A New Symmetric-Object Oriented Approach for Motion Estimation in Wireless Multimedia Sensor Networks

Thanaa Jbeily, Mothanna Alkubeily, Iyad Hatem

Department of Mechanical and Electrical Engineering, Tishreen University, Lattakia, Syria

Abstract: Motion estimation is a process of determining motion vectors. It is an essential challenge in digital video processing and computer vision. Therefore, it does not fit critical condition applications and constricted resource systems such as Wireless Multimedia Sensor Networks (WMSN). The main objective of the motion estimation is powerfully reducing temporal redundancy between successive frames to achieve significant video compression. In this paper, we propose a novel symmetric-object oriented approach for motion estimation in WMSN, called SYMO-ME. Our goal is to reduce the high complexity of motion estimation process. We also introduce a new motion estimation energy consumption model for block matching algorithms (BMAs) in WMSNs. This model depends on the energy consumption value of different executed instructions. Simulation results confirm the efficiency of our proposition in reducing the total number of search points for finding the motion vectors over a video frame compared with a group of considered BMAs, and improving the WMSN life time by saving the limited resources.

Keywords: Wireless Multimedia Sensor Networks, Energy consumption, Motion Estimation, block-matching algorithm, Symmetry.

1. Introduction

Wireless Multimedia Sensor Network (WMSN) [1]-[3] is a group of tiny size, self-organized, and constricted resources sensory devices. Video signal processing over WMSN is an extremely difficult task because of the complexity of video data on one hand and the critical characters of this kind of networks on another hand. The most computationally-expensive and resources greedy operation over the video processing system is the motion estimation (ME) process [4]-[6]. However, ME is the heart of video processing. It determines motion vectors (MV) which describe the transformation from adjacent frames in a video sequence. There are different classes of motion estimation algorithms [7],[8]. Fast block matching algorithms (FBMAs) class exploits the correlation between the pixels within a block and the correlation between neighboring blocks [7]-[9]. Unfortunately, FBMAs still find separately and exhaustingly the motion vector of every individual block all over the frame.

Symmetry is an omnipresence visual, geometrical, and physical phenomenon which has motivated many related studies and applications [10]-[12]. However, to the best of our knowledge, symmetry property is not exploited yet in motion estimation domain in any previous cited paper. All existing motion estimation algorithms do the searching process over the frame without taking into consideration studying or analysing the frame objects properties or shapes.

In this paper, we propose a new SYMmetric-Object oriented approach for Motion Estimation, called SYMO-ME. It is a WMSN fitted butterfly symmetry based block matching motion estimation. SYMO-ME aims to reduce the high complexity of motion estimation process. It depends on the fact that parts of symmetric objects tend to submit symmetric motions. Consequently, motion vectors of one of that symmetric parts can be exploited in motion estimation of the other remaining parts.

Moreover, we introduce a new motion estimation energy consumption model for BMAs in WMSNs. This model depends on the energy consumption values of different executed instructions. Each instruction requires a certain number of processor cycles, in which the consumed energy per each cycle is fixed [13],[14]. The cycle count is a measure of the complexity and also the energy consumption of an implementation on a specific processor [15],[16].

The reminder of this paper is organized as follows: In section II, we study motion estimation concept with emphasis on block matching motion estimation class criteria and algorithms. We also present a view on symmetry property types, applications, and examples. In section III, we propose our method SYMO-ME, and introduce a new motion estimation energy consumption model for BMAs in WMSNs. In section IV, several extensive simulations are conducted to evaluate the efficiency of SYMO-ME and to measure its impact on several well-known BMAs. Finally, conclusion and future work are presented in section V.

2. Related Works

2.1 Motion Estimation

Motion estimation (ME) is the key of video signal processing [5],[17],[18]. It benefits from the high correlation among successive video frames and exploits their information redundancy - called temporal domain redundancy. It determines the shift of a particular region in the current frame by considering a suitable region in a reference frame. This shift is represented by displacement vector which is commonly known by motion vector (MV) [4],[17].
There are different ways to classify motion estimation algorithms. Based on the particular region whose motion vector is needed to be found, ME is classified into two basic classes [7]: (1) Pixel based algorithms which suffer of heavy complexity load and huge associate size of data caused by the use of motion vector for each individual pixel, and (2) Block based algorithms which overcome the above problem by exploiting the high correlation between each pixel and its neighbors. It separately finds the motion vector of each block in the current frame according to matching process. Consequently, algorithms related to this class is called Block Matching Algorithms (BMAs) [7],[8].

