RTT TECHNOLOGY TOPIC
March 2002

Source coding

Some time ago (February 2000!) we reviewed progress with low bandwidth video compression techniques including MP4 compliant encoders.

As predicted, the MP4 standard has moved on to embrace object based encoding/decoding and many new pre-processing and post processing techniques.

In this month's HOT TOPIC, we set out to show how compression techniques and technologies will change future content value.

Compression is **not** a new technique. First generation cellular phones used analogue compression technologies (companding to reduce dynamic range, pre-emphasis/de-emphasis to reduce high frequency noise). Analogue compression delivered improved audio quality from a bandwidth limited noise limited channel.

Second generation cellular handsets introduced digital codecs on the basis (hope and expectation) that a better voice quality/bandwidth quality trade off could be achieved by compressing and filtering in the digital rather than analogue domain.

Additionally it was decided to use codecs that coded in the **frequency domain** - ie to use speech synthesis codecs rather than the existing simpler waveform **time domain** ADPCM codecs used in wireline and digital cordless phones.

The time domain/frequency domain transform provided a more effective way of exposing the redundancy and predictability inherent in the input audio waveform and provided the basis for a lower rate codec that delivered acceptable/tolerable voice quality.

This basic premise still holds true for digital cellular vocoders today - refinements have centred on the development of variable rate codecs that are switchable (the 3GPP1 multi-rate vocoder) and/or adaptive (the 3GPP2 variable rate vocoder).

Compression is achieved by expressing the harmonic content of speech as a series of digital filter coefficients, differences from sample to sample are sent as a difference residual. If a speech frame sample is severely errored, the sample is discarded and the prior sample re-used.

These techniques are similarly applied to image and video encoding. Voice codecs (vocoders) exploit the psycho acoustic properties of human beings - we can tolerate a relatively large loss of audio 'information' without perceiving a significant loss of quality.

Image and video encoders exploit similar limitations in our psycho-visual perception.

M-PEG4 encoders/decoders presently available encode pixel blocks (typically 16 x 16 pixel macroblocks) using a discrete cosine transform to describe the frequency content of the block (discrete cosine transforms replace the sine components of the frequency coefficients with a single number to reduce source coding overhead).

Differences from macroblock to macroblock are encoded and sent as difference information. Similarly, image to image differences and similarities are described, encoded and compressed. Typically, compression ratios of 40:1 are achievable with tolerable image quality.

However, one of the interesting things about M-PEG4 (and what differentiates M-PEG4 from M-PEG2 and previous M-PEG standards) is the support for object coding.

Object coding separates complex objects (for example a person moving across a background) from primitive objects (for example a table, a road - anything inanimate that doesn't move about). The complex and primitive objects can be separately encoded.

Complex objects do not need to be particularly complex. An object moving across a background only changes if it deforms, moves into shadow or rotates.

When an object (eg person) moves, it/he/she moves along an 'optic flow axis', ie the axis and direction of movement can be predicted (and reconstructed or 'rendered' in a decoder).

How an object moves and how an object changes perspective can be described using 'mesh coding', a method for coding **where objects are** within a co-ordinate system **and where they are going** - a process known as motion estimation.

Part of the data sent is a rendering instruction. The decoder uses this data to move (i.e. 'render') the object along and across the screen.

Similar techniques can be used in (very) low bit rate video conferencing. The decoder contains a generic face. Only differentiation parameters are transmitted (the difference between the face encoded and the generic face stored in the decoder). The face on the decoder is a neutral expressionless face. It is only the animation parameters that are transmitted (smiles, eye movements, lip movements).

Effectively in this example, we are trading increased memory bandwidth in the decoder (and processor overhead) for reduced delivery bandwidth.

Motion description and motion estimation together provide the basis for motion compensation. 3G handsets with integral cameras are relatively light devices that suffer from camera shake. Camera shake creates a moving image and increases entropy - i.e. the encoder bit rate increases. Motion compression can be used to eradicate or reduce camera shake prior to encoding.

Most parents with children see mesh coding and motion estimation used in practice every day on Play station/game console devices.

The techniques are well understood.

It is only processor power budget and memory constraints that to date have prevented large scale application of mesh coding in digital cellular phones.

Mesh coding is effectively a mechanism for improving content quality (and thereby content value). It trades handset resident memory and processor bandwidth against (a reduced need for) radio and network bandwidth.

Summary

Source coding has evolved as part of the 2G to 3G digital cellular transition to embrace object coding and mesh coding techniques. These techniques involve a substantial increase in **decoder** (handset resident) memory and processor power.

The optimisation of mesh coding will be a key performance differentiator in next generation handset design.

At network level, source coding will directly influence traffic properties and offered traffic distribution and will need to be considered as an integral part of the network planning process.

About RTT Technology Topics

RTT Technology Topics reflect areas of research that we are presently working on.

We aim to introduce new terminology and new ideas to clarify present and future technology and business issues.

Do pass these Technology Topics on to your colleagues, encourage them to join our Push List and respond with comments.

Contact RTT

RTT, the Shosteck Group and The Mobile World are presently working on a number of research and forecasting projects in the cellular, two way radio, satellite and broadcasting industry.

If you would like more information on this work then please contact

geoff@rttonline.com

00 44 208 744 3163