

Research Article

A Novel Compression Technique for Multi-Camera Nodes through Directional Correlation

**Saima Shaheen,¹ M. Younus Javed,¹ Muid Mufti,² Shehzad Khalid,³
Aasia Khanum,⁴ Shoab A. Khan,¹ and M. Usman Akram¹**

¹Department of Computer Engineering, College of Electrical & Mechanical Engineering, National University of Sciences & Technology, Islamabad 44000, Pakistan

²ID Technologies, 6205-10th Street S.E., Calgary, AB, Canada T2H 2Z9

³Bahria University, Islamabad 44000, Pakistan

⁴Department of Computer Science, Forman Christian College, Lahore 54000, Pakistan

Correspondence should be addressed to Saima Shaheen; saima1282@yahoo.com

Received 8 September 2014; Revised 17 February 2015; Accepted 24 February 2015

Academic Editor: Yun Liu

Copyright © 2015 Saima Shaheen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A major fraction of multimedia stream contents tends to be redundant and leads to wastage of storage capacity and channel bandwidth. In order to eliminate surplus data, standard video compression algorithms exploit spatial and temporal correlation present in video sequence. However, in case of a multisensor network, intersensor statistical redundancy is the most significant factor in acquiring efficient link utilization as well as making resultant findings valuable to the end user. In this paper, an extension to our previously proposed scheme has been presented to accomplish performance goals of a multisensor environment. Standard MPEG codec has been used to accomplish distributed motion compensation in prespecified directions known as directional correlation. Video frame correlation has been estimated locally at the camera node as well as across different nodes, defined as node communication strategies. Further, receiver feedback assists in quality control after reconstitution by decoder assessment. Results estimated have been analyzed for saving ratios and multimedia quality. Results analysis illustrates increased gains in frame quality and compression saving, achieved through reducing node displacement from the reference node (N_R).

1. Introduction

In recent era of deploying multimedia applications in distributed environment, advancements have been made in video compression algorithms to attain high transmission rates. Multimedia data constitutes rich audio and video patterns extensively redundant with useless repetitions. A digital video is a sequence of frames, normally presented at regular time interval so that human eye can perceive fluid motion. Each individual frame is an image with basic element called a pixel. A digital image is obtained by quantizing a continuous image in a two-dimensional frame.

Further, involving a third temporal dimension to a two-dimensional digital image t that denotes time gives a frame sequence as shown in the following:

$$f(m, n, t) = Q[g(x_0 + m\Delta x, y_0 + n\Delta y, t_0 + t\Delta t)], \quad (1)$$

where Q is a quantization operator, x_0 and y_0 are the origins of image plane, while x and y are quantization plans in horizontal and vertical dimensions, and m , n , and t are the discrete values with range $0, 1, 2, 3, \dots, n$ for individual frames, respectively.

Digitization of the spatial coordinates and of the amplitude called image sampling and gray-level quantization, respectively, results in high spatial and temporal correlation (spatio-temporal sampling). Encoding of these similarities is further carried out by registering the differences within a frame (spatial) and between frames (temporal). Spatial encoding is performed by exploiting the fact that the human eye is unable to distinguish fractional color differences than changes in brightness. Correspondingly, temporal compression takes into account the change in the pixel value from one frame to the next based on the interframe resemblance. At the root level, compression is carried out when an input

video stream is analyzed from the viewer perspective and useless information is discarded. Bit codes are assigned to various video events; more frequent events are assigned; short length codes and long codes are associated with rare occurrences. The processing steps of signal analysis, quantization, and variable length encoding are collectively called video compression. Deploying video compression in a multisensor network is a critical task; sensor nodes require enormous care to design an optimum performance methodology due to inherent limitations of these tiny, energy constrained nodes. In a video sensor network, huge amount of multimedia data from large number of terminals requires significant link bandwidth. Normally, a video compression algorithm deployed on each video sensor independently; it involves eradication of temporal as well as spatial redundancy from each video sequence. However this still possesses great repetition due to intersequence statistical redundancy among the data load coming from different video sensor nodes. Eliminating this repetition leads to improved compression efficiency as well as link saving. Likewise efficiency is further assured by keeping the communication among the video sensor nodes at minimum.

The concept of distributed source coding (DSC) has been deployed in a number of ways to resolve performance issues of multisensor setup. In DSC, encoders are designed to be simple and fast, while a central decoder is deployed as an entity strong enough to do fast and efficient joint decoding. This encoder-decoder complexity encapsulation is the major essence of DSC concept. Previously, we proposed an idea as a blend of DSC methodology and an extension of a standard video codec MPEG for a multiterminal environment [1]. For experimentation purposes, a multicamera setup was deployed where camera sensor nodes in a wireless network were allowed to interact at different levels (communication strategies 1 and 2). In strategy 1, minimal interaction among the sensor nodes was allowed, while sensor nodes were allowed to interact more in strategy 2. These communication strategies were meant to exploit the frame temporal correlation. Communication strategy 1 was supported with maximum link saving and optimum reconstructed video quality.

