

# Predictive Block-Matching Algorithm for Wireless Video Sensor Network Using Neural Network

Zhugue Yan, Siu-Yeung Cho, Sherif Welsen Shaker

Department of Electrical and Electronic Engineering, Faculty of Science and Engineering, University of Nottingham Ningbo China, Ningbo, China  
Email: david.cho@nottingham.edu.cn

**How to cite this paper:** Yan, Z.G., Cho, S.-Y. and Shaker, S.W. (2017) Predictive Block-Matching Algorithm for Wireless Video Sensor Network Using Neural Network. *Journal of Computer and Communications*, 5, 66-77.

<https://doi.org/10.4236/jcc.2017.510007>

**Received:** June 15, 2017

**Accepted:** August 25, 2017

**Published:** August 28, 2017

Copyright © 2017 by authors and Scientific Research Publishing Inc.  
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

This paper proposed a back propagation neural network model for predictive block-matching. Predictive block-matching is a way to significantly decrease the computational complexity of motion estimation, but the traditional prediction model was proposed 26 years ago. It is straight forward but not accurate enough. The proposed back propagation neural network has 5 inputs, 5 neurons and 1 output. Because of its simplicity, it requires very little calculation power which is negligible compared with existing computation complexity. The test results show 10% - 30% higher prediction accuracy and PSNR improvement up to 0.3 dB. The above advantages make it a feasible replacement of the current model.

## Keywords

Wireless Sensor Network, Predictive Block-Matching, Neural Network, High Efficaciously Video Coding

---

## 1. Introduction

A good wireless multimedia sensor network (WMSN) must be energy efficient, for the possibility of operating in harsh environment and lack of maintenance [1] [2] [3]. However, video encoding requires massive amount of calculation power. Taking H.264 standard as an example, processing an HDTV (720P) video signal in real time requires a 3600 MIPS processor with 5570 GB/s memory band-width [4] [5], which exceeds the computing power of most desktop processor. Apart from implementing more advanced processors with higher energy efficiency ratio, which will increase the hardware cost, using more efficient encoding algorithm is a more

practical option.

For a video encoding system, predictive coding is the key part and requires more calculation than other parts such as entropy coding. Motion estimation plays an important role in predictive coding [6] [7]. Therefore, in the field of video stream transmission for WMSNs, an energy-efficient block-matching algorithm is of high importance. There are variety of block-matching search algorithms for motion estimation, such as three-step search, four-step search, diamond search, modified diamond-square search and gradient descent search, etc. [6] [7]. Prediction technique proposed by Zhang and Zafar [8], [9] uses a simple prediction model to reduce the search windows size for block-matching, therefore significantly reduced the calculation complexity of motion estimation, with only small sacrifice on image quality. It has been widely used by many researchers in different block-matching algorithms [10] [11] [12]. The prediction model proposed by Zhang and Zafer is straight forward, simply assuming that the current block has the same motion vector as the block at the same place in the previous frame or around it in the same frame. However, in our test the model did not provide high prediction accuracy. Instead of taking a mean value, we believe the reference motion vectors have certain weights. Therefore, we thought of artificial neural network (ANN).

An ANN mimics somewhat the learning process of a human brain. Instead of complex rules and mathematical routines, ANNs are able to learn the key information patterns within a multidimensional information domain [13] [14]. In this report we aim to use a simple back propagation neural network instead of the traditional prediction model, to improve the prediction accuracy. The results show that the proposed neural network can significantly improve the prediction accuracy and leads to better image quality. Because of the simplicity of the neural network, the required calculation is little.

