

NGHIÊN CỨU ẢNH HƯỞNG CỦA ĐIỂM LẤY NÉT ĐỐI VỚI CÁC MÔ HÌNH ĐÁNH GIÁ CHẤT LƯỢNG ẢNH LIGHT FIELD

Phí Công Huy^{*,#}, 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

Tóm tắt: Dữ liệu ảnh Light Field (LF) đã thu hút được sự chú ý đáng kể trong những năm gần đây do khả năng thay đổi điểm nét sau khi chụp và tái tạo hình ảnh 3D. Quá trình thể hiện và trình chiếu trên LF thường dẫn đến giảm chất lượng hình ảnh, do đó, các phương pháp đánh giá chất lượng hình ảnh đóng một vai trò quan trọng để đảm bảo trải nghiệm của người dùng. Do đó, đối với dữ liệu ảnh LF, tìm ra các phương pháp đánh giá chất lượng ảnh phù hợp nhất là vô cùng cần thiết. Trong bài báo này, chúng tôi đưa ra một phân tích đầu tiên liên quan tới hiệu năng của các phương pháp đánh giá khách quan đối với đặc tính thay đổi điểm nét của ảnh LF. Dữ liệu ảnh LF được sử dụng để đánh giá bao gồm 228 hình ảnh từ nguồn EPFL với hai mức lấy nét khác nhau (lấy nét tiền cảnh và lấy nét hậu cảnh). Kết quả được so sánh với điểm số từ phương pháp đánh giá chủ quan, nhằm so sánh sự tương quan giữa các phương pháp đánh giá khách quan với MOS. Hai phương pháp đánh giá khách quan phổ biến như PSNR, SSIM, được kiểm thử và đưa ra kết luận cụ thể trong bài báo này.

Từ khóa: Mã hóa dữ liệu Light Field, chất lượng ảnh Light Field, phương pháp đánh giá khách quan.

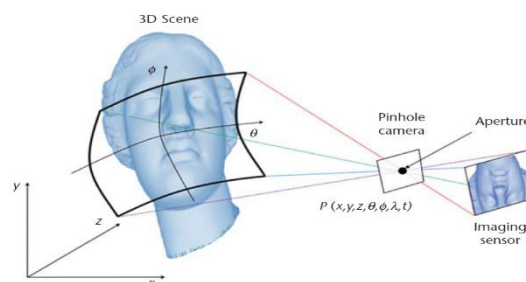
I. GIỚI THIỆU CHUNG

Hình ảnh LF là một công nghệ mới nổi cho phép chụp và mô phỏng thông tin hình ảnh ở nhiều khoảng cách tiêu cự và các vị trí của ảnh. Bằng cách thu thập bức xạ của các tia sáng từ các hướng khác nhau, LF có thể cung cấp độ sâu và góc nhìn để đưa cho người dùng trải nghiệm hình ảnh phong phú hơn mà nhiếp ảnh truyền thống không thể có.

Biểu diễn ảnh LF được thể hiện thông qua dữ liệu hình ảnh bao gồm các thông tin [1] 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 sắp 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 hình ảnh LF có dung lượng khá lớn bởi cấu trúc của nó, một bức ảnh LF có dung lượng khoảng 50MB. Để giải quyết vấn đề này, vấn đề mã hóa LF đã thu hút được sự chú ý gần đây. Các kỹ thuật xử lý khác nhau chẳng hạn như nén dữ liệu LF và các phương pháp tái tạo chế độ xem đều được sử dụng nhằm giải quyết nhu cầu lưu trữ. Nhóm nghiên cứu chuẩn hóa JPEG đã đưa ra mục tiêu nghiên cứu về nén 4D LF vào năm 2017 [2] và chuẩn hóa JPEG Pleno đã được giới thiệu, trong đó có dữ liệu ảnh LF [3].

Trong quá trình xử lý ảnh LF có thể tạo ra các yếu tố làm nhiễu loạn thị giác của người xem. Do đó, cần xem xét cẩn thận trong quá trình thiết kế các thuật toán xử lý ảnh LF để duy trì chất lượng hình ảnh. Phương pháp đánh giá chất lượng chủ quan được coi là phương pháp đáng tin cậy để đo lường chất lượng của nội dung đa phương tiện, nhưng lại tốn nhiều tài nguyên, như: tiền bạc để thiết lập thiết bị, tập trung người dùng để thử nghiệm. Ngược lại, các phương pháp đánh giá chất lượng khách quan lại có khả năng dự đoán chất lượng nhanh hơn và chúng có thể được triển khai hiệu quả trong dự đoán chất lượng hình ảnh, tuy nhiên các phương pháp này lại không phù hợp với cách tiếp nhận hình ảnh của mắt người. Sự chính xác của một phương pháp đánh giá chất lượng khách quan được đo lường dựa trên sự đồng bộ với điểm số của phương pháp chủ quan, và điểm số này được coi là điểm mốc để so sánh.

Trong khi các phương pháp đánh giá chất lượng khách quan cho dữ liệu LF vẫn đang phát triển [4-7], điều quan trọng là phải kiểm tra hiệu năng của các phương pháp đánh giá, từ đó có thể thiết lập thước đo tốt nhất để sử dụng rộng rãi cho dữ liệu LF. Nhiều phương pháp dự đoán chất lượng hình ảnh và video đã được đề xuất gần đây để áp dụng đánh giá chất lượng LF. Độ chính xác của các phương pháp này thường được đánh giá chỉ trên một tập dữ liệu duy nhất với

Tác giả liên hệ: Phí Công Huy,

Email: huyphc@ptit.edu.vn

Đến tòa soạn: 12/10/2021, chỉnh sửa: 8/11/2021, chấp nhận đăng: 29/11/2021.

một thang điểm chủ quan cụ thể như trong các đề xuất tại [8-10]. Tuy nhiên, hiệu suất của các chỉ số có thể khác nhau qua nhiều tập dữ liệu với các kiểu biến dạng và cách tính điểm khác nhau.

