An example of the partition filtering process (see Section 8.7.2) used for partition down-sampling is presented in Figure 9.14.

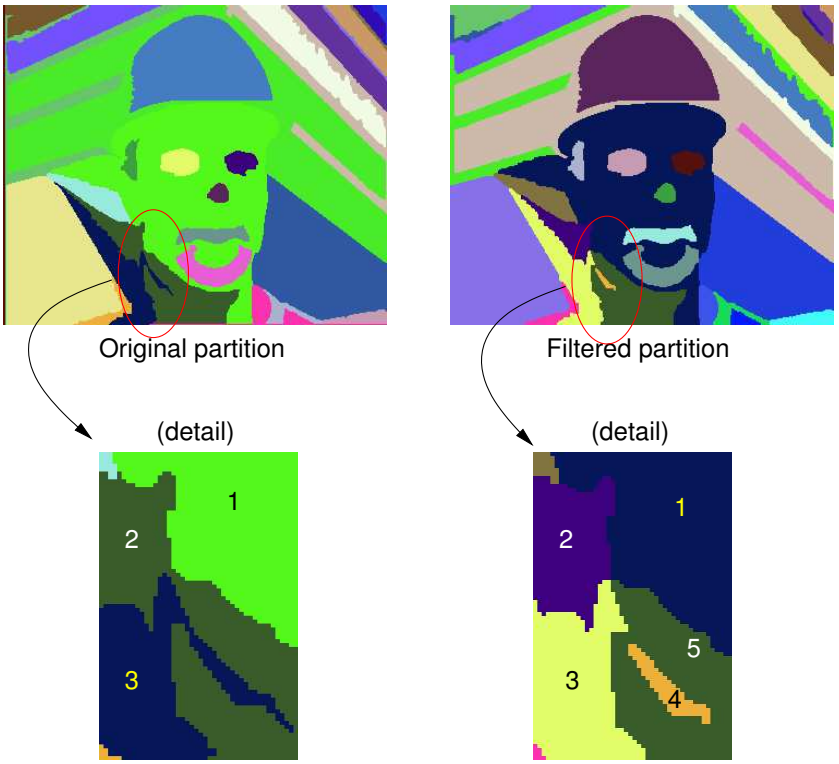

Figure 9.14: Example of partition filtering for down-sampling

Note that this filtering process can result in the elimination of small regions and the introduction of new regions in the filtered partition. In the example, thin connections in region #2 and #3 are removed and the resulting connected components are relabeled, resulting in two new regions labeled as #4 and #5. Some minor changes in the contour positions can be observed in the details in Figure 9.14. These changes are very small and are not relevant for the coding step as they do not noticeably affect the homogeneity of the regions.

In the Projection step, the double projection construction (see Section 8.3) is done similarly as in the PSNR scalability case. In Figure 9.15 an example of this process is given for the frame #93 of the *News* sequence. Differences arise from the fact that in the Spatial scalability case, base and enhancement layer projected partitions do have different resolutions. Thus, the previous frame base layer partition has to be up-sampled before it can be merged with the previous frame enhancement layer partition. Similarly, when constructing the final base layer projected partition, the projected partition has to be down-sampled in order to produce a partition at the desired base layer resolution.

