A FLUID MODEL FOR ERROR PROPAGATION CHARACTERIZATION IN VIDEO CODING

Xiaoming Sun and C.-C. Jay Kuo

Integrated Media Systems Center and Department of Electrical Engineering
University of Southern California, Los Angeles, CA 90089-2564
sunx@pollux.usc.edu, cckuo@sipi.usc.edu

ABSTRACT

An error corruption model (ECM) to describe the interframe error propagation phenomenon in a motion-compensated predictive video codec using fluid flow characteristics is proposed in this research. First, we derive a diffusion differential equation and discuss its solution, which captures the damping effect of error propagation. Then, we propose a tracking quadrilateral (TQ) mechanism to capture the shaping and drifting effects of error propagation. Finally, we integrate these building blocks to form an adaptive fluid-based ECM (F-ECM). The accuracy of the proposed F-ECM is verified by experimental results.

1. INTRODUCTION

The motion-compensated predictive coding principle has been widely used in modern video codecs to achieve high compression performance [1]. However, the predictive framework establishes strong dependencies between adjacent frames so that transmission errors in one frame may result in severe propagating errors in later frames. Some error corruption models (ECMs) have been studied before. They can be used to design an effective error-resilient video codec in the presence of hostile channels. Generally an error in a local region propagates along both spatial and temporal domains. It is observed in experiments that, similar to a flowing fluid, the interframe error propagation can be characterized by three effects; namely, damping, shaping and drifting. By the damping effect, we mean that the error variance at the center of the corruption area is monotonically decreasing. The shaping effect means that the boundary of the affected erroneous region can be deformed from a normal circular shape. Finally, the center of the erroneous region may drift along time.

Previous work on ECM is reviewed below. Färber *et al.* [2] performed pioneering research on characterizing interframe error propagation under two simplifying assumptions. First, the overall effect of spatial filtering at the decoder can be modeled by a separable average loop filter. Second, the impulse response of the loop filter is assumed to be of the Gaussian form using the Central Limit Theorem (CLT). A formula was derived in [2] to estimate the error variance of an erroneous macroblock (MB) along the time, which we call the FSG model. However, the propagation of errors along the spatial domain in later frames was not discussed in the FSG model. Later, Kim *et al.* [3] generalized the FSG model so as to estimate error propagation along both temporal and spatial domains using parallel and cascade trajectories. This generalized ECM can be used for short-term prediction of propagating errors. However, it primarily offers a mechanical tool for error resilience, but does not provide a global view of the error propagation behavior. The work in [2] focused on the understanding of the damping effect. Even though shaping and drifting effects were addressed in [3], the treatment was somehow implicit.

In this work, we provide an explicit model to characterize the damping, shaping and drifting effects. First, we derive a diffusion equation and discuss its solution, which captures the damping effect in Sec. 2. Then, we propose a tracking quadrilateral (TQ) algorithm to capture the shaping and drifting effects in Sec. 3. Finally, we integrate these building blocks to form an adaptive fluid-based ECM (F-ECM), and verify the accuracy of the proposed F-ECM by experimental results in Sec. 4.

2. F-ECM FOR DAMPING

The overall distortion D in compressed video transmission arises from two types of errors [2]; namely, the encoder error D_e and the transmission error D_v. In this work, we mainly focus on estimating D_v. Let $v[x,y,t]$ be the error signal and $u[x,y] = v[x,y,0]$ be the initial error. The variance $\sigma^2[x,y,t]$ of $v[x,y,t]$ can be used as the distortion measure. Intuitively, the propagation of errors is influenced by the underlying motion pattern of coded video. Thus, we relate motion prediction schemes to ECM below.

This research was funded by the Integrated Media Systems Center, a National Science Foundation Engineering Research Center, Cooperative Agreement No. EEC-9529152.

0-7803-8554-3/04/$20.00 ©2004 IEEE.

Sub-pixel (*e.g.* half-pixel or quarter-pixel) motion estimation (ME) and motion compensation (MC) schemes have been used in latest video coding standards such as MPEG-4 and H.264. The overlapped block motion compensation (OBMC) is an extension of the classic MC method. Our discussion is based on OBMC, which can however be extended to other MC methods. Although ME is done at the MB or block level for integer pixel or sub-pixel precision, MC is always applied to each block, where either a block has its own motion vector (MV) or four blocks share the same MV in the MB.

