframe threshold are interframe coded. After the MVs and the DCT information are transmitted, the remaining bits for this frame are divided by the computed amount of bits per interframe blocks. This gives the number of blocks, which could be coded in this frame.

In contrast to the H.261 concept, when an interframe-DCT is implemented, the block differences are coded in the time domain. This involves an adaptive quantization (AQ) and the coding of special structures (SC) belonging to the block [7].

In a first step, the probability density function of the pixel difference inside the blocks is determined. Next, a three-step symmetric quantizer is devised by computing the thresholds and the replacement values. The three replacement values are transmitted to the decoder.

Every  $8 \times 8$  differential block is quantized with this quantizer function and then divided into sixteen  $2 \times 2$  blocks. Inside these  $2 \times 2$  blocks, only  $3^4 = 81$  combinations of replacement values are possible and are called structures. Former subjective investigations have shown that these 81 structures can be represented by only 31. In this case, only a 5-bit wide index must be transmitted to reconstruct the 31 structures on the receiver side. The described process is a special form of vector quantization.

The source-coded sequence is obtained by multiplexing (MUX) the individually coded parameters: mask of moving objects (in blocks), motion vectors, address information of coded blocks, DCT coefficients, replacement values, and quadtree information of structure coded  $2 \times 2$  blocks and their indices.

Every parameter group exhibits a different probability density. The introduction of an entropy coder (EC) for every group allows an additional bit reduction up to 50 percent. In typical applications, the entropy coding is realized by different Huffman tables or by means of an arithmetic coder, as in case of the presented codec.

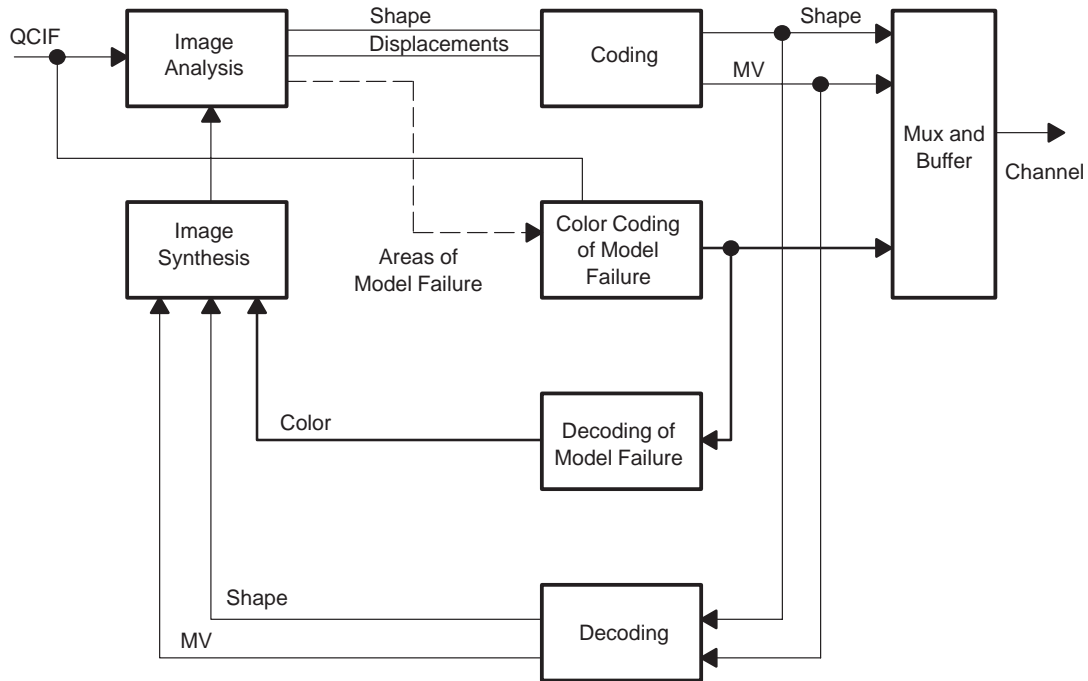
The source decoder is part of the source coder. It consists of the DPCM loop for the interframe coded blocks, an inverse DCT, and the prediction memory. After entropy decoding of the received sequence, the motion compensation is calculated in the prediction memory by using the decoded motion vectors. The reconstructed image is displayed after the decoding of the intraframe-coded blocks (IDCT) and the interframe-coded blocks in  $SC^{-1}$  at the coder and the decoder side.