In this paper, node interaction strategy 1 has been embedded with decoder feedback to further enhance video picture quality after being reconstituted. Receiver has been functionalized to assist in statistical quality estimation which is fed back to encoder to readjust the correlation as well as camera displacement parameters. This modification to the basic idea better achieves the performance objectives of a multiterminal environment, since node communication strategy 1 assures the minimum node interaction that reduces the power consumption and link saving is achieved. Moreover, improved video quality is another accomplishment that magnifies the algorithm productivity at lower bitrates.

This paper is organized as follows. Section 2 summarizes the related work and Section 3 provides theoretical background. The proposed methodology and results are discussed in Sections 3 and 4, respectively. Finally Section 5 concludes the idea with identification of future research directions.

2. Related Work

Domain of multiterminal video compression has remained a foremost target for research contributions.

About three decades ago, Slepian and Wolf introduced the idea of DSC with its practical limitations identified [2]. Later on, researchers made valuable additions to the basic idea through extensive experimentations in this domain. A number of practices are proposed to tackle various design issues of multiterminal video compression using DSC approach. In some methodologies [3–5], each encoder independently deploys the DSC approach to utilize temporal correlation for a single video sequence only. Such encoder designed is simple enough to get better error resilience. Likewise, some researchers suggest the scheme of image resolution to take advantage of intersensor statistical redundancy [6]. Images are encoded at low resolution while superresolution methodologies are used for reconstruction. However, estimating correlation between low resolution images results in the lowest coding gains that can better be achieved through deploying the theory of multiterminal source coding [2, 7, 8].

Other methodologies define multiterminal setup with some limitations on various settings of multisensor objects like camera; for instance, cameras are bound to be located along a horizontal line and field objects are assumed to be within some defined camera range. Others [9, 10] have defined a lower bound on a number of cameras for some video scene reconstruction. Some have considered camera sensors to be mounted to specific locations to compute correlation among sensors findings [11]. Researchers have also worked out the mathematical models for depth estimation at decoder ends. In [12], the idea of depth estimation has been proposed to compute correlation among camera sensor views. An approach based on Whyner-Ziv coding has encapsulated the concept of DSC as a simple and fast encoder is coupled with a computationally complex decoder [13]. This scheme has made use of image based rendering applications but has ignored the factor of camera position correspondence. Further, feedback approach has also been proven much productive in eliminating internode redundancy [14]. In such feedback approach, central decoder is designed to provide feedback to one of the encoders.

One of the remarkable contributions is the concept of MTVC (multiterminal video coding) [15, 16]. Numbers of ideas have been associated with this scheme; in their pioneer work, the idea of epipolar geometry has been applied to get node correlation. Subsequently, model-based approach has been recommended to estimate corresponding points of two node findings [17]. They have defined an MTVC framework which shows high performance gains at low bitrates. A remarkable effort is MATLAB simulation of epipolar geometry [18] that can better be utilized to compute epipolar lines and camera premises. One of the recent ideas proposed is the multiterminal video coding with the help of low-resolution depth camera [19]. Though it is a fresh idea, it acquires low performance gains.

Yang et al. presented their idea of two terminal video coding [20]. It exhibits multiterminal source coding of two correlated video sequences to save the sum rate over independent

coding of both sequences. Previously, we have proposed a framework for multiterminal video coding by adapting standard video codecs [1]. That piece of work involves the idea of motion compensation among the data findings of different network nodes that resulted in irredundant data stream that was further compressed to achieve high compression ratios. The idea in this paper is the modification to one of the interaction strategies proposed earlier that has shown the optimum performance in the last experimentation supported through high quality gains and extensive link saving.

3. Proposed Methodology

The major essence of the DSC lies in the design of low complexity simple encoders deployed independently at each video sensor node (VSNo) and a computationally intense and complex decoder sufficient to make joint decoding of commutative bit stream coming out of a video sensor network (VSN). The list of performance constraints for a VSN nominates the lesser communication among the VSNoS to be of most significance. The idea that we proposed earlier [1] introduced two ways of VSNo intercommunication. One of these has been proven to involve lesser interaction among VSNoS to attain lower power consumption as well as better media quality. Such idea of minimum collaboration among VSNoS has been extended in this paper with some modifications embedded in system core functionalities as well as an innovation of receiver feedback introduced.

3.1. Node Interaction Strategy. Being the better interaction scheme with valuable outcomes, communication strategy 1 has been chosen for further extension. The scheme describes VSNo interaction in a way that one node is initialized as the reference node (N_R) undergoes MPEG encoding. Initially, a group of pictures (GOP) of N_R is generated and its first frame is intracoded, later used to estimate the subsequent frames of its own GOP. This first intracoded frame of N_R is further used in motion compensation of other nodes video sequences in a distributed fashion as shown in Figure 1. It should be noted that only initial frame of N_R is intracoded and assists in reconstruction of initial frames of other nodes video sequences. These initial frames of other VSNoS tend to estimate the subsequent frames of their own video GOPs. These all local GOPs constitute a grand GOP; size is estimated using relation given in

$$\text{size}(\text{GOP}_D) = \text{size}(\text{GOP}_L) * N_V, \quad (2)$$

where GOP_D and GOP_L are GOP for the whole distributed environment and local to each VSNo, respectively, and N_V represents number of VSNoS implanted.

