Journal of Zhejiang University-SCIENCE C (Computers & Electronics)
ISSN 1869-1951 (Print); ISSN 1869-196X (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Statistically uniform intra-block refresh algorithm for very low delay video communication*

Rui SONG†, Yang-li WANG, Ye HAN, Yun-song LI
(*State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an 710071, China*)
†E-mail: ruiScientif@gmail.com

Received Nov. 23, 2012; Revision accepted Mar. 14, 2013; Crosschecked Apr. 9, 2013

Abstract: This paper focuses on the mechanism underlying the overall delay of a real-time video communication system from the time of capture at the encoder to the time of display at the decoder. A detailed analysis is presented to illustrate the delay problem. We then describe a statistically uniform intra-block refresh scheme for very low delay video communication. By scattering intra-blocks uniformly into continuous frames, the overall delay is significantly decreased, and object changes in the scene could be presented to the end user instantly. For comparison, the overall delay and the peak signal-to-noise ratio (PSNR) performance are tested. The experiment results show that an average of approximately 0.1 dB PSNR gain on average is obtained relative to random intra-macroblock refresh algorithm in H.264 JM, and the end-to-end delay performance is significantly improved.

Key words: Video coding, Intra refresh, Low delay
doi:10.1631/jzus.C1200333 **Document code:** A **CLC number:** TN919

1 Introduction

H.264, like other existing video coding standards, employs hybrid coding techniques including discrete cosine transform (DCT) and motion compensated prediction, and provides high efficiency coding tools that make real-time video transmission possible on a band-limited channel (Yu and Wang, 2010). However, it does not consider end-to-end delay. Rate control algorithms (Liu *et al.*, 2007; Hong *et al.*, 2010; Chang *et al.*, 2012; Lee and Sull, 2012) and network optimization frameworks (Song, 2011; Asioli *et al.*, 2012; Bobarshad *et al.*, 2012; Choi *et al.*, 2012; Faucheux *et al.*, 2012) for low-delay video coding and transmission have been comprehensively studied in recent years. However, such approaches remain insufficient for delay-sensitive video communication in certain applications, such as human-aided machine control and early warning (Murata *et al.*, 2011). One typical example of video-based human-aided machine control is docking of spacecraft with the spacestation. An example of early warning is video-based fire warning. End-to-end delay and quick-look performance are important indicators in such applications. Operators prefer being provided with rough image information instantly, to a clearer but tardy video image. Typically, end-to-end delay should be no greater than 200 ms (Hong *et al.*, 2010) and a lower delay is preferred.

Successive pictures are supposed to be taken together as a group of pictures (GOP) in H.264, as well as other conventional video coding schemes. A GOP always begins with an intra-coded frame, which is followed by inter-coded frames. Only a small number of macroblocks (MBs) in an inter-frame are selected as intra by the rate distortion optimization (RDO) algorithm, because of the temporal redundance between consecutive frames. Therefore, the number of bits consumed for encoding an inter-frame is expected to be significantly smaller than that of an intra-frame. Rate control algorithms always exploit

---

* Project partially supported by the Fundamental Research Funds for the Central Universities, and the 111 Project Program for Changjiang Scholars and Innovative Research Team in University (No. B08038), China

this property, and thus allocate more bits to an intra-frame. Thus, the delay attributed to the transmission of an intra-frame becomes the primary component.

Intra-block refresh is a well-known technique to mitigate error propagation. An adaptive intra update scheme was first introduced by Liao and Villasenor (1996) in H.263, and a significant improvement in the error-prone channel was achieved. Subsequently, the periodic intra-coding of contiguous blocks, random blocks, or the entire frame was proposed. Zhang *et al.* (2000) further improved the rate-distortion based mode selection methods and proposed a recursive optimal per-pixel estimate algorithm (ROPE). The algorithm achieves improved performance because of its accurate estimation of the overall distortion of a decoded frame at the encoder. ROPE was then extended by Zhang *et al.* (2007) to include sub-pixel prediction. Ali *et al.* (2011; 2012) proposed a cyclic-line MB refresh algorithm to mitigate the lossy channel, and another intra-refresh algorithm based on epitomic analysis was presented by Wang *et al.* (2010). However, the algorithms focused on error-resilience performance and had nothing to do with end-to-end delay.

In this paper, the overall delay of a real-time system was analyzed from the time of capture at the encoder to the time of display at the decoder. Analysis shows that the maximum delay stems from the frame with the maximum encoded bit size. A statistically uniform intra-block refresh (SUIBR) scheme was then proposed. The number of intra-coded MBs was statistically uniformly distributed in each frame, and for each MB location in a frame, the temporal refresh interval was approximately equal to the GOP size. The end-to-end delay was evidently decreased. A comparison between the proposed algorithm and the random intra-refresh algorithm shows an average of 0.1 dB gain. Additionally, the proposed algorithm outperforms the conventional coding scheme in terms of quick-look and error-resilience performance.

