ÔN ĐỊNH CHẤT LƯỢNG ẢNH LIGHT FIELD DỰA TRÊN BỘ MÃ HÓA VIDEO PHÂN TÁN THẾ HỆ MỚI

Phí Công Huy***, Stuart Perry*, Eva Cheng*, Trịnh Anh Vũ*, Hoàng Văn Xiêm* *Khoa Đa phương tiện, Học Viện Công Nghệ Bưu Chính Viễn Thông *Khoa Điện tử Viễn thông, Trường Đại học Công nghệ, Đại học Quốc gia Hà Nội #University of Technology Sydney

Tóm tắt—Mã hóa video phân tán (DVC) là một giải pháp mã hóa đầy hứa hẹn cho các ứng dụng video yêu cầu độ phức tạp thấp như mạng cảm biến không dây hoặc hệ thống giám sát hình ảnh. Yêu cầu tính ổn định chất lượng hình ảnh là một trong những vấn đề quan trọng trong các hệ thống hiển thị tiên tiến. Tuy nhiên, hầu hết các giải pháp mã hóa video phân tán gần đây, được phát triển dựa trên tiêu chuẩn H.264/AVC hoặc H.265/HEVC, không thể cung câp video với chất lượng ổn định. Trong bài báo này, chúng tôi đề xuất một giải pháp mã hóa video phân tán với ma trân lượng tử (QMs) áp dụng cho dữ liệu Light Field, đảm bảo sự ổn định trong chất lượng hiện thị và cải thiện hiệu năng nén. Trong giải pháp DVC được đề xuất, chuẩn mã hóa video mới nhất, mã hóa video đa nhiệm (Video Versatile Coding – VVC) được lựa chon phù hợp để nén các khung hình chính (Key frame). Trong khi đó, để đạt được sự ổn định trong chất lượng video, các tham số lượng tử hóa (QPs) chọn để nén các khung chính và ma trận lượng tử hóa (QMs) chọn để nén các khung hình Wyner-Ziv (WZ frame) sẽ được giải thích chi tiết trong bài báo. Kết quả thử nghiệm cho thây giải pháp mã hóa video phân tán tích hợp VVC codec (D-VVC) được để xuất vượt trội hơn đáng kế so với các giải pháp khác mã hóa DVC liên quan, đặc biệt là bộ mã hóa phân tán DISCOVER và giải pháp DVC-HEVC gần đây, về hiệu suất nén đồng thời cung cấp tính ốn định cho chất lượng hình ảnh tốt hơn cho video được giải mã.

Từ khóa— Mã hóa dữ liệu Light Field, chất lượng ảnh Light Field, mã hóa video phân tán.

I. GIỚI THIỆU CHUNG

Công nghệ mã hóa video đã và đang đóng một vai trò quan trọng trong bối cảnh các dịch vụ nghe nhìn như truyền hình kỹ thuật số, video di động và truyền trực tuyến Internet đáp ứng các yêu cầu nén cao. Hầu hết các chuẩn mã hóa video hiện có, đặc biệt là tiêu chuẩn ITU-T H.26x và ISO/IEC MPEG-x [1], áp dụng cái gọi là mô hình mã hóa video dự đoán. Trong đó các mối tương quan thời gian và không gian được khai thác tại bộ mã hóa bằng cách sử

Tác giả liên hệ: Hoàng Văn Xiêm,

Email: xiemhoang@vnu.edu.vn

Đến tòa soạn: 31/8/2021, chỉnh sửa: 19/10/2021, chấp nhận đăng: 22/10/2021.

dụng ước tính chuyển động/vòng bù chuyển động và biến đổi không gian tương ứng. Kết quả các chuẩn mã hóa này thường sẽ phức tạp tại bộ mã hóa và đơn giản tại bộ giải mã.

Tuy nhiên, với sự bùng nổ của các ứng dụng hiện nay, chẳng hạn như giám sát video không dây công suất thấp, mạng cảm biến hình ảnh, camera PC không dây và điện thoại camera di động, độ phức tạp thấp của bộ mã hóa trở thành một tính năng cần thiết cho các chương trình mã hóa video [2]. Trong bối cảnh này, mã hóa video phân tán (DVC), được phát triển dựa trên các định lý Slepian-Wolf và Wyner-Ziv [3, 4] để cho phép khai thác mối tương quan thời gian tại bộ giải mã thay vì tại bộ mã hóa mà không bị mất hiệu suất nén, đây có thể được coi như một giải pháp mã hóa video đầy hứa hẹn cho các ứng dụng nói trên.

Công nghệ tiên tiến nhất về mã hóa video phân tán là bộ mã hóa phân tán DISCOVER [5] được phát triển bởi F. Pereira và cộng sự. Giải pháp mã hóa này tuân theo phương pháp tiếp cận DVC của Stanford [6, 7], trong đó video đầu vào được chia thành hai chuỗi con: một chuỗi các khung hình chính (Key frame) được nén với tiêu chuẩn H.264/AVC truyền thống [8] và một chuỗi các khung hình Wyner-Ziv (WZ frame) được nén bằng giải pháp mã hóa phân tán. Khung chính và khung WZ được mã hóa riêng biệt nhưng được giải mã chung để khai thác mối tương quan thời gian giữa hai khung.

Gần đây, lấy cảm hứng từ việc đạt được tiêu chuẩn Mã hóa video hiệu quả cao (HEVC) [9], đặc biệt là tiết kiệm được khoảng 50% tốc độ bit khi so sánh với chuẩn H.264/AVC trước đây. C. Brites và cộng sự đã trình bày trong [10] một giải pháp DVC dựa trên HEVC trong đó cấu hình HEVC Intra có độ phức tạp thấp được sử dụng để nén các Key frame. Bài báo này cũng trình bày một đánh giá hiệu suất chi tiết cho bộ mã hóa DVC-HEVC. Tuy nhiên, hầu hết các sản phẩm trước đây, bao gồm cả DVC-HEVC không thể cung cấp video đã giải mã với chất lượng ổn đinh. DVC-HEVC cho phép kiểm soát chất lượng (hoặc tốc độ bit) của khung chính được giải mã bằng tham số lượng tử hóa (QP), đồng thời khung WZ được dùng với ma trận lượng tử hóa (QM). Trong [10], một bộ (QP, QM) được sử dụng trong bộ mã hóa DISCOVER DVC [5], trước đó được sử dụng lại cho phần mềm tham chiếu DVC-HEVC. Tuy nhiên, với tốc độ nén tương tự, chất lượng của các khung hình chính được mã hóa theo tiêu chuẩn HEVC tốt hơn

nhiều so với tiêu chuẩn H.264/AVC. Do đó, việc sử dụng khuyến nghị tương tự (QP, QM) có thể không phù hợp với nâng cấp DVC-HEVC, để đảm bảo chất lượng ổn định của hình ảnh.

