

MÔ HÌNH NHIỀU TƯƠNG QUAN CHO HỆ THỐNG MÃ HÓA VIDEO PHÂN TÁN

Nguyễn Thị Hương Thảo, Vũ Văn San

Học viện Công nghệ Bưu chính Viễn thông

Tóm tắt: Mã hóa video phân tán (DVC) là sơ đồ mã hóa video mới phù hợp cho các ứng dụng đường lên như hệ thống giám sát video không dây, mạng cảm biến không dây. Các kết quả nghiên cứu trước đã chỉ ra rằng mặc dù thực hiện mã hóa độc lập và giải mã kết hợp, hiệu năng của hệ thống mã hóa video phân tán có thể đạt mức tương đương so với các hệ thống mã hóa video dự đoán hiện nay. Tuy nhiên, hiệu năng méo – tốc độ (RD) của các hệ thống mã hóa video phân tán này phụ thuộc rất lớn vào việc mô hình hóa nhiều tương quan giữa thông tin gốc ở phía mã hóa và thông tin phụ trợ tương ứng ở phía giải mã. Trong các nghiên cứu trước, hầu hết đều sử dụng mô hình Laplacian để mô hình hóa nhiều tương quan mà không tính đến đặc tính thống kê và đặc tính chuyển động của chuỗi. Bài báo này đề xuất phương pháp mô hình hóa nhiều tương quan mới thay đổi thích ứng cho các hệ số DC dựa vào đặc tính của chuỗi. Các kết quả thực nghiệm cho thấy hiệu năng của phương pháp đề xuất đã được cải thiện hơn so với mô hình Laplacian trước đây.

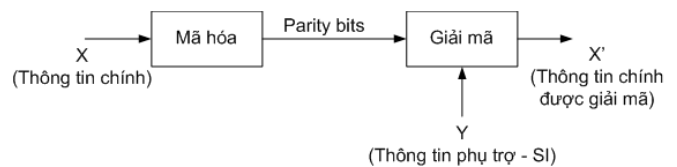
Từ khóa: Video phân tán, nhiều tương quan.

I. GIỚI THIỆU

Trong các hệ thống truyền thông và đa phương tiện hiện nay, các kỹ thuật nén video số đóng một vai trò vô cùng quan trọng vì sự hạn chế của độ rộng băng tần. Ngoài các kỹ thuật mã hóa video truyền thống sử dụng mã hóa dự đoán và biến đổi cosine rời rạc còn có một kỹ thuật mã hóa video khác được gọi là mã hóa video phân tán. Điểm khác biệt chính giữa hai phương pháp này là nơi thực hiện khai thác tương quan thời gian, điều đó dẫn đến sự khác nhau về độ phức tạp của bộ mã hóa và bộ giải mã. Kiến trúc mã hóa video dự đoán hướng đến các ứng dụng video mà ở đó video được mã hóa một lần và được giải mã nhiều lần, ví dụ truyền hình quảng bá. Vì vậy, bộ mã hóa có độ phức tạp cao hơn từ 5 đến 10 lần so với bộ giải mã. Tuy nhiên kiến trúc này lại không phù hợp với các ứng dụng mới ví dụ như mạng giám sát video không dây, mạng cảm biến không dây vì các ứng dụng này có rất nhiều bộ mã hóa trong khi chỉ có một vài bộ giải mã. Giải pháp cho tình huống này là sử dụng mã hóa video phân tán. DVC thực hiện khai thác tương quan thời gian, một phần hoặc toàn phần tại phía giải mã, do đó làm giảm độ phức tạp cho bộ mã hóa. Nói cách khác, DVC thực hiện mã hóa độc lập và giải mã kết hợp. Điều này giúp dịch chuyển bớt độ phức tạp từ phía mã hóa sang phía giải mã trong khi hiệu suất nén vẫn giữ tương đương so với kiến trúc mã hóa video dự đoán truyền thống.

Định lý Slepian-Wolf [1] và định lý Wyner-Ziv [2] từ lý thuyết thông tin đã phát biểu rằng đối với nén không tổn thất và nén có tổn thất, có thể mã hóa độc lập và giải mã kết hợp cho các nguồn có tương quan với nhau mà vẫn giữ được tốc độ tương đương với trường hợp mã hóa và giải mã kết hợp như trong mã hóa video truyền thống.