2 Delay analysis

In a typical hybrid video coding scheme including H.264, video is usually processed on a frame basis. Once a frame is captured, it is sent to the video encoder, and the encoder then splits it into MBs. MBs are coded either in intra-mode or in inter-mode. The bits generated after coding are later transmitted to the decoder. At the decoder side, once bits for a whole frame are received, they are parsed and decoded; the decoded frame is then sent to display. Obviously, a delay in such a configuration comes from two parts: processing and transmission. For time-sensitive applications, to reduce the aforementioned processing delay, the video encoder should work on a scan line basis; i.e., once a line of MBs are captured, these MBs are encoded using the same technique as in the frame-based scheme, and bits for the line of MBs are then transmitted. Next, we use the latter scenario and analyze the overall end-to-end delay. Suppose the video frame rate is r, the video frame size in MB is $w\times h$, the delays induced by encoding and decoding a row of MBs are respectively T_{be} and T_{bd}, and the channel transmission time and delay are respectively T_{t} and D_{c}. Then the overall delay could be formulated as

$$T_{\mathrm{o}}=\frac{1}{r}\times\frac{1}{h}+T_{\mathrm{be}}+T_{\mathrm{t}}+D_{\mathrm{c}}+T_{\mathrm{bd}}+\frac{1}{r}, \tag{1}$$

where the last term $1/r$ accounts for display and it indicates that, although encoding is carried out on a scan line basis, display is performed only when a whole frame is ready. As an example, we show typical delays when the common intermediate format (CIF) resolution video is encoded at the frame rate of 25 frames/s. In such a scenario, $1/r$ is 40 ms, T_{be} and T_{bd} are of the same order as $(1/r)\times(1/h)$ or about 2 ms in this case, and D_{c} is negligible. While T_{t} is dependent on channel bandwidth and the counter of bits generated by the encoder, and should contribute to the most significant part of the end-to-end delay.

2.1 Delay formation

When coding is carried out on a scanline basis, the processing delay can be made as small as possible (assuming that capturing, encoding, and decoding can all be performed in real-time). However, generally, a transmission channel has a limited capacity; additionally, the bit error or erasure is generally unavoidable, and this in turn limits the transmission bandwidth. This way, a transmission delay is definitely unavoidable. It is well known that in a typical hybrid video coding scheme, a video sequence is first divided into groups called GOP. The first frame in a GOP is intra-coded, while the remaining ones are inter-coded. The use of this structure is to avoid error propagation, which occurs when one of the reference frames is lost or errors occur during transmission. The GOP structure will stop error propagation from one GOP to another. In addition, intra-MB refresh can be adopted to further prevent the error propagation in pictures within a GOP. However, in a GOP, the intra-coded frame will consume a large portion of the bit budget for the whole GOP; typically, an intra-frame is about 10 times as large as any inter-frame in it. As such transmission of bits for intra-coded frame will result in extra end-to-end delay, we consider a video sequence with frame numbers $\{k|k=0, 1, 2, \cdots, k\in\mathbb{N}\}$, and predefine a number of abbreviations for convenience (Table 1).

Table 1 Predefined abbreviations for end-to-end delay

Parameter	Description
f^i	ith input frame
$p^i{}_\mathrm{c}$	Capturing of frame f^i starts
$p^i{}_\mathrm{e}$	Encoding of frame f^i starts
$p^i{}_\mathrm{t}$	Transmitting of the bits for frame f^i starts
$p^i{}_\mathrm{r}$	Receiving of frame f^i starts
$p^i{}_\mathrm{d}$	Displaying of frame f^i starts
c_i	Bit counts for frame f^i after compression
B	Bandwidth of the transmission channel
D_c	Channel delay

The formation of end-to-end delay is depicted in Fig. 1. Fig. 1 shows that the time interval from capture to the transmission of $p^i{}_\mathrm{t}-p^i{}_\mathrm{c}$ is rationally equal to $(1/r)(1/h)+T_\mathrm{be}$, and the interval from the end of decoding to the start of displaying $p^{i+1}{}_\mathrm{r}-p^i{}_\mathrm{d}$ is device-dependent. Thus, more attention should be paid to the delay from $p^{i+1}{}_\mathrm{r}$ to $p^i{}_\mathrm{t}$.