Mặc dù HEVC [9] có thể đạt được sự cải thiện đáng kể về hiệu quả nén khi so sánh với tiêu chuẩn H.264/AVC trước đây [8], có thể thấy trước rằng sẽ cần phải có khả năng nén tốt hơn nữa trong tương lai, cho cả ứng dụng truyền thống và ứng dung mới. Về vấn đề này, một hoạt động khảo sát chung đã được bắt đầu bởi ISO/IEC (Tổ chức tiêu chuẩn hóa quốc tế/Ủy ban kỹ thuật điện quốc tế) MPEG (Nhóm chuyên gia hình ảnh chuyển động) và ITU-T (Liên minh viễn thông quốc tế- Lĩnh vực tiêu chuẩn hóa viễn thông) VCEG (Nhóm chuyên gia mã hóa video) dưới sự bảo trợ của Nhóm nghiên cứu video (JVET). Việc tiêu chuẩn hóa trong tương lai như vậy có thể bắt đầu từ các phần mở rộng bổ sung của HEVC hoặc một chuẩn hoàn toàn mới [11]. Các hoạt động của JVET đã dẫn đến đề xuất một chuẩn video tiên tiến mới, mã hóa video đa nhiệm (VVC), có thể đạt được khoảng 50% tiết kiệm tốc độ bit khi so sánh với tiêu chuẩn HEVC [11].

Trong bài báo này, để hoàn thiện được bộ mã hóa DVC chất lượng hiệu quả kèm theo ổn định về khung nhìn, chúng tôi đề xuất một giải pháp DVC mới, khai thác ưu điểm từ các nâng cấp của bộ mã hóa video đa nhiệm của JVET [11]. Trong giải pháp DVC được để xuất, Video Versatile Coding (VVC) được sử dụng để nén các khung chính thay vì sử dụng H.264/AVC hoặc HEVC như các công trình trước đây [5, 6, 10]. Ngoài ra, để duy trì độ phức tạp thấp của kiểu mã hóa phân tán, chúng tôi đề xuất lựa chọn kỹ lưỡng một số công cụ mã hóa Intra được đề xuất cho cầu trúc VVC. Cuối cùng, để đảm bảo chất lượng ốn định giữa Key frame được giải mã và WZ frame, đặc biệt lần đầu áp dụng cho loại dữ liệu mới Light Field, chúng tôi để xuất một bộ cấu hình (QP, QM) mới cho bộ mã hóa D-VVC. Kết quả thử nghiệm cho thấy hiệu suất nén của D-VVC được đề xuất tốt hơn đáng kể so với DVC dựa trên H.264/AVC trước đây (tức là giải pháp DISCOVER DVC) [12] và giải pháp DVC-HEVC, đồng thời làm cho chất lượng khung hình được giải mã ổn định hơn đối với người dùng khi xem.

Cấu trúc của bài báo này như sau. Phần 2 tóm tắt nền tảng lý thuyết và các công trình gần đây liên quan tới dữ liệu Light Field và mã hóa video phân tán. Phần 3 trình bày kiến trúc D-VVC được đề xuất. Sau đó, Phần 4 mô tả việc lựa chọn (QP, QM) để đạt được sự ổn định về chất lượng cho video được giải mã, và Phần 5 đánh giá hiệu suất của giải pháp D-VVC. Cuối cùng, Phần 6 đưa ra một số nhận xét kết luận và các công việc trong tương lai.

II. LÝ THUYẾT CƠ SỞ VÀ NHỮNG NGHIÊN CỨU LIÊN QUAN

DVC đã thu hút nhiều sự chú ý trong thập kỷ qua do các tính năng mã hóa đặc thù của nó, đặc biệt là sự phân bố linh hoạt độ phức tạp của bộ mã hóa, khả năng phục hồi lỗi và khả năng mở rộng độc lập của bộ mã hóa vì không sử dụng vòng lặp dự đoán như tại mã hóa dự đoán [13]. Trong phần này, chúng tôi sẽ mô tả ngắn gọn bộ mã hóa DVC tiên tiến nhất và các hoạt động gần đây trên DVC, đồng thời đặc tính của dữ liệu đặc biệt Light Field.

A. Đặc tính dữ liệu ảnh Light Field

Dữ liệu plenoptic-light field (LF) được biết đến như một hình thức kết xuất dựa trên hình ảnh (IBR), nó thu thập lượng lớn dữ liệu hình ảnh vì cường độ của tia sáng truyền theo mọi góc độ tại mọi điểm trong không gian 3D đều được ghi nhận. Do đó, dữ liệu hình ảnh LF bao gồm các thông tin [14] như không gian (x, y, z), thời gian (t), vị trí (θ, ϕ) bước sóng của ánh sáng (λ). Nó được tổng hợp chung trong tham số $P_{LF}(x, y, z, \theta, \phi, \lambda, t)$ như Hình 1. Để đơn giản hóa tham số của LF gốc, chức năng LF đã được giảm xuống còn 4 tham số hay còn gọi là 4D-LF, $P_{LF} = L(u, v, x, y)$, với (x, y) là điểm nhìn còn (u, v) là góc nhìn. Một hình ảnh LF gốc bao gồm các hình ảnh cực nhỏ (MI) và một tập hợp các hình ảnh khẩu độ kèm theo (SAI) thu được bằng cách sấp xếp lại các pixel cùng vị trí từ mỗi MI.



Hình 1. Các loại thông tin thể hiện cho dữ liệu Light Field

Dữ liệu Ligh Field được tạo ra hoặc ghi lại bằng một máy ảnh chuyên dụng - Lytro Illum, trong đó nó có một hệ thống những ống kính nhỏ được sắp xếp trước cảm biến của máy ảnh, để ghi lại những hình ảnh từ các góc độ và chiều sâu khác nhau như Hình 2. Máy ảnh có thể xử lý đữ liệu 4D-LF với mảng 4D có kích thước $15 \times 15 \times 434 \times 625 \times 4$ trong đó 15×15 biểu thị số lượng chế độ xem các ảnh nhỏ (SAI), 434×625 biểu thị độ phân giải không gian của mỗi chế độ xem và 4 đại diện cho không gian màu của RGB và trọng số đi kèm như ở Hình 3 [15]. Vì vậy, dữ liệu của light field dung lượng và kích thước khá lớn với mỗi bức hình đuôi file là .LFR (Light Field Raw) cùng kích thước 7728x5368, dung lượng sấp xỉ 50MB.