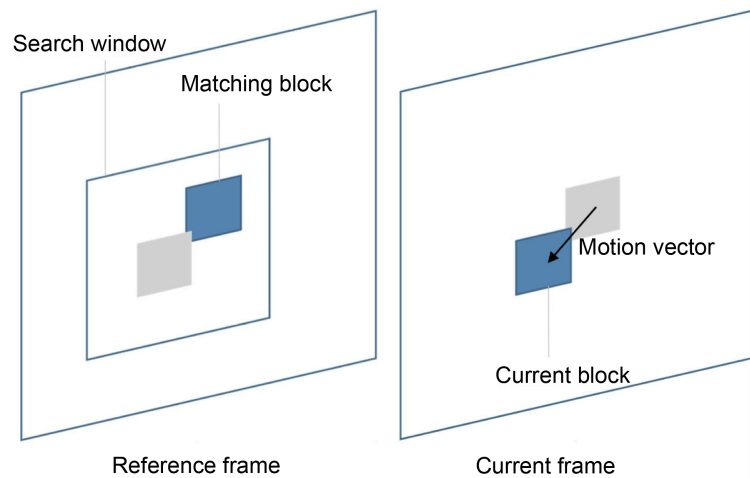
## 2. Background Research

### 2.1. Motion Estimation

The principle of block-matching motion estimation is to find out the movement of each microblock compared with the previous frame. Getting motion vector is the first step to do predictive encoding for video compression. To find out the movement information, the current frame is divided into  $16 \times 16$  pixels microblocks. Each microblock is compared to the reference frame in a search window of a certain size, therefore locate a microblock which has a minimum MAD (mean absolute difference) value as is shown in **Figure 1**. The lengths of the microblock moved along the x axis and y axis in pixels are the motion vectors.

#### 2.1.1. Three Step Search

Three-step search (TSS) algorithm was first introduced in 1981 by Koga *et al.* [15]. It is a classic search algorithm for motion estimation. It is an efficient algorithm which needs only 3 three steps to finish the search in a search window



**Figure 1.** Block-matching motion estimation.

of  $\pm 7$ . If the search window is larger than  $\pm 7$ , then this algorithm needs more than 3 steps to finish the search. TSS uses a coarse-to-fine search pattern. Starting from the initial point, TSS compares 8 search points around it with a certain search step and finds the direction of block movement, then perform a more accurate search by reducing the search step. The description of TSS algorithm is detailed as follows:

Step 1: The step size is set as half of (search window size  $-1$ ). Together with the initial point, the SAD values of 8 search points at the distance of step size are compared.

Step 2: The point which has the minimum SAD value is picked as the new initial point, the step size is halved.

Step 1 and 2 are repeated until the step size is smaller than 1.

**Figure 2** shows the search pattern of TSS algorithm.

### 2.1.2. Enhanced Modified Orthogonal Search

Orthogonal search (OS) algorithm was introduced by Soongsathitanon *et al.* [16] in 2005. The search pattern is horizontally and vertically conducted alternatively (**Figure 3(a)**). It may be considered as improved TSS algorithm, which takes less number of search points. However, this algorithm has the same disadvantage as TSS does: it is inefficient for small motion estimation.

Modified Orthogonal search (MOS) algorithm was introduced by Metkar *et al.* [17] in 2010, meant to solve the small motion estimation problem. It adds 8 new search points as a  $3 \times 3$  square around the initial point in the first search step (**Figure 3(b)**). The added 8 search points improve its performance on small motion estimation.

Enhanced modified orthogonal search (EMOS) algorithm was introduced by Pandian *et al.* [18] in 2011. It replaced the  $3 \times 3$  square by a small diamond search pattern (**Figure 3(c)**). EMOS further improved the motion estimation performance compared with MOS, especially in reducing search points.

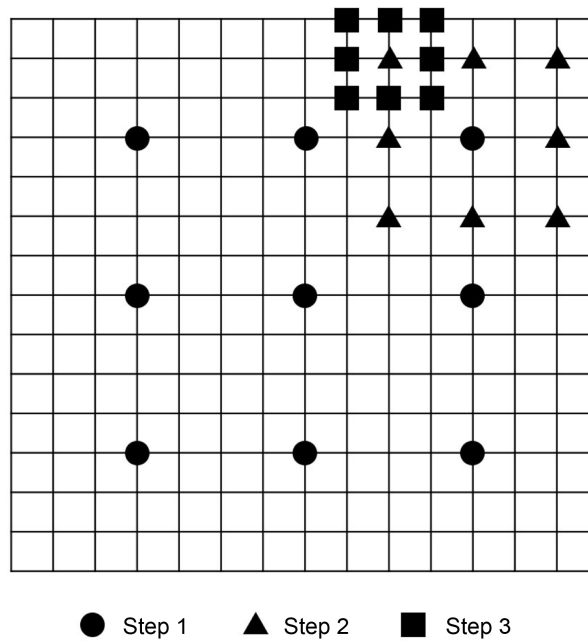


Figure 2. Three-step search algorithm (search window  $\pm 7$ ).

