Estimation and Compensation of Video Motion - A Review

1K.Vidyavathi, 2Dr.R.S.Sabeenian
1Assistant Professor – ECE & King College of Technology, vidyavathiece2010@gmail.com
2Professor – ECE & Centre Head SONA SIPRO, sabeenian@sonatech.ac.in

Abstract

Motion estimation and compensation are the main factors in the video compression due to the ability of realizing high compression by removing the temporal redundancies. This paper mainly focused on the motion estimation techniques. A survey on papers related to different motion estimation is also included in the paper.

Keywords: Block matching, motion estimation, AVC, HD Videos, adaptive motion estimation algorithm

1. Introduction

The use of video data in our daily life is increasing day by day. The transmission of large size video is facing limitation due to the limited bandwidth and storage capacity. The solution for this is the video compression. In video coding for compression, the basic idea is to exploit redundant data. There are two types of redundancies in a moving picture (temporal redundancy and spatial redundancy). For reducing the temporal redundancies motion estimation and compensation is used. H.264/MPEG-4 part 10 or Advanced Video Coding (AVC), a video compression format and is currently one of the most commonly used formats for the recording, compression, and distribution of video content. It is a Motion compensated-based video compression standard developed by the by the ITU-T Coding Experts Group (VCEG) together with the ISO/IEC JTC1 Moving Picture Experts Group (MPEG). The figure 1 shows the video compression process.

2. Motion estimation and classification

Motion estimation and compensations are the key factors of efficient video compression process. Efficient reduction in temporal redundancy is needed to obtain a high compression. The main causes for temporal redundancy are the from frame to frame in a moving picture the picture element have motion, object of one frame move within the frame to form object of other frame. The motion can be in form of zoom, rotation and translation. For compression the frame is not taken one by one. In the decoding and encoding of video compression a prediction method is used. Two types of prediction are there: Intra prediction and Inter prediction. If reference frame is used for prediction, it is called inter prediction.
The figure 2 shows the diagram for the prediction block. It consists of reference frames, current MB, residual MB. Motion estimation is the most difficult operation and computationally expensive module in the entire computation process. For estimating the motion of objects between the frames, estimate the motion occurred between the reference frame and the current frame. This process is called motion estimation. A motion vector is obtained as a result of the motion estimation process. The motion vector represents the estimated motion. Motion compensation uses the knowledge of the obtained motion vectors to achieve data compression.

There are many motion estimation techniques available. The broad classification of motion estimation algorithm is represented in figure 3. In this paper more focus is given to the block matching algorithm. Most of the fast motion estimation schemes are based on matching algorithm which is composed of one or more of these strategies. Distance criteria are used for measuring the distance between the previous block and search area block.

![Image of Flow Diagram for Prediction Block](image1)

**Figure 2.** Flow Diagram for Prediction Block

The fastness of the algorithm depends on the search strategy used. All fast motion estimation search algorithms use search area sub-sampling technique. The search area is divided into two types. They are fixed search area and adaptive search area. The two main classification of block matching algorithm are time domain and frequency domain algorithms. The frequency domain technique is based on relationship between transformed coefficients of shifted images, and they are not widely used for image sequence coding. In this, the motion estimation is done by taking the transform of the block first in frequency domain (e.g. by DCT or by wavelet). In time domain the feature matching and block matching are entirely different. In feature matching, matching of Meta information extracted from the current block and search area picture elements are used. And the matching process is performed by morphological filters and projection methods. In block matching, matching of (all/some) pixels of current block with the candidate block in search area is performed according to distance criterion described.

![Image of Broad Classification of ME Algorithms](image2)

**Figure 3.** Broad Classification of ME Algorithms
Since the mesh based models employ interpolation for obtaining motion vectors of the picture elements within a given range, this gives in general a more continuous effect than BMAs. So, in terms of prediction accuracy, mesh based models can give visually more pleasing prediction, especially in the presence of non-translational motions, such as head rotation and turning. While in terms of computational complexity the BMAs certainly have an edge over Mesh based ME, since mesh based models involve interpolation of motion vectors which requires more complex architecture.