Hình 2. Cấu tạo cơ bản bên trong máy ảnh để thu dữ liệu Ligh-Field



Hình 3: Cấu trúc dữ liệu LF.



Hình 5. Cấu trúc của mô hình D_VVC

Từ những đặc thù của cấu trúc dữ liệu LF, những tính năng mà dữ liệu LF có thể cung cấp như: xuất ảnh 3D, thay đổi điểm lấy nét, hay tạo ra các video dựa trên các ảnh nhỏ (sub-aperture image – SAI), v.v

B. Những kiến thức nền tảng liên quan tới mã hóa video phân tán

Bộ mã hóa phân tán DISCOVER [5] là một trong những chương trình DVC hiệu quả nhất trong lĩnh vực phân tán và được áp dụng rộng rãi làm tài liệu tham khảo cho sự phát triển của các bộ mã hóa DVC sau này. Để đánh giá hiệu suất, kết quả thông qua việc đánh giá tại [12] thể hiện hiệu suất tốt hơn của giải pháp DVC so với H.264/AVC, với độ phức tạp mã hóa thấp thông qua hai loại mã hóa, đó là H.264/AVC Intra và H.264/AVC No Motion. Bộ giải mã DISCOVER DVC được phát triển dựa trên cách tiếp cận DVC ban đầu của Stanford [6, 7]. Trong giải pháp này, video đầu vào được chia thành các khung hình chính và khung hình WZ thông qua cái gọi là Group Of Picture (GOP). Khung đầu tiên của mỗi GOP, được gọi là khung chính, được mã hóa bằng chuẩn mã hóa H.264/AVC Intra truyền thống [8] trong khi các khung WZ còn lai được mã hóa bằng cách tiếp cận DVC. Tại bộ giải mã, hai khung chính được giải mã liên tiếp được sử dụng để tạo ra cái gọi là thông tin phụ, một phiên bản nhiễu của khung WZ gốc. Biết được mối tương quan giữa khung WZ ban đầu và thông tin phụ của nó thông qua mô hình nhiễu tương quan, một số bit chẳn lẻ được gửi từ bộ mã hóa đến bộ giải mã đế giúp tái tạo lại khung WZ. Hình 4 minh họa kiến trúc của codec DISCOVER DVC phổ biến.



Hình 4. Cấu trúc bộ mã hóa video phân tán DISCOVER DVC [5]

C. Những nghiên cứu liên quan tới mã hóa video phân tán

Trong DVC, việc tạo thông tin phụ và mô hình nhiễu

tương quan đóng vai trò quan trọng vì chúng ảnh hưởng trực tiếp đến hiệu suất nén DVC cuối cùng. Do đó, nhiều nghiên cứu đã được thực hiện để tăng chất lượng thông tin phụ - Side Information (SI) [16-24] hoặc ước tính chính xác mối tương quan giữa WZ gốc và thông tin phụ [25-27].

Kênh phản hồi được trình bày trong cách tiếp cận DVC của Stanford có thể không phù hợp với nhiều ứng dụng có yêu cầu độ trễ thấp, một số nghiên cứu [28 - 31] đã được giới thiệu để loại bỏ sự hiện diện của kênh phản hồi trong khi vẫn duy trì hiệu suất nén của codec DVC.

Trong cách tiếp cận mã hóa DVC của Stanford, bộ giải mã bao gồm một số thành phần tính toán, tức là tạo SI, mô hình nhiễu tương quan và giải mã WZ lặp lại [5-7]. Trong bối cảnh này, bộ giải mã DVC thường đòi hỏi độ phức tạp tính toán cao, đáng chú ý là khi so sánh với các tiêu chuẩn mã hóa video dự đoán truyền thống [8, 9]. Do đó, một số nghiên cứu đã được thực hiện để giảm độ phức tạp của bộ giải mã DVC, bao gồm [32, 33].

Cuối cùng, vì mã hóa khung chính ảnh hưởng trực tiếp đến hiệu suất nén DVC tổng thể, bài báo tại [10] đã giới thiệu một giải pháp mã hóa khung chính được cải thiện với các nâng cấp HEVC. So với bộ mã hóa DISCOVER DVC hiện đại nhất, DVC-HEVC được đề xuất tiết kiệm đáng kế tốc độ bit trong khi cung cấp chất lượng khung hình được giải mã theo cách tương tự. Tuy nhiên, cần lưu ý rằng vì hiệu suất nén của tiêu chuẩn HEVC đương nhiên vượt trội hơn so với chuẩn H.264/AVC, với cài đặt thông số lượng tử hóa (QP) tương tự, chất lượng của khung hình chính được giải mã bằng DVC-HEVC sẽ cao hơn nhiều so với các khung WZ được giải mã. Điều này dẫn đến sự dao động chất lượng trong video được giải mã và ảnh hưởng đến chất lượng xem video của người dùng. Đặc biệt với dữ liệu ảnh Light Field, chưa có nghiên cứu nào thực sự để xuất liên quan.

Để khắc phục vấn đề dao động chất lượng với khung chính được giải mã và khung WZ gốc, chúng tôi đề xuất trong phần tiếp theo, một tham số lượng tử hóa mới và thiết lập ma trận lượng tử hóa cho khung chính và mã hóa WZ. Hơn nữa, lấy cảm hứng từ nghiên cứu gần đây của mã hóa video JVET [11], chúng tôi đề xuất chuẩn mã hóa VVC có độ phức tạp thấp nhưng hiệu quả để mã hóa các khung hình chính, sử dụng cho dữ liệu ảnh Light Field. Trong trường hợp này, D-VVC được đề xuất sẽ có thể kế thừa hiệu suất nén cao của giải pháp VVC đồng thời đảm bảo chất lượng ổn định trong khung chính được giải mã và khung WZ.

III. ĐỀ XUẤT MÔ HÌNH MÃ HÓA VIDEO PHÂN TÁN D_VVC

Trong phần này, chúng tôi sẽ mô tả kiến trúc D_VVC được đề xuất. Sau đó, chúng tôi trình bày các công cụ mã hóa VVC đã chọn để nén khung hình chính.