Trong bài báo này, mục tiêu là đánh giá một cách khách quan chất lượng của nội dung LF với điểm nét thay đổi, đồng thời kiểm thử hiệu năng của một số phương pháp đánh giá khách quan phổ biến trên bộ dữ liệu đó.

Cấu trúc của bài báo này như sau. Phần 2 tóm tắt các công trình gần đây liên quan tới phương pháp đánh giá dữ liệu Light Field. Phần 3 phân tích hiệu năng của các phương pháp đánh giá khách quan. Sau đó, Phần 4 mô tả các kết quả và kiểm thử, cuối cùng Phần 5 đưa ra một số nhận xét kết luận và các công việc trong tương lai.

II. NHỮNG NGHIÊN CỨU LIÊN QUAN

Phương pháp đánh giá chất lượng khách quan có thể cung cấp dự đoán nhanh hơn về chất lượng hình ảnh, tuy nhiên lại không có sự tin cậy cao với cảm nhận của mắt người. Do đó, dựa trên đánh giá chất lượng chủ quan, có nhiều mô hình đánh giá chất lượng hình ảnh với dữ liệu LF (LF-IQA) được giới thiệu. Tương tự như các mô hình đánh giá dữ liệu ảnh 2D, mô hình LF-IQA được định nghĩa thành ba loại, đó là mô hình tham chiếu đầy đủ (FR), mô hình tham chiếu một phần (RR) và mô hình không tham chiếu (NR). Gần đây, các đề xuất trong [11, 12] đã giới thiệu một mô hình FR để đo lường sự suy giảm chất lượng do quá trình biểu diễn LF gây ra. Các đề xuất này xem xét các đặc điểm của LF bằng cách khai thác thông tin về độ sâu, góc nhìn và điểm nhìn, ví dụ: trích xuất đặc trưng đối xứng, trích xuất đặc trưng về chiều sâu, chất lượng của toàn không gian ảnh, chất lượng cục bộ của không gian ảnh và ước tính chất lượng góc. Tương tự như vậy, trong bài báo [13], tác giả trình bày một mô hình NR khai thác các đặc điểm 2D và 3D (đó là các đặc điểm về độ sáng, màu sắc, độ bão hòa và kết cấu) của LF với sự hỗ trợ của mô hình hồi quy vector (SVR), và thông tin về độ sâu để có được dự đoán được điểm số cuối cùng. Ngoài ra, mô hình RR được đề xuất trong [14], lấy bản đồ độ sâu của LF làm ý tưởng chính. Mô hình này đo lường sự giống nhau về cấu trúc của bản đồ độ sâu giữa điểm gốc và độ biến dạng để dự đoán chất lượng của hình ảnh LF.

Tuy nhiên, các phương pháp đánh giá LF-IQA thường được coi là phức tạp bởi mô hình tính toán. Do đó, các phương pháp đánh giá truyền thống có thể là một lựa chọn khôn ngoan, chẳng hạn như PSNR [15], SSIM [16], FSIM [17], MS-SSIM [18], IW-PSNR, IW-SSIM [19], v.v. Đáng chú ý, một nghiên cứu cập nhật về hiệu suất của các phương pháp đánh giá khách quan cho LF [20] đã thử nghiệm chi tiết trên các mô hình IQA được sử dụng rộng rãi hiện nay. Trong nghiên cứu này, MAD [21], IW-SSIM [19] và GMSD [22] có hiệu năng cao nhất trong số các mô hình khác về việc đánh giá chất lượng LF, trong khi các phương pháp đánh giá rất phổ biến (như là PSNR và SSIM) chỉ là thứ hạng lần lượt là 13 và 14. Ngoài ra, công trình này cũng kết luận rằng các mô hình NR không phải là lựa chọn phù hợp cho tập dữ liệu LF. Nhưng bài báo này thiếu sự phân tích về điểm lấy nét của hình ảnh LF, nó có thể ảnh hưởng đến hiệu suất của các phương pháp đánh giá khách quan.

Dựa trên cơ sở này, cần phải thử nghiệm nhiều đánh giá và phân tích hơn về các tính năng của hình ảnh LF để cải thiện các chỉ số IQA hiện tại.

III. CÁC MÔ HÌNH ĐÁNH GIÁ CHẤT LƯỢNG VÀ BỘ DỮ LIỆU LF

A. Các phương pháp đánh giá khách quan

Trong bài báo này 2 phương pháp đánh giá khách quan được thực hiện đó là PSNR và SSIM.

Phương pháp đánh giá khách quan PSNR được sử dụng phổ biến nhất để đo chất lượng của việc tái tạo lại các tín hiệu từ các bộ nén dữ liệu có tổn thất (ví dụ: để nén hình ảnh). Tín hiệu trong trường hợp này là dữ liệu gốc và nhiễu là lỗi do nén [15]. Công thức tính PSNR được mô tả như sau

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right) \quad (1)$$

Với

$$MSE(x, y) = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (x_{ij} - y_{ij})^2 \quad (2)$$

Ở đây, x_{ij} và y_{ij} là vector các giá trị quan sát được và giá trị dự báo, và M và N là tổng số mẫu.