2.2 Block Matching Motion Estimation

Block Matching for motion estimation has been widely adopted by current video coding standards such as H.261, H.263, MPEG-1, MPEG-2, MPEG-4, and H.264 due to its effectiveness and simplicity for both hardware and software implementation. This process searches for the best matched block, i.e. the best cost function among a group of candidate blocks positioned in a search region that is called search window. This window is limited by a search parameter (p), i.e. the search will be performed within a square region of [-p, +p] around the position of the current block [8] and each candidate block is seen as a Search Point (SP) as shown in Figure 1. In terms of block matching two major issues must be taken into consideration, the block matching criterion and the searching algorithm.

![Search window](image)

Figure 1: Block motion estimation.

2.2.1 Block Matching Criterion

Current block is compared with the search window blocks to find its best matched one. This is done based on calculating a cost function which determines the block matching criterion. Indeed, there are several cost functions [7],[8], and we can classify them into the following two basic classes.

1) **Block divergence based cost functions**: Functions of this class measure the difference between the current block and each of the candidate blocks [3],[8]. The block with minimum dissimilarity i.e. minimum matching error is the best matched one. The most popular criterion related to this class is Mean Absolute Differences (MAD) which is represented as in (1) [8].

\[
MAD = \frac{1}{n \times n} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} |C_{ij} - R_{ij}|
\]  

(1)

Where \(n \times n\) is the block size, \(C_{ij}\) is the value of pixel (i, j) in the current block, \(R_{ij}\) is the value of pixel (i, j) in the candidate block.

2) **Block correlation based cost functions**: This cost functions depend on finding the maximum similarity between the current block and each of the candidate blocks [3],[8]. The block with maximum similarity is the best matched one. The most popular criterion related to this class is Peak Signals to Noise Ratio (PSNR) which is given as in (2) [8].

\[
PSNR = 20 \log_{10} \left( \frac{255}{MAD} \right)
\]

(2)

2.2.2 Block Matching Algorithm

According to the considered search points, BMAs are classified into two essential categories [7],[9]: (1) Exhaustive Block Matching Algorithm (EBMA): It is called the full search algorithm due to the fact that it examines all possible search points over the whole search window to choose the best matching one. (2) Fast Block Matching Algorithms (FBMAs): Unlike EBMA, FBMAs try to reduce the complexity associated with EBMA [7],[8]. Therefore, they examine only some positions through the search window based on a search pattern. As a result, FBMAs do not offer the same level of quality as EBMA, but they approximate it and moreover achieve a noticeable load reduction [3],[7],[9],[19]. Consequently, FBMAs are widely adopted. There are different fast block matching algorithms for searching proposed in the literature. Well-known examples are, Three-Step Search (TSS) [3], New Three Step Search (NTSS) [8], Four Step Search (FSS) [20], Diamond Search (DS) [9], and Adaptive Rood Pattern Search (ARPS) [21], etc.

2.3 Symmetry Surrounds Us

Symmetry is an intrinsic visual, geometrical, and physical phenomenon [10]-[12]. It is variously defined as "proportion," "perfect, or harmonious proportions," and "a structure that allows an object to be divided into parts of an equal shape and size" [10]. Plenty magic symmetric scenes are manifesting themselves all around us. Symmetry can be found in various aspects and it is occurring naturally such as butterflies and faces, in manufactured artefacts such as carpets, and in architecture such as buildings [11].

The omnipresence of symmetry has motivated many studies and numerous applications [22]-[24], such as object recognition and identification, facial image analysis, vehicle detection, 3D reconstruction, visual attention and inspection, shape representation, and biological vision segregation, etc. Basically, there are three types of symmetry, they are [10]-[12]:

1) **Reflectional symmetry**: It is the most common symmetry type, in which if a central line is drawn, the object or the scene can be divided into two matching halves. Based on this, many names can be used for it such as bilateral, line, or mirror symmetry. The line of symmetry can be in any direction not just up-down or left-right. Butterflies are the best example of this symmetry type. Butterfly and other examples are shown in Figure 2.

2) **Rotational symmetry**: A shape has rotational symmetry if it fits onto itself two or more times in one turn. The order of rotational symmetry is the number of times the shape fits onto itself in one turn. Figure 3 presents some
aspects of this type of symmetry which is also called radial symmetry.

3) Transitional symmetry: This type outcomes from a parallel shift of an object to form translations that have no fixed points. Examples of transitional symmetric scenes are shown in Figure 4.

