

BỘ GIÁO DỤC VÀ ĐÀO TẠO

BỘ QUỐC PHÒNG

HỌC VIỆN KỸ THUẬT QUÂN SỰ

NGUYỄN VĂN PHƯƠNG

PHÁT TRIỂN MỘT SỐ KỸ THUẬT PHÁT HIỆN DỊ
THƯỜNG TRÊN ẢNH VIỄN THÁM QUANG HỌC
PHỤC VỤ CÔNG TÁC TÌM KIẾM CỨU NẠN

LUẬN ÁN TIẾN SĨ KỸ THUẬT

HÀ NỘI - 2023

BỘ GIÁO DỤC VÀ ĐÀO TẠO

BỘ QUỐC PHÒNG

HỌC VIỆN KỸ THUẬT QUÂN SỰ

NGUYỄN VĂN PHƯƠNG

PHÁT TRIỂN MỘT SỐ KỸ THUẬT PHÁT HIỆN DỊ
THƯỜNG TRÊN ẢNH VIỄN THÁM QUANG HỌC
PHỤC VỤ CÔNG TÁC TÌM KIẾM CỨU NẠN

Chuyên ngành: CƠ SỞ TOÁN HỌC CHO TIN HỌC

Mã số: 9 46 01 10

LUẬN ÁN TIẾN SĨ KỸ THUẬT

NGƯỜI HƯỚNG DẪN KHOA HỌC:

1. TS ĐÀO KHÁNH HOÀI
2. TS TỔNG MINH ĐỨC

HÀ NỘI - 2023

LỜI CAM ĐOAN

Tôi xin cam đoan các kết quả trình bày trong luận án là công trình nghiên cứu của tôi dưới sự hướng dẫn của cán bộ hướng dẫn. Các số liệu, kết quả trình bày trong luận án là hoàn toàn trung thực và chưa được ai công bố trong bất kỳ công trình nào trước đây. Các kết quả sử dụng tham khảo đều đã được trích dẫn đầy đủ và theo đúng quy định.

Hà Nội, ngày 14 tháng 3 năm 2023

NGHIÊN CỨU SINH

Nguyễn Văn Phương

LỜI CẢM ƠN

Trong quá trình học tập, nghiên cứu và thực hiện luận án, nghiên cứu sinh đã nhận được sự hướng dẫn, giúp đỡ tận tình, các ý kiến đóng góp quý báu của các Thầy, Cô, các nhà khoa học; Sự động viên, chia sẻ của bạn bè, đồng nghiệp và gia đình.

Nghiên cứu sinh xin bày tỏ lòng biết ơn sâu sắc đến Thầy giáo hướng dẫn **TS Đào Khánh Hoài** và **TS Tống Minh Đức**. Các thầy đã nhiệt tình, tận tâm định hướng, hướng dẫn, giúp đỡ nghiên cứu sinh trong suốt quá trình nghiên cứu và hoàn thành luận án này.

Nghiên cứu sinh trân trọng cảm ơn quý Thầy, Cô giáo Khoa Công nghệ Thông tin, Học viện Kỹ thuật Quân sự đã tận tình giảng dạy, giúp đỡ trong thời gian nghiên cứu sinh học tập, nghiên cứu tại đây.

Nghiên cứu sinh gửi lời cảm ơn đến Học viện Lục quân; Phòng Sau đại học, Bộ môn Hệ thống Thông tin, Học viện Kỹ thuật Quân sự đã giúp đỡ, tạo điều kiện cho nghiên cứu sinh được đi học tập, nghiên cứu, và hoàn thành luận án này.

Cuối cùng, nghiên cứu sinh gửi lời cảm ơn chân thành tới gia đình, bạn bè và đồng nghiệp, những người đã luôn ủng hộ, tạo niềm tin, động viên, chia sẻ những khó khăn với nghiên cứu sinh trong suốt thời gian vừa qua.

Hà Nội, tháng 3 năm 2023

Nguyễn Văn Phương

MỤC LỤC

MỤC LỤC	iii
DANH MỤC CÁC KÝ HIỆU	vi
DANH MỤC CÁC KÝ HIỆU THUẬT TOÁN	vii
DANH MỤC TỪ VIẾT TẮT, Ý NGHĨA	ix
DANH MỤC CÁC HÌNH VẼ	xii
DANH MỤC CÁC BẢNG BIỂU	xiv
MỞ ĐẦU	1

CHƯƠNG 1.

TỔNG QUAN VỀ PHÁT HIỆN DỊ THƯỜNG TRÊN ẢNH VIỄN THÁM QUANG HỌC

	9
1.1 Ảnh viễn thám quang học	9
1.1.1 Phổ của ảnh viễn thám quang học	10
1.1.2 Các loại ảnh viễn thám quang học	11
1.1.3 Độ phân giải không gian	14
1.1.4 Cấu trúc dữ liệu của ảnh viễn thám quang học	14
1.1.5 Thư viện quang phổ và thư viện ảnh viễn thám quang học	20
1.2 Bài toán phát hiện dị thường trên ảnh viễn thám quang học phục vụ công tác tìm kiếm cứu nạn	22
1.2.1 Khái niệm dị thường	22
1.2.2 Công tác tìm kiếm cứu nạn	24
1.2.3 Phát biểu bài toán	27
1.3 Một số tiếp cận về phát hiện dị thường trên ảnh đa phổ, siêu phổ	27
1.3.1 Phương pháp dựa trên mô hình xác suất thống kê	28
1.3.2 Phương pháp dựa trên phân cụm và phân đoạn	32
1.3.3 Phương pháp dựa trên biến đổi phi tuyến	34
1.3.4 Phương pháp dựa trên biến đổi tuyến tính	36
1.3.5 Phương pháp dựa trên học máy	39
1.4 Tiêu chí đánh giá độ chính xác của các thuật toán phát hiện dị thường	45

1.5	Phương pháp tính độ phức tạp thuật toán	47
1.5.1	Khái niệm	47
1.5.2	Đánh giá độ phức tạp thuật toán	48
1.5.3	Thời gian thực hiện các câu lệnh trong các ngôn ngữ lập trình	49
1.6	Thách thức và hướng phát triển	49
1.6.1	Thách thức về tăng độ chính xác phát hiện dị thường	49
1.6.2	Thách thức về phát hiện dị thường với các kích cỡ khác nhau	51
1.6.3	Thách thức về tốc độ tính toán	52
1.7	Kết luận	53

CHƯƠNG 2.

KỸ THUẬT NÂNG CAO ĐỘ CHÍNH XÁC PHÁT HIỆN DỊ THƯỜNG TRÊN ẢNH VIỄN THÁM QUANG HỌC

		54
2.1	Giới thiệu chung	54
2.2	Thuật toán phát hiện các điểm ảnh dị thường dựa trên KDE	55
2.3	Giải pháp tăng độ chính xác của thuật toán phát hiện các điểm ảnh dị thường dựa trên KDE	58
2.4	Tính đúng đắn của thuật toán IKDE	70
2.5	Đánh giá độ phức tạp tính toán của thuật toán KDE và IKDE	71
2.5.1	Độ phức tạp tính toán của thuật toán KDE	71
2.5.2	Độ phức tạp tính toán của thuật toán IKDE	72
2.6	Thử nghiệm và đánh giá kết quả đề xuất	73
2.6.1	Dữ liệu thử nghiệm	73
2.6.2	Phương pháp thử nghiệm	79
2.6.3	Kết quả thử nghiệm	79
2.7	Kết luận	89

CHƯƠNG 3.

KỸ THUẬT TĂNG TỐC ĐỘ PHÁT HIỆN DỊ THƯỜNG TRÊN ẢNH VIỄN THÁM QUANG HỌC

91

3.1	Giới thiệu chung	91
3.2	Tăng tốc độ tính toán của thuật toán IKDE	92
3.3	Kỹ thuật nhóm các điểm ảnh và sắp xếp	96
3.4	Tính toán PDF và phân loại các điểm ảnh	98
3.4.1	Ứng dụng cây kd-tree hỗ trợ tính toán PDF	99
3.4.2	Tính toán PDF đa luồng trên CPU	105
3.4.3	Tính toán PDF song song trên GPU	111
3.5	Đánh giá độ phức tạp của các thuật toán	115
3.5.1	Thuật toán nhóm các điểm ảnh (CreateGroupPixels)	115
3.5.2	Thuật toán tìm kiếm những nhóm điểm ảnh (Search-GroupPixels)	115
3.5.3	Cây kd-tree	115
3.5.4	Thuật toán phát hiện các điểm ảnh dị thường trên ảnh màu RGB sử dụng kỹ thuật nhóm các điểm ảnh và cây kd-tree	116
3.5.5	Thuật toán phát hiện các điểm ảnh dị thường trên ảnh màu MSI, HSI sử dụng kỹ thuật nhóm các điểm ảnh và cây kd-tree	117
3.5.6	Thuật toán phát hiện các điểm ảnh dị thường trên ảnh màu RGB sử dụng kỹ thuật nhóm các điểm ảnh và tính toán đa luồng trên CPU hoặc song song trên GPU	117
3.5.7	Thuật toán phát hiện các điểm ảnh dị thường trên ảnh màu MSI, HSI sử dụng kỹ thuật nhóm các điểm ảnh và tính toán đa luồng trên CPU hoặc song song trên GPU	118
3.6	Thử nghiệm và đánh giá kết quả đề xuất	120
3.6.1	Dữ liệu thử nghiệm	120
3.6.2	Phương pháp thử nghiệm	122
3.6.3	Kết quả thử nghiệm	123
3.7	Kết luận	130
	KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN	132
	CÁC CÔNG TRÌNH KHOA HỌC ĐÃ CÔNG BỐ	135
	TÀI LIỆU THAM KHẢO	136

DANH MỤC CÁC KÝ HIỆU

A	Mảng hai chiều lưu chỉ số các điểm ảnh
C_0	Tập các điểm ảnh
C_1	Tập các điểm ảnh "bình thường"
C_2	Tập các điểm ảnh "dị thường"
G	Tập các nhóm điểm ảnh
$K(u)$	Hàm nhân [122]
L	Số kênh phổ
M	Số lượng các nhóm điểm ảnh
N	Số điểm ảnh
P	Tập các điểm ảnh thỏa mãn điều kiện $K(u) \neq 0$
S	Tập các điểm ảnh đặc trưng
X	Ma trận điểm ảnh
W	Cửa sổ giới hạn
$\hat{f}(\cdot)$	Hàm mật độ xác suất
q	Số phần đều nhau của ma trận A
h	bảng thông
r	Bán kính siêu cầu chứa các điểm ảnh thỏa mãn điều kiện $K(u) \neq 0$ ($r = h \times \epsilon$)
x_i	Điểm ảnh thứ i
ϵ	Ngưỡng giới hạn để $K(u) \neq 0$
σ	Hệ số làm mịn
η	Ngưỡng phát hiện các điểm ảnh dị thường
$\Phi(\cdot)$	Hàm ánh xạ phi tuyến
*	Toán hạng tích chập

DANH MỤC CÁC KÝ HIỆU THUẬT TOÁN

TT	Thuật toán	Ký hiệu	Mô tả
1	Thuật toán 2.1	KDE	Thuật toán phát hiện các điểm ảnh dị thường dựa trên KDE
2	Thuật toán 2.2	GaussianBlur	Thuật toán làm mịn ảnh
3	Thuật toán 2.3	IKDE	Cải tiến thuật toán phát hiện các điểm ảnh dị thường dựa trên KDE
4	Thuật toán 3.1	CreateGroupPixels	Thuật toán nhóm các điểm ảnh
5	Thuật toán 3.2	SearchGroupPixels	Thuật toán tìm kiếm những nhóm điểm ảnh nằm trong bán kính r , tâm là PUT
6	Thuật toán 3.3	CreateKdTree	Thuật toán tạo cây kd-tree
7	Thuật toán 3.4	SearchKdTree	Thuật toán tìm kiếm trên cây kd-tree
8	Thuật toán 3.5	GP-KDT1	Thuật toán phát hiện các điểm ảnh dị thường trên ảnh màu RGB sử dụng kỹ thuật nhóm các điểm ảnh cùng giá trị và cây kd-tree

- | | | | |
|----|-----------------|---------|---|
| 9 | Thuật toán 3.6 | GP-KDT2 | Thuật toán phát hiện các điểm ảnh dị thường trên ảnh đa phổ, siêu phổ sử dụng kỹ thuật nhóm các điểm ảnh cùng giá trị và cây kd-tree |
| 10 | Thuật toán 3.7 | GP-MC1 | Thuật toán phát hiện các điểm ảnh dị thường trên ảnh màu RGB sử dụng kỹ thuật nhóm các điểm ảnh cùng giá trị và tính toán đa luồng trên CPU |
| 11 | Thuật toán 3.8 | GP-MC2 | Thuật toán phát hiện các điểm ảnh dị thường trên ảnh đa phổ, siêu phổ sử dụng kỹ thuật nhóm các điểm ảnh cùng giá trị và tính toán đa luồng trên CPU |
| 12 | Thuật toán 3.9 | GP-PC1 | Thuật toán phát hiện các điểm ảnh dị thường trên ảnh màu RGB sử dụng kỹ thuật nhóm các điểm ảnh cùng giá trị và tính toán song song trên GPU |
| 13 | Thuật toán 3.10 | GP-PC2 | Thuật toán phát hiện các điểm ảnh dị thường trên ảnh đa phổ, siêu phổ sử dụng kỹ thuật nhóm các điểm ảnh cùng giá trị và tính toán song song trên GPU |

DANH MỤC TỪ VIẾT TẮT, Ý NGHĨA

Viết tắt	Nghĩa tiếng Anh	Nghĩa tiếng Việt
AD	Anomaly Detection	Bộ dò dị thường
AIC	Akaike Information Criterion	Tiêu chuẩn Thông tin Akaike
BIC	Bayesian Information Criterion	Tiêu chuẩn Thông tin Bayesian
CAP	Civil Air Patrol	Tuần tra hàng không dân dụng
CBAD	Cluster Based Anomaly Detector	Bộ dò dị thường trên cụm
CFAR	Constant False alarm Rate	Tỉ lệ cảnh báo sai bất biến
CKM	Constrained Kurtosis Maximization	Tối đa hóa Kurtosis cưỡng bức
CPU	Central Processing Unit	Bộ xử lý trung tâm
DSPAD	Deterministic Signal Subspace Processing AD	AD xử lý không gian tín hiệu xác định
EM	Expectation-Maximisation	
FA	False Alarm	Báo động sai
FAR	False Alarm Ratio	Tỷ lệ cảnh báo sai
FGT	Fast Gauss Transform	Biến đổi Gauss nhanh
FPR	False Positive Rate	Tỷ lệ dương tính sai
GIC	Generalized Information Criterion	Tiêu chuẩn Thông tin tổng quát
GLRT	Generalised Likelihood Ratio Test	Thử nghiệm tỉ lệ hợp lý tổng quát
GPU	Graphics Processing Unit	Bộ xử lý đồ họa
GMM	Gaussian-Mixture Model	Mô hình hỗn hợp Gaussian
GMRF	Gaussian Markov Random Field	Trường ngẫu nhiên Markov Gaussian

GNM	Global Normal Model	Mô hình phân phối chuẩn toàn cục
GSNR	Generalized Signal to Noise Ratio	Tỉ lệ nhiễu-tín hiệu tổng quát
GP-KDT	Group Pixels and use KD-Tree	Nhóm các điểm ảnh và sử dụng cây kd-tree
GP-MC	Group Pixels and Multithreading Computation	Nhóm các điểm ảnh và tính toán đa luồng
GP-PC	Group Pixels and Parallel Computation	Nhóm các điểm ảnh và tính toán song song
IIR	Infinite Impulse Response	Đáp ứng xung vô hạn
IFGT	Improved Fast Gauss Transform	Cải tiến thuật toán biến đổi Gauss nhanh
IKDE	Improved KDE	Cải tiến thuật toán KDE
KR	Likelihood Ratio	Tỷ lệ hợp lý
KDE	Kernel Density Estimation	Ước lượng mật độ nhân
LMM	Linear Mixing Model	Mô hình trộn tuyến tính
LNM	Local Normal Model	Mô hình phân phối chuẩn cục bộ
LRT	Likelihood Ratio Test	Thử nghiệm tỉ lệ hợp lý
LRXD	Local RXD	Thuật toán RX cục bộ
MAP	Maximum A Posteriori	Quy tắc Posteriori tối đa
MD	Mahalanobis Distance	Khoảng cách Mahalanobis
MDL	Minimum Description Length	Độ dài mô tả tối thiểu
ML	Maximum Likelihood	Hợp lý tối đa
MLE	Maximum Likelihood Estimation	Ước lượng hợp lý tối đa
NP	Neyman-Pearson	Neyman-Pearson
OSP	Orthogonal Subspace Projection	Phép chiếu không gian con trực giao
PCA	Principal Component Analysis	Phân tích thành phần chính
PDF	Probability density function	Hàm mật độ xác suất
PR	Precision-Recall	Độ chính xác - Độ bao phủ

PUT	Pixel Under Test	Điểm ảnh đang xét
ROC	Receiver Operating Characteristics	Đường cong đặc trưng
ROI	Regions of Interest	Khu vực quan tâm
RX	Reed-Xiaoli	Thuật toán RX
SET	Stochastic Expectation Maximization	Tối đa hóa kì vọng ngẫu nhiên
SIFT	Scale-invariant feature transform	Thuật toán trích rút các điểm đặc trưng
SURF	Speeded up robust features	Thuật toán trích rút các điểm đặc trưng
SSP	Signal Subspace Processing	Xử lý không gian con tín hiệu
SSTD	Spectral Signature-based Target Detector	Bộ dò mục tiêu dựa trên đặc trưng quang phổ
SVD	Singular Value Decomposition	Phân rã giá trị kì dị
SVDD	Support Vector Data Description	Phương pháp mô tả SVD
TPR	true positive rate	Tỷ lệ dương tính đúng
UAV	Unmanned aerial vehicle	Phương tiện bay không người lái
VQ	Vector Quantization	Vector lượng tử hóa

DANH MỤC CÁC HÌNH VẼ

Hình 1.1	Nguyên lý của ảnh quang phổ [94]	10
Hình 1.2	Biểu đồ mối quan hệ giá trị phản xạ và bước sóng [85] .	12
Hình 1.3	Ví dụ một ảnh toàn sắc, ảnh có mã là 6010_1_2_P do DSTL cung cấp [24]	12
Hình 1.4	Các kênh đa phổ và độ rộng của một bộ cảm biến đa phổ TM/ETM+ [2]	13
Hình 1.5	Cấu trúc của một bộ cảm siêu phổ. Các kênh ảnh được thu nhận đồng thời trên nhiều kênh phổ hẹp [2]	14
Hình 1.6	Khuôn viên của Đại học Maryland tại College Park với bốn phân giải không gian khác nhau [66]. Ảnh (a) có độ phân giải không gian là 1m/pixel; ảnh (b) có độ phân giải không gian là 10m/pixel; ảnh (c) có độ phân giải không gian là 30m/pixel và ảnh (d) có độ phân giải không gian là 250m/pixel [2]	16
Hình 1.7	Ví dụ minh họa hệ tọa độ ảnh số	18
Hình 1.8	Kênh Pan trên ảnh SPOT được mã hóa 8 bit [2]	18
Hình 1.9	Ảnh Envisat ASAR được mã hóa 16 bit [2]	19
Hình 1.10	Mô hình hoạt động của hệ thống tìm kiếm và cứu hộ hàng hải dựa trên ứng dụng công nghệ viễn thám [92]	26
Hình 1.11	Đường cong ROC [6]	46
Hình 2.1	Các bước nhằm tăng độ chính xác phát hiện dị thường.	60
Hình 2.2	Ảnh được làm mịn với các hệ số làm mịn (σ) khác nhau	60
Hình 2.3	Ví dụ minh họa việc thực hiện hàm DoG để tìm những điểm đặc trưng bất biến cao. (a) là ảnh đã được làm mờ với hệ số làm mờ (σ) khác nhau, (b) thể hiện giá trị của hàm DoG trong một quãng tám (octave).	63
Hình 2.4	Ví dụ minh họa việc tìm các cực trị cục bộ [71]	64
Hình 2.5	Ví dụ minh họa việc tìm các cực trị cục bộ trên ảnh. (a) là ảnh đã được làm mờ với hệ số làm mờ (σ) khác nhau, (b) thể hiện giá trị của hàm DoG trong một quãng tám (octave), (c) thể hiện giá trị của việc tìm các cực trị cục bộ.	65

Hình 2.6	Hình (a) là biểu đồ 3D thể hiện giá trị hàm DoG của các điểm ảnh trong không gian tỷ lệ $\sigma=1.6$, hình (b) là ảnh gốc. Các vị trí được đánh dấu (A, B và C) ở hình (a) tương ứng với các vị trí được đánh dấu ở hình (b)	66
Hình 2.7	Kết quả sau khi thực hiện giai đoạn một và giai đoạn hai của thuật toán SIFT, ảnh (a) là ảnh gốc, các chấm màu đỏ trên ảnh (b) là vị trí các điểm ảnh đặc trưng được tìm thấy trong giai đoạn hai của thuật toán SIFT	68
Hình 2.8	Biểu diễn các điểm ảnh đặc trưng (các chấm màu xanh) và cửa sổ W	68
Hình 2.9	Ba ảnh đã được cấy các mẫu dị thường	73
Hình 2.10	Các mẫu dị thường được bố trí trên địa hình	75
Hình 2.11	Ảnh (a), (b) và (c) đã được thêm 2% nhiễu Gaussian	75
Hình 2.12	Ảnh (d), (e) và (f) đã được thêm 2% nhiễu Gaussian	76
Hình 2.13	Ảnh (a), (b) và (c) đã được thêm 5% nhiễu Gaussian	76
Hình 2.14	Ảnh (d), (e) và (f) đã được thêm 5% nhiễu Gaussian	76
Hình 2.15	Biểu đồ so sánh độ chính xác phát hiện dị thường của thuật toán KDE (thuật toán gốc) và thuật toán IKDE (thuật toán cải tiến) trên 36 ảnh gốc	85
Hình 2.16	Biểu đồ so sánh độ chính xác phát hiện dị thường của thuật toán KDE (thuật toán gốc) và thuật toán IKDE (thuật toán cải tiến) trên 36 ảnh đã được thêm 2% nhiễu vào ảnh gốc	86
Hình 2.17	Biểu đồ so sánh độ chính xác phát hiện dị thường của thuật toán KDE (thuật toán gốc) và thuật toán IKDE (thuật toán cải tiến) trên 36 ảnh đã được thêm 5% nhiễu vào ảnh gốc	87
Hình 3.1	Ý tưởng giảm dữ liệu cần tính toán việc ước lượng PDF.	94
Hình 3.2	Các bước để tăng tốc độ tính toán của thuật toán phát hiện các điểm ảnh dị thường dựa trên KDE	95
Hình 3.3	Tìm kiếm những nhóm điểm ảnh nằm trong hoặc giao cắt với hình siêu cầu bán kính r , tâm là PUT	99
Hình 3.4	a) Minh họa phân chia miền không gian, b) Minh họa cây kd-tree đã được xây dựng từ dữ liệu đã cho.	101
Hình 3.5	Minh họa những điểm được chọn để tính $K(u)$	103

Hình 3.6	Các ảnh đa phổ: (a) Ảnh 3 kênh phổ mã số 6010_1_2; (b) Ảnh 8 kênh phổ mã số 6010_1_2_M (hình hiển thị là sự tổ hợp 3 kênh, kênh số 1, kênh số 2 và kênh số 3); (c) thể hiện những điểm ảnh dị thường (màu trắng) và những điểm ảnh bình thường (màu đen).	121
Hình 3.7	Ảnh siêu phổ: (a) Kênh 220 của ảnh siêu phổ Salinas 224 kênh phổ; (b) thể hiện những điểm ảnh dị thường (màu trắng) và những điểm ảnh bình thường (màu đen).	122
Hình 3.8	Biểu đồ so sánh thời gian chạy của thuật toán tính toán trực tiếp các điểm ảnh theo Thuật toán KDE và thuật toán tính toán đa luồng trên CPU do Michailidis và các cộng sự đề xuất tại [86].	125
Hình 3.9	Biểu đồ so sánh thời gian chạy của thuật toán tính toán trực tiếp các điểm ảnh theo Thuật toán KDE và thuật toán tính toán song song trên GPU do Michailidis và các cộng sự đề xuất tại [87].	125
Hình 3.10	Biểu đồ so sánh thời gian chạy của Thuật toán KDE và thuật toán GP-KDT1, GP-KDT2	127
Hình 3.11	Biểu đồ so sánh thời gian chạy của thuật toán Intel TBB và thuật toán GP-MC1, GP-MC2	128
Hình 3.12	Biểu đồ so sánh thời gian chạy của thuật toán GPU CUDA và thuật toán GP-PC1, GP-PC2	128

DANH MỤC CÁC BẢNG BIỂU

Bảng 1.1	Bảng so sánh các loại ảnh viễn thám	15
Bảng 1.2	Độ phân giải không gian của một số loại bộ cảm vệ tinh [66]	17
Bảng 1.3	Đặc điểm của một số phương pháp phát hiện dị thường	41
Bảng 2.1	Một số nhân điển hình [122]	57
Bảng 2.2	Bảng so sánh sự khác nhau giữa thuật toán KDE và IKDE	71
Bảng 2.3	Danh sách những ảnh được chọn trong thư viện do DSTL cung cấp [24]	78
Bảng 2.4	Độ đo AUC của thuật toán KDE và thuật toán IKDE trên 36 ảnh gốc (ảnh chưa thêm nhiễu)	80
Bảng 2.5	Độ đo AUC của thuật toán KDE và thuật toán IKDE trên 36 ảnh đã được thêm 2% nhiễu vào ảnh gốc	81
Bảng 2.6	Độ đo AUC của thuật toán KDE và thuật toán IKDE trên 36 ảnh đã được thêm 5% nhiễu vào ảnh gốc	83
Bảng 2.7	Thời gian thực thi của thuật toán KDE và thuật toán IKDE trên 36 ảnh gốc, thời gian tính bằng giây (s)	87
Bảng 3.1	Bảng so sánh độ phức tạp tính toán của các thuật toán	119
Bảng 3.2	Thời gian thực thi của các thuật toán trên các ảnh (tính bằng giây).	124
Bảng 3.4	Thời gian và độ chính xác phát hiện các điểm ảnh dị thường trên ảnh 6010_1_2 khi áp dụng thuật toán IFGT tính toán PDF cho PUT	126
Bảng 3.5	Thời gian và độ chính xác phát hiện các điểm ảnh dị thường trên ảnh 6010_1_2_M khi áp dụng thuật toán IFGT tính toán PDF cho PUT	126
Bảng 3.6	So sánh thời gian thực thi của các thuật toán trên ảnh màu RGB mã số 6010_1_2.	129
Bảng 3.7	So sánh thời gian thực thi của các thuật toán trên ảnh đa phổ 8 kênh phổ mã số 6010_1_2_M.	129

Bảng 3.8 So sánh thời gian thực thi của các thuật toán trên ảnh
siêu phổ phổ 224 kênh phổ mã Salinas. 130

MỞ ĐẦU

1. Tính cấp thiết của Luận án

Theo số liệu của Ủy ban Quốc gia Ứng phó sự cố thiên tai và tìm kiếm cứu nạn [3], từ năm 2004 đến năm 2021, tại Việt Nam đã có 54.964 vụ thiên tai, sự cố, hậu quả làm chết 14.937 người, mất tích 2.616 người, bị thương 15.502 người, làm hư hỏng hàng chục ngàn trang thiết bị. Hiện nay, do biến đổi khí hậu toàn cầu và những hoạt động thiếu ý thức của con người, nhất là trong bảo vệ môi trường sống nên dự báo trong những năm tới, tình hình thiên tai, sự cố diễn biến phức tạp, bất thường, theo chiều hướng cực đoan, cả về tính chất, quy mô, cường độ và mức độ tàn phá. Theo đó, công tác tìm kiếm cứu nạn cũng phải được đầu tư đúng mức để theo kịp tình hình thực tế.

Một công cụ đang ngày càng được chứng minh có hiệu quả cao không chỉ trong công tác tìm kiếm cứu nạn mà còn trong nhiều ứng dụng khác như trong nông nghiệp và lâm nghiệp [5], tài nguyên nước [34], rà phá bom mìn [8, 55], công nghiệp thực phẩm [35],... đó là sử dụng ảnh viễn thám quang học. Bởi công nghệ Viễn thám có nhiều tính năng ưu việt so với phương pháp truyền thống như độ phủ trùm rộng, thời gian cập nhật ngắn, dải phổ rộng, có thể thu nhận ảnh ở những khu vực rất khó tiếp cận như vùng núi cao, rừng rậm, hải đảo, khu vực ngoài lãnh thổ,... Đặc biệt, trong những năm gần đây, các thiết bị bay không người lái (UAV) đã và đang phát triển rất mạnh mẽ, được ứng dụng rộng rãi cả trong dân sự và quân sự. Nó thực sự là một nguồn lực rất lớn cho sứ mệnh tìm kiếm, cứu nạn bởi thiết bị này có thể mang các bộ cảm biến để thu thập hình ảnh có độ phân giải cao với phạm vi hoạt động rộng lớn, địa hình đa dạng mà không cần quá nhiều nhân lực và chi phí cho quá trình tìm kiếm [89].

Hoạt động tìm kiếm và cứu nạn bao gồm việc tìm kiếm và giải cứu người, phương tiện bị mắc kẹt trong các tình huống khó khăn hoặc được báo nạn. Cách tiếp cận sử dụng ảnh viễn thám quang học là sử dụng ảnh đa phổ [102], siêu phổ [15, 36] có độ phân giải cao được các cảm biến gắn trên máy bay, khinh khí cầu hoặc vệ tinh thu nhận hoàn toàn khả thi và mang lại hiệu quả cao. Tuy nhiên, các ảnh hưởng bất lợi gây ra bởi đặc trưng của địa hình, điều

kiện thời tiết khắc nghiệt làm cho vị trí báo nạn có dung sai lớn. Các đối tượng cần tìm kiếm đôi khi bị che khuất bởi địa hình, mật độ dày đặc của lá cây hoặc bị chìm một phần dưới nước. Trực tiếp tìm ra người đang gặp nạn sẽ là lý tưởng, nhưng trong một số trường hợp, các đồ vật đi kèm như quần áo, chăn mền, va ly, túi xách, lều bạt nơi cắm trại, các mảnh vỡ phương tiện,... (gọi tắt là những đối tượng nhân tạo) có thể cung cấp một số thông tin hữu ích [12, 80, 89, 95]. Ngoài ra, các thiết bị cảm biến thu dữ liệu phải quét trên một diện rộng và dung lượng dữ liệu lớn là một rào cản đối với việc tìm kiếm thủ công bằng mắt thường. Các kỹ thuật tự động phát hiện mục tiêu là giải pháp phù hợp giúp người quan sát nâng cao hiệu quả và tốc độ tìm kiếm [89].

Phát hiện mục tiêu trong ảnh đa phổ, siêu phổ được tiếp cận theo ba hướng [74]. Thứ nhất, xác định xem các đối tượng là mục tiêu mong muốn hay sự đa dạng tự nhiên dựa trên tín hiệu quang phổ đã biết của mục tiêu mong muốn thu được từ thư viện quang phổ hoặc từ một tập hợp dữ liệu về mục tiêu đã được huấn luyện. Thứ hai, phát hiện các mục tiêu dựa trên phát hiện sự thay đổi, được sử dụng để xác định các thay đổi trong một cảnh theo thời gian, có hoặc không có bất kỳ kiến thức nào về tín hiệu mục tiêu. Thứ ba, phương pháp phát hiện các điểm ảnh dị thường (ngoại lai), xác định các vector phổ từ các điểm ảnh bị lệch hoặc có sự khác biệt đáng kể so với các điểm ảnh xung quanh.

Đối với cách tiếp cận thứ nhất, do có sẵn các thư viện đặc trưng quang phổ cho nhiều loại vật liệu, các SSTD (Spectral Signature-based Target Detector) được kiểm tra rộng rãi [77, 78]. Cách tiếp cận này giả định đặc trưng quang phổ mục tiêu đã biết và cố gắng phát hiện mọi điểm ảnh có phổ cho thấy mức độ tương quan cao với phổ tham chiếu, chúng được gọi là thuật toán so khớp phổ. Tuy nhiên, các kỹ thuật này dựa trên phổ tham chiếu, điển hình là độ phản xạ phổ được đo bằng máy quang phổ cầm tay, trong khi các điểm ảnh thu được tại cảm biến thường được đo bằng bức xạ quang phổ, do đó phải trải qua bước bù trừ khí quyển. Khi đó, kết quả phát hiện các phổ mục tiêu sẽ bị ảnh hưởng rất lớn [74]. Ngoài ra, hiệu suất của các SSDT có thể bị hạn chế bởi sự không chắc chắn như bù trừ khí quyển, điều chỉnh bức xạ quang phổ, hình dạng tiếp nhận và nhiễu từ các vật liệu liền kề [44, 113]. Trong công tác tìm kiếm cứu nạn, các đối tượng cần tìm kiếm rất phong phú và

trong nhiều trường hợp phổ mục tiêu không được xác định trước nên phương pháp này chỉ áp dụng cho một số trường hợp cụ thể.

Đối với cách tiếp cận thứ hai, dữ liệu ảnh đa phổ, siêu phổ thu được từ trên cao trong khoảng thời gian dài, được sử dụng để phát hiện các thay đổi trên bề mặt của khu vực cần quan tâm. Các nhà nghiên cứu đã phát triển một số kỹ thuật [30, 81, 107] để xác định các thay đổi trên bề mặt địa hình theo thời gian, từ đó có thể phát hiện được mục tiêu cần tìm kiếm. Hình ảnh phải được đăng ký trước khi thực hiện việc phát hiện sự thay đổi (đăng ký hình ảnh là quá trình sắp xếp hình ảnh vào cùng một khung tọa độ, các phương pháp đăng ký hình ảnh đã được Zitova and Flusser tổng hợp và đánh giá trong công bố "Image registration methods: A survey" [138]). Phát hiện sự thay đổi là một kỹ thuật mạnh để giảm cảnh báo sai trong các kỹ thuật phát hiện mục tiêu và có thể được áp dụng trong các tình huống có hình ảnh tham chiếu trước của khu vực tìm kiếm mục tiêu. Tuy nhiên, phương pháp này phải đối phó với một số vấn đề, điển hình như: khoảng thời gian thu thập ảnh tham chiếu có sự thay đổi của khí quyển, độ nhạy của cảm biến, chiếu sáng nền, sự thay đổi tự nhiên của bề mặt mặt đất; sự sai lệch khi đăng ký hình ảnh, ... có thể ảnh hưởng đến hiệu suất phát hiện các đối tượng cần tìm kiếm. Ngoài ra, trong một số trường hợp, khu vực cần tìm kiếm không có đầy đủ ảnh tham chiếu đã lưu trữ trước đó dẫn đến phương pháp này không thực hiện được.

Cách tiếp cận thứ ba, phương pháp phát hiện dị thường được sử dụng để tìm những điểm ảnh dị thường. Các điểm ảnh dị thường trên ảnh đa phổ, siêu phổ được xác định là những điểm ảnh có phổ nổi bật hoặc khác biệt so với những điểm ảnh lân cận, những điểm ảnh này thưa thớt và hiếm khi đại diện cho ảnh. Tập hợp của các điểm ảnh dị thường liền kề nhau tạo thành một vùng rất nhỏ về mặt không gian và tồn tại với xác suất thấp trong một cảnh, những điểm ảnh này chứa mục tiêu cần tìm kiếm hoặc là nhiễu. Khác với cách tiếp cận thứ nhất là phải biết trước phổ mục tiêu, trong cách tiếp cận này, các thuật toán sẽ cố gắng để xác định vị trí bất cứ thứ gì khác biệt về mặt không gian hoặc quang phổ (màu sắc) từ môi trường xung quanh mà không cần biết trước về phổ mục tiêu. Trong công tác tìm kiếm cứu nạn, ngoài các mục tiêu đã biết trước như con người, các loại phương tiện như ô tô, tàu thuyền, máy bay, ... các mục tiêu nhân tạo có thể chứa đựng thông

tin cho công tác này. Các đối tượng như vậy về cơ bản chúng ta không biết rõ về phổ phản xạ của chúng, cũng không biết rõ về đặc trưng hình học hay màu sắc để có thể sử dụng phương pháp so khớp quang phổ hay từ một tập hợp dữ liệu về mục tiêu đã được huấn luyện. Do đó, việc lựa chọn phương pháp phát hiện dị thường phổ (hoặc màu sắc) để xác định các vector phổ từ các điểm ảnh bị lệch hoặc có sự khác biệt phổ đáng kể so với các điểm ảnh nền xung quanh tương đối phù hợp cho công tác tìm kiếm cứu nạn.

Trong hơn 20 năm qua, nhiều phương pháp phát hiện các điểm ảnh dị thường trên ảnh viễn thám quang học đã được xây dựng như: phương pháp phát hiện dựa trên mô hình xác suất thống kê; phương pháp phát hiện dựa trên các kỹ thuật phân cụm và phân đoạn; phương pháp phát hiện dựa trên biến đổi phi tuyến; phương pháp phát hiện dựa trên biến đổi tuyến tính và phương pháp phát hiện dựa trên học máy. Mỗi phương pháp đều có những ưu điểm riêng, tuy nhiên, vẫn còn tồn tại ba thách thức cần được giải quyết. Thách thức đầu tiên liên quan đến việc tăng tỷ lệ phát hiện trong khi giảm tỷ lệ báo động giả. Mỗi một điểm ảnh là dị thường thật, nó luôn chứa đựng nhiều thông tin bổ ích cho công tác tìm kiếm cứu nạn. Nếu như những điểm ảnh này lỡ bị bỏ qua, một phần nào đó làm giảm cơ hội phát hiện sớm vị trí người bị nạn, giảm cơ hội sống sót của nạn nhân. Thách thức thứ hai liên quan đến việc phát hiện dị thường với các hình dạng và kích cỡ khác nhau. Trong thực tế, kích thước của dị thường có thể dao động từ một phần điểm ảnh đến vài điểm ảnh và việc phát hiện các dị thường với kích thước khác nhau cùng một kỹ thuật là một thách thức lớn. Thách thức thứ ba, đó là tốc độ tính toán, các kỹ thuật phát hiện thường có thời gian tính toán gần với thời gian thực sẽ rất hữu ích trong các ứng dụng thực tế, đặc biệt là trong công tác tìm kiếm cứu nạn. Chính vì vậy, nghiên cứu sinh (NCS) đã lựa chọn và thực hiện đề tài "*Phát triển một số kỹ thuật phát hiện dị thường trên ảnh viễn thám quang học phục vụ công tác tìm kiếm cứu nạn*". Trong khuôn khổ một luận án tiến sĩ, rất khó để giải quyết cùng lúc cả ba thách thức trên nên luận án chỉ tập trung vào hai hướng nghiên cứu chính, phù hợp với công tác tìm kiếm cứu nạn, đó là đề xuất thuật toán nhằm nâng cao độ chính xác phát hiện các điểm ảnh dị thường và tăng tốc độ tính toán của thuật toán phát hiện các điểm ảnh dị thường dựa trên ước lượng mật độ nhân (KDE - Kernel Density Estimation) [120, 121].

Trong hướng nghiên cứu thứ nhất, luận án đề xuất thuật toán nhằm nâng cao độ chính xác phát hiện các điểm ảnh dị thường bằng cách giảm tỷ lệ dương tính giả. Do các hạn chế vật lý vốn có của các thiết bị chụp ảnh cũng như điều kiện khí quyển, thời tiết hay các thuật toán nén, hình ảnh có xu hướng bị nhiễu trong quá trình thu nhận và lưu trữ. Các điểm ảnh là nhiễu thường chiếm một phần nhỏ trong ảnh, trong khi các thuật toán phát hiện dị thường hoạt động giống như một bộ lọc nên nó thường được phân vào lớp dị thường (là những điểm dị thường giả hay còn gọi là dương tính giả), điều này làm giảm độ chính xác phát hiện dị thường của các thuật toán [14]. Để giảm số lượng các điểm ảnh dương tính giả, phương án luận án đề xuất là khử nhiễu dữ liệu đầu vào và định vị những khu vực nghi ngờ có chứa điểm dị thường. Trong những khu vực nghi ngờ chứa điểm ảnh dị thường, sử dụng thuật toán KDE để phát hiện các điểm ảnh dị thường, các điểm ảnh ở những vùng khác được xem là những điểm ảnh bình thường. Để khử nhiễu dữ liệu đầu vào, luận án sử dụng phương pháp làm mịn ảnh, bằng cách nhân tích chập ảnh gốc với toán tử Gaussian Blur. Để định vị những khu vực nghi ngờ có chứa các điểm ảnh dị thường, luận án sử dụng phương pháp trích rút các điểm ảnh đặc trưng bất biến cục bộ SIFT [71]. Như vậy, với việc giảm nhiễu đối với dữ liệu đầu vào và thu hẹp phạm vi tìm kiếm đã tăng độ chính xác phát hiện các điểm ảnh dị thường của thuật toán KDE.

Trong hướng nghiên cứu thứ hai, luận án đề xuất thuật toán nhằm tăng tốc độ tính toán của thuật toán đã đề xuất trong hướng nghiên cứu thứ nhất. Bằng cách kết hợp một số kỹ thuật về xử lý dữ liệu như nhóm các điểm ảnh có phổ (màu sắc) trùng nhau, sắp xếp nhóm các điểm ảnh hay sử dụng cây kd-tree để tìm kiếm nhanh nhóm điểm ảnh cần tham gia tính toán. Quá trình tính toán hàm mật độ xác suất cho các điểm ảnh sử dụng phương pháp tính toán đa luồng trên CPU (Central Processing Unit), song song trên GPU (Graphics Processing Unit).

Với hai thuật toán đề xuất của luận án sẽ góp phần xây dựng cơ sở lý thuyết ứng dụng ảnh viễn thám quang học trong công tác tìm kiếm cứu nạn.

2. Mục tiêu nghiên cứu

Đề xuất thuật toán nhằm tăng độ chính xác và giảm thời gian tính toán phát hiện các điểm ảnh dị thường trên ảnh viễn thám quang học theo phương

pháp thống kê phục vụ công tác tìm kiếm cứu nạn.

3. Nội dung nghiên cứu

- Đề xuất thuật toán tăng độ chính xác phát hiện các điểm ảnh dị thường của thuật toán KDE [120, 121] bằng cách giảm tỷ lệ dương tính giả (gọi tắt là thuật toán IKDE).
- Đề xuất thuật toán tăng tốc độ tính toán của thuật toán IKDE bằng cách sử dụng kỹ thuật nhóm điểm ảnh, sắp xếp điểm ảnh, áp dụng cây kd-tree và tính toán hàm mật độ xác suất đa luồng trên CPU hoặc song song trên GPU.
- Cài đặt, thực nghiệm nhằm kiểm chứng tính đúng đắn của các thuật toán đã đề xuất.

4. Đối tượng và phạm vi nghiên cứu

- Đối tượng nghiên cứu:
 - Các điểm ảnh dị thường trên ảnh viễn thám quang học.
 - Các thuật toán phát hiện điểm ảnh dị thường trên ảnh viễn thám quang học.
 - Các mô hình tính toán đa luồng trên CPU và tính toán song song trên GPU.
- Phạm vi nghiên cứu:

Với mục đích hỗ trợ tìm kiếm nhanh các điểm ảnh hoặc nhóm điểm ảnh chứa đựng thông tin hữu ích cho công tác tìm kiếm cứu nạn nhưng nó dễ bị bỏ qua do có kích thước nhỏ và bị trộn lẫn với nền hoặc bị chìm một phần dưới nước nên mắt người rất khó phát hiện. Vì vậy, luận án chỉ tập trung nghiên cứu phát hiện các điểm ảnh dị thường có màu sắc hoặc phổ bất kỳ trên ảnh đa phổ, siêu phổ thu được từ máy ảnh gắn trên các thiết bị bay.

5. Phương pháp nghiên cứu

- Về lý thuyết: dùng phương pháp khảo sát, phân tích, đánh giá, tổng hợp các kỹ thuật phát hiện dị thường trên ảnh đa phổ, siêu phổ. Từ đó đưa

ra những định hướng có thể nghiên cứu, phát triển nhằm cải thiện độ chính xác phát hiện cũng như tốc độ tính toán. Bên cạnh đó, luận án cũng nghiên cứu, vận dụng các kiến thức liên quan đến xử lý dữ liệu lớn, cấu trúc dữ liệu và giải thuật nâng cao, lý thuyết độ phức tạp tính toán.

- Về thực nghiệm: Luận án sử dụng phương pháp thực nghiệm để chứng minh tính đúng đắn của các phương pháp đề xuất. Sử dụng các ngôn ngữ lập C++ và C# để cài đặt các thuật toán. Sử dụng bộ dữ liệu mẫu do Phòng thí nghiệm khoa học và công nghệ quốc phòng (DSTL)- Vương quốc Anh cung cấp [24] và dữ liệu do NCS thu thập làm dữ liệu nguồn để đánh giá các thuật toán đã đề xuất.

6. Ý nghĩa khoa học và thực tiễn của luận án

Việc nghiên cứu, phát triển các thuật toán có khả năng triển khai, áp dụng trong thực tế nhằm nâng cao hiệu quả trong công tác tìm kiếm cứu nạn luôn là mục tiêu của các nhà khoa học trong nước và trên thế giới. Với các nghiên cứu được trình bày, luận án đã đóng góp về ý nghĩa khoa học và thực tiễn như sau:

- Ý nghĩa khoa học: luận án đã đề xuất thuật toán tăng độ chính xác và tốc độ tính toán của thuật toán phát hiện các điểm ảnh dị thường trên ảnh viễn thám quang học dựa trên hàm ước lượng mật độ nhân.
- Ý nghĩa thực tiễn: kết quả đề xuất của luận án làm tăng độ chính xác phát hiện các điểm ảnh dị thường của thuật toán KDE, nhất là trong trường hợp ảnh có nhiễu tăng cao. Thời gian tính toán của thuật toán đã giảm hơn 80% đối với ảnh đa phổ và siêu phổ, 99% đối với ảnh màu RGB, trong khi độ chính xác phát hiện các điểm ảnh dị thường tương đương với thuật toán gốc. Đây là cơ sở cho nghiên cứu phát triển phục vụ công tác tìm kiếm cứu nạn. Vì vậy, luận án có ý nghĩa thực tiễn.

7. Bố cục của luận án

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

Chương 1: TỔNG QUAN VỀ PHÁT HIỆN DỊ THƯỜNG TRÊN ẢNH VIỄN THÁM QUANG HỌC.

Chương này trình bày tổng quan những kiến thức liên quan đến ảnh viễn thám quang học, một số kỹ thuật phát hiện các điểm ảnh dị thường trên ảnh viễn thám quang học, tiêu chí để đánh giá độ chính xác phát hiện dị thường của các thuật toán. Đưa ra một số nhận định về hướng nghiên cứu. Ngoài ra, những kiến thức cơ sở cũng được trình bày, làm nền tảng cho các nghiên cứu, đề xuất trong các chương 2 và 3.

Chương 2: KỸ THUẬT NÂNG CAO ĐỘ CHÍNH XÁC PHÁT HIỆN DỊ THƯỜNG TRÊN ẢNH VIỄN THÁM QUANG HỌC.

Trong chương này, luận án đề xuất thuật toán tăng độ chính xác phát hiện các điểm ảnh dị thường của thuật toán phát hiện các điểm ảnh dị thường dựa trên KDE [120, 121]. Bằng cách loại bỏ bớt nhiều khối tập các điểm ảnh dị thường được tìm thấy của thuật toán trên (giảm tỷ lệ dương tính giả). Cài đặt thuật toán, thử nghiệm, đánh giá kỹ thuật đề xuất.

Chương 3: KỸ THUẬT TĂNG TỐC ĐỘ PHÁT HIỆN DỊ THƯỜNG TRÊN ẢNH VIỄN THÁM QUANG HỌC.

Chương này trình bày đề xuất thuật toán tăng tốc độ tính toán của thuật toán phát hiện các điểm ảnh dị thường đã đề xuất trong Chương 2 thông qua việc thực hiện thuật toán theo bốn bước. Bước đầu tiên là nhóm các điểm ảnh có cùng giá trị, bước thứ hai là sắp xếp các nhóm điểm ảnh này tăng dần theo chiều dữ liệu thứ nhất. Bước thứ ba, tính toán hàm mật độ xác suất của các nhóm điểm ảnh bằng một trong ba kỹ thuật: sử dụng cây kd-tree để quản lý và tìm kiếm nhanh những nhóm điểm ảnh cần tham gia tính toán; sử dụng phương pháp tính toán đa luồng trên CPU (Central Processing Unit) hoặc song song trên GPU (Graphics Processing Unit). Bước cuối cùng, phân loại các điểm ảnh dựa theo giá trị mật độ xác suất đã được tính toán ở bước thứ ba. Cài đặt thuật toán, thử nghiệm, đánh giá thuật toán đề xuất.

Chương 1

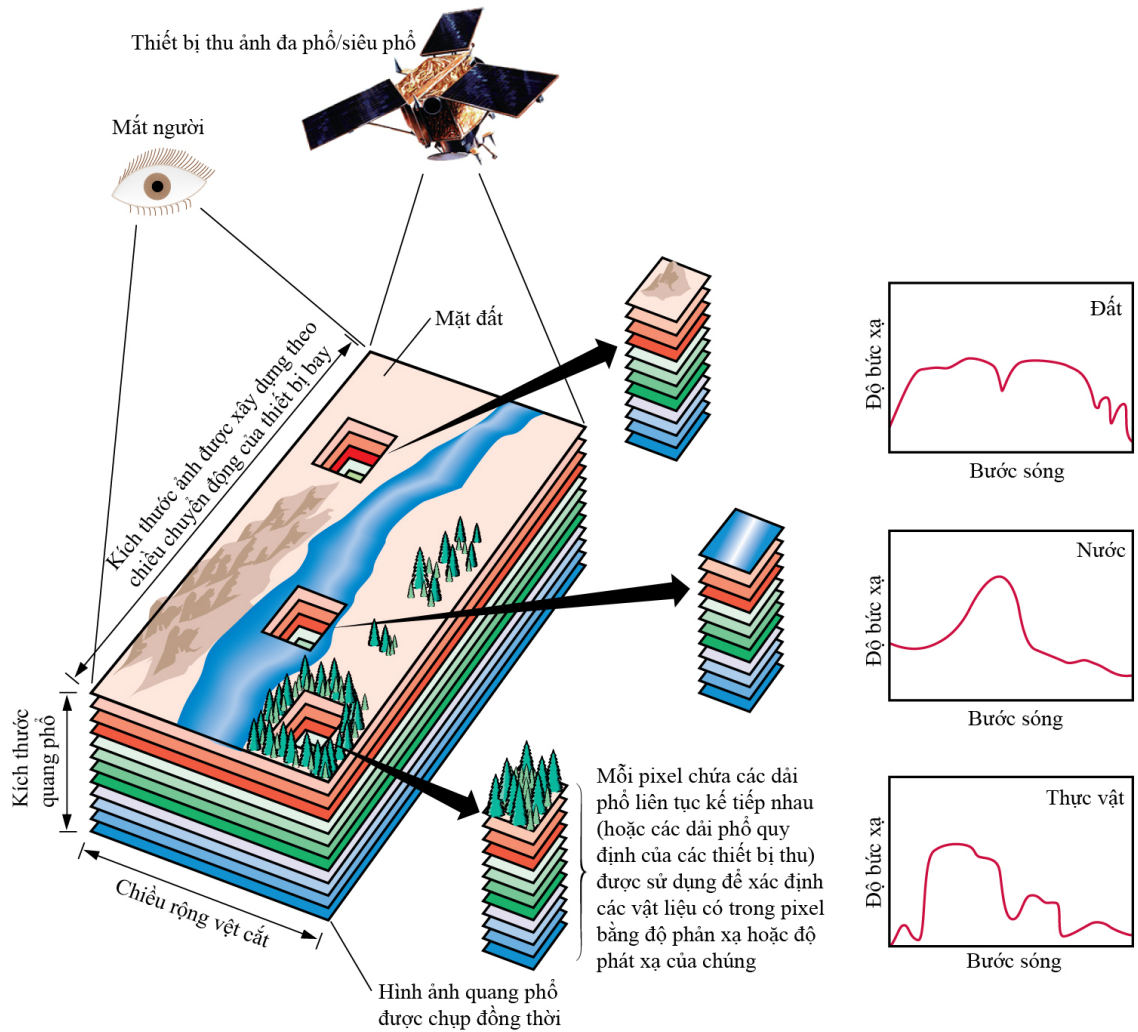
TỔNG QUAN VỀ PHÁT HIỆN DỊ THƯỜNG TRÊN ẢNH VIỄN THÁM QUANG HỌC

Nội dung trong chương trình bày những kiến thức cơ bản liên quan đến ảnh viễn thám quang học, một số phương pháp phát hiện các điểm ảnh dị thường trên ảnh viễn thám quang học, một số nhận định về hướng nghiên cứu và tiêu chí để đánh giá hiệu quả phát hiện dị thường của các thuật toán. Những nội dung này đã được công bố trong [CT1].

1.1. Ảnh viễn thám quang học

Viễn thám (Remote sensing) là khoa học thu nhận thông tin về một vật thể, một vùng, hoặc một hiện tượng qua phân tích dữ liệu thu được bởi phương tiện không tiếp xúc với vật, vùng hoặc hiện tượng khi khảo sát [67]. Hiểu đơn giản, viễn thám là thăm dò từ xa về một đối tượng hoặc một hiện tượng mà không có sự tiếp xúc trực tiếp với đối tượng hoặc hiện tượng đó. Điều này thực hiện được do các vật thể khác nhau hấp thụ và phát ra bức xạ điện từ (photon) theo cách đặc trưng của thành phần hóa học cấu tạo nên nó. Nếu đo năng lượng của bức xạ này bằng một thiết bị theo bước sóng, kết quả thu được là tín hiệu quang phổ và như vậy, quang phổ có thể được sử dụng để mô tả đặc trưng cho bất kỳ vật liệu nào. Sự kết hợp giữa quang phổ, công nghệ hình ảnh và phương pháp để thu thập thông tin quang phổ trên các khu vực rộng lớn được gọi là quang phổ hình ảnh. Nguyên lý phổ hình ảnh cơ bản được minh họa trong Hình 1.1.

Ảnh viễn thám quang học chính là loại ảnh được tạo ra bởi việc thu nhận các bước sóng trong dải ánh sáng nhìn thấy ($0,4 - 0,7 \mu m$), cận hồng ngoại ($0,7 - 1,3 \mu m$) và hồng ngoại ($1,3 - 2,7 \mu m$) qua các thiết bị bay (vệ tinh, kính khí cầu, máy bay,...) sử dụng các ống kính quang học. Từ khi ảnh viễn thám quang học ra đời đã và đang mang lại rất nhiều lợi ích phục vụ đời sống xã hội, nó được công nhận trong nhiều ứng dụng, chẳng hạn như trong nông nghiệp và lâm nghiệp [5], công nghiệp thực phẩm [35], rà phá bom mìn [8, 55], y học [72], tìm kiếm cứu nạn [74, 56] và tài nguyên nước [34].