Còn đối với phương pháp đánh giá khách quan SSIM, nó là một phương pháp để đo mức độ giống nhau giữa hai hình ảnh thông qua dự đoán chất lượng hình ảnh gốc và hình ảnh nhiễu. Phương pháp này được coi là phù hợp với hệ thống thị giác của con người [16]. Phương pháp đánh giá khách quan SSIM được tính toán dựa trên các đặc tính của hình ảnh như: độ sáng, độ tương phản, và độ tương đồng. Công thức SSIM được tính như sau:

$$SSIM(x, y) = l_{(x,y)} c_{(x,y)} s_{(x,y)} \quad (3)$$

Với

$$\begin{aligned} l_{(x,y)} &= 2\mu_x\mu_y + C_1/\mu_x^2 + \mu_y^2 + C_1 \\ c_{(x,y)} &= 2\sigma_x\sigma_y + C_2/\sigma_x^2 + \sigma_y^2 + C_2 \\ s_{(x,y)} &= \sigma_{xy} + C_3/\sigma_x\sigma_y + C_3 \end{aligned} \quad (4)$$

B. Phương pháp xây dựng bộ dữ liệu LF

Trong nghiên cứu này, một phương pháp kết hợp kỹ thuật thể hiện hình ảnh chủ động và bị động được thực hiện, nhằm có một bộ dữ liệu đánh giá chất lượng chủ quan cho F hiệu quả nhất. Phương pháp thể hiện chủ động nghĩa là người dùng có thể chủ động lựa chọn điểm lấy nét trong ảnh LF và đưa ra điểm số. Ngược lại, phương pháp bị động, người dùng chỉ có thể xem các ảnh hoặc video có sẵn và đưa ra kết quả nhận xét.

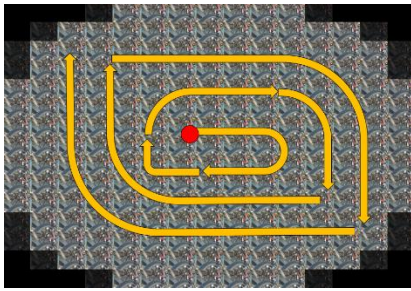
Dữ liệu LF trong bài báo này được thu thập bằng phương pháp đánh giá chủ quan thông qua phương pháp bị động, bằng kỹ thuật tạo video giả lập (PV) [23], được chọn từ 101 SAI của hình ảnh LF thông qua cách thức quét xoắn ốc như Hình 2, với 30 khung hình/giây (fps). Đồng thời, kỹ thuật thay đổi điểm nét của ảnh (RI) [24] cũng được sử dụng kết hợp.

Bảng 1. Tổng quan về định nghĩa dữ liệu LF

Loại lấy nét LF	Mô tả	Số lượng mẫu
CWF	Lớp phức tạp với FOV rộng và lấy nét vào Tiền cảnh	10
CWB	Lớp phức tạp với FOV rộng và lấy nét vào Hậu cảnh	10
CNF	Lớp phức tạp với FOV hẹp và lấy nét vào Tiền cảnh	10
CNB	Lớp phức tạp với FOV hẹp và lấy nét vào Hậu cảnh	10
SWF	Lớp đơn giản với FOV rộng và lấy nét vào Tiền cảnh	8
SWB	Lớp đơn giản với FOV rộng và lấy nét vào Hậu cảnh	8
SNF	Lớp đơn giản với FOV hẹp và lấy nét vào Tiền cảnh	10
SNB	Lớp đơn giản với FOV hẹp và lấy nét vào Hậu cảnh	10

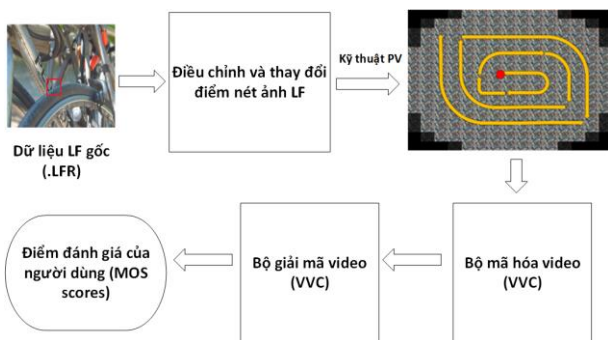
Bảng 2. Cấu hình thay đổi điểm nét cho dữ liệu LF

Tên LF	LF slope	Nét hậu cảnh	Nét tiền cảnh	Nét trung tâm	Tên LF	LF slope	Nét hậu cảnh	Nét tiền cảnh	Nét trung tâm
House_Lake	SW	-	-0.3	0	Backlight_1	CW	0.4	-0.4	-
Overexposed_Sky		0.3	-0.3	-	Bench_in_Paris		0.4	-0.4	-
Paved_Road		0.2	-0.3		Black_Fence		-	-0.4	0
Red_White_Building		-	-0.3	0	Fountain_Vince nt_2		0.6	-0.3	-
Reeds		-	-0.3	0	Friends_3		0.6	-0.3	-
Semi_reflecting_Structure_1		-	-0.3	0	Palais_du_Luxembourg		0.3	-0.6	-
Semi_reflecting_Structure_2		-	-0.3	0	Parc_du_Luxembourg		0.3	-0.6	-
Wheat_Silos		0.3	-0.6	-	Pillars		-	-0.6	0
Chain_link_Fence_1	SN	1.6	-	0	Rolex_Learning_Center	CN	0.6	-0.3	-
					Swans_1		0.4	-0.3	-
					Bikes		0.5	-0.6	-
Color_Chart_3		-	-0.3	0	Books		1	-0.7	-
Fountain_1		0.3	-0.3	-	Bridge		1	-0.7	-
Geometric_Sculpture		-	-0.5	0	Danger_de_Mort		0.6	-0.3	-
ISO_Chart_16		-	-0.3	0	Mirabelle_Prune_Tree		0.6	-0.7	-
Perforated_Metal_1		1.6	-0.5	-	Poppies		0.5	-0.7	-
Perforated_Metal_3		0.3	-0.3	-	Rusty_Fence		0.4	-0.5	-
Sewer_Drain		0.5	-0.3	-	Stone_Pillars_Inside		0.7	N/A	0
Wall_Decoration		0.7	-0.7	-	Stone_Pillars_Outside		0.4	-0.4	-
Water_Drops		0.3	-0.3	-	University		0.3	-0.4	-



