

PHƯƠNG PHÁP TẠO THÔNG TIN PHỤ TRỢ MỚI CHO HỆ THỐNG MÃ HÓA VIDEO LIÊN LỚP PHÂN TÁN

Nguyễn Thị Hương Thảo, Vũ Văn San

Học viện Công nghệ Bưu chính Viễn thông

Tóm tắt: Video ngày càng trở nên quan trọng trong cuộc sống. Mã hóa video đã tiến một bước rất dài với rất nhiều các cải tiến quan trọng. Với các yêu cầu mới, trong những năm gần đây, mã hóa video liên lớp ngày càng được quan tâm nhiều hơn. Có thể kể đến các chuẩn mã hóa video liên lớp hiện nay như SVC, mở rộng khả năng liên lớp của H264/AVC hay SHVC, mở rộng khả năng liên lớp cho chuẩn mã hóa video HEVC. Bên cạnh đó, các giải pháp mã hóa video liên lớp phân tán (DSVC) đã được đề xuất. Các kết quả nghiên cứu cho thấy DSVC mang đến một giải pháp hoàn hảo cho các ứng dụng vì có nhiều đặc điểm nổi trội như bộ mã hóa có độ phức tạp thấp, khả năng chống lỗi cao trong khi vẫn tương thích với các chuẩn mã hóa video hiện thời. Cũng giống như mã hóa video phân tán (DVC), thông tin phụ trợ (SI) được tạo ra ở bộ giải mã đóng vai trò quan trọng đối với hiệu năng hệ thống DSVC. Vì vậy, bài báo này giới thiệu một phương pháp tạo thông tin phụ trợ mới cho hệ thống DSVC. Phương pháp đề xuất sử dụng kỹ thuật kết hợp ở mức khối giữa khung hình ở lớp cơ sở và khung hình ở lớp tăng cường để tạo ra một khung hình SI. Kết quả thực nghiệm cho thấy phương pháp mới này tạo ra thông tin phụ trợ có chất lượng tốt hơn so với một số phương pháp tạo thông tin phụ trợ trước đó.

Từ khóa: Video phân tán, mã hóa video liên lớp, thông tin phụ trợ.

I. GIỚI THIỆU

Nhu cầu sử dụng video tăng nhanh trong những năm gần đây và điều này càng đặt ra nhiều yêu cầu đối với các kỹ thuật mã hóa video. Mục đích của mã hóa video là nén dữ liệu video với số lượng bit ít hơn mà vẫn đáp ứng các yêu cầu liên quan của một ứng dụng cụ thể nào đó. Các tiêu chuẩn mã hóa video hiện nay chủ yếu dựa vào các kỹ thuật sau đây: 1) các kỹ thuật ước lượng và bù chuyển động để loại bỏ dư thừa thời gian; 2) các kỹ thuật mã hóa và lượng tử hóa để loại bỏ dư thừa không gian và tận dụng các đặc điểm của hệ thống thị giác người và 3) kỹ thuật mã hóa entropy để loại bỏ dư thừa thống kê. Hầu hết các tác vụ này đều được thực hiện tại phía bộ mã hóa và vì vậy, độ phức tạp bộ mã hóa là rất cao so với độ phức tạp bộ giải mã. Điều này trở thành trở ngại cho các ứng dụng mới yêu cầu độ phức tạp bộ mã hóa thấp và khả năng chống lỗi cao. Giải pháp đầy hứa hẹn có thể hỗ trợ cho các ứng dụng này là mã hóa video phân tán (DVC) [1]. Giờ đây, mã hóa video phân tán (DVC) đã không còn xa lạ với rất nhiều các giải pháp được đề xuất với các kết quả rất khả quan. Nhắc đến mã hóa video phân tán là nói đến các ưu điểm mà các kỹ thuật mã hóa video truyền thống không có được ví dụ như bộ mã hóa độ phức tạp thấp, khả năng chống lỗi cao. Các kết quả thực nghiệm của các

codec video phân tán thực tế cho thấy sự phù hợp của DVC đối với các ứng dụng này.

Một điều quan trọng cần phải nhắc đến là thói quen giải trí video đã thay đổi. Với sự phát triển nhanh chóng của các thiết bị thông minh khác nhau như điện thoại di động, máy tính bảng hay máy tính cầm tay, có thể nói người xem có thể xem video mọi lúc, mọi nơi bất cứ khi nào họ muốn. Cùng một nội dung video có thể được truyền trên các đường truyền có băng thông khác nhau, chất lượng khác nhau và được hiển thị trên các thiết bị có kích thước khác nhau, độ phân giải khác nhau và khả năng xử lý khác nhau. Người ta gọi đó là hiện trạng mạng và thiết bị hỗn tạp. Trước tình hình này, một yêu cầu khác đặt ra cho các chuẩn mã hóa video hiện thời là khả năng liên lớp. Làm sao để nội dung video đến được với nhiều người sử dụng trong nhiều tình huống mạng khác nhau và được xem trên các thiết bị đầu cuối khác nhau.