Hình 1.1: Nguyên lý của ảnh quang phổ [94]

1.1.1. Phổ của ảnh viễn thám quang học

1.1.1.1. Thu nhận phổ phản xạ của các vật thể

Bộ cảm (sensor) giữ nhiệm vụ thu nhận các năng lượng bức xạ do vật thể phản xạ từ nguồn cung cấp tự nhiên (mặt trời) hoặc nhân tạo (do chính vệ tinh phát). Năng lượng này được chuyển thành tín hiệu số tương ứng với năng lượng bức xạ ứng với từng bước sóng do bộ cảm nhận được trong dải phổ đã được xác định.

Trong bộ cảm, kính lọc phổ được sử dụng để tách năng lượng bức xạ ứng với từng bước sóng khác nhau, và năng lượng này được dẫn đến các tế bào quang điện để biến đổi quang năng thành điện năng. Độ lớn của dòng điện

tạo ra trong tế bào quang điện tỷ lệ thuận với cường độ của sóng phản xạ từ vật thể. Do đó, sự thay đổi của dòng điện có thể được sử dụng để đo lường sự thay đổi năng lượng của ánh sáng mạnh hay yếu ứng với từng bước sóng khác nhau. Năng lượng sóng điện từ sau khi tới tế bào quang điện, được chuyển thành tín hiệu có giá trị thay đổi liên tục theo thời gian (số thực) và để ghi nhận thông tin, việc chuyển đổi tín hiệu điện liên tục thành một số nguyên hữu hạn được gọi là số hoá. Giá trị độ sáng ghi nhận được phụ thuộc vào số bit dùng trong quá trình số hoá và toàn bộ năng lượng của sóng điện từ thu được (tín hiệu nhập) chuyển sang tín hiệu số (tín hiệu xuất) chỉ sử dụng phần biến đổi tuyến tính của bộ cảm.

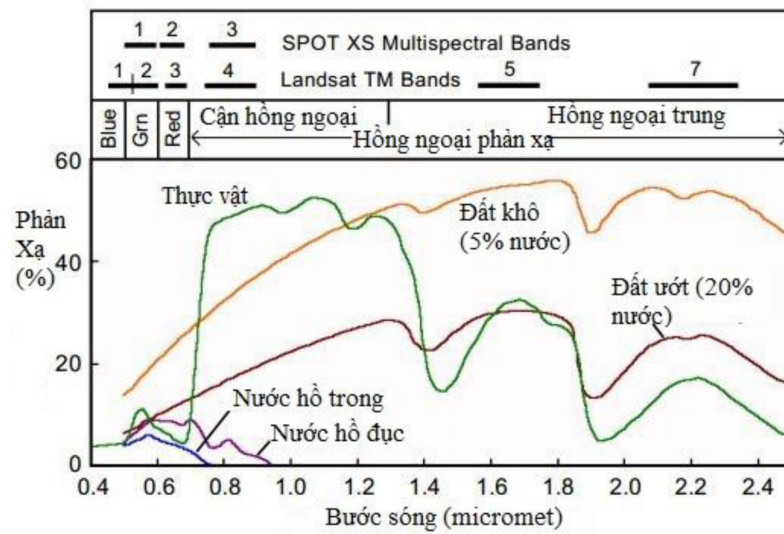
1.1.1.2. Đặc trưng phổ phản xạ một số đối tượng

Tỉ số giữa năng lượng phản xạ và năng lượng tới là một hàm của bước sóng. Giá trị phản xạ sẽ thay đổi theo giá trị bước sóng đối với hầu hết các vật liệu bởi vì năng lượng ở mỗi bước sóng nhất định sẽ tán xạ hoặc hấp thụ ở các mức độ khác nhau, chúng thể hiện rõ khi chúng ta so sánh các đường cong quang phổ phản xạ (biểu đồ phản xạ so với bước sóng) cho các vật liệu khác nhau. Đường vồng xuống của các đường cong quang phổ đánh dấu dải bước sóng dao động mà vật liệu hấp thụ năng lượng tới. Các đặc điểm này thường được gọi là dải hấp thụ. Hình dạng tổng thể, vị trí, khả năng hấp thụ của đường cong quang phổ trong nhiều trường hợp có thể được sử dụng để xác định và phân biệt các vật liệu khác nhau. Ví dụ: thực vật có giá trị phản xạ cao hơn trong phạm vi dải hồng ngoại gần và thấp hơn trong vùng ánh sáng đỏ so với đất. Trên Hình 1.2, những đường cong phản xạ đại diện cho một số vật liệu phổ biến trên bề mặt Trái đất trên phạm vi ánh sáng nhìn thấy và quang phổ hồng ngoại.

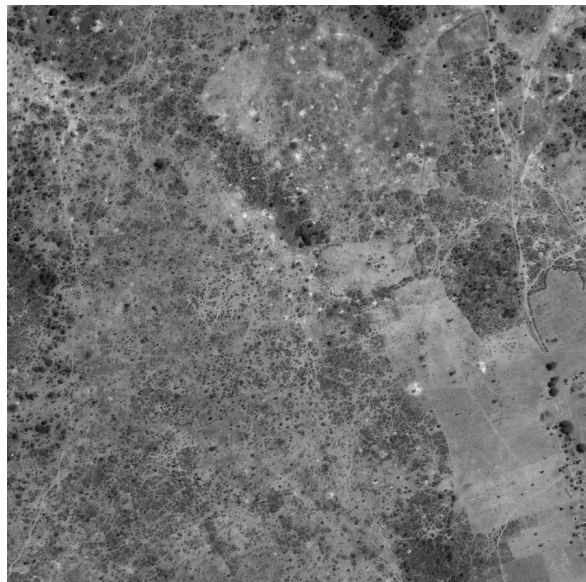
1.1.2. Các loại ảnh viễn thám quang học

1.1.2.1. Ảnh toàn sắc

Ảnh toàn sắc là loại ảnh được tạo ra bởi các cảm biến đo năng lượng phản xạ chỉ trong một phần cụ thể của phổ điện từ, thường là ảnh đen trắng được thu nhận trong dải sóng nhìn thấy. Bộ cảm toàn sắc thường giới hạn bước sóng nhìn thấy trong khoảng từ 0,5 - 0,8 μm , nó cung cấp dữ liệu ảnh có độ phân giải không gian thường tốt hơn dữ liệu ảnh đa phổ trên cùng một thiết



Hình 1.2: Biểu đồ mối quan hệ giá trị phản xạ và bước sóng [85]



Hình 1.3: Ví dụ một ảnh toàn sắc, ảnh có mã là 6010_1_2_P do DSTL cung cấp [24]

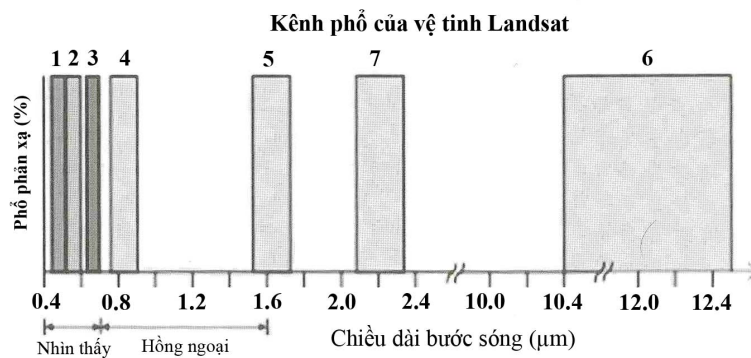
bị mang. Ví dụ: vệ tinh QuickBird cung cấp dữ liệu ảnh toàn sắc có độ phân giải không gian là 0,6m, trong khi đó độ phân giải của ảnh đa phổ là 2,4m.

Ảnh toàn sắc thường được kết hợp với tư liệu ảnh đa phổ để tạo ra ảnh kết hợp có độ phân giải cao như của tư liệu toàn sắc. Bộ cảm toàn sắc thường ứng dụng trong địa chất, sinh học, đo đạc khảo sát hoặc thành lập bản đồ. Ngoài ra, tư liệu ảnh toàn sắc được sử dụng trong thành lập mô hình số bề

mặt DSM (Digital Surface Model) có độ chính xác cao.

1.1.2.2. Ảnh đa phổ

Ảnh đa phổ được tạo ra bởi các cảm biến đo năng lượng phản xạ trong một số phần cụ thể (còn gọi là các dải) của phổ điện từ. Cảm biến thu ảnh đa phổ thường có từ 3 đến 10 (hoặc nhiều hơn) băng phổ khác nhau trong mỗi pixel của hình ảnh chúng tạo ra. Một loại ảnh đa phổ nổi tiếng là ảnh màu RGB, ảnh này gồm ba kênh phổ đó là kênh phổ màu đỏ có bước sóng trong khoảng $0,63 - 0,69\mu\text{m}$, màu xanh lá cây có bước sóng trong khoảng $0,52 - 0,60\mu\text{m}$ và màu xanh lam có bước sóng trong khoảng $0,45 - 0,52\mu\text{m}$. Landsat, Quickbird và vệ tinh Spot là các cảm biến vệ tinh nổi tiếng sử dụng cảm biến đa phổ để tạo ra các ảnh đa phổ.



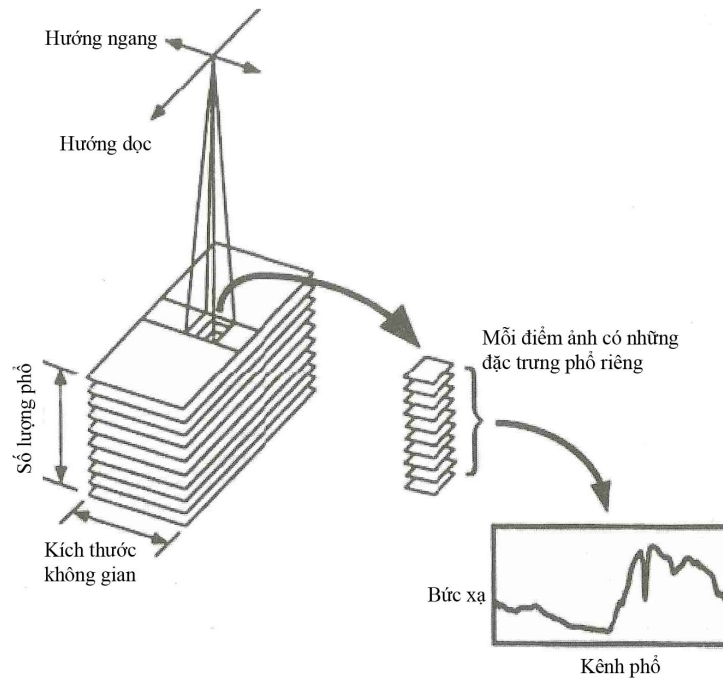
Hình 1.4: Các kênh đa phổ và độ rộng của một bộ cảm biến đa phổ TM/ETM+ [2]

Mỗi một điểm ảnh trên ảnh đa phổ, tín hiệu thu được sẽ bao gồm các tín hiệu trong vùng quang phổ nhìn thấy được (từ $0,4 - 0,7 \mu\text{m}$) và sẽ đến bước sóng hồng ngoại (từ $0,8 - 10 \mu\text{m}$) hoặc dài hơn được phân loại là vùng gần hồng ngoại (NIR), hồng ngoại trung (MIR) và hồng ngoại xa (FIR).

1.1.2.3. Ảnh siêu phổ

Ảnh siêu phổ được tạo ra bởi các cảm biến siêu phổ, cảm biến này đo năng lượng trong khoảng hẹp hơn và nhiều dải hơn so với cảm biến đa phổ. Các cảm biến siêu phổ lấy mẫu trong một phần của phổ điện từ ánh sáng có thể nhìn thấy ($0,4 - 0,7\mu\text{m}$) đến cận hồng ngoại và hồng ngoại (khoảng $2,7\mu\text{m}$) trong hàng trăm dải tiếp giáp hẹp rộng khoảng 10nm (mỗi một băng phổ có độ rộng là 10nm). Độ phân giải phổ cao như vậy có thể phân biệt giữa các

vật liệu khác nhau trên mặt đất.



Hình 1.5: Cấu trúc của một bộ cảm siêu phổ. Các kênh ảnh được thu nhận đồng thời trên nhiều kênh phổ hẹp [2]

1.1.3. Độ phân giải không gian

Độ phân giải không gian là kích thước nhỏ nhất của một vật mà bộ cảm có thể nhận biết được về một đối tượng không gian phân cách được với đối tượng không gian khác nằm kề đối tượng này. Độ lớn của điểm ảnh sẽ là đơn vị xác định độ phân giải không gian của hệ thống. Độ phân giải không gian của các ảnh số có thể từ 0,6 m (vệ tinh Quickbird), 1 m (SPIN2 và IKONOS), đến 6.4 m, 10 m (SPOT), và 1 km (ảnh vệ tinh NOAA). Một ảnh số có độ phân giải càng cao là ảnh có kích thước pixel càng nhỏ. Hình 1.6 cho thấy khuôn viên của Đại học Maryland tại College Park ở bốn độ phân giải không gian khác nhau. Bảng 1.7 cho thấy độ phân giải không gian của một số cảm biến phổ biến.

1.1.4. Cấu trúc dữ liệu của ảnh viễn thám quang học

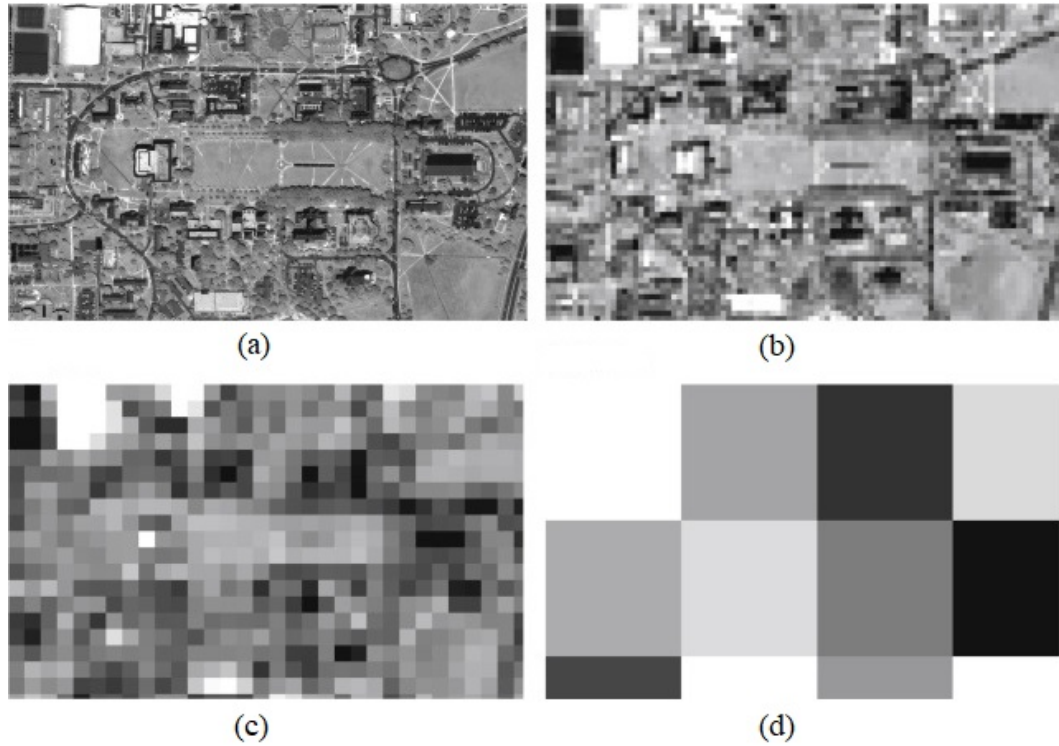
Thông thường, ảnh số được tổ chức dưới dạng một ma trận với số hàng và số cột hữu hạn. Nếu ký hiệu số hàng là h và số cột là w thì số điểm ảnh là $h \times w$. Dung lượng của một ảnh đa phổ, siêu phổ có k kênh phổ sẽ là

Bảng 1.1: Bảng so sánh các loại ảnh viễn thám

Loại ảnh	Ưu điểm	Hạn chế
Ảnh toàn sắc	Có độ phân giải không gian thường tốt hơn dữ liệu ảnh đa phổ, siêu phổ trên cùng một thiết bị mang; dung lượng cần thiết để lưu trữ thấp hơn ảnh đa phổ và ảnh siêu phổ.	Do bộ cảm chỉ thu nhận trong một khoảng giới hạn nên tư liệu ảnh toàn sắc thiếu thông tin phổ các đối tượng lớp phủ trên bề mặt, dẫn đến khó khăn khi giải đoán các đối tượng trên địa hình.
Ảnh đa phổ	Lợi thế của ảnh đa phổ là có thể sử dụng tích hợp các kênh phổ khác nhau để phân tích giải đoán các đối tượng theo các đặc trưng bức xạ phổ; ảnh đa phổ dễ sử dụng và xử lý ít phức tạp hơn ảnh siêu phổ vì số lượng dải tần ít hơn.	Có độ phân giải không gian thường thấp hơn ảnh toàn sắc. Vì vậy, để tăng chất lượng của ảnh thì phải kết hợp với ảnh toàn sắc để tạo ra ảnh kết hợp có độ phân giải cao như của tư liệu ảnh toàn sắc.
Ảnh siêu phổ	Ưu điểm chính của ảnh siêu phổ đó là thông tin phổ của các đối tượng lớp phủ bề mặt. Xử lý thông tin cho phép phân loại chính xác được các đối tượng.	Nhược điểm chính của tư liệu siêu phổ là có chi phí cao và xử lý hình ảnh phức tạp. Ngoài ra dung lượng lưu trữ dữ liệu lớn nên khả năng lưu trữ trên các thiết bị mang và dưới mặt đất cũng là yêu cầu quan trọng đối với loại ảnh này.

$k \times h \times w$, nếu như bỏ qua các phần thông tin bổ trợ và coi mức lượng tử hóa của mỗi một điểm ảnh là 8 bit. Trong trường hợp mức lượng tử hóa của mỗi một điểm ảnh nhiều hơn 8 bit và nhỏ hơn hoặc bằng 16 bit thì dung lượng của ảnh đa phổ, siêu phổ sẽ là $2 \times k \times h \times w$. Trong viễn thám hiện đại, mức lượng tử hóa của ảnh quang học thường từ 8 đến 14 bit, khi đó dung lượng của ảnh sẽ được tính theo byte theo công thức sau:

$$V = h \times w \times k \times q \quad (1.1)$$



Hình 1.6: Khuôn viên của Đại học Maryland tại College Park với bốn phân giải không gian khác nhau [66]. Ảnh (a) có độ phân giải không gian là $1\text{m}/\text{pixel}$; ảnh (b) có độ phân giải không gian là $10\text{m}/\text{pixel}$; ảnh (c) có độ phân giải không gian là $30\text{m}/\text{pixel}$ và ảnh (d) có độ phân giải không gian là $250\text{m}/\text{pixel}$ [2]

trong đó: V là dung lượng theo byte, h là số hàng, w là số cột, k số kênh phổ, q là mức lượng tử hóa bằng 1 (8 bit) hoặc bằng 2 (16 bit).

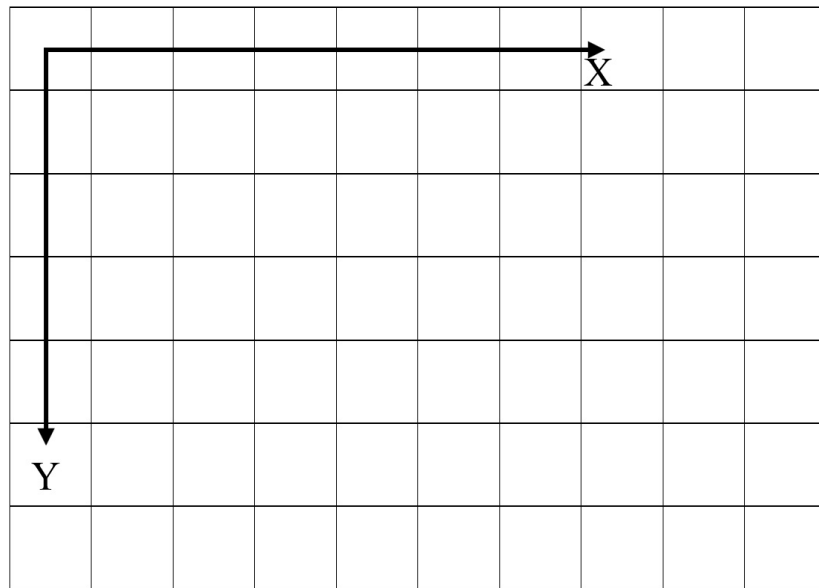
Mỗi điểm ảnh được xác định trong hệ tọa độ ảnh bằng một cặp chỉ số hàng và cột. Hệ tọa độ ảnh thông thường có gốc nằm ở góc trên bên trái (xem Hình 1.7). Như vậy, chỉ số cột tăng dần từ trái sang phải và chỉ số hàng tăng dần từ trên xuống dưới. Tọa độ ảnh là số nguyên và thường bắt đầu bằng số hàng số cột bằng 1 (tùy theo ngôn ngữ lập trình sử dụng). Đối với mỗi loại tư liệu, cần tìm hiểu cách bố trí điểm gốc hệ tọa độ thông qua các tài liệu kỹ thuật kèm theo.

Nếu biết được tọa độ trong hệ tọa độ phẳng Universal Transverse Mercator (UTM) của ô trên cùng bên trái trong mảng và khoảng cách các điểm ảnh (tính theo đơn vị m) thì chúng ta có thể tính toán vị trí của bất kỳ điểm nào trong ảnh. Giá trị độ xám của các điểm ảnh được lưu trữ trong các phần

Bảng 1.2: Độ phân giải không gian của một số loại bộ cảm vệ tinh [66]

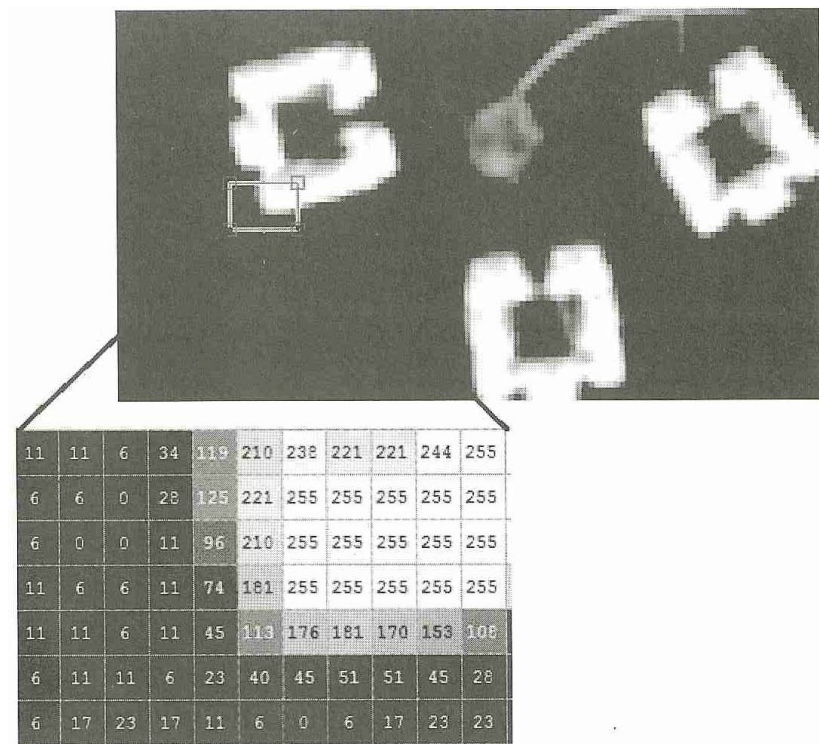
Loại	Bộ cảm vệ tinh	Giải quang phổ	Độ phân giải (m)	Mức lượng tử hóa (bit)
Độ phân giải thấp (>1000m)	POLDER	B1-B9	6000×7000	12
Độ phân giải trung bình (100 - 1000m)	MODIS	B1-B2	250	12
		B3-B7	500	
		B8-B36	1000	
	AVHRR	B1-B5	1100	10
Độ phân giải cao (5-100m)	ALI/EO1 ASTER/Terra	B1	15	8
		B5-B9	30	
		B11-B14	90	12
	ETM+/ Landsat 7	Pan	15	8
		B1-B5, B7 B6	30 60	
	HRV/ SPOT5	Pan	2.5 hoặc 5	8
		B1-B3	10	
SW-IR		20		
Độ phân giải siêu cao (<5m)	Ikonos	Pan	0.82	11
		B1-B4	3.2	
	Quickbird	Pan	0.61	11
		B1-B4	2.44	
	World view	Pan	0.5	11
	Geoeye-1	Pan	1.41	11
B1-B4		1.65		

tử mảng nằm trong một phạm vi xác định, thường là 0-255, tương ứng với phạm vi độ sáng của màu cơ bản liên kết với mảng hình ảnh đó. Giá trị 0 là



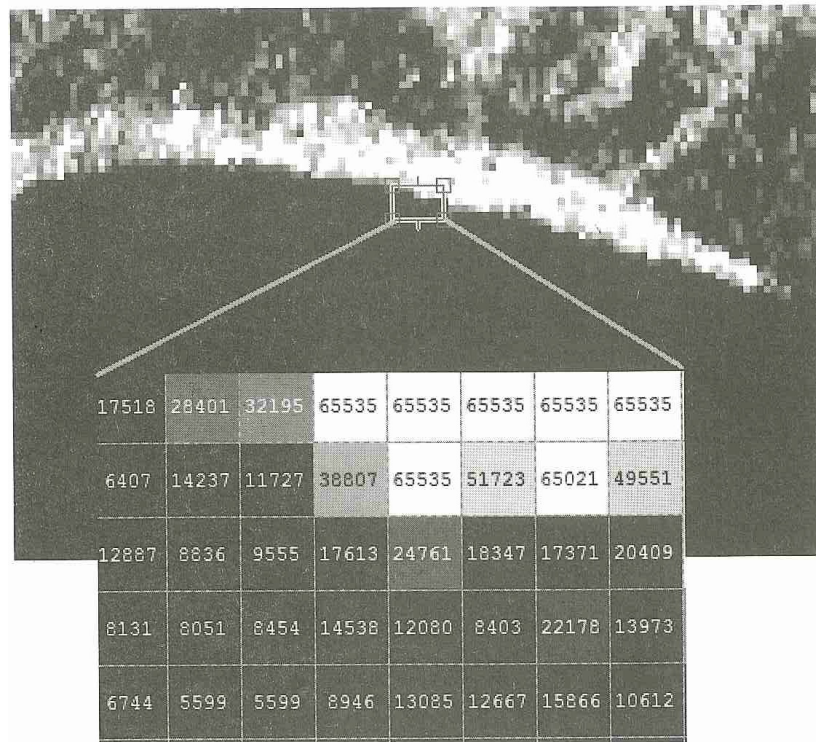
Hình 1.7: Ví dụ minh họa hệ tọa độ ảnh số

mức thấp nhất và giá trị 255 là mức sáng nhất mà màu đó được hiển thị.



Hình 1.8: Kênh Pan trên ảnh SPOT được mã hóa 8 bit [2]

Mức lượng tử hóa của mỗi điểm ảnh thường là 8 bit (0-255) tương ứng



Hình 1.9: Ảnh Envisat ASAR được mã hóa 16 bit [2]

với phạm vi độ sáng của màu sắc liên kết với màu hình (xem Hình 1.8). Tuy nhiên, không phải tất cả các ảnh vệ tinh đều có giá trị pixel nằm trong khoảng 0-255. Dữ liệu AVHRR (Advanced Very High Resolution Radiometer) sử dụng một mức lượng tử từ 0-1023. Giá trị điểm ảnh IKONOS (là một vệ tinh thương mại quan sát Trái đất, thu thập hình ảnh ở độ phân giải cao, được thiết kế bởi Lockheed Martin Space Systems) nằm trong phạm vi 0-2047 và các kênh nhiệt của tư liệu ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) được nằm trong khoảng 0-4095. Trong dữ liệu vệ tinh ASTER, ví dụ “0” và “4095”, chúng được sử dụng để chỉ ra “dữ liệu xấu” và giá trị bão hòa. Khoảng giá trị của cường độ điểm ảnh được gọi là miền giá trị ảnh - dynamic range của tư liệu ảnh. Các mức lượng tử hóa của pixel được thể hiện trong máy tính bởi một tập hợp các chữ số nhị phân có thể được coi là kỹ thuật công tắc “bật/tắt” hoặc dấu chấm và dấu gạch ngang trong mã Morse. Ví dụ, nếu mức lượng tử hóa của pixel là 8 chữ số nhị phân (binary digits-bits) thì giá trị màu được lưu trữ trong mỗi pixel có giá trị tối thiểu là 0 và tối đa là 255, tương ứng trong hệ nhị phân là 00000000

và 11111111. Trong ảnh màu, nếu mỗi màu cơ bản được lưu trong phạm vi 8 bit thì hình ảnh được tạo ra từ việc kết hợp 3 màu cơ bản tạo nên “hình ảnh 24 bit”. Hình ảnh băng tần đơn 10 bit cung cấp 1024 mức xám, trong khi hình ảnh 16 bit có thể đại diện cho dữ liệu dương và âm trong phạm vi +32767 đến -32768, đây là trường hợp dữ liệu số nguyên 16 bit có dấu. Nếu là dữ liệu số nguyên không dấu 16 bit sẽ có giá trị trong phạm vi 0-65535 (xem Hình 1.9).

1.1.5. Thư viện quang phổ và thư viện ảnh viễn thám quang học

1.1.5.1. Thư viện quang phổ

Một số thư viện phổ phản xạ của các vật liệu tự nhiên hay nhân tạo có sẵn để phục vụ cộng đồng những nhà nghiên cứu. Các thư viện này cung cấp một nguồn tài liệu tham khảo về quang phổ có thể lý giải hình ảnh đa phổ và siêu phổ, cụ thể:

- **Thư viện phổ ASTER:** là thư viện được NASA tạo ra, bao gồm các phổ được biên soạn từ phòng thí nghiệm Jet Propulsion của NASA, Đại học Johns Hopkins, và cục khảo sát địa chất Hoa Kỳ. Thư viện phổ ASTER có gần 2000 quang phổ bao gồm các khoáng sản, đá, đất, vật liệu nhân tạo, nước và tuyết. Nhiều phổ trong thư viện bao gồm toàn bộ khu vực bước sóng 0.4-1.4 μm . Thư viện này có thể truy cập thông qua website <http://speclib.jpl.nasa.gov>. Có thể tìm kiếm phổ theo các thể loại, có thể xem trước biểu đồ quang phổ và tải về các dữ liệu dưới dạng file text.
- **Thư viện phổ USGS:** phòng thí nghiệm khảo sát phổ địa chất tại Denver, Colorado đã biên soạn một thư viện khoảng 500 quang phổ phản xạ khoáng sản và một số thảm thực vật trên phạm vi bước sóng 0.2-0.3 μm . Có thể truy cập vào thư viện này thông qua địa chỉ trực tuyến <http://speclab.cr.usgs.gov/spectral.lib04/spectral-lib04.html>.

Một điểm cần lưu ý, các thư viện quang phổ trên được tạo ra từ các phòng thí nghiệm, sử dụng máy quang phổ cầm tay do đó, nó có thể khác với quang phổ thu được từ vệ tinh hay các thiết bị bay khác. Lý do có sự khác nhau bởi sự ảnh hưởng của khí quyển, thời tiết, ánh sáng, vật liệu liền kề, ...

1.1.5.2. Thư viện ảnh viễn thám quang học trong phát hiện dị thường

Có một thực tế rằng, cho đến nay, thế giới chưa có một thư viện ảnh viễn thám quang học nào dùng chung cho bài toán phát hiện các điểm ảnh dị thường, dữ liệu ảnh đa phổ hoặc siêu phổ thu chụp trong các tình huống tìm kiếm cứu nạn thực tế không được phát hành công khai bởi đây là dữ liệu mật của mỗi quốc gia. Do đó, để kiểm nghiệm đề xuất trong các công trình đã công bố về phát hiện các điểm ảnh dị thường trên ảnh viễn thám quang học các nhà khoa học thường sử dụng 3 loại sau:

- Loại dữ liệu thứ nhất, cấy các mẫu dị thường vào trong ảnh, mục đích của việc cấy các dị thường vào trong ảnh nhằm đánh giá hoạt động của các thuật toán phát hiện trong một môi trường được kiểm soát hoàn toàn. Các công trình đã công bố sử dụng loại dữ liệu này như: [12, 25, 39, 61, 76, 80, 82, 96, 124, 125, 127, 130, 133, 137].

- Loại dữ liệu thứ hai, bố trí các mẫu vật, phương tiện, ... trên thực địa, sau đó sử dụng máy ảnh để chụp khu vực đã được bố trí. Các mẫu vật, phương tiện, ... được xem là dị thường trên ảnh. Các công trình sử dụng loại dữ liệu này như: [9, 38, 40, 48, 58, 59, 61, 62, 63, 69, 75, 88, 89, 98, 112, 114, 119, 124, 131, 134, 137].

- Loại dữ liệu thứ ba, sử dụng ảnh được cung cấp bởi những đơn vị chuyên về ảnh viễn thám như NASA hoặc những ảnh được cung cấp miễn phí trên internet như:

- + Thư viện ảnh đa phổ do Phòng thí nghiệm khoa học và công nghệ quốc phòng (DSTL)- Vương quốc Anh cung cấp tại địa chỉ <https://www.kaggle.com/c/dstl-satellite-imagery-feature-detection>. Đây là những ảnh được dùng trong cuộc thi "Dstl Satellite Imagery Feature Detection" do DSTL tổ chức [24]. Trong thư viện có 3 loại ảnh, ảnh 3 kênh phổ (ảnh màu RGB), ảnh 8 kênh phổ được chụp trong dải hồng ngoại gần (400 - 1040nm) và ảnh 8 kênh phổ chụp ở dải hồng ngoại ngắn (1195 - 2365nm). Những ảnh này được thu từ bộ cảm biến WorldView 3, phạm vi không gian trong một ảnh là $1km^2$ ($1km \times 1km$).

- + Thư viện ảnh siêu phổ do M. Graña và các cộng sự thu thập, cung cấp miễn phí tại địa chỉ https://www.ehu.es/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes. Trong thư viện ảnh siêu phổ này có 05 ảnh gồm:

1. Indian Pines: cảnh này được cảm biến AVIRIS thu chụp ở Indian Pines, Tây Bắc Indiana, bao gồm 145×145 điểm ảnh và 224 kênh phổ, thu chụp trong dải sóng $0.4\text{--}2.5 \mu\text{m}$.
2. Salinas: cảnh này được cảm biến AVIRIS thu chụp ở Thung lũng Salinas, California, bao gồm 512×217 điểm ảnh và 224 kênh phổ, ảnh có độ phân giải không gian $3.7\text{m}/\text{điểm ảnh}$.
3. Pavia Centre and University: đây là hai cảnh được cảm biến ROSIS thu chụp khi bay qua Pavia, miền Bắc nước Ý. Số kênh phổ là 102 đối với Trung tâm Pavia và 103 đối với Đại học Pavia. Trung tâm Pavia có kích thước 1096×1096 điểm ảnh và Đại học Pavia là 610×610 điểm ảnh, độ phân giải không gian là $1.3\text{m}/\text{điểm ảnh}$.
4. Kennedy Space Center: cảnh này được cảm biến AVIRIS thu chụp ở Trung tâm Không gian Kennedy (KSC), Florida, vào ngày 23 tháng 3 năm 1996, bao gồm 224 kênh phổ, thu chụp trong dải $0.4\text{--}2.5 \mu\text{m}$.
5. Botswana: cảnh này được Vệ tinh EO-1 của NASA thu chụp ở đồng bằng Okavango, Botswana trong năm 2001-2004. Cảm biến Hyperion trên EO-1 thu thập dữ liệu ở độ phân giải $30\text{m}/\text{điểm ảnh}$, 242 kênh phổ trong dải $0.4\text{--}2.5 \mu\text{m}$.

Trong các công trình đã công bố, thông thường, những ảnh được lựa chọn chứa những công trình nhân tạo, phương tiện giao thông, máy bay, tàu, thuyền, ...; những đối tượng này được xem là các dị thường đối với môi trường xung quanh nó. Các công trình sử dụng loại dữ liệu này như: [9, 25, 28, 29, 39, 45, 51, 59, 61, 63, 65, 68, 75, 76, 82, 88, 93, 112, 114, 119, 120, 124, 125, 126, 127, 130, 132, 133, 134, 137].

Để kiểm nghiệm các đề xuất, luận án sẽ sử dụng cả 3 loại dữ liệu như mô tả ở trên.

1.2. Bài toán phát hiện dị thường trên ảnh viễn thám quang học phục vụ công tác tìm kiếm cứu nạn

1.2.1. Khái niệm dị thường

Theo M. Hawkins, dị thường là một sự quan sát (observation) xa rời những quan sát khác tạo nên sự ngờ vực rằng nó được tạo ra bởi một cơ chế khác

[6]. Dị thường còn được gọi là bất thường hay dị biệt trong khai phá dữ liệu và xác suất thống kê. Trong hầu hết các ứng dụng, dữ liệu được tạo bởi các tiến trình có thể phản ánh hoạt động của hệ thống hoặc các hoạt động quan sát về các thực thể. Khi tiến trình diễn ra một cách bất thường, nó đưa ra các dữ liệu bất thường.

Theo K. S. Ettabaa và cộng sự [32], trong viễn thám, phát hiện dị thường nhằm phát hiện các điểm ảnh dị thường trong các ảnh đa phổ và siêu phổ. Dị thường thường là các điểm ảnh có dấu hiệu quang phổ khác biệt đáng kể so với nền cục bộ và có thể chứa thông tin hữu ích. Ví dụ, dị thường có thể là loài thực vật quý hiếm, dị thường trong tăng trưởng thực vật, thực vật bất hợp pháp liên quan đến buôn bán ma túy, khu vực bị ô nhiễm ở vùng nước ven biển, các nhà thám hiểm bị lạc, cấu trúc khảo cổ bị chôn vùi, vượt biên bất hợp pháp hoặc xe quân sự dưới lớp phủ thực vật, v.v. Các đối tượng như vậy thường xuất hiện dưới dạng dị thường là sự khác biệt với môi trường xung quanh [14]. Lý do của sự dị thường là bởi vì chúng không được biết trước, mặt khác, các dị thường thường xảy ra với xác suất thấp. Do đó, sự tồn tại của chúng nói chung không thể được phát hiện bằng bất kỳ phương pháp được giám sát hoặc kiểm tra trực quan nào. Đối với sự hiện diện, phạm vi không gian của dị thường khá hạn chế vì chúng có thể chỉ nằm trong một phần của điểm ảnh ngoài ra nó còn bị trộn lẫn với nền hoặc các chất khác quanh nó.

Như vậy, dị thường trên ảnh viễn thám quang học có bốn đặc điểm chính [39]:

1. Không có kiến thức về sự tồn tại (các điểm ảnh dị thường có màu sắc hoặc phổ bất kỳ);
2. Có dấu hiệu quang phổ khác biệt đáng kể so với nền cục bộ (đây là đặc điểm quan trọng nhất);
3. Xác suất xuất hiện thấp;
4. Phạm vi không gian của dị thường khá hạn chế.

Như đề cập trong phần mở đầu, trong công tác tìm kiếm cứu nạn, trực tiếp tìm ra người đang gặp nạn sẽ là lý tưởng, nhưng trong một số trường

hợp, các đồ vật đi kèm như quần áo, chăn mền, va ly, túi xách, lều bạt nơi cắm trại, các mảnh vỡ phương tiện,... có thể cung cấp một số thông tin hữu ích [12, 80, 89, 95]. Các đối tượng như vậy thường có màu sắc hoặc phổ bất kỳ (không biết trước), tương đối khác biệt với môi trường xung quanh và có mật độ thấp nên nó được xem là các dị thường.

1.2.2. Công tác tìm kiếm cứu nạn

Tại Điều 3 Nghị định số 30/2017/NĐ-CP ngày 21 tháng 3 năm 2017 của Chính phủ quy định tổ chức, hoạt động, ứng phó sự cố, thiên tai và tìm kiếm cứu nạn thì: tìm kiếm là sử dụng nhân lực, trang, thiết bị để xác định vị trí của người, phương tiện bị nạn; cứu nạn là các hoạt động cứu người bị nạn thoát khỏi nguy hiểm đang đe dọa đến tính mạng của họ, bao gồm cả tư vấn biện pháp y tế ban đầu, các biện pháp khác để đưa người bị nạn đến vị trí an toàn.

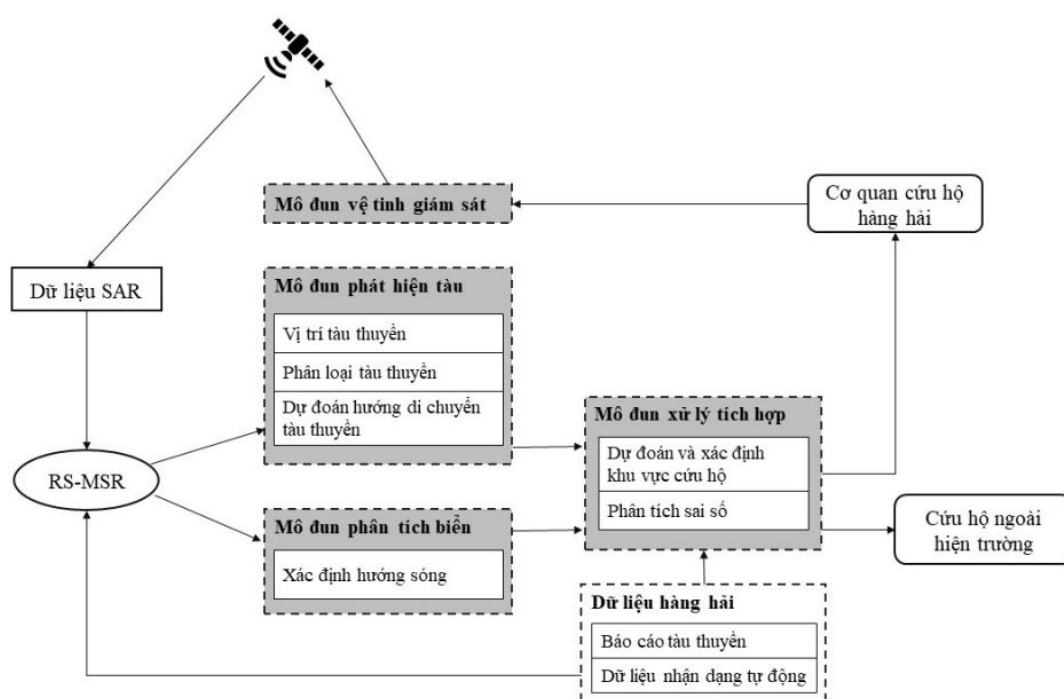
Như vậy, tìm kiếm cứu nạn là toàn bộ các hoạt động của lực lượng cứu nạn, cứu hộ, được tiến hành khi có một sự cố, tai nạn xảy ra nhằm tìm kiếm và giải cứu một người nào đó đang gặp nạn, bị mất tích, bị thương,... ở một khu vực xa xôi hoặc khó tiếp cận, chẳng hạn như ở rừng, núi, sa mạc hoặc trên biển. Với điều kiện địa hình như vậy, cùng với những giới hạn của mắt người, lực lượng cứu nạn, cứu hộ khó có thể tìm thấy mục tiêu nhỏ trong điều kiện ánh sáng bất lợi, ban đêm hoặc điều kiện thời tiết mưa to, gió lớn. Quan sát thời gian dài liên tục cũng gây ra sự mệt mỏi của mắt người, dẫn đến độ nhạy phát hiện kém. Tất cả những yếu tố trên đã làm ảnh hưởng đến kết quả của hoạt động tìm kiếm cứu nạn. Trong khi đó, phát hiện sớm là một yếu tố quan trọng nhất để cứu sống nạn nhân, cho dù lực lượng cứu nạn, cứu hộ có đến hiện trường nhanh, có chuyên môn giỏi và trang thiết bị hiện đại nhưng nếu không kịp thời phát hiện ra vị trí người bị nạn thì khả năng người bị nạn nguy hiểm tới tính mạng vẫn có thể xảy ra. Để cải thiện hiệu quả của hoạt động tìm kiếm cứu nạn trong điều kiện địa hình và thời tiết bất lợi, công nghệ viễn thám là một phương pháp tiềm năng có thể khắc phục các hạn chế trên. Bởi công nghệ viễn thám sở hữu khả năng hoạt động trong bất kì thời gian và điều kiện thời tiết nào, với phạm vi bao trùm rộng và môi trường không thể tiếp cận trực tiếp.

Trên thế giới, đặc biệt là các nước sở hữu công nghệ vệ tinh, việc ứng dụng

công nghệ viễn thám phục vụ công tác tìm kiếm cứu nạn đã được áp dụng và mang lại hiệu quả to lớn. Một số nước phát triển như Mỹ, Nhật, Nga đều có các trung tâm quốc gia hoặc Bộ các tình huống khẩn cấp phản ứng nhanh với các tình huống thảm họa, thiên tai trong đó công nghệ thông tin vệ tinh, thông tin địa lý và thông tin viễn thám là các hợp phần không thể thiếu. Đi kèm với nó là các phân hệ phần mềm xử lý thông tin địa lý, ảnh viễn thám kết xuất các thông tin hữu ích góp phần đẩy nhanh quá trình phản ứng với thảm họa, tìm kiếm cứu nạn và giảm nhẹ thiệt hại. Theo Wallacea và các cộng sự [123], hệ thống Sarsat kết hợp với đối tác Cospas của Nga và NASA đã đạt được thành công lớn khi đã giải cứu thành công hơn 8.000 người trên thế giới trong 16 năm hoạt động. Các cuộc tìm kiếm hiện đang được thực hiện với sự hỗ trợ của thông tin vô tuyến, cảm biến quang học và radar trong các nền tảng trên không. Đội tuần tra hàng không dân dụng CAP (Civil Air Patrol) của Mỹ đã phát triển hệ thống cảm biến ARCHER [111], với mục đích triển khai rộng rãi hệ thống hỗ trợ tự động phát hiện sự bất thường, tín hiệu đặc trưng và các thay đổi liên quan đến máy bay bị rơi. Hệ thống này đã sử dụng ảnh siêu phổ xử lý thời gian thực để hỗ trợ các nhà khai thác nhanh chóng phát hiện các mục tiêu cần quan tâm dưới mặt đất.

Năm 2010, Corbane và các cộng sự [22] đã nghiên cứu và thử nghiệm thành công hệ thống phát hiện các thuyền cỡ nhỏ trên ảnh vệ tinh quang học SPOT5 của Pháp độ phân giải 5m. Năm 2012, Harris và các cộng sự [41] đã thử nghiệm thành công và công bố quy trình phân tích ảnh vệ tinh quang học độ phân giải siêu cao (0,5m) Worldview-2 có nhiều kênh phổ để phát hiện các mảnh vỡ nhỏ trên biển ứng dụng trong tìm kiếm cứu hộ trên biển trong các tình huống khẩn cấp như sóng thần, máy bay rơi. Cũng trong năm này, School [101] đã công bố nghiên cứu phát hiện đối tượng mìn kích thước dưới điểm ảnh dựa trên kỹ thuật phân tích phổ dữ liệu ảnh vệ tinh Worldview-2, thực nghiệm trên biển. Meng và Kerekes [84] đã công bố nghiên cứu ứng dụng ảnh vệ tinh độ phân giải siêu cao WorldView-2 để dò vết đối tượng. Năm 2014, Gao và các cộng sự [33] đã công bố ứng dụng phân tích ảnh vệ tinh Worldview-2 để chiết tách thông tin về đối tượng chuyển động. Grossman làm việc cho cơ quan tình báo Mỹ, trong công trình luận án tiến sĩ “Phương pháp tự động tìm kiếm phổ cho mục đích phát hiện mục tiêu nhỏ” [37] đã đưa ra được khái niệm cụ thể, định lượng được về mục tiêu nhỏ, đề

xuất và cài đặt thử nghiệm thành công các thuật toán tự động tìm kiếm có hướng và thuật toán lọc khớp phổ cải tiến làm nổi lên các mục tiêu nhỏ cần quan tâm trên ảnh vệ tinh thương mại độ phân giải siêu cao có nhiều kênh phổ worldview-2. Nhóm các nhà nghiên cứu của Đại học Boston, Mỹ trong các năm 2012, 2013 và 2015 [12, 80, 95] đã tập trung nghiên cứu các kỹ thuật phát hiện ngoại lai màu trên ảnh UAV ứng dụng cho công tác tìm kiếm cứu nạn, hiệu suất phát hiện các điểm ảnh ngoại lai trên tập dữ liệu mẫu của các thuật toán đều lớn hơn 95%.



Hình 1.10: Mô hình hoạt động của hệ thống tìm kiếm và cứu hộ hàng hải dựa trên ứng dụng công nghệ viễn thám [92]

Một ví dụ về ứng dụng ảnh viễn thám trong tìm kiếm và cứu hộ hàng hải [92], Jing Peng và cộng sự cho thấy, mô hình hoạt động của hệ thống tìm kiếm và cứu hộ hàng hải dựa trên ứng dụng công nghệ viễn thám bao gồm bốn mô-đun (Xem Hình 1.10), mô-đun vệ tinh giám sát, mô-đun phát hiện tàu thuyền, mô-đun phân tích trạng thái biển và mô-đun xử lý tích hợp. Mô-đun phát hiện tàu có ba chức năng là định vị vị trí tàu, xác định/phân loại loại tàu và dự đoán hướng di chuyển của tàu. Phân tích trạng thái biển chủ yếu ước tính hướng độ cao sóng và hướng dòng chảy. Mô-đun xử lý tích

hợp nhận thông tin từ mô-đun phát hiện tàu thuyền và mô-đun phân tích trạng thái biển. Dựa vào kết quả về vị trí của tàu gặp nạn và hướng sóng, kết hợp thời gian nhận dữ liệu, nó ước tính vị trí của tàu gặp nạn và kết hợp tham số vệ tinh, nó có thể hiệu chỉnh kết quả thu được thông qua phát hiện tàu. Kết quả phân tích bằng mô-đun xử lý tích hợp có thể được truyền đến Cơ quan an toàn hàng hải và tàu cứu hộ ngoài hiện trường, để đưa ra các quyết định cho công tác cứu hộ. Phát hiện các điểm ảnh dị thường trên ảnh viễn thám là một trong những phương pháp dùng để định vị vị trí của các đối tượng cần tìm kiếm.

1.2.3. Phát biểu bài toán

Như được nhấn mạnh trong công bố "Detection algorithms for hyperspectral imaging applications" của D. Manolakis và cộng sự [79], trên quan điểm lý thuyết, phát hiện các điểm ảnh dị thường có thể được xem như là vấn đề phân lớp nhị phân, có nghĩa rằng mục đích của nhiệm vụ này là phân loại các điểm ảnh vào lớp "dị thường" và lớp "bình thường". Các điểm ảnh được coi "dị thường" là những điểm ảnh hiếm, nếu những điểm ảnh này nằm kề nhau cũng chỉ tạo một vùng rất nhỏ của khu vực nền rộng lớn trong một cảnh. Do đó, phát hiện dị thường là vấn đề phân lớp nhị phân mà lớp mục tiêu được giả định là không có nhiều, trong khi lớp nền thì phổ biến rộng rãi.

Bài toán phát hiện dị thường màu trên ảnh viễn thám quang học được phát biểu như sau:

- Cho một ảnh đa phổ, siêu phổ có N điểm ảnh và L kênh phổ: $\mathbf{X} = \{\mathbf{x}_i\}_{i=1}^N \in \mathbf{R}^{L \times N}$.

- Xét một điểm ảnh $\mathbf{x}_i \in \mathbf{R}^L, i = 1, 2, \dots, N$, hãy phân đúng \mathbf{x}_i vào một trong hai lớp là lớp "dị thường" hoặc lớp "bình thường".

1.3. Một số tiếp cận về phát hiện dị thường trên ảnh đa phổ, siêu phổ

Mục đích của phát hiện dị thường (AD – anomaly detection) là tìm kiếm các đối tượng trong ảnh, những đối tượng này có sự bất thường đối với nền. Các AD giả định rằng không có kiến thức biết trước về đặc trưng phổ mục tiêu và khảo sát đơn giản bộ dữ liệu để tìm những điểm ảnh có phổ khác biệt đáng kể so với nền. Dựa trên các phương pháp phát hiện các điểm ảnh dị thường trên ảnh đa phổ, siêu phổ khác nhau, năm nhóm phương pháp được

phân chia như sau:

- (1) Phương pháp dựa trên mô hình xác suất thống kê;
- (2) Phương pháp dựa trên phân cụm và phân đoạn;
- (3) Phương pháp dựa trên biến đổi phi tuyến;
- (4) Phương pháp dựa trên biến đổi tuyến tính;
- (5) Phương pháp dựa trên học máy.

1.3.1. Phương pháp dựa trên mô hình xác suất thống kê

1.3.1.1. Thuật toán phát hiện dị thường RX và các biến thể của RX

Thuật toán phát hiện dị thường trên ảnh đa phổ, siêu phổ thực sự bắt nguồn từ nghiên cứu của Reed và Xiaoli [32, 100] công bố vào năm 1990 [96], thường được gọi là phương pháp phát hiện dị thường RX (RXD). RXD được coi là phương pháp phát hiện dị thường chuẩn cho ảnh đa phổ, siêu phổ [52, 108] và nó đã trở thành thuật toán chuẩn được dùng để so sánh với các thuật toán phát hiện điểm ảnh dị thường mới [61, 76, 112, 114]. Reed và Xiaoli đã đề xuất thuật toán tỷ lệ cảnh báo sai không đổi (CFAR) do Chen và Reed xây dựng [19]. Nó được xem như là thuật toán phân lớp nhị phân theo phương pháp Neyman-Pearson [60] sử dụng kiểm nghiệm tỷ lệ khả năng tổng quát (GLRT - generalised likelihood ratio test). Trong đó, giả thuyết đầu tiên, dữ liệu hình ảnh nền được mô hình hóa như một phân phối chuẩn với trung bình bằng 0 và giả thuyết thứ hai, tín hiệu của mục tiêu là sự kết hợp tuyến tính của nền với nhiễu tạp dư. Cốt lõi của quy tắc phát hiện các điểm ảnh dị thường được tính bằng cách tối đa hóa xác suất phát hiện với ràng buộc duy trì xác suất cảnh báo sai không đổi ở một giá trị mong muốn. Để đảm bảo dữ liệu nền tuân theo phân phối chuẩn, RXD được tính toán cục bộ thông qua một cửa sổ lấy PUT (pixel under test) làm trung tâm gọi là LRXD (Local RXD). Do LRXD dựa trên ước lượng cục bộ của các tham số nền nên nó phù hợp cho xử lý hình ảnh trong quá trình thu nhận, có thể được sử dụng cho các ứng dụng thời gian thực. Trong trường hợp này, một cửa sổ không trượt trên ảnh, mà một bộ đệm có thể được dùng để lưu trữ một vài dòng tiếp theo, được dùng cho các quá trình ước lượng ma trận khử và ma trận hiệp phương sai. Một phiên bản thời gian thực được phát triển bởi Stellman và các cộng sự [110] và được thử nghiệm trên máy ảnh độ phân giải cao trong thử nghiệm bay Dark HORSE 1. Trong hệ thống này, các tham số nền được

tính toán tại mỗi điểm ảnh bằng cách sử dụng đệ quy đáp ứng xung vô hạn trong lúc thu nhận từng dòng dữ liệu ảnh.