Hình 2. Kỹ thuật quét hình xoắn ốc để tạo video giả lập cho ảnh LF

Thông thường, phương pháp thể hiện hình ảnh chủ động có thể cung cấp QoE tốt hơn bằng cách trao quyền kiểm soát các điểm lấy nét cho người đánh giá. Tuy nhiên, việc không kiểm soát được số lượng điểm nhìn mà mỗi người đánh giá lựa chọn là một hạn chế của phương pháp này. Trong tập dữ liệu LF của bài báo này, người dùng không thể kiểm soát các điểm lấy nét, nhưng chúng tôi cung cấp nhiều điểm lấy nét khác nhau (từ hậu cảnh đến tiền cảnh), các điều kiện độ sâu khác nhau (từ hẹp tới rộng) và độ phức tạp khác nhau của đối tượng trong mỗi ảnh LF, nhằm thu thập tối đa các đánh giá chủ quan của người dùng. Quy trình tạo ra dữ liệu ảnh LF để phục vụ đánh giá chất lượng ảnh có thể tham khảo dựa trên Hình 3.



Hình 3. Quy trình tạo dữ liệu LF để phục vụ đánh giá chất lượng ảnh

Dữ liệu LF để đánh giá được lựa chọn từ EPFL [25] gồm 38 thể loại nội dung khác nhau, với các loại độ sâu (DOF), và khung nhìn (FOV) khác nhau, được mã hóa và giải mã bằng bộ mã hóa VVC [26]. Các định nghĩa về bộ dữ liệu được mô tả trong Bảng 1. Dữ liệu này được chia thành hai phần dựa trên độ phức tạp của LF-DOF, tức là lớp DOF phức tạp và lớp DOF đơn giản (khái niệm lớp phức tạp và đơn giản được định nghĩa thông qua số lượng đối tượng mà

ảnh LF có thể lấy nét, nếu hơn 2 đối tượng được coi là lớp phức tạp). Mỗi loại DOF chứa 2 loại FOV, tức là FOV hẹp và FOV rộng. Các thông số thay đổi cấu hình của các điểm lấy nét của từng hình ảnh LF được tính toán trong dải tham số từ -0.7 cho tới 1 như trong Bảng 2. Để thay đổi điểm lấy nét của hình ảnh LF, chúng ta thay đổi tham số của độ dốc, gọi là “slope” của hình ảnh LF bằng LF Toolbox v.0.4 [27]. Có hai loại lấy nét, lấy nét tiền cảnh và lấy nét hậu cảnh. Tuy nhiên, một số hình ảnh LF chỉ có thể lấy nét với tham số độ dốc bằng 0, trường hợp này, nội dung bức ảnh sẽ quyết định đó là bức ảnh lấy nét tiền cảnh hay hậu cảnh.

IV. KẾT QUẢ VÀ KIỂM THỬ

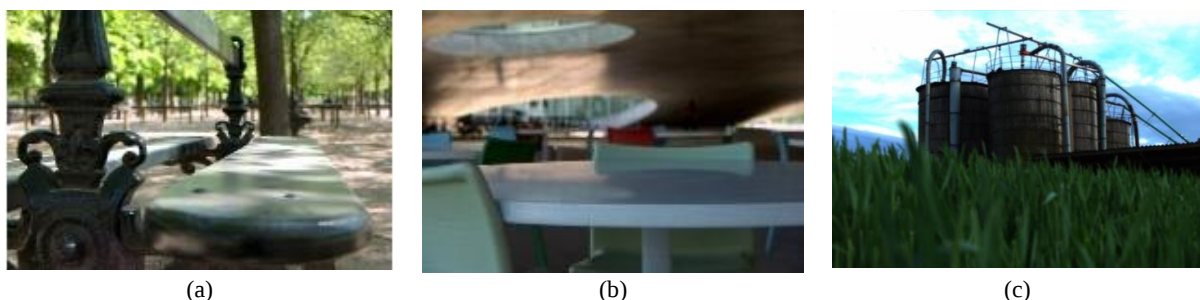
A. Phương pháp thống kê

Để có được kết quả đánh giá trực quan và chính xác nhất, các phương pháp thống kê sau đây sẽ được sử dụng nhằm thể hiện hiệu năng của các mô hình đánh giá khách quan cho dữ liệu ảnh LF. Các phương pháp thống kê gồm có, hệ số tương quan của Kendall (KRCC), hệ số tương quan của Spearman (SRCC) [28], hệ số tương quan tuyến tính Pearson (PLCC) [29], sai số tuyệt đối trung bình (MAE), lỗi bình phương trung bình (RMSE) [30] và R bình phương (R2) [31]. KRCC và SRCC được sử dụng để đánh giá và dự đoán tính đơn điệu, trong khi PLCC được sử dụng để đánh giá độ chính xác của mô hình dự đoán, MAE và RMSE được sử dụng để đánh giá tính ổn định của mô hình dự đoán và R2 được sử dụng để đánh giá chất lượng tương quan của mô hình dự đoán.