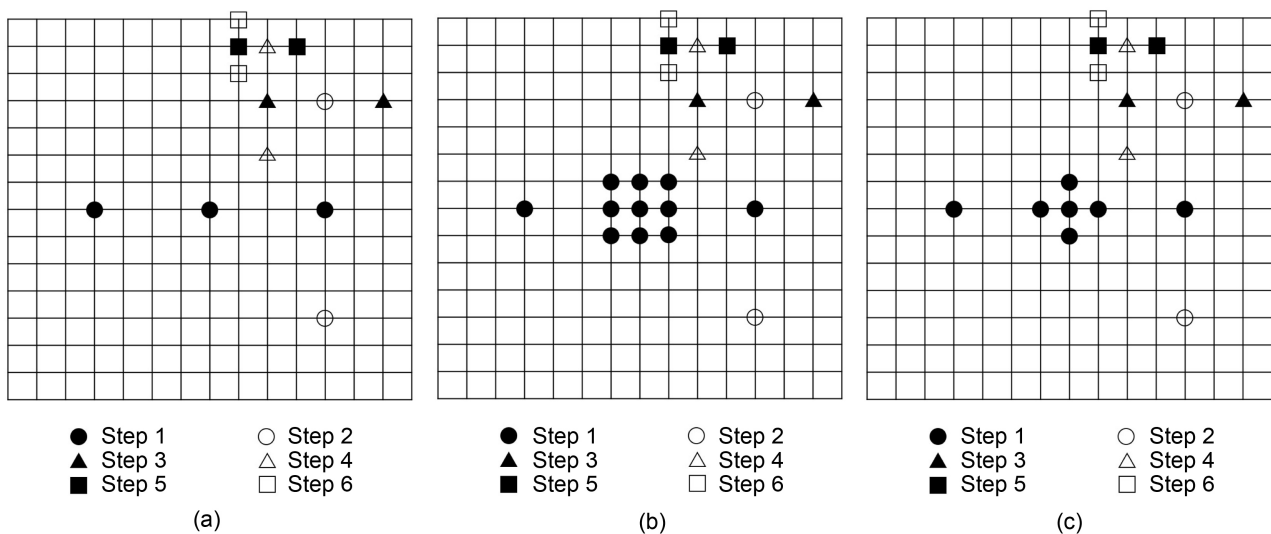


Figure 3. Search algorithms for motion estimation [18]. (a) Orthogonal search algorithm; (b) Modified orthogonal search algorithm; (c) Enhanced modified orthogonal search algorithm.

Step 1: An extra small diamond search pattern around the initial point is added in addition to the original horizontal search points in the first step of OS algorithm. If the initial point (0, 0) of the search window has the minimum SAD value, the block is assumed to have zero movement and search will be terminated. Otherwise, the point which has the minimum SAD value is selected as the new initial point.

Step 2: Step size is halved if the minimum SAD point is any one of the points in the small diamond pattern. 2 new search points at a distance of step size in the horizontal direction from the initial point together with the initial point itself are

compared. The initial point is moved to the winning point.

Step 3: Two points in the vertical direction at a distance of the step size from the initial point are compared with the initial point. The initial point is moved to the winning one.

Step 4: Halve the step size until it is smaller than 1, etc.

The early termination of the algorithm at step 1 helped to reduce the computations significantly also retain the simplicity and regularity of EMOS algorithm. The number of search points needed to find the motion vector for EMOS as compared with the MOS, OS and TSS algorithms.

### 2.2. Predictive Block-Matching Motion Estimation Scheme

Predictive block-matching motion estimation scheme was introduced in 1991, namely predictive pattern search (PPS) [8] [9]. PPS has two prediction modes: inter-block prediction and inter-frame prediction. Inter-block prediction assumes neighboring blocks usually have the same or similar moving direction and distance. Therefore, via the reference of a nearby motion vector, the search window for the current block can be reduced, hence reduces the number of search points. Inter-frame prediction assumes a current block moves to the same or similar direction at the same or a similar distance compared with the corresponding block in the previous frame. **Figure 4** demonstrates the way inter-block and inter-frame prediction works.