3. Symmetric-Object Oriented Approach for Motion Estimation

3.1 Overview of Our Proposal

All studies insure that motion estimation is the most computationally expensive and resource hungry operation over the video compression/coding system [4]-[6]. This results in an inappropriate load which does not fit critical condition applications and constricted resource systems such as WMSNs. Many attempts tried to reduce this deep complexity such as avoiding EBMA and depending only on FBMAs. FBMAs exploit the correlation between the pixels within the block on a side and the correlation between neighbouring blocks on another side [7],[9]. Unfortunately, FBMAs still both separately and exhaustingly find the motion vector of every individual block all over the frame from the top-left corner to the bottom-right one [3]. From our point of view, this is still too heavy. Moreover, we believe that additional possible intra frame redundancy may be exploited. This kind of redundancy is related to the frame content, i.e. it is linked to the objects of the considered scene. In fact, processing on objects level is very logical and it rises in various domains starting from Object Oriented Programming (OOP) such as C++ and Java languages, to Object Oriented Coding (OOC) such as H.264 video coding standard.

Essentially, the picked video frames extremely reflect what we see around us in everyday life. However paying little attention to most of the objects around us, we directly notice that the most major object related property is the symmetry property [10]. Actually human naturally is attracted to symmetric views, and likely has an uncontrollable urge to look for symmetry in everything [10]-[12]. We can say that it is really rare, if we do not say impossible, to find a scene without any symmetry aspect.

Although there has been considerable studies in literature on symmetry in computer vision [22]-[24], but they are limited with still images only. To the best of our knowledge, symmetry property is not exploited yet in motion estimation domain in any previous cited research. All existing motion estimation algorithms do the searching process over the frame without taking into consideration studying or analysing the frame objects properties or shapes.

We propose, based on the preceding discussion, a new SYMMetric-Object oriented approach for Motion Estimation in WMSN, called SYMO-ME. Our proposal aims to reduce the high complexity of motion estimation process. Where, we believe that parts of symmetric objects tend to submit symmetric motions. Consequently, motion vectors of one of that symmetric parts can be exploited in motion estimation of the other remaining parts. As a result, we expect SYMO-ME will noticeably decrease the complexity of any motion estimation algorithm and make it more efficient for real time applications and limited resources networks such as WMSNs.

Our proposal is basically inspired from the butterfly image which is the best example of reflectional symmetry, as explained in Figure 5.
In this paper, we introduce a new motion estimation energy based on the difference between the two cases. (2) Coding program running, and then calculating the consumed measuring the current consumption in the case of coding/not divergence based cost functions, we choose Mean Absolute difference (MAD) which is the most frequently used block matching criterion due to its simplicity, low computational and memory requirements comparing to other cost functions [8],[19].

3.3 Our Proposed Motion Estimation Energy Consumption Model

Motion estimation energy consumption is constant for a given search algorithm and different from one to another [25]. We did an extensive investigation to find a direct model for the motion estimation energy consumption. However, all researches present it based on either aided or arbitrary metrics. For example: (1) Authors in [25] depended on measuring the current consumption in the case of coding/not coding program running, and then calculating the consumed energy based on the difference between the two cases. (2) Other studies relied on average number of search points per block either to directly represent the consumed energy [19] or the execution time of the motion estimation algorithm [8]. (3) Both [7] and [26] calculated the number of computational operations to represent the consumed energy, and (4) in [3] we adopted the metrics of average number of search points per block, both complexity degree and execution time of motion estimation algorithm, and computational complexity.

In this paper, we introduce a new motion estimation energy consumption model for BMAs in WMSNs. Our proposal depends on the energy consumption values of different executed arithmetic operations. Each operation requires a certain number of processor cycles and the energy per each cycle is fixed [13],[14]. The cycle count is a measure of the complexity and also of the energy consumption of an implementation on a specific processor [15],[16]. It is also inversely proportional to the throughput which describes how fast an algorithm works [16].

Actually, BMA consumes a given amount of energy due to the computational overhead that it creates during the block matching process, i.e. during the calculation of the matching criteria [3],[8]. Mean absolute difference-MAD, which is the matching criterion in our study, has to be evaluated for each search point. At each search point, we compare \( n \times n \) pixels, i.e. as block size, and each pixel comparison requires three operations, namely: subtraction, addition and absolute value calculation.

Assuming SPpF is the total number of checked search points per frame, the following operations are performed for each frame:

1. \( \text{SPpF} \times n \times n \) subtraction operations.
2. \( \text{SPpF} \times n \times n \) addition operations.
3. \( \text{SPpF} \times n \times n \) absolute value calculation operations.

