

CHAPTER 6

Results from Simulations of Subband Video Coding

6.1 Quality and output rate

6.2 Error concealment

6.3 Statistics of the output rate

6.3.1 Simulation

6.3.2 Results

6.4 Discussion on quality of service

6.5 Summary

We have obtained encouraging results pertaining to the compression and the associated quality which can be obtained with subband coding of video. The results include both lossless transmission and transmission with simulated packet loss. The coding scheme as described in the previous chapter has been implemented in software and its operation has been simulated for two different image sequences. The sequences will be referred to as *Miss A* and *Miss B*. The first sequence is commonly called *Miss America*, so we shortened that name, and the second sequence got its name in recognition of Bell Communications Research by which it was provided. Each sequence contains a head-and-shoulders view of a talking woman, which has been captured by a single stationary camera without zoom or pan. Fig. 6.1 shows representative images of the two sequences used in the simulation and the statistics of the sequences are tabulated in Table 6.1. Generally speaking, the *Miss B* sequence contains more motion than does the *Miss A* sequence, which is manifested in the compression results.

Figure 6.1 Pictures representing the video sequences used in the simulations. (a) One frame from Miss A, (c) one from Miss B, taken from segment with little motion, and (d) a frame from a segment with substantial motion.

(a)



(b)



(c)



	<i>Miss A</i>	<i>Miss B</i>
Width	360	512
Height	288	480
Gray levels	256	256
Format	Progressive	Interlaced
Length	150 frames	300 frames
Frame rate	30 frames / s	30 frames / s
Bit rate	24.88 Mbps	58.98 Mbps

Table 6.1 Statistics of the two image sequences.

The interlaced sequence has been processed sequentially which affects vertically high-pass filtered subbands during periods of motion in the scene. However, these subbands would be equally affected if the processing was field-based since vertical correlation would be reduced during periods of low motion in the scene. The vertical-temporal sampling structure is quincux for interlaced sequences and it could be processed non-separably as such. It would however complicate implementation and appropriate filters would have to be designed.

In order to collect substantial amounts of statistics of the output rates, a simplified version of the subband coder was simulated as well. This coder was used to process ten different sequences, each two minutes in duration.

6.1 Quality and output rate

We have found that the simulations of the subband coder justify many of our assumptions regarding hierarchical source coding for packet video. Firstly, the quality of the reconstructed video is good and adequate for the intended services of video-telephony. The quality conveyed by the pictures in Fig. 6.2 is representative for the entire sequences: only slightly noticeable impairment seems present. Secondly, the compression yields an average data rate of 1.0 Mbps for *Miss A* which is roughly $\frac{1}{25}$ th of its input rate, and for *Miss B* the average rate is 3.1 Mbps, which is roughly $\frac{1}{19}$ th of the input rate.

The complete statistics produced by the simulations have been tabulated in Table 6.2, where average bit rates and signal-to-noise ratios (SNR) are given for the sequences together with standard deviations and minimum and maximum values. A bit rate is calculated as the number of bits needed to encode a single frame times the rate of 15 frames per second (half of the input's rate due to the temporal subsampling). Thus, the minimum rate is calculated from the minimum number of bits needed to encode any single frame in the sequence. The maximum rate is calculated analogously. The SNR is calculated as

$$SNR = 10 \log_{10} \left(\frac{256^2}{mse} \right), \quad (6.1)$$

where 256 is the intensity range of the input and *mse* is the mean-square error between input and output frames. Fig. 6.3 is serving as a reminder of the subband indices.

Figure 6.2 The pictures in Fig. 6.1 shown after compression: (a) Miss A, (b) and (c) Miss B.

(a)



(b)



(c)

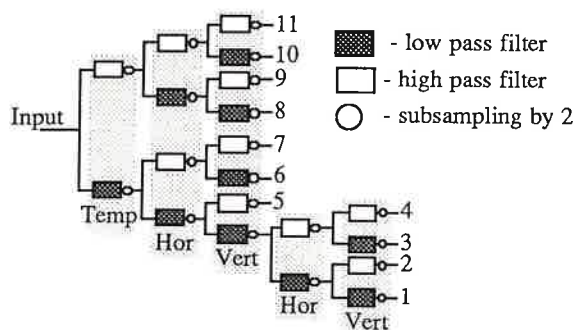


Figure 6.3 The eleven-band subband coding system.

According to our results, the argument that varying rate may give a constant quality appears to hold. For instance, *Miss B* is an especially active sequence, but still its SNR exhibits a standard deviation of only 0.45 dB! Actually, the fluctuations in the SNR are fully contained within 1.8 dB for *Miss A* and 2.3 dB for *Miss B*, which appears as a constant quality when the reconstructed sequences are viewed at video rate. It is worth restating, varying bit rate transmission can yield constant

	Miss A				Miss B			
	Mean	Std	Min	Max	Mean	Std	Min	Max
Band 1	341.6	2.1	338.0	347.6	848.1	21.7	775.0	886.9
2	128.8	3.2	119.9	136.7	365.0	38.4	234.0	425.9
3	158.3	6.8	136.5	169.4	441.0	45.7	283.0	514.5
4	24.5	5.3	10.1	33.5	61.7	19.4	13.0	96.3
5	81.9	4.5	70.3	89.5	281.1	140.6	108.6	903.4
6	46.3	13.8	10.0	76.7	78.0	32.0	2.3	129.9
7	0.6	0.0	0.6	0.7	0.2	0.3	0.0	1.0
8	103.1	73.0	8.7	285.5	693.8	357.5	65.4	1,908.1
9	11.7	7.4	0.7	30.3	228.9	184.1	8.6	961.6
10	44.1	12.8	25.8	75.8	63.1	19.2	20.0	103.7
11	70.6	3.0	62.1	77.2	14.9	10.7	1.3	54.2
Total	1,011.8	65.6	921.6	1,143.5	3,075.7	573.0	2,150.6	5,244.4
SNR	34.6	0.5	33.9	35.7	36.5	0.4	35.5	37.8

Table 6.2 Bit rates, in kilo-bits per second, of the subbands and the total output, as well as the signal-to-noise ratio of the reconstructed sequences (in decibels).

perceptual video quality, and the constant quality / varying bit rate paradigm of Section 2.1.1 is therefore assumed to be valid.

