

Phân tích, Đánh giá Hiệu năng Mã hóa Video với Chuẩn H.265/HEVC

Tác giả: Bùi Thanh Hương^{1,2}, Phí Công Huy^{2,3,4}, Đinh Triều Dương², Hoàng Văn Xiêm²

¹ Khoa Công nghệ Thông tin, trường Đại học Xây dựng

² Khoa Điện tử - Viễn thông, trường Đại học Công nghệ, Đại học Quốc gia Hà Nội

³ Khoa Đa phương tiện, Học viện Công nghệ Bưu chính Viễn Thông

⁴ Faculty of Engineering and Information Technology, University of Technology Sydney
xiemhoang@vnu.edu.vn

Tóm tắt—Ngày nay, song song với sự ra đời của các loại video chất lượng cao (độ phân giải lên đến 8k×4k) thì nhu cầu về dịch vụ video đáp ứng được khả năng truyền dẫn (chưa thể tăng ngay về thông lượng) trên hạ tầng mạng hiện tại cũng như trên các thiết bị cá nhân hiện đại (như điện thoại di động, máy tính bảng) ngày càng gia tăng. Việc cải tiến các chuẩn mã hóa (nén) video là nhu cầu tất yếu đã và đang được quan tâm phát triển trong vài thập kỷ qua. Chuẩn mã hóa video hiện đại nhất cho đến thời điểm này là chuẩn H.265/HEVC (High Efficiency Video Coding) với nhiều cải tiến về hiệu quả mã hóa, tích hợp hệ thống truyền tải, khôi phục mất dữ liệu cũng như thực hiện kiến trúc xử lý song song. Trên cơ sở đó, bài báo trình bày kiến trúc tổng quan của chuẩn H.265/HEVC, phân tích những điểm mới, nổi bật so với các chuẩn mã hóa video cũ và mô phỏng đánh giá hiệu năng của chuẩn H.265/HEVC trên cơ sở hai yếu tố: hiệu năng nén và độ phức tạp thuật toán. Kết quả nghiên cứu sẽ góp phần thúc đẩy việc khai thác chuẩn H.265/HEVC trong các ứng dụng thực tế tại Việt Nam và trên thế giới.

Từ khóa—Chuẩn H.265/HEVC, Mã hóa video, hiệu năng mã hóa, độ phức tạp

I. GIỚI THIỆU

Trong nhiều năm qua, nhu cầu sử dụng các dịch vụ video ngày càng cao, đòi hỏi chất lượng video tăng về độ phân giải đã đặt ra thách thức trong giới truyền thông là cần có kỹ thuật nén video hiệu quả để giảm được lượng bit cần biểu diễn cho một video, đồng thời đảm bảo người dùng vẫn thu được video có chất lượng cao như mong muốn. Có như vậy, mặc dù có nhiều hạn chế về sự phát triển của hệ thống kênh truyền video (bị giới hạn về băng thông) và của các thiết bị từ phía người dùng (hạn chế về nguồn năng lượng vận hành thiết bị) thì các dịch vụ đáp ứng nhu cầu ngày càng cao trong sử dụng video vẫn tăng trưởng nhanh chóng. Đây cũng là lý do cho thấy trong hơn 3 thập kỷ gần đây, các chuẩn nén video liên tục được nâng cấp và chuẩn hóa.

Sau 4 năm, kể từ khi chuẩn H.120 là chuẩn nén video đầu tiên do một đơn vị trong tổ chức viễn thông quốc tế ITU-T (the International Telecommunications Union Telecommunication Standardization Sector) [1] giới thiệu vào năm 1984 thì chuẩn H.261 ra đời. Kể từ đó, ngoài ITU-T còn có một tổ chức khác cùng đồng hành trong việc cải tiến chuẩn nén video, đó chính là tổ chức các tiêu chuẩn thế giới ISO/IEC (the International Standardization Organization/International Electrotechnical Commission) [2]. Minh chứng cho điều này, các chuẩn cải tiến tiếp theo của chuẩn H.261 là các chuẩn H.262/MPEG-2 [3], H.263, H.263++/MPEG-4 [4], H.264/AVC [5] lần lượt được hai tổ chức này đồng thời (và sau đó là phối hợp) công bố vào các năm tương ứng: 1995, 1996, 1999-2000, 2003. Các chuẩn này được cải tiến để lần lượt đáp ứng các nhu cầu phát triển các dịch vụ như: hội nghị truyền hình với độ phân giải video tiêu chuẩn SD (Standard Definition) có tốc độ 1.5-2 Mbit/s (chưa phát quang bá và phân tán), hội nghị truyền hình quốc tế (với tốc

độ bitrate và độ phân giải tiêu chuẩn được mở rộng hơn), các dịch vụ DVD phổ rộng, truyền dẫn truyền hình vệ tinh số và các hệ thống HDTV và phân tán video qua Internet, truyền thông video thời gian thực.