A. Cấu trúc mô hình D_VVC

Hình 5 minh họa kiến trúc D_VVC được đề xuất với những điểm mới, những phần quan trọng đã được bôi vàng. D_VVC được đề xuất tuân theo cách tiếp cận mã hóa DVC của Stanford thường được áp dụng trong bộ mã hóa DISCOVER DVC [5]. Trong giải pháp D_VVC được đề xuất, mã hóa video VVC [11] được chọn để nén các khung hình chính thay vì sử dụng các tiêu chuẩn H.264/AVC hoặc HEVC.

Cấu trúc của mô hình D_VVC được mô tả như sau:

- Quá trình mã hóa:
 - Hình thành và tách khung: Dữ liệu LF trước hết được giải nén và giải mã và biểu diễn dưới dạng 4D-LF. Sau đó, các chuỗi ảnh nhỏ Sub-aperture image (SAI) trong 4D-LF được nhóm lại thành một chuỗi giả sử dụng cách quét xoáy ốc như hình 6. Từ đây, những vấn đề của việc nén ảnh đã được chuyển sang thành vấn đề của nén video. Khung đầu tiên của mọi nhóm ảnh (GOP), được gọi là khung chính (Key), được mã hóa bằng cách tiếp cận mã hóa nội bộ H.266/VVC tiêu chuẩn mã hóa mới nhất [11, 34], chỉ sử dụng tương quan không gian; do đó, có thể đạt được độ phức tạp thấp. Đối với các khung WZ còn lại, các bước sau được thực hiện theo trình tự.



Hình 6. Phương thức quét xoáy ốc để tạo video giả lập cho LF

- Mã hóa khung chính (Key): Trong hệ thống DVC, khung chính đóng một vai trò quan trọng vì thông tin được giải mã của nó sẽ được sử dụng tại bộ giải mã để giúp tái tạo lại các khung WZ ban đầu. Đối với mã hóa khung chính, chuẩn mã hóa VVC Intra áp dụng trong D_VVC được đề xuất. Trong mã hóa khung chính, chỉ sử dụng mối tương quan không gian; do đó, có thể đạt được độ phức tạp thấp và độ bền của lỗi.
- Mã hóa khung WZ: Đối với mã hóa khung WZ, biến đổi cosin rời rạc, bộ lượng tử đồng nhất và mã Kiểm tra chẵn lẻ mật độ thấp (LDPC) tương ứng được áp dụng để nén khung WZ gốc thường được sử dụng trong bộ mã hóa DISCOVER DVC [5]
- Quá trình giải mã:
 - Giải mã khung chính: Tại bộ giải mã, dòng bit được mã hóa trước hết được giải mã bằng cách sử dụng bộ giải mã VVC để đạt được các khung chính đã được giải mã. Các khung chính được giải mã sau đó được sử dụng để giải mã các khung WZ phía sau.
- Giải mã khung WZ: Như đã nêu trong khái niệm mã hóa phân tán [3, 4], DVC cho phép mã hóa riêng video đầu vào, tức là khung chính và khung WZ, nhưng cùng giải mã các khung WZ bằng cách sử dụng cả thông tin của khung chính và WZ. Việc giải mã chung được thực hiện theo các bước sau:
- Tạo thông tin phụ (SI): Để khai thác mối tương quan thời gian giữa khung chính và khung WZ, một bước gọi là tạo thông tin phụ được thực hiện để tạo thông tin phía bộ giải mã, một phiên bản nhiễu của khung WZ gốc. Trong bài báo này, thuật toán nội suy thời gian bù chuyển động nâng cao [35, 36] được sử dụng để tạo SI chất lượng cao.
- Mô hình nhiễu tương quan: Trong DVC, mối tương quan giữa các khung WZ và SI ban đầu được mô hình hóa dưới dạng phân phối Laplacian và các khung chính được giải mã được sử dụng để xây



Hình 7. Hình minh họa chất lượng chủ quan không ổn định nếu không lựa chọn đúng (QP, QM): dữ liệu Bikes



Hình 8. Dữ liệu kiểm thử của hình ảnh LF

	QPs	PSNR_Key -	QM							
Dữ liệu			1	2	3	4	5	6	7	8
Bikes	46	23.38	24.81	25.08	25.34	27.10	27.25	28.07	30.20	33.79
	40	27.04	27.74	27.94	27.98	28.82	28.95	29.66	31.09	34.23
	34	31.27	31.68	31.82	31.83	32.08	32.11	32.51	33.17	35.34
	32	32.77	33.05	33.19	33.19	33.36	33.38	33.70	34.21	36.02
	27	36.81	36.45	36.61	36.60	36.71	36.71	36.98	37.32	38.46
	22	41.22	39.36	39.61	39.58	39.69	39.68	40.02	40.40	41.45
Books Black Fence	46	25.30	26.81	27.09	27.35	29.25	29.41	30.44	32.63	36.44
	40	28.63	29.31	29.52	29.60	30.64	30.81	31.71	33.33	36.74
	34	32.99	32.81	32.98	33.01	33.45	33.54	34.15	34.98	37.43
	32	34.52	33.95	34.12	34.17	34.57	34.63	35.15	35.84	37.96
	27	38.40	36.29	36.50	36.53	36.93	36.97	37.48	38.04	39.62
	22	42.27	37.73	37.98	38.01	38.48	38.52	39.08	39.70	41.21
	46	24.63	25.83	26.09	26.21	27.71	27.80	28.70	30.68	34.51
	40	28.82	29.50	29.73	29.75	30.34	30.42	31.04	32.23	35.31
	34	33.31	33.47	33.66	33.66	33.95	33.98	34.40	34.95	36.88
	32	34.85	34.82	35.01	35.00	35.24	35.26	35.61	36.06	37.62
	27	38.86	37.80	38.04	38.02	38.27	38.27	38.60	39.03	40.08
	22	43.01	40.07	40.41	40.36	40.68	40.67	41.07	41.65	42.78
Poppies	46	22.95	24.74	25.02	25.23	27.11	27.19	27.81	30.19	33.92
	40	27.02	27.70	27.95	28.06	29.20	29.25	29.81	31.52	34.70
	34	31.45	30.84	31.11	31.17	31.95	31.99	32.48	33.70	36.19
	32	32.99	31.78	32.07	32.13	32.88	32.92	33.41	34.53	36.85
	27	37.04	33.60	33.97	34.02	34.84	34.87	35.39	36.50	38.58
	22	41.30	34.60	35.02	35.08	36.03	36.04	36.65	37.88	40.02

BẢNG I. ĐỒNG BỘ GIÁ TRỊ PSNR (DB) GIỮA QPS VÀ QMS

dựng xác suất như vậy.

