

BỘ THÔNG TIN VÀ TRUYỀN THÔNG
HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



NGUYỄN THỊ HƯƠNG THẢO

NGHIÊN CỨU CẢI TIẾN HIỆU NĂNG HỆ THỐNG
MÃ HÓA/GIẢI MÃ VIDEO PHÂN TÁN

CHUYÊN NGÀNH: KỸ THUẬT ĐIỆN TỬ

MÃ SỐ: 9.52.02.03

TÓM TẮT LUẬN ÁN TIẾN SĨ KỸ THUẬT

HÀ NỘI - 2020

BỘ THÔNG TIN VÀ TRUYỀN THÔNG
HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



NGUYỄN THỊ HƯƠNG THẢO

NGHIÊN CỨU CẢI TIẾN HIỆU NĂNG HỆ THỐNG
MÃ HÓA/GIẢI MÃ VIDEO PHÂN TÁN

CHUYÊN NGÀNH: KỸ THUẬT ĐIỆN TỬ

MÃ SỐ: 9.52.02.03

TÓM TẮT LUẬN ÁN TIẾN SĨ KỸ THUẬT

HÀ NỘI - 2020

Công trình được hoàn thành tại:

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Người hướng dẫn khoa học:

- 1. PGS.TS. Vũ Văn San**
- 2. TS. Nguyễn Ngọc Minh**

Phản biện 1:

Phản biện 2:

Phản biện 3:

Luận án sẽ được bảo vệ trước hội đồng chấm luận văn tại:

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

vào hồi: **giờ** , ngày **tháng** năm

Có thể tìm hiểu luận án tại:

- 1. Thư viện Quốc gia Việt Nam**
- 2. Thư viện Học viện Công nghệ Bưu chính Viễn thông**

MỞ ĐẦU

Lý do nghiên cứu

Trong nhiều ứng dụng hiện nay, video đóng một vai trò hết sức quan trọng. Trong các nghiên cứu về video, nén là một trong những khía cạnh được quan tâm nhiều nhất. Mục đích của nén video là tạo ra một cách biểu diễn dữ liệu video sao cho chiếm ít dung lượng nhất.

Tuy nhiên, nén cũng đi đôi với những trả giá. Thứ nhất, nén càng nhiều sẽ càng làm giảm chất lượng hình ảnh. Thứ hai, độ phức tạp của các thuật toán nén sẽ tăng lên khi muốn tỷ lệ nén tốt hơn.

Trong các chuẩn mã hóa video truyền thống, bộ mã hóa thường phức tạp hơn rất nhiều so với bộ giải mã. Cách thiết kế này thường phù hợp với các ứng dụng video truyền thống như truyền hình quảng bá.

Với sự phát triển mạnh mẽ của mạng không dây, xuất hiện hàng loạt các ứng dụng mới mà ở đó không tuân thủ theo sơ đồ một - nhiều như các ứng dụng trước đó. Đối với các ứng dụng mới này, thường sẽ có nhiều bộ mã hóa gửi dữ liệu tới một vài bộ giải mã trung tâm. Vì vậy thách thức trong các hệ thống này là làm sao có thể thực hiện nén video trên các thiết bị có tài nguyên hạn chế, có khả năng chống lại hiện tượng mất gói, hiệu suất nén cao và độ trễ thấp.

Một giải pháp thay thế có thể đáp ứng các yêu cầu này là mã hóa video phân tán (DVC). Trong mã hóa video phân tán, ước lượng chuyển động và do đó là độ phức tạp của nó được chuyển từ bộ mã hóa sang bộ giải mã. Phải nhấn mạnh rằng mục đích của DVC không phải là thay thế cho H.264/AVC hay HEVC bằng cách cung cấp khả năng nén tốt hơn. Trên thực tế, DVC tốt nhất được dự kiến là có thể thực hiện với hiệu năng nén tương đương. Tuy nhiên, các kết quả thực tế cho thấy DVC vẫn chưa đạt được hiệu suất mong muốn như các chuẩn nén video dự đoán hiện nay, đặc biệt trong bối cảnh giữ cho bộ mã hóa có độ phức tạp thấp.

Mục tiêu nghiên cứu

Các phân tích ở trên cho thấy rằng DVC đóng một vai trò quan trọng đối với các ứng dụng mới như mạng cảm biến hình ảnh hay mạng giám sát không dây. Tuy nhiên thách thức chính đối với DVC là phải đạt được hiệu năng nén ở

mức tiệm cận với các hệ thống mã hóa video dự đoán hiện có. Vì vậy mục tiêu nghiên cứu của luận án là nghiên cứu đề xuất kỹ thuật cải thiện hiệu năng nén của mã hóa video phân tán DVC và mã hóa video phân tán liên lớp DSVC.

Nội dung nghiên cứu

Với mục tiêu nghiên cứu ở trên, phần nội dung nghiên cứu của Luận án tập trung vào các nội dung sau:

- Đề xuất các phương pháp mới nhằm cải thiện hiệu năng nén cho mã hóa video phân tán bao gồm phương pháp được thực hiện tại phía mã hóa và nhóm các phương pháp thực hiện tại phía giải mã.

- Đề xuất các kỹ thuật tạo thông tin phụ trợ mới để cải thiện hiệu năng nén cho mã hóa video phân tán liên lớp.

Đối tượng và phạm vi nghiên cứu

Đối tượng nghiên cứu là mã hóa video phân tán và mở rộng của mã hóa video phân tán là mã hóa video phân tán liên lớp. Phạm vi nghiên cứu của luận án là các kỹ thuật nhằm nâng cao hiệu năng nén của mã hóa video phân tán và mã hóa video phân tán liên lớp.

Ý nghĩa khoa học và thực tiễn của đề tài nghiên cứu

Ý nghĩa khoa học của luận án mà nghiên cứu sinh hướng đến là đề xuất các kỹ thuật mới nhằm nâng cao hiệu năng nén của mã hóa video phân tán và mã hóa video phân tán liên lớp.

Ý nghĩa thực tiễn của luận án mà nghiên cứu sinh hy vọng đạt được là những kết quả nghiên cứu của luận án góp phần vào việc rút ngắn khoảng cách về hiệu suất mã hóa của mã hóa video phân tán với hệ thống mã hóa video dự đoán hiện nay, sớm đưa mã hóa video phân tán vào các ứng dụng thực tế.

Phương pháp nghiên cứu

Để thực hiện được các nội dung nghiên cứu đề ra, phương pháp nghiên cứu được sử dụng trong Luận án này bao gồm:

Phương pháp nghiên cứu lý luận: Phân tích, tổng hợp và đánh giá các công trình nghiên cứu, sách chuyên khảo và các nguồn tài liệu khoa học có liên quan đến mã hóa video phân tán và mã hóa video phân tán liên lớp. Từ đó đề xuất các kỹ thuật mới nhằm cải thiện hiệu năng nén cho mã hóa video phân tán và mã hóa video phân tán liên lớp.

Phương pháp mô phỏng: Các đề xuất cải thiện hiệu năng nén cho mã hóa video phân tán và mã hóa video phân tán liên lớp được thử nghiệm và đánh giá so sánh với các giải pháp trước đó trên các phần mềm tham chiếu.

Cấu trúc của luận án

Ngoài phần mở đầu và phần kết luận, kiến nghị, Luận án được chia thành 3 chương với bố cục như sau:

Chương 1 giới thiệu tổng quan về mã hóa video bao gồm mã hóa video dự đoán và mã hóa video phân tán, trong đó tập trung vào các vấn đề liên quan đến mã hóa video phân tán. Chương 2 trình bày các đề xuất cải thiện hiệu năng nén cho mã hóa video phân tán tại phía bộ mã hóa và bộ giải mã [1], [2], [3]. Chương 3 luận án trình bày các đề xuất cải thiện hiệu năng nén cho mã hóa video phân tán liên lớp [4],[5].