The subband coding yields a separation of the image information, which results in vastly different behavior of the output rates. In Fig. 6.4, the mean rate and standard deviation have been plotted for all subbands. As shown, the temporal low-pass bands are generally less variant than the temporal high-pass bands, with the exception of band 5 of *Miss B*. This band is vertically high-pass filtered, which yields a large variance due to the sequential processing in conjunction with the interlaced format of the video. Fig. 6.5 shows the temporal behavior (*i.e.*, burstiness and constancy) of three subbands. Owing to the temporal subsampling, the sequence length is halved. Note that the bursts last for several frames. Hence, only limited smoothing can be obtained by buffering, if the end-to-end delay should meet the real-time constraint (Section 2.4.1). In addition, the figure hints that rates from the subbands are correlated. For *Miss A*, the standard deviation of the total rate is 65.6 kbps. If the bands were independent, the sum-of-variances would give 76.5

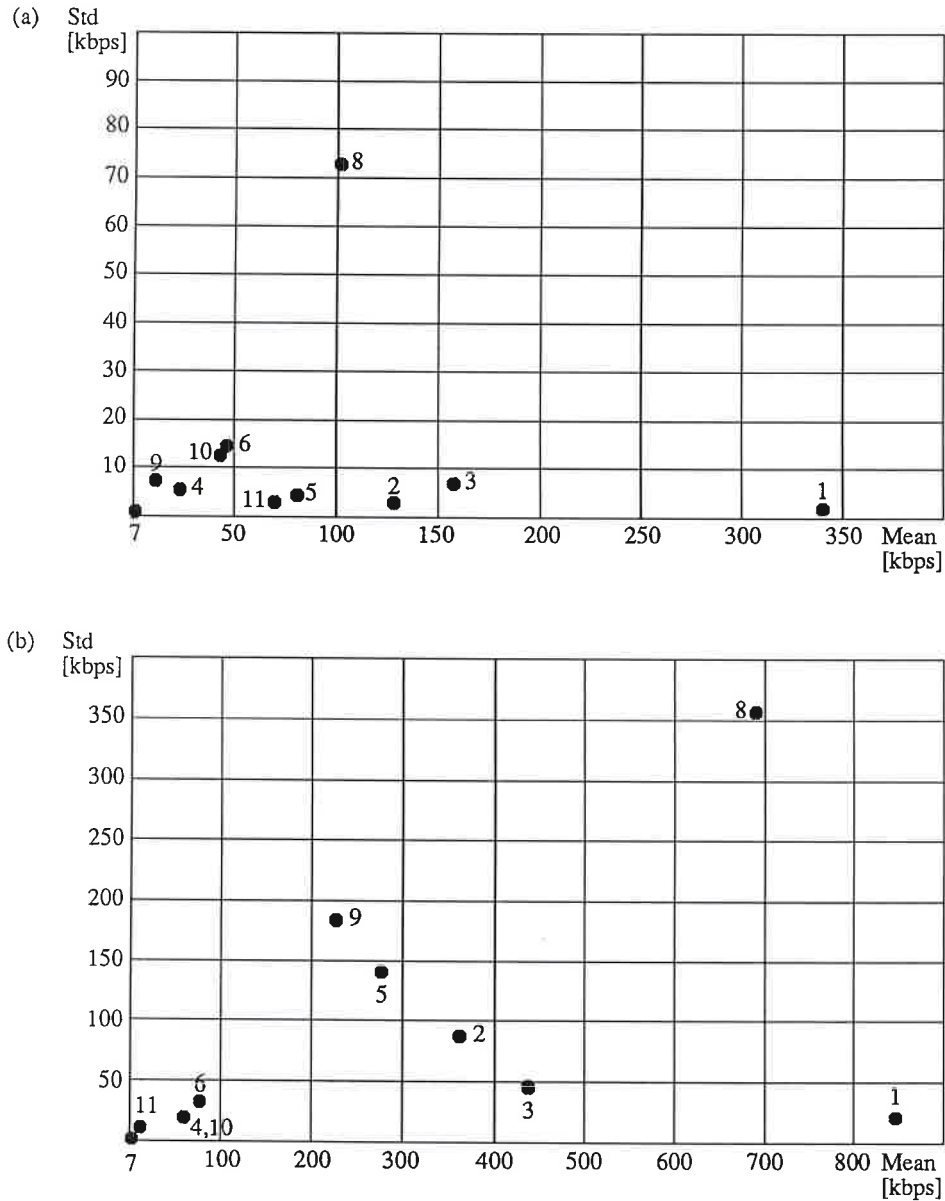
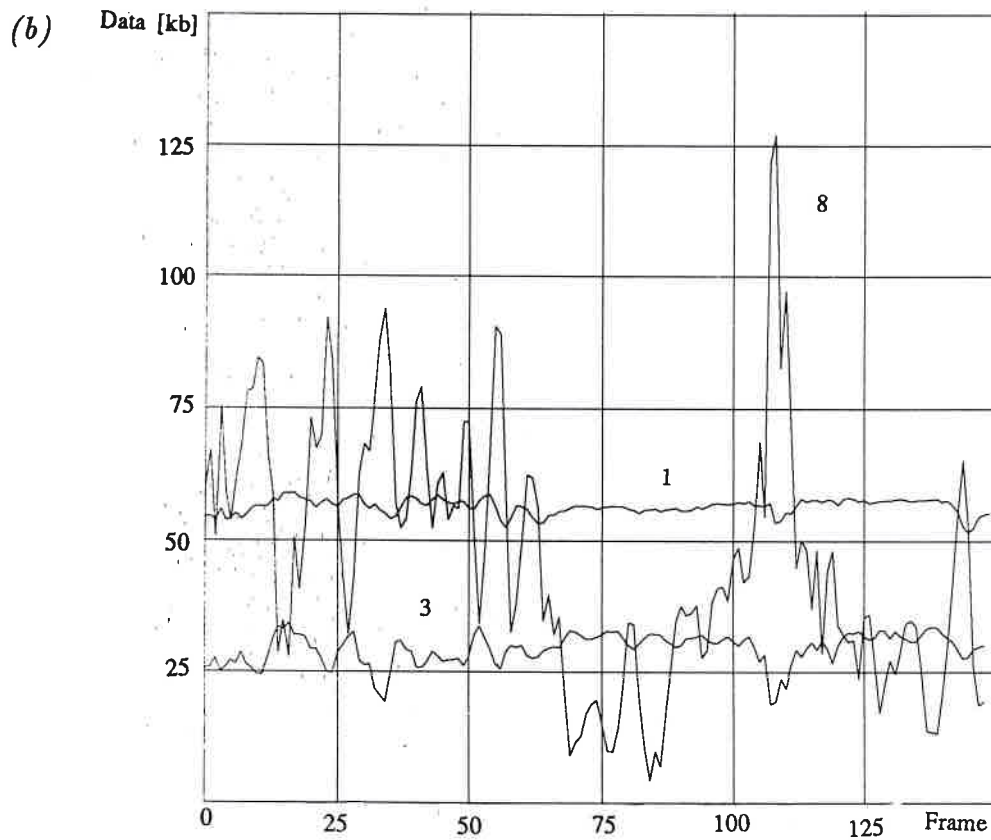
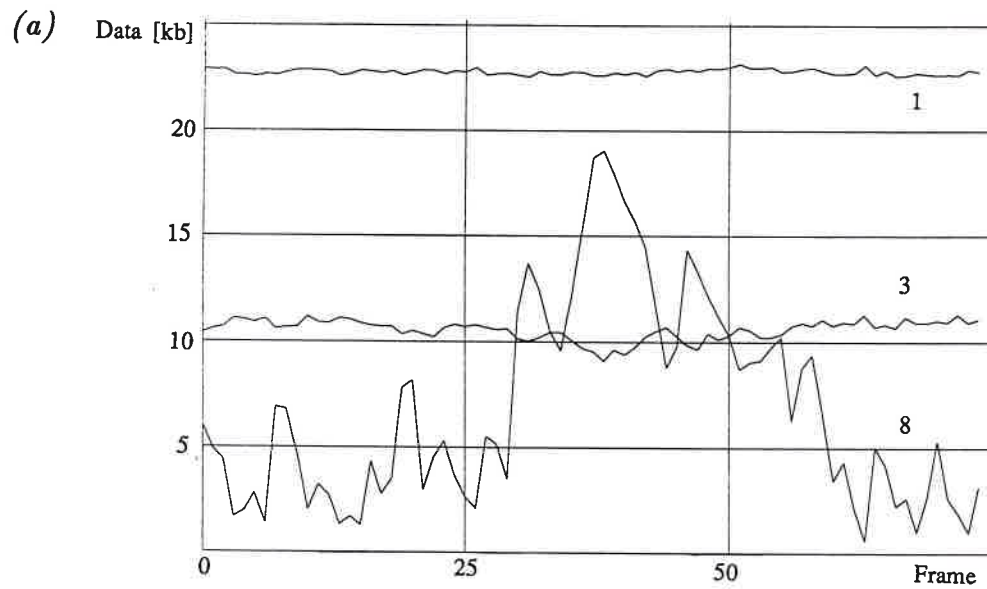


