A new block matching algorithm specially proper for a large search area, is proposed. The algorithm uses spatial correlation of motion field and hierarchical search. Motion vectors of causal neighboring blocks can be credible MV candidates of the current block, if the motion field has high spatial correlation. However, complex or discontinuous motion can be found better by using a hierarchical search. Our scheme consists of the higher level search that uses spatial correlation for continuous motion and adopts hierarchical structure for random or complex motion, and the lower level search for the final motion vector refinement. Compared to the conventional hierarchical BMA, the scheme reduces the local minimum phenomenon. It also alleviates the error propagation due to the use of spatial correlation when the complex motion is involved. Simulation results show that the proposed algorithm drastically reduces the computational complexity to about 2.9% of that of FS-BMA, with the minor PSNR degradation of 0.25dB even in the worst case.

Keywords: motion estimation, spatial correlation, block matching, hierarchical search, center-biased search, video image compression.

1. INTRODUCTION

Recent development in video image compression have been the main focus in many applications such as multi-media transmission, videophones, teleconferencing, high-definition television (HDTV), CD-ROM storage, etc. The effective method generally used for reducing the temporal redundancy existing in video sequences is motion estimation/compensation techniques. Most of current video communication systems and standards such as CCITT H.261 (px64)[1] and MPEG[2] employ a block matching scheme[3] for motion estimation. A straightforward block matching algorithm (BMA) is the full search BMA (FS-BMA) which searches every possible displaced candidate within a search area in order to find the block with minimum difference. However, the implementation of FS-BMA requires massive computational efforts or huge hardware cost to achieve significant speed. As a result, many fast BMA’s [4-7] have been developed to alleviate the heavy computation of FS-BMA. Basically, these algorithms reduce the number of searching operations by selectively checking only a small number of positions under the assumption that distortion increases monotonically as the searching location moves away from the best matching position. However, this assumption does not always hold in practice and thus a motion vector often gets trapped in a local minimum. Therefore, none of the fast search algorithms mentioned above would always give reliable performance. Furthermore, as the search area increases, the performance of these algorithm deteriorates drastically because their scheme cannot solve the higher local minimum problem. In order to improve the performance, several fast algorithms have been proposed [8-11]. In the zero waiting-cycle algorithm (ZWCA) [8], two-step hierarchical search technique has been proposed by using a block subsampling scheme. Because of its regularity and simplicity on the data flow and matching process, the ZWCA is appropriate for VLSI implementation. However, the matching performance tends to be unstable due to strong limitation of computing complexity. Liu’s methods (Liu’s) [9] used motion field and pixel subsampling techniques. Though Liu’s methods [9]
provide more robust matching performance than that of the ZWCA, their computational complexity is somewhat increased. In the SBMA(S+T) [10], a set of good MV candidates is constructed and chosen from them the one which satisfies a certain spatio-temporal correlation rule so as to reduce the computing complexity of FS-BMA. Due to its excessive dependence on spatio-temporal correlation, however, the SBMA(S+T) has the risk of bringing about the error propagation phenomenon in a large search area. This is because the algorithm doesn’t use the spatio-temporal correlation efficiently.

A number of MV estimation schemes using the spatial and temporal correlations of MVs have been studied. The main idea is to select a set of initial candidates from spatially and/or temporally neighboring blocks and choose the best one as the initial estimate for further refinement. Theoretically, the initial estimate can be obtained by using an autoregressive (AR) model. However, a simple model has been adopted in experiments by choosing only one candidate as the initial estimate. Then, the refinement has been performed by doing the local search, for example FS-BMA, 3SHS, and etc., with reduced search area. In order to reduce the computational burden and acquire a correct MV in the local search, the accuracy of the initial estimate should be high. For the video sequences dominated by continuous motion, it is very effective to exploit the spatial correlation in order to reduce the local search area because there is strong spatial correlation between the MVs of adjacent blocks. However, if motion field is random or nonhomogeneous between the MVs of adjacent blocks, it rather results in error propagation problem to use the spatial correlation. In this case, it is very desirable to adopt the hierarchical structure. Hierarchical search gives the good performance for the video sequences with random or complex motion because it can find the global motion direction approximately in the higher levels. Our goal in this work is to develop the algorithm that exploit both spatial correlation and the merit of hierarchical search in a computationally simple way, maintaining a similar rate-distortion performance to that of FS-BMA, for the video sequences with random or complex motion as well as continuous motion.