Predictive Enhanced Modified Orthogonal Search (PEMOS), performs both inter-block and inter-frame prediction. The estimation of the current vector can be calculated using an autoregressive model given by [8] [9]:

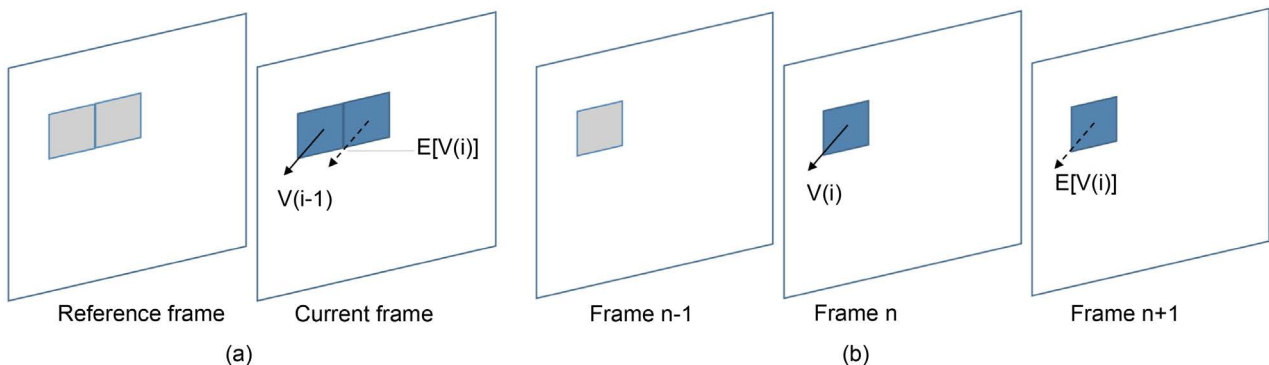
$$E\{V_{i+1}(m,n)\} = \sum_{p,q \in \theta} \alpha_{p,q} V_i(m-q, n-q) \tag{1}$$

where  $\{\alpha_{p,q}\}$  is a set of prediction coefficients and  $\theta$  can be a causal, semi-causal or noncausal set defined by [8]:

Causal set:  $\theta = \{p, q : q > 0 \forall p\} \cup \{p, q : q = 0, p > 0\}$

Semi-causal set:  $\theta = \{p, q : q > 0 \forall p\} \cup \{p, q : q = 0, \forall p \neq 0\}$

Noncausal set:  $\theta = \{p, q : \forall (p, q) \in \delta\}$



**Figure 4.** Predictive pattern search. (a) Inter-block predictive pattern search scheme; (b) Inter-frame predictive pattern search scheme.

According to [8] and [9], for inter-block prediction, a casual model is used. For inter-frame prediction, any one of the three might be used.

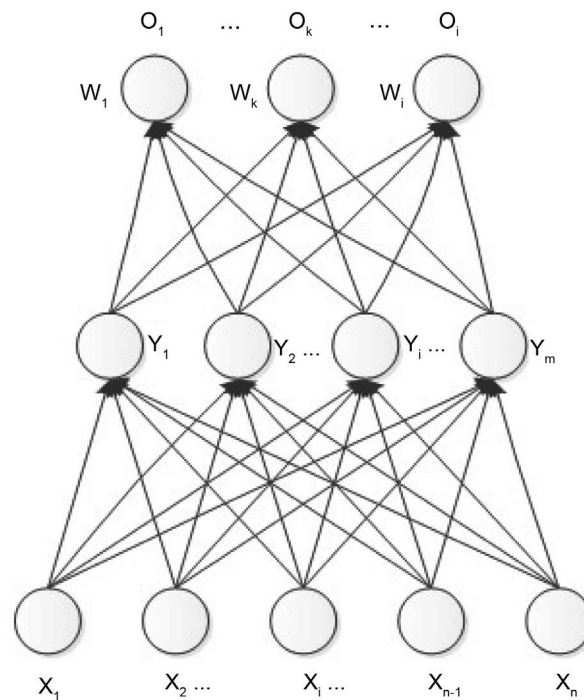
Predictive search scheme has been added to different search algorithms and proved effective. For example, [10] proposed predictive three step search, which adds predictive block-matching scheme to three step search. The results show that predictive three step search significantly reduces computational complexity, with negligible decrease on image quality.