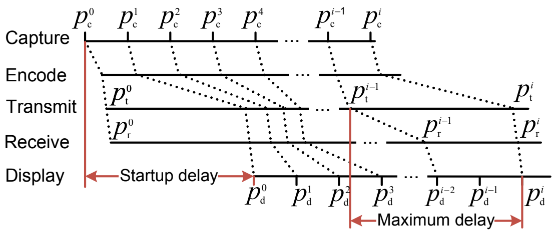


Fig. 1 End-to-end delay in the traditional coding scheme

We first focus on the startup delay. The first frame of one sequence f^0 is always encoded as an intra-frame. The first bit is transmitted at $p^0{}_\mathrm{t}$, and the last pixel is decoded at $p^1{}_\mathrm{r}$. Assuming that the channel is stable with bandwidth B and delay D_c, the time interval for the transmission of one bit is $1/B$. Thus, the transmission delay for frame f^0 is $c_0/B+D_\mathrm{c}$. Note that c_0/B is the T_t term in Eq. (1). Given that c_0/B is usually several times greater than $1/r$, the transmission of bits for f^0 will last $c_0\times r/B$ frames, and the display is postponed.

The second frame f^1 is encoded instantly after being captured. However, the stream bits are queued in the buffer and could not be transmitted until bits for f^0 have been transmitted. Thus, all the following frames are held off.

Thereafter, when the subsequent intra-frame f^i is encoded, the bit stream will be buffered for a while and then transmitted at time point $p_\mathrm{t}{}^i$.

The bit stream size is reasonably larger than those of other P frames. With a large probability, the frame will not have been totally received when it should have been displayed. Therefore, f^i will be displayed at the next tick time, resulting in a further delay. The delay will be increased at all times unless display techniques are applied. An important fact is that, if $c_i>c_0$, the maximum delay on the channel is enlarged to $c_i/B+D_\mathrm{c}$ for any i.

To conclude, the maximum end-to-end delay depends on the maximum single frame bit size $\max\{c_i\}$.

3 Statistically uniform intra-block refresh (SUIBR) algorithm

To cut off the delay caused by the transmission of continuing intra-coded MBs in an I frame, we can scatter the intra-MBs to inter-frames in the GOP with statistically uniform distribution. Suppose the GOP length is g. If the following constraints are met, we could achieve a comparable video quality and evidently decrease the end-to-end delay:

1. The compulsory intra-coded MBs in one frame are subject to 2D uniform distribution.
2. The average number of intra-MBs at each frame is equal to $w\times h/g$.
3. For co-located MBs in consecutive frames, the average intra-refresh interval is equal to g.
4. For MBs that are not forced to be intra-coded, the coding mode is determined by the mode-selection algorithm.

Conditions 1 and 2 ensure that every spatial location within a video frame is intra-refreshed with equal priority. Condition 3 ensures that, statistically, there is one intra-coded MB at a spatial location within the GOP length. Condition 4 ensures that the use of the new algorithm will not result in a dramatic performance loss.

3.1 Algorithm model

In H.264, if MB $B_{p,q}$ with location $[p, q]$ is intra-coded using intra-prediction, it will impact its direct neighbors $B_{p+1,q}$, $B_{p-1,q+1}$, $B_{p,q+1}$, and $B_{p+1,q+1}$, if they are also intra-coded. This effect will further propagate to their neighbors by a diminishing factor $B_{p,q}$, and will also affect MBs in the next frame if they are inter-coded and choose $B_{p,q}$ as their temporal prediction. Suppose the range of the motion vector is $[-v/2, v/2]$ for both the horizontal and vertical directions. The maximum direct effect range of $B_{p,q}$ in the next frame will be $[-v/2, v/2]$. The impact factor is high at the matched position and decreases with distance.

The effect of one intra-coded MB could be modeled as follows: for a given video sequence of frame resolution $w\times h$ in MB size, a refresh factor matrix $\{R[x][y] | 1\leq y\leq w, 1\leq x\leq h\}$ and a threshold T_r is defined. For any MB at position $[p, q]$, a refresh factor $r_{p,q}=R[p][q]$ is assigned. The value is reset to an initial value when the corresponding MB is intra-refreshed. This value is continuously decreased during the encoding process when MB at location $[p, q]$ is inter-coded. The refresh factor is a measurement of how imperative it is for the MB to be refreshed; a smaller value indicates a greater urgency.

When MB $B_{p,q}$ is being encoded, its initial coding mode is determined by the refresh factor. If $r_{p,q}<T_r$, the MB is to be intra-coded, and only intra prediction modes are evaluated. Otherwise, its mode is left to be determined by the mode decision algorithm considering both intra- and inter-prediction. If MB $B_{p,q}$ is decided to be intra-coded, its refresh factor and the refresh factors of the surrounding MBs will be updated.

3.2 Refresh factor initialization

The refresh factor is initialized at the startup of the encoding process. Given that the refresh process should not be sensitive to the initial value in a random access case, the initial values in $\boldsymbol{R}$ could be random integers ranging from 0 to the maximum allowed value. For simplicity, all the initial values could be set to zero, and the subsequent values could be left for the algorithm to determine.