Đứng trước các thay đổi này, một yêu cầu đặt ra đối với các giải pháp mã hóa video cho các ứng dụng mới là phải đáp ứng được các yếu tố sau: độ phức tạp bộ mã hóa thấp, khả năng chống lỗi cao và có khả năng liên lớp. Câu trả lời là hệ thống mã hóa video liên lớp phân tán (DSVC) [2,3]. DSVC đề cập tới cả các giải pháp mã hóa video mà ở đó có sử dụng các nguyên tắc mã hóa video phân tán đồng thời vẫn cung cấp khả năng liên lớp. Cũng giống như các hệ thống mã hóa video liên lớp khác, DSVC gồm một lớp cơ sở (BL) và một hoặc nhiều lớp tăng cường (EL). Để có thể tương thích với các chuẩn mã hóa video hiện nay, thông thường lớp cơ sở được mã hóa bằng các chuẩn mã hóa video truyền thống như H264/AVC hay HEVC và lớp tăng cường được mã hóa theo nguyên tắc mã hóa video phân tán.

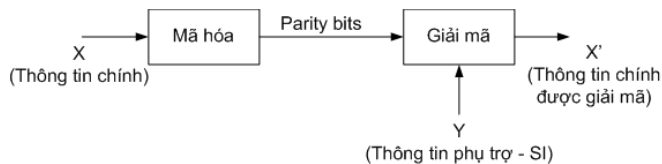
Để hiểu hơn về DSVC, trước hết bài báo giới thiệu về mã hóa nguồn phân tán vì đây là nền tảng cho mã hóa video phân tán. Mã hóa nguồn phân tán dựa trên hai kết quả lý thuyết thông tin quan trọng là định lý Slepian-Wolf và Wyner-Ziv. Các định lý phát biểu rằng khi so sánh với sơ đồ mã hóa truyền thống sẽ không có sự tổn thất về hiệu suất nén nếu hai hoặc nhiều nguồn độc lập thống kê được mã hóa độc lập và giải mã kết hợp. Hai định lý này đã mở ra các cơ hội mới cho mã hóa video hay còn gọi là mã hóa video phân tán. Do việc mã hóa được thực hiện độc lập và việc giải mã thực hiện kết hợp nên dư thừa thời gian chỉ được khai thác tại bộ giải mã thay vì bộ mã hóa, do đó bộ mã hóa có độ phức tạp rất thấp nếu so với mã hóa video truyền thống. Trong DVC, chuỗi X là thông tin

Tác giả liên hệ: Nguyễn Thị Hương Thảo

email: thaonth@ptit.edu.vn

Đến tòa soạn: 7/2017, chỉnh sửa: 8/2017, chấp nhận đăng: 9/2017.

chính, thông tin được mã hóa Wyner-Ziv và chuỗi Y là thông tin phụ trợ (SI). Thông tin phụ trợ là một dự đoán của thông tin chính và được tạo ra ở phía giải mã. Bộ giải mã sẽ sử dụng thông tin phụ trợ này kết hợp với một số thông tin bổ sung từ phía mã hóa gửi đến để dự đoán ra thông tin chính X (Hình 1). Do đó, nếu tương quan giữa chuỗi X và chuỗi Y càng cao thì số lượng bit bộ mã hóa cần gửi sang cho bộ giải mã càng ít. Nói cách khác, hiệu quả các kỹ thuật tạo thông tin phụ trợ sẽ ảnh hưởng đáng kể đến hiệu năng tổng thể của codec video phân tán hay còn gọi là codec video Wyner-Ziv (WZ). Cho đến hiện nay, rất nhiều các nghiên cứu tập trung vào các kỹ thuật tạo thông tin phụ trợ. Thông tin phụ trợ có thể được tạo ra theo cách tiếp cận kiểu dự đoán [4,5]. Trong [4], Aaron và các cộng sự đã đề xuất một phương pháp tạo thông tin phụ trợ bằng cách lấy trung bình hai khung hình chính gần nhất và thực hiện nội suy bù chuyển động dựa trên các vector chuyển động đối xứng. Trong [5], SI được tạo ra theo bốn bước: 1) ước lượng chuyển động trước; 2) ước lượng chuyển động song hướng; 3) làm mịn chuyển động không gian và 4) bù chuyển động song hướng. Một cách tiếp cận khác trong việc tạo thông tin phụ trợ là dựa trên kỹ thuật học [6]. Thay vì giữ nguyên thông tin phụ trợ trong toàn bộ quá trình giải mã khung hình WZ, ở đây thông tin phụ trợ được cải thiện liên tục khi bộ giải mã nhận được càng nhiều bit chẵn lẻ từ phía bộ mã hóa. Các tác giả trong [7,8] đã đề xuất cách tiếp cận dựa trên kỹ thuật thử trong đó quá trình giải mã và lựa chọn thông tin phụ trợ được thực hiện đồng thời. Trong các phương pháp này, bộ mã hóa gửi các thông tin bổ sung tới bộ giải mã nhưng bộ giải mã chỉ sử dụng các thông tin này để xác minh tính đúng đắn của thông tin được giải mã.

