

## NGHIÊN CỨU ẢNH HƯỞNG CỦA CẤU TRÚC GOP LÊN CHẤT LƯỢNG TRUYỀN VIDEO H.265 TRONG MÔI TRƯỜNG WLAN

**Lương Thị Thảo Hiếu\*, Lê Thị Thu Hiền**  
*Trường Đại học Kinh tế - Kỹ thuật Công nghiệp*

### TÓM TẮT

Những tiến bộ trong truyền thông không dây và công nghệ di động đã gia tăng đáng kể các dịch vụ video cho người dùng đầu cuối, tuy nhiên, trải nghiệm video chất lượng cao trên môi trường không dây luôn là một thách thức. Xu hướng này làm tăng nhu cầu nghiên cứu, đánh giá chất lượng video trong môi trường mạng đặc biệt là mạng không dây. Chất lượng hình ảnh video phụ thuộc vào nhiều tham số khác nhau trong đó có nhóm hình ảnh GOP. Trong bài báo này chúng tôi tập trung nghiên cứu ảnh hưởng của cấu trúc GOP đến chất lượng video H.265 trong môi trường mạng không dây biến đổi, việc đánh giá được thực hiện dựa trên mô phỏng truyền các chuỗi video được thiết kế với các cấu trúc GOP khác nhau trên môi trường WLAN trong ns2 tích hợp Evalvid. Các thí nghiệm chứng minh rằng chất lượng video được cải thiện khi kích thước của GOP tăng.

**Từ khóa:** GOP; HEVC/H.265; Evalvid; ns2; PSNR

*Ngày nhận bài: 26/02/2020; Ngày hoàn thiện: 04/5/2020; Ngày đăng: 11/5/2020*

## RESEARCH EFFECTIVE OF GOP STRUCT ON QUALITY OF H.265 IN WLAN ENVIRONMENT

**Lương Thị Thao Hiếu\*, Lê Thị Thu Hiền**  
*University of Economic and Technical Industry*

### ABSTRACT

Advances in wireless communications and mobile technology have dramatically increased video services for end user, however, transferring high quality video over wireless is always a challenge. This trend increases the demand for researching and assessing video quality in network environments, especially wireless. The video quality depends on a number of parameters including the GOP (Group of Picture). This paper focuses on the impact of GOP structure on H.265 video quality. The evaluation was carried out based on the transmission simulation of video sequences designed with different GOP structures on WLAN environment in ns2 integrate Evalvid framework. Experiments demonstrate that video quality will improve as the size of the GOP increase.

**Keywords:** GOP; HEVC/H.265; Evalvid; ns2; PSNR

*Received: 26/02/2020; Revised: 04/5/2020; Published: 11/5/2020*

\* Corresponding author. Email: ltthieu@uneti.edu.vn

## 1. Mở đầu

Dịch vụ đa phương tiện phát triển mạnh, cùng với nhu cầu sử dụng các ứng dụng video thời gian thực với độ phân giải cao đã thúc đẩy ITU-T và ISO/IEC phát triển chuẩn mã hóa video thế hệ mới - mã hóa video hiệu quả cao HEVC (High Efficiency Video Coding) hay còn gọi là chuẩn H.265. H.265 là chuẩn mã hóa video mới nhất (có mặt tháng 1/2013), cải thiện hơn 50% khả năng nén so với H.264 nhưng vẫn cho hình ảnh chất lượng [1], hơn nữa H.265 được thiết kế để đáp ứng truyền thông video chất lượng cao ngay cả trong điều kiện băng thông thấp, chuẩn này tiêu thụ băng thông giảm một nửa so với H.264 và hướng tới các ứng dụng video có độ phân giải cao như Tivi 4k (4096x2160) và độ phân giải rất cao như Tivi 8k (7680x4320) - Ultra High Definition Television-UHDTV [1]. Sự khác biệt về chất lượng hình ảnh giữa H.264 và H.265 được minh họa trong hình 1.



