

## Research Article

# A Data Hiding Technique to Synchronously Embed Physiological Signals in H.264/AVC Encoded Video for Medicine Healthcare

Raul Peña, Alfonso Ávila, David Muñoz, and Juan Lavariega

*Tecnológico de Monterrey, 64849 Monterrey, NL, Mexico*

Correspondence should be addressed to Raul Peña; [raul.p.ortega@gmail.com](mailto:raul.p.ortega@gmail.com)

Received 2 January 2015; Accepted 24 February 2015

Academic Editor: Cheng-Hong Yang

Copyright © 2015 Raul Peña 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.

The recognition of clinical manifestations in both video images and physiological-signal waveforms is an important aid to improve the safety and effectiveness in medical care. Physicians can rely on video-waveform (VW) observations to recognize difficult-to-spot signs and symptoms. The VW observations can also reduce the number of false positive incidents and expand the recognition coverage to abnormal health conditions. The synchronization between the video images and the physiological-signal waveforms is fundamental for the successful recognition of the clinical manifestations. The use of conventional equipment to synchronously acquire and display the video-waveform information involves complex tasks such as the video capture/compression, the acquisition/compression of each physiological signal, and the video-waveform synchronization based on timestamps. This paper introduces a data hiding technique capable of both enabling embedding channels and synchronously hiding samples of physiological signals into encoded video sequences. Our data hiding technique offers large data capacity and simplifies the complexity of the video-waveform acquisition and reproduction. The experimental results revealed successful embedding and full restoration of signal's samples. Our results also demonstrated a small distortion in the video objective quality, a small increment in bit-rate, and embedded cost savings of  $-2.6196\%$  for high and medium motion video sequences.

## 1. Introduction

Video technology continues to improve the safety and effectiveness in healthcare. Today, physicians and engineers rely on rigorous video-based studies to improve medical practices and procedures. These studies are necessary for the identification of clinical manifestations in patients and for the reduction of errors during medical procedures. Extending the video-based studies to incorporate the analysis of physiological-signal waveforms further enhanced the recognition of clinical manifestations and the reduction of false positive cases. Physicians can rely on simultaneous video-waveform observations to recognize difficult-to-spot signs and symptoms. These observations can also expand the recognition coverage to abnormal health conditions.

The synchronization between the video images and the physiological-signal waveforms is fundamental for enhanced recognition of the clinical manifestations. The identification of signs and symptoms invisible during specific diagnosis is possible with synchronized video-waveform observations.

Physicians are able to diagnose seizures in neonates after creating and analyzing a permanent record [1]. The permanent record needed for this diagnosis contains synchronized video and electroencephalographic (EEG) waveforms.

Three commonly used techniques are suitable for video-signal synchronization [2]. The timestamp-based technique is the most common alternative for video-audio synchronization. This technique inserts time codes at each signal stream. These time codes are also useful for future browsing, storage, and reproduction of permanent records [1]. However, the timestamp-based synchronization involves complex tasks such as the capture/compression of the video signal, the acquisition/compression of each physiological signal, the insertion of timestamps into individual streams, and the use of specialized software for stream synchronization [2].

The second technique relies on synchronization marks. The technique sends synchronization marks from a transmission node. The disadvantage of this technique is the need of an additional assistant communication channel to transmit the synchronization mark. The third technique is

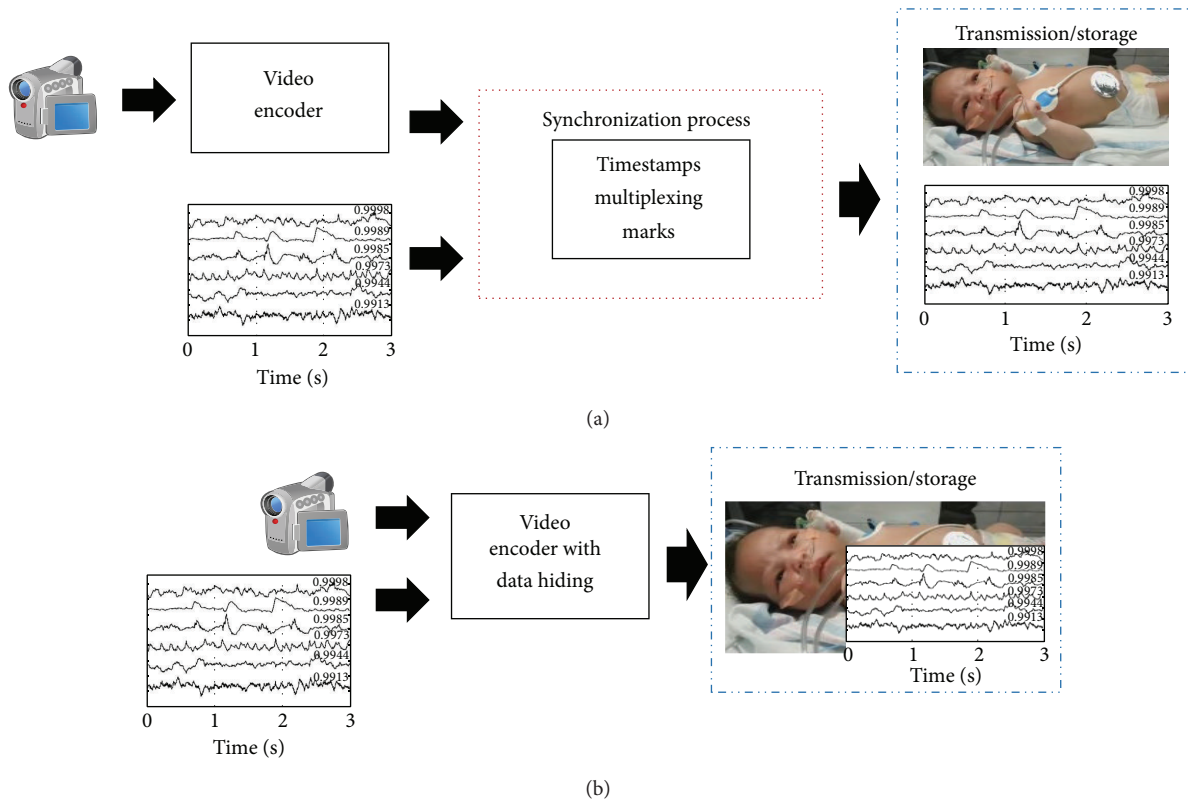


FIGURE 1: A comparison between (a) conventional synchronization process and (b) our proposed approach.

multiplexing. This technique maintains the correlation among the media streams during the transmission process. However, multiplexing-based synchronization usually results in a loss of agility and integrality.