Figure 6.4 Mean rate and standard deviation of all subbands (for the band's indices see Fig. 6.3): (a) for Miss A and (b) for Miss B.

kbps as standard deviation. For *Miss B*, the standard deviation of the total rate is 0.57 Mbps and the sum-of-variances would give 0.43 Mbps as standard deviation. Hence, *Miss A* yields noticeable negative correlation between subbands while *Miss B* give more positive correlation between high bit-rate subbands. This may be attributed to the interlacing which makes bands 5, 8 and 9 to behave similarly. The

Figure 6.5 The number of bits per frame for three subbands (1, 3, and 8) for (a) Miss A and (b) for Miss B. (Note that the sequence length is halved due to the temporal subsampling.)



statistics of the rates from a subband coder is explored in Section 6.3 for several longer sequences.

Note that without the temporal subband analysis, the coding scheme could at best be seen as consisting of subbands 1 to 7 which would operate at full frame rate. From the simulation results above, this would correspond to an average rate of 1.6 Mbps for *Miss A* and 4.2 Mbps for *Miss B*. This gives an indication of the gain in compression which may be obtained by incorporating temporal subband division. The gain of about 30% is less than what could have been obtained by sophisticated motion compensation methods, but considering its simplicity, temporal subband analysis yields nevertheless a good improvement over intraframe processing.

Subband	<i>Miss A</i>		<i>Miss B</i>	
	Mean	Max	Mean	Max
1	2.9	2.9	1.2	1.3
2	7.6	8.1	2.7	4.2
3	6.2	7.2	2.2	3.4
4	39.9	94.4	15.8	73.1
5	11.9	13.9	3.5	9.0
6	21.1	93.5	12.5	177.3
7	1.5 [s]	1.5 [s]	4.9 [s]	~7 [s]
8	9.5	79.1	1.4	14.9
9	83.3	267.5	4.3	96.5
10	22.1	37.9	15.5	48.7
11	13.8	15.7	65.5	317.2

Table 6.3 Packetization delays for the subbands (in ms, except where indicated).

The packetization delays, corresponding to the subbands' output rates have been tabulated in Table 6.3 for 1024-bit packets with 976-bit information fields (as used in [LAZ85, LAZ87]). The table illustrates a potential problem with hierarchical coding: excessive packetization delay. When calculating the permissible end-to-end delay, the maximum packetization delay has to be used. Bands such as 7

and 11 of *Miss B* may be either multiplexed, or omitted permanently. However, bursty subbands which at any point in time may yield a low bit rate are potential problems, such as *Miss B*'s subbands 4, 6, 9, and 10 in the table. While multiplexing can eliminate the maximum delay, the added rates of these bursty bands may require unacceptable processing needs at rate-peaks. One solution would be to avoid transmission of a packet that required more than the allowed time for its packetization.