CHƯƠNG 1

TỔNG QUAN VỀ MÃ HÓA VIDEO

Tóm tắt: *Nội dung của chương trình bày tổng quan về mã hóa video bao gồm mã hóa video dự đoán và mã hóa video phân tán (DVC). Trong chương này giới thiệu một số các chuẩn nén video dự đoán và lý thuyết nền tảng của DVC cũng như các bộ mã hóa video phân tán thực tế được nghiên cứu trong cộng đồng nghiên cứu DVC. Phần cuối của chương 1 đề cập đến các vấn đề đang được nghiên cứu trong DVC cũng như các hạn chế của các nghiên cứu trước đây và đề xuất hướng nghiên cứu, phạm vi nghiên cứu và phương thức tiếp cận của luận án.*

1.1. Mã hóa video dự đoán

Phần này trình bày tổng quan về các chuẩn mã hóa video dự đoán bao gồm các chuẩn mã hóa video H.261, MPEG-1, H.262/MPEG-2, H.263 và H.264/AVC, HEVC.

1.2. Mã hóa video phân tán

Phần này giới thiệu hai định lý trong lý thuyết thông tin là cơ sở của mã hóa nguồn phân tán và mã hóa video phân tán là định lý Slepian-Wolf và định lý Wyner-Ziv.

1.3. Các kiến trúc mã hóa video phân tán

Phần này giới thiệu kiến trúc cơ bản của mã hóa video phân tán và ba kiến trúc DVC thực tế được đề xuất và sử dụng trong các nghiên cứu về DVC gồm kiến trúc Stanford của đại học Stanford, kiến trúc PRISM của đại học California và kiến trúc DISCOVER là cải tiến của kiến trúc Stanford..

1.4. Các vấn đề đang được nghiên cứu trong mã hóa video phân tán

Phần này giới thiệu về các vấn đề đang được nghiên cứu đối với DVC bao gồm: cải thiện hiệu năng nén, cung cấp khả năng chống lỗi, cung cấp khả năng liên lớp, mã hóa video phân tán đa góc nhìn. Các nghiên cứu cải thiện hiệu năng nén tập trung vào cải thiện bộ mã hóa Slepian-Wolf, tạo thông tin phụ trợ, mô hình nhiễu tương quan, lựa chọn chế độ mã hóa, tái tạo.

Khảo sát cho thấy hầu hết các nghiên cứu tập trung vào cải thiện hiệu năng nén cho mã hóa video phân tán nhưng hiệu năng nén của các hệ thống này vẫn còn một khoảng cách xa với mã hóa video dự đoán đang được sử dụng hiện nay. Với nhu cầu về các hệ thống mã hóa video liên lớp, nên ngoài cải thiện hiệu năng nén, một số hướng nghiên cứu cũng tập trung vào mã hóa video liên lớp sử dụng nguyên tắc phân tán. Tuy nhiên các nghiên cứu này chưa nhiều và hiệu năng nén của mã hóa video phân tán liên lớp còn chưa cao.

Nhận thấy các vấn đề cũng như các tồn tại trong các nghiên cứu DVC, trong luận án này đề xuất giải pháp nhằm cải thiện hiệu năng nén cho các hệ thống mã hóa video phân tán.

1.5. Các hướng nghiên cứu trong luận án

Với mục tiêu cải thiện hiệu năng nén cho mã hóa video phân tán trong khi vẫn duy trì độ phức tạp thấp cho bộ mã hóa, nghiên cứu sinh tập trung vào các vấn đề sau:

- Nghiên cứu đề xuất phương pháp thực hiện tại phía mã hóa để cải thiện hiệu năng nén của mã hóa video phân tán nhưng không làm tăng đáng kể độ phức tạp mã hóa.

- Nghiên cứu đề xuất các phương pháp có thể phức tạp hơn thực hiện tại phía giải mã để cải thiện hiệu năng nén cho mã hóa video phân tán.

- Nghiên cứu đề xuất các phương pháp cải thiện hiệu năng nén cho mở rộng của mã hóa video phân tán là mã hóa video phân tán liên lớp.

1.6. Kết luận chương

Chương này giới thiệu về mã hóa video nói chung, và đặc biệt là cung cấp nền tảng lý thuyết và các kết quả nghiên cứu về DVC. Dựa trên những khảo sát cũng như đánh giá sơ bộ về tình hình nghiên cứu DVC hiện nay, nghiên cứu sinh đã đưa ra hướng nghiên cứu cụ thể nhằm cải thiện hiệu năng nén cho mã hóa video phân tán và mã hóa video phân tán liên lớp.

CHƯƠNG 2

CẢI TIẾN HIỆU NĂNG NÉN CHO MÃ HÓA VIDEO PHÂN TÁN

Tóm tắt: *Chương 2 trình bày các kết quả nghiên cứu về cải tiến hiệu năng nén cho mã hóa video phân tán bao gồm: (1) cải tiến hiệu năng nén cho DVC tại phía mã hóa sử dụng phương pháp thay đổi kích thước nhóm ảnh; (2) cải tiến hiệu năng nén cho DVC tại phía giải mã với mô-đun cải tiến phương pháp tạo thông tin phụ trợ và mô-đun ước lượng tham số α cho mô hình nhiễu tương quan.*

2.1. Giới thiệu chung

Với mục tiêu cải tiến hơn nữa hiệu năng nén của DVC, chương này nghiên cứu và đề xuất một phương pháp thực hiện tại phía mã hóa và hai phương pháp được thực hiện tại phía giải mã nhằm cải tiến hiệu năng nén của DVC.

2.2. Đề xuất 1: Cải tiến hiệu năng nén cho DVC tại phía mã hóa

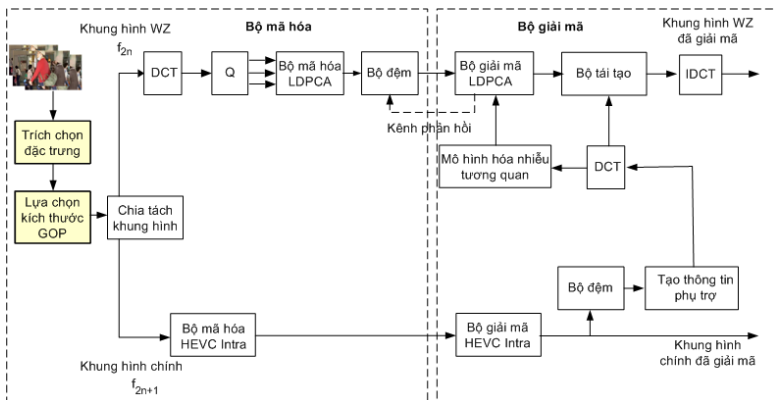
2.2.1. Đặt vấn đề

Phần này giới thiệu mối tương quan giữa nội dung chuyển động của chuỗi video với kích thước nhóm ảnh (GOP). Việc lựa chọn kích thước GOP cũng quyết định đến chất lượng ước lượng thông tin phụ trợ (SI) và mô hình nhiễu tương quan (CNM). Do đó việc thay đổi kích thước GOP theo nội dung chuyển động của chuỗi video có thể cải thiện hiệu năng nén của hệ thống mã hóa video phân tán.

Tuy nhiên, việc lựa chọn kích thước GOP chính xác cho mỗi phân đoạn video khó đạt được độ chính xác. Vì vậy, luận án đề xuất một phương pháp xác định kích thước GOP cho mỗi phân đoạn video dựa trên thuật toán cây quyết định.

2.2.2. Phương pháp thay đổi kích thước GOP đề xuất

Phương pháp thay đổi kích thước GOP đề xuất được thực hiện tại bộ mã hóa DVC. Kiến trúc mã hóa video phân tán AGOP-DVC với khối đề xuất mới được tô màu đậm được mô tả trong hình 2.1 dưới đây.



Hình 2. 1: Kiến trúc mã hóa video phân tán AGOP-DVC

Trong đề xuất này, mỗi phân đoạn gồm 5 khung hình video đầu vào được đánh giá để lựa chọn kích thước GOP phù hợp bằng cách trích chọn các đặc trưng. Các đặc trưng này được đưa vào cây quyết định để phân loại mỗi phân

đoạn này là chuyển động nhanh hay chuyển động chậm để lựa chọn kích thước GOP phù hợp. Để tiết kiệm thời gian xử lý, quá trình huấn luyện được thực hiện ngoại tuyến trước khi sử dụng cho quá trình mã hóa

1) Xác định đặc trưng

Như đề cập ở trên, các đặc trưng được lựa chọn phải phản ánh đầy đủ bản chất của nội dung video. Vì vậy, một số đặc trưng được lựa chọn liên quan đến chuyển động trong khi một số khác lại liên quan đến kết cấu.

