Tracking-Optimal Error Control Schemes for H.264 Compressed Video For Vehicle Surveillance

Z. Chen, E. Soyak, S. A. Tsaftaris, A. K. Katsaggelos

Image & Video Processing Lab
Northwestern University
Aug. 30, 2012
Outline

• Background and motivation
• Proposed system
  • Performance metrics
  • Preprocessing
    • Temporal deviation thresholding (TDT) filtering
    • Kurtosis-based ROI extraction
  • Forward error correction
  • Tracking-aware error concealment
• Numerical Examples
• Conclusions
Background & Motivation

- Remote sensors are commonly deployed for transportation monitoring and surveillance.
- Captured videos often need to be sent to central offices for processing
  - Wireline transmission: inflexible, high cost, etc.
  - Wireless transmission: limited bandwidth, lossy
- The applications specifically require:
  - Video compression with special characteristics of transportation video taken into account
  - Error control techniques to mitigate channel degradations.
System Overview

• Block diagram

- Preprocessing identifies tracking-critical video content and generates region-of-interest (ROI) map.
- Forward error control (FEC) protects video packets against channel losses
- Error concealment recovers lost video information for decoding
Performance Metrics

• How to measure tracking quality?
  • Closeness of detected trajectories
  • Total targets detected
  • Missed targets and falsely identified targets

• Comparing the ground truth (GT, obtained from raw video) and algorithm result (AR, obtained from preprocessed/received/decoded video)
  • Overlap, Precision and Sensitivity
  • Overall accuracy defined as a convex combination

\[ A = (\alpha \times \text{Overlap}) + (\beta \times \text{Precision}) + (\gamma \times \text{Sensitivity}) \]
Preprocessing: TDT filtering

• Goals:
  • remove temporal noise from video sequence to reduce bitrate
  • Retain video content most important to tracking

• Temporal deviation thresholding (TDT) algorithm
  • Iteratively generates frame masks and applies masks to input frames

\[
\hat{\sigma}_t = \text{mode}(\text{std}([F_{t-T+1}, \ldots, F_t]))
\]

\[
\Delta_t = |F_t - F_{t-1}|
\]

\[
M_t = \Delta_t > \tau \hat{\sigma}_t
\]

\[
\tilde{F}_t = M_t \times F_t + (1 - M_t) \times \tilde{F}_{t-1}
\]
**Preprocessing: ROI Extraction**

- **Goal:** generate ROI map for error protection
- **Kurtosis-based ROI extraction**
  - Per-pixel kurtosis is defined as ($T=$training duration):
    \[
    \kappa(n_1, n_2) = \frac{1}{T} \sum_{t=0}^{T-1} \left( f_t(n_1, n_2) - \bar{f}(n_1, n_2) \right)^4
    \left( \frac{1}{T} \sum_{t=0}^{T-1} \left( f_t(n_1, n_2) - \bar{f}(n_1, n_2) \right)^2 \right)^2 - 3,
    \]
- **Classification:**
  - Capture noise (Gaussian) -> kurtosis = 0
  - Periodic motion, such as trees and water reflection (Mixture of Gaussians) -> kurtosis = 0 (additivity)
  - Motion of tracking interest, such as vehicles and pedestrians (Poisson) -> kurtosis = 6
Preprocessing: ROI Extraction

- Kurtosis-based ROI extraction
  - Thresholding per-pixel kurtosis value at the mid-point between:
    - 0 (uninteresting intensity variation due to noise and periodic motion)
    - 6 (interesting intensity variation due to object motion)
  - Macroblocks are classified by “majority vote rule”, i.e.
    - Foreground if 128 or more pixels have kurtosis >= 3
    - Background if less than 128 pixels have kurtosis >= 3
Forward Error Correction (FEC)