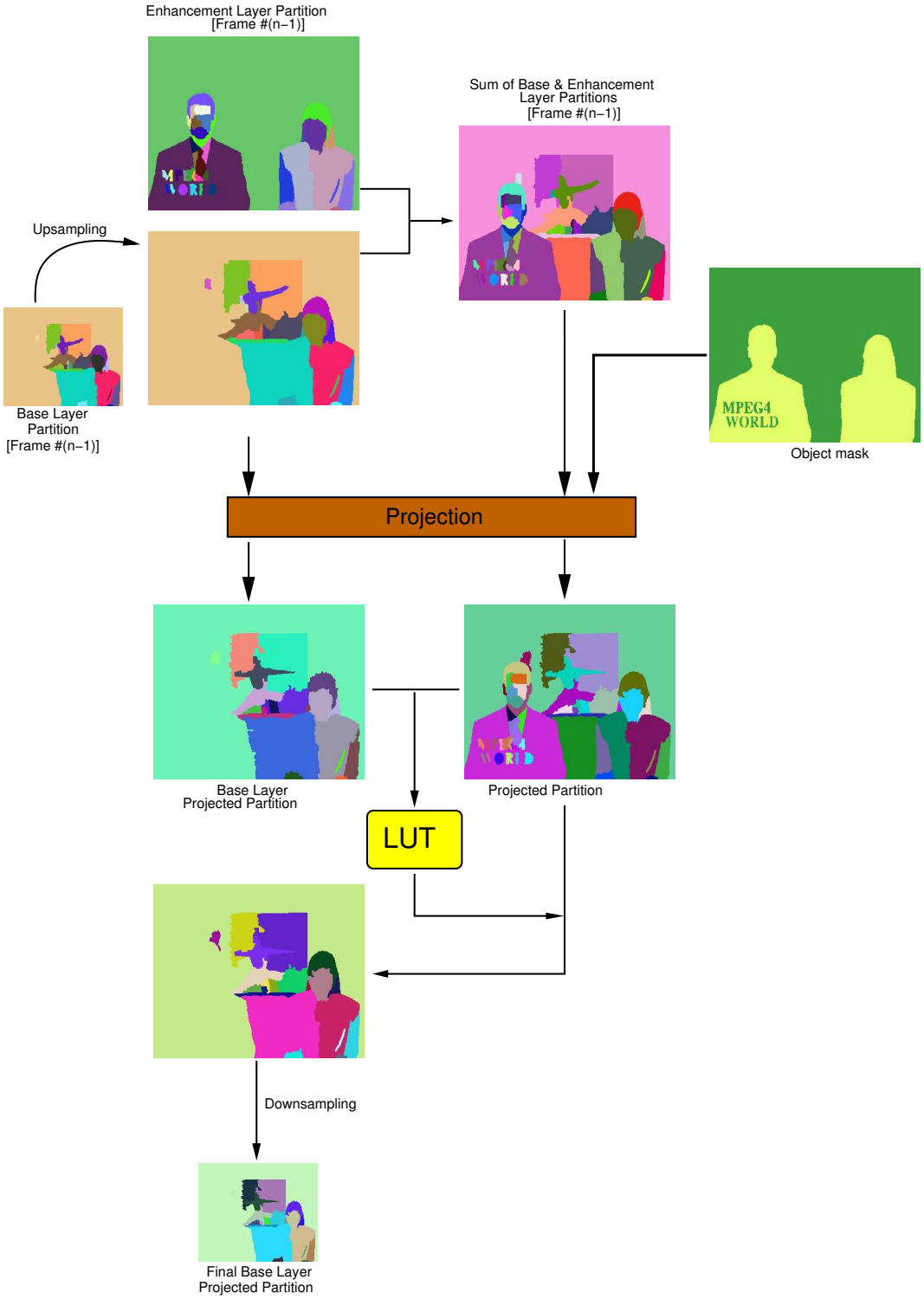

Figure 9.15: Partition compensation process in Spatial scalability

Results for Spatial scalability

Figure 9.16 compares the performance of the scalable encoder with the MoMuSys implementation of MPEG-4. As in the PSNR scalability case the performance of XSESAME can not match the performance of the MoMuSys encoder.