Data hiding is another alternative for video-waveform synchronization. Data hiding has the goal of embedding information into encoded video sequences with a minimum amount of perceivable degradation [3]. The embedded information can be text, pictures, or physiological-signal samples. Data hiding synchronizes the video and signals after hiding each physiological sample at its corresponding video frame in time.

This paper introduces an improved data hiding technique with a larger data-hiding capacity in the context of medical healthcare. Our data hiding technique synchronously embeds physiological-signal samples into H.264/AVC-encoded video sequences. The implementation of data hiding is simpler than other synchronization techniques. The data-hiding technique requires only one encoder to process video and signal's samples. Data hiding also offers the advantage of a unique communication channel for video-audio transmission and requires no complex tasks related to timestamps, synchronization marks, or multiplexing [4, 5]. Figure 1 illustrates a comparison between the commonly used synchronization processes and our synchronization approach based on data hiding. Other important advantage is that our technique is strongly related to trends in secure handling of medical data during signal transmission. Data hiding makes possible

the secure transmission of patient's information over the internet. Important features for secure transmission of personal information are authentication, integrity, and confidentiality [6].

## 2. Background

**2.1. Data Hiding Techniques.** Existing data hiding techniques are able to hide information in the video frames during the video encoding process. In [3], authors proposed a data hiding scheme based on the macroblock's size needed by the H.264/AVC interprediction process. The scheme is able to hide two bits per macroblock and requires the following partitions types:  $16 \times 16$ ,  $16 \times 8$ ,  $8 \times 16$ , and  $8 \times 8$ . The scheme loses no hidden data and may result in bit-rate increments.

In [7], the data hiding scheme relies on constraints associated with the H.264/AVC inter/intraprediction modes. In the interprediction mode, the scheme hides 0 bits at the interprediction mode using the block sizes  $16 \times 8$ ,  $8 \times 16$ ,  $8 \times 4$ , and  $4 \times 8$ . The scheme also hides a 1 bit using the block sizes  $16 \times 16$ ,  $8 \times 8$ , and  $4 \times 4$ . In the intraprediction mode, the scheme hides 0 bits using the block sizes  $16 \times 16$  and  $4 \times 4$ . Hiding a 1-bit value requires the  $8 \times 8$  block size. The scheme has minimum impact on the video quality and controls the distortion degradation by hiding no data in  $4 \times 4$  blocks.

In [8], the data hiding scheme embeds information during context-based adaptive variable length coding (CAVLC). The scheme hides one bit using the trailing ones parity of

TABLE 1: Qualitative metrics and statistics of the data hiding (DH) schemes reviewed.

Reference	DH schemes	DH capacity	Max. PSNR <sub>Y</sub> distortion	Max. bit-rate distortion
[3]	Forcing block-type partitions	2 bits per MB (interframes)	-1.4 dB (40 kbps)	7000 bps
[7]	Grouping block-type partitions	1 bit per MB (inter/intraframes)	-0.09 dB	4.65%
[8]	Parity of trailing ones CAVLC	1 bit per MB (intraframes)	-2.57 dB	0.039%
[9]	Quarter-pixel search positions	1-4 bits per MB (interframes)	-0.03 db	1.31%
[10]	Last nonzero coefficient parity	1 bit per MB (interframes)	Approx. -0.01 dB	1200 bps
[11]	Motion vectors and mode sel.	1 bit per frame (interframes)	0.03 dB	0.98%

the CAVLC code-word. The CAVLC process results in moderate visual degradation and maintains the overall size of the video stream.

In [9], the proposed technique exploited quarter-pixel motion estimation process to hide data. The scheme hides one bit by modulating the best search points of a subblock. The rate-distortion cost is introduced to reduce both the impact on the video quality and the increment in the bit-rate after search point adjustments. The hiding capacity is dependent on the content of the video sequence.

In [10], the authors proposed a data hiding scheme based on an adaptive method. The method hides one bit using the last nonzero coefficient parity after quantization of a  $4 \times 4$  luma block. The scheme relies on an adaptive rather than a fixed point for data embedding. The scheme results in a proportionally direct behavior between the bit-rate and capacity size and between the bit-rate difference and the amount of embedded data.

In [11], the proposed data hiding scheme relies on motion vectors and mode selection. The scheme only hides one bit per frame. This scheme embeds data using the macroblock search regions with a left area restriction for a 1 bit and a right area restriction for 0 bits.

In [4, 5], the authors demonstrated that existing data hiding schemes in [8, 10] successfully embedded audio into encoded video sequences with minor impact on video image quality and bit-rate.

The desirable features of a data hiding technique suitable for embedding physiological signals into encoded video sequences are a large data hiding capacity, a low impact in video quality, and a minor effect in bit-rate of the video. Table 1 compares the data hiding schemes previously reviewed considering three metrics: hiding data capacity, the maximum PSNR<sub>Y</sub> (objective video quality), and bit-rate distortion. These schemes offer relatively low values for the three metrics. Therefore, the main limitation resides in the data hiding capacity. In [9], the quarter-pixel motion estimation scheme offers the highest data hiding capacity using an  $8 \times 8$  partition. This scheme also offers very low PSNR<sub>Y</sub> distortion and less than 1.32% of bit-rate distortion. Our proposed technique extends the quarter-pixel motion estimation scheme to satisfy the data hiding capacity needs and to ensure the low PSNR<sub>Y</sub> and bit-rate distortions.