3.3 Factor update for neighboring MBs

For factor update, a matrix $U[v][v]$ is defined, where v is the motion vector range. Values in U are subject to 2D zero mean Gaussian distribution, with σ related to the motion intensity. If MB $B_{p,q}$ is intra-refreshed, the factors centered at $B_{p,q}$ and falling within the range $[-v/2, v/2]$ in $R[x][y]$ at both axes are updated. That is, if $B_{p,q}$ is intra-coded, a squared box of refresh factors from $B_{p-v/2,q-v/2}$ to $B_{p+v/2,q+v/2}$ will be processed. For any MB $B_{p+i,q+j}$ with offset $[i, j]$, if $R[p+i][q+j]\geq U[i][j]$, $R[p+i][q+j]$ will be left unchanged; otherwise, it will be set to $U[i][j]$.

3.4 Factor decay with time

As MB $B_{p,q}$ may be chosen as a temporal prediction for MBs in the succeeding frame, it will also affect the MBs in the succeeding frame. This effect will last several frames. However, the impact factor should decrease with time.

To map this property to the refresh factor matrix, a decaying constant r_t is defined. When the encoding process moves on to the next frame, all the factors will be reduced by r_t. To keep the value meaningful, the factor is clamped to 0.

3.5 Constraints on refresh parameters

To make the encoder adopting the proposed intra-refresh algorithm work in a similar way as it does when GOP structure is used, all parameters in the proposed algorithm should be related to the GOP length g.

In accordance with the factor update rule for neighboring MBs, if $r_{p,q}<T_r$, $B_{p,q}$ will be intra-refreshed, and its neighboring refresh factors will be updated. We construct a local coordinate centered at the $B_{p,q}$ as shown in Fig. 2. If $B_{p,q}$ is intra-refreshed, d subject to $U_{d,0}<T_r$ exists, and MB $B_{p+d,q}$ will subsequently be intra-refreshed. The same condition holds for vertical axis. In this manner, one MB out of d^2 MBs will be intra-refreshed on average, and a total of $w\times h/d^2$ MBs will be intra-refreshed for each frame. To satisfy Eq. (2), d should be equal to $\sqrt{g}$.

$$U[d][0]\leq T_r, \quad \text{for } \lfloor d \rfloor=\sqrt{g}. \tag{2}$$

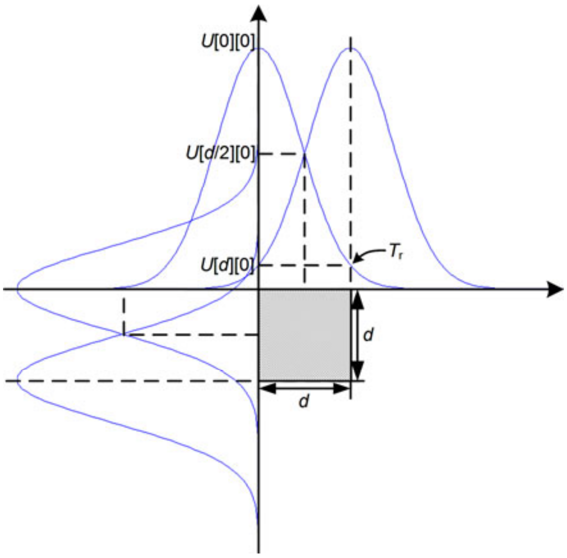


Fig. 2 Restriction on refresh parameters

When the encoding process moves on to the next frame, all factors in $\boldsymbol{U}$ are reduced by r_t, and the MB $B_{p+d/2,\,q}$ is expected to be refreshed to achieve a uniform distribution. Then, Eq. (3) should be satisfied:

$$U\left[\frac{d}{2}\right][0]-r_{\mathrm{t}} \le T_{\mathrm{r}}. \tag{3}$$

Considering Eq. (3), the refresh interval for an arbitrary MB should be equivalent to g, such that the central update factor U[0][0] satisfies

$$U[0][0]=r_{\mathrm{t}}\times g. \tag{4}$$

3.6 Encoding and decoding processes

In the proposed Algorithm 1, only the first frame is intra-coded and all the following frames are inter-coded.

However, the bits for the intra-coded frame are discarded and not transmitted. The first reconstructed frame is used for inter coding of the following frames as in normal coding. Obviously, there is a mismatch between the encoder and decoder. This mismatch is mitigated by the proposed intra-refresh algorithm, which also reduces the end-to-end delay. For MBs in an inter-frame, their coding modes are initially determined by the refresh factor. If the MB is not forced intra, the mode selection algorithm will take effect.