Với mục tiêu đáp ứng nhu cầu ngày càng cao cho việc phát triển các ứng dụng và dịch vụ truyền thông video có độ phân giải lớn (HD, High Definition: 1280×720; full HD: 1920×1080) tới siêu lớn (Ultra HD 4K: 3840×2160), năm 2013, chuẩn mã hóa video H.265/HEVC (High Efficiency Video Coding) ra đời [6]. Cho đến nay, chuẩn H.265/HEVC vẫn là chuẩn mã hóa nén video tốt nhất và vẫn tuân theo kiến trúc lai kết hợp giữa mã hóa dự đoán và mã hóa biến đổi [6 - 8]. So với chuẩn H.264/AVC, H.265/HEVC tập trung giải quyết được việc nén video có độ phân giải tăng cao và kiến trúc xử lý song song, cụ thể là cải tiến các phép dự đoán trong khung, liên khung cho kết quả dự đoán tốt hơn, sử dụng phép biến đổi cosin rời rạc với nhiều kích thước hơn [6]. Kết quả là H.264/HEVC cải thiện được hiệu suất nén, so với H.264/AVC thì H.265/HEVC tiết kiệm được từ 30% đến 50% lượng bitrate mà vẫn cho video có chất lượng tương đương [9,10]. HEVC giúp phát triển các ứng dụng truyền tải video có độ phân giải, tốc độ khung hình và dải động cao cũng như gam màu rộng hơn [8].

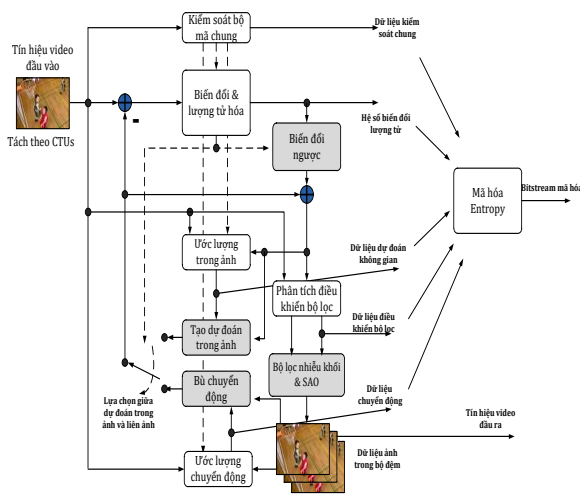
Để có được các cải tiến mạnh mẽ về mã hóa nén video trong chuẩn HEVC, nhiều nhà nghiên cứu đã đưa ra nhiều đề xuất mới nhằm nâng cao chất lượng mã hóa theo hai khía cạnh là cải thiện hiệu năng nén và giảm độ phức tạp của các thuật toán nén dữ liệu video. Trong khuôn khổ của bài báo này, chúng tôi tập trung mô tả chuẩn mã hóa nén video HEVC (phần II), từ đó, trong phần III mô tả một số đề xuất mới trong chuẩn mã hóa video H.265/HEVC cho truyền thông đa phương tiện theo hai hướng chính, đó là nâng cao hiệu năng mã hóa và giảm độ phức tạp trong mã hóa video. Tiếp theo, trước khi đưa ra các kết luận, trong phần IV chúng tôi trình bày các đánh giá cụ thể và phân tích cho từng đề xuất mới thông qua các kết quả chạy thử nghiệm cho các đề xuất như bộ tối ưu phép lượng tử (RDOQ: Rate-Distortion Optimized Quantization), bộ bù mẫu thích ứng (SAO: Sample Adaptive Offset) và phân chia khối chuyển động bất đối xứng (AMP: Asymmetric Motion Partitioning) hay phương thức tìm kiếm nhanh TZSearch (Test Zone Search), FDM (Fast Decision for Merge RD Cost) và TransformSkip.