Tuy nhiên, RXD cũng gặp phải nhiều vấn đề: thứ nhất, mặc dù đã tính toán RXD trong một cửa sổ cục bộ nhưng dữ liệu nền trong cửa sổ cục bộ này thực tế rất hiếm khi phân bố của dữ liệu là phân bố chuẩn bởi các cảnh ảnh rất đa dạng và chứa nhiều lớp đối tượng khác nhau [9, 16, 40]. Thứ hai, RXD tính toán cục bộ trong một cửa sổ, do đó, ma trận hiệp phương sai được ước lượng từ một số lượng nhỏ các mẫu dữ liệu có số chiều lớn dẫn đến nghịch đảo của chúng thường không ổn định (cần chuẩn tắc hóa để giảm các lỗi thống kê) [90]. Thứ ba, RXD có tỷ lệ dương tính giả cao, nó phát hiện các điểm ảnh dị thường trong vùng cục bộ quanh điểm ảnh đang xét nhưng không phải trong toàn cục (một cái cây đơn lẻ trong đồng cỏ đồng nhất cục bộ được coi là dị thường cục bộ ngay cả khi toàn bộ ảnh có cả một khu rừng)[9, 40, 109, 82].

Để giảm tỷ lệ dương tính giả, Chang và Chiang [16] đã trình bày bốn biến thể của RXD. Hai biến thể đầu tiên là NRXD (Normalized RXD) và MRXD (Modified RXD), hai biến thể này hoạt động tương tự như bộ lọc thích nghi. Cách tiếp cận thứ ba là RXD-UTD (RXD - Uniform Target Detector), sự thích ứng của quá trình phát hiện mục tiêu đồng nhất [42] cho RXD. Biến thể thứ tư là CRXD (Causal RXD), CRXD tập trung vào việc xử lý thời gian thực. Trong CRXD, ma trận hiệp phương sai trong RXD được thay thế bằng ma trận tương quan cho bốn biến thể và phép trừ giá trị trung bình của dữ liệu được loại bỏ.

Một phương pháp tiếp cận khác sử dụng ngưỡng tự động, dựa trên việc sử dụng biểu đồ của hình ảnh để xác định xác suất loại bỏ. Cách tiếp cận này đề xuất phân loại các điểm ảnh dị thường được phát hiện bằng biểu đồ của biện pháp lọc tương quan phù hợp và được xử lý theo thời gian thực. Bộ phân loại hoạt động theo nguyên lý phương sai tối thiểu cưỡng bức tuyến tính [18]. Du và Zhang [25] đã đề xuất một biến thể của RXD bằng cách lựa chọn tùy ý các điểm ảnh đại diện nền cho mỗi khối dữ liệu thay vì sử dụng toàn bộ hình ảnh. Khoảng cách Mahalanobis được tính giữa PUT và điểm ảnh đại diện của mỗi khối. Cách tiếp cận này ít nhạy cảm hơn với nhiễu và thể hiện sự phân tách tốt hơn giữa dị thường và nền.

Guo và các cộng sự [39] đã đề xuất hai biến thể của RXD nhằm cải thiện đặc tính nền để có kết quả phát hiện tốt hơn. Biến thể đầu tiên là WRXD (Weighted RXD), WRXD thực hiện giảm trọng số của các điểm ảnh bất thường hoặc tín hiệu nhiễu và tăng trọng số của các điểm ảnh nền. Biến thể thứ hai là LF-RXD (Linear Filter - RXD), việc tính toán RXD dựa trên bộ lọc tuyến tính sử dụng xác suất của từng điểm ảnh nền để lọc các trường hợp dị thường hoặc nhiễu. Molero và các cộng sự [88] đã đề xuất triển khai song song cho GRXD (Global RXD) và LRXD (Local RXD) dựa trên các nền tảng đa lõi. Cách tiếp cận này trình bày tối ưu hóa phương pháp của Chang và Chiang [16] bằng cách sử dụng đại số tuyến tính, thực hiện phân tách hệ phương trình ma trận để thoát khỏi nghịch đảo của ma trận tương quan để chỉ tính một nửa. Đối với LRXD, mối quan hệ lặp lại giữa các ma trận liên quan đến các điểm ảnh lân cận được khai thác tốt. Liu và Chang [69], Ren và các cộng sự [98] đã đề xuất cách tiếp cận dựa trên nhiều cửa sổ để phát hiện tốt hơn các dị thường có kích thước và hình dạng khác nhau. Cách tiếp cận này phù hợp để thực hiện trên các bộ phát hiện dị thường: RXD; phương pháp tiếp cận dựa trên cửa sổ không gian lồng nhau để phát hiện mục tiêu [69]; biến đổi phân tách eigen (được thiết kế để tối đa hóa việc tách dữ liệu hai lớp trong không gian con có số chiều dữ liệu thấp [58]) dựa trên cửa sổ kép [58].

Imani [51] cho rằng, các điểm ảnh liên kề điểm ảnh đang xét thường có các đặc trưng phổ, hình dạng và kết cấu tương đối giống nhau. Trong khi đó hầu hết các nghiên cứu bỏ qua các đặc điểm không gian có giá trị này trong việc tính toán đưa ra quyết định. Để giải quyết vấn đề này, Imani đề xuất hợp nhất thông tin về quang phổ và mối liên hệ không gian của các điểm ảnh lân cận bằng cách áp dụng bộ lọc Gabor 2D và 3D. Sau khi ảnh được lọc qua Gabor thuật toán RXD sẽ được áp dụng để phát hiện các điểm ảnh dị thường. Wu và các cộng sự [125] đã đề xuất phương pháp SSAC (spectral and spatial approximate computing) để phát hiện nhanh các điểm ảnh dị thường trên ảnh siêu phổ thông qua ba giai đoạn. Giai đoạn đầu tiên là loại bỏ một số kênh phổ để giảm số chiều của dữ liệu thông qua việc phân tích mối tương quan của dữ liệu theo chiều không gian và chiều phổ. Giai đoạn thứ hai sử dụng thuật toán k-mean để phân cụm dữ liệu, và giai đoạn cuối cùng sử dụng

RXD để phát hiện các điểm ảnh dị thường.

1.3.1.2. Phát hiện dị thường theo phương pháp thống kê bậc cao

Ren và các cộng sự [97] đã chỉ ra rằng, bằng các thí nghiệm sử dụng số liệu thống kê bậc hai để phát hiện mục tiêu xảy ra với xác suất thấp và phân bố nhỏ sẽ kém hiệu quả hơn so với sử dụng thống kê bậc cao. Cách tiếp cận của Ren và các cộng sự dựa trên quy trình làm trắng hóa để trừ các tín hiệu nền sau đó sử dụng phép chiếu không gian con trực giao. Một phát triển của phương pháp này được trình bày trong [17], trong đó chiều ảo được sử dụng để phát hiện tự động số lượng tín hiệu khác biệt phổ có trong dữ liệu. Đối với phương pháp này, phép chiếu được thực hiện bằng phân tích thành phần độc lập nhanh [49], với tiêu chí phủ định để xấp xỉ thông tin lẫn nhau do đó tăng tốc độ phát hiện.

Gu và các cộng sự [38] đã đề xuất phương pháp phát hiện dị thường bằng cách phân tích thành phần chính của hạt nhân chọn lọc để khai thác tốt hơn mối tương quan của dải phổ bậc cao, và sau đó để tính toán ALS (average local singularity) dựa trên thống kê thứ tự cao, thành phần có ALS tối đa được coi là dị thường.

1.3.1.3. Ước lượng mật độ xác suất phi tham số nền

Khác với phương pháp GLRT là đi tìm các tham số thỏa mãn yêu cầu của các hàm mật độ xác suất (PDF - probability density function), phương pháp ước lượng mật độ phi tham số không cần giả sử bất kỳ dạng hàm số nào cho PDF có điều kiện. Hàm ước lượng phi tham số được cài đặt nhiều nhất cho PDF là hàm ước lượng Parzen (PE) [91]. Một PDF được ước lượng bằng cách đặt một hàm kernel cho mỗi mẫu huấn luyện hoặc dữ liệu.

Năm 2011, Veracini và các cộng sự [120, 121] đề xuất sử dụng phương pháp Parzen Window (PW) để ước tính PDF nền. Sau khi PDF nền được xấp xỉ thông qua PW, nó được dùng làm đầu vào để phát hiện các dấu hiệu dị thường trên ảnh dựa trên phương pháp kiểm nghiệm tỷ lệ khả năng. Năm 2012, Bolukbasi và Tran [12] đã xây dựng kiểm nghiệm giả thuyết nhị phân cho phát hiện dị thường và sử dụng thuật toán KNN để tìm k láng giềng gần nhất để tính hàm mật độ xác suất phi tham số cho điểm ảnh đang xét. Năm 2014, Matteoli và các cộng sự [83] đã đưa ra chiến lược để quyết định một điểm ảnh có phải là dị thường hay là nền dựa trên định lý Neyman-Pearson

sử dụng các hàm PDF. Trong đó các tác giả đã kiểm nghiệm trên ba hạt nhân PDF đó là hạt nhân Gaussian cố định bằng thông, hạt nhân Gaussian không cố định bằng thông và tìm kiếm k láng giềng gần nhất để ước lượng hàm mật độ giống như trong [12]. Năm 2017, Zhao và các cộng sự [135] đã xây dựng kỹ thuật phát hiện dị thường là sự kết hợp giữa ước lượng mật độ phi tham số và biểu diễn mối quan hệ tương quan để nâng cao độ chính xác phát hiện dị thường.

Các phương pháp phát hiện dị thường dựa trên ước lượng mật độ xác suất phi tham số không cần phải biết trước về phân bố của dữ liệu cũng không cần phải tính toán ma trận hiệp phương sai như RXD. Phương pháp này được tính toán toàn cục trên cả ảnh nên đã giảm được tỷ lệ dương tính giả so với RXD. Kết quả đã chứng minh rằng phương pháp này vượt qua RX truyền thống và KRX về độ chính xác phát hiện dị thường, nhưng chi phí tính toán của nó rất lớn (có độ phức tạp tính toán là $O(N^2)$, vẫn còn xa so với xử lý thời gian thực.

1.3.2. Phương pháp dựa trên phân cụm và phân đoạn

1.3.2.1. Bộ dò dị thường dựa trên cụm

Năm 2005, Carlotto [13] đã đề xuất phương pháp phát hiện dị thường dựa trên cụm (CBAD - cluster based anomaly detector), được coi là bộ phát hiện dị thường chuẩn cho các phương pháp tiếp cận dựa trên phân cụm [32, 100]. Cách tiếp cận này là phân cụm các điểm ảnh để thu được bản đồ chủ đề theo phương pháp lượng tử hóa biểu đồ của các thành phần chính hình ảnh. Trong mỗi cụm, dữ liệu được cho là có phân phối theo mô hình hỗn hợp Gaussian (GMM - Gaussian-mixture model). Sau đó, sử dụng phương pháp đo khoảng cách Mahalanobis để kiểm tra khoảng cách từ PUT đến tâm của mỗi cụm. Điểm ảnh vượt quá ngưỡng được coi là dị thường. Cách tiếp cận này có khả năng phát hiện các đối tượng có kích thước khác nhau và tính toán hiệu quả hơn so với cách tiếp cận dựa trên cụm khác. Hiệu quả của phương pháp này đã được Hyta và các cộng sự [48] chứng minh khi nghiên cứu so sánh các mô hình hỗn hợp Gaussian.

Một phiên bản CBAD có hiệu quả về thời gian được phát triển trong nghiên cứu [28, 29], trong đó kỹ thuật phân cụm dựa trên bản đồ tự tổ chức (SOM - Self Organizing Maps) [27, 93] được sử dụng. Trong bước đầu, sử

dụng một tỷ lệ nhỏ ngẫu nhiên các điểm ảnh để phân cụm, do đó làm tăng tốc độ tính toán; sau đó phát hiện dị thường được thực hiện theo từng dòng, cùng với cập nhật động SOM và do đó có các tham số nền. Phân cụm SOM được chứng minh là chính xác, linh hoạt và hiệu quả về thời gian hơn phương pháp phân cụm k-means [28]. Sử dụng kỹ thuật này, thời gian thực hiện gần với RXD. Ngoài ra, hệ thống DARK HORSE 1 đã được đề cập [110] cung cấp một AD thời gian thực dựa trên CBAD. Xử lý thời gian thực thu được bằng cách hình thành các cụm ban đầu theo bước đệm đầu tiên, và sau đó bằng cách cập nhật đệ quy các tham số các cụm trong quá trình thu nhận ảnh.

Phương pháp Bayesian có thể tiếp cận dùng cho phân lớp không giám sát [7, 43] bằng cách áp dụng quy tắc Posteriori tối đa để gán nhãn. Hazel [43] không sử dụng mô hình nền khác nhau mà sử dụng mô hình trường ngẫu nhiên Gauss-Markov (GMRF) đa biến cho mỗi cụm, để tính đến sự tương quan không gian ở mức điểm ảnh. Tác giả tuyên bố rằng ngay cả mô hình kết cấu đẳng hướng bậc nhất được giả định đơn giản nhất cũng đưa ra cải thiện trong phân đoạn và phát hiện khi so sánh với phân cụm phổ đơn thuần.

1.3.2.2. Bộ phát hiện dị thường dựa trên phân đoạn không gian

GMRF được sử dụng để phân đoạn ảnh trước khi phát hiện dị thường, do đó Schweizer và Moura [103] cho rằng sự lộn xộn nền có thể được xem là GMRF. Mô hình này cho phép cắt hình ảnh thành các vùng Gauss-Markov có màu đại diện cho các vật liệu khác nhau. Ba phương pháp được đề xuất để ước tính các tham số của GMRF: khả năng tối đa, bình phương nhỏ nhất và khả năng tối đa gần đúng. So với RXD, cách tiếp cận này thoát khỏi sự đảo ngược của ma trận hiệp phương sai bằng cách tham số hóa trực tiếp nghịch đảo của nó dẫn đến chi phí tính toán thấp hơn.

Li và các cộng sự [64] đã đề xuất phân vùng modulo dựa trên khối bằng cách sử dụng phân tích thành phần chính của phép chiếu nền để duy trì sự kết hợp không gian cục bộ. Kết quả phân đoạn không gian được kết hợp với phân loại điểm ảnh với quy tắc bỏ phiếu đa số. Kết quả đã chứng minh rằng việc sử dụng trình phát hiện RXD hoạt động với phân chia modulo dựa trên khối theo sau là không gian con được chiếu ngẫu nhiên vượt trội hơn so với trình phát hiện hoạt động với phân chia modulo dựa trên điểm ảnh.

Yuan và các cộng sự [131] đã đề xuất một kỹ thuật phát hiện dị thường

mới dựa trên sự phân kỳ thừa thớt cục bộ. Phương pháp đề xuất bắt đầu bằng cách xây dựng các từ điển phổ và không gian địa phương bằng cách sử dụng một cửa sổ kép cho phép trích xuất các hệ số thừa thớt của từng điểm ảnh. Bước tiếp theo là về tính toán của bản đồ phân kỳ thừa thớt cục bộ ở mỗi dải quang phổ riêng biệt và sau đó một phân đoạn chung của các bản đồ phân kỳ thừa thớt cục bộ trên các dải khác nhau được thực hiện để phát hiện dị thường dẫn đến hiệu suất tốt hơn.

1.3.3. Phương pháp dựa trên biến đổi phi tuyến

1.3.3.1. Thuật toán Kernel - RX

Vấn đề chính với các kỹ thuật phát hiện dị thường dựa trên thống kê là giả định về tính tuyến tính của nền. Trong thực tế, sự phân bố của hàm mật độ xác suất của hình ảnh đa phổ, siêu phổ là rất xa so với phân phối chuẩn do cảnh ảnh rất đa dạng và chứa nhiều lớp đối tượng khác nhau [9, 16, 40]. Do đó, các kỹ thuật phát hiện dị thường phi tuyến được đề xuất sử dụng các chiến lược hạt nhân.

Một ý tưởng được Kwon và Nasrabadi [59] phát triển được gọi phương pháp kernel và thủ thuật kernel [104]. Triết lý cơ bản là thể hiện một mô hình phức tạp bằng một mô hình đơn giản hơn trong một không gian đặc trưng có số chiều lớn, thu được thông qua ánh xạ phi tuyến tính $\Phi(\cdot)$. Ánh xạ dữ liệu trong không gian đặc trưng N chiều với $N \gg L$ (có thể cho $N \rightarrow \infty$), cho phép các lớp nền và mục tiêu được rời rạc. Sau đó, do tính không đồng nhất nền làm giảm hiệu quả của mô hình phân phối chuẩn (NM - Normal Model) trong không gian ban đầu, nên ý tưởng chính là áp dụng NM trong không gian đặc trưng chiều cao hơn, trong đó được kỳ vọng là phù hợp hơn cho việc đặc trưng hóa tính không đồng nhất nền. Việc cài đặt trực tiếp của thuật toán RX trong không gian đặc trưng là không khả thi do có số chiều cao. Tuy nhiên, các tác giả chứng minh rằng thuật toán có thể được “kernelized”, tức là được biểu diễn dưới dạng tích vô hướng trong không gian đặc trưng, sau đó có thể được tính toán trong miền dữ liệu thông qua thủ thuật kernel. Thủ thuật này [104] cho phép việc tính toán tích vô hướng trong không gian đặc trưng chiều cao bằng các hàm đơn giản, gọi là các hàm kernel, được xác định trên các cặp dữ liệu đầu vào. Thực tế, thủ thuật này cho phép xây dựng các biến thể phi tuyến tính của bất kỳ thuật toán nào có thể được biểu diễn

dưới dạng tích vô hướng. Phiên bản kernelized của thuật toán trong không gian đặc trưng được biểu diễn dưới dạng các ma trận Gram, có các phần tử là các hàm kernel được tính từ các cặp điểm ảnh đầu vào. Trong [59], hàm kernel Gaussian RBF (radial basis function) được sử dụng.

1.3.3.2. Ước lượng vùng hỗ trợ nền

Một phương pháp phát hiện dị thường dựa trên cách tiếp cận vector hỗ trợ được đưa ra và áp dụng cho ảnh siêu phổ [9]. Phương pháp này dựa trên ý tưởng [105]: thay vì ước tính giá trị của PDF, phương pháp này sẽ ước lượng hình dạng và kích thước của “vùng hỗ trợ” nền, tức là khu vực trong không gian đầu vào trong đó hầu hết dữ liệu nền đều nằm trong đó. Ý tưởng này đã được áp dụng rộng rãi để thực hiện phát hiện dị thường trong một số lĩnh vực đã được trình bày trong [105, 116, 117].

Trong [9], phương pháp SVDD [116, 117] được sử dụng. SVDD là một trình phân lớp cho phép ước lượng vùng hỗ trợ của dữ liệu huấn luyện bằng cách tìm ra “siêu cầu kín” tối thiểu chứa dữ liệu huấn luyện. Ở đây, dữ liệu huấn luyện là các điểm ảnh nền được bắt bởi các cửa sổ trượt kép. Các dị thường có thể được xác định bằng cách kiểm thử xem PUT có nằm trong hoặc ngoài của siêu cầu đó. Thật vậy, khoảng cách bình phương giữa PUT và tâm của siêu cầu được dùng làm thống kê kiểm tra ngưỡng. Trên thực tế, vấn đề tối ưu của việc tìm siêu cầu kín tối thiểu được thực hiện trong không gian đặc trưng chiều cao hơn, thu được qua ánh xạ phi tuyến tính $\Phi(\cdot)$, do đó xác định được ranh giới quyết định rất chặt chẽ trong không gian đầu vào ban đầu. Thuật toán có thể dễ dàng được diễn giải dạng tích vô hướng trong không gian đặc trưng và do đó có thể áp dụng thủ thuật kernel.

Mặc dù thuật toán Kernel-RX (KRX) có độ chính xác phát hiện các điểm ảnh dị thường tốt hơn thuật toán RX, nhưng nó vẫn còn đó những điểm hạn chế giống như LRX. Do đó, Zhao và các cộng sự [136] đề xuất phương pháp làm cho dữ liệu tuân theo phân phối Gaussian và một phương pháp không sử dụng ma trận nghịch đảo. Ngoài ra, KRX rất ít được áp dụng rộng rãi trong thực tế, bởi vì nó tiêu tốn quá nhiều về không gian bộ nhớ và tiêu tốn quá nhiều thời gian để tính toán. Vì vậy, Hidalgo và các cộng sự [45] đã đề xuất phương pháp cải thiện hiệu quả không gian bộ nhớ và thời gian tính toán của KRX; bằng cách sử dụng các kỹ thuật hiệu quả dựa trên các tính chất

của Fourier ngẫu nhiên.

1.3.4. Phương pháp dựa trên biến đổi tuyến tính

1.3.4.1. Phân tích thành phần độc lập và phép chiếu

Phân tích thành phần độc lập [49] và theo đuổi phép chiếu [46] tuân theo cách tiếp cận chiếu dữ liệu siêu phổ chiều cao vào không gian dữ liệu chiều thấp, trong khi vẫn giữ lại những thông tin quan tâm trong các ứng dụng phát hiện dị thường và mục tiêu, thường được xác định là sự hiện diện của các điểm ảnh dị thường đi lệch khỏi phân bố nền. Cách tiếp cận này được áp dụng rộng rãi trong phát hiện dị thường và mục tiêu [18, 26, 50]. Trong các trường hợp như vậy, các chỉ mục phép chiếu được sử dụng để tìm ra các phép chiếu phù hợp nhất thường là các mô-men có thứ tự cao. Các kỹ thuật này cho phép phát hiện và phân biệt các lớp mục tiêu khác nhau, đồng thời chiếu dữ liệu theo các hướng được đưa ra bằng cách tối ưu hóa các chỉ mục được chọn. Một số thuật toán như FastICA [46] và Tối đa hóa Kurtosis cưỡng bức [26] rất hiệu quả về mặt tính toán, cho nên chúng có thể phát hiện mục tiêu quan trọng với ít phép chiếu. Tuy nhiên, các vấn đề cần chú ý là thiết lập các điều kiện ban đầu và các điều kiện dừng. Hơn nữa, công việc tiền xử lý ảnh là rất cần thiết để kiểm tra các phép chiếu.

1.3.4.2. Biểu diễn thưa và biểu diễn hạng thấp

Trong thời gian gần đây, các phương pháp dựa trên biểu diễn đã thu hút được nhiều sự chú ý. Các phương pháp này giả định rằng tín hiệu phổ có thể được biểu diễn bằng cách sử dụng từ điển [130, 134, 137]. Với các ràng buộc khác nhau về các hệ số biểu diễn có thể thu được các phương pháp phát hiện dị thường khác nhau. Phương pháp phát hiện dị thường dựa trên biểu diễn thưa được giới thiệu trong phát hiện mục tiêu siêu phổ có giám sát [20, 21] và giả định rằng mỗi mẫu có thể được đại diện bởi một vài phần tử trong từ điển. Các kỹ thuật phát hiện dị thường dựa trên biểu diễn hợp tác được Li và Du [63], Ling và các cộng sự [68] đề xuất. Nó dựa trên khái niệm rằng mỗi điểm ảnh trong nền có thể được biểu diễn xấp xỉ bởi các vùng lân cận không gian của nó, trong khi dị thường không thể. Biểu diễn là sự kết hợp tuyến tính của các điểm ảnh lân cận và sự cộng tác giữa các điểm ảnh này được củng cố bằng cách giảm thiểu chuẩn l_2 của vector trọng số biểu diễn. Tuy nhiên, không có phương pháp nào trong số các phương pháp này xem

xét mối tương quan của tất cả các điểm ảnh trong ảnh siêu phổ; do đó, thông tin toàn cầu không được tính trong các phương pháp này.

Một phương pháp phát hiện dị thường mới dựa trên biểu diễn hạng thấp và biểu diễn thưa được Xu và các cộng sự [127] đề xuất. Trái ngược với các phương pháp dựa trên biểu diễn khác, phương pháp được đề xuất được xây dựng dựa trên sự tách biệt giữa phần bất thường và phần nền, thông tin cơ sở được chứa trong biểu diễn hạng thấp nhất của các điểm ảnh siêu phổ. Biểu diễn hạng thấp có thể được sử dụng để tìm biểu diễn hạng thấp nhất của tất cả các điểm ảnh cùng chung tính chất. Sau đó, phần bất thường có thể thu được bằng phần dư của ảnh gốc và phần nền được phục hồi. Li và các cộng sự [61] đã có sự kết hợp giữa LRX và biểu diễn thưa để phát hiện các điểm ảnh dị thường trên ảnh siêu phổ. Đầu tiên, LRX được sử dụng để thu được các điểm ảnh nền, những điểm ảnh nền này được sử dụng để xây dựng từ điển. Sau khi có được từ điển, nó được dùng để biểu diễn các điểm ảnh, độ lệch giữa vector điểm ảnh gốc và vector điểm ảnh được tái tạo sau biểu diễn thưa được dùng để phát hiện các điểm ảnh dị thường.

Một biểu diễn thưa có trọng số thích ứng và AD dựa trên ước lượng nền được trình bày trong [137], nó sử dụng phương pháp chiết xuất endmember (người ta thấy rằng trong diện tích của 1 điểm ảnh thường có nhiều yếu tố khác nhau, ví dụ: 30% là nước, 30% là lúa, 40% là đất trồng. Mỗi yếu tố được gọi là một endmember [1]) để mô tả các phổ; biểu diễn thưa có trọng số có thể áp dụng được cả toàn cục và cục bộ. Một kỹ thuật học đặc trưng phân biệt với nhiều từ điển đã được giới thiệu trong công trình của Ma và các cộng sự tại [75].

Vafadar và Ghassemian [119] đã đề xuất phương pháp CRBORAD (collaborative-representation-based with outlier removal anomaly detector) và CRBORAD-Gaussian-kernel sử dụng cả thông tin phổ và không gian để phát hiện các điểm ảnh dị thường. Các tác giả cho rằng, mỗi một điểm ảnh nếu là điểm ảnh nền thì nó có thể được xấp xỉ bởi các điểm ảnh lân cận trong vùng không gian xung quanh nó, ngược lại các điểm ảnh dị thường không thể được tạo bởi các điểm ảnh liền kề. Chính vì vậy, phải đi tìm một vector để biểu diễn điểm ảnh đang xét. Việc tìm vector này bằng cách nhân ma trận các điểm ảnh xung quanh điểm ảnh đang xét sau khi trừ đi các điểm ảnh dị thường trong một cửa sổ kép với một vector trọng số. Hiệu của vector điểm

ảnh đang xét với vector vừa tìm được chính là giá trị dị thường của điểm ảnh đang xét, nếu nó lớn hơn một ngưỡng nào đó nó là điểm ảnh dị thường, ngược lại nó là nền.

Dựa trên ý tưởng của Vafadar và Ghassemian, Hou và các cộng sự [132] đã thay đổi cách tính toán để đưa ra một vector trọng số mới, kết quả kiểm nghiệm đã vượt qua CRBORAD về hiệu suất phát hiện các điểm ảnh dị thường. Xiong và các cộng sự [126] cũng cho rằng, các điểm ảnh nền có thể được biểu diễn thưa bởi các vùng lân cận của chúng, trong khi các điểm ảnh dị thường thì không thể. Thông thường để tìm nghiệm thưa, người ta thường sử dụng thuật toán OMP (orthogonal matching pursuit), tuy nhiên, các điểm ảnh siêu phổ, thông thường tín hiệu phổ được trộn phi tuyến nên OMP không có hiệu quả tốt. Do đó Xiong và các cộng sự [126] đề xuất sử dụng phương pháp KOMP (kernel orthogonal matching pursuit) để giải bài toán tìm nghiệm thưa trên. Nguyên tắc cơ bản là sử dụng hàm phi tuyến $\Phi(\cdot)$ để chiếu dữ liệu vào không gian đặc trưng có số chiều cao hơn, làm cho dữ liệu có thể phân tách và cho kết quả thu được bằng thuật toán KOMP phù hợp hơn cho dữ liệu đa phổ.

Ma và các cộng sự [75] đề xuất một phương pháp phát hiện dị thường từ việc trích chọn đặc trưng. Các tác giả cho rằng, vector phổ của các điểm ảnh có thể được biểu diễn thưa bởi một vài kênh phổ nổi bật, sử dụng phương pháp biểu diễn thưa thông qua việc xây dựng đa từ điển sẽ được một vector phổ đặc trưng mới. Với vector phổ đặc trưng mới này sẽ gia tăng sự chênh lệch giữa nền và các dị thường từ đó phân tách giữa nền và dị thường dễ dàng và chính xác hơn.

Tan và các cộng sự [115] đã đề xuất hai kỹ thuật để phát hiện các điểm ảnh dị thường. Thứ nhất, đó là đề xuất LSUNRSORAD (local summation unsupervised nearest regularized subspace with an outlier removal anomaly detector), các tác giả cho rằng sự hiện diện của các điểm ảnh dị thường sẽ ảnh hưởng đến độ chính xác phát hiện dị thường của thuật toán. Để tăng độ chính xác phát hiện, chiến lược loại bỏ các điểm ảnh dị thường được đưa ra (các điểm ảnh có giá trị cường độ lớn hơn ngưỡng tối đa hoặc nhỏ hơn ngưỡng tối thiểu được coi là dị thường) trước khi sử dụng thuật toán UNRS [62] để phát hiện các điểm ảnh dị thường. Thứ hai, đề xuất LSAD-CR-IDW (local summation anomaly detection based on collaborative representation

and inverse distance weight), để sử dụng đầy đủ thông tin thống kê nền cục bộ, Tan và các cộng sự [115] đề xuất sử dụng cửa sổ kép tổng hợp cục bộ. Đối với mỗi điểm ảnh thử nghiệm, chín cửa sổ được tạo ra bằng cách sử dụng chiến lược trượt một cửa sổ xung quanh điểm đang xét. Chín cửa sổ cục bộ này được sử dụng để tính toán khoảng cách Mahalanobis của điểm ảnh thử nghiệm, từ đó phân tách được điểm ảnh nền và điểm ảnh dị thường.

1.3.5. Phương pháp dựa trên học máy

Trong thời gian gần đây, phương pháp học sâu đã thu hút sự chú ý của các nhà khoa học và áp dụng nó ngày càng nhiều trong phân tích ảnh siêu phổ. Li và các cộng sự [65] đã đề xuất sử dụng mạng học sâu CNN (convolutional neural network) để phát hiện các điểm ảnh dị thường gọi là CNND (CNN-base detector). Để thực hiện phương pháp phát hiện dị thường này, các tác giả sử dụng một cảnh ảnh tham chiếu (từ cùng một cảm biến) với các mẫu đã được gán nhãn. Sử dụng các cặp điểm ảnh từ dữ liệu tham chiếu đo mức độ tương tự giữa điểm ảnh đang xét và môi trường xung quanh bằng CNN. Trung bình giá trị đầu ra của CNN làm đầu ra để phát hiện các điểm ảnh dị thường. Ma và các cộng sự [76] đề xuất kỹ thuật DBN (deep belief network) để phát hiện các điểm dị thường. Bộ mã hóa tự động dựa trên DBN được sử dụng để trích rút các đặc trưng và các lỗi tái cấu trúc của ảnh siêu phổ mà không cần giả định phân phối của dữ liệu. Khi đó, các điểm ảnh sẽ được biểu diễn dưới dạng một loạt các mã ngắn. Để phân loại các điểm ảnh dị thường, sử dụng phương pháp tính khoảng cách giữa điểm ảnh đang xét và các điểm ảnh lân cận của nó.

Yan và các cộng sự [128] đề xuất phương pháp 2SHyperNet (two-stage deep learning – base hyperspectral neural network) bằng cách sử dụng một mạng học sâu hai giai đoạn trên ảnh siêu phổ để phát hiện người trên biển. Giai đoạn đầu thông tin về phổ được sử dụng để phân loại, giai đoạn hai là sự kết hợp thông tin về không gian để loại bỏ các vùng không chắc chắn dẫn đến sự cải thiện về hiệu suất phát hiện các điểm ảnh dị thường. Zhang và Cheng [133] đề xuất phương pháp SAEASM (stacked autoencoders-based adaptive subspace model) sử dụng ba cửa sổ lấy điểm ảnh đang xét làm trung tâm, những điểm ảnh nằm trong cửa sổ trong cùng (gần với điểm ảnh đang xét nhất) gọi là những điểm ảnh hàng xóm cục bộ, những điểm ảnh nằm trong

cửa sổ giữa gọi là những điểm ảnh nền cục bộ và những điểm ảnh nằm trong cửa sổ ngoài cùng được gọi là những điểm ảnh từ điển cục bộ. SAE (stacked autoencoders) thực hiện nhiệm vụ tìm các đặc trưng khác biệt giữa điểm ảnh đang xét và các điểm ảnh từ điển cục bộ. Sau khi thu được các đặc trưng, sử dụng 2-norm dựa trên mô hình không gian con thích ứng để phát hiện các điểm ảnh dị thường.

Jiang và các cộng sự [54] đã đề xuất một phương pháp học bán giám sát để phát hiện dị thường trên ảnh siêu phổ dựa trên mạng sáng tạo đối nghịch (GAN- Generative Adversarial Network). Wang và các cộng sự [124] đề xuất một mạng Auto-AD phát hiện dị thường trên ảnh siêu phổ, trong kỹ thuật này, dữ liệu nền được mạng Auto-AD tái tạo lại và các điểm dị thường xuất hiện dưới dạng lỗi trong quá trình tái tạo lại.

Bảng 1.3: Đặc điểm của một số phương pháp phát hiện dị thường

Cách tiếp cận	Dữ liệu	Kích thước mục tiêu	Thời gian thực thi/Độ phức tạp tính toán	Đóng góp của thuật toán và những điểm hạn chế
RXD và các biến thể của RXD	Quang phổ	Có kích thước lớn	Ngoại trừ LRXD và CRXD thực hiện theo thời gian thực, các thuật toán còn lại có độ phức tạp tính toán là $O(L^3)^*$	<ul style="list-style-type: none"> - Không cần biết trước phổ mục tiêu. - Một số điểm còn tồn tại: Có tỷ lệ dương tính giả cao nếu dữ liệu không tuân theo phân bố chuẩn, ảnh bị nhiễu; việc tính toán nghịch đảo ma trận hiệp phương sai với số chiều dữ liệu cao sẽ rất phức tạp.
KRXD	Quang phổ	Những dị thường có kích thước lớn	$O(N^3)^{**}$	<ul style="list-style-type: none"> - Không cần biết trước phổ mục tiêu. - Có độ chính xác phát hiện các điểm ảnh dị thường tốt hơn RXD. - Một số điểm còn tồn tại: độ phức tạp tính toán cao, có thời gian tính toán lớn nên khó ứng dụng trong thực tiễn; việc lựa chọn tham số tối ưu cho hàm nhân khó khăn.

Bộ phát hiện dị thường theo phương pháp thống kê bậc cao	Quang phổ	Từ một phần điển ảnh đến vài điểm ảnh	Theo thời gian thực	<ul style="list-style-type: none"> - Không cần biết trước phổ mục tiêu. - Phát hiện mục tiêu xảy ra với xác suất thấp và phân bố nhỏ. - Có tốc độ và độ chính xác phát hiện dị thường cao hơn RXD.
Ước lượng mật độ xác suất phi tham số	Quang phổ	Từ một phần điển ảnh đến vài điểm ảnh	$O(N^2)$	<ul style="list-style-type: none"> - Không cần biết trước phổ mục tiêu. - Không cần giả định phân bố của dữ liệu, không cần tính toán ma trận hiệp phương sai và nghịch đảo của nó. - Độ chính xác phát hiện dị thường cao hơn RXD. - Một số điểm còn tồn tại: thời gian tính toán cao nên khó triển khai trong các ứng dụng thực tế. Việc xác định băng thông cho hàm nhân để hiệu quả phát hiện tốt nhất gặp nhiều khó khăn.
Ước lượng vùng hỗ trợ nền	Quang phổ	Những dị thường có kích thước lớn	$O(N^3)**$	<ul style="list-style-type: none"> - Không cần biết trước phổ mục tiêu. - Không cần giả định phân bố của dữ liệu. - Việc tìm ra “siêu cầu kín” rất khó khăn.

Bộ phát hiện dị thường dựa trên cụm	Quang phổ	Từ một phần điểm ảnh đến vài điểm ảnh	$O(L^3)^*$	<ul style="list-style-type: none"> - Không cần biết trước phổ mục tiêu. - Không cần giả định phân bố của dữ liệu. - Độ chính xác phát hiện dị thường cao hơn RXD. - Một số điểm còn tồn tại: thời gian tính toán cao do phải trải qua giai đoạn phân cụm trước; có tỷ lệ dương tính giả cao nếu dữ liệu trong các cụm không tuân theo phân bố chuẩn.
Bộ phát hiện dị thường dựa trên phân đoạn không gian	Dữ liệu phổ và dữ liệu không gian	Những dị thường có kích thước lớn	$O(L^3)^*$	<ul style="list-style-type: none"> - Không cần biết trước phổ mục tiêu. - Không cần giả định phân bố của dữ liệu. - Độ chính xác phát hiện dị thường cao hơn RXD; - Một số điểm còn tồn tại: thời gian tính toán cao do phải trải qua giai đoạn phân đoạn trước; có tỷ lệ dương tính giả cao nếu dữ liệu trong các cụm không tuân theo phân bố chuẩn.

Biểu diễn thưa và biểu diễn hạng thấp	Dữ liệu phổ và dữ liệu không gian	Từ một phần điểm ảnh đến vài điểm ảnh	$O(N^2)$ **	<ul style="list-style-type: none"> - Không cần biết trước phổ mục tiêu. - Không cần giả định phân bố của dữ liệu. - Hiệu suất phát hiện dị thường cao. - Thời gian tính toán cao do phải trải qua giai đoạn tìm nghiệm thưa hoặc tìm ma trận hạng thấp.
Phương pháp dựa trên học máy	Dữ liệu phổ và dữ liệu không gian	Mục tiêu có kích thước lớn	Không theo thời gian thực	<ul style="list-style-type: none"> - Không cần giả định phân bố của dữ liệu; - Hiệu suất phát hiện dị thường cao; - Phải biết trước phổ mục tiêu, dữ liệu đào tạo và dữ liệu kiểm tra phải được thu thập từ cùng một loại sensor.

* Độ phức tạp tính toán của mỗi điểm ảnh tỷ lệ với số kênh phổ.

** Độ phức tạp tính toán của mỗi điểm ảnh tỷ lệ với số lượng các điểm ảnh trong nền cục bộ.

1.4. Tiêu chí đánh giá độ chính xác của các thuật toán phát hiện dị thường

Từ phát biểu của bài toán, mô hình để phân các điểm ảnh vào một trong hai lớp: "dị thường" hoặc "bình thường" là mô hình dự báo sự kiện với 2 khả năng positive (dương tính) và negative (âm tính). Do đó, có bốn trường hợp xảy ra như sau:

- Trường hợp thứ nhất: nếu như PUT là điểm ảnh "dị thường" thật và thuật toán phân PUT vào lớp "dị thường", trường hợp này gọi là dương tính thật (True Positive, viết tắt là TP);

- Trường hợp thứ hai: nếu như PUT là điểm ảnh "bình thường" thật nhưng thuật toán phân PUT vào lớp "dị thường", trường hợp này gọi là dương tính giả (False Positive, viết tắt là FP);

- Trường hợp thứ ba: nếu như PUT là điểm ảnh "bình thường" thật và thuật toán phân PUT vào lớp "bình thường", trường hợp này gọi là âm tính thật (True Negative, viết tắt là TN);

- Trường hợp thứ tư: nếu như PUT là điểm ảnh "dị thường" thật nhưng thuật toán phân PUT vào lớp "bình thường", trường hợp này gọi là âm tính giả (False Negative, viết tắt là FN);

Như vậy, TP và TN là những case dự báo đúng; FP, FN là những case dự báo sai. FN tương đương với mắc sai lầm loại I (Bác bỏ sự kiện là positive và gán cho nó là negative) và FP tương đương với mắc sai lầm loại II (Chấp nhận một sự kiện là positive khi bản chất sự kiện là negative).

Để đánh giá độ chính xác của một thuật toán phát hiện dị thường xem nó tốt hay xấu là một vấn đề khó, bởi vì những điểm ảnh dị thường là những điểm ảnh hiếm. Hầu hết các thuật toán đưa ra giá trị dị thường và ngưỡng để kết luận các điểm ảnh có là dị thường hay không. Việc chọn ngưỡng sẽ ảnh hưởng đến việc bỏ sót (dương tính giả) hoặc nhầm lẫn với các điểm ảnh không phải là dị thường (âm tính giả). Để tránh chủ quan khi chỉ lựa chọn một ngưỡng để đánh giá mô hình, có một cách là duyệt qua hết tất cả các ngưỡng có thể được, và quan sát ảnh hưởng lên tỷ lệ dương tính thật (True Positive Rate, viết tắt là TPR) với tỷ lệ dương tính giả (False Positive Rate, viết tắt là FPR). Khi đó ta có tất cả các điểm (TPR, FPR) theo các ngưỡng khác nhau sẽ tạo thành đường ROC (Receiver operating characteristic). TPR

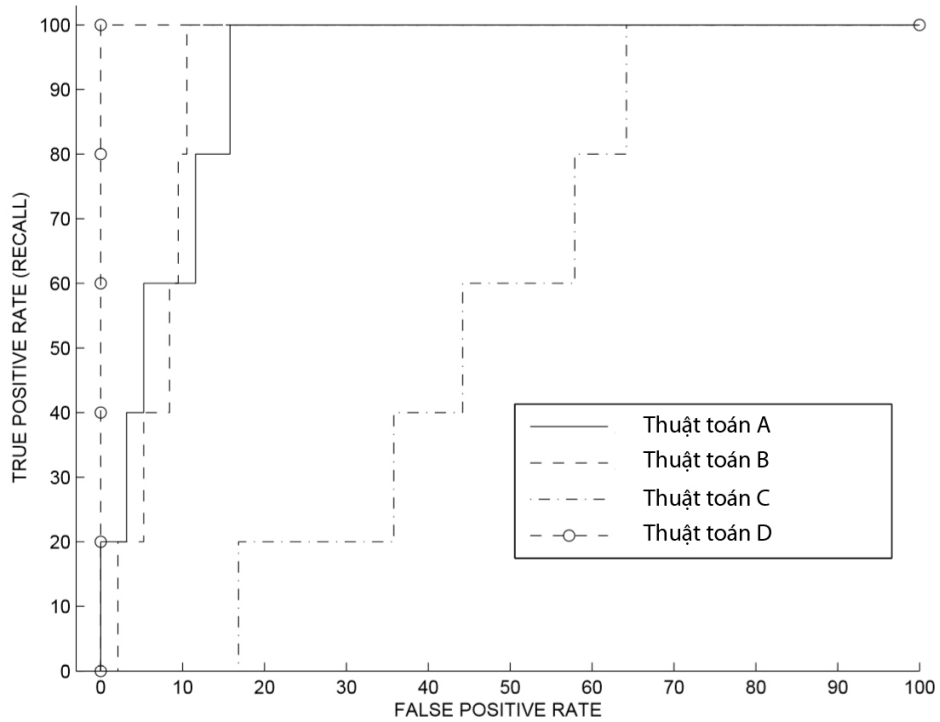
và FPR tại một ngưỡng t được xác định như sau:

Gọi D là tập tất cả các điểm ảnh, G là tập các điểm ảnh dị thường thật trên ảnh, $S(t)$ là tập các điểm ảnh dị thường thu được sau khi chạy thuật toán, với ngưỡng là t . Khi đó,

$$TPR(t) = \frac{|S(t) \cap G|}{|G|} \quad (1.2)$$

$$FPR(t) = \frac{|S(t) - (S(t) \cap G)|}{|D - G|} \quad (1.3)$$

Đường cong ROC là đường biểu diễn các cặp chỉ số $TPR(t)$, $FPR(t)$ tại mỗi



Hình 1.11: Đường cong ROC [6]

ngưỡng t với TPR là trục tung và FPR là trục hoành (xem Hình 1.11). Dựa trên đường cong ROC, có thể chỉ ra rằng một thuật toán có độ chính xác hay không dựa vào giá trị của FPR và TPR, một thuật toán độ chính xác cao khi có FPR thấp đồng thời TPR cao. Tuy đường cong ROC là một công cụ tốt cho ta biết một thuật toán tốt hay xấu, nhưng nó không nói gì đến độ chính xác của thuật toán đạt được. Vì vậy, diện tích dưới đường cong ROC (AUC

- Area Under Curve) được sử dụng để đánh giá độ chính xác của các thuật toán hay chỉ số AUC chính là con số thể hiện độ chính xác phân loại của mô hình. Công thức để tính AUC như sau:

$$AUC = \sum_{t=2}^n \frac{(FPR(t) - FPR(t-1)) * (TPR(t) + TPR(t-1))}{2} \quad (1.4)$$

Trong đó, n là số ngưỡng. Chỉ số AUC càng cao thì thuật toán phát hiện các điểm ảnh dị thường càng tốt; lý tưởng nhất $AUC=1$; AUC càng gần 0,5 thì độ chính xác phân loại các điểm ảnh dị thường của các thuật toán càng tệ; nếu AUC gần 0, các thuật toán phân loại ngược kết quả (phân loại dương tính thành âm tính và ngược lại).

Trong công bố "Facing Imbalanced Data Recommendations for the Use of Performance Metrics" [53], Jeni và các cộng sự đã sử dụng một số độ đo như Accuracy, F-score, AUC, Precision-Recall để đánh giá độ chính xác của các thuật toán nhận biết các đơn vị hành động trên khuôn mặt của các tập dữ liệu mất cân bằng; kết quả cho thấy rằng, chỉ có AUC phản ánh chính xác độ chính xác của các thuật toán, các công cụ còn lại đều bị suy giảm do dữ liệu mất cân bằng, đặc biệt là trong các trường hợp mất cân bằng nghiêm trọng. Vì vậy, trong luận án này, để kiểm nghiệm độ chính xác phát hiện các điểm ảnh dị thường của các thuật toán, chỉ số AUC được sử dụng làm thước đo để đánh giá.

1.5. Phương pháp tính độ phức tạp thuật toán

1.5.1. Khái niệm

Thời gian mà máy tính thực hiện một thuật toán không chỉ phụ thuộc vào bản thân thuật toán đó, ngoài ra còn tùy thuộc từng máy tính. Để đánh giá hiệu quả của một thuật toán, có thể xét số các phép tính phải thực hiện khi thực hiện thuật toán này. Thông thường số các phép tính được thực hiện phụ thuộc vào cỡ của bài toán, tức là độ lớn của đầu vào. Vì thế, độ phức tạp thuật toán là một hàm phụ thuộc đầu vào.

Khi đi đánh giá độ phức tạp của một thuật toán, nghĩa là đang tìm ra một đánh giá về thời gian và không gian cần thiết để thực hiện thuật toán đó.

Không gian ở đây được hiểu là các yêu cầu về bộ nhớ, thiết bị lưu trữ, ...

của máy tính để thuật toán có thể làm việc. Việc xem xét về không gian của thuật toán phụ thuộc phần lớn vào cách tổ chức dữ liệu của thuật toán.

Trong luận án, đánh giá độ phức tạp thuật toán dựa trên đánh giá về mặt thời gian. Đánh giá về thời gian thực hiện thuật toán là xác định mối liên quan giữa dữ liệu đầu vào (input) của thuật toán và chi phí (số thao tác, số phép tính cộng, trừ, nhân, chia, căn, ...) để thực hiện thuật toán. Một cách tổng quát, thời gian thực hiện thuật toán là một hàm số phụ thuộc vào dữ liệu đầu vào: $T = f(\text{input})$.

Tuy vậy, khi phân tích thuật toán, thông thường chỉ chú ý đến độ lớn (kích thước) của dữ liệu đầu vào, độ lớn này thường được thể hiện bằng một số nguyên n . Chẳng hạn: sắp xếp một dãy n số, tìm giá trị lớn nhất trong n số, tính điểm trung bình của n học sinh, ...

Lúc này, thời gian thực hiện thuật toán là một hàm số phụ thuộc vào n : $T = f(n)$.

Việc xây dựng một hàm T tổng quát như trên trong mọi trường hợp của thuật toán là một việc rất khó khăn, nhiều lúc không thể thực hiện được. Chính vì vậy, chỉ đi xây dựng hàm T cho một số trường hợp đáng chú ý nhất của thuật toán, thường là trường hợp tốt nhất, xấu nhất hoặc trung bình.

1.5.2. Đánh giá độ phức tạp thuật toán

1.5.2.1. Ký hiệu Big-O

Khi đánh giá độ phức tạp của một thuật toán, thường dùng 2 ký hiệu O-lớn (Big-O) và Theta (Θ). Tuy nhiên, luận án sẽ sử dụng ký hiệu O-lớn.

Thuật toán A có thời gian thực hiện là $T(n) = O(f(n))$. Khi đó, thuật toán A có độ phức tạp $f(n)$.

1.5.2.2. Các quy tắc tính độ phức tạp

*** Quy tắc bỏ hằng số:**

$T(n) = O(C.f(n)) = O(f(n))$ với C là một hằng số dương.

*** Quy tắc cộng:**

Hai chương trình P_1 và P_2 nối tiếp nhau. $T_1(n) = O(f(n))$ là thời gian thực hiện chương trình 1 và $T_2(n) = O(g(n))$ là thời gian thực hiện chương trình 2. Thời gian thực hiện tổng thể 2 chương trình này là $O(f(n) + g(n)) = O(\text{Max}(f(n), g(n)))$.

***Quy tắc nhân:**

Hai chương trình P_1 và P_2 lồng nhau. $T_1(n) = O(f(n))$ là thời gian thực hiện chương trình P_1 và $T_2(n) = O(g(n))$ là thời gian thực hiện chương trình P_2 . Khi đó, thời gian thực hiện tổng thể 2 chương trình này là $O(f(n).g(n))$.

1.5.3. Thời gian thực hiện các câu lệnh trong các ngôn ngữ lập trình

1.5.3.1. Các câu lệnh đơn giản

Các câu lệnh đơn giản như gán, so sánh, return, đọc/ghi, ... có thời gian thực hiện là $O(1)$.

1.5.3.2. Vòng lặp

Trong phần lớn các trường hợp, thời gian thực hiện vòng lặp bằng thời gian thực hiện thân vòng lặp nhân với số vòng lặp.

1.5.3.3. Câu lệnh rẽ nhánh

Tổng thời gian thực hiện lớn nhất bằng thời gian kiểm tra điều kiện if (thường là hằng số) cộng với thời gian thực hiện then cộng với thời gian thực hiện else.

1.6. Thách thức và hướng phát triển

Các thách thức đặt ra trong việc phát hiện các điểm ảnh dị thường trên ảnh đa phổ, siêu phổ gồm: thách thức đầu tiên liên quan đến việc tăng tỷ lệ phát hiện trong khi giảm tỷ lệ dương tính giả. Thách thức thứ hai liên quan đến việc phát hiện dị thường với các hình dạng và kích cỡ khác nhau. Trong thực tế, kích thước của dị thường có thể dao động từ một phần điểm ảnh đến vài điểm ảnh và việc phát hiện các dị thường với kích thước khác nhau cùng một kỹ thuật phát hiện dị thường là một thách thức lớn. Thách thức thứ ba, đó là tốc độ tính toán, các kỹ thuật phát hiện dị thường có thời gian tính toán gần với thời gian thực sẽ rất hữu ích trong các ứng dụng thực tế, đặc biệt là trong công tác tìm kiếm cứu nạn.

1.6.1. Thách thức về tăng độ chính xác phát hiện dị thường

Ngay từ đầu, các nhà nghiên cứu đã tiếp cận với vấn đề không có bất kỳ kiến thức trước về dữ liệu được xử lý. Do đó, các nhà khoa học đã cố gắng sử dụng các phương pháp thống kê để so sánh giữa PUT và nền. Đối với các

phương pháp thống kê, nền được mô hình hóa với phân bố tuyến tính của PDF và giả sử tính đồng nhất của nó. Giả định này làm tăng tỷ lệ dương tính giả, đặc biệt đối với các hình ảnh có độ phân giải cao trong đó giả định về tính đồng nhất dường như không phù hợp vì sự đa dạng lớn của các vật liệu hiện diện trong ảnh. Để giảm tỷ lệ dương tính giả, tăng độ chính xác phát hiện dị thường, nhiều thuật toán được đề xuất như những giải pháp để xử lý vấn đề này.

Để cải thiện tỷ lệ dương tính giả, các phương pháp thống kê bậc cao và ước lượng mật độ xác suất phi tham số nền do Ren và các cộng sự [97], Gu và các cộng sự [38], Chang và các cộng sự [17], Veracini và các cộng sự [120, 121], Matteoli và các cộng sự [83],... được đề xuất. Cách tiếp cận này bắt nguồn từ thực tế là sự đóng góp của dị thường vào số liệu thống kê bậc cao có ý nghĩa hơn so với sự đóng góp của nó vào số liệu thống kê thứ hai. Chúng có khả năng phát hiện dị thường có kích thước khác nhau nhưng rất tốn thời gian. Một giải pháp khác để giảm tỷ lệ dương tính giả là sử dụng các kỹ thuật phát hiện dị thường dựa trên nhân có thể mô hình hóa nền với phân phối phi tuyến tính. Chúng cho kết quả phát hiện tốt hơn so với RXD, chủ yếu cho hình ảnh có độ phân giải cao nhưng việc xác định các thông số tối ưu của hạt nhân là một vấn đề khó khăn và có thời gian tính toán vô cùng lớn nên khó áp dụng trong những ứng dụng thực tế.

Các kỹ thuật phát hiện dị thường dựa trên phân cụm và phân đoạn được đề xuất để giảm tỷ lệ dương tính giả. Trong thực tế, phân cụm và phân đoạn cho phép chúng ta tạo ra một bản đồ chuyên đề trước khi phát hiện dị thường, bổ sung kiến thức nội dung hiện có về hình ảnh và phân phối của chúng tạo điều kiện cho quá trình phát hiện. Những phương pháp này đã giảm tỷ lệ dương tính giả, cho kết quả tốt hơn phương pháp phát hiện dị thường dựa trên thống kê cho cảnh có độ phân giải cao và có thể phát hiện dị thường có kích thước khác nhau. Tuy nhiên, xử lý thời gian thực là một chủ đề cần đầu tư nghiên cứu vì cần nhiều thời gian cho phân cụm hoặc phân đoạn.

