

Nâng cao chất lượng ảnh bằng phương pháp trộn ảnh từ hai nguồn ảnh khác nhau

Lê Khánh Thành, Lê Bá Tuấn, Vũ Quốc Huy*

Viện Tự động hóa Kỹ thuật quân sự, Viện Khoa học và Công nghệ quân sự, 89B Lý Nam Đế, Cửa Đông, Hoàn Kiếm, Hà Nội, Việt Nam.

*Email: maihuyvu@gmail.com

Nhận bài: 26/12/2023; Hoàn thiện: 13/3/2024; Chấp nhận đăng: 14/3/2024; Xuất bản: 01/4/2024.

DOI: <https://doi.org/10.54939/1859-1043.j.mst.CAPITI.2024.161-167>

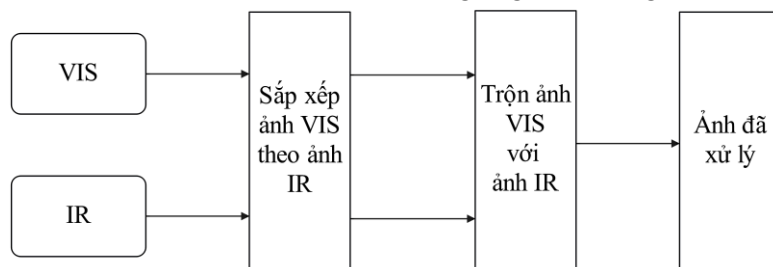
TÓM TẮT

Bài báo mô tả một phương pháp nâng cao chất lượng ảnh hồng ngoại bằng cách trộn ảnh từ hai nguồn ảnh camera thường và ảnh hồng ngoại cùng quan sát một cảnh. Hai ảnh từ hai camera khác nhau không có cấu trúc dữ liệu thống nhất: độ phân giải khác nhau, trường nhìn khác nhau (FOV) và điều chỉnh ống kính khác nhau. Để thực hiện việc nâng cao chất lượng một nguồn ảnh, trước tiên các hình ảnh phải được biến đổi về cùng điều kiện. Sau đó, hai ảnh được trộn với nhau để nâng cao chất lượng ảnh hồng ngoại. Kết quả thực nghiệm cho thấy tính hiệu quả của phương pháp khi chất lượng ảnh hồng ngoại được nâng cao.

Từ khoá: Thị giác máy; Xử lý ảnh; Khớp ảnh; Cải thiện chất lượng ảnh.

1. ĐẶT VẤN ĐỀ

Nâng cao chất lượng ảnh [1-5] thường là bước tiền xử lý cho các quá trình xử lý tiếp theo. Nó đóng một vai trò quan trọng trong kết quả xử lý cuối cùng. Việc nâng cao chất lượng ảnh thường áp dụng các thuật toán tăng cường trên một ảnh gốc để cho ra một ảnh có chất lượng tốt hơn. Trong hệ thống quang điện tử có sử dụng đa cảm biến hình ảnh như các đài chỉ huy trung tâm của các hệ thống phòng pháo phòng không, các hệ thống này thường dùng kết hợp camera thường (VIS) và camera hồng ngoại (IR), việc tăng cường chất lượng ảnh hiện đang được thực hiện trên mỗi cảm biến ảnh. Ảnh VIS có thể lưu trữ hiệu quả các thông tin chi tiết của cảnh nhưng chúng cũng dễ bị ảnh hưởng bởi các yếu tố như sự thay đổi của cường độ sáng và thời tiết khắc nghiệt. Ảnh IR thu lại bức xạ nhiệt phát ra trong cảnh, có thể phân biệt hiệu quả mục tiêu và hậu cảnh trong cảnh, đồng thời không bị ảnh hưởng bởi vật cản và thời tiết khắc nghiệt. Tuy nhiên, ảnh IR không chứa thông tin chi tiết của cảnh như ảnh VIS, đây là thông tin quan trọng trong các thuật toán bám bắt mục tiêu bằng xử lý ảnh. Vì thế, việc nâng cao chất lượng ảnh IR của hệ thống quang điện tử đài chỉ huy trung tâm hệ thống phòng pháo phòng không là vấn đề khoa học cần được giải quyết. Nghiên cứu [16] chỉ ra rằng có thể kết hợp các đặc trưng ảnh nổi trội từ mỗi cảm biến ảnh để tạo ra một ảnh có chất lượng tốt hơn. Để thực hiện tổng hợp các ảnh đúng cách, điều quan trọng là phải đảm bảo rằng thông tin từ mỗi cảm biến thỏa mãn có đặc điểm giống nhau trong môi trường [6-7].