3. Survey


In this paper [1] a multiplication-free C 1BT (MF-1BT) is proposed. Using this method 1BT is carried with integer arithmetic using addition and shift operation only. In the 1BT proposed in [2], the new diamond shaped convolution kernel is used to filter image frames to construct the 1BT. As a result of the 1BT, image frames are converted into a single bit plane representation. While this representation is generally successful in discriminating pixels with different values, one aspect that reduces the ME accuracy is the fact that pixel values on directly opposite sides of the threshold are categorized into separate classes and counted as a non match in the matching process, even if original pixel values are actually close.

The proposed approach improves the performance of 1BT-based ME schemes using simple and low complexity operations. Experimental results show that the proposed approach gives improved ME accuracy compared to 1BT [3], MF-1BT [4], as well as 2BT [5]. The proposed approach is particularly suitable for applications that require portable and mobile video encoding due to the low complexity and low power consumption needs in such cases.

3.2. Enhanced transform domain intra prediction for MPEG-2 to H.264/AVC transcoding (2008)

An enhanced transform-domain intra prediction method for H.264/AVC is proposed in this paper. Two filters are used H filter and V filter. Using this, intra prediction can be decomposed in to some sparse matrices and hence the computational complexity can be reduced. This scheme can reduce the number of multiplication and addition operations more than Chen’s [7] by 22.4% and 21.6%, respectively, in each 4×4 block. Also matrix multiplication can be simply processed using a look-up table which consists of 9 types of 4×4 constant matrices corresponding to 9 prediction modes in H.264. The proposed scheme in this paper can be used for fast transform domain intra prediction in the area of transcoding MPEG-2 streams into H.264/AVC streams.

3.3. Sub pixel interpolation architecture for multi standard video motion estimation (2009)

In this paper new domain-specific reconfigurable sub pixel interpolation architecture for multi standard video motion estimation or motion compensation is presented. The mixed use of parallel-input and serial-input FIR filters achieves high throughput rate and efficient silicon area utilization. Flexibility is achieved by a detailed examination of the interpolation filters used in these standards and by using a multiplexed reconfigurable data-path controlled by a selection signal. This compares very favorably with existing fixed solutions that are based on the H.264 standard. A detailed design study shown that this can be implemented using 34.8 × 10^3 gates. As discussed in references [7] and [8], it is the critical path of the interpolation unit that determines the throughput rate of FME architecture. For example, in [8] a reduction from 10 to 4.04 ns in the filter critical path delay leads to an increase by a factor of over two in the block throughput rate.

This paper presents a gradient based motion estimation algorithm based on shape-motion prediction, which takes advantage of the correlation between neighboring Binary Alpha Blocks (BABs), to match with the Mpeg-4 shape coding case and speed up the estimation process. The PSNR and computation time achieved by the proposed algorithm seem to be better than those obtained by most popular motion estimation techniques.

In fact, many researchers have focused on ME algorithms especially based on texture coding. However, one of the most important concepts introduced by the Mpeg-4 visual standard is the use of video object (VO) as an entity the user can access and manipulate. The instance of a VO at a particular point of time is called video object plane (VOP) [9]. Therefore, in MPEG-4 video coding, ME of shape is also imperative for real-time VOP-based encoding. Several papers have proposed software implementation methods for shape coders [10], [11] where shape information is used to reduce search point per macro block and only valid predecessors are evaluated [10] for boundary macro-blocks. Results show that the algorithm presents a good PSNR result with a net decreasing in the number of iterations and computation time.