Việc giảm tỷ lệ dương tính giả cũng đồng nghĩa với tăng độ chính xác phát hiện dị thường của các thuật toán. Ngoài việc giả sử tính đồng nhất của nền gây ra tỷ lệ dương tính giả, còn hai nguyên nhân sau cũng tạo nên tỷ lệ dương tính giả: thứ nhất, các thuật toán phát hiện dị thường tính toán cục bộ nên không xét đến mối tương quan của các điểm ảnh trong toàn bộ cảnh làm cho

tỷ lệ dương tính giả cao. Thứ hai, quá trình thu nhận hình ảnh của các bộ cảm biến thường có nhiễu được thêm vào (nhiều ánh sáng, nhiễu khí quyển, nhiễu tín hiệu chuyển đổi của bộ cảm biến,...). Trong khi đó, hầu hết các thuật toán hoạt động như một bộ lọc dẫn đến các điểm ảnh là nhiễu sẽ được đưa vào lớp dị thường làm cho tỷ lệ dương tính giả cao. Vì vậy, quá trình tiền xử lý dữ liệu, trích chọn các đặc trưng trước khi phân tích dữ liệu có thể làm tăng độ chính xác phát hiện dị thường của các thuật toán.

Xem xét một điểm ảnh bất kỳ trong một cảnh chúng ta thấy rằng, điểm ảnh này ít nhiều đều có mối quan hệ với những điểm ảnh xung quanh nó. Tuy nhiên, ngoại trừ một số ít các nghiên cứu sử dụng phương pháp biểu diễn thưa, còn lại đều bỏ qua vấn đề này. Nghiên cứu phân tích mối quan hệ, ràng buộc dữ liệu, mối tương quan về mặt không gian của các điểm ảnh có thể làm tăng độ chính xác phát hiện dị thường trong thời gian tới.

1.6.2. Thách thức về phát hiện dị thường với các kích cỡ khác nhau

Rõ ràng rằng, một mục tiêu lớn tới mức có thể xác định được bằng mắt thường sẽ không được coi là dị thường. Với cách giải thích này, kích thước của dị thường phải tương đối nhỏ. Vậy mục tiêu đó nhỏ như thế nào để được xem là dị thường? Có một điều thú vị là nó liên quan đến tỷ lệ của toàn bộ kích thước hình ảnh với kích thước của dị thường. C.-I. Chang đã đưa ra một ngưỡng để xác định kích thước của một dị thường $n = \frac{N}{100}$, trong đó N là tổng số điểm ảnh trong hình ảnh [14].

Các kỹ thuật được Chang và Chiang mô tả trong [16], Du và Zhang [25], Molero và các cộng sự [88] đã thành công trong việc xử lý thời gian thực. Tuy nhiên, họ bỏ qua các dị thường có kích thước nhỏ, các dị thường kích thước nhỏ có thể được phát hiện bởi RXD cục bộ, tuy nhiên thách thức là phát hiện các dị thường có kích thước khác nhau với cùng một cách tiếp cận, do đó Liu và Chang [70] đã đề xuất một cách tiếp cận dựa trên việc sử dụng nhiều cửa sổ. Cách tiếp cận này cung cấp một giải pháp xử lý gần như theo thời gian thực cho từng điểm ảnh.

Chen và các cộng sự [21, 20] sử dụng phương pháp biểu diễn thưa để phát hiện các điểm ảnh dị thường, trong nghiên cứu [127] Xu và các cộng sự đề xuất phương pháp phát hiện dị thường dựa trên biểu diễn hạng thấp. Các phương pháp này sử dụng cho cảnh có độ phân giải cao và có thể phát hiện

dị thường có kích thước khác nhau. Tuy nhiên, xử lý thời gian thực rất khó thực hiện do phải trải qua giai đoạn tìm nghiệm thưa hoặc tìm ra ma trận hạng thấp sau đó mới có thể tìm được các điểm ảnh dị thường.

Trong thời gian gần đây, phương pháp học sâu đã thu hút sự chú ý của các nhà khoa học và áp dụng nó ngày càng nhiều trong phân tích ảnh siêu phổ. Li và các cộng sự [65] đã đề xuất sử dụng mạng học sâu CNN, Yan và các cộng sự [128] đề xuất phương pháp 2SHyperNet, Zhang và Cheng [133] đề xuất phương pháp SAEASM để phát hiện các điểm ảnh dị thường. Phương pháp học sâu đã cho thấy sự hiệu quả trong vấn đề này, tuy nhiên phải có dữ liệu đào tạo được gán nhãn lớn và tín hiệu phổ mục tiêu phải biết trước. Ngoài ra, dữ liệu đào tạo và dữ liệu kiểm tra phải được thu thập cùng một loại cảm biến.

Các phương pháp dựa trên phân cụm và phân đoạn sẽ là tương lai của phát hiện dị thường vì khả năng phát hiện của chúng ngay cả trong các cảnh có độ phân giải cao, đặc biệt là khi sử dụng đặc tính phổ không gian của dữ liệu. Ngoài ra, phương pháp dựa trên biểu diễn thưa hoặc biểu diễn hạng thấp đã và đang được nhiều nhà nghiên cứu quan tâm. Bởi phương pháp này giả định rằng tín hiệu phổ có thể được biểu diễn bằng cách sử dụng từ điển, với các ràng buộc khác nhau về các hệ số biểu diễn, chúng ta có thể thu được các kỹ thuật phát hiện dị thường khác nhau.

1.6.3. Thách thức về tốc độ tính toán

Một số ít thuật toán đã được đề xuất có thể tính toán theo thời gian thực như đề xuất của Chang và cộng sự [16], Du và cộng sự [25] và Molero và các cộng sự [88]. Tuy nhiên, các thuật toán này phải đánh đổi bằng việc đã bỏ qua các dị thường có kích thước nhỏ. Hiện nay, các thiết bị phần cứng đã và đang rất phát triển, có hiệu năng và tốc độ tính toán cao. Tuy nhiên, đến một giai đoạn nào đó việc tính toán chỉ dựa trên tốc độ tính toán của phần cứng sẽ không đáp ứng được yêu cầu của các ứng dụng bởi nó sẽ dần đạt đến ngưỡng trong khi khối lượng dữ liệu tính toán ngày một tăng. Do đó nghiên cứu, phát triển các thuật toán vừa mang lại độ chính xác vừa có tốc độ tính toán cao sẽ rất hữu ích.

Từ các thách thức đã nêu, chúng ta thấy rằng, để giải quyết cả ba thách thức trong cùng một thuật toán là một điều hết sức khó khăn và cho đến nay

chưa có đề xuất nào có thể thực hiện được. Chỉ có thể căn cứ vào mức độ ưu tiên của từng ứng dụng để vận dụng một cách linh hoạt các thuật toán đã công bố. Trong công tác tìm kiếm cứu nạn, vì liên quan đến tính mạng của con người nên yêu cầu cao về thời gian xử lý và độ chính xác phát hiện. Vì vậy, luận án lựa chọn hướng nghiên cứu để giải quyết thách thức thứ nhất và thách thức thứ ba của phương pháp phát hiện dị thường dựa trên hàm mật độ xác suất phi tham số.

1.7. Kết luận

Một số vấn đề đã được trình bày trong chương này gồm:

- *Thứ nhất*, trình bày một số kiến thức tổng quan về ảnh viễn thám quang học và ứng dụng ảnh viễn thám quang học phục vụ công tác tìm kiếm cứu nạn.
- *Thứ hai*, phát biểu về bài toán phát hiện dị thường trên ảnh viễn thám quang học, là cơ sở để nghiên cứu tìm các giải pháp để giải quyết bài toán này.
- *Thứ ba*, nghiên cứu các cách tiếp cận trên thế giới về bài toán phát hiện dị thường trên ảnh viễn thám quang học. Trên cơ sở đó, luận án đã chỉ ra những kết quả đạt được và nêu lên ba thách thức cần được giải quyết.

Từ ba thách thức đã được đưa ra, trong Chương 2 luận án đề xuất các nghiên cứu nhằm giải quyết vấn đề liên quan đến thách thức đầu tiên, đó là tăng độ chính xác phát hiện các điểm ảnh dị thường của thuật toán phát hiện các điểm ảnh dị thường dựa trên ước lượng mật độ nhân [120, 121]. Chương 3 đề xuất thuật toán nhằm giải quyết vấn đề liên quan đến thách thức thứ ba, đó là đề xuất thuật toán tăng tốc độ tính toán của thuật toán phát hiện các điểm ảnh dị thường đã đề xuất ở Chương 2.

Chương 2

KỸ THUẬT NÂNG CAO ĐỘ CHÍNH XÁC PHÁT HIỆN DỊ THƯỜNG TRÊN ẢNH VIỄN THÁM QUANG HỌC

Trong chương trình bày giải pháp tăng độ chính xác phát hiện các điểm ảnh dị thường của thuật toán phát hiện các điểm ảnh dị thường dựa trên ước lượng mật độ nhân (KDE - Kernel Density Estimation) [120, 121]. Đề xuất được thực hiện bằng cách loại bỏ một phần nhiều khỏi tập các điểm ảnh dị thường được tìm thấy của thuật toán trên (giảm tỷ lệ dương tính giả). Cài đặt thuật toán, thử nghiệm, đánh giá giải pháp đề xuất và kết luận. Kết quả nghiên cứu tăng độ chính xác phát hiện các điểm ảnh dị thường của thuật toán phát hiện các điểm ảnh dị thường dựa trên KDE được công bố trong các công trình [CT2], [CT3].

2.1. Giới thiệu chung

Phát hiện dị thường trên ảnh viễn thám quang học nhằm mục đích hỗ trợ tìm kiếm nhanh các điểm ảnh hoặc nhóm điểm ảnh chứa đựng thông tin hữu ích cho công tác tìm kiếm cứu nạn. Trong một số trường hợp, thông tin có giá trị cho công tác này chỉ nằm trong một cụm vài điểm ảnh (do độ phân giải của ảnh không cao, do góc chụp của thiết bị, do đối tượng bị che lấp một phần bởi địa hình, mật độ dày đặc của lá cây hay bị chìm một phần dưới nước,...), bằng mắt thường khó phát hiện được chúng. Để tránh bỏ sót các đối tượng như vậy, các thuật toán phát hiện các điểm ảnh dị thường phải thay đổi ngưỡng phát hiện, dẫn đến có nhiều điểm ảnh là nhiễu được thêm vào (làm tăng tỷ lệ dương tính giả) gây ra tình trạng nhiễu loạn thông tin, làm khó khăn thêm cho quá trình tìm kiếm. Vì vậy, việc nghiên cứu các giải pháp tăng độ chính xác phát hiện các điểm ảnh dị thường của các thuật toán là hết sức cần thiết mang nhiều ý nghĩa.

Trong hơn 20 năm qua, nhiều phương pháp tìm các điểm ảnh dị thường trên ảnh viễn thám quang học đã được xây dựng như phương pháp dựa trên mô hình xác suất thống kê, dựa trên phân cụm và phân đoạn, trên biến đổi

phi tuyến, biến đổi tuyến tính và dựa trên học máy. Thuật toán phát hiện các điểm ảnh dị thường dựa trên KDE [120, 121] là một trong những phương pháp phát hiện các điểm ảnh dị thường dựa trên mô hình xác suất thống kê, thuật toán này đã được xây dựng và kiểm chứng trong những năm gần đây, trong các nghiên cứu [12, 83, 135] đã cho thấy được độ chính xác trong việc phát hiện các điểm ảnh dị thường trên ảnh đa phổ, siêu phổ và độ chính xác phát hiện dị thường của nó đã vượt RXD (là một thuật toán chuẩn để phát hiện các điểm ảnh dị thường, ngày nay, nó vẫn được sử dụng để so sánh với các thuật toán phát hiện điểm ảnh dị thường mới [61, 76, 112, 114]). Chính những lý do đó, thuật toán phát hiện các điểm ảnh dị thường dựa trên KDE được lựa chọn để nghiên cứu và cải tiến nhằm tăng độ chính xác phát hiện các điểm ảnh dị thường trên ảnh đa phổ, siêu phổ ứng dụng cho công tác tìm kiếm cứu nạn.

2.2. Thuật toán phát hiện các điểm ảnh dị thường dựa trên KDE

Phát hiện các điểm ảnh dị thường trên ảnh đa phổ, siêu phổ có thể được coi là nhiệm vụ phân các điểm ảnh vào một trong hai lớp "*bình thường*" (lớp C_1) hoặc "*dị thường*" (lớp C_2). Xét một ảnh có L kênh phổ, ảnh này được tạo thành từ tập hợp gồm N điểm ảnh $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$. Quan sát thứ i trên \mathbf{X} là $\mathbf{x}_i = (\mathbf{x}_{i1}, \mathbf{x}_{i2}, \dots, \mathbf{x}_{iL})^T, i = 1, 2, \dots, N$, để có thể phân \mathbf{x}_i vào lớp "*bình thường*" hoặc "*dị thường*" theo phương pháp thống kê là giải quyết vấn đề kiểm nghiệm giả thuyết nhị phân bằng cách sử dụng quy tắc Neyman-Pearson dựa trên tỷ lệ khả năng (LR) của hàm mật độ xác suất có điều kiện theo hai giả thuyết:

$$H(\mathbf{x}_i) = \begin{cases} H_0 : \mathbf{x}_i \in C_1; \\ H_1 : \mathbf{x}_i \in C_2. \end{cases} \quad i = 1, 2, \dots, N \quad (2.1)$$

Theo quy tắc Neyman-Pearson, ta có:

$$\Lambda_{NP}(\mathbf{x}_i) = \frac{\hat{f}_{\mathbf{X}|H_1}(\mathbf{x}_i)}{\hat{f}_{\mathbf{X}|H_0}(\mathbf{x}_i)} \underset{H_0}{\overset{H_1}{\geq}} \eta, i = 1, 2, \dots, N \quad (2.2)$$

Trong đó $\hat{f}(\cdot)$ là hàm mật độ xác suất có điều kiện, η là ngưỡng thích hợp để phân \mathbf{x}_i vào lớp "*bình thường*" hoặc lớp "*dị thường*". Tuy nhiên, trong thực

tế, các tham số để tính toán các hàm mật độ xác suất có điều kiện trong công thức (2.2) thường không có sẵn bởi vì chúng ta thiếu kiến thức lớp "*di thường*", hoàn toàn chưa có các nghiên cứu thống kê về nó, hơn nữa các điểm ảnh dị thường có giá trị phổ bất kỳ, chỉ phụ thuộc vào cảnh của ảnh, chúng là những điểm độc lập hoặc cụm điểm ảnh nhưng có kích thước rất nhỏ so với ảnh và có mật độ thưa thớt. Vì vậy, chúng ta có thể giả định rằng $\hat{f}_{X|H_1}(x_i)$ là một hằng số, khi đó công thức (2.2) được rút ngắn thành:

$$\Lambda(\mathbf{x}_i) = -\log\{\hat{f}_{\mathbf{X}|H_0}(\mathbf{x}_i)\} \begin{matrix} \underset{H_1}{\geq} \\ \underset{H_0}{<} \end{matrix} \eta, i = 1, 2, \dots, N \quad (2.3)$$

Do mật độ xác suất của dữ liệu nền $\hat{f}_{\mathbf{X}|H_0}(\mathbf{x}_i)$ chưa được biết nên nó phải được ước tính từ dữ liệu có sẵn. Với giả thuyết rằng, các điểm ảnh dị thường rất ít, mật độ thưa nên tất cả các điểm ảnh $\mathbf{x}_i \in \mathbf{X}, i = 1, 2, \dots, N$ có thể được sử dụng cho ước tính này.

Trong các nghiên cứu [83, 120, 121, 135], các tác giả đã sử dụng phương pháp ước tính mật độ xác suất phi tham số để ước lượng $\hat{f}_{\mathbf{X}|H_0}(\mathbf{x}_i)$ bởi nó có ưu điểm là không cần phải đưa ra bất kỳ một giả định nào về phân phối của dữ liệu. Trong phương pháp này, công cụ chính của nó là KDE đã được Rosenblatt công bố vào năm 1956 [99] và sau đó được Parzen phát triển, công bố vào năm 1962 [91]. Đối với dữ liệu một chiều, xét vector ngẫu nhiên $\mathbf{x} = (x_1 x_2 \dots x_N)^T$ của biến ngẫu nhiên \mathbf{x} có N phần tử. Điều này có nghĩa rằng có N quan sát của biến ngẫu nhiên \mathbf{x} và x_i là quan sát thứ i của biến ngẫu nhiên \mathbf{x} . Khi đó, ước tính mật độ nhân của biến ngẫu nhiên $\mathbf{x} = (x_1 x_2 \dots x_N)^T$ như sau:

$$\hat{f}(x_i) = \frac{1}{N} \sum_{j=1}^N \frac{1}{h_j} K\left(\frac{x_i - x_j}{h_j}\right), i = 1, 2, \dots, N \quad (2.4)$$

Trong đó, $\hat{f}(\cdot)$ gọi là hàm mật độ xác suất (PDF - Probability Density Function), $K(u)$ được gọi là hàm nhân thỏa mãn điều kiện $\int_{-\infty}^{\infty} K(u) d(u) = 1$ và h_j là hệ số tỷ lệ quyết định "*khoảng rộng*" của hàm nhân hay còn gọi là băng thông. Thảo luận mở rộng về các thuộc tính thống kê của $\hat{f}(\cdot)$ có thể được tìm thấy trong [23], $K(u)$ có thể là các hàm nhân điển hình do Hardle

trình bày trong [122] được thể hiện trên Bảng 2.1.

Bảng 2.1: Một số nhân điển hình [122]

Tên hàm nhân	$K(u)$	Điều kiện
Uniform	$\frac{1}{2}$	$ u \leq 1$
Hypercube	1	$ u \leq \frac{1}{2}$
Triangular	$1 - u $	$ u \leq 1$
Epanechnikov	$\frac{3}{4\sqrt{5}} - \frac{3u^2}{20\sqrt{5}}$	$ u \leq \sqrt{5}$
Quartic	$\frac{15}{16}(1 - u^2)^2$	$ u \leq 1$
Triweight	$\frac{35}{32}(1 - u^2)^3$	$ u \leq 1$
Tricube	$\frac{70}{81}(1 - u ^3)^3$	$ u \leq 1$
Gaussian	$\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}$	
Cosine	$\frac{\pi}{4} \cos\left(\frac{\pi}{2}u\right)$	$ u \leq 1$

Trong trường hợp dữ liệu có L chiều, quan sát thứ i của \mathbf{X} sẽ là $\mathbf{x}_i = (\mathbf{x}_{i1}, \mathbf{x}_{i2}, \dots, \mathbf{x}_{iL})^T, i = 1, 2, \dots, N$, công thức ước tính mật độ nhân của dữ liệu đa biến được định nghĩa trong [122] là:

$$\hat{f}(\mathbf{x}_i) = \frac{1}{N} \sum_{j=1}^N \left\{ \prod_{d=1}^L \frac{1}{h_d} K\left(\frac{\mathbf{x}_{id} - \mathbf{x}_{jd}}{h_d}\right) \right\} \quad (2.5)$$

Đối với các ảnh đa phổ, siêu phổ, dữ liệu thuộc dạng đa biến nên luận án sử dụng công thức (2.5) để cài đặt thuật toán. Không làm mất tính tổng

quát, sẽ cố định băng thông, đặt $h = h_1 = h_2 = \dots = h_d$ với $d = 1, 2, \dots, L$. Thuật toán 2.1 được xây dựng để phân các điểm ảnh vào một trong hai lớp C_1 và C_2 .

Thuật toán 2.1: Phát hiện các điểm ảnh dị thường dựa trên KDE (gọi tắt là Thuật toán KDE)[120, 121]

input: X : Ma trận các điểm ảnh, N : số điểm ảnh, L : số kênh phổ, h : băng thông, η : ngưỡng phát hiện dị thường.

output: C_1 : tập các điểm ảnh "bình thường", C_2 : tập các điểm ảnh "dị thường".

```

1  $C_1 \leftarrow \emptyset$ ;
2  $C_2 \leftarrow \emptyset$ ;
3 for  $i \leftarrow 1$  to  $N$  do
4    $sum\_ker \leftarrow 0$ ;
5   for  $j \leftarrow 1$  to  $N$  do
6      $mul\_ker \leftarrow 1$ ;
7     for  $d \leftarrow 1$  to  $L$  do
8        $mul\_ker \leftarrow mul\_ker \times K \left( \frac{X[i][d] - X[j][d]}{h} \right)$ ;
9      $sum\_ker \leftarrow sum\_ker + mul\_ker$ ;
10    if  $\frac{sum\_ker}{N} \leq \eta$  then
11       $C_2 \leftarrow C_2 \cup \{X[i]\}$ ;
12    else
13       $C_1 \leftarrow C_1 \cup \{X[i]\}$ ;

```

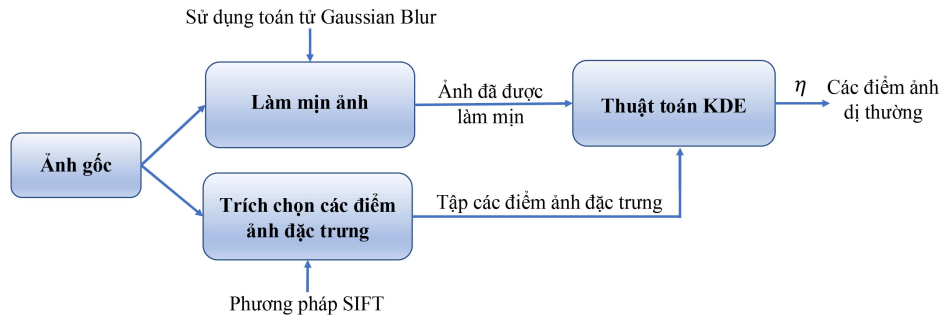
2.3. Giải pháp tăng độ chính xác của thuật toán phát hiện các điểm ảnh dị thường dựa trên KDE

Thuật toán phát hiện các điểm ảnh dị thường dựa trên KDE (Thuật toán 2.1) hoạt động theo nguyên tắc dựa trên giá trị của hàm mật độ xác suất để phân lớp. Nếu điểm ảnh đang xét có mật độ xác suất thấp (nhỏ hơn ngưỡng η), nó sẽ được phân về lớp "dị thường" và ngược lại. Trong khi đó, các điểm ảnh là nhiễu thường chiếm một phần nhỏ trong ảnh nên nó sẽ được phân vào lớp dị thường làm tăng tỷ lệ dương tính giả, dẫn đến giảm độ chính xác phát hiện dị thường của thuật toán.

Do các hạn chế vật lý vốn có của các thiết bị chụp khác nhau, hình ảnh có xu hướng bị nhiễu ngẫu nhiên trong quá trình thu nhận hình ảnh. Nhiễu có thể hiểu là hiện tượng méo tín hiệu gây cản trở quá trình quan sát hình ảnh và trích xuất thông tin. Việc chụp ảnh trong điều kiện khí quyển, ánh sáng không thuận lợi hay các phương pháp nén, khôi phục hình ảnh cũng tạo nên nhiễu. Bên cạnh đó, nhiễu xuất hiện trong ảnh bởi nhiều nguyên nhân như do sự thay đổi độ nhạy của sensor, do sự biến đổi của môi trường, do chính bản thân chất liệu sinh ra (nhiều từ vật liệu liền kề), do sai số lượng tử hóa hay sai số truyền,... Tất cả các nguyên nhân gây ra nhiễu ở trên đã sinh ra nhiễu, nhiễu trong ảnh được phân thành các loại chính như sau:

- Nhiễu Gauss: Nhiễu này có được do bản chất rời rạc của bức xạ (hệ thống ghi ảnh bằng cách đếm các photon (lượng tử ánh sáng)). Mỗi pixel trong ảnh nhiễu là tổng giá trị pixel đúng và pixel ngẫu nhiên.
- Nhiễu muối – tiêu (Salt & Pepper noise): Nhiễu này sinh ra do xảy ra sai số trong quá trình truyền dữ liệu. Những pixel đơn được đặt luân phiên mang giá trị bằng 0 hay giá trị cực đại tạo ra hình chấm dạng muối tiêu trên ảnh.
- Nhiễu Shot hay nhiễu Poisson: Nhiễu này sinh ra do trong quá trình thu nhận, số lượng lớn hạt photon đã tập trung vào một điểm và chúng đã tạo ra nhiễu tại điểm đó. Nhiễu được đặc trưng bởi hàm mật độ phân bố xác suất Poisson, nên được gọi là nhiễu Poisson.
- Nhiễu Speckle hay nhiễu đốm: Là loại nhiễu phát sinh do ảnh hưởng của điều kiện môi trường lên cảm biến hình ảnh trong quá trình thu nhận hình ảnh. Nhiễu lốm đốm hầu như được phát hiện trong trường hợp ảnh y tế, ảnh Radar.

Trong ảnh viễn thám quang học cũng có các loại nhiễu như đã trình bày ở trên, ngoài ra còn có các nhiễu mang tính đặc trưng như sương mù, mây... Với các nhiễu đặc thù cần có phương pháp khử nhiễu đặc thù như phương pháp loại bỏ sương mù, đám mây sử dụng thuật toán Mallat [11]. Đối với nhiễu thông thường, sẽ dùng các phương pháp khử nhiễu thông thường như lọc nhiễu hay làm trơn ảnh (làm mịn ảnh) [4]. Luận án đã sử dụng phương pháp làm mịn ảnh để giảm nhiễu trước khi thực hiện thuật toán.



Hình 2.1: Các bước nhằm tăng độ chính xác phát hiện dị thường.

Trong nghiên cứu về tăng độ chính xác phát hiện dị thường trên ảnh UAV ứng dụng trong công tác tìm kiếm cứu nạn [CT3], NCS và tập thể cán bộ hướng dẫn đã đề xuất phương pháp giảm tỷ lệ dương tính giả của thuật toán phát hiện các điểm ảnh dị thường dựa trên KDE bằng cách thực hiện theo ba bước (xem minh họa trên Hình 2.1), gồm: trích chọn các điểm ảnh đặc trưng và làm mịn ảnh; tính toán thuật toán phát hiện các điểm ảnh dị thường dựa trên KDE cho những điểm ảnh xung quanh các điểm ảnh đặc trưng; phân loại các điểm ảnh đã được tính toán.

Bước đầu tiên, trích chọn các điểm ảnh đặc trưng và làm mịn ảnh. Để giảm một phần nhiễu trên ảnh, luận án sử dụng dụng phương pháp làm mịn ảnh bằng cách tích chập ảnh gốc với toán tử Gaussian Blur:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma} e^{-(x^2+y^2)/2\sigma^2}, \quad (2.6)$$



Hình 2.2: Ảnh được làm mịn với các hệ số làm mịn (σ) khác nhau

Trong đó, σ là hệ số làm mịn, σ càng cao thì ảnh càng mịn. Tuy nhiên, σ càng cao thì độ sắc nét (chi tiết) của các đối tượng càng giảm (xem Hình 2.2), có thể làm mất tính đặc trưng của các điểm ảnh dị thường (có màu sắc nổi bật so với những điểm ảnh lân cận) dẫn đến độ chính xác phát hiện các

điểm ảnh dị thường của thuật toán giảm so với ảnh chưa qua làm mịn. Thuật toán 2.2 được xây dựng theo công thức 2.6 để làm mịn ảnh.

Thuật toán 2.2: Thuật toán làm mịn ảnh (GaussianBlur)

input: X : ma trận các điểm ảnh, h : số hàng, w : số cột, L : số kênh phổ, $k=3$: kích thước ma trận hạt nhân, σ : hệ số làm mịn.

output: X' : ma trận các điểm ảnh đã được làm mịn.

```

1  $ker \leftarrow \text{array}[k][k]$ ;
2  $f \leftarrow (k-1)/2$ ;
3  $constant \leftarrow \frac{1}{2 \times \pi \times \sigma^2}$ ;
4 for  $y \leftarrow -f$  to  $f$  do
5   for  $x \leftarrow -f$  to  $f$  do
6      $distance \leftarrow \frac{x^2 + y^2}{2 \times \sigma^2}$ ;
7      $ker[y + f + 1][x + f + 1] = constant \times e^{-distance}$ ;
8 for  $y \leftarrow 1$  to  $h$  do
9   for  $x \leftarrow 1$  to  $w$  do
10    for  $i \leftarrow y - f$  to  $y + f$  do
11      for  $j \leftarrow x - f$  to  $x + f$  do
12        if  $i > 0$  And  $i \leq h$  And  $j > 0$  And  $j \leq w$  then
13          for  $d \leftarrow 1$  to  $L$  do
14             $X'[(y-1) \times w + x][d] = X[(y-1) \times w + x][d] +$ 
               $X[(i-1) \times w + j][d] \times ker[i-(y-f)+1][j-(x-f)+1]$ ;
15 for  $i \leftarrow 1$  to  $h \times w$  do
16    $X'[i] = \frac{X'[i]}{k^2}$ ;
17 retun  $X'$ ;

```

Dựa vào tính chất của các điểm ảnh dị thường là có màu sắc nổi bật hoặc có sự khác biệt lớn so với các điểm ảnh xung quanh và chúng rất thưa thớt, hiếm khi đại diện cho ảnh, luận án sẽ sử dụng một công cụ để định vị các điểm ảnh này. Tập những điểm ảnh được định vị này có thể không chứa tất cả các điểm ảnh dị thường và có thể có thêm các điểm ảnh bình thường nhưng chúng là những điểm ảnh có màu sắc nổi bật (có thể có giá trị lớn nhất hoặc nhỏ nhất) trong một vùng cục bộ, những điểm ảnh này gọi là tập

những điểm đặc trưng. Từ các điểm ảnh đặc trưng, sử dụng thuật toán KDE để tính toán các điểm xung quanh trong một phạm vi nhất định để tìm các điểm ảnh dị thường thực sự, các điểm ảnh còn lại sẽ được bỏ qua và coi nó là những điểm ảnh bình thường. Như vậy, thay vì phải tính toán thuật toán KDE cho toàn bộ các điểm ảnh, chúng ta chỉ đi tính toán cho một tập hợp nhỏ các điểm ảnh xung quanh các điểm đặc trưng, từ đó giảm được nhiều bị gán vào tập giá trị dị thường làm giảm tỷ lệ dương tính giả.

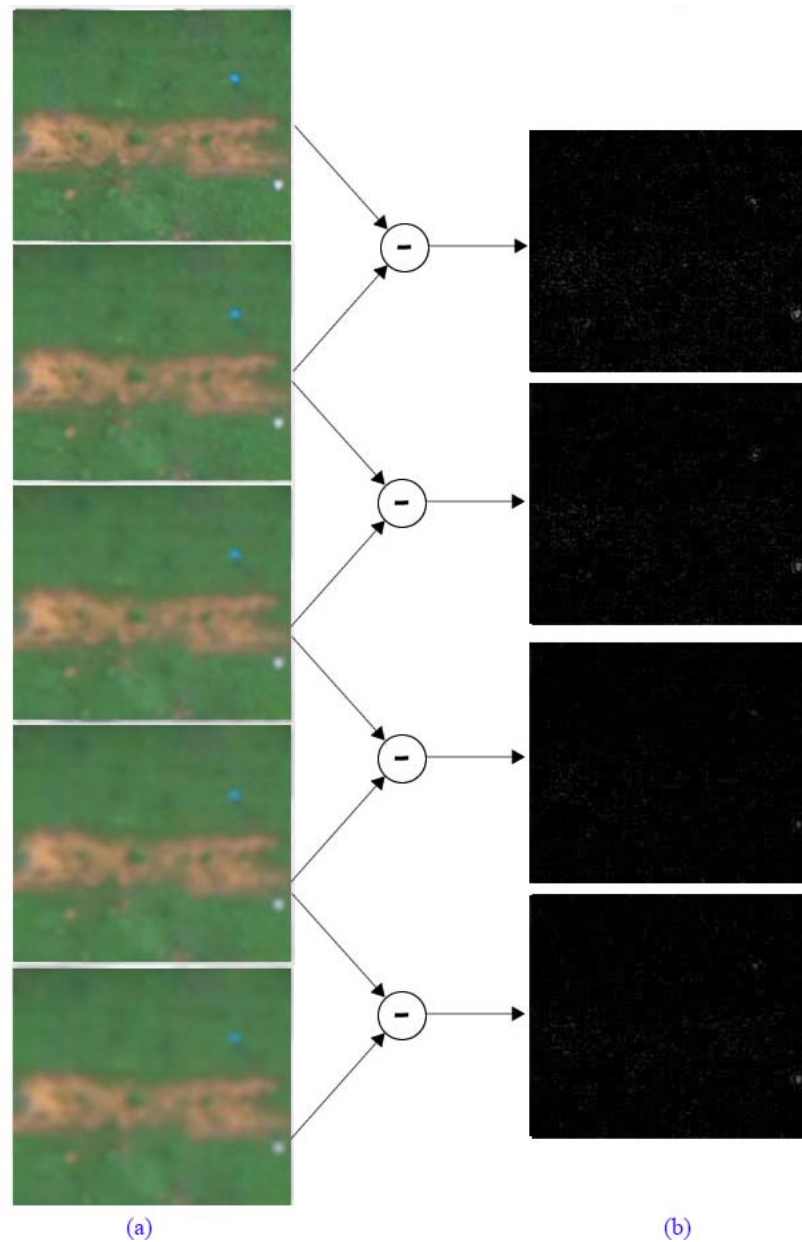
Qua nghiên cứu phương pháp trích rút các điểm ảnh đặc trưng bất biến cục bộ SIFT [71] cho thấy, trong một cụm các điểm ảnh dị thường nằm cạnh nhau sẽ có ít nhất một điểm ảnh được chọn là điểm đặc trưng. Điều này hoàn toàn phù hợp bởi các điểm ảnh dị thường có màu sắc nổi bật hoặc khác biệt so với những điểm ảnh xung quanh, trong khi đó SIFT tìm những điểm ảnh có giá trị lớn nhất hoặc nhỏ nhất trong một vùng cục bộ làm điểm ảnh đặc trưng. SIFT (Scale-invariant Feature Transform) là thuật toán trích rút đặc trưng cục bộ bất biến trong ảnh do David Lowe [71] công bố vào năm 2004. Đặc trưng được trích chọn trong SIFT là các điểm đặc biệt (keypoint), có bốn giai đoạn chính để thực hiện thuật toán SIFT bao gồm: (1) Phát hiện cực trị không gian tỷ lệ (Scale-space Extrema Detection), (2) Lọc và trích xuất các điểm đặc biệt (Keypoint localization), (3) Gán hướng cho các điểm đặc trưng (Oriented Assignment) và (4) Bộ mô tả điểm đặc trưng (Keypoint Description).

Giai đoạn đầu tiên, thuật toán sẽ áp dụng hàm sai khác Gaussian (DoG - Difference of Gaussian) để tìm ra các điểm có khả năng làm điểm đặc trưng tiềm năng (candidate keypoints), đó là những điểm rất ít phụ thuộc (bất biến) vào sự thu phóng ảnh và xoay ảnh. Hàm DoG được tính theo công thức:

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \quad (2.7)$$

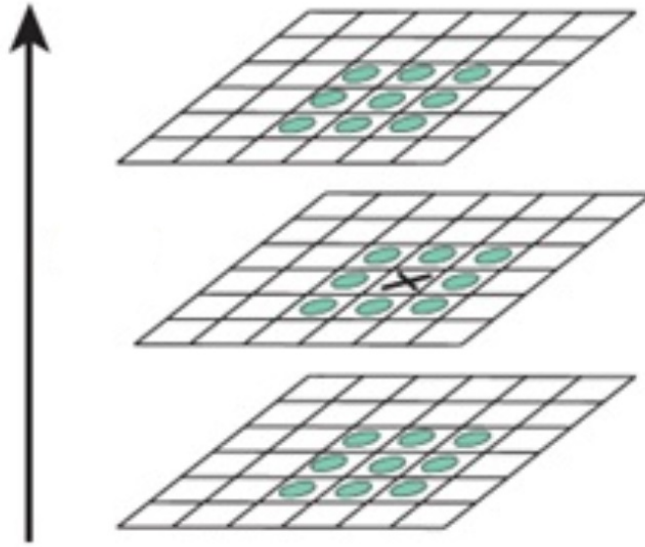
Trong đó, G là toán tử Gaussian Blur, I là ảnh gốc, σ là hệ số làm mịn (có thể xem nó như là số lượng mờ, giá trị càng lớn, độ mờ càng lớn), (x, y) là tọa độ vị trí điểm ảnh đang xét, và toán hạng $*$ là tích chập áp dụng Gaussian Blur G lên ảnh I tại điểm ảnh có tọa độ là (x, y) , k là hằng số bất kỳ (thông thường $k = \sqrt{2}$).

Hàm DoG được tính toán từ sự sai khác giữa 2 không gian đo cạnh nhau



Hình 2.3: Ví dụ minh họa việc thực hiện hàm DoG để tìm những điểm đặc trưng bất biến cao. (a) là ảnh đã được làm mờ với hệ số làm mờ (σ) khác nhau, (b) thể hiện giá trị của hàm DoG trong một quang tám (octave).

của một ảnh với hệ số làm mờ lệch nhau một hằng số k . Việc tính toán hàm sai khác đơn giản chỉ với phép trừ ma trận điểm ảnh, mục đích là tăng khả năng hiển thị của các cạnh và các chi tiết. Kết quả sau khi thực hiện hàm sai khác cho ta thấy, những điểm ảnh dị thường nổi bật hơn hẳn so với những điểm ảnh khác bởi nó có giá trị cao hơn hoặc thấp hơn những điểm ảnh nền (xem ví dụ minh họa tại Hình 2.3).

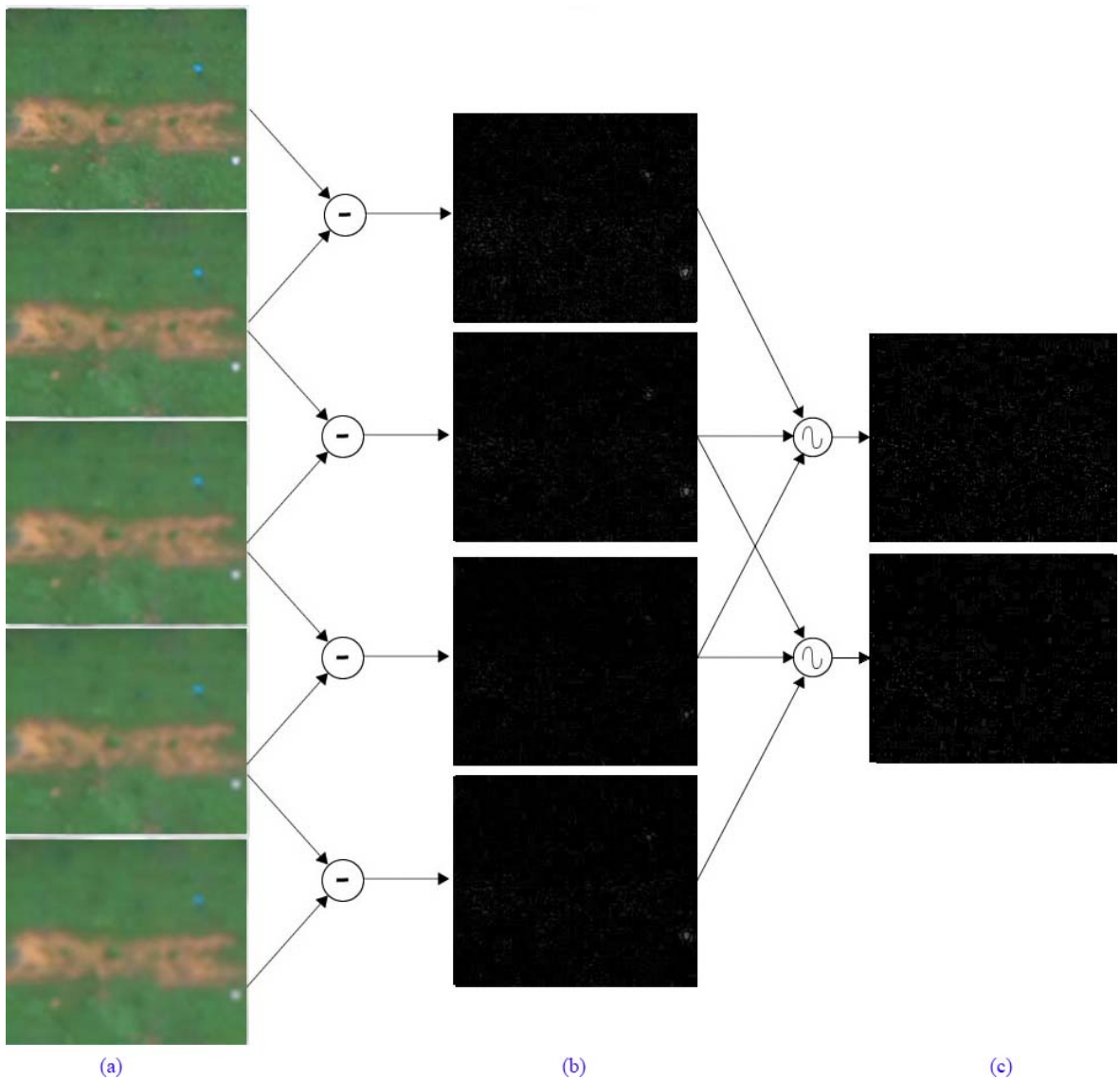


Hình 2.4: Ví dụ minh họa việc tìm các cực trị cục bộ [71]

Tiếp theo, tìm các cực trị trong các lớp kết quả của hàm DoG theo từng miền cục bộ. Cụ thể là tại mỗi điểm trên các lớp kết quả sẽ được so sánh với 8 điểm lân cận trên cùng lớp và 18 điểm lân cận của hai lớp liền kề trên và dưới (xem ví dụ minh họa tại Hình 2.4). Nếu điểm đang xét có giá trị hoặc là lớn nhất hoặc là nhỏ nhất so với 26 điểm đó, nó sẽ là điểm cực trị và có tiềm năng là điểm đặc trưng (xem ví dụ minh họa tại Hình 2.5). Quan sát Hình 2.5, Hình 2.6 cho thấy, khi thực hiện hàm sai khác các điểm ảnh dị thường sẽ cho giá trị cao nhất trong vùng cục bộ nên khi thực hiện tìm giá trị cực bộ, điểm ảnh dị thường sẽ đạt cực trị trong vùng cục bộ nên nó sẽ là điểm ảnh tiềm năng, nằm trong danh sách trong giai đoạn sau.

Giai đoạn thứ hai, vì số lượng các cực trị là rất lớn, một số trong chúng có thể không cần thiết. Ở giai đoạn này sẽ loại bỏ các điểm có độ tương phản kém (nhạy cảm với nhiễu) hoặc tính đặc trưng cục bộ ít hơn các điểm khác hoặc có xu hướng là đường biên đối tượng. Bước thực hiện này gồm 3 công đoạn:

- Phép nội suy lân cận cho vị trí đúng của điểm tiềm năng: phép nội suy



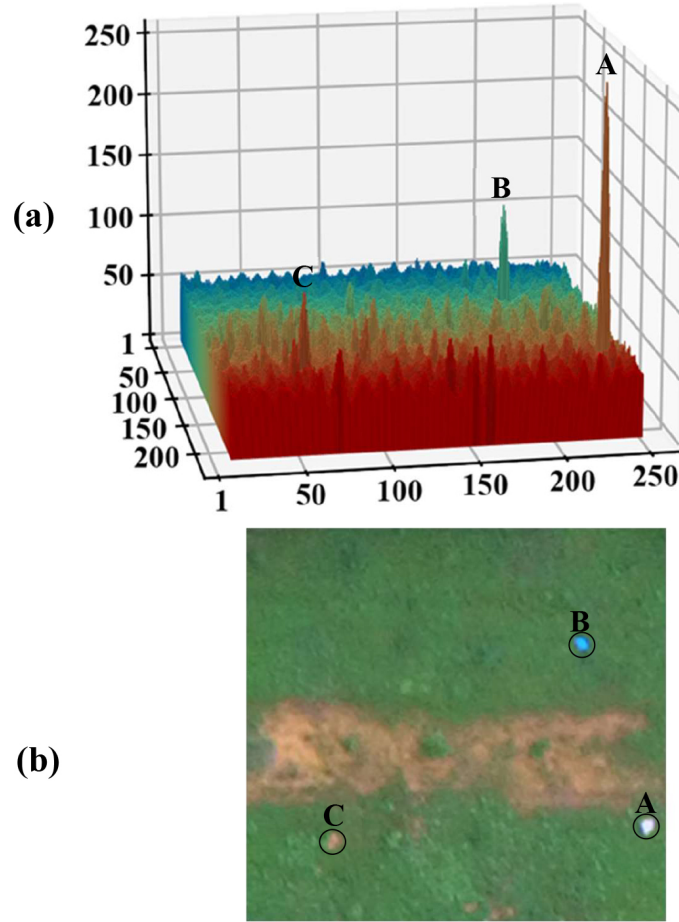
Hình 2.5: Ví dụ minh họa việc tìm các cực trị cục bộ trên ảnh. (a) là ảnh đã được làm mờ với hệ số làm mờ (σ) khác nhau, (b) thể hiện giá trị của hàm DoG trong một quãng tám (octave), (c) thể hiện giá trị của việc tìm các cực trị cục bộ.

lân cận sử dụng khai triển Taylor cho hàm DoG:

$$D(X) = D + \frac{\partial D}{\partial X} X + \frac{1}{2} X^T \frac{\partial^2 D}{\partial X^2} X \quad (2.8)$$

Với $X = (x, y, \sigma)$ là độ dịch so với các điểm lân cận của điểm lấy mẫu. Vùng chứa điểm nổi bật được xác định qua:

$$\hat{X} = -\frac{\partial^2 D^{-1}}{\partial X^2} \frac{\partial D}{\partial X} \quad (2.9)$$



Hình 2.6: Hình (a) là biểu đồ 3D thể hiện giá trị hàm DoG của các điểm ảnh trong không gian tỷ lệ $\sigma=1.6$, hình (b) là ảnh gốc. Các vị trí được đánh dấu (A, B và C) ở hình (a) tương ứng với các vị trí được đánh dấu ở hình (b)

Nếu $\hat{X} > 0.5$, điểm nổi bật nằm ở gần điểm lấy mẫu khác, nếu $\hat{X} < 0.5$, những điểm này sẽ được đưa vào tập hợp mẫu tốt nhất để tiếp tục phân tích. Khi đó, giá trị của hàm $D(\hat{X})$ có thể được tính như sau:

$$D(\hat{X}) = D + \frac{1}{2} \frac{\partial D^T}{\partial X} \hat{X} \quad (2.10)$$

- Loại trừ các điểm có tính tương phản kém: Nếu điểm tiềm năng nào có giá trị $|D(\hat{X})| < 0.03$ thì điểm đó sẽ bị loại bỏ [71].
- Loại bỏ các điểm dư thừa theo biên: Sử dụng giá trị của ma trận Hessian

cấp 2:

$$\mathbf{H} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (2.11)$$

Với α đủ lớn và β đủ nhỏ, có thể tính được tổng các giá trị riêng của \mathbf{H} và giá trị của ma trận Hessian như sau:

$$Tr(\mathbf{H}) = D_{xx} + D_{yy} = \alpha + \beta \quad (2.12)$$

$$Det(\mathbf{H}) = D_{xx}D_{yy} - D_{xy}^2 = \alpha\beta \quad (2.13)$$

với $r = \frac{\alpha}{\beta}$ là tỷ lệ giữa giá trị riêng lớn nhất α và giá trị riêng nhỏ nhất β . Khi đó:

$$R = \frac{Tr(\mathbf{H})^2}{Det(\mathbf{H})} = \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(r\beta + \beta)^2}{r\beta^2} = \frac{(r + 1)^2}{r} \quad (2.14)$$

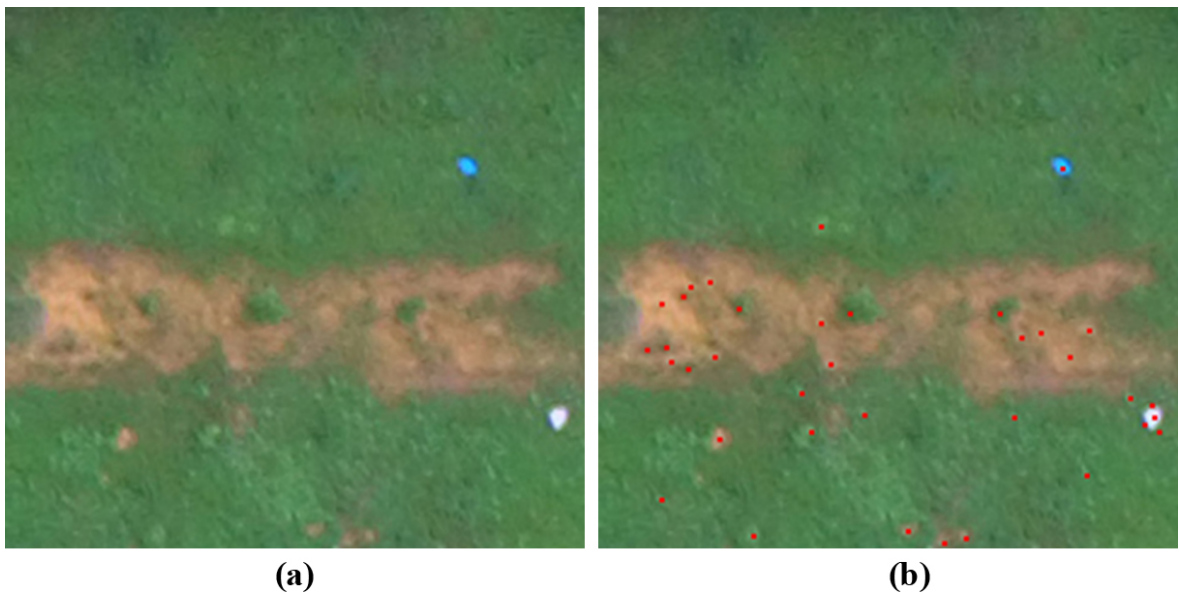
Có thể thấy R phụ thuộc vào giá trị của α và β hơn là các giá trị riêng. Do đó, để kiểm tra tỷ số cong nằm dưới một ngưỡng r nào đó, chỉ cần kiểm tra:

$$\frac{Tr(\mathbf{H})^2}{Det(\mathbf{H})} < \frac{(r + 1)^2}{r} \quad (2.15)$$

Theo kết quả thực nghiệm với nhiều nguồn dữ liệu khác nhau, David G.Lowe đã cho thấy các đặc trưng có tỷ lệ cong r lớn hơn 10 đều bị loại bỏ [71].

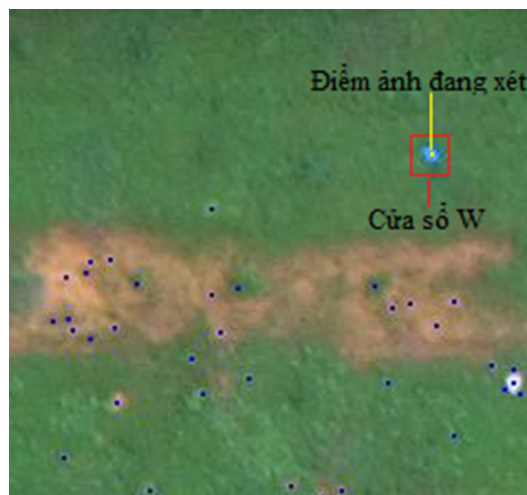
Trong giai đoạn trích xuất các điểm đặc trưng, thuật toán SIFT đã loại bỏ những điểm ảnh tiềm năng có độ tương phản kém hoặc tính đặc trưng cục bộ ít hơn các điểm khác hoặc có xu hướng là đường biên. Do tính chất của các điểm ảnh dị thường, nó là những điểm ảnh có độ tương phản cao hơn những điểm ảnh nền nên nó được chọn là điểm đặc trưng. Tuy nhiên, SIFT sẽ loại bỏ những điểm ảnh nổi bật nếu nó ở gần một điểm đặc trưng khác (nếu $\hat{X} > 0.5$ theo công thức (2.9)) nên SIFT không chọn tất cả các điểm ảnh dị thường trong một cụm điểm ảnh dị thường là những điểm ảnh đặc trưng (xem Hình 2.7).

Như vậy, đến đây đã xác định được vị trí các điểm đặc trưng nên giai đoạn ba và giai đoạn bốn (gán hướng cho các điểm đặc trưng và mô tả điểm đặc



Hình 2.7: Kết quả sau khi thực hiện giai đoạn một và giai đoạn hai của thuật toán SIFT, ảnh (a) là ảnh gốc, các chấm màu đỏ trên ảnh (b) là vị trí các điểm ảnh đặc trưng được tìm thấy trong giai đoạn hai của thuật toán SIFT

trung) không phải thực hiện nữa. Từ sự mô tả của thuật toán SIFT, rõ ràng rằng, SIFT là công cụ phù hợp để tìm tập các điểm ảnh đặc trưng phù hợp với mục đích của ý tưởng tăng độ chính xác phát hiện các điểm ảnh dị thường đã trình bày ở trên.



Hình 2.8: Biểu diễn các điểm ảnh đặc trưng (các chấm màu xanh) và cửa sổ W

Bước thứ hai, thuật toán phát hiện các điểm ảnh dị thường dựa trên KDE cho những điểm ảnh xung quanh các điểm ảnh đặc trưng. Sau khi tìm được tập hợp những điểm đặc trưng, tại mỗi một điểm đặc trưng, chỉ tính toán thuật toán phát hiện các điểm ảnh dị thường dựa trên KDE cho một số điểm ảnh xung quanh điểm đặc trưng thông qua một cửa sổ giới hạn gọi là cửa sổ W , những điểm ảnh nằm ngoài cửa sổ W sẽ được gán luôn cho lớp bình thường mà không phải tính toán. W có tâm là điểm ảnh đặc trưng đang xét, W phải bao hết toàn bộ các điểm ảnh dị thường (nếu có) xung quanh điểm đặc trưng. Nhưng W cũng không được quá lớn, nếu lớn quá sẽ tương đương với việc tính toán Thuật toán 2.1 cho toàn bộ các điểm ảnh (Xem Hình 2.8).

Bước thứ ba, phân loại các điểm ảnh đã được tính toán. Những điểm ảnh được tính toán giá trị dị thường trong bước thứ hai sẽ được phân loại khi so sánh với một ngưỡng phát hiện dị thường η . Những điểm ảnh không được tính toán trong bước thứ hai sẽ được xác định là những điểm ảnh bình thường.