2) Quá trình huấn luyện

Tập dữ liệu sử dụng cho quá trình huấn luyện của đề xuất này được trích xuất từ năm chuỗi video *Foreman*, *Soccer*, *Hall Monitor*, *Akiyo*, *Carphone* và *News* với các đặc tính chuyển động đa dạng. Tập dữ liệu gồm 352 phân đoạn video, mỗi phân đoạn gồm 5 khung hình.

Nhãn GOP2 và GOP4 tương ứng cho mỗi phân đoạn được xác định như sau:

- Với mỗi phân đoạn 5 khung hình thực hiện mã hóa DVC với hai kích thước GOP=2 và GOP=4.

- Tính BD-PSNR tương ứng với hai cách mã hóa.

- Kích thước GOP nào tương ứng với BD-PSNR lớn hơn sẽ được chọn là nhãn tương ứng cho phân đoạn đó.

Các đặc trưng cùng với nhãn tương ứng của 352 phân đoạn được đưa vào để huấn luyện bằng thuật toán tiêu biểu của cây quyết định là C4.5.

3) Áp dụng vào quá trình lựa chọn kích thước GOP

Sau quá trình huấn luyện, mô hình cây quyết định xây dựng ở trên sẽ được sử dụng để lựa chọn kích thước GOP. Chuỗi video đầu vào sẽ được chia thành các phân đoạn gồm 5 khung hình và các đặc trưng tương ứng với mỗi phân đoạn được trích xuất. Dựa trên các đặc trưng này, sử dụng mô hình cây quyết định đã được huấn luyện để chọn ra nhãn tương ứng (GOP2 hay GOP4) cho mỗi phân đoạn. Sau khi lựa chọn được kích thước GOP phù hợp sẽ thực hiện chia tách các khung hình thành khung hình WZ (WZF) và khung hình chính KF tùy theo kích thước GOP.

2.2.3. Đánh giá hiệu năng của phương pháp đề xuất AGOP-DVC

Phương pháp thay đổi kích thước GOP trong mục 2.2.2 được đánh giá trong ngữ cảnh áp dụng vào kiến trúc mã hóa video phân tán AGOP-DVC như trong Hình 2.1. Hiệu năng được đánh giá ở đây là hiệu năng méo – tốc độ bit (RD) của AGOP-DVC với hai tham số PSNR (dB) và Bitrate (bps).

Bộ mã hóa tham chiếu được sử dụng để so sánh ở đây dựa trên kiến trúc DISCOVER. Đây là bộ mã hóa video phân tán DISCOVER với KF được mã hóa bằng chuẩn mã hóa HEVC. Bộ mã hóa này khi sử dụng kích thước cố định GOP=2 và GOP=4 có tên gọi lần lượt là GOP2-DVC và GOP4-DVC. Hai bộ mã hóa này sẽ được so sánh với bộ mã hóa đề xuất AGOP-DVC trên khía cạnh hiệu năng RD.

A. Điều kiện thử nghiệm

Việc đánh giá hiệu năng nén được thực hiện trên bốn chuỗi video thử nghiệm *Coastguard*, *Suzie*, *Pamphlet* và *Harbour*. Các chuỗi này được lựa chọn vì tính đa dạng về chuyển động và kết cấu trong mỗi chuỗi.

B. Đánh giá hiệu năng RD

Hiệu năng RD được đo bằng tốc độ bit (bps) và PSNR (dB) đối với bốn chuỗi video. Như đã dự đoán trước, các kết quả cho thấy các giá trị PSNR của phương pháp đề xuất AGOP-DVC tốt hơn các giá trị PSNR của GOP4-DVC và xấp xỉ các giá trị PSNR của GOP2-DVC. Các giá trị bitrate của phương pháp đề xuất cao hơn các giá trị của GOP4-DVC và thấp hơn các giá trị của GOP2-DVC. Mức tiết kiệm tốc độ bit của phương pháp đề xuất trung bình là 3,37% và 9,62% khi so với GOP2-DVC và GOP4-DVC.

2.3. Đề xuất 2: Cải tiến hiệu năng nén cho DVC tại phía giải mã

2.3.1. Đặt vấn đề

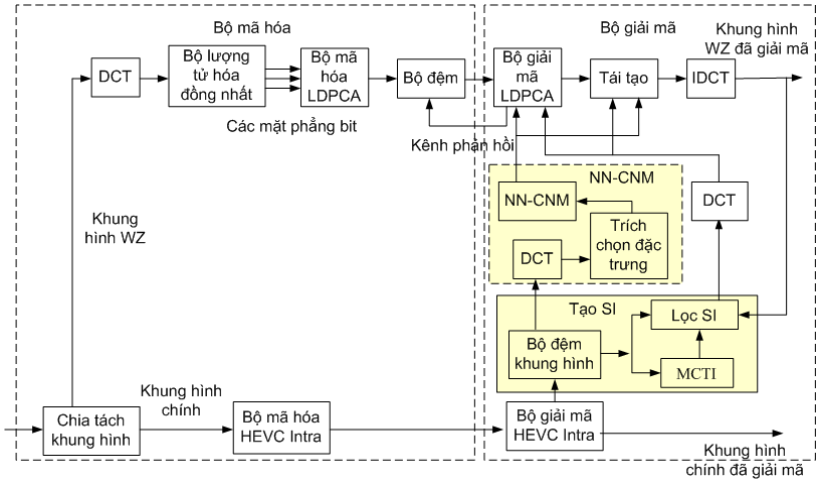
Trong DVC, SIF tương ứng với WZF gốc được tạo ra tại phía giải mã bằng cách sử dụng các KF đã giải mã. Vì vậy, hiệu năng nén của DVC phụ thuộc rất nhiều các mô-đun tại bộ giải mã như mô-đun tạo SI và mô hình nhiễu tương quan.

Luận án này đề xuất mới một phương pháp tạo SI bằng cách lọc liên tục SI trên các khung hình đã giải mã trước đó và chính SI được tạo ra ban đầu và một

phương pháp ước lượng được tham số α của mô hình hóa nhiễu tương quan sử dụng mạng nơ-ron.

2.3.2. Kiến trúc bộ mã hóa đề xuất Adv-DVC

Sơ đồ kiến trúc mã hóa DVC đề xuất (Adv-DVC) với hai công cụ mã hóa mới được tô đậm là tạo thông tin phụ trợ và mô hình hóa nhiễu tương quan được mô tả trong Hình 2.2.



Hình 2.2: Kiến trúc mã hóa DVC đề xuất Adv-DVC

Nguyên lý hoạt động của kiến trúc này cũng tương tự như kiến trúc DVC được giới thiệu trong Hình 2.1 với một số thay đổi như sau: 1) kích thước GOP=2; 2) Mô-đun Tạo SI tạo thông tin phụ trợ bằng phương pháp lọc liên tục trong quá trình giải mã; và 3) Mô-đun NN-CNM ước lượng tham số của CNM dựa trên mạng nơ-ron.

2.3.3 Các mô-đun mã hóa đề xuất mới

Trong phần này giới thiệu hai mô-đun mã hóa mới thực hiện tại bộ giải mã và được sử dụng trong kiến trúc DVC đề xuất bao gồm:

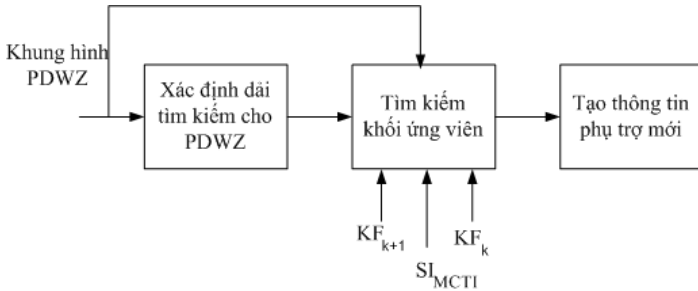
- Tạo thông tin phụ trợ bằng cách lọc liên tục trong suốt quá trình giải mã.
- Ước lượng tham số α cho mô hình nhiễu tương quan dựa trên mạng nơ-ron.