Dựa trên các kết quả lý thuyết của các định lý Slepian-Wolf và Wyner-Ziv, các kiến trúc DVC thực tế đã được đề xuất trong [3,4] bởi nhóm nghiên cứu nhóm của giáo sư Bernd Girod tại Đại học Stanford hay còn gọi là kiến trúc Stanford mà sau này được cải tiến thành codec DISCOVER [3] và nhóm của giáo sư Kannan Ramchandran tại Berkeley (Đại học California) còn được biết đến là kiến trúc PRISM [4]. Hình 1 mô tả phương pháp nén video phân tán với thông tin phụ trợ được tạo ra tại phía phát. Trong kiến trúc Stanford, chuỗi video được chia thành các khung hình chính (KF) và các khung hình Wyner-Ziv (WZF). Tại phía mã hóa, các KF sẽ được mã hóa kỹ thuật mã hóa video truyền thống như H.264/AVC Intra hoặc HEVC Intra. Các WZF được biến đổi cosine rời rạc (DCT) và lượng tử hóa, sau đó áp dụng mã hóa kênh để tạo ra các bit kiểm tra. Tuy nhiên, chỉ các bit kiểm tra này được gửi tới bên thu tùy theo yêu cầu còn các bit hệ thống bị loại bỏ nhằm hạn chế số lượng bit cần gửi đi. Tại phía giải mã, các KF đã mã hóa sẽ được giải mã. Các khung hình này sẽ được sử dụng như các khung hình tham chiếu để tạo ra khung hình thông tin phụ trợ (SI), một phiên bản ‘nhiều’ của khung hình WZ gốc. Bộ giải mã kênh sẽ sử dụng các bit kiểm tra được gửi tới để ‘sửa sai’ cho các SI để thu được các khung hình WZ ban đầu. Như vậy, để số lượng bit kiểm tra phải gửi tới phía thu càng ít thì ngoài việc tạo ra SI có chất lượng tốt, điều quan trọng là phải dự đoán đúng mô hình nhiều tương quan giữa khung hình WZ gốc và SI được tạo ra tại phía giải mã. Tuy nhiên, đây là một công việc rất phức tạp vì thông tin phụ trợ chỉ có tại bộ giải mã và chất lượng SI thay đổi theo chuỗi và thay đổi trong bản thân mỗi khung hình. Nói cách khác, nhiều tương quan không chỉ dừng theo thời gian mà còn dừng theo không gian. Khi chuỗi chuyển động nhanh, rất khó để dự đoán khung hình WZ và sai lỗi trong SI tăng lên đáng kể.



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

Trong hầu hết các nghiên cứu về DVC đều lựa chọn mô hình Laplacian để mô tả nhiều tương quan. Tuy nhiên qua quan sát thực nghiệm cho thấy không phải lúc nào mô hình Laplacian cũng chính xác. Vì vậy bài báo này nghiên cứu một mô hình lai ghép kết hợp mô hình Laplacian và mô hình Gaussian để mô tả nhiều tương quan. Việc lựa chọn mô hình nào tùy thuộc vào nội dung của khung hình tương ứng.

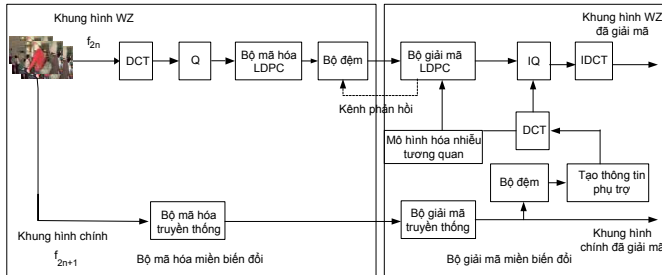
Tác giả liên hệ: Nguyễn Thị Hương Thảo

Email: thaonh@ptit.edu.vn

Đến tòa soạn: 10/2018, chỉnh sửa: 12/2018, chấp nhận đăng: 12/2018.

Phần tiếp theo của bài báo được tổ chức như sau. Phần II giới thiệu về kiến trúc DVC được sử dụng trong bài báo. Mô hình nhiều tương quan được đề xuất trong Phần III. Các kết quả và thảo luận được giới thiệu trong phần IV và phần V là kết luận.

II. KIẾN TRÚC MÃ HÓA VIDEO WYNER-ZIV MIỀN BIẾN ĐỔI



Hình 2. Kiến trúc bộ mã hóa DVC miền biến đổi

Hình 2 mô tả kiến trúc bộ mã hóa DVC miền biến đổi.