II. MÔ TẢ CHUẨN MÃ HÓA VIDEO THẾ HỆ MỚI H.265/HEVC

A. Kiến trúc chuẩn H.265/HEVC

Hình 1 mô tả kiến trúc tổng quan của bộ mã hóa H.265/HEVC. Một ảnh gốc sau khi được chia thành các khối (block) có kích thước 8×8, mỗi khối có 64 điểm ảnh 8 bit thì được đưa vào bộ mã hóa nén video cùng với ảnh dự đoán (được tạo ra ở cả phía mã hóa và phía giải mã). Bộ mã hóa nén video chỉ mã hóa thông tin khác biệt giữa ảnh gốc và ảnh dự đoán. Tức là lúc này, thông tin khác biệt lần lượt được tác

động bởi các phép biến đổi, lượng tử hóa (xuôi và ngược) rồi được thực hiện dự đoán (trong khung hoặc liên khung), lọc vòng lặp, sau đó cho được ảnh tái tạo ở đầu ra bộ mã hóa. Dữ liệu ảnh tái tạo này cũng được đưa vào lưu trữ ở bộ đệm (phía bộ giải mã) đồng thời được dùng cho việc tạo thông tin dự đoán ở phía bộ mã hóa.



Hình 1: Kiến trúc bộ mã hóa video H.265/HEVC

B. Các thành phần cơ bản

Là bộ mã hóa video lai, chuẩn H.265/HEVC có các khối chính sau: khối biến đổi và lượng tử hóa, mã hóa entropy, mã hóa dự đoán trong khung và dự đoán liên khung, bộ lọc vòng lặp. Dưới đây là các mô tả chức năng cho từng khối.

• Phép biến đổi và lượng tử hóa

Đầu tiên, thông tin khác biệt này sẽ qua khối biến đổi và lượng tử hóa. Mục tiêu của phép biến đổi là biến đổi các hệ số dư thừa thành các hệ số càng nhỏ càng tốt trong vùng biến đổi, do đó, phô bày biên độ lớn, phần còn lại coi như không đáng kể. Hay chính là giảm sự tương đồng vẫn tồn tại ở khối thông tin khác biệt (là sự khác biệt (hiệu) giữa các khối ban đầu với các khối dự đoán).

Phép biến đổi hay dùng trong mã hóa nén video là phép biến đổi cosin rời rạc DCT (Discrete Cosine Transform) loại II₃ được dùng nhiều (kể từ H.261/MPEG-1). Còn phép biến đổi sin rời rạc DST (Discrete Sine Transform) loại VII, được chọn bổ sung trong HEVC [11 - 15].

Sau khi biến đổi, để giảm bớt số lượng đầu ra (số trọng số) có thể, phép lượng tử hóa được áp dụng. Đây là phép toán xấp xỉ nên có gây ra tổn thất dữ liệu. HEVC sử dụng phép lượng tử vô hướng (mỗi hệ số xấp xỉ độc lập với các giá trị xung quanh). Việc lượng tử hóa là thao tác không thể hủy bỏ trong toàn bộ tiến trình mã hóa video lai (là loại mã hóa có mất mát dữ liệu).

• Mã hóa Entropy

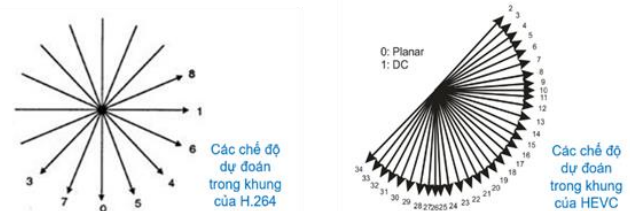
Mã hóa entropy giúp giảm lượng bit truyền và là hoạt động không mất mát dữ liệu vì không có tính toán xấp xỉ trong giai đoạn này. Trong chuẩn HEVC, mã hóa entropy vẫn dùng mã hóa số học nhị phân thích ứng ngữ cảnh CABAC (Context Adaptive Binary Arithmetic Coding) nhưng có 1 số cải tiến so với chuẩn trước đó, để cải thiện tốc độ truyền dữ liệu (có kiến trúc xử lý song song) và tăng hiệu năng nén cũng như giảm yêu cầu bộ nhớ ngữ cảnh [6].

• Mã hóa dự đoán

Tùy vào thông tin tham chiếu ở trên cùng một ảnh hay trên ảnh đã có trước đó, ta có hai phương thức dự đoán là dự đoán trong khung hay dự đoán liên khung. Các khối mẫu lân cận của khối hiện tại trong chính ảnh đang được mã hóa có thể được dùng để bộ mã hóa thực hiện dự đoán, cách thức này gọi là dự đoán trong khung. Còn các ảnh đã được mã hóa xong có thể được bộ mã hóa tham chiếu để thực hiện dự đoán cho ảnh hiện tại, cách thức này chính là dự đoán liên khung. Đối với ảnh đầu tiên của video khi mã hóa chỉ áp dụng dự đoán trong khung vì không có các ảnh nào khác phía trước đó để tham chiếu. Đối với các ảnh tiếp theo, bộ mã hóa sẽ quyết định lựa chọn cách thức dự đoán liên khung hay trong khung tùy thuộc vào một số chỉ số như tối ưu hóa tỉ lệ méo [6, 7].