A. Tạo thông tin phụ trợ bằng cách lọc liên tục

Trong mô-đun tạo SI này, SIF ban đầu được tạo ra sử dụng kỹ thuật MCTI sử dụng các KF đã giải mã trước đó gồm KF_{k+1}, KF_{k-1} . SIF ban đầu này được gọi là SI_{MCTI} .

Mỗi khi một mặt phẳng bit được giải mã LDPCA thành công, chúng được đưa tới mô-đun tái tạo để khôi phục lại hệ số DCT gốc với sự trợ giúp của SIF tương ứng. Các hệ số DCT đã tái tạo sẽ thay thế cho các hệ số cùng vị trí trong SIF trước đó để tạo ra SIF mới được gọi là WZF được giải mã một phần (PDWZ).

Quá trình tạo SI bằng cách lọc liên tục được mô tả trong Hình 2.3 dưới đây.



Hình 2.3 Quá trình tạo SI

1) Xác định dải tìm kiếm cho PDWZ

Sau khi giải mã xong một mặt phẳng bit, khung hình PDWZ mới được tạo ra. Khung hình PDWZ này được ước lượng chuyển động trong các vùng tìm kiếm khác nhau tùy thuộc vào nội dung chuyển động của khung hình đó. Trong đề xuất này, mức độ chuyển động của khung hình được đánh giá dựa trên tham số biên độ vector chuyển động trung bình. Tùy thuộc vào giá trị biên độ chuyển động trung bình này, dải tìm kiếm (SR) được xác định cho mỗi PDWZ phụ thuộc vào điều kiện mà nó thỏa mãn.

2) Tìm kiếm khối ứng viên

Khung hình PDWZ sẽ được chia thành các khối kích thước 4×4 và việc tìm kiếm ứng viên sẽ được thực hiện cho từng khối trên ba khung hình tham chiếu (RF) gồm KF_{k+1} , KF_{k-1} và SI_{MCTI} . Mỗi khối 4×4 trong khung hình PDWZ sẽ so sánh với các khối 4×4 trong dải tìm kiếm tương ứng trong ba

khung hình tham chiếu bằng cách tính tổng các trị tuyệt đối sai số (MAD). Khối nào trong khung hình tham chiếu tương ứng với MAD nhỏ nhất sẽ được coi là giống với khối trong khung hình PDWZ nhất. Các giá trị MAD tương ứng với các khung hình tham chiếu là $\alpha_n(KF_{k+1})$, $\alpha_n(KF_{k-1})$ và bSI_{MCTI} . Các ứng viên này được đưa vào bước kết hợp để tạo ra khối SI mới.

3) Kết hợp các khối ứng viên

Sau bước 2) ở trên nhận được ba khối ứng viên trong ba khung hình tham chiếu tương ứng với ba giá trị MAD. Theo lẽ tự nhiên, sai số đối sánh càng nhỏ chứng tỏ khối đó có khả năng là một ứng viên tốt và có thể đóng vai trò lớn hơn trong việc tạo SI cuối cùng và ngược lại. Vì vậy, trọng số $\beta_n(RF) = \frac{1}{\alpha_n(RF)}$ được sử dụng để xác định đóng góp của mỗi khối ứng viên trong cơ chế kết hợp. SIF mới được tạo ra lại được sử dụng để giải mã cho mặt phẳng bit kế tiếp. Khi tất cả các dải hệ số DCT được giải mã, SI cuối cùng sẽ được tạo ra để thực hiện tái tạo lần cuối cùng để nhận được WZF gốc.

B. Ước lượng tham số α cho mô hình nhiễu tương quan

Tham số α của mô hình nhiễu tương quan nếu được ước lượng đúng phải sử dụng WZF gốc tại phía mã hóa và SIF tại phía giải mã. Tuy nhiên, tại phía giải mã không sẵn có WZF nên nhiễu tương quan được xấp xỉ bằng cách sử dụng các KF đã giải mã và được bù chuyển động.

Giá trị α được ước lượng ở mức dải DCT, do đó các đặc trưng được trích xuất trong miền DCT.

1) Xác định đặc trưng

Biến đổi khung hình dư thừa $R_k(x, y)$ sang miền DCT. Các hệ số DCT của khung hình dư thừa được nhóm thành 16 dải chứa các hệ số ở cùng vị trí của mỗi khối.

Trong mỗi dải, bốn đặc trưng *Min*, *Max*, *Trung bình*, *Phương sai* được tính cho mỗi dải.

2) Xây dựng tập dữ liệu huấn luyện

Để xây dựng tập dữ liệu huấn luyện cho mạng nơ-ron, luận án sử dụng 10 chuỗi video *Coastguard*, *Hall-Monitor*, *News*, *Container*, *Flower Garden*, *Mobile*, *Mother*, *Claire*, *Grandma*, *Harbour* với độ phân giải không gian 176×144 (QCIF) và số lượng các khung hình cho mỗi chuỗi là 300. Lý do để lựa chọn các chuỗi video này để huấn luyện là sự đa dạng của các đặc tính chuyển động của chuỗi. Các chuỗi này được mã hóa và giải mã HEVC Intra với bốn tham số lượng tử và xây dựng được 1490 khung hình dư thừa.

Đối với mỗi dải DCT sẽ có 1490 mẫu tương ứng, mỗi mẫu có 4 đặc trưng. Đầu ra tương ứng, hay giá trị α tham chiếu, được gán cho mỗi mẫu được tính dựa trên khung hình dư thừa đúng.

3) Kiến trúc mạng nơ-ron

Để huấn luyện cho tập dữ liệu xây dựng ở trên, trong nghiên cứu này sử dụng một mạng nơ-ron nhiều lớp. Để quyết định cấu trúc mạng nơ-ron với số lớp ẩn và số nơ-ron trong mỗi lớp, nhiều cấu trúc mạng nơ-ron khác nhau đã được thử nghiệm. Dựa trên độ chính xác của các mô hình, cấu trúc mạng nơ-ron $\{4,4,2,1\}$ được chọn. Mạng nơ-ron này gồm có một lớp đầu vào, hai lớp ẩn và một lớp đầu ra với tên gọi NN-CNM.

Lớp đầu vào với bốn giá trị X_1, X_2, X_3, X_4 tương ứng là bốn đặc trưng Min, Max, Trung bình và Phương sai của các dải DCT của khung hình dư thừa. Tất cả các lớp trong mạng được kết nối đầy đủ. Trong các lớp ẩn 1 và 2, hàm kích hoạt ReLU được sử dụng. Tại lớp đầu ra, hàm kích hoạt tuyến tính được sử dụng để dự đoán ra giá trị $\hat{\alpha}$ cho dải hệ số DCT.

4) Huấn luyện mô hình

Tập dữ liệu chuẩn bị ở trên được chia thành tập huấn luyện và tập kiểm tra theo tỷ lệ 70/30 để huấn luyện cho mạng nơ-ron 2 lớp ẩn. Kết quả của quá trình huấn luyện là tập các trọng số tương ứng.

5) Áp dụng mô hình đã huấn luyện cho bộ mã hóa DVC

Tại phía bộ giải mã DVC, các KF đã giải mã được sử dụng để tạo ra khung hình nhiều tương quan. Các đặc trưng tương ứng cho mỗi dải DCT của khung

hình này được trích xuất. Sử dụng tập các trọng số của mô hình huấn luyện để áp dụng cho tập đặc trưng này để dự đoán giá trị đầu ra \hat{x} .

2.3.4 Đánh giá hiệu năng

Phần này giới thiệu và phân tích các kết quả về hiệu năng nén của các công cụ mã hóa đề xuất so với các kiến trúc mã hóa video có liên quan.

A. Điều kiện thử nghiệm

Đánh giá hiệu năng được thực hiện trên bốn chuỗi video *Akiyo*, *Carphone*, *Foreman* và *Soccer*.