Another example is [11], which added predictive block-matching scheme to EMOS. EMOS is believed to be one of the most efficient block-matching algorithms [6]. According to [11], the proposed algorithm used Zhang and Zafer's prediction model to estimate the motion vectors, therefore the first step of EMOS can be skipped, *i.e.*, a smaller search window is used. If the estimated motion vector is very small, an even smaller search pattern will be used. As a result, PEMOS can reduce computational complexity up to 30%, also with very slight decrease on image quality.

### 2.3. Back Propagation Neural Network

Rumelhart and McClland published a book namely Parallel Distributed Processing [19], discussed the algorithm Error Back Propagation in detail. Today back propagation neural network has become one of the most widely used artificial neural network models [13]. It is widely used in the field of function approximation, pattern recognition, data mining, system identification, automation technology and so on [14].

**Figure 5** shows a typical single hidden layer back propagation neural network



**Figure 5.** An example of back propagation neural network.

model. It includes an input layer, a hidden layer and an output layer.

For the hidden layer, we have:

$$y_i = f(\text{net}_i), j = 1, 2, \dots, m \tag{2}$$

where

$$\text{net}_j = \sum_{i=0}^n v_{ij} x_i, j = 1, 2, \dots, m \tag{3}$$

For the output layer, we have:

$$o_k = f(\text{net}_k), k = 1, 2, \dots, l \tag{4}$$

where

$$\text{net}_k = \sum_{j=0}^m w_{jk} y_j, k = 1, 2, \dots, m \tag{5}$$

In Equations (2) and (3), the transform function  $f(x)$  can be sigmoid functions similar as Equations (7) or (8), according specific applications.

$$f(x) = \frac{1}{1 + e^{-x}} \tag{6}$$

$$f(x) = \frac{1 - e^{-x}}{1 + e^{-x}} \tag{7}$$

### 3. Proposed Neural Network for Predictive Block-Matching

A back propagation neural network model with 5 inputs, 1 hidden layer which contains 5 neurons, and 1 output is used in this report (Figure 6).

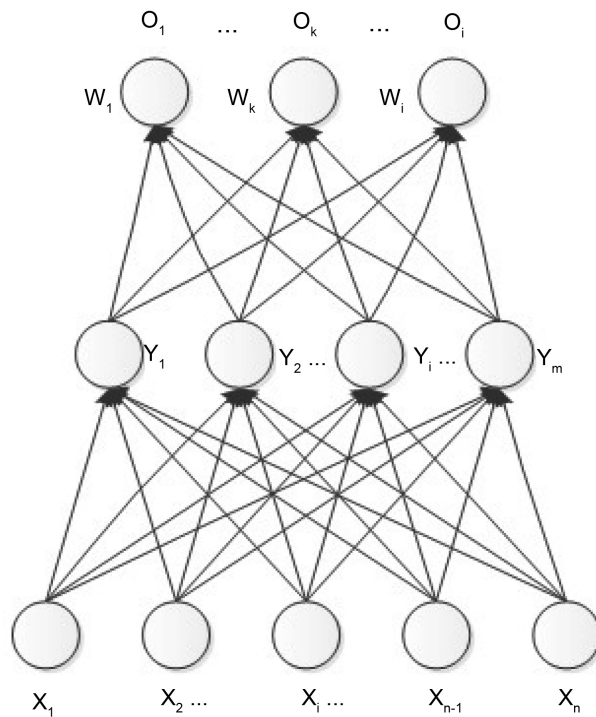


Figure 6. Proposed back propagation neural network.

