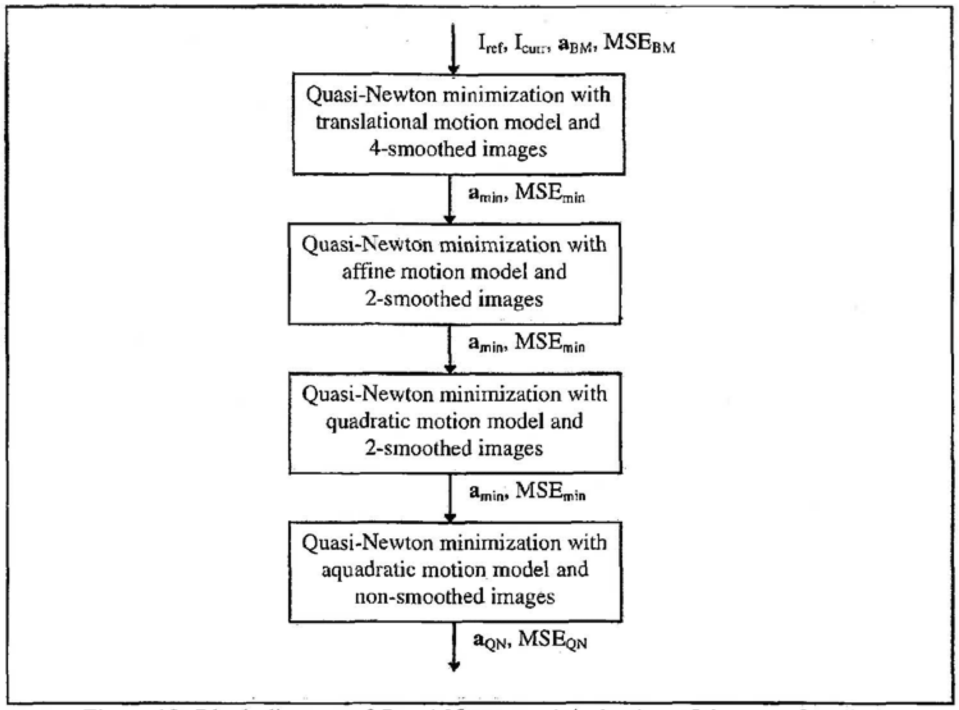

Figure 19: Block diagram of Quasi-Newton minimization of the error function

After each minimization, error function value (evaluated in non-smoothed image domain) is compared to the lowest value obtained so far. If the value has decreased the motion parameters are stored and the next minimization begins at this point in the parameter space. Otherwise the motion parameters of the last minimization are discarded and the next minimization is initialized by the parameters yielding the smallest value of the error function. In the first three minimizations bilinear interpolation of the subpixel values is applied to speed up the minimization. Cubic interpolation is then used only in the last minimization to guarantee maximum performance.

After the termination of the Quasi-Newton minimizations error function value MSE_{QN} is compared to the one after the minimizations of the Taylor expanded error functions MSE_{Taylor} and the motion parameters yielding smaller value are chosen for the motion parameters for this block.

4.3 Merging

Splitting and motion estimation results in a large number of blocks (usually) with good predictions. However many of these blocks could have been predicted with the same set of motion parameters without affecting the quality of the prediction too much. To remove this redundancy, blocks with similar motion characteristics are merged together to form larger segments. Following subsections describe how to decide which blocks are to be merged and how to obtain motion parameters for the merged segment knowing the parameters for the original ones.

4.3.1 Finding the final segmentation

To find the optimal segmentation, Lagrangian cost

$$L = \frac{D + \lambda R}{P} \tag{50}$$

of the image is to be minimized [Kar96]. In (50), the square error between the original and the decoded frame is denoted by D, the number of bits needed to code the frame is denoted by R and the number of pixels in the frame is denoted by P. λ is the Lagrangian multiplier specifying the cost of one bit with respect to the distortion measure. In order to find the minimum in L, all the possible combinations of segments should be checked. To avoid the computational complexity an approximative method is presented in this section.