Trong phần đánh giá này sử dụng các bộ mã hóa sau: 1) *HEVC Intra*: Bộ mã hóa này sử dụng phần mềm tham chiếu HM với chế độ mã hóa Intra. 2) *DISCOVER-HEVC*: Đây là kiến trúc DISCOVER với khung hình chính được mã hóa bởi HEVC Intra thay vì H.264/AVC Intra. 3) *ReSI-DVC*: Bộ mã hóa này là bộ mã hóa DISCOVER-HEVC với mô-đun tạo SI cải tiến được đề xuất trong phần A của mục 2.3.3. và 4) *Adv-DVC*: Đây là bộ mã hóa được xây dựng trên DISCOVER-HEVC với mô-đun tạo SI cải tiến và mô-đun NN-CNM mới như mô tả trong hình 2.2.

Phần tiếp theo sẽ thực hiện ba đánh giá: 1) đánh giá chất lượng SI được tạo ra theo phương pháp lọc liên tục, 2) độ chính xác của phương pháp ước lượng tham số α đề xuất và 3) đánh giá về hiệu năng nén tổng thể của kiến trúc DVC đề xuất khi thực hiện đồng thời cả hai cải tiến về tạo SI và CNM so với các bộ mã hóa khác có liên quan.

B. Chất lượng của SI theo phương pháp đề xuất

Phần này đánh giá chất lượng của SI theo phương pháp lọc đề xuất bằng cách so sánh với giải pháp tạo SI được sử dụng nhiều trong kiến trúc DVC là MCTI. Từ các kết quả có thể rút ra một số nhận xét như sau:

- Phương pháp tạo SI đề xuất ReSI đạt được chất lượng SI tốt hơn so với MCTI cho tất cả các tham số lượng tử của các chuỗi video thử nghiệm.
- Mức cải thiện không đồng đều giữa các chuỗi có nội dung chuyển động khác nhau. Mức cải thiện lớn nhất đạt được đối với chuỗi chuyển động chậm *Akiyo* và mức cải thiện ít hơn đối với các chuỗi chuyển động nhanh như *Soccer*.

Đó là bởi vì với các chuỗi chuyển động nhanh, khó tạo ra SI chất lượng tốt vì sự sai khác giữa các KF là rất lớn so với trường hợp chuỗi chuyển động chậm.

- Mặc dù phương pháp đề xuất cho chất lượng SI tốt hơn so với MCTI cho cả bốn tham số lượng tử nhưng kết quả thường tốt hơn đối với các QP thấp. Đó là bởi vì với QP thấp, các KF thường có chất lượng tốt hơn và điều đó giúp tạo ra các SI có chất lượng tốt hơn.

C. Độ chính xác của phương pháp ước lượng tham số α đề xuất

Trong tiểu mục này, tham số α được ước lượng bởi phương pháp NN-CNM được so sánh với tham số α được tính như trong DISCOVER-HEVC. Quy ước tên gọi các phương pháp như sau: 1) *NN-CNM*: giá trị α được ước lượng bằng mạng nơ-ron; 2) *DISCOVER-HEVC*: giá trị α được ước lượng như trong kiến trúc DISCOVER-HEVC. 3) *Oracle*: giá trị α tham chiếu hay giá trị α được tính dựa trên nhiễu tương quan thực.

Nếu tham số ước lượng nào (NN-CNM hoặc DISCOVER-HEVC) càng gần giá trị tham chiếu (Oracle), việc ước lượng được coi là càng chính xác.

Các kết quả cho thấy NN-CNM gần với Oracle hơn so với DISCOVER-HEVC, đặc biệt với các chuỗi video chuyển động chậm như *Akiyo* và *Carphone*. Điều này cho thấy phương pháp ước lượng tham số α đề xuất dựa trên mạng nơ-ron đã cải thiện tính chính xác cho CNM so với phương pháp trước đó được sử dụng trong DISCOVER.

D. Đánh giá chất lượng của khung hình giải mã

Trong phần này, chất lượng của khung hình giải mã được tạo ra bởi bộ mã hóa đề xuất Adv-DVC được so sánh với khung hình giải mã của các bộ mã hóa khác. Từ các kết quả có thể rút ra một số nhận xét như sau:

- *Adv-DVC* với các bộ mã hóa DVC khác:

Các bộ mã hóa DVC khác ở đây bao gồm DISCOVER-HEVC và ReSI-DVC. Kiến trúc đề xuất Adv-DVC đạt được kết quả tốt hơn so với tất cả các bộ mã hóa DVC khác cho tất cả các chuỗi video thử nghiệm. So với DISCOVER-HEVC, PSNR của Adv-DVC có mức cải thiện lên đến 3,55 dB, và mức cải thiện ít hơn khi gặp chuỗi chuyển động nhanh và phức tạp như *Soccer*. So với ReSI-DVC, Adv-DVC cũng đạt được những mức cải thiện đáng kể.

- *Adv-DVC* với *HEVC Intra*:

HEVC Intra được sử dụng để so sánh vì đây là bộ mã hóa video dự đoán có độ phức tạp thấp. Các kết quả cho thấy Adv-DVC đạt được giá trị PSNR cao hơn so với HEVC Intra cho hầu hết các chuỗi ngoại trừ chuỗi *Carphone*. Các mức cải thiện đối với các chuỗi chuyển động chậm và nhanh cũng khác nhau. Đối với các chuỗi chuyển động chậm, ví dụ *Akiyo*, mức cải thiện PSNR có thể lên đến 1,37 dB nhưng kết quả không được tốt đối với các chuỗi chuyển động nhanh *Carphone*. Lý do là vì chuỗi *Carphone* được coi là chuỗi chuyển động phức tạp với sự thay đổi bất thường về nội dung. Với chế độ mã hóa Intra của HEVC, do không khai thác tương quan thời gian nên chất lượng của khung hình giải mã có chất lượng tốt hơn.

Tổng thể, kiến trúc mã hóa đề xuất Adv-DVC cho chất lượng khung hình giải mã tốt hơn so với cả ba bộ mã hóa còn lại ReSI-DVC, DISCOVER-HEVC và HEVC-Intra.

E. Hiệu năng nén của bộ mã hóa DVC đề xuất Adv-DVC

Từ các kết quả có thể rút ra một số nhận xét sau:

- *Adv-DVC so với HEVC Intra*: Hiệu năng RD của Adv-DVC tốt hơn so với HEVC Intra cho hầu hết các chuỗi ngoại trừ chuỗi chuyển động phức tạp *Soccer*. Đối với chuỗi chuyển động chậm, Adv-DVC cho kết quả tốt hơn bởi vì SI có chất lượng tốt và CNM chính xác. Tổng thể, Adv-DVC đạt được mức tiết kiệm tốc độ bit 22,01 % và mức cải thiện BD-PSNR 2,47 dB.

- *Adv-DVC với các bộ mã hóa DVC khác*: Hiệu năng RD của Adv-DVC được cải thiện đáng kể so với các bộ mã hóa DVC khác đối với tất cả các chuỗi video thử nghiệm. Mức cải thiện đối với các chuỗi chuyển động chậm cao hơn so với các chuỗi chuyển động phức tạp. Tổng thể, Adv-DVC đạt được mức giảm tốc độ bit trung bình là 35,27 % khi so với DISCOVER-HEVC và 21,03 % khi so sánh với ReSI-DVC.

2.4. Kết luận chương

Chương 2 đã trình bày kết quả nghiên cứu mới bao gồm giải pháp lựa chọn chế độ mã hóa được thực hiện tại phía mã hóa và nhóm giải pháp được thực hiện tại phía giải mã bao gồm cải thiện chất lượng của thông tin phụ trợ và cải thiện tính chính xác của mô hình nhiễu tương quan.

Các đóng góp được công bố trong các bài báo [1], [2], [3] tập trung vào hai nội dung chính:

- Phương pháp thay đổi kích thước nhóm ảnh được thực hiện tại phía mã hóa;
- Nhóm giải pháp được thực hiện tại phía giải mã bao gồm tạo SI bằng cách lọc liên tục và ước lượng tham số α cho mô hình hóa nhiễu tương quan dựa trên mạng nơ-ron.

CHƯƠNG 3

CẢI TIẾN HIỆU NĂNG NÉN CHO MÃ HÓA

VIDEO PHÂN TÁN LIÊN LỚP

Tóm tắt: Trong chương này, Luận án trình bày các kết quả nghiên cứu về cải tiến hiệu năng nén cho mã hóa video phân tán liên lớp bao gồm: (1) tạo thông tin phụ trợ dựa trên kỹ thuật kết hợp ảnh; và (2) tạo thông tin phụ trợ dựa trên SVM.