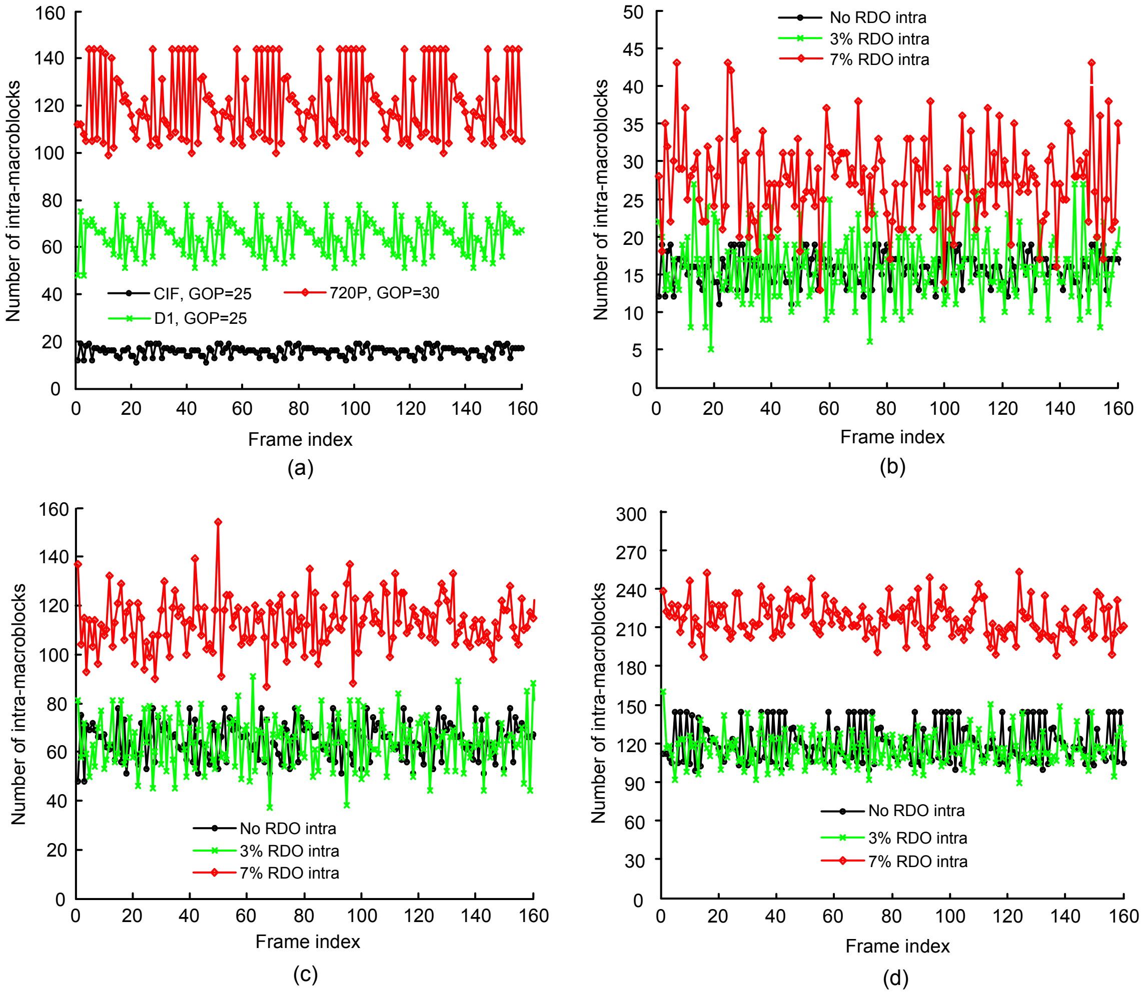
Algorithm 1 Statistically uniform intra-block refresh (SUIBR) algorithm

```
Initialize $\boldsymbol{R}$
Encode the 1st frame as I frame, and discard the bits.
For subsequent frames:
loop
  for $q$=1 to $h$ and $p$=1 to $w$ do
    $R[p][q]=R[p][q]-r_t$
    if $R[p][q]\le T_r$, then
      force to encode $B[p][q]$ with intra-mode
      for $i=-v/2$ to $v/2$ and $j=-v/2$ to $v/2$ do
        update refresh factor $R[p+i][q+j]$
      end for
    else
      encode $B[p][q]$ using the H.264 rate distortion
        optimization (RDO) algorithm
    end if
  end for
end loop
```

4 Simulation results and analysis

The proposed algorithm was integrated in H.264 reference software JM14.2, and also integrated into a Texas Instruments (TI) DSP based video communication system. The performance of the proposed algorithm was evaluated. The experiments included: intra-refresh performance, end-to-end delay, rate-distortion performance, quick-look, and error-resilience. Experiments were carried out either using test video sequences or using real scene video with different resolutions.