Let ΔL_{kl} be the reduction in Lagrangian cost when two segments S_k and S_l are merged to form S_{kl}. ΔL_{kl} can be calculated by

$$\Delta L_{kl} = (L(S_k) + L(S_l)) - L(S_{kl}) \tag{51}$$

The segment S_k can be coded in three different modes (i.e. in INTER, INTRA or UNCHANGED mode). The cost of the segment is assumed to be the minimum of the costs of these three modes, since it minimizes (50). Thus the Lagrangian costs for all the modes have to be evaluated

In calculation of the INTER mode cost the number of bits for the motion information is needed. However this information will be available only after the merging and motion model adaptation. To overcome the problem a constant number of bits per motion parameter is assumed. By simulations a value of three bits per parameter is found to be a good approximation. Prediction error of the segment is coded with 8x8 DCTs similarly as in ITU's H.263 recommendation [ITU95]. Bits for the prediction error coding are added to the bits for motion information to yield $R(S_k)$. $D(S_k)$ is the squared error between the coded segment and the corresponding segment in the current image. These values are then placed to the equation (46) and $L_{INTER}(S_k)$ is evaluated.

Calculations of the Lagrangian costs for INTRA and UNCHANGED mode are simpler. INTRA mode cost $L_{INTRA}(S_k)$ is obtained by coding the segment with the codec's intra coding method (in this implementation that is the one specified in H.263 recommendation) and the cost for the UNCHANGED mode $L_{UNCHANGED}(S_k)$ is equal to the mean squared difference between the reference and the current image in the segment area. Lagrangian cost for the segment S_k is then

$$L(S_k) = \min\{L_{INTER}(S_k), L_{INTRA}(S_k), L_{UNCHANGED}(S_k)\} \quad (52)$$

INTRA and UNCHANGED mode costs for combined segment S_{kl} are calculated simply by adding costs for the same mode of the segments S_k and S_l. To obtain Lagrangian cost for the INTER mode, joint motion parameters must be found, motion estimated prediction must be build and the prediction error must be coded.

Merging procedure is initialized by building an adjacency graph of the segments. Each node in the graph is associated with the Lagrangian cost of the segment and each arc with the reduction in cost when adjacent segments are merged. Figure 20 shows an example of such a graph before and after segments S_k and S_l have been merged together.

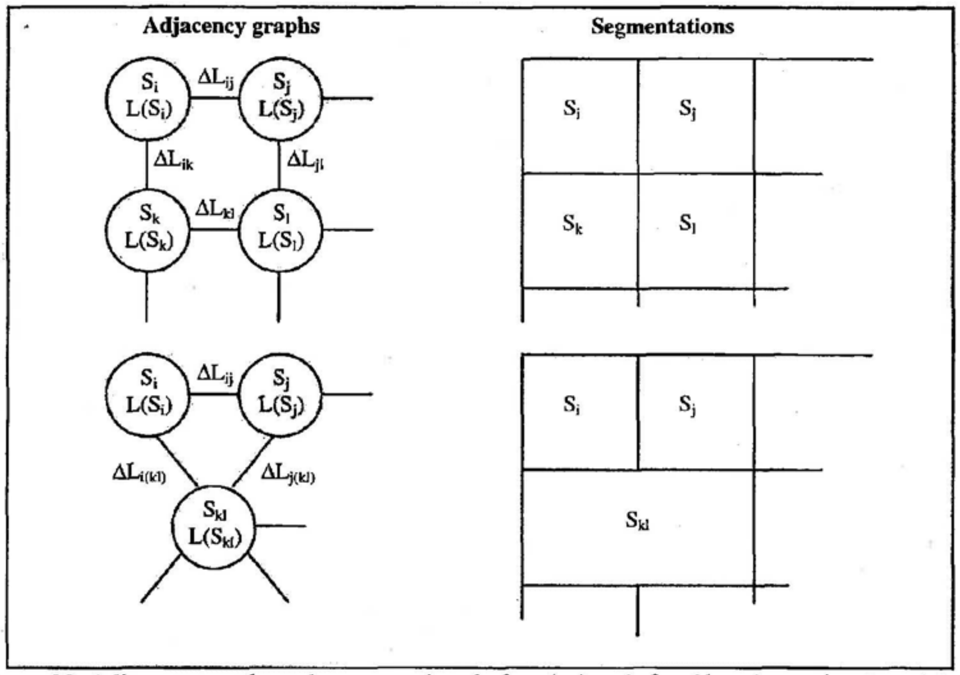

Figure 20: Adjacency graphs and segmentations before (up) and after (down) merging S_k and S_l