3.1. Giới thiệu chung

Phần này trình bày tổng quan về khái niệm và vai trò của mã hóa video liên lớp.

3.1.1 Mã hóa video liên lớp

Phần này trình bày về nguyên tắc cơ bản của mã hóa video liên lớp đang được sử dụng trong các chuẩn mã hóa video dự đoán.

3.1.2 Mã hóa video phân tán liên lớp

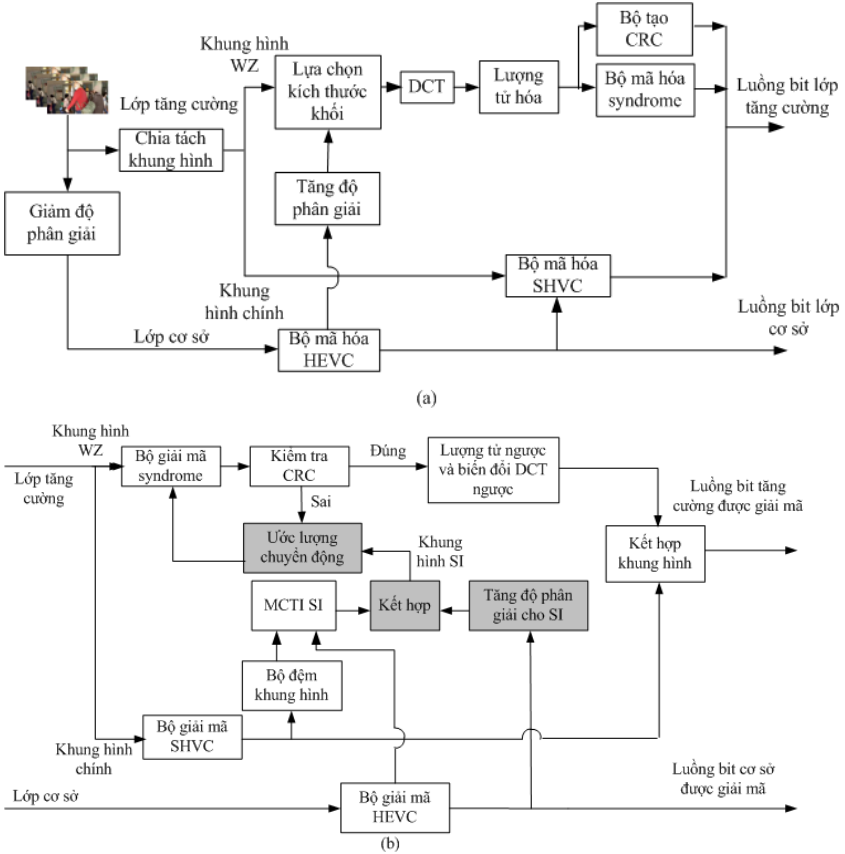
Phần này trình bày về khái niệm và khảo sát một số nghiên cứu liên quan đến mã hóa video phân tán liên lớp (DSVC).

3.2. Đề xuất 3: Tạo thông tin phụ trợ dựa trên kỹ thuật kết hợp ảnh cho DSVC

3.2.1 Kiến trúc bộ mã hóa video phân tán liên lớp đề xuất với tên gọi DSVC-Fusion

Kiến trúc bộ mã hóa video phân tán liên lớp đề xuất với tên gọi DSVC-Fusion được mô tả trong Hình 3.1. Kiến trúc này có hai lớp gồm một lớp BL và một lớp EL. Bộ mã hóa DSVC ở đây cung cấp khả năng mở rộng không gian.

Vì vậy mỗi khung hình trong chuỗi video gốc được mã hóa hai lần ở lớp BL và lớp EL. Đối với lớp BL, trước khi đưa vào bộ mã hóa HEVC Intra, khung hình gốc sẽ được giảm độ phân giải không gian.



Hình 3.1: Sơ đồ kiến trúc bộ mã hóa và giải mã video phân tán liên lớp DSVC-Fusion

3.2.2 Phương pháp tạo thông tin phụ trợ dựa trên kỹ thuật kết hợp ảnh

Đối với bộ mã hóa DSVC-Fusion, khối thông tin phụ trợ (SIB) có thể được tạo ra tại bộ giải mã bằng cách sử dụng thêm thông tin từ khung hình độ phân giải thấp ở lớp BL thay vì chỉ sử dụng KF giải mã. Nói cách khác, trong trường hợp này cả tương quan thời gian và tương quan không gian được sử dụng để

ước lượng SIB. Điều này có thể giúp cải thiện chất lượng của SIB trong bộ mã hóa DSVC.

A. Tạo các ứng viên thông tin phụ trợ

Trong bộ mã hóa DSVC-Fusion này, có hai ứng viên SIB được tạo ra qua khai thác thông tin tương quan thời gian và tương quan không gian. Bằng cách khai thác tương quan thời gian, SI thứ nhất được tạo ra sử dụng kỹ thuật nội suy thời gian bù chuyển động (MCTI) được gọi là SIB_{MCTI} . Khai thác tương quan không gian, SI thứ hai được tạo ra bằng cách tăng độ phân giải của khung hình lớp BL đã giải mã với tên gọi là SIB_{UpBL} .

B. Cải thiện chất lượng cho SIB_{MCTI}

Ứng viên SIB_{MCTI} được cải thiện chất lượng bằng cách sử dụng SIB_{UpBL} như một khung hình tham chiếu. Trong phương pháp này, SIB_{UpBL} được chọn làm khung hình tham chiếu vì hai lý do sau: (1) chất lượng của SIB_{UpBL} cao hơn so với khung hình MCTI và (2) khung hình SIB_{UpBL} sẵn có tại bộ giải mã.

Sau khi khung hình MCTI được tạo ra từ các KF trước và sau đã bù chuyển động song hướng, các pixel của ba khung hình được so sánh với các pixel tương ứng trong khung hình đã tăng độ phân giải. Pixel ở khối nào có sai khác nhỏ nhất sẽ được coi là giá trị pixel mới của khung hình MCTI.

C. Kết hợp các SIB ứng viên

Sau khi cải thiện chất lượng của SIB_{MCTI} , khối mới SIB_{New_MCTI} này sẽ được kết hợp với SIB_{UpBL} để tạo ra khối mới SIB_{Fusion} .

3.2.3 Đánh giá hiệu năng

Phần này giới thiệu kết quả mô phỏng của phương pháp tạo SI đề xuất được thực hiện trên bốn chuỗi video thử nghiệm.

A. Điều kiện thử nghiệm

Việc đánh giá hiệu năng được thực hiện trên bốn chuỗi video *BasketBall*, *BQMall*, *PartyScene*, và *RaceHorse* với sự đa dạng về đặc tính chuyển động và cấu trúc bề mặt.

B. Đánh giá chất lượng SI

Để đánh giá hiệu năng của phương pháp đề xuất, PSNR của khung hình SI được tạo ra bằng phương pháp đề xuất SI_{Fusion} được so sánh với khung hình SI

được tạo ra bằng kỹ thuật MCTI, SI_{MCTI} và khung hình SI được tạo ra bởi khung hình lớp BL sau khi tăng độ phân giải SI_{UpBL} .

Các kết quả thu được cho thấy chất lượng của SI_{UpBL} thường tốt hơn chất lượng của SI_{MCTI} và trừ chuỗi *PartyScene*. Sau quá trình kết hợp, chất lượng của SI_{Fusion} đạt được tốt hơn so với các SI ứng viên khác cho tất cả các chuỗi video thử nghiệm.

C. Đánh giá chất lượng khung hình giải mã

PSNR của WZF tái tạo khi sử dụng các phương pháp tạo SI khác nhau được sử dụng để so sánh và đánh giá. Các kết quả chỉ ra rằng $DSVC_{Fusion}$ tốt hơn $DSVC_{Previous}$ cho hầu hết các chuỗi ngoại trừ chuỗi *RaceHorse*. Tổng thể, phương pháp đề xuất đạt được mức cải thiện trung bình tốt hơn. Vì thế, chất lượng của khung hình tái tạo được cải thiện.

