4.4 Media Scaling and Media Filtering

4.4.1 Media Scaling

Definition

Scaling = adaptation of the data rate to the capacity of the network

Requirements for Scaling Algorithms

- Quick and precise adaptation to the capacity of the network
- Minimum possible visual/audible effect
- Robustness in case of packet loss

Resource Reservation vs. Stream Adaptation

Reservation of Resources

- The application requests the necessary QoS (and thus network resources) in advance. Those resources will be guaranteed for the duration of the stream.

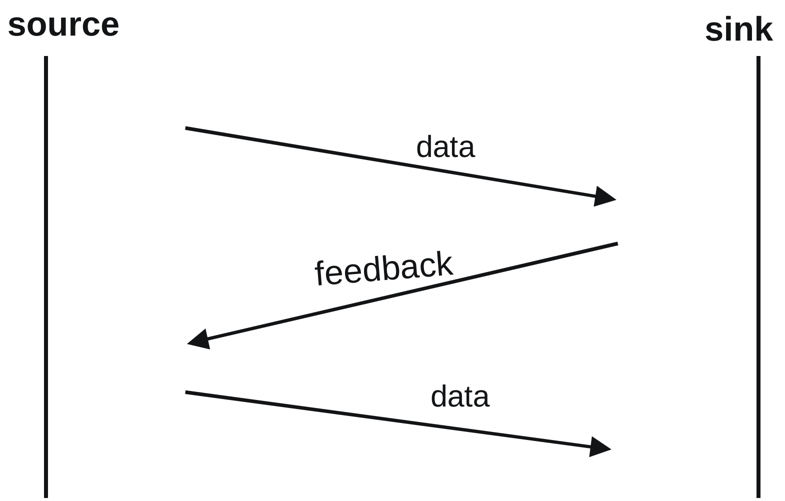
Adaptation

- The **application** adapts dynamically to the resources/capacity currently available from the network.
- This works for existing „best effort“ networks, in particular for the Internet with IP Version 4.
- Adaptive applications must implement audio resp. Video codecs that allow scaling at run-time.

Reservation and adaptation can be combined, in particular for probabilistic QoS guarantees: A reasonable amount of resources is reserved in advance, and in short phases of exceedingly high bit rates, the application scales down the stream.

Messages Exchanged for Scaling

- The source starts to send at some initial data rate.
- The network or the destination sends feedback messages indicating their current load.
- The source adapts the data stream by dynamically changing the parameters of the codec.



Scaling Example

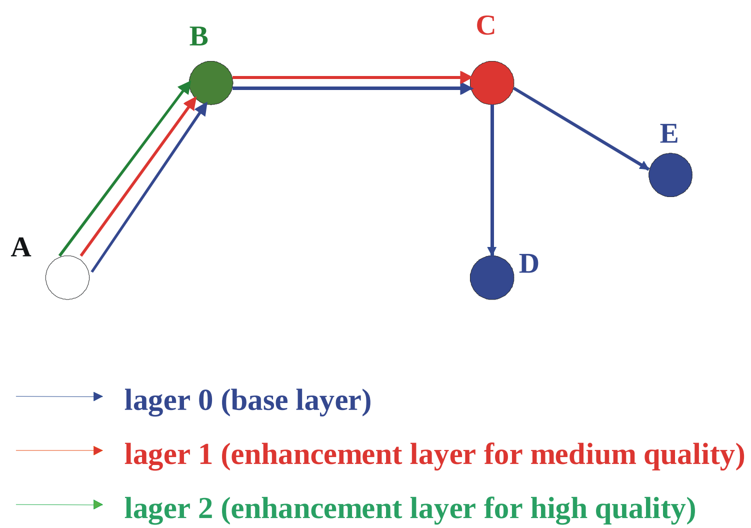
While an MPEG video is being played from a video server to a user the network signals a congestion. As a consequence, the video server **changes the quantization table** of the MPEG codec (increases the quantization step size) and thus generates VBR video at a lower bit rate (and a lower image quality).