• Dự đoán trong khung (dự đoán không gian)

Các mẫu ranh giới đã giải mã trước đó của các khối lân cận sẽ là dữ liệu tham khảo cho dự đoán trong khung. Trong HEVC, dự đoán trong khung có 35 chế độ hay hướng, trong khi chuẩn H.264/MPEG-4 AVC có 8 chế độ hay hướng. Cụ thể: 33 hướng [2÷34]: định nghĩa cho PU (vuông) kích cỡ từ 4×4 đến 32×32 và 2 chế độ: planar[0] và DC[1]. Các góc độ được thiết kế có chủ đích (tăng hiệu quả dự đoán tín hiệu).



Hình 2: Các chế độ dự đoán trong khung [6, 16]

• Dự đoán liên khung (dự đoán thời gian)

Thay vì mã hóa trực tiếp các giá trị pixel thô cho mỗi block đó, người ta tìm một khối tương tự trên khung đã được mã hóa trước đó để tham chiếu. Nếu tìm kiếm thành công thì khối này được mã hóa bởi một vector chuyển động (là vector chỉ ra vị trí của khối phù hợp với khung tham chiếu). Quá trình xác định vector chuyển động được gọi là ước lượng chuyển động. Khối tìm thấy có thể không phải là khối phù hợp chính xác với khối mã hóa nên cần tính toán sự khác biệt giữa chúng (các giá trị dư thừa này gọi là lỗi dự báo, được chuyển đổi và gửi tới bộ giải mã). Vậy để tìm được khối phù hợp trên khung tham chiếu, có 2 phần tử có thể phục hồi các pixel thô của khối là vector chuyển động trỏ tới khối phù hợp và lỗi dự báo (tức thông tin khác biệt).

• Bộ lọc

Mục đích của bộ lọc là giúp đạt hiệu quả tốt trong mã hóa video và nâng cao chất lượng hình ảnh tái tạo. Ảnh được lọc sẽ được lưu trữ trong bộ đệm ảnh giải mã dùng để tham khảo cho dự đoán liên khung.

Có 2 loại bộ lọc là bộ lọc vòng giải mã khối DBF (In-loop deblocking filtering) và bộ lọc mẫu bù đắp thích nghi SAO (Sample Adaptive Offset). Bộ lọc thứ nhất là DBF, được áp dụng trên một khu vực hoặc toàn bộ ảnh. Còn bộ lọc thứ hai là bộ lọc SAO, được áp dụng trên phạm vi không gian cục bộ của ảnh [6 - 8].

III. CÁC ĐỀ XUẤT MỚI TRONG CHUẨN H.265/HEVC

A. Các đề xuất nâng cao hiệu năng mã hóa

Như đã nói ở trên, việc cải tiến chuẩn mã hóa nén video thường được chú trọng theo hai hướng chính là nâng cao hiệu năng mã hóa và giảm độ phức tạp cho quy trình thực hiện mã hóa nén thông tin video. Dưới đây là các đề xuất hướng đến việc nâng cao hiệu năng mã hóa.

- **RDOQ (Rate-Distortion Optimized Quantization)**

Mục đích của RDOQ là tìm ra tập hợp các hệ số biến đổi lượng tử tối ưu để biểu diễn dữ liệu dư thừa trong khối được mã hóa. RDOQ tính toán biến dạng hình ảnh (bằng cách lượng tử hóa các hệ số biến đổi) trong khối được mã hóa và tính toán một lượng bit cần thiết để mã hóa hệ số biến đổi lượng tử tương ứng [17]. Dựa trên hai giá trị này, bộ mã hóa chọn hệ số tốt hơn, bằng cách tính chi phí RD.

- **SAO (Sample Adaptive Offset)**

Một số nhiễu xuất hiện trong các tiêu chuẩn mã hóa trước đó cũng xảy ra trong HEVC vì việc dự đoán trong khung và liên khung, phép biến đổi và lượng tử hóa trong HEVC vẫn được xử lý dựa trên khối. Các nhiễu này là nhiễu khối, nhiễu lặp và nhiễu gây mờ [18].