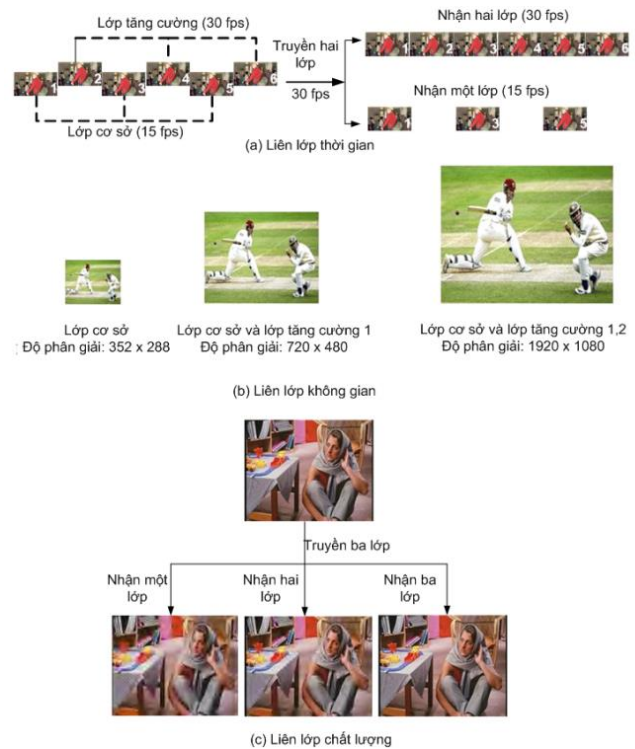


Hình 1. Phương pháp nén video phân tán với thông tin phụ trợ phía giải mã

Nói về khái niệm khả năng liên lớp, đó chính là việc chia luồng video truyền thống thành nhiều luồng bit con và thường gọi là các lớp. Luồng bit con có chất lượng thấp nhất gọi là lớp cơ sở và các luồng bit con còn lại gọi là các lớp tăng cường với chất lượng tốt nhất thuộc về lớp tăng cường cao nhất. Độ phân giải và chất lượng của luồng bit có thể thay đổi tùy theo khả năng xử lý, kích thước màn hình của thiết bị giải mã và tốc độ truyền tải của mạng bằng cách loại bỏ bớt các gói tin từ luồng bit gốc ban đầu. Thông thường có ba khả năng liên lớp là liên lớp thời gian, không gian và chất lượng (hình 2).

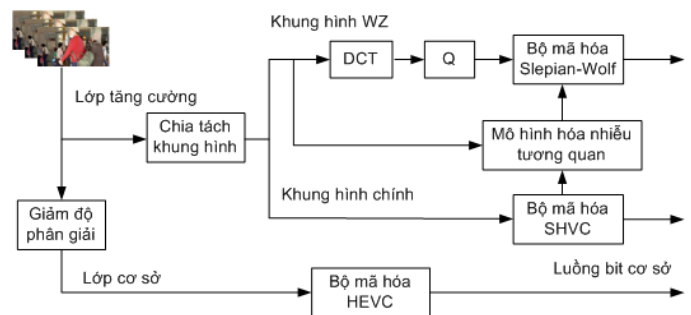
Trong [2,3], Hoàng Văn Xiêm và các cộng sự đã đề xuất kiến trúc mã hóa video phân tán liên lớp thời gian và chất lượng. Bài báo này tập trung vào kiến trúc mã hóa video phân tán liên lớp không gian. Trong kiến trúc DSVC này, các khung hình ở lớp cơ sở có kích thước nhỏ hơn so với các khung hình ở lớp tăng cường. Vì vậy, so với các kiến trúc DSVC liên lớp thời gian và chất lượng, tạo thông tin phụ trợ trong hệ thống DSVC liên lớp không gian có rất nhiều điểm khác biệt. Bài báo này giới thiệu một phương pháp tạo thông tin phụ trợ mới cho hệ thống DSVC liên lớp không gian dựa trên kỹ thuật kết hợp khung hình.

Phần tiếp theo của bài báo được tổ chức như sau. Phần II giới thiệu về kiến trúc DSVC liên lớp không gian được sử dụng trong bài báo. Phương pháp tạo thông tin phụ trợ đề xuất được mô tả trong Phần III. Các kết quả và thảo luận được giới thiệu trong phần IV và phần V là kết luận.