Let $Z(x,y,t)$ denote the decoded pixel value at position (x,y) and frame t. Let $MV_i(x,y)$, $i=1,2,3,4,5$, denote the MV of the current block that covers pixel (x,y), and its upper, lower, left and right neighboring overlapped blocks, respectively. The OBMC predictive coding in the temporal domain can be described as

$$\begin{aligned} Z(x,y,t) = \sum_{i=1}^{5} Z_i(x+MV_i(x,y)[x], \\ y+MV_i(x,y)[y], t-1)H_i + \Delta(x,y,t), \end{aligned} \tag{1}$$

where H_i is the normalized weighting matrices satisfying $\sum_{\text{All pixels}} H_i = 1$ and $\Delta(x,y,t)$ is the residual data. The pixel difference between the original error-free block and the reconstructed error-concealed block is computed via

$$\begin{aligned} \delta Z(x,y,t) = \sum_{i=1}^{5} \delta Z_i(x+MV_i(x,y)[x], \\ y+MV_i(x,y)[y], t-1)H_i, \end{aligned} \tag{2}$$

where δ denotes the difference operator. The residual $\Delta(x,y,t)$ is filtered by difference operator δ. It is worthwhile to point out that only the MC operation carries errors in previous reconstructed frames to the current frame to be decoded. Let pixel (x,y) be located in $Blk(x,y)$ to be decoded and the notation VAR or σ^2 be the variance value of $Blk(x,y)$, where $Blk(x,y)$ represents the block covering the pixel (x,y). With Eq. (2), the variance of $Blk(x,y)$ can be calculated as

$$\begin{aligned} VAR(\delta Z(x,y,t)) = \sum_{i=1}^{5}(1-W_i(x,y,t)) \times \\ \sigma^2(x,y,t-1)H_i + \sum_{i=1}^{5} W_i(x,y,t)\sigma_i^2(x,y,t-1)H_i, \end{aligned} \tag{3}$$

where W_i is the weighting coefficient that characterizes the contribution of Blk_i of frame $t-1$ to $Blk(x,y)$ of frame t due to OBMC. Generally, we define *dependency weight*(DW) $W_{(i,j)}$ to be the normalized number of pixels in $Blk_i(MB_i)$ of the previous reconstructed frame used to predict $Blk_j(MB_j)$ of the current frame. The differentiation operation can be approximated by taking the difference of variance σ^2 with the time increment $dt=1$ for two consecutive frames and with the spatial displacement $dx=1$ for two consecutive blocks, respectively. Then, based on Eq. (3), we can obtain

$$\begin{aligned} &\frac{\partial}{\partial t}\sigma^2(x,y,t) \approx \sigma^2(x,y,t) - \sigma^2(x,y,t-1) \\ &= \sum_{i=1}^{5} W_i(x,y,t)\left(\sigma_i^2(x,y,t-1) - \sigma^2(x,y,t-1)\right)H_i. \end{aligned} \tag{4}$$

Note that $\sigma_i^2(x,y,t-1)$ is the amount of variance σ^2 propagating to $Blk_i(x,y)$, which is actually $\frac{\partial\sigma^2}{\partial x_i}$ at time $t-1$. Hence,

$$\sigma_i^2(x,y,t-1) - \sigma^2(x,y,t-1) \approx \frac{\partial}{\partial x_i}\frac{\partial\sigma^2}{\partial x_i} = \frac{\partial^2\sigma^2}{\partial x_i^2}$$

for $i=1,2,3,4,5$. The overall effect can be approximated by averaging the five derivatives. Let

$$\begin{aligned} \overline{\sigma_r^2(x,y,t-1)} &= \frac{\sum \sigma_i^2(x,y,t-1)}{5}, \\ \kappa &= \sum_{i=1}^{5} W_i(x,y,t)H_i, \end{aligned} \tag{5}$$

$$\sigma_i^2(x,y,t-1) \approx \overline{\sigma_r^2(x,y,t-1)}. \tag{6}$$

Plugging (5) and (6) in (4) yields the diffusion partial differential equation (PDE)