 Tái tạo: Cuối cùng, các bit chẵn lẻ, thu được từ bộ giải mã LDPC, cùng với SI và thông tin nhiễu tương quan được ước tính từ các bước trước được sử dụng để tái tạo lại khung WZ. Để đạt được sự tái tạo tốt, việc xây dựng lại WZ dựa trên sai số bình phương trung bình (MMSE) tối thiểu [37] được áp dụng trong cấu trúc D_VVC.

B. Mã hóa khung hình chính bằng VVC Intra

H.266/VVC là tiêu chuẩn mã hóa video mới nhất, được ITU-T và ISO/IEC cùng xuất bản. Dự án mã hóa video này đã được bắt đầu từ năm 2015 trong giai đoạn thăm dò; ở giai đoạn phát triển, nhóm cộng tác chung bắt đầu soạn thảo phiên bản đầu tiên của bộ mã hóa này vào năm 2018. Vào tháng 7 năm 2020, bộ mã hóa VVC được công bố là chuẩn nén video thế hệ tiếp theo. Tiêu chuẩn video này được thiết kế để hiệu quả và linh hoạt để đáp ứng nhu cầu truyền thông hiện đại.

Đối với hiệu quả mã hóa, H.266/VVC được bố sung nhiều công cụ mã hóa mới so với tiêu chuẩn H.265/HEVC [9] trước đó như độ phân giải vecto chuyển động thích ứng (AMVR), chế độ phân vùng tam giác (TPM), ma trận dựa trên nội dự đoán (MIP), phân vùng khối linh hoạt với cây đa loại (MTT), dự đoán trước sau với trọng số CU (BCW), biến đổi khối phụ (SBT), và nhiều cải tiến gia tăng của thiết kế mã hóa video lai cổ điển [34]. Dựa trên sự cải tiến của các công cụ mã hóa mới, mức giảm tốc độ bit được báo cáo là khoảng 50% so với H.265/HEVC, đặc biệt, H.266/VVC có thể sử dụng cho video độ nét cực cao, ví dụ: 4K hoặc 8K, cũng với dải động cao và gam màu rộng [34].

Đối với tính linh hoạt, chuẩn mã hóa mới bao gồm ba khía cạnh quan trọng, được giải quyết cho các ứng dụng cụ thể, đó là mã hóa nội dung màn hình cho một số ứng dụng (ví dụ: chơi game, chia sẻ màn hình), lấy mẫu hình ảnh tham chiếu để phát trực tuyến tương thích với chuyển đổi độ phân giải và hình ảnh phụ độc lập để phát trực tuyến video 360 đô [34].

Về quy trình mã hóa WZ, chuỗi video giả lập 4D-LF được mã hóa và giải mã bằng chế độ cấu hình trong khung (Intra), được nâng cấp lên 67 hướng dự đoán, so với 35 hướng dự đoán của HEVC. Do đó, có một lợi thế cho hiệu suất nén của D_VVC được đề xuất dựa trên tiêu chuẩn H.266/VVC

IV. TẠO TÍNH ÔN ĐỊNH TRONG CHẤT LƯỢNG ẢNH LIGHT FIELD

Chất lượng video đồng bộ đóng một vai trò quan trọng trong các hệ thống hiện thi hiện nay. Hình 7 minh họa video được giải mã bằng D_VVC nhưng chưa lựa chọn bộ đề xuất (QP, QM). Ở đây, QP và QM được lựa chọn ngẫu nhiên, cụ thể QP = 30 và QM = 0 cho dữ liệu Bikes.

Như đã trình bày, với cài đặt (QP, QM) khi chưa được lựa chọn và sự dụng ngẫu nhiên, chất lượng của các khung WZ được giải mã sẽ thấp hơn chất lượng của các khung chính được giải mã. Vấn đề này đương nhiên sẽ ảnh hưởng đến chất lượng xem, khiến chất lượng video bị dao động không mong muốn. Để giải quyết vấn đề dao động chất lượng của video, chúng tôi trình bày trong bài báo này một thử nghiệm mở rộng để chọn cấu hình (QP, QM) cho mã hóa khung chính và khung WZ tương ứng.

Với mục đích này, bốn dữ liệu LF khác nhau được sử dụng để kiểm tra, Bikes, Books, Black Fence và Poppies, những dữ liệu LF này đã được giải mã và biến đổi để tận dụng các hình ảnh SAIs của LF để tạo ra video giả như Hình 8.

Để có được sự đa dạng của chất lượng video được giải mã, các QP được đặt từ 22 đến 46, và kết quả chất lượng PSNR được giải mã nằm trong khoảng từ 30dB đến 40dB. Vì mã hóa khung chính ảnh hưởng đến hiệu suất mã hóa khung WZ, do đó, các QM được lựa chọn cẩn thận dựa trên cài đặt QP và dưới sự dao động tối ưu của chất lượng. Ở đây, sự lựa chọn tối ưu cho sự dao động của chất lượng được đo lường thông qua số liệu Delta PSNR được tính toán dựa trên trung bình của khung chính và chất lượng khung hình được giải mã WZ.

$$Delta_{PSNR} = |PSNR_{Key} - PSNR_{WZ}|$$
(1)

Có thể thấy, Delta PSNR nhỏ hơn, mức độ dao động chất lượng video thấp hơn. Để đạt được khả năng chất lượng nhất quán và ổn định, đối với mỗi QP, QM có thể được chọn Delta PSNR thấp nhất. Điều này có nghĩa là

$$QM = \underset{QM=\{1,2,..8\}}{\operatorname{argmin}} Delta_{PSNR}$$
(2)

Từ các kết quả thu được trong Bảng I, để đạt được sự ổn định về chất lượng, có thể chọn ma trận (QP, QM) như trong Bảng II

BÁNG II. GIÁ TRỊ QP VÀ QM LỰA CHỌN CHO DỮ LIỆU
LF

Bikes		Bo	oks	Bla Fe	ack nce	Poppies		
QP	QM	QP	QМ	QP	QM	QP	QM	
46	1	46	1	46	1	46	1	
40	1	40	1	40	1	40	1	
34	1	34	2	34	1	34	3	
32	1	32	4	32	1	32	5	
27	5	27	7	27	7	27	7	
22	7	22	8	22	8	22	8	

V. ĐÁNH GIÁ VÀ KIỄM THỬ

A. Phương pháp đánh giá

Để đánh giá chất lượng ổn định và hiệu suất của bộ mã hóa D_VVC, hiệu suất tốc độ - sự biến dạng (RD) của bộ mã hóa D_VVC được so sánh với các chuẩn mã hóa khác nhau, cùng với bộ đề xuất (QP, QM) để nén khung chính và khung WZ. Bốn chuỗi thử nghiệm phổ biến, Bikes, Books, Black Fence và Poppies được sử dụng với định dạng YUV 4: 2: 0, kích thước QCIF và ở độ phân giải 15Hz.