3.2. Directional Correlation. Deploying the idea of multi-terminal video coding through MPEG extension involves distributed motion compensation to compute internode correspondence among the dispersed camera nodes. This idea has been used to avail the benefits of nodes correlation as a function of their angle of vision (AOV) with respect to the scene captured.

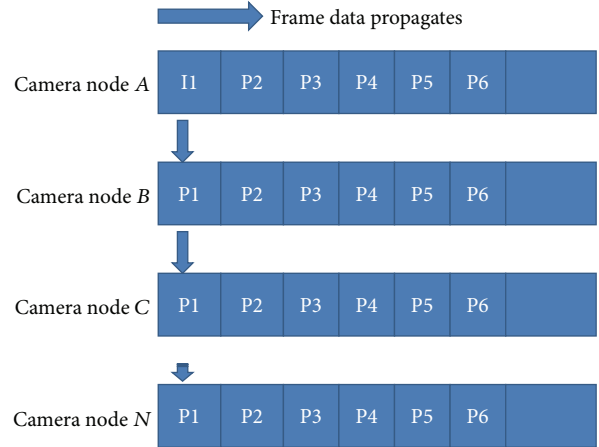


FIGURE 1: Node interaction strategy 1. Nodes exchange data of I-frames only.

The idea of directional correlation (DiC) elaborates the fact that higher degree of correlation is seen among the camera nodes placed in a specific frame of reference capturing the scene at a particular AOV. The mechanism of DiC begins with selection of a node as a reference node N_R . The camera nodes implanted around the N_R are correlated specifically with other nodes present in the same proximity. For instance, camera nodes placed in the right of N_R are correlated with all the nodes present in the right domain and vice versa. Motion compensation through DiC can detect slight variations in macroblocks in the subsequent frames; hence it attains better estimation. This is attributed to the fact that parametric values belonging to a specific camera view exhibit higher degree of association to parameters in the same domain or the views in the same direction, captured later.

While testing the idea of DiC, we use a camera grid, with each camera node placed at a precomputed displacement from N_R . Precalibrated camera nodes placed in left and right of N_R begin to capture the video scene and stopped after a specific interval. From each sensor, data captured is fed to encoder module where video sequences captured from the right of N_R are processed with a right reference view and left frame sequences are estimated using the left reference view. Here, N_R must be selected with immense care. It should be a central point in the camera setup (capturing maximum possible details of the video scene), for instance, a node facing the front view of any subject. Selection of best possible N_R guarantees the optimum estimation and better detection of slight variations in frame contents. In deploying the concept of DiC, the most significant factor is AOV of a candidate node that also derives the node displacement “ d ” factor (will be used later). For instance, consider N_R as a stationary node capturing the front view of a subject, and a candidate node capturing right view of subject with higher values of AOV exhibits minimum correlation and it descends further on increasing the angle. Below the threshold AOV, further reduction tends to increase the degree of correlation or similarity between the camera views unless it becomes zero for N_R . Smaller angle yields more compression gain and

ultimate improved frame quality. So if it is desired to capture more details and to attain high saving ratio, then a denser sensor network has to be built with camera nodes closely spaced, facing the N_R with small AOVs. In fact, it shows a tradeoff between commodity cost (number of camera nodes implanted) and saving ratio or media quality. DiC module fits in the encoder layer where each sensor data is directionally correlated and a collective bit stream is fed to the channel from the VSN.

Figure 2 presents the algorithm processing steps. Encoder layer transmits the cumulative bit stream to central decoder for efficient decoding that assists in quality control through feedback loop. A multiterminal domain has been modeled as a collection of camera nodes facing a subject/scene at a specific AOV. N_R is selected in accordance with the specifications indicated earlier. Initial frame of the GOP of N_R is used to compute frame differences for its own GOP as well as for the initial frames of all the surrounding camera nodes. These initial frames act as the anchor frames for the GOPs of their neighbor nodes in the same direction. Finally, all the motion vectors and frame residues are compressed and a collective bit stream is fed to the channel for approaching the receiver end deployed with central decoder.

3.3. Receiver Feedback Mechanism. Following the concept of DSC, all the encoder nodes encode their video sequences (locally) separately in accordance with internode communication whereas a single central decoder is designed for joint decoding of the cumulative bit stream on receiver end. This central decoder is functionally equipped with efficient decoding as well as quality estimation tool in order to maintain an optimum quality level for better reconstruction.

Frame quality is assessed against a threshold quality level, computed as peak signal to noise ratio (PSNR) given as the ratio of original source signal power to the power of noise induced. Quality levels laying on the threshold edge and below are reported to encoder layer on transmitter end to restore fluctuations. Encoder layer instantly responds to this feedback event through varying the nodes displacements with respect to N_R . Video frame quality can further be enhanced through reducing the factor “ N ” that defines interdistance of the two anchor frames. This suffices well for the quality control at the receiving end.

Decoder feedback events are categorized as severe quality alarms and slight alarm characterized by statistical measure of reconstructed video quality. Severe alarms are medicated through varying the angle of displacement where video sequences are captured for the new node position, acknowledged through decoder notification of quality improvement. Slight quality degradation may protrude as a result of channel noise or some missed frame while capturing. These quality alarms may well be eliminated through refreshing the reference information for current video interval; that is, instead of sending frame differences only, video frames are intracoded.