- FEC is appropriate for transportation video transmission due to the “live” nature of continuously captured video sequence
- Limited computing power at remote sensors necessitates computationally efficient approach
- Proposed solution:
  - Redundant slice – based approach
  - Unequal error protection (UEP) by assigning higher protection levels to more important packets (slices)
  - Linearly increases bandwidth consumption
  - Exponentially decreases packet (slice) loss probability (for channels characterized by IID loss patterns)
Forward Error correction (FEC)

- Unequal error protection (UEP)
- ROI map provides FG/BG separation information
- Flexible macroblock ordering (FMO) feature of H.264 allows arbitrary mapping from MBs to slice groups
- Slices in FG slice group $\Rightarrow$ assign protection level $H$
- Slices in BG slice group $\Rightarrow$ assign protection level $L$

$$c_{\text{target}} = \left\lfloor \log_{P_{\text{unprot}}} (P_{\text{target}}) \right\rfloor;$$

**if** $c_{\text{target}} \sum_{i \in I_H} R_i + \sum_{i \in I_L} R_i > \hat{R}$ **then**

$$H = \left\lfloor (\hat{R} - \sum_{i \in I_L} R_i) / \sum_{i \in I_H} R_i \right\rfloor;$$

$L = 1$;

**else**

$$H = c_{\text{target}};$$

$$L = \left\lfloor (\hat{R} - c_{\text{target}} \sum_{i \in I_H} R_i) / \sum_{i \in I_L} R_i \right\rfloor;$$

**end if**
Error Concealment

• With FEC, lost packets are NOT retransmitted, hence errors must be concealed for decoding
• Motion information in transportation videos is critical for tracking applications
• Conventional boundary matching algorithms (BMAs) do NOT explicitly retain motion information
• Proposed solution:
  • Motion-copy (MC) algorithm
  • Copies motion information (motion vectors and reference picture index) from co-located blocks in previous frames
  • Retains motion consistency
Numerical Examples

• Experimental setup:
  • Open-source JM encoder with FMO enabled
  • JM decoder is modified to implement the MC and frame-copy (FC) error concealment algorithms
  • OpenCV implementation of mean-shift tracking
  • Publicly available test sequences are used
    • “Camera6” sequence under NGSIM license shows an intersection with light traffic
    • “dt_passat” sequence by courtesy of KOGS/IAKS Universitat Karlsruhe shows a busy intersection interrupted by traffic signal and urban rail crossing
Numerical Examples

- Effect of TDT
  - Compared with baseline H.264 encoding
  - Packet loss probability = 0.1
Numerical Examples

- Comparison of error concealment strategies
  - Boundary matching algorithm (BMA)
  - Motion copy algorithm (MC)
  - Frame copy algorithm (FC, simplified MC with zero motion vectors)

![Graphs showing tracking accuracy vs. bitrate for Camera6 and dt_passat](image-url)
Numerical Examples

- ROI extraction
  - “dt_passat” sequence with identified ROI map and slice group boundary

(a) ROI mapping

(b) Frame w/ slice group mapping
Numerical Examples

- Integrated system comparison
  - Proposed system: UEP (H=3, L=2), MC
  - Reference system: EEP (protection=3), BMA
- Intermediate system-1: UEP (H=3, L=2), BMA
- Intermediate system-2: EEP (protection=3), MC
- All sequences are TDT-filtered to remove temporal noise
- The experiments show 60% bitrate reduction for maintaining the similar tracking accuracy, compared with the state-of-the-art system.
Numerical Examples

- Integrated system comparison

![Graph showing Tracking Accuracy vs. Bitrate for different systems: UEP32-MC, UEP32-BMA, EEP3-MC, EEP3-BMA. The graph is labeled (b) dt_passat.](http://example.com/graph.png)
Numerical Examples

- Integrated system comparison
Conclusions

• We proposed video coding and transmission system specifically tailored to automatic vehicle tracking and surveillance.
• Characteristics of transportation video and lossy wireless channels were taken into consideration.
• Unequal error protection in conjunction with tracking-aware error concealment strategies were proposed to mitigate negative channel impairments.
• The effectiveness of the proposed system was demonstrated and performance improvement over the state-of-the-art system was shown.
Thank you