Để đánh giá hiệu suất nén D_VVC và chất lượng ổn định của video, chúng tôi trình bày 2 điểm trọng tâm trong phần này: i) đánh giá tổng thể hiệu suất RD của D_VVC; ii) đánh giá chất lượng ổn định của video.

B. Đánh giá hiệu năng mã hóa

Về hiệu suất nén, hiệu suất RD được sử dụng rộng rãi cho các lược đồ mã hóa video của Bjøntegaard Delta (BD) -PSNR và Bjøntegaard Delta (BD) -Rate [38]. Hình 9 trình bày so sánh đường cong RD giữa giải pháp mã hóa D_VVC và các chuẩn mã hóa liên quan khác, đó là DVC dựa trên HEVC được gọi là DVC-H.265/HEVC [10] và DISCOVER codec [5]; trong khi BD-Rate và BD-PSNR được tính trong Bảng III.

BẢNG III. HIỆU NĂNG BD-RATE VÀ BD-PSNR GIỮA D VVC VÀ CÁC CHUÂN KHÁC

	<u></u>	n ente et			
	D_VVC	vs DVC-	D_VVC vs DISCOVER [5]		
LF	H.265/	HEVC			
sequences	[10]				
	BD-	BD-	BD-	BD-	
	Rate	PSNR	Rate	PSNR	
Black fence	-35.78	3.78	-47.02	5.42	
Bikes	-29.01	2.38	-31.47	3.48	
Poppies	-42.71	3.44	-51.08	4.41	
Books	-40.40	3.29	-50.16	4.42	
Average	-36.97	3.22	-44.93	4.43	

Từ các kết quả quan sát được, có thể rút ra một số kết luận như sau.

- D_VVC so với DVC-H.265/HEVC: Với DVC-H.265/HEVC, giải pháp D_VVC được đề xuất cải thiện đáng kể hiệu suất RD cho tất cả các chuỗi video giả lập 4D-LF, với nhiều loại nội dung khác nhau. Như quan sát thấy, mức giảm BD-Rate tương ứng của loại nội dung là 35,78%, 29.01%, 42,71% và 40,40% thu được đối với các loại nội dung là Black fence, Bikes, Poppies, và Books. Do đó, cải tiến hiệu suất thu được chứng minh hiệu quả của chuẩn mã hóa VVC được nâng cấp trong kiến trúc D_VVC.
- D_VVC so với DISCOVER codec: D_VVC đề xuất đã đạt được kết quả ấn tượng về hiệu suất nén so với DISCOVER codec với mức giảm tổng cộng của bitrate khoảng 44,93%. Ngoài ra, tính đến mối tương quan cao giữa các SAI và độ phức tạp thấp của giải pháp mã hóa WZ, kiến trúc D_VVC đạt được hiệu năng nén đáng kể, cụ thể tăng thêm 5.42dB cho Black fence và trung bình khoảng 4,4dB đối với các loại dữ liệu khác.

C. Đánh giá sự ổn định chất lượng video

Trong bài báo này, để đạt được sự ổn định của chất lượng video, đề xuất mới cho bộ (QP, QM) được trình bày trong



Hình 10. Minh họa chất lượng chủ quan ổn định với cài đặt (QP, QM) được đề xuất: dữ liệu Bikes

Phần 4 để có thể áp dụng tương ứng như khuyến nghị cài đặt của bộ mã hóa DISCOVER [5, 12].

Để đánh giá chất lượng video được giải mã liên tiếp theo các khung hình với bộ QP, QM mới được đề xuất, Hình 10 cho thấy chất lượng chủ quan khá ổn định và nhất quán của 5 khung hình video liên tiếp với bộ (QP = 46, QM = 3) đã được đề xuất cho dữ liệu Bikes. Trong thử nghiệm này, bản đồ lỗi liên quan đến mỗi khung hình đã giải mã được tính bằng cách sử dụng thông tin gốc, và Delta PSNR được tính cho mỗi cài đặt (QP, QM).

VI. KẾT LUẬN

Trong bài báo này, một giải pháp mã hóa video phân tán mới đã được đề xuất áp dụng chuẩn mã hóa video mới nhất VVC. Các công cụ mã hóa trong khung VVC Intra được sử dụng để nén các khung hình chính trong khi mã hóa DVC được sử dụng để nén các khung hình WZ. Để đạt được chất lượng nhất quán và ổn định cho chuỗi video được giải mã, một bộ tham số lượng tử (QP) và ma trận lượng tử (QM) dã được đề xuất để mã hóa tương ứng cho khung hình chính và mã hóa khung hình WZ. Cuối cùng, một đánh giá hiệu suất tổng quát được thực hiện để kiểm chứng hiệu năng của giải pháp D_VVC so sánh với bộ mã hóa DISCOVER DVC và giải pháp DVC-HEVC gần đây. Kết quả cho thấy bộ mã hóa D_VVC đã cung cấp một hiệu năng nén tốt, kèm theo với sự ổn định của chất lượng hiện thị thông qua bộ (QM, QP), đặc biệt áp dụng cho dữ liệu ảnh Light Field, so với các chuẩn mã hóa liên quan.

Công việc trong tương lai có thể nhóm tác giả sẽ tập trung vào việc tìm ra một thuật toán kiểm soát chất lượng ổn định mang tính tự động, có thể ước tính chất lượng của khung hình WZ dựa trên nội dung đầu vào video.