Besides, block matching algorithms (BMAs) parameters may undergo reconfiguration to pinpoint the finer details of distributed video frames. For instance, macroblock size is reduced to assure better matching and search space is

broadened to facilitate exhaustive search for the best matching candidate block.

4. Experimental Results and Discussion

A variety of experiments have been conducted for in-depth assessment of proposed algorithm. Efficacy of the algorithm is analyzed through compression metrics like saving ratio (compression percentage, CP) and objective quality measures like mean squared error (MSE), PSNR, and so forth. PSNR is an objective quality metric used to measure the quality of reconstructed video frame. Higher values indicate better quality. It is computed via MSE, for instance, for two $M \times N$ video frames f and f' where the former is original frame and the latter is the reconstructed version or approximation of the original. MSE is defined in

$$\text{MSE} = \frac{1}{MN} \left(\sum_x \sum_y [f'(x, y) - f(x, y)] \right). \quad (3)$$

The following equation shows the relation between PSNR and MSE:

$$\text{PSNR} = 10 \log_{10} \left(255 * \frac{255}{\text{MSE}} \right). \quad (4)$$

Similarly, CP or saving ratio achieved by the compression system can be computed as

$$\text{CP} = 100 - \left(\frac{f'}{f} \right) \times 100, \quad (5)$$

where f' is compressed and f is original frame data bits. Higher CP denotes higher compression gain achieved and optimum link saving.

4.1. Data Set. Data set used in the experimentation exhibits a variety in background and foreground contents variation. Algorithm is fed with real time captured video samples reflecting stationary multiview images as well as image sequences with high motion contents. A subject face sequence was captured by three camera nodes implanted at variable displacements from the preselected N_R , to exploit the higher coding gains of DiC.

Video sequences captured were categorized in accordance with ascending the internode displacement, shown in graphical results identified through various threshold relations. Besides, camera nodes are also positioned to capture the scene with moving objects and stationary images from different AOVs. In other video samples, subjects were told to gesture slightly with moving hands to create minor variations in frame contents and vice versa.

4.2. Experimental Setup. Camera nodes are made to capture subject face sequences to assess performance of the proposed algorithm against finer details detection. Experimental setup comprises a grid of camera nodes, entrenched in a way such that each node is capturing the video scene with various AOVs characterized by certain parametric displacements from the reference point.