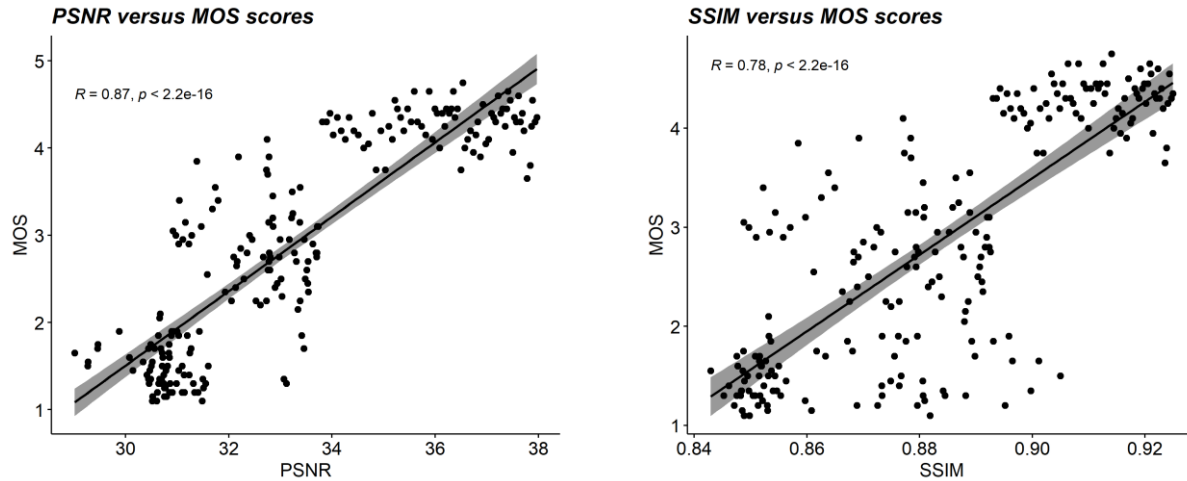
Chỉ số cho các phương pháp thống kê được xác định như sau: hiệu năng được coi là tốt nếu kết quả KRCC, SRCC, PLCC, R2 có chỉ số cao và ngược lại, đối với MAE, RMSE có chỉ số thấp, phạm vi giá trị của các phương pháp thống kê được đo từ 0 đến 1.

B. Đánh giá hiệu năng tổng thể dựa trên sự tương quan

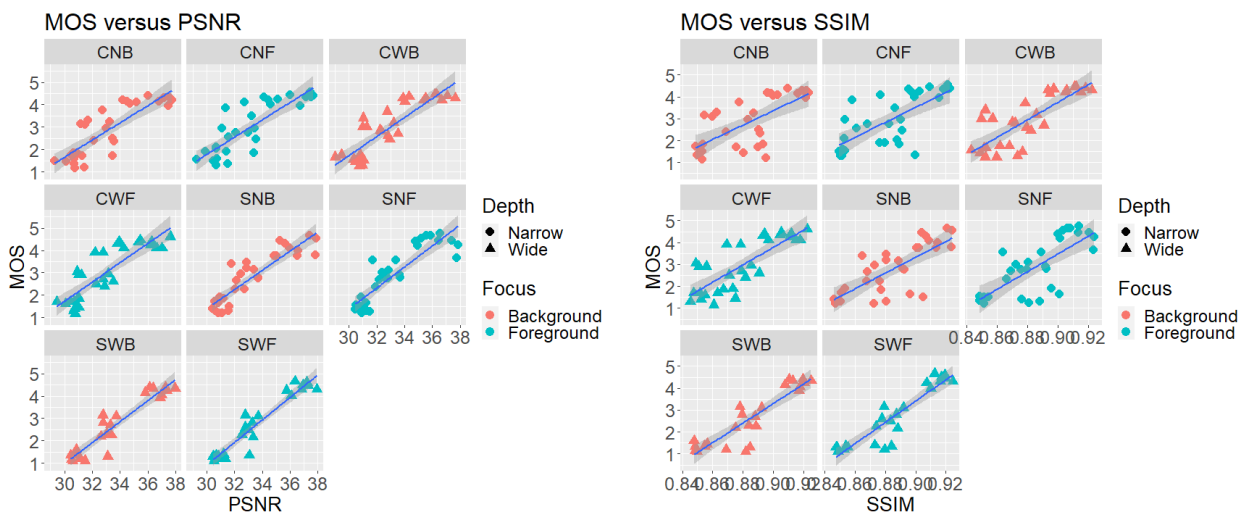
Hình 5 cho thấy khả năng tương quan của 2 phương pháp đánh giá khác nhau SSIM và PSNR so với điểm trung bình của cách đánh giá chủ quan MOS trên toàn bộ dữ liệu LF. Hiệu năng của các phương pháp đánh giá khách quan được tính bằng R bình phương với phương pháp thống kê PLCC, cụ thể là 87% và 78% tương ứng với PSNR và SSIM. Ngoài ra, hiệu năng của SSIM và PSNR so với MOS đối với các điều kiện khác nhau được trình bày trong Hình 6, nó thể hiện giá trị của PSNR và SSIM có sự trải rộng và thay đổi đối với các dữ liệu của LF. Do đó, điều này kết luận rằng hai phương pháp đánh giá khách quan PSNR và SSIM hiện không phù hợp cho sử dụng đối với dữ liệu LF, nhưng đây cũng là cơ hội để phát triển các mô hình mới



Hình 4. Hình ảnh của LF sử dụng để đánh giá kiểm thử, trong đó: a) Bench in Paris, b) Rolex Learning Center, và c) Wheat Silos



Hình 5. Sự tương quan của phương pháp đánh giá SSIM và PSNR so với MOS



Hình 6. Sự tương quan của SSIM và PSNR so với MOS trong các điều kiện khác nhau

phục vụ cho dữ liệu LF dựa trên các phương pháp đánh giá trên.