TÀI LIỆU THAM KHẢO

- D. Turaga and T. Chen, 'ITU-T Video Coding Standards', in Electrical and Computer Engineering, Carnegie Mellon University, 2000
- [2] C. Brites and F. Pereira, 'Distributed video coding: Bringing new applications to life', 5th Conference on

Telecommunications – ConfTele, Tomar, Portugal, Apr. 2005

- [3] D. Slepian and J. Wolf, 'Noiseless coding of correlated information sources', IEEE Transactions on Information Theory, 1973, 19, (4), pp. 471-480
- [4] A. Wyner and J. Ziv, 'The rate-distortion function for source coding with side information at the decoder', IEEE Transactions on Information Theory, 1976, 22, (1), pp. 1-10
- [5] X. Artigas et al, 'The discover codec: architecture, techniques and evaluation', in Proceedings of Picture Coding Symposium (PCS'07), Lisboa, Portugal, Nov 2007.
- [6] A. Aron, R.U.I Zhang, and B. Girod, 'Wyner-Ziv coding of motion video', in Proceedings of the 36th Asilomar Conference on Signals Systems and Computers, California, USA, Nov. 2002, pp. 240-244
- [7] B. Girod et al, 'Distributed video coding', in Proceedings of the IEEE, 2005, 93, (1), pp. 71-83
- [8] T. Wiegand et al, 'Overview of the H.264/AVC video coding standard', IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13, (7), pp. 560-576
- [9] G.J. Sullivan et al, 'Overview of the High Efficiency Video Coding (HEVC) Standard', IEEE Transactions on Circuits and Systems for Video Technology, 2012, 22, (12), pp. 1649-1668
- [10] C. Brites and F. Pereira, 'Distributed Video Coding: Assessing the HEVC upgrade', Signal Processing: Image Communication, 2015, 32, (3), pp. 81-105
- [11] B. Benjamin, C. Jianle, L. Shan, W. Ye-Kui, "Versatile Video Coding (Draft 10)", 19th Meeting: by teleconference, Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, 22 June – 1 July 2020.
- [12] Discover DVC Final Results, http://www.img.lx.it.pt/~discover/home.html
- [13] F. Pereira et al, 'Distributed video coding: Selecting the most promising application scenarios', Signal Processing: Image Communication, 2008, 23, (5), pp. 339-352
- [14] T. Ebrahimi, F. Pereira, P. Schelkens, "JPEG Pleno: Toward an Efficient Representation of Visual Reality", Industry and Standards, IEEE Multimedia, 2016
- [15] Ren Ng, M. Levoy, et al, "Light Field Photography with a Hand-held Plenoptic Camera", Stanford Tech Report, 2005
- [16] L. Wei, Y. Zhao, and A. Wang, 'Improved side-information in distributed video coding', in Proceedings of International Conference on Innovative Computing, Information and Control, Beijing, China, Sept. 2006
- [17] Hoang Van, Quang and Dao Thi Hue, Le and Du dinh, Vien and Nguyen Hong, Vu and Hoang, Van Xiem, "Complexity Controlled Side Information Creation for Distributed Scalable Video Coding". in: International Conference on Recent Advances in Signal Processing, Telecommunications & Computing, HaNoi, Vietnam, 2019.
- [18] X. Hoang Van, J. Ascenso and F. Pereira, "Correlation modeling for a distributed scalable video codec based on the HEVC standard," 2014 IEEE 16th International Workshop on Multimedia Signal Processing (MMSP), 2014,
- [19] X. HoangVan, J. Park, and B. Jeon. "A flexible side information generation scheme using adaptive search range and overlapped block motion compensation." In Proceedings of the 5th International Conference on Ubiquitous Information Management and Communication (ICUIMC '11). Association for Computing Machinery, New York, NY, USA, 2011.
- [20] J. Ascenso, C. Brites, and F. Pereira, 'A flexible side information generation framework for distributed video coding', Multimedia Tools and Applications, 2010, 48, (3), pp. 381-409
- [21] X. HoangVan et al, 'A flexible side information generation scheme using adaptive search range and overlapped block motion compensation', in Proceeding of the 5th

International Conference on Ubiquitous Information Management and Communication, Seoul, Korea, Feb. 2011

- [22] Y.C. Shen, P.S. Wang and J.L. Wu, 'Progressive side information refinement with non-local means based denoising process for Wyner-Ziv video coding', Data Compression Conference, Snowbird, UT, Apr. 2012, pp. 219-226
- [23] H. Van Luong et al, 'Side information and noise learning for distributed video coding using optical flow and clustering', in IEEE Transactions in Image Processing, 2012, 21, (12), pp. 4782-4796
- [24] B. Dash et al, 'Decoder driven side information generation using ensemble of MLP networks for distributed video coding', Multimedia Tools and Applications, Aug. 2017, pp. 1-30
- [25] C. Brites and F. Pereira, 'Correlation noise modeling for efficient pixel and transform domain Wyner–Ziv video coding', IEEE Transactions on Circuits and Systems for Video Technology, 2008, 18, (9), pp. 1177-1190
- [26] N.M. Cheung, H. Wang and A. Ortega, 'Sampling-Based correlation estimation for distributed source coding under rate and complexity constraints', IEEE Transactions on Image Processing, 2008, 17, (11), pp. 2122-2137
- [27] Y.M. Taheri, M.O. Ahmad and M.N.S. Swamy, 'A joint correlation noise estimation and decoding algorithm for distributed video coding', Multimedia Tools and Applications, 2018, 77, (6), pp. 7327-7355
- [28] J. Slowack et al, 'Distributed video coding with feedback channel constraints', IEEE Transactions on Circuits and Systems for Video Technology, 2012, 22, (7), pp. 1014-1026
- [29] C. Yaacoub, J. Farah and B.J. Pesquet-Popescu, 'Feedback channel suppression in distributed video coding with adaptive rate allocation and quantization for multiuser applications', EURASIP Journal on Wireless Communications and Networking, 2008, 2008, (1), pp. 1-13
- [30] C. Brites and F. Pereira, 'An efficient encoder rate control solution for transform domain Wyner–Ziv video coding', IEEE Transactions on Circuits and Systems for Video Technology, 2011, 21, (9), pp. 1278-1292
- [31] J. Q. Pedro et al, 'Studying error resilience performance for a feedback channel based transform domain Wyner-Ziv video codec', in Picture Coding Symposium, Lisbon, Portugal, Nov. 2007
- [32] J. Ascenso and F. Pereira, 'Complexity efficient stopping criterion for LDPC based distributed video coding', in Proceedings of the Mobimedia, Kingston, UK, Sept. 2009, pp. 28:1-28:7
- [33] X. HoangVan and B. Jeon, 'Flexible complexity control solution for transform domain Wyner-Ziv video coding', IEEE Transactions on Broadcasting, 2012, 58, (2), pp. 209-220
- [34] Fraunhofer Heinrich Hertz Institute, "VVC Overview," [Online]. Available: https://www.hhi.fraunhofer.de/en/departments/vca/technolo gies-and-solutions/h266-vvc.html
- [35] T. Chen, 'Adaptive temporal interpolation using bidirectional motion estimation and compensation', IEEE International Conference of Image Processing, New York, USA, Sept. 2002, pp.313-316
- [36] C. K. Wong, O. C. Au, and C. W. Tang, 'Motion compensated temporal interpolation with overlapping', in proceeding of IEEE International Symposium on Circuits and Systems, May. 1996, 2, pp. 608 – 611
- [37] D. Kubasov et al, 'Optimal reconstruction in Wyner-Ziv video coding with multiple side information', IEEE Workshop on Multimedia Signal Processing, Crete, Greece, Oct. 2007, pp. 183-186
- [38] G. Bjøntegaard, "Calculation of average PSNR differences between RD-curves", Doc. ITU-T SG16 VCEG-M33, Austin, TX, USA, Apr. 2001.

