CCITT Study Group XV
Working Party on Visual Telephony
Specialists Group on Coding for Visual Telephony

Document No SHARP 18

Date 22 March 1985

Source: British Telecom

Title: Comments on Source Coding

Document D9 (Tokyo meeting) summarised those activities in Europe which are relevant to the work of the specialists group. This document provides some recent information concerning source coding activities in Europe and raises some points that require discussion within the specialists group.

OVERVIEW OF CURRENT STUDIES

The work on source coding for 'second generation' videoconference codecs in the COST 211 bis project is still at the stage where a range of algorithms are being compared. Around 7 laboratories are involved in the computer simulation of coding schemes. In the case of most algorithms two or more laboratories are studying the same fundamental process and several laboratories have more than one algorithm under consideration. The aim is that by June or July of this year, the number of basic coding concepts meriting further study will be reduced and the remaining candidate algorithms will be presented to the specialists group for joint studies, and further optimisation.

Coding approaches currently under study are:

1. Various forms of hybrid coding using the Discrete Cosine Transform
2. Human Observer Oriented Transforms
3. Hybrid coding using pyramid transforms
4. Vector quantisation
5. Recursive Binary Nesting combined with predictive coding
6. Conditional Replenishment
7. Classical Transform Coding

To date no one or two schemes are emerging as clear winners. Comparisons are being performed using the common set of test sequences demonstrated at the Tokyo meeting, and using a common, reduced resolution, format for coding.

OBSERVATIONS

It is apparent that at 384 kbit/s and below it will not be sufficient to rely on conventional source coding alone, as most schemes are only providing acceptable pictures down to 0.3 bits per pel. A considerable degree of pre-processing is therefore likely to be needed, although at this early stage we should not rule out the possibility that much of this could be motion adaptive.

Accepting that pre-processing will be required, with regard to source coding there are advantages in adopting a non-interlaced image structure, as this provides a straightforward distinction between temporal and spatial domains. It would of course be necessary to compensate for any temporal shift involved in the formation of the non-interlaced image by means of interpolation. It might not however be necessary to carry out the inverse operation at the receiver.

There is also an argument that one or more intermediate image formats involving perhaps less lines and/or fields/second than normal video could be derived that were common to both 625/50 and 525/60 codecs. This would then provide the key to standards conversion (see Document SHARP 17).

Solely for the purpose of comparing algorithms, the simulation work in the COST 211 bis project on 384 kbit/s coding is being carried out using an intermediate format of:

12.5Hz non-interlaced field rate

288 lines

360 pels per line (luminance)

Apart from easing the source coding and standards conversion problems, the use of a lower resolution coding format can also have a considerable influence on the implementation, where at lower clock frequencies one can begin to consider the use of microprocessors or advanced signal processors for real time coding.

Disadvantages are that the resolution loss could limit the applications of the codec to face to face conferencing only, so careful consideration needs to be given to methods of including means of transmission of pictures of detailed documents or objects. Also it is likely that there will be requirement for using primary bit rates for 'economy' broadcast applications such as ENG where permanent temporal or spatial resolution loss may not be acceptable.

Participants within the specialists group are urged to consider the above in their deliberations on source coding.

RCN/045