A. Quá trình mã hóa

Chia tách khung hình: Các khung hình của chuỗi video 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.

DCT: Để khai thác dư thừa không gian, các khung hình WZ sẽ được biến đổi DCT 4×4 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 đồng nhất với các hệ số lượng tử tùy theo chất lượng mong muốn. Các ký tự sau lượng tử hóa được chia thành các mặt phẳng bit, sau đó được mã hóa độc lập sử dụng bộ mã hóa kiểm tra chẵn lẻ mật độ thấp (LDPC). Các bit chẵn lẻ sẽ được lưu lại trong bộ đệm và được gửi (tùy theo yêu cầu của bộ giải mã) tới bộ giải mã trong khi các bit hệ thống bị loại bỏ.

B. Quá trình giải mã

Các khung hình chính được giải mã bởi bộ giải mã truyền thống. 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 thông tin phụ trợ: Thông tin phụ trợ được coi như một phiên bản nhiễu của thông tin gốc. Nó được tạo ra dựa trên các thông tin đã giải mã từ bộ đệm gửi tới. Trong kiến trúc DISCOVER, SI được tạo ra sử dụng kỹ thuật nội suy thời gian bù chuyển động (MCTI). Kiến trúc MCTI có thể tóm tắt như sau:

1) **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.

2) **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.

3) **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.

4) **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.

5) **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.

DCT: Thông tin phụ trợ được biến đổi DCT khối 4×4 để nhận được các hệ số DCT nguyên.

Mô hình hóa nhiều 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 nhiều 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 giải mã LDPC.

Bộ giải mã LDPC: Bộ giải mã LDPC thực hiện sửa các sai lỗi trong thông tin phụ trợ sử dụng các bit chẵn lẻ từ phía mã hóa gửi tới. Thông thường, đại lượng tin cậy dựa trên tỉ số xác suất tiên nghiệm được sử dụng làm tiêu chí phát hiện sai để xác định xác suất sai lỗi của mặt phẳng bit. Các mặt phẳng bit quan trọng sẽ có ngưỡng xác suất sai lỗi thấp hơn so với các mặt phẳng bit ít quan trọng. Nếu chưa đạt được xác suất lỗi cho phép, bộ giải mã sẽ gửi yêu cầu đến bộ mã hóa qua kênh phản hồi để gửi thêm các bit chẵn lẻ.

Giải lượng tử và biến đổi ngược IDCT: Sau khi giải mã LDPC, thông tin được tái tạo bằng cách giải lượng tử và biến đổi DCT ngược. Toàn bộ khung hình được khôi phục lại ở miền pixel.

III. MÔ HÌNH NHIỀU TƯƠNG QUAN TRONG MÃ HÓA VIDEO PHÂN TÁN

Để tận dụng tốt nhất thông tin phụ trợ có được, bộ giải mã cần phải có hiểu biết tin cậy về mô hình mô tả nhiều tương quan giữa khung hình WZ gốc và khung hình thông tin phụ trợ tương ứng. Nhiều tương quan ($WZ - SI$) có thể được coi là kênh ảo với mẫu lỗi được đặc trưng bởi phân bố thống kê nào đó bởi vì SI có thể được coi như phiên bản 'nhiều' của thông tin gốc. Trong kiến trúc codec miền biến đổi, sự sai khác này chính là sự sai khác giữa các dải DCT tương ứng của khung hình WZ và SI. Nếu mô hình mô tả chính xác ($WZ - SI$), hiệu suất mã hóa sẽ tăng lên, ngược lại, sự tổn thất về hiệu suất mã hóa sẽ xảy ra. Mục đích của phần này là giới thiệu ngắn gọn về mô hình nhiều tương quan được sử dụng trong

các nghiên cứu mã hóa video phân tán và đề xuất phương pháp mới nhằm cải tiến mô hình nhiễu tương quan.