Hình 1. So sánh hình ảnh H.264 với H.265

Tuy nhiên, khi phân phối video với tốc độ cao, thời gian thực trong môi trường không dây dẫn đến tỉ lệ mất gói lớn làm giảm chất lượng trải nghiệm người dùng đầu cuối (quality of experience - QoE). Hơn nữa, đặc điểm chung của mạng không dây (WLAN) là chất lượng kênh truyền thay đổi theo thời gian và hiệu suất kém hơn so với mạng có dây (LAN), từ đó để cung cấp video hiệu quả, chất lượng tốt trên môi trường không dây luôn là một thách thức. Trong bài báo này, tác giả nghiên cứu các đặc điểm chính của chuẩn mã hóa H.265, tập trung vào nhóm hình ảnh GOP, nghiên cứu ảnh hưởng của cấu trúc

GOP đến chất lượng video, tạo video H.265 với các GOP độ dài khác nhau, sau đó mô phỏng, phân tích, đánh giá chất lượng truyền video trong môi trường mạng không dây dựa trên khung làm việc Evalvid tích hợp trong ns 2.35, từ đó đề xuất cấu trúc GOP phù hợp với video H.265 khi sử dụng trong môi trường WLAN.

## 2. Một số chuẩn MPEG tiên tiến

### 2.1. H.264/AVC

Chuẩn H.264 còn được gọi là MPEG-AVC (Advanced Video Coding), được công bố năm 2003, hiện là chuẩn hỗ trợ công nghệ nén tiên tiến và hiệu quả. Chuẩn này kế thừa những ưu điểm nổi trội của những chuẩn nén trước đây, đồng thời sử dụng những thuật toán nén và phương thức truyền hình ảnh mới phức tạp. H.264 đã làm giảm đáng kể dữ liệu và băng thông truyền đi của video.

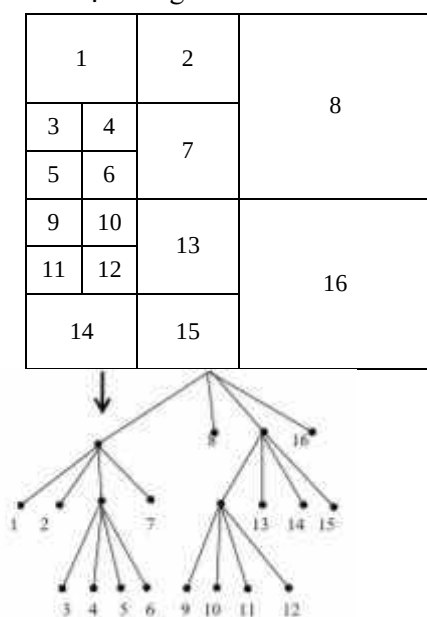
### 2.2. H.265/HEVC

Chuẩn H.265/HEVC được thiết kế đạt nhiều mục tiêu, bao gồm mã hóa hiệu quả, dễ dàng tích hợp hệ thống, khả năng phục hồi mất mát dữ liệu cũng như tăng khả năng thực hiện bằng cách sử dụng kiến trúc xử lý song song. Với cùng chất lượng hình ảnh, một chuỗi video H.265 chiếm ít dung lượng lưu trữ hoặc ít dung lượng đường truyền hơn chuỗi video H.264 tương ứng. Khi cùng băng thông đường truyền hoặc không gian lưu trữ, thì chất lượng độ phân giải của H.265 cao hơn H.264. Lớp mã hóa video của H.265 còn sử dụng các kỹ thuật hỗn hợp bao gồm cả dự đoán nội suy và mã hóa biến đổi 2 chiều. Với những cải tiến này, H.265 đạt hiệu quả mã hóa cao hơn và khả năng phục hồi mất mát dữ liệu tốt hơn so với các chuẩn trước đây, đồng thời nó hỗ trợ video chất lượng Ultra HD (điều này khó thực hiện với H.264), hơn nữa nó có khả năng thích ứng tốt hơn trong điều kiện mạng biến đổi [2].