Hình 1. Quá trình nâng cao ảnh IR.

Hai nguồn ảnh mà bài báo sử dụng là ảnh VIS và ảnh IR được thu từ hai camera khác nhau, chúng có chức năng thu nhận thông tin và hệ quang học khác nhau, ảnh thu được sẽ khác nhau. Vì vậy, để trộn hai ảnh từ hai camera này, đầu tiên, các ảnh này phải được sắp xếp sao cho chúng

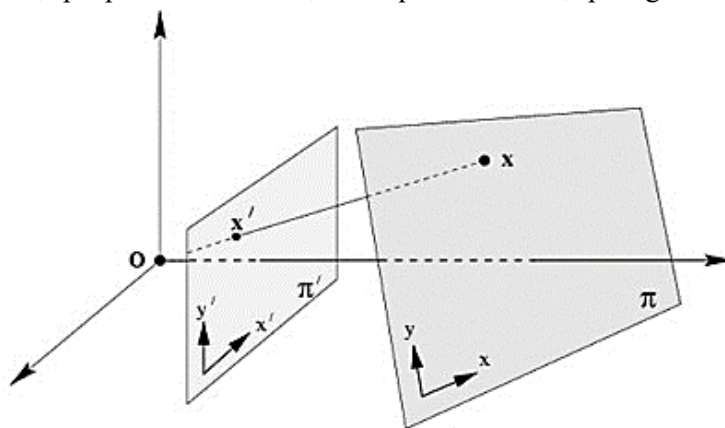
giống nhau về trường nhìn (FOV), độ phân giải, không gian màu. Sau khi ảnh VIS và ảnh IR đã được biến đổi về cùng các điều kiện yêu cầu, các ảnh này sẽ được sử dụng để trộn và nâng cao chất lượng ảnh IR. Bài báo này sử dụng kết hợp độ tương phản của ảnh VIS vào trong ảnh IR.

2. THUẬT TOÁN SẮP XẾP VÀ TRỘN ẢNH

2.1. Biến đổi và sắp xếp ảnh

Trong xử lý ảnh số, ảnh có thể được biến đổi theo các biến đổi khác nhau như biến đổi tương quan, biến đổi afin và biến đổi phối cảnh. Các phép biến đổi ảnh này đã được nghiên cứu bởi nhiều công trình nghiên cứu [8-10]. Trong trường hợp biến đổi hai ảnh từ hai cảm biến ảnh khác nhau của hệ quang điện tử đài chỉ huy trung tâm hệ thống pháo phòng không cùng quan sát một đối tượng về cùng trường nhìn và độ phân giải, quá trình biến đổi này được gọi là sắp xếp ảnh. Sắp xếp ảnh là quá trình căn chỉnh các hình ảnh được thu từ các cảm biến khác nhau hoặc từ các góc nhìn khác nhau sao cho tất cả các điểm tương ứng trong các ảnh đều khớp nhau. Một phép biến đổi phải được xác định để liên kết các điểm trong một hình ảnh với các điểm tương ứng của chúng trong một hình ảnh khác. Sự biến đổi này phụ thuộc vào đặc điểm của sự khác biệt giữa các hình ảnh cần biến đổi và được tính toán dựa trên hình ảnh tham chiếu hoặc hình ảnh cơ sở. Nếu lấy cảm biến ảnh thứ nhất là gốc (ảnh IR trong bài báo) thì ảnh từ cảm biến ảnh thứ hai (ảnh VIS trong bài báo) được coi là đã bị biến đổi (biến dạng) so với ảnh từ cảm biến gốc.

Phép biến đổi phối cảnh (perspective transform hay còn gọi là homography) từ ảnh đã bị biến đổi về ảnh gốc là một phép biến đổi hình học liên quan tới hai mặt phẳng khác nhau.



Hình 2. Quan hệ không gian giữa hai mặt phẳng ảnh.

Quan sát trong hình 2, mỗi điểm $f_X = (x, y)$ của mặt phẳng π liên quan tới điểm $f_{X'} = (x', y')$ của mặt π' như mô tả trong công thức (1) trên hệ tọa độ đồng nhất.