Note the difference to the reaction you would observe on a network without adaptive (scaling) applications: the congestion control algorithm of the network (e.g., the one contained in TCP) would **reduce the packet rate**, leading to a lower video frame rate at the receiver. The network has no alternative; only the source (the application) can handle the overload situation in a more intelligent way.

4.4.2 Media Layering and Filtering

Layering and filtering are especially relevant in **multicast trees:**

- Different receivers require different bit rates and stream qualities.
- Only those parts of the stream are transported to a receiver that it can really use.
- Irrelevant parts of the stream are removed **by the inner nodes of the network.**



Layered Video Encoding

The two most important layering techniques for video are **temporal scaling** and **spatial scaling**.

Temporal Scaling

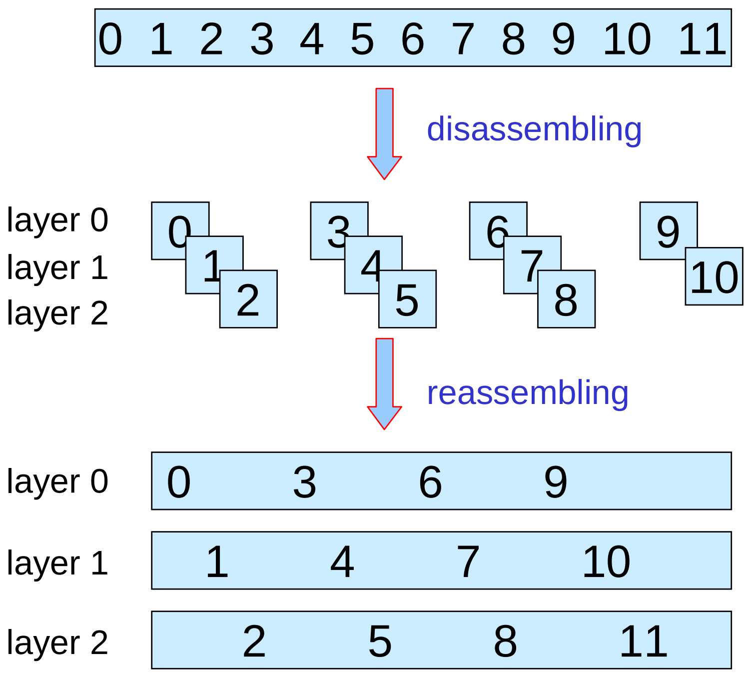
Assign frames to layers in a round-robin fashion. The more layers the receiver gets, the higher the frame rate will be.

Spatial Scaling

Assign data values to layers such that for each frame, a low-quality version can be reproduced from layer 0; adding higher layers will add "visible quality" to the frame.

There are other scaling/layering schemes, e.g., spatio-temporal techniques, which we will not discuss here.

Temporal Scaling for Video

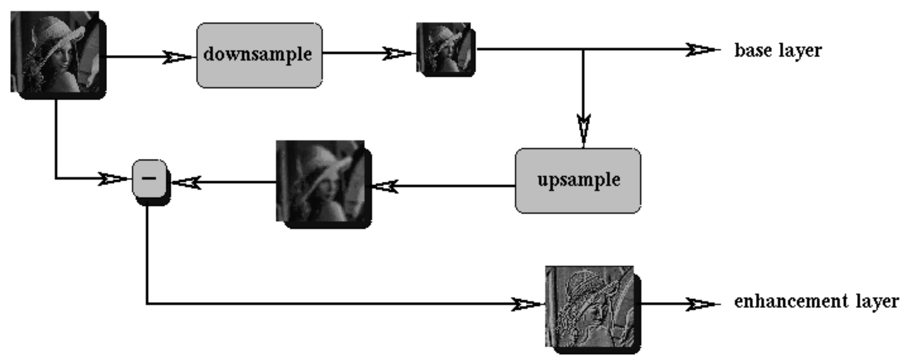


The base layer (i.e, layer 0) transmits every third frame. So do the first and the second enhancement layers. Together, the full frame rate is transmitted.