Ngoài ra, hiệu năng của 2 phương pháp đánh giá SSIM và PSNR được so sánh tại Bảng 3. Kết quả cho thấy khá thú vị, đó là với dữ liệu LF thay đổi điểm nét, thì hiệu năng của PSNR đều cho kết quả tốt hơn so với SSIM, mặc dù SSIM luôn được đánh giá là phương pháp thể hiện tốt đối với thị giác con người. Điều này có thể giải thích rằng, các đặc trưng ảnh hưởng tới thị giác mắt người trong phương pháp SSIM được thiết kế để đánh giá cho các dữ liệu của ảnh tự nhiên thông thường, chứ không phục vụ được cho

dữ liệu ảnh LF. Đối với phương pháp đánh giá PSNR, kỹ thuật của nó đơn giản chỉ trừ hai ảnh để tìm điểm khác nhau, nên không bị ảnh hưởng bởi dữ liệu thiết kế cho ảnh tự nhiên thông thường. Do vậy, PSNR có thể là điểm nghiên cứu khởi đầu để phát triển các mô hình đánh giá dữ liệu ảnh LF thay vì dựa trên phương pháp SSIM như trước đây.

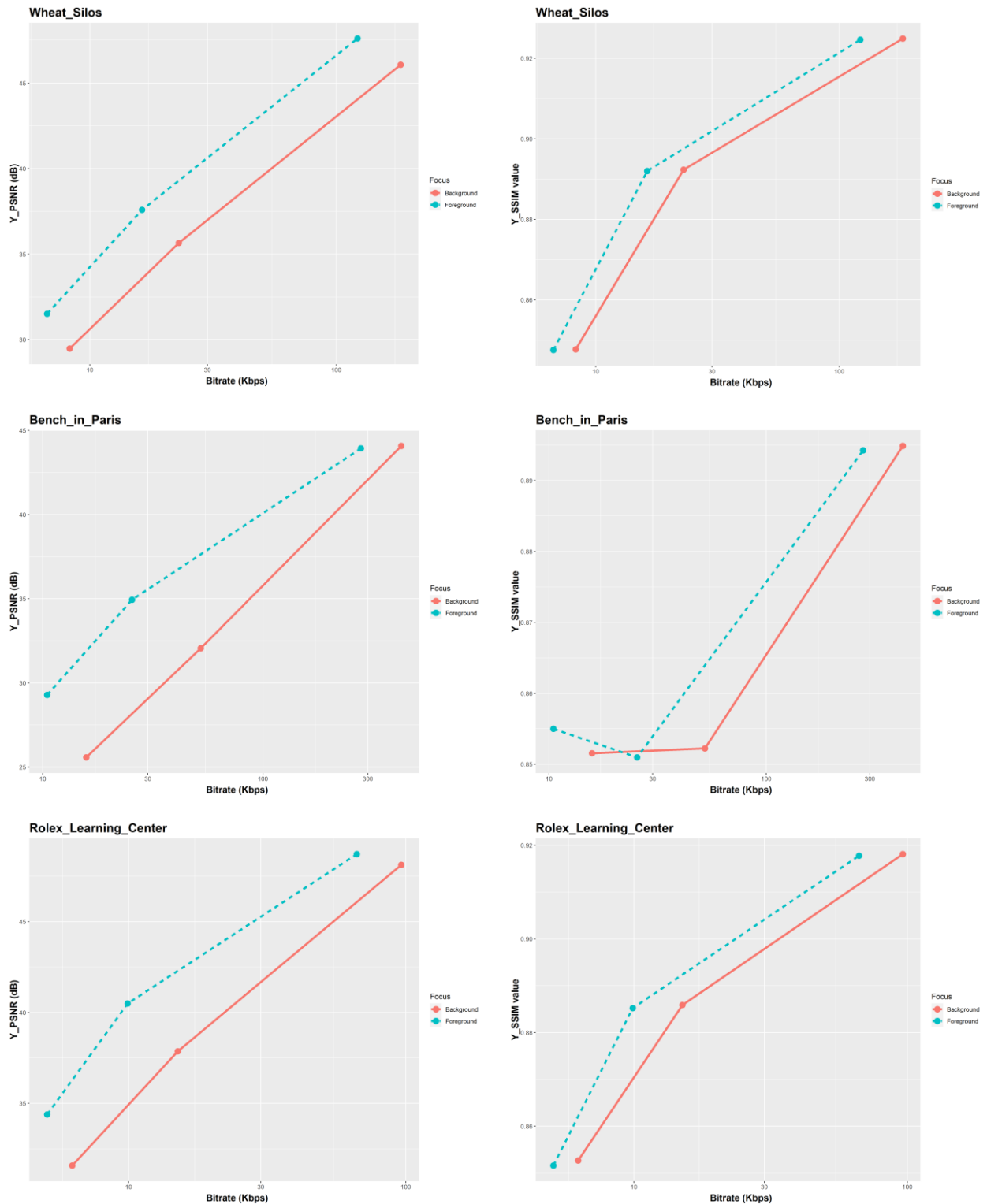
C. Đánh giá hiệu năng dựa trên điều kiện lấy nét

Trong hình 7, phương pháp đánh giá SSIM và PSNR được so sánh dựa trên các điểm nét trước và sau của dữ liệu LF có tên là Wheat Silos, Bench in Paris, và Rolex Learning Center (như Hình 4).