### **Advanced Source Codec Architectures**

Advanced codec architectures are under investigation worldwide. The main goal is to change from block-oriented to object-oriented algorithms. The main disadvantages of block-oriented codecs are the

visible errors like blocking and the mosquito effect [8]. Figure 3 depicts a proposal for an image sequence coding scheme, which is based on an object-oriented analysis-synthesis approach. With this approach, the original input image is decomposed into objects, each described by a set of shape, motion, and color (luminance and chrominance) parameters in the image analysis-synthesis stage. Therefore, different model types (i. e., 2-dimensional, 3-dimensional, rigid, and nonrigid objects) are possible. The areas in which no modeling is possible are denoted as model failure.

**Figure 3. Object-Oriented Analysis-Synthesis Codec for  $p \times 8$  kbps**



The shape information has pixel accuracy and is coded by polygon and spline approximation. This leads to a nonvisible shape error of between one and two pixels. The motion information has half-pixel accuracy. For coding the model failure, different methods are under investigation. The (decoded) parameter sets of any object are forced as the input information for the image synthesis stage. If the objects in case of the analysis-synthesis codec are interpreted as a set of blocks, then the analysis-synthesis codec equals the hybrid codec scheme.

### Coding of Still Pictures

The source coding on still images can be executed with the same codec architecture. A motion estimation is not necessary. The input image format is CIF (common intermediate format).

### Speech Coding

Speech source coding is accomplished by means of an LPC (linear predictive coding) codec with a data rate of 2 to 4 kbps. The corresponding algorithms can be found in [9].

## Realization of a Source Codec Based on a Multiprocessor System

In this section, several alternatives for the realization of the source codec presented in *A Hybrid Codec for Moving Pictures* are discussed.

A system for implementation of complex algorithms cannot be conceived with standard ICs, such as ALU and multipliers or with programmable ICs, such as erasable and nonerasable PLDs (programmable logic devices), and LCAs (logic cell arrays), because it would not be compact, reasonable, and cost effective. Moreover, fast prototyping with standard components becomes time consuming as soon as unavoidable modifications become necessary. Even a demonstrator is not practical, because it does not lead to higher integration.

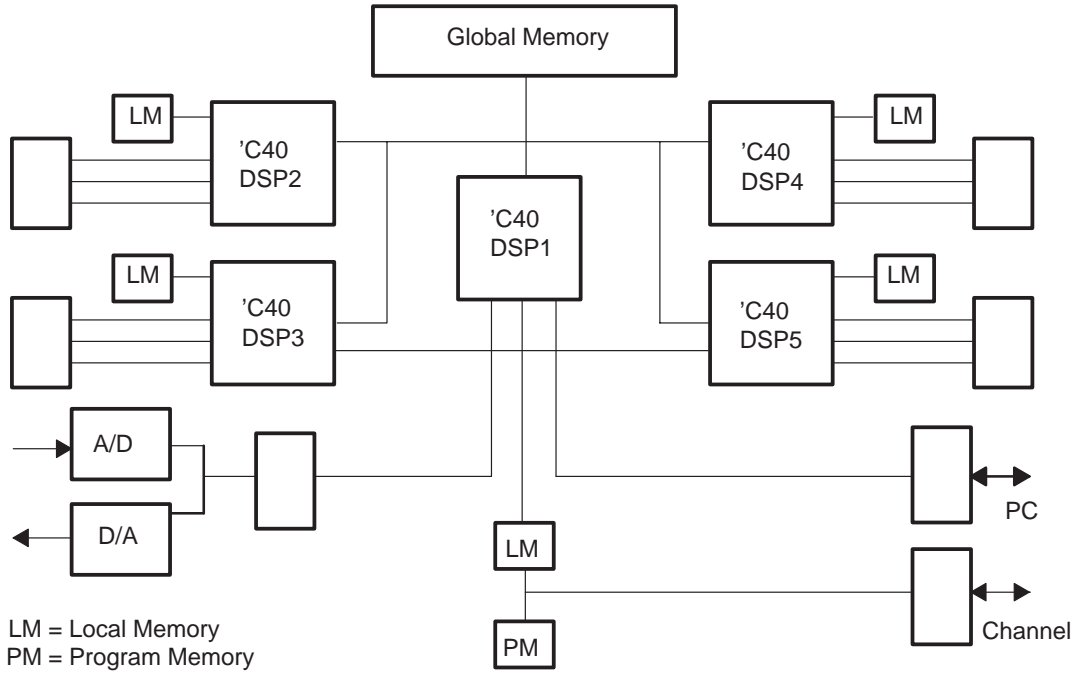
Flexible hardware should involve programmable signal processors. One solution could be the use of specialized single instruction multiple data (SIMD) architectures for high data rates. These processors are supplied with microcode programs for standard video algorithms. Today, available systems that are built with SIMD and a higher hierarchical level with multiple instruction multiple access (MIMD) processors do not have the necessary computational power. To implement the algorithms of the image codec, high-performance digital signal processors and their development tools are required. The DSPs should also be programmable in a high-level language (C) and should be provided with libraries supporting multisignal processor systems. This enables easy system extension with added processors for more computational power and guarantees picture coding in real time.