We first evaluate the intra-refresh performance of the proposed algorithm. This is carried out to see if the proposed algorithm can achieve a truly uniform distribution of bits for each frame, and whether the number of intra-MBs in each frame is approximately equal. As the RDO-based mode decision in H.264 will also choose an MB as intra, it is meaningful to evaluate the impact of the RDO-based mode decision on the performance of the proposed algorithm. For this purpose, three video sequences with resolutions of CIF (352×288), D1 (720×576), and 720P (1280 ×720) are used, and two sets of tests are carried out. First, RDO-based mode decision is forced to choose only from inter-modes. The results are shown in Fig. 3a. It can be seen that the allocated intra-MBs number for each frame is approximately equal, with slight fluctuations attributed to the frame boundary effect.



(a) (b) (c) (d)

Fig. 3 Numbers of intra-MBs per frame with and without rate distortion optimization (RDO) at different resolutions: (a) without RDO; (b) CIF with RDO; (c) D1 with RDO; (d) 720P with RDO

Next, random intra-MBs are inserted with a definite probability to imitate the effect of RDO. The results for the three test video sequences with probabilities of 3% and 7% are shown in Figs. 3b–3d and Table 2. The results showed that although the number of intra-MBs in each frame exhibits a relatively large fluctuation, the curve generally remains stable. Moreover, the mean value approaches the target value (Table 2). Here the target refresh interval should be equal to the assumed GOP size or g, while the target number of MBs refreshed per frame should be equal to $w\times h/g$; i.e., each location will be refreshed once in GOP size g. In Table 2, the frame statistics indicates the average number of intra-MBs per frame (avg.) as well as the standard deviation (std.) for all MBs.

Table 2 Statistical data on intra-refreshed MBs per frame and refresh interval for each MB

Frame resolution	Sequence	Frame stat.[a]		MB stat.[b]	
		Avg.	Std.	Max.	Avg.
CIF	Clarie	17.38	6.34	61	20.92
	Container	16.35	3.67	46	22.96
	Paris	17.01	2.64	36	21.97
	Student	16.33	3.23	57	23.14
D1	Children	66.15	14.56	33	23.03
	Household	67.83	13.41	31	23.51
720P	Bigships	129.17	20.01	51	26.74
	Shuttlestart	134.37	17.25	77	25.35

[a] Number of intra-refreshed MBs per frame (Avg.) and the standard deviation (Std.); [b] Maximum (Max.) and average (Avg.) refresh interval for each MB

4.1 Transmission delay

To evaluate the end-to-end delay of the SUIBR algorithm, we use a TI DSP based real-time video communication system. The system included a complementary metal-oxide-semiconductor (CMOS) camera, an H.264 encoder board, and an H.264 decoder board. An Ethernet was established between the two boards. In addition, both sides contain display devices. Again, we used three different resolution videos for this test, but the videos were from real scene with time in ms unit overlaid on the top-left corner of each frame. The resolutions are the same as before, namely CIF, D1, and 720P. The bit rates for them are 0.5, 2.0, and 4.0 Mb/s, respectively. Accordingly, the logical transmission channel is also made of the same width by consuming the remaining bandwidth with dummy data. The delay was measured by capturing the video displayed at the encoder and decoder sides with a digital camera, and calculating the difference of time displayed on the two videos.

The proposed algorithm and a reference one are used in this test. The reference method is with a default encoding configuration in JM14.2 (JM Online, 2010), and for fair comparison, it was also improved to work on a scan line basis, but without intra-refresh.

Table 3 shows the test results for the proposed algorithm and the reference method. The results are obtained by averaging results over 30 runs. It can be seen that the end-to-end delay of the proposed algorithm is dramatically decreased.

Table 3 End-to-end delay test results

Method	End-to-end delay (ms)		
	CIF	D1	720P
Statistically uniform intra-block refresh (SUIBR)	77	79	82
Reference	329	326	326

4.2 Rate-distortion comparison

The rate-distortion (RD) performance of the proposed algorithm is compared with that of the random intra-refresh (RndIBR) algorithm applied in H.264 JM14.2 (JM Online, 2010), in which each MB is assigned a random number. At the mode decision stage of the encoding process, MBs with a specific assigned number are forced to be intra-refreshed. The results of the two algorithms are shown in Fig. 4. Compared to RndIBR, our algorithm could achieve 0.1 dB gain on PSNR.

4.3 Quick-look performance

To test the quick-look performance, the experiment is set as follows: test video sequence Paris with CIF resolution is separately encoded using H.264 JM and the SUIBR algorithm. The bits for the I frame are then discarded, whereas the remaining bits are decoded. The fifth decoded frame could facilitate an impartial comparison of the two schemes.