## 3. Đặc tính nổi bật của H.265

Chuẩn H.265 được mã hóa dựa trên nguyên lý mã hóa lai block-based: mỗi ảnh trong suốt

quá trình mã hóa được chia thành các block vuông có cùng kích thước, mỗi block này đóng vai trò là gốc của cây mã hóa khối (coding tree block - CTB). Các CTB lại được chia nhỏ thành các coding block (CB) - là các thực thể mà bộ mã hóa phải quyết định dùng dự đoán intra-picture (dự đoán bù chuyển động) hoặc inter-picture (khai thác dư thừa không gian giữa các block lân cận). Bản đồ phân vùng của các block được đưa đến bộ giải mã, ảnh đầu tiên trong chuỗi video được mã hóa bằng dự đoán bù chuyển động, các ảnh còn lại được mã hóa bằng dự đoán khai thác dư thừa không gian [3]. Do mỗi ảnh trong HEVC được phân chia thành các CTB, một CTB của các mẫu luma kết hợp với 2 CTB của các mẫu chroma tạo thành một đơn vị Cây mã hóa - Coding Tree Unit (CTU). Mỗi CTU được chia thành nhiều đơn vị mã hóa - coding unit (CU) có kích thước  $L \times L$ ,  $L$  có thể biến đổi 16, 32 hoặc 64 (trong khi đó H.264 sử dụng kích thước  $L$  cố định là 16 x16). Kích thước CU lớn hơn giúp HEVC hiệu quả khi mã hóa video với độ phân giải cao; tuy nhiên cũng có thể làm tăng độ trễ mã hóa/giải mã cũng như tăng độ phức tạp tính toán và bộ nhớ. Phân hoạch CTU thành các CU minh họa trong hình 2.



Hình 2. Phân hoạch cây CTU thành các CU [3]

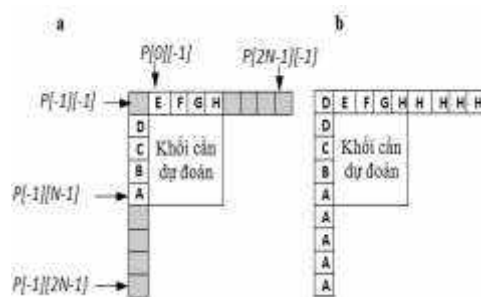
HEVC sử dụng khung dự đoán nội suy, thực hiện bằng cách nội suy các giá trị mẫu từ các mẫu tham chiếu, quá trình bao gồm ba bước: xây dựng mảng mẫu tham chiếu, dự đoán mẫu, và hoàn tất quá trình xử lý. Các bước này được thiết kế để đạt hiệu quả mã hóa cao. Điều hình sử dụng phương pháp dự đoán angular - góc, và planar - phẳng: phương pháp dự đoán góc sử dụng 33 hướng để dự đoán, với độ chính xác 1/32 thay cho 8 hướng (trong H.264). Thao tác cơ bản thực hiện dự đoán góc trong một mẫu được mô tả như sau:  $p = (u * a + v * b + 16) \gg 5$  (1)

với  $u$  và  $v$  là hai hằng số cho trước,  $a$  và  $b$  là các giá trị mẫu tham chiếu. Điều này có nghĩa là dự đoán góc cần 5 thao tác: hai phép nhân, hai phép cộng, một phép dịch chuyển để dự đoán mẫu. Trong H.264 dự đoán một mẫu được thực hiện như sau:

$$p = (a + 2 * b + c + 2) \gg 2 \quad (2)$$

với  $a, b, c$  là các mẫu tham chiếu, điều này có nghĩa là để dự đoán một mẫu H.264, cần ba phép cộng, một phép nhân và một phép dịch chuyển, như vậy độ phức tạp ít hơn HEVC.

Dự đoán phẳng trong HEVC được sử dụng để tạo ra một bề mặt dự đoán khi có ranh giới giữa các biên trong khối - làm mịn biên - thực hiện bằng cách lấy trung bình dự đoán tuyến tính theo chiều ngang và dọc trên cơ sở mẫu (H.264 không hỗ trợ làm mịn biên). Trong trường hợp một số mẫu tham chiếu lân cận không có sẵn, để dự đoán HEVC sử dụng chế độ dự đoán thay thế các mẫu tham chiếu (minh họa trong hình 3), điều này trong H.264 không hỗ trợ,



Hình 3. Minh họa thay thế các mẫu tham chiếu [3]

- Các mẫu tham chiếu trước khi thay thế
- Các mẫu tham chiếu sau khi thay thế

Ngoài ra, H.265 còn sử dụng bộ lọc in-loop de-blocking (tương tự H.264), tuy nhiên được thiết kế tối ưu hóa để xử lý song song. Điểm khác biệt giữa đặc tính mã hóa của H.265 và H.264 được thể hiện qua các thông số trong bảng 1.