Hình 2. Các khả năng liên lớp

II. KIẾN TRÚC MÃ HÓA VIDEO PHÂN TÁN LIÊN LỚP KHÔNG GIAN



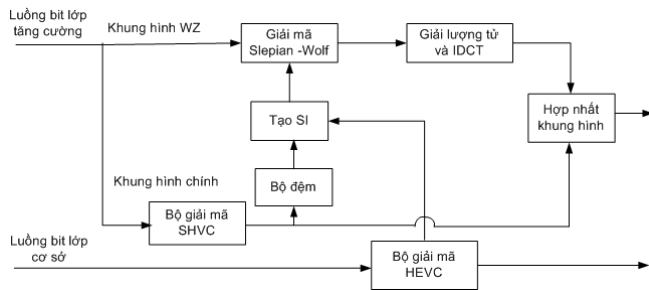
Hình 3. Bộ mã hóa DSVC

Giống như các kiến trúc mã hóa video truyền thống, kiến trúc mã hóa video liên lớp phân tán DSVC cũng bao gồm một lớp cơ sở và một/hoặc nhiều lớp tăng cường. Để đơn giản, trong bài báo này mô tả kiến trúc DSVC gồm một lớp cơ sở và một lớp tăng cường. Để có khả năng tương thích với các chuẩn mã hóa hiện nay, lớp cơ sở được mã hóa theo phương pháp mã hóa video truyền thống, trong mô hình này sử dụng bộ mã hóa và giải mã HEVC. Lớp tăng cường được mã hóa theo phương pháp mã hóa video phân tán hay còn gọi là mã hóa Wyner-Ziv (WZ). Nhóm các khung hình (GOP) được thiết lập bằng 2 nghĩa là cứ một khung hình WZ nằm giữa hai khung hình

chính. Sơ đồ bộ mã hóa và giải mã DSVC được mô tả trong hình 3,4 tương ứng.

A. Quá trình mã hóa

Hình 3 mô tả kiến trúc bộ mã hóa DSVC liên lớp không gian. Chuỗi video gốc sẽ được đưa qua module giảm độ phân giải để tạo ra các khung hình có độ phân giải không gian nhỏ hơn cho lớp cơ sở. Các khung hình lớp cơ sở sẽ được mã hóa bằng bộ mã hóa HEVC. Đối với lớp tăng cường, quá trình mã hóa như sau.



Hình 4. Bộ giải mã DSVC

Chia tách khung hình: Các khung hình lớp tăng cường được chia thành các khung hình WZ (WZF) và khung hình chính (KF) xen kẽ nhau. Các khung hình lẻ là các khung hình chính còn các khung hình chẵn là các khung hình WZ.

Bộ mã hóa SHVC: Các khung hình chính được mã hóa bằng cách sử dụng bộ mã hóa SHVC, chuẩn mở rộng khả năng liên lớp của chuẩn mã hóa video HEVC.

Mô hình hóa nhiễu tương quan: Phương pháp mã hóa video WZ thực hiện mã hóa sự sai khác giữa thông tin gốc tại bộ mã hóa và thông tin phụ trợ được tạo ra tại phía giải mã. Do đó, cả bộ mã hóa và bộ giải mã cần phải biết về tương quan thống kê giữa thông tin gốc và thông tin phụ trợ. Vì thế khối này thực hiện mô hình hóa nhiễu tương quan giữa khung hình WZ gốc và khung hình thông tin phụ trợ tương ứng. Thông tin đầu ra sẽ được gửi tới bộ mã hóa Slepian-Wolf để mã hóa cho khung hình WZ.

DCT: Để khai thác dư thừa không gian, các khung hình WZ sẽ được biến đổi DCT để tạo ra các ma trận hệ số tương ứng.

Lượng tử hóa: Để loại bỏ dư thừa không gian và tận dụng các đặc điểm của thị giác người, các ma trận hệ số DCT được đưa qua bộ lượng tử hóa với các hệ số lượng tử tùy theo chất lượng mong muốn.

Bộ mã hóa Slepian-Wolf: Bộ mã hóa này thực hiện mã hóa các hệ số DCT đã lượng tử hóa có sử dụng thông tin từ khối mô hình hóa nhiễu tương quan. Bộ mã hóa này có thể là bộ mã hóa kênh hoặc bộ mã hóa syndrome tùy thuộc vào các cách tiếp cận khác nhau.

B. Quá trình giải mã

Kiến trúc bộ giải mã DSVC liên lớp không gian được mô tả trong Hình 4. Trước hết, luồng bit cơ sở được giải mã bởi bộ giải mã HEVC. Luồng bit tăng cường được chia thành các khung hình chính và khung hình WZ. Các khung hình chính

được giải mã bởi bộ giải mã SHVC. Các khung hình WZ được giải mã như sau.

Bộ đệm: Bộ đệm lưu giữ các khung hình chính trước và sau khi đã được giải mã. Các khung hình chính này được sử dụng để hỗ trợ cho việc tạo thông tin phụ trợ.

Tạo SI: Thông tin phụ trợ được coi như một phiên bản nhiễu của thông tin gốc. Nó được tạo ra dựa trên các thông tin sẵn có ở phía giải mã. Trong kiến trúc DSVC liên lớp không gian, thông tin phụ trợ có thể được xây dựng dựa trên các khung hình chính trước và sau và các khung hình tương ứng ở lớp cơ sở.