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \mathbf{H} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (1)$$

Trong đó, \mathbf{H} là ma trận 3×3 đã chuẩn hóa với $h_{33} = 1$ hay $h_{11}^2 + h_{12}^2 + h_{13}^2 + h_{21}^2 + h_{22}^2 + h_{23}^2 + h_{31}^2 + h_{32}^2 + h_{33}^2 = 1$.

Để xác định ma trận \mathbf{H} , cần xác định các điểm tương đồng giữa hai ảnh (tồn tại giữa hai ảnh). Các điểm tương đồng được xác định dựa trên các điểm đặc trưng của hai ảnh. Việc trích xuất các đặc trưng của hai ảnh có thể được thực hiện bằng một trong các thuật toán SHIFT, SURF hay ORB [11-13]. Thuật toán [14] được sử dụng để tìm ra các điểm tương đồng giữa hai ảnh từ các điểm đặc trưng. Giả sử $(x_1, y_1), \dots, (x_n, y_n)$ là tọa độ các điểm tương đồng ở ảnh gốc và $(x'_1, y'_1), \dots, (x'_n, y'_n)$ là tọa độ của các điểm tương đồng ở ảnh cần biến đổi. Để thực hiện ước lượng ma trận \mathbf{H} , ta cần tính hệ phương trình (2).

$$Ah = \begin{bmatrix} -x'_1 & -y'_1 & -1 & 0 & 0 & 0 & x_1x'_1 & x_1y'_1 & x_1 \\ 0 & 0 & 0 & -x'_1 & -y'_1 & -1 & y_1y'_1 & y_1y'_1 & y_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -x'_n & -y'_n & -1 & 0 & 0 & 0 & x_nx'_n & x_ny'_n & x_n \\ 0 & 0 & 0 & -x'_n & -y'_n & -1 & y_ny'_n & y_ny'_n & y_n \end{bmatrix} h = 0 \quad (2)$$

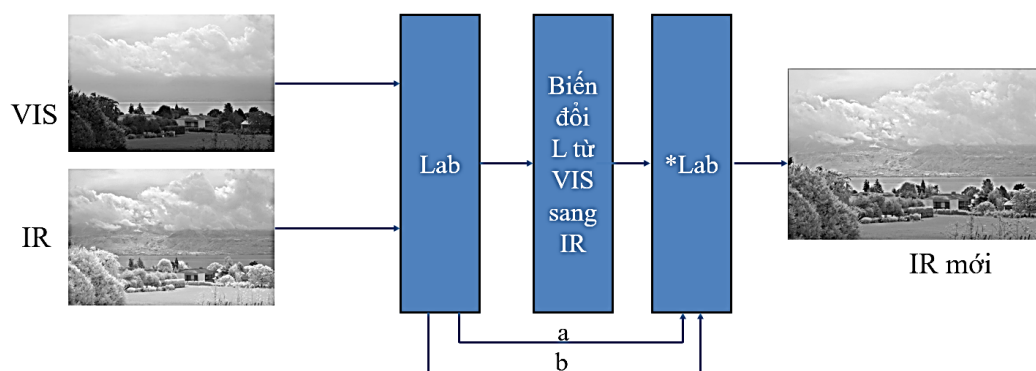
Trong đó, $h = [h_{11} \ h_{12} \ h_{13} \ h_{21} \ h_{22} \ h_{23} \ h_{31} \ h_{32} \ h_{33}]^T$

Việc giải phương trình (2) được thực hiện bởi giải thuật RANSAC [15]. Sau khi ước lượng được ma trận H, ma trận H sẽ được chuẩn hóa để đảm bảo tránh các vấn đề như thay đổi tỷ lệ biến dạng mong muốn trong ảnh kết quả, giảm ảnh hưởng của nhiễu và sai số trong dữ liệu điểm tương đồng, đặc biệt là khi sử dụng [15] để loại bỏ các điểm ngoại lệ. Sau khi xác định được ma trận H, sử dụng công thức (1) để biến đổi ảnh bị biến dạng về ảnh gốc.

Sau khi ảnh VIS và IR được đồng bộ về trường nhìn và độ phân giải, thông tin vị trí mỗi điểm ảnh ở hai ảnh VIS và IR là tương đồng, thỏa mãn điều kiện bài toán trộn ảnh.