Để loại bỏ các nhiễu này, HEVC sử dụng hai bộ lọc vòng lặp, đó là DBF và SAO. DBF được áp dụng cho các ranh giới của khối đã tái tạo lại để giảm các thành phần nhiễu khối. Ngoài DBF ra, một kỹ thuật mới SAO được áp dụng sau khi thực hiện DBF xong. SAO rất hữu ích để giảm nhiễu lặp, là loại nhiễu chủ yếu đến từ các lỗi lượng tử hóa của các hệ số biến đổi [18].

Nhiệm vụ của SAO là thực hiện bù cho các mẫu được tái tạo bằng cách thêm một phần bù vào từng pixel, do đó, sự biến dạng giữa hình ảnh tái tạo lại và ảnh gốc có thể được giảm. Vấn đề chính của phương pháp này là làm thế nào để phân loại các mẫu đã được tái tạo lại và cách chọn các độ lệch cho mỗi loại [18]. Có hai loại SAO khác nhau để lựa chọn là bù dải (BO-band offset) và bù cạnh (EO-edge offset).

- **AMP (Asymmetric Motion Partitioning)**

Việc phân vùng CB thành PB thông thường là phân vùng đối xứng, đây là cách phân vùng đã được áp dụng từ chuẩn H.264/AVC, nhưng HEVC cho phép bổ sung thêm bốn phân vùng không đối xứng, gọi là phân vùng chuyển động bất đối xứng (AMP). AMP giúp phân vùng hiệu quả trong các trường hợp có sự chồng chéo một phần của một đối tượng nền phía trước và nền phía sau trong khu vực của một khối mã hóa [7]. Kích thước khối nhỏ nhất để dự đoán bù chuyển động là 4×8 và 8×4 (do đó, dự đoán liên khung với các khối 4×4 bị loại trừ).

B. Các đề xuất giảm độ phức tạp thuật toán

Bên cạnh các nghiên cứu về nâng cao hiệu năng mã hóa video, một số đề xuất sau đây đã giúp giảm độ phức tạp trong nén video, tức giảm thiểu thời gian mã hóa.

- **Tìm kiếm nhanh (Fast Search)**

Thuật toán TZSearch (Test Zone Search) là sự pha trộn giữa tìm kiếm theo vùng và các mẫu tìm kiếm raster. Thuật toán tuân theo các bước sau:

- **Dự đoán vector chuyển động:** Thuật toán TZSearch sử dụng công cụ dự đoán trung vị cho các lân cận trái, trên

và phải. Việc tối thiểu các dự đoán được chọn làm vị trí bắt đầu cho các bước tìm kiếm tiếp theo.

- **Tìm kiếm lưới ban đầu:** chọn cửa sổ tìm kiếm bằng cách sử dụng mô hình kim cương hoặc hình vuông với kích thước khác nhau từ 1 đến 64. Các mẫu được sử dụng là tìm kiếm kim cương tám điểm hoặc tìm kiếm hình vuông tám điểm. Điểm có vector chuyển động với SAD nhỏ nhất được lấy làm điểm tìm kiếm trung tâm cho các bước tiếp theo.

- **Tìm kiếm raster:** Tìm kiếm raster là hoạt động tìm kiếm quét toàn bộ (full search) một phiên bản được lấy mẫu xuống của cửa sổ tìm kiếm. Cửa sổ tìm kiếm sẽ bắt đầu quét theo tuần tự từ trái qua phải và từ trên xuống dưới.

- **Tinh chỉnh Raster/Star:** sàng lọc tốt của các vector chuyển động thu được từ bước trước theo một trong hai tinh chỉnh raster hoặc sàng lọc mẫu hình vuông/kim cương (tinh chỉnh sao) có thể được kích hoạt. Lúc này, mẫu hình vuông tám điểm hoặc mô hình kim cương tám điểm được sử dụng. Hai phương pháp sàng lọc khác nhau ở hoạt động tìm kiếm của chúng. Tổng số điểm tìm kiếm phụ thuộc vào từng chuỗi video khác nhau.

- **FDM (Fast Decision for Merge RD Cost)**

FDM là việc ra quyết định nhanh và hiệu quả cho thuật toán hợp nhất chi phí RD nhằm tránh ước tính tất cả các chi phí biến dạng tỷ lệ của các ứng cử viên hợp nhất. Hay nói cách khác là để không phải thực hiện một số đánh giá chi phí biến dạng tỷ lệ cho một số ứng cử viên hợp nhất, một quy tắc chấm dứt sớm việc ước tính chi phí biến dạng tỷ lệ của các dự đoán vector chuyển động được thực hiện ở phía bộ mã hóa [19].