Các bước để tăng độ chính xác phát hiện dị thường của Thuật toán KDE được thể hiện trên Hình 2.1. Đầu tiên, ảnh đầu vào sẽ trải qua bước làm mịn và cũng từ ảnh gốc này được dùng để trích rút các điểm ảnh đặc trưng. Ảnh sau khi làm mịn, kết hợp với tập các điểm ảnh đặc trưng sẽ được sử dụng để tính toán theo Thuật toán 2.1, cuối cùng là phân loại để chọn ra những điểm ảnh dị thường thông qua ngưỡng phát hiện dị thường η .

Thuật toán 2.3 được xây dựng theo ý tưởng đã được trình bày ở trên. Bước đầu tiên, khởi tạo tập C_0 , tập này chứa tất cả các điểm ảnh, khi một điểm ảnh bất kỳ được đưa ra tính toán nó sẽ được loại bỏ khỏi tập C_0 . Bước thứ hai, tìm tập hợp các điểm ảnh đặc trưng theo thuật toán SIFT [71], lưu trữ vào tập P . Bước thứ ba, làm mịn ảnh gốc bằng cách tính tích chập ảnh gốc với hàm Gaussian Blur (Thuật toán 2.2), ảnh sau khi được làm mịn sẽ được dùng để tính toán tìm các điểm ảnh dị thường. Bước thứ tư, xét từng điểm đặc trưng trong P , tính toán giá trị dị thường của các điểm ảnh nằm trong cửa sổ W (lấy điểm đặc trưng đang xét làm tâm của W) nếu như điểm ảnh đó vẫn nằm trong C_0 (tránh phải tính toán nhiều lần cho điểm ảnh đang xét) theo công thức (2.5). Bước cuối cùng là phân loại điểm ảnh dựa trên giá trị dị thường được tính toán tại bước thứ tư và ngưỡng phát hiện dị thường η . Sự khác nhau giữa thuật toán KDE và thuật toán IKDE được thể hiện trên Bảng 2.2.

Thuật toán 2.3: Cải tiến thuật toán phát hiện các điểm ảnh dị thường dựa trên KDE (gọi tắt là Thuật toán IKDE)

input: X : ma trận các điểm ảnh, N : số điểm ảnh, L : số kênh phổ, h : băng thông, η : ngưỡng phát hiện dị thường, σ : hệ số làm mịn, W : cửa sổ.

output: C_1 : tập các điểm ảnh bình thường, C_2 : tập các điểm ảnh dị thường.

```

1  $P \leftarrow$  tập các điểm ảnh đặc trưng (sử dụng thuật toán SIFT);
2  $X' \leftarrow$  GaussianBlur( $X, \sigma$ ); // Sử dụng Thuật toán 2.2
3  $C_0 \leftarrow \{X'_i\}, i=1,2,\dots,N; C_1 \leftarrow \emptyset; C_2 \leftarrow \emptyset;$ 
4 foreach  $p \in P$  do
5      $W \leftarrow$  tập hợp các điểm ảnh nằm trong một cửa sổ lấy  $p$  là tâm;
6     foreach  $X'_i \in W$  do
7         if  $X'_i \in C_0$  then
8              $sum\_ker \leftarrow 0;$ 
9             for  $j \leftarrow 1$  to  $N$  do
10                 $mul \leftarrow 1;$ 
11                for  $d \leftarrow 1$  to  $L$  do
12                     $mul \leftarrow mul \times K \left( \frac{X'[i][d] - X'[j][d]}{h} \right);$ 
13                 $sum\_ker \leftarrow sum\_ker + mul;$ 
14                if  $\frac{sum\_ker}{N} \leq \eta$  then
15                     $C_2 \leftarrow C_2 \cup \{X_i\};$ 
16                else
17                     $C_1 \leftarrow C_1 \cup \{X_i\};$ 
18                 $C_0 \leftarrow C_0 - \{X_i\};$ 
19  $C_1 \leftarrow C_1 \cup C_0;$ 

```

2.4. Tính đúng đắn của thuật toán IKDE

Để phân một điểm ảnh vào lớp "bình thường" C_1 hay lớp "dị thường" C_2 (từ hàng 8 đến hàng 18) phải căn cứ trên hàm mật độ xác suất của điểm ảnh đang xét và được thực hiện như thuật toán gốc (thuật toán KDE).

Vì vậy, cần phải chứng minh tất cả các điểm dị thường đều được xét, hay nói cách khác, các điểm dị thường phải nằm trong các cửa sổ W lấy các điểm

Bảng 2.2: Bảng so sánh sự khác nhau giữa thuật toán KDE và IKDE

Các bước	Thuật toán KDE	Thuật toán IKDE
Tiền xử lý	Không có	- Tìm các điểm ảnh đặc trưng - Làm mịn ảnh
Khởi tạo	$C_1 \leftarrow \emptyset; C_2 \leftarrow \emptyset$	$C_0 \leftarrow \{X_i\}; C_1 \leftarrow \emptyset; C_2 \leftarrow \emptyset$
Vòng lặp đầu	for $i \leftarrow 1$ to N	foreach $p \in P$ (P là tập điểm ảnh đặc trưng)
Vòng lặp thứ hai	for $j \leftarrow 1$ to N	foreach $X_i \in W$ (W là cửa sổ lấy p làm tâm)

đặc trưng làm tâm.

Thật vậy:

- Trong một cụm điểm ảnh dị thường, có ít nhất 1 điểm ảnh được chọn làm điểm ảnh đặc trưng theo Thuật toán SIFT như đã trình bày tại Mục 2.3.

- Thuật toán cải tiến IKDE đã xét tất cả các điểm đặc trưng (hàng 4), tại mỗi một điểm đặc trưng, xét toàn bộ các điểm ảnh nằm trong một cửa sổ W lấy điểm ảnh đặc trưng đó làm tâm (hàng 6), kích thước cửa sổ W dài và rộng gấp 2 lần cụm điểm ảnh dị thường (do các điểm ảnh đặc trưng thường nằm trên biên của cụm ảnh dị thường) nên tất cả các điểm ảnh dị thường đều được xét. Do đó, tất cả các điểm ảnh dị thường được phân vào lớp C_2 nên thuật toán IKDE có tính đúng đắn.

2.5. Đánh giá độ phức tạp tính toán của thuật toán KDE và IKDE

2.5.1. Độ phức tạp tính toán của thuật toán KDE

Thuật toán KDE có ba vòng lặp lồng nhau, vòng lặp thứ nhất và thứ hai cùng thực hiện N lần (N là số điểm ảnh), vòng lặp thứ ba thực hiện L lần (L là số kênh phổ). Độ phức tạp tính toán mỗi lần là $O(1)$ do đó:

$$T(N) = \sum_{i=1}^N \sum_{j=1}^N \sum_{d=1}^L C = C.L.N^2.$$

Như vậy, độ phức tạp tính toán của thuật toán KDE là $O(N^2)$ (do C và L là hằng số).

2.5.2. Độ phức tạp tính toán của thuật toán IKDE

2.5.2.1. Giai đoạn tiền xử lý dữ liệu

- Độ phức tạp tính toán của thuật toán lọc nhiễu là $T_1(N) = O(N)$ [57].

- Độ phức tạp tính toán của thuật toán tìm các điểm ảnh đặc trưng SIFT: do thuật toán IKDE chỉ sử dụng hai giai đoạn trong SIFT nên độ phức tạp tính toán được tính trong hai giai đoạn này. Trong công trình "Analyzing and exploring feature detectors in images" [8], các tác giả đã tính toán độ phức tạp của từng bước như sau:

+ Độ phức tạp tính toán của bước tính hàm DoG là $O(m \times n - \frac{m \times n}{\text{Max}(m, n)})$.

Trong đó m là chiều cao và n là chiều rộng của ảnh, $m \times n = N$. Do đó, độ phức tạp tính toán của bước tính toán hàm DoG là $O(N)$.

+ Độ phức tạp tính toán của bước tìm các điểm cực trị là $O(N)$.

+ Độ phức tạp tính toán giai đoạn lọc và trích xuất các điểm đặc biệt là $O(N)$.

Do đó, độ phức tạp tính toán để đi tìm các điểm ảnh đặc trưng là:

$$T_2(N) = O(N) + O(N) + O(N) = O(N).$$

2.5.2.2. Giai đoạn phân lớp các điểm ảnh

- Như thuật toán KDE, thuật toán IKDE cũng có ba vòng lặp lồng nhau, vòng lặp thứ nhất thực hiện M lần ($M=|P|$ là số điểm ảnh đặc trưng), vòng lặp thứ hai thực hiện N lần và vòng lặp thứ ba thực hiện L lần. Mỗi lần có độ phức tạp tính toán là $O(1)$, do đó:

$$T_3(N) = \sum_{i=1}^M \sum_{j=1}^N \sum_{d=1}^L C = C.L.M.N.$$

Như vậy, độ phức tạp tính toán của thuật toán IKDE là:

$T(N) = \text{Max}(T_1(N), T_2(N), T_3(N)) = C.L.M.N = O(N^2)$ (vì $M \leq N$, C và L là hằng số).

2.5.2.3. Tính hiệu quả của thuật toán

Về lý thuyết, độ phức tạp tính toán của Thuật toán KDE bằng với độ phức tạp tính toán của thuật toán IKDE. Xét trường hợp tồi nhất thì $M=N$ (tất cả các điểm ảnh đều là điểm ảnh đặc trưng). Tuy nhiên, thực tế thì $M \ll N$ bởi giai đoạn lọc và trích xuất điểm ảnh đặc biệt của thuật toán SIFT đã loại bỏ những điểm ảnh đặc trưng gần nhau, là đường biên, những điểm tiềm

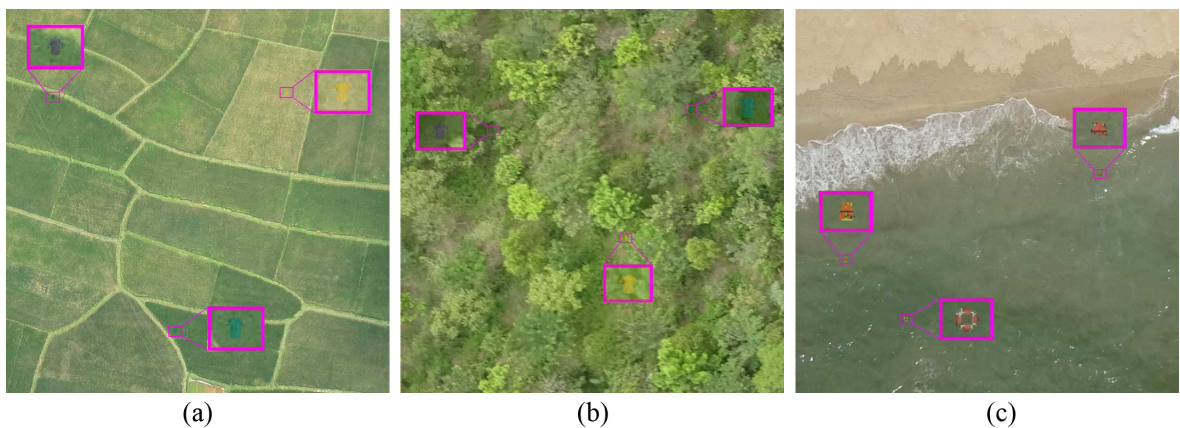
năng có độ tương phản kém,... Do đó, thời gian tính toán của Thuật toán IKDE đã giảm đáng kể so với Thuật toán KDE.

Một ví dụ cụ thể: ảnh ở Hình 2.8 có kích thước là $563 \times 525 = 295.575$ điểm ảnh, trong khi số điểm ảnh đặc trưng được tìm thấy là 36 điểm. Có nghĩa rằng $N \approx 8210 \times M$.

2.6. Thử nghiệm và đánh giá kết quả đề xuất

2.6.1. Dữ liệu thử nghiệm

Dữ liệu ảnh thu chụp trong các tình huống tìm kiếm cứu nạn thực tế ở Việt Nam còn rất hiếm và cơ bản là không được phát hành công khai. Để kiểm nghiệm kết quả nghiên cứu, luận án tiếp cận theo ba cách: thứ nhất, cấy một số mẫu dị thường vào các ảnh chụp từ thiết bị UAV trên ba loại địa hình khác nhau (địa hình đồng bằng, địa hình rừng thưa và địa hình vùng biển). Thứ hai, để tạo tình huống gần sát với thực tế, bố trí một số mẫu dị thường trên không gian quan sát, sử dụng thiết bị UAV thu chụp. Những mẫu được cấy đã được chỉnh sửa kích thước sao cho tương đồng với độ phân giải mặt đất của các ảnh. Tất cả mẫu cấy và mẫu dải trên địa hình đều có kích thước rất nhỏ so với ảnh, tổng số điểm ảnh dị thường trong một ảnh không vượt quá 0.05% tổng số điểm ảnh trong một ảnh. Các ảnh đều được chụp vào ban ngày, trời nắng nhẹ. Thứ ba, sử dụng các ảnh được dùng cho mục đích phân lớp đã được công bố trên thế giới phù hợp với mục đích bài toán phát hiện dị thường trên ảnh.



Hình 2.9: Ba ảnh đã được cấy các mẫu dị thường

Trong trường hợp thứ nhất, cấy mẫu dị thường lên ảnh chụp. Cảnh chụp

ở địa hình đồng bằng (gọi là ảnh (a)) và cảnh chụp ở địa hình rừng thưa (gọi là ảnh (b)) được cấy ba mẫu áo có màu sắc và kích cỡ khác nhau trên mỗi ảnh. Cảnh chụp tại vùng biển ven bờ (gọi là ảnh (c)) được cấy hai mẫu áo phao cứu sinh có màu sắc khác nhau và một phao cứu sinh. Ảnh (a) thể hiện trên Hình 2.9 (a), chụp ở địa hình đồng bằng bằng máy ảnh Canon IXUS 127 HS ở độ cao 190m, độ phân giải mặt đất là 6,3cm/1 điểm ảnh; ảnh (b) và (c) được thể hiện trên Hình 2.9(b) và Hình 2.9(c), chụp ở địa hình rừng thưa và địa hình vùng biển ven bờ bằng máy ảnh Sony DSC-WX220 ở độ cao 200m, độ phân giải mặt đất là 6,4cm/1 điểm ảnh. Mỗi ảnh có kích thước 1000×1000 điểm ảnh và được cấy ba mẫu dị thường khác nhau tùy theo các địa hình khác nhau.

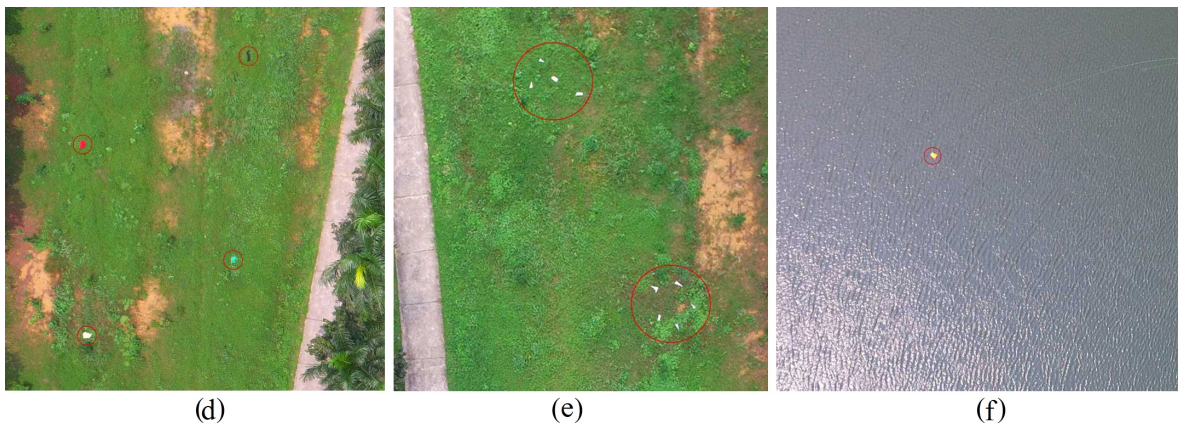
Phương pháp cấy như sau: đối với ảnh (a) và ảnh (b) được cấy 3 mẫu áo có màu sắc khác nhau; mẫu áo thứ nhất (mẫu 1) có kích thước 750x683 mm, mẫu áo thứ hai (mẫu 2) có kích thước 670x787 mm và mẫu áo cuối cùng (mẫu 3) có kích thước 700x750 mm. Khi cấy vào các ảnh, các mẫu này cũng phải thu nhỏ theo tỉ lệ tương đồng với độ phân giải của ảnh, cụ thể:

+ Đối với ảnh (a), mẫu 1 kích thước co lại còn 12x11 pixel; mẫu 2 kích thước co lại là 11x13 pixel; mẫu 3 kích thước co lại là 11x12 pixel;

+ Đối với ảnh (b), mẫu 1 kích thước co lại là 12x11 pixel; mẫu 2 kích thước co lại là 11x12 pixel; mẫu 3 kích thước co lại là 11x12 pixel.

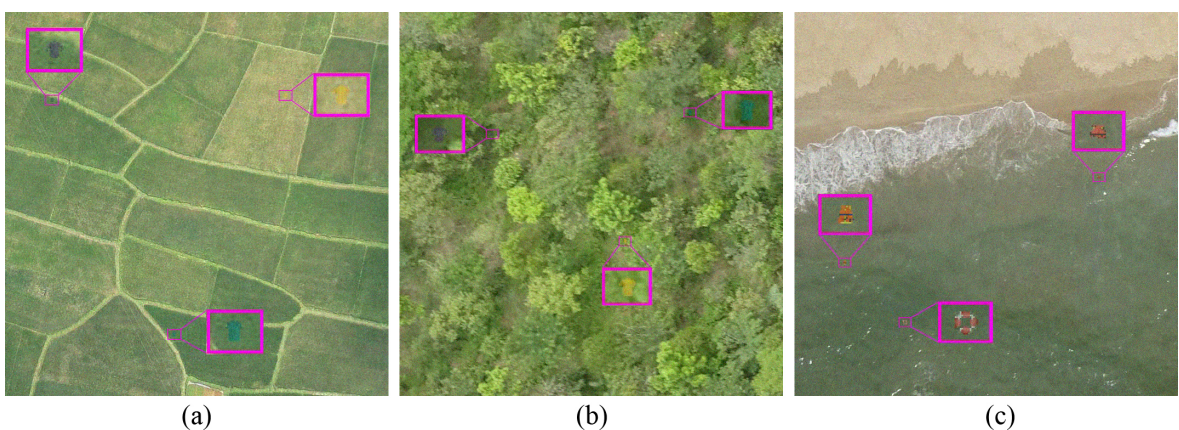
Đối với ảnh (c), đây là ảnh chụp địa hình vùng biển, hai mẫu áo phao cứu sinh có màu sắc khác nhau và một mẫu phao cứu sinh được cấy vào ảnh. mẫu áo phao thứ nhất (mẫu 4) có kích thước 610x530 mm, mẫu áo phao thứ hai (mẫu 5) có kích thước 610x708 mm và mẫu phao cứu sinh (mẫu 6) có kích thước 900x900 mm. Khi cấy vào ảnh (c), các mẫu này cũng phải thu nhỏ theo tỉ lệ tương đồng với độ phân giải của ảnh, cụ thể: mẫu 4 kích thước co lại còn 10x8 pixel; mẫu 5 kích thước co lại là 10x11 pixel; mẫu 6 kích thước co lại là 14x14 pixel.

Ngoài ra, để cả 6 mẫu nhìn giống tự nhiên hơn, các mẫu đã được tăng độ trong suốt lên 50% trước khi cấy vào. Điều này hoàn toàn phù hợp với tự nhiên là các mục tiêu thường bị lẫn vào các đối tượng lân cận. Cả ba ảnh này đã được công bố trong nghiên cứu "Một số kỹ thuật phát hiện dị thường trên ảnh UAV ứng dụng trong công tác tìm kiếm cứu nạn" [CT2].



Hình 2.10: Các mẫu dị thường được bố trí trên địa hình

Trường hợp thứ hai, bố trí mẫu dị thường trên bề mặt địa hình sau đó sử dụng thiết bị UAV chụp. Trong khu vực của Học viện Kỹ thuật Quân sự tại Láng - Hòa Lạc, Thạch Thất, Hà Nội, đã bố trí một số quần, áo có màu sắc và kích cỡ khác nhau; một số miếng nhựa dẻo (mô phỏng là một số mảnh vỡ của phương tiện); bố trí một áo phao cứu sinh trên một hồ nước. Sử dụng thiết bị bay DJI Inspire 1 gắn camera X3 model FC350, bay ở độ cao 254m và có độ phân giải mặt đất là 4cm/1 điểm ảnh. Hình 2.10(d) là hình ảnh chụp trong khu vực bố trí một số quần, áo (gọi là ảnh (d)), Hình 2.10(e) là hình ảnh chụp trong khu vực bố trí một số mảnh nhựa dẻo được xem như là các mảnh vỡ của phương tiện (gọi là ảnh (e)) và Hình 2.10(f) là hình ảnh chụp trong khu vực bố trí áo phao cứu sinh (gọi là ảnh (f)).

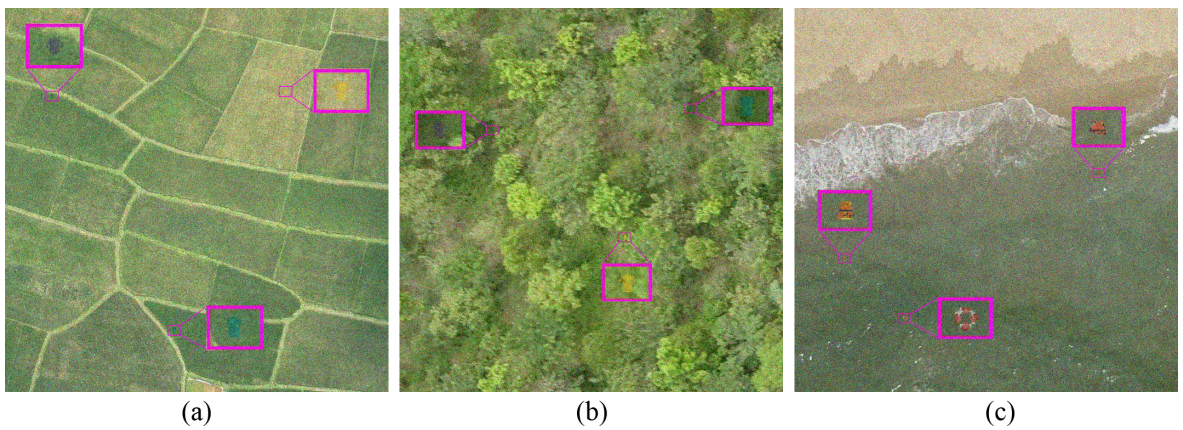


Hình 2.11: Ảnh (a), (b) và (c) đã được thêm 2% nhiễu Gaussian

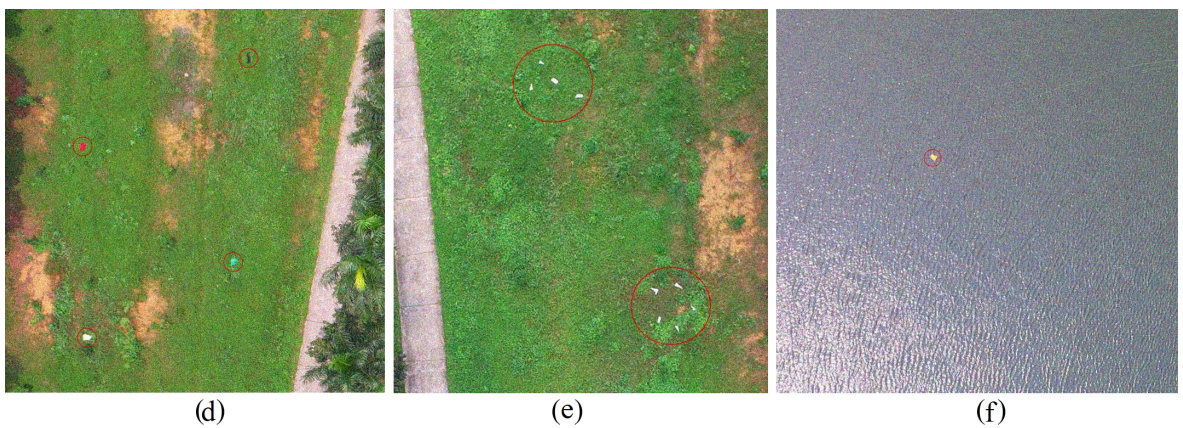
Trường hợp thứ ba, lựa chọn cách tiếp cận sử dụng các ảnh được dùng



Hình 2.12: Ảnh (d), (e) và (f) đã được thêm 2% nhiễu Gaussian



Hình 2.13: Ảnh (a), (b) và (c) đã được thêm 5% nhiễu Gaussian



Hình 2.14: Ảnh (d), (e) và (f) đã được thêm 5% nhiễu Gaussian

cho mục đích phân lớp đã được công bố phù hợp với mục đích bài toán phát hiện dị thường trên ảnh. Đó là sử dụng ảnh 3 kênh phổ và ảnh 8 kênh phổ do

Phòng thí nghiệm khoa học và công nghệ quốc phòng (DSTL)- Vương quốc Anh cung cấp tại địa chỉ <https://www.kaggle.com/c/dstl-satellite-imagery-feature-detection>. Đây là những ảnh được dùng trong cuộc thi "Dstl Satellite Imagery Feature Detection" do DSTL tổ chức [24]. Trong thư viện có 3 loại ảnh, ảnh 3 kênh phổ (ảnh màu RGB), ảnh 8 kênh phổ được chụp trong dải hồng ngoại gần (400 - 1040nm) và ảnh 8 kênh phổ chụp ở dải hồng ngoại ngắn (1195 - 2365nm). Những ảnh này được thu từ bộ cảm biến WorldView 3; phạm vi không gian trong một ảnh là $1km^2$ ($1km \times 1km$). Loại ảnh 3 kênh phổ là ảnh màu RGB với độ phân giải mặt đất là 0,31m, kích thước ảnh là 3396×3349 điểm ảnh; đặc điểm nhận biết, tên file ảnh là: "tên tệp = ImageId.tif". Loại ảnh 8 kênh phổ được chụp từ dải hồng ngoại gần (400 - 1040nm) có độ phân giải mặt đất là 1,24m, kích thước ảnh là 849×837 điểm ảnh; đặc điểm nhận biết, tên file là: "tên tệp = ImageId_M.tif". Loại ảnh 8 kênh phổ được chụp ở dải hồng ngoại ngắn (1195 - 2365nm) có độ phân giải mặt đất là 7,5m, kích thước ảnh là 135×134 điểm ảnh; đặc điểm nhận biết, tên file là: "tên tệp = ImageId_A.tif". Những ảnh có "ImageId" giống nhau là những ảnh được chụp cùng thời gian, cùng địa điểm; chỉ khác ở số kênh phổ và dải phổ thu nhận trên mỗi ảnh.

Trong mỗi ảnh chứa rất nhiều đối tượng khác nhau như đường xá, tòa nhà, xe cộ, nông trại, cây cối, đường nước, v.v. DSTL đã gán nhãn 10 lớp khác nhau:

- Lớp 1: công trình (tòa nhà lớn, khu dân cư, cơ sở lưu trữ nhiên liệu, ...).
- Lớp 2: các công trình nhân tạo.
- Lớp 3: đường.
- Lớp 4: đường mòn; lối đi bộ.
- Lớp 5: cây (rừng, hàng rào, nhóm cây, cây độc lập).
- Lớp 6: cây ngũ cốc, cây trồng theo hàng như: khoai tây, củ cải, ...
- Lớp 7: sông ngòi.
- Lớp 8: ao, hồ.

Bảng 2.3: Danh sách những ảnh được chọn trong thư viện do DSTL cung cấp [24]

TT	Tên file ảnh	TT	Tên file ảnh	TT	Tên file ảnh
1	6010_1_2.tif	11	6110_1_2_A.tif	21	6120_2_0_M.tif
2	6010_1_2_A.tif	12	6110_1_2_M.tif	22	6140_1_2.tif
3	6010_1_2_M.tif	13	6110_3_1.tif	23	6140_1_2_A.tif
4	6090_2_0.tif	14	6110_3_1_A.tif	24	6140_1_2_M.tif
5	6090_2_0_A.tif	15	6110_3_1_M.tif	25	6170_0_4.tiff
6	6090_2_0_M.tif	16	6110_4_0.tif	26	6170_0_4_A.tif
7	6100_2_2.tif	17	6110_4_0_A.tif	27	6170_0_4_M.tif
8	6100_2_2_A.tif	18	6110_4_0_M.tif	28	6170_2_4.tif
9	6100_2_2_M.tif	19	6120_2_0.tif	29	6170_2_4_A.tif
10	6110_1_2.tif	20	6120_2_0_A.tif	30	6170_2_4_M.tif

- Lớp 9: xe cỡ lớn, ví dụ: xe tải, xe buýt, xe container, ...
- Lớp 10: xe cỡ nhỏ, ví dụ: ô tô con, xe van, xe máy, ...

Trong thư viện có 75 ảnh (gồm 25 ảnh 3 kênh phổ, 25 ảnh 8 kênh phổ hồng ngoại gần và 25 ảnh 8 kênh phổ hồng ngoại gần) đã được DSTL gán nhãn. Luận án đã lựa chọn được 30 ảnh (gồm 10 ảnh 3 kênh phổ, 10 ảnh 8 kênh phổ hồng ngoại gần và 10 ảnh 8 kênh phổ hồng ngoại gần) đã được gán nhãn phù hợp với bài toán phát hiện các điểm ảnh dị thường. Những ảnh được lựa chọn là những ảnh được thu chụp ở địa hình bán sa mạc, rừng núi và những nơi dân cư thưa thớt. Những điểm ảnh là các đối tượng như: công trình nhân tạo, phương tiện vận tải (xe cỡ lớn, xe cỡ nhỏ) chúng ta có thể coi nó là những điểm ảnh dị thường bởi chúng thỏa mãn điều kiện: là những điểm ảnh có phổ (hoặc màu sắc) khác biệt nhiều so với những điểm ảnh lân cận, những điểm ảnh này thưa thớt, cụm điểm ảnh tạo thành đối tượng có diện tích chiếm một phần rất nhỏ về mặt không gian và tồn tại với xác suất thấp trong một ảnh. Danh sách những ảnh được chọn thể hiện trên Bảng 2.3.

Để kiểm tra khả năng chống lại nhiễu của kỹ thuật đã đề xuất, lần lượt thêm 2% và 5% nhiễu vào tất 36 ảnh trên (3 ảnh cây mẫu dị thường, 3 ảnh rải mẫu trên địa hình và 30 ảnh của DSTL). Trên Hình 2.11 và Hình 2.12 là sáu ảnh (ảnh (a), (b), (c), (d), (e), (f)) đã được thêm vào 2% nhiễu. Hình 2.13 và Hình 2.14 là sáu ảnh (ảnh (a), (b), (c), (d), (e), (f)) đã được thêm

vào 5% nhiều.

Cách thức chuẩn bị ground-truth: đối với trường hợp thứ nhất, ba ảnh cây các mẫu dị thường, tại các vị trí được cây mẫu dị thường, những điểm ảnh ở đó được gán nhãn "1" (nhãn lớp "dị thường"), những điểm ảnh còn lại được gán nhãn "0" (nhãn lớp "bình thường"). Đối với trường hợp thứ hai, ba ảnh thu chụp trên địa hình có rải các mẫu dị thường ở thực địa; NCS sử dụng phần mềm Adobe Photoshop để tách các cụm điểm ảnh của các mẫu đã rải trên địa hình khỏi nền, những điểm ảnh này được gán nhãn "1", những điểm ảnh còn lại được gán nhãn là "0". Đối với trường hợp thứ ba, như đã trình bày ở trên, những điểm ảnh là các đối tượng như: công trình nhân tạo (thuộc lớp 1), xe cỡ lớn (thuộc lớp 9), xe cỡ nhỏ (thuộc lớp 10) là những điểm ảnh dị thường; căn cứ vào phân lớp do DSTL cung cấp, NCS đã gán nhãn "1" cho những điểm ảnh thuộc các lớp 2, 9 và 10, những điểm ảnh còn lại được gán nhãn là "0".

2.6.2. Phương pháp thử nghiệm

Như vậy, chúng ta có sáu ảnh là những ảnh được chụp ở những địa hình khác nhau, được cây hoặc dải trên địa hình những mẫu vật khác nhau và 30 ảnh trong thư viện của DSTL (tổng số ảnh là 36 ảnh). Từ 36 ảnh gốc, thêm lần lượt 2% và 5% nhiễu vào những ảnh này. Tương ứng với mỗi một ảnh như vậy, sẽ chạy thuật toán KDE và thuật toán IKDE. Như vậy, chúng ta có tất cả 2 thuật toán thực thi trên tập dữ liệu gồm 108 ảnh. Độ chính xác phát hiện dị thường của các thuật toán được tính bằng độ đo AUC.

2.6.3. Kết quả thử nghiệm

Trường hợp đầu tiên, tính toán các thuật toán trên 36 ảnh gốc (ảnh chưa thêm nhiễu vào), độ chính xác phát hiện dị thường của các thuật toán dựa trên độ đo AUC được thể hiện trên Bảng 2.4. Nhìn vào Bảng 2.4 chúng ta thấy, độ chính xác phát hiện các điểm ảnh dị thường của các thuật toán trên 3 ảnh được cây mẫu dị thường (ảnh (a), (b) và (c)), trên 3 ảnh rải các mẫu ở thực địa (ảnh (d), (e) và (f)) và trên 10 ảnh 3 kênh phổ của DSTL cao hơn những ảnh còn lại. Ảnh 8 kênh phổ hồng ngoại gần là những ảnh cho độ chính xác phát hiện của các thuật toán thấp nhất, phần lớn độ chính xác phát hiện các điểm ảnh dị thường của các thuật toán trên loại ảnh này nhỏ hơn 85%, nhiều ảnh độ chính xác phát hiện chỉ ở mức hơn 70%. Quan sát

tập dữ liệu đầu vào cho thấy: từ ảnh (a) đến (f) và 10 ảnh 3 kênh phổ của DSTL có độ phân giải cao (độ phân giải $\leq 1,24m/pixel$), những ảnh còn lại có độ phân giải tương đối thấp ($7,5m/pixel$). Như vậy, chất lượng hình ảnh sẽ ảnh hưởng đến độ chính xác phát hiện các điểm ảnh dị thường của các thuật toán.

Bảng 2.4: Độ đo AUC của thuật toán KDE và thuật toán IKDE trên 36 ảnh gốc (ảnh chưa thêm nhiễu)

STT	Tên file ảnh	Thuật toán KDE	Thuật toán IKDE
1	Ảnh (a)	0,9981	0,9982
2	Ảnh (b)	0,9969	0,9976
3	Ảnh (c)	0,9719	0,9883
4	Ảnh (d)	0,9591	0,9600
5	Ảnh (e)	0,9887	0,9947
6	Ảnh (f)	0,9996	0,9999
7	6010_1_2.tif	0,9034	0,9114
8	6010_1_2_A.tif	0,7192	0,7262
9	6010_1_2_M.tif	0,8656	0,8726
10	6090_2_0.tif	0,8106	0,8196
11	6090_2_0_A.tif	0,9564	0,9664
12	6090_2_0_M.tif	0,9693	0,9773
13	6100_2_2.tif	0,8466	0,8556
14	6100_2_2_A.tif	0,8519	0,8549
15	6100_2_2_M.tif	0,8726	0,8826
16	6110_1_2.tif	0,9272	0,9332
17	6110_1_2_A.tif	0,9480	0,9552
18	6110_1_2_M.tif	0,8940	0,9052
19	6110_3_1.tif	0,9130	0,9220
20	6110_3_1_A.tif	0,7064	0,7074
21	6110_3_1_M.tif	0,8673	0,8753
22	6110_4_0.tif	0,9263	0,9313
23	6110_4_0_A.tif	0,8503	0,8533
24	6110_4_0_M.tif	0,9509	0,9529

Bảng 2.4 – Tiếp theo ở trang trước

STT	Tên file ảnh	Thuật toán KDE	Thuật toán IKDE
25	6120_2_0.tif	0,9082	0,9152
26	6120_2_0_A.tif	0,7758	0,7858
27	6120_2_0_M.tif	0,9278	0,9308
28	6140_1_2.tif	0,9177	0,9217
29	6140_1_2_A.tif	0,8350	0,8420
30	6140_1_2_M.tif	0,9240	0,9302
31	6170_0_4.tiff	0,8121	0,8181
32	6170_0_4_A.tif	0,9882	0,9912
33	6170_0_4_M.tif	0,9759	0,9819
34	6170_2_4.tif	0,8524	0,8574
35	6170_2_4_A.tif	0,9999	0,9999
36	6170_2_4_M.tif	0,9783	0,9803
Trung bình		0,9052	0,9110

Trên Bảng 2.4 và Hình 2.15 cho ta thấy, phương án đề xuất (thuật toán IKDE) có độ chính xác phát hiện các điểm ảnh dị thường tốt hơn thuật toán gốc (thuật toán KDE). Cụ thể, trên 36 ảnh gốc, thuật toán KDE cho độ chính xác phát hiện các điểm ảnh dị thường trung bình tốt hơn thuật toán KDE 0.58%. Như vậy, đề xuất của luận án để loại bỏ một phần nhiều được phân vào lớp "dị thường" của các thuật toán hoàn toàn chính xác; do nhiều trong mỗi ảnh không phải là nhiễu nên không có chênh lệch lớn về độ chính xác phát hiện giữa thuật toán KDE và thuật toán IKDE.

Bảng 2.5: Độ đo AUC của thuật toán KDE và thuật toán IKDE trên 36 ảnh đã được thêm 2% nhiễu vào ảnh gốc

STT	Tên file ảnh	T.toán KDE	T.toán IKDE
1	Ảnh (a) + 2% nhiễu	0,9871	0,9937
2	Ảnh (b) + 2% nhiễu	0,9904	0,9976
3	Ảnh (c) + 2% nhiễu	0,9280	0,9829
4	Ảnh (d) + 2% nhiễu	0,9453	0,9544

Bảng 2.5 – Tiếp theo ở trang trước

STT	Tên file ảnh	T.toán KDE	T.toán IKDE
5	Ảnh (e) + 2% nhiễu	0,9785	0,9956
6	Ảnh (f) + 2% nhiễu	0,9976	0,9999
7	6010_1_2.tif + 2% nhiễu	0,8913	0,9087
8	6010_1_2_A.tif + 2% nhiễu	0,7060	0,7237
9	6010_1_2_M.tif + 2% nhiễu	0,8523	0,8679
10	6090_2_0.tif + 2% nhiễu	0,7975	0,8168
11	6090_2_0_A.tif + 2% nhiễu	0,9427	0,9636
12	6090_2_0_M.tif + 2% nhiễu	0,9536	0,9746
13	6100_2_2.tif + 2% nhiễu	0,8286	0,8531
14	6100_2_2_A.tif + 2% nhiễu	0,8338	0,8520
15	6100_2_2_M.tif + 2% nhiễu	0,8611	0,8800
16	6110_1_2.tif + 2% nhiễu	0,9101	0,9303
17	6110_1_2_A.tif + 2% nhiễu	0,9267	0,9528
18	6110_1_2_M.tif + 2% nhiễu	0,8751	0,9030
19	6110_3_1.tif + 2% nhiễu	0,9006	0,9189
20	6110_3_1_A.tif + 2% nhiễu	0,6940	0,7047
21	6110_3_1_M.tif + 2% nhiễu	0,8560	0,8724
22	6110_4_0.tif + 2% nhiễu	0,9088	0,9286
23	6110_4_0_A.tif + 2% nhiễu	0,8357	0,8507
24	6110_4_0_M.tif + 2% nhiễu	0,9362	0,9500
25	6120_2_0.tif + 2% nhiễu	0,8970	0,9125
26	6120_2_0_A.tif + 2% nhiễu	0,7650	0,7836
27	6120_2_0_M.tif + 2% nhiễu	0,9153	0,9280
28	6140_1_2.tif + 2% nhiễu	0,9005	0,9184
29	6140_1_2_A.tif + 2% nhiễu	0,8208	0,8388
30	6140_1_2_M.tif + 2% nhiễu	0,9075	0,9263
31	6170_0_4.tiff + 2% nhiễu	0,7970	0,8150
32	6170_0_4_A.tif + 2% nhiễu	0,9737	0,9888
33	6170_0_4_M.tif + 2% nhiễu	0,9602	0,9777
34	6170_2_4.tif + 2% nhiễu	0,8390	0,8551
35	6170_2_4_A.tif + 2% nhiễu	0,9849	0,9970

Bảng 2.5 – *Tiếp theo ở trang trước*

STT	Tên file ảnh	T.toán KDE	T.toán IKDE
36	6170_2_4_M.tif + 2% nhiễu	0,9637	0,9785
	Trung bình	0,8905	0,9083

Trường hợp thứ hai, tính toán các thuật toán trên 36 ảnh gốc đã được thêm vào 2% nhiễu, độ chính xác phát hiện dị thường của các thuật toán được thể hiện trên Bảng 2.5. Nhìn vào Bảng 2.5 và Hình 2.16 chúng ta thấy, độ chính xác phát hiện dị thường của các thuật toán đã giảm đi so với tính toán trên 36 ảnh gốc; độ chính xác phát hiện các điểm ảnh dị thường của thuật toán KDE đã giảm trung bình là 1,47%, thuật toán IKDE đã giảm trung bình là 0,27%. Lúc này, thuật toán luận án đề xuất đã có khoảng cách rõ ràng về độ chính xác phát hiện dị thường so với thuật toán gốc; độ chính xác phát hiện dị thường trung bình của thuật toán IKDE đã vượt thuật toán KDE 1,78%.

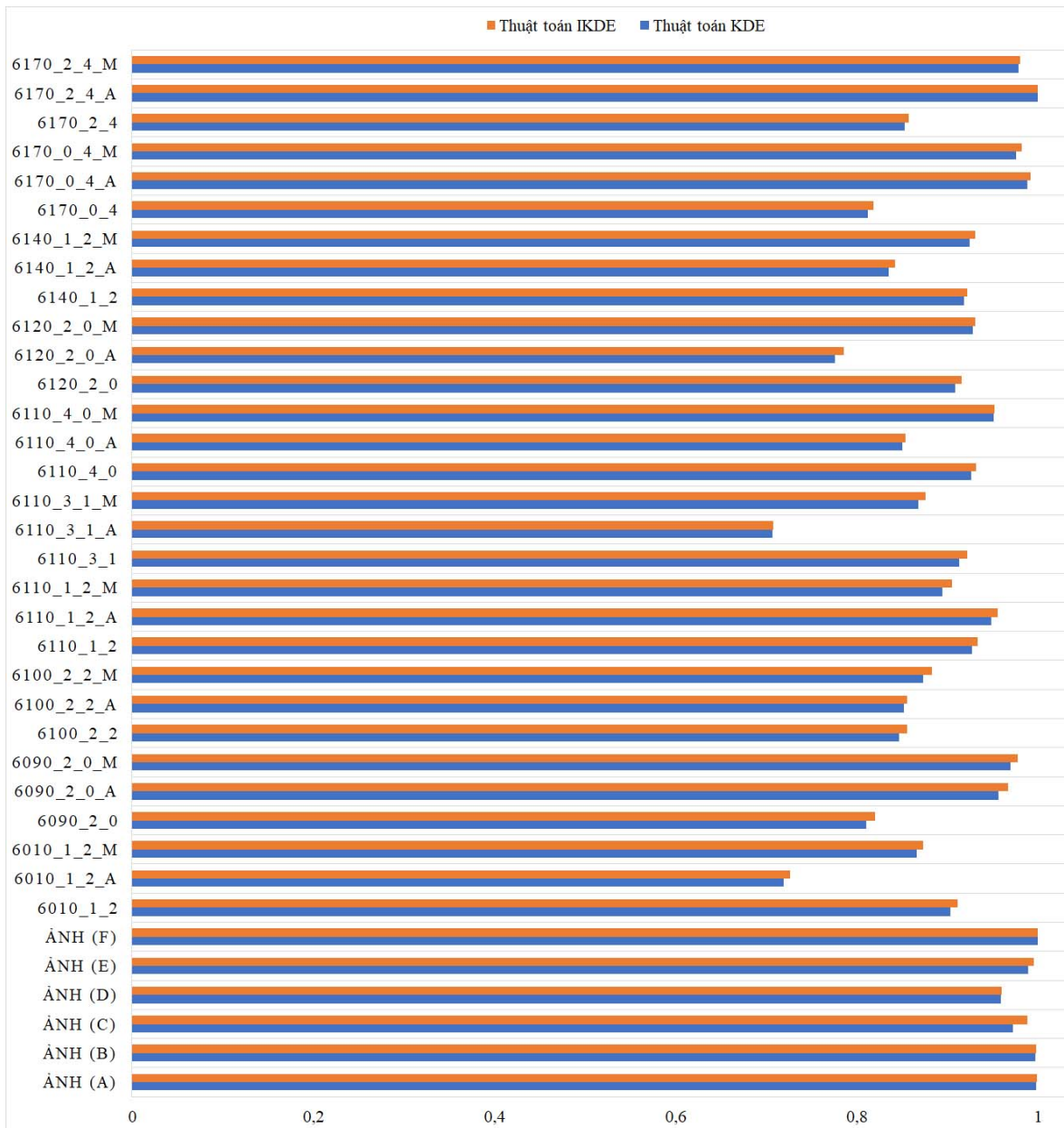
Bảng 2.6: *Độ đo AUC của thuật toán KDE và thuật toán IKDE trên 36 ảnh đã được thêm 5% nhiễu vào ảnh gốc*

STT	Tên file ảnh	T.toán KDE	T.toán IKDE
1	Ảnh (a) + 5% nhiễu	0,9508	0,9834
2	Ảnh (b) + 5% nhiễu	0,9300	0,9956
3	Ảnh (c) + 5% nhiễu	0,8866	0,9727
4	Ảnh (d) + 5% nhiễu	0,9272	0,9532
5	Ảnh (e) + 5% nhiễu	0,9600	0,9922
6	Ảnh (f) + 5% nhiễu	0,9829	0,9994
7	6010_1_2.tif + 5% nhiễu	0,8554	0,9047
8	6010_1_2_A.tif + 5% nhiễu	0,6762	0,7187
9	6010_1_2_M.tif + 5% nhiễu	0,8225	0,8657
10	6090_2_0.tif + 5% nhiễu	0,7674	0,8118
11	6090_2_0_A.tif + 5% nhiễu	0,9113	0,9586
12	6090_2_0_M.tif + 5% nhiễu	0,9236	0,9706
13	6100_2_2.tif + 5% nhiễu	0,7986	0,8481

Bảng 2.6 – Tiếp theo ở trang trước

STT	Tên file ảnh	T.toán KDE	T.toán IKDE
14	6100_2_2_A.tif + 5% nhiễu	0,8038	0,8470
15	6100_2_2_M.tif + 5% nhiễu	0,8311	0,8570
16	6110_1_2.tif + 5% nhiễu	0,8800	0,9263
17	6110_1_2_A.tif + 5% nhiễu	0,9067	0,9478
18	6110_1_2_M.tif + 5% nhiễu	0,8451	0,8980
19	6110_3_1.tif + 5% nhiễu	0,8706	0,9149
20	6110_3_1_A.tif + 5% nhiễu	0,6840	0,7007
21	6110_3_1_M.tif + 5% nhiễu	0,8360	0,8684
22	6110_4_0.tif + 5% nhiễu	0,8788	0,9246
23	6110_4_0_A.tif + 5% nhiễu	0,8157	0,8467
24	6110_4_0_M.tif + 5% nhiễu	0,9062	0,9460
25	6120_2_0.tif + 5% nhiễu	0,8670	0,9085
26	6120_2_0_A.tif + 5% nhiễu	0,7305	0,7786
27	6120_2_0_M.tif + 5% nhiễu	0,8853	0,9240
28	6140_1_2.tif + 5% nhiễu	0,8805	0,9144
29	6140_1_2_A.tif + 5% nhiễu	0,7908	0,8348
30	6140_1_2_M.tif + 5% nhiễu	0,8775	0,9233
31	6170_0_4.tiff + 5% nhiễu	0,7870	0,8110
32	6170_0_4_A.tif + 5% nhiễu	0,9437	0,9838
33	6170_0_4_M.tif + 5% nhiễu	0,9302	0,9750
34	6170_2_4.tif + 5% nhiễu	0,8080	0,8497
35	6170_2_4_A.tif + 5% nhiễu	0,9549	0,9930
36	6170_2_4_M.tif + 5% nhiễu	0,9337	0,9735
Trung bình		0,8622	0,9039

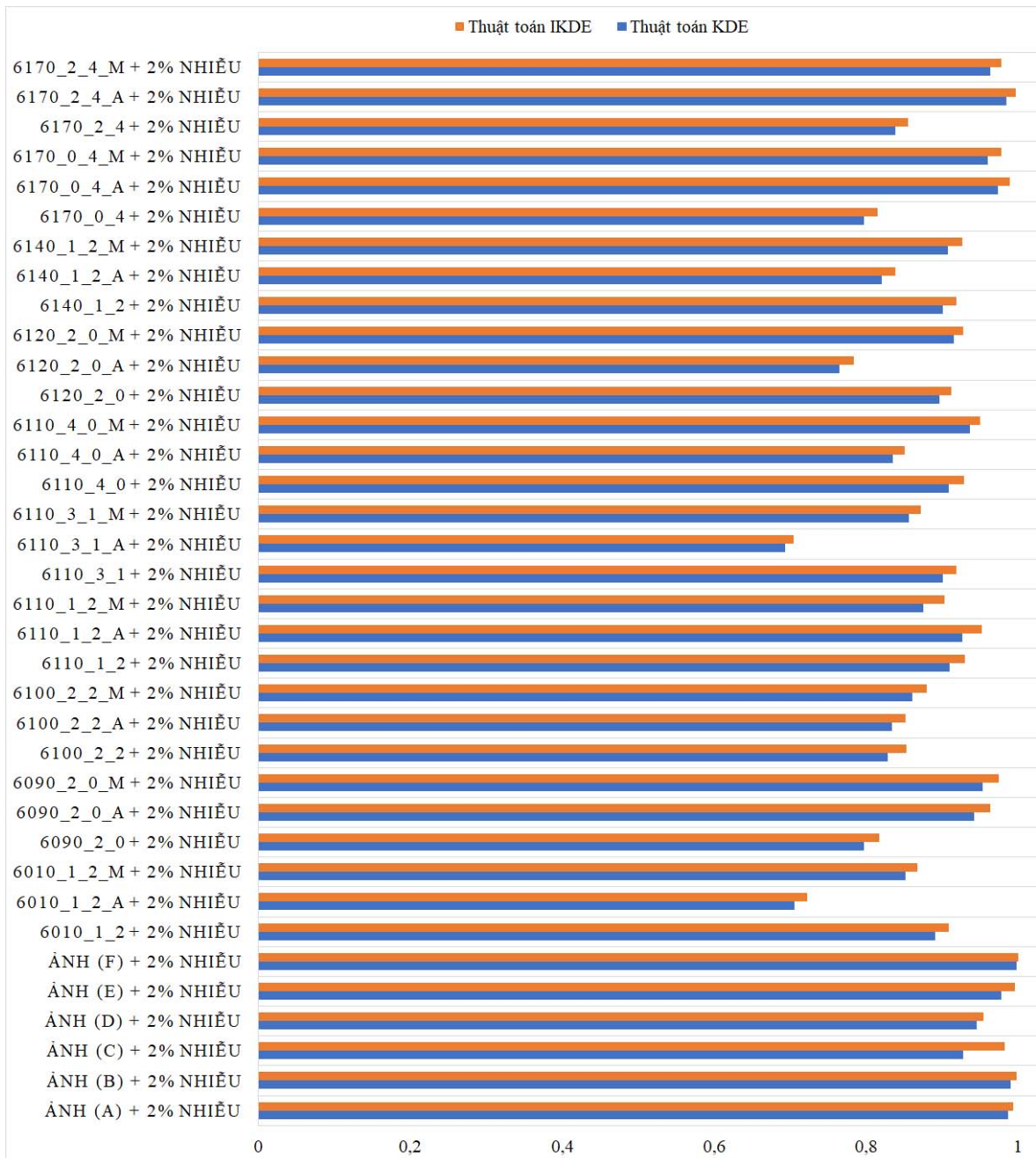
Trường hợp thứ ba, tính toán các thuật toán trên 36 ảnh gốc đã được thêm vào 5% nhiễu, độ chính xác phát hiện dị thường của các thuật toán được thể hiện trên Bảng 2.6. Nhìn vào Bảng 2.6 và Hình 2.17 chúng ta thấy, độ chính xác phát hiện dị thường của các thuật toán đã giảm đi so với tính toán trên 36 ảnh gốc, đặc biệt là thuật toán KDE, độ chính xác phát hiện dị thường đã giảm trung bình là 4,3%. Trong khi đó, thuật toán IKDE chỉ giảm độ chính xác phát hiện các điểm ảnh dị thường trung bình là 0,71%



Hình 2.15: Biểu đồ so sánh độ chính xác phát hiện dị thường của thuật toán KDE (thuật toán gốc) và thuật toán IKDE (thuật toán cải tiến) trên 36 ảnh gốc

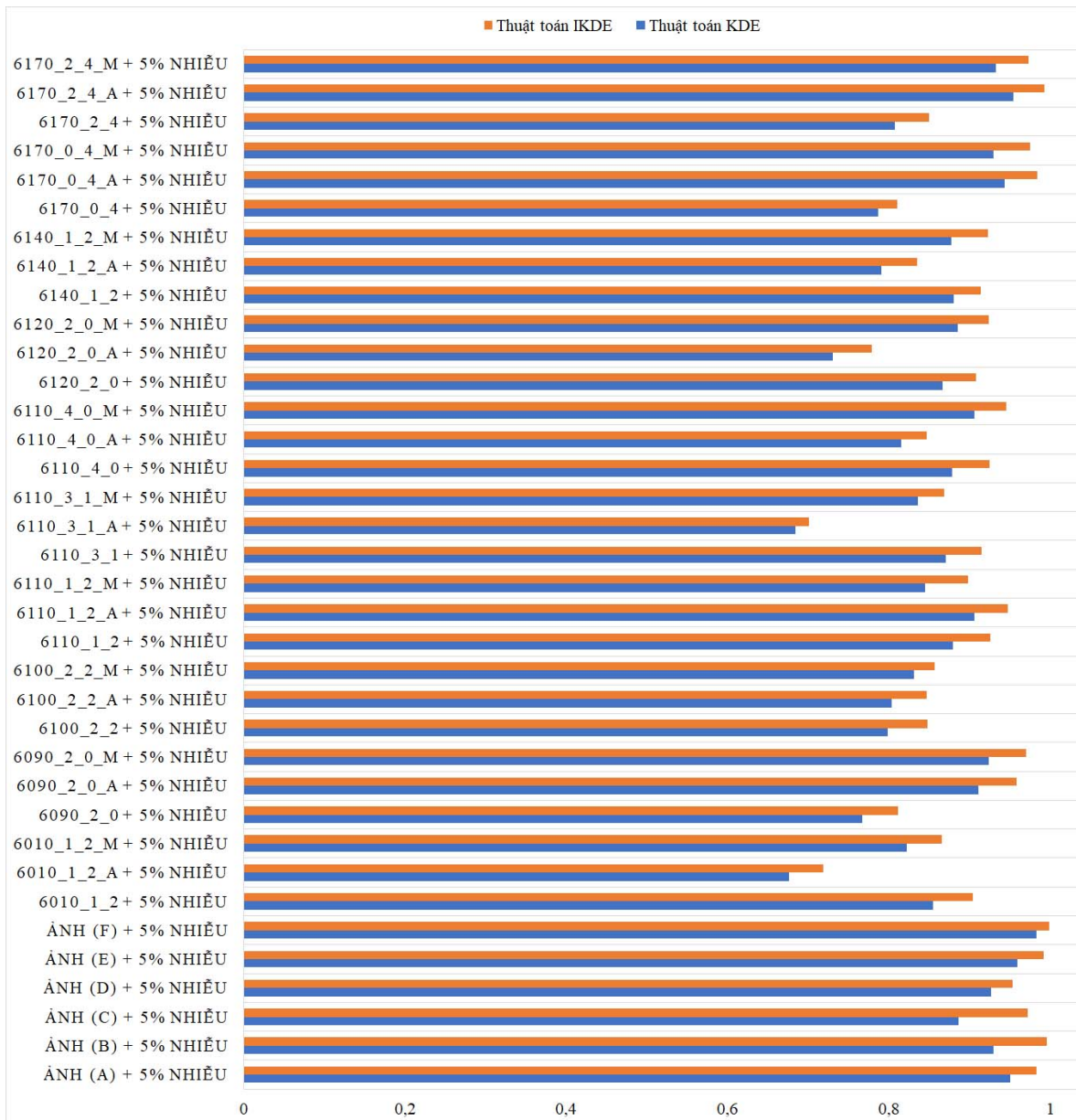
so với tính toán trên 36 ảnh gốc. Thuật toán luận án đã đề xuất đã thực sự có khoảng cách rõ ràng về độ chính xác phát hiện dị thường so với thuật toán gốc; độ chính xác phát hiện dị thường trung bình của thuật toán IKDE đã tốt hơn thuật toán KDE 4,17%.