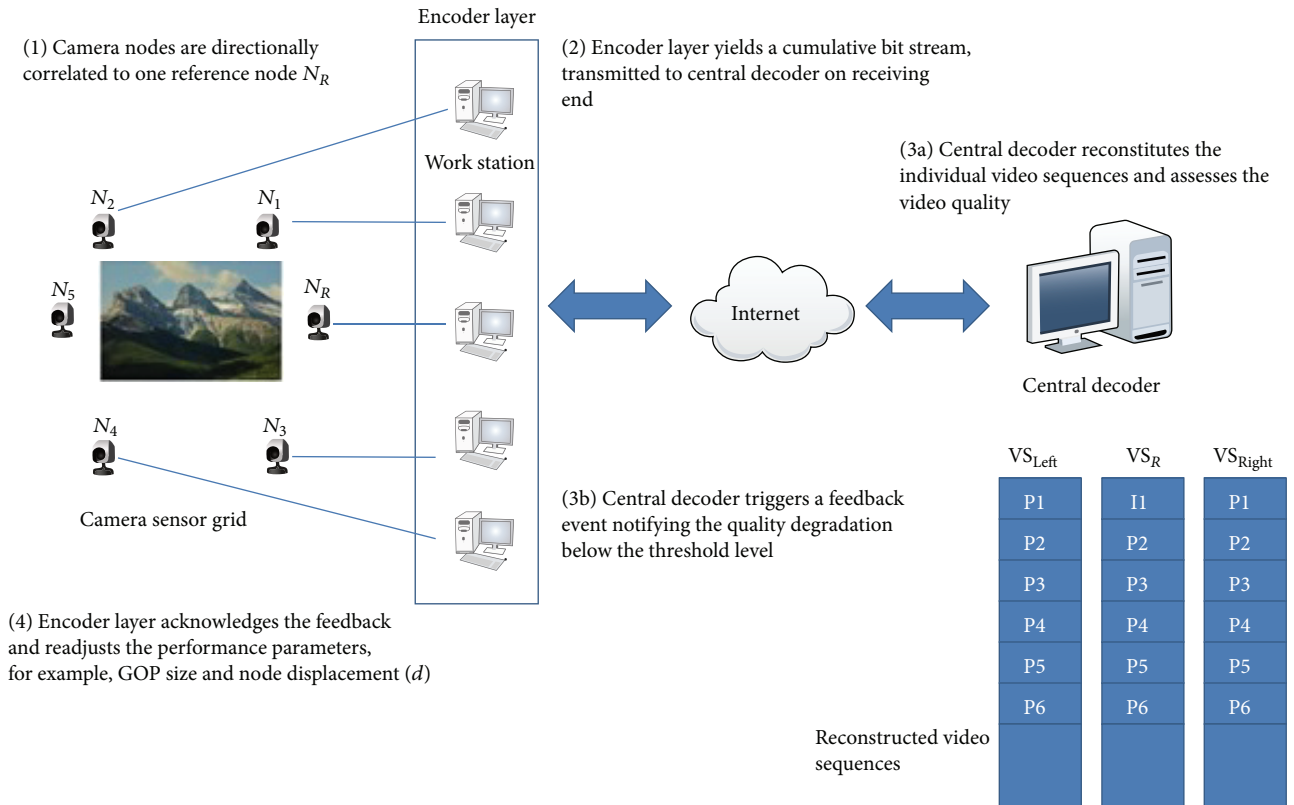


FIGURE 2: Algorithm processing steps describe distributed source coding implementation with individual system components interacting to process multiterminal video results.

4.2.1. *Reconstructed Video Quality.* To evaluate the mechanism of DiC, face sequences are primarily chosen as the video samples. Figure 3 shows the video quality gain measured with respect to the camera node displacements from the N_R , where d and T denote candidate node displacement and threshold displacement from reference for attaining acceptable quality level, respectively. T denotes the point with maximum possible distance from the reference point to measure directional correlation with N_R .

Typical saw-tooth formation of the quality curve is apparent for sample video sequence. It is a snapshot of quality estimates for two GOPs of right camera node video sequence.

It indicates that, with the beginning of a GOP, first frame undergoes motion estimation using reference information from the N_R and attains a quality gain of 32.3 dB for the top most curve. Later on, it becomes an anchor frame for subsequent frames of the GOP. All the successive frames exhibit an exceptional quality trend as a sudden rise of 0.4 dB and then drop gradually for each following frame. This saw-tooth formation is justified through the fact that frames belonging to the same GOP show higher correlation. Quality gets its lower bound at 32.35 dB for the last frame of this GOP and slightly descends more for next GOP anchor frame at 32.21 dB. This new anchor frame gets estimated through N_R reference information and node correlation parameters are restored. Additionally, this figure shows the

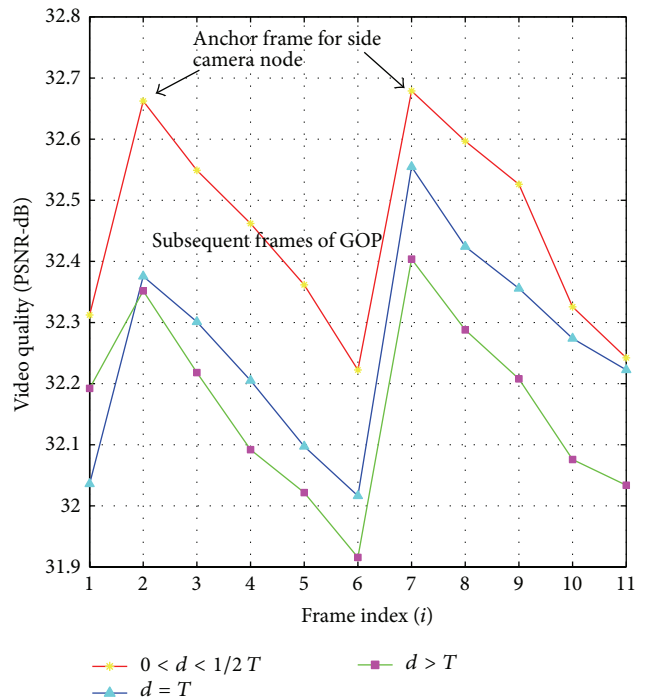


FIGURE 3: Quality of reconstructed video sequences for various node displacements (d) in accordance with maximum possible displacement for acceptable video quality (T).

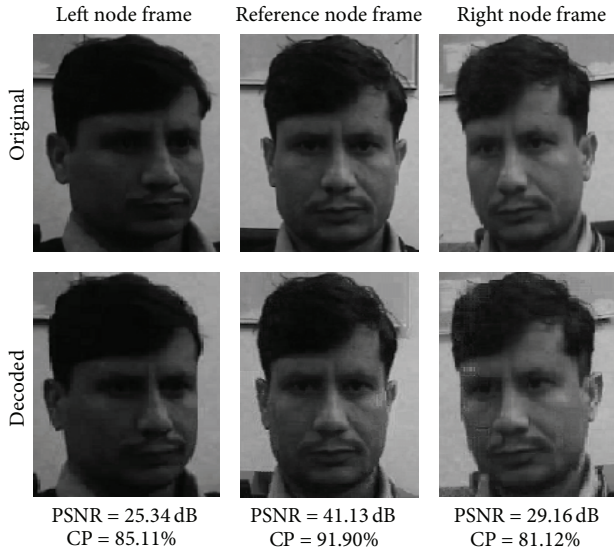


FIGURE 4: Video frames compared for visual perception with statistical quality estimation. The top row shows original frames and the bottom one represents reconstructed format.

node quality trend for three camera positions. It is illustrated that comparatively higher quality gains are achieved for the video data of the nodes with less displacements. Bottom curve shows the result for nodes comparatively placed at longer displacements. Conclusively, camera nodes placed within the proximity of T attain the higher quality gains, shown through the top most curve, and the gain drops down as the nodes move away from T (from 32.7 dB to 32 dB). T is derived from AOV; it is the distance between a candidate node and N_R that varies widely depending upon the nature of video scene captured. Generally, higher angles result in less node data correlation and reduced media quality whereas lower angles result in higher node correspondence that yields maximum saving ratio, since two highly correlated camera views yield large motion vector contents and small fraction of frame differences that result in high compression gain (i.e., lesser number of bits to compress the frame in question).