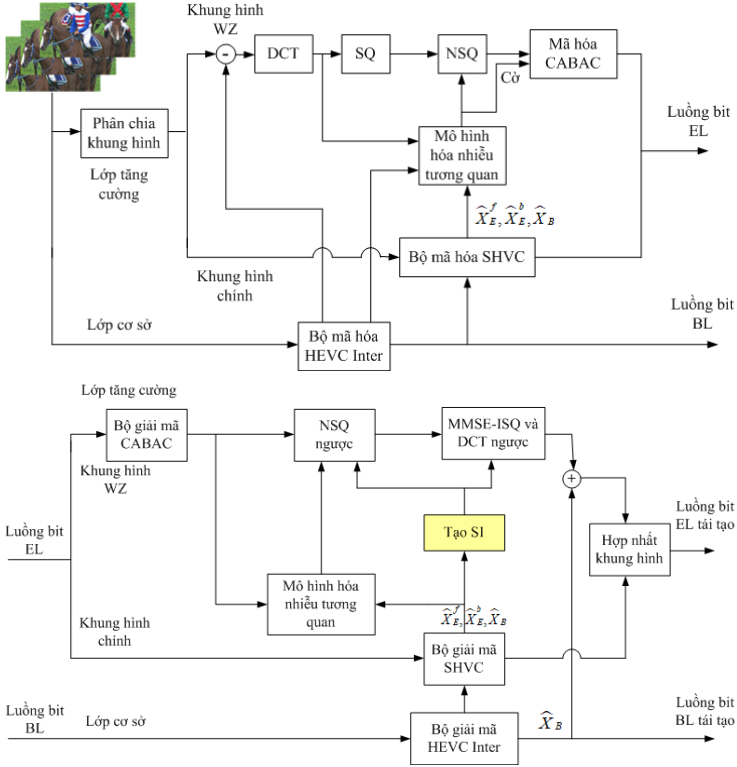
3.3 Đề xuất 4: Tạo thông tin phụ trợ dựa trên kỹ thuật SVM

3.3.1 Đặt vấn đề

Trong phần này đề xuất một phương pháp tạo SI mới bằng cách khai thác một cách hiệu quả thông tin chuyển động BL và khung hình EL giải mã. Các thông tin này được sử dụng linh hoạt để tạo ra ba SI ứng viên. Các SI ứng viên này thể hiện chất lượng vượt trội khác nhau trong các ngữ cảnh khác nhau, do đó cần một cơ chế phân loại thích hợp để chọn ra một SI ứng viên phù hợp trong hoàn cảnh cụ thể. Kỹ thuật phân loại được lựa chọn ở đây là kỹ thuật máy vec-tơ hỗ trợ (SVM) với quá trình huấn luyện trực tuyến.

3.3.2 Kiến trúc bộ mã hóa $DSVC-SVM$

Phần này sẽ giới thiệu kiến trúc tổng thể của bộ mã hóa đề xuất $DSVC-SVM$ như trong Hình 3.2. Trong kiến trúc này có hai lớp, một lớp BL và một lớp EL. Các khung hình lớp BL được nén bằng bộ mã hóa HEVC Inter. Các khung hình BL giải mã cùng với dư thừa lớp BL và thông tin chuyển động lớp BL được sử dụng để nén các khung hình EL. Trong kiến trúc này, dư thừa EL tương ứng với sự sai khác giữa WZF gốc và khung hình BL giải mã. Chỉ có phần dư thừa này được mã hóa và truyền tới bộ giải mã. Lượng dư thừa EL được mã hóa là số lượng các bit LSB, n_{LSB} , đây là phần không thể rút ra từ SI tại bộ giải mã.



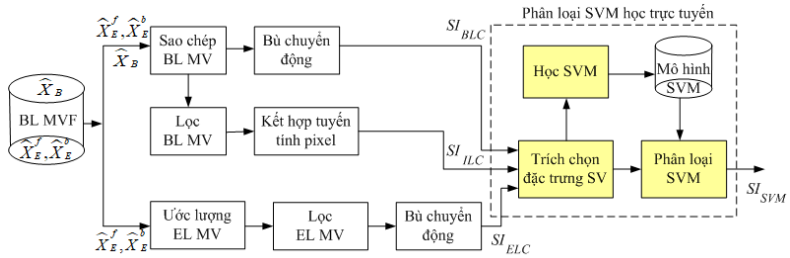
Hình 3.2: Sơ đồ kiến trúc bộ mã hóa và bộ giải mã video phân tán liên lớp
DSVC-SVM

3.3.3 Phương pháp tạo thông tin phụ trợ dựa trên kỹ thuật SVM

Kiến trúc tạo SI gồm hai mô-đun chính là mô-đun tạo ba ứng viên SI và mô-đun lựa chọn ứng viên SI được mô tả trong Hình 3.3. Như được mô tả trong Hình 3.3, quá trình tạo SI gồm ba nhánh để tạo ra ba ứng viên SI và một quá trình lựa chọn một trong ba ứng viên.

A. Quá trình tạo ứng viên SI

Ba SI ứng viên tương ứng với với các loại tương quan thời gian và tương quan liên lớp được tạo ra bao gồm: SI được tạo ra dựa trên trường chuyển động BL, SI_{BLC} ; - SI được tạo ra dựa trên lọc trường chuyển động BL, SI_{ILC} và SI được tạo ra dựa trên lọc trường chuyển động EL, SI_{ELC} .



Hình 3.3: Kiến trúc tạo SI

A. Quá trình tạo ứng viên SI

Ba SI ứng viên tương ứng với với các loại tương quan thời gian và tương quan liên lớp được tạo ra bao gồm: SI được tạo ra dựa trên trường chuyển động BL, SI_{BLC} ; - SI được tạo ra dựa trên lọc trường chuyển động BL, SI_{ILC} và SI được tạo ra dựa trên lọc trường chuyển động EL, SI_{ELC} .

B. Quá trình lựa chọn SI

Để lựa chọn một trong ba ứng viên SI một cách hiệu quả, quá trình lựa chọn SI được coi là một bài toán phân loại. Trong ngữ cảnh này, SVM được lựa chọn để giải bài toán phân loại. Tiêu chí để lựa chọn ở đây là phải chọn được khối SI ứng viên nào giống với khối tương ứng trong khung hình WZ gốc nhất. Việc lựa chọn ấy được quyết định dựa vào thuật toán SVM và mục tiêu cuối cùng là khi xây dựng lại hoàn chỉnh, SIF phải có chất lượng tốt.

Để quá trình chọn SI hiệu quả nhất cũng như để giảm thiểu chi phí tính toán, việc lựa chọn và định nghĩa tập các vector hỗ trợ (SV) cần phải được thực hiện một cách cẩn thận.

1) Xác định đặc trưng

Trong ngữ cảnh này, các đặc trưng được lựa chọn bao gồm: đặc trưng đánh giá trường véc-tơ chuyển động, đặc trưng sai khác SI, đặc trưng tương quan thời gian và đặc trưng tương quan giữa các lớp.

2) Huấn luyện mô hình SVM

Bài toán phân loại ở đây là với ba đầu vào là ba khối SI ứng viên $SI_{BLC}, SI_{ELC}, SI_{ILC}$ cần phải chọn đầu ra là một trong ba ứng viên đó.

Mô hình SVM này cần được huấn luyện trước khi sử dụng cho giai đoạn phân loại. Bởi vì nội dung video rất đa dạng, nên cần sử dụng SVM trực tuyến mà ở đó mô hình huấn luyện được tạo ra một cách thích ứng cho mỗi chuỗi video tùy thuộc vào nội dung thông tin của nó. Để làm được điều này, các đặc trưng SVM được mô tả như ở trên được trích xuất từ khung hình SI đầu tiên của mỗi chuỗi video thử nghiệm. Để cân bằng giữa độ chính xác mô hình và thời gian xử lý để tạo SI nên trong đóng góp này, huấn luyện SVM được đề xuất thực hiện chỉ một lần, cụ thể là cho khung hình SI đầu tiên của mỗi chuỗi thử nghiệm.

3) Trích xuất các đặc trưng

Đối với các khung hình tiếp theo bắt đầu từ khung hình thứ 2, ba SI ứng viên lại được tạo ra và từ đó trích xuất ra tám đặc trưng SV tương ứng theo từng khối.