A. Các nghiên cứu về mô hình nhiễu tương quan

Nếu gọi WZ là thông tin gốc và SI là thông tin phụ trợ thì sự sai khác giữa WZ và SI được biểu thị là $N=WZ-SI$ gọi là nhiễu tương quan. Nhiễu tương quan có thể được mô hình hóa bằng cách sử dụng phân bố Laplacian [5,6] hoặc phân bố Gaussian. Tuy nhiên, phân bố Laplacian được sử dụng rộng rãi hơn vì có sự cân bằng tốt giữa độ chính xác mô hình và độ phức tạp. Sử dụng phân bố Laplacian, mối tương quan giữa WZ và SI được mô tả như sau:

$$p(WZ-SI) = \frac{\alpha}{2} e^{-\alpha(WZ-SI)} \quad (1)$$

Ở đó, $p(.)$ là hàm mật độ xác suất và α là tham số phân

$$\text{bố Laplacian được xác định bởi công thức: } \alpha = \sqrt{\frac{2}{\sigma^2}} \quad (2)$$

ở đó σ^2 là phương sai của sự sai khác giữa WZ và SI . α có thể thay đổi theo thời gian và không gian. Có rất nhiều các nghiên cứu khác nhau nhằm ước lượng giá trị α này [7, 8, 9, 10].

Tùy thuộc vào nơi thực hiện, ước lượng nhiễu tương quan có thể phân loại thành ba nhóm: ước lượng nhiễu tương quan tại phía mã hóa (ECNE), ước lượng nhiễu tương quan tại phía giải mã (DCNE) và ước lượng nhiễu tương quan hai phía (HCNE) nghĩa là nhiễu tương quan được ước lượng tại cả hai phía mã hóa và giải mã.

Ước lượng nhiễu tương quan tại phía mã hóa

Thực hiện ước lượng nhiễu tương quan tại phía mã hóa cho phép điều khiển tốc độ tại bộ mã hóa. Các tham số nhiễu tương quan có thể tính được dựa vào khung hình gốc và bản sao của thông tin phụ trợ. Tuy nhiên, do ước lượng chuyển động để tạo thông tin phụ trợ được thực hiện tại phía mã hóa nên độ phức tạp bộ mã hóa sẽ tăng. Mặt khác, các tham số nhiễu tương quan được tính toán tại bộ mã hóa phải được gửi tới bộ giải mã để hỗ trợ quá trình giải mã. Trong [11,12], nhiễu tương quan được tính trong miền tần số tại bộ mã hóa. Module tìm kiếm chuyển động được sử dụng để tìm ra thông tin phụ trợ tốt nhất và tương quan giữa khối hiện thời và thông tin phụ trợ này được tính toán. Thông tin chỉ báo về thông tin phụ trợ và số lượng các bit ít quan trọng nhất (các bit này được rút ra từ thông tin phụ trợ) được gửi tới bộ giải mã để giải mã thành công cho khối hiện thời. Brites và Pereira [5] đề xuất các phương pháp ước lượng nhiễu tương quan tại phía mã hóa trong cả hai miền pixel và miền biến đổi. Mỗi khối của thông tin phụ trợ được tạo ra bằng cách lấy trung bình của các khối tham chiếu trước và sau được bù chuyển động. Vì vậy, sự tương đồng giữa các khối này được sử dụng để ước lượng nhiễu tương quan giữa dữ liệu gốc và thông tin phụ trợ. Trong miền pixel, có ba mức được đề xuất là mức khung hình, mức khối và mức pixel. Đối với miền biến đổi, dải DCT và hệ số là hai mức đề xuất. Các giá trị α được tính toán cho mỗi mức ở cả hai miền. Năm 2008, Martinez và các cộng sự [13] đề xuất kiến trúc mã hóa video Wyner-Ziv miền pixel ở đó nhiễu

tương quan được tính bằng cách sử dụng ước lượng thông tin phụ trợ nhanh và khung hình WZ hiện tại. Từ khung hình sai khác này, một số tham số thống kê tin cậy được trích xuất để ước lượng chất lượng của thông tin phụ trợ tại phía giải mã, từ đó ước lượng số lượng bit cần thiết để sửa các lỗi trong thông tin phụ trợ. Năm 2010, Sheng và các cộng sự trong [14] đề xuất một phương pháp ước lượng nhiễu tương quan tại phía mã hóa dựa trên phân bố Laplacian. Để tránh làm tăng độ phức tạp bộ mã hóa và sử dụng kênh phản hồi, giá trị α của phân bố Laplacian được tính là sai số bình phương trung bình (MSE) giữa khung hình WZ hiện thời và khung hình tham chiếu.