Spatial Scaling for Video (1)

Pyramid Encoding

The encoder first down-samples the image, compresses it according to the chosen encoding technique, and then transmit the result in the base layer. At the same time, it decompresses the image, up-samples it and gets a coarse version of the original. To compensate for the difference, the encoder subtracts the resulting copy from the original and sends the encoded difference in the enhancement layer. The MPEG-2 standard supports pyramid encoding.



Spatial Scaling for Video (2)

Layered Frequencies

Each 8x8 block is DCT-transformed and quantized. After quantization, the coefficients are stored in different layers; e.g., the first three coefficients in the base layer, the next three in the first enhancement layer, the next four in the second enhancement layer, etc.

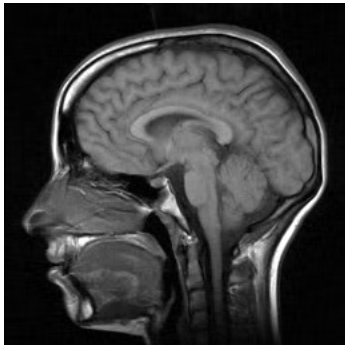
Spatial Scaling for Video (3)

Layered Quantization

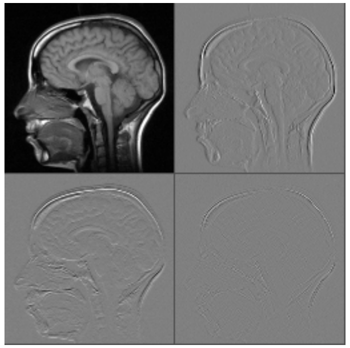
Each 8x8 block is DCT-transformed. The bits of the DCT coefficients are distributed over several layers. For example, the base layer may contain the first two bits of all coefficients (corresponding to a very coarse quantization), the first enhancement layer the next two bits of all coefficients, etc. Receiving more layers is then equivalent to choosing a finer quantization.

Spatial Scaling for Video (4)

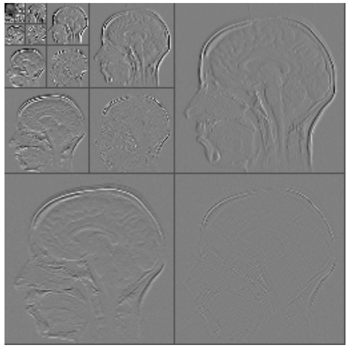
Layered Encoding of Wavelet-Transformed Video



Original frame



Wavelet-transformed, 1 iteration

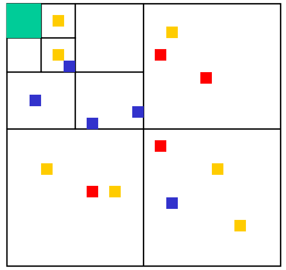


Wavelet-transformed, 5 iterations

Spatial Scaling for Video (5)

Layered Wavelet Encoding

The approximation (i.e., the low-pass filtered part) of a wavelet-transformed frame is stored in the base layer. The coefficients of the details (i.e., the band-pass or high-pass filtered part) are stored in enhancement layers in decreasing order of their absolute value.



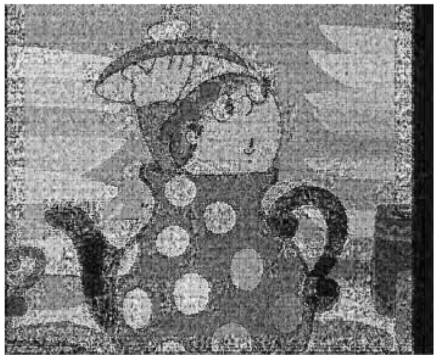
- base layer l_0
- enhancement layer l_1
- enhancement layer l_2
- enhancement layer l_3

Examples of Spatial Layering

Sample frames of a video sequence (a cartoon). Note that the images are shown at different quality levels in order to illustrate the characteristic artefacts.

Pyramid encoding

Layered DCT frequencies



Layered DCT quantization

Layered wavelet encoding

Scaling and Layering Summary

- Scaling techniques for digital video were developed in recent years and work well.
- Layering (i.e., the decomposition of the video into separate layers) is not yet available in most commercial video codecs.
- The integration of layer filters into internal network nodes of multicast trees is still a research issue.