This paper proposes a new fast hierarchical motion estimation algorithm which is especially suitable for a large search area. It uses not only the spatial correlation of motion field, but also adopts the structure of hierarchical search in order to predict an approximate motion direction. At the higher level, the global motion direction can be found by concurrently checking the four MVs of causal neighboring blocks and the locations with fixed distance in the whole search area. At the lower level, with the MV chosen in the higher level as the search center, the MV refinement procedure is performed so as to find the final accurate MV close to true MV. Our proposed scheme outperforms other fast algorithms for various test image sequences and its computational requirement is lower than other fast algorithms.

This paper is organized as follows. In section 2, we present a fast hierarchical search algorithm using the spatial correlation. Then, experimental results are provided and compared with the existing schemes in section 3. Conclusions are given in section 4.

2. FAST MOTION ESTIMATION ALGORITHM USING SPATIAL CORRELATION AND HIERARCHICAL SEARCH

In most video sequences, motion field may be dominated by continuous, random, or complex motion. If an object moves continuously, the MVs of blocks included in the moving object will be similar to the MVs of their spatially adjacent blocks. This property is very useful for initial estimate. The merit of the property is demonstrated in Fig. 1. We call the property motion continuity. The MV of block A can be predict from those of the spatially adjacent blocks because the strong spatial correlation exists between these MVs caused by motion continuity. Using this motion continuity, the computational requirement of motion estimation can be reduced considerably, compared with that of FS-BMA. The MV with minimum MAD is chosen as the approximate direction of current block among four MVs of the neighboring blocks. Subsequently, the local motion search with the small search area can be performed with the chosen MV as a search center in order to find an accurate MV. Therefore, using motion continuity, we can not only find the MV of current block accurately but also the computational complexity can be reduced very much.
However, motion continuity is always not useful. On the contrary, to use motion continuity may result in the degradation of performance. In case that the motion of current block is complex or random, the MV of current block is not similar to the MVs of spatially neighboring blocks any more. If the MV of current block is chosen among the four neighboring MVs and local motion search proceeds in the reduced search area with the chosen MV as a search center, the probability of being trapped in a local minimum will increase. This is the reason that motion estimation using motion continuity can not keep the good performance in image sequences having random or complex motion any more. Hierarchical search algorithm is very useful for searching random or complex motion with a small amount of computation. Hierarchical block matching provides reliable estimates of the true MVs. Due to the adaptive parameters, the hierarchical block matching method is able to cope with large displacements and to provide the true motion vector field with higher resolution and accuracy. Hierarchical block matching algorithm has two kinds of patterns. The first pattern is that with the fixed measurement block, the subsampled searching points are checked in the fixed search area of each level. Three step hierarchical search is one example with this pattern. In the second one, measurement block size is varied in each hierarchy level. At the higher levels, the measurement block of larger size is used with subsampling, and low-pass filtering reduces the risk of converging to a local minimum. The filtering should be reduced from one level to the next level of the hierarchy, in order to get a sufficiently accurate estimate at the last level. The measurement block size is reduced at the lower levels. Of course, lower hierarchy level is, smaller the corresponding search area is. Basically, hierarchical block matching searches the whole search area in the first level. Therefore, this method has the merit of looking for random or complex motion efficiently in a large search area. However, in order to keep the performance close up to that of FS-BMA, the hierarchical block matching needs a large amount of computation. So this method is not sufficiently fast enough to cover a large search area. Our scheme uses both the motion continuity and the merit of hierarchical search algorithm so as to track continuous motion as well as random or complex motion accurately.