- **Transform Skip**

Việc thực hiện thông thường của biến đổi 2D trong bộ mã hóa video tiêu chuẩn dựa trên phép biến đổi DCT và DST với số nguyên - vì các biến đổi 1D được áp dụng trên các hàng và cột của các đơn vị biến đổi. Trong khi đối với một số loại tín hiệu nhất định, việc thực hiện các biến đổi theo cả hai hướng có thể là một sự lựa chọn tối ưu về mặt nén năng lượng, cho tín hiệu với các tính chất khác nhau, một số chiến lược khác có thể dẫn đến hiệu suất Tỷ lệ-Biến dạng (RD) tốt hơn. Điều này có thể xảy ra và phổ biến trong một số trường hợp tín hiệu 2D ở đầu vào của phép biến đổi để hiển thị thông kê khác nhau theo trục dọc và trục ngang. Phương pháp này đã được giải quyết trong HEVC bằng cách áp dụng DST đối với thông tin dư thừa trong khung liên quan đến các chế độ dự đoán theo hướng nhất định. Tuy nhiên, một số loại thông tin dư thừa có thể được hưởng lợi từ việc bỏ qua hoàn toàn biến đổi trong khung. Đối với tín hiệu bù chuyển động thưa (sparser) mà cũng thể hiện các đặc tính khác nhau theo chiều ngang và dọc thì có thể được xử lý bằng phương pháp bỏ qua biến đổi các khối theo chiều ngang và/hoặc dọc, để cải thiện nén [20].

IV. ĐÁNH GIÁ VÀ PHÂN TÍCH

Để đánh giá hiệu năng mã hóa chuẩn H.265/HEVC, chúng tôi thực hiện việc mô phỏng và kiểm thử trên 04 chuỗi video với độ phân giải 832×840 và 416×240 tương ứng [21]. Chế độ mã hóa Random Access được thiết lập trên bộ codec HM16.10 với nền tảng phần cứng Intel core i7, RAM 8GB. Các thuật toán mới được kiểm thử và đánh giá lần lượt. Thông số cụ thể và tên gọi các chuỗi video được mô tả ở Bảng 1. Hiệu năng mã hóa được xác định thông qua phép

tính đại lượng BD-Rate (Bjontegaard Delta rate) [22], trong đó nếu BD-Rate mang giá trị âm thì phương pháp đề xuất nâng cao hiệu năng mã hóa và ngược lại.

BẢNG 1: CÁC ĐIỀU KIỆN CHẠY KIỂM THỬ

TT	Chuỗi video	Tổng số frame	Tốc độ khung hình	Nhóm video
1	RaceHorses	-f 297	30 frames/s	D
2	BlowingBubbles	-f 497	50 frames/s	
3	PartyScene	-f 497	50 frames/s	C
4	BasketballDrill	-f 497	50 frames/s	

A. Đánh giá hiệu năng mã hóa

Bảng 2 trình bày kết quả đánh giá so sánh hiệu năng mã hóa video của từng thuật toán với trường hợp kích hoạt toàn bộ các thuật toán.

BẢNG 2: KẾT QUẢ NÂNG CAO HIỆU NĂNG NÉN VIDEO

TT	Tool	ClassD		ClassC	
		RaceHorses	BlowingBubbles	PartyScene	BasketballDrill
		BD-Rate (%)			
1	RDOQ	-4.61	-4.78	-5.25	-4.11
2	HadamardME	-2.21	-1.76	-1.39	-1.23
3	AMP	-1.26	-1.07	-0.82	-0.70
4	SAO	-1.37	0.51	-0.23	-2.10
5	FastSearch	1.45	0.77	1.01	0.88
6	FEN	0.51	0.36	0.51	0.74
7	FDM	0.08	0.06	0.07	0.13
8	TransformSkip	0.02	-0.53	-0.71	-0.28

Kết quả thực nghiệm cho thấy:

Đề xuất tối ưu các hệ số biến đổi lượng tử RDOQ giúp cải thiện hiệu năng mã hóa đáng kể nhất khi chạy thực nghiệm các khối mới (bảng 2), kết quả cải thiện được từ 4-5% bitrate. Tiếp theo là các đề xuất HadamardME và AMP cũng cho các kết quả tương đối khả quan, tiết kiệm được 1-2% bitrate. Ngoài ra, bộ lọc SAO cải thiện kết quả tốt cho một số video, một số video khác thì bộ lọc này không phát huy được ưu điểm của nó (xem Bảng 2).