6.2 Error concealment

Throughout this presentation we have pointed to the issue of robustness to transmission loss. Our study covers error concealment through the use of visual redundancy, as discussed in Sections 2.3.2 and 5.4. It does not include the use of error correcting codes. We have found that subband coding offers a good framework for performing error recovery [KAR87, KAR88b]. Firstly, synchronization flags are added according to the method of Section 5.4 so that the location and number of lost packets are known. As pointed out before, only subband 1 requires some form of computed concealment of the erasures. The other subbands recover to near invisibility of the errors by setting the erasures equal to zero (the mean). For subband 1, the erasures were patched over by using the corresponding area from the previous frame. The performance of this error recovery method was simulated by randomly erasing 1024-bit packets. For lack of a better model, the simulated loss was taken to be uncorrelated over time and between subbands.

Although the visibility of errors is diminished when the sequence is viewed at video rate, compared to the scrutiny of individual frames, we will try to make our point with the representative pictures from *Miss B* in Fig. 6.6: in (a) five randomly chosen subbands, not including subband 1, have suffered loss of one 1024-bit packet

Figure 6.6 The effect of packet loss. (a) Five packets have been lost in all but subband 1. (b) One packet lost in subband 1, and five packets in the other subbands.

(a)



(b)



each, and (b) shows a case when subband 1 has been affected by one lost packet in addition to five packets for subbands 2 to 11. Note that the lost data appear in different locations in (a) and (b). Apparently, the method works well for areas without motion, but subband 1 leaves visible error in areas with motion. However, when the sequence is viewed at real video rate the errors are less visible. This is because the errors are randomly scattered with different spatial locations in each affected frame, and the loss may affect different combinations of the bands for each frame. By the latter, the appearance of the loss is not constant over time. Note that the single packet lost per frame in subband 1 corresponds to an average of 1.7% packet loss for that band, and the corresponding number for the five lost packets in the other bands is 2.4% (assuming equal priority transmission of these bands), which is considered high for most networks. Again, the information field of the packet was taken to be 976 bits.

The unsatisfactorily concealment of subband 1 at times of high motion can be eliminated by using motion estimation to find the appropriate area to be used as concealment, as described in Section 5.4.* The improvement of motion compensated concealment is illustrated in Fig. 6.7. In the test 10 scan-lines, appearing at the height of the eyes, have been erased in every fourth frame of the *Miss B* sequence. Then these lines are concealed by, firstly, the corresponding lines from the previous frame and, secondly, by the area corresponding to the estimated motion. There is thus no compression involved. The temporal discontinuities, caused by the erased scan-lines, are dramatically reduced by the motion compensated replacement and only a slight flickering is noticeable when viewed at video rate. This is caused by the slight spatial discontinuity at the upper and lower limits of the concealed area, which is still visible in Fig. 6.7 (b). Spatial filtering could further reduce the

* The work was done by Kamil Metin Uz.

Figure 6.7 Comparison of previous frame substitution (a) without motion estimation and (b) with motion estimation.

(a)



(b)



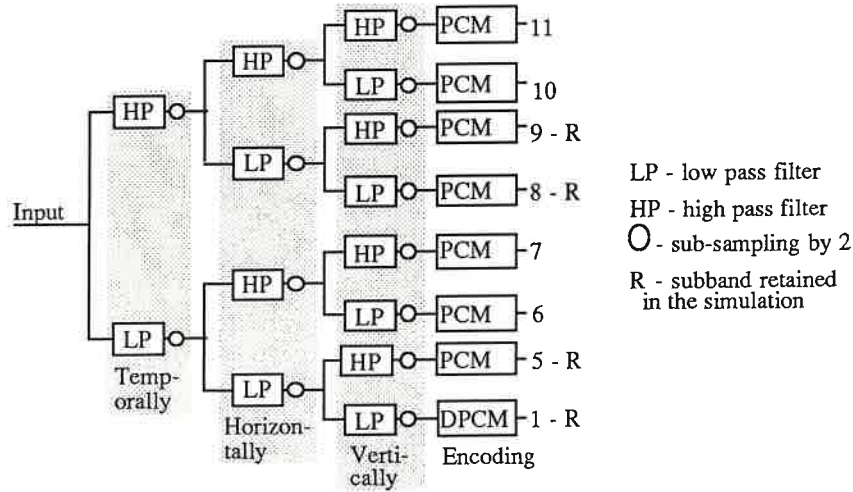


Figure 6.8 An eight-band subband analysis system. Note that the band indices are chosen to correspond to those of the eleven-band coder, shown in Fig. 6.3.

visibility of a concealment by smoothing along the border of the affected area.

6.3 Statistics of the output rate*

The coding method used in the previous simulation has also been implemented in a simplified version. The simplifications of the coding scheme have been necessary in order to reduce the amount of computation and thus making practical simulations with long video sequences. A block diagram of the subband division, the analysis, is shown in Fig. 6.8. The analysis has created eight subbands, which each has half the frame rate of the input sequence, and half of the spatial dimensions.

The temporal low-pass and high-pass filters are the same as those of Eq. (5.8). The spatial analysis filters are given in the z -transform domain by [LEG88b]

$$H_0(z) = \frac{1}{4}(-1 + 3z^{-1} + 3z^{-2} - z^{-3}), \text{ and } H_1(z) = \frac{1}{4}(-1 + 3z^{-1} - 3z^{-2} + z^{-3}). \quad (6.2)$$

* This work is part of a collaboration with Paul Douglas, which was reported in [DOU88].