Như vậy, từ kết quả kiểm nghiệm này chúng ta thấy rằng, thuật toán luận



Hình 2.16: Biểu đồ so sánh độ chính xác phát hiện dị thường của thuật toán KDE (thuật toán gốc) và thuật toán IKDE (thuật toán cải tiến) trên 36 ảnh đã được thêm 2% nhiễu vào ảnh gốc

án đề xuất đã cho độ chính xác phát hiện dị thường trên ảnh tốt hơn thuật toán gốc. Đặc biệt là trong trường hợp ảnh có nhiễu, nhiễu này có thể được thêm vào trong quá trình thu nhận hình ảnh từ thiết bị.



Hình 2.17: Biểu đồ so sánh độ chính xác phát hiện dị thường của thuật toán KDE (thuật toán gốc) và thuật toán IKDE (thuật toán cải tiến) trên 36 ảnh đã được thêm 5% nhiễu vào ảnh gốc

Bảng 2.7: Thời gian thực thi của thuật toán KDE và thuật toán IKDE trên 36 ảnh gốc, thời gian tính bằng giây (s)

STT	Tên file ảnh	Thuật toán 2.1	Thuật toán 2.3
1	Ảnh (a)	11.600	6.595

Bảng 2.7 – *Tiếp theo ở trang trước*

STT	Tên file ảnh	Thuật toán 2.1	Thuật toán 2.3
2	Ảnh (b)	12.600	10.887
3	Ảnh (c)	9.500	2.364
4	Ảnh (d)	13.700	12.473
5	Ảnh (e)	9.200	8.847
6	Ảnh (f)	24.400	20.072
7	6010_1_2.tif	1.819.712	1.277.438
8	6010_1_2_A.tif	3	3
9	6010_1_2_M.tif	23.279	16.063
10	6090_2_0.tif	1.801.515	1.298.892
11	6090_2_0_A.tif	3	3
12	6090_2_0_M.tif	23.512	15.753
13	6100_2_2.tif	1.856.107	1.308.555
14	6100_2_2_A.tif	3	3
15	6100_2_2_M.tif	23.046	16.247
16	6110_1_2.tif	1.856.106	1.319.691
17	6110_1_2_A.tif	4	3
18	6110_1_2_M.tif	23.414	17.092
19	6110_3_1.tif	1.783.318	1.248.323
20	6110_3_1_A.tif	3	3
21	6110_3_1_M.tif	23.745	16.147
22	6110_4_0.tif	1.801.515	1.286.282
23	6110_4_0_A.tif	4	3
24	6110_4_0_M.tif	23.065	16.215
25	6120_2_0.tif	1.837.909	1.328.808
26	6120_2_0_A.tif	3	3
27	6120_2_0_M.tif	23.408	16.409
28	6140_1_2.tif	1.801.815	1.291.901
29	6140_1_2_A.tif	3	3
30	6140_1_2_M.tif	23.742	16.667
31	6170_0_4.tiff	1.783.259	1.250.065
32	6170_0_4_A.tif	3	3

Bảng 2.7 – Tiếp theo ở trang trước

STT	Tên file ảnh	Thuật toán 2.1	Thuật toán 2.3
33	6170_0_4_M.tif	23.146	16.272
34	6170_2_4.tif	1.802.918	1.244.013
35	6170_2_4_A.tif	4	3
36	6170_2_4_M.tif	23.513	15.989
Tổng		18.459.077	13.078.090

Xét về thời gian tính toán, nhìn vào Bảng 2.7 chúng ta thấy, thuật toán KDE có thời gian tính toán chậm nhất; thời gian tính toán của thuật toán IKDE đã giảm được trung bình 29,15% tổng thời gian tính toán trên 36 ảnh so với thuật toán KDE. Lý do giảm được thời gian tính toán của thuật toán IKDE do thuật toán này chỉ tính toán những điểm ảnh xung quanh các điểm ảnh đặc trưng, những điểm ảnh còn lại không cần phải tính toán (mặc định được xem là những ảnh bình thường). Như vậy, thuật toán đề xuất là sự kết hợp của làm mịn ảnh, tìm các điểm ảnh đặc trưng theo phương pháp SIFT sau đó tính toán phát hiện các điểm ảnh dị thường dựa trên KDE vừa tăng độ chính xác phát hiện các điểm ảnh dị thường vừa giảm được thời gian tính toán.

2.7. Kết luận

Một số vấn đề đã được trình bày trong chương này gồm:

- *Thứ nhất*, trình bày thuật toán phát hiện dị thường dựa trên ước lượng mật độ nhân.
- *Thứ hai*, trình bày thuật toán tăng độ chính xác phát hiện dị thường dựa trên ước lượng mật độ nhân. Cài đặt các thuật toán dựa trên ý tưởng đã đề xuất.
- *Thứ ba*, kiểm nghiệm đề xuất, sử dụng 36 ảnh để kiểm tra độ chính xác phát hiện các điểm ảnh dị thường của thuật toán KDE và thuật toán đề xuất (thuật toán IKDE). Cụ thể: thuật toán IKDE có độ chính xác cao hơn thuật toán KDE là 0,58%; thuật toán IKDE giảm được trung bình 29,15% tổng thời gian tính toán so với thuật toán 2.1. Thuật toán luận án đã đề xuất cho thấy độ chính xác trong trường hợp ảnh đầu vào có

nhiều tăng cao, thuật toán đề xuất chỉ giảm 0,71% so với ảnh gốc, ngược lại, thuật toán gốc đã giảm tới 4,17% độ chính xác phát hiện dị thường so với khi thực thi trên ảnh gốc.

Để tránh bỏ sót các điểm ảnh dị thường trong các khu vực nghi ngờ thì W phải đủ lớn; phần lớn, các điểm đặc trưng được tìm thấy bởi SIFT nằm ở đường biên của một cụm ảnh dị thường, do đó chiều dài và chiều rộng của W phải gấp đôi chiều dài và chiều rộng của cụm điểm ảnh dị thường (dự kiến). Tuy nhiên, W cũng không được quá lớn, nếu W lớn quá sẽ tương đương với việc tính toán cho toàn bộ các điểm ảnh. Ngoài ra, SIFT có thể phát hiện thiếu các điểm ảnh đặc trưng là điểm ảnh dị thường thật do độ tương phản của nó kém hoặc tính đặc trưng cục bộ ít hơn các điểm ảnh khác (các điểm ảnh dị thường này không có sự khác biệt lớn so với các điểm ảnh xung quanh nó). Do đó, để tránh bỏ sót các điểm đặc trưng, có thể giảm giá trị ngưỡng của hàm $D(\hat{X})$ (theo David Lowe[71], nếu một điểm ảnh có giá trị $|D(\hat{X})| < 0.03$ tính theo công thức (2.10) sẽ được coi có độ tương phản kém và bị loại).

Do SIFT có thể phát hiện thiếu keypoint nên việc chứng minh tính đúng đắn của giải thuật nhằm loại trừ dương tính giả nhưng có thể bỏ sót là một việc nên xem xét kỹ lưỡng. NCS nên có những nhận định để việc bắt nhầm còn hơn bỏ sót, và vì thế, phân tích kỹ hơn về các giá trị ngưỡng nhằm giữ lại các keypoint dù “yếu” (theo các tiêu chí hiện tại của SIFT) nhưng có thể vẫn có chứa dị thường.

Độ chính xác phát hiện dị thường của thuật toán IKDE được thể hiện trên Bảng 2.4, 2.5 và 2.6 cho thấy, thuật toán này phát hiện các điểm ảnh dị thường tốt hơn thuật toán KDE. Tuy nhiên, thời gian tính toán của thuật toán IKDE chưa đáp ứng được với yêu cầu thực tế nên rất khó để triển khai, ứng dụng, đặc biệt là trong công tác tìm kiếm cứu nạn. Chương sau sẽ trình bày giải pháp tăng tốc độ phát hiện các điểm ảnh dị thường của thuật toán IKDE mà không làm thay đổi độ chính xác phát hiện dị thường của thuật toán này.

Chương 3

KỸ THUẬT TĂNG TỐC ĐỘ PHÁT HIỆN DI THƯỜNG TRÊN ẢNH VIỄN THÁM QUANG HỌC

Chương này trình bày giải pháp tăng tốc độ tính toán của thuật toán phát hiện các điểm ảnh dị thường dựa trên ước lượng mật độ nhân đã đề xuất trong Chương 2 (thuật toán IKDE) thông qua việc thực hiện thuật toán theo bốn bước. Bước đầu tiên là nhóm các điểm ảnh có cùng giá trị, bước thứ hai là sắp xếp các nhóm điểm ảnh này tăng dần theo chiều dữ liệu thứ nhất. Bước thứ ba là tính toán hàm mật độ xác suất của các nhóm điểm ảnh xung quanh các điểm ảnh đặc trưng bằng một trong ba kỹ thuật: sử dụng cây kd-tree để quản lý và tìm kiếm nhanh những nhóm điểm ảnh cần tham gia tính toán; sử dụng phương pháp tính toán đa luồng trên CPU hoặc song song trên GPU. Bước cuối cùng phân loại các điểm ảnh dựa theo giá trị mật độ xác suất đã được tính toán ở bước thứ ba. Cài đặt thuật toán, thử nghiệm, đánh giá giải pháp đề xuất và kết luận chương. Nghiên cứu này được công bố trong [CT4], [CT5].

3.1. Giới thiệu chung

Như đã đề cập trong phần mở đầu, tốc độ tìm kiếm là yếu tố quan trọng hàng đầu trong công tác tìm kiếm cứu nạn, nghiên cứu về khả năng sống sót của con người sau tai nạn máy bay cho thấy rằng, người bị nạn có khả năng sống sót nhỏ hơn 10% nếu việc cứu hộ bị trễ quá 2 ngày, và tỷ lệ sống sót lên tới 60% nếu việc cứu hộ được thực hiện kịp thời trong vòng 8 tiếng. Sự khẩn cấp tương tự cũng được áp dụng trong các tình huống cấp cứu hàng hải hay trên đất liền. Vì vậy, thời gian tìm ra người bị nạn và tổ chức giải cứu kịp thời là nhân tố quyết định đến sự sống còn của người bị nạn.

Công tác tìm kiếm cứu nạn là công việc khó khăn và phức tạp, đòi hỏi sự khẩn trương về thời gian tìm kiếm. Các công cụ hỗ trợ ra đời với mục đích nhằm ngày càng nâng cao hiệu quả của công tác này. Ứng dụng ảnh viễn thám vào công tác này là một trong những công cụ đó, các thuật toán được sử dụng cũng phải đáp ứng được yêu cầu về thời gian thực thi. Nếu một thuật toán có

độ chính xác phát hiện mục tiêu cần tìm kiếm rất tốt nhưng thời gian thực thi chậm cũng sẽ khó để được áp dụng vào thực tế. Như đã trình bày trong Chương 2, thuật toán phát hiện các điểm ảnh dị thường dựa trên KDE rất hiệu quả, dễ dàng cài đặt, độ chính xác phát hiện các điểm ảnh dị thường tốt hơn thuật toán RXD. Tuy nhiên, độ phức tạp tính toán của thuật toán này là $O(N^2)$, N là số điểm ảnh. Trong phần thực nghiệm của nghiên cứu [31], tác giả sử dụng 1.000.000 điểm dữ liệu để kiểm nghiệm đã phải tính toán gần hết 10^6 giây (hơn 11 ngày). Trong phần thực nghiệm của nghiên cứu [CT4], một ảnh màu 3 kênh RGB, kích thước 3396×3349 điểm ảnh tốn 21 ngày để tính toán, dẫn đến việc ứng dụng thuật toán này vào trong công tác tìm kiếm cứu nạn là không khả thi. Mặc dù thời gian thực thi của thuật toán IKDE có thể giảm tới 29.15% thời gian tính toán nhưng thời gian thực thi trên các ảnh có kích thước lớn vẫn còn quá cao nên vẫn rất khó để áp dụng ngoài thực tế.

Để áp dụng thuật toán IKDE nhằm phát hiện các điểm ảnh dị thường phục vụ trong công tác tìm kiếm cứu nạn, luận án đề xuất giải pháp có thể giảm thời gian tính toán tới hơn 99% so với thuật toán IKDE và có thể ứng dụng ngoài thực tế.

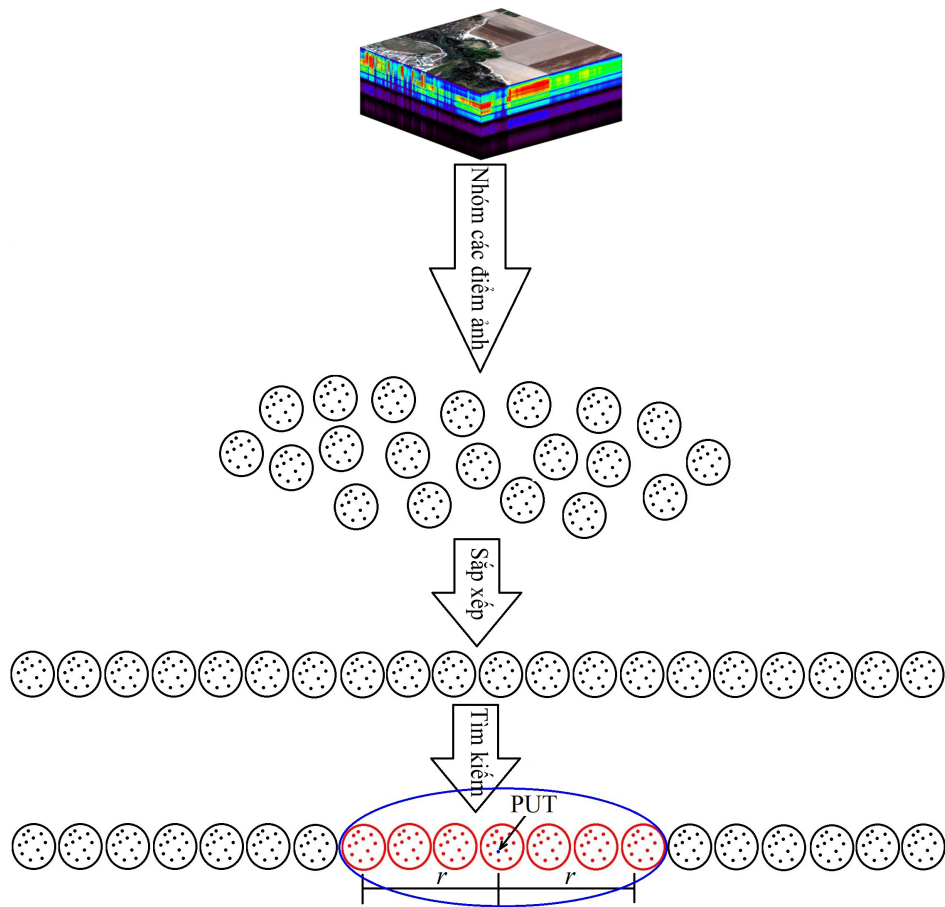
3.2. Tăng tốc độ tính toán của thuật toán IKDE

Trọng tâm tính toán của phương pháp phát hiện dị thường dựa trên KDE là việc phải tính toán KDE, độ phức tạp tính toán là $O(N^2)$. Để tăng tốc độ tính toán, giảm thời gian xử lý, một số kỹ thuật gần đúng đã được đề xuất. Đầu tiên, đó là đề xuất của Silverman trong nghiên cứu [106], tác giả đã sử dụng biến đổi Fourier nhanh (FFT - Fast Fourier Transform) để ước lượng mật độ xác suất. Nó làm giảm đáng kể yêu cầu tính toán của phương pháp ước tính mật độ, đã giảm độ phức tạp tính toán từ $O(N^2)$ xuống còn $O(N \log N)$. Một phương pháp khác là áp dụng biến đổi Gaussian nhanh (FGT - Fast Gauss Transform) được Elgammal và các cộng sự đề xuất trong nghiên cứu [31], phương pháp này đã giảm độ phức tạp tính toán từ $O(NM)$ xuống còn $O(N + M)$. Trong đó, N là kích thước dữ liệu, M là số lượng mục tiêu cần tính PDF. FFT và FGT chỉ phù hợp với số chiều dữ liệu thấp, với số chiều của dữ liệu cao, sự phức tạp của tính toán và lưu trữ của FGT tăng theo cấp số nhân theo số chiều của dữ liệu [129]. Yang và các cộng sự [129] đã cải tiến FGT gọi là IFGT (Improved Fast Gauss Transform), IFGT đã tốt

hơn FGT với số chiều dữ liệu lớn hơn. Tuy nhiên, IFGT cũng chỉ hiệu quả với số chiều của dữ liệu nhỏ hơn 12 [129], với số chiều lớn như ảnh đa phổ có số kênh phổ lớn hơn 12 hoặc ảnh siêu phổ, IFGT có độ phức tạp tính toán và lưu trữ vô cùng lớn, yêu cầu bộ nhớ dùng để lưu trữ vượt quá sự quản lý của hệ điều hành hệ thống. Mặc dù cả hai phương pháp FFT và FGT, IFGT đã giảm độ phức tạp tính toán PDF nhưng do đây là hai kỹ thuật tính toán gần đúng nên nó đã làm giảm độ chính xác phát hiện dị thường của thuật toán, đây là điều không mong muốn trong nghiên cứu này.

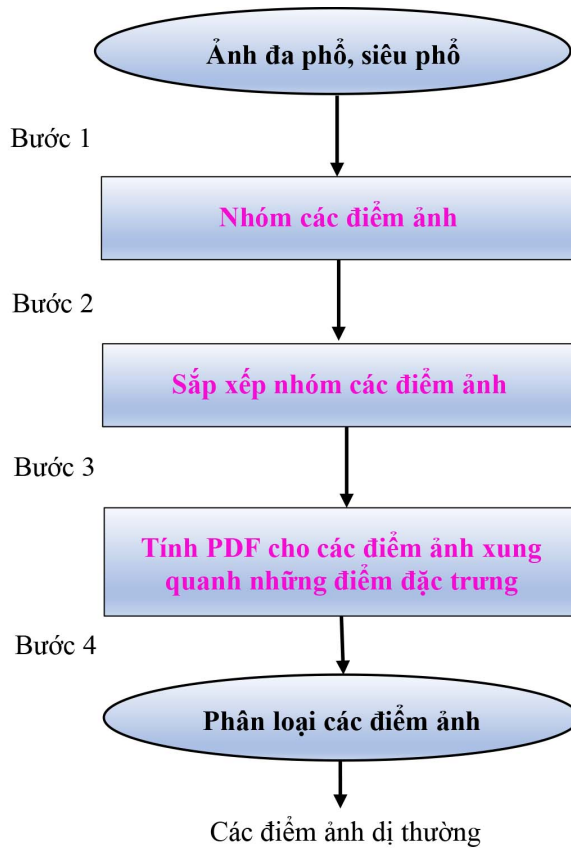
Ngoài ra, một cách tiếp cận khác để giảm thời gian tính toán là song song hóa quá trình ước tính mật độ hàm nhân trên mạng máy tính, trên CPU hoặc GPU. Trong nghiên cứu [73], Lukasik đã đề xuất sử dụng thư viện MPI để song song hóa việc ước lượng hàm mật độ xác suất. Năm 2013 Michailidis và các cộng sự [86] đã song song hóa ước lượng mật độ hàm nhân trên các khung lập trình khác nhau như Pthreads, OpenMP, Intel Cilk ++, Intel TBB và SWARM. Cũng trong năm 2013 Michailidis và các cộng sự [87] tiếp tục song song hóa ước lượng hàm mật độ nhân trên nền tảng GPU CUDA. Ưu điểm của các phương pháp này là không làm thay đổi độ chính xác phát hiện dị thường của thuật toán. Tuy nhiên, độ phức tạp tính toán PDF không thay đổi vẫn là $O(N^2)$, thời gian tính toán giảm do các phương pháp này đã chia tổng khối lượng công việc làm nhiều phần và tính toán đồng thời.

Qua quá trình nghiên cứu cho thấy rằng, khi tính hàm mật độ xác suất theo công thức (2.4) hoặc (2.5), việc tìm những điểm ảnh trong phạm vi băng thông để $K(u) \neq 0$ tiêu tốn rất nhiều thời gian, những điểm ảnh nằm ngoài băng thông sẽ làm cho $K(u) = 0$ và như vậy sẽ không làm thay đổi giá trị của $\hat{f}(\cdot)$ trong công thức (2.4) hoặc (2.5) nên chúng ta sẽ bỏ qua những điểm ảnh này. Vì vậy, phương pháp đầu tiên nghĩ đến là làm thế nào để giảm dữ liệu tham gia tính toán mà không làm thay đổi độ chính xác phát hiện các điểm ảnh dị thường của thuật toán IKDE. Lợi dụng sự giống nhau trên miền giá trị của nhiều điểm ảnh trên cùng một ảnh, đặc biệt là trên ảnh màu RGB, sẽ nhóm những điểm ảnh này thành một nhóm. Trong tự nhiên, lớp phủ thực địa luôn có tính chất phân lớp đối tượng, lớp phủ càng đồng nhất số lượng nhóm càng ít. Bước tiếp theo, các nhóm điểm ảnh sẽ được sắp xếp, với mục đích để chọn nhanh các nhóm điểm ảnh có thể chứa các điểm ảnh thỏa mãn điều kiện $K(u) \neq 0$. Trên Hình 3.1 thể hiện việc lựa chọn các



Hình 3.1: Ý tưởng giảm dữ liệu cần tính toán việc ước lượng PDF.

nhóm điểm ảnh như vậy, chỉ những nhóm điểm ảnh màu đỏ nằm trong hình ellipse màu xanh là những nhóm điểm ảnh được lựa chọn để tham gia việc ước lượng PDF cho PUT, những nhóm khác bị loại bỏ. Do các nhóm điểm ảnh đã được sắp xếp nên lấy PUT làm trung tâm, lần lượt lựa chọn những nhóm điểm ảnh ở bên trái và bên phải của nhóm điểm ảnh chứa PUT nếu như khoảng cách của các nhóm đó đến PUT (tính khoảng cách theo những chiều dữ liệu được nhóm) nhỏ hơn hoặc bằng r ($r = h \times \epsilon$, h là băng thông, ϵ hằng số điều kiện để $K(u) \neq 0$). Như vậy, bước phân nhóm cơ bản làm giảm đáng kể số lượng dữ liệu cần tính toán, bước sắp xếp các nhóm điểm ảnh sẽ chọn nhanh được những nhóm điểm ảnh trong đó có các điểm ảnh có thể làm thỏa mãn điều kiện $K(u) \neq 0$. Những nhóm điểm ảnh không được lựa chọn để tham gia tính toán chúng ta không cần quan tâm đến nó nên đã giảm được khối lượng dữ liệu cần tính toán, giảm được thời gian tính toán.



Hình 3.2: Các bước để tăng tốc độ tính toán của thuật toán phát hiện các điểm ảnh dị thường dựa trên KDE

Sau khi những nhóm điểm ảnh được lựa chọn để tham gia ước lượng PDF cho PUT, chúng ta có thể dùng một trong ba phương pháp: sử dụng cây kd-tree để tìm nhanh những điểm ảnh thỏa mãn điều kiện $K(u) \neq 0$; tính toán đa luồng trên CPU giống như đề xuất của Michailidis và các cộng sự trong [86] hoặc tính toán song song trên GPU CUDA như trong đề xuất của Michailidis và các cộng sự [87].

Như vậy, để tăng tốc độ phát hiện các điểm ảnh dị thường trên ảnh đa phổ, siêu phổ luận án đề xuất tính toán qua bốn bước (xem Hình 3.2), bước đầu tiên là nhóm các điểm ảnh có cùng giá trị, bước thứ hai là sắp xếp nhóm các điểm ảnh (theo chiều dữ liệu thứ nhất được nhóm), bước ba là ước lượng PDF của PUT bằng cách sử dụng cây kd-tree để hỗ trợ tìm nhanh những điểm ảnh thỏa mãn điều kiện $K(u) \neq 0$, sử dụng phương pháp tính toán đa luồng trên CPU, song song trên GPU CUDA. Bước cuối cùng là phân lớp các điểm ảnh dựa vào giá trị đã được tính toán ở bước thứ ba theo công

thức (2.3).

3.3. Kỹ thuật nhóm các điểm ảnh và sắp xếp

Nhóm trực tiếp các điểm ảnh có cùng giá trị trên ảnh đa phổ, siêu phổ mất rất nhiều thời gian, độ phức tạp tính toán là $O(MN)$. Trong đó, M là số nhóm các điểm ảnh có phổ trùng nhau, N là số điểm ảnh. Trong trường hợp xấu nhất $M=N$, có nghĩa rằng không có điểm ảnh nào trong ảnh trùng nhau, khi đó độ phức tạp tính toán là $O(N^2)$. Do giá trị của các điểm ảnh thông thường được lưu trữ ở dạng số nguyên và giá trị nhỏ hơn hoặc bằng 12bit nên để giảm độ phức tạp tính toán, sẽ xây dựng một mảng hai chiều, gọi là mảng \mathbf{A} . Kích thước chiều thứ nhất của \mathbf{A} là số lượng tổ hợp màu của 3 kênh phổ đầu tiên (ngoài ảnh màu RGB chỉ có 3 kênh, ảnh đa phổ có nhiều hơn 3 kênh phổ và ảnh siêu phổ cũng chỉ chọn 3 kênh đầu tiên để tính tổ hợp), ví dụ ảnh màu RGB 24 bit, chiều thứ nhất của mảng \mathbf{A} sẽ có kích thước là 16.777.216 phần tử. Qua kiểm nghiệm thực tế cho thấy rằng, chỉ nên sử dụng tối đa 3 kênh phổ để tính toán tổ hợp màu, nếu chọn số kênh phổ lớn hơn 3, số nhóm được tạo ra sẽ rất lớn dẫn đến giảm hiệu quả về thời gian tính toán. Hơn nữa, nếu sử dụng số kênh ảnh lớn hơn 3 để tổ hợp sẽ tạo ra một số vô cùng lớn dẫn đến tình trạng hệ điều hành không quản lý được mảng \mathbf{A} (đối với hệ điều hành Windows 10 64bit, công cụ lập trình là C++, kích thước chiều thứ nhất của mảng \mathbf{A} không nên vượt quá 2.147.483.647). Kích thước chiều thứ hai của \mathbf{A} sẽ cấp phát linh động để lưu trữ vị trí không gian trên ảnh của các điểm ảnh thuộc về nhóm đó.

Để xác định vị trí của các điểm ảnh trong mảng \mathbf{A} , thực hiện tính toán giá trị tổ hợp 3 kênh phổ theo công thức (3.1):

$$index_i = \sum_{d=1}^3 (m+1)^{3-d} \times x_{id}, i = 1, 2, \dots, N \quad (3.1)$$

trong đó,

$$m = \max x_{id}, d = 1, 2, 3; i = 1, 2, \dots, N \quad (3.2)$$

là giá trị lớn nhất của ba kênh phổ đầu tiên của tất cả các điểm ảnh (thông thường, đối với ảnh lưu trữ 8bit/1 kênh $m = 256$, 10bit/kênh $m = 1024, \dots$). $index_i$ tương ứng với chỉ số trên chiều thứ nhất của mảng \mathbf{A} . Chiều thứ hai của mảng \mathbf{A} sẽ tự động tăng thêm một ô nhớ để lưu trữ vị trí không gian

của điểm ảnh đó. Sau khi phân hết các điểm ảnh vào mảng \mathbf{A} , một số phần tử của mảng \mathbf{A} không có bất cứ điểm ảnh nào, chúng ta sẽ loại bỏ các phần tử này và như vậy sẽ được các nhóm chứa các điểm ảnh. Thuật toán nhóm được trình bày cụ thể tại Thuật toán 3.1.

Thuật toán 3.1: Nhóm các điểm ảnh (CreateGroupPixels)

input: X : ma trận của các điểm ảnh, N : số điểm ảnh.

output: A : mảng chứa các nhóm điểm ảnh.

```

1  $m \leftarrow 0$ 
2 for  $i \leftarrow 1$  to  $N$  do
3   for  $d \leftarrow 1$  to 3 do
4     if  $m < X[i][d]$  then
5        $m = X[i][d]$ ;
6  $A \leftarrow \text{sizeof}((m + 1)^3)$ ; //Khởi tạo mảng A (khởi tạo chiều thứ nhất
   của mảng A có kích thước bằng  $\text{size}A$ )
7 for  $i \leftarrow 1$  to  $N$  do
8    $\text{index} \leftarrow 0$ ;
9   for  $d \leftarrow 1$  to 3 do
10     $\text{index} \leftarrow \text{index} + (m + 1)^{3-d} \times X[i][d]$ ;
11     $\mathbf{A}[\text{index}] \leftarrow \mathbf{A}[\text{index}] \cup \{i\}$ ;
12 return  $\mathbf{A}$ ;
```

Do sử dụng phương pháp tổ hợp 3 kênh phổ theo công thức (3.1) để tạo chỉ số của nhóm nên hiển nhiên các nhóm của các điểm ảnh đã được sắp xếp tăng dần theo chiều thứ nhất của dữ liệu. Và như vậy, chúng ta không phải tiêu tốn thêm thời gian để sắp xếp lại danh sách nhóm các điểm ảnh.

Sau khi danh sách nhóm các điểm ảnh được tạo ra, hình dạng của các nhóm này được hình dung như một khối đa diện, kích thước của 3 mặt tương ứng với 3 chiều dữ liệu đầu tiên bằng 1 đơn vị. Các chiều còn lại của khối đa diện bị giới hạn bởi các điểm ảnh thuộc nhóm đó. Các cạnh còn lại có điểm đầu và điểm cuối trùng với các giá trị cực trị trên các kênh phổ tương ứng của các điểm ảnh thuộc nhóm đang được xét. Toàn bộ các điểm ảnh trong một nhóm có giá trị giống nhau trên ba kênh phổ đầu tiên.

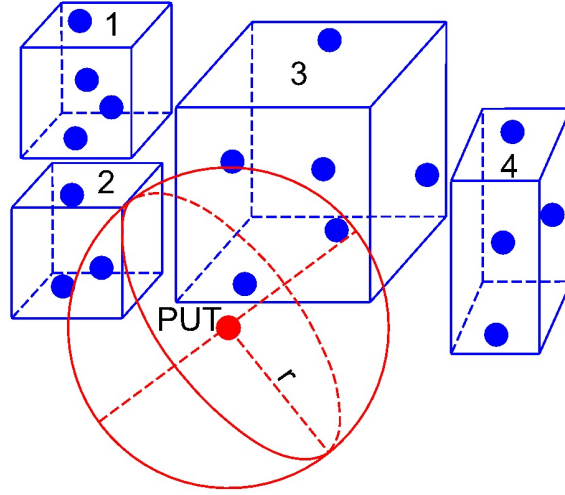
3.4. Tính toán PDF và phân loại các điểm ảnh

Tính toán PDF cho PUT là chức năng chính và quan trọng nhất của thuật toán phát hiện các điểm ảnh dị thường trên ảnh đa phổ, siêu phổ. Chính vì vậy, thời gian để ước lượng hàm mật độ xác suất theo phương pháp phi tham số cho PUT cũng chiếm phần lớn thời gian xử lý của thuật toán. Điểm mấu chốt tiêu tốn nhiều thời gian trong phương pháp ước lượng mật độ xác suất phi tham số là tìm kiếm những điểm dữ liệu thỏa mãn điều kiện $K(u) \neq 0$.

Trong bảng 2.1, ngoại trừ hàm nhân Gaussian là không có điều kiện cho $K(u)$, các hàm nhân còn lại đều có điều kiện $|u| = \left| \frac{x_{id} - x_{jd}}{h} \right| \leq \epsilon$, $i = 1, 2, \dots, n$; $j = 1, 2, \dots, n$; $d = 1, 2, \dots, k$ thì $K(u) \neq 0$, ngược lại $K(u) = 0$. Trong công trình [CT5] cho thấy rằng, hàm nhân Gaussian không phải là hàm nhân phù hợp để áp dụng nhằm phát hiện các điểm ảnh dị thường trên ảnh viễn thám quang học, do đó nó sẽ không được sử dụng trong luận án này. Các hàm nhân còn lại, tùy thuộc vào các hàm nhân khác nhau, ϵ sẽ được gán các giá trị khác nhau (ví dụ hàm nhân Uniform $\epsilon = 1$, Hypercube $\epsilon = \frac{1}{2}, \dots$). Đặt $r = h \times \epsilon$, khi đó để thỏa mãn điều kiện $K(u) \neq 0$ thì $|x_{id} - x_{jd}| \leq r$. Điều này có ý nghĩa rằng với PUT là \mathbf{x}_i , những nhóm các điểm ảnh nào nằm trong hoặc giao cắt với hình siêu cầu có bán kính r , tâm là \mathbf{x}_i sẽ được chọn để tham gia tính toán PDF cho PUT. Những điểm ảnh nằm trong các nhóm điểm ảnh không được lựa chọn sẽ không được tham gia tính toán PDF cho PUT bởi những điểm ảnh này làm cho $K(u) = 0$. Trong ví dụ minh họa trên Hình 3.3, có thể coi các nhóm điểm ảnh được bao bọc trong một hình hộp, nhóm 1 và nhóm 4 sẽ bị loại ngay từ đầu còn những điểm ảnh hoặc nhóm điểm ảnh trong nhóm 2 và nhóm 3 là những ứng cử viên có thể thỏa mãn điều kiện nhân $K(u) \neq 0$, chúng ta chỉ còn phải kiểm tra chúng để tìm ra được những điểm ảnh thực sự thỏa mãn điều kiện để nhân $K(u) \neq 0$.

Như vậy, sẽ loại bỏ một số lượng lớn các điểm ảnh khi tính toán PDF cho PUT nên đã giảm thời gian tính toán của thuật toán. Thuật toán để tìm những nhóm điểm ảnh là ứng viên tham gia tính toán PDF cho PUT được thể hiện trong Thuật toán 3.2.

Bước đầu tiên của Thuật toán 3.2 là tìm kiếm nhóm điểm ảnh chứa PUT. Do các nhóm điểm ảnh đã được sắp xếp tăng dần theo chiều dữ liệu thứ nhất nên lấy PUT làm trung tâm, lần lượt lựa chọn những nhóm điểm ảnh nằm



Hình 3.3: Tìm kiếm những nhóm điểm ảnh nằm trong hoặc giao cắt với hình siêu cầu bán kính r , tâm là PUT

ở bên trái và bên phải của nhóm điểm ảnh chứa PUT nếu như khoảng cách của các nhóm đó đến PUT (tính khoảng cách theo chiều thứ nhất) nhỏ hơn hoặc bằng r .

Để tăng tốc độ tính toán PDF cho PUT, luận án đề xuất sử dụng một trong ba phương pháp: sử dụng cây kd-tree để hỗ trợ tìm nhanh những điểm ảnh thỏa mãn điều kiện $K(u) \neq 0$, sử dụng phương pháp tính toán đa luồng trên CPU, song song trên nền tảng GPU CUDA.

3.4.1. Ứng dụng cây kd-tree hỗ trợ tính toán PDF

Ứng dụng cây kd-tree để hỗ trợ tìm nhanh những điểm ảnh thỏa mãn điều kiện $K(u) \neq 0$, những điểm ảnh này sẽ được dùng để tính toán PDF cho PUT mà không cần phải kiểm tra điều kiện $|u| \leq \epsilon$. Cây kd-tree được phát triển và công bố bởi Bentley [10] vào năm 1975. Về bản chất, nó là một cây nhị phân (do mỗi node của cây có tối đa 2 nhánh con), trong đó mỗi nút biểu diễn một phân vùng không gian k chiều. Node gốc đại diện cho toàn bộ không gian, các node lá đại diện cho không gian con chứa một tập con độc nhất của tập dữ liệu đầu vào. Điểm đặc biệt của cây kd-tree là các đỉnh của cây là những điểm phân chia không gian thành hai phần. Việc phân chia không gian như vậy sẽ thuận tiện cho tìm kiếm những điểm trong cây gần nhất với một điểm nào đó trong vùng không gian. Điều này có nghĩa rằng, để tìm những điểm thuộc cây gần với một điểm nào đó trong không gian sẽ

Thuật toán 3.2: Tìm kiếm những nhóm điểm ảnh nằm trong bán kính r , tâm là PUT (SearchGroupPixels)

input: X : ma trận của các điểm ảnh, A : mảng chứa các nhóm điểm ảnh, $query$: PUT, r : bán kính.

output: P : danh sách nhóm các điểm ảnh cần tìm kiếm.

```

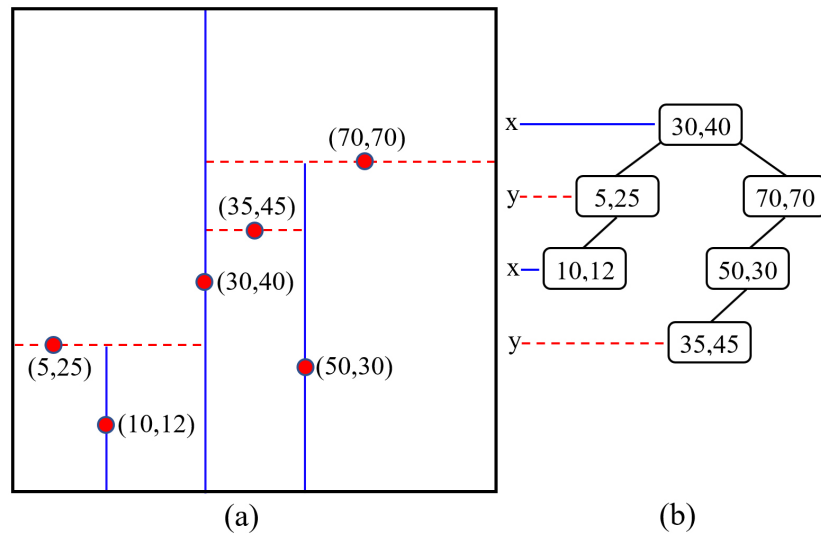
1  $L \leftarrow \emptyset$ ;
2  $m \leftarrow 0$ 
3 for  $i \leftarrow 1$  to  $N$  do
4   for  $d \leftarrow 1$  to 3 do
5     if  $m < X[i][d]$  then
6        $m = X[i][d]$ ;
7    $index \leftarrow 0$ 
8   for  $d \leftarrow 1$  to 3 do
9      $index \leftarrow index + (m + 1)^{3-d} \times query[d]$ ;
10   $P \leftarrow P \cup \{A[index]\}$ 
11  for  $i \leftarrow index - 1$  to 1 do
12    if khoảng cách từ query đến  $A[i]$  trên chiều dữ liệu thứ nhất vẫn
        nhỏ hơn hoặc bằng  $r$  then
13       $P \leftarrow P \cup \{A[i]\}$ ;
14    else
15      break;
16  for  $i \leftarrow index + 1$  to  $|A|$  do
17    if khoảng cách từ query đến  $A[i]$  trên chiều dữ liệu thứ nhất vẫn
        nhỏ hơn hoặc bằng  $r$  then
18       $P \leftarrow P \cup \{A[i]\}$ ;
19    else
20      break;
21 return  $P$ ;

```

dựa trên một số phép phân hoạch không gian để loại bỏ các vùng không gian không cần thiết, như vậy sẽ thu hẹp được không gian tìm kiếm.

Để hiểu rõ hơn về cây kd-tree, chúng ta xét một ví dụ xây dựng cây kd-tree từ bộ dữ liệu 2 chiều (30,40), (5,25), (10,12), (70,70), (50,30), (35,45), chi tiết được thể hiện trong Hình 3.4. Hình 3.4(a) thể hiện các vùng không gian đã

được chia, những đường thẳng liền nét là những đường chia không gian dữ liệu theo chiều thứ nhất (chiều x), những đường thẳng nét đứt chia không gian dữ liệu theo chiều thứ hai (chiều y). Hình 3.4(b) thể hiện cây kd-tree được xây dựng từ bộ dữ liệu trên. Quy tắc xây dựng cây và phân chia không gian như sau: Chọn một điểm dữ liệu làm gốc, tại gốc chia toàn bộ không gian dữ liệu theo chiều thứ nhất làm hai phần (trong ví dụ, điểm gốc được chọn là $(30,40)$, điểm này chia toàn bộ không gian dữ liệu theo chiều x thành 2 phần, phần bên trái là những điểm dữ liệu có chiều x nhỏ hơn hoặc bằng 30, phần bên phải là những điểm dữ liệu có chiều x lớn hơn hoặc bằng 30). Tiếp theo, xét lần lượt từng điểm dữ liệu, so sánh điểm dữ liệu này với các nút trên cây, bắt đầu từ gốc. Quy tắc so sánh như sau: giả sử dữ liệu của chúng ta có k chiều (quy định bắt đầu từ 0 đến $k - 1$), lấy bậc của nút cần so sánh chia cho k được phần dư, phần dư này chính là chiều dữ liệu được dùng để so sánh. Nếu điểm dữ liệu nhỏ hơn hoặc bằng với nút so sánh thì điểm dữ liệu sẽ nằm bên trái nút so sánh, ngược lại nằm bên phải. Tiếp tục so sánh đến khi gặp lá thì thêm điểm dữ liệu vào cây.



Hình 3.4: a) Minh họa phân chia miền không gian, b) Minh họa cây kd-tree đã được xây dựng từ dữ liệu đã cho.

Thuật toán 3.3 xây dựng cây kd-tree từ các điểm ảnh của ảnh đầu vào. Trong thuật toán này, đầu tiên phải xây dựng cấu trúc một node của cây kd-tree để lưu trữ dữ liệu và một số thuộc tính khác nữa của node, sau đó xây dựng hàm InsertNode để chèn một node vào cây, cuối cùng xây dựng

hàm CreateKDTree để xây dựng thành một cây hoàn chỉnh.

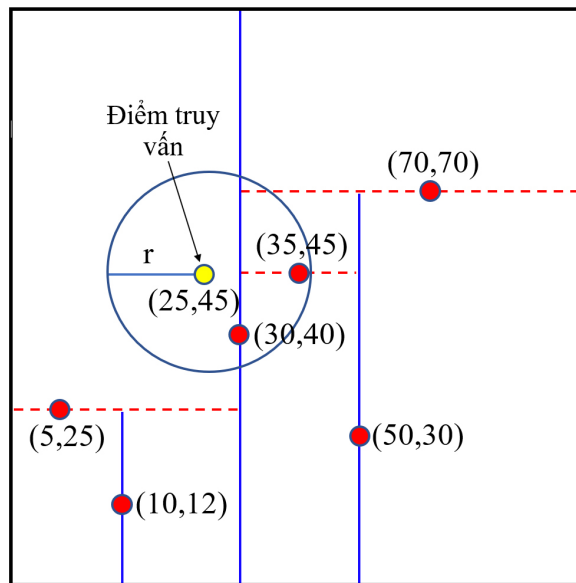
Thuật toán 3.3: Tạo cây kd-tree (CreateKdTree)

```

1 Function InsertNode()
2   input: nút gốc root, điểm dữ liệu point, số chiều dữ liệu L, mức
   của cây level;
3   output: điểm dữ liệu được thêm vào cây;
4   Tìm chiều của không gian dữ liệu:  $axis \leftarrow level \% L$ ;
5   if  $root = \emptyset$  then
6     | - Khởi tạo một nút mới;
7     | - Gán nút mới khởi tạo cho root;
8   else
9     | if  $point[axis] < root.data[axis]$  then
10    | | InsertNode(root.left, point, L, level + 1);
11    | else
12    | | InsertNode(root.right, point, L, level + 1);
13    | end
14  end
15 Function CreateKDTree()
16   input: Ma trận điểm ảnh X, số điểm ảnh N, số kênh phổ L;
17   output: Cây Kd-tree hoàn chỉnh;
18   KDNode* root  $\leftarrow \emptyset$ ;
19   for  $i \leftarrow 0$  to  $N - 1$  do
20     | InsertNode(root,  $\mathbf{X}_i$ , L, 0);
21   end
22   return root;

```

Sau khi xây dựng cây kd-tree (theo Thuật toán 3.3), mục tiêu của chúng ta là sử dụng cây này để tìm tất cả những điểm ảnh đáp ứng điều kiện $K(u) \neq 0$. Với điểm x_i đang được xem xét, những điểm x_j nằm trong hình siêu cầu bán kính r (sử dụng thước đo khoảng cách Chebyshev[118] để đo khoảng cách từ điểm x_i tới điểm x_j) sẽ là những điểm được chọn để tính $K(u)$, xem minh họa trong hình 3.5. Trong hình 3.5, điểm dữ liệu cần tính toán PDF là (25,45) với $r = 10$ thì những điểm dữ liệu (35,45) và (30,40) nằm trong hình tròn bán kính r sẽ thỏa mãn điều kiện để $K(u) \neq 0$. Vì vậy, những điểm dữ liệu này được tham gia tính toán PDF cho điểm dữ liệu đang xét.



Hình 3.5: Minh họa những điểm được chọn để tính $K(u)$.

Thuật toán 3.4 sẽ tìm kiếm những điểm ảnh nằm trong hình siêu cầu có bán kính r , có tâm là x_i . Thuật toán sử dụng phương pháp đệ quy để tìm kiếm danh sách các điểm dữ liệu nằm trong hình siêu cầu bán kính r có tâm là điểm đang xét. Những điểm dữ liệu thỏa mãn yêu cầu được lưu trong danh sách *list*. Đầu tiên, thuật toán kiểm tra node gốc *root*, nếu nó rỗng thì thoát khỏi thuật toán. Tiếp đến tính toán khoảng cách từ điểm đang xét đến *root* (sử dụng phương pháp tính khoảng cách Chebyshev), tìm chiều dữ liệu để so sánh. Kiểm tra nếu *root* nằm trong hình siêu cầu đó thì thêm *root* vào *list*. So sánh điểm đang xét với *root*, nếu nhỏ hơn *root* tìm nhánh bên trái của *root*, ngược lại tìm nhánh bên phải của *root*. Chỉ xét riêng trên một chiều dữ liệu so sánh, nếu khoảng cách từ điểm đang xét đến *root* trên chiều dữ liệu đó mà nhỏ hơn hoặc bằng r bắt buộc phải tìm cả nhánh bên trái và nhánh bên phải của *root*. Ví dụ trên hình 3.5, rõ ràng rằng điểm truy vấn nằm bên phần không gian bên trái của node gốc là (30,40), thông thường chúng ta chỉ tìm những điểm nằm trên nhánh trái của *root* sẽ bỏ qua những điểm dữ liệu thỏa mãn yêu cầu của hàm nhân $K(u) \neq 0$.

Có hai trường hợp xảy ra khi áp dụng cây kd-tree hỗ trợ quá trình tính toán PDF cho PUT. Trường hợp đầu tiên, dữ liệu đầu vào là ảnh màu RGB, sau khi nhóm các điểm ảnh được tạo ra từ Thuật toán 3.1 thì trong cùng một nhóm các điểm ảnh sẽ có giá trị trên từng kênh phổ hoàn toàn giống

Thuật toán 3.4: Tìm kiếm trên cây kd-tree (SearchKdTree)

input: Node gốc $root$, điểm cần tính PDF $query$, số chiều dữ liệu L , bán kính siêu cầu r ;

output: Danh sách các điểm được tìm thấy $list$;

```

1 if  $root = NULL$  then
2   | return
3 Tính khoảng cách từ  $query$  đến  $root$ , gán vào  $d$ ;
4 Tìm chiều dữ liệu so sánh:  $axis \leftarrow root.level \% L$ ;
5 if  $d \leq r$  then
6   |  $list \leftarrow list + \langle root \rangle$ ;
7 if  $query[axis] < root.data[axis]$  then
8   | if  $root.left \neq \emptyset$  then
9     | SearchKdTree( $root.left, query, r, list$ );
10  if  $root.right \neq \emptyset$  then
11    | if  $|query[axis] - root.data[axis]| \leq r$  then
12      | SearchKdTree( $root.right, query, r, list$ );
13 else
14   | if  $root.right \neq \emptyset$  then
15     | SearchKdTree( $root.right, query, r, list$ );
16   if  $root.left \neq \emptyset$  then
17     | if  $|query[axis] - root.data[axis]| \leq r$  then
18       | SearchKdTree( $root.left, query, r, list$ );
19 return  $list$ ;
```

nhau. Do đó chúng ta có thể coi một nhóm tương đương với một điểm ảnh. Vì mỗi một nhóm các điểm ảnh tương đương với một điểm ảnh nhưng vai trò của các điểm ảnh trong nhóm không thay đổi nên công thức (2.5) được biến đổi thành:

$$\hat{f}_A(\mathbf{A}_i) = \frac{1}{N} \sum_{j=1}^{MC_i} \left\{ \prod_{d=1}^L \frac{1}{h_d} K \left(\frac{A_{id} - P_{jd}}{h_d} \right) \right\} \times |P_j| \quad (3.3)$$

$i = 1, 2, \dots, M$, M là số nhóm các điểm ảnh có các giá trị phổ trùng nhau, P là danh sách nhóm điểm ảnh nằm trong hình siêu cầu bán kính r , tâm A_i tìm được trong Thuật toán 3.4, P_j là nhóm điểm ảnh thứ j trong danh sách

các điểm ảnh P . $MC_i = |P|$ là tổng số nhóm điểm ảnh trong danh sách P , A_{id} là giá trị của chiều thứ d của nhóm điểm ảnh A_i , $|P_j|$ số lượng các điểm ảnh trong nhóm thứ j của danh sách P .

Thuật toán 3.5 được xây dựng theo công thức (3.3) để phát hiện các điểm ảnh dị thường trên ảnh màu RGB sử dụng phương pháp nhóm các điểm ảnh cùng giá trị và cây kd-tree để hỗ trợ quá trình tính toán PDF. Ngoài ra, Thuật toán 3.5, đã kế thừa ý tưởng tăng độ chính xác phát hiện các điểm ảnh dị thường của thuật toán IKDE (Thuật toán 2.3) trong Chương 2. Trong thuật toán này, cây kd-tree đóng vai trò là cây tìm kiếm nhanh những điểm ảnh nằm trong hình cầu bán kính r , tâm là PUT, sử dụng thước đo khoảng cách Chebyshev để tính khoảng cách từ PUT đến các điểm ảnh lân cận. Do các điểm ảnh trong cùng một nhóm giống nhau hoàn toàn trên các kênh nên mật độ xác suất của các điểm ảnh trong nhóm đó giống nhau.

Trường hợp thứ hai, dữ liệu đầu vào là ảnh đa phổ (MSI), ảnh siêu phổ (HSI) có số kênh phổ lớn hơn 3. Lúc này các điểm ảnh trong cùng một nhóm chỉ giống nhau trên 3 kênh đầu tiên, những kênh còn lại khác nhau. Trước khi áp dụng cây kd-tree để giảm thời gian tính toán PDF cho PUT, sử dụng Thuật toán 3.2 để tìm kiếm những nhóm điểm ảnh chứa những điểm ảnh có thể thỏa mãn điều kiện $K(u) \neq 0$. Những điểm ảnh trong những nhóm điểm ảnh vừa tìm được sẽ được dùng để xây dựng cây kd-tree. Thuật toán 3.6 được xây dựng để phát hiện các điểm ảnh dị thường trên MSI, HSI sử dụng phương pháp nhóm các điểm ảnh cùng giá trị và cây kd-tree để hỗ trợ quá trình tính toán PDF, thuật toán này cũng đã kế thừa ý tưởng của thuật toán IKDE (Thuật toán 2.3) để tăng độ chính xác phát hiện các điểm ảnh dị thường. Thuật toán 3.5 sử dụng tất cả các nhóm điểm ảnh được dùng để xây dựng cây kd-tree và chỉ xây dựng một cây kd-tree duy nhất cho mọi nhóm điểm ảnh, Thuật toán 3.6 có sự khác biệt. Tương ứng với mỗi một nhóm các điểm ảnh sẽ xây dựng một cây kd-tree với dữ liệu đầu vào là tất cả các điểm ảnh trong các nhóm các điểm ảnh tìm được trong Thuật toán 3.2. Vai trò của cây kd-tree trong Thuật toán 3.6 tương tự như trong Thuật toán 3.5.