Figure 4 further demonstrates the visual quality of a face video sequence. It assesses the algorithm capabilities for detecting finer details in video picture with high degree of correlation available. N_R is capturing the front view of the subject while left and right camera nodes are placed along its two sides. Right camera is placed slightly closer to N_R , shown through a minor ascend of approximately 4 dB in media quality and 4% in saving ratio. Intracoded frame of N_R is shown with maximum media quality of 41.1 dB. In DiC, video frames estimated from the frames of N_R possess lower quality as well as link saving as compared with subsequent frames of their GOP due to availability of higher correlation among the same GOP data.

4.2.2. Compression Saving Estimation. In this experiment, again camera nodes are placed at certain distances from the N_R to quantify the upper bound of the link saving achieved.

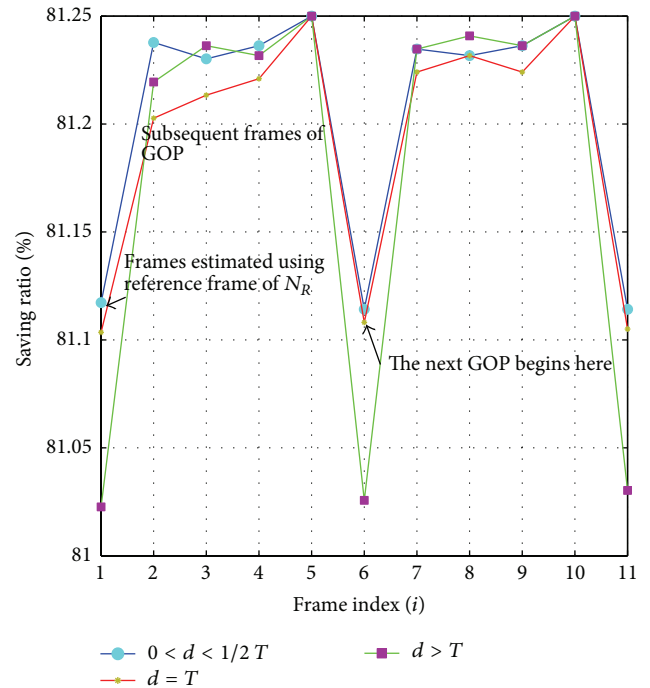


FIGURE 5: Comparative analysis of saving ratio versus node displacements.

For instance, right camera node is selected and its correlation with N_R is computed at various camera positions with certain AOV. Camera node displacement d values are varied with respect to T ; it shows that saving ratio is favored by lesser node displacements (i.e., for $d < 1/2T$) and adding more to displacement (i.e., for $d > T$) reduces the link saving and adds to video data rate as well.

We have constructed a three-terminal video setup in a way that two camera nodes were implanted at certain positions relative to the fixed N_R . Video sequences are captured from all the camera nodes and, later on, compressed through the proposed algorithm. Video frame size is taken as 352×288 pixels and compressed at a frame rate of 25 f/s (inclusive of three camera nodes). Bit rate computed for the above parameters varies in relation to side camera positions (displacement) from the N_R ; that is, as camera nodes are placed closer to the N_R , degree of correlation increases between the frames. Ultimately encoded stream will add lesser number of bits to cumulative bit stream, shown through CP comparison of all three camera placements in Figure 5. Our proposed algorithm comparatively exhibits higher saving ratio achieved than one of the other schemes proposed as MTVC algorithm [17]. Existing methods have worked on synthesized video sequences, created at various angles of separation from zero-degree version. They computed the saving ratios at various bit rates and showed that high saving ratios are achieved at lower bit rates for all the angles. However, in their results, maximum link saving reached up to 60–70% and gradually dropped for higher degree of separation. We have captured the video scene with three different camera positions and assessed each video sequence

TABLE 1: Saving ratio comparison.

	Saving ratio (CP-%)	Bit rate (kb/s)
DiC	81.01	437 kb/s
	81.2	461 kb/s
	81.33	483 kb/s
MTVC	62	150 kb/s
	41	200 kb/s
	19	610 kb/s

captured in accordance with our algorithm. For instance, right camera node is considered; three camera positions are denoted as points A, B, and C. When camera node is at A, B, and C, video bit rates achieved are 437, 461, and 483 kb/s, respectively. Similarly, saving ratio varies as the bit rate and camera position vary but does not drop sharply for higher bit rates, as shown in Table 1.