3.5. Algorithm and architecture co-design of hardware-oriented, modified diamond search for fast motion estimation in H.264/AVC (2011)

In this paper, a new hardware-oriented, integer-pel, fast ME algorithm called HMDS, along with its supporting hardware architecture is presented. At lower bit rates, HMDS has insignificant RD losses when compared to full search, while it has superior RD performance compared to some state-of-the-art fast ME algorithms. HMDS superior RD performance and its suitability for hardware implementation arise from the fact that with a fixed and regular search pattern, HMDS is robust enough to find the globally optimum motion vector because it simultaneously tracks large and local motions through its satellite search points and last two steps of intensive local search. It is shown that HMDS has superior speedup than many fast algorithms, while its loss in speedup is minimal when compared to SUMH.


This paper mainly focuses on the two main classification of motion estimation algorithms used for video compression. Motion Estimation (ME) algorithms vary with respect to the a priori information and constraints they employ, as well as the method of computation they use to obtain the estimate. The classifications for ME algorithms are based on feature/region matching, gradient based methods, spatiotemporal energy methods, deterministic model based methods and stochastic model based methods. This paper focuses more on block matching algorithms which comes under feature/region matching and gradient based methods. This paper compares the two important ME algorithms such as Full Search (FS) and a newly found adaptive algorithm called AMEA. In FS every candidate points were evaluated and more time has been taken for predicting the suitable motion vectors. Based on the above noted drawback, the above said adaptive algorithm called AMEA is proposed.

3.7. Frame rate up-conversion for high-definition video applications (2013)

In this paper, a novel frame rate up-conversion (FRUC) scheme to achieve better visual quality for high definition videos is proposed. The proposed algorithm has three main advantages over conventional FRUC algorithms. First, to increase the accuracy of motion estimation, this paper considers the diversity of motions, the retiming motion vectors of the previous motion field, and an adaptive update strategy. Second, the motion vector that is selected by using the sum of the absolute differences criterion is outputted after a geometric-based motion vector restriction, thereby improving the consistency of the motion vectors. Third, post-processing is performed on the estimated motion vectors based on localized global motions which are calculated by using motion histograms. The figure 4 shows the framework of the proposed algorithm.

This paper proposed a novel frame rate up conversion algorithm using adaptive extended bilateral motion estimation. Conventionally, extended bilateral motion estimation (EBME) conducts dual ME processes on the same region, therefore involves high complexity. However, in this proposed scheme, a novel block type matching procedure has suggested to accelerate the ME procedure [16]. In this work, the edge information was calculated using Sobel mask, and the calculated edge information is used in block type matching procedure. Based on the block type matching, decision were made whether to use EBME. Motion vector smoothing is adopted to detect outliers and correct outliers in the motion vector field.

3.9. Motion compensated frame rate up-conversion using modified adaptive extended bilateral motion estimation (2014)

In this paper, a novel frame rate up conversion using modified adaptive bilateral motion estimation has been proposed. In the conventional algorithms, extended bilateral motion estimation carried out bilateral motion estimation twice on the same region and hence high complexity. Adaptive extended bilateral motion estimation algorithm has also proposed to reduce complexity and increase visual quality by using block type matching process and considering frame motion activity [17]. In this algorithm, the calculated edge information was used to detect a global scene cut change, and then is used in block type matching process whether to use Extended Bilateral ME. Finally, this paper adopted to interpolate intermediate frames by overlapped block motion compensation and motion compensated frame interpolation.

4. Conclusion

In this paper a brief study of motion estimation is done. As Motion estimation has various promises in applications like video telephony, HDTV, automatic video tracker and computer vision etc. Thus, Extensive research is has been done over years to develop new algorithms and designing cost-effective and massively parallel hardware architecture suitable for current VLSI technology. So, till now there are unlimited numbers of algorithms being claimed by different researchers in world. From all the previous types of algorithms discussed, Block Matching Algorithms are the simplest way for motion estimation in terms of hardware and software implementations. A survey on various papers on motion estimation is also done. The researcher’s contribution to the motion estimation and .Improvement in image quality and compression rates are visible in each paper.

5. References