Ước lượng nhiễu tương quan tại bộ giải mã

Để giữa cho bộ mã hóa có độ phức tạp thấp thì quá trình ước lượng và bù chuyển động không nên thực hiện tại bộ mã hóa. Vì vậy, trong hầu hết các hệ thống mã hóa video Wyner-Ziv, các module ước lượng nhiễu tương quan thường tồn tại bên phía giải mã. Nhiễu thuật toán ước lượng nhiễu [5,10,15] đã được đề xuất sử dụng phân bố Laplacian cho các hệ số DCT. Các giá trị α khác nhau trong công thức (1) được ước lượng cho các mức khác nhau là mức khung hình, mức dải và mức hệ số. Khác với thuật toán trong [5], các tác giả trong [10] ước lượng nhiễu miền biến đổi bằng cách chuyển đổi các ước lượng nhiễu trong miền pixel. Các tham số nhiễu miền pixel được ước lượng sử dụng thông tin sẵn có từ khung hình Wyner-Ziv được giải mã trước đó cũng như các dải hệ số đã được giải mã trước. Một cách ước lượng nhiễu tương quan miền biến đổi cũng được đề xuất bởi Huan và Forchhammer trong [15] bằng cách sử dụng tương quan chéo dải. Dựa trên các quan sát và phân bố thống kê của nhiễu tương đối với các chuỗi có chuyển động khác nhau và với các dải hệ số DC, AC khác nhau, mô hình nhiễu tương quan thích ứng được đề xuất trong [16,17]. Bằng cách sử dụng thông tin hỗ trợ từ phía mã hóa gửi tới, mô hình nhiễu tương quan trong [16] được xây dựng. Sau đó, bộ giải mã lựa chọn thích ứng phân bố nhiễu tương quan Laplacian hoặc Gaussian cho các hệ số DC dựa trên nội dung chuỗi video. Một cách tiếp cận khác là sử dụng mô hình nhiễu tương quan hỗn hợp [17]. Hai phân bố khác nhau được sử dụng cho các hệ số AC và DC tùy thuộc vào sai số khác nhau của các hệ số DC và AC trong khung hình thông tin phụ trợ và đặc tính phân bố của nhiễu tương quan ở mức lượng tử khác nhau. Huynh Van Luong và các cộng sự đã đề xuất các mô hình nhiễu trong [18,19]. Trong [18], kỹ thuật học nhiễu được đề xuất để tận dụng các dư thừa của các khung hình được giải mã trước đó. Để tạo ra ước lượng nhiễu tương quan chính xác hơn, kỹ thuật bù chuyển động cho dư thừa được đề xuất trong [19] sử dụng thông tin từ các khung hình được giải mã trước đó và tương quan giữa khung hình trước và khung hình thông tin phụ trợ được ước lượng hiện thời.

Ước lượng nhiễu tương quan tại hai phía

Một cách tiếp cận khác là thực hiện ước lượng nhiễu tương quan tại cả phía mã hóa và giải mã [20,21]. Ước lượng nhiễu tương quan trong [20] là cách tiếp cận đối xứng ở đó cả bộ mã hóa và giải mã sử dụng cùng một mô hình tương quan. Vì vậy, cách tiếp cận này chỉ khai thác thông tin giải mã sẵn có ở cả hai phía mã hóa và giải mã. Vì vậy, mặc dù tránh được sự ước

lượng không giống nhau ở hai phía nhưng độ phức tạp bộ mã hóa trong trường hợp này lại tăng vì quá trình ước lượng và bù chuyển động để tạo ra thông tin phụ trợ cần phải thực hiện tại phía mã hóa. Mặt khác, hiệu năng méo – tốc độ (RD) có thể bị giảm đi vì chỉ tận dụng được thông tin đã giải mã. Để khắc phục nhược điểm này, mô hình hóa nhiều tương quan bất đối xứng được đề xuất trong [21] trong đó bộ mã hóa sẽ sử dụng giải pháp tạo thông tin phụ trợ độ phức tạp thấp và bộ giải mã sử dụng giải pháp tạo thông tin phụ trợ độ phức tạp cao hơn. Tuy nhiên, cách tiếp cận này có thể dẫn đến các kết quả khác nhau tại bộ mã hóa và giải mã. Vì vậy cần phải thực hiện một số các kỹ thuật phụ thêm để loại bỏ sự ước lượng sai của mô hình tương quan.

Với các phân tích ở trên, có thể thấy có rất nhiều cách tiếp cận đối với mô hình nhiều tương quan.

Bài báo này tập trung vào điều chỉnh mô hình Laplacian hay Gaussian tùy thuộc vào nội dung của chuỗi video. Mô hình nhiều tương quan sẽ được thực hiện tại bộ giải mã - giải pháp thực tế nhất cho các kiến trúc mã hóa video phân tán.

