SUBJECT: Object Analysis Techniques Applied to a Hybrid DPCM/Transform
            Coder.
SOURCE: UK, NL, FRG

INTRODUCTION

It would appear from recent simulation results that conventional hybrid
DPCM/Transform video coders are reaching their performance limits. The picture
quality that these coders produce at 64kbit/s can be described as 'just acceptable',
therefore any significant improvement in quality is desirable.

Conventional hybrid DCT coders use no 'scene content' information or 'knowledge'
of the scene to appropriately code the data within the scene. One sub-block is
operated on with the same visual importance as the next. It is clear that certain areas
within a scene may be visually more important than other areas and the objects
within the scene obey a set of rules with respect to motion. By using this information
the hybrid DCT coder can be significantly augmented. The generic term used to
describe this approach will be refered to as 'object analysis'.

Starting from the most general set of possible images it is thought a typical sub-set
of images that would be used in a 64kbit/s viewphone would consist of a head and
shoulders against a background. Thus any object analysis can be optimised to this
sub-set of images.

OBJECT ANALYSIS APPLICATIONS

Two distinct areas have been investigated. The first is the location of objects within
the scene such that special visual importance may be attributed to them. The second
is the location of objects within the scene and the tracking of their motion so as to
significantly improve the impression of smooth motion even at very low frame rates.

Head and Shoulders Identification

It is thought that recognition of the head and shoulders within a scene will allow the
encoder to allocate a higher quality to the head within the scene at the expense of
poor background quality. This also leads to the ability to define 'background' with a
high degree of confidence. Demonstrations have been produced indicating a 'head
location' algorithm and a 'head and shoulders' identification within a hybrid DCT
loop.

Object Identification for Motion Prediction.

It is thought that if the average frame rate of any coding scheme falls below 10
frames per second, problems of lip synchronisation and jerky motion will occur.
However an advantage of low frame rates is that the spatial resolution is good. By
using object identification and motion estimation it is possible to produce artificial
frames in between the true coded frames. These artificial frames are constructed
using data only from the last coded frame and translating the objects therein inorder
to retain the impression of true motion. Further the mouth and eyes within the
artificial frame may be specially treated in order to retain good lip synchronisation

1

and retain natural eye motion. Demonstrations have been produced to illustrate this technique.

CONCLUSION

The concept of object analysis is very young as far as image coding is concerned. The above techniques represent only a few of the many possibilities for improvements in this direction. Ultimately a significant improvement in coded picture quality is thought to be possible using these techniques.