2.1 The proposed fast hierarchical search algorithm using spatial correlation

The hierarchical block matching of the above-mentioned first pattern is applied to our scheme. The reason is that the fixed measurement block, such as 16 x 16 macroblock, is used at each level without low-pass filtering for subsampling. The approximate motion of current block is predicted in the first step by considering continuous motion as well as random or complex motion. At lower levels, with the predicted MV as the search center, an accurate MV is found within much smaller fixed search area than the whole search area. In the first step of our scheme, in order to predict the motion direction of current block, the MVs of four causal neighboring blocks and the searching points located with proper distance regularly within the whole search area, are exploited as the MV candidates of current block. Fig. 2 shows the MV candidates in the first step. The first candidate set consists of the MVs for continuous motion. The second one consists of the MV candidates for the case that the spatial correlation of motion field is not useful any more because random or complex motion are dominant around the current block. In Fig. 2, R x R is the whole search area, 63 pels x 63 pels in our scheme. D depends on the search area of the lower step. The MV close to true MV of current block is chosen among these MV candidates by comparing MADs. That is to say, the MV with the smallest MAD is used as the search center of the next steps. Two kinds of MV candidate sets have reciprocal help and improve the performance. In case that the spatial correlation is considered only, if the MV of current block is not similar to the MVs of causal neighboring blocks because of random or complex motion, the risk of error propagation increases and performance deteriorates. In this case, if the second MV candidate set is adopted in the first step of our scheme, although the MV of current block is not correlated spatially with the MVs of adjacent blocks, the correct motion direction of current block can be found among the second MV candidate set. Additionally, this results in the effect that the spatial correlation of next block is also improved. Therefore, the proposed algorithm can have almost the same performance as FS-BMA.

The dominant factor of computational requirement is the number of searching points. The computational complexity in the first step of the proposed algorithm is the sum of 4 searching points which is the MVs of causal neighboring blocks, and the element number of the second candidate set which is
For example, if R is 63 pixels and D is 8 pixels, the element number of the second candidate set is 49. Thus, in the first step of our scheme, 53 MADs calculation are needed. This demonstrates that the approximate motion direction can be found through a small amount of computation. The chosen MV with minimum MAD among this candidate set may be somewhat different from the true MV of current block. In order to compensate this error, local search is performed with a small search area, [-D+1 ~ D-1]. In this lower step, two step search is used for the reduction of computational complexity. Even though two step search has small search area, the probability of being trapped in a local minimum is still high. Therefore, we propose the center-biased two step search proper for our scheme as the lower level search method.

2.2 The two step center-biased algorithm in the lower level.

This search method must be made considering the following two possibilities. As the first one, if there are strong spatial correlation between adjacent blocks, the differential distance between the true MV of the current block and the predicted MV in the first step is very small. Therefore, the search range of the lower level don't care. In this case, so as to decrease the risk of being trapped in a local minimum, the search around the center need be reinforced. In other words, the lower level search should be center-biased. However, as the second possibility, if the motion variation is large, the differential distance between the true MV of the current block and the predicted MV in the first step will be also large. Thus the appropriate search range must be secured in order to find correct MV. Fig. 3 explains this lower level search method in case of D=8. In order to provide for the above-mentioned first possibility, MADs of nine locations including the search center must always be calculated. One MV with minimum MAD and its MAD are stored in some memory. Then the locations with 2 pixel distance, horizontally and vertically, are checked in lower level search range, -D+1 ~ D-1, except the search center. Another MV with minimum MAD is chosen among these locations. This minimum MAD is compared with the above-stored MAD. The MV with minimum MAD is chosen among these as the final MV. In case of D=8, the search range of this two step center-biased method is -7~7 and the number of searching points is 65. Compared with the computational complexity of FS-BMA, that of our scheme amounts to only 29% of FS-BMA. The final MV can be fall outside the given search area, R x R. In order to prevent the final MV from fall outside R x R, the predicted MV in the first step of our scheme is moved by appropriate location. And then the lower level search is performed.

The detailed implementation of the proposed algorithm is described below. Only forward prediction is considered where the previous frame is used as a reference frame.