Each absolute value calculation operation is assumed as a comparison and a multiplication operation. Therefore, the operations for each frame will be:

1. \( \text{SPpF} \times n \times n \) subtraction operations.
2. \( \text{SPpF} \times n \times n \) addition operations.
3. \( \text{SPpF} \times n \times n \) comparison operations.
4. \( \text{SPpF} \times n \times n \) multiplication operations.

If \( \epsilon \text{ (instruction)} \) is the consumed energy of instruction operation, the total energy consumption during motion estimation per frame is given by (3):

\[
\text{Motion estimation energy consumption / frame} = \text{SPpF} \times n \times n \times \epsilon \text{(sub)} + \text{SPpF} \times n \times n \times \epsilon \text{(add)} + \text{SPpF} \times n \times n \times \epsilon \text{(comp)} + \text{SPpF} \times n \times n \times 2 \times \epsilon \text{(mult)} \tag{3}
\]

Based on the importance of energy consumption over limited energy supplied systems like WMSNs, our motion estimation energy consumption model is investigated on Mica2 sensor node, developed by Crossbow Technology Inc with Atmel AtMega128L microcontroller. From the Atmel AtMega128L microcontroller datasheet [27] and [13], we obtained the cycle counts and energy consumption values of different operations as shown in Table 1.

Table 1: Energy consumption per operation for ATMEL ATMEGA128L microcontroller

<table>
<thead>
<tr>
<th>Arithmetic Operation</th>
<th>Cycle Count</th>
<th>Energy Consumption value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \epsilon \text{(sub)} )</td>
<td>1</td>
<td>3.3 nJ/byte</td>
</tr>
<tr>
<td>( \epsilon \text{(add)} )</td>
<td>1</td>
<td>3.3 nJ/byte</td>
</tr>
<tr>
<td>( \epsilon \text{(mult)} )</td>
<td>2</td>
<td>6.6 nJ/byte</td>
</tr>
<tr>
<td>( \epsilon \text{(cmp)} )</td>
<td>1</td>
<td>3.3 nJ/byte</td>
</tr>
</tbody>
</table>

Using the energy model formulated in (3) and the energy cost values in Table 1, and assuming \( \epsilon \) is the energy cost of one processor cycle, the motion estimation energy consumption can be calculated as in (4).

\[
\text{Motion estimation energy consumption / frame} = \text{SPpF} \times n \times n \times \epsilon + \text{SPpF} \times n \times n \times \epsilon + \text{SPpF} \times n \times n \times 2 \times \epsilon \text{(mult)} \tag{4}
\]

By grouping the terms, our final formula of energy cost model will be as in (5).

\[
\text{Motion estimation energy consumption / frame} = 5 \times \epsilon \times (\text{SPpF} \times n \times n) \tag{5}
\]
The analysis of the proposed formula shows that, for a fixed block size $n \times n$, the consumed energy is directly proportional to the total number of checked search points per frame. For the processor used in this analysis, Atmel AtMega128L on the Mica2, with $\varepsilon = 3.3 \text{nJ/byte}$, the last formula will result in (6).

$$\text{Motion estimation energy consumption /frame} = 16.5 \times (\text{SPpF} \times n \times n)$$

(6)

4. Results and Analysis

4.1 Simulation Environment and Setup

We used MATLAB 7.12.0 (R2011) to simulate our proposal. „Akiyo” video sequence with one object depicting reflectional symmetry was used to generate the frame-by-frame results. The still of the analyzed video is illustrated in Figure 6. We used the first 150 frames supported by the quarter common interchange format (QCIF, 176 pixels × 144 pixels), with GOP equal to 5, and with MAD as a block matching criterion. These parameters are suitable for the buffer of video sensors in WMSNs [3].

Several well known BMAs: TSS, NTSS, FSS, DS, and ARPS were implemented. To evaluate the performance of SYMO-ME and measure its impact on the mentioned algorithms, we considered three basic evaluation metrics: (1) Quality of the BMA in terms of speed and complexity (Total number of search points per frame measured by search point per frame-SPpF), (2) Motion estimation energy cost (motion estimation energy consumption measured by nJ). and (3) Quality of video (Peak Signal to Noise Ratio-PSNR per frame measured by dB).

4.2 Simulation Results

To evaluate the efficiency of SYMO-ME, several extensive simulations were conducted. We got the following results for the considered performance evaluation metrics.