**Bảng 1.** So sánh đặc tính mã hóa giữa H.265 và H.264

| Đặc tính                   | H.265           | H.264                        |
|----------------------------|-----------------|------------------------------|
| Kích thước CU              | 16, 32, 64      | 16                           |
| Bộ lọc in-loop de-blocking | Xử lý song song | Không hỗ trợ xử lý song song |
| Dự đoán góc                | 33 hướng        | 8 hướng                      |
| Làm mịn biên               | Có              | Không                        |
| Thay thế mẫu tham chiếu    | Có              | Không                        |

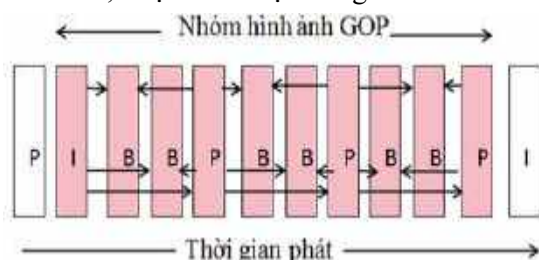
#### 4. Xây dựng video H.265 với các cấu trúc GOP khác nhau

#### 4.1. Cấu trúc GOP

Với phương pháp mã hoá video hiện đại sẽ mã hóa mỗi khung hình thành: khung I (khung chính), P (khung dự đoán) hoặc B (hai chiều), có một trật tự và tần số xuất hiện xác định của các khung I, P, B trong các tập tin video. Khung I, xuất hiện đều đặn tại các đoạn, phân chia video thành "nhóm các hình ảnh" hay GOP [4], GOP đóng vai trò quan trọng trong chất lượng video. Thông thường, GOP bắt đầu bằng khung I và chỉ tồn tại một khung hình I.

Có hai tham số quan trọng để mô tả cấu trúc khung trong GOP, m: số lượng khung B giữa hai khung P liên tiếp, n: số lượng khung B và P giữa hai khung hình I liên tiếp, đây cũng chính là độ dài GOP.

Một mẫu GOP điển hình theo thứ tự là IBBPBBPBBPBB, và được mô tả với  $n = 12$  và  $m = 3$ , được minh họa trong hình 4.

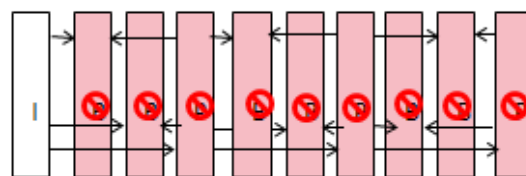


**Hình 4. Nhóm hình ảnh GOP [4]**

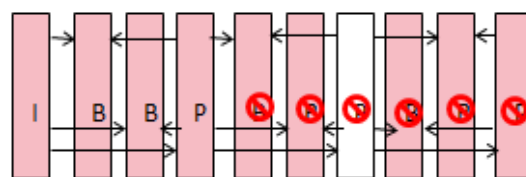
Khung I là khung hình cần thiết nhất, còn được gọi là khung “neo”, nếu khung I bị mất thì các khung khác trong GOP không được giải mã. Mỗi liên hệ giữa sự mất khung hình I, P, B đến quá trình mã hóa được minh họa trong hình 5.

Về mặt lý thuyết lẫn thực nghiệm, chúng ta có thể thay đổi số lượng khung I trong GOP (thay đổi độ dài của GOP), để thu được video với cấu trúc và chất lượng khác nhau, từ đó tùy từng môi trường truyền, lựa chọn GOP phù hợp sẽ cho chất lượng trải nghiệm người dùng đầu cuối tốt nhất.

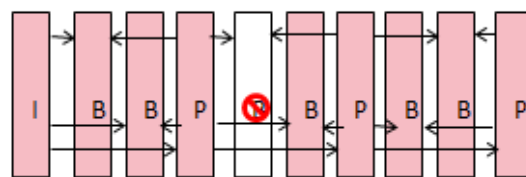
Mất khung I =không giải mã



Mắt khung P



### Mắt khung B



**Hình 5.** Ảnh hưởng của mất khung I, P, B đến quá trình mã hóa [4]

#### 4.2. Quy trình tạo HEVC với cấu trúc GOP khác nhau

Trong phần này, giới thiệu qui trình tạo file video theo định dạng HEVC, xuất phát từ file video thô yuv, kết quả tạo ra file video HEVC có cấu trúc GOP khác nhau.