The reconstruction of the signal, the synthesis, is the reciprocal system of the analysis, with the filters given by $G_0(z) = H_1(-z)$ and $G_1(z) = -H_0(-z)$. The system gives perfect reconstruction when there is no coding nor transmission loss. These new spatial filters were chosen to simplify the implementation since they have equal lengths and gain factors, but they do however have inferior characteristics compared to the filters of Eqs. (5.9) and (5.10).

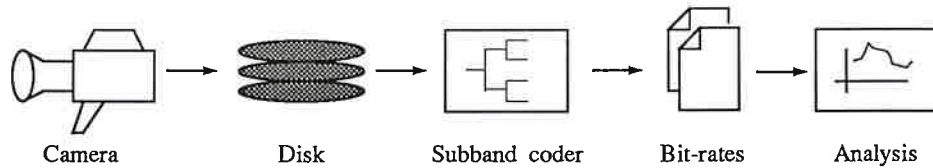


Figure 6.9 *The coding simulation: a sequence is acquired in real-time onto a storage device; the sequence is coded to yield the bit-rates, and the bit-rate statistics are computed and analyzed.*

6.3.1 Simulation

The set-up of the simulation is shown in Fig. 6.9. The camera used gives a composite (*i.e.*, NTSC) output with a vertical resolution of 300 lines. This signal is fed through an analog demodulator to obtain an analog luminance signal. The luminance signal is then digitized in an image processing system and the sequence is stored on a hard-disk. The disk storage allows sequences of length 3600 frames to be used, which corresponds to two minutes of video. Although, the sequences are of interlaced format they are processed sequentially. This causes the bands which are vertically high-pass filtered to fluctuate with the motion in the scene. Four of the eight subbands from the coder showed negligible importance in terms of image quality, which was also reflected in their bit-rates. These subbands were therefore excluded from the simulation. The retained subbands are 1, 5, 8, and 9 (see Fig. 6.8 for the meaning of the band indices).

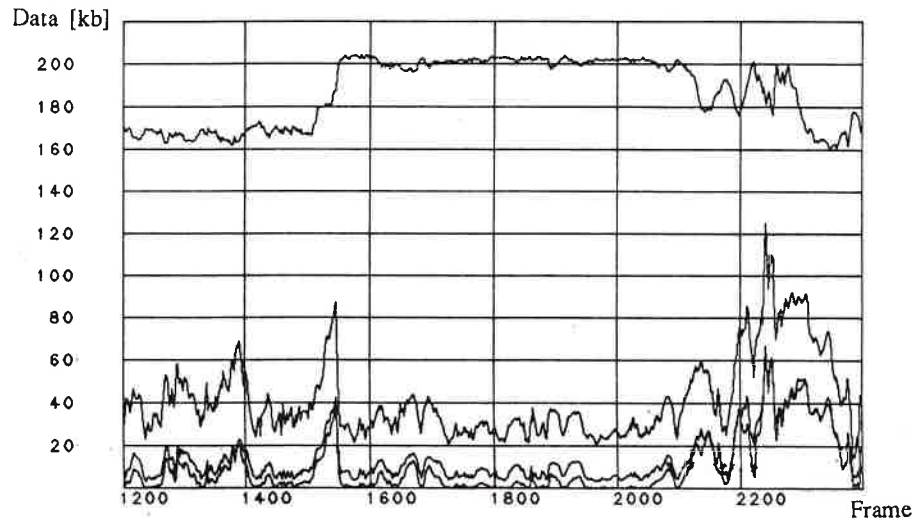


Figure 6.10 The bit-rates, in kbps, of subbands 1, 5, 8, and 9 for a third of the sequence Xuejie (see Table 6.5). Band 1 shows the highest rate followed by bands 8, 5 and 9. Due to the temporal subsampling, the bit-rates are only available for odd frames.

All the 10 sequences captured for the simulation contained the head and shoulders of a talking person in front of a stationary background. A different person appears in each sequence. The sequences also contained actions which would complicate the compression, both personal movement, such as hand waving, and camera changes, such as panning and zooming. The statistics collected consisted of the output bit-rates for each subband. The bit-rate was taken to be the number of bits per frame in a subband (15 values per second), resulting in a total of 1800 values for a subband and sequence. Note that the number of bits per frame is the smallest entity available since all images are coded in their entirety. The bit-rates of the subbands are shown in Fig. 6.10 for a segment of a sequence. The effect of the interlacing is evident in band 5, which is as variant as a temporally high-pass filtered subband.

The simulation had to be performed with a simplified coding scheme because of the high computational load. However, a concern with the simplified coding scheme could be that it does not capture the nature of the output rate from the more

BAND	μ		σ		σ / μ		Peak / μ		ρ_{cross}
1	114	161	7	5	1.1	1.1	.06	.03	.91
5	19	12	9	6	3.2	2.9	.50	.46	.99
8	46	45	24	13	2.7	1.9	.51	.29	.96
9	15	8	12	7	4.2	4.2	.80	.81	1.0

Table 6.4 Comparison between the output from a complex subband coding scheme and its simplified counter part. The values given in the shaded fields pertain to the complex coder. The parameters are the time-average (μ), the standard deviation over time (σ), the standard deviation to time-average ratio, the peak to time-average ratio, and the cross-correlation (ρ_{cross}).

sophisticated eleven-band coder. This, however, has been shown not to be the case. The *Miss B* sequence of 300 frames has been used for simulation by the simplified coding scheme and its complex counter part, and their corresponding subbands have been compared. Table 6.4 shows the results from the system in Fig. 6.8 compared with that of Chapter 5. (Subband 1 in the scheme of Chapter 5 was taken as the sum of the four lowest bands.) As the table clearly shows, the schemes have similar characteristics, and the cross-correlation between corresponding bands is very strong. In conclusion, it is believed that the simplified coding scheme captures the statistic properties of the more sophisticated scheme.

6.3.2 Results

The time-average (μ), the standard deviation over time (σ), the standard deviation to time-average ratio, and the peak (p) to time-average ratio are tabulated for all ten sequences in Table 6.5. The peak to time-average ratio indicates the upper limit of the variations. It is worth noting that the bands with the highest time-average vary the least, *i.e.*, they have the lowest standard deviation (or peak value) to time-average ratio. Also, the peak rates appear to remain below five times the mean even for the most varying subbands.