$$\begin{aligned} &\frac{\partial}{\partial t}\sigma^2(x,y,t) \approx \\ &\sum_{i=1}^{5} W_i(x,y,t)H_i\left(\overline{\sigma_r^2(x,y,t-1)} - \sigma^2(x,y,t-1)\right) \\ &\approx \kappa\frac{\partial^2}{\partial r^2}\sigma^2(x,y,t), \end{aligned} \tag{7}$$

where $\vec{\mathbf{r}}$ satisfying $r^2 = x^2 + y^2$ is the error energy propagation direction determined by the overall vector average of the five prediction blocks. In essence, the above equation models the *isotropic* fluid flow. If the initial wave shape at time $t=0$ is assumed to be Gaussian

$$\sigma^2(\vec{\mathbf{r}},t)|_{t=0} = \sigma_u^2 e^{-r^2/a^2},$$

where a denotes the standard deviation, Eq. (7) has the following solution

$$\sigma^2(\vec{\mathbf{r}},t) = \frac{\sigma_u^2}{\sqrt{1+\frac{4\kappa t}{a^2}}} e^{-\frac{r^2}{a^2+4\kappa t}}, \tag{8}$$

where κ characterizes the damping effect, which is related to the averaging of MVs in Eq. (5). The variance at the center of the erroneous region characterizes the damping effect. Hence, the variance $\frac{\sigma_u^2}{\sqrt{1+\frac{4\kappa t}{a^2}}}$ at time t with $r=0$ can be used to define the cost function to determine the optimal κ. We propose an algorithm to estimate κ below.

Algorithm 1 Determine κ

Let X_i, $i = 0, \ldots, I$, be a sequence of variances of MBs or blocks sampled at time $t = 0, \ldots, I$ at the center of the erroneous region.

Find parameters $\hat{\xi}$ and $\hat{\zeta}$ to minimize a cost function $G(\xi,\zeta) = \sum_{i=0}^{I}\left(\frac{\xi}{\sqrt{1+\zeta i}} - X_i\right)^2$. Then, we can get $\hat{\zeta} = 4\hat{\kappa}/a^2$.

3. F-ECM FOR SHAPING AND DRIFTING

Eq. (8) provides a closed-form description of interframe error propagation with short-range motion. In the isotropic fluid flow field, the corruption area forms a sequence of concentric circles centered at the MB where the error is initially introduced. However, Eq. (8) only offers an explanation for the damping effect, but not the shaping and drifting effects. Actually, it has only a local view of error propagation using the PDE. The tool to describe the long-range motion is however lacking in the above model. Essentially, the shaping and drifting effects arise from the time-varying MV field. To capture the long-range motion which results in the shaping and drifting effects at the block level, we propose a tracking quadrilateral (TQ) method, as given in Algo.2.

Let $TQ(P_1,P_2,P_3,P_4)$ be a quadrilateral defined by four corner points P_i $(i = 1,2,3,4)$ such that the corruption area by a single error in a frame is enclosed by TQ. The centroid C of TQ is computed as the center of the line segment connecting P_1P_3 and P_2P_4.

Algorithm 2 Tracking Quadrilateral

Let Q_i be the corner point of the initially erroneous MB and P_i' be the updated point of P_i, $i = 1,2,3,4$.

1. At time $t = 0$, $TQ(P_i)$ is initialized as $P_i = Q_i$.

2. If P_i is the left-top corner of $Blk_i(x,y)$ at time t, then $P_i' = P_i - MV_i$ at time $t+1$, where MV_i is for $Blk_i(x,y)$.

2'. If P_i falls into the interior of $Blk_i(x,y)$ at time t, then $P_i' = P_i - \overline{MV_i}$ at time $t+1$, where $\overline{MV_i} = (MV_C + MV_R + MV_B + MV_{RB})/4$, and MV_C, MV_R, MV_B and MV_{RB} are the MVs of the current $Blk_i(x,y)$, and its right, bottom, and right-bottom neighboring block, respectively.

3. Repeat Step 2 or 2' until the desired t is reached.

An illustration of the TQ algorithm is shown in Fig. 1 (left). It is observed in our experiments that the TQ algorithm can accurately track the shaping and drifting effects of a single-error corruption region in the long term. In Fig. 1 (right), we show the tracked error region in the 10^{th} frame, where a single error is initially introduced to $MB[4,5]$ in the first frame. The MV field is also highlighted.