After building the adjacency graph and calculation of the initial $L(S_k)$s and ΔL_{kl}s, arc with the largest reduction in Lagrangian cost is chosen and the segments associated with it are merged. Adjacency graph is updated and the algorithm proceeds by choosing again the arc which yields the largest reduction in cost. This is iterated until there exists no longer a pair of segments which yields reduction in cost when merged. Algorithm is summarized in figure 21 and in the figure 22 there is shown resulting segmentation for the first frames of both Akiyo and Carphone test sequences.

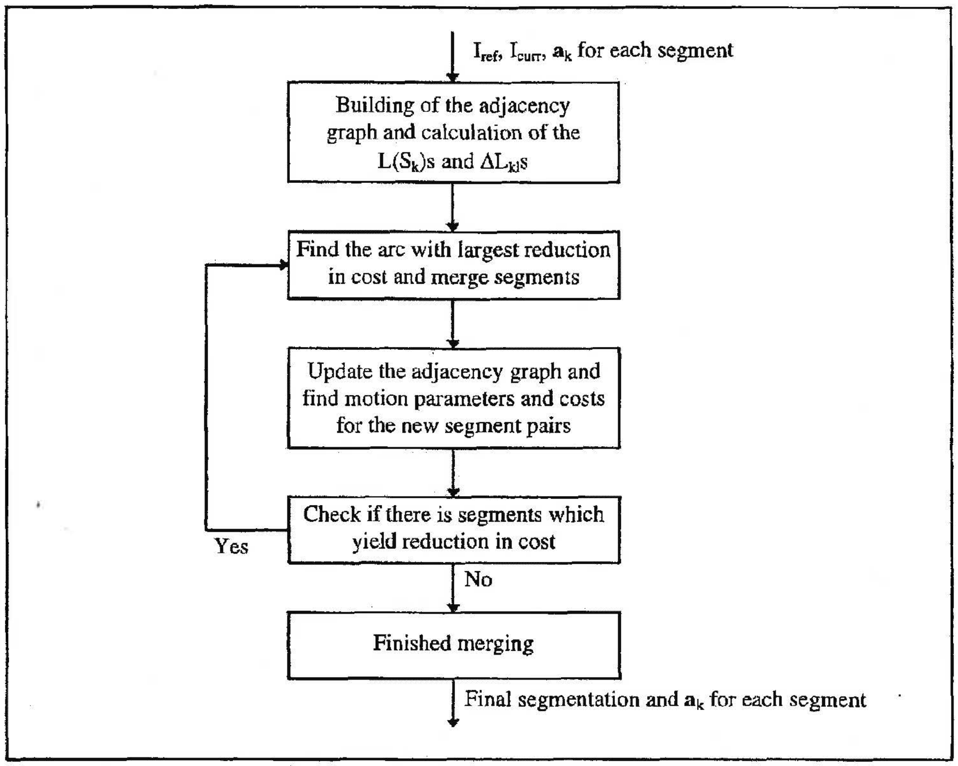

Figure 21: Block diagram of the merging

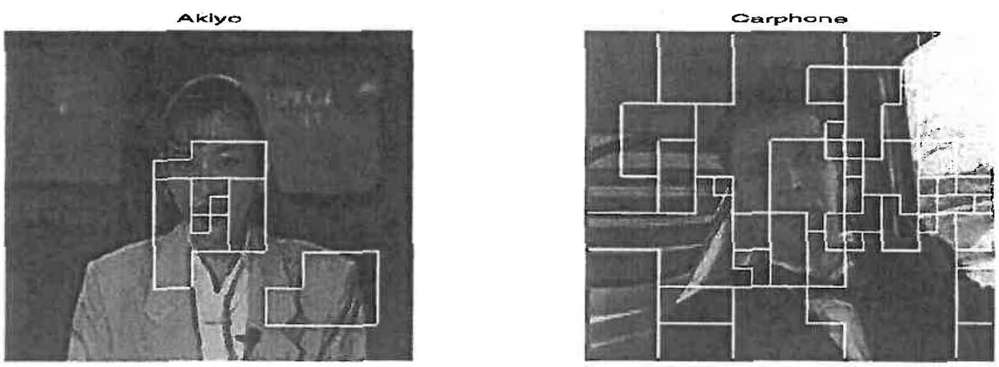

Figure 22: Resulting segmentations after merging for the first frames of Akiyo and Carphone sequences