The five inputs are: the motion vector at the same place in the previous frame, the motion vectors at top left, top, top right and left of the current block in the current frame. The number of hidden neurons is decided applying the following equation [20]:

$$\begin{aligned} & \text{Number of hidden neurons} \\ & = \frac{1}{2}(\text{inputs} + \text{outputs}) \cdot \sqrt{\text{number of training patterns}} \end{aligned} \quad (8)$$

The number of training patterns was assumed to be 4. Although the actual training patterns are more than 4, in order to keep the simplicity of neural network model, we consider motion vector values (positive, positive), (positive, negative), (negative, positive), (negative, negative) as 4 patterns. The training data contains 2000 sets of values generated using full search in the platform MATLAB. A linear transform function is used for computation simplicity.

#### 4. Results and Discussion

In order to find out the effectiveness of neural network prediction, three block-matching algorithms introduced in Section 0 were tested, respectively full search, three step search and enhanced modified orthogonal search. We recorded prediction accuracy with different prediction schemes, image quality in terms of peak signal to noise ratio (PSNR) and energy efficiency in terms of average search points required per block.

In our test, for traditional prediction scheme, the mean value of inter-block and inter-frame prediction is used for the estimation of motion vector  $E[V_{i+1}(m, n)]$ .

Several videos were tested using different block-matching search algorithms. **Table 1** shows the prediction accuracy of the proposed back propagation neural network compared with traditional prediction scheme. **Tables 2-7** illustrate average search points needed per block, and the average peak signal-to-noise ratio (PSNR) which represents the image quality for different video and different search algorithms.

**Table 1.** Prediction accuracy comparison between neural network and traditional method.

Video	Bus	Akiyo	Claire	Foreman	Stefan	Highway	Mobile
Neural network	54.3%	91.1%	52.4%	74.5%	39.6%	85.9%	55.7%
Traditional	84.9%	97.3%	78.0%	91.9%	61.9%	95.2%	86.1%

**Table 2.** PSNR (dB) performance with full search (search window  $\pm 7$ ).

Video	Bus	Akiyo	Claire	Foreman	Stefan	Highway	Mobile
Without prediction	37.01	47.77	46.70	40.29	38.49	41.38	37.68
Neural network	36.47	47.77	46.70	39.96	38.19	41.31	37.67
Traditional	36.12	47.77	46.70	39.86	37.66	41.23	37.66



**Table 3.** Average search points per block with full search (search window  $\pm 7$ ).

Video	Bus	Akiyo	Claire	Foreman	Stefan	Highway	Mobile
Without prediction	116.1	183.9	166.1	166.1	183.9	183.9	183.9
Neural network	39.18	41.03	37.39	37.39	42.14	41.03	41.03
Traditional	39.16	41.03	47.39	37.37	32.14	41.03	41.03

**Table 4.** PSNR (dB) performance with three step search (search window  $\pm 7$ ).

Video	Bus	Akiyo	Claire	Foreman	Stefan	Highway	Mobile
Without prediction	36.54	47.74	46.69	40.05	38.22	41.13	37.59
Neural network	36.28	47.56	46.51	39.75	37.71	41.14	37.54
Traditional	36.22	47.62	46.55	39.74	37.55	41.14	37.32

**Table 5.** Average search points per block with three step search (search window  $\pm 7$ ).

Video	Bus	Akiyo	Claire	Foreman	Stefan	Highway	Mobile
Without prediction	20.05	20.89	19.34	19.60	21.03	21.20	20.96
Neural network	15.91	15.72	15.27	15.51	15.84	16.02	15.79
Traditional	15.93	15.72	15.27	15.50	15.84	16.02	15.78

**Table 6.** PSNR (dB) performance with EMOS (search window  $\pm 7$ ).

Video	Bus	Akiyo	Claire	Foreman	Stefan	Highway	Mobile
Without prediction	36.22	47.62	46.55	39.75	37.55	41.08	37.58
Neural network	36.22	47.70	46.63	39.73	37.73	41.08	37.59
Traditional	36.14	47.70	46.63	39.75	37.54	41.06	37.29

**Table 7.** Average search points per block with EMOS (search window  $\pm 7$ ).

Video	Bus	Akiyo	Claire	Foreman	Stefan	Highway	Mobile
Without prediction	11.03	6.35	6.72	9.74	10.61	11.48	110.49
Neural network	8.60	4.67	5.02	7.80	7.70	7.80	6.75
Traditional	8.80	4.67	5.01	8.03	7.87	8.18	7.75

**Table 3** and **Table 6** show that using prediction schemes does have negative effect on image quality, causing the drop of average PSNR. Videos with larger movements such as “Bus” and “Stefan” appears to be more affected compared

with videos with smaller movements such as “Akiyo” and “Claire”. However the average search points decrease to around a quarter using predicting schemes. Compared with traditional prediction scheme, neural network prediction can provide slightly higher image quality, because of its higher prediction accuracy as is shown in **Table 1**.