3.4.2. Tính toán PDF đa luồng trên CPU

Tính toán PDF đa luồng trên CPU đã được Michailidis và các cộng sự trình bày chi tiết trong [86]. Ý tưởng của phương pháp này là phân chia một

Thuật toán 3.5: Phát hiện các điểm ảnh dị thường trên ảnh màu RGB sử dụng kỹ thuật nhóm các điểm ảnh cùng giá trị và cây kd-tree (gọi tắt là Thuật toán GP-KDT1).

input: X : ma trận các điểm ảnh, N : số điểm ảnh, h : băng thông, ϵ : hằng số điều kiện để $K(u) \neq 0$, η : ngưỡng phát hiện dị thường.

output: C_1 : tập hợp những điểm ảnh bình thường, C_2 : tập hợp những điểm ảnh dị thường.

```

1  $X' \leftarrow \mathbf{GaussianBlur}(X, \sigma)$ ; // sử dụng Thuật toán 2.2
2  $S \leftarrow$  tập các điểm ảnh đặc trưng (sử dụng SIFT để tìm các điểm ảnh
   đặc trưng);
3  $D \leftarrow$  tập hợp các điểm ảnh nằm trong các cửa sổ lấy các điểm trong
    $S$  làm tâm;
4  $A \leftarrow \mathbf{CreateGroupPixels}(X', N, 3)$ ; // sử dụng Thuật toán 3.1;
5  $T \leftarrow \mathbf{CreateKdTree}(A, |A|, 3)$ ; // sử dụng Thuật toán 3.3
6  $C_1 \leftarrow \emptyset$ ;  $C_2 \leftarrow \emptyset$ ;  $r \leftarrow h \times \epsilon$ ;
7 for  $i \leftarrow 1$  to  $|A|$  do
8   if  $A[i] \in D$  then
9      $P \leftarrow \mathbf{T.SearchKdTree}(A[i], r)$ ; // sử dụng Thuật toán 3.4;
10     $sum\_ker \leftarrow 0$ ;
11    for  $j \leftarrow 1$  to  $|P|$  do
12       $mul \leftarrow 1$ ;
13      for  $d \leftarrow 1$  to 3 do
14         $mul \leftarrow mul \times K\left(\frac{A[i][d] - P[j][d]}{h}\right)$ ;
15       $sum\_ker \leftarrow sum\_ker + mul \times |P[j]|$ ;
16    for  $j \leftarrow 1$  to  $|A[i]|$  do
17      if  $\frac{sum\_ker}{N} \leq \eta$  then
18         $C_2 \leftarrow C_2 \cup \{X'[A[i][j]]\}$ ;
19      else
20         $C_1 \leftarrow C_1 \cup \{X'[A[i][j]]\}$ ;
21  else
22    for  $j \leftarrow 1$  to  $|A[i]|$  do
23       $C_1 \leftarrow C_1 \cup \{X'[A[i][j]]\}$ ;

```

Thuật toán 3.6: Phát hiện các điểm ảnh dị thường trên MSI, HSI sử dụng kỹ thuật nhóm các điểm ảnh cùng giá trị và cây kd-tree (gọi tắt là Thuật toán GP-KDT2).

input: X : ma trận các điểm ảnh, N : số điểm ảnh, L : số kênh phổ, h : băng thông, ϵ : hằng số điều kiện để $K(u) \neq 0$, η : ngưỡng phát hiện dị thường.

output: C_1 : tập hợp những điểm ảnh bình thường, C_2 : tập hợp những điểm ảnh dị thường.

```

1  $S \leftarrow$  tập các điểm ảnh đặc trưng (sử dụng SIFT để tìm các điểm ảnh
   đặc trưng);
2  $D \leftarrow$  tập hợp các điểm ảnh nằm trong các cửa sổ lấy các điểm trong
    $S$  làm tâm;
3  $X' \leftarrow$  GaussianBlur( $X, \sigma$ ); // sử dụng Thuật toán 2.2
4  $A \leftarrow$  CreateGroupPixels( $X', N, L$ ); // sử dụng Thuật toán 3.1;
5  $C_1 \leftarrow \emptyset$ ;  $C_2 \leftarrow \emptyset$ ;  $r \leftarrow h \times \epsilon$ ;
6 for  $i \leftarrow 1$  to  $|A|$  do
7    $query \leftarrow X[A[i][1]]$ ;
8    $K \leftarrow$  SearchGroupPixels( $A, query, r$ ); // sử dụng Thuật toán 3.2
9    $T \leftarrow$  CreateKDTree( $L, |K|, L$ ); // sử dụng Thuật toán 3.3
10  for  $j \leftarrow 1$  to  $|A[i]|$  do
11    if  $X'[A[i][j]] \in D$  then
12       $P \leftarrow$  T.SearchKdTree( $X'[A[i][j]], r$ ); // theo Thuật
        toán 3.4;
13      for  $l \leftarrow 1$  to  $|P|$  do
14         $mul\_ker \leftarrow 1$ ;
15        for  $d \leftarrow 1$  to  $L$  do
16           $mul\_ker \leftarrow$ 
             $mul\_ker \times K\left(\frac{X'[A[i][j]][d] - P[l][d]}{h}\right)$ ;
17         $sum\_ker \leftarrow sum\_ker + mul\_ker$ ;
18      if  $\frac{sum\_ker}{N} \leq \eta$  then
19         $C_2 \leftarrow C_2 \cup \{X'[A[i][j]]\}$ ;
20      else
21         $C_1 \leftarrow C_1 \cup \{X'[A[i][j]]\}$ ;
22    else
23       $C_1 \leftarrow C_1 \cup \{X'[A[i][j]]\}$ ;

```

vấn đề có quy mô lớn thành một số tác vụ nhỏ hơn có thể được giải quyết đồng thời trên các máy tính độc lập. Việc song song hóa nhiệm vụ tính toán PDF trong nghiên cứu này sẽ được tiếp cận theo hướng phân vùng dữ liệu. Có nghĩa rằng, tất cả các luồng thực hiện các hoạt động giống nhau trên các phần dữ liệu khác nhau của một tập dữ liệu lớn. Dữ liệu đầu vào là một mảng chứa các nhóm điểm ảnh A được tạo ra từ Thuật toán 3.1, trên chiều thứ nhất của A sẽ được chia lần lượt thành q phần tương ứng với q luồng, từ phần đầu tiên đến phần $q - 1$ có kích thước là phần nguyên của $\frac{M}{q}$ nhóm các điểm ảnh ($M = |A|$), phần thứ q có kích thước là $M - (q - 1) \times \frac{M}{q}$ nhóm các điểm ảnh. Chỉ số đầu tiên (*startindex*) của các phần trên A sẽ là $startindex_i = i \times \frac{M}{q}$, $i = 0, 2, \dots, q - 1$, chỉ số cuối (*endindex*) của các phần trên A sẽ là $endindex_i = startindex_i + \frac{M}{q}$, riêng chỉ số cuối của phần thứ q sẽ bằng chỉ số cuối cùng của A .

Tương tự như phương pháp đầu tiên sử dụng cây kd-tree hỗ trợ quá trình tính toán PDF cho PUT, phương pháp tính toán PDF đa luồng trên CPU cho PUT cũng xảy ra hai trường hợp. Trường hợp đầu tiên, dữ liệu đầu vào là ảnh màu RGB, chúng ta có thể coi một nhóm tương đương với một điểm ảnh. Công thức (2.5) được biến đổi thành:

$$\hat{f}_A(\mathbf{A}_i) = \frac{1}{N} \sum_{j=1}^M \left\{ \prod_{d=1}^L \frac{1}{h_d} K \left(\frac{A_{id} - A_{jd}}{h_d} \right) \right\} \times |A_j| \quad (3.4)$$

Trong đó, M là số nhóm các điểm ảnh, $|A_j|$ là số lượng điểm ảnh trong nhóm điểm ảnh thứ j .

Thuật toán 3.7 được xây dựng để thực thi trên tất cả các luồng với dữ liệu đầu vào khác nhau nhằm phát hiện các điểm ảnh dị thường trên ảnh màu RGB. Trước khi thực thi Thuật toán 3.7 trên các luồng, chương trình chính phải thực hiện việc nhóm các điểm ảnh theo Thuật toán 3.1 trước, làm mịn ảnh và tìm kiếm các điểm ảnh đặc trưng theo phương pháp SIFT hoặc SURF. Do các điểm ảnh trong cùng một nhóm giống nhau hoàn toàn trên các kênh nên mật độ xác suất của các điểm ảnh trong nhóm đó giống nhau.

Trường hợp thứ hai, dữ liệu đầu vào là MSI, HSI có số kênh phổ lớn hơn

Thuật toán 3.7: Phát hiện các điểm ảnh dị thường trên ảnh màu RGB sử dụng kỹ thuật nhóm các điểm ảnh cùng giá trị và tính toán đa luồng (gọi tắt là Thuật toán GP-MC1).

input: X' : ma trận các điểm ảnh đã được làm mịn, S : tập các điểm ảnh đặc trưng, D : tập hợp các điểm ảnh nằm trong các cửa sổ lấy các điểm trong S làm tâm, A : ma trận chứa nhóm các điểm ảnh, $startindex$: chỉ số đầu tiên trên chiều thứ nhất của A , $endindex$: chỉ số cuối trên chiều thứ nhất của A , h : băng thông, ϵ : hằng số để $K(u) \neq 0$, η : ngưỡng phát hiện dị thường.
output: C_1 : tập các điểm ảnh bình thường, C_2 : tập các điểm ảnh dị thường.

```

1  $C_1 \leftarrow \emptyset$ ;  $C_2 \leftarrow \emptyset$ ;  $r \leftarrow h \times \epsilon$ ;
2 for  $i \leftarrow startindex$  to  $endindex$  do
3   if  $A[i] \in D$  then
4      $P \leftarrow \text{SearchGroupPixels}(A, A[i], r)$ ; //theo Thuật toán 3.2
5      $sum\_ker \leftarrow 0$ ;
6     for  $j \leftarrow 1$  to  $|P|$  do
7        $mul\_ker \leftarrow 1$ ;
8       for  $d \leftarrow 1$  to 3 do
9          $mul\_ker \leftarrow mul\_ker \times K\left(\frac{A[i][d] - P[j][d]}{h}\right)$ ;
10         $sum\_ker \leftarrow sum\_ker + mul\_ker \times |P[j]|$ ;
11      for  $j \leftarrow 1$  to  $|A[i]|$  do
12        if  $\frac{sum\_ker}{N} \leq \eta$  then
13           $C_2 \leftarrow C_2 \cup \{X'[A[i][j]]\}$ ;
14        else
15           $C_1 \leftarrow C_1 \cup \{X'[A[i][j]]\}$ ;
16      else
17        for  $j \leftarrow 1$  to  $|A[i]|$  do
18           $C_1 \leftarrow C_1 \cup \{X'[A[i][j]]\}$ ;

```

Thuật toán 3.8: Phát hiện các điểm ảnh dị thường trên MSI, HSI sử dụng kỹ thuật nhóm các điểm ảnh cùng giá trị và tính toán đa luồng (gọi tắt là Thuật toán GP-MC2).

input: X' : ma trận các điểm ảnh đã được làm mịn, S : tập các điểm ảnh đặc trưng, D : tập hợp các điểm ảnh nằm trong các cửa sổ lấy các điểm trong S làm tâm, A : ma trận chứa nhóm các điểm ảnh, $startindex$: chỉ số đầu tiên trên chiều thứ nhất của A , $endindex$: chỉ số cuối trên chiều thứ nhất của A , h : băng thông, ϵ : hằng số để $K(u) \neq 0$, η : ngưỡng phát hiện dị thường.

output: C_1 : tập các điểm ảnh bình thường, C_2 : tập các điểm ảnh dị thường.

```

1  $C_1 \leftarrow \emptyset$ ;  $C_2 \leftarrow \emptyset$ ;  $r \leftarrow h \times \epsilon$ ;
2 for  $i \leftarrow startindex$  to  $endindex$  do
3   for  $j \leftarrow 1$  to  $|A[i]|$  do
4     if  $X'[A[i][j]] \in D$  then
5        $P \leftarrow \text{SearchGroupPixels}(A, X'[A[i][j]], r)$ ; //theo Thuật
        toán 3.2
6        $sum\_ker \leftarrow 0$ ;
7       for  $l \leftarrow 1$  to  $|P|$  do
8         for  $f \leftarrow 1$  to  $|P[l]|$  do
9            $mul \leftarrow 1$ ;
10          for  $d \leftarrow 1$  to  $L$  do
11             $mul \leftarrow$ 
12               $mul \times K\left(\frac{X'[A[i][j]][d] - X'[P[l][f]][d]}{h}\right)$ ;
13             $sum\_ker \leftarrow sum\_ker + mul$ ;
14          if  $\frac{sum\_ker}{n} \leq \eta$  then
15             $C_2 \leftarrow C_2 \cup \{X'[A[i][j]]\}$ ;
16          else
17             $C_1 \leftarrow C_1 \cup \{X'[A[i][j]]\}$ ;
18          else
19             $C_1 \leftarrow C_1 \cup \{X'[A[i][j]]\}$ ;

```

ba. Thuật toán 3.8 được xây dựng nhằm phát hiện các điểm ảnh dị thường sử dụng phương pháp nhóm các điểm ảnh cùng giá trị và tính toán PDF đa luồng trên CPU. Lúc này các điểm ảnh trong cùng một nhóm chỉ giống nhau trên 3 kênh đầu tiên, những kênh còn lại khác nhau. Do đó, phải tính toán PDF cho từng điểm ảnh trong từng nhóm. Tương ứng với mỗi một điểm ảnh chúng ta sẽ tìm được danh sách nhóm các điểm ảnh (theo Thuật toán 3.2) mà những điểm ảnh trong các nhóm này có thể làm thỏa mãn $K(u) \neq 0$, những nhóm điểm ảnh không được lựa chọn chúng ta không phải quan tâm tới nó bởi chắc chắn các điểm ảnh trong các nhóm này sẽ làm cho $K(u) = 0$. Những điểm ảnh trong danh sách nhóm các điểm ảnh vừa được chọn được tham gia tính toán PDF cho PUT theo công thức (2.5). Tuy nhiên, lúc này n trong công thức (2.5) là số lượng điểm ảnh trong danh sách các nhóm điểm vừa được lựa chọn theo Thuật toán 3.2.

3.4.3. Tính toán PDF song song trên GPU

Tương tự như tính toán PDF đa luồng trên CPU, phương pháp tính toán PDF song song trên GPU CUDA cũng đã được Michailidis và các cộng sự trình bày chi tiết trong [87]. Luận án sử dụng ý tưởng của Michailidis và các cộng sự để cài đặt thuật toán, tuy nhiên có sự thay đổi, bao gồm: dữ liệu đầu vào của thuật toán là danh sách nhóm các điểm ảnh; sử dụng Thuật toán 3.2 để tìm kiếm nhanh những điểm ảnh tiềm năng có thể làm cho $K(u) \neq 0$, những điểm ảnh còn lại sẽ được bỏ qua không phải tham gia tính toán PDF cho PUT. Tương tự như trong phương pháp thứ nhất và thứ hai, phương pháp tính toán PDF song song trên nền tảng GPU CUDA cho PUT cũng xảy ra hai trường hợp. Trường hợp đầu tiên, dữ liệu đầu vào là ảnh màu RGB, tính toán PDF cho PUT theo thức (3.4).

Thuật toán 3.9 được xây dựng để phát hiện các điểm ảnh dị thường trên ảnh màu RGB, dữ liệu đầu vào là danh sách nhóm các điểm ảnh được tạo ra từ Thuật toán 3.1. Kết quả thu được là một danh sách chứa giá trị dị thường của các điểm ảnh, nếu nhận giá trị là 0, điểm ảnh đó là "bình thường", ngược lại nhận giá trị 1 điểm ảnh đó là điểm ảnh "dị thường". Thuật toán 3.10 được xây dựng để phát hiện các điểm ảnh dị thường trên MSI, HSI sử dụng phương pháp nhóm các điểm ảnh cùng giá trị và tính toán PDF song song sử dụng GPU CUDA. Thuật toán 3.10 có ý tưởng và trình tự thực hiện tương

Thuật toán 3.9: Phát hiện các điểm ảnh dị thường trên ảnh màu RGB sử dụng kỹ thuật nhóm các điểm ảnh cùng giá trị và tính toán song song (gọi tắt là Thuật toán GP-PC1).

input: X' : ma trận các điểm ảnh đã được làm mịn, A : ma trận chứa nhóm các điểm ảnh, S : tập các điểm ảnh đặc trưng, D : tập hợp các điểm ảnh nằm trong các cửa sổ lấy các điểm trong S làm tâm, h : băng thông, ϵ : hằng số để $K(u) \neq 0$, η : ngưỡng phát hiện dị thường.

output: *anomaly_value*: giá trị dị thường của các điểm ảnh (nhận giá trị 0 là điểm ảnh bình thường, nhận giá trị 1 là những điểm ảnh dị thường).

```

1  $r \leftarrow h \times \epsilon$ ;
2  $id \leftarrow blockDim.x * blockIdx.x + threadIdx.x$ ;
3 if  $A[id] \in D$  then
4    $P \leftarrow \text{SearchGroupPixels}(A, A[id], r)$ ; //theo Thuật toán 3.2
5    $sum\_ker \leftarrow 0$ ;
6   for  $j \leftarrow 1$  to  $|P|$  do
7      $mul\_ker \leftarrow 1$ ;
8     for  $d \leftarrow 1$  to 3 do
9        $mul\_ker \leftarrow mul\_ker \times K\left(\frac{A[id][d] - P[j][d]}{h}\right)$ ;
10     $sum\_ker \leftarrow sum\_ker + mul\_ker \times |P[j]|$ ;
11  for  $j \leftarrow 1$  to  $|A[id]|$  do
12    if  $\frac{sum\_ker}{N} \leq \eta$  then
13       $\text{atomicAdd}(\&anomaly\_valalue[A[id][j]], 1)$ ;
14    else
15       $\text{atomicAdd}(\&anomaly\_valalue[A[id][j]], 0)$ ;
16 else
17   for  $j \leftarrow 1$  to  $|A[id]|$  do
18      $\text{atomicAdd}(\&anomaly\_valalue[A[id][j]], 0)$ ;

```

Thuật toán 3.10: Phát hiện các điểm ảnh dị thường trên MSI, HSI sử dụng kỹ thuật nhóm các điểm ảnh cùng giá trị và tính toán song song (gọi tắt là Thuật toán GP-MC2).

input: X' : ma trận các điểm ảnh đã được làm mịn, A : ma trận chứa nhóm các điểm ảnh, S : tập các điểm ảnh đặc trưng, D : tập hợp các điểm ảnh nằm trong các cửa sổ lấy các điểm trong S làm tâm, h : băng thông, ϵ : hằng số để $K(u) \neq 0$, η : ngưỡng phát hiện dị thường.

output: *anomaly_value*: giá trị dị thường của các điểm ảnh (nhận giá trị 0 là điểm ảnh bình thường, nhận giá trị 1 là những điểm ảnh dị thường).

```

1  $r \leftarrow h \times \epsilon$ ;
2  $id \leftarrow blockDim.x * blockIdx.x + threadIdx.x$ ;
3 for  $i \leftarrow 1$  to  $|A[id]|$  do
4   if  $X'[A[id][i]] \in B$  then
5      $P \leftarrow \text{SearchGroupPixels}(A, X'[A[id][i]], r)$ ; //theo Thuật
        toán 3.2
6      $sum\_ker \leftarrow 0$ ;
7     for  $j \leftarrow 1$  to  $|P|$  do
8       for  $l \leftarrow 1$  to  $|P[j]|$  do
9          $mul \leftarrow 1$ ;
10        for  $d \leftarrow 1$  to  $L$  do
11           $mul \leftarrow mul \times K\left(\frac{X'[A[id][i]][d] - X'[P[j][l]][d]}{h}\right)$ ;
12           $sum\_ker \leftarrow sum\_ker + mul$ ;
13        if  $\frac{sum\_ker}{N} \leq \eta$  then
14           $\text{atomicAdd}(\&anomaly\_valvalue[A[id][i]], 1)$ ;
15        else
16           $\text{atomicAdd}(\&anomaly\_valvalue[A[id][i]], 0)$ ;
17      else
18         $\text{atomicAdd}(\&anomaly\_valvalue[A[id][i]], 0)$ ;

```

tự Thuật toán 3.9, chỉ khác là phải tính PDF riêng cho từng điểm ảnh trong cùng một nhóm các điểm ảnh (Thuật toán 3.9 tính toán PDF chung cho một nhóm do các điểm ảnh trong cùng một nhóm giống nhau hoàn toàn trên ba kênh ảnh) do các điểm ảnh trong cùng một nhóm chỉ giống nhau trên 3 kênh đầu tiên, những kênh còn lại khác nhau.

Như vậy, luận án đã trình bày chi tiết các bước đã đề xuất để tăng tốc độ tính toán cho Thuật toán 2.3:

- Bước 1, nhóm các điểm ảnh có cùng giá trị. Mục đích của việc nhóm các điểm ảnh nhằm giảm thiểu dữ liệu tính toán PDF cho PUT;
- Bước 2, sắp xếp nhóm các điểm ảnh. Mục đích nhằm hỗ trợ tìm kiếm nhanh các nhóm điểm ảnh nằm phía trước và nằm phía sau PUT trong phạm vi giới hạn. Những nhóm điểm ảnh này chứa các điểm ảnh có thể làm cho $K(u) \neq 0$, những nhóm điểm ảnh không được lựa chọn sẽ không phải tham gia tính toán PDF cho PUT.
- Bước 3, tính toán PDF cho các điểm ảnh. Đề xuất ba kỹ thuật để hỗ trợ quá trình tính toán PDF cho các điểm ảnh. Kỹ thuật thứ nhất là ứng dụng cây kd-tree để hỗ trợ tìm nhanh những điểm ảnh nằm trong hình siêu cầu có tâm là PUT, bán kính r , những điểm ảnh này sẽ tham gia tính toán PDF cho PUT, những điểm ảnh còn lại sẽ không phải tham gia tính toán PDF cho PUT nên đã giảm thời gian tính toán. Kỹ thuật thứ hai là tính toán PDF đa luồng trên CPU và kỹ thuật thứ ba là tính toán PDF song song trên GPU sử dụng nền tảng GPU CUDA. Ý tưởng chính của kỹ thuật thứ hai và thứ ba là phân chia một vấn đề có quy mô lớn thành một số tác vụ nhỏ hơn và tính toán đồng thời từ đó giảm được thời gian tính toán;
- Bước 4, phân lớp các điểm ảnh. Sau khi có được kết quả của bước thứ ba, dựa vào ngưỡng để phân các điểm ảnh vào lớp bình thường hoặc lớp dị thường.

3.5. Đánh giá độ phức tạp của các thuật toán

3.5.1. Thuật toán nhóm các điểm ảnh (*CreateGroupPixels*)

- Trong Thuật toán 3.1, có hai lệnh gán (hàng 1 và 2), những lệnh gán này đều có độ phức tạp tính toán là $O(1)$.

- Thuật toán có hai vòng lặp lồng nhau, vòng lặp thứ nhất (hàng 3) thực hiện N lần, vòng lặp thứ hai (hàng 5) thực hiện 3 lần, mỗi lần thực hiện có độ phức tạp tính toán là $O(1)$, do đó:

$$T(N) = \sum_{i=1}^N \sum_{j=1}^3 C = 3.C.N$$

Như vậy, độ phức tạp tính toán của thuật toán nhóm các điểm ảnh là $O(N)$.

3.5.2. Thuật toán tìm kiếm những nhóm điểm ảnh (*SearchGroupPixels*)

- Thuật toán 3.2 có ba lệnh gán (hàng 1, 2 và 3), những lệnh gán này đều có độ phức tạp tính toán là $O(1)$.

- Thuật toán có hai vòng lặp riêng biệt, vòng lặp thứ nhất (hàng 4) thực hiện tối đa M lần ($M=|A|$ là số nhóm các điểm ảnh),

$$T_1(M) = \sum_{i=1}^M C = C_1.M.$$

Vòng lặp thứ hai (hàng 9) thực hiện tối đa M lần,

$$T_2(M) = \sum_{i=1}^m C_2 = C_2.M.$$

$$T(M) = T_1(M) + T_2(M) = (C_1 + C_2)M.$$

Như vậy, độ phức tạp tính toán của thuật toán tìm kiếm nhóm các điểm ảnh là $O(N)$ (vì $M \leq N$), N là số điểm ảnh.

3.5.3. Cây kd-tree

Trong công trình "Multidimensional Binary Search Tree Used for Associative Seaching" [10], J.L.Bentley đã đánh giá độ phức tạp tính toán để tạo cây và tìm kiếm trên cây kd-tree, cụ thể:

- Tạo cây kd-tree: độ phức tạp tính toán là $O(N \log N)$.
- Tìm kiếm trên cây: độ phức tạp tính toán là $O(N)$.

3.5.4. Thuật toán phát hiện các điểm ảnh dị thường trên ảnh màu RGB sử dụng kỹ thuật nhóm các điểm ảnh và cây kd-tree

- Độ phức tạp tính toán của thuật toán nhóm các điểm ảnh (hàng 1) là $T_1(N) = O(N)$.

- Độ phức tạp tính toán tạo cây kd-tree (hàng 2) là $T_2(N) = C.M.\log M$, trong đó M là số nhóm điểm ảnh được tạo ra của Thuật toán 3.1.

- Độ phức tạp tính toán của thuật toán làm mịn ảnh (hàng 3) là $T_3(N) = O(N)$

- Độ phức tạp tính toán của thuật toán tìm các điểm ảnh đặc trưng (hàng 4) là $T_4(N) = O(N)$

- Thuật toán có bốn lệnh gán (hàng 5 và 6), những lệnh gán này đều có độ phức tạp tính toán là $O(1)$.

- Độ phức tạp tính toán của thuật toán tìm kiếm trên cây kd-tree (hàng 9) là $O(N)$.

- Thuật toán có ba vòng lặp lồng nhau, vòng lặp thứ nhất (hàng 7) thực hiện M lần, vòng lặp thứ hai (hàng 11) và thứ tư (hàng 16) thực hiện tối đa M lần, vòng lặp thứ ba (hàng 13) thực hiện L lần. Ta có:

$$T_5(N) = \sum_{i=1}^M (C_{51} \cdot M + \sum_{j=1}^M \sum_{d=1}^L C_{52}) = C_{51} \cdot M^2 + C_{52} \cdot L \cdot M^2 = (C_{51} + L \cdot C_{52}) M^2.$$

$$T(N) = \text{Max}(T_1(N), T_2(N), T_3(N), T_4(N), T_5(N)) = (C_{51} + L \cdot C_{52}) M^2.$$

Như vậy, độ phức tạp tính toán của thuật toán là $O(N^2)$ (vì $M \leq N$).

*** Tính hiệu quả của thuật toán:** trong trường hợp xấu nhất, $M=N$, có nghĩa rằng tất cả các điểm ảnh trong một ảnh có màu sắc khác nhau, không có bất kỳ điểm ảnh nào giống nhau. Điều này rất khó xảy ra bởi trong tự nhiên, lớp phủ thực địa luôn có tính chất phân lớp đối tượng, lớp phủ càng đồng nhất số lượng nhóm càng ít nên $M \ll N$.

Một ví dụ cụ thể: ảnh ở Hình 3.6 có tất cả $N=11.373.204$ điểm ảnh, trong khi số nhóm điểm ảnh được tìm thấy là 65.607 nhóm. Có nghĩa rằng $M \approx 0.006 \times N$.

3.5.5. Thuật toán phát hiện các điểm ảnh dị thường trên ảnh màu MSI, HSI sử dụng kỹ thuật nhóm các điểm ảnh và cây kd-tree

- Độ phức tạp tính toán của thuật toán nhóm các điểm ảnh (hàng 1) là $T_1(N) = O(N)$.

- Độ phức tạp tính toán của thuật toán làm mịn ảnh (hàng 2) là $T_2(N) = O(N)$

- Độ phức tạp tính toán của thuật toán tìm các điểm ảnh đặc trưng (hàng 3) là $T_3(N) = O(N)$

- Thuật toán có bốn lệnh gán (hàng 4 và 5), những lệnh gán này đều có độ phức tạp tính toán là $O(1)$.

- Độ phức tạp tính toán tạo cây kd-tree (hàng 8) là $O(N \log N)$.

- Độ phức tạp tính toán của thuật toán tìm kiếm trên cây kd-tree (hàng 9) là $O(N)$.

- Thuật toán có ba vòng lặp lồng nhau, vòng lặp thứ nhất (hàng 7) thực hiện M lần, vòng lặp thứ hai (hàng 11) và thứ tư (hàng 16) thực hiện tối đa M lần, vòng lặp thứ ba (hàng 13) thực hiện L lần. Ta có:

$$T_4(N) = \sum_{i=1}^M (C_{41} \cdot M + C_{42} \cdot M + \sum_{j=1}^M \sum_{d=1}^L C_{43}) = (C_{41} + C_{42})M^2 + C_{43}LM^2 = (C_{41} + C_{42} + C_{43}L)M^2.$$

$$T(N) = \max(T_1(N), T_2(N), T_3(N), T_4(N)) = (C_{41} + C_{42} + C_{43}L)M^2.$$

Như vậy, độ phức tạp tính toán của thuật toán là $O(N^2)$ (vì $M \leq N$).

* **Tính hiệu quả của thuật toán:** thông thường, $M \ll N$, nhưng trong trường hợp xấu nhất, $M=N$, có nghĩa rằng tất cả các điểm ảnh trong một ảnh có màu sắc khác nhau, không có bất kỳ điểm ảnh nào giống nhau (điều này rất khó xảy).

3.5.6. Thuật toán phát hiện các điểm ảnh dị thường trên ảnh màu RGB sử dụng kỹ thuật nhóm các điểm ảnh và tính toán đa luồng trên CPU hoặc song song trên GPU

- Độ phức tạp tính toán của thuật toán nhóm các điểm ảnh (hàng 1) là $T_1(N) = O(N)$.

- Độ phức tạp tính toán của thuật toán làm mịn ảnh (hàng 2) là $T_2(N) = O(N)$

- Độ phức tạp tính toán của thuật toán tìm các điểm ảnh đặc trưng (hàng 3) là $T_3(N) = O(N)$

- Thuật toán có bốn lệnh gán (hàng 4 và 5), những lệnh gán này đều có độ phức tạp tính toán là $O(1)$.

- Độ phức tạp tính toán của thuật toán tìm kiếm nhóm điểm ảnh (hàng 8) là $O(N)$.

- Thuật toán có ba vòng lặp lồng nhau, vòng lặp thứ nhất (hàng 6) thực hiện M lần (M là số nhóm các điểm ảnh tạo ra từ Thuật toán 3.1, tổng của các luồng là M), vòng lặp thứ hai (hàng 10) và thứ tư (hàng 15) thực hiện tối đa M lần, vòng lặp thứ ba (hàng 12) thực hiện L lần. Ta có:

$$T_4(N) = \sum_{i=1}^M (C_{41} \cdot M + \sum_{j=1}^M \sum_{d=1}^L C_{42}) = C_{41} \cdot M^2 + C_{42} \cdot L \cdot M^2 = (C_{41} + C_{42} \cdot L) M^2.$$

$$T(N) = \text{Max}(T_1(N), T_2(N), T_3(N), T_4(N)) = (C_{41} + C_{42} \cdot L) M^2.$$

Như vậy, độ phức tạp tính toán của thuật toán là $O(N^2)$ (vì $M \leq N$).

* **Tính hiệu quả của thuật toán:** tương tự các thuật toán trên, thông thường $M \ll N$, trường hợp xấu nhất $M=N$, nhưng rất khó xảy ra.

3.5.7. Thuật toán phát hiện các điểm ảnh dị thường trên ảnh màu MSI, HSI sử dụng kỹ thuật nhóm các điểm ảnh và tính toán đa luồng trên CPU hoặc song song trên GPU

- Độ phức tạp tính toán của thuật toán nhóm các điểm ảnh (hàng 1) là $T_1(N) = O(N)$. Vậy số phép tính là $f_1(N) = C_1 \cdot N$.

- Độ phức tạp tính toán của thuật toán làm mịn ảnh (hàng 2) là $T_2(N) = O(N)$. Vậy số phép tính là $f_2(N) = C_2 \cdot N$.

- Độ phức tạp tính toán của thuật toán tìm các điểm ảnh đặc trưng (hàng 3) là $T_3(N) = O(N)$. Vậy số phép tính là $f_3(N) = C_3 \cdot N$.

- Thuật toán có bốn lệnh gán (hàng 4 và 5), những lệnh gán này đều có độ phức tạp tính toán là $O(1)$. Vậy số phép tính là $f_4(N) = C_4$.

- Tổng số lần thực hiện trong vòng lặp thứ nhất và thứ hai (hàng 6 và hàng 7) là M ($M \leq N$). Vậy số phép tính là $f_5(N) = C_5 \cdot N$.

- Trong thân vòng lặp (từ hàng 9 đến hàng 20) thực hiện:

+ Thuật toán tìm kiếm nhóm điểm ảnh (hàng 9) có độ phức tạp là $O(N)$. Vậy số phép tính là $f_{51}(N) = C_{51} \cdot N$.

+ Từ hàng 10 đến hàng 20 là phân nhóm điểm ảnh (như thuật toán KDE), có độ phức tạp tính toán là $O(N)$. Vậy số phép tính là $f_{52}(N) = C_{52} \cdot N$.

Vậy độ phức tạp tính toán của thuật toán là:

$$\begin{aligned} f(N) &= f_1(N) + f_2(N) + f_3(N) + f_4(N) + f_5(N)(f_{51}(N) + f_{52}(N)) \\ &= C_1N + C_2N + C_3N + C_4 + C_5N(C_{51}N + C_{52}N) \\ &= C_1N + C_2N + C_3N + C_4 + (C_{51} + C_{52})C_5N^2 = O(N^2) \end{aligned}$$

* **Tính hiệu quả của thuật toán:** tương tự các thuật toán trên, thông thường $M \ll N$, trường hợp xấu nhất $M=N$, nhưng rất khó xảy ra.

Bảng 3.1: Bảng so sánh độ phức tạp tính toán của các thuật toán

Thuật toán	Độ phức tạp	Ghi chú
KDE (Thuật toán 2.1)	$O(N^2)$	Thời gian tính toán của Thuật toán KDE phụ thuộc hoàn toàn vào số điểm ảnh trong một ảnh (N).
GP-KDT1 (Thuật toán 3.5)	$O(N^2)$	Thời gian tính toán của thuật toán phụ thuộc vào số nhóm điểm ảnh (M). Trong trường hợp xấu nhất $M=N$ (trong một ảnh không có bất cứ điểm ảnh nào giống nhau), điều này khó xảy ra bởi trong tự nhiên, lớp phủ thực địa luôn có tính chất phân lớp đối tượng, lớp phủ càng đồng nhất số nhóm các điểm ảnh càng ít nên $M \ll N$. Ví dụ: Hình 3.6 có $N=11.373.204$ điểm ảnh, số nhóm điểm ảnh là 65.607 nhóm, có nghĩa rằng $M \approx 0.006 \times N$.
GP-KDT2 (Thuật toán 3.6)	$O(N^2)$	Tương tự Thuật toán GP-KDT1, thời gian tính toán của Thuật toán GP-KDT2 cũng phụ thuộc vào số nhóm điểm ảnh (M). Trong trường hợp xấu nhất $M=N$, có nghĩa là trong một ảnh không có bất cứ điểm ảnh nào giống nhau.

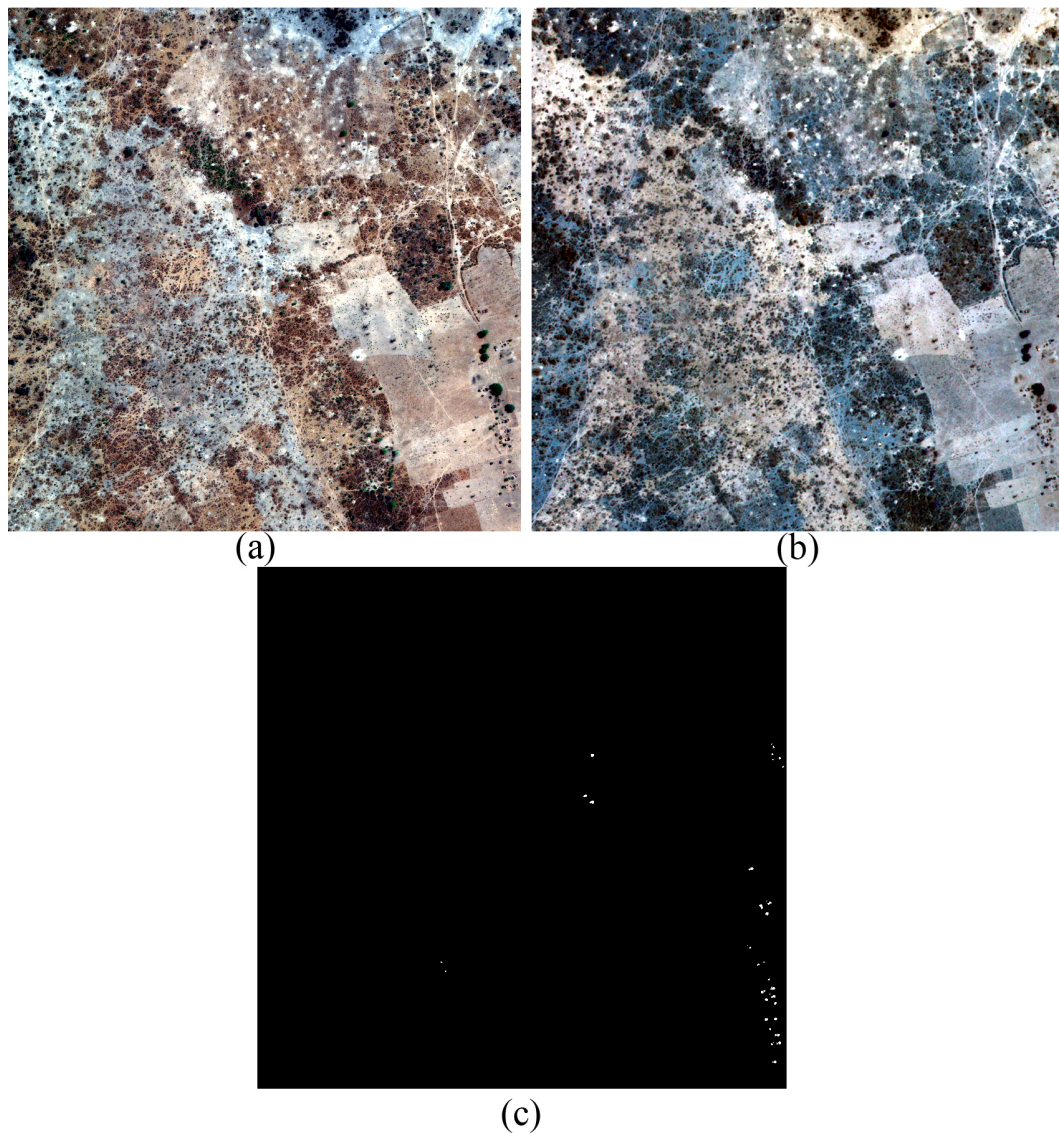
Bảng 3.1 – *Tiếp theo ở trang trước*

Thuật toán	Độ phức tạp	Ghi chú
GP-MC1 (Thuật toán 3.7)	$O(N^2)$	Thời gian tính toán của Thuật toán GP-MC1 cũng phụ thuộc vào số nhóm điểm ảnh (M).
GP-MC2 (Thuật toán 3.8)	$O(N^2)$	Thời gian tính toán của Thuật toán GP-MC2 cũng phụ thuộc vào số nhóm điểm ảnh (M).
GP-PC1 (Thuật toán 3.9)	$O(N^2)$	Thời gian tính toán của Thuật toán GP-PC1 cũng phụ thuộc vào số nhóm điểm ảnh (M).
GP-PC2 (Thuật toán 3.10)	$O(N^2)$	Thời gian tính toán của Thuật toán GP-PC2 cũng phụ thuộc vào số nhóm điểm ảnh (M).

3.6. Thử nghiệm và đánh giá kết quả đề xuất

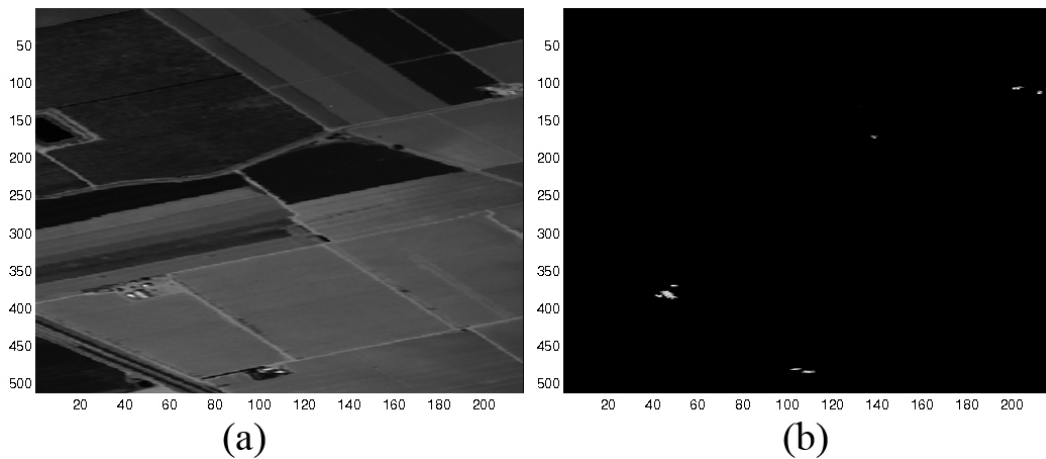
3.6.1. Dữ liệu thử nghiệm

Dữ liệu thử nghiệm gồm ba loại ảnh khác nhau, loại ảnh đầu tiên là ảnh màu có 3 kênh phổ (RGB), loại ảnh thứ 2 là ảnh đa phổ có 8 kênh phổ, và loại ảnh thứ 3 là ảnh siêu phổ 224 kênh phổ. Đầu tiên, sử dụng ảnh 3 kênh phổ và ảnh có 8 kênh phổ do DSTL cung cấp [24] đã được mô tả trong Chương 2. Ảnh 3 kênh phổ có mã là 6010_1_2 (gọi tắt là ảnh 3 kênh phổ) và ảnh 8 kênh phổ có mã là 6010_1_2_M.tif (gọi tắt là ảnh 8 kênh phổ). Ảnh 3 kênh phổ được thể hiện trên Hình 3.6(a)), Ảnh 8 kênh phổ thể hiện trên Hình 3.6(b); Hình 3.6(c) hiển thị những điểm ảnh dị thường (màu trắng, những điểm ảnh dị thường này là những công trình nhân tạo hỗn hợp trên thực địa) và những điểm ảnh bình thường (màu đen) của hai ảnh 6010_1_2 và 6010_1_2_M.



Hình 3.6: Các ảnh đa phổ: (a) Ảnh 3 kênh phổ mã số 6010_1_2; (b) Ảnh 8 kênh phổ mã số 6010_1_2_M (hình hiển thị là sự tổ hợp 3 kênh, kênh số 1, kênh số 2 và kênh số 3); (c) thể hiện những điểm ảnh dị thường (màu trắng) và những điểm ảnh bình thường (màu đen).

Sau đó, luận án thu thập ảnh siêu phổ 224 kênh phổ tại [47] (Hình 3.7(a)), cảnh ảnh này được thu tại Thung lũng Salinas, California bởi bộ cảm biến AVIRIS với 224 kênh phổ, độ phân giải không gian là 3,7 mét, kích thước ảnh là 512×217 pixels (gọi tắt là ảnh Salinas). Hình 3.7(b) thể hiện những điểm ảnh dị thường (màu trắng, những điểm ảnh dị thường này là những công trình nhân tạo được bao quanh bởi cánh đồng trồng các loại thực vật) và những điểm ảnh không dị thường (màu đen).



Hình 3.7: Ảnh siêu phổ: (a) Kênh 220 của ảnh siêu phổ Salinas 224 kênh phổ; (b) thể hiện những điểm ảnh dị thường (màu trắng) và những điểm ảnh bình thường (màu đen).

3.6.2. Phương pháp thử nghiệm

Tương ứng với mỗi một ảnh đã trình bày ở trên, chạy các thuật toán: Thuật toán KDE (Thuật toán 2.1); thuật toán tính toán đa luồng trên CPU sử dụng ngôn ngữ lập trình Intel TBB do Michailidis và các cộng sự đề xuất tại [86] (gọi tắt là thuật toán Intel TBB); thuật toán tính toán song song trên GPU CUDA cũng do Michailidis và các cộng sự đề xuất tại [87] (gọi tắt là thuật toán GPU CUDA); thuật toán IFGT do Yang và các cộng sự đề xuất trong [129] trên ảnh 3 kênh phổ và ảnh 8 kênh phổ. Thuật toán IFGT không thực hiện được trên ảnh 224 kênh phổ do ảnh này có số chiều dữ liệu lớn nên thuật toán đòi hỏi bộ nhớ lớn vượt khỏi khả năng quản lý của hệ điều hành. Cuối cùng, tương ứng với mỗi ảnh, thực thi các thuật toán do luận án đề xuất, bao gồm: nhóm các điểm ảnh và sử dụng cây kd-tree hỗ trợ quá trình tính toán PDF cho PUT theo Thuật toán 3.5 (gọi tắt là Thuật toán GP-KDT1 cho ảnh đầu vào là ảnh màu RGB) và Thuật toán 3.6 (gọi tắt là Thuật toán GP-KDT2 cho ảnh đa phổ 8 kênh phổ và ảnh siêu phổ 224 kênh phổ); Tính toán PDF đa luồng cho PUT trên CPU theo Thuật toán 3.7 (gọi tắt là Thuật toán GP-MC1 cho ảnh đầu vào là ảnh màu RGB) và Thuật toán 3.8 (gọi tắt là Thuật toán GP-MC2 cho ảnh đa phổ 8 kênh phổ và ảnh siêu phổ 224 kênh phổ). Tính toán PDF song song cho PUT trên nền GPU CUDA theo Thuật toán 3.9 (gọi tắt là Thuật toán GP-PC1 cho ảnh đầu vào

là ảnh màu RGB) và Thuật toán 3.10 (gọi tắt là Thuật toán GP-PC2 cho ảnh đa phổ 8 kênh phổ và ảnh siêu phổ 224 kênh phổ).

Sử dụng hàm nhân Hypercube (trong Bảng 2.1) với băng thông cố định $h = 10$ để kiểm nghiệm các thuật toán trên 3 ảnh này. Cấu hình máy tính dùng để chạy các thuật toán tính PDF là như sau:

- CPU: Intel Core i5-7400 3.00 GHz;
- Mainboard: MSI B150M MORTAR ARCTIC;
- RAM: DDR4 16GB;
- HDD: SSD BIOSTAR S100 - 240GB;
- Graphic: NVIDIA GeForce GTX 1070 Ti.
- Hệ điều hành: Windows 10 Enterprise 64bit.

3.6.3. Kết quả thử nghiệm

Thời gian thực thi của các thuật toán trên cả ba ảnh như đã mô tả trong phần 3.6.1 được thể hiện trên bảng 3.2. Các kết quả về thời gian tính toán của những thuật toán cho ta thấy rằng: thời gian tính toán của Thuật toán KDE trên ảnh màu RGB có 11.373.204 điểm ảnh tiêu tốn hết 21 ngày; trên ảnh 8 kênh phổ có 710.613 điểm ảnh tiêu tốn hết hơn 3 giờ; trên ảnh 224 kênh phổ có 111.104 điểm ảnh tiêu tốn hết hơn 9 phút.

Thời gian tính toán đa luồng trên CPU sử dụng công cụ lập trình Intel TBB do Michailidis và các cộng sự đề xuất tại [86] đã được rút ngắn hơn so với Thuật toán KDE. Nhìn vào Bảng 3.2 và Hình 3.8 chúng ta thấy, thời gian tính toán trung bình của thuật toán Intel TBB đã giảm 1,24 lần so với Thuật toán KDE. Tuy nhiên, phương pháp này vẫn tiêu tốn rất nhiều thời gian nên rất khó ứng dụng vào thực tế.

Đối với phương pháp tính toán song song trên GPU sử dụng công cụ lập trình GPU CUDA do Michailidis và các cộng sự đề xuất tại [87], thời gian tính toán đã giảm đi rất nhiều so với Thuật toán KDE. Nhìn vào Bảng 3.2 và Hình 3.9 chúng ta thấy, thời gian tính toán trung bình của phương pháp tính toán song song trên GPU CUDA đã giảm 55,88 lần so với Thuật toán

Bảng 3.2: Thời gian thực thi của các thuật toán trên các ảnh (tính bằng giây).

Thuật toán	6010_1_2	6010_1_2_M	Salinas
KDE (Thuật toán 2.1)	1.819.712	11.038	557
Intel TBB [86]	1.404.590	9.052	458
GPU CUDA [87]	21.609	253	14
GP-KDT1 (Thuật toán 3.5)	21		
GP-KDT2 (Thuật toán 3.6)		3.744	40
GP-MC1 (Thuật toán 3.7)	15		
GP-MC2 (Thuật toán 3.8)		548	8
GP-PC1 (Thuật toán 3.9)	5		
GP-PC2 (Thuật toán 3.10)		45	2

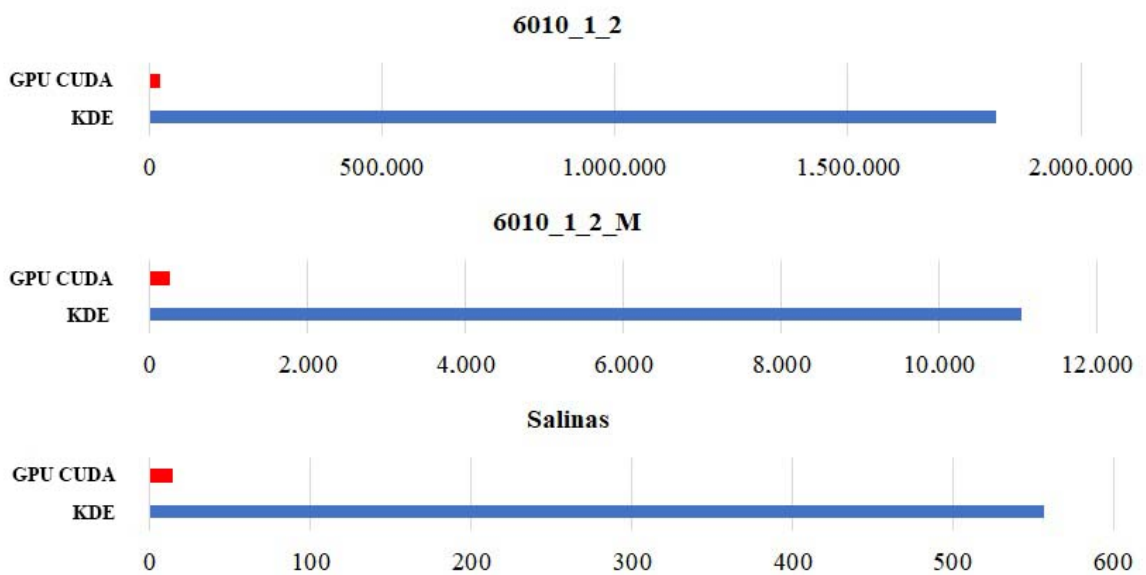
KDE. Cũng như phương pháp tính toán đa luồng trên CPU, thời gian tính toán của phương pháp này vẫn còn khá cao, thời gian tính toán trên ảnh 6010_1_2 hết 6 giờ nên vẫn rất khó ứng dụng vào thực tế.

Áp dụng thuật toán IFGT để hỗ trợ quá trình tính toán PDF cho PUT, đầu tiên phải chuẩn hóa dữ liệu về khoảng $[0, 1]$. Sau đó sử dụng thuật toán IFGT để tính toán PDF cho PUT với các thông số đầu vào như sau: bảng thông $h = 0.05$ số cụm $K = 100$ cho cả hai ảnh 6010_1_2 và 6010_1_2_M, lần lượt chọn $p = 5, 10, 15, 20, 25, 30$ và 35 (p là giá trị để ước tính lỗi nhỏ hơn độ chính xác mong muốn theo công thức (24) trong [129]) để tính toán cho ảnh 6010_1_2, $p = 5, 10, 15$ và 20 để tính toán cho ảnh 6010_1_2_M. Thời gian tính toán và hiệu suất phát hiện dị thường trên hai ảnh 6010_1_2 và 6010_1_2_M được thể hiện trên Bảng 3.4 và Bảng 3.5.

Quan sát trên Bảng 3.4 và Bảng 3.5 chúng ta thấy rằng độ chính xác phát



Hình 3.8: Biểu đồ so sánh thời gian chạy của thuật toán tính toán trực tiếp các điểm ảnh theo Thuật toán KDE và thuật toán tính toán đa luồng trên CPU do Michailidis và các cộng sự đề xuất tại [86].



Hình 3.9: Biểu đồ so sánh thời gian chạy của thuật toán tính toán trực tiếp các điểm ảnh theo Thuật toán KDE và thuật toán tính toán song song trên GPU do Michailidis và các cộng sự đề xuất tại [87].

hiện các điểm ảnh dị thường khi áp dụng thuật toán IFGT để tính PDF cho PUT đã giảm đi rất nhiều so với Thuật toán KDE và nó chịu ảnh hưởng lớn

Bảng 3.4: Thời gian và độ chính xác phát hiện các điểm ảnh dị thường trên ảnh 6010_1_2 khi áp dụng thuật toán IFGT tính toán PDF cho PUT

	Thời gian (s)	độ chính xác phát hiện
p=5	430	0,5814
p=10	3.094	0,6111
p=15	5.036	0,6260
p=20	11.011	0,6764
p=25	21.085	0,6901
p=30	35.432	0,7031
p=35	55.460	0,7098

Bảng 3.5: Thời gian và độ chính xác phát hiện các điểm ảnh dị thường trên ảnh 6010_1_2_M khi áp dụng thuật toán IFGT tính toán PDF cho PUT

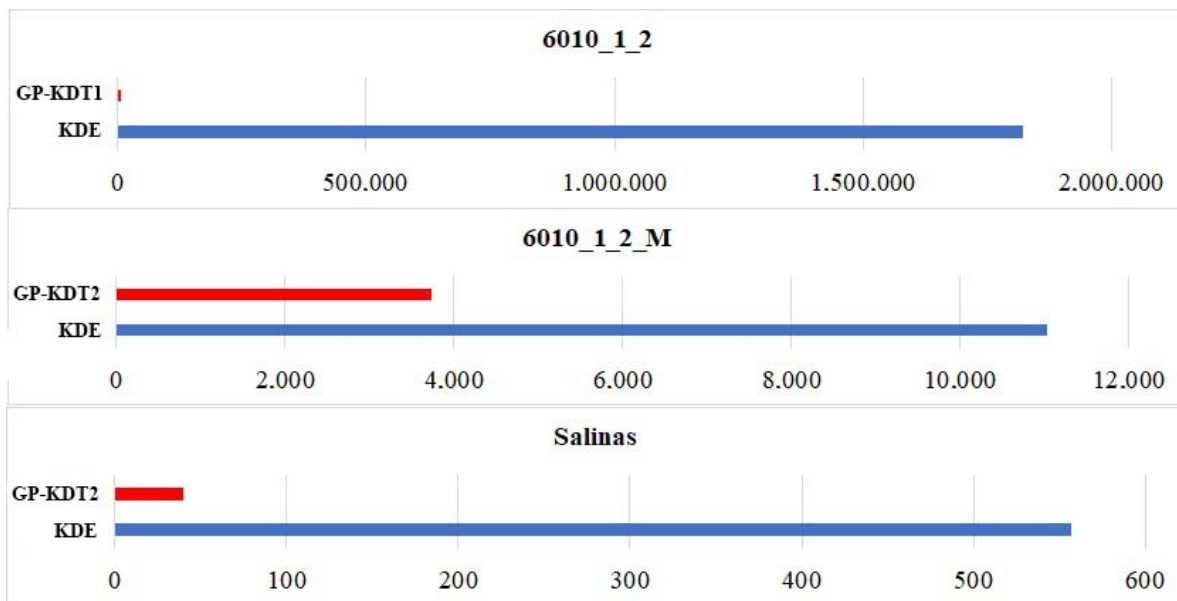
	Thời gian (s)	độ chính xác phát hiện
p=5	272	0,5122
p=10	9.518	0,5232
p=15	113.925	0,5294
p=20	780.355	0,5353

bởi tham số p . p tăng đồng nghĩa với độ chính xác của việc ước lượng hàm mật độ xác suất của PUT tăng dẫn đến độ chính xác phát hiện dị thường của thuật toán cũng tăng theo. Tuy nhiên, khi lựa chọn p cao thời gian để tính toán cũng tăng theo hàm số mũ.