**2.2. Motion Estimation in H.264/AVC.** Motion estimation (ME) is an important element in the H.264/AVC interpretation process. For a given frame, the ME goal is to find

TABLE 2: Examples of medical applications.

Reference	Medical application	Phys. signals
[1]	Silent neonatal seizures	EEG
[12]	Sleep-related breathing disorders	EEG, ECG, SpO <sub>2</sub> , and CO <sub>2</sub>
[13]	Physical activity and congestive heart failure	Accelerometers
[14]	Patient safety in anesthesia operating rooms	Vital signs

the best predictions for both levels: macroblock (MB) selection and motion vector (MV) estimation. A MB is an array of  $16 \times 16$  pixels. The MB selection process assumes the partitioning illustrated in Figure 2. Each partition contains a MV value. Equation (1) shows how to select the best block partition by calculating the Lagrangian rate distortion ( $J_{\text{mode}}$ ) optimization. In this equation,  $\lambda_{\text{mode}}$  is the Lagrangian multiplier, SSD is the sum of the squared difference between the original and the reconstructed block, and  $R$  is the number of bits of MB parameters such as quantization parameter, header, motion vectors, and residue coefficients:

$$J_{\text{mode}} = \text{SSD} + (\lambda_{\text{mode}})(R). \quad (1)$$

The motion estimation process computes motion vectors for each macroblock partition found in each video frame. At a given frame, the ME process searches for the new MB position of each MB located in the reference frame. The ME process calculates motion vectors based on these new MB positions and encodes these vectors in the encoded frames. Figure 2 illustrates the three ME stages to compute a MV. The first ME stage identifies the best MB position at the integer-pixel mesh. The second ME stage identifies the best MB position at the half-pixel mesh based on the best integer-pixel position. The third ME stage identifies the best MB position at the quarter-pixel mesh based on the best half-pixel position. The selected position becomes the final MV value.

**2.3. Application Examples of Video-Based Medical Care.** This section presents additional examples of medical applications related to the simultaneous observation of video and physiological-signal waveforms. Table 2 shows the names of the applications and the specific physiological signals needed for simultaneous correlation with the video. In [1], physicians take advantage of synchronized EEG recordings with video to

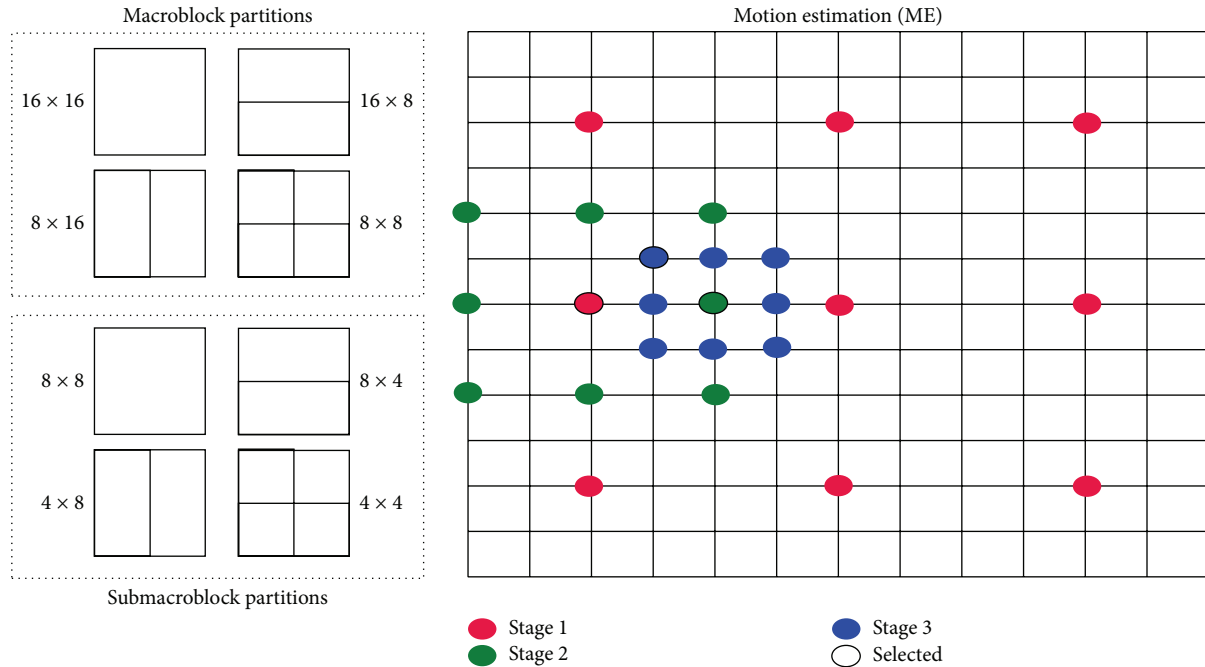


FIGURE 2: Macroblock partitions and search points in motion estimation.

correlate clinical manifestations such as lip smacking, fixing of eyeballs, and cyclic leg movements.

In [12], physicians take advantage of synchronized digital video recordings to identify nocturnal breathing anomalies usually undetected by standard polysomnography. Successful identification of these anomalies requires the correlation among EEG recordings, oxygen saturation ( $\text{SpO}_2$ ), endtidal  $\text{CO}_2$  level, in-video leg movement, and in-video rapid eye movement (REM).

In [13], the purpose is to assess the validity of a new physical-activity monitor in the context of congestive heart failure. This monitor utilizes body-fixed accelerometers to distinguish among activities such as body postures, sitting, standing, normal walking, stairs walking, cycling, and wheelchair driving. These in-video activities are correlated with the accelerometer waveforms to assess the correct operation of the activity monitor.

In [14], clinical investigators perform rigorous studies to enhance patient safety in operating rooms. The investigators first elaborate a permanent feedback record containing the in-video health delivery process, vital signs, and other signals. Then, the investigators reproduce this permanent record to observe and to assess the health delivery process. The synchronization of the video and the physiological-signal waveforms is fundamental for the identification of factors resulting in adverse events.