2.2. Trộn ảnh VIS và ảnh IR

Độ tương phản của ảnh IR thường thấp hơn so với ảnh được chụp bằng ánh sáng nhìn thấy. Lí do là vì phần lớn các vật thể trong ảnh IR có xu hướng phản chiếu ánh sáng hồng ngoại một cách tương đối đồng đều, dẫn đến sự thiếu đối lập giữa các vùng trong ảnh. Điều này có thể dẫn đến việc mất đi một số chi tiết và cấu trúc trong ảnh, làm giảm sự phân biệt giữa các đối tượng. Để cải thiện chất lượng ảnh IR, cần cải thiện độ tương phản của nó. Thực nghiệm cho thấy, nếu chỉ tăng giảm độ tương phản chung của riêng ảnh IR sẽ không phản ánh đúng độ tương phản cục bộ giữa các vùng trong ảnh. Ảnh VIS chứa đựng thông tin độ tương phản giữa các vùng trong ảnh tốt hơn ảnh IR, vì thế để cải thiện độ tương phản của ảnh IR, trong bài báo này sẽ sử dụng kết hợp độ tương phản của ảnh VIS vào trong ảnh IR. Trong các không gian màu, không gian màu Lab là không gian màu thể hiện rõ sự tương phản của ảnh. Trong không gian màu này, kênh L thể hiện độ sáng của của màu sắc trong khi kênh a và b thể hiện thông tin về màu sắc và sắc độ của màu. Không gian Lab sẽ được lựa chọn là không gian màu trung gian của ảnh VIS và IR. Quá trình trộn ảnh VIS và ảnh IR được mô tả như trong hình 3.



Hình 3. Quá trình trộn ảnh VIS với ảnh IR.

Để thực hiện việc trộn ảnh và nâng cao chất lượng ảnh IR, đầu tiên, ảnh IR thu được là ảnh mức xám được chuyển sang thành ảnh RGB trong đó mỗi điểm ảnh RGB được tính theo công thức (3).

$$P_{rgb} = (I_{gray}, I_{gray}, I_{gray}) \quad (3)$$

Tiếp theo, ảnh VIS và ảnh IR được biến đổi sang không gian màu Lab. Ảnh VIS được chuyển sang không gian Lab theo công thức (4):

$$\begin{bmatrix} L \\ a \\ b \end{bmatrix} = \begin{bmatrix} 116fY - 16 \\ 500 \left[f\left(\frac{X}{0.95047}\right) - f(Y) \right] \\ 200 \left[f(Y) - f\left(\frac{Z}{1.08883}\right) \right] \end{bmatrix} \quad (4)$$

Trong đó, X, Y, Z được tính như trong công thức (5).

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4124564 & 0.3575761 & 0.1804375 \\ 0.2126729 & 0.7151522 & 0.0721570 \\ 0.0193339 & 0.1191920 & 0.9503041 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (5)$$

Và f được tính như trong công thức (6):

$$f(t) = \begin{cases} t^{\frac{1}{3}} & \text{nếu } t \geq \left(\frac{6}{29}\right)^3 \\ \frac{1}{3}\left(\frac{29}{6}\right)^2 + \frac{4}{29} & \text{nếu } t \leq \left(\frac{6}{29}\right)^3 \end{cases} \quad (6)$$

Mục đích của việc chuyển đổi ảnh VIS và ảnh IR sang không gian Lab là để lấy về độ sáng (kênh L, L càng lớn thì điểm ảnh càng sáng và ngược lại) và độ sắc của ảnh (kênh a và b). Sau khi chuyển đổi sang ảnh Lab, thay thế kênh L của ảnh IR với kênh L của ảnh VIS. Cuối cùng, chuyển đổi ảnh IR từ không gian màu Lab về không gian màu RGB theo công thức (7).