Figure 4 shows this type of system. It is composed of five 'C40s, which are linked together through their communication ports, and a fast global memory for storing the images [10]. The ports are connected together in the form of a "spoked wheel", the hub of the wheel being the master processor. Every slave processor has three parallel ports, which can be used to communicate to other 'C40s or to dedicated picture-coding components to form a more complex parallel processor system. The master DSP uses one of its 6 communication ports to communicate with the PC. Another port is used for data transmission between the video codec and A/D and D/A converters. The other four communication ports are tied to the slave processors. The host PC is used during the debug phase as development platform. The JTAG interface ties all processors together and is controlled by emulator software running on the XDS510 board [11]. In a future version, the PC will be used as a control interface to the picture codec. Also, compressed video sequences could be stored on the PC disk for several postprocessing operations.

Before the algorithm is implemented on a multi-DSP system, it must be divided into tasks. A self-written operating system supports the distribution of the tasks on the multi-DSP system. The master DSP controls the process, which is determined by the picture frame rate. Control words are sent over the communication ports. The data is exchanged over the global memory or over the parallel ports. After the completion of the described picture coding system at the end of 1992, further investigations concerning the application of the system in a mobile environment will be conducted.



Figure 4. Multi-DSP System



## References

- [1] Murota, K., and Hirade, K. "GMSK Modulation for Digital Radio Telephony", *IEEE Transactions in Communications*, vol. COM-29, pp. 1044–1050, 1981.
- [2] Jakes, W.C. *Microwave Mobile Communications*, John Wiley, 1974.
- [3] Ungerböck, G. "Channel Coding with Multilevel/Phase Signals", *IEEE Transactions in Information Theory*, 1982, pp. 55–67.
- [4] Mann Pelz, R., and Biere, D. "Video and Speech Transmission in Mobile Telecommunication Systems", *Nachrichtentechnik Elektronik*, vol. 1, pp. 7–12, 1992.
- [5] Hagenauer, J. "Rate Compatible Punctured Convolutional Codes (RCPC) and Their Applications", *IEEE Transactions in Communications*, vol. COM-36, pp. 389–400, 1988.
- [6] Stiller, C. "Motion Estimation for Coding of Moving Video at 8 kbit/s with Gibbs Modeled Vectorfield Smoothing", *Proc. SPIE, Lausanne*, pp. 468–476, 1990.
- [7] Amor, H. "Quellencodierung der Feinsturkturkomponente hochaufgelöster Bilder", *Fernseh und Kinotechnik* 37, pp. 15– 20, 1983.
- [8] Musmann, H.G, Hötter, H., and Ostermann, J. "Object-Oriented Analysis-Synthesis Coding of Moving Images", *Image Communication*, vol. 1, 1989.
- [9] Tremain, T. E. "Government Standard Linear Predictive Coding Algorithm: LPC-10", *Speech Technology*, pp. 40–49, 1982.
- [10] *MPS40 Hardware Reference*, SKALAR Computer GmbH, Göttingen.
- [11] *TMS320C4x User's Guide*, Texas Instruments, 1991.