### 3. The Proposed Data Hiding Technique

Our proposed technique hides streams of data samples into encoded video sequences. The implementation of our technique is a set of software routines added to the original

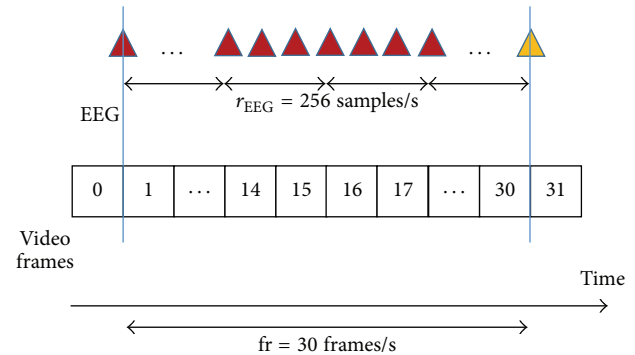


FIGURE 3: Synchronization process.

H.264/AVC codec. Figure 3 illustrates the synchronization between the video and the physiological-signal waveforms of EEG samples. Our technique synchronizes video and EEG signals by hiding samples of these signals at their corresponding frame in time.

**3.1. Encoding Process in Our Data Hiding Technique.** Our technique modifies the H.264/AVC interprediction process to hide the samples of the physiological signals. Our technique relies upon the quarter-pixel motion estimation process, part of the H.264/AVC encoder, to hide these samples into encoded video frames. Our technique divides the search points into four groups to improve the data hiding capacity of the original technique [9]. Equation (2) presents the proposed partitioning of search points and the binary assignment for each partition. The expression in (3) indicates how to select

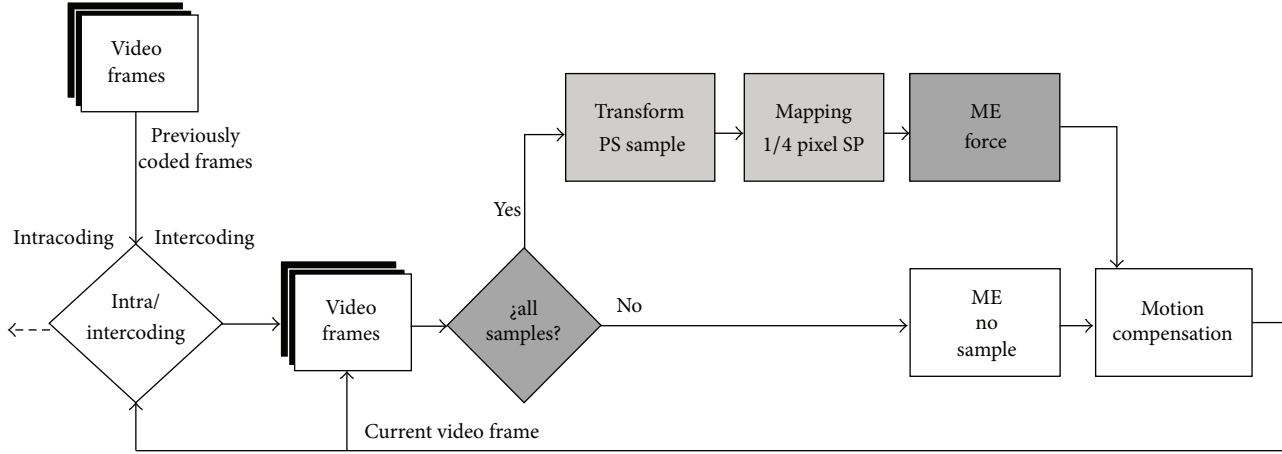


FIGURE 4: The block diagram of our proposed data hiding technique based on H.264/AVC encoder.

**Algorithm:** Embed a sample.  
**Input:** Data of the sample in bit-pairs according to (2).  
**Result:** Embedded sample in a frame.

- (1) **for** NewSample **do**
- (2)   curblock  $\rightarrow$  current macroblock
- (3)   **for** MacroblockPartition **do**
- (4)     **switch** Mapping **do**
- (5)       case 00: Quarter-Pixel Search Position =  $i \mid \min(Ji), i \in \{5, 6\}$
- (6)       case 01: Quarter-Pixel Search Position =  $i \mid \min(Ji), i \in \{1, 2\}$
- (7)       case 10: Quarter-Pixel Search Position =  $i \mid \min(Ji), i \in \{3, 4, 7, 8\}$
- (8)       case 11: Quarter-Pixel Search Position =  $\{0\}$
- (9)     **end switch**
- (10)   **end for**  $\rightarrow$  until complete all partitions of the current macroblock
- (11) **end for**  $\rightarrow$  until complete all samples

ALGORITHM 1: Algorithm of our data hiding encoding technique.

the search point in each group with the minimum distortion cost. In (3),  $Ji$  is the Lagrangian rate distortion parameter:

$$\text{searchpos} \in \begin{cases} G1 = \{5, 6\} & \text{if "00"} \\ G2 = \{1, 2\} & \text{if "01"} \\ G3 = \{3, 4, 7, 8\} & \text{if "10"} \\ G4 = \{0\} & \text{if "11"}, \end{cases} \quad (2)$$

$$\text{searchpos} = \{i \mid \min(Ji), i \in (G1, G2, G3, G4)\}. \quad (3)$$

Our proposed technique, illustrated in Figure 4, repeats the gray blocks until no samples are available. The blocks in gray are the additional routines needed to implement our proposed data hiding algorithm. The alternate path executes the original H.264/AVC motion estimation process. Algorithm 1 presents details of our data hiding algorithm for encoding. Our technique hides the signal samples in the motion vectors of each block partition located at a frame. Each motion vectors hides two bits of a signal sample. Our technique hides no samples in the block types I4 MB and I16 MB due to their association with intraframe prediction.

Our technique is also unable to hide samples into PSKIP blocks due to the lack of motion vectors.