Kết quả thể hiện cho thấy, điều kiện lấy nét tại tiền cảnh (foreground) luôn cho hiệu năng tốt hơn với điều kiện lấy nét hậu cảnh (background) ở cả 2 phương pháp đánh giá SSIM hay PSNR. Điều này chứng tỏ rằng, dữ liệu LF với các ảnh được lấy nét ở tiền cảnh thể hiện dữ liệu tốt hơn như: hình ảnh rõ nét, chủ thể có kích thước đủ lớn, hoặc đơn giản là xu hướng người xem thích chú

Bảng 3. So sánh hiệu năng giữa SSIM và PSNR đối với dữ liệu LF

Phương pháp thống kê	Phương pháp đánh giá	
	PSNR	SSIM
KRCC	0.6328	0.5583
SRCC	0.8509	0.7659
PLCC	0.8661	0.7761
RMSE	0.5977	0.7541
MAE	0.4797	0.5690
R2	0.7502	0.6023



Hình 7. Hiệu năng của SSIM và PSNR so với các điểm nét trước và nét sau

ý những vật thể ở gần mình hơn. Thêm nữa, phương pháp đánh giá PSNR một lần nữa cho thấy hiệu năng thể hiện tốt hơn SSIM tại bộ dữ liệu LF này, thông qua khoảng cách giữa tiền cảnh và hậu cảnh lớn hơn như thể hiện tại hình 7.

Dựa vào đây, ta có thể đưa ra kết luận rằng, phương pháp đánh giá PSNR sẽ là lựa chọn tốt nhất so với SSIM

khi áp dụng để đánh giá dữ liệu ảnh LF, đặc biệt sử dụng để đánh giá hiệu năng lấy nét tiền cảnh.

V. KẾT LUẬN

Bài báo này trình bày một đánh giá đầy đủ và chi tiết về hiệu năng của hai mô hình đánh giá SSIM và PSNR đối với bộ dữ liệu LF, đặc biệt khả năng thay đổi điểm nét của bộ dữ liệu này. Với phương pháp xây dựng bộ dữ

liệu LF đặc thù, nhằm thu thập kết quả đánh giá chủ quan đầy đủ nhất từ người dùng, 228 dữ liệu với các điểm lấy nét và các điều kiện khác nhau đã được sử dụng. Độ chính xác của các phương pháp đánh giá khách quan được so sánh với điểm số chủ quan ứng với mỗi phương pháp. Hơn nữa, thống kê sự tương quan và hiệu năng của các phương pháp với MOS đã được so sánh và thảo luận. Từ những kết quả đánh giá, cho thấy phương pháp đánh giá PSNR sẽ phù hợp để sử dụng hơn SSIM cho bộ dữ liệu LF.

Dựa trên kết quả nghiên cứu này, chúng tôi có thể phát triển thêm với các mô hình đánh giá khách quan khác nhau, nhằm đưa ra một phương pháp đánh giá khách quan tốt nhất sử dụng cho bộ dữ liệu LF, đồng thời có thể xây dựng được một mô hình dự đoán chính xác hơn trong tương lai.