B. Mô hình nhiều tương quan đề xuất

Khi thử với rất nhiều khung hình của các chuỗi video khác nhau, kết quả cho thấy phân bố Laplacian này không hoàn toàn phù hợp. Cụ thể là, với các hệ số AC thì nhiều tương quan tuân theo phân bố Laplace khá chính xác nhưng với hệ số DC, khi chuỗi chuyển động phức tạp thì nhiều này tuân theo phân bố Laplace nhưng khi chuỗi chuyển động chậm thì lại phù hợp với phân bố Gaussian hơn. Vì vậy, để cải thiện tính chính xác của mô hình nhiều tương quan trực tuyến, đề tài này đề xuất một thuật toán xây dựng mô hình nhiều tương quan thích ứng cho hệ thống mã hóa video miền biến đổi. Thuật toán đề xuất sử dụng hai loại phân bố Laplacian và Gaussian cho hệ số DC tùy thuộc vào nội dung chuỗi video.

Bước 1: Trước tiên tính khung hình dư thừa R giữa khung hình WZ và khung hình SI tương ứng bằng cách xấp xỉ hiệu giữa các phiên bản bù chuyển động của các khung hình chính trước và sau X_B và X_F sử dụng công thức dưới đây:

$$R(x, y) = \frac{X_F(x+dx_f, y+dy_f) - X_B(x+dx_b, y+dy_b)}{2} \quad (3)$$

ở đó $X_F(x+dx_f, y+dy_f)$ và $X_B(x+dx_b, y+dy_b)$ là các khung hình chính trước và sau được bù chuyển động tương ứng và (x, y) là vị trí pixel trong khung hình dư thừa R . (dx_b, dy_b) và (dx_f, dy_f) mô tả vector chuyển động của các khung hình X_B và X_F tương ứng.

Bước 2: Biến đổi DCT cho khung hình R bằng cách áp dụng biến đổi cosine rời rạc 8×8 cho khung hình R để nhận được các hệ số DCT của khung hình T .

$$T(u, v) = DCT[R(x, y)] \quad (4)$$

Bước 3:

Đối với mỗi hệ số của biến đổi DCT sẽ áp dụng các mô hình nhiều khác nhau tùy thuộc vào giá trị của chúng như biểu thức (5) dưới đây:

$$f(n) = \begin{cases} \sqrt{\frac{2}{\sigma^2}} e^{-\sqrt{\frac{2}{\sigma^2}}|n|} & \text{với tất cả các hệ số AC} \\ \sqrt{\frac{2}{\sigma^2}} e^{-\sqrt{\frac{2}{\sigma^2}}|n|} & \text{với các hệ số DC } n < \gamma \\ \sqrt{\frac{2}{\sigma^2}} e^{-\sqrt{\frac{2}{\sigma^2}}n^2} & \text{với các hệ số DC } n \geq \gamma \end{cases} \quad (5)$$

đó $f(n)$ là hàm phân bố xác suất của nhiều tương quan. μ và σ^2 là trung bình và phương sai của nhiều tương quan tương ứng. Giá trị ngưỡng γ thực nghiệm được chọn là 0.02.

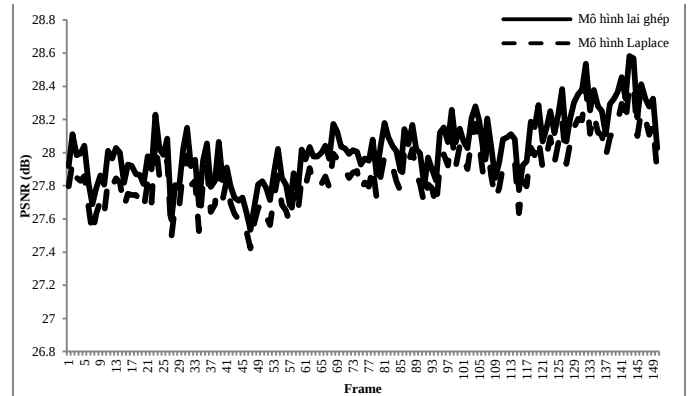
IV. ĐIỀU KIỆN THỬ NGHIỆM VÀ PHÂN TÍCH KẾT QUẢ

A. Điều kiện thử nghiệm

Để đánh giá hiệu năng của giải pháp đề xuất so với các phương pháp khác, ba chuỗi video sau được sử dụng là Aikyo, Foreman và Carphone với các đặc điểm được mô tả trong Bảng I. Cấu trúc GOP được sử dụng trong hệ thống là “KWKW...” ở đó K là khung hình chính và W là khung hình WZ tương ứng. Khung hình chính được mã hóa theo chế độ intramode của phần mềm HM. Các thông số của chuỗi được mô tả trong Bảng I.