Our technique also incorporates an approach to overcome the data-hiding capacity limitation found in low-motion video sequences. PSKIP blocks are the most common block partition found in encoded low-motion video sequences. A large number of PSKIP blocks limit the data hiding capacity of our technique due to few number of motion vectors found in the low-motion video sequence. Therefore, our proposed technique forces the H.264/AVC encoder to replace the PSKIP block by a  $P16 \times 16$  block partition. This PSKIP replacement adds a motion vector to the data hiding capacity of our technique. The PSKIP replacement also contributes to maintaining the synchronization between the encoded video and the physiological signals. However, this replacement may also result in an increment in the bit-rate of the sequence. Algorithm 2 presents details of the algorithm for low-motion sequences.

**3.2. Decoding Process of Our Data Hiding Technique.** Our data hiding technique extracts the samples of the EEG signal; at the same time, conventional video H.264/AVC decoding

**Algorithm:** Embed a sample in low-motion sequences  
**Input:** Data of the sample in bit-pairs according to (2).  
**Result:** Embedded sample in a frame.

- (1) **for** NewSample **do**
- (2) curblock  $\rightarrow$  current macroblock
- (3) **if** PSKIP partition **do**
- (4) curblock = P16  $\times$  16 partition type
- (5) **switch** Mapping **do**
- (6) case 00: Quarter-Pixel Search Position =  $i \mid \min(Ji), i \in \{5, 6\}$
- (7) case 01: Quarter-Pixel Search Position =  $i \mid \min(Ji), i \in \{1, 2\}$
- (8) case 10: Quarter-Pixel Search Position =  $i \mid \min(Ji), i \in \{3, 4, 7, 8\}$
- (9) case 11: Quarter-Pixel Search Position =  $\{0\}$
- (10) **end switch**
- (11) **end if**  $\rightarrow$  until complete all partitions of the current macroblock
- (12) **end for**  $\rightarrow$  until complete all samples

ALGORITHM 2: Algorithm of our data hiding encoding technique for low motion sequences.

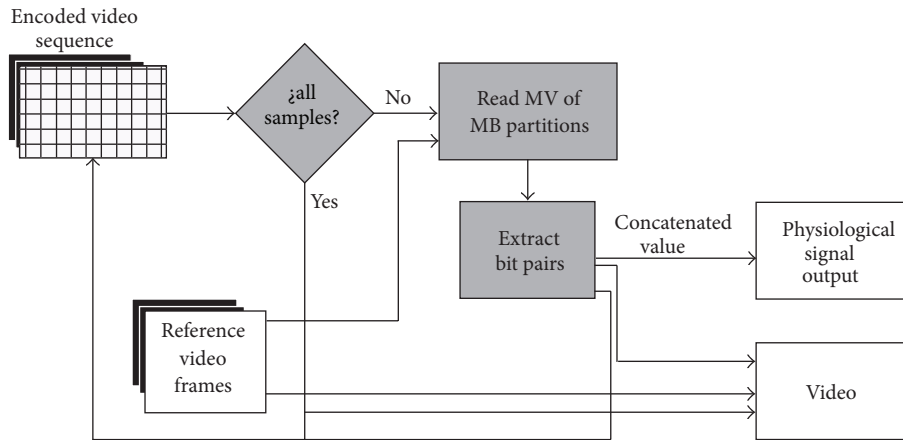


FIGURE 5: Sample extraction during the decoding process.

process takes place. This is an important feature at the time to do synchronized playback of both video and physiological-signal waveforms. The sample extraction algorithm in our technique catches and gathers sample data. This extraction algorithm, illustrated in Figure 5, is repeated as many times as needed to extract all the samples embedded in the video sequence. To do this, our data hiding technique interacts with the H.264/AVC decoding process.

The routine reads the motion vector (MV) of every macroblock partition and inputs the  $MV_x$  and  $MV_y$  components into (4) to identify a binary combination. The  $|MV|\%2$  terms output the quarter-pixel search point values. The routine obtains the bit-pair samples after identifying the G1 to G4 values mapped into (2):

$$\text{type} \in \begin{cases} G1 & \text{if } (|MV_x|\%2 = 1), (|MV_y|\%2 = 0) \\ G2 & \text{if } (|MV_x|\%2 = 0), (|MV_y|\%2 = 1) \\ G3 & \text{if } (|MV_x|\%2 = 1), (|MV_y|\%2 = 1) \\ G4 & \text{if } (|MV_x|\%2 = 0), (|MV_y|\%2 = 0). \end{cases} \quad (4)$$

Algorithm 3 presents details of our proposed decoding algorithm. This decoding process extracts the physiological samples from the encoded video sequences.

#### 4. Results

The experimental setup, illustrated in Figure 6, included a PC, an EEG database, and a set of seven video test sequences. The EEG samples were extracted from the CHB-MIT Scalp EEG Database. This database is located at the PhysioBank digital recordings (<http://www.physionet.org>). The experimental setup included 6 signal electrodes at 256 samples per second and 12-bit sample resolution [1]. The EEG samples generated in one second were embedded into the first 30 frames of each test video sequence to establish synchronization.

The implementation of encoding and decoding processes needed the modification of the JM reference software version 16.1. Table 3 shows the JM configuration. A program was developed to provide and convert samples from the database to the encoder.

The video test sequences had a CIF (352  $\times$  288) resolution and a 4:2:0 YUV format. The name of the test sequences are

**Algorithm:** Extract a sample.  
**Input:** Embedded sample in a synchronous frame.  
**Result:** Data of the sample in bit-pairs according to (2).  
 (1) **for** ExtractSample **do**  
 (2) curblock  $\rightarrow$  current macroblock  
 (3) **for** MacroblockPartition **do**  
 (4) **switch** Mapping **do**  
 (5) case ( $|MV_x| \% 2 = 1$ ) & ( $|MV_y| \% 2 = 0$ ): Bit-pair = 00  
 (6) case ( $|MV_x| \% 2 = 0$ ) & ( $|MV_y| \% 2 = 1$ ): Bit-pair = 01  
 (7) case ( $|MV_x| \% 2 = 1$ ) & ( $|MV_y| \% 2 = 1$ ): Bit-pair = 10  
 (8) case ( $|MV_x| \% 2 = 0$ ) & ( $|MV_y| \% 2 = 0$ ): Bit-pair = 11  
 (9) **end switch**  
 (10) **end for**  $\rightarrow$  until complete all partitions of the current macroblock  
 (11) **end for**  $\rightarrow$  until complete all samples

ALGORITHM 3: Algorithm of our data hiding decoding technique.