B. Đánh giá độ phức tạp mã hóa

BẢNG 3: KẾT QUẢ GIẢM ĐỘ PHỨC TẠP NÉN VIDEO

TT	Tool	ClassD		ClassC	
		RaceHorses	BlowingBubbles	PartyScene	BasketballDrill
		ΔT (%)			
1	RDOQ	10.40	10.26	11.18	9.78
2	HadamardME	10.34	12.94	12.40	14.34
3	AMP	20.26	14.84	15.63	14.66
4	SAO	0.43	0.7	0.61	0.50
5	FastSearch	-95.54	-96.48	-96.20	-95.64
6	FEN	-23.38	-21.06	-21.27	-23.41
7	FDM	-5.27	-8.69	-7.43	-10.52
8	TransformSkip	7.89	7.98	7.33	7.76

Bảng 3 tổng hợp kết quả so sánh độ phức tạp mã hóa video của từng công cụ với trường hợp tắt cả các thuật toán được kích hoạt, với đại lượng ΔT được tính bởi công thức sau:

$$\Delta T = \frac{T_{on} - T_{off}}{T_{off}} * 100 \quad (1)$$

Trong đó: T_{on} là thời gian mã hóa video khi kích hoạt tất cả các công cụ; T_{off} là thời gian mã hóa video khi tắt công cụ tương ứng.

Kết quả thực nghiệm cho thấy:

Với các đề xuất giảm độ phức tạp của thuật toán, việc đưa ra được giải thuật tối ưu thời gian tìm kiếm trong dự đoán véc tơ chuyển động là một đề xuất có tính đột phá, đóng góp rất lớn cho cải tiến chuẩn nén video. Kết quả thực nghiệm ở bảng 3 chỉ ra đề xuất này đã giúp giảm đến trên 95% thời gian cần để mã hóa nén video trong khi chất lượng nén video gần như giảm không đáng kể (1%). Còn đề xuất FEN và đề xuất giảm thiểu việc ước tính các chi phí biến dạng tỷ lệ của các ứng cử viên hợp nhất FDM, kết quả thực nghiệm cho thấy hai kỹ thuật này đã góp phần giảm lần lượt tương ứng khoảng 22% và gần 8% thời gian nén video.

BẢNG 4: ĐÁNH GIÁ TỔNG HỢP TỪNG ĐỀ XUẤT

TT	Tool	Nhận xét khi bật tool	
		Hiệu năng nén	Thời gian mã hóa
1	RDOQ	tăng đáng kể nhất	tăng (10%)
2	HadamardME	tăng (1,65%)	tăng (11%)
3	AMP	tăng (gần 1%)	tăng (16%)
4	SAO	có sequence tăng, có sequence giảm	gần như không tăng
5	FastSearch	giảm không đáng kể	giảm đáng kể nhất (96%)
6	FEN	giảm rất ít (0,5%)	giảm nhiều (30%)
7	FDM	gần như không giảm	giảm (5-10%)
8	TransformSkip	tăng (0,38%)	tăng (7%)

Ngoài ra, việc chạy thực nghiệm với chuỗi video PartyScene, khi thay phiên tắt các thuật toán, đa số đều cho kết quả thay đổi không đáng kể.

KẾT LUẬN

Trong bài báo này, chúng tôi đã tổng hợp và đưa ra những đánh giá về hiệu năng mã hóa và khả năng giảm độ phức tạp của chuẩn mã hóa video H.265/HEVC. Với các đóng góp đáng kể, các công cụ trong chuẩn mã hóa H.265/HEVC đã khắc phục và giải quyết được những vấn đề còn tồn tại ở chuẩn mã hóa video H.264/AVC. Cụ thể là đóng góp của đề xuất RDOQ đã cải thiện được 4-5% lượng bitrate qua việc tối ưu các hệ số biến đổi lượng tử và bên cạnh đóng góp giảm khoảng 95% thời gian mã hóa của thuật toán TZSearch thì đề xuất FEN cũng giúp giảm độ phức tạp thuật toán, tiết kiệm đến 22% thời gian mã hóa. Dựa trên đánh giá tổng quan này, bài báo cũng sẽ là cơ sở để phân tích và so sánh hiệu năng của H.265/HEVC với các chuẩn mã hóa khác trong tương lai.

LỜI CẢM ƠN

Nghiên cứu này được tài trợ bởi Trường Đại học Công nghệ, Đại học Quốc gia Hà Nội theo đề tài mã số CN18.13.