MTVC results for 15-degree separation version are shown. According to Table 1, MTVC shows comparatively better saving at lower bit rates but it sharply drops at higher bit rates like 19% at 610 kb/s. For DiC results, three different camera positions are assessed and exhibit remarkable link saving. It can be seen that saving ratio tends to increase for closer camera positions but at the cost of network commodities (nodes closely spaced). Numerical values shown are average values for two GOPs. Here, sharp variations in CP for different camera displacements are not seen. Except the first frame of a GOP that is compensated through N_R frame, all the remaining frames show higher link saving. This is due to presence of lesser correlation between two nodes frames than that between the frames of the same GOP. Consequently, large differences in CP of primary GOP frame are averaged out and this seems as slight variation for different camera positions.

4.2.3. Correlation Parameter Tradeoffs. To attain improved link savings, video frames of each side of VSNo are correlated with N_R frame, only once, whereas anchor frames at each VSNo keep on estimating the subsequent frames of their own GOP. Such correlation incorporates the BMA [21]. These BMAs involve motion estimation parameters that possess dramatic effects on frame quality as well as link saving ratio and computational cost. In fact, a tradeoff can be set between these parameters to attain desired performance gains for a VSN. Motion estimation parameters primarily involved are candidate matching block size and search space limits, denoted as correlation size (C_S) and correlation width (C_W), respectively. Figures 6 and 7 show the ultimate distortion in frame quality against different values of C_S and C_W . Keeping smaller C_S assists in capturing the minor variations across frame sequences and ultimately in-depth macroblock search is facilitated.

Minimum distortion and better quality reconstruction are only possible with smaller block size but they add to computational cost. Analysis of search space size sets a threshold at higher C_W values. Increasing C_W minimizes the frame distortion but at higher values (width limits of 11–15) almost constant distortion is seen with slight variations. Like

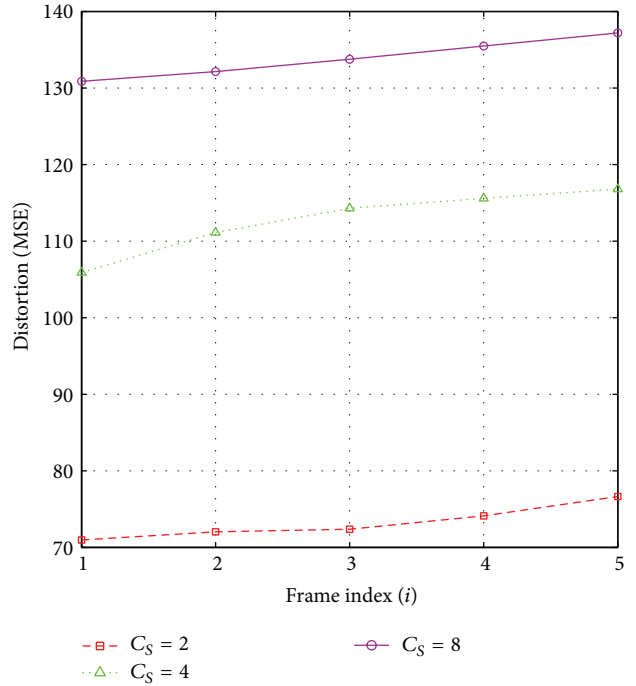


FIGURE 6: Frame distortion as a function of macroblock size.

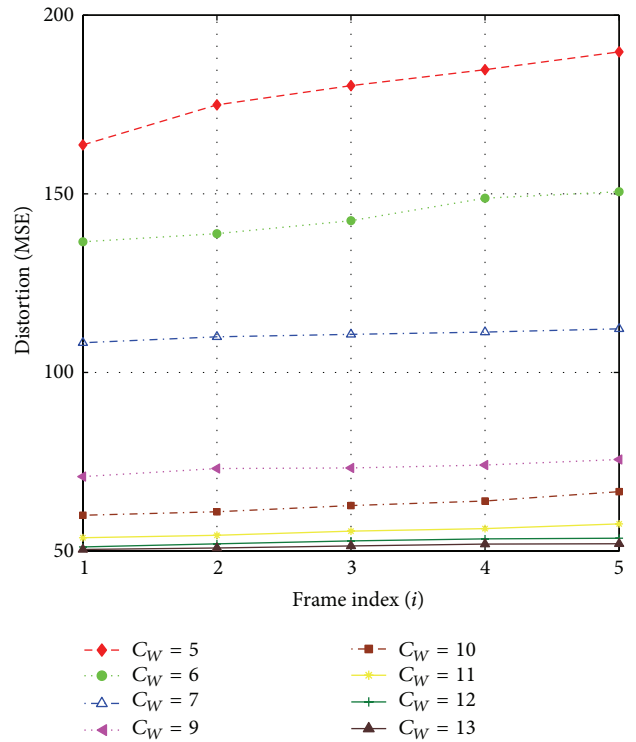


FIGURE 7: Frame distortion as a function of search space size.

C_S , for higher C_W computational time adds up to delay and VSNo power consumption. This cost is mitigated through selecting a smart block size C_S to be searched in suitable search space width C_W for enhanced frame quality as well as power saving.