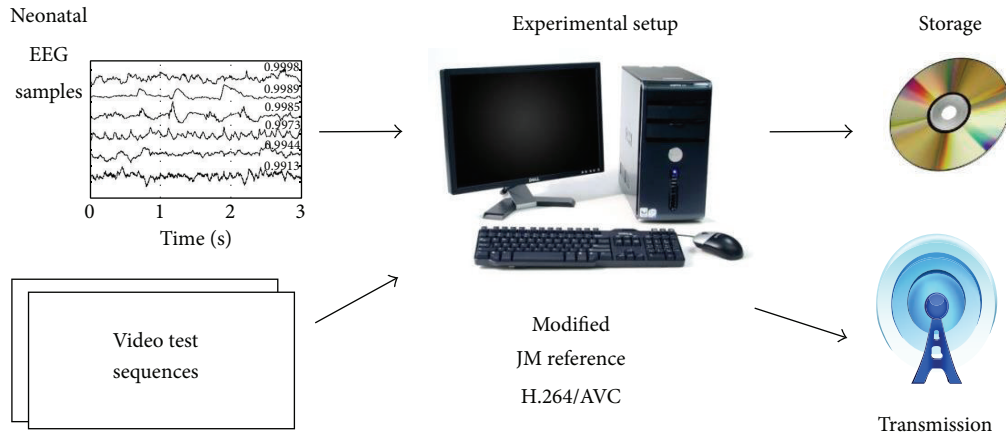


FIGURE 6: Experimental setup for our synchronization data hiding scheme.

TABLE 3: Configuration parameters for JM reference software.

Parameter	Resolution
Profile	Baseline
Frames	30
Motion estimation algorithm	Full search
RD optimization and rate control	Disabled
8 × 8 subblocks	Disabled
Number of reference frames	1
Quantization parameter (QP)	28

akiyo, bridge-far, carphone, football, foreman, mobile, and neonatal. Neonatal is not considered a standard video test sequence. Neonatal was introduced to match the context of the application example related to EEG seizures on neonates. The selected coding structure of the bit stream is “IPPP...” to have an intraframe encoded in the first frame and interframes encoded in the remaining frames.

Metrics to evaluate the effectiveness of our proposed technique are video objective quality, bit-rate difference, embedding cost, and perceptual quality of the image. The peak

signal-to-noise ratio (PSNR), illustrated in (5), is an objective quality metric to report video image degradation.  $M$  and  $N$  are the height and the width of the video frame, respectively.  $I_{ij}$  and  $I_{ij}'$  represent the original pixel and the processed pixel, respectively [5]. The  $PSNR_{Y,diff}$  metric, illustrated in (6), indicates how the luma ( $Y$ ) samples impact the video quality after embedding the physiological samples.  $PSNR_Y'$  represents the impact of luma samples generated by embedding the samples, and  $PSNR_Y$  represents the impact of luma samples generated with the original H.264/AVC encoder:

$$PSNR = 10 \log \frac{\max(I_{ij}^2)}{(1/M) \sum_i^M \sum_j^N [I_{ij} - I_{ij}']^2}, \quad (5)$$

$$PSNR_{Y,diff} = PSNR_Y' - PSNR_Y. \quad (6)$$

Equation (7) shows how to calculate the change of bit-rate (BRI).  $R$  is the original bit-rate and  $R'$  is the embedding samples bit-rate. Equation (8) shows how to estimate the embedded cost  $O_e$ ,  $O_v$  is the data volume generated by the original video coder,  $DHe$  is the data volume generated by embedding the samples, and  $EEGe$  is the embedded data volume [7]. The  $EEGe$  term refers to the data coming from

TABLE 4: Experimental results for seven video test sequences.

Video sequences	Embedded capacity (bits)	Modified MBs	Modified QP-MV	PSNR <sub>Y,diff</sub> (dB)	BRI (%)	Oe (%)
Football	45598	0	6957	0.011	0.4459	-0.3787
Mobile	43598	0	7017	-0.002	0.5837	-0.1273
Foreman	28550	0	6702	0	0.9341	-2.6196
Carphone	23454	0	6699	-0.007	1.3446	-2.2459
Neonatal	16544	944	6364	0.029	5.604	-1.1826
Akiyo	5718	6357	5145	-0.003	33.3143	15.8725
Bridge-far	3650	7391	4662	-0.002	36.575	19.3017

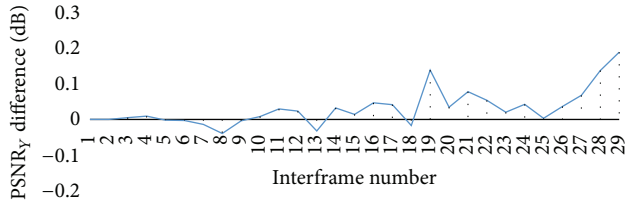


FIGURE 7: PSNR luma difference values for neonatal video sequence.

