CCITT Study Group XV
Working Party XV/1
Specialists Group on Coding for Visual Telephony

Doc #120
June 1986

Source: UK

Title: A Hardware Aspect of Multi-Pel Predictors

From a hardware viewpoint, the predictors under consideration for the nx384kbit/s video codec can be separated into two types.

Type 1. The prediction value for each pel is obtained by using the value of one pel in the previous decoded picture. This is the situation where no filtering is applied in the loop and motion compensation vector components have integer pel values.

Type 2. The prediction value for each pel is obtained by using some combination of the values of more than one pel in the previous decoded picture. This covers the schemes which incorporate a spatial filter in the loop and it also applies to motion compensation with non-integer pel displacements.

This contribution examines an additional consequence of adopting predictors of the second type instead of the first.

Since the first type can also be regarded as a special case of the second, then if the hardware is designed for the second type, it can, in principle, also perfom the first type of prediction.

Picture Store Considerations

Only the picture store in the decoder loop is considered below. The encoder one may be simpler if the optional motion compensation is not incorporated, otherwise it will have the same or greater complexity than the decoder store.

The decoder picture store is required to provide a prediction from several pels in the region pointed to by the motion vector. In the case of a cross shaped spatial filter five pel values are used and for vectors with fractional components four pel values might suffice. Under some circumstances the contribution from some of these four or five pels may be zero, but in the general case they are all required.

In a scheme which does not have motion compensation it is possible to access several pels by shortening the delay of the picture store and adding a tapped delay line with pel and line stores. Unfortunately, the introducton of motion compensation prevents this simple procedure as the process can no longer be pipelined. This means that the picture store must give n parallel accesses, where n is 4 or 5. Possible methods are:

1. To increase the speed of the store so that n read cycles can occur instead of one.
2. To have n standard speed stores in parallel.

Neither of these is an attractive proposition. However, bearing in mind that the motion vector range is restricted, it is possible to have most of picture period delay in one standard speed store and implement the multiple accesses on auxilary storage - about 32 lines for ±15 motion vectors. So two more methods are:

3. To have an auxilary store with n times faster read access.
4. To have n auxilary standard speed stores.

The fourth method listed above is feasible and represents a incease in total storage of 32% for n=4 and 42% for n=5.

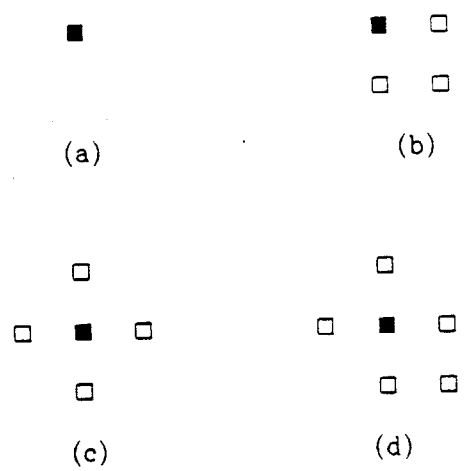
The discussion above has made no assumptions on the number of different ways the values of the n pels are combined to form the prediction. Clearly, if a shaping filter with invariant characteristic is used then it is possible to place this filter before the picture store. Then only one pel read access is required to get each prediction pel.

If the pel combination function has several clauses, then it is possible to perform all of these operations in parallel and put each result in a separate picture store. A spatial filter with two characteristics - one for zero motion vector and another for non-zero for example - would need 2 picture stores. This compares favourably with the 5 picture stores of method #2 above. On the other hand, for motion vectors with half pel resolution on each component there are 4 clauses and hence 4 picture stores would be needed - the same as for method #2.

It is worth noting that for the two cases of multi-pel predictors used above, several of the pels sites are the same, and, of course, they contain the one pel needed in a single pel predictor - see figure 1. Therefore, if required for experimentation, the flexible hardware could be constructed to access all 6 pels. The lowest overhead to achieve this is 52% more storage using method #4, plus the an arithmetic unit which can perform the calculations for each of the different combinations.

Conclusion

Multi-pel predictors in which the pels are combined in more than one way lead to extra complexity in the picture stores, which although not trivial can be accommodated if necessary to obtain acceptable picture quality. Simulation results have not yet produced convincing results one way or the other. Therefore it is proposed that the hardware specification should initially be for a single pel access picture store, but with the proviso that this will be changed later if required.



(a) Single pel prediction
(b) Fractional component motion vectors
(c) Cross shaped spatial filter
(d) All pels in (a), (b) and (c)

The filled square represents the pel pointed to by the integer parts of the motion vector components.

Figure 1