5. Conclusion and Future Work

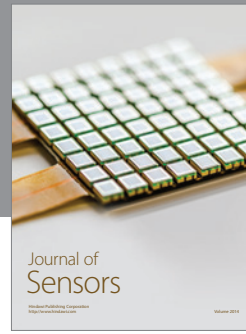
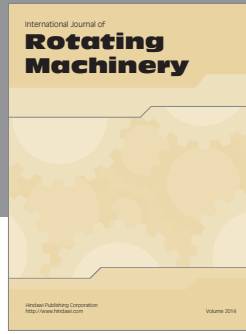
Video processing in multiterminal domain has remained a food for thought for researchers. Various ideas have been proposed for focusing on different performance areas of video processing in distributed domain. One of these ideas is the implementation of standard MPEG codec in multicamera setup that we proposed previously, where nodes correspondence is computed through distributed motion estimation techniques. Idea proposed in this paper is based on MPEG implementation with new concepts of DiC and receiver feedback mechanism supported through extensive experimentations. Experimental outcomes reveal that DiC assures better quality and link saving. Moreover, receiver feedback assists in timely parameter adjustment for ultimate quality enhancement.

Conflict of Interests

The authors declare that they have no conflict of interests regarding the publication of this paper.

References

- [1] S. Shaheen and M. Y. Javed, "An implementation of MPEG in multi-terminal video coding," in *Proceedings of the 6th IEEE International Conference on Emerging Technologies (ICET '10)*, pp. 144–149, IEEE, Islamabad, Pakistan, October 2010.
- [2] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. 19, no. 4, pp. 471–480, 1973.
- [3] B. Girod, A. M. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed video coding," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 71–83, 2005.
- [4] A. Sehgal, A. Jagmohan, and N. Ahuja, "Wyner-Ziv coding of video: an error-resilient compression framework," *IEEE Transactions on Multimedia*, vol. 6, no. 2, pp. 249–258, 2004.
- [5] R. Puri, A. Majumdar, and K. Ramchandran, "PRISM: a video coding paradigm with motion estimation at the decoder," *IEEE Transactions on Image Processing*, vol. 16, no. 10, pp. 2436–2448, 2007.
- [6] R. Wagner, R. Nowak, and R. Baraniuk, "Distributed image compression for sensor networks using correspondence analysis and super-resolution," in *Proceedings of the International Conference on Image Processing (ICIP '03)*, vol. 1, pp. 597–600, September 2003.
- [7] T. Burger, "Multi-terminal source encoding," in *The Information Theory Approach to Communications*, G. Longo, Ed., CISM Courses and Lectures 229, Springer, New York, NY, USA, 1978.
- [8] Y. Oohama, "Gaussian multi-terminal source coding," *IEEE Transactions on Information Theory*, vol. 43, no. 6, pp. 1912–1923, 1997.
- [9] N. Gehrig and P. L. Dragotti, "Different: distributed and fully flexible image encoders for camera sensor networks," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '05)*, pp. 690–693, September 2005.
- [10] N. Gehrig and P. L. Dragotti, "Distributed sampling and compression of scenes with finite rate of innovation in camera sensor networks," in *Proceedings of the Data Compression Conference (DCC '06)*, pp. 83–92, March 2006.
- [11] X. Zhu, A. Aaron, and B. Girod, "Distributed compression for large camera arrays," in *Proceedings of the IEEE Workshop on Statistical Signal Processing*, pp. 30–33, September 2003.
- [12] J. C. Dagher, M. W. Marcellin, and M. A. Neifeld, "A method for coordinating the distributed transmission of imagery," *IEEE Transactions on Image Processing*, vol. 15, no. 7, pp. 1705–1717, 2006.
- [13] Y. Yang, V. Stankovic, W. Zhao, and Z. Xiong, "Multiterminal video coding," in *Proceedings of the Information Theory and Applications Workshop*, San Diego, Calif, USA, January 2007.
- [14] M. Flierl and B. Girod, "Coding of multi-view image sequences with video sensors," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '06)*, pp. 609–612, IEEE, Atlanta, Ga, USA, October 2006.
- [15] B. Song, O. Bursalioglu, A. K. Roy-Chowdhury, and E. Tuncel, "Towards a multi-terminal video compression algorithm using epipolar geometry," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '06)*, May 2006.
- [16] B. Song, A. Roy-Chowdhury, and E. Tuncel, "A multi-terminal model-based video compression algorithm," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '06)*, pp. 265–268, October 2006.
- [17] B. Song, E. Tuncel, and A. K. Roy-Chowdhury, "Towards a multi-terminal video compression algorithm by integrating distributed source coding with geometrical constraints," *Journal of Multimedia*, vol. 2, no. 3, pp. 9–16, 2007.
- [18] <http://www.mathworks.com/matlabcentral/fileexchange/13020>.
- [19] Y. Zhang, Y. Yang, and Z. Xiong, "Depth camera assisted multiterminal video coding," in *Proceedings of the 13th IEEE International Workshop on Multimedia Signal Processing (MMSP '11)*, pp. 1–6, IEEE, Hangzhou, China, November 2011.
- [20] Y. Yang, V. Stanković, Z. Xiong, and W. Zhao, "Two-terminal video coding," *IEEE Transactions on Image Processing*, vol. 18, no. 3, pp. 534–551, 2009.
- [21] <http://www.mathworks.com/matlabcentral/fileexchange/8761-block-matching-algorithms-for-motion-estimation>.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