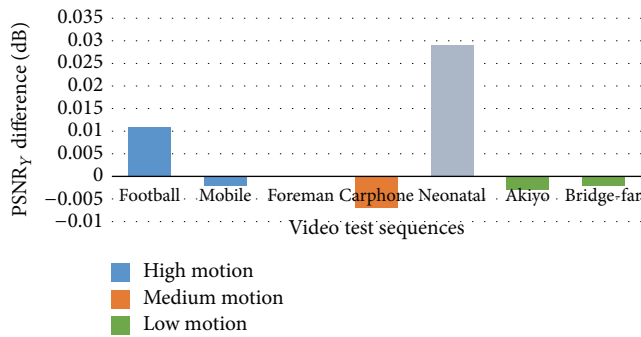


FIGURE 8: PSNR luma difference for seven video test sequences.

the EEG signals. The perceptual quality provides an estimation of subjective quality of the image obtained by visual inspection:

$$BRI = \frac{R' - R}{R} \times 100\%, \quad (7)$$

$$Oe = \frac{DHe - Ov - EEGe}{Ov + EEGe} \times 100\%. \quad (8)$$

Figure 7 compares the PSNR<sub>Y</sub> difference between the original and the embedding data encoding processes. The graph shows a very small difference between PSNR<sub>Y</sub> values for the first 30 frames of the neonatal video sequence. The largest increment in quality was 0.187 dB. This increment appeared in the 29th frame. The largest decrement in quality was -0.039 dB. This decrement appeared in the 8th frame.

Figure 8 demonstrates that the embedding process has minor effect on the video objective quality of seven test sequences. The y-axis represents the PSNR luma difference values expressed in decibels. The positive values above average, 0.011 dB and 0.029 dB, represent a very small improvement in video quality. The negative values, from -0.002 dB to

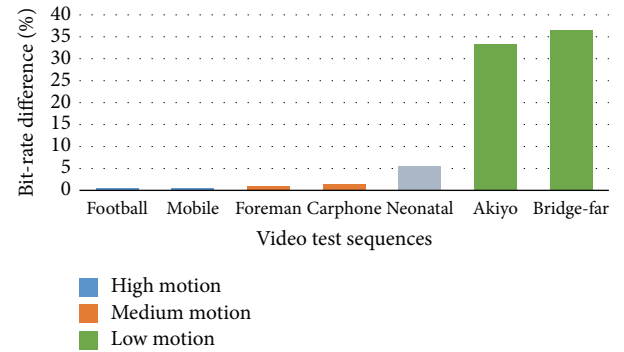


FIGURE 9: Change of bit-rate for seven video test sequences.

-0.007, represent a small degradation in video quality. These degradations and improvements are due to the lack of bit-rate constraint.

Embedding the EEG data into the video sequences produced increments of bit-rate for the seven video test sequences as illustrated in Figure 9. Our experimental setup included high motion, medium motion, and low motion video test sequences. The amount motion of the neonatal video sequence was considered between medium and low. The graph presents the neonatal sequence with gray color to indicate that it is not a standard video test sequence. For the high and medium motion sequences, there are small bit-rate increments. The bit-rate increments ranged from 0.45% to 1.34%. For the neonatal and low motion sequences, the PSKIP block replacement occurred and resulted in bigger bit-rate increments. The bit-rate increments ranged from 5.6% to 36.58%. However, the changes in the bit-rate had no effect on the video-waveform synchronization.

Table 4 shows the results for the seven test sequences in terms of embedded capacity, modified macroblocks, modified motion vectors, PSNR<sub>Y</sub> difference, bit-rate, and the embedded cost. For the high, medium motion and neonatal sequences, the embedded cost ranged from -0.1273% to -2.6196% representing savings in data volume. For the low motion sequences, the embedded cost had an increment in 15.8726% and 19.3017%. These results indicate that our data hiding technique offers both an adequate efficiency for video-signal transmission and savings in the storage of high, medium, and neonatal sequences.

The inspection of subject quality indicated minimum visual artifacts or distortion between original and data-hiding



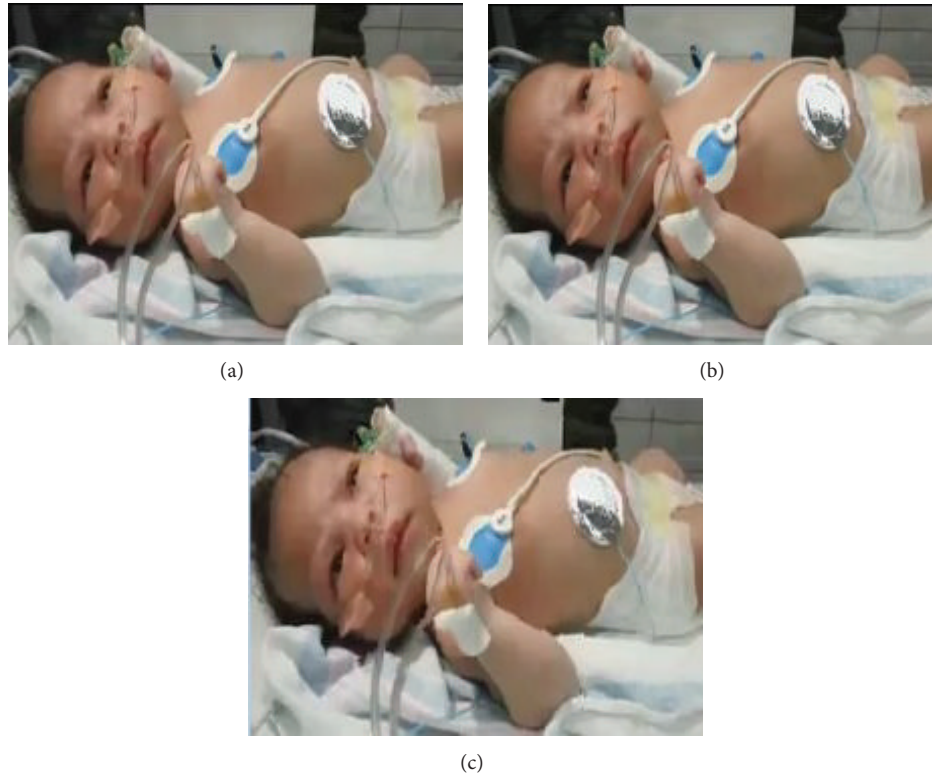


FIGURE 10: 10th frame of neonatal video sequence. (a) Original image, (b) image using video original coding, and (c) image with EEG data.

images. Figure 10 illustrates the perceptual quality of the video frames for the 10th frame of the neonatal sequence. Thus, data hiding generates no significant difference in quality from a human eye perception.