1) **Quality of the BMA**: We plot the P-frame-by-P-frame total number of search points results for the considered BMAs: TSS, NTSS, FSS, DS, and ARPS in Figure 7 to Figure 11, respectively. It is obvious that our proposal did not add neither additional time nor additional complexity to motion estimation. On the contrary, SYMO-ME saved the limited resources over WMSN. It required less total number of search points per frame at all frames. The average percentage reduction rates were 52%, 57%, 56%, 59%, and 60% for the above mentioned BMAs, respectively.
2) Motion estimation energy cost: Here we mapped from Figure 12 to Figure 16 the results of motion estimation energy consumption per P-frames by the Atmega128L on the Mica2. We can clearly notice the excellent performance of our proposal of adopting SYMO-ME. It consumes less energy at all frames and for all analyzed BMAs. SYMO-ME saves the energy obviously where the average percentage energy reduction rates are 52%, 57%, 56%, 59%, and 60% respectively compatible with the studied BMAs due to the proportionality of this metric with total number of search points.
3) Quality of video: We simulated PSNR results for each P-frame of the considered BMAs: TSS, NTSS, FSS, DS, and ARPS which are shown from Figure 17 to Figure 21, respectively. However, although our proposal of using SYMO-ME offers a distinct performance of the two former metrics, it still maintains the video quality. Where, its PSNR results compete the results of each one of the considered algorithms. It does not cause a deep degradation of the average PSNR value of the analyzed video as demonstrated in Table 2.

![Figure 16: Motion estimation energy consumption (nJ): ARPS&ARPS-SYMO.](image)

![Figure 17: Peak signal to noise ratio (dB): TSS&TSS-SYMO.](image)

![Figure 18: Peak signal to noise ratio (dB): NTSS&NTSS-SYMO](image)

![Figure 19: Peak signal to noise ratio (dB): FSS&FSS-SYMO](image)

![Figure 20: Peak signal to noise ratio (dB): DS&DS-SYMO.](image)

![Figure 21: Peak signal to noise ratio (dB): ARPS&ARPS-SYMO.](image)

<table>
<thead>
<tr>
<th>BMA</th>
<th>BMA</th>
<th>BMA-SYMO ME</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSS</td>
<td>36.57</td>
<td>32.05</td>
</tr>
<tr>
<td>NTSS</td>
<td>39.41</td>
<td>33.61</td>
</tr>
<tr>
<td>FSS</td>
<td>36.19</td>
<td>32.25</td>
</tr>
<tr>
<td>DS</td>
<td>36.30</td>
<td>32.35</td>
</tr>
<tr>
<td>ARPS</td>
<td>37.29</td>
<td>32.69</td>
</tr>
</tbody>
</table>

Table 2: Average PSNR value of the analyzed video
5. Conclusion and Future Work

In this paper we have proposed a new symmetric-object oriented approach for motion estimation in WMSN, called SYMO-ME. It reduces the total number of search points for finding the motion vectors over a video frame compared with a group of considered algorithms. It then saves the limited resources and improves the WMSN life time. We also introduced a new motion estimation energy consumption model in WMSNs to compare different block matching algorithms energy cost which is very critical issue for these networks.

As a future work, we plan to develop SYMO-ME to treat multi objects with multi types of symmetry. We also look to exploit the achieved energy reduction of our proposal to improve the video quality.

Acknowledgment

We would like to express our gratitude to Dr. Boushra Maala for her help, invaluable discussions, and encouragement.

References


Author Profile

Thanaa Jbeily is currently a Ph.D. candidate at Department of Communications and Electronics Engineering, Faculty of Mechanical and Electrical Engineering, Tishreen University, Latakia, Syria. She received her M.S. degree in Communication Engineering from Tishreen University in 2015. Her research interest includes Wireless Multimedia Sensor Networks, Digital Image and Video Processing and Internet of Multimedia Things.

Mothanna Alkubeily is currently a Doctor in Department of Communications and Electronics Engineering, Faculty of Mechanical and Electrical Engineering, Tishreen University, Latakia, Syria. He received the M.S. degree and PhD degree in Application-Level Multicast Protocols from the University of Technology of Compiègne (UTC, France) in 2006 and 2009 respectively. His research interest includes Wireless Sensor Networks, VANET and Application-Level Multicast.

Iyad Hatem is currently a Assistant Professor in Department of Mechatronics Engineering, Faculty of Mechanical and Electrical Engineering, Tishreen University, Latakia, Syria. He received the M.S. degree and PhD degree in Biological Engineering (Image Analysis and Pattern Classification) from the University of Missouri, Columbia, Missouri, USA in 1998 and 2003 respectively. His research interest includes Digital Image processing, Computer Vision, and Logic programming.