Giải mã Slepian-Wolf: Bộ giải mã Slepian-Wolf thực hiện giải mã thông tin gốc bằng cách sửa sai thông tin phụ trợ dựa trên các thông tin nhận được của bộ mã hóa Slepian-Wolf từ phía mã hóa gửi tới.

Giải lượng tử và IDCT: Sau khi giải mã Slepian-Wolf, thông tin được tái tạo bằng cách giải lượng tử và biến đổi DCT ngược.

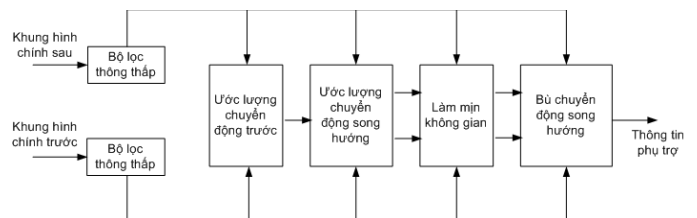
Hộp nhốt khung hình: Các khung hình WZ và khung hình chính được hợp nhất với nhau tạo nên lớp tăng cường.

III. PHƯƠNG PHÁP TẠO THÔNG TIN PHỤ TRỢ

Với kiến trúc DSVC liên lớp không gian được đề cập ở trên, có thể thấy vai trò quan trọng của thông tin phụ trợ đối với hiệu năng của hệ thống. Thông tin phụ trợ được xây dựng càng giống với thông tin gốc thì lượng bit cần gửi bổ sung từ phía mã hóa đến phía giải mã càng ít. Vì vậy bài báo này đề xuất một giải pháp tạo thông tin phụ trợ dựa trên kỹ thuật kết hợp các khung hình ở mức khối. Trong kiến trúc này, có hai ứng viên thông tin phụ trợ được tạo ra: 1) SI_{MCTI} - thông tin phụ trợ được tạo nên bởi kỹ thuật rất phổ biến trong mã hóa video phân tán có tên gọi nội suy thời gian bù chuyển động (MCTI)[5] và 2) SI_{BL} - thông tin phụ trợ được tạo ra bằng cách tăng độ phân giải cho khung hình lớp cơ sở tương ứng. Sau đó một kỹ thuật kết hợp được thực hiện giữa hai ứng viên SI này để tạo ra thông tin phụ trợ cuối cùng SI_{Fusion} .

A. Tạo SI_{MCTI}

Kỹ thuật nội suy thời gian bù chuyển động thực hiện ước lượng khung hình thông tin phụ trợ dựa vào các khung hình chính trước và sau. Hình 5 mô tả các bước thực hiện của kỹ thuật này.



Hình 5. Kỹ thuật nội suy bù chuyển động - MCTI

Bộ lọc thông thấp: Khởi đầu, các khung hình chính được đưa qua bộ lọc thông thấp để cải thiện độ tin cậy của các vector chuyển động.

Ước lượng chuyển động trước: Bước này thực hiện ước lượng một vector chuyển động cho mỗi khối trong khung hình chính sau với tham chiếu tới khung hình chính trước.

Ước lượng chuyển động song hướng: Sử dụng các phép chiếu của vector chuyển động, đối với mỗi khối trong khung hình SI sẽ lựa chọn vector chuyển động nào đi qua khối đó và gần với tâm khối nhất và coi đó là vector chuyển động của khối. Vector chuyển động được lựa chọn sẽ được chia thành hai vector chuyển động trước và sau với giả định chuyển động không đổi.

Làm mịn không gian: Tại bước này, bộ lọc trung vị được thực hiện trên hai trường vector chuyển động để loại bỏ các vector chuyển động ở biên.

Bù chuyển động song hướng: Thực hiện lấy trung bình hai khối đã bù chuyển động trong khung hình chính trước và sau để tạo ra thông tin phụ trợ.

B. Tạo SI_{BL}

Đối với kiến trúc DSVC liên lớp không gian, lớp cơ sở và lớp tăng cường có độ phân giải khác nhau. Lớp cơ sở có thể coi như phiên bản độ phân giải thấp của lớp tăng cường. Do đó SI_{BL} được tạo ra bằng cách tăng độ phân giải cho khung hình lớp cơ sở tương ứng. Trong bài báo này sử dụng kỹ thuật tăng độ phân giải giống như trong SHVC [9]. Mỗi khung hình lớp cơ sở được tăng độ phân giải với bộ lọc đáp ứng xung chiều dài hữu hạn FIR. Khi tăng độ phân giải một khung hình lên 16 lần kích thước của nó và sau đó giảm kích thước với tỉ lệ M ở đó $M = 16$, theo cả hai hướng x và y . Gọi $F_{SL,L}$ và $F_{SL,C}$ các khung hình thông tin phụ trợ với các kênh chói và kênh màu. $F_{BL,L}$ và $F_{BL,C}$ là các khung hình lớp cơ sở với các kênh chói và kênh màu tương ứng. Khi đó mỗi pixel (x,y) trong khung hình SI_{BL} được tính như sau:

$$F_{SL,L}(x,y) = \frac{1}{64^2} \sum_{i=0}^7 f_L(y_{phase}, i) \times \left(\sum_{j=0}^7 f_L(x_{phase}, j) F_{BL,L}(x_{BL} + j - 3, y_{BL} + i - 3) \right) \quad (1)$$

$$F_{SL,C}(x,y) = \frac{1}{64^2} \sum_{i=0}^7 f_C(y_{phase}, i) \times \left(\sum_{j=0}^3 f_C(x_{phase}, j) F_{BL,C}(x_{BL} + j - 1, y_{BL} + i - 1) \right) \quad (2)$$

Trong đó $x_{BL}, y_{BL}, x_{phase}, y_{phase}$ được xác định như sau:

$$\begin{aligned} x_{BL} &= \text{round}(x.N) \\ y &= \text{round}(y.N) \\ x_{phase} &= \text{round}(x.N.16)\%16 \\ y_{phase} &= \text{round}(y.N.16)\%16 \end{aligned} \quad (3)$$

N là hệ số tỉ lệ giữa lớp cơ sở ban đầu và sau khi tăng độ phân giải. $\%$ là toán tử chia lấy phần dư và 'round' là hàm làm tròn đến số nguyên gần nhất. f_L and f_C được định nghĩa trong [9].

Tại biên của khung hình, khi tính toán công thức 1 và 2 mà cần các giá trị bên ngoài khung hình thì sẽ sử dụng các giá trị tại biên đó. Công thức như sau:

$$\begin{aligned} x'_{BL} &= \max(0, \min(F_{width} - 1, x_{BL})) \\ y'_{BL} &= \max(0, \min(F_{height} - 1, y)) \end{aligned} \quad (4)$$

ở đó F_{width}, F_{height} là độ rộng và chiều cao của khung hình.

C. Kết hợp khung hình

Trong phần này, SI_{MCTI} và SI_{BL} được kết hợp với nhau để tạo thành thông tin phụ trợ cuối cùng SI_{Fusion}

Bước 1: Cải thiện chất lượng của khung hình MCTI

Sau khi khung hình MCTI được tạo ra từ các khung hình chính trước và sau, các khối (block) có kích thước 8×8 của ba khung hình này được so sánh với các pixel tương ứng trong khung hình đã tăng độ phân giải. Các khối có sự sai khác nhỏ nhất (Mean Absolute Difference) sẽ được ấn định mang giá trị khối của khung hình MCTI. Giả định rằng $P_{i,j}^B, P_{i,j}^F, P_{i,j}^M$ và $P_{i,j}^U$ là giá trị của các pixel trong các khung hình trước, sau, MCTI và khung hình đã tăng độ phân giải tương ứng tại các tọa độ i và j . Giá trị của khung hình MCTI mới được tính toán như sau:

$$M = \min(\sum_{i,j=1}^{64} |P_{i,j}^B - P_{i,j}^U|, \sum_{i,j=1}^{64} |P_{i,j}^F - P_{i,j}^U|, \sum_{i,j=1}^{64} |P_{i,j}^M - P_{i,j}^U|) \quad (5)$$

$$B_{r,c}^M = \begin{cases} B_{r,c}^B & \text{Nếu } M = \sum_{i,j=1}^{64} |P_{i,j}^B - P_{i,j}^U| \\ B_{r,c}^F & \text{Nếu } M = \sum_{i,j=1}^{64} |P_{i,j}^F - P_{i,j}^U| \end{cases} \quad (6)$$

trong đó $B_{r,c}^M, B_{r,c}^B, B_{r,c}^F$ tương ứng là các khối hình có kích thước 8×8 của khung hình MCTI, khung hình trước và khung hình sau tại tọa độ r,c .

Bước 2: Tạo khung hình kết hợp

Khung hình kết hợp được tạo ra bằng cách kết hợp giữa khung hình đã tăng độ phân giải và khung hình MCTI được tạo ra ở bước 1. Cụ thể, các giá trị pixel trong khung hình kết hợp được tính như sau:

$$P_{i,j}^{Fusion} = \alpha P_{i,j}^M + (1-\alpha) P_{i,j}^U \quad (7)$$

Trong công thức 7, α là trọng số đo lường sự đóng góp của khung hình MCTI và khung hình đã tăng độ phân giải vào khung hình kết hợp. α được lựa chọn bằng 0.1.