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 3.2404542 & -1.5371385 & -0.4985314 \\ -0.9692660 & 1.8760108 & 0.0415560 \\ 0.0556434 & -0.2040259 & 1.0572252 \end{bmatrix} \begin{bmatrix} \frac{L+16}{116} + \frac{a}{500} \\ \frac{L+16}{116} \\ \frac{L+16}{116} - \frac{b}{200} \end{bmatrix} \quad (7)$$

Ảnh IR thu được sau khi chuyển từ không gian màu Lab về RGB là ảnh đã được xử lý chất lượng. Quá trình thực hiện nâng cao chất lượng ảnh bằng việc trộn ảnh từ hai nguồn hình ảnh khác nhau được thực hiện theo các bước như sau:

Bước 1: Đọc các ảnh từ camera;

Bước 2: Trích xuất điểm đặc trưng của mỗi ảnh theo [11-13];

Bước 3: Khớp các điểm tương đồng giữa hai ảnh theo [14] (chọn ra các điểm giống nhau giữa hai ảnh);

Bước 4: Xác định ma trận H và thực hiện biến đổi ảnh theo công thức (1);

Bước 5: Trộn hai ảnh sau khi đã được biến đổi để thu được ảnh cuối cùng.

3. KẾT QUẢ THỰC NGHIỆM

Trước khi phân tích kết quả xử lý ảnh IR, cần lựa chọn phương pháp để đánh giá. Để đánh giá hiệu quả và chất lượng của ảnh được xử lý, bài báo sử dụng phương pháp tính entropy để đánh giá chất lượng ảnh. Phương pháp tính entropy thường được sử dụng để đánh giá mức độ đa dạng hoặc phức tạp của thông tin trong hình ảnh kết quả sau khi thực hiện việc kết hợp các hình ảnh từ nhiều nguồn khác nhau. Giá trị entropy lớn thể hiện ảnh thu được đa dạng thông tin, có nhiều chi tiết hơn so với ảnh gốc. Entropy bậc nhất tương ứng với entropy toàn cục để phân ngưỡng ảnh mức xám. Tuy nhiên, entropy bậc một có một nhược điểm là nó không tính đến mối tương quan không gian của hình ảnh như với entropy bậc 2. Vì thế entropy bậc 2 tính theo (8) sẽ được sử dụng trong đo lường chất lượng ảnh, với p_i là xác suất xuất hiện giá trị màu thứ i trong ảnh và được tính theo công thức (9). N là tổng số điểm ảnh có trong ảnh.

$$Entropy = - \sum_{i=1}^N p_i \log_2(p_i) \quad (8)$$

$$p_i = \frac{\text{Số lần xuất hiện của giá trị điểm ảnh } i}{N} \quad (9)$$



Hình 4. Ảnh thường (trái), ảnh hồng ngoại (giữa), ảnh đã xử lý (phải).

Nhóm tác giả thực nghiệm nâng cao chất lượng ảnh bằng cách trộn hai ảnh từ hai nguồn ảnh khác nhau với hai nguồn ảnh từ camera thường và ảnh hồng ngoại có độ phân giải 640x480. Phần mềm kiểm tra được viết bằng ngôn ngữ C++ với Qt IDE và chạy trên máy tính tính Dell Vostro Core i7-4510U. Một số hình ảnh kết quả nâng cao chất lượng ảnh hồng ngoại được đưa ra trong hình 4.

Bảng 1. Kết quả đánh giá chất lượng ảnh IR sau xử lý.

Ảnh	Entropy (ảnh IR ban đầu)	Entropy (ảnh IR đã xử lý)
1	6,22	11,03
2	6,28	11,24
3	3,90	13,61
4	6,69	11,60
5	6,87	11,37

Các kết quả tính toán giá trị đo lường Entropy cho một số cặp ảnh để đánh giá chất lượng ảnh IR sau xử lý trong bảng 1 cho thấy chất lượng IR sau xử lý có kết quả tốt hơn so với ảnh IR ban đầu. Ảnh IR sau xử lý chứa nhiều lượng thông tin hơn ảnh IR ban đầu.

4. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Bài báo đã trình bày phương pháp nâng cao chất lượng ảnh cho đài chỉ huy trung tâm của hệ thống pháo phòng không bằng cách trộn hai ảnh từ hai nguồn ảnh camera thường và camera hồng ngoại. Đây là các nguồn ảnh thường có chất lượng ảnh cần thiết phải nâng cao để phục vụ bài toán phát hiện và bám chuyển động của đối tượng trong ảnh. Với yêu cầu tốc độ xử lý phải nhanh, việc lựa chọn phương pháp trộn ảnh đơn giản giúp việc đảm bảo tốc độ xử lý. Các kết quả ban đầu cho thấy việc nâng cao chất lượng ảnh bằng phương pháp này cho đối tượng đài chỉ huy trung tâm của hệ thống pháo phòng không là khả thi. Các thuật toán tổng hợp ảnh phức tạp hơn như PCA, Wavelet, Deep Learning sẽ được nghiên cứu sâu hơn tiếp sau kết quả nghiên cứu này để có thể ứng dụng trong thực tế trong đài chỉ huy trung tâm hệ thống pháo phòng không.