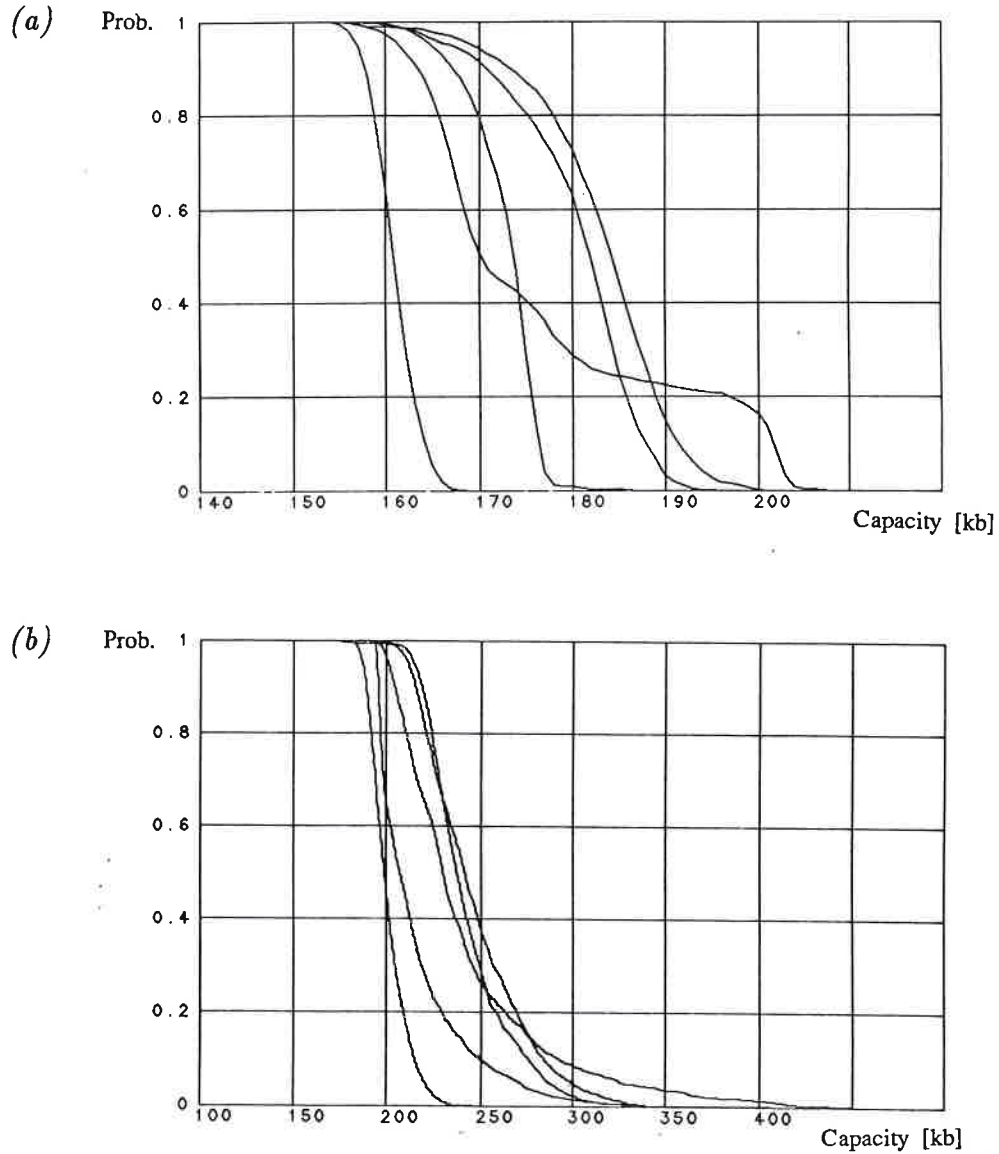
Seq.	Band 1				Band 5				Band 8				Band 9				Total			
	μ	σ	σ/μ	ρ/μ	μ	σ	σ/μ	ρ/μ	μ	σ	σ/μ	ρ/μ	μ	σ	σ/μ	ρ/μ	μ	σ	σ/μ	ρ/μ
Gail	183	7.3	.04	1.1	10.3	6.1	.59	5.4	43.2	14.3	.33	2.6	5.6	7.5	1.3	12.0	242	23.1	.10	1.7
Wesley	180	6.7	.04	1.1	13.0	7.4	.57	3.7	43.9	16.0	.37	2.9	9.0	8.6	.96	5.6	246	27.7	.11	1.6
Xuejie	177	14.1	.08	1.2	13.8	11.3	.82	5.2	41.3	18.2	.44	3.0	9.8	13.5	1.4	7.8	242	41.5	.17	1.9
Betsy	173	3.7	.02	1.1	7.2	6.5	.90	5.6	32.4	14.6	.45	3.0	3.5	6.2	1.8	10.3	216	24.8	.11	1.6
Jackie	161	2.4	.02	1.0	5.8	2.5	.44	2.9	32.3	6.9	.21	1.8	1.8	2.1	1.2	7.6	201	10.2	.05	1.2
Peter	196	12.4	.06	1.1	10.1	6.4	.63	6.5	41.1	13.4	.33	2.9	5.4	7.1	1.3	13.0	252	22.0	.09	1.8
Salman	172	9.5	.06	1.2	11.6	8.3	.71	11.1	42.0	14.7	.35	5.2	7.7	8.6	1.1	13.3	234	29.3	.13	2.6
Colin	175	7.2	.04	1.1	17.8	11.7	.67	4.8	58.0	19.4	.33	2.8	13.7	13.4	.98	8.1	264	40.8	.15	1.0
Habib	170	8.0	.05	1.1	12.0	8.6	.71	4.3	38.0	16.8	.44	2.8	8.3	10.1	1.2	7.1	228	31.9	.14	1.6
Gunnar	162	6.0	.04	1.1	6.7	4.7	.70	7.3	29.2	13.5	.46	3.9	3.7	5.5	1.5	16.9	202	19.5	.10	1.8

Table 6.5 Statistics from the simulation.