Using predictive block-matching schemes on three step search and EMOS shows similar results, as is illustrated in **Tables 4-7**. Orthogonal search (OS) can be seen as an improved version of TSS, and EMOS is improved OS. EMOS is already one of the most efficient block-matching algorithms. However, using predictive schemes is still an effective way to further improve its efficiency. Using neural network to predict the motion vector again is more accurate than the traditional approach, and leads to higher image quality without adding too much computation complexity.

**Table 8** illustrates the computational complexity of the neural network for each block. Take EMOS as an example, assuming number of required search points per block is 7. The required numerical operation includes 3584 add or sub and 1792 absolute value for each block. The additional calculation complexity for the neural network is negligible compared with the existing calculation. Considering the improvement on prediction accuracy and image quality, it is very feasible to use neural network for motion vector prediction.

## 5. Future Work

As the proposed back propagation neural network requires little calculation, in the future, implementation on real-time system might be carried out. The simplicity of the neural network model makes it not difficult to be implemented on either FPGA systems or embedded systems with general processors. Its performance in practice is to be found out and compared with the MATLAB simulation results after that.

## 6. Conclusion

This report proposed a predictive model for block-matching based on neural network. This model is a back propagation neural network with 5 neurons, aiming to improve the prediction accuracy of predictive block-matching schemes, with little computation cost. The results approve that the proposed model is effective. Firstly, Compared with the traditional predictive scheme, neural network shows great advantage in prediction accuracy, especially when the video has large movements. In addition, we did three sets of tests, respectively: without using prediction,

**Table 8.** Computational complexity of back propagation neural network.

Video	Multiplication	Add/Sub
Hidden layer	25	20
Output layer	5	4
Total	30	24

using neural network prediction, and using traditional prediction scheme. For each case, three widely used block-matching algorithms were tested. Using neural network results in better image quality than using traditional prediction, and the value average search points stay the same or become less. In some cases, the image quality is better than without using prediction. Because of the simplicity of the neural network model, it can be easily implemented on existing designs, and requires very little computational power. In conclusion, using the proposed neural network is an effective way to improve the image quality of video encoding systems using predictive block-matching algorithms, and the cost on computational complexity is very small.

## Acknowledgements

The work was partly supported by the Zhejiang Natural Science Foundation under grant No. LY13F010014.

## References

- [1] Akyildiz, I.F., Melodia, T. and Chowdhury, K.R. (2007) A Survey on Wireless Multimedia Sensor Networks. *Computer Networks*, **51**, 921-960. <https://doi.org/10.1016/j.comnet.2006.10.002>
- [2] Akyildiz, I.F., Melodia, T. and Ghosal, D. (2008) Wireless Sensor Network Survey. *Computer Networks*, **14**, 32-39.
- [3] Yick, J., Mukherjee, B. and Chowdhury, K.R. (2007) Wireless Multimedia Sensor Networks: A Survey. *IEEE Wireless Communications*, **52**, 2292-2330.
- [4] Chen, T.C., Lian Jr, C. and Chen, L.G. (2007) Hardware Architecture Design of an H. 264/AVC Video Codec. *Proceedings of the 2006 Asia and South Pacific Design Automation Conference*, Yokohama, 24-27 January 2006, 750-757.
- [5] Chen, T.C., Chien, S.Y. and Huang, Y.W. (2006) Analysis and Architecture Design of an HDTV720p 30 Frames/s H. 264/AVC Encoder. *IEEE Transactions on Circuits and Systems for Video Technology*, **16**, 673-688. <https://doi.org/10.1109/TCSVT.2006.873163>
- [6] Chee, C.S., Jambek, A.B. and Hussin, R. (2012) Review of Energy Efficient Block-Matching Motion Estimation Algorithms for Wireless Video Sensor Networks. 2012 *IEEE Symposium on Computers & Informatics (ISCI)*, Penang, 18-20 March 2012, 241-246. <https://doi.org/10.1109/ISCI.2012.6222702>
- [7] Huang, Y.W., Chen, C.Y. and Tsai, C.H. (2006) Survey on Block Matching Motion Estimation Algorithms and Architectures with New Results. *Journal of VLSI Signal Processing Systems for Signal, Image and Video Technology*, **42**, 297-320. <https://doi.org/10.1007/s11265-006-4190-4>
- [8] Zafar, S., Zhang, Y.Q. and Baras, J.S. (1991) Predictive Block-Matching Motion Estimation for TV coding. I. Inter-Block Prediction. *IEEE Transactions on Broadcasting*, **37**, 97-101. <https://doi.org/10.1109/TBC.1991.1492730>
- [9] Zhang, Y.Q. and Zafar, S. (1991) Predictive Block-Matching Motion Estimation for TV Coding. II. Inter-Frame Prediction. *IEEE Transactions on Broadcasting*, **37**, 102-105. <https://doi.org/10.1109/11.99095>
- [10] Amirpour, H., Mousavinia, A. and Shamsi, N. (2013) Predictive Three Step Search (PTSS) Algorithm for Motion Estimation. 2013 *8th Iranian Conference on Machine*