Test results show that we could only obtain a vague and unrecognizable image using H.264 JM

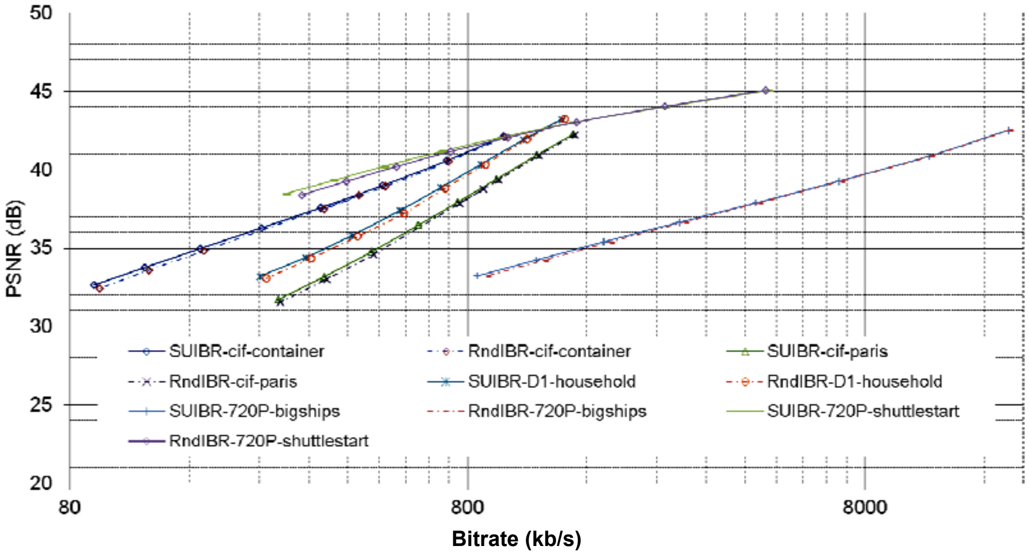


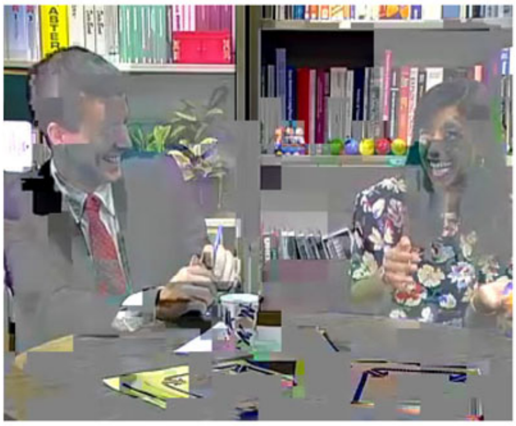
Fig. 4 PSNR comparison of the RndIBR scheme and the proposed SUIBR scheme

(Fig. 5a). On the other hand, the proposed algorithm could generate an image with uniformly distributed intra-MBs (Fig. 5b), which clearly depicts the whole scene.

4.4 Error-resilience performance

Given that errors are unavoidable in the transmission channel, the spatial-temporal error propagation in hybrid video coding has to be limited. By employing the SUIBR algorithm, scattered intra-block breaks the correlations of stream bits and refreshes the regional information to achieve improved visual performance.

(a)



(b)

Fig. 5 Quick-look snapshot at 20-ms delay for the fifth decoded frames using the traditional coding scheme (a) and the proposed scheme (b)

The experimental results for error-resilience performance are shown in Fig. 6. In this experiment, video sequences are encoded using H.264 JM and the SUIBR algorithm. A random bit error is then added. The noisy stream is decoded, and a random frame is selected to demonstrate the effect.

For a binary symmetric channel, when the bit error rate (BER) is below 10^{-6}, both schemes generate a similar decoded frame with favorable presence. When BER increases to 10^{-5}, Fig. 6 shows that the object encoded by H.264 JM becomes difficult to identify, but that encoded by our algorithm is still recognizable.

(a)

(b)

Fig. 6 Comparison of error-resilience performance at BER=10^{-5}
(a) Traditional coding scheme; (b) Proposed scheme

5 Conclusions

The end-to-end delay, quick-look, and error-resilience are three of the most important indicators of real-time video communication systems for human-aided control or early warning. Conventional video coding schemes like H.264 could achieve competitive compression performance which is suitable for video recording and other similar applications that are less sensitive to delay. However, conventional schemes fall short of the aforementioned indicators.

The proposed coding scheme scatters the intra-blocks uniformly across all inter-coded frames. Therefore, the bits consumed by each frame are almost equal. The end-to-end delay is significantly decreased, and the delay for each frame become smooth and steady. Simulation results show that the proposed algorithm outperforms the RndIBR at an average of approximately 0.1 dB. Moreover, the quick-look and error-resilience performances are improved compared with those of the conventional coding scheme.