Our technique will offer an adequate performance in the context of the application examples presented in Section 2.3. The video sequences of these applications exhibit sufficient amount of motion. In the neonatal-seizures application, a sufficient amount of motion is needed to identify clinical manifestations like epileptic attacks. In the application about breathing disorders, respiratory and abnormal movements are needed for accurate diagnosis. In the application about congestive heart failures, the motion is associated with the physical activities of the patient. Finally, the amount of motion is associated with the medical staff activity rather than the patient for the application example related to the improvement of medical practices in operating rooms.

## 5. Conclusions

The proposed data hiding technique was demonstrated to be suitable for the medicine healthcare context. Our technique successfully embedded samples of six EEG signals into encoded video sequences with high, medium, and low motion. Our technique also extracted the hidden samples from the encoded video sequences without loss of information. The implementation of our technique required simpler tasks compared to other existing synchronization techniques: (1) less number of encoders and decoders, (2) no timestamps needed, (3) no software needed for synchronization of video and signal streams, and (4) higher data capacity compared

to other data hiding techniques, especially for high motion sequences.

The experimental results demonstrated minimum degradation in video quality and data savings in terms of storage/transmission. The experimental results for high and medium motion video test sequences ranged from  $-0.007$  dB to  $0.011$  dB in PSNR luma difference, from  $0.4459\%$  to  $1.3446\%$  in the bit-rate difference, and from  $-2.6196\%$  to  $-0.1273\%$  in embedded cost. The changes in PSNR<sub>Y</sub> difference, and bit-rate resulted in both no impacts in video-waveform synchronization and minimum distortions in video quality. For storage and transmission purposes, the embedded cost for high and medium motion video sequences represent savings. For low motion video sequences the experimental results ranged from  $-0.003$  dB to  $-0.002$  dB in PSNR<sub>Y</sub> difference, from  $33.3143\%$  to  $36.575\%$  in the bit-rate difference, and from  $15.8725\%$  to  $19.3017\%$  in embedded cost. The changes in bit-rate were higher compared to the high and medium video sequences.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgment

This work was supported by Latin American and Caribbean Collaborative ICT Research Federation (LACCIR) through a research grant (R1209LAC001).

## References

- [1] S. Bhattacharyya, A. Roy, D. P. Dogra et al., "Summarization of neonatal video EEG for seizure and artifact detection," in *Proceedings of the 3rd National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG '11)*, pp. 134–137, Hubli, India, December 2011.
- [2] J. Zhang, Y. Li, and Y. Wei, "Using timestamp to realize audio-video synchronization in real-time streaming media transmission," in *Proceedings of the International Conference on Audio, Language and Image Processing (ICALIP '08)*, pp. 1073–1076, Shanghai, China, July 2008.
- [3] S. K. Kapotas, E. E. Varsaki, and A. N. Skodras, "Data hiding in H. 264 encoded video sequences," in *Proceedings of the 9th IEEE Workshop on Multimedia Signal Processing (MMSP '07)*, pp. 373–376, October 2007.
- [4] X. Li, H. Chen, D. Wang, and X. Qi, "Audio-video synchronous coding based on mode selection in H.264," in *Proceedings of the 4th International Congress on Image and Signal Processing (CISP '11)*, vol. 1, pp. 113–117, October 2011.
- [5] B. Li and M.-Q. Shi, "Audio-video synchronization coding approach based on H.264/AVC," *IEICE Electronics Express*, vol. 6, no. 22, pp. 1556–1561, 2009.
- [6] H.-M. Chao, C.-M. Hsu, and S.-G. Miaou, "A data-hiding technique with authentication, integration, and confidentiality for electronic patient records," *IEEE Transactions on Information Technology in Biomedicine*, vol. 6, no. 1, pp. 46–53, 2002.
- [7] C.-H. Liu and O. T.-C. Chen, "Data hiding in inter and intra prediction modes of H.264/AVC," in *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS '08)*, pp. 3025–3028, May 2008.
- [8] K. Liao, D. Ye, S. Lian, Z. Guo, and J. Wang, "Lightweight information hiding in H.264/AVC video stream," in *Proceedings of the International Conference on Multimedia Information Networking and Security*, vol. 1, pp. 578–582, November 2009.
- [9] H. Zhu, R. Wang, and D. Xu, "Information hiding algorithm for H.264 based on the motion estimation of quarter-pixel," in *Proceedings of the 2nd International Conference on Future Computer and Communication (ICFCC '10)*, vol. 1, pp. V1-423–V1-427, IEEE, Wuhan, China, May 2010.
- [10] Y. Li, H.-X. Chen, and Y. Zhao, "A new method of data hiding based on H.264 encoded video sequences," in *Proceedings of the IEEE 10th International Conference on Signal Processing (ICSP '10)*, pp. 1833–1836, October 2010.
- [11] P. Wang, Z. Zheng, and L. Li, "A video watermarking scheme based on motion vectors and mode selection," in *Proceedings of the International Conference on Computer Science and Software Engineering (CSSE '08)*, vol. 5, pp. 233–237, Wuhan, China, December 2008.
- [12] K. Banno and M. H. Kryger, "Use of polysomnography with synchronized digital video recording to diagnose pediatric sleep breathing disorders," *Canadian Medical Association Journal*, vol. 173, no. 1, pp. 28–30, 2005.
- [13] H. J. G. van den Berg-Emons, J. B. J. Bussmann, H. M. M. Balk, and H. J. Stam, "Validity of ambulatory accelerometry to quantify physical activity in heart failure," *Scandinavian Journal of Rehabilitation Medicine*, vol. 32, no. 4, pp. 187–192, 2000.
- [14] M. B. Weinger, D. C. Gonzales, J. Slagle, and M. Syeed, "Video capture of clinical care to enhance patient safety," *Quality and Safety in Health Care*, vol. 13, no. 2, pp. 136–144, 2004.