The main complexity of the algorithm presented here comes from finding the motion parameters for segment pairs which are candidates for merging and from the calculation of the INTER mode Lagrangian costs for these segment pairs. In the next section a low complexity method for finding the joint motion parameters for two segments is shown. To decrease the burden caused by the calculation of INTER mode costs, following assumptions were made. First it is assumed that if adjacent segments S_k and S_l are both in INTRA or in UNCHANCED mode (i.e. $\min\{L_{INTER}(S), L_{INTRA}(S), L_{UNCHANGED}(S)\}$ is either $L_{INTRA}(S)$ or $L_{UNCHANGED}(S)$) the resulting segment S_{kl} will be in the same mode and thus there is no need to calculate $L_{INTER}(S_{kl})$. It is also assumed that the lower the increase in prediction error due to merging the higher the reduction in Lagrangian cost will be. An estimate for the increase in prediction error is obtained from the calculation of the joint motion parameters as will be seen in the following section. This information is used to choose whether or not the Lagrangian cost for a segment pair S_{kl} is to be calculated.

Assume that S_{kl} is a segment which was formed by merging two segments S_k and S_l. To update the adjacency graph, joint motion parameters for S_{kl} and all its neighbours have to be calculated. Also the approximated increases in prediction error have to be obtained for all the segment pairs. Letting ΔG_{min} be the smallest increase in prediction error, the INTER costs are calculated only for those segment pairs for which $\Delta G < T \cdot \Delta G_{min}$. Otherwise the INTER mode Lagrangian cost is set to infinity. In the implementation a value of 3.0 is used for the multiplier T. With this value, on average less than a half of the segment pairs need to be examined and the best candidate for merging is found in over 95% of the cases [Kar96].

4.3.2 Finding the joint set of motion parameters for two segments

Given two segments S_k and S_l, the problem is to find a set of motion parameters minimizing the error function for the merged segment $S_{kl} = S_k \cup S_l$. The approach is very similar to the one in Taylor expanded minimization of the error function. Again Taylor expansion of the reference image I_{ref} is employed. The expansion is done around the points

$$\begin{cases} x' = x + d_x(\mathbf{a}_k, x, y), & y' = y + d_y(\mathbf{a}_k, x, y), \quad (x,y) \in S_k \\ x' = x + d_x(\mathbf{a}_l, x, y), & y' = y + d_y(\mathbf{a}_l, x, y), \quad (x,y) \in S_l \end{cases} \tag{53}$$

in image coordinate space. Using (35) and (36), the problem can be written as

$$\begin{cases} \mathbf{E}_k \mathbf{a}_{kl} = \mathbf{y}_k \\ \mathbf{E}_l \mathbf{a}_{kl} = \mathbf{y}_l \end{cases} \tag{54}$$

where $\mathbf{a}_{kl}$ is a set of joint motion parameters and the matrices $\mathbf{E}_k$ and $\mathbf{E}_l$ and vectors $\mathbf{y}_k$ and $\mathbf{y}_l$ are given by (38) and (39), respectively. Substituting both $\mathbf{E}_k$ and $\mathbf{E}_l$ by their QR decompositions (54) becomes

$$\begin{cases} \mathbf{Q}_k \mathbf{R}_k \mathbf{a}_{kl} = \mathbf{y}_k \\ \mathbf{Q}_l \mathbf{R}_l \mathbf{a}_{kl} = \mathbf{y}_l \end{cases} \tag{55}$$

Multiplying both sides of the equations above by the transposes (inverses) of the orthogonal $\mathbf{Q}$ matrices yields

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.

WHAT WILL YOU BUILD? | **sales@docketalarm.com** | **1-866-77-FASTCASE**

fastcase®
Smarter legal research.