Bảng I. ĐIỀU KIỆN THỬ NGHIỆM

Chuỗi Video thử nghiệm	Độ phân giải không gian	Độ phân giải thời gian	Số lượng khung hình
Aikyo	176 x 144	50 Hz	150
Foreman		60 Hz	150
Carphone		50 Hz	150



Hình 3. PSNR của chuỗi Aikyo

B. Phân tích kết quả

Để đánh giá kết quả của thuật toán, tham số PSNR (tỷ số tín hiệu/nhiều đỉnh trung bình) của khung hình WZ giải mã được sử dụng khi áp dụng hai phương pháp: mô hình Laplacian thuần túy và mô hình nhiều tương quan thích ứng.

Bảng II. PSNR CỦA CÁC KHUNG HÌNH WZ GIẢI MÃ (dB)

Chuỗi video	Mô hình Laplacian	Mô hình đề xuất
Aikyo	28.0	27.8
Foreman	28.8	29.5
Carphone	29.2	29.4

Kết quả của các chuỗi được trình bày trong Bảng II. Bảng II mô tả sự so sánh PSNR trung bình của các khung hình WZ giải mã trong hai trường hợp: sử dụng mô hình Laplacian để mô hình hóa nhiễu tương quan và khi sử dụng mô hình lai ghép đề xuất. Minh họa hình ảnh của chuỗi Aikyo được trình bày trong Hình 3.

Các kết quả cho thấy phương pháp đề xuất cho mức độ cải thiện trung bình lên tới 0.7 dB cho chuỗi Foreman và 0.2 dB cho chuỗi Carphone tuy nhiên lại giảm 0.2 dB đối với chuỗi Aikyo. Điều này cho thấy kết quả đạt được khá tốt với chuỗi có nhiều chuyển động nhưng lại chưa hiệu quả đối với chuỗi ít chuyển động.

V. KẾT LUẬN

Bài báo này giới thiệu mô hình nhiễu tương quan cho mã hóa video Wyner-Ziv miền biến đổi. Bằng cách sử dụng thông tin phụ được gửi từ bộ mã hóa, thuộc tính thống kê của nhiễu tương quan và đặc tính chuyển động của chuỗi video, thuật toán đề xuất có thể được xây dựng dựa trên phân bố Laplacian hoặc Gaussian.

Các kết quả thực nghiệm so sánh phương pháp đề xuất với các phương pháp sử dụng phân bố Laplacian thông thường cho thấy phương pháp đề xuất có thể cải thiện đáng kể chất lượng của khung hình giải mã với độ phức tạp tăng lên không đáng kể. Trong các nghiên cứu tiếp theo sẽ tập trung áp dụng các kỹ thuật trí tuệ nhân tạo vào xây dựng mô hình nhiễu tương quan cho kiến trúc DVC nhằm cải thiện hơn nữa hiệu năng tổng thể của hệ thống.