4) Phân loại SVM

Có nhiều cách để thực hiện thuật toán SVM nhiều lớp nhưng trong đề xuất này sử dụng ba bộ phân loại SVM nhị phân cho ba cặp SI ứng viên như sau:

- Phân loại SVM nhị phân cho mỗi cặp SI: Trong bước này, ba quá trình phân loại SVM nhị phân được thực hiện cho mỗi cặp SI ứng viên.
- Tìm ra SI tốt nhất: Sau khi lựa chọn nhãn đầu ra cho ba bộ phân loại SVM nhị phân ở bước trước, ta được ba nhãn tương ứng với ba bộ phân loại. Nhãn nào được lựa chọn nhiều nhất sẽ được coi là SI cuối cùng.

3.3.4 Đánh giá hiệu năng

Trong phần này sẽ đánh giá chất lượng của SI được tạo ra bằng phương pháp đề xuất và so sánh với một số phương pháp tạo SI khác.

A. Điều kiện thử nghiệm

Đánh giá hiệu năng được thực hiện trên bốn chuỗi video thử nghiệm *RaceHorses*, *BlowingBubbles*, *Basketball Pass*, và *BQSquare*. Phần mềm *SVM^{light}* được sử dụng để lựa chọn các SI ứng viên. Phần mềm tham chiếu HEVC HM phiên bản 14.0 được sử dụng để mã hóa BL và phần mềm tham chiếu SHVC SHM phiên bản 6.0 được sử dụng để mã hóa các KF lớp EL.

B. Đánh giá chất lượng tạo SI

Các giải pháp tạo SI được sử dụng để so sánh với phương pháp đề xuất SI_{SVM} gồm SI_{BLC} , SI_{ELC} , SI_{ILC} và giải pháp tạo SI trong nghiên cứu trước đó SI_{SPIC} . Từ các kết quả có thể rút ra một số kết luận như sau:

- Giải pháp tạo SI đề xuất đạt kết quả tốt hơn so với các phương pháp có liên quan. Điều này đến từ thực tế rằng phương pháp lựa chọn SI dựa trên SVM có thể lựa chọn một cách hiệu quả một SI từ một số các SI ứng viên.

- Giải pháp tạo SI đề xuất có thể đạt được SI với chất lượng tốt hơn, đáng chú ý là mức cải thiện trung bình lên đến 4,57 dB khi so sánh với SI_{ELC} .

- Chất lượng của các SI ứng viên có thể thay đổi tùy vào nội dung chuỗi video. Trong các chuỗi chuyển động nhanh và trung bình, SI_{ELC} vượt trội so với hai ứng viên còn lại. Còn đối với chuỗi chuyển động chậm, SI_{BLC} lại chiếm ưu thế. Điều đó cho thấy vai trò của mỗi ứng viên trong các ngữ cảnh khác nhau là khác nhau và cần một cơ chế lựa chọn phù hợp với nội dung chuỗi video.

- Thêm vào đó, kết quả cũng cho thấy rằng phương pháp đề xuất đạt mức cải thiện cao trong trường hợp chuỗi video chuyển động chậm. Mức cải thiện chủ yếu đến từ việc sử dụng thông tin chuyển động BL và độ chính xác của giải pháp lựa chọn SI dựa trên học máy.

3.4. Kết luận chương

Trong chương này, Luận án đã đề xuất:

- Phương pháp tạo thông tin phụ trợ dựa trên kỹ thuật kết hợp ảnh cho mã hóa video phân tán liên lớp không gian và thời gian. Trong đề xuất này, thông tin phụ trợ được tạo ra bằng cách khai thác thông tin tương quan thời gian và thông tin giữa các lớp. Các SI ứng viên này được kết hợp lại với nhau để tạo nên thông tin phụ trợ cuối cùng có chất lượng tốt hơn.

- Luận án cũng đã đề xuất một phương pháp tạo thông tin phụ trợ dựa trên kỹ thuật SVM cho mã hóa video phân tán liên lớp thời gian và chất lượng. Các kết quả mô phỏng cho thấy chất lượng của SI tạo ra bằng phương pháp đề xuất luôn tốt hơn so với các ứng viên đơn lẻ. Đó là nhờ sử dụng thuật toán SVM huấn luyện trực tuyến một cách hiệu quả.

KẾT LUẬN

Các kết quả nghiên cứu chính của luận án bao gồm các nội dung sau đây:

- *Cải tiến hiệu năng nén cho mã hóa video phân tán tại phía mã hóa* với mô-đun thay đổi kích thước nhóm ảnh. Bộ mã hóa DVC đề xuất lựa chọn kích thước $GOP=2$ hay $GOP=4$ tùy thuộc vào nội dung của từng phân đoạn video gồm 5 khung hình video. Các kết quả mô phỏng cho thấy bộ mã hóa DVC đề xuất có thể tiết kiệm tốc độ bit so với khi sử dụng kích thước GOP cố định cho toàn bộ chuỗi video.

- *Cải tiến hiệu năng nén cho mã hóa video phân tán tại phía giải mã*. Giải pháp mà luận án đề xuất tại phía giải mã bao gồm phương pháp tạo thông tin phụ trợ mới dựa trên lọc liên tục trong quá trình giải mã và phương pháp ước lượng tham α cho mô hình nhiễu tương quan dựa trên mạng nơ-ron. Khi kết hợp hai mô-đun vào bộ mã hóa, các kết quả mô phỏng cho thấy hiệu năng nén tổng thể của bộ mã hóa DVC đề xuất đạt được kết quả tốt hơn khi so với một số bộ mã hóa video liên quan.

- *Cải tiến hiệu năng nén cho mã hóa video phân tán liên lớp bằng cách cải tiến chất lượng của thông tin phụ trợ*. Có hai phương pháp tạo thông tin phụ trợ được đề xuất cho DSVC. Phương pháp thứ nhất đề xuất sử dụng kỹ thuật kết hợp ảnh để kết hợp hai SI ứng viên cho mã hóa video phân tán liên lớp thời gian và không gian. Phương pháp thứ hai đề xuất sử dụng thuật toán SVM huấn luyện trực tuyến để lựa chọn một trong số ba SI ứng viên cho mã hóa video phân tán liên lớp thời gian và chất lượng. Các kết quả mô phỏng cho thấy cả hai phương pháp tạo SI cho DSVC đều có chất lượng tốt hơn so với các ứng viên riêng rẽ cũng như tốt hơn so với một số phương pháp trước đó.

CÁC CÔNG TRÌNH ĐÃ CÔNG BỐ CỦA TÁC GIẢ

[1] (2018) **T. N. Thi Huong**, H. Phi Cong, T. V. Huu, and X. Hoang Van, *Artificial Intelligence Based Adaptive GOP Size Selection for Effective Wyner-Ziv Video Coding*, International Conference on Advanced Technologies for Communications, vol. 2018-October, pp. 120–124, 2018, doi: 10.1109/ATC.2018.8587573.

[2] (2019) T. V. Huu, **T. Nguyen Thi Huong**, M. N. Ngoc, and X. Hoangvan, *Improving performance of distributed video coding by consecutively refining of side information and correlation noise model*, Proceedings - 2019 19th International Symp Symposium on Communications and Information Technologies (ISCIT 2019), pp. 502–506, 2019, doi: 10.1109/ISCIT.2019.8905187.

[3] (2020) Tien Vu Huu, **Thao Nguyen Thi Huong**, Xiem Hoang Van, San Vu Van, *Improving TDWZ Correlation Noise Estimation: A Deep Learning based Approach*, REV Journal on Electronics and Communications: Article scheduled for publication in Vol. 10, No. 1–2, January–June, pp. 11–20, 2020, ISSN 1859 – 378X.

[4] (2017) **N. T. H. Thao**, V. Van San, and V. H. Tien, *Fusion Based Side Information Creation Method for Distributed Scalable Video Coding*, Tạp chí Khoa học và Công nghệ, vol. 121, pp. 48–53, 2017, ISSN 2354-1083.

[5](2017) X. Hoangvan and **T. N. T. Huong**, *An online SVM based side information creation for efficient distributed scalable video coding*, in International Conference on Advanced Technologies for Communications, 2017, vol. 2017-October, pp. 225–228, doi: 10.1109/ATC.2017.8167622.