TÀI LIỆU THAM KHẢO

- [1]. Ying, Zhenqiang and Li, Ge and Ren, Yurui and Wang, Ronggang, and Wang, Wenmin, “A New Image Contrast Enhancement Algorithm Using Exposure Fusion Framework”, International Conference on Computer Analysis of Images and Patterns, Springer, pp. 36-46, (2017).
- [2]. Fu, Qingtao and Jung, Cheolkon and Xu, Kaiqiang, “Retinex-based perceptual contrast enhancement in images using luminance adaptation”, IEEE Access, volume 6, pp. 61277—61286, (2018).
- [3]. Wang, Wencheng and Chen, Zhenxue and Yuan, Xiaohui and Wu, Xiaojin, “Adaptive image enhancement method for correcting low-illumination images”, Information Sciences, Elsevier, Volume 496, pp. 25—41, (2019).
- [4]. Agrawal, Sanjay and Panda, Rutuparna and Mishro, PK and Abraham, Ajith, “A novel joint histogram equalization based image contrast enhancement”, Journal of King Saud University-Computer and Information Sciences, Elsevier, (2019).
- [5]. Hessel, Charles and Morel, Jean-Michel, “An extended exposure fusion and its application to single image contrast enhancement”, The IEEE Winter Conference on Applications of Computer Vision, pp. 137-146, (2020).
- [6]. R. Brooks, S. Iyengar, “Multi-Sensor Fusion: Fundamentals and Applications”, Prentice Hall, (1998).
- [7]. R. Luo and M. Kay, “Multisensor Integration and Fusion for Intelligent Machines and Systems”, Ablex, Norwood, NJ, (1995).
- [8]. H. Li and Y. Zhou, “Automatic visual/IR image registration,” Optical Engineering, 35(2), pp. 391-400, (1996).
- [9]. R. Sharma and M.Pavel, “Multisensor image registration,” SID Digest, Society for Information Display, XXVIII, pp. 951–954, (1997).
- [10]. H. Li, B. Manjunath, and S. Mitra, “A contour-based approach to multisensor image registration,” IEEE Transactions on Image Processing, 4, No. 3, (1995).
- [11]. D. G. Lowe, “Object Recognition from Local Scale-Invariant Features”, Proceedings of the Seventh 422 IEEE International Conference on Computer Vision, Vol. 2, pp. 1150–1157, (1999).
- [12]. Herbert Bay, Andreas Ess, Tinne Tuytelaars, Luc Van Gool, “Speeded-Up Robust Features”, Computer Vision and Image Understanding (CVIU), Vol. 110, No. 3, pp. 346–359, (2008).

- [13]. Rublee E, Rabaud V, Konolige K, Bradski G, “ORB: An efficient alternative to SIFT or SURF”, 2011 International Conference on Computer Vision, pp. 2564–2571, (2011).
- [14]. Marius Muja, David G. Lowe, “FLANN: Fast Library for Approximate Nearest Neighbors”, ACM International Conference on Multimedia (ACM MM), pp. 951-954, (2009).
- [15]. Martin A. Fischler và Robert C. Bolles, “Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography”, Communications of the ACM (ACM), pp.120-133, (1981).
- [16]. Jiayi Ma, Yong Ma, Chang Li, “Infrared and visible image fusion methods and applications: A survey”, Information Fusion Volume 45, pp. 153-178, (2019).

ABSTRACT

A method to improve image quality by mixing images from two different image sources

This article describes a method to improve the quality of infrared images by mixing images from two sources of conventional camera images and infrared images to observe the same scene. Two images from two different cameras do not have a consistent data structure: different resolutions, different fields of view (FOV), and different lens adjustments. To improve the quality of an image source, the images must first be converted to the same conditions. The two images are then mixed to enhance the quality of the infrared image. Experimental results show the effectiveness of the method when infrared image quality is improved.

Keywords: Machine vision; Image processing; Improve image quality; Image matching.