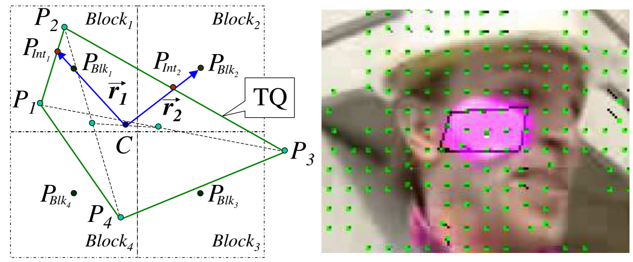

Fig. 1. Illustration of the TQ algorithm (left) and an example (the 10^{th} frame) tracked using the TQ algorithm (right).

Algorithm 3 Determine λ_r

Given $TQ(P_1,P_2,P_3,P_4)$ that has the centroid C and the center of block Blk is denoted by P_{Blk}. Let $\vec{\mathbf{r}} = \overrightarrow{CP}_{Blk}$.

1. Determine the two nearest points P_i and P_{i+1} for P_{Blk}.

2. Compute the interception point P_{Int} of the two lines P_iP_{i+1} and CP_{Blk}.

3. Determine if Blk belongs to the interior of TQ. If $|CP_{Int}| < |CP_{Blk}|$, then Blk is outside. Otherwise, Blk is inside.

4. If Blk belongs to the interior of the TQ, then λ_r is computed by $\lambda_r = \frac{\theta\sqrt{a^2+4\kappa t}}{|CP_{Int}|}$, where θ is a parameter to measure the variance σ^2 at P_{Int}. In practice, θ is set to 2 or 3 to ensure that $\sigma^2(P_{Int})$ is small enough.

4. ADAPTIVE F-ECM FOR ERROR TRACKING

It is worthwhile to point out that there exists interaction between the damping, shaping and drifting effects. In the isotropic F-ECM, the damping speeds are homogeneous along different propagation directions. With the shaping and drifting effects, we can capture the *directional propagation* phenomenon, *i.e.*, a different propagation speed along a directional vector $\vec{\mathbf{r}}$ radiating from the error center C. The actual damping of the propagating error along each direction is related to the non-isotropic fluid flow speed in the local area. For the non-isotropic fluid flow, we can modify the solution in Eq. (8) by introducing a scaling factor λ_r. That is,

$$\sigma_v^2(\vec{\mathbf{r}},t) = \frac{\sigma_u^2}{\sqrt{1+\frac{4\kappa t}{a^2}}} e^{-\frac{\lambda_r^2 r^2}{a^2+4\kappa t}}, \qquad (9)$$

where $r^2 = x^2 + y^2$. The overall damping effects are determined by both the isotropic damping (κ) and the directional damping (λ_r). Algo. 3 can be used to estimate λ_r for some given TQ tracking result. Based on Algos. 1, 2 and 3 as building blocks, we can obtain an adaptive F-ECM to track the interframe error propagation phenomenon in Algo. 4.

The proposed adaptive F-ECM is a generalization of the FSG model [2]. By ignoring the difference of the square-root operation, the F-ECM is reduced to the FSG model examined at the error origin with $r = 0$. However, the proposed F-ECM can also be used to track the error propagation along both spatial and temporal domains, which is not covered in the FSG model.

Algorithm 4 Adaptive F-ECM

1. Apply Algo. 1 to obtain the optimal $\hat{\kappa}$ and $\hat{\xi}$.

2. Apply Algo. 2 to compute the TQ.

3. For any given $Blk(x,y)$, whose left-top corner is (x,y), apply Algo. 3 to decide if the block is inside the TQ. If the block is outside, then its error variance is set to 0. Otherwise, two nearest control points P_i and P_{i+1} of $TQ(P_1,P_2,P_3,P_4)$ are determined.

4. Use P_i, P_{i+1} and Algo. 3 to compute $\hat{\lambda}_r$.

5. Plug $\hat{\kappa}$ and $\hat{\lambda}_r$ in Eq. (9) to estimate the variance of the propagating error for given x, y and t, where $\sigma_u^2 \approx \hat{\xi}$.