Fig. 6.11 (a) shows the probability function $P(s_i(n) > c)$, where $s_i(n)$ is the bit-rate sequence of a subband 1 of sequence i , and c is a capacity level [KOG81]. Although subband 1 of most sequences decreases smoothly, one of them, Xuejie, exhibits a plateau with nearly constant probability for a range of capacity levels. This behavior is attributed to a long term trend in the bit-rate (*i.e.*, non-stationarities), which could be the result of a pan or a zoom during the sequence. Note that it cannot be attributed to a burst of high activity during a short period of time, since the peak to time-average ratio is modest for subband 1 of all sequences. In Fig. 6.11 (b) the same probability function is shown for the total rate of the same five sequences. The plateau behavior has now disappeared due to the multiplexing of the subbands. The same behavior is shown by the second half of the sequences which are not included in the figure. The probability function will be used next to evaluate possible gains by statistical multiplexing, in a similar way as used in [VER86, VER88c].

A high ratio of the standard deviation to the time-average indicates that an

Figure 6.11 The probability that the bit-rate of a sequence exceeds a given capacity (given in kilo-bits) for (a) band 1 of the first five sequences, and (b) for the total rates of the same sequences.



advantage can be obtained by statistical multiplexing over a fixed capacity allocation. In order to quantify this advantage the following measure was used:

$$Adv = \frac{\sum_{i=1}^N C(s_i(n))}{C(\sum_{i=1}^N s_i(n))}, \quad (6.3)$$

where $s_i(t)$ is the bit-rate function of a source, and N is the number of sources.

$\mathcal{C}(\cdot)$ gives the capacity limit which the sequence, given as argument, exceeds by a probability of 0.1% (*i.e.*, $P(s_i(n) > \mathcal{C}(s_i(n))) = 0.001$). In other words, the advantage measure was taken to be the ratio of the sum of the capacity limits for each source to the capacity limit for the total of the sources. This is plotted in Fig. 6.12 and it corresponds well with Table 6.5. Since the simulation only contained ten sources (*i.e.*, sequences) the higher number of sources were generated by reusing the ten with random time-displacements.

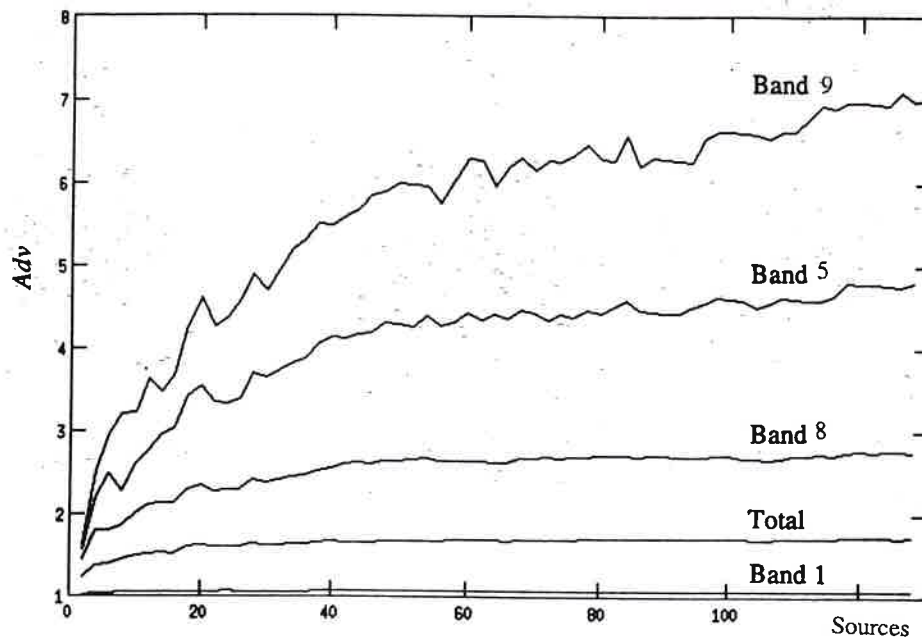


Figure 6.12 The statistical multiplexing gain as a function of the number of sources. The plots, given bottom-up, are for band 1, the total of all four bands, band 8, band 5, and band 9.

The results in Fig. 6.12 indicate that subband 1 can be well handled on the basis of fixed capacity allocation. Subbands 5, 8 and 9, however, are most advantageously multiplexed statistically since the gain in channel utilization can be as high as 7 times within their respective traffic class. For the total rate the statistical multiplexing gain is limited to 2.5 which is reached for about 40 sources. The reason for the lower gain is that subband 1, which is nearly constant, dominates in

the total rate. It has a bit-rate of nearly three times the rate of the other bands together (see Table 6.5). Also, some averaging has already taken place between the subbands of a source, but it may not contribute to much of the low multiplexing gain. This smoothing is evident from the strong correlations between bit-rates of the subbands: for the sequence Wesley (see Table 6.5 for its statistics) bands 1 and 5 show cross-correlation of -0.59, bands 5 and 8 are correlated by 0.89, 5 and 9 by 0.98, and 8 and 9 by 0.88 (all values are for a zero-lag). Many of the bursts in the higher bands are offset by a slightly lowered rate in subband 1. This may also be intuitively clear: an increase in temporal activity will result in lowered spatial activity (*i.e.*, motion blur in the camera) which gives the discussed behavior. Note that due to real-time constraints, it is not feasible to delay subbands 5, 8 and 9 in respect to one another in order to get substantially lower cross-correlation; a lag of 10 frames (*i.e.*, 0.67 s) is necessary to bring the correlation below 0.5 for any two of the bands. In the following section we will use these simulation results in a discussion on resource allocation.