TÀI LIỆU THAM KHẢO

- [1] D. Slepian and J. Wolf, *Noiseless Coding of Correlated Information Sources*, IEEE Transactions on Information Theory, vol. 19, no. 4, pp.471-480, July 1973.
- [2] A. Wyner and J. Ziv, *The Rate-Distortion Function for Source Coding with Side Information at the Decoder*, IEEE Transactions on Information Theory, vol. 22, no. 1, pp.1-10, January 1976.
- [3] X. Artigas, J. Ascenso, M. Dalai, S. Klomp, D. Kubasov, and M. Ouaret, "The DISCOVER codec: Architecture, techniques and evaluation" in Proc. Picture Coding Symp., Lisbon, Portugal, Oct. 2007, pp. 1-5.
- [4] 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.
- [5] 18. Catarina Brites, Fernando Pereira, *Correlation noise modeling for efficient pixel and transform domain Wyner-Ziv video coding*, [IEEE Transactions on Circuits and Systems for Video Technology](#), Volume: 18, Issue: 9, Sept. 2008.
- [6] B. Girod, A. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed Video Coding," Proceedings of the IEEE, vol. 93, no. 1, pp. 71-83, January 2005.
- [7] Catarina Brites, Joao Ascenso, Fernando Pereira, *Studying temporal correlation noise modeling for pixel based Wyner-Ziv video coding*, [Image Processing, 2006 IEEE International Conference on](#), ICIP 2006.
- [8] Catarina Brites, Fernando Pereira, *Correlation noise modeling for multiview transform domain Wyner-Ziv video coding*, [Image Processing \(ICIP\), 2014 IEEE International Conference on](#).
- [9] Xiem Hoang Van, Joao Ascenso, Fernando Pereira, *Adaptive scalable video coding: a HEVC based framework combining the predictive and distributed paradigms*, [IEEE Transactions on Circuits and Systems for Video Technology](#), Volume: 27, Issue: 8, Aug. 2017.
- [10] Jürgen Slowack, Jozef Škorupa, Stefaan Mys, Nikos Deligiannis, Peter Lambert, Adrian Munteanu, and Rik Van de Walle (2011). *Correlation Noise Estimation in Distributed Video Coding*. Effective Video Coding for Multimedia Applications, pp. 133- 156, Intech Publishing, 2011, ISBN 978- 953- 307- 177- 0.
- [11] S. Minali and G. Calvagno, "A distributed video coder based on the H.264/AVC standard," in EUSIPCO, Poznan, Poland, Sep. 2007.
- [12] Minali, J. Wang, and K. Ramchandran, "Achieving H.264-like compression efficiency with distributed video coding," in SPIE VCIP, San Jose, CA, USA, Jan. 2007
- [13] J. L. Martínez, G. Fernández-Escribano, H. Kalva, W. A. R. J. Weerakkody, W. A. C. Fernando, and A. Garrido, "Feedback free DVC architecture using machine learning," in Proc. IEEE ICIP, Oct. 2008, pp. 1140-1143.
- [14] T. Sheng, X. Zhu, G. Hua, H. Guo, J. Zhou, and C. W. Chen, "Feedback free rate-allocation scheme for transform domain Wyner-Ziv video coding," *Multimedia Syst.*, vol. 16, no. 2, pp. 127-137, 2010.
- [15] X. Huang and S. Forchhammer, "Cross-band noise model refinement for transform domain Wyner-Ziv video coding," *Signal Process., Image Commun.*, vol. 27, no. 1, pp. 16-30, 2012.
- [16] Hao Qin, Bin Song, Yue Zhao, and Haihua Liu, *Adaptive Correlation Noise Model for DC Coefficients in Wyner-Ziv Video Coding*, ETRI Journal, Volume 34, Number 2, April 2012
- [17] Hu Xiaofei, Zhu Xiuchang, "A Wyner-Ziv video coding method utilizing mixture correlation noise model", *Journal of electronics (China)*, Vol.29, No.3/4, pp. 197-203, July 2012.
- [18] Huynh Van Luong, Lars Lau Raket, Xin Huang, and Soren Forchhammer, *Side information and noise learning for distributed video coding using optical flow and clustering*, [IEEE Transactions on Image Processing](#) (Volume: 21, Issue: 12, Dec. 2012)
- [19] Huynh Van Luong, Lars Lau Raket, Xin Huang, and Soren Forchhammer, *Re-estimation of motion and reconstruction for distributed video coding*, [IEEE Transactions on Image Processing](#) (Volume: 23, Issue 7, July. 2014).
- [20] Xiem Hoang Van, Joao Ascenso, Fernando Pereira, *Adaptive scalable video coding: a HEVC based framework combining the predictive and distributed paradigms*, [IEEE Transactions on](#)

[Circuits and Systems for Video Technology](#), Volume: 27, [Issue: 8](#), Aug. 2017.

- [21] X. HoangVan et al., “*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.

Abstract: Distributed video coding is a new paradigm which is suitable for uplink applications such as wireless sensor networks, video surveillance systems. Previous research results have shown that despite of independent encoding and joint decoding, distributed video coding can achieve equivalent performance to predictive video coding. However, the Rate - Distortion(RD) performance of these distributed video encoding systems depends greatly on correlation noise modeling between the original information and corresponding side information at the decoder. In previous works, most use Laplacian distribution to model correlation noise and don't take into account statistical property of the transform domain correlation noise and the motion characteristic of the frame. This paper proposes a new method in which models for the DC coefficients are adaptively adjusted depending on the motion characteristics of sequence. The experimental results show that the performance of the proposed method has been improved compared to the previous Laplacian model.

Keyword: DVC, Wyner-Ziv Coding, correlation noise model



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.