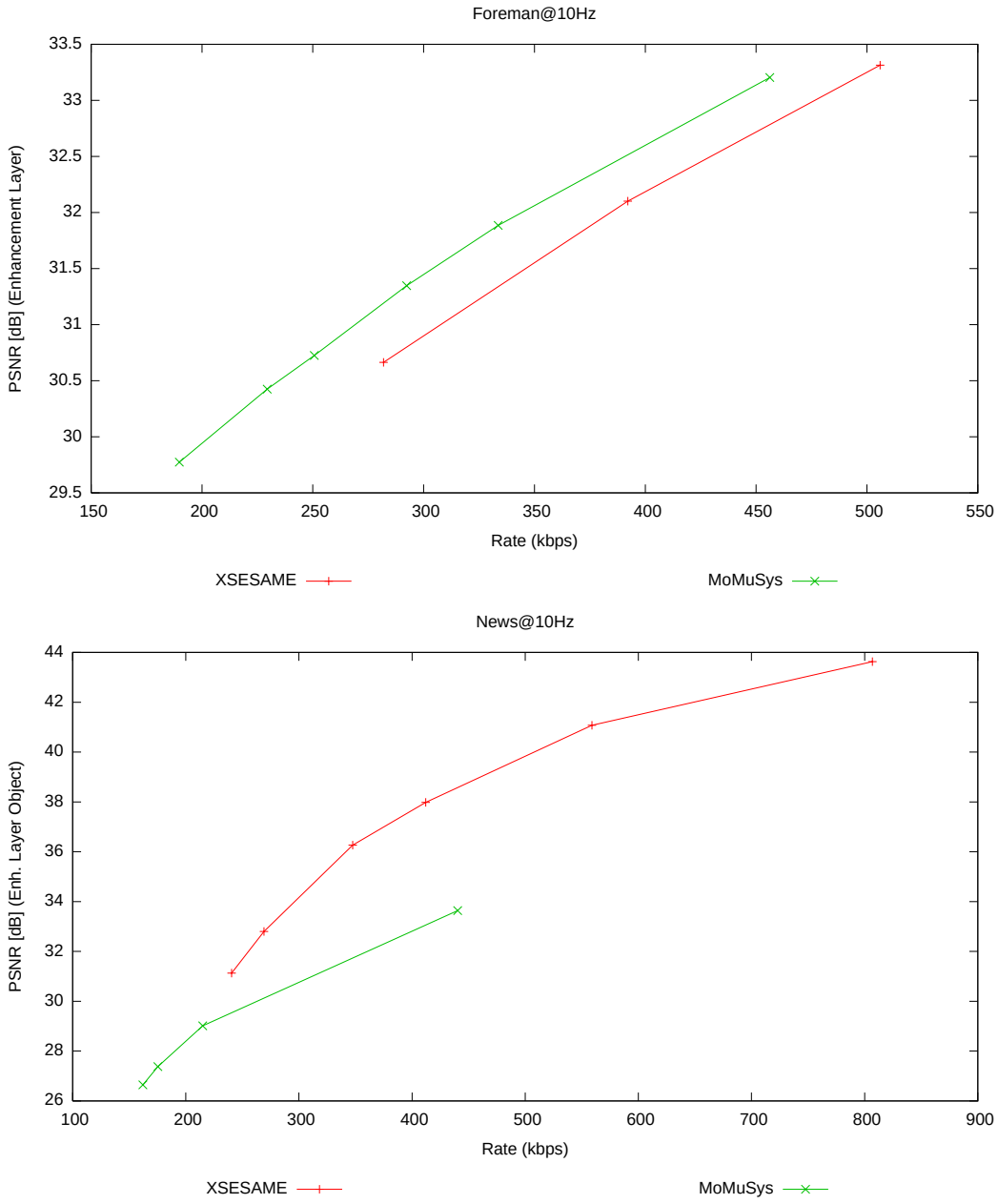

Figure 9.16: Comparison between XSESAME and MoMuSys coders for full frame (top) and object functionalities (bottom) spatial scalability

Examples of decoded images in the Spatial scalability mode can be seen in Figures 9.17 and 9.18 for full frame mode (*Foreman* sequence, CIF@10Hz). The figures represent for each frame the base (left) and enhancement layer (right) decoded images, as well as the corresponding partitions. The base layer (QCIF) is encoded at 128 kbps while the full enhancement layer (CIF) is encoded at 512 kbps.

Full frame mode - Summary of results: For the base layer, the same conclusions as in the PSNR case concerning the number of regions of the partition and the selected texture coding techniques can be derived. At the enhancement layer, inter-frame techniques are even more used than in the PSNR scalability case to the detriment of layer intra techniques. This is because the interpolation step necessary for layer intra techniques makes these technique less efficient. The inverse relationship between the bit-rate and number of regions already observed in PSNR scalability is also observed in Spatial scalability. Increasing the bit-rate results in less regions per frame.

In Figures 9.19, 9.20 and 9.21 there are some examples for spatial object scalability mode. Images have been obtained from the *News* sequence (QCIF@10Hz). The base layer (QCIF) has been encoded at 128 kbps and the enhancement layer (CIF) has been encoded at 384 kbps, with all the bit budget in the enhancement layer used to code the selected object (the two anchor people in the foreground).

Full frame mode - Summary of results: As in the PSNR case, the number of regions in the enhancement layer partition in the object functionalities mode does not depend as much on bitrate as in the full frame mode. This is because the forced inclusion of the contours of the object determines to a great extent the number of regions in the partition.

For texture coding, layer intra techniques are slightly more used than in the full frame mode due to the increased number of regions, but its use is not as important as in the object functionalities mode for PSNR scalability.