A CONSISTENT LIGHT FIELD IMAGE QUALITY CONTROL BASED ON THE DISTRIBUTED VIDEO CODING AND H.266/VVC

Abstract: Distributed video coding (DVC) is a promising encoding solution for low complexity video applications such as wireless sensor networks or video surveillance systems. The requirement for image quality stability is one of the important issues in advanced display systems. However, most recent distributed video encoding solutions, which are developed based on H.264/AVC or H.265/HEVC standards, cannot provide video with stable quality. In this paper, we propose a distributed video encoding solution with quantization matrices (QMs) applied to Light Field data, ensuring stability in display quality and improving compression performance. In the proposed DVC solution, the latest video coding standard, Video Versatile Coding - VVC is selected appropriately to compress Key frames. Meanwhile, to achieve stability in video quality, quantization parameters (QPs) choose to compress key frames and quantization matrices (QMs) choose to compress Wyner-Ziv frames (WZ) will be explained in detail in the article. The test results show that the proposed distributed video encoding solution with integrated VVC codec (D-VVC) is significantly superior to other related DVC encoding solutions, especially distributed encoders. DISCOVER and recent DVC-HEVC solution, in terms of compression performance while providing stability for better picture quality for decoded video.

Keywords: Light Field coding, Distributed video coding, Image quality assessment.



Phí Công Huy tốt nghiệp đại học Oxford Brookes, UK; lấy bằng thạc sĩ tại Monash, Úc, và đang làm nghiên cứu sinh trong chương trình hợp tác giữa đại học quốc gia Hà nội và đại học công nghệ Sydney. Hiện đang công tác tại Khoa Đa phương tiện, Học viện Công nghệ Bưu chính Viễn thông, nghiên cứu chuyên sâu các lĩnh vực

liên quan tới xử lý ảnh, video, kiểm soát chất lượng dữ liệu ảnh Light Field.

Email: huypc@ptit.edu.vn; 17028025@vnu.edu.vn



Stuart Perry là giảng viên đại học UTS, giám đốc phụ trách kỹ thuật tại Pi lab, có hơn 20 năm kinh nghiệm thực hiện nghiên cứu về xử lý hình ảnh, tâm sinh lý, xử lý tín hiệu, chất lượng hình ảnh và các mô hình để định lượng mức độ ưa thích và thẩm mỹ hình ảnh cho cả chính phủ, ngành công nghiệp và học viện. Ông có bằng

Tiến sĩ tại Đại học Sydney năm 1999. Email: <u>stuart.perry@uts.edu.au</u>,



Eva Cheng là Giám đốc Hiệp hội phụ nữ trong Kỹ thuật và Công nghệ Thông tin, và Giảng viên Cao cấp của Trường Thực hành Chuyên nghiệp và Lãnh đạo tại UTS. Nền tảng nghiên cứu của cô là về xử lý tín hiệu đa phương tiện, bao gồm xử lý tín hiệu giọng nói và âm thanh và thị giác máy tính. Các dự án nghiên cứu hợp tác bao gồm làm việc với các nhạc sĩ, nghệ sĩ thị giác, kiến trúc sư, nhà sinh vật học và các tổ chức phi lợi nhuận. Email: eva.cheng@uts.edu.au.



Trịnh Anh Vũ Cử nhân Vật lý Vô tuyến (1983) tại Đại học Tổng hợp Hà Nội, Tiến sĩ Toán - lý (1994) tại Đại học Tổng hợp Matxcova (mang tên M.V.Lomonosov) tại Nga. Trao đổi nghiên cứu (năm 2002) tại Đại học Tasmania, Australia.

Hiện là Phó Giáo sư Chủ nhiệm bộ môn Thông tin vô tuyến tại Khoa Điện tử Viễn thông (FET), Trường Đại học Công nghệ

(UET), Đại học Quốc gia Hà Nội (ĐHQG Hà Nội). Thành viên hội vô tuyến điện tử Việt nam. Các lĩnh vực nghiên cứu quan tâm: Truyền thông không dây, Hệ thống truyền thông 5G, 6G. Kết nối IoT.

Email: vuta@vnu.edu.vn



Hoàng Văn Xiêm là Phó Chủ nhiệm phụ trách Bộ môn Kỹ thuật Robot, Khoa Điện tử Viễn thông, Trường Đại học Công nghệ, ĐHQGHN. TS. Xiêm hoàn thành chương trình nghiên cứu sinh tại Đại học Lisbon, Bồ Đào Nha năm 2015, chương trình cao học tại Đại học Sungkyukwan, Hàn Quốc, năm 2011, và nhận bằng kỹ sư tại Trường

ĐH Bách Khoa Hà Nội, năm 2009. TS. Xiêm là thành viên ủy ban hội đồng của trung tâm nghiên cứu và sáng tạo công nghệ ĐHQGHN-ĐH Công nghệ Sydney, Úc. Cố vấn công ty công nghệ Aimesoft. Hướng nghiên cứu quan tâm bao gồm học máy, xử lý và truyền thông ảnh, video và thị giác Robot. Email: xiemhoang@vnu.edu.vn