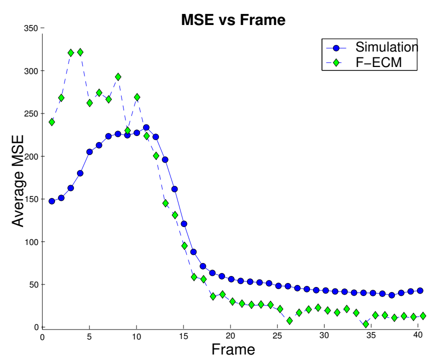

Fig. 2. The error tracking performance of adaptive F-ECM.

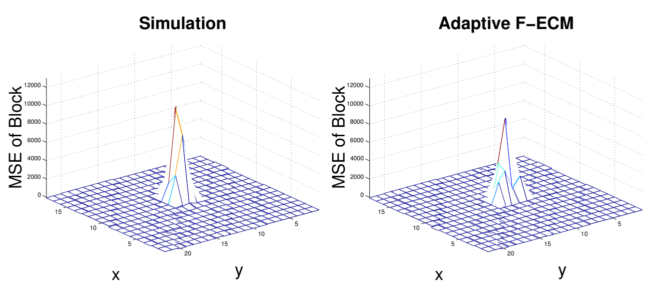

Fig. 3. Comparison of MSE in the 15^{th} frame.

Generally, the mean square error (MSE) is adopted to measure the error variance at the block/MB level. In our experiments, the testing sequence is "Foreman" in format of YUV 4:2:0, 176×144 pixels/frame (QCIF), which is encoded into the MPEG-4 format at a target bit rate of 48 Kbps and a frame rate of 6 frame/sec in the Video Packet (VP) coding mode. The initial error is introduced to MB[4,5].

The comparison of the simulated average MSE and the predicted value using the adaptive F-ECM in the temporal domain is shown in Fig. 2. The value of σ^2 is considerably large near the singularity point $t = -\frac{a^2}{4\kappa}$ of Eq. (8), which causes the fitting discrepancy for the first few frames as shown in Fig. 2. The actual MSE of block by the simulation (left) and the predicted MSE by the adaptive F-ECM (right) for the 15^{th} frame are compared in Fig. 3. Our experiments show that the proposed adaptive F-ECM provides good estimation of interframe error propagation along both spatial and temporal domains.

5. CONCLUSION

An ECM to describe the interframe error propagation phenomenon, called the F-ECM, was proposed in this work. For error-resilient video coding, we can perform the error propagation analysis using the proposed F-ECM for individual block/MB to determine its importance for unequal error protection. In [4, 5], we proposed a dependency weight (DW) approach to enhance the error-resilient video encoding. This DW approach is closely related to the F-ECM introduced in this work. That is, we see from Eq. (5) that κ, which characterizes the damping speed, is approximated by a linear combination of DWs from the previous reconstructed frame. Thus, DWs can be used to measure the importance of blocks/MBs in controlling interframe error propagation.

6. REFERENCES

[1] Y. Wang, S. Wenger, J. Wen, and A. K. Katsaggelos, "Error resilient video coding techniques," *IEEE Signal Processing Magazine*, vol. 17, pp. 61–82, 2000.

[2] N. Färber, K. Stuhlmüller, and B. Girod, "Analysis of error propagation in hybrid video coding with applicaiton to error resilience," *Proc. of IEEE Int. Conf., Imaging Processing (ICIP'99)*, vol. 2, pp. 550–554, Oct. 1999.

[3] J.-G. Kim, J. Kim, and C.-C. J. Kuo, "Corruption model of loss propagation for relative prioritized packet video," *Proc. of SPIE, Applications of Digtial Image Processing XXIII*, vol. 4115, pp. 214–224, 2000.

[4] X. Sun and C.-C. J. Kuo, "Performance evaluation of MPML-based error resilient video transmission," *Proc. of SPIE, 17th Annual Int. Symp. on AeroSense*, vol. 5108, pp. 263–274, Apr. 2003.

[5] X. Sun and C.-C. J. Kuo, "MPML-based error resilient wireless video transmission," *Proc. of SPIE, ITCOM*, vol. 5241, pp. 111–122, Sep. 2003.