*Vision and Image Processing (MVIP)*, Zanjan, 10-12 September 2013, 48-52.

<https://doi.org/10.1109/IranianMVIP.2013.6779948>

- [11] Yan, Z.G. and Welsen, S. (2016) Predictive Enhanced Modified Orthogonal Search Algorithm with Two Patterns for Wireless Video Sensor Network. *Asia Pacific Conference on Multimedia and Broadcasting*, Bali, 17-19 November 2016, 60-65.
- [12] Ahmed, Z., Hussain, A.J. and Al-Jumeily, D. (2011) Mean Predictive Block Matching (MPBM) for Fast Block-Matching Motion Estimation. *3rd European Workshop on Visual Information Processing*, Paris, 4-6 July 2011, 67-72.
- [13] Demuth, H.B., Beale, M.H. and De Jess, O. (2014) *Neural Network Design*, Martin Hagan.
- [14] Schmidhuber, J. (2015) Deep Learning in Neural Networks: An Overview. *Neural Networks*, **61**, 85-117.
- [15] Koga, T. (1981) Motion-Compensated Interframe Coding for Video Conferencing. *Proceedings of National Telecommunication Conference*, New Orleans, 29 November-3 December 1981, G5.3.1-G5.3.5.
- [16] Soongsathitanon, S., Woo, W.L. and Dlay, S.S. (2005) Fast Search Algorithms for Video Coding Using Orthogonal Logarithmic Search Algorithm. *IEEE Transactions on Consumer Electronics*, **51**, 552-559.
- [17] Smetkar, S.P. and Talbar, S.N. (2010) Fast Motion Estimation Using Modified Orthogonal Search Algorithm for Video Compression. *Signal, Image and Video Processing*, **4**, 123-128. <https://doi.org/10.1007/s11760-009-0104-9>
- [18] Pandian, S., Bala, G.J. and Anitha, J. (2011) Enhanced Modified Orthogonal Search for Motion Estimation. *Recent Advances in Intelligent Computational Systems*, Trivandrum, 22-24 September 2011, 907-910.
- [19] Rumelhart, D.E. and McClelland, J.L. (1988) *Parallel Distributed Processing*.
- [20] Kalogirou, S.A. and Bojic, M. (2000) Artificial Neural Networks for the Prediction of the Energy Consumption of a Passive Solar Building. *Energy*, **25**, 479-491.



**Submit or recommend next manuscript to SCIRP and we will provide best service for you:**

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.

A wide selection of journals (inclusive of 9 subjects, more than 200 journals)

Providing 24-hour high-quality service

User-friendly online submission system

Fair and swift peer-review system

Efficient typesetting and proofreading procedure

Display of the result of downloads and visits, as well as the number of cited articles

Maximum dissemination of your research work

Submit your manuscript at: <http://papersubmission.scirp.org/>

Or contact [jcc@scirp.org](mailto:jcc@scirp.org)