6.3 Discussion on quality of service

Let us define a traffic class by its allocated resources (capacity guarantees) and its transmission priority. In most networks a traffic class would consist of only one of the two, but we include both for the generality of our discussion. Then we may differentiate two transmission scenarios: firstly, one traffic class for all subbands and sources, and secondly, one separate traffic class for each subband. In the former, statistical multiplexing gain can be achieved between separate sources and between subbands of a source, as shown in the previous section. With separate classes, only the subbands in a class are multiplexed, thus barring any gain that can be achieved between different subbands. However, for the sake of error resilience, different traffic

classes appears to be a necessity for packet video transmission.

Clearly to avoid any packet loss allocated capacity has to exceed the peak rate of the source (for a buffer free network). For a highly varying source, this may however lead to wasted resources and unnecessarily high blocking probabilities if an admission control is enforced. Consider a sequence like subband 8 of Salman in Table 6.5 which has a very high peak to time-average ratio; up to five times the average rate has to be available in the network in order to accept the transmission and a large part of that capacity is only going to be used for a split second.

The DPCM encoded subband 1 will require a traffic class with the highest possible transmission priority and, if possible, guaranteed capacity according to its maximum rate. It is a fortunate circumstance that the band, with its low peak to time-average ratio, lends itself to such an allocation without wasting resources. The other, more varying, subbands can be transmitted with lower priority since they are not necessary for the continuity of the session; only its quality. So, an alternative resource allocation for these subbands, which is more viable than capacity assignment according to their peak rates, is to allow limited packet loss and to provide a bound on the loss during the periods that the rate exceeds the guaranteed capacity. The bound should reflect the importance of carried information. In order to provide such a bound, the bit-rates of the subbands have to be modeled so that one can determine the amount of capacity necessary to meet the bound on packet loss. Bit-rate models which may serve this purpose have been suggested in [MAG87, SEN87, HUA88, MAG88, DOU88] and are discussed in Appendix A. To illustrate, an allocation for some subbands could guarantee capacity equal to their mean rates. Thus, they will have to compete for unassigned portions of the network capacity when needed. An even lower allocation-class would correspond

to subbands which have no guaranteed capacity and therefore have to vie for their entire capacity needs. Priority assignment and amount of guaranteed capacity for a subband could be made according to mean, variance and visual importance. The former two are illustrated in the plot of Fig. 6.4. The visual importance will be discussed next and may be determined by aid of Fig. 6.13.

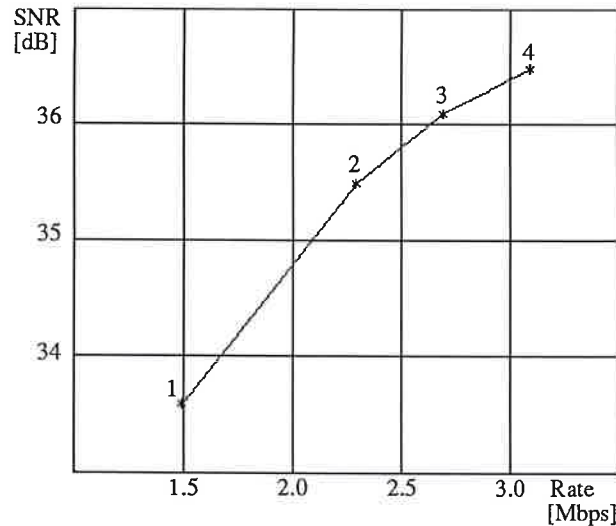


Figure 6.13 The SNR as a function of bit rate. Case 1 consists of subbands 1 and 8; case 2: 1 through 3, and 8; case 3: 1 through 6, and 8; and case 4 includes all subbands.

In Fig. 6.13, the output rate is plotted against the SNR for selected subsets of all subbands of the eleven-band coder. This can serve as an indication of how degradation can vary with available transmission capacity. The pictures in Fig. 6.14 represent the quality associated with the four cases in Fig. 6.13. Cost/quality tuning (see Section 3.3.2) may be achieved by omitting less important subbands in order to yield a desired output rate. The possible range of quality and output rates corresponds to those of Fig. 6.13. Note that packet loss is a temporary quality decrease along this curve, while cost/quality is a permanent one for the session. To summarize, for the quality of service, as described in Section 3.3.2, the video quality under ideal transmission conditions is set by the number of subbands transmitted

Figure 6.14 Pictures representing the cases in Fig. 6.13. (a) The original and (b) through (e) show cases 1 to 4.

(a)



(b)



(c)



(d)



(e)



and it would be described by the curve in Fig. 6.13. The allowed deterioration under lossy transmission will depend on the traffic classes that the subbands have been assigned and the sophistication of the error recovery function.

6.5 Summary

It appears that subband coding is a promising way of doing hierarchical source coding. The simulation results indicate that a good video quality can be obtained at compression ratios around 20 times with a low complexity coder. The simulations also verify that the quality level remains virtually constant when the transmission rate is allowed to vary. Error recovery by concealment was simulated and shown to be a good remedy for packet loss. However, for subband 1 the concealment should be based on motion compensated replacement from the previous frame, since without it the replacement performs insufficiently in areas of motion.

The transmission rates of subband coded video have been investigated for long video sequences. Ten sequences, each two minutes in duration, have been acquired and used to produce bit-rate statistics. The analysis of the output rates shows that the subbands have vastly different characteristics, ranging from nearly constant rate to highly varying. The basic statistic properties have been tabulated and the gain achievable with statistical multiplexing has been calculated. It shows that statistical multiplexing may yield a channel utilization of up to 7 times that possible with fixed multiplexing for a highly varying subband.

Finally, in light of all these results, the quality of service was discussed. It is shown that subband 1, which exhibits a near-constant output rate, can be assigned capacity resources according to its peak rate without leading to any substantial loss of channel utilization. This is of great importance since this subband is the most vital for the quality of the reconstructed video. Any other subband should have partial or no guaranteed capacity, depending on its importance and statistical behavior. Transmissions of subsets of all subbands were also simulated. These cases show that a lowered transmission rate can be obtained simply by excluding subbands from transmission which also reduces the quality level.