Figure 9.17: Example of Spatial scalability: Foreman sequence (CIF) coded at 10Hz, full frame scalability, 128 kbps (base layer, QCIF) + 512 kbps (enhancement layer, CIF). The figure shows base (up-sampled by a factor two) and enhancement layers in the first row and the respective partitions in the second row for frames #0 and #60

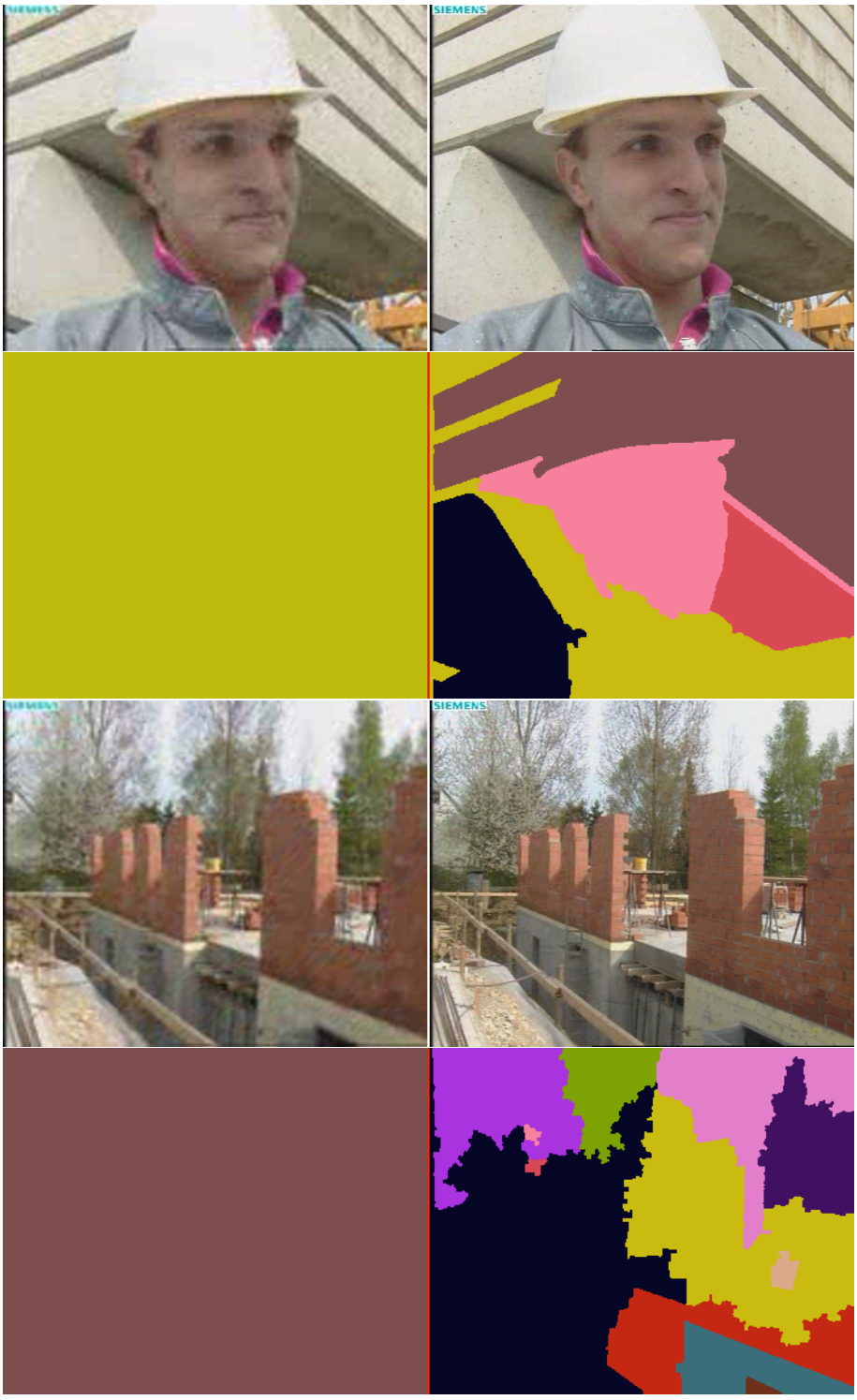

Figure 9.18: Example of Spatial scalability: Foreman sequence (CIF) coded at 10Hz, full frame scalability, 128 kbps (base layer, QCIF) + 512 kbps (enhancement layer, CIF). Figure shows base (up-sampled by a factor two) and enhancement layers in the first row and the respective partitions in the second row for frames #120 and #297