IV. ĐIỀU KIỆN THỬ NGHIỆM VÀ PHÂN TÍCH KẾT QUẢ

A. Điều kiện thử nghiệm

Để đánh giá hiệu năng của giải pháp tạo thông tin phụ trợ đề xuất so với các phương pháp khác, bốn chuỗi video sau được sử dụng là BasketballDrill, BQMall, PartyScene và

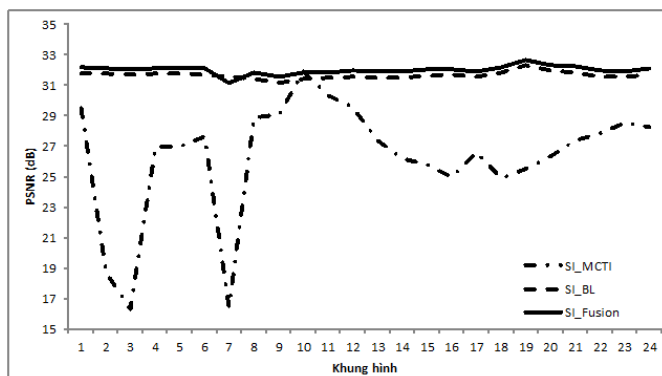
RaceHorses với các đặc điểm được mô tả trong Bảng I. Hình 6 mô tả các khung hình đầu tiên của bốn chuỗi video thử nghiệm.



Hình 6. Các khung hình đầu tiên của 4 chuỗi video thử nghiệm

Bảng I. ĐIỀU KIỆN THỬ NGHIỆM

Chuỗi Video thử nghiệm	Độ phân giải không gian	Độ phân giải thời gian	Số lượng khung hình	Hệ số lượng tử
Basketball Drill	EL: 832 x 480 BL: 416 x 240	50 Hz	50	EL: 32 BL: 30
BQMall		60 Hz	50	EL: 32 BL: 30
PartyScene		50 Hz	50	EL: 32 BL: 30
RaceHorses		30 Hz	50	EL: 32 BL: 30

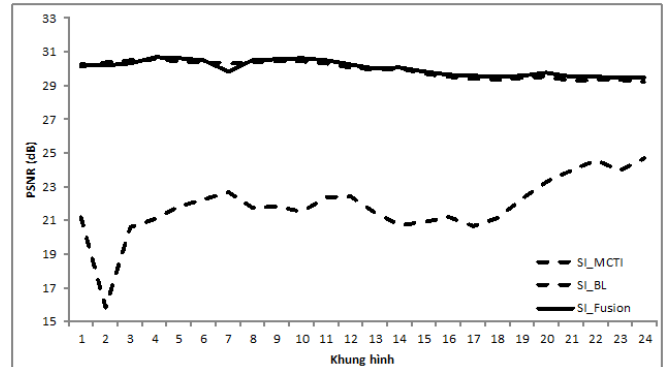


Hình 7. PSNR của khung hình thông tin phụ trợ trong chuỗi BasketballDrill.

B. Phân tích kết quả

Hình 7, 8 mô tả PSNR của các khung hình thông tin phụ trợ trong bốn chuỗi video thử nghiệm. Như quan sát trong hình,

phương pháp MCTI cho kết quả PSNR thấp nhất. Lý do là vì trong phương pháp này, thông tin phụ trợ được tạo ra bởi nội suy chuyển động từ các khung hình trước và sau tại lớp tăng cường. Tuy nhiên, hiệu quả của nội suy phụ thuộc vào chuyển động của khung hình. Vì vậy, chất lượng của khung hình bị thay đổi theo từng khung hình. Trong khi đó, phương pháp tăng độ phân giải từ lớp cơ sở và phương pháp kết hợp cho giá trị PSNR cao hơn, đặc biệt là giá trị PSNR của phương pháp kết hợp. Điều này đạt được là do trong phương pháp kết hợp, khung hình kết hợp được tạo ra từ các khối hình có chất lượng cao nhất được chọn từ khung hình MCTI và khung hình của lớp cơ sở.



Hình 8. PSNR của khung hình thông tin phụ trợ trong chuỗi RaceHorse.

Bảng II mô tả sự so sánh PSNR trung bình của các khung hình SI được tạo ra trong ba phương pháp. Như được chỉ ra trong hình, phương pháp đề xuất đạt được PSNR cao hơn so với các phương pháp MCTI và phương pháp tăng độ phân giải cho lớp cơ sở. Sự cải thiện này đạt được chủ yếu từ việc sử dụng kỹ thuật nâng cao chất lượng của khung hình MCTI và sau đó là sự kết hợp giữa khung hình cơ sở và khung hình MCTI.