### Bước 1: Mã hóa file yuv thành dạng .H265

```
./ffmpeg -s cif -r 30 -f test.yuv -vcodec  
x265 -g 12 testtx265.h265
```

**Bước 2:** Sử dụng MP4Box để chứa các frame được tạo ra.

```
$.MP4Box -hint -mtu 1024 -fps 30 testltx265.h265 testltx265.mp4
```

**Bước 3:** Tạo file vết video (st)

```
$/mp4trace -f -s 192.168.0.2 12346 testltx265.mp4 > st
```

Trong qui trình xây dựng, chúng tôi thay đổi cấu trúc GOP của HEVC ở bước thứ 2, sau đó thực hiện mô phỏng truyền video để đánh giá chính xác ảnh hưởng của cấu trúc GOP lên chất lượng hình ảnh khi sử dụng trong môi trường mạng không dây biến đổi.

## 5. Thực nghiệm và kết quả

### Yêu cầu:

Hệ điều hành Ubuntu 18.04, Evalvid 2.6, ns 2.35. Tích hợp Evalvid vào ns2: sửa đổi code nguồn trong file packet.h, agent.h, agent.cc trong thư mục ns-allinone-2.35/ns-2.35. Thay đổi nội dung file makefile trong thư mục /ns-allinone-2.35/ ns-2.35, sau đó biên dịch lại ns bằng lệnh `$.configure -with-tcl-ver=8.5.10; make clean; make [5]`.

Thực hiện viết kịch bản tcl mô phỏng, đánh giá truyền video HEVC với độ dài GOP khác nhau, thông qua 2 thông số: độ trễ trung bình của một gói tin và PSNR - tỉ số tín hiệu cực đại trên nhiễu (peak signal - to - noise ratio), đây là một thuật ngữ dùng để tính tỉ lệ giữa giá trị năng lượng tối đa của một tín hiệu và năng lượng nhiễu ảnh hưởng đến độ chính xác của thông tin, PSNR được biểu diễn bởi đơn vị logarithm decibel, giá trị thông thường của PSNR nằm từ 30 đến 50 dB, giá trị càng cao chất lượng video càng tốt. Mỗi quan hệ giữa PSNR và MOS được minh họa trong bảng 2.

**Bảng 2.** Mối quan hệ giữa PSNR và MOS [6]

| PSNR [DB] | MOS            |
|-----------|----------------|
| >37       | 5 (rất tốt)    |
| 31 – 37   | 4 (tốt)        |
| 25 – 31   | 3 (trung bình) |
| 20 – 25   | 2 (tồi)        |
| < 20      | 1 (rất tồi)    |

Kịch bản sử dụng mô hình mạng không dây gồm 8 nút được minh họa trong hình 6.

+ Nút  $n_0$  gửi video - truyền các gói video (dữ liệu trong file st) đến nút  $n_1$  - nhận video.

+ Nút  $su_1$  gửi dữ liệu CBR truyền với giao thức UDP đến đích  $ru_1$  với tốc độ 265 Kbps và kích thước gói là 1000 byte.

+ Nút  $su_2$  gửi dữ liệu FTP sử dụng giao thức TCP đến đích  $ru_2$ . Liên kết giữa  $r_0$  và AP (Access Point) có băng thông 5 Mb và độ trễ 10 ms.

+ Thời gian tiến hành mô phỏng là thời gian để truyền hết toàn bộ video (250s), hàng đợi đặt tại các đường truyền là DropTail.



**Hình 6.** Mô hình mạng không dây

Các thông số chi tiết minh họa trong bảng 3.

**Bảng 3.** Thông số cấu hình mạng

| Tham số              | Chi tiết    |
|----------------------|-------------|
| Diện tích mô phỏng   | 670 x 670   |
| Thời gian mô phỏng   | 250s        |
| Tính linh động node  | Động        |
| Băng thông           | 5Mbs        |
| Kích thước gói video | 1052Byte    |
| Lớp MAC              | IEEE 802.11 |
| Tốc độ CBR           | 0.265Mbps   |
| Kiểu Antenna         | Omni        |
| Giao tiếp mạng       | Wireless    |
| Giao thức truyền tải | TCP/UDP     |

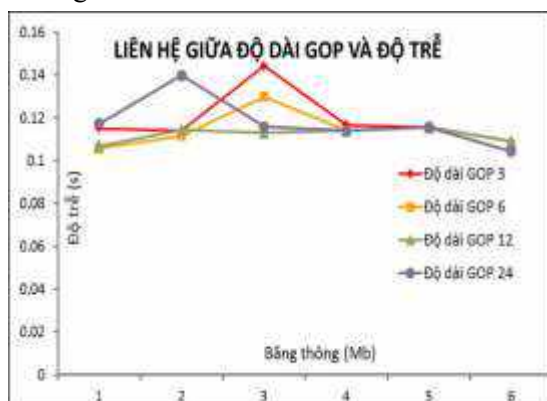
Kết quả mô phỏng được thể hiện qua hai đồ thị sau:



**Hình 7.** Mối liên hệ giữa độ dài GOP và PSNR



Qua hình 7 cho thấy, trong môi trường mạng không dây biến đổi, khi tốc độ bit thấp (dưới 5Mbps) độ dài GOP có ảnh hưởng rõ rệt đến chất lượng video (các đoạn đồ thị có khoảng cách tách biệt). GOP càng lớn cho chất lượng video càng tốt, khi tốc độ bit cao (trên 5Mbps), các đường đồ thị có xu hướng xẹp gần nhau. Điều này có nghĩa là sự khác biệt giá trị PSNR của các mẫu GOP khác nhau được thu hẹp lại, và chất lượng của video có GOP 12, 24 là gần như nhau.



Hình 8. Ảnh hưởng của GOP tới độ trễ video

Qua hình 8 cho thấy, khi băng thông đường truyền thấp dưới 3Mb, độ trễ trung bình của một gói tin video của các độ dài GOP 3, 6, 24 là không ổn định, lúc nhanh, lúc chậm, độ trễ trung bình của gói tin video có độ dài GOP 12 tương đối ổn định. Khi băng thông từ 3Mb trở lên, độ trễ trung bình của một gói tin video bắt đầu giảm dần đều khi độ dài GOP tăng dần.

## 6. Kết luận

Trong bài báo này, chúng tôi nghiên cứu các đặc điểm mã hóa tiên tiến của chuẩn nén H.265, ảnh hưởng của cấu trúc GOP đến chất lượng video, nghiên cứu tạo video H.265 với

các cấu trúc GOP độ dài khác nhau, sau đó đánh giá chất lượng truyền video H.265 qua môi trường mạng WLAN để thấy được ảnh hưởng của độ dài GOP đến chất lượng hình ảnh. Kết quả, dựa trên một số thang đo: PSNR và độ trễ trung bình của một gói tin nhận được trên đường truyền mạng không dây cho thấy chất lượng video HEVC/H.265 tăng dần khi tăng độ dài GOP. Trong điều kiện mạng không dây có băng thông thấp, tốc độ bit nhỏ, video H.265 với GOP độ dài 12 cho kết quả trải nghiệm thực sự hiệu quả.

## TÀI LIỆU THAM KHẢO/ REFERENCES

- [1]. K. E. Psannis, "HEVC in wireless environments," *Journal of Real-Time Image Processing*, vol. 14, pp. 1-9, 2015.
- [2]. J. F. Miroslav, "UHRINA1 Impact of H.264/AVC and H.265/HEVC Compression," *Digital image processing and computer graphics*, pp. 368-376, 2014.
- [3]. V. Sze, *High Efficiency Video Coding (HEVC) Algorithms and Architectures*, Springer International Publishing Switzerland, 2014, pp. 91-104.
- [4]. S. I. A. Huszák, "Analysing GOP Structure and Packet Loss Effects on Error", *Control and signal processing*, ISCCSP 2010, 2010, pp. 1-4.
- [5]. S. Nefti, "PSNR and Jitter Analysis of Routing Protocols for Video Streaming in Sparse MANET Networks, using NS2 and the Evalvid Framework", *International Journal of Computer Science Information Security*, vol. 14, pp. 1-6, 2016.
- [6]. U. S. Zulpratita, "GOP Length Effect Analysis on H.264/AVC Streaming Transmission Quality over LTE Network", *3rd International Conference on Computer Science and Information Technology*, 2013, pp. 5-7.