1. For each macroblock of current frame the candidate set is constructed as Fig 3 shows. If there are no causal neighboring blocks of current block like boundary blocks of frame, the MVs of causal neighboring blocks are set to (0,0).

2. The MV with the smallest MAD is chosen among the candidate set. This MV becomes the search center of the lower level search.

3. With the chosen MV as the search center, two step center-biased search is performed and the final MV is found.

3. EXPERIMENTAL RESULTS

All four test video sequences are used for comparison in the experiment. They are ‘Football’ (fb), ‘Car’ (car), ‘Cheer leaders’ (cheer), and ‘Mobile and calendar’ (mob). Each frame has a resolution of 720 pels x 480 lines and 30 frames/sec as the ITU-R 601 format. A 16 x 16 square block is considered as the macroblock for MV estimation as specified by the MPEG. The performance of the proposed algorithm is
measured in terms of peak-to-peak signal-to-ratio (PSNR), and the MAD distortion function is used as a matching criterion. We adopted 63 pels x 63 pels \((63 = 2 \times 31 + 1)\) as the maximum search area. And \(D\) is set to 8 pels. Therefore, the search range of the lower level is -7 pels \(\sim\) 7 pels.

The comparisons are made among four algorithms; FS-BMA to establish an upper bound of BMA performance, the full search, ZWCA, Liu’s method1 and 2, and SBMA \((S+T)\), and the proposed algorithm. In the simulation, the alternate subsampling technique was not applied to SBMA\((S+T)\) to keep the computational complexity similar to the proposed method. The decimation factor \(d\) of the ZWCA and the minimum search window size \(W_0\) of SBMA \((S+T)\) were fixed to 4 and 7, respectively. In Table I, it is observed that the performance of the proposed algorithm closely follows that of FS-BMA, and our scheme outperforms recently developed fast search algorithms with large search area in the all test sequences. Compared with FS-BMA, the proposed algorithm shows the PSNR degradation of 0.25dB for ‘Football’ sequence in the worst case. Especially, it is worth to note that our scheme shows far superior performance, compared with other algorithms, for ‘Cheer leaders’ which have random or complex motion as well as fast motion. This is because, by using the hierarchical search structure for random or complex motion and spatial correlation of motion fields in order to choose a search center of the lower level, the proposed scheme is more robust against being trapped in local minima than any other algorithms. In addition, Fig. 4 shows PSNR comparisons for the ‘Football’ and ‘Cheer leaders’ sequences, respectively.

The computational complexity of motion estimation is often measured as the number of searching points, or the number of MAD calculations. In the proposed algorithm with \(D=8\) and \(R=63\), the number of searching points is 118. Compared with FS-BMA, the computational complexity of our scheme is less than 2.9% of FS-BMA. This numerical value is the smallest one among the algorithms for comparison. It is clearly seen that the proposed algorithm always provides much better performance with the minimal computational requirement than other fast search algorithms.

4. CONCLUSIONS

In this paper, a fast hierarchical search algorithm is proposed by using spatial correlation of motion field and the merit of hierarchical search. As the search range increases, the conventional fast search algorithms have a drawback that their performances decrease substantially, compared with that of FS-BMA. The tradeoff between performance and computational complexity is problematic because these algorithms need a large amount of computation to keep the performance close up to FS-BMA. In order to give almost the same performance as FS-BMA in a large search area, with as a small amount of computation as possible, we use not only the spatial correlation of motion field for continuous motion, but also adopt the hierarchical structure for random or complex motion. The simulation results evaluates the performance of our scheme. The proposed algorithm keeps outstanding performance of minor 0.25dB PSNR degradation even in the worst case within a large search area, 63 pels x 63 pels, and outperforms any other fast algorithms by maximum 2dB for ‘Football’ sequence. The computational complexity of our scheme is no more than 2.9% of FS-BMA. Moreover, that of our scheme is less than those of any other algorithms for performance comparison. Since the proposed scheme uses the spatial correlation basically, the correlation between MVs of adjacent blocks is strong than FS-BMA. Therefore, this property may potentially help the saving of coding bits and lead a better overall rate-distortion performance.

5. REFERENCES