Bảng II. PSNR (dB) TRUNG BÌNH CỦA CÁC CHUỖI VIDEO THỬ NGHIỆM

	SI _{MCTI}	SI _{BL}	SI _{Fusion}
BasketballDrill	26.33	31.67	32.01
BQMall	27.70	28.12	28.38
PartyScene	25.27	24.99	25.33
RaceHorses	21.84	29.94	30.04

V. KẾT LUẬN

Bài báo này giới thiệu mô hình kiến trúc DSVC liên lớp không gian và đề xuất kỹ thuật kết hợp khung hình vào việc tạo thông tin phụ trợ. Trong phương pháp này, khung hình thông tin phụ trợ tại bộ giải mã được tạo ra bằng cách sử dụng kỹ thuật kết hợp ở mức khối giữa khung hình được tạo từ kỹ thuật MCTI và khung hình được tạo từ kỹ thuật tăng độ phân giải cho khung hình lớp cơ sở. Các kết quả thực nghiệm so sánh

phương pháp đề xuất với các phương pháp MCTI và phương pháp tăng độ phân giải cho thấy phương pháp đề xuất có thể cải thiện đáng kể chất lượng của khung hình thông tin phụ trợ. Trong các nghiên cứu tiếp theo sẽ tập trung cải thiện các module khác trong kiến trúc DSVC nhằm cải thiện hơn nữa hiệu năng tổng thể của hệ thống.

TÀI LIỆU THAM KHẢO

- [1] P.L. Dragotti and M. Gastpar, "Distributed Source Coding: Theory, Algorithms and Applications," Academic Press, Feb. 2009.
- [2] X. HoangVan, J. Ascenso, and F. Pereira, "HEVC backward compatible scalability: A low encoding complexity distributed video coding based approach," *Signal Process.: Image Commun.*, vol. 33, no. 4, pp. 51-70, Apr. 2015.
- [3] X. HoangVan, J. Ascenso, and F. Pereira., "Adaptive Scalable Video Coding: a HEVC based Framework Combining the Predictive and Distributed Paradigms", *IEEE TCSVT*, vol. 99, no. 00, pp. 1-14, Mar. 2016.
- [4] A. Aaron, R. Thang, and B. Girod (2002) "Wyner-Ziv Coding of Motion Video", in *Proc. Asilomar Conference on Signals and Systems*, Pacific Grove, CA, USA, November 2002.
- [5] J. Ascenso, C. Brites, and F. Pereira, Improving Frame Interpolation with Spatial Motion Smoothing for Pixel Domain Distributed Video Coding" in *EURASIP Conference on Speech and Image Processing, Multimedia Communications and Services (EC-SIPMCS)*, Smolenice, Slovak Republic, June 2005.
- [6] J. Ascenso, C. Brites, and F. Pereira, "Motion Compensated Refinement for Low Complexity Pixel based Distributed Video Coding," in *IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS)*, Como, Italy, September 2005.
- [7] R. Puri and K. Ramchandran, PRISM: A new robust video coding architecture based on distributed compression principles, *40th Allerton Conf. Communication, Control and Computing*, Allerton, IL, USA, 2002.
- [8] R. Puri, A. Majumdar, and K. Ramchandran, "PRISM: a video coding paradigm with motion estimation at the decoder," *IEEE Transactions on Image Processing*, vol. 16, no. 10, pp. 2436-2448, Oct. 2007.
- [9] Jianle Chen; Jill Boyce; Yan Ye; Miska M. Hannuksela; Gary J. Sullivan; Ye kui Wang. HEVC Scalable Extensions (SHVC) Draft Text 7. document JCTVC-R1008 v7. Oct. 2014. url: http://phenix.intevry.fr/jct/doc_end_user/current_document.php?id=9465..

Abstract: In recent years, video entertainment demand has significantly changed. Video content is transmitted through different bandwidth connections and played on many devices that have different processing capabilities and screen sizes. For this reason, scalable extensions of video coding standards have been released, e.g SHVC, scalable extension of HEVC. Beside high compression efficiency, SHVC has disadvantages including high encoder complexity and weakness in error resilience. These are not suitable for emerging applications such as wireless sensor networks, video surveillance systems or remote sensing that have limited processing capabilities, low energy and low network bandwidth. A potential solution

supporting for these systems is Distributed Scalable Video Coding (DSVC). In DSVC system, Side Information (SI) creation plays a critical role in deciding system overall performance. Therefore, this paper proposes a spatially scalable DSVC architecture and a new side information creation technique for this DSVC system. Results show that the proposed method generates better quality SI when compared to some previous SI creation methods. Consequently, the system performance is improved when compared with the previous methods.

Keyword: DVC, Wyner-Ziv Coding



Nguyễn Thị Hương Thảo, Nhận bằng tốt nghiệp đại học và thạc sỹ Học viện Công nghệ Bưu chính Viễn thông vào các năm 2003 và 2010. Hiện giảng dạy và làm Nghiên cứu sinh tại Khoa Kỹ thuật Điện tử 1 - Học viện Công nghệ Bưu chính Viễn thông. Lĩnh vực nghiên cứu: Xử lý tín hiệu Video, Xử lý Ảnh, Lý thuyết thông tin.



Vũ Văn San, Nhận học vị Tiến sỹ năm 2000 tại Viện Điện tử Viễn thông, Hàn quốc. Hiện TS. Vũ Văn San công tác tại Học viện Công nghệ Bưu chính Viễn thông. Lĩnh vực nghiên cứu: Truyền dẫn và xử lý tín hiệu số.