Trên ảnh màu RGB mã số 6010_1_2, IFGT có thời gian tính toán nhanh gấp 4.232 lần so với thời gian tính toán của Thuật toán KDE khi chọn $p = 5$. Tuy nhiên độ chính xác phát hiện dị thường kém tới 32% so với Thuật toán KDE. Khi chọn $p = 35$ thời gian tính toán của IFGT đã tăng lên tương đối nhiều và chỉ nhanh hơn 33 lần so với Thuật toán KDE, độ chính xác phát hiện các điểm ảnh dị thường đã được cải thiện, chỉ còn kém hơn 19% so với Thuật toán KDE. Lúc này thời gian tính toán đã vượt quá 15 giờ, với thời gian tính toán như vậy sẽ không phù hợp cho công tác tìm kiếm cứu nạn nên luận án không thử tiếp các trường hợp p cao hơn.

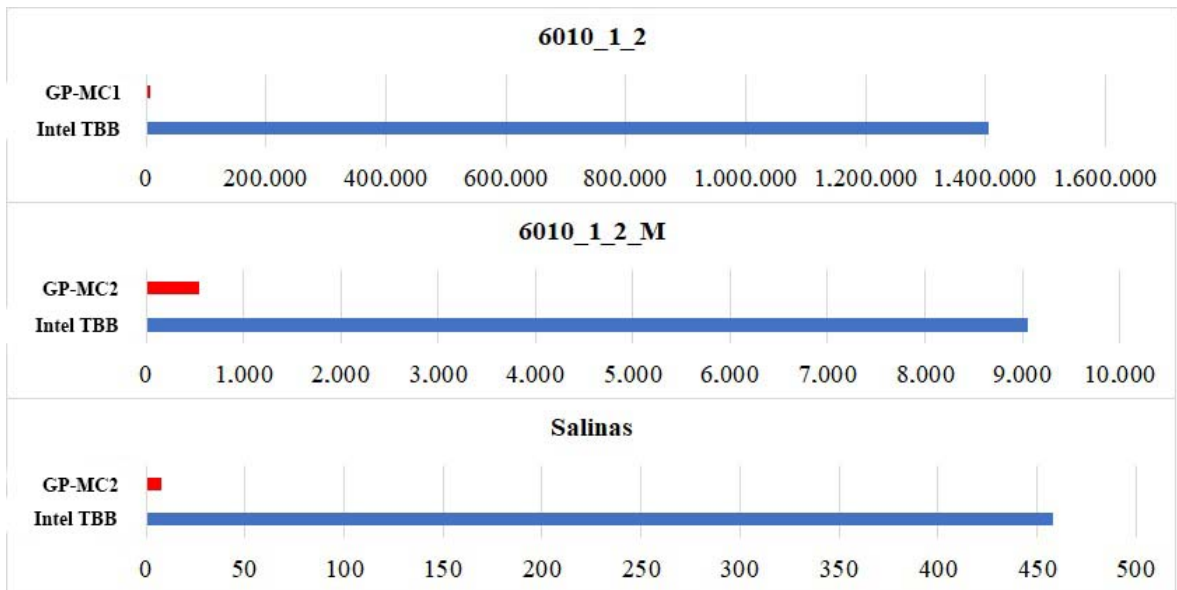
Trên ảnh đa phổ số 6010_1_2_M, IFGT có thời gian tính toán nhanh hơn 41 lần so với thời gian tính toán của Thuật toán KDE khi chọn $p = 5$.

Độ chính xác phát hiện dị thường giảm tới 35% so với Thuật toán KDE. Khi chọn $p = 20$ thời gian tính toán của IFGT đã tăng lên rất cao (thời gian chạy hết hơn 9 ngày) và chậm hơn 71 lần so với Thuật toán KDE, độ chính xác phát hiện các điểm ảnh dị thường không được cải thiện nhiều so với khi chọn $p = 5$. Như vậy, khi áp dụng thuật toán IFGT để hỗ trợ tính toán nhanh PDF cho PUT không hiệu quả với số chiều dữ liệu cao. Thời gian tính toán giảm cũng kéo theo độ chính xác phát hiện các điểm ảnh dị thường trên MSI giảm theo.

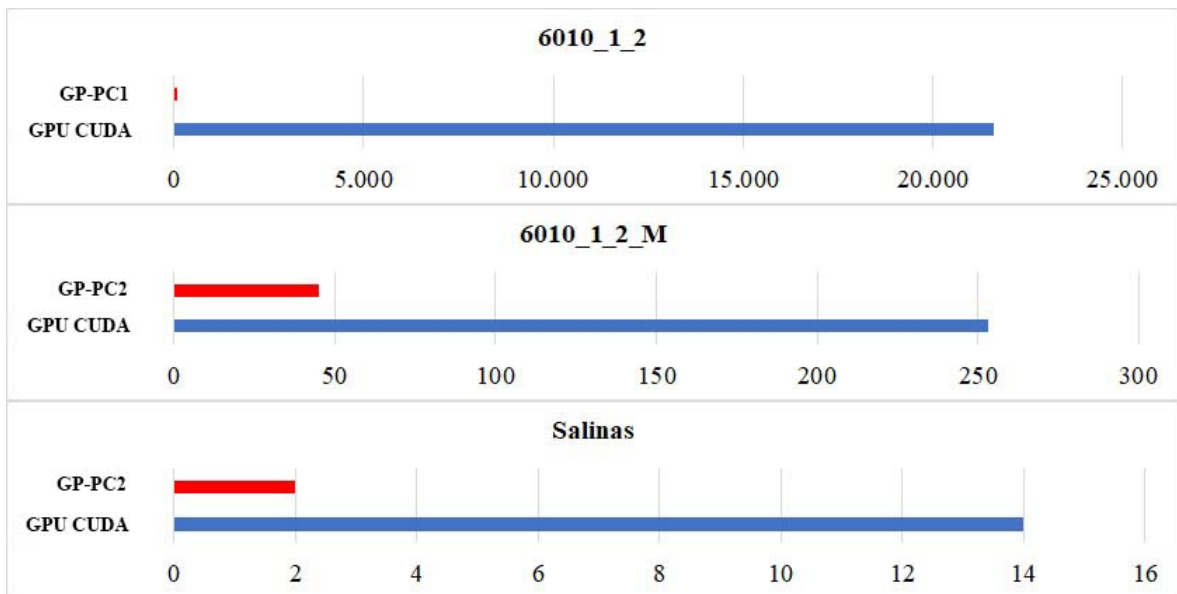


Hình 3.10: Biểu đồ so sánh thời gian chạy của Thuật toán KDE và thuật toán GP-KDT1, GP-KDT2

Xét hiệu quả về thời gian tính toán của ba thuật toán đã đề xuất, quan sát kết quả trên Bảng 3.2, Hình 3.10, Hình 3.11 và Hình 3.12 cho thấy thuật toán GP-KDT1, GP-KDT2, GP-MC1, GP-MC2, GP-PC1 và GP-CP2 đã có sự vượt trội hoàn toàn so với các thuật toán còn lại. Đặc biệt là trên ảnh màu RGB mã số 6010_1_2, thời gian tính toán của thuật toán GP-KDT1, GP-MC1, GP-PC1 đã giảm tới 99,999% so với Thuật toán KDE. So với thời gian tính toán của thuật toán Intel TBB, thuật toán GP-KDT1, GP-MC1, GP-PC1 đã giảm ít nhất 99,99%. Thời gian tính toán của các thuật toán GP-KDT1, GP-MC1, GP-PC1 đã giảm ít nhất 99,90% so với thuật toán GPU CUDA. Như vậy, trên ảnh màu RGB, các thuật toán đã đề xuất đạt hiệu quả cao về thời gian tính toán, có thể giảm thời gian tính toán tới 99,90% so với



Hình 3.11: Biểu đồ so sánh thời gian chạy của thuật toán Intel TBB và thuật toán GP-MC1, GP-MC2



Hình 3.12: Biểu đồ so sánh thời gian chạy của thuật toán GPU CUDA và thuật toán GP-PC1, GP-PC2

các thuật toán đã đề cập ở trên (xem chi tiết trên Bảng 3.6).

Trên ảnh đa phổ 8 kênh phổ mã 6010_1_2_M, thuật toán GP-KDT2 có thời gian tính toán chậm nhất nhưng cũng đã giảm được 66,08% so với thời gian tính toán của Thuật toán KDE, giảm được 58,64% so với thời gian tính

Bảng 3.6: So sánh thời gian thực thi của các thuật toán trên ảnh màu RGB mã số 6010_1_2.

Thuật toán	KDE	Intel TBB	GPU CUDA	GP-KDT1	GP-MP1
Intel TBB	↓22,8%				
GPU CUDA	↓98,81%	↓98,46%			
GP-KDT1	↓99,998%	↓99,990%	↓99,90%		
GP-MC1	↓99,999%	↓99,998%	↓99,90%	↓28,57%	
GP-PC1	↓99,999%	↓99,999%	↓99,98%	↓76,19%	↓66,67%

toán của Thuật toán Intel TBB. Thuật toán GP-MC2 có thời gian tính toán giảm được 94,71% so với thời gian tính toán của Thuật toán KDE, giảm được 93,55% so với thời gian tính toán của Thuật toán Intel TBB. Thuật toán GP-PC2 có thời gian tính toán nhanh nhất, giảm được 99,56% so với thời gian tính toán của Thuật toán KDE, giảm được 99,50% so với thời gian tính toán của Thuật toán Intel TBB và giảm được 82,21% so với thời gian tính toán của Thuật toán GPU CUDA. Như vậy, trên ảnh đa phổ 8 kênh phổ, các thuật toán đã đề xuất cũng đã đạt hiệu quả cao về thời gian tính toán, sử dụng thuật toán GP-PC2 có thể giảm thời gian tính toán tới 82,21% so với các thuật toán đã đề cập ở trên (xem chi tiết trên Bảng 3.7).

Bảng 3.7: So sánh thời gian thực thi của các thuật toán trên ảnh đa phổ 8 kênh phổ mã số 6010_1_2_M.

Thuật toán	KDE	Intel TBB	GPU CUDA	GP-KDT2	GP-MC2
Intel TBB	↓17,99%				
GPU CUDA	↓97,71%	↓97,21%		↓92,24%	↓53,83%
GP-KDT2	↓66,08%	↓58,64%			
GP-MC2	↓95,04%	↓93,95%		↓85,36%	
GP-PC2	↓99,59%	↓99,50%	↓82,21%	↓98,80%	↓91,79%

Trên ảnh siêu phổ 224 kênh phổ mã Salinas, thuật toán GP-KDT2 cũng có thời gian tính toán chậm nhất nhưng cũng đã giảm được 92,82% so với thời gian tính toán của Thuật toán KDE, giảm được 91,27% so với thời gian tính toán của Thuật toán Intel TBB. Thuật toán GP-MC2 có thời gian tính toán giảm được 98,56% so với thời gian tính toán của Thuật toán KDE, giảm

được 98,25% so với thời gian tính toán của Thuật toán Intel TBB và giảm được 42,86% so với thời gian tính toán của Thuật toán GPU CUDA. Thuật toán GP-PC2 vẫn là thuật toán có thời gian tính toán nhanh nhất, giảm được 99,64% so với thời gian tính toán của Thuật toán KDE, giảm được 99,56% so với thời gian tính toán của Thuật toán Intel TBB và giảm được 85,71% so với thời gian tính toán của Thuật toán GPU CUDA. Như vậy, trên ảnh siêu phổ 224 kênh phổ, các thuật toán đã đề xuất cũng đã đạt hiệu quả cao về thời gian tính toán, sử dụng thuật toán GP-PC2 có thể giảm thời gian tính toán tới 85,71% so với các thuật toán đã đề cập ở trên (xem chi tiết trên Bảng 3.8).

Bảng 3.8: So sánh thời gian thực thi của các thuật toán trên ảnh siêu phổ phổ 224 kênh phổ mã Salinas.

Thuật toán	KDE	Intel TBB	GPU CUDA	GP-KDT2	GP-MC2
Intel TBB	↓17,78%				
GPU CUDA	↓97,49%	↓96,94%		↓65%	
GP-KDT2	↓92,82%	↓91,27%			
GP-MC2	↓98,56%	↓98,25%	↓42,86%	↓80%	
GP-PC2	↓99,64%	↓99,56%	↓85,71%	↓95%	↓75%

Nếu so sánh thời gian tính toán giữa các chế độ đơn luồng trên CPU, đa luồng trên CPU và song song trên GPU chúng ta thấy rằng: ở chế độ đơn luồng, trên Hình 3.10 thể hiện thuật toán GP-KDT1, GP-KDT2 đã nhanh hơn Thuật toán KDE ít nhất 66,10% (trên ảnh 6010_1_2_M) và cao nhất là 99,999% (trên ảnh 6010_1_2). Trong chế độ đa luồng trên CPU, trên Hình 3.11 thể hiện thuật toán GP-MC1, GP-MC2 đã nhanh hơn thuật toán Intel TBB ít nhất 93,55% (trên ảnh 6010_1_2_M) và cao nhất là 99,999% (trên ảnh 6010_1_2). Trong chế độ song song trên GPU, trên Hình 3.12 thể hiện thuật toán GP-PC1, GP-PC2 đã nhanh hơn thuật toán GPU CUDA ít nhất 82,21% (trên ảnh 6010_1_2_M) và cao nhất là 99,999% (trên ảnh 6010_1_2).

3.7. Kết luận

Trong công tác tìm kiếm cứu nạn, việc rút ngắn thời gian xử lý dữ liệu và ra quyết định đồng nghĩa với việc giảm phí tổn tài chính, sức lực, tinh thần

và nâng cao khả năng sống sót của nạn nhân. Một số vấn đề đã được trình bày trong chương này gồm:

- *Thứ nhất*, khảo sát một số phương pháp nhằm tăng tốc độ tính toán PDF cho các điểm ảnh, làm nổi bật những ưu, khuyết điểm của các phương pháp. Đề xuất các bước và ba kỹ thuật tính toán PDF cho PUT để làm giảm thời gian phát hiện các điểm ảnh dị thường trên ảnh siêu phổ và đa phổ.
- *Thứ hai*, đã trình bày chi tiết bốn bước để tăng tốc độ tính toán của Thuật toán IKDE, gồm: nhóm các điểm ảnh có cùng giá trị; sắp xếp các nhóm điểm ảnh; tính toán PDF cho các điểm ảnh và phân lớp các điểm ảnh. Trong bước thứ ba đã trình bày ba kỹ thuật tính toán PDF cho PUT, gồm: ứng dụng cây kd-tree để hỗ trợ tìm nhanh những điểm ảnh thỏa mãn điều kiện $K(u) \neq 0$; tính toán PDF đa luồng trên CPU và tính toán PDF song song trên GPU.
- *Thứ ba*, kết quả cho thấy: kỹ thuật đề xuất mang lại hiệu quả cao về thời gian tính toán. Đặc biệt trên ảnh màu RGB, thời gian tính toán đã giảm tới hơn 99% so với các thuật toán như Thuật toán KDE, thuật toán Intel TBB, GPU CUDA. Trên ảnh đa phổ 8 kênh phổ và ảnh siêu phổ 224 kênh phổ thuật toán GP-PC2 đã giảm được ít nhất 82.21% so với thuật toán GPU CUDA do Michailidis và các cộng sự đề xuất năm 2013 [87]. Điểm đặc biệt trong đề xuất này, độ chính xác phát hiện các điểm ảnh dị thường của thuật toán không thay đổi so với Thuật toán KDE, đây là điều hết sức quan trọng khi áp dụng hỗ trợ công tác tìm kiếm cứu nạn.

KẾT LUẬN

Đề tài luận án "Phát triển một số kỹ thuật phát hiện dị thường trên ảnh viễn thám quang học phục vụ công tác tìm kiếm cứu nạn" đã giải quyết được một số vấn đề lý luận và thực tiễn đặt ra cho bài toán phát hiện những điểm ảnh dị thường trên ảnh viễn thám quang học, làm cơ sở phục vụ cho công tác tìm kiếm cứu nạn.

Luận án đã trình bày những nghiên cứu tổng quan về phát hiện dị thường trên ảnh viễn thám quang học, nêu lên những ưu, nhược điểm của các kỹ thuật phát hiện dị thường trên ảnh đa phổ, siêu phổ. Từ đó đưa ra những thách thức cần giải quyết. Trên cơ sở nghiên cứu lý thuyết và thử nghiệm, luận án đã đề xuất một số kỹ thuật nhằm tăng độ chính xác phát hiện dị thường và tăng tốc độ tính toán. Kết quả thử nghiệm đã cho thấy tính hiệu quả của các kỹ thuật đề xuất, từ đó có thể áp dụng vào giải quyết bài toán thực tế.

Những đóng góp mới của đề tài luận án:

- (1) Đề xuất kỹ thuật giảm tỷ lệ dương tính giả (làm tăng độ chính xác phát hiện) của thuật toán phát hiện dị thường dựa trên ước lượng mật độ nhân [120, 121] bằng cách khử bớt nhiễu trên ảnh gốc và xác định các khu vực giới hạn tính toán của Thuật toán KDE. Kết quả thử nghiệm cho thấy: Thuật toán IKDE tăng 0,58% về độ chính xác phát hiện dị thường so với Thuật toán KDE. Ngoài tăng về độ chính xác phát hiện dị thường, Thuật toán IKDE đã giảm được 29,15% tổng thời gian tính toán so với Thuật toán KDE. Giải pháp đề xuất cho thấy hiệu quả trong trường hợp ảnh đầu vào có nhiễu tăng cao, thuật toán đề xuất chỉ giảm 0,71% so với ảnh gốc, ngược lại, Thuật toán KDE bị giảm tới 4,17% độ chính xác phát hiện dị thường so với khi thực thi trên ảnh gốc. Như vậy, khả năng kháng nhiễu của kỹ thuật đề xuất tốt hơn so với thuật toán gốc.
- (2) Đề xuất kỹ thuật tăng tốc độ tính toán của Thuật toán IKDE bằng cách nhóm các điểm ảnh có giá trị giống nhau làm giảm kích thước dữ liệu cần tính toán, sắp xếp nhóm các điểm ảnh để tiện lợi cho quá trình tìm kiếm, áp dụng cây kd-tree để tìm kiếm nhanh các điểm ảnh, nhóm điểm

ảnh thỏa mãn hàm nhân; tính toán PDF đa luồng hoặc song song trên GPU. Từ kết quả thử nghiệm cho thấy kỹ thuật đề xuất mang lại hiệu quả cao về thời gian tính toán, đặc biệt trên ảnh màu RGB, thời gian tính toán đã giảm tới hơn 99% so với các thuật toán như Thuật toán KDE, thuật toán Intel TBB, GPU CUDA. Trên ảnh đa phổ 8 kênh phổ và ảnh siêu phổ 224 kênh phổ, thuật toán GP-PC2 đã giảm được ít nhất 82,21% so với thuật toán GPU CUDA do Michailidis và các cộng sự đề xuất năm 2013 [87]. Độ chính xác phát hiện các điểm ảnh dị thường vẫn được giữ nguyên so với Thuật toán KDE.

Kiến nghị về hướng nghiên cứu tiếp theo:

Trong tương lai, xử lý ảnh đa phổ, siêu phổ để phát hiện dị thường trong các ứng dụng viễn thám vẫn là một lĩnh vực nghiên cứu đặc biệt, xứng đáng được đồng đẳng cộng đồng các nhà khoa học trên thế giới quan tâm. Bởi vì dữ liệu ảnh viễn thám quang học, đặc biệt là ảnh siêu phổ mang nhiều thông tin có giá trị có thể hữu ích trong nhiều ứng dụng khác nhau.

Mặc dù luận án đã tập trung nghiên cứu lý thuyết cũng như đề xuất các kỹ thuật tăng độ chính xác phát hiện dị thường và tăng tốc độ tính toán nhằm giải quyết các thách thức. Tuy nhiên, còn một số vấn đề cần tiếp tục được nghiên cứu:

- (1) Do tính chất của các điểm ảnh "dị thường" là thưa thớt và hiếm khi đại diện cho ảnh, dẫn đến tình trạng mất cân bằng nghiêm trọng về số lượng giữa lớp "dị thường" và lớp "bình thường" trên ảnh. Sử dụng độ đo phù hợp để đánh giá độ chính xác phát hiện các điểm ảnh dị thường của các thuật toán như luận án đã sử dụng là một hướng tiếp cận [53], tuy nhiên, hướng tiếp cận này không giải quyết được gốc của vấn đề là dữ liệu bị mất cân bằng thực sự. Giải quyết được vấn đề mất cân bằng dữ liệu có thể làm tăng độ chính xác phát hiện các điểm ảnh dị thường của các thuật toán, bằng cách: thu thập thêm dữ liệu (Collect More Data); lấy lại mẫu trong tập dữ liệu (Resampling), gồm Over sampling (làm gia tăng mẫu của lớp "dị thường") và Under sampling (làm giảm số lượng quan sát của lớp "bình thường"); tạo mẫu tổng hợp (Generate Synthetic Samples); ... Ngoài ra, có thể sử dụng phương pháp phạt mô hình (Penalized Models), có nghĩa rằng: việc dự báo sai một quan sát

thuộc mẫu "bình thường" sẽ ít nghiêm trọng hơn so với dự báo sai một quan sát thuộc mẫu "dị thường", do đó, sẽ phạt nặng hơn đối với sai số dự báo thuộc nhóm "dị thường" bằng cách gán cho nó một trọng số lớn hơn trong công thức của hàm loss function.

- (2) Xem xét một điểm ảnh bất kỳ trong ảnh cho thấy: điểm ảnh này ít nhiều đều có mối quan hệ với những điểm ảnh xung quanh nó. Do đó, nghiên cứu, phân tích mối quan hệ ràng buộc dữ liệu, mối tương quan về mặt không gian của các điểm ảnh có thể làm tăng độ chính xác phát hiện dị thường của các thuật toán.
- (3) Trong thuật toán IKDE, việc lựa chọn một vùng cửa sổ cố định W có thể lãng phí thời gian tính toán (nếu dị thường có kích thước nhỏ) hoặc không phát hiện hết điểm dị thường (nếu dị thường có kích thước lớn). Quan sát Hình 2.6 có thể thấy: giá trị DoG của các điểm ảnh dị thường cao hơn (hoặc thấp hơn) những điểm ảnh xung quanh không phải là dị thường. Do đó, có thể sử dụng giá trị này để tìm kích thước cửa sổ W thích nghi cho PUT, tránh lãng phí thời gian tính toán.
- (4) Phát hiện cụm điểm ảnh dị thường cần kết hợp các đặc trưng hình dạng, kích thước trên cùng một kỹ thuật phát hiện các điểm ảnh dị thường.
- (5) Xây dựng bộ dữ liệu đặc trưng đầy đủ, phong phú sẽ cải thiện đáng kể việc nghiên cứu trong lĩnh vực này, cho phép so sánh tốt hơn giữa các phương pháp theo các kịch bản khác nhau.
- (6) Phát hiện các điểm ảnh dị thường theo thời gian thực có thể được thực hiện bằng cách xử lý dữ liệu trực tuyến hoặc cung cấp kết quả phát hiện trong thời gian thực. Những vấn đề này vẫn còn bỏ ngỏ và đã thu hút sự chú ý của cộng đồng khoa học, thể hiện trong các công trình nghiên cứu gần đây đã tập trung vào các kỹ thuật dựa trên biểu diễn và triển khai các phương pháp tiếp cận dựa trên mạng nơ-ron [54, 124].

CÁC CÔNG TRÌNH KHOA HỌC ĐÃ CÔNG BỐ

- [CT1] **Nguyễn Văn Phương**, Đào Khánh Hoài, Tống Minh Đức, Cao Thị Vinh (2020), "Phương pháp phát hiện dị thường trên ảnh viễn thám quang học ứng dụng trong công tác tìm kiếm cứu nạn", trong *Tạp chí Khoa học Trường Đại học Cần Thơ*, Tập 56, số 1A, tr. 21–36, ISSN: 1859-2333.
- [CT2] **Nguyễn Văn Phương**, Đào Khánh Hoài (2018), "Một số kỹ thuật phát hiện dị thường trên ảnh UAV ứng dụng trong công tác tìm kiếm cứu nạn", *Chuyên san Các công trình nghiên cứu phát triển Công nghệ thông tin và Truyền thông - Bộ Thông tin và Truyền thông*, Tập V-1, Số 39, tr. 1–8, ISSN: 1859-3534, DOI: 10.32913/rd-ict.vol1.no39.581.
- [CT3] **Nguyễn Văn Phương**, Đào Khánh Hoài, Tống Minh Đức (2020), "Tăng hiệu quả phát hiện dị thường trên ảnh UAV ứng dụng trong công tác tìm kiếm cứu nạn", trong *Tạp chí Khoa học và Công nghệ - Đại học Thái Nguyên*, Tập 225, Số 06, tr. 58 – 65, ISSN: 1859-2171.
- [CT4] **Nguyễn Văn Phương**, Đào Khánh Hoài, Tống Minh Đức (2019), "Tăng tốc độ phát hiện dị thường trên ảnh đa phổ và siêu phổ ứng dụng trong công tác tìm kiếm cứu nạn", *Chuyên san Các công trình nghiên cứu phát triển Công nghệ thông tin và Truyền thông - Bộ Thông tin và Truyền thông*, Tập 2019, Số 2, tr. 70-82, ISSN: 1859-3534, DOI: 10.32913/mic-ict-research-vn.v2019.n2.866.
- [CT5] **Nguyễn Văn Phương**, Đào Khánh Hoài, Tống Minh Đức (2020), "The Anomaly Detection Efficiency of Kernel Density Estimation Functions on UAV Images", trong Chuyên san *Section on Information and Communication Technology (ICT)*, No. 15, pp. 7-22, ISSN: 1859 – 0209.

TÀI LIỆU THAM KHẢO

Tiếng Việt:

- [1] Hoàng Việt Anh (2001), “Ứng dụng ảnh viễn thám siêu phổ (hyperspectral) vào việc theo dõi quá trình sa mạc hoá. Nghiên cứu thử nghiệm tại Tabernas, Tây Ban Nha”, *Vietnam Journal of Soil Science*, 5, pp. 34–38.
- [2] Nguyễn Đình Dương and Lê Minh Hằng (2021), *Xử lý ảnh số viễn thám*, Nhà xuất bản Khoa học tự nhiên và Công nghệ, 218 pp.
- [3] Cổng thông tin điện tử Ủy ban Quốc gia ứng phó sự cố thiên tai và tìm kiếm cứu nạn (2021), “Số liệu công tác ứng phó sự cố, thiên tai và tìm kiếm cứu nạn từ năm 2004 – 2021”, URL: <http://tkcn.gov.vn/?p=11625>.
- [4] N. T. Trung, V. V. Thỏa, and Đ. V. Đức (2015), “Một phương pháp tăng cường độ tương phản ảnh viễn thám dựa trên tiếp cận cục bộ”, *Các công trình nghiên cứu, phát triển và ứng dụng CNTT-TT*, V-2 (34), 83–97.

Tiếng Anh:

- [5] T. Adao et al. (2017), “Hyperspectral Imaging: A Review on UAV-Based Sensors, Data Processing and Applications for Agriculture and Forestry”, *Remote Sensing*, 9 (11), 30 pages.
- [6] C. C. Aggarwal (2017), *Outlier Analysis Second Edition*, Springer, Cham.
- [7] E. A. Ashton (1998), “Detection of Subpixel Anomalies in Multi-spectral Infrared Imagery Using an Adaptive Bayesian Classifier”, *IEEE Trans. Geosci. Remote Sens.*, 36 (2), pp. 506–517.
- [8] M. Bajic (2021), “Modeling and Simulation of Very High Spatial Resolution UXOs and Landmines in a Hyperspectral Scene for UAV Survey”, *Remote Sensing*, 13 (5).

- [9] A. Banerjee, P. Burlina., and C. Diehl (2006), “A Support Vector Method for Anomaly Detection in Hyperspectral Imagery”, *IEEE Trans. Geosci. Remote Sens.*, 44 (8), pp. 2282–2291.
- [10] J. L. Bentley (1975), “Multidimensional Binary Search Trees Used for Associative Searching”, *Communications of the ACM*, 18 (9), pp. 509–517.
- [11] James Bezdek (1981), *Pattern Recognition With Fuzzy Objective Function Algorithms*, ISBN: 978-1-4757-0452-5, DOI: 10.1007/978-1-4757-0450-1.
- [12] T. Bolukbasi and P. Tran (2012), *Outline Color Identification For Search And Rescue*, ECE-2012-07, Technical Report of Department of Electrical and Computer Engineering, Boston University.
- [13] M. J. Carlotto (2005), “A Cluster-Based Approach for Detecting Man-Made Objects and Changes in Imagery”, *IEEE Trans. Geosci. Remote Sens.*, 43 (4), pp. 374–387.
- [14] C. Chang and M. Hsueh (2006), “Characterization of anomaly detection in hyperspectral imagery”, *Sensor Review - SENS REV*, 26, pp. 137–146, DOI: 10.1108/02602280610652730.
- [15] C-I. Chang (2003), *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*, Springer Science and Business Media, 370 pp.
- [16] C-I. Chang and S-S. Chiang (2002), “Anomaly Detection and Classification for Hyperspectral imagery”, *IEEE Transactions on Geoscience and Remote Sensing*, 40 (6), pp. 1314–1325.
- [17] C. I. Chang, W. Xiong, and C.-H. Wen (2014), “A Theory of High-Order Statistics-Based Virtual Dimensionality for Hyperspectral Imagery”, *IEEE Transactions on Geoscience and Remote Sensing*, 52 (1), pp. 188–208.
- [18] S-S. Chang, C-I. Chang, and I. W. Ginsberg (2001), “Unsupervised Target Detection in Hyperspectral Images Using Projection Pursuit”, *IEEE Trans. Geosci. Remote Sens.*, 39 (7), pp. 1380–1391.

- [19] J. Y. Chen and I. S. Reed (1987), “A Detection Algorithm for Optical targets in Clutter”, *IEEE Trans. Aerosp. Electron. Syst.*, 23 (1), pp. 394–405.
- [20] Y. Chen, N. M. Nasrabadi, and T. D. Tran, “Hyperspectral image classification via kernel sparse representation”, in: *Proc. IEEE Int. Conf. Image Process., Brussels, Belgium*, 2011, 1233—1236.
- [21] Y. Chen, N. M. Nasrabadi, and T. D. Tran (2011), “Simultaneous joint sparsity model for target detection in hyperspectral imagery”, *IEEE Geoscience and Remote Sensing Letters*, 8 (4), pp. 676–680.
- [22] C. Corbane, L. Najman, and E. Pecoul (2010), “A complete processing chain for ship detection using optical satellite imagery”, *International Journal of Remote Sensing*, 31 (22), 5837—5854.
- [23] L. Devroye (1985), *Nonparametric Density Estimation: The L1 View*.
- [24] *Dstl Satellite Imagery Feature Detection*, URL: <https://www.kaggle.com/c/dstl-satellite-imagery-feature-detection>.
- [25] B. Du and L. Zhang (2011), “Random-Selection-Based Anomaly Detector for Hyperspectral Imagery”, *IEEE Transactions on Geoscience and Remote Sensing*, 49 (5), pp. 1578–1589.
- [26] Q. Du and I. Kopriva (2008), “Automated Target Detection and Discrimination Using Constrained Kurtosis Maximization”, *IEEE Geosci. Remote Sens. Letters*, 5 (1), pp. 1380–1391.
- [27] O. Duran and M. Petrou (2005), “A time-efficient clustering method for pure class selection”, *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 1, pp. 510–513.
- [28] O. Duran and M. Petrou (2007), “A Time-Efficient Method for Anomaly Detection in Hyperspectral Images”, *IEEE Trans. Geosci. Remote Sens.*, 45 (12), pp. 3894–3904.
- [29] O. Duran et al., “Anomaly Detection Through Adaptive Background Class Extraction From Dynamic Hyperspectral Data”, in: *Proceedings of the 7th Nordic Signal Processing Symposium - NORSIG 2006, Reykjavik*, 2006, pp. 234–237.

- [30] M. T. Eismann, J. Meola, and R. C. Hardie (2008), “Hyperspectral change detection in the presence of diurnal and seasonal variations”, *IEEE Transactions on Geoscience and Remote Sensing*, 46, 237–249.
- [31] A. Elgammal, R. Duraiswami, and LS Davis (2003), “Efficient Kernel density estimation using the Fast Gauss Transform with applications to color modeling and tracking”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25, pp. 1499–1504.
- [32] K. S. Ettabaa and M. B. Salem (2016), *Anomaly detection in hyperspectral imagery: an overview*, pp. 1587–1606.
- [33] F. Gao et al. (2014), “Moving Vehicle Information Extraction from Single-Pass WorldView-2 Imagery Based on ERGAS-SNS Analysis”, *Remote Sensing*, 6 (7), pp. 6500–6523.
- [34] M. Govender, K. Chetty, and H. Bulcock (2007), “A review of hyperspectral remote sensing and its application in vegetation and water resource studies”, *Water SA*, 33 (2), pp. 145–151.
- [35] A.A. Gowen et al. (2007), “Hyperspectral imaging – an emerging process analytical tool for food quality and safety control”, *Trends in Food Sci. Technol.*, 18 (12), pp. 590–598.
- [36] H. Grahn and P. Geladi (2007), *Techniques and Applications of Hyperspectral Image Analysis*, John Wiley and Sons, 390 pp.
- [37] S. I. Grossman (2015), *An automated directed spectral search methodology for small target detection*, ProQuest Dissertations And Theses, Thesis (Ph.D.), George Mason University, 191 pp.
- [38] Y. Gu, Y. Liu, and Y. Zhang (2008), “A selective KPCA algorithm based on high-order statistics for anomaly detection in hyperspectral imagery”, *Geoscience and Remote Sensing Letters*, 5 (1), pp. 43–47.
- [39] Q. Guo et al. (2014), “Weighted-RXD and Linear Filter-Based RXD: Improving Background Statistics Estimation for Anomaly Detection in Hyperspectral Imagery”, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7 (6), pp. 2351–2366.

- [40] P. Gurram and H. Kwon (2011), “Support-vector-based hyperspectral anomaly detection using optimized kernel parameters”, *IEEE Geoscience and Remote Sensing Letters*, 8 (6), 1060—1064.
- [41] T. Harris et al., “Spectral Target Detection for Detecting and Characterizing Floating Marine Debris”, in: *American Geophysical Union Fall Meeting*, 2012.
- [42] J. C Harsanyi (1993), *Detection and classification of subpixel spectral signatures in hyperspectral image sequences*, University of Maryland Baltimore County, 232 pp.
- [43] G. H. Hazel (2000), “Multivariate Gaussian MRF for Multispectral Scene Segmentation and Anomaly Detection”, *IEEE Trans. Geosci. Remote Sens.*, 38 (3), pp. 1199–1211.
- [44] G. Healey and D. Slater (1999), “Models and methods for automated material identification in hyperspectral imagery acquired under unknown illumination and atmospheric conditions”, *IEEE Transactions on Geoscience and Remote Sensing*, 37 (11), pp. 2706–2717.
- [45] J. A. Padrón Hidalgo et al. (2021), “Efficient Nonlinear RX Anomaly Detectors”, *IEEE Geoscience and Remote Sensing Letters*, 18 (2), pp. 231–235.
- [46] P. J. Huber (1985), “Projection Pursuit”, *Ann. Statist.*, 13 (2), pp. 435–475.
- [47] *Hyperspectral Remote Sensing Scenes*, URL: [http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral _Remote _Sensing_Scenes](http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral%20Remote%20Sensing%20Scenes).
- [48] P. Hytla et al. (2007), “Anomaly Detection in Hyperspectral Imagery: A Comparison of Methods Using Seasonal Data”, *Proceedings of SPIE - The International Society for Optical Engineering*, 2 (1), pp. 656506–1–11.
- [49] JA. Hyvärinen and E. Oja (2000), “Independent Component Analysis: Algorithms and Applications”, *Neural Networks Research Centre*, 13 (4-5), pp. 411–430.

- [50] A. Ifarraguerri and C-I. Chang (2000), “Unsupervised Hyperspectral Image Analysis with Projection Pursuit”, *IEEE Trans. Geosci. Remote Sens.*, 38 (6), pp. 2529–2538.
- [51] M. Imani (2018), “3D Gabor based hyperspectral anomaly detection”, *AUT Journal of Modeling and Simulation*, 50 (2), pp. 189–194.
- [52] R. Ivan and A. Krtalić (2021), “Systematic Review of Anomaly Detection in Hyperspectral Remote Sensing Applications”, *Applied Sciences*, 11 (11).
- [53] L. Jeni, J. Cohn, and D. T. Fernando, “Facing Imbalanced Data - Recommendations for the Use of Performance Metrics”, in: vol. 2013, Sept. 2013, DOI: 10.1109/ACII.2013.47.
- [54] K. Jiang et al. (2020), “Semisupervised Spectral Learning With Generative Adversarial Network for Hyperspectral Anomaly Detection”, *IEEE Transactions on Geoscience and Remote Sensing*, 58 (7), pp. 5224–5236.
- [55] A. Krtalić and M. Bajić (2019), “Development of the TIRAMISU Advanced Intelligence Decision Support System”, *European Journal of Remote Sensing*, 52 (1), pp. 40–55.
- [56] A. Krtalić et al. (2020), “The AIDSS Module for Data Acquisition in Crisis Situations and Environmental Protection”, *Sensors*, 20 (5).
- [57] Ivan Kutsir, *Fastest Gaussian Blur (in linear time)*, URL: <http://blog.ivank.net/fastest-gaussian-blur.html> (visited on 03/20/2022).
- [58] H. Kwon, S. Z. Der, and N. M. Nasrabadi (2003), “Adaptive anomaly detection using subspace separation for hyperspectral imagery”, *Opt. Eng.*, 42 (11), pp. 3342–3351.
- [59] H. Kwon and N.M. Nasrabadi (2005), “Kernel RX-Algorithm: A Nonlinear Anomaly Detector for Hyperspectral Imagery”, *IEEE Trans. Geosci. Remote Sens.*, 43 (2), pp. 388–397.

- [60] E. L. Lehmann (1993), “The Fisher, Neyman-Pearson theories of testing hypotheses: One theory or two?”, *Journal of the American Statistical Association*, 88 (424), pp. 1242–1249.
- [61] F. Li et al. (2018), “Exploiting structured sparsity for hyperspectral anomaly detection”, *IEEE Transactions on Geoscience and Remote Sensing*, 56 (7), pp. 4050–4064.
- [62] W. Li and Q. Du, “Unsupervised nearest regularized subspace for anomaly detection in hyperspectral imagery”, in: *In Proceedings of the Geoscience and Remote Sensing Symposium, Melbourne, Australia*, 2014, 1055—1058.
- [63] W. Li and Q. Du (2015), “Collaborative representation for hyperspectral anomaly detection”, *IEEE Trans. Geosci. Remote Sens.*, 53 (3), 1463—1474.
- [64] W. Li, S. Prasad, and J. E. Fowler (2013), “Integration of Spectral–Spatial Information for Hyperspectral Image Reconstruction From Compressive Random Projections”, *IEEE Geoscience and Remote Sensing Letters*, 10 (6), pp. 1379–1383.
- [65] W. Li, G. Wu, and Q. Du (2017), “Transferred deep learning for anomaly detection in hyperspectral imagery”, *IEEE Geoscience and Remote Sensing Letters*, 14 (5), pp. 597–601.
- [66] “Chapter 1 - A systematic view of remote sensing”, in: *Advanced Remote Sensing (Second Edition)*, ed. by S. Liang and J. Wang, Second Edition, Academic Press, 2020, pp. 1–57, ISBN: 978-0-12-815826-5, DOI: <https://doi.org/10.1016/B978-0-12-815826-5.00001-5>, URL: <https://www.sciencedirect.com/science/article/pii/B9780128158265000015>.
- [67] T. Lillesand, Kiefer R. W., and J. Chipman (2015), *Remote Sensing and Image Interpretation, 7th Edition*, Wiley, 736 pp.
- [68] Q. Ling et al. (2019), “A Constrained Sparse Representation Model for Hyperspectral Anomaly Detection”, *IEEE Transactions on Geoscience and Remote Sensing*, 57 (4), pp. 2358–2371.

- [69] W. Liu and C-I. Chang, “A nested spatial window-based approach to target detection for hyperspectral imagery”, in: *IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium, Anchorage, AK*, 2004, p. 268.
- [70] W-M. Liu and Ch-I. Chang (2013), “Multiple-Window Anomaly Detection for Hyperspectral Imagery”, *Selected Topics in IEEE Journal of Applied Earth Observations and Remote Sensing*, 6 (2), pp. 644–658.
- [71] D.G. Lowe (2004), “Distinctive image features from scale-invariant keypoints”, *Int. J. Comput. Vision*, 60 (2), 91—110.
- [72] Guolan Lu and Baowei Fei (2014), “Medical hyperspectral imaging: a review”, *Journal of Biomedical Optics*, 19 (1), pp. 1 –24.
- [73] S. Lukasik, “Parallel Computing of Kernel Density Estimates with MPI”, in: *Conference: Computational Science - ICCS 2007, 7th International Conference, Beijing, China, 2007*.
- [74] N. M. Nasrabadi M. T. Eismann A. D. Stocker (2009), “Automated Hyperspectral Cueing for Civilian Search and Rescue”, *Proceedings of the IEEE*, 97 (6), pp. 1031–1055.
- [75] D. Ma, Y. Yuan, and Q. Wang (2018), “Hyperspectral anomaly detection via discriminative feature learning with multiple-dictionary sparse representation”, *Remote Sensing*, 10 (5), p. 745.
- [76] N. Ma et al. (2018), “An unsupervised deep hyperspectral anomaly detector”, *Sensors (Basel)*, 18 (3), p. 693.
- [77] D. Manolakis (2005), “Taxonomy of detection algorithms for hyperspectral imaging applications”, *Optical Engineering*, 44 (6), pp. 066403 1–11.
- [78] D. Manolakis, D. Marden, and G. A. Shaw (2003), “Hyperspectral Image Processing for Automatic Target Detection Applications”, *IEEE Signal Process. Mag.*, 14 (1), pp. 79–116.
- [79] D. Manolakis and G. Shaw (2002), “Detection algorithms for hyperspectral imaging applications”, *IEEE Signal Process. Mag.*, 19, 29—43.

- [80] T. Marshall and L. N. Perkins (2015), *Color Outline Detection For Search And Rescue*, ECE-2015-01, Technical Report of Department of Electrical and Computer Engineering, Boston University.
- [81] J. F. Mas (1999), “Monitoring land-cover changes: A comparison of change detection techniques”, *Int. J. Remote Sens.*, 20 (1), 139–152.
- [82] S. Matteoli et al. (2013), “A Locally Adaptive Background Density Estimator: An Evolution for RX-Based Anomaly Detectors”, *IEEE Geoscience and Remote Sensing Letters*, 11 (1), pp. 323–327.
- [83] S. Matteoli et al. (2014), “Background Density Nonparametric Estimation With Data-Adaptive Bandwidths for the Detection of Anomalies in Multi-Hyperspectral Imagery”, *IEEE Geoscience and Remote Sensing Letters*, 11, pp. 163–167.
- [84] L. Meng and J. P. Kerekes (2012), “Object Tracking Using High Resolution Satellite Imagery”, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5, pp. 146–152.
- [85] S. Mercan and M. Alam (2011), “Anomaly detection in hyperspectral imagery using Stable Distribution”, *Proceedings of SPIE - The International Society for Optical Engineering*, 8049.
- [86] P. D. Michailidis and K. G. Margaritis, “Parallel Computing of Kernel Density Estimation with Different Multi-core Programming Models”, in: *2013 21st Euromicro International Conference on Parallel, Distributed, and Network-Based Processing*, 2013, pp. 77–85.
- [87] P. D. Michailidis and K. G. Margaritis (2013), “Accelerating Kernel Density Estimation on the GPU Using the CUDA Framework”, *Applied Mathematical Sciences*, 7 (30), pp. 1447–1476.
- [88] J. M. Molero et al. (2013), “Analysis and optimizations of global and local versions of the rx algorithm for anomaly detection in hyperspectral data”, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6 (2), pp. 801–814.

- [89] B.S. Morse, D. Thornton, and M. A. Goodrich, “Color anomaly detection and suggestion for wilderness search and rescue”, in: *2012 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2012, pp. 455–462, DOI: 10.1145/2157689.2157837.
- [90] N. M. Nasrabadi, “Penalized Spectral Matched Filter for Target Detection in Hyperspectral Imagery”, in: *2007 IEEE International Geoscience and Remote Sensing Symposium, Barcelona, Spain, 2007*, pp. 4830–4833.
- [91] E. Parzen (1962), “On the estimation of a probability density function and mode”, *The Annals of Mathematical Statistics*, 33 (3), pp. 1065–1076.
- [92] Jing Peng and Chaojian Shi, “Remote Sensing Application in the Maritime Search and Rescue”, in: *Remote Sensing*, ed. by Boris Escalante-Ramirez, Rijeka: IntechOpen, 2012, chap. 17.
- [93] B. Penn, “Using self-organizing maps for anomaly detection in hyperspectral imagery”, in: *Proceedings, IEEE Aerospace Conference, Big Sky, MT, USA, 2002*, pp. 1531–1535.
- [94] J. Plaza et al. (2009), “Massively Parallel Processing of Remotely Sensed Hyperspectral Images”, *Proceedings of SPIE - The International Society for Optical Engineering*, 7455.
- [95] M. Ramachandran and W. Moik (2013), *Outline Color Identification For Search And Rescue*, ECE-2013-03, Technical Report of Department of Electrical and Computer Engineering, Boston University.
- [96] I. S. Reed and X. Yu (1990), “Adaptive Multiple-Band CFAR Detection of An Optical Pattern with Unknown Spectral Distribution”, *IEEE Trans. Acoust. Speech Signal Process.*, 38 (10), pp. 1760–1770.
- [97] H. Ren et al. (2006), “Automatic target recognition for hyperspectral imagery using high-order statistics”, *IEEE Transactions on Aerospace and Electronic Systems*, 42 (4), pp. 1372–1385.

- [98] Lang Ren, Liaoying Zhao, and Yulei Wang (2020), “A Superpixel-Based Dual Window RX for Hyperspectral Anomaly Detection”, *IEEE Geoscience and Remote Sensing Letters*, 17 (7), pp. 1233–1237.
- [99] M. Rosenblatt (1956), “Remarks on Some Nonparametric Estimates of a Density Function”, *The Annals of Mathematical Statistics*, 27 (3), pp. 832–837.
- [100] M. B. Salem, K. S. Ettabaa, and M. A. Hamdi, “Anomaly detection in hyperspectral imagery: an overview”, in: *International Image Processing, Applications and Systems Conference*, 2014, pp. 1–6.
- [101] N. P. School (2012), *Detection of Subpixel Submerged Mine-Like Targets in Worldview-2 Multispectral Imagery*, CreateSpace Independent Publishing Platform, 108 pp.
- [102] R. A. Schowengerdt (2006), *Remote Sensing: Models and Methods for Image Processing Third Edition*, Academic Press, 560 pp.
- [103] S. M. Schweizer and J. M. F. Moura (2000), “Hyperspectral Imagery: Clutter Adaptation in Anomaly Detection”, *IEEE Trans. Geosci. Remote Sens.*, 46 (5), pp. 1855–1871.
- [104] B. Schölkopf and A. J. Smola (2001), *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*, The MIT Press, 644 pp.
- [105] B. Schölkopf et al. (2001), “Estimating the support of a High Dimensional Distribution”, *Neural Computation*, 13, pp. 1443–1471.
- [106] B. Silverman (1982), “Algorithm AS 176: Kernel density estimation using the fast Fourier transform”, *Applied Statistics*, 31 (1), pp. 93–99.
- [107] A. Singh (1989), “Digital change detection techniques using remotely sensed data”, *Int. J. Remote Sens.*, 10 (6), 989–1003.
- [108] T. E. Smetek and K. W. Bauer, “Finding Hyperspectral Anomalies Using Multivariate Outlier Detection”, in: *2007 IEEE Aerospace Conference, Big Sky, MT*, 2007, pp. 1–24.

- [109] D. W. J. Stein et al. (2002), “Anomaly Detection from Hyperspectral Imagery”, *IEEE Signal Process. Mag.*, 19 (1), pp. 58–69.
- [110] C. M. Stellman et al. (2000), “Real-time hyperspectral detection and cuing”, *Opt. Eng.*, 39 (7), pp. 1928–1935.
- [111] B. Stevenson et al. (2005), “Design and performance of the civil air patrol ARCHER hyperspectral processing system”, *Proceedings of SPIE - The International Society for Optical Engineering*, 5806, 731–742.
- [112] Hongjun Su et al. (2018), “Hyperspectral Anomaly Detection Using Collaborative Representation With Outlier Removal”, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11 (12), pp. 5029–5038.
- [113] P-H. Suen, G. Healy, and D. Slater (2001), “The Impact of Viewing Geometry on Material Discriminability in Hyperspectral Images”, *IEEE Transactions on Geoscience and Remote Sensing*, 39 (7), pp. 352–358.
- [114] Ashkan Taghipour and Hassan Ghassemian (2019), “Hyperspectral anomaly detection using spectral–spatial features based on the human visual system”, *International Journal of Remote Sensing*, 40 (23), pp. 8683–8704.
- [115] K. Tan et al. (2019), “Anomaly detection for hyperspectral imagery based on the regularized subspace method and collaborative representation”, *Remote Sensing*, 11 (11), p. 138.
- [116] D. M. J. Tax and R. P. W. Duin (1999), “Support Vector Domain Description”, *Pattern Recognition Letters*, 20, pp. 1191–1199.
- [117] D. M. J. Tax and R. P. W. Duin (2004), “Support Vector Data Description”, *Machine Learning*, 54 (1), pp. 45–66.
- [118] P. Trebuña and J. Halčinová (2012), “Experimental Modelling of the Cluster Analysis Processes”, *Procedia Engineering*, 48, 673–678.
- [119] M. Vafadar and H. Ghassemian (2018), “Anomaly detection of hyperspectral imagery using modified collaborative representation”, *IEEE Geoscience and Remote Sensing Letters*, 15 (4), 577–581.

- [120] T. Veracini et al., “An anomaly detection architecture based on a data-adaptive density estimation”, in: *011 3rd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Lisbon, 2011*, pp. 1–4.
- [121] T. Veracini et al. (2011), “Nonparametric Framework for Detecting Spectral Anomalies in Hyperspectral Images”, *IEEE Geoscience and Remote Sensing Letters*, 8 (4), pp. 666–670.
- [122] M. Muller W. Hardle A. Werwatz and S. Sperlich (2004), *Nonparametric Density Estimation, In: Nonparametric and Semiparametric Models*, pp. 39–83.
- [123] R. G. Wallacea, D. W. Affensa, and S. W. McCandless (1998), “Search and rescue from space”, *Part of the SPIE Conference on Automatic Target Recognition VIII*, 3371, pp. 174–184.
- [124] S. Wang et al. (2021), “Auto-AD: Autonomous Hyperspectral Anomaly Detection Network Based on Fully Convolutional Autoencoder”, *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–14.
- [125] Y. Wu et al. (2019), “Approximate computing for onboard anomaly detection from hyperspectral images”, *Journal of Real-Time Image Processing*, 16 (1), pp. 99–114.
- [126] J. Xiong et al., “Kernel sparse representation for anomaly detection in hyperspectral imagery”, in: *International Conference on Advances in Image Processing (ICAIP'18), Chengdu, China, 2018*, pp. 106–110.
- [127] Y. Xu et al. (2016), “Anomaly Detection in Hyperspectral Images Based on Low-Rank and Sparse Representation”, *IEEE Transactions on Geoscience and Remote Sensing*, 54 (4), pp. 1990–2000.
- [128] L. Yan et al. (2019), “A novel two-stage deep learning-based small-object detection using hyperspectral images”, *Optical Review*, 26, 597–606.
- [129] C. Yang et al., “Improved fast gauss transform and efficient kernel density estimation”, in: *Proceedings Ninth IEEE International Conference on Computer Vision, Nice, France, 2003*.

- [130] Y. Yang et al. (2019), “Hyperspectral Anomaly Detection via Dictionary Construction-Based Low-Rank Representation and Adaptive Weighting”, *Remote Sensing*, 11 (2).
- [131] Z. Yuan et al. (2014), “Local Sparsity Divergence for Hyperspectral Anomaly Detection”, *IEEE Geoscience and Remote Sensing Letters*, 11 (10), pp. 1697–1701.
- [132] K. Tan Z. Hou Y. Chen and P. Du (2018), “Novel hyperspectral anomaly detection methods base on unsupervised nearest regularized subspace”, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-3, pp. 539–54.
- [133] L. Zhang and B. Cheng (2019), “A stacked autoencoders-based adaptive subspace model for hyperspectral anomaly detection”, *Infrared Physics & Technology*, 96, 52–60.
- [134] X. Zhang et al. (2021), “Spectral-Difference Low-Rank Representation Learning for Hyperspectral Anomaly Detection”, *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–14.
- [135] C. Zhao, X. Wang, and G. Zhao (2017), “Detection of hyperspectral anomalies using density estimation and collaborative representation”, *Remote Sensing Letters*, 8 (11), 1025–1033.
- [136] C. Zhao, X. F. Yao, and Y. Yan (2017), “Modified Kernel RX Algorithm Based on Background Purification and Inverse-of-Matrix-Free Calculation”, *IEEE Geoscience and Remote Sensing Letters*, 14 (4), pp. 544–548.
- [137] L. Zhu and G. Wen (2018), “Hyperspectral Anomaly Detection via Background Estimation and Adaptive Weighted Sparse Representation”, *Remote Sensing*, 10 (2).
- [138] B. Zitova and J. Flusser (2003), “Image registration methods: A survey”, *Image Vision Comput.*, 21 (11), 977–1000.