References

Ali, I., Moiron, S., Ghanbari, M., Fleury, M., 2011. Prioritized Packetization for Video with Intra-Refresh Macroblock Line. 12th IEEE Int. Conf. on Multimedia and Expo, p.1-6. [doi:10.1109/icme.2011.6011978]

Ali, I., Fleury, M., Ghanbari, M., 2012. Content-Aware Intra-Refresh for Video Streaming over Lossy Links. IEEE Int. Conf. on Consumer Electronics, p.118-119. [doi:10.1109/icce.2012.6161768]

Asioli, S., Ramzan, N., Izquierdo, E., 2012. A game theoretic approach to minimum-delay scalable video transmission over P2P. *Signal Process. Image Commun.*, **27**(5): 513-521. [doi:10.1016/j.image.2012.02.012]

Bobarshad, H., van der Schaar, M., Aghvami, A.H., Dilmaghani, R.S., Shikh-Bahaei, M.R., 2012. Analytical modeling for delay-sensitive video over WLAN. *IEEE Trans. Multim.*, **14**(2):401-414. [doi:10.1109/TMM.2011.2173477]

Chang, C.Y., Chou, C.F., Chan, D.Y., Lin, T.N., Chen, M.H., 2012. Accurate bitrate model and greedy-based rate controller for low delay video transmission. *IEEE Syst. J.*, **6**(3):414-425. [doi:10.1109/JSYST.2011.2163986]

Choi, K., Choi, S.G., Choi, J.K., 2012. Hybrid video transmission scheme for minimizing server bandwidth consumption with zero start-up delay in video-on-demand service. *IEEE Commun. Lett.*, **16**(1):6-8. [doi:10.1109/LCOMM.2011.110711.111959]

Faucheux, F., Morel, M.L.A., Kerboeuf, S., Roullet, L., 2012. Variable FEC decoding delay and playout slowdown method for low start delay and fast channel change for video streaming in DVB-SH mobile broadcast systems. *Bell Labs Techn. J.*, **16**(4):63-83. [doi:10.1002/bltj.20534]

Hong, D., Horowitz, M., Eleftheriadis, A., Wiegand, T., 2010. H.264 Hierarchical P Coding in the Context of Ultra-Low Delay, Low Complexity Applications. 28th Picture Coding Symp., p.146-149. [doi:10.1109/pcs.2010.5702445]

JM Online, 2010. Citing Electronic Sources of Information. Fraunhofer Heinrich Hertz Institute. Available from http://iphome.hhi.de/suehring/tml/download/old_jm/jm14.2.zip [Accessed on June 12, 2010].

Lee, H., Sull, S., 2012. A VBR video encoding for locally consistent picture quality with small buffering delay under limited bandwidth. *IEEE Trans. Broadc.*, **58**(1):47-56. [doi:10.1109/TBC.2011.2164308]

Liao, J.Y., Villasenor, J.D., 1996. Adaptive Intra Update for Video Coding over Noisy Channels. Proc. 3rd IEEE Int. Conf. on Image Processing, p.763-766. [doi:10.1109/ICIP.1996.560837]

Liu, Y., Li, Z.G.G., Soh, Y.C., 2007. A novel rate control scheme for low delay video communication of H.264/AVC standard. *IEEE Trans. Circ. Syst. Video Technol.*, **17**(1):68-78. [doi:10.1109/tcsvt.2006.887081]

Murata, K., Hattori, M., Shibuya, Y., 2011. Visual feedback to reduce influence of delay on video chatting. *Human Interf. Manag. Inf. Interact. Inf.*, **6772**:157-164. [doi:10.1007/978-3-642-21669-5_19]

Song, W., 2011. Delay Analysis for Compressed Video Traffic over Two-Hop Wireless Moving Networks. 54th Annual IEEE Global Telecommunication Conf. on Energizing Global Communications, p.5-10. [doi:10.1109/GLOCOM.2011.6133700]

Wang, Q., Hu, R., Wang, Z., Hang, B., 2010. Intra Coding and Refresh Based on Video Epitomic Analysis. IEEE Int. Conf. Multimedia and Expo, p.452-455. [doi:10.1109/icme.2010.5583074]

Yu, L., Wang, J.P., 2010. Review of the current and future technologies for video compression. *J. Zhejiang Univ.-Sci. C. (Comput. & Electron.)*, **11**(1):1-13. [doi:10.1631/jzus.C0910684]

Zhang, R., Regunathan, S.L., Rose, K., 2000. Video coding with optimal inter/intra-mode switching for packet loss resilience. *IEEE J. Sel. Areas Commun.*, **18**(6):966-976. [doi:10.1109/49.848250]

Zhang, Y., Gao, W., Lu, Y., Huang, Q.M., Zhao, D.B., 2007. Joint source-channel rate-distortion optimization for H.264 video coding over error-prone networks. *IEEE Trans. Multimedia*, **9**(3):445-454. [doi:10.1109/TMM.2006.887989]