TÀI LIỆU THAM KHẢO

- [1] T. Ebrahimi, F. Pereira, P. Schelkens, "JPEG Pleno: Toward an Efficient Representation of Visual Reality", *Industry and Standards*, IEEE Multimedia, 2016
- [2] ISO/IEC JTC1/SC29/WG1, "JPEG Pleno Call for Proposals on Light Field Coding," wG1N74014, in: 74th JPEG Meeting, Geneva, Switzerland, 2017.
- [3] P. Schelkens, et al, "JPEG Pleno: a standard framework for representing and signaling plenoptic modalities", in A.G. Tescher (Ed.), *Proceedings of SPIE Optics Photonics: Applications of Digital Image Processing XLI*, in: *Applications of Digital Image Processing XLI*, vol. 10752, SPIE, United States, p. 107521P, 2018.
- [4] L. Shi, W. Zhou, Z. Chen, J. Zhang, "No-reference light field image quality assessment based on spatial-angular measurement," *IEEE Trans. Circuits Syst. Video Technol.* Vol. 1, 2019 <http://dx.doi.org/10.1109/tcsvt.2019.2955011>.
- [5] Z. Luo, W. Zhou, L. Shi, Z. Chen, "No-reference light field image quality assessment based on micro-lens image," in 2019 *Picture Coding Symposium, PCS*, pp. 1–5, 2019
- [6] Y. Tian, H. Zeng, J. Hou, J. Chen, K. Ma, "Light field image quality assessment via the light field coherence," *IEEE Trans. Image Process.* Vol. 29, pp. 7945–7956, 2020
- [7] W. Zhou, L. Shi, Z. Chen, J. Zhang, *Tensor oriented no-reference light field image quality assessment*, *IEEE Trans. Image Process.* Vol. 29, pp. 4070–4084, 2020.
- [8] V. Kiran Adhikarla, M. Vinkler, D. Sumin, R.K. Mantiuk, K. Myszkowski, H.-P. Seidel, P. Didyk, "Towards a quality metric for dense light fields," in: *IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2017
- [9] C. Perra, P. Astola, E.A. da Silva, H. Khanmohammad, C. Pagliari, P. Schelkens, I. Tabus, "Performance analysis of JPEG pleno light field coding," in: *Applications of Digital Image Processing XLII*, SPIE, 2019.
- [10] H. Amirpour, A.M. Pinheiro, M. Pereira, M. Ghanbari, "Reliability of the most common objective metrics for light field quality assessment", in: *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP*, IEEE, pp. 2402–2406, 2019
- [11] X. Min, J. Zhou, G. Zhai, P. Le Callet, X. Yang and X. Guan, "A Metric for Light Field Reconstruction, Compression, and Display Quality Evaluation," in *IEEE Transactions on Image Processing*, vol. 29, pp. 3790–3804, 2020
- [12] Y. Tian, H. Zeng, J. Hou, J. Chen, J. Zhu and K. Ma, "A Light Field Image Quality Assessment Model Based on Symmetry and Depth Features," in *IEEE Transactions on Circuits and Systems for Video Technology*, doi: 10.1109/TCSVT.2020.2971256
- [13] L. Shan, P. An, C. Meng, X. Huang, C. Yang and L. Shen, "A No-Reference Image Quality Assessment Metric by Multiple Characteristics of Light Field Images," in *IEEE Access*, vol. 7, pp. 127217–127229, 2019
- [14] I. Viola, M. Re`r`abek, and T. Ebrahimi, "A new approach to subjectively assess quality of plenoptic content," in *SPIE Optical Engineering and Applications. International Society for Optics and Photonics*, pp. 99710X–99710X, 2016
- [15] A. Horé and D. Ziou, "Image Quality Metrics: PSNR vs. SSIM," 2010 20th International Conference on Pattern Recognition, pp. 2366–2369, 2010.
- [16] Z. Wang, A. Bovik, H. Sheikh, E. Simoncelli, "Image quality assessment: From error visibility to structural similarity", *IEEE Trans. Image Process.* vol. 13, no. 4, pp. 600–612, 2004
- [17] L. Zhang, L. Zhang, X. Mou, D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.* vol. 20, no. 8, pp. 2378–2386, 2011
- [18] Z. Wang, E.P. Simoncelli, A.C. Bovik, "Multiscale structural similarity for image quality assessment," in: *The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers*, Vol. 2, IEEE, pp. 1398–1402, 2003
- [19] Z. Wang, Q. Li, "Information content weighting for perceptual image quality assessment," *IEEE Trans. Image Process.* vol. 20, no. 5, pp. 1185–1198, 2010
- [20] S. Mahmoudpour, P. Schelkens, "On the performance of objective quality metrics for light fields," *Signal Processing: Image Communication*, vol. 93, pp. 116179, 2021
- [21] E.C. Larson, D.M. Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy," *J. Electron. Imaging*, vol. 19, no. 1, pp. 011006, 2010
- [22] W. Xue, L. Zhang, X. Mou, A.C. Bovik, "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index", *IEEE Trans. Image Process.* vol. 23, no. 2, pp. 684–695, 2013
- [23] F. Battisti, M. Carli, and P. L. Callet, "A study on the impact of visualization techniques on light field perception," in 2018 26th European Signal Processing Conference (EUSIPCO), pp. 2155–2159, 2018
- [24] P. Paudyal, F. Battisti, M. Sjöström, R. Olsson, and M. Carli, "Towards the perceptual quality evaluation of compressed light field images," *IEEE Trans. Broadcast.*, vol. 63, no. 3, pp. 507522, Sep. 2017
- [25] M. Re`r`abek and T. Ebrahimi, "New Light Field Image Dataset," 8th International Conference on Quality of Multimedia Experience, Lisbon, Portugal, 2016
- [26] Fraunhofer Heinrich Hertz Institute, "VVC Overview," [Online]. Available: <https://www.hhi.fraunhofer.de/en/departments/vca/technologies-and-solutions/h266-vvc.html>
- [27] D. Dansereau, "Light Field Toolbox v0.4," <http://www.mathworks.com/matlabcentral/fileexchange/49683-light-field-toolbox-v0-4>
- [28] H. R. Sheikh, M. F. Sabir, and A. C. Bovik. A statistical evaluation of recent full reference image quality assessment algorithms. *IEEE Trans. Image Process.*, 15(11):3440–3451, 2006
- [29] "Pearson's Correlation Coefficient". In: Kirch W. (eds) *Encyclopedia of Public Health*. Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-5614-7_2569, 2008.
- [30] Wang, Weijie & Lu, Yanmin. "Analysis of the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE)", in *Assessing Rounding Model. IOP Conference Series: Materials Science and Engineering*. 324. 012049. 10.1088/1757-899X/324/1/012049, 2018.
- [31] Figueiredo, Dalson & Júnior, Silva, & Rocha, Enivaldo. "What is R2 all about?." *Leviathan-Cadernos de Pesquisa Política*. Vol. 3. pp. 60–68, 2011.

FOCUS POINT BASED EFFECTS ON LIGHT FIELD IMAGE QUALITY ASSESSMENTS

Abstract: Light Field (LF) image data has attracted considerable attention in recent years due to its ability to change focus point after capture and reproduce 3D images. The process of rendering and displaying on LF often leads to a loss of image quality, so image quality assessment methods play an important role in ensuring user experience. Therefore, for LF image data, finding the most suitable image quality assessment methods is extremely necessary. In this paper, we present a first analysis related to the performance of objective evaluation methods for the sharpness change characteristics of LF images. The LF image data used for the evaluation consisted of 228 images from the EPFL source with two different focus levels (foreground focus and background focus). The results are compared with the scores from the subjective assessment method, to compare the correlation between the objective assessment methods and the MOS. Two popular objective assessment methods such as PSNR, SSIM, are tested and given specific conclusions in this paper.

Keywords: Light Field image, Objective quality assessment, Light Field coding.



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.



TS. 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.