TÀI LIỆU THAM KHẢO

- [1] ITU-T VCEG : <http://www.itu.int/en/ITU-T/studygroups/com16/video/Pages/default.aspx>
- [2] ISO/IEC MPEG: <https://mpeg.chiariglione.org/>
- [3] ITU-T and ISO/IEC JTC 1, "Generic coding of moving pictures and associated audio information – Part 2: Video," ITU-T Rec. H.262-ISO/IEC 13818-2, Nov. 1994.

- [4] A. Olivier, E. Alexandros, H. Carsten, R. Ganesh, W. Liam, "MPEG-4 Systems: Overview," *Signal Processing: Image Communication*, pp.281-298, vol. 15, no. 4-5, 2000.
- [5] T. Wiegand, G.J. Sullivan, G. Bjontegaard, A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560-576, 2003.
- [6] G.J. Sullivan, et al., "Overview of the High Efficiency Video Coding (HEVC) Standard", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649-1668, 2012.
- [7] M. Wien, "High Efficiency Video Coding", *Signals and Communication Technology*, Springer-Verlag Berlin Heidelberg, 2015.
- [8] V. Sze, M. Budagavi, G. J. Sullivan, "High Efficiency Video Coding (HEVC) Algorithms and Architectures", *Integrated Circuits and Systems*, Springer International Publishing Switzerland, 2014.
- [9] Z. Milicevic, and Z. Bojkovic, "HEVC vs. H.264/AVC through performance and complexity comparison," in *Proceedings of the 13th International Conference on Applications of Computer Engineering (ACE'14)*, Lisbon, Portugal, 2014.
- [10] K. Zeng, A. Rehman, J. Wang, and Z. Wang, "From H.264 to HEVC: coding gain predicted by objective video quality assessment models," in *Proceedings of International Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM2013)*, Scottsdale, AZ, 2013.
- [11] N. Ahmed, T. Natarajan, and K. Rao, "Discrete Cosine Transform," *IEEE Trans. Comput.*, vol. C-23, no. 1, pp. 90-93, Jan. 1974.
- [12] M. Budagavi, A. Fuldseth, G. Bjontegaard, V. Sze, and M. Sadafale, "Core Transform Design in the High Efficiency Video Coding (HEVC) Standard," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 6, pp. 1029-1041, Dec. 2013.
- [13] V. Britanak, P. C. Yip, and K. R. Rao, "Discrete Cosine and Sine Transforms: General Properties, Fast Algorithms and Integer Approximations," Elsevier, Sep. 2006.
- [14] R. K. Chivukula and Y. A. Reznik, "Fast Computing of Discrete Cosine and Sine Transforms of Types VI and VII," in *Proc. SPIE 8135, Applications of Digital Image Processing XXXIV*, no. 813505, pp. 1-10, Sep. 2011.
- [15] Y. A. Reznik, "Relationship between DCT-II, DCT-VI, and DST-VII transforms," in *Proc. 2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 5642-5646, May. 2013.
- [16] G. J. Sullivan, P. Topiwala, and A. Luthra, "The H.264/AVC Advanced Video Coding Standard: Overview and Introduction to the Fidelity Range Extensions", *SPIE Conference on Applications of Digital Image Processing XXVII*, Aug. 2004.
- [17] J. Stankowski, C. Korzeniewski, M. Domański, T. Grajek "Rate-Distortion Optimized Quantization in HEVC: Performance Limitations", *2015 Picture Coding Symposium (PCS)*, Cairns, QLD, pp. 85-89, Jun. 2015.
- [18] H. Zhang, O. C. Au, Y. Shi, W. Zhu, V. Jakhetiya, L. Jia, "Improved Sample Adaptive Offset for HEVC", *2013 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, Kaohsiung, pp. 1-4, 2013.
- [19] B. G. Kim, K. Goswami, "Basic Prediction Techniques in Modern Video Coding Standards", *SpringerBriefs in Electrical and Computer Engineering*, 2016.
- [20] M. Mrak, J. Z. Xu, "Improving screen content coding in HEVC by transform skipping", *European Signal Processing Conference (EUSIPCO)*, 2012.
- [21] HEVC Video test sequences, [Online]. Available: <ftp://hevc@ftp.tnt.uni hannover.de/testsequences/>.
- [22] G. Bjontegaard, "Calculation of average PSNR differences between RD curves", in: *ITU-T SC16/Q6 13